

# Encyclopedia of Sociology

---

Edgar F. Borgatta  
Rhonda J.V. Montgomery

Second Edition

# Encyclopedia of Sociology

---

Second Edition

# Encyclopedia of Sociology

---

Second Edition

VOLUME **1**

**Edgar F. Borgatta**

*Editor-in-Chief*

*University of Washington, Seattle*

**Rhonda J. V. Montgomery**

*Managing Editor*

*University of Kansas, Lawrence*

**Macmillan Reference USA**

*an imprint of the Gale Group*

*New York • Detroit • San Francisco • London • Boston • Woodbridge, CT*

Encyclopedia of Sociology  
Second Edition

Copyright © 2000 by Edgar F. Borgatta and Rhonda J. V. Montgomery

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the Publisher.

Macmillan Reference USA  
an imprint of The Gale Group  
1633 Broadway  
New York, NY 10019

**Library of Congress Catalog in Publication Data**

Encyclopedia of Sociology / Edgar F. Borgatta, editor-in-chief, Rhonda Montgomery, managing editor.—2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-02-864853-6 (set: alk paper)—ISBN 0-02-864849-8 (v. 1: alk. paper)—  
0-02-864850-1 (v. 2)—0-02-86485-1 (v. 3)—0-02-864852-8 (v. 4)—0-02-865581-8 (v. 5)

I. Sociology—Encyclopedias. I. Borgatta, Edgar F., 1924- II. Montgomery, Rhonda J. V.

HM425 .E5 2000  
301'.03—dc21

00-028402  
CIP

Printed in the United States of America by the Gale Group  
Gale Group and Design is a trademark used herein under license.

**Staff**

*Publisher*

Elly Dickason

*Project Editors*

Timothy Prairie

Pamela Proffitt

*Editorial Assistants*

Shawn Beall

Wayne Yang

*Assistant Manager, Composition*

Evi Seoud

*Buyer*

Rhonda Williams

*Senior Art Director*

Michelle DiMercurio

---



# Contents

---

Preface .....	v
Preface for Second Edition .....	ix
Acknowledgments .....	xi
List of Articles .....	xiii
List of Authors .....	xxiii
Encyclopedia of Sociology.....	1
Index .....	3291

# Preface

---

The idea for this *Encyclopedia of Sociology* was in gestation for a long time. Probably the notion arose when, as Sociology Advisory Editor for Rand McNally and Company, I arranged for a series of handbooks that were published in the 1960s and 1970s. This influential group of volumes covered most of sociology, especially with the *Handbook of Modern Sociology* (Robert E. L. Faris, 1964) as a key volume. Other titles in the list included: *Handbook of Marriage and the Family* (Harold T. Christensen, 1964); *Handbook of Organizations* (James G. March, 1965); *Handbook of Socialization Theory and Research* (David A. Goslin, 1968); *Handbook of Personality Theory and Research* (Edgar F. Borgatta & William W. Lambert, 1968); *Handbook on the Study of Social Problems* (Erwin O. Smigel, 1971); and *Handbook of Criminology* (Daniel Glaser, 1974). Effectively, the series functioned as an encyclopedia, especially since there was additional related coverage already provided by the *Handbook of Social Psychology* (Gardner Lindzey and Elliot Aronson, 1968). At that time Macmillan's *International Encyclopedia of the Social Sciences* (David L. Sills, ed., 1968) was also available, and a separate encyclopedia for sociology seemed superfluous.

With time, however, as social-science research and professional involvement grew, along with the proliferation of subfields, each of the social and behavioral sciences and, indeed, other specialties, such as statistics, area studies, and applied areas, developed useful encyclopedias. In the late 1970s I talked about an encyclopedia of sociology with F. E. (Ted) Peacock (F. E. Peacock Publishers, Inc.), who encouraged the development of the project.

However, since it takes time for these things, it was not until the early 1980s that I actually started reflecting actively on what would need to be done, and I sought advice on what actually would be involved in such a project. Fortunately, Raymond J. Corsini, a good friend with whom I had worked on other matters, invited me to be an Associate Editor for the *Encyclopedia of Psychology* (Corsini, 1984). I got a close look at what was involved in undertaking a project of this magnitude and I was persuaded that the task would be a feasible one for sociology.

The field of sociology had been growing and evolving rapidly in the post-World War II period. Possibly the decades of the 1960s and 1970s will be seen in retrospect as one of the periods of great change for the discipline. Of course, different people will judge past developments differently, but some of the changes that have to be recognized as important include the following:

*First.* Sociology, which August Comte had blessed with the title of the “Queen of the Social Sciences,” seemed to be losing much of the empire. In particular, applied fields dealing with social behavior blossomed, but as they did so, sociology seemed indifferent, uninvolved. The field of social work developed its advanced degree programs and established research interests that sociology relinquished as uninteresting because they were “applied.” The field of industrial sociology virtually disappeared as the interest in research flourished in several specialties in psychology and in schools of business and management. Interest in the key institution, the family, was largely lost to

the special applied organizations in that area. And so it went in a number of other fields. The “Queen” appeared indifferent, possibly with the exception of the field of medical sociology, in which there was considerable development.

*Second.* Technical training in sociology became increasingly more demanding. When I taught the first graduate course in statistics for sociology at the New York University Graduate School in 1954, it included regression analysis and factor analysis. The reception and reputation was a bit like that greeting the arrival of extraterrestrials. The title (or epithet) “Factor Analyst” was definitely not meant to be complimentary. Nevertheless, in the 1950s, the Social Science Research Council (SSRC) and others supported the idea that the formal theory and technical bases of the social sciences required attention, and programs were initiated to foster a greater appreciation of mathematics and statistics. Particularly with the support of the National Institutes of Mental Health (NIMH) graduate training grants, the University of Wisconsin, the University of Michigan, and other centers concentrated on “research methods” during and following the 1960s. The discipline reflected this focus in its journals. Sociology also became known as the leader in research training in the social sciences, with the new generation of scholars becoming conversant with statisticians, econometricians, and psychometricians, and providing service to history, political science, and anthropology. The “Queen” again had some empire.

*Third.* The 1960s experienced the civil-rights movement, the student movement, the feminist movement, and, implicitly or explicitly, sociologists reacted to and sometimes participated actively in these social movements. Challenges arose to the “traditional” values of objective and “value free” science in sociology. These challenges ranged from positions asserting knowledge by intuition to the posing of more serious epistemological questions. Attention was drawn to the fact that sociology apparently had little utility in solving social problems, aside from assisting in exposing them, but, further, sociologists were often accused of not studying complex problems because they were limited and hampered by their methodologies. A resurgence of interest in “qualitative” approaches developed, which also provided a stimulus for a reexamination of existing research approaches.

*Fourth.* At the same time, the scope of what sociologists could accomplish more generally expanded with technical development. Two of the more prohibitive cost factors in research and scholarship have progressively been reduced, since the development of computing packages made possible the elimination of computing clerks at the same time that it made possible complex numerical and statistical analyses. Additionally, this development eliminated time losses as the labor intensive aspects were eliminated. Also, the availability of word processing packages made it possible for even the most helpless scholar to by-pass the secretary or typing pool and get materials into a readable and revisable format. As these earlier “barriers” to productivity were removed, presumably the social sciences responded accordingly. In any event, there has been a proliferation of journals, and increasing collateral publication continues in various media.

*Fifth.* The continued development of the field of sociology can be marked by the increase of special subfields. Aside from the increases in publication, the number of specialization sections in the American Sociological Association (ASA) continues to grow, as do the Research Committees in the International Sociological Association (ISA). A reflection of this may be seen by glancing at the topical coverage of *Contemporary Sociology*, the ASA journal of book reviews.

This broadening of the field of sociology affected the way topics were chosen for the *Encyclopedia of Sociology*. In the early stages, a broad set of topics was used to accumulate the important concepts and subfields included in sociology. Initially, the objective was to be as inclusive as possible and to avoid errors of omission. A constant problem in the process was that topics did not fit neatly into only one broad category. Often they could fit as easily into two, three, or four. In fact, the number of broad categories became increasingly elastic, but eventually these were reduced to seventeen, corresponding to no known system of organization other than expedience. The broad categories did not have any obvious theoretical basis of division, which was disconcerting, but represented the pragmatic result of many revisions. Our Advisory and Associate Editors participated in reviews of the total set of categories or of selected subsets for a few of the broad categories. It is fair to report that while we often saw consensus in the process,

sometimes we felt that there was no effective way to manage the procedure for selection of topics or to satisfy every piece of advice, sound as it might seem. At one point we had more than 1,700 potential entry titles. These eventually were consolidated into about 400 titles with notations of how overlapping concepts were handled, how related concepts were to be combined, and so forth. In making the arrangements with authors, further consolidation brought the final number of entries to the 370 in this 4-volume set.

The process of defining topics, thus, while driven by theoretical interests and strategic representations of the field, ultimately resulted in a pragmatic and eclectic product. Thus some topics became very comprehensive while others have more specific content. In areas where there is intensive attention by sociologists, such as social stratification, race and ethnic studies, gender, medical sociology, and aging, coverage by authors may overlap in a way that provides emphasis.

Other factors that guided the formulation of entry topics included defining the audience for whom the encyclopedia was intended. It was expected that sociologists would read about areas with which they were not familiar, but we wanted the materials to be useful to other scholars and professionals who need information about topics in sociology. Further, encyclopedias are gold mines for students, and so a central concern was that articles could be read and understood by younger and uninitiated persons looking for a first introduction to a sociological topic. This latter message was communicated to authors, and in large part it has been possible to provide presentations that

will reach a broad range of literate audiences. There are some obvious exceptions. In some technical areas the presentations, while self-contained and elegantly presented, do require a preexisting knowledge base in order to be fully understood by the readers.

OCTOBER 1991

EDGAR F. BORGATTA, *EDITOR-IN-CHIEF*

REFERENCES

- Borgatta, Edgar F., and William W. Lambert 1968 (eds.) *Handbook of Personality Theory and Research*. Chicago: Rand McNally and Company.
- Christensen, Harold T. 1964 (ed.) *Handbook of Marriage and the Family*. Chicago: Rand McNally and Company.
- Corsini, Raymond J. 1984 (ed.) *Encyclopedia of Psychology*. New York: John Wiley & Company.
- Faris, Robert E. L. 1964 (ed.) *Handbook of Modern Sociology*. Chicago: Rand McNally and Company.
- Glaser, Daniel 1974 (ed.) *Handbook of Criminology*. Chicago: Rand McNally and Company.
- Goslin, David A. 1968 (ed.) *Handbook of Socialization Theory and Research*. Chicago: Rand McNally and Company.
- Lindzey, Gardner, and Elliot Aronson 1968 (eds.) *Handbook of Social Psychology*, 2nd ed. Boston: Addison-Wesley Publishing Company.
- March, James G. 1965 (ed.) *Handbook of Organizations*. Chicago: Rand McNally and Company.
- Sills, David L. 1968 (ed.) *International Encyclopedia of the Social Sciences*. New York: Macmillan Company and Free Press.
- Smigel, Erwin O. 1971 (ed.) *Handbook on the Study of Social Problems*. Chicago: Rand McNally and Company.

# Preface for Second Edition

---

After the *Encyclopedia of Sociology* had been in print only three years, we began to receive inquiries about when there would be a revised edition. This was surprising given that the *Encyclopedia* was so well received, that its distribution had been much broader than even optimistic supporters of the project had anticipated, and that the articles were largely broad reviews and summaries of areas of knowledge in sociology. However, some areas in sociology changed quickly during the last decade as we approached the Millennium so that interest in a recapitulation and updating did not seem inappropriate. In addition, the social sciences appear to have softened their borders, and thus we realized that a substantial and thoughtful addition of titles would add breadth and depth to the *Encyclopedia*.

August Comte's description of sociology as the "queen" of the social sciences seems to have been awakened in a new generation, and the relevance of sociology to the social and behavioral sciences has been renewed. We took seriously our obligation to improve the representation of the areas of sociology in this edition of the *Encyclopedia*. The *Encyclopedia* was greatly improved through the input of Advisory Editors and authors who identified new content areas and titles that should be included and indicated which titles could be eliminated or consolidated. Some provided comprehensive reviews of the *Encyclopedia's* scope and coverage, as well as reviews of the content of many individual articles. Suggestions for additional titles for the revised *Encyclopedia* accumulated to a list of over 80 concepts and themes, resulting in the

addition of 66 new titles, but in addition some of the revised articles also included substantially new and expanded topics.

With the help of the Advisory Editors and quite a few authors, we reviewed articles and sectors of coverage to determine what changes would be important in a new edition. We distilled the major points of emphasis provided by reviews and user comments, and incorporated them into the guidelines for revision provide to authors.

Reflecting the kind of question that comes up so often in sociology doctoral exams, reviewers repeatedly asked us why a particular article was included in an encyclopedia of sociology. Authors who are expert in a particular subject area assume too frequently that readers will know their topic's relevance for sociology. To guard against this we asked authors to note the sociological relevance of the topic and to show how it fit into not only the scheme of sociological knowledge but also social and behavioral knowledge in general. As a consequence, most articles have been expanded.

Authors, experts in their fields, often concentrate on the knowledge and the issues within their field but do not give sufficient attention to the *practical value* of that knowledge, particularly how it is important for policy formation and in applications to everyday life. Of course, this is a comment often made about academic scholars in general, namely, that they sometimes forget that an important reason for research and the accumulation of knowledge is to provide bases for useful and informed applications.

Reviewers raised another theme. Articles often did a wonderful job of summarizing knowledge but did not indicate what to expect from future endeavors in the field. In other words, what areas need more attention in scholarship and research to expand the knowledge in a given field? While this kind of presentation is speculative, we reminded authors of the need to give direction for future work.

An additional theme for the revised edition is one that is temporally controlled. There is no way that references can provide more information than what already has been published. Updating content is important, but equally important is providing information about easily accessible general resources for those who want to go beyond the relatively brief discussions in the *Encyclopedia*. We reminded authors that the purpose of the reference section is to provide users with an opportunity to explore the area further. Academic scholars can too easily become exhaustive bibliographers. Thus, we asked authors to give special attention to providing direction rather than overwhelming the reader, and we are impressed that most authors have been extremely successful in this task. In addition to the work of the authors, the professional sociological staff of the *Encyclopedia* prepared for some article a short list of additional references to broaden the scope of coverage and

provide additional transitions to related concepts. We updated and provided new references for 20 articles from the earlier edition of the *Encyclopedia*.

Finally, reviewers commented that some of the presentations in the first edition were too brief, and some topics were too narrowly drawn. Thus, some topics have been combined, some topics have been eliminated and the content incorporated into related broader articles, and many articles have been expanded to cover neglected aspects of a topic and to provide greater detail for a more well-rounded presentation. Thirty-nine titles were eliminated and incorporated into more substantial articles, but some additional titles were changed when the original topic was expanded. In summary, there were 370 articles in the original edition, 39 were eliminated and 66 new articles were added, resulting in 397 in this revised edition.

In short, we have greatly improved the breadth and depth of coverage in the *Encyclopedia*, and we have paid particular attention to those articles that relate to other social and behavioral sciences. We have substantially increased the content of the *Encyclopedia* in this edition, and we have made every effort to ensure that the content is current, accurate, and representative of the field.

EDGAR F. BORGATTA, *EDITOR-IN-CHIEF*

# Acknowledgments

---

The revision of the *Encyclopedia of Sociology* began with the optimistic hope that following reasonably on the original edition it would turn out to be easier to do. It did not turn out that way, but for good reasons. With the support and encouragement of Elly Dickason, Macmillan Reference USA's Publisher, the research and development for the project was more extensive than anticipated. On this score, our debt to the Advisory Editors in providing broad reviews and suggestions can hardly be emphasized enough. Further, as we proceeded in inviting old and new authors it became apparent that the Encyclopedia was getting a great deal of use and authors provided guidance from their experience on directions for development. Our authors are, of course, truly responsible for the Encyclopedia, but in addition through much interaction they provided support that was both welcome and of enormous value.

Marie L. Borgatta was the managing editor for the original publication, and for this edition she requested a less demanding role as a consulting editor. It was a nice idea, but she inevitably became involved and delivered support in quantities and quality that can only be described as exploitation.

Many ideas and innovations are due to her participation in the project. As Managing Editor for the revision, Rhonda J. V. Montgomery provided a work and intellectual partnership for the project that contributed enormously to the breadth and quality of the Revised Edition. Her initiatives kept the progress of the project smooth and on a reasonable schedule in spite of often unanticipated demands.

During the preparation of this Revised Edition, the Gale Group acquired Macmillan Reference USA, and production moved from New York to Michigan. This was done with only minimal loss of time and effort because of the planning and management of Elly Dickason, the Publisher, but also because in the New York office Timothy Prairie generated a very effective program for copyediting manuscripts and managing the records of progress. Of course, the shift to Michigan was obviously not effortless, but Linda Hubbard guided the project into the capable hands of Pamela Proffitt, who took on the mammoth task of making the Revised Edition of the *Encyclopedia of Sociology* a reality. The devoted efforts of the publisher's staff are greatly appreciated.



# List of Articles

---

ADOLESCENCE Michael J. Shanahan	ALCOHOL Ronald L. Akers Thomas R. Heffington	APPLIED SOCIOLOGY David W. Britt
ADULT EDUCATION John Houghton	ALIENATION Clarence Y. H. Lo	ART AND SOCIETY Gladys Lang
ADULTHOOD Jeylan T. Mortimer Pamela Aronson	ALTERNATIVE LIFESTYLES Karen Seccombe	ASIAN AMERICAN STUDIES William T. Liu Elena S. H. Yu
AFFECT CONTROL THEORY AND IMPRESSION FORMATION David R. Heise	ALTRUISM Judith Howard Jane Allyn Piliavin	ATTITUDES John D. DeLamater
AFFIRMATIVE ACTION Nathan Glazer	AMERICAN FAMILIES Laurie Russell Hatch	ATTRIBUTION THEORY Judith Howard Danielle Kane
AFRICAN-AMERICAN STUDIES Joe R. Feagin	AMERICAN INDIAN STUDIES C. Matthew Snipp	BANKRUPTCY AND CREDIT Teresa A. Sullivan
AFRICAN STUDIES Walter R. Allen Shani O'Neal	AMERICAN SOCIETY Robin Williams, Jr.	BEHAVIORISM W. David Pierce
AGGRESSION Craig A. Anderson	AMERICAN SOCIOLOGICAL ASSOCIATION AND OTHER SOCIOLOGICAL ASSOCIATIONS Carla B. Howery	BIRTH AND DEATH RATES Deborah A. Sullivan
AGING AND THE LIFE COURSE Linda K. George	ANALYSIS OF VARIANCE AND COVARIANCE Richard A. Bulcroft	BRITISH SOCIOLOGY Vaneeta D'Andrea Peter Robbins
AGRICULTURAL INNOVATION David Gartrell John Gartrell	ANOMIE Robert Crutchfield Kristin A. Bates	BUREAUCRACY David G. Nickinovich
		CAPITALISM Patrick H. Mooney

LIST OF ARTICLES

---

CASE STUDIES Gale E. Miller Richard S. Jones	COMMUNITY Gunnar Almgren	COURT SYSTEMS AND LAW Edward Gross
CASTE AND INHERITED STATUS Rita Jalali	COMPARABLE WORTH Nancy E. Durbin Barbara Melber	CRIME, THEORIES OF Don C. Gibbons
CAUSAL INFERENCE MODELS Hubert M. Blalock, Jr.	COMPARATIVE HEALTH-CARE SYSTEMS Mary Zimmerman	CRIME RATES Robert M. O'Brien
CENSORSHIP AND THE REGULATION OF EXPRESSION Jeffrey Strange Melanie Green Timothy Brock	COMPARATIVE-HISTORICAL SOCIOLOGY Larry Griffin	CRIMINAL AND DELINQUENT SUBCULTURES James F. Short
CENSUS Karl E. Taeuber	COMPLEX ORGANIZATIONS Howard E. Aldrich Peter V. Marsden	CRIMINALIZATION OF DEVIANCE Jackson Toby
CHILDHOOD SEXUAL ABUSE Heather Cecil	COMPLIANCE AND CONFORMITY Cecilia L. Ridgeway	CRIMINAL SANCTIONS Peter Cordella
CHINA STUDIES Martin K. Whyte	COMPUTER APPLICATIONS IN SOCIOLOGY Ronald E. Anderson	CRIMINOLOGY Robert Crutchfield
CITIES Maurice D. Van Arsdol, Jr.	CONFLICT THEORY Randall Collins	CRITICAL THEORY Dwight B. Billings Patricia Jennings
CIVIL LIBERTIES Howard P. Greenwald	CONTENT ANALYSIS Philip J. Stone	CROSS CULTURAL ANALYSIS William A. Corsaro Arthur Alderson
CLASS AND RACE Edna Bonacich	CONVERGENCE THEORIES Richard M. Coughlin	CROWDS AND RIOTS David Snow Ronnelle Paulsen
CLINICAL SOCIOLOGY Jan Fritz	CONVERSATION ANALYSIS Jack Whalen Geoff Raymond	CULTURE Samuel L. Gilmore
COALITIONS Theodore Caplow	CORPORATE ORGANIZATIONS Cora B. Marrett	DATA BANKS AND DEPOSITORIES Guido Martinotti Sonia Stefanizzi
COGNITIVE CONSISTENCY THEORIES Todd D. Nelson	CORRELATION AND REGRESSION ANALYSIS Herbert L. Costner	DEATH AND DYING John W. Riley, Jr.
COHORT PERSPECTIVES Matilda White Riley	COUNTERCULTURES Patricia A. Adler Peter Adler	DECISION-MAKING THEORY AND RESEARCH Evan Pritchard
COMMUNITARIANISM Amitai Etzioni	COURTSHIP Mary Riege Laner	DEMOCRACY Andrew L. Creighton

---

LIST OF ARTICLES

---

DEMOGRAPHIC METHODS Neil G. Bennett	EDUCATIONAL ORGANIZATION Kevin Dougherty Floyd Hammack	EVOLUTION: BIOLOGICAL, SOCIAL, CULTURAL Maurice N. Richter, Jr.
DEMOGRAPHIC TRANSITION Avery M. Guest Gunnar Almgren	EDUCATION AND DEVELOPMENT Yih-Jin Young	EXPECTATION STATES THEORY Martha Foschi
DEMOGRAPHY Halliman H. Winsborough	EDUCATION AND MOBILITY John Katsillis J. Michael Armer	EXPERIMENTS Gerald Marwell
DEPENDENCY THEORY Jon Hendricks	EMOTIONS Theodore D. Kemper	EXTREME INFLUENCE: THOUGHT REFORM, HIGH CONTROL GROUPS, INTERROGATION, AND RECOVERED MEMORY PSYCHOTHERAPY Richard Ofshe
DEPRESSION Regan A. R. Gurung	ENVIRONMENTAL EQUITY Glynis Daniels	FACTOR ANALYSIS Duane F. Alwin
DESCRIPTIVE STATISTICS Daphne Kuo	ENVIRONMENTAL SOCIOLOGY Riley E. Dunlap Eugene A. Rosa	FAMILY AND HOUSEHOLD STRUCTURE Gary R. Lee Padma Srinivasan
DEVIANCE THEORIES George S. Bridges Scott A. Desmond	EPIDEMIOLOGY Charles E. Holzer, III	FAMILY AND POPULATION POLICY IN LESS DEVELOPED COUNTRIES Julie Brines
DIFFUSION THEORIES Ronald W. Perry	EPISTEMOLOGY Costantino Cipolla Guido Giarelli	FAMILY AND RELIGION Jeffrey G. Pankhurst Susan K. Houseknecht
DISASTER RESEARCH Enrico L. Quarantelli	EQUALITY OF OPPORTUNITY James M. McPartland Jomills Henry Braddock, II	FAMILY LAW Lisa J. McIntyre
DISCRIMINATION Thomas F. Pettigrew Marylee C. Taylor	ETHICS IN SOCIAL RESEARCH Howard P. Greenwald	FAMILY PLANNING Jacqueline Darroch Jennifer Frost
DIVISION OF LABOR Paul Morgan Baker	ETHNICITY Richard D. Alba	FAMILY POLICY IN WESTERN SOCIETIES Irving Tallman Ginna M. Babcock
DIVORCE Steven L. Nock Alison Burke Albers	ETHNOGRAPHY Carol A. B. Warren	FAMILY SIZE Lauri Steel
DRUG ABUSE Charles Winick Leonard Saxe	ETHNOMETHODOLOGY John Heritage	FAMILY VIOLENCE Murray A. Straus
ECONOMIC DETERMINISM William DiFazio	EVALUATION RESEARCH Karl Kosloski	
ECONOMIC INSTITUTIONS Fred Block	EVENT HISTORY ANALYSIS Paul D. Allison	
ECONOMIC SOCIOLOGY Richard Swedberg		

---

LIST OF ARTICLES

---

FEMININITY/MASCULINITY Peter J. Burke Jan E. Stets	GROUP SIZE David R. Schmitt	IDENTITY THEORY Sheldon Stryker
FEMINIST THEORY Gaye Tuchman	HEALTH AND ILLNESS BEHAVIOR David Mechanic Donna McAlpine	ILLEGITIMACY Frank F. Furstenberg, Jr.
FERTILITY DETERMINANTS Susan Cotts Watkins	HEALTH AND THE LIFE COURSE Donald E. Stull	IMPERIALISM, COLONIALISM, AND DECOLONIZATION David Strang
FIELD THEORY Richard B. Kettner-Polley	HEALTH CARE UTILIZATION AND EXPENDITURES Charlene Harrington David N. Pellow	INCEST Brenda J. Vander Mey
FILIAL RESPONSIBILITY Rhonda J. V. Montgomery	HEALTH POLICY ANALYSIS Carroll L. Estes Christopher Wellin	INCOME DISTRIBUTION IN THE UNITED STATES Beth B. Hess
FRENCH SCHOOL OF SOCIOLOGY, THE Bernard Valade	HEALTH PROMOTION AND HEALTH STATUS Rosalie F. Young	INDIAN SOCIOLOGY Vimal P. Shah
FUNCTIONALISM AND STRUCTURALISM Alexandra Maryanski Jonathan H. Turner	HIERARCHICAL LINEAR MODELS A. George Alder	INDIGENOUS PEOPLES Joane Nagel Thomas D. Hall
FUTURES STUDIES AS HUMAN AND SOCIAL ACTIVITY Eleonora Barbieri Masini	HIGHER EDUCATION Steven G. Brint	INDIVIDUALISM Lawrence E. Hazelrigg
GAME THEORY AND STRATEGIC INTERACTION H. Andrew Michener	HISPANIC AMERICANS Marta Tienda	INDUSTRIALIZATION IN LESS DEVELOPED COUNTRIES R. Alan Hedley
GENDER Judith Lorber	HISTORICAL SOCIOLOGY Stephen A. Kent	INDUSTRIAL SOCIOLOGY Ivar Berg
GENOCIDE David Smith	HOMELESSNESS Barrett A. Lee	INFANT AND CHILD MORTALITY Georgina Rojas-Garcia Samuel Echevarria
GERMAN SOCIOLOGY Erwin K. Scheuch	HUMAN ECOLOGY AND ENVIRONMENTAL ANALYSIS Lakshmi Bharadwaj	INFORMAL ECONOMY, THE Ann R. Tickamyer Stephanie Bohon
GLOBALIZATION AND GLOBAL SYSTEMS ANALYSIS Sean O Riain Peter B. Evans	HUMANISM Joseph Scimecca	INFORMATION SOCIETY Don A. Dillman
GOVERNMENT REGULATION Robin Stryker	HUMAN NATURE Alan Wolfe	INHERITANCE Marvin B. Sussman
GROUP CONFLICT RESOLUTION Lynne Woehrle	HUMAN RIGHTS/CHILDREN'S RIGHTS Edgar F. Borgatta	INTELLECTUALS Seymour Martin Lipset

---

LIST OF ARTICLES

---

INTELLIGENCE Linda S. Gottfredson	JAPANESE SOCIOLOGY Masamichi Sasaki	LIFE CYCLE Dennis Hogan
INTERGENERATIONAL RELATIONS Karl Pillemer J. Jill Sutor Shirley A. Keeton	JUVENILE DELINQUENCY, THEORIES OF John Hagan	LIFE EXPECTANCY Eileen M. Crimmins
INTERGENERATIONAL RESOURCE TRANSFERS Donald E. Stull Barbara Hirshorn	JUVENILE DELINQUENCY (AND JUVENILE CRIME) James E. Teele	LIFE HISTORIES AND NARRATIVE David R. Maines
INTERGROUP AND INTERORGANIZATIONAL RELATIONS Oscar Grusky Jeff Erger	KINSHIP SYSTEMS AND FAMILY TYPES Bernard Farber	LIFESTYLES AND HEALTH Janet Hankin
INTERMARRIAGE Grace J. Yoo Robert E. Staples	LABOR FORCE Valerie K. Oppenheimer	LITERATURE AND SOCIETY Sarah M. Corse
INTERNAL MIGRATION Charles F. Longino, Jr.	LABOR MOVEMENTS AND UNIONS William Form	LONGITUDINAL RESEARCH Richard T. Campbell
INTERNATIONAL ASSOCIATIONS IN SOCIOLOGY David Sciulli	LATIN AMERICAN STUDIES Carlos H. Waisman	LONG-TERM CARE David R. Phillips
INTERNATIONAL LAW Mary Ellen O'Connell	LAW AND LEGAL SYSTEMS Joseph Sanders	LONG TERM CARE FACILITIES Eva Kahana
INTERNATIONAL MIGRATION David M. Heer	LAW AND SOCIETY William M. Evan	LOVE Felix M. Berardo Erica Owens
INTERNET Christopher W. Gunn Diana Odom Gunn	LEADERSHIP Roya Ayman	MACROSOCIOLOGY Ann R. Tickamyer Jieli Li
INTERPERSONAL ATTRACTION Ardyth Stimson	LEGISLATION OF MORALITY Augustine Brannigan	MAJOR PERSONALITY THEORIES Raymond J. Corsini
INTERPERSONAL CONFLICT RESOLUTION Anupama Joshi	LEISURE John R. Kelly	MARGINAL EMPLOYMENT Leif Jensen Timothy A. Slack
INTERPERSONAL POWER Morris Zelditch, Jr.	LEVELS OF ANALYSIS Robert M. O'Brien	MARITAL ADJUSTMENT Yoshinori Kamo
ITALIAN SOCIOLOGY Alberto Gasparini	LIBERALISM/CONSERVATISM Robert Boguslaw	MARRIAGE Marilyn Ihinger-Tallman
	LIBRARY RESOURCES AND SERVICES FOR SOCIOLOGY Timothy D. Jewell	MARRIAGE AND DIVORCE RATES Paul C. Glick
	LIFE COURSE, THE Glen H. Elder, Jr.	MARXIST SOCIOLOGY Raymond A. Morrow

---

LIST OF ARTICLES

---

MASS MEDIA RESEARCH Janice M. Engsberg Gladys Lang Kurt Lang	MILITARY SOCIOLOGY David J. Armor	PARENTAL ROLES Constance L. Shehan
MASS SOCIETY Richard F. Hamilton	MODERNIZATION THEORY J. Michael Armer John Katsillis	PARTICIPATORY RESEARCH Francesca Cancian
MATERIALISM Robert J. Antonio	MONEY Viviana Zelizer	PEACE Louis Kreisberg
MATE SELECTION THEORIES Kris Bulcroft	MORAL DEVELOPMENT Thomas J. Figurski	PENOLOGY J. Forbes Farmer
MATHEMATICAL SOCIOLOGY Barbara F. Meeker Robert K. Leik	MULTIPLE INDICATOR MODELS Kyle Kercher	PERSONAL AUTONOMY Kerry Marsh
MEASUREMENT Edgar F. Borgatta Yoshinori Kamo	MUSIC Richard A. Peterson	PERSONAL DEPENDENCY Mario Mikulincer
MEASURES OF ASSOCIATION Herbert L. Costner	NATIONAL BORDER RELATIONS, MANAGEMENT OF Raimondo Strassoldo	PERSONALITY AND SOCIAL STRUCTURE Kenneth I. Spenner
MEDICAL-INDUSTRIAL COMPLEX Carroll L. Estes David N. Pellow Charlene Harrington	NATIONALISM Saad Nagi	PERSONALITY MEASUREMENT Gerard Saucier
MEDICAL SOCIOLOGY Mary Zimmerman	NEGOTIATION OF POWER Bruno Tellia	PERSONALITY THEORY Frank Dumont
MENTAL ILLNESS AND MENTAL DISORDERS Suzanne T. Ortega Sharon L. Larson	NONPARAMETRIC STATISTICS Subhash Sonnad	PERSUASION Susan McWilliams
META-ANALYSIS Blair T. Johnson	OBSERVATION SYSTEMS Richard B. Kettner-Polley	PHENOMENOLOGY Myron Orleans
METATHEORY Walter L. Wallace	OCCUPATIONAL AND CAREER MOBILITY Margaret Mooney Marini	POLICE Carl B. Klockars
MEXICAN STUDIES Guillermo de la Pena Bryan R. Roberts	OCCUPATIONAL PRESTIGE Keiko Nakao	POLISH AND EASTERN EUROPEAN SOCIOLOGY Jacek Szmataka Kinga Wysienska
MIDDLE EASTERN STUDIES George Sabagh	ORGANIZATIONAL STRUCTURE Rodolfo Alvarez Leah Robin	POLITICAL AND GOVERNMENTAL CORRUPTION Richard Mcgrath Skinner
	ORGANIZED CRIME Marilyn E. Walsh	POLITICAL CORRECTNESS Pierre L. van den Berghe
	PARADIGMS AND MODELS Bernard P. Cohen	POLITICAL CRIME Austin T. Turk

---

LIST OF ARTICLES

---

POLITICAL ORGANIZATIONS David Knoke	PROBATION AND PAROLE Joseph G. Weis	REMARRIAGE J. Jill Suitor Shirley A. Keeton
POLITICAL PARTY SYSTEMS Frederick D. Weil	PROFESSIONS Patricia A. Roos	REPLICATION Lowell J. Hargens
POLITICAL SOCIOLOGY J. Craig Jenkins	PROTEST MOVEMENTS Leonard Gordon	RESEARCH FUNDING IN SOCIOLOGY Howard P. Greenwald
POPULAR CULTURE Chandra Mukerji	PUBLIC OPINION Kurt Lang	RETIREMENT Melissa A. Hardy Kim Shuey
POPULATION Thomas W. Pullum	PUBLIC POLICY ANALYSIS James D. Wright	REVOLUTIONS Robert C. Liebman
PORNOGRAPHY Charles Winick	QUALITATIVE METHODS Laura L. Ellingson Carolyn Ellis	ROLE THEORY Bruce J. Biddle
POSITIVE MENTAL HEALTH Janet Hankin	QUALITATIVE MODELS Alex Durig David R. Heise	ROLE THEORY: FOUNDATIONS, EXTENSIONS, AND APPLICATIONS John H. Yost James R. Bailey
POSITIVISM Jonathan H. Turner	QUALITY OF LIFE Kyriakos S. Markides	RURAL SOCIOLOGY Gene F. Summers Frederick H. Buttel
POSTINDUSTRIAL SOCIETY Robert A. Fiala	QUASI-EXPERIMENTAL RESEARCH DESIGNS Kyle Kercher Karl Kosloski	SAMPLE SELECTION BIAS Christopher Winship Robert D. Mare
POSTMODERNISM Augustine Brannigan	RACE Susan R. Pitchford	SAMPLING PROCEDURES Lowell J. Hargens
POVERTY Wayne J. Vellemez	RATIONAL CHOICE THEORY Russell Hardin	SCANDINAVIAN SOCIOLOGY Carl Gunnar Janson
PRAGMATISM David R. Maines	RELIABILITY Richard A. Zeller	SCIENCE Warren Hagstrom
PREDICTION AND FUTURES STUDIES Alberto Gasparini	RELIGION, POLITICS, AND WAR John H. Simpson	SCIENTIFIC EXPLANATION Linda Derksen John Gartrell
PREGNANCY AND PREGNANCY TERMINATION Lynn Borgatta	RELIGIOUS MOVEMENTS Jeffrey K. Hadden	SECONDARY DATA ANALYSIS AND DATA ARCHIVES Toby L. Parcel
PREJUDICE Marylee C. Taylor Thomas F. Pettigrew	RELIGIOUS ORGANIZATIONS James R. Wood	
PROBABILITY THEORY Ronald J. Schoenberg	RELIGIOUS ORIENTATIONS Robert Wuthnow	

---



LIST OF ARTICLES

---

SECULARIZATION N.J. Demerath, III	SMALL GROUPS Robert W. Shotola	SOCIAL NETWORKS Peter V. Marsden
SEGREGATION AND DESEGREGATION Albert W. Black, Jr.	SOCIAL AND POLITICAL ELITES Patrick Akard	SOCIAL ORGANIZATION Herman Turk
SEGREGATION INDICES Douglas S. Massey	SOCIAL BELONGING Gabriele Pollini	SOCIAL PERCEPTION Barry Markovsky
SELF-CONCEPT Richard B. Felson	SOCIAL CAPITAL Tracy Karner	SOCIAL PHILOSOPHY Alan Wolfe
SELF-ESTEEM Todd F. Heatherton Kathleen D. Vohs	SOCIAL CHANGE Craig Calhoun	SOCIAL PROBLEMS Armand L. Mauss Valerie Jenness
SENTIMENTS Bernardo Cattarinussi	SOCIAL COMPARISON PROCESSES Karen A. Hegvedt	SOCIAL PSYCHOLOGY Howard B. Kaplan
SEX DIFFERENCES Monica Biernat Kay Deaux	SOCIAL CONTROL Allen E. Liska	SOCIAL PSYCHOLOGY OF STATUS ALLOCATION Archibald O. Haller Edward L. Fink Laura Janusik
SEXUAL BEHAVIOR IN MARRIAGE AND CLOSE RELATIONSHIPS Marion C. Willetts Janis C. Weber	SOCIAL DYNAMICS Nancy Brandon Tuma	SOCIAL RESOURCES THEORY Nan Lin
SEXUAL BEHAVIOR PATTERNS Kendal Broad John H. Scanzoni William Marsiglio	SOCIAL EXCHANGE THEORY Karen S. Cook	SOCIAL SECURITY SYSTEMS Fred C. Pampel
SEXUALLY TRANSMITTED DISEASES Mary Gilmore Pepper Schwartz	SOCIAL FORECASTING John Stimson	SOCIAL STRATIFICATION David B. Grusky
SEXUAL ORIENTATION William Simon	SOCIAL INDICATORS Kenneth C. Land	SOCIAL STRUCTURE Steve Rytina
SEXUAL VIOLENCE AND EXPLOITATION Sharon Araji	SOCIAL INEQUALITY Angela M. O'Rand	SOCIAL VALUES AND NORMS Margaret Mooney Marini
SLAVERY AND INVOLUNTARY SERVITUDE David R. James Sara Heiliger	SOCIALISM AND COMMUNISM Szonja Szelenyi	SOCIAL WORK Carol Austin Robert McClelland
	SOCIALIZATION Viktor Gecas	SOCIETAL STRATIFICATION Archibald O. Haller
	SOCIAL JUSTICE Duane F. Alwin	SOCIETY AND TECHNOLOGICAL RISKS Lee Clarke
	SOCIAL MOBILITY Thomas A. DiPrete	SOCIOBIOLOGY, HUMAN Joseph Lopreato
	SOCIAL MOVEMENTS Robert D. Benford Timothy B. Gongaware Danny L. Valadez	

---

LIST OF ARTICLES

SOCIOCULTURAL ANTHROPOLOGY Lin Poyer	STATUS ATTAINMENT Donald J. Treiman	TRANSNATIONAL CORPORATIONS R. Alan Hedley
SOCIOLINGUISTICS Allen D. Grimshaw	STATUS INCONGRUENCE Mattei Dogan	TYPOLOGIES Kenneth D. Bailey
SOCIOLOGY AMONG THE SOCIAL SCIENCES Mattei Dogan	STRESS Nan Lin	URBAN SOCIOLOGY Lee J. Haggerty
SOCIOLOGY OF EDUCATION Paul W. Kingston	STRUCTURAL LAG (AND CULTURAL LAG) Anne Foner	URBAN UNDERCLASS John D. Kasarda
SOCIOLOGY OF ISLAM Riaz Hassan	STUDENT MOVEMENTS Leonard Gordon	UTOPIAN ANALYSIS AND DESIGN Robert Boguslaw
SOCIOLOGY OF KNOWLEDGE E. Doyle McCarthy	SUBURBANIZATION John R. Logan	VALIDITY George W. Bohrnstedt
SOCIOLOGY OF LAW James M. Inverarity	SUICIDE Ronald W. Maris	VALUE THEORY AND RESEARCH David R. Karp
SOCIOLOGY OF RELIGION Rodney Stark	SURVEY RESEARCH Howard Schuman	VOLUNTARY ASSOCIATIONS William V. D'Antonio
SOUTHEAST ASIA STUDIES Charles Hirschman	SYMBOLIC INTERACTION THEORY Sheldon Stryker	VOTING BEHAVIOR Manfred Kuechler
SOVIET AND POST-SOVIET SOCIOLOGY Igor S. Kon Vladimir A. Yadov	SYSTEMS THEORY Robert Boguslaw	WAR Edgar Kiser
SPORT Janet Lever Jay Coakley	TABULAR ANALYSIS Jae-On Kim Myoung-Jin Lee	WHITE COLLAR CRIME James E. Teele
STANDARDIZATION Christine L. Himes	TERRITORIAL BELONGING Renzo Gubert	WIDOWHOOD Felix M. Berardo Donna H. Berardo
STATE, THE Jack A. Goldstone	TERRORISM Brian Michael Jenkins	WORK AND OCCUPATIONS Barbara F. Reskin
STATISTICAL GRAPHICS John Fox	TIME SERIES ANALYSIS Richard Berk	WORK ORIENTATION Charles W. Mueller
STATISTICAL INFERENCE Alan C. Acock	TIME USE RESEARCH Elke Koch-Weser Ammassari	WORLD RELIGIONS Larry D. Shinn
STATISTICAL METHODS J. Scott Long	TOURISM Giovanni Delli Zotti	

# List of Authors

---

- Alan C. Acock  
Department of Human  
Development and Family  
Sciences  
Oregon State University  
STATISTICAL INFERENCE
- Patricia A. Adler  
University of Colorado,  
Boulder  
COUNTERCULTURES
- Patrick Akard  
Department of Sociology,  
Anthropology, and  
Social Work  
Kansas State University  
SOCIAL AND POLITICAL  
ELITES
- Ronald L. Akers  
Criminology and Law  
University of Florida,  
Gainesville  
ALCOHOL
- Richard D. Alba  
Department of Sociology  
State University of New  
York, Albany  
ETHNICITY
- A. George Alder  
Department of Psychology  
University College of  
the Cariboo  
HIERARCHICAL LINEAR  
MODELS
- Howard E. Aldrich  
Department of Sociology  
University of North Carolina  
COMPLEX ORGANIZATIONS
- Walter R. Allen  
Department of Sociology  
University of California, Los  
Angeles  
AFRICAN STUDIES
- Paul D. Allison  
Sociology Department  
University of Pennsylvania  
EVENT HISTORY ANALYSIS
- Gunnar Almgren  
School of Social Work  
University of Washington  
COMMUNITY
- Rodolfo Alvarez  
Department of Sociology  
University of California, Los  
Angeles  
ORGANIZATIONAL  
STRUCTURE
- Duane F. Alwin  
Department of Sociology  
University of Michigan  
FACTOR ANALYSIS  
SOCIAL JUSTICE
- Elke Koch-Weser Ammassari  
University of Rome "La  
Sapienza"  
TIME USE RESEARCH
- Craig A. Anderson  
Department of Psychology  
Iowa State University  
AGGRESSION
- Ronald E. Anderson  
University of Minnesota  
COMPUTER APPLICATIONS  
IN SOCIOLOGY
- Robert J. Antonio  
Department of Sociology  
University of Kansas  
MATERIALISM
- Sharon K. Araj  
Department of Sociology  
University of Alaska  
SEXUAL VIOLENCE AND  
EXPLOITATION
- J. Michael Armer  
Department of Sociology  
Florida State University  
MODERNIZATION THEORY
- David J. Armor  
George Mason University  
MILITARY SOCIOLOGY
- Carol D. Austin  
Department of Social Work  
The University of Calgary  
SOCIAL WORK

LIST OF AUTHORS

- 
- |  |   |   |
|--|---|---|
| Roya Ayman<br>Illinois Institute of<br>Technology<br>LEADERSHIP  | Monica Biernat<br>Department of Psychology<br>University of Kansas<br>SEX DIFFERENCES   | Lynn Borgatta<br>Department of Obstetrics<br>and Gynecology<br>Boston University School of<br>Medicine<br>PREGNANCY AND<br>PREGNANCY<br>TERMINATION |
| Kenneth D. Bailey<br>Department of Sociology<br>University of California, Los<br>Angeles<br>TYPOLOGIES                         | Dwight B. Billings<br>Department of Sociology<br>University of Kentucky<br>CRITICAL THEORY                                    | Augustine Brannigan<br>Department of Sociology<br>The University of Calgary<br>LEGISLATION OF MORALITY<br>POSTMODERNISM                             |
| Paul Morgan Baker<br>Department of Sociology<br>University of Victoria<br>DIVISION OF LABOR                                    | Albert W. Black<br>Department of Sociology<br>University of Washington<br>SEGREGATION AND<br>DESEGREGATION                    | George S. Bridges<br>Department of Sociology<br>University of Washington<br>DEVIANCE THEORIES   |
| Robert D. Benford<br>Department of Sociology<br>University of Nebraska<br>SOCIAL MOVEMENTS                                     | Hubert M. Blalock<br>CAUSAL INFERENCE<br>MODELS   | Julie Brines<br>Department of Sociology<br>University of Washington<br>FAMILY AND POPULATION<br>POLICY IN LESS<br>DEVELOPED COUNTRIES               |
| Neil G. Bennett<br>Branch School of Public<br>Affairs and Columbia<br>University<br>DEMOGRAPHIC METHODS                        | Fred Block<br>Department of Sociology<br>University of California,<br>Davis<br>ECONOMIC INSTITUTIONS                          | Steven G. Brint<br>Department of Sociology<br>University of California,<br>Riverside<br>HIGHER EDUCATION  |
| Felix M. Berardo<br>Department of Sociology<br>University of Florida<br>LOVE<br>WIDOWHOOD                                      | Robert Boguslaw<br>SYSTEMS THEORY<br>UTOPIAN ANALYSIS AND<br>DESIGN<br>LIBERALISM/<br>CONSERVATISM                            | David W. Britt<br>Department of Sociology<br>Wayne State University<br>APPLIED SOCIOLOGY  |
| Ivar Berg<br>University of Pennsylvania<br>INDUSTRIAL SOCIOLOGY  | George W. Bohrnstedt<br>American Institutes for<br>Research<br>VALIDITY   | Kendal Broad<br>Department of Sociology<br>University of Florida<br>SEXUAL BEHAVIOR<br>PATTERNS   |
| Richard Berk<br>Departments of Statistics<br>and Sociology<br>University of California, Los<br>Angeles<br>TIME SERIES ANALYSIS | Edna Bonacich<br>Department of Sociology<br>University of California,<br>Riverside<br>CLASS AND RACE                          | Kris Bulcroft<br>Department of Sociology<br>Western Washington<br>University<br>MATE SELECTION<br>THEORIES  |
| Lakshmi Bharadwaj<br>University of Wisconsin<br>HUMAN ECOLOGY AND<br>ENVIRONMENTAL<br>ANALYSIS                                 | Edgar F. Borgatta<br>Department of Sociology<br>University of Washington<br>HUMAN RIGHTS/<br>CHILDREN'S RIGHTS<br>MEASUREMENT |   |
| Bruce J. Biddle<br>Center for Research in<br>Social Behavior<br>University of Missouri<br>ROLE THEORY                          |   |   |
-

LIST OF AUTHORS

- 
- |  |   |  |
|--|---|--|
| Richard A. Bulcroft<br>Department of Sociology<br>Western Washington<br>University<br>ANALYSIS OF VARIANCE<br>AND COVARIANCE     | Bernard P. Cohen<br>Department of Sociology<br>Stanford University<br>PARADIGMS AND MODELS  | Eileen M. Crimmins<br>Andrus Gerontology Center<br>University of Southern<br>California, Los Angeles<br>LIFE EXPECTANCY  |
| Peter J. Burke<br>Department of Sociology<br>Washington State University<br>FEMININITY/MASCULINITY                               | Randall Collins<br>CONFLICT THEORY  | Robert Crutchfield<br>Department of Sociology<br>University of Washington<br>ANOMIE<br>CRIMINOLOGY                       |
| Craig Calhoun<br>Social Science Research<br>Council<br>SOCIAL CHANGE   | Karen S. Cook<br>Department of Sociology<br>Stanford University<br>SOCIAL EXCHANGE<br>THEORY  | Vaneeta D'Andrea<br>Roehampton Educational<br>Development Centre<br>Roehampton Institute,<br>London<br>BRITISH SOCIOLOGY |
| Richard T. Campbell<br>LONGITUDINAL RESEARCH   | Peter Cordella<br>Department of Sociology<br>Saint Anselm College<br>CRIMINAL SANCTIONS   | William V. D'Antonio<br>Catholic University<br>VOLUNTARY ASSOCIATIONS  |
| Francesca Cancian<br>Department of Sociology<br>University of California,<br>Irvine<br>PARTICIPATORY RESEARCH                    | William A. Corsaro<br>Department of Sociology<br>Indiana University<br>CROSS CULTURAL<br>ANALYSIS   | Glynis Daniels<br>Department of Sociology<br>Pennsylvania State<br>University<br>ENVIRONMENTAL EQUITY                    |
| Theodore Caplow<br>Department of Sociology<br>University of Virginia<br>COALITIONS   | Sarah M. Corse<br>Department of Sociology<br>University of Virginia<br>LITERATURE AND SOCIETY   | Jacqueline Darroch<br>The Alan Guttmacher<br>Institute<br>FAMILY PLANNING  |
| Bernardo Cattarinussi<br>Dipartimento di Economia<br>Societa e Territorio<br>University of Udine, Italy<br>SENTIMENTS            | Raymond J. Corsini<br>Honolulu, Hawaii<br>MAJOR PERSONALITY<br>THEORIES   | Guillermo de la Pena<br>CIESAS del Occidente<br>MEXICAN STUDIES  |
| Heather Cecil<br>Department of Maternal and<br>Child Health<br>University of Alabama,<br>Birmingham<br>CHILDHOOD SEXUAL<br>ABUSE | Herbert L. Costner<br>Department of Sociology<br>University of Washington<br>CORRELATION AND<br>REGRESSION ANALYSIS<br>MEASURES OF<br>ASSOCIATION | John D. DeLamater<br>Center for Demography and<br>Ecology<br>University of Wisconsin,<br>Madison<br>ATTITUDES            |
| Costantino Cipolla<br>Universita di Parma<br>EPISTEMOLOGY  | Richard M. Coughlin<br>University of New Mexico<br>CONVERGENCE THEORIES   | Giovanni Delli Zotti<br>Department of Human<br>Sciences<br>University of Trieste, Italy<br>TOURISM                       |
| Lee Clarke<br>Rutgers University<br>SOCIETY AND<br>TECHNOLOGICAL RISKS   | Andrew L. Creighton<br>University of California,<br>Berkeley<br>DEMOCRACY   | N. J. Demerath<br>Department of Sociology<br>University of Massachusetts<br>SECULARIZATION                               |
-

LIST OF AUTHORS

- 
- |   |  |  |
|---|--|--|
| Linda Derksen<br>Department of Sociology<br>University of Alberta<br>SCIENTIFIC EXPLANATION                                       | Glen H. Elder<br>Carolina Population Center<br>University of North<br>Carolina, Chapel Hill<br>LIFE COURSE, THE  | Richard B. Felson<br>Department of Sociology<br>Pennsylvania State<br>University<br>SELF-CONCEPT                                     |
| William DiFazio<br>Department of Sociology<br>St. John's University<br>ECONOMIC DETERMINISM                                       | Laura L. Ellingson<br>Department of Sociology<br>University of South Florida<br>QUALITATIVE METHODS  | Robert A. Fiala<br>Department of Sociology<br>University of New Mexico<br>POSTINDUSTRIAL SOCIETY                                     |
| Don A. Dillman<br>Washington State University<br>INFORMATION SOCIETY  | Janice M. Engsberg<br>Department of Journalism<br>and Communication<br>Xiamen University, People's<br>Republic of China<br>MASS MEDIA RESEARCH                 | Thomas J. Figurski<br>Department of Psychology<br>Eastern Michigan University<br>MORAL DEVELOPMENT                                   |
| Thomas A. DiPrete<br>Department of Sociology<br>Duke University<br>SOCIAL MOBILITY  | Carroll L. Estes<br>Institute for Health and<br>Aging<br>University of California, San<br>Francisco<br>HEALTH POLICY ANALYSIS<br>MEDICAL-INDUSTRIAL<br>COMPLEX | Anne Foner<br>Rutgers University<br>STRUCTURAL LAG (AND<br>CULTURAL LAG)   |
| Mattei Dogan<br>National Center of Scientific<br>Research, Paris<br>SOCIOLOGY AMONG THE<br>SOCIAL SCIENCES<br>STATUS INCONGRUENCE | Amitai Etzioni<br>George Washington<br>University<br>COMMUNITARIANISM  | William Form<br>Department of Sociology<br>Ohio State University<br>LABOR MOVEMENTS AND<br>UNIONS                                    |
| Kevin Dougherty<br>Manhattan College<br>EDUCATIONAL<br>ORGANIZATION   | William M. Evan<br>Department of Sociology<br>University of Pennsylvania<br>LAW AND SOCIETY  | Martha Foschi<br>Department of<br>Anthropology and<br>Sociology<br>University of British<br>Columbia<br>EXPECTATION STATES<br>THEORY |
| Frank Dumont<br>Educational and Counseling<br>Psychology Department<br>McGill University<br>PERSONALITY THEORY                    | Bernard Farber<br>Department of Sociology<br>Arizona State University<br>KINSHIP SYSTEMS AND<br>FAMILY TYPES   | John Fox<br>Department of Sociology<br>McMaster University<br>STATISTICAL GRAPHICS   |
| Riley E. Dunlap<br>Department of Sociology<br>Washington State University<br>ENVIRONMENTAL<br>SOCIOLOGY                           | J Forbes Farmer<br>Division of Behavioral<br>Sciences<br>Franklin Pierce College<br>PENOLOGY   | Jan Fritz<br>University of Cincinnati<br>CLINICAL SOCIOLOGY  |
| Nancy E. Durbin<br>University of Washington<br>COMPARABLE WORTH   | Joe R. Feagin<br>Department of Sociology<br>University of Florida<br>AFRICAN-AMERICAN<br>STUDIES   | Frank F. Furstenberg<br>Department of Sociology<br>University of Pennsylvania<br>ILLEGITIMACY  |
| Alex Durig<br>Sociology Program<br>California State University,<br>San Marcos<br>QUALITATIVE MODELS                               |  | David Gartrell<br>Department of Sociology<br>University of Victoria<br>AGRICULTURAL<br>INNOVATION                                    |
-

LIST OF AUTHORS

- 
- |   |   |  |
|---|---|--|
| Alberto Gasparini<br>Dipartimento di Scienze<br>dell 'Uomo<br>Universita degli Studi di<br>Trieste, Italy<br>PREDICTION AND FUTURES<br>STUDIES<br>ITALIAN SOCIOLOGY | Leonard Gordon<br>Office for Academic<br>Programs<br>Arizona State University<br>STUDENT MOVEMENTS<br>PROTEST MOVEMENTS   | Avery M. Guest<br>University of Washington<br>DEMOGRAPHIC<br>TRANSITION  |
| Viktor Gecas<br>Department of Sociology<br>Washington State University<br>SOCIALIZATION   | Linda S. Gottfredson<br>Department of Educational<br>Studies<br>University of Delaware<br>INTELLIGENCE  | Christopher W. Gunn<br>University of Kansas<br>INTERNET  |
| Linda K. George<br>Center for the Study of<br>Aging and Human<br>Development<br>Duke University<br>AGING AND THE LIFE<br>COURSE                                     | Howard P. Greenwald<br>Sacramento Center, School<br>of Public Administration<br>University of Southern<br>California<br>ETHICS IN SOCIAL<br>RESEARCH<br>RESEARCH FUNDING IN<br>SOCIOLOGY<br>CIVIL LIBERTIES | Regan A. R. Gurung<br>Department of Psychology<br>University of Wisconsin<br>DEPRESSION  |
| Don C. Gibbons<br>Portland State University<br>CRIME, THEORIES OF   | Larry Griffin<br>Vanderbilt University<br>COMPARATIVE-HISTORICAL<br>SOCIOLOGY   | Jeffrey K. Hadden<br>Department of Sociology<br>University of Virginia<br>RELIGIOUS MOVEMENTS  |
| Mary Gilmore<br>School of Social Work<br>University of Washington<br>SEXUALLY TRANSMITTED<br>DISEASES   | Allen D. Grimshaw<br>Indiana University<br>SOCIOLINGUISTICS   | John Hagan<br>Faculty of Law<br>Northwestern University and<br>American Bar Foundation,<br>Chicago<br>JUVENILE DELINQUENCY,<br>THEORIES OF                           |
| Samuel L. Gilmore<br>Department of Sociology<br>University of California,<br>Irvine<br>CULTURE  | Edward Gross<br>Department of Sociology<br>University of Washington<br>COURT SYSTEMS AND LAW  | Lee J. Haggerty<br>Department of Sociology<br>Portland State University<br>URBAN SOCIOLOGY   |
| Nathan Glazer<br>Harvard University<br>AFFIRMATIVE ACTION   | David B. Grusky<br>Department of Sociology<br>Cornell University<br>SOCIAL STRATIFICATION   | Warren Hagstrom<br>Department of Sociology<br>University of Wisconsin,<br>Madison<br>SCIENCE   |
| Paul C. Glick<br>Arizona State University<br>MARRIAGE AND DIVORCE<br>RATES  | Oscar Grusky<br>Department of Sociology<br>University of California, Los<br>Angeles<br>INTERGROUP AND<br>INTERORGANIZATIONAL<br>RELATIONS   | Archibald O. Haller<br>Sociology and Rural<br>Sociology<br>University of Wisconsin,<br>Madison<br>SOCIAL STRATIFICATION<br>SOCIAL PSYCHOLOGY OF<br>STATUS ALLOCATION |
| Jack A. Goldstone<br>Department of Sociology<br>University of California,<br>Davis<br>STATE, THE  | Renzo Gubert<br>Dipartimento di Sociologia e<br>Ricerca Sociale<br>Universita di Trento, Italy<br>TERRITORIAL BELONGING   | Richard F. Hamilton<br>The Mershon Center<br>The Ohio State University<br>MASS SOCIETY   |
-



LIST OF AUTHORS

---

- |   |   |  |
|---|---|--|
| Janet Hankin<br>Wayne State University<br>LIFESTYLES AND HEALTH<br>POSITIVE MENTAL HEALTH   | R. Alan Hedley<br>Department of Sociology<br>University of Victoria<br>INDUSTRIALIZATION IN<br>LESS DEVELOPED<br>COUNTRIES<br>TRANSNATIONAL<br>CORPORATIONS | Dennis Hogan<br>Department of Sociology<br>Brown University<br>LIFE CYCLE  |
| Russell Hardin<br>Department of Political<br>Science<br>Stanford University<br>RATIONAL CHOICE THEORY   | David M. Heer<br>University of Southern<br>California<br>INTERNATIONAL<br>MIGRATION   | Charles E. Holzer III<br>Department of Psychiatry<br>and Behavioral Science<br>University of Texas<br>Medical Branch<br>EPIDEMIOLOGY                           |
| Melissa A. Hardy<br>Department of Sociology<br>Florida State University<br>RETIREMENT   | Karen A. Hegvedt<br>Department of Sociology<br>Emory University<br>SOCIAL COMPARISON<br>PROCESSES   | John Houghton<br>St. Edwards University<br>ADULT EDUCATION   |
| Lowell L. Hargens<br>Ohio State University<br>SAMPLING PROCEDURES<br>REPLICATION  | David R. Heise<br>Indiana University<br>AFFECT CONTROL THEORY<br>AND IMPRESSION<br>FORMATION  | Judith Howard<br>Department of Sociology<br>University of Washington<br>ATTRIBUTION THEORY<br>ALTRUISM   |
| Charlene Harrington<br>Social and Behavioral<br>Sciences<br>University of California, San<br>Francisco<br>HEALTH CARE UTILIZATION<br>AND EXPENDITURES     | Jon Hendricks<br>Department of Sociology<br>Oregon State University<br>DEPENDENCY THEORY  | Carla B. Howery<br>American Sociological<br>Association, Washington,<br>D.C.<br>AMERICAN SOCIOLOGICAL<br>ASSOCIATION AND OTHER<br>SOCIOLOGICAL<br>ASSOCIATIONS |
| Riaz Hassan<br>Department of Sociology<br>University of California, Los<br>Angeles and Flinders<br>University of South<br>Australia<br>SOCIOLOGY OF ISLAM | John Heritage<br>Department of Sociology<br>University of California, Los<br>Angeles<br>ETHNOMETHODOLOGY  | Marilyn Ihinger-Tallman<br>Department of Sociology<br>Washington State University<br>MARRIAGE  |
| Laurie Russell Hatch<br>Department of Sociology<br>University of Kentucky<br>AMERICAN FAMILIES  | Beth B. Hess<br>County College of Morris<br>INCOME DISTRIBUTION IN<br>THE UNITED STATES   | James M. Inverarity<br>Department of Sociology<br>Western Washington<br>University<br>SOCIOLOGY OF LAW   |
| Lawrence E. Hazelrigg<br>Department of Sociology<br>Florida State University<br>INDIVIDUALISM   | Christine L. Himes<br>Center for Policy Research<br>and Department of<br>Sociology<br>Syracuse University<br>STANDARDIZATION                                | Rita Jalali<br>Bethesda, Maryland<br>CASTE AND INHERITED<br>STATUS   |
| Todd F. Heatherton<br>Department of Psychology<br>Dartmouth College<br>SELF-ESTEEM  | Charles Hirschman<br>Russell Sage Foundation<br>University of Washington<br>SOUTHEAST ASIA STUDIES  | David R. James<br>Department of Sociology<br>Indiana University<br>SLAVERY AND<br>INVOLUNTARY SERVITUDE  |

LIST OF AUTHORS

- 
- |   |   |   |
|---|---|---|
| Carl Gunnar Janson<br>Department of Sociology<br>Stockholm University<br>SCANDINAVIAN SOCIOLOGY   | Howard B. Kaplan<br>Department of Sociology<br>Texas A & M University<br>SOCIAL PSYCHOLOGY  | Jae-On Kim<br>The Center for Asian and<br>Pacific Studies<br>University of Iowa<br>TABULAR ANALYSIS       |
| J. Craig Jenkins<br>Department of Sociology<br>Ohio State University<br>POLITICAL SOCIOLOGY   | Tracy Karner<br>Gerontology Center<br>University of Kansas<br>SOCIAL CAPITAL  | Paul W. Kingston<br>Department of Sociology<br>University of Virginia<br>SOCIOLOGY OF EDUCATION           |
| Brian Michael Jenkins<br>Kroll Associates, Los<br>Angeles<br>TERRORISM  | David R. Karp<br>Department of Sociology<br>Skidmore College<br>VALUE THEORY AND<br>RESEARCH  | Edgar Kiser<br>Department of Sociology<br>University of Washington<br>WAR                                 |
| Leif Jensen<br>Agricultural Economics and<br>Rural Sociology<br>Pennsylvania State<br>University<br>MARGINAL EMPLOYMENT                           | John D. Kasarda<br>University of North<br>Carolina, Chapel Hill<br>URBAN UNDERCLASS   | Carl B. Klockars<br>Department of Sociology<br>University of Delaware<br>POLICE                           |
| Timothy D. Jewell<br>Allen Library<br>University of Washington<br>LIBRARY RESOURCES AND<br>SERVICES FOR SOCIOLOGY                                 | John M. Katsillis<br>Department of Education<br>University of Patras, Greece<br>EDUCATION AND MOBILITY  | David Knoke<br>Department of Sociology<br>University of Minnesota<br>POLITICAL<br>ORGANIZATIONS           |
| Blair T. Johnson<br>University of Connecticut<br>META-ANALYSIS  | John R. Kelly<br>Department of Sociology<br>University of Illinois<br>LEISURE   | Igor S. Kon<br>Institute of Ethnography,<br>Moscow<br>SOVIET AND POST-SOVIET<br>SOCIOLOGY                 |
| Anupama Joshi<br>Department of Child and<br>Family Studies<br>California State University,<br>Los Angeles<br>INTERPERSONAL CONFLICT<br>RESOLUTION | Theodore D. Kemper<br>Department of Sociology<br>St. John's University<br>EMOTIONS  | Karl Kosloski<br>Department of Gerontology<br>University of Nebraska<br>EVALUATION RESEARCH               |
| Eva Kahana<br>Department of Sociology<br>Case Western Reserve<br>University<br>LONG TERM CARE<br>FACILITIES                                       | Stephen A. Kent<br>Department of Sociology<br>University of Alberta<br>HISTORICAL SOCIOLOGY   | Louis Kriesberg<br>Maxwell School of<br>Citizenship and Public<br>Affairs<br>Syracuse University<br>PEACE |
| Yoshinori Kamo<br>Department of Sociology<br>Louisiana State University<br>MARITAL ADJUSTMENT   | Kyle Kercher<br>Department of Sociology<br>Case Western Reserve<br>University<br>QUASI-EXPERIMENTAL<br>RESEARCH DESIGNS<br>MULTIPLE INDICATOR<br>MODELS | Manfred Kuechler<br>Department of Sociology<br>Hunter College, CUNY<br>VOTING BEHAVIOR                    |
|   | Richard B. Kettner-Polley<br>Business Administration<br>Jones International<br>University<br>OBSERVATION SYSTEMS<br>FIELD THEORY                        | Daphne Kuo<br>Department of Sociology<br>University of Washington<br>DESCRIPTIVE STATISTICS               |
-

LIST OF AUTHORS

- 
- |  |   |  |
|--|---|--|
| Mary Riege Laner<br>Department of Sociology<br>Arizona State University<br>COURTSHIP                             | William T. Liu<br>National University of<br>Singapore and University<br>of Illinois, Chicago<br>ASIAN AMERICAN STUDIES                            | Kyriakos S. Markides<br>Division of Sociomedical<br>Sciences<br>University of Texas Medical<br>Branch, Galveston<br>QUALITY OF LIFE        |
| Gladys Engel Lang<br>University of Washington<br>ART AND SOCIETY   | Clarence Y. H. Lo<br>Department of Sociology<br>University of Missouri<br>ALIENATION  | Barry Markovsky<br>Program Director for<br>Sociology<br>University of Iowa<br>SOCIAL PERCEPTION  |
| Kurt Lang<br>University of Washington<br>PUBLIC OPINION  | John R. Logan<br>Department of Sociology<br>State University of New<br>York<br>SUBURBANIZATION  | Cora B. Marrett<br>CORPORATE<br>ORGANIZATIONS  |
| Barrett A. Lee<br>Department of Sociology<br>Pennsylvania State<br>University<br>HOMELESSNESS                    | J. Scott Long<br>Department of Sociology<br>Indiana University<br>STATISTICAL METHODS   | Peter V. Marsden<br>Department of Sociology<br>Harvard University<br>SOCIAL NETWORKS   |
| Gary R. Lee<br>Department of Sociology<br>Bowling Green State<br>University<br>FAMILY AND HOUSEHOLD<br>STRUCTURE | Charles F. Longino, Jr.<br>Wake Forest University<br>INTERNAL MIGRATION   | Kerry Marsh<br>Department of Psychology<br>University of Connecticut<br>PERSONAL AUTONOMY  |
| Janet Lever<br>Department of Behavioral<br>Sciences<br>California State University,<br>Los Angeles<br>SPORT      | Joseph Lopreato<br>University of Texas<br>SOCIOBIOLOGY, HUMAN   | Guido Martinotti<br>Departimento di Sociologia<br>Universita degli Studi di<br>Milano<br>DATA BANKS AND<br>DEPOSITORIES                    |
| Nan Lin<br>Department of Sociology<br>Duke University<br>STRESS<br>SOCIAL RESOURCES<br>THEORY                    | Judith Lorber<br>City University of New York<br>and Brooklyn College<br>GENDER  | Gerald Marwell<br>Department of Sociology<br>University of Wisconsin<br>EXPERIMENTS  |
| Robert C. Liebman<br>Department of Sociology<br>Portland State University<br>REVOLUTIONS                         | David R. Maines<br>Department of Sociology<br>Oakland University<br>PRAGMATISM<br>LIFE HISTORIES AND<br>NARRATIVE                                 | Alexandra Maryanski<br>Department of Sociology<br>University of California,<br>Riverside<br>FUNCTIONALISM AND<br>STRUCTURALISM             |
| Seymour Martin Lipset<br>INTELLECTUALS   | Margaret Mooney Marini<br>Department of Sociology<br>University of Minnesota<br>OCCUPATIONAL AND<br>CAREER MOBILITY<br>SOCIAL VALUES AND<br>NORMS | Eleonora Barbieri Masini<br>Faculty of Social Sciences<br>Gregorian University, Rome<br>FUTURES STUDIES AS<br>HUMAN AND SOCIAL<br>ACTIVITY |
| Allen E. Liska<br>SOCIAL CONTROL   | Ronald W. Maris<br>Center for the Study of<br>Suicide<br>University of South Carolina<br>SUICIDE  |  |
-

LIST OF AUTHORS

- 
- |  |   |   |
|--|---|---|
| Douglas S. Massey<br>Department of Sociology<br>University of Pennsylvania<br>SEGREGATION INDICES  | Mario Mikulincer<br>Department of Psychology<br>Bar-Ilan University, Israel<br>PERSONAL DEPENDENCY              | Keiko Nakao<br>Department of Sociology<br>Tokyo Metropolitan<br>University<br>OCCUPATIONAL PRESTIGE   |
| Armand L. Mauss<br>Department of Sociology<br>Washington State University<br>SOCIAL PROBLEMS   | Gale E. Miller<br>Department of Social and<br>Cultural Sciences<br>Marquette University<br>CASE STUDIES         | Todd D. Nelson<br>Psychology Department<br>California State<br>University, Turlock<br>COGNITIVE CONSISTENCY<br>THEORIES   |
| E. Doyle McCarthy<br>Department of Sociology<br>Fordham University<br>SOCIOLOGY OF<br>KNOWLEDGE  | Rhonda J. V. Montgomery<br>Gerontology Center<br>University of Kansas<br>FILIAL RESPONSIBILITY                  | David G. Nickinovich<br>NRC, Inc.<br>BUREAUCRACY  |
| Lisa J. McIntyre<br>Department of Sociology<br>Washington State University<br>FAMILY LAW   | Patrick H. Mooney<br>CAPITALISM   | Steven L. Nock<br>Department of Sociology<br>University of Virginia<br>DIVORCE  |
| James M. McPartland<br>Johns Hopkins University<br>EQUALITY OF<br>OPPORTUNITY  | Raymond A. Morrow<br>Department of Sociology<br>University of Alberta<br>MARXIST SOCIOLOGY                      | Robert M. O'Brien<br>Department of Sociology<br>University of Oregon<br>LEVELS OF ANALYSIS<br>CRIME RATES   |
| Susan McWilliams<br>Department of Sociology<br>University of Southern<br>Maine<br>PERSUASION   | Jeylan T. Mortimer<br>Life Course Center<br>University of Minnesota<br>ADULTHOOD                                | Mary Ellen O'Connell<br>INTERNATIONAL LAW   |
| David Mechanic<br>Institute for Health, Health<br>Care Policy, and Aging<br>Research<br>Rutgers University<br>HEALTH AND ILLNESS<br>BEHAVIOR | Charles W. Mueller<br>Obermann Center for<br>Advanced Studies<br>University of Iowa<br>WORK ORIENTATION         | Angela M. O'Rand<br>Department of Sociology<br>Duke University<br>SOCIAL INEQUALITY   |
| Barbara F. Meeker<br>Department of Sociology<br>University of Maryland,<br>College Park<br>MATHEMATICAL<br>SOCIOLOGY                         | Chandra Mukerji<br>Department of<br>Communications<br>University of California,<br>San Diego<br>POPULAR CULTURE | Richard Ofshe<br>Department of Sociology<br>University of California,<br>Berkeley<br>EXTREME INFLUENCE:<br>THOUGHT REFORM, HIGH<br>CONTROL GROUPS,<br>INTERROGATION, AND<br>RECOVERED MEMORY<br>PSYCHOTHERAPY |
| H. Andrew Michener<br>Department of Sociology<br>University of Wisconsin,<br>Madison<br>GAME THEORY AND<br>STRATEGIC INTERACTION             | Joane Nagel<br>Department of Sociology<br>University of Kansas<br>INDIGENOUS PEOPLES                            | Valerie K. Oppenheimer<br>Department of Sociology<br>University of California, Los<br>Angeles<br>LABOR FORCE  |
|  | Saad Nagi<br>The Ohio State University<br>NATIONALISM   |   |
-

LIST OF AUTHORS

- 
- |   |   |  |
|---|---|--|
| Myron Orleans<br>Department of Sociology<br>California State University,<br>Fullerton<br>PHENOMENOLOGY  | W. David Pierce<br>Department of Sociology<br>University of Alberta<br>BEHAVIORISM  | Sean Ó Riain<br>University of California,<br>Davis<br>GLOBALIZATION AND<br>GLOBAL SYSTEMS<br>ANALYSIS                                      |
| Suzanne T. Ortega<br>Department of Sociology<br>University of Nebraska,<br>Lincoln<br>MENTAL ILLNESS AND<br>MENTAL DISORDERS                          | Karl Pillemer<br>Department of Human<br>Development<br>Cornell University<br>INTERGENERATIONAL<br>RELATIONS                           | Maurice N. Richter, Jr.<br>Sociology Department<br>State University of New<br>York at Albany<br>EVOLUTION: BIOLOGICAL,<br>SOCIAL, CULTURAL |
| Fred C. Pampel<br>Department of Sociology<br>University of Colorado<br>SOCIAL SECURITY SYSTEMS  | Susan R. Pitchford<br>University of Washington<br>RACE  | Cecilia L. Ridgeway<br>Department of Sociology<br>Stanford University<br>COMPLIANCE AND<br>CONFORMITY                                      |
| Jeffrey G. Pankhurst<br>Department of Sociology<br>Wittenberg University<br>FAMILY AND RELIGION   | Gabriele Pollini<br>Dipartimento di Sociologia e<br>Ricerca Sociale<br>Universita degli Studi di<br>Trento, Italy<br>SOCIAL BELONGING | Matilda White Riley<br>National Institutes of Health<br>COHORT PERSPECTIVES  |
| Toby L. Parcel<br>Social and Behavioral<br>Sciences<br>Ohio State University<br>SECONDARY DATA<br>ANALYSIS AND DATA<br>ARCHIVES                       | Lin Poyer<br>Department of<br>Anthropology<br>University of Wyoming<br>SOCIOCULTURAL<br>ANTHROPOLOGY                                  | John W. Riley, Jr.<br>Chevy Chase, Maryland<br>DEATH AND DYING   |
| Ronald W. Perry<br>Arizona State University<br>DIFFUSION THEORIES   | Evan Pritchard<br>Department of Psychology<br>University of Winnipeg<br>DECISION-MAKING THEORY<br>AND RESEARCH                        | Georgina Rojas-Garcia<br>Population Research Center<br>University of Texas<br>INFANT AND CHILD<br>MORTALITY                                |
| Richard A. Peterson<br>Department of Sociology<br>and Social Policy<br>University of Leeds<br>MUSIC   | Thomas W. Pullum<br>Department of Sociology<br>University of Texas<br>POPULATION  | Patricia A. Roos<br>Rutgers University<br>PROFESSIONS  |
| Thomas F. Pettigrew<br>Stevenson College<br>University of California,<br>Santa Cruz<br>DISCRIMINATION   | E. L. Quarantelli<br>Disaster Research Center<br>University of Delaware<br>DISASTER RESEARCH  | Steve Rytina<br>Department of Sociology<br>McGill University<br>SOCIAL STRUCTURE   |
| David R. Phillips<br>Asia-Pacific Institute of<br>Ageing Studies, Faculty of<br>Social Sciences<br>Lingnan University,<br>Hong Kong<br>LONG-TERM CARE | Barbara F. Reskin<br>Department of Sociology<br>Harvard University<br>WORK AND OCCUPATIONS  | George Sabagh<br>Department of Sociology<br>University of California<br>MIDDLE EASTERN STUDIES   |
|   |   | Joseph Sanders<br>University of Houston<br>LAW AND LEGAL SYSTEMS   |
|   |   | Masamichi Sasaki<br>Hyogo Kyoiku University,<br>Japan<br>JAPANESE SOCIOLOGY  |
-

LIST OF AUTHORS

- 
- |   |   |   |
|---|---|---|
| Gerard Saucier<br>Department of Psychology<br>University of Oregon<br>PERSONALITY<br>MEASUREMENT                                  | Constance L. Shehan<br>Department of Sociology<br>University of Florida<br>PARENTAL ROLES                             | Subhash Sonnad<br>Department of Sociology<br>Western Michigan University<br>NONPARAMETRIC<br>STATISTICS   |
| Erwin K. Scheuch<br>GERMAN SOCIOLOGY  | Larry D. Shinn<br>Berea College<br>WORLD RELIGIONS  | Kenneth I. Spenner<br>Department of Sociology<br>Duke University<br>PERSONALITY AND SOCIAL<br>STRUCTURE   |
| David R. Schmitt<br>Department of Sociology<br>University of Washington<br>GROUP SIZE   | James F. Short<br>Department of Sociology<br>Washington State University<br>CRIMINAL AND<br>DELINQUENT<br>SUBCULTURES | Rodney Stark<br>Department of Sociology<br>University of Washington<br>SOCIOLOGY OF RELIGION  |
| Ronald J. Schoenberg<br>Department of Sociology<br>University of Washington<br>PROBABILITY THEORY                                 | Robert W. Shotola<br>Department of Sociology<br>Portland State University<br>SMALL GROUPS                             | Lauri Steel<br>American Institutes for<br>Research<br>FAMILY SIZE   |
| Howard Schuman<br>Institute for Social Research<br>University of Michigan<br>SURVEY RESEARCH                                      | William Simon<br>Department of Sociology<br>University of Houston<br>SEXUAL ORIENTATION                               | Ardyth Stimson<br>Kean University<br>INTERPERSONAL<br>ATTRACTION  |
| Joseph Scimecca<br>Department of Sociology<br>George Mason University<br>HUMANISM   | John H. Simpson<br>Department of Sociology<br>University of Toronto<br>RELIGION, POLITICS, AND<br>WAR                 | John Stimson<br>Sociology Department<br>William Paterson University<br>SOCIAL FORECASTING   |
| David Sciulli<br>Department of Sociology<br>Texas A & M University<br>INTERNATIONAL<br>ASSOCIATIONS IN<br>SOCIOLOGY               | Richard Mcgrath Skinner<br>University of Virginia<br>POLITICAL AND<br>GOVERNMENTAL<br>CORRUPTION                      | Philip J. Stone<br>Department of Sociology<br>Harvard University<br>CONTENT ANALYSIS  |
| Karen Seccombe<br>Department of Sociology<br>Portland State University<br>ALTERNATIVE LIFESTYLES                                  | David Smith<br>University of Kansas<br>GENOCIDE   | David Strang<br>Center for Advanced Study<br>in the Behavioral Sciences<br>Cornell University<br>IMPERIALISM,<br>COLONIALISM, AND<br>DECOLONIZATION |
| Vimal P. Shah<br>Gujarat University, India<br>INDIAN SOCIOLOGY  | C. Matthew Snipp<br>Department of Sociology<br>Stanford University<br>AMERICAN INDIAN STUDIES                         | Jeffrey J. Strange<br>Public Insight<br>CENSORSHIP AND THE<br>REGULATION OF<br>EXPRESSION   |
| Michael J. Shanahan<br>Center for Advanced Study<br>in the Behavioral Sciences<br>Pennsylvania State<br>University<br>ADOLESCENCE | David Snow<br>Department of Sociology<br>University of Arizona<br>CROWDS AND RIOTS                                    |   |
-

LIST OF AUTHORS

- 
- |   |  |  |
|---|--|--|
| Raimondo Strassoldo<br>Dipartimento EST<br>University of Udine, Italy<br>NATIONAL BORDER<br>RELATIONS,<br>MANAGEMENT OF                       | Richard Swedberg<br>Department of Sociology<br>Stockholm University<br>ECONOMIC SOCIOLOGY  | Marta Tienda<br>Office of Population<br>Research<br>Princeton University<br>HISPANIC AMERICANS                           |
| Murray A. Straus<br>University of New<br>Hampshire<br>FAMILY VIOLENCE   | Szonja Szelenyi<br>Department of Sociology<br>Cornell University<br>SOCIALISM AND<br>COMMUNISM                                       | Jackson Toby<br>Sociology Department<br>The State University of New<br>Jersey, Rutgers<br>CRIMINALIZATION OF<br>DEVIANCE |
| Robin Stryker<br>Department of Sociology<br>University of Iowa<br>GOVERNMENT<br>REGULATION  | Jacek Sztatka<br>Jagiellonian University,<br>Poland<br>POLISH AND EASTERN<br>EUROPEAN SOCIOLOGY                                      | Donald J. Treiman<br>Department of Sociology<br>University of California at<br>Los Angeles<br>STATUS ATTAINMENT          |
| Sheldon Stryker<br>Department of Sociology<br>Indiana University<br>IDENTITY THEORY<br>SYMBOLIC INTERACTION<br>THEORY                         | Karl E. Taeuber<br>Department of Sociology<br>University of Wisconsin<br>CENSUS  | Gaye Tuchman<br>Department of Sociology<br>University of Connecticut<br>FEMINIST THEORY                                  |
| Donald E. Stull<br>Department of Sociology<br>University of Akron<br>HEALTH AND THE LIFE<br>COURSE<br>INTERGENERATIONAL<br>RESOURCE TRANSFERS | Irving Tallman<br>Department of Sociology<br>Washington State University<br>FAMILY POLICY IN<br>WESTERN SOCIETIES                    | Nancy Brandon Tuma<br>Department of Sociology<br>Stanford University<br>SOCIAL DYNAMICS                                  |
| J. Jill Suitor<br>Department of Sociology<br>Louisiana State University<br>REMARRIAGE   | Marylee C. Taylor<br>Department of Sociology<br>Pennsylvania State<br>University<br>PREJUDICE  | Austin T. Turk<br>Department of Sociology<br>University of California,<br>Riverside<br>POLITICAL CRIME                   |
| Teresa A. Sullivan<br>University of Texas at Austin<br>BANKRUPTCY AND CREDIT  | James E. Teele<br>Department of Sociology<br>Boston University<br>JUVENILE DELINQUENCY<br>(AND JUVENILE CRIME)<br>WHITE COLLAR CRIME | Herman Turk<br>SOCIAL ORGANIZATION   |
| Deborah A. Sullivan<br>Department of Sociology<br>Arizona State University<br>BIRTH AND DEATH RATES   | Bruno Tellia<br>Dipartimento Economia,<br>Società, Territorio<br>Università di Udine, Italy<br>NEGOTIATION OF POWER                  | Jonathan H. Turner<br>Department of Sociology<br>University of California,<br>Riverside<br>POSITIVISM                    |
| Gene F. Summers<br>Department of Rural<br>Sociology<br>University of Wisconsin,<br>Madison<br>RURAL SOCIOLOGY                                 | Ann R. Tickamyer<br>Department of Sociology<br>The Ohio University<br>MACROSOCIOLOGY<br>INFORMAL ECONOMY, THE                        | Ralph H. Turner<br>University of California, Los<br>Angeles<br>COLLECTIVE BEHAVIOR                                       |
| Marvin B. Sussman<br>INHERITANCE  |  | Bernard Valade<br>The Sorbonne<br>FRENCH SCHOOL OF<br>SOCIOLOGY, THE   |
-

LIST OF AUTHORS

---

- |   |   |   |
|---|---|---|
| Gerard Saucier<br>Department of Psychology<br>University of Oregon<br>PERSONALITY<br>MEASUREMENT                                  | Constance L. Shehan<br>Department of Sociology<br>University of Florida<br>PARENTAL ROLES                             | Subhash Sonnad<br>Department of Sociology<br>Western Michigan University<br>NONPARAMETRIC<br>STATISTICS   |
| Erwin K. Scheuch<br>GERMAN SOCIOLOGY  | Larry D. Shinn<br>Berea College<br>WORLD RELIGIONS  | Kenneth I. Spenner<br>Department of Sociology<br>Duke University<br>PERSONALITY AND SOCIAL<br>STRUCTURE   |
| David R. Schmitt<br>Department of Sociology<br>University of Washington<br>GROUP SIZE   | James F. Short<br>Department of Sociology<br>Washington State University<br>CRIMINAL AND<br>DELINQUENT<br>SUBCULTURES | Rodney Stark<br>Department of Sociology<br>University of Washington<br>SOCIOLOGY OF RELIGION  |
| Ronald J. Schoenberg<br>Department of Sociology<br>University of Washington<br>PROBABILITY THEORY                                 | Robert W. Shotola<br>Department of Sociology<br>Portland State University<br>SMALL GROUPS                             | Lauri Steel<br>American Institutes for<br>Research<br>FAMILY SIZE   |
| Howard Schuman<br>Institute for Social Research<br>University of Michigan<br>SURVEY RESEARCH                                      | William Simon<br>Department of Sociology<br>University of Houston<br>SEXUAL ORIENTATION                               | Sonia Stefanizzi<br>Department of Sociology<br>and Social Research<br>University of Milan-Bicocca<br>DATABANKS AND<br>DEPOSITORIES                  |
| Joseph Scimecca<br>Department of Sociology<br>George Mason University<br>HUMANISM   | John H. Simpson<br>Department of Sociology<br>University of Toronto<br>RELIGION, POLITICS, AND<br>WAR                 | Ardyth Stimson<br>Kean University<br>INTERPERSONAL<br>ATTRACTION  |
| David Sciulli<br>Department of Sociology<br>Texas A & M University<br>INTERNATIONAL<br>ASSOCIATIONS IN<br>SOCIOLOGY               | Richard Mcgrath Skinner<br>University of Virginia<br>POLITICAL AND<br>GOVERNMENTAL<br>CORRUPTION                      | John Stimson<br>Sociology Department<br>William Paterson University<br>SOCIAL FORECASTING   |
| Karen Seccombe<br>Department of Sociology<br>Portland State University<br>ALTERNATIVE LIFESTYLES                                  | David Smith<br>University of Kansas<br>GENOCIDE   | Philip J. Stone<br>Department of Sociology<br>Harvard University<br>CONTENT ANALYSIS  |
| Vimal P. Shah<br>Gujarat University, India<br>INDIAN SOCIOLOGY  | C. Matthew Snipp<br>Department of Sociology<br>Stanford University<br>AMERICAN INDIAN STUDIES                         | David Strang<br>Center for Advanced Study<br>in the Behavioral Sciences<br>Cornell University<br>IMPERIALISM,<br>COLONIALISM, AND<br>DECOLONIZATION |
| Michael J. Shanahan<br>Center for Advanced Study<br>in the Behavioral Sciences<br>Pennsylvania State<br>University<br>ADOLESCENCE | David Snow<br>Department of Sociology<br>University of Arizona<br>CROWDS AND RIOTS                                    |   |



LIST OF AUTHORS

- 
- |   |  |  |
|---|--|--|
| Jeffrey J. Strange<br>Public Insight<br>CENSORSHIP AND THE<br>REGULATION OF<br>EXPRESSION   | Gene F. Summers<br>Department of Rural<br>Sociology<br>University of Wisconsin,<br>Madison<br>RURAL SOCIOLOGY                        | Bruno Tella<br>Dipartimento Economia,<br>Società, Territorio<br>Universita di Udine, Italy<br>NEGOTIATION OF POWER       |
| Raimondo Strassoldo<br>Dipartimento EST<br>University of Udine, Italy<br>NATIONAL BORDER<br>RELATIONS,<br>MANAGEMENT OF                       | Marvin B. Sussman<br>INHERITANCE   | Ann R. Tickamyer<br>Department of Sociology<br>The Ohio University<br>MACROSOCIOLOGY<br>INFORMAL ECONOMY, THE            |
| Murray A. Straus<br>University of New<br>Hampshire<br>FAMILY VIOLENCE   | Richard Swedberg<br>Department of Sociology<br>Stockholm University<br>ECONOMIC SOCIOLOGY  | Marta Tienda<br>Office of Population<br>Research<br>Princeton University<br>HISPANIC AMERICANS                           |
| Robin Stryker<br>Department of Sociology<br>University of Iowa<br>GOVERNMENT<br>REGULATION  | Szonja Szelenyi<br>Department of Sociology<br>Cornell University<br>SOCIALISM AND<br>COMMUNISM                                       | Jackson Toby<br>Sociology Department<br>The State University of New<br>Jersey, Rutgers<br>CRIMINALIZATION OF<br>DEVIANCE |
| Sheldon Stryker<br>Department of Sociology<br>Indiana University<br>IDENTITY THEORY<br>SYMBOLIC INTERACTION<br>THEORY                         | Jacek Szmataka<br>Jagiellonian University,<br>Poland<br>POLISH AND EASTERN<br>EUROPEAN SOCIOLOGY                                     | Donald J. Treiman<br>Department of Sociology<br>University of California at<br>Los Angeles<br>STATUS ATTAINMENT          |
| Donald E. Stull<br>Department of Sociology<br>University of Akron<br>HEALTH AND THE LIFE<br>COURSE<br>INTERGENERATIONAL<br>RESOURCE TRANSFERS | Karl E. Taeuber<br>Department of Sociology<br>University of Wisconsin<br>CENSUS  | Gaye Tuchman<br>Department of Sociology<br>University of Connecticut<br>FEMINIST THEORY                                  |
| J. Jill Suitor<br>Department of Sociology<br>Louisiana State University<br>INTERGENERATIONAL<br>RELATIONS<br>REMARRIAGE                       | Irving Tallman<br>Department of Sociology<br>Washington State University<br>FAMILY POLICY IN<br>WESTERN SOCIETIES                    | Nancy Brandon Tuma<br>Department of Sociology<br>Stanford University<br>SOCIAL DYNAMICS                                  |
| Teresa A. Sullivan<br>University of Texas at Austin<br>BANKRUPTCY AND CREDIT  | Marylee C. Taylor<br>Department of Sociology<br>Pennsylvania State<br>University<br>PREJUDICE  | Austin T. Turk<br>Department of Sociology<br>University of California,<br>Riverside<br>POLITICAL CRIME                   |
| Deborah A. Sullivan<br>Department of Sociology<br>Arizona State University<br>BIRTH AND DEATH RATES   | James E. Teele<br>Department of Sociology<br>Boston University<br>JUVENILE DELINQUENCY<br>(AND JUVENILE CRIME)<br>WHITE COLLAR CRIME | Herman Turk<br>SOCIAL ORGANIZATION   |
|   |  | Jonathan H. Turner<br>Department of Sociology<br>University of California,<br>Riverside<br>POSITIVISM                    |
-

LIST OF AUTHORS

- 
- |   |  |   |
|---|--|---|
| Ralph H. Turner<br>University of California, Los Angeles<br>COLLECTIVE BEHAVIOR   | Frederick D. Weil<br>Department of Sociology<br>Louisiana State University<br>POLITICAL PARTY SYSTEMS  | Alan Wolfe<br>Boston College<br>SOCIAL PHILOSOPHY<br>HUMAN NATURE   |
| Bernard Valade<br>The Sorbonne<br>FRENCH SCHOOL OF SOCIOLOGY, THE   | Joseph G. Weis<br>Department of Sociology<br>University of Washington<br>PROBATION AND PAROLE  | James R. Wood<br>Department of Sociology<br>Indiana University<br>RELIGIOUS ORGANIZATIONS                                     |
| Maurice D. Van Arsdol, Jr.<br>Institute of International Studies<br>University of Southern California and Monterey CITIES | Jack Whalen<br>Xerox Palo Alto Research Center<br>CONVERSATION ANALYSIS  | James D. Wright<br>Department of Sociology<br>Tulane University<br>PUBLIC POLICY ANALYSIS                                     |
| Pierre L. van den Berghe<br>Department of Sociology<br>University of Washington<br>POLITICAL CORRECTNESS                  | Martin K. Whyte<br>Department of Sociology<br>George Washington University<br>CHINA STUDIES  | Robert Wuthnow<br>Department of Sociology<br>Princeton University<br>RELIGIOUS ORIENTATIONS                                   |
| Brenda J. Vander Mey<br>Department of Sociology<br>Clemson University<br>INCEST   | Marion C. Willetts<br>Department of Sociology and Anthropology<br>Illinois State University<br>SEXUAL BEHAVIOR IN MARRIAGE AND CLOSE RELATIONSHIPS | Grace J. Yoo<br>Asian-American Studies Department<br>San Francisco State University<br>INTERMARRIAGE                          |
| Wayne J. Vilemez<br>Department of Sociology<br>University of Connecticut<br>POVERTY                                       | Robin M. Williams, Jr.<br>University of California, Irvine and Cornell University<br>AMERICAN SOCIETY  | John H. Yost<br>Department of Psychology<br>John Carroll University<br>ROLE THEORY: FOUNDATIONS, EXTENSIONS, AND APPLICATIONS |
| Carlos H. Waisman<br>Department of Sociology<br>University of California, San Diego<br>LATIN AMERICAN STUDIES             | Charles Winick<br>The Graduate School and University Center<br>City University of New York<br>PORNOGRAPHY<br>DRUG ABUSE                            | Rosalie F. Young<br>Department of Community Medicine<br>Wayne State University<br>HEALTH PROMOTION AND HEALTH STATUS          |
| Walter L. Wallace<br>Department of Sociology<br>Princeton University<br>METATHEORY  | Halliman H. Winsborough<br>DEMOGRAPHY  | Yih-Jin Young<br>Nassau Community College<br>EDUCATION AND DEVELOPMENT  |
| Marilyn E. Walsh<br>ORGANIZED CRIME   | Christopher Winship<br>Department of Sociology<br>Harvard University<br>SAMPLE SELECTION BIAS  | Morris Zelditch, Jr.<br>Department of Sociology<br>Stanford University<br>INTERPERSONAL POWER                                 |
| Carol A. B. Warren<br>Department of Sociology<br>University of Kansas<br>ETHNOGRAPHY                                      | Lynne Woehrle<br>Wilson College<br>GROUP CONFLICT RESOLUTION   |   |
| Susan Cotts Watkins<br>Department of Sociology<br>University of Pennsylvania<br>FERTILITY DETERMINANTS                    |  |   |
-

LIST OF AUTHORS

---

Viviana Zelizer  
Department of Sociology  
Princeton University  
MONEY

Richard A. Zeller  
Department of Sociology  
Bowling Green State  
University  
RELIABILITY

Mary Zimmerman  
Health Policy and  
Management  
University of Kansas  
MEDICAL SOCIOLOGY  
COMPARATIVE HEALTH-  
CARE SYSTEMS

# Editorial Board

---

**Editor-in-Chief**, Edgar F. Borgatta  
**Managing Editor**, Rhonda J. V. Montgomery  
**Consulting Editor**, Marie L. Borgatta

## ADVISORY EDITORS

Duane Alwin  
George W. Bohrnstedt  
Karen S. Cook  
Herbert L. Costner  
Eileen Crimmons  
William D'Antonio  
Doris R. Entwistle  
Amatai Etzioni  
Archibald O. Haller  
Maureen Hallinan  
David Heise  
Beth B. Hess  
Charles Hirschman  
Joan Huber  
Yoshinori Kamo  
David Knoke  
Gladys E. Lang  
Stanley Lieberman  
Seymour Martin Lipset  
J. Scott Long  
Coral B. Marrett

David Mechanic  
Margaret Mooney Marini  
Peter V. Marsden  
Douglas S. Massey  
Jill Quadagno  
Matilda White Riley  
Pepper Schwartz  
William H. Sewell  
James F. Short, Jr.  
Teresa A. Sullivan  
James Teele  
Marta Tienda  
Nancy B. Tuma  
Robin M. Williams, Jr.

## INTERNATIONAL ADVISORY EDITORS

Vaneeta D'Andrea (Great Britain)  
Mattei Dogan (France)  
Alberto Gasperini (Italy)  
Carl-Gunner Janson (Scandinavia)  
Igor S. Kon (Soviet)  
Judah Matra (Israel)  
Georges Sabagh (Middle East)  
Masamichi Sasaki (Japan)  
Jacek Szmata (Poland, Eastern Europe)  
Vimal Shah (India)

# A

---

## **ABORTION**

*See* Family Planning; Pregnancy and Pregnancy Termination.

## **ACCULTURATION**

*See* Ethnicity.

## **ADDICTION**

*See* Alcohol; Drug Abuse.

## **ADOLESCENCE**

Recognition of the life stage between childhood and adulthood as a subject of modern scientific inquiry began in the early twentieth century with the publication of Antonio Marro's *La Puberta* (1898) and G. Stanley Hall's highly influential compendium *Adolescence* (1904). Although Hall's book represented an initial effort to describe adolescence, it nevertheless resonated with themes already familiar among scholars and the public. In Europe, romantic conceptions of a sexually charged, troubled youth (e.g., in Rousseau's *Émile*) circulated among the socially concerned. In America, an established tradition of cautionary literature emphasized the impressionable nature of young people and their vulnerability to sin (e.g., in the essays and sermons of Cotton Mather). Hall incorporated many of these ideas into a Darwinian framework to conjure an "adolescence" recognizable to

his readers (Ross 1972). Although the work is viewed as a curious and difficult amalgam today, it nevertheless emphasized themes that continue to shape the study of youth.

Hall viewed adolescence through the lens of Ernst Haeckel's biogenetic principle, which holds that the human life span recapitulates the phases of human biological and social evolution (Gould 1977). Hall maintained that late childhood corresponds to a period of peaceful savagery in the distant past, whereas adolescence represents a "neo-atavistic" period of migration into a challenging environment, which prompted physical, social, and psychological conflict and growth. This characterization of the adolescent, as troubled by all-encompassing turmoil, was contested early in the twentieth century by prominent behavioral scientists such as Edwin Thorndike (1917) and has been repeatedly challenged since then, perhaps most famously by Margaret Mead's *Coming of Age in Samoa* (1928) (but see Freeman 1983; Côté 1994). Likewise, sociologists such as Robert and Helen Lynd (1929) and August Hollingshead (1949) found little evidence of pervasive trouble among the youth of Middletown or Elmtown. Contemporary behavioral scientists take a more moderate view than Hall's, depicting adolescence as a time of both change and continuity (e.g., Douvan and Adelson 1966). Nevertheless, the study of adolescence has been indelibly marked by the "storm and stress" motif.

Hall also maintained that adolescents are highly responsive to adult guidance. Drawing on work by

Edward Cope, a leading American proponent of the biogenetic principle, he believed that the influence of the environment in producing acquired characteristics that were then transmissible by heredity was greatest during adolescence. The implication of this Lamarckian view was momentous: The future development of the human race depended on improvements in the adolescent (Hall 1904, v. 1, p. 50). Indeed, as a leader of the Child Study Movement, Hall forcefully argued for collaborative efforts between pedagogy and the emerging discipline of psychology, creating schools that push adolescents to their physical and mental limits, and effect the “moral rejuvenation” of youth, society, and indeed the human race. A view of adolescence as a source of manifold revitalization was especially appealing to Hall’s readership, a Gilded Age middle class weary from concern over urbanization and the perceived cultural and eugenic threats posed by large-scale immigration into the United States (Kett 1977; Ross 1972). The view that adults can constructively regulate the socialization of youth is reflected in continuing scientific and public interest in the settings of youth (e.g., the workplace) and their implications for development.

Hall’s *Adolescence* was an interdisciplinary work, and drew from a wide range of sources, including writings by early sociologists such as Auguste Comte, Herbert Spencer, Gustave Le Bon, and Adolphe Quételet. The interdisciplinary study of youth remains an important theme, with many fields recognizing adolescence as a significant area of inquiry, including psychology (Petersen 1988), history (Modell and Goodman 1990), and anthropology (Schlegel and Barry 1991). Yet each discipline has its unique presuppositions and focal points. Psychologists tend to focus on adolescents’ cognitive, motivational, and emotional capacities; their maturation (often along universalistic lines, as one finds in the work of Piaget and Erikson), their interrelationships, and how they are shaped by experiences in proximal settings, including the family, peer group, school, and workplace. Anthropologists and historians focus on the range of experiences that adolescence encompasses across cultures and through historical time: the existence

of the adolescent life phase, its distinctive social and cultural traditions, and interrelationships among youth, other age groups, and social institutions.

Sociological studies of adolescence often overlap with these concerns, reflecting interests in the social settings of youth and their implications for the self, as well as variability in this stage of life across societies and through historical time. Yet sociologists have also maintained a unique view by drawing on the life course paradigm as an analytic framework. The life course focuses on age-graded roles, opportunities, and constraints; how these differ through historical time, and how they shape the biography. The analytic focus is on the structural complexity and diversity of social settings through time and place, as well as the plasticity of humans in these settings (Dannefer 1984).

The remainder of this entry will focus on three distinctive features of adolescence as viewed from a life-course perspective (see Table 1). The first feature concerns adolescence as a life phase in historical perspective: Has adolescence been a recognized part of the life course through historical time? And how have the factors that mark the transition both into and out of adolescence changed? Implicit in the concept of markers that distinguish adolescence from childhood and adulthood is the rate of movement from one phase to the next. Accordingly, the second feature concerns how quickly young people move through the adolescent role set and the social circumstances that promote an accelerated life course.

The third feature focuses on the central role of institutionalized pathways through adolescence. In this context, pathways refer to routes from childlike dependence on the family of origin to the autonomies of adulthood. At the same time, individuals actively construct their lives. Within the structured pathways from childhood to adulthood, how do adolescents actively shape their biographies? Throughout this essay, social historical accounts are presented to underscore the highly variable nature of adolescence in the last two centuries; in turn, these accounts are juxtaposed with current sociological efforts to understand the social worlds of youth. The entry concludes by considering the dual role of sociologists in the study of adolescence: To contribute to substantive debates about the place of youth in society, but

**Adolescence As a Phase of the Life Course**

Features of Adolescence	Core Idea
1. Adolescence in social historical perspective	Variability in the adolescent experience can be studied through the social history of youth.
A. Historical permanence of adolescence	Adolescence is a semi-autonomous phase of life that is not of modern origin. Adolescence is always changing in response to social forces.
B. The boundaries of adolescence	Adolescence is differentiated from childhood and adulthood by transition markers and roles.
(1) From childhood to adolescence	The pubertal transition was not always a critical marker between childhood and adolescence.
(2) From adolescence to adulthood	The transition markers have been compressed and their sequence has become more complex.
2. Pace of movement through adolescent roles	Social stressors may promote rapid movement into, through, and out of adolescent roles.
3. Pathways through adolescence	Pathways direct youth through social positions in organizations.
A. Pathways in the school	This pathway is defined by the transition to 8th grade, tracks, and transitions out of high school.
B. Pathways in the workplace	This pathway is defined by the adolescent work career: extent of work involvement, quality of work, and fit with other roles and life goals.
C. Agency in pathways	Adolescent planfulness is a critical resource with which to actively negotiate the life course.

Table 1

also to identify how the contours of these debates are themselves the products of social forces.

Two additional features of adolescence are not covered in this entry. One involves the social relationships of youth, a subject that has been examined from several vantage points. Considerable attention has been devoted to the “sociometric” properties of peer relationships, mapping out affiliations among young people in high schools (see Hallinan and Smith, 1989 for a contribution to this tradition). Relatedly, sociologists have also examined the typical personalities, behavioral patterns, and group identities of youth as they reflect responses to the social organization of the high school and this phase of life (e.g., Matza 1964). Sociologists have also focused on youth and their intergenerational relationships: How youth are integrated into adult society (for example, see Coleman 1994), how they and their parents interrelate (for a useful review, see Dornbusch 1989), and how youth serve as agents of social change (for

a classic statement, see Mannheim 1928/1952). The second feature is juvenile delinquency (see *JUVENILE DELINQUENCY*).

**ADOLESCENCE AS A LIFE PHASE**

Each phase of life reflects social norms and institutional constraints and serves as a principal source of identity for the individual by specifying appropriate behaviors and roles (Elder 1980). The study of adolescence as a life phase requires that it be situated in the life course, that its distinctive features be identified in comparison to both childhood and adulthood. Indeed, adolescence is frequently depicted as a transitional period of semiautonomy, reflecting movement from the complete dependence of children on their parents to the establishment of one’s own livelihood and family in adulthood (e.g., Kett 1974; Katz 1979; Gillis 1974). Yet the study of adolescence as a life phase also requires that it be situated in history,

that the changing norms and institutions that shape adolescence be identified, and correspondingly, that the changing nature of adolescent semiautonomy be recognized. Within this frame of life-stage analysis, social scientists have studied the historical permanence of adolescence and the factors that have circumscribed this phase of life, marking its beginning and end.

**“Adolescence” in historical time.** Initial efforts to interpret the social history of youth concluded that adolescence did not exist before the modernization of societies (modernization refers to a constellation of societal changes thought to mark a break with previous forms of social organization: rapid technological changes, the emergence of market economies, urbanization, industrialization, the decline of agricultural life, secularization, broad-based political participation, the use of currency, and the spread of science [Kleiman 1998]). Norbert Elias (1994) suggested that children and adults became increasingly distinct in their behaviors as etiquette became more widespread and refined, particularly with the collapse of feudal societies. Indeed, Elias implied that the life span recapitulates the history of manners, an instance of Haeckel’s biogenetic law.

More focused on youth was Philippe Ariés’s path-setting *Centuries of Childhood* (1962). Drawing on a diverse array of evidence—including art history, linguistics, and literature—Ariés argued that in medieval times children merged directly into adult roles starting at around seven years of age. Medieval society distinguished between adults and nonadults, but, in the latter category, distinctions were not maintained between children and adolescents. Most medieval and premodern children did not attend school, but were incorporated into adult life as quickly as possible by way of daily interactions with their elders in tightly knit communities. The few youth who did attend school remained integrated in adult society by way of a vocational curriculum designed largely to train lawyers and the clergy. According to Ariés, beginning as early as the sixteenth century, a wide range of factors—from Cartesianism to technological advancements—led to the prolongation of childhood and the emergence of adolescence as a life phase. Youth were to be educated in age-segregated settings according to curricula that were less concerned with vocational training. With this prolongation of education and segregation from the

adult world, adolescence emerged as a distinct age-graded identity.

Like Hall’s *Adolescence*, Ariés’s work was read by a receptive audience (Ben-Amos 1995). The prominent functionalist Kingsley Davis (1944) had already argued that the transmission of adult norms and values took less time in “simpler” societies. Similarly, in his highly influential *The Lonely Crowd*, David Riesman (1950) argued that children could assume adult roles in tradition-directed societies (see also Eisenstadt 1956). But Ariés was unique in his use of the historical record, and his work was a point of departure for the social history of children and adolescence that subsequently emerged in the 1970s (e.g., Demos 1970; Gillis 1974). Accordingly, some commentators maintained that adolescence was “discovered” or “invented” in the eighteenth century, as shown in part by more precise distinctions among words like “child,” “adolescent,” and “youth” (e.g., Musgrove 1964). Yet a linguistic analysis says little about the historical permanence of adolescence as a set of transitional experiences marked by semiautonomy between childhood and adulthood. Indeed, many of Ariés’s arguments have been seriously contested, including his description of education in medieval and Renaissance Europe and his lack of appreciation for the semiautonomous roles youth often played in these societies as servants or apprentices (e.g., Davis 1975).

The guiding principle of most contemporary social historical research is that adolescence reflects ever-changing values and societal structures, encompassing demographic, political, economic, and social realities. For example, Kett (1977) views American adolescence as a set of behaviors imposed on youth beginning in the late nineteenth century. This imposition was justified by “psychological laws” (e.g., the necessity of religious decision during adolescence) freighted with middle-class concerns over the dangers of cities, immigrants, bureaucracies, and the pace of social change. Not surprisingly, social historians have detected “different adolescences” in diverse settings defined by historical time and place.

Reflecting a continuing engagement with Ariés, a focal point of social historical research has been the experience of adolescence before, during, and after the emergence of industrialism. An overview



of research on England highlights the contingent nature of adolescence as a social category. Two realities were prominent in defining adolescence in premodern England (Gillis 1974). First, children of all social classes were often sent out of their household of origin to become servants for another family at about age seven. "Binding out" coincided with the adoption of adult dress and codes of behavior, although the young person was viewed as neither fully adult nor child. Rather, this practice marked a form of semiautonomy: they participated in the labor market and resided outside their parents' home, but they did not marry and they were not financially independent.

Second, marriage was often linked with the establishment of an independent household. The timing of this transition, which marked full adulthood, was in turn determined by when the father conveyed dowries for daughters and annuities or land to sons. Although the specifics of inheritance were often far more complex than the rule of primogeniture would suggest (Stone 1979), the net result was that marriage was typically postponed among the poor and lower-middle classes (that is, most of society) until young people were in their late twenties, with males marrying about two years later than females. For the proportion of the population that never married (about one in five), the commencement of adulthood hinged on occupational achievements and financial independence, which probably took place in the late twenties as well.

The coming of industrialism changed the adolescent experience dramatically. Reactions to industrialism differed greatly by class and were complicated by a wide array of factors. For many strata of society, however, economic livelihood was often enhanced by encouraging several wage laborers within the family (Gillis 1974). In turn, wage labor was a strong force in creating new and popular pathways into marriage. Concerns over inheritance were less common than in the earlier period, and kin ties were defined along more pragmatic lines that allowed youth greater freedom to marry and to establish a household. The absence of strong patriarchal control and new-found pocket money led many youth to courting and consumption patterns that shocked their elders. A new adolescence had emerged.

This broad-brush view varied in important ways from place to place, among social classes, and

by gender (for related accounts of the English experience, see Anderson 1971; Musgrove 1964; Smelser 1959; for the Continental experience, see Mitterauer and Sieder 1982; for the American experience see Demos 1970; Handlin and Handlin 1971; Hareven 1982; Kett 1977; Prude 1983). Yet it is instructive for two reasons. First, most scholars now agree that as a transitional stage of semiautonomy, adolescence existed before the emergence of modern societies, although it had distinctive "premodern" characteristics. For example, premodern adolescence was typically not a period of identity formation, as Erik Erikson's (1963) putatively universal model of psychosocial development maintains (Mitterauer 1992). Young people knew their occupational and educational futures, their parents arranged both their marriages and home-leaving, and the realities of inheritance, fecundity, and infant mortality dictated their reproductive behaviors. Furthermore, for most youth, few real political or religious options presented themselves. Although there are recorded instances of youth riots in urban areas and many adolescents and young adults were active in the Protestant Reformation, political and religious beliefs generally reflected the traditions and customs of the locale.

Second, although adolescence existed in preindustrial times, historians such as Ariés maintain too sharp a distinction between premodern and modern phases of the life course (Ben-Amos 1995). The adolescences of both the preindustrial and contemporary West are not entirely dissimilar, suggesting that there are distinctly "modern" features of preindustrial adolescence and "traditional" features of contemporary adolescence. For example, many adolescents of both periods lack a parent. In seventeenth-century England, life expectancy was approximately thirty-two years, so that many youth, born when the mother was in her early to mid-twenties, lacked at least one parent. In contemporary society, parental separation is not uncommon through the early life course. For example, among cohorts born between 1967 and 1973, about 20 percent of white males and 60 percent of black males have lived in a mother-only family between birth and age 15 (Hill et al. 1999). Parental separation and absence today is a substitute for the parental mortality of the premodern period.

Similarly, the adolescence of both historical eras was a drawn-out process. Because of early home-leaving and late marriage, preindustrial adolescence in England spanned two decades. In contemporary times, it is commonly asserted that the span between puberty and the transition to adulthood is excessive because of delays in school completion, marriage, and home-leaving. Indeed, some sociologists have suggested adding a “postadolescence” stage to the life course (e.g., Hurrelmann 1989). Although the duration of adolescence today appears extended against the backdrop of the early to mid-twentieth century, it is nevertheless brief when compared to preindustrial adolescence (about twelve versus twenty years).

In short, adolescence traces back to at least the High Middle Ages in the West, but its form and content have been remarkably responsive to social setting (Mitterauer 1992). Furthermore, one often observes both similarities and differences among the “adolescences” defined by historical time and place.

**The transition from childhood to adolescence.** Sociologists have not studied the transition to adolescence extensively, perhaps because it is typically equated with the onset of puberty, which strongly reflects individual differences in genetics, nutrition, and physical exertion (Tanner 1978). This lack of interest is unfortunate, as the effects of these factors on pubertal timing have always been conditioned by social circumstances (e.g., improvements in nutrition diffused through many societies on the basis of class and urban-rural distinctions; see Mitterauer 1992).

In any event, historical analyses suggest that puberty was not always the primary marker of the transition to adolescence. By today’s standards, physical changes associated with puberty occurred notably later in the premodern and early modern periods. For example, the average age of menarche was about fifteen for girls in early eighteenth-century America and final height was not attained among men until around age twenty-five (Kett 1977). Sources from mid-sixteenth century Europe suggest even later dates and a much more gradual progression of physical changes than is observed today (for a review of earlier sources and their critical evaluation, see Tanner 1981).

In the premodern period, young people were probably viewed as semi-autonomous when they

were sent to other households as servants or apprentices (often between ages seven and ten). Other local customs (such as religious confirmation and conversion, and membership in a wide array of village groups) also marked the end of childhood, and these frequently occurred before the pubertal transition. Thus historical evidence suggests that physical changes associated with puberty were not prominent factors that distinguished children from adolescents, and this generalization may be valid into the mid-nineteenth century, when improvements in nutrition began to take hold for large segments of society. Indeed, the pubertal transition often represented an important step into adult roles. Before the mid-nineteenth century, puberty in America was associated with a sense of rising power and energy and the ability to assume adult work responsibilities. It was largely after the Civil War that puberty came to represent a vulnerable and awkward stage closely associated with the adolescence of today (Kett 1977).

**The transition from adolescence to young adulthood.** A range of “transition markers” are typically used to indicate movement out of adolescence and into adulthood. These include leaving school, starting a full-time job, leaving the home of origin, getting married, and becoming a parent. These markers can not be used uncritically, however, because their relevance in defining stages of the life course changes through historical periods (Mitterauer 1992). Furthermore, although young people in the contemporary West rely on these markers to distinguish between adolescence and adulthood, they also draw on other criteria, including cognitive self-sufficiency, emotional self-reliance, and behavioral self-control (Arnett and Taber 1994). Nevertheless, most sociological research has focused on these transition markers and has generated valuable insights about the changing life course. Paradoxically, many commentators argue that markers of the transition from adolescence to adulthood have become both more standardized and variable.

**Standardization: The compression of transition markers.** Standardization reflects the increasing importance of age-grading and is seen in the increasing “compactness” of transition markers, particularly the ages of school completion, first job, and marriage. Theorists argue that the organization of public services, transfer payments, and

employment opportunities according to age renders the life course more orderly and calculable (Beck 1992; Kohli 1986). Also, as the state increased the number of rights that an individual could claim on a universalistic, standardized basis, it also restricted the individual's right to organize many aspects of life (e.g., with respect to education and entry into, and exit from, the labor market) (Buchmann 1989).

Evidence from historical demography suggests that the transition to adulthood has indeed become more standardized. Examining the prevalence of different female life-course patterns (e.g., spinster, dying mother, widowed mother) among cohorts of women born between 1830 and 1920, Uhlenberg (1969) observes a convergence on the "typical" pattern, involving marriage, having children, and surviving with husband until age 55. Among women born in 1830, about 21 percent experienced this "typical" pattern in contrast to about 57 percent of women born in 1920. In addition, the age range in which women typically married and had children narrowed. The primary factor promoting standardization of the life course was improvement in mortality due to the management of contagious and infectious diseases such as smallpox, typhoid and scarlet fever, diphtheria, and measles. Similarly, the time it took 80 percent of both men and women to leave the household of origin, marry, and establish their own households decreased markedly between 1880 and 1970 among those who experienced these transitions (Modell et al. 1976). A considerable body of evidence suggests that the transition to adulthood was standardized between about 1830 and 1960, as measured by a constriction of the time in which most people pass through a range of transition markers (for further discussion, see Shanahan forthcoming).

Theoreticians have emphasized the critical role of "modernity" in explaining this long-term pattern, but this formulation, with its connotation of a monotonic pace and continuous process, has not been supported by empirical study. Instead, age standardization has been affected by historically specific conditions, including improvements in health (Uhlenberg 1969) and age-grading in the school system (Hogan 1981). And as historians have noted, legal reforms, public debates about the rights and responsibilities of age groups, and cultural innovations have come into play at different times and with varying degrees of import (Kett

1977; Modell 1989; Zelizer 1994). Evidence thus points to a long-term trend of compression of the transition markers, but that trend reflects manifold factors proceeding at an uneven pace.

**Variability: The complex sequencing of transition markers.** Variability is found in the increasing complexity of role overlap and sequencing during the transition to adulthood. Theorists of modernity maintain that as individuals were freed from the traditional constraints of family and locale, they were able to exercise more agency in the construction of their biographies (e.g., Beck 1992; Giddens 1991). Consistent with these arguments, Modell, Furstenberg, and Hershberg (1976) observe that as transition markers occurred in briefer periods of time, they exhibited greater diversity in their sequencing. Between 1880 and 1970, the familial and nonfamilial transition markers increasingly overlapped, creating variability in the transition to adulthood in the form of more sequence patterns of school completion, leaving home, starting a family and career, and becoming a parent.

Hogan (1981) provides important empirical evidence for variability in the sequencing of markers among cohorts born between 1907 and 1946. The percentage of men experiencing an "intermediate nonnormative" order of transition markers (beginning work before completing school or marriage before beginning work but after school completion) increased from about 20 percent in the cohorts born between 1907 and 1912 to about 30 percent for men born in 1951. The prevalence of "extreme nonnormative" ordering (marriage before school completion) increased from less than 10 percent among cohorts born between 1907 and 1911 to over 20 percent for cohorts born between 1924 and 1947. "Modernity" has a large negative effect on the prevalence of the normative pattern, but a large positive effect on the prevalence of the extreme nonnormative pattern. That is, in historical times marked by greater educational attainment, lower infant mortality, greater longevity, and fewer youth in the adult labor market, men are more likely to make extremely nonnormative transitions to adulthood. Evidence thus suggests a trend toward individualization of the life course as found in the increased variability in the sequencing and overlap of transitions (for further discussion, see Shanahan forthcoming).

### THE ACCELERATED LIFE COURSE AND ADOLESCENCE

A conception of adolescence as bounded by markers implies a normative rate of movement through roles that indicate childhood, adolescence, and adulthood. Accordingly, influential accounts of youth view adolescence as a transitory period, marking normatively paced movement from childlike to adult roles (e.g., Linton 1942). The individual is construed as more or less adultlike depending on the acquisition of symbols, the provision of opportunities, and demands for responsibility that indicate adulthood. With respect to the second decade of life, behavioral scientists recognize three forms of an accelerated life course, reflecting the nonnormatively rapid (1) transition into adolescence, (2) movement through adolescence, and (3) transition into adulthood. These manifestations of the accelerated life course are frequently linked to stressors operating on the parents and the young person.

Precocious youth represent an early form of accelerated life course in American history. Since at least the mid-eighteenth century, some young people have been portrayed as astonishingly adultlike in their intellectual, moral, physical, and social capacities. Early American history is replete with admiration for youth who grew up in log cabins only to rise to high levels of prominence while still young, a rise often linked to exceptional talents exhibited in childhood. Yet precocity was also viewed as an inconvenience. Thus, the father of an eleven-year-old graduate from Yale at the end of the eighteenth century lamented that his son was “in no way equipped to do much of anything” (Graff 1995, p. 47). Instances such as these became rare as the social institutions of youth became age-graded, beginning in the mid-nineteenth century.

In contemporary times, an accelerated life course may reflect an early transition to adolescence. For example, Belsky and his colleagues (1991) have proposed a sociobiological model of early menarche and sexual activity. According to this model, high levels of stress during childhood—reflecting marital discord, inconsistent and harsh parenting, and inadequate financial resources—lead to aggression and depression in late childhood, which in turn foster early puberty and sexual activity. In contrast, children whose families

enjoy spousal harmony, adequate financial resources, and sensitive, supportive parenting tend to experience a later onset of puberty and sexual activity. That is, depending on environmental cues about the availability and predictability of resources (broadly defined), development follows one of two distinct reproductive strategies.

Empirical studies provide partial support for this model. Menarche occurs earlier among girls who live in mother-only households or with stepfathers. Conflict in parent-child relationships or low levels of warmth are also associated with earlier menarche (Ellis and Graber forthcoming; Graber, Brooks-Gunn, and Warren 1995; Surbey 1990). Drawing on longitudinal data, Moffitt and her colleagues (1992) report that family conflict and the absence of a father in childhood lead to earlier menarche, although this relationship is not mediated by any psychological factors examined in their study. Research findings to date, however, are open to genetic interpretation. It may be that early-maturing mothers transmit a genetic predisposition toward early puberty and the same genes produce traits in the mother that affect parenting (Rowe forthcoming) finds some support for both models. Maccoby (1991) suggests an additional class of explanations, namely that these findings reflect social psychological processes (e.g., the adolescent’s imitation of the mother’s permissiveness).

The accelerated life course may also involve the rapid assumption of autonomy not typically associated with adolescent roles. This form may appear in the adoption of the parent role by adolescents in their family of origin, what Minuchin (1974) refers to as the “parental child.” Children and adolescents may respond to the family’s emotional and practical needs through activities such as serving as a confidant to a parent, mediating family disputes, and the extensive parenting of younger siblings. Young people may assume responsibilities such as these in single, working-parent homes, which can have positive consequences for adolescents but detrimental effects for younger children (Weiss 1979). The contextual and interpersonal factors that promote “parentification,” however, are potentially numerous and complex, perhaps encompassing family structure, sibship size, marital dysfunction, and the employment status of the parents (Jurkovic 1997).

Likewise some critics of youth employment suggest that extensive involvement in the workplace during the high school years can promote “pseudomaturity,” the appearance of adult status that nevertheless lacks the full set of rights and responsibilities that accompany adulthood. (Psychologists have also expressed concern about “pseudomaturity” among contemporary youth, although they define it as a disjunction between the apparent ability to play adult roles and a lack of “commensurate psychological differentiation” [e.g., Erikson 1959].) For example, although youth may earn considerable amounts of money during high school, they spend a relatively high percentage of their income on entertainment because they lack the financial responsibilities of true adults (e.g., insurance, housing). In turn, this “premature affluence” may interfere with the development of realistic financial values (Bachman 1983; Bachman and Schulenberg 1993). However, studies show that youth spend their earnings on a wide range of things, not all of which are concerned with leisure, including savings for future education, car insurance, and even loans and contributions to parents (Shanahan et al. 1996). Moreover, studies of families during the Great Depression suggest that economic hardship can lead to the assumption of more adultlike work responsibilities among children and adolescents, which can benefit the latter group (Elder 1974).

Finally, the accelerated life course may reflect a rapid transition to adulthood. For example, because young black men have markedly shorter life expectancies than white men, they often accelerate their transition behaviors (Burton et al. 1996). There may also be an important element of intentionality in the “search for role exit” from adolescence (Hagan and Wheaton 1993). Although the desire to leave home, marry, and have children may be normative, the intent to exit the adolescent role set too early is nonnormative and often associated with deviant acts. Indeed, the search for adolescent role exits significantly predicts frequency of dating and the timing of first marriage and parenthood. That is, some adolescents may seek rapid transition to adulthood because they are substantially dissatisfied with their experience of adolescence as a life phase. In any event, sociologists have offered numerous explanations for adolescent parenthood including, for example, the lack of role models and opportunities that would

otherwise encourage postponing intercourse and pregnancy (Brewster 1994).

### PATHWAYS AND AGENCY THROUGH ADOLESCENCE

The preceding discussion highlights the variable meanings of adolescence in the life course; a related issue is the structured pathways that constitute likely sequences of social positions through which the adolescent moves. Pathways reflect institutional arrangements that both provide and restrict opportunities, channeling youth from one social position to another. Pathways are a prominent feature in the school and workplace—and between these institutions (see *ADULTHOOD*)—where social forces match individuals to social opportunities and limitations. Pathways are also evident in the family, as a sequence of roles that the child assumes and that offer progressively greater autonomy.

Before the mid-nineteenth century in America, the immediate environments of youth were casual and unstructured; mortality and frequent moves from the home placed limits on the direct and sustained application of parental discipline, and schools were decidedly unstructured settings marked by violence and informality (Kett 1977). Between 1840 and 1880, a different viewpoint emerged in both Britain and the United States, a viewpoint that emphasized character formation in planned, “engineered” environments. For example, Horace Bushnell’s influential *Christian Nurture* (1848) argued for carefully controlled settings that would promote in youth qualities necessary to succeed in a world threatened by urbanization and non-Protestant immigrants. By the end of the nineteenth century, efforts to standardize the settings of youth—particularly schools—led to the emergence of recognized tracks of educational and occupational experiences for young men entering the medical and legal professions. Educational and occupational pathways from adolescence to adulthood were becoming standardized.

Although pathways sort individuals and assign them to various positions in social systems, people are also active agents who attempt to shape their biographies. In life-course perspective, agency at the level of the person can be defined as the ability

to formulate and pursue life plans. Young people are constrained and enabled by opportunity structures of school and work, but they also construct their life course through their active efforts. This section provides a set of examples that highlight structured pathways that describe likely sequences of social positions in terms of education and work. It also discusses sociological efforts to understand the active efforts of adolescents to shape their biographies in these structured settings.

#### **Educational pathways through adolescence.**

At the turn of the nineteenth century, youth fortunate enough to attend school typically started their educations late and attended class sporadically. Academies of education, not uncommonly a single room in a private residence, were eager to accommodate the seasonal demands of agriculture. Consequently student bodies encompassed a wide range of pupils, whose ages often said little about their academic accomplishments. This situation was more pronounced in district schools, where attendance was said to fluctuate from hour to hour. Furthermore, teaching was not yet an occupation that required credentials, and teachers frequently resorted to violence to impart lessons or to maintain order, as did the students in response to frequent humiliation.

At the college level, the violence was pronounced. There were riots at Harvard, Yale, and Princeton, nightly stoning of the president's house at Brown through the 1820s, and frequent beating of blacks, servants, fellow students, and professors (Kett 1977). During the early years of the republic, little in the educational system was standardized. Beginning in the mid-nineteenth century, the educational system changed as Horace Mann, Henry Barnard, and Calvin Stowe led the common school revival, aimed at creating regulated, controlled school settings. In addition to the efforts of early reformers, many social, economic, political, and cultural forces contributed to the standardization of schools, particularly in the late nineteenth century (Tyack 1974; Tyack, James, and Benavot 1987).

Today, the educational system is highly standardized and adolescent pathways through school are readily identifiable: Adolescent pathways through the educational system can be described in terms of the transition to the seventh grade and tracking through the high school years. A considerable

body of sociological research also examines the many structured connections among secondary education, tertiary institutions such as colleges and vocational schools, and the workplace.

At the transition to seventh grade, children may remain in the same school (a "K-8" system) or change to a junior high school. The latter structure is thought to be more stressful by sociologists, as it typically brings with it a disruption in social networks and the student's first exposure to a bureaucratic setting with a high degree of specialization. Students in K-8 systems continue to have one teacher and one set of classmates for all subjects, while students attending junior high school often have a different teacher and classmates for each subject. In fact, students in K-8 systems are more influenced by peers, date more, and prefer to be with close friends more than students making the transition to junior high school. The latter report higher levels of anonymity. Further, girls who make the transition to junior high school appear especially vulnerable to low levels of self-esteem when compared with girls in a K-8 system and boys in either system (Blyth et al. 1978; Simmons and Blyth 1987; Simmons et al. 1979).

It may be that the negative effects of the transition to junior high school are amplified as the number of transitions experienced by a young person increases. A "focal theory of change" maintains that young people are better able to cope with significant life events serially rather than simultaneously (Coleman 1974). Some evidence supports this view: As the number of transitions—including school change, pubertal change, early dating, geographic mobility, and major family disruptions such as death—that students must cope with concurrently increase, their grades, extracurricular involvements, and self-esteem decrease (Simmons et al. 1987). These associations are particularly deleterious for girls. How the transition to seventh grade is organized can thus have pervasive implications for the well-being of youth.

Within secondary schools students are frequently assigned to "tracks," different curricula for students of differing talents and interests. Sociologists have identified several noteworthy features of tracks, including selectivity (the extent of homogeneity within tracks), electivity (the extent to which students choose their tracks), inclusiveness (the extent to which tracks leave open options

for future education), and scope (the extent to which students are assigned to the same track across subjects and through time). Scope—particularly the extent to which students are in the same track across grade levels—predicts math and verbal achievement, when earlier scores of achievement are controlled. (Gamoran 1992).

In turn, these differences in educational achievement largely reflect socialization and allocation processes (for a useful review, see Gamoran 1996). Socialization refers to systematically different educational experiences across tracks. Allocation refers to the decisions made by teachers to assign students to tracks, assignments that provide information to students about their abilities and that elicit differential responses from others. That is, students of differing ability are assigned to different educational opportunities, which in turn creates inequalities in outcomes, even if initial differences in ability are taken into account. Unfortunately, tracking systems are often unfair in the sense that students of similar ability are assigned to different tracks, and assignments may be based on factors other than intellectual talents and interests (see Entwisle and Alexander 1993). Furthermore, status allocations from primary through tertiary school and to the labor force are remarkably consistent (Kerckhoff 1993). The advantages or disadvantages of one's position in the educational and occupational systems cumulate as individuals increasingly diverge in their educational and labor market attainments. Thus, educational tracks can exert substantial influence on socioeconomic achievements throughout the life course.

**Pathways in the workplace.** Work responsibilities have always indicated one's status in the life course. Through the early nineteenth century in America, young people began performing chores as early as possible in childhood and often assumed considerable work responsibilities by age seven, either on the farm or as a servant in another household. Many youth were fully incorporated into the workforce with the onset of physical maturity, in the mid- to late teens (Kett 1977). During this same period, however, agricultural opportunities waned in the Northeast while expansions in commerce, manufacturing, and construction provided new employment for youth in and around cities. Many families adopted economic strategies whereby parents and children were involved in

complex combinations of farming, work in factories, and other sources of wage labor (Prude 1983) or whereby entire families were recruited into factory work (Hareven 1982; for the case of England, see Anderson 1971; Smelser 1959).

The second Industrial Revolution, commencing after the Civil War and extending to World War I, led in part to less reliance on children as factory workers (Osterman 1979). Technological innovations in the workplace—the use of internal combustion engines, electric power, and continuous-processing techniques—created an economic context conducive to the consolidation of primary schools in the life course as many youth jobs were eliminated by mechanization (Troen 1985) and manufacturers required more highly skilled employees (Minge-Kalman 1978). These economic factors operated in concert with progressive political movements, as well as cultural, demographic, social, and legal changes (Hogan 1981; Zelizer 1985), all of which fueled debates about the appropriate role of youth in the workplace. These debates continue to the present, particularly focusing on the work involvements of high school students (for useful overviews, see Institute of Medicine/National Research Council 1998; Mortimer and Finch 1996). Today, almost all adolescents work in paid jobs and time commitments to the workplace can be substantial (Bachman and Schulenberg 1993; Manning 1990). Whereas youth work at the beginning of the twentieth century often centered around agriculture and involved family, kin, and neighbors, today's adolescent is more frequently employed in “entry-level” jobs among unrelated adults in the retail and restaurant sectors. These changes have prompted arguments that contemporary adolescent work interferes significantly with the basic developmental tasks of youth.

Yet the transition to paid work represents a large step toward autonomy and can promote a sense of contribution, of egalitarianism, and of being “grown up” among youth. Although very little research has examined the adolescent work career, several generalizations are currently plausible. First, adolescents have work careers in that they typically progress from informal work (e.g., babysitting and yard work for neighbors) to a surprisingly diverse set of occupations as seniors in high school, a trend that is accompanied by an

increase in earnings. That is, work tends to become more complex and to produce more earnings through the high school years (Mortimer et al. 1994).

Second, adolescent work can have positive or negative consequences, depending on the extent and the quality of the experience, as well as its meaning. For example, work of high intensity (that is, exceeding twenty hours of work on average across all weeks employed) curtails postsecondary education among boys and increases alcohol use and smoking among high school girls (Mortimer and Johnson 1998). But adolescents engaged in low-intensity work during high school have favorable outcomes with respect to schooling (for boys) and part-time work (for both boys and girls).

Third, the quality of work matters. For example, several studies show that jobs that draw on or confer skills deemed useful in the workplace are associated with feelings of efficacy during the high school years as well as success in the job market three years after high school (Finch et al. 1991; Stern and Nakata 1989). Similarly, adolescent work experiences can have positive implications for development depending on how they fit into the adolescent's life. Thus, when earnings are saved for college, working actually has a positive effect on grades (Marsh 1991), and nonleisure spending (e.g., spending devoted to education or savings) may enhance relationships with parents (Shanahan et al. 1996). In short, the adolescent work career can pose positive or negative implications for subsequent attainment and adjustment.

**Life course agency in adolescence.** Whereas the concept of pathways reflects an interest in how organizations and institutions allocate youth to social positions and their attendant opportunities and limitations, young people are also active agents attempting to realize goals and ambitions. It is unlikely that most youth were agents in this sense before or during the founding of the republic. In a detailed study of autobiographical life histories between 1740 and 1920, Graff (1995) observes that lives marked by conscious choice and self-direction, a search for opportunities including social mobility, the instrumental use of further education, and risk-taking in the commercial marketplace, were atypical before the nineteenth century. In some accounts this emerging orientation

was expressed in explicit emulation of the widely circulated autobiography of Benjamin Franklin, who emphasized planning, thrift, decision making, and independence. Before this period, the major features of the life course—including education, occupation, and family life—were largely determined by family circumstances. (The notable exception to a lack of decision making about one's life was religious "rebirth." Departing from Catholic and Lutheran doctrine, religious sects such as the Anabaptists emphasized the adolescent's conscious decision to be baptized and then lead an appropriately Christian life [Mitterauer 1992].)

This active orientation toward the life course gained currency as "conduct-of-life" books in the first decades of the nineteenth century departed from Puritan hostility to assertiveness in order to emphasize the building of a decisive character marked by "a strenuous will" (e.g., as found in John Foster's popular essay "Decision of Character" of 1805). The message was especially appealing to the large number of young men leaving rural areas for the city in search of jobs. Since that time, themes such as self-reliance, decisiveness, willpower, and ambition have recurred in popular advice books and social commentaries directed to adolescents and their parents (Kett 1977). Indeed, a view of "youth as shapers of youth" has become prominent among social historians of the twentieth century (for a superb example, see Modell 1989).

In contemporary times, a conception of life as shaped through decision-making, planning, and persistent effort is common. A particularly useful concept to study this phenomenon is *planful competence*, the thoughtful, assertive, and self-controlled processes that underlie selection into social institutions and interpersonal relationships (Clausen 1991a). Although these traits can be found in approaches to personality (e.g., conscientiousness), planful competence is uniquely concerned with the ability to select social settings that best match a person's goals, values, and strengths. That is, planful competence describes the self's ability to negotiate the life course as it represents a socially structured set of age-graded opportunities and limitations. Clausen (1991b, 1993) maintains that a planful orientation in mid-adolescence (about ages 14 and 15) is especially relevant to the life course because it promotes realistic decision-making about the roles and relationships of adulthood. That is, one's self-reflexivity, confidence, and self-regulation at



mid-adolescence lead to better choices during the transition to adulthood, choices that in turn have implications for later life. Clausen reasons that children are not capable of planful competence, but most adults possess at least some; therefore interindividual differences at mid-adolescence are most likely to differentiate people during the transition to adulthood and through later life. Adolescents who are planfully competent “better prepare themselves for adult roles and will select, and be selected for, opportunities that give them a head start” (1993, p. 21).

Drawing on extensive longitudinal archives from the Berkeley and Oakland samples at the Institute of Child Welfare, Clausen (1991b, 1993) demonstrated that planful competence in high school (age 15 to 18 years) had pervasive effects on functioning in later life. Planfulness significantly predicted marital stability, educational attainment for both males and females, occupational attainment and career stability for males, and life satisfaction in later adulthood. Satisfaction with marriage was often associated with adolescent planfulness among men and women, especially among men who were capable of interpersonal warmth. Men who were more planful earlier in life reported greater satisfaction with their careers, more job security, and better relationships with their coworkers.

Some research suggests that the effects of planfulness are conditioned by historical experience, an insight that joins the concepts of pathways and agency. Drawing on the Terman Sample of Gifted Children, Shanahan and his colleagues (1997) examined the lifetime educational achievement of two cohorts of men who grew up during the Great Depression. The older cohort, those born between 1900 and 1910, often were in college or had just begun their careers when the Great Depression hit. The younger cohort, those born between 1910 and 1920, typically attended college after the Depression and began their careers in the post-World War II economic boom. For the older cohort, it was hypothesized that adolescent planful competence would not predict adult educational attainment. Rather, very high levels of unemployment during the Depression would support a prolonged education for this cohort—through continuity in or return to school—regardless of their planfulness. In contrast, planful competence in adolescence was expected to predict adult educational attainment in the younger

cohort, which was presented with practical choices involving employment opportunities and further education.

As expected, planfulness at age fourteen positively predicted educational attainment, but only for the men born between 1910 and 1920, who often finished school during the postwar economic boom. Planful competence did not predict educational attainment for men from the older cohort, who typically remained in school or returned to school after their nascent careers floundered. Thus, for the older cohort, the lack of economic opportunity precluded entry into the workplace and under these circumstances, personal agency did not predict level of schooling. In short, the Terman men’s lives reflect “bounded agency,” the active efforts of individuals within structured settings of opportunity.

#### ADOLESCENCE IN RETROSPECT AND PROSPECT

A historically sensitive inquiry into the sociology of adolescence reveals several prominent features of this life phase. First, although adolescence as a state of semiautonomy between childhood and adulthood has long been part of the Western life course, it has nevertheless been highly responsive to social, political, economic, and cultural forces. Indeed, adolescence acts like a “canary in the coal mine.” As successive generations of youth encounter adult society for the first time, their reactions tell us much about the desirability of social arrangements. These reactions have ranged from enthusiastic acceptance to large-scale revolt and have often led to the emergence of new social orders.

Second, the many “adolescences” revealed through historical time and place are both different and similar in significant ways. Many supposedly universalistic accounts of social and psychological development would not describe adolescent experiences and interpretive frames in the past (Mitterauer 1992). On the other hand, few, if any, aspects of the adolescent experience are without precedent. Statistics on adolescent sexual behavior are met with great alarm today, and yet alarm was already sounded in pre-Colonial times: Leaving England for America in the 1630s, the Puritans hoped to establish an “age-relations utopia” between young people and their elders, which had

not been possible in the Old World because of the “licentiousness of youth” (Moran 1991; for the related problem of adolescent pregnancy, see Smith and Hindus 1975; Vinovskis 1988). Likewise, contemporary debates about whether adolescents should be allowed to engage in paid work echo exchanges over child labor laws a century ago (Zelizer 1985). Analogous observations could be made about drug abuse, violence in schools, gangs, and troubled urban youth with diminished prospects.

Yet each generation of adults insists that its adolescence is uniquely vulnerable to social change and the problems that come with it. Indeed, this orientation has spawned a large and often alarmist and contradictory literature about the American adolescent (see Graff 1995, especially Chap. 6). Although many strands of scientific evidence show that adolescence as a state of semiautonomy has lengthened over the past several decades, a torrent of books nonetheless warn of “the end of adolescence.” And although many studies show that adolescence is not a period of sudden and pervasive distress, scientific journals devoted to adolescence are filled with contributions examining psychological disorder, substance use, antisocial behavior, sexually transmitted disease, delinquency, troubled relationships with parents, and poor academic performance. Denials of “storm and stress” frequently accompany implicit statements of “doom and gloom.”

Unfortunately, these negative images run the risk of defining adolescence (Graff 1995). Studies and social commentaries that highlight the troubled nature of youth may alert the public to social problems in need of redress, but they may also create dominant cultural images that ultimately define what it means to be an adolescent in negative terms, breeding intergenerational mistrust and, among youth, alienation (Adelson 1986).

In this context of research and representation, sociology has much to contribute to an understanding of adolescence. What is needed are balanced accounts of their lived experiences and how these vary by social class, race, gender, and other indicators of inequality. These descriptions need to be situated in both place and time. Place refers to the many contexts of youth, encompassing families and neighborhoods, urban and rural distinctions, ideologies, modes of production, and national and international trends. Time refers to the

history of youth, with its many continuities and discontinuities, but it also refers to the life course, how adolescence fits into the patterned sequence of life’s phases. Through such efforts, sociologists can contribute to public discourse about youth and comprehend this discourse as being itself a product of social forces.

(SEE ALSO: *Adulthood; Juvenile Delinquency and Juvenile Crime; Juvenile Delinquency, Theories of; The Life Course*)

REFERENCES

- Adelson, Joseph 1986 *Inventing Adolescence: The Political Psychology of Everyday Schooling*. New Brunswick, N. J.: Transaction Books.
- Anderson, Michael 1971 *Family Structure in Nineteenth Century Lancashire*. Cambridge, Eng.: Cambridge University Press.
- Ariés, Philippe 1962 *Centuries of Childhood: A Social History of Family Life*. Translated by Robert Baldick. New York: Vintage Books.
- Arnett, Jeffrey J., and Susan Taber 1994 “Adolescence Terminable and Interminable: When Does Adolescence End?” *Youth and Society* 29:3–23.
- Bachman, Jerald G. 1983 “Premature Affluence: Do High School Students Earn Too Much?” *Economic Outlook* 10:64–67.
- , and John Schulenberg 1993 “How Part-time Work Intensity Relates to Drug Use, Problem Behavior, Time Use, and Satisfaction Among High School Seniors: Are These Consequences or Merely Correlates?” *Developmental Psychology* 29:220–235.
- Beck, Ulrich 1992 *Risk Society: Towards a New Modernity*. Translated by Mark Ritter. London: Sage.
- Belsky, Jay, Laurence Steinberg, and Patricia Draper 1991 “Childhood Experiences, Interpersonal Development, and Reproductive Strategy: An Evolutionary Theory of Socialization.” *Child Development* 62:647–670.
- Ben-Amos, Ilana Krausman 1995 “Adolescence as a Cultural Invention: Philippe Ariés and the Sociology of Youth.” *History of the Human Sciences* 8:69–89.
- Blyth, Dale A., Roberta G. Simmons, and Diane Mitbush 1978 “The Transition into Early Adolescence: A Longitudinal Comparison of Youth in Two Educational Contexts.” *Sociology of Education* 51:149–162.
- Brewster, Karen 1994 “Race Differences in Sexual Activity Among Adolescent Women: The Role of Neighborhood Characteristics.” *American Sociological Review* 59:408–424.

- Buchmann, Marlis 1989 *The Script of Life in Modern Society: Entry into Adulthood in a Changing World*. Chicago: University of Chicago Press.
- Burton, Linda M., Dawn A. Obeidallah, and K. Allison 1996 "Ethnographic Insights on Social Context and Adolescent Development Among Inner-city African-American Teens." In Richard Jessor, Anne Colby, and Richard Shweder, eds., *Ethnography and Human Development*. Chicago: University of Chicago Press.
- Bushnell, Horace 1848 *Views of Christian Nurture, and of Subjects Adjacent Thereto*. Hartford, Conn.: E. Hunt.
- Clausen, John 1993 *American Lives: Looking Back at the Children of the Great Depression*. Berkeley: University of California Press.
- 1991a "Adolescent Competence and the Shaping of the Life Course." *American Journal of Sociology* 96:805–842.
- 1991b "Adolescent Competence and the Life Course, Or Why One Social Psychologist Needed a Concept of Personality." *Social Psychology Quarterly* 54:4–14.
- Coleman, James S. 1994 "Social Capital, Human Capital, and Investment in Youth." In Anne C. Petersen and Jeylan T. Mortimer, eds., *Youth Unemployment and Society*. New York: Cambridge University Press.
- Coleman, J. C. 1974 *Relationships in Adolescence*. Boston: Routledge and Kegan Paul.
- Côté, James E. 1994 *Adolescent Storm and Stress: An Evaluation of the Mead-Freeman Controversy*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Dannefer, Dale 1984 "Adult Development and Social Theory: A Paradigmatic Reappraisal." *American Sociological Review* 49:100–116.
- Davis, Kingsley 1944 "Adolescence and the Social Structure." *The Annals of the American Academy of Political and Social Science* 236:8–16.
- Davis, Natalie Z. 1975 *Society and Culture in Early Modern France*. Stanford, Calif.: Stanford University Press.
- Demos, John 1970 *A Little Commonwealth: Family Life in Plymouth Colony*. New York: Oxford University Press.
- Dornbusch, Sanford 1989 "The Sociology of Adolescence." *Annual Review of Sociology* 15:233–259.
- Douvan, Elizabeth, and Joseph Adelson 1966 *The Adolescent Experience*. New York: Wiley.
- Eisenstadt, Shmuel N. 1956 *From Generation to Generation: Age Groups and Social Structure*. Glencoe, Ill.: Free Press.
- Elder, Glen H., Jr. 1980 "Adolescence in Historical Perspective." In Joseph Adelson, ed., *Handbook of Adolescent Psychology*. New York: Wiley.
- 1974 *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press.
- Elias, Norbert 1994 *The Civilizing Process*. Translated by Edmund Jephcott. Oxford: Blackwell.
- Ellis, Bruce J., and Judy Graber forthcoming. "Psychosocial Antecedents of Variation in Girls' Pubertal Timing: Maternal Depression, Stepfather Presence, and Marital and Family Stress." *Child Development*.
- Entwisle, Doris R., and Karl L. Alexander 1993 "Entry into Schools: The Beginning School Transition and Educational Stratification in the United States." In *Annual Review of Sociology*, vol. 19. Palo Alto, Calif.: Annual Reviews, Inc.
- Erikson, Eric H. 1963 *Childhood and Society*. New York: Norton.
- 1959 *Identity and the Life-Cycle*. New York: International Universities Press.
- Finch, Michael D., Michael J. Shanahan, Jeylan T. Mortimer, and Seongryeol Ryu 1991 "Work Experience and Control Orientation in Adolescence." *American Sociological Review* 56:597–611.
- Freeman, Derek 1983 *Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth*. Cambridge, Mass.: Harvard University Press.
- Gamoran, Adam 1996 "Educational Stratification and Individual Careers." In Alan C. Kerckhoff, ed., *Generating Social Stratification: Toward a New Research Agenda*. Boulder, Colo.: Westview Press.
- 1992 "The Variable Effects of High School Tracking." *American Sociological Review* 57:812–828.
- Giddens, Anthony 1991 *Modernity and Self-identity: Self and Society in the Late Modern Age*. Stanford, Calif.: Stanford University Press.
- Gillis, John R. 1974 *Youth and History: Tradition and Change in European Age Relations 1770–Present*. New York: Academic Press.
- Gould, Stephen Jay 1977 *Ontogeny and Phylogeny*. Cambridge, Mass.: Belknap Press of Harvard University.
- Graber, Judy, Jeanne Brooks-Gunn, and M. P. Warren 1995 "The Antecedents of Menarcheal Age: Heredity, Family Environment, and Stressful Life Events." *Child Development* 66:346–359.
- Graff, Harvey J. 1995 *Conflicting Paths: Growing Up in America*. Cambridge, Mass.: Harvard University Press.
- Hagan, John, and Blair Wheaton 1993 "The Search for Adolescent Role Exits and the Transition to Adulthood." *Social Forces* 71:955–980.
- Hall, G. Stanley 1904 *Adolescence: Its Psychology and its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion, and Education*. New York: D. Appleton and Co.

- Hallinan, Maureen, and Steven S. Smith 1989 "Classroom Characteristics and Student Friendship Cliques." *Social Forces* 67:898-919.
- Handlin, Oscar, and Mary F. Handlin 1971 *Facing Life: Youth and the Family in American History*. Boston: Little, Brown.
- Hareven, Tamara 1982 *Family Time and Industrial Time: The Relationship Between the Family and Work in a New England Industrial Community*. New York: Cambridge University Press.
- Hill, Martha S., W. Yueng, and Greg J. Duncan 1999 "How Has the Changing Structure of Opportunities Affected Transitions to Adulthood?" In Alan Booth, Nan Crouter, and Michael Shanahan, eds., *Transitions to Adulthood in a Changing Economy*. Greenwood, Conn.: Greenwood Publishing.
- Hogan, Dennis P. 1981 *Transitions and Social Change: The Early Lives of American Men*. New York: Academic Press.
- Hollingshead, August 1949 *Elmstown's Youth, The Impact of Social Classes on Adolescents*. New York: Wiley.
- Hurrelmann, Klaus 1989 "The Social World of Adolescents: A Sociological Perspective." In Klaus Hurrelmann and Uwe Engel, eds., *The Social World of Adolescents: International Perspectives*. Berlin: Walter de Gruyter.
- Institute of Medicine/National Research Council 1998 *Protecting Youth at Work: Health, Safety, and Development of Working Children and Adolescents in the United States*. Washington, D.C.: National Academy Press.
- Jurkovic, Gregory J. 1997 *Lost Childhoods: The Plight of the Parentified Child*. New York: Brunner Mazel.
- Katz, Michael B. 1975 *The People of Hamilton, Canada West: Family and Class in a Mid-Nineteenth Century City*. Cambridge, Mass.: Harvard University Press.
- Kerckhoff, Alan C. 1993 *Diverging Pathways: Social Structure and Career Deflections*. Cambridge, Eng.: Cambridge University Press.
- Kett, Joseph F. 1977 *Rites of Passage: Adolescence in America, 1790 to the Present*. New York: Basic Books.
- Kleiman, Jordan 1998 "Modernization." In Richard W. Fox and James T. Kloppenberg, eds., *Companion to American Thought*. Malden, Mass.: Blackwell.
- Kohli, Martin 1986 "The World We Forgot: A Historical Review of the Life Course." In Victor Marshall, ed., *Later Life*. Beverly Hills, Calif.: Sage.
- Linton, Ralph 1942 "Age and Sex Categories." *American Sociological Review* 7:589-603.
- Lynd, Robert S., and Helen M. Lynd 1929 *Middletown: A Study of American Culture*. New York: Harcourt, Brace and Co.
- Maccoby, Eleanor E. 1991 "Differential Reproductive Strategies in Males and Females." *Child Development* 62:676-681.
- Mannheim, Karl (1928) 1952 "The Problem of Generations." In Paul Kecskemeti, ed., *Essays on the Sociology of Knowledge*. New York: Oxford University Press.
- Manning, Wendy 1990 "Parenting Employed Teenagers." *Youth & Society* 22:184-200.
- Marro, Antonio 1898 *La Puberta*. Torino: Fratelli Bocca Editori.
- Marsh, Herbert W. 1991 "Employment During High School: Character Building or a Subversion of Academic Goals?" *Sociology of Education* 64:172-189.
- Matza, David 1964 "Position and Behavior Patterns of Youth." In Robert E. L. Faris, ed., *Handbook of Modern Sociology*. Chicago: Rand McNally & Company.
- Mead, Margaret 1928 *Coming of Age in Samoa: A Psychological Study of Primitive Youth for Western Civilization*. New York: Blue Ribbon Books.
- Minge-Kalman, Wanda 1978 "The Industrial Revolution and the European Family: The Institutionalization of 'Childhood' as a Market for Family Labor." *Comparative Studies in Society and History* 20:454-468.
- Minuchin, Salvador 1974 *Families and Family Therapy*. Cambridge, Mass.: Harvard University Press.
- Mitterauer, Michael 1992 *A History of Youth*. Translated by Graeme Dunphy. Oxford: Basil Blackwell.
- , and Reinhard Sieder 1982 *The European Family: Patriarchy to Partnership from the Middle Ages to the Present*. Oxford, Eng: Basil Blackwell.
- Modell, John 1989 *Into One's Own*. Berkeley: University of California.
- , Frank F. Furstenberg Jr., and Theodore T. Hershberg 1976 "Social Change and Transitions to Adulthood in Historical Perspective." *Journal of Family History* 1:7-32.
- , and Madeline Goodman 1990 "Historical Perspectives." In S. Shirley Feldman and Glen R. Elliott, eds., *At the Threshold: The Developing Adolescent*. Cambridge: Harvard University Press.
- Moffitt, Terrie E., Avshalom Caspi, Jay Belsky, and Phil A. Silva 1992 "Childhood Experience and Onset of Menarche: A Test of a Sociobiological Model." *Child Development* 63:47-58.
- Moran, Gerald F. 1991 "Adolescence in Colonial America." In Richard M. Lerner, Anne C. Petersen, and Jeanne Brooks-Gunn, eds., *Encyclopedia of Adolescence*. New York: Garland Publishers.
- Mortimer, Jeylan T., and Michael D. Finch 1996 "Work, Family, and Adolescent Development." In Jeylan T.

- Mortimer and Michael D. Finch, eds., *Adolescents, Work, and Family: An Intergenerational Developmental Analysis*. Thousand Oaks, Calif.: Sage Publications.
- , Katherine Dennehey, Chaimun Lee, and Timothy Beebe 1994 "Work Experience in Adolescence." *Journal of Vocational Behavior* 19:39–70.
- Mortimer, Jeylan T., and Monika K. Johnson 1998 "New Perspectives on Adolescent Work and the Transition to Adulthood." In Richard Jessor, ed., *New Perspectives on Adolescent Risk Behaviors*. New York: Cambridge University Press.
- Musgrove, Frank 1964 *Youth and the Social Order*. Bloomington: Indiana University Press.
- Osterman, Paul 1979 "Education and Labor Markets at the Turn of the Century." *Politics and Society* 9:103–122.
- Petersen, Anne C. 1988 "Adolescent Development." *Annual Review of Psychology* 39:583–607.
- Prude, Jonathan 1983 *The Coming of the Industrial Order: Town and Factory Life in Rural Massachusetts, 1810–1860*. Cambridge, Eng.: Cambridge University Press.
- Riesman, David, in collaboration with Reuel Denney, and Nathan Glazer 1950 *The Lonely Crowd: A Study of the Changing American Character*. New Haven, Conn.: Yale University Press.
- Ross, Dorothy 1972 *G. Stanley Hall: The Psychologist as Prophet*. Chicago, Ill.: University of Chicago Press.
- Rowe, David C. forthcoming "Environmental and Genetic Influences on Pubertal Development: Evolutionary Life History Traits?" In J. Rogers and David Rowe, eds., *Genetic Influences on Fertility and Sexuality*. Boston: Kluwer.
- Schlegel, Alice, and Herbert Barry III 1991 *Adolescence: An Anthropological Inquiry*. New York: Free Press.
- Shanahan, Michael J. forthcoming "Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective." *Annual Review of Sociology*.
- , Glen H. Elder Jr., Margaret Burchinal, and Rand D. Conger 1996 "Adolescent Paid Labor and Relationships with Parents: Early Work-Family Linkages." *Child Development* 67:2183–2200.
- , Glen H. Elder Jr., and Richard A. Miech 1997 "History and Agency in Men's Lives: Pathways to Achievement in Cohort Perspective." *Sociology of Education* 70:54–67.
- Simmons, Roberta G., and Dale A. Blyth 1987 *Moving Into Adolescence: The Impact of Pubertal Change and School Context*. New York: Aldine de Gruyter.
- , Edward F. Van Cleave, and Diane Mitsch Bush 1979 "Entry into Early Adolescence: The Impact of School Structure, Puberty, and Early Dating on Self-Esteem." *American Sociological Review* 44:948–967.
- Simmons, Roberta G., Richard Burgeson, Steven Carlton-Ford, and Dale A. Blyth 1987 "The Impact of Cumulative Change in Early Adolescence." *Child Development* 58:1220–1234.
- Smelser, Neil J. 1959 *Social Change in the Industrial Revolution: An Application of Theory to the British Cotton Industry*. Chicago: University of Chicago Press.
- Smith, Daniel S., and Michael S. Hindus 1975 "Premarital Pregnancy in America, 1640–1971: An Overview and Interpretation." *Journal of Interdisciplinary History* 4:537–570.
- Stern, D., and Y. F. Nakata 1989 "Characteristics of High School Students, Paid Jobs, and Employment Experience After Graduation." In D. Stern and Dorothy Eichorn, eds., *Adolescence and Work: Influences of Social Structures, Labor Markets, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Stone, Lawrence 1977 *The Family, Sex and Marriage in England, 1500–1800*. New York: Harper & Row, Co.
- Surbey, M. 1990 "Family Composition, Stress, and Human Menarche." In F. Bercovitch and T. Zeigler, eds., *Socioendocrinology of Primate Reproduction*. New York: Liss.
- Tanner, James M. 1981 *A History of the Study of Growth*. Cambridge, Eng.: Cambridge University Press.
- 1978 *Foetus into Man: Physical Growth from Conception to Maturity*. Cambridge, Mass.: Harvard University Press.
- Thorndike, Edwin 1917 "Magnitude and Rate of Alleged Changes at Adolescence." *Educational Review* 54:140–147.
- Troen, Selwyn 1985 "Technological Development and Adolescence: The Early 20th Century." *Journal of Early Adolescence* 5:429–439.
- Tyack, David 1974 *The One Best System: A History of American Urban Education*. Cambridge, Mass.: Harvard University Press.
- Tyack, David, Thomas James, and Aaron Benavot 1987 *Law and the Shaping of Public Education, 1785–1954*. Madison: University of Wisconsin Press.
- Uhlenberg, Peter 1969 "A Study of Cohort Life Cycles: Cohorts of Native Born Massachusetts Women, 1830–1920." *Population Studies* 23:407–420.
- Vinovskis, Maris A. 1988 *An "Epidemic" of Adolescent Pregnancy: Some Historical and Policy Perspectives*. New York: Oxford.
- Weiss, R. S. 1979 "Growing Up a Little Faster: The Experience of Growing Up in a Single-parent Household." *Journal of Social Issues* 35:97–111.

Zelizer, Viviana 1985 *Pricing the Priceless Child*. New York: Basic Books.

MICHAEL J. SHANAHAN

## ADULT EDUCATION

Most basically defined, adult education is the intentional, systematic process of teaching and learning by which persons who occupy adult roles acquire new values, attitudes, knowledge, skills, and disciplines. As a concept, “adult education” demarcates a subfield of education that is distinct from the latter’s historical and still general identification with the formal schooling of youth in primary and secondary schools, colleges, and universities. Once lacking the central social significance long recognized for this formal schooling, adult education expanded rapidly after 1950. Changing social, economic, and demographic forces occasioned new educational forms and organizations and new levels of adult participation in existing forms. Adult education is now so widespread and important a feature of societies worldwide that it increasingly occupies the attention of social scientists, policy makers, businesses, and the public.

### TYPES OF ADULT EDUCATION

Adult education now permeates modern societies. It does not do so, however, with the kind of public funding, legislative sanction, organizational cohesion, and standardization of practice that have made universal schooling a highly visible and central institution. Precise substantive definition and classification of adult education is frustrated by the great and changing variety that characterizes the field (Courtney 1989). The complex circumstances of adult life and development lead to the informal, nonformal, and formal pursuit of education for many different purposes. In response to an intricate array of social, economic, and political conditions, formal and nonformal organizations—from multi-state international agencies to corporations to local recreational clubs—support and develop adult education programs. In consequence, an eclectic set of professions, occupations, disciplines, and practices forms the division of adult education labor.

Adults seek a wide variety of educational goals. These include basic literacy and work readiness skills; knowledge and technical competencies required for entering and improving performance of occupational, avocational, and recreational roles; credentials for status attainment; information for the improvement of family life, health, and psychological well-being; knowledge, values, and disciplines for spiritual growth and intellectual enrichment; and tools for addressing community problems and advancing political and social-action agendas. An equally diverse set of organizations and groups provides such education. Publishers and producers of print and electronic educational media serve a growing market for informal adult education with products that range from golf tutorials to taped lectures on the history of philosophy. A large and rapidly expanding nonformal sector (i.e., educational organizations that are not a part of the formal school and college system), now mobilizes very considerable resources to educate adults. Businesses, government agencies, and non-profit organizations train employees to enhance productivity, organizational effectiveness, and client satisfaction, to spur innovation in products and services, and as an employment benefit to attract workers. Proprietary schools and training companies seek profits by providing similar training to both businesses and individuals. National, regional, and local governments fund adult education programs to reduce welfare dependency and promote economic development. Political parties and special-interest groups deliver adult education designed to foster either dominant or insurgent civic values, knowledge, and action. Professional associations sponsor and certify continuing education to maintain and enhance member competence, ensure the value of their credentials, and maintain market advantages for their members.

Other major providers of nonformal education programs that expressly target adults include unions; churches; libraries, and museums; the armed forces; prison systems; charitable, fraternal, service, and cultural associations; and the health care industry. The formal educational system itself no longer serves only the young. Community school adjuncts to primary and secondary schools teach basic literacy, prepare adults for high school equivalency exams, and offer classes in subjects ranging from the latest computer software to traditional arts and crafts. Colleges and

universities now educate almost as many adults as youth; in the United States almost half of all college students are adults above twenty-five years of age. With increasing frequency, these students study in divisions of colleges and universities specifically dedicated to adults.

The functional and organizational diversity of adult education is mirrored in its professions and occupations. Those working in the field include the administrators, researchers, and professors in university graduate programs that train adult educators and maintain adult education as an academic discipline. Teachers and student-service personnel in university, governmental, and proprietary organizations deliver graduate, undergraduate, and continuing education to adults. Managers, trainers, and associated marketing and support personnel staff the employee, technical, and professional training industry. Adult literacy and basic education practitioners form a specialization of their own. Professional activists, organizers, and volunteers consciously include adult education in the portfolio of skills that they apply to pursuits ranging across the full spectrum of ideologies and interests. Policy analysts, planners, researchers, and administrators staff the adult education divisions of international organizations, national and regional governments, and independent foundations and development agencies.

While those who occupy these professional statuses and roles are clearly doing adult education, not all identify themselves as adult educators or, even when sharing this identity, see themselves as engaged in similar practice. The field is conceptually, theoretically, and pedagogically heterogeneous both within and among its many sectors. Role identity and performance differences based in organizational setting and population served are compounded by differences in fundamental aims and methods. One of the sharpest divides is between many in the “training and development” industry and those in academia and elsewhere who identify with “adult education” as a discipline, as a profession and, sometimes, as a social movement.

Training and development specialists tend to define their task as cultivating human resources and capital that can be used productively for the purposes of businesses, armed forces, government agencies, and other formal organizations. For training line employees, and all employees in technical

areas, they tend to emphasize teaching and learning methodologies that maximize the efficiency and effectiveness of individual acquisition of skill sets that can be easily and usefully applied to well-established performance objectives. For executive and managerial development, they tend to emphasize leadership, team, problem-solving, strategy, and change management competencies in the context of developing general organizational learning, effectiveness, and continuous quality improvement (Craig 1996).

Those who identify themselves as adult educators, on the other hand, tend to emphasize theories and methods designed to “facilitate” individual transformation and development. Variants of the facilitation model advocated by the adult education profession include the following: one perspective focuses on the special characteristics of learning in adulthood, and on “andragogy” as a new type of specifically adult teaching and learning distinct from pedagogy, as critical to successful adult education (Knowles 1980); another emphasizes adult life circumstances and experiences as key variables (Knox 1986); a third sees facilitating new critical and alternative thinking as the key to successful adult transformation (Brookfield 1987); and yet another sees adult education as active, consciousness-transforming engagement with social conditions to produce individual liberation and progressive social change (Coben, Kincheloe, and Cohen 1998). Common to all of the facilitation approaches is the ideal of adult education as a democratic, participatory process wherein adult educators facilitate active learning and critical reflection for which adult learners themselves assume a large measure of responsibility and direction.

Theoretical and ideological differences among adult education practitioners draw sharper lines than does their actual practice. Active learning techniques through which concepts, information, and skills are acquired in the course of real or simulated practical problem solving, strongly advocated by professional adult educators, have been embraced by the corporate training and development industry. Educational technology tools such as interactive, computer-based learning modules and Web-based tutorials, tools most robustly developed by the training industry, enable precisely the kind of independent, self-directed learning celebrated by the adult education profession. In

most of its settings and branches, adult educators of all types deploy the entire array of pedagogies from rote memorization to classic lecture-recitation to the creation of self-sustaining “learning communities.” Few central methodological tendencies demarcate distinct factions within the field, or the field itself from other types of education. Visible variants, such as the training and facilitation models, serve only partially to distinguish different segments of the field; and even these differences stem more from the particular histories, conditions, aims, and clients of those segments than from distinct disciplines, theories, or methods. Other methodological tendencies, such as widespread reliance on adult experience and self-direction as foundational for instructional design and delivery, reflect differences between adult and childhood learning more than a distinctly adult pedagogy (Merriam and Cafferella 1991).

Adult learning has several well-established characteristics that distinguish it from learning earlier in the life cycle: greater importance of clear practical relevance for learning, even of higher-order reasoning skills; the relatively rich stock of experience and knowledge to which adults relate new learning; “learner” or “student” as a role secondary to and embedded in adult familial, occupational and social roles; and the application of adult levels of responsibility and self-direction to the learning process. Indeed, at the level of practice, the learning characteristics of those being educated (i.e., adults) serve to distinguish adult education as a distinct field much more clearly than do distinctively adult educators, organizations, or methods.

### **DEVELOPMENT OF ADULT EDUCATION**

The remarkable variety of contemporary adult education directly reflects complexity in the social environment and the necessities and rewards associated with mastering that complexity. Like schooling, adult education emerged and developed in response to the social, economic, political, cultural, and demographic forces that produced increasing structural and functional differentiation as one of the few clear trends in human social evolution. As social roles and practices proliferated, conveying the skills, knowledge, and disciplines that they embodied required the intentional and

organized teaching and learning that is education. As productivity, wealth, power, and status became more dependent on the mastery and application of knowledge, education to acquire it increasingly occupied the interest and resources of individuals and groups.

The long and discontinuous trajectory of increasing social complexity within and among human societies yielded very little formal and nonformal adult education before the advent of industrialism. Prior to the Neolithic revolution, education of any type was rare; informal socialization without conscious, systematic intent to train or study sufficed for most cultural transmission and role acquisition. Agrarian, state-organized societies, especially the early and late classical civilizations of the Middle East, Asia, and Europe, developed the first formal and nonformal schooling in response to the increasing complexity of knowledge, administration, social control, and production. This schooling was delivered by professional tutors and early versions of primary and secondary schools to educate the children of political, military, and religious aristocracies for their rulership roles; in schools and colleges to train bureaucratic and religious functionaries, professionals in law and medicine, and the elite in the liberal arts; and by nonformal systems of apprenticeship, such as the medieval guilds, for specialized crafts and trades. Although there are many examples of adults seeking informal education from adepts in the arts, religion, and natural philosophy, the educational systems of agrarian societies were devoted mostly to preparation for adult roles.

The widespread and diverse adult education of the present era emerged in response to the development of modern, urban, scientifically and technologically complex societies. Education became an important and dynamic institutional sector, one that gradually extended its territory from basic schooling for the literacy, numeracy, and general knowledge necessary to market relationships, industrial production, and democratic politics to adult continuing education for advanced professional workers in the theoretical and applied sciences. With some exceptions, the pattern of extension was from earlier to later stages of the life cycle (finally yielding education for learning that is “lifelong”) from the upper to the lower reaches of class and status hierarchies (yielding



“universal education”) and from informal education occurring in avocational and domestic contexts, to nonformal education conducted by voluntary organizations to increasingly institutionalized formal systems.

Prior to industrialization, informal and nonformal adult education was relatively widespread in Europe, especially in England and North America, and especially among the growing urban middle class of artisans and merchants. As this new middle class sought to acquire its share of the growing stock of culture, and as literacy spread and became intrinsic to social and economic participation, adults increasingly engaged in self-directed study (aided by a publishing explosion that included a growing number of “how-to” handbooks) participated in informal study groups, and established cultural institutes and lyceums that delivered public lectures and evening courses of study.

Systematic efforts to spread adult education to the working class and the general population emerged during the process of industrialization. Until the last decades of the nineteenth century, these were almost entirely voluntary efforts devoted to democratization, social amelioration, and social movement goals. In both Britain and the United States, mechanics’ institutes, some with libraries, museums, and laboratories, delivered education in applied science, taught mechanical skills, and conducted public lectures on contemporary issues. In Scandinavia, “folk high schools” performed similar functions. Religious groups conducted literacy campaigns among the new urban masses and established adult educational forums in organizations such as the Young Men’s Christian Association (YMCA). Women’s suffrage groups, labor unions, abolitionists, socialists, and many others developed educational programs both to develop their members and as an organizing tool. After 1860 and until World War I, efforts to popularize education among adults continued in various ways: in an extensive network of lyceums, rural and urban Chautauquas and settlement homes; in educational efforts to aid and acculturate immigrants to the burgeoning industrial cities; and in post-slavery self-improvement efforts of African Americans. These middle- and working-class adult education activities received considerable support and extension from the spread of public libraries.

In most societies, state support for and sponsorship of adult education has always been scant in comparison to that for schooling children and adolescents. Late nineteenth- and twentieth-century imperialism included modest adult educational efforts designed to selectively spread literacy and educate indigenous peoples as functionaries for colonial administrations. Some higher-level adult education of the type delivered by the mechanics’ and folk institutes received public funding, especially in Nordic societies where leisure time and adult continuing education were well provisioned by the state. With the Hatch Act of 1887 and enabling legislation in 1914, American land-grant universities were charged with developing “extension” services to deliver a wide range of educational and technical assistance to farmers and rural populations. The agricultural extension model later expanded to include technical training and assistance for industry, and night schools and continuing education for adults. Public funding for these developments was minimal; extension services beyond the land-grant mandate, to the present, have usually been expected to operate as fiscally self-supporting units.

Robust state support for adult education occurred only in the socialist societies that emerged after the Russian Revolution of 1917. Beginning with the mass literacy campaign initiated by the Bolsheviks, adult education was an important priority of the Soviet regime (Lee 1998). It established a large and wide-ranging formal adult continuing education system, with branches in virtually all institutional sectors of Soviet life. It was intended to foster ideological allegiance to the Soviet system, develop the Soviet workforce, spread socialist culture, and foster progress in sports and the arts. A nonformal, decentralized system of “people’s universities” paralleled and reinforced the formal system. After World War II, Soviet-style adult education was transplanted to most of its Eastern European satellites and, usually in severely truncated form, to the developing societies of Asia, Africa, and Latin America that followed the Soviet model. After the communist victory in 1949, China began its own mass literacy campaign for adults. Later it developed a near-universal system of “spare-time” and other adult schools that trained students for very specific political, economic, and cultural roles in the new society, and for political and ideological conformity to the tenets and goals

of the revolution. In the decade following 1966, China conducted one of the largest and most disastrous adult education programs in history. The Cultural Revolution spawned a severe and mandatory form of experiential learning that replaced much of higher education by sending professionals and officials with suspect class pedigrees, intellectuals, professors, and students into agricultural villages and factories to be reeducated by proletarian labor and incessant exposure to Mao-sanctioned communist propaganda.

In spite of the social control service to the dominant regime that it was designed to deliver, socialist adult education played a very significant role in the rapid industrialization and modernization of agrarian Russia and China. In the market societies of the West, the unplanned and longer-term realization of such development did not entail centralized efforts to transform adults for new roles. The systems of political and social control did not generally require the systematic and continuous indoctrination of adults. Thus, until the second half of the twentieth century, adult education in the West remained largely a function of individual and small group informal education, voluntary and philanthropic organizations, relatively small government efforts, specialty units of the system of formal schooling, and small training programs in proprietary schools and businesses.

#### **CONTEMPORARY STRUCTURES AND PARTICIPATION**

After 1950, adult education grew in scale and organization. The pattern of this growth involved a shift from adult education delivered by community-based organizations to that provided by formal educational institutions and the training industry. Demand for education among adults grew because of increasing rates of technological innovation, professional specialization, organizational complexity, knowledge intensiveness in goods-and-services production, and rising credential requirements for employment. Adults returned to secondary and postsecondary education in steadily increasing numbers. In the United States, the GI Bill sent the first large wave of these new students into colleges and universities after World War II (Olson 1974). In the 1960s and 1970s new or expanded units of formal education were organized to accommodate and recruit adult students in

both Europe and the United States: college adult degree programs, open universities, evening colleges, programs for accrediting experientially based adult learning, and government and school-district sponsored adult high school completion programs.

Higher education attendance among adults continued to accelerate in the 1980s and 1990s as a result of several factors: multiple stop-in, stop-out college career paths; rapid obsolescence of knowledge and associated multiple career changes throughout a lengthening cycle of life and work; increased requirements and rewards for specialized graduate, certificate, and technical education; and, most significantly, the large returning student population among Baby Boom adults. In turn, the growing market of adult students spawned new forms of higher education. In the 1990s, the for-profit, fully accredited University of Phoenix became the largest private university in the United States, with over 65,000 adult graduate and undergraduate students (Winston 1999). Phoenix provides a library that is entirely electronic, uses both branch campuses in strip malls and office buildings and Internet-delivered courses to provide education nationwide, and employs only forty-five full-time faculty and more than 4,500 adjuncts. Internationally, “mega-universities,” such as Britain’s Open University and China’s TV University—100,000 and 500,000 students respectively—deliver education to adults through diffuse networks of distance education technologies and part-time local mentors. Collectively, the new virtual (i.e. based entirely on the Internet), distance, for-profit, and mega-universities, all oriented primarily to adults, are beginning to significantly alter the structure and practice of higher education worldwide.

Training and development programs witnessed similar growth after World War II under the impact of a combination of forces. New technologies, especially those associated with information technology production and use, and increased rates of innovation, required education for entirely new skill sets and for continuous upgrading of existing knowledge. Increased job mobility required more new worker training. More intense, global economic competition and new understandings of the contribution of training to productivity and profitability led to greater training investments by firms, agencies, and individuals.

The growing volume of adult training is delivered in many forms and sectors of society and has important consequences. In the advanced industrial societies of Europe and Asia, training and development systems, usually involving employers, unions, and governments, provide training in a substantial majority of firms and to most workers (Lynch 1994). Yearly participation rates by both firms and individuals are highest in the European and Japanese systems and may be an important factor in their high rates of productivity growth. Even in the more diffuse U.S. system with lower participation, over 60 percent of adults between twenty-eight and fifty years of age had participated in some variety of part-time adult education or training by the early 1990s (Hight 1998). Workplace training alone became a \$60-billion a year industry with 53,000 providers (Martin 1998). Some firms supply most of their own training through another new institutional form of adult education, the "corporate university." With curricula ranging from basic employee orientation to the most advanced technical and business subjects, several of these universities are now fully accredited to grant baccalaureate and graduate degrees. As the U.S. military's reliance on technologically complex strategies, tactics, logistics and equipment grew, it became the largest single provider of training in the United States. Even with post-Cold War downsizing, the military's adult training activities remain significant contributors to labor-force development, especially in computer, electronics, and mechanical specialties (Barley 1998). Higher education institutions are also major training providers. Collectively, college and university adult continuing and professional education now rivals or surpasses the volume of military training (Gose 1999). Harvard serves over 60,000 adults each year in continuing education classes. New York University's School of Continuing and Professional Studies offers more than one hundred certificate programs and has revenues of over \$90-million a year. The involvement of universities in the expansion of training presents something of a paradox. In many instances, the new training complex is endowing its adult students with professional and industry certifications that are beginning to rival standard degrees as the credentials of both individual and employer choice. This training may also be working significant changes in the social structure. Considerable evidence suggests that a large

measure of the growing income inequality characteristic of advanced industrial societies is a function of technological change and associated wage premiums paid for workers with the kind of technical credentials and competencies that much of the new training is designed to deliver (Bassie 1999).

Contemporary developing societies of Africa, Asia, and Latin America contain all of the forms of adult education found in the West, as well as forms of popular adult education now only minimally present there. While virtually all of the less developed societies have succeeded in establishing systems of primary, secondary, and higher education for youth, participation rates vary widely among and within them. Some Latin American countries, for example, have higher college attendance rates than their European counterparts while many poor children never enter or complete primary school (Arnové et al. 1996). In such contexts, adult education is often bifurcated (Torres 1990). The poor are served by literacy programs, popular education efforts developed by local nongovernmental organizations, and the health and technical education programs of international agencies. Middle and upper classes participate in corporate training, government-sponsored adult higher education, and professional continuing education programs that are little different from and, in the case of training delivered distance education technologies and global firms, exactly the same as those found in the most advanced societies. However, under the impact of globalization, the distinctions among higher education, human resource training, and popular education are softening. There is a growing recognition among education professionals and policy makers that widespread and inclusive lifelong learning is a critical common good (Walters 1997).

Lifelong learning and adult education that is on a par with schooling are increasingly articulated goals of social policy. The United Nations concluded its fifth Conference on Adult Education in 1997 with a call for worldwide lifelong learning to promote social and economic development, empower women, support cultural diversity, and incorporate new information technologies. Adult education and training is now a central focus of the United Nations Education Scientific and Cultural Organization's (UNESCO) development policy, planning, and programs. The European Union

designated 1996 the “European Year of Lifelong Learning” and established the European Lifelong Learning Initiative to research, coordinate, and develop lifelong learning policies and programs throughout Europe. In 1997, the Commission for a Nation of Lifelong Learners, representing U.S. business, labor, government, and education interests, called for policies that would ensure equity of access to lifelong learning, optimize the use of new technologies, reorganize the education and training system into one providing comprehensive lifelong learning, and commit private and public resources sufficient to these ends. The federal administration established a yearly summit on lifelong learning to pursue implementation of these goals. Similar policy formation and advocacy is now widespread at the local level, and cities recognize the importance of lifelong learning to workforce development and urban economic viability in the new “knowledge economy.” It remains to be seen whether the new focus on adult education policy will result in the kind of resource deployment or accessible, comprehensive system for lifelong learning recommended by most analysis.

### SOCIOLOGY OF ADULT EDUCATION

The sociology of education has traditionally focused on the processes, structures, and effects of schooling, with very little attention to adult education. Thus, there is no scholarly or applied field, no distinct body of theory or research that can be properly labeled “the sociology of adult education.” Expositions of a sociology of adult education that do exist tend to be general characterizations of the field in terms drawn from general sociological and schooling literature (e.g., Rubenson 1989; Jarvis 1985). Certainly much of this is applicable. Adult education presents many opportunities for sociological interpretation in terms of the functionalist, conflict, reproductive, or postmodernist educational models applied to schooling; in terms of adult education’s impact on social mobility, status attainment, and distributional equity; or in terms of its contribution to organizational effectiveness, social welfare, or economic development. The small body of sociological research in the area takes just such an approach, focusing on issues such as: the patterns and causes of adult enrollment in higher education (e.g., Jacobs and Stoner-Eby 1998); the consequences of women’s return to

higher education (e.g., Felmlee 1988); and the patterns of access to and provision of employer-provided training (e.g., Jacobs, Lukens, and Useem 1996). However, neither the extent nor the depth of sociological study in the area matches its potential importance.

Adult education is clearly no longer a social activity marginal to schooling. The information-technology revolution, the knowledge society, the learning organization, and the increasingly critical concern of individuals and organizations with the development and control of human and knowledge capital are representative of trends that are bringing adult education into the mainstream of social interest, analysis, and policy. Responses to these trends for which adult education is central include new and powerful forms of higher education, new determinants of the structure and process of social inequality, new bases for economic and social development, and new conceptions of the relationship of education to the human life cycle. Many opportunities for the application of sociological theory, research, and practice to adult education await realization.

### REFERENCES

- Arnove, Robert F., Alberto Torres, Stephen Franz, and Kimberly Morse 1996 “A Political Sociology of Education and Development in Latin America.” *International Journal of Comparative Sociology* 37:140–158.
- Barley, Stephen R. 1998 “Military Downsizing and the Career Prospects of Youths.” *The Annals* 559:41–157.
- Bassi, Laurie J. 1999 “Are Employers’ Recruitment Strategies Changing?” In N. G. Stacey, project manager, *Competence Without Credentials*. Jessup, M.: U.S. Department of Education.
- Brookfield, Stephen D. 1987 *Developing Critical Thinkers: Challenging Adults to Explore Alternative Ways of Thinking and Acting*. San Francisco: Jossey-Bass.
- Coben, Diana, Joe L. Kincheloe, and Diana Cohen 1998 *Radical Heroes: Gramsci, Freire and the Politics of Adult Education*. New York: Garland Publishing.
- Courtney, Sean 1989 “Defining Adult and Continuing Education.” In S. B. Merriam and P.M. Cunningham, eds., *Handbook of Continuing and Adult Education*. San Francisco: Jossey-Bass.
- Craig, Robert L. (ed.) 1996 *The ASTD Training and Development Handbook: A Guide to Human Resource Development*, 4th ed. New York: McGraw Hill Text.

- Felmlee, Diane 1988 "Returning to School and Women's Occupational Attainment." *Sociology of Education* 61:29-41.
- Gose, Ben 1999 "Surge in Continuing Education Brings Profits for Universities." *The Chronicle of Higher Education* 45(February 9):A, 24.
- Hight, Joseph E. 1998 "Young Worker Participation in Post-School Education and Training." *Monthly Labor Review* 121:14-21.
- Jacobs, Jerry A., Marie Lukens, and Michael Useem 1996 "Organizational, Job, and Individual Determinants of Workplace Training: Evidence from the National Organizations Survey." *Social Science Quarterly* 77:159-177.
- Jacobs, Jerry A., and Scott Stoner-Eby 1998 "Adult Enrollment and Educational Attainment." *The Annals* 559:91-108.
- Jarvis, Peter 1985 *The Sociology of Adult and Continuing Education*. London: Croom-Helm.
- Knowles, Malcolm S. 1980 *The Modern Practice of Adult Education: From Pedagogy to Andragogy*, 2nd ed. Chicago: Association/Follett.
- Knox, Alan B. 1986 *Helping Adults Learn*. San Francisco: Jossey-Bass.
- Lee, David Currie 1998 *The People's Universities of the USSR*. Westport, Conn.: Greenwood.
- Lynch, Lisa M. 1994 "Payoffs to Alternative Training Strategies at Work." In Richard B. Freeman, ed., *Working Under Different Rules*. New York: Russell Sage.
- Martin, Justin 1998 "Lifelong Learning Spells Earnings." *Fortune* 138(July 6):197-200.
- Merriam, Sharan B., and Rosemary S. Cafferella 1991 *Learning in Adulthood*. San Francisco: Jossey-Bass.
- Olson, Keith W. 1974 *The GI Bill, the Veterans, and the Colleges*. Lexington, Ky.: The University Press of Kentucky.
- Rubenson, Kjell 1989 "The Sociology of Adult Education." In S. B. Merriam and P. M. Cunnigham, eds., *Handbook of Adult and Continuing Education*. San Francisco: Jossey-Bass.
- Torres, Carlos Alberto 1990 *The Politics of Nonformal Education in Latin America*. New York: Praeger.
- Walters, Shirley (ed.) 1997 *Globalization, Adult Education and Training: Impacts and Issues*. Cape Town: Zed.
- Winston, Gordon C. 1999 "For Profit Higher Education." *Change* 31(January/February):13-19.

JOHN HOUGHTON

## ADULTHOOD

Becoming an adult is a life-cycle transition signified by multiple markers (Hogan and Astone 1986). These include the completion of education, the establishment of an independent residence, the attainment of economic self-sufficiency, marriage, parenthood, permission to vote and serve in the military, and the entry into full-time work. These multifaceted, objective markers of adult status, variability in the ages at which they occur, and differences in both preadult and adult roles, make the character of this transition variable across societies, historically relative, and subject to diverse interpretations and subjective meanings. Each new generation's experience of transition to adulthood is somewhat unique, dependent on the particular economic, political and social currents of the time (Mannheim 1952). Institutional contexts (cultural, social, educational, economic) determine the pathways through which the transition to adulthood occurs, as well as the competencies that enable successful adaptation to adult roles.

In addition to the formal markers of transition, there are clearly recognized prerogatives of adult status (e.g., smoking, alcohol use, and sexuality) that are widely frowned upon or legally prohibited when engaged in by minors. Youth's engagement in these "problem behaviors" can be attempts to affirm maturity, gain acceptance by peers, or to negotiate adult status (Jessor and Jessor 1977, p. 206; Maggs 1997). Finally, there are even more subtle, subjective indicators of adulthood—for example, the development of "adult-like" psychological orientations or the acquisition of an adult identity. Considering oneself as an adult may or may not coincide with the formal markers of transition.

There has been a trend in the United States and Western Europe toward earlier assumption of full adult civil rights (e.g., the age at which it is legal to vote or to marry without parental consent) from the age of twenty-one to eighteen (Coleman and Husen 1985). Youth must be age eighteen to vote in France, Germany, Great Britain, and the United States, but must be twenty to vote in Japan (The World Factbook 1998).

Within countries, legal restrictions on the age at which certain events can occur, signifying adulthood to a greater or lesser degree, vary depending

on the marker in question. For example, in the United States, a person has the right to vote and serve in the military at the age of eighteen, but must be twenty-one to purchase alcoholic beverages. At age sixteen, a young person can leave school and work in paid jobs without federal restriction on hours of work. (However, restrictions on work hours of sixteen and seventeen year olds exist in many states; see Committee on the Health and Safety Implications of Child Labor 1998). The legal age of marriage varies widely across U.S. states, from none to twenty-one (The World Factbook 1998), but the age of consent is sixteen to eighteen in most states ([www.ageofconsent.com](http://www.ageofconsent.com)).

However, many eighteen, and even twenty-one year olds in the United States and other modern countries would not be socially recognized as “adults,” since they have not yet accomplished other key markers of transition. The law only sets minimum standards, and may have little relation to the age at which young people actually make the various transitions in question. Despite prohibitions on the use and purchase of alcohol until the age of twenty-one, “it is more normative to drink during adolescence than it is not to drink” (Maggs 1997, p. 349). The majority of young people in the United States stay in school beyond the age of sixteen; they graduate from high school and receive at least some postsecondary education. With the extension of formal education, contemporary youth in modern countries delay the acquisition of full-time employment, and remain economically dependent for longer periods of time. Moreover, the legal age of marriage hardly reflects the average age of first marriage. Since the 1950s, it has been increasing (from 20.3 for women and 22.8 for men in 1950 to 24.4 for women and 26.5 for men in 1992; Spain and Bianchi 1996).

#### **HISTORICAL AND SOCIETAL VARIATION IN THE TRANSITION TO ADULTHOOD**

Neither the timing nor the process of becoming adult are universalistic or biologically determined. Throughout Western history, there has been increasing differentiation of early life stages, postponement of entry to adulthood, and change in the status positions from which adulthood is launched (Klein 1990). Some scholars argue that in medieval times persons moved directly from a

period of infancy, when small size and limited strength precluded productive work, to adulthood, at which time younger persons began to work alongside their elders (Ariés 1962).

A new stage of childhood, between infancy and adulthood, arose with the emergence of schools. As economic production shifted from agriculture to trade and industry, persons increasingly entered adulthood after a stage of apprenticeship or “child labor.” By the beginning of the twentieth century, with schooling extended and child labor curtailed, adolescence gained recognition as the life stage preceding adulthood (Hall 1904). The adolescent, though at the peak of most biological and physiological capacities, remained free of adult responsibilities.

With more than 60 percent of contemporary young people obtaining some postsecondary education (Halperin 1998), a new phase of “youth” or “postadolescence” has emerged, allowing youth in their mid-to-late twenties, and even older youth, to extend the preadult “moratorium” of continued exploration. This youth phase is characterized by limited autonomy but continued economic dependence and concern about the establishment of adult identity (Keniston 1970; Coleman and Husen 1985; Buchmann 1989).

Various forms of independent residence are now common in the United States before marriage (Goldscheider and Goldscheider 1993). Youth’s residence in dormitories (or, less commonly, military barracks) allows independence from familial monitoring, while a formal institution assumes some control (Klein 1990). Even greater freedom from supervision occurs when young people, still economically dependent on parents, live in their own apartments. For contemporary young people in the United States who enter the labor force after high school, a continuing period of “moratorium” (Osterman 1989) lasts several years. During this time, youth hold jobs in the secondary sector of the economy to satisfy immediate consumption needs. They experience high unemployment and job instability (Borman 1991). At the same time, employers express preference for low-wage workers who do not require fringe benefits and are not likely to unionize. When filling adultlike “primary” jobs, such employers seek evidence of stability or “settling down.”

But youth “irresponsibility” and employer reluctance to offer desirable jobs to youthful recruits are not universal in modern societies. Instead, they derive from particular institutional arrangements. In the United States, the absence of clear channels of mobility from education to the occupational sector, and youths’ lack of occupation-specific educational credentials, fosters a prolonged period of trial and instability in the early career (Kerckhoff 1996). In contrast, the highly developed institution of apprenticeship in Germany implies the full acceptance of young workers as adults, and encourages employers to invest in their human capital (Hamilton 1990, 1999; Mortimer and Kruger forthcoming).

The onset of adulthood has thus been delayed through historical time by the emergence of successive preadult life stages. This historical progression of preadult phases refers to the normative, legitimate pathways to adulthood. A highly problematic, but increasingly prevalent way station in the transition is supervision by the criminal justice system; indeed, it is estimated that a full ten percent of U.S. males, aged twenty to twenty-nine, is in jail, in prison, on parole, or on probation on any one day (Halperin 1998).

Obstacles to “growing up” are sometimes presented when assuming adultlike statuses threatens adult interests and values. The extension of required schooling was motivated, at least in part, by a desire to curb competition for jobs with older workers (Osterman 1980). “Warehousing” the young in secondary and tertiary education has reduced adolescent unemployment during times of economic contraction; earlier movement out of school into the workforce is promoted by economic expansion (Shanahan et al. 1998).

Societal wealth may also encourage postponement of adulthood and the extension of “youthful” values and life styles to older ages. Japanese young people, traditionally oriented to the extended family, obedience, educational achievement, and hard work, seem to be becoming more rebellious and interested in immediate enjoyment as delayed gratification becomes more difficult to sustain in a more affluent society (Connor and De Vos 1989; White 1993).

Although age-related increases from birth through the second decade of life—in strength,

cognitive capacity, and autonomy (Shanahan, this volume)—are probably recognized in some form in all human societies, the social construction of the early life course clearly reflects societal diversity and institutional change. Processes of modernization, encompassing changes in education, the labor force, and the emergence of the welfare state, produce age standardization of the early life course (Shanahan forthcoming). For example, the social differentiation of children and adolescents was less pronounced in the educational system in nineteenth- and early twentieth-century America than in the contemporary period (Graff 1995). Secondary school students have become increasingly homogeneous in age; this more pronounced age grading promotes “adolescent” life styles and orientations. Similarly, the extension of postsecondary schooling promotes the perpetuation of “youth.” However, during the past century in the United States the changes marking the transition have taken place first in quicker, and then in more lengthy, succession (Shanahan forthcoming; Modell et al. 1976).

As a result of technological and economic change and increasing educational requirements (Hogan and Mochizuki 1988; Arnett and Taber 1994; Côté and Allahaar 1994), as well as family instability (Buchmann 1989), the entry to adulthood has been characterized as increasingly extended, diversified, individualized (Buchmann 1989; Shanahan forthcoming), “disorderly” (Rindfuss et al. 1987), variable (Shanahan forthcoming), and less well defined (Buchmann 1989). For example, while the acquisition of full-time work is widely considered to mark the transition to adulthood, distinctions between “youth work” and “adult work” blur as young people increasingly combine, and alternate, student and occupational roles, in various forms and levels of intensity through a lengthy period of adolescence and youth (Mortimer and Johnson 1998, 1999; Mortimer et al. 1999; Morris and Bernhardt 1998; Arum and Hout 1998). Markers of family formation likewise become less clear as young people move in and out of cohabiting and marital unions (Spain and Bianchi 1996).

Orderly sequences of transition events have become less common (Shanahan forthcoming). For example, many youth return to their parental homes after leaving for college or other destinations (Cooney 1994). Divergence from normative

timing or sequencing sometimes generates public alarm, becomes defined as a social problem, and leads to difficulties for the young people whose lives exhibit such patterns. Whereas parenting is becoming more common outside of marriage (approximately one-third of births in the United States occur to unmarried women; Spain and Bianchi 1996), it is widely thought that nonmarital birth in the teenage years marks a too-early transition to adulthood. Adolescent parenting is linked to school dropout, difficulties in the job market, restricted income and marital instability (Furstenberg, et al. 1987). (However, Furstenberg and colleagues' 1987 study of a panel of black adolescent mothers sixteen to seventeen years after their children were born actually showed considerable diversity in maternal socioeconomic outcomes.)

Still, assertions that the transition to adulthood is especially indeterminate, or becoming more difficult, in contemporary Western societies remain controversial. Graff (1995) documents the extended and equivocal nature of this transition in eighteenth- to twentieth-century America. Foner and Kertzer (1978) noted ambiguity even in premodern contexts, where elders and the rising adult generation often struggled over the timing of age-set transitions and corresponding transfers of power, wealth, and privilege. It is certain, however, that the transition to adulthood assumes a quite different character historically and across national contexts, and may be more clear in some than in others.

### THE SUBJECTIVE TRANSITION TO ADULTHOOD

A series of psychological, or subjective, changes are expected to occur as young people move into adulthood. The adolescent is said to be oriented to fun, sports, popular music, and peers; receptive to change; and ready to experiment with alternative identities and sometimes, mood-altering substances (Hall 1904). Youth are encouraged to enjoy themselves as they continue to explore their interests and potentials. Osterman (1989) describes out-of-school employed youth as lacking career orientation; instead, they emphasize peer relationships, travel, adventure, and short-term jobs.

Young adults, in contrast, are expected to relinquish such dependent, playful, experimental,

carefree, and even reckless stances of adolescence and youth, so that they can address the "serious business" of life. Those who become financially and emotionally independent, productive, hard-working, and responsible are considered "adult" (Klein 1990). Moreover, they themselves are expected to "feel like" adults (Aronson 1998).

The concept of "maturity" is integrally tied to adulthood. Most generally, it refers to the psychological competencies deemed necessary to adapt to the roles and responsibilities of adulthood (Galambos and Ehrenberg 1997). For Greenberger and Sorenson (1974), the term signifies autonomy, the capacity to make decisions on the basis of life goals and to function independently in work and other spheres; skills in communicating and relating to others; and social responsibility, the motivation and ability to contribute to the wider society. Greenberger and Steinberg (1986) worry that teenagers who have paid jobs become "pseudomature," as they take on adult identities and behavioral prerogatives without being psychologically equipped for them. Such pseudomaturity can precipitate disengagement from more beneficial, albeit dependent and childlike, roles (especially the role of student). Bachman and Schulenberg (1993) similarly note syndromes of adultlike roles, activities, and identities that promote premature entry to adult family and occupational roles.

But like the objective markers of transition, beliefs about the specific attributes that define maturity, and how these competencies may be fostered and recognized, vary across time and social space (Burton et al. 1996). The extent to which youth feel that they possess such capacities, and the likelihood that adults will attribute these qualities to them, may be highly variable across social situations. Furthermore, in postmaterialist societies and especially in the more highly educated and affluent social niches within them, orientations promoting success in the economic sphere may recede in importance as criteria of maturity, in favor of continued "youthful" emphases on freedom and self-actualization throughout adulthood (Vinken and Ester 1992; Inglehart 1990).

Meanings and interpretations of the various transitions signifying acquisition of adult status, such as those linked to family formation (Modell



1989), also exhibit historical and cross-societal variability. In contemporary modern societies, structural differences in the link between school and work (Shavit and Muller 1998) are reflected in the phenomenal experience of the transition into the labor force. Variation in the institutional connection between school and work in the United States and Germany influences the ages at which young people begin to actively prepare for the highly consequential decisions (especially regarding postsecondary schooling and career) that lie ahead of them, and the degree of stress and uncertainty they encounter in doing so (Mortimer and Kruger forthcoming).

It is widely believed that age norms, specifying the timing and order of key life events, influence the subjective passage, as well as the objective trajectory, through the life course (Neugarten et al. 1965; Neugarten and Danan 1973). "On-time" transitions are "culturally prepared" by socialization and institutional arrangements (Model 1989, p. 13), and are thereby rendered psychologically salutary. Those who are "off-time," too early or too late, are thought to be the target of negative social sanctions and to experience psychological strain (Rossi 1980). National polls on the ideal age to marry and become a parent yield age distributions that cluster around the modal ages at which these changes generally occur. But consistency in expectations, "ideal ages," or even in the actual timing of transitions, may have more to do with institutionally determined pathways and other structural constraints than personal norms.

The notion that norms control timing behavior is contradicted by evidence that age preferences ("ideal" ages for marker events) lag behind behavioral change (Modell 1980; McLaughlin et al. 1988, ch. 9). Marini (1984) notes that there is little direct evidence regarding the existence or content of social norms governing the timing of life events. She asserts that if norms (and associated sanctions) do exist, they probably vary by population subgroup (e.g., by socioeconomic status, gender, and ethnicity), and encompass such a wide range of acceptable ages that they lack causal import. In one study, college students' expectations about the ages at which they would most likely traverse various markers of transition were found to be more variable than those of high school students (Greene 1990).

Older ages of marriage and finishing school, coupled with the increasing reversibility of status changes (serial marriages and cohabiting unions, adult education, etc.) could erode age norms, to the extent that they do exist, since such passages are "no longer viewed in a linear way as transitions passed at a given point and left behind permanently" (Arnett 1997, p. 20). Historically, as objective markers of adulthood have become increasingly variable in their timing and sequencing, age norms may have lost whatever force they once had. But in fact, we do not know whether the failure to make particular transitions at an age when one's contemporaries have mostly done so engenders a subjective sense of being "too early" or "too late" and creates stress, either for earlier cohorts or for contemporary young people.

Arnett (1997) found that youth themselves were more likely to choose psychological traits indicative of individualism and responsibility as necessary for a person to be considered an adult; as criteria for adult status, youth reject role transitions or "objective markers" of adulthood, such as finishing school, marriage, or parenthood. Rituals, such as marriage and graduation, have traditionally allowed public recognition of successful passage to adulthood. If such objective markers of transition have less psychological salience than in the past, they may be less important in signifying and reinforcing adult status.

However, in Arnett's study, two such transitions were endorsed by the majority: living outside the parental household and becoming financially independent of parents. These linked changes, signifying autonomy from parents, may have important symbolic meaning for young people as they contemplate becoming adults. So too, three-fourths of the participants endorsed "decide on personal beliefs and values independently of parents or other influences" (p. 11) as criteria of adulthood.

Consistent with Arnett's research, there is apparently no clear correspondence between particular objective markers (e.g. graduating from college, beginning a full-time job) and young people's subjective identification of themselves as "adults." Aronson (1999b) found that some contemporary young women in their mid-twenties do not "feel like" adults, even when occupying roles

widely considered to be indicative of adulthood. Interviews revealed a great deal of ambiguity about adult identity, as well as career uncertainty, despite having already graduated from college and begun full-time work linked to postsecondary areas of study.

Asynchronies in the age-grading systems of different societal institutions generate status inconsistencies (Buchmann 1989) with important psychological implications. If adultlike identities are confirmed in some contexts, such as the workplace, but not in others, such as the family, this discrepancy could produce strain. Moreover, the "loosening" of age-grading and structured sequencing in transitional activities decreases the ability to predict the future from current circumstances (Buchmann 1989), and may thereby engender both stressful situations and depressive reactions (Seligman 1988).

It is widely believed that adolescence and youth are stressful life stages, and that problems diminish with successful acquisition of adult roles (Modell et al. 1976). Consistent with this supposition, youth have been found to exhibit less depressed mood from late adolescence to early adulthood, as they move from high school into postsecondary education and, especially, into full-time work roles (Gore et al. 1997). Moreover, there is evidence that men's self-perceptions of personal well-being and competence decline during college, but rise during the following decade (Mortimer et al. 1982).

However, women's morale may follow a less sanguine trajectory from adolescence through midlife (Cohler et al. 1995). Contemporary women are more likely than those in the historical past to take on multiple and conflicting roles. Difficulties in balancing work and family may be particularly acute and stressful during the transition to adulthood (Aronson 1999a), as increasing numbers of young women attempt to balance the conflicting demands of single motherhood, work, and postsecondary education.

Though the literature tends to focus on problematic outcomes of historical change and the loosening of age-graded social roles, greater diversity in the sequencing and combination of roles (schooling and work, parenting and employment) have promoted more diversified, and autonomous,

courses of action. Indeed, the allowance of more diverse sequences of transitions enables some youth to escape from dissatisfying circumstances in adolescence, for example, by leaving home (Cooney and Mortimer 1999).

The greater individualization and lengthening of the adult transition and early life course in recent times may increase the potential for freedom, and the effective exercise of choice, as well as stress. The extension of formal education allows more time for the exploration of vocational and other life-style alternatives (Maggs 1997). Although change in occupational choice in the years after high school (Rindfuss et al. 1990) may be seen as indicative of a kind of "floundering" and instability, it also may reflect youth's increasing capacity to assess alternatives before making a firm vocational commitment. Modell's (1989) social history of the transition to adulthood in twentieth-century America finds youth increasingly taking charge of their heterosexual relationships and the formation of new families, becoming ever freer of adult surveillance and control. Aronson (1998) found that contemporary young women appreciate their life-course flexibility (Aronson 1998).

#### **THE TRANSITION TO ADULTHOOD AS A CRITICAL PERIOD OF HUMAN DEVELOPMENT**

The transition to adulthood is a highly formative period for the crystallization of psychological orientations relating to work, leisure (Inglehart 1990), and politics (Glenn 1980). Alwin and colleagues' (1991) study of a panel of Bennington college women from the 1930s to the 1980s reports extraordinary persistence of political attitudes formed while in college over an approximately fifty-year period (a stability coefficient of .781). Work orientations also become more stable following early adulthood (Lorence and Mortimer 1985; Mortimer et al. 1988). Three explanations have been put forward to account for this pattern: The first implicates the environment; the second, features of the person; the third combines both elements. According to the first line of reasoning, the relatively dense spacing of major life events during the transition to adulthood generates external pressures to form new attitudes or to change previous

views (Glenn 1980). While similar events can occur later in life (e.g., a job or career change, remarriage, or entry into an adult education program), they are usually spaced more widely and are like those experienced previously, rather than wholly new circumstances requiring adaptation. Similarly, experiences at work generally assume greater constancy after an initial period of job instability (Osterman 1980). Primary relationships that provide support for attitudes are often in flux during the transition to adulthood; thereafter, stable primary groups may provide continuing support for previously crystallized attitudinal positions (Sears 1981; Backman 1981; Alwin et al. 1991). According to this perspective, there may be continuing capacity to change throughout life (Baltes et al. 1980), but if environments become more stable after the adult transition, there will be less impetus for such change (Moss and Susman 1980).

A second explanation links the “aging stability” pattern to intrapersonal processes that increasingly support the maintenance of existing personality traits, and resistance to change. Mannheim’s (1952) classic concept of generation implies that the young are especially receptive to influences generated by the key historical changes of their time (economic upheaval, war, or political revolution). Analysis of data from the European Values Survey shows that younger cohorts are less religiously traditional, are more sexually and morally permissive, and value personal development more highly than older cohorts (Vinken and Ester 1992).

Alwin and colleagues (1991) find evidence for a “generational/persistence model,” which similarly combines notions of vulnerability in youth and persistence thereafter. Before role and character identities are formed, the person may be quite malleable. However, preserving a consistent, stable sense of self is a major motivational goal (Rosenberg 1979); and once self-identities are linked to key attitudes and values, the person’s self may become inextricably tied to those views (Sears 1981). Moreover, feelings of dissonance (Festinger 1957) arise when attitudes and beliefs that provide a sense of understanding are threatened (Glenn 1980). Consistent with the notion that young adults may be more ready to change, occupational experiences (i.e., related to autonomy) have stronger influences on the work orientations of younger

workers (ages sixteen to twenty-nine) than of those who are older (Lorence and Mortimer 1985; Mortimer et al. 1988).

According to a third point of view, an interaction between the young person and the environment fosters a process of “accentuation” of preexisting traits. While early experiences provide initial impetus for personal development, attitudes and values formed in childhood or adolescence are later strengthened through the individual’s selection, production, and/or maintenance of environmental circumstances that support earlier dispositions. According to this view, youth making the transition to adulthood select and/or mold their environments (Lerner and Busch-Rossnagel 1981) often so as to maintain (or to reinforce) initial psychological states. This process typically results in an “increase in emphasis of already prominent characteristics during social transitions in the life course” (Elder and Caspi 1990, p. 218; see also Elder and O’Rand 1995).

Such processes of accentuation take many forms. For example, students choosing particular college majors become increasingly similar in interests and values over time (Feldman and Weiler 1976). Mortimer and colleagues’ (1986) study of a panel of young men showed that competence measured in the senior year of college predicted work autonomy ten years later, which in turn, fostered an increasing sense of competence. Similarly, intrinsic, extrinsic and people-oriented values prior to adult entry to the workforce led to the selection of occupational experiences that served to strengthen these value preferences. A similar pattern was found among high school students seeking part-time jobs; intrinsic values predicted opportunities to acquire skills and to help others at work, and further opportunities for skill development strengthened intrinsic values (Mortimer et al. 1996). Alwin and colleagues’ (1991) follow-up study of women who attended Bennington College in the 1930s indicated that the choice of associates and the formation of supportive reference groups—e.g., spouses, friends, and children—played a substantial part in maintaining the women’s political values.

Elder’s longitudinal study of young people growing up during the Great Depression amassed considerable evidence that successful encounters

with problems in adolescence can build confidence and resources that promote effective coping with events later in life, fostering personality continuity (Elder 1974; Elder et al. 1984). Thus, early achievements and difficulties can give rise to “spiralling” success and failure. For members of the Oakland cohort, early economic deprivation provided opportunity to help the family in a time of crisis; the consequent increase in self-efficacy, motivation, and capacity to mobilize effort fostered adult work and family security. Elder and Caspi (1990) similarly find that adolescents with more resilient personalities reacted more positively as young adults to combat in World War II. Such reciprocities are also evident among women in the National Longitudinal Survey of Youth; early self-esteem predicted educational attainment and more substantively complex employment, which, in turn, further enhanced their sense of worth (Menahgan 1997).

In contrast, negative early events may set in motion processes that accentuate problems. Early failures can produce psychological reactions and cognitive attributes that perpetuate poor outcomes. There is evidence from several studies that early unemployment fosters distress, self-blame, and negative orientations toward work in general, engendering continued failure in the labor market (Mortimer 1994). Traumatic war experiences in early adulthood can threaten marriage and thereby reinforce a cycle of irritability (Elder and Caspi 1990, p. 235).

Whereas sociologists emphasize the social determination of early adult outcomes, and social psychologists have noted enduring personality traits that influence the process of transition to adulthood, changes in both socioeconomic and personal trajectories do occur, frequently at times of life-course transition. Change can occur as a result of “fortuitous events” that intervene in the developmental process rather than reinforcing patterns of preadult behavior (Elder and Caspi 1988, p. 102). For example, marriage to a nondeviant spouse, the quality of a first marriage, or of economic self-sufficiency from a full-time job may lead to change in direction of a previously “disorderly” or otherwise problematic early life course (Rutter and Quinton 1984; Sampson and Laub 1993; Gore et al. 1997). “Identity transformations” (Wells and

Stryker 1988) can also result when “turning points” (Strauss 1959) in personal history intersect with a rapidly changing historical context to alter a previously held worldview (Aronson forthcoming).

#### **FACTORS INFLUENCING SUCCESSFUL ACQUISITION AND ADAPTATION TO ADULT ROLES**

Social scientists are giving increasing attention to processes of individual agency, including goal setting, choice among alternative lines of action, and the mobilization of effort, which influence trajectories of attainment throughout the life course. Early orientations toward, and expectancies about, competent action are critical for later adult success (Mainquist and Eichorn 1989). Jordaan and Super (1974) report that adolescents’ planfulness, responsibility, and future orientation predicted their level of occupational attainment at the age of twenty-five. The more explorative adolescents, who were actively engaging of the environment, had more positive early adult outcomes. “Planful competence,” denoting ambition, productivity, and dependability in adolescence, has also been linked to men’s adult occupational status and marital stability (Clausen 1991, 1993). These attributes imply planfulness, delayed gratification, an intellectual orientation, and a sense of control over goal attainment. Planfully competent adolescents actively explore future options and opportunities, and select those that match their developing proclivities and potentialities. This process gives rise to a better fit between the person and the environment, fostering satisfaction and stability in adult social roles (see Shanahan and Elder 1999).

However, the institutional structure, and the individual’s place within hierarchies establishing unequal resources and opportunities, may either facilitate or limit the effective exercise of agency. Shanahan (forthcoming) speaks of “limited strategic action,” resulting from “the dynamic tension between selection and assignment.” Because the relation between courses of study and higher educational outcomes is often obscure, students may have limited ability to alter negative educational trajectories (Dornbusch 1994). The consequences of planfulness may also be historically variable, depending on the degree of opportunity available

to a cohort at critical phases of its life course (Shanahan and Elder 1999).

As the process of acquiring adultlike markers of transition becomes more complex and ambiguous, a wide variety of psychological orientations—including aspirations, values, goals, life plans, self-concepts, and identities—may become increasingly determinative of subsequent outcomes (Mortimer 1996). In fact, both general (Bandura 1996) and facet-specific (Grabowski et al. 1998) dimensions of efficacy are predictive of goal-directed behavior and achievement. Fact-specific orientations include expectations relevant to particular spheres, such as school and work.

The social structural conditions and circumstances that enable some young people to continue to pursue their goals despite obstacles, and others to relinquish them, deserve further study. However, in some circumstances, more favorable outcomes will accrue to those who are more flexible. Kerckhoff and Bell (1988) find that the achievement of postsecondary educational credentials in the form of some vocational certificates yield higher earnings than attaining only some college. The relative merit of tenacious goal pursuit (e.g. a baccalaureate degree) versus the substitution of new, more realistic, alternative objectives (“assimilation” and “accommodation” in Brandtstadter’s [1998] terminology) may be determined by structurally constrained resources as well as opportunities.

The character and outcomes of the transition to adulthood are clearly dependent on diverse resources that are differentially distributed among young people (Shanahan forthcoming). There are social class differences in the age at which adult roles are acquired, in the character of marking events, and even in the availability of opportunities to assume adult-status positions (Meijers 1992). The socioeconomic background of the family of origin sets the level of available resources, fostering intergenerational continuity in attainment (Blau and Duncan 1967; Sewell and Hauser 1976; Kerckhoff 1995). Relative advantage or disadvantage can derive from placement in familial and other networks that provide information (Granovetter 1974; Osterman 1989), for example, about higher educational opportunities, jobs (Lin 1992), or even prospective marital partners. A lack of resources, as well as instability (McLanahan and Sandefur 1994),

in the family of origin is associated with disadvantaged transitions to adulthood. Moreover, structural sources of cumulative advantage, such as advantageous placement in ability groups, high school tracks or secondary schools, increase the likelihood of higher education (Kerckhoff 1993; Garmoran 1996).

Personal resources facilitating the educational and occupational attainment process have been linked to social class background. Adolescents’ educational and occupational aspirations, and their educational attainments, are important mediators of the effects of occupational origins on destinations (Featherman 1980; Featherman and Spenner 1988). The transmission of self-directed values may also constitute a mechanism through which socioeconomic status is perpetuated across generations (Kohn and Schooler 1983; Kohn et al. 1986). Close father-son relationships in late adolescence engender continuity in paternal occupations and values, and sons’ work values and early adult-occupational destinations (Mortimer and Kumka 1982; Ryu and Mortimer 1996). Parents of higher socioeconomic level typically engage in more supportive child-rearing behavior (Gecas 1979), which fosters the development of personality traits such as competence, work involvement, and positive work values, that facilitate early adult socioeconomic attainments (Mortimer et al. 1986).

Gender differences in future orientations can foster differences in achievement. For example, if young women view their futures as contingent on the needs of future spouses, children, and others, this will diminish their propensity to make firm plans (Hagestad 1992), thereby diminishing their attainment. Despite dramatic changes in adult women’s employment (McLaughlin et al. 1988; Moen 1992), and in contrast to young men’s occupational aspirations, many young women are “talking career but thinking job” (Machung 1989, pp. 52–53).

Geissler and Kruger (1992) have identified different patterns of contemporary “biographical continuity” among young German women, each having important implications for career achievement. Traditionally-oriented women expect to have limited labor force participation and to be economically dependent on a husband. Career-oriented young women emphasize the acquisition of professional qualifications and delay marriage. Still

others are actively concerned with both work and family spheres. Such divergence in future orientations and planning influences the intensity of striving for achievement and attainment during the transition to adulthood.

Moreover, to the extent that young women are aware of the difficulties adult women face, resulting from employer discrimination and the unequal division of family work, this would likely depress expectations for labor market success. Ambivalence about work could exacerbate the detrimental consequences of other psychological differences (relative to males) for socioeconomic attainment, such as lower self-efficacy, lower self-esteem, and higher levels of depressive affect (Simmons and Blyth 1987; Gecas 1989; Finch et al. 1991; Shanahan et al. 1991; Mortimer 1994). It is therefore not surprising that although women have narrowed the historical gap in educational attainment (McLaughlin et al. 1988), they often get “diverted” from their initial plans, emphasizing romance over academics (Holland and Eisenhart 1990). For today’s women, taking on adult work and family roles is more likely than for men to result in the termination of schooling (Pallas 1993).

If the family is an important institutional context for the acquisition of economic and other resources for the adult transition, family poverty or disintegration may be expected to have negative consequences. Experiences in youth may thus set in motion a train of events that have a profound impact on the early life course. Disruption and single parenthood in the family of origin, and the economic loss and emotional turmoil that frequently ensue, may jeopardize parental investment in children and youth. However, Coleman (1994) argues that declining investments in the next generation may occur even in more favorable and affluent circumstances. As the functions of the family are transferred to other agencies in welfare states (e.g., as the government takes over education, welfare, support of the aged, and other functions), there is a declining economic dependence of family members on one another. As the multigenerational organization and functions of the family weaken, parental motivation to invest attention, time, and effort in the younger generation may also decrease throughout the population.

However, Shanahan (forthcoming) speaks of “knifing off” experiences during the transition

to adulthood, which can enable some youth to escape from poverty, familial conflict, stigmatization, and other debilitating circumstances of their childhood and teen years. For example, service in the military enabled many men from disadvantaged backgrounds to extricate themselves from the stressful circumstances of their families, mature psychologically, and therefore be in a better position to take advantage of educational benefits for veterans after World War II (Elder and Caspi 1990).

The availability of opportunities for anticipatory socialization or practice of adult roles may also affect adaptation to them. Contemporary youth spend much of their time in schools, cut off from meaningful contact with adult workers (excepting their teachers). Some have expressed concern that this isolation from adult work settings reduces opportunities for career exploration and encourages identification with the youth subculture (Panel on Youth 1974). Many parents encourage adolescent children to work, believing that this experience will help them to become responsible and independent, to learn to handle money, and to effectively manage their time (Phillips and Sandstrom 1990). However, “youth jobs” are quite different from adult work, involving relatively simple tasks and little expectation of continuity.

The impacts of part-time employment during adolescence for the transition to adulthood is the subject of much controversy (Committee on the Health and Safety Implications of Child Labor 1998, ch. 4). While the full impact of employment in adolescence is not known, there is evidence that increasing investment in paid work (as indicated by the number of hours spent working per week) is associated with reduced educational attainment (Marsh 1991; Chaplin and Hannaway 1996; Carr et al. 1996). However, several studies have shown that employment during high school also predicts more stable work histories and higher earnings in the years immediately following (Mortimer and Finch 1986; Marsh 1991; Ruhm 1995, 1997; Mortimer and Johnson 1998). Stern and Nakata (1989), using data from noncollege youth in the National Longitudinal Survey of Youth, report that more complex work activity in adolescence is associated with lower incidence of unemployment and higher earnings three years after high school.

The meaning of adolescent work also influences subsequent educational and occupational

outcomes. For example, employment has been found to have a positive effect on high school students' grades when the workers are saving their earnings to go to college (Marsh 1991; Ruscoe et al. 1996). Consistently, youth who effectively balance their part-time jobs with school, working near continuously during high school but restricting the intensity of their employment to twenty hours a week or fewer, were found to have high early-achievement orientations and obtained more months of post-secondary education (Mortimer and Johnson 1998).

There is further evidence that the quality of adolescent work experience matters for psychological outcomes that are likely to influence adult attainment (Mortimer and Finch 1996). For example, adolescent boys who felt that they were obtaining useful skills and who perceived opportunities for advancement in their jobs exhibited increased mastery (internal control) over time; girls who thought that they were being paid well for their work manifested increasing levels of self-efficacy (Finch et al. 1991). Exposure to job stressors, in contrast, heightened depressive affect (Shanahan et al. 1991).

For the most disadvantaged segments of society there is concern that poor educational opportunities and a rapidly deteriorating economic base in the inner cities preclude access to youth jobs as well as to viable adult work roles (Wilson 1987), irrespective of personal efficacy, ambition, or other traits. The shift from an industrial- to a service-based economy has lessened the availability of entry-level employment in manufacturing. Economic and technological change has created a class of "hard-core unemployed"; those whose limited education and skills place them at a severe disadvantage in the labor market (Lichter 1988; Halperin 1998). Black males' lack of stable employment in U.S. inner cities limits their ability to assume the adult family role (as male provider) and fosters the increasing prevalence and legitimacy of female-headed families.

Research on African-American adolescents suggests that the transition to adulthood may differ significantly from that of non-minority teens. Ogbu (1989) implicates beliefs about success as critical to understanding the paradox of high aspirations among black adolescents and low subsequent

achievement. The "folk culture of success," fostered by a history of discrimination and reinforced by everyday experience (e.g., the observation of black career ceilings, inflated job qualifications, housing discrimination, and poor occupational achievement despite success in school), convinces some young blacks that desired occupational outcomes will not be assured by educational attainment. Given the belief that external forces controlled by whites determine success, alternative strategies for achievement may be endorsed—hustling, collective action, or dependency on a more powerful white person. These, in turn, may diminish the effort in school that is necessary to obtain good grades and educational credentials.

In the context of urban poverty and violence, youth expectations for a truncated life expectancy (an "accelerated life course"), may lessen the salience of adolescence as a distinct life stage (Burton et al. 1996). With few employment and educational opportunities available, these youth often focus on alternative positive markers of adulthood, such as becoming a parent, obtaining material goods, or becoming involved in religious activities. Under such conditions, even a menial job can engender and reinforce "mainstream" identities and "possible selves" as economically productive working adults (Newman 1996). However, structural opportunities pervasively influence achievement orientations and outcomes throughout the social class hierarchy; such processes are clearly not limited to any particular societal stratum (Kerckhoff 1995).

Supportive families and other social bonds are predictive of successful adjustment in the face of poverty and other disadvantages (Ensminger and Juon 1998). Families who are successful in these circumstances place great emphasis on connecting their adolescent children with persons and agencies outside the immediate household, in their neighborhoods and beyond, thereby enhancing their social networks and social capital, invaluable resources in the transition to adulthood (Furstenberg et al. 1999; Sullivan 1989).

In summary, becoming an adult involves changes in objective status positions and in subjective orientations. The transition to adulthood has changed through historical time as a result of economic, political, and social trends. Contemporary young people face increasingly extended and

individualized paths in the timing and sequencing of their movement into adult roles. The transition to adulthood is found to be a critical period for the crystallization of key orientations and values. Adult status placement and adjustment are influenced by socioeconomic background and poverty, gender, early employment experiences, and psychological resources.

The multiplicity of transition markers, the variety of macrostructural and personal influences, and the little-studied subjective component give rise to important challenges and complexities in the study of this phase of life.

## REFERENCES

- Alwin, Duane F., Ronald L. Cohen, and Theodore M. Newcomb 1991 *Political Attitudes over the Life Span: The Bennington Women after Fifty Years*. Madison: University of Wisconsin Press.
- Ari@acute;s, Philippe 1962 *Centuries of Childhood: A Social History of Family Life*. New York: Vintage.
- Arnett, Jeffrey Jensen 1997 "Young People's Conceptions of the Transition to Adulthood." *Youth and Society* 29:3-23.
- , and Susan Taber 1994 "Adolescence Terminable and Interminable: When Does Adolescence End?" *Journal of Youth and Adolescence* 23:5, 517-537.
- Aronson, Pamela 1999a "The Balancing Act: Young Women's Expectations and Experiences of Work and Family." In Toby Parcel, ed., *Research in the Sociology of Work*. Greenwich, Conn.: JAI Press.
- 1999b Blurring life course stages: Young women's transition from adolescence to adulthood in the contemporary era. Paper under review.
- 1998 "Coming of age in the 1990s: Women's identities, life paths, and attitudes towards feminism." Ph.D. diss., Minneapolis: University of Minnesota.
- Forthcoming "The Development and Transformation of Feminist Identities under Changing Historical Conditions." In Timonhy J. Owens, ed., *Advances in Life Course Research: Identity across the Life Course in Cross-Cultural Perspective*. Greenwich, Conn.: JAI Press.
- Arum, Richard, and Michael Hout 1998 "The Early Returns: The Transition from School to Work in the United States." In Yossi Shavit and Walter Muller, eds., *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford: Clarendon Press.
- Bachman, Jerald G., and John Schulenberg 1993 "How Part-time Work Intensity Relates to Drug Use, Problem Behavior, Time Use, and Satisfaction among High School Seniors: Are These Consequences or Merely Correlates?" *Developmental Psychology* 29:220-235.
- Backman, Carl W. 1981 "Attraction in Interpersonal Relationships." In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Baltes, Paul B., Hayne W. Reese, and Lewis P. Lipsitt 1980 "Life-Span Developmental Psychology." *Annual Review of Psychology* 31:65-110.
- Bandura, Albert 1996 *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Borman, Kathryn M. 1991 *The First 'Real' Job: A Study of Young Workers*. Albany: State University of New York Press.
- Brandtstadter, Jochen 1998 "Action Perspectives on Human Development." In Richard M. Lerner, ed., *Handbook of Child Psychology*. Vol. 1, *Theoretical Models of Human Development*. New York: John Wiley & Sons, Inc.
- Buchmann, Marlis 1989 *The Script of Life in Modern Society: Entry into Adulthood in a Changing World*. Chicago: University of Chicago Press.
- Burton, Linda M., Dawn A. Obeidallah, and Kevin Allison 1996 "Ethnographic Insights on Social Context and Adolescent Development among Inner-City African-American Teens." In R. Jessor, A. Colby, and R. Shweder, eds., *Ethnography and Human Development: Context and Meaning in Social Inquiry*. Chicago: University of Chicago Press.
- Carr, R. V., J. D. Wright, and C. J. Brody 1996 "Effects of High School Work Experience a Decade Later: Evidence from the National Longitudinal Survey." *Sociology of Education* 69 (January):66-81.
- Chaplin, D., and J. Hannaway 1996 High school enrollment: meaningful connections for at-risk youth. Paper presented at the Annual Meeting of the American Educational Research Association, New York. Washington, D.C.: Urban Institute.
- Clausen, John A. 1993 *American Lives: Looking Back at the Children of the Great Depression*. New York: Free Press.
- 1991 "Adolescent Competence and the Shaping of the Life Course." *American Journal of Sociology* 96:805-842.
- Cohler, Bertram J., Frances M. Stott, and Judith S. Musick 1995 "Adversity, Vulnerability, and Resilience: Cultural and Developmental Perspectives." In



- Dante Cicchetti and Donald J. Cohen, eds., *Developmental Psychopathology*. Vol. 2, *Risk, Disorder, and Adaptation*. New York: John Wiley & Sons, Inc.
- Coleman, James S. 1994 "Social Capital, Human Capital, and Investment in Youth." In Anne C. Petersen and Jeylan T. Mortimer, eds., *Youth Unemployment and Society*. Cambridge, Eng.: Cambridge University Press.
- , and Torsten Husen 1985 *Becoming Adult in a Changing Society*. Paris: Organization for Economic Co-operation and Development.
- Committee on the Health and Safety Implications of Child Labor, National Research Council 1998 *Protecting Youth at Work: Health, Safety, and Development of Working Children and Adolescents in the United States*. Washington, D.C.: National Academy Press.
- Connor, John W., and George A. De Vos 1989 "Cultural Influences on Achievement Motivation and Orientation toward Work in Japanese and American Youth." In D. Stern and D. Eichorn, eds., *Adolescence and Work: Influences of Social Structure, Labor Markers, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum.
- Cooney, Teresa M. 1994 "Leaving Home and Coming Back: The Influence of Parental Divorce and Family Dynamics on Young Adults' Residential Transitions." *Sociological Studies of Children* 6:159–179.
- , and Jeylan T. Mortimer Forthcoming "Family Structure Differences in the Timing of Leaving Home: Exploring Mediating Factors." *Journal of Research on Adolescence*.
- Côté, James E., and Anton L. Allahaar 1994 *Generation on Hold: Coming of Age in the Late Twentieth Century*. New York: New York University Press.
- Dornbusch, Sanford M. 1994 Off the track. Presidential address at the meeting of the Society for Research on Adolescence, 10–13 February, at San Diego, Calif.
- Elder, Glen H., Jr. 1974 *Children of the Great Depression*. Chicago: University of Chicago Press.
- , and Avshalom Caspi 1990 "Studying Lives in a Changing Society: Sociological and Personological Explorations." In A. I. Rabin, R. A. Zucker, and S. Frank, eds., *Studying Persons and Lives*. New York: Springer-Verlag.
- 1988 "Human Development and Social Change: An Emerging Perspective on the Life Course." In N. Bolger, A. Caspi, G. Downey, and M. Moorehouse, eds., *Persons in Context: Developmental Processes*. New York: Cambridge University Press.
- Elder, Glen H., Jr., J. K. Liker, and C. E. Cross 1984 "Parent-Child Behavior in the Great Depression: Life Course and Intergenerational Influences." In P. B. Baltes, ed., *Life-Span Development and Behavior*. New York: Academic Press.
- Elder, Glen H., Jr., and Angela M. O'Rand 1995 "Adult Lives in a Changing Society." In K. S. Cook, G. A. Fine and J. S. House, eds., *Sociological Perspectives on Social Psychology*. New York: Allyn and Bacon.
- Ensminger, Margaret E., and Hee Soon Juon 1998 "Transition to Adulthood among High-Risk Youth." In Richard Jessor, ed., *New Perspectives on Adolescent Risk Behavior*. New York: Cambridge University Press.
- Featherman, David L. 1980 "Schooling and Occupational Careers: Constancy and Change in Worldly Success." In O. G. Brim, Jr. and J. Kagan, eds., *Constancy and Change in Human Development*. Cambridge, Mass.: Harvard University Press.
- , and Kenneth I. Spenner 1988 "Class and the Socialization of Children: Constancy, Change, or Irrelevance?" In E. Mavis Hetherington, R. M. Lerner, and M. Perlmutter, eds., *Child Development in Perspective*. Hillsdale, N.J.: Lawrence Erlbaum.
- Feldman, Kenneth A., and John Weiler 1976 "Changes in Initial Differences among Major-Field Groups: An Exploration of the 'Accentuation Effect.'" In W. H. Sewell, R. M. Hauser, and D. L. Featherman, eds., *Schooling and Achievement in American Society*. New York: Academic Press.
- Festinger, Leon 1957 *A Theory of Cognitive Dissonance*. Stanford, Calif.: Stanford University Press.
- Finch, Michael D., Michael J. Shanahan, Jeylan T. Mortimer, and Seongryeol Ryu 1991 "Work Experience and Control Orientation in Adolescence." *American Sociological Review* 56 (October):597–611.
- Foner, Anne, and David Kertzer 1978 "Transitions over the Life Course: Lessons from Age-Set Societies." *American Journal of Sociology* 83:1081–1104.
- Furstenberg, Frank F., Jr., J. Brooks-Gunn, and S. Philip Morgan 1987 *Adolescent Mothers in Later Life*. Cambridge, Eng.: Cambridge University Press.
- Furstenberg, Frank F., Jr., and Thomas D. Cook, Jacquelynne Eccles, Glen H. Elder, Jr., and Arnold Sameroff 1999 *Managing to Make It: Urban Families and Adolescent Success*. Chicago: University of Chicago Press.
- Galambos, Nancy L., and Marion F. Ehrenberg 1997 "The Family as Health Risk and Opportunity: A Focus on Divorce and Working Families." In John Schulenberg, Jennifer L. Maggs, and Klaus Hurrelmann, eds., *Health Risks and Developmental Transitions during Adolescence*. Cambridge, Eng.: Cambridge University Press.

- Gamoran, Adam 1996 "Educational Stratification and Individual Careers." In Alan C. Kerckhoff, ed., *Generating Social Stratification: Toward a New Research Agenda*. Boulder, Colo.: Westview Press.
- Gecas, Viktor 1989 "The Social Psychology of Self-Efficacy." *Annual Review of Sociology* 15:291-316.
- 1979 "The Influence of Social Class on Socialization." In W. R. Burr, R. Hill, F. I. Nye, and I. L. Reiss, eds., *Contemporary Theories about the Family*. Vol. 1, *Research-Based Theories*. New York: Free Press.
- Geissler, B., and Helga Kruger 1992 "Balancing the Life Course in Response to Institutional Arrangements." In Walter Heinz, ed., *Institutions and Gatekeeping in the Life Course*. Weinheim, Germany: Deutscher Studien Verlag.
- Glenn, Norval D. 1980 "Values, Attitudes, and Beliefs." In O. G. Brim, Jr. and J. Kagan, eds., *Constancy and Change in Human Development*. Cambridge, Mass.: Harvard University Press.
- Goldscheider, Frances K., and Calvin Goldscheider 1993 *Leaving Home Before Marriage: Ethnicity, Familism, and Generational Relationships*. Madison: University of Wisconsin Press.
- Gore, Susan, Robert Aseltine, Jr., Mary Ellen Colten, and Bin Lin 1997 "Life after High School: Development, Stress, and Well-Being." In Ian H. Gotlib and Blair Wheaton, eds., *Stress and Adversity over the Life Course: Trajectories and Turning Points*. Cambridge, Eng.: Cambridge University Press.
- Grabowski, Lorie Schabo, Kathleen T. Call, and Jeylan T. Mortimer 1998 Global and economic self-efficacy in the attainment process. Paper presented at the annual meeting of the American Sociological Association, August, San Francisco, Calif.
- Graff, Harvey 1995 *Conflicting Paths: Growing Up in America*. Cambridge, Mass.: Harvard University Press.
- Granovetter, Mark S. 1974 *Getting a Job*. Cambridge, Mass.: Harvard University Press.
- Greenberger, Ellen, and A. B. Sorenson 1974 "Toward a Concept of Psychosocial Maturity." *Journal of Youth and Adolescence* 3:329-358.
- Greenberger, Ellen, and Laurence Steinberg 1986 *When Teenagers Work*. New York: Basic Books.
- Greene, A. L. 1990 "Great Expectations: Constructions of the Life Course During Adolescence." *Journal of Youth and Adolescence* 19:289-306.
- Hagestad, Gunhild 1992 "Assigning Rights and Duties: Age, Duration, and Gender in Social Institutions." In Walter Heinz, ed., *Institutions and Gatekeeping in the Life Course*. Weinheim, Germany: Deutscher Studien Verlag.
- Hall, G. S. 1904 *Adolescence: Its Psychology and its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion, and Education*. New York: Appleton.
- Halperin, Samuel (ed.) 1998 *The Forgotten Half Revisited: American Youth and Young Families, 1988-2008*. Washington, D.C.: American Youth Policy Forum, Inc.
- Hamilton, Stephen F. 1990 *Apprenticeship for Adulthood. Preparing Youth for the Future*. New York: Free Press.
- Hogan, Dennis P., and Nan Marie Astone 1986 "The Transition to Adulthood." *Annual Review of Sociology* 12:109-130.
- Hogan, Dennis P., and Takashi Mochizuki 1988 "Demographic Transitions and the Life Course: Lessons from Japanese and American Comparisons." *Journal of Family History* 13:291-305.
- Holland, Dorothy, and Margaret Eisenhart 1990 *Educated in Romance: Women, Achievement, and College Culture*. Chicago: University of Chicago Press.
- Inglehart, R. 1990 *Culture Shift in Advanced Industrial Society*. Princeton, N.J.: Princeton University Press.
- Jessor, Richard, and Shirley L. Jessor 1977 *Problem Behavior and Psychosocial Development. A Longitudinal Study of Youth*. New York: Academic Press.
- Jordaan, Jean Pierre, and Donald E. Super 1974 "The Prediction of Early Adult Vocational Behavior." In D. F. Ricks, A. Thomas, and M. Roff, eds., *Life History Research in Psychopathology*. Minneapolis: University of Minnesota Press.
- Keniston, Kenneth 1970 "Youth as a Stage of Life." *American Scholar* 39:631-654.
- Kerckhoff, Alan C. 1996 "Building Conceptual and Empirical Bridges between Studies of Educational and Labor Force Careers." In A. C. Kerckhoff, ed., *Generating Social Stratification: Toward a New Research Agenda*. Boulder, Colo.: Westview Press.
- 1995 "Social Stratification and Mobility Processes: Interaction between Individuals and Social Structures." In Karen S. Cook, Gary Alan Fine, and James S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- 1993 *Diverging Pathways. Social Structure and Career Deflections*. New York: Cambridge University Press.
- and Lorraine Bell 1998 "Hidden Capital: Vocational Credentials and Attainment in the United States." *Sociology of Education* 71:152-174.
- Klein, Hugh 1990 "Adolescence, Youth and Young Adulthood: Rethinking Current Conceptualizations of Life Stage." *Youth and Society* 21:446-471.

- Kohn, Melvin L., and Carmi Schooler 1983 *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, N.J.: Ablex Publishing.
- Kohn, Melvin L., Kazimierz M. Slomczynski, and Carrie Schoenbach 1986 "Social Stratification and the Transmission of Values in the Family: A Cross-National Assessment." *Sociological Forum* 1:73-102.
- Lerner, Richard M., and N. A. Busch-Rossnagel (eds.) 1981 *Individuals as Producers of Their Development: A Life-Span Perspective*. New York: Academic Press.
- Lichter, Daniel T. 1988 "Racial Differences in Underemployment in American Cities" *American Journal of Sociology* 93:771-792.
- Lin, Nan 1992 "Social Resources Theory." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*. New York: Macmillan.
- Lorence, Jon, and Jeylan T. Mortimer 1985 "Job Involvement through the Life Course: A Panel Study of Three Age Groups." *American Sociological Review* 50:618-638.
- McLanahan, Sara, and Gary Sandefur 1994 *Growing up with a Single Parent*. Cambridge, Mass.: Harvard University Press.
- McLaughlin, Steven D., Barbara D. Melber, John O. G. Billy, Denise M. Zimmerle, Linda D. Wings, and Terry R. Johnson 1988 *The Changing Lives of American Women*. Chapel Hill: University of North Carolina Press.
- Machung, Anne 1989 "Talking Career, Thinking Job: Gender Differences in Career and Family Expectations of Berkeley Seniors." *Feminist Studies* 15:35-39.
- Maggs, Jennifer L. 1997 "Alcohol Use and Binge Drinking as Goal-Directed Action during the Transition to Postsecondary Education." In John Schulenberg, Jennifer L. Maggs, and Klaus Hurrelmann, eds., *Health Risks and Developmental Transitions during Adolescence*. Cambridge, Eng.: Cambridge University Press.
- Mainquist, Sheri, and Dorothy Eichorn 1989 "Competence in Work Settings." In D. Stern and D. Eichorn, eds., *Adolescence and Work: Influences of Social Structure, Labor Markets, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum.
- Mannheim, Karl 1952 "The Problem of Generations." In P. Kecskemeti, ed., *Essays in the Sociology of Knowledge*. London: Routledge and Kegan Paul.
- Marini, Margaret Mooney 1984 "Age and Sequencing Norms in the Transition to Adulthood." *Social Forces* 63:229-244.
- Marsh, H. W. 1991 "Employment during High School: Character Building or Subversion of Academic Goals?" *Sociology of Education* 64:172-189.
- Meijers, Frans 1992 "'Being Young' in the Life Perceptions of Dutch, Moroccan, Turkish and Surinam Youngsters." In Wim Meeus, Martijn de Goede, Willem Kox, and Klaus Hurrelmann, eds., *Adolescence, Careers, and Cultures*. Berlin: Walter de Gruyter.
- Menaghan, Elizabeth G. 1997 "Intergenerational Consequences of Social Stressors: Effects of Occupational and Family Conditions on Young Mothers and Their Children." In Ian Gotlib and Blair Wheaton, eds., *Stress and Adversity over the Life Course: Trajectories and Turning Points*. Cambridge, Eng.: Cambridge University Press.
- Modell, John 1989 *Into One's Own. From Youth to Adulthood in the United States, 1920-1975*. Berkeley: University of California Press.
- 1980 "Normative Aspects of American Marriage Timing since World War II." *Journal of Family History* 5:210-234.
- , Frank Furstenberg, and Theodore Hershberg 1976 "Social Change and Transitions to Adulthood in Historical Perspective." *Journal of Family History* 1:7-32.
- Moen, Phyllis 1992 *Women's Two Roles: A Contemporary Dilemma*. Westport, Conn.: Auburn House.
- Morris, Martina, and Annette Bernhardt 1998 The transition to work in the post-industrial labor market. Paper presented at the annual meeting of the American Sociological Association, August, San Francisco, Calif.
- Mortimer, Jeylan T. 1996 "Social Psychological Aspects of Achievement." In Alan C. Kerckhoff, ed., *Generational Social Stratification: Toward a New Generation of Research*. Boulder, Colo.: Westview.
- 1994 "Individual Differences as Precursors of Youth Unemployment." In Anne C. Petersen and Jeylan T. Mortimer, eds., *Youth Unemployment and Society*. New York: Cambridge University Press.
- , Carolyn Harley, and Pamela J. Aronson 1999 "How do Prior Experiences in the Workplace Set the Stage for Transitions to Adulthood?" In Alan Booth and Nan Crouter, eds., *Transitions to Adulthood in a Changing Economy: No Work, No Family, No Future?* New York: Greenwood Press.
- Mortimer, Jeylan T., and Michael D. Finch 1986 "The Effects of Part-time Work on Self-Concept and Achievement." In K. Borman and J. Reisman, eds., *Becoming a Worker*. Norwood, N.J.: Ablex Publishing.
- 1996 "Work, Family, and Adolescent Development." In J. T. Mortimer and M. D. Finch, eds., *Adolescents, Work and Family: An Intergenerational Developmental Analysis*. Thousand Oaks, Calif.: Sage Publications.

- , and Donald S. Kumka 1982 "Persistence and Change in Development: The Multidimensional Self-Concept." In P. D. Baltes and O. G. Brim, Jr., eds., *Life-Span Development and Behavior*. New York: Academic Press.
- Mortimer, Jeylan T., Michael D. Finch, and Geoffrey Maruyama 1988 "Work Experience and Job Satisfaction: Variation by Age and Gender." In J. T. Mortimer and K. M. Borman, eds., *Work Experience and Psychological Development through the Life Span*. Boulder, Colo.: Westview Press.
- Mortimer, Jeylan T. and Monica K. Johnson 1999. "Adolescent Part-Time Work and Post-Secondary Transition Pathways in the United States." Walter Heinz, ed., *From Education to Work: Cross National Perspectives*. New York: Cambridge University Press.
- 1998 "New Perspectives on Adolescent Work and the Transition to Adulthood." In Richard Jessor, ed., *New Perspectives on Adolescent Risk Behavior*. New York: Cambridge University Press.
- Mortimer, Jeylan T., and Helga Kruger (Forthcoming) "Transition from School to Work in the United States and Germany: Formal Pathways Matter." In M. Hallinan, ed., *Handbook of the Sociology of Education*. New York: Plenum Publishing Company.
- Mortimer, Jeylan T., and Donald S. Kumka 1982 "A Further Examination of the 'Occupational Linkage Hypothesis.'" *Sociological Quarterly* 23:3-16.
- Mortimer, Jeylan T., Jon Lorence, and Donald S. Kumka 1986 *Work, Family, and Personality: Transition to Adulthood*. Norwood, N.J.: Ablex Publishing.
- Mortimer, Jeylan T., Ellen Efron Pimentel, Seongryeol Ryu, Katherine Nash, and Chaimun Lee 1996 "Part-Time Work and Occupational Value Formation in Adolescence." *Social Forces* 74:1405-1418.
- Moss, Howard A., and Elizabeth J. Susman 1980 "Longitudinal Study of Personality Development." In O. G. Brim, Jr., and J. Kagan, eds., *Constancy and Change in Human Development*. Cambridge, Mass.: Harvard University Press.
- Neugarten, Bernice L., and Nancy Datan 1973 "Sociological Perspectives on the Life Cycle." In P. B. Baltes and K. Warner Schaie, eds., *Life Span Developmental Psychology: Personality and Socialization*. New York: Academic Press.
- Neugarten, Bernice L., Joan W. Moore, and John C. Lowe 1965 "Age Norms, Age Constraints, and Adult Socialization." *American Journal of Sociology* 70:710-717.
- Newman, Katherine 1996 "Working Poor: Low Wage Employment in the Lives of Harlem Youth." In J. Graber, J. Brooks-Gunn, and A. Petersen, eds., *Transitions through Adolescence: Interpersonal Domains and Context*. Hillsdale, N.J.: Erlbaum Associates.
- Ogbu, John U. 1989 "Cultural Boundaries and Minority Youth Orientation toward Work Preparation." In D. Stern and D. Eichorn, eds., *Adolescence and Work: Influences of Social Structure, Labor Markets, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum.
- Osterman, Paul 1989 "The Job Market for Adolescents." In D. Stern and D. Eichorn, eds., *Adolescence and Work: Influences of Social Structure, Labor Markets, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum.
- 1980 *Getting Started: The Youth Labor Market*. Cambridge, Mass.: MIT Press.
- Pallas, Aaron 1993 "Schooling in the Course of Human Lives: Social Context of Education and Transition to Adulthood in Industrial Society." *Review of Educational Research* 63, 4, Winter: 409-447.
- Panel on Youth of the President's Science Advisory Committee 1974 *Youth: Transition to Adulthood*. Chicago: University of Chicago Press.
- Phillips, Sarah, and Kent Sandstrom 1990 "Parental Attitudes toward 'Youthwork'" *Youth and Society* 22:160-183.
- Rindfuss, Ronald R., Elizabeth C. Cooksey, and R. L. Sutterlin 1990 Young adult occupational achievement: early expectations versus behavioral reality. Paper presented at the World Congress of Sociology, July, Madrid.
- Rindfuss, Ronald R., C. Gray Swicegood, and Rachel Rosenfeld 1987 "Disorder in the Life Course: How Common and Does it Matter?" *American Sociological Review* 52:785-801.
- Rosenberg, Morris 1979 *Conceiving the Self*. New York: Basic Books.
- Rossi, Alice S. 1980 "Life-Span Theories and Women's Lives." *Signs: Journal of Women in Culture and Society* 6:4-32.
- Ruhm, C. J. 1997 "Is High School Employment Consumption or Investment?" *Journal of Labor Economics* 15:735-776.
- 1995 "The Extent and Consequences of High School Employment." *Journal of Labor Research* 16:293-303.
- Ruscoe, G., J. C. Morgan, and C. Peebles 1996 "Students Who Work." *Adolescence* 31:625-632.
- Rutter, Michael, and David Quinton 1984 "Long-Term Follow-Up of Women Institutionalized in Childhood: Factors Promoting Good Functioning in Adult Life." *British Journal of Developmental Psychology* 2:191-204.

- Ryu, Seongryeol, and Jeylan T. Mortimer 1996 "The 'Occupational Linkage Hypothesis' applied to Occupational Value Formation in Adolescence." In Jeylan T. Mortimer and Michael D. Finch, eds., *Adolescents, Work, and Family: An Intergenerational Developmental Analysis*. Thousand Oaks, Calif.: Sage.
- Sampson, Robert J., and John H. Laub 1993 *Crime in the Making: Pathways and Turning Points through Life*. Cambridge: Harvard University Press.
- Sears, David O. 1981 "Life-Stage Effects on Attitude Change, Especially among the Elderly." In B. Kiesler, J. N. Morgan, and V. Kincaid Oppenheimer, eds., *Aging: Social Change*. New York: Academic Press.
- Seligman, Martin 1988 "Boomer Blues." *Psychology Today* (October): 50–55.
- Sewell, William H., and Robert M. Hauser 1976 "Causes and Consequences of Higher Education: Models of the Status Attainment Process." In W. H. Swell, R. M. Hauser, and D. L. Featherman, eds., *Schooling and Achievement in American Society*. New York: Academic Press.
- Shanahan, Michael J. (Forthcoming) "Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective." *Annual Review of Sociology*.
- This volume "Adolescence." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*. New York: Macmillan.
- , and Glen H. Elder, Jr. 1999 History, agency and the life course. Paper presented at the Lincoln, Nebraska Symposium on Motivation.
- Shanahan, Michael J., Richard A. Miech, and Glen H. Elder, Jr. 1998 "Changing Pathways to Attainment in Men's Lives: Historical Patterns of School, Work, and Social Class." *Social Forces* 77:231–256.
- Shanahan, Michael J., Michael D. Finch, Jeylan T. Mortimer, and Seongryeol Ryu 1991 "Adolescent Work Experience and Depressive Affect." *Social Psychology Quarterly* 54:299–317.
- Shavit, Yossi, and Walter Muller 1998 *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford: Clarendon Press.
- Simmons, Roberta, and Dale Blyth 1987 *Moving into Adolescence: The Impact of Pubertal Change and School Context*. New York: Aldine.
- Spain, Daphne, and Suzanne M. Bianchi 1996 *Balancing Act: Motherhood, Marriage, and Employment among American Women*. New York: Russell Sage Foundation.
- Stern, David, and Yoshi-Fumi Nakata 1989 "Characteristics of High School Students' Paid Jobs, and Employment Experience after Graduation." In D. Stern and D. Eichorn, eds., *Adolescence and Work. Influences of Social Structure, Labor Markets, and Culture*. Hillsdale, N.J.: Lawrence Erlbaum.
- Strauss, Anselm 1959 *Mirrors and Masks: The Search for Identity*. Glencoe, Ill.: The Free Press.
- Sullivan, Mercer L. 1989 *Getting Paid: Youth Crime and Work in the Inner City*. Ithaca, NY: Cornell University Press.
- Vinken, Henk, and Peter Ester 1992 "Modernization and Value Shifts: A Cross-Cultural and Longitudinal Analysis of Adolescents' Basic Values." In Wim Meeus, Martijn de Goede, Willem Kox, and Klaus Hurrelmann, eds., *Adolescence, Careers, and Cultures*. Berlin: Walter de Gruyter.
- Wells, L. Edward, and Sheldon Stryker 1988 "Stability and Change in Self over the Life Course." In Paul Baltes, David Featherman, and Richard Lerner, eds., *Life Span Development and Behavior*. Hillsdale, N.J.: Lawrence Erlbaum.
- White, Merry 1993 *The Material Child: Coming of Age in Japan and America*. New York: Free Press.
- Wilson, William Julius 1987 *The Truly Disadvantaged. The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- World Factbook, The 1998 Compiled by the Central Intelligence Agency. Washington, D.C.: U.S. Government Printing Office.

JEYLAN T. MORTIMER  
PAMELA ARONSON

## AFFECT CONTROL THEORY AND IMPRESSION FORMATION

Sociologist Erving Goffman (1969) argued that people conduct themselves so as to generate impressions that maintain the identities, or "faces," that they have in social situations. Human action—aside from accomplishing tasks—functions expressively in reflecting actors' social positions and in preserving social order. Affect control theory (Heise 1979; Smith-Lovin and Heise 1988; MacKinnon 1994) continues Goffman's thesis, providing a mathematized and empirically grounded model for explaining and predicting expressive aspects of action.

### AFFECTIVE MEANING

Cross-cultural research among people speaking diverse languages in more than twenty-five nations around the world (Osgood, May, and Miron 1975) revealed that any person, behavior, object, setting, or property of persons evokes an affective response consisting of three components. One component consists of approval or disapproval of the entity—an evaluation based on morality (good versus bad), aesthetics (beautiful versus ugly), functionality (useful versus useless), hedonism (pleasant versus unpleasant), or some other criterion. Whatever the primary basis of evaluation, it tends to generalize to other bases, so, for example, something that is useful tends also to seem good, beautiful, and pleasant.

Another component of affective responses is a potency assessment made in terms of physical proportions (large versus small, deep versus shallow), strength (strong versus weak), influence (powerful versus powerless), or other criteria. Again, judgments on the basis of one criterion tend to generalize to other criteria, so, for example, a powerful person seems large, deep (in a metaphorical sense), and strong.

The third component of affective responses—an appraisal of activity—may depend on speed (fast versus slow), perceptual stimulation (noisy versus quiet, bright versus dim), age (young versus old), keenness (sharp versus dull), or other criteria. These criteria also generalize to some degree so, for example, a young person often seems metaphorically fast, noisy, bright, and sharp.

The evaluation, potency, and activity (EPA) structure in subjective responses is one of the best-documented facts in social science, and an elaborate technology has developed for measuring EPA responses on “semantic differential scales” (Heise 1969). The scales consist of adjectives separated by a number of check positions. For example, a standard scale has Good–Nice at one end and Bad–Awful at the other end, and intervening positions on the scale allow respondents to record the direction and intensity of their evaluations of a stimulus. The middle rating position on such scales represents neutrality and is coded 0.0. Positions moving outward are labeled “slightly,” “quite,” “extremely,” and “infinitely,” and they are coded 1.0, 2.0, 3.0, and 4.3 respectively—positive on the good, potent, and active sides of the scales; and

negative on the bad, impotent, and inactive sides (Heise 1978).

EPA responses tend to be socially shared within a population (Heise 1966, 1999), so a group’s average EPA response to an entity indexes the group sentiment about the entity. Group sentiments can be computed from as few as thirty ratings since each rater’s transient response originates from a single shared sentiment rather than a separately held position (Romney, Weller, and Batchelder 1986). Sentiments vary across cultures. For example, potent authorities such as an employer are evaluated positively in U.S. and Canadian college populations, but German students evaluate authorities negatively; small children are evaluated positively in Western nations but neutrally in Japan.

### IMPRESSION FORMATION

Combinations of cognitive elements bring affective meanings together and create outcome impressions through psychological processes that are complex, subtle, and yet highly predictable (Gollob 1974; Heise 1979; Anderson 1996).

One kind of impression formation amalgamates a personal attribute with a social identity resulting in a sense of how a person is different from similar others (Averett and Heise 1987). For example, among U.S. college students (the population of raters for examples henceforth), someone who is rich is evaluatively neutral, very powerful, and a little on the quiet side. Meanwhile, a professor is fairly good, fairly powerful, and a bit quiet. The notion of a “rich professor” combines these sentiments and yields a different outcome. A rich professor is evaluated somewhat negatively, mainly because the personalized power of wealth generates an uneasiness that is not overcome by esteem for academic status. A rich professor seems very powerful because the average potency of wealth and of professors is high, and the mind adds an extra increment of potency because of the personalized power deriving from wealth. A rich professor seems even quieter than the component statuses because activity connotations do not merely average, they summate to some degree.

The processes that are involved in combining a social identity with a status characteristic like “rich” also are involved in combining a social

identity with personal traits. Thus, an authoritarian professor evokes an impression roughly similar to a rich professor since the affective association for authoritarian is similar to that for rich.

Another example involving emotion illustrates additional processes involved in combining personal characteristics with social identities (Heise and Thomas 1989). Being outraged implies that one is feeling quite bad, somewhat potent, and somewhat lively. A child is felt to be quite good, quite impotent, and very active. The combination "outraged child," seems fairly bad, partly because the child is flaunting personalized power deriving from an emotion and partly because the mind discounts customary esteem for a person if there is a personal basis for evaluating the person negatively: The child's negative emotion undercuts the regard one usually has for a child. The child's impotency is reduced because of a bad and potent emotional state. And the child's activity is greater than usual because the activity of the emotion and the activity of the identity combine additively.

The combining of attributions and identities involves somewhat different processes in different domains. For example, North American males and females process attributions in the same way, though both make attributions about females that are more governed by morality considerations than attributions about males (Heise 1999). Meanwhile, Japanese males are socialized to be more concerned than Japanese females with evaluative consistency in attributions and with matching goodness and weakness, so Japanese males process attributions about everything with more concern for morality than do Japanese females (Smith, Matsuno, and Ike 1999).

Events are another basis for impression formation. A social event—an actor behaving on an object or person within some setting—amalgamates EPA impressions of the elements comprising the event and generates a new impression of each element. For example, when an athlete strangles a coach, the athlete seems bad, and the coach seems cowardly because the coach loses both goodness and potency as a result of the event. The complex equations for predicting outcome impressions from input impressions have been found to be similar in different cultures (Heise 1979; Smith-Lovin 1987a; Smith-Lovin 1987b; Smith, Matsuno, and Umino 1994).

To a degree, the character of a behavior diffuses to the actor who performs the behavior. For example, an admired person who engages in a violent act seems less good, more potent, and more active than usual. Impressions of the actor also are influenced by complex interplays between the nature of the behavior and the nature of the object. For example, violence toward an enemy does not stigmatize an actor nearly so much as violence toward a child. That is because bad, forceful behaviors toward bad, potent objects seem justified while such behaviors toward good, weak objects seem ruthless. Moreover, the degree of justification or of ruthlessness depends on how good the actor was in the first place; for example, a person who acts violently toward a child loses more respect if initially esteemed than if already stigmatized.

Similarly, diffusions of feeling from one event element to another and complex interplays between event elements generate impressions of behaviors, objects, and settings. The general principle is that initial affective meanings of event elements combine and thereby produce new impressions that reflect the meaning of the event. Those impressions are transient because they, in turn, are the meanings that are transformed by later events.

### IMPRESSION MANAGEMENT

Normal events produce transient impressions that match sentiments, whereas events that generate impressions deviating widely from sentiments seem abnormal (Heise and MacKinnon 1987). For example, "a parent assisting a child" creates impressions of parent, child, and assisting that are close to sentiments provided by our culture, and the event seems normal. On the other hand, "a parent harming a child" seems abnormal because the event produces negative impressions of parent and child that are far different than the culturally-given notions that parents and children are good.

According to affect control theory, people manage events so as to match transient impressions with sentiments and thereby maintain normality in their experience. Expressive shaping of events occurs within orderly rational action, and ordinarily the expressive and the rational components of action complement each other because

cultural sentiments incite the very events that are required by the logic of social institutions like the family, law, religion, etc.

Having adopted an appropriate identity at a scene and having cast others in complementary identities, a person intuitively behaves that will create normal impressions. For example, if a person in the role of judge is to act on someone who is a proven crook, then she must do something that confirms a judge's power and that confirms the badness of a crook, and behaviors like "convict" and "sentence" produce the right impressions. Fitting behaviors may change in the wake of prior events. For example, a father who is fulfilling his role in a mediocre manner because his child has disobeyed him strives to regain goodness and power by controlling the child or by dramatizing forgiveness. Other people's identities may serve as resources for restoring a compromised identity (Wiggins and Heise 1987); for example, a father shaken by a child's disobedience might recover his poise by supporting and defending mother.

Behaviors that confirm sentiments are the intrinsically motivated behaviors in a situation. Actors sometimes comply to the demands of others and thereby forego intrinsically motivated behaviors. Yet compliance also reflects the basic principle since compliant behavior is normal in one relationship even though it may be abnormal in another relationship. For example, a child acts normally when calling on a playmate though also abnormally if his mother has ordered him not to do so and he disobeys her. The actor maintains the relationship that is most salient.

Sometimes other people produce events that do not confirm sentiments evoked by one's own definition of a situation. Affect control theory suggests several routes for restoring consistency between impressions and sentiments in such cases. First, people may try to reinterpret other's actions so as to optimize expressive coherence. For example, an actor's movement away from another person can be viewed as departing, leaving, escaping, fleeing, deserting—and one chooses the interpretation that seems most normal, given participants' identities and prior events (Heise 1979). Of course, interpretations of a behavior are bound by determinable facts about the behavior and its consequences, so some behaviors cannot be interpreted in a way that completely normalizes an event.

Another response to disturbing events is construction of new events that transform abnormal impressions back to normality. Restorative events with the self as actor might be feasible and enacted, as in the example of a father controlling a disobedient child. Restorative events that require others to act might be elicited by suggesting what the other should do. For example, after a child has disobeyed his mother, a father might tell the child to apologize.

Intractable disturbances in interaction that cannot be handled by reinterpreting others' actions or by instigating new events lead to changes in how people are viewed, such as attributing character traits to people in order to form complex identities that account for participation in certain kinds of anomalous events. For example, a father who has ignored his child might be viewed as an inconsiderate person.

Changing base identities also can produce the kind of person who would participate in certain events. For example, an employee cheating an employer would be expressively coherent were the employee known to be a lawbreaker, and a cheating incident may instigate legal actions that apply a lawbreaker label and that withdraw the employee identity. The criminal justice system changes actors in deviant events into the kinds of people who routinely engage in deviant actions, thereby allowing the rest of us to feel that we understand why bad things happen.

Trait attributions and labels that normalize particular incidents are added to conceptions of people, and thereafter the special identities may be invoked in order to set expectations for a person's behavior in other scenes or to understand other incidents. Everyone who interacts with a person builds up knowledge about the person's capacities in this way, and a person builds up knowledge about the self in this way as well.

## EMOTION

Affect control theory is a central framework in the sociology of emotions (Thoits 1989; Smith-Lovin 1994), and its predictions about emotions in various situations match well with the predictions of real people who imagine themselves in those situations (Heise and Weir 1999). According to affect control theory, spontaneous emotion reflects the



state a person has reached as a result of events and also how that state compares to the ideal experience of a person with a particular social identity. For example, if events make a person seem neutral on goodness, potency, and activity, then the tendency is to feel emotionally neutral, but someone in the sweetheart role ends up feeling blue because he or she is experiencing so much less than one expects in a romantic relationship.

Because emotions reflect the impressions that events have generated, they are a way of directly sensing the consequences of social interaction. Because emotions simultaneously reflect what kinds of identities people are taking, emotions also are a way of sensing the operative social structure in a situation. Moreover, because displays of emotion broadcast a person's subjective appraisals to others, emotions contribute to intersubjective sharing of views about social matters.

People sometimes mask their emotions or display emotions other than those that they feel spontaneously in order to hide their appraisal of events from others or to conceal personal definitions of a situation. For example, an actor caught in misconduct might display guilt and remorse beyond what is felt in order to convince others that he believes his behavior is wrong and that he is not the type who engages in such activity. Such a display of negative emotion after a deviant act makes the actor less vulnerable to a deviant label—an hypothesis derived from the mathematics of affect control theory (Heise 1989).

### APPLICATIONS

Affect control theory provides a comprehensive social-psychological framework relating to roles, impression formation, behavior, emotion, attribution, labeling, and other issues (Stryker and Statham 1985). Consequently it is applicable to a variety of social-psychological problems. For example:

- Smith-Lovin and Douglass (1992) showed that sentiments about relevant identities and behaviors are positive in a deviant subculture, and therefore subcultural interactions are happier than outsiders believe.
- MacKinnon and Langford (1994) found that moral evaluations determine the prestige of occupations with low and middle

but not high levels of education and income; and they found that income affects occupational prestige partly by adjusting feelings about the potency of the occupation.

- Robinson and Smith-Lovin (1992) found that people with low self-esteem prefer to associate with their critics rather than their flatterers. Robinson (1996) showed that networks can emerge from self-identities, with cliques reflecting differing levels of self-esteem, and dominance structures reflecting differing levels of self-potency.
- Francis (1997a, 1997b) showed that therapists often promote emotional healing by embedding clients in a social structure where key identities have particular EPA profiles; the identities are associated with different functions in different therapeutic ideologies.
- Studies of courtroom scenarios (Robinson, Smith-Lovin, and Tsoudis 1994; Tsoudis and Smith-Lovin 1998) showed that people (such as jurors) deal more leniently with convicted criminals who show remorse and guilt over their crimes, as predicted by affect control theory. In a related study, Scher and Heise (1993) suggested that perceptions of injustice follow justice-related emotions of anger or guilt, so social interactional structures that keep people happy can prevent mobilization regarding unjust reward structures.
- Heise (1998) suggested that solidarity comes easier when a group identity is good, potent, and lively so that group members engage in helpful actions with each other and experience emotions in parallel. Britt and Heise (forthcoming) showed that successful social movements instigate a sequence of member emotions, culminating in pride, which reflect a good, potent, and lively group identity.

Affect control theory's mathematical model is implemented in a computer program that simulates social interactions and predicts the emotions and interpretations of interactants during expected or unexpected interpersonal events. Simulations can be conducted with EPA measurements

obtained in a variety of nations—Northern Ireland, Canada, Germany, Japan, and the United States. The computer program, the datasets, and other materials are available at the affect control theory web site at: [www.indiana.edu/~socpsy/ACT](http://www.indiana.edu/~socpsy/ACT).

### CONCLUSION

Goffman (1967, p. 9) called attention to the expressive order in social relations:

*By entering a situation in which he is given a face to maintain, a person takes on the responsibility of standing guard over the flow of events as they pass before him. He must ensure that a particular expressive order is sustained—an order that regulates the flow of events, large or small, so that anything that appears to be expressed by them will be consistent with his face.*

Affect control theory's empirically based mathematical model offers a rich and productive foundation for studying the expressive order.

### REFERENCES

- Anderson, Norman H. 1996 *A Functional Theory of Cognition*. New York: Lawrence Erlbaum.
- Averett, C. P., and D. R. Heise 1987 "Modified Social Identities: Amalgamations, Attributions, and Emotions." *Journal of Mathematical Sociology* 13:103–132.
- Britt, Lory, and D. Heise (Forthcoming) "From Shame to Pride in Identity Politics." Forthcoming. In Sheldon Stryker, Tim Owens, and Robert White, eds., *Self Identity, and Social Movements*.
- Francis, Linda 1997a "Ideology and interpersonal Emotion Management: Redefining identity in Two Support Groups." *Social Psychology Quarterly* 60:153–171.
- 1997b "Emotion, Coping, and Therapeutic Ideologies." *Social Perspectives on Emotion* 4:71–101.
- Goffman, Erving 1967 *Interaction Ritual: Essays on Face-to-Face Behavior*. Garden City, N.Y.: Anchor Books.
- Gollob, Harry F. 1974 "A Subject-Verb-Object approach to social cognition." *Psychological Review* 81:286–321.
- Heise, David R. 1966 "Social Status, Attitudes, and Word Connotations." *Sociological Inquiry* 36:227–239.
- 1969 "Some Methodological Issues in Semantic Differential Research." *Psychological Bulletin* 72:406–422.
- 1979 *Understanding Events: Affect and the Construction of Social Action*. New York: Cambridge University Press.
- 1989 "Effects of Emotion Displays on Social Identification." *Social Psychology Quarterly* 52:10–21.
- 1998 "Conditions for Empathic Solidarity." In Patrick Doreian and Thomas Fararo, eds., *The Problem of Solidarity: Theories and Models*, 197–211 Amsterdam: Gordon and Breach.
- 1999 "Controlling Affective Experience Interpersonally." *Social Psychology Quarterly*. Forthcoming.
- , and Neil MacKinnon 1987 "Affective Bases of Likelihood Judgments." *Journal of Mathematical Sociology* 13:133–151.
- , and Lisa Thomas 1989 "Predicting Impressions Created by Combinations of Emotion and Social Identity." *Social Psychology Quarterly* 52:141–148.
- , and Brian Weir 1999 "A Test of Symbolic Interactionist Predictions About Emotions in Imagined Situations." *Symbolic Interaction*. Forthcoming.
- MacKinnon, Neil J. 1994 *Symbolic Interactionism as Affect Control*. Albany, N.Y.: SUNY Press.
- , and Tom Langford 1994 "The Meaning of Occupational Prestige Scores: A Social Psychological Analysis and Interpretation." *Sociological Quarterly* 35:215–245.
- Osgood, C. H., W. H. May, and M. S. Miron 1975 *Cross-Cultural Universals of Affective Meaning*. Urbana: University of Illinois Press.
- Robinson, Dawn T. 1996 "Identity and friendship: Affective Dynamics and Network Formation." *Advances in Group Process* 13:91–111.
- , and Lynn Smith-Lovin 1992 "Selective Interaction as a Strategy for Identity Maintenance: An Affect Control Model." *Social Psychology Quarterly* 55:12–28.
- , Lynn Smith-Lovin, and Olga Tsoudis 1994 "Heinous Crime or Unfortunate Accident? The Effects of Remorse on Responses to Mock Criminal Confessions." *Social Forces* 73:175–190.
- Romney, A. Kimball, Susan C. Weller, and William H. Batchelder 1986 "Culture as Consensus: A Theory of Culture and Informant Accuracy." *American Anthropologist* 88:313–338.
- Scher, Steven J., and D. R. Heise 1993 "Affect and the Perception of Injustice." *Advances in Group Process* 10:223–252.
- Smith, Herman W., Takanori Matsuno, and Michio Umino 1994 "How Similar are Impression-Formation Processes among Japanese and Americans?" *Social Psychology Quarterly* 57:124–139.
- Smith, Herman W., Takanori Matsuno, and Shuichirou Ike 1999 "The Social Construction of Japanese and American Attributional Expectations." St. Louis University of Missouri Press.

- Smith-Lovin, Lynn 1987a "Impressions from Events." *Journal of Mathematical Sociology* 13:35-70.
- 1987b "The Affective Control of Events within Settings." *Journal of Mathematical Sociology* 13:71-101.
- 1994 "The Sociology of Affect and Emotion." In K. Cook, G. Fine, and J. House, eds., *Sociological Perspectives on Social Psychology*. New York: Allyn and Bacon.
- , and D. R. Heise 1988 *Analyzing Social Interaction: Advances in Affect Control Theory*. New York: Gordon and Breach.
- , and William Douglass 1992 "An Affect-Control Analysis of Two Religious Groups." In V. Gecas and D. Franks, eds., *Social Perspectives on Emotion*, vol. 1. Greenwich, Conn.: JAI Press.
- Stryker, S., and A. Statham 1985 "Symbolic Interaction and Role Theory." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, vol.1, 3rd ed. New York: Random House.
- Thoits, P. A. 1989 "The Sociology of Emotions." *Annual Review of Sociology* 15:317-342.
- Tsoudis, Olga, and Lynn Smith-Lovin 1998 "How Bad Was It? The Effects of Victim and Perpetrator Emotion on Responses to Criminal Court Vignettes." *Social Forces* 77:695-722.
- Wiggins, Beverly, and D. R. Heise 1987 "Expectations, Intentions, and Behavior: Some tests of Affect Control Theory." *Journal of Mathematical Sociology* 13:153-169.

DAVID R. HEISE

## AFFIRMATIVE ACTION

The term *affirmative action* has been used in the United States since the late 1960s to refer to policies that go beyond the simple prohibition of discrimination on grounds of race, national origin, and sex in employment practices and educational programs. These policies require some further action, "affirmative action," to make jobs and promotions and admissions to educational programs available to individuals from groups that have historically suffered from discrimination in gaining these opportunities or are, whether discriminated against or not by formal policies and informal practices, infrequently found in certain occupations or educational institutions and programs.

Affirmative action policies may be policies of governments or governmental units, affecting their own procedures in employment or in granting contracts; or they may be policies of governments, affecting the employment procedures of companies or nonprofit agencies and organizations over whom the governments have power or with whom they deal; or they may be the policies of profit and nonprofit employers, adopted voluntarily or under varying degrees of public or private pressure. Affirmative action policies may include the policies of philanthropic foundations, when they affect the employment policies of their grantees, or educational accrediting agencies, when they affect the employment or admissions policies of the institutions they accredit.

The range of policies that can be called affirmative action is wide, but the term also has a specific legal meaning. It was first used in a legal context in the United States in an executive order of President John F. Kennedy. Subsequent presidential executive orders and other administrative requirements have expanded its scope and meaning, and since 1971 affirmative action so defined has set employment practice standards for contractors of the United States, that is, companies, colleges, universities, hospitals, or other institutions that have business with the U.S. government. These standards are enforced by an office of the Department of Labor, the Office of Federal Contract Compliance Programs. Because of the wide sweep of the executive order and its reach into the employment practices of almost every large employer, affirmative action policies have become extremely controversial.

Affirmative action, under other names, is also to be found in other countries to help groups, whether majority or minority, that have not fared as well as others in gaining employment in higher status occupations or admissions to advanced educational programs.

Affirmative action has been controversial because it appears to contradict a central objective of traditional liberalism and the U.S. civil rights movement, that is, the treatment of individuals on the basis of their individual talents and not on the basis of their color, race, national origin, or sex. Affirmative action, as it has developed, requires surveys by employers of the race, national origin,

and sex of their employees to uncover patterns of “underutilization” and to develop programs to overcome this underutilization and thus to take account of the race, national origin, and sex of applicants for employment and of candidates for promotion. To many advocates of expanded civil rights, this is seen as only the next and a most necessary step in achieving equality for groups that have in the past faced discrimination. To others, who may also deem themselves advocates of civil rights and of the interests of minority groups, affirmative action, in the form in which it has developed, is seen as a violation of the first requirement for a society that promises equal opportunity, that is, to treat individuals as individuals independent of race, national origin, or sex.

This apparent contradiction between civil rights and affirmative action may be glimpsed in the very language of the Civil Rights Act of 1964, the central piece of legislation that banned discrimination in government programs, public facilities, and employment. In the debate over that act, fears were expressed that the prohibition of discrimination in employment, as codified in Title VII, would be implemented by requiring certain numbers of employees to be of a given race. This fear was dealt with by placing language in the act that was understood at the time specifically to forbid the practices that are required under affirmative action since the late 1960s and early 1970s. Title 703 (j) reads:

*Nothing contained in this title shall be interpreted to require any employer . . . to grant preferential treatment to any individual or to any group because of the race, color, religion, sex, or national origin of such individual or group on account of an imbalance which may exist with respect to the total number or percentage of persons of any race, color, religion, sex, or national origin employed by any employer.*

However, federal executive orders governing how the federal government does its business may set their own standards, independent of statutory law. The first executive order using the term *affirmative action* was issued by President John F. Kennedy in 1961. It created a President’s Committee on Equal Employment Opportunity to monitor the obligations contractors undertook to extend affirmative action. At this time, the general understanding of affirmative action was that it

required such things as giving public notice that the employer did not discriminate, making the availability of positions and promotions widely known, advertising in minority media, and the like. With the Civil Rights Act of 1964—which not only prohibited discrimination on grounds of race, color, and national origin but also on grounds of sex—a new executive order, no. 11,246, was formulated by President Lyndon B. Johnson and came into effect. It replaced the President’s Committee on Equal Employment Opportunity with an Office of Federal Contract Compliance Programs (which still operates). Subsequent federal regulations of the late 1960s and early 1970s specified what was meant by affirmative action in the executive order, and the meaning of affirmative action was considerably expanded into the full-fledged program that has existed since 1971. Revised order no. 4 of that year, which is part of the *Code of Federal Regulations* and is still in effect, reads in part:

*An affirmative action program is a set of specific and result-oriented procedures to which a contractor commits itself to apply every good effort. The objective of those procedures plus such efforts is equal employment opportunity. Procedures without efforts to make them work are meaningless; and effort, undirected by specific and meaningful procedures, is inadequate. An effective affirmative action program must include an analysis of areas within which the contractor is deficient in the utilization of minority groups and women, and further, goals and timetables to which the contractor’s good faith efforts must be directed to correct the deficiencies and, [sic] thus to achieve prompt and full utilization of minorities and women, at all levels and in all segments of its work force where deficiencies exist. (Code of Federal Regulations 1990, pp. 121–122)*

Much of the controversy over affirmative action is over the term *goals and timetables*: Are these “quotas”? Supporters of affirmative action say not—only good faith efforts are required, and if they fail the contractor is not penalized. Further controversy exists over the term *utilization*: What is the basis on which a group is found “underutilized,” and to what extent is this evidence of discrimination? Similarly, there is considerable dispute over how to label these programs. “Affirmative action”

has a positive air, and in public opinion polls will receive considerable support. Label the same programs “racial preference”—which indeed is specifically what they are—and public support drops radically. In the 1990s, as campaigns were launched to ban affirmative action programs by popular referendum, just what language could or should be used in these referenda became hotly disputed.

Controversy also arises over the categories of employees that contractors must report on and over whose utilization they must be concerned. The executive order lists four categories: blacks, Spanish-surnamed Americans, American Indians, and Orientals. (These are the terms in the order as of 1971 and are still used in the *Code of Federal Regulations*.) The preferred names of these groups have changed since then to Afro- or African Americans, Hispanics or Latinos, Native Americans, and Asians. While the original executive order and the Civil Rights Act was a response to the political action of black civil rights groups, and it was the plight of blacks that motivated both the executive order and the Civil Rights Act, it was apparently deemed unwise in the mid-1960s to limit affirmative action requirements to blacks alone. The Civil Rights Act bans discrimination against any person on grounds of race, national origin, and sex and specifies no group in particular for protection; but the Equal Employment Opportunity Commission, set up by the Civil Rights Act to monitor discrimination in employment, from the beginning required reports on the four groups listed above, despite the fact that even in the 1960s it could be argued that discrimination against Asians was far less acute and much less of a problem than discrimination against blacks, that discrimination against American Indians also differed in severity and character from discrimination against blacks, and that discrimination against Spanish-surnamed Americans ranged from the nonexistent or hardly existent (Spaniards from Spain? Cubans? Sephardic Jews?) to the possibly significant. Nevertheless, these four categories set up in the mid-1960s are still the groups that governmental programs of affirmative action target for special attention (Glazer 1987).

Since affirmative action is a governmental program operated by government agencies that grant contracts and is overseen by the Office of

Federal Contract Compliance Programs, one major issue of controversy has been over the degree to which these programs are really enforced. It is generally believed that enforcement is more severe under Democratic administrations than under Republican administrations, even though the program was first fully developed under the Republican administration of President Richard Nixon. President Ronald Reagan was an avowed opponent of affirmative action, but despite his eight-year administration no modification of the program took place. Changes were proposed by some parts of the Republican administration but opposed by others. Business, in particular big business, had learned to live with affirmative action and was not eager to upset the apple cart (Belz 1990).

One of the most controversial areas in which affirmative action is applied is in the employment and promotion of police, firefighting and sanitation personnel, and teachers and other local government employees. Here strict racial quotas often do apply. They are strongly resented by many employees when new employees are hired by race and even more when promotions are given out by race and layoffs are determined by race. The basis of these quotas is not the presidential executive order but rather consent decrees entered into by local government on the basis of charges of discrimination brought by the federal government. These charges are brought on the basis of the Civil Rights Act; under this act, if discrimination is found, quotas can be required by courts as a remedy. Since local government employment is generally on the basis of tests, one very controversial aspect of such cases is the role of civil service examinations. Blacks and Hispanics characteristically do worse than white applicants. Are these poorer results to be taken as evidence of discrimination? A complex body of law has been built up on the basis of various cases determining when a test should be considered discriminatory. In the Civil Rights Act of 1964, one provision read “it shall not be an unlawful employment practice . . . for an employer to give and act upon the results of any professionally developed ability test provided that such test . . . is not designed, intended, or used to discriminate.” But the courts decide whether the test is “designed, intended, or used to discriminate.” Because of the frequency with which courts have found tests for the police, fire, or

sanitation force discriminatory, and because state and local governments believe they will lose such cases, many have entered into “consent decrees” in which they agree to hire and promote on the basis of racial and sex criteria.

Affirmative action is also used in the granting of government contracts on the basis of either statutes (federal, state, or local) or administrative procedures. It is in this area that the edifice of affirmative action was first effectively attacked in the 1980s and 1990s, in the wake of the failure of Republican administrations to take any action limiting affirmative action. The first major crack came in the U.S. Supreme Court’s decision in *City of Richmond v. Croson*, in which the Court ruled against a Richmond, Virginia, city program to set aside 30 percent of city contracts for minority-owned businesses. The Court ruled that such programs could only stand, under the judicial “strict scrutiny” standard triggered by apparent government discrimination, if it could be demonstrated there had been discrimination against these groups by the city in the past. The response of cities to this judicial limitation on their minority “set-aside” programs was often to commission studies to demonstrate they had indeed discriminated, in order to save the set-aside programs (LaNoue 1993). In a later decision, *Adarand v. Peña* (1995), the Court ruled against a federal statute requiring that 10 percent of public works contracts be set aside for minorities. While these programs still continue in many jurisdictions, including the federal government, they are all legally threatened.

One issue in minority contract set-asides has been that of possible fraud, as various contractors find it to their advantage to take on black partners so as to present themselves as minority contractors and thus to get whatever advantages in bidding that status provides. In this area, as in other areas where advantage might follow from minority status, there have been debates over what groups may be included as minorities. It was unclear, for example, whether Asian Indians—immigrants and American citizens of Indian origin—were to be considered Asian. Asian Indian Americans were divided among themselves on this question, but during the Reagan administration they were reclassified as Asian, presumably in part for the modest political advantage this gave the Republican administration.

Affirmative action also governs the employment practices of colleges and universities, whether public or private, because they all make use of federal grants and loans for their students, and many have government research contracts. Colleges and universities therefore must also survey their faculties and other staffs for underutilization, and they develop elaborate affirmative action programs. Affirmative action applies to women as well as to racial and ethnic minority groups. There has been, perhaps in part because of affirmative action programs, a substantial increase in female faculty. But there has been little increase in black faculty during the 1980s. The numbers of blacks taking doctorates in arts and sciences has been small and has not increased. The higher rewards of law, business, and medicine have attracted into those fields black students who could prepare themselves for an academic career. Many campuses have been shaken by controversies over the small number of black faculty, with administrators arguing that few were available and protestors, often black students, arguing that greater effort would change the situation.

In the 1990s, the most controversial area of affirmative action became admissions to selective colleges and universities and professional schools. Affirmative action in admissions is not required by government regulations, as in the case of employment, except in the special case of southern public higher education institutions. There parallel and separate black and white institutions existed, and while all of these institutions have been open to both white and black students since at least the early 1970s, an extended lawsuit has charged that they still preserve their identity as traditionally black and traditionally white institutions. As a result of this litigation, many of these institutions must recruit a certain number of black students. But the major pressure on many other institutions to increase the number of black students has come from goals voluntarily accepted by administrators or as a result of black student demands. (In one case, that of the University of California, the state legislature has called on the institution to mirror in its racial-ethnic composition the graduating classes of California high schools.) Voluntary affirmative action programs for admission of students, targeted on black, Hispanic, and Native American students, became widespread in the late

1960s and early 1970s, particularly after the assassination of Martin Luther King.

The first affirmative action cases to reach the Supreme Court challenged such programs of preference for black and other minority students. A rejected Jewish applicant for admission to the University of Washington Law School, which had set a quota to increase the number of its minority students, sued for admission, and his case reached the Supreme Court. It did not rule on it. The Court did rule on a subsequent case, in 1978, that of Allan Bakke, a rejected applicant to the University of California, Davis, Medical School, which also had set a quota. The Court, splitting into a number of factions, rejected fixed numerical quotas but asserted race was a factor that could be taken into account in admissions decisions for purposes of promoting academic diversity. Under the protection offered by this complex decision, most colleges and universities and professional schools do grant preferences to black and Hispanic students. Asian Americans, also considered a minority, did not receive preference, but this was hardly necessary since their academic achievement is high. Indeed, by the 1980s, Asian American students were protesting that it was more difficult for them than whites to gain admission to selective institutions (Bunzel and Au 1987).

Matters turned around in the mid-1990s. In 1995, the regents of the University of California banned any consideration of race or ethnicity in admission to the university. In 1996 the voters of California approved the California Civil Rights Initiative, which banned the use of race or ethnic criteria in state government action, in employment and contracts as well as college admissions. This initiative, launched by one academic and one former academic, became the basis of a movement to extend the ban on affirmative action. In 1998 the State of Washington became the second state to pass such an initiative. In Texas, the assault on affirmative action in admissions led to a wide-ranging decision by the Fifth Circuit Court of Appeals in 1996 banning the use of race and ethnicity in admissions to the University of Texas Law School. This decision affects all institutions of higher education in the states covered by the fifth circuit. Massive changes have followed in the admissions procedures of the University of California, the peak institution in the system of public

higher education in California, and the University of Texas, which holds a similar position in Texas.

Initially, there were substantial drops in enrollment of black and Hispanic students, but energetic action by the university administrations has stemmed this fall-off, and the decline statistically is not as drastic as originally projected or feared. The commitment by university administrators to maintain a substantial representation of black and Hispanic students is so strong that they have devised new admissions practices and procedures designed to keep the number up, and have had some success in doing so. Further, the Texas legislature has voted that the top 10 percent of every high school graduating class be eligible for admission to the University of Texas, and the Board of Regents of the University of California has similarly voted that the top 4 percent of California high school graduating classes should be eligible for enrollment in the University of California. The effect of such actions, in view of the high concentration of black and Hispanic students in low-achieving high schools, which ordinarily send few students to the selective state institutions, is to keep up the number of black and Hispanic students.

Affirmative action in admissions has become perhaps the best-researched area of affirmative action as a result of these controversies. An important study by William Bowen and Derek Bok of admissions procedures in selective institutions has argued effectively that "race-sensitive" admissions have been good for the students, good for the institutions, and good for the country. Their conclusions, however, have been sharply disputed by critics of affirmative action. (Bowen and Bok 1998; Trow 1999). At this writing (Fall, 1999) the Supreme Court has not yet ruled on the issue of the degree to which public colleges and universities may take race into account. In view of the fact that neither Congress nor state legislatures will take decisive action on race preference owing to the political sensitivity of the issue, it is clear the key decisions in this area will have to be taken by the Supreme Court. It is possible that the age of affirmative action in American race relations and race policy is coming to an end.

Affirmative action has been a divisive issue in American political life and has sometimes been raised effectively in political campaigns. It has

divided former allies on civil rights issues, in particular American Jews, normally liberal, from blacks. Jews oppose quotas in admissions to medical and law schools because they were in the past victims of very low quotas imposed by American universities.

Affirmative action under various names and legal arrangements is found in many countries: in India, to provide opportunities to scheduled castes, scheduled tribes, and other backward classes, where different requirements operate at the national level and within the states and where some degree of preference has existed in some areas and for some purposes as far back as the 1920s (Galanter 1984); in Malaysia, to protect the native Malay population; in Sri Lanka, to benefit the majority Sri Lankan population (Sowell 1990); and in Australia and Canada, where milder forms of affirmative action than those found in the United States operate. It is now being raised in the countries of Western Europe, which have received since World War II large numbers of immigrants who now form distinctive communities who lag behind the native populations in education and occupational status. The policies called affirmative action in the United States are called "reservations" in India, and "positive discrimination" in some other countries.

There is considerable debate as to the effects of affirmative action policies and how weighty these can be as against other factors affecting employment, promotion, and educational achievement (Leonard 1984a; 1984b). A summary judgment is difficult to make. Black leaders generally consider affirmative action an essential foundation for black progress, but some black intellectuals and publicists have been skeptical. Black leaders often denounce opponents of affirmative action as racists, hidden or otherwise, yet it is clear that many opponents simply find the use, required or otherwise, of racial and sexual characteristics to determine job and promotion opportunities and admission to selective college programs in contradiction with the basic liberal principles of treating individuals without regard to race, national origin, color, and sex. Affirmative action has undoubtedly increased the number of blacks who hold good jobs and gain admission to selective programs. But it has also had other costs in the form of increased racial tensions. It has coincided with a period in which a pattern of black advancement occupationally and educationally since World War II has been

surprisingly slowed. The defenders of affirmative action argue that this is because it has not yet been applied vigorously enough. The opponents argue that the concentration on affirmative action encourages the neglect of the key factors that promote educational and occupational progress, which are basically the acquisition of qualifications for better jobs and superior educational programs. (SEE ALSO: *Discrimination; Equality of Opportunity*)

## REFERENCES

- Belz, Herman 1990 *Equality Transformed: A Quarter-Century of Affirmative Action*. New Brunswick, N.J.: Transaction Books.
- Bowen, William G., and Derek Bok 1990 *The Shape of the River: Long-term Consequences of Considering Race in College and University Admissions*. Princeton, N.J.: Princeton University Press.
- Bunzel, John H., and Jeffrey K. D. Au 1987 "Diversity or Discrimination? Asian Americans in College." *The Public Interest* 87:49-62.
- Code of Federal Regulations* 1990 Title 41, section 60-2.10, pp. 121-122.
- Galanter, Marc 1984 *Competing Equalities: Law and the Backward Classes in India*. Berkeley: University of California Press.
- Glazer, Nathan 1987 *Affirmative Discrimination: Ethnic Inequality and Public Policy*. Cambridge, Mass.: Harvard University Press. (Originally published 1975 and 1978, New York: Basic Books.)
- LaNoue, George R. 1993 "Social Science and Minority Set-Asides." *The Public Interest* 110:49-62.
- Leonard, Jonathan 1984a "The Impact of Affirmative Action on Employment." Working Paper No. 1,310. Cambridge, Mass.: National Bureau of Economic Research.
- 1984b "What Promises Are Worth: The Impact of Affirmative Action Goals." Working Paper no.1,346. Cambridge, Mass.: National Bureau of Economic Research. (Also published in *Journal of Human Resources* (1985) 20:3-20.)
- McWhirter, Darien A. 1996 *The End of Affirmative Action: Where Do We Go From Here?* New York: Birch Lane Press.
- Orfield, Gary, and Edward Miller (eds.) 1998 *Chilling Admissions: The Affirmative Action Crisis and the Search for Alternatives*. Cambridge, Mass.: Harvard Education Publishing Group.
- Skrentny, John David 1996 *The Ironies of Affirmative Action*. Chicago: University of Chicago Press.



Sowell, Thomas 1990 *Preferential Policies: An International Perspective*. New York: William Morrow.

Trow, Martin 1999 "California After Racial Preferences." *The Public Interest* 135:64–85.

NATHAN GLAZER

## AFRICAN-AMERICAN STUDIES

Research on African Americans covers many important areas, only a few of which will be discussed here: theories of white-black relations; the enslavement of African Americans; the development of an antiblack ideology; the creation of white wealth with black labor; the idea of whiteness; racial discrimination today; and possibilities for social change.

### THEORIES OF WHITE-BLACK RELATIONS

Explanatory theories of U.S. racial relations can be roughly classified into *order-deficit theories* and *power-conflict theories*. Order-deficit theories accent the gradual inclusion and assimilation of an outgroup such as African Americans into the dominant society and emphasize the barriers to progress that lie within the outgroup. Power-conflict theories, in contrast, emphasize past and present structural barriers preventing the full integration of African Americans into the society's institutions, such as the huge power and resource imbalance between white and black Americans. They also raise larger questions about the society's historically racist foundations.

Milton Gordon (1964) is an order-deficit scholar who distinguishes several types of initial encounters between racial and ethnic groups and an array of subsequent assimilation outcomes ranging from acculturation to intermarriage. In his view the trend of immigrant adaptation in the United States has been in the direction of substantial conformity—of immigrants giving up much of their heritage for the Anglo-Protestant core culture. Gordon and other scholars apply this scheme to African Americans, whom they see as substantially assimilated at the cultural level (for example, in regard to language), with some cultural differences remaining because of "lower class subculture" among

black Americans. This order-deficit model and the accent on defects in culture have been popular among many contemporary analysts who often seek to downplay discrimination and accent instead problems internal to black Americans or their communities. For example, Jim Sleeper (1990) has argued that there is no institutionalized racism of consequence left in United States. Instead, he adopts cultural explanations—for example, blacks need to work harder—for present-day difficulties in black communities.

In contrast, power-conflict analysts reject the assimilationist view of black inclusion and eventual assimilation and the inclination to focus on deficits within African-American individuals or communities as the major barriers to racial integration. From this perspective, the current condition of African Americans is more oppressive than that of any other U.S. racial or ethnic group because of its roots in centuries-long enslavement and in the subsequent semislavery of legal segregation, with the consequent low-wage jobs and poor living conditions. Once a system of extreme racial subordination is established historically, those in the superior position in the hierarchy continue to inherit and monopolize disproportionate socioeconomic resources over many generations. One important power-conflict analyst was Oliver C. Cox. His review of history indicated that from the 1500s onward white-on-black oppression in North America arose out of the European imperialistic system with its profit-oriented capitalism. The African slave trade was the European colonists' "Way of recruiting labor for the purpose of exploiting the great natural resources of America" (Cox 1948, p. 342). African Americans provided much hard labor to build the new society—first as slaves, then as sharecroppers and tenant farmers, and later as low-wage laborers and service workers in cities.

In the late 1960s civil rights activist Stokely Carmichael (later renamed Kwame Ture) and historian Charles Hamilton (1967) documented contemporary patterns of racial discrimination and contrasted their power-conflict perspective, which accented institutional racism, with an older approach focusing only on individual whites. They were among the first to use the terms "internal colonialism" and "institutional racism" to describe discrimination by whites as a group against African Americans as a group. Adopting a similar

power-conflict perspective, sociologist Bob Blauner (1972) argued that there are major differences between black Americans and the white immigrant groups at the center of assimilationist analysis. Africans brought across the ocean became part of an internally subordinated colony; white slaveowners incorporated them against their will. Black labor built up white wealth. European immigrants, in contrast, came more or less voluntarily.

More recently, sociologist Molefi Kete Asante (1988) has broken new ground in the development of a new power-conflict perspective, termed *afrocentricity*, that analyzes the Eurocentric bias in U.S. culture and rejects the use of the concepts such as “ethnicity,” “minority,” and “ghetto” as antithetical to developing a clear understanding of racism and to building antiracist movements. Similarly, anthropologist Marimba Ani (1994) has shown how from the beginning European colonialism was supported by a well-developed theory of white supremacy, a worldview that attempted to destroy the cultures of other non-European peoples. Yet other analysts (Feagin 2000) have argued for a power-conflict framework that understands white oppression of black Americans—with its racist ideology—as the foundation of U.S. society from the beginning. From this perspective much in the unfolding drama of U.S. history is viewed as a continuing reflection of that foundation of systemic racism.

### THE ENSLAVEMENT OF AFRICAN AMERICANS

Research on slavery emphasizes the importance of the power-conflict perspective for understanding the history and conditions of African Americans. Manning Marable (1985, p. 5) demonstrates that before the African slave trade began, Europeans were predisposed to accept slavery. Western intellectuals from Aristotle to Sir Thomas More defended slavery.

Research by sociologists, historians, and legal scholars emphasizes the point that slavery is the foundation of African-American subordination. Legal scholar Patricia Williams has documented the dramatic difference between the conditions faced by enslaved Africans and by European immigrants: “The black slave experience was that of lost

languages, cultures, tribal ties, kinship bonds, and even of the power to procreate in the image of oneself and not that of an alien master” (Williams 1987, p. 415). Williams, an African American, discusses Austin Miller, her great-great grandfather, a white lawyer who bought and enslaved her great-great-grandmother, Sophie, and Sophie’s parents. Miller forced the thirteen-year-old Sophie to become the mother of Williams’s great grandmother Mary. Williams’s white great-great grandfather was thus a rapist and child molester. African Americans constitute the only U.S. racial group whose heritage involves the forced mixing of its African ancestors with members of the dominant white group.

Eugene Genovese (1974) has demonstrated that the enslavement of African Americans could produce both servile accommodation and open resistance. Even in the extremely oppressive slave plantations of the United States, many peoples from Africa—Yorubas, Akans, Ibos, and others—came together to create one African-American people. They created a culture of survival and resistance by drawing on African religion and values (Stuckey 1987, pp. 42–46). This oppositional culture provided the foundation for many revolts and conspiracies to revolt among those enslaved, as well as for later protests against oppression.

In all regions many whites were implicated in the slavery system. Northern whites built colonies in part on slave labor or the slave trade. In 1641, Massachusetts was the first colony to make slavery legal; and the state’s merchants and shippers played important roles in the slave trade. Not until the 1780s did public opinion and court cases force an end to New England slavery. In 1786 slaves made up 7 percent of the New York population; not until the 1850s were all slaves freed there. Moreover, an intense political and economic subordination of free African Americans followed abolition (Higginbotham 1978, pp. 63–65, 144–149). As Benjamin Ringer (1983, p. 533) puts it, “despite the early emancipation of slaves in the North it remained there, not merely as fossilized remains but as a deeply engrained coding for the future.” This explains the extensive system of antiblack discrimination and “Jim Crow” segregation in northern states before the Civil War. Later, freed slaves and their descendants who migrated there from the South came into a socioeconomic system already coded to subordinate African Americans.

### THE DEVELOPMENT OF AN ANTIBLACK IDEOLOGY

Truly racist ideologies—with “race” conceptualized in biologically inferiority terms—appear only in modern times. St. Clair Drake (1987) has shown that in the Greek and Roman periods most Europeans attached greater significance to Africans’ culture and nationality than to their physical and biological characteristics. Beginning with Portuguese and Spanish imperialism in the fifteenth century, a racist ideology was gradually developed to rationalize the brutal conquest of the lands and labor undertaken in the period of European imperialism (Snowden 1983).

The system of antiblack racism that developed in the Americas is rooted deeply in European and Euro-American consciousness, religion, and culture. Europeans have long viewed themselves, their world, and the exploited “others” within a parochial perspective, one that assumes European culture is superior to all other cultures, which are ripe for exploitation (Ani 1994). For the colonizing Europeans it was not enough to bleed Africa of its labor. A well-developed anti-African, antiblack ideology rationalized this oppression and thus reduced its moral cost for whites. As it developed, this ideology accented not only the alleged physical ugliness and mental inferiority of Africans and African Americans, but also their supposed immorality, family pathologies, and criminality. Notions that African Americans were, as the colonial settlers put it, “dangerous savages” and “degenerate beasts,” were apparently an attempt by those who saw themselves as civilized Christians to avoid blame for the carnage they had created. As historian George Frederickson put it, “otherwise many whites would have had to accept an intolerable burden of guilt for perpetrating or tolerating the most horrendous cruelties and injustices” (Frederickson 1971, p. 282).

### CREATING WHITE WEALTH WITH BLACK LABOR

In his masterpiece *The Souls of Black* (1903), the pioneering sociologist W. E. B. Du Bois anticipated current research on the significance of African Americans for U.S. prosperity and development:

“Your country. How came it yours? Before the Pilgrims landed we were here. Here we have brought our three gifts and mingled them with yours: a gift of story and song—soft, stirring melody in an ill-harmonized and unmelodious land; the gift of sweat and brawn to beat back the wilderness, conquer the soil, and lay the foundations of this vast economic empire two hundred years earlier than your weak hands could have done it; the third, a gift of the spirit” (Du Bois 1989 [1903], pp. 186–187). The past and present prosperity of the nation is substantially the result of the enforced labor of millions of African Americans under slavery and segregation.

Under common law, an innocent individual who benefits unknowingly from wealth gained illegally or by unjust actions in the past generally cannot, if the ill-gotten gains are discovered, claim a right to keep them (Cross 1984, p. 510; Williams 1991, p. 101). A coerced taking of one’s possessions by an individual criminal is similar to the coerced taking of one’s labor by a white slaveholder or discriminator. Over centuries great wealth was unjustly created for white families from the labor of those enslaved, as well as from the legal segregation and contemporary racist system that came after slavery.

Some researchers (see America 1990) have examined the wealth that whites have unjustly gained from four hundred years of the exploitation of black labor. Drawing on James Marketti (1990, p. 118), one can estimate the dollar value of the labor taken from enslaved African Americans from 1620 to 1861—together with lost interest from then to the present—as between two and five *trillion* dollars (in current dollars). Adding to this figure the losses to blacks of the labor market discrimination in place from 1929 to 1969 (plus lost interest) would bring the total figure to the four to nine trillion dollar range (see Swinton 1990, p. 156). Moreover, since the end of legal segregation African Americans have suffered more economic losses from continuing discrimination. For more than two decades now the median family income of African-American families has been about 55 to 61 percent of the median family income of white families. Compensating African Americans for the value of the labor stolen would clearly require a very large portion of the nation’s

current and future wealth. And such calculations do not take into account the many other personal and community costs of slavery, segregation, and modern racism.

Research has shown that for centuries whites have benefited from large-scale government assistance denied to African Americans. These programs included large-scale land grants from the 1600s to the late 1800s, a period when most blacks were ineligible. In the first decades of the twentieth century many government-controlled resources were given away, or made available on reasonable terms, to white Americans. These included airline routes, leases on federal lands, and access to radio and television frequencies. During the 1930s most federal New Deal programs heavily favored white Americans. Perhaps the most important subsidy program benefiting whites was the Federal Housing Administration (FHA) loan insurance that enabled millions of white families to secure homes and accumulate equity later used for education of their children (Sitkoff 1978). During the long segregation period from the 1890s to the late 1960s, black families received much less assistance from government programs and were unable to build up wealth comparable to that of white families.

#### **WHITENESS DEVELOPS IN RELATION TO BLACKNESS**

From the beginning of the nation, the ideas of “whiteness” and “blackness” were created by whites as an integral part of the increasingly dominant racist ideology. The first serious research on whiteness was perhaps that of Du Bois (1992 [1935]), who showed how white workers have historically accepted lower monetary wages in return for the *public and psychological wage* of white privilege. In return for not unionizing or organizing with black workers, and thus accepting lower monetary wages, white workers were allowed by employers to participate in a racist hierarchy where whites enforced deference from black Americans. Several social scientists (Roediger 1991; Allen 1994; Brodtkin 1998) have shown how nineteenth-century and twentieth-century immigrants from Europe—who did not initially define themselves as “white” but rather as Jewish, Irish, Italian, or other European

identity—were pressured by established elites to view themselves and their own groups as white. Racial privileges were provided for new European immigrants as they aligned themselves with the native-born dominant Anglo-American whites and actively participated in antiblack discrimination.

Today, whites still use numerous myths and stereotypes to defend white privilege (Frankenberg 1993; Feagin and Vera 1995). Such fictions often describe whites as “not racist” or as “good people” even as the same whites take part in discriminatory actions (for example, in housing or employment) or express racist ideas. In most cases, the positive white identity is constructed against a negative view of black Americans.

#### **RACIAL DISCRIMINATION TODAY**

Racial discrimination involves the practices of dominant group members that target those in subordinate groups for harm. Discrimination maintains white wealth and privilege. A National Research Council report noted that by the mid-1970s many white Americans believed “the Civil Rights Act of 1964 had led to broad-scale elimination of discrimination against African Americans in public accommodations” (Jaynes and Williams 1989, p. 84). However, social science research still finds much racial discrimination in housing, employment, education, and public accommodations.

Many scholars accent current black economic progress. In the 1980s and 1990s the number of black professional, technical, managerial, and administrative workers has increased significantly. Yet African Americans in these categories have been disproportionately concentrated in those jobs with lower status. Within the professional-technical category, African Americans today are most commonly found in such fields as social and recreational work, public school teaching, vocational counseling, personnel, dietetics, and health-care work; they are least often found among lawyers and judges, dentists, writers and artists, engineers, and university teachers. Within the managerial-administrative category African Americans are most commonly found among restaurant and bar managers, health administrators, and government officials; they are least commonly found among top corporate executives, bank and financial managers, and wholesale sales managers. Such patterns

of job channeling indicate the effects of intentional and indirect racial discrimination over several centuries.

Studies have shown that housing segregation remains very high in U.S. metropolitan areas, North and South. This is true for both low-income and middle-income African Americans. Change in residential balkanization has come very slowly. Census data indicate that from 1980 to 1990 there were only small decreases in the level of residential segregation in thirty major U.S. metropolitan areas; fewer changes than for the previous decade. Two-thirds of the black residents of the southern metropolitan areas and more than three-quarters in northern metropolitan areas would have to move from their present residential areas if one wished to create proportional desegregation in housing arrangements in these cities (Massey and Denton 1993, pp. 221–223). U.S. cities remain highly segregated along racial lines.

At all class levels, African Americans still face much discrimination. After conducting pioneering interviews with forty black women in the Netherlands and the United States, social psychologist Philomena Essed (1990) concluded that racial discrimination remains an omnipresent problem in both nations. She showed that black accounts of racism are much more than discrete individual accounts, for they also represent systems of knowledge that people collectively accumulate to make sense of the racist society. Research by U.S. social scientists has confirmed and extended these findings. Lois Benjamin (1991), Kesho Y. Scott (1991), and Joe Feagin and Melvin Sikes (1994) conducted in-depth interviews to move beyond the common litany of black underclass pathologies to document the racial discrimination that still provides major barriers to black mobility in U.S. society. These field studies have revealed the everyday character of the racial barriers and the consequent pain faced by blacks at the hands of whites in employment, housing, education, and public accommodations.

Researchers Nancy Krieger and Stephen Sidney (1996) gave about 2,000 black respondents a list of seven settings where there might be discrimination. Seventy percent of the female respondents and 84 percent of the male respondents reported facing discrimination in at least one area.

The majority reported discrimination from whites in at least three settings. Several national surveys as well have found substantial racial discrimination. For example, a 1997 Gallup survey (1997, pp. 29–30, 108–110) inquired of black respondents if they had faced discrimination in five areas (work, dining out, shopping, with police, in public transportation) during the last month. Forty-five percent reported discrimination in one or more of these areas in that short period.

Discrimination for most black Americans entails much more than an occasional discriminatory act, but rather a lifetime of thousands of blatant, covert, and subtle acts of differential treatment by whites—actions that cumulate to have significant monetary, psychological, family, and community effects. African Americans contend against this discrimination in a variety of ways, ranging from repressed rage to open resistance and retaliation (Cobbs 1988). This cumulative and persisting discrimination is a major reason for the periodic resurgence of civil rights organizations and protest movements among African Americans (see Morris 1984).

#### **ANTIBLACK RACISM AND OTHER AMERICANS OF COLOR**

During the 1990s numerous researchers from Latino, Asian, and Native-American groups accentuated their own group's perspective on racial and ethnic relations and their experience with discrimination in the United States. They have often criticized a binary black-white paradigm they feel is dominant in contemporary research and writing about U.S. racial-ethnic relations (see Perea 1997). From this perspective, the binary black-white paradigm should be abandoned because each non-European group has its distinctive experiences of oppression.

However, a few scholars (Feagin 2000; Ani 1994) have shown the need to adopt a broader view of the long history and current realities of U.S. racism. The racist foundation of the nation was laid in the 1600s by European entrepreneurs and settlers as they enslaved Africans and killed or drove off Native Americans. By the middle of the seventeenth century African Americans were treated by whites, and by the legal system, as chattel

property—a position held by *no other group* in the four centuries of American history. This bloody article-of-property system created great wealth for whites and was soon rationalized in the aforementioned racist ideology. Ever since, whites have remained in firm control of all major U.S. institutions; they have perpetuated a racist system that is still imbedded in all these institutions.

White-on-black oppression is a comprehensive system originally designed for the exploitation of African Americans, one that for centuries has shaped the lives of every American, regardless of background, national origin, or time of entry. This long-standing white-racist framework has been extended and tailored for each new non-European group brought into the nation. Immigrants from places other than Europe, such as Chinese and Japanese immigrants from the mid-1800s to the 1950s and Mexican immigrants after 1900, were often oppressed for white gain and constructed as racialized inferiors without citizenship rights (Takaki 1990). Other types of white racism have been important in U.S. history, but white-on-black racism is the most central case. While it has changed in some ways as time has passed, this systemic racism has stayed roughly constant in its fundamentals. U.S. society is not a multiplicity of disconnected racisms, but has a central white-racist core that was initially developed by whites as they drove Native Americans off their lands and intensely exploited enslaved African Americans. Scholarship (Takaki 1990; Feagin 2000) has shown how this framework was gradually extended and adapted for the oppression of all other non-European groups.

### POSSIBILITIES FOR CHANGE

Some African-American scholars have expressed great pessimism about the possibility of significant racial change in the United States. The constitutional scholar Derrick Bell (1992) argues that racism is so fundamental that white Americans will never entertain giving up privileges and thus that black Americans will never gain equality.

There is a long history of African Americans and other people of color resisting racism. The development of resistance movements in the 1950s and 1960s was rooted in activism in local organizations, including churches, going back for centuries

(Morris 1984). Given these deep and persisting roots, many other analysts, black and nonblack, remain optimistic about the possibility of civil rights action for social change. Thus, legal scholar Lani Guinier (1994) has spelled out new ideas for significantly increasing the electoral and political power of black Americans. While the Voting Rights Act (1965) increased the number of black voters and elected officials, it did not give most of these officials an adequate or substantial influence on political decisions in their communities. Guinier suggests new strategies to increase black influence, including requiring supermajorities (a required number of votes from black officials who are in the minority) on key political bodies when there are attempts to pass major legislation.

Some scholars and activists are now pressing for a two-pronged strategy that accents both a continuing civil rights struggle outside black communities and an internal effort to build up self-help projects within those communities. A leading scholar of civil rights, Roy Brooks (1996), has documented the failures and successes of the traditional desegregation strategy. Though still supportive of integration efforts, Brooks has argued that blacks must consider separatist, internally generated, community development strategies for their long-term economic, physical, and psychological success. Working in the tradition of W. E. B. Du Bois and Malcolm X, numerous African-American scholars and community leaders have reiterated the import of traditional African and African-American values for the liberation of their communities from continuing discrimination and oppression by white Americans.

### REFERENCES

- Allen, Theodore W. 1994 *The Invention of the White Race: Racial Oppression and Social Control*. London: Verso.
- America, Richard F. (ed.) 1990 *The Wealth of Races: The Present Value of Benefits from Past Injustices*. New York: Greenwood Press.
- Ani, Marimba 1994 *Yurugu: An African-Centered Critique of European Cultural Thought and Behavior*. Trenton, N.J.: Africa World Press.
- Asante, Molefi Kete 1988 *Afrocentricity*. Trenton, N.J.: Africa World Press.
- Bell, Derrick 1992 *Faces at the Bottom of the Well: The Permanence of Racism*. New York: Basic Books.

- Benjamin, Lois 1991 *The Black Elite: Facing the Color Line in the Twilight of the Twentieth Century*. Chicago: Nelson-Hall.
- Blauner, Bob 1972 *Racial Oppression in America*. New York: Harper and Row.
- Brodkin, Karen 1998 *How The Jews Became White Folks: And What That Says About Race in America*. New Brunswick, N.J.: Rutgers University Press.
- Brooks, Roy L. 1996 *Integration or Separation? A Strategy for Racial Equality*. Cambridge, Mass.: Harvard University Press.
- Carmichael, Stokely, and Charles Hamilton 1967 *Black Power*. New York: Random House/Vintage Books.
- Cross, Theodore 1984 *Black Power Imperative: Racial Inequality and the Politics of Nonviolence*. New York: Faulkner.
- Drake, St. Clair 1987 *Black Folk Here and There*. Los Angeles: UCLA Center for Afro-American Studies.
- Du Bois, William E. B. (1935) 1992 *Black Reconstruction in America: An Essay Toward a History of the Part which Black Folk Played in the Attempt to Reconstruct Democracy in America, 1860–1880*. New York: Atheneum.
- (1903) 1989 *The Souls of Black Folk*. New York: Bantam Books.
- Essed, Philomena 1990 *Everyday Racism: Reports from Women of Two Cultures*. Claremont, Calif.: Hunter House.
- Feagin, Joe R. 2000 *Racist America*. New York: Routledge.
- , and Melvin P. Sikes 1994 *Living with Racism: The Black Middle Class Experience*. Boston: Beacon.
- Feagin, Joe R., and Hernan Vera 1995 *White Racism: The Basics*. New York: Routledge.
- Frankenberg, Ruth 1993 *White Women, Race Matters: The Social Construction of Whiteness*. Minneapolis: University of Minnesota.
- Fredrickson, George M. 1971 *The Black Image in the White Mind: The Debate on Afro-American Character and Destiny, 1817–1914*. Hanover, N.H.: Wesleyan University Press.
- Gallup Organization 1997 *Black/White Relations in the United States*. Princeton, N.J.: Gallup Organization.
- Gordon, Milton 1964 *Assimilation in American Life*. New York: Oxford University Press.
- Guinier, Lani 1994 *The Tyranny of the Majority: Fundamental Fairness in Representative Democracy*. New York: Free Press.
- Higginbotham, A. Leon 1978 *In the Matter of Color*. New York: Oxford University Press.
- Jaynes, Gerald D., and Robin Williams, Jr. (eds.) 1989 *A Common Destiny: Blacks and American Society*. Washington, D.C.: National Academy Press.
- Krieger, Nancy, and Stephen Sidney 1996 “Racial Discrimination and Blood Pressure.” *American Journal of Public Health* 86:1370–1378.
- Marketti, James 1990 “Estimated Present Value of Income Diverted during Slavery.” In Richard F. America, ed., *The Wealth of Races: The Present Value of Benefits from Past Injustices*. New York: Greenwood Press.
- Massey, Douglas S., and Nancy A. Denton 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Morris, Aldon D. 1984. *The Origins of the Civil Rights Movement*. New York: Free Press.
- Perea, Juan F. 1997 “The Black/White Binary Paradigm of Race: The ‘Normal Science’ of American Racial Thought.” *California Law Review* 85:1219–1221.
- Ringer, Benjamin B. 1983 *“We the People” and Others*. New York: Tavistock.
- Roediger, David R. 1991 *The Wages of Whiteness*. London and New York: Verso.
- Scott, Kesho Y. 1991 *The Habit of Surviving: Black Women’s Strategies for Life*. New Brunswick, N.J.: Rutgers University Press.
- Sitkoff, Harvard 1978 *A New Deal for Blacks*. New York: Oxford.
- Sleeper, Jim 1990 *The Closest of Strangers: Liberalism and the Politics of Race in New York*. New York: Norton.
- Snowden, Frank 1983 *Color Prejudice*. Cambridge, Mass.: Harvard University Press.
- Stuckey, Sterling 1987 *Slave Culture*. New York: Oxford University Press.
- Swinton, David H. 1990 “Racial Inequality and Reparations.” In Richard F. America, ed., *The Wealth of Races: The Present Value of Benefits from Past Injustices*. New York: Greenwood Press.
- Takaki, Ronald 1990 *Iron Cages: Race and Culture in 19th Century America*. New York: Oxford University Press.
- Williams, Patricia 1991 *The Alchemy of Race and Rights*. Cambridge, Mass.: Harvard University Press.
- 1987 “Alchemical Notes: Reconstructing Ideals from Deconstructed Rights.” *Harvard Civil Rights and Civil Liberties Review* 22:401–434.

JOE R. FEAGIN

## AFRICAN STUDIES

African studies simply defined is the systematic, scientific study of African peoples, and their institutions, culture, and history. But such a simple definition fails to encompass adequately the complexity of this important but often neglected arena for sociological theory and research.

### THE GEOGRAPHY OF AFRICA

In sociology, the definition of one's unit of study is prerequisite to undertaking any research project. However, facts of history make the task of defining what encompasses African studies elusive. Geographic or spatial definitions are generally clear-cut, so if one asks, "Where is Africa?" a concise answer is expected. The continent of Africa is easily identifiable on any atlas or globe. By association African studies could be defined as all research falling within the identified physical boundaries of this continental landmass. This is a simple and neat solution—or so it would seem. In reality, however, even the task of defining the physical boundaries of Africa can be daunting. As a vast continent of rich cultural, linguistic, political, and historical diversity, Africa is subject to considerable geographical disaggregation. Thus, what is one continent is approached and conceived of as several subcontinents or subregions. For example, many people—both lay and professional alike—are not accustomed to thinking of Egypt and the northern Islamic states (e.g., Algeria, Libya, Algeria, Morocco) as part of the African continent. These African states are routinely separated in scholarly discourse from Africa and African studies, and are generally treated as parts of the Middle East or Mediterranean.

The tendency toward geographic disaggregation in the conceptualization and study of Africa is also apparent at levels beyond the "North African" versus "sub-Saharan African" distinction. African studies in the so-called "sub-Saharan" context is usually divided into presumably distinct regions of this vast continent: East Africa (e.g., Kenya, Somalia, Tanzania, Uganda); West Africa (e.g., Nigeria, Ghana, Senegal); Central Africa (e.g., Zaire, Congo, Central African Republic); and Southern Africa (e.g., South Africa, Zimbabwe, Lesotho).

An additional complication in approaching the geography of Africa is represented by the partitioning of the continent into arbitrarily designated and imposed nation-states. Climaxing during the critical decades between 1870 and 1914, Western European imperialists divided Africa among themselves in order to share in the exploitation of the continent's rich natural and human resources. The Berlin Conference, a meeting among various European powers, was held from November 15, 1884 to February 26, 1885 in response to envy and mistrust spurred by competing attempts to claim and colonize Africa. The gathering produced negotiated guidelines for the division and governance of Africa by Europeans. As one striking example, the Congo River basin, envied and hotly contested by several European powers (e.g., Portugal, Britain, France, Germany, Belgium), was by consensus ceded to King Leopold II and Belgium. As another example, while some contested claims were settled, West Africa continued to be divided unequally among the colonizing countries that were fighting amongst themselves for the land and the people: Portugal claimed 14,000 square miles; Germany took possession of 33,000 square miles; Britain declared ownership of 450,000 square miles; and France laid claim to 1.8 million square miles of African soil. In claiming the land, these European powers also declared dominion over the people who occupied these lands.

By 1898 Europeans had colonized most of Africa. Since the lines of demarcation were drawn with European rather than African interests in mind, these artificially imposed geographic boundaries often dissected cultural and national groups or tribes that had been unified entities for centuries before the arrival of Europeans. The political or national boundaries arbitrarily established by European conquest had profound, far-reaching consequences. These boundaries are largely responsible for contemporary borders and nation-states in Africa. Moreover, these artificial divisions or mergers, or both, damaged historic patterns and relationships and, by so doing, contributed in some degree to ongoing ethnic conflicts in Africa. As with icebergs formed over centuries where we only can see the tip, these historic machinations provided the impetus for many "modern" conflicts in Africa. In these cases, colonial influence persists long after independence was proclaimed.



As a strategy, colonizers chose and groomed new African leadership to reinforce the ideals of newly formed, imposed nation-states, thereby creating a native elite and fostering division and mistrust among the colonized. The strategic approach to implementation varied for different colonial powers. France and Britain were the most successful Western European colonizing powers. Interestingly, their approaches to subduing and exploiting the African colonies differed vastly. The French saw traditional African heads-of-state as occupying the least important position within their new administrative system. Existing African governments were viewed as obstacles to the ultimate integration of Africans with French culture and society. Thus, achieving the French goal necessitated the utter destruction of established governmental systems as well as the absolute eradication of traditional cultures. Similar to American laws during the era of slavery, practicing ancestral religions, speaking traditional languages, and participating in customs such as dancing were deemed unlawful and were punishable by whipping, torture, and even death. The British, on the other hand, were more indirect. They sought to mold traditional governmental structures to accomplish their goals at the district level, while closely regulating the colonial administration nationally. The intention was for their African colonies to follow the examples of Canada and Australia, eventually emerging as self-governing extensions of the British Empire.

The institutionalized practice of misrepresenting the scale of Africa in relation to other continents is another point of contention. A European-inspired and -dominated cartography has successfully institutionalized blatantly misrepresentative views of African topography. The traditional world map portrays Europe's landmass as much larger than its true physical reality. Since the 1700s, the Mercator map scale, the most widely used cartographic scale in the world, has distorted the sizes of continents to favor the Northern Hemisphere. While traditional mapmaking and representations have instilled a picture of North America as equal in size to Africa, the Sahara Desert alone is in fact roughly the same size as the United States. The African continent has nearly four times the landmass of North America and comprises approximately twenty percent of the world's landmass.

More subtly, the world's geographic view of Africa has evolved to attribute a unidimensional image of Africa as consisting wholly of lush, impenetrable, tropical forests. This view of African geography fails to do justice to the rich, variegated landscape of this vast continent. Tropical rainforests represent only the smallest fraction of Africa's myriad landscape, which ranges from snow-capped mountains to deserts to high plains to hardwood forests, from rippling fields of grain to placid lakes. In fact the world's largest desert, the world's longest river, and natural wonders from the spectacular Victoria Falls to snow-capped Mount Kilimanjaro, all characterize the diversified topography of this continent.

### THE PEOPLE OF AFRICA

"Where is Africa?" We see that the answer to this, the most essential or rudimentary of originating questions for sociological research, can be quite elusive. Facts of history and perception combine to make what should be a simple interrogatory quite complicated. Equally elusive, if not more so, is the task of defining "Who is African?" The answer to this seemingly straightforward question seems obvious. More often than not, answers to this question conform to the widespread view of Africans as a race of black people characterized by dark skin, curly hair, broad noses and numerous other physical features. In fact, the biological diversity of Africans matches and, at points, surpasses Africa's vast geographic diversity. Few other continents in the world approach or match the breadth of human biology and physiology historically found in Africa. Africans run the gamut of the human color spectrum, encompassing the range of human prototypes—the Negroid, the Mongoloid, the Caucasoid. For centuries the continent of Africa has been home not only to people of traditional biophysical description, but also to people descended from Europeans, Asians, Arabs, and Hispanics. Moreover, centuries of biological intermingling have produced a continent of hybrid people. Africa is truly a continent where the human reality defies attempts to neatly categorize race and racial identity.

Traditional images of race fail to embrace or represent the African reality adequately. Modern

conceptions of race are derived from a hierarchical European worldview that assigns the lowest status to Africans, who are equated with blackness—the opposite of whiteness or European ancestry. When carefully examined in either an African or European context, racial identity loses much of its force as a concept. “White” and “black” are more political designations than physical ones. Thus, although there are indigenous Africans who have lighter complexions than some indigenous Europeans, widespread views of Europe as white and Africa as black persist. By the same token, race is often reified and incorrectly attributed characteristics that are in point of fact intellectual, social, cultural, or economic rather than biological. Generally speaking, race is incorrectly presumed to incorporate characteristics that extend far beyond human physical traits.

Apartheid in South Africa provides an excellent example of the politically motivated, exclusionary nature of racial classification systems (as does the historical case of “Jim Crow” segregation in the U.S. South). The assignment of people to categories of White, Coloured, and African (Black) in South Africa, coupled with the subdivision of each racial category into smaller racial groups (e.g., Whites = Afrikaner and English-speaking; Asians = East Indian and Malaysian; Coloureds = European/African and Asian/African; and Blacks = Zulu, Xhosa, Sotho), was largely determined by white efforts, as a demographic minority of the population, to maintain their historic political, economic, and social dominance. Apartheid relied on an elaborate system of “racial markers,” such as hair texture, skin color, and parentage, to classify people into distinct racial groups. Associated with each racial group were certain privileges and restrictions. Restrictions were heaviest and privileges least for black Africans, the group at the bottom of the South African color hierarchy. Thus race, as a socially constructed, politically manipulated reality in South Africa, exerted overwhelming force in limiting black and coloured Africans’ access and opportunities. The myth of the color hierarchy became—to some degree—a self-fulfilling, self-perpetuating prophecy in that many people, irrespective of race, internalized this value system. The result was often self-imposed ranking, with people of color creating additional levels within the established hierarchy, thereby constructing

intricate designations for race that extended far beyond traditional definition.

Discussions of African racial identity are additionally complicated by the vast global population of people of African ancestry. People of African descent are present in sizeable numbers in the Americas, Western Europe, and in parts of Asia and the South Pacific. In most instances, the dispersion of African people around the world—notably in Brazil, the United States, the Caribbean, and Britain—is directly traceable to the European conquest, domination, and distribution of African people. Europe created, installed, and operated a system of racial slave trade that fueled the economic, agricultural, and industrial development of the Americas and of Europe. The slave trade displaced millions of Africans and struck a crippling blow to social, economic, political, and cultural life on the African continent. The traffic in slaves was a demographic disaster for the African continent, taking away people in the prime of their reproductive and productive lives. The African slave trade left in its wake destroyed villages, ruined crops, disrupted cultures, and crumbling social institutions. So traumatic was the devastation of this trade in human lives, and the subsequent colonial exploitation of Africa for natural resources, that four centuries later Africa has not yet fully recovered.

Dramatic and extensive dispersion of Africans confounded questions of race and racial identity because of the extensive intermingling of Africans with other so-called racial groups. This was particularly true in the Americas, where systematic and legalized rape was characteristic of the era of enslavement. The manipulation of race and racial identity was a central feature in this drama. White men used sexual subjugation as another means of reinforcing their domination, resulting in the established pattern of African hybrid identity. Yet under the legal guidelines of American enslavement, the children who resulted from this institutionalized rape were considered black, and thereby referred to by Chief Justice Taney (*Dred Scott v. Sandford*, U.S. Supreme Court, 1857) as a people with “no rights which the white man was bound to respect.” African ancestry combined with Native American, Asian, and European bloodlines to further diversify the already rich biological heritage of Africans. The extension of the African diaspora

to the Americas produced a people represented by black Chinese in Mississippi, Jamaica, and Trinidad; black Amerindians in Florida (Seminoles), Oklahoma (Cherokee), and Surinam (Arawaks); black Irish in Virginia, black English in Barbados, black Portuguese in Brazil; and black East Indians in Guyana, black French Canadians in Montreal; and black Mexicans in Vera Cruz. A veritable human rainbow resulted from the transplantation of Africans to the “New World” and points beyond. Yet throughout the diaspora, Social Darwinism assigned Africans the lowest position on the evolutionary hierarchy and laid the foundation for elaborate ideology and pseudoscience that offered justifications for the continued subjugation of Africans. Ideologies founded upon the rationalization that lent credence to this travesty persist today, further plaguing Africans throughout the diaspora.

#### THE CULTURAL AND SOCIAL INSTITUTIONS OF AFRICA

We have seen the complexity that underlies the seemingly simple questions of “Where is Africa” and “Who is African.” The task of defining for sociological study a people who are both geographically and genetically dispersed is extremely challenging. Additional complication results when we ask the next question, “What is African?” Here we simply raise the logical question of which institutions, customs, values, institutions, cultural features, and social forms can be characterized or labeled as distinctively African.

The survival (albeit in evolved form) of indigenous African customs, values, and institutions is remarkable, especially considering the abundance of historical barriers and complications. First among these is the reality of African conquest and domination by other cultures, most notably European, but also including cultures from the Islamic world. The experience of conquest and domination by external powers, often designed to annihilate African civilization, makes the myriad existing retentions all the more amazing. One of the best examples of these retentions is what Africa and people of African ancestry have done with the abundance of nonindigenous languages that were imposed upon them during colonization.

Language plays a vital role in the process of cultural and personal affirmation. Therefore, it is not surprising to note that the conquest experience of Africa and Africans in the diaspora was commonly associated with systematic attempts to suppress or eliminate indigenous languages. The African continent can be divided into European language communities that parallel the geographic regions associated with European domination and partition of Africa (e.g., Anglophone, Francophone). In a similar fashion, members of the African diaspora have adopted the dominant languages of the cultures and regions where they found themselves, speaking Arabic, Portuguese, Spanish, English, French, or Dutch. However, before European languages were introduced, there were well over 800 languages spoken by various African ethnic groups, most of which can be classified between three of the principal language families—Niger-Kordofian, Nilo-Saharan, and Khoisian. Numerically, these African-selected language groups incorporate approximately 300 million people.

Many if not most Africans are minimally bilingual, routinely speaking several languages in addition to their own. However, colonial tongues are generally recognized as the country’s “official” language. Part of the devastation of colonization is the demise of original languages as the primary means of communication. These languages are forgotten, ignored, and sometimes even mocked by those who would assimilate into Westernized ideals. Strikingly, many independent African nations have embarked on programs aimed at the regeneration of indigenous languages, often creating written forms for languages that were previously solely oral. Other African nations have substituted indigenous languages for European languages (derived from the country’s colonial experience) as the country’s “official” language. South Africa is an example of a country where multiple national languages were officially established in a process that validated the myriad of mother tongues as well as the colonizers’ language.

Attempting to strip Africans of their language was an essential feature of the move to supplement military, economic, and political domination with cultural domination. However, understanding the essential connection between language, worldview, personal identity, and cultural survival,

many Africans fought to retain their indigenous languages. They believed their survival as a unique people hinged on successful retention of their original languages, and the culture, values, self-affirmation, and history embodied within those native tongues. Interesting variations in the retention of African language forms are observable throughout Africa and the African diaspora. The patois (patwa) spoken in Jamaica contains many words that need no translation from Twi, the language spoken by the Ghanaian Ashanti who constituted the majority of the slaves brought in to work on Jamaican plantations. French has been transformed in both Louisiana and Haiti with varieties of Kreyol. In North America we can observe the Gullah people, primarily found in South Carolina and Georgia, who infuse their English-based Creole with the language of their enslaved African ancestors. Standard black American English (sometimes referred to as Ebonics) also retains evidence of similar African-derived language structures. A related case is provided by places like Bahia in Brazil where Yoruba, an indigenous African language, was successfully transplanted and maintained outside the continent. There are nations on the African continent where indigenous languages were practiced throughout the colonial experience and reinstated after independence. However, as we have observed, the most common case across the diaspora is the synthesis of African syntax with the dominant European language syntax to create a new language.

The pattern of adaptive acculturation or synthesis that occurs with language also characterizes other sociocultural institutions in Africa and across the diaspora. Art and music provide distinctive examples. African people traditionally use visual art as a means of communication. Cloth making, for example, is not simply an aesthetic undertaking. Often, messages are inscribed or embedded within the very combinations of patterns and colors chosen to create the fabric. Among the Ashanti in Ghana, funeral cloth is often adorned with what are known as Adinkra symbols. The symbols impressed upon the cloth convey different meanings, thereby transmitting messages to those who are knowledgeable. The Touaregs, a people concentrated in Niger, are also known to transcribe messages on cloth in the distinctive alphabet that characterizes their language. Within Rastafarian

culture in Jamaica, sculpture is used as another method of communicating and reinforcing the beliefs of the practitioners. Moreover, African-American artists often draw from the diaspora in their creation of visual art. Lois Mailou Jones traversed the French-speaking African world, and her work was influenced by the various cultures she studied.

Artistic products of African people, especially those living outside of the continent, reflect the mixture of cultural influences—part indigenous, part nonindigenous—that characterize the African experience. The synthesis contains elements that are obviously African-derived alongside elements that are obviously European or Islamic. Yet the final tune or drawing is often truly syncretic, possessing emergent qualities greater than the sum of the influences. It is in the realm of cultural production, music and art, where the African influence on world society has been most readily apparent. During the period of colonization, ruling cultures often defined the oppressed as “others” in an effort to justify their subhuman treatment. Historic views of African-American culture provide powerful examples of objectification and stereotypes such as “Sambo,” the happy darkie, or “Jezebel,” the promiscuous black female. These false perceptions permeated the fabric of American society, allowing whites to embrace the racial stereotypes and notions that blacks were content in their status as second-class citizens. In order to perpetuate these false images, blacks were allowed to prosper in media and occupations that coincided with and reinforced views of their subservient status. To do otherwise was to risk retribution from whites. Thus, the entertainment industry (and its related variant, professional sports) has historically provided socially acceptable, non-threatening roles for blacks.

Many blacks recognized the power inherent in these media and used them for empowerment and to accomplish larger goals. For example, we can note the themes of resistance embodied in traditional spirituals sung during slavery. It is also interesting to observe the tremendous impact of music during the civil rights movement. In many instances, a syncretic merger has been achieved that fuses indigenous African “authenticity” with nonindigenous traits to create new forms of music

(e.g., rap). People of African descent have facilitated the growth of myriad musical genres, such as reggae, zouk, jazz, blues, soca, zaico, calypso, go-go, Hip Hop, and gospel. In yet other instances, Africans have chosen to embrace, master, and operate within the unmodified European form (e.g., opera, classical music).

Religion is another excellent example of a sociocultural institution that challenges us to define what is distinctively African. Forms of indigenous African religion are as diverse as the continent and people themselves. Christianity, Islam, and Judaism all have long histories on the continent. The same is true for traditional, animistic African religions that imbue all features of the environment with spiritual qualities. It is perhaps the long tradition of polytheism—or, more correctly, the belief in, and acceptance of a Supreme God who is in turn represented in the world by several intermediary spirits—that made Africa and African people such fertile ground for the spread of diverse religions.

In addition, colonial powers felt a need to control all forms of thought and expression in order to ensure their continued dominion over the indigenous people. Religion was recognized as an especially powerful medium with immense capacity to control the daily behavior of people. Thus, religion was a central component in the colonial arsenal and pattern of domination. In spite of this, many traditional African religions have survived the test of time. Both the original forms and syncretic amalgamations of Western and traditional religion are found throughout the diaspora. On the island of Jamaica, one can observe how the Ashanti religion Kumina has flourished and grown over time into the distinctly Jamaican Pukumina (Pocomania). Linkages such as the one between the Gabonese sect known as Bwiti and Haitian vodun continue, in which the words for traditional healer are almost the same. Continuities are also seen in Shango/Santeria (Brazil) and Rastafarianism (Jamaica and internationally). Religious forms throughout the African diaspora range from the classical European through the traditional African into the Islamic, with a large variety of syncretic forms interspersed throughout.

The organization of social and community life in Africa and across the African diaspora also runs

the spectrum from indigenous to nonindigenous. In a now familiar pattern, there exist places in Africa or in the African diaspora where traditional African forms of family life, dating back centuries, persist. Similarly, there are places where family forms are closer to the European or the Islamic model. The same observations can be offered about education and the organization of schools among Africans in a given society.

Predictably, given Africa's history of subjugation, it is in realms of international power and influence, economics, government, and the military, that African development has been most restricted. Historical deprivation has resulted in the ranking of Africa and people of African ancestry at the bottom of all indicators of economic development. While other formerly colonized countries have managed to advance economically (e.g., Brazil, Singapore, and Korea), much of Africa and the African diaspora continue to be economically dependent. Coupled with this economic dependence are persistently high rates of unemployment, disease, educational disadvantage, and population growth. Africans continue to be without significant representation at the centers of international power, judged by economic clout, political influence, and military might. In each of these aspects, Africa remains the suppressed giant, unable to exert world influence commensurate to her nearly billion people, strategic geopolitical location, and rich mineral and human resources. The emergence of Africa onto the world stage, like Japan's in 1960 and China's in 1980, will have to wait.

#### UNIFYING THEMES IN AFRICAN STUDIES

Due to historical context, we have seen there are no simple answers to the basic questions for orienting any scientific, sociological study of Africa and her people. The African diaspora has been uniquely shaped by experiences of conquest and domination, resistance and survival. The experiences of slavery, colonization, and external domination have left in their wake considerable devastation. Yet one also sees the amazing strength of a people able to adapt and diversify, and to love and live, and ultimately continue to be.

Studies of African people and institutions commonly reveal the creative retention of authentic or

indigenous traits. Such creative responses have been, in their own way, acts of resistance enabling cultural perpetuation. These adaptive responses have assured the ultimate survival of many aspects of African culture and institutions. It must also be noted that with time comes transition. Many things traditionally African have been altered, progressing from their original, indigenous form to completely new and different forms. It is perhaps in this tension, the reconciliation of the old with the new, the indigenous with the nonindigenous, that African studies will find its most exciting terrain for future inquiry. The challenge will be to discover the cultures and the people who have been historically distorted by the twin activities of concealment from within, and degradation and misperception from without.

#### **AMERICAN SCHOLARSHIP AND AFRICAN STUDIES**

The dismissal of the importance of African studies preceded its acceptance as a scholarly discipline. While clearly diminished today, the white male Eurocentric focus has historically dominated university curricula. Both grassroots and academic movements pointed to the need for recognition of the African contribution to world and American culture. The early to mid-1900s saw the emergence of black intellectuals in Africa (e.g., Léopold Sédar Senghor, Patrice Lumumba, Cheikh Anta Diop), the Caribbean (Frantz Fanon, Aimé Césaire, Claude McKay), and America (W.E.B. Du Bois, Langston Hughes, Anna Julia Cooper), who spoke of parallel movements based on the assumption that black people in all parts of the world were a community with shared interests and identity (e.g., Negritude, Pan-Africanism). The goals of these scholars emphasized the uplift of people of African ancestry worldwide through education, research, politics, and cultural activities. These international movements among black intellectuals were in direct response to theories of black inferiority and the systematic oppression of Africans (both on the continent and abroad). Herbert Spencer, among others, launched the ideology of Social Darwinism, which developed into a popular, pseudoscientific justification for racial hierarchies. Social Darwinism assigned Africans, and thereby African Americans, to the lowest rung in the evolutionary ladder.

A nineteenth-century phenomenon, Social Darwinism and its associated beliefs persisted well into the twentieth century. Its philosophy is still heard in the remnants of eugenics, in the present-day interest in sociobiology, and in recurring assertions of black innate intellectual inferiority.

Before the incorporation of African studies into predominantly white colleges, black educators and leaders had stressed the value of black America's ties to Africa for decades. Notable activists from William Monroe Trotter to Booker T. Washington and Marcus Garvey called for "Back to Africa" movements. Zora Neale Hurston was an anthropologist whose studies of African Americans emphasized retentions and links with original African cultures. The distinguished sociologist W.E.B. Du Bois founded the Pan-African Congress in 1921, supported by notables such as the author Jessie Redmon Fauset, in order to explicitly link the problems and fortunes of African Americans to those of blacks in Africa and elsewhere. Moreover, much of his work over a long and illustrious career focused on Africa. Ultimately Du Bois renounced his American citizenship and accepted Kwame Nkrumah's invitation to settle in Ghana.

Within white-dominated institutions, the value of African studies had strange origins. Egyptologists and anthropologists gathered information from Africa during colonial rule. In such instances, sociology was both friend and foe. Early sociologists promoted cross-cultural studies as well as research into the social conditions under which blacks lived. However, many of these sociologists embraced Social Darwinism and its belief in the inherent inferiority of blacks. Rarely did these early academics speak out or take active stances against the oppression of Africa and African people by Europeans.

Some American colleges remained racially segregated until the 1960s. The inclusion of African-American students, and later African-American studies classes, came in response to student activism, which occurred against the backdrop of the push for civil rights and amidst significant racial unrest. The black power movement of this time strongly influenced many African Americans to reclaim their heritage in everyday life and to demand that black history and culture be included in school and university curricula. There was—and

continues to be—significant struggle around these questions, since what is at risk here was (is) the control and validation of knowledge in the society. The black studies movement paved the way for further interest in African studies. The publication of Alex Haley's *Roots*, a volume which traced his family lineage back to Gambia, West Africa, followed by a popular television miniseries, was a signal moment in the development of black Americans' interest in Africa. In addition, many African Americans celebrate Kwanzaa, a holiday founded upon principles of the African harvest, as another conscious link with their cultural roots. Ironically, the growing interest of black Americans in their cultural roots in Africa helped to fuel a resurgence of ethnic pride and the search for roots among other racial groups in this country. Interestingly, in contemporary universities, the field of African-American studies tends to be dominated by African Americans, while African-studies programs tend to be dominated by European Americans.

### CONCLUSION

African studies is a woefully underdeveloped area of institutionalized research in the field of sociology. Researchers need to mount aggressive programs determined to "ask new questions and to question old answers." For this research to be successful, it must be located in broader context, recognizing the unique historical, economic, social, cultural, political, and academic relationships that determine reality for Africans on the continent and throughout the diaspora. In each of these areas, relationships are generally structured hierarchically, with African worldviews, values, institutional forms, methodologies, and concerns being considered subordinate to those of Europeans or whites. Such distorted structural relations lead inevitably to distortions in research and conclusions. The irony is that clear understanding of African people and institutions will help pave the way to better understanding Whites and European heritage, institutions, and experiences.

### REFERENCES

- Appiah, Anthony 1992 *In My Father's House: Africa in the Philosophy of Culture*. New York: Oxford University Press.
- Asante, Molefi 1987 *The Afrocentric Idea*. Philadelphia: Temple University Press.
- Bates, Robert H., V. Y. Mudimbe, and Jean O'Barr, (eds.) 1993 *Africa and the Disciplines: The Contributions of Research in Africa to the Social Sciences and Humanities*. Chicago, Ill.: University of Chicago Press.
- Césaire, Aimé 1972 *Discourse on Colonialism* New York: Macmillan Reference USA.
- Chinweizu 1975 *The West and the Rest of Us*. New York: Vintage.
- Diawara, Manthia 1998 *In Search of Africa*. Cambridge, Mass.: Harvard University Press.
- Diop, Cheikh Anta 1989 *The Cultural Unity of Black Africa: The Domains of Patriarchy and of Matriarchy in Classical Antiquity*. London: Karnak House.
- Fanon, Frantz 1986 *The Wretched of the Earth*. New York: Grove Press.
- Franklin, John Hope, and Alfred A. Moss, Jr. 1994 *From Slavery to Freedom: A History of African Americans*, 7th ed. New York: McGraw-Hill.
- Fyfe, Christopher (ed.) 1976 *African Studies Since 1945: A Tribute to Basil Davidson*. London: Longman.
- Gilroy, Paul 1993 *The Black Atlantic: Modernity and Double Consciousness*. Cambridge, Mass.: Harvard University Press.
- Gromyko, Anatoly, and C. S. Whitaker (eds.) 1990 *Agenda for Action: African-Soviet-U.S. Cooperation*. Boulder, Colo.: L. Rienner.
- Haley, Alex 1976 *Roots*. Garden City, N.Y.: Doubleday.
- Hochschild, Adam 1998 *King Leopold's Ghost: A Story of Greed, Terror and Heroism in Colonial Africa*. Boston, Mass.: Houghton-Mifflin.
- Jackson, Henry 1984 *From the Congo to Soweto: U.S. Foreign Policy Toward Africa Since 1960*. New York: Quill.
- James, C. L. R. 1969 *A History of Pan-African Revolt*, 2nd ed. Washington, D.C.: Drum and Spear Press.
- Mazrui, Ali 1986 *The Africans: A Triple Heritage*. London: BBC Publications.
- Mbiti, J. S. 1969 *African Religions and Philosophy*. London: Heinemann.
- Mudimbe, V. Y. 1988 *The Invention of Africa: Gnosis, Philosophy, and the Order of Knowledge*. Bloomington, Ind.: Indiana University Press.
- Nkrumah, Kwame 1965 *Neo-Colonialism: The Last Stage of Imperialism*. New York: International.
- Rodney, Walter 1974 *How Europe Underdeveloped Africa*. Washington, D.C.: Howard University Press.
- Soyinka, Wole 1976 *Myth, Literature and the African World*. New York: Cambridge University Press.
- Sunshine, Catherine 1988 *The Caribbean: Survival, Struggle and Sovereignty*. Boston: South End Press.

Thompson, Robert Farris 1984 *Flash of the Spirit: African and Afro-American Art and Philosophy*. New York: Random House.

Thompson, Vincent 1969 *Africa and Unity: The Evolution of Pan-Africanism*. London: Longman.

Winston, James 1998 *Holding Aloft the Banner of Ethiopia: Caribbean Radicalism in Early Twentieth-Century America*. London: Verso.

WALTER R. ALLEN  
SHANI O'NEAL

## AGGRESSION

In its most extreme forms, aggression is human tragedy unsurpassed. Hopes that the horrors of World War II and the Holocaust would produce a worldwide revulsion against the taking of another human's life, resulting in the end of genocidal practices and a reduction in homicide rates, have been dashed by the realities of increasing homicide and genocide in the last half of the twentieth century. The litany of genocidal events is both long and depressing, including major massacres in Uganda, Cambodia, Rwanda, Burundi, Zaire, Bosnia, Serbia, Croatia, and Herzegovina, among others. Homicide rates have risen in a number of industrialized countries since World War II, most notably in the United States.

We have seen slight declines in the homicide rate in the United States during the 1990s. But despite six consecutive years of decreases, the 1997 homicide rate was still 133 percent of the 1965 rate, and 166 percent of the 1955 rate. For these and related reasons, interest in understanding the causes of aggression remains high, and there have been major advances in the social psychology of aggression.

### WHAT IS AGGRESSION?

Definitions have varied widely over time and across research domains. However, a consensus has emerged among most social psychologists studying human aggression about what constitutes "aggression" in general and what constitutes the major forms or "ideal types" of aggression. (See the following books for current definitions and perspectives on aggression: Baron and Richardson

1994; Berkowitz 1993; Geen 1990; Geen and Donnerstein 1998; Tedeschi and Felson 1994).

**Basic Definitions.** *Aggression vs. Assertiveness vs. Violence.* Human aggression is behavior performed by one person (the aggressor) with the intent of harming another person (the victim) who is believed by the aggressor to be motivated to avoid that harm. "Harm" includes physical harm (e.g., a punch to the face), psychological harm (e.g., verbal insults), and indirect harm (e.g., destroying the victim's property).

Accidental harm is not "aggressive" because it is not intended. Harm that is an incidental by-product of actions taken to achieve some superordinate goal is also excluded from "aggression" because the harm-doer's primary intent in such cases is to help the person achieve the superordinate goal and because the harm-recipient doesn't actively attempt to avoid the harm-doer's action. For example, pain delivered during a dental procedure is not "aggression" by the dentist against the patient.

In their scientific usages "aggressiveness" is very different from "assertiveness" even though the general public frequently uses these words interchangeably. When people say that someone is an "aggressive" salesperson they typically mean that he or she is assertive—pushy or confident or emphatic or persistent—but they do not truly mean "aggressive" unless, of course, they believe that the salesperson intentionally tries to harm customers. Similarly, coaches exhorting players to "be more aggressive" seldom mean that players should try to harm their opponents; rather, coaches want players to be more assertive—active and confident.

Violence, on the other hand, is a subtype of aggression. The term "violence" is generally used to denote extreme forms of aggression such as murder, rape, and assault. All violence is aggression, but many instances of aggression are not violent. For example, one child pushing another off a tricycle is considered aggressive but not violent. For example, one child pushing another off a tricycle is considered aggressive but not violent.

*Affective vs. Instrumental Types of Aggression.* "Affective" aggression has the primary motive of harming the target, and is thought to be based on



anger. It is sometimes labeled hostile, impulsive, or reactive aggression, though these labels often carry additional meaning. When aggression is merely a tool to achieve another goal of the aggressor, it is labeled “instrumental” aggression. Most robberies are primarily instrumental, whereas most murders and assaults are affective. Similarly, Jack may hit Jim merely to obtain a desirable toy, a case of instrumental aggression. Jim may get angry and respond by hitting Jack in order to hurt him, a case of affective aggression.

*Proactive vs. Reactive Types of Aggression.* “Proactive” aggression occurs in the absence of provocation. It is usually instrumental, as when Jack hit Jim to get the toy. “Reactive” aggression is a response to a prior provocation, such when Jim retaliated. There is an asymmetrical relation between proactive and reactive aggression. Children who are high on proactive aggression usually are high on reactive aggression as well, but many children who are high on reactive aggression engage in little proactive aggression.

*Thoughtful vs. Thoughtless Aggression.* A more recent distinction among types of aggression concerns whether the aggressive act resulted from thoughtful or thoughtless (impulsive) psychological processes. In past work, instrumental aggression has usually been seen as thoughtful, involving the careful weighing of potential costs and benefits. But more recent work reveals that frequent use of aggression to obtain valued goals can become so automatized that it also becomes thoughtless. Affective aggression has usually been seen as thoughtless, but people sometimes consider various possible courses of action and decide that an angry outburst is the best way to achieve those goals. This distinction between thoughtful and thoughtless aggression has important implications for the development of and intervention in aggression.

Distinguishing among types of aggression is difficult because underlying motives and psychological processes must be inferred. Is Jim’s angry attack on Jack purely anger-based, solely intended to harm Jack, or is there also some instrumental component? There is a growing realization that these ideal types of aggression rarely exist in pure form in the real world of human interaction. Indeed, a few scholars have argued that all aggression is instrumental, serving goals such as social control, public-image management, private-image

management (i.e., self-esteem), and social justice. Nonetheless, most aggression scholars still find these distinctions helpful for theoretical, rhetorical, and application-oriented reasons.

### WHAT CAUSES AGGRESSION?

The causes of aggression can be analyzed at two different levels: the proximal causes (in the immediate situation) and the more distal causes that set the stage for the emergence and operation of proximate causes.

**Distal Causes: Biological Factors.** Distal causes of aggression are those that make people ready and capable of aggression. Some are structural, built into the human species. Others are developmental, based on the particular environmental history of the individual, and result in individual differences in preparedness to aggress.

*Genetics.* In the broadest sense aggression is a species characteristic. That is, the human species has physical, cognitive, and emotional systems capable of intentionally inflicting harm on other humans. The genetic basis of aggression is easier to identify in nonhuman species, in which fighting behaviors can be produced by stimulating certain regions of the limbic system. Similar physiological systems exist in humans, but human behavior is much more complexly determined.

In the more usual sense genetic influences refer to individual differences in aggressiveness that are linked to genetic differences within the species. Human twin studies have yielded mixed results in estimates of the genetic contribution to human aggression. Miles and Carey (1997) did a meta-analysis (i.e., statistical review) on twenty-four “genetically informative” studies. Two important conclusions were: (1) up to 50 percent of variation in self- or parent-reported aggression was attributable to genetic effects; and (2) when aggressiveness was measured by careful observation of laboratory behaviors, the genetic effect disappeared and a strong family-environment effect emerged. These contradictory findings highlight the complexity of human aggression as well as the need for additional studies.

*Mechanisms.* Several biological mechanisms appear plausible as potential causes of individual differences in aggressiveness. Hormones (e.g.,

testosterone), neurochemicals (e.g., serotonin), attention deficit hyperactivity disorder, and general levels of arousal have all been linked to aggression. For example, Eysenck and Gudjonsson (1989) proposed that individuals whose nervous system is relatively insensitive to low levels of environmental stimulation seek out high-risk activities, including criminal ones, to increase their arousal.

But many biological effects on aggression are neither as strong nor as consistent as the general public believes. For example, testosterone is frequently cited as the explanation for male/female differences in violence rates, but the human literature on testosterone effects is far from clear. Testosterone levels in humans seems more closely linked to social dominance, which in turn may well influence aggression under some limited circumstances (Campbell, Muncer, and Odber 1997; Geary 1998).

Other psychological variables with links to aggression also appear to have some genetic basis. Empathy, behavioral inhibition, negative affectivity, extraversion, neuroticism, and psychoticism all have yielded evidence of some genetic heritability, and have obvious links to aggression. General intelligence may also link biological variation to aggressiveness; low intelligence increases the occurrence of frustrating failures and aversive conditions, which might increase the likelihood of a person developing an aggressive personality.

**Distal Causes: Environmental and Psychological Factors.** Numerous social, environmental, and psychological factors contribute to the development of habitual aggressiveness. Learning stands out as the most important factor of all.

*Learning.* Bandura's social-learning theory of aggression (1973) has been most influential. One key idea in this and all modern learning approaches is that much of human development is based on learning by observing how other people behave. Patterson, DeBaryshe, and Ramsey (1989) presented a detailed look at the maladaptive social-learning processes found in families of aggressive children. Among the key problems are parental use of poor disciplinary measures and inadequate monitoring of their children's activities. Similarly, Olweus (1995) has identified a number of child-rearing factors that are conducive to creating bullies: caretakers with indifferent attitudes toward the child; permissiveness for aggressive behavior

by the child; and the use of physical punishment and other power-assertive disciplinary techniques.

Cognitive psychology has also been crucial in the present understanding of the aggressive personality, as can be seen in books by Berkowitz (1993) and Geen (1990), and in Huesmann's (1998) information-processing theory of aggressive personality development. In brief, humans begin learning from infancy how to perceive, interpret, judge, and respond to events in the physical and social environment. We learn perceptual schemata that help us decide what to look for and what we "see." We learn rules for how the social world works. We learn behavioral scripts and use them to interpret events and actions of others and to guide our own behavioral responses to those events. These various *knowledge structures* develop over time. They are based on the day-to-day observations of and interactions with other people: real (as in the family) and imagined (as in the mass media). For example, the long-term exposure to media violence can increase later aggressive behavior by influencing a variety of aggression-related knowledge structures. Such long-term media violence effects have been shown to be substantial in size and long lasting in duration (Huesmann and Miller 1994).

As knowledge structures develop, they become more complex, interconnected, and difficult to change. Developing knowledge structures are like slowly hardening clay. Environmental experiences shape the clay. Changes are relatively easy to make at first, when the clay is soft, but later on changes become increasingly difficult. Longitudinal studies suggest that aggression-related knowledge structures begin to harden around age eight or nine, and become more perseverant with increasing age.

People learn specific aggressive behaviors, the likely outcome of such behaviors, and how and when to apply these behaviors. They learn hostile perception, attribution, and expectation biases, callous attitudes, and how to disengage or ignore normal empathic reactions that might serve as aggression inhibitors.

The pervasiveness, interconnectedness, and accessibility of any learned knowledge structure is largely determined by the frequency with which it is encountered, imagined, and used. With great

frequency even complex perception-judgment-behavior knowledge structures can become *automatized*—so overlearned that they are applied automatically with little effort or awareness. Frequent exposure to aggressive models is particularly effective in creating habitually aggressive people, whether those models are in the home, neighborhood, or mass media. Once the use of any particular knowledge structure has become automatized, it becomes very difficult for the person to avoid using it because the perceptions and behavioral impulses it produces seem to be based on “how the world really is.”

*Social Processes.* Several common social processes contribute to disproportionate exposure to and learning of aggression-related knowledge structures. Low intellect (social or academic) creates excessive failures and frustration in a variety of developmental contexts. Low social intelligence, for example, leads to problems in interpersonal interactions, whereas low academic intelligence creates problems in school settings. Problems in either context typically lead to higher-than-normal levels of aggression, which lead to further frustrating encounters with parents, teachers, and peers. The resulting social ostracism often forces children to spend more time with other social misfits who also have highly aggressive behavior patterns. This “gang” can impede further intellectual development and reward additional antisocial tendencies.

*Environments.* Many social environments foster the development of an aggressive personality. Such factors include poverty; living in violent neighborhoods; deviant peers; lack of safe, supervised child recreational areas; exposure to media violence; bad parenting; and lack of social support. Growing up in a culture of fear and hate, as in many ethnic-minority communities around the world, may well be the most extreme version of an aggressive-personality-fostering environment, and may well account for the generation after generation of ethnic and religious hatreds and genocidal tendencies that occasionally erupt into genocidal wars (Keltner and Robinson 1996; Staub 1989, 1998). The perceptual knowledge structures modeled and explicitly taught in these contexts guarantee continued mistrust, misunderstanding, and hatred of key outgroups.

Even in its simplest form, poverty is associated with more frustrations, bad role models, and lack

of good role models. Bad parenting includes several particularly common and damaging factors such as lack of parental attention, inconsistent discipline, harsh and abusive discipline, and inattention to nonaggressive efforts at problem solving by the child. Privation, victimization, and violence in a social milieu of long-standing ethnic/religious conflicts provide a powerful learning environment that is mightily resistant to change.

Short-term impoverishment, such as that brought on by a general decline in economic activity (e.g., a recession or depression), has been proposed as a causal factor in aggression directed against ethnic minorities. The dominant model is that the frustration engendered by such economic downturns leads to increased aggression against relatively powerless target groups. However, research casts considerable doubt on this hypothesis. For example, Green, Glaser, and Rich (1998) reanalyzed data on lynchings and data on “gay-bashing,” and showed no evidence of short-term fluctuations in economic conditions and violence directed at minorities.

*Child abuse and neglect.* Child abuse and neglect itself are self-perpetuating problems. Abused or neglected children are particularly likely to become abusing and neglecting parents and violent criminal offenders. Children learn maladaptive beliefs, attitudes, and values from their abusive or neglectful parents (Azar and Rohrbeck 1986; Peterson, Gable, Doyle, and Ewugman 1997).

**Proximate Causes: Individual Differences.** The distal causes described in earlier sections set the stage for human aggression of various types. Proximate causes are those that are present in the current situation. One type of proximate cause consists of individual differences between people that have been created by their biological and social pasts. People differ widely in readiness for aggressing. These differences show considerable consistency across time and situations (Huesmann and Moise 1998).

*Hostility Biases.* Hostility biases have been identified in aggressive adults and children, some as young as six years. The *hostile perception bias* is the tendency of aggression-prone people to perceive social behaviors as more aggressive than do normal people, whereas the *hostile expectation bias* is the tendency of aggression-prone people to expect

and predict others to behave relatively more aggressively (Dill, Anderson, Anderson, and Deuser 1997). The more widely studied *hostile attribution bias* is the tendency of aggression-prone people to attribute hostile intent to others' accidentally harmful behaviors. For example, Dodge (1980) had aggressive and nonaggressive children listen to a story about a boy who hurt another boy by hitting him with a ball. When asked, aggressive children attributed more hostile intent to the boy who threw the ball than did nonaggressive children.

*Attitudes and Beliefs.* Aggression-prone people hold favorable attitudes toward aggression, believing that aggressive solutions to problems are effective and appropriate. Aggressive thoughts and aggressive solutions come to mind quickly and easily. However, creating nonaggressive alternatives is particularly difficult for the aggressive person.

For example, Malamuth, Linz, Heavey, Barnes, and Acker (1995) found that sexually aggressive males hold relatively positive attitudes toward the use of aggression against women, believe in numerous rape myths, engage in more impersonal sex, and are likely to aggress against women in nonsexual contexts as well. Research (Anderson and Anderson 1999) reveals that sexually aggressive men are specifically aggressive only against women, in both sexual and nonsexual contexts, but are not unusually aggressive against other men.

*Narcissism and Self-Esteem.* The predominant view of the link between self-esteem and violence has been that low self-esteem contributes to high violence. However, research from several perspectives has demonstrated a very different pattern. Certain individuals with high self-esteem are most prone to anger and are most aggressive when their high self-image is threatened. Specifically, it is high self-esteem people who react most violently to threats to their self-esteem—if their high self-esteem is inflated (undeserved), unstable, or tentative. In other words, narcissists are the dangerous people, not those with low self-esteem or those who are confident in their high self-image (Baumeister, Smart, and Boden 1996; Bushman and Baumeister 1998; Kernis, Grannemann, and Barclay 1989).

*Sex.* Males and females differ in aggressive tendencies, especially in the most violent behaviors of homicide and aggravated assault. The ratio of male to female murderers in the United States is

almost 10:1. Laboratory studies show the same type of sex effect, but provocation has a greater effect on aggression than does sex. Bettencourt and Miller (1996) used meta-analytic procedures and found that sex differences in aggression practically disappear under high provocation.

Men and women also appear to differ in what provokes them. Bettencourt and Miller showed that males are particularly sensitive to negative intelligence provocations whereas females are particularly sensitive to insults by a peer and to physical attacks. Geary, Rumsey, Bow-Thomas, and Hoard (1995) showed that males are more upset by sexual infidelity of their mates than by emotional infidelity, whereas the opposite pattern occurs for females. Buss and Shackelford (1997) showed similar sex differences in the effects of infidelity on mate-retention tactics, including use of violence.

*Biology.* Other biological differences that people bring with them to the current situation may also contribute to aggression, but as noted earlier many biological effects on aggression are neither as strong nor as consistent as the general public believes. For example, testosterone is frequently cited as the explanation for male/female differences in violence rates, but the human literature on testosterone effects is mixed.

**Proximate Causes: Situational Factors.** The second type of proximate causes of aggression consists of the situational factors currently present. Some of these factors are so powerful that even normally nonaggressive individuals can be made to behave aggressively.

*Provocation.* Most aggressive incidents can be directly linked to some type of perceived provocation. Some are direct and obvious, such as verbal insults and physical assaults. Some are less direct, as when an expected pay raise fails to materialize. Most murders and assaults in normal (i.e., nonwar) contexts are the result of provocations of one kind or another, usually in a series of escalatory provocations, threats, and counterthreats. Federal Bureau of Investigation data reveal that most murders in the United States occur during arguments among family, friends, or acquaintances. The tendency for stranger-based homicides to be relatively rare is even more pronounced in other industrialized cultures than in the United States. Frequently, the provocations involve sexual or emotional infidelity, or perceived insults to one's honor.

*Frustration.* Frustration is both an event and an emotional reaction. It occurs when something blocks the attainment or threatens the continued possession of a valued goal objective. For example, a supervisor's bad report may prevent a promotion, a spouse's infidelity may threaten the continued existence of a marriage, or a flood may destroy one's home. If the frustrating agent is another person, then the frustrating event is also a provocation.

The original form of the frustration-aggression hypothesis by Dollard, Doob, Miller, Mowrer, and Sears (1939) stated that: (1) all acts of aggression are the result of previous frustration; (2) all frustration leads to aggression. But some frustrations do not yield aggression, and some aggression is not the result of a prior frustration. Indeed, many contemporary scholars believe that if a frustrating event is fully justified, the frustrated person would show no residual inclination to aggress. However, Berkowitz (1989) claimed that even fully justified frustration can produce aggressive tendencies. This prediction was recently confirmed by Dill and Anderson (1995).

In a similar vein, Miller and Marcus-Newhall (1997) have shown that provocations can lead to increased aggressive tendencies against individuals who were not part of the frustrating event at all, a phenomenon typically labeled *displaced aggression*. Miller and Marcus-Newhall also suggest that such displaced aggression is increased if the displacement target provides a minor "triggering" provocation, and if the displacement target is a member of a disliked outgroup.

*Incentives.* Incentives are the rewards or benefits a person expects for having performed a particular action. Many situations in politics, the business world, and sports encourage aggression by their incentives. People often expect their chances of winning an election, getting a contract, or defeating an opponent to be enhanced by harming their competitor. Research on television violence has shown that seeing a character rewarded (or not punished) for aggressing increases subsequent aggression by the viewer more so than does unrewarded (or punished) television violence, presumably by increasing the perceived incentive value of aggressive behavior.

The prototypical incentive-based example of individual aggression is the contract killer, who

murders purely for money. The Iraqi assault and takeover of Kuwait, as well as NATO's subsequent attack on Iraq are clear examples of incentive-based institutional aggression (though other factors also clearly played a role). Contract murders account for only a small percentage of homicide totals, but they nicely illustrate the concept of relatively anger-free instrumental aggression.

*Aversive Stimulation and Stress.* Almost any form of aversive stimulation can increase the likelihood of aggression—noise, pain, crowding, cigarette smoke, heat, daily hassles, and interpersonal problems illustrate a few such aversive factors. When the cause of an aversive stimulus is an identifiable person, such as a smoker, these factors are also provocations. As such, they can increase aggression directed at the person identified as the provocateur, as well as against other "displaced" targets.

In cases where there is no identifiable human agent causing the aversive stimulation the effects on aggression are often less noticeable, but much research demonstrates their reality. The most studied of these effects, with relevant data gathered for over one hundred years, is the heat effect. Anderson and Anderson (1998) showed that a wide array of studies across time, culture, and method converge on the conclusion that hot temperatures increase aggressive tendencies. People who live in hotter cities have higher violent crime rates than those in cooler cities. This effect persists even when controlling for poverty, education, and culture. Violent crime rates are higher during hotter years, seasons, months, and days. When people are hot, they think more aggressive thoughts, feel more hostile, and behave more aggressively.

*Alcohol and Drugs.* Bushman (1993) reviewed studies on alcohol and drug effects on aggression, and found that central nervous system depressants increase aggression. Neither actual alcohol consumption nor the mere belief that one has consumed alcohol were individually sufficient to produce reliable increases in aggression, but when research participants believed they had consumed alcohol and had actually consumed alcohol, aggression increased. The exact mechanisms underlying these drug effects are not yet fully understood. Steele and Josephs (1990) proposed an "alcohol myopia" explanation, in which alcohol

impairs key perceptual processes necessary to normal inhibitions against extreme and risky behavior. Bushman's review (1997) confirmed this view.

*Aggression Cues.* Objects or events associated with aggression in semantic memory can cue or "prime" aggression-related thoughts, affects, and behavior programs also stored in memory. For instance, seeing a gun can prime aggressive thoughts (Anderson, Benjamin, and Bartholow 1998) and increase aggressive behavior. This phenomenon, labeled the "weapons effect" by Berkowitz and LePage (1967), has been found in field and laboratory studies, in several different countries, with pictures of weapons and with real weapons.

As mentioned earlier, one prevalent source of aggressive cues in modern society is the mass media. Television shows, movies, and video games are filled with violence. Over 1,000 empirical comparisons, compiled by Paik and Comstock (1994) have conclusively demonstrated that even short-term exposure to media violence increases aggression. The immediate impact of viewing violent media is more pronounced for people with strong aggressive tendencies (Bushman 1995). Unfortunately, aggressive people also are the most likely to seek out violent media.

Many people in modern society believe that viewing aggression (e.g., on television) or behaving in a mildly aggressive way within protected environments (e.g., playing football) will reduce later aggressive behavior. This catharsis hypothesis, though, has been thoroughly debunked (Bushman, Baumeister, and Stack in press; Geen and Quanty 1977).

*Opportunity.* Some situations restrict opportunities to aggress; others provide "good" opportunities. Church service situations have many impediments to aggression—there are witnesses, strong social norms against aggression, and specific nonaggressive behavioral roles for everyone in attendance. Country and Western bars on Saturday nights present better opportunities for aggression, because many aggression facilitators are present: alcohol, aggression cues, aggression-prone individuals, males competing for the attention of females, and relative anonymity.

**Removal of Self-Regulatory Inhibitions.** One often-neglected facet of human aggression has

garnered increased attention; the aggression inhibitions that normally operate in most people. Several different research groups have independently identified and discussed how these inhibitions are sometimes overridden (Bandura, Barbaranelli, Caprara, and Pastorelli 1996; Keltner and Robinson 1996; Staub 1989, 1998). Most people do not commit extreme acts of violence even if they could do so with little chance of discovery or punishment. Such self-regulation is due, in large part, to the fact that people cannot easily escape the consequences that they apply to themselves. Self-image, self-standards, and sense of self-worth—in other words moral standards—are used in normal self-regulation of behavior.

However, people with apparently normal moral standards sometimes behave reprehensibly toward others, including committing such actions as murder, torture, even genocide. Two particularly important mechanisms that allow people to disengage their normal moral standards involve moral justification and dehumanizing the victim. Common justifications for extreme and mass violence include "it is for the person's own good," or the good of the society, or that personal honor demands the violent action. These justifications can be applied at multiple levels, from a parent's abuse of a child to genocidal war. Dehumanizing the victim operates by making sure that one's moral standards are simply not applicable. War propaganda obviously fits this mechanism, but people also use this mechanism at an individual level. Potential victims are placed in the ultimate outgroup—one that has no human qualities.

**The Escalation Cycle.** Many proximate causal factors seem too trivial or weak to contribute to serious aggression. How can seeing a weapon, being uncomfortably hot, or watching a violent movie increase murder rates? The answer lies in the escalation cycle. As noted earlier, assaults and homicides do not typically result from one brief encounter or provocation. The parties involved usually know each other and have had a series of unpleasant exchanges. The final encounter may well begin as a relatively minor dispute, but one person escalates the level of aggression. The other person responds in kind and subsequently increases the aggressiveness of the next response. A shouting match can quickly become a shoving match, which can lead to fists, guns, and death. Seemingly trivial factors increase the likelihood of violence by

increasing the accessibility of aggressive thoughts, affect, and behavioral acts at each turn of the escalation cycle.

### INTERVENTION: PREVENTION AND TREATMENT

The knowledge structure approach explains the difficulty of rehabilitating adults who repeatedly commit violent crimes, or of changing the genocidal climate of groups that have long histories of hate and violence. At the individual level, a lifetime of developing aggressive behavior scripts and automatized hostile perception, expectation, and attribution biases cannot be unlearned easily. However, this approach also reveals that preventing the development of an aggressive or genocidal personality is a more reasonable goal if appropriate steps are taken prior to full maturation.

**Preventing and Treating Aggressive Personality.** There are three main loci for preventing a child from developing into an aggressive adult. First, one can reduce exposure to events that teach aggressive behaviors or scripts. This would include direct modeling (e.g., by abusive or violent parents) as well as indirect modeling (e.g., exposure to media violence). Second, one can reduce exposure to events that teach that aggression is rewarding. For example, most media violence is highly rewarding for the perpetrator, especially when it is the protagonist who is committing the violence. Similarly, adult violence against children (e.g., by parents or school officials) appears highly rewarding to the child because the adult “wins” the encounter and there are no obvious costs to the adult for harming the child. Third, one can reduce exposure to events that teach hostile perception, expectation, and attribution biases. Once again, the entertainment media is one source of violence exposure that increases the perception that the world is a dangerous place. A heavy dose of media violence (e.g., television, movies, video games, music) can increase all three hostility biases. Witnessing high levels of violence in one’s neighborhood also increases these biases.

At all three loci, reducing exposure to aggression-enhancing factors would seem much easier to do in the context of a normal and relatively non-violent culture than in the context of a genocidal culture. Though the following statements focus on

dealing with the aggressive personality, the general principles apply to dealing with the genocidal personality.

Furthermore, treating people who have already developed a strong and stable aggressive personality is much more difficult than preventing the development of such a personality. People with aggressive personalities must learn new nonhostile knowledge structures ranging from perceptual schemata through attributional ones to behavioral scripts. The knowledge structure approach outlined earlier explains why it is easiest to intervene successfully in younger children whose personalities are still malleable, harder to succeed with violent juvenile offenders and young abusive parents, and hardest of all to succeed with habitually violent adult criminals.

*Child Abuse: Treatment and Prevention.* Early intervention attempts relied primarily on intensive dynamic psychotherapy with the abuser, but this approach has repeatedly failed. Cognitive behavioral interventions have had much greater success, largely because they deal directly with the knowledge structure issues that are so important in this domain (Wolf 1994). This approach succeeds by teaching abusive caregivers to use nonaggressive child compliance techniques, personal anger control, and developmentally appropriate beliefs about childhood abilities.

*Reducing Exposure to Aggressive Social Models.* Reducing children’s exposure to aggressive social models would reduce the percentage who grow up believing in and using aggressive tactics. One way of doing this is to reduce exposure to violent media, especially television and video games. The research literature on television violence has conclusively demonstrated that early and repeated exposure to violent television causes children to develop into aggressive adults. For example, kids who watch a lot of violent television at age eight are more likely to have criminal records at age thirty, even after statistically controlling for a variety of other relevant social variables. Research has suggested that exposure to violent video games has a similar effect.

Reducing other types of exposure to violent social models would also help. Reducing parental violence towards children, reducing the frequency

and visibility of violence in children's neighborhoods, reducing violence in schools—including violence by school authorities in attempts to control children—would all have a positive impact on the overall level of aggressiveness in society.

*Treating Violent Juvenile Offenders.* Many treatments have been tried with violent juvenile offenders, including such things as "boot camps," individual therapy, "scared straight" programs, and group therapy; there is little evidence of sustained success for any of these approaches. One problem is that these standard approaches do not address the wide range of factors that contribute to the development and maintenance of violent behavior. However, there is evidence that treatment can have a significant beneficial impact on violent juvenile offenders (e.g., Simon 1998). Tate, Reppucci, and Mulvey (1995) drew attention to one approach with impressive results—the Multisystemic Therapy developed by Henggeler and Borduin (e.g., Henggeler, Schoenwald, Borduin, Rowland, and Cunningham 1998). Multisystemic Therapy is a family-based approach that first identifies the major factors contributing to the delinquent and violent behaviors of the particular individual undergoing treatment. Biological, school, work, peers, family, and neighborhood factors are examined. Intervention is then tailored to fit the individual constellation of contributing factors. Opportunities to observe and commit further violent and criminal offenses are severely restricted, whereas prosocial behavior opportunities (including studying school subjects, developing hobbies) are greatly enhanced, and are rewarded. Both the long-term success rate and the cost/benefit ratio of this approach have greatly exceeded other attempts at treating this population.

*Adults.* Attempts at treatment or "rehabilitation" of violent adults, usually done in the context of prison programs, have led to a general consensus of failure. However, several studies have yielded some evidence of a positive effect of treatment on the behavior of violent adults (e.g., Simon 1998). Rice (1997) reported that an intensive program for violent offenders cut recidivism rates in half for nonpsychopathic offenders. Unfortunately, the recidivism rate for psychopathic offenders was significantly increased by this particular treatment program.

## MAKING MODERN SOCIETIES LESS VIOLENT

Several controversial suggestions for social change emerge from the past forty years of research on human aggression. These suggestions, designed to decrease aggression and violence levels generally rather than to treat already-violent individuals, are controversial for political rather than scientific reasons. Research results clearly support each of them.

1. Reduce exposure to media violence and other aggressive role models, especially for children and adolescents.
2. Replace the use of corporal punishment with more positive child-control techniques.
3. Reduce social rewards for aggressive activities, including those previously thought to be cathartic.
4. Increase social rewards and social support for nonaggressive prosocial activities (e.g., learning in school) while making success at such activities possible (e.g., reducing class sizes).
5. Increase the quality of prenatal and postnatal care, to decrease the proportion of the population suffering from developmental difficulties that interfere with normal learning and socialization processes (Anderson in press).
6. Increase the quality of parenting, by providing instruction, social support, and economic support.

### REFERENCES

- Anderson, C. A. (in press) "Aggression and Violence." *The Encyclopedia of Psychology*. Washington, D.C.: American Psychological Association.
- , A. J. Benjamin, and B. D. Bartholow 1998 "Does the Gun Pull the Trigger? Automatic Priming Effects of Weapon Pictures and Weapon Names." *Psychological Science* 9:308–314.
- Anderson, C. A., and K. B. Anderson 1998 "Temperature and Aggression: Paradox, Controversy, and a (Fairly) Clear Picture." In R. Geen and E. Donnerstein, eds., *Human Aggression: Theories, Research and Implications for Social Policy*. New York: Academic Press.



- Anderson, K. B., and C. A. Anderson 1999 *Men who Target Women: Specificity of Target, Generality of Aggressive Behavior*. Unpublished manuscript.
- Azar, S. T., and C. A. Rohrbeck 1986 "Child Abuse and Unrealistic Expectations: Further Validation of the Parent Opinion Questionnaire." *Journal of Consulting and Clinical Psychology* 54:867-868.
- Bandura, A. 1973 *Aggression: A Social Learning Analysis*. Englewood Cliffs, N.J.: Prentice-Hall.
- , C. Barbaranelli, G. V. Caprara, and C. Pastorelli 1996 "Mechanisms of Moral Disengagement in the Exercise of Moral Agency." *Journal of Personality and Social Psychology* 71:364-374.
- Baron, R. A., and D. R. Richardson 1994 *Human Aggression*, 2nd ed. New York: Plenum Press.
- Baumeister, R. F., L. Smart, and J. M. Boden 1996 "Relation of Threatened Egotism to Violence and Aggression: The Dark Side of High Self-Esteem." *Psychological Review* 103:5-33.
- Berkowitz, L. 1989 "Frustration-Aggression Hypothesis: Examination and Reformulation." *Psychological Bulletin* 106:59-73.
- 1993 *Aggression: Its Causes, Consequences, and Control*. New York: McGraw-Hill.
- , and A. LePage 1967 "Weapons as Aggression-Eliciting Stimuli." *Journal of Personality and Social Psychology* 7:202-207.
- Bettencourt, B. A., and N. Miller 1996 "Gender Differences in Aggression as a Function of Provocation: A Meta-Analysis." *Psychological Bulletin* 119:422-447.
- Bushman, B. J. 1993 "Human Aggression While Under the Influence of Alcohol and Other Drugs: An Integrative Research Review." *Current Directions in Psychological Science* 2:148-152.
- 1995 "Moderating Role of Trait Aggressiveness in the Effects of Violent Media on Aggression." *Journal of Personality and Social Psychology* 9:950-960.
- 1997 "Effects of Alcohol on Human Aggression: Validity of Proposed Explanations." In M. Galanter, et al., eds., *Recent Developments in Alcoholism: Alcohol and Violence*, vol. XIII. New York: Plenum Press.
- , and R. F. Baumeister 1998 "Threatened Egotism, Narcissism, Self-Esteem, and Direct and Displaced Aggression: Does Self-Love or Self-Hate Lead to Violence?" *Journal of Personality and Social Psychology* 75:219-229.
- , and A. D. Stack 1999 "Catharsis, Aggression, and Persuasive Influence: Self-Fulfilling or Self-Defeating Prophecies?" *Journal of Personality and Social Psychology* 76:367-376.
- Buss, D. M., and T. K. Shackelford 1997 "From Vigilance to Violence: Mate Retention Tactics in Married Couples." *Journal of Personality and Social Psychology* 72:346-361.
- Campbell, A., S. Muncer, and J. Odber 1997 "Aggression and Testosterone: Testing a Bio-Social Model." *Aggressive Behavior* 23:229-238.
- Dill, K. E., C. A. Anderson, K. B. Anderson, and W. E. Deuser 1997 "Effects of Aggressive Personality on Social Expectations and Social Perceptions." *Journal of Research in Personality* 31:272-292.
- Dodge, K. A. 1980 "Social Cognition and Children's Aggressive Behavior." *Child Development* 51:162-170.
- Dollard, J., L. Doob, N. Miller, O. H. Mowrer, and R. R. Sears 1939 *Frustration and Aggression*. New Haven, Conn.: Yale University Press.
- Eysenck, H. J., and G. Gudjonsson 1989 *The Causes and Cures of Criminality*. New York: Plenum Press.
- Geary, D. C. 1998 *Male, Female: The Evolution of Human Sex Differences*. Washington, D.C.: American Psychological Association.
- , M. Rumsey, C. C. Bow-Thomas, and M. K. Hoard 1995 "Sexual Jealousy as a Facultative Trait: Evidence from the Pattern of Sex Differences in Adults from China and the United States." *Ethology and Sociobiology* 16:355-383.
- Geen, R. G. 1990 *Human Aggression*. Pacific Grove, Calif.: Brooks/Cole.
- , and E. Donnerstein (eds.) 1998 *Human Aggression: Theories, Research and Implications for Social Policy*. New York: Academic Press.
- Geen, R. G., and M. B. Quanty 1977 "The Catharsis of Aggression: An Evaluation of a Hypothesis." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 10. New York: Academic Press.
- Green, D. P., J. Glaser, and A. Rich 1998 "From Lynching to Gay Bashing: The Elusive Connection between Economic Conditions and Hate Crime." *Journal of Personality and Social Psychology* 75:82-92.
- Goldstein, A. P. 1994 *The Ecology of Aggression*. New York: Plenum.
- Henggeler, S. W., S. K. Schoenwald, C. M. Borduin, M. D. Rowland, and P. B. Cunningham 1998 *Multisystemic Treatment of Antisocial Behavior in Children and Adolescents*. New York: The Guilford Press.
- Huesmann, L. R. 1998 "The Role of Social Information Processing and Cognitive Schema in the Acquisition and Maintenance of Habitual Aggressive Behavior." In R. Geen and E. Donnerstein, eds., *Human Aggression: Theories, Research and Implications for Social Policy*. New York: Academic Press.

- (ed.) 1994 *Aggressive Behavior: Current Perspectives*. New York: Plenum.
- , and L. S. Miller 1994 “Long-Term Effects of Repeated Exposure to Media Violence in Childhood.” In L. R. Huesmann, ed., *Aggressive Behavior: Current Perspectives*. New York: Plenum Press.
- Huesmann, L. R., and J. F. Moise 1998 “The Stability and Continuity of Aggression from Early Childhood to Young Adulthood.” In D. J. Flannery and C. F. Huff, eds., *Youth Violence: Prevention, Intervention, and Social Policy*. Washington, D.C.: American Psychiatric Press.
- Keltner, D., and R. J. Robinson 1996. “Extremism, Power, and the Imagined Basis of Social Conflict.” *Current Directions in Psychological Science* 5:101–105.
- Kernis, M. H., B. D. Grannemann, and L. C. Barclay 1989 “Stability and Level of Self-Esteem as Predictors of Anger Arousal and Hostility.” *Journal of Personality and Social Psychology* 56:1013–1022.
- Malamuth, N. M., D. Linz, C. L. Heavey, G. Barnes, and M. Acker 1995 “Using the Confluence Model of Sexual Aggression to Predict Men’s Conflict with Women: A 10-Year Follow-Up Study.” *Journal of Personality and Social Psychology* 69:353–369.
- Miles, D. R., and G. Carey 1997 “Genetic and Environmental Architecture of Human Aggression.” *Journal of Personality and Social Psychology* 72:207–217.
- Miller, N., and A. Marcus-Newhall 1997 “A Conceptual Analysis of Displaced Aggression.” In R. Ben-Ari and Y. Rich, eds., *Enhancing Education in Heterogeneous Schools: Theory and Application*. Ramat-Gan: Bar-Ilan University Press.
- Olweus, D. 1995 “Bullying or Peer Abuse at School: Facts and Intervention.” *Current Directions in Psychological Science* 4:196–200.
- Paik, H., and G. Comstock 1994 “The Effects of Television Violence on Antisocial Behavior: A Meta-Analysis.” *Communication Research* 21:516–546.
- Patterson, G. R., B. D. DeBaryshe, and E. Ramsey 1989 “A Developmental Perspective on Antisocial Behavior.” *American Psychologist* 44:329–335.
- Peterson, L., S. Gable, C. Doyle, and B. Ewugman 1997 “Beyond Parenting Skills: Battling Barriers and Building Bonds to Prevent Child Abuse and Neglect.” *Cognitive and Behavioral Practice* 4:53–74.
- Rice, M. E. 1997 “Violent Offender Research and Implications for the Criminal Justice System.” *American Psychologist* 52:414–423.
- Rubin, J. Z., D. G. Pruitt, and S. H. Kim 1994 *Social Conflict: Escalation, Stalemate, and Settlement*, 2nd ed. New York: McGraw-Hill.
- Simon, L. M. J. 1998 “Does Criminal Offender Treatment Work?” *Applied & Preventive Psychology* 7:137–159.
- Staub, E. 1989 *The Roots of Evil: The Origins of Genocide and Other Group Violence*. New York: Cambridge University Press.
- 1998 “Breaking the Cycle of Genocidal Violence: Healing and Reconciliation.” In J. Harvey, ed., *Perspectives on Loss: A Sourcebook*. Philadelphia: Taylor and Francis.
- Steele, C. M., and R. A. Josephs 1990 “Alcohol Myopia: Its Prized and Dangerous Effects.” *American Psychologist* 45:921–933.
- Strasburger, V. C. 1995 *Adolescents and the Media: Medical and Psychological Impact*. Thousand Oaks, Calif.: Sage.
- Straus, M. A., and R. J. Gelles 1990 *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*. New Brunswick, N.J.: Transaction.
- Tate, D. C., N. D. Reppucci, and E. P. Mulvey 1995 “Violent Juvenile Delinquents: Treatment Effectiveness and Implications for Future Action.” *American Psychologist* 50:777–781.
- Tedeschi, J. T., and R. B. Felson 1994 *Violence, Aggression, & Coercive Actions*. Washington, D.C.: American Psychological Association.
- Wolfe, D. A. 1994 “The Role of Intervention and Treatment Services in the Prevention of Child Abuse and Neglect.” In G. B. Melton and F. D. Barry, eds., *Protecting Children from Abuse and Neglect: Foundations for a New National Strategy*. New York: Guilford Press.

CRAIG A. ANDERSON

## AGING

See Aging and the Life Course; Cohort Perspectives; Filial Responsibility; Intergenerational Relations; Intergenerational Resource Transfer, Long Term Care, Long Term Care Facilities; Retirement; Widowhood.

## AGING AND THE LIFE COURSE

Social gerontology, or the sociology of aging, has two primary foci: (a) social factors during late life, and (b) social antecedents and consequences of aging. Thus, social gerontology includes examination of both the status of being old and the process

of becoming old. Increasingly, theories and methods of the life course are replacing the earlier emphasis on late life as a separate topic of inquiry. This is a vast arena, and the sociology of aging is appropriately informed by the theories and methods of many sociological subspecialties ranging from macrohistorical and demographic perspectives to the microorientations of social psychology and interpretive sociology.

### HISTORY OF THE FIELD

Historically, social gerontology emerged from a social-problems orientation and focused on the deprivations and losses that were expected to characterize late life (e.g., Burgess 1960; Cain 1959). Early research in the field focused on issues such as poverty during late life; old age as a marginal status, reflecting problems of social integration; the negative effects of institutionalization and poor quality of long-term care; and ageism and age discrimination. Early on, however, some investigators saw the dangers of allowing a crisis orientation to dominate the study of aging and focused attention on patterns of “normal” and “successful” aging (Havighurst 1963; Palmore 1970). A significant proportion of research also focuses on the problems of late life. Investigators remain concerned about social integration and adaptation to loss. The majority of funding for aging research is provided by the National Institute on Aging, which is mandated to support health-related research. Consequently, much aging research focuses on illness and the health care delivery system. The dramatic aging of the population (U.S. Bureau of the Census 1987)—a trend that will peak with the aging of the Baby Boom cohorts—leads to questions about the capacities of social institutions and public policies to meet the needs of an unprecedented number and proportion of older adults. Scholars using political economy theories focus on the ways in which societies respond to the dependency needs of older adults and the social implications of those responses.

Although much research remains focused on the problems of late life, sociologists now recognize the broader importance and implications of old age and aging. Two primary factors appear to have been the driving forces that account for this broader and more complex view. First, despite the social-problems orientation of most early research,

empirical data failed to confirm a uniformly bleak picture of old age. For example, in spite of higher rates of illness and disability, the vast majority of older adults are competent and able to live independent lives (Kunkel and Applebaum 1992). Similarly, rather than representing involuntary loss of a treasured role, retirement is actively sought by the majority of older workers and seldom poses adaptive problems (e.g., Hardy and Quadagno 1995). In addition, some of the problems observed in early studies of older adults have been remedied by the increased resources that recent cohorts have brought to late life, as well as to effective public policies. Thus, although health care costs remain a burden for many older adults, Medicare and Medicaid substantially reduced barriers to health care among older people. Similarly, as a result of improvements in Social Security benefits and increased participation in private pensions, older Americans now are no more likely to live in poverty than younger adults and, indeed, are less likely to live in poverty than children (U.S. Bureau of the Census 1997). Such findings pushed social gerontology toward more complex and empirically defensible perspectives on old age and aging.

Second, sociologists came to recognize that age plays a fundamental role in social structure and social organization. As a parameter of social organization, age affects the allocation of social resources and social roles. Along with sex and race, age is an ascribed status. But age is unique among ascribed statuses in that it changes over time, and movement across age categories results in changing expectations for behavior, changing access to social resources, and changing personal and social responsibilities. The structural quality of age is best articulated in *age stratification theory* (Riley 1987). Age stratification refers to the division of society into meaningful age groups that differ in social value and the allocation of social resources. The concept of age stratification has proven to be useful in a variety of ways. At the broadest level, it reminds us that age is a fundamental parameter of social organization. Age stratification has been particularly useful in highlighting age-related roles and norms. It has a social-psychological facet as well: age consciousness or awareness and identification with members of one's age group (Day 1990). Moreover, the structural and individual facets of age stratification operate in both directions. On the one hand, the structural

component of age stratification allocates roles and resources and assigns differential social value to age strata. Thus, ageism is largely an effect of age stratification. On the other hand, by promoting age consciousness, age stratification sets the stage for age-based public policies and collective efforts by older adults to protect or increase their share of societal resources (e.g., voting and lobbying efforts based on the self-interests of the elderly).

In the 1990s, the sociology of aging focused on change and stability across the life course. Life-course perspectives have enriched aging research in several ways (e.g., Elder 1995; George 1993). First, a life-course approach is attractive because it recognizes that the past is prologue to the future. That is, status and personal well-being in late life depend in large part on events and achievements experienced earlier in the life course. Second, life-course perspectives emphasize relationships across life domains, recognizing that, for example, family events affect and are affected by work and health. Traditionally, sociological research has focused on specific life domains (e.g., the sociology of work, the sociology of the family); life-course perspectives, in contrast, are person-centered rather than domain-centered. Third, life-course perspectives focus on the intersection of history and personal biography. Although the macro-micro schism remains difficult to bridge, life-course research has documented some of the complex ways that historical conditions affect personal lives both contemporaneously and over subsequent decades.

### THE AGE-PERIOD-COHORT PROBLEM

Isolating the effects of age and characterizing the aging process are difficult tasks. Because many factors affect social structure and individual behavior, it is always difficult to isolate the effects of a specific factor. But this task is especially difficult with regard to age, because it is inherently confounded with the effects of two other factors: cohort and period. *Age*, of course, refers to time since birth, and *age effects* refers to patterns resulting from the passage of time or sheer length of life. *Cohort* refers to the group of persons born at approximately the same time (e.g., the 1920 cohort, the 1940–1944 cohort). There are two primary kinds of *cohort effects*. One type results from historical factors. For example, cohorts who lived through the Great Depression or World War II

had different life experiences than cohorts who were not exposed to those historical events. And, as a further complication, the effects of historical events vary depending on the ages of those who experience them. The second type of cohort effect reflects compositional characteristics. For example, large cohorts (such as the Baby Boomers) may face greater competition for social resources than smaller cohorts (e.g., those born during the Great Depression, when fertility rates were low). Both types of cohort effects can have persistent effects on the life course, and, hence, late life. *Period effects* (also called *time of measurement effects*) result from events or situations that happen at a specific time, and tend to affect individuals regardless of age or cohort. For example, faith in government decreased in all Americans (regardless of age or cohort) at the time of the Watergate scandal.

Age, cohort, and period effects are intertwined. If one knows when an individual was born and also knows the time of measurement, simple subtraction provides accurate information about the individual's age. Similarly, if one knows an individual's age and time of measurement, one can easily calculate date of birth or birth cohort. Statistically, there are no easy methods for disentangling age, period, and cohort. In general, however, the most compelling research results are those that are based on examination of multiple cohorts at multiple times of measurement. If the same age patterns are observed across different cohorts measured at different times, those patterns are likely to reflect age effects. If patterns are not similar across cohorts and times of measurement, however, they are likely to reflect cohort or period effects.

The issue of age-period-cohort effects has lost some of its appeal; critics point out that simply knowing, for example, that there is a cohort effect leaves unanswered what it is about those cohorts that generated the observed differences. This is an appropriate criticism; nonetheless, it is immensely helpful in searching for causal explanations to know whether the underlying mechanism is consistent across time and place (an aging effect), had strong contemporaneous effects on persons of all ages (a period effect), or affected only specific cohorts (a cohort effect). Age, cohort, and period effects are all important in aging research. Age effects provide information about human development as it unfolds in social context. Cohort effects permit us to observe the social implications

of shared history and cohort composition. Period effects provide information about the effects of contemporaneous events and situations on social structure and individual behavior. Distinguishing among age, period, and cohort effects also has important implications for the generalization of research results. Age effects are the same or highly similar across time and place; thus, they are generalizable. In contrast, cohort and period effects are, by definition, variable across time. Consequently, generalization is limited.

Examination of age-period-cohort effects requires large data bases in which multiple cohorts are observed on multiple occasions over long periods of time. Because of these stringent requirements, few aging studies focus specifically on disentangling these confounded factors. But recognition of these sources of confounding appropriately temper investigators' generalizations. In addition, this issue has sensitized researchers to the need to examine change over time, with the result that longitudinal studies have become the dominant research design in efforts to characterize the aging process.

No single theme nor easily summarized list of topics does justice to the scope and diversity of research on aging and the life course. However, three major research domains can provide a general sense of the current major avenues of aging and life-course research: aged heterogeneity, life-course dynamics, and life-course trajectories.

### **AGED HETEROGENEITY**

The majority of research on older adults focuses on heterogeneity in that population. In many ways, this is the legacy of the problem orientation upon which initial aging research in the social sciences rested. Even now, most aging research focuses on differences among older adults and the complex configurations of social factors that can account for individual differences during late life. And, although there are many exceptions, a majority of research in this domain focuses on the social factors that explain individual behavior and personal well-being.

The specific topics that are examined in research on aged heterogeneity are extensive. Health and disability are major concerns in this research tradition, including physical illness; mental illness,

especially depression; cognitive status and dementing illness; functional status and disability; and health behaviors. Health service utilization is a corollary emphasis; there are numerous studies on both utilization in general (e.g., doctors visits, hospitalizations) and specialized health care settings and providers (e.g., long-term care facilities, emergency room visits, screening programs and other preventative services, mental health professionals, dentists). There also has been extensive research on the use of social and community services and living arrangements—both of which are strongly driven by health during late life.

Along with health, socioeconomic status has been a high-volume, long-standing focus in research on aged heterogeneity. Economic status, in the form of income and, to a lesser extent, assets has been the primary emphasis of this research, with studies of the antecedents and consequences of retirement ranking as a close second. Pensions have been shown to have strong effects upon the adequacy of postretirement income, which has spurred substantial investigation of the determinants of pension acquisition and value.

From a more social-psychological perspective, there are strong research traditions examining multiple psychosocial states that can be subsumed under the umbrella of "quality of life." Life satisfaction and morale have received paramount attention in this regard. But several dimensions of the self (e.g., self-esteem, locus of control, sense of mastery) also have rich research traditions.

Two common elements of the wide range of research conducted on the topic of aged heterogeneity merit note. First, a common thread in this research is the desire to understand the processes that render some older adults advantaged and others disadvantaged. This is often described as the "applied" character of aging research. However, it also is a focus on stratification in the broadest sense, which is one of the most persistent and cherished issues in mainstream sociology. Research on aging and the life course has contributed much to our understanding of how broadly life chances and life quality are linked to social factors, as well as to the multiple powerful social bases of stratification.

Second, examination of heterogeneity among the elderly has proven to be a strategic site for

testing and refining multiple middle-range theories and concepts that are central to sociological research. A few examples will illustrate this point. Research on health and disability during late life has been enriched by attention to issues of social stress and social support; conversely, the older population has proven to be ideal for testing theoretical propositions about the effects of stress and social support. Similarly, the issue of social integration, which has an honored tradition in sociology, has been reactivated in studies of older adults, who vary widely in number and quality of links to social structure. And issues such as the “feminization of poverty” and the “double/triple jeopardy hypothesis” (i.e., the potential interacting deprivations associated with being female, nonwhite, and old) have highlighted the extent to which socioeconomic stratification rests on ascribed rather than achieved statuses.

### LIFE-COURSE DYNAMICS

While the issue of heterogeneity focuses attention on differences across individuals, life-course dynamics focus on the persisting effects of social factors over time and stages of life; that is, on intraindividual change. At least three types of life-course studies have received substantial attention in aging research.

**The Intersection of History and Personal Biography.** There now is substantial evidence that historical events can permanently alter personal lives. Most research to date has focused on two historical events: the Great Depression and World War II; both historical events have been linked to life circumstances during late life. Elder (1974, 1999) has compellingly documented, for example, that entering the labor market during the Depression had a permanent negative effect on occupation and income, which in turn affected socioeconomic status during late life. In contrast, younger men, who entered the labor force immediately after World War II experienced historically unparalleled occupational opportunities (and, if they were veterans, government-subsidized college educations).

Other studies by Elder and colleagues focused on the life-course consequences of combat experience during World War II (Elder, Shanahan, and Clipp 1994, 1997). Exposure to combat was strongly related to subsequent health problems, not only

immediately after the war, but also during late life. Social factors loom large in these dynamics. First, social factors were strongly predictive of which soldiers were exposed to combat. Second, social resources and deprivations were strong predictors of the onset or avoidance of health problems.

The effects of historical events on the life course appear on two levels. When an historical event is pervasive, one method of observing its effects is via cohort differences. That is, an event with wide-ranging effects will render those who experience it discernibly different from cohorts that come before and after it. At the same time, historical events do not have the same immediate or long-term consequences for all members of a cohort. For example, families were differentially affected by the economic dislocations of the Depression, with some suffering extensive economic deprivation and others experiencing little or no change in their economic fortunes. Thus, not everyone who “lives through” a major historical event will experience serious life-course consequences. One of the tasks of the life-course scholar is to identify the social factors that determine whether or not an historical event alters personal biography.

**The Persisting Effects of Early Traumas and Deprivations.** A second research focus that illustrates the power of life-course perspectives is investigation of the consequences of early traumas and deprivations on well-being in middle and late life. A growing body of research documents the far-reaching effects of severe childhood experiences—including parental divorce and, to a lesser extent, parental death; childhood poverty; and childhood physical and sexual abuse—on the course of adulthood. Most research to date has focused on the implications of these traumas and deprivations on physical and mental health (e.g., Krause 1998; Landerman, George, and Blazer 1991). The robust relationships between events experienced during childhood and health fifty to sixty years later, controlling on contemporaneous risk factors for morbidity, is strong evidence of the power of fateful events to alter the life course.

Again, social factors are strongly implicated in the processes that account for the relationships between childhood traumas and health and well-being during late life. First, the availability of social resources *at the time* of the trauma can dampen its negative effects in both the short and the long

term. For example, financial security and adequate supervision ameliorate most of the negative effects of parental divorce on subsequent socioeconomic achievements and physical and mental health (e.g., Kessler and Magee 1994). Second, two of the primary mechanisms by which childhood traumas generate poor health in later life are socioeconomic achievement and high-quality social relationships (e.g., McLeod 1991). That is, childhood traumas are often associated with lower socioeconomic status and poor-quality relationships during adulthood—both of which are risk factors for physical and mental illness. If, however, individuals who experienced childhood traumas manage to achieve adequate financial resources and supportive social ties during adulthood, their excess risk of illness in middle and late life is reduced substantially.

**The Persisting Effects of Early Life Decisions.** There now is substantial evidence that the decisions that individuals make during early adulthood have important consequences for their life circumstances in late life. Studies of retirement income provide perhaps the best illustration of this research domain. The strongest predictor of retirement income is occupational history. Throughout adulthood, individuals “sort themselves” into jobs that differ not only in income, but also in benefits (i.e., total compensation packages). Of these, the availability and quality of pensions is most important for retirement income. There is strong evidence that the provision of pensions differs not only by occupation, but also by industrial sector (Quadagno 1988). Thus, when individuals make occupational choices—including job changes throughout adulthood—they are inevitably determining, in part, their retirement incomes.

Research on women’s retirement income has broadened our understanding of the life-course consequences of early decisions. Women and men tend to be concentrated in different occupations and different industrial sectors—and those in which women dominate have, on average, lower earnings and lower likelihood of pension coverage (O’Rand 1988). Moreover, family formation decisions strongly affect women’s job histories. Compared to men, women are less likely to work full-time and work fewer total years, largely as a result of parental responsibilities. All of these factors combine to produce substantially lower retirement incomes for women than for men (O’Rand and Landerman 1984).

## LIFE-COURSE TRAJECTORIES AND PERSON-CENTERED RESEARCH

The two research domains described above focus on *interindividual differences* in late life and *intraindividual change* over the life course. A third domain, less developed than the others but exciting in its scope, attempts to examine interindividual differences and intraindividual change simultaneously. At this point, two emerging research traditions illustrate the nature and potential of this approach.

**Life-Course Trajectories.** Trajectories refer to long-term patterns of stability and change. They can be examined at both the aggregate (e.g., the “typical” career, the modal pathway to nursing home placement) and individual levels. Thus, trajectories capture patterns of intraindividual change. Examination of heterogeneity can be pursued in two ways. In the first, the trajectory that best describes the sample or population is constructed. Subsequently, using techniques such as hierarchical linear modeling or growth-curve analysis, investigators can examine the extent to which factors of interest alter the shape of the trajectory. For example, a trajectory of earnings across adulthood can be constructed for a given sample. Investigators can then examine the degree to which factors such as sex, education, and race affect the shape of the earnings trajectory. In the second approach, multiple common trajectories are identified and investigators then determine the characteristics associated with those trajectories. Using this approach, for example, several common trajectories of earnings during adulthood could be identified (e.g., consistently increasing earnings, earnings peaking during mid-life and then decreasing, consistently decreasing earnings, a relatively flat earnings history). Factors such as sex, education, and race could then be examined to determine their association with these distinctive earnings trajectories.

Trajectory-based research, with its focus on both interindividual differences and intraindividual change, is very attractive. The major limitation to the use of this approach is the availability of data, because longitudinal data covering long periods of time are required if one wishes to understand life-course patterns. It should be noted, however, that trajectory-based research can also be useful for studying shorter processes (e.g., patterns of illness

outcome, with relatively short-term trajectories of death, chronicity, recovery, and relapse).

Trajectory-based research relevant to our understanding of middle and late life is gradually accumulating. This is especially true for long-term patterns of health and functioning, providing evidence about both the dynamics of disability during late life (e.g., Maddox and Clark 1992; Verbrugge, Reoma, and Gruber-Baldini 1994) and long-term patterns of stability, improvement, and decline in health over the course of adulthood (e.g., Clipp, Pavalko, and Elder 1992). Important trajectory-based research on pathways to retirement (Elder and Pavalko 1993) and place of death (Merrill and Mor 1993) is also available.

The concept of trajectories has been valuable in theoretical development, as well as empirical inquiry. The theory of *cumulative advantage/disadvantage* has intersected well with studies of life-course trajectories. This theory posits that heterogeneity is greater in late life than earlier in the life course as a result of the accumulation of assets (advantage) or liabilities (disadvantage) over time. The theory of cumulative advantage/disadvantage has been especially useful in understanding socioeconomic heterogeneity in late life, especially differences in total net worth (e.g., Crystal and Shea 1990; O’Rand 1996). However, it can be applied to other sources of heterogeneity in late life as well (e.g., health). Investigators who prefer examination of the multiple distinctive trajectories within a sample would offer a caution to cumulative advantage/disadvantage theory, however. They would note that although trajectories of increasing and decreasing advantage are undoubtedly common, there are likely to be other important trajectories as well—e.g., a trajectory of cumulating advantage that is reversed as a result of a personal (e.g., serious illness) or societal (e.g., severe economic downturn) catastrophic event.

**Person-centered research.** A more recent contribution to the sociological armamentarium for understanding aging and the life course is *person-centered research*. As the label implies, the focus of this emerging research strategy is to analyze “people” rather than “variables.” In practice, this means that members of a sample are first grouped into subsets on the dependent variable of interest. Subsequently, those categories are further disaggregated into groups based on distinctive

pathways or life histories associated with the outcome of interest.

Work by Singer, Ryff, and Magee (1998) provides the richest example of person-centered aging research to date. The dependent variable of interest was mental health. In the first stage of their research, groups of middle-aged women were divided into groups that differed on levels of mental health. Subsequently, the large archive on longitudinal data obtained from these women over the previous three decades was examined to identify distinctive pathways associated with mid-life mental health. For example, the group of women who exhibited high levels of mental health and well-being were further subdivided into two groups that Singer et al. (1999) labeled the healthy and the resilient. Healthy women were those who had life histories that were relatively free of major stressors or traumas, enjoyed adequate or higher levels of social and economic resources, and exhibited stable patterns of robust mental health. Resilient women were those who had achieved robust mental health at mid-life despite earlier evidence of poor mental health and/or histories of stress and/or inadequate social and economic resources. Clearly the life histories of women who were mentally healthy at mid-life varied in important ways. Moreover, the healthy women illustrate the benefits of leading relatively “charmed” lives, while the resilient women help us to understand the circumstances under which “risky” life histories can be turned around to produce health and personal growth.

There are clear similarities between person-centered research and trajectory-based research. Both are based on long-term patterns of change and stability and both are designed to understand both life-course dynamics and heterogeneity in those dynamics. But there also are important differences between the two approaches. In trajectory-based research, the trajectory or pathway itself is the “dependent variable” of interest and the “independent variables” are factors that have the potential to alter the shape(s) of those pathways. In person-centered research, the “dependent variable” is an outcome of interest (e.g., mental health), and trajectories of the “independent variables” are constructed to explain that outcome. Both are currently at the cutting edge of research that attempts to simultaneously examine life-course dynamics and life-course heterogeneity.



Aging and the life course is an important sociological specialty. It provides us with information about aging and being "old," and about the antecedents and consequences of stability and change during late life. The increasing life-course focus of aging research is especially important in that it concentrates sociologists' attention on the dynamics of intraindividual change and on the intersections of social structure, social change, and personal biography. At its best, aging and life-course research effectively link processes and dynamics at the macrohistorical and societal levels with individual attitudes and behaviors. In addition, the life course in general, and late life in particular, provide excellent contexts for testing and, indeed, challenging some commonly held assumptions and hypotheses about the dynamics of social influence. In this way, it offers valuable contributions to the larger sociological enterprise.

## REFERENCES

- Burgess, Ernest W. 1960 "Aging in Western Culture." In E. W. Burgess, ed., *Aging in Western Societies*. Chicago: The University of Chicago Press.
- Cain, Leonard D., Jr. 1959 "The Sociology of Aging: A Trend Report and Bibliography." *Current Sociology* 8:57-133.
- Clipp, Elizabeth C., Eliza K. Pavalko, and Glen H. Elder, Jr. 1992 "Trajectories of Health: In Concept and Empirical Pattern." *Behavior, Health, and Aging* 2:159-177.
- Crystal, Stephen, and Dennis Shea 1990 "Cumulative Advantage, Cumulative Disadvantage, and Inequality among Elderly People." *The Gerontologist* 30:437-443.
- Day, Christine L. 1990 *What Older People Think*. Princeton, N.J.: Princeton University Press.
- Elder, Glen H., Jr. 1999 *Children of the Great Depression* (twenty-fifth anniversary ed.). Boulder, Colo.: Westview Press.
- 1995 "The Life Course Paradigm: Historical, Comparative, and Developmental Perspectives." In P. Moen, G. H. Elder, and K. Luscher, eds., *Examining Lives in Context: Perspectives on the Ecology of Human Development*. Washington, D.C.: American Psychological Association Press.
- 1974 *Children of the Great Depression*. Chicago: University of Chicago Press.
- , and Eliza K. Pavalko 1993 "Work Careers in Men's Later Years: Transitions, Trajectories, and Historical Change." *Journal of Gerontology: Social Sciences* 48:S180-S191.
- Elder, Glen H., Jr., Michael J. Shanahan, and Elizabeth C. Clipp 1997 "Linking Combat and Physical Health: The Legacy of World War II in Men's Lives." *American Journal of Psychiatry* 154:330-336.
- 1994 "When War Comes to Men's Lives: Life-Course Patterns in Family, Work, and Health." *Psychology and Aging* 9:5-16.
- George, Linda K. 1993 "Sociological Perspectives on Life Transitions." *Annual Review of Sociology* 19:353-373.
- Hardy, Melissa A., and Jill Quadagno 1995 "Satisfaction with Early Retirement: Making Choices in the Auto Industry." *Journal of Gerontology: Social Sciences* 50:S217-S228.
- Havighurst, Robert A. 1963 "Successful Aging." In R. Williams, C. Tibbets, and W. Donahue, eds., *Processes of Aging*. New York: Atherton.
- Kessler, Ronald C., and William J. Magee 1994 "Childhood Family Violence and Adult Recurrent Depression." *Journal of Health and Social Behavior* 35:13-27.
- Krause, Neal 1998 "Early Parental Loss, Recent Life Events, and Changes in Health." *Journal of Aging and Health* 10:395-421.
- Kunkel, Susan R., and Robert A. Applebaum 1992 "Estimating the Prevalence of Long-Term Disability for an Aging Society." *Journal of Gerontology: Social Sciences* 47:S253-S260.
- Landerman, Richard, Linda K. George, and Dan G. Blazer 1991 "Adult Vulnerability for Psychiatric Disorders: Interactive Effects of Negative Childhood Experiences and Recent Stress." *Journal of Nervous and Mental Disease* 179:656-663.
- Maddox, George L., and Daniel O. Clark 1992 "Trajectories of Functional Impairment in Later Life." *Journal of Health and Social Behavior* 33:114-125.
- McLeod, Jane D. 1991 "Childhood Parental Loss and Adult Depression." *Journal of Health and Social Behavior* 32:205-220.
- Merrill, Deborah M., and Vincent Mor 1993 "Pathways to Hospital Death Among the Oldest Old." *Journal of Aging and Health* 5:516-535.
- O'Rand, Angela M. 1996 "The Precious and the Precocious: Understanding Cumulative Disadvantage and Cumulative Advantage Over the Life Course." *The Gerontologist* 36:230-238.
- 1988 "Convergence, Institutionalization, and Bifurcation: Gender and the Pension Acquisition Process." *Annual Review of Gerontology and Geriatrics* 8:132-155.
- , and L. Richard Landerman 1984 "Women's and Men's Retirement Income: Early Family Role Effects." *Research on Aging* 6:25-44.

- Palmore, Erdman B. 1970 *Normal Aging: Reports from the Duke Longitudinal Studies, 1955-1969*. Durham, N.C.: Duke University Press.
- Quadagno, Jill 1988 "Women's Access to Pensions and the Structure of Eligibility Rules: Systems of Production and Reproduction." *Sociological Quarterly* 29:541-558.
- Riley, Matilda White 1987 "On the Significance of Age in Sociology." *American Sociological Review* 52:1-14.
- Singer, Burton, Carol D. Ryff, Deborah Carr, and William J. Magee 1998 "Life Histories and Mental Health: A Person-Centered Strategy." In A. Rafferty, ed., *Sociological Methodology, 1988*. Washington, D.C.: American Sociological Association.
- U. S. Bureau of the Census 1997 *Statistical Abstracts of the United States, 1997*. Washington, D. C.: U. S. Government Printing Office.
- U. S. Bureau of the Census 1987 *An Aging World*. Washington, D. C.: U. S. Government Printing Office.
- Verbrugge, Lois M., Joseto M. Reoma, and Ann L. Gruber-Baldini 1994 "Short-Term Dynamics of Disability and Well-Being." *Journal of Health and Social Behavior* 35:97-117.

LINDA K. GEORGE

## AGRICULTURAL INNOVATION

Getting a new idea adopted can be very difficult. This is all the more frustrating when it seems to the proponents of the new idea that it has very obvious advantages. It can be a challenge to try to introduce new ideas in rural areas, particularly in less-developed societies, where people are somewhat set in their ways—ways that have evolved slowly, through trial and error. It's all the more difficult when those introducing new ideas don't understand why people follow traditional practices. Rural sociologists and agricultural extension researchers who have studied the diffusion of agricultural innovations have traditionally been oriented toward speeding up the diffusion process (Rogers 1983). Pro-innovation bias has sometimes led sociologists to forget that "changing people's customs is an even more delicate responsibility than surgery" (Spicer 1952).

Although innovation relies on invention, and although considerable creativity often accompanies the discovery of how to use an invention,

innovation and invention are not the same thing. Innovation does, however, involve more than a change from one well-established way of doing things to another well-established practice. As with all innovations, those in agriculture involve a change that requires significant imagination, break with established ways of doing things, and create new production capacity. Of course, these criteria are not exact, and it is often difficult to tell where one innovation stops and another starts. The easiest way out of this is to rely on potential adopters of an innovation to define ideas that they perceive to be new.

Innovations are not all alike. New ways of doing things may be more or less compatible with prevalent norms and values. Some innovations may be perceived as relatively difficult to use and understand (i.e., complex), while others are a good deal simpler. Some can be experimented with in limited trials that reduce the risks of adoption (i.e., divisible). Innovations also vary in the costs and advantages they offer in both economic and social terms (e.g., prestige, convenience, satisfaction). In the economists' terms, innovation introduces a new production function that changes the set of possibilities which define what can be produced (Schumpeter 1950). Rural sociologists have studied the adoption of such agricultural innovations as specially bred crops (e.g., hybrid corn and high-yield wheat and rice); many kinds of machines (e.g., tractors, harvesters, pumps); chemical and biological fertilizers, pesticides, and insecticides; cropping practices (e.g., soil and water conservation); and techniques related to animal husbandry (e.g., new feeds, disease control, breeding). Often they have relied upon government agencies such as the U.S. Department of Agriculture to tell them what the recommended new practices are.

The diffusion of agricultural innovations is a process whereby new ways of doing things are spread within and between agrarian communities. Newness implies a degree of uncertainty both because there are a variable number of alternatives and because there is usually some range of relative probability of outcomes associated with the actions involved. Rogers (1983) stresses that the diffusion of innovations includes the communication of information, by various means, about these sets of alternative actions and their possible outcomes. Information about innovations may come via impersonal channels, such as the mass

media, or it may pass through social networks. From an individual's point of view, the process of innovation is usually conceived to start with initial awareness of the innovation and how it functions. It ends with adoption or nonadoption. In between these end points is an interactive, iterative process of attitude formation, decision making, and action. The cumulative frequency of adopters over time describes an S-shaped (logistic) curve. The frequency distribution over time is often bellshaped and approximately normal.

Individual innovativeness has been characterized in five ideal-type adopter categories (Rogers 1983). The first 2 to 3 percent to adopt an innovation, the "innovators," are characterized as venturesome. The next 10 to 15 percent, the "early adopters," are characterized as responsible, solid, local opinion leaders. The next 30 to 35 percent are the "early majority," who are seen as being deliberate. They are followed by the "late majority" (30 to 35 percent), who are cautious and skeptical, and innovate under social and economic pressures. Finally, there are the "laggards," who comprise the bottom 15 percent. They are characterized as "traditional," although they are often simply in a precarious economic position. Earlier adopters are likely to have higher social status and better education, and to be upwardly mobile. They tend to have larger farms, more favorable attitudes toward modern business practices (e.g., credit), and more specialized operations. Earlier adopters are also argued to have greater empathy, rationality, and ability to deal with abstractions. They are less fatalistic and dogmatic, and have both positive attitudes toward change and science, and higher achievement motivation and aspirations. Early adopters report more social participation and network connections, particularly to change agents, and greater exposure to both mass media and interpersonal communication networks.

Although Rogers (1983) provides dozens of such generalizations about the characteristics of early and late adopters, he admits that the evidence on many of these propositions is somewhat mixed (Downs and Mohr 1976). Even the frequently researched proposition that those with higher social status and greater resources are likely to innovate earlier and more often has garnered far less than unanimous support (Cancian 1967, 1979; Gartrell 1977). Cancian argues that this is a result of "upper middle-class conservatism," but

subsequent meta-analysis has clearly demonstrated that the relationship between status and innovation is indeed linear (Gartrell and Gartrell 1985; Lewis et al. 1989). If anything, those with very high status or resources show a marked tendency to turn their awareness of innovations into trial at a very high rate (Gartrell and Gartrell 1979).

Ryan and Gross (1942) provide a classic example of diffusion research. Hybrid corn seed, developed by Iowa State and other land-grant university researchers, increased yields 20 percent over those of open-pollinated varieties. Hybrid corn also was more drought-resistant and was better suited to mechanical harvesting. Agricultural extension agents and seed company salesmen promoted it heavily. Its drawback was that it lost its hybrid vigor after only one generation, so farmers could not save the seed from the best-looking plants. (Of course, this was not at all a drawback to the seed companies!)

Based on a retrospective survey of 259 farmers in two small communities, Ryan and Gross found that 10 percent had adopted hybrid corn after five years (by 1933). Between 1933 and 1936 an additional 30 percent adopted, and by the time of the study (1941) only two farmers did not use the hybrid. Early adopters were more cosmopolitan, and had higher social and economic status. The average respondent took nine years to go from first knowledge to adoption, and interpersonal networks and modeling were judged to be critical to adoption. In other cases diffusion time has been much shorter. Beginning in 1944, the average diffusion time for a weed spray (also in Iowa) was between 1.7 years for innovators and 3.1 years for laggards (Rogers 1983, p. 204). Having adopted many innovations, farmers are likely to adopt others more quickly.

Adoption-diffusion research in rural sociology has dominated all research traditions studying innovation. Rural sociology produced 791 (26 percent) of 3,085 studies up to 1981 (Rogers 1983, p. 52). Most of this research relied upon correlational analysis of survey data based on farmers' recall of past behaviors. This kind of study reached its peak in the mid 1960s. By the mid 1970s the farm crisis in the United States and the global depression spurred rural sociologists to begin to reevaluate this tradition. By the 1980s global export markets had shrunk, farm commodity prices had fallen, net

farm incomes had declined, and high interest rates had resulted in poor debt-to-asset ratios. What followed was a massive (50 percent) decapitalization of agriculture, particularly in the Midwest and Great Plains.

Criticisms of adoption-diffusion research include (1) pro-innovation bias; (2) a lack of consideration of all the consequences of innovation; (3) an individual bias; (4) methods problems; (5) American ethnocentric biases; (6) the passing of the dominant modernization-development paradigm. The pro-innovation bias of researchers has led them to ignore the negative consequences of innovation (van Es 1983). Indeed, innovativeness itself is positively valued (Downs and Mohr 1976). The agencies that fund research and the commercial organizations (e.g., seed companies) that support it have strong vested interests in promoting diffusion. Furthermore, successful innovations leave visible traces and can be more easily studied using retrospective social surveys, so researchers are more likely to focus on successful innovations.

Since most researchers are well aware of this problem, it can be addressed by deliberately focusing on unsuccessful innovations, and by studying discontinuance and reinvention. It can also be avoided by the use of prospective research designs, including qualitative comparative case studies, that track potential innovation and innovators' perceptions and experiences. This should facilitate the investigation of noncommercial innovations and should result in a better understanding of the reasons why people and organizations decide to use new ideas. Moreover, these methods will likely lead to a better understanding of the system context in which innovations diffuse.

One of the most strident critiques of the pro-innovation bias of the "land-grant college complex" was voiced by Hightower (1972). Agricultural scientists at Davis, California, worked on the development and diffusion of hard tomatoes and mechanized pickers (Friedland and Barton 1975). They ignored the effects of these innovations on small farms and farm labor, except in the sense that they designed both innovations to solve labor problems expected when the U.S. Congress ended the bracero program through which Mexican workers were brought in to harvest the crops. In the six years after that program ended (1964 to 1970) the mechanical harvester took over the industry. About

thirty-two thousand former hand pickers were out of work. They were replaced by eighteen thousand workers who rode machines and sorted tomatoes. Of the four thousand farmers who produced tomatoes in California in 1962, only six hundred were still in business in 1971. The tomato industry honored the inventor for saving the tomato for California, and consumers got cheaper, harder tomatoes—even if they preferred softer ones.

Several other classic examples of agricultural innovation illustrate problems that result from not fully considering the consequences of innovation (Fliegel and van Es 1983). Until the late 1970s rural sociologists, among others, studiously ignored Walter Goldschmidt's 1940s study (republished in 1978) of the effects of irrigation on two communities in California's San Joaquin Valley. Dinuba had large family farms, and it also had more local business, greater retail sales, and a greater diversity of social, educational, recreational, and cultural organizations. Arrin was surrounded by large industrial corporate farms supported by irrigation. These farms had absentee owners and Mexican labor. This produced a much lower quality of life that was confirmed three decades later (Buttel et al. 1990, p. 147).

The enforced ban on earlier chemical innovations in agriculture by the U.S. Food and Drug Administration provides another interesting example. Chemical innovations such as DDT insecticide, 2,4-D weed spray, and DES cattle feed revolutionized farm production in the 1950s and 1960s. In 1972, DDT was banned because it constituted a health threat (Dunlap 1981), and 2,4-D, DES, and similar products were banned soon afterward. Finally, in 1980 the U.S. Department of Agriculture reversed its policy and began to advise farmers and gardeners to consider alternative, organic methods that used fewer chemicals.

The impact of technical changes in U.S. agriculture, particularly the rapid mechanization begun in the Great Depression, put farmers on the "treadmill of technology" (Cochran 1979; LeVein 1978). Larger farmers who are less risk-averse adopt early, reap an "innovation rent," reduce their per-unit costs, and increase profits. After the innovation spreads to the early majority, aggregate output increases dramatically. Prices then fall disproportionately, since agricultural products have low elasticity of demand. Lower, declining prices

force the late majority to adopt, but they gain little. They have to adopt to stay in business, and some late adopters may be forced out because they cannot compete. This treadmill increases concentration of agricultural production and benefits large farmers, the suppliers of innovations, and consumers. Indeed, it helps to create and to subsidize cheap urban labor. When it comes to environmental practices, however, large farms are not early innovators (Pampel and van Es 1977; Buttel et al. 1990).

The individual bias of adoption-diffusion research is evident in its almost exclusive focus upon individual farmers rather than upon industrial farms or other agribusiness. There is also a tendency to blame the victim if anything goes wrong (Rogers 1983). Change agents are too rarely criticized for providing incomplete or inaccurate information, and governments and corporations are too infrequently criticized for promoting inappropriate or harmful innovations. Empirical surveys of individual farmers also lead to a number of methodological problems. As noted above, if surveys are retrospective, recall relies on fallible memory and renders unsuccessful innovations difficult to study. These surveys are commonly combined with correlational analysis that makes it difficult to address issues of causality. After all, the farmer's attitudes and personality are measured at the time of the interview, and the innovation probably occurred some time before. As we have pointed out, these issues can be addressed by prospective designs that incorporate other methods, such as qualitative case studies and available records data, and focus on the social context of innovation.

Taking into account the social context of innovation involves shifting levels of analysis from individual farmers to the social, economic, and political structures in which they are embedded. Contextual analysis of social structures has evolved in two directions, both of which have been inspired by the adoption-diffusion paradigm. The first considers social structure as a set of social relations among farmers, that is, a social network. Typically, social network structure has been studied from the point of view of individual farmers. For instance, farmers are more likely to innovate if they are connected to others with whom they can discuss new farming ideas (Rogers and Kincaid 1981; Rogers 1983; Warriner and Moul 1992). In

this type of analysis, "connectedness" becomes a variable property of individual farmers that is correlated with their innovativeness. It is much less common to find studies that consider how agricultural innovation is influenced by structural properties of entire networks, such as the presence of subgroups or cliques, although other types of innovation have been studied within complete networks (see, e.g., Rogers and Kincaid 1981 on the diffusion of family planning in Korean villages). Field studies of subcultural differences in orientations to innovation report what amount to network effects, though networks are rarely measured directly. For instance, studies of Amish farmers have revealed that members of this sect restrict certain kinds of social contacts with outsiders in order to preserve their beliefs, which include environmental orientations based on religious beliefs (for a review, see Sommers and Napier 1993). Given the growing importance of social network analysis in contemporary sociology (Wasserman and Faust 1995), and the demonstrated importance of networks of communication and influence in innovation research (Rogers and Kincaid 1981), future studies of agricultural innovation could profitably incorporate network models and data in their research designs.

A second type of social structural analysis has considered how agricultural innovation is influenced by distributions of resources within farming communities. Much of this research has focused on the so-called "Green Revolution." This term refers to the increases in cereal-grain production in the Third World, particularly India, Pakistan, and the Philippines, in the late 1960s, through the use of hybrid seeds and chemical fertilizer. In Indian villages where knowledge of new farming technology and agricultural capital were highly concentrated, the rate at which individual farmers translated their knowledge into trial was higher (Gartrell and Gartrell 1979). Yet overall levels of innovation tended to be lower in such high-inequality villages. Had the primary goal of India's development programs been to maximize the rate at which knowledge of new farming practices is turned into innovation, then these results could have been seen as a vindication of a development strategy that concentrated on well-to-do cultivators and high-inequality villages. Yet, this and other assessments of the Green Revolution in the

1970s suggested that development would likely exacerbate rural inequality; that most of the benefits of innovation would accrue to farmers who were wealthy enough to afford the new inputs (Frankel 1971; Poleman and Freebairn 1973); and that the rural poor would be further marginalized and forced to seek employment in the increasingly capital-intensive industries developing in cities. The Green Revolution, so the thinking went, contained the seeds of civil unrest in the cities—an urban Red Revolution (Sharma and Poleman 1993).

Recent research in agricultural economics paints a more optimistic portrait of the long-run distributional effects of the Green Revolution. Once small farmers were given the necessary infrastructural support, their productivity and incomes increased. Growing rural incomes and the resulting growth in consumption demand stimulates the development of a wide variety of off-farm and noncrop employment opportunities. Through participation in these “second generation” effects of the Green Revolution, the incomes of landless and near-landless households have increased dramatically (Sharma and Poleman 1993). These indirect consequences of agricultural innovation are not limited to the Green Revolution. Innovation in agriculture and the expansion of rural, nonagricultural manufacturing were strongly associated in the development of Western Europe and East Asia as far back as the eighteenth century (Grabowski 1995). We tend to think of the Industrial Revolution as an urban phenomenon in which agricultural surplus labor was transferred from occupations of low productivity in agriculture to those of high productivity in urban manufacturing. Yet in this early phase of industrialization, the flow of people and economic activity went the other way—from town to country. Then, as now, agricultural innovation influenced and was influenced by the growth of rural manufacturing. Expanding commercialization of rural areas fosters innovation by providing many of the inputs needed by agriculture, as well as sources of credit for farm operations and alternative sources of income to buffer the risks associated with innovations. These reciprocal effects have been observed in the experience of China in the 1990s (Islam and Hehui 1994).

Economists have argued that agricultural innovation is induced by changes in the availability

and cost of major factors of production, particularly land and labor (Binswanger and Ruttan 1978). Historical, cross-national studies show that countries differently endowed with land and labor have followed distinct paths of technological change in agriculture. Population pressures on land resources have impelled technological change and development (Boserup 1965, 1981; Binswanger 1986). Rather than focusing on individual farmers’ adoption decisions, research in this tradition has examined variation in innovation by region according to demographic and other conditions.

The social context of innovation also has an important political dimension. Political structures powerfully influence the path of innovation. When family farms were turned over to collective management by the Democratic Republic of Vietnam, local farmers in both the lowland and upland areas were unable to use their own knowledge of farm management (Jamieson et al. 1998). Traditional knowledge passed down over many generations became lost to new generations of farmers. Yet local knowledge systems are often crucial to the successful implementation of modern farming technologies brought in from the outside. Such problems are compounded when the power to make decisions about the course of development is centralized in national agencies, as is the case in Vietnam.

Innovation can also call forth new political structures. India’s Green Revolution became politicized as the terms and prices at which agricultural inputs could be obtained and the price at which agricultural products could be sold were determined by government and its local agencies. Peasant movements arose as a response to concern about access to the new farming technology. The incidence of improved agricultural practices has been associated with the rise of political parties such as the Lok Dal of Uttar Pradesh, which most clearly articulated rural interests (Duncan 1997). By asserting new identities and interests created in the changed circumstances brought about by the Green Revolution, the Lok Dal was able to mobilize across traditional lines of caste and locality.

The social, economic, and political structures of the social context of innovation do not exist in isolation from one another. In any development setting, a contextually informed understanding of

agricultural innovation must consider the relationships among these different types of structures (Jamieson et al. 1998). While it may no longer be as fashionable as it once was, the adoption-diffusion model still has much to offer in such efforts. The model refers implicitly to structural effects of socioeconomic status and communication behavior, though these are conceptualized at an individual level (Black and Reeve 1992). Structural analysis has recently moved more firmly into this interdisciplinary realm, particularly in economics (see "Economic Sociology"). With the appropriate structural tools, rural sociologists could make notable contributions to our understanding of how the social structures of markets influence innovation.

Technological change in agriculture is still vitally important throughout the world and, correctly applied, diffusion research can assist in its investigation. It is important to consider the consequences of technological change as well as the determinants of adoption of innovation. It is critical to apply the model to environmental practices and other "noncommercial" innovations in agriculture. In-depth case studies over time are needed to further our understanding of how and why individuals and agricultural social collectives adopt technological change. Above all, the social, economic, and political contexts of innovation must be studied with the models and methods of modern structural analysis. All this provides a basis for continuing to build on a wealth of research materials.

(SEE ALSO: *Diffusion Theories; Rural Sociology*)

REFERENCES

Binswanger, Hans P. 1986 "Evaluating Research System Performance and Targeting Research in Land-abundant Areas of Sub-Saharan Africa." *World Development* 14:469-475.

—, and Vernon W. Ruttan 1978 *Induced Innovation: Technology, Institutions, and Development*. Baltimore: Johns Hopkins University Press.

Black, Alan W., and Ian Reeve 1993 "Participation in Landcare Groups: the Relative Importance of Attitudinal and Situational Factors." *Journal of Environmental Management* 39:51-71.

Boserup, Ester 1965 *The Conditions of Agricultural Growth*. Chicago: Aldine.

— 1981 *Population and Technological Change: A Study of Long-term Trends*. Chicago: University of Chicago Press.

Buttel, Frederick, Olav Larson, and Gilbert Gillespie, Jr. 1990 *The Sociology of Agriculture*. New York: Greenwood Press.

Cancian, Frank 1967 "Stratification and Risk Taking: A Theory Tested on Agricultural Innovations." *American Sociological Review* 32:912-927.

— 1979 *The Innovator's Situation: Upper-Middle Class Conservatism in Agricultural Communities*. Stanford, Calif.: Stanford University Press.

Cochran, Willard 1979 *The Development of American Agriculture*. Minneapolis: University of Minnesota Press.

Downs, George, and Lawrence Mohr 1976 "Conceptual Issues in the Study of Innovations." *Administrative Science Quarterly* 21:700-714.

Duncan, Ian 1997 "Agricultural Innovation and Political Change in North India: The Lok Dal in Uttar Pradesh." *The Journal of Peasant Studies* 24:246-248.

Dunlap, Thomas 1981 *DDT: Scientists, Citizens and Public Policy*. Princeton, NJ.: Princeton University Press.

Fliegel, Frederick, and John van Es 1983 "The Diffusion-Adoption Process in Agriculture: Changes in Technology and Changing Paradigms." In Gene Summers, ed., *Technology and Social Change in Rural Areas*. Boulder, Colo.: Westview Press.

Frankel, Francine R. 1971 *India's Green Revolution: Economic Gains and Political Costs*. Princeton: Princeton University Press.

Friedland, William, and Amy Barton 1975 *Destalking the Wily Tomato: A Case Study in Social Consequences in California Agricultural Research*. Research Monograph no. 15. Davis: Department of Applied Behavioral Sciences, University of California Press.

Gartrell, David, and John Gartrell 1985 "Social Status and Agricultural Innovation: A Meta-analysis." *Rural Sociology* 50(1):38-50.

Gartrell, John 1977 "Status, Inequality and Innovation: The Green Revolution in Andhra Pradesh, India." *American Sociological Review* 42:318-337.

—, and David Gartrell 1979 "Status, Knowledge and Innovation." *Rural Sociology* 44:73-94.

Goldschmidt, Walter 1978 *As You Sow*. New York: Harcourt, Brace.

Grabowski, Richard 1995 "Commercialization, Nonagricultural Production, Agricultural Innovation, and

- Economic Development." *The Journal of Developing Areas* 30:41-62.
- Hightower, James 1972 *Hard Tomatoes, Hard Times: The Failure of America's Land-Grant College Complex*. Cambridge, Mass.: Schenkman.
- Islam, Rizwanul, and Jin Hehui 1994 "Rural Industrialization: An Engine of Prosperity in Postreform Rural China." *World Development* 22:1643-1659.
- Jamieson, Neil L., Le Trong Cuc, and A. Terry Rambo 1998 *The Development Crisis in Vietnam's Mountains*. Honolulu: East-West Center Special Reports No. 6.
- LeVein, Phillip 1978 "The Prospects for Small-Scale Farming in an Industrial Society: A Critical Appraisal of Small Is Beautiful." In Richard Dorf and Yvonne Hunter, eds., *Appropriate Visions*. San Francisco: Boyd and Fraser.
- Lewis, Scott, David Gartrell, and John Gartrell 1989 "Upper-Middle-Class Conservatism in Agricultural Communities: A Meta-analysis." *Rural Sociology* 54 (3):409-419.
- Pampel, Fred, and J. C. van Es 1977 "Environmental Quality and Issues of Adoption Research." *Rural Sociology* 42:57-71.
- Poleman, Thomas T., and D. K. Freebairn 1973 *Food, Population and Employment: The Impact of the Green Revolution*. New York: Praeger.
- Rogers, Everett 1983 *The Diffusion of Innovations*, 3rd ed. New York: Free Press.
- , and D. Lawrence Kincaid 1981 *Communication Networks: Toward a New Paradigm for Research*. New York: Free Press.
- Ryan, Bryce, and Neal Gross 1942 "The Diffusion of Hybrid Corn Seed in Two Iowa Communities." *Rural Sociology* 8:15-24.
- Schumpeter, Joseph 1950 *Capitalism, Socialism, and Democracy*. New York: Harper and Row.
- Sharma, Rita, and Thomas T. Poleman 1993 *The New Economics of India's Green Revolution*. Ithaca, N.Y.: Cornell University Press.
- Sommers, David G., and Ted L. Napier 1993 "Comparison of Amish and Non-Amish Farmers: A Diffusion/Farm-Structure Perspective." *Rural Sociology* 58:130-145.
- Spicer, Edward 1952 *Human Problems in Technological Change*. New York: Russell Sage Foundation.
- van Es, J. C. 1983 "The Adoption/Diffusion Tradition Applied to Resource Conservation: Inappropriate Use of Existing Knowledge." *The Rural Sociologist* 3:76-82.
- Warriner, G. Keith, and Trudy M. Moul 1992 "Kinship and Personal Communication Network Influences on the Adoption of Agriculture Conservation Technology." *Journal of Rural Studies* 8:279-291.
- Wasserman, Stanley, and Katherine Faust 1994 *Social Network Analysis*. Cambridge, Eng.: Cambridge University Press.

JOHN GARTRELL  
DAVID GARTRELL

## ALCOHOL

### INTRODUCTION

The sociological study of alcohol in society is concerned with two broad areas. (1) The first area is the study of alcohol behavior, which includes: (a) social and other factors in alcohol behavior, (b) the prevalence of drinking in society, and (c) the group and individual variations in drinking and alcoholism. (2) The second major area of study has to do with social control of alcohol, which includes: (a) the social and legal acceptance or disapproval of alcohol (social norms), (b) the social and legal regulations and control of alcohol in society, and (c) efforts to change or limit deviant drinking behavior (informal sanctions, law enforcement, treatment, and prevention). Only issues related to the first area of study, sociology of alcohol behavior, will be reviewed here.

### PHYSICAL EFFECTS OF ALCOHOL

There are three major forms of beverages containing alcohol (ethanol) that are regularly consumed. Wine is made from fermentation of fruits and usually contains up to 14 percent of ethanol by volume. Beer is brewed from grains and hops and contains 3 to 6 percent ethanol. Liquor (whisky, gin, vodka, and other distilled spirits) is usually 40 percent (80 proof) to 50 percent (100 proof) ethanol. A bottle of beer (12 ounces), a glass of wine (4 ounces), and a cocktail or mixed drink with a shot of whiskey in it, therefore, each have about the same absolute alcohol content, one-half to three-fourths of an ounce of ethanol.

Alcohol is a central nervous system depressant, and its physiological effects are a direct function of the percentage of alcohol concentrated in



the body's total blood volume (which is determined mainly by the person's body weight). This concentration is usually referred to as the BAC (blood alcohol content) or BAL (blood alcohol level). A 150-pound man can consume one alcoholic drink (about three-fourths of an ounce) every hour essentially without physiological effect. The BAC increases with each additional drink during that same time, and the intoxicating effects of alcohol will become noticeable. If he has four drinks in an hour, he will have an alcohol blood content of .10 percent, enough for recognizable motor-skills impairment. In almost all states, operating a motor vehicle with a BAC between .08 percent and .10 percent (determined by breathalyzer or blood test) is a crime and is subject to arrest on a charge of DWI (driving while intoxicated). At .25 percent BAC (about ten drinks in an hour) the person is extremely drunk, and at .40 percent BAC the person loses consciousness. Excessive drinking of alcohol over time is associated with numerous health problems. Cirrhosis of the liver, hepatitis, heart disease, high blood pressure, brain dysfunction, neurological disorders, sexual and reproductive dysfunction, low blood sugar, and cancer, are among the illnesses attributed to alcohol abuse (National Institute on Alcohol Abuse and Alcoholism 1981, 1987; Royce 1990; Ray and Ksir 1999).

### **SOCIAL FACTORS IN ALCOHOL BEHAVIOR**

Alcohol has direct effects on the brain, affecting motor skills, perception, and eventually consciousness. The way people actually behave while drinking, however, is only partly a function of the direct physical effects of ethanol. Overt behavior while under the influence of alcohol depends also on how they have learned to behave while drinking in the setting and with whom they are drinking with at the time. Variations in individual experience, group drinking customs, and the social setting produce variations in observable behavior while drinking. Actions reflecting impairment of coordination and perception are direct physical effects of alcohol on the body. These physical factors, however, do not account for "drunken comportment"—the behavior of those who are "drunk" with alcohol before reaching the stage of impaired muscular coordination (MacAndrew and Edgerton 1969). Social, cultural, and psychological factors are more

important in overt drinking behavior. Cross-cultural studies (MacAndrew and Edgerton 1969), surveys in the United States (Kantor and Straus 1987), and social psychological experiments (Marlatt and Rohsenow 1981), have shown that both conforming and deviant behavior while "under the influence" are more a function of sociocultural and individual expectations and attitudes than the physiological and behavioral effects of alcohol. (For an overview of sociocultural perspectives on alcohol use, see Pittman and White 1991)

Sociological explanations of alcohol behavior emphasize these social, cultural, and social psychological variables not only in understanding the way people act when they are under, or think they are under, the influence of alcohol but also in understanding differences in drinking patterns at both the group and individual level. Sociologists see all drinking behavior as socially patterned, from abstinence, to moderate drinking, to alcoholism. Within a society persons are subject to different group and cultural influences, depending on the communities in which they reside, their group memberships, and their location in the social structure as defined by their age, sex, class, religion, ethnic, and other statuses in society. Whatever other biological or personality factors and mechanisms may be involved, both conforming and deviant alcohol behavior are explained sociologically as products of the general culture and the more immediate groups and social situations with which individuals are confronted. Differences in rates of drinking and alcoholism across groups in the same society and cross-nationally reflect the varied cultural traditions regarding the functions alcohol serves and the extent to which it is integrated into eating, ceremonial, leisure, and other social contexts. The more immediate groups within this sociocultural milieu provide social learning environments and social control systems in which the positive and negative sanctions applied to behavior sustain or discourage certain drinking according to group norms. The most significant groups through which the general cultural, religious, and community orientations toward drinking have an impact on the individual are family, peer, and friendship groups, but secondary groups and the media also have an impact. (For a social learning theory of drinking and alcoholism that specifically incorporates these factors in the social and cultural context see Akers 1985, 1998; Akers and La

Greca 1991. For a review of sociological, psychological, and biological theories of alcohol and drug behavior see Goode 1993.)

### SOCIAL CHARACTERISTICS AND TRENDS IN DRINKING BEHAVIOR

**Age.** Table 1 shows that by time of high school graduation, the percentages of current teen-age drinkers (still under the legal age) is quite high, rivaling that of adults. The peak years for drinking are the young adult years (eighteen to thirty-four), but these are nearly equaled by students who are in the last year of high school (seventeen to eighteen years of age). For both men and women, the probability that one will drink at all stays relatively high from that time up to age thirty-five; about eight out of ten are drinkers, two-thirds are current drinkers, and one in twenty are daily drinkers. The many young men and women who are in college are even more likely to drink (Berkowitz and Perkins 1986; Wechsler et al. 1994). Heavy and frequent drinking peaks out in later years, somewhat sooner for men than women. After that the probability for both drinking and heavy drinking declines noticeably, particularly among the elderly. After the age of sixty, both the proportion of drinkers and of frequent or heavy drinkers decrease. Studies in the general population have consistently found that the elderly are less likely than younger persons to be drinkers, heavy drinkers, and problem drinkers (Cahalan and Cisin 1968; Fitzgerald and Mulford 1981; Meyers et al. 1981-1982; Borgatta et al. 1982; Holzer et al. 1984; Akers 1992).

**Sex.** The difference is not as great as it once was, but more men than women drink and have higher rates of problem drinking in all age, religious, racial, social class, and ethnic groups and in all regions and communities. Teenage boys are more likely to drink and to drink more frequently than girls, but the difference between male and female percentages of current drinkers at this age is less than it is in any older age group. Among adults, men are three to four times more likely than women (among the elderly as much as ten times more likely) to be heavy drinkers and two to three times more likely to report negative personal and social consequences of drinking (National Institute on Alcohol Abuse and Alcoholism 1987).

### Percentages Reporting Drinking by Age Group (1997)

Age Group	Lifetime	Past Year	Past Month
12-17	39.7	34	20.5
High School Seniors	81.7	74.8	52.7
18-25	83.5	75.1	58.4
26-34	88.9	74.6	60.2
35+	87	64.1	52.8

Table 1

SOURCE: Substance Abuse and Mental Health Services Administration 1998; University of Michigan (for high school seniors) 1998.

**Social Class.** The proportion of men and women who drink is higher in the middle class and upper class than in the lower class. The more highly educated and the fully employed are more likely to be current drinkers than the less educated and unemployed. Drinking by elderly adults increases as education increases, but there are either mixed or inconsistent findings regarding the variations in drinking by occupational status, employment status, and income (Holzer 1984; Borgatta 1982; Akers and La Greca 1991).

**Community and Location.** Rates of drinking are higher in urban and suburban areas than in small towns and rural areas. As the whole country has become more urbanized the regional differences have leveled out so that, while the South continues to have the lowest proportion of drinkers, there is no difference among the other regions for both teenagers and adults. Although there are fewer of them in the South, those who do drink tend to drink more per person than drinkers in other regions (National Institute on Alcohol Abuse and Alcoholism 1998a).

**Race, Ethnicity, and Religion.** The percent of drinking is higher among both white males and females than among African-American men and women. Drinking among non-Hispanic whites is also higher than among Hispanic whites. The proportion of problem or heavy drinkers is about the same for African Americans and white Americans (Fishburne et al. 1980; National Institute on Drug Abuse, 1988; National Institute on Alcohol Abuse and Alcoholism 1998a). There may be a tendency for blacks to fall into the two extreme categories, heavy drinkers or abstainers (Brown and Tooley

1989), and black males suffer the highest rate of mortality from cirrhosis of the liver (National Institute on Alcohol Abuse and Alcoholism 1998b). American Indians and Alaskan Natives have rates of alcohol abuse and problems several times the rates in the general population (National Institute on Alcohol Abuse and Alcoholism 1987).

Catholics, Lutherans, and Episcopalians have relatively high rates of drinking. Relatively few fundamentalist Protestants, Baptists, and Mormons drink. Jews have low rates of problem drinking, and Catholics have relatively high rates of alcoholism. Irish Americans have high rates of both drinking and alcoholism. Italian Americans drink frequently and heavily but apparently do not have high rates of alcoholism (see Cahalan et al. 1967; Mulford 1964). Strong religious beliefs and commitment, regardless of denominational affiliation, inhibit both drinking and heavy drinking among teenagers and college students (Cochran and Akers 1989; Berkowitz and Perkins 1986).

**Trends in Prevalence of Drinking.** There has been a century-long decline in the amount of absolute alcohol consumed by the average drinker in the United States. There was a period in the 1970s when the per capita consumption increased, and the proportion of drinkers in the population was generally higher by the end of the 1970s than at the beginning of the decade, although there were yearly fluctuations up and down. The level of drinking among men was already high, and the increases came mainly among youth and women. But in the 1980s the general downward trend resumed (Keller 1958; National Institute on Alcohol Abuse and Alcoholism 1981, 1987, 1998). Until the 1980s, this per capita trend was caused mainly by the increased use of lower-content beer and wine and the declining popularity of distilled spirits rather than a decreasing proportion of the population who are drinkers.

Alcohol-use rates were quite high in the United States throughout the 1970s and into the 1980s (see table 2). Since then, there have been substantial declines in use rates in all demographic categories and age groups. In 1979 more than two-thirds of American adolescents (twelve to seventeen years of age) had some experience with alcohol and nearly four out of ten were current drinkers (drank within the past month). In 1988, these proportions

had dropped to one-half and one-fourth respectively. In 1997, adolescent rates had dropped even lower to four out of ten having ever used alcohol and only two out of ten reporting use in the past month. Current use in the general U.S. population (aged twelve and older) declined from 60 percent in 1985 to 51 percent in 1997. Among the adult population eighteen years of age and older, current use declined from 71 percent in 1985 to 55 percent in 1997. Lifetime use rates have also declined from 88.5 percent in 1979 to 81.9 percent in 1997 (aged twelve and older). Generally, there have been declines in both annual (past year) prevalence of drinking (decreases of 3 to 5 percent) and current (past month) prevalence of drinking (decreases of 7 to 10 percent) among high school seniors, young adults, and older adults. Although lifetime prevalence is not a sensitive measure of short-term change in the adult population (since the lifetime prevalence is already fixed for the cohort of adults already sampled in previous surveys), it does reflect an overall decline in alcohol use. It should be remembered, however, that most of this is light to moderate consumption; the modal pattern of drinking for all age groups in the United States has long been and continues to be nondeviant, light to moderate social drinking.

The relative size of the reductions in drinking prevalence over the last two decades have been rather substantial; however, the proportions of drinkers remains high. By the time of high school graduation, one-half of adolescents are current drinkers and the proportion of drinkers in the population remains at this level throughout the young adult years. Three-fourths of high school seniors and young adults and two-thirds of adults over the age of thirty-five have consumed alcohol in the past year (see table 1).

Although lifetime use and current use rates appear to be continuing a slight decline in all categories, there have been some slight increases in rates of frequent (daily) drinking among high school seniors and young adults. While these increases do not approach the rates observed in the 1980s they may indicate that the overall rates are stabilizing and hint of possible increases in alcohol use rates in the future.

**Estimates of Prevalence of Alcoholism.** In spite of these trends in lower levels of drinking,

**Percentages Reporting Lifetime, Past Year, and Past Month Use of Alcohol  
in the U.S. Population Aged 12 and Older (1979-1997)**

	1979	1985	1991	1993	1997
<b>Lifetime</b>	88.5	84.9	83.6	82.6	81.9
<b>Past Year</b>	72.9	72.9	68.1	66.5	64.1
<b>Past Month</b>	63.2	60.2	52.2	50.8	51.4

*Table 2*

SOURCE: Substance Abuse and Mental Health Services Administration 1998.

alcoholism remains one of the most serious problems in American society. Alcohol abuse and all of the problems related to it cause enormous personal, social, health, and financial costs in American society. Cahalan et al. (1969) in a 1965 national survey characterized 6 percent of the general adult population and 9 percent of the drinkers as "heavy-escape" drinkers, the same figures reported for a 1967 survey (Cahalan 1970). These do not seem to have changed very much in the years since. They are similar to findings in national surveys from 1979 to 1988 (National Institute on Alcohol Abuse and Alcoholism 1981, 1987, 1988, 1989; Clark and Midanik 1982), which support an estimate that 6 percent of the general population are problem drinkers and that about 9 percent of those who are drinkers will abuse or fail to control their intake of alcohol. Royce (1989) and Vaillant (1983) both estimate that 4 percent of the general population in the United States are "true" alcoholics. This estimate would mean that there are perhaps 10.5 million alcoholics in American society (see also Liska 1997). How many alcoholics or how much alcohol abuse there is in our society is not easily determined because the very concept of alcoholism (and therefore what gets counted in the surveys and estimates) has long been and remains controversial.

### THE CONCEPT OF ALCOHOLISM

The idea of alcoholism as a sickness traces back at least 200 years (Conrad and Schneider 1980). There is no single, unified, disease concept, but the prevailing concepts of alcoholism today revolve around the one developed by E. M. Jellinek (1960) from 1940 to 1960. Jellinek defined alcoholism as a disease entity that is diagnosed by the "loss of control" over one's drinking and that progresses through a series of clear-cut "phases."

The final phase of alcoholism means that the person is rendered powerless by the disease to drink in a controlled, moderate, nonproblematic way.

The disease of alcoholism is viewed as a disorder or illness for which the individual is not personally responsible for having contracted. It is viewed as incurable in the sense that alcoholics can never truly control their drinking. That is, sobriety can be achieved by total abstention, but if even one drink is taken, the alcoholic cannot control how much more he or she will consume. It is a "primary" self-contained disease that produces the problems, abuse, and "loss of control" over drinking by those suffering from this disease. It can be controlled through proper treatment to the point where the alcoholic can be helped to stop drinking so that he or she is in "remission" or "recovering." "Once an alcoholic, always an alcoholic" is a central tenet of the disease concept. Thus, one can be a sober alcoholic, still suffering from the disease even though one is consuming no alcohol at all. Although the person is not responsible for becoming sick, he or she is viewed as responsible for aiding in the cure by cooperating with the treatment regimen or participation in groups such as Alcoholics Anonymous.

The disease concept is the predominant one in public opinion and discourse on alcohol (according to a 1987 Gallup Poll, 87 percent of the public believe that alcoholism is a disease). It is the principal concept used by the vast majority of the treatment professionals and personnel offering programs for alcohol problems. It receives widespread support among alcohol experts and continues to be vigorously defended by many alcohol researchers (Keller 1976; Vaillant 1983; Royce 1989). Alcoholics Anonymous, the largest single program for alcoholics in the world, defines alcoholism as a disease (Rudy 1986). The concept of

alcoholism as a disease is the officially stated position of the federal agency most responsible for alcohol research and treatment, the National Institute of Alcoholism and Alcohol Abuse (National Institute on Alcohol Abuse and Alcoholism 1987).

Nonetheless, many sociologists and behavioral scientists remain highly skeptical and critical of the disease concept of alcoholism (Trice 1966; Cahalan and Room 1974; Conrad and Schneider 1980; Rudy 1986; Fingarette 1988, 1991; Peele 1989). The concept may do more harm than good by discouraging many heavy drinkers who are having problems with alcohol, but who do not identify themselves as alcoholics or do not want others to view them as sick alcoholics, from seeking help. The disease concept is a tautological (and therefore untestable) explanation for the behavior of people diagnosed as alcoholic. That is, the diagnosis of the disease is made on the basis of excessive, problematic alcohol behavior that seems to be out of control, and then this diagnosed disease entity is, in turn, used to explain the excessive, problematic, out-of-control behavior.

In so far as claims about alcoholism as a disease can be tested, "Almost everything that the American public believes to be the scientific truth about alcoholism is false" (Fingarette 1988, p.1; see also Peele 1989; Conrad and Schneider 1980; Fingarette 1991; Akers 1992). The concept preferred by these authors and by other sociologists is one that refers only to observable behavior and drinking problems. The term *alcoholism* then is nothing more than a label attached to a pattern of drinking that is characterized by personal and social dysfunctions (Mulford and Miller 1960; Conrad and Schneider 1980; Rudy 1986; Goode 1993). That is, the drinking is so frequent, heavy, and abusive that it produces or exacerbates problems for the drinker and those around him or her including financial, family, occupational, physical, and interpersonal problems. The heavy drinking behavior and its attendant problems are themselves the focus of explanation and treatment. They are not seen as merely symptoms of some underlying disease pathology. When drinking stops or moderate drinking is resumed and drinking does not cause social and personal problems, one is no longer alcoholic. Behavior we label as alcoholic is problem drinking that lies at one extreme end of a continuum of drinking behavior with

abstinence at the other end and various other drinking patterns in between (Cahalan et al. 1969). From this point of view, alcoholism is a disease only because it has been socially defined as a disease (Conrad and Schneider 1980; Goode 1993).

**Genetic Factors in Alcoholism.** Contrary to what is regularly asserted, evidence that there may be genetic, biological factors in alcohol abuse is evidence neither in favor of nor against the disease concept, any more than evidence that there may be genetic variables in criminal behavior demonstrates that crime is a disease. Few serious researchers claim to have found evidence that a specific disease entity is inherited or that there is a genetically programmed and unalterable craving or desire for alcohol. It is genetic susceptibility to alcoholism that interacts with the social environment and the person's drinking experiences, rather than genetic determinism, that is the predominant perspective.

The major evidence for the existence of hereditary factors in alcoholism comes from studies that have found greater "concordance" between the alcoholism of identical twins than between siblings and from studies of adoptees in which offspring of alcoholic fathers were found to have an increased risk of alcoholism even though raised by nonalcoholic adoptive parents (Goodwin 1976; National Institute on Alcohol Abuse and Alcoholism 1982; U.S. Department of Health and Human Services 1987; for a review and critique of physiological and genetic theories of alcoholism see Rivers 1994). Some have pointed to serious methodological problems in these studies that limit their support for inherited alcoholism (Lester 1987). Even the studies finding evidence for an inherited alcoholism report that only a small minority of those judged to have the inherited traits become alcoholic and an even smaller portion of all alcoholics have indications of hereditary tendencies. Whatever genetic variables there are in alcoholism apparently come into play in a small portion of cases. Depending upon the definition of alcoholism used, the research shows that biological inheritance either makes no difference at all or makes a difference for only about one out of ten alcoholics. Social and social psychological factors are the principal variables in alcohol behavior, including that which is socially labeled and diagnosed as alcoholism (Fingarette 1988; Peele 1989).

## REFERENCES

- Akers, Ronald L. 1998 *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston: Northeastern University Press.
- 1992 *Drugs, Alcohol, and Society*. Belmont, Calif.: Wadsworth.
- 1985 *Deviant Behavior: A Social Learning Approach*. Belmont, Calif.: Wadsworth.
- , and Anthony J. La Greca 1991 “Alcohol Use among the Elderly: Social Learning, Community Context, and Life Events.” In David J. Pittman and Helene White, eds., *Society, Culture and Drinking Patterns Re-examined*. New Brunswick, N.J.: Rutgers University Press.
- Berkowitz, Alan D., and H. Wesley Perkins 1986 “Problem Drinking among College Students: A Review of Recent Research.” *Journal of American College Health* 35:21–28.
- Borgatta, Edgar F., Rhonda J. V. Montgomery, and Marie L. Borgatta 1982 “Alcohol Use and Abuse, Life Crisis Events, and the Elderly.” *Research on Aging* 4:378–408.
- Brown, Frieda, and Joan Tooley 1989 “Alcoholism in the Black Community.” In Gary W. Lawson and Ann W. Lawson, eds. *Alcohol and Substance Abuse in Special Populations*, 115–130. Rockville, Md.: Aspen Publishers.
- Cahalan, Don 1970 *Problem Drinkers: A National Survey*. San Francisco: Jossey-Bass.
- , Ira H. Cisin, and Helen M. Crossley 1967 *American Drinking Practices*. Washington, D.C.: George Washington University Press.
- Cahalan, Don, and Robin Room 1974 *Problem Drinking Among American Men*. New Brunswick, N.J.: College and University Press.
- Clark, Walter B., and Lorraine Midanik 1982 “Alcohol Use and Alcohol Problems among U.S. Adults: Results of the 1979 Survey.” *Alcohol Consumption and Related Problems*. Department of Health and Human Services, No. 82–1190. Washington, D.C.: U.S. Government Printing Office.
- Cochran, John K., and Ronald L. Akers 1989 “Beyond Hellfire: An Exploration of the Variable Effects of Religiosity on Adolescent Marijuana and Alcohol Use.” *Journal of Research on Crime and Delinquency* 26:198–225.
- Conrad, Peter, and Joseph W. Schneider 1980 *Deviance and Medicalization*. St. Louis: C.V. Mosby.
- Fingarette, Herbert 1991 “Alcoholism: The Mythical Disease.” In David J. Pittman and Helene White, eds., *Society, Culture and Drinking Patterns Re-examined*. New Brunswick, N.J.: Rutgers University Press.
- 1988 *Heavy Drinking: The Myth of Alcoholism as a Disease*. Berkeley: University of California Press.
- Fishburne, Patricia, Herbert I. Abelson, and Ira Cisin 1980 *National Survey on Drug Abuse: Main Findings*. Washington, D.C.: U.S. Government Printing Office.
- Fitzgerald, J. L., and Harold A. Mulford 1981 “The Prevalence and Extent of Drinking in Iowa, 1979.” *Journal of Studies on Alcohol* 42:38–47.
- Goode, Erich 1993 *Drugs in American Society*, 4th ed. New York: McGraw-Hill.
- Goodwin, Donald 1976 *Is Alcoholism Hereditary?* New York: Oxford University Press.
- Holzer, C., Lee Robins, Jerome Meyers, M. Weissman, G. Tischler, P. Leaf, J. Anthony, and P. Bednarski 1984 “Antecedents and Correlates of Alcohol Abuse and Dependence in the Elderly.” In George Maddox, Lee Robins, and Nathan Rosenberg, eds., *Nature and Extent of Alcohol Abuse Among the Elderly*. Department of Health and Human Services No. 84–1321. Washington, D.C.: U.S. Government Printing Office.
- Jellinek, E. M. 1960 *The Disease Concept of Alcoholism*. New Haven, Conn.: Hillhouse Press.
- Johnston, Lloyd D., Patrick M. O’Malley, and Jerald G. Bachman 1989 *Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young Adult Populations, 1975–1988*. Washington, D.C.: U.S. Government Printing Office.
- Kantor, Glenda K., and Murray A. Straus 1987 “The ‘Drunken Bum’ Theory of Wife Beating.” *Social Problems* 34:213–230.
- Keller, Mark 1976 “The Disease Concept of Alcoholism Revisited.” *Journal of Studies on Alcohol* 37:1694–1717.
- 1958 “Alcoholism: Nature and Extent of the Problem.” *Annals* 315:1–11.
- Lester, David 1987 “Genetic Theory: An Assessment of the Heritability of Alcoholism.” New Brunswick, N.J.: Center of Alcohol Studies, Rutgers University.
- Liska, Ken 1997 *Drugs and the Human Body*, 5th ed. Upper Saddle River, N.J.: Prentice Hall.
- MacAndrew, Craig, and Robert B. Edgerton 1969 *Drunk and Comportment: A Social Explanation*. Chicago: Aldine.
- Marlatt, G. Alan, and Damaris J. Rohsenow 1981 “The Think-Drink Effect.” *Psychology Today* (Dec):60–69, 93.
- Meyers, A. R., E. Goldman, R. Hingson, N. Scotch, and T. Mangione 1981 “Evidence of Cohort and Generational Differences in Drinking Behavior of Older Adults.” *International Journal of Aging and Human Development* 14:31–44.

- Mulford, Harold A. 1964 "Drinking and Deviant Drinking, USA, 1963." *Quarterly Journal of Studies on Alcohol* 25:634-650.
- , Donald Miller 1960 "Drinking in Iowa IV: Pre-occupation with Alcohol and Definitions of Alcoholism, Heavy Drinking, and Trouble Due to Drinking." *Quarterly Journal of Studies on Alcohol* 21:279-296.
- National Institute on Alcohol Abuse and Alcoholism 1998a *Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 1977-1996*. (Gerald D. Williams, Frederick S. Stinson, Lorna L. Sanchez, and Mary C. Dufour) Rockville, Md.: National Institute on Alcohol Abuse and Alcoholism.
- 1998b *Surveillance Report No. 48: "Liver Cirrhosis Mortality in the United States, 1970-1995*. (Forough Saadatmand, Frederick S. Stinson, Bridget F. Grant, and Mary C. Dufour) Rockville, Md.: National Institute on Alcohol Abuse and Alcoholism.
- 1987 *Alcohol and Health: Sixth Special Report to the U. S. Congress from the Secretary of Health and Human Services*. Rockville, Md.: National Institute on Alcohol Abuse and Alcoholism.
- 1982 "Researchers Investigating Inherited Alcohol Problems, NIAAA Information and Feature Service No. 99, August 30. Rockville, Md.: National Clearinghouse for Alcohol Information, National Institute on Alcohol Abuse and Alcoholism.
- 1981 *Fourth Special Report to the U.S. Congress on Alcohol and Health*. Washington, D.C.: U.S. Government Printing Office.
- National Institute on Drug Abuse 1989 "Highlights of the 1988 Household Survey on Drug Abuse." *NIDA Capsules*, August. Rockville, Md.: National Institute on Drug Abuse.
- 1989 *Results from High School, College, and Young Adult Populations, 1975-1988*. Washington, D.C.: U.S. Government Printing Office.
- 1985 *National Household Survey on Drug Abuse: Main Findings*. Rockville, Md.: National Institute on Drug Abuse.
- Peele, Stanton 1989 *Diseasing of America: Addiction Treatment Out of Control*. Lexington, Mass.: Lexington Books.
- Pittman, David J., and Helene White (eds.) 1991 *Society, Culture, and Drinking Patterns Re-examined*. New Brunswick, N.J.: Rutgers University Press.
- Ray, Oakley, and Charles Ksir 1999 *Drugs, Society, and Human Behavior*. 8th ed. Boston: WCB/McGraw-Hill.
- Rivers, P. Clayton 1994 *Alcohol and Human Behavior: Theory, Research, and Practice*. Englewood Cliffs, N.J.: Prentice Hall.
- Royce, James E. 1989 *Alcohol Problems and Alcoholism*, rev. ed. New York: Free Press.
- Rudy, David 1986 *Becoming Alcoholic: Alcoholics Anonymous and the Reality of Alcoholism*. Carbondale, Ill.: Southern Illinois University Press.
- Substance Abuse and Mental Health Services Administration 1998 *Preliminary Results from the 1997 National Household Survey on Drug Abuse*. Rockville, Md.: (SAMHSA). <http://www.samhsa.gov>.
- Trice, Harrison 1966 *Alcoholism in America*. New York: McGraw-Hill.
- University of Michigan 1998 *Monitoring the Future Survey*. <http://www.isr.umich.edu/src/mtf/>.
- U. S. Department of Health and Human Services 1987 *Sixth Special Report to the U. S. Congress on Alcohol and Health from the Secretary of Health and Human Services*. Rockville, Md.: National Institute on Alcohol Abuse and Alcoholism.
- Vaillant, George 1983 *The Natural History of Alcoholism*. Cambridge, Mass.: Harvard University Press.
- Wechsler, H., A. Davenport, G. W. Dowdall, B. Moeykens, and S. Castillo 1994 "Health and Behavioral Consequences of Binge Drinking in College: A National Survey of Students at 140 Campuses." *Journal of the American Medical Association* 272:1672-1677.

RONALD L. AKERS  
THOMAS R. HEFFINGTON

## ALIENATION

Since 1964, many commentators have been speaking of a crisis of confidence in the United States, a malaise marked by widespread public belief that major institutions—businesses, labor unions, and especially the government, political parties, and political leaders—are unresponsive, remote, ineffective, and not to be trusted (Lipset and Schneider 1983). *Alienation* became the catchword for these sentiments, detected among discontented workers, angry youth, and militant minority groups. American leaders concerned about the increase in alienation found new relevance in ongoing discussions among sociologists and other social scientists, who have defined alienation, used survey research to measure the level of alienation in society, and have debated the causes, significance, and consequences of alienation and particularly, political alienation.

### DIMENSIONS OF ALIENATION AND POLITICAL ALIENATION

Theorists and sociological researchers have developed different definitions of alienation (Seeman 1975). Scholars influenced by the philosophical writings of Karl Marx have used the word *alienation* to mean self-estrangement and the lack of self-realization at work (Blauner 1964; Hodson 1996). Marx argued that although humans by their very nature are capable of creative and intrinsically rewarding work, the Industrial Revolution alienated workers from their creative selves and reduced workers to the unskilled tenders of machines (Braverman 1974). The worker produced machinery and other commodities that formed the capitalist system of workplace hierarchies and global markets, which the worker could not control. Rather, the system dominated workers as an alienated, “reified” force, apart from the will and interests of workers (Meszaros 1970). Whereas this oldest definition links alienation to the development of capitalism in modern society, some scholars see alienation as a characteristic reaction to the postmodern condition of fragmented multiple images and loss of individual identities and any shared meanings (Geyer 1996).

Alienation can also refer to the isolation of individuals from a community—a detachment from the activities, identifications, and ties that a community can provide. In addition, the concept of alienation has included the notion of cultural radicalism or estrangement from the established values of a society. Inglehart (1981) has argued that the highly educated generation that came of age in the counterculture of the 1960s rejected their elders’ traditional values of materialism, order, and discipline. Easterlin (1980, pp. 108–111) suggests that it is the relatively large cohort size of the Baby Boom generation that led them to suffer competition for jobs, psychological stress, discontent, and hence, generalized political alienation. On the contrary, Inglehart (1997) argues that baby boomers and succeeding generations will only express alienation against specific authoritative institutions, such as the police, the military, and churches. With succeeding generations increasingly espousing “postmaterialist” values such as the quality of life, self-realization, and participatory democracy, Inglehart finds a worldwide increase in some activities that reduce alienation such as petition-signing and political conversation.

Much of the literature on alienation in the 1990s focused on alienation from political institutions, and some writers have examined how alienation has changed in former authoritarian nations such as Argentina and South Africa and in Eastern Europe (Geyer 1996; Geyer and Heinz 1992). Sociologists interested in the political well-being of the United States have measured the extent to which individuals feel powerless over government (i.e., unable to influence government) and perceive politics as meaningless (i.e., incomprehensible; Seeman 1975). Such attitudes may be connected to a situation of normlessness, or *anomie*, which occurs when individuals are no longer guided by the political rules of the game (Lipset and Raab 1978). Social scientists have been concerned that alienation might reduce political participation through institutional channels such as voting, and might lead to nonconventional activity like protest movements and collective violence.

### MEASUREMENT AND CONSEQUENCES OF POLITICAL ALIENATION

Political alienation consists of attitudes whereby citizens develop (or fail to develop) meanings and evaluations about government and about their own power (or powerlessness) in politics. Specifically, political alienation is composed of the attitudes of distrust and inefficacy. Distrust (also called cynicism) is a generalized negative attitude about governmental outputs: the policies, operations, and conditions produced by government. Compared to the simple dislike of a particular policy or official, distrust is broader in scope. Whereas distrust is an evaluation of governmental outputs, inefficacy is an expectation about inputs, that is, the processes of influence over government. People have a sense of inefficacy when they judge themselves as powerless to influence government policies or deliberations (Gamson 1971).

Researchers have sought to find opinion poll questions that yield responses consistently correlated to only one underlying attitude, of distrust for example. Mason, House, and Martin (1985) argue that the most “internally valid” measures of distrust are two questions: “How much of the time do you think you can trust the government in Washington to do what is right—just about all of the time, most of the time, or only some of the time?” and “Would you say that the government is



pretty much run by a few big interests looking out for themselves or that it is run for the benefit of all people?" Similarly, a person's sense of inefficacy can be measured by asking the person to agree or disagree with the following statements, which contain the words "like me": "People like me don't have any say about what the government does" and "I don't think public officials care much what people like me think."

During election years since the 1950s, the Center for Political Studies at the University of Michigan at Ann Arbor has posed these and other questions to national samples of citizens. Those replying that you can trust the government only some of the time or none of the time comprised 22 percent in 1964 but 73 percent in 1980. This percentage fell during President Ronald Reagan's first term but then increased through 1994 (reaching 78 percent) before falling to 67 percent in 1996. Those disagreeing with the statement that public officials care rose from 25 percent in 1960, to 52 percent in 1980 and 66 percent in 1994 (Orren 1997; Poole and Mueller 1998).

In addition, polls indicate that in the same time period, increasing numbers of citizens felt that government was less responsive to the people (Lipset and Schneider 1983, pp. 13–29). This attitude, which can be termed *system unresponsiveness*, was measured by asking questions that did not use the words "like me." Responses to the questions thus focused not on the respondents' evaluations of their own personal power, but rather on their judgments of the external political system. (Craig 1979 conceptualizes system unresponsiveness as "output inefficacy").

What are the consequences of the increase in political alienation among Americans? Social scientists have investigated whether individuals with highly alienated attitudes are more likely to withdraw from politics, engage in violence, or favor protest movements or extremist leaders. Research findings have been complicated by the fact that the same specific alienated attitudes have been linked to different kinds of behaviors (Schwartz 1973, pp. 162–177).

The alienated showed little tendency to support extremist candidates for office. (The only exception was that high-status alienated citizens supported Goldwater's presidential campaign in 1964. See Wright 1976, pp. 227, 251; Herring

1989, p. 98.) Social scientists have generally agreed that politically alienated individuals are less likely to participate in conventional political processes. During four presidential elections from 1956 to 1968, citizens with a low sense of efficacy and a low level of trust were less likely to vote, attend political meetings, work for candidates, contribute money, or even pay attention to the mass media coverage of politics. Although some studies fail to confirm that those with low *trust* are likely to be apathetic (Citrin 1974, p. 982), those with a low sense of political *efficacy* are indeed likely to be nonvoters, mainly because they are also less educated (Lipset and Schneider 1983, p. 341). In the United States, the percentage of eligible voters who actually cast ballots declined between 1960 and 1980, while the percentage who expressed political inefficacy rose in the same period; Abramson and Aldrich (1982) estimate that about 27 percent of the former trend is caused by the latter. (See Shaffer 1981 for confirmation but Cassel and Hill 1981 and Miller 1980 for contrary evidence).

Piven and Cloward (1988) vigorously dispute the notion that the alienated attitudes of individuals are the main cause for the large numbers of nonvoters in the United States. Piven and Cloward construct a historical explanation—that in the early twentieth century, political reformers weakened local party organizations in cities, increased the qualifications for suffrage, and made voting registration procedures more difficult. Legal and institutional changes caused a sharp decrease in voting, which only then led to widespread political alienation. Voting participation, especially among the minority poor in large cities, continues to be low because of legal requirements to register in advance of election day and after a change in residence, and because of limited locations to register.

Some researchers have found that the politically alienated are more likely to utilize nonconventional tactics such as political demonstrations or violence. College students who participated in a march on Washington against the Vietnam War, compared to a matched sample of students from the same classes at the same schools, expressed more alienated attitudes, stemming from an underlying sense of inefficacy and system unresponsiveness (Schwartz 1973 pp. 138–142). Paige's (1971) widely influential study drew on Gamson's distinctions between trust and efficacy and showed

that Blacks who participated in the 1967 riot in Newark, New Jersey, had low levels of political trust but high levels of political efficacy (i.e., high capabilities and skills to affect politics, measured indirectly in this instance by the respondents' level of political knowledge). However, Sigelman and Feldman's (1983) attempt to replicate Paige's findings in a seven-nation study discovered that the participants and supporters of unconventional political activity were only slightly more likely to feel both efficacious and distrusting. Rather than being generally distrusting, participants and supporters in some nations were more likely to be dissatisfied about specific policies. (See also Citrin 1974, p. 982 and Craig and Maggiotto 1981 for the importance of specific dissatisfactions).

Even though politically alienated individuals may sometimes be found in social movements, the alienation of individuals is not necessarily the cause of social movements. McCarthy and Zald (1977) have argued that alienation and indeed policy dissatisfactions and other grievances are quite common in societies. Whether or not a social movement arises depends on the availability of resources and the opportunities for success. The civil rights movement, according to McAdam (1982), succeeded not when blacks believed that the political system was unresponsive, but rather when blacks felt that some national leaders showed signs of favoring their cause.

#### DISTRIBUTION AND SIGNIFICANCE OF POLITICAL ALIENATION

Social scientists have argued that political alienation is concentrated in different types of groups—among those who dislike politicians of the opposing political party, in certain economic and racial groups, and among those dissatisfied with government policy. Each of these findings supports a different assessment of the causes and the importance of political alienation.

**Partisan Bickering?** First of all, high levels of political distrust can be found among those who have a negative view of the performance of the presidential administration then in office. Citrin (1974) concludes that widespread expressions of political distrust (cynicism) merely indicate Democratic versus Republican Party rivalries as usual.

Cynicism, rather than being an expression of deep discontent, is nothing more than rhetoric and ritual and is not a threat to the system. Even partisans who intensely distrust a president from an opposing party are proud of the governmental system in the United States and want to keep it. However, King (1997) argues that distrust stems from a more serious problem, that congressional leaders and activists in political parties have become more ideological and polarized (see Lo and Schwartz 1998 on conservative leaders). The public has remained in the center but is becoming alienated from politicians whose ideologies are seen as far removed from popular concerns.

**The alienation of social classes and minorities.** Second, other researchers interested in finding concentrations of the politically alienated have searched not among people with varying partisan identifications, but rather in demographic groups defined by such variables as age, gender, education, and socioeconomic class. Many public opinion surveys using national samples have found alienation only weakly concentrated among such groups (Orren 1997). In the 1960s, the sense of inefficacy increased uniformly throughout the entire U.S. population, rather than increasing in specific demographic groups such as blacks or youth (House and Mason 1975). Using a 1970 survey, Wright (1976) noted that feelings of inefficacy and distrust were somewhat concentrated among the elderly, the poorly educated, and the working class. Still, Wright's conclusion was that the alienated were a diverse group that consisted of both rich and poor, black and white, and old and young, making it very unlikely that the alienated could ever become a unified political force.

Research on the gender gap, that is, the differing political attitudes between women and men (Mueller 1988), indicates that women are not more politically alienated than men (Poole and Muller 1998). In fact, a higher percentage of women compared to men support more government spending on social programs and a more powerful government with expanded responsibilities (Clark and Clark 1996) and thus are less distrustful of the broad scope of government. Some studies have shown that the politically alienated are indeed concentrated among persons with less education and lower income and occupational status (Wright

1976, p. 136; Lipset and Schneider 1983, pp. 311–315; Finifter 1970; Form and Huber 1971). Research that directly focuses on obtaining the opinions of minority and poor respondents has uncovered high degrees of political alienation among these groups. A survey of roughly equal numbers of blacks and whites in metropolitan Detroit in 1992 showed that blacks, compared to whites, evaluated schools and the police more negatively, distrusted local government more, and thought that participation in local politics was less efficacious (Bledsoe et al. 1996). Bobo and Hutchings (1966) oversampled minority residents of Los Angeles County and found that higher percentages of blacks compared to whites expressed “racial alienation,” that is, the opinion that blacks faced inferior life chances, fewer opportunities, and unfair treatment. Blacks living in Detroit neighborhoods where over 20 percent of the residents are poor, were more likely than other blacks to say that they had little influence in community decisions and that community problems were complex and unsolvable (Cohen and Dawson 1993).

Wright argues that even though sizable numbers of persons express alienated attitudes, these people pose little threat to the stability of regimes, because they rarely take political action and even lack the resources and skills to be able to do so. Lipset (1963) has argued that apathy is a virtue because it allows elites in democratic societies to better exert leadership. (For a critique see Wolfe 1977, p. 301.) For many social scientists in the 1950s, widespread apathy was a welcome alternative to the alleged mass activism that had produced the fascist regimes in Germany and Italy. However, Wright (1976, pp. 257–301) counters that since the alienated masses actually pose no threat to the contemporary political system, an increase in mass democratic participation, perhaps the mobilization of workers on the issues of class division, could very well be beneficial.

But the class mobilization that Wright envisions might turn out to be a middle-class affair (Teixeira 1996) rather than a working-class revolt. Whereas Lipset and Wright have been concerned about the concentration of political alienation in the lower socioeconomic strata, Warren (1976) emphasizes the alienation among “Middle American Radicals,” who believe that they are disfavored by a government that gives benefits to the poor and to the wealthy. Feelings of inefficacy and

distrust have increased the most among the middle strata—private-sector managers, middle-income workers, and a “new layer” of public-sector professionals (Herring 1989).

Unlike the poor, the middle strata have the resources to protest and to organize social movements and electoral campaigns, exemplified by protests against the property tax that culminated with the passage of Proposition 13 in California and Proposition 2 1/2 in Massachusetts. Property tax protesters were middle-class homeowners who expressed their political alienation when they condemned “taxation without representation.” Citizens who felt cut off from political decision making were the most likely to support the tax revolt (Lowery and Sigelman 1981). Protests centered around unresponsive government officials who continued to increase assessments and tax rates, without heeding the periodic angry protests of homeowners. Movement activists interpreted their own powerlessness and power in community and metropolitan politics, thereby shaping the emerging tactics and goals of a grass-roots citizens’ movement (Lo 1995).

**A Crisis for Democracy?** Finally, other social scientists have found intense alienation among those with irreconcilable dissatisfactions about government policy, thus threatening to make effective government impossible. Miller (1974) argued that between 1964 and 1970, political distrust (cynicism) increased simultaneously among those favoring withdrawal and those favoring military escalation in the Vietnam War. Similarly, distrust increased both among blacks who thought that the civil rights movement was making too little progress, and among white segregationists who held the opposite view. The 1960s produced two groups—cynics of the left and cynics of the right, each favoring polarized policy alternatives (see also Lipset and Schneider 1983, p. 332). Cynics of the right, for example, rejected both the Democratic and Republican parties as too liberal. (Herring 1989 has developed a similar “welfare split” thesis, that more social spending has different effects on the distrust level of various groups but, overall, raises political distrust.) Miller concludes that increased cynicism, along with a public bifurcated into extreme stances on issues, makes it difficult for political leaders to compromise and build support for centrist policies. While agreeing with Wright that the alienated are divided amongst themselves,

Miller argues that this fragmentation does indeed constitute a crisis of legitimacy for American politics.

For some social theorists, widespread political alienation is a sign of even deeper political contradictions. Throughout American history, as citizens have fought to extend their democratic freedoms and personal rights, businesses have used the notion of property rights to protect their own interests and stifle reform (Bowles and Gintis 1987). Wolfe (1977) sees political alienation as a symptom of how the democratic aspirations of the citizenry have been frustrated by the state, which has attempted to foster the growth of capitalism while at the same time maintaining popular support.

### OVERCOMING ALIENATION

One diagnosis for overcoming alienation has been proposed by Sandel, Etzioni (1996), and others from the communitarian perspective, which promotes the values of civic commitment. Sandel (1996) argues that citizens today feel powerless over their fate and disconnected from politics because of an excessively individualist culture in the United States. America's leaders must encourage devotion to the common good, attachments to communities, volunteerism, and moral judgments and dialogues. Political alienation can be overcome through the associations and networks of civil society.

Others also trace political alienation back to its roots in society, but focus on work and economic hardships, which Marx long ago characterized as alienating and that now prevent the emergence of the caring and democratic public life envisioned by the communitarians (Bennett 1998). According to Lerner (1991), people experience a deep and debilitating sense of inefficacy (what he terms "surplus powerlessness") in their personal and family lives. Surplus powerlessness can be overcome through such measures as communities of compassion, occupational stress groups, and family support groups, which will build the attachments and religious values sought by the communitarians, while at the same time compassionate unions will seek to change power relations at work and in the society at large.

Alienation, originally a Marxist concept depicting the economic deprivations of industrial workers, is now a political concept portraying the

plight of citizens increasingly subjected to the authority and the bureaucracy of the state in advanced capitalist societies. Perhaps returning to the original theorizing about alienation in the economic sphere can deepen our analysis of contemporary political alienation.

### REFERENCES

- Abramson, Paul, and John Aldrich 1982 "Decline of Electoral Participation in America." *American Political Science Review* 76:502-521.
- Bennett, W. Lance 1998 "The Uncivic Culture: Communication, Identity, and the Rise of Lifestyle Politics." *PS: Political Science and Politics* 31:740.
- Blauner, Robert 1964 *Alienation and Freedom: The Factory Worker and His Industry*. Chicago: University of Chicago Press.
- Bledsoe, Timothy, Lee Sigelman, Susan Welch, and Michael W. Combs 1996 "The Social Contract in Black and White: Racial Differences in Evaluations of Government." In Stephen C. Craig, ed., *Broken Contract? Changing Relationships between Americans and their Government*. Boulder, Colo.: Westview Press.
- Bobo, Lawrence, and Vincent L. Hutchings 1996 "Perceptions of Racial Group Competition: Extending Blumer's Theory of Group Position to a Multiracial Social Context." *American Sociological Review* 61:951-972.
- Bowles, Samuel, and Herbert Gintis 1987 *Democracy and Capitalism: Property, Community, and the Contradictions of Modern Social Thought*. New York: Basic Books.
- Braverman, Harry 1974 *Labor and Monopoly Capital*. New York: Monthly Review Press.
- Cassel, Carol A., and David B. Hill 1981 "Explanations of Turnout Decline: A Multivariate Test." *American Politics Quarterly* 9:181-195.
- Citrin, Jack 1974 "Comment: The Political Relevance of Trust in Government." *American Political Science Review* 68:973-988.
- Clark, Janet, and Cal Clark 1996 "The Gender Gap: A Manifestation of Women's Dissatisfaction with the American Polity?" In Stephen C. Craig, ed., *Broken Contract? Changing Relationships between Americans and their Government*. Boulder, Colo.: Westview Press.
- Cohen, Cathy J., and Michael C. Dawson 1993 "Neighborhood Poverty and African American Politics." *American Political Science Review* 87: 286-303.
- Craig, Stephen C. 1979 "Efficacy, Trust, and Political Behavior: An Attempt to Resolve a Lingering Conceptual Dilemma." *American Politics Quarterly* 7:225-239.

- , and Michael Maggiotto 1981 "Political Discontent and Political Action." *Journal of Politics* 43:514–522.
- Easterlin, Richard A. 1980 *Birth and Fortune: The Impact of Numbers on Personal Welfare*. New York: Basic Books.
- Etzioni, Amitai 1996 *The New Golden Rule: Community and Morality in a Democratic Society*. New York: Basic Books.
- Finifter, Ada W. 1970 "Dimensions of Political Alienation." *American Political Science Review* 64:389–410.
- Form, William H., and Joan Huber Rytina 1971 "Income, Race, and the Ideology of Political Efficacy." *Journal of Politics* 33:659–688.
- Gamson, William 1971 "Political Trust and its Ramifications." In Gilbert Abcarian and John W. Soule, eds., *Social Psychology and Political Behavior: Problems and Prospects*. Columbus, Oh.: Charles E. Merrill.
- Geyer, Felix (ed.) 1996 *Alienation, Ethnicity, and Postmodernism*. Westport, Conn.: Greenwood Press.
- , and Walter R. Heinz (eds.) 1992 *Alienation, Society, and the Individual: Continuity and Change in Theory and Research*. New Brunswick, N.J.: Transaction Publishers.
- Herring, Cedric 1989 *Splitting the Middle: Political Alienation, Acquiescence, and Activism Among America's Middle Layers*. New York: Praeger.
- Hodson, Randy 1996 "Dignity in the Workplace under Participative Management: Alienation and Freedom Revisited." *American Sociological Review*, 61:719–738.
- House, James, and William Mason 1975 "Political Alienation in America, 1952–1968." *American Sociological Review* 40:123–147.
- Inglehart, Ronald 1997 "Postmaterialist Values and the Erosion of Institutional Authority." In Joseph P. Nye, Jr., Philip D. Zelikow, and David C. King, eds., *Why People Don't Trust Government*. Cambridge, Mass.: Harvard University Press.
- 1981 "Post-Materialism in an Environment of Insecurity." *American Political Science Review* 75:880–900.
- King, David C. 1997 "The Polarization of American Parties and Mistrust of Government." In Joseph P. Nye, Jr., Philip D. Zelikow, and David C. King, eds., *Why People Don't Trust Government*. Cambridge, Mass.: Harvard University Press.
- Lerner, Michael 1991 *Surplus Powerlessness: The Dynamics of Everyday Life and the Psychology of Individual and Social Transformation*. Atlantic Highlands, N.J.: Humanities Press International.
- Lipset, Seymour Martin 1963 *Political Man*. Garden City, N.Y.: Anchor.
- , and Earl Raab 1978 *The Politics of Unreason: Right-Wing Extremism in America, 1790–1977*. Chicago: University of Chicago Press.
- , and William Schneider 1983 *The Confidence Gap: Business, Labor, and Government in the Public Mind*. New York: Free Press.
- Lo, Clarence Y.H. 1995 *Small Property versus Big Government: Social Origins of the Property Tax Revolt*. Berkeley and Los Angeles: University of California Press.
- , and Michael Schwartz 1998 *Social Policy and the Conservative Agenda*. Oxford, Eng. and Cambridge, Mass.: Blackwell Publishers.
- Lowrey, David, and Lee Sigelman 1981 "Understanding the Tax Revolt: Eight Explanations." *American Political Science Review* 75: 963–974.
- McAdam, Doug 1982 *Political Process and the Development of Black Insurgency*. Chicago: University of Chicago Press.
- McCarthy, John D., and Mayer N. Zald 1977 "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82: 1112–1141.
- Mason, William W., James S. House, and Steven S. Martin 1985 "On the Dimensions of Political Alienation in America." In Nancy Brandon Tuma, ed., *Sociological Methodology*. San Francisco: Jossey-Bass.
- Meszaros, I. 1970 *Marx's Theory of Alienation*. London: Merlin.
- Miller, Arthur 1974 "Political Issues and Trust in Government: 1964–1970." *American Political Science Review* 68:951–972.
- Miller, Warren E. 1980 "Disinterest, Disaffection, and Participation in Presidential Politics." *Political Behavior* 2:7–32.
- Mueller, Carol M. (ed.) 1988 *The Politics of the Gender Gap: The Social Construction of Political Influence*. Beverly Hills, Calif.: Sage.
- Orren, Gary 1997 "Fall from Grace: The Public's Loss of Faith in Government." In Joseph P. Nye, Jr., Philip D. Zelikow, and David C. King, eds., *Why People Don't Trust Government*. Cambridge, Mass.: Harvard University Press.
- Paige, Jeffery 1971 "Political Orientation and Riot Participation." *American Sociological Review* 36:801–820.
- Piven, Frances Fox, and Richard A. Cloward 1988 *Why Americans Don't Vote*. New York: Pantheon.
- Poole, Barbara L., and Melinda A. Mueller 1998 "Alienation and the 'Soccer Mom': A Media Creation or a New Trend in Voting Behavior?" In Thomas J. Johnson, Carol E. Hays, and Scott P. Hays, eds.,

- Engaging the Public: How Government and the Media and Reinvent American Democracy.* Lanham, Md.: Rowman and Littlefield.
- Sandel, Michael 1996 *Democracy's Discontent: America in Search of a Public Philosophy.* Cambridge, Mass.: Harvard University Press.
- Schwartz, David C. 1973 *Political Alienation and Political Behavior.* Chicago: Aldine.
- Seeman, Melvin 1975 "Alienation Studies." In Alex Inkeles, James Coleman, and Neil Smelser, eds. *Annual Review of Sociology*, vol. 8. Palo Alto, Calif.: Annual Reviews.
- Shaffer, Stephen D. 1981 "A Multivariate Explanation of Decreasing Turnout in Presidential Elections, 1960-1976." *American Journal of Political Science* 25:68-95.
- Sigelman, Lee, and Stanley Feldman 1983 "Efficacy, Mistrust, and Political Mobilization: A Cross-National Analysis." *Comparative Political Studies* 16:118-143.
- Teixeira, Ruy A. 1996 *Economic Changes and the Middle-Class Revolt Against the Democratic Party.* In Stephen C. Craig, ed., *Broken Contract? Changing Relationships between Americans and their Government.* Boulder, Colo.: Westview Press.
- Warren, Donald I. 1976 *The Radical Center: Middle Americans and the Politics of Alienation.* Notre Dame, Ind.: University of Notre Dame Press.
- Wolfe, Alan 1977 *The Limits of Legitimacy: Political Contradictions of Contemporary Capitalism.* New York: Free Press.
- Wright, James D. 1976 *The Dissent of the Governed: Alienation and Democracy in America.* New York: Academic Press.

CLARENCE Y.H. LO

## ALTERNATIVE LIFESTYLES

Lifestyles that were considered "alternative" in the past are becoming less unusual and increasingly normative. Many people, for example, experience cohabitation, divorce, and remarriage. Other lifestyles, such as singlehood, gay and lesbian relationships, or remaining childfree may not be rising drastically in frequency, but they are less stigmatized and more visible than they were in recent decades.

It was during the 1960s and 1970s that the utility and the structure of many social institutions were seriously questioned. This included the institution of the family. What was the purpose of

family? Was it a useful social institution? Why or why not? How can it be improved? The given cultural milieu of the period, such as resurgence of the women's movement, concerns about human rights more generally, and improvements in our reproductive and contraceptive technology, exacerbated these questions. In increasing numbers individuals began to experiment with new and alternative ways in which to develop meaningful relationships, sometimes outside the confines of marriage. Literature soon abounded among both the academic community and the popular press describing and deliberating on these new lifestyles. In 1972, a special issue of *The Family Coordinator* was devoted to the subject of alternative lifestyles, with a follow-up issue published in 1975. The subject was firmly entrenched within the field of family sociology by 1980 when the *Journal of Marriage and Family* devoted a chapter to alternative family lifestyles in their decade review of research and theory.

Despite the increased visibility and tolerance for a variety of lifestyles during the 1990s, concern is voiced from some people over the "demise" of the family. The high divorce rate, increased rates of premarital sexuality, cohabitation, and extramarital sex are pointed to as both the culprits and the consequences of the deterioration of family values. Popenoe and Whitehead write about cohabitation, for example: "Despite its widespread acceptance by the young, the remarkable growth of unmarried cohabitation in recent years does not appear to be in children's or the society's best interest. The evidence suggests that it has weakened marriage and the intact, two-parent family and thereby damaged our social well-being, especially that of women and children" (1999, p. 16). What do the authors mean by "society's best interest"? And what type of family relationship would be in our "best interest"? What invariably comes to mind is the married, middle-class, traditional two-parent family in which the father works outside the home and the mother stays at home to take care of the children, at least while they are young. This monolithic model, however, excludes the majority of the population; indeed, a growing number of persons do not desire such a model even if it were attainable. It is based on the false notion of a single and uniform intimate experience that many argue has racist, sexist, and classist connotations.

This distress about the demise of the family is not particularly new. For at least a century American observers and social critics have warned against the negative consequences of changes in the family. Yes, the family has indeed changed, but the vast majority of the population, both then and now, still prefers to marry, have children, and live in a committed, monogamous relationship. The most profound changes to date have not occurred in alternatives *to* marriage but rather in alternatives *prior* to marriage, and alternative ways in *structuring* marriage itself, while keeping the basic institution and its purposes intact. For example, nonmarital sex, delayed marriage and childbearing, and cohabitation are practiced with increasing frequency and are tolerated by a larger percentage of the population than ever before. And within marriage itself, new behaviors and ideologies are becoming increasingly popular, such as greater equality between men and women (although gender equality is more an ideal than a reality in most marriages).

#### NEVER-MARRIED SINGLES

A small but growing percentage of adult men and women remain single throughout their lives. In the United States, approximately 5 percent never marry (U.S. Bureau of the Census 1998). These individuals experience life without the support and obligations of a spouse and usually children. While often stereotyped as either “swingers” or “lonely losers,” Stein reports that both categorizations are largely incorrect (1981). Instead, singles cannot be easily categorized and do not constitute a single social type. Some have chosen singlehood as a preferred option, perhaps due to career decisions, sexual preference, or other family responsibilities. Others have lived in locations in which demographic imbalances have affected the pool of eligibles for mate selection. And others have been lifelong isolates, have poor social skills, or have significant health impairments that have limited social contacts.

Attitudes toward singlehood have been quite negative historically, especially in the United States, although change has been noted. Studies report that during the 1950s, remaining single was viewed as pathology, but by the mid-1970s singlehood was not only tolerated but also viewed by many as an

avenue for enhancing one’s happiness. In the early 1990s, when asked about the importance of being married, approximately 15 percent of unmarried white males and 17 percent of unmarried white females between the ages of 19 and 35 did not agree with the statement that they “would like to marry someday.” The percentage of blacks that did not necessarily desire marriage was even higher, at 24 percent and 22 percent of black males and females, respectively. Interestingly, the gap in attitudes between males and females was the largest among Latinos, with only 9 percent of Latino males, but 25 percent of Latina females claiming that they did not necessarily want to marry (South 1993).

Despite this gender gap, single males are viewed more favorably than are single females. Males are stereotyped as carefree “bachelors,” while single women may still be characterized as unattractive and unfortunate “spinsters.” In the popular card game “Old Maid,” the game’s loser is the one who is stuck with the card featuring an old and unattractive unmarried woman. Oudijk (1983) found that the Dutch population generally affords greater lifestyle options to women, and only one-quarter of his sample of married and unmarried persons reported that married persons are necessarily happier than are singles.

Shostak (1987) has developed a typology in which to illustrate the divergence among the never-married single population. It is based on two major criteria: the voluntary versus involuntary nature of their singlehood, and whether their singlehood is viewed as temporary or stable. *Ambivalents* are those who may not at this point be seeking mates but who are open to the idea of marriage at some time in the future. They may be deferring marriage for reasons related to schooling or career, or they may simply enjoy experimenting with a variety of relationships. *Wishfuls* are actively seeking a mate but have been unsuccessful in finding one. They are, generally, dissatisfied with their single state and would prefer to be married. The *resolved* consciously prefer singlehood. They are committed to this lifestyle for a variety of reasons; career, sexual orientation, or other personal considerations. A study of 482 single Canadians reported that nearly half considered themselves to fall within this category (Austrom and

Hanel 1985). They have made a conscious decision to forgo marriage for the sake of a single lifestyle. Small but important components of this group are priests; nuns; and others who, for religious reasons, choose not to marry. Finally, *regretfuls* are those who would rather marry but who have given up their search for a mate and are resigned to singlehood. They are involuntarily stable singles.

While the diversity and heterogeneity among the never-married population is becoming increasingly apparent, one variable is suspected to be of extreme importance in explaining at least some variation: gender. Based on data gathered in numerous treatises, the emerging profiles of male and female singles are in contrast. As Bernard (1973) bluntly puts it, the never-married men represent the “bottom of the barrel,” while the never-married women are the “cream of the crop.” Single women are generally thought to be more intelligent, are better educated, and are more successful in their occupations than are single men. Additionally, research finds that single women report to be happier, less lonely, and have a greater sense of psychological well-being than do their single male counterparts.

One reason why single women are more likely to be the “cream of the crop” as compared to men is that many well-educated and successful women have difficulty finding suitable mates who are their peers, and therefore have remained unmarried. Mate-selection norms in the United States encourage women to “marry up” and men to “marry down” in terms of income, education, and occupational prestige. Thus, successful women have fewer available possible partners, because their male counterparts may be choosing from a pool of women with less education and income. A second reason for the gender difference is that some highly educated and successful women do not *want* what they interpret as the “burdens” of a husband and children. Career-oriented women are not rewarded, as are career-oriented men, for having a family. Someone who is described as a “family man” is thought to be a stable and reliable employee; there is no semantic equivalent for women. Thus, well-educated, career-oriented women may see singlehood as an avenue for their success, whereas well-educated, career-oriented men may see marriage as providing greater benefits than singlehood.

Demographers predict that the proportion of singles in our population is likely to increase slightly in the future. As singlehood continues to become a viable and respectable alternative to marriage, more adults may choose to remain single throughout their lives. Others may remain single not out of choice but due to demographic and social trends. More people are postponing marriage, and it is likely that some of these will find themselves never marrying. For example, the number of women between the ages of forty and forty-four today who have never married is double the number in 1980, at approximately 9 percent (U.S. Bureau of the Census 1998). Some of these women may marry eventually, but many will likely remain unmarried. Moreover, the increasing educational level and occupational aspirations of women, coupled with our continued norms of marital homogamy, help to ensure that the number of never-married single persons—women in particular—is likely to increase somewhat into the twenty-first century.

### COHABITATION

Cohabitation, or the sharing of a household by unmarried intimate partners, is quickly becoming commonplace in the United States. Some people suggest that it is now a normative extension of dating. According to the U.S. Bureau of the Census, the number of cohabiting couples topped 4 million in 1997, up from less than one-half million in 1960. Approximately one-half of unmarried women between the ages of twenty-five and thirty-nine have lived with a partner or are currently doing so, and over half of all first marriages are now preceded by cohabitation. Approximately one-third of these unions contain children (U.S. Bureau of the Census 1998). Cohabitation is now seen as an institutionalized component to the larger progression involving dating, courtship, engagement, and marriage.

Given the attitudes of even younger persons, we expect trends in cohabitation to continue to increase in the United States. Nearly 60 percent of a representative sample of high school seniors “agreed,” or “mostly agreed” with the statement “it is usually a good idea for a couple to live together before getting married in order to find out whether they really get along” (Survey Research Center, University of Michigan 1995).



Cohabitation is not a recent phenomenon, nor one unique to the United States. In Sweden and other Scandinavian countries, for example, cohabitation has become so common that it is considered a social institution in and of itself. It is a variant of marriage rather than of courtship; approximately 30 percent of all couples in Sweden who live together are unmarried (Tomasson 1998). People who cohabit have the same rights and obligations as do married couples with respect to taxation, inheritance, childcare, and social welfare benefits. Many of these unions have children born within them, and there is little stigma attached to being born "out of wedlock." Another study of eighty-seven Canadian couples, located through newspaper wedding announcements, reported that 64 percent of the couples had cohabited for some period, 43 percent of these for over three months. In contrast, cohabitation is relatively rare in more traditional and Roman Catholic nations such as Italy.

Cohabitors tend to differ from noncohabitators in a variety of sociodemographic characteristics. For example, cohabitators tend to see themselves as being more androgynous and more politically liberal, are less apt to be religious, are more experienced sexually, and are younger than married persons. Although cohabitators may argue that living together prior to marriage will enhance the latter relationship by increasing their knowledge of their compatibility with day-to-day living prior to legalizing the union, such optimism is generally not supported. While some studies indicate no differences in the quality of marriages among those who first cohabited and those that did not, others find that those people who cohabit have higher rates of divorce. This may, however, have nothing to do with cohabitation per se but rather may be due to other differences in the personalities and expectations of marriage between the two groups.

A wide variety of personal relationships exist among cohabiting couples. Several typologies have been created to try to capture the diversity found within these relationships. One particularly useful one, articulated by Macklin (1983), is designed to exemplify the diversity in the stability of such relationships. She discusses four types of cohabiting relationships. These include: (1) *temporary or casual* relationships, in which the couple cohabits for

convenience or for pragmatic reasons; (2) *going together*, in which the couple is affectionately involved but has no plans for marriage in the future; (3) *transitional*, which serves as a preparation for marriage; and (4) *alternative to marriage*, wherein the couple opposes marriage on ideological or other grounds.

Although attitudes toward cohabitation have become increasingly positive, especially among younger persons, Popenoe and Whitehead (1999) remind us that cohabitation was illegal throughout the United States before 1970 and remains illegal in a number of states based on a legal code outlawing "crimes against chastity" (Buunk and Van Driel 1989). These laws, however, are rarely if ever enforced. In the Netherlands, or in other countries where cohabitation is institutionalized, the majority of the population sees few distinctions between cohabitation and marriage. Both are viewed as appropriate avenues for intimacy, and the two lifestyles resemble one another much more so than in the United States in terms of commitment and stability.

The future of cohabitation, and the subsequent changes in the attitudes toward it, is of considerable interest to sociologists. Many predict that cohabitation will become institutionalized in the United States to a greater degree in the near future, shifting from a pattern of courtship to an alternative to marriage. Whether it will ever achieve the status found in other countries, particularly in Scandinavia, remains to be seen.

### CHILDFREE ADULTS

There is reason to believe that fundamental changes are occurring in the values associated with having children. As economic opportunities for women increase; as birth control, including abortion, becomes more available and reliable; and as tolerance increases for an array of lifestyles, having children is likely to become increasingly viewed as an option rather than a mandate. Evidence is accumulating to suggest that both men and women are reevaluating the costs and benefits of parenthood. Approximately 9 percent of white and African-American women and 6 percent of Latina women indicate that they would like to have no children (U.S Bureau of the Census 1998).

This trend is occurring not only in the United States but in many industrialized countries in Europe as well. The decline in childbearing there has been referred to as the “second demographic transition” (Van de Kaa 1987). Davis (1987) posits that features of industrial societies weaken the individual’s desire for children. He lists several interrelated traits of industrialization, including the postponement of marriage, cohabitation, and high rates of divorce, claiming that these trends decrease the need for both marriage and childbearing.

Remaining childfree is not a new phenomenon, however. In 1940, for example, 17 percent of married white women between the ages of thirty-five and thirty-nine were childfree. Some of these women were simply delaying parenthood until their forties; however, many remained childfree. This percentage began to drop considerably after World War II, and by the late 1970s only 7 percent of women in the thirty-five to thirty-nine age group did not have children. Today the figure has risen to over 13 percent among this age group (U.S. Bureau of the Census 1998). This increase is due to a multitude of factors: delayed childbearing, infertility, and voluntary childlessness.

An important distinction to make in the discussion of childlessness is whether the decision was voluntary or involuntary. *Involuntary* childlessness involves those who are infecund or subfecund. For them, being childfree is not a choice. Unless they adopt or create some other social arrangement, they are inevitably committed to this lifestyle. *Voluntary* childlessness, the focus of this discussion, involves those who choose to remain childfree. Large differences exist within members of this group; *early articulators* have made their decision early in their lives and are committed to their choice. *Postponers*, on the other hand, begin first by delaying their childbearing, but wind up being childfree due to their continual postponement. Early articulators generally exhibit less stereotypical gender roles, are more likely to cohabit, and enjoy the company of children less than do postponers. Seccombe (1991) found that among married persons under age forty who have no children, wives are more likely than their husbands to report a preference for remaining childfree (19 percent and 13 percent, respectively).

Despite increasing rates of voluntary childlessness, most research conducted within the United States

documents the pervasiveness of pronatalist sentiment. Those who voluntarily opt to remain childfree are viewed as selfish, immature, lonely, unfulfilled, insensitive, and more likely to have mental problems than are those who choose parenthood. In the past, females, persons with less education, those with large families of their own, Catholics, and residents of rural areas were most apt to judge the childfree harshly. However, more recently, data from a nationally representative sample suggest that women are more likely to want to remain childfree than are men (Seccombe 1991).

Most studies report that those persons who opt to remain childfree are well aware of the sanctions surrounding their decision yet are rarely upset by them (see Houseknecht 1987 for a review). In her review of twelve studies, Houseknecht found only three that reported that childfree individuals had trouble dealing with the reaction from others. Sanctions apparently are not strong enough to detract certain persons from what they perceive as the attractiveness of a childfree lifestyle. Houseknecht (1987), in a content analysis of twenty-nine studies reporting the rationales for remaining childfree, identified nine primary motivations. These are, in order of the frequency in which they were found: (1) freedom from child-care responsibilities: greater opportunity for self-fulfillment and spontaneous mobility; (2) more satisfactory marital relationship; (3) female career considerations; (4) monetary advantages; (5) concern about population growth; (6) general dislike of children; (7) negative early socialization experience and doubts about the ability to parent; (8) concern about physical aspects of childbirth and recovery; and (9) concern for children given world conditions. Gender differences were evidenced in a number of areas. Overall, females were more likely to offer altruistic rationales (e.g., concern about population growth, doubts about the ability to parent, concern for children given world conditions). The male samples, conversely, were more apt to offer personal motives (e.g., general dislike of children, monetary advantages).

The consequences of large numbers of persons in industrialized societies forgoing parenthood are profound. For example, the demographic structure in many countries is in the process of radical change; populations are becoming increasingly aged. More persons are reaching old age than ever before, those persons are living longer,

and birth rates are low. The cohort age eighty-five or older, in fact, is the fastest-growing cohort in the United States. The question remains: Who will care for the elderly? Some Western European countries provide a variety of services to assist elderly persons in maintaining their independence within the community as long as possible. But social policies in other countries, including the United States, rely heavily on adult children to provide needed care to elderly parents. Formal support services, when available, tend to be uncoordinated and expensive. The question of who will provide that needed care to the large numbers of adults who are predicted to have no children has yet to be answered.

### GAY AND LESBIAN RELATIONSHIPS

Not long ago homosexuality was viewed by many professionals as an illness or a perversion. It was only as recently as 1973, for example, that the American Psychiatric Association removed homosexuality from its list of psychiatric disorders. Today, due in large part to the efforts of researchers such as Kinsey and associates (1948, 1953), Masters and Johnson (1979), and to organizations such as the Gay Liberation Front during the late 1960s, homosexuality is slowly but increasingly being viewed as a lifestyle rather than an illness. The work of Kinsey and associates illustrated that a sizable minority of the population, particularly males, had experimented with same-sex sexual relationships, although few considered themselves exclusively homosexual. Thirty-seven percent of males, they reported, had experienced at least one homosexual contact to the point of orgasm, although only 4 percent were exclusively homosexual. Among females, 13 percent had a same-sex sexual contact to the point of orgasm, while only 2 percent were exclusively homosexual in their orientation.

Obviously not all people who have had a homosexual experience consider themselves to be gay or lesbian. Most do not. One national probability sample of adult males interviewed by telephone found that 3.7 percent reported to have either a homosexual or bisexual identity (Harry 1990). Others suggest that the percentages are higher, that perhaps 3 to 5 percent of adult women and 5 to 10 percent of adult men are exclusively lesbian or gay (Diamond 1993).

Cross-cultural evidence suggests that the majority of cultures recognize the existence of homosexual behavior, particularly in certain age categories such as adolescence, and most are tolerant of homosexual behavior. Culturally speaking, it is rare to find an actual *preference* for same-sex relations; a preference tends to occur primarily in societies that define homosexuality and heterosexuality as mutually exclusive, as in many industrial countries.

There still remains a tremendous degree of hostility to homosexual relationships by large segments of society. Many regions in the United States, particularly in the South and in the West, still have laws barring homosexual activity among consenting adults. The results of national polls indicate that the majority of adults believe that homosexuals should still be restricted from certain occupations, such as elementary-school teacher, and should not be allowed to marry. Gays and lesbians report that their sexual orientation has caused a variety of problems in securing housing and in the job market. They often report that negative comments or acts of violence have been levied on them in public.

This contrasts sharply with the view toward homosexuality in the Scandinavian countries. In 1989 Denmark lifted the ban on homosexual marriages, the first country to do so. Norway and Sweden followed suit in the 1990s. These changes extend to gays and lesbians the advantages that heterosexual married couples experience with respect to inheritance, taxation, health insurance, and joint property ownership.

There is a growing amount of research illuminating various aspects of homosexual relationships, such as gender roles; degree of commitment; quality of relationships; and the couples' interface with other relationships, such as children, ex-spouses, or parents. However, because of unique historical reactions to gays and lesbians, and the different socialization of men and women in our society, it is important to explore the nature of lesbian and gay male relationships separately. Gender differences emerge in homosexual relations within a variety of contexts; for example, lesbians are more apt to have monogamous, stable relationships than are gay men, although the popular stereotype of gays as sexually "promiscuous"

has been exaggerated. The majority of gay men, just like lesbians, are interested in monogamous, long-term relationships. The lack of institutional support for gay and lesbian relationships and the wide variety of obstacles not encountered among heterosexuals, such as prejudice and discriminatory behavior, take their toll on these relationships, however.

The AIDS epidemic has had an enormous impact on the gay subculture. While the impact on lesbians is significantly less, they have not been untouched by the social impact of the devastating medical issue, despite the slow response of the world's governments.

### **DIVORCE AND REMARRIAGE**

Throughout most of history in America, the primary reason that marriages ended was because of death. In 1970, the trends were reversed, and for the first time in our nation's history, more marriages ended by divorce than by death (Cherlin 1992). The United States now has the highest rate of divorce in the world, at approximately 20 divorces per 1,000 married women aged fifteen and over in the population (U.S. Bureau of the Census 1998). This is twice as high as the divorce rate found in Canada, four times that of Japan, and ten times as high as Italy. But, it's important to note that the divorce rate has actually leveled off, or even declined somewhat in the United States after peaking in the early 1980s.

Who divorces? Research has shown that those at greatest risk for divorcing are people who marry young, especially after only a brief dating period; those who have lower incomes, although very high-earning women are also more likely to divorce; African Americans, who are about 25 percent more likely to divorce than whites; people who have been divorced before or whose parents divorced; and those who are not religious or claim no religious affiliation. The likelihood of divorce is particularly high among couples who marry in their teens because of an unplanned pregnancy. Women are almost twice as likely as men to petition for divorces, reflecting the fact that women are more often dissatisfied with their marriages than are men.

There are a variety of reasons why the rate of divorce has increased so dramatically in many

Western nations during the twentieth century: (1) there is increasing emphasis on individualism and individual happiness over the happiness of the group—or a spouse and children; (2) divorce is more socially acceptable and less stigmatized than in the past; (3) divorce is easier to obtain, as “fault” is generally no longer required; (4) women are less financially dependent on men, on average, and therefore can end an unhappy marriage more easily; (5) there is an increase in the number of adult children who grew up in divorced households, who are more likely to see divorce as a mechanism to end an unhappy marriage; and (6) today's marriages experience increased stress, with outside work consuming the time and energy people used to devote to their marriages and families.

One of the consequences of divorce is that many children will live at least a portion of their lives in single-parent households. Single-parent households are becoming increasingly normative, with approximately half of all children under age eighteen spending some portion of their lives living with only one parent, usually their mother. Single-parent households are more likely to be poor and often lack the social capital available in two-parent households, and consequently place the child at greater risk for a variety of negative social and health outcomes. Commonly the absentee parent does not pay the child support that is due, and fails to see the children with regularity.

Although divorce has become common in many industrialized nations, it would be incorrect to assume that this represents a rejection of marriage. Four out of five people who divorce remarry, most often within five years. Men are more likely to remarry than are women. This difference is due to the greater likelihood that children will be living with their mothers full-time, rather than their fathers; the cultural pattern of men marrying younger women; and the fact that there are fewer men than women in the population in general.

Remarriage often creates “blended families” composed of stepparents and possibly stepsiblings. It is estimated that approximately one-third of all children will live with a stepparent for at least one year prior to reaching age eighteen. Despite the increasing commonality of this type of family, it has been referred to as an “incomplete institution” because the social expectations are less clear

than in other family structures. There are no well-established rules for how stepparents and children are supposed to relate, or the type of feelings they should have for one another. Without a clear-cut set of norms, blended families may be seen as "alternative lifestyles," but yet they are becoming increasingly common in modern industrialized societies where divorce is common.

### CONCLUDING COMMENTS

There is considerable accumulating evidence to suggest that family lifestyles are becoming more varied and that the public is becoming increasingly tolerant of this diversity. The data indicate that traditional marriage itself *per se* may be less important in sanctioning intimacy. The review by Buunk and Hupka (1986) of seven countries reveals that individualism, as expressed by following one's own personal interests in intimate relationships, was more prevalent in affluent democratic countries such as the United States and in most of Western Europe than in poorer and nondemocratic nations.

This does not mean, however, that people are discarding the institution of marriage. In the United States, as elsewhere, the vast majority of the population continues to endorse marriage and parenthood in general, and for themselves personally. Most still plan to marry and have children, and optimism remains high that theirs will be a lasting union despite high national frequencies of divorce.

Alternative lifestyles are not replacing marriage. Instead, they are gaining acceptance because they involve, for some, modifications of the family structure as an adaptation to changing conditions in society. The lifestyles discussed here, as well as others, reflect the broader social changes in values, relationships, and even technology that are found within society as a whole. As Macklin notes, the family is not disappearing, but "continuing its age-old process of gradual evolution, maintaining many of its traditional functions and structures while adapting to changing economic circumstances and cultural ideologies" (1987, p. 317). This process has merely accelerated during the past several decades, and these changes have caught the attention of the general public. College classes and their corresponding textbooks within this discipline of sociology are still often titled *Marriage and the*

*Family*, as if there were only one model of intimacy. Yet perhaps a more appropriate title would be *Marriages, Families, and Intimate Relationships*. This would reflect not only the diversity illustrated here but would also acknowledge the tremendous ethnic and class variations that make for rich and meaningful personal relationships.

(SEE ALSO: *American Families; Courtship; Divorce and Remarriage; Marriage; Mate Selection Theories; Sexual Orientation*)

### REFERENCES

- Austrom, Douglas, and Kim Hanel 1985 "Psychological Issues of Single Life in Canada: An Exploratory Study." *International Journal of Women's Studies* 8:12-23.
- Bernard, Jessie 1973 *The Future of Marriage*. New York: Bantam Books.
- Buunk, Bram P., and R. B. Hupka 1986 "Autonomy in Close Relationships: A Cross-Cultural Study." *Family Perspective* 20:209-221.
- , and Barry Van Driel 1989 *Variant Lifestyles and Relationships*. Newbury Park, Calif.: Sage Publications.
- Cherlin, Andrew 1992 *Marriage, Divorce, Remarriage*, rev. ed. Cambridge Mass.: Harvard University Press.
- Davis, Kingsley 1987 "Low Fertility in Evolutionary Perspective." In K. Davis, M. S. Bernstam, and R. Ricardo-Campbell, eds., *Below Replacement Fertility in Industrial Societies*. Cambridge, Eng.: Cambridge University Press.
- Diamond, M. 1993 "Homosexuality and Bisexuality in Different Populations." *Archives Of Sexual Behavior*, 22:291-310.
- Harry, Joseph 1990 "A Probability Sample of Gay Men." *Journal of Homosexuality* 19:89-104.
- Houseknecht, Sharon K. 1987 "Voluntary Childlessness." In M. B. Sussman and S. K. Steinmetz, eds., *Handbook of Marriage and the Family*. New York: Plenum Press.
- Kinsey, Alfred, W. Pomeroy, P. H. Gebhard, and C. E. Martin 1953 *Sexual Behavior in the Human Female*. Philadelphia: W. B. Saunders.
- Kinsey, Alfred, W. Pomeroy, and C. E. Martin 1948 *Sexual Behavior in the Human Male*. Philadelphia: W. B. Saunders.
- Macklin, Eleanor D. 1987 "Nontraditional Family Forms." In M. B. Sussman and S. K. Steinmetz, eds., *Handbook of Marriage and the Family*. New York: Plenum Press.
- 1983 "Nonmarital Heterosexual Cohabitation: An Overview." In E. D. Macklin and R. H. Rubin,

- eds., *Contemporary Families and Alternative Lifestyles: Handbook on Research and Theory*. Beverly Hills, Calif.: Sage Publications.
- Masters, William H., and Virginia E. Johnson 1979 *Homosexuality in Perspective*. Boston: Little, Brown.
- Oudijk, C. 1983 *Social Atlas of Women* (in Dutch). The Hague: Staatsuitgeverij.
- Popenoe, David, and Barbara Dafoe Whitehead 1999 *Should We Live Together? What Young Adults Need to Know About Cohabitation Before Marriage*. New Brunswick, N.J.: The National Marriage Project.
- Seccombe, Karen 1991 "Assessing the Costs and Benefits of Children: Gender Comparisons among Child-free Husbands and Wives." *Journal of Marriage and the Family* 53:191-202.
- Shostak, Arthur B. 1987 "Singlehood." In M. B. Sussman and S. K. Steinmetz, eds., *Handbook of Marriage and the Family*. New York: Plenum Press.
- South, Scott 1993 "Racial and Ethnic Differences in the Desire to Marry." *Journal Of Marriage and the Family* 55:357-370.
- Stein, Peter J. 1981 *Single Life: Unmarried Adults in Social Context*. New York: St. Martin's Press.
- Survey Research Center, University of Michigan 1995 *Monitoring the Future Survey*. Ann Arbor, Mich.: Survey Research Center.
- Tomasson, Richard F. 1998 "Modern Sweden: The Declining Importance of Marriage." *Scandinavian Review* August, 1998:83-89.
- United States Bureau of the Census 1998 *Statistical Abstract of the United States, 1998* 118th ed. Washington, D.C.: U.S. Bureau of the Census.
- Van de Kaa, Dick J. 1987 "Europe's Second Demographic Transition." *Population Bulletin* 42:1-57.

KAREN SECCOMBE

## ALTRUISM

Three terms are commonly used in the broad research area that investigates positive interpersonal action: prosocial behavior, helping behavior, and altruism. "Prosocial behavior" is the broadest of the three; it refers to any behavior that can be construed as consistent with the norms of a given society. Thus, murder, when enacted on behalf of one's country on a battlefield, is as prosocial a behavior as intervening to prevent a crime. "Helping behavior" refers simply to any behavior that

provides some benefit to its recipient. "Altruism" is the narrowest of the three concepts. Altruism is behavior that not only provides benefits to its recipient but also provides no benefits to the actor and even incurs some costs. If one conceives of psychological rewards as benefits to the actor, this definition of altruism is so narrow that it excludes virtually all human behavior. Hence, many social psychologists maintain simply that altruistic behavior need exclude only the receipt of material benefits by the actor. Some theorists require as part of the definition that the act be motivated "with an ultimate goal of benefiting someone else" (Batson 1991, p. 2), but do not rule out the incidental receipt of benefits by the actor.

Related terms include philanthropy, charity, volunteering, sharing, and cooperating. Philanthropy and charity have largely come to mean donation of money or material goods. Volunteering, similarly, generally refers to giving time for the ultimate purpose of benefiting others, under the aegis of some nonprofit organization. Sharing and cooperating refer to coordinated actions among members of a group or collectivity in the service of better outcomes for the group as a whole. All of these terms may be subsumed under the generic term "prosocial behavior," and often under "helping behavior," although they would seldom meet the stringent criteria for altruism.

## HISTORY

The origins of the contemporary study of altruism have been traced back to August Comte, who explored the development of altruism and "sympathetic instincts." The existence of an altruistic instinct was emphasized in McDougall's *Introduction to Social Psychology* (1908) but argued against by the naturalistic observational research of Lois Murphy (1937). Early symbolic interactionists attributed altruistic behavior to the capacity to "take the role of the other"—to imagine oneself in another person's situation (Mead 1934). The developmental study of altruism has built on the theoretical work of Piaget (1932), who explored stages in the development of sharing behavior, as well as on the work of Kohiberg (1969) on the development of moral judgment. Hartshorne and May conducted one of the earliest series of empirical studies (1928-30), focusing on honesty and

altruism in children. Sorokin (1950, 1970[1950]) wrote extensively on love and altruism, and carried out the first empirical work on informal helping and volunteering by studying individuals nominated by others as “good neighbors.” It is only since the mid-1960s, however, that altruism has been extensively examined through systematic research.

Most social psychology textbooks attribute this upsurge of interest in altruism and helping behavior to the murder of Kitty Genovese in 1964 and the failure of the thirty-nine witnesses to intervene. The subject of widespread media coverage, this incident motivated Latané and Darley’s experimental investigations of bystander inaction, published in *The Unresponsive Bystander: Why Doesn’t He Help?* (1970). During the 1970s, helping behavior became one of the most popular topics in social psychological research, although this emphasis declined considerably through the 1980s and 1990s (see Batson 1998 for figures on the number of published studies by decade). Because of this beginning, the vast majority of the studies deal with intervention in the momentary problems of strangers. Only since the 1990s has there been much attention to informal and formal volunteering, charitable donation, and blood donation. Virtually all textbooks now have a chapter on altruism and helping behavior, and a number of books on the topic have been published in the past three decades.

### THEORIES OF ALTRUISM AND HELPING BEHAVIOR

Helping behavior has been explained within a variety of theoretical frameworks, among them evolutionary psychology, social learning, and cognitive development. One sociobiological approach maintains that helping behavior and altruism have developed through the selective accumulation of behavioral tendencies transmitted genetically. Three mechanisms have been suggested: kin selection, reciprocal altruism, and group selection (see Sober and Wilson 1998). These mechanisms explain the evolution of altruistic behavior as a function of, in turn: the greater likelihood that altruists would save kin, through perpetuating an altruistic gene shared among them; a tendency to help others who have engaged in helpful acts, presumably based on a reciprocity gene; and the greater

likelihood of survival of entire groups that include a higher proportion of altruists. A second sociobiological theory maintains that helping behavior has developed through sociocultural evolution, the selective accumulation of behavior retained through purely social modes of transmission. (See Krebs and Miller [1985] for an excellent review of this literature.) The cognitive-developmental approach to the development of helping behavior in children emphasizes the transformation of cognitive structures and experiential role-taking opportunities as determinants. Social learning theory explains altruism and helping behavior as learned through interaction with the social environment, mainly through imitation and modeling, but also through reinforcement. Reflecting the same behaviorist principles, exchange theory suggests that individuals perform helping acts while guided by the principles of maximizing rewards and minimizing costs. Helping behavior is instrumental in acquiring rewards that may be material, social, or even self-administered. A more explicitly sociological framework suggests that individuals help out of conformity to social norms that prescribe helping. Three norms have received special attention: the norm of giving, which prescribes giving for its own sake; the norm of social responsibility, which prescribes helping others who are dependent; and the norm of reciprocity, which prescribes that individuals should help those who have helped them.

Reflecting the contemporary social psychological emphasis on cognition, several decision-making models of helping behavior have guided much of the research into adult helping behavior (Latané and Darley 1970; Piliavin et al. 1981; Schwartz and Howard 1981). These models specify sequential decisions that begin with noticing a potential helping situation and end with a decision to help (or not). Research has focused on identifying those personality and situational variables that influence this decision-making process and specifying how they do so. There also has begun to be more attention to the social and sociological aspects of helping—to the context in which helping occurs, to the relationship between helper and helped, and to structural factors that may affect these interactions (Gergen and Gergen 1983; Callero 1986). A very active focus of work, mainly identified with Batson (see, e.g., Batson 1991) and Cialdini and colleagues, has been the attempt to

demonstrate (or refute) the existence of “true” altruistic motivations for helping.

### RESEARCH ON ALTRUISM AND HELPING BEHAVIOR

**Person Variables.** There has been an extensive and confused debate, due both to definitional and measurement problems, about the existence of an altruistic personality (see Schroeder et al. 1995). There is now good evidence of a pattern of prosocial personality traits that characterize individuals whose behavior involves long-term, sustained forms of helping behavior (e.g., community mental health workers, see Krebs and Miller 1985; volunteers who work with AIDS patients, see Penner et al. 1995; Penner and Finkelstein 1998). The traits that make up the prosocial personality include empathy, a sense of responsibility, concern for the welfare of others, and a sense of self-efficacy. With regard to helping in emergencies, the evidence is stronger for person by situation effects; that is, interactions of characteristics of both individuals and situations that influence helping in emergencies. For example, self-confidence and independence can predict differentially how individuals will behave in emergency situations when there are others present or when the person is alone (Wilson 1976). And Batson and his colleagues have found that prosocial personality characteristics correlate with helping, but only when helping is egoistically motivated, not when true altruism is involved. The general proposal that individual difference factors are most effective when situational pressures are weak seems generally applicable in the helping area.

Internalized values as expressed in personal norms have also been shown to influence helping. Personal norms generate the motivation to help through their implications for self-based costs and benefits; behavior consistent with personal norms creates rewards such as increased self-esteem, whereas behavior that contradicts personal norms generates self-based costs such as shame. This influence has been demonstrated in high-cost helping such as bone marrow donation (Schwartz 1977). Other personality correlates of helping are less directly related to the costs and benefits of the helping act itself. For example, information-processing styles such as cognitive complexity influence helping.

Clary and Snyder (1991) have pursued a functional approach to understanding motivations for helping. They have developed a questionnaire measure that distinguishes six potential motives for long-term volunteering (e.g. value expression, social motivation, career orientation) and have demonstrated both predictive and discriminant validity for the instrument (Clary, Snyder et al. 1998). In one study, they showed that it was not the more purely altruistic motivations that predicted long-term commitment. Temporary emotional states or moods may also affect helping. A series of studies by Isen (1970) and her colleagues demonstrate that the “glow of good will” induces people to perform at least low-cost helping acts such as helping someone pick up a pile of dropped papers, and research by Cialdini and colleagues has shown that helping can be motivated by the need to dispel a bad mood.

**Situation Variables.** Characteristics of the situation also influence the decision to help. The salience and clarity of a victim’s need influence both the initial tendency to notice need and the definition of the perceived need as serious. Salience and clarity of need increase as the physical distance between an observer and a victim decreases; thus, victims of an emergency are more likely to be helped by those physically near by. Situational cues regarding the seriousness of another’s need influence whether need is defined as serious enough to warrant action. Bystanders are more likely to offer aid when a victim appears to collapse from a heart attack than from a hurt knee, for example, presumably because of perceived seriousness. The presence of blood, on the other hand, can deter helping, perhaps because it suggests a problem serious enough to require medical attention.

One of the most strongly supported findings in the area of helping is that the number of others present in a potential helping situation influences an individual’s decision to help. Darley and Latané (1968) demonstrated experimentally that the higher the number of others present, the lower the chance of any one individual helping. One process underlying this effect involves the diffusion of responsibility: the higher the number of potential helpers, the less any given individual perceives a personal responsibility to intervene. The presence of an individual who may be perceived as having special competence to help also reduces the felt



responsibility of others to help. Thus, when someone in medical clothing is present at a medical emergency, others are less likely to help. A second process underlying the effect, when bystanders can see each other, involves definition of the situation. If no one moves to intervene, the group may collectively provide a social definition for each other that the event is not one that requires intervention.

**Social Variables.** Research has also demonstrated the influence of other social variables on helping. Darley and Latané (1968) showed experimentally that people were more likely to provide help in an emergency in the presence of a friend rather than in the presence of strangers. They reasoned that in emergency situations in which a friend does not respond, one is not likely to attribute this to lack of concern, but rather will seek other explanations. In addition, bystanders who are acquainted are more likely to talk about the situation. Thus, preexisting social relationships among bystanders affect helping. Individuals are also more likely to help others who are similar to them, whether in dress style or in political ideology. The perceived legitimacy of need, a variable defined by social norms, also affects rates of helping. In one field study of emergency intervention, bystanders were more likely to help a stranger who collapsed in a subway car if the distress was attributed to illness rather than to drunkenness (Piliavin, Rodin, and Piliavin 1969).

Other demographic variables, such as sex, age, socioeconomic status, and race have also been investigated. Race appears to affect helping mainly when the costs for helping are relatively high, or when failure to help can be attributed to factors other than prejudice (Dovidio 1984). In the study cited above (Piliavin et al. 1969) the rate of helping by white and black bystanders was unrelated to the race of the victim who appeared to be ill, but help offered to the drunk was almost always by people of his own race. Females are usually helped more than males, but who helps more depends heavily on the nature of the help required. Males tend to help females more than they help males, whereas females are equally helpful to females and males (see Piliavin and Unger 1985). This pattern may reflect stereotypic gender roles: Females are stereotyped as dependent and weaker than males. Other studies of the effect of social statuses on helping

indicate, consistent with social categorization theory, that members of one's own group tend to be helped more than outgroup members. Studies of reactions following natural disasters show that people tend to give aid first to family members, then to friends and neighbors, and last to strangers.

### THE SOCIOLOGICAL CONTEXT OF HELPING

Gergen and Gergen (1983) call for increased attention to the social structural context of helping and to the interactive history and process of the helping relationship (see also Piliavin and Charng 1990). Social structure is clearly important as a context for helping. Social structure specifies the pool of social roles and meaning systems associated with any interaction (Callero, Howard, and Piliavin 1987). Social structure also influences the distribution of resources that may be necessary for certain helping relationships. One needs money to be able to donate to a charity and medical expertise to be able to help earthquake victims. Wilson and Musick (1997) have presented data in support of a model using both social and cultural capital as predictors of involvement in both formal and informal volunteering. Social structure also determines the probability of both social and physical interaction among individuals and thus influences the possibility of helping.

Interaction history is also crucial to understanding helping behavior. If a relationship has been positive and mutually supportive, this context suggests that beneficial actions should be defined as helping. If a relationship has been characterized by competition and conflict, this context does not support defining beneficial action as helping. In this case, alternative, more self-serving motivations may underlie helping. Thus the provision of U.S. foreign aid to countries with which the United States has had conflict is often viewed as a strategic tool, whereas when such aid has been provided to countries with which the United States has had positive relationships, it is viewed generally as genuine helping. Such patterns illustrate this influence of interaction history on the interpretation of helping behavior.

Another sociological approach emphasizes helping as role behavior and is guided by Mead's

(1934) conception of roles as patterns of social acts framed by a community and recognized as distinct objects of the social environment. Roles define individual selves and thus also guide individual perception and action. Helping behavior has been shown to express social roles. A series of studies of blood donors (Callero, Howard, and Piliavin 1987; Piliavin and Callero 1991) demonstrate that role-person merger (when a social role becomes an essential aspect of self) predicts blood donation, independent of the effects of both personal and social norms, and is more strongly associated with a history of blood donation than are social or personal norms. This study demonstrates the importance of helping for self-validation and reproduction of the social structure as expressed in roles. More recent research has shown similar effects for identities tied to volunteering time and giving money (Grube and Piliavin in press; Lee, Piliavin, and Call in press). This attention to concepts such as roles, interaction history, and social structure is evidence of the sociological significance of altruism and helping behavior.

#### CROSS-CULTURAL RESEARCH IN ALTRUISM

Until the last few decades, little work had been done systematically comparing altruism and helping behavior across cultures. Beginning in the 1970s, researchers have compared helping in rural and urban areas, rather consistently finding that helping of strangers, although not of kin, is more likely in less densely populated areas all around the world. In a real sense, urban and rural areas have different "cultures"; small towns are more communal or collective, while cities are more individualistic. A review of other cross-cultural comparisons (Ting and Piliavin forthcoming) examines similarities and differences not only in the helping of strangers but also in the development of moral reasoning, socialization of prosocial behavior, and participation in "civil society." The collectivism-individualism distinction across societies provides a good organizing principle for understanding many of the differences that are found. Not only do societies differ in the level of helping, but in the pattern. For example, in communal societies, the difference between the amount of help offered to ingroup and outgroup members is exaggerated in comparison with the more individualistic societies.

#### ALTRUISM RESEARCH IN OTHER FIELDS

Scholars from many fields other than social psychology have also addressed the question of altruism. The long debate in evolutionary biology regarding the possibility that altruism could have survival value appears to have been answered in the affirmative (Sober and Wilson 1998). Some authors (e.g. Johnson 1986; Rushton 1998) in fact view patriotism or ethnic conflict, or both, as rooted in altruism fostered by kin selection. Game theorists have discovered that in repeated prisoner's dilemma games and public goods problems, some individuals consistently behave in more cooperative or altruistic ways than do others (Liebrand 1986). Even economists and political scientists, who have long held to the belief that all motivation is essentially selfish, have begun to come to grips with evidence (such as voting behavior and the public goods issue) that indicates that this is not true (see Mansbridge 1990; Clark 1998).

**Recommended reading.** The interested reader is referred Schroeder et al., *The Psychology of Helping and Altruism* (1995) for a relatively nontechnical overview of the field, or Batson, "Altruism and Prosocial Behavior" (1998) for a briefer, more technical approach emphasizing work demonstrating that "true altruism" can be a motivation for helping. For an excellent examination of approaches to the topic of altruism by economists and political scientists, read Mansbridge, *Beyond Self-Interest* (1990). For an engaging read on the topic of both evolutionary and psychological altruism, try Sober and Wilson's *Unto Others* (1998). Finally, for a view toward the practical application of ideas from altruism research, read Oliner et al., *Embracing the Other* (1992). (Full citations for these works are in the references that follow.)

(SEE ALSO: *Social Psychology*)

#### REFERENCES

- Batson, C. Daniel 1998 "Altruism and Prosocial Behavior." In Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed., Boston, Mass.: McGraw-Hill.
- 1991 *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, N.J.: Erlbaum.
- Callero, Peter L. 1986 "Putting the Social in Prosocial Behavior: An Interactionist Approach to Altruism." *Humboldt Journal of Social Relations* 13:15-32.

- , Judith A. Howard, and Jane A. Piliavin 1987 "Helping Behavior as Role Behavior: Disclosing Social Structure and History in the Analysis of Prosocial Action." *Social Psychology Quarterly* 50:247–256.
- Cialdini, Robert B., Donald J. Baumann, and Douglas T. Kenrick 1981 "Insights from Sadness: A Three-Step Model of the Development of Altruism as Hedonism." *Developmental Review* 1:207–223.
- Clark, Jeremy 1998 "Fairness in Public Good Provision: an Investigation of Preferences for Equality and Proportionality." *Canadian Journal of Economics* 31:708–729.
- Clary, E. Gil, and Mark Snyder 1991 "Functional Analysis of Altruism and Prosocial Behavior: the Case of Volunteerism." In M.S. Clark, ed., *Prosocial Behavior* Newbury Park, Calif.: Sage.
- , Robert D. Ridge, John Copeland, Arthur A. Stukas, Julie Haugen, and Peter Miene 1998 "Understanding and Assessing the Motivations of Volunteers: A Functional Approach." *Journal of Personality and Social Psychology* 74:1516–1530.
- Darley, John M., and Bibb Latané 1968 "Bystander Intervention in Emergencies: Diffusion of Responsibility." *Journal of Personality and Social Psychology* 8:377–383.
- Dovidio, John F. 1984 "Helping Behavior and Altruism: An Empirical and Conceptual Overview." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 17. New York: Academic.
- Gergen, Kenneth J., and Mary M. Gergen 1983 "Social Construction of Helping Relationships." In J. F. Fisher, A. Nadler, and B. M. DePaulo, eds., *New Directions in Helping*, vol. I. New York: Academic.
- Grube, Jean, and Jane A. Piliavin forthcoming "Role Identity and Volunteer Performance." *Personality and Social Psychology Bulletin*.
- Hartshorne, H., and M. A. May 1928–30 *Studies in the Nature of Character*, vols. 1–3. New York: Macmillan.
- Isen, Alice M. 1970 "Success, Failure, Attention and Reaction to Others: The Warm Glow of Success." *Journal of Personality and Social Psychology* 15:294–301.
- Johnson, G. R. 1986 "Kin Selection, Socialization, and Patriotism: An Integrating Theory." *Politics and the Life Sciences* 4:127–154.
- Kohlberg, Lawrence 1969 "Stage and Sequence: The Cognitive-Developmental Approach to Socialization." In D. Goslin, ed., *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- Krebs, Dennis L., and Dale T. Miller 1985 "Altruism and Aggression." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*, 3rd ed., vol. 2. Hillsdale, N.J.: Erlbaum.
- Latané, Bibb, and John M. Darley 1970 *The Unresponsive Bystander: Why Doesn't He Help?* New York: Appleton-Crofts.
- Lee, Lichang, Jane A. Piliavin, and Vaughn Call forthcoming "Giving Time, Money, and Blood: Similarities and Differences." *Social Psychology Quarterly*.
- Liebrand, Wim B.G. "The Ubiquity of Social Values in Social Dilemmas." In Henk A.M. Wilke, David M. Messick, and Christel G. Rutte, eds., *Experimental Social Dilemmas*, 113–133. Frankfurt am Main: Verlag Peter Lang.
- Mansbridge, Jane J. (ed.) 1990 *Beyond Self-Interest*. Chicago: University of Chicago Press.
- McDougall, William 1908 *Introduction to Social Psychology*. London: Methuen.
- Mead, George Herbert 1934 *Mind, Self and Society from the Standpoint of a Social Behaviorist*, edited by C. W. Morris. Chicago: University of Chicago Press.
- Murphy, Lois B. 1937 *Social Behavior and Child Personality: An Exploratory Study of Some Roots of Sympathy*. New York: Columbia University Press.
- Oliner, Pearl M., Samuel P. Oliner, Lawrence Baron, Lawrence A. Blum, Dennis L. Krebs, and M. Zuzanna Smolenska 1992 *Embracing the Other: Philosophical, Psychological, and Historical Perspectives on Altruism*. New York: New York University Press.
- Oliner, Samuel P., and Pearl M. Oliner 1988 *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press.
- Penner, Louis A., and Marcia A. Finkelstein 1998 "Dispositional and Structural Determinants of Volunteerism." *Journal of Personality and Social Psychology* 74:525–537.
- , Barbara A. Fritzsche, J. Philip Craiger, and Tamara S. Freifeld 1995 "Measuring the Prosocial Personality." In James N. Butcher, Charles D. Spielberger, et al., eds., *Advances in Personality Assessment*, vol. 10. Hillsdale, N.J.: Erlbaum.
- Piaget, Jean 1932 *The Moral Judgment of the Child*. London: Routledge and Kegan Paul.
- Piliavin, Irving M., Judith Rodin, and Jane A. Piliavin 1969 "Good Samaritanism: An Underground Phenomenon?" *Journal of Personality and Social Psychology* 13:289–299.
- Piliavin, Jane A., and Peter L. Callero 1991 *Giving Blood: The Development of an Altruistic Identity*. Baltimore: Johns Hopkins University Press.
- Piliavin, Jane A., and Hong-wen Charng 1990 "Altruism: A Review of Recent Theory and Research." *Annual Review of Sociology* 16:25–65.

- Piliavin, Jane A., John F. Dovidio, Samuel Gaertner, and Russell D. Clark 1981 *Emergency Intervention*. New York: Academic.
- Piliavin, Jane A., and Rhoda Kesler Unger 1985 "The Helpful but Helpless Female: Myth or Reality?" In Virginia E. O'Leary, Rhoda K. Unger, and Barbara S. Wallston, eds., *Women, Gender, and Social Psychology*. Hillsdale, N.J.: Erlbaum.
- Rushton, J. Philippe 1998 "Genetic Similarity Theory and the Roots of Ethnic Conflict." *The Journal of Social, Political, and Economic Studies* 23:477-486.
- Schroeder, David A., Louis A. Penner, John F. Dovidio, and Jane Allyn Piliavin 1995 *The Psychology of Helping and Altruism: Problems and Puzzles*. New York: McGraw-Hill.
- Schwartz, Shalom H. 1977 "Normative Influences on Altruism." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 10. New York: Academic Press.
- , and Judith A. Howard 1981 "A Normative Decision-Making Model of Altruism." In J. P. Rushton and R. M. Sorrentino, eds., *Altruism and Helping Behavior: Social, Personality, and Developmental Perspectives*. Hillsdale, N.J.: Erlbaum.
- Sober, Elliott, and David S. Wilson 1998 *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, Mass.: Harvard University Press.
- Sorokin, Pitirim A. 1950 *Altruistic Love: A Study of American 'Good Neighbors' and Christian Saints*. Boston: The Beacon Press.
- (ed.) 1970 [1950] *Explorations in Altruistic Love and Behavior: A Symposium*. New York: Kraus Reprint Co. [Boston: The Beacon Press]
- Ting, Jen-Chieh, and Jane A. Piliavin forthcoming "Altruism in Comparative International Perspective." In James Phillips, ed., *Charities: Between State and Market*. Montreal: McGill-Queen's University Press.
- Wilson, John P. 1976 "Motivation, Modeling, and Altruism: A Person X Situation Analysis." *Journal of Personality and Social Psychology* 34:1078-1086.
- Wilson, John, and Marc A. Musick 1997 "Who Cares? Toward and Integrated Theory of Volunteer Work." *American Sociological Review* 62:694-713.

JUDITH A. HOWARD  
JANE ALLYN PILIAVIN

## AMERICAN FAMILIES

Many long-standing assumptions about American families have been challenged by family scholars.

Among these assumptions is the belief that in colonial times the American family was extended in its structure, with three generations living together under one roof. It has been commonly believed that the nuclear family came about as a result of industrialization, with smaller families better able to meet the demands of an industrialized economy. However, historical data show that the extended family was not typical in the colonial era and that the earliest families arriving from Great Britain and other western European countries were already nuclear in structure (Demos 1970; Laslett and Wall 1972).

Some commonly held views about contemporary families also have been debunked in the research literature. For example, it has been commonly thought that American families neglect their elder members and are quick to place them in nursing homes. However, research shows that most older Americans have frequent contact with one or more members of their family, and that families typically provide extensive, long-term care to older persons when such care is needed. In most cases, families turn to placement in a nursing home as a last resort rather than a first option when elder members grow ill or frail.

More generally, family scholars have successfully challenged the notion of "the American family." As Howe (1972, p. 11) puts it, "the first thing to remember about the American family is that it doesn't exist. Families exist. All kinds of families in all kinds of economic and marital situations." This review will show the great diversity of family patterns that characterize the United States of the past, the present, and the foreseeable future.

### HISTORICAL OVERVIEW

It is unfortunate that textbooks intended for courses on the family rarely include a discussion of Native Americans, for a historical examination of these groups shows a striking range of variation in family patterns. In fact, virtually all the variations in marriage customs, residence patterns, and family structures found the world over could be found in North America alone (Driver 1969). Though some of these traditional family patterns have survived to the present day, others have been disrupted over the course of U.S. history. It is important to

note however, that research has not confirmed the commonly held assumption that Native-American societies were placid and unchanging prior to European contact and subsequent subjugation. Important changes were taking place in Native-American societies long before the arrival of Europeans (Lurie 1985).

As has been noted, European immigrants to the American colonies came in nuclear rather than extended families (and also came as single persons—for example, as indentured servants). It was long believed that colonial families were very large, with some early writers claiming an average of ten to twelve children per family, and that most people lived in extended rather than nuclear families. Family scholars, however, have cited evidence showing somewhat lower numbers of children, with an average of eight children born to colonial women (Zinn and Eitzen 1987). Scholars also have distinguished between number of children born per woman and family size at a given point in time. Average family size was somewhat smaller than the average number of children born, due to high infant mortality and because the oldest children often left home prior to the birth of the last child. Evidence suggests an average family size of five to six members during colonial times (Nock 1987). Thus, although the average size of colonial families was somewhat larger than today's families, they were not as large as has been commonly assumed. Furthermore, most people in colonial America lived in nuclear rather than extended family settings.

To understand the size and composition of colonial households, consideration must be given to nonrelated persons living in the home. Servants often lived with prosperous colonial families, and other families took in boarders and lodgers when economic conditions required such an arrangement (Zinn and Eitzen 1987). Households might also include apprentices and other employees. The presence of nonfamily members has important implications for family life. Laslett (1973) has argued that the presence of nonkin meant that households offered less privacy to families and hence provided the opportunity for greater scrutiny by "outsiders." Colonial homes also had fewer rooms than most American homes today, which also contributed to the relatively public nature of these households.

Colonial communities placed great importance on marriage, particularly in New England, where sanctions were imposed on those who did not marry (for example, taxes were imposed on single men in some New England colonies). However, historical records indicate that colonists did not marry at especially young ages. The average age at marriage was twenty-four to twenty-five for men and twenty-two to twenty-three for women (Leslie and Korman 1989). Older ages at marriage during this era reflect parental control of sons' labor on the farm (with many sons delaying marriage until their fathers ceded land to them) and also reflect the lower relative numbers of women (Nock 1987). Parents also typically exerted strong influence over the process of mate selection but did not control the decision. Divorce was rare during this period. The low divorce rate cannot be equated with intact marriages, however. Spousal desertion and early widowhood were far more common experiences than they are today.

The population for the American colonies came primarily from Great Britain, other western European countries, and from western Africa. Initially brought to the colonies in 1619 as indentured servants, hundreds of thousands of Africans were enslaved and transported to America during the colonial period (by 1790, the date of the first U.S. census, African Americans composed almost 20 percent of the population; Zinn and Eitzen 1987). It has been commonly assumed that slavery destroyed the cultural traditions and family life of African Americans. The reasoning behind this assumption was that slave families often were separated for sale to other masters, males were unable to provide for or protect their families, and slave marriages were not legal. The stereotype of "matriarchal" black families, in which women are the family heads and authorities, usually assumes that slavery produced this family form. Empirical research challenges these assumptions. Though slave families lived in constant fear of separation (Genovese 1986), many slave marriages were strong and long-lasting (Gutman 1976). Marriages were legitimized within the slave community (symbolized, for example, by the ritual of "jumping over a broomstick"; Boris and Bardaglio 1987), and two-parent families were common among slaves as well as among free blacks in the North and the South (Gutman 1976). A variety of family structures,

including female-headed households, were found in the slave community and attested to the importance placed on kin ties. Rather than the “absent family” assumed to characterize slave life, slaves were connected to one another through extensive kinship networks (Genovese 1986). Extended kin ties continue to be an important aspect of African-American families today.

For decades, the heritage of slavery and its presumed effects on family life have been invoked to explain social problems in poor black communities (e.g., Moynihan 1965). The historical evidence described above does not lend support to this explanation. Most writers today argue that social problems experienced in poor black communities can more accurately be attributed to the effects of discrimination and the disorganizing effects of mass migration to the urbanized North rather than to the heritage of slavery (e.g., Staples 1986).

Societal changes associated with the Industrial Revolution profoundly affected all types of American families, though the specific nature and extent of these effects varied by social class, race, ethnic origins, and geographic region. Prior to the Industrial Revolution, family members worked together as an economic unit. Their workplace was the home or family farm, with families producing much of what they consumed. Family life and economic life were one and the same, and the boundaries between “private life” in the family and “public life” in the community were blurred. With the development of a commercial economy, the workplace was located away from the family unit. Separation of work and family life created a sharp distinction between the “public” realm of work and the “private” realm of family. Particularly for women’s roles, changes initiated by the Industrial Revolution have been long-lasting and far-reaching. Increasingly, women’s roles were defined by activities assumed to be noneconomic, in the form of nurturing and caring for family members. This was especially true for middle-class women, and married women were excluded from many jobs. Poor and working-class women often participated in wage labor, but this work was generally seen as secondary to their family roles. Men were viewed as having primary responsibility for the economic welfare of their families. No longer an economically interdependent unit, families were

transformed such that women and children became economically dependent on the primary wage earner.

Thus, children’s roles and family relationships also changed with industrialization. In contrast to earlier times, in which children were viewed as miniature adults and engaged in many of the same tasks they would also perform as adults, childhood came to be seen as a special stage of life characterized by dependence in the home. And although children in working-class homes were more likely to work for pay, the evidence suggests that these families also viewed childhood as a stage of life distinct from adulthood (Zinn and Eitzen 1987). Overall, the family became increasingly defined as a private place specializing in the nurturance of children and the satisfaction of emotional needs, a “haven in a heartless world” (Lasch 1977).

Family structures also changed during the 1800s. Family size declined to an average of four to five members. (Of course, average numbers obscure variation in family sizes across social classes and other important dimensions such as race and ethnicity.) Though the average size of nineteenth-century American families was close to that of today’s families, women bore more children during their lifetimes than do American women today. Infant and child mortality was higher and births were spaced further apart, thus decreasing the average size of families at a given point in time. Household size also declined, with fewer households including nonrelated persons such as boarders or apprentices. The average ages at which women and men married were similar to those of colonial times, with an average of twenty-two for women and twenty-six for men. However, greater life expectancy meant that marriages typically lasted longer than they did during the colonial period (Nock 1987).

From 1830 to 1930 the United States experienced two large waves of immigration. The first wave, from 1830 to 1882, witnessed the arrival of more than ten million immigrants from England, Ireland, Germany, and the Scandinavian countries. During the second wave, from 1882 to 1930, over twenty-two million people immigrated to the United States. Peoples from northern and western Europe continued to come to the United States during this second wave, but a large proportion of

immigrants came from southern and eastern Europe as well (Zinn and Eitzen 1987). Immigrants' family lives were shaped by their ethnic origins as well as by the diverse social and economic structures of the cities and communities in which they settled.

Ethnic traditions also helped Mexican-American families adapt to changing circumstances. Annexation of territory from Mexico in 1848 and subsequent immigration from Mexico produced sizable Mexican-American communities in the Southwest. Immigrants from Mexico often reconstructed their original family units within the United States, typically including extended as well as nuclear family members. Extended family households are more common today among Mexican Americans than among non-Hispanic whites, reflecting Mexican Americans' strong family orientation (or "familism") as well as their less advantaged economic circumstances (Zinn and Eitzen 1987).

Imbalanced sex ratios among Chinese and Japanese immigrants greatly influenced the family experiences of these groups. First coming to the United States in the early 1900s, Chinese immigrants were predominantly male. The Chinese Exclusion Act of 1882 barred further immigration, and only wealthy Chinese men were able to bring brides to the United States (Boris and Bardaglio 1987). As of 1910, there were 1,413 Chinese men in the United States to every 100 Chinese women. This sex ratio was still skewed in 1940, when there were 258 men to every 100 women. In contrast to the extended family networks typical of traditional Chinese culture, many Chinese-American households consisted of single men living alone (Marden and Meyer 1973).

Substantial immigration from Japan took place between 1885 and 1924. Like traditional Chinese families, Japanese families were based on strong extended kin networks. As was true for Chinese immigration, most Japanese immigrants were male. In addition to immigration restrictions, Japanese-American families (especially those on the West Coast) were disrupted by property confiscation and the mass relocations that took place during World War II (Marden and Meyer 1973).

In addition to the "old" immigration of the mid-nineteenth century and the "new" immigration of the late nineteenth and early twentieth

centuries, a third wave of large-scale immigration to the United States began in the mid-1960s. In contrast to the earlier waves, when most immigrants came from European countries, most immigration in this third wave has been from Latin America and Asia. However, as has been true of earlier periods of immigration, public controversies surround the economic and social absorption of these new groups (Marger 1991). In addition to occupational and economic challenges facing immigrant families, social challenges include the continuing debate over whether schools should provide bilingual education to non-English-speaking children.

### **MARRIAGE AND FAMILY TRENDS IN THE UNITED STATES**

The separation of paid work and family life, associated with the transition to an industrialized society, gave rise to profound changes in family life. Over the course of the twentieth century, women's roles were defined primarily by family responsibilities within the "private sphere" of the home, but except for a brief period following World War II, women's labor-force participation rose steadily. As of 1997, nearly half (46 percent) of all employed workers were women. Increases in labor-force participation were especially great among married women. In 1900, less than 4 percent of married women were in the labor force. By 1997, that figure had risen to 62 percent (U.S. Bureau of the Census 1998). In contrast to earlier eras of American history, when African-American women were more likely to work for pay than white women, rates of labor force participation are now nearly the same for women in these racial groups, for both married and unmarried women. Married Hispanic women are slightly less likely to be employed than married African-American and non-Hispanic white women (U.S. Bureau of Labor Statistics 1991). As discussed below, women's labor-force participation has important implications for many dimensions of family life.

Though American families have changed in important ways over the past 100 years, examination of historical trends also reveals continuation of some family patterns begun long ago. Notably, the period of the 1950s is commonly thought to

mark the end of a golden age of family life. However, historical data show that for a number of family patterns, the 1950s was an unusual period. Lower rates of divorce, lower ages at marriage, and higher rates of childbearing observed during the 1950s have been attributed mainly to greater economic prosperity following the Great Depression and World War II (Cherlin 1992).

**Age at First Marriage.** According to U.S. census data, the average (median) age at first marriage in the United States was twenty-five years for women and 26.7 years for men in 1998 (U.S. Bureau of the Census 1998). These average ages are higher than for most U.S. population censuses of the past century, but the current median age for men is comparable to that documented for (white) men near the turn of the twentieth century. In 1890, the median age at first marriage was twenty-two for women and 26.1 for men. The average age at marriage declined until 1956, when the median age at first marriage was 20.1 for women and 22.5 for men. The average age at marriage subsequently began to rise (Saluter 1996). Comparison of non-Hispanic whites, blacks, and Hispanic whites shows that age at first marriage climbed more rapidly for blacks than for non-Hispanic whites, with blacks marrying later than non-Hispanic whites. In contrast, Hispanics marry at younger ages than do other groups. It is difficult to assess whether Hispanics' lower age at marriage reflects long-term trends within the United States due to the large numbers of Mexican Americans who immigrated in the 1980s and 1990s (Sweet and Bumpass 1987).

Factors promoting later age at marriage include greater societal acceptance of singlehood and cohabitation as well as greater emphasis on educational attainment. The relationship between age at marriage and level of education is nearly linear for non-Hispanic white men and women, with more education associated with later age at marriage. This relationship is more complex for minority groups, and especially for black and Hispanic men. For these men, later age at marriage is associated both with lower and higher educational levels, producing a U-shaped relationship between education and age at marriage. Minority men with low education are likely to have especially poor job prospects, which in turn affect prospects for marriage. Overall, less racial and ethnic diversity in age

at marriage is shown for those with higher educational attainment (Sweet and Bumpass 1987).

**Interracial Marriage.** Most American marriages are homogamous with regard to race. In 1997, there were 1.3 million interracial marriages, representing only 2.3 percent of all marriages. Many Americans equate interracial marriage with black-white marriage, but only one-quarter of American interracial marriages are between blacks and whites. Currently, 4 percent of black men and 2 percent of black women are married to a white partner. Marriage between African Americans and Asian Americans is even less common than black-white marriages (U.S. Bureau of the Census 1998). Asian Americans are more likely than African Americans to marry a partner of another race. A 1990 study found that 23 percent of Asian Americans intermarry, and that over 40 percent of Japanese-American women are married to a man of a different race (Lee and Yamanaka 1990).

Although the total number of interracial marriages is quite small, the rate of growth in these marriages has been increasing. The number of interracial marriages in the United States increased fourfold between 1970 and 1998. During this period, the total number of marriages grew by only 24 percent (U.S. Bureau of the Census 1998).

**Singles.** Throughout U.S. history a proportion of individuals have remained unmarried. High levels of singlehood were recorded in the late 1800s, when 15 percent of women and 27 percent of men had not married by the age of thirty-four. In 1940, 15 percent of women and 21 percent of men had not married by this age, but by 1970, these proportions had dropped dramatically. By 1970, just 6 percent of women and 9 percent of men had not married by the age of thirty-four (Kain 1990). The size of the unmarried population has been increasing since the 1970s. In 1994, 20 percent of women and 30 percent of men had not married by age thirty-four (Saluter 1996). Women's changing roles have been linked with the rise in singleness: Women with higher education and higher personal income are less likely to marry or have children. Also, in contrast to earlier eras, there is greater societal tolerance of singlehood, providing greater freedom for both women and men to choose a single lifestyle.

Since few individuals marry for the first time at the age of sixty-five or older, a more accurate



picture of the never-married population is provided when attention is restricted to the older population. At present, 3.8 percent of men and 4.7 percent of women age sixty-five and older have never married (Lugaila 1998). Due to the continuing stigmatization of homosexuality, it is difficult to ascertain the numbers of single persons who are gay or lesbian. Researchers have estimated that 4 percent of men and 2 percent of women are exclusively homosexual (Collins 1988). Though homosexual marriages are not legally recognized, many gay and lesbian couples form lasting unions.

**Childbearing.** Childbearing patterns have varied somewhat over the past century. Women born in 1891 had an average of three children. Women born in 1908, who bore children during the Great Depression, had an average of two children. This figure increased to three children per mother during the 1950s and has since declined to two children per mother on average today (Sweet and Bumpass 1987; U.S. Bureau of the Census 1998). In addition to fewer numbers of children born, current trends in childbearing include higher age at first childbirth and longer intervals of time between births. These trends are interrelated. Waiting longer to have a first child and spacing births further apart decrease the average number of children born per mother. The timing of childbearing also has important effects on other life experiences, including educational and occupational attainment. Lower rates of childbearing are associated with higher educational levels and higher incomes.

Fewer married couples are having their first child in the period immediately following marriage, but there are some important differences by race. In 1960, 54 percent of non-Hispanic white couples had children within twelve to seventeen months of marriage. In 1980 this figure dropped to 36 percent. Little change was shown over this period for black couples, who had children within twelve to seventeen months of marriage. Compared to the total population, black couples are likely to have more children on average and to have a child present at the time of marriage. For whites (Hispanics and non-Hispanics) as well as blacks, nine-tenths of all couples have children within seven to eight years of marriage (Sweet and Bumpass 1987).

Childbearing among single women has increased greatly over the past several decades. While just 5 percent of all births were to single women in 1960, approximately one-third (32.2 percent) of all births were to single women in 1995 (Saluter 1996; U.S. Bureau of the Census 1998). As is true for many family patterns, there are substantial variations across racial-ethnic groups. Asian Americans and Pacific Islanders show a relatively low rate of nonmarital births, at 16.3 percent of all births in this group. In contrast, 25.3 percent of all births among whites are to unmarried women, and the rate for Hispanics (who may be of any race) is 40.8 percent. Nearly 70 percent of African-American children are born to single women (U.S. Bureau of the Census 1998). Although socioeconomic factors do not account completely for births to single women, nonmarital childbearing in the United States tends to be higher among those who are poor. Socioeconomic factors can help to explain why African Americans, who are disproportionately likely to be poor, have had higher rates of nonmarital childbearing.

**Divorce.** A rising divorce rate has been a feature of U.S. society since the Civil War. At that time, the divorce rate per 1,000 existing marriages was just 1.2 (Jacobson 1959). By the early 1990s, the rate had increased to more than 20 divorces per 1,000 existing marriages (Cherlin 1996). This long-term trend does not show a smooth and progressive rise, however. Divorce rates have risen more sharply after every major war during this century. Divorce also increased following the Great Depression, apparently reflecting stresses associated with unemployment and economic deprivation. Economic prosperity as well as a greater emphasis on family life have been linked to the lower divorce rate observed from 1950 to 1962. Following 1962, dramatic increases in the divorce rate occurred, with a 100 percent increase between 1963 and 1975 (Cherlin 1992). By the early 1970s, the chance of eventual divorce reached almost 50 percent. The divorce rate has more or less stabilized since that time, such that approximately 50 percent of all first marriages are projected to terminate eventually in divorce. The chances of divorce are higher for second marriages, of which 60 percent are projected to end in divorce (Olson and DeFrain 1994).

Population trends have been linked with the increased rate of divorce. Among these trends is

greater longevity, with an average life expectancy at birth of approximately eighty years for women and seventy-three years for men who were born in 1991. When they reach the age of sixty-five, women who were born in 1991 can expect to live an additional nineteen years, while men of that birth cohort can expect to live fifteen more years (U.S. Bureau of the Census 1996). (These figures are for the total population. Life expectancies are lower for members of racial and ethnic minorities.) In contrast, the life expectancy at birth for those born in 1900 was forty-eight years for women and forty-six years for men (Grambs 1989). Unsatisfactory marriages that formerly may have been terminated by the death of one partner are now more likely to be dissolved by divorce (Uhlenberg 1986). Scholars have also noted the apparent connection between women's increasing levels of labor-force participation and the increased rate of divorce in the United States. Divorce is more likely to occur in couples where the wife is able to support herself financially.

The risk of divorce also varies with age at marriage, duration of the marriage, education, race, and ethnicity. Age at marriage is one of the most important factors, with the likelihood of divorce twice as great among couples where the wife was seventeen or younger than among couples where the wife was in her early twenties. Further, most divorces take place within the first few years of marriage. The longer a couple has been together, the less likely they will be to get divorced. This pattern also holds among couples in which one or both partners have remarried. Education also seems to be an important factor, with a higher divorce rate observed among high school dropouts than among college graduates. However, the effect of education is due in large part to the fact that college graduates tend to marry at later ages. Looking across racial and ethnic groups, the risk of divorce is greater among African Americans than among whites, and especially high divorce rates are observed for Hispanics (Puerto Ricans in particular), Native Americans, and Hawaiians. Divorce is less common among Asian Americans (Sweet and Bumpass 1987).

**Widowhood.** Rising life expectancies have increased the average age of widowhood. Among women, the median age at widowhood was fifty-one in 1900, compared to sixty-eight in 1979. The median age at widowhood for men was forty-five

in 1900 and seventy-one in 1979 (Grambs 1989). Lower average ages of bereavement for men compared to women in 1900 are linked with women's risks for death in childbirth in that era. Women today can expect to live longer in a widowed status compared to widowed men. This gender gap is explained primarily by higher female life expectancy and lower rates of remarriage among women, and also because women tend to marry men several years older than themselves. Nearly one-half (45 percent) of all American women who are age sixty-five or older are widowed, while 15 percent of men in this age group are widowed. Among the oldest-old—those who are eighty-five years of age and older—these figures rise to 77 percent for women and 42 percent for men (Lugaila 1998).

**Remarriage.** Due in large part to the fact that widowhood tends to occur later in life, fewer men and women remarry following the death of a spouse, compared to those who remarry following a divorce. Most people who divorce eventually remarry, but the likelihood of remarriage varies greatly according to gender, age, and race. Approximately five out of six men eventually remarry following a divorce, compared to two out of three women who do so (Cherlin 1992). As noted above, men are also more likely than women to remarry following the death of a spouse. The probability of remarriage declines with age, especially among women. Only one in four women who divorce at age forty or older eventually remarry (Levitan, Sar, and Gallo 1988). Race differences also are observed for remarriage. The proportion of women who remarry following divorce is approximately 50 percent for African Americans and 75 percent for whites (Bumpass, Sweet, and Castro Martin 1990).

Increased rates of divorce and remarriage are transforming American families. "Blended" families or stepfamilies are becoming increasingly common, whereby one or both spouses bring children into a remarriage. One in four children will spend some time in a blended family (Furstenberg and Cherlin 1991). Nearly all of these children live with their biological mothers.

**Household Structure.** As defined by the U.S. Bureau of the Census, "family households" contain persons who are related to the household head (the person in whose name the home is

owned or rented) by blood ties, marriage, or adoption. “Nonfamily households” consist of individuals who live alone or with one or more unrelated persons. Historically, most American households have been family households, and most of these have included married couples. In 1910, 80 percent of all households included married couples. By 1998 this percentage had declined to 53 percent (Casper and Bryson 1998). In contrast, the proportion of single-person households has risen dramatically over the century. In 1890, only 4 percent of all households were of this type (Sweet and Bumpass 1987). As of 1998, single-person households accounted for one-quarter of all U.S. households. Nearly 12 percent of all households consist of men living alone, 15 percent consist of women living alone (Lugaila 1998).

Breakdowns of family structure by race and ethnicity have shown that Americans of Korean, Filipino, Vietnamese, and Mexican heritage are most likely to live in family households (for each group, about 84 percent reported living in family households). African Americans and non-Hispanic whites are somewhat less likely to live in family households. Also, compared to other racial and ethnic groups, Puerto Ricans are most likely to live in a household consisting of a mother and one or more children (with 23 percent living in this type of household), followed by African Americans, Native Americans, and Hawaiians (Sweet and Bumpass 1987).

Type of household is tied closely with economic status. While the “typical” dual-earner couple with children earned an average annual income of \$46,629 in 1991, the average income for mother-only households was only \$13,012 (McLanahan and Casper 1998). Of all household types, those headed by a woman with no husband present have the highest poverty rate. In 1997, nearly 32 percent of these households had incomes that fell below the poverty line (Dalaker and Naifeh 1998).

Due largely to women’s risks for poverty and the rise in female-headed households, children are more likely to be poor today than they were several decades ago. The proportion of American children living in poor families declined during the 1960s—from 26.5 percent in 1960 to 15 percent in 1970—but has since increased. Nearly one-quarter of all American children live in poor families. For children who live in a female-headed household,

the chances of living in poverty rise to 46 percent for white-Hispanic and non-Hispanic children, and to 82 percent for African-American children (Ollenburger and Moore 1998).

### IMPACT OF FAMILY TRENDS ON CHILDREN AND OLDER PERSONS

The majority of the family patterns described here represent long-term trends in the United States and are unlikely to represent the demise of American families, as has been decried by some social observers. Indeed, marriage remains as “popular” as ever. Although Americans are marrying somewhat later on average than they did in the 1950s and 1960s, the majority of women and men continue to marry and have children. Those who divorce tend to remarry. Although these trends did not originate in the late twentieth century, it is nonetheless true that divorce, single-parent households, employed mothers, and nonmaternal childcare are more typical features of American life today than they were in the past. The impact of these family patterns on children’s development and well-being has been a matter of great concern to researchers and policymakers. The impact of women’s employment and changes in family composition also have given rise to concerns regarding the provision of informal care to elderly parents.

**Divorce.** Parents’ divorce has been linked with a range of negative outcomes for children in the areas of psychological adjustment, life satisfaction, academic achievement, and social relationships. These effects are strongest in the first year or two following the divorce, but some long-term consequences also have been found. The experience of parents’ divorce can continue to have negative effects on a child’s well-being as she or he grows into young adulthood (Cherlin, Chase-Lansdale, and McRae 1998).

Although divorce typically is a stressful experience, studies have found a great deal of variation in how children adapt to parents’ divorce. Among the factors identified as important to consider are the family’s socioeconomic status, race-ethnicity, the child’s gender and his or her age at the time of the parents’ divorce or separation. For example, the effects of divorce apparently are more acute for children of school age than for preschool children. Marital disruption also brings greater

risks to children when the parent-child relationship suffers as a result of the divorce and when one or both parents experience multiple divorces (Amato and Booth 1991).

The level and type of conflict present in the family prior to a divorce also is important in understanding the effects of divorce on children. Children whose families were highly conflictual may show increased well-being following a divorce. Conversely, problems of adjustment have been found among children whose parents did not divorce, but whose family lives were characterized by high levels of conflict (Furstenberg and Cherlin 1991).

Individuals who experienced "low-stress" parental divorces do not appear to differ significantly from those who grew up in happy, intact families (Amato and Booth 1991). In general, negative consequences of divorce are not found for children when parents maintain a positive relationship with the child and with one another, and when the child is provided with adequate social and economic resources. As discussed below, economic difficulties pose a central challenge for single-parent (usually single-mother) households.

**Single-parent households.** Just as the impact of divorce on children depends upon multiple factors, no one pattern characterizes how children's well-being is influenced by living in a single-parent household. The circumstances of single parenthood are diverse—single parents can be divorced, widowed, or never-married—and they have access to varying levels and types of social and economic resources. Households may include only the single parent and one or more children, or may include extended family or other household members. Noncustodial parents may or may not be part of the child's life.

It has been noted that single-parent households are more likely than dual-parent households to be poor, and that mother-headed households are especially likely to be poor. Following a divorce, women's subsequent income declines 27 percent on average, while men's average income increases by 10 percent (Peterson 1996). Some of this income gap is due to the fact that children are more likely to live with their mothers than their fathers following a divorce. Currently one child in four lives in a single-parent family, and women head 83 percent of these families (U.S. Bureau of the Census 1998).

Many of the problems associated with single-parent households are associated with problems of poverty. The economic difficulties faced by single parents are compounded when child support payments from the noncustodial parent are not provided regularly. About one-half of custodial mothers were awarded child support payments in 1992. Of the women who had been awarded child support, 76 percent received full or partial payment. A significant amount of awarded child support is not received by the custodial parent: one-third of all awarded child support was not paid in 1991. Among custodial parents, mothers have higher child support award rates and payment rates than fathers, but they are also much more likely to be poor. Single custodial mothers are two and one-half times more likely to be poor than single custodial fathers (U.S. Bureau of the Census 1995).

In addition to economic strains, single parents report problems arising from role and task overload, in coordinating a social life and parenting responsibilities, and difficulties with former spouses. Strains experienced by the parent can in turn impact the well-being of children (Richards and Schmiede 1993). Some evidence also suggests that single parents provide less supervision of children compared to two-parent families (married or cohabiting) (e.g., Astone and McLanahan 1991).

Although research has focused primarily on problems of single-parent households, some potential benefits of these households have been identified for children's development and well-being. For example, children in single-parent households have been found to take greater responsibility for household tasks than children in two-parent households. Along with increased responsibilities within the home, children in one-parent households also apparently develop higher levels of personal autonomy and independence (see Richards and Schmiede 1993).

**Nonmaternal Childcare.** With more American women working for pay than ever before, more preschool-age children are receiving care from their fathers, other family members, or from nonrelatives during their mothers' hours of employment. As of 1991, 9.9 million children age five or younger required care during the hours their mothers were employed. Of these children, the majority were cared for in a home environment:

36 percent were cared for in their own homes (usually by the father or another relative); 31 percent were cared for in another home (usually by a nonrelative in the care provider's home). A further 9 percent were cared for by their mothers while they worked—this was generally in a home environment as well, as most of these mothers had home-based paid work. Approximately one-quarter of the children who required care while their mothers worked were enrolled in an organized day care facility (U.S. Bureau of the Census 1995).

With the increasing use of childcare provided by individuals other than the mother, concerns have been raised regarding the impact of nonmaternal care on children's physical, cognitive, and psychological development. The majority of studies on this topic have concluded that nonmaternal childcare, and nonparental childcare more generally, is not in itself harmful to children. In fact, day care settings apparently aid the development of children's social skills. However, the quality of the care—whether in a family or nonfamily setting—is extremely important for children's development and well-being. A comprehensive report published by the National Research Council found that children whose families were undergoing psychological stress or economic deprivation were more likely to receive care in lower quality settings. Hence, these children are vulnerable not only to poverty and psychological stress within the home, but they are also likely to suffer from the effects of poor-quality care outside the home (Hayes, Palmer, and Zaslow 1990).

**Eldercare.** Most older Americans live independently in the community and are in relatively good health. Most also provide economic and other support to their adult children, such as providing babysitting or other services. It is only when the elder's health or economic status is severely compromised that the balance of exchange tips in the direction of receiving a greater amount of support than elders provide to their children.

The bulk of care provided to older Americans who are frail or ill is provided by family members, usually women, who are typically spouses or children of the elder (Stone, Cafferata, and Sangl 1987; Wolf, Freedman, and Soldo 1997). If a spouse or child is not available, other relatives or a friend may provide care (though assistance from these latter sources is typically less intensive and for

brief periods of time). Much research attests to the extensive and prolonged care provided by family members to elders within a home environment (see Dwyer and Coward 1992). In many if not most cases, it is only after families have exhausted their physical, economic, and emotional resources that a decision is made to place an elder in a long-term care facility. Indeed, only 5 percent of all Americans age sixty-five and older reside in a nursing home or other long-term care facility at a given point in time. The chances of nursing home placement rise with age: Nearly 25 percent of those age eighty-five or older live in a nursing home. However, even among this "oldest-old" age group, the majority of individuals reside in the community rather than a nursing home (Morgan and Kunkel 1998).

Recognizing that women typically are the primary care providers to older family members, researchers and policymakers have raised concerns that such care will be curtailed in the future due to trends in women's employment and an increasing older population. The prevalence of elder care among the employed population is difficult to estimate due to variations in sampling across studies and how caregiving is defined. However, an averaged estimate from studies on this topic is that about one in five employees has some elder care responsibilities (Gorey, Rice, and Brice 1992).

Research has shown that caregiving can be emotionally gratifying for individuals who are providing care to a loved one (Lechner 1992). However, intensive, long-term caregiving can produce serious physical, economic, and emotional strains, especially when caregiving interferes with the provider's work or household responsibilities (Gerstel and Gallagher 1993). A number of studies have found that employment status does not alter the type or amount of care provided to the elderly by their families, but some family members reduce or terminate their employment in order to provide elder care. A national survey found that 9 percent of employees in the study had quit their jobs and a further 20 percent had altered their work schedules to accommodate their caregiving responsibilities. Women were more likely than men to rearrange their work schedules or to reduce or terminate their employment in order to provide family care (Stone, Cafferata, and Sangl 1987).

The strains of intensive caregiving can compromise the provider's ability to continue in the caregiving role. A number of researchers have concluded that greater assistance from the government as well as from the workplace is needed to assist family members and others who provide care to older adults in the community (e.g. Lechner 1992). Such assistance not only benefits the health and well-being of care providers and recipients, but also allows elders to continue living in the community setting for a longer period of time, as is preferred by most elders and their families.

### CROSS-NATIONAL COMPARISONS

The family trends described here are not unique to the United States. Other western societies also have witnessed a rise in the age at first marriage, an increased divorce rate, increased numbers of children born to unmarried women, and an increase in the numbers of women who participate in paid labor. Although the trajectories of these trends are similar for most western societies, differing proportions of individuals and families are represented in these trends across societies. For example, although the average age at marriage has been rising in the United States and most other western societies, the average age at first marriage is lower in the United States than in most of the European countries. And although a rising divorce rate has been a feature of the majority of western societies, the rate of divorce is highest in the United States. Out-of-wedlock births also have increased in western societies over the course of the century. However, in contrast to Scandinavian countries including Sweden and Denmark, more children in the United States live in a single-parent household than with cohabiting parents.

A striking gap between the United States and a number of other countries is found in its relatively high rate of infant mortality. Infant mortality is linked with poverty and the lack of adequate nutrition and health care. The wealth enjoyed by the United States as a nation belies the economic deprivation experienced by subgroups within the society. In 1988, the U.S. infant mortality rate was 10 babies per 1,000 live births. This rate was higher not only as compared with other western societies, but also compared with many nonwestern countries. In a ranking of infant mortality rates worldwide, which were ordered from the lowest to the

highest rates, the United States ranked eighteenth from the bottom. This rate was higher than that found in countries including Singapore, Spain, and Ireland, which are poorer relative to the United States, but which have lower levels of economic inequality within the society (Aulette 1994). A statement from the Children's Defense Fund illustrates the high risks for infant mortality faced by racial-ethnic minority groups within the United States: "A black child born in the inner-city of Boston has less chance of surviving the first year of life than a child born in Panama, North or South Korea or Uruguay" (Children's Defense Fund 1990, p. 6, cited in Aulette 1994, p. 405). (The famine experienced within North Korea during the 1990s likely would remove that country from this listing.)

The United States is unique among western societies in its lack of a national health care program, which helps to explain its higher infant mortality rate. In 1997, 43.4 million Americans, or slightly over 16 percent of the U.S. population, had no health insurance coverage for the entirety of that calendar year. The lack of health insurance is especially acute among the poor. Although the Medicaid program is intended to provide health coverage to the poor, 11.2 million poor Americans, or one-third of all poor people in the United States, had no health coverage in 1997 (Bennefield 1998).

In the absence of concerted social policy measures, infant mortality and other risks faced by poor Americans are unlikely to diminish in the near future. Using the Gini index (a measure of income concentration), the U.S. Bureau of the Census reported that income inequalities within the country increased by 16 percent between 1968 and 1992. Even greater inequalities have developed since 1992: Between 1968 and 1994 the rate of increase in U.S. income inequality was over 22 percent (Weinberg 1996). At the same time, the percentage of Americans with incomes below the poverty line also increased—from 11.7 percent in 1979 to 13.3 percent in 1997 (Dalaker and Naifeh 1997).

### AMERICAN FAMILIES AND THE FUTURE

Traditional distinctions between "family" and nonfamily" are increasingly challenged. Though

still a relatively small proportion of all households (about 4 percent; U.S. Bureau of the Census 1998), the number of cohabiting heterosexual couples has increased greatly in the past several decades. Marriage between homosexuals is not legally recognized, but some have elected to adopt each other legally, and a growing number are raising children. In addition to the “traditional” nuclear family form of two parents with children, other family types can be expected to continue in the future. These include single parents, blended families resulting from remarriage, and households in which other relatives such as grandparents reside.

Increased longevity has brought about some of the most important changes in American family life over the past century. Children are more likely than ever before to interact with their grandparents. Further, many persons are becoming grandparents while their own parents are still alive (Uhlenberg 1986). Research has documented the prevalence and importance of social interaction, emotional support, financial help, and other assistance between the generations. For all types of American families, indications are that high levels of interaction and assistance between the generations will continue in the future. It remains to be seen how U.S. social policy will respond to the needs of family care-providers, and to the needs of poor Americans and their families, in the years to come.

## REFERENCES

- Amato, Paul R., and Alan Booth 1991 “Consequences of Parental Divorce and Marital Unhappiness for Adult Well-Being.” *Social Forces* 69:895–914.
- Astone, Nan Marie, and Sara S. McLanahan 1991 “Family Structure, Parental Practices and High School Completion.” *American Sociological Review* 56:309–320.
- Aulette, Judy Root 1994 *Changing Families*. Belmont, Calif.: Wadsworth.
- Bennefield, Robert L. 1998 *Health Insurance Coverage: 1997*. U.S. Bureau of the Census, Current Population Reports, Series P60–202. Washington, D.C.: U.S. Government Printing Office.
- Boris, Eileen, and Peter Bardaglio 1987 “Gender, Race, and Class: The Impact of the State on the Family and the Economy, 1790–1945.” In Naomi Gerstel and Harried Engel Gross, eds., *Families and Work*. Philadelphia: Temple University Press.
- Bumpass, Larry, James A. Sweet, and Teresa Castro-Martin 1990 “Changing Patterns of Remarriage.” *Journal of Marriage and the Family* 52:747–756.
- Casper, Lynn M., and Ken Bryson 1998 *Household and Family Characteristics: March 1998 (Update)*. U.S. Bureau of the Census, Current Population Reports, Series P20–515. Washington, D.C.: U.S. Government Printing Office.
- Cherlin, Andrew J. 1996 *Public and Private Families*. New York: McGraw-Hill.
- 1992 *Marriage, Divorce, Remarriage*, rev. ed. Cambridge, Mass.: Harvard University Press.
- , Frank F. Furstenberg, Jr., P. Lindsay Chase-Lansdale, Kathleen E. Kiernan, Philip K. Robins, Donna Ruane Morrison, and Julien O. Teitler 1991 “Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States.” *Science* 252:1386–1389.
- Collins, Randall 1988 *Sociology of Marriage and the Family: Gender, Love, and Property*, 2d ed. Chicago: Nelson-Hall.
- Dalaker, Joseph, and Mary Naifeh 1998 *Poverty in the United States: 1997*. U.S. Bureau of the Census, Current Population Reports, Series P60–201. Washington, D.C.: U.S. Government Printing Office.
- Demos, John 1970 *A Little Commonwealth*. New York: Oxford University Press.
- Driver, Harold E. 1969 *Indians of North America*, 2d ed. Chicago: University of Chicago Press.
- Dwyer, Jeffrey W., and Raymond T. Coward 1992 “Gender, Family, and Long-Term Care of the Elderly.” In Jeffrey W. Dwyer and Raymond T. Coward, eds., *Gender, Families, and Elder Care*. Newbury Park, Calif.: Sage.
- Furstenberg, Frank F., Jr., and Andrew J. Cherlin 1991 *Divided Families: What Happens to Children When Parents Part?* Cambridge, Mass.: Harvard University Press.
- Genovese, Eugene D. 1986 “The Myth of the Absent Family.” In Robert Staples, ed., *The Black Family: Essays and Studies*, 3d ed. Belmont, Calif.: Wadsworth.
- Gerstel, Naomi, and Sally K. Gallagher 1993 “Kinkeeping and Distress: Gender, Recipients of Care, and Work-Family Conflict.” *Journal of Marriage and the Family* 55:598–607.
- Gorey, Kevin M., Robert W. Rice, and Gary C. Brice 1992 “The Prevalence of Elder Care Responsibilities Among the Work Force Population.” *Research on Aging* 14(3):399–418.
- Grambs, Jean Dresden 1989 *Women over Forty: Visions and Realities*. New York: Springer.

- Gutman, Herbert 1976 *The Black Family in Slavery and Freedom, 1750–1925*. New York: Pantheon.
- Hayes, Cheryl D., John L. Palmer, and Martha J. Zaslow (eds.) 1990 *Who Cares for America's Children? Child Care Policy for the 1990s*. Washington, D.C.: National Academy Press.
- Howe, Louise Knapp 1972 *The Future of the Family*. New York: Simon and Schuster.
- Jacobson, Paul H. 1959 *American Marriage and Divorce*. New York: Rinehart.
- Kain, Edward 1990 *The Myth of Family Decline: Understanding Families in a World of Rapid Social Change*. Lexington, Mass.: Heath.
- Lasch, Christopher 1977 *Haven in a Heartless World: The Family Besieged*. New York: Basic Books.
- Laslett, Barbara 1973 "The Family as a Public and Private Institution: An Historical Perspective." *Journal of Marriage and the Family* 35:480–494.
- Laslett, Peter, and Richard A. Wall (eds.) 1972 *Household and Family in Past Time*. New York: Cambridge University Press.
- Lechner, Viola M. 1992 "Predicting Future Commitment to Care for Frail Parents Among Employed Caregivers." *Journal of Gerontological Social Work* 18:69–84.
- Lee, Sharon M., and Keiko Yamanaka 1990 "Patterns of Asian American Intermarriage and Marital Assimilation." *Journal of Comparative Family Studies* 21:287–305.
- Leslie, Gerald R., and Sheila K. Korman 1989 *The Family in Social Context*, 7th ed. New York: Oxford University Press.
- Levitan, Sar A., Richard S. Belous, and Frank Gallo 1988 *What's Happening to the American Family? Tensions, Hopes, Realities*. Baltimore, Md.: Johns Hopkins University Press.
- Lugaila, Terry A. 1998 *Marital Status and Living Arrangements: March 1998 (Update)*. U.S. Bureau of the Census, Current Population Reports, Series P20–514. Washington, D.C.: U.S. Government Printing Office.
- Lurie, Nancy Oestreich 1985 "The American Indian: Historical Background." In N.R. Yetman, ed., *Majority and Minority*, 4th ed. Boston: Allyn and Bacon.
- Marden, Charles F., and Gladys Meyer 1973 *Minorities in American Society*, 4th ed. New York: D. Van Nostrand.
- Marger, Martin N. 1991 *Race and Ethnic Relations: American and Global Perspectives*. Belmont, Calif.: Wadsworth.
- McLanahan, Sara, and Lynne Casper 1998 "Growing Diversity and Inequality in the American Family." In Andrew J. Cherlin, ed., *Public and Private Families*. Boston: McGraw-Hill.
- Morgan, Leslie, and Suzanne Kunkel 1998 *Aging: The Social Context*. Thousand Oaks, Calif.: Pine Forge Press.
- Moynihan, Daniel P. 1965 *The Negro Family: The Case for National Action*. Washington, D.C.: Office of Policy Planning and Research, U.S. Department of Labor.
- Nock, Steven L. 1987 *Sociology of the Family*. Englewood Cliffs, N.J.: Prentice Hall.
- Ollenburger, Jane C., and Helen A. Moore 1998 *A Sociology of Women: the Intersection of Patriarchy, Capitalism, and Colonization*, 2d ed. Upper Saddle River, N.J.: Prentice Hall.
- Olson, David H., and John DeFrain 1994 *Marriage and the Family: Diversity and Strengths*. Mountain View, Calif.: Mayfield.
- Peterson, Richard R. 1996 "A Re-Evaluation of the Economic Consequences of Divorce." *American Sociological Review* 61:528–536.
- Richards, Leslie N., and Cynthia J. Schmiede 1993 "Problems and Strengths of Single-Parent Families." *Family Relations* 42:277–285.
- Saluter, Arlene F. 1996 *Marital Status and Living Arrangements: March 1994*. Current Population Reports, Series P20–484. Washington, D.C.: U.S. Government Printing Office.
- Staples, Robert (ed.) 1986 *The Black Family: Essays and Studies*, 3d ed. Belmont, Calif.: Wadsworth.
- Stone, Robyn, Gail Lee Cafferata, and Judith Sangl 1987 "Caregivers of the Frail Elderly: A National Profile." *The Gerontologist* 27:616–626.
- Sweet, James A., and Larry L. Bumpass 1987 *American Families and Households*. New York: Russell Sage.
- Uhlenberg, Andrew 1986 "Death and the Family." In A.S. Skolnick and J.H. Skolnick, eds., *Family in Transition*. Boston: Little, Brown.
- U.S. Bureau of the Census 1999 *Interracial Married Couples: 1960 to Present*. <http://www.census.gov/population/socdemo/ms-la/tabms-3.txt>
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States, 1998*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of the Census 1996 *65+ in the United States*. Current Population Reports, Special Studies, Series P–23, No. 190. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of the Census 1995 *Population Profile of the United States: 1995*. Current Population Reports, Series P23–189. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of Labor Statistics 1991 *Employment and Earnings*. Washington, D.C.: U.S. Government Printing Office.



Weinberg, Daniel H. 1996 *A Brief Look at U.S. Income Inequality*. U.S. Bureau of the Census, Current Population Reports, Series P60-191. Washington, D.C.: U.S. Government Printing Office.

Wolf, Douglas A., Vicki Freedman, and Beth J. Soldo 1997 "The Division of Family Labor: Care for Elderly Parents." *Journals of Gerontology* 52B:102-109.

Zinn, Maxine Baca, and D. Stanley Eitzen 1987 *Diversity in American Families*. New York: Harper and Row.

LAURIE RUSSELL HATCH

## AMERICAN INDIAN STUDIES

American Indian Studies blends many fields in the social sciences and humanities; history and anthropology have been especially prominent, along with education, sociology, psychology, economics, and political science. For convenience, this literature can be grouped into several subject areas: demographic behavior, socioeconomic conditions, political and legal institutions, and culture and religion. Of course, there is a great deal of overlap. To date, this literature deals mostly with aboriginal North Americans and their descendants. As the field has evolved, little attention has been devoted to the natives of South America or the Pacific Islanders. However, there is a growing interest in common experiences of indigenous peoples around the world (e.g. Fleras and Elliot 1992).

### DEMOGRAPHY

**Historical Demography.** Historical demography is important for understanding the complexity of indigenous North American societies and for assessing the results of their contacts with Europeans. For example, complex societies require large populations to generate economic surpluses for trade, and large populations often entail highly developed systems of religion, culture, and governance. Because American Indians almost disappeared in the late nineteenth century, large numbers of pre-Columbian Indians would indicate that devastating mortality rates and profound changes in native social organization followed the arrival of Europeans.

No one knows with certainty when populations of *Homo sapiens* first appeared in the Western

Hemisphere. The first immigrants to North America probably followed game from what is now Siberia across the Beringia land bridge, now submerged in the Bering Sea. This land bridge has surfaced during several ice ages, leading to speculation that the first populations arrived as early as 40,000 years ago or as recently as 15,000 years ago—25,000 years ago is a credible estimate (Thornton 1987, p. 9).

In 1918 a Smithsonian anthropologist, James Mooney, published the first systematic estimates of the American Indian population. He reckoned that 1.15 million American Indians were living around 1600. Alfred Kroeber (1934) subsequently reviewed Mooney's early estimates and deemed them correct, though he adjusted the estimate downward to 900,000 (Deneven 1976). The Mooney-Kroeber estimates of approximately one million American Indians in 1600 have been the benchmark for scholars throughout most of this century. These estimates were flawed, however, because they failed to take epidemic disease into account; European pathogens devastated native populations.

Noting the shortcomings of the Mooney-Kroeber figures, Henry Dobyns (1966) revised the estimate for the 1492 precontact population, suggesting that it was as large as twelve million. His article ignited an intense debate that is still not fully resolved. Conservative estimates now number the indigenous 1492 population at approximately three to five million (Snipp 1989). Dobyns (1983) later raised his estimate to eighteen million.

Population estimates substantially larger than the Mooney-Kroeber figures are consistent with the archaeological record, which indicates that relatively complex societies occupied the Southwest, the Pacific Northwest, and the Mississippi River valley before the Europeans arrived (Thornton 1987). The effects of European contact were certainly greater than once believed. European diseases, slavery, genocidal practices, and the intensification of conflicts nearly exterminated the native people. Huge population losses undoubtedly caused large-scale amalgamation and reorganization of groups struggling to survive and wrought profound changes in their cultures and social structures.

Despite long and heated debates about the likely number of pre-Columbian North Americans, there is relatively little consensus about this figure. It seems likely that it will never be known

with certainty. As scholars realize the elusiveness of this number, there is less interest in trying to establish the definitive estimate and more in attempting to understand the complex demographic behavior related to depopulation (Thornton et al. 1991; Verano and Ubelaker 1992).

**Contemporary Demography.** During the twentieth century, the American Indian population grew very quickly, from about 228,000 in 1890 to about 1.96 million in 1990 (Shoemaker 1999). American Indian fertility is exceedingly high (Snipp 1996). Indians often have better access to health care (from the Indian Health Service) than other equally impoverished groups, and they are experiencing diminishing infant mortality and increasing longevity (Young 1994; Snipp 1996).

A peculiar characteristic of American Indian population growth, at least since 1970, is that a large share of the increase has resulted from persons switching the racial identification they report to the census from another category (such as black or white) to American Indian (Passell and Berman 1986; Harris 1994). The U.S. census, virtually the only comprehensive source of data for American Indians, depends on voluntary racial self-identification. Declining racial discrimination, growing ethnic pride, and resurgence in tribal organization have been cited as reasons that persons of mixed heritage may choose to report themselves as American Indian (Passell and Berman 1986). Evidence indicates that persons who change their identity so they may claim their Native-American heritage tend to be relatively well-educated (Eschbach et al. 1998).

The fluidity of the American Indian population underscores a particularly problematic concern for demographers: namely, defining population boundaries. Definitions abound, and there is no single agreed-upon standard. Some federal agencies and a number of tribes use an arbitrary measure of descent, such as one-fourth blood quantum; standards for tribal membership vary greatly from one-half to one-sixty-fourth Indian blood.

For many other applications, genealogical verification of blood quantum standards is too complex. Agencies such as the U.S. Bureau of the Census thus simply rely on self-identification. By default, most studies of American Indians also rely

on self-identification, especially if they use secondary data from federal government sources. To complicate the matter, the Canadian government uses a somewhat different set of standards to define the boundaries of its native Indian population (Boldt 1993).

Beyond the complexities of counting, studies show that American Indians, more than other minorities, are concentrated in rural areas; slightly less than one-half reside in cities. Most live west of the Mississippi River, primarily because nineteenth-century removal programs were directed at eastern American Indians. A large number of studies document that American Indians are one of the least educated, most often unemployed, poorest, and least healthy groups in American society (see Sandefur et al. 1996). Nonetheless, American Indians are more likely than other groups, especially blacks, to live in a large husband-wife household, and about one-third of them speak an Indian language—provisional evidence of the continuing influence of traditional culture in family organization and language use (Sandefur and Liebler 1996).

#### STUDIES OF SOCIAL AND ECONOMIC STATUS

Studies of the early social and economic status of American Indians focus on the historical development of so-called dependency relations between them and Euro-Americans (White 1983, 1991; Chase-Dunn and Mann 1998). Dependency theory, a variant of neo-Marxist World Systems Theory, has been widely criticized for its shortcomings, but it has gained some acceptance among scholars of white-Indian relations (Wolf 1982; White 1983; Hall 1989). In this view, economic dependency arose from trade relations in which Euro-Americans enjoyed disproportionate economic advantage stemming from a near monopoly over items such as manufactured goods and rum (Wolf 1982; White 1983). European business practices, such as the use of credit, also fostered dependency.

Dependency relations promoted highly exploitative conditions that were a frequent source of conflict and periodically erupted into serious violence. Unscrupulous traders and a growing commerce in Indian captives, for example, spawned the Yamasee War, which ended Indian slavery in

the Southeast (Merrell 1989). Early colonial officials frequently complained about the conflicts created by the unethical practices of frontier traders and sought to curb their abuses, though with little success (White 1991).

Nevertheless, European traders introduced innovations that altered cultures and lifestyles forever. In the Southwest, for example, guns and horses revolutionized relations between nomadic and sedentary groups and allowed the Spanish to exploit traditional antagonisms (Hall 1989).

The emergence of industrial capitalism, large-scale manufacturing, growing urbanization, and an influx of immigrants from Europe and slaves from Africa changed dramatically the relations between Euro-Americans and indigenous peoples. Trading with Indians subsided in favor of policies and measures designed to remove them from lands desired for development (Jacobsen 1984). Throughout the nineteenth century, American Indians were more or less forcibly induced to cede their lands for the development of agriculture, timber, and water. In the late nineteenth century, U.S. corporations began to develop petroleum, coal, and other minerals on tribal lands (Miner 1976).

Exploitation of Indian lands has continued, prompting some scholars to argue that American Indian tribes have a quasi-colonial status within the U.S. economy (Snipp 1986). Natural resources such as timber, water, and minerals are extracted from reservations and exported to distant urban centers where they are processed. In exchange, manufactured goods are imported for consumption. The value of the imported goods typically exceeds the value of the exported resources. The deficit between imports and exports contributes to the persistent poverty and low levels of economic development on many reservations.

The Meriam Report, published in 1928, furnished the first systematic empirical assessment of the economic status of American Indians. Since its publication, numerous studies have documented the disadvantaged status of American Indians (Levitan and Miller 1993). Although many reports have described economic conditions in detail, fewer have attempted to isolate the causes of poverty and unemployment. Clearly, a number of factors can be blamed. American Indians have very little formal education, limiting their access to jobs.

Whether racial discrimination limits opportunities is unclear. Some research suggests that discrimination is not a significant disadvantage for American Indians (Sandefur and Scott 1983), but other studies disagree with this conclusion (Gwartney and Long 1978).

Conditions on reservations, where about one-third of American Indians live, are particularly harsh. Unemployment rates above 50 percent are not unusual. Studies of reservation economies usually blame the isolated locales for many of their woes. The collision of traditional native values and the ethics of capitalism (Cornell and Kalt 1992) frequently complicates economic development in Indian country. In the 1990s, some reservations have enjoyed limited (and in a few instances, spectacular) success in spurring economic development, especially in tourism, gambling, and light manufacturing (Snipp 1995).

Urban American Indians enjoy a higher standard of living than their counterparts in reservation areas (Snipp 1989). Even so, there is disagreement about the benefits of rural-urban migration for American Indians; earlier studies have identified tangible benefits for urban immigrants (Clinton, Chadwick, and Bahr 1975; Sorkin 1978), but later research found contrary evidence (Gundlach and Roberts 1978; Snipp and Sandefur 1988). Federal programs that encouraged urban immigration for American Indians in the 1950s and 1960s were abandoned amid controversies over their effectiveness and overall results (Fixico 1986).

The economic hardships facing rural and urban American Indians alike have been a major source of other serious distress. Alcoholism, suicide, and homicide are leading causes of death for American Indians (Snipp 1996).

#### **POLITICAL ORGANIZATION AND LEGAL INSTITUTIONS**

The political and legal status of American Indians is an extremely complicated subject, tangled in conflicting treaties, formal laws, bureaucratic regulations, and court decisions. Unlike any other racial or ethnic group in U.S. society, American Indians have a distinctive niche in the legal system. As a result of this legal history, a separate agency within the federal government (the Bureau of Indian Affairs [BIA]), a volume of the Code of Federal

Regulations, and a multiplicity of other rules exist for dealing with American Indians.

The political status of American Indian tribes is difficult to characterize. In 1831, Chief Justice John Marshall described tribes as “domestic, dependent nations,” setting forth the principle that tribes are autonomous political entities that enjoy a quasi-sovereignty yet are subject to the authority of the federal government (Pommersheim 1995; Boldt 1993). The limits on tribal political autonomy have fluctuated as a result of court decisions and federal legislation curtailing or extending tribal powers. Since the early 1900s, tribal governments have greatly increased their autonomy (Pommersheim 1995).

One of the most significant political developments in this century for American Indians was the passage of the Wheeler-Howard Indian Reorganization Act (IRA) of 1934. This legislation made it possible for tribes legally to reconstitute themselves for the purpose of limited self-government (Prucha 1984, ch. 37). Subject to the democratic precepts imposed by the federal government, tribes were allowed to have representative governments with judicial, executive, and legislative branches. Other forms of tribal governance—based on the inheritance of authority, for example—were not permitted by the IRA legislation. Today, virtually every reservation has a form of representative government (O’Brien 1989).

Tribal sovereignty is a complex legal doctrine affecting the political autonomy of tribal governments. It is distinct from a closely aligned political principle known as self-determination. The principle of self-determination, unlike tribal sovereignty, is relatively recent in origin and was first posed as a claim for administrative control of reservation affairs. As a political ideology, self-determination developed in response to the unilateral actions of the federal government in implementing policies such as the Termination legislation of the 1950s. In the 1960s, it was a rallying theme for promoting greater tribal involvement in federal policies affecting American Indians. The principle was formally enacted into public law with the passage of the Indian Self-Determination and Educational Assistance Act of 1975, P.L. 97-638. Since its passage, federal agencies have gradually divested control over programs and services such as those once administered by the BIA. For example, many tribal

governments have contracts to provide social services similar to the arrangements made with state and local governments.

In recent years, ideas about self-determination have developed to the point where self-determination is nearly indistinguishable from tribal sovereignty (O’Brien 1989). The most influential statement merging the two is a report presented to the U.S. Senate by the American Indian Policy Review Commission (AIPRC) in 1976. The AIPRC report was a comprehensive, though highly controversial evaluation of federal Indian policy. Every presidential administration since Richard Nixon’s has endorsed the principle of tribal sovereignty. Shortly after taking office, the Clinton Administration endorsed this principle and there is no indication of a reversal in the foreseeable future.

The political revitalization of American Indians accelerated with the civil rights movement. Some observers have suggested that Indian political activism in the 1960s was a response to postwar termination policies (Cornell 1988), which tried to dissolve the federal reservation system and liquidate the special status of the tribes. Relocation programs in the 1950s accelerated the urbanization of American Indians and, at the very least, may have contributed to the political mobilization of urban Indians, as well as their reservation counterparts (Fixico 1986). Though often complementary, the political agendas of urban and reservation Indians are not always in strict accord.

The diverse tribal composition of urban Indian populations has meant that it is virtually impossible to organize them around issues affecting only one or a few tribes. In the face of this constraint, the ideology of “pan-Indianism” is particularly appealing to urban Indian groups (Hertzberg 1971; Nagel 1996). Pan-Indianism is a supratribal ideology committed to broad issues such as economic opportunity and social justice and to cultural events such as intertribal pow-wows.

The roots of modern pan-Indian organizations can be traced first to the Ottawa leader Pontiac and later to the Shawnee leader Tecumseh and Joseph Brant, a Mohawk. These men led pan-Indian movements opposing Euro-American frontier settlement in the late eighteenth and early nineteenth centuries (e.g., Pontiac’s Revolt 1763). In the late nineteenth century, pan-Indian messianic

movements known as Ghost Dances swept across the West (Thornton 1986).

Pan-Indian organizations have been active throughout the twentieth century, but urbanization hastened their development in the 1950s and 1960s (Cornell 1988; Nagel 1996). Some, such as the National Congress of American Indians (founded in 1944), have moderate political agendas focused on lobbying; others, such as the American Indian Movement, are highly militant. The latter was involved in the sacking of the Washington, D.C., BIA office in 1972 and in the armed occupation of Wounded Knee, South Dakota, in 1973 (Smith and Warrior 1996). Today, most cities with large Indian populations have pan-Indian organizations involved in political organization, cultural events, and social service delivery (Johnson et al. 1997).

### CULTURE AND RELIGION

The cultures of American Indians are extremely diverse, and the same can be said, in particular, about their religious beliefs. Not much is known about the spiritual life of American Indians before the fifteenth century. Only from archaeological evidence is such knowledge available, and this seldom captures the rich complexity of religious symbol systems. Most of what is known about American Indian religions is based on the later reports of explorers, missionaries, traders, and anthropologists (Brown 1982).

Contemporary spiritual practices reflect several different types of religious observances: Christian, neotraditional, and traditional. Participation in one type does not necessarily preclude participation in another. Furthermore, there is a great deal of tribal variation.

American Indians who are practicing Christians represent the legacy of European missionaries. The Christian affiliation of many, perhaps most, American Indians reflects their tribal membership and the denomination of the missionaries responsible for their tribe's conversion. Numerical estimates are not available, but there are many Catholic Indians in the Southwest, and American Indians in the Midwest are often Lutheran, to mention only two examples.

American Indians who participate in neotraditional religions often belong to a branch of the

Native American Church (NAC). NAC is a pan-Indian religion practiced throughout the United States and Canada. It combines elements of Christianity with traditional religious beliefs and practices.

Traditional religions are often practiced in informally organized groups such as sweatlodge or feasting societies. Some of these groups are remnants of older religious movements such as the Ghost Dance. Not much is written about them because they are ordinarily not open to outsiders; the Sun Dance is an exception. It is perhaps best known for the ritual scarification and trances of its participants (Jorgenson 1972).

The secrecy in which many traditional religions are practiced may be due to the intense repression once directed at their observances by the federal government. In 1883, the BIA established Courts of Indian Offenses that prosecuted people for practicing native religions. Among other things, the courts forbade traditional medicines, shaman healers, and all traditional ceremonial observances. Despite their dubious legal foundation, the Courts of Indian Offenses were active until their mandate was rewritten in 1935 (Prucha 1984).

In 1935, the federal government ended its official repression of tribal culture and religion. But the conflicts between government authorities and American Indians trying to practice non-Christian religions did not end. Many Indians regard freedom of religion as an elusive promise. Most controversies involve NAC ceremonies, the preservation of sacred areas, and the repatriation of religious artifacts and skeletal remains in museum collections (Loftin 1989; Echohawk 1993). NAC ceremonies are controversial because they sometimes involve the use of peyote (a hallucinogen) as a sacrament. Although peyote was once outlawed, the NAC won the right to use it within narrowly defined limits prescribed by the courts. The U.S. Supreme Court upheld a case in which Oregon banned the use of peyote, however, raising concerns about how the conservative court will interpret freedom of religion cases in the future (Echohawk 1993).

Preservation of sacred areas places Indian groups at odds with land developers, property owners, local governments, and others who would use sites deemed sacred by spiritual leaders. In one

case, the Navajo and Hopi in 1983 went to court to petition against the development of a ski resort that intruded on sacred grounds. In this case and several similar ones, the courts ruled against the Indians (Loftin 1989). Similar conflicts have arisen over the repatriation of religious artifacts and skeletal remains in museums. These issues pit academics such as scientists and museum curators against Indian groups. In some instances, remains and artifacts have been returned to tribes; Stanford University returned burial remains to the Ohlone tribe in California, for example. Other institutions have opposed repatriation or are studying the matter. The Smithsonian has developed a complex policy for repatriation, and the University of California appointed a committee to develop a policy. For the foreseeable future, the controversy is likely to linger in the courts, Congress, and academic institutions.

Compared to repatriation, cultural studies are a less controversial though no less important domain of American Indian Studies. Indian religion represents one of the central forms of native culture, but cultural studies also emphasize other elements of Indian lifestyles, values, and symbol systems. Some of these studies focus on the content of tribal culture; other research deals with the consequences of tribal culture.

For decades, ethnologists recording for posterity details about Indian culture, especially material culture, or documenting the ways that European contact influenced the content of tribal culture dominated studies of American Indians. The popularity of this type of research has declined significantly, partly because there are few "pristine" cultures left anywhere in the world, much less in North America. Another reason, perhaps more damaging, is the growing realization that studies purporting to document precontact Indian culture were based on secondhand accounts of groups that were not truly pristine. The influence of European diseases and trade goods often arrived far in advance of Europeans (Ramenofsky 1987).

Many studies of American Indian culture now resemble literary or artistic criticism. Others focus on how European innovations have been incorporated into tribal culture in unique ways; silversmithing and rug weaving are two well-known examples. A related set of studies deals with the

resurgence of traditional culture, such as the increase in the use of American Indian languages (Leap 1988).

The behavioral consequences of culture are perhaps most prominent in a large literature on American Indian mental health, education, and rehabilitation (Bennett and Ames 1985; Foster 1988). Many studies show that education and rehabilitation efforts can be made more effective if they are sensitive to cultural nuances (LaFromboise 1995). In fact, many specialists take this idea as a point of departure and focus their research instead on the ways in which Euro-American educational and therapeutic practices can be adapted to the cultural predisposition of American Indian clients (Lafromboise 1995).

Like the American Indian population, American Indian Studies is a highly diverse and growing field of inquiry. It is interdisciplinary and eclectic in the perspectives it uses. Once primarily the domain of historians and anthropologists, American Indian Studies has rapidly expanded beyond the bounds of these disciplines with contributions from scholars in a wide variety of fields.

## REFERENCES

- Bennett, Linda A., and Genevieve M. Ames (eds.) 1985 *The American Experience with Alcohol: Contrasting Cultural Perspectives*. New York: Plenum.
- Boldt, Menno 1993 *Surviving as Indians: The Challenge of Self-Government*. Toronto: University of Toronto Press.
- Brown, Joseph Epes 1982 *The Spiritual Legacy of the American Indian*. New York: Crossroad.
- Chase-Dunn, Christopher, and Kelley M. Mann 1998 *The Wintu and Their Neighbors: A Very Small World-System in Northern California*. Tucson: University of Arizona Press.
- Clinton, Lawrence, Bruce A. Chadwick, and Howard M. Bahr 1975 "Urban Relocation Reconsidered: Antecedents of Employment among Indian Males." *Rural Sociology* 40:117-133.
- Cornell, Stephen 1988 *The Return of the Native. American Indian Political Resurgence*. New York: Oxford University Press.
- , and Joseph P. Kalt (eds.) 1992 *What Can Tribes Do? Strategies and Institutions in American Indian Economic Development*. Los Angeles: American Indian Studies Center, University of California, Los Angeles.

- Deneven, William M. (ed.) 1976 *The Native Population of the Americas in 1492*. Madison: University of Wisconsin Press.
- Dobyns, Henry F. 1983 *Their Number Become Thinned: Native American Population Dynamics in Eastern North America*. Knoxville: University of Tennessee Press.
- 1966 "Estimating Aboriginal American Population: An Appraisal of Techniques with a New Hemispheric Estimate." *Current Anthropology* 7:395-416.
- Echohawk, Walter R. 1993 "Native American Religious Liberty: Five Hundred Years After Columbus." *American Indian Culture and Research Journal* 17:33-52.
- Eschbach, Karl, Khalil Supple, and C. Matthew Snipp 1998 "Changes in Racial Identification and the Educational Attainment of American Indians, 1970-1990." *Demography* 35:35-43.
- Fixico, Donald L. 1986 *Termination and Relocation, 1945-1960*. Albuquerque: University of New Mexico Press.
- Fleras, Augie, and Jean Leonard Elliott 1992 *The Nations Within: Aboriginal-State Relations in Canada, the United States, and Australia*. New York: Oxford University Press.
- Foster, Daniel V. 1988 "Consideration of Treatment Issues with American Indians in the Federal Bureau of Prisons." *Psychiatric Annals* 18:698-701.
- Gundlach, James H., and Alden E. Roberts 1978 "Native American Indian Migration and Relocation: Success or Failure." *Pacific Sociological Review* 21:117-128.
- Gwartney, James D., and James E. Long 1978 "The Relative Earnings of Blacks and Other Minorities." *Industrial and Labor Relations Review* 31:336-346.
- Hall, Thomas D. 1989 *Social Change in the Southwest, 1350-1880*. Lawrence: University Press of Kansas.
- Harris, David 1994 "The 1990 Census Count of American Indians: What do the Numbers Really Mean?" *Social Science Quarterly* 75:580-593.
- Hertzberg, Hazel W. 1971 *The Search for an American Indian Identity*. Syracuse, N.Y.: Syracuse University Press.
- Jacobsen, Cardell K. 1984 "Internal Colonialism and Native Americans: Indian Labor in the United States from 1871 to World War II." *Social Science Quarterly* 65:158-171.
- Johnson, Troy, Joane Nagel, and Duane Champagne (eds.) 1997 *American Indian Activism: Alcatraz to the Longest Walk*. Urbana: University of Illinois Press.
- Jorgensen, Joseph G. 1972 *The Sun Dance Religion: Power for the Powerless*. Chicago: University of Chicago Press.
- Kroeber, Alfred L. 1934 "Native American Population." *American Anthropologist* 36:1-25.
- LaFromboise, Teresa D. 1995 *American Indian Life Skills Development*. Madison: University of Wisconsin Press.
- Leap, William L. 1988 "Indian Language Renewal." *Human Organization* 47:283-291.
- Levitan, Sar A., and Elizabeth I. Miller 1993 *The Equivocal Prospects for Indian Reservations*. Occasional Paper 1993-2 (May). Washington, D.C.: Center for Social Policy Studies, The George Washington University.
- Loftin, John D. 1989 "Anglo-American Jurisprudence and the Native American Tribal Quest for Religious Freedom." *American Indian Culture and Research Journal* 13:1-52.
- Merrell, James H. 1989 *The Indian's New World: Catawbas and Their Neighbors from European Contact Through the Era of Removal*. Chapel Hill: University of North Carolina Press.
- Miner, H. Craig 1976 *The Corporation and the Indian: Tribal Sovereignty and Industrial Civilization in Indian Territory, 1865-1907*. Columbia: University of Missouri Press.
- Nagel, Joane 1996 *American Indian Ethnic Renewal: Red Power and the Resurgence of Identity and Culture*. New York: Oxford University Press.
- O'Brien, Sharon 1989 *American Indian Tribal Governments*. Norman: University of Oklahoma Press.
- Passel, Jeffrey S., and Patricia A. Berman 1986 "Quality of 1980 Census Data for American Indians." *Social Biology* 33:163-182.
- Pommersheim, Frank 1995 *Braid of Feathers: American Indian Law and Contemporary Tribal Life*. Berkeley: University of California Press.
- Prucha, Francis Paul 1984 *The Great Father: The United States Government and the American Indians*. Lincoln: University of Nebraska Press.
- Ramenoffsky, Ann F. 1987 *Vectors of Death: The Archaeology of European Contact*. Albuquerque: University of New Mexico Press.
- Sandefur, Gary D., and Carolyn Liebler 1996 "The Demography of American Indian Families." In Gary D. Sandefur, Ronald R. Rindfuss, and Barney Cohen, eds., *Changing Numbers, Changing Needs: American Indian Demography and Public Health*. Washington, D.C.: National Academy Press.
- Sandefur, Gary D., Ronald R. Rindfuss, and Barney Cohen (eds.) 1996 *Changing Numbers, Changing Needs: American Indian Demography and Public Health*. Washington, D.C.: National Academy Press.
- Sandefur, Gary D., and Wilbur Scott 1983 "Minority Group Status and the Wages of Indian and Black Males." *Social Science Research* 12:44-68.

- Shoemaker, Nancy 1999 *American Indian Population Recovery in the Twentieth Century*. Albuquerque: University of New Mexico Press.
- Smith, Paul Chaat, and Robert Allen Warrior 1996 *Like a Hurricane: The Indian Movement from Alcatraz to Wounded Knee*. New York: The New Press.
- Snipp, C. Matthew 1996 "The Size and Distribution of the American Indian Population: Fertility, Mortality, Migration, and Residence." In Gary D. Sandefur, Ronald R. Rindfuss, and Barney Cohen, eds., *Changing Numbers, Changing Needs: American Indian Demography and Public Health*. Washington, D.C.: National Academy Press.
- 1995 "American Indian Economic Development." In Emery N. Castle, ed., *The Changing American Countryside*. Lawrence: University Press of Kansas.
- 1989 *American Indians: The First of This Land*. New York: Russell Sage Foundation.
- 1986 "The Changing Political and Economic Status of American Indians: From Captive Nations to Internal Colonies." *American Journal of Economics and Sociology* 45:145–157.
- , and Gary D. Sandefur 1988 "Earnings of American Indians and Alaska Natives: The Effects of Residence and Migration." *Social Forces* 66:994–1008.
- Sorkin, Alan L. 1978 *The Urban American Indian*. Lexington, Mass.: D. C. Heath.
- Thornton, Russell 1987 *American Indian Holocaust and Survival: A Population History Since 1492*. Norman: University of Oklahoma Press.
- 1986 *We Shall Live Again: The 1870 and 1890 Ghost Dance Movements as Demographic Revitalization*. New York: Cambridge University Press.
- , Jonathon Warren, and Tim Miller 1992 "Depopulation in the Southeast after 1492." In John W. Verano and Douglas H. Ubelaker, eds., *Disease and Demography in the Americas*. Washington, D.C.: Smithsonian Institution.
- Verano, John W., and Douglas H. Ubelaker (eds.) 1992 *Disease and Demography in the Americas*. Washington, D.C.: Smithsonian Institution.
- White, Richard 1991 *The Middle Ground: Indians, Empires, and Republics in the Great Lakes Region, 1650–1815*. New York: Cambridge University Press.
- 1983 *The Roots of Dependency: Subsistence, Environment, and Social Change among the Choctaws, Pawnees, and Navajos*. Lincoln: University of Nebraska Press.
- Wolf, Eric R. 1982 *Europe and the People Without History*. Berkeley: University of California Press.

Young, T. Kue 1994 *The Health of Native Americans: Towards a Biocultural Epidemiology*. New York: Oxford University Press.

C. MATTHEW SNIPP

## AMERICAN SOCIETY

The term *American society* is used here to refer to the society of the United States of America. This conventional usage is brief and convenient and implies no lack of recognition for other societies of North, Central, and South America.

Boundaries of modern national societies are permeable and often socially and culturally fuzzy and changeable. Lines on maps do not take into account the cross-boundary flows and linkages of trade, tourists, information, workers, diseases, military arms and personnel, ethnic or linguistic affiliations, and the like. As a large, heterogeneous country, the United States well illustrates such interdependence and cultural diversity.

During the second half of the twentieth century it became increasingly plain that an understanding of American society required analysis of its place in a global system. National societies have become highly interdependent through extensive flows of capital, technology, goods and services, ideas and beliefs, cultural artifacts, and symbols. A world system of politico-military relationships ( blocs and hierarchies) interacts with a global system of trade, finance, and population transfers, and both systems are influenced by cultural interpenetration (including organizational forms and procedures).

While these developments were under way, the American people became healthier, life expectancy increased, educational levels rose, income and wealth increased, major new technologies developed (e.g., the so-called Information Revolution), and ethnic and racial minorities gained in income, occupational status, and political participation (Farley 1996). On the other hand, economic inequality increased sharply, the prison population grew rapidly (and became disproportionately made up of African Americans), divorce rates remained at high levels, single-parent households increased, and infant mortality remained high, as did violent crime (Farley 1996).



Containing less than 5 percent of the world's population, the United States is a polyglot nation of nations that now accepts a greater and more diverse inflow of legal immigrants than any other country—an average of about one million a year from 1990 to 1995. It is often called a young nation, but elements of its culture are continuous with the ancient cultures of Europe, Asia, and Africa, and its political system is one of the most long-enduring constitutional democracies. Many writers have alleged that its culture is standardized, but it continues to show great diversity of regions, ethnic groupings, religious orientations, rural-urban contrasts, age groupings, political views, and general lifestyles. It is a society in which many people seem convinced that it is undergoing rapid social change, while they hold firmly to many values and social structures inherited from the past. Like all other large-scale societies, in short, it is filled with ambiguities, paradoxes, and contradictions.

This society emerged as a product of the great period of European expansion. From 1790 to 1990, the U.S. land area expanded from fewer than 900,000 square miles to well over three million; its populations from fewer than four million to about 265 million. The United States has never been a static social system in fixed equilibrium with its environment. Peopled primarily by history's greatest voluntary intercontinental migration, it has always been a country on the move. The vast growth of metropolitan areas is the most obvious sign of the transformation of a rural-agricultural society into an urban-industrial society. In 1880, the nation had four million farms; in 1992 it had 1.9 million. From 1949 to 1979 the index of output per hour of labor went from about twenty to about 200. The most massive change in the occupational structure, correspondingly, has been the sharply decreasing proportion of workers in agriculture—now less than 1 percent of the labor force.

The technological transformations that have accompanied these trends are familiar. The total horsepower of all prime movers in 1940 was 2.8 billion; in 1963 it was 13.4 billion; by 1978 it was over 25 billion. Productive capacities and transportation and communication facilities show similar long-term increases. For example, from 1947 to 1995, the annual per capita energy consumption in the United States went from 230 to 345 million

BTUs—an increase of about 50 percent. The American people are dependent to an unprecedented degree on the automobile and the airplane. (As of the late 1980s, the average number of persons per passenger car was 1.8; by 1995, there were 201 million motor vehicles in a population of some 263 million—1.6 persons per vehicle.) Mass transit is only weakly developed. During the single decade of the 1960s there was a 50 percent increase in the number of motor vehicles, and in many cities such vehicles account for 75 percent of the outdoor noise and 80 percent of the air pollution. With about 200 million motor vehicles in 1998, it is even possible to imagine an ultimate traffic jam—total immobilization from coast to coast.

All indicators of what we may call “heat and light” variables have increased greatly: energy consumption, pieces of mail handled by the U.S. Postal Service (from 106 billion in 1980 to 183 billion in 1996), televisions (2.3 sets per household in 1995), telephones (93.9 percent of households had one in 1995), radios, electronic mail (29 million persons using the Internet), cellular phones, and fax. A vast flood of messages, images, and information criss-crosses the continent.

In American households, the average number of hours that the television was on increased from 5.6 in 1963 to 6.8 in 1976, and continues to slowly increase. The effects of television viewing are complex, although there is general agreement among researchers that mass exposure is selective, does focus attention on some matters rather than others (raising public awareness), influence attitudes on specific issues—especially those on which information is scanty—and probably has cumulative effects on a variety of beliefs and preferences (cf. Lang and Lang 1992).

In short, this is a society of high technology and extremely intensive energy use. It is also a country that has developed a tightly organized and elaborately interdependent economy and social system, accompanied by vast increases in total economic productivity. Thus the real gross national product doubled in just two decades (1959–1979), increasing at an average rate of 4.1 percent per year (Brimmer 1980, p. 98). But beginning with the sharp increases in oil prices after 1973, the society entered a period of economic stagnation and low productivity that was marked in the 1980s by large trade deficits, greatly increased

federal budget deficits, and increased problems of international competitiveness. Coming after a long period of sustained growth, the changes of the 1980s resulted in an economy of low savings, high consumption, and low investment—a situation of “living beyond one’s means.” In the 1990s, a sustained period of low inflation and rising stock markets marked a somewhat uneasy sense of prosperity, seen as insecure as the decade ended.

### MAJOR INSTITUTIONS

“Institution” here means a definite set of interrelated norms, beliefs, and values centered on important and recurrent social needs and activities (cf. Williams 1970, chap. 3). Examples are family and kinship, social stratification, economic system, the polity, education, and religion.

**Kinship and Family.** American kinship patterns are essentially adaptations of earlier European forms of monogamous marriage, bilateral descent, neolocal residence, and diffuse extended kinship ties. All these characteristics encourage emphasis on the marriage bond and the nuclear family. In an urbanized society of great geographical and social mobility and of extensive commercialization and occupational instability, kinship units tend to become small and themselves unstable. Since the 1950s, the American family system has continued its long-term changes in the direction of greater instability, smaller family units, lessened kinship ties, greater sexual (gender) equality, lower birth rates, and higher rates of female-headed households. The percentage of married women in the labor force rose from 32 in 1960 to 61 in the 1990s. The so-called “traditional” family of husband, wife, and children under eighteen that comprised 40 percent of families in 1970 had declined to 25 percent in 1995. Over one-fifth of households are persons living alone. Marriage rates have decreased, age at marriage has increased, and rates of divorce and separation continue to be high. The percentage of children under eighteen years of age living in mother-only families increased in the years between 1960 and 1985, from 6 to 16 among white, and from 20 to 51 among black Americans (Jaynes and Williams 1989, p. 522). As individuals, Americans typically retain commitment to family life, but external social and cultural forces are producing severe family stresses.

**Social Stratification.** Stratification refers to structural inequalities in the distribution of such scarce values as income, wealth, power, authority, and prestige. To the extent that such inequalities result in the clustering of similarly situated individuals and families, “strata” emerge, marked by social boundaries and shared styles of life. When succeeding generations inherit positions similar to previous generations, social classes can result.

The American system is basically one of open classes, with relatively high mobility, both within individual lifetimes and across generations. A conspicuous exception has been a caste-like system of racial distinctions, although this has eroded substantially since the civil rights movement of the 1960s (Jaynes and Williams 1989). The 1980s and 1990s were marked by growing income inequality, as the rich became richer and the poor did not. Large increases in earnings inequality accrued during the 1980s; meanwhile labor unions had large membership losses and labor markets were deregulated (unions lost over 3 million members between 1980 and 1989—see Western 1998, pp. 230–232). Union membership declined from 20.1 percent of workers in 1983 to 14.5 percent in 1996 (*Statistical Abstract of the United States, 1997*, p. 441).

In the early 1990s, the United States had the highest income inequality of any of twenty-one industrialized countries (the income of individuals of the highest decile was over six times the income of the lowest decile). The United States differs from other industrialized countries in the especially great disadvantage of its poorest people. This result arises from very low wages at the bottom of the distribution and low levels of income support from public programs (Smeeding and Gottschalk 1998, pp. 15–19).

The end-of-the-century levels of inequality are less than in the early decades of the 1900s, but represent large increases since the lower levels at the end of the 1960s (Plotnick, Smolensky, Evenhouse, and Reilly 1998, p. 8). By 1996, the richest 20 percent of Americans received about 47 percent of total income, while the poorest got 4.2 percent.

The dominant ideology remains individualistic, with an emphasis on equality of opportunity and on individual achievement and success. Although extremes of income, power, and privilege

produce strong social tensions, the system has shown remarkable stability.

The history of stratification has included conquest of Native Americans, slavery of African peoples, and extensive discrimination against Asians, Hispanics, and various immigrants of European origin. Assimilation and other processes of societal inclusion have moved the whole society increasingly toward a pluralistic system, but deep cleavages and inequalities continue. A fundamental tension persists between principles of equality of opportunity and individual merit, on the one hand, and practices of ascribed status and group discrimination, on the other (cf. Myrdal, Sterner, and Rose 1944).

As a consequence of increased openness since the Immigration Act of 1965, the population contains increased proportions of persons whose backgrounds are in Asia and Latin America; this development complicates ethnic/racial boundaries and alignments, including political formations (Edmonston and Pasell 1994). The increased receptivity to immigrants was enhanced by the Refugee Act of 1980, the Immigration and Control Act of 1986, and the Immigration Act of 1990. The result is that the number of immigrants admitted per year has soared to an average of about a million during the period 1990–1995 (*Statistical Abstract of the United States, 1997*, p. 10). Although vigorous political controversy surrounds immigration, the country remains committed to a general policy of acceptance and to citizenship by residence rather than by ethnic origin.

Although much reduced in its most obvious forms, discrimination against African Americans continues to be widespread, in housing (Yinger 1995), credit, and employment (Wilson 1996; Jaynes and Williams 1989). Residential segregation continues at high levels, although the 1980s brought a small movement toward more residential mixing, primarily in smaller Southern and Western cities (Farley and Frey 1994).

**The Economy.** The American economic system is a complex form of “high capitalism” characterized by large corporations, worldwide interdependence, high levels of private consumption, and close linkages with the state.

Increased specialization leads both to increased complexity and to increased interdependence, two

sides of the same coin. In the United States, as in all industrialized countries, the movement from the primary extractive and agricultural industries to manufacturing was followed by growth of the tertiary exchange-facilitating activities, and then to expansion of occupations having to do with control and coordination and those ministering directly to the health, education, recreation, and comfort of the population. As early as 1970, nearly two-thirds of the labor force was in pursuits other than those in “direct production” (primary and secondary industries).

Since the late 1970s, there has been rapid growth in involuntary part-time jobs and in other insecure and low-paying employment as the economy has shifted toward trade and services and corporations have sought to lower labor costs (Tilly 1996).

As the economy has thus shifted its focus, the dominance of large corporations has become more and more salient. In manufacturing, the total value added that is accounted for by the 200 largest companies went from 37 percent in 1954 to 43 percent in 1970. Of all employees in manufacturing, the percent working in multi-unit firms increased from 56 percent in 1947 to 75 percent by 1972 (Meyer 1979, pp. 27–28). The top 500 industrial companies account for three-fourths of industrial employment (Wardwell 1978, p. 97).

Meanwhile, organized labor has not grown correspondingly; for decades, overall unionization has remained static, increasing only in the service, technical, and quasi-professional occupations. The importance of the great corporations as the primary focus of production and finance continues to increase. Widespread dispersion of income rights in the form of stocks and bonds has made the giant corporation possible, and this same dispersion contributes directly to the concentration of control rights in the hands of salaried management and minority blocs of stockholders. With widened markets for mass production of standardized products, strong incentives were created for effective systems of central control. Although such tendencies often overreached themselves and led to a measure of later decentralization, the modern corporation, not surprisingly, shows many of the characteristics of the most highly developed forms of bureaucracy.

The interpenetration of what were previously regarded as separate political and economic affairs is a central fact. The interplay takes many different forms. For a long time government has set rules for maintaining or lessening business competition; it has regulated the plane or mode of competition, the conditions of employment, and the place and functioning of labor unions. Pressure groups, based on economic interests, ceaselessly attempt to influence law-making bodies and executive agencies. Governmental fiscal and monetary policies constitute a major factor influencing economic activity. As the economic role of the state has expanded, economic forces increasingly affect government itself and so-called private corporations increasingly have come to be “public bodies” in many ways, rivaling some sovereign states in size and influence. The post-1980s political movements for a smaller role for the central government resulted in a partial dismantling of the “welfare state” but did not remove the important linkages of state and economy.

**Political Institutions.** In ideology and law the American polity is a parliamentary republic, federal in form, marked by a strong central executive but with a tripartite separation of powers. From the highly limited state of the eighteenth century, the actual government has grown in size and scope and has become more complex, centralized, and bureaucratized. Partly because of pervasive involvement in international affairs, since World War II a large permanent military establishment has grown greatly in size and importance. In 1996, the Department of Defense included 3.2 million persons, and total defense and veterans outlays amounted to \$303 billion. The executive agencies, especially the presidency, became for decades increasingly important relative to the Congress, although the 1990s brought a resurgence of congressional power. Among other changes, the following appear to be especially consequential:

1. Continuing struggles over the character of the “welfare state,” dedicated to maintaining certain minimal safeguards for health and economic welfare;
2. High development of organized interest groups, which propose and “veto” nearly all important legislation. The unorganized general public retains only an episodic and delayed power to ratify or reject whole programs of government action. A rapid increase in the number of Political Action Committees—from 2551 in 1980 to 4016 in 1995—is only one indication of the importance of organized interests;
3. Decreased cohesion and effectiveness of political parties in aggregating interests, compromising parties in conflict, and reaching clear public decisions;
4. Increasingly volatile voting and diminished party regularity and party commitment (split-ticket voting, low rates of voting, large proportion of the electorate with no firm party reference).

American political parties are coalitions of diverse actors and interests, with accompanying weak internal discipline, but they remain relatively stable under a system of single-member districts and plurality voting—“first past the post.” Although the polity is subject to the hazards of instability associated with a presidential rather than parliamentary system, the national federal system, the separation of powers, and the centrality of the Constitution and the judiciary combine to support the traditional two-party electoral arrangement, although support for a third party appears to be growing (Lipset 1995, p. 6).

Historically, political parties in the United States have been accommodationist: They have served to articulate and aggregate interests through processes of negotiation and compromise. The resulting “packages” of bargains have converted diverse and diffuse claims into particular electoral decisions. To work well, such parties must be able to plan nominations, arrange for representativeness, and sustain effective competition. In the late twentieth century, competitiveness was weakened by volatile elections—for example, landslides and deadlocks with rapidly shifting votes—and by party incoherence. In the nominating process the mass media and direct primaries partly replaced party leaders and patronage. Representativeness was reduced by polarization of activists, single-issue voting, and low turnouts in primaries. And the inability of parties to protect legislators seemed to increase the influence of single-issue organizations and to enlarge the scope of “symbolic” actions. Hard choices, therefore, tended to be deferred (cf. Fiorina 1980, p. 39).

The existence of an “interest-group” polity was clearly indicated. The political system readily expressed particular interests but found difficulties in articulating and integrating partly incompatible demands into long-term national programs.

As the century drew to a close, many commentators expressed concerns about the increasing expense of political campaigns, the increasing importance of very large contributions through Political Action Committees, the potential influence upon voters of “vivid soundbites” on television, and the increasing centralization of control of the mass media. At the same time, the conspicuous behavior of so-called independent counsels (special prosecutors) raised the fears that a “Fourth Branch” of government had arisen that would be relatively free of the checks and balances, traditional in the tripartite system of governance. Public opinion polls showed increased disaffection with political institutions and processes, and lower voting turnouts indicated much apathy in the electorate. As investigations, prosecutions, and litigation have escalated and have been rendered omnipresent by the media, erosion of trust in government has likewise increased greatly (Lipset 1995).

Yet detailed analysis of data from national public opinion surveys in the last decade of the century failed to find the alleged extreme polarization that had been suggested by acrimonious partisanship between the political parties and in the Congress. Thus, a major study (DiMaggio, Evans, and Bryson 1996) found little evidence of extreme cleavages in social opinions between 1974 and 1994, with two exceptions: attitudes toward abortion diverged sharply, and the attitudes of those who identify with the Democratic and Republican parties have become more polarized. Instead of moderating dissension, the party system between 1970 and the 1990s appears to have sharpened cleavages. There is a possibility that some political leaders have been pulled away from centrist positions by militant factions within their own party. The total picture seemed to be that of extreme contentiousness within the central government while the wider society showed much greater tolerance, consensus, and stability.

**Education.** In addition to diffuse processes of socialization found in family and community, specialized educational institutions now directly involve one-fifth of the American people as teachers,

students, and other participants. In the twentieth century, an unparalleled expansion of mass education occurred. Nearly 80 percent of the appropriate age group graduate from secondary school and 62 percent of these attend college; in 1993, 21.3 percent had completed four years of college or more.

Historically, the educational system was radically decentralized, with thousands of school districts and separate educational authorities for each state (Williams 1970, chap. 8). In contrast to countries with strong central control of education and elitist systems of secondary and higher education, the United States for most of its history has had a weak central state and a mass education system. Education was driven by demands for it rather than by state control of standards, facilities, tests, curricula, and so on (cf. Garnier, Hage, and Fuller 1989).

These characteristics partly derive from widespread faith in education as a means of social advancement as well as from commitments to equality of opportunity and to civic unity. Inequalities of access were long enforced by involuntary racial segregation, now somewhat reduced since 1954, when the Supreme Court declared such segregation unconstitutional. Inequalities of access due to social class and related factors, of course, continue (Jencks et al. 1979). Formal educational attainments have come to be so strongly emphasized as a requirement for employment and advancement that some observers speak of the development of a “credential society” (Collins 1979). Meanwhile the slow but steady decline in students’ test scores has aroused much concern but little agreement as to remedial measures.

**Religious Institutions.** Major characteristics of institutionalized religion include: formal separation of church and state, freedom of religious expression and practice, diversity of faiths and organizations, voluntary support, evangelism, high rates of membership and participation, widespread approval of religion and acceptance of religious beliefs, complex patterns of partial secularization, frequent emergence of new religious groupings, and important linkages between religious affiliations and social class and ethnicity (cf. Williams 1970, chap. 9; Wilson 1978).

Many of these characteristics are causally interrelated. For example, earlier sectarian diversity

encouraged separation of church and state and religious toleration, which, in turn, favored further diversity, voluntarism, evangelism, and religious innovations. Self-reported religious affiliations in social surveys shows these percentages: Protestant, 60; Catholic, 25; Jewish, 2; other, 4; none, 9. These broad categories cover hundreds of diverse groupings (General Social Surveys 1994).

Changes include growth in membership of evangelical Protestant denominations (now one-fifth of the population, Hunter 1997), closer ties between religious groupings and political activities, and the rise of many cults and sects. Separation of church and state was increasingly challenged in the 1990s, and religious militancy in politics increased. Nevertheless, national surveys (1991) showed that the religiously orthodox and theological progressives were not polarized into opposing ideological camps across a broad range of issues—although there were sharp divisions on some particular issues (Davis and Robinson 1996).

Among industrialized Western countries, the United States manifests extraordinary high levels of membership and participation. Thus, although there has been extensive secularization, both of public life and of the practices of religious groups themselves, religious influence remains pervasive and important (Stark and Bainbridge 1985).

### **SOCIAL ORGANIZATION**

The long-term increase in the importance of large-scale complex formal organizations, salient in the economy and polity, is evident also in religion, education, and voluntary special-interest associations. Other trends include the reduced autonomy and cohesion of small locality groupings and the increased importance of special-interest formal organizations and of mass publics and mass communication. The long-term effects of the saturation of the entire society with advertising, propaganda, assorted information, and diverse and highly selective world views remain to be ascertained. Local communities and kinship groupings have been penetrated more and more by formal, centralized agencies of control and communication. (Decreasing localism shows itself in many forms. A well-known and striking example is the continuous decrease in the number of public school districts.)

These changes have moved the society as a whole in the direction of greater interdependence, centralization, formality, and impersonality.

### **VALUES AND BELIEFS**

Beliefs are conceptions of realities, of how things are. Values are conceptions of desirability, of how things should be (Williams 1970, chap. 11). Through shared experience and social interaction, communities, classes, ethnic groupings, or whole societies can come to be characterized by similarities of values and beliefs.

The weight of the evidence for the United States is that the most enduring and widespread value orientations include an emphasis on personal achievement (especially in occupational activity), success, activity and work, stress on moral principles, humanitarianism, efficiency and practicality, science, technology and rationality, progress, material comfort, equality, freedom, democracy, worth of individual personality, conformity, nationalism and patriotism; and, in tension with most other values, values of group superiority and racism. Mixed evidence since the 1970s seems to indicate complex shifts in emphasis among these orientations—primarily in the direction of success and comfort, with lessened commitment to more austere values. Some erosion in the emphasis placed on work and some lessening in civic trust and commitment may have occurred.

In contrast to many images projected by the mass media, national surveys show that most Americans still endorse long-standing beliefs and values: self-reliance, independence, freedom, personal responsibility, pride in the country and its political system, voluntary civic action, anti-authoritarianism, and equality within limits (Inkeles 1979; Williams 1970, chap. 11). And for all their real disaffections and apprehensions, most Americans see no other society they prefer: As late as 1971, surveys in eight countries found that Americans were less likely than persons in any other country to wish to live elsewhere (Campbell, Converse, and Rodgers 1976, pp. 281–285). Americans in national public opinion surveys (1998) ranked second among twenty-three countries in pride in the country and in its specific achievements. Thus, popular attitudes continue to reflect a perennial satisfaction and positive nationalism (Smith and Jarkko 1998).

## REFERENCES

- Brimmer, Andrew F. 1980 "The Labor Market and the Distribution of Income." In Norman Cousins, ed., *Reflections of America*. Washington D.C.: U.S. Government Printing Office.
- Campbell, Angus, Phillip E. Converse, and Willard L. Rodgers 1976 *The Quality of American Life*. New York: Russell Sage Foundation.
- Cherlin, Andrew J. 1992 *Marriage, Divorce, Remarriage*. Cambridge, Mass.: Harvard University Press.
- Collins, Randall 1979 *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Academic Press.
- Davis, Nancy S., and Robert V. Robinson 1996 "Are the Rumors Exaggerated? Religious Orthodoxy and Moral Progressivism in America." *American Journal of Sociology* 102(3):756-787.
- DiMaggio, Paul, John Evans, and Bethany Bryson 1996 "Have Americans' Social Attitudes Become more Polarized?" *American Journal of Sociology* 102(3):690-755.
- Edmonston, Barry, and Jeffrey S. Pasell 1994 *Immigrational Ethnicity: The Integration of America's Newest Arrivals*. Washington D.C.: The Urban Institute Press.
- Farley, Reynolds 1996 *The New American Reality: How We Got There, Where We Are Going*. New York: Russell Sage Foundation.
- , and William H. Frey 1994 "Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps Toward a More Integrated Society." *American Sociological Review* 59(1):23-45.
- Fiorina, Morris P. 1980 "The Decline of Collective Responsibility in American Politics." *Daedalus*, Summer: 25-45.
- Frank, Robert H., and Phillip J. Cook 1995 *The Winner-Take-All Society*. New York: Free Press.
- Garnier, Maurice, Jerald Hage, and Bruce Fuller 1989 "The Strong State, Social Class, and Controlled Expansion in France, 1881-1975." *American Journal of Sociology* 95:279-306.
- Hunter, James Davidson 1987 *Evangelism: The Coming Generation*. Chicago: University of Chicago Press.
- Inkeles, Alex 1979 "Continuity and Change in the American National Character." In Seymour Martin Lipset, ed., *The Third Century: America as a Post-Industrial Society*. Stanford, Calif.: Hoover Institution Press.
- Jaynes, Gerald David, and Robin M. Williams, Jr. (eds.) 1989 *A Common Destiny: Blacks and American Society*. Washington, D.C.: National Academy Press.
- Jencks, Christopher, Susan Bartlett, Mary Corcoran, James Crouse, David Eaglesfield, Gregory Jackson, Kent McClelland, Peter Mueser, Michael Olneck, Joseph Schwartz, Sherry Ward, and Jill Williams 1979 *Who Gets Ahead: The Determinants of Economic Success in America*. New York: Basic Books.
- Lang, Gladys Engel, and Kurt Lang 1992 "Mass Media Research." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*, vol. 3. New York: Macmillan Publishing Company.
- Lipset, Seymour Martin 1995 "Malaise and Resiliency in America." *Democracy* 6(3):4-18.
- Meyer, Marshall W 1979 "Debureaucratization?" *Social Science Quarterly* 60:25-34.
- Myrdal, Gunnar, Richard Sterner, and Arnold Rose 1944 *An American Dilemma*. New York: Harper and Row.
- Plotnick, Robert D., Eugene Smolensky, Eirik Evenhouse, and Siobhan Reilly 1998 "Inequality and Poverty in the United States: The Twentieth-Century Record." *Focus* 19(3):7-14.
- Smeeding, Timothy M., and Peter Gottshalk 1998 "Cross-National Income Inequality: How Great Is It and What Can We Learn From It?" *Focus* 19(3):15-19.
- Smith, Tom W., and Lars Jarkko 1998 "National Pride: A Cross-National Analysis." *GSS Cross-National Report No. 19*. Chicago: National Opinion Research Center.
- Stark, Rodney, and William Sims Bainbridge 1985 *The Future of Religion: Secularization, Revival, and Cult Formation*. Berkeley: University of California Press.
- Tilly, Cris 1996 *Half a Job: Bad and Good Part-time Jobs in a Changing Labor Market*. Philadelphia: Temple University Press.
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States, 1997* (117th ed.). Washington, D.C.:
- Wardwell, Nancy Needham 1978 "The Corporation." *Daedalus*, Winter: 97-110.
- Western, Bruce 1998 "Institutions and the Labor Market." In Mary C. Brinton and Victor Nee, eds., *The New Institutionalism in Sociology*. New York: Russell Sage Foundation.
- Williams, Robin M., Jr. 1979 "Change and Stability in Value Systems: A Sociological Perspective." In Milton Rokeach, ed., *Understanding Human Values*. New York: Free Press.
- 1970 *American Society: A Sociological Interpretation*. New York: Alfred A. Knopf.
- Wilson, John N. 1978 *Religion in American Society: The Effective Presence*. Englewood Cliffs, N.J.: Prentice-Hall.
- Wilson, William Julius 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf.

Yinger, John 1995 *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discriminations*. New York: Russell Sage Foundation.

ROBIN M. WILLIAMS, JR.

## AMERICAN SOCIOLOGICAL ASSOCIATION AND OTHER SOCIOLOGICAL ASSOCIATIONS

The American Sociological Association (ASA) will celebrate its centennial year in 2005; since its inception, it has grown in size, diversity, programs, and purpose. Current ASA goals are as follows:

- Serving sociologists in their work,
- Advancing sociology as a science and as a profession,
- Promoting the contributions and use of sociology to society.

While the first goal remains the *raison d'être* for the membership organization, over the ASA's 100 years, there have been ebbs and flows, support and controversy, about the latter two goals and how the association embodies them.

### ASA MEMBERSHIP TRENDS

An interesting perspective on the ASA's history is revealed through an examination of membership trends. Table 1 shows fairly slow but stable growth up until 1931. During the years of the Great Depression, there were substantial declines. Despite these declines, however, sociologists were becoming very visible in government agencies such as the U.S. Department of Agriculture and the U.S. Bureau of the Census. Between 1935 and 1953, for example, there were an estimated 140 professional social scientists, the great majority of them sociologists, employed in the Division of Farm Population and Rural Life. This activity reached its peak between 1939 and 1942, when there were approximately sixty professionals working in Washington, D.C. and in regional offices. Sociologists are well placed in many federal agencies and non-profit organizations in Washington; however, they are "undercover," working under a variety of job titles.

The years following World War II saw a rapid increase in ASA membership—the number nearly quadrupled between 1944 (1,242) and 1956 (4,682). Between 1957 and 1967, membership more than doubled, from 5,223 to 11,445, and continued upward to 15,000 during the heights of the social protest and anti-Vietnam War movements. However, during the latter half of the 1970s, membership gradually drifted downward and reached a seventeen-year low of 11,223 in 1984. In the next fifteen years, the membership increased by 2,000 and has remained stable at over 13,000 in the 1990s.

The growth and decline in the ASA can be accounted for in part by a combination of ideological and demographic factors as well as the gradually changing nature of work in American society, particularly since the end of World War II. For example, the GI bill made it possible for an ordinary veteran to get a college education. The college population jumped from one-half million in 1945 to several million within three years. Gradually, while urban and metropolitan populations grew, the number and percentage of people in the manufacturing sector of the labor force declined, and the farm population declined even more dramatically, while the service sector grew. Within the service sector, information storage, retrieval, and exchange grew in importance with the coming of the computer age. These societal changes helped to stimulate a growth in urban problems involving areas such as family, work, and drugs, and these changes led to a growth of these specialty areas in sociology.

Membership in the ASA rapidly increased in the 1960s and early 1970s, an era of many social protest movements. Sociology was seen as offering a way of understanding the dynamic events that were taking place in this country. Substantive areas within the ASA and sociology were also affected by these social changes. As Randall Collins (1989) points out, the social protest movements of the 1960s and 1970s coincided with the growth within the ASA of such sections as the Marxist, environmental, population, world systems, collective behavior and social movements, and racial and ethnic minorities sections. In addition, the growing public concerns in the late 1970s and early 1980s about aging and equality for women were reflected within the ASA by new sections on sex and gender and aging. Similarly, the "me" generation, in the aftermath of the protest movements and the



**ASA Official Membership Counts  
1906-1999**

1906	115	1931	1,567	1954	4,350	1977	13,755
1909	187	1932	1,340	1955	4,450	1978	13,561
1910	256	1933	1,149	1956	4,682	1979	13,208
1911	357	1934	1,202	1957	5,233	1980	12,868
1912	403	1935	1,141	1958	5,675	1981	12,599
1913	621	1936	1,002	1959	6,323	1982	12,439
1914	597	1937	1,006	1960	6,875	1983	11,600
1915	751	1938	1,025	1961	7,306	1984	11,223
1916	808	1939	999	1962	7,368	1985	11,485
1917	817	1940	1,034	1963	7,542	1986	11,965
1918	810	1941	1,030	1964	7,789	1987	12,370
1919	870	1942	1,055	1965	8,892	1988	12,382
1920	1,021	1943	1,082	1966	10,069	1989	12,666
1921	923	1944	1,242	1967	11,445	1990	12,841
1922	1,031	1945	1,242	1968	12,567	1991	13,021
1923	1,141	1946	1,651	1969	13,485	1992	13,072
1924	1,193	1947	2,057	1970	14,156	1993	13,057
1925	1,086	1948	2,450	1971	14,827	1994	13,048
1926	1,107	1949	2,673	1972	14,934	1995	13,254
1927	1,140	1950	3,582	1973	14,398	1996	13,134
1928	1,352	1951	3,875	1974	14,654	1997	13,082
1929	1,530	1952	3,960	1975	13,798	1998	13,273
1930	1,558	1953	4,027	1976	13,958	1999	13,056

Table 1

disillusionment that set in after the Vietnam War, may have contributed both to a decline in student enrollments in sociology courses and in ASA membership. The growth of the college student population and some disillusionment with purely vocational majors, as well as sociology's intrinsic interest to students, led to a gradual rise in membership in the 1990s.

ASA membership trends can also be examined in the context of the availability of research money. Postwar federal support for sociology grew with the development of sponsored research and the growth of research labs and centers on university campuses. Coincident with an increase in ASA membership, research funding from federal agencies during the 1960s and 1970s grew steadily. In the early 1980s, however (particularly in 1981 and 1982), cutbacks in research funding for the social sciences were especially noticeable. In the 1990s, major efforts to educate Congress on the importance of social science research won support for sociology and the other social sciences. The result has been a reversal of the negative trend and a slow but steady improvement in funding, not only for basic research but also in greater amounts for research with an applied or policy orientation.

**ASA SECTIONS, PUBLICATIONS, AND PROGRAMS**

The increase in the number of sections in the ASA and of membership in them is another sign of growth within the association in the 1990s. Despite a decline in overall ASA membership in the early 1980s, the number of sections increased from nineteen in 1980 to twenty-six in 1989 and to thirty-nine a decade later. The new sections represent some new fields of study (or at least a formal nomenclature for these specialties) such as sociology of emotions, sociology of culture, rational choice, and sociology of sexualities. Furthermore, this overall increase in sections was not achieved by simply redistributing members already in sections but resulted from an actual growth in section members from 8,000 to 11,000 to 19,000 in the three time periods. More than half the ASA members belong to at least one section; the modal membership is in two sections. The ASA has learned from other associations, such as psychology and anthropology, about the possible pitfalls of subgroups within "the whole." The ASA continues to require members to belong to ASA as a condition for joining a section, so that everyone has a connection to the discipline at large as well as to their

specialty groups. This approach has prevented the ASA from becoming a federation of sections and probably has minimized “split off” groups. The annual meeting grew by a thousand participants in the decade of the 1990s, now topping 5,000 people who find professional development in the broad program as well as in section involvement.

The growth of the ASA is also reflected in the growth of the number of journal publications. Since 1936, when the first issue of the *American Sociological Review* was published, ASA publications have expanded to include seven additional journals: *Contemporary Sociology*; *Journal of Health and Social Behavior*; *Social Psychology Quarterly*; *Sociological Methodology*; *Sociological Theory*; *Sociology of Education*; and *Teaching Sociology*. In addition, a *Rose Series* (funded from the estate of Arnold and Caroline Rose) publishes monographs that are important “small market” books in sociology. That series shifted from very specialized academic monographs to integrative pieces of broad appeal. The ASA’s newest journal is a general perspectives journal, aimed at the social science community (including students), and the educated lay public.

The Sydney S. Spivack Program in Applied Social Research and Social Policy (funded from a donation from Spivack’s estate) sponsors Congressional seminars and media briefings on timely topics for which there is a body of sociological knowledge. From these events, the ASA has published a series of issue briefs on topics ranging from youth violence to welfare-to-work to immigration to affirmative action. These publications are useful to ASA members but also to a wider public audience.

The teaching of sociology in kindergarten through high school and in undergraduate and graduate schools has received varying degrees of emphasis over the course of the ASA’s history. In particular, in the late 1970s and early 1980s, when sociology enrollments and membership in the ASA were at a low point, the ASA Teaching Services Program was developed. Now part of the Academic and Professional Affairs Program (APAP), the Teaching Services Program includes providing opportunities through seminars and workshops to improve classroom teaching and to examine a wide range of new curricula for almost all sociology courses. The Teaching Resources Center in the

ASA’s executive office produces over a hundred resources, including syllabi sets and publications on topics such as classroom techniques, curriculum, departmental management, and career information. APAP works concertedly with departments and chairs to build strong departments of sociology and excellent curricula. The ASA sponsors an annual conference for chairs, a meeting of directors of graduate study, and an electronic broadcast, CHAIRLINK, for chairs.

Odd as it may seem, the ASA often did not collect or have access to data on the profession. In 1993, the Research Program on the Discipline and Profession remedied the situation by conducting surveys of members and departments, and a tracking survey of a cohort of Ph.D.s. The program routinely publishes “research briefs” that share these data and aid departments and individuals with planning and trend analysis.

#### **TRANSITION FROM SECRETARIAT TO A PROFESSIONAL ASSOCIATION**

The ASA has undergone an organizational transformation over its century of serving the professional interests of sociologists. The shape and mission of the executive office reflects the shifts. In the early years, the office was essentially a secretariat—a place where records were kept, and dues and payments were processed. The first executive officer, Matilda White Riley, appointed in 1963, jokes that the office was a file card box on her kitchen table.

As the American Sociological Society (as it was named until 1959) grew and flourished, it adopted the model of a “learned society,” primarily concerned with the production of new disciplinary knowledge. The society/association centered its resources on the annual meeting and the journals. The executive office personnel, primarily clerical, staffed those functions.

The expansion period of the 1960s and 1970s, and the societal context of those times, led the ASA to add many programs and activities (for a detailed description, see Simpson and Simpson 1994). The ASA transformed into a professional association, with a wider range of services and benefits to its members as well as to a broader

**1998 Section Totals**

	Total	Low Income	Student	Regular
1. Undergraduate Education	419	12	62	345
2. Methodology	414	13	118	283
3. Medical Sociology	1,021	35	306	680
4. Crime, Law, and Deviance	634	16	236	382
5. Sociology of Education	579	12	182	385
6. Family	759	30	216	513
7. Organizations, Occupations, and Work	1,062	27	346	689
8. Theory	691	19	180	492
9. Sex and Gender	1,114	32	405	677
10. Community and Urban Sociology	553	16	175	362
11. Social Psychology	666	19	260	387
12. Peace, War and, Social Conflict	274	7	69	198
13. Environment and Technology	401	15	126	260
14. Marxist Sociology	362	11	103	248
15. Sociological Practice	308	13	36	259
16. Sociology of Population	406	8	84	314
17. Political Economy of the World System	416	9	151	256
18. Aging and the Life Course	563	16	147	400
19. Mental Health	408	17	116	275
20. Collective Behavior and Social Movements	568	10	222	336
21. Racial and Ethnic Minorities	685	15	227	443
22. Comparative Historical Sociology	522	5	150	367
23. Political Sociology	580	9	207	364
24. Asia/Asian America	325	6	117	202
25. Sociology of Emotions	279	9	96	174
26. Sociology of Culture	843	27	315	501
27. Science, Knowledge, and Technology	390	13	147	230
28. Computers, Sociology and	263	11	67	185
29. Latino/a Sociology	247	5	86	156
30. Alcohol and Drugs	247	10	57	180
31. Sociology of Children	330	10	85	235
32. Sociology of Law	316	3	123	190
33. Rational Choice	183	3	37	143
34. Sociology of Religion	535	19	173	343
35. International Migration	283	12	83	188
36. Race, Gender, and Class	830	20	382	428
37. Mathematical Sociology	206	4	76	126
38. Section on Sexualities	281	12	137	132
39. History of Sociology	215	5	42	168
<b>Totals</b>	<b>19,178</b>	<b>535</b>	<b>6,147</b>	<b>12,496</b>

Table 2

public. As Simpson and Simpson (1994) describe the change: “ASA reacted to the pressures of the 1960s and 1970s mainly by absorbing the pressure groups into its structure. A result has been to expand the goals and functions of the association beyond its initial disciplinary objective. Functions are more differentiated now, encompassing more professional and activist interests” (p. 265). The executive office grew slightly, with the growth in Ph.D.-level sociology staff who led these new ventures. In addition to the executive officer, these sociologists had titles (e.g., Staff Sociologist for Minorities, Women, and Careers) that reflected their work.

In the last ten years, some significant organizational changes have occurred. The executive office has been professionalized, with new hires often having at least a B.A. in sociology. The senior sociology staff direct the core programs (see below) and no longer have fixed terms of employment. As such, the office is more programmatic and proactive.

Key changes in the ASA’s governance include:

- Passage of an ASA mission and goals statement, with six core programs in the executive office (the six programs are:

Minority Affairs; Academic and Professional Affairs; Public Affairs; Public Information; Research on the Discipline and Profession; and the Spivack Program),

- The launch of the ASA's Spivack Program in applied social research and social policy and a continuous, intentional effort to bring research to bear on public policy issues,
- The beginning, and end, of the ASA's certification program,
- The beginning, and end, of the journal *Sociological Practice Review*,
- The beginning of the MOST program (Minority Opportunities for Summer Training for the first five years and Minority Opportunities through School Transformation, for the next five years, funded by the Ford Foundation),
- Greater autonomy for sections and an increase in the number of sections,
- New policies on ASA resolutions and policymaking,
- Revision of the Code of Ethics, as an educative document, which serves as a model for aligned sociological groups,
- Restructuring of ASA committees to a more targeted "task force" model,
- Addition of a new "perspectives" journal,
- Increased attention to science policy and funding, including collaborations with many other groups,
- Increased attention to sociology departments as units, and to chairs as their leaders, as shown in the Department Affiliates program,
- Active collaboration between the ASA and higher education organizations such as the American Association for Higher Education and the American Association of Colleges and Universities.

Members (and nonmember sociologists) have differing views about the current form of the ASA, which is discussed at the end of this article.

The ASA is but one organization in a network of sociological organizations or associations in

which sociologists comprise a significant part of the membership. These groups operate in a complementary way to the ASA; some were formed in juxtaposition to the ASA to fulfill a need the ASA was not serving or to pressure the ASA to change. The genres of these organizations are briefly discussed below.

### REGIONAL AND STATE ASSOCIATIONS

In many disciplines, a national association includes state and regional chapters. In sociology, those regional and state groups have always been independent entities, with their own dues, meetings, and journals. Nonetheless, the ASA has worked collaboratively with these associations. The ASA sends staff representatives to their annual meetings, offers to serve on panels and meet with their councils, sends materials and publications, and convenes a meeting of regional and state presidents at the ASA's annual meeting. In the 1960s regional representatives sat on the ASA council. Everett Hughes (1962) provided a sociological critique of this approach to governance, suggesting that the ASA was a disciplinary not a professional association. He argued that the ASA should not be organized as a federation of such representatives, and this regional delegate format ended in 1967. Later, candidates for the Committee on Committees and the Committee on Nominations were nominated by district (not identical to regional associations) to ensure regional representation. That approach ended in 1999.

Twenty-four states (or collaborations among states) have sociological associations. Some are extremely active (e.g., Wisconsin, Minnesota, North Carolina, Illinois, Georgia) and have some special foci that link to their state-based networks. The Georgia Sociological Association, for example, sponsors a workshop for high school teachers; the association also honors a member of the media for the best presentation of sociological work. Wisconsin sociologists used the Wisconsin Sociological Association to organize to defeat a licensure bill that would have prevented sociologists from employment in certain social service jobs. The Minnesota sociologists have made special outreach efforts to practitioners and include these colleagues on their board.

**Regional Sociology Associations**

Regional Organization	Founding Date	# of Members (1999)	Journal
Eastern Sociological Society	1930	1,000	<i>Sociological Forum</i>
Mid-South Sociological Association	1975	293	<i>Sociological Spectrum</i>
Midwest Sociological Society	1936	1,250	<i>The Sociological Quarterly</i>
New England Sociological Association		250	none
North Central Sociological Association	1925	442	<i>Sociological Focus</i>
Pacific Sociological Association	1929	1,350	<i>Sociological Perspectives</i>
Southern Sociological Society	1935	1,748	<i>Social Forces</i>
Southwestern Sociological Association	1923	507	<i>Supports Social Science Quarterly</i>
District of Columbia Sociological Society	1934	200	none

Table 3

**ALIGNED ASSOCIATIONS**

Many of the aligned associations offer a small, vital intellectual home for sociologists interested in a particular specialty. Over time, a few of these organizations have consciously decided to form a section within the ASA. The various sociology of religion groups did so recently, to have a “place” within the ASA as well as their own meeting and publications. Many of the aligned groups have members from many disciplines including sociology.

**INTERDISCIPLINARY TIES**

The most significant interdisciplinary organization is the Consortium of Social Science Associations (COSSA), formed in 1980. The budget cutting of the Reagan administration served as a catalyst for the major social science associations to establish this umbrella organization to lobby for funding for social science research. In the 1980s and 1990s, COSSA, with its own professional staff, has become a well-respected voice on social science policy, federal funding, and the professional concerns of social scientists (e.g., data archiving, confidentiality protection, and support for research on controversial topics). COSSA is, of course, an organization of organizations.

In 1997, the ASA offered membership discounts with other societies, so that individuals could join several of these groups. The interdisciplinary discounts now apply to: the American Political Science Association, the American Educational Research Association, the Society for Research in Child Development, and the Academy of Management.

**POLITICAL PRESSURES ON AND IN ASA: FEMINISM AND ACTIVISM**

Of the many aligned associations, two organizations provide association missions that the ASA does not (or does not sufficiently) satisfy, and have an agenda to change the ASA.

Sociologists for Women in Society (SWS), founded in 1970, has had a two-pronged mission: to use the tools and talents in sociology to improve the lives of women in society; and secondly, to enhance the participation, status, and professional contributions of feminist sociologists. Most SWS members are also ASA members. Originally named the Women’s Caucus, the group’s 1970 statement of demands summarized the gender issues in the ASA quite clearly: “What we seek is effective and dramatic action; an unbiased policy in the selection of stipend support of students; a concerted commitment to the hiring and promotion of women sociologists to right the imbalance that is represented by the current situation in which 67 percent of the women graduate students in this country do not have a single woman sociology professor of senior rank during the course of their graduate training, and when we participate in an association of sociologists in which NO woman will sit on the 1970 council, NO woman is included among the associate editors of the *American Sociological Review*, and NO woman sits on the thirteen member committee on publications and nominations” (Roby, p. 24). Over time, as the founding mothers of SWS moved through their career trajectories, more and more SWS members became part of the leadership of the ASA. In 2000, nine of twenty ASA council members are women. Since its founding, SWS has sought to pressure the ASA in more feminist directions, and to supplement what the

**State and Aligned Sociological Organizations**

<b>State Associations</b>	<b>International Associations</b>
National Council of State Sociological Associations	Asia Pacific Sociological Association
Alabama/Mississippi Sociological Association	Australian Sociological Association
Arkansas Sociological Association	British Sociological Association
California Sociological Association	Canadian Sociology & Anthropology Association
Florida Sociological Association	European Society for Rural Sociology
Georgia Sociological Association	European Sociological Association
Great Plains Association (North and South Dakota)	International Institute of Sociology
Hawaii Sociological Association	International Network for Social Network Analysis
Illinois Sociological Association	International Society for the Sociology of Religion
Iowa Sociological Association	International Sociological Association
Kansas Sociological Association	International Visual Sociology Association
Anthropologists and Sociologists of Kentucky	Sociological Association of Aotearoa (New Zealand)
Michigan Sociological Society	<b>Social Science/Interdisciplinary Associations</b>
Sociologists of Minnesota	Consortium of Social Science Associations
Missouri State Sociological Association	Academy of Management
Nebraska Sociological Association	American Association for the Advancement of Science
New York State Sociological Association	American Association for Public Opinion Research
North Carolina Sociological Association	American Educational Research Association
Oklahoma Sociological Association	American Evaluation Association
Pennsylvania Sociological Society	American Society of Criminology
South Carolina Sociological Association	American Statistical Association
Virginia Social Science Association	Association of Gerontology in Higher Education
Washington Sociological Association	Council of Professional Associations on Federal Statistics
West Virginia State Sociological Association	Gerontological Society of America
Wisconsin Sociological Association	Law and Society Association
<b>Aligned Associations</b>	Linguistic Society of America
Alpha Kappa Delta	National Council on Family Relations
Anabaptist Sociology and Anthropology Association	National Council for the Social Studies
Association of Black Sociologists	Popular Culture Association
Association of Christians Teaching Sociology	Population Association of America
Association for Humanist Sociology	Religious Research Association
Association for the Sociology of Religion	Social Science History Association
Chicago Sociological Practice Association	Society for Research in Child Development
Christian Sociological Society	Society for the Scientific Study of Religion
North American Society for the Sociology of Sport	Society for Social Studies of Science
Rural Sociological Society	<b>Commissions</b>
Society for the Advancement of Socio-Economics	Commission on Applied and Clinical Sociology
Society for Applied Sociology	
Society for the Study of Social Problems	
Society for the Study of Symbolic Interaction	
Sociological Practice Association	
Sociologists' AIDS Network	
Sociologists for Women in Society	
Sociologists' Lesbian, Gay, Bisexual, and Transgender Caucus	
Sociology of Education Association	

Table 4

ASA seemed not to offer. SWS runs its annual meeting program parallel to the ASA's annual meeting. In the earlier years, many sessions concentrated on work in sex and gender, as well as on informal networking, and mentoring workshops. Over time, as the ASA's section on sex and gender has become the largest in the association, the SWS program has downplayed scholarly papers on sex and gender, and has emphasized instead informal networking, socializing, and political organizing.

SWS proposed that the ASA publish a journal on *Sex & Gender*, but the ASA declined, due to a rather full publication portfolio. SWS entered an agreement with Sage Publishers to start such a journal, which has been a successful intellectual and business venture. At various points in its history, SWS has been explicit in its watchdog role over the ASA. Members came to observe the ASA council meetings; the membership endorsed candidates for ASA offices; and candidates were asked

to complete a survey that was sent to all SWS members.

The Society for the Study of Social Problems (SSSP), founded in 1951, pursued the insider-outsider strategy as well. The SSSP meets prior to (often with a day of overlap) the ASA annual meeting. As the name implies, the topics for sessions and for the divisions deal with social problems, what sociologists know about them, and their solutions. The SSSP journal, *Social Problems*, is well regarded and well subscribed.

In the 1950s, the ASA centered on positivism and “objective” scientific pursuits. Twenty years earlier, a group of ASA members had warned that the achievement of scientific status and academic acceptance were hindered by the application of sociology to social problems. The motion read, in part: “The undersigned members, animated by an ideal of scientific quality rather than of heterogeneous quantity, wish to prune the Society of its excrescences and to intensify its scientific activities. This may result in a reduction of the members and revenues of the society, but this is preferable to having many members whose interest is primarily or exclusively other than scientific” (Rhoades 1981, pp. 24–25). The SSSP has been a vital counterpoint to those views, keeping sociology’s leftist leanings alive.

### THE APPLIED SIDE

In 1978, two new sociological associations formed to meet the needs of sociological practitioners. The Clinical Sociology Association (CSA), now called the Sociological Practice Association (SPA), centered on sociologists engaged in intervention work with small groups (e.g., family counseling) or at a macro-level (e.g., community development). This group emphasized professional training and credentials. Most of the members were employed primarily outside of the academe; many felt they needed additional credentials to meet state licensure requirements or to receive third-party payments, or both. The SPA established a rigorous certification program, where candidates with a Ph.D. in sociology and substantial supervised experience in clinical work, would present their credentials and make a presentation as part of the application for certification. Those who passed this review could use the title Certified Clinical Sociologist.

The Society for Applied Sociology (SAS) was formed by a group of colleagues in Ohio, most of whom are primarily academics, but who engage in extensive consulting and applied work. The core of the SAS centers on applied social research, evaluation research, program development, and other applications of sociological ideas to a variety of organizational settings. The SAS has worked extensively with curriculum and program development to prepare the next generation of applied sociologists. The SAS has also focused on the master’s-level sociologist much more than other organizations.

Both of these practice organizations hold an annual meeting and sponsor a journal. At various times in the twenty years of each group, members have advocated a merger to reduce redundancy, strengthen the membership base, and use resources together. One place where the two groups have worked in tandem is through their joint Commission on Applied and Clinical Programs, which accredits sociology programs that meet the extensive criteria set forth by the commission. In this sense, these applied sociology programs (usually a part or a track within a regular sociology department) are modeling professional programs such as social work, which have an accrediting mechanism. Both societies held a joint meeting prior to the ASA meeting in 2000, which may portend future collaboration.

The history of the ASA shows the ebb and flow in interest in applied sociology, certainly going back to President Lester Frank Ward, and evident again with the election of contemporary Presidents William Foote Whyte, Peter Rossi, and Amitai Etzioni. Within the ASA, there is an active, though not large, section on sociological practice, drawing overlapping membership with the SPA and the SAS. The ASA published a journal, *Sociological Practice Review*, as a five-year experiment (1990–1995) but dropped the publication when there were insufficient subscribers and few manuscripts. In the early 1980s, in response to member interest, the ASA began a certification program, through which Ph.D.-level sociologists could be certified in six areas (demography, law and social control, medical sociology, organizational analysis, social policy and evaluation research, and social psychology). At the master’s level, sociologists could take an exam; passing the test would result in



certification in applied social research. The certification program received few applicants and was terminated by the ASA council in 1998.

Since the 1980s, there have been forces pushing for more involvement of practitioners in the ASA, and thus more membership benefits to serve non-academic members, as the paring down of those benefits (as in the case of the *Sociological Practice Review* and certification) when interest wanes. In part, there is greater “within group” variance among practitioners than “between group” variance between practitioners and academics. Thus while there is clearly a *political constituency* for applied work, it is less clear there is a core intellectual constituency.

In 1981, the ASA held a conference on applied sociology, from which a book of proceedings was published by Jossey-Bass. That event was a springboard for the introduction of applied issues within the ASA. A committee on sociological practice, a section on sociological practice, and an ASA award for a distinguished career in sociological practice were created. In 1999, about 25 percent of ASA members had primary employment in nonacademic positions; the estimate of the number of Ph.D.s (some of whom were nonmembers) in sociological practice was higher. The diversity of the nonacademic membership and their professional needs has been a challenge to the ASA. In a 1984 article in the *American Sociological Review*, Howard Freeman and Peter Rossi wrote of the significant changes that might be needed in departments of sociology and in the reward structure of the profession to reduce the false dichotomy between applied and basic research.

### COMMITMENT TO DIVERSITY

The ASA has made concerted efforts to be inclusive of women and minorities in its activities and governance. Since 1976, the ASA has undertaken a biennial report on the representation of women and minorities in the program (invited events and open submissions), on editorial boards, in elections, and in the governance (committees) of the ASA.

In 1987, the ASA appointed a task force on participation designed to identify ways to more fully enfranchise colleagues in two- and four-year colleges. That task force held a number of open hearings and met for five years before issuing a

report of recommendations to the council. As a spin-off from the committee on teaching, a task force on community colleges made recommendations to the council in 1997 and 1998 about how to more actively involve these colleagues in ASA affairs.

The ASA council adopted the following policy in 1997:

*Much of the vitality of ASA flows from its diverse membership. With this in mind, it is the policy of the ASA to include people of color, women, sociologists from smaller institutions or who work in government, business, or other applied settings, and international scholars in all of its programmatic activities and in the business of the Association.*

At the same time, the demographics of the profession have been shifting (Roos and Jones 1993). Over 55 percent of new Ph.D.s are women, and about 45 percent of ASA members are female. The Minority Fellowship Program, begun in 1974, has provided predoctoral funding and mentoring support for minority sociologists. The program boasts an astounding graduation record of 214 Ph.D.s; many of these colleagues from the early cohorts are now senior leaders in departments, organizations, and in the ASA.

### DEMOCRATIZATION OR CONSOLIDATION?

The ASA membership has diverse views about the extent to which the current organizational structure and goals are optimal. Simpson and Simpson argue that core disciplinary concerns have taken a back seat at the ASA; they speak of the “disciplinary elite and their dilution” (1994, p. 271). Their analysis of ASA budgets, as a indicator of priorities, shows shifts from disciplinary concerns (e.g., journals and meetings) to professional priorities (e.g., jobs, teaching, applied work, and policy issues). Other segments of the profession allege that the ASA leadership is too elite (Reynolds 1998) and has a falsely rosy view of the field (p. 20). Demographically and programmatically, the ASA has changed in its century of service to sociology. With a solid membership core and generally positive trends in the profession, the ASA will continue to sit at the hub of a network of sections and aligned organizations.



## REFERENCES

- Collins, Randall 1989 "Future Organizational Trends of the American Sociological Association." *Footnotes* 17:1-5, 9.
- Freeman, Howard, and Peter H. Rossi 1984 "Furthering the Applied Side of Sociology." *American Sociological Review*, vol. 49. 4 (August): 571-580.
- Hughes, Everett C. 1962 "Association or Federation?" *American Sociological Review* 27:590.
- Reynolds, Larry T. 1998 "Two Deadly Diseases and One Nearly Fatal Cure: The Sorry State of American Sociology." *The American Sociologist*. vol. 29. 1:20-37.
- Rhoades, Lawrence J. 1981 *A History of the American Sociological Association: 1905-1980*. Washington, D.C.: American Sociological Association.
- Roby, Pamela 1992 "Women and the ASA: Degendering Organizational Structures and Processes, 1964-1974." *The American Sociologist* (Spring):18-48.
- Roos, Patricia A., and Katharine Jones 1993 "Shifting Gender Boundaries: Women's Inroads into Academic Sociology." *Work and Occupations*, vol. 20. 4 (November): 395-428.
- Simpson, Ida Harper, and Richard Simpson 1994 "The Transformation of ASA." *Sociological Forum*, vol. 9. 2 (June):259-278.

CARLA B. HOWERY

## ANALYSIS OF VARIANCE AND COVARIANCE

Analysis of variance (ANOVA) and analysis of covariance (ANACOVA) are statistical techniques most suited for the analysis of data collected using experimental methods. As a result, they have been used more frequently in the fields of psychology and medicine and less frequently in sociological studies where survey methods predominate. These techniques can be, and have been, used on survey data, but usually they are performed within the analysis framework of linear regression or the "general linear model." Given their applicability to experimental data, the easiest way to convey the logic and value of these techniques is to first review the basics of experimental design and the analysis of experimental data. Basic concepts and procedures will then be described, summary measures and assumptions reviewed, and the applicability of these techniques for sociological analysis discussed.

## EXPERIMENTAL DESIGN AND ANALYSIS

In a classical experimental design, research subjects are randomly assigned in equal numbers to two or more discrete groups. Each of these groups is then given a different treatment or stimulus and observed to determine whether or not the different treatments or stimuli had predicted effects on some outcome variable. In most cases this outcome variable has continuous values rather than discrete categories. In some experiments there are only two groups—one that receives the stimulus (the experimental group) and one that does not (the control group). In other studies, different levels of a stimulus are administered (e.g., studies testing the effectiveness of different levels of drug dosages) or multiple conditions are created by administering multiple stimuli separately and in combination (e.g., exposure to a violent model and reading pacifist literature).

In all experiments, care is taken to eliminate any other confounding influences on subjects' behaviors by randomly assigning subjects to groups. As a result of random assignment, preexisting differences between subjects (such as age, gender, temperament, experience, etc.) are randomly distributed across groups making the groups equal in terms of the potential effects of these preexisting differences. Since each group contains approximately equal numbers of subjects of any given age, gender, temperament, experience, etc., there should be no differences between the groups on the outcome variable that are due to these confounding influences. In addition, experiments are conducted in standardized or "physically controlled" situations (e.g., a laboratory), thus eliminating any extraneous external sources of difference between the groups. Through random assignment and standardization of experimental conditions, the researcher is able to make the qualifying statement "Other things being equal. . ." and assert that any differences found between groups on the outcome measure(s) of interest are due solely to the fact that one or more groups received the experimental stimulus (stimuli) and the other group did not.

**Logic of analysis procedures.** Analysis of variance detects effects of an experimental stimulus by first computing means on the outcome variable for the experimental and control groups and then comparing those means. If the means are the

same, then the stimulus didn't "make a difference" in the outcome variable. If the means are different, then, because of random assignment to groups and standardization of conditions, it is assumed that the stimulus caused the difference. Of course, it is necessary to establish some guidelines for interpreting the size of the difference in order to draw conclusions regarding the strength of the effect of the experimental stimulus. The criterion used by analysis of variance is the amount of random variation that exists in the scores within each group. For example, in a three-group comparison, the group mean scores on some outcome measure may be 3, 6, and 9 on a ten-point scale. The meaningfulness of these differences can only be assessed if we know something about the variation of scores in the three groups. If everyone in group one had scores between 2 and 4, everyone in group two had scores between 5 and 7, and everyone in group three had scores between 8 and 10, then in every instance the variation within each group is low and there is no overlap across the within group distributions. As a result, whenever the outcome score of an individual in one group is compared to the score of an individual in a different group, the result of the comparison will be very similar to the respective group mean score comparison and the conclusions about who scores higher or lower will be the same as that reached when the means were compared. As a result, a great deal of confidence would be placed in the conclusion that group membership makes a difference in one's outcome score. If, on the other hand, there were several individuals in each group who scored as low or as high as individuals in other groups—a condition of high variability in scores—then comparisons of these subjects' scores would lead to a conclusion opposite to that represented by the mean comparisons. For example, if group one scores varied between 1 and 5 and group two scores varied between 3 and 10, then in some cases individuals in group one scored higher than individuals in group two (e.g., 5 vs. 3) even though the mean comparisons show the opposite trend (e.g., 3 vs. 6). As a result, less confidence would be placed in the differences between means.

Although analysis of variance results reflect a comparison of group means, conceptually and computationally this procedure is best understood through a framework of explained variance. Rather than asking "How much difference is there

between group means?" the question becomes "How much of the variation in subjects' scores on the outcome measure can be explained or accounted for by the fact that subjects were exposed to different treatments or stimuli?" To the extent that the experimental stimulus has an effect (i.e., group mean differences exist), individual scores should differ from one another because some have been exposed to the stimulus and others have not. Of course, individuals will differ from one another for other reasons as well, so the procedure involves a comparison of how much of the total variation in scores is due to the stimulus effect (i.e., group differences) and how much is due to extraneous factors. Thus, the total variation in outcome scores is "decomposed" into two elements: variation due to the fact that individuals in the different groups were exposed to different conditions, experiences, or stimuli (explained variance); and variation due to random or chance processes (error variance). Random or chance sources of variation in outcome scores can be such things as measurement error or other causal factors that are randomly distributed across groups through the randomization process. The extent to which variation is due to group differences rather than these extraneous factors is an indication of the effect of the stimulus on the outcome measure. It is this type of comparison of components of variance that provides the foundation for analysis of variance.

### BASIC CONCEPTS AND PROCEDURES

The central concept in analysis of variance is that of *variance*. Simply put, variance is the amount of difference in scores on some variable across subjects. For example, one might be interested in the effect of different school environments on the self-esteem of seventh graders. To examine this effect, random samples of students from different school environments could be selected and given questionnaires about their self-esteem. The extent to which the students' self-esteem scores differ from each other both within and across groups is an example of the variance.

Variance can be measured in a number of ways. For example, simply stating the range of scores conveys the degree of variation. Statistically, the most useful measures of variation are based on the notion of the *sums of squares*. The sums of squares is obtained by first characterizing a sample

or group of scores by calculating an average or mean. The mean score can be thought of as the score for a “typical” person in the study and can be used as a reference point for calculating the amount of differences in scores across all individuals. The difference between each score and the mean is analogous to the difference of each score from every other score. The variation of scores is calculated, then, by subtracting each score from the mean, squaring it, and summing the squared deviations (squaring the deviations before adding them is necessary because the sum of nonsquared deviations from the mean will always be 0). A large sum of squares indicates that the total amount of deviation of scores from a central point in the distribution of scores is large. In other words, there is a great deal of variation in the scores either because of a few scores that are very different from the rest or because of many scores that are slightly different from each other.

**Decomposing the sums of squares.** The total amount of variation in a sample on some outcome measure is referred to as the *total sums of squares* ( $SS_{\text{total}}$ ). This is a measure of how much the subjects’ scores on the outcome variable differ from one another and it represents the phenomenon that the researcher is trying to explain (e.g., Why do some seventh graders have high self-esteem, while others have low or moderate self-esteem?). The procedure for calculating the total sums of squares is represented by the following equation:

$$SS_{\text{TOTAL}} = \sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 \quad (1)$$

where,  $\sum_i \sum_j$  indicates to sum across all individuals ( $i$ ) in all groups ( $j$ ), and  $(Y_{ij} - \bar{Y}_{..})^2$  is the squared difference of the score of each individual ( $Y_{ij}$ ) from the grand mean of all scores ( $\bar{Y}_{..} = \sum_{ij} Y_{ij} / N$ ). In terms of explaining variance, this is what the researcher is trying to account for or explain.

The total sums of squares can then be “decomposed” or mathematically divided into two components: the *between-groups sums of squares* ( $SS_{\text{BETWEEN}}$ ) and the *within-groups sums of squares* ( $SS_{\text{WITHIN}}$ ).

The between-groups sums of squares is a measure of how much variation in outcome scores exists between groups. It uses the group mean as the best single representation of how each individual in the group scored on the outcome measure. It essentially assigns the group mean score to

every subject in the group and then calculates how much total variation there would be from the grand mean (the average of all scores regardless of group membership) if there was no variation within the groups and the only variation comes from cross-group comparisons. For example, in trying to explain variation in adolescent self-esteem, a researcher might argue that junior high schools place children at risk because of schools’ size and impersonal nature. If the school environment is the single most powerful factor in shaping adolescent self-esteem, then when adolescents are compared to each other in terms of their self-esteem, the only comparisons that will create differences will be those occurring between students in different school types, and all comparisons involving children in the same school type will yield no difference. The procedure for calculating the between-groups sums of squares is represented by the following equation:

$$SS_{\text{BETWEEN}} = \sum_j N_j (\bar{Y}_{.j} - \bar{Y}_{..})^2 \quad (2)$$

where,  $\sum_j$  indicates to sum across all groups ( $j$ ), and  $N_j (\bar{Y}_{.j} - \bar{Y}_{..})^2$  is the number of subjects in each group ( $N_j$ ) times the difference between the mean of each group ( $\bar{Y}_{.j}$ ) and the grand mean ( $\bar{Y}_{..}$ ).

In terms of the comparison of means, the between-groups sums of squares directly reflects the difference between the group means. If there is no difference between the group means, then the group means will be equal to the grand mean and the between-groups sums of squares will be 0. If the group means are different from one another, then they will also differ from the grand mean and the magnitude of this difference will be reflected in the between-groups sums of squares. In terms of explaining variance, the between-groups sums of squares represents only those differences in scores that come about because the individuals in one group are compared to individuals in a different group (e.g., What if all students in a given type of school had the same level of self-esteem, but students in a different school type had different levels?). By multiplying the group mean difference score by the number of subjects in the group, this component of the total variance assumes that there is no other source of influence on the scores (i.e., that the variance within groups is 0). If this assumption is true, then the  $SS_{\text{BETWEEN}}$  will be equal to the  $SS_{\text{TOTAL}}$  and the group effect could be

said to be extremely strong (i.e., able to overshadow any other source of influence). If, on the other hand, this assumption is not true, then the  $SS_{\text{BETWEEN}}$  will be small relative to the  $SS_{\text{TOTAL}}$  because other influences randomly dispersed across groups are generating differences in scores not reflected in mean score differences. This situation indicates that group differences in experimental conditions add little to our ability to predict or explain differences in outcome scores. In the extreme case, when there is no difference in the group means, the  $SS_{\text{BETWEEN}}$  will equal 0, indicating no effect.

The within-groups sums of squares is a measure of how much variation exists in the outcome scores within the groups. Using the same example as above, adolescents in a given type of school might differ in their self-esteem in spite of the esteem-inflating or -depressing influence of their school environment. For example, elementary school students might feel good about themselves because of the supportive and secure nature of their school environment, but some of these students will feel worse than others because of other factors such as their home environments (e.g., the effects of parental conflict and divorce) or their neighborhood conditions (e.g., wealth vs. poverty). The procedure for calculating the within-groups sums of squares is represented by the following equation:

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_j)^2 \quad (3)$$

where  $\sum_i \sum_j$  indicates to sum across all individuals ( $i$ ) in all groups ( $j$ ), and  $(Y_{ij} - \bar{Y}_j)^2$  is the squared difference between the individual scores ( $Y_{ij}$ ) and their respective group mean scores ( $\bar{Y}_j$ ).

Both in terms of comparing means and explaining variance, the within-groups sums of squares represents the variance due to other factors or “error”; it is the degree of variation in scores despite the fact that individuals in a given group were exposed to the same influences or stimuli. This component of variance can also serve as an estimate of how much variability in outcome scores occurs in the population from which each group of respondents was drawn. If the within-groups sums of squares is high relative to the between-groups sums of squares (or the difference between the means), then less confidence can be placed in the conclusion that any group differences are meaningful.

## SUMMARY MEASURES

Analysis of variance procedures produce two summary statistics. The first of these— $ETA^2$ —is a measure of how much effect the predictor variable or factor has on the outcome variable. The second statistic— $F$ —tests the *null hypothesis* that there is no difference between group means in the larger population from which the sample data was randomly selected.

**$ETA^2$ .** As noted above, a large between-groups sums of squares is indicative of a large difference in the mean scores between groups. The meaningfulness of this difference, however, can only be judged against the overall variation in the scores. If there is a large amount of variation in scores relative to the variation due exclusively to between-groups differences, then group effects can only explain a small proportion of the total variation in scores (i.e., a weak effect).  $ETA^2$  takes into account the difference between means and the total variation in scores. The general equation for computing  $ETA^2$  is as follows:

$$ETA^2 = \frac{SS_{\text{BETWEEN}}}{SS_{\text{TOTAL}}} \quad (4)$$

As can be seen from this equation,  $ETA^2$  is the proportion of the total sums of squares explained by group differences. When all the variance is explained, there will be no within-group variance, leaving  $SS_{\text{TOTAL}} = SS_{\text{BETWEEN}}$  ( $SS_{\text{TOTAL}} = SS_{\text{BETWEEN}} + 0$ ). Thus,  $ETA^2$  will be equal to 1, indicating a perfect relationship. When there is no effect, there will be no difference in the group means ( $SS_{\text{BETWEEN}} = 0$ ) and  $ETA^2$  will be equal to 0.

**F Tests.** Even if  $ETA^2$  indicates that a sizable proportion of the total variance in the sample scores is explained by group differences, the possibility exists that the sample results do not reflect true differences in the larger population from which the samples were selected. For example, in a study of the effects of cohabitation on marital stability, a researcher might select a sample of the population and find that, among those in his or her sample, previous cohabitators have lower marital satisfaction than those without a history of cohabitation. Before concluding that cohabitation has a negative effect on marriage in the broader population, however, the researcher must assess the probability that, by chance, the sample used in

this study included a disproportionate number of unhappy cohabitators or overly satisfied noncohabitators, or both. There is always some probability that the sample will not be representative, and the  $F$  statistic utilizes probability theory (under the assumption that the sample was obtained through random selection) to assess that likelihood.

The logic behind the  $F$  statistic is that chance fluctuations in sampling are less likely to account for differences in sample means if the differences are large, if the variation in outcome scores in the population from which the sample was drawn is small, if the sample size is large, or if all these situations have occurred. Obviously, large mean differences are unlikely due to chance because they would require many more extremely unrepresentative cases to be selected into the sample. Selecting extreme cases, however, is more likely if there are many extremes in the population (i.e., the variation of scores is great). Large samples, however, reduce the likelihood of unrepresentative samples because any extreme cases are more likely to be counteracted by extremes in the opposite direction or by cases that are more typical.

The general equation for computing  $F$  is as follows:

$$F = \frac{MS_{\text{BETWEEN}}}{MS_{\text{WITHIN}}} \quad (5)$$

The  $MS_{\text{BETWEEN}}$  is the mean square for between-group differences. It is an adjusted version of the  $SS_{\text{BETWEEN}}$  and reflects the degree of difference between group means expressed as individual differences in scores. An adjustment is made to the  $SS_{\text{BETWEEN}}$  because this value can become artificially high by chance as a function of the number of group comparisons being made. This adjustment factor is called the *degrees of freedom* ( $DF_{\text{BETWEEN}}$ ) and is equal to the number of group comparisons ( $k-1$ , where  $k$  is the number of groups). The formula for the  $MS_{\text{BETWEEN}}$  is:

$$MS_{\text{BETWEEN}} = SS_{\text{BETWEEN}} / (k-1) \quad (6)$$

The larger the  $MS_{\text{BETWEEN}}$  the greater the value of  $F$  and the lower the probability that the sample results were due to chance.

The  $MS_{\text{WITHIN}}$  is the mean square for within-group differences. This is equivalent to the  $SS_{\text{WITHIN}}$ ,

with an adjustment made for the size of the sample minus the number of groups ( $DF_{\text{WITHIN}}$ ). Since the  $SS_{\text{WITHIN}}$  represents the amount of variation in scores within each group, it is used in the  $F$  statistic as an estimate of the amount of variation in scores that exists in the populations from which the sample groups were drawn. This is essentially a measure of the potential for error in the sample means. This potential for error is reduced, however, as a function of the sample size. The formula for the  $MS_{\text{WITHIN}}$  is:

$$MS_{\text{WITHIN}} = SS_{\text{WITHIN}} / (N-k) \quad (7)$$

As can be seen in equations six and seven, when the number of groups is high, the estimate of variation between groups is adjusted downward to account for the greater chance of variation. When the number of cases is high, the estimation of variation within groups is adjusted downward. As a result, the larger the number of cases being analyzed, the higher the  $F$  statistic. A high  $F$  value reflects greater confidence that any differences in sample means reflect differences in the populations. Using certain assumptions, the possibility that any given  $F$  value can be obtained by chance given the number of groups ( $DF1$ ) and the number of cases ( $DF2$ ) can be calculated and compared to the actual  $F$  value. If the chance probability is only 5 percent or less, then the null hypothesis is rejected and the sample mean differences are said to be “significant” (i.e., not likely due to chance, but to actual effects in the population).

## ADJUSTING FOR COVARIATES

Analysis of variance can be used whenever the predictor variable(s) has a limited number of discrete categories and the outcome variable is continuous. In some cases, however, an additional continuous predictor variable needs to be included in the analysis or some continuous source of extraneous effect needs to be “controlled for” before the group effects can be assessed. In these cases, analysis of covariance can be used as a simple extension of the analysis of variance model.

In the classical experimental design, the variable(s) being controlled for—the covariate(s)—is frequently some background characteristics or pre-test scores on the outcome variable that were not



adequately randomized across groups. As a result, group differences in outcome scores may be found and erroneously attributed to the effect of the experimental stimulus or group condition, when in fact the differences between groups existed prior to, or independent of, the presence of the stimulus or group condition.

One example of this situation is provided by Roberta Simmons and Dale Blyth (1987). In a study of the effects of different school systems on the changing self-esteem of boys and girls as they make the transition from sixth to seventh grade, these researchers had to account for the fact that boys and girls in these different school systems had different levels of self-esteem in sixth grade. Since those who score high on a measure at one point in time (T1) will have a statistical tendency to score lower at a later time (T2) and vice versa (a negative relationship), these initial differences could lead to erroneous conclusions. In their study, if boys had higher self-esteem than girls in sixth grade, the statistical tendency would be for boys to experience negative change in self-esteem and girls to experience positive change even though seventh-grade girls in certain school systems experience more negative influences on their self-esteem.

The procedure used in adjusting for covariates involves a combination of analysis of variance and linear regression techniques. Prior to comparing group means or sources of variation, the outcome scores are adjusted based upon the effect of the covariate(s). This is done by computing predicted outcome scores based on the equation:

$$\hat{Y} = a + b_1 X_1 \quad (8)$$

where  $\hat{Y}$  is the new adjusted outcome score,  $a$  is a constant, and  $b_1$  is the linear effect of the covariate ( $X_1$ ) on the outcome score ( $Y$ ). The difference between the actual score and the predicted score ( $Y_{ij} - \hat{Y}_{ij}$ ) is the residual. These residuals represent that part of the individuals' scores that is not explained by the covariate. It is these residuals that are then analyzed using the analysis of variance techniques described above. If the effect of the covariate is negative, then those who scored high on the covariate will have their scores adjusted upward. Those who scored lower on the covariate would have their scores adjusted downward. This would counteract the reverse effect that the covariate

has had. Group differences could then be assessed after the scores have been corrected.

This model can be expanded to include any number of covariates and is particularly useful when analyzing the effects of a discrete independent variable (e.g., gender, race, etc.) on a continuous outcome variable using survey data, where other factors cannot be randomly assigned and where conditions cannot be standardized. In such situations, other preexisting difference between groups (often, variables measured on continuous scales) need to be statistically controlled for. In these cases, researchers often perform analysis of variance and analysis of covariance within the context of what has been termed the "general linear model."

### GENERAL LINEAR MODEL

The *general linear model* refers to the application of the linear regression equation to solve analysis problems that initially do not meet the assumptions of linear regression analysis. Specifically, there are three situations where the assumptions of linear regression are violated but regression techniques can still be used: (1) the use of nominal level measures (e.g., race, religion, marital status) as independent variables—a violation of the assumption that all variables be measured at the interval or ratio level; (2) the existence of interaction effects between independent variables—a violation of the assumption of additivity of effects; and (3) the existence of a curvilinear effect of the independent variable on the dependent variable—a violation of the assumption of linearity. The linear regression equation can be applied in all of these situations provided that certain procedures and operations on the variables are carried out. The use of the general linear model for performing analysis of variance and analysis of covariance is described in greater detail below.

**Regression with dummy variables.** In situations where the dependent variable is measured at the interval level of measurement (ordered values at fixed intervals) but one or more independent variables are measured at the nominal level (no order implied between values), analysis of variance and covariance procedures are usually more appropriate than linear regression. Linear regression analysis can be used in these circumstances, however, as long as the nominal level variables are

first “dummy coded.” The results will be consistent with those obtained from an analysis of variance and covariance, but can also be interpreted within a regression framework.

Dummy coding is a procedure where a separate dichotomous variable is created for each category of the nominal level variable. For example, in a study of the effects of racial experience, the variable for race can have several values that imply no order or degree. Some of the categories might be white, black, Latino, Indian, Asian, and other. Since these categories imply no order or degree, the variable for race cannot be used in a linear regression analysis.

The alternative is to create five dummy variables—one for each race category except one. Each of these variables measures whether or not the respondent is the particular race or not. For example, the first variable might be for the category “white,” where a value of 0 is assigned if the person is any other race but white, and a value of 1 is assigned if the person is white. Similarly, separate variables would be created to identify membership in the black, Latino, Indian, and Asian groups. A dummy variable is not created for the “other” category because its values are completely determined by values on the other dummy variables (e.g., all persons with the racial category “other” will have a score of 0 on all of the dummy variables). This determination is illustrated in the table below.

Since there are only two categories or values for each variable, the variables can be said to have interval-level characteristics and can be entered into a single regression equation such as the following:

$$Y = a + b_1D1 + b_2D2 + b_3D3 + b_4D4 + b_5D5 \quad (9)$$

where  $Y$  is the score on the dependent variable,  $a$  is the constant or  $Y$ -intercept,  $b_1, b_2, b_3, b_4,$  and  $b_5$  are the regression coefficients representing the effects of each category of race on the dependent variable, and  $D1, D2, D3, D4,$  and  $D5$  are dummy variables representing separate categories of race. If a person is black, then his or her predicted  $Y$  score would be equal to  $a + b_2$ , since  $D2$  would have a value of 1 ( $b_2D2 = b_2 * 1 = b_2$ ) and  $D1, D3, D4,$  and  $D5$  would all be 0 (e.g.,  $b_1D1 = b_1 * 0 = 0$ ). The effect of race would then be the addition of each of the

Respondent's Race	Values on Dummy Variables				
	D1	D2	D3	D4	D5
White	1	0	0	0	0
Black	0	1	0	0	0
Latino	0	0	1	0	0
Indian	0	0	0	1	0
Asian	0	0	0	0	1
Other	0	0	0	0	0

dummy variable effects. Other dummy or interval-level variables could then be included in the analysis and their effects could be interpreted as “controlling for” the effects of race.

**Estimating analysis of variance models.** The use of dummy variables in regression makes it possible to estimate analysis of variance models as well. In the example above, the value of  $a$  is equal to the mean score on the dependent variable for those who had a score of “other” on the race variable (i.e., the omitted category). The mean scores for the other racial groups can then be calculated by adding the appropriate  $b$  value (regression coefficient) to the  $a$  value. For example, the mean for the Latino group would be  $a + b_3$ . In addition, the squared multiple correlation coefficient ( $R^2$ ) is equivalent to the measure of association used in analysis of variance ( $ETA^2$ ), and the  $F$  test for statistical significance is also equivalent to the one computed using conventional analysis of variance procedures. This general model can be further extended by adding additional terms into the prediction equation for other control variables measured on continuous scales. In effect, such an analysis is equivalent to an analysis of covariance.

### APPLICABILITY

In sociological studies, the researcher is rarely able to manipulate the stimulus (or independent variable) and tends to be more interested in behavior in natural settings rather than controlled experimental settings. As a result, randomization of preexisting differences through random assignment of subjects to experimental and control groups is not possible and physical control over more immediate outside influences on behavior cannot be attained. In sociological studies, “other things” are rarely equal and must be ruled out as possible alternative explanations for group differences through “statistical control.” This statistical control is best accomplished through correlational

techniques, although in certain circumstances the analysis of covariance can be used to statistically control for possible extraneous sources of influence.

Where analysis of variance and covariance are more appropriate in sociological studies is: (1) where the independent variable can be manipulated (e.g., field experiments investigating natural reactions to staged incidences, or studies of the effectiveness of different modes of intervention or teaching styles); (2) analyses of typologies or global variables that capture a set of unspecified, interrelated causes or stimuli (e.g., comparisons of industrialized vs. nonindustrialized countries, or differences between white collar vs. blue collar workers); or (3) where independent (or predictor) variables are used that have a limited number of discrete categories (e.g., race, gender, religion, country, etc.).

#### REFERENCES

- Blalock, Hubert M. 1972 *Social Statistics*, 2nd ed. New York: McGraw-Hill.
- Bornstedt, G. W., and D. Knoke 1995 *Statistics for Social Data Analysis*. Itasca, Ill.: F. E. Peacock Publishers, Inc.
- McClendon, M. J. 1994 *Multiple Regression and Causal Analysis*. Itasca, Ill.: F. E. Peacock Publishers, Inc.
- Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent 1975 *SPSS: Statistical Package for the Social Sciences*, 2nd ed. New York: McGraw-Hill.
- Scheffe, H. A. 1959 *The Analysis of Variance*. New York: Wiley.
- Simmons, Roberta G., and Dale A. Blyth 1987 *Moving Into Adolescence: The Impact of Pubertal Change and School Context*. New York: Aldine de Gruyter.
- Tabachnick, Barbara G., and Linda S. Fidell 1996 *Using Multivariate Statistics*, 3rd ed. New York: Harper-Collins.

RICHARD BULCROFT

## ANDROGYNY

See Femininity/Masculinity.

## ANOMIE

The concept *anomie* was used by early sociologists to describe changes in society produced by the Industrial Revolution. The demise of traditional

communities and the disruption of norms, values, and a familiar way of life were major concerns of nineteenth-century philosophers and sociologists. For sociologists, anomie is most frequently associated with Emile Durkheim, although others used it differently even during his lifetime (Wolff 1988).

Durkheim ([1893] 1956) used the French word *anomie*, meaning “without norms,” to describe the disruption that societies experienced in the shift from agrarian, village economies to those based on industry. Anomie is used to describe a state of society, referring to characteristics of the social system, not of individuals, although individuals were affected by this force. Increasingly, this term has taken on a more social psychological meaning. This is not to say that it no longer has uses consistent with the initial definition, but its meaning has been broadened considerably, at times consistent with Durkheim’s usage, at times at substantial variance with it.

There are, no doubt, sociologists who cringe at any expanded usage of this and other concepts, but the fact of the matter is that we have no more control over its usage than Thomas Kuhn (1970) has over abominable uses of the concept “paradigm,” or than computer engineers have over those who say “interface” when they mean “meet with.” Although we cannot completely stop the misappropriation of such terms as anomie we can be careful that sociological extensions of anomie are logically derived from their early uses.

### DURKHEIM ON ANOMIE

According to Durkheim, village life based on agriculture had consistent, well-established norms that governed the day-to-day lives of individuals. Norms prescribed patterns of behavior, obligation, and expectations. Durkheim called this pattern of social life mechanical solidarity. Communities characterized by “mechanical solidarity” were self-contained units in which the family and the village provided for all of the needs of their members. With the emergence of industrial capitalism and the beginnings of population shifts from the hinterland to cities, mechanical solidarity could not successfully structure social life. Durkheim believed that a new, “organic solidarity” based on a division of labor would emerge, with a regulating normative structure that would be as functional as mechanical solidarity. The emergence of organic



solidarity would take time, however. The transitional period, characterized by normative disorganization, Durkheim described as anomic. By this he did not mean to imply literal normlessness but, rather, a state of *relative* normative disorder (Coser 1977). Compared with communities characterized by mechanical solidarity, developing larger towns and cities would have a less regulated, less structured, less ordered pattern of social life.

Release from the restraining influence of norms was not a liberating circumstance, according to Durkheim. In this state, without adequate normative direction, people did not know what to expect or how to behave. Many of the social problems that Durkheim witnessed in rapidly changing industrializing Europe, he blamed on inadequate normative regulation. In his classic *Suicide*, Durkheim ([1897] 1951) identifies “anomic suicide” as occurring when the values and norms of the group cease to have meaning or serve as anchors for the individual, leading to feelings of isolation, confusion, and personal disorganization.

#### CONTEMPORARY USES OF “ANOMIE”

Anomie continues to be used as defined by Durkheim, but it has also been extended during the twentieth century. In addition to extensions similar to past uses of this concept, social psychological conceptions of anomie have become widespread. Robert Merton’s use of “anomie” is very similar to that described by Durkheim. His application (1949) has been the core theoretical statement in one of the twentieth century’s major criminological traditions. “Anomia” is a social psychological derivative used to represent a state of disaffection or disconnectedness.

**Merton on Anomie.** Merton (1949) used the concept anomie to describe how social structure produced individual deviance. According to Merton, when there existed within a society a disjuncture between the legitimate goals that members of a society were aspiring to and the legitimate means of achieving these goals, then that society was in a state of anomie. For both Durkheim and Merton, frustrated aspirations were an important cause of norm violations, or deviance. They differed in what they saw as the source of aspirations. For Durkheim, it was human nature to have limitless desires, growing from a natural “wellspring” within. Merton argued that desires did not come from

within us, but were advanced by a widely held conception of what constitutes “the good life.”

Durkheim believed that when a society was characterized by anomie, there were inadequate normative constraints on the desires and expectations of people. Peasants could come to believe, even expect, that they could rise to live like the aristocracy, or become captains of newly developing industry. Part of mechanical solidarity was the norms that constrained these expectations, that ordered the intercourse between social classes, that checked the natural wellsprings of desires and encouraged peasants to be happy with their lot in life. Without these checks, desires exceeded reasonable hope of attainment, producing frustration and potentially deviance.

Merton’s conception of anomie placed the society itself in the position of determining both the legitimate goals that people should aspire to and the legitimate means of pursuing these goals. While this goal has often been expressed by researchers as wealth attainment, Merton (1997) believed that wealth attainment was only one example of many societal goals. Unfortunately, society frequently caused people to have grandiose expectations without providing all of its members with reasonable opportunities to pursue them legitimately. This circumstance, where the goals and the means were not both universally available to the members of a society, Merton called anomie.

When individuals were faced with anomie, they had to choose whether to forgo the socially advanced goals, their society’s shared vision of the good life, or to seek these objectives by means not defined as legitimate. Merton described five choices available to these individuals. With “conformity,” the individual uses the socially prescribed means to obtain the goals advanced by that society. “Innovation” is the choice to use illegitimate means to achieve the legitimate goals; much criminal behavior is an example of innovation. When a person goes through the motions of using the legitimate means, fully aware that the socially advanced goals are beyond his reach, this is “ritualism.” “Retreatism” is the choice neither to use the legitimate means nor to strive for the legitimate goals of a society. Finally, “rebellion” is rejecting the society’s means and goals and replacing them with ones defined by the individual as superior.

A common mistake is interpreting Merton as arguing that an individual chooses to live his or her life as a conformist, or innovator, or retreatist. To the contrary, Merton's position is that we all are constantly making choices when faced with behavioral alternatives. At one point during the day one might choose to act as a conformist, but later, when confronted with another choice, one may choose to innovate. For example, a person who engages in robbery, innovation, is not always an innovator; he or she may also have a job, which indicates conformity. While one of these choices may predominate with some people, they should be seen as alternatives that people choose from in deciding how to act in a particular instance, not identities that they assume.

In applying Merton's perspective to Western nations, sociologists have argued that most of these societies are characterized by some degree of anomie, which manifests itself as a lack of equal opportunity. The extent of anomie, the degree of disjuncture between goals and means in a society, can be used to predict the level of crime and deviance that society will experience. The high crime rates of the United States can be linked to great inequalities in income, education, and job opportunities (Loftin and Hill 1974). To explain individual propensity to deviate from norms, one must consider the extent to which individuals have accepted the society's conception of "the good life," and the legitimate means individuals can use to attain it (Cloward and Ohlin 1960). As an explanation of crime, this theory has given way to different approaches, but anomie has been absorbed into larger perspectives to explain the relationship between poverty and crime (Messner 1983).

**New Approaches to Anomie.** Anomie saw a theoretical resurgence in the late 1980s and 1990s (Agnew and Passas 1997), especially in criminological research. This resurgence first occurred with Agnew's (1985) *general strain theory* and later, with macro-variations such as *institutional anomie* (Messner and Rosenfeld 1994).

**Strain Theory.** While many have been critical of anomie and strain theories of the past (Hirschi 1969; Kornhauser 1978), Agnew (1985, 1992) argues that research in the areas of medical sociology, social psychology, and psychology can help create new directions for this theory. Agnew (1992)

has proposed a micro-level theory; that "Adolescents are pressured into delinquency by the negative affective states—most notably anger and related emotions—that often result from negative relationships" (p. 48). His extension of traditional strain theories focuses on more than one form of strain or anomie. Agnew (1992) suggests there are three major types of strain that can be experienced by individuals: strain (1) "as the failure to achieve positively valued goals," (2) "as the removal of positively valued stimuli," and (3) "as the presentation of negative stimuli." This extension of anomie or strain theories allows our understanding of the creation of anomie to move even farther away from that first misconception that it *must* be connected to wealth attainment.

**Institutional Anomie Theory.** Most research on anomie has been at the micro-level (Agnew and Passas 1997). For example, variations such as Agnew's general strain theory have ignored the theoretical implications at the macro-level (Agnew and Passas 1997; see also Bernard 1987; Messner 1988; Messner and Rosenfeld 1994; Rosenfeld 1989). Institutional anomie theory posits that in order to understand any social phenomena we must understand the basics of social organization. These basics are culture and social structure and are best understood by their linking mechanism, social institutions (Messner and Rosenfeld 1994; Rosenfeld and Messner 1997). Rosenfeld and Messner (1994, 1997) suggest that social institutions are both interdependent and in conflict with one another, which leads to a constant, necessary balancing of institutional demands. According to Rosenfeld and Messner (1997) the economy is at the center of this balancing act. Institutional anomie theory helps explain the effect of the domination of the economy over other institutions by suggesting that "economic dominance stimulates the emergence of anomie at the cultural level, and . . . erodes the structural restraints against crime associated with the performance of institutional roles" (Rosenfeld and Messner 1997, p. 213). Institutional anomie theory which, up until this stage, has been used to explain trends in crime, could successfully be extended to other social phenomena.

**Social Psychological Conceptions of Anomia.** Items designed to measure individual feelings of anomia are now frequently included in surveys such as the General Social Survey, an annual national survey conducted by the National Opinion

Research Center (NORC) at the University of Chicago. Examples of these items illustrate the current uses of the concept, as in the following anomia items from the 1988 NORC survey (respondents were instructed to indicate the extent of their agreement with each statement): "Most public officials (people in public office) are not really interested in the problems of the average man," and "It's hardly fair to bring a child into the world with the way things look for the future" (NORC 1988, pp. 215–216). (Nearly 40 percent of the respondents to the second question agreed, and 68 percent agreed with the first.)

### IN SUMMARY

Anomie has been and will continue to be a mainstay concept in sociology. Papers discussing the meanings and uses of this concept continue to be written (see, for example, Adler and Laufer 1995; Bjarnason 1998; Deflem 1989; de Man and Labreche-Gauthier 1993; Hackett 1994; Hilbert 1989; Menard 1995; Passas and Agnew 1997; Wolff 1988). The basic meaning of the term anomie, though—both in its initial usage as a description of society and in its modern extensions—is well established and widely understood within the discipline. Students new to sociology should take care to understand that the definitions of the word may not be as broad for sociologists as for the general public. The utility of the concept for the study of society is best maintained by extending it in ways that are consistent with its original definition.

### REFERENCES

- Adler, Freda, and William S. Laufer 1995 *Advances in Criminological Theory*, vol. 6: *The Legacy of Anomie*. New Brunswick, N.J.: Transaction.
- Agnew, Robert 1992 "Foundations For a General Strain Theory of Crime and Delinquency." *Criminology* 30:47–87.
- 1985 "A Revised Strain Theory of Delinquency." *Social Forces* 64:151–167.
- , and Nikos Passas 1997 Introduction. In Nikos Passas and Robert Agnew, eds., *The Future of Anomie Theory*, Boston: Northeastern University Press.
- Bernard, Thomas J. 1987 "Testing Structural Strain Theories." *Journal of Research in Crime and Delinquency* 24:262–280.
- Bjarnason, Thoroddur 1998 "Parents, Religion and Perceived Social Coherence: A Durkheimian Framework of Adolescent Anomie." *Journal for the Scientific Study of Religion* 37:742–755.
- Cloward, Richard A., and Lloyd E. Ohlin 1960 *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press.
- Coser, Lewis A. 1977 *Masters of Sociological Thought: Ideas in Historical and Social Context*, 2nd ed. New York: Harcourt Brace Jovanovich.
- Deflem, Mathiew 1989 "From Anomie to Anomia and Anomic Depression: A Sociological Critique on the Use of Anomie in Psychiatric Research." *Social Science and Medicine* 29:627–634.
- de Man, A.F., and L. Labreche-Gauthier 1993 "Correlates of Anomie in French-Canadian Adolescents." *Journal of Social Psychology* 133:141–146.
- Durkheim, Emile (1897) 1951 *Suicide*. New York: Free Press.
- (1893) 1956 *The Division of Labor in Society*. New York: Free Press.
- Hackett, Edward J. 1994 "A Social Control Perspective on Scientific Misconduct." *Journal of Higher Education* 65:242–261.
- Hilbert, Richard A. 1989 "Durkheim and Merton on Anomie: An Unexplored Contrast and Its Derivatives." *Social Problems* 36:242–250.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- Kornhauser, Ruth 1978 *Social Sources of Delinquency*. Chicago: University of Chicago Press.
- Kuhn, Thomas 1970 *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- Loftin, Colin, and Robert H. Hill 1974 "Regional Subculture and Homicide." *American Sociological Review* 39:714–724.
- Menard, Scott 1995 "A Developmental Test of Mertonian Anomie Theory." *Journal of Research in Crime and Delinquency* 32:136–175.
- Merton, Robert K. 1997 Forward. In Nikos Passas and Robert Agnew, eds., *The Future of Anomie Theory*. Boston: Northeastern University Press.
- 1949 *Social Theory and Social Structure: Toward the Codification of Theory and Research*. New York: Free Press.
- Messner, Steven F. 1988 "Merton's 'Social Structure and Anomie': The Road Not Taken." *Deviant Behavior* 9:33–53.
- 1983 "Regional and Racial Effects on the Urban Homicide Rate: The Subculture of Violence Revisited." *American Journal of Sociology* 88:997–1007.

- , and Richard Rosenfeld 1994 *Crime and the American Dream*. Belmont, Calif.: Wadsworth.
- National Opinion Research Center 1988 *General Social Surveys, 1972–1988: Cumulative Codebook*. Chicago: National Opinion Research Center, University of Chicago.
- Passas, Nikos, and Robert Agnew 1997 *The Future of Anomie*. Boston: Northeastern University Press.
- Rosenfeld, Richard 1989 “Robert Merton’s Contribution to the Sociology of Deviance.” *Sociological Inquiry* 59:453–466.
- , and Steven F. Messner 1997 “Markets, Morality, and an Institutional-Anomie Theory of Crime.” In Nikos Passas and Robert Agnew, eds., *The Future of Anomie Theory*. Boston: Northeastern University Press.
- Wolff, Kurt H. 1988 “Anomie and the Sociology of Knowledge, in Durkheim and Today.” *Philosophy and Social Criticism* 14:53–67.

ROBERT CRUTCHFIELD  
KRISTIN A. BATES

## ANTI-SEMITISM

See Discrimination; Prejudice; Race.

## APARTHEID

See African Studies; Discrimination; Race; Segregation and Desegregation.

## APPLIED SOCIOLOGY

Applied sociology is sociology in use. It is policy-oriented, action-directed, and intends to assist people and groups to think reflectively about what it is they do, or how it is they can create more viable social forms capable of adapting to changing external and internal conditions. The roots of applied sociology in the United States go back to the publication in 1883 of Lester Ward’s *Dynamic Sociology: or Applied Social Science*, a text in which he laid the groundwork for distinguishing between an understanding of causal processes and how to intervene in them to foster social progress. Today applied sociology has blossomed in every arena of sociological endeavor (Olsen and Micklin 1981).

The nature of applied sociology can more easily be grasped by examining those characteristics that distinguish it from basic sociology. Different audiences are involved (Coleman 1972). Basic sociology is oriented toward those who have a concern for the advancement of sociological knowledge. The quality of such work (quantitative or qualitative,) is evaluated in accordance with agreed-upon standards of scientific merit. Applied sociology is oriented more toward those who are making decisions, developing or monitoring programs, or concerned about the accountability of those who are making decisions and developing programs. The quality of applied work is evaluated in accordance with a dual set of criteria: (1) how useful it is in informing decisions, revealing patterns, improving programs, and increasing accountability; and (2) whether its assumptions and methods are appropriately rigorous for the problems under investigation.

If we were to imagine a continuum between pure research and pure practice, applied sociology would occupy a space in the middle of this continuum. This space is enlarged along one boundary when practitioners and applied sociologists collaborate to explain patterns of behavior or develop causal models for predicting the likely impact of different courses of action. It is enlarged on the other boundary when applied and basic sociologists collaborate in the elaboration of abstract theory so as to make it more useful (Lazersfeld and Reitz 1975). There is always tension between the two, however. In part, the tension is attributable to the analytic distance that should characterize sociological analysis more generally (Lofland 1997). In part, it is attributable to the infusion of an ethical position in applied analysis, a posture which, if nothing else, is sensitive to the operation of power.

The boundaries of applied sociology may also be specified by enumerating the activities that play a central role in what it is that applied sociologists do. Freeman and Rossi (1984) have suggested three activities: (1) mapping and social indicator research, (2) modeling social phenomena, and (3) evaluating purposive action. To this could be added at least one more activity: (4) conceptualizing, studying, and facilitating the adaptability of alternative social forms. Examining these activities also permits considering some of the presumed trade-offs that are commonly thought to distinguish basic from applied sociology.

**Mapping and social indicator research.** Such studies are primarily descriptive, and designed to provide estimates of the incidence and prevalence of phenomena that are social in nature. There may be interest in how these phenomena are distributed in different social categories (e.g., by ethnic group affiliation, lifestyles, or social classes) or are changing over time. For example, corporations may wish to know how consumption patterns for various goods are changing over time for different groups to facilitate the development of marketing strategies. Federal and state agencies may wish to understand how the incidence and prevalence of diseases for different social groups is changing over time to develop more effective prevention and treatment strategies. Neighborhood groups may wish to know how citizen complaints regarding the police, high school graduation, or cervical cancer rates are distributed in relation to income characteristics of neighborhoods.

It is often assumed that applied sociology is less rigorous than basic sociology. Freeman and Rossi (1984) suggest that it is more appropriate to argue that the norms governing the conduct of basic sociological research are universally rigorous, while the norms governing the conduct of applied sociological research, of which mapping and social indicator research are but two examples, have a sliding scale of rigor. For critical decisions, complex phenomena that are either difficult to measure or disentangle, or where precise projections are needed, sophisticated quantitative or qualitative measures, or both, may be required and sophisticated analytic techniques may be needed, or may need to be developed. But as time and budget constraints increase, and the need for precision decreases, "quick and dirty" measures might be more appropriate. The level of rigor in applied sociology, in short, is driven by the needs of the client and the situation as well as the nature of the problem under investigation.

**Modeling social phenomena.** The modeling of social phenomena is an activity common to both basic and applied sociology. Sociologists of both persuasions might be interested in modeling the paths by which adolescents develop adaptive or maladaptive coping strategies, or the mechanisms by which social order is maintained in illicit drug networks. The applied sociologist would need to go beyond the development of these causal models. For the adolescents, applied sociologists would

need to understand how various interventions might increase the development and maintenance of adaptive strategies. In the case of drug networks, they might need to understand the relative effectiveness of different crime-control strategies in reducing the capacity of drug networks to protect themselves.

Just as with mapping and social indicator research, while there are norms supporting rigor to which basic researchers should adhere in the development of causal models, there is usually a sliding scale of rigor in applied research. There is not a necessary trade-off between doing applied work and levels of rigor. It usually happens, however, that applied problems are relatively more complex and tap into concepts that are less easy to quantify. In trying to capture more of the complexity, precision in the specification of concepts and elegance of form of the overall model may be traded off against an understanding of dynamics that is sufficient to inform decisions.

For both basic and applied work it is imperative that the mechanisms by which controllable and uncontrollable concepts have an impact on the phenomenon of interest are properly specified. Specification decisions must be made with respect to three aspects of the dynamics of situations (Britt 1997). First, the nature of what concepts are (and what they are not) must be specified with regard to the contextually specific indicators that give them meaning. Second, the concepts that are considered sufficiently important to be included in a model (as well as those that are deemed sufficiently unimportant to be left out) must be specified. Finally, the nature of relationships that exist (and do not exist) among the included concepts must be specified. Since these specification decisions interpenetrate one another (i.e., have implications for one another) over time, it may be prudent to think of models as vehicles for summarizing what we think we know about the dynamics of situations (Britt 1997).

The clients of applied researchers, however, are not interested in how elegant models are, but in how well the implications of these models assist them in reducing the uncertainty associated with decisions that must be made. As time and budget constraints increase, or researchers become more confident in their ability to understand which controllable concepts are having the biggest (or

most unanticipated) impact, less formal techniques may be used to develop and evaluate the models under consideration. Elegance must be balanced against usability. Complexity must be balanced against communicability. Theoretical sophistication must be balanced against the capacity of a model to interpret the lives of individuals and groups living through situations.

**Evaluating purposive action.** Evaluation research is an applied activity in which theories and methods of the social sciences are used to ascertain the extent to which programs are being implemented, services are being delivered, goals are being accomplished, and these efforts are being conducted in a cost-effective manner. These may be relatively small-scale efforts with finite and specific research questions. A manufacturing company may be interested in evaluating the impact of a new marketing program. A drug rehabilitation center may be interested in evaluating the cost-effectiveness of a new treatment modality. A movement organization may be interested in the situational effectiveness of particular strategies for fostering policies that are conducive to the reduction in infant-mortality differentials or the production of higher graduation rates in high-risk areas, or to a more equitable allocation of tax revenues (Maines and McCallion 2000).

These programs may or may not be of national importance, may or may not have large sums of money contingent on the outcomes, and may or may not require an understanding of anything but gross effects. It may be necessary to perform such analyses with limited personnel, time, and money. Under such circumstances, relatively unsophisticated methods are going to be used to conduct the evaluations and reanalysis will not be likely.

On the other hand, programs may involve the lives of many people, and deal with critical and complex social issues. The Coleman report on the equality of educational opportunity was presumably intended to establish once and for all that gross differences in school facilities did exist for black and white children in the United States (Coleman et al. 1966). The report, carried out by a team of sophisticated social scientists in a relatively short time, unleashed a storm of reanalyses and critiques (e.g., Mosteller and Moynihan 1972). These reanalyses attempted to apply the most

sophisticated theoretical and methodological weapons in the sociological arsenal to the task of evaluating the implications of the Coleman report. Similarly, econometric analyses initially conducted to evaluate the impact of capital punishment on the homicide rate spawned very painstaking and sophisticated applied research (e.g., Bowers and Pierce 1975) in an attempt to evaluate the robustness of the conclusions. In these latter two cases, a high level of rigor by any standard was maintained.

**Conceptualizing, studying, and facilitating the adaptability of alternative social forms.** A legitimate test of applied sociology is whether it can be used as a basis for designing and implementing better social institutions (Street and Weinstein 1975). An element of critical theory is involved here that complicates the distinction between basic and applied sociology, for it challenges applied sociologists and their clients to imagine: (1) alternative social forms that might be more adaptive in the face of changing social, environmental, and technological trends; and (2) alternative environments that must be sought if equitable social forms are to exist. At the level of families, this may mean asking what new role relationships could create more adaptive family structures. Alternatively, it might mean asking in what normative environments egalitarian family roles might exist. Or perhaps, asking what the impact of pressure for reproductive rights might be on the influence of women in families and communities. At the level of work groups faced with changing technologies and dynamic environments, it is appropriate to ask whether flatter organizational structures and more autonomous work groups might better serve both organizational goals and those of its members (e.g., Myers 1985). At the community level, exploring the viability of alternative interorganizational relationships, or how communities respond to the threat of drug dealing are among a host of legitimate questions. Whatever the specific focus of such questions and the role of applied sociologists in working with clients to answer them, questions of power and ethics should never be far away.

In applied sociology, problems drive the development of both theory and method. When problems and their dynamics cannot be explained by existing theories, new assumptions are added (Lazerefeld and Reitz 1975), new ways of thinking about concepts like adaptability are developed

(e.g., Britt 1989), or more fundamental theoretical shifts take place in a manner described by Kuhn (1961). When problems cannot be studied using existing methodological and statistical techniques, new techniques are developed. For example, the computer was developed under a contract from the Census Bureau so that the 1950 census could be conducted, and advances in area sampling theory were stimulated by the Department of Commerce needing to get better unemployment estimates (Rossi 1986). Ragin (1987) developed qualitative comparative analysis to permit the rigorous analysis of relatively rare events such as revolutions. Yet the applied implications of being able to study the alternative combinations of conditions that might give rise to particular outcomes has immense applied value (Britt 1998).

The continuing pressure on applied sociology to adapt to the needs of clients has had two important second-order effects beyond the development of theory and method. The nature of graduate training for applied sociologists is changing by virtue of the wider repertoire of skills needed by applied sociologists, and the dilemmas faced by sociologists vis-à-vis their clients are being confronted, and norms regarding the appropriateness of various courses of action are being developed.

A universal component of graduate applied education is the internship. Learning by doing, and experiencing the array of problems associated with designing and conducting research under time and budget constraints, while still having supportive ties with the academic program, are very important. The range and depth of coursework in qualitative and quantitative methods and statistics is increasing in applied programs to prepare students for the prospect of needing to employ techniques as varied as structural equations, focus groups, archival analysis, and participant observation in order to deal with the complexity of the problems requiring analysis.

There have been corresponding changes in the area of theory, with more emphasis given to moving back and forth from theory to applied problem. And there have been increases in courses designed to train students in the other skills required for successful applied work: networking, problem decomposition, and dealing with client-sociologist dilemmas.

## REFERENCES

- Bowers, W., and G. Pierce 1975 "The Illusion of Deterrence in Isaac Erlich's Research on Capital Punishment." *Yale Law Journal* 85: 187-208.
- Coleman, J. G. 1972 *Policy Research in the Social Sciences*. Morristown, N.J.: General Learning Press.
- Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfield, R. L. York 1966 *Equality of Educational Opportunity*, 2 vols. Washington, D.C.: U.S. Government Printing Office.
- Freeman, H. E., and P. H. Rossi 1984 "Furthering the Applied Side of Sociology." *American Sociological Review* 49:571-580.
- Lazersfeld, P. F., and J. G. Reitz 1975 *An Introduction to Applied Sociology*. New York: Elsevier.
- Mosteller, F., and D. P. Moynihan 1972 *On Equality of Educational Opportunity*. New York: Random House.
- Myers, J. B. "Making Organizations Adaptive to Change: Eliminating Bureaucracy at Shenandoah Life." *National Productivity Review*, (Spring) 131-138.
- Olsen, M.E., and M. Micklin 1981 *Handbook of Applied Sociology*. New York: Praeger.
- Rossi, P.H. 1986 "How Applied Sociology can Save Basic Sociology." *Journal of Applied Sociology* 3:1-6.
- Street, D., and E. Weinstein 1975 "Prologue: Problems and Prospects of Applied Sociology." *The American Sociologist* 10:65-72.
- Maives, D. R. and M. J. McCallion (2000) Urban Inequality and the Possibilities of Church-based Intervention. *Studies in Symbolic Interaction* 23, (forthcoming).

DAVID BRITT

## ART AND SOCIETY

There is no consensus as to what art is nor, until the 1970s, had sociologists expended much energy on its study or on the development of a sociology of the arts. While in Europe art had longer been of interest to sociologists than in the United States, even there it had not developed into an identifiable field with clear and internationally accepted parameters. As recently as 1968 the term *sociology of art* was not indexed in the *International Encyclopaedia of the Social Sciences*, which sought to sum and assess the thinking and accomplishments in the rapidly expanding social sciences of the post-World War II period. Yet by the end of the century the study of art had moved into the mainstream of sociological theory and was rapidly becoming a

avored subject for empirical investigation not only in the countries of Central and Western Europe but also in the United States.

Why should art, as a subject for sociological study, have been so neglected as to have virtually disappeared from mention in American textbooks for half a century after World War I? In large part this reflected the inherent tension between sociology and art, which, as noted by Pierre Bourdieu, make an “odd couple.” Artists, believing in the uniqueness of the original creator, resented the social scientist’s attempt to demystify their achievements by dissecting the role of the artist in society, by questioning to what extent artists are “born” rather than “made,” by conceptualizing artistic works as the products of collective rather than individual action, by anthropologically approaching art institutions, by studying the importance of networks in artistic success, and by investigating the economic correlates of artistic productivity. Many scholars in the humanities were also skeptical. For them the appeal of art is something of a mystery and best left that way; they could hardly relate to the attempts of social scientists who, in their quest for objectivity, sought to eliminate any evaluative component from their own research. This practice of disregarding one’s own personal preferences and tastes hardly seemed legitimate to aestheticians. Moreover, in pursuing a rigorous methodology, many sociologists chose to study only those problems that could yield readily to statistical analysis, and art did not seem to be one of those. They also preferred to focus on subjects that were important in the solution of social problems, and, in the United States, the arts were not generally regarded as high on this list.

Nonetheless, there has been—especially since the late 1960s—a slow but steady movement toward the development of a sociology of the arts. This is due, in part, to a narrowing of the intellectual gulf between the humanistic and sociological approaches. On the one hand, art historians have legitimated the study of art within its social context, and, on the other, mainstream sociology has become more hospitable to the use of other than purely “scientific” methodology. In part this progress resulted from the expanded contacts of American sociologists after World War II with their counterparts in other countries where art is regarded as a vital social institution and a public good. And just as art must be understood and

studied within its social context, so too the growing sociological interest in the arts reflects the growing importance of the arts within American society and the recognition of this importance by the government. Despite the concerted opposition of those who believe there is no role for government in funding the arts, at every level of government—federal, state, and local—arrangements for the support of grassroots arts have become institutionalized.

A small but dedicated number of scholars can be credited with sparking this postwar advancement of theory and research in the sociology of culture and, more specifically, in the sociology of the arts. The latter term, though not unknown before then, began to surface with some frequency in the 1950s; thus, in 1954 a session on the sociology of art was listed in the program of the annual meeting of the American Sociological Association. In 1957 a symposium on the arts and human behavior at the Center for Advanced Study in the Behavioral Sciences served as a catalyst for the production of a book, based partially on papers presented there. In its preface, the editor, Robert N. Wilson, concluded that a sociology of art, though in the early stages of its development, was not yet ripe for formalization. Nonetheless, Wilson’s book included a number of articles based on empirical research that attracted attention. In one, Cynthia White, an art historian, and Harrison White, a Harvard sociologist, reported on their investigation into institutional change in the French painting world and how this affected artistic careers. Later expanded and published in book form as *Canvases and Careers* (1965), this research provides a working example of how a changing art form might best be studied and understood within its historical and social context.

Perhaps the most important step toward the development of the field in these postwar years came with the publication of a collection of readings edited by Milton Albrecht, James Barnett, and Mason Griff (1970). Clearly titled as to subject matter—*The Sociology of Art and Literature*—it was intended to serve a classroom purpose but also to advance an institutional approach to its study. In one article, originally published in 1968, Albrecht oriented the reader to art as an institution, using *art* as a collective term for a wide variety of aesthetic products, including literature, the visual arts, and music. In another (“The Sociology of Art”),



Barnett reprinted his state-of-the-field synthesis as it stood in 1959, and, in yet another, Griff published a seminal article on the recruitment and socialization of artists, drawing in some part on his earlier empirical studies of art students in Chicago. Though here, too, the editors spoke of the sociology of art as being still in its infancy, they helped it to take its first steps by including in their reader a large number of empirically grounded articles—by scholars in the humanities as well as the social sciences. Divided into six subjects—forms and styles, artists, distribution and reward systems, tastemakers and publics, methodology, history and theory—it served for many years as an exemplary resource both for those attempting to set up courses on the arts and society and those embarking on research.

Beginning in the 1970s the sociology of art moved toward formalization and started to come into its own. Speeding this development in the new age of television dominance was a growing sociological interest in the mass media, in visual communications, and in the popular arts. The debate about mass versus popular culture was revitalized by new fears about the effects of commercialization, but some scholars began to wonder about the terms in which the debate was being cast. The assumption that art forms could be categorized as “high” or “low” or, put another way, as “mass” or “elite”—an assumption that had fueled the critiques of Theodor Adorno and other members of the Frankfurt School—came into question as reputable researchers looked more closely at the empirical evidence (Gans 1974). Howard S. Becker’s conceptualization of art as collective action (1982) did not so much mute the debate as turn attention away from the circumstances surrounding the production of any particular work—that is, what kind of an artist produced work for what kind of audience under what system of rewards—toward the collective (cooperative) nature of the activity whereby works regarded as art are produced as well as to that collective process itself. As attention turned to the production of culture, the arts came to be widely regarded by sociologists as socially constructed entities whose symbolic meanings reside not in the objects themselves but change as circumstances change.

Recent American studies in the sociology of art have taken varied approaches, both as to subject matter and methodology. Some have focused

on genres that are considered marginal to established categories of fine art, including such “outsider art” as that produced by asylum inmates, “naive artists,” African primitives, and Australian aborigines. Others have researched the process whereby “outsider artists” may make the transition to being “insiders” while still others, extending their interests to the politics of art, have considered what happens to “insiders” (and the art they have produced) when shifts in the political culture recast them as “outsiders.” Sociologists have also extended their inquiries to the economics of art as they have considered the influence of funding and the structure of museums on the creation, production, preservation, and dissemination of art works; case studies of “arts management” abound.

Some inquiries involving genres marginal to established categories of fine art have adopted methodologically unusual approaches. These include Wendy Griswold’s studies of the social factors influencing the revival of Renaissance plays (1986); Robert Crane’s study of the transformation of art styles in post-World War II New York (1987); Liah Greenfeld’s study of taste, choice, and success in the Israeli art worlds (1989); Vera Zolberg’s studies of art patronage and new art forms (1990); and Gladys Lang and Kurt Lang’s study of the building and survival of artistic reputations (1990).

These and other empirical studies that have already appeared in print or are under way are helping to clarify what is meant by a sociology of art. While there still may be no consensus as to what art is—nor need there be—some consensus is shaping up as to the direction in which the field should be moving. Leading theoreticians—Vera Zolberg, Janet Wolff, Paul DiMaggio, Richard Peterson, and Anne Bowler among them—agree on the need to keep the art itself at the center of theoretical concern but continue to disagree on the proper methodological approach to that “centering.” Essentially, this pits the case for focusing on the institutions in which aesthetic objects are produced and received—an analytical approach—against one that emphasizes criticism and textual interpretation of the objects themselves. Zolberg has voiced the need to avoid the narrowness of both social science and aesthetic disciplines, accepting the premise that art should be contextualized in terms of time and place in a general sense as

well as in a specific sense, that is, in terms of institutional norms, professional training, reward, and patronage or other support. To approach art as a part of society's culture, Zolberg argues, is no more potentially reductionist than to treat it, as aestheticians do, as an activity that only restricted groups with special interests and knowledge can hope to comprehend. Literature in the field is growing at a rapid rate as intellectual barriers between humanistic and social science approaches to the study of art begin to crumble.

REFERENCES

Adorno, Theodor (1962) 1976 *Introduction to the Sociology of Music*. New York: Seabury Press.

Albrecht, M. C., J. H. Barnett, and M. Griff (eds.) 1970 *The Sociology of Art and Literature: A Reader*. New York: Praeger.

Alexander, Victoria D. 1996 *Museums and Money: The Impact of Funding on Exhibitions, Scholarship, and Management*. Bloomington: Indiana University Press.

Balfe, Judith Huggins (ed.) 1993 *Paying the Piper: Causes and Consequences of Art Patronage*. Urbana: University of Illinois Press.

Becker, Howard S. 1982 *Art Worlds*. Berkeley: University of California Press.

Bourdieu, Pierre (1979) 1984 *Distinction: A Social Critique of the Judgment of Taste*. Cambridge, Mass.: Harvard University Press.

Bowler, Anne 1994 "Methodological Dilemmas in the Sociology of Art." In Diana Crane, ed., *Sociology of Culture*. Oxford: Blackwell.

Crane, Diana 1987 *The Transformation of the AvantGarde: The New York Art World 1940-85*. Chicago: University of Chicago Press.

Foster, Arnold W., and Judith R. Blau (eds.) 1989 *Art and Society: Readings in the Sociology of the Arts*. Albany: State University of New York Press.

Gans, Herbert J. 1974 *Popular Culture and High Culture: An Analysis and Evaluation of Taste*. New York: Basic Books.

Greenfeld, Liah 1989 *Different Worlds: A Sociological Study of Taste, Choice and Success in Art*. Cambridge, England: Cambridge University Press.

Griswold, Wendy 1986 *Renaissance Revivals: City Comedy and Revenge Tragedy in the London Theater, 1576-1980*. Chicago: University of Chicago Press.

Lang, Gladys Engel, and Kurt Lang 1990 *Etched in Memory: The Building and Survival of Artistic Reputation*. Chapel Hill: University of North Carolina Press.

— 1996 "Banishing the Past: The German Avant-Garde." *Qualitative Sociology* 19:323-343.

Peterson, Richard A. (ed.) 1976 *The Production of Culture*. Beverly Hills, Calif.: Sage Publications.

Rueschemeyer, Marilyn 1993 "State Patronage in the German Democratic Republic." In Judith Huggins Balfe, ed., *Paying the Piper*. Urbana: University of Illinois Press.

White, Harrison C., and Cynthia A. White 1965 *Canvases and Careers: Institutional Change in the French Painting World*. New York: Wiley.

Wilson, Robert N. (ed.) 1964 *The Arts in Society*. Englewood Cliffs, N.J.: Prentice-Hall.

Wolff, Janet 1992 "Excess and Inhibition: Interdisciplinarity in the Study of Art." In Lawrence Grossberg, Cary Wilson, and Paula A. Treichler, eds., *Cultural Studies*. New York: Routledge Press.

Zolberg, Vera L. 1990 *Constructing a Sociology of the Arts*. New York: Cambridge University Press.

—, and Charbo, Joni M. (eds.) 1997 *Outsider Art*. New York: Cambridge University Press.

GLADYS ENGEL LANG

ASIAN-AMERICAN STUDIES

The term *Asian American* is used in the United States by federal, state, and local governments to designate people of Asian descent, including Pacific Islanders (residents from the Pacific islands that are under U. S. jurisdiction, such as Guam, American Samoa, and the Marshall Islands). Although historically relevant and geographically appropriate, inclusion of the Pacific islands in the generic term *Asian American* stemmed from administrative convenience for the federal government rather than from race or ethnic identifications.

Reflecting deep-seated prejudices against people of color, in 1917 the Congress of the United States created the Asiatic Barred Zone, which stretched from Japan in the east to India in the west. People from within the zone were banned from immigration. The geographic concept was incorporated into the Immigration Act of 1924 (Oriental Exclusion Act), a law that had a profound impact on the demographic structure of Asian-American communities as well as on U.S. foreign policy. Although it is generally assumed that the term *Asian American* has a racial basis,

particularly from the perspective of U.S. immigration history, the racial overtone is muted by the inclusion in the 1980 census of people from India in the "Asian and Pacific Islander" category; they had been classified as "white" prior to 1980.

### IMMIGRATION AND RESTRICTIONS OF ASIAN AMERICANS

Asian immigration can be divided into two periods: the old and the new. The old immigration period was marked by nonoverlapping waves of distinct Asian populations who came largely in response to the sociopolitical conditions in their homelands and to the shortage of unskilled labor experienced by special-interest groups in the United States. The new immigration was characterized by the simultaneous arrival of people from the Asia-Pacific Triangle, spurred principally by the 1965 legislative reforms in U.S. immigration policy, shortages of certain skilled and professional labor, the involvement of United States in Asia, and the sociopolitical situations in Asia in the context of the Cold War. In between these two waves, there was another wave of Asian immigrants, who came between the end of World War II and the mid 1960s, though the number was small and involved principally Filipinos and their families because of their services in the US military. This group came outside the fifty-person quota allowed for Filipinos as a result of the *Tydings-McDuffie Act of 1935* when the Republic of the Philippines was granted independence.

The year 1848 marked the beginning of Asian immigration to the United States when the coastal Chinese—mostly from Guangdong—responded to the California gold rush and failures in the rural economy of China. Within fewer than thirty-five years, the Chinese became the first group in U.S. history to be legally barred from becoming citizens because of race. The 1882 Anti-Chinese Exclusion Act was followed by an influx of immigrants from the southern prefectures of Japan during the last decade of the nineteenth century—until that flow ended abruptly with the so-called "Gentlemen's Agreement" of 1907–1908. Unlike the termination of Chinese immigration, and reflecting Japan's position as a world power, cessation of entry by Japanese was accomplished through a diplomatic compromise between the two governments rather than through an act of Congress. Without a

continuous flow of Japanese farm workers to ease the labor shortage on the Hawaiian plantations, contractors turned to the Philippine Islands—which had been a U.S. possession since 1898—for cheap labor. From 1906 to the independence of the Republic of the Philippines in 1946, over 125,000 (predominantly single) Filipino males, the majority of them from the Ilocos region, labored on Hawaiian sugar plantations.

The exclusions of Asians enacted into the National Origins Act of 1924 essentially remained in effect until 1965. By Act of Congress in 1943, however, 105 Chinese were permitted to immigrate annually, and in 1952, under the McCarran-Walter Act, a token one hundred persons from each Asian country were allowed entry. The symbolic opening of immigration doors to Asians was attributed to Walter Judd, a congressman from Minnesota who had spent many years in China as a medical missionary. The provision of a quota of one hundred persons seemed to be an important moral victory for those who wanted the elimination of the exclusion act, but it was in fact a restatement of the 1924 national origin quota basis for immigration.

The new stream of Asian immigrants to the United States reflected the 1965 legislative reform that allowed an equal number of persons (20,000) from each country outside the Western Hemisphere to immigrate. Furthermore, family unification and needed skills became the major admission criteria, replacing national origin. Besides China and the Philippines, Korea and the Indian sub-continent became, and continue to be, the major countries of origin of many newly arrived Asian immigrants. Refugees from Vietnam, Cambodia, and Laos began to enter the United States in 1975, and by 1990, peoples from the former Indochina (Vietnam, Cambodia and Laos) had become the third-largest Asian group, following Chinese and Filipinos. In contrast, Japan's immigration to the United States practically ceased from 1945 to 1965, when it resumed at a much lower rate than those reported for other Asian countries.

### SOCIAL CONSEQUENCES OF IMMIGRATION RESTRICTIONS

Several distinct demographic characteristics illustrate most graphically past restrictions and the

1965 revision of the immigration laws. Earlier immigrants from China and the Philippines were predominantly single males. As a result of racial prejudice that culminated in the passage of antimiscegenation laws directed primarily against people of color in many western and southwestern states, the majority of these earlier Asian immigrants remained unmarried. The lack of family life caused unattached immigrants to depend on one another, creating an apparent great solidarity among people of the same ethnic group. Many of the earlier studies of Chinese and Filipino communities depicted themes of social isolation and loneliness, which did not apply to the Japanese community. Paul Siu (1952) portrayed the extreme social isolation of Chinese laundrymen in Chicago in his doctoral dissertation, that was published at the time only as a paper in the *American Journal of Sociology*, with the title "The Sojourner." Although Siu's work was written under the direction of Robert E. Park and Ernest W. Burgess, it was not included in the Chicago School sociological series published by the University of Chicago Press that focused on urban and ethnic social structure of the time allegedly because Siu argued that the Chinese immigrants did not fit the Park-Burgess assimilation model because of the "race factor." Thus a major piece of Asian-American research, *The Chinese Laundryman: A Study in Social Isolation* (Siu 1987) remained unpublished until after the author's death in the mid-1980s.

The existence of single-gender communities of Filipinos and Chinese is clearly demonstrated in the U.S. censuses between 1860 and 1970. In 1860, the sex ratio for Chinese was 1,858 men for every 100 women. By 1890, following the peak of Chinese immigration during the previous decade, the ratio was 2,678 males for every 100 females—the highest recorded. Skewed sex ratios for the Chinese population later declined steadily as the result of legislative revisions in 1930 (46 U.S. Stat. 581) and 1931 (46 U.S. Stat. 1511) that enabled women from China to enter the United States.

A second factor that helped to balance the sex ratio in the Chinese community, particularly among the younger age cohorts, was the presence of an American-born generation. In 1900, U.S.-born persons constituted only 10 percent of the Chinese-American population. By 1970, the figure was 52 percent. Nevertheless, in the 1980 census, the sex ratio remained high for some age groups within

certain Asian-American subpopulations: among Filipinos, for example, the highest sex ratio was found in those sixty-five and older.

The demographic characteristics of Japanese Americans present yet another unusual feature. Under the "Gentlemen's Agreement" between Japan and the United States, Japanese women were allowed to land on the West Coast to join their men though the immigration of male laborers was curtailed. The majority of the women came as picture brides (Glenn 1986, pp. 31-35) within a narrow span of time. Thus, the years following 1910 were the decade of family building for the first (*issei*) generation of Japanese Americans. Since almost all *issei* were young and their brides were chosen from a cohort of marriageable applicants of about the same age, it was not surprising that *issei* began their families at about the same time after marriage. The historical accident of controlled migration of brides resulted in a uniform age cohort of the second-generation Japanese Americans (*nisei*). The relatively homogeneous age group of the *nisei* generation meant that their children, the third generation (known as *sansei*), were also of about the same age. The fourth generation followed the same pattern. The amazingly nonoverlapping age and generational cohorts among Japanese Americans is not known to have had parallels in other population groups.

Fourth, while Asian Americans in general continue to grow in number as a result of new immigration, the size of the Japanese American population increases primarily by the addition of new generations of U.S.-born babies. It is generally believed that the offspring of Japanese women who marry Caucasians have lost their Japanese identity, even though there are no estimates of the impact of intermarriages upon the shrinkage of the Japanese-American community. An educated guess would be that about two-thirds of Japanese Americans marry non-Japanese partners. Given the fact that Japanese immigrants had lower fertility rates than women in Japan during the period prior to and shortly after World War II, and that the number of new immigrants since the war has remained small, Japanese communities have larger percentages of older people than do other ethnic minority populations, including other Asian Americans. In short, Japanese Americans will be a much smaller ethnic minority in the future. The

plurality ranking for all Asian groups placed the Japanese at the top of the list in 1970; they dropped to the third place in 1980, are expected to place fourth in 1990, and to be ranked last by the year 2000.

One more demographic fact is worthy of note. Hawaii and the West Coast states continue to draw large numbers of new immigrants from Asia. Through a process known as "chain migration," relatives are likely to follow the immigrants soon after their arrival. This leads to sudden increases in population within the ethnic enclaves. The post-1965 pattern of population growth in many Chinatowns, for instance, is an example of the renewal and revitalization of ethnic communities—which prior to 1965 were experiencing a decline—as are the formation and expansion of Koreantowns, Filipinotowns, and Little Saigons. Moreover, the settlement of post-1965 immigrants from Asia is more dispersed than that of the earlier groups, owing to the fact that the need for professional and skilled manpower is widely distributed throughout the United States. The emergence of Thai, Malaysian, and Vietnamese communities in major metropolitan areas has added a new dimension to the ethnic composition of Asian Americans.

Two separate chains of immigration resulted from the new immigration legislation of 1965. One chain, largely found in Chinese and Filipino communities, is kin-selective in that the process of settlement follows the family ties of earlier immigrants. The other process is occupation-selective, based on skills and professional qualifications. These two processes created significantly different immigrant populations, with clearly discerned bimodal distributions of status characteristics. It is therefore common to find recent immigrants from Asia among the high-income groups as well as among the families living below the poverty level; some find their homes in the ethnic enclaves of central cities while others live in high-income suburban communities. Any attempt to describe Asian Americans by using average measures of social status characteristics, such as income, education, and occupation, can produce a distorted and misleading profile that fits no particular group, which can be misused by researchers and planners. A more useful description would be the use of standard errors to show the polarities or deviations of the immigrant group from the norm of the majority.

In short, the sociodemographic and socioeconomic characteristics of all Asian American communities since 1850 have been greatly influenced by federal immigration legislation. A clear grasp of the structure and change of Asian-American communities must begin with an understanding of the history of immigration legislation.

### ASIAN-AMERICAN RESEARCH

Asian-American research may be divided into six periods: (1) the early period before World War II, which was influenced by, and was a part of, the Chicago School of Sociology that focused on the emergence of a multiethnic urban America. (2) the World War II period, which saw a preponderance of Japanese-American studies that also centered around the University of Chicago; (3) the postwar era, with a strong emphasis on culture and personality studies related to Japanese and Japanese Americans; (4) a shift toward "ethnic" experience as a result of the civil rights movement; and the emergence of a new academic course of Asian-American studies and finally (5) the integration of social science theories and concepts of ethnic inequalities and the changing Asian-American vista in the 1980s and 1990s.

**The Early Period.** The pioneer sociological studies on the assimilation of immigrants in American urban communities may be attributed to the work of Robert E. Park. Although Park had done little empirical investigation, he had supervised a large number of graduate students and had formulated what was known as the theory of race cycle, which stressed the unidirectional process of competition, accommodation, and assimilation as the basis of race relations in urban America. Park later led a group of researchers to study Chinese and Japanese communities on the Pacific Coast. The results failed to prove the race-cycle theory. In defending his views, according to Lyman (1977, p.4), Park employed the Aristotelian doctrine of "obstacles," which suggests that among Chinese and Japanese the assimilation progress in the hypothesized direction was only delayed.

Early published sociological research on Asian Americans included the works of Bogardus (1928, 1930), who attempted to delineate degrees of prejudice against minorities through an operational concept of "social distance." Other topics were

chosen randomly such as "Oriental crime" in California (Beach 1932); school achievement of Japanese-American children (Bell 1935); and anti-Asian sentiments (Sandmeyer 1939; Ichihashi 1932). A noted pioneer community study of Japanese Americans conducted by Frank Miyamoto (1939) in Seattle in the late 1930s paved the way for the long and significant bibliography on Japanese-American studies that followed.

Perhaps the most significant and ambitious piece of work during the prewar era was the study of social isolation of Chinese immigrants, which took more than a decade to complete. The author, Paul Siu, working under a condition of extreme poverty for a decade, observed the life of Chinese laundrymen. The product of his research endeavors offers a classic text in the study of "unmeltable" immigrants, from which the concept of "sojourner" independently complemented the earlier work of G. Simmel (Siu 1952, 1987).

**World War II and Japanese-American Studies.** Large-scale systematic studies on Asian Americans began shortly after the Japanese attack on Pearl Harbor, when the United States declared war on Japan. The U.S. government stripped Japanese Americans of their property, relocated them, and housed them in internment camps for several years. Alexander Leighton, a psychiatrist, recruited nisei social science graduates to assist in his work in the camps, monitoring the morale and loyalty of internees; this perhaps was the pioneer work in assessing their group cohesion and structure. A few of Leighton's nisei assistants completed their doctoral studies after the war, maintaining a close and affectionate relationship with him. All had made their own contributions as social scientists and as Asian-American specialists. Leighton's work on the internment of these civilians (both citizens and non citizens) resulted in the publication of a classic text on loyalty (Leighton 1945).

Thomas and Nishimoto (1946), Thomas (1952), and Broom and Kitsuse (1955) also carefully documented the situation and the people of Japanese-American communities. The focus of these studies was the question of loyalty on the part of Japanese-American citizens and their offspring in spite of their brilliant wartime combat duties on behalf of their adopted country in Europe. It was the American home front conditions had sparked an area of

development in social science research that paralleled some of the classic work on the study of the Polish peasant (Thomas and Znaniecki 1946). As a result it increased the general knowledge base on Japanese Americans, including their families and communities, and their sacrifices and contributions to America's wartime efforts.

**Culture and Personality Studies in the Post-war Era.** During World War II, the U.S. government had reason and the opportunity to question the suitability of Asians as American citizens in regard to loyalty and civic responsibilities. It was also a time to test the myth that Asian immigrants could not assimilate into American society. Social scientists were intrigued by the way culture shapes the personality. Ruth Benedict's classic work on the Japanese personality and society (Benedict 1946) opened a new vista for research. A cohort of young scholars at the University of Chicago, which included Japanese-American graduate students, became known for their pioneer work in studying Japanese behavioral patterns. It had a profound effect on a generation of interested social scientists and resulted in the publication of many classic works on culture and personality (Caudill 1952; Jacobson and Rainwater 1953; Caudill and DeVos 1956; DeVos 1955; Kitano 1961, 1962, 1964; Caudill and Scarr 1961; Babcock and Caudill 1958; Meredith 1966; and Vogel 1961). Similar studies on other Asian-American groups are conspicuously absent.

**Ethnic Studies and the Civil Rights Movement.** In the 1960s, the civil rights movement, sparked by the death of Martin Luther King, Jr., contributed to the passage of an unprecedented immigration-legislation reform. At the time there existed among Asian Americans on the Pacific Coast, principally in California, a collective search for identity that shared many of the goals and rhetoric of the black movement. Research into the ethnic (Asian) U.S. communities had added two dimensions. The first was the need to raise ethnic consciousness as a part of the social movement. Personal testimonials of experiences as members of an oppressed minority provided insight into the psychology of ethnic minorities. The emphasis on the cathartic, as well as the cathetic quality in much of the writings of the civil-rights era reflected the mood of the period: that there is a need for alternative theories against the early assimilation

model in standard texts on racial and ethnic studies. Second, consistent with the radical theme, was the apparent influence of Marxian views on race and ethnic relations, which posited that African-American and other minorities are victims of oppression in a capitalist society.

Expectedly, the civil rights movement began a renewed interest in research on the experiences of the earliest Asian Americans. With time the titles ranged from well-documented academic publication to insightful popular readings for the lay public (Chen 1980; Daniels 1988; Choy 1979; Ichioka 1988; Miller 1969; Nee and de Bary 1973; Saxton 1971; Sung 1971; Takaki 1989; Wilson and Hosokawa 1980).

Asian-American studies was established in the 1970s as an academic discipline in a number of institutions of higher learning, particularly in California, at a time when there were only a few major publications as sources of information for undergraduates (e.g., Kitano 1961–1976; Lyman 1974; Petersen 1971). The birth of an academic specialty was marked by the conspicuous absence of available materials, particularly on Filipinos, Koreans, Vietnamese, and the peoples of the Indian subcontinent (Min 1989). In response to this void, the Asian-American Studies Center at the University of California at Los Angeles published two collections of papers (*Roots and Counterpoint*) and a quarterly journal, *The Amerasia*. On the Atlantic Coast, a group of U.S.-born professionals published an intellectual nonacademic monthly, *The Bridge*, for nearly a decade. In the 1970s and 1980s, these publications were recommended as collateral readings for college students interested in Asian-American studies. *Amerasia* has since become more academic in the 1980s and 1990s while *Bridge*, for the lack of funds, quietly folded after nearly one decade. In its place, two new academic publications appeared in the 1990s: the *Journal of Asian American Studies*, which is a U.S. East Coast complement (at Cornell University) to *Amerasia*; and the *Journal of Asian American Health*, that serves health research readers as well as the general public. In addition, funded research on health and mental health by the U. S. Department of Health and Human Services has greatly enhanced the research productivities on Asian Americans as well as cumulative bibliographies.

**The New Age of Asian-American Research: Emerging Theories and Concepts.** Stanley Lyman at the University of California at Berkeley, and S. Frank Miyamoto at the University of Washington, are generally acknowledged as pioneers in Asian-American research. Through his numerous papers and books, Lyman has maintained a theoretical relevance and has demonstrated an historical insight into the origin and growth of Asian-American communities, especially those of the Chinese and Japanese. As a social historian, he based his research, by and large, on archival documents (see Lyman 1970b). Miyamoto, on the other hand, belongs to a founding generation of Japanese-American researchers whose long-time association with the University of Washington and his training under Professor Herbert Blumer at the University of Chicago gave him a mix of symbolic interactionist bend and logical positivist approach to the establishment of sound theoretical propositions in the study of Japanese Americans.

In the 1980s and 1990s, a few well-trained sociologists began to emerge, many of them foreign born and foreign educated, with American postgraduate training—the “first-generation new immigrants”—scholars who arrived at a time when America had become sensitive to diverse cultures. The new breed of Asian-American researchers are increasingly more vocal, questioning traditional sociological theories and concepts based on studies of European immigrants in the early 1920s and 1930s. There is also a reflection of the postmodern and world community perspectives that have become influential intellectual trends of the time. These new studies of Asian-American communities have added much to a field that had been underserved by the social sciences. In addition, as Asian-American studies became a new academic discipline in line with the African-American and other ethnic minority and gender studies, the need grew for more social science information. The lack of usable and more accurate estimates on many attributes of Asian Americans based on adequate sample design and culturally appropriate survey instruments employed in federal surveys has frustrated many Asian-American researchers, in spite of the special publication of information on Asian Americans and Pacific Islanders in the 1980 and 1990 censuses. Asian-American researchers constructively point out how such statistics can

be improved to the advantage of both legislators and research academicians (see Yu and Liu 1992).

Much information on Asian Americans, particularly new immigrants, still depends on small sample backyard research by individual investigators. In the post-civil rights movement decades, studies on ethnic identities and race and ethnic relations continue to flourish, adding new Asian immigrants into the sample of observations and analyses. The newly arrived Korean businessmen and professionals gave birth to fresh topics in the research literature. Publications on Korean professionals and small businesses provided opportunities to support the old "middleman" theory that argued Asians succeeded in finding business opportunities by being middleman between white and black supplier and customer relations; (see Light and Bonacich 1988) and also advanced an argument that newly emerged Asian-American ethnic business communities can best be understood in the context of a global business community (Min 1987, and 1988; Yu, E.Y. 1983, Light and Sanchaz 1987).

Newer research also includes a collection of refugee studies that involved Vietnamese, Cambodians, and Laotian Hmongs. Publications on the Vietnamese community centered largely on the initial movement of refugees after the withdrawal of American troops from Vietnam on April 25, 1975. Liu and his associates made a survey of the first wave of refugees who were brought in directly from Vietnam to Camp Pendleton in California (Liu, Lamanna, and Murata 1979). He directed the reader's attention to the special status of refugees vis-à-vis immigrants, and the pathway to becoming refugees from previously nonrefugee status as a result of both political and military decisions. The distinction between refugees and immigrants was later further elaborated by Haines (1989a, 1989b). Gold (1992) compared Vietnamese with Jewish refugees' adaptive experience, and showed commonalities of refugee experience in general. Freeman (1984) on the other hand, collected life histories and reported refugees' own accounts of flight and adjustment. Taken together, studies on Vietnamese, Cambodians, and Laotians focused more on their cohort experience as refugees, rather than their transition to an emerging new ethnic community.

The smaller number of Cambodian and Laotian Hmong people, relative to the Vietnamese, is a major reason why they remained as a part of the underserved Asian-American population in the Asian-American literature. Much of the writings on the Hmong people were based on clinical interviews by practicing health providers and academic psychiatrists (Beiser 1987, 1988; Boehnlein 1987; Mollicca et al. 1987). There is also the reflection of a time when the concern was on wartime trauma, refugee experience, and cultural adjustment (see Chan 1994). These publications tend to leave an impression that Hmong people had the most difficult time in America, and are psychologically maladjusted. A popular book that portrayed the experience between American physicians and a Hmong child revealed how culture interfered in clinical judgment, no matter how well intentioned (Fadiman 1997)

There remained a scarcity of systematic studies, except for Agarwal's (1991) work on Indian immigrants. Agarwal's survey on immigrant Asian Indians in America was based on a rather small sample of professionals. Available studies on immigrant Americans from the Indian subcontinent, however, remained inadequate and woefully few.

The arrival of voluminous publications on new immigrants from Asia did not slow down studies of prior to 1965 immigrants. A cumulative bibliography began to focus on the aftermath of Japanese nisei resettlement (Ichioka 1989a; Miyamoto 1989; Warren 1989), and sensei scholars began to search for the Japanese identity of their parents' generation (Ichioka 1989b). A well-publicized book by Ronald Tadaki (1989), *Strangers from a Different Shore: A History of Asian Americans*, contained verbatim recorded life histories. The book was well received by the public but was poorly reviewed in the social science literature because it lacks a new conceptual framework for an old topic. However, it did arouse renewed interest in Asian Americans and their American experience, particularly at a time of race consciousness.

While the third-generation Japanese Americans were busy finding their identities, nisei continued to search for the meaning of "being American." In September of 1987, on the campus at the University of California at Berkeley, a group of scholars who belonged to the original Japanese American Evacuation and Resettlement Study



(JERS) at the Berkeley campus convened, and reminisced about lost valuable information of that period (see JERS 1989). Among those who attended include Ichioka, Miyamoto, L.R and J.H. Hirabayashi, Suzuki, Spencer, Kikuchi, Takagi, and Sakoda. Later researchers learned from discussion notes that there was a diary kept by Charles Kikuchi, known as the *Kikuchi Diary*, but it was destroyed. A manuscript entitled *American Betrayed* by Professor Morton Grodzin was suppressed from being published. Furthermore, another manuscript reportedly written by Violet de Christoforo described internees' refusal to leave camps as camps were closed. The book also exposed the harshness of expulsion. The existence of this manuscript, if true, shed additional light on similar experiences reported in Camp Pendelton as the Vietnamese camps were closed, when refugees were labeled as suffering from *campetitis* by the federal administrator in charge of closing all camps that housed refugees from Indochina (see Liu et al. 1979).

Clearly Asian-American research could not be separated from the general field of race and ethnic studies. In the 1980s and 1990s, researchers continued to be fascinated by Asians who had mutual ill feelings in their home countries because of neighboring dominations and the World War II experience, but appeared to be united behind a cause of advancing the status of Asian-Americans as a group. Whether there is a single identity of all Asian-Americans, or multiple subidentities of separate entities depends, of course, on issues. But the persistence of a multiple subidentities seem to be present not only among the first-generation immigrants, which is understandable, it also appears to be real among second- or third-generation ethnics. Fugita and O'Brien's (1991) volume on Japanese Americans is a good illustration; they seemed to have suggested that a single identity of Asian Americans is but an abstract rather than a reality. Among the Asian-American professionals, perhaps this issue is more complex and deep than is commonly expected (see Espirita 1992).

There also had begun some long-awaited new social science contributions to the literature on Filipino immigrants in the 1980s that were absent in the earlier generation. These studies included Filipino-American income levels (Cabezas, Shinagawa, and Kawaguchi 1986-1987), Filipino assimilation (Pido 1986), and Filipino health practices (Montepio 1989).

As more practicing professionals took part in studying Asian immigrants, there appeared in the 1970s and 1980s a bifurcation of writings on Asian Americans: those that catered to an academic audience and those that served as "voices" of the grassroots community (see Radhakrishnan 1996). It is understandable that, when more and more publications contained a whole spectrum of topics, the quality of writings would also vary widely. At the conceptual level, there remained in these writings a contrast between neo-Marxian and neo-classical economic approaches to status attainment, income, and employment opportunities for Asian Americans.

For example, the censuses of 1980 and 1990 confirmed the fact that U.S.-born Japanese and Chinese men came closest to parity with white men with respect to individual income. Japanese and Chinese women have significantly lower income than white women. The rest of those Asian men and women continued to have lower incomes than those of whites, Japanese, and Chinese. One explanation seemed to have suggested that Asians as a group valued education and scholastic achievement or individual effort, known as the *human capital theory* (Becker 1975). But within various Asian subgroups, the disturbing question is why some Asian immigrants attained higher income than others in the same ethnic group. Here both Becker's *human capital* and the *labor market segmentation* explanations are applicable. The Lahei suggests a two tier wage system prevail (see Cabezas, Shinagawa, and Kawaguchi 1986-1987; Bonacich 1975, 1988, 1992).

Perhaps a third explanation of the differential pace of status attainment came from some writings about the Filipino immigrant-labor market structure that related to the manner in which newcomers came to join their kin-relatives. The so-called "network recruitment" tended to form clusters in the same occupation area, resembling the earlier immigration patterns of Chinese in California (who came from four or five villages in Guangdong Province) and Japanese farm workers (who were recruited from southern prefectures of Japan).

Unlike blacks and Native Americans, Asian Americans varied significantly in terms of language proficiency, which had an enormous impact on income and the kinds of occupations in which they tended to cluster together. Barry Chiswick

(1974) has demonstrated that English proficiency is a major factor in income parity. P.G. Min (1986) showed that Filipino and Korean immigrants, an even match for income, engage in quite different occupations. Koreans engage in business but Filipinos do not. Min pointed out that the language barrier forces Koreans to be self-employed, no matter how hazardous their business turns out to be. Furthermore, it was Korean immigrants' linkage to industrial development in Korea that led to the development of some of the Korean communities in the United States. Filipino communities, on the other hand, did not develop similar enterprises to form a visible ethnic enclave in America. Instead, they depended more on ethnic professional societies and trade associations as basis of ethnic cohesion.

The same rationale was also applied in the case of "new" Chinatown studies. The earlier descriptive studies of Chinatown either over-romanticized the exotic quality of an ethnic enclave, or highlighted the worst side of humanity. To see Chinatown as a unique economic system is perhaps a new approach that is not without controversy. Zhou (1992), herself a new immigrant scholar from China, took the conceptual framework of her major advisor A. Portes, and wrote and published her dissertation on New York's Chinatown. She ably showed how a separate economic system had developed and embedded in urban America. Zhou's book on Chinatown is refreshingly different from several other volumes on the same topic published between the latter part of the 1970s to the early part of the 1990s that continued to portray patterns of conflict and cleavage (Wong 1977; Kuo 1977; Kwon 1987; Kinkead 1992).

The rise of Asian-American Studies programs, either as a part of multicultural or ethnic studies, or as an independent program, certainly has been instrumental in the increasing research and number of publications that deal with Asian-American communities. Given the interdisciplinary nature of the field, the multitudes of topics to be covered, and the unusual personal experiences of writers, the uneven quality of these publications is expected and sometimes unavoidable. Among studies on new communities, on income and status attainment, and on racial and ethnic relations, there appeared to be a genuine effort to formulate theoretically applicable concepts, and to use existing theories and models, which compared to fairly

good and important descriptive studies of an earlier era, both before and after World War II. A good example is the analysis of the Korean rotating credit associations in Los Angeles (Light, Kwuon, and Zhong 1990), in which the authors examined the functions of financing and savings groups based on informal trust in Korean business communities.

The most noticeable achievement of the Asian-American Studies programs in the 1970s through the 1990s is the abundant effort to compile important bibliographies, by searching for research materials that are normally scattered through scientific journals, unpublished files, doctoral and masters' dissertations, working papers, diaries, and conference proceedings. The federal government is credited with making bibliographies on Indochinese refugees, either through the U.S. Government Printing Office, or at the University of Minnesota Southeast Asian Refugee Studies Project (see references at the end of this section). Finally, the rise of Asian-American Studies programs increased efforts to establish libraries and documentation centers in a number of universities throughout the country.

## REFERENCES

- Agarwal, Priya 1991 *Passage from India: Post 1965 Indian Immigrants and Their Children*. Palos Verdes, Calif.: Yavati Publishers.
- Becker, G. S. 1975 *Human Capital*. Chicago: University of Chicago Press.
- Beiser, Morton 1987, "Changing Time Perspective and Mental Health among Southeast Asian Refugees." *Culture, Medicine & Psychiatry* 11(4):437-464.
- 1988 "Influence of Time, Ethnicity, and Attachment on Depression in Southeast Asian Refugees." *American Journal of Psychiatry* 145(1):46-51.
- Boehnlein, J. K. 1987 "Clinical Relevance of Grief and Mourning Among Cambodian Refugees." *Social Science and Medicine* 25:765-772.
- Cabezas, A. L., H. Shinagawa, and G. Kawaguchi 1986-1987 "New Inquiries into the Socioeconomic Status of Filipino Americans in California." *Amerasia* 12(1):21.
- Caces, Fe Immigrant Recruitment into the Labor Force: Social Networks Among Filipinos in Hawaii, AA 12:23-38.
- Chan, Sucheng (ed.) 1994 *Hmong Means Free: Life in Laos and America*. Philadelphia: Temple University Press.

- Chiswick, Barry 1974 *Income Inequality*.
- Chuong, Chung Hoang 1987 *The Vietnam Conflict and the Vietnamese in the U.S.: A Bibliography*. Berkeley Calif: University of California Press.
- ERIC Clearinghouse on Urban Education 1988 "Asian/Pacific American Education. *Brief Bibliographies*," March, 1988.
- Espirita, Yen Le 1992 *Asian American Panethnicity: Bridging Institutions and Identities*. Philadelphia: Temple University Press.
- Fadiman, Anne *The Spirit Catches You and You Fall Down*. New York: The Noonday Press.
- Freeman, James M. 1984 *Hearts of Sorrow: Vietnamese American Lives*. Stanford, Calif.: Stanford University Press.
- Fugita, S., and J. O'Brien *Japanese American Ethnicity: The Persistence of a Community*. Seattle: The University of Washington Press.
- Gold, Steven J. 1992 *Refugee Communities: A Comparative Study*. Newbury Park, Calif.: Sage Publications.
- Haines, D.W. 1989 "Southeast Asian Refugee Resettlement in the United States and Canada." *Amerasia Journal* 15(2):141-156.
- (ed) 1989b *Refugees as Immigrants: Cambodians, Laotians, and Vietnamese in America*. Totowa, N.J.: Rowman and Littlefield.
- Hammond, R. E., and G. L. Hendricks 1988 *Southeast Refugee Youth: An Annotated Bibliography*. Minneapolis University of Minnesota Center for Urban and Regional Affairs.
- Ichioka, Juji "Coming of Age: From Rural 30s to Cities." *Amerasia Journal*.
- "A Study in Dualism: Japanese American Courier: 1928-1942" *Amerasia Journal*.
- Japanese American Evacuation and Resettlement Study 1989 *Views from Within*. Los Angeles: UCLA Asian American Studies Center.
- Kim, Hyung-chan 1989 *Asian American Studies: An Annotated Bibliography and Research Guide*. Westport, Conn.: Greenwood Press.
- Kinkead, Gwen 1992 *Chinatown: A Portrait of a Closed Society*.
- Kuo, Chia-liong 1977 *Social and Political Change in New York's Chinatown*. New York: New York University Press.
- Kwong, Peter 1987 *New Chinatown*.
- Light, Ivan, and Edna Bonacich 1988 *Immigrant Entrepreneurs: Koreans in Los Angeles, 1965-1982*. Berkeley, Calif.: University of California Press.
- Light, I. M., J. Kwuon, and D. Zhong "Korean Rotating Credit Association in Los Angeles." *Amerasia* 16(2):35-54.
- Light, I., and A. A. Sanchez 1987 "Immigrant Entrepreneurs in 272 SMSAs." *Sociological Perspectives* 30(4):373-399.
- Liu, William T., Maryanne Lamanna, and Alice Murata 1979 *Transitions to Nowhere: Vietnamese Refugees in America*. Nashville, Tenn: The Charter House Publishers.
- Marston, J. 1987 *An Annotated Bibliography of Cambodia and Cambodian Refugees*. Minneapolis: University of Minnesota Center for Urban and Regional Affairs.
- Min, Pyong Gap 1989 "Textbooks for Asian American Studies: A Review Essay." *Amerasia* 15(1):259-266.
- 1988 *Ethnic Business Enterprise: Korean Small Business in Atlanta*. New York: Center for Migration Studies.
- 1986-1987 "Filipino and Korean Immigrants in Small Business: A Comparative Analysis." *Amerasia* 12:53-71.
- Miyamoto, A. S. 1988 "Problems of Interpersonal Style Among Nisei: From Self Restraint to Self Expression." *Amerasia* 13:29-45.
- Mollicca, R. F., G. Wyshak, and J. Lavelle 1987 "The Psychosocial Impact of War Trauma and Torture on Southeast Asian Refugees." *American Journal of Psychiatry* 144(12):1567-1572.
- Montepio, S. N. 1989 "Folk Medicine in the Filipino American Experience." *Amerasia* 13(1):151-162.
- 1987-1988 "Magical Medicine and the Filipino Healer: Floklore and Mythology Studies. 11, 12:36-46.
- National/Pacific Asian Resources Center 1988 *On Aging: Annotated Bibliography*. Seattle: National Pacific Asian Resources Center.
- Pido, Antonio J. A. 1986 Center for Migration Studies: New York. *The Filipinos in America: Macro/Micro Dimensions of Immigration and Integration*.
- Radhakrishnan R. 1996 *Diasporic Mediations: Between Home and Location*. St. Paul: University of Minnesota Press.
- Sawada, Mitziko "After the Camps: Seabrook Farms, N.J. and the Resettlement of Japanese Americans, 1944-1947." *Amerasia*.
- Shin, H. E., and K. S. Chang 1988 "Peripherization of Immigrant Professionals: The Case of Korean Physicians in the United States." *International Migration Review* 22:4 609-626.
- Smith, J. C. 1988 *The Hmong. An Annotated Bibliography, 1983-1987*. Minneapolis University of Minnesota Southeast Asian Refugee Studies Project.

- Southeast Asian Asian Refugee Studies Project 1983 *Bibliography of Hmongs*. Minneapolis: University of Minnesota Press.
- U.S. Government Printing Office 1989 *An Annotated Bibliography on Refugee Mental Health*. Washington, D.C.: U.S. Government Printing Office.
- Warren, W. H. 1989 "Maps: A spacial Approach to Japanese American Communities in Los Angeles." *Amerasia* 13:137-151.
- Wong, Bernard 1987 "The Role of Ethnicity in Enclave Enterprises: A Study of the Chinese Garment Factories in New York City." *Human Organization* 46(2):120-130.
- 1988 *Patronage, Brokerage, Entrepreneurship and the Chinese Community of New York*. New York: AMS Press.
- 1977 *Chinese American Community*.
- Wong, Paul, Meora Manvi, and Takeo Hirota Wong 1998 "Asiacentrism and Asian American Studies?" *Amerasia*. 21 (1, 2):137-147.
- Yochum, G., and V. Agarwal 1988 "Permanent Labor Certifications for Alien Professionals, 1975-1982". *International Migration Review* 22(2):365-381.
- 1983 "Korean Communities in America: Past, Present, and Future." *Amerasia* 10(2):23-51.
- Yu, Elena, and William T. Liu 1992 "U.S. National Health Data on Asian/Pacific Islanders: A Research Agenda for the 1990s." *American Journal of Public Health* 82:12.
- Zhou, M. 1992 *Chinatown: The Socioeconomic Potentials of An Urban Enclave*. Philadelphia: Temple University Press.

WILLIAM T. LIU  
ELENA S. H. YU

## ASSIMILATION

See Ethnicity.

## ATHEISM

See Religious Orientations.

## ATTITUDES

*Attitude* "is probably the most distinctive and indispensable concept in contemporary American social psychology" (Allport 1985, p. 35). Hundreds

of books and thousands of articles have been published on the topic. A review of this literature may be found in Eagly and Chaiken (1998). Despite this popularity, there is considerable disagreement about such basics as terminology. Several terms are frequently used as synonyms for attitude, including *opinion* and *belief*. Contemporary writers often distinguish attitudes from *cognitions*, which is broader and includes attitudes as well as perceptions of one's environment. Most analysts distinguish attitude from *value*, the latter referring to a person's ultimate concerns or preferred modes of conduct.

An attitude is a learned predisposition to respond to a particular object in a generally favorable or unfavorable way. Every attitude is about an object, and the object may be a person, product, idea, or event. Each attitude has three components: (1) a belief, (2) a favorable or unfavorable evaluation, and (3) a behavioral disposition. This definition is used by most contemporary writers. However, a small minority define attitude as consisting only of the positive or negative evaluation of an object.

A *stereotype* is one type of attitude. Originally, the term referred to a rigid and simplistic "picture in the head." In current usage, a stereotype is a belief about the characteristics of members of some specified social group. A stereotype may be positive (Asian Americans are good at math) or negative (women are bad at math). Most stereotypes are resistant to change.

Attitudes link the person to other individuals, groups, and social organizations and institutions. Each person has literally hundreds of attitudes, one for each significant object in the person's physical and social environment. By implication, the individual's attitudes should reflect his or her location in society. Thus, attitudes are influenced by gender, race, religion, education, and social class. Considerable research on the relationship between social position and attitudes has been carried out; this literature is reviewed by Kiecolt (1988).

## ATTITUDE FORMATION

Many attitudes are learned through direct experience with the object. Attitudes toward one's school, job, church, and the groups to which one belongs

are examples. Attitudes toward the significant persons in one's life are also learned in this way. More often, attitudes are learned through interactions with others. Socialization by parents, explicit teaching in educational and religious settings, and interactions with friends are important sources of attitudes. Research shows that children's attitudes toward a variety of objects, including gender roles and political issues, are similar to those held by their parents.

Another source of attitudes is the person's observations of the world. A topic of continuing interest is the impact of mass media on the attitudes (and behavior) of users. A thorough review of the literature on this topic (Roberts and Maccoby 1985) concludes that television viewing affects both children's and adolescents' attitudes about gender roles. Further, the viewing of programs intentionally designed to teach positive attitudes toward racial or ethnic minorities does increase children's acceptance of such persons. With regard to adults, evidence supports the "agenda setting" hypothesis; the amount and quality of coverage by the media (press, radio, and television) of an issue influences the public's perception of the importance of that issue. The effects of mass media exposure on aggression are discussed by Felson (1996) and Geen (1998).

Stereotypes are also learned. A stereotype may arise out of direct experience with a member of the stereotyped group, for example, a person who encounters a musically talented black person may create a stereotype by overgeneralizing, inferring that all African Americans are gifted musically. More often, however, stereotypes are learned from those with whom we interact such as parents. Other stereotypes may be acquired from books, television, or film. Research indicates that television programming portrays women, the elderly, and members of some ethnic minorities in negative ways and that these portrayals create (or reinforce) misperceptions and negative stereotypes in viewers (McGuire 1985).

Social institutions influence the attitudes one learns in several ways. Adults' ties to particular ethnic, religious, and other institutions influence the attitudes they teach their children. The instruction given in schools reflects the perspectives of the dominant political and economic institutions

in society. The amount and quality of media coverage of people and events reflects the interests of particular groups in society. Through these mechanisms, the individual's attitudes reflect the society, institutions, and groups of which she or he is a member.

Each attitude fulfills one or more of four functions for the individual. First, some attitudes serve an instrumental function: An individual develops favorable attitudes toward objects that aid or reward the individual and unfavorable attitudes toward objects that thwart or punish the individual. For example, a person who earns a large salary will have a positive attitude toward the job. Second, attitudes often serve a knowledge function. They provide the person with a meaningful and structured environment. Third, some attitudes express the individual's basic values and reinforce self-image. Whites' attitudes toward black Americans reflect the importance that whites place on the values of freedom and equality. Fourth, some attitudes protect the person from recognizing certain thoughts or feelings that threaten his or her self-image or adjustment.

Stereotypes also serve several functions. The act of classifying oneself as a member of a group (males, Republicans, whites) elicits the image of a contrasting group (females, Democrats, Latinos). Thus, stereotypes contribute to social identity. They also reduce the demands on the perceiver to process information about individual members of a stereotyped group; instead, one can rely on a stereotype. Finally, stereotypes may be used to justify the political and economic status quo.

### ATTITUDE MEASUREMENT

Because attitudes are mental states, they cannot be directly observed. Social scientists have developed a variety of methods for measuring attitudes, some direct and some indirect.

**Direct Methods.** These methods involve asking the person questions and recording the answers. Direct methods include various rating scales and several sophisticated scaling techniques.

The three most frequently used rating scales are single item, Likert scales, and the semantic differential. The *single-item scale* usually consists of

a direct positive or negative statement about the object, and the respondent indicates whether he or she agrees, disagrees, or is unsure. Such a measure is easy to score, but is not precise. A *Likert scale* typically involves several statements, and the respondent is asked to indicate the degree to which she or he agrees or disagrees with each. By analyzing differences in the pattern of responses across respondents, the investigator can order individuals from greatest agreement to greatest disagreement. Whereas Likert scales assess the denotative (literal) meaning of an object to a respondent, the *semantic differential* technique assesses the connotative (personal) meaning of the object. Here, an investigator presents the respondent with a series of bipolar adjective scales. Each of these is a scale whose poles are two adjectives having opposite meanings, for example, good–bad, exciting–boring. The respondent rates the attitude object, such as “my job,” on each scale. After the data are collected, the researcher can analyze them by various statistical techniques.

A variety of more sophisticated scaling techniques have been developed. These typically involve asking a series of questions about a class of objects, for example, occupations, crimes, or political figures, and then applying various statistical techniques to arrive at a summary measure. These include magnitude techniques (e.g., the Thurstone scale), interlocking techniques (e.g., the Guttman scale), proximity techniques (e.g., smallest space analysis), and the unfolding technique developed by Coombs. None of these has been widely used.

**Indirect Methods.** Direct methods assume that people will report honestly their attitudes toward the object of interest. But when questions deal with sensitive issues, such as attitudes toward members of minority groups or abortion, respondents may not report accurately. In an attempt to avoid such reactivity, investigators have developed various indirect methods.

Some methods involve keeping respondents unaware of what is being measured. The “lost letter” technique involves dropping letters in public areas and observing the behavior of the person who finds it. The researcher can measure attitudes toward abortion by addressing one-half of the letters to a prochoice group and the other half to a prolife group. If a greater percentage of letters to

the latter group are returned, it suggests people have prolife attitudes. Another indirect measure of attitude is pupil dilation, which increases when the person observes an object she or he likes and decreases when the object is disliked.

Some indirect measures involve deceiving respondents. A person may be asked to sort a large number of statements into groups, and the individual’s attitude may be inferred from the number or type of categories used. Similarly, a respondent may be asked to write statements characterizing other people’s beliefs on an issue, and the content and extremity of the respondent’s statements are used to measure his or her own attitude. A third technique is the “bogus pipeline.” This involves attaching the person with electrodes to a device and telling the person that the device measures his or her true attitudes. The respondent is told that some signal, such as a blinking light, pointer, or buzzer, will indicate the person’s real attitude, then the person is asked direct questions.

While these techniques may reduce inaccurate reporting, some of them yield measures whose meaning is not obvious or is of questionable validity. Does mailing a letter reflect one’s attitude toward the addressee or the desire to help? There is also evidence that measures based on these techniques are not reliable. Finally, some researchers believe it is unethical to use techniques that involve deception. Because of the importance of obtaining reliable and valid measures, research has been carried out on how to ask questions. This research is reviewed by Schuman and Presser (1996).

For a comprehensive discussion of attitude measurement techniques and issues, see Dawes and Smith (1985).

### ATTITUDE ORGANIZATION

An individual’s attitude toward some object usually is not an isolated psychological unit. It is embedded in a cognitive structure and linked with a variety of other attitudes. Several theories of attitude organization are based on the assumption that individuals prefer consistency among the elements of cognitive structure, that is, among attitudes and perceptions. Two of these are balance theory and dissonance theory.

**Balance Theory.** Balance theory, developed by Heider, is concerned with cognitive systems composed of two or three elements. The elements can be either persons or objects. Consider the statement “I will vote for Mary Sweeney; she supports parental leave legislation.” This system contains three elements—the speaker, P; another person (candidate Mary Sweeney), O; and an impersonal object (parental leave legislation), X. According to balance theory, two types of relationships may exist between elements. *Sentiment relations* refer to sentiments or evaluations directed toward objects and people; a sentiment may be either positive (liking, endorsing) or negative (disliking, opposing). *Unit relations* refer to the extent of perceived association between elements. For example, a positive unit relation may result from ownership, a relationship (such as friendship or marriage), or causality. A negative relation indicates dissociation, like that between ex-spouses or members of groups with opposing interests. A null relation exists when there is no association between elements.

Balance theory is concerned with the elements and their interrelations from P’s viewpoint. In the example, the speaker favors parental leave legislation, perceives Mary Sweeney as favoring it, and intends to vote for her. This system is balanced. By definition, a *balanced state* is one in which all three relations are positive or in which one is positive and the other two are negative. An *imbalanced state* is one in which two of the relationships between elements are positive and one is negative or in which all three are negative. For example, “I love (+) Jane; Jane loves (+) opera; I hate (-) opera” is imbalanced.

The theory assumes that an imbalanced state is unpleasant and that when one occurs, the person will try to restore balance. There is considerable empirical evidence that people do prefer balanced states and that attitude change often occurs in response to imbalance. Furthermore, people maintain consistency by responding selectively to new information. There is evidence that people accept information consistent with their existing attitudes and reject information inconsistent with their cognitions. This is the major mechanism by which stereotypes are maintained.

**Dissonance Theory.** Dissonance theory assumes that there are three possible relationships

between any two cognitions. Cognitions are consistent, or *consonant*, if one naturally or logically follows from the other; they are *dissonant* when one implies the opposite of the other. The logic involved is psycho logic—logic as it appears to the individual, not logic in a formal sense. Two cognitive elements may also be irrelevant; one may have nothing to do with the other.

*Cognitive dissonance* is a state of psychological tension induced by dissonant relationships between cognitive elements. There are three situations in which dissonance commonly occurs. First, dissonance occurs following a decision whenever the decision is dissonant with some cognitive elements. Thus, choice between two (or more) attractive alternatives creates dissonance because knowledge that one chose A is dissonant with the positive features of B. The magnitude of the dissonance experienced is a function of the proportion of elements consonant and dissonant with the choice. Second, if a person engages in a behavior that is dissonant with his or her attitudes, dissonance will be created. Third, when events disconfirm an important belief, dissonance will be created if the person had taken action based on that belief. For example, a person who buys an expensive car in anticipation of a large salary increase will experience dissonance if she or he does not receive the expected raise.

Since dissonance is an unpleasant state, the theory predicts that the person will attempt to reduce it. Usually, dissonance reduction involves changes in the person’s attitudes. Thus, following a decision, the person may evaluate the chosen alternative more favorably and the unchosen one more negatively. Following behavior that is dissonant with his or her prior attitude, the person’s attitude toward the behavior may become more positive. An alternative mode of dissonance reduction is to change the importance one places on one or more of the attitudes. Following a decision, the person may reduce the importance of the cognitions that are dissonant with the choice; this is the well-known “sour grapes” phenomenon. Following disconfirmation of a belief, one may increase the importance attached to the disconfirmed belief. A third way to reduce dissonance is to change behavior. If the dissonance following a choice is great, the person may decide to choose B instead of A. Following disconfirmation, the person may change behaviors that were based on the belief.

Numerous books and hundreds of articles about dissonance theory have been published since it was introduced by Festinger in 1957. There is a substantial body of research evidence that supports various predictions from and elaborations of the theory. Taken together, this literature has produced a detailed taxonomy of situations that produce dissonance and of preferred modes of dissonance reduction in various types of situations.

### ATTITUDE STABILITY AND ATTITUDE CHANGE

Both balance and dissonance theories identify the desire for consistency as a major source of stability and change in attitudes. The desire to maintain consistency leads the individual either to interpret new information as congruent with his or her existing cognitions (assimilation) or to reject it if it would challenge existing attitudes (contrast). This process is very important in preserving stability in one's attitudes. At the same time, the desire for consistency will lead to attitude change when imbalance or dissonance occurs. Dissonance theory explicitly considers the link between behavior and attitudes. It predicts that engaging in counterattitudinal behavior may indirectly affect attitudes. This is one mechanism by which social influences on behavior may indirectly affect attitudes. This mechanism comes into play when the person experiences changes in roles and the requirements of the new role are inconsistent with his or her prior attitudes.

The classic perspective in the study of attitude change is the *communication-persuasion paradigm*, which grew out of the work by Hovland and his colleagues at Yale University. *Persuasion* is defined as changing the beliefs or attitudes of a person through the use of information or argument. Attempts at persuasion are widespread in everyday interaction, and the livelihood of advertisers and political consultants. According to the paradigm, each attempt involves source, message, target, and context. Thousands of empirical studies, many of them experiments done in laboratory settings, have investigated the influence of variations in these four components on the outcome of an attempt. In general, if the source is perceived as an expert, trustworthy, or physically attractive, the message is more likely to produce attitude change.

Thus, it is no accident that magazine and television commercials feature young, attractive models. Message variables include the extent of discrepancy from the target's attitude, whether it arouses fear, and whether it presents one or both sides. Under certain conditions, highly discrepant, fear-arousing, and one-sided messages are more effective. Target factors include intelligence, self-esteem, and prior experience and knowledge. The most researched contextual factor is mood. For a review of this research, see Perloff (1993).

In the 1990s considerable work built upon the *elaboration-likelihood model* proposed by Petty and Cacioppo (1986). This model identifies two basic routes through which a message may change a target's attitudes: the central and the peripheral. Persuasion via the *central route* occurs when a target scrutinizes the arguments contained in the message, interprets and evaluates them, and integrates them into a coherent position. This process is termed *elaboration*. In elaboration, attitude change occurs when the arguments are strong, internally coherent, and consistent with known facts. Persuasion via the *peripheral route* occurs when, instead of elaborating the message, the target pays attention primarily to extraneous cues linked to the message. Among these cues are characteristics of the source (expertise, trustworthiness, attractiveness), superficial characteristics of the message such as length, or characteristics of the context such as response of other audience members. Several factors influence whether elaboration occurs. One is the target's involvement with the issue; if the target is highly involved with and cares about the issue addressed by the message, he or she is more likely to elaborate the message. Other factors include whether the target is distracted by noise or some other aspect of the situation, and whether the target is tired. Which route a message elicits in attitude change is important. Attitudes established by the central route tend to be more strongly held and more resistant to change because the target has thought through the issue in more detail.

The literature on attitude change flourished in the 1990s. The effects of many variables have been studied experimentally. A review of the research conducted in 1992 to 1995 concluded that any one variable may have multiple effects, depending upon other aspects of the persuasion attempt (Petty, Wegener, and Fabrigar 1997). For



example, consider the effect of mood on attitude change. Assuming that happy people are more open to new information, one might predict that persuasive messages would lead to greater attitude change in happy persons than in sad ones. However, research indicates that happy people spend less time processing persuasive messages than persons in neutral moods, and so may be less influenced by them. The *hedonic contingency* hypothesis states that the effect of mood depends upon the hedonic tone of the message. Happy people want to maintain their happy mood, so they are likely to scrutinize and process happy messages but not sad ones. Sad people often want to change their mood, and most messages will improve it, so their processing is not affected by messages' hedonic tone. The results of two experiments support the hypothesis (Wegener, Petty, and Smith 1995). Thus, the effect of mood on persuasion depends on whether the message is uplifting or depressing.

#### ATTITUDE-BEHAVIOR RELATION

The attitude-behavior relation has been the focus of considerable research since the early 1970s. This research has identified a number of variables that influence the extent to which one can predict a person's behavior from his or her attitudes.

Some of these variables involve the measurement of the attitude and of the behavior. The correspondence of the two measures is one such variable: one can predict behavior more accurately if the two measures are at the same level of specificity. An opinion poll can predict the outcome of an election because there is high correspondence between the attitude ("Which candidate do you prefer for mayor in next month's election?") and the behavior (voting for a candidate in that election). The length of time between the measure of attitude and the occurrence and measure of the behavior is also an important variable. The shorter the time, the stronger the relationship. The longer the elapsed time, the more likely the person's attitude will change, although some attitudes are stable over long periods, for example, twenty years.

The characteristics of the attitude also influence the degree to which one can predict behavior from it. In order for an attitude to influence behavior, it must be activated, that is, brought

from memory into conscious awareness. An attitude is usually activated by a person's exposure to the attitude object. Attitudes vary in *accessibility*, the ease with which they are activated. The more accessible an attitude is, the more likely it is to guide future behavior (Kraus 1995). Another variable is the source of the attitude. Attitudes based on direct experience with the object are more predictive of behavior. The certainty or confidence with which the person holds the attitude also moderates the attitude-behavior relationship.

The attitude-behavior relation is also influenced by situational constraints—the social norms governing behavior in a situation. An attitude is more likely to be expressed in behavior when the behavior is consistent with these norms.

An important attempt to specify the relationship between attitude and behavior is the *theory of reasoned action*, developed by Fishbein and Ajzen (1975). According to this theory, behavior is determined by behavioral intention. Behavioral intention is determined by two factors: attitude and subjective norm. Attitude is one's beliefs about the likely consequences of the behavior and one's evaluation—positive or negative—of each of those outcomes. Subjective norm is the person's belief about other important persons' or groups' reactions to the behavior and the person's motivation to comply with the expectations of each. One of the strengths of the theory is this precise specification of the influences on behavioral intention. It is possible to measure quantitatively each of the four components (likely consequences, evaluation, likely reactions, motivation to comply) and use these to make precise predictions of behavior. Many empirical studies report results consistent with such predictions. The theory applies primarily to behavior that is under conscious, volitional control.

On the other hand, researchers have shown that attitudes can affect behavior without being brought into conscious awareness (Bargh 1996). Attitudes toward objects influence our judgments and behavior toward those objects without conscious awareness or intent. Stereotypes of social groups are often activated automatically, as soon as an individual is perceived as a member of the group. Automatic processing is more likely when the individual experiences information overload, time pressures, or is not interested in engaging in effortful processing.

### ATTITUDE AS INDICATOR

Increasingly, attitudes are employed as indicators. Some researchers use attitude measures as indicators of concepts, while others study changes in attitudes over time as indicators of social change.

**Indicators of Concepts.** Measures of specific attitudes are frequently used as indicators of more general concepts. For example, agreement with the following statement is interpreted as an indicator of powerlessness: "This world is run by the few people in power, and there is not much the little guy can do about it." Powerlessness is considered to be a general orientation toward the social world and is a sense that one has little or no control over events. Feelings of powerlessness may be related to such varied behaviors as vandalism, not voting in elections, and chronic unemployment.

Attitude measures have been used to assess many other concepts used in the analysis of political attitudes and behavior. These include the liberalism-conservatism dimension, political tolerance (of radical or unpopular groups), trust in or disaffection with national institutions, and relative deprivation. (For a review of this literature, see Kinder and Sears 1985.) Attitude measures are used to assess many other characteristics of persons. In the realm of work these include occupational values, job satisfaction, and leadership style.

A major concern when attitudes are employed as indicators is *construct validity*, that is, whether the specific items used are valid measures of the underlying concept. In the powerlessness example, the connection between the content of the item and the concept may seem obvious, but even in cases like this it is important to demonstrate validity. A variety of analytic techniques may be used, including interitem correlations, factor analysis, and LISREL.

**Indicators of Social Change.** Two methodological developments have made it possible to use attitudes to study social change. The first was the development of probability sampling techniques, which allow the investigator to make inferences about the characteristics of a population from the results obtained by surveying a sample of that population. The second is the use of the same attitude measures in surveys of representative samples at two or more points in time.

A major source of such data is the General Social Survey (GSS), an annual survey of a probability sample of adults. The GSS repeats a core set of items on a roughly annual basis, making possible the study of changes over a period of thirty years. Many of these items were drawn directly from earlier surveys, making comparisons over a forty- or fifty-year timespan possible. A published book describes these items and presents the responses obtained each time the item was used (Niemi, Mueller, and Smith 1989). Other sources of such data include the National Election Studies and the Gallup Polls.

This use of attitude items reflects a general concern with social change at the societal level. The investigator uses aggregate measures of attitudes in the population as an index of changes in cultural values and social institutions. Two areas of particular interest are attitudes toward race and gender roles. In both areas, efforts have been made to improve access to educational programs, jobs, and professions, increase wages and salaries, and provide greater opportunity for advancement. The availability of responses to the same attitude items over time allows us to assess the consistency between these social changes and attitudes in the population. Consider the question "Do you think civil rights leaders are trying to push too fast, are going too slowly, or are moving at about the right speed?" This question was asked in surveys of national samples every two years from 1964 to 1976 and in 1980. The percentage of whites replying "too fast" declined from 74 percent in 1964 to 40 percent in 1980 (Bobo 1988), suggesting increased white support for the black movement. In general, research indicates that both racial and gender-role attitudes became more liberal between 1960 and 1990, and this finding is consistent with the social changes in these areas. Other topics that have been studied include attitudes toward abortion, social class identification, and subjective quality of life.

There are several issues involved in this use of attitude items. The first is the problem of "nonattitudes." Respondents may answer survey questions or endorse statements even though they have no attitude toward the object. In fact, when respondents are questioned about fictional objects or organizations, some of them will express an opinion. Schuman and Kalton (1985) discuss

this issue in detail and suggest ways to reduce the extent to which nonattitudes are given by respondents.

The second issue involves the interpretation of responses to items. In the example above, the analyst assumes that white respondents who reply "Too fast" feel threatened by the movement. However, there is evidence that small changes in the wording of survey items can produce substantial changes in aggregate response patterns. This evidence and guidelines for writing survey items are discussed in Schuman and Presser (1996).

Finally, there is the problem of equivalence in meaning over time. In order to make meaningful comparisons across time, the items need to be the same or equivalent. Yet over time the meaning of an item may change. Consider the item "Are you in favor of *desegregation*, strict *segregation*, or something in between?" This question was asked of national samples in 1964 and every two years from 1968 to 1978. From 1964 to 1970, the percentage of white, college-educated adults endorsing desegregation increased; from 1970 to 1978, the percentage decreased steadily. Until 1970, desegregation efforts were focused on the South; after 1970, desegregation efforts focused on school integration in northern cities. Evidence suggests that endorsement of desegregation changed because the meaning of the question for white adults changed (Schuman, Steeh, and Bobo 1985).

## REFERENCES

- Allport, Gordon W. 1985 "The Historical Background of Social Psychology." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Bargh, John A. 1996 "Automaticity in Social Psychology." In E.T. Higgins and A. Kruglanski, eds., *Social Psychology: Handbook of Basic Principles*. New York: The Guilford Press.
- Bobo, Lawrence 1988 "Attitudes toward the Black Political Movement: Trends, Meaning, and Effects on Racial Policy Preference." *Social Psychology Quarterly* 51:287-302.
- Dawes, Robyn M., and Tom L. Smith 1985 "Attitude and Opinion Measurement." In G.Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Eagly, Alice, and Shelly Chaiken 1998 "Attitude Structure and Function." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *Handbook of Social Psychology*, 4th ed. New York: McGraw-Hill.
- Felson, Richard B. 1996 "Mass Media Effects on Violent Behavior." In J. Hagen and K. Cook, eds., *Annual Review of Sociology*, vol. 22. Palo Alto, Calif.: Annual Reviews.
- Festinger, Leon 1957 *A Theory of Cognitive Dissonance*. Stanford, Calif.: Stanford University Press.
- Fishbein, Martin, and Icek Ajzen 1975 *Belief, Attitude, Intention, and Behavior*. Reading, Mass.: Addison-Wesley.
- Geen, Russel 1998 "Aggression and Antisocial Behavior." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *Handbook of Social Psychology*, 4th ed. New York: McGraw-Hill.
- Kiecolt, K. Jill 1988 "Recent Developments in Attitudes and Social Structure." In W. R. Scott and J. Blake, eds., *Annual Review of Sociology*, vol. 14. Palo Alto, Calif.: Annual Reviews.
- Kinder, Donald R., and David O. Sears 1985 "Public Opinion and Political Action." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Kraus, S.J. 1995 "Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature." *Personality and Social Psychology Bulletin* 18:152-162.
- McGuire, William J. 1985 "Attitudes and Attitude Change." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Niemi, Richard G., John Mueller, and Tom W. Smith 1989 *Trends in Public Opinion: A Compendium of Survey Data*. New York: Greenwood Press.
- Perloff, R. M. 1993 *The Dynamics of Persuasion*. Hillsdale, N.J.: Lawrence Erlbaum.
- Petty, Richard, and John Cacioppo 1986 "The Elaboration-Likelihood Model of Persuasion." *Advances in Experimental Social Psychology* 19:123-205.
- Petty, Richard, Duane Wegener, and Leandre Fabrigan 1997 "Attitudes and Attitude Change." *Annual Review of Psychology* 48:609-647.
- Roberts, Donald F., and Nathan Maccoby 1985 "Effects of Mass Communication." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Schuman, Howard 1995 "Attitudes, Beliefs, and Behavior." In K. S. Cook, G. A. Finc, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- , and Stanley Presser 1996 *Questions and Answers in Attitude Surveys*. Thousand Oaks, Calif.: Sage Publications.

- Schuman, Howard, and Graham Kalton 1985 "Survey Methods." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Schuman, Howard, Charlotte Steeh, and Lawrence Bobo 1985 *Racial Attitudes in America*. Cambridge, Mass.: Harvard University Press.
- Wegener, Duane T., Richard E. Petty, and Stephen M. Smith 1995 "Positive Mood can Increase or Decrease Message Scrutiny: The Hedonic Contingency View of Mood and Message Processing." *Journal of Personality and Social Psychology* 69:5-15.

JOHN D. DELAMATER

## ATTRIBUTION THEORY

Attribution is a cognitive process that entails linking an event to its causes. Attribution is one of a variety of cognitive inferences that are included within social cognition, which is one of several theoretical models within social psychology. Social cognition has been the most dominant social psychological perspective within psychology since the 1960s, and this is evident in the popularity of research on attribution. In the mid-1970s, as much as 50 percent of the articles in major social psychology journals concerned attributional processes, in part because attribution theory is relevant to the study of person perception, event perception, attitude change, the acquisition of self-knowledge, and a host of applied topics including therapeutic interventions, close relationships, legal and medical decision making, and so forth. Although the proportion of published research that focused on this topic declined during the 1980s, attribution remains one of the more popular fields of social psychological research.

### DEFINITION

An attribution is an inference about why an event occurred. More generally, "attribution is a process that begins with social perception, progresses through a causal judgment and social inference, and ends with behavioral consequences" (Crittenden 1983, p. 426). Although most theories of and research on attribution focus on causal inference, empirical research has dealt with attributions not

only of cause but also of blame and responsibility. Although these types of attributions are closely related, they are not conceptually identical. Furthermore, because personality characteristics constitute a major category of potential causes of behavior, attributions about individuals' traits (both one's own traits and those of others) have received explicit theoretical attention.

### MAJOR THEORIES OF ATTRIBUTION

Even though attribution has been one of the most popular social psychological research topics in the social sciences, only a few theories of attribution have been developed. The study of attribution began with Fritz Heider's (1958) original attempt to provide a systematic, conceptual explanation of "naive" psychology. Heider maintained that people strive to understand, predict, and control events in their everyday lives in much the same way as scientists do in their professional lives. On the basis of observation, individuals form theories about their social worlds, and new observations then serve to support, refute, or modify these theories. Because people act on the basis of their beliefs, Heider argued that it is important to understand this layperson's psychology. Although Heider did not develop an explicit theory of attribution, he did assert several principles that have guided all subsequent theorizing on this topic.

Primary among these principles is the notion that people are inclined to attribute actions to stable or enduring causes rather than to transitory factors. Heider also stressed the importance of distinguishing unintentional from intentional behavior, a distinction that has been particularly influential in theories of the attribution of responsibility. He identified environmental and personal factors as two general classes of factors that produce action and hypothesized that an inverse relationship exists between these two sets of causes. He also suggested that the "covariational principle" is fundamental to attribution: An effect is attributed to a factor that is present when the effect is present and to a factor that is absent when the effect is absent. Heider's early analyses of social perception represent a general conceptual framework about common sense, implicit theories people use in understanding events in their daily lives. The two most influential theories of attribution are based on Heider's work but go beyond it

in the development of more systematic statements about attributional processes.

**Covariational model.** Harold Kelley's (1967, 1973) covariational model of attribution addresses the question of whether a given behavior is caused by an actor or, alternatively, by an environmental stimulus with which the actor engages. According to this model, the attribution of cause is based on three types of information: consensus, distinctiveness, and consistency. Consensus refers to the similarity between the actor's behavior and the behavior of other people in similar circumstances. Distinctiveness refers to the generality of the actor's behavior: Does she or he behave in this way toward stimuli in general, or is the behavior specific to this stimulus? Consistency refers to the actor's behavior toward this stimulus across time and modality. There are many possible combinations of these three types of information, but Kelley makes explicit predictions about just three. The combination of high consensus, high distinctiveness, and high consistency supports an attribution to the environmental stimulus, whereas a profile of low consensus, low distinctiveness, and high consistency supports an attribution to the actor. When the behavior is inconsistent, regardless of the level of consensus or distinctiveness, an attribution to circumstances is predicted.

Empirical tests of Kelley's model have focused either on the effects of a particular type of information or, more in keeping with his formulation, on the effects of particular patterns of information (McArthur 1972). In an innovative analysis, Miles Hewstone and Jos Jaspars (1987) proposed a different logic (although one consistent with Kelley's model), suggesting that potential attributers consider whether different causal loci are necessary and sufficient conditions for the occurrence of an effect. They conclude that the notion of causality is flexible and thus assert that there may be some advantage to conceiving of situation-specific notions of causality. More recent work has examined the universality and external validity of Kelley's model. Irina Anderson and Geoffrey Beattie (1998), for example, analyzed actual conversations between men and women talking about rape. They found that men tended to use the reasoning outlined in Kelley's model by making reference to consensus, distinctiveness, and consistency, as well as by using these types of information to formulate

attributions for behavior. Women, on the other hand, made less use of these variables and introduced the variable of "foreseeability" into their analyses.

**Correspondent inference.** The theory of correspondent inference (Jones and Davis 1965; Jones and McGillis 1976) addresses the attribution of personality traits to actors on the basis of their behavior and focuses on attributions about persons in greater depth than does Kelley's covariational model. These two theories thus address different questions. Kelley asks: When do we attribute an event to an actor or to some stimulus in the environment? Edward Jones asks: When do we attribute a trait to an actor on the basis of her or his behavior? The theory of correspondent inference focuses more narrowly on the actor but also yields more information about the actor in that it specifies what it is about the actor that caused the behavior. Jones and his coauthors predict that two factors guide attributions: (1) the attributer's prior expectancies for behavior, specifically, expectancies based either on knowledge of earlier behaviors of the actor (target-based) or on the actor's social category memberships (category-based), and (2) the profile of effects that follow from the behavioral choices available to the actor.

Edward Jones and Daniel McGillis propose that expectancies determine the degree of confidence with which a particular trait is attributed; the lower the expectancy of behavior, the more confident the attribution. The profile of effects helps the attributer identify what trait might have produced the behavior in question. Noncommon effects—effects that follow from only one of the behavioral options—provide information about the particular disposition. The fewer the noncommon effects, the clearer the attribution. Thus, behavior that contradicts prior expectancies and a profile of behavioral choices with few noncommon effects combine to maximize the possibility of attributing a disposition to the actor (a correspondent inference). Empirical research generally has supported these predictions.

These two models share some attributional principles. Expectancy variables (target-based and category-based) are analogous to Kelley's types of information. Although the predicted effects of consensus information and its analogue, category-based expectancies, are compatible, the predicted

effects of consistency and distinctiveness information and their analogue, target-based expectancies, present some incompatibilities. This contradiction has been evaluated both conceptually and empirically (Howard and Allen 1990).

**Attribution in achievement situations.** Bernard Weiner (1974) and his colleagues have applied attributional principles in the context of achievement situations. According to this model, we make inferences about an individual's success on the basis of the individual's ability to do the task in question, how much effort is expended, how difficult the task is, and to what extent luck may have influenced the outcome. Other possible causal factors have since been added to this list. More important perhaps is Weiner's development of, first, a structure of causal dimensions in terms of which these causal factors can be described and, second, the implications of the dimensional standing of a given causal factor (Weiner, Russell, and Lerman 1978). The major causal dimensions are locus (internal or external to the actor), stability or instability, and intentionality or unintentionality of the factor. Thus, for example, ability is internal, stable, and unintentional. The stability of a causal factor primarily affects judgments about expectancies for future behavior, whereas locus and intentionality primarily affect emotional responses to behavior. This model has been used extensively in educational research and has guided therapeutic educational efforts such as attribution retraining. Cross-cultural research has explored the cultural generalizability of these models. Paul Tuss, Jules Zimmer, and Hsiu-Zu Ho (1995) and Donald Mizokawa and David Rickman (1990) report, for instance, that Asian and Asian American students are more likely to attribute academic failure and success to effort than are European American students, who are more likely to attribute performance to ability. European American students are also more likely to attribute failure to task difficulty. Interestingly, as Asian Americans spend more time in the United States, they place less emphasis on the role of effort in performance.

**Attribution biases.** The theoretical models described above are based on the assumption that social perceivers follow the dictates of logical or rational models in assessing causality. Empirical research has demonstrated, not surprisingly, that there are systematic patterns in what has been variously conceived of as bias or error in the

attribution process (Ross 1977). Prominent among these is what has been called the "fundamental attribution error," the tendency of perceivers to overestimate the role of dispositional factors in shaping behavior and to underestimate the impact of situational factors. One variant of this bias has particular relevance for sociologists. This is the general tendency to make inadequate allowance for the role-based nature of much social behavior. That is, perceivers fail to recognize that behavior often derives from role memberships rather than from individual idiosyncrasy. Again, cross-cultural research has called into question the generalizability of this bias. Joan Miller (1984) shows that the tendency to attribute behavior to persons is markedly more prominent in the United States, among both adults and children, whereas a tendency to attribute behavior to situational factors is more prominent among Indian Hindus, both adults and children, calling into question the "fundamentality" of this attributional pattern.

A second systematic pattern is the actor-observer difference, in which actors tend to attribute their own behavior to situational factors, whereas observers of the same behavior tend to attribute it to the actor's dispositions. A third pattern concerns what have been called self-serving or egocentric biases, that is, attributions that in some way favor the self. According to the false consensus bias, for example, we tend to see our own behavioral choices and judgments as relatively common and appropriate whereas those that differ from ours are perceived as uncommon and deviant. There has been heated debate about whether these biases derive from truly egotistical motives or reflect simple cognitive and perceptual errors.

#### METHODOLOGICAL AND MEASUREMENT ISSUES

The prevalent methodologies and measurement strategies within attributional research have been vulnerable to many of the criticisms directed more generally at social cognition and to some directed specifically at attribution. The majority of attributional research has used structured response formats to assess attributions. Heider's original distinction between person and environmental cause has had a major influence on the development of these structured measures. Respondents typically are asked to rate the importance of situational and

dispositional causes of events. These ratings have been obtained on ipsative scales as well as on independent rating scales. Ipsative measures pose these causes as two poles on one dimension; thus, an attribution of cause to the actor's disposition is also a statement that situational factors are not causal. This assumed inverse relationship between situational and dispositional causality has been rejected on conceptual and empirical grounds. In more recent studies, therefore, respondents assign each type of causality separately. (Ipsative measures are appropriate for answering some questions, however, such as whether the attribution of cause to one actor comes at the expense of attribution to another actor or to society.) In theory, then, both dispositional and situational variables could be identified as causal factors.

The breadth of these two categories has also been recognized as a problem. Dispositional causes may include a wide variety of factors such as stable traits and attitudes, unstable moods and emotions, and intentional choices. Situational cause is perhaps an even broader category. It is quite possible that these categories are so broad as to render a single measure of each virtually meaningless. Thus, researchers often include both general and more specific, narrower responses as possible choices (e.g., choices of attributing blame to an assailant or to the situation might be refined to the assailant's use of a weapon, physical size, and psychological state, on the one hand, and the location, time of day, and number of people nearby, on the other).

Structured measures of attributions are vulnerable to the criticism that the categories of causes presented to respondents are not those they use in their everyday attributions. Recognizing this limitation, a few researchers have used open-ended measures. Comparative studies of the relative utility of several different types of measures of causal attributions conclude that scale methods perform somewhat better in terms of their inter-test validity and reliability, although open-ended measures are preferable when researchers are exploring causal attributions in new situations. Some researchers have attempted to overcome some of the limitations of existing scales; Curtis McMillen and Susan Zuravin (1997), for example, have developed and refined "Attributions of Responsibility" and "Blame Scales" for use in clinical research that may be useful for other kinds of attributional situations.

The great majority of attribution studies use stimuli of highly limited social meaning. Generally, the behavior is represented with a brief written vignette, often just a single sentence. Some researchers have shifted to the presentation of visual stimuli, typically with videotaped rather than actual behavioral sequences, in order to ensure comparability across experimental conditions. Recognizing the limitations of brief, noncontextualized stimuli, a few researchers have begun to use a greater variety of more extended stimuli including newspaper reports and published short stories. Most of these stimuli, including the videotaped behavioral sequences, rely heavily on language to convey the meaning of behavior. Conceptual attention has turned recently to how attribution relies on language and to the necessity of considering explicitly what that reliance means. Some researchers (Anderson and Beattie 1998; Antaki and Leudar 1992) have attempted to use more naturalistic approaches, incorporating into the study the analysis of spontaneous conversations among study participants.

#### WHEN DO WE MAKE ATTRIBUTIONS?

Long after attribution had attained its popularity in social psychology, a question that perhaps should have been raised much earlier began to receive attention: When do we make attributions? To what extent are the attributions in this large body of research elicited by the experimental procedures themselves? This is a question that can be directed to any form of social cognition. It is particularly relevant, however, to attribution. Most people, confronted with a form on which they are to answer the question "Why?," do so. There is no way of knowing, within the typical experimental paradigm, whether respondents would make attributions on their own. In a sense there are two questions: Do people make attributions spontaneously, and if they do, under what circumstances do they do so?

In response to the first question, Weiner (1985) has marshaled impressive evidence that people do indeed make attributions spontaneously. Inventive procedures have been developed for assessing the presence of attributional processing that is not directly elicited. This line of evidence has dealt almost entirely with causal attributions. Research

suggests that trait attributions may be made spontaneously much more often than causal attributions. In response to the second question, a variety of studies suggest that people are most likely to make attributions when they encounter unexpected events or events that have negative implications for them.

### SOCIOLOGICAL SIGNIFICANCE OF ATTRIBUTION

Attribution is a cognitive process of individuals; much of the extant research on attribution is, accordingly, highly individualistic. In the 1980s, however, researchers began to pay increasing attention to the sociological relevance of attribution. The process of attribution itself is fundamentally social. Attribution occurs not only within individuals but also at the interpersonal, intergroup, and societal levels. Moreover, the process of attribution may underlie basic sociological phenomena such as labeling and stratification.

**Interpersonal attribution.** At the interpersonal level, attribution is basic to social interaction. Interpersonal encounters are shaped in many ways by attributional patterns. Behavioral confirmation, or self-fulfilling prophecies, illustrate the behavioral consequences of attribution in social interaction; attribution of specific characteristics to social actors creates the expectancies that are then confirmed in behavior. Considering attribution at this interpersonal level demonstrates the importance of different social roles and perspectives (actors vs. observers) as well as how attribution is related to evaluation. The self is also important to the attribution process; the evidence for attributional egotism (self-esteem enhancing attributional biases), self-presentation biases, and egocentrism is persuasive. Attributions also affect social interaction through a widespread confirmatory attribution bias that leads perceivers to conclude that their expectancies have been confirmed in social interaction.

Research has challenged the universality of egocentric biases by examining differences in attributional styles according to race, class, and gender. Several studies (Reese and Brown 1995; Wiley and Crittenden 1992; Broman 1992; Andrews and Brewin 1990) have shown that women

are less likely than men to make self-serving dispositions. In a study of academics' accounts of their success in the profession, Mary Glenn Wiley and Kathleen Crittenden (1992) argue that women explain their success in a more modest manner in order to preserve a feminine identity, at the expense of their professional identities. This attributional style may make it more likely for women to blame themselves for various types of negative situations. Consistent with this reasoning, Bernice Andrews and Chris Brewin (1990) found that female victims of marital violence tended to blame themselves for their experience of violence. Childhood experiences of physical or sexual abuse increased women's chances of characterological self-blame when, as adults, they found themselves in abusive relationships. (Other research, however, notes that the relationship between attributions and adjustment is complex and not always so straightforward; see McMillen and Zuravin 1997.)

Race and class characteristics also create distinct attributional patterns; research on these factors has been more likely than research on gender to consider possible interactive effects on attributions. In a study conducted by Fathali Moghaddam and colleagues (1995) on attributional styles of whites, blacks, and Cubans in Miami, for example, middle-class black respondents were more likely to blame negative outcomes on discrimination than were lower-class blacks. Lower-class whites were the only group to attribute failure to themselves personally.

The great preponderance of research on social interaction has been based on relationships between strangers in experimental contexts, which may seem to undermine the claim that attribution is significant for interpersonal interaction. The best evidence of this significance, then, is the increasingly large body of research on the role of attribution in the formation, maintenance, and dissolution of close relationships. There is substantial evidence that attributions are linked to relationship satisfaction and behaviors such as conflict resolution strategies. There is also evidence that distressed and nondistressed couples make differing attributions for significant events in their relationships; these patterns may actually serve to maintain marital distress among troubled



couples, thus ultimately influencing marital satisfaction. Attributions also play an important role in relationship dissolution. Attributions are a critical part of the detailed accounts people provide for the dissolution of their relationships, and these accounts go beyond explanation to rationalize and justify the loss of relationships.

Research on interpersonal attribution extends the intrapersonal approach in several ways. When people who interact have substantial knowledge of and feelings about each other, the attribution process involves evaluation as well as cognition. Issues of communication, and hence potential changes in preexisting attributions for recurrent relationship events, also become salient at the interpersonal level. There is very little research on how attributions change through interaction and relationships, but this is clearly a significant topic. Attributions at the interpersonal level also entail greater concern with accountability for action; causality at this level also raises issues of justification.

**Intergroup attribution.** Intergroup attribution refers to the ways in which members of different social groups explain the behavior of members of their own and other social groups. At this level, social categorization has a direct impact on attribution. Studies using a variety of subjects from different social groups and often different countries show consistent support for an ingroup-serving attributional pattern, for example, a tendency toward more dispositional attributions for positive as opposed to negative behavior for ingroup actors. The evidence for the converse pattern, more dispositional attributions for negative as opposed to positive behavior for outgroup actors, is not as strong. (Moreover, these patterns are stronger in dominant than in dominated groups.) Social desirability biases can counteract this pattern; Steven Little, Robert Sterling, and Daniel Tingstrom (1996), for example, found that white respondents held black actors less responsible for participating in a bar fight than did black respondents. The authors suggest that this finding may be due to the desire of the white sample—undergraduates from a suburban area—to appear racially progressive.

Parallel studies of a group's success and failure show a consistent pattern of ingroup protection. Outgroup failure is attributed more to lack of ability than is ingroup failure. Effects of group

membership on attributions about success are not as strong. Interestingly, there is also some evidence of outgroup-favoring and/or ingroup-derogating attributions among widely recognized lower-status, dominated groups such as migrant labor populations (Hewstone 1989). A third form of evidence of intergroup attribution is provided by studies of attributions about social positions occupied by existing groups. In general, these studies, like those cited above, show higher ratings of ingroup-serving as opposed to outgroup-serving attributions.

**Societal attribution.** At the societal level, those beliefs shared by the members of a given society form the vocabulary for social attributions. The concept of social representations, which has its origins in Durkheim's concept of "representation collectives," was developed by Serge Moscovici (1976) to represent how knowledge is shared by societal members in the form of common-sense theories about that society. Social representations are intimately connected to the process of attribution. Not only are explanation and accountability part of a system of collective representations, but such representations determine when we seek explanations. Social representations serve as categories that influence the perception and processing of social information; moreover, they underscore the emphasis on shared social beliefs and knowledge. Social representations are useful in interpreting research on laypersons' explanations of societal events such as poverty and wealth, unemployment, and racial inequality. Poverty tends to be attributed to individualistic factors, for example, whereas unemployment tends to be attributed to societal factors. Not surprisingly, these patterns may be qualified by attributors' own class backgrounds; although middle-class people attribute poverty more often to internal factors, those who are themselves poor attribute poverty more often to external factors such as governmental policies (Singh 1989). Gender and racial stereotyping can also shape attributions. Cynthia Willis, Marianne Hallinan, and Jeffrey Melby (1996), for example, found that respondents with a traditional sex-role orientation showed a favorable bias toward the male perpetrator in domestic violence situations. When the female victim was African American and married, both egalitarians and traditionalists were less likely to attribute blame to the man.

Moving away from an emphasis on normative stereotyping, research has considered possible effects of counter-stereotyping on attributional patterns, and of attributions on changes in stereotypes. Portrayals of structurally subordinate groups that counter stereotypical expectations can increase the perceived credibility of members of those groups (Power, Murphy, and Coover 1996). Moreover, an attribution of counterstereotypic behavior to dispositional factors of clearly typical outgroup members can modify outgroup stereotypes (Wildner, Simon, and Faith 1996).

Cross-cultural research illustrates another aspect of societal attributions. This work compares the extent and type of attributional activity across cultures. Although there has been some support for the applicability of Western models of attribution among non-Western cultures, most of this research has demonstrated in a variety of ways the cultural specificity of particular patterns of attribution (Bond 1988). As noted above, for example, Joan Miller (1984) provides cross-cultural empirical evidence that the fundamental attribution error, the tendency to attribute cause more to persons than to situations, is characteristic of Western but not non-Western societies. Attributions about the self also vary across cultures; members of some non-Western cultures attribute performance more to effort than to ability; the opposite pattern has been found in the United States. Cultural patterns also shape responsibility attributions; V. Lee Hamilton and Shigeru Hagiwara (1992) found that U.S. (and male) respondents were more likely to deny responsibility for their own inappropriate behaviors than were Japanese (and female) respondents, who were more likely to apologize for the behaviors. Hamilton and Hagiwara argue that the Japanese (and women) were more concerned with maintaining the quality of the interaction between the accuser and the accused.

### SOCIOLOGICAL APPLICATIONS

It may be useful to identify several sociological phenomena to which theories of attribution are relevant. A number of scholars have suggested integrating attribution theory and labeling theory (Crittenden 1983). Attribution occurs within an individual; labels are applied by a group. Judith Howard and Randy Levinson (1985) offer empirical evidence that the process of applying a label is

directly analogous to attribution. They report that the relationships between attribution information and jury verdicts are consistent with predictions based on a labeling perspective.

Richard Della Fave (1980) demonstrates the importance of attribution for understanding a key but neglected aspect of stratification, namely, how it is that stratification systems become legitimated and accepted by those disadvantaged as well as by those advantaged by those systems. He draws heavily on attribution theory in developing a theory of legitimation and in identifying possible sources of delegitimation.

Attribution is a significant social process that ranges widely from cognitive processes to collective beliefs. The field is still imbalanced; more work has been done at the intrapersonal and interpersonal levels. There is evidence for both intergroup and societal attributions, however, and research at these two levels is steadily increasing (Hewstone 1989). Recent research has demonstrated connections across areas as diverse as social cognition, social interaction, intergroup relations, and social representations; these connections provide increasing evidence of the importance of attribution for sociological phenomena. Indeed, in the next decade it may be that the fruits of attribution theories will be evident more in research on these other topics, than in research on attribution alone.

### REFERENCES

- Anderson, Irina, and Geoffrey Beattie 1996 "How Important is Kelley's Model of the Attribution Process When Men and Women Discuss Rape in Conversation?" (*Semiotica* 1/2:1-21).
- Andrews, Bernice, and Chris R. Brewin 1990 "Attributions of Blame for Marital Violence: A Study of Antecedents and Consequences." *Journal of Marriage and the Family* 52:757-767.
- Antaki, Charles, and Ivan Leudar 1992 "Explaining in Conversation: Towards an Argument Model." *European Journal of Social Psychology* 22:181-194.
- Bond, Michael H. (ed.) 1988 *The Cross-Cultural Challenge to Social Psychology*. Newbury Park, Calif.: Sage.
- Broman, Clifford L. 1992 "The Black Experience: Attributing the Causes of Labor Market Success." *National Journal of Sociology* 6:77-90.
- Crittenden, Kathleen S. 1983 "Sociological Aspects of Attribution." *Annual Review of Sociology* 9:425-446.

- Della Fave, L. Richard 1980 "The Meek Shall Not Inherit the Earth: Self-Evaluation and the Legitimacy of Stratification." *American Sociological Review* 45:955-971.
- Hamilton, V. Lee, and Shigeru Hagiwara 1992 "Roles, Responsibility, and Accounts Across Cultures." *International Journal of Psychology* 27:157-179.
- Heider, Fritz 1958 *The Psychology of Interpersonal Relations*. New York: Wiley.
- Hewstone, Miles 1989 *Causal Attribution: From Cognitive Processes to Collective Beliefs*. Oxford: Basil Blackwell.
- , and Jos M. F. Jaspars 1987 "Covariation and Causal Attribution: A Logical Model of the Intuitive Analysis of Variance." *Journal of Personality and Social Psychology* 53:663-672.
- Howard, Judith A., and Carolyn Allen 1990 "Making Meaning: Revealing Attributions Through Analyses of Readers' Responses." *Social Psychology Quarterly* 52:280-298.
- Howard, Judith A., and Randy Levinson 1985 "The Overdue Courtship of Attribution and Labeling." *Social Psychology Quarterly* 48:191-202.
- Jones, Edward E., and Keith E. Davis 1965 "From Acts to Dispositions: The Attribution Process in Person Perception." *Advances in Experimental Social Psychology* 2:219-266.
- Jones, Edward E., and Daniel McGillis 1976 "Correspondent Inferences and the Attribution Cube: A Comparative Reappraisal." In J. H. Harvey, W. Ickes, and R. F. Kidd, eds., *New Directions in Attribution Research*, Vol. 1. Hillsdale, N.J.: Erlbaum.
- Kelley, Harold H. 1967 "Attribution Theory in Social Psychology." *Nebraska Symposium of Motivation* 15:192-238.
- 1973 "The Processes of Causal Attribution." *American Psychologist* 28:107-128.
- Little, Steven G., Robert C. Sterling, and Daniel H. Tingstrom 1996 "The Influence of Geographic and Racial Cues on Evaluation of Blame." *The Journal of Social Psychology* 136:373-379.
- McArthur, Leslie Z. 1972 "The How and What of Why: Some Determinants and Consequences of Causal Attribution." *Journal of Personality and Social Psychology* 22:171-197.
- McMillen, Curtis, and Susan Zuravin 1997 "Attributions of Blame and Responsibility for Child Sexual Abuse and Adult Adjustment." *Journal of Interpersonal Violence* 12:30-48.
- Miller, Joan G. 1984 "Culture and the Development of Everyday Social Explanation." *Journal of Personality and Social Psychology* 46:961-978.
- Mizokawa, Donald T., and David B. Ryckman 1990 "Attributions of Academic Success and Failure: A Comparison of Six Asian-American Ethnic Groups." *Journal of Cross-Cultural Psychology* 21:434-451.
- Moghaddam, Fathali M., Donald M. Taylor, Wallace E. Lambert, and Amy E. Schmidt 1995 "Attributions and Discrimination: A Study of Attributions to the Self, the Group, and External Factors among Whites, Blacks, and Cubans in Miami." *Journal of Cross-Cultural Psychology* 26:209-220.
- Moscovici, Serge 1976 *La Psychanalyse: Son image et son public*, 2nd ed. Paris: Presses Universitaires de France.
- Moskowitz, G. B. 1996 "The Mediational Effects of Attributions and Information Processing in Minority Social Influence." *British Journal of Social Psychology* 35:47-66.
- Power, J. Gerard, Sheila T. Murphy, and Gail Coover 1996 "Priming Prejudice: How Stereotypes and Counter-Stereotypes Influence Attribution of Responsibility and Credibility among Ingroups and Outgroups." *Human Communication Research* 23:36-58.
- Reese, Laura T., and Ronald E. Brown 1995 "The Effects of Religious Messages on Racial Identity and System Blame among African Americans." *The Journal of Politics* 57:24-43.
- Ross, Lee 1977 "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process." *Advances in Experimental Social Psychology* 19:174-220.
- Singh, Anup K. 1989 "Attribution Research on Poverty: A Review." *Psychologia* 32:143-148.
- Tuss, Paul, Jules Zimmer, and Hsiu-Zu Ho 1995 "Causal Attributions of Underachieving Fourth-Grade Students in China, Japan, and the United States." *Journal of Cross-Cultural Psychology* 26:408-425.
- Weiner, Bernard 1974 *Achievement Motivation and Attribution Theory*. Morristown, N.J.: General Learning Press.
- 1985 "'Spontaneous' Causal Thinking." *Psychological Bulletin* 97:74-84.
- , Dan Russell, and David Lerman 1978 "Affective Consequences of Causal Ascription." In J. H. Harvey, W. Ickes., and R. F. Kidd, eds., *New Directions in Attribution Research*, Vol. 3. Hillsdale, N.J.: Erlbaum.
- Wilder, David A., Andrew F. Simon, and Myles Faith 1996 "Enhancing the Impact of Counterstereotypic Information: Dispositional Attributions for Deviance." *Journal of Personality and Social Psychology* 71:276-287.
- Wiley, Mary Glenn, and Kathleen S. Crittenden 1992 "By Your Attributions You Shall Be Known: Consequences of Attributional Accounts for Professional and Gender Identities." *Sex Roles* 27:259-276.

Willis, Cynthia E., Marianne N. Hallinan, and Jeffrey Melby 1996 "Effects of Sex Role Stereotyping among European American Students on Domestic Violence Culpability Attributions." *Sex Roles* 34:475-491.

JUDITH A. HOWARD  
DANIELLE KANE

# B

---

## **BALANCE THEORY**

*See* Attitudes; Cognitive Consistency Theories.

## **BANKRUPTCY AND CREDIT**

Every society must resolve the tension between debtors and creditors, especially if the debtors cannot pay or cannot pay quickly enough. Many of the world's religions have condemned lending money for interest, at least among co-religionists. In traditional societies, money lenders, although necessary for ordinary commerce, were often viewed as morally suspect. The development of a robust capitalism was based upon raising capital by paying interest or dividends, and so it has been important for capitalist societies to develop institutions and mechanisms for handling debt and credit.

The inherent tension between creditors and debtors turns upon the creditor's claim to justice as the property owner and the debtor's interest in fairness in the terms of repayment. To be sure, lenders sometimes used their position to create social control mechanisms such that debtors often could never work their way out of debt. Such arrangements as sharecropping and the use of the company store often tied laborers to employers through the bonds of debt. With some exceptions, states and legal regimes upheld property rights against the claims of the debtors.

Through the years, societies have sanctioned creditors' use of slavery, debt-prison, transportation to debtors' colonies, debt-peonage, seizure of assets or garnishment of wages to control debtors.

Most such methods, however, work to the advantage only of the first creditor or the most aggressive creditor to demand payment. Bankruptcy, as used in the United States and a number of other countries, provides a means for resolving not only the debtor-creditor conflict but also the potential conflict among creditors.

Bankruptcy is a very old concept. The word itself comes from an Italian phrase meaning "broken bench," because a bankrupt merchant's work bench would be broken by his creditors if he could not repay the debt. In the United States, each state has laws governing debtor-creditor relations, but the enactment of bankruptcy statutes is reserved to the Congress by the U.S. Constitution. United States bankruptcy occurs in the federal courts and is regulated by statutes enacted by Congress. Special bankruptcy courts are located in each federal judicial district, and specially appointed bankruptcy judges oversee the caseload. Bankruptcy decisions may be appealed to the federal district court and subsequently to the federal appeals court and the U.S. Supreme Court.

Bankruptcy is technically different from insolvency. Insolvency refers to a financial situation in which a person or business has liabilities that exceed assets. To be bankrupt, the debtor must file for bankruptcy protection in the federal court. Although the overwhelming majority of individuals filing for bankruptcy are also insolvent, occasionally a solvent business will file for bankruptcy because of anticipated liabilities that will exceed its assets. An example of such a bankruptcy is that of pharmaceutical manufacturer A. H. Robbins, whose

Dalkon Shield intrauterine device (a contraceptive) was judged the cause of many injuries and some deaths among women who used it. As the financial judgments against the company mounted, its management sought the protection of the bankruptcy courts. Bankruptcy allowed the company to hold its creditors at bay for a period of time to allow the company to propose a financial settlement.

### THE ROLE OF THE STATES

Although bankruptcy is a federal matter, states also have laws to govern debtor-creditor relations. Each state's debtor-creditor laws are affected by its history and by the debt conventions that are part of its history. Before becoming a state, for example, Georgia was a debtor's colony. Before the Civil War, debtors in such southern states often fled harsh debt-collection laws by going to Texas, indicated by "G.T.T." in sheriffs' records. Texas has traditionally retained pro-debtor statutory provisions, especially its generous exemption.

An exemption is the property that a debtor may keep despite bankruptcy or a judgment for nonpayment. The federal bankruptcy statutes recognize the right of state law to prescribe exemptions. There is also a federal exemption, but state legislatures may require their citizens to claim only the state exemption, which in some states is smaller than the federal exemption. State exemption laws vary widely, with some states allowing a substantial exemption and other states exempting very little property. Even a generous exemption law, however, is not a guarantee that a debtor will keep a lot of property. A home is not exempt if there is a mortgage on the home, and other goods are similarly not protected from a secured creditor if they have been used as collateral for a debt.

The states with a Spanish heritage often followed the Spanish tradition that a bankrupt's family should have the means to continue to make a living. Thus, the state laws of Florida, Texas, and California, for example, have traditionally been liberal in permitting debtors to keep their homesteads and some other assets, such as the tools of their trade and current wages. Other states exempt items that are believed necessary for the family's well being, such as children's school books, certain farm equipment, sewing machines, and

funeral plots. There is a recent trend toward substituting dollar limitations for exemptions instead of listing specific items of property.

The state exemption is the principal determinant of the resources a bankrupt debtor will have following the bankruptcy. The rest of the debtor's postbankruptcy status depends upon the type of bankruptcy the debtor declares.

### TYPES OF BANKRUPTCY

Bankruptcy may be entered either on a voluntary or on an involuntary basis. U.S. bankruptcy law arranges several ways by which debtors may voluntarily declare bankruptcy. Either individuals or corporate actors, including incorporated and unincorporated businesses, not-for-profit agencies, and municipalities may declare bankruptcy. The law makes special provision for the bankruptcies of railroads and stockbrokers. Creditors may in some circumstances initiate an involuntary bankruptcy. Only a small proportion of all bankruptcies is involuntary, and nearly all of those cases are targeted toward a business.

The Clerk of the Bankruptcy Court classifies each case filed as a business bankruptcy or a nonbusiness bankruptcy. These distinctions are not always clear-cut. As many as one in every five "nonbusiness" debtors reports currently owning a business or having recently owned a business. Moreover, many of the "business" bankruptcies are small family-owned enterprises. Whether classified as business or nonbusiness, the bankruptcy of a small business owner typically affects the family's welfare as well as that of the business.

Individual, noncorporate debtors typically have two choices in bankruptcy: *Chapter 7 liquidation* or a *Chapter 13 repayment plan*. In a Chapter 7 case, the debtor's assets over and above the exempt property will be sold and the creditors will be paid *pro rata*. Creditors with secured debts (those debts with collateral) will be allowed to have the collateral. All remaining debt will be discharged, or wiped away by the court. The debtor will not be able to file a Chapter 7 bankruptcy again for six years.

The Chapter 13 repayment plan is available only to individual debtors with a regular source of income. There are also other legal limitations,

such as maximum limits on the amount of debt owed. The debtor files with the court a plan for repaying debts over a three-to-five year schedule. The debtor must report all income to the court, and must also include a budget that provides the necessities of rent, food, clothing, medical treatment, and so on. The difference between the budgeted amount and the monthly income is the disposable income, which becomes the amount of the monthly payment.

The Chapter 13 petitioner must pay to a trustee, who is appointed by the court, all disposable income for the period of the plan. The debtor gets to keep all property. The trustee disburses the funds to the creditors. For a judge to confirm a Chapter 13 plan, the unsecured creditors must receive more money through the plan payments than they would have received in a Chapter 7 liquidation. At the conclusion of the plan, any remaining debt is discharged. Following this discharge, the Chapter 13 petitioner is also barred from further Chapter 7 bankruptcies for six years.

About two of every three Chapter 13 filers do not complete the proposed plan. Although the Chapter 13 may provide some benefits to the debtor—for example, he might have time to reorganize his finances—once plan payments are missed the court may dismiss the Chapter 13 and the debtor loses the protection of bankruptcy. Repayment plans such as Chapter 13 were begun by bankruptcy judges in northern Alabama before Congress wrote them into law. Even today, Chapter 13 is disproportionately popular in some judicial districts in the South. About one-third of all nonbusiness bankruptcies are filed in Chapter 13.

Some debts survive bankruptcy, regardless of the chapter in which the bankruptcy is filed. Child support, alimony, federally-backed educational loans, and some kinds of taxes are among those effectively nondischargeable in Chapter 7. Creditors may also object to the discharge of a specific debt if the debt arose from certain misbehavior, including fraud and drunkenness. Creditors may also object to the discharge generally on the grounds of debtor misbehavior, including hiding assets, disposing of assets before the bankruptcy, and lying to the court. A judge may object to a debtor filing in Chapter 7 if the judge believes that the filing represents a “substantial abuse.” A Chapter 13 plan must be filed in “good faith.”

For persons with substantial assets and for businesses there is an additional choice, *Chapter 11*, also called a *reorganization*. In Chapter 11, a debtor business seeks the protection of the court to reorganize itself in such a way that its creditors can be repaid (although perhaps not one hundred cents on the dollar). Creditors are given an opportunity to review the plan and to vote on its acceptability. If a sufficient number of creditors agree, the others may be forced to go along. In large cases, a committee of creditors is often appointed for the duration of the Chapter 11. Reorganizations may provide a means to save jobs while a business reorganizes.

In recent years, critics have charged that undoing business obligations has become an important motive for companies to reorganize under Chapter 11. Various companies have been accused of trying to avoid labor contracts, to avoid product liability, or to insulate management from challenges. These strategic uses of bankruptcy are potentially available to solvent companies. Although strategic uses of bankruptcies by individuals probably occur, most studies have found that the individuals in bankruptcy are in poor financial condition, often with debts in excess of three years' income.

There is also *Chapter 12* bankruptcy, another type of *reorganization* specifically for farmers. It was introduced in 1986 and in most years between one and two thousand cases are filed in Chapter 12.

Federal law permits the fact of a bankruptcy to remain on a credit record for ten years, but the law also prohibits discrimination against bankrupt debtors by governments or private employers.

## BANKRUPTCY TRENDS

The number of business bankruptcies remains relatively small, usually fewer than 75,000 in a year, with exceptions in a few years. The number of nonbusiness bankruptcies, however, has generally risen for about two decades, from about 313,000 in 1981 to 811,000 in 1991. Nonbusiness bankruptcies passed the millionmark in 1996, and by 1998 there were more than 1.4 million nonbusiness bankruptcies in the United States.

Embedded within the general increase in bankruptcy are substantial regional variations in filing

rates that seem to follow the business cycle. Some analysts believe that bankruptcy is a lagging indicator; that is, some months after an economic slowdown, bankruptcies begin to rise as laid-off workers run out of resources or small business owners find they cannot remain in business.

Each bankruptcy affects at least one household, and one estimate puts the average number of creditors affected at eighteen. Some creditors, such as credit card issuers, are affected by numerous bankruptcies within a single year. Some bankruptcies initiate a chain of events: The bankruptcy of a single business may lead later to the bankruptcies of employees who lost their jobs, of creditors whose accounts receivable were never received, and of suppliers who could not find new customers. Creditors often argue that the rising numbers of bankruptcies cause them costs that are then passed on to all consumers in the form of higher interest rates.

A bankruptcy may be filed either by a single individual or jointly by a couple. Since the early 1980s, there has been a sharp increase in the proportion of bankruptcies filed by adult women. Several studies indicate between one-third and one-half of bankruptcies are now filed by women. Over half of bankrupt debtors are between the ages of thirty and fifty. Their mean occupational prestige is approximately the same as that of the labor force as a whole, and their educational level is also similar to that of the adult population. The proportion of immigrants in the bankrupt population is approximately the same as their proportion in the general population. There are conflicting findings about the extent to which minority populations are overrepresented or underrepresented in bankruptcy. Despite their educational and occupational levels, however, the median incomes of bankrupt debtors are less than half the median income of the general population.

### CAUSES OF BANKRUPTCY

Much of what social scientists know about bankruptcy comes from reviewing the bankruptcy petitions filed in courts, and from interviewing judges, lawyers, clerks, and others who work in the bankruptcy courts. Information about the causes of bankruptcy, however, typically comes from interviews with the debtors themselves. Because many

debtors are ashamed of their bankruptcies, and because American debtor populations are geographically mobile, interview studies often have somewhat low response rates. For learning the cause of a bankruptcy, however, there is probably no substitute for asking the debtor.

Debtors often report “inability to manage money” or “too much debt” as the reason for their bankruptcy. In probing a little deeper, however, researchers have found five major issues involved in a large fraction of all bankruptcies: job problems, divorce and related family problems, medical problems, homeowner problems, and credit card debts.

Many debtors have been laid off completely, lost overtime, or had their hours of work or their pay rates reduced. Some debtors have had long periods of unemployment; others have lost their jobs because of the closure of the plant or the store where they worked. The job loss causes an unplanned loss of income, and this fact in turn often creates a mismatch between the petitioner’s debt obligations and the ability to meet those obligations. During the recession of the early 1990s, when many industries were downsizing and otherwise restructuring their labor forces, about two in every five bankrupt debtors reported a job-related reason for bankruptcy.

Divorce and other family problems may result in lower incomes for the two ex-spouses (especially the ex-wife). Moreover, in most divorce settlements the debts are also divided, and the debt burden may be too great for one of the ex-spouses. Sometimes an ex-spouse will file bankruptcy knowing that any jointly-incurred debts discharged by the bankrupt spouse will have to be paid by the other ex-spouse. Although alimony and child support cannot be discharged as debts, for custodial parents who do not receive the payments the financial consequences may be grave. Similarly, for parents who must make the payments, a reduction in other debts may make the payments more possible. Finally, it is harder to support two households on the income that used to support just one household. For all of these reasons, the recently divorced may find themselves in bankruptcy.

Medical problems may lead to higher debts if the debtor is uninsured or if insurance is insufficient to pay for needed medical procedures, pharmaceuticals, and professional services. Medical



problems may at the same time lead to lower incomes if a person is too ill or injured to work. Either way, a spell of illness or an automobile accident may push a previously solvent person into a situation in which debts are no longer manageable. Medical problems rise in frequency as a reported cause of bankruptcy for adults aged fifty-five to sixty-four, years in which medical problems may increase but before Medicare coverage becomes available.

Earlier efforts to examine medical debts as an indicator of the causes of bankruptcy showed medical debt to be a fairly insignificant cause, whereas direct interviews of debtors are more likely to reveal medical issues as causative. There are reasons to believe that the debt indicators underestimate this reason for bankruptcy. Insurance will often pay medical providers but will not replace income, so that debts to hospitals or physicians might not appear in the records even though the illness or injury is nevertheless the reason that the debtor cannot repay debts. Moreover, some medical providers accept credit cards, so that the medical expenses are hidden within credit card debt. And some debtors will make efforts to pay off their medical creditors first for fear of being denied services. On the other hand, there may be some overestimate of the impact of medical issues in interviews, because respondents may believe a medical reason may be seen as a socially acceptable cause for bankruptcy.

Bankrupt debtors are somewhat less likely to be homeowners than the general population, but a substantial number of homeowners file for bankruptcy, often to prevent the foreclosure of their mortgage. Chapter 13 permits homeowners to pay the arrearages (missed payments) on their mortgage along with their current payments. Home ownership may also be an issue for a worker who is transferred to a different city and buys a second home before the first home is sold. Equity is the portion of the home's value for which the homeowner has made payment. The recent proliferation of home-equity loans, which allow homeowners to use the equity in their homes as collateral, has also put homes at risk even if the payments on the principal mortgage are current. Small business owners are often required to use their homes as collateral for business loans, which means that a failing business may also entail the threat of the family losing its home. Home-equity lending is

being extended to a number of other situations, including college loans and credit cards, with the result that many Americans are risking their homes, often without realizing it.

A final reason for bankruptcy is large credit card debts, often at high rates of interest. All-purpose cards (such as Visa, MasterCard, and American Express) are now used for many consumption purposes, including payment of federal income taxes, payment of college tuition, and many goods and services, with the result that credit card debt is assumed for many different purposes. Most ordinary living expenses can now be charged with a credit card, so that interpreting a high credit card debt is difficult. High rates of interest accelerate the credit card debt quickly. Credit card debt is the fastest-growing reason for bankruptcy being given by debtors.

### **BANKRUPTCY MYTHS**

Empirical studies have refuted a number of myths about bankruptcy. One such myth is that some people file for bankruptcy again and again. Several studies have been unable to confirm this myth, but there are a number of people who are unsuccessful at Chapter 13 who eventually file Chapter 7. These people have not really repeated bankruptcy, because they have never received a discharge from their debts in their Chapter 13 filings.

Another myth is that large numbers of bankrupt debtors have high debts because of alcoholism, drug abuse, and gambling. Although most empirical studies have identified a few isolated cases in which one of these problems plays a role, the great increase in bankruptcy numbers cannot be attributed to a great increase in addictive behavior.

A third myth is that there is widespread abuse of the bankruptcy process because many debtors could allegedly pay all of their debts. Most studies of the repayment possibilities for the debtors find that the debtors are unable to repay their debts, even if their families live on very modest budgets (such as the model low-income budget of the U.S. Bureau of Labor Statistics). The few studies that have found debtors able to repay have often eliminated from repayment whole categories of debt that the debtors themselves cannot eliminate. There are, however, documented cases of fraudulent or

abusive use of bankruptcy, some of which are prosecuted by the government as criminal matters.

### STRUCTURAL SOURCES OF BANKRUPTCY

Several changes in American financial life have contributed to the imbalance of debt with income and may have increased the numbers of people who file for bankruptcy.

Consumer credit has been available for decades, principally as a means to help households finance a large purchase over a period of time. Individual sellers extended credit, often on the basis of personal knowledge of the borrower and the borrower's ability to repay. Later, banks, credit unions, and personal finance companies helped finance the purchase of homes, cars, and small appliances. These institutional arrangements offered more protection to the seller and made a personal acquaintanceship less important in the lending decision.

The development of credit cards, with preapproved credit limits, allowed consumers to buy a large variety of goods or services on credit, not just large or expensive items. It was not just the invention of the credit card, however, but some later developments in its use that have changed consumer financial patterns in a way that may have influenced higher bankruptcy rates.

First, beginning in the late 1970s, most states repealed their laws making usury an offense, and Congress made state usury laws largely ineffective. Usury, the charging of excessive interest, was formerly regulated by the states, most of which set limits on the maximum amount of interest that could be charged to borrowers. With the repeal of these laws, high rates of interest—some as high as 20 percent or more—could legally be charged. Prior to this time, such high rates of interest were usually defined as criminal and were associated with loan-sharking and the lending practices of organized crime.

Second, credit cards became a major profit center for many banks, especially because of the high interest rates that could be charged. Marketing of credit cards mushroomed. New markets, including relatively low-income families, became the targets of sophisticated mail and telephone campaigns. Even though many of the families might not be able to repay everything they charged,

the large interest payments made the lending profitable. Credit cards are now extensively marketed to young people, especially college students, who are relatively unfamiliar with financial matters.

Third, credit devices proliferated to include gold and platinum levels, cash advances, and many other features that were both profitable to card issuers and attractive to card holders. Sometimes card holders did not understand how these features worked; for example, many card holders did not realize that there is no grace period for repaying a cash advance, and the interest rate on a cash advance is often higher than the rate for purchases.

Fourth, merchants who accepted credit cards enjoyed increased business because the buyers did not have to carry a large supply of cash. Advertising by stores and restaurants increasingly emphasized the acceptability of credit cards, so that the advertising for goods and services reinforced the advertising of the credit card issuers. Meanwhile, the credit card issuers provided protection against nonpayment to the merchant who accepted a credit card, and they also provided some protection for the card-holder against the fraudulent use of the cards.

Fifth, a major restructuring of the U.S. economy occurred at the beginning of the 1990s, in which millions of workers were laid off or could find only contingent jobs. Many of these workers found credit cards to be a convenient way to maintain consumption levels and their family's lifestyle even though their expected income stream was interrupted.

These trends contributed to an increase in the numbers of people who had high debt-to-income ratios. Some people incurred high levels of debt, often at high interest rates, and simultaneously experienced declining or stagnant income. While not every person with a high debt-to-income ratio filed for bankruptcy, those who did file for bankruptcy had very high ratios. Changes in the American economy during the decade of the 1990s probably influenced the increase in the number of bankruptcies. In particular, the increasing indebtedness of Americans closely tracked the rise in bankruptcies.

**Rising Indebtedness.** The economic changes of the 1990s, in addition to the well-entrenched borrowing for home mortgages, car purchases,

and college expenses, have resulted in an increase in consumer credit outstanding from 298 billion dollars in 1980 to 1,025 billion in 1995. Thus, while the U.S. population increased by 16 percent, indebtedness increased by 244 percent. These are nominal dollars, but even accounting for inflation (constant dollars) the debt burden doubled. During this same time frame, mortgage lending increased by 223 percent, from 1,463 billion dollars in 1980 to 4,724 billion dollars in 1995. Credit card debt, which was only 81.2 billion dollars in 1980, had increased by 351 percent to 366.4 billion dollars in 1995. Thus, while there was an increase in all forms of indebtedness, credit card debt—which began from a smaller base—showed the largest percentage increase.

Industry sources estimate that in 1990 there were 1.03 billion credit cards in circulation in the United States, and that by 2000 a projected 1.34 billion cards will be in circulation. In 1990, these cards accounted for 466 billion dollars of spending and resulted in 236 billion dollars of debt. By 2000, the projected spending is 1,443 billion dollars, with 661 billion dollars as debt.

In its survey of how consumers use credit cards, the Board of Governors of the Federal Reserve System reported that 54.5 percent of the sample always paid off the balance on their credit cards each month. Another 19.1 percent sometimes pay off the balance, and the remaining 26.4 percent hardly ever pay off the balance. It is from this latter group that the credit card issuers run the greatest risk of eventual nonpayment but also have the possibility of earning the greatest amount of interest. About 30 percent of the consumers earning less than \$50,000 a year “hardly ever” pay off the balance, compared with only 10.5 percent of those who earn more than \$100,000. Over 35 percent of the consumers under the age of thirty-five hardly ever pay off their balance, compared with 10.5 percent of people over the age of seventy-five.

Credit cards alone do not pose a major financial problem for most people, even for those who maintain a balance; the median balance is about \$1,000, with a median charge of \$200 a month. For those people with larger balances, however, and with interest charges and sometimes penalties and late fees added to the principal, indebtedness may

become a serious problem. And although facing this problem can be postponed by making small monthly payments, the balance can quickly become too large ever to be handled on most salaries.

The ubiquity and convenience of credit cards has led to their growing use for additional purposes: as a form of identification, as security for returning rented cars and videotapes, and—perhaps ironically—as an indicator of creditworthiness for additional extensions of credit.

### INTERNATIONAL CREDIT ISSUES

Besides the indebtedness of the individuals in a population, there is also growing concern about other forms of indebtedness. Many countries, especially those with less developed economies, have borrowed large sums of money from more developed countries. The repayment of these loans, and the terms that are demanded, have provided a source of tension between the richer nations and the poorer nations. International agencies such as the International Monetary Fund often prescribe austerity measures to help a country meet its repayment obligations.

The development of transnational corporations has also raised issues of which set of debtor-creditor laws govern transactions that may span several countries. Debtor-creditor laws and the laws of insolvency and bankruptcy vary dramatically from country to country, and an important issue in international commerce is how to handle problems of nonpayment. Many of these issues remain highly contested and largely unsettled.

#### REFERENCES

- Caplovitz, David 1974 *Consumers In Trouble: A Study of Debtors in Default*. New York: Free Press.
- Delaney, Kevin J. 1992 *Strategic Bankruptcy: How Corporations and Creditors Use Chapter 11 to Their Advantage*. Berkeley: University of California Press.
- Jacob, Herbert 1969 *Debtors in Court; The Consumption of Government Services*. Chicago: Rand McNally.
- Ritzer, George 1995 *Expressing America: A Critique of the Global Credit Card Society*. Thousand Oaks, Calif.: Pine Forge Press.
- Ryan, Martin 1995 *The Last Resort: A Study of Consumer Bankrupts*. Aldershot, Eng.: Avebury.

- Stanley, David T., and Marjorie Girth 1971 *Bankruptcy: Problem, Process, Reform*. Washington, D.C.: The Brookings Institution.
- Sullivan, Teresa A., Elizabeth Warren, and Jay Lawrence Westbrook 2000 *The Fragile Middle Class*. New Haven, Conn.: Yale University Press.
- 1999 *As We Forgive Our Debtors: Bankruptcy and Consumer Credit in America*. Reprint. Chevy Chase, Md.: Beard Books.
- 1995 “Bankruptcy and the Family.” *Marriage and Family Review* 21, 3/4 (1995):193–215.
- 1994 “Consumer Debtors Ten Years Later: A Financial Comparison of Consumer Bankrupts 1981–1991.” *American Bankruptcy Law Journal* 68,2 (Spring):121–154.
- 1994 “The Persistence of Local Legal Culture: Twenty Years of Evidence from the Federal Bankruptcy Courts.” *Harvard Journal of Law and Public Policy* 17,3 (Summer):801–865.

TERESA A. SULLIVAN

## BEHAVIORISM

Behaviorism is the conceptual framework underlying the science of behavior. The science itself is often referred to as the *experimental analysis of behavior or behavior analysis*. Modern behaviorism emphasizes the analysis of conditions that maintain and change behavior as well as the factors that influence the acquisition or learning of behavior. Behaviorists also offer concepts and analyses that go well beyond the common-sense understanding of reward and punishment. Contemporary behaviorism provides an integrated framework for the study of human behavior, society, and culture.

Within the social sciences, behaviorism has referred to the social-learning perspective that emphasizes the importance of reinforcement principles in regulating social behavior (McLaughlin 1971). In addition, sociologists such as George Homans and Richard Emerson have incorporated the principles of behavior into their theories of elementary social interaction or exchange (Emerson 1972; Homans 1961). The basic idea in social exchange approaches is that humans exchange valued activities (e.g., giving respect and getting help) and that these transactions are “held together” by the principle of reinforcement. That is,

exchange transactions that involve reciprocal reinforcement by the partners increase in frequency or probability; those transactions that are not mutually reinforcing or are costly to the partners decrease in frequency over time. There is a growing body of research literature supporting social exchange theory as a way of understanding a variety of social relationships.

### SOME BASIC ISSUES

The roots of behaviorism lie in its philosophical debate with introspectionism—the belief that the mind can be revealed from a person’s reports of thoughts, feelings, and perceptions. Behaviorists opposed the use of introspective reports as the basic data of psychology. These researchers argued for a natural-science approach and showed how introspective reports of consciousness were inadequate. Reports of internal states and experiences were said to lack objectivity, were available to only one observer, and were prone to error. Some behaviorists used these arguments and also others to reject cognitive explanations of behavior (Skinner 1974; Pierce and Epling 1984; but see Bandura 1986 for an alternative view).

The natural-science approach of behaviorism emphasizes the search for general laws and principles of behavior. For example, the *quantitative law of effect* is a mathematical statement of how the rate of response increases with the rate of reinforcement (Herrnstein 1970). Under controlled conditions, this equation allows the scientist to predict precisely and to regulate the behavior of organisms. Behavior analysts suggest that this law and other behavior principles will eventually account for complex human behavior (McDowell 1988).

Contemporary behaviorists usually restrict themselves to the study of observable responses and events. Observable events are those that are directly sensed or are made available to our senses by instruments. The general strategy is to manipulate aspects of the environment and measure well-defined responses. If behavior reliably changes with a manipulated condition, the researcher has established an environment-behavior relationship. Analysis of such relationships has often resulted in behavioral laws and principles. For example, the principle of discrimination states that an organism

will respond differently to two situations if its behavior is reinforced in one setting but not in the other. You may talk about politics to one person but not to another, because the first person has been interested in such conversation in the past while the second has not. The principle of discrimination and other behavior principles account for many aspects of human behavior.

Although behaviorism usually has been treated as a uniform and consistent philosophy and science, a conceptual reconstruction indicates that there are many branches to the behavioral tree (Zuriff 1985). Most behavior analysts share a set of core assumptions; however, there are internal disputes over less central issues. To illustrate, some behaviorists argue against hypothetical constructs (e.g., memory) while others accept such concepts as an important part of theory construction.

Throughout the intellectual history of behaviorism, a variety of assumptions and concepts has been presented to the scientific community. Some of these ideas have flourished when they were found to further the scientific analysis of behavior. Other formulations were interesting variations of behavioristic ideas, but they became extinct when they served no useful function. For instance, one productive assumption is that a person's knowledge of emotional states is due to a special history of verbal conditioning (Bem 1965, 1972; Skinner 1957). Self-perception and attributional approaches to social psychology have built on this assumption, although researchers in this field seldom acknowledge the impact. In contrast, the assumption that thinking is merely subvocal speech was popular at one time but is now replaced by an analysis of problem solving (Skinner 1953, 1969). In this view, thinking is behavior that precedes and guides the final performance of finding a solution. Generally, it is important to recognize that behaviorism continues to evolve as a philosophy of science, a view of human nature, and an ideology that recommends goals for behavioral science and its applications.

### THE STUDY OF BEHAVIOR

Behaviorism requires that a scientist study the behavior of organisms for its own sake. Behaviorists do not study behavior in order to make inferences

about mental states or physiological processes. Although most behaviorists emphasize the importance of biology and physiological processes, they focus on the interplay of behavior and environment.

In order to maintain this focus, behaviorists examine the evolutionary history and physiological status of an organism as part of the context for specific environment-behavior interactions. For example, a biological condition that results in blindness may have profound behavioral effects. For a newly sightless individual, visual events, such as watching television or going to a movie no longer support specific interactions, while other sensory events become salient (e.g., reading by braille). The biological condition limits certain kinds of behavioral interactions and, at the same time, augments the regulation of behavior by other aspects of the environment. Contemporary behaviorism therefore emphasizes what organisms are doing, the environmental conditions that regulate their actions, and how biology and evolution constrain or enhance environment-behavior interactions.

Modern behaviorists are interested in voluntary action, and they have developed a way of talking about purpose, volition, and intention within a natural-science approach. They note that the language of intention was pervasive in biology before Darwin's functional analysis of evolution. Although it appears that giraffes grow long necks in order to obtain food at the tops of trees, Darwin made it clear that the process of evolution involved no plan, strategy of design, or purpose. Natural variation ensures that giraffes vary in neck size. As vegetation declines at lower heights, animals with longer necks obtain food, survive to adulthood, and reproduce; those with shorter necks starve to death. In this environment (niche), the frequency of long-necked giraffes increases over generations. Such an increase is called natural selection. Contemporary behaviorists insist that selection, as a causal mode, also accounts for the form and frequency of behavior during the lifetime of an individual. A person's current behavior is therefore composed of performances that have been selected in the past (Skinner 1987).

An important class of behavior is selected by its consequences. The term *operant* refers to behavior that operates upon the environment to produce effects, outcomes, or consequences. Operant

behavior is said to be emitted because it does not depend on an eliciting stimulus. Examples of operant behavior include manipulation of objects, talking with others, problem solving, drawing, reading, writing, and many other performances. Consequences select this behavior in the sense that specific operants occur at high frequency in a given setting. To illustrate, driving to the store is operant behavior that is likely to occur when there is little food in the house. In this situation, the operant has a high probability if such behavior has previously resulted in obtaining food (i.e. the store is open). Similarly, the conversation of a person also is selected by its social consequences. At the pub, a student shows high probability of talking to his friends about sports. Presumably, this behavior occurs at high frequency because his friends have previously "shown an interest" in such conversation. The behavior of an individual is therefore adapted to a particular setting by its history of consequences.

A specific operant, such as opening a door, includes many performance variations. The door may be opened by turning the handle, pushing with a foot, or even by asking someone to open it. These variations in performance have a common effect upon the environment in the sense that each one results in the door being opened. Because each variation produces similar consequences, behaviorists talk about an operant as a response class. Operants such as opening a door, talking to others, answering questions, and many other actions are each a response class that includes a multitude of forms, both verbal and nonverbal.

In the laboratory, the study of operant behavior requires a basic measure that is sensitive to changes in the environment. Most behaviorists use an operant's rate of occurrence as the basic data for analysis. Operant rate is measured as the frequency of an operant (class) over a specified period of time. Although operant rate is not directly observable, a cumulative recorder is an instrument that shows the rate of occurrence as changes in the slope (or rise) of a line on moving paper. When an operant is selected by its consequences, the operant rate increases and the slope becomes steeper. Operants that are not appropriate to the requirements of the environment decrease in rate of occurrence (i.e., decline in slope). Changes in operant rate therefore reflect the basic causal process of *selection by consequences* (Skinner 1969).

Behavior analysts continue to use the cumulative recorder to provide an immediate report on a subject's behavior in an experimental situation. However, most researchers are interested in complex settings where there are many alternatives and multiple operants. Today, microcomputers collect and record a variety of behavioral measures that are later examined by complex numerical analysis. Researchers also use computers to arrange environmental events for individual behavior and provide these events in complex patterns and sequences.

### CONTINGENCIES OF REINFORCEMENT

Behaviorists often focus on the analysis of environment-behavior relationships. The relationship between operant behavior and its consequences defines a *contingency of reinforcement*. In its simplest form, a two-term contingency of reinforcement may be shown as  $R(Sr)$ . The symbol  $R$  represents the operant class, and  $Sr$  stands for the reinforcing stimulus or event. The arrow indicates that "if  $R$  occurs, then  $Sr$  will follow." In the laboratory, the behavior analyst arranges the environment so that a contingency exists between an operant (e.g., pecking a key) and the occurrence of some event (e.g., presentation of food). If the presentation of the event increases operant behavior, the event is defined as a positive reinforcer. The procedure of repeatedly presenting a positive reinforcer contingent on behavior is called positive reinforcement (see Pierce and Epling 1999).

A contingency of reinforcement defines the probability that a reinforcing event will follow operant behavior. When a person turns the ignition key of the car (operant), this behavior usually has resulted in the car starting (reinforcement). Turning the key does not guarantee, however, that the car will start; perhaps it is out of gas, the battery is run down, and so on. Thus, the probability of reinforcement is high for this behavior, but reinforcement is not certain. The behavior analyst is interested in how the probability of reinforcement is related to the rate and form of operant behavior. For example, does the person continue to turn the ignition key even though the car doesn't start? Qualities of behavior such as persistence, depression, and elation reflect the probability of reinforcement.

Reinforcement may depend on the number of responses or the passage of time. A schedule of reinforcement is a procedure that states how consequences are arranged for behavior. When reinforcement is delivered after each response, a continuous schedule of reinforcement is in effect. A child who receives payment each time she mows the lawn is on a continuous schedule of reinforcement. Continuous reinforcement produces a very high and steady rate of response, but as any parent knows, the behavior quickly stops if reinforcement no longer occurs.

Continuous reinforcement is a particular form of ratio schedule. Fixed-ratio schedules state the number of responses per reinforcement. These schedules are called *fixed ratio* since a fixed number of responses are required for reinforcement. In a factory, piece rates of payment are examples of fixed-ratio schedules. Thus, a worker may receive \$1 for sewing twenty pieces of elastic wristband. When the ratio of responses to reinforcement is high (value per unit output is low), fixed-ratio schedules produce long pauses following reinforcement: Overall productivity may be low, leading plant managers to complain about “slacking off” by the workers. The problem, however, is the schedule of reinforcement that fixes a high number of responses to payment.

Reinforcement may be arranged on a variable, rather than fixed, basis. The schedule of payoff for a slot machine is a *variable-ratio* schedule of reinforcement. The operant involves putting in a dollar and pulling the handle, and reinforcement is the jackpot. The jackpot occurs after a variable number of responses. Variable-ratio schedules produce a high rate of response that takes a long time to stop when reinforcement is withdrawn. The gambler may continue to put money in the machine even though the jackpot rarely, if ever, occurs. Behavior on a variable-ratio schedule is said to show *negative utility* since people often invest more than they get back.

Behavior may also be reinforced only after an interval of time has passed. A *fixed-interval* schedule stipulates that the first response following a specified interval is reinforced. Looking for a bus is behavior that is reinforced after a fixed time set by the bus schedule. If you just missed a bus, the probability of looking for the next one is quite low. As time passes, the rate of response increases with

the highest rate occurring just before the bus arrives. Thus, the rate of response is initially zero but gradually rises to a peak at the moment of reinforcement. This response pattern is called *scalloping* and is characteristic of fixed-interval reinforcement. In order to eliminate such patterning, a *variable-interval* schedule may be stipulated. In this case, the first response after a variable amount of time is reinforced. If a person knows by experience that bus arrivals are irregular, looking for the next bus will occur at a moderate and steady rate because the passage of time no longer signals reinforcement (i.e., arrival of the bus).

The schedules of reinforcement that regulate human behavior are complex combinations of ratio and interval contingencies. An *adjusting* schedule is one example of a more complex arrangement between behavior and its consequences (Zeiler 1977). When the ratio (or interval) for reinforcement changes on the basis of performance, the schedule is called adjusting. A math teacher who spends more or less time with a student depending on the student’s competence (i.e., number of correct solutions) provides reinforcement on an adjusting-ratio basis. When reinforcement is arranged by other people (i.e., social reinforcement), the level of reinforcement is often tied to the level of behavior (i.e., the greater the strength of response the less the reward from others). This adjustment between behavior and socially arranged consequences may account for the flexibility and variability that characterize adult human behavior.

Human behavior is regulated not only by its consequences. Contingencies of reinforcement also involve the events that precede operant behavior. The preceding event is said to “set the occasion” for behavior and is called a discriminative stimulus or *Sd*. The ring of a telephone (*Sd*) may set the occasion for answering it (operant), although the ring does not force one to do so. Similarly, a nudge under the table (*Sd*) may prompt a new topic of conversation (operant) or cause the person to stop speaking. Discriminative stimuli may be private as well as public events. Thus, a headache may result in taking a pill or calling a physician. A mild headache may be discriminative stimulus for taking an aspirin, while more severe pain sets the occasion for telephoning a doctor.

Although discriminative stimuli exert a broad range of influences over human behavior, these

events do not stand alone. These stimuli regulate behavior because they are an important part of the contingencies of reinforcement. Behaviorism has therefore emphasized a three-term contingency of reinforcement, symbolized as  $Sd:R(r)Sr$ . The notation states that a specific event ( $Sd$ ) sets the occasion for an operant ( $R$ ) that produces reinforcement ( $Sr$ ). The discriminative stimulus regulates behavior only because it signals past consequences. Thus, a sign that states “Eat at Joe’s” may set the occasion for your stopping at Joe’s restaurant because of the great meals received in the past. If Joe hires a new cook, and the meals deteriorate in quality, then Joe’s sign will gradually lose its influence. Similarly, posted highway speeds regulate driving on the basis of past consequences. The driver who has been caught by a radar trap is more likely to observe the speed limit.

### CONTEXT OF BEHAVIOR

Contingencies of reinforcement, as complex arrangements of discriminative stimuli, operants, and reinforcements, remain a central focus of behavioral research. Contemporary behaviorists are also concerned with the *context of behavior*, and how context affects the regulation of behavior by its consequences (Fantino and Logan 1979). Important aspects of context include the biological and cultural history of an organism, its current physiological status, previous environment–behavior interactions, alternative sources of reinforcement, and a history of deprivation (or satiation) for specific events or stimuli. To illustrate, in the laboratory food is used typically as an effective reinforcer for operant behavior. There are obvious times, however, when food will not function as reinforcement. If a person (or animal) has just eaten a large meal or has an upset stomach, food has little effect upon behavior.

There are less obvious interrelations between reinforcement and context. Recent research indicates that depriving an organism of one reinforcer may increase the effectiveness of a different behavioral consequence. As deprivation for food increased, animals worked harder to obtain an opportunity to run on a wheel. Additionally, animals who were satiated on wheel running no longer pressed a lever to obtain food. These results imply that eating and running are biologically interrelated. Based on this biological history, the supply or

availability of one of these reinforcers alters the effectiveness of the other (Pierce, Epling, and Boer 1986). It is possible that many reinforcers are biologically interrelated. People commonly believe that sex and aggression go together in some unspecified manner. One possibility is that the availability of sexual reinforcement alters the reinforcing effectiveness of an opportunity to inflict harm on others.

### CHOICE AND PREFERENCE

The emphasis on context and reinforcement contingencies has allowed modern behaviorists to explore many aspects of behavior that seem to defy a scientific analysis. Most people believe that choice and preference are basic features of human nature. Our customary way of speaking implies that people make decisions on the basis of their knowledge and dispositions. In contrast, behavioral studies of decision making suggest that we choose an option based on its rate of return compared with alternative sources of reinforcement.

Behaviorists have spent the last thirty years studying choice in the laboratory using concurrent schedules of reinforcement. The word *concurrent* means “operating at the same time.” Thus, concurrent schedules are two (or more) schedules operating at the same time, each schedule providing reinforcement independently. The experimental setting is arranged so that an organism is free to alternate between two or more alternatives. Each alternative provides a schedule of reinforcement for choosing it over the other possibilities. A person may choose between two (or more) response buttons that have different rates of monetary payoff. Although the experimental setting is abstract, concurrent schedules of reinforcement provide an analogue of choice in everyday life.

People are often faced with a variety of alternatives, and each alternative has its associated benefits (and costs). When a person puts money in the bank rather than spending it on a new car, television, or refrigerator, we speak of the individual choosing to save rather than spend. In everyday life, choice often involves repeated selection of one alternative (e.g. putting money in the bank) over the other alternatives considered as a single option (e.g. buying goods and services). Similarly, the criminal chooses to take the property of others rather than take the socially acceptable



route of working for a living or accepting social assistance. The arrangement of consequences for crime and legitimate ways of making a living is conceptually the same as concurrent schedules of reinforcement (Hamblin and Crosbie 1977).

Behaviorists are interested in the distribution or allocation of behavior when a person is faced with different rates of reinforcement from two (or more) alternatives. The distribution of behavior is measured as the relative rate of response to, or relative time spent on, a specific option. For example, a student may go to school twelve days and skip eight days each month (not counting weekends). The relative rate of response to school is the proportion of the number of days at school to the total number of days, or  $12/20 = 0.60$ . Expressed as a percentage, the student allocates 60 percent of her behavior to school. In the laboratory, a person may press the left button twelve times and the right button eight times each minute.

The distribution of reinforcement may also be expressed as a percentage. In everyday life, it is difficult to identify and quantify behavioral consequences, but it is easily accomplished in the laboratory. If the reinforcement schedule on the left button produces \$30 an hour and the right button yields \$20, 60 percent of the reinforcements are on the left. There is a fundamental relationship between relative rate of reinforcement and relative rate of response. This relationship is called the *matching law*. The law states that the distribution of behavior to two (or more) alternatives matches (equals) the distribution of reinforcement from these alternatives (Herrnstein 1961; de Villiers 1977).

Although it is difficult to identify rates of reinforcement for attending school and skipping, the matching law does suggest some practical solutions (Epling and Pierce 1988). For instance, parents and the school may be able to arrange positive consequences when a child goes to school. This means that the rate of reinforcement for going to school has increased, and therefore the relative rate of reinforcement for school has gone up. According to the matching law, a child will now distribute more behavior to the school.

Unfortunately, there is another possibility. A child may receive social reinforcement from friends for skipping, and as the child begins to spend more time at school, friends may increase their rate of reinforcement for cutting classes. Even though the

absolute rate of reinforcement for going to school has increased, the relative rate of reinforcement has remained the same or decreased. The overall effect may be no change in school attendance or even further decline. In order to deal with the problem, the matching law implies that interventions must increase reinforcement for attendance and maintain or reduce reinforcement for skipping, possibly by turning up the cost of this behavior (e.g., withdrawal of privileges).

The matching law has been tested with human and nonhuman subjects under controlled conditions. One interesting study assessed human performance in group discussion sessions. Subjects were assigned to groups discussing attitudes toward drug abuse. Each group was composed of three confederates and a subject. Two confederates acted as listeners and reinforced the subject's talk with brief positive words and phrases, provided on the basis of cue lights. Thus, the rate of reinforcement by each listener could be varied depending on the number of signals arranged by the researchers. A third confederate asked questions but did not reinforce talking. Results were analyzed in terms of the relative time subjects spent talking to the two listeners. Speakers matched their distribution of conversation to the distribution of positive comments from the listeners. Apparently, choosing to speak to others is behavior that is regulated by the matching law (Conger and Kileen 1974).

Researchers have found that exact matching does not always hold between relative rate of reinforcement and relative rate of response. A more general theory of behavioral matching has been tested in order to account for the departures from perfect matching. One source of deviation is called *response bias*. Bias is a systematic preference for an alternative, but the preference is not due to the difference in rate of reinforcement. For example, even though two friends provide similar rates of reinforcement, social characteristics (e.g., status and equity) may affect the distribution of behavior (Sunahara and Pierce 1982). Generalized matching theory is able to address many social factors as sources of bias that affect human choice and preference (Baum 1974; Pierce and Epling 1983; Bradshaw and Szabadi 1988).

A second source of deviation from matching is called *sensitivity to differences* in reinforcement.

Matching implies that an increase of 10 percent (e.g., from 50 to 60 percent) in relative rate of reinforcement for one alternative results in a similar increase in relative rate of response. In many cases, the increase in relative rate of response is less than expected (e.g., only 5 percent). This failure to discriminate changes in relative rate of reinforcement is incorporated within the theory of generalized matching. To illustrate, low sensitivity to changes in rate of reinforcement may occur when an air-traffic controller rapidly switches between two (or more) radar screens. As relative rate of targets increases on one screen, relative rate of detection may be slow to change. Generalized matching theory allows behaviorists to measure the degree of sensitivity and suggests procedures to modify it (e.g., setting a minimal amount of time on a screen before targets can be detected).

Matching theory is an important contribution of modern behaviorism. In contrast to theories of rational choice proposed by economists and other social scientists, matching theory implies that humans may not try to maximize utility (or reinforcement). People (and animals) do not search for the strategy that yields the greatest overall returns; they equalize their behavior to the obtained rates of reinforcement from alternatives. Research suggests that matching (rather than maximizing) occurs because humans focus on the immediate effectiveness of their behavior. A person may receive a per-hour average of \$10 and \$5 respectively from the left and right handles of a slot machine. Although the left side generally pays twice as much, there are local periods when the left option actually pays less than the right. People respond to these changes in *local rate of reinforcement* by switching to the lean alternative (i.e., the right handle), even though they lose money overall. The general implication is that human impulsiveness ensures that choice is not a rational process of getting the most in the long run but a behavioral process of doing the best at the moment (Herrnstein 1990).

#### MATHEMATICS AND BEHAVIOR MODIFICATION

The matching law suggests that operant behavior is determined by the rate of reinforcement for one alternative relative to all other sources of reinforcement. Even in situations that involve a single

response on a schedule of reinforcement, the behavior of organisms is regulated by alternative sources of reinforcement. A rat that is pressing a lever for food may gain additional reinforcement from exploring the operant chamber, scratching itself, and so on. In a similar fashion, rather than work for teacher attention a pupil may look out the window, talk to a friend, or even daydream. Thus in a single-operant setting, multiple sources of reinforcement are functioning. Herrnstein (1970) argued this point and suggested an equation for the single operant that is now called the *quantitative law of effect*.

Carr and McDowell (1980) applied Herrnstein's equation to a clinically relevant problem. The case involved the treatment of a 10-year-old boy who repeatedly and severely scratched himself. Before treatment the boy had a large number of open sores on his scalp, face, back, arms, and legs. In addition, the boy's body was covered with scabs, scars, and skin discoloration. In their review of this case, Carr and McDowell demonstrated that the boy's scratching was operant behavior. Careful observation showed that the scratching occurred more often when he and other family members were in the living room watching television. This suggested that a specific situation set the occasion for the self-injurious behavior. Further observation showed that family members repeatedly and reliably reprimanded the boy when he engaged in self-injury. Reprimands are seemingly negative events, but adult attention (whether negative or positive) can serve as reinforcement for children's behavior.

In fact, McDowell (1981) showed that the boy's scratching was in accord with Herrnstein's equation (i.e., the quantitative law of effect). He plotted the reprimands per hour on the *x*-axis and scratches per hour on the *y*-axis. When applied to this data, the equation provided an excellent description of the boy's behavior. The quantitative law of effect also suggested how to modify the problem behavior. In order to reduce scratching (or any other problem behavior), one strategy is to increase reinforcement for alternative behavior. As reinforcement is added for alternative behavior, problem behavior must decrease; this is because the reinforcement for problem behavior is decreasing (relative to total reinforcement) as reinforcement is added to other (acceptable) behavior.

### APPLIED BEHAVIOR ANALYSIS AND EDUCATION

The application of behavior principles to improve performance and solve social problems is called *applied behavior analysis* (Baer, Wolf, and Risley 1968). Principles of behavior change have been used to improve the performance of university students, increase academic skills in public and high school students, teach self-care to developmentally delayed children, reduce phobic reactions, get people to wear seat belts, prevent industrial accidents, and help individuals stop cocaine abuse, among other things. Behavioral interventions have had an impact on such things as clinical psychology, medicine, education, business, counseling, job effectiveness, sports training, the care and treatment of animals, environmental protection, and so on. Applied behavioral experiments have ranged from investigating the behavior of psychotic individuals to designing contingencies of entire institutions (see Catania and Brigham 1978; Kazdin 1994).

One example of applied behavior analysis in higher education is the method of personalized instruction. Personalized instruction is a self-paced learning system that contrasts with traditional lecture methods that often are used to instruct college students. In a university lecture, a professor stands in front of a number of students and talks about his or her area of expertise. There are variations on this theme (e.g., students are encouraged to be active rather than passive learners), basically the lecture method of teaching is the same as it has been for thousands of years.

Dr. Fred Keller (1968) recognized that the lecture method of teaching was inefficient and in many cases a failure. He reasoned that anyone who had acquired the skills needed to attend college was capable of successfully mastering most or all college courses. Some students might take longer than others to reach expertise in a course, but the overwhelming majority of students would be able to do so. If behavior principles were to be taken seriously, there were no bad students, only ineffective teaching methods.

In a seminal article, titled "Good-bye, teacher. . .," Keller outlined a college teaching method based on principles of operant conditioning (Keller 1968). Keller's personalized system of instruction (PSI) involves arranging the course material in a sequence of graduated steps (units or modules).

Each student moves through the course material at his or her own pace and the modules are set up to ensure that most students have a high rate of success learning the course content. Some students may finish the course in a few weeks, others require a semester or longer.

Course material is broken down into many small units of reading and (if required) laboratory assignments. Students earn points (conditioned reinforcement) for completing unit tests and lab assignments. Mastery of lab assignments and unit tests is required. If test scores are not close to perfect, the test (in different format) must be taken again after a review of the material for that unit. The assignments and tests build on one another so they must be completed in a specified order.

Comparison studies have evaluated student performance on PSI courses against the performance of students given computer-based instruction, audio-tutorial methods, traditional lectures, visual-based instruction, and other programmed instruction methods. College students instructed by PSI outperform students taught by these other methods when given a common final examination (see Lloyd and Lloyd 1992 for a review). Despite this positive outcome, logistical problems in organizing PSI courses such as teaching to mastery level (most students get an A for the course), and allowing students more time than the allotted semester to complete the course, have worked against widespread adoption of PSI in universities and colleges.

### SUMMARY

Modern behaviorism emphasizes the context of behavior and reinforcement. The biological history of an organism favors or constrains specific environment-behavior interactions. This interplay of biology and behavior is a central focus of behavioral research. Another aspect of context concerns alternative sources of reinforcement. An individual selects a specific option based on the relative rate of reinforcement. This means that behavior is regulated not only by its consequences but also by the consequences arranged for alternative actions. As we have seen, the matching law and the quantitative law of effect are major areas of basic research that suggest new intervention strategies for behavior modification. Finally, applied behavior

analysis, as a technology of behavior change, is having a wide impact on socially important problems of human behavior. A personalized system of instruction is an example of applied behavior analysis in higher education. Research shows that mastery-based learning is more effective than alternative methods of instruction, but colleges and universities its implementation.

## REFERENCES

- Baer, D.M., M.M. Wolf, and T. R. Risley 1968 "Some Current Dimensions of Applied Behavioral Analysis." *Journal of Applied Behavior Analysis* 1:91-97.
- Bandura, A. 1986 *Social Foundations of Thought and Action*. Englewood Cliffs, N.J.: Prentice Hall.
- Baum, W. M. 1974 "On Two Types of Deviation from the Matching Law: Bias and Undermatching." *Journal of the Experimental Analysis of Behavior* 22:231-242.
- Bem, D. J. 1965 "An Experimental Analysis of Self-Persuasion." *Journal of Experimental Social Psychology* 1:199-218.
- 1972 "Self-Perception Theory." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, Vol. 6. New York: Academic Press.
- Bradshaw, C. M., and E. Szabadi 1988 "Quantitative Analysis of Human Operant Behavior." In G. Davey and C. Cullen, eds. *Human Operant Conditioning and Behavior Modification*. New York: Wiley.
- Carr, E. G., and J. J. McDowell 1980 "Social Control of Self-Injurious Behavior of Organic Etiology." *Behavior Therapy* 11:402-409.
- Catania, C. A., and T. A. Brigham, eds. 1978 *Handbook of Applied Behavior Analysis: Social and Instructional Processes*. New York: Irvington Publishers.
- Conger, R., and P. Killeen 1974 "Use of Concurrent Operants in Small Group Research." *Pacific Sociological Review* 17:399-416.
- De Villiers, P. A. 1977 "Choice in Concurrent Schedules and a Quantitative Formulation of the Law of Effect." In W. K. Honig and J. E. R. Staddon, eds., *Handbook of Operant Behavior*. Englewood Cliffs, N.J.: Prentice-Hall.
- Emerson, R. M. 1972 "Exchange Theory Part 1: A Psychological Basis for Social Exchange." In J. Berger, M. Zelditch, Jr., and B. Anderson, eds. *Sociological Theories in Progress* 38-57. Boston: Houghton Mifflin Co.
- Epling, W. F., and W. D. Pierce 1988 "Applied Behavior Analysis: New Directions from the Laboratory." In G. Davey and C. Cullen, eds., *Human Operant Conditioning and Behavior Modification*. New York: Wiley.
- Fantino, E., and C.A. Logan 1979 *The Experimental Analysis of Behavior: A Biological Perspective*. San Francisco: W. H. Freeman.
- Herrnstein, R. J. 1961 "Relative and Absolute Response Strength as a Function of Frequency of Reinforcement." *Journal of the Experimental Analysis of Behavior* 4:267-272.
- 1970 "On the Law of Effect." *Journal of the Experimental Analysis of Behavior* 13:243-266.
- 1990 "Rational Choice Theory: Necessary but Not Sufficient." *American Psychologist* 45:356-367.
- Homans, G. C. 1974 *Social Behavior: Its Elementary Forms*, rev. ed. New York: Harcourt Brace Jovanovich.
- Kazdin, A. E. 1994 *Behavior Modification in Applied Settings*. New York: Brooks/Cole Publishing.
- Keller, F. S. 1968 "Good-bye teacher. . ." *Journal of Applied Behavior Analysis* 1:79-89.
- Lloyd, K. E., and M. E. Lloyd 1992 "Behavior Analysis and Technology of Higher Education." In R. P. West and L. A. Hamerlynck, eds., *Designs for Excellence in Education: The Legacy of B. F. Skinner* 147-160. Longmont, Colo.: Spores West, Inc.
- McDowell, J. J. 1981 "On the Validity and Utility of Herrnstein's Hyperbola in Applied Behavior Analysis." In C. M. Bradshaw, E. Szabadi and C. F. Lowe, eds., *Quantification of Steady-State Operant Behaviour* 311-324. Amsterdam: Elsevier/North-Holland.
- 1988 "Matching Theory in Natural Human Environments." *The Behavior Analyst* 11:95-109.
- McLaughlin, B. 1971 *Learning and Social Behavior*. New York: Free Press.
- Pierce, W. D., and W. F. Epling 1983 "Choice, Matching, and Human Behavior: A Review of the Literature." *The Behavior Analyst* 6:57-76.
- 1984 "On the Persistence of Cognitive Explanation: Implications for Behavior Analysis." *Behaviorism* 12:15-27.
- 1999 *Behavior Analysis and Learning*. Upper Saddle River, N.J.: Prentice-Hall.
- , W.F. Epling, and D. Boer 1986 "Deprivation and Satiation: The Interrelations Between Food and Wheel Running." *Journal of the Experimental Analysis of Behavior* 46:199-210.
- Skinner, B.F. 1953 *Science and Human Behavior*. New York: Free Press.
- 1957 *Verbal Behavior*. New York: Appleton Century Crofts.

— 1969 *Contingencies of Reinforcement: A Theoretical Analysis*. New York: Appleton Century Crofts.

— 1974 *About Behaviorism*. New York: Knopf.

— 1987 "Selection by Consequences." In B.F. Skinner, *Upon Further Reflection*. Englewood Cliffs, N.J.: Prentice-Hall.

Sunahara, D., and W. D. Pierce 1982 "The Matching Law and Bias in a Social Exchange Involving Choice Between Alternatives." *Canadian Journal Of Sociology* 7:145-165.

Zeiler, M. 1977 "Schedules of Reinforcement: The Controlling Variables." In W.K. Honig and J.E.R. Staddon, eds., *Handbook of Operant Behavior*. Englewood Cliffs, N.J.: Prentice-Hall.

Zuriff, G.E. 1985 *Behaviorism: A Conceptual Reconstruction*. New York: Columbia University Press.

W. DAVID PIERCE

## BIRTH AND DEATH RATES

Much of the birth and death information published by governments is in absolute numbers. These raw data are difficult to interpret. For example, a comparison of the 42,087 births in Utah with the 189,392 in Florida in 1996 reveals nothing about the relative levels of fertility because Florida has a larger population (Ventura et al. 1998, p. 42).

To control for the effect of population size, analyses of fertility and mortality usually use rates. A rate measures the number of times an event such as birth occurs in a given period of time divided by the population at risk to that event. The period is usually a year, and the rate is usually expressed per 1,000 people in the population to eliminate the decimal point. Dividing Florida's births by the state's population and multiplying by 1,000 yields a birth rate of 13 per 1,000. A similar calculation for Utah yields 21 per 1,000, evidence that fertility makes a greater contribution to population growth in the state with the large Mormon population.

### BIRTH RATES

The crude birth rate calculated in the preceding example,

Crude birth rate per 1,000= (1)

$$\frac{\text{Live births in year } x}{\text{Midyear population in year } x} \times 1,000$$

is the most common measure of fertility because it requires the least amount of data and measures the impact of fertility on population growth. Crude birth rates at the end of the twentieth century range from over 40 per 1,000 in many African countries and a few Asian countries such as Yemen and Afghanistan to less than 12 per 1,000 in the slow-growing or declining countries of Europe and Japan (Population Reference Bureau 1998).

The crude birth rate is aptly named when used to compare childbearing levels between populations. Its estimate of the population at risk to giving birth includes men, children, and postmenopausal women. If women of childbearing age compose different proportions in the populations under consideration or within the same population in a longitudinal analysis, the crude birth rate is an unreliable indicator of the relative level of childbearing. A portion of Utah's 61 percent higher crude birth rate is due to the state's higher proportion of childbearing-age women, 23 percent, versus 20 percent in Florida (U.S. Bureau of the Census 1994). The proportion of childbearing-age women varies more widely between nations, making the crude birth rate a poor choice for international comparisons.

Other rates that more precisely specify the population at risk are better comparative measures of childbearing, although only the crude birth rate measures the impact of fertility on population growth. If the number of women in childbearing ages is known, general fertility rates can be calculated:

General fertility rate per 1,000= (2)

$$\frac{\text{Live births in year } x}{\text{Women 15-44 in year } x} \times 1,000$$

This measure reveals that a thousand women in Utah of childbearing age produce more births in a year than the same number in Florida, 89 versus 64 births per 1,000 women between ages fifteen and forty-four (Ventura et al. 1998, p. 42). The 39 percent difference in the two states' general fertility rates is substantially less than that indicated by

BIRTH AND DEATH RATES

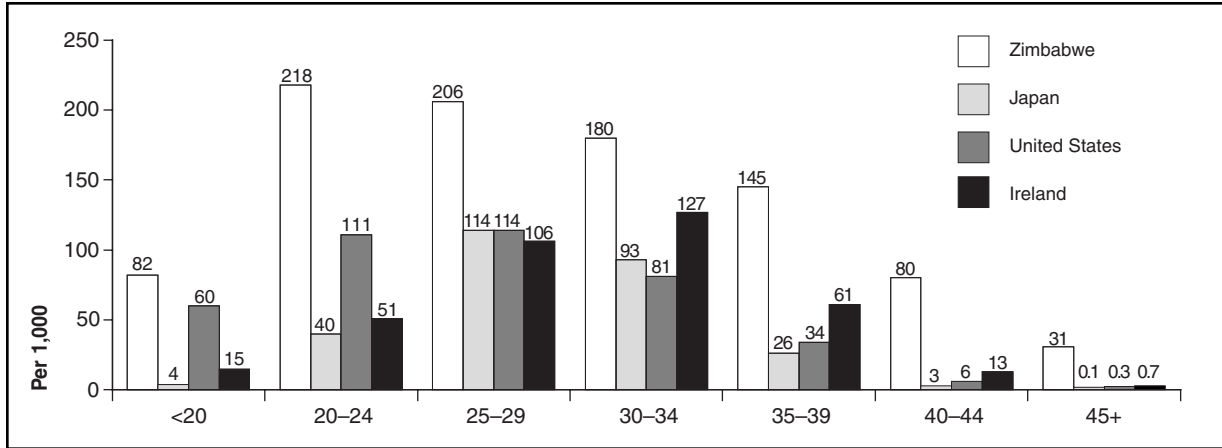


Figure 1. Age-Specific Fertility Rates: Selected Countries, Mid-1990s.

SOURCE: United Nations, *Demographic Yearbook*, 1996.

their crude birth rates which are confounded by age-composition differences.

Although the general fertility rate is a more accurate measure of the relative levels of fertility between populations, it remains sensitive to the distribution of population across women of child-bearing ages. When women are heavily concentrated in the younger, more fecund ages, such as in developing countries today and in the United States in 1980, rather than the less fecund older ages, such as in the United States and other developed countries today, the general fertility rate is not the best choice for fertility analysis. It inflates the relative level of fertility in the former populations and deflates the estimates in the latter populations.

Age-specific fertility rates eliminate potential distortions from age compositions. These rates are calculated for five-year age groups beginning with ages fifteen to nineteen and ending with ages forty-five to forty-nine:

$$\text{Age-specific fertility rate per 1,000} = \frac{\text{Live births in year } x \text{ to women age } a}{\text{Women age } a \text{ in year } x} \times 1,000 \quad (3)$$

Age-specific fertility rates also provide a rudimentary measure of the tempo of childbearing. The four countries in Figure 1 have distinct patterns. Zimbabwe, a less-developed country with high fertility, has higher rates at all ages. At the

other extreme, Japan's low fertility is highly concentrated between the ages of twenty-five and thirty-four, even though the Japanese rates are well below those in Zimbabwe. In contrast, teenage women in the United States continue to have much higher fertility than teenage women in Japan and other industrialized countries, despite declines in the 1990s. The younger pattern of American fertility also is evident in the moderately high rates for women in their early twenties. At the other extreme, Irish women have an older pattern of fertility.

More detailed analyses of the tempo of childbearing require extensive information about live birth order to make fertility rates for each age group specific for first births, second births, third births, and so forth (Shryock and Siegel 1976, p. 280). A comparison of these age-order-specific rates between 1975 and 1996 reveals an on-going shift toward later childbearing in the United States (Ventura et al. 1998, p. 6). The first birth rate increased for women over thirty while it decreased for women ages twenty to twenty-four. As a result, 22 percent of all first births in 1996 occurred to women age thirty and over, compared to only 5 percent in 1975.

When the tempo of fertility is not of interest, the advantages of age-specific fertility rates are outweighed by the cumbersome task comparing many rates between populations. As an alternative, each population's age-specific rates can be

condensed in an age–sex adjusted birth rate (Shryock and Siegel 1976, pp. 284–288). The most frequently used age–sex adjusted rate is calculated:

$$\text{Total fertility rate} = \frac{\text{Sum (age-specific fertility rates per 1,000 X 5)}}{1,000} \quad (4)$$

if the age-specific fertility rates are for five-year age groups. Single-year age-specific rates are summed without the five-year adjustment. When expressed per single woman, as in equation (4), the total fertility rate can be interpreted as the average number of births that a hypothetical group would have at the end of their reproduction if they experienced the age-specific fertility observed in a particular year over the course of their childbearing years.

Age-specific rates in real populations that consciously control fertility can be volatile. For example, the fertility rate of American women ages thirty to thirty-four fell to 71 per 1,000 in the middle of the Great Depression and climbed back to 119 during the postwar Baby Boom, only to fall again to 53 in 1975 and rebound to 84 in 1996 (U.S. Bureau of the Census 1975, p. 50; Ventura et al. 1998, p. 32). Consequently, a total fertility rate calculated from one year's observed age-specific rates is not a good estimate of the eventual completed fertility of childbearing-age women. It is, however, an excellent index of the level of fertility observed in a year that is unaffected by age composition.

The total fertility rate also can be interpreted as an estimate of the reproductivity of a population. Reproductivity is the extent to which a generation exactly replaces its eventual deaths. While the total fertility measures the replacement of both sexes, other measures of reproductivity focus only on the replacement of females in the population (Shryock and Siegel 1976, pp. 314–316). The gross reproduction rate is similar to the total fertility rate except that only female births are included in the calculation of the age-specific rates. It is often approximated by multiplying the total fertility rate by the ratio of female births to all births.

Theoretically, women need to average two births, one female and one male, and female newborns need to live long enough to have their own

female births at the same ages that their mothers gave birth to them to maintain a constant population size. In real populations some female newborns die before their mothers' ages at their births and the tempo of fertility fluctuates. As a result, both the total fertility rate and the gross reproduction rate overestimate reproductivity. The net reproduction rate adjusts for mortality, although it remains sensitive to shifts in the tempo of childbearing. This measure of reproductivity is calculated by multiplying the age-specific fertility rates for female births by the corresponding life-table survival rates that measure the probability of female children surviving from birth to the age of their mothers, and summing across childbearing ages. If the tempo of fertility is constant, a net reproduction rate of greater than one indicates population growth; less than one indicates decline; and one indicates a stationary population.

Because of the impact of female mortality, the total fertility rate must exceed two for a generation to replace all of its deaths. In industrialized countries with a low risk of dying before age fifty, the total fertility rate needs to be about 2.1 for replacement. Developing countries with higher mortality need a higher total fertility rate for replacement. Malawi, for example, with an infant mortality rate of 140 per 1,000 live births and a life expectancy of only thirty-six needs a rate of about three births per woman for replacement.

The fertility of most industrialized countries has fallen below replacement (Population Reference Bureau 1998). Some, including the United States (Figure 2), Ireland, Iceland, and New Zealand, are barely below replacement. Others have declined to unprecedented low rates. Spain (Figure 2), Portugal, Italy, Greece, Germany, most Eastern European countries, and Japan have total fertility rates of 1.1 to 1.4 births per woman. Some of these populations are already declining. Without constant, substantial net in-migration, all will decline unless their fertility rates rebound to replacement levels or above.

In contrast, the total fertility rates of most developing countries whose economies are still dependent primarily on agriculture exceed replacement (Population Reference Bureau 1998). The highest rates are found in Africa where most countries have rates greater than five births per



## BIRTH AND DEATH RATES

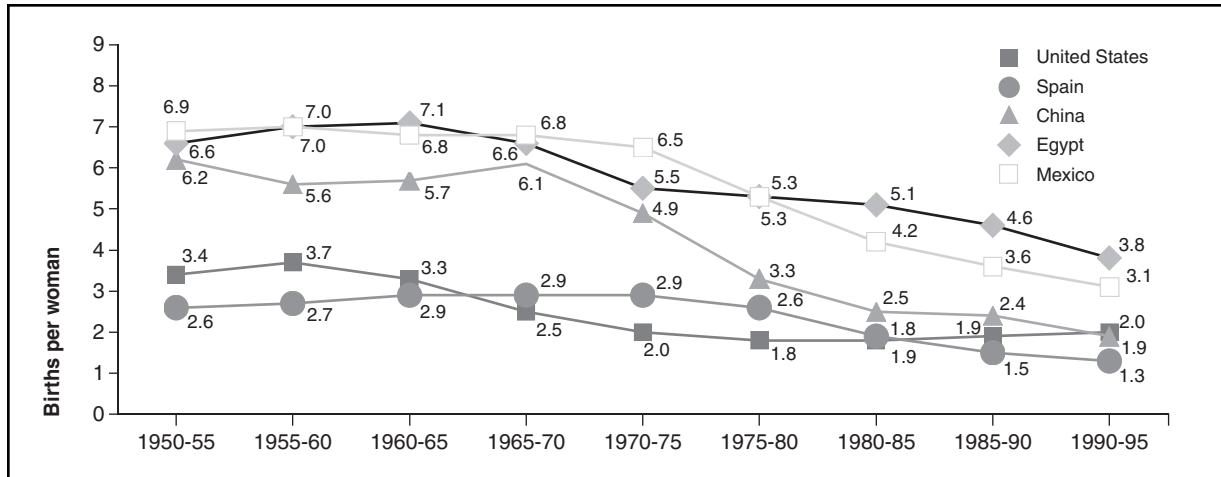


Figure 2. Total Fertility Rates: Selected Countries, 1950–1995.

SOURCE: United Nations, *World Population Prospects*, 1996.

woman. A few, including Ethiopia, Somalia, Niger, and Angola, equal or exceed seven births per woman. As a result, sub-Saharan Africa is growing at about 2.6 percent per year. If unchanged, Africa's population would double in twenty-seven years. Change, however, appears underway in most African countries. The largest fertility declines since 1980 have occurred in North African countries like Egypt (Figure 2) and in Kenya, Zimbabwe, and South Africa. Nevertheless, fertility in all African countries remains well above replacement. The continent's 1998 total fertility rate was 5.6 births per woman.

Fertility has declined to lower levels in other developing regions. Led by three decades of decline in Mexico (Figure 2), Brazil, Ecuador, Peru, and Venezuela, fertility rates for Central and South America have declined to 3.4 and 2.8 births per woman, respectively. Most Caribbean countries have near replacement-level fertility. Haiti, the Dominican Republic, and Jamaica are the major exceptions.

Rapid fertility declines also have occurred throughout much of Asia. China's aggressive birth control policy and nascent economic growth reduced its fertility to 1.8 births per woman (Figure 2), about the same as its Korean and Taiwanese neighbors. A number of other Asian countries, including the large populations of Bangladesh, Iran, Thailand, Vietnam, and Turkey, experienced

more than a 50 percent decline in the last two decades of the twentieth century. The even larger populations of India and Indonesia continued their previous slow downward trend, yielding 1998 rates of 3.4 and 2.7 births per woman, respectively. Only a few Asian countries, other than traditional Moslem societies such as Afghanistan, Iraq, and Pakistan, continue to have more than five births per woman.

The fertility measures discussed up to this point are period rates. They are based on data for a particular year and represent the behavior of a cross-section of age groups in the population in that year. Fertility also can be measured over the lifetime of birth cohorts. Cumulative fertility rates can be calculated for each birth cohort of women by summing the age-specific fertility rates that prevailed as they passed through each age (Shryock and Siegel 1976, p. 289). This calculation yields a completed fertility rate for birth cohorts that have reached the end of their reproductive years. It is the cohort equivalent of the period total fertility rate.

## DEATH RATES

The measurement of mortality raises many of the same issues discussed with fertility. Rates are more informative than absolute numbers, and those rates that more precisely define the population at



risk to dying are more accurate. Unlike fertility, however, the entire population is at risk to dying, and this universal experience happens only once to an individual.

The impact of mortality on population growth can be calculated with a crude death rate:

$$\text{Crude death rate per 1,000} = \frac{\text{Deaths in year } x}{\text{Midyear population in year } x} \times 1,000 \quad (5)$$

Crude death rates vary from over 20 per 1,000 in some African countries to as low as 2 per 1,000 (Population Reference Bureau 1998). The lowest crude death rates are not in the developed countries of Europe, North America, and Oceania, which have rates between 7 and 14 per 1,000. Instead, the lowest crude death rates are found in oil-rich Kuwait, Qatar, and United Arab Emirates, where guest workers inflate the proportion of young adults, and in the young populations of developing countries experiencing declining mortality. Developed countries that underwent industrialization and mortality decline before 1950 have old-age compositions. The proportion of people age sixty-five and over ranges between 11 and 17 percent in these countries compared to less than 5 percent in most African, Asian, and Latin American populations. Although there is a risk of dying at every age, the risk rises with age after childhood. Consequently, older populations have higher crude death rates.

To control for the strong influence of age on mortality, age-specific rates can be calculated for five-year age groups:

$$\text{Age-specific mortality rate per 1,000} = \frac{\text{Deaths in year } x \text{ to the population age } a}{\text{Population age } a \text{ in year } x} \times 1,000 \quad (6)$$

Before age five, the age-specific mortality rate usually is subdivided to capture the higher risk of dying immediately after birth. The rate for one- to four-year olds, like other age-specific rates, is based on the midyear estimate of this population. The conventional infant mortality rate, however, is based on the number of live births:

$$\text{Infant mortality rate per 1,000} = \frac{\text{Deaths under age 1 in year } x}{\text{Live births in year } x} \times 1,000 \quad (7)$$

The infant mortality rate is often disaggregated into the neonatal mortality rate for the first month of life and the postneonatal rate for the rest of the year.

$$\text{Neonatal mortality rate per 1,000} = \frac{\text{Deaths under 29 days of life in year } x}{\text{Live births in year } x} \times 1,000 \quad (8)$$

$$\text{Postneonatal mortality rate per 1,000} = \frac{\text{Deaths from 29 days to age 1 in year } x}{\text{Live births in year } x} \times 1,000 \quad (9)$$

Infant mortality varies widely throughout the world. Iraq, Afghanistan, Cambodia, and many African countries have 1998 rates that still exceed 100 per 1,000 live births, although most have declined (Population Reference Bureau 1998). Infant mortality rates in most other African, Asian, and all other Latin American countries have declined as well. They now range in developing countries from 7 in Cuba to 195 in Sierra Leone. In contrast, developed countries have rates at or below 10, led by Japan with under 4 infant deaths per 1,000 live births.

The declining American infant mortality rate, 7 per 1,000 live births in 1998, continues to lag behind Japan, Australia, New Zealand, Canada, and most Northern and Western European countries due to a higher prevalence of prematurity and low-birth-weight babies which are major causes of infant death (Peters et al. 1998, pp. 12–13). The prematurity and low-birth-weight rates for the country's largest minority, African Americans, are double those of non-Hispanic whites (Ventura 1998, pp. 57–58). Not all minority mothers have higher rates than non-Hispanic whites. Low-birth-weight rates are lower for Americans of Chinese, Mexican, Central and South American origin, about the same for Native Americans and those of Cuban origin, and higher for Puerto Rican, Filipino, and Japanese Americans.

BIRTH AND DEATH RATES

**Death Rates per 100,000 for the Fourteen Leading Causes of Death in the Population 45-64 Years of Age, by Sex and Selected Race and Ethnic Categories: United States, 1997**

	Males			Females		
	Non-Hispanic Whites	Blacks	Hispanic	Non-Hispanic Whites	Blacks	Hispanic
Heart diseases	254	469	160	93	231	69
Cancers	254	431	147	219	285	129
Strokes	25	91	32	19	58	20
Respiratory diseases	26	31		24	22	7
Accidents						
Motor vehicle	19	30	21	9	11	9
Other	23	49	26	8	14	6
Pneumonia and flu	11	32	11	8	15	6
Diabetes	20	60	33	16	54	29
HIV	6	72	25		18	5
Suicide	25		12	8		
Liver disease and cirrhosis	26	44	53	10	15	14
Kidney disease						4
Septicemia				4	14	
Homicide and legal intervention		28	12			

Table 1. Death Rates per 100,000 for the Leading Causes of Death in the Population 45-64 Years of Age, by Sex and Selected Race and Ethnic Categories: United States, 1997.

SOURCE: Hoyert, D. *Deaths: Final Data for 1997*.

Infant mortality also varies by sex. The rate for male infants born to white mothers in the United States is 6.7 per 1,000 live births, compared to 5.4 for female infants (Peters 1998, p. 80). Similarly, the rate for male infants born to African-American mothers is 16 per 1,000 live births, compared to 13 for female infants.

Like race and ethnic differences, the sex difference in mortality is evident at all ages. The greatest gap is among young adults; American males are two and one-half to three times more

likely to die between the ages of fifteen and twenty-nine than females due largely to behavioral causes. These include motor vehicle and other accident fatalities, homicides, and suicides. Higher male death rates from congenital anomalies at younger ages and from heart disease in middle age (Table 1) suggest that biological factors also contribute to the sex difference in mortality. Behavioral causes continue to play a significant role in the sex difference in mortality in middle age with chronic liver disease and cirrhosis and human immunodeficiency virus (HIV) adding premature male mortality.

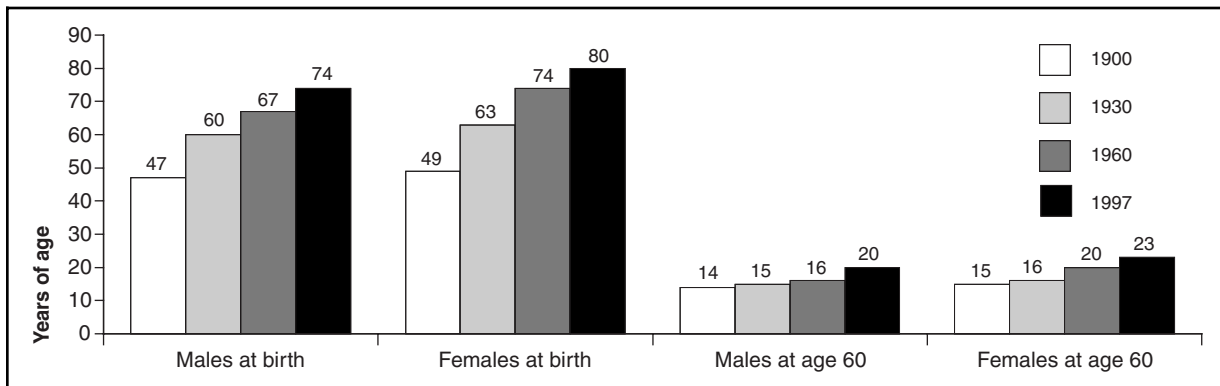


Figure 4. Life Expectancy for Whites by Sex: United States, 1900-1996.

SOURCE: U.S. Bureau of the Census, *Historical Statistics of the United States, Colonial Times to 1970*, and D. Hoyert, *Deaths: Final Data for 1997*.

BIRTH AND DEATH RATES

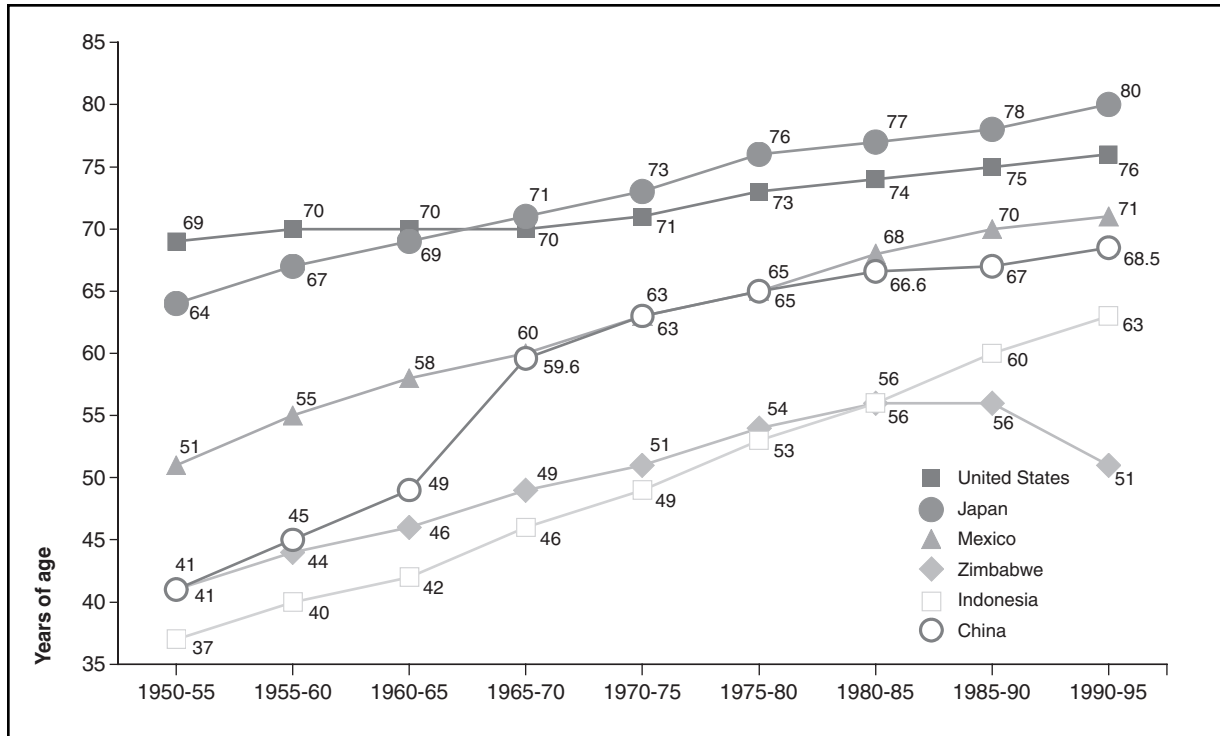


Figure 3. Life Expectancy at Birth: Selected Countries, 1950–1995.

SOURCE: United Nations, *World Population Prospects*, 1996.

Note that cause-specific death rates are calculated per 100,000 population in a specified group, rather than per 1,000 to avoid working with decimals.

Age-specific mortality rates usually are specific for sex and race or ethnicity because of the large differences evident in Table 1. Using five-year age categories results in thirty-eight rates for each racial or ethnic group. Analyses of such data can be unwieldy. When the age pattern of mortality is not of interest, an age-adjusted composite measure is preferable. The most readily available of these for international comparisons is life expectancy.

Life expectancy is the average number of years that members of an age group would live if they were to experience the age-specific death rates prevailing in a given year. It is calculated for each age group in a life table (Shryock and Siegel 1976, pp. 249–268). Life expectancy has increased since the nineteenth century in developed countries like the United States in tandem with economic growth and public health reforms. Japan has the longest

life expectancy at birth, eighty years, due to extremely low infant mortality. With the exception of Russia and some of the other former Soviet-bloc countries of Eastern Europe that have had declines in life expectancy, all other developed countries have life expectancies at birth that equal or exceed seventy years (Population Reference Bureau 1998).

Figure 3 presents trends in life expectancy since World War II in selected countries. Some of the most rapid increases have been in Asia and Latin America. China, Sri Lanka, Malaysia, Mexico, and Venezuela, for example, have 1998 life expectancies at birth of over seventy years. Increases in life expectancy in African countries, in contrast, where there is a relatively high prevalence of poor nutrition, unsanitary conditions, and infectious diseases, generally have lagged behind Asia and Latin America. Most have 1998 life expectancies in the forties or fifties, similar to those of developed countries at the end of the nineteenth century. The future trend in life expectancy in many African countries is uncertain.

Progress made in some is being undermined by the spread of HIV and political conflict. The drug therapies that have reduced mortality from HIV in developed countries have so far proven too costly for widespread use in Africa and political unrest continues.

Declines in infant and child mortality from infectious diseases have been largely responsible for the historical increase in life expectancy around the world. As a result, life expectancy has increased far more at birth than at age sixty (Figure 4). Heart disease, cancer, and cerebrovascular diseases have become the major causes of death in the United States (Peters et al. 1998, p. 7), other developed countries, and some developing countries. Bringing these chronic diseases under control has proven to be difficult. Research suggests that both genetic composition and unhealthy behaviors affect the development of such chronic diseases over time. In the United States where health education campaigns have stressed the link between these diseases and obesity, use of tobacco, and lack of sufficient physical activity, mortality has declined from heart disease since 1950 and from cancer since 1990. If these trends continue, gains in life expectancy at older ages will accelerate.

## REFERENCES

- Hoyert, D., Kenneth Kochanek, and Sherry Murphy 1999 "Deaths: Final Data for 1997." *National Vital Statistics Reports*, vol. 47, no. 19. Hyattsville, Md.: National Center for Health Statistics.
- Population Reference Bureau 1998 *1998 World Population Data Sheet*. Washington, D.C.: Population Reference Bureau.
- 1980 *1980 World Population Data Sheet*. Washington, D.C.: Population Reference Bureau.
- Shryock, Henry, and Jacob Siegel 1976 *The Methods and Materials of Demography*. San Francisco: Academic Press.
- United Nations 1998 *Demographic Yearbook, 1996*. New York: United Nations.
- 1998 *World Population Prospects, 1996*. New York: United Nations.
- U.S. Bureau of the Census 1994 "Population Projections for States by Age, Sex, Race, and Hispanic Origin: 1993 to 2020." *Current Population Reports*, series P-25, no. 1111. Washington, D.C.: Government Printing Office.

— 1975 *Historical Statistics of the United States, Colonial Times to 1970*. Washington D.C.: Government Printing Office.

Ventura, Stephanie, Joyce Martin, Sally Curtin, and T. J. Mathews 1998 "Report of Final Natality Statistics, 1996." *Monthly Vital Statistics Report*, vol. 46, no. 11, Supp. Hyattsville, Md.: National Center for Health Statistics.

DEBORAH A. SULLIVAN  
DEPARTMENT OF SOCIOLOGY  
ARIZONA STATE UNIVERSITY

## BISEXUALITY

See Sexual Orientation, Sexually Transmitted Diseases.

## BLACK STUDIES

See African-American Studies; Affirmative Action; Discrimination; Prejudice; Race.

## BRITISH SOCIOLOGY

In a global age, the concept of British sociology poses an interesting question with regard to the viability of national sociologies. Neither academic disciplines nor the subjects studied fit easily into national boundaries. An academic's closest colleague may be in New York or Delhi rather than in Lancaster or Birmingham. Key figures in British sociology, such as Dahrendorf, Westergaard, and Bauman are not British but have spent some or all of their careers working in British institutions (Halsey 1989). As sociologists working in Britain they were well placed to investigate questions related to British society. Then there are the British sociologists who have left Britain to research and teach elsewhere; John Goldthorpe to Sweden and Germany, and John Hall and Michael Mann to the United States, for example. British sociologists have often studied other nations too: Ronald Dore focuses on Japan, David Lane on Russia, and John Torrance on Austria to name a few. With all of these international influences exemplifying the present status of sociology in Britain, how "British" then is British sociology? This entry briefly explores the range of sociology that has developed in

Britain from its origins to the present day, and ends by noting possible implications for its future.

### FOUNDATION

The discipline of sociology in Great Britain has a history that stretches back to the early 1900s. Martin White and the London School of Economics (LSE) figure prominently in the development of British sociology. In 1907, White effectively founded the study of sociology in Britain by investing about £1,000 to fund a series of lectures at the LSE, as well as to establish the Sociological Society. The first annual report of the society indicated 408 members distributed throughout Great Britain, and thirty-two overseas. Early members of the society included an interesting variety of prominent public and literary figures, such as H.H. Asquith, Hilaire Belloc, and the Bishop of Stepney; British academics including Bertrand Russell, Graham Wallas, and Beatrice Webb; as well as international academics such as Emile Durkheim and Ferdinand Tonnies, among others. Also in 1907, White gave the University of London £10,000 for a permanent chair in sociology to be located at LSE. White also donated additional funds for lectureships, bursaries, and scholarships in sociology. Because of White's prominence in supporting these early initiatives, Dahrendorf has argued that "it is not too much to say that one man, Martin White, established the discipline of sociology in Britain. . . . Moreover, sociology came to life at LSE." (Dahrendorf 1995, p. 103).

Despite this promising start, by 1945 the LSE remained the only university with a department of sociology in Britain. Several reasons have been identified for this late development. Among these was the long-standing opposition to the creation of sociology as a university subject by the universities of Oxford and Cambridge, which were at the top of the educational establishment in Britain. In addition, two other disciplines had claims on similar social research that predated the emergence of sociology. Anthropology and political economy both focused on social research that suited the interests of Britain at the time. Studies of foreign shores while Britain was still a major empire was of greater interest than social research focused on issues closer to home. Empirically based scholarship on the political economy was preferable to

the theoretical emphasis of many sociologists because of its perceived lack of application to the real world. The purported lack of credibility of those promoting the study of sociology, many of whom were either located on the outside or on the margins of academe, did not lend a helping hand to the development of sociology either. But, the most persistent obstacle was the hierarchical social structure of British society that prevented the effective interrogation of its social structures (Albrow 1989).

### EXPANSION

Following World War II, the fortunes of sociology changed dramatically, in line with the social changes in Britain at this time. In these years there was a general feeling of optimism regarding possible changes in social relationships as well as increasing expectations of education and science to create a better life. In this context the study of sociology represented a commitment to social reorganization. "There was a demand [for sociology]. . . irrespective of what was on offer" (Albrow 1989, p. 202).

Around this same time, Edward Shils moved to London from the University of Chicago. His teaching of classical European sociology has been described "as magnificent a professorial presentation of the social science scenery as could be found in the Western world," and had a great deal to do with shaping and inspiring the first generation of British sociologists (Halsey 1999). This first generation of "career sociologists" in Britain completed their education at the LSE between 1950 and 1952. They included: J.A. Banks, Olive Banks, Michael Banton, Basil Bernstein, Percy Cohen, Norman Dennis, Ralf Dahrendorf, A.H. Halsey, David Lockwood, Cyril Smith, J.H. Smith, Asher Tropp, and John Westergaard. All became prominent sociologists in Britain by the mid-1960s. Members of this cohort have described the sociological analysis they developed as grounded in their own experiences of social inequality and informed by critical reflections on Marx, Parsons, and Popper (Halsey 1985; Dahrendorf 1995). Two of the many publications from members of this group are: David Lockwood's *Some Remarks on The Social System*, published in 1956, and Dahrendorf's *Class*

and *Class Conflict in Industrial Society*, published in 1959. Lockwood's paper is notable because it became the most widely read single paper by a British sociologist at the time. Dahrendorf's book, a critical reflection on Marx and Parsons, argued that Weber's ideas were useful in explaining the positive effects of conflict in Great Britain, and was dedicated to this first cohort at the LSE.

British sociology was further developed in the next decade by the work of John Rex at the University of Leeds. His work *Key Problems in Sociological Theory* (1961) became the most popular British sociological theory textbook in the 1960s. Rex used a Weberian action framework in contrast to the functionalist orthodoxy prominent at the time. In addition to Leeds, departments of sociology had been established at the universities of Leicester, Birmingham, Liverpool, Nottingham, and Hull. In 1967, three important works became available in Britain, Garfinkel's *Studies in Ethnomethodology*, Berger and Luckmann's *The Social Construction of Reality*, and Schutz's *The Phenomenology of the Social World*. These books contained an implicit critique of the kind of sociology that had been pursued in Britain up to this time. They emphasized a qualitative analysis of the ways in which social actors create meaning and acquire social positions in the context of language. The growth of the discipline continued throughout this time with nearly every institution of higher education in Britain housing a department of sociology. This led some to make the observation that it was growing with "explosive force" (Heyworth 1965, p.11). Sociology as a discipline was beginning to come into its own in Britain.

By the 1970s, there were many theoretical currents in British sociology; not only were phenomenological and Marxist arguments being pursued, but also the works of Althusser, the French structuralists, the "Frankfurt School," and Habermas and Gramsci. Links between British sociology and European social theory were becoming stronger. In addition there were important empirical studies being done; the best example is the 1972 Oxford Mobility Studies undertaken by a group of sociologists at Nuffield College, Oxford, and greatly influenced by the work of one of their team, John Goldthorpe. Around this same time a dialogue between the debates in social theory and empirical understanding was initiated by Anthony Giddens

in his analysis of the founders of sociology, *Capitalism and Modern Social Theory* (1971). By 1979 he had developed structuration theory to reconcile the previous theoretical debates. Giddens's systematic analysis of modernity is the most extensive and widely disseminated work of any British sociologist to date.

Starting in the 1980s, the emerging theoretical *zeitgeist* of British sociology was an increasing interest in the study of culture. Cultural studies as developed in the 1960s through the rather different works of Richard Hoggart, E.P. Thompson, and Stuart Hall emphasized the social aspects of culture and led to the specialization in the 1980s and 1990s of the areas of media, feminism, and ethnicity. Cultural analyses were developed as part of an overall analysis of fundamental shifts in modern societies around production, consumption, and social interaction. Influenced by European theories of politics, ideology, and discourse, existing Marxist and functionalist theories were seen as unable to adequately theorize the modern social world.

The work of sociologists based at the Birmingham Centre for Contemporary Cultural Studies represents an example of this analysis. In 1983, the group consisting of John Solomos, Bob Findlay, Simon Jones, and Paul Gilroy outlined a neo-Marxist approach to racism in a series of articles entitled *The Empire Strikes Back*. Gilroy's book, which followed the original articles, *There Ain't No Black in the Union Jack*, developed these themes and stressed the contest over the meaning of "race" that occurs in the public sphere between different social actors. Gilroy's analysis is an example of the wider move in British sociology away from a Marxist analysis of class conflict toward an emphasis on new social movements, such as the women's movement, youth movements, peace movements, green movements, and others they claimed were not easily reducible to class-based politics. There is, of course, a continuing debate between cultural-studies thinkers and Marxists, such as Robert Miles, who argue that class conflict is still far from dead, and remains a central feature in explaining the production of racism. The diversity and pluralism within sociology in Britain at this time has been described as one of its great strengths (Albrow 1989).

### CHALLENGES AND OPPORTUNITIES

The growth of sociology in Britain was abruptly impeded by a governmental investigation of higher-education research funding in 1981. One recommendation following the investigation was the withdrawal of funding for any new departments of sociology and those seen as "substandard." Sociologists were also limited in their representation on the Social Science Research Council (SSRC), with only two of the ninety-four members from sociology departments. As a reflection of this change in academic focus, the SSRC was renamed the Economic and Social Research Council in 1984. Albrow (1989) argues that no other academic discipline was singled out in this way.

Governmental scrutiny of sociology programs and departments has become an integral part of British higher education. It is, however, not because it continues to be singled out for special treatment as it was in previous decades but rather it is part of the larger process of fiscal accountability and quality assurance instituted by the government. The first two Research Assessment Exercises (RAE) were conducted in 1992 and 1996 respectively. Each RAE rated every department of every subject on its research output. In 1996, sociology at the universities of Essex and Lancaster received 5\* (the highest rank) ratings, while close behind were the universities of Loughborough, Manchester, Oxford, Surrey, Warwick, and Edinburgh, and the college of Goldsmiths, with ratings of 5 each (RAE 1996). Most important about the ratings is that the British government "... funding bodies use the ratings to inform the distribution of grants for research to HEIs" (higher education institutions) (RAE 1996). It has been reported that, "Some 20 percent of funding from the Higher Education Funding Council for England is a reward for research excellence, as a measured by the "research assessment exercise" (Wolf 1999). In fact, these funds are extremely important to the day-to-day running of departments and in some cases make them viable or not.

If research was to be assessed, then teaching could not be far behind. The Teaching Quality Assessment (TQA) exercise was instituted in 1993 with various subjects coming up for review in each country of Britain over an eight-year period. In England, sociology was reviewed between 1995

and 1996. At the time of TQA there were approximately 20,000 students studying sociology in eighty-one institutions of higher education in England alone. The Overview Report of all departments reviewed in England noted that, "The overall picture that emerges from the assessment process of the quality of education in sociology is a positive one. . . . On the other hand there is little room for complacency." It was further noted that, "The majority of institutions use a suitable range and variety of teaching, learning and assessment methods. The assessors found considerable evidence of innovative approaches to teaching, which encouraged and enthused students (HEFCE 1996, p. 14).

Funding does not follow from high TQA scores. Instead, high-scoring departments can bid for money from the Fund for the Development of Teaching and Learning. In sociology five consortium projects were funded to a total of over £1,000,000. The five projects publish a regular newsletter and there are plans to launch an online journal for teaching in 2000 (Middleton 1999).

### THE NEXT CENTURY

As British sociology enters a global age it is challenged with the ironies and complexities of understanding the current state of modernity that includes the increasing interconnectedness of the world as a whole, the increasing choices each individual has, and the resilience of national identities as evidenced by the increasing number of nations that emerge year after year. Each of these issues has been the subject of research by British sociologists in the past decade. The 1990s saw an "explosion" of interest in studies of culture informed by debates around the high-modern or postmodern condition, albeit influenced by the works of the French sociologists, Pierre Bourdieu, and Michel Foucault (Swingewood 1998). Both the study of nations and trends toward globalization find British sociology in the lead as well. Ernest Gellner (1983) and Anthony Smith (1986) have claimed the territory of nations and nationalism. At the same time, Anthony Giddens (1990), Roland Robertson (British by birth) (1992), Leslie Sklair (1995), and Martin Albrow (1996) have advanced theories of globalization. Emerging from work on globalization is the sociology of human rights. Two examples include: Anthony Woodiwiss's (1998) work on the compatibilities between Asian

values and human rights discourses, and Kevin Bales's (1999) study of economic globalization and the development of modern slavery. These seem to highlight some important issues to be debated within sociology (worldwide) in this millennium.

The first century of British sociology saw its ups and downs. At times it was a popular subject to study and at other times it faced significant threats to its existence. Probably there was no threat more dramatic than the former Prime Minister Margaret Thatcher's declaration that, "There is no such thing as society." However, despite the attacks and the loss of funding, sociology as a discipline has managed not only to survive but is poised to lead policy making in Britain. Currently there are increasing numbers of Members of Parliament and Lords who have a sociology background and there are nearly a dozen Vice Chancellors of universities who are sociologists. "Indeed, the subject somehow seems in tune with the times. With a government that not only accepts, but explicitly seeks to understand, and manage society, it is surely no coincidence that Tony Blair's favourite academic, Anthony Giddens. . . is another sociologist-cum-VC (or Director as the position is called at the London School of Economics)" (Brown 1999, p. iii). The reference here is to Giddens's latest contribution to British sociology, his book entitled *The Third Way: The Renewal of Social Democracy* (1998). In many ways sociology has come full circle, with LSE academics once again playing an important role in its resurgence. Despite all attempts to deny its value, sociology in Britain can once again try to fulfill the claim made a century ago and an ocean away, that "Sociology has a foremost place in the thought of modern men [and women]" (Small 1895, p.1).

## REFERENCES

- Albrow, Martin 1996 *The Global Age*. Cambridge, Eng.: Polity.
- 1989 "Sociology in the United Kingdom after the Second World War.", In N. Genov, ed., *National Traditions in Sociology*. London: Sage.
- Bales, Kevin 1999 *Disposable People: Modern Slavery in the Global Economy*. Berkeley: University of California Press.
- Berger, Peter L., and Thomas Luckmann 1967 *The Social Construction of Reality*. London: Allen Lane.
- Brown, Matthew 1999 "Time to Take Us Seriously." *Guardian Higher* 19 January, ii–iii.
- Centre for Contemporary Cultural Studies (eds.) 1983 *The Empire Strikes Back*. London: Hutchinson.
- Dahrendorf, Ralph 1995 *A History of the London School of Economics*. Oxford: Pergamon.
- 1959 *Class and Class Conflict in Industrial Society*. London: Routledge and Kegan Paul.
- Garfinkel, Harold 1967 *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.
- Gellner, Ernest 1983 *Nations and Nationalism*. Oxford: Basil Blackwell.
- Giddens, Anthony 1998 *The Third Way: The Renewal of Social Democracy*. Cambridge, Eng.: Polity.
- 1990 *The Consequences of Modernity*. Cambridge, Eng.: Polity.
- 1971 *Capitalism and Modern Social Theory*. Cambridge, Eng.: Cambridge University Press.
- Gilroy, Paul 1987 *There Ain't No Black in the Union Jack*. London: Hutchinson.
- Goldthorpe, John H. 1980 *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon.
- Halsey, Albert H. 1999 "Edward Albert Shils." Personal correspondence, 18 June.
- 1989 "A Turning of the Tide? The Prospects for Sociology in Britain." *British Journal of Sociology* 40(3):353–373.
- 1985 "Provincials and Professionals: The British Post-war Sociologists." In Martin Bulmer, ed., *Essays on the History of British Sociological Research*. Cambridge, Eng.: Cambridge University Press.
- Heyworth, Lord 1965 *Report of the Committee on Social Studies*. London: HMSO, Cmnd 2660.
- Higher Education Funding Council for England 1996 *Subject Overview Report: Quality Assessment of Sociology 1995–1996*. Bristol, Eng.: HEFCE.
- Lockwood, David 1956 "Some Remarks on The Social System." *British Journal of Sociology* VII(2):134–145.
- Middleton, Chris 1999 Personal correspondence, 9 June.
- Research Assessment Exercise 1996 "Unit of Assessment: 42 Sociology." [http://www.niss.ac.uk/education/hefc/rae96/1\\_96t42.html](http://www.niss.ac.uk/education/hefc/rae96/1_96t42.html).
- 1996 "The Outcome." [http://www.niss.ac.uk/education/hefc/rae96/cl\\_96.html](http://www.niss.ac.uk/education/hefc/rae96/cl_96.html).
- Rex, John 1961 *Key Problems in Sociological Theory*. London: Routledge and Kegan Paul.
- Robertson, Roland 1992 *Globalization: Social Theory and Global Culture*. London: Sage.



- Schutz, Alfred 1967 *The Phenomenology of the Social World*. Evanston, Ill.: Northwestern University Press.
- Sklair, Leslie 1995 *Sociology of the Global System*. Baltimore, Md.: Johns Hopkins University Press.
- Small, Albion 1895 "The Era of Sociology." *American Journal of Sociology*. 1(1):1-15.
- Smith, Anthony 1986 *The Ethnic Origins of Nations*. Oxford: Basil Blackwell.
- Swingewood, Alan 1998 *Cultural Theory and the Problem of Modernity*. London: Macmillan.
- Wolf, Martin 1999 "More Meant Worse." *Financial Times*, 21 June, 22.
- Woodiwiss, Anthony 1998 *Globalization, Human Rights and Labour Law in Pacific Asia*. Cambridge, Eng.: Cambridge University Press.

PETER THAYER ROBBINS  
VANEETA-MARIE D'ANDREA

## BUREAUCRACY

The origin of the term *bureaucracy* can be traced to eighteenth-century French literature (Albrow 1970). The early usage referred to an official workplace (bureau) in which individual activities were routinely determined by explicit rules and regulations. As modern systems of management, supervision, and control, bureaucracies are designed to rationally coordinate the duties and responsibilities of officials and employees of organizations. The delineation of official duties and responsibilities, by means of formal rules and programs of activity (March and Simon 1958), is intended to displace and constrain the otherwise private, idiosyncratic, and uniquely personal interests and actions of individuals. Bureaucratic systems of administration are designed to ensure that the activities of individuals rationally contribute to the goals and interests of the organizations within which they work.

### THE CONTRIBUTION OF MAX WEBER

The historical trend toward increasing bureaucratization throughout modern Western Europe, highlighted by the changing structure of military organizations, is documented by the work of Karl

Marx ([1852] 1963) and of Alexis de Tocqueville (1877). Michels' (1949) analysis of the dynamics of power distribution within bureaucratic organizations and the concomitant development of oligarchic tendencies (fundamentally detrimental to democratic principles) provided an understanding of one of the more important unintended consequences of bureaucratic processes. However, the study of bureaucratic structure and processes as a prominent sociological topic is based on the intellectual legacy of Max Weber (1864-1920). Although Weber conducted his studies at the turn of the nineteenth century, wide recognition of his work by English-speaking theorists had to await later translations of his work (Weber 1946, 1947).

Weber's thorough and richly informative work included consideration of Chinese, Egyptian, Roman, Prussian, and French administrative systems. In his comparative analysis of this vast array of diverse cultural systems of administration, Weber recognized an inexorable relationship between power and authority, on the one hand, and differing systems of administration on the other. In Weber's analysis, the bureaucratic form of administration reflects one of three ways in which power can be legitimated. To gain a clear appreciation of the unique features of bureaucratic structure and processes, it is necessary to briefly address Weber's proclaimed relationship between power, authority, and systems of administration. *Power*, for Weber, represents the ability or capacity to have other people behave in accordance with certain orders or dictates, irrespective of whether those affected regard its application as rightful or legitimate. *Authority* represents the legitimation of this power by those individuals whose activities are so ordered such that the application of power and its impact on them are perceived to be proper and acceptable. In Weber's historical analysis, three different types of authority are identified: traditional, charismatic, and rational-legal (Weber 1947, p. 328).

**Three Types of Authority.** *Traditional authority* represents a system in which the use of power is regarded as legitimate when it is predicated upon belief in the sanctity of time-honored traditions and patterns of behavior. By contrast, *charismatic authority* is based on situationally specific acts of personal devotion. Under these circumstances, authority is conferred upon a specific individual

who is regarded by the devoted followers to exhibit exceptional, sacred, and/or heroic characteristics. Charismatic authority is therefore limited to a specific setting and time. *Rational-legal* authority derives from the belief in the legitimacy of law, specifically in the legality of rules and the authority of officials and employees to perform certain legally sanctioned and mandated duties to which they have been formally assigned. Official rational-legal authority often permits individuals to perform tasks that otherwise would not be permitted. For example, police officers may engage in certain behaviors in the course of their official duties, such as the use of lethal force, that if performed by an ordinary citizen would expose the citizen to legal liability. Thus, it is not the act or set of behaviors per se that is critical to an understanding of rational-legal authority, but rather who performs the act or behavior and whether the individual performing the act is legally authorized to engage in such behavior.

**Three Systems of Administration.** Each type of authority is associated with a distinctive system of administration. Over the course of premodern social history, traditional authority was the principal means by which social organization was achieved. This type of authority structure resulted in the development of a wide variety of highly stable but nonetheless particularistic systems of administration, in which personal relations of dependence or loyalty provided the bases for authority. In general, these systems of administration are most clearly exemplified by patrimonial and feudal systems of administration. Unlike traditional authority, charismatic authority results in highly unstable systems of administration, because the foundations of these systems, the profoundly personal relationships between charismatic leaders and followers, are decisively limited in both time and circumstance. Given the tenuous foundation of charismatic authority structures, it is not surprising that the systems of administration that arise in such situations encounter difficulties in generating stable administrative practices. In this regard, problems pertaining to the routinization of authority and leadership succession are particularly salient and acute (Weber 1947, pp. 358–373). Thus, administrative systems predicated upon charismatic authority tend to be inherently transitory and most likely arise during periods of crisis or unprecedented change.

In contrast to the personal bases of authority inherent in both traditional and charismatic systems of administration, rational-legal systems of administration rely on impersonal rules and regulations, culminating in the emergence of highly precise and universalistic systems of administration that are most clearly exemplified by the modern rational bureaucracy. Weber clearly regards such administrative practices to be relatively recent in their development: “Bureaucracy . . . is fully developed in political and ecclesiastical communities only in the modern state, and in the private economy only in the most advanced institutions of capitalism” (Weber 1978, p. 956).

### THE FORMAL CHARACTERISTICS OF BUREAUCRACIES

The most distinguishing feature of modern rational bureaucracies is the formal control, prescription, and regulation of individual activity through the enforcement of rules. The explicit intent of enforcing these rules is to efficiently achieve specific organizational goals. In orchestrating individual action, a succinct and unambiguous specification of the official duties and responsibilities is provided to minimize, if not to eliminate, the influence of personal interests and ambitions upon the performance of official duties. The official or employee is then able to concentrate exclusively upon the technical aspects of the work, in particular the efficient and rational completion of assigned tasks. In addition to this attempt to separate individuals’ private concerns from their official duties and responsibilities, other distinguishing characteristics of bureaucracies include:

1. The hierarchical ordering of authority relations, limiting the areas of command and responsibility for subordinate as well as superordinate personnel
2. The recruitment and promotion of individuals on the basis of technical expertise and competence
3. A clearly defined division of labor with specialization and training required for assigned tasks
4. A structuring of the work environment to ensure continuous and full-time employment, and the fulfillment of individual career expectations within the organization

5. The impersonality and impartiality of relationships among organization members and with those outside the organization
6. The importance of "official records" in the form of written documents.

### **THE RATIONALITY OF BUREAUCRATIC RULES**

Within a bureaucratic organization, rules serve to direct individual action in ways that promote the technical efficiency of the organization. The distinctive feature of bureaucratic organization is not the use of rules per se but, rather, the type of rules employed within an organization as well as the justification for the use of rules. Rules have been, and continue to be, used in other forms of administration to control individual action; however, the rules used in the other administrative forms (such as rules based on tradition in feudal administration or on dependency in a patrimonial system) are not necessarily based on technical knowledge and the rational achievement of specific goals. By contrast, within a modern bureaucracy explicit rules are designed to assure uniformity of performance in accordance with technical requirements. Bureaucracies denote systems of control that attempt to ensure that the technical abilities of individuals are effectively utilized. A concerted effort is made to systematically exclude any factor or element that would reduce the prospect of an official's or employee's performance being anything other than organizationally focused, affectively neutral, and achievement oriented (Parsons and Shils 1951, pp. 76-91).

### **THE IDEAL TYPE CONSTRUCT**

By underscoring the unique features of modern bureaucratic systems of administration, Weber provides an ideal-type characterization of bureaucracies. Despite subsequent confusion as to the value of the ideal-type portrayal of bureaucracies, Weber did not intend the characterization to represent an accurate empirical description of bureaucracies. As noted in the introduction to one of his edited works, "Situations of such pure type have never existed in history. . . . The ideal types of Weber's sociology are simply mental constructs to serve as categories of thought, the use of which will

help us to catch the infinite manifoldness of reality . . . "(Rhinestein 1954, pp. xxix-xxx). The ideal-type characterization is not suggestive of either extreme cases or distinct logical categories of phenomena (Mouzelis 1967). Rather, it provides a simplification and exaggeration of empirical events so that one can appreciate more clearly the features of the phenomena in question. Actual bureaucratic organizations may exhibit only a limited number of these properties or may possess them in varying degrees, a point well understood by Weber.

### **A GENERAL THEORETICAL PERSPECTIVE ON BUREAUCRATIC ORGANIZATIONS**

As a general line of inquiry, the study of bureaucratic structure and processes has tended to represent a "closed" or "rational-system" approach to the study of organizations (Scott 1992). Predicated upon Weber's conception of bureaucracy, the "legal-rational" features of organizations such as the use of formal rules, the creation of specific personnel positions or roles, and personnel policies directed at the efficient achievement of organizational goals have been the focus of study. By directing attention toward the internal properties of bureaucratic organizations, interest in a number of topics has arisen, including the impact of technology on organizational structures (Thompson 1967), the interrelationships among organizational features (Hage et al. 1968), the integration of individuals within organizations (Argyris 1964; Herzberg et al. 1959), and the increase of worker productivity and organizational efficiency (Taylor 1911; Mayo 1933). Across this array of topics, a general theoretical perspective prevails in which the internal properties of bureaucratic organizations are regarded to be of cardinal importance. By directing primary attention toward the internal properties of bureaucratic organizations, the extent to which organizations are embedded in the more global structure and context of society is not fully realized. However, as bureaucratic organizations have grown in function and complexity, assuming an ever-expanding role in society, and displacing previous, more particularistic and personalized forms of administration, management, and control, the vibrant interplay between the internal properties of organizations and the environments in which they are situated has become more salient. Correspondingly, a more "open" theoretical orientation is warranted. Organizations

are viewed as “natural systems” (Scott 1992) in which the interaction between an organization and its environment involves a series of potential two-way relationships. While many of the processes within a bureaucracy are designed to insulate the organization from external influences, invariably environmental influences can appreciably impact an organization and its internal operations. Similarly, the bureaucratic organization, with its expanding role in modern society, has unquestionably had a major impact on society.

### EMPIRICAL ASSESSMENT

Stanley Udy (1959) was among the first to propose that, rather than regarding the specification of bureaucracies to be strictly a matter of definition, we need to ascertain empirically the extent to which bureaucratic characteristics are associated with one another in actual organizations. In the subsequent efforts at empirical assessment of the extent to which organizations exhibit bureaucratic properties, the works of Richard Hall (1963) and the Aston University research group (Pugh et al. 1968) are especially noteworthy. Hall’s (1963) findings suggest that among samples of U.S. organizations, bureaucratic features of organizations may vary independently of one another. As is illustrated by the negative relationship between an emphasis on technical qualifications and other bureaucratic features—in particular hierarchy of authority and rule enforcement—Hall’s study suggests that bureaucratic systems of organization may be indicative of multidimensional rather than unitary processes.

The Aston University research group (Pugh et al. 1968) reported similar findings for organizations in Britain. On the basis of their measurements of the bureaucratic characteristics of organizations specified by Weber, these researchers found four mutually independent dimensions of organizational structure rather than a single overarching bureaucratic dimension. With this finding, the authors concluded that bureaucracy is a multidimensional phenomenon and that it “. . . is not unitary, but that organizations may be bureaucratic in any number of ways” (Pugh et al. 1968, p. 101). However, these findings have not been unchallenged. The contentious nature of inquiry into the precise structure of bureaucratic organization is succinctly reflected in the work of

Blau (1970) and of Child (1972), the adoption of a modified position by one of the investigators in the original Aston group (Hickson and McMillan 1981), and the subsequent reply by Pugh (1981).

### THE RISE OF BUREAUCRATIC CONTROL

Although bureaucracy existed in imperial Rome and ancient China, as well as in various national monarchies, the complexity of legislative issues arising within the modern state has caused an enormous growth of administrative function within both government and the private sector. Consequently, the power and authority of bureaucratic administrative officials to control policy within an organization as well as the modern state has, over time, increased significantly. The rise to power of bureaucratic officials means that, without expressly intending to achieve power, nonelected officials can and do have a significant impact on a broad spectrum of activities and future developments within society. As Weber noted,

*The bureaucratic structure is everywhere a late product of historical development. The further back we trace our steps, the more typical is the absence of bureaucracy and of officialdom in general. Since bureaucracy has a rational character, with rules, means–ends calculus, and matter-of-factness predominating, its rise and expansion has everywhere had revolutionary results. (1978, p. 1002)*

Moreover, Weber contended that traditional authority structures have been, and will continue to be, replaced by the rational–legal authority structures of modern bureaucracies, given their “purely technical superiority over any other form of organization” (Weber 1946, p. 214). Regardless of the apparent technical benefit, the increased prominence of bureaucracies in both the public and the private sectors is not without its problems. As Weber acknowledged, certain negative consequences may follow the development of bureaucratic systems of administration to include:

1. The monopolization of information and the creation of “official secrets”
2. The inability to change bureaucratic structure because of vested incentive and reward systems, and the dependency of society on the specialization and expertise provided by the bureaucracy

3. The tendency for bureaucracies to act in an autocratic manner, indifferent to variations or changes not previously articulated or anticipated within the bureaucracy. (Weber 1947, pp. 224–233)

A frequently expressed concern is that bureaucracies are often unresponsive to the individuals and groups they were designed to serve. To exacerbate the situation still further, few, if any, techniques of control are available outside the bureaucracies to make officials more responsive. Additionally, a number of practical problems may arise to potentially undermine bureaucratic efficiency. These difficulties can include the unwarranted application of rules and regulations, the duplication of effort, and an indifferent and even cavalier attitude among officials. Nonetheless, bureaucracies are relatively efficient and technically superior forms of administration proven to be indispensable to large, complex organizations and modern society. As Perrow has noted (1972), criticism of bureaucracies frequently relates to the fact that the actions of officials are not bureaucratic enough and personal interests may not be fully insulated from official duties.

In viewing the operation and function of bureaucracies, it is imperative that the operational efficiency of bureaucratic procedures be recognized but not at the cost of neglecting a more “humanistic” orientation (Kamenka 1989, ch. 5). The situational constraints faced by individuals and groups both within and outside bureaucracies warrant attention. As Martin Albrow has noted, bureaucracy is “a term of strong emotive overtones and elusive connotations” (1970, p. 13), and as such it represents more than a straightforward technical process and deserves an eclectic perspective in order to fully appreciate its complexities.

### BUREAUCRATIC DYSFUNCTIONS

The classic works of Merton (1940) and of Gouldner (1954) illustrate how unanticipated developments can adversely impact the intended effectiveness of bureaucratic procedures. Merton notes that, commencing with the need for bureaucratic control, individual compliance with rules is enforced, thereby allowing for the development of routinely prescribed, reliable patterns of activity. However, when

this agenda of rule compliance is implemented in a dynamic and fluctuating environment requiring more spontaneous responses, these prescribed patterns of bureaucratic activity can lead to adverse unintended consequences. Even though the circumstances require a different type of response, prescribed and fixed patterns of response may still be adopted because such responses are legitimated and defensible within the bureaucracy, given the extent to which they enhance individual reliability. Consequently, officials and employees do not accommodate the unique features of the situation, efficiency is undermined, and difficulties with clients and customers may ensue. Eventually, troublesome experiences with customers and clients may contribute to an even greater emphasis on bureaucratically reliable behavior rather than attenuating this encapsulated and counterproductive type of behavior. As Merton (1940) notes, “Adherence to the rules, originally conceived of as a means, becomes transformed into an end in itself; there occurs the familiar process of displacement of goals whereby an instrumental value becomes a terminal value. . .” While more highly adaptive and flexible behavior is required and permitted within a bureaucracy, powerful structural constraints may operate to promote situationally inappropriate rule-bound behavior (Blau and Meyer 1987; Allinson 1984).

Like Merton, Gouldner (1954) is concerned with possible unintended effects of formal rule enforcement. In Gouldner’s model, the implications of using general and impersonal rules as a means of enforcing organizational control are investigated. The intention of using such rules is to mask or partially conceal differential power relations between subordinates and their superiors. In societies with egalitarian norms, such as the United States, this serves to enhance the legitimacy of supervisory positions, thereby reducing the prospect of tension between groups with differing power. However, the use of general and impersonal rules also has the unintended consequence of providing only minimal guidelines regarding acceptable organizational behavior. In turn, if only minimum standards of performance are specified and if individuals conform only to these standards, then a disparity arises between the stated goals of the organization, which require a level of performance beyond minimally acceptable and specified standards, and actual individual performance. Since

a greater level of individual output and performance is required, more personal and closer forms of supervision need to be employed. This closer supervision increases the visibility of power relations, a consequence that may undermine the effect previously achieved by the use of general and impersonal rules.

The satirical, but nonetheless insightful, work of Peter and Hull (1969) indicates that bureaucratic processes may not always reflect the principle of personal expertise. Peter and Hull contend that even though individual talent and expertise are the formal requisites for recruitment and promotion within a bureaucracy, individuals often are promoted to positions that exceed their level of competence. Since the act of promotion often requires the assumption of new duties, an element of uncertainty over whether individuals will be as efficient in their new positions as they were in their previous positions is ever present. Not until an individual has had the position for a period of time is it apparent if he will be able to perform the newly assigned tasks in an efficient and competent manner. Most individuals are able to achieve an acceptable level of competence in their new duties; but what about those individuals who are unable to perform their new duties competently? Peter and Hull contend that, instead of being relocated to more suitable positions, these individuals remain in the elevated positions due, in all likelihood, to the difficulties and organizational costs associated with removing and reassigning them. Consequently, some individuals may be promoted to and remain in positions that exceed their level of competence, a portrayal clearly at variance with the image of operational efficiency.

#### **THE PERVASIVENESS OF BUREAUCRACY IN SOCIETY**

Throughout the nineteenth and twentieth centuries, a notable increase occurred in the number and type of functions performed within or regulated by a bureaucratic organization. Bureaucracies have superseded other more personalized, particularistic operations (ostensibly due to the technical superiority of the bureaucratic process) or, as new processes developed, they developed within the context of a bureaucratic structure. The ubiquitous presence of bureaucracies in modern society is indicated by the common usage of the term

“red tape.” The term originally denoted unnecessary and prolonged bureaucratic paperwork and procedures, and has since been expanded to encompass a more generic sense of unproductive effort. The level of everyday exposure to bureaucracies is sufficient such that notions of what a bureaucracy is (or perhaps more importantly, what it is not) have become a part of the popular culture of modern society. Fictional settings in which bureaucratic procedures interfere with the good intentions of people, promote an unproductive or unhealthy outcome, or lead to an uninspired work environment are common themes of movies, novels, syndicated cartoon features, and television shows.

The expansion of bureaucratic procedures has clearly impacted developments in the areas of health care and labor. In the area of health care, changes in the technological base of medicine, the methods by which medical services are financed, the regulatory role of government, as well as efforts to enhance the efficiency and accountability of health care providers has culminated in the bureaucratization of medical care (Mechanic 1976). In contrast to past medical practice settings characterized by a private and dyadic relationship between a doctor and a patient, the delivery of medical care in bureaucratized settings, intended to provide broader and more uniform care, increases the likelihood of conflicts of interest and the diminished salience of individualized care. When medical care is delivered in a bureaucratic setting, the specific features of patient care may be based on a number of nonclinical elements in addition to the clinically pertinent factors. Nonclinical criteria such the determination of what services will be provided and the avoidance of “unnecessary” services, patient case management by nonmedical personnel, a predetermined referral protocol and network, and the disclosure of sensitive patient information within the bureaucratic organization all limit the role of the doctor. Each of these nonclinical, bureaucratically derived elements introduce different and potentially conflicting viewpoints in an evaluation of effective patient care and may direct attention more toward resolving organizational conflicts than enhancing patient care. The issues facing bureaucratic medical care will likely become more acute with the “aging” of society and the need to attend to a greater number of patients suffering from debilitating, protracted, and nonrecoverable medical conditions.

The logic underlying the structure of bureaucratic organization is to control and regulate work routines on the basis of formal rules and regulations. In intending to use formal rules to administer workers' duties and responsibilities, it is assumed that key components of a job can be identified to the extent that rules can effectively monitor and be used to evaluate individual performance. The precise specification of work routines for the purpose of enhancing worker productivity and organizational efficiency has been an ongoing interest throughout the twentieth century, as indicated by the time-motion studies and efficiency "experts" of the early part of the century and process evaluation and quality assessments of the later part of the century. In the course of precisely specifying work responsibilities and routines, the de-skilling of the labor force (Taylor 1911) is, if not an intended outcome, a likely development. If a job can be defined in terms of specific component tasks and if the component tasks are rudimentary enough to be regulated by formal rules, then the skill requirements of job are reduced. As this process continues, the skill levels of workers in the labor force may exhibit similar reduction. Organizations can hire an individual with the nominal skill sets needed to adhere to the formally specified rules instead of hiring someone who is capable of displaying individual judgment working effectively in an environment characterized by job variability and uncertainty.

Bureaucracies have become both indispensable and despised fixtures of modern society. In the drive for overall organizational efficiency and accountability, individuals within and outside the bureaucracy often experience the unsettling feeling that they have no influence over a process that may directly impact them. The aggravation of wanting to have a problem solved or question answered but being unable to find anyone who is in the position to respond to the expressed concern fuels criticism of bureaucracies.

While criticism of bureaucratic procedures (or lack of "appropriate" procedures) is frequent, it is also apparent many of the everyday functions in society are possible only because of the operation of bureaucracies. The major institutions of society, with the multitude of diverse projects performed every day, depend on the administrative and regulatory features of bureaucracies. The

faceless bureaucratic official who may be criticized for indifference is nonetheless an essential element in such consequential areas as national defense, public education and welfare, and communications. The assessment of bureaucratic structure and procedures therefore needs to attend to both the overall scope of bureaucratic functions and the manner in which particular tasks are performed.

"Red tape" and inefficiency are two of the most heavily criticized elements of public bureaucracies, and instigate appropriate calls for reform. However, bureaucracies contain complex structures and multifaceted procedures intended to serve a number of purposes. Criticism often reflects common misperceptions rather than the true workings of bureaucracies. Regarding red tape, the bureaucracy does not necessarily create these protracted procedures in an effort to either insulate the organization from competition or to ensure that little if anything gets done without the organization's official approval. Rather, it is more likely the result of disclosure and accountability requirements imposed on the bureaucracy (Kaufman 1977, p. 29). Similarly, inefficiency needs to be placed in the more global perspective of what in addition to the specific task is being accomplished by the bureaucracy. If the focus is limited to a specific task (e.g., construction of a highway), then the process can reasonably be viewed as inefficient if the task could have been completed sooner or cost less if done by someone else. However, there may be a number of additional goals that the bureaucracy is attempting to achieve while completing the specific task in question. In the public sector, a bureaucracy is often required to be responsive to a number of outside interest groups to ensure that its operation is socially fair, accountable, and impartial, in addition to being technically able to complete a particular task (Wilson 1989, pp.315-325). The determination of efficiency is compounded by this collateral concern regarding the attainment of social goals and specific technical goals. Bureaucracies are thus faced with the dual challenge of being socially responsible and technically proficient and therefore have to be assessed on the basis of the extent to which they concurrently achieve these two sets of goals and not just one set at the exclusion of the other.

SEE ALSO: *Complex Organizations; Organizational Effectiveness; Organizational Structure*

## REFERENCES

- Albrow, Martin 1970 *Bureaucracy*. London: Pall Mall Press.
- Allinson, Christopher W. 1984 *Bureaucratic Personality and Organization Structure*. Aldershot, Eng.: Gower.
- Argyris, Chris 1964 *Integrating the Individual and the Organization*. New York: Wiley.
- Blau, Peter M. 1970 "Decentralization in Bureaucracies." In Mayer N. Zald, ed., *Power in Organizations*. Nashville, Tenn.: Vanderbilt University Press.
- , and Marshall W. Meyer 1987 *Bureaucracy in Modern Society*, 3rd ed. New York: Random House.
- Child, John 1972 "Organization, Structure, and Strategies of Control." *Administrative Science Quarterly* 17:163–177.
- Gouldner, Alvin W. 1954 *Patterns of Industrial Bureaucracy*. Glencoe, Ill.: Free Press.
- Hage, Jerald, Derek Pugh, David Hickson, Robert Hinings, and Chris Turner 1968 "Dimensions of Organizational Structure." *Administrative Science Quarterly* 13:65–104.
- Hall, Richard 1963 "The Concept of Bureaucracy: An Empirical Assessment." *American Journal of Sociology* 69:32–40.
- Herzberg, Frederick, Bernard Mausner, and Barbara Snyderman 1959 *The Motivation to Work*. New York: Wiley.
- Hickson, David, and C. J. McMillan 1981 *Organization and Nation: The Aston Programme IV*. Westmead, Eng.: Gower.
- Kamenka, Eugene 1989 *Bureaucracy*. Oxford: Basil Blackwell.
- Kaufman, Herbert 1977 *Red Tape*. Washington, D.C.: The Brookings Institute.
- March, James G., and Herbert A. Simon 1958 *Organizations*. New York: Wiley.
- Marx, Karl (1852) 1963 *The Eighteenth Brumaire of Louis Bonaparte*. New York: International Publishing Co.
- Mayo, Elton 1933 *The Human Problems of Industrial Civilization*. New York: Macmillan.
- Mechanic, David 1976 *The Growth of Bureaucratic Medicine: An Inquiry into the Dynamics of Patient Behavior and the Organization of Medical Care*. New York: Wiley.
- Merton, Robert K. 1940 "Bureaucratic Structure and Personality." *Social Forces* 18:560–568.
- Michels, Robert 1949 *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy*. New York: Free Press.
- Mouzelis, Nicos P. 1967 *Organization and Bureaucracy*. London: Routledge and Kegan Paul.
- Parsons, Talcott, and E. A. Shils 1951 *Toward a General Theory of Action*. Cambridge, Mass.: Harvard University Press.
- Perrow, Charles 1972 *Complex Organizations: A Critical Essay*. Glenview, Ill.: Scott Foresman.
- Peter, Laurence F., and Raymond Hull 1969 *The Peter Principle*. New York: Morrow.
- Pugh, D. S. 1981 "The Aston Programme of Research: Retrospect and Prospect." In Andrew H. Van de Ven and W. F. Joyce, eds., *Perspectives on Organization Design and Behavior*. New York: Wiley.
- , et al. 1968 "Dimensions of Organization Structure." *Administrative Science Quarterly* 13:65–105.
- Rhinestein, Max 1954 *Max Weber on Law in Economy and Society*. New York: Simon and Schuster.
- Scott, W. Richard 1992 *Organizations: Rational, Natural, and Open Systems*, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
- Taylor, Frederick W. 1911 *Principles of Scientific Management*. New York: Harper and Row.
- Thompson, James 1967 *Organizations in Action*. New York: McGraw-Hill.
- Tocqueville, Alexis de (1877) 1955 *L'ancien regime and the French Revolution*, J. P. Mayer and A. P. Kerr, eds. Garden City, N.Y.: Doubleday.
- Udy, Stanley H. Jr. 1959 "'Bureaucracy' and 'Rationality' in Weber's Organization Theory." *American Sociological Review* 24:591–595.
- Weber, Max 1946 *From Max Weber: Essays in Sociology*, H. H. Gerth and C. W. Mills, eds. London: Oxford University Press.
- Wilson, James Q. 1989 *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- 1978 *Economy and Society: An Outline of Interpretive Sociology*, Guenther Roth and Claus Wittich, eds. Berkeley: University of California Press.
- 1947 *The Theory of Social and Economic Organization*, A. M. Henderson and T. Parsons, eds. Glencoe, Ill.: Free Press.

DAVID G. NICKINOVICH



# C

---

## CAPITALISM

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

Sociology has no complete, formal consensus on a specific definition of capitalism. The discipline of sociology itself arose as an attempt to understand and explain the emergence and nature of modern capitalist societies. Sociology's founding theorists were very much concerned with the development of capitalism. Émile Durkheim sought to find the bases of new forms of morality and social solidarity in the division of labor, which capitalism both expanded and accelerated (Durkheim 1984). Karl Marx, of course, spent his adult life analyzing and criticizing capitalist society. Marx's project was guided by his hope and expectation that capitalism would be displaced as history moved toward a socialist, and then communist, future. Max Weber, too, devoted considerable attention to the origins of modern capitalism and the historically specific character of Western society under capitalist expansion. Contemporary sociology's treatment of capitalism is grounded in the works of these theorists. The works of Marx and Weber, insofar as they more explicitly focused attention on the dynamics of capitalism, provide a point of departure for discussing modern sociology's approaches to capitalism.

The term *capitalism* is sometimes used to refer to the entire social structure of a capitalist society.

Unless otherwise indicated, it is used here with specific reference to a form of economy to which multiple social institutions are effectively bound in relatively compatible ways. Weber used the term *capitalism* in a very general way: "wealth used to gain profit in commerce" (Weber 1976, p. 48). This understanding of capitalism permits the discovery of capitalism in a wide variety of social and historical settings. Weber describes this general form of capitalism in traditional India and China, ancient Babylon, Egypt, and Rome and in medieval and modern Europe. However, Weber also constructs a more specific typology that pertains to the form that capitalism has taken in more contemporary Western society. This form of capitalism is referred to as modern, or Western, capitalism. In *The Protestant Ethic and the Spirit of Capitalism*, Weber (1958 pp. 21–22) contends that this is "a very different form of capitalism which has appeared nowhere else" and that it is unique in its rational "organization of formally free labor." Other important characteristics of modern capitalism, such as the separation of business from the household and rational bookkeeping, derive their significance from this peculiar organization of labor. In this emphasis on the importance of free labor, or the creation of a labor market, Weber's definition of capitalism moves much closer to Marx's use of the term.

For Marx, it is the creation of a market for human labor that is the essence of capitalism. Marx wrote that capitalism can "spring into life only when the owner of the means of production and subsistence meets in the market with the free

laborer selling his labor power” (Marx, quoted in Sweezy 1970 pp. 56–57). The emergence of the free laborer represents the destruction of other noncapitalist economic forms. Feudal or slave economies, for example, are not characterized by the recognized right of laborers to sell their own labor power as a commodity. Simple commodity production, or economies in which laborers own their own means of production (tools, equipment, etc.), are not characterized by the need for laborers to sell their labor power as a commodity. In the latter case, Weber concurs with Marx that this freedom is only formal since such laborers are compelled to sell their labor by the “whip of hunger.”

The sociological conception of capitalism also varies with particular theoretical understandings of the nature of history. Marxists, guided by an evolutionlike vision of history, tend to see capitalism as a stage in humanity’s progressive movement to a communist future. In this manner, Marxist sociology also often refers to various phases of capitalism. Wright (1978 pp. 168–169), for example, describes six stages of capitalist development: primitive accumulation, manufacture, machinofacture, monopoly capital, advanced monopoly capital, and state-directed monopoly capitalism. The implicit assumptions of law-like forces at work in the historical process are evident in the Marxist confidence that capitalism, like all previous socioeconomic orders, will eventually be destroyed by the internal contradictions it generates. References to the current stage of capitalism as “late capitalism” (e.g., Mandel 1978), for instance, reveal a belief in the inevitability of capitalism’s demise.

The Weberian tradition, on the other hand, rejects the assumption of history’s governance by “iron laws.” This leads to a recognition of various types of capitalism but without the presumption that capitalism must eventually be eliminated. The Weberian tradition discovers in the history of Western capitalism a process of rationalization toward depersonalization, improved monetary calculability, increased specialization, and greater technical control over nature as well as over persons (Brubaker 1984). However, while the Weberian tradition can expect the *probability* of continued capitalist rationalization, it does not predict the *inevitability* of such a course for history. It is important to note that, for Weber, a transition from

capitalism to socialism would probably only further this rationalization. Such developments were seen as associated with industrial society and bureaucratic forms of domination rather than with capitalism per se.

This background permits a more detailed examination of contemporary sociology’s treatment of capitalism. Already, it can be seen that sociology’s understanding of capitalism is more specific than popular conceptions of capitalism as simply “free-market” or “free-enterprise” systems. This is especially so insofar as sociology focuses its attention on modern society. It is the emergence of a “free market” in human labor that sociology tends to recognize as the distinguishing characteristic of modern capitalism. For Durkheimian sociology, this market guides the normal division of labor that is the basis of social solidarity. In this view, the absolute freedom of such a market is necessary to generate the conditions of equal opportunity that are required to guarantee norms by which people come to accept capitalism’s highly developed division of labor. Under conditions of a truly free labor market, the stratification system is seen as legitimate since individuals attain their position through their own achievement and not by means of some ascribed status (e.g., caste, gender, race, ethnicity, nepotism) or political patronage. For Marxian sociology, it is in this labor market that the two fundamental and opposing classes of capitalism meet: the owners of the means of production or capitalists (*bourgeoisie*) and the workers (*proletariat*). In this view, the struggle between these two classes is the dynamic force behind capitalist development. For Weberian sociology, this market for human labor is necessary for the development of the advanced and superior calculability of capitalist economic action. This calculability is, in turn, a fundamental component of the rationalization process in modern Western society.

This transformation of human labor into a commodity is the force behind both capitalism’s internal dynamism as well as its outward expansion. On the one hand, capitalism is constantly driven to enhance its productivity. This compulsion of modern capitalism continuously to develop its technical capacity to produce is not driven simply by competition among capitalists but is related to the unique role that human labor plays in capitalist production. Prior to the emergence of a market for human labor, premodern forms of

capitalism exhibited no such pressure constantly to revolutionize the technical means of production. Modern capitalism's dependence on human labor as a commodity, however, demands that this cost of production be kept as low as possible. First, technological development can lower the cost of labor as a commodity by vastly increasing the production of mass consumer goods. The subsequent reduction in the cost of items like food and clothing translates into reductions in the cost of wages to sustain laborers and their families. Another means to this end is automation, the creation of technology that can replace or enhance human labor. Such technological development also permits capitalists to circumvent the natural limits of the human body to labor and the tendency of laborers to organize and demand higher wages, especially important spurs to technological development in capitalist societies characterized by a shortage of labor. However, under such conditions the capitalist's demand for profitability may limit the internal expansion of technology as a means of increasing production (i.e., capital-intensive production) in favor of an outward expansion that draws upon new sources of labor. This expansion of capitalism can take two basic forms. On the one hand, there is a drive toward proletarianization, or the inclusion of more and more of society's population segments that have previously escaped the labor market. On the other hand, there is a tendency to reach outside of the society itself toward other societies, thus incorporating ever larger regions of the world into the sphere of capitalism.

In the early development of capitalist societies, peasants are freed from feudal relations and slaves are freed from slave relations to add to the available pool of labor. This transformation is rarely a smooth one. The great revolutions in Western Europe and the U.S. Civil War forced these precapitalist classes to surrender their workers to the capitalist labor market. Another major source of labor for capitalist expansion has been independent laborers or persons who "work for themselves." Farmers who own their own land and equipment and work without hired labor are a good example of the type of self-employed producer that sociologists commonly refer to as simple commodity producers. There are other occupations in this general classification, of course. Highly skilled laborers have at times been able to

retain independence from capitalist labor markets. However, capitalism has displayed a powerful capacity to bring these laborers into the sphere of capitalist, wage labor relations. For example, carpenters, mechanics, butchers, even doctors and lawyers increasingly find themselves working for wages or a salary in a capitalist firm rather than working for themselves. From time to time, the number of self-employed appears actually to rise in certain capitalist societies (Bechhofer and Elliott 1985). This is usually the result of the introduction of some new technology or new service. Sometimes persons who strongly wish to be their "own bosses" are able to take advantage of specific market conditions or are willing to sacrifice potential income to achieve this status. But clearly, if capitalist societies are examined over the course of the last 200 to 300 years, the tendency is strongly toward increased absorption of persons into the capitalist labor market.

In recent times, labor markets in nations like the United States have found another major source of labor power in women. The traditional role of homemaker impeded the inclusion of women in the labor market. That role of women within the home has changed somewhat, but the role played by women in expanding the pool of labor available to capital has increased tremendously. According to Christensen (1987), for example, in 1960 30 percent of American mothers were employed; in 1986, 62 percent were employed. Ethnic minorities have often performed a similar function in expanding the size of the labor market. Succeeding waves of immigrants have frequently played an initially marginal role in the labor market, only to be gradually absorbed into more routine participation as time passes.

Capitalism's inherent expansionary tendencies also push the capitalist society to reach beyond the borders of the nation. This expansion occurs as capitalism seeks markets for its products but also in the search for raw materials and cheaper labor to produce goods for the home market. Eventually, capitalism may simply seek profitable investment outlets outside the nation of origin. Sociology has analyzed capitalism's transnational expansion with two general but conflicting theoretical approaches. Modernization theory views this expansion in a positive way, seeing it as a means by which undeveloped societies are enabled to begin the process of development that the

developed societies have already achieved. This theoretical orientation has been especially important in shaping development policies directed at the “third world” by many Western governments, the World Bank, and even the United Nations (Giddens 1987).

Sociology’s other basic approach to the emergence of a world-scale capitalist economy involves a more critical interpretation. This view tends to see capitalist expansion as having actually caused underdevelopment. The underdevelopment approach sees the lack of development in less-developed societies (periphery) as a consequence of systematic exploitation of their people and resources by the advanced societies (core). This process of underdevelopment is generally viewed as having occurred in three stages. The first stage, that of merchant capitalism, persisted from the sixteenth century to the late nineteenth century. Merchant capitalism, supported by military force, transferred vast amounts of wealth from the periphery to the European nations to help finance initial industrial development in what are now the advanced capitalist societies. The second stage, colonialism, persisted until about ten to fifteen years after World War II, when many colonial nations were granted formal independence. In the colonial stage, the developed societies organized economic and political institutions in the less-developed nations to serve the needs of industrial capitalism in the advanced nations. In the postcolonial period, formal political independence has been granted, but the persistent economic inequalities between developed and underdeveloped nations strongly favor the more-developed capitalist societies. Even when raw materials and finished products remain in the lesser-developed nation, the profits derived from such production are taken from the periphery and returned to the advanced core societies. Thus, the pattern of underdevelopment continues in the face of formal independence.

Traditionally, much of sociology’s attention to international capitalist expansion has focused on relations between nations. Increasingly, sociology is examining these matters with greater attention to relations between classes. This shift of emphasis reflects the increasing importance of the transnational corporation in recent times. The transnational corporation’s greater capacity to use several international sites for component production and the shift of much industrial production to

underdeveloped regions are generating a process of deindustrialization in the advanced capitalist societies. While capitalism has long been a world system, many sociologists contend that the transnational, or “stateless,” corporation has significantly less commitment or loyalty to any specific nation. Capital flows ever more rapidly throughout the world, seeking the cheapest source of labor. Modern computer technology has facilitated this trend. Asian, European, and North American capital markets are increasingly interdependent. Sociology’s shift of emphasis reflects this tendency for the U.S. capitalist to have more in common, sociologically, with the Japanese or German capitalist than the U.S. investor has in common with the American worker. Sociology’s new attention to the internationalization of capital may present a need for rethinking the usefulness of the nation as the typical boundary of a society. The emergence of the “new Europe” and the demise of state socialism in Eastern Europe and the USSR may also lead to a more flexible notion of what constitutes a society.

The preceding discussion has focused on capitalism as a specific form of economy that is defined by the expansion of a labor market in which propertyless workers sell their labor power for money. Capitalist societies are, of course, far more complex than this. There are a number of distinct economic forms that coexist with capitalism in both complementary and conflictual relations. The capitalist economy itself may be broken down into two basic sectors, one representing big business and one representing small business. Sociologists usually refer to these as the monopoly sector and the competitive sector. This dual economy is reflected in a segmented, or dual, labor market. Monopoly sector workers are more likely to be male, unionized, receive better wages and benefits, have greater job security, and work in a more clearly defined hierarchy of authority based on credentials. Competitive sector workers are less likely to possess strong credentials, more likely to be female, receive lower pay, work under more dangerous conditions, and work without union protection, benefits, or job security.

Noncapitalist economic forms also exist alongside this dual capitalist economy. The self-employed reflect a form distinct from modern capitalism that sociologists commonly refer to as simple

commodity producers or petite bourgeoisie. These people produce goods and services (commodities) for sale on the market, but they work for themselves and use only their own labor in production. Most capitalist societies also contain cooperative economic organizations that are distinct from capitalist enterprises. Cooperatives are commercial, nonprofit enterprises, owned and democratically controlled by the members. The nonprofit status and democratic distribution of control (one member, one vote) set cooperatives apart. The cooperative is an especially important form of economic organization for those simple commodity producers, like farmers, described above. All capitalist societies today also contain elements of socialist economy. Key social services, such as health care, and even some commodity production are provided by the state in many societies commonly recognized as capitalist. The United States is perhaps among the most resistant to this movement toward mixed economy. Yet even the United States has, to some extent, socialized education, mail delivery, libraries, police and fire protection, scientific research and technological development, transportation networks (e.g., highways, airports, urban transit), military production, industrial infrastructure provision, and so on. The mixed character of this economy is further indicated by the common practice of the government contracting capitalist firms to produce many goods and services. In a related way, the development of welfare has influenced the nature of capitalist society. The increased intervention of government into the economy has generated the notion of the welfare state. While popular conceptions of the welfare state tend to focus on the role that government plays in alleviating the impacts of poverty on individuals, sociology also recognizes that welfare reduces the cost of reproducing the commodity-labor. In this sense, welfare functions, in the long run, as a subsidy to capital. Further, many of the socialized sectors of capitalist economies function to ensure a profitable environment for capitalist firms (O'Connor 1973; Offe 1984). In many instances, such subsidization of capital is biased toward the largest corporations.

Some sociologists have contended that the rise of the large capitalist corporation, with its dispersed stock ownership and bureaucratic form of organization, has eroded the power of individual capitalists to control the corporation in which

they have invested their capital. Instead, it is argued, bureaucratic managers have gained control over corporate capital. Such managers are thought to be relatively free of the drive to maximize profits and are willing to accept average rates of profit. In this way, the ruthless character of earlier forms of capitalism are seen as giving way to a capitalism that is managed with broader interests in mind. Further, the dispersal of ownership and control is said to eliminate the misuse of capital in the interests of an elite and wealthy minority. This "managerialist" position complements pluralist political theory by providing greater authority for institutions of representative democracy to control the allocation of society's resources.

This view is challenged in a variety of ways. Other sociologists contend that while management may have gained some formal independence, its job requires devotion to profit maximization, and its performance is assessed on this criteria. In this view, the larger companies' executives and owners are able to use interlocking directorates and their common class background (e.g., elite schools, private clubs, policy-planning organizations) to minimize price competition and thus sustain high profit levels. Others argue that those informal ties are of less consequence than their common dependency on banks. In this view, banks, or finance capital, play a disproportionately powerful role in centralizing control over the allocation of capital resources. This position strongly opposes the managerialist arguments and those who envision an independent corporate elite whose power lies in the control of corporate bureaucracies rather than in personal wealth (Glasberg and Schwartz 1983). Further, to the extent that it is valid, this argument is particularly important given the increasing internationalization of capital discussed above, since a great deal of that development has occurred in the sphere of finance capital.

These sorts of arguments reflect an important shift in sociology's understanding of the relationship between capitalism and democracy. Sociology's Enlightenment roots provided a traditional legacy of viewing capitalism and democracy as intertwined in the process of modernization. While this parallel development is certainly recognizable in early forms of modern capitalism, more recent forms of capitalism call into question the extent to which capitalism and democracy are inevitably bound together. Indeed, recent works suggest that capitalism

and democracy are now opposed to one another (e.g., Piven and Cloward 1982). This view holds that the tremendous inequality of wealth generated by modern capitalism impedes the possibility of political equality. Even conservative writers (e.g., Huntington 1975) have noted the problematic relationship between contemporary capitalism and democracy, suggesting a retrenchment of democratic forms in defense of capitalism. The future of capitalism will be shaped by this tension with democracy, but that tension itself is located in an increasingly global economy. At the moment, the politics of capitalist society seem to lag behind the economic changes. The individual is increasingly pressured to sell his or her labor as a commodity on a world market, yet at the same time that individual remains a citizen, not of the world, but of the nation. The transnational capitalist enterprise and the flow of capital, on the other hand, are ever more free of such national borders.

(SEE ALSO: *Marxist Sociology*; *Postindustrial Society*)

## REFERENCES

- Bechhofer, F., and B. Elliott 1985 "The Petite Bourgeoisie in Late Capitalism." *Annual Review of Sociology* 11:181-207.
- Brubaker, Rogers 1984 *The Limits of Rationality*. London: Allen and Unwin.
- Christensen, Kathleen 1987 "Women and Contingent Work." *Social Policy* 17:15-18.
- Dicken, Peter 1997 "Transnational Corporations and Nation-States." *International Social Science Journal* 47:77-89.
- Durkheim, Emile 1984 *The Division of Labor in Society*. New York: Free Press.
- Giddens, Anthony 1987 *Sociology: A Brief but Critical Introduction*, 2d ed. New York: Harcourt Brace Jovanovich.
- Glasberg, Davita Silfen, and Michael Schwartz 1983 "Ownership and Control of Corporations." *Annual Review of Sociology* 9:311-332.
- Haller, William J., and Vijai P. Singh 1996 "Technology, Producer Services, and the New International Division of Labor." *Journal of Developing Societies* 12:4-18.
- Huntington, Samuel P. 1975 "The United States." In Michael Crozier, Samuel P. Huntington, and Joji Watanuki, eds., *The Crisis of Democracy: Report on the Governability of Democracies to the Trilateral Commission*. New York: New York University Press.
- London, Bruce, and Robert J. S. Ross 1995 "The Political Sociology of Foreign Direct Investment: Global Capitalism and Capital Mobility, 1965-1980." *International Journal of Comparative Sociology* 36:3-4.
- Mandel, Ernest 1978 *Late Capitalism*. London: Verso.
- Melling, Joseph 1991 "Industrial Capitalism and the Welfare of the State: The Role of Employers in the Comparative Development of Welfare States: A Review of Recent Research." *Sociology* 25:219-239.
- Morss, Elliott R. 1991 "The New Global Players: How They Compete and Collaborate." *World Development* 19:55-64.
- O'Connor, James 1973 *The Fiscal Crisis of the State*. New York: St. Martin's Press.
- Offe, Claus 1984 *Contradictions of the Welfare State*. Cambridge, Mass.: M.I.T. Press.
- Piven, Frances Fox, and Richard A. Cloward 1982 *The New Class War*. New York: Pantheon Books.
- Polychroniou, Constantine G. 1990 "Global Capitalism: A Perspective of Convergence." *International Review of Modern Sociology* 20:239-251.
- Prashad, Vijay 1994 "Contract Labor: The Latest Stage of Illiberal Capitalism." *Monthly Review* 46:19-26.
- Rogers, Hudson P., Alphonso O. Ogbuehi, and C. M. Kochunny 1995 "Ethics and Transnational Corporations in Developing Countries: A Social Contract Perspective." *Journal of Euromarketing* 4:11-38.
- Ross, Robert J. S. 1995 "The Theory of Global Capitalism: State Theory and Variants of Capitalism on a World Scale." In D. Smith and B. J. Jozsef, eds., *A New World Order? Global Transformations in the Late Twentieth Century*. Westport, Conn.: Praeger Publishers.
- Scott, Alan 1996 "Bureaucratic Revolutions and Free Market Utopias." *Economy and Society* 25:89-110.
- Sklair, Leslie 1995 *Sociology of the Global System* 2nd ed. Baltimore, Md.: Johns Hopkins University Press.
- Sklair, Leslie 1998 "The Transnational Capitalist Class and Global Capitalism: The Case of the Tobacco Industry." *Political Power and Social Theory* 12:3-43.
- Sweezy, Paul M. 1970 *The Theory of Capitalist Development*. New York: Monthly Review Press.
- Valocchi, Steve 1990 "The Capitalist State and the Origins of National Welfare Policy in Nineteenth-Century Germany." *Critical Sociology* 17:91-109.
- Weber, Max 1958 *The Protestant Ethic and the Spirit of Capitalism*. Talcott Parsons, trans. New York: Scribners.
- 1976 *The Agrarian Sociology of Ancient Civilizations*. R. I. Frank, trans. London: New Left Books.

Wright, Erik Olin 1978 *Class, Crisis, and the State*. London: New Left Books.

PATRICK H. MOONEY

## CASE STUDIES

There is a sense in which virtually every activity that we associate with sociology might be called “case studies.” These activities include the generation of samples (which are made up of individual cases) for statistical analysis, the use of empirical examples (or cases) to illustrate aspects of general sociological theories, and comparative analyses of the interconnected events (again cases) that form historical and cultural patterns. Indeed, Charles Ragin and Howard S. Becker (1992) have edited an anthology dedicated to defining case studies, entitled *What is a Case?* Historically, the answer that sociologists have usually given to this question is that case studies are in-depth analyses of single or a few communities, organizations, or persons’ lives. They involve detailed and often subtle understandings of the social organization of everyday life and persons’ experiences. Because they focus on naturally occurring events and relationships (not laboratory experiments or survey data), case studies are sometimes described as naturalistic. Case studies usually involve extensive interviews about persons’ lives, or direct observation of community or organization members’ activities, or both.

The case-study approach is not unique to sociology but is a general approach to social life that is used by social scientists (especially anthropologists and historians), psychotherapists and family therapists, and journalists. All such uses of case studies involve idiographic interpretation that emphasizes how social action and relationships are influenced by their social contexts. Case studies are unique within sociology because they require that researchers immerse themselves in the lives and concerns of the persons, communities, or organizations they study, or all of these. While case studies are based on the general scientific method and are intended to advance the scientific goals of sociology, they are also humanistic because they offer readers insight into the concerns, values, and relationships of persons making up diverse social worlds.

Social scientists use the understandings developed in case studies to introduce the general

public to the unique ways of life or problems of communities, or both, to apply and build theories, and to develop policy interventions concerned with individual and social problems. Two classic examples of how case studies may be used to achieve these ends are Elliott Liebow’s (1967) *Tally’s Corner* and Helen MacGill Hughes’s (1961) *Fantastic Lodge*. Liebow’s book reports on his experiences as a participant-observer within a poor, male, African-American urban community. Liebow describes the practical problems faced by these men in living their everyday lives, and the practical strategies they used to deal with life’s pressing problems. The study challenged many of the prevailing assumptions held by policy makers and academics during the 1960s, and has been used to reassess how the social service and mental health needs of inner city, minority groups are best addressed.

Hughes’s study, which is subtitled *The Autobiography of a Drug Addict*, details the life experiences of Janet Clark, a young white woman living in a poor urban area. Hughes describes Ms. Clark as speaking from a marginal urban world made up of drugs and drug addicts, arrest and incarceration, and the urban “sporting” life. The book raises themes that were later systematized by Howard S. Becker (1963)—the person who conducted the interviews of Janet Clark—in developing his version of the labeling approach to the sociology of deviance. The labeling perspective emphasizes that deviance is not just a matter of rule breaking, but is always created through the official responses of others. It has had profound and enduring implications for how sociologists define, study, and analyze societal responses to rule-breaking behavior.

### CASE STUDIES AND THE DEVELOPMENT OF SOCIOLOGY IN THE UNITED STATES

Sociological case studies are most associated with the ethnographic traditions established at the University of Chicago during the first half of the twentieth century. The Chicago school’s emphasis on case studies partly reflects the influence of Robert E. Park, who joined the Sociology Department in 1916 and later served as department chair. Park taught his students that case studies should emphasize how persons’ lives and the organization of communities are shaped by general social processes and structures. For example, Park analyzed

the ways in which cities develop as interrelated territories involving distinctive ways of life and opportunities. He described such territories as “natural areas” and stressed that they emerged based on social and economic competition.

Many of the best-known and most influential sociological case studies done in the United States were conducted in the 1920s and 1930s by students and faculty members at the University of Chicago who were interested in the distinctive ways of life in diverse natural areas (e.g., Anderson 1923; Cressey 1932; Shaw 1930; Thomas 1923; Wirth 1928; and Zorbaugh 1929). While it was done later, H. M. Hughes’s (1961) study of Janet Clark’s life is also an example of the Chicago approach to case studies and urban life. Perhaps the best-known and most influential case study done in the early Chicago school tradition is William Whyte’s (1943) *Street Corner Society*, a participant-observer study of a poor Italian-American community. The significance of the study is related to Whyte’s use of his observations to identify and explicate some basic sociological issues involving social relations and social control in small groups.

Park’s approach to case studies has been modified and refined over the years, particularly by Everett C. Hughes (1970), who developed a comparative approach to work groups and settings. The approach uses case studies to identify and analyze comparatively generalized aspects of work groups and settings, such as work groups’ definitions of “dirty work” and their interest in controlling the conditions of their work. Becker’s (1963) use of multiple case studies in articulating the labeling perspective is a notable example of how the comparative strategy advocated by E. C. Hughes may be applied to develop sociological theory. Another major contributor to the Chicago school of sociology is Erving Goffman (1959), who used case studies to develop a dramaturgical perspective on social interaction. The perspective treats mundane interactions as quasi-theatrical performances involving scripts, stages, and characters.

Although less influential than the Chicago school, a second source for case studies in American sociology was structural-functionalist theorists who analyzed communities and organizations as stability-seeking social systems. Structural-functional case studies were influenced by anthropological studies of nonindustrial communities and the more

abstract theories of Talcott Parsons and Robert Merton. These studies analyze how social systems are maintained and adapt to changing environmental circumstances. For example, structural functionalists have used case studies to analyze the consequences of organizational activities and relationships for maintaining organizational systems (Blau 1955; Gouldner 1954; Sykes 1958).

An exemplary structural-functionalist case study is Rosabeth Moss Kanter’s (1977) analysis of a multinational corporation, which she calls Indsco. The study examines how social relations in corporations are shaped by social structures which produce feelings of uncertainty and marginality among managers and secretaries. Kanter uses her case study to illustrate how gender segregation is produced and maintained in Indsco as secretaries and managers cope with the practical problems emergent from corporate power and opportunity structures. She also makes some practical suggestions for altering these structures in order to better address the needs of managers and secretaries, and to produce more egalitarian relations between corporate members. Many of Kanter’s suggestions have been adopted by diverse American corporations. This is another way in which case studies may influence social policies.

#### RECENT DEVELOPMENTS IN CASE-STUDIES RESEARCH

While many qualitative sociologists continue to work within the Chicago school and structural-functional traditions, several important developments have occurred in qualitative sociology in the past twenty-five years. For example, many of the case studies of ethnic communities done by early Chicago school sociologists deal with the problems and social organization of European-American communities. Many more recent case studies, on the other hand, explore these issues within the context of nonEuropean-American communities. Ruth Horowitz’s (1983) case study of a Hispanic community and Elijah Anderson’s (1990) research in an African-American community are major contributions to this development.

Age-based communities have also emerged as subjects for case studies by sociologists. One focus in this literature involves the distinctive life circumstances and coping strategies of communities of older people. Arlie Hochschild’s (1973) study of



a group of widowed women as an “unexpected community” is an important example of this focus, as is Jennie Keith’s (1977) analysis of the social construction of community in a retirement facility in France. Another approach to case-study research is Jaber Gubrium’s (1993) analysis of the life stories told to him by a group of elderly nursing home residents. Equally significant are recent case studies of communities of young people, particularly of communities organized around shared interests in popular culture. A useful example is Sarah Thornton’s (1996) case study of dance clubs and raves in London. Thornton’s analysis extends traditional analysis of communities as subcultures by showing how this community was created and is maintained through the actions of (not so youthful) members of the London mass media.

A related change has been case studies of experiential communities. That is, social groups that share common social experiences, but not a common territory. For example, Richard Majors and Janet Mancini Billson (1992) analyze African-American males as a racial and gender community that is organized around shared practical problems and strategies for managing them. Using a semiotic perspective, Dick Hedbridge (1979) also analyzes an experiential community in using several case studies to show how young people form a subculture that is organized around the expression of cultural styles, which symbolically resist dominant forces in society. Another focus in this literature is on the distinctive life experiences of members of gay and lesbian communities. A notable example is Carol A. B. Warren’s (1974) case study of a gay community in which she raises some important questions about the labeling perspective. Thus, both Warren’s and Hedbridge’s research show how case studies contribute to theoretical development in sociology.

A different, but complementary, use of the case-study research strategy is auto-ethnography. These are case studies conducted within and about one’s own group. The researcher acts as both a sociologist seeking information about a community, and as a subject of the research. A groundbreaking auto-ethnography is David Hayano’s (1982) study of professional card players. This case study provides readers with distinctive access to, and insights into, the experiences of professional card players. Two other notable auto-ethnographies are Helen Rose Fuchs Ebaugh’s (1977) case study

of her own and others’ exit from the social role of Roman Catholic nun, and Thomas Schmid’s and Richard S. Jones’s (1991) analysis of Jones’s and other inmates’ adaptations to prison life.

Auto-ethnographers’ concern for taking account of their own and others’ experiences is extended in another trend in sociological case-studies research. This trend focuses on emotions as a basic and important aspect of social life. Many case studies done by sociologists of emotions deal with other persons’ emotional experiences (Hochschild 1983; Katz 1988). Other sociologists, however, adopt a more auto-ethnographic strategy by treating their own lives and feelings as topics for sociological analysis. The sociologist’s feelings become the case or part of the case under study. An important example of the latter approach to case studies is David Karp’s (1996) discussion of his own experiences with depression, and his linking of them to the emotional experiences reported by other members of this community of sufferers.

In addition to the changes discussed above, three other changes in case-studies research warrant special notice. They are the emergence of radical case studies, a focus on reality construction, and concern for the politics and poetics of writing case studies.

### **RADICAL CASE STUDIES**

While early sociological case studies were sometimes associated with reformist movements, they were seldom intended to advance politically radical causes or to build radical social theory. This orientation may be contrasted with that of radical journalists of the same era (such as Upton Sinclair) who used case studies to highlight social problems and raise questions about the legitimacy of capitalism. Radical sociologists have begun to use case studies to develop their sociological goals. One source for radical case studies is Marxist sociologists concerned with the ways in which worker-management relations are organized in work settings, how social classes are perpetuated, and how capitalism is justified.

Marxist sociologists use their case studies to challenge more conservative case studies that, from the radicals’ standpoint, do not take adequate account of the ways in which persons’ everyday

lives are shaped by political and economic structures. They are also used to analyze historical changes in capitalism and the consequences of the changes for workers. For Marxist theorists, a major change has been the rise of a new of society—monopoly capitalism—dominated by multinational corporations. These theorists use case studies of work settings to illustrate and analyze the effects of monopoly capitalism on workers (Burawoy 1970).

Although it is not as well developed in sociology as in anthropology, another focus of radical case studies involves the social and personal consequences of the international division of labor. The studies are concerned with the ways in which multinational corporations internationalize the production process by exporting aspects of production to Third World countries. Case studies of this trend show the profound impact of global economic changes for gender and family roles in Third World countries (Ong 1987). Finally, radical sociologists have used the case-study method to analyze the ways in which capitalist institutions and relationships are justified and perpetuated by noneconomic institutions such as schools (Willis 1977).

A second source for radical case studies is feminist sociology. While it is diverse and includes members who hold many different political philosophies, all forms of feminist sociology are sensitive to the politics of human relationships. A major theme in feminist case studies involves the ways in which women's contributions to social relationships and institutions go unseen and unacknowledged. Feminist sociologists use case studies to call attention to women's contributions to society and analyze the political implications of their invisibility. A related concern involves analyzing relationships and activities that are typically treated as apolitical and private matters as matters of public concern. One way in which feminist sociologists do so is by treating aspects of their own lives as politically significant and making them matters for sociological analysis (McCall 1993).

The case-study approach is central to the feminist sociology of Dorothy E. Smith (1987). Smith treats case studies as points of entry for studying general social processes that shape persons' experiences and lives. Her approach to case studies emphasizes how the seemingly insignificant activities of everyday life are related by general social

processes (such as patriarchy and market relationships) and how they help to perpetuate the processes. Smith describes, for example, how the commonplace activity of dining in a restaurant is organized within, and perpetuates, capitalist commodity relations. Smith's analysis might also be seen as an example of the general analytic strategy, which Michael Burawoy (1998) calls the "extended case method." This approach to case studies involves four major steps:

- Researcher observations of, and immersion in, a social setting,
- Analysis of the researcher's observations as aspects of general social processes that shape life in diverse contemporary social settings,
- Historicizing the observations and social processes by showing how they are embedded in historical forces that guide and structure the evolution of capitalist society, and
- Linking the observations and analyses to formal sociological theories that explain the researcher's initial observations and their larger social and historical contexts.

This is one way in which case studies may be used as a springboard for developing systematic, general, and formal sociological analyses.

### **CASE STUDIES OF REALITY CONSTRUCTION**

Basic to the case-study method and idiographic interpretation is a concern for human values and culture. Beginning in the 1960s, one focus of this concern has been with the ways in which social realities are produced in social interactions. The new focus is generally based on the social phenomenology of Alfred Schutz and ethnomethodology. While different in some ways, both are concerned with the folk methods that people use to construct meanings (Silverman 1975). For example, Harold Garfinkel (1967), the founder of ethnomethodology, used a case study of a patient seeking a sex-change operation to analyze how we orient to ourselves and others as men and women.

Case studies of reality construction emphasize how social realities are created, sustained, and

changed through language use. The emphasis has given rise to new orientations to traditional areas of sociological research, such as science. Until recently, sociologists of science have treated scientific work as a simple, noninterpretive process centered in scientists' adherence to the rules of the scientific method, which emphasizes how "facts" and "truth" emerge from observations of the "real" world. Viewed this way, scientific facts are not matters of interpretation or social constructions. This view of science has been challenged by case studies of scientific work that focus on the ways in which scientific facts are socially produced based on scientists' interpretations of their experiments (Latour and Woolgar 1979).

A major branch of ethnomethodology is conversation analysis, which focuses on the turn-by-turn organization of social interactions (Sacks, Schegloff, and Jefferson 1974). For conversation analysts, social reality is collaboratively constructed within turn-taking sequences, which may be organized around such linguistic practices as questions and answers or charges and rebuttals. A major contribution of conversation-analytic studies has been to show how IQ and other test results are shaped by the ways in which test-givers and test-takers interact (Marlaire and Maynard 1990). Also, sociologists influenced by conversation analysis theory have considered how power and dominance are interactionally organized and accomplished.

Some qualitative sociologists have extended and reformulated ethnomethodological case studies by analyzing the relationships between the interpretive methods used by interactants in concrete situations and the distinctive meanings they produce in their interactions. This approach to case studies focuses on the practical and political uses of meanings in social institutions. These ethnographers analyze meanings as rhetorics which interactants use to persuade others and to assign identities to themselves and others. An important contribution to this approach to case-study research is Donileen Loseke's (1992) study of decision making and social relations in a shelter for battered women. Loseke details the practical contexts within which shelter workers, shelter residents, and applicants to the shelters encounter each other and manage their social relationships. It is in these encounters, Loseke states, that shelter

workers and residents give concrete meaning to the abstract cultural category of battered women.

Two other important, and related, trends in case studies of reality construction involve comparative analysis of two or more cases, and the use of new theoretical perspectives in analyzing case-study data. One example is Jaber Gubrium's (1992) comparative analysis of two different family therapy sites. Gubrium uses observational data from the sites to show how different therapy approaches are organized to "detect" and remedy different kinds of family and personal problems. He also shows how aspects of Weberian theory may be used in analyzing the sociological significance of family therapy. Gale Miller and David Silverman (1995) have also contributed to this development by comparatively analyzing social interactions in two different counseling centers located in the United States and England. This study is based on the ethnography of institutional discourse perspective (Miller 1994), which is a strategy for developing general theoretical statements from case-study research.

#### **THE POLITICS AND POETICS OF WRITING CASE STUDIES**

While some qualitative sociologists focus on the folk methods of description and interpretation used by others in creating realities, other sociologists are reconsidering the reality constructing methods used by sociologists in writing case studies. Their interest centers in the question, How do we do our work? The question raises issues about the relationships between qualitative researchers and the people they study. For example, is it enough that case studies inform the public and sociologists about aspects of contemporary society, or should they also help the persons and communities studied? A related issue involves editorial control over the writing of case studies. That is, should the subjects of case studies have a voice in how they are described and analyzed?

Such questions have given rise to many answers, but most of them involve analyzing case studies as narratives or stories that sociologists tell about themselves and others. For example, John van Maanen (1988) divides ethnographic writing into several types of "tales" involving different

orientations to the persons described, readers, and authorship. Other analyses focus on the various rhetorical devices used by ethnographers to write their case studies (Bruyn 1966). Such rhetorical devices include metaphor, irony, paradox, synecdoche, and metonymy, which ethnographers use both to describe others and to cast their descriptions as objective and authoritative.

Qualitative sociologists' interest in writing case studies may be part of a larger interdisciplinary movement involving social scientists and humanists. The movement is concerned with analyzing the rhetorics of social scientific inquiry (Nelson et al. 1987) as well as how to write better narratives. The latter issue is basic to efforts by British sociologists to develop new forms of writing that better reflect the ways in which their case studies are socially constructed (Woolgar 1988) and other sociologists' interest in developing alternatives to the logico-scientific writing style that better express their theoretical perspectives, political philosophies, and experiences (DeVault 1990; Richardson 1990). A related change in this area is some sociologists' interest in using performance art to represent their research data. For example, Carolyn Ellis and Arthur Bochner (1992) have written a play about the experiences and dilemmas faced by women and their male partners in deciding whether to abort the woman's pregnancy. Ellis and Bochner state that this presentational form allows for a wider range of communication devices than does the usual textual approach to reporting case-study findings.

In sum, the case-study method is a dynamic approach to studying social life, which sociologists modify and use to achieve diverse political and theoretical goals. While the popularity of the case-study method waxes and wanes over time, it is likely to always be a major research strategy of humanistically oriented sociologists.

(SEE ALSO: *Ethnomethodology*; *Field Research Methods*; *Life Histories*; *Qualitative Methods*)

## REFERENCES

Anderson, Elijah 1990 *Streetwise: Race, Class, and Change in an Urban Community*. Chicago: University of Chicago Press.

Anderson, Nels 1923 *The Hobo*. Chicago: University of Chicago Press.

Becker, Howard S. 1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.

Blau, Peter M. 1955 *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press.

Bruyn, Severyn T. 1966 *The Human Perspective in Sociology*. Englewood Cliffs, N.J.: Prentice-Hall.

Burawoy, Michael 1970 *Manufacturing Consent*. Chicago: University of Chicago Press.

—1998 "Critical Sociology: A Dialogue Between Two Sciences." *Contemporary Journal of Sociology* 27:12–20.

Cressey, Paul 1932 *The Taxi-Dance Hall*. Chicago: University of Chicago Press.

DeVault, Marjorie L. 1990 "Talking and Listening from Women's Standpoint." *Social Problems* 37:96–116.

Ebaugh, Helen Rose Fuchs 1997 *Out of the Cloister: A Study of Organizational Dilemmas*. Austin: The University of Texas Press.

Ellis, Carolyn and Arthur P. Bochner 1992 "Investigating Subjectivity." In Carolyn Ellis and Michael G. Flaherty, eds., *Telling and Performing Personal Stories*. Newbury Park, Calif.: Sage Publications.

Garfinkel, Harold 1967 *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.

Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, N.Y.: Doubleday.

Gouldner, Alvin 1954 *Patterns of Industrial Democracy*. New York: Free Press.

Gubrium, Jaber F. 1992 *Out of Control: Family Therapy and Domestic Disorder*. Newbury Park, Calif.: Sage.

— 1993 *Speaking of Life: Horizons of Meaning for Nursing Home Residents*. Hawthorne, N.Y.: Aldine de Gruyter.

Hayano, David M. 1982 *Poker Faces: The Life and Work of Professional Card Players*. Berkeley: University of California Press.

Hedbridge, Dick 1979 *Subculture: The Meaning of Style*. New York: London.

Hochschild, Arlie Russell 1973 *The Unexpected Community*. Englewood Cliffs, N.J.: Prentice-Hall.

— 1983 *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press.

Horowitz, Ruth 1983 *Honor and the American Dream: Culture and Identity in a Chicano Community*. New Brunswick, N.J.: Rutgers University Press.

- Hughes, Everett C. 1970 "The Humble and the Proud." *The Sociological Quarterly* 11:147-156.
- Hughes, Helen MacGill 1961 *The Fantastic Lodge: The Autobiography of a Drug Addict*. Greenwich, Conn.: A Fawcett Premier Book.
- Kanter, Rosabeth Moss 1977 *Men and Women of the Corporation*. New York: Basic Books.
- Katz, Jack 1988 *Seductions of Crime: Moral and Sensual Attractions in Doing Evil*. New York: Basic Books.
- Keith, Jennie 1973 *Old People, New Lives: Community Creation in a Retirement Residence*. Chicago: University of Chicago Press.
- Latour, Bruno and Steven Woolgar 1979 *Laboratory Life*. Beverly Hills, Calif.: Sage Publications.
- Liebow, Elliot 1967 *Tally's Corner: A Study of Negro Streetcorner Men*. Boston: Little, Brown.
- Loseke, Donileen R. 1992 *The Battered Woman and Shelters: The Social Construction of Wife Abuse*. Albany, N.Y.: SUNY Press.
- Majors, Richard, and Janet Mancini Billson 1992 *Cool Pose: The Dilemmas of Black Manhood in America*. New York: Touchstone Books.
- Marlaire, Courtney L., and Douglas W. Maynard 1990 "Standardized Testing as an Interactional Phenomenon." *Sociology of Education* 63:83-101.
- McCall, Michal A. 1993 "Social Constructionism and Critical Feminist Theory and Research." In Gale Miller and James A. Holstein, eds., *Constructionist Controversies: Issues in Social Problems Theory*. Hawthorne, N.Y.: Aldine de Gruyter.
- Miller, Gale 1994 "Toward Ethnographies of Institutional Discourse: Proposal and Suggestions." *Journal of Contemporary Ethnography* 23:280-306.
- , and David Silverman 1995 "Troubles Talk and Counseling Discourse: A Comparative Study." *Sociological Quarterly* 36:725-747.
- Nelson, John S., Allan Megill, and Donald N. McCloskey (eds.) 1987 *The Rhetoric of the Human Sciences*. Madison: University of Wisconsin Press.
- Ong, Aihwa 1987 *Spirits of Resistance and Capitalist Discipline*. Albany, N.Y.: SUNY Press.
- Ragin, Charles C. and Howard S. Becker (eds.) 1992 *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge, UK: Cambridge University Press.
- Richardson, Laurel 1990 "Narrative and Sociology." *Journal of Contemporary Ethnography* 19:116-135.
- Sacks, Harvey, Emmanuel A. Schegloff, and Gail Jefferson 1974 "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50:696-735.
- Schmid, Thomas, and Richard S. Jones 1991 "Suspended Identity Dialectic: Identity Transformation in a Maximum Security Prison." *Symbolic Interaction* 14:415-432.
- Shaw, Clifford 1930 *The Jack-Roller*. Chicago: University of Chicago Press.
- Silverman, David 1975 "Accounts of Organizations: Organizational 'Structures' and the Accounting Process." In John B. McKinlay, ed., *Processing People: Cases in Organizational Behavior*. London: Holt, Rinehart and Winston.
- Smith, Dorothy E. 1987 *The Everyday World as Problematic*. Boston: Northeastern University Press.
- Sykes, Gresham M. 1958 *The Society of Captives*. Princeton, N.J.: Princeton University Press.
- Thomas, W. I. 1923 *The Unadjusted Girl*. Boston: Little, Brown.
- Thornton, Sarah 1996 *Club Cultures: Music, Media and Subcultural Capital*. Hanover, N.H. and London: Wesleyan University Press.
- Van Maanen, John 1988 *Tales of the Field*. Chicago: University of Chicago Press.
- Warren, Carol A. B. 1974 *Identity and Community in the Gay World*. New York: John Wiley & Sons.
- Whyte, William Foote 1943 *Street Corner Society: The Social Structure of an Italian Slum*. Chicago: University of Chicago Press.
- Willis, Paul 1977 *Learning to Labour*. New York: Columbia University Press.
- Wirth, Louis 1928 *The Ghetto*. Chicago: University of Chicago Press.
- Woolgar, Steven 1988 *Science, the Very Idea*. New York: Tavistock.
- Zorbaugh, Harvey 1929 *The Gold Coast and the Slum*. Chicago: University of Chicago Press.

GALE MILLER  
RICHARD S. JONES

## CASTE AND INHERITED STATUS

The study of social inequalities is one of the most important areas of sociology as inequalities in property, power, and prestige have long-term consequences on access to the basic necessities of life

such as good health, education, and a well-paying job. Unequal resources between groups also influence their interaction with all the basic institutions of society: the economy, the political system, and religion, among others. An understanding of how inequalities arise, their forms and consequences, and the processes that tend to sustain them is essential for the formulation of policies that will improve the well-being of all in society.

All societies in the world are socially stratified (i.e. wealth, power, and honor are unequally distributed among different groups). However, they vary in the ways in which inequality is structured. One of the most frequently used bases for categorizing different forms of stratification systems is the way status is acquired. In some societies, individuals acquire status on the basis of their achievements or merit. In others, status is accorded on the basis of ascribed, not achieved characteristics. One is born into them or inherits them, regardless of individual abilities or skills. A person's position is unalterable during his or her lifetime. The most easily understood example is that of the prince who inherits the status of king because he is the son of a king.

Sociologists who study stratification use the idea of ascribed and achieved status to contrast caste systems with class systems. In class systems one's opportunities in life, at least in theory, are determined by one's actions, allowing a degree of individual mobility that is not possible in caste systems. In caste systems a person's social position is determined by birth, and social intercourse outside one's caste is prohibited.

The term "caste" itself is often used to denote large-scale kinship groups that are hierarchically organized within a rigid system of stratification. Caste systems are to be found among the Hindus in India and among societies where groups are ranked and closed as in South Africa during the apartheid period. Toward the end of this article, examples of caste systems will be given from other non-Hindu societies to illustrate how rigid, ranked systems of inequality, where one's position is fixed for life, are found in other areas that do not share India's religious belief system.

Early Hindu literary classics describe a society divided into four *varnas*: Brahman (poet-priest),

Kshatriya (warrior-chief), Vaishya (traders), and Shudras (menials, servants). The *varnas* formed ranked categories characterized by differential access to spiritual and material privileges. It excluded the Untouchables, who were despised because they engaged in occupations that were considered unclean and polluting.

The *varna* model of social ranking persisted throughout the Hindu subcontinent for over a millennia. The basis of caste ranking was the sacred concept of purity and pollution with Brahmins, because they were engaged in priestly duties considered ritually pure, while those who engaged in manual labor and with ritually polluting objects were regarded as impure. Usually those who had high ritual status also had economic and political power. Beliefs about pollution generally regulated all relations between castes. Members were not allowed to marry outside their caste; there were strict rules about the kind of food and drink one could accept and from what castes; and there were restrictions on approaching and visiting members of another caste. Violations of these rules entailed purifactory rites and sometimes expulsion from the caste (Ghurye 1969).

The *varna* scheme refers only to broad categories of society, for in reality the small endogamous group or subcaste (*jati*) forms the unit of social organization. In each linguistic area there are about two thousand such subcastes. The status of the subcaste, its cultural traditions, and its numerical strength vary from one region to another, often from village to village.

Field studies of local caste structures revealed that the caste system was more dynamic than the earlier works by social scientists had indicated. For example, at the local level, the position of the middle castes, between the Brahmins and the Untouchables, is often not very clear. This is because castes were often able to change their ritual position after they had acquired economic and political power. A low caste would adopt the Brahminic way of life, such as vegetarianism and teetotalism, and in several generations attain a higher position in the hierarchy. Upward mobility occurred for an entire caste, not for an individual or the family. This process of upward mobility, known as Sanskritization (Srinivas 1962), did not however, affect the movement of castes at the

extremes. Brahmans in most parts of the country were found at the top, and Untouchables everywhere occupied a degrading status because of their economic dependency and low ritual status.

The operation of this hierarchical society was justified with reference to traditional Hindu religious beliefs about *samsara* (reincarnation) and *karma* (quality of actions). A person's position in this life was determined by his or her actions in previous lives. Persons who were born in a Brahman family must have performed good deeds in their earlier lives. Being born a Shudra or an Untouchable was punishment for the sinful acts committed in previous lives.

Some scholars (Leach 1960; Dumont 1970) saw the caste system as a cooperative, inclusive arrangement where each caste formed an integral part of the local socioeconomic system and had its special privileges. In a *jajmani* system, as this arrangement between castes was known, a village was controlled by a dominant caste, which used its wealth, numerical majority, and high status to dominate the other castes in the village. Most other castes provided services to this caste. Some worked on the land as laborers and tenants. Others provided goods and services to the landowning households and to other castes. A village would thus have a potter, blacksmith, carpenter, tailor, shoemaker, barber, sweeper, and a washerman, with each caste specializing in different occupations. These were hereditary occupations. In return for their services castes would be paid in kind, usually farm produce. These patron-client relationships continued for generations, and it was the religious duty of the *jajman* (patron) to take care of others.

Although the system did provide security for all, it was essentially exploitative and oppressive (Berreman 1981; Beidelman 1959; Freeman 1986), particularly for the Untouchables, who were confined to menial, despised jobs, working as sweepers, gutter and latrine cleaners, scavengers, watchmen, farm laborers, and curers of hides. They were denied access to Hindu temples; were not allowed to read religious Sanskrit books and remained illiterate; could not use village wells and tanks; were forced to live in settlements outside the village; and were forbidden to enter the residential areas of the upper castes.

## CHANGES IN THE CASTE SYSTEM

British rule profoundly affected the Indian social order. The ideas of Western culture; the opening of English educational institutions; the legal system, which introduced the principle of equality before the law; and the new economic activities and the kind of employment they generated all brought new opportunities for greater advancement. Although these new developments resulted in greater mobility and opened doors for even the low castes, those castes that benefited most were the ones already in advantageous positions. Thus, Brahmans with a tradition of literacy were the first to avail themselves of English education and occupy administrative positions in the colonial bureaucracy.

The spread of communications enabled local subcastes to link together and form caste associations. These organizations, although initially concerned with raising the caste status in terms of Brahmanical values, later sought educational, economic, and social benefits from the British (Rudolph and Rudolph 1960). When the colonial authorities widened political participation by allowing elections in some provinces, castes organized to make claims for political representation. In some regions, such as the South, the non-Brahman castes were successful in restricting entry of Brahmans in educational institutions and administrative services.

To assuage the fears of communities about upper-caste Hindu rule in independent India and also to weaken the nationalist movement, the British granted special political representation to some groups such as the Untouchables. They had become politically mobilized under the leadership of Dr. B. R. Ambedkar and had learned, like other castes and communities, the use of political means to secure status and power (Zelliot 1970).

After the country became independent from British rule in 1947, the Indian leaders hoped that legislative and legal measures would reorder an entrenched social structure. A new Constitution was adopted, which abolished untouchability and prohibited discrimination in public places. In addition, special places were reserved for Untouchables in higher educational institutions, government services, and in the lower houses of the central and state legislatures.

What progress has the country made toward improving the lives of the Untouchables, who now form 16.48 percent (according to the 1991 Indian Census) of the population? Has the traditional caste system disintegrated?

The movement from a traditional to a modern economy—increase in educational facilities; expansion of white-collar jobs, especially in the state sector; expansion of the transportation and communication networks; increase in agricultural production (known as the Green Revolution)—has had a significant impact on the institution of caste. However, political factors have been equally if not more important in producing changes in the caste system. One is the democratic electoral system. The other is the state's impact on intercaste relations through its policy of preferences for selected disadvantaged castes.

The close association between caste and traditional occupation is breaking down because of the expansion of modern education and the urban-industrial sector. In India, an urban middle class has formed whose members are drawn from various caste groups. This has reduced the structural and cultural differences between castes, as divisions based on income, education, and occupation become more important than caste cleavages for social and economic purposes. However, the reduction is most prominent among the upper socioeconomic strata—the urban, Western-educated, professional, and higher-income groups, whose members share a common lifestyle (Beteille 1969).

For most Indians, especially those who live in rural areas (73 percent of the Indian population is still rural), caste factors are an integral part of their daily lives. In many parts of the country Dalits (the term means “oppressed” and is now preferred by the members of the Untouchable community rather than the government-assigned label “Scheduled Castes”) are not allowed inside temples and cannot use village water wells. Marriages are generally arranged between persons of the same caste.

With the support of government scholarships and reservation benefits, a small proportion of Dalits has managed to gain entry into the middle class—as school teachers, clerks, bank tellers, typists, and government officials. Reservation of seats in the legislature has made the political arena

somewhat more accessible, although most politicians belonging to the Dalit community have little say in party matters and government policymaking. The majority of Dalits remain landless agricultural laborers, powerless, desperately poor, and illiterate.

Modern economic forces are changing the rural landscape. The increase in cash-crop production, which has made grain payments in exchange for services unprofitable; the introduction of mechanized farming which has displaced manual labor; the preference for manufactured goods to handmade ones; and the migration to cities and to prosperous agricultural areas for work and better wages have all weakened the traditional patron-client ties and the security it provided. The Dalits and other low castes have been particularly affected as the other sectors of the economy have not grown fast enough to absorb them.

The rural social structure has been transformed in yet another way. The dominant castes are no longer from the higher castes but belong to the middle and lower peasant castes—the profit maximizing “bullock capitalists” (Rudolph and Rudolph 1987) who were the chief beneficiaries of land reform and state subsidies to the agricultural sector (Blair 1980; Brass 1985). They have displaced the high-caste absentee landlords who have moved to cities and taken up modern occupations.

Modern political institutions have also brought about changes in the traditional leadership and power structure of local communities. Relations between castes are now governed by rules of competitive politics, and leaders are selected for their political skills and not because they are members of a particular caste. The role of caste varies at different levels of political action. At the village and district levels caste loyalties are effectively used for political mobilization. But at the state and national levels, caste factors become less important for political parties because one caste rarely commands a majority at these levels. The rise of a Dalit political party, the Bahujan Samaj Party, is evidence that Dalits are finally gaining some political power. They are particularly strong in the northern Indian state of Uttar Pradesh where they received 20.61 percent of the votes in the 1996 general elections. However, at the national level they have fared poorly, capturing only 11 seats (and 3.64 percent of the votes) in the 1996 general elections.



In the 1990s there were numerous instances of confrontations between the middle peasant castes and Dalits in rural areas. Violence and repression against Dalits has increased as they have begun to assert themselves. With the support of Communist and Dalit movements, they are demanding better wages, the right to till government-granted land, and the use of village wells.

In urban areas, caste conflict has mainly centered around the issue of "reservation." The other backward castes (who belong mainly to the Shudra caste and form about 50 percent of the country's population) have demanded from the government benefits similar to those available to Dalits in government service and educational institutions. Under electoral pressures the state governments have extended these reservation benefits to the other backward castes, leading to discontent among the upper castes.

Extension of preferential treatment from Dalits to the more numerous and in some states somewhat better-off backward castes has not only created great resentment among the upper castes but also has reduced public support for the policy of special benefits for the Dalits. In cities they have often been victims during anti-reservation agitations. That this is happening at the very time when the preferential programs have gradually succeeded in improving the educational and economic conditions for Dalits is not accidental (Sheth 1987).

As education and the meaning of the vote and the ideas of equality and justice spread, the rural and urban areas will witness severe intercaste conflicts. What is significant, however, is that these conflicts are not over caste beliefs and values, but like conflicts elsewhere between ethnic groups, have to do with control over political and economic resources.

#### CASTE IN OTHER SOCIETIES

Do castes exist outside India? Is it a unique social phenomenon distinct from other systems of social stratification? Opinion among scholars is divided over this issue. Castelike systems have been observed in the South Asian subcontinent and beyond (in Japan, Africa, Iran, and Polynesia). Caste has also been used to describe the systems of racial stratification in South Africa and the southern United States.

Whether the term "caste" is applicable to societies outside the South Asian region depends on how the term is defined. Those who focus on its religious foundations argue that caste is a particular species of structural organization found only in the Indian world. Louis Dumont (1970), for example, contends that caste systems are noncomparable to systems of racial stratification because of the differences in ideology—one based on the ideology of hierarchy, the other on an equalitarian ideology.

Cultural differences notwithstanding, caste as a ranked system exists in many societies. In fact, wherever ethnic groups stand in a hierarchical or ranked relationship to each other they resemble castes (Weber 1958; Horowitz 1985; Berreman 1981). As in caste systems, the identity of an ethnic group is regarded as being a consequence of birth or ancestry and hence immutable; mobility opportunities are restricted; and members of the subordinate group retain their low social position in all sectors of society—political, economic, and social. Social interactions between groups remain limited and are suffused with deference. Given these similarities in their social structures and social processes, caste stratification is congruent with race stratification and ranked ethnic systems. Below are some examples of castelike systems in countries outside South Asia.

In Japan, during the Tokugawa period (from the early 1600s to the middle 1800s), the Shogun rulers established a very rigid, hierarchical system that was maintained by force of law and other means. At the top were the *shogunate* warrior-bureaucrats, their *samurai* military elite, and the higher aristocracy. This was followed by peasants, then artisans, and then merchants. At the bottom and separated from the rest of the populace there was a group of outcasts called *eta* (meaning heavily polluted) or *hinin* (meaning nonhuman) who were treated much like the Indian Untouchables described above. *Eta* were legally barred from marrying outside their group or from living outside their designated hamlets. These hamlets were called *buraku*, and their residents known as *Burakumins*.

*Burakumins* are indistinguishable in appearance from other Japanese. They faced discrimination because they inherited their status from people whose jobs were considered polluting and undesirable like butchering animals, tanning skins,

digging graves, handling corpses, and guarding tombs.

The outcasts were formally emancipated in 1871, by the Meiji government (1868–1912). The descendants of *Burakumins* were identified as “new common people.” However, they continued to face discrimination as their identity could be revealed through the household register system that included the ancestry of all Japanese families. Although now the household register is not made available to the public without the permission of the family, the identity of individuals is frequently revealed when families and employers conduct investigations for marriage purposes and hiring (Ishida 1992).

Rwanda, a country just south of the Equator in Africa and which has witnessed violent ethnic conflict is another example of a system of caste stratification. Before European colonization, political power was concentrated in the hands of the king and the pastoral aristocracy (Tutsi). The Tutsis constituted only about 10 percent of the population. Hutus, the lower caste of agriculturalists, formed the vast majority of the population. The lowest caste, known as Twa, were a small minority and worked as potters, court jesters, and hunters. No intermarriage was permitted between the groups. The Tutsis used their political and military power to maintain the hierarchical system (Southall 1970), particularly in central parts of the country. This hierarchical relationship was later reinforced under colonial rule and lasted until it was brought to an end in the 1950s (Newbury 1988).

As Gerald Berreman (1981) has argued, the blacks in America and South Africa, the Burakumin of Japan, the Dalit of India, and the Hutu and Twa of Rwanda, all live in societies that are alike in their structure and in their effect on the life experiences of those most oppressed. However, those at the bottom do not accept their condition willingly. Often too powerless to revolt openly, structurally similar forms of domination create common forms of *infrapolitics*—of surreptitious resistance. Rituals of aggression, tales of revenge, use of carnival symbolism, gossip, and rumor are all examples of the strategic form of resistance the subordinates use for open defiance under severely repressive conditions (Scott 1990). Finally, similarities also exist in the political consequences of preferential policies that culturally distinct societies such as the

United States and India have adopted to reduce group disparities (Weiner 1983).

REFERENCES

- Beidelman, T. O. 1959 *A Comparative Analysis of the Jajmani System*. Monograph VII of the Association of Asian Studies. Locust Valley, N.Y.: J.J. Augustin.
- Berreman, Gerald 1981 *Caste and Other Inequities* Delhi: Manohar.
- Beteille, Andre 1969 *Caste: Old and New*. Bombay: Asia Publishing House.
- Blair, Harry W. 1980 “Rising Kulaks and Backward Castes in Bihar: Social Change in the Late 1970’s.” *Economic and Political Weekly* 12 (Jan.):64–73.
- Brass, Paul R. 1983 *Caste, Faction and Party in Indian Politics. Faction and Party*, vol.I. Delhi: Chanakya Publications.
- 1985 *Caste, Faction and Party in Indian Politics. Election Studies*, vol. II. Delhi: Chanakya Publications.
- Breman, Jan 1974 *Patronage and Exploitation*. Berkeley: University of California Press.
- Dumont, Louis 1970 *Homo Hierarchicus*. Chicago: University of Chicago Press.
- Freeman, James M. 1986 “The Consciousness of Freedom Among India’s Untouchables.” In D. K. Basu and R. Sisson, eds., *Social and Economic Development in India*. Beverly Hills, Calif.: Sage.
- Galanter, Marc 1984 *Competing Equalities: Law and the Backward Classes in India*. Berkeley: University of California Press.
- Gerth, H. H., and C. Wright Mills 1958 *From Max Weber: Essays in Sociology*. New York: Oxford University Press.
- Ghurye, G. S. 1969 *Caste and Race in India*, 5th ed. Bombay: Popular Prakashan.
- Gould, Harold A. 1987 *The Hindu Caste System: The Sacralization of a Social Order*. Delhi: Chanakya Publications.
- Horowitz, Donald L. 1985 *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Hutton, J. H. 1961 *Caste in India*, 4th ed. Oxford, Eng.: Oxford University Press.
- Ishida, Hiroshi 1992 “Stratification and Mobility: The Case of Japan.” In Myron L. Cohen, ed., *Asia Case Studies in the Social Sciences*. New York: M.E. Sharpe.
- Kolenda, Pauline 1978 *Caste in Contemporary India: Beyond Organic Solidarity*. Menlo Park, Calif.: Benjamin/Cummings.

- Kothari, Rajni 1970 *Caste in Indian Politics*. New Delhi: Orient Longman.
- Leach, E. R. (ed.) 1960 *Aspects of Caste in South India, Ceylon, and North-West Pakistan*. Cambridge, Eng.: Cambridge University Press.
- Mahar, J. M. (ed.) 1972 *The Untouchables in Contemporary India*. Tucson: University of Arizona Press.
- Rudolph, L. I., and S. H. Rudolph 1960 "The Political Role of India's Caste Associations." *Pacific Affairs* 33(1):5-22.
- 1987 *In Pursuit of Laxmi*. Chicago: University of Chicago Press.
- Scott, James C. 1990 *Domination And The Arts Of Resistance: Hidden Transcripts*. New Haven, Conn.: Yale University Press.
- Sheth, D. L. 1987 "Reservation Policy Revisited." *Economic and Political Weekly*, vol. 22, Nov.14.
- Southall, Aidan W. 1970 "Stratification In Africa." In Leonard Plotnicov and Arthur Tuden, eds., *Essays in Comparative Social Stratification*. Pittsburg, Pa.: University of Pittsburg Press.
- Srinivas, M.N. 1962 *Caste in Modern India*. London: Asia Publishing House.
- 1969 *Social Change in Modern India*. Berkeley: University of California Press.
- Weiner, Myron 1983 "The Political Consequences of Preferential Policies: A Comparative Perspective." *Comparative Politics* 16(1):35-52.
- Zelliot, Eleanor 1970 "Learning the Uses of Political Means: The Mahars of Maharashtra." In Rajni Kothari, ed., *Caste in Indian Politics*. Delhi: Orient Longman.

RITA JALALI

## CAUSAL INFERENCE MODELS

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

The notion of causality has been controversial for a very long time, and yet neither scientists, social scientists, nor laypeople have been able to think constructively without using a set of explanatory concepts that, either explicitly or not, have

implied causes and effects. Sometimes other words have been substituted, for example, *consequences*, *results*, or *influences*. Even worse, there are vague terms such as *leads to*, *reflects*, *stems from*, *derives from*, *articulates with*, or *follows from*, which are often used in sentences that are almost deliberately ambiguous in avoiding causal terminology. Whenever such vague phrases are used throughout a theoretical work, or whenever one merely states that two variables are correlated with one another, it may not be recognized that what purports to be an "explanation" is really not a genuine theoretical explanation at all.

It is, of course, possible to provide a very narrow definition of causation and then to argue that such a notion is totally inadequate in terms of scientific explanations. If, for example, one defines causation in such a way that there can be only a single cause of a given phenomenon, or that a necessary condition, a sufficient condition, or both must be satisfied, or that absolute certainty is required to establish causation, then indeed very few persons would ever be willing to use the term. Indeed, in sociology, causal terminology was almost deliberately avoided before the 1960s, except in reports of experimental research. Since that time, however, the notion of *multivariate causation*, combined with the explicit allowance for impacts of neglected factors, has gradually replaced these more restrictive usages.

There is general agreement that causation can never be proven, and of course in a strict sense *no* statements about the real world can ever be "proven" correct, if only because of indeterminacies produced by measurement errors and the necessity of relying on evidence that has been filtered through imperfect sense organs or fallible measuring instruments. One may accept the fact that, strictly speaking, one is always dealing with causal *models* of real-world processes and that one's inferences concerning the adequacy of such models must inevitably be based on a combination of *empirical evidence* and *untested assumptions*, some of which are about underlying causal processes that can never be subject to empirical verification. This is basically true for all scientific evidence, though the assumptions one may require in making interpretations or explanations of the underlying reality may be more or less plausible in view of

supplementary information that may be available. Unfortunately, in the social sciences such supplementary information is likely to be of questionable quality, thereby reducing the degree of faith one has in whatever causal assertions have been made.

In the causal modeling literature, which is basically compatible with the so-called structural equation modeling in econometrics, equation systems are constructed so as to represent as well as possible a presumed real-world situation, given whatever limitations have been imposed in terms of omitted variables that produce unknown biases, possibly incorrect functional forms for one's equations, measurement errors, or in general what are termed *specification errors* in the equations. Since such limitations are always present, any particular equation will contain a disturbance term that is assumed to behave in a certain fashion. One's assumptions about such disturbances are both critical for one's inferences and also (for the most part) inherently untestable with the data at hand. This in turn means that such inferences must always be tentative. One never "finds" effects, for example, but only infers them on the basis of findings about covariances and temporal sequences *and* a set of untested theoretical assumptions. To the degree that such assumptions are hidden from view, both the social scientist and one's readers may therefore be seriously misled to the degree that these assumptions are also incorrect.

In the recursive models commonly in use in sociology, it is assumed that causal influences can be ordered, such that one may designate an  $X_1$  that does not depend on any of the remaining variables in the system but, presumably, varies as a result of exogenous causes that have been ignored in the theory. A second variable,  $X_2$ , may then be found that may depend upon  $X_1$  as well as a different set of exogenous factors, but the assumption is that  $X_2$  does *not* affect  $X_1$ , either directly or through any other mechanism. One then builds up the system, equation by equation, by locating an  $X_3$  that may depend on either or both of  $X_1$  or  $X_2$ , plus still another set of independent variables (referred to as exogenous factors), but with the assumption that neither of the first two  $X$ 's is affected by  $X_3$ . Adding still more variables in this recursive fashion, and for the time being assuming linear and additive relationships, one arrives at the system of equations shown in equation system 1,

$$\begin{aligned} X_1 &= \varepsilon_1 \\ X_2 &= \beta_{21}X_1 + \varepsilon_2 \\ X_3 &= \beta_{31}X_1 + \beta_{32}X_2 + \varepsilon_3 \\ &\vdots \\ X_k &= \beta_{k1}X_1 + \beta_{k2}X_2 + \beta_{k3}X_3 + \cdots + \beta_{k,k-1}X_{k-1} + \varepsilon_k \end{aligned} \quad (1)$$

in which the disturbance terms are represented by the  $\varepsilon_i$  and where for the sake of simplicity the constant terms have been omitted.

The essential property of recursive equations that provides a simple causal interpretation is that changes made in any given equation may affect subsequent ones but will *not* affect any of the prior equations. Thus, if a mysterious demon were to change one of the parameters in the equation for  $X_3$ , this would undoubtedly affect not only  $X_3$  but also  $X_4$ ,  $X_5$ , through  $X_k$ , but could have no effect on either of the first two equations, which do not depend on  $X_3$  or any of the later variables in the system. As will be discussed below, this special property of recursive systems does not hold in the more general setup involving variables that may be reciprocally interrelated. Indeed, it is this recursive property that justifies one's dealing with the equations separately and sequentially as single equations. The assumptions required for such a system are therefore implicit in all data analyses (e.g., log-linear modeling, analysis of variance, or comparisons among means) that are typically discussed in first and second courses in applied statistics.

Assumptions are always critical in causal analyses or—what is often not recognized—in *any* kind of theoretical interpretation of empirical data. Some such assumptions are implied by the forms of one's equations, in this case linearity and additivity. Fortunately, these types of assumptions can be rather simply modified by, for example, introducing second- or higher-degree terms, log functions, or interaction terms. It is a mistake to claim—as some critics have done—that causal modeling requires one to assume such restrictive functional forms.

Far more important are two other kinds of assumptions—those about measurement errors and those concerning the disturbance terms representing the effects of all omitted variables. Simple causal modeling of the type represented by

equation system 1 requires the naive assumption that all variables have been perfectly measured, an assumption that is, unfortunately, frequently ignored in many empirical investigations using path analyses based on exactly this same type of causal system. Measurement errors require one to make an auxiliary set of assumptions regarding both the sources of measurement-error bias and the causal connections between so-called true scores and measured indicators. In principle, however, such measurement-error assumptions can be explicitly built into the equation system and empirical estimates obtained, provided there are a sufficient number of multiple indicators to solve for the unknowns produced by these measurement errors, a possibility that will be discussed in the final section.

In many instances, assumptions about one's disturbance terms are even more problematic but equally critical. In verbal statements of theoretical arguments one often comes across the phrase "other things being equal," or the notion that in the ideal experimental design all causes except one must be literally held constant if causal inferences are to be made. Yet both the phrase "other things being equal" and the restrictive assumption of the perfect experiment beg the question of how one can possibly know that "other things" are in fact equal, that all "relevant" variables have been held constant, or that there are no possible sources of measurement bias. Obviously, an alert critic may always suggest another variable that indeed does vary across settings studied or that has not been held constant in an experiment.

In recursive causal models this highly restrictive notion concerning the constancy of all possible alternative causes is relaxed by allowing for a disturbance term that varies precisely because they are *not* all constant. But if so, can one get by without requiring any other assumptions about their effects? Indeed not. One must assume, essentially, that the omitted variables affecting any one of the  $X$ 's are uncorrelated with those that affect the others. If so, it can then be shown that the disturbance term in each equation will be uncorrelated with each of the independent variables appearing on the right-hand side, thus justifying the use of ordinary least-squares estimating procedures. In practical terms, this means that if one has had to omit any important causes of a

given variable, one must also be willing to assume that they do not systematically affect any of its presumed causes that have been explicitly included in our model. A skeptic may, of course, be able to identify one or more such disturbing influences, in which case a modified model may need to be constructed and tested. For example, if  $\epsilon_3$  and  $\epsilon_4$  contain a common cause that can be identified and measured, such a variable needs to be introduced explicitly into the model as a cause of both  $X_3$  and  $X_4$ .

Perhaps the five-variable model of Figure 1 will help the reader visualize what is involved. To be specific, suppose  $X_5$ , the ultimate dependent variable, represents some behavior, say, the actual number of delinquent acts a youth has perpetrated. Let  $X_3$  and  $X_4$ , respectively, represent two internal states, say, guilt and self-esteem. Finally, suppose  $X_1$  and  $X_2$  are two setting variables, parental education and delinquency rates within the youth's neighborhood, with the latter variable being influenced by the former through the parents' ability to select among residential areas.

The fact that the disturbance term arrows are unconnected in Figure 1 represents the assumption that they are mutually uncorrelated, or that the omitted variables affecting any given  $X_i$  are uncorrelated with any of its explicitly included causes among the remaining  $X$ 's. If ordinary least squares is used to estimate the parameters in this model, then the empirically obtained residuals  $e_i$  will indeed be uncorrelated with the independent  $X$ 's in their respective equations, but since this automatically occurs as a property of least-squares estimation, it cannot be used as the basis for a test of our a priori assumptions about the true disturbances.

If one is unwilling to accept these assumptions about the behavior of omitted variables, the only way out of this situation is to reformulate the model and to introduce further complexities in the form of additional measured variables. At some point, however, one must stop and make the (untestable) assumption that the revised causal model is "closed" in the sense that omitted variables do not disturb the patterning of relationships among the included variables.

Assuming such theoretical closure, then, one is in a position to estimate the parameters, attach

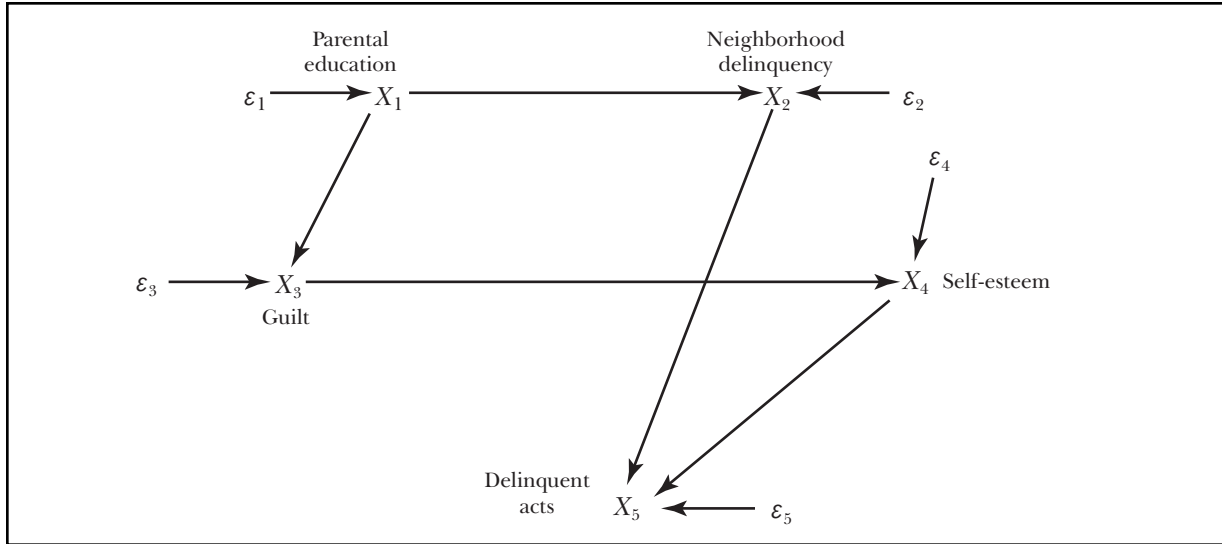


Figure 1. Simple Recursive Model

their numerical values to the diagram, and also evaluate the model in terms of its consistency with the data. In the model of Figure 1, for instance, there are no direct arrows between  $X_2$  and  $X_3$ , between  $X_4$  and both  $X_1$  and  $X_2$ , and between  $X_5$  and both  $X_1$  and  $X_3$ . This means that with controls for all prior or intervening variables, the respective partial correlations can be predicted to be zero, apart from sampling errors. One arrives at the predictions in equation system 2.

$$\begin{aligned}
 r_{23.1} &= 0 & r_{14.23} &= 0 & r_{24.13} &= 0 \\
 r_{15.234} &= 0 & r_{35.124} &= 0 & & 
 \end{aligned}
 \tag{2}$$

Thus, for each omitted arrow one may write out a specific “zero” prediction. Where arrows have been drawn in, it may have been possible to predict the signs of direct links, and these directional predictions may also be used to evaluate the model. Notice a very important property of recursive models. In relating any pair of variables, say,  $X_2$  and  $X_3$ , one expects to control for antecedent or intervening variables, but it is *not* appropriate to introduce as controls any variables that appear as subsequent variables in the model (e.g.,  $X_4$  or  $X_5$ ). The simple phrase “controlling for all relevant variables” should therefore not be construed to mean variables that are presumed to depend on both of the variables being studied. In an experimental setup, one would presumably be unable to

carry out such an absurd operation, but in statistical calculations, which involve pencil-and-paper controlling only, there is nothing to prevent one from doing so.

It is unfortunately the case that controls for dependent variables can sometimes be made inadvertently through one’s research design (Blalock 1985). For example, one may select respondents from a list that is based on a dependent variable such as committing a particular crime, entering a given hospital, living in a certain residential area, or being employed in a particular factory. Whenever such improper controls are introduced, whether recognized explicitly or not, our inferences regarding relationships among causally prior variables are likely to be incorrect. If, for example,  $X_1$  and  $X_2$  are totally uncorrelated, but one controls for their common effect,  $X_3$ , then even though  $r_{12} = 0$ , it will turn out that  $r_{12.3} \neq 0$ .

Recursive models also provide justifications for common-sense rules of thumb regarding the conditions under which it is *not* necessary to control for prior or intervening variables. In the model of Figure 1, for example, it can be shown that although  $r_{24.13} = 0$ , it would be sufficient to control for *either*  $X_1$  or  $X_3$  but not both in order for the partial to disappear. Similarly, in relating  $X_3$  to  $X_5$ , the partial will be reduced to zero if one controls for either  $X_2$  and  $X_4$  or  $X_1$  and  $X_4$ . It is not necessary

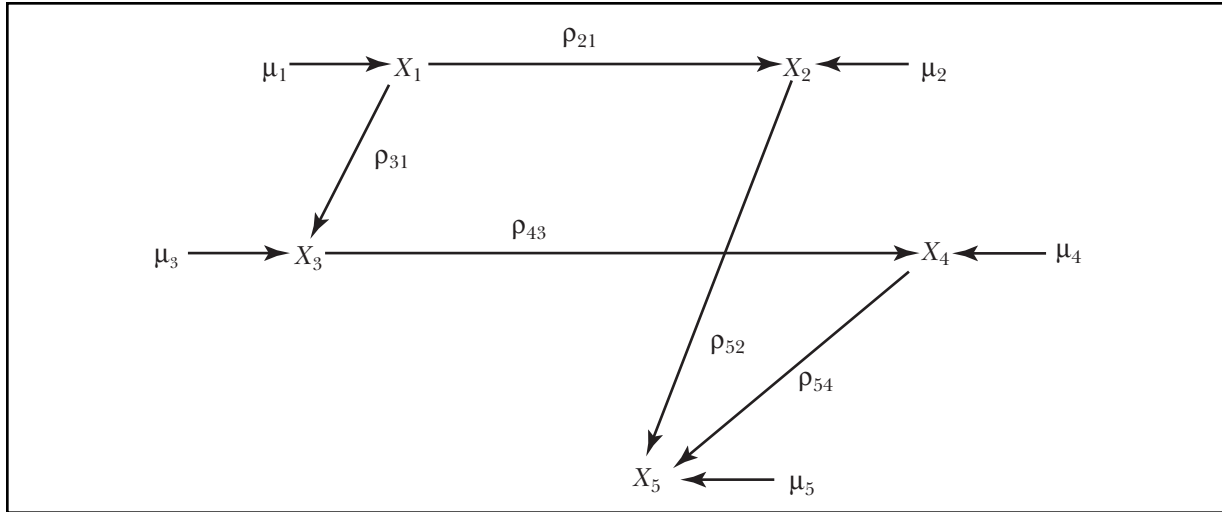


Figure 2. Model of Figure 1, with Path Coefficients and Standardized Variables

to control for all three simultaneously. More generally, a number of simplifications become possible, depending on the patterning of omitted arrows, and these simplifications can be used to justify the omission of certain variables if these cannot be measured. If, for example, one could not measure  $X_3$ , one could draw in a direct arrow from  $X_1$  to  $X_4$  without altering the remainder of the model. Without such an explicit causal model in front of us, however, the omission of variables must be justified on completely ad hoc grounds. The important point is that pragmatic reasons for such omissions should not be accepted without *theoretical* justifications.

**PATH ANALYSIS AND AN EXAMPLE**

Sewall Wright (1934, 1960) introduced a form of causal modeling long before it became fashionable among sociologists. Wright, a population geneticist, worked in terms of standardized variables with unit variances and zero means. Expressing any given equation in terms of what he referred to as path coefficients, which in recursive modeling are equivalent to beta weights, Wright was able to derive a simple formula for decomposing the correlation between any pair of variables  $x_i$  and  $x_j$ . The equation for any given variable can be written as  $x_i = p_{i1}x_1 + p_{i2}x_2 + \dots + p_{ik}x_k + u_i$ , where the  $p_{ij}$  represent standardized regression coefficients and where

the lower-case  $x$ 's refer to the standardized variables. One may then multiply both sides of the equation by  $x_j$ , the variable that is to be correlated with  $x_i$ . Therefore,  $x_i x_j = p_{i1}x_1 x_j + p_{i2}x_2 x_j + \dots + p_{ik}x_k x_j + u_i x_j$ . Summing over all cases and dividing by the number of cases  $N$ , one has the results in equation system 3.

$$r_{ij} = \frac{\sum x_i x_j}{N} = p_{i1} \frac{\sum x_1 x_j}{N} + p_{i2} \frac{\sum x_2 x_j}{N} + \dots + p_{ik} \frac{\sum x_k x_j}{N} + \frac{\sum u_i x_j}{N}$$

$$= p_{i1} r_{1j} + p_{i2} r_{2j} + \dots + p_{ik} r_{kj} + 0 = \sum_k p_{ik} r_{kj} \quad (3)$$

The expression in equation system 3 enables one to decompose or partition any total correlation into a sum of terms, each of which consists of a path coefficient multiplied by a correlation coefficient, which itself may be decomposed in a similar way. In Wright's notation the path coefficients are written without any dots that indicate control variables but are indeed merely the (partial) regression coefficients for the standardized variables. Any given path coefficient, say  $p_{54}$ , can be interpreted as the change that would be imparted in the dependent variable  $x_5$ , in its standard deviation units, if the other variable  $x_4$  were to change by one of its standard deviation units, with the remaining explicitly included independent variables (here  $x_1$ ,  $x_2$ , and  $x_3$ ) all held constant. In working with standardized variables one is able to simplify these expressions owing to the fact that  $r_{ij}$

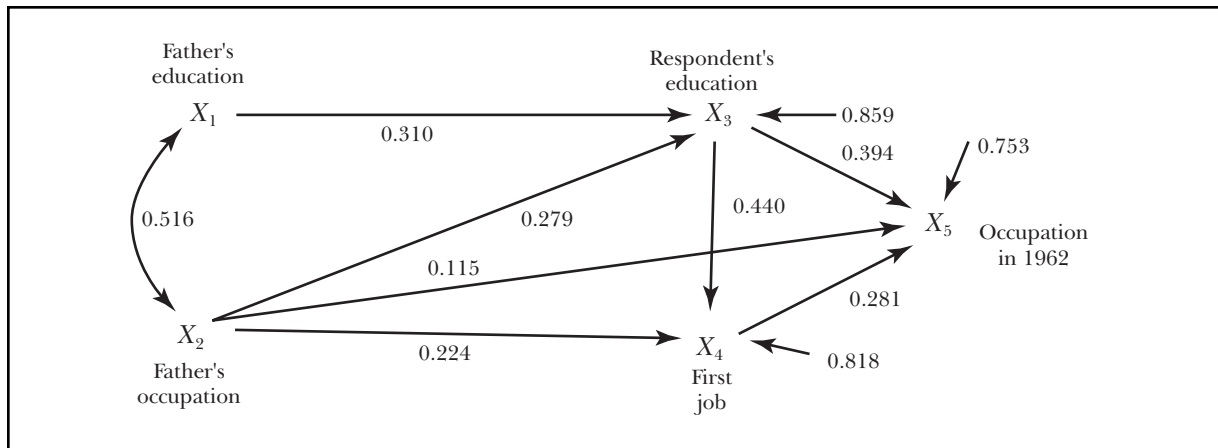


Figure 3. Path Diagram for Blau-Duncan Model

$= \sum x_i x_j / N$ , but one must pay the price of then having to work with standard deviation units that may vary across samples or populations. This, in turn, means that two sets of path coefficients for different samples, say men and women, cannot easily be compared since the standard deviations (say, in income earned) may be different.

In the case of the model of Figure 2, which is the same causal diagram as Figure 1, but with the relevant  $p_{ij}$  inserted, one may write out expressions for each of the  $r_{ij}$  as shown in equation system 4.

$$\begin{aligned}
 r_{12} &= p_{21}r_{11} = p_{21}(1) = p_{21} \\
 r_{13} &= p_{31}r_{11} = p_{31} \\
 r_{23} &= p_{31}r_{12} = p_{31}p_{21} = r_{13}r_{12} \quad (\text{or } r_{23.1} = 0) \\
 r_{14} &= p_{43}r_{13} = p_{43}p_{31} \\
 r_{24} &= p_{43}r_{23} = p_{43}p_{31}p_{21} \\
 r_{34} &= p_{43}r_{33} = p_{43} \\
 r_{15} &= p_{52}r_{21} + p_{54}r_{41} = p_{52}p_{21} + p_{54}p_{43}p_{31} \\
 r_{25} &= p_{52}r_{22} + p_{54}r_{42} = p_{52}p_{54}p_{43}p_{31}p_{21} \\
 r_{35} &= p_{52}r_{23} + p_{54}r_{43} = p_{52}p_{31}p_{21} + p_{54}p_{43} \\
 r_{45} &= p_{52}r_{24} + p_{54}r_{44} = p_{52}p_{43}p_{31}p_{21} + p_{54}
 \end{aligned}
 \tag{4}$$

In decomposing each of the total correlations, one takes the path coefficients for each of the arrows coming into the appropriate dependent variable and multiplies each of these by the total correlation between the variable at the source of

the arrow and the “independent” variable in which one is interested. In the case of  $r_{12}$ , this involves multiplying  $p_{21}$  by the correlation of  $x_1$  with itself, namely  $r_{11} = 1.0$ . Therefore one obtains the simple result that  $r_{12} = p_{21}$ . Similar results obtain for  $r_{13}$  and  $r_{34}$ . The decomposition of  $r_{23}$ , however, results in the expression  $r_{23} = p_{31}r_{12} = p_{31}p_{21} = r_{12}r_{13}$ , which also of course implies that  $r_{23.1} = 0$ .

When one comes to the decomposition of correlations with  $x_5$ , which has two direct paths into it, the expressions become more complex but also demonstrate the heuristic value of path analysis. For example, in the case of  $r_{35}$ , this total correlation can be decomposed into two terms, one representing the indirect effects of  $x_3$  via the intervening variable  $x_4$ , namely the product  $p_{54}p_{43}$ , and the other the spurious association produced by the common cause  $x_1$ , namely the more complex product  $p_{52}p_{31}p_{21}$ . In the case of the correlation between  $x_4$  and  $x_5$  one obtains a similar result except that there is a direct effect term represented by the single coefficient  $p_{54}$ .

As a numerical substantive example consider the path model of Figure 3, which represents the basic model in Blau and Duncan’s classic study, *The American Occupational Structure* (1967, p. 17). Two additional features of the Blau-Duncan model may be noted. A curved, double-headed arrow has been drawn between father’s education and father’s occupation, indicating that the causal paths between these two exogenous or independent variables have not been specified. This means that



there is no  $p_{21}$  in the model, so that  $r_{12}$  cannot be decomposed. Its value of 0.516 has been inserted into the diagram, however. The implication of a failure to commit oneself on the direction of causation between these two variables is that decompositions of subsequent  $r_{ij}$  will involve expressions that are sometimes combinations of the relevant  $p$ 's and the unexplained association between father's education and occupation. This, in turn, means that the indirect effects of one of these variables "through" the other cannot be assessed. One can determine the direct effects of, say, father's occupation on respondent's education, or its indirect effects on occupation in 1962 through first job, but not "through" father's education. If one had, instead, committed oneself to the directional flow from father's education to father's occupation, a not unreasonable assumption, then all indirect effects and spurious connections could be evaluated. Sometimes it is indeed necessary to make use of double-headed arrows when the direction of causation among the most causally prior variables cannot be specified, but one then gives up the ability to trace out those indirect effects or spurious associations that involve these unexplained correlations.

The second feature of the Blau-Duncan diagram worth noting involves the small, unattached arrows coming into each of the "dependent" variables in the model. These of course represent the disturbance terms, which in a correctly specified model are taken to be uncorrelated. But the *magnitudes* of these effects of outside variables are also provided in the diagram to indicate just how much variance remains unexplained by the model. Each of the numerical values of path coefficients coming in from these outside variables, when squared, turns out to be the equivalent of  $1 - R^2$ , or the variances that remain unexplained by *all* of the included explanatory variables. Thus there is considerable unexplained variance in respondent's education (0.738), first job (0.669), and occupation in 1962 (0.567), indicating, of course, plenty of room for other factors to operate. The challenge then becomes that of locating additional variables to improve the explanatory value of the model. This has, indeed, been an important stimulus to the development of the status attainment literature that the Blau-Duncan study subsequently spawned.

The placement of numerical values in such path diagrams enables the reader to assess, rather easily, the relative magnitudes of the several direct effects. Thus, father's education is inferred to have a moderately strong direct effect on respondent's education, but none on the respondent's occupational status. Father's occupation is estimated to have somewhat weaker direct effects on both respondent's education and first job but a much weaker direct effect on his later occupation. The direct effects of respondent's education on first job are estimated to be only somewhat stronger than those on the subsequent occupation, with first job controlled. In evaluating these numerical values, however, one must keep in mind that all variables have been expressed in standard deviation units rather than some "natural" unit such as years of schooling. This in turn means that if variances for, say, men and women or blacks and whites are not the same, then comparisons across samples should be made in terms of unstandardized, rather than standardized, coefficients.

### SIMULTANEOUS EQUATION MODELS

Recursive modeling requires one to make rather strong assumptions about temporal sequences. This does not, in itself, rule out the possibility of reciprocal causation provided that lag periods can be specified. For example, the actions of party A may affect the later behaviors of party B, which in turn affect still later reactions of the first party. Ideally, if one could watch a dynamic interaction process such as that among family members, and accurately record the temporal sequences, one could specify a recursive model in which the behaviors of the same individual could be represented by distinct variables that have been temporally ordered. Indeed Strotz and Wold (1960) have cogently argued that many simultaneous equation models appearing in the econometric literature have been misspecified precisely because they do not capture such dynamic features, which in causal models should ideally involve specified lag periods. For example, prices and quantities of goods do not simply "seek equilibrium." Instead, there are at least three kinds of autonomous actors—producers, customers, and retailers or wholesalers—who react to one another's behaviors with varying lag periods.

In many instances, however, one cannot collect the kinds of data necessary to ascertain these lag periods. Furthermore, especially in the case of aggregated data, the lag periods for different actors may not coincide, so that macro-level changes are for all practical purposes continuous rather than discrete. Population size, literacy levels, urbanization, industrialization, political alienation, and so forth are all changing at once. How can such situations be modeled and what additional complications do they introduce?

In the general case there will be  $k$  mutually interdependent variables  $X_i$  that may possibly each directly affect the others. These are referred to as endogenous variables, with the entire set having the property that there is no single dependent variable that does not feed back to affect at least one of the others. Given this situation, it turns out that it is not legitimate to break the equations apart in order to estimate the parameters, one equation at a time, as one does in the case of a recursive setup. Since any given variable may affect the others, this also means that its omitted causes, represented by the disturbance terms  $\varepsilon_i$ , will also directly or indirectly affect the remaining endogenous variables, so that it becomes totally unreasonable to assume these disturbances to be uncorrelated with the “independent” variables in their respective equations. Thus, one of the critical assumptions required to justify the use of ordinary least squares cannot legitimately be made, meaning that a wide variety of single equation techniques discussed in the statistical literature must be modified.

There is an even more serious problem, however, which can be seen more readily if one writes out the set of equations, one for each of the  $k$  endogenous variables. To this set are added another set of what are called predetermined variables,  $Z_j$ , that will play an essential role to be discussed below. Our equation set now becomes as shown in equation system 5.

The regression coefficients (called “structural parameters”) that connect the several endogenous variables in equation system 5 are designated as  $\beta_{ij}$  and are distinguished from the  $\gamma_{ij}$  representing the direct effects of the predetermined  $Z_j$  on the relevant  $X_i$ . This notational distinction is made because the two kinds of variables play different roles

$$\begin{aligned}
 x_1 &= \beta_{12}x_2 + \beta_{13}x_3 + \cdots + \beta_{1k}x_k + \gamma_{11}z_1 + \gamma_{12}z_2 \\
 &\quad + \cdots + \gamma_{1m}z_m + \varepsilon_1 \\
 x_2 &= \beta_{21}x_1 + \beta_{23}x_3 + \cdots + \beta_{2k}x_k + \gamma_{21}z_1 + \gamma_{22}z_2 \\
 &\quad + \cdots + \gamma_{2m}z_m + \varepsilon_2 \\
 x_3 &= \beta_{31}x_1 + \beta_{32}x_2 + \cdots + \beta_{3k}x_k + \gamma_{31}z_1 + \gamma_{32}z_2 \\
 &\quad + \cdots + \gamma_{3m}z_m + \varepsilon_3 \\
 &\vdots \\
 x_k &= \beta_{k1}x_1 + \beta_{k2}x_2 + \cdots + \beta_{k,k-1}x_{k-1} + \gamma_{k1}z_1 + \\
 &\quad \gamma_{k2}z_2 + \cdots + \gamma_{km}z_m + \varepsilon_k
 \end{aligned} \tag{5}$$

in the model. Although it cannot be assumed that the disturbances  $\varepsilon_i$  are uncorrelated with the endogenous  $X$ 's that appear on the right-hand sides of their respective equations, one may make the somewhat less restrictive assumption that these disturbances are uncorrelated with the predetermined  $Z$ 's.

Some  $Z$ 's may be truly exogenous, or distinct independent variables, that are assumed not to be affected by any of the endogenous variables in the model. Others, however, may be lagged endogenous variables, or prior levels of some of the  $X$ 's. In a sense, the defining characteristic of these predetermined variables is that they be uncorrelated with any of the omitted causes of the endogenous variables. Such an assumption may be difficult to accept in the case of lagged endogenous variables, given the likelihood of autocorrelated disturbances, but we shall not consider this complication further. The basic assumption regarding the truly exogenous variables, however, is that these are uncorrelated with all omitted causes of the  $X$ 's, though they may of course be correlated with the  $X$ 's and also possibly each other.

Clearly, there are more unknown parameters than was the case for the original recursive equation system (1). Turning attention back to the simple recursive system represented in equation system 1, one sees that the matrix of betas in that equation system is triangular, with all such coefficients above the main diagonal being set equal to zero on a priori grounds. That is, in equation system 1, half of the possible betas have been set equal to zero, the remainder being estimated using ordinary least squares. It turns out that in the more general equation system 5, there will be too

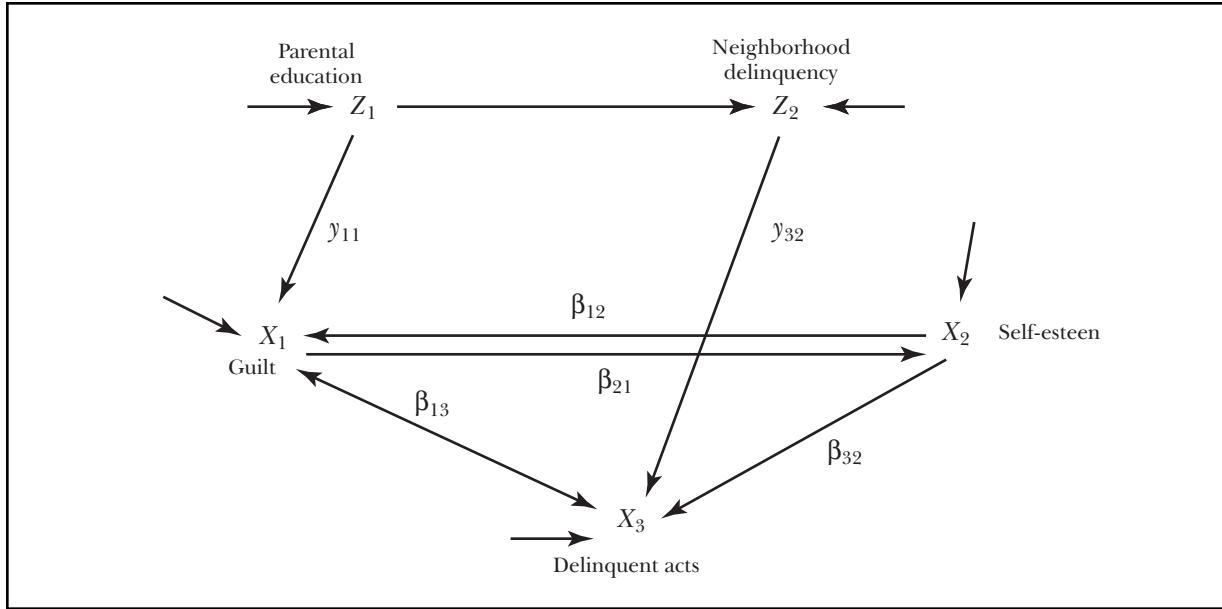


Figure 4. Nonrecursive Modification of Figure 1

many unknowns unless additional restrictive assumptions are made. In particular, in each of the  $k$  equations one will have to make a priori assumptions that at least  $k - 1$  coefficients have been set equal to zero or some other known value (which cannot be estimated from the data). This is why one needs the predetermined  $Z_i$  and the relevant gammas. If one is willing to assume that, for any given endogenous  $X_j$ , certain direct arrows are missing, meaning that there are no *direct* effects coming from the relevant  $X_j$  or  $Z$  variable, then one may indeed estimate the remaining parameters. One does not have to make the very restrictive assumptions required under the recursive setup, namely that if  $X_j$  affects  $X_i$ , then the reverse cannot hold. As long as one assumes that *some* of the coefficients are zero, there is a chance of being able to identify or estimate the others.

It turns out that the *necessary* condition for identification can be easily specified, as implied in the above discussion. For any given equation, one must leave out at least  $k - 1$  of the remaining variables. The *necessary and sufficient* condition is far more complicated to state. In many instances, when the necessary condition has been met, so will the sufficient one as well, unless some of the

equations contain exactly the same sets of variables (i.e., exactly the same combination of omitted variables). But since this will not always be the case, the reader should consult textbooks in econometrics for more complete treatments.

Returning to the substantive example of delinquency, as represented in Figure 1, one may revise the model somewhat by allowing for a feedback from delinquent behavior to guilt, as well as a reciprocal relationship between the two internal states, guilt and self-esteem. One may also relabel parental education as  $Z_1$  and neighborhood delinquency as  $Z_2$  because there is no feedback from any of the three endogenous variables to either of these predetermined ones. Renumbering the endogenous variables as  $X_1$ ,  $X_2$ , and  $X_3$ , one may represent the revised model as in Figure 4.

In this kind of application one may question whether a behavior can ever influence an internal state. Keeping in mind, however, that the concern is with *repeated* acts of delinquency, it is entirely reasonable to assume that earlier acts feed back to affect subsequent guilt levels, which in turn affect future acts of delinquency. It is precisely this very frequent type of causal process that is ignored whenever behaviors are taken, rather simply, as “dependent” variables.

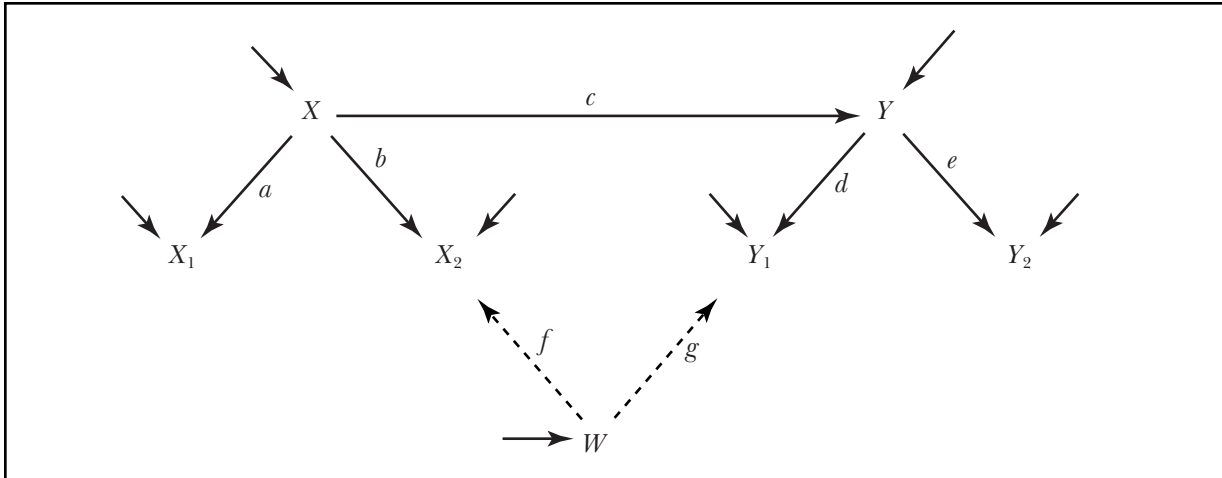


Figure 5. Simple Measurement-Error Model, with Two Unmeasured Variables  $X$  and  $Y$ , Two Indicators of Each, and Source of Bias  $W$

Here,  $k = 3$ , so that at least two variables must be left out of each equation, meaning that their respective coefficients have been set equal to zero. One can rather simply check on the necessary condition by counting arrowheads coming to each variable. In this instance there can be no more than two arrows into each variable, whereas in the case of guilt  $X_1$  there are three. The equation for  $X_1$  is referred to as being “underidentified,” meaning that the situation is empirically hopeless. The coefficients simply cannot be estimated by *any* empirical means. There are exactly two arrows coming into delinquency  $X_3$ , and one refers to this as a situation in which the equation is “exactly identified.” With only a single arrow coming into self-esteem  $X_2$ , one has an “overidentified” equation for which one actually has an excess of empirical information compared to the number of unknowns to be estimated. It turns out that overidentified equations provide criteria for evaluating goodness of fit, or a test of the model, in much the same way that, for recursive models, one obtains an empirical test of a null hypothesis for each causal arrow that has been deleted.

Since the equation for  $X_1$  is underidentified, one must either remove one of the arrows, on a priori grounds, or search for at least one more predetermined variable that does *not* belong in this equation, that is, a predetermined variable that is assumed not to be a direct cause of level of

guilt. Perhaps school performance can be introduced as  $Z_3$  by making the assumption that  $Z_3$  directly affects both self-esteem and delinquency but not guilt level. A check of this revised model indicates that all equations are properly identified, and one may proceed to estimation. Although space does not permit a discussion of alternative estimation methods that enable one to get around the violated assumption required by ordinary least squares, there are various computer programs available to accomplish this task. The simplest such alternative, two-stage least squares (2SLS), will ordinarily be adequate for nearly all sociological applications and turns out to be less sensitive to other kinds of specification errors than many of the more sophisticated alternatives that have been proposed.

### CAUSAL APPROACH TO MEASUREMENT ERRORS

Finally, brief mention should be made of a growing body of literature—closely linked to factor analysis—that has been developed in order to attach measurement-error models to structural-equation approaches that presume perfect measurement. The fundamental philosophical starting point of such models involves the assumption that in many if not most instances, measurement errors can be conceived in causal terms. Most often, the indicator or measured variables are taken as *effects*

of underlying or “true” variables, plus additional factors that may produce combinations of random measurement errors, which are unrelated to all other variables in the theoretical system, and systematic biases that are explainable in causal terms. Thus, measures of “true guilt” or “true self-esteem” will consist of responses, usually to paper-and-pencil tests, that may be subject to distortions produced by other variables, including some of the variables in the causal system. Perhaps distortions in the guilt measure may be a function of amount of delinquent behavior or parental education. Similarly, measures of behaviors are likely to overestimate or underestimate true frequencies, with biases dependent on qualities of the observer, inaccuracies in official records, or perhaps the ability of the actor to evade detection.

In all such instances, we may be able to construct an “auxiliary measurement theory” (Blalock 1968; Costner 1969) that is itself a causal model that contains a mixture of measured and unmeasured variables, the latter of which constitute the “true” or underlying variables of theoretical interest. The existence of such unmeasured variables, however, may introduce identification problems by using more unknowns than can be estimated from one’s data. If so, the situation will once more be hopeless empirically. But if one has available *several* indicators of each of the imperfectly measured constructs, *and* if one is willing to make a sufficient number of simplifying assumptions strategically placed within the overall model, estimates may be obtainable.

Consider the model of Figure 5 (borrowed from Costner 1969), which contains only two theoretical variables of interest, namely, the unmeasured variables  $X$  and  $Y$ . Suppose one has two indicators each for both  $X$  and  $Y$  and that one is willing to make the simplifying assumption that  $X$  does not affect either of  $Y$ ’s indicators,  $Y_1$  and  $Y_2$ , and that  $Y$  does not affect either of  $X$ ’s indicators,  $X_1$  and  $X_2$ . For the time being ignore the variable  $W$  as well as the two dashed arrows drawn from it to the indicators  $X_2$  and  $Y_1$ . Without  $W$ , the nonexistence of other arrows implies that the remaining causes of the four indicators are assumed to be uncorrelated with all other variables in the system, so that one may assume measurement errors to be strictly random.

If one labels the path coefficients (which all connect measured variables to unmeasured ones) by the simple letters  $a, b, c, d$ , and,  $e$ , then with  $4(3)/2 = 6$  correlations among the four indicators, there will be six pieces of empirical information (equation system 5) with which to estimate the five unknown path coefficients.

$$\begin{aligned} r_{x_1x_2} &= ab & r_{x_1y_2} &= ace \\ r_{y_1y_2} &= de & r_{x_2y_1} &= bcd \\ r_{x_1y_1} &= acd & r_{x_2y_2} &= bce \end{aligned} \tag{6}$$

One may now estimate the correlation or path coefficient  $c$  between  $X$  and  $Y$  by an equation derived from equation system 6.

$$c^2 = \frac{abc^2de}{(ab)(de)} = \frac{r_{x_1y_1}r_{x_2y_2}}{r_{x_1x_2}r_{x_1y_2}} = \frac{r_{x_1y_2}r_{x_2y_1}}{r_{x_1x_2}r_{y_1y_2}} \tag{7}$$

Also notice that there is an excess equation that may be used to check on the consistency of the model with the data, namely the prediction that  $r_{x_1y_1}r_{x_2y_2} = r_{x_1x_2}r_{x_1y_2} = abc^2de$ .

Suppose next that there is a source of measurement error bias  $W$  that is a common cause of one of  $X$ ’s indicators (namely  $X_2$ ) and one of  $Y$ ’s (namely  $Y_1$ ). Perhaps these two items have similar wordings based on a social survey, whereas the remaining two indicators involve very different kinds of measures. There is now a different expression for the correlation between  $X_2$  and  $Y_1$ , namely  $r_{x_2y_1} = bcd + fg$ . If one were to use this particular correlation in the estimate of  $c^2$ , without being aware of the impact of  $W$ , one would obtain a biased estimate. In this instance one would be able to detect this particular kind of departure from randomness because the consistency criterion would no longer be met. That is,  $(acd)(bce) \neq (ace)(bcd + fg)$ . Had  $W$  been a common cause of the two indicators of either  $X$  or  $Y$  alone, however, it can be seen that one would have been unable to detect the bias even though it would have been present.

Obviously, most of one’s measurement-error models will be far more complex than this, with several (usually unmeasured) sources of bias, possible nonlinearities, and linkages between some of the important variables and indicators of *other* variables in the substantive theory. Also, some indicators may be taken as *causes* of the conceptual variables, as for example often occurs when one is

attempting to get at experience variables (e.g., exposure to discrimination) by using simple objective indicators such as race, sex, or age. Furthermore, one's substantive models may involve feedback relationships so that simultaneous equation systems must be joined to one's measurement-error models.

In all such instances, there will undoubtedly be numerous specification errors in one's models, so that it becomes necessary to evaluate alternative models in terms of their goodness of fit to the data. Simple path-analytic methods, although heuristically helpful, will no longer be adequate. Fortunately, there are several highly sophisticated computer programs, such as LISREL, that enable social scientists to carry out sophisticated data analyses designed to evaluate these more complex models and to estimate their parameters once it has been decided that the fit to reality is reasonably close. (See Joreskog and Sorbom 1981; Long 1983; and Herting 1985.)

In closing, what needs to be stressed is that causal modeling tools are highly flexible. They may be modified to handle additional complications such as interactions and nonlinearities. Causal modeling in terms of attribute data has been given a firm theoretical underpinning by Suppes (1970), and even ordinal data may be used in an exploratory fashion, provided that one is willing to assume that dichotomization or categorization has not introduced substantial measurement errors that cannot be modeled.

Like all other approaches, however, causal modeling is heavily dependent on the assumptions one is willing to make. Such assumptions need to be made as explicit as possible—a procedure that is unfortunately often not taken sufficiently seriously in the empirical literature. In short, this set of tools, properly used, has been designed to provide precise meaning to the assertion that neither theory nor data can stand alone and that any interpretations of research findings one wishes to provide must inevitably also be based on a set of assumptions, many of which cannot be tested with the data in hand.

Finally, it should be stressed that causal modeling can be very useful in the process of theory construction, even in instances where many of the variables contained in the model will remain unmeasured in any given study. It is certainly a

mistake to throw out portions of one's theory merely because data to test it are not currently available. Indeed, without a theory as to how missing variables are assumed to operate, it will be impossible to justify one's assumptions regarding the behavior of disturbance terms that will contain such variables, whether explicitly recognized or not. Causal modeling may thus be an important tool for guiding future research and for providing guidelines as to what kinds of neglected variables need to be measured.

(SEE ALSO: *Correlation and Regression Analysis*; *Epistemology*; *Multiple Indicator Models*; *Scientific Explanation*; *Tabular Analysis*)

## REFERENCES

- Allison, Paul D. 1995 "Exact Variance of Indirect Effects in Recursive Linear Models." *Sociological Methodology* 25:253-266.
- Arminger, Gerhard 1995 "Specification and Estimation of Mean Structure: Regression Models." In G. Arminger, C. C. Clogg, and M. E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum Press.
- Blalock, Hubert M. 1968 "The Measurement Problem: A Gap Between the Languages of Theory and Research." In H. M. Blalock and A. B. Blalock, eds., *Methodology in Social Research*. New York: McGraw-Hill.
- 1985 "Inadvertent Manipulations of Dependent Variables in Research Designs." In H. M. Blalock, ed., *Causal Models in Panel and Experimental Designs*. New York: Aldine.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Blossfeld, Hans-Peter, and Gotz Rohwer 1997 "Causal Inference, Time and Observation Plans in the Social Sciences." *Quality and Quantity* 31:361-384.
- Bollen, Kenneth A. 1989 *Structural Equations with Latent Variables*. New York: Wiley.
- Costner, Herbert L. 1969 "Theory, Deduction, and Rules of Correspondence." *American Journal of Sociology* 75:245-263.
- Herting, Jerald R. 1985 "Multiple Indicator Models Using LISREL." In H. M. Blalock, ed., *Causal Models in the Social Sciences*, 2nd ed. New York: Aldine.
- Joreskog, Karl G., and Dag Sorbom 1981 *LISREL V: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood; User's Guide*. Uppsala, Sweden: University of Uppsala Press.

- Long, J. Scott 1983 *Confirmatory Factor Analysis: A Preface to LISREL*. Beverly Hills, Calif.: Sage.
- Pearl, Judea 1998 "Graphs, Causality, and Structural Equation Models." *Sociological Methods and Research* 27:226–284.
- Sobel, Michael E. 1995 "Causal Inference in the Social and Behavioral Sciences." In G. Arminger, C. C. Clogg, and M. E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum Press.
- Sobel, Michael E. 1996 "An Introduction to Causal Inference." *Sociological Methods and Research* 24:353–379.
- Spirtes, Peter, Thomas Richardson, Christopher Meek, Richard Scheines, and Clark Glymour 1998 "Using Path Diagrams as a Structural Equation Modeling Tool." *Sociological Methods and Research* 27:182–225.
- Strotz, Robert H., and Herman O. A. Wold 1960 "Recursive Versus Nonrecursive Systems." *Econometrica* 28:417–427.
- Suppes, Patrick 1970 *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.
- Von Eye, Alexander, and Clifford C. Clogg (eds.) 1994 *Latent Variables Analysis: Applications for Developmental Research*. Thousand Oaks, Calif.: Sage Publications.
- Wright, Sewall 1934 "The Method of Path Coefficients." *Annals of Mathematical Statistics* 5:161–215.
- 1960 "Path Coefficients and Path Regressions: Alternative or Complementary Concepts?" *Biometrics* 16:189–202.

HUBERT M. BLALOCK, JR.

## CENSORSHIP AND THE REGULATION OF EXPRESSION

Modern discussions of censorship center on the legitimacy of the regulatory structures and actions through which expression and communication are governed, and the extent to which these structures meet the requirements of democratic societies. In this entry, we first survey prominent historical examples of centralized censorship systems in the West. This history provides a context for a discussion of modern conceptions of censorship and issues associated with the term itself. We then turn to legal structures regulating speech and press in the United States, to government control of political speech, and to some modern-day controversies over the regulation of expression.

## EARLY SYSTEMS OF GOVERNMENT CENSORSHIP

Traditional conceptions of censorship are rooted in rigid systems of ecclesiastic and governmental control over discourse and printing. The term itself derives from the office of the census in early Rome, where the censor served as both census taker and as supervisor of public conduct and morals. Before the advent of the printing press in the fifteenth century, most manuscripts in Europe were produced in monasteries, which controlled their production. Centralized systems of control over books developed largely in response to the invention of the printing press, which both church and state perceived as threats to their authority. In the mid-sixteenth century the Catholic Church issued the Index of Forbidden Books, which was enforced through compliance of the faithful, pre-publication screening of books, the burning of heretical tracts, and the persecution of heretics. In Protestant countries, the State generally assumed control over the publication of books. The English monarchy published its first list of prohibited books in 1529 and exercised its control through a contractual arrangement with the Stationers' Company, which, in 1557, was granted a monopoly on the production and distribution of printed materials. This charter remained in effect until 1694 when it was allowed to expire, primarily because of difficulties in its administration. These centralized mechanisms of control were replaced by less systematic methods, such as laws against seditious libel, through which speech that merely criticized government policies could be punished.

Although systems similar to these proliferated worldwide and continue to exist in modern-day authoritarian regimes, they are consensually viewed as incompatible with the practice of democracy. The history of free speech principles in the West coincides with the rise of democratic thought, as expressed in the writings of the eighteenth-century Enlightenment philosophers in France and in the influential political philosophies of John Locke, John Milton, and John Stuart Mill in England. The turn of mind that gripped Europe during this period is reflected in Locke's dictum that governments are the servants of the people, not the reverse, and in the insistence by each of these philosophers that self-governance cannot function under regimes where the circulation of ideas is dependent upon the whims of rulers. In modern



political theory, Habermas (1989) considers the demise of systematic state censorship to be a precondition of the rise of a “public sphere,” an admittedly idealized realm of discourse—independent of both state and market—in which public issues can be deliberated in an environment where reason, not the status of speakers is honored.

### THE RHETORIC OF “CENSORSHIP”

The use of language plays a critical role in framing thought and discussion about the legitimate control of speech. The term “censorship” most frequently arises in debates over the desirability of restricting access to one or another form of communicative content. However, the term is fraught with the complexities of multiple meanings, uses, and understandings. These complexities arise from the historical baggage it carries, the multiple contexts—popular, legal, and scholarly—in which it is used, and the highly contested nature of the debates in which it is invoked. Although systematic empirical analyses of the term have yet to be conducted, its meanings appear to differ along two dimensions, connotative and descriptive. Underlying the connotative force of the term is the strong conviction that suppression of speech is at best a necessary evil. As such, the term typically carries with it a highly pejorative connotation and a strong air of illegitimacy. Although First Amendment scholar Smolla (1991) informs us that “censorship was not always a dirty word,” this facet of the term is now found in Webster’s Dictionary, which states that it refers especially to control that is “exercised repressively.”

Descriptive uses of the term differ according to the breadth of the domain covered. Under strict uses of the term, censorship means the prior restraint of information by government. It is this meaning that enables the Federal Communications Commission (FCC) to enact regulations that impose post hoc penalties for some forms of speech while at the same time declaring that “nothing in the Act [that governs the FCC] shall be . . . construed to give the Commission the power of censorship.” In a second form of use, censorship refers to any form of government regulations that restrict or disable speech. Fines imposed by the FCC for “indecent” speech on radio fit this use, as the intent is to deter further indecency. In a third, less conventional usage, the term is modified to

refer to nongovernmental restrictions on speech. For example, in one of the few instances in which the Supreme Court has applied the term to private concerns (*Red Lion v. FCC*, 395 U.S. 367, 1969), the Court stated that “The First Amendment does not protect private censorship by broadcasters who are licensed by the Government to use a scarce resource which is denied to others.”

In each of these three descriptive domains the term “censorship” almost invariably carries its pejorative inflection. In its broadest sense, censorship signifies control over the means of expression; the determination of what content will not be communicated. This sense of the term is found in the work of Bourdieu (1991), for whom censorship is located not only in explicit prohibitions, but also in the everyday practices and power relationships that determine what is and is not said. It includes not only the sixteenth-century Inquisitor who decides which books will be burned, but also the twentieth-century film editor who leaves a scene on the cutting room floor. It includes both the intentional actions of individuals and, as importantly, the de facto results of impersonal forces that lead some ideas not to be expressed. This expansive meaning, which has yet to make significant inroads into popular discourse, is incompatible with the pejorative connotation found in everyday usage: If censorship is an integral part of everyday life, it cannot always, or even typically, be evil.

The prototypical understanding of “censorship” is firmly anchored at the point where the term’s pejorative connotation intersects with governmental restrictions on speech. This can be seen in how the term is and is not used in legal and popular discourse. First, there are numerous forms of government restrictions that tend not to be thought of as censorship at all. Examples include punishment for perjury, threats, and libel, and prohibitions on misleading advertising. The core meaning of the term “censorship” tends to be applied to restrictions that are consensually deemed to be illegitimate; it tends not to be applied to restrictions that are consensually deemed to be legitimate, although they fall within the descriptive scope of the term. As such, the term typically serves to label those policies and actions that have been deemed undesirable rather than to describe a set of activities whose legitimacy could then be judged. A thorough sociological understanding of



“censorship” would require a mapping of the shifting boundaries around which this legitimacy is withheld or conferred. For example, punishment for blasphemy was once considered legitimate, whereas penalties for inciting hatred against ethnic groups may not be thought of as “censorship” in the future. Second, in areas where the legitimacy of a restriction is under dispute, those who oppose the restriction tend to label it as censorship, whereas those who support the restriction attempt to distance themselves from the term.

This core understanding of the term—as illegitimate government action—leads to a number of consequences, two of which will be mentioned here. First, discussions of the legitimate control of expression are typically framed as battles between censorship, on the one hand, and Free speech, on the other: Free speech is the absence of governmental control. This framing of the communicative needs of a democracy—codified in the First Amendment—places a wholesome burden on governments to justify regulatory action. On the other hand, this framing excludes from discussion the positive role that governments can (and do) play in supporting these needs. It also excludes consideration of the instances in which private concerns and impersonal market forces can produce deleterious effects; ones that exclude particular viewpoints from the marketplace of ideas. The restrictions placed by health maintenance organizations (HMOs) upon what physicians can say to their patients (i.e., “gag rules”) are but one of countless areas in which private institutions “censor” valuable speech.

Research is needed to uncover the role that the rhetoric of “censorship” plays in maintaining the conception of free speech as the absence of government regulation. It is apparently with the goal of disabling this entrenched dichotomy that some theorists have used the term—e.g., “soft censorship,” “de facto censorship” “private censorship”—to refer to those private and impersonal forces that lead some forms of expression to be systematically excluded from the marketplace of ideas (e.g., Barbur 1996; cf. Post 1998). Barbur, for example, argues that “monopoly is a polite word for uniformity, which is a polite word for virtual censorship—censorship not as a consequence of political choices, but as a consequence of inelastic markets, imperfect competition, and economies of scale . . .” (1996, pp. 137–138). Similarly, theorists

like Baker (1998) use the term “structural regulation” to refer to government interventions designed to increase viewpoint diversity in the public forum. The goal of such rhetorical interventions is to reframe the discussion of “free speech” in a way that is meticulously sensitive to the threats of government incursions on speech, yet, that at the same time allow this discussion to incorporate both the speech-restricting characteristics of private action and the speech-expanding possibilities of government.

The second consequence of our core understanding of “censorship” is linked primarily to the distaste the term evokes. This pejorative connotation renders the term useful in singling out illegitimate expressive restrictions, and as a persuasive device in popular and legal debate. However, the pejorative nature of the term also disables it as a useful construct in discussions that aim to present a balanced account of contemporary debates over the control of expression. For example, to frame a controversy concerning the legitimacy of restrictions on hate speech as a debate over whether such speech should be censored is to ask, in effect, whether the illegitimate suppression of hate speech is or is not legitimate. This framing, though a popular one, clearly biases the discussion toward one side of the debate.

In this entry we attempt to use less loaded terms to describe the concerns that the term “censorship” evokes and the underlying controversies in which the term is typically used. Correspondingly, we strive to reserve the term “censorship” for those uses in which it is integral to the viewpoint being expressed.

#### LEGAL PRINCIPLES GOVERNING THE REGULATION OF EXPRESSION IN THE UNITED STATES

The legal principles governing freedom of expression in the United States are based largely on interpretations of the First Amendment to the Constitution, whose “speech” and “press” clauses are combined in the statement that “Congress shall make no law . . . abridging the freedom of speech, or of the press.” The amendment refers only to actions taken by the U.S. Congress, and it was not held to apply to laws made by the individual states until 1925 (*Gitlow v. New York*, 2688 U.S.

652, 1925) when the Court ruled that the fourteenth Amendment required that state laws not be in conflict with federal law. In addition to state and constitutional law, other major sources of legal regulation originate in common law (e.g., sedition, privacy) and administrative law (e.g., rulemaking of the FCC).

Despite the apparent clarity of the First Amendment language, both Congress and the states have crafted many laws that regulate freedoms of speech and press. These regulations stem primarily from three sources. First, the First Amendment is subject to widely varying interpretations. Some constitutional scholars conclude that at time the Bill of Rights was drafted, conceptions of freedom of expression referred only to prior restraints on speech. Others note ambiguities in the words “abridge” and “speech,” and further question the intent of the framers of the Constitution. For example, do regulation of defamatory tracts, false commercial advertising, perjury, threats, violent obscenity, or nude dancing necessarily count as abridgments of speech (Sunstein 1993)? Some have also argued that the speech clause governs only political speech or that the press clause should not be interpreted to apply to tabloid entertainment that masquerades as news.

Second, the modern world differs from the world of the late eighteenth century. The framers could not have foreseen the technological advances that have brought us telephones, television, and the Internet. Neither could they have had in mind the concentrated corporate ownership or the advertising-driven programming that characterizes today’s media. Some argue that government regulation is required in order to achieve the principles embodied in the Constitution.

Third, speech rights do not stand alone. For example, the rights of some to speak may conflict with the rights of others to be let alone or to be treated with dignity and equality. Also, the speech rights asserted by some inevitably conflict with the speech rights claimed by others.

In light of such issues, the clarity and simplicity that at first seem to characterize the First Amendment turn out to be illusory. As a result, scholars have developed theories of the underlying values that they believe free speech should promote in a democracy and should govern interpretation of

the First Amendment. Although no theory is universally accepted as dominant, most agree with the judgment of constitutional scholar Emerson (1973) that the value of free expression lies in its ability to promote participation in the decision-making process by all members of society; to advance knowledge and the pursuit of truth; to promote individual self-fulfillment; and, by allowing dissent to be publicly vented, to promote a balance between stability and change.

Disputes over the role of government in regulating speech turn, in part, on which of these values is given precedence. Those who emphasize the role of speech in promoting deliberative democracy often favor an affirmative role of government to promote viewpoint diversity or to ensure that the distribution of viewpoints in the public sphere is somewhat reflective of their distribution in the community. In practical terms, this may lead to restrictions on the amount of money that corporations can contribute to political campaigns, to the enforcement of rights of access to media by citizens, to the strategic placement of “public forum” spaces on the Internet, to stricter enforcement of the antitrust laws as they pertain to media, or to government subsidies for valued forms of speech (Baker 1998). In many instances, the promotion of viewpoint diversity and balance may require the government to regulate or infringe upon the speech or property rights of some in order to advance the needs of deliberative democracy.

#### **DIFFERENTIAL JUDICIAL TREATMENT OF DIFFERENT FORMS OF EXPRESSION**

The corpus of First Amendment decisions shows that all speech is not equal in the protections it is afforded. Political discourse—broadly defined to include speech about social, cultural, and religious issues—is considered speech of the highest value. As a result, the government must demonstrate a “compelling” interest to warrant its restriction. Particularly, political speech cannot be directly restricted unless it is “directed to inciting or producing imminent lawless action and is likely to incite or produce such action.” Advertising was long considered outside the purview of First Amendment protections and was subject to strict regulation. Commercial speech has experienced tremendous leaps in its status under the Court as

many local restrictions have been declared unconstitutional. Obscenity lies at the lowest rung of speech and receives no protection under the First Amendment. Although the requirements for demonstrating that speech is obscene have become progressively more demanding, Congress and the states are permitted to ban the public dissemination of speech that meets these criteria (see below). Depending upon how they are classified in terms of value to society, other forms of speech receive greater or lesser protection. For example, unlicensed medical advice and misleading advertising can be restricted through a “balancing test” which shows that the harm that stems from suppressing them are fewer than the harm they cause. A “balancing test” which shows that the harm that stems from suppressing them are less than the harm they cause.

Current First Amendment doctrine also reflects a second, cross-cutting mode of classification according to which proposed regulations are categorized as content neutral, content based, or viewpoint based. Of the three, content neutral regulations must pass the lowest constitutional hurdle whereas viewpoint-based regulations must pass the highest. An example of content-neutral regulations might include rules that designate sound-level restrictions on expressive activities, such as music in Central Park. Examples of viewpoint-based restrictions might include a ban on a gay parade or on either Democrat or Republican billboards. Regulations on expression that are content based, but viewpoint neutral face a high, but not insurmountable, constitutional hurdle. For example, in their contractual relations with cable companies, municipalities have been allowed to require the companies to provide local news and sports, and the federal government has been allowed to prohibit all partisan political campaigning on army bases.

The application of First Amendment law is also a function of the technological environment, or medium, in which expression is conveyed. Except through the sporadic use of antitrust laws, the Court has been least likely to permit regulations of print-based news and most likely to allow regulations of broadcast media. Cable television has fallen somewhere in between, with regulations reflecting the monopolistic control that companies exert over cable access to individual households and the fact that municipalities own the

property through which television cables are distributed. Telephone companies are classified as “common carriers,” which interdicts their editorial control over the information that passes through their wires. The newest form of communication, the Internet, has so far been granted the highest rung of protection from government regulation.

The willingness of the Court to allow government regulation of broadcast media lies largely on three rationales: public ownership of the airwaves, scarcity of the broadcast spectrum, and “pervasiveness” of the broadcast signal. First, the airwaves through which broadcast signals are transmitted are owned by the public and licensed on a renewable basis to radio and television broadcasters. Second, only a limited number of broadcast signals can coexist in any given segment of airspace (scarcity principle). In other words, the broadcast spectrum is a scarce resource, one that has historically required some regulatory body to decide which of the many interested broadcasters will be allotted the frequencies that exist. The management of these tasks is the function of the FCC, which was authorized to regulate broadcasters in the “public convenience, interest, or necessity.” In interpreting the “public-interest” clause, the FCC has issued a number of requirements, including, for example, a modicum of public-interest programming such as local news, rights of reply to those attacked in political editorials, and requirements to air political advertisements. Regulations governing the latter are governed by a “no censorship” clause, whereby the FCC deprives stations of editorial rights over political advertisements. Although most of these regulations are aimed to increase viewpoint diversity, broadcasters often argue that FCC rules infringe upon their rights to free speech. Finally, the Supreme Court’s judgment that radio and television broadcasts are “pervasive” refers to the assumed inability of viewers and listeners to fully control their own access, or that of their children, to unexpected program content. The susceptibility of audience members to be caught unawares by programming that offends them, or that they deem harmful to their children, has led the court to characterize such broadcasts as an “uninvited intruder” into the privacy of a viewer’s abode. Primarily to provide a “safe haven,” for children, the Court has allowed time restrictions on the broadcast of “indecent” or “patently offensive” programming.

Given the different levels of protection afforded by the First Amendment for different forms of expression, the determination of which category of speech, which kind of regulation, and which form of media a given case will be deemed to embody, is of critical importance. Levels of protection for different forms of speech have changed over time and there is no reason to think that they will not continue to do so. Movies were long considered a form of crass entertainment, outside of First Amendment protections, and birth control information was once classified as obscene. In the not-distant future, the plethora of programming opportunities on the Internet is likely to result in Supreme Court challenges of the scarcity principle upon which public-interest broadcast regulations are based. Particularly in the area of new media, the difficulties inherent in determining how to categorize a particular speech situation often leads the categorical approach to cede to an approach in which the harms of competing outcomes are less formally “balanced” against each other. Projects for sociology might include examination of the social underpinnings of the origin of these categories, of the application of these categories to different forms of speech, and of the evidentiary criteria used to assess both the harms of speech and the harms of regulation. For instance, a near consensus has been reached in the scientific community that media violence leads to social aggression and violence (e.g., Donnerstein and Smith 1997), whereas evidence of harms done by “indecent” broadcasts is anecdotal. What then are the social and political processes that lead one and not the other to be regulated?

#### **PUBLIC SPEECH, PRIVATE SPACES**

As can be seen in the language of the First Amendment, the central concern of the framers of the Constitution was to protect the private realm from domination by the state. The First Amendment itself has little to say about the control that private organizations exert over the expressive rights of individuals, or the control that private media companies exert over the communication of public issues. Examples of private-domain controversies include disputes over the “gag rules” that prohibit physicians from discussing alternate, more expensive treatments; corporate contracts that prevent employees from publicizing questionable employer practices; and control over the correspondence

that employees send over company e-mail systems. An evolving legal controversy surrounds the growing replacement of public forums—sidewalks, streets, and public parks—by private ones, such as shopping malls, condominiums, and gated communities. Should the political protesters who once convened in front of the local store be barred from entry to the shopping mall where the store now stands? Should those gathering signatures for political initiatives be excluded from gated communities? These issues pit the property rights of owners against the expressive needs of communities. Supreme Court decisions have given individual states some leeway in deciding which and when one of these interests is the more compelling. Related issues arise in the domain of media, where some point to the growing concentration of media ownership (Bagdikian 1997), the presence of only one daily newspaper in most U.S. cities, and the “skewing” of media content away from the interests of the poor (Baker 1998) as evidence that the marketplace of ideas is not fully served by unregulated economic markets (Sunstein 1993).

A strict reading of the First Amendment suggests that the government has no right to intercede to regulate communication between a corporation and its employees (harassment rules provide an exception) or to intervene when privately owned media fail to meet the needs of a community. Many argue that this is as it should be: employees can always seek employment elsewhere and those whose speech is barred in one forum can always seek another. Others insist that differences between the conditions of modern society and those at the time the Constitution was drafted warrant the extension of communicative rights beyond those provided in the First Amendment.

#### **THE LEGAL SUPPRESSION OF SPEECH DEEMED TO THREATEN THE ESTABLISHED ORDER**

A central rationale for laws restricting the expressive activities of people and press is to preserve national security. The chief mechanisms of enforcement under democratic governments have been restrictions on governmentally controlled information, laws prohibiting seditious libel, and, in times of war, systems of prepublication clearance of press dispatches. Most agree that there are some circumstances, particularly in times of war,

in which some national security restrictions are necessary. However, a strong proclivity exists to abuse these laws to protect policies or governments that are losing, or have lost, their base of popular support.

The prototypical conception of censorship is firmly rooted in governmental restrictions on speech that is critical of government, yet poses significant threats to security. The law of seditious libel was imported to the United States from England, where, during the eighteenth century, it served as a principal vehicle through which governments attempted to stave off criticism and to control public opinion. The power of seditious libel laws lay largely in their breadth—advocacy of an “ill opinion” of government was considered actionable—and from the selectivity with which they could be enforced. During most of the eighteenth century, these cases were decided by judges rather than juries, and the truth of a charge was not a defense until the nineteenth century.

The climate in which the Constitution’s press clause was drafted was one in which a central threat to viewpoint diversity was posed by governments who could use their monopolies on legitimized force to control the marketplace of ideas. The most noted illustration of this in colonial America was the case of John Peter Zenger, whose newspaper, the *New York Weekly* was launched in 1734 as the only voice of opposition to the widely reviled colonial Governor, William Cosby. When the newspaper’s opening salvo attacked Cosby’s arbitrary exercise of power, Zenger was promptly charged with seditious libel. After spending eight months in jail, Zenger was acquitted by a jury (reluctantly allowed in this case), who ignored the judge’s instructions that the truth of a charge was irrelevant. Although many thought the First Amendment would change such abuse, Congress soon passed the Sedition Act, which was used by President John Adams, a Federalist, to silence criticism launched by Republican supporters of Thomas Jefferson.

The Sedition Act was allowed to lapse when Jefferson took the Presidency and would not return until World War I. However, in the years during which the Constitution did not apply to the states, statutory attacks on specific viewpoints could be fierce at the state level. Louisiana’s pre-Civil

War statute prosecuting speech that sowed “discontent among the free population or insubordination among the slaves” provides just one example of how law could be enlisted in the service of majoritarian prejudices. When the federal government returned to the sedition business with passage of the Espionage and Sedition Acts, the country was characterized by a climate of isolationism and a receptive ear to the tenets of Socialists, who opposed the entry of the United States into World War I. The Sedition Act rendered it illegal to speak against the draft or to advocate strikes that might hinder wartime production. It resulted in more than 800 convictions, including that of Socialist presidential candidate Eugene Debs. Prosecutions for political speech continued after World War II, when provisions of the 1940 Smith Act were used as the legal arm of an extensive and popularly backed campaign to suppress Communist viewpoints in the press, the workplace, and the entertainment media.

A significant judicial outcome of cases stemming from the enforcement of these acts was the evolution of the criteria used to determine whether political speech warrants conviction. Whereas it had previously sufficed that expression result in only a “bad tendency” to cause harm, these cases eventually produced today’s criterion of “imminent lawlessness.” Had the expressive restrictions of the first part of the century been in place during the civil rights movement and the Vietnam War, much of what was said and written during these periods would have resulted in federally sanctioned jail sentences.

National security interests have also been invoked to justify secrecy classification systems, the labeling of foreign “propaganda,” and the use of contract law by agencies including the Central Intelligence Agency and the Voice of America to require prepublication clearance of communications by both current and former employees.

### WARTIME RESTRICTIONS

Compulsory systems of prepublication clearance during wartime were first used during the Civil War and, with the notable exception of Vietnam, have continued to be used in all major military conflicts. The exception of Vietnam was prompted

by the government's fear that prior review systems would alienate the press and by the fact that readily available nonmilitary air transportation in the area would have rendered enforcement difficult. However, there is credible evidence that military officials systematically misrepresented progress by the United States during the war and that media routines and practices rendered them susceptible to this manipulation (Hallin 1986).

Despite the consensus on the need for some form of press management system to protect the lives of soldiers during wartime, controversies arise over the scope and mechanisms of enforcement. Of particular concern to press and public is the fear that controls ostensibly designed to protect the lives of soldiers are used instead to manage public opinion at home.

### PORNOGRAPHY

The term pornography comes from the Greek words for "prostitute" and "write," and originally referred to writings about prostitutes and their activities. Today, pornography is broadly used to mean material with explicit sexual content. Like the word censorship itself, however, the definition of pornographic or obscene material is often contested. For example, feminist writers often draw a distinction between pornography, which combines sexuality with abuse or degradation, and erotica, which is sexually arousing material that respects the human dignity of the participants. A distinction is also sometimes drawn between hard-core pornography, which shows actual sexual intercourse or penetration, and soft-core pornography, which may be only suggestive of these activities. Child pornography is prohibited in most nations, and restrictions on its production and distribution tend to be noncontroversial.

In the United States, the Commission on Obscenity and Pornography (1970) and the Attorney General's Commission on Pornography (1986) have provided recommendations for government action regarding pornography. The first report suggested that pornography should not be regulated by law, while the second rejected the claims that pornographic material was harmless and urged prosecution especially for materials that contained violence or degradation. These differing conclusions highlight the ongoing tensions between individual rights and perceived community needs, as

well as changes in the progress and interpretation of research on the effects of pornography. The uses of the conclusions drawn in these and other social-scientific reports are of particular sociological interest. For example, the Nixon Administration was quick to dismiss the conclusions of a report (which it had itself commissioned) that were contrary to its political goals. Here, as in other contested areas of expression, the political climate in which research findings are interpreted is often the deciding factor in how research is used.

Regarding the related question of obscenity, in 1973 the Supreme Court (*Miller v. California*, 413, U.S. 15, 1973) established three criteria for determining whether a given work was obscene: An average person, applying contemporary local community standards, finds that the work, taken as a whole, appeals to prurient interest; the work depicts in a patently offensive way sexual conduct specifically defined by applicable state law; and the work in question lacks serious literary, artistic, political, or scientific value.

Political struggles over the regulation of pornography have been especially intriguing because they have brought together conservatives and feminists, groups traditionally on opposite sides of the ideological spectrum. Although both sides may support banning pornography, they have different reasons for doing so, and the scope of the material they wish to have regulated differs as well. Feminists focus on the degrading character of pornography, whereas conservatives view pornography as morally corrosive. Antipornography activists such as Catherine MacKinnon have defined pornography as a civil rights issue, arguing that pornography itself is a form of sexual discrimination; by presenting women in dehumanizing ways, pornography subordinates them.

Although there is wide variation in pornographic content, pornography often presents what many consider to be an unrealistic view of sexual relations. Encounters take place most often between strangers, not in the context of enduring relationships; participants are sex objects rather than complete individuals. Sexual activity always results in ecstasy, and consequences (such as pregnancy or disease) are nonexistent.

Empirical research on the effects of pornography has shown a range of negative effects. In

typical experimental studies, participants are randomly assigned to pornographic stimuli or nonpornographic stimuli, and their attitudes, behaviors, or physiological reactions are then assessed. There are also paradigms assessing prolonged exposure, including ones in which participants return to the laboratory for multiple sessions of exposure to pornography, to better simulate real-life consumption patterns. Among other effects, studies have shown that after exposure to pornography, participants viewed rape as a less serious crime, overestimated the popularity of less common sexual practices, and showed greater callousness toward women. Furthermore, pornography consumers have shown weaker beliefs in the desirability of marriage and having children, and stronger beliefs in the normality of sexual promiscuity. Ironically, pornography has also been shown to reduce viewers' satisfaction with their own sex lives and partners.

Although these outcomes have been found in a variety of studies using different research procedures, not all studies have confirmed these findings. Some critics of this research argue that rape and other antisocial sexual behavior existed even before pornography became widely available, and that forces besides pornography play a greater causal role in antisocial behavior. Others argue that the increased levels of aggression, hostility, or bias often found in experimental studies of pornography might be traced to differences between the experimental and naturalistic environments in which pornography is viewed: particularly, it is argued that these effects may be due to the lack of opportunity for men to ejaculate in the experimental setting. They also caution that the political biases of investigators may influence the interpretation of results. Others believe that even if the effects research is accurate, the costs of suppressing pornographic material—as measured in state encroachment on individual autonomy—outweigh any benefits that might be gained.

An argument sometimes raised for controlling or eliminating pornography in the form of images (rather than words) is that the individuals pictured in the pornographic photos and films suffered harm or coercion during the creation of the materials. A counterpoint to this argument is provided by Stoller and Levine (1993), who have

conducted in-depth ethnographic style interviews with producers, performers, and other employees in the pornography industry. In this work, the researchers allowed the people who create pornography to speak in their own words, providing both defenses of pornography and insight into why individuals choose, for better or for worse, to work in the pornography industry.

### HATE SPEECH

From the advent of the printing press onward, most efforts to legally regulate expression have focused on various forms of mass communication. Efforts to restrict speech can also target communication in interpersonal settings. One such example can be found in attempts to place legal limits on hate speech; defined as harassing or intimidating remarks that derogate the hearer's race, gender, religion, or sexual orientation.

Those who support regulating hate speech compare it to "fighting words," a category of speech not protected by the First Amendment, or argue that it creates a "hostile environment," which violates provisions of the Civil Rights Acts of 1964 and 1990. Champions of hate speech regulations cite the harms its victims suffer: these may range from feelings of exclusion from a community, to the experience of debasement that leads students to skip classes or distress that leads them to leave school. In short, the liberty of a speaker to harass may deny the hearer's right to equality. Critics of hate speech codes cite the administrative excesses they allow. For example, the student guide to the University of Michigan code stated that students could be punished for making a comment "in a derogatory way about someone's appearance."

Following an increase in reported incidents of hate speech in the late 1980s and early 1990s, many universities in the United States adopted policies forbidding discriminatory verbal harassment. However, as they affect public institutions, these rules have been declared unconstitutional in the courts, in part for being either too broad—affecting too many forms of speech—or for being viewpoint based (see above).

The problems inherent in drafting laws that affect only the speech these rules target can be

seen in comparing the provisions of the University of Michigan code with provisions in the bill advanced by Senator Jesse Helms to limit what kinds of works could receive funding from the National Endowment for the Arts. Among the Helms bill's other planks (aimed at homoerotic art, for example), it sought to prohibit funding of "material which denigrates . . . a person, group, or class of citizens on the basis of race, creed, sex, age . . ." Similarly, the Michigan code sought to bar speech that "stigmatizes or victimizes an individual on the basis of race, ethnicity, religion, sex . . ." and also "creates a hostile environment." The projects targeted by the Helms bill included one with a crucifix immersed in a bottle of urine. This work could also be a violation of the Michigan code if it were to be hung in an area frequented by offended Christians. Difficulties of this sort have prompted those supporting hate speech codes to argue that this is one area in which viewpoint-based speech regulations should be allowed: that, similar to laws in Germany that selectively ban Holocaust denial, speech laws in the United States should be allowed to account for the history of discrimination experienced by some groups, but not others. Critics say that viewpoint discrimination of this sort would open the floodgates for future laws that selectively target viewpoints that happen to be deemed undesirable.

The hate speech debates highlight a fundamental dilemma faced by modern democracies: to what extent are we willing to tolerate speech whose very goal is to silence the speech of others, or to deny their rights to equal education or employment? Efforts continue to craft codes that are narrow in scope and that meet constitutional muster, with even staunch civil libertarians favoring prosecution when hate speech poses a clear and present danger of violence (e.g., Smolla 1992). The hate speech debates also mark a significant turn in the rhetoric of "censorship," one in which advocates of speech codes have attempted to enlist the power of this term in their favor. This rhetorical strategy is exemplified by the argument of Catharine MacKinnon, one of the staunchest code supporters, that "the operative definition of censorship . . . shifts from government silencing what powerless people say, to powerful people violating powerless people into silence and hiding behind state power to do it" (1993, p. 10).

### **BATTLES FOR CONTROL OVER SCHOOL CURRICULA**

Schools are the arena in which the values of open deliberation and the constructive exchange of ideas are most prized in democratic societies. It is with some irony, then, that schools are also the domain in which some of the most concerted efforts to limit the scope of deliberation have been focused. The content of textbooks used by most students in the United States has been selectively tailored to the ideological viewpoints of organized pressure groups, and at least one-third of high school students are not exposed to books and films that parents and pressure groups have successfully purged from their educational experience (Davis 1979).

From the vantage point of most teachers, parental and school board mandates to omit works of literature from the curriculum are viewed as illegitimate challenges to their expertise and infringements on their rights as professionals. In effect, teachers believe that they are in the best positions to judge the educational needs of the students they teach. However, the tradition of academic freedom that characterizes both public and private universities is rarely present in elementary and secondary education. Public school curricula are generally approved by the state or by school boards that actively regulate what is taught and what students read. School board members tend to be elected to their positions and are particularly responsive to what teachers view as unreasonable censorial demands.

Parents, on the other hand, see their interventions in the school curricula as a legitimate exercise of control over what their children read in school. They justify their decision, in part, with respect to the compulsory nature of early schooling, where they see their children as captive audiences. A number of well-funded organizations have exerted strong influences on both school boards and textbook adoption processes to influence what does and does not gain entry to the schoolroom.

In money terms, the largest impact of pressure groups has been on textbook publication and adoptions. Most states place orders for textbooks for schools in the entire state. Because publishers cannot provide multiple versions for different regions or states, texts are geared to the largest



markets, usually Texas and California (Del Fattore 1992). As a result, publishers provide their writers with guidelines that govern topics and viewpoints that are currently deemed objectionable (or desirable) in the largest states that engage in statewide adoptions. In this way, pressure groups in one state, such as Texas, often exercise veto power over the entire country's textbook market. Local communities in small and mid-size states exercise a relatively small influence on content.

Community protest and lawsuits are the principal vehicles used to challenge textbooks and reading lists in literature, social studies, and health education. An illustrative controversy—at McClintock High School in Tempe, Arizona in 1996—highlights the susceptibility of teachers and schools to parental and community demands. In this case, a parent of a McClintock High student protested the assignment of Mark Twain's classic *The Adventures of Huckleberry Finn* because the word "nigger" is included in its dialogue. In the ensuing controversy, some held that if schools were forced to shield students from exposure to the term, such intervention would undermine the legitimate authority of teachers and would endorse ignorance of American history and the practice of censorship.

We may ask whether the students grasp the distinctions among Twain's stance on racism, the use of the term "nigger" by a character in the novel, and societal endorsement of the term then and today. Of course the crux of such conflicts also rests upon the educational and social contexts in which sensitive, taboo, or potentially affronting topics that appear in the world of fiction are discussed. For instance, if *The Adventures of Huckleberry Finn* or any other literature is misused to communicate racism, the problem is larger than the choice of which books are read in the classroom.

Unfortunately, the social context of this controversy added complexity. School benches at the same high school were etched with the words "I hate niggers." Of course, there is an important distinction between eliminating the word "nigger" from *The Adventures of Huckleberry Finn* and removing "I hate niggers" from a school bench. The etching creates a hostile educational and social environment in a way that the assignment of book should not do, as long as teachers and schools are doing their jobs well. In controversy after controversy, these distinctions are as important as

they have been difficult to establish. Often teachers simply capitulate to perceptions of the most conservative or protective community elements and use readings that are sure to be "safe" (Davis 1979).

In 1990s most lawsuits were mounted by fundamentalists who condemned many books on school lists as anti-Christian, antiparent, antigovernment, immoral, and obscene. The challenged books change from year to year, but frequently banned books have included *Of Mice and Men* (Steinbeck), *Catch 22* (Heller), and *Catcher in the Rye* (Salinger). Fundamentalists are effective out of all proportion to their numbers because of the intense dedication they bring to their cause.

Research has yet to provide a sound basis for confidently assessing the effects of reading on school children's beliefs. When a story contains a character's arguments against the existence of God, might students question their faith? Qualitative research, particularly, has shown that students take widely divergent readings from stories and that these understandings are frequently critical of the viewpoints adopted by both characters and writers.

Can we maintain that literature gives readers insights into life? That it can change lives? Parents—former pupils, after all—have been encouraged to believe that great works contain great truths. It is not surprising that some of those parents who discern arguments against the social order in their children's readings seek to eliminate the threat from children's lives. Eventually, these ongoing and recurring controversies may be illuminated by scientifically grounded understanding of the social and cognitive foundations of narrative impact (Green, Strange, and Brock 2000).

One of the strongest arguments for keeping school curricula open to ideas to which we currently object or to heinous artifacts of historical prejudices is that one of the important functions of education is the development of critical inquiry. From this perspective, to turn the schoolhouse into an environment where only "safe" ideas are encountered is to do students and society a disservice. This position is rendered more compelling in the context of the current media landscape which makes it certain that students will also encounter these ideas outside the context of a critical learning environment. This argument, of course,

rests on the assumption that schools are, in fact, places where critical inquiry is fostered.

### THE SOCIAL PSYCHOLOGY OF “CENSORSHIP”

When information (e.g. a book, a film, a speech, etc.) is explicitly “censored” do people desire that information much less—in accordance with the presumed intention of the censoring agents? Are people less likely to be influenced by a point of view when its expression has been prohibited by those in authority? To answer such questions, the impact of the label “censored” can be studied experimentally by exposing people to a communication (or a “taste” of a communication) to which “censored” is affixed and by then observing the consequences of such a juxtaposition of label and communication on their values and beliefs. When this careful experimental work has been done, the answer to both of these questions is “no.” Numerous experiments have shown that explicit censorship often *backfires*: the prohibition of communications leads people to covet them and to change their attitudes in the directions the censored material advocates.

A social-commodity theory (Brock 1968) accounts for the backfire effect of censorship in terms of the psychology of supply and demand. All public communications are perceived to have some limits on their distribution, and message recipients usually have some dim awareness of the extent of these limits. However, if the distribution is explicitly limited by government statute, or other overt interventions, awareness of scarcity is sharply heightened and the desirability and impact of the communication are consequently increased. When some information becomes less available, this unavailability augments the information’s value. A decrease in supply causes an increase in demand, and perceived censorship invariably entails a perceived decrease in supply.

Experimental social psychology has illuminated the impact of censorship by showing that (a) an individual can be influenced by a censored communication, even without actually receiving it, (b) the backfire effect increases as the penalties for violating censorship are increased; and, somewhat paradoxically, (c) the backfire effect is greater when the prohibition applies selectively to one group of people. For example, for a seventeen

year old, a film that was prohibited to persons under eighteen years of age would be more desirable than a film that was prohibited to all persons. In the former case, the censorship has more impact because it is viewed as violating a personal freedom; freedom to view a film is more threatened if there are many other people (all persons over eighteen years, say) who do have that freedom.

The experimentally grounded insights from social psychology—that explicit censorship usually doesn’t work as planned, and that it often backfires—is often not factored into the planning and interventions of censoring agents. In some instances this may be good, as it reduces the effectiveness of those who attempt to limit information that is vital to public deliberation. In some instances, public information officers appear to have become aware that perceptions of censorship can hamper their efforts. In the 1970s, the Department of Defense directed that the term be completely dropped. This seemingly cosmetic change affects the way news coverage labels the sources of news reports: what was once attributed to the Office of Censorship, became, during the Gulf War, a report from the Joint Information Bureau.

Research on related forms of content advisory labels, such as “parental discretion advised,” “viewer discretion advised,” and the more specific codes signaling levels of violence, nudity, sex, and profanity on television, has begun to reveal the extent to which different kinds of labels affect the selection choices of boys and girls of different ages (Cantor and Harrison, 1994–1995). In several cases, these studies have revealed a “backfire” or “forbidden fruit” effect similar to the pattern found in reactions to the censorship label. The finding that restricted materials become more attracted—which depended on age, gender, and the kinds of labels used—have been extended to video game and music advisories.

Advocates of requiring such labels argue that their goal is not to limit communication at its source, but rather to provide parents and children a greater degree of control over their communication choices. Some critics argue that, regardless of intent, these labels lead producers to alter their productions in order to avoid undesirable labeling. As with the “censored” label, the effects on consumption of the targeted content will depend not only on how the label affects desirability of the

materials, but also on the ultimate availability of the labeled content. If parents use advisory labels to preselect their children's television viewing choices, these labels will effectively place greater selective control in the hands of parents. The probability that parents will do so is heightened by technologies like the "V-chip," through which parental choices can be automated. Similar labels may have a much different effect on the selection of music and video games, where children are more likely to make their own media selections.

Research has also begun to examine the social and psychological origins of calls to restrict deemed harmful to children or to society at large. Studies of the "third person effect" suggest that support for restrictions on such content may sometimes be grounded in systematic overestimates of the effects of media on other persons. Sociological treatments of the rise of public alarm are particularly relevant here. Nicola Beisel's (1992) account of the late-nineteenth-century attack on obscenity that resulted in the highly suppressive Comstock Laws is one example of a fruitful direction that research might take.

### INTERNATIONAL ISSUES

Whereas most democratic nations have codified and enforced protections against the most blatant forms of government suppression of sociopolitical speech, governments around the world continue to engage in harsh repression. The bounty placed on the life of Salman Rushdie in retribution for his book, *The Satanic Verses*, is a well-publicized instance of the most egregious form of government-sanctioned repression. However, a variety of both legal and extra-legal tactics continue to be used worldwide to protect those in power. To cite just two typical examples from 1999: A court in Kazakhstan suspended publication of an opposition newspaper for the three months preceding presidential elections on the pretext that it failed to indicate the place of printing on its masthead; and, shortly after the director of the Cameroon newspaper *Le Messager-Populi* was released from a ten-month prison sentence for spreading "false information," the paper's founder was driven from the country upon threats to his life (see generally, Webb and Bell 1998). Typical extra-legal means of suppressing speech critical of the government include awarding and withdrawing governmental

advertising contracts, the licensing of journalists, controls on distribution of newsprint, and, more generally, government ownership and control of broadcast media.

Controls at the site of reception provide another means of suppressing viewpoints that lack the official stamp of approval. For example, a number of countries pose outright bans on ownership of satellite receiving dishes. However, new technologies, particularly the Internet, pose serious threats to the control of political information by authoritarian regimes. The Internet's impending threat is linked to the growing dependency upon computer-based communications in the marketplace. Authoritarian governments appear to be faced with a Hobson's choice: allow connections to the vast and diverse information resources of the Internet or face economic stagnation.

Particularly when looking beyond suppressive controls on political speech, it is important to recognize the distinction between government-controlled media and government-funded media. The BBC is but one example of the government-funded, public television stations that have developed reputations for independent news coverage that gives voice to diverse perspectives of both majority and minority interests. Publicly-subsidized television in Holland instantiates an intriguing model of "free speech" through a system that allots television channels and channel space on a proportional basis to key institutions and sociocultural groups. Public television stations in many European nations seek independence from the demands of both government and marketplace through nongovernmental boards of trustees and through stable funding mechanisms not subject to annual tampering by politically driven legislatures.

### INTERNATIONAL INFORMATION FLOW

A handful of news agencies, led by the Associated Press (AP), Reuters, and Agence-France Presse dominate the news that people around the world read about issues and events originating outside their borders. The dominance of both print and broadcast news by a small group of companies is particularly felt in developing nations, ones lacking the capital to compete with the economies of scale enjoyed by entrenched foreign corporations. The concern generated in developing nations focuses not only on the news they receive from

abroad, but also on the images that news agencies project about them. This concern is compounded by news agency values that typically stress large-scale violence and disruption—"coups and earthquakes" as the saying goes—and ignore the cultural achievements and developmental efforts for which these countries would like to be known.

Those who control international news flows view their position of dominance as a desirable and natural outcome of free-market practice. Those in developing nations view the imbalances as a form of de facto neo-imperialism. This view finds its historical roots in the late-nineteenth and early-twentieth centuries, when three European news agencies, Havas (France), Wolff (Germany), and Reuters (England) carved the world into three areas of monopoly control, ones that built upon their respective colonial empires. Some insight into the present-day perspective of developing nations can be gained by examining how news about the United States was communicated in the early part of the century. During those years, the "Ring Cartel" prevented AP from sending news about the United States beyond its borders. According to Kent Cooper, the AP manager who fought to dismantle it, the cartel decided "what the people of each nation would be allowed to know of the people of other nations and in what shade of meaning the news was to be presented." It was Reuters who "told the world about Indians on the war path in the West, lynchings in the South, and bizarre crimes in the North. . . . The charge for years was that nothing credible to America was ever sent" (cited in Frederik 1993, p.39). To those in developing nations, it is the AP who now plays the role of cartel.

Several proposals have been advanced to correct the imbalance in international news flows. Notable are those advanced through the United Nations Educational, Scientific, and Cultural Organization (UNESCO) during the late 1970s and early 1980s. These calls for a "New World Information Order" (NWIO) generated immediate and prolonged controversy, and provoked the United States to withdraw from UNESCO in 1985. The contours of these debates are as intriguing as they are complex, reflecting sharp differences in news values, international economic principles, and the very definition of "free speech." Whereas NWIO supporters liken the values of developmental journalism to current trends in the United States, such

as "civic" journalism, its detractors suggest that the values of developmental journalism cater to the needs of local authority. And whereas NWIO critics argue for a "free flow" of information across international boundaries, NWIO supporters advocate a "balanced flow." Although these issues remain unresolved, promising developments include the rise of regional news agencies, programs such as CNN's World Report, in which locally produced broadcasts gain worldwide exposure; and prospects that the Internet will facilitate the international communication of locally produced news.

The increasing global dominance of U.S. movies and television programming have prompted the European Economic Community and Canada, among other regional and national entities, to seek ceilings on imported cultural goods as exceptions to free trade principles in international trade agreements. These demands reflect the belief that when applied to the area of expression, free market principles support forms of "soft" or "de facto censorship" in which locally produced speech is driven from the marketplace of ideas (cf. Frederik 1993). More generally, these demands reflect the growing tension between the ongoing course of globalization and the threats to cultural sovereignty it entails.

## REFERENCES

- Bagdikian, B. 1997 *The Media Monopoly*. Boston: Beacon Press.
- Baird, R.M., and S.E. Rosenbaum 1991 *Pornography: Private Right or Public Menace?* Buffalo, N.Y.: Prometheus Books.
- Baker, C.E. 1998 "The Media that Citizens Need." *University of Pennsylvania Law Review* 147, 317.
- Barbur, B. R. 1996 *Jihad vs. McWorld: How Globalism and Tribalism are Reshaping the World*. New York: Ballantine Books.
- Beisel, Nicola. 1992 "Constructing a Shifting Moral Boundary: Literature and Obscenity in Nineteenth-Century America." In M. Lamont and M. Fournier, Eds., *Cultivating Differences: Symbolic Boundaries and the Making of Inequality*. Chicago: University of Chicago Press.
- Bourdieu, P. 1991. "Censorship and the Imposition of Form." In J.B. Thompson, Ed. *Language and Symbolic Power*. Cambridge, Mass.: Harvard University Press.
- Brock, T.C. 1968. "Implications of Commodity Theory for Value Change." In A.G. Greenwald, T.C. Brock,

- and T.M.Ostrom, eds. *Psychological Foundations of Attitudes*. New York: Academic Press.
- Cantor, J., and K. Harrison 1994–1995 Ratings and Advisories for Television Programming.” In *National Television Violence Study: Scientific Papers*. Los Angeles: Mediascope Inc.
- Christenson, Peter G., and Donald F. Roberts 1998. “It’s Not Only Rock and Roll: Popular Music in the Lives of Adolescents.” Cresswell, N.J.: Hampton Press.
- Court Case Citation, *Gitlow vs. New York*, 268 U.S. 652, 1925.
- Court Case Citation, *Red Lion Broadcasting vs. FCC*, 395 U.S. 367, 1969.
- Court Case Citation, *Miller vs. California*, 413 U.S. 15, 1973.
- Curran, J., and M. Gurevitch 1996 *Mass Media and Society*. London: Edward Arnold.
- Davis, J. Ed. 1979. *Dealing with Censorship*. Urbana, Ill.: National Council of Teachers of English.
- Deetz, S. A. 1992. *Democracy in an Age of Corporate Colonization: Developments in Communication and the Politics of Everyday Life*. New York: State University of New York Press.
- Del Fattore, J. 1992. *What Johnny Shouldn’t Read: Textbook Censorship in America*. New Haven, Conn.: Yale University Press.
- Dennis, E. E., et al. 1991. *The Media at War: The Press and the Persian Gulf Conflict*. New York: Gannett Foundation Media Center.
- Donnerstein, E., and S.L. Smith 1997. “Impact of Media Violence on Children, Adolescents, and Adults.” In S. Kirschner and D. A. Kirschner, Eds., *Perspectives on Psychology and the Media*. Washington, D.C.: American Psychological Association.
- Emerson, T. I. 1973 *The System of Free Expression*. New York: Random House.
- Foerstel, H.N. 1997. *Free Expression and Censorship in America: An Encyclopedia*. Westport, Conn.: Greenwood Press.
- Frederick, H.H. 1993 *Global Communication and International Relations*. Belmont, Calif.: Wadsworth.
- Green, J. 1990 *The Encyclopedia of Censorship*. New York: Facts on file.
- Green, M.C., J.J. Strange, and T.C. Brock (eds.) 2000. *Narrative Impact: Social and Cognitive Foundations*. Mahwah, N.J.: Erlbaum Associates.
- Habermas, J. 1989 *The Structural Transformation of the Public Sphere*. Cambridge Mass.: Polity.
- Hallin, D. C. 1986. *The “Uncensored War”: The Media and Vietnam*. New York: Oxford University Press.
- Jensen, C. 1997 *20 Years of Censored News*. New York: Seven Stories Press.
- MacKinnon, C. A. 1993 *Only Words*. Cambridge, Mass.: Harvard University Press.
- Post, R.C. (ed.) 1998 *Censorship and Silencing: Practices of Cultural Regulation*. Los Angeles: The Getty Research Institute for the History of Art and the Humanities.
- Shoemaker, P.J., and S.D. Reese 1996 *Mediating the Message: Theories of Influences on Mass Media Content*. New York: Longman.
- Smolla, R.A. 1993 *Free Speech in an Open Society*. New York: Vintage Books.
- Stoller, R.J, and I.S. Levine 1993 *Coming Attractions: The Making of an X-Rated Video*. New Haven, Conn.: Yale University Press.
- Sunstein, C. 1993. *Democracy and the Problem of Free Speech*. New York: Free Press.
- Webb, W.L, and R. Bell (eds.) 1998 *An Embarrassment of Tyrannies: Twenty-Five Years of Index on Censorship*. New York: George Braziller.
- Zillmann, D., and J. Bryant (eds.) 1989 *Pornography: Research Advances and Policy Considerations*. Hillsdale, NJ: Erlbaum.

JEFFREY J. STRANGE  
MELANIE C. GREEN  
TIMOTHY C. BROCK

## CENSUS

A national census of population is “the total process of collecting, compiling, evaluating, analysing and publishing or otherwise disseminating demographic, economic, and social data pertaining, at a specified time, to all persons in a country” (United Nations 1988, p. 3). The United Nations encourages its members to take regular censuses and provides technical assistance. It also publishes an annual demographic yearbook and maintains a web site with selected census data and other population information for most countries.

A nation may conduct censuses on other subjects such as housing, business firms, agriculture, and local governments. Because of the close link between families and housing units, censuses of population and housing are usually combined.

The actual enumeration for a national population census is usually spread over a period of weeks

or months, but an attempt is made to record circumstances as of a designated census day (April 1 for the 2000 U.S. census). In some censuses, persons are reported in their legal or usual place of residence (*a de jure enumeration*). According to the Bible, a census decree issued by the government in Rome ordered that persons be counted and taxed in their home towns. In response, Joseph and his pregnant wife, Mary, traveled from Nazareth to Bethlehem, where Mary gave birth to Jesus. Other censuses record people where they are on the census day (*a de facto enumeration*). Whatever residency rule is followed, all conceivable special circumstances must be anticipated, questionnaires well designed, and census staff thoroughly trained.

The U.S. census uses the concept of “usual residence,” where the person lives and sleeps most of the time, but some exceptions are made. A few examples from the 2000 U.S. census illustrate the range of circumstances. Some persons staying in hotels and rooming houses are transients temporarily absent from their usual homes; they completed special census forms that were later compared to the census forms received from their home addresses to confirm that they were properly reported by other household members. An operation called “Service-Based Enumeration” was designed to locate homeless persons and others with no usual residence. College students were presumed to be residents at the place they lived while attending college, even if they were at home on spring break on census day. Efforts were made to avoid having an away-at-college child erroneously double-counted as a member of the home household. Pre-college students at away-from-home schools were assigned to the parental household. Citizens working in another country for the U.S. government, and their families, were enumerated (primarily through the use of administrative records), but other citizens living abroad at the census date were not counted. Short-term tourists, however, were supposed to be reported by other members of their households, or by late enumeration upon their return home. Persons not legally resident in the United States were supposed to be counted, but many feared all contact with authorities and took steps to avoid detection by census procedures.

The total population count is only one result of a modern census. Information is also obtained

on many characteristics, such as age, sex, and relationship to others living in the same household, educational level, and occupation. The census then reports on population size and characteristics for provinces, states, counties, cities, villages, and other administrative units. To obtain all this information, the 2000 U.S. census (U.S. Census Bureau 1999) asked every household to complete a “short form” with six questions about each person and one question about each housing unit. A carefully selected sample of one of every six households received a “long form” that included additional questions on fourteen topics for each person and thirteen topics for each housing unit.

Reporting information for the entire country and for the thousands of cities and other subareas results in many volumes of printed reports. Many nations issue additional data on computer-readable files designed for convenient use by national and local governments, organizations, and individuals. Increasing quantities of data are available to anyone with access to the Internet.

In the United States, the idea for a regular census emerged during debates about the problems of creating a representative form of government. The U.S. Constitution (Article I, Section 2) directs that membership in the House of Representatives be based on population: “The actual enumeration shall be made within three years after the first meeting of the Congress of the United States, and within every subsequent term of ten years, in such manner as they shall by law direct.”

The questions asked in a census reflect political and social issues of the day and often provoke spirited debates. For example, slavery was a recurrent divisive issue for the Constitutional Convention. One of the two newly created legislative bodies, the House of Representatives, was to have the number of legislators from each state proportional to the population of the state. Including slaves in the population count would increase the legislative power of southern states. A compromise (Article I, Section 2) provided that: “Representatives. . . shall be apportioned among the several states. . . according to their respective numbers, which shall be determined by adding to the whole number of free persons, excluding Indians not taxed, three-fifths of all other persons.” The distinction between “free persons” and “all other

persons” (slaves) was eliminated after the Civil War (Amendment 14, Section 2).

The 1790 census was modest in scope. Assistants to federal marshals made lists of households, recording for each household the number of persons in five categories: free white males over sixteen and under sixteen, free white females, other free persons, and slaves (Anderson 1988, p. 13).

Censuses from 1790 through 1840 were conducted with little central organization or statistical expertise. A temporary federal census office was established to conduct the 1850 census. The individual rather than the household became the focus of the enumeration, and the content of the census was expanded to include occupation, country or state of birth, and other items. Experienced statisticians were consulted, and the United States participated in the first International Statistical Congress in 1853.

The temporary office for the 1890 census became one of the largest federal agencies, employing 47,000 enumerators and 3,000 clerical workers. To help with the enormous task of tallying data, census officials supported the development by a young inventor, Herman Hollerith, of an electrical tabulating machine. His punched-card system proved effective in census operations and later contributed to the growth of the IBM corporation.

A permanent census office was established in 1902. An increasingly professional staff assumed responsibility for the population censuses and a broad range of other statistical activities. Deficiencies in the federal statistical system became apparent during the 1930s as the nation tried to assess the effects of the Great Depression and to analyze an array of new programs and policies. Several federal agencies successfully advocated expansion of the social and economic content of the 1940 census. New questions in 1940 asked about participation in the labor force, earnings, education, migration (place of residence in 1935), and fertility. A housing census was paired with the population census to provide information about housing values and rents, mortgages, condition of dwellings, water supply, and other property issues. The practice of providing population and housing information for subareas of large cities was greatly expanded.

Governmental statistical agencies are often set in their ways, concerned with continuity rather than innovation and removed from the higher levels of policymaking. For several decades beginning in the late 1930s, the U.S. Census Bureau was extraordinarily creative. Social scientists and statisticians employed by the Census Bureau were active in research and scholarly publication and played leadership roles in professional organizations. Census Bureau personnel pioneered in development of the theory and practice of population sampling. To accommodate large increases in content and geographic scope, the 1940 census questionnaire was divided into two parts. A set of basic questions was asked of everyone, while a set of supplementary questions was asked of a one-in-twenty sample. Subsequent U.S. censuses have continued to use sampling, as with the “long-form” versions of the 1990 and 2000 questionnaires that went to one of every six households. Sampling theory was also applied to the development of periodic sample surveys to provide timely information between censuses and to provide information on additional topics.

For the 1950 census, the Bureau participated actively in the development and utilization of another new technology, the computer. In later censuses, the schedules were microfilmed and optically scanned for direct input of information (without names and addresses) into a computer for error-checking, coding, and tabulating.

Through 1950, census enumeration in the United States was accomplished almost entirely by having an enumerator conduct a personal interview with one or more members of each household and write the information on a special form. By 1990, most census questionnaires were distributed and returned to the Census Bureau by mail, with telephone follow-up replacing many personal visits. Large-scale use of the Internet and other innovative response technologies may be practicable for widespread use in the 2010 census.

National censuses have so far been the best method for obtaining detailed information about the entire population. Only censuses provide counts and characteristics for small administrative and statistical areas, such as villages, voting districts, city blocks, and census tracts. Only censuses provide reliable information on small population

groups, such as the income distribution of female plumbers. Only censuses provide the large numbers of cases needed for complex multivariate analyses. But censuses are so large and difficult to process that by the time the data are released they are, for some purposes, out of date, and this problem increases throughout the interval, commonly ten years, until the next census.

To address the timeliness problem, most national statistical agencies conduct a series of large sample surveys that provide select information on key topics at timely intervals. In the United States, for example, the latest unemployment rate, released on the first Friday of each month, is headline news. It is based on the Current Population Survey, a monthly sample survey that provides current information on many social and economic characteristics.

A few countries keep population registers, continuous records of where people live along with some of their characteristics. If reasonably complete and accurate, these may be used to supplement and update some census information. In many nations, administrative records such as social security, national health, tax, and utility company files, may be adapted to provide periodic estimates of population size and a few characteristics, but difficult questions arise about data quality and comparability with census- and survey-based information. Citizens of many nations are wary about letting the government and corporations compile extensive personal data. Legislative restrictions are increasingly being placed on what information may be gathered and stored and how it may be used.

The questions asked in successive censuses typically change more slowly than the procedures. Keeping the same topics and the same wordings of questions help a government measure change from one census to the next. This appeals to researchers and policy analysts, but policymakers and administrators, whose attention is focused mainly on current programs and next year's budget, often plead for new wording to better serve current concepts. Another reason tending to stifle innovation is that the census is an expensive and visible tool. A lengthy review process confronts any agency that seeks to add, delete, or alter a question. In the United States, both the executive branch and the

Congress must approve the final census schedule. A question on pet ownership has been regularly proposed but rejected because there is no compelling governmental interest in such a question and private sample surveys can provide the desired information. A third reason the content of the questionnaire changes little is that proponents of new topics and questions must compete for questionnaire space, while questions previously asked tend to already have a network of users who have a vested interest in retaining the topic. This competition spurs extensive lobbying and mobilization of support from federal agencies, congressional committees, and interest groups.

An innovative approach to providing timely data is "continuous measurement" using a "rolling-sample" survey. The *American Community Survey* has been implemented by the U.S. Census Bureau (1999), with plans to expand to three million households a year in 2003. Because the survey uses a small, permanent, well-trained and supervised staff (as compared to the large, temporary, briefly trained staff utilized for a census), data quality may be higher than in a census. Another advantage of a monthly survey over a decennial census is the ease of introducing new topics and testing the effects of changes in the wording of questions.

*American Community* survey results will be cumulated over a year to provide estimates of numbers and characteristics for the nation, states, and places or groups of 65,000 or more people. For smaller places and groups, or for more reliable and detailed data for larger places, data will be cumulated for periods of up to five years. The five-year cumulated sample is designed to be approximately equivalent to the census "long-form" sample and may eliminate the need for a "long form" in the 2010 census. Moving averages could provide annual updates and be used for time series. There are many issues to be worked through in developing and evaluating this type of survey and in determining how well information averaged over a long period compares to standard statistical measures based on time-specific censuses and surveys.

In many nations, questions on race and ethnicity are a sensitive and contentious topic. In the United States, the groups recognized have changed



from each census to the next. For example, special tallies were made from the 1950 and 1960 censuses of persons with Spanish surnames, but only for five southwestern states. Beginning in 1970, and elaborated in later censuses, persons were asked if they were of Spanish/Hispanic origin. Many responses to this question have seemed to be inconsistent with responses to the separate question on race. The Bureau has conducted sample surveys to test various question wordings, and has sponsored field research and in-depth interviews to provide insights into how people interpret and respond to questions about ethnicity. For the 2000 census the basic concepts were retained, but the wording of each was adjusted, provision was made for individuals to report more than one race, and the question on Hispanic ethnicity was placed before the question on race.

In recent decades, extensive social science research has been conducted in many parts of the world on issues of racial and ethnic identity. These identities are almost always more flexible and more complex than can be captured with simple questions. A further complication is that many persons have ancestral or personal links to two or more racial and ethnic groups; how they respond depends on how they perceive the legitimacy and purposes of the census or survey.

In the 1990s, the U.S. Office of Management and Budget conducted an extensive review of its policy statement that specifies the standard race and ethnic classifications to be used by all government agencies, including the Census Bureau (Edmonston and Schulze 1995). Serious methodological problems have arisen as a result of inconsistencies among classifications of individuals on repeated interviews and in different records. A dramatic example occurs with infant mortality rates for race and ethnic groups, which are based on the ratio of counts based on birth registration to counts based on death registration of infants less than one year old. Comparing birth and death certificates for the same person, recorded less than one year apart, revealed large numbers for whom the race/ethnic classification reported at birth differed from that reported at death. The high-level government review also heard from many interest groups, such as Arab-Americans, multiracial persons, and the indigenous people of Hawaii,

who were anxious for the government to enumerate and classify their group appropriately.

In the United Kingdom, controversy about a proposed question on ethnic identity led to the question's omission from the 1981 census schedule, thus hindering analyses of an increasingly diverse population.

The most politically intense and litigious controversy about recent U.S. censuses is not that of content but of accuracy. States, localities, and many interest groups have a stake in the many billions of dollars of government funds that are distributed annually based in part on census numbers.

President George Washington commented about the first census that "our real numbers will exceed, greatly, the official returns of them; because the religious scruples of some would not allow them to give in their lists; the fears of others that it was intended as the foundation of a tax induced them to conceal or diminish theirs; and through the indolence of the people and the negligence of many of the Officers, numbers are omitted" (quoted in Scott 1968, p. 20).

A perfect census of a school, church, or other local organization is sometimes possible, if membership is clearly defined and the organization has up-to-date and accurate records, or if a quick and easily monitored enumeration is feasible. A national census, however, is a large-scale social process that utilizes many organizations and depends on cooperation from masses of individuals. Planning, execution, and tabulation of the 2000 U.S. Census of Population extended over more than ten years. Hundreds of thousands of people were employed at a cost of several billion dollars. A discrepancy between the results of a national census and "our real numbers" is inevitable.

Statisticians, recognizing that error is ubiquitous, have developed many models for identifying, measuring, and adjusting or compensating for error. Census statisticians and demographers around the world have participated in these developments and in trying to put professional insights to practical use.

Following the 1940 U.S. census, studies comparing birth certificates and selective service records to census results documented a sizable net

undercount. To provide more information on census coverage and accuracy, a post-enumeration survey was conducted following the 1950 census using specially selected and closely supervised enumerators. One finding from the survey was that the undercount of infants in the census did not arise, as first thought, from a tendency for new parents to forget to mention a new baby to the census enumerator. The problem, rather, was that many young couples and single parents had irregular and difficult-to-find living arrangements, and entire households were missed by census enumerators. Based on this and related findings, much effort was given in subsequent censuses to precensus and postcensus review of lists of dwelling units.

Special surveys preceding, accompanying, and following censuses are increasingly used to provide evidence on magnitude of error, location of error in specific groups or places, and procedural means for reducing error. Techniques of “demographic analysis” have also been developed and continually refined to provide evidence of census error. The basic technique is to analyze the numbers of persons of each age, sex, and race in successive censuses. For example, the number of twenty-eight year-old white women in the 1990 census should be consistent with the number of eighteen-year-olds in 1980 and eight-year-olds in 1970. Information on births, deaths, immigration, and emigration is taken into account. Based on demographic analysis, the estimated net undercount for the total U.S. population was about 5.6 percent in 1940, 1.4 percent in 1980 (Fay, Passel, and Robinson 1988, Table 3.2) and 1.8 percent in 1990. Estimates of net undercount have also been made for specific age, sex, and race groupings.

If the net undercount were uniform for all population groups and geographic regions, it would not affect equity in the distribution of seats in Congress or public funds. But census errors are not uniform. The estimates for 1990 show net undercounts exceeding 10 percent of adult black males and small net overcounts for some age and race groups.

In 1980, the mayor of Detroit, the City of Detroit, New York City, and others sued the federal government, alleging violation of their constitutional rights to equal representation and fair distribution of federal funds (Mitroff, Mason, and

Barabba 1983). More litigation occurred with respect to the 1990 census. In neither case were the initially reported census counts adjusted to reflect estimated errors in enumeration, although in the case of the 1990 census the Census Bureau technical staff and director thought their methodology would improve the accuracy of the counts.

The Census Bureau developed innovative plans for the 2000 census to integrate an evaluation survey with the census, to use sampling to increase quality and reduce costs of enumerating nonresponsive households, and to produce official counts that already incorporate the best available procedures for minimizing error.

These Census Bureau plans engendered major political and budgetary battles between the Republican-controlled Congress and the Democratic administration. Many of the persons hardest to reach using traditional census methods are poor, often minorities. Many members of Congress assumed that a census with near-zero undercount would increase the population reported for states and localities that tend to vote more for Democrats, and hence reduce, relatively, the representation for states and localities that tend to vote more for Republicans. At the national level, this could alter the reapportionment among states of seats in the House of Representatives. Legislatures in states, counties, and cities are also subject to decennial redistricting according to the “one-person, one-vote” rule, and similar shifts in relative political power could occur.

The controversy between legislative and administrative branches over methods for the 2000 census was taken to the Supreme Court, with plaintiffs seeking a ruling that the Bureau’s plans to use sample-based estimates would be unconstitutional. The Supreme Court avoided the constitutional argument, but ruled that current law did not permit such use of sampling. The Democrats lacked the power to rewrite that part of the basic census law. The two parties compromised by appropriating additional dollars to cover the extra costs of complete enumeration while retaining funds for many of the methodological innovations that permit preparation of improved estimates. Reapportionment among states of seats in Congress could therefore use the results of traditional complete enumeration. Most scholars, analysts, and other

users of census data are likely to consider the adjusted estimates more accurate and hence more useful.

The percentage net undercount understates the total coverage error in the census. If one person is omitted and another person counted twice, the total count is correct but there are two “gross” errors that may affect the counts for specific places and characteristics. Identification of gross errors is receiving increased attention, to improve information on the nature of census error and potential error-reducing methods, and to facilitate development of techniques of data analysis that compensate for known and unknown errors.

One of the difficulties in debates about undercount and other census error arises from confusion between the idea of a knowable true count and the reality that there is no feasible method for determining with perfect accuracy the size and characteristics of any large population. A national census is a set of procedures adopted in a political, economic, and social context to produce population estimates. It is politically convenient if all parties accept these results as the common basis for further action, much as sports contests use decisions by umpires or referees. The more that policy and budgets depend on census results and the more aware that politicians and citizens become of the importance of the census, the more contentious census taking will be.

Every nation confronts political and social problems with its censuses. Regularly scheduled censuses are often postponed or abandoned because of international conflict. The United States and the United Kingdom canceled plans for mid-decade censuses because of national budget constraints. West Germany was unable to take a census for several years because of citizen fears about invasions of privacy. Ethnic conflict has interfered with census taking in India, Lebanon, and other nations.

The processes by which census procedures are determined, the ways in which census figures are used, and the conflicts that occur about these procedures and numbers are not merely “technical” but are embedded in a broader social process. The character of a nation’s census and the conflicts that surround it are core topics for the “sociology of official statistics” (Starr 1987).

(SEE ALSO: *Demography; Population*)

#### REFERENCES

- Anderson, Margo J. 1988 *The American Census: A Social History*. New Haven: Yale University Press.
- and Stephen E. Fienberg 1999 *Who Counts? The Politics of Census-Taking in Contemporary America*. New York: Russell Sage Foundation.
- Edmonston, Barry, and Charles Schultze 1995 *Modernizing the U.S. Census*. Washington, D.C.: National Academy Press.
- Mitroff, Ian I., Richard O. Mason, and Vincent P. Barabba 1983 *The 1980 Census: Policymaking and Turbulence*. Lexington, Mass.: Lexington Books.
- Scott, Ann H. 1968 *Census, USA: Fact Finding for the American People, 1790–1970*. New York: Seabury Press.
- Starr, Paul 1987 “The Sociology of Official Statistics.” In W. Alonso and P. Starr, eds., *The Politics of Numbers*. New York: Russell Sage Foundation.
- Steffey, Duane, and Norman Bradburn (eds.) 1994 *Counting People in the Information Age*. Washington, D.C.: National Academy Press.
- United Nations 1998 *Principles and Recommendations for Population and Housing Censuses*. ST/ESA/STAT/SER.M/67/Rev.1. New York: United Nations.
- United States Census Bureau. *Census Bureau Home Page*. On the Internet at <http://www.census.gov>. See links to *Census 2000* and *American Community Survey*.

KARL TAEUBER

## CHANGE MEASUREMENT

See Experiments; Longitudinal Research; Quasi-Experimental Research Design; Measurement.

## CHILD ABUSE

See Childhood Sexual Abuse; Family Violence; Incest; Sexual Violence.

## CHILDBEARING

See Family Planning; Family Size; Fertility Determinants; Pregnancy and Pregnancy Termination.

## CHILDHOOD SEXUAL ABUSE

Although child abuse is probably as old as childhood itself, serious research into child abuse arguably began in 1962 with the publication of Kempe and colleagues' seminal paper, "The Battered Child Syndrome" (Kempe, Silverman, Steele, Droegemuller, and Silver 1962). Not long afterwards, health care professionals began to direct their attention to the specific problem of childhood sexual abuse (Cosentino and Collins 1996). In the past few decades, numerous instances have been documented in detailed case histories, and important research into the causes and consequences of childhood sexual abuse has been initiated.

Numerous extensive reviews have been published that summarize what is presently known about childhood sexual abuse, focusing on the following domains.

- Short-term effects (Beitchman et al. 1991; Beitchman et al. 1992; Briere and Elliott 1994; Browne and Finkelhor 1986; Finkelhor 1990; Gomes-Schwartz et al. 1990; Green 1993; Kelley 1995; Kendall-Tackett et al. 1993; Trickett and McBride-Chang, 1995).
- Long-term consequences (Beitchman et al. 1991; Briere 1988; Briere and Elliott 1994; Briere and Runtz 1991; Cahill et al. 1991a; Collings 1995; Ferguson 1997; Finkelhor 1987; Gibbons 1996; Glod 1993; Green 1993; Murray 1993; Polusny and Follette 1995; Trickett and McBride-Chang 1995; Wolfe and Birt, 1995).
- Prevention of abuse (Adler and McCain 1994; Berrick and Barth 1992; MacMillan et al. 1994; Olsen and Widom 1993; Wolfe et al. 1995).
- Treatment of both survivors and abusers (Cahill et al. 1991b; Cosentino and Collins 1996; Faller 1993; Finkelhor and Berliner 1995; O'Donohue and Elliot 1993).

Consequently, this chapter is not intended to provide a detailed analysis and review of the literature on childhood sexual abuse. Rather, it is meant to serve as a brief overview of, and introduction to, this area of inquiry.

## DEFINITION OF THE PROBLEM

Before delving into the field of childhood sexual abuse, one must first understand what is meant by the term. Indeed, the lack of a standard definition has been a major criticism of the field and controversies abound (Finkelhor 1994a; Gibbons 1996; Gough 1996a; Green 1993). In general, the legal and practical research definitions of child sexual abuse require the following two elements: (1) sexual activities involving a child (sometimes construed as including adolescence), and (2) the existence of an "abuse condition" indicating lack of consensuality (Faller 1993; Finkelhor 1994a). "Sexual activities" refer to behaviors intended for sexual stimulation; such activities need not involve physical contact, however, leading to separate definitions for "contact sexual abuse" and "noncontact sexual abuse." Contact sexual abuse includes both penetrative (e.g., insertion of penis or other object into the vagina or anus) and nonpenetrative (e.g., unwanted touching of genitals) acts. Noncontact sexual abuse refers to activities such as exhibitionism, voyeurism, and child pornography (see Faller 1993 and Finkelhor 1994a for reviews on definitions of childhood sexual abuse).

An abuse condition is said to exist when there is reason to believe that the child either did not, or was incapable of, consenting to sexual activity. Three main conditions can be distinguished: (1) the perpetrator has a large age or maturational advantage over the child; or (2) the perpetrator is in a position of authority or in a caretaking relationship with the child; or (3) activities are carried out against the child's will using force or trickery. As evident from this brief summary, "Childhood Sexual Abuse" covers a wide range of acts and situations, and therefore is open to considerable subjective interpretation.

## INCIDENCE AND PREVALENCE OF CHILDHOOD SEXUAL ABUSE

Increasing attention has been directed toward childhood sexual abuse not only because of the psychosocial sequelae associated with its occurrence, but also because it now appears to be more widespread than previously thought (Adler and McCain 1994). However, accurate estimates of the occurrence of childhood sexual abuse are difficult

to obtain because many cases are never reported. Thus, available statistics represent only those cases reported to child protection agencies or to law enforcement, and therefore underestimate the true magnitude of the problem.

There are two official sources of incidence data for cases in the United States: The National Incidence Study of Child Abuse and Neglect (NIS), a federally funded research project; and statistics from state child protection agencies (Finkelhor 1994a). The NIS conducted in 1993 documented over 300,000 cases of sexual abuse among children known to professionals in the course of a year, or a rate of approximately forty-five cases per 10,000 children (Sedlak and Broadhurst 1996). Child protective services data suggest that approximately 140,000 cases of childhood sexual abuse occur annually, or twenty-one cases per 10,000 children (Leventhal 1998). An editorial by Finkelhor (1998) suggested that the incidence may be declining. However, this remains to be substantiated. In general, reports suggest that the rate of childhood sexual abuse has increased substantially over the past decades (Sedlak and Broadhurst 1996).

Retrospective surveys provide a second source of information on the occurrence of childhood sexual abuse. Reported rates depend upon how it is defined and operationalized in any given survey. When childhood sexual abuse has been assessed with a single item that narrowly defines it as rape or sexual intercourse, reported prevalence rates tend to be low (e.g., less than 12 percent). Conversely, when it is defined more broadly (e.g., touching genitalia) and assessed using multiple items, prevalence rates tend to be much higher, but with a wide range. The discrepancies observed in the estimated prevalence of childhood sexual abuse points to the need for increased standardization and the use of better assessment instruments in this research.

Because rates of childhood sexual abuse are substantially greater among females than males (e.g., Cosentino and Collins 1996; Finkelhor et al. 1990; Sedlak and Broadhurst 1996), the majority of this research has focused on women. Surveys suggest that anywhere from 18 percent to 30 percent of college women and 8 percent to 33 percent of male and female high school students report having experienced this abuse at some point in

their lives (Ferguson 1997; Finkelhor 1994a; Gibbons 1996; Green 1993; Gorey and Leslie 1997). Among the general adult female population, prevalence rates range from 2 percent to 62 percent (Finkelhor 1987; Finkelhor et al. 1990; Saunders et al. 1992). Rates are even higher within various clinical populations, with 35 percent to 75 percent of female clients reporting a history of some form of sexual abuse during childhood (Gibbons 1996; Wurr and Partridge 1996).

The number of males who have been sexually abused is difficult to estimate because it has been the subject of fewer high-quality studies (Holmes and Slap 1998; Watkins and Bentovim 1992). Nonetheless, a few studies have examined this issue among men and estimate that approximately 3 percent to 16 percent of all men in the United States were sexually abused during childhood (Finkelhor et al. 1990; Gibbons 1996; Holmes and Slap 1998). As is the case for women, rates are higher when studying clinical populations, with estimates ranging from 13 percent to 23 percent (Holmes and Slap 1998; Metcalfe et al. 1990).

In sum, based on the evidence available in the literature, Finkelhor (1994a) estimates that 20 percent of adult women and 5 percent to 10 percent of adult men are survivors of childhood sexual abuse. Similarly, a synthesis of sixteen cross-sectional surveys on the prevalence of it among nonclinical populations reported unadjusted estimates of 22.3 percent for women and 8.5 percent for men (Gorey and Leslie 1997).

### SEQUELAE

Research conducted over the past decade indicates that a wide range of psychological and interpersonal problems are more prevalent among those people who have been sexually abused than among those who have not been sexually abused. Although a definitive casual relationship between such problems and sexual abuse cannot be established using retrospective or cross-sectional research methodologies (Briere 1992a; Plunkett and Oates 1990), the aggregate of consistent findings in this literature has led many investigators and health care providers to conclude that childhood sexual abuse is a major risk factor for a variety of problems, both in the short-term and in later

adulthood (Briere and Elliott 1994; Faller 1993). The various problems and symptoms described in the literature can be categorized as follows: (1) posttraumatic stress; (2) cognitive distortions; (3) emotional distress; (4) impaired sense of self; (5) avoidance phenomena; (6) personality disorders; and (7) interpersonal difficulties. In the remainder of this section, each of these categories is defined and illustrated using examples from the research literature.

Posttraumatic stress refers to certain enduring psychological symptoms that occur in reaction to a highly distressing, psychiatric disruptive event. To be diagnosed with posttraumatic stress disorder (PTSD) an individual must experience not only a traumatic event, but also the following problems: (1) frequent re-experiencing of the event through nightmares or intrusive thoughts; (2) a numbing of general responsiveness to, or avoidance of, current events; and (3) persistent symptoms of increased arousal, such as jumpiness, sleep disturbances, or poor concentration (American Psychiatric Association (APA) 1994). In general, researchers have found that children and adults who have been sexually abused are significantly more likely to receive a PTSD diagnosis than their nonabused peers (McLeer et al. 1992; Murray 1993; Rowan and Foy 1993; Saunders et al. 1992; Wolfe and Birt 1995).

Cognitive distortions are negative perceptions and beliefs held with respect to oneself, others, the environment, and the future. In the abused individual, this type of thought process is reflected in his or her tendency to overestimate the amount of danger or adversity in the world and to underestimate his or her worth (Briere and Elliott 1994; Dutton et al. 1994; Janoff-Bulman 1992). Numerous studies document such feelings and perceptions such as helplessness and hopelessness, impaired trust, self-blame, and low self-esteem among children who have been sexually abused (Oates et al. 1985). Moreover, these cognitive distortions often continue on into adolescence and adulthood (Gold 1986; Shapiro and Dominiak 1990).

Emotional distress or pain, which typically manifests itself as depression, anxiety, anger, or all of these, is reported by many survivors of childhood sexual abuse. Depression is the most commonly reported symptom among adults with a

history of childhood sexual abuse (Beitchman et al. 1992; Browne and Finkelhor 1986; Cahill et al. 1991a; Polusny and Follette 1995). Greater depressive symptomatology is also found among children who have been abused (Lipovsky et al. 1989; Yama et al. 1993).

Similarly, elevated anxiety levels have been documented in child victims of sexual abuse and adults who have a history of it (Gomes-Schwartz et al. 1990; Mancini et al. 1995; Yama et al. 1993). Adults with a history of childhood sexual abuse are more likely than their nonabused counterparts to meet the criteria for generalized anxiety disorder, phobias, panic disorder, obsessive compulsive disorder, or all of these (Mancini et al. 1995; Mulder et al. 1998; Saunders et al. 1992).

Another common emotional sequel of childhood sexual abuse is anger; specifically, chronic irritability, unexpected or uncontrollable feelings of anger, difficulties associated with expressing anger, or all of these (Briere and Elliott 1994; Van der Kolk et al. 1994). During childhood and adolescence, anger is most likely to be reflected in behavioral problems such as fighting, bullying, or attacking other children (Chaffin et al. 1997; Garnefski and Diekstra 1997).

Sexual abuse also can damage a child's developing sense of self, with adverse long-term consequences. The development of a sense of self is thought to be one of the earliest developmental tasks of the infant and young child, typically unfolding in the context of early relationships (Alexander 1992). Thus, how a child is treated early in life critically influences his or her growing self-awareness. Childhood sexual abuse can interfere with this process, preventing the child from establishing a strong self image (Cole and Putnam 1992). Without a healthy sense of self, the person is unable to soothe or comfort himself or herself adequately, which may lead to overreactions to stressful or painful situations, and to an increased likelihood of re-victimization (Messman and Long 1996). Indeed, numerous studies have found high rates of sexual re-victimization (e.g., rape, coercive sexual experience) among individuals reporting a history of childhood sexual abuse (Fergusson et al. 1997; Urquiza and Goodlin-Jones 1994; Wyatt et al. 1992).

Avoidance is another major response to having been sexually abused. Avoiding activities among

victims may be viewed as attempts to cope with the chronic trauma and dysphoria induced by the abuse. Among the dysfunctional activities associated with efforts to avoid recalling or reliving specific memories are: dissociative phenomena, such as losing time, repression of unpleasant memories, detachment, or body numbing (APA 1994; Cloitre et al. 1997; Mulder et al. 1998; Nash et al. 1993); substance abuse and addiction (Arellano 1996; Briere 1988; Wilsnack et al. 1997); suicide or suicidal ideation (Briere and Runtz 1991; Saunders et al. 1992; Van der Kolk et al. 1991); inappropriate, indiscriminate, and/or compulsive sexual behavior (McClellan et al. 1996; Widom and Kuhns, 1996), (for reviews see Friedrich 1993; Kendall-Tackett et al. 1993; Tharinger 1990); eating disorders (Connors and Morse 1993; Douzinas et al. 1994; Schwartz and Cohn 1996; Wonderlich et al. 1997); and self-mutilation (Briere 1988; Briere and Elliott 1994). Each of these behaviors serve to prevent the individual from experiencing the considerable pain of abuse-specific awareness and thus reduces the distress associated with remembrance. However, avoidance and self-destructive methods of coping with the abuse may ultimately lead to higher levels of symptomatology, lower self-esteem, and greater feelings of guilt and anger (Leitenberg et al. 1992).

Numerous investigations have found a greater rate of borderline personality and dissociative identity (formerly multiple personality disorder) disorders among adult female survivors of childhood sexual abuse (Green 1993; Polusny and Follette 1995; Silk et al. 1997). Borderline personality disorder includes symptoms of impulsiveness associated with intense anger or suicidal, self-mutilating behavior, and affective instability with depression which are typical sequelae in sexually abused children and in adult survivors of sexual abuse (APA 1994). Childhood trauma, especially continued sexual abuse, is an important etiological factor in many cases of dissociative identity disorder, the most extreme type of dissociative reaction (APA 1994).

Finally, interpersonal difficulties and deficient social functioning often are observed among individuals who have been sexually abused (Briere 1992b; Cloitre et al. 1997). Such difficulties stem from the immediate cognitive and conditioned responses to victimization that extend into the longer term (e.g., distrust of others), as well as the

accommodation responses to ongoing abuse (e.g., passivity). Often, victims of childhood sexual abuse know the perpetrator, who might be a family member, clergy, or friend of the family. The abuse is thus a violation and betrayal of both personal and interpersonal (relationship) boundaries. Hence, it is not surprising that many children and adults with a history of childhood sexual abuse are found to be less socially competent, more aggressive, and more socially withdrawn than their nonabused peers (Cloitre et al. 1997; Mannarino et al. 1991; Mullen et al. 1994).

Among adults, interpersonal difficulties are manifested in difficulties in establishing and maintaining relationships (Finkelhor et al. 1989; Liem et al. 1996) and in achieving sexual intimacy (Browne and Finkelhor 1986; Mullen et al. 1994). Children who have been sexually abused are more likely to exhibit increased or precocious sexual behavior, such as kissing and inappropriate genital touching (Cosentino et al. 1995).

The consequences of childhood sexual abuse can be quite severe and harmful. However, it is important to note that an estimated 10 to 28 percent of persons with a history of it report no psychological distress (Briere and Elliott 1994; Kendall-Tackett et al. 1993). This raises the question of why some persons exhibit difficulties while others do not. Research addressing this question has found the following to be important mediators of individual reactions to childhood sexual abuse:

- *Age at onset of abuse.* Although further elucidation is needed, the available evidence suggests that postpubertal abuse is associated with greater trauma and more severe adverse sequelae than is prepubertal abuse (Beitchman et al. 1992; Browne and Finkelhor 1986; McClellan et al. 1996; Nash et al. 1993);
- *Gender of the victim.* One of the main findings in this area is that male victims appear to show greater disturbances of adult sexual functioning (Beitchman et al. 1992; Dube and Herbert 1988; Garnefski and Diekstra 1997);
- *Relationship to perpetrator.* Abuse involving a father or father figure (e.g., stepfather), which accounts for an estimated 25 percent of all cases (Sedlak and Broadhurst

1996), is associated with greater long-term harm (Browne and Finkelhor 1986; Gold 1986);

- *Duration and frequency of abuse.* In general, the greater the frequency and duration of abuse, the greater the impact on later psychological and social functioning (Nash et al. 1993);
- *Use of force.* The use of force or threat of force is associated with more negative outcomes (Kendall-Tackett et al. 1993);
- *Penetration or invasiveness.* Penetrative abuse is generally associated with greater long-term harm than are most other forms of abuse (Kendall-Tackett et al. 1993); and
- *Family characteristics and response to abuse disclosure.* Individuals who have been abused are more likely to originate from single-parent families, families with a high level of marital conflict, and families with pathology (e.g., parent is an alcoholic, violence between parents, maternal disbelief, and lack of support); all of which are associated with a poorer outcome and greater levels of distress (Beitchman et al. 1992; Draucker 1996; Green 1996; Romans et al. 1995).

### CROSS-CULTURAL ISSUES

This review has focused exclusively on studies done in the United States. This is important to keep in mind because the nature of abuse varies according to one's cultural belief system (Gough 1996). As a result, attempts to compare societies on the basis of their care of children or the extent of violence in family relations are fraught with problems (Gough 1996b; Levinson 1989). In addition, methodological factors, such as the questions asked and how childhood sexual abuse is defined and measured, hamper the ability to make direct comparisons among the rates across different countries. All that can be surmised, to date, is that it is not a phenomena just of the United States, but is an international problem (Finkelhor 1994b). Finkelhor's synthesis indicates that most countries have rates similar to those found in the United States and that females are abused at a greater rater than males.

Nonetheless, cross-cultural studies can shed new light on the origins and impact of sexual abuse (Leventhal 1998; Runyan 1998). Such investigations may enable us to understand better the relative importance of different factors that influence the occurrence of abuse and teach us about societies that have been successful at protecting children. Intercultural comparative investigations can also help us appreciate the range of sexual behaviors that are, or can be considered "normal," and thereby contribute to a better understanding of abnormality in childhood sexual behavior and adult behaviors toward children.

It is not surprising that definitions of abuse not only vary within a culture, but also between cultures. The meaning of "abuse," especially, depends upon ideas of individual rights and roles and responsibilities between people and groups within society (Gough 1996a). How a child is viewed will influence what is evaluated as abuse. Definitions of child abuse would be quite different, for example, in a feudal state in which children are considered to be their parents' possessions. Cultural expectations about sexual interactions among adolescents and between different age groups will also affect whether particular practices are defined as abuse (Abramson and Pinkerton 1995). Among the Sambia people of the highlands of Papua, New Guinea, for example, all young boys are expected to participate in ritualized fellatio with older boys as part of their initiation into manhood (Abramson and Pinkerton 1995). In essence, the younger boys are forced to submit; thus, this practice fulfills the two main criteria for childhood sexual abuse as defined above (coercive sex with a child or adolescent). Nevertheless, the Sambia consider ritualized fellatio to be critical for survival (they believe that semen is the source of manly strength and that it must be obtained through ingestion).

Definitions of abuse can also be expected to change over time to reflect societal changes. For instance, anecdotal evidence suggests that in Victorian England it was considered acceptable for a nurse or nanny to quiet a male infant by putting his penis in her mouth (Abramson and Pinkerton 1995). Today, this practice would clearly be considered childhood sexual abuse. The recent broadening of definitions of abuse has been accompanied by a greater sensitivity to signs of abuse that fit these changing definitions, and a greater willingness for professionals and others to intervene



into the private family lives of others, or beyond the walls of institutional life, and so to offer greater visibility of children's experiences (Gough 1996a).

### DIRECTIONS FOR FUTURE RESEARCH

Perusal of the literature in this domain suggests several avenues for future studies. First and foremost, large-scale, longitudinal investigations of child victims are needed that examine global functioning, abuse-specific functioning, and attributional and coping strategies along with abuse, child, family, and community factors. Such studies could provide information about both the initial effects of sexual abuse and the factors that influence adjustment at later developmental stages and into adulthood. Because not all victims of childhood sexual abuse develop adjustment problems, a better understanding of who is likely to experience such problems is needed.

Second, further research is needed regarding the efficacy of various approaches for the treatment of related problems (see Cosentino and Collins 1996; O'Donohue and Elliot 1993 for reviews). The utility or effectiveness of any particular treatment modality has yet to be demonstrated using a large-scale, randomized study in which treatment outcomes are measured using standardized instruments and untreated control groups. Consequently, treatment decisions often are made by clinicians without empirically tested guidelines.

Third, greater attention needs to be paid to methodological rigor in the conduct of childhood sexual abuse studies (Green 1993; Plunkett and Oates 1990; Trickett and McBride-Change 1995). In particular, investigators need to: (1) define it clearly and consistently; (2) use instruments with documented psychometric properties; (3) use control or comparison groups, when appropriate; and (4) employ large sample sizes whenever this is feasible. Studies of treatment modalities should also conduct multiple follow-up assessments to determine whether intervention effects are sustained over time. Finally, studies are needed that clearly disentangle the effects of sexual abuse from other forms of abuse or maltreatment.

### SUMMARY

Childhood sexual abuse is a significant and widespread problem in our society, no matter how it is

defined. Since it was first brought to the public's attention in the 1960s, a vast amount of research has been conducted in an effort to understand its impact on victims. From the various reviews on the topic, several conclusions can be drawn. First, despite the considerable heterogeneity among sexual abuse victims, as a group, sexually abused children and adolescents tend to display significantly higher levels of symptomatology than their nonabused, nonclinic-referred peers. Second, compared to other clinic-referred children, two problem areas appear to differentiate sexually abused children and adolescents: post-traumatic stress disorder symptomatology and sexuality problems. Third, the type and severity of sequelae experienced by victims depends on the specific characteristics of the abuse situation and the perpetrator. Fourth, research on adult survivors suggest that abuse-related problems tend to persist into adulthood. Indeed, childhood sexual abuse histories are common among several clinical populations, including patients initially diagnosed as depressed or as having a borderline personality disorder.

Taken together, these findings suggest that clinicians and health care providers should screen for sexual abuse among children, adolescents, and adults. For instance, questions about childhood sexual abuse could become part of routine intake procedures. Moreover, it is important that a wide range of service providers are sensitive to, and have staff trained to deal with, issues of childhood sexual abuse when it emerges. Service providers' policy and practice guidelines should explicitly acknowledge the prevalence and impact of it on women, men, and children.

Schools also need to have greater involvement in the prevention and diagnosis of childhood sexual abuse, especially since it generally occurs between the ages of six and twelve (Finkelhor 1994a; Sedlak and Broadhurst 1996). School-based education programs may be a useful vehicle for intervention and prevention. Teachers need to be educated and trained about their role in recognizing and reporting suspected cases.

Finally, further community awareness is needed to help prevent it from occurring. Greater community efforts toward providing treatment services for persons with a history of childhood sexual abuse are needed as are programs targeting

potential perpetrators. Society must work together to stamp out this abhorrence and to assist its survivors.

*Preparation of this article was supported in part by a training grant from the Maternal and Child Health Bureau, U.S. Department of Health and Human Services (MCJ 9040). The author thanks Steven D. Pinkerton, Ph.D. for his helpful comments on this manuscript.*

## REFERENCES

- Abramson, P.R., and S.D. Pinkerton 1995 *With pleasure: Thoughts on the Nature of Human Sexuality*. New York: Oxford University Press.
- Adler, N.A., and J.L. McCain 1994 "Prevention of Child Abuse. Issues for the Mental Health Practitioner." *Child and Adolescent Psychiatric Clinics of North America* 3(4):679-693.
- Alexander, P.C. 1992 "Application of Attachment Theory to the Study of Sexual Abuse." *Journal of Consulting and Clinical Psychology* 60:185-195.
- American Psychiatric Association 1994 *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. Washington, D.C.: Author.
- Arellano, C.M. 1996 "Child Maltreatment and Substance use: A Review of the Literature." *Substance Use and Misuse* 31(7):927-935.
- Beitchman, J.H., K.J. Zucker, J.E. Hood, G.A. DaCosta, and D. Akman 1991 "A Review of the Short-term Effects of Childhood Sexual Abuse." *Child Abuse and Neglect* 15:537-556.
- , and E. Cassavia 1992 "A review of the long-term effects of child sexual abuse." *Child Abuse and Neglect* 16:101-118.
- Berrick, J.D., and R.P. Barth 1992 "Child Sexual Abuse Prevention: Research Review and Recommendations." *Social Work Research and Abstracts* 28(4):6-15.
- Briere, J. 1988 "The Long-term Clinical Correlates of Childhood Sexual Victimization." *Annals of the New York Academy of Sciences* 528:327-334.
- 1992a "Methodological Issues in the Study of Sexual Abuse Effects." *Journal of Consulting and Clinical Psychology* 60:196-203.
- 1992b *Child Abuse Trauma: Theory and Treatment of the Lasting Effects*. Newbury Park, Calif.: Sage Publications.
- , and D.M. Elliott 1994 "Immediate and Long-term Impacts of Child Sexual Abuse." *Sexual Abuse of Children* 4(2):54-69.
- , and M. Runtz 1991 "The Long-term Effects of Sexual Abuse: A Review and Synthesis." *New Directions for Mental Health Services* 51:3-13.
- Browne, A., and D. Finkelhor 1986 "Impact of Child Sexual Abuse: A Review of the Research." *Psychological Bulletin* 99(1):66-77.
- Cahill, C., S.P. Llewelyn, and C. Pearson 1991a "Long-term effects of Sexual Abuse Which Occurred in Childhood: A Review." *British Journal of Clinical Psychology* 30(2):117-130.
- 1991b "Treatment of sexual abuse which occurred in childhood: A review." *British Journal of Clinical Psychology* 30(1):1-12.
- Chaffin, M., J.N. Wherry, and R. Dykman 1997 "School Age Children's Coping with Sexual Abuse: Abuse Stresses and Symptoms Associated with Four Coping Strategies." *Child Abuse and Neglect* 21(2):227-240.
- Cloitre, M., P. Scarvalone, and J.A. Difede 1997 "Posttraumatic Stress Disorder, Self- and Interpersonal Dysfunction Among Sexually Retraumatized Women." *Journal of Traumatic Stress* 10(3):437-452.
- Cole, P.M., and F.W. Putnam 1992 "Effect of Incest of Self and Social Functioning: A Developmental Psychopathology Perspective." *Journal of Consulting and Clinical Psychology* 60:174-184.
- Collings, S.J. 1995 "The Long-term Effects of Contact and Noncontact Forms of Child Sexual Abuse in a Sample of University Men." *Child Abuse and Neglect* 19(1):1-6.
- Connors, M.E., and W. Morse 1993 "Sexual Abuse and Eating Disorders: A review." *International Journal of Eating Disorders* 13(1):1-11.
- Cosentino, C.E., and M. Collins 1996 "Sexual Abuse of Children: Prevalence, Effects, and Treatment." *Annals of the New York Academy of Sciences* 789:45-65.
- Cosentino, C.E., H.F. Meyer-Bahlburg, J.L. Alpert, S.L. Weinberg, and R. Gaines 1995 "Sexual Behavior Problems and Psychopathology Symptoms in Sexually Abused Girls." *Journal of the American Academy of Child and Adolescent Psychiatry* 34(8):1033-1042.
- Douzinas, N., V. Fornari, B. Goodman, T. and Sitnick 1994 "Eating Disorders and Abuse." *Child and Adolescent Psychiatric Clinics of North America* 3(4):777-796.
- Draucker, C.B. 1996 "Family-of-Origin Variables and Adult Female Survivors of Childhood Sexual Abuse: A review of the Research." *Journal of Child Sexual Abuse* 5(4):35-63.
- Dube, R., and M. Hebert 1988 "Sexual Abuse of Children Under 12 Years of Age: A Review of 511 Cases." *Child Abuse and Neglect* 12(3):321-330.

- Dutton, M.A., K.J. Burghardt, S.G. Perrin, K.R. Chrestman, and P.M. Halle 1994 "Battered Women's Cognitive Schemata." *Journal of Traumatic Stress* 7(2):237-255.
- Faller, K.C. 1993 *Child Sexual Abuse: Intervention and Treatment Issues*. Washington, D.C.: U.S. Dept. of Health and Human Services, Administration for Children and Families, Administration on Children, Youth, and Families, National Center on Child Abuse and Neglect.
- Ferguson, A.G. 1997 "How Good is the Evidence Relating to the Frequency of Childhood Sexual Abuse and the Impact Such Abuse has on the Lives of Adult Survivors?" *Public Health* 111(6):387-391.
- Fergusson, D.M., L.J. Horwood, and M.T. Lynskey 1997 "Childhood Sexual Abuse, Adolescent Sexual Behaviors and Sexual Revictimization." *Child Abuse and Neglect* 21(8):789-803.
- Finkelhor, D. 1987 "The Sexual Abuse of Children: Current Research Reviewed." *Psychiatric Annals* 17:233-241.
- 1990 "Early and Long-term Effects of Child Sexual Abuse: An Update." *Professional Psychology* 21(5):325-330.
- 1994a "Current Information on the Scope and Nature of Child Sexual Abuse." *Future of Children* 4(2):31-53.
- 1994b "The International Epidemiology of Child Sexual Abuse." *Child Abuse and Neglect* 18(5):409-417.
- 1998 "Improving Research, Policy, and Practice to Understand Child Sexual Abuse." *Journal of the American Medical Association* 280(21):1864-1865.
- , and L. Berliner 1995 "Research on the Treatment of Sexually Abused Children: A Review and Recommendations." *Journal of the American Academy of Child and Adolescent Psychiatry* 34(11):1408-1423.
- Finkelhor, D., G. Hotaling, I.A. Lewis, and C. Smith 1990 "Sexual Abuse in a National Survey of Adult Men and Women: Prevalence, Characteristics, and Risk Factors." *Child Abuse and Neglect* 14(1):19-28.
- 1989 "Sexual Abuse and its Relationship to Later Sexual Satisfaction, Marital Status, Religion, and Attitudes." *Journal of Interpersonal Violence* 4:279-399.
- Friedrich, W.N. 1993 "Sexual Victimization and Sexual Behavior in Children: A Review of Recent Literature." *Child Abuse and Neglect* 17(1):59-66.
- Garnefski, N., R.F. and Diekstra 1997 "Child Sexual Abuse and Emotional and Behavioral Problems in Adolescence: Gender Differences." *Journal of the American Academy of Child and Adolescent Psychiatry* 36(3):323-329.
- Gibbons, J. 1996 "Services for Adults who Have Experienced Child Sexual Assault: Improving Agency Response." *Social Science and Medicine* 43(12):1755-1763.
- Glod, C.A. 1993 "Long-term Consequences of Childhood Physical and Sexual Abuse." *Archives of Psychiatric Nursing* 7(3):163-173.
- Gold, E.R. 1986 "Long-term Effects of Sexual Victimization in Childhood: An Attributional Approach." *Journal of Consulting and Clinical Psychology* 54:471-471.
- Gomes-Schwartz, B., J.M. Horowitz, and A.P. Cardarelli 1990 *Child Sexual Abuse: The Initial Effects*. Newbury Park, Calif.: Sage Publications.
- Gorey, K.M., and D.R. Leslie 1997 "The Prevalence of Child Sexual Abuse: Integrative Review Adjustment for Potential Response and Measurement Biases." *Child Abuse and Neglect* 21(4):391-398.
- Gough, D. 1996a "Defining the Problem: Comment." *Child Abuse and Neglect* 20(11):993-1002.
- 1996b "Child Abuse in Japan." *Child Psychology and Psychiatry Review* 1:12-18.
- Green, A.H. 1993 "Child Sexual Abuse: Immediate and long-term Effects and Intervention." *Journal of the American Academy of Child and Adolescent Psychiatry* 32(5):890-902.
- 1996 "Overview of Child Sexual Abuse." In S. J. Kaplan, ed., *Family Violence: A Clinical and Legal Guide*. Washington, D.C.: American Psychiatric Press, Inc.
- Holmes, W.C., and G.B. Slap 1998 "Sexual Abuse of Boys. Definition, Prevalence, Correlates, Sequelae, and Management." *Journal of the American Medical Association* 280(21):1855-1862.
- Janoff-Bulman, R. 1992 *Shattered Assumptions: Towards a new Psychology of Trauma*. New York: Free Press.
- Kelley, S.J. 1995 "Child Sexual Abuse: Initial Effects." *Annual Review of Nursing Research* 13:63-85.
- Kempe, C.H., F.N. Silverman, B.F. Steele, W. Droegemuller, and A.K. Silver 1962 "The Battered Child Syndrome." *Journal of the American Medical Association* 187:17-24.
- Kendall-Tackett, K.A., L.M. Williams, and D. Finkelhor 1993 "Impact of Sexual Abuse on Children: A Review and Synthesis of Recent Empirical Studies." *Psychological Bulletin* 113:164-180.
- Leitenberg, H., E. Greenwald, and S. Cado 1992 "A Retrospective Study of Long-term Methods of Coping with Having Been Sexually Abused During Childhood." *Child Abuse and Neglect* 16(3):399-407.
- Leventhal, J.M. 1998 "Epidemiology of Sexual Abuse of Children: Old Problems, New Directions." *Child Abuse and Neglect* 22(6):481-491.

- Levinson, D. 1989 *Family Violence in Cross Cultural Perspective*. Newbury Park, Calif.: Sage.
- Liem, J.H., J.G. O'Toole, and J.B. James 1996 "Themes of Power and Betrayal in Sexual Abuse Survivors' Characterizations of Interpersonal Relationships." *Journal of Traumatic Stress* 9(4):745-761.
- Lipovsky, J.A., B.E. Saunders, and S.M. Murphy 1989 "Depression, Anxiety, and Behavior Problems Among Victims of Father-child Sexual Assault and Nonabused Siblings." *Journal of Interpersonal Violence* 4:452-468.
- MacMillan, H.L., J.H. MacMillan, D.R. Offord, L. Griffith, and A. MacMillan 1994 "Primary Prevention of Child Sexual Abuse: A Critical Review: II." *Journal of Child Psychology and Psychiatry and Allied Disciplines* 35(5):857-876.
- Mancini, C., M. Van Ameringen, and H. MacMillan 1995 "Relationship of Childhood Sexual and Physical Abuse to Anxiety Disorders." *Journal of Nervous and Mental Disease* 183(5):309-314.
- Mannarino, A.P., J.A. Cohen, J.A. Smith, and S. Moore-Motily 1991 "Six- and Twelve-month Follow-up of Sexually Abused Girls." *Journal of Interpersonal Violence* 6:494-511.
- McClellan, J., C. McCurry, M. Ronnei, J. Adams, A. Eisner, and M. Storck 1996 "Age of Onset of Sexual Abuse: Relationship to Sexually Inappropriate Behaviors." *Journal of the American Academy of Child and Adolescent Psychiatry* 35(10):1375-1383.
- McLeer, S.V., E. Deblinger, D. Henry, and H. Orvaschel 1992 "Sexually Abused Children at High Risk for Post-traumatic Stress Disorder." *Journal of the American Academy of Child and Adolescent Psychiatry* 31:875-879.
- Messman, T.L., and P.J. Long 1996 "Child Sexual Abuse and its Relationship to Revictimization in Adult Women: A Review." *Clinical Psychology Review* 16(5):397-420.
- Mulder, R.T., A.L. Beautrais, P.R. Joyce, and D.M. Fergusson 1998 "Relationship Between Dissociation, Childhood Sexual Abuse, Childhood Physical Abuse, and Mental Illness in a General Population Sample." *American Journal of Psychiatry* 155(6):806-811.
- Mullen, P.E., J.L. Martin, J.C. Anderson, S.E. Romans, and G.P. Herbison 1994 "The Effect of Child Sexual Abuse on Social, Interpersonal and Sexual Function in Adult life." *British Journal of Psychiatry* 165(2):35-47.
- Murray, J.B. 1993 "Relationship of Childhood Sexual Abuse to Borderline Personality Disorder, Posttraumatic Stress Disorder, and Multiple Personality Disorder." *Journal of Psychology* 127(6):657-676.
- Nash, M.R., O.A. Zivney, and T. Hulsey 1993 "Characteristics of Sexual Abuse Associated with Greater Psychological Impairment Among Children." *Child Abuse and Neglect* 17(3):401-408.
- Oates, R.K., D. Forest, and A. Peacock 1985 "Self-esteem of Abused Children." *Child Abuse and Neglect* 9:159-163.
- O'Donohue, W., A. and Elliot 1993 "Treatment of the Sexually Abused Child: A review." *Journal of Clinical Child Psychology* 3:218-228.
- Olsen, J.L., and C.S. Widom 1993 "Prevention of Child Abuse and Neglect." *Applied and Preventive Psychology* 2(4):217-229.
- Plunkett, A., and R.K. Oates 1990 "Methodological Considerations in Research on Child Sexual Abuse." *Pediatric and Perinatal Epidemiology* 4(3):351-360.
- Polusny, M.A., and V.M. Follette 1995 "Long-term Correlates of Child Sexual Abuse: Theory and Review of the Empirical Literature." *Applied and Preventive Psychology* 4:143-166.
- Romans, S.E., J.L. Martin, J.C. Anderson, M.L. O'Shea, and P.E. Mullen 1995 "Factors that Mediate Between Child Sexual Abuse and Adult Psychological Outcome." *Psychological Medicine* 25(1):127-142.
- Rowan, A.B., and D.W. Foy 1993 "Post-traumatic Stress Disorder in Child Sexual Abuse Survivors: A Literature Review." *Journal of Traumatic Stress* 6(1):3-20.
- Runyan, D.K. 1998 "Prevalence, Risk, Sensitivity, and Specificity: A Commentary on the Epidemiology of Child Sexual Abuse and the Development of a Research Agenda." *Child Abuse and Neglect* 22(6):493-498.
- Saunders, B.E., L.A. Villepontoux, J.A. Lipovsky, D.G. Kilpatrick, and J.L. Veronen 1992 "Child Sexual Assault as a Risk Factor for Mental Disorders Among Women: A community Survey." *Journal of Interpersonal Violence* 7:189-204.
- Schwartz, M.F., and L. Cohn 1996 *Sexual Abuse and Eating Disorders: A Clinical Overview*. New York: Brunner/Mazel.
- Sedlak, A.J., and D.D. Broadhurst 1996 *Third national incidence study of child abuse and neglect*. Washington, D.C.: U.S. Dept. of Health and Human Services, Administration for Children and Families, Administration on Children, Youth, and Families, National Center on Child Abuse and Neglect.
- Shapiro, S., and G. Dominiak 1990 "Common Psychological Defenses Seen in the Treatment of Sexually Abused Adolescents." *American Journal of Psychotherapy* 44(1):68-74.
- Silk, K.R., J.T. Nigg, D. Esten, and N.E. Lohr 1997 "Severity of Childhood Sexual Abuse, Borderline Symptoms, and Familial Environment." In M.C.

- Zanarini, ed., *Role of Sexual Abuse in the Etiology of Borderline Personality Disorder* Washington, D.C.: American Psychiatric Press, Inc.
- Tharinger, D. 1990 "Impact of Child Sexual Abuse on Developing Sexuality." *Professional Psychology: Research and Practice* 21(5):331-337.
- Trickett, P.K., and C. McBride-Chang 1995 "The Developmental Impact of Different Forms of Child Abuse and Neglect." *Developmental Review* 15(3):311-337.
- Urquiza, A.J., and B.L. Goodlin-Jones 1994 "Child Sexual Abuse and Adult Revictimization with Women of Color." *Violence and Victims* 9(3):223-232.
- Van der Kolk, B.A., J.C. Perry, and J.L. Herman 1991 "Childhood Origins of Self-destructive Behavior." *American Journal of Psychiatry* 148(12):1665-1671.
- Van der Kolk, B.A., A. Hostenler, N. Herron, and R.E. Fisler 1994 "Trauma and the Development of Borderline Personality Disorder." *Psychiatric Clinics of North America* 17(4):715-730.
- Watkins, B., and A. Bentovim 1992 "The Sexual Abuse of Male Children and Adolescents: A Review of Current Research." *Journal of Child Psychology and Psychiatry and Allied Disciplines* 33(1):197-248.
- Widom, C.S., and J.B. Kuhns 1996 "Childhood Victimization and Subsequent Risk for Promiscuity, Prostitution, and Teenage Pregnancy: A Prospective Study." *American Journal of Public Health* 86(11):1607-1612.
- Wilsnack, S.C., N.D. Vogeltanz, A.D. Klassen, and T.R. Harris 1997 "Childhood Sexual Abuse and Women's Substance Abuse: National survey findings." *Journal of Studies on Alcohol* 58(3):264-271.
- Wolfe, D.A., N.D. Reppucci, and S. Hart 1995 "Child Abuse Prevention: Knowledge and Priorities." *Journal of Clinical Child Psychology* 24(Suppl):5-22.
- Wolfe, V.V., and J.A. Birt 1995 "The Psychological Sequelae of Child Sexual Abuse." *Advances in Clinical Child Psychology* 17:233-263.
- Wonderlich, S.A., T.D. Brewerton, Z. Jovic, B.S. Dansky, and D.W. Abbott 1997 "Relationship of Childhood Sexual Abuse and Eating Disorders." *Journal of the American Academy of Child and Adolescent Psychiatry* 36(8):1107-1115.
- Wurr, C.J., and I.M. Partridge 1996 "The Prevalence of a History of Childhood Sexual Abuse in an Acute Adult Inpatient Population." *Child Abuse and Neglect* 20(9):867-872.
- Wyatt, G.E., D. Guthrie, and C.M. Notgrass 1992 "Differential Effects of Women's Child Sexual Abuse and Subsequent Sexual Revictimization." *Journal of Consulting and Clinical Psychology* 60(2):167-173.
- Yama, M.F., S.L. Tovey, and B.S. Fogas 1993 "Childhood Family Environment and Sexual Abuse as Predictors of Anxiety and Depression in Adult Women." *American Journal of Orthopsychiatry* 63(1):136-141.

HEATHER CECIL

## CHILDREARING

See Parental Roles; Socialization.

## CHINA STUDIES

The sociological study of China has a complex history. It could be argued that Confucian thought embodied a native tradition of sociological thinking about such things as the family, bureaucracy, and deviant behavior. However, modern Chinese sociology initially had foreign origins and inspiration. The appearance of the field in China might be dated from the translation of parts of Herbert Spencer's *The Study of Sociology* into Chinese in 1897. The earliest sociology courses and departments in China were established in private, missionary colleges, and Western sociologists such as D. H. Kulp, J. S. Burgess, and Sidney Gamble played central roles in initiating sociology courses and research programs within China (see Wong 1979).

The Chinese Sociological Association was established in 1930, and in the following two decades a process of Sinification progressed. Chinese sociologists, trained both at home and in the West, emerged. Sociology courses and departments began to proliferate in government-run, public colleges. Increasingly, texts using material from Chinese towns and villages displaced ones that focused on Chicago gangs and the assimilation of Polish immigrants into America. As this process of setting down domestic roots continued, Chinese sociology developed some distinctive contours. Inspiration for the field came as much from British social anthropology as from sociology. The crisis atmosphere of the 1930s and 1940s fostered a strong emphasis on applied, problem-solving research. As a result of these developments, Chinese sociology in those years came to include what in the United States might be considered social anthropology and social work. Within these broad contours, Chinese sociologists/anthropologists produced some of the finest early ethnographies of

“nonprimitive” communities—Chinese peasant villages (see Fei 1939; Yang 1944; Lin 1948).

Although many of the earliest leaders in establishing Chinese sociology were foreigners, by the time of the Chinese communist victory in 1949, its leading practitioners were Chinese. There were only a small number of Western sociologists who concentrated their research on Chinese society (see Lang 1946; Levy 1949), and their methods, problems, and data were not distinctive from those being employed by their Chinese counterparts. Originally a foreign transplant, Chinese sociology had become a thriving enterprise with increasingly strong domestic roots.

When the Chinese Communist Party (CCP) swept to national power in 1949, some Chinese sociologists left the country but most remained. (The development of sociology in the Republic of China on Taiwan after 1949 will not be dealt with here.) Initially, those who remained were optimistic that their skills would be useful to the new government. Experience in community fieldwork and an orientation toward studying social problems seemed to make Chinese sociologists natural allies of those constructing a planned social order. These hopes were dashed in 1952 when the CCP abolished the field of sociology. That decision was motivated by the CCP’s desire to follow the Soviet model. Joseph Stalin had earlier denounced sociology as a “bourgeois pseudo-science” and banned the field from Soviet academe. More to the point, there was no room in the People’s Republic of China (PRC) for two rival sciences of society, Marxism-Leninism-Maoism and sociology. The CCP argued that it had developed its own methods of “social investigations” during the revolutionary process, with Mao Zedong playing a leading role in this development (see Mao [1927] 1971; Thompson 1990). This approach stressed grass roots investigations designed to further official revolutionary or economic goals of the CCP rather than any sort of attempted value-free search for objective truth. Chinese sociologists, trained in a different tradition and able to use professional claims to raise questions about CCP policies, were seen as a threat to the ideological hegemony of the new regime.

The ban on sociology in China was to last twenty-seven years, until 1979. One major attempt was made by Chinese sociologists in 1956 and

1957 to get the ban lifted. In the more relaxed political atmosphere prevailing at that time, ushered in by the CCP’s slogan “let 100 flowers bloom and 100 schools of thought contend” and encouraged by the revival of sociology in the Soviet Union, leading figures such as Fei Xiaotong, Pan Guangdan, Wu Jingchao, and Chen Da wrote articles and made speeches arguing that sociology could be useful to socialist China in the study of social change and social problems—precisely the same rationale used prior to 1952. When the political atmosphere turned harsh again in the latter part of 1957, with the launching of the “anti-rightist campaign,” Fei and many of the other leaders of this revival effort were branded “rightist elements” and disappeared from view.

From 1952 to 1979 some of those trained in sociology were assigned to work that had some links to their abolished field. Throughout this period ethnology remained an established discipline, devoted to the study of China’s various ethnic minorities. When Fei Xiaotong emerged from his political purgatory in 1972, it was as a leading figure in ethnology. But Fei and others could not do research on the 94 percent of the population that is Han Chinese, and they also had to renounce publicly all their former work and ideas as “bourgeois.” Psychology survived better than sociology by becoming defined as a natural, rather than a social, science. Some areas of social psychology (for example, educational psychology) remained at least somewhat active through the 1960s and 1970s (see Chin and Chin 1969). Other components of sociology disappeared in 1952 but reappeared prior to 1979. Demography was established as a separate field in the early 1970s and remains a separate discipline in Chinese academe today. With these partial exceptions, however, Chinese sociology ceased to exist for these twenty-seven years.

While sociology was banned in China, the sociological study of that society developed gradually in the West. Starting in the late 1950s, American foundations and government agencies, and to some extent their European and Japanese counterparts, provided funds for the development of the study of contemporary China—for fellowships, research centers, journals, language training facilities, and other basic infrastructure for the field. Within this developmental effort, sociology was

targeted as an underdeveloped discipline (compared to, say, history, literature, and political science) deserving of special priority. Stimulated by this developmental effort, during the 1960s and 1970s the number of Western sociologists active in the study of China increased from a handful to perhaps two dozen.

Prior to the early 1970s, with a few minor exceptions (e.g., Geddes 1963), it was impossible for foreign sociologists to travel to China, much less to do research there. Even when such travel became possible during the 1970s, it was initially restricted to very superficial “academic tourism.” In this highly constrained situation, foreign sociologists studying China developed distinctive research methods.

Much research was based on detailed culling of Chinese mass media reports. Special translation series of Chinese newspapers and magazines and of monitored radio broadcasts developed in the 1950s and became the mainstays of such research. The Chinese media provided fairly clear statements about official policy and social change campaigns. By conveying details about local “positive and negative models” in the implementation of official goals, the media also provided some clues about the difficulties and resistance being encountered in introducing social change. In a few instances the media even yielded apparently “hard” statistical data about social trends. A minor industry developed around scanning local press and radio reports for population totals, and these were pieced together to yield national population estimates (typically plus or minus 100 million).

Some researchers developed elaborate schemes to code and subject to content analysis such media reports (see, for example, Andors 1977; Cell 1977). However, this reliance on media reports to study Chinese social trends had severe limits. Such reports revealed much more about official goals than they did about social reality, and the researcher might erroneously conclude that an effort to produce major changes was actually producing them. The sorts of unanticipated consequences that sociologists love to uncover were for the most part invisible in the highly didactic Chinese media. It was also very difficult to use media reports to analyze variations across individuals and communities, the staple of sociological analysis elsewhere. Studies based on media reports tended to convey

an unrealistic impression of uniformity, thus reinforcing the prevailing stereotype of China as a totalitarian society.

Partly in reaction to these limitations, a second primary method of research on China from the outside developed—the refugee interview. Scholars in sociology and related fields spent extended periods of time in Hong Kong interviewing individuals who had left the PRC for the British colony. Typically, these individuals were used not as respondents in a survey but as ethnographic informants “at a distance” about the corners of the Chinese social world from which they had come. Although the individuals interviewed obviously were not typical of the population remaining in the PRC, for some topics these refugees provided a rich and textured picture of social reality that contrasted with the unidimensional view conveyed by Chinese mass media. Elaborate techniques were developed by researchers to cope with the problems of atypicality and potential bias among refugee informants (see Whyte 1983).

By combining these methods with brief visits to China, sociologists and others produced a large number of useful studies of contemporary mainland society (Whyte, Vogel, and Parish 1977). Even though the arcane methods they used often seemed to nonspecialists akin to Chinese tea-leaf reading, the best of these studies have stood the test of time and have been confirmed by new research made possible by changes in China since 1979. However, the exclusion from meaningful field research in China did have unfortunate effects on the intellectual agendas of researchers. Much energy was focused on examining the “Maoist model” of development and whether or not China was succeeding in becoming a revolutionary new type of social order. Since the reality of Chinese social organizations in the Maoist period was quite different from the slogans and ideals upon which such constructions were based, this effort appears to have been a waste of scarce intellectual resources. Both this unusual intellectual agenda and the distinctive research methods sociologists studying China were forced to use reinforced the isolation of these researchers from “mainstream” work in sociology.

The year 1979 was significant in two ways for the sociological study of China. First, as already noted, that was the year in which China’s leaders

gave approval for the “rehabilitation” of sociology. After the death of Mao Zedong in 1976, and given the authority crisis that followed in the wake of the Cultural Revolution, the post-Mao leadership reopened many issues. China faced a number of serious social problems that had been ignored previously, and in this situation long-standing arguments about the utility of sociology for the study of social problems could be raised once again.

The second major change that occurred in 1979 was the “normalization” of diplomatic relations between the United States and China. That development led to approval for American and other Western researchers to conduct extended field research in China for the first time since 1949. The days of frustrating confinement to research from outside and academic tourism were over. It also became possible for U.S. sociologists to meet Chinese counterparts and to begin to discuss common research interests and plans for collaboration. Westerners began coming to China to teach sociology courses, Chinese began to be sent abroad for training in sociology, and collaborative conferences and publications were initiated. Leading figures in the West, including Peter Blau, Alex Inkeles, Nan Lin, W. J. Goode, Hubert Blalock, and Jeffrey Alexander, lectured in China and recruited Chinese counterparts for collaborative research and for training in their home institutions. All of the kinds of intellectual interchange between Chinese sociologists and their foreign counterparts that had been impossible for a generation were resumed.

The dangerous political aura surrounding sociology in the Mao era and the lesson of the abortive 1957 attempt to revive the field might have been expected to make it difficult to attract talented people to Chinese sociology after 1979. However, the field revived very rapidly and showed surprising intellectual vigor. A number of surviving pre-1949 sociologists reappeared to play leading roles in the revival, considerable numbers of middle-aged individuals trained in related fields were recruited (or ordered) to become sociologists, and many young people expressed eagerness to be selected for training in the field. The Chinese Sociological Association was formally revived, with Fei Xiaotong as its head. (Fei “unrenounced” his pre-1949 works, weakened his ties with ethnology, and resumed advocacy of the program of Chinese

rural development he had championed a generation earlier.) A number of departments of sociology were established in leading universities, and new sociological journals and a large number of monographs and textbooks began to appear. Institutes of sociology were established within the Chinese Academy of Social Science in Peking as well as in many provincial and city academies. Apparently, the novelty of the field, the chance to make a place for oneself, and perhaps even the chance to study abroad without having to face competition from layers of senior people more than compensated for the checkered political past of sociology. In any case, the speed and vigor with which the field revived surprised the skeptics (see Whyte and Pasternak 1980).

The intellectual focus of the revived sociology was fairly broad. Social problems and the sociology of development were two major foci. Within these and other realms, two sets of issues loomed large. One was the study of the after-effects of the Cultural Revolution and the radical policies pursued under Mao Zedong in the period 1966 to 1976. The other major issue concerned the impact of the opening to the outside world, decollectivization of agriculture, and the market-oriented reforms launched in 1978. A large number of studies began on these and other issues, and by the end of the 1980s the problem facing foreign sociologists studying China was not so much the scarcity of data on their object of study but its abundance. The publication of large amounts of statistical material, the availability of census and survey data dealing with demographic and other matters, and the mushrooming of social science publications threatened to overwhelm those who tried to keep track of Chinese social trends.

In pursuing their new intellectual agendas, Chinese sociologists easily found common ground with Western specialists. The latter were no longer so entranced with the Maoist model and shared a fascination with the social impact of the post-Mao changes. Many Western specialists readily adapted to the chance to conduct research in China and to collaborate with Chinese colleagues. Hong Kong did not die out as a research site and remained vital for some topics, but work from outside China was increasingly seen as supplementary to research conducted within the country. (After 1997 and Hong Kong’s reversion to China this was still the



case, as the special autonomy enjoyed by the former British colony produced a freer research atmosphere than in the rest of the country.) China specialists now had to share the stage with many sociologists who did not have area studies training. With data more readily available and collaboration with Chinese counterparts possible, area studies training and familiarity with the arcane techniques developed in the 1960s and 1970s were no longer so important. Many non-China specialists found the opportunity to study the world's most populous society increasingly attractive. These changes have been beneficial for China specialists within sociology. Given their small numbers and the complexity of Chinese society, there is more than enough "turf" to be shared with sociologists not trained in China studies. And these changes have reduced the intellectual isolation of the specialists, from both colleagues in China and those in their own departments.

As a result of these changes, Chinese sociologists ended their isolation from world sociology, and Western sociologists conducting research on China escaped from a highly constrained environment and began to produce work that was less idiosyncratic. It should be noted, however, that sociological research in China is not without limits, either for foreign researchers or for Chinese. Even in the most liberal phases of the initial reform decade from 1979 to 1989, some topics remained taboo, some locales were off limits, and bureaucratic obstruction and interference were constant facts of life. Nonetheless, even for foreigners it became possible to conduct extended ethnographic studies, carry out probability sample surveys, and gain access for secondary analysis to the raw data from census and survey studies conducted by Chinese, experiences that those who were studying the Soviet Union before its collapse could only envy. The result was a new spurt of research activity and publication that enriched our knowledge of Chinese social life immensely. (For a review of Western research during the decade after 1979, see Walder 1989. For an overview of developments within Chinese sociology during that period, see Lu 1989.)

However, the political fallout from the 1989 Tiananmen massacre raised new problems for the sociological study of China. At least some conservatives within the Chinese leadership used the

crackdown following the massacre to attack sociology and other newly revived social sciences in terms reminiscent of 1952 and 1957. The field was once again accused of being an ideological Trojan horse designed to spread doubts about socialist orthodoxy. Some sociology departments were not allowed to enroll new students, while others had their enrollment targets cut. While these restrictions eased after a year or two, new and serious threats to collaborative research projects continued through the 1990s. Edicts designed to restrict and maintain strict political control over survey projects involving foreign collaboration were passed in secret in 1990 and 1995, resulting in the confiscation of survey data already collected in one instance and delays and torpedoing of projects in others. (The data confiscated after the 1990 secret edict were eventually released to the researchers involved, and other stalled projects resumed, but only after several years of tense negotiations.) In 1998 an international collaborative survey project in social gerontology was publicly attacked within China for allegedly endangering the health of China's elderly and having a concealed aim of enabling foreigners to profit from the "genetic secrets" of the Chinese people. Only after a vigorous public defense from both foreign and Chinese researchers was the disputed project saved. These developments mean that researchers who embark on sociological research within China on even quite nonpolitical topics face some risk that they will not be able to complete their projects due to political interference.

Despite the unpredictable political environment, Chinese and foreign sociologists pressed ahead with a wide-ranging variety of research activities during the 1990s. The success of this effort is a testament to the increasingly robust ties linking Chinese and foreign sociologists. During the 1980s large numbers of aspiring young Chinese sociologists were sent abroad for doctoral training, primarily to the United States. By the latter part of that decade and into the 1990s, a majority of the Western sociology doctoral theses dealing with China were being produced by students from China. After the 1989 Tiananmen massacre most of these newly minted specialists on Chinese society did not return to China, but instead embarked on research careers in the West. However, mindful of the desirability of leaving the door open for their eventual return, China did not treat this

decision as disloyalty, but instead in most instances encouraged and invited these foreign-based Chinese sociologists to visit China to assist in the development of the discipline there and to undertake collaborative research projects. These young, foreign-trained sociologists are able to come and go and to navigate the complex political and bureaucratic terrain surrounding sociological research within China much more easily than their non-Chinese counterparts. Furthermore, the relatively smaller number of Chinese sociologists trained abroad who have returned to the PRC usually have found their new skills and contacts valued and rewarded, rather than treated with envy and suspicion. Many of these individuals are moving up rapidly into positions of responsibility and leadership within Chinese research institutes and universities, where they are generally able to keep research projects moving ahead despite bureaucratic obstacles. As a result of these trends, an increasing share of the new sociological research initiatives and publications that are having an impact on China studies in the West involve sociologists from China, whether currently based abroad or within China.

The changing sources of funding for research on Chinese society also are having an impact. Whereas in the past most research by Chinese social scientists was funded by the Chinese government and was responsive to state priorities and concerns, in the reform era Chinese sociologists and others in China face increasing constraints on government funding of research. New competitive peer review procedures are being introduced to determine who gets scarce research funds, and Chinese authorities encourage research institutes and researchers to find ways to raise or obtain funds on their own to support their research. In this altered environment a variety of Western and international sources of funding—from United Nations agencies, the World Bank, the Ford Foundation, foreign nongovernmental organizations, etc.—have become important in financing research that Chinese sociologists otherwise could not carry out. This changed funding picture thus reinforces other trends that are strengthening the links between sociology within China and abroad.

While the topics pursued within the sociological study of China in the 1990s were varied, the dominant focus was still on the process of market reforms. Some of the impetus for this focus came

from a comparison with developments in the former Soviet Union and its satellites. Sociologists have joined others in trying to understand why to date the dismantling of central planning and the introduction of market reforms have generally been much more successful in China than in Eastern Europe (see, for example, Rozman 1992; Walder 1996a). The bulk of this research, however, looks not at the causes, but at the social consequences of the reforms, and here the concerns are varied and wide-ranging. For example, spirited debates have arisen about whether or in what respects the reforms are contributing to increased inequality within China, and about whether prereform political elites are largely monopolizing the gains produced by the reforms or are being displaced increasingly by new elites (see, for example, Nee 1996; Xie and Hannum 1996; Walder 1996b; Bian and Logan 1996; Zhou, Tuma, and Moen 1996; Buckley 1997). Sociologists and others are also examining the new entrepreneurial class arising within China as a result of market reforms to determine what sorts of social origins they have, and researchers on this topic debate whether these entrepreneurs are a force for political change or are, instead, co-opted and controlled as a result of their dependence upon the state (see, for example, Unger 1996; Nevitt 1999). Research on the latter theme is part of a larger exploration of whether China's reforms are leading to the development of institutions of a "civil society" that might eventually lead to the democratization of China's political system (see Whyte 1992; White 1993). Within research on Chinese stratification there is also a growing body of work considering whether, on balance, the reforms have been harmful or beneficial in terms of the quest for gender equality (see Bauer et al. 1992; Croll 1994; Lee 1998; Entwisle and Henderson forthcoming). Another major focus of research concerns changes in Chinese family life, with the investigation of topics such as rising divorce and premarital sex rates and whether such trends indicate that Chinese family patterns are increasingly "converging" toward the patterns found in other societies (e.g. Xu and Whyte 1990; Davis and Harrell 1993; Logan, Bian, and Bian forthcoming). Still another major focus is on tracing the impact of particular reform changes—for example, the relaxation of migration restrictions that have allowed tens of millions of rural migrants to flood into China's cities, the attempt to revive labor markets as a way of

distributing jobs, and the layoffs and unemployment being generated by reforms of China's inefficient state-owned enterprises (e.g. Bian 1994; Guthrie 1997; Matthews 1998) Solinger 1995; Sensenbrenner 1996.

Another primary research focus of China sociologists, in addition to the causes and consequences of the post-1978 economic reforms, involves China's dramatic fertility decline and its social consequences. The transformation of China from a society with moderately high fertility (total fertility rates of 5+) to extraordinarily low rates (total fertility rates under 2) has occurred in roughly the same time period as China's economic reforms (from the 1970s through the 1990s), but these two fundamental changes have been produced by contradictory processes. Whereas in economic affairs the state has increasingly withdrawn from bureaucratic management and has allowed markets to govern distribution, in regard to population the state has shifted away from indirect management and persuasion in favor of direct bureaucratic mandating of fertility quotas. The high national priority of this fertility reduction program has complex implications, in terms of sociological research. High priority, combined with recurring foreign criticism of China's human rights record in the realm of family planning (see Whyte 1998), produce political sensitivities that often obstruct and constrain research on this topic. However, that same priority has made research on topics related to fertility so important that foreign as well as Chinese demographers and others have been recruited to help plan and carry out a large number of high-quality surveys of China's demographic dynamics. These data have also in most instances been made available for examination by foreign as well as Chinese researchers. On balance the result has been a boom in research on China's population and a much better understanding of current trends in fertility, mortality, and other demographic domains than existed prior to the 1980s (see, for example, Banister 1987; Coale 1984; Feeney and Yuan 1994; Riley and Riley 1997).

In sum, recurring political turbulence in China has in some cases obstructed and delayed sociological research, but both the field and its practitioners have continued to forge ahead despite the uncertainties involved. The rapid and complex social changes that have occurred in China in the

1980s and 1990s have provided many new research questions for sociologists to explore, and a maturing network of Chinese sociologists and foreign China specialists (and nonspecialist sociologists interested in studying the world's most populous and most rapidly developing society) have taken advantage of new research opportunities. As a result of these changes the sociological study of China at the end of the 1990s was much better established than it was two decades earlier, with a wide variety of social processes and trends better studied and more fully understood.

## REFERENCES

- Andors, Stephen 1977 *China's Industrial Revolution: Politics, Planning, and Management, 1949 to the Present*. New York: Pantheon.
- Banister, Judith 1987 *China's Changing Population*. Stanford, Calif.: Stanford University Press.
- Bauer, J., Wang Feng, N. Riley, and Zhao Xiaohua 1992 "Gender Inequality in Urban China." *Modern China* 18:333-369.
- Bian, Yanjie 1994 *Work and Inequality in Urban China*. Albany, N.Y.: SUNY Press.
- , and John Logan 1996 "Market Transition and the Persistence of Power: The Changing Stratification System in Urban China." *American Sociological Review* 61:739-758.
- Buckley, Christopher 1997 Social mobility and stratification in urban China from Maoism through reform. Ph.D. diss. Australian National University.
- Cell, Charles P. 1977 *Revolution at Work*. New York: Academic Press.
- Chin, R., and A. Chin 1969 *Psychological Research in Communist China 1949-1966*. Cambridge, Mass.: MIT Press.
- Coale, Ansley 1984 *Rapid Population Change in China, 1952-1982*. Washington, D.C.: National Academy Press.
- Croll, Elisabeth 1994 *From Heaven to Earth: Images and Experiences of Development in China*. London: Routledge.
- Davis, Deborah, and Stevan Harrell (eds.) 1993 *Chinese Families in the Post-Mao Era*. Berkeley: University of California Press.
- Entwisle, Barbara, and Gail Henderson (eds.) forthcoming *Redrawing Boundaries: Work, Household, and Gender in China*. Berkeley: University of California Press.
- Feeney, Griffith, and Yuan Jianhua 1994 "Below Replacement Fertility in China? A Close Look at Recent Evidence." *Population Studies* 48:381-394.

- Fei Xiaotong [Fei Hsiao-t'ung] 1939 *Peasant Life in China*. London: G. Routledge.
- Geddes, W. R. 1963 *Peasant Life in Communist China*. Ithaca, N.Y.: Society for Applied Anthropology.
- Guthrie, Douglas 1997 Strategy and structure in Chinese firms: Organizational action and institutional change in Shanghai. Ph.D. diss. University of California, Berkeley.
- Lang, Olga 1946 *Chinese Family and Society*. New Haven, Conn.: Yale University Press.
- Lee, Ching Kuan 1998 *Gender and the South China Miracle*. Berkeley: University of California Press.
- Levy, Marion J., Jr. 1949 *The Family Revolution in Modern China*. Cambridge, Mass.: Harvard University Press.
- Lin Yaohua [Lin Yüeh-hwa] 1948 *The Golden Wing: A Sociological Study of Chinese Familism*. New York: Oxford University Press.
- Logan, John, Bian Fuqin, and Bian Yanjie forthcoming "Tradition and Change in the Urban Chinese Family: The Case of Living Arrangements." *Social Forces*.
- Lu Xueyi 1989 *China Yearbook of Sociology, 1979–1989* (in Chinese). Peking: Chinese Encyclopedia Press.
- Mao Zedong (1927) 1971 "Report on an Investigation of the Peasant Movement in Hunan." In Mao Tsetung, *Selected Readings from the Works of Mao Tsetung*. Peking: Foreign Languages Press.
- Matthews, Rebecca 1998 Where do labor markets come from? The emergence of urban labor markets in the People's Republic of China. Ph.D. diss. Cornell University, Ithaca, N.Y.
- Nee, Victor 1996 "The Emergence of a Market Society: Changing Mechanisms of Stratification in China." *American Journal of Sociology* 101:908–949.
- Parris, Kristen 1999 "Entrepreneurs and Citizenship in China." *Problems of Post-Communism* 46:43–61.
- Riley, Nancy, and Robert Riley 1997 *China's Population: A Review of the Literature*. Liege: International Union for the Scientific Study of Population.
- Rozman, Gilbert (ed.) 1992 *Dismantling Communism: Common Causes and Regional Variations*. Washington, D.C.: Woodrow Wilson Center Press.
- Sensenbrenner, Julia 1996 Rust in the iron rice bowl: Labor reforms in Shanghai state enterprises, 1992–93. Ph.D. diss. Johns Hopkins University.
- Solinger, Dorothy 1995 "China's Urban Transients in the Transition from Socialism and the Collapse of the Urban Public Goods Regime." *Comparative Politics* 27:127–147.
- Thompson, Roger 1990 *Mao Zedong: Report from Xunwu*. Stanford, Calif.: Stanford University Press.
- Unger, Jonathan 1996 "'Bridges': Private Business, the Chinese Government, and the Rise of New Associations." *The China Quarterly* 147:795–819.
- Walder, Andrew (ed.) 1996a *China's Transitional Economy*. London: Oxford University Press.
- 1996b "Markets and Inequality in Transitional Economies: Toward Testable Theories." *American Journal of Sociology* 101:1060–1073.
- 1989 "Social Change in Post-Revolution China." *Annual Review of Sociology* 15:405–424.
- White, Gordon 1993 "Prospects for Civil Society in China: A Case Study of Xiaoshan City." *Australian Journal of Chinese Affairs* 29:63–87.
- Whyte, Martin King 1998 "Human Rights Trends and Coercive Family Planning in the PRC." *Issues and Studies* 34:1–29.
- 1992 "Urban China: A Civil Society in the Making?" In A. Rosenbaum, ed., *State and Society in China*. Boulder, Colo.: Westview Press.
- 1983 "On Studying China at a Distance." In Anne Thurston and Burton Pasternak, eds., *The Social Sciences and Fieldwork in China*. Boulder, Colo.: Westview.
- , and Burton Pasternak 1980 "Sociology and Anthropology." In Anne Thurston and Jason Parker, eds., *Humanistic and Social Science Research in China*. New York: Social Science Research Council.
- Whyte, Martin K., Ezra F. Vogel, and William L. Parish, Jr. 1977 "Social Structure of World Regions: Mainland China." *Annual Review of Sociology* 3:179–207.
- Wong, Siu-lun 1979 *Sociology and Socialism in Contemporary China*. London: Routledge and Kegan Paul.
- Xie, Yu, and Emily Hannum 1996 "Regional Variation in Earnings Inequality in Reform-Era Urban China." *American Journal of Sociology* 101: 950–992.
- Xu, Xiaohe, and Martin King Whyte 1990 "Love Matches and Arranged Marriages: A Chinese Replication." *Journal of Marriage and the Family* 52:709–722.
- Yang, C. K. 1944 *A North China Local Market Economy*. New York: Institute of Pacific Relations.
- Zhou, Xueguang, Nancy Tuma, and Phyllis Moen 1996 "Stratification Dynamics under State Socialism: The Case of Urban China, 1949–1993." *Social Forces* 74:759–796.

MARTIN KING WHYTE

## CITIES

A *city* is a relatively large, dense, permanent, heterogeneous, and politically autonomous settlement whose population engages in a range of nonagricultural occupations. Definitions of cities and their associated phenomena vary by time and place, and by population size, area, and function (Shryock, Siegel, and associates 1976, pp. 85–104). The city is often defined in terms of administrative area, which may be larger than, smaller than, or equal to the area of relatively dense settlement that comprises what is otherwise known as the city proper. The suburb is a less dense but permanent settlement that is located outside the city proper and contains populations that usually have social and economic ties to the city.

Definitions of urban vary by nation; in the United States the term refers to populations of 2,500 or more living in towns or cities and to populations living in urbanized areas, including suburbs. In other nations, the lower limits for settlements defined as urban vary between 200 and 50,000 persons. United Nations definitions of urban areas emphasize a population of 20,000 or more, and cities a population of 100,000 or more. *Urbanization* refers to the economic and social changes that accompany population concentration in urban areas and the growth of cities and their surrounding areas.

Cities reflect other areas with which they are linked and the civilizations of which they are a part. Cities are centers of markets, governments, religion, and culture (Weber 1958, pp. 65–89). A *community* is a population sharing a physical environment and leading a common and interdependent life. The size, density, and heterogeneity of the urban community have been described as leading to “urbanism as a way of life,” which includes organizational, attitudinal, and ecological components different from those of rural areas (Wirth 1938).

### THE CITY IN HISTORY

Town and city development has been described as tied to a technological revolution in agriculture that increased food production, thereby freeing agriculturalists to engage in nonagricultural occupations. This resulted in an evolution to urban

living and eventually to industrial production (Childe 1950). A second view is that some towns and cities first developed as trade centers, and were then nourished by agricultural activity in their hinterlands (Jacobs 1970).

Increasing complexity of social organization, environmental adaptation, and technology led to the emergence of cities (Child 1950; Duncan 1964). Excluding preagricultural settlements, towns and then cities were first established in the fourth to third millennium B.C. in ancient Mesopotamia within the Tigris and Euphrates river valleys, in the Harappa civilization in the valleys of the Indus River and its tributaries, and in the Egyptian Old Kingdom in the lower Nile valley. Other centers appeared in the Huang Ho basin on the east coast of China and the Peruvian Andes in the second to first millennium B.C., and in Mesoamerica in the first millennium B.C. (Phillips 1997, pp. 82–85).

Small agricultural surpluses and limits on transportation meant that the first towns were small and few in number, and contained only a small proportion of the populations of their regions. Economic activities of the earliest towns were tied largely to their surrounding areas. After the rise of towns in the Middle East, trading centers appeared on the shores and islands of the Mediterranean; some, such as Athens, became city-states. After developing more effective communication and social organization, some Western city-states expanded and acquired empires, such as those of Alexander the Great and of Rome. Following the decline of Rome, complex city life continued in the East and in the West in the Byzantine and Muslim empires, while the population of Europe declined and reverted, for the most part, to subsistence agriculture and organized into small territories held together by the Catholic Church (Hawley 1981, pp. 1–35).

Sjoberg (1960) has described preindustrial cities as feudal in nature and sharing social, ecological, economic, family, class, political, religious, and educational characteristics different from those in modern industrial cities. In the former, the city center, with its government and religious and economic activities, dominated the remainder of the city and was the locale of the upper social classes. Homogeneous residential areas were found throughout the city, but nonresidential activities

were not confined to distinct neighborhoods (Sjoberg 1960).

Beginning in the tenth century, further town development in the West was facilitated by increases in agricultural technology, population, trade, and communication; the rise of an entrepreneurial class; and an expanding web of social norms regarding economic activity. Communication and manufacturing were revived, which led to the growth of towns with local autonomy and public administration, and eventually to networks of cities. Surplus rural populations migrated to towns and cities, which grew because of their specialization and larger markets, becoming focal points of European societies (Hawley 1981, pp. 37–83).

The emergence of a global economy structured city development (Lo and Yeung 1996, pp. 1–13). In the thirteenth and fourteenth centuries, some European central place and port cities, such as those of the Hanseatic League of Northern Europe, established commercial links with others. Meanwhile, other cities in the Eastern Hemisphere and some cities in the Western Hemisphere were linked together by long-distance trade routes transcending the boundaries of empires and states (Wolfe 1982, p. 250). Military-commercial alliances facilitated the incorporation of territories into states. Commercial ties expanded during the late-fifteenth and the sixteenth centuries, as Europeans developed merchant capitalism, and traded with and colonized peoples on other continents. European and North American states developed more complex economies during the next three centuries, facilitating development of a “European world economy” with a global market, a global division of labor, and a global system of cities (Wallerstein 1974).

### **WESTERN CITIES SINCE THE INDUSTRIAL REVOLUTION**

In most regions of the more developed world, urban growth and urbanization have occurred at an increasing rate since the beginning of the Industrial Revolution. After 1820, the numbers and sizes of urban areas and cities in the United States increased as a result of employment concentration in construction and manufacturing, so urban areas began to grow more rapidly than rural ones.

The population classified as urban by the U.S. Census Bureau (based on aggregations of 2,500 or more and including the surrounding densely populated territories) increased from 5 percent in 1790 to 75 percent in 1990. In 1790 the largest urban place in the United States had fewer than 50,000 inhabitants. As late as 1840, not a single urban place in the United States had more than half a million inhabitants. There were four such places in 1890, fourteen in 1940, twenty-two in 1980, and twenty-three in 1990 (U.S. Bureau of the Census 1997, p. 44).

City growth in the United States during the nineteenth century was driven by migration, since there were sometimes excesses of urban deaths over births in the early part of that century and lower birthrates in the last part of that century. Population concentration facilitated greater divisions of labor within and among families and individuals, as well as increasing numbers of voluntary associations centered on new urban interests and problems. At first, cities were compact, growth was vertical, and workers resided near their workplaces. The outward expansion of the residential population was facilitated in the last part of the century by steam and electric railways, the outward expansion of industry, and the increasing role of the central business district in integrating economic activities (Hawley 1981, pp. 61–145).

The increasing scale of production in the United States led to the development of a system of cities that organized activities in their hinterlands. Cities were differentiated from each other by their degree of dominance over or subordination to other cities, some of which were engaged in centralized manufacturing, others of which depended on transportation, commercial, administrative, or other functions. According to Hawley (1986), cities may have certain key functions that dominate other cities, that is, integrating, controlling, or coordinating activities with these cities. Examples of key functions are administration, commerce, finance, transportation, and communication (Duncan et al. 1960). Since each city exists within its own organizational environment, the expansion of linkages of urban organizations has accompanied the development of a system of cities (Turk 1977).

In the United States, the nature of key functions in city systems changed with the expansion of

settlements westward, as colonial seaports, river ports, Midwestern railway towns, central places on the Great Plains, extractive centers, and government centers were integrated into an urban system. A nationwide manufacturing base was established by 1900, as were commercial and financial centers. By 1960 a fully developed system of differentiated urban centers existed within the United States (Duncan and Lieberman 1970).

### METROPOLITAN AREAS

In the United States, metropolitan areas are defined as being of two kinds. *Metropolitan statistical areas* (MSAs) are areas including one or more central cities with a population of 50,000 or more, or areas with a less densely populated central city but with a combined urban population of at least 100,000 (75,000 in New England), including surrounding counties or towns. *Consolidated metropolitan statistical areas* (CMSAs) contain at least one million population and may have subareas called *primary metropolitan statistical areas* (PMSAs) (Frey 1990, p. 6). In 1990 a majority of the population of the United States resided in the thirty-nine major metropolitan areas: those with more than one million population (Frey 1995, p. 276).

New urban-population-density patterns appeared in the United States with the development of metropolitan areas. Growth was characterized by increases in population density in central cities, followed by increases in density in the metropolitan ring. Transport and communication technologies facilitated linkages of diverse neighborhoods into metropolitan communities dominated by more densely populated central cities. These linkages in turn organized relationships between central cities and less densely populated hinterlands and subcenters. Older metropolitan areas became the centers of CMSAs, while newer areas were the frontiers of expansion, growing by natural increase and especially by net migration. Metropolitan sprawl extended beyond many former nonurban functions, as well as more centrally located older features of the cityscape.

The interior of large Western metropolitan areas represents a merging of urban neighborhoods into complex overlapping spatial patterns, which reflect to some extent the dates when urban neighborhoods were settled and built-up. The blurring of neighborhood distinctions, facilitated by

freeway and mass transit networks, facilitate interaction among “social circles” of people who are not neighborhood-based. Meanwhile, urban neighborhoods organized around such factors as status, ethnicity, or lifestyle, also persist. As the city ages, so does suburban as well as centrally located housing—a delayed consequence of the spread of urban settlement.

Kasarda (1995, p. 239) states that major cities in the United States have been transformed from centers of manufacturing and wholesale and retail trade, to centers of finance, administration, and information processing. Frey (1995, p.272) indicates that during the 1970s and 1980s populations concentrated in metropolitan areas with diverse economies that emphasized service and knowledge industries, as well as in recreation and retirement areas. In the 1970s there was a metropolitan and nonmetropolitan population turnaround. Populations in smaller metropolitan areas grew more rapidly than did those in larger metropolitan areas and population in developing countries grew more rapidly than in developed countries. In the 1980s, metropolitan population shifts, enhanced by international and internal migration, led to an urban revival, increasing regional racial, skill, and age divisions, the suburban dominance of growth and employment and more isolation of the urban core (Frey 1995, pp. 272–275).

Residential moves—the individual counterpart of population redistribution processes—are affected by population characteristics and other factors. Elements that influence mobility decisions are socioeconomic and psychological factors pertaining to the family and the family life cycle, housing and the local environment, and occupational and social mobility (Sabagh et al. 1969). For each of these factors, conditions may restrain persons from moving, or push or pull them to new locations. Information concerning new housing opportunities, the state of the housing market, and the availability of resources may impede or facilitate moves; subsequent moves may stem from recent migration or residential turnover in the metropolitan areas (Long 1988, pp. 219–224).

### TRADITIONAL EXPLANATIONS OF CITIES

During the nineteenth century the character of the Western city was seen as different from that of

noncity areas. Beginning with the public-health movement and concerns with urban housing, social scientists documented “pathologies” of urban life through the use of social surveys, the purpose of which was to provide policy-relevant findings. But since “urbanism as a way of life” has permeated the United States, many of the social and economic problems of cities have spread into smaller towns and rural areas, thus rendering the notion of unique urban pathologies less valid than when it was formulated.

In the first six decades of the twentieth century, explanations of city development were based largely on Western human ecology perspectives, and disagreements have arisen (Sjoberg 1968, p. 455). Theories based on Social Darwinism (Park [1916–1939] 1952) and economics (Burgess 1925, pp. 47–52) were first used to explain the internal structures of cities. “Subsocial” aspects of social and economic organization—which did not involve direct interpersonal interaction—were viewed as generating population aggregation and expansion. Market related competitive-cooperative processes—including aggregation-thinning out, expansion-contraction, centralization, displacement, segregation, migration, and mobility—were believed to determine the structures and patterns of urban neighborhoods (Quinn 1950). The competition of differing urban populations and activities for optimal locations was described as creating relatively homogeneous communities, labeled “natural areas,” which display gradient patterns of decreasing densities of social and economic activities and problems with increasing distance from the city center.

On the basis of competitive-cooperative processes, city growth was assumed to result in characteristic urban shapes. The Burgess hypothesis specifies that in the absence of countervailing factors, the American city takes the form of a series of concentric zones, ranging from the organizing central business district to a commuters’ zone (Burgess 1925, pp. 47–52). Other scholars emphasized star-shaped, multiple-nuclei, or cluster patterns of development. These views were descriptive rather than theoretical; assumed a Western capitalist commercial-industrial city; were distorted by topography, and street and transportation networks; and generally failed to take into account use of land for industrial purposes, which is found in most city zones.

Theodorson (1982) and Michaelson (1976, pp. 3–32) have described research on American cities as reflecting neo-orthodox, social-area analysis, and sociocultural approaches, each with its own frame of reference and methods. Neo-orthodox approaches have emphasized the interdependence of components of an ecological system, including population, organization, technology, and environment (Duncan 1964). This view has been applied to larger ecological systems that can extend beyond the urban community. Sustenance organization is an important focus of study (Hawley 1986). While neo-orthodox ecologists did not integrate the notion of Social Darwinism into their work, the economic aspects of their approach have helped to guide studies of population phenomena, including population shifts and urban differentiation (Frey 1995, pp. 271–336; White 1987), residence changes (Long 1988, pp. 189–251), racial and ethnic diversity (Harrison and Bennett 1995, pp. 141–210), segregation (Farley and Allen 1987, pp. 103–159; Lieberman and Waters 1988, pp. 51–93), housing status (Myers and Wolch 1995, pp. 269–334), and industrial restructuring and the changing locations of jobs (Kasarda 1995, pp. 215–268).

Social-area analysis, in contrast, regards modern industrial society as based on increasing scale, which represents “increased rates and intensities of social relation,” “increased functional differentiation,” and “increased complexity of social organization” (Shevky and Bell 1955). These concepts are related to neighborhood dimensions of social rank, urbanization, and segregation, which are delimited by the factor analysis of neighborhood or census tract measures. “Factorial ecology” has been widely used for classifying census tracts, sometimes for planning purposes. There are disagreements concerning whether factor-analysis studies of urban neighborhood social structure, which are associated with social-area analysis and similar approaches, result in theoretically useful generalizations (Janson 1980, p. 454; White 1987, pp. 64–66)

Sociocultural ecology has used social values such as sentiment and symbolism to explain land use in central Boston (Firey 1947) and other aspects of city life. While values and culture are relevant to explanations of city phenomena, this perspective has not led to a fully developed line of investigation.



Michaelson (1976, pp. 17–32) has argued that none of these aforementioned ecological approaches to the city explicitly study the relationship between the physical and the social environment, for the following reasons: (1) their incomplete view of the environment; (2) their focus on population aggregates; (3) their failure to consider contributions of other fields of study; and (4) the newness of the field. Since Michaelson wrote his critique, sociologists have given more attention to the urban environment.

### NEW EXPLANATIONS OF CITIES

Traditional explanations of cities did not adequately account for the economic, political, racial and social upheavals in U.S. cities during the 1960s. Further, urbanization in less developed regions does not necessarily follow the Western pattern. City growth may absorb national population increments and reflect a lack of rural employment, a migration of the rural unemployed to the city, and a lack of urban industrial development. Increasing concentration of population in larger cities in less developed regions may be followed by the emergence of more “Western-style” hierarchies of cities, functions, and interorganizational relationships. Many cities in less developed regions experience the environmental hazards of Western cities, a compartmentalization of life, persistent poverty and unemployment, a rapidly worsening housing situation, and other symptoms of Western social disorganization. The juxtaposition of local urbanism and some degree of Western urbanization may vitiate a number of traditional Western solutions to urban problems.

Kasarda and Crenshaw (1991, pp. 467–500) have summarized and compared theoretical perspectives concerning how urbanization occurs in less developed regions. Modernization-human ecology perspectives portray city building as the result of changes in social organization and applications of technology. Rural-urban migration, resulting from urban industrialization and excess rural fertility, can eventually lead to the reduction of rural-urban economic and social differences (Hawley 1981). Dependency-world-system perspectives describe the capitalist world-system as guiding change in the less developed regions so as to maintain the dominance of the more developed core (Wallerstein 1974). Urban bias perspectives

emphasize the role of the state, sometimes governed by urban elites, which relies on urban resources, and favors urban over rural development (Lipton 1977). Kasarda and Crenshaw (1991) regard these perspectives as underspecified and as lacking empirical confirmation.

Political economy perspectives explain city growth in more developed as well as less developed regions as a product of globalization, including capital accumulation and nation-state formation (Tilly 1975). Europe and the United States colonized non-European peoples and obtained raw materials from the colonies which they processed and exchanged with their colonies as manufactured products. Colonial areas gained political independence after World War II. Economic restructuring then resulted in the globalization of economic and cultural life, and changes in the international division of labor between cities.

Sassen (1991) indicates that following World War II communications technology facilitated the dispersion of manufacturing by multinational corporations to low-wage cities, some in former colonial areas. A globally integrated organization of economic activity then supplanted world economic domination by the United States. Economic activities were integrated beyond national urban hierarchies into a small number of key global cities, which are now international banking centers, with transnational corporate headquarters, sophisticated business services, information processing, and telecommunications. These cities (London, New York, and to a lesser degree Tokyo) command the global economy and are supported by worldwide hierarchies of decentralized specialized cities (Sassen 1991). International networks of cities sometimes provide opportunities for multinational corporations from less developed regions to penetrate more developed regions (Lo and Yeung 1998, p. 2).

Feagan and Smith (1987, pp. 3–33) describe economic restructuring in U.S. cities in the 1980's as including plant closures and start-ups, the development and expansion of corporate centers, and corporate movement to outlying areas, all of which impact different groupings of cities. They portray economic restructuring as a product of interactions between governmental components of nation-states; multinational, national, and local corporations and businesses; and nongovernment

organizations. These interactions guide policies affecting local taxation, regulation, implementation, and public-private partnerships. Some city neighborhoods have become expendable locations for rapidly shifting economic activity. Policies to accommodate to shifting sites of economic activity influence international and intranational migration, family life, as well as the use of urban, suburban, and rural space.

Aspects of political-economy perspectives have been taken into account in more traditional studies of cities (Frey 1995, pp. 271–336; Kasarda 1995, pp. 215–268). Walton (1993, pp. 301–320) maintains that political-economy perspectives have made the following contributions to the study of cities: (1) showing that urbanization and urbanism are contingent upon the development of social and economic systems; (2) generating comparative studies, particularly in developing countries; (3) elucidating the operation of the informal economy in cities; (4) outlining a political economy of place; (5) showing how globalization relates to ethnicity and community; and (6) relating urban political movements to changes in the global economy.

### RETROSPECT AND PROSPECT

Globalization of economic and cultural life is associated with new urban trends. While cities in less developed regions lack resources when compared with those in more developed regions, cities in both types of regions are becoming more responsive to changes in worldwide conditions.

During the twenty-first century the majority of human population is expected to be urban residents (United Nations 1998). In 1995 approximately 46 percent (2.6 billion) of the world's population lived in urban areas. By the year 2005, approximately half of the population is projected to be urban, compared with fewer than one of every three persons in 1950. By 2030 approximately six-tenths (5.1 billion) of the population is projected to be urban.

Urban areas in less developed areas are projected to dominate the growth of the less developed regions and of the world during the twenty-first century. Reasons are the high population growth rates and high numerical population growth of less developed regions, and high rural-urban migration to cities in these regions. In 1995, only

38 percent of the populations in the less developed regions were urban dwellers, compared with 75 percent of the populations in the more developed regions. Urban areas in less developed areas are expected to account for approximately 90 percent (2.4 billion) of the total of 2.7 billion persons that the United Nations projects will be added to the earth's population from 1995 to 2030 (United Nations 1998). (Cities in more developed regions are expected to account for only 140 million of the population increase in the same period.)

There will continue to be significant differences in urbanization between the less developed regions. About three fourths of Latin-American population was urban in 1995, roughly the average level of industrialized regions in Europe, Japan, Australia, New Zealand, the United States and Canada. The most extensive future urban growth will be in Asia and Africa, which are now only about one-third urban (O'Mera 1999; United Nations 1998).

An increasing portion of urban population is residing in giant urban agglomerations. An *urban agglomeration*, according to the United Nations (1998), is the population within a contiguous territory inhabited at urban levels without regard to administered boundaries. In 1995 the fifteen largest ranged in size from 27 million (Tokyo) to 9.9 million (Delhi). Less developed countries are sometimes characterized by primate cities, that is, the largest city in a country dominates other cities and are larger than would be expected on the basis of a "rank-size" rule, which indicates that the rank of a population aggregation times its size equals a constant (Shryock, Siegel, and associates 1976), thus resulting in an inadequate supporting hierarchy of smaller cities. Primate cities often appear in small countries, and in countries with a dual economy, but are not as apparent in large countries or those with long urban histories (Berry 1964).

*Megacities*, defined by the United Nations (1998) as those cities with 10 million or more inhabitants, are increasing in number and are concentrated in less developed regions. There were fourteen megacities in 1995, including ten in less developed regions. Twenty-six megacities are projected in 2015, including twenty-two in less developed regions, of which sixteen will be in Asia. Megacities have both assets and liabilities. Brockerhoff and

Brennan (1998, pp. 75–114) have suggested that cities' size and rates of population growth are inversely related to welfare. Kasarda and Crenshaw (1991, pp. 471–474) note that megacities, which may have serious social and economic problems, can be driving engines for industrial production in developing regions, and contribute disproportionately to their countries' economies (Kasarda and Crenshaw 1991, pp. 467–501).

Berry (1978) indicates that urbanization may be accompanied by “counter-urbanization” or decreasing city size and density. City growth sometimes occurs in cycles, which begin with rapid growth in the urban core, followed by rapid growth in the suburban ring, a decline in growth in both the core and the ring, and then rapid growth in the core (United Nations 1998). These cycles—of urbanization, suburbanization, counter-urbanization, and reurbanization—appear to be associated with concentrations of services in city centers, followed by improvements in commuting by the labor force and increased suburban home ownership by urban labor forces. The 1970s deurbanization (metropolitan turnaround) in the United States was followed by reurbanization in the 1980s, and the start of another period of deurbanization in the 1990s (United Nations 1998). During the 1970s and 1980s counter-urbanization occurred in other more developed regions and in less developed regions, including a slow-down of population growth rates in some megacities in developing regions, particularly in Latin America (United Nations 1998).

Changes in the global economy and in the aforementioned city growth cycles are associated with new forms of urban land use in the United States, which are laid over preexisting urban patterns. Businesses, which provide the economic base for cities, move between optimum locations in different cities (Wilson 1997, p. 8). Globalization has enhanced the growth of suburbs and “edge cities” organized around outlying business and high-technology centers linked by telecommunications networks to other cities (Castells, 1989; Muller 1997). Unregulated informal sectors of the economy develop in a variety of intraurban locations. High-income native and immigrant populations, which profit from the new global economy, sometimes cluster in protected enclaves. Low-income immigrant and ethnic populations often occupy areas inhabited by earlier cohorts of the

urban poor. Such trends may be related to a spread of ghetto-underclass neighborhoods and to increasingly polarized intracity neighborhood differences of poverty and affluence (Morenoff and Tienda 1997, pp. 59–72).

Trans-border cites are new urban forms that transcend the boundaries of nation-states and reflect increasingly borderless economies. They result from globalization trends, including the economic integration of regions, reductions of trade barriers, and the establishment of multinational free trade zones. They link cities adjacent to a border, and sometimes metamorphose into trans-border systems, complete with specialized functions and populations, and extensive cross-border social and economic ties (Rubin-Kurtzman et al. 1993). Examples are Southern California (U.S.)-Baja California (Mexico), the Singapore-Johore (Malaysia)-Riau (Indonesia) region, and the Beijing (China)-Pyongyang (North Korea)-Seoul (South Korea)-Tokyo (Japan) urban corridor. Trans-border cities pose questions regarding the limits of national sovereignty, but can also integrate human and economic resources and thus enhance international stability.

Cities will continue to exhibit extremes of affluence and poverty, but the extent and consequences of these extremes are unclear. Massey (1996, pp. 395–412) has argued that both the affluent and the poor are concentrating in cities in the United States; consequences may include increased densities of crime, addictions, diseases, and environmental degradation, the emergence of oppositional subcultures, and enhanced violence. Massey assumes these trends apply to less developed regions as well as to the United States. Farley (1996, pp. 417–420) has advanced a counterargument that while economic inequality is increasing in the United States, as may be the geographic segregation of the poor, the continuing (1996) rise in prosperity increases welfare at all income levels. Firebaugh and Beck (1994, pp. 631–653) maintain that economic growth, even in dependent less developed countries, may eventually increase welfare. Hout et al. (1998) state that political institutions, which are partially responsible for growing inequality, can provide appropriate remedies.

As cities account for increasing shares of the earth's population, they will consume increasing shares of the earth's resources, produce increasing

shares of pollution, and their populations will be more subject to negative feedbacks from human impacts on the environment. O'Mera (1999, p. 137) estimates that populations of cities, while occupying only two percent of the earth's surface, consume 75 percent of the earth's resources. Cities are often sited on areas that contain prime agricultural land; urban expansion then inhibits and degrades crop production. Conversion of rural land to urban use intensifies natural hazards including floods, forest and brush fires, and earth slides. Further development concentrates and then disburse to outlying locations such artificial hazards as air pollution, land and water pollution, and motor vehicle and air traffic noise. Cities are also sited on the shorelines of oceans, thus increasing numbers of city residents are subject to the impacts of storm surges and erosion. Lowry (1992) argues that environmental impacts attributed to cities reflect population growth, industrialization and prosperity rather than city growth itself.

Some of the aforementioned trends are apparent, for example, in Los Angeles and Mexico City, which ranked seventh and second in size, respectively, among the world's largest megacities in 1995. After World War II and until the 1970s, Los Angeles's industrial and population growth and suburban sprawl made it a prototype for urban development that brought with it many inner-city, energy, suburban, environmental, state-management, ethnic, and capital-accumulation problems.

Since the early 1970s, Los Angeles has become a radically changed global megacity based on accumulation of global capital, economic restructuring, communications, and access to new international markets. The reorganization of Los Angeles has affected labor-force demands and the character of both native and immigrant populations. Immigration transformed Los Angeles into a multicultural metropolis (Waldinger and Bozorgmehr 1996, pp. 3-37). The area's economy emphasizes high-income service-sector jobs and low-income service-sector and industrial-sector jobs. Los Angeles has a great concentration of industrial technology and great international cultural importance, is dependent on the private automobile, has environmental hazards (earthquakes, brush fires, and air pollution), limited water supplies, housing shortages, crime, and has experienced uncontrolled intergroup conflicts in 1965 and 1992.

Mexico City, while not commanding the same global economic stature as Los Angeles, is the primate city of Mexico. Mexico City shares many of Los Angeles's problems, including environmental hazards (earthquakes, and air pollution), water shortages, lack of housing and services, expansion of low-income settlements, dependence on the private automobile, lack of sufficient transportation, earthquakes, crime, and rising ethnic violence. Effects of policy responses to these problems have been dampened by economic reversals (Ward 1998).

These illustrations suggest that cities at somewhat similar levels of influence within their respective countries share similar characteristics, whether in the more developed or in the less developed regions. This would support a view that determinants of city development are rooted in the global economy and are influenced by similar trends, but vary according to the city's place in the system.

## REFERENCES

- Berry, Brian J. L. 1978 "The Counter-Urbanization Process: How General?" In Niles M. Hansen, ed., *Human Settlement Systems: International Perspectives on Structure, Change and Public Policy*. Cambridge, Mass.: Ballinger.
- 1964 "Cities as Systems Within Systems of Cities." In J. Friedman and W. Alonso, eds., *Regional Development and Planning*. Cambridge, Mass.: MIT Press.
- Brockhoff, Martin, and Ellen Brennan 1998 "The Poverty of Cities in Developing Regions." *Population and Development Review* 24:1:75-114.
- Burgess, Ernest W. (1924) 1925 "The Growth of the City, an Introduction to a Research Project." In Robert E. Park, Ernest W. Burgess, and R. D. McKenzie, eds., *The City*. Chicago: University of Chicago Press.
- Castells, Manuel 1989 *The Informational City: Information Technology, Economic Restructuring, and the Urban-Regional Process*. Cambridge, Mass.: Basil Blackwell, Ltd.
- Childe, V. Gordon 1950 "The Urban Revolution." *Town Planning Review* 21:3-17.
- Duncan, Beverly, and Stanley Lieberman 1970 *Metropolis and Region in Transition*. Beverly Hills, Calif.: Sage.
- Duncan, Otis D. 1964 "Social Organization and the Ecosystem." In Robert E. L. Faris, ed., *Handbook of Modern Sociology*. Chicago: Rand McNally.

- , William R. Scott, Stanley Lieberman, Beverly Duncan, and Haliman H. Winsborough 1960 *Metropolis and Region*. Baltimore: Johns Hopkins University Press.
- Farley, Reynolds 1996 “The Age of Extremes: A Revisionist Perspective.” *Demography* 37:417–420.
- , and Walter R. Allen 1987 *The Color Line and the Quality of Life in America*. New York: Russell Sage Foundation.
- Feagin, Joe R., and Michael Peter Smith 1987 “Cities and the New International Division of Labor: An Overview.” In Michael Peter Smith and Joe R. Feagin, eds., *The Capitalist City*. New York: Basil Blackwell, Ltd.
- Firebaugh, Glen, and Frank D. Beck 1994 “Does Economic Growth Benefit the Masses? Growth Dependence and Welfare in the Third World.” *American Sociological Review* 59:631–653.
- Firey, Walter 1947 *Land Use in Central Boston*. Cambridge, Mass.: Harvard University Press.
- Frey, William 1995 “The New Geography of Population Shifts.” In Reynolds Farley, ed., *State of the Union: America in the 1990s, Volume Two: Social Trends*. New York: Russell Sage Foundation.
- 1990 “Metropolitan America: Beyond the Transition.” *Population Bulletin* 45, 2:1–53.
- 1987 “Migration and Depopulation of the Metropolis: Regional Restructuring or Rural Renaissance?” *American Sociological Review* 52:240–257.
- Harrison, Roderick, and Claudette E. Bennett 1995 “Racial and Ethnic Diversity.” In Reynolds Farley, ed., *State of the Union: America in the 1990s, Volume Two: Social Trends*. New York: Russell Sage Foundation.
- Hawley, Amos H 1986 *Human Ecology: A Theoretical Essay*. Chicago: University of Chicago Press.
- 1981 *Urban Society: An Ecological Approach*, 2nd ed. New York: Wiley.
- Hout, Michael, Richard Arons, and Kim Voss 1996 “The Political Economy of Inequality in the ‘Age of Extremes.’” *Demography* 33:421–425.
- Jacobs, Jane 1970 *The Economy of Cities*. New York: Vintage.
- Janson, Carl-Gunnar 1980 “Factorial Social Ecology: An Attempt at Summary and Evaluation.” In A. Inkeles, N. J. Smelser, and R. H. Turner, eds., *Annual Review of Sociology*, vol. 6. Palo Alto, Calif.: Annual Reviews.
- Kasarda, John 1995 “Industrial Restructuring and the Changing Nature of Jobs.” In Reynolds Farley, ed., *State of the Union: America in the 1990s, Volume One: Economic Trends*. New York: Russell Sage Foundation.
- , and Edward M. Crenshaw 1991 “Third World Urbanization: Dimensions, Theories, and Determinants.” In W. Richard Scott, ed., *Annual Review of Sociology* 17:467–500.
- Lieberman, Stanley, and Mary C. Waters 1988 *From Many Strands: Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage Foundation.
- Lipton M. 1977 *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge, Mass.: Harvard University Press.
- Lo, Fu-chen, and Yoe-Man Yeung 1996 “Introduction.” In Fu-chen Lo and Yoe-Man Yeung, eds., *Emerging World Cities in Pacific Asia*. Tokyo: United Nations University Press.
- Long, Larry 1988 *Migration and Residential Mobility in the United States*. New York: Russell Sage Foundation.
- Lowry, Ira S. 1991 “World Urbanization in Perspective.” In Kingsley Davis and Mikhail S. Bernstram, eds., *Resources, Environment, and Population: Present Knowledge: Future Options, Population and Development Review, a supplement to Volume 16, 1990*.
- Massey, Douglas S. 1996 “The Age of Extremes: Concentrated Affluence and Poverty in the Twenty-first Century.” *Demography* 33:395–412.
- Michaelson, William 1976 *Man and His Urban Environment: A Sociological Approach*. Reading, Mass.: Addison-Wesley.
- Morenoff, Jeffrey, and Marta Tienda 1997 “Underclass Neighborhoods in Temporal and Ecological Perspective.” In Alan W. Heston, ed., David Wilson, special ed., *The Annals of the American Academy of Political and Social Science Globalization and the American City* 551:59–72.
- Muller, Peter O. 1997 “The Suburban Transformation of the Globalizing American City.” In Alan W. Heston, ed., David Wilson, special ed., *The Annals of the American Academy of Political and Social Science: Globalization and the Changing American City* 551: 44–50.
- Myers, Dowell, and Jennifer R. Wolch 1995 “The Polarization of Housing Status.” In Reynolds Farley, ed., *State of the Union: America in the 1990s, Volume Two: Social Trends*. New York: Russell Sage Foundation.
- O’Mera, Molly 1999 “Exploring a New Vision for Cities.” In Linda Stack, ed., *State of the World 1999*. New York: W.W. Norton and Company.
- Park, R. E. (1916–1939) 1952 “Human Communities: The City and Human Ecology.” *Collected Papers*, vol. 2. Glencoe, Ill.: Free Press.
- Phillips, E. Barbara 1996 *City Lights: Urban-Suburban Life in the Global Society*. New York: Oxford University Press.

- Quinn, James 1950 *Human Ecology*. New York: Prentice-Hall.
- Rubin-Kurtzman, Jane R., Roberto Ham Chande, and Maurice D. Van Arsdol, Jr. 1993 "Population in Trans-Border Regions: The Southern California-Baja, California Urban System." *International Migration Review* 30:1020–1045.
- Sabagh, Georges, Maurice D. Van Arsdol, Jr., and Edgar W. Butler 1969 "Determinants of Intrametropolitan Residential Mobility: Conceptual Considerations." *Social Forces* 48, 1:88–97.
- Sassen, Saskia 1991 *The Global City*. Princeton, N.J.: Princeton University Press.
- Shevky, Eshref, and Wendell Bell 1955 *Social Area Analysis: Theory, Illustrative Applications and Computational Procedures*. Stanford Sociological Series, no. 1. Stanford, Calif.: Stanford University Press.
- Shryock, Henry S., Jacob S. Siegel, and associates 1976 *The Materials and Methods of Demography*, condensed edition ed. by Edward G. Stockwell. San Diego: Academic Press/Harcourt Brace Jovanovich.
- Sjoberg, Gideon 1968 "The Modern City." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 1. New York: MacMillan and Free Press.
- 1960 *The Pre-Industrial City: Past and Present*. New York: Free Press.
- Theodorson, George A. 1982 *Urban Patterns: Studies in Human Ecology*. University Park: Pennsylvania State University Press.
- Tilly, Charles 1974 *The Formation of National States in Western Europe*. Princeton, N.J.: Princeton University Press.
- Turk, Herman 1977 *Organizations in Modern Life: Cities and Other Large Networks*. San Francisco: Jossey-Bass.
- United Nations 1998 *World Urbanization Prospects: The 1996 Revision*. New York: United Nations.
- U.S. Bureau of the Census 1997 *Statistical Abstract of the United States*. Washington, D.C.: Government Printing Office.
- Waldinger, Roger, and Mehdi Bozorgmehr 1996 "The Making of a Multicultural Metropolis." In Roger Waldinger and Mehdi Bozorgmehr, eds., *Ethnic Los Angeles*. New York: Russell Sage Foundation.
- Wallerstein, Immanuel 1974 *The Modern World System: Capitalist Agriculture and the Origin of the European World Economy in the Sixteenth Century*. New York: Academic Press.
- Walton, John 1993 "Urban Sociology: The Contribution and Limits of Political Economy." In Judith Blake, ed., *Annual Review of Sociology*. Palo Alto, Calif.: Annual Reviews Inc.
- Ward, Peter M. 1998 *Mexico City: Revised Second Edition*. Chichester, N.Y.: John Wiley.
- Weber, M. 1958 "The Nature of the City." In Don Martindale and Gertrude Neuwirth, eds., *The City*. New York: Free Press.
- White, Michael J. 1987 *American Neighborhoods and Residential Differentiation*. New York: Russell Sage Foundation.
- Wilson, David 1997 "Preface." In Alan W. Heston, ed., David O. Wilson, special ed., *The Annals of the American Academy of Political and Social Science: Globalization and the Changing U.S. City*, 551: 8–16.
- Wirth, Louis 1938 "Urbanism as a Way of Life." *American Journal of Sociology* 44:1–24.
- Wolf, Eric R. 1982 *Europe and the People Without History*. Berkeley: University of California Press.

MAURICE D. VAN ARSDOL, JR.

## CIVIL DISOBEDIENCE

See Protest Movements; Student Movements.

## CIVIL LIBERTIES

Civil liberties and associated controversies reflect the basic sociological issue of what may comprise the requirements of a free yet sustainable society. Classical interests of social thought directly or indirectly concern civil liberties because they address the degree to which individuals may exercise autonomy within the bounds of enduring social relations and community needs. Some of sociology's most venerable research has focused directly on civil liberties. Stouffer's *Communism, Conformity, and Civil Liberties* (1955) served as an intellectual punctuation mark on the McCarthy era, the period during which public discord over civil liberties reached its most intense state in the post-World War II world. Sociological thinking and research has helped American intellectuals and policy makers frame the issues associated with civil liberties and understand the implications of decisions regarding civil liberties for the well-being of society.

Civil liberties may be understood as legally protected areas in which the individual may function without interference by the state or the broader community of citizens. “Civil liberties” are analytically distinct from “civil rights.” Civil liberties concern the individual’s freedom from the broader society and its laws. Civil rights derive from the individual’s claim on society and the state to give him or her equal protection through the state’s police power and equal rights regarding public facilities, services, and largesse. Civil liberties concern the individual’s rights to think, speak, and act outside the state’s apparatus and jurisdiction. Civil rights address the individual’s claim on equal access to public resources such as buses and schools, protection from harm by state agencies, and participation in government and politics.

Sociology offers several core capabilities to promote the citizens’ and policy makers’ understanding of civil liberties and the implications of related public decisions. Classical and contemporary work by sociologists has pertinence in three areas. First, sociological theory and commentary informed by theory helps specify the central dilemmas raised by civil liberties. Second, sociological research enables observers to discover the state of public opinion regarding civil liberties, the dynamics by which public opinion has developed and changed in the past, and the manner in which public opinion may unfold in the future. Finally, sociological thinking and research serves as a resource for understanding the potential consequences of public decisions regarding civil liberties in the years to come. This last capability can aid public decision making and help lay groundwork for achieving the broadest range of civil liberty in society while maintaining the social cohesion necessary to ensure stability, continuity, and affirmation of individual life by core social institutions.

Debate regarding civil liberties has traditionally concerned freedom of expression and due process. Positions regarding freedom of expression have sought to protect the right of individuals to publicly support politically unpopular causes or to display or publish material others may view as objectionable (e.g., pornography). Due process issues have focused on the rights of defendants in criminal cases and claimants in civil and administrative proceedings. Civil liberties advocates have drawn core support from the Bill of Rights and subsequent amendments to the U.S. Constitution

safeguarding free speech, prohibiting unreasonable search and seizure, and limiting the criminal justice system’s ability to require citizens to give self-incriminating testimony. Civil liberties advocates depend heavily on legal doctrines and devices derivative of the Bill of Rights such as fairness, equal protection of the law, and the right of privacy.

The late twentieth century saw a vast extension of activities to which the status of a civil liberty was applied. The 1998 edition of the American Civil Liberties Union’s (ACLU) *The Year in Civil Liberties*, for example, reports challenges by ACLU units to practices and policies such as:

- Religious celebration in public settings (viz. school prayer and holiday displays),
- Youth curfews,
- Prohibition of marijuana use for medical purposes,
- School vouchers,
- Wrongful dismissal from employment due to politics or sexual preference,
- Restraint and corporal punishment of prisoners and “high-risk” legal defendants,
- The death penalty,
- Prohibition of public funding for abortions,
- Legal barriers to adoption of children by lesbian or gay individuals or couples,
- Restriction of legal marriage to heterosexuals only,
- “Sodomy” laws.

According to the ACLU document, these laws and policies belong to a common category of threats to “fairness, freedom of expression, equality, and keeping the government out of our private lives.”

The ACLU does not unilaterally speak for those concerned with civil liberties. But controversy surrounding the ACLU’s extended definition of civil liberties recapitulates central issues in sociological theory, social thought, and public policy. Early social theorists emerging from free-market economics and utilitarianism encountered (wittingly or by implication) the question of what

holds society together, given that individuals “naturally” behave in an atomized manner. Modern participants in civil liberties controversies confront (again wittingly or by implication) a tension between unrestricted individual liberty and the needs of the community and requirements for viable social institutions.

Etzioni, in an essay commenting on the ACLU’s expansion of concerns, emphasizes contradictions between individual liberties and community needs (Etzioni 1991). He stresses the necessity of modifying constitutionally protected individual rights in instances of compelling social exigency. Examples of such modification in the late twentieth century included x-raying of luggage at airports, conducting voluntary fingerprinting of children to facilitate their identification if kidnapped, contact-tracking for people infected with HIV, and mandatory drug testing of workers whose impairment endangers others, such as train engineers. Although these measures have enjoyed public support and none has materially affected the basic rights of the general population, each has been the focus of civil liberties controversies and actions.

Etzioni characterizes opposition to measures such as these as “radical individualism,” encouraged at late century by an imbalance between “excessive individual rights and insufficient social responsibility.” His analysis characterizes the U.S. Constitution as broader than a code of legal provisions to protect the individual from government. The law of the land is also a reflection of “public morality, social values, and civic virtue.” Eclipse of these elements of civil society, Etzioni implies, precludes even marginal modification of legal traditions in the face of compelling social need. He warns that resulting governmental paralysis may ultimately give rise to popular disillusionment, social distress, and abandonment of safeguards to personal liberty on a far greater scale than the marginal modifications initially proposed.

Another critic of ACLU positions alleges that the imbalance between one-sided civil liberties protection and community needs has already affected American social institutions adversely and to a significant degree. Siegel (1991) writes that “the libertarians and their allies in the courts have . . . reshaped virtually every American public institution in the light of their understanding of due

process and equal protection under the Fourteenth Amendment.” This reshaping has had an “elitist” quality, proceeding through abstract legal reasoning and argument but materially harmful, particularly to the economically disadvantaged. According to Siegel’s argument, civil liberties victories in court place burdens on social institutions and prevent them from responding to social reality. Siegel writes:

*Civil liberties have become an economic issue as those who can afford it either flee the cities or buy out of public institutions. For those who can’t afford to pay for private school, or private vacations, and are left with junkie infested parks, who can’t afford the private buses which compete with public transportation, and are unable to pay for private police protection, the rights revolution has become a hollow victory. The imposition of formal equality, the sort that makes it almost impossible, for instance, to expel violent high school students, has produced great substantive inequality as would-be achievers are left stranded in procedurally purified, but failing institutions. (Siegel 1991)*

Both Etzioni’s and Siegel’s critique of the civil liberties movement reflect sociology’s core perspective and concern, the essential tension between individualistic and social forces. More concretely, sociology’s traditional concern with civil liberties has focused on the citizen’s thinking regarding tolerance of deviation. Survey research has served as the primary source of such information.

Stouffer’s above-referenced classic sounded an optimistic note at the conclusion of the McCarthy era. His study focused on tolerance of people espousing communism and atheism, “nonconformist” ideologies that excited widespread public hostility at the time. In separate surveys conducted by the National Opinion Research Center (NORC) and the Gallup organization, Stouffer asked respondents whether communists and atheists should be allowed to speak in their communities, whether they should be allowed to teach in colleges or universities, and whether their books should be removed from public libraries.

The Stouffer study is remembered largely for reporting relationships between two focuses of



social distinction and tolerance for the nonconformist ideologies. Community leaders and people with advanced education were more likely to score in the “more tolerant” range than the national cross section. On this basis, Stouffer concluded that Americans would become more tolerant of nonconformity in the decades to follow, since the average American was receiving more years of education than his or her parents. Level of education correlated strongly with tolerance in every age group except sixty and over.

A subtheme in the heritage left by Stouffer was evidence for personality-based causes of intolerance regarding civil liberties for deviants. Experience with European totalitarianism had led psychologists to develop the theory of the “authoritarian personality.” Measured according to a device known as the “F scale,” personalities of this kind were distinguished by a simplistic world view, respect for power, and obedience to authority (Adorno 1950). Statistically significant relationships were found between F scale items and intolerance in the Stouffer data. Consistent with these findings, sociologists such as Lipset (1981) claimed that authoritarianism was more likely to be found in the working rather than the middle or upper classes. It is tempting to conclude that a negative relationship between education and basic authoritarianism explains the greater willingness of educated people to extend civil liberties to the politically unpopular, and to speculate that greater education will reduce, if it has not reduced already, personality-related proclivities toward intolerance.

Later research, though, has shown the sociology of public opinion regarding civil liberties to be more complex. Early critics pointed out technical flaws in the F scale. The scale’s items, for example, were all worded in the same direction, encouraging positive responses. Critics raised the possibility that reported relationships between F scale scores and education merely reflected a positive response bias which was particularly strong among working-class respondents. Members of the working class, it was theorized, have a tendency to acquiesce to strong, positive assertions, particularly when these are presented by higher-status individuals such as pollsters.

Reanalysis of the Stouffer data and analysis of data from NORC’s 1990 General Social Survey (GSS) by Schuman, Bobo, and Krysan (1992) casts

doubt upon the causal chain implied above: that low social status (indicated by education) “causes” authoritarian personality, and that authoritarian personality subsequently “causes” intolerance of civil liberties for nonconformists and deviates. Reanalyzing Stouffer’s data, these investigators found relationships between authoritarianism and intolerance of communists and atheists only among the more highly educated. In the 1990 GSS data, they found relationships between authoritarianism and intolerance for blacks and Jews again confined to the educated. The investigators conclude that there is no substantive relationship between class and authoritarianism. Evidence does emerge for a relationship between personality factors and both support for civil liberties for nonconformists and tolerance of minorities. But the roots of these personality factors are unknown and presumably much more complex than class-based socialization.

Changes in public concerns since the 1950s make it risky to apply the findings of Adorno, Stouffer, and others of their era to today’s citizens and social issues. By the end of the twentieth century, communism and atheism had ceased to be mainstream public concerns in the United States. Analysis of civil liberties issues regarding crime had risen to prominence. Remedies such as permanent incarceration of habitual criminals and community notification regarding sex offenders (“Megan’s Law”) had been widely adopted. Increased latitude by police for searching and surveillance of citizens was widely discussed.

Public attitudes favoring compromise of civil liberties in the interests of aggressive law enforcement seemed stable during the 1980s and 1990s. Comparison over time of poll results on the trade-off between aggressive policing and civil liberties indicates growing support for warrantless police searches of cars and drivers. Decided majorities of respondents to Roper and Gallup polls in 1985 and 1986 approved of school officials’ searching students’ belongings for drugs or weapons, again without a warrant. The late twentieth century, though, saw no large-scale support for abandonment of civil liberties in pursuit of greater security. One trend showed a modest rise in support for surveillance of citizens, but another indicated just the opposite: the public did not think it was necessary to “give up some civil liberties” to prevent terrorism (Shaw 1998).

Civil libertarians might feel more alarmed by the polls' findings regarding public ignorance about constitutional rights. According to one survey, only 56 percent of Americans were aware of the innocent-until-proven-guilty principle (Parisi 1979). In another study, only one-third of the respondents correctly indicated the truth or falsehood of a statement regarding double jeopardy (McGarrell and Flanagan 1985).

Review of the studies cited above implies two major conclusions about public opinion regarding civil liberties. First, social determinants of support for civil liberties are likely to be complex and to change over time. Second, the specific focus of concern surrounding civil liberties—for example, the rights of communists versus those of crooks—may predominantly affect their support among citizens. Continual exercise of pertinent sociological research tools is required to maintain awareness of civil liberties-related attitudes and trends; associated theories appear in periodic need of reconstruction.

There is good evidence that sociological thinking and research techniques can promote understanding of the consequences of public decisions regarding civil liberties and help balance civil liberties and community needs. Etzioni's critique includes a recommendation for "limited adjustment" of civil liberties in the interests of society. Criteria for activation of limited adjustment include a "clear and present danger" of sufficient gravity to "endanger large numbers of lives, if not the very existence of our society," and a "direct link between cause and effect." As illustrations, Etzioni cites nuclear weapons, crack cocaine, and AIDS. He recommends minimal interference with constitutional rights, seeking remedies whenever possible that do not actually involve civil liberties.

Even the most measured approach to "adjustment" of civil liberties, though, raises issues for social theory and research. Designation of "clear and present danger" is as much a social fact as one of nuclear physics, pharmacology, and epidemiology. Civil libertarians may justifiably ask what makes crack cocaine a potential threat to society while other narcotics, while causing significant human misery, have not brought society down. Similar issues may be raised regarding AIDS, a biologically less-contagious disease than the traditional scourges

of syphilis and gonorrhea. To what extent, researchers should ask, may the objective significance of these threats have been exaggerated by public emotion?

The impacts of small modifications of civil liberties on the problems these measures are intended to ameliorate should also be viewed as empirical issues. Does contact-tracking of people with AIDS actually drive some underground, making their disease invisible to society and hence more dangerous? If so, how many go underground, for how long, and by what means? Research on likely behavior of people with AIDS and other stigmatized diseases is an essential adjunct to related public decision making.

Finally, the degree to which minor adjustment may ultimately weaken the fabric of civil liberties is a necessary direction for research. Etzioni puts aside the notion that minor modification may initiate a slide down the "slippery slope" toward government or communitarian domination by citing the innocuous nature of procedures such as child fingerprinting. But systematic examination of many seemingly small adjustments may indicate that some indeed result in cascades of increasingly pernicious modifications. Study of the conditions under which the minor modification of traditions has in fact led to their eventual collapse could form the basis of a relevant theory.

## REFERENCES

- Adorno, Theodore W. et al. 1950 *The Authoritarian Personality*. New York: Harper.
- Etzioni, Amitai 1991 "Too Many Rights, Too Few Responsibilities." *Society* 28:41-48.
- Lipset, Seymour M. 1981 *Political Man*. New York: Doubleday.
- McGarrell, E.F., and T. Flanagan (eds.) 1985 *Sourcebook of Criminal Justice Statistics-1984*. Washington, D.C.: Department of Justice.
- Parisi, Nicolette et al. (eds) 1979 *Sourcebook of Criminal Justice Statistics-1978*. Washington, D.C.: Department of Justice.
- Schuman, Howard, Lawrence Bobo, and Maria Krysan 1992 "Authoritarianism in the General Population: The Education Interaction Hypothesis." *Social Psychology Quarterly* 55:379-387.
- Shaw, Greg M. et al. 1998 "Crime, the Police and Civil Liberties." *Public Opinion Quarterly* 62:405-426.

Siegel, Fred 1991 "Individualism, Etatism, and the ACLU." *Society* 28:20-22.

Stouffer, Samuel A. 1955 *Communism, Conformity, and Civil Liberties*. Garden City, N.Y.: Doubleday.

Whitfield, Emily, and Amy Weil 1998 *The Year in Civil Liberties: 1998*. New York: American Civil Liberties Union.

HOWARD P. GREENWALD

## CIVIL RIGHTS

See Protest Movements; Student Movements.

## CLANS

See Indigenous Peoples.

## CLASS AND RACE

There is considerable debate in the sociology of race relations over how social inequality based on class and that based on race intertwine or intersect. Are these separate dimensions of inequality that simply coexist? Or are they part of the same reality?

Efforts to develop an understanding of the relationship between class and race have a long history in sociology. In the 1930s and 1940s it was common to conceptualize the issue as "caste and class" (Davis, Gardner, and Gardner 1941). Studies were conducted in southern towns of the United States, and a parallel was drawn between the southern racial order and the Indian caste system. Class differentiation was observed within each of the two racial "castes," but a caste line divided them, severely limiting the social status of upper-class African Americans. This view, while descriptively illuminating, was challenged by Cox (1948), who saw U.S. race relations as only superficially similar to caste and based on a very different dynamic.

The relationship between class and race remains hotly debated today. Wilson (1980, 1987) has argued that class has superceded race as a factor in the continuing disadvantage of black

inner-city communities. Wilson argues that economic forces, including the exodus of major industries, have more to do with the social problems of the inner city than do race-based feelings and actions. On the other hand, Omi and Winant (1986) assert the independence of race from class and resist the reduction of race to class forces. They claim that the United States is organized along racial lines from top to bottom and that race is a more primary category than class.

For most Marxist sociologists of race relations, class and race cannot be treated as separate dimensions of inequality that somehow intersect. Rather, they argue that race and class are both part of the same system and need to be understood through an analysis of the system as a whole. Modern race relations are seen as distinctive products of the development of world capitalism. Both racism and capitalism developed together reinforcing one another in a single, exploitative system. The central question then becomes: How has capitalism, as a system based on class exploitation, shaped the phenomena of race and racism?

### CAPITALISM AND RACISM

Conventional thinking tends to follow the line that the development of capitalism should eliminate racism. People holding this position argue that racism is an unfortunate leftover from more traditional social systems. Capitalism, based on rational criteria such as efficiency, should gradually eliminate the irrational features of the past. The market is "colorblind," it should only select on the basis of merit. For example, in the area of job allocation, selecting on the basis of such irrelevant criteria as skin color or the race of one's great grandparents, would lead those firms that so choose to perform less well than those that select purely on the basis of ability, and they would go out of business. Only the rational, colorblind firms would survive and racism would disappear in the labor market.

Unfortunately, this idealized theoretical model of the way capitalism works has not proved true in practice. We continue to live in a highly segregated society, with a continuing racial division of labor, and with a high degree of racial inequality on every social and economic dimension. White families, on average, control much higher levels of wealth than African-American families, for example (Oliver and Shapiro 1995). The continuation

of racism within advanced capitalist societies requires further explanation.

Whereas cultural differences have served as a basis for intergroup conflicts for the entire history of humanity, the expansion of Europe, starting in the sixteenth century, set the stage for a new form of intergroup relations. Never before was conquest so widespread and thorough. Nor was it ever associated with such a total ideology of biological and cultural inferiority. Modern racism, with its pseudoscientific claims of inferiority, is a unique phenomenon.

An understanding of European expansion, and its impact on people of color, begins with an analysis of capitalism as it developed in Europe. Capitalism is a system that depends on the private ownership of productive property. In order to earn profits on property, the owners depend on the existence of a nonowning class that has no alternative but to sell its labor-power to the owners. The owners accumulate wealth through profit, that is, the surplus they extract from labor. Hence, a class struggle develops between capitalists and workers over the rights of capitalists to the surplus.

In Europe, labor came to be “free,” that is, people were no longer bound by serfdom or other forms of servitude but were free to sell their labor-power on an open market to the highest bidder. Being free in this sense gave European workers a certain political capacity, even though they were often driven to conditions of poverty and misery.

Capitalism is an expansionary system. Not only does it unleash great economic growth, but it also tends to move beyond national boundaries. The expansionist tendencies lie in a need for new markets and raw materials, a search for investment opportunities, and a pursuit of cheaper labor in the face of political advances by national labor forces. European capitalism thus developed into an imperialistic system (Lenin 1939).

European imperialism led to a virtually total conquest of the globe. Europe carved up the entire world into spheres of influence and colonial domination. The idea and ideology of race and racism emerged from this cauldron. Europeans constructed a kind of folk-scientific view of human differences, dividing the world’s human population into semi-species or “races.” Of course, this division

has no basis in fact, and racial categorization has been completely discredited. Nevertheless, the idea of race, and its use in structuring societies along hierarchical lines, remains exceedingly robust. In sum, race is strictly a social construction, but one with profound implications for the way society is organized.

European domination took multiple forms, from unequal treaties, unfair trade relations, conquest, and the establishment of alien rule to annihilation and white settlement in places where once other peoples had thrived. Imperialism received ideological justification in beliefs that non-European cultures were primitive, uncivilized, barbaric, and savage, and their religions were pagan and superstitious. Europeans were convinced that they had the true religion in Christianity and that all other peoples needed to be “saved.” The denigration of other cultures was accompanied by beliefs in natural, biological inferiority. Dark skin color was a mark of such inferiority, while white skin was viewed as more highly evolved. Africans, in particular, were seen as closer to the apes. These kinds of ideas received pseudoscientific support in the form of studies of cranial capacity and culturally biased intelligence tests (Gould 1981). The totalizing oppression and dehumanization of colonial domination is well captured in Memmi’s *The Colonizer and the Colonized* (1967).

European economic domination had many aspects, but a major feature was the exploitation of colonized workers. Unlike white labor, which was free (in the sense of unbound), colonial labor was typically subjected to various forms of coercion. As conquered peoples, colonized nations could be denied any political rights and were treated openly as beings whose sole purpose was to enhance white wealth. Throughout the colonial world, various forms of slavery, serfdom, forced migrant labor, indentured servitude, and contract labor were common.

Not only did European imperialists exploit colonized workers in their homelands, but they also moved many people around to other areas of the colonial world where they were needed. The most notorious instance was the African slave trade, under which Africans were brought in bondage to the Caribbean area and sections of North and South America. However, other examples

include the movement of contract workers from China and India all over the colonial world. Britain, as the chief imperialist power, moved Indians to southern Africa, Fiji, Trinidad, Mauritius, and other places to serve as laborers in remote areas of the British Empire. These movements created “internal colonies” (Blauner 1972), where workers of color were again subject to special coercion.

Even seemingly free immigrants of color have been subject to special constraints. For example, Chinese immigrants to the United States in the late nineteenth century were denied naturalization rights, in contrast to European immigrants, and as a result, were subjected to special legal disabilities. In the United States, Australia, Canada, and elsewhere, Chinese were singled out for “exclusion” legislation, limiting their access as free immigrants.

African slavery had the most profound effects on the shaping of racial thought and racial oppression. Even though slavery contradicted the basic premise of capitalism as based on a free labor market, it nevertheless flourished within world capitalism, and was an essential feature of it. In a seminal book, Williams (1944; 1966) argued that capitalism could not have developed without slavery, a position elaborated upon by Blackburn (1997). The coerced labor of African slaves enabled the western European nations to accumulate capital and import cheap raw materials that served as a basis for industrialization.

### WORKING CLASS DIVISIONS

Within the United States, the coexistence of free labor in the North and slavery in the South, proved to be disastrous, drawing an especially harsh race line between blacks and whites. The very concept of whiteness became associated with the notion of freedom and free labor, while blacks were seen as naturally servile (Roediger 1991). White workers divided themselves from blacks (and other racially defined workers), believing that capitalists could use coerced and politically disabled workers to undermine their interests. Thus a deep division emerged in the working class, along racial lines. The racism of the white working class can be seen as a secondary phenomenon, arising from the ability of capitalists to engage in the super-exploitation of workers of color.

The sections of the world with the worst racial conflicts are the “white settler colonies.” In the British Empire, these include the United States, South Africa, Zimbabwe (Rhodesia), Canada, Australia, and New Zealand. These societies established large white working classes that came into conflict with colonial capitalists over the use of coerced labor (Harris 1964).

From the point of view of workers of color, the distinction between white property owners and white workers seems minimal. Although class conflict raged within the white community, people of color experienced the effects as a uniform system of white domination. All whites appeared to benefit from racism, and all whites appeared to collude in maintaining segregation, job discrimination, and the disenfranchisement of people of color. In this sense, race appears to override class. Nevertheless, it should be recognized that people of color were and are exploited as *labor*, in order to enhance capitalist profits. Thus their relationship to capital includes both race and class elements. Racial oppression intensifies their class oppression as workers.

### MIDDLE CLASSES

So far we have talked only of the relations between white capital, white labor, and colonized labor. The colonial world was, of course, more complex than this. Not only did colonized people have their own middle or upper classes, but sometimes outside peoples immigrated or were brought in and served as indirect rulers of the colonized.

Middle strata from among the colonized peoples can play a dualistic role in the system. On the one hand, they can help the imperialists exploit more effectively. Examples include labor contractors, police, or small business owners who make use of ethnic ties to exploit members of their own group. In these types of situations, the dominant white group can benefit by having members of the colonized population help to control the workers primarily for the dominant whites while taking a cut of the surplus for themselves. On the other hand, middle strata can also be the leaders of nationalist movements to rid their people of the colonial yoke.

Outside middle strata, sometimes known as middleman minorities, can be invaluable to the

colonial ruling class. As strangers to the colonized, they have no ambivalence about the aspirations of the colonized for self-determination. They take their cut of profits while not seriously threatening to take over from the Europeans. Because middleman groups tend to serve as the chief interactors with the colonized, they often become a major butt of hostility, deflecting the hostility that would otherwise be directed at the colonial elite. Thus the class and race relations resulting from the development of European capitalism and imperialism have been complex and world-shaping.

### CONTEMPORARY RACE RELATIONS

Even though formal colonialism as outright political domination has been successfully challenged by national liberation movements, and even though the most oppressive forms of coerced labor have been legally banned in most of the world, neocolonialism and racial oppression continue in various guises.

For example, African Americans in the United States remain a relatively disenfranchised and impoverished population. Although illegal, racial discrimination persists in everyday practice, and racist ideology and attitudes pervade the society. Many whites continue to believe that blacks are innately inferior and object to social integration in the schools or through intermarriage. African Americans are almost totally absent from positions of power in any of the major political, economic, and social institutions of the society. Meanwhile, they suffer from every imaginable social deprivation in such areas as housing, health care, and education.

The capitalist system maintains racism in part because racially oppressed populations are profitable. Racial oppression is a mechanism for obtaining cheap labor. It allows private owners of capital to reduce labor costs and increase their share of the surplus derived from social production. This is very evident in Southern California today, where the large, immigrant Latino population provides virtually all the hard labor at exceptionally low wages. Their political status as noncitizens, a typical feature of racist social systems, makes them especially vulnerable to the dehumanization of sweatshops and other forms of super-exploitation.

With an increasingly globalized world capitalism, these processes have taken an international dimension. Not only do capitalists take advantage of oppressed groups in their own nation-states, but they seek them out wherever in the world they can be found. Such people are, once again, of color. Of course, the rise of Japan as a major capitalist power has changed the complexion of the ruling capitalist elite, but the oppressed remain primarily African, Latin American, and Asian.

It is common today for people to assume that racism goes both ways and that everyone is equally racist, that African Americans have just as much animosity toward whites as whites have toward blacks. According to this thinking, whites should not be singled out for special blame because racism against those who are different is a universal human trait: We are all equally guilty of racism. This view denies the importance of the history described above. To the extent that peoples of color are antiwhite, it is a reaction to a long history of abuse. Claiming that the antiwhite sentiments of blacks are equally racist and on the same level as white racism is an attempt to negate the responsibility of Europeans and their descendents for a system of domination that has tried to crush many peoples.

At the foundation of the problem of race and class lies the value system of capitalism, which asserts that pursuit of self-interest in a competitive marketplace will lead to social enhancement for all and that therefore the social welfare need not be attended to directly. This assumption is patently untrue. The United States, perhaps the worst offender, has let this social philosophy run amok, resulting in the creation of a vast chasm between excessive wealth and grinding poverty, both heavily correlated with color. Without severe intervention in "free market" processes, the United States is heading toward increased racial polarization and even possible violence.

Movements for social change need to address racial oppression and disadvantage directly. Changing the system of capitalist exploitation will not eliminate racism, since the power and resources available to white workers are so much greater than those of workers of color. The whole system of inequality based on appropriation of surplus wealth by a few, mainly white, private property

owners needs to be challenged, along with its racial aspects. Major redistribution, based on racial disadvantage, would be required. Neither class-based nor racial inequality can be attacked alone. They are linked with each other and must be overthrown together.

## REFERENCES

- Blackburn, Robin 1997 *The Making of New World Slavery: From the Baroque to the Modern, 1492–1800*. London: Verso.
- Blauner, Robert 1972 *Racial Oppression in America*. New York: Harper and Row.
- Cox, Oliver Cromwell 1948 *Caste, Class, and Race*. New York: Modern Reader.
- Davis, Alliso, Burleigh B. Gardner, and Mary R. Gardner 1941 *Deep South*. Chicago: University of Chicago Press.
- Gould, Stephen Jay 1981 *The Mismeasure of Man*. New York: W.W. Norton.
- Harris, Marvin 1964 *Patterns of Race in the Americas*. New York: Walker.
- Lenin, V. I. 1939 *Imperialism: The Highest Stage of Capitalism*. New York: International.
- Memmi, Albert 1967 *The Colonizer and the Colonized*. Boston: Beacon Press.
- Oliver, Melvin L., and Thomas M. Shapiro 1985 *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge.
- Omi, Michael, and Howard Winant 1986 *Racial Formation in the United States*. New York: Routledge and Kegan Paul.
- Roediger, David R. 1991 *The Wages of Whiteness: Race and the Making of the American Working Class*. London: Verso.
- Williams, Eric 1966 *Capitalism and Slavery*. New York: Capricorn.
- Wilson, William J. 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- 1980 *The Declining Significance of Race: Blacks and Changing American Institutions*. Chicago: University of Chicago Press.

EDNA BONACICH

## CLASSIFICATION

See Tabular Analysis; Typologies.

## CLINICAL SOCIOLOGY

Clinical sociology is a humanistic, multidisciplinary specialization that seeks to improve the quality of people's lives. Clinical sociologists assess situations and reduce problems through analysis and intervention. Clinical analysis is the critical assessment of beliefs, policies, and/or practices with an interest in improving a situation. Intervention, the creation of new systems as well as the change of existing systems, is based on continuing analysis.

Clinical sociologists have different areas of expertise—such as health promotion, sustainable communities, social conflict, or cultural competence—and work in many capacities. They are, for example, community organizers, sociotherapists, mediators, focus group facilitators, social policy implementers, action researchers, and administrators. Many clinical sociologists are full-time or part-time university professors, and these clinical sociologists undertake intervention work in addition to their teaching and research.

The role of the clinical sociologist can be at one or more levels of focus from the individual to the intersocietal. Even though the clinical sociologist specializes in one or two levels of intervention (e.g., marriage counseling, community consulting), the practitioner will move among a number of levels (e.g., individual, organization, community) in order to analyze or intervene or both.

## THE HISTORY OF AMERICAN CLINICAL SOCIOLOGY

When sociology emerged as a discipline in the 1890s, the nation was struggling with issues of democracy and social justice. There was rural and urban poverty, women were still without the vote, and there were lynchings. Farmers and workers in the late 1800s were frustrated because they could see the centralization of economic and political power in the hands of limited groups of people. This kind of frustration led to public protests and the development of reform organizations. In this climate, it is not surprising that many of the early sociologists were scholar-practitioners interested in reducing or solving the pressing social problems that confronted their communities.

**The First University Courses** . While many of the early sociologists were interested in practice,

the earliest known proposal using the words “clinical sociology” was by Milton C. Winternitz (1885–1959), a physician who was dean of the Yale School of Medicine from 1920 through 1935. At least as early as 1929, Winternitz began developing a plan to establish a department of clinical sociology within Yale’s medical school. Winternitz wanted each medical student to have a chance to analyze cases based on a medical specialty as well as a specialty in clinical sociology.

Winternitz vigorously sought financial support for his proposal from the Rosenwald Fund, but he was unable to obtain the necessary funds for a department of clinical sociology. He did note, however, the success of a course in the medical school’s section on public health that was based on the clinical sociology plan.

The first course using the words “clinical sociology” in the title was taught by Ernest W. Burgess (1886–1966) at the University of Chicago. Burgess taught the course in 1928 and then offered it twice in 1929. During these years, the course was considered to be a “special” course and did not appear in the university’s catalog. Burgess offered the clinical sociology course, as a regular course, five times from 1931 through 1933. The course continued to be listed in the catalog for the next several years but was not taught after 1933.

The University of Chicago catalogs did not include a description of the clinical sociology course, but the course always was listed under the social pathology grouping. All courses in this section dealt with topics such as criminality, punishment, criminal law, organized crime, and personal disorganization. Several of the students enrolled in these first clinical sociology courses were placed in child guidance clinics. Clarence E. Glick, for instance, was the staff sociologist at Chicago’s Lower North Side Child Guidance Clinic and Leonard Cottrell was the clinical sociologist at the South Side Child Guidance Clinic.

Two other universities offered clinical courses in the 1930s—Tulane University in Louisiana and New York University. The Tulane University course was designed to give students the opportunity to learn about behavior problems and social therapy by conferences and fieldwork in a child guidance clinic. Louis Wirth (1897–1952), a full-time faculty member and director of the New Orleans Child Guidance Clinic, was scheduled to teach the course

in the spring of 1930. Wirth was unable to teach the course because he accepted a one-year Social Science Research Council Fellowship to work in Europe. The course was taught in his absence, but the university’s course information does not identify the professor who took Wirth’s place.

When Wirth returned to the United States in 1931, he joined the faculty of the University of Chicago. In the spring of 1932 he taught a “minor” course in clinical sociology but by then he no longer was working with child guidance clinics.

New York University also offered clinical sociology courses in the early 1930s. Harvey Warren Zorbaugh (1896–1965) was a faculty member there in the School of Education which provided undergraduate and graduate preparation for visiting teachers, educational counselors, clinicians, social workers, and school guidance administrators. The major focus of the program was the solution of educational problems and other social dilemmas.

Zorbaugh, along with Agnes Conklin, offered “Seminar in Clinical Practice” in 1930. The course was intended to qualify students as counselors or advisers to deal with behavioral difficulties in schools. From 1931 through 1933 the clinical practice course was called “Seminar in Clinical Sociology.” The course was one of the highest numbered courses in educational sociology and was offered both terms of each year. The course was open to graduate students who were writing theses or engaged in research projects in the fields of educational guidance and social work.

Zorbaugh, author of *The Gold Coast and the Slum: A Sociological Study of Chicago’s Near North Side*, had been involved with clinics at least since 1924. That was the year Zorbaugh and Clifford Shaw organized two sociological clinics in Chicago—the Lower North and South Side Child Guidance Clinics. Zorbaugh was associate director of the Lower North Side Child Guidance Clinic in 1925.

Zorbaugh was a founder, in 1928, of the Clinic for the Social Adjustment of the Gifted at New York University. He was director of this clinic at its inception and was actively involved in its work for over fifteen years. The clinic was for intellectually gifted and talented preadolescents. The clinic gave graduate students the opportunity to have supervised experiences in teaching, clinical diagnosing and treating of children with behavioral problems.



During the 1953–54 academic year, Alvin W. Gouldner (1920–1980) was teaching in the Department of Sociology and Anthropology at Antioch College in Ohio. Before joining the faculty, Gouldner had been a university teacher for four years and then worked, for one year, as a consultant to Standard Oil of New Jersey.

Gouldner offered “Foundations of Clinical Sociology” at Antioch. The course was taught at the highest undergraduate level, and students who enrolled in the course were expected to have completed the department’s course in social pathology. The college bulletin provided the following description of the course:

*A sociological counterpart to clinical psychology with the group as the unit of diagnosis and therapy. Emphasis on developing skills useful in the diagnosis and therapy of group tensions. Principles of functional analysis, group dynamics, and organizational and small group analysis examined and applied to case histories. Representative research in the area assessed.*

**The term “clinical sociology” first appears in print**. The first known published linking of the words *clinical* and *sociology* was in 1930 when Milton C. Winternitz, a pathologist and dean of the Yale Medical School, wanted to establish a department of clinical sociology. After working on the idea at least as early as 1929, he wrote about it in a report to the president of the Yale Medical School and the report was published in the 1930 Yale University *Bulletin*. That same year saw the publication of a speech Winternitz had given at the dedication of the University of Chicago’s new social science building. The speech also mentioned clinical sociology.

Abraham Flexner, a prominent critic of medical education and director of the Institute for Advanced Study at Princeton, mentioned clinical sociology in 1930 in his *Universities: American, English, German*. Flexner did not approve of the Institute of Human Relations that Winternitz was establishing at Yale. In the pages of criticism devoted to the institute, Flexner briefly mentioned clinical sociology: “Only one apparent novelty is proposed: a professor of clinical sociology” (Flexner 1930).

Winternitz continued to write about the value of clinical sociology until 1936 when his last report

as dean was filed. One of Winternitz’s (1932) most forceful statements in support of the field was the contemporary-sounding statement that appeared in his 1930–1931 annual report:

*The field for clinical sociology does not seem by any means to be confined to medicine. Within the year it has become more and more evident that a similar development may well be the means of bringing about aid so sorely needed to change the basis of court action in relation to crime. . .*

*Not only in medicine and in law, but probably in many other fields of activity, the broad preparation of the clinical sociologist is essential. . .*

The first discussion of clinical sociology by a sociologist was Louis Wirth’s 1931 article, “Clinical Sociology,” in *The American Journal of Sociology*. Wirth wrote at length about the possibility of sociologists working in child development clinics, though he did not specifically mention his own clinical work in New Orleans. Wirth wrote “it may not be an exaggeration of the facts to speak of the genesis of a new division of sociology in the form of clinical sociology” (Wirth 1930).

In 1931, Wirth also wrote a career development pamphlet, which stated:

*The various activities that have grown up around child-guidance clinics, penal and correctional institutions, the courts, police systems, and similar facilities designed to deal with problems of misconduct have increasingly turned to sociologists to become members of their professional staffs (Wirth 1931).*

Wirth “urged (sociology students) to become specialists in one of the major divisions of sociology, such as social psychology, urban sociology. . . or clinical sociology” (Wirth 1931).

In 1931, Saul Alinsky was a University of Chicago student who was enrolled in Burgess’s clinical sociology course. Three years later, Alinsky’s article, “A Sociological Technique in Clinical Criminology,” appeared in the *Proceedings of the Sixty-Fourth Annual Congress of the American Prison Association*. Alinsky, best known now for his work in community organizing, was, in 1934, a staff sociologist and member of the classification board of the Illinois State Penitentiary.

In 1944 the first formal definition of clinical sociology appeared in H.P. Fairchild's *Dictionary of Sociology*. Alfred McClung Lee, the author of that definition, was known as one of the founders of the Society for the Study of Social Problems, the Association for Humanist Sociology, and the Sociological Practice Association. Lee later used the word "clinical" in the title of two articles—his 1945 "Analysis of Propaganda: A Clinical Summary" and the 1955 article "The Clinical Study of Society."

Also appearing in 1944 was Edward McDonagh's "An Approach to Clinical Sociology." McDonagh had read Lee's definition of clinical sociology but had not seen Wirth's 1931 article. McDonagh, in his *Sociology and Social Research* article, proposed establishing social research clinics that had "a group way of studying and solving problems" (McDonagh 1944).

In 1946 George Edmund Haynes's "Clinical Methods in Interracial and Intercultural Relations" appeared in *The Journal of Educational Sociology*. Haynes was a cofounder of the National Urban League (1910) and the first African American to hold a U.S. government subcabinet post. His 1946 article, written while he was executive secretary of the Department of Race Relations at the Federal Council of the Churches of Christ in America, discussed the department's urban clinics. The clinics were designed to deal with interracial tensions and conflicts by developing limited, concrete programs of action.

**Contemporary contributions.** While publications mentioning clinical sociology appeared at least every few years after the 1930s, the number of publications increased substantially after the founding of the Clinical Sociology Association in 1978. The association, now called the Sociological Practice Association, made publications a high priority. Individuals were encouraged to publish and identify their work as clinical sociology, and the association established publication possibilities for its members. The *Clinical Sociology Review* and the theme journal *Sociological Practice* were published by the association beginning in the early 1980s. These annual journals were replaced in the 1990's by *Sociological Practice: A Journal of Clinical and Applied Sociology*, a quarterly publication.

The Sociological Practice Association has had a central role in the development of American

clinical sociology. The association helped make available the world's most extensive collection of teaching, research, and intervention literature under the label of clinical sociology and it introduced the only clinical sociology certification process.

The Sociological Practice Association's rigorous certification process for clinical sociologists is available at the Ph.D. and M.A. levels. The Ph.D.-level process was adopted in 1983 and certification was first awarded in 1984. The association began to offer M.A.-level certification in 1986. Successful candidates at both the doctoral and master's level are awarded the same designation—C.C.S. (Certified Clinical Sociologist).

Experienced clinical sociologists are encouraged to apply for certification, which is given for intervention work (assessing and changing social systems). As part of the application process, a candidate is required to identify her or his area of specialization (e.g., community, family counseling) and level of intervention (e.g., organization, individual). The certification process requires membership in the Sociological Practice Association, documentation of appropriate education and supervised training, documentation of interdisciplinary training, essays about ethics and theory, and a demonstration before peers and a reviewing committee.

The Sociological Practice Association, along with the Society for Applied Sociology, also has put in place a Commission on Applied and Clinical Sociology. The commission has set standards for the accreditation of clinical and applied sociology programs at the baccalaureate level and intends to do the same for graduate programs.

## CLINICAL SOCIOLOGY AND SOCIOLOGICAL PRACTICE

The practical sociology of the 1890s and early 1900s is now referred to as sociological practice. This general term *sociological practice* involves two areas, clinical sociology and applied sociology. Clinical sociology emphasizes hands-on intervention while applied sociology emphasizes research for practical purposes. Both specialties require different kinds of specialized training.

Some sociological practitioners are "clinical" in that they only or primarily do intervention

work; others are “applied” in that they only or primarily conduct research that is of practical interest. Some practitioners do both. Clinical sociologists, for instance, may conduct research before beginning an intervention project to assess the existing state of affairs, during an intervention (e.g., to study the process of adaptation), and/or after the completion of the intervention to evaluate the outcome of that intervention. For some clinical sociologists, the research activity is an important part of their own clinical work. These sociologists have appropriate research training and look for opportunities to conduct research. Other clinical sociologists prefer to concentrate on the interventions and leave any research to other team members. Those clinical sociologists who decide not to engage in research may have research skills but prefer to conduct interventions, may not have enough expertise in the conduct of research, or may know that other team members have more expertise in research.

#### **THEORIES, METHODS, AND INTERVENTION STRATEGIES**

Clinical sociologists are expected to have education and training in at least one area in addition to sociology. This means that not only are clinical sociologists exposed to the range of theories (e.g., symbolic interaction, structural-functionalism, conflict, social exchange) and quantitative and qualitative research methods generally taught in sociology programs, but they also have additional influences from outside of their own programs. The result is that clinical sociologists integrate and use a broad range of theoretical and methodological approaches.

Clinical sociologists use existing theory to formulate models that will be helpful in identifying and understanding problems and also to identify strategies to reduce or solve these problems. Clinical sociologists also have shown that practice can have an influence on existing theories and help in the development of new ones.

While clinical sociologists use a wide variety of research methods and techniques (e.g., participatory action research, geographic information systems, focus group analysis, surveys), they probably are best known for their case studies. Case studies involve systematically assembling and analyzing

detailed, in-depth information about a person, place, event, or group. This methodological approach involves many data-gathering techniques such as document analysis, life histories, in-depth interviews, and participant observation.

Clinical sociologists who have been in the field for ten or twenty years probably learned about intervention strategies primarily through courses and workshops given outside of sociology departments as well as through their work and community experiences. Clinical sociologists who have more recently entered the field also may have learned intervention techniques as part of their sociology programs. These sociology programs might include courses, for instance, on focus groups, mediation, or administration, as well as require supervised residencies or internships.

#### **CLINICAL SOCIOLOGY IN INTERNATIONAL SETTINGS**

Clinical sociology is as old as the field of sociology and its roots are found in many parts of the world. The clinical sociology specialization, for instance, often is traced back to the fourteenth-century work of the Arab scholar and statesperson Abd-al-Rahman ibn Khaldun (1332–1406). Ibn Khaldun provided numerous clinical observations based on his varied work experiences such as Secretary of State to the rule of Morocco and Chief Judge of Egypt.

Auguste Comte (1798–1857) and Emile Durkheim (1858–1917) are among those whose work frequently is mentioned as precursors to the field. Comte, the French scholar who coined the term “sociology”, believed that the scientific study of societies would provide the basis for social action. Emile Durkheim’s work on the relation between levels of influence (e.g., social compared to individual factors) led Alvin Gouldner (1965, p.19) to write that “more than any other classical sociologist (he) used a clinical model.”

Interest in clinical sociology has been growing in a number of countries. For example, French is the predominant language of many, if not most, of the current international clinical sociology conferences, and books and articles have appeared with clinical sociology in the title in France and French-speaking Canada. The French-language clinical

sociologists emphasize clinical analysis. They have a solid international network and have done an excellent job of attracting nonsociologists to that network. Their literature is substantial. Particularly notable is the work of Jacques van Bockstaele and Maria van Bockstaele; Robert Sevigny, Eugene Enriquez, Vincent de Gaulejac, and Jacques Rheaume.

Beginning in the mid-1990's, Italians hosted clinical sociology conferences, published clinical sociology books and articles and ran numerous clinical sociology training workshops. If one is interested in learning about clinical sociology in Italy, one would want to review the work of Michelina Tosi, Francesco Battisti, and Lucio Luisson. Luisson's 1998 book, *Introduzione alla Sociologia clinica* (Introduction to Clinical Sociology), contains thirteen articles written by Americans. One is an original article written for the volume but all the others are translations of articles that appeared in the Sociological Practice Association's *Clinical Sociology Review* or *Sociological Practice*. The volume concludes with the Sociological Practice Association's code of ethics.

Clinical sociology also is found in other parts of the world. Of particular interest would be developments in Greece, Brazil, Mexico, Uruguay, and South Africa. In South Africa, for instance, one university's sociology department has put a sociological clinic in place and another sociology department has developed a graduate specialization in counseling.

The international development of clinical sociology has been supported primarily by two organizations. The clinical sociology division of the International Sociological Association (ISA) was organized in 1982 at the ISA World Congress in Mexico City. The other major influence is the clinical sociology section of the Association internationale des Sociologues de Langue Francaise (International Association of French Language Sociologists).

It is clear that a global clinical sociology is beginning to emerge. American clinical sociology had a strong role in the early development of the global specialization but now it is only one of many influences. It will be interesting to see if the thrust of the international field will be as explicitly humanistic and intervention-oriented as American clinical sociology.

## REFERENCES

- Bruhn, John G., and Howard M. Rebach 1996 *Clinical Sociology: An Agenda for Action*. New York: Plenum.
- Clark, Elizabeth J., Jan Marie Fritz, and P.P. Rieker (eds.) 1990 *Clinical Sociological Perspectives on Illness & Loss: The Linkage of Theory and Practice*. Philadelphia: The Charles Press.
- Enriquez, Eugene 1997 "The Clinical Approach: Genesis and Development in Western Europe." *International Sociology*. 12/2(June):151-164.
- 1992 "Remarques Terminales Vers une Sociologie Clinique d'Inspiration Psychanalytique." *L'Organisation en Analyse*. Fevrier. Paris: P.U.F.
- , Gilles Houle, Jacques Rheaume, and Robert Sevigny (eds.) 1993 *L'Analyse Clinique dans les Sciences Humaines*. Montreal: Editions Saint-Martin.
- Fritz, Jan Marie (ed.) 1996 *The Clinical Sociology Resource Book*. 4th ed. Washington, D.C.: American Sociological Association Teaching Resources Center and the Sociological Practice Association.
- Fritz, Jan Marie 1991a "The History of American Clinical Sociology: The First Courses." *Clinical Sociology Review* 9:5-26.
- Fritz, Jan Marie 1991b "The Emergence of American Clinical Sociology." Pp. 17-32 In H. Rebach and J. Bruhn, eds., *Handbook of Clinical Sociology*. New York: Plenum.
- Fritz, Jan Marie 1985 *The Clinical Sociology Handbook*. New York: Garland.
- Gaulejac, Vincent de and Shirley Roy (ed.), 1993 *Sociologies Cliniques*. Paris: Hommes et Perspectives.
- Giorgino, Enzo 1998 "Per un Ridefinizione del Lavoro Professionale in Sociologia." *Sociologia e Professione*. 29 (Marzo):8-23.
- Glassner, Barry, and Jonathan A. Freedman 1979 *Clinical Sociology*. New York: Longman.
- Gouldner, Alvin 1965 "Explorations in Applied Social Science." *Social Problems*. 3/3(January):169-181. Reprinted 1965 in Alvin Gouldner and S.M. Miller, eds., *Applied Sociology*. 5-22 New York: Free Press.
- Luisson, Lucio (ed.) 1998 *Introduzione alla Sociologia Clinica: Teorie, Metodi e Tecniche di Intervento*. Milano: FrancoAngeli.
- Rebach, Howard M., and John G. Bruhn (eds.), 1991 *Handbook of Clinical Sociology*. New York: Plenum
- Rheaume, Jacques 1997 "The Project of Clinical Sociology in Quebec." *International Sociology*. 12/2(June):165-174.

- Sevigny, Robert 1997 "The Clinical Approach in the Social Sciences." *International Sociology*. 12/2(June):135-150.
- Straus, Roger A. (ed.), 1999 *Using Sociology: An Introduction from the Applied and Clinical Perspectives*. Third edition. New York: General Hall.
- Straus, Roger A. (ed.), 1979 "Special Issue on Clinical Sociology." *The American Behavioral Scientist*. March/April.
- Tosi, Michelina, and Francesco Battisti (eds.) 1995 *Sociologia Clinica e Sistemi Socio-Sanitari: Dalle Premesse Epistemologiche Allo Studio di Casi e Interventi*. Milano: FrancoAngeli.
- van Bockstaele, Jacques, Maria van Bockstaele, Colette Barrot, and Cl. Magny 1963 "Travaux de Sociologie Clinique: Quelques Conditions d'une Intervention de Type Analytique en Sociologie." *L'Annee Sociologique*. Paris: Presses Universitaires de France.
- van Bockstaele, Jacques, Maria van Bockstaele, Colette Barrot, Jacques Malbos, and Pierrette Schein 1968 "Problemes de la Sociologie Clinique: Nouvelles Observations sur la Definition de la Socioanalyse." *L'Annee Sociologique*. Paris: Presses Universitaires de France.
- Winternitz, Milton Charles 1932 "Clinical Sociology." In Report of the dean of the School of Medicine, *Bulletin of Yale University*.
- Wirth, Louis 1931a "Clinical Sociology." *American Journal of Sociology*, 37:49-66.
- 1931b *Sociology: Vocations for those Interested in It*. Pamphlet. Vocational Guidance Series, No. 1. Chicago: University of Chicago. Louis Wirth Collection, University of Chicago, Department of Special Collections. Box LVI, Folder 6.

JAN MARIE FRITZ

## CLUSTER ANALYSIS

See Correlation and Regression Analysis; Factor Analysis.

## COALITIONS

Originally a word for union or fusion, the term *coalition* came in the eighteenth century to mean a temporary alliance of political parties. In modern social science, the meaning has broadened to include any combination of two or more social

actors formed for mutual advantage in contention with other actors in the same social system. In most contemporary theories of coalition formation, it is taken for granted that the principles governing coalition formation are not much affected by the size of the actors, who may be small children or large nations, but are significantly affected by the number of actors in the system. In the sociological and social-psychological literature, interest has focused on coalition formation in social systems containing three actors, commonly known as *triads*, and on the factors that influence the formation of coalitions in that configuration. Coalitions in triads have certain properties that are very useful in the analysis of power relationships in and among organizations. Moreover, tetrads, pentads, and higher-order social systems can be viewed for analytical purposes as clusters of linked triads. In the literature of political science, the principal topic has been the formation of electoral and legislative coalitions in multi-party and two-party systems.

The social science perspective on coalitions derives from two major sources: the formal sociology of Georg Simmel (1902) and the *n*-person game theory of John Von Neumann and Oskar Morgenstern (1944). Simmel had the fundamental insight that conflict and cooperation are opposite sides of the same coin so that no functioning social system can be free of internal conflicts or of internal coalitions. Simmel also proposed that the geometry of social relationships is independent of the size of the actors in a social system but heavily influenced by their number; that social systems are held together by internal differentiation; that relationships between superiors and subordinates are intrinsically ambivalent; that groups of three tend to develop coalitions of two against one; and that, in stable social systems, coalitions shift continually from one situation to another.

While the basic ideas are attributable to Simmel, the analytical framework for most of the empirical research on coalitions that has been undertaken so far is that of Von Neumann (and his collaborator Oskar Morgenstern). Any social interaction involving costs and rewards can be described as an *n*-person game. In two-person games, the problem for each player is to find a winning strategy, but in games with three or more players, the formation of a winning coalition is likely to be the major

strategic objective. The theory distinguishes between zero-sum games, in which one side loses whatever the other side gains, and non-zero-sum games with more complex payoff schedules. And it provides a mathematical argument for the equal division of gains among coalition partners, the gist of which is that any essential member of a winning coalition who is offered less than an equal share of the joint winnings can be induced to desert the coalition and join an adversary who offers more favorable terms. In the various experimental and real-life settings in which coalitions are studied, this solution has only limited application, but game theory continues to furnish the vocabulary of observation.

Some recent writers identify coalition theory as that branch of game theory involving zero-sum games with more than two players and game theory as a branch of rational choice theory (Wood and McLean 1995). The basic assumption of rational choice—by voters, lobbyists, legislators, and managers—has been vigorously attacked (see Green and Shapiro 1994) and as strongly defended (Nicholson 1992, among many others). The critics argue that rational choice theory is essentially self-contained; its elaborate intellectual apparatus does not provide a clear view of political action. The defenders say, in effect, that judgment should be withheld.

Meanwhile, game theory (and its coalition branch) have been developing new ideas, largely based on the key concept of equilibrium. Equilibrium in a game is that condition in which none of the players have incentives to deviate from their chosen strategies. It is called *Nash equilibrium*, after its formulator (Nash 1951), and has been extended to include two interesting varieties: subgame perfect equilibrium and Bayesian equilibrium. The former requires that rational players refrain from incredible threats. The latter replaces the players' initial knowledge about payoff schedules with a set of probabilistic statements, subject to change by additional information. Another interesting innovation is the concept of *nested games* (Tsebelis 1990), in which the apparent irrationality of players' moves in a given game is a rational consequence of their concurrent involvement in other games.

Modern empirical work on coalitions falls into two major categories: (1) experimental studies of

outcomes in games played by small groups—games that have been devised by the experimenter to test hypotheses about the choice of coalition partners and the division of coalition winnings under specified conditions, and (2) observational studies of coalitions in the real world. Stimulated by the publication of divergent theories of coalition formation (Mills 1953; Caplow 1956; Gamson 1961) in the *American Sociological Review*, coalition experiments became part of the standard repertory of social psychology in the 1960s and continue to be so to this day (Bottom, Eavey, and Miller 1996). A great deal has been learned about how the choice of coalition partners and the division of coalition winnings are affected by variations in game rules and player attributes. Much, although by no means all, of this work has focused on three-player games in which the players have unequal resources and any coalition is a winning coalition, the distribution of resources falling into one of three types: (1)  $A > B > C$ ,  $A < B + C$ ; (2)  $A = B$ ,  $B > C$ ,  $A < B + C$ ; and (3)  $A > B$ ,  $B = C$ ,  $A < B + C$ . With respect to the choice of coalition partners, the central question has been whether subjects will consistently choose the partner with whom they can form the minimum winning coalition, or the stronger partner, or the partner who offers the more favorable terms, or the partner who resembles themselves in attributes or ideology. The general finding is that each of these results can be produced with fair consistency by varying the rules of the experimental game. The division of winnings between coalition partners has attracted even more attention than the choice of partners. The question has been whether winnings will be divided on the principle of *equality*, as suggested by game theory; or of *parity*, proportionate to the contribution of each partner, as suggested by exchange theory; or at an intermediate ratio established by *bargaining*. Although many experimenters have claimed that one or the other of these principles is primary, their collective results seem to show that all three modes of division occur spontaneously and that subjects may be tilted one way or another by appropriate instructions. Additional nuances of coalition formation have been explored in games having more than three players, variable payoffs, or incomplete information. Non-zero-sum games and sequential games with continually changing weights have been particularly instructive. The findings readily lend themselves to mathematical

expression (Kahan and Rapoport 1984; Prasnika and Roth 1992).

The explicit application of coalition analysis to real-life situations began with William Riker's (1962) study of political coalitions in legislative bodies; he discerned a consistent preference for minimal winning coalitions and emphasized the pivotal role of weak factions. Theodore Caplow (1968) showed how the developing theory of coalitions in triads could be used to analyze conflict and competition in nuclear and extended families, organizational hierarchies, primate groups, revolutionary movements, international relations, and other contexts. The initial development of observational studies was relatively slow, compared with the proliferation of laboratory studies, but there were some notable achievements, particularly in family dynamics and international relations, where coalition models fit gracefully into earlier lines of investigation. Coalition theory was also applied, albeit in a more tentative way, to work groups, intra- and interorganizational relationships, litigation and criminal justice, class and ethnic conflict, and military strategy. However, the bulk of empirical research after 1980 was undertaken by political scientists and focused on international relations, with particular emphasis on nuclear deterrence (Powell 1990) and on the formation of legislative coalitions (Laver and Schofield 1990; Shepsle 1991; Krebbiel 1991; Cox and McCubbins 1993). Some investigators have shifted their focus from coalition formation to coalition breaking (Lupia and Strom 1995; Horowitz and Just 1995; Mershon 1996), which appears to follow a quite different dynamic. Economists have studied customs unions, trading blocs, and other forms of economic combination (Burbidge et al. 1995; Yi 1996). But with a few notable exceptions (e. g. Lemieux 1997), sociologists have tended to neglect the study of coalitions since the promising beginnings of the 1970s.

Whatever the field of application, the examination of coalitions, especially the simple coalition of two against one, provides a key to the social geometry of innumerable situations involving conflict, competition, and cooperation. In nearly every conflict, each of the contending parties seeks the support of relevant third parties, and the side that gains that support is likely to prevail. In very many competitive situations, the outcome is eventually decided by the formation of a winning coalition. And any system of cooperation that involves

a status order must rely on the routine formation of coalitions of superiors against subordinates and be able to counter coalitions of subordinates against superiors.

All of these situations are susceptible to coalitions of two against one, which tend to transform strength into weakness and weakness into strength. Under many conditions, in the first of the triads mentioned above ( $A > B > C$ ,  $A < B + C$ ), both A and B will prefer C as a coalition partner; his initial weakness ensures his inclusion in the winning coalition. When  $A > B$ ,  $B = C$ ,  $A < B + C$ , B and C will often prefer each other as coalition partners; A's initial strength ensures his exclusion from the winning coalition. When  $A = B$ ,  $A > C$ , C's initial weakness again makes him a likely winner. The first purpose of any hierarchy must be to restrain in one way or another the inherent tendency of subordinates to combine against superiors. Although force and ritual are often deployed for this purpose, the stability of complex status orders depends on certain interactive effects that appear in triads with overlapping membership, called *linked triads*. In such clusters, the choice of coalition partners in one triad influences the choices made in other triads. The natural rules that seem to govern the formation of coalitions in linked hierarchical triads are that a coalition adversary in one triad may not be chosen as a coalition partner in another triad, and that actors offered a choice between incompatible winning coalitions will choose the one in the higher-ranking triad. The net effect favors conservative coalitions of superiors against subordinates without entirely suppressing revolutionary coalitions of subordinates against superiors.

Cross-cutting the coalition preferences that arise from unequal distributions of power and resources are preferences based on affinity, compatibility, and prior experience with potential partners. These other bases of coalition formation are conspicuous in intimate groups such as the family, where same-sex coalitions alternate with same-generation coalitions.

The study of coalitions in nuclear families is particularly rewarding because the distribution of power in the triad of mother-father-child changes so dramatically as the child grows, and because same-sex coalitions are differently valued than

cross-sex coalitions. The initial distribution of power between husband and wife is always transformed by the arrival of children; most cultures encourage certain patterns, such as the Oedipus and Electra complexes dear to Freudians: coalitions of mother and son against father and of father and daughter against mother. Research on the contemporary American family suggests that parental coalitions are quite durable, both mother-daughter and mother-son coalitions against the father are very common, father-daughter coalitions against the mother much less so, and father-son coalitions against the mother comparatively rare. Sibling coalitions are most likely among same-sex siblings adjacent in age. Sibling aggression is endemic in families of this type, especially in the presence of parents. An interesting study by Richard Felson and Natalie Russo (1988) suggests that parents usually take side with the weaker child in these incidents, and this leads to more frequent aggression by the excluded child. There are very few family conflicts that cannot be instructively described by a coalition model.

The application of coalition theory to international relations was particularly rewarding with respect to the “strategic triangle” of the United States, China, and the Soviet Union during the Cold War era of 1950–1985. In one of the many studies that have examined the internal dynamics of this triad, James Hsiung (1987) concluded that China as the weak player in this triad benefitted much more than either of the superpowers from the various coalitional shifts that occurred over time, as would be theoretically expected in a triad of this type ( $A=B$ ,  $B>C$ ,  $A<B+C$ ). A study by Caplow (1989) explained the failure of peace planning in 1815, 1919, and 1945, by showing how efforts to put an end to the international war system were undermined by the formation of coalitions to prevent the domination of the peacekeeping organization by the strongest of the victorious powers. Many older studies of international balances of power visualize international relations as a game in which the first priority of every major player is to block the domination of the entire system by any other player. Frank C. Zagare’s (1984) analysis of the Geneva Conference on Vietnam in 1954 as a three-player game compared the preference schedules of the three players and showed how they combined to produce the unexpected outcome of the negotiations.

Both family dynamics and international relations in peacetime exemplify situations of continuous conflict, wherein relationships have long histories and are expected to persist indefinitely, and the opposition of interests is qualified by the necessity for cooperation. The choice of coalition partners and the division of winnings is strongly influenced by the past transactions of the parties and by the fact that payoffs are not completely predictable. Continuous conflict triads with  $A>B>C$ ,  $A<B+C$  often alternate the three possible coalitions according to circumstances: the *conservative coalition*  $AB$  reinforces the existing status order; the *revolutionary coalition*  $BC$  challenges it; and the *improper coalition*  $AC$  subverts it.

Episodic conflicts, by contrast, involve discrete zero-sum games played under strict rules. The passage of any measure in a legislative body necessarily involves the formation of a coalition. Even when one party has a solid majority, its members will seldom be in complete agreement on an issue. The formation of a coalition for the passage of a specific measure usually involves hard bargaining and payoffs negotiated in advance. Under these conditions, the tendency to minimize costs by forming the minimal winning coalitions is very strong. When  $A>B>C$ ,  $A<B+C$ , a  $BC$  coalition is highly probable. Empirical studies of legislative voting bear this out, although more than minimal coalitions also occur, for various reasons.

The resolution of disputes by civil and criminal litigation is another variety of episodic conflict that can be studied as a coalition process. Donald Black (1989) explored the triad of judge and courtroom adversaries and discovered a clear tendency for judges to favor the litigant to whom they are socially closer, ordinarily the litigant of higher status—a tacit conservative coalition. But in forms of dispute resolution where the third party is less authoritative, the weaker adversary may be favored. Marital counselors, for example, often side with wives against husbands, and ombudsmen and other relatively powerless mediators normally incline toward the weaker party.

In terminal conflicts, the object is the permanent destruction of adversaries, and the formation of coalitions is a delicate matter. In the triad where  $A>B>C$ ,  $A<B+C$ , a successful  $BC$  coalition that destroys  $A$  leaves  $C$  at the mercy of  $B$ . Indeed, any winning coalition is hazardous for the weaker



partner. A fragile peace can be maintained if  $A > B > C$  and  $A = B + C$ ; the BC coalition forms as a matter of course, creating what is known as a balance of power. This has been the key configuration in European affairs for the past several centuries. The balance breaks down with any significant shift in the relative power of the parties; for example, if A grows stronger than the BC coalition, it will be tempted to conquer them. If B becomes equal to A, an AB coalition may be tempted to attack and partition C. If C grows stronger and the triad assumes the form  $A > B$ ,  $B = C$ ,  $B + C > A$ , the formation of a BC coalition to overthrow A is likely. In the eighteenth century, the breakdown of a balance of power led to war without delay. Under current conditions, the breakdown of a balance of power among major industrialized states does not involve an automatic resort to arms, but in several regional arenas, such as the Middle East, the old mechanism is still intact.

Terminal conflicts occur also within nations as coups, resistance movements, and revolutions. One common pattern is the urban uprising against a dictatorial regime, in which the players are the government, the army, and the populace. If the army continues to support the government and is willing to fire on the populace, the uprising fails, as in China in 1989. If the army sides with the populace, the government is overthrown, as in Indonesia in 1998. Often the issue is undecided until the moment when the troops confront the demonstrators. At a more fundamental level, successful revolutions require a coalition of formerly separate factions against the ruling group.

Every organization generates both internal and boundary coalitions. Internal coalitions are activated whenever persons or groups of unequal status interact before witnesses. In general, the presence of a high-status witness reinforces the authority of a superior, while the presence of a low-status witness reduces it; examined in detail, these catalytic effects are delicate and precise. Boundary coalitions occur whenever one organization has permanent relations with another. Their respective agents must form a coalition with each other to perform their functions, and that coalition pits them both against their own colleagues, always with interesting consequences.

In a long-term perspective, the three bodies of coalition studies, theoretical, experimental, and

observational, have developed unevenly. The theories are elaborate and elegant. The experimental studies have explored nearly every possibility suggested by the theories, run down every lead, manipulated every variable. But in sociology, as distinct from political science and economics, the observational studies have scarcely tapped the rich possibilities suggested by the available theories. The most important work remains to be done.

(SEE ALSO: *Decision-Making Theory and Research*)

## REFERENCES

- Adams, Wesley J. 1985 "The Missing Triad: The Case of Two-Child Families." *Family Process* 24:409-413.
- Black, Donald 1989 *Sociological Justice*. New York: Oxford University Press.
- Bonacich, Phillip, Oscar Grusky, and Mark Peyrot 1985 "Family Coalitions: A New Approach and Method." *Social Psychological Quarterly* 44:42-50.
- Bottom, William P., Cheryl L. Eavey, and Gary J. Miller 1996 "Coalitional Integrity as a Constraint on the Power of Agenda Setters," *Journal of Conflict Resolution* 40:2:298-319.
- Burbidge, John B., James A. DePater, Gordon M. Myers, and Abhjit Sengupta 1997 "A Coalition-Formation Approach to Equilibrium Federations and Trading Blocs" *American Economic Review* 87:940-56.
- Caplow, Theodore 1956 "A Theory of Coalitions in the Triad." *American Sociological Review* 21:480-493.
- 1968 *Two Against One: Coalitions in Triads*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1989 *Peace Games*. Middletown, Conn.: Wesleyan University Press.
- Felsen, Richard B. and Natalie Russo 1988 "Parental Punishment and Sibling Aggression." *Social Psychology Quarterly* 51:1:11-18.
- Gamson, William A. 1961 "A Theory of Coalition Formation." *American Sociological Review* 26:565-573.
- Green, D.P. and I. Shapiro 1994 *Pathologies of Rational Choice Theory*. New Haven, Conn.: Yale University Press.
- Horowitz, John K. and Richard E. Just 1995 "Political Coalition Breaking and Sustainability of Policy Reform." *Journal of Development Economics* 47:271-286.
- Hsiung, James C. 1987 "Internal Dynamics in the Sino-Soviet-U.S. Triad," In I.J. Kim, ed., *The Strategic Triangle*. New York: Paragon.

- Kahan, James P., and Amnon Rapoport 1984 *Theories of Coalition Formation*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Komorita, S.S., and Charles E. Miller 1986 "Bargaining Strength as a Function of Coalition Alternatives." *Journal of Personality and Social Psychology* 51:325-332.
- Lemieux, Vincent 1997 "Réseaux et Coalitions." *L'Année Sociologique* 47:1:55-71.
- Lupia, Arthur, and Kaare Strom 1995 "Coalition Termination and the Strategic Timing of Parliamentary Elections." *American Political Science Review* 89:648-665.
- Mershon, Carol 1996 "The Costs of Coalition: Coalition Theories and Italian Governments." *American Political Science Review* 90:3:534-554.
- Miller, Charles E., and R. Crandall 1980 "Experimental Research on the Social Psychology of Bargaining and Coalition Formation." In R. Paulus, ed., *Psychology of Group Influence*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Mills, Theodore M. 1953 "Power Relations in Three-Person Groups." *American Sociological Review* 18:351-357.
- Murnighan, J. Keith 1978 "Models of Coalition Behavior: Game Theoretic, Social Psychological, and Political Perspectives." *Psychological Bulletin* 85:1130-1153.
- Nail, Paul, and Steven G. Cole 1985 "Three Theories of Coalition Behavior: A Probabilistic Extension." *British Journal of Social Psychology* 24:181-190.
- Nash, John 1951 "Non-cooperative games." *Annals of Mathematics* 54:286-95.
- Nicholson, Michael 1992 *Rationality and the Analysis of International Conflict*. Cambridge, U.K.: Cambridge University Press.
- Prasnikar, V., and A. E. Roth 1992 "Considerations of fairness and strategy: Experimental evidence from sequential games." *Quarterly Journal of Economics* 107:865-868.
- Riker, William H. 1962 *The Theory of Political Coalitions*. New Haven, CT: Yale University Press.
- Rodgers, Joseph Lee, and Vaida D. Thompson 1986 "Towards a General Framework of Family Structure: A Review of Theory-Based Empirical Research." *Population and Environment* 8:143-171.
- Rose, Irene Kathryn 1986 "Testing Coalition Theory in the Great Gatsby and the Rabbit Trilogy." Ph.D. diss., University of Oklahoma.
- Russo, Richard B., and Natalie Russo 1988 "Parental Punishment and Sibling Aggression." *Social Psychological Quarterly* 51:11-18.
- Sened, Itai 1996 "A Model of Coalition Formation: Theory and Evidence." *Journal of Politics* 58:2:350-372.
- Simmel, Georg 1902 "The Number of Members as Determining the Sociological Form of the Group." *American Journal of Sociology* 8:1-46; 158-196.
- Tsebilis, G. 1990 *Nested Games; Rational Choice in Comparative Politics*. Berkeley: University of California Press.
- Von Neumann, John and Oskar Morgenstern 1944 *Theory of Games and Economic Behavior*. 2d ed. Princeton, N.J.: Princeton University Press.
- Wood, Stewart, and Iain McLean 1995 "Recent Work in Game Theory and Coalition Theory." *Political Studies* 43:703-717.
- Yi, Sang-Seung 1996 "Endogenous Formation of Customs Unions under Imperfect Competition." *Journal of International Economics* 41:1-2; 153-177.
- Zagare, Frank C. 1984 *Game Theory: Concepts and Applications*. Beverly Hills, Calif.: Sage Publications.

THEODORE CAPLOW

## COGNITIVE CONSISTENCY THEORIES

Cognitive consistency theories have their origins in the principles of Gestalt psychology, which suggests that people seek to perceive the environment in ways that are simple and coherent (Köhler 1929). Cognitive consistency theories have their beginnings in a number of seemingly unrelated research areas (Eagly and Chaiken 1993). Early consistency theorists drew upon theories of conflict (Lewin 1935; Miller 1944), memory (Miller 1956), and the intolerance for ambiguity by those with an authoritarian personality (Adorno, Frenkel-Brunswick, Levinson, and Stanford 1950). According to Newcomb (1968a), social scientists should not have been surprised at the rise of cognitive consistency theories. He points to a truism that in any field of scientific inquiry, there is an inevitable movement from description of the elements of the field, to understanding the relationships between them. At the heart of cognitive consistency theories is the assumption that people are motivated to seek coherent attitudes, thoughts, beliefs, values, behaviors, and feelings. If these are inconsistent, they will produce a "tension state" in the individual, and motivate the individual to reduce this tension. Individuals reduce this tension, according

to consistency theories, by making their relevant cognitions consistent.

Cognitive consistency theories gained tremendous popularity in the social sciences in the 1950s, and generated hundreds of studies. Toward the end of the 1960s, however, research interest waned. In 1968, Abelson and colleagues published a massive handbook, entitled *Theories of Cognitive Consistency: A Sourcebook*. The book was a thorough chronicle of cognitive consistency theories, and it addressed these theories from virtually any angle the reader could imagine. Ironically, the scholarly detail in which the editors and authors carefully described their research seemed to have been the death knell of cognitive consistency theories. Virtually no research on cognitive consistency theories took place during the 1970s.

With the end of the 1960s, theories of behavior that centered around motivational and affective forces (which certainly described cognitive consistency theories) were “out of vogue” with researchers. Many point to the simultaneous rise of social cognition approaches in general, and attribution theory in particular as helping to divert interest from cognitive consistency theories. Research on cognitive consistency theories was supplanted by more complex (some believed) “cold cognition” approaches that strictly dealt with how cognitive processes work together, not accounting for “hot” forces such as feelings and motivation.

Theories of cognitive consistency theory did not die, they just went away for a while. Abelson (1983) noted the reemergence of cognitive consistency theories in the early 1980s with the observation that authors were beginning to write about social cognition theories in light of the renewed interest in the nature of affect (e.g., Fiske 1982; Hamilton 1981). Toward the end of the 1980s, researchers began to take a closer look at the influence of affect on cognitive processes (e.g., Forgas 1990; Isen 1987; Schwarz 1990). This change was precipitated by the development of several theoretical perspectives concerning the nature and structure of emotion (e.g., Frijda 1988; Ortony, Clore, and Collins 1988). In an influential article, Zanna and Rempel (1988) argued that attitudes toward different attitude objects may be more or less determined by affective, rather than cognitive, sources. Consistency theories were not only back, they were thriving (Harary 1983).

This chapter will discuss five major theories of cognitive consistency that have had the most impact on the behavioral sciences. They are (in no particular order) balance theory (Heider 1946, 1958), strain toward symmetry (Newcomb 1953, 1968b), congruency theory (Osgood and Tannenbaum 1955), affective-cognitive consistency model (Rosenberg 1956), and finally, what Eagly and Chaiken (1993) refer to as the “jewel in the consistency family crown,” (p. 456) cognitive dissonance theory (Festinger 1957).

### BALANCE THEORY

The earliest consistency theory is Heider’s balance theory (1946, 1958). This approach is concerned with an individual’s perceptions of the relationships between himself (*p*) and (typically) two other elements in a triadic structure. In Heider’s formulation, the other elements are often another person (*o*) and another object (e.g., an issue, object, a value). The attitudes in the structure are designated as either positive or negative. The goal of assessing the structure of a triad is to ascertain whether the relationships (attitudes) between the actors and the other elements are balanced, or consistent. According to Heider (1958), a balanced triad occurs when all the relationships are positive, or two are negative and one is positive (i.e., two people have a negative attitude toward an issue, but they like each other), and the elements in the triad fit together with no stress. Imbalance occurs when these outcomes are not achieved (i.e., all three relationships are negative, or you have a negative attitude toward an issue that your friend favors). Heider assumed that people prefer balanced states to imbalanced ones, because imbalance results in tension and feelings of unpleasantness. Balance, according to Heider, is rewarding.

Interestingly, imbalanced states can also be rewarding and exciting. Heider said that sometimes balance can be “boring” and that “The tension produced by unbalanced situations often has a pleasing effect on our thinking and aesthetic feelings” (1958, p. 180). In other words, imbalance stimulates us to think further, to solve the problem, to imagine, and to understand the mystery of the imbalance. According to balance theory, there are three ways to restore balance to an imbalanced triad: (1) one may change one’s attitude toward either the object or the other person, in order to

restore balance; (2) one might distort reality to perceive that the relationships are balanced (e.g., your friend doesn't really favor something you dislike, she really dislikes it); and (3) one might cognitively differentiate the relationship one has with a friend, so that the friend's opposing attitude toward something one favors is separated from one's positive attitude toward the friend as a person (e.g., you might compartmentalize a friend's opposite political views apart from your attitude toward her, in order to maintain your friendship and maintain balance, in most other areas where she is concerned) (Eagly and Chaiken 1993).

A limitation of Heider's balance theory is that it did not account for the strength of attitudes between persons and objects in the triad. It merely categorized the relationships as either positive or negative, and it therefore assumed that tension that is produced by imbalance was objectively of the same strength and effect on the individuals in the triad. Because some attitudes are held with more conviction and are more meaningful and important to us, it stands to reason that triads that involve imbalance with such strongly-held attitudes ought to evoke more tension (Eagly and Chaiken 1993). Another shortcoming of the theory is that it only deals with relationships between three entities. To address this latter concern, Cartwright and Harary (1956) published a paper that nicely generalized Heider's theory to account for structures of any size.

#### STRAIN TOWARD SYMMETRY MODEL

Newcomb (1953, 1968b) suggested that there are three, rather than two types of balance relationships in a triad. First, a structure that does not motivate modification (or acceptance) is termed a "nonbalanced" structure. These situations are characterized by indifference. Here, disagreement with another individual about an issue or object does not arouse tension if that other individual is devalued (or otherwise not important). However, when the other person is valued (i.e., a friend or significant other), then agreement with him or her about an object results in a "positively balanced" structure, while disagreement results in a "positively imbalanced" structure. The term "positive" in the latter two types of structure denotes the valence of the relationship between  $p$  and  $o$ , who is a valued other. The important focus in Newcomb's

approach is the relationship of  $o$  to  $p$ , and  $p$ 's view of  $o$  as a valued person, and "suitability as a source of information, or support, or of influence concerning the object" (Newcomb 1968b, p. 50). Newcomb's experiments supported his idea that the tension that is aroused when  $p$  and  $o$  have strong attitudes in the structure is much greater than when their attitudes are held with little conviction. He also found that positively balanced situations are the most preferred structures, followed by nonbalanced structures, with the positively imbalanced situations being the least preferred.

#### CONGRUENCY THEORY

A particular advantage of Osgood and Tannenbaum's (1955; Tannenbaum 1968) congruency theory is its precision in assessing: 1) the strength of the relationships between  $p$  and  $o$ , 2) the strength of the motivation to change an incongruent triad, and 3) the degree of attitude change that is necessary to balance a triad. Another advantage of this theory is that, like Newcomb's approach, it takes into account the strength of the attitudes of  $p$  and  $o$  in evaluating the degree of incongruity in the structure. Osgood and Tannenbaum discuss the Heider triad in terms of  $p$ , another individual, termed the source ( $s$ ) and  $s$ 's attitude (termed an "assertion") toward another object or concept ( $x$ ). According to the theory, attitudes can be quantified along a seven-unit evaluative scale, from extremely negative (-3) to neutral (0) to extremely positive (+3).

When  $p$ 's attitude toward  $s$  and  $x$  are positive, and  $s$ 's assertion is equally strong and of the same valence, there is a "congruous" structure to the triad. There is no motivation to change one's attitude toward the object or toward the source. When  $p$ 's attitude toward  $s$  is positive, and  $p$  has an equally positive attitude toward  $x$  that  $s$  later negatively evaluates, an incongruous structure is established. In this situation,  $p$  is motivated to change his or her attitude toward  $s$ , or  $x$ , or both, in the direction of congruity. Consider the following example. If  $p$ 's attitude toward  $s$  is a +2, and  $p$ 's attitude toward  $x$  is a -2, the structure would be congruent if  $s$ 's assessment is a -2. If, however, the assessment is a +2, the structure is imbalanced. In this case,  $p$ 's attitude toward either  $x$  or  $s$  needs to change four units to make the triad congruent. Of course, if the relationships are weaker, the degree

of attitude change to make the triad congruent is that much less (by the exact amount denoted in the quantitative calculation of all the relations of  $p$ ,  $s$ , and  $x$ ). Osgood and Tannenbaum also argued that strongly held attitudes would be less likely to be modified in incongruent triads. This was supported in subsequent research (Tannenbaum 1968).

#### AFFECTIVE-COGNITIVE CONSISTENCY MODEL

This approach suggests that people seek consistency in order to satisfy a general motivation toward simplicity in cognition, and/or to adhere to norms, traditions, customs, or values that reinforce consistency in one's cognitions and behavior (Rosenberg 1956, 1968). Another interesting twist on the consistency approach is that in the affective-cognitive consistency model, Rosenberg (1956, 1968) proposed that people are more motivated to maintain cognitive consistency so that *other people* perceive that they are consistent. In other words, while the individual may occasionally feel some tension as a result of inconsistency, other people find the inconsistency more aversive, because it represents a conflict for those around the individual. Specifically, if  $o$  has a positive attitude toward  $p$ , but  $p$  dislikes  $x$ , which  $o$  likes,  $o$  is caught between being friendly with, and avoiding,  $p$ . In this model,  $o$  feels tension at this conflict, and must reduce the tension by changing attitudes toward  $p$  (e.g., increasing attraction toward  $p$ , which would thereby outweigh any conflict with  $p$ 's negative attitude toward  $x$ ) or toward  $x$  (e.g.,  $o$  devalues  $x$ , so that  $p$ 's dislike of  $x$  does not result in  $o$  feeling conflicted).

Rosenberg's model also considers the relationship between the individual, his or her values, and an attitude object. For example, consider that  $p$  also has various other important values, denoted as  $y_1$ ,  $y_2$ ,  $y_3$ , etc. Rosenberg suggests that the  $p$ - $x$ - $y$  triad is just as important in understanding cognitive consistency as the traditional  $p$ - $o$ - $x$  triad. In the affective-cognitive consistency approach, we must consider  $p$ 's attitudes toward each of his or her values, how  $p$  feels about  $x$ , and  $p$ 's perception of the relationship between  $x$  and each of the values. When all or most of the  $p$ - $x$ - $y$  triads are consistent, the individual has achieved cognitive consistency. When most or all of the  $p$ - $x$ - $y$  triads are inconsistent, the individual experiences cognitive inconsistency.

The reason Rosenberg's approach is called the affective-cognitive consistency model is that it proposes that inconsistency results when one's feelings are inconsistent with one's beliefs. That is, when the way we think and feel about an object or person are at odds, we will modify one or both to make the attitude consistent. Thus, this model attempts to address consistency within one's own attitudes toward other people and objects, but also consistency in how one's value system relates to other people and objects. The model is also unique in suggesting that other people experience more tension as a result of one's own inconsistency. Generally speaking, the model has been supported by experiments (Rosenberg 1964), and is considered a very useful addition to the family of cognitive consistency theories.

#### COGNITIVE DISSONANCE THEORY

Of all the cognitive consistency theories, none has had more influence on researchers and subsequent theories than cognitive dissonance theory (Festinger 1957). A conservative estimate suggests that at least 1,000 articles have been published in which researchers present data bearing upon the theory and their own revisions of the theory (Cooper and Fazio 1984). Many agree with Jones's (1976) assessment that cognitive dissonance theory is "the most important development in social psychology to date" (p. x). Along the way, the theory has been hailed for its elegant simplicity, and its powerful range of utility (Collins 1992). It has also been criticized for its lack of specificity (Lord 1992; Schlenker 1992).

In formal terms, Festinger's theory states that two elements (behaviors or thoughts, or both) "...are in a dissonant relation if, considering these two alone, the obverse of one element would follow from the other" (1957, p. 13). Dissonance, then, refers to a negative arousal brought about by one's inconsistent thoughts or actions, or both. Essentially, this translates into the following assumptions. If one has opposing thoughts or behaviors, or both, this brings about an aversive state of tension, akin to a drive state like hunger or thirst. This tension motivates the individual to seek relief by eliminating the tension. The tension can be dissipated by changing: 1) either a thought or attitude to make it consonant with the opposing thought or behavior, or 2) one's behavior, to make

it consonant with the opposing behavior or thought. Because it is often much easier to change one's thoughts rather than one's behaviors, these are typically the elements that get modified by the person in dissonance reduction.

As an example, Festinger (1957) talked about the dissonance experienced by most smokers at some point in their lives. Smokers engage in behavior (smoking) that is harmful to their health. This is at odds with our desire to avoid harming ourselves. This arouses tension in the individual. The smoker could reduce it by changing his or her behavior (quit smoking) or changing the way he or she thinks about the smoking behavior. As mentioned above, changing behavior is often more difficult than changing cognitions, and, as most smokers will affirm, quitting smoking is certainly no exception to this axiom. In this instance, Festinger suggests, smokers eliminate their dissonance by changing their thoughts about smoking. They may: 1) disbelieve the validity of the health consequences of smoking, or distort the information about smoking by thinking that smoking is only harmful if you smoke so many packs a day, or if you inhale cigar smoke, etc., or more fatalistically, 2) convince themselves that "we all die of something, and I might as well die doing something I enjoy." All of these changes in thoughts eliminate the dissonance for the smoker.

It should be noted that Festinger was not talking about logical inconsistencies. There are certainly conditions under which people think and do logically inconsistent things, yet feel no dissonance, or they feel dissonance, yet are not in a situation where a logical inconsistency is present. Festinger recognized what has become a truism in psychology, that a person's reaction to a stimulus is not a function of the objective properties of the stimulus itself, but rather the individual's *construal*, or *perception of*, that stimulus. This explains why the presence or absence of logical inconsistencies may or may not be accompanied by dissonance in an individual. The most important and reliable way to predict a person's behavior in a dissonance situation is to understand how he or she construes the potential dissonance arousing thoughts or behaviors, or both.

**History.** Cognitive dissonance theory came onto the scene in the 1950s when reinforcement

theories of behavior were very dominant in virtually all areas of inquiry in psychology. According to reinforcement principles, behavior that is followed by a reward is more likely to be repeated. Behavior that is followed by a strong reward should be more likely to be learned and repeated than behavior followed by a weak (or no) reward. Reinforcement theory was such a simple yet very powerful principle that it seemed to explain virtually all behavior in any context. For that reason, it was extremely popular among behavioral scientists. An experiment by Festinger and Carlsmith (1959) showed that reinforcement theory was not the all-purpose theory it appeared to be. In their experiment, Festinger and Carlsmith (1959) had participants do boring tasks (i.e., turning pegs one-quarter turn on a cribbage board) for an hour. Participants randomly assigned to a control group were then given a short questionnaire in which they were asked to rate how much they enjoyed the task. In other conditions, the experimenter then told participants that his research assistant had not yet arrived for a different version of the experiment, and he asked the participant if he would do the research assistant's job of telling the next participant (in the hall, who was in reality a confederate) that he enjoyed the experiment tasks. This was, of course, a lie, because the tasks were boring. These participants were assigned to one of two conditions. Some were given \$1 to tell the lie, and others were given \$20 to tell the lie. After participants had told the lie and were leaving, the experimenter ran up to the participant, explaining that he forgot to have the participant complete the ratings of the attitudes toward the experiment tasks.

The experiment pitted reinforcement theory against the predictions made from dissonance theory. Reinforcement theory suggests that the participants who were given \$20 should find the tasks more rewarding (pleasant), and should have a more positive attitude toward the tasks (and the experiment) than those only given a weak (or no) reward. Dissonance theory suggests that those in the control condition would feel no dissonance because they did a boring task, and would rate the tasks as such on the questionnaire. However, counter to intuition (and reinforcement theory) those in the \$20 condition should feel little (or no) dissonance, because although they did boring tasks, and disliked the tasks, saying that the tasks were fun is not an inconsistent behavior if one has

adequate justification (\$20) for doing so. They could attribute their lying to the incentive, and they would not feel hypocritical. The \$1 participants experienced significant dissonance because they did boring tasks, but yet they said they thought the tasks were fun. As Festinger and Carlsmith predicted, the \$1 was an insufficient justification for the lie, so the dissonance remained unless the participants changed their attitudes toward the task, and convinced themselves (as shown in their ratings of the tasks) that maybe the tasks were *not* boring, and in fact, they rather enjoyed them! The results were precisely as predicted, and this paved the way for a flurry of research that tested the exciting, often dramatic, and counterintuitive predictions that arose from cognitive dissonance theory.

#### **Alternate Versions of Dissonance Theory.**

Very soon after the publication of Festinger's theory, research revealed that the theory might need to be revised somewhat, to account for more of the data that were being published, which didn't quite fit with the theory. In one notable revision, one of Festinger's protégés, Aronson (1969) posited that the theory would be strengthened if it stated that dissonance would be most clearly aroused when the self-concept of the person is engaged. In other words, dissonance is stronger and more clearly evoked when the way we think about ourselves is at odds with our cognitions or behavior. This modification was supported by much subsequent research (Aronson 1980). Less an alternate version and more of a theoretical competitor, Bem's (1967) self-perception theory was the first major theory that offered a plausible account of the dissonance data, and pointed to different causal mechanisms. Unlike cognitive dissonance theory, Bem's approach did not invoke reference to hypothetical motivational processes, but rather tried to account for the person's behavior in terms of the stimuli present in the individual's environment and his or her related behavior. Bem's theory proposed that attitudinal change in dissonance experiments happens not due to an aversive tension (or other motivation), but due to a person's perceptions of his or her own behavior. Specifically, Bem said that people infer their attitudes from their actions, in much the same way that observers of our behavior infer the nature of our attitudes from our behavior. Attitude change occurs when their most recent behavior is different from their previous attitudes.

This behavioral approach to dissonance phenomena recasts the Festinger and Carlsmith experiment in a very different light. In a replication of the Festinger and Carlsmith study, Bem asked participants to listen to a tape describing a person named Bob, who did some boring motor tasks. Control condition participants then were asked to assess Bob's attitude toward the tasks. Other subjects then learned that Bob was given \$1 or \$20 to say to the next participant that the motor tasks were fun. Participants then listened to a recording of Bob enthusiastically telling a subsequent woman participant how enjoyable the motor tasks were. Participants were then asked to evaluate Bob's attitude toward the motor tasks. Those who were told that Bob was given \$20 to tell the lie inferred that the only reason he told the lie was because he was paid a lot of money. They assumed that he didn't really have a positive attitude toward the motor tasks. Those who were told that Bob received \$1 didn't think Bob had a good reason for lying, so his behavior (lying) told participants that Bob must really feel positively about the motor tasks. Control condition participants inferred that Bob negatively evaluated the motor tasks. As can be seen, these results are virtually identical to those obtained in the Festinger and Carlsmith experiment. Thus, according to self-perception, the Festinger and Carlsmith participants inferred their attitudes toward the boring tasks based on their recent behavior.

Subsequent research on self-perception theory was aimed at testing the self-perception theory contention that no arousal exists as a result of the dissonance situation. Zanna and Cooper (1976) found that arousal did indeed accompany counterattitudinal advocacy, so it was apparent that self-perception did not apply to all dissonance situations. Fazio, Zanna, and Cooper (1977) suggested that dissonance accounted for attitude change when behavior is truly counterattitudinal, but that self-perception can account for situations where behavior is only mildly counterattitudinal. For most researchers, this seems to have settled the debate about the situations to which each theory may be applied (Abelson 1983).

A final major revision was proposed by Cooper and Fazio (1984). They suggested that dissonance does not result from mere cognitive inconsistency, but only is evoked when the person feels personally responsible for causing an aversive event.

The theory suggests, then, that aversive consequences are necessary for dissonance to occur. In subsequent experiments, however, Aronson and his colleagues (Aronson, Fried, and Stone 1991) induced participants to make an educational video advocating the use of condoms for safe sex, to be shown in high schools. Then, the participant was subsequently reminded of situations in which he or she had not used condoms in the past. According to Cooper and Fazio, there should be no dissonance because the participant had not produced an aversive event, but rather, a positive one (advocating safe sex in an educational video). According to Aronson, however, the participant should feel dissonance because he or she was saying one thing and doing another (acting like a hypocrite). Aronson predicted that if dissonance was aroused in those hypocrites, they should be more likely to (if given the opportunity) take more free pamphlets and condoms at the end of the study, as a way of regaining consistency with their advocated position (i.e. they were making an effort to change their behavior to be consistent with their advocated message). Results showed that this is precisely what occurred. Subsequent research has demonstrated strong support for Aronson and colleagues' (1991) contention that the production of aversive consequences is not necessary to create dissonance (Harmon-Jones, Brehm, Greenberg, Simon, and Nelson 1996).

#### CURRENT STATUS OF COGNITIVE CONSISTENCY THEORIES

In the 1990s, cognitive consistency theories experienced a rebirth, primarily through renewed interest in cognitive dissonance theory (Aronson 1992). With the renewed interest in motivation, and the interaction of cognition and affect, researchers are once again taking up the questions (and there are many) left unanswered by earlier cognitive dissonance researchers. For example, while research has shown that dissonance evokes (as Festinger theorized) psychological discomfort (in comparison to physiological arousal), researchers know little about what occurs between the onset of the psychological discomfort and the start of the discomfort reduction process (Elliot and Devine 1994). Some have called for an effort to more fully explicate the conditions under which dissonance occurs, and when it does not occur (Aronson 1992), and researchers are beginning to

do just that (Shultz and Lepper 1996). Cognitive consistency theories have been a cornerstone of psychology for over four decades, and while they receded into the background in the 1970s, they are experiencing a strong resurgence of empirical and theoretical interest.

#### REFERENCES

- Abelson, R. P. 1983 "Whatever Became of Consistency Theory?" *Personality and Social Psychology Bulletin* 9:37-54.
- , E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum (eds.) 1968 *Cognitive Consistency Theories: A Sourcebook*. Chicago, Ill.: Rand McNally.
- Adorno, T. W., E. Frenkel-Brunswick, D. J. Levinson, and R. N. Sanford 1950 *The Authoritarian Personality*. New York: Harper and Row.
- Aronson, E. 1969 "The Theory of Cognitive Dissonance: A Current Perspective." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, Vol. 4. New York: Academic Press.
- 1980 "Persuasion via Self-Justification: Large Commitments for Small Rewards." In L. Festinger, ed., *Retrospections on Social Psychology*. New York: Oxford University Press.
- 1992 "The Return of the Repressed: Dissonance Theory Makes a Comeback." *Psychological Inquiry* 3 (4):303-311.
- , C. Fried, and J. Stone 1991 "AIDS Prevention and Dissonance: A New Twist on an Old Theory." *American Journal of Public Health* 81:1636-1638.
- Bem, D. J. 1967 "Self-perception: An Alternative Interpretation of Cognitive Dissonance Phenomena." *Psychological Review* 74 (3):183-200.
- Cartwright, D., and F. Harary 1956 "Structural Balance: A Generalization of Heider's Theory." *Psychological Review* 63:277-293.
- Collins, B. 1992 "Texts and Subtexts." *Psychological Inquiry* 3 (4):315-320.
- Cooper, J., and R. H. Fazio 1984 "A New Look at Dissonance Theory." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, Vol. 17. New York: Academic Press.
- Eagly, A. H., and S. Chaiken 1993 *The Psychology of Attitudes*. New York: Harcourt, Brace, Jovanovich.
- Elliot, A. J., and P. G. Devine 1994 "On the Motivational Nature of Cognitive Dissonance: Dissonance as Psychological Discomfort." *Journal of Personality and Social Psychology* 67(3):382-394.



- Fazio, R. H., M. P. Zanna, and J. Cooper 1977 "Dissonance and Self-Perception: An Integrative View of Each Theory's Proper Domain of Application." *Journal of Experimental Social Psychology* 13:464-479.
- Festinger, L. 1957 *A Theory of Cognitive Dissonance*. Palo Alto, Calif.: Stanford University Press.
- , and M. Carlsmith 1959 "Cognitive Consequences of Forced Compliance." *Journal of Abnormal and Social Psychology* 58:203-210.
- Fiske, S. T. 1982 "Schema-Triggered Affect: Applications to Social Perception." In M. S. Clark and S. T. Fiske, eds., *Affect and Cognition*. Hillsdale, N.J.: Erlbaum.
- Forgas, J. P. 1990 "Affect and Social Judgments: An Introductory Review." In J. P. Forgas, ed., *Emotion and Social Judgments*. Oxford: Pergamon Press.
- Frijda, N. H. 1988 "The Laws of Emotion." *American Psychologist* 43:349-358.
- Hamilton, D. L. 1981 "Illusory Correlation as a Basis for Stereotyping." In D. L. Hamilton, ed., *Cognitive Processes in Stereotyping and Intergroup Behavior*. Hillsdale, N.J.: Erlbaum.
- Harary, F. 1983 "Consistency Theory is Alive and Well." *Personality and Social Psychology Bulletin* 9:60-64.
- Harman-Jones, E., J. Brehm, J. Greenberg, L. Simon, and D. E. Nelson 1996 "Evidence that the Production of Dissonance is Not Necessary to Create Cognitive Dissonance." *Journal of Personality and Social Psychology* 70:5-16.
- Heider, F. 1946 "Attitudes and Cognitive Organizations." *Journal of Psychology* 21:107-112.
- 1958 *The Psychology of Interpersonal Relations*. Hillsdale, N.J.: Erlbaum.
- Isen, A. M. 1987 "Positive Affect, Cognitive Processes, and Social Behavior." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, Vol. 20. New York: Academic Press.
- Köhler, W. 1929 *Gestalt Psychology*. New York: Liveright.
- Lewin, K. 1935 *A Dynamic Theory of Personality*. New York: McGraw-Hill.
- Lord, C. G. 1992 "Was Cognitive Dissonance Theory a Mistake?" *Psychological Inquiry* 3 (4):339-342.
- Miller, G. A. 1956 "The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63:81-97.
- Miller, N. E. 1944 "Experimental Studies of Conflict." In J. Hunt, ed., *Personality and the Behavioral Disorders*, Vol. 1. New York: Ronald Press.
- Newcomb, T. M. 1953 "An Approach to the Study of Communicative Acts." *Psychological Review* 60:393-404.
- 1968a "Introduction." In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum, eds., *Cognitive Consistency Theories: A Sourcebook*. Chicago, Ill.: Rand McNally.
- 1968b "Interpersonal Balance." In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum, eds., *Cognitive Consistency Theories: A Sourcebook*. Chicago, Ill.: Rand McNally.
- Ortony, A., G. L. Clore, and A. Collins 1988 *The Cognitive Structure of Emotions*. New York: Cambridge University Press.
- Osgood, C. E., and P. H. Tannenbaum 1955 "The Principle of Congruity in the Prediction of Attitude Change." *Psychological Review* 62:42-55.
- Rosenberg, M. J. 1956 "Cognitive Structure and Attitudinal Affect." *Journal of Abnormal and Social Psychology* 53:367-372.
- 1964 "A Structural Theory of Attitude Dynamics." In E. E. Sampson, ed., *Approaches, Contexts, and Problems of Social Psychology*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1968 "Hedonism, Inauthenticity, and Other Goads Toward Expansion of a Consistency Theory." In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum, eds., *Cognitive Consistency Theories: A Sourcebook*. Chicago, Ill.: Rand McNally.
- Schlenker, B. R. 1992 "Of Shape Shifters and Theories." *Psychological Inquiry* 3 (4):342-344.
- Schultz, T. R., and M. R. Lepper 1996 "Cognitive Dissonance Reduction as Constraint Satisfaction." *Psychological Review* 103 (2):219-240.
- Schwarz, N. 1990 "Feelings as Information: Informational and Motivational Functions of Affective States." In R. Sorrentino and E. T. Higgins, eds., *Handbook of Motivation and Cognition*, Vol. 2. New York: Guilford.
- Tannenbaum, P. H. 1968 "The Congruity Principle: Retrospective Reflections and Recent Research." In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum, eds., *Cognitive Consistency Theories: A Sourcebook*. Chicago, Ill.: Rand McNally.
- Zanna, M. P., and J. Cooper 1976 "Dissonance and the Attribution Process." In J. H. Harvey, W. J. Ickes, and R. F. Kidd, eds., *New Directions in Attribution Research*. Hillsdale, N.J.: Erlbaum.
- Zanna, M.P., and J. K. Rempel, 1988 "Attitudes: A New Look at an Old Concept." In D. Bar-Tal and A.

Kruglanski, eds., *The Social Psychology of Knowledge*. Cambridge, Eng.: Cambridge University Press.

TODD D. NELSON

## COGNITIVE DISSONANCE THEORY

See Cognitive Consistency Theories; Social Psychology.

## COHABITATION

See Alternative Life-Styles; Courtship.

## COHORT ANALYSIS

See Cohort Perspectives; Longitudinal Research; Quasi-Experimental Research Design.

## COHORT PERSPECTIVES

The birth cohort, or set of people born in approximately the same period of time, has a triple reference as an analytical tool in sociology: (1) to cohorts of people who are *aging* and succeeding each other in particular eras of history; (2) to the *age composition* of the population and its changes; and (3) to the interplay between cohorts of people and the age-differentiated *roles and structures* of society. Diverse sociological studies illustrate the use of these cohort perspectives (i.e., both theoretical and empirical approaches) to investigate varied aspects of aging and cohort succession, population composition, and the reciprocal relationships between cohorts and social structures.

### CONCEPTUAL FRAMEWORK

Figure 1 is a rough schematization of the major conceptual elements implicated in these interrelated cohort perspectives as they have relevance for sociology (for an overview, see Riley, Johnson, and Foner 1972; Riley, Foner and Riley 1999).

**Aging and Cohort Succession.** The *diagonal bars* represent cohorts of people born at particular

time periods who are aging from birth to death—that is, moving across time and upward with age. As they age, the people in each cohort are changing socially and psychologically as well as biologically; they are actively participating with other people; and they are accumulating knowledge, attitudes, and experiences. The series of diagonal bars (as in the selected cohorts A, B, and C) denote how successive cohorts of people are continually being born, grow older through different eras of time, and eventually die.

**Age Composition of the Population.** The *perpendicular* lines direct attention to the people simultaneously alive in the society at particular dates. A single cross-sectional slice through the many coexisting cohorts (as in 1990) demonstrates how people who differ in cohort membership also differ in age—they are stratified by age from the youngest at the bottom to the oldest at the top. Over time, while society moves through historical events and changes, this vertical line should be seen as moving across the space from one period to the next. At different time periods the people in particular age strata are no longer the same people; inevitably, they have been replaced by younger entrants from more recent cohorts with more recent life experiences.

**Cohorts and Social Structures.** Corresponding to the age strata in the population, the perpendicular lines also denote the age-related role opportunities and normative expectations available in the various social structures (e.g., in schools for the young, in work organizations for those in the middle years, in nursing homes for the old, in families for all ages, etc.). People and structures are interdependent: changes in one influence changes in the other. Yet the two are often out of alignment, causing problems for both individuals and society.

This three-fold heuristic schematization, though highly oversimplified, aids interpretation and design of sociological work that takes cohort perspectives into account. (For simplicity, the discussion is limited here to cohorts in the larger society, with entry into the system indexed by date of birth. Parallel conceptualization refers also to studies of cohorts entering other systems, such as hospitals, with entry indexed by date of admission, or the community of scientists, with entry indexed by date of the doctoral degree—e.g., Zuckerman and

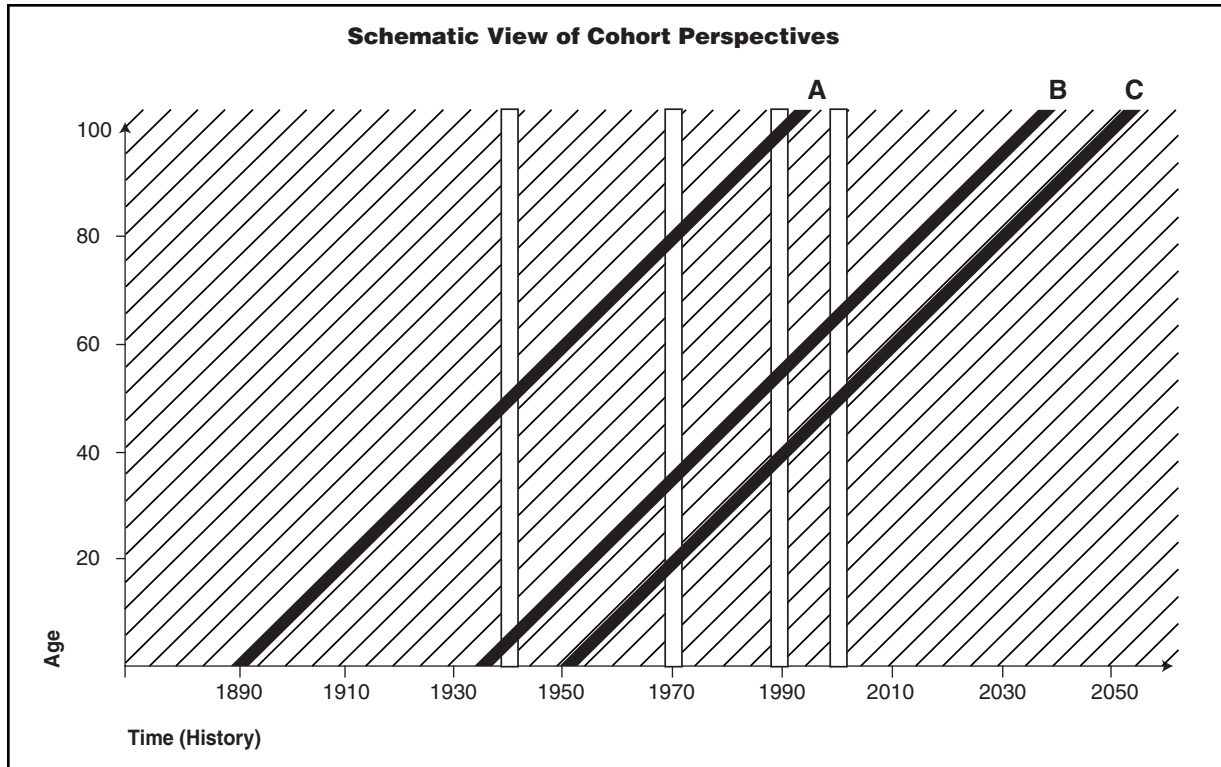


Figure 1

SOURCE: Riley, Foner, and Waring, 1998, p. 245 (adapted).

Merton 1972; here “aging” refers to duration in the particular system.)

### AGING PERSPECTIVES

Research on the processes of aging within and across cohorts illuminates the interrelated aspects of people’s lives and the particular characteristics and historical backgrounds of the cohorts to which they belong.

**Intracohort Perspectives.** Many empirical studies and much conceptual work uses the “life-course approach” to trace over time the lives of members of a single cohort (e.g., Clausen 1986). As one familiar example, studies of “status attainment” investigate lifelong trajectories of achievement behaviors, using longitudinal and causal modeling to examine the interconnections among such variables as family background, scholastic achievement, succession of jobs, and employment and unemployment (cf. Featherman 1981). The intracohort perspective is used in many forms—micro-

and macro-level, objective and subjective—in a range of studies on how people as they develop and grow older move through diverse paths in the changing society (Dannefer 1987), how sequences of role transitions are experienced, and how aging people relate to the changing environment. Psychologists as well as sociologists study interindividual changes in performance over the life course (e.g., Schaie 1996), while Nesselroade (1991) looks at intra-individual fluctuations over shorter periods of time.

Longitudinal studies of aging in a single cohort can contribute importantly to causal analysis by establishing the time order of correlated aspects of people’s lives and environmental events. However, this perspective is vulnerable to possible misinterpretation through the *fallacy of “cohort-centrism,”* that is, erroneously assuming that members of all cohorts will age in exactly the same fashion as members of the cohort under study (Riley 1978). Yet in fact, members of different cohorts, as they respond to different periods of

history, usually age in different ways. For example, the enjoyment of “midlife” experienced at around age 50 by cohort members studied in the 1990s may not be felt by some future cohort until age 85.

**Intercohort Perspectives.** Broader than the intracohort focus, is a focus on the lives of members of two or more successive cohorts who are growing older under differing historical or socio-cultural conditions. Studies of intercohort differences in the late-twentieth century demonstrated for other sciences what sociologists had learned early: the central principle that the process of aging is not immutable or fixed for all time, but varies across and within cohorts as society changes (Riley 1978). Such studies have shown that members of cohorts already old differ markedly from those in cohorts not yet old in such respects as standard of living, education, work history, age of menarche, experience with acute vs. chronic diseases, and perhaps most importantly the number of years they can expect to live. These cohort differences cannot be explained by evolutionary changes in the human genome, which remains much the same from cohort to cohort; instead, they result from a relatively unchanging genetic background combined with a continually changing society (Riley and Abeles 1990, p.iii). Thus the finding of cohort differences has pointed to possible linkages of lives with particular social or cultural changes over historical time, or with particular “period” events such as epidemics, wars, or depressions (e.g., Elder and Rockwell 1979). These linkages are useful in postulating explanations for changes—or absence of changes—in the process of aging.

Studies of cohort differences focus on aging processes at either the individual or the collective level. At the individual level, cohort membership is treated as a *contextual characteristic* of the individual, and then analyzed together with education, religion, and other personal characteristics to investigate how history and other factors affect the heterogeneous ways individuals grow older (e.g., Messeri 1988; but see Riley 1998). At the collective level, the lives of members are aggregated within each cohort to examine alterations in average patterns of aging. Striking advances have recently been made in the data banks available for intercohort comparisons. Archived data from many large-scale studies now cover long periods of history, multiple societies, and multidisciplinary aspects of the life

course; and repeated longitudinal studies are being launched, such as the National Institute on Aging’s Health and Retirement Study (HRS) and Asset and Health Dynamics Among the Oldest Old (AHEAD) (cf. Campbell 1994; O’Rand and Campbell 1999).

Cohort perspectives are useful, not only in explaining past changes in aging processes, but also in improving forecasts of future changes. Unlike the more usual straight projections of cross-sectional information, forecasts based on cohorts can be informed by established facts about the past lives of people in each of the cohorts already alive (e.g., Manton 1989). Thus, if cohorts of teenagers today are on the average less healthy, less cared for, or less prepared for life than their parents were at the same age (National Association of State Boards of Education 1990), the lives of both offspring and parents will predictably also differ in the future when both have grown older.

#### COMPOSITIONAL PERSPECTIVES.

Complementing sociological work on cohort differences (or similarities) in the aging process are studies of how cohort succession contributes to formation and change in the age composition of the population. Thus in Figure 1, the perpendicular lines indicate how cohorts of people fit together at given historical periods to form the cross-sectional *age strata* of society; and how, as society changes, new cohorts of people are continually aging and entering these strata, replacing the previous incumbents.

**Single Period of Time.** In Figure 1, as indicated above, a single vertical line at a given period (as in 1990) is a cross-sectional slice through all the coexisting cohorts, each with its unique size, composition, earlier life experiences, and historical background. This familiar cross-sectional view of all the age strata is often denigrated because its misinterpretation is the source of the *life-course fallacy*—that is, the erroneous assumption that cross-sectional age differences refer directly to the process of aging, hence disregarding the cohort differences that may also be implicated (Riley 1973). That people who are differentially located in the age composition of society differ not only in age but also in cohort membership was dramatized early by Mannheim ([1928] 1952) and Ryder

(1968); yet persistent failure to comprehend this duality has perpetuated numerous false stereotypes (e.g., that intelligence or physical functioning begin inevitable declines at very early ages).

Properly interpreted, of course, a cross-sectional perspective has its special uses: for describing current differences and similarities, social relationships, and interactions among coexisting people who differ in age-cum-cohort membership. Thus, for example, issues of “intergenerational equity” require explication by both age and cohort, as a larger share of the federal budget is reportedly spent on cohorts of people now old than on cohorts of children (Duncan, Hill, and Rodgers 1986; Preston 1984).

**Across Time.** Comprehension of the underlying dynamics of the age strata requires going beyond the single cross-sectional snapshot to a sequence of cross-sections (the moving perpendicular line in Figure 1), as successive cohorts interact with historical trends in the society (Ryder 1965; Riley 1982). Historical change means not only that new cohorts are continually entering the system through birth or immigration, while others are leaving it through death or emigration (e.g., men tend to die earlier than women, and blacks earlier than whites). Historical change also means that the members of *all* existing cohorts are simultaneously aging and thus moving from younger to older strata. As successive cohorts move concurrently through the system, they affect the age strata in several ways. They can alter the numbers and kinds of people in particular strata, as each cohort starts the life course with a characteristic size, genetic makeup, sex ratio, racial and ethnic background, and other properties that are subsequently modified through migration, mortality, and environmental contact. The succession of cohorts can also affect the capacities, attitudes, and actions of people in particular strata as the members of each cohort bring to society their experiences with the social and environmental events spanned by their respective lifetimes.

The most significant alterations in the age composition of modern societies stem from the dramatic and unprecedented increases in the longevity of successive cohorts. Age pyramids diagramming the age composition of the United States in 2010 compared with 1955, for example, demonstrate that entirely new strata have been added at

the oldest ages (Taeuber 1992)—strata of old people who are healthier and more competent than their predecessors (Manton, Corder, and Stallard 1997). The advent of these “new” old people is already having untold consequences: Individuals now have time to spread education, work, family activities, and leisure more evenly over their long lives, and wider structural opportunities are needed in society for the age-heterogenous population.

### SOCIAL STRUCTURAL PERSPECTIVES

Cohorts, as described above, are composed of people, who age and fit together in strata to form the composition of the population; but cohorts also shape, and are shaped by social structures—the surrounding families, communities, work organizations, educational institutions, and the like. Against the backdrop of history, social structures, like lives, tend to change, and two “dynamisms”—changing structures and changing lives—are in continuing interplay, as each influences the other. Thus full understanding of cohorts requires understanding their reciprocal relations with structures (as in Foner and Kertzer 1978; Mayer 1988). Toward this end, some studies examine how the processes of aging and cohort flow relate to structures, while other studies examine the congruence—or lack of congruence—between age composition and social structures.

**Aging and Structures .** Because cohorts differ in size and character, and because their members age in new ways (the diagonal lines in Figure 1), they exert collective pressures for adjustments—not only in people’s ideas, values, and beliefs—but also in role opportunities throughout the social institutions.

As one example, the influences of cohort differences in *size* were defined early by Joan Waring’s (1975) powerful analysis of “disordered cohort flow.” This disordered flow has been dramatically brought to attention as the Baby Boom cohorts first pressed for expansions in the school systems and the labor force, and will become the twenty-first century “senior boom” that will exacerbate the inadequacy of roles for the elderly. Later, as large cohorts were followed by smaller successors, ways were sought to reduce these expanded structures again. Meanwhile, as structures changed, the lives of people moving through these structures also changed.

In another example, the influences of cohort differences in *norms* has been analyzed as the process of “cohort norm formation” (Riley 1978). As members of a cohort respond to shared historical experiences, they gradually and subtly develop common patterns of response, common definitions, and common beliefs, that crystallize into new norms and become institutionalized in altered social structures. For instance, over the past century many individual women in successive cohorts have responded to common social changes by making many millions of separate but similar personal decisions to move in new directions: to go to college, have a career, or form their families in innovative ways. Such decisions, beginning in one cohort and transmitted from cohort to cohort, can feed back into the social structures and gradually pervade entire segments of society. Thus, many new age norms have become expectations that women should work, and have stimulated the demand for new role opportunities at work and in the family for people of all ages.

**Age Composition and Structures.** At any given period of history, the coexisting cohorts of people that form the age strata coincide with the existing role structures (both indicated by the perpendicular lines in Figure 1). People who differ in age and experience confront the available age-related opportunities, or lack of opportunities, in work, education, recreation, the family, and elsewhere. However, people and structures rarely fit together smoothly: there is a mismatch or “lag” of one dynamism behind the other.

*Structural Lag.* While people sometimes lag behind structures as technology advances, more frequent in modern society is the failure of structural changes to keep pace with the increasing numbers of long-lived and competent people (Riley, Kahn, and Foner 1994). Cohorts of those who are young today have few “real-world” opportunities; those in the middle years are stressed by the combined demands of work and family; and those who have reached old age are restive in the prolonged “roleless role” of retirement. Cohorts of people now old are more numerous, better educated, and more vigorous than their predecessors were in 1920 or 1950; but few changes in the places for them in society have been made. Capable people and empty role structures cannot long coexist. Thus, implicit in the lag are perpetual pressures toward structural change.

*Age Integration.* Among societal responses to structural lag are current tendencies toward “age integration” (Riley, Foner, and Riley 1999, p.338). With pressures from the expanded numbers of age strata, many age barriers dividing education, work and family, and retirement are gradually becoming more flexible. People of different ages are more often brought together, as lifelong education means that old and young study together, as new entrepreneurs hire employees of mixed ages, as in many families four generations are alive at the same time, or as the age segregation of nursing homes is replaced by home health care with wide access to others. Where such tendencies may lead in the future is not yet known. But the interdependence between cohorts and structure is clear.

## RESEARCH METHODS.

When aspects of these broad cohort perspectives are translated into empirical studies, a variety of research methods are required for specific objectives: from analyses of historical documents and subjective reports, to panel analyses and mathematical modeling, to rigorous tests of specific hypotheses. This brief overview can only hint at the diverse research designs involved in analyses of the multiple factors affecting lives of people in particular cohorts; or in the shifting role opportunities for cohort members confronting economic, religious, political, and other social institutions (for one example, see Hendricks and Cutler 1990).

**Cohort Analysis.** The tool most widely used in large-scale studies is “cohort analysis,” which takes the intercohort aging perspective—in contrast to “period analysis” (Susser 1969), which takes the cross-sectional perspective. (The difference is illustrated in Figure 1 by comparison of the diagonal cohort lines, in contrast to comparison of a sequence of vertical compositional slices). In his 1992 formulation of the technical aspects of cohort analysis, Ryder defines the term as “the parameterization of the life cycle behavior of individuals over personal time, considered in the aggregate, and the study of change in those parameters over historical time” (p. 230). He conceptualizes the cohort as “providing a macro-analytic link between movements of individuals from one to another status, and movements of the population composition from one period to the next.”

His classic work on “demographic translation” sets out the mathematical procedure for moving between the cohort and the period “modes of temporal aggregation” (Ryder 1963, 1983).

Central to this method is the *identification problem* confounding many attempts at cohort analysis. This problem occurs from efforts to interpret the separate effects of three concepts—cohort, period, and age (C, P, and A)—when only two variables, such as date and years of age, are indexed. Apart from various procedures that assume one of the parameters is zero, the most appropriate solution to this problem is to specify and measure directly the three concepts used in the particular analysis (Cohn 1972; Rodgers 1982; Riley, Foner, and Waring 1988, pp. 260–261). After all, as Ryder puts it (1992, p. 228), the cohort (C) is a set of actors, the age (A) is their age, and the period (P) stands for the social context at the time of observation.

**Wide Range of Methods.** Ryder’s exegesis of this particular method illustrates, through its strengths and limitations, the utility of the tripartite conceptualization of cohort perspectives as a heuristic guide. The strengths in Ryder’s formulation focus exclusively on the significance of cohorts as a macro-analytic vehicle for social change. In the broader cohort perspectives outlined here, many other methods are useful for specific objectives where cohort analysis is inappropriate. Some employ a cross-sectional approach. Others utilize intercohort comparisons to focus on aging processes at the individual as well as the collective levels. Still others complement demographic analyses of populations with examination of the related structures of social roles and institutions.

Thus, despite its signal contributions, neither cohort analysis nor any other single method can comprehend the full power of cohorts as ingredients of aging processes, age composition, and the complex interplay with social structure.

(SEE ALSO: *Structural Lag*)

#### REFERENCES

- Campbell, Richard T. 1994 “A Data-Based Revolution in the Social Sciences.” *ICPSR Bulletin*, XIV, 1–4.
- Clausen, John A. 1986 *The Life Course: A Sociological Perspective*. Englewood Cliffs, N.J.: Prentice-Hall.

- Cohn Richard 1972 “On Interpretation of Cohort and Period Analyses: A Mathematical Note.” In Matilda White Riley, Marilyn Johnson, and Anne Foner, eds., *Aging and Society: A Sociology of Age Stratification*, Vol. III. New York: Russell Sage Foundation.
- Dannefer, Dale 1987 “Aging as Intracohort Differentiation: Accentuation, the Matthew Effect, and the Life Course.” *Sociological Forum* 2 (spring):211–236.
- Duncan, Greg J., Martha Hill, and Willard Rodgers 1986 “The Changing Fortunes of Young and Old.” *American Demographics* 8:26–34.
- Elder, Glenn H., Jr., and R. C. Rockwell 1979 “Economic Depression and Postwar Opportunity in Men’s Lives: A Study of Life Patterns and Mental Health.” In R. G. Simmons, ed., *Research in Community and Mental Health*, Vol. 1. Greenwich, Conn.: JAI Press.
- Featherman, David L. 1981 “The Life-Span Perspective.” In *The National Science Foundation’s 5-Year Outlook on Science and Technology*, Vol. 2. Washington, D.C.: Government Printing Office.
- Foner, Anne and David I. Kertzer 1978 “Transitions Over the Life Course: Lessons from Age-Set Societies.” *American Journal of Sociology* 83:1081–1104.
- Hendricks, Jon, and Stephen J. Cutler 1990 “Leisure and the Structure of our Life Worlds.” *Aging and Society* 10:85–94.
- Mannheim, Karl (1928) 1952 “The Problem of Generations.” In Paul Kecskemeti, ed., *Essays on the Sociology of Knowledge*. London: Routledge & Kegan Paul.
- Manton, Kenneth G. 1989 “Life-Style Risk Factors.” In M. W. Riley and J. W. Riley, Jr., eds., *The Quality of Aging: Strategies for Interventions*, Special issue of *The Annals* (503). Newbury Park, Calif.: Sage.
- , Larry Corder, and Eric Stallard 1997 “Chronic Disability Trends in Elderly United States Populations: 1982–1994.” *Proceedings of the National Academy of Sciences* 94:2593–2598
- Mayer, Karl Ulrich 1988 “German Survivors of World War II: The Impact of the Life Course of the Collective Experiences of Birth Cohorts.” In M. W. Riley, B. J. Huber, and B. B. Hess, eds., *Social Structures and Human Lives*. Newbury Park, Calif.: Sage.
- Messeri, Peter 1988 “Age, Theory Choice, and the Complexity of Social Structure.” In M. W. Riley, B. J. Huber, and B. B. Hess, eds., *Social Structures and Human Lives*. Newbury Park, Calif.: Sage.
- National Association of State Boards of Education 1990 *Code Blue: Uniting for Healthier Youth*. Alexandria, Va.: National Association of State Boards of Education.
- Nesselroade, John R. 1991 “The Warp and the Woof of the Development Fabric.” In R. M. Downs, L. S.

- Liben, and D. S. Palermo, eds., *Visions Of Aesthetics*. Hillsdale, N.J.: Lawrence Erlbaum.
- O'Rand, Angela M. and Richard T. Campbell 1999 "On Reestablishing the Phenomenon and Specifying Ignorance: Theory Development and Research Design in Aging." In Vern L. Bengtson and K. Warner Schaie, eds., *Handbook of Theories of Aging*. New York: Springer.
- Preston, Samuel 1984 "Children and the Elderly: Diverse Paths for America's Dependents." *Demography* 21:435-457.
- Riley, Matilda White 1973 "Aging and Cohort Succession: Interpretations and Misinterpretation." *Public Opinion Quarterly* 37:35-49.
- 1978 "Aging, Social Change, and the Power of Ideas." *Daedalus* 107:39-52.
- 1982 "Aging and Social Change." In M. W. Riley, R. P. Abeles, and M. S. Teitelbaum, eds., *Aging from Birth to Death: Sociotemporal Perspectives*, Vol. II, AAAS Selected Symposium 79. Boulder, Colo.: Westview Press.
- 1988 "The Aging Society: Problems and Prospects." In *Proceedings of the American Philosophical Society*, Vol. 132, No. 2, 148-153.
- 1994 "A Life Course Approach." In J. Z. Giele and G. H. Elder, Jr., eds., *Methods of Life Course Research*. Thousand Oaks: Sage.
- , and Ronald P. Abeles 1990 *The Behavioral and Social Research Program at the National Institute of Aging: History of a Decade*. Bethesda, Md.: National Institute on Aging, National Institutes of Health.
- Riley, Matilda White, Anne Foner, and John W. Riley, Jr. 1999 In Vern L. Bengtson and K. Warner Schaie, eds., *Handbook of Theories of Aging*. New York: Springer.
- Riley, Matilda White, Robert L. Kahn, and Anne Foner 1994 *Age and Structural Lag: Society's Failure to Provide Meaningful Opportunities in Work, Family, and Leisure*. New York: Wiley.
- Riley, Matilda White, Anne Foner, and Joan Waring 1988 "Sociology of Age." In Neil J. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage Publications.
- Riley, Matilda White, Marilyn Johnson, and Anne Foner 1972 *Aging and Society, Vol. III: A Sociology of Age Stratification*. New York: Russell Sage Foundation.
- Rodgers, Willard L. 1982 "Estimable Functions of Age, Period, and Cohort Effects." *American Sociological Review* 47:774-787.
- Ryder, Norman B. 1963 "The Translation Model of Demographic Change." In Norman B. Ryder, ed., *Emerging Techniques in Population Research*. New York: Milbank Memorial Fund.
- 1965 "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30:843-861.
- 1968 "Cohort Analysis." *International Encyclopedia of the Social Sciences* 5:546-550.
- 1983 "Cohort and Period Measures of Changing Fertility." In Rodolfo A. Bulatao and Ronald D. Lee, eds., *Determinants of Fertility in Developing Countries*. New York: Academic Press.
- 1992 "Cohort Analysis." In E. F. Borgatta and M. L. Borgatta, eds., *Encyclopedia of Sociology*, Vol. 1. New York: MacMillan.
- Schaie, K. Warner 1996 *Intellectual Development in Adulthood: The Seattle Longitudinal Study*. New York: Cambridge University Press.
- Susser, Mervyn 1969 "Aging and the Field of Public Health." In M. W. Riley, J. W. Riley, Jr., and M. E. Johnson, eds., *Aging and Society, Vol. II, Aging and the Professions*. New York: Russell Sage Foundation.
- Tauber, Cynthia M. 1992 "Sixty-five Plus in America." *Current Population Reports*, P23-178RV. Washington, D.C.: U.S. Bureau of the Census.
- Waring, Joan M. 1975 "Social Replenishment and Social Change." *American Behavioral Scientist* 19:237-256.
- Zuckerman, Harriet and Robert K. Merton 1972 "Age, Aging and Age Structure in Science." In M. W. Riley, M. Johnson, and A. Foner, eds., *Aging and Society: Volume II, A Sociology of Age Stratification*. New York: Russell Sage Foundation.

MATILDA WHITE RILEY

## COLLECTIVE BEHAVIOR

Collective behavior consists of those forms of social behavior in which the usual conventions cease to guide social action and people collectively transcend, bypass, or subvert established institutional patterns and structures. As the name indicates, the behavior is collective rather than individual. Unlike small group behavior, it is not principally coordinated by each-to-each personal relationships, though such relationships do play an important part. Unlike organizational behavior, it is not coordinated by formally established goals, authority, roles, and membership designations, though emergent leadership and an informal role structure are important components. The best known forms of collective behavior are rumor, spontaneous collective responses to crises such as natural disasters; crowds, collective panics,



crazes, fads, fashions, publics (participants in forming public opinion), cults, followings; and reform and revolutionary movements. Social movements are sometimes treated as forms of collective behavior, but are often viewed as a different order of phenomena because of the degree of organization necessary to sustain social action. This essay will include only those social movement theories that also have relevance for the more elementary forms of collective behavior.

Theories of collective behavior can be classified broadly as focusing on the behavior itself (microlevel) or on the larger social and cultural settings within which the behavior occurs (macrolevel or structural). An adequate theory at the microlevel must answer three questions, namely: How is it that people come to transcend, bypass, or subvert institutional patterns and structures in their activity; how do people come to translate their attitudes into significant overt action; and how do people come to act collectively rather than singly? Structural theories identify the processes and conditions in culture and social structure that are conducive to the development of collective behavior. Microlevel theories can be further divided into *action* or *convergence* theories and *interaction* theories.

### MICROLEVEL CONVERGENCE THEORIES

*Convergence* theories assume that when a critical mass of individuals with the same disposition to act in a situation come together, collective action occurs almost automatically. In all convergence theories it is assumed that: "The individual in the crowd behaves just as he would behave alone, *only more so*," (Allport 1924, p. 295), meaning that individuals in collective behavior are doing what they wanted to do anyway, but could not or feared to do without the "facilitating" effect of similar behavior by others. The psychological hypothesis that frustration leads to aggression has been widely applied in this way to explain racial lynchings and riots, rebellion and revolution, and other forms of collective violence. Collective behavior has been conceived as a collective pursuit of meaning and personal identity when strains and imbalances in social institutions have made meaning and identity problematic (Klapp 1972). In order to explain the convergence of a critical mass of people experiencing similar frustrations, investigators

posit deprivation shared by members of a social class, ethnic group, gender group, age group, or other social category. Because empirical evidence has shown consistently that the most deprived are not the most likely to engage in collective protest, more sophisticated investigators assume a condition of *relative deprivation* (Gurr 1970), based on a discrepancy between expectations and actual conditions. Relative deprivation frequently follows a period of rising expectations brought on by improving conditions, interrupted by a setback, as in the J-curve hypothesis of revolution (Davies 1962). Early explanations for collective behavior, generally contradicted by empirical evidence and repudiated by serious scholars, characterized much crowd behavior and many social movements as the work of criminals, the mentally disturbed, persons suffering from personal identity problems, and other deviants.

**Rational decision theories.** Several recent convergence theories assume that people make rational decisions to participate or not to participate in collective behavior on the basis of selfinterest. Two important theories of this sort are those of Richard Berk and Mark Granovetter.

Berk (1974) defines collective behavior as the behavior of people in crowds, which means activity that is transitory, not well planned in advance, involving face-to-face contact among participants, and considerable cooperation, though he also includes panic as competitive collective behavior. Fundamental to his theory is the assumption that crowd activity involves rational, goal-directed action, in which possible rewards and costs are considered along with the chances of support from others in the crowd. Rational decision making means reviewing viable options, forecasting events that may occur, arranging information and choices in chronological order, evaluating the possible consequences of alternative courses of action, judging the chances that uncertain events will occur, and choosing actions that minimize costs and maximize benefits. Since the best outcome for an individual in collective behavior depends fundamentally on what other people will do, participants attempt to advance their own interests by recruiting others and through negotiation. Berk's theory does not explain the origin and nature of the proposals for action that are heard in the crowd, but describes the process by which these proposals

are sifted as the crowd moves toward collaborative action, usually involving a division of labor. To explain decision making, he offers a simple equation in which the probability of a person beginning to act (e.g., to loot) is a function of the product of the net anticipated personal payoff for acting (e.g., equipment or liquor pilfered) and the probability of group support in that action (e.g., bystanders condoning or joining in the looting.)

Granovetter's (1978) application of rational decision theory focuses on the concept of *threshold*. He assumes that each person, in a given situation, has a threshold number or percentage of other people who must already be engaging in a particular action before he or she will join in. Since it can be less risky for the individual to engage in collective behavior (riotous behavior, for example) when many others are doing so than when few are involved, the benefit-to-cost ratio improves as participation increases. Based on the personal importance of the action in question, individual estimation of risk, and a host of other conditions, individual thresholds will vary widely in any situation. Collective behavior cannot develop without low-threshold individuals to get it started, and development will stop when there is no one with the threshold necessary for the next escalation step. Collective behavior reaches an equilibrium, which can be ascertained in advance from knowing the distribution of thresholds, when this point is reached. Like Berk, Granovetter makes no effort to explain what actions people will value. Furthermore, intuitively appealing as the theory may be, operationalizing and measuring individual thresholds may be, for all practical purposes, impossible.

#### MICRO-LEVEL INTERACTION THEORIES

**Contagion Theories.** Early interaction theories, which lay more emphasis on what happens to people in the context of a crowd or other collectivity than on the dispositions people bring to the collectivity, stressed either the emergence of a group mind or processes of imitation, suggestion, or social contagion. Serge Moscovici (1985a, 1985b) is a defender of these early views, stressing that normal people suffer a lowering of intellectual faculties, an intensification of emotional reactions, and a disregard for personal profit in a crowd. The fundamental crowd process is suggestion, emanating from charismatic leaders. During the twentieth

century, the breakdown of social ties has created masses who form larger and larger crowds that are controlled by a few national and international leaders, creating an historically new politics or appeal to the masses.

Herbert Blumer (1939) developed a version of the contagion approach that has been the starting point for theories of collective behavior for most American scholars. Blumer explains that the fitting together of individual actions in most group behavior is based on shared understandings under the influence of custom, tradition, conventions, rules, or institutional regulations. In contrast, collective behavior is group behavior that arises spontaneously, and not under the guidance of preestablished understandings, traditions, or rules of any kind. If sociology in general studies the social order, collective behavior consists of the processes by which that order comes into existence. While coordination in publics and social movements and involving more complex cognitive processes called *interpretation interaction*, coordination in the crowd and other elementary forms of collective behavior is accomplished through a process of *circular reaction*. Circular reaction is a type of interstimulation in which the response by others to one individual's expression of feeling simply reproduces that feeling, thereby reinforcing the first individual's feeling, which in turn reinforces the feelings of the others, setting in motion an escalating spiral of emotion. Circular reaction begins with individual restlessness, when people have a blocked impulse to act. When many people share such restlessness, and are already sensitized to one another, circular reaction can set in and create a process of *social unrest* in which the restless state is mutually intensified into a state of *milling*. In milling, people move or shift their attention aimlessly among each other, thereby becoming preoccupied with each other and decreasingly responsive to ordinary objects and events. In the state of *rappport*, collective excitement readily takes over, leading to a final stage of *social contagion*, the "relatively rapid, unwitting, and non-rational dissemination of a mood, impulse, or form of conduct." (Blumer 1939) Social unrest is also a prelude to the formation of publics and social movements. In the case of the public, the identification of an issue rather than a mood or point of view converts the interaction into discussion rather than circular reaction. Social movements begin with circular

reaction, but with persisting concerns they acquire organization and programs, and interpretative interaction prevails.

**Emergent Norm Theory.** Ralph Turner and Lewis Killian [(1957) 1989] criticize convergence theories for underemphasizing the contribution of interaction processes in the development of collective behavior, and found both convergence and contagion theories at fault for assuming that participants in collective behavior become homogeneous in their moods and attitudes. Instead of emotional contagion, it is the emergence of a norm or norms in collective behavior that facilitates coordinated action and creates the illusion of unanimity. The emergent norm is characteristically based on established norms, but transforms or applies those norms in ways that would not ordinarily be acceptable. What the emergent norm permits or requires people to believe, feel, and do corresponds to a disposition that is prevalent but not universal among the participants. In contrast to convergence theories, however, it is assumed that participants are usually somewhat ambivalent, so that people could have felt and acted in quite different ways if the emergent norm had been different. For example, many rioters also have beliefs in law and order and fair play that might have been converted into action had the emergent norm been different. Striking events, symbols, and keynoting—a gesture or symbolic utterance that crystallizes sentiment in an undecided and ambivalent audience—shape the norm and supply the normative power, introducing an element of unpredictability into the development and direction of all collective behavior.

Emergent norm theory differs from contagion theories in at least six important and empirically testable ways. First, the appearance of unanimity in crowds, social movements, and other forms of collective behavior is an illusion, produced by the effect of the emergent norm in silencing dissent. Second, while the collectivity's mood and definition of the situation are spontaneously induced in some of the participants, many participants experience group pressure first and only later, if at all, come to share the collectivity's mood and definition of the situation. Third, unlike collective excitement and contagion, normative pressure is as applicable to quiet states such as dread and sorrow as it is to excited states. Fourth,

according to emergent norm theory, a conspicuous component in the symbolic exchange connected with the development of collective behavior should consist of seeking and supplying justifications for the collectivity's definition of the situation and action, whereas there should be no need for justifications if the feelings were spontaneously induced through contagion. Fifth, a norm not only requires or permits certain definitions and behaviors; it also sets acceptable limits, while limits are difficult to explain in terms of a circular reaction spiral. Finally, while contagion theories stress anonymity within the collectivity as facilitating the diffusion of definitions and behavior that deviate from conventional norms, emergent norm theory asserts that familiarity among participants in collective behavior enhances the controlling effect of the emergent norm.

Emergent norm theory has been broadened to make explicit the answers to all three of the key questions microlevel theories must answer: The emergent normative process as just described provides the principal answer to the question, why people adopt definitions and behavior that transcend, bypass, or contravene established social norms; participants translate their attitudes into overt action rather than remaining passive principally because they see action as feasible and timely; and action is collective rather than individual primarily because of preexisting groupings and networks and because an event or events that challenge conventional understandings impel people to turn to others for help in fashioning a convincing definition of the problematic situation. In addition, these three sets of processes interact and are mutually reinforcing in the development and maintenance of collective behavior. This elaboration of the emergent norm approach is presented as equally applicable to elementary forms of collective behavior such as crowds and to highly developed and organized forms such as social movements.

**Other Interaction Theories.** Although all interactional theories presume that collective behavior develops through a cumulative process, Max Heirich (1964) makes this central to his theory of collective conflict, formulated to explain the 1964–1965 year of spiraling conflict between students and the administration at the University of California, Berkeley. Common action occurs when observers perceive a situation as critical, with limited time for action, with the crisis having a simple

cause and being susceptible to influence by simple acts. Heirich specifies determinants of the process by which such common perceptions are created and the process by which successive redefinitions of the situation take place. Under organizational conditions that create unbridged cleavages between groups that must interact regularly, conflict escalates through successive encounters in which cleavages become wider, issues shift, and new participants join the fray, until the conflict becomes focused around the major points of structural strain in the organization.

Also studying collective conflict as a cumulative process, Bert Useem and Peter Kimball (1989) developed a sequence of stages for prison riots, proceeding from pre-riot conditions, to initiation, expansion, siege, and finally termination. While they identify disorganization of the governing body as the key causative factor, they stress that what happens at any one stage is important in determining what happens at the next stage.

Clark McPhail (1991), in an extensive critique of all prior work, rejects the concept of collective behavior as useless because it denotes too little and fails to recognize variation and alternation within assemblages. Instead of studying collective behavior or crowds, he proposes the study of *temporary gatherings*, defined as two or more persons in a common space and time frame. Gatherings are analyzed in three stages, namely, assembling, gathering, and dispersing. Rather than positing an overarching principle such as contagion or norm emergence, this approach uses detailed observation of individual actions and interactions within gatherings and seeks explanations at this level. Larger events such as campaigns and large gatherings are to be explained as the "repetition and/or combination of individual and collective sequences of actions." (McPhail 1991, p. 221). These elementary actions consist of simple observable actions such as clustering, booing, chanting, collective gesticulation, "locomotion," synchroclapping, and many others. The approach has been implemented by precise behavioral observation of people assembling for demonstrations and other preplanned gatherings. A promised further work will help determine how much this approach will contribute to the understanding of those fairly frequent events usually encompassed by the term collective behavior.

### MACROLEVEL OR STRUCTURAL THEORIES

Microlevel theories attempt first to understand the internal dynamics of collective behavior, then use that understanding to infer the nature of conditions in the society most likely to give rise to collective behavior. In contrast, macrolevel or structural theories depend primarily on an understanding of the dynamics of society as the basis for developing propositions concerning when and where collective behavior will occur. Historically, most theories of elementary collective behavior have been microlevel theories, while most theories of social movements have been structural. Neil Smelser's 1993 *value-added* theory is primarily structural but encompasses the full range from panic and crazes to social movements.

Smelser attempted to integrate major elements from the Blumer and Turner/Killian tradition of microtheory into an action and structural theory derived from the work of Talcott Parsons. Smelser describes the normal flow of social action as proceeding from values to norms to mobilization into social roles and finally to situational facilities. Values are the more general guides to behavior; norms specify more precisely how values are to be applied. Mobilization into roles is organization for action in terms of the relevant values and norms. Situational facilities are the means and obstacles that facilitate and hinder attainment of concrete goals. The four "components of social action" are hierarchized in the sense that any redefinition of a component requires readjustment in the components below it, but not necessarily in those above. Each of the four components in turn has seven levels of specificity with the same hierarchical ordering as the components. Types of collective behavior differ in the level of the action components they aim to restructure. Social movements address either values, in the case of most revolutionary movements, or norms, in the case of most reform movements. Elementary collective behavior is focused at either the mobilization or the situational facilities level. Collective behavior is characterized formally as "an uninstitutionalized mobilization for action in order to modify one or more kinds of strain on the basis of a generalized reconstitution of a component of action." (Smelser 1963 p. 71). The distinguishing feature of this action is a shortcircuiting of the normal flow of

action from the general to the specific. There is a jump from extremely high levels of generality to specific, concrete situations, without attention to the intervening components and their levels of specificity. Thus, in Smelser's view, collective behavior is intrinsically irrational.

In order for collective behavior to occur, six conditions must be met, each of which is necessary but insufficient without the others. Smelser likens the relationship among the six determinants to the value-added process in economics, with each adding an essential component to the finished product. The first determinant is *structural conduciveness*, meaning that the social structure is organized in a way that makes the particular pattern of action feasible. The second determinant is *structural strain*, consisting of ambiguities, deprivations, conflicts, and discrepancies experienced by particular population segments. Third (and central in Smelser's theorizing), is the growth and spread of a *generalized belief* that identifies and characterizes the supposed source of strain and specifies appropriate responses. The generalized belief incorporates the short-circuiting of the components of action that is a distinctive feature of collective behavior. Fourth are *precipitating factors*, usually a dramatic event or series of events that give the generalized belief concrete and immediate substance and provide a concrete setting toward which collective action can be directed. The fifth determinant is *mobilization of participants for action*, in which leadership behavior is critical. The final determinant is the *operation of social control*. Controls may serve to minimize conduciveness and strain, thus preventing the occurrence of an episode of collective behavior, or they may come into action only after collective behavior has begun to materialize, either dampening or intensifying the action by the way controls are applied. These determinants need not occur in any particular order.

Addressing a more limited range of phenomena, David Waddington, Karen Jones, and Chas Critcher (1989) have formulated a *flashpoint* model to explain public disorders that bears some resemblance to the value-added component of Smelser's theory. Public disorders typically begin when some ostensibly trivial incident becomes a flashpoint. The flashpoint model is a theory of the conditions that give a minor incident grave significance. Explanatory conditions exist at six levels. At

the *structural* level are conflicts inherent in material and ideological differences between social groups that are not easily resolvable within the existing social structure, meaning especially the state. At the *political/ideological* level, dissenting groups are unable to express their dissent through established channels, and their declared ends and means are considered illegitimate. At the *cultural* level, the existence of groups with incompatible definitions of the situation, appropriate behavior, or legitimate rights can lead to conflict. At the *contextual* level, a history of past conflicts between a dissenting group and police or other authorities enhances the likelihood that a minor incident will become a flashpoint. At the *situational* level, immediate spatial and social conditions can make public control and effective negotiation difficult. Finally, at the *interactional* level, the dynamics of interaction between police and protesters, as influenced by meanings derived from the other five levels, ultimately determine whether there will or will not be public disorder and how severe it will be. Unlike Smelser, Waddington and associates make no assumption that all levels of determinants must be operative. Also, they make no explicit assumption that disorderly behavior is irrational, though their goal is to formulate public policy that will minimize the incidence of public disorders.

*Resource mobilization* theories have been advanced as alternatives to Smelser's value-added theory and to most microlevel theories. Although they have generally been formulated to explain social movements and usually disavow continuity between social movements and elementary collective behavior, they have some obvious implications for most forms of collective behavior. There are now several versions of resource mobilization theory, but certain core assumptions can be identified. Resource mobilization theorists are critical of prior collective behavior and social movement theories for placing too much emphasis on "structural strain," social unrest, or grievances; on "generalized beliefs," values, ideologies, or ideas of any kind; and on grass-roots spontaneity in accounting for the development and characteristics of collective behavior. They assume that there is always sufficient grievance and unrest in society to serve as the basis for collective protest (McCarthy and Zald 1977), and that the ideas and beliefs exploited in protest are readily available in the culture (Oberschall 1973). They see collective protest as

centrally organized, with the bulk of the participants “mobilized” much as soldiers in an army are mobilized and directed by their commanders. In explaining the rise of collective protest they emphasize the availability of essential resources, such as money, skills, disposable time, media access, and access to power centers, and prior organization as the base for effectively mobilizing the resources. Resource mobilization theorists favor the use of rational decision models to explain the formulation of strategy and tactics, and emphasize the role of social movement professionals in directing protest.

There has been some convergence between resource mobilization theorists and the theorists they criticize. The broadened formulation of emergent norm theory to incorporate resources (under “feasibility”) and prior organization as determinants of collective behavior takes account of the resource mobilization contribution without, however, giving primacy to these elements. Similarly, many resource mobilization theorists have incorporated social psychological variables in their models.

Alberto Melucci (1989) offers a *constructivist* view of collective action which combines macro- and micro-orientations. Collective action is the product of purposeful negotiations whereby a plurality of perspectives, meanings, and relationships crystallize into a pattern of action. Action is constructed to take account of the goals of action, the means to be utilized, and the environment within which action takes place, which remain in a continual state of tension. The critical process is the negotiation of collective identities for the participants. In the current post-industrial era, conflicts leading to collective action develop in those areas of communities and complex organizations in which there is greatest pressure on individuals to conform to the institutions that produce and circulate information and symbolic codes.

## REFERENCES

- Allport, Floyd H. 1924 *Social Psychology*. Boston: Houghton Mifflin.
- Berk, Richard A. 1974 “A Gaming Approach to Crowd Behavior.” *American Sociological Review* 39:355–73.
- Blumer, Herbert 1939 “Collective Behavior.” in Robert E. Park ed., *An Outline of the Principles of Sociology* 221–280. New York: Barnes & Noble.

- Davies, James C. 1962 “Toward a Theory of Revolution.” *American Journal of Sociology* 27:5–19.
- Granovetter, Mark 1978 “Threshold Models of Collective Behavior.” *American Journal of Sociology* 83:1420–43.
- Gurr, Ted R. 1970 *Why Men Rebel*. Princeton: Princeton University Press.
- Heirich, Max 1971 *The Spiral of Conflict: Berkeley, 1964*. Berkeley: University of California Press.
- Klapp, Orrin E. 1972 *Currents of Unrest: An Introduction to Collective Behavior*. New York: Holt, Rinehart, & Winston.
- McPhail, Clark 1991 *The Myth of the Madding Crowd* New York: Aldine de Gruyter.
- Melucci, Alberto 1989 *Nomads of the Present: Social Movements and Individual Needs in Contemporary Society*. Philadelphia: Temple University Press.
- Moscovici, Serge 1985a *The Age of the Crowd: A Historical Treatise on Mass Psychology*. Cambridge: Cambridge University Press.
- 1985b “The Discovery of the Masses,” In C. F. Graumann and Moscovici, eds., *Changing Conceptions of Crowd Mind and Behavior*, 5–25 New York: Springer-Verlag.
- Smelser, Neil J. 1963 *Theory of Collective Behavior*. New York: Free Press.
- Turner, Ralph H., and Lewis M. Killian (1957) 1987 *Collective Behavior*. Englewood Cliffs, N. J.: Prentice-Hall.
- Turner, Ralph H. 1994 “Race Riots Past and Present: A Cultural-Collective Behavior Approach.” *Symbolic Interaction* 17:309–24.
- 1996 “Normative Emergence in Collective Behavior and Action” *Solidarity: Research in Social Movements*. 1:1–16.
- Useem, Bert, and Peter Kimball 1989 *States of Siege: U. S. Prison Riots, 1971–1986*. New York: Oxford University Press.
- Waddington, David, Karen Jones, and Chas Critcher 1989 *Flashpoints: Studies in Public Disorder*. London: Routledge.

RALPH H. TURNER

**COLONIZATION**

See Imperialism, Colonialism, and Decolonization.

## COMMUNES

See Alternative Life-Styles.

## COMMUNISM

See Socialism and Communism.

## COMMUNITARIANISM

Communitarianism is a social philosophy that core assumption is the required shared (“social”) formulations of the good. The assumption is both empirical (social life exhibits shared values) and normative (shared values ought to be formulated). While many sociologists may consider such an assumption as subject to little controversy, communitarianism is in effect a highly contested social philosophy. It is often contrasted with liberalism (based on the works of John Locke, Adam Smith, and John Stuart Mill, not to be confused with liberalism as the term is used in contemporary American politics). Liberalism’s core assumption is that what people consider right or wrong, their values, should strictly be a matter for each individual to determine. To the extent that social arrangements and public policies are needed, these should not be driven by shared values but by voluntary arrangements and contracts among the individuals involved, thus reflecting their values and interests. Communitarians, in contrast, see social institutions and policies as affected by tradition and hence by values passed from generation to generation. These become part of the self through nonrational processes, especially internalization, and are changed by other processes such as persuasion, religious or political indoctrination, leadership, and moral dialogues.

In addition, communitarianism emphasizes particularism, the special moral obligations people have to their families, kin, communities, and societies. In contrast, liberalism stresses the universal rights of all individuals, regardless of their particular membership. Indeed, liberal philosopher Jeremy Bentham declared that the very notion of a society is a fiction.

Until 1990, sociological and social psychological researchers and theorists and communitarian

philosophers often ignored one another’s works, despite the fact that they dealt with closely related issues. It should be noted, though, that communitarians were much more inclined to be openly and systematically normative than many social scientists.

### HISTORY

Like many other schools of thought, communitarianism has changed considerably throughout its history, and has various existing camps that differentiate significantly. As far as can be determined, the term “communitarian” was not used until 1841, when Goodwyn Barmby, an official of the Communist Church, founded the Universal Communitarian Association.

Communitarian issues were addressed long before that date, however, for instance in Aristotle’s comparison of the isolated lives of people in the big metropolis to close relationships in the smaller city. Both the Old and the New Testament deal with various issues one would consider communitarian today, for instance the obligations to one’s community. The social teaching of the Catholic Church (for instance, concerning subsidiarity) and of early utopian socialism (for example, regarding communal life and solidarity), all contain strong communitarian *elements*, although these works are not comprehensive communitarian statements and are not usually considered as communitarian works per se.

Among early sociologists whose work is strongly communitarian, although this fact is as a rule overlooked by social philosophers, are Ferdinand Tönnies, especially his comparison of the *Gemeinschaft* and *Gessellschaft*, (or community and society); Emile Durkheim, especially his concerns about the integrating role of social values and the relations between the individual and society; and George Herbert Mead. These works are extensively examined elsewhere in this encyclopedia and hence are not discussed here.

A communitarian who combined social philosophy and sociology Martin Buber. Especially relevant are Buber’s contrast between I-It and I-Thou relations, his interest in dialogue, and his distinction between genuine communal relationships and objectified ones.

Other sociologists whose work contains communitarian elements are Robert E. Park, William Kornhauser, and Robert Nisbet. Philip Selznick, Robert Bellah and his associates, and Amitai Etzioni wrote books that laid the foundations for new (or responsive) communitarianism, which Etzioni launched as a “school” and somewhat of a social movement in 1990.

Selznick’s *The Moral Commonwealth* is especially key in forming a strong intellectual grounding for new communitarian thinking, and presents an integration of moral and social theory in a synthesis of “communitarian liberalism.” According to Selznick, communitarianism does not reject basic liberal ideals and achievements; it seeks reconstruction of liberal perspectives to mitigate the excesses of individualism and rationalism, and to encourage an ethic of responsibility (in contrast to liberalism, where the concept of responsibility has no major role). In a community there is an irrepressible tension between exclusion and inclusion, and between civility and piety. Thus community is not a restful idea, a realm of peace and harmony. On the contrary, competing principles must be recognized and dealt with.

#### AUTHORITARIAN, POLITICAL THEORETICAL, AND RESPONSIVE COMMUNITARIANISM

Different communitarian camps are no closer to one another than National Socialists (Nazis) are to Scandinavian Social Democrats (also considered socialists). It is hence important to keep in mind which camp one is considering. The differences concern the normative relations between social order and liberty, and the relations between the community and the individual.

**Authoritarian Communitarians.** Authoritarian communitarians (some of whom are often referred to as “Asian” or “East Asian” communitarians) are those who argue that to maintain social order and harmony, individual rights and political liberties must be curtailed. Some believe in the strong arm of the state (such as former Singapore Prime Minister Lee Kuan Yew and Malaysian head of state Mahathir Bin Mohamad), and some in strong social bonds and the voice of the family and community (especially the kind of society Japan had, at

least until 1990). Among the arguments made by authoritarian communitarians is that social order is important to people, while what the West calls “liberty” actually amounts to social, political, and moral anarchy; that curbing legal and political rights is essential for rapid economic development; and that legal and political rights are a Western idea, which the West uses to harshly judge other cultures that have their own inherent values. The extent to which early sociological works, for instance, by Tönnies and *Community and Power* by Robert Nisbet, include authoritarian elements, is open to question.

**Political Theoreticians** . In the 1980s communitarian thinking became largely associated with three scholars: Charles Taylor, Michael Sandel, and Michael Walzer. They criticized liberalism for its failure to realize that people are socially “situated” or contextualized, and its negligence of the greater common good in favor of individualistic self-interests. In addition, as Chandran Kukathas relates in *The Communitarian Challenge to Liberalism* (Paul, Miller, Paul, eds. 1990, p.90), communitarians argue that political community is an important value which is neglected by liberal political theory. Liberalism, they contend, views political society as a supposedly neutral framework of rules within which a diversity of moral traditions coexist. . . . [Such a view] neglects the fact that people have, or can have, a strong and ‘deep’ attachment to their societies—to their nations.

While for many outside sociology, especially until 1990, these three scholars were considered the founding fathers of communitarian thinking, none of them uses the term in their work, possibly to avoid being confused with authoritarian communitarians. These scholars almost completely ignored sociological works that preceded them, and were largely ignored by sociologists.

**New or Responsive Communitarians.** Early in 1990, a school of communitarianism was founded in which sociologists played a key role, although it included scholars from other disciplines such as William A. Galston (political theory), Mary Ann Glendon (law), Thomas Spragens, Jr. (political science), and Alan Ehrenhalt (writer) to mention but a few. The group, founded by Amitai Etzioni, took communitarianism from a small and somewhat esoteric academic discipline and introduced



it into public life, and recast its academic content. Its tools were *The Responsive Communitarian Platform: Rights and Responsibilities*, a joint manifesto summarizing the guiding principles of the group; an intellectual quarterly, *The Responsive Community*, whose editors include several sociologists; several books; position papers on issues ranging from a communitarian view of the family to organ donation and to bicultural education; and numerous public conferences, op-eds, and a web site ([www.gwu.edu/~ccps](http://www.gwu.edu/~ccps)).

**Key Assumptions and Concepts.** Responsive communitarianism methodologically is based on the macro-sociological assumption that societies have multiple and not wholly compatible needs and values, in contrast to philosophies that derive their core assumptions from one overarching principle, for instance liberty for libertarianism. Responsive communitarianism assumes that a “good society” is based on a carefully crafted balance between liberty and social order, between individual rights and social responsibilities, between particularistic (ethnic, racial, communal) and society-wide values and bonds. In that sense, far from representing a Western model, the communitarian good society combines ‘Asian’ values (also reflecting tenets of Islam and Judaism that stress social responsibilities) with a Western concern with political liberty and individual rights.

While the model of the good society is applicable to all societies, communitarianism stresses that different societies, during various historical periods, may be off balance in rather contrasting ways and hence may need to move in *different* directions in order to approximate the *same* balance. Thus, contemporary East Asian societies require much greater tolerance for individual and communal differences, while in the American society—especially at the end of the 1980s—excessive individualism needs to be reigned in. To put it differently, communitarianism suggests that the specific normative directives that flow from the good society model are historically and culturally contingent.

Responsive communitarians stress that the relationship between liberty and social order is not a zero-sum situation; up to a point they are mutually supportive. Thus, in situations such as those prevailing in late-1990s Moscow, where liberty and social order are neglected, increasing order might well also enhance people’s autonomy and life

choices. The same might be said about reducing crime in American cities when it reached the point where people did not venture into parks, and were reluctant to ride the subway or walk the streets after dark. Moreover, totalitarian regimes, the ultimate loss of freedom, are said to arise when order is minimized.

While up to a point social order and liberty enhance one another, if the level of social order is increased further and further, responsive communitarians expect it to reach a level where it will erode people’s liberty. And, if the scope of liberty is extended ever more, it will reach a point where it will undermine the social order. This idea is expressed in the term *inverting symbiosis*, which indicates that up to a point liberty and order nourish one another, and beyond it they turn antagonistic.

The same point applies to the relationship between the self and the community. Political theorists have tended to depict the self as “encumbered,” “situated,” or “contextualized,” all of which imply that it is constrained by social order. Responsive communitarians stress that individuals within communities are able to be more reasonable and productive than isolated individuals, but if social pressure to conform reaches a high level, such pressures undermine the development and expression of the self.

The next question is: Under what conditions can the zone of symbiosis be expanded, and that of antagonism between liberty and order be minimized? To answer that question the communitarian view of human nature must be introduced. While sociologists tend to avoid this term, on the grounds that it is not testable and can lead to racism (as evident in the notion that some groups of people are more intelligent by nature), communitarians use the term with less reluctance.

The view of human nature most compatible with responsive communitarian thinking is a *dynamic* (developmental) view, which holds that people at birth are akin to animals. But unlike social conservatives, who tend to embrace a dour view of human nature, and tend to view even adults after socialization as impulsive, irrational, dangerous, or sinful—communitarians maintain that people can become increasingly virtuous if the proper processes of value-internalization and reinforcement of undergirding social institutions, the “moral

infrastructure,” are in place. At the same time, communitarians do not presume that people can be made as virtuous as liberals assume them to be from the onset. (Liberals tend to assume that crime and forms of deviant behavior reflect social conditions, especially government interventions that pervert good people, rather than criminals’ innate nature.)

The moral infrastructure, an essential foundation of a good society, draws on four social formations: families, schools, communities, and the community of communities. The four core elements of the moral infrastructure are arranged like Chinese nesting boxes, one within the other, and in a sociological progression. Infants are born into *families*, which communitarians stress have been entrusted throughout human history with beginning the process of instilling values and launching the moral self. *Schools* join the process as children grow older, further developing the moral self (“character”), or trying to remedy character neglect suffered under family care. Schools are hence viewed not merely or even primarily as places of teaching, where the passing of knowledge and skills occur, but as educational institutions in the broadest sense of the term.

Human nature, communitarians note, is such that even if children are reared in families dedicated to child raising and moral education, and children graduate from strong and dedicated schools, these youngsters are still not sufficiently equipped for a good, communitarian society. This is a point ignored by social philosophers who often assume that once people have acquired virtue and are habituated, they will be guided by their inner moral compass. The very concept of “conscience” assumes the formation of a perpetual inner gyroscope.

In contrast, communitarians—following standard sociological positions—assume that the good character of those who have acquired it tends to degrade. If left to their own devices, individuals gradually lose much of their commitments to their values, unless these are continuously reinforced. A major function of the *community*, as a building block of the moral infrastructure, is to reinforce the character of its members. This is achieved by the community’s “moral voice,” the informal sanction of others, built into a web of informal affect-laden relationships, which communities provide.

In general, the weaker the community—because of high population turnover, few shared core values, high heterogeneity, etc.—the thinner the social web and the slacker the moral voice. The strength of the moral voice and the values it speaks for have been studied using a series of questions such as, Should one speak up if child abuse is witnessed? Or if children are seen painting swastikas? What about less dire situations, such as insisting that friends wear their seatbelts, or admonishing a nondisabled person one witnesses parking in a handicap space?

Informal surveys show that Americans in the 1980s were very reluctant to raise their moral voice; many accepted the liberal ideology that what is morally sound is to be determined by each individual, and one should not pass judgments over others. Alan Wolfe’s study, *One Nation After All*, found that Americans, even in conservative parts of the country, have grown very tolerant of a great variety of social behavior. Increase in tolerance is of course by itself virtuous; communitarians, though, raise the question: At which point does such increased tolerance engendering an amoral culture where spousal abuse, discrimination, child neglect, drunk drivers, obsessive materialism, and other forms of antisocial behavior become matters the community should ignore, leaving them to individual discretion or the law.

More specifically, communitarians inquire various elements of the moral infrastructure whether they reinforce, neglect, or undermine it. In this context, the special communitarian perspective of voluntary associations is especially important. Previously, the significance of these associations has been highlighted as protecting individuals from the state (a protection they would not have if they faced the state as isolated or “atomized” individuals), and as intermediating bodies that aggregate, transmit, and underwrite individual signals to the state.

Communitarians argue that, in addition, the very same voluntary associations often fulfill a rather different role: They serve as social spaces in which members of communities reinforce their social webs and articulate their moral voice. That is, voluntary associations often constitute a basis of communal relationships. Thus, the members of a local chapter of the Masons, Elks, or Lions care

about one another and reinforce each other's particular brand of conservative views. Similarly, the members of the New York City Reform Clubs, Americans for Democratic Action, and local chapters of the ACLU reinforce one another's particular brand of liberal views.

Communitarians pay special attention to the condition of public spaces as places communities happen (as distinct from private places like homes and cars). Even though one may carpool with friends or have them over for a visit, these are mainly activities of small friendship groups (what Robert Putnam calls "bowling alone"). Communities need more encompassing webs, and those are formed and reinforced in public gathering places—from school assembly halls to parks, from plazas to promenades. To the extent that these spaces become unsafe, communities lose one of their major sources of reinforcement; recapturing them for community use is hence a major element of community regeneration.

Most important, drawing again on sociology, and particularly on what has been called the "consensus" rather than the "conflict" model, communitarians tend to maintain that if in addition to strong families and schools that build character, a society has communities, where social webs are intact and that moral voice is clearly articulate, that society will be able to base its social order largely on moral commitments rather than the forces of the state. This is the case, communitarians argue, drawing on sociological assumptions and studies, because once moral commitments are internalized and reinforced they help shape people's preferences in favor of prosocial behavior—thus reducing the need for coercion by the state and diminishing the tension between liberty and social order.

Many discussions of community and of the moral infrastructure stop at this point, having explored the moral agency of family, school, and community. However, social and moral communities are not freestanding; they are often parts of more encompassing social entities. Moreover, communitarians note that unless communities are bound socially and morally into more encompassing entities, they may war with one another. Hence, the importance of *communities of communities*, the society.

Communitarians argue that one should not view society as composed of millions of individuals, but as pluralism within unity. They further maintain that subcultures and loyalties are not a threat to the integrity of society as long as a core of shared values and institutions (such as the Constitution and its Bill of Rights, the democratic way of life, and mutual tolerance) are respected.

Communitarians draw on the four elements of the moral infrastructure—families, schools, communities, and communities of communities—as a sort of a checklist to help determine the state of the moral infrastructure in a given society. They argue that the decline of the two-parent family (due to high divorce rates, growing legitimation of single-parent families, and psychological disinvestment of parents in children), the deterioration of schools (due to automatic promotions and deterioration of social order in schools), the decline of communities (due to modernization), and the decline of the community of communities (as a result of excessive emphasis on diversity without parallel concern over shared bonds) resulted in the decline of moral and social order in the American society during the 1970s and 1980s. This was evidenced by the sharp rise in violent crime, drug abuse, teen pregnancy, and in the decline of voluntarism, among other factors. The fact that some of these trends slowed down and reversed in the 1990s is viewed in part by communitarians as a reflection of changes in social thinking and practices they helped champion.

Here lies a great difference between the communitarian position and that of various religious social conservatives. Both groups recognize the need to regenerate the moral infrastructure, but conservatives favor returning to traditional social formations while communitarians point to new ways of shoring up society's ethical framework. For instance, many social conservatives favor women "graciously submitting to their husbands" and returning to homemaking, while communitarians argue for peer marriage, a concept introduced by Pepper Schwartz. Peer marriage suggests equal rights and responsibilities for mothers and fathers, but favors marriages that last, as compared to the liberal argument that single-parent families or child care centers can socialize children as well if not better than two-parent families. (Among the sociologists who have struck a

communitarian position in this matter are Linda Waite, Glen Elder, Alice Rossi, and David Popenoe).

The communitarian argument over the role of communities in maintaining social order is strongly supported by sociological research of the kind conducted by Robert Sampson on the role of communities in fighting crime and drug abuse. David Karp and Todd Clear have also studied community involvement in criminal justice, focusing on ideas of restorative justice and policies that are concerned more with reintegrating offenders into their communities than merely punishing them.

Other communitarian themes examined by sociologists include topics explored by Edward W. Lehman, especially his writing on macro-sociology; Martin Whyte's work on the family; and Richard Coughlin's comparison of communitarian thinking to socioeconomics.

#### **CIVIL SOCIETY, THE THIRD WAY, AND THE GOOD SOCIETY**

Much of the normative debate in the West, at least since the middle of the nineteenth century, has focused on the merit of the free market (or capitalism) versus the role of the state in securing the citizens' well-being. Communitarians have basically leapfrogged this debate, focusing instead on the importance of the third element of social life, that of the civic society, which is neither state nor market. Communitarians have played a key role in the debate over the condition of civic society in the West, such as examining whether participation in voluntary associations, voting, and trust in institutions have declined, and to what effect. The work of Robert Bellah and his associates has been particularly influential here, demonstrating the rise of first expressive and then instrumental individualism, and their ill effects.

Communitarians have argued that rather than dumping people (often the most vulnerable members of society) into the marketplace as the welfare state is curtailed, civic society's various institutions can empower these individuals to help one another in attending to some of their social needs. Communal institutions (including places of worship) can shoulder important parts of care previously provided by state agencies, although the state will have to continue to shoulder an important part of the burden.

Communitarians stress that mutuality, rather than charity, is the basis for community-wide action that is not solely limited to helping one particular vulnerable group or another. The CPR training of some 400,000 Seattle citizens, who are thus able to help one another without public costs or private charges, is held up as a key case in point. Other examples include voluntary recycling programs, crime watch patrols, and above all the massive assistance given to immigrants by members of their own ethnic group. Communitarians have also pointed to the importance of a culture of civility in maintaining a society's ability to work out differences without excessive conflict.

Communitarians have argued that a civic society is good, but not good enough. Civic society tends to be morally neutral on many matters other than values concerning its own inherent virtue and the attributes citizens need to make them into effective members of a civic society, for instance, to be able to think critically. Thus, all voluntary associations, from the KKK to the Urban League, from militias to Hadassah, are considered to have the same basic standing. In contrast, a good society seeks to promote a core of substantive values, and thus views some social associations and activities as more virtuous than others. In the same vein, communitarians have stressed that while everyone's legal right to free speech should be respected, there is no denying that some speech—seen from the community's viewpoint—is morally sound while other speech is abhorrent. For instance, the (legal) right to speak does not make hate speech (morally) right. Communitarians would not seek to suppress hate speech by legal means, however, but they urge communities to draw on their moral voice to chastise those who speak in ways that are offensive.

#### **CRITICS AND RESPONSES**

Critics of responsive communitarianism argue that the concept of community is vague; indeed that the term "community" itself cannot be well defined. In response, *community* has been defined as a combination of two elements: a) A web of affect-laden relationships among a group of individuals, relationships that often crisscross and reinforce one another (rather than merely one-on-one or chainlike individual relationships) and b) A measure of commitment to a set of shared values,

norms, and meanings, and a shared history and identity to a particular culture.

Some critics also contend that the quest for community is anachronistic, that contemporary societies are urban and populations geographically highly mobile, and thus bereft of community. Communitarians respond that communities exist in contemporary societies in small towns, suburbs, campuses and within city neighborhoods, often based on ethnic ties in places such as Korea Town in Los Angeles, Little Italy in New York City, the Irish section of South Boston, and so on. Moreover, communitarians point out that communities need not be geographic, members can be spread among nonmembers. For instance, homosexual groups often constitute communities even if they are not all neighbors.

Critics maintain that communities are authoritarian and oppressive, and have charged communitarians with seeking "Salem without witches." Communitarians respond that communities vary regarding this assessment. Contemporary communities tend to be relatively freer, given the relative ease of intercommunity mobility as well as shifting loyalties and psychic investments among various groups of which the same person is a member.

Critics also maintain that communities are exclusionary, and hence bigoted. Communitarians respond that communities must respect the laws of the society in which they are situated, but do tend to thrive on a measure of homogeneity and on people's desire to be with others of their own kind. Moreover, given the human benefit of community membership, a measure of self-segregation should be tolerated.

Critics of responsive communitarianism claim that communitarians ignore matters of power and injustice as well as economic considerations, and are generally inclined to adopt a consensus rather than a conflict model. Communitarians agree that they ought to pay more attention to the effects of these factors on communities. However, they do envision the possibilities of conflict within communities, and responsive communitarians do propose that one should not treat conflict and community as mutually exclusive. Extending this idea to the treatment of diversity and multiculturalism, communitarians argue in favor of a society in which many differences can be celebrated as long

as a set of commitments to the overarching society is upheld.

### VALUES AND VIRTUES

While sociologists greatly altered and enriched communitarian thinking, communitarian thinking's main contribution to sociology is the challenge of facing issues raised by the moral standing of various values, and the related question of cross-cultural moral judgments. Sociologists tend to treat all values as conceptually equal; a sociologist may refer to racist Afrikaners' beliefs and to humanitarian beliefs using the same "neutral" term, calling both "values." Communitarians use the term virtue to denote that some values (or belief systems) have a higher standing than others because they are compatible with the good society, while other values are not (and hence aberrant).

In the same vein, communitarians do not shy away from passing cross-cultural moral judgments, rejecting cultural relativism's claim that all cultures have basically equal moral standing. Thus, they view female circumcision, sex slaves, and *hudud* (chopping off the right hand of thieves) as violations of liberty and individual rights, and abandoning children, violating implicit contracts built into communal mutuality, or neglecting the environment, as evidence of a lack of commitment to social order and neglect of social responsibilities.

### IMPACT

So far this examination has focused on the place of communitarian thinking in academic, conceptual, theoretical, and imperial works. Responsive communitarians have also been playing a considerable public role, presenting themselves as the founders of a new kind of environmental movement, one dedicated to shoring up society (not the state) rather than nature. Like environmentalism, communitarianism appeals to audiences across the political spectrum, although it has found greater acceptance with some groups rather than others. British Prime Minister Tony Blair is reported to have adopted the communitarian platform, and German Social Democrat Rudolf Scharping has suggested that his party should meet the communitarians "half way." President Bill Clinton and his wife Hillary Rodham Clinton (author of *It*

*Takes a Village*) have combined communitarian with welfare-liberal themes. Among conservatives, Jack Kemp, a group of Tory members of the British parliament (especially David Willet), and German governor Kurt Biedenkopf are often listed as influenced by communitarianism. While this is only a partial list, it serves to illustrate the scope of communitarianism's influence and its cross over traditional ideological lines.

Communitarian terms have become part of the public vocabulary in the 1990s, especially references to "assuming responsibilities to match rights," while "communitarianism" itself is used much less often. The number of articles about communitarian thinking in the popular press increased twelvefold during the last decade of the twentieth century. The increase in the number of books, articles, and Ph.D. dissertations in academia for the same period, has been about eightfold. Interestingly, taking communitarianism from academia to the "streets" in the early 1990s resulted in the middle and late 1990s into a significant increase in academic interest. In the process, bridges have been built between social philosophers, sociologists, and community members and leaders, although they still sometimes travel on parallel rather than convergent pathways.

*An extensive bibliography of communitarian works is also listed on The Communitarian Network's website, <http://www.gwu.edu/~ccps>.*

## REFERENCES

- Bellah, Robert et al. 1991 *The Good Society*. New York: Knopf.
- 1986 *Habits of the Heart*. New York: Harper & Row.
- Coughlin, Richard 1991 *Morality, Rationality, and Efficiency: New Perspectives on Socio-economics*. Armonk, N.Y.: M.E. Sharpe.
- Etzioni, Amitai 1998 *The Essential Communitarian Reader*. New York: Roman & Littlefield.
- 1996 *The New Golden Rule*. New York: Basic Books.
- 1995 *New Communitarian Thinking*. Charlottesville, Va.: The University of Virginia Press.
- 1993 *The Spirit of Community*. New York: Touchstone.
- Paul, Ellen Franken, Fred Miller, Jr., and Jeffrey Paul (eds.) 1996 *The Communitarian Challenge to Liberalism*. New York: Cambridge University Press.
- Selznick, Philip 1992 *The Moral Commonwealth*. Berkeley: University of California Press.
- The Responsive Community: Rights and Responsibilities*, a quarterly journal.

AMITAI ETZIONI

## COMMUNITY

The sociology of community has been a dominant source of sociological inquiry since the earliest days of the discipline. Each of the three most influential nineteenth century sociologists (Marx, Durkheim, and Weber) regarded the social transformation of community in its various forms to be a fundamental problem of sociology and sociological theory. Thomas Bender (1978) suggests that as early social thinkers observed the disruption of the traditional social order and traditional patterns of social life associated with industrialization, urbanization, and the rise of capitalism, significant attention was focused on the social transformation of community and communal life. It should be emphasized that contemporary sociology remains, at its core, a discipline largely concerned with the definition and persistence of community as a form of social organization, social existence, and social experience.

The definition of community in sociology has been problematic for several reasons, not the least of which has been nostalgic attachment to the idealized notion that community is embodied in the village or small town where human associations are characterized as *Gemeinschaft*: that is, associations that are intimate, familiar, sympathetic, mutually interdependent, and reflective of a shared social consciousness (in contrast to relationships that are *Gesellschaft*—casual, transitory, without emotional investment, and based on self-interest). According to this traditional concept of community, the requirements of community or communal existence can be met only in the context of a certain quality of human association occurring within the confines of limited, shared physical territory.

The classic perspective on community offered by Carle Zimmerman (1938) is consistent with this theme, in that the basic four characteristics argued by Zimmerman to define community (social fact,

specification, association, and limited area) require a territorial context. George Hillary (1955), in a content analysis of ninety-four definitions of community advanced in sociological literature, discovered basic consensus on only three definitional elements: social interaction between people, one or more shared ties, and an area context. However, Hillary noted that area context was the least required of these three definitional elements. Others (e.g., Lindeman 1930; Bender 1978; McMillan and Chavis 1986) argue that community can be achieved independently of territorial context where social networks exist sufficiently to sustain a *Gemeinschaft* quality of interaction and association. According to this point of view, territory is neither a necessary nor a sufficient condition to define the existence of community. In this vein, David McMillan and David Chavis suggest a state of community exists when four elements co-exist: membership, influence, integration and fulfillment of needs, and shared emotional connections. They argue that communities can be defined either in relational terms or territorial terms as long as these four elements are present together.

### MAJOR QUESTIONS IN THE SOCIOLOGY OF COMMUNITY

The major questions that concern the sociology of community include the distinguishing characteristics and definition of community, the bases of communal experience and integration, the unique functions and tasks of community, the units of social structure within the community and the relationships and interactions between structural units, the economic and social bases of the community social structure, the relationship and distinction between internal community social structure and macrosocial structures external to the community, the relationship between individual experience and behavior and communal experience and behavior, the causes and processes of transformation from *Gemeinschaft* to *Gesellschaft* states of social existence, and processes of community persistence and adaptation in the face of social change.

### COMMUNITY STUDIES

Community studies undertaken by sociologists over the past sixty years have to a large extent

sought to address some if not all of these issues. The most famous and controversial include Robert and Helen Lynd's Middletown studies (1929, 1937) and the Yankee City series by W. Lloyd Warner and his associates (1963). The more well-known studies that have focused on the problem of community within large cities have included William Whyte's *Street Corner Society* and Gerald Suttles's *The Social Order of the Slum*, which themselves are aligned with earlier Robert Park and Ernest Burgess conceptions of the "natural community" arising within the confines of a seemingly faceless, anonymous, large city (Suttles 1972, pp. 7-9). Descriptive studies that emphasize field work and examine social structure as a spatial phenomenon are the hallmark of the highly influential Chicago School that arose and flourished under Robert Park and Ernest Burgess of the University of Chicago Department of Sociology during the 1920s and 1930s.

Robert and Helen Lynd carried out two studies (1929, 1937) involving extensive personal field work on the town of Muncie, Indiana: *Middletown* (1929) and *Middletown in Transition* (1937). In the first Middletown study, the Lynds spent the years 1924-1925 participating in and observing the community life of Middletown (population 36,500), and performing extensive survey work. Their objective was to address all aspects of social life and social structure of the community. A fundamental focus of the Lynds' earlier analysis concerned the consequences of technological change (industrialization) on the social structure of Middletown, in particular the emergence of social class conflicts subsequent to turn-of-the-century industrialization.

Although the Lynds found distinctions between the living conditions and opportunity structures of business and working-class families that were consistent with their Marxist expectations (i.e., children of the working class were more likely to drop out of school to help support the family, working-class families labored longer hours for less pay and less financial security, and living conditions in general were more harsh for working-class families), they failed to discover a disparate value structure or alienation among the working class. At all levels of social class, the Middletown of 1925 shared a common conservative value structure that entailed self-reliance, faith in the future, and a belief in hard work. The subsequent study, undertaken in 1935 by Robert Lynd and a staff of

five assistants, addressed the effects on Muncie of certain events during the period between 1925 and 1935, some of which were economic boom times, a thirty-seven percent population increase, and the emergence of the Great Depression. The Lynds' fundamental questions in the later study addressed the persistence of the social fabric and culture of the community in the face of the "hard times" and other aspects of social change, the stability of community values concerning self-reliance and faith in the future when confronted by structurally-induced poverty and dependence, whether the Depression promoted a sense of community or undermined community solidarity by introducing new social cleavages, and the outcomes of latent conflicts observed in the mid-1920s (Lynd and Lynd 1937, p. 4).

The conclusion reached by the Lynds was that the years of depression did little to diminish or otherwise change the essentially bourgeois value structure and way of life in Middletown, and that in almost all fundamental respects the community culture of Middletown remained much as it did a tumultuous decade earlier: "In the main, a Rip Van Winkle, fallen asleep in 1925 while addressing Rotary or the Central Labor Union, could have awakened in 1935 and gone right on with his interrupted address to the same people with much the same ideas" (Lynd and Lynd 1937, p. 490). Although this remark seems to reflect some amount of disappointment on the Lynds' part that Middletown's bourgeois value system and class structure remained so unchanged in the face of widespread and unprecedented destitution, the Lynds still remained convinced that the Middletown family was in jeopardy, as evidenced by (among other things) an ever-widening generation gap. The Lynds' apprehensions concerning the survival of the American family were (and are) in keeping with the popular belief concerning the decline of the American family as its socialization functions are assumed by other formal social institutions external to the family.

The Middletown III study undertaken from 1976 to 1978 by Theodore Caplow, Howard M. Bahr, and Bruce A. Chadwick, attempted to closely replicate the methodology of the Lynds. However, their findings were at odds with the more pessimistic predictions of the Lynds. In contrast to the Lynds' foreboding in 1935 and popular sociology since that time, Caplow and his associates

contend that Middletown's families of the 1970s have "increased family solidarity, a smaller generation gap, closer marital communication, more religion, and less mobility" (Caplow et al. 1982, p. 323). The conclusions derived from the third Middletown studies also reject similar assumptions concerning consistent linear trends in equalization, secularization, bureaucratization, and depersonalization consistent with the relentless *Gemeinschaft* to *Gesellschaft* theme (Bahr, Caplow, and Chadwick 1983).

Although many of the Lynds' predictions concerning the social transformation of Middletown failed to pan out as history unfolded, their work and remarkable powers of observation remain unparalleled in many respects. Of equal importance, the early Middletown studies helped to inspire such other works as *Street Corner Society* and the Yankee City studies, and they remain the standard by which all other community studies are judged.

The largest scale community study undertaken remains W. Lloyd Warner's Yankee City, published in a five volume series from 1941 through 1959 (W. Lloyd Warner and Paul S. Lund, *The Social Life of a Modern Community* 1941; W. Lloyd Warner and Paul S. Lund, *The Status System of a Modern Community* 1942; W. Lloyd Warner and Leo Srole, *The Social Systems of American Ethnic Groups* 1945; W. Lloyd Warner and J. O. Low, *The Social System of the Modern Factory* 1947; and W. Lloyd Warner *The Living and the Dead* 1959). The Yankee City project was undertaken by Warner and his associates in Newburyport, Massachusetts during the late 1930s. Warner, an anthropologist, attempted to obtain a complete ethnographic account of a "representative" American small community with a population range from 10,000 to 20,000. To accomplish this task, Warner's staff (numbering in the thirties) conducted aerial surveys of Newburyport and its surrounding communities, gathered some 17,000 "social personality" cards on every member of the community, gathered data on the professed and de facto reading preferences of its citizens, and even subjected plots of local plays to content analysis (Thernstrom 1964).

Warner's conception of Yankee City was that of a stable, rather closed community with a social structure being transformed in very negative ways



by the latter stages of industrialization. According to Warner's vision of Yankee City, the loss of local economic control over its industries through a factory system controlled by "outsiders" disrupted traditional management-labor relations and communal identification with local leadership. Moreover, the factory system was seen by Warner as promoting an increasingly rigid class structure and decreased opportunities for social mobility. In particular, Warner's discussion of the loss of local economic control through horizontal and vertical affiliation, orientation, and delegation of authority seems to have offered a prophetic glimpse into the future for many American communities.

Although the Yankee City study produced a voluminous ethnographic record of an American city that has remained untouched in scale, Warner found little support for his contention that the ethnographic portrait of Newburyport produced by the Yankee City series could be generalized to other small American communities. Moreover, Warner's contention that social mobility is reduced by industrial change was not supported by the quality of his data and was less true in Warner's time than it probably is today. Other critiques of Warner's Yankee City study primarily concern his nearly exclusive reliance on ethnographic information as the basis for all measures of social structure and his disdain for historical data (Thernstrom 1964).

Both Gerald Suttles's *The Social Order of the Slum* (1968) and William Whyte's *Street Corner Society* (1943) provide sociology with unparalleled ethnographic accounts of neighborhood social structure and communal life in urban environs. Suttles's work focuses on the territorial relationships, neighborhood social structure, and communal life among Italian, Latino, and African American inhabitants of the slums of Chicago's Near West Side in the 1960s. Whyte's *Street Corner Society*, based on Whyte's residence in a Chicago Italian slum district a generation earlier, provides sociology with an understanding of the complex and stable social organization that existed within slum neighborhoods conventionally believed to have epitomized social disorganization. Whyte's observations and keen insights concerning small group behavior are pioneering contributions to that area of sociology. Both studies, in method, theory, and substance are classic examples of Chicago School

sociology. More contemporary community studies, while they draw heavily on Chicago School sociological traditions in theory and method, have as their primary concern aspects of community that relate to poverty, juvenile delinquency, and violence. For example, Robert Sampson and his colleagues conducted a large-scale study of Chicago neighborhoods that linked subjective definitions of neighborhood with social cohesion and violence inhibiting actions on the part of neighbors (Sampson 1997a, 1997b). Theoretically, Sampson's work draws heavily on Robert Park's conception of the social cohesion that exists within the "natural areas" of the city that are defined by both physical and sentimental boundaries (Park 1925), while methodologically Sampson and his colleagues employed the classic Chicago School preference for field observation—albeit with the modern advantages of a video camera located within a slowly moving van in place of the shoe leather sociology of their predecessors.

#### COMMUNITY IN THE CONTEXT OF SOCIAL REFORM

Efforts to enhance the social context of human existence continuously take place at all levels of social organization, from the microcontext of the nuclear family to the macrocontext of relationships between nation-states. No unit of social organization has received more attention in theories and activities linked to social reform than the human community, however it is defined and measured. Community in the context of social reform is typically viewed or employed in one of the following four ways: as a unit of analysis for the purpose of broad generalization, as a critical mediating influence between the organization of mass society and individual outcomes, as a specific target of social reform efforts, or as a symbolic conception of whatever is "right" or "wrong" with society at large.

The use of rural and urban communities as the basis for reform-motivated generalization emerged in the United States during the Progressive era, when social activists affiliated with Chicago's famous settlement house, Hull House, and the newly formed Department of Sociology at the University of Chicago shared a common geographically-rooted conception of social science and concern for the living conditions of Chicago's urban

poor (Sklar 1998). *Hull House Maps and Papers*, published in 1895, provides a remarkably detailed description of the lives and living conditions within Chicago's ethnic neighborhoods based on the field research of university students and full-time residents of Hull House. In fact, the beginnings of Chicago School sociology were clearly rooted in the methods if not the concerns of Hull House social reformers, despite the fact that as women they were excluded from holding academic appointments until the creation of the Chicago School of Civics and Philanthropy in 1907 (Muncy 1991).

In later years, small towns and neighborhoods within larger towns became a common unit of analysis as community studies conducted by the federal Children's Bureau tried to assess the incidence and causes of infant mortality. The first of these studies, conducted in Johnstown, Pennsylvania in 1913, was the earliest scientific study of the incidence of infant mortality by social class, education, occupation, and specific living conditions conducted in the United States (Duke 1915). Reform-driven community studies since then have addressed such social issues as child labor, juvenile delinquency, education, industrial working conditions, and adequate housing. Although single communities are still used by social reformers as a basis for broad generalization, such studies are enormously expensive to conduct, often of limited use for generalization, and altogether less frequent with the availability of national survey data.

Despite the conceptual ambiguities involved, the geographically-bounded community (e.g., census tract, city ward, resident defined neighborhood) is generally viewed by liberal and conservative social reformers alike as the critical mediating social context between the well-being of individuals, the effective functioning and stability of families, and society at large. Robert Hauser and his associates identify such factors as the physical infrastructure, the quality and quantity of neighborhood institutions, the demographic composition, and the degree of "social capital" present as measurable aspects of neighborhoods that are critical to the well-being of children and families (Hauser, Brown, and Prosser 1997). The concept of social capital, introduced by James Coleman, refers to the beneficial normative context that arises in some neighborhoods based on the social ties among neighboring households and local institutions (Coleman 1988). In this vein, Robert

Sampson demonstrates that neighborhoods with evidence of more social capital are more effective in inhibiting adolescent delinquency (Sampson, 1997b). Speaking from the opposite perspective of social disorganization, William Julius Wilson (1987, 1996), and Douglas Massey and Nancy Denton (1993) emphasize the deleterious effects on the normative context within communities created by such macrolevel exogenous factors as economic restructuring, concentrated poverty, and racial segregation. A related concern is the extent to which increased spatial stratification by social class may be promoting neighborhoods and larger communities, which function as distinct social worlds with ever more divergent values and opportunity structures (Massey 1996).

Efforts to effect social reform at the community level have a long tradition of sentiment, failure, and mixed success. There are a variety of reasons for failure, not the least of which is that communities, like individuals, both mirror and are shaped by complex exogenous social processes. Contemporary issues include conceptual differences in the measures used for community, the fact that community effects on individual outcomes are difficult to isolate and often weaker than popular theory would suggest (Plotnick and Hoffman 1999; Brooks-Gunn, Duncan, and Aber 1997), and the limiting effect of macrolevel processes on community-level reform efforts (Wallace and Wallace 1990; Halpern 1991). Recent efforts at community-level social reform have placed more emphasis on understanding the unique social context within a target community before applying social prescriptions that appear logically appealing or may have worked elsewhere. For example, David Hawkins and Richard Catalano, in their work on juvenile drug and alcohol abuse, focus on a community assessment process that considers the risk and protective factors that are unique to each community before deciding upon specific community-level interventions (Hawkins and Catalano 1992).

#### **SOCIAL THEORY AND THE TRANSFORMATION OF COMMUNITY**

Every generation of sociologists since the time of Durkheim have concerned themselves with the social transformation and meaning of community in the face of industrial change and urbanization.

Roland Warren (1978) describes the modern social transformation of community as a change of orientation by the local community units toward the extracommunity systems of which they are a part, with a corresponding decrease in community cohesion and autonomy (Warren 1978, pp. 52–53). Warren identifies seven areas through which social transformation can be analyzed: division of labor, differentiation of interests and association, increasing systemic relationships to the larger society, bureaucratization and impersonalization, transfer of functions to profit enterprise and government, urbanization and suburbanization, and changing value.

Bender (1978) proposes that the observations by various community scholars at different points in historical time, each suggesting that theirs is the historical tipping point from community to mass society, contradict linear decline or an interpretation of history that stresses the collapse of community. He suggests that a “bifurcation of social experience” or sharpening of the distinction between *Gemeinschaft* and *Gesellschaft* realms of social interaction is a more accurate interpretation of the historical transformation of community than that provided by the linear *Gemeinschaft-to-gesellschaft* framework.

Barry Wellman and Barry Leighton (1979) suggest that there are three essential arguments concerning the fate of community in mass society: community “lost,” community “saved,” and community “liberated.” The community “lost” argument emerged during the Industrial Revolution, as traditional communal modes of production and interaction gave way to centralized, industrialized sources of production and dependence. According to this hypothesis and its variations, the intimate, sustained, and mutually interdependent human associations based on shared fate and shared consciousness observed in traditional communal society are relentlessly giving way to the casual, impersonal, transitory, and instrumental relationships based on self-interest that are characteristic of social existence in modern industrial society.

The classic essay in the community “lost” tradition is Louis Wirth’s “Urbanism as a Way of Life” (1938). Wirth’s eloquent essay presents a perspective of urban existence that continues to capture sociological thinking about the emergence

of a heterogeneous urban mass society characterized by a breakdown of informal, communal ways of meeting human need and the rise of human relationships that are best characterized as “largely anonymous, superficial, and transitory” (Wirth 1938, p. 1). Another important contribution in the decline of community tradition is the “community of limited liability” thesis (Janowitz 1952; Greer 1962). According to this thesis, networks of human association and interdependence exist at various levels of social organization, and there are social status characteristics associated with differentiated levels of participation in community life (e.g., family life-cycle phase). The idea of “limited liability” poses the argument that, in a highly mobile society, the attachments to community tend to be based on rationalism rather than on sentiment and that even those “invested” in the community are limited in their sense of personal commitment.

In direct contrast to the community “lost” perspective, the community “saved” argument suggests that communities and communal relationships continue to exist within industrialized bureaucratic urban societies as people are increasingly motivated to seek “safe communal havens” (William and Lieghton 1979, p. 373). For example, Bahr, Caplow, and Chadwick (1983), forty years after the Lynds’ *Middletown in Transition* study, failed to find the singular trends in bureaucratization, secularization, mobility, and depersonalization that would be predicted from a linear decline of community hypothesis. Their observation of Middletown in the 1970s was more consistent with the perspective proposed by Robert Redfield (1955); that both urban ways and folkways coexist within contemporary small towns and cities: “In every isolated little community there is civilization; in every city there is the folk society” (Redfield 1955, p. 146).

The community “liberated” argument concedes and to some extent qualifies key aspects of both the community “lost” and community “saved” perspectives. While it acknowledges that neighborhood-level communal ties have been weakened in the face of urbanization, it argues that communal ties and folkways still flourish, albeit in alternative non-spatial forms. The community “liberated” argument suggests that the spatial dependence of communal ties have been replaced by ease of mobility and communication across boundaries of

both geographic and social distance. Although the community “liberated” argument preceeded the development of the Internet and cyberspace “chatrooms” by decades, it emphasizes the role of communication technology in the creation of such future manifestations of community and in that sense was strikingly prophetic.

Despite the fact that the decline of community or the community “lost” perspective continues to hold broad appeal (e.g., Robert Putnam’s “The Strange Disappearance of Civic America,” *The American Prospect*, Winter, 1996), the empirical evidence for a real decline in community is far from conclusive. A time trend analysis of national surveys concerning the persistence of community in American society by Avery Guest and Susan Wierzbicki (1999) suggests that while traditional intra-neighborhood forms of socializing have slowly declined over the past two decades, they have been largely replaced by communal ties outside the local neighborhood. Moreover, Guest and Wierzbicki’s findings also suggest that spatial “with-in neighborhood” communal ties continue to be important to a large segment of the population. Their findings imply that the major questions about the place of community in mass society should not be about whether communal forms of relationships will continue to exist, but rather under what conditions and in what form.

## REFERENCES

- Bahr, Howard M., Theodore Caplow, and Bruce Chadwick 1983 “Middletown III: Problems of Replication, Longitudinal Measurement, and Triangulation.” *Annual Review of Sociology* 9:249–258.
- Bender, Thomas 1978 *Community and Social Change in America*. New Brunswick, N.J.: Rutgers University Press.
- Brooks-Gunn, Jeanne, Greg Duncan, and Lawrence Aber 1997 *Neighborhood Poverty: Context and Consequences for Children, Volume I*. New York: Russell Sage Foundation.
- Caplow, Theodore, Howard M. Bahr, Bruce A. Chadwick, Reuben Hill, and Margaret Holmes Williamson 1982 *Middletown Families: Fifty Years of Change and Continuity*. Minneapolis: University of Minnesota Press.
- Coleman, James 1988 “Social Capital in the Creation of Human Capital.” *American Journal of Sociology*, 94 (supp): S95–S120.
- Duke, Emma 1914 *Infant Mortality: Results of a Field Study in Johnstown, P.A., Based on Births in One Calendar Year*. U.S.D.L. Children’s Bureau Publication No. 9. Washington D.C.: U.S. Government Printing Office.
- Greer, Scott 1962 *The Emerging City: Myth and Reality*. New York: Free Press of Glencoe.
- Guest, Avery, and Susan Wierzbicki 1999 “Social Ties at the Neighborhood Level: Two Decades of GSS Evidence.” *Urban Affairs Review* 35: forthcoming.
- Halpern, Robert 1991 “Supportive Services for Families in Poverty: Dilemmas of Reform.” *Social Service Review* 65 (3):343–365.
- Hauser, Robert, Brett Brown, and William Prosser 1997 *Indicators of Children’s Well-Being*. New York: Russell Sage Foundation.
- Hillary, George A., Jr. 1955 “Definitions of Community: Areas of Agreement.” *Rural Sociology* 20:111–123.
- Janowitz, Morris 1952 *The Community Press in an Urban Setting: The Social Elements of Urbanism*. Chicago: University of Chicago Press.
- Jencks, Christopher and Susan Mayer 1996 “The Social Consequences of Growing Up in a Poor Neighborhood.” In Lynn Le and McGeary MGH, eds., *Inner-city Poverty in the United States*. Washington D.C.: National Academy Press.
- Lindeman, E. C. 1930 “Community.” In Edwin R. A. Seligman, ed., *The Encyclopedia of Social Sciences*. New York: Macmillan.
- Lynd, Robert, and Helen Lynd 1929 *Middletown: A Study in American Culture*. New York: Harcourt, Brace and Company.
- 1937 *Middletown in Transition: A Study in Cultural Conflicts*. New York: Harcourt, Brace and Company.
- Massey, Douglas 1996 “The Age of Extremes: Concentrated Affluence and Poverty in the Twenty-First Century.” *Demography* 33 (4):395–413.
- and Nancy Denton 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- McMillan, David and David Chavis 1986 “Sense of Community: A Definition and Theory.” *Journal of Community Psychology* 14:6–23.
- Muncy, Robyn 1991 *Creating a Female Dominion in American Reform 1890–1935*. New York: Oxford University Press.
- Park, Robert 1925 “The City: Suggestions for the Investigation of Human Behavior in the Urban Environment.” In Robert Park, Ernest Burgess, and Roderick McKenzie, eds., *The City*. Chicago: University of Chicago Press.

- Plotnick, Robert and Saul Hoffman 1999 "The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates." *Social Science Quarterly* 80 (1):forthcoming.
- Putnam, Robert 1996 "The Strange Disappearance of Civic America." *The American Prospect*. Winter, 1996: 34-48.
- Redfield, Robert 1955 *The Little Community*. Chicago: University of Chicago Press.
- Sampson, Robert 1997a "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918-925.
- 1997b "Collective Regulation of Adolescent Misbehavior: Validation Results from Eighty Chicago Neighborhoods." *Journal of Adolescent Research* 12 (2):227-244.
- Sklar, Kathryn 1998 "Hull-House Maps and Papers: Social Science as Women's Work in the 1890's." In Helene Silverberg, ed., *Gender and American Social Science: The Formative Years*. Princeton, N.J.: Princeton University Press.
- Suttles, Gerald 1972 *The Social Construction of Communities*. Chicago: University of Chicago Press.
- 1968 *The Social Order of the Slum: Ethnicity and Territory in the Inner City*. Chicago: The University of Chicago Press.
- Thernstrom, Stephen 1964 *Poverty and Progress*. Cambridge, Mass.: Harvard University Press.
- Wallace, Roderick, and Deborah Wallace 1990 "Origins of Public Health Collapse in New York City: The Dynamics of Planned Shrinkage, Contagious Urban Decay, and Social Disintegration." *Bulletin of the New York Academy of Medicine* 66:391-434.
- Warner, W. Lloyd 1959 *The Living and the Dead: A Study of the Symbolic Life of Americans*. New Haven: Yale University Press.
- , and J.O. Low 1947 *The Social System of the Modern Factory: The Strike, A Social Analysis*. New Haven: Yale University Press.
- , Paul S. Lund, and Leo Srole 1963 *Yankee City*. New Haven: Yale University Press.
- Warner, W. Lloyd, and Leo Srole 1945 *The Social Systems of American Ethnic Groups*. New Haven: Yale University Press.
- Warner, W. Lloyd, and Paul S. Lund 1941 *The Social Life of a Modern Community*. New Haven: Yale University Press.
- 1942 *The Status System of a Modern Community*. New Haven: Yale University Press.
- Warren, Roland 1978 *The Community in America*. Chicago: Rand McNally.
- Wellman, Barry, and Barry Leighton 1979 "Networks, Neighborhoods, and Community: Approaches to the Study of the Community Question." *Urban Affairs Quarterly* 14 (3):363-390.
- Whyte, William Foote 1943 *Street Corner Society: The Social Structure of an Italian Slum*. Chicago: University of Chicago Press.
- Wilson, William 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.
- 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Wirth, Louis 1938 "Urbanism as a Way of Life." *American Journal of Sociology* 44:1-24.
- Zimmerman, Carle C. 1938 *The Changing Community*. New York: Harper and Brothers.

GUNNAR ALMGREN

## COMMUNITY HEALTH

See Community; Comparative Health-Care Systems; Health Promotion and Health Status; Medical Sociology.

## COMPARABLE WORTH

"Comparable worth" is one approach to increasing pay equity between jobs done primarily by women and minorities and those done primarily by majority men. It refers to equalizing compensation for jobs requiring comparable levels of effort, skill, and responsibility. The concept of comparable worth defines "equal work" as "work of equivalent value," a broader concept than limiting "equal work" to meaning the "same work." Comparable worth uses job-evaluation methods to establish equivalencies of different jobs in order to identify and correct disparities in pay between jobs held primarily by women and minority men and jobs held primarily by nonminority men. It is worth noting that the term *pay equity* is often used in conjunction with discussions of comparable worth; it usually refers not only to comparable worth but also to other approaches to achieving equity and justice in wages. Pay equity is the best subject heading or "key word" to use when doing further research on the topic of comparable worth.

Proponents of using comparable worth to establish pay equity argue that 1) at least some part of the lower wages in female-dominated jobs and occupations is due to pay discrimination against women, and 2) job evaluation systems can determine the equivalency of different jobs and thus identify the jobs and occupations where this form of discrimination exists.

The issue of comparable worth arose in response to obvious differences in the rewards for jobs held primarily by women and those held primarily by men—even when those jobs required the same or similar levels of education, skill, and responsibility. Advocates of comparable worth and pay equity argue that these wage differences are based on historical and current discrimination in the setting of wages for jobs held primarily by women and minorities. That is, jobs in which workers are primarily women and minorities have been, and remain, systematically undervalued relative to equivalent jobs held by majority men. This undervaluation depresses the wages in jobs done by women and minorities relative to the wages for jobs historically performed by white men. Thus, discrimination is embedded in the current wage structure of jobs (Remick 1984; Marini 1989). Continued job and occupational segregation by sex, combined with the continued systematic undervaluation of jobs held by women and minorities, is thereby a cause of continued inequality and a form of labor market discrimination.

The argument that discrimination is the basis for the inequality of wages between men's and women's jobs is not new. In the 1922 presidential address to the British Association of Economists, F.Y. Edgeworth spoke on "Equal Pay to Men and Women for Equal Work." Edgeworth outlined three major conclusions regarding wage inequality between men and women. First, that men and women work in different jobs, albeit jobs that often require similar levels of effort and skill. Second, that jobs held by women are paid far less than those held by men. Third, that removing overt discrimination would be unlikely to equalize wages for men and women fully (Edgeworth 1922). The findings of recent empirical studies have generally supported these three conclusions.

There is considerable evidence demonstrating that men and women work in different occupations. Comparisons of occupational segregation

by sex in the United States since 1900 show that levels of segregation have been persistent through the 1960s and 1980s (Gross 1968; Jacobs 1989). During the 1980s more than half of the workers of one sex would have had to change occupations in order to equalize the distribution of men and women across all occupations (Jacobs 1989). Researchers using more specific job titles within firms have found that almost no men and women work together in the same job in the same firm (Bielby and Baron 1986).

On average, women continue to earn substantially less than men. Women earned 62 percent of what men earned in 1975 and just over 75 percent of what men earned in 1995 (Figart and Kahn 1997). The concentration of women in low-paying, female-dominated occupations has been found to account for a substantial portion of this income difference. The amount of the income gap between men and women explained by the sex segregation of occupations varies from 25 percent to over 33 percent across many studies of this issue (Sorensen 1986; Figart and Kahn 1997). The remainder of the difference between men's and women's average earnings is due to other factors, such as differences in the overall skills and experience individual men and women bring with them to the labor market.

Empirical studies have examined whether the gap in wages between female-dominated and male-dominated occupations is based on differences in the occupations with regard to the skills required or the work environments. Treiman and colleagues (1984), in an evaluation of the effects of differences in characteristics of male and female occupations on wages, found that "about 40 percent of the earnings gap between male- and female-dominated occupations can be attributed to differences in job characteristics and 60 percent to differences in the rate of return on these characteristics." That is, he found that the premium paid for skills in male-dominated occupations was higher than that paid for the same skills in female-dominated occupations. Other research has confirmed the findings that specific skills (such as dealing with the public) or requirements (such as having a high school diploma) increase wages more in a male-dominated occupation than in a female-dominated occupation (McLaughlin 1978; Beck and Kemp 1986). Other possible explanations for differences

in wages between male-dominated and female-dominated occupations—such as that female-dominated occupations have greater nonmonetary compensations (such as vacation, sick leave, flexibility in working hours) (Jencks, Perman and Rainwater 1988) or that they are more accommodating to intermittent careers (England 1982)—have not been supported by empirical studies.

The conclusion of J.S. Mill in 1865, as quoted by Edgeworth in 1922—“The remuneration of the peculiar employments of women is always, I believe, greatly below that of employments of equal skill and equal disagreeableness carried on by men.”—is similar to the conclusion reached by the National Research Council/National Academy of Sciences committee report in 1981:

*“[Such] differential earnings patterns have existed for many decades. They may arise in part because women and minority men are paid less than white men for doing the same (or very similar) jobs within the same firm, or in part because the job structure is substantially segregated by sex, race, and ethnicity and the jobs held mainly by women and minority men pay less than the jobs held mainly by non-minority men”* (Treiman and Hartmann 1981, p.92).

The evidence is fairly conclusive that occupational-level wage discrimination exists—that is, that skills and requirements are less well rewarded in jobs held primarily by women than they are in jobs held primarily by men. Comparable worth advocates argue that applying job-evaluation methods is a viable method for reducing this form of wage discrimination. This is clarified in Helen Remick’s proposed working definition of comparable worth as “the application of a single, bias-free point factor job evaluation system within a given establishment, across job families, both to rank-order jobs and to set salaries” (1984, p.99).

Job-evaluation methods are well established and have been used for decades to establish equivalencies across jobs. Actual methods of job evaluation differ, but the usual approach is to start by describing all jobs within a given organization. Next, a list of important job requirements is developed and jobs are rated on each requirement. For example, one requirement could be the use of mathematics. In this case each job would be rated from “low” (e.g. addition and subtraction of whole

numbers) to “high” (e.g., the use of differential equations). Most job-evaluation methods include job requirements such as level of education, skills, level of responsibility, and the environment in which the work is performed. Some job-evaluation methods also include such characteristics of job incumbents as average education, training, and experience. More complex job-evaluation systems also consider how jobs rank with regard to fringe benefits (e.g. sick leave), hours (e.g. shift work), training and promotion opportunities, hazards, autonomy (e.g. can leave work without permission), authority (e.g. supervises others), and organizational setting (e.g. organizational size) (see Jencks et al. 1988). After each job is rated and given a certain number of “points” for each requirement, the points are then added into an overall “score” for each job. These scores are then weighted based on the importance assigned to a particular job attribute. Each job then receives a total number of points based on all of the appropriate factors in order to compare the value to the firm of different jobs. These composite scores are then used to rank jobs in order to help determine appropriate wages (Blau and Ferber 1986). This makes it possible to compare wages paid for jobs with very different—but comparable—content. In addition to considering the training and work requirements for jobs within the firm, systems of job evaluation often also take into account whatever information is available on prevailing wages for different types of labor. The use of job evaluation is neither new nor unusual, and currently job evaluation is often used to determine pay scales by governments and by many businesses. Job evaluations are primarily used when employers cannot rely on the market to establish wages. (See Spilerman 1986, for a discussion of the types of organizations that determine wages based on nonmarket mechanisms.) Employers must determine wages, for example, when positions are filled entirely from within an organizational unit (e.g. through promotion of an existing workforce) or when they fill jobs that are unique to a particular firm. In these cases, “going rates” for all jobs are not always available in local labor markets.

There are at least two critical limitations to using job-evaluation methods in establishing comparable worth. First, it is difficult to eliminate the effects of past practices on the identification and the weighting of important job characteristics.

Existing job-evaluation schemes have been criticized for undervaluing, or not even considering, the skills and abilities that are emphasized in some female jobs (Beatty and Beatty 1984; Stienberg 1992). For example, at one time the coding in the job-evaluation system used in the Dictionary of Occupational Titles rated the primarily male occupation "Dog Pound Attendant" as requiring a higher level of complexity with regard to working with data, people, and things than the primarily female occupations of "Nursery School Teacher" and "Practical Nurse"—which were rated as having minimal or no relationship with data, people, or things (Miller et al. 1980). Second, because job-evaluation methods are used within a particular firm or organization, they do not address wage inequalities *across* firms or organizations. This particularly limits the scope of comparable worth because, with the exception of governments (which often employ individuals across a wide range of occupational categories), most organizations are staffed by individuals in a relatively narrow span of occupations. For example, jobs in the textile and poultry-processing industries, usually held by women, and jobs in the lumber industry, usually held by men, could have the same overall scores in terms of job characteristics. However, because these jobs are usually not in the same organization, it is unlikely that job-evaluation methods could be used to equalize wages for these jobs.

Comparable-worth methods have been used in a number of legal actions in attempts to increase the equivalencies of wages for jobs held primarily by women and those held primarily by men. The outcomes of these cases have been mixed (see Remick 1984; Heen 1984; Steinberg 1987; and Figart and Kahn 1997 for reviews and discussions of cases). In the United States, the right of "equal pay for equal work" is provided by Title VII of the Civil Rights Act of 1964. An important legal action based on the premise of comparable worth was the Supreme Court ruling in the *County of Washington v Gunther*, 452 U.S. 161 (1981). This ruling removed a major legal obstacle to comparable worth as the basis of equalizing wages. Although it did not endorse the comparable worth approach, it did rule that a man and woman need not do "equal work" in order to establish pay discrimination under Title VII (Heen 1984). Subsequent lower court rulings have not resulted in clear-cut decisions regarding comparable worth, and at present

it seems unlikely that, in the United States, court decisions will mandate the comparable-worth approach. Legal and legislative actions based on comparable worth and pay equity also have emerged in countries other than the United States. Canada included a provision for equal pay for work of equal value in the Canadian Human Rights Act of 1977 (see Cadieux 1984; and Ontario Pay Equity Commission 1998). The implementation of this legislation has resulted in significant decisions regarding the need to increase wages in female-dominated occupations—both in private industry and in the government. However, the implementation of these decisions has not been without difficulty. For example, the government and the government-workers union of the Northwest Territories were, in early 1998, in disagreement with regard to whether the evaluation system used by the government to establish comparable worth was biased (Government of the Northwest Territories 1998).

Despite its limitations and difficulties in implementation, the concept of comparable worth as a basis for pay equity remains important in public policy initiatives. The National Committee on Pay Equity (1999), the American Federation of State, County, and Municipal Employees (1999), and other organizations provide information and support for advocates of comparable worth. As one approach to increasing pay equity, comparable-worth applications have the potential for identifying and correcting one of the most persistent bases for the disparity between earnings for majority men and earnings for women and minorities.

## REFERENCES

- Beatty, R. W., and J. R. Beatty 1984 "Some Problems with Contemporary Job Evaluation Systems." In H. Remick, ed., *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities*. Philadelphia: Temple University Press.
- Bielby, W. T., and J. N. Baron 1986 "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91:759–99.
- Blau, F. D., and M. A. Ferber 1986 *The Economics of Women, Men and Work*. New York: Prentice-Hall.
- Cadieux, Rita 1984 "Canada's Equal Pay for Work of Equal Value Law." In H. Remick, ed., *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities*. Philadelphia: Temple University Press.



- Edgeworth, F. Y. 1922 "Equal Pay to Men and Women for Equal Work." *The Economic Journal* 32:431-56.
- England, P. 1982 "The Failure of Human Capital Theory to Explain Occupational Sex Segregation." *Journal of Human Resources* 17:358-70.
- Figart, Sarah M., and Peggy Kahn 1997 *Contesting the Market: Pay Equity and the Politics of Economic Restructuring*. Detroit, Mich.: Wayne State University Press.
- Government of the Northwest Territories [http://www.gov.nt.ca/Executive/Pay\\_Equity](http://www.gov.nt.ca/Executive/Pay_Equity), last edited April, 1998.
- Gross, E. 1968 "Plus ça change. . . ? The Sexual Structure of Occupations Over Time." *Social Problems* 16:198-208.
- Heen, M. 1984 "A Review of Federal Court Decisions Under Title VII of the Civil Rights Act of 1964." In H. Remick, ed., *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities*. Philadelphia: Temple University Press.
- Jacobs, J. A. 1989 "Long Term Trends in Occupational Segregation by Sex." *American Journal of Sociology* 95:160-73.
- Jencks, C., L. Perman, and L. Rainwater 1988 "What is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93:132-257.
- Kemp, A. A., and E. M. Beck 1986 "Equal Work, Unequal Pay: Gender Discrimination Within Work-similar Occupations." *Work and Occupations* 13:324-347.
- Marini, M. M. 1989 "Sex Differences in Earnings in the U.S." In W. R. Scott, ed., *Annual Review of Sociology*, vol. 15. Palo Alto, Calif.: Annual Reviews.
- McLaughlin, S. 1978 "Occupational Sex Identification and the Assessment of Male and Female Earnings Inequality." *American Sociological Review* 43:909-921.
- Miller, Ann R., D. J. Treiman, P. S. Cain, and P. A. Roos 1980 *Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles*. Washington, D.C.: National Academy Press.
- National Committee on Pay Equity, <http://www.feminist.com/fairpay.htm>, as of February, 1999.
- Ontario Pay Equity Commission, <http://www.gov.on.ca/LAB/pec/acte.htm>, last modified December 28, 1998.
- Remick, H. (ed.) 1984 *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities*. Philadelphia: Temple University Press.
- 1984 "Major Issues in a priori Applications." In H. Remick, ed., *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities* 99-117. Philadelphia: Temple University Press.
- Sorensen, E. 1986 "Implementing Comparable Worth: A Survey of Recent Job Evaluation Studies." *American Economic Review Papers and Proceedings* 76:364-367.
- Spilerman, S. 1986 "Organizational Rules and the Features of Work Careers." *Research in Social Stratification and Mobility* 5:41-102.
- Steinberg, Ronnie, et al. 1985 *The New York State Comparable Worth Study: Final Report*. New York: Center for Women in Government.
- Treiman, D. J., and H. I. Hartman 1981 *Woman, Work and Wages: Equal Pay for Jobs of Equal Value*. Washington, D.C.: National Academy Press.
- U.S. Bureau of the Census 1987 *Male-Female Differences in Work Experience Occupation and Earnings: 1984*. Current Population Report P-70, no. 10. Washington, D.C.: U.S. Government Printing Office.
- 1986 *Statistical Abstract of the United States: 1987*, 107th ed. Washington, D.C.: U.S. Government Printing Office.

NANCY E. DURBIN  
BARBARA MELBER

## COMPARATIVE HEALTH-CARE SYSTEMS

At the dawn of the twenty-first century, access to health-care services, their cost and quality constitute key social, political, and economic issues for virtually every country in the world. Identifying the conditions under which health-care systems function most effectively has become a vital, albeit elusive, goal. One point is certain: It is impossible to fully understand the dynamics of health-care systems without comparative health-care research. Knowledge of systems other than one's own provides the observer with multiple vantage points from which to gain a fresh perspective on strengths and weaknesses at home. Studying other systems, including their successful as well as failed health-reform efforts, provides a global laboratory for health-systems development. While some countries have been quick to draw upon the health-care innovations of their neighbors, the United States has been relatively slow to look internationally for health-reform ideas. Fortunately, the proliferation of comparative health-care studies promises that such insularity will be much less likely in the future.

The comparative study of health-care systems focuses on two broad types of issues. The first

involves *describing* the range of health-care services in populations or societies, particularly their organization and functioning. By far the most common type of research, descriptive studies, brings together statistical indicators and factual explanations about how various national systems operate (van Atteveld et al. 1987). Some of this work incorporates an analytical dimension by categorizing systems in terms of conceptual schemes or typologies. Less attention has been paid to the second type of research, which looks more closely at the dynamics of how health-care systems behave. The intent in this type of research is to *analyze* the patterns among system characteristics, especially with the idea of anticipating the outcomes that are likely with specific types of system arrangements. Still in its infancy, this work promises an in-depth yet practical understanding of how health-care can be organized and financed to achieve desirable levels of both quality and access.

Interest in cross-national studies of health-care systems increased dramatically in the early 1990s as a result of national debates over reorganizing American health-care. Rapidly aging populations in many advanced, capitalistic countries, in combination with the expanding scope of high technology medicine, resulted in increased public demand for health-care. At the same time, poverty and other forms of social inequality as well as ineffective societal institutions created major public health problems in many developing countries, such as hazardous water, inadequate or harmful food supplies, poor air quality, unsafe homes and workplaces, and the swift spread of infectious diseases. In both cases, health-care systems have been severely challenged and often cannot meet the needs of citizens. Because of these problems and also due to enhanced global cooperation, social scientists and policy makers are increasingly turning their attention to the experience of other countries.

#### KEY CHARACTERISTICS FOR THE COMPARISON OF HEALTH-CARE SYSTEMS

Drawing on the work of Anderson (1989) and Frenk (1994), a health-care system can be defined as the combination of health-care institutions, supporting human resources, financing mechanisms, information systems, organizational structures that link institutions and resources, and management

structures that collectively culminate in the delivery of health services to patients. Within this broad framework, the methodology for comparing health-care systems can vary widely. A standard approach would include some or all of the dimensions outlined below.

The most fundamental comparative dimension is the *organization, financing, and control* of a health-care system. This involves comparing which health services are provided; how they are paid for; how they are configured, planned, and regulated; and how citizens gain access to them. Among countries with advanced economies, health-care services today look much the same to the casual observer; however, the financing arrangements and policy-making mechanisms that underlie them vary widely. The role of government is perhaps the most significant organizational variable in international health-care. All governments, with the notable exception of the U.S. government, accept responsibility for the health-care of citizens (Evans 1997). Some governments take it upon themselves to actually provide health services and, therefore, own their own clinics and hospitals and hire their own physicians and staff—examples are Sweden and Denmark. Within the U.S., the Veteran's Administration operates its own health-care system this way. In a variation of this model, the government acts as the purchaser (but not the owner) of health-care services, obtaining services from private providers on behalf of patients, such as in Canada or the reformed health-care system in Britain. In Finland, local governments can purchase from either public or private providers. In another model, illustrated by Germany and Japan, the government avoids acting as the major payer, and instead takes the role of an overseer, setting mandates for health coverage, including the type and level of coverage, and regulating the terms of what is largely a private system. Due to economic pressures, national governments in both Germany and Japan provided increasing subsidies to support their systems in the 1990s.

It is important also to compare health-care systems in terms of *physician characteristics and provider arrangements for primary care and prevention*. The supply of medical personnel (e.g., the number of physicians per unit of population) is a key comparative indicator. Interestingly, there is significant variation in the number of physicians in advanced economy nations, ranging per 10,000

population from fifty-five in Italy to thirty-four in Germany to twenty-six in the United States, which is more typical, to seventeen in the United Kingdom (Anderson and Poullier 1999). Equally interesting is that there is no apparent corresponding variation in the health status of these populations. A more complex issue is how different systems organize and divide medical work between various professions and occupations. In some countries, such as Sweden, Finland, and the Netherlands, midwives or nurse midwives have primary responsibility for normal prenatal care and childbirth; in others, such as the United States, physicians have responsibility for these tasks and midwives are relatively rare. Practice arrangements between generalist and specialist physicians are another point of comparison. The United States is unique among its peers because primary care physicians working in ambulatory settings also have hospital privileges, and, therefore, have the right to admit patients and to treat hospitalized patients. In Britain, Sweden, Germany, and many other countries, on the other hand, only specialists comprise the hospital medical staff and only they can treat patients there.

*Hospitals and long-term care arrangements* constitute another dimension in comparing health-care systems. Countries vary widely in how they use hospitals, as well as in how and where citizens with chronic illnesses and other debilitating conditions receive ongoing, nonacute care. Many Western countries today are in the process of shifting from institutional to community-based care. There are many factors affecting this transition, including whether alternatives to institutional care, such as home health and support programs, are readily available. In countries such as Japan and Germany, that through the 1990s have relied on informal family caregiving arrangements rather than institutions or community-based services, sociodemographic changes (elders living longer as well as changes in the work force participation of caregivers) are producing a long-term care crisis.

Health-systems also can be compared according to the degree to which care is *integrated and coordinated* between various sectors and levels of services. Systems that have rigid boundaries between ambulatory and hospital services, or between general medical and mental health services and that are so highly specialized that patients are required to transfer among multiple providers, illustrate arrangements where communication

across "borders" is vital for optimal patient care. The Nordic countries, in particular, have been highly conscious of gaps in the coordination and continuity of care, and have developed reforms to bridge them. Finland has adopted a model that integrates the basic education of both health-care and social care personnel, a strategy intended to link the biomedical and social aspects of health and health-care at the very beginning of professional education.

Perhaps the most common basis for comparing countries' health-care systems is various outcome statistics such as *economic characteristics, personnel resources, utilization rates, and population health status measures*. The proportion of the population covered by government-assured health insurance stands as the indicator with the least international variation. Of the twenty-nine countries analyzed by the Organization for Economic Cooperation and Development in 1998, governments in twenty-four countries assured coverage to 99 percent to 100 percent of their population. The five exceptions were Germany with 92 percent coverage, Mexico and the Netherlands with 72 percent, Turkey with 66 percent, and the U.S. with 33 percent (Anderson and Poullier 1999). Other indicators in the same report vary widely. In per capita health-care spending, the United States spends just less than \$4,000, an amount that is substantially more than for the other nations, which range from Switzerland's \$2547 to Turkey's \$260. Contrary to what one might think, the amount a country spends on health-care is not reflected in the longevity of its population. Men in twenty-one countries live longer than do men in the United States (up to an average of 77 years in Japan compared to 72.7 in the United States). Similarly, women in nineteen countries live longer than American women (up to an average of 83.6 years in Japan versus 79.4 in the United States). In fact, relatively few of the high-spending countries appear among those with the greatest longevity.

Detailed comparisons of statistical characteristics have limited value when neither issues of measurement nor the contextual meanings associated with each measure are immediately clear. Each country has its own distinctive set of *sociohistorical and cultural characteristics*; these should be considered in comparative research, since they often play an important role in explaining the origin and

development of health-care systems (Payer 1996). For example, Starr (1982) has posed and explored the question of why the United States ignored national health insurance at the same time most European countries were adopting such programs. National insurance, he argues, was a form of social protectionism and was, therefore, most likely to be enacted by paternalistic regimes such as Germany and only later by more liberal states such as France and the United Kingdom. In addition, the United States' long history of rejecting national health insurance reflects its decentralized government, the relative lack of domestic unrest, and the failure of major interest groups (labor, business, and medicine) to provide support. Regarding the same issue, Steinmo and Watts (1995) contend that the structure of American political institutions creates conditions that work against the adoption of national health insurance.

A final means of comparing health-care systems is on the basis of *specific problems*, including assessments of citizen satisfaction (Donelan et al. 1999). Closely related comparisons focus on the various *strategies for reform* that countries have implemented in an attempt to address their problems (Graig 1993). Much can be learned by studying and comparing which solutions seem to work best for certain types of problems, as well as observing patterns of failure across multiple systems.

#### FRAMEWORKS FOR COMPARING HEALTH-CARE SYSTEMS

While it is useful to compare health-care systems on the basis of a series of characteristics, in-depth understanding of the variation in systems requires a more analytical approach. Most frameworks proposed by comparative health-system researchers can be characterized as typologies that, though they add a conceptual basis for distinguishing among systems, offer limited analytical complexity or depth. These typologies typically emphasize political or economic criteria, and several of the more complex schemes attempt to integrate the two. Graig (1993) organized the six countries she analyzed on a continuum with public systems (e.g., the United Kingdom) at one end, private systems (the United States) at the other, and mixed or "convergence" systems, such as Japan, Germany, and the Netherlands, in the middle. Roemer

(1991) developed a sixteen-cell typology that combines political and economic elements, describing systems in relation to government policies (entrepreneurial/permissive, welfare oriented, universal/comprehensive, and socialist/centrally planned) as they intersect with economic conditions (affluent, developing, poor and resource rich). These and similar approaches offered by Anderson (1989) and Light (1990) offer descriptive distinctions, but do not emphasize the theoretical basis for the particular category schemes.

Mechanic (1996) hypothesizes that health-systems internationally are converging, based on the idea that medicine forms a world culture in which knowledge and health-care ideas are quickly disseminated. He proposes six areas of convergence: nations' concerns with controlling cost and improving the efficiency and effectiveness of health-care, nations' realization that population health outcomes are largely a product of circumstances outside the medical care system, nations' concern and attempts to address inequalities in health outcomes and access to care, nations' growing interest in patient satisfaction and consumer choice, and the increasing attention nations are placing on the linkage between medical and social factors in health-care, and the struggle all nations are having between technology, specialization, and the need to develop primary care. Field (1980) also theorizes a progression of health-system development with a scheme placing systems along a continuum according to the extent to which health-care is seen as a social good. He identifies five system types in order of their progression toward a socialized system: anomic, pluralistic, insurance/social security, national health service, and socialized health-care.

Elling (1994) takes issue with the notion of "automatic convergence" and with typologies that posit unidirectional evolution or change, pointing out that countries are involved in ongoing, dynamic class struggles in the development of health-systems. His neo-Marxist framework, influenced by Wallerstein's world-systems theory, allows for countries to move back and forth among five types: core capitalist (e.g., the United States and Germany), core capitalist/social welfare (e.g., Canada, Japan, Sweden, United Kingdom), industrialized socialist-oriented (e.g., pre-1990s Soviet and

eastern European systems), capitalist dependencies in the periphery and semi-periphery (e.g., Brazil and India), and socialist-oriented, quasi-independent of the world system (e.g., China and Cuba).

In 1990, Esping-Andersen (1990) contributed an important theoretical framework to the field of comparative welfare states research, that can also be useful in distinguishing among health-care systems in advanced capitalist nations. Esping-Andersen conceptualizes all welfare activities, including health-care, as a product of a state-market-family nexus. His typology organizes “welfare regimes” around this nexus, specifically as it results in the decommodification of workers in a country; that is, the degree to which a citizen can obtain basic health and social welfare services outside of the market. In terms of health-care, this would mean a citizen’s ability to access health-care services without having to purchase them “out-of-pocket.” Esping-Andersen identifies three types of systems: conservative or corporatist, liberal, and socialist or social democratic. In *The Three Worlds of Welfare Capitalism*, he constructs measures of decommodification that he applies to data from eighteen nations in order to assess where each nation’s welfare regime ranks according to the three types. Under Esping-Andersen’s scheme, social democratic systems include the Nordic countries and the Netherlands; liberal systems include Canada, Japan, and the United States; and the conservative or corporatist welfare states include France, Germany, and Italy. These placements appear similar to what Graig proposed using a much simpler public-private dimension; however, Esping-Andersen’s methodology is unique in its theoretical approach and because it can be operationalized quantitatively to allow precise distinctions among a large number of nations.

Reviewing various health-care system frameworks shows a wide variety of approaches and scholarly emphases. Comparative research is in the process of moving beyond simple typologies to focus on the conditions under which different types of systems emerge and under which they change. This work demonstrates the value of a broad approach, integrating the key aspects that affect the development of health-care systems, including historical, cultural, political, and economic factors.

## REVIEW OF SELECTED HEALTH-CARE SYSTEMS

Comparative international research on health-care systems requires information that is both detailed and current. Obtaining data is complicated by the fact that health-systems throughout the world operate in a state of constant flux. These summaries present the most recent information available for the health-care systems of selected countries. Among countries with advanced economies, Sweden and the United States often occupy opposite extremes when it comes to health-care organization and financing. For this reason, Sweden is the first country discussed along with three countries that have similar systems—Finland, the United Kingdom, and Canada—followed next by Germany, Japan, Russia, and China, with additional discussions of France, Mexico, Argentina, Chile, Colombia, and Ghana.

Swedish health-care reflects three basic principles: *equality* among citizens in access to health-care; *universality* in the nature of services (the idea that everyone should receive the same quality of services); and *solidarity*, the concept of one social group sacrificing for another group in the interest of the whole society (Zimmerman and Halpert 1997). Solidarity in this context refers to taxing those who use fewer services at the same rate as those who use more—for example, similar health-care taxation for younger persons or affluent persons versus the elderly or the poor). The Swedish health-care system is predominantly a publicly owned and funded system; thus, approximately 85 percent of Swedish health-care is publicly funded, whereas in the United States the public portion is just under 50 percent (Lassey, Lassey, and Jinks 1997). The differences between the two systems are more startling in terms of the growth of overall spending. In the early 1980s, both countries were spending approximately 9.5 percent of their Gross Domestic Product on health-care. By the end of the century, the situation had changed dramatically; in 1999, Sweden was spending 8.6 percent (\$1728 per capita) compared to 13.5 percent (\$3924 per capita) in the United States (Anderson and Poullier 1999). Populations in the two nations also differ markedly on several basic health status indicators. Infant mortality Sweden is four deaths per thousand live births compared to 7.8 in the United States. Swedish men live nearly four years longer

on average than American men, and Swedish women live two years longer than their American counterparts. These favorable statistical indicators compel a closer look into how the Swedish health-care system is organized.

The Swedish welfare state, including health-care and social services, is one of the most comprehensive and universal in the world. Health-care in Sweden is the responsibility of the state, which delegates it, in turn, to each of Sweden's twenty-one county councils (Swedish Institute 1999). Elected officials in each county are charged with providing comprehensive health services for residents, and with levying the taxes to finance them. The system is decentralized; each county by law must provide the same generous common core of services to all residents, although just how they decide to do it can vary. In the 1990s, Sweden embarked on a series of reforms in order to increase health-care quality and efficiency. As a result, Swedish citizens now have greater freedom in choosing their own primary care physicians. The vast majority of these doctors are employed by the county councils to practice in small group clinics and health centers distributed geographically throughout the country. Specialist physicians practice in hospitals where they also see outpatients on both a referral and self-referral basis. The medical division of labor also includes district nurses, physical therapists, and midwives, all of whom are used extensively to deliver care through local health centers (maternity clinics in the case of midwives). Midwives also work in hospitals where they have responsibility for normal cases of labor and delivery. Sweden's elderly constitute an increasing proportion of the population, creating significant challenges for both current and future health and social services. Sweden's social policy emphasizes that citizens should be able to live in their own homes for as long as possible, meaning that nursing home placement occurs only when absolutely necessary. Services for the elderly may involve as many as five or six home nursing visits per day in order for the disabled and elderly to remain at home in the community.

Swedish citizens are taxed heavily to maintain the quality and level of services they expect; at the same time, they have shown high levels of political support for maintaining their expensive system. In the 1980s, many services were entirely free; however, today there typically is a modest copayment.

The copayment for a primary care physician visit, for example, currently ranges from \$12 to \$17 depending on the county council. For specialist physician visits the copayment ranges from \$15 to \$31, and for hospital stays it is fixed at \$10/day. The Swedish system includes a high-cost ceiling so that, after a person spends approximately \$113 out-of-pocket each year, health-care services are free. Medications must be purchased by the individual until they have reached a threshold of a little more than \$100. Prescriptions are then discounted until the patient has spent \$225, at which point medications become free (Swedish Institute 1999). These amounts have increased somewhat during the 1990s, but due to Sweden's already high taxation rate, county councils were hesitant to ask patients to pay more. Reforms during the same period included establishing internal performance incentives or "public competition" (Saltman and Von Otter 1992), a structural arrangement that arguably enabled Sweden to maintain the basic features of its health-care system without large tax or out-of-pocket increases. Some have expressed concern that the system is stretched to its limits. Regardless of which view is correct, the Swedish system requires a healthy economy in order to continue, given continuing cost pressures and an increasingly aged population.

The Swedish model of a publicly owned and financed health-care system shares common features with systems in several other countries, including the United Kingdom, Finland, and Canada. The British National Health Service (NHS), like its Swedish counterpart, provides publicly funded, comprehensive health-care to the population, and enjoys a solid base of citizen support, albeit under ongoing criticism. What distinguishes the NHS from health-care systems in other Western countries is its frugality. Characterized by long waiting lists and what Klein (1998) refers to as "rationing by professionally defined need," Britain runs the cheapest health-care system in Europe, outside of Spain and Portugal. In the early 1990s, the NHS went through a series of dramatic changes, creating "internal markets," a system of inside competition intended to increase productivity and further decentralize its historically large and unwieldy bureaucracy. These reforms—several of which were adopted by Sweden—reorganized primary care practices and shifted physician payment from fixed salary to capitation. NHS

hospitals also entered into new arrangements where they have greater structural independence and compete among themselves. In 1997, the British government introduced new proposals that changed course yet again. The direction of the NHS in the twenty-first century is unclear (Klein 1998).

Finland's health-care system also is publicly owned, financed through general taxation, and decentralized. In fact, Finland operates with more health-care decentralization than Sweden. Since 1993, Finnish funding for health-care has been incorporated into block grants from the national government that are given annually to each of the country's 455 municipalities, some of which form partnerships for purposes of delivering health-care (Hermanson, Aro, and Bennett 1994). Within the parameters of national guidelines, elected officials in each of these jurisdictions (similar to the county councils in Sweden) have the responsibility to obtain and deliver health-care services to the population. High standards, a comprehensive array of services, as well as modest out-of-pocket payments, make the Finnish system comparable to the Swedish system. Canada also provides health-care to its entire population. The Canadian system is administered by the provinces and is financed largely by public taxation, roughly three-quarters of which is from the provincial government. While basic services remain constant, some specific provisions of health-care in Canada vary considerably from province to province. Unlike in Sweden, Finland, and the United Kingdom, Canadian physicians are paid on a fee-for-service basis. Furthermore, Canadian primary care physicians act as gatekeepers to specialists and hospitals (much as they do in Finland and the United Kingdom, but not in Sweden). As is the case with all the countries in this group of four, Canadians have extreme pride in their health-care system. They appear resolute in maintaining it, although a number of controversial cost-cutting measures and as yet unsolved problems raise considerable concern about the future (Lassey, Lassey, and Jinks 1997).

Germany and Japan, as well as France, have achieved comprehensive and universal coverage (92 percent in the case of Germany) with a model that is closer to that of the United States than the social democratic model of Sweden discussed above (U.S. General Accounting Office 1991). In all three of these countries, medical care is provided by

private physicians, by both private and public hospitals, and patients can choose their physicians. Benefits are comprehensive and mandated by the national government, which also regulates enrollment, premiums, and reimbursement of providers. In contrast to the four countries discussed earlier, financing in Germany, Japan, and France is predominantly private with multiple payers. Workplace-based insurance (financed typically through payroll deductions) covers most employees and their dependents while other payers cover the remainder of the population. Patients make copayments for physician visits and hospital stays, ranging from a nominal amount in Germany to as much as 20 percent or 30 percent of the fee in France and Japan. There is national regulation to ensure consistency. Coverage and care conditions vary from fund to fund, resulting in greater inequalities of benefits compared to the public systems of Sweden, Finland, Canada, and the United Kingdom. Arguing that such systems actually help maintain social divisions and inequality, Esping-Andersen (1990) refers to them as conservative or corporatist.

Although Germany privately finances much of its health-care system, the national government plays a strong role. All but the most wealthy of German citizens are required by law to join one of Germany's 750 insuring organizations, called "sickness funds." In practice, all but about 10 percent of the population opt to join the system, encouraged to participate by the great difficulty of getting back in later on. Sickness funds are private, non-profit organizations that collect premiums or "contributions" for each member—half paid by the individual and half by the employer—and, in turn, contract for health services with physician organizations and hospitals. One of the biggest problems for the German system is continuing cost pressures, exacerbated by the reunification of East and West Germany in 1991. The system uses fixed budgets for hospitals and strict fee schedules for physicians, with punitive measures for inordinate increases in volume, to combat the problem. These were tightened in a major 1993 reform that imposed strict three-year budgets on all major sectors of the system as well as longer-term structural reforms. Early results have been positive (U.S. General Accounting Office 1994).

Japanese health-care follows much the same private, multi-payer model as in Germany, with health-care provided to all citizens through 5,000

independent insurance plans. The plans fall into three major groups, each enrolling about a third of the population: large-firm employees, small-firm employees, and self-employed persons and pensioners. In the case of the first two plans, as is the case with the German sickness funds, the employer pays approximately half of the premium and the employee pays the remaining portion. The similarity to the German system is no accident; Japan has consciously patterned its health-care system after Germany's, dating back to its modernization in the late nineteenth and early twentieth centuries (Lassey, Lassey, and Jinks 1997). Despite these similarities, Japan's health-care system presents some unique and somewhat startling features compared to the other systems discussed here.

Compared to other systems with a significant private component, the Japanese system costs considerably less (Andersen and Poullier 1999). In 1997, for example, per capita health-care spending in Japan was \$1741, less than in the United States, Canada, France, Germany, and many other European countries. Of the nations discussed here, only the austere British system (\$1347 per capita) and efficiency-conscious Finnish system (\$1492 per capita) spent less. Japan's economical approach to delivering health-care raises two paradoxes. First, the low levels of spending would seem to contradict the fact that Japan currently has the longest life expectancy and the lowest infant mortality in the world. In addition, utilization rates in Japan are high, which some would argue indicates a sicker rather than a healthier population. Specifically, the Japanese visit physicians two to three times more frequently, stay in the hospital three to four times longer, and devote considerably more health spending to pharmaceuticals than the other nations discussed here. How can these contradictions be explained? Ikegami and Campbell (1999) argue that, in part, the paradox can be explained by the country's much lower incidence of social problems related to health, such as crime, drug use, high-speed motor vehicle accidents, teen-age births, and HIV infections. Less aggressive medicine and lower hospital staffing and amenities also are thought to keep down costs in Japan.

Health-care arrangements in nations where the political economy is neither "advanced" nor a stronghold of capitalistic democracy can also be instructive. During the latter part of the twentieth

century and until today, both China and Russia have been faced with the monumental challenge of providing health-care with very limited financial resources to huge, diverse populations, many of which live under poor social conditions. Their health-care systems, however, are quite different and their current problems reflect the distinctive political and economic trajectories of the two countries. Russian health-care at the dawn of the twenty-first century is a system in crisis. Based on socialist principles of universal and free access, the old Soviet system included a primary care network of local clinics ("polyclinics") typically connected to a general hospital, as well as more specialized hospitals. This means that a regional city might have separate hospitals for emergencies, maternity, children, and various infectious diseases (Albrecht and Salmon 1992). Funding came directly from the central government until 1993 when a new health insurance law was approved, shifting the source of financing to employer payroll deductions. There are major questions, however, as to whether such a system can be effective during the current period of resource scarcity and instability in major social institutions (Lassey, Lassey, and Jinks 1997).

Chinese health-care, for most Westerners, evokes images of acupuncture and other forms of Eastern medicine, as well as the idealized "barefoot doctors" of the 1960s and 1970s, practitioners with basic medical training who provided primary care in rural areas. In reality, traditional Chinese medicine exists alongside an increasingly dominant Western medical establishment, and the barefoot doctors have all but disappeared. Since new leadership took over the Communist Party in the late 1970s, China has encouraged privatization and decentralization in health-care. By the 1990s, nearly half of all village health-care was provided by private practitioners (Lassey, Lassey, and Jinks 1997). These changes reportedly have been accompanied by a decline in preventive care and public health efforts in rural areas. At the same time, the situation in urban areas seems to have improved. China's revolutionary-era network of local and regional clinics and hospitals has been modernized, although resource shortages continue to limit the level of technological advancement. The most significant change in China is the growing impact of privatization, which appears to be bringing China many of the same problems that



have plagued privatized systems elsewhere: lack of insurance coverage, increasing costs, maldistribution of providers, and inequalities in the overall quality of care (Liu, Liu, and Meng 1994). As in the United States, the gap between the health-care received by the rich and that received by the poor is growing (Shi 1993).

Huge disparities between the rich and poor are characteristic of Latin America where they constitute a significant barrier to universal health coverage. Latin-American health-systems vary considerably, reflecting socioeconomic differences between countries as well as historical and political contingencies. The Mexican system illustrates many of the obstacles faced by developing nations, whether in Latin America or elsewhere. The Mexican constitution established federal responsibility for health-care in 1917, along with a centralized administrative tradition that still exists; yet, to date, the two major government insurance schemes cover only about 47 percent of the population, with another 7 percent insured privately. Ostensibly, there are programs for the remaining 46 percent of the population, most of whom are low income, but in reality, many low income areas and impoverished communities are poorly served. Some have argued that a basic health-care infrastructure is in place and that there are sufficient numbers of well-trained health-care professionals available (Lassey, Lassey, and Jinks 1997). They contend that, had it not been for several national crises in the 1980s and 1990s, coupled with a lack of political will, more of the Mexican population would now be covered by the health-care system.

Bertranou (1999) has compared Argentina, Chile, and Columbia, all of which have employed various ways to reform health insurance arrangements in recent years. Chile's reforms date back to its military dictatorship in the early 1980s. At that time, private health insurers were allowed to compete for worker payroll contributions. There was little regulation of the system which encouraged adverse selection, resulting in significant inequities within the system. Even so, approximately 70 percent of the Chilean population today is covered by insurance, compared to 64 percent in Argentina and only 43 percent in Colombia. Argentina faces numerous obstacles in reforming its complex and confusing system of three types of health-care arrangements (social insurance organizations,

private health insurers and providers, and the public health-system). Its goals include universal coverage and a standard benefits package. In the case of Columbia, reform goals are more related to the relative poverty in the country and the fact that large segments of the population cannot afford health insurance. Per capita expenditures for health-care in Columbia are 42 percent of what they are in Chile and 20 percent of the expenditure level in Argentina. Instead of giving free access to public facilities, Columbia's reforms involve providing vouchers that allow low-income families to join the health organization of their choice. To be successful, all these reform efforts require social equilibrium where governments are able to maintain political and economic stability.

The political instability and socioeconomic inequality that have characterized Latin America are also a hindrance to health-care systems in Africa. An even more fundamental problem in Africa, however, is the formidable lack of resources to address overwhelming health-care needs (Schieber and Maeda 1999). Even where clinics, hospitals, and medical personnel exist, there is likely to be a lack of the required equipment and medicines. In Africa as a whole, 80 percent of the physicians live and practice in the cities where less than 20 percent of the population lives. According to Sanneh (1999), this continuing situation supports the prominent role of traditional healers, 85 percent of whom live in rural areas. The difficulties encountered by Ghana in implementing a system of primary care illustrate the situation affecting many African nations. In 1983, soon after the primary care system was adopted, the government attempted cost containment by introducing "user fees" in all public health facilities, clinics as well as hospitals. Two years later the fees were increased and, subsequently, a health insurance program was instituted. One must question the practical significance of these developments in a country where over half of rural residents and nearly half of those in urban areas live below poverty (Anyinam 1989). These circumstances are a reminder that developing countries contain 84 percent of the world's population, yet account for only 11 percent of global health-care spending (Schieber and Maeda 1999), making the task of designing strategies for effective health-care delivery in the developing world the true challenge for comparative health-care system researchers.

Comparing health-care systems entails a vast array of information, including historical background, cultural patterns and beliefs, geographic considerations, as well as social, economic, and political factors. It involves detailed descriptions of policies and procedures, complex statistical profiles, as well as an understanding of conceptual frameworks, theory, and comparative methods. The potential rewards of comparative work, however, balance the challenges. Whether there is convergence in the structure and functioning of health-care systems or not, many of the problems faced by nations in delivering health-care to citizens are similar. There are lessons to be learned from comparing health-care systems internationally that can only aid in addressing these problems.

## REFERENCES

- Albrecht, Gary L., and J. Warren Salmon 1992 "Soviet Health-care in the Glastnost Era." In Marilyn Rosenthal and Marcel Frenkel, eds., *Health-care Systems and their Patients: An International Perspective*. Boulder, Co: Westview Press.
- Anderson, Gerard F., and Jean-Pierre Poullier 1999 "Health Spending, Access, and Outcomes: Trends in Industrialized Countries." *Health Affairs* 18:178-192.
- Anderson, Odin W. 1989 *The Health Service Continuum in Democratic States: An Inquiry into Solvable Problems*. Ann Arbor, Mich.: Health Administration Press.
- Anyinam, C. A. 1989 "The Social Costs of the International Monetary Fund's Adjustment Programs for Poverty: The Case of Health-Care Development in Ghana." *International Journal of Health Services* 19:531-547.
- Bertranou, Fabio M. 1999 "Are Market-Oriented Health Insurance Reforms Possible in Latin America? The Cases of Argentina, Chile and Colombia." *Health Policy* 47:19-36.
- Donelan, Karen, Robert J. Blendon, Cathy Schoen, Karen Davis, and Katherine Binns 1999 "The Cost of Health-System Change: Public Discontent in Five Nations." *Health Affairs* 18:206-216.
- Elling, Ray H. 1994 "Theory and Method for the Cross-National Study of Health Systems" *International Journal of Health Services* 24:285-309.
- Esping-Andersen, Gosta 1990 *The Three Worlds of Welfare Capitalism*. Princeton, N.J.: Princeton University Press.
- Evans, Robert G. 1997 "Going for the Gold: The Redistributive Agenda behind Market-Based HealthCare Reform." *Journal of Health Politics, Policy and Law* 22:427-465.
- Field, Mark G. "The Health-System and the Polity: A Contemporary American Dialectic." *Social Science and Medicine* 14:397-413.
- Frenk, Julio 1994 "Dimensions of Health-System Reform." *Health Policy* 27:19-34.
- Graig, Laurene A. 1993 *Health of Nations: An International Perspective on U.S. Health Care Reform*. 2nd Ed. Washington, D.C.: Congressional Quarterly.
- Hermanson, Terhi, Seppo Aro, and Charles L. Bennett 1994 "Finland's Health-Care System: Universal Access to Health-Care in a Capitalistic Democracy." *Journal of the American Medical Association* 271:1957-1962.
- Ikegami, Naoki, and John Creighton Campbell 1999 "Health-Care Reform in Japan: The Virtues of Muddling Through." *Health Affairs* 18:56-75.
- Klein, Rudolf 1998 "Why Britain is Reorganizing Its National Health Service—Yet Again." *Health Affairs* 17:111-125.
- Lassey, Marie L., William R. Lassey, and Martin J. Jinks 1997 *Health-Care Systems Around the World: Characteristics, Issues, Reforms*. Upper Saddle River, N.J.: Prentice Hall.
- Light, Donald W. 1990 "Comparing Health-Care Systems: Lessons from East and West Germany." In P. Conrad and R. Kern eds., *The Sociology of Health & Illness: Critical Perspectives*, 3rd ed. New York: St. Martin's Press.
- Liu, Gordon, Xingszhu Liu, and Qingyue Meng 1994 "Privatization of the Medical Market in Socialist China: A Historical Approach." *Health Policy* 27:157-174.
- Mechanic, David, and David A. Rochefort 1996 "Comparative Medical Systems." *Annual Review of Sociology* 22:239-270.
- Payer, Lynn 1996 *Medicine and Culture*, 2nd ed. New York: Owl Books/Henry Holt.
- Roemer, Milton I. 1991 *National Health-Systems of the World*, vols. 1 and 2 New York: Oxford University Press.
- Saltman, Richard B., and Casten Von Otter 1992 *Planned Markets and Public Competition: Strategic Reform in Northern European Health Systems*. Philadelphia: Open University Press.
- Sanneh, Kanja B. A. S. 1999 "The Role of Traditional Healers in Primary Health-Care: Roadblocks to Health Education in Africa." Unpublished paper. Department of Health Policy and Management, University of Kansas.

- Schieber, George, and Akiko Maeda 1999 "Health-Care Financing and Delivery in Developing Countries." *Health Affairs* 18:193-205.
- Shi, Leiyu 1993 "Health-Care in China: A Rural-Urban Comparison After the Reforms." *Bulletin of the World Health Organization* 71:723-736.
- Starr, Paul 1982 *The Social Transformation of American Medicine*. New York: Basic Books.
- Steinmo, Sven, and Jon Watts 1995 "It's the Institutions, Stupid! Why Comprehensive National Health Insurance Always Fails in America." *Journal of Health Politics, Policy and Law* 20:329-372.
- Swedish Institute 1999 "Fact Sheet on Health Care." Stockholm, Sweden: Swedish Institute
- U.S. General Accounting Office 1991 "Health-Care Spending Control: The Experience of France, Germany, and Japan." GAO/HRD-92-9 Washington, D.C.: U.S. General Accounting Office.
- U.S. General Accounting Office 1994 "German Health-reforms: Changes Result in Lower Health Costs in 1993." GAO/HEHS-95-27 Washington, D. C.: U.S. General Accounting Office.
- Van Atteveld, L., C. Broeders, and R. Lapre 1987 "International Comparative Research in Health-Care: A Study of the Literature." *Health Policy* 8:105-136.
- Zimmerman, Mary K., and Burton Halpert 1997 "Health and Welfare Reforms in Sweden: 1990-1996: How Safe is the Safety Net?" Paper presented at the annual meeting of the Midwest Sociological Society Des Moines, Ia.

MARY KO ZIMMERMAN

## COMPARATIVE-HISTORICAL SOCIOLOGY

Explicit analytic attention to time and space as the context, cause, or outcome of fundamental social processes distinguishes comparative-historical analysis from other forms of social research. Historical processes occurring in or across geographic, political, or economic units (e.g., regions, nation-states, multi-state alliances, or entire world systems) are systematically compared for the purposes of more generally understanding patterns of social stability and social change (Abrams 1982; Skocpol 1984a; Tilly 1984; Mahoney 1999). Three very different and influential studies illustrate both

the kinds of questions comparative-historical sociologists address and the approaches they use.

First is the classic study by Reinhard Bendix ([1956] 1974) on work and authority in industry. Bendix initially observed that all industrial societies must authoritatively coordinate productive activities. Yet by systematically comparing how this was done in four countries during particular historical periods—pre-1917 Russia, post-World War II East Germany, and England and United States during epochs of intense industrialization—he showed that national variation in ideologies of workplace dominance were related to differences in the social structures of the countries studied.

Second is the analysis of the historical origins and development of the modern world system by Immanuel Wallerstein (1974). Wallerstein took as his unit of analysis the entire sixteenth-century capitalist world economy. Through comparing instances of the geographic division of labor, and especially the increasing bifurcation of global economic activity into "core" and "peripheral" areas, Wallerstein suggested that the economic interdependence of nation-states likely conditions their developmental trajectories.

Third is the analysis of social revolutions by Theda Skocpol (1979). Skocpol compared the histories of revolution in three ancient regime states: pre-1789 France, czarist Russia, and imperial China, and found that revolutionary situations emerged in these states because international crises exacerbated problems induced by their agrarian class structures and political institutions. She then buttressed her causal generalizations by comparing similar agrarian societies—Meiji Japan, Germany in 1806 and 1848, and seventeenth century England—that witnessed failed revolutions.

Each of these influential studies combined theoretical concepts and nonexperimental research methods to compare and contrast historical processes occurring within and across a number of geographic cases or instances. By using such research methods, Bendix, Wallerstein, Skocpol, and many others are following in the footsteps of the founders of sociology. In their attempts to understand and explain the sweeping transformations of nineteenth-century Europe, Alexis de Tocqueville, Karl Marx, Emile Durkheim, and Max Weber all employed and contributed to the formulation of this broad analytic frame (Smelser 1976; Abrams

1982). Comparative-historical analysis in sociology—and, as we shall see, debate over how it should be conducted—thus is as old as sociology itself.

### PURPOSES, PROMISE, ACHIEVEMENTS

The analytic power of comparative-historical strategies stems from the uniquely paradoxical quality of the perspectives, data, and procedures in comparative-historical research. On the one hand, historical comparisons have the potential to harness and exploit the huge variation in social processes and institutions. Some scholars believe this essential to the development of truly general theory and to transcultural/transhistorical explanation (Przeworski and Tuene 1970; Kiser and Hechter 1991). On the other hand, historical comparisons have the potential to exploit the “time-space boundedness” of social life and its historical antecedents and specificity. Others view this as equally essential to theoretical development and to concrete, “real world” explanation (Moore 1966; Skocpol 1984b; Tilly 1984; Stryker 1996). Most comparative-historical sociologists capitalize in some fashion on this paradox, finding diversity in the midst of uniformity and producing regularities from differences.

One of the great advantages of historical comparisons is that they reduce bias induced by culturally and historically limited analyses and interpretations of the social world. Social structures and processes in the past were generally quite different from those observed today, and contemporary institutional arrangements and social relations differ substantially across cultures, regions, and states. Patterns of historical change and continuity, moreover, have varied from one country or culture to another. Some scholars, such as Bendix (1963, [1956] 1974), relish this diversity and use historical comparisons to emphasize and interpret the peculiarities of each case. Even seeming uniformities across cases may mask important differences, and comparative-historical inquiry may be used to detect these “false similarities” (Bloch [1928] 1969).

Examination of historical or national differences also can lead to the detection of previously unknown facts that may suggest a research problem or pose a hypothesis amenable to empirical exploration. Marc Bloch ([1928] 1969), for example, tells of how his knowledge of the English land

enclosures led him to discover similar events in France. Comparing histories also aids in the formation of concepts and the construction of “ideal types” (Weber 1949; Bendix 1963; Smelser 1976). For example, Weber’s ([1904] 1958) concepts of “the Protestant ethic” and “the spirit of capitalism” and Wallerstein’s (1974) notion of the “world system” derive from comparative-historical inquiry. Linking apparently disparate historical and geographical phenomena, as Karl Polanyi ([1944] 1957) does when he relates the gold standard to the relative geopolitical tranquility throughout much of the nineteenth century, also is one of the fruits of comparative-historical analysis. New information, conceptual development and the discovery of unlikely commonalities, linkages, and previously unappreciated differences are necessary for the generation, elaboration, and historical grounding of social theory.

The analysis of comparative-historical patterns can allow for more adequate testing of established theory than does study of a single nation, culture, or time period. Plausible theories of large-scale social change, for example, are intrinsically processual and so require historical analysis for a genuine elaboration or examination of their hypotheses (Tilly 1984). Historical comparisons also can be used to assess the generality of putative “universal” explanations for social structure and social action (e.g., functionalism, Marxism). Abstract propositions thought operative across time and space, for example, can be directly confronted through the analysis of parallel cases that should display the same theoretical process (Skocpol and Summers 1980), thereby specifying a theory’s generality and empirical scope.

Comparative-historical analysis sometimes is directed toward developing explanations that explicitly are “relative” to space and time (Beer 1963) or that represent historically or culturally “limited generalizations” (Joynt and Rescher 1960). Skocpol’s (1979) analysis of social revolutions in France, Russia, and China, for example, resulted in limited causal generalizations deemed valid for these three cases only. Exceptions to theoretical and empirical generality, moreover, can be conceptualized as deviant cases—*anomalies* or *puzzles*. For example, Werner Sombart ([1906] 1976) posed the question “Why is there no socialism in the United States?” precisely because the United

States, when contrasted to the European experience and stacked up against available explanations for the development of class-conscious labor politics, appeared both historically and theoretically anomalous. To explain such puzzles, the original theory is modified and new concepts and theories with greater explanatory power are formulated. The improved theory then serves as the new starting point for subsequent inquiry, so that cumulation of knowledge is facilitated (Stryker 1996). Thus historical patterns are comparatively situated, the “inexplicable” residue of time and place therefore is potentially “explicable” (Sewell 1967), and sociological theory is further developed.

### ANALYTIC TYPES OF HISTORICAL COMPARISONS

Contemporary comparative-historical research practice rests upon diverse, even contradictory, epistemologies and research strategies (Bonnell 1980, Skocpol 1984b, Tilly 1984; Ragin 1987; McMichael 1990; Kiser and Hechter 1991; Griffin 1992; Sewell 1996; Stryker 1996; Mahoney 1999). These can be initially grouped into two basic approaches, labeled here “analytical formalism” and “interpretivism.” Each displays considerable internal diversity, and the two can overlap and be combined in practice. Moreover, they can be synthesized in creative ways, thereby capitalizing on the strengths of each approach and reducing their respective weaknesses. The synthesis, “causal interpretivism,” is sufficiently distinct as to constitute a third approach to the analysis of comparative history.

**Analytically formal comparison.** *Formal comparison* conforms generally to conventional scientific practice in that causal explanation is the goal. It is therefore generally characterized by the development and empirical examination or testing of falsifiable theory of wide historical scope, by the assumption of the preexistence of discrete and identifiable cases, and by the use of formal logical or statistical tools and replicable analytic procedures. Historical narration and the “unities of time and place” (Skocpol 1984b, p. 383) are deliberately replaced by the language of causal analysis. Analytic formal comparison can be used to generalize across time and space, to uncover or produce limited causal regularities among a set of carefully

chosen cases, and to establish a theory’s scope conditions. There are two major procedural subtypes in this genre, statistical comparisons and formal qualitative comparisons.

*Statistical analyses* of comparative-historical phenomena are logically and inferentially identical to statistical analyses of any other social phenomena. Thus, they rely on numerical counts, theoretical models, and techniques of statistical inference to assess the effects of theoretically salient variables, to test the validity of causal arguments, and to develop parsimonious, mathematically precise generalizations and explanations. One important way in which the statistical method appears in comparative-historical inquiry is in the form of “comparative time-series” analysis, in which statistical series charting historical change are systematically compared across countries. Charles, Louise, and Richard Tilly (1978) and Bruce Western (1994), for example, first use statistical time-series procedures to map and explain historical variation in working-class activity within European nations. They then compare, either qualitatively (the Tillys) or quantitatively (Western), these national statistical patterns to detect and explain historical differences and similarities across these countries. The second major use of statistical procedures relies on quantitative data from many social units for only one or a few time points. Emphasis in this “cross-national” tradition (e.g., Chase-Dunn 1979; Jackman 1984; Korpi 1989) is on detecting causal generalizations that are valid for a large sample or even an entire population of countries. Though the results of such studies typically resemble a static, cross sectional snapshot of historical process, this analytic strategy sometimes is necessitated because complete time-series data do not exist for very many nations, especially those now undergoing economic and social development (e.g., Pempel and Williamson 1989). Due both to their data limitations and to their focus on statistical regularities and theory testing, cross-national studies typically slight historical processes and homogenize cultural differences across cases. The understanding of cases as real social units deserving of explanation in their own right consequently is sometimes lost (Tilly 1984; Skocpol 1984b; Ragin 1987; Stryker 1996). In the 1990s, however, some cross-national statistical analysts, though continuing their search for general patterns, began to

incorporate a more thorough appreciation of history and cultural difference into their analyses (e.g., Western 1994).

*Formal qualitative comparison*, by way of contrast, views cases “holistically,” as qualitatively distinct and independent units that cannot (or should not) be decomposed into scores on quantitative variables as in statistical analysis. This strategy adopts a “case-oriented” approach that pervades the entire research process (Ragin 1987). The explanation of a few carefully chosen cases, for example, generally is the rationale for and product of the analysis, and explicit strategies have been designed to select proper instances to be compared and analyzed (Przeworski and Teune 1970; Frensdreis 1983; Lijphart 1971; Stryker 1996). Detection of causal regularities is often inferred through the application of John Stuart Mill’s ([1843] 1967) inductive canons or “method of agreement” and “method of difference” to comparative data (Skocpol 1984b; Ragin 1987). Using the method of agreement, analysts select cases that have positive outcomes on the phenomenon under study but that differ on putative explanatory conditions. These cases are then compared to see what causal factor they share. Alternative explanations are eliminated if antecedent conditions representing those claims do not occur in all cases with positive outcomes. The “method of difference” is used to guard against false inferences adduced from the method of agreement. Here analysis is conducted with cases instancing both positive and negative outcomes but that are as similar as possible on the putative causal factors. The objective of Mill’s methods is to find the one condition that is present in all positive cases and absent in all negative cases (Skocpol 1984b; Ragin 1987).

Mill’s methods are marred because they presuppose that one causal factor or configuration holds for all cases with positive instances. Historical patterns displaying “causal heterogeneity” or “multiple causal conjunctures”—two or more distinct combinations of causal forces generating the same outcome (Ragin 1987)—are therefore logically ruled out. This shortcoming results from the excessive weight Mill’s canons give negative cases: Because exceptions exist to almost any general process, the inability to find causal universals that are doubly confirmed by the twin logics of agreement and difference can rule out virtually any nontrivial explanation (Lijphart 1971; Ragin 1987;

Mahoney 1999). However, Charles Ragin’s (1987) alternative comparative algorithm—“qualitative comparative analysis” (QCA)—allows for the detection of causal heterogeneity in a large number of cases. QCA uses a data reduction strategy rooted in Boolean algebra to search for similarities among positive instances and to exploit the inferential utility of negative cases. Negative or “deviant” cases may be explained by, or give rise to, an alternative causal process, but they are not allowed to invalidate all causal generalizations. Ragin argues that QCA’s Boolean logic most closely approximates the mode of reasoning—including the use of logically possible “historical hypotheticals” and “historical counterfactuals”—employed by Weber (1949) and Barrington Moore (1966) in their powerful but less formalized comparative-historical studies.

Formal qualitative comparison can provide both historically grounded explanation and theoretical generalization. QCA is especially useful because it permits multiple causal configurations to emerge from comparative historical data and largely removes the inferential problems induced by the analysis of a small number of cases of unknown representativeness. Nonetheless, detractors—who otherwise disagree about preferred research strategies—believe that formal comparison via either Mill’s methods or QCA often is fraught with hidden substantive assumptions, unable to exert sufficient analytical control over the multitude of competing explanations, and compromised by its roots in inductive logic (Burawoy 1989; Kiser and Hechter 1991). Moreover, formal comparisons rest on a key assumption that is seriously challenged by proponents of holistic interpretive comparison—that the historical cases themselves are not systematically interrelated (Wallerstein 1974; McMichael 1990).

**Interpretive comparisons.** Interpretive comparisons are most concerned with developing a meaningful understanding of broad cultural or historical patterns (Skocpol 1984b). Two very different comparative logics, “holistic” and “individualizing,” are used to construct historical interpretations. Although neither logic relies extensively on formal analytic procedures, nor is geared toward testing theory, interpretive comparisons are not necessarily atheoretical or lacking in rigor. Indeed, concepts and theories, often of sweeping scope and grandeur, are extensively developed

and deployed, but for the most part as interpretive and organizing frames, or as lenses through which history is understood and represented. Interpretive comparisons often are thought impressive, but sometimes of questionable validity due either to their self-validating logic or to their lack of explicit scientific criteria for evaluating the truth content of the interpretation (Bonnell 1980; Skocpol 1984b; Tilly 1984; Stryker 1996; Mahoney 1999).

Holistic comparisons usually are tied to particular theories and are used when a “social whole” such as a world system (Wallerstein 1974) is methodologically posited. Conceptualizing the entire world system in this manner suggests that there is but one theoretical unit of analysis, and that unit is the world system. What are considered to be separate “units of analysis” or “cases” in most comparative-historical strategies are, in the holistic methodological frame, really only interrelated and interdependent historical realizations of a singular emergent process or social system. Whether nation-states, regions, or cultures, these “moments” therefore are not the discrete and independent units demanded by analytic formalism. Analysts using holistic historical comparisons in this manner often depend on an “encompassing” functional logic that explains similarities or differences among parts of a whole by the relationship the parts have to the whole (Tilly 1984). Thus, Wallerstein’s (1974) comparisons explain the differential development of temporally and spatially specific (though interdependent) economic units in the world system—the core, periphery, and semi-periphery—by their differential positions and roles in the world economy. Such explanations are often circular in their reasoning and always difficult to test, but they may serve as useful illustrations of the workings of a social whole or of the inner logic of a theory (Bonnell 1980; Skocpol 1984b).

Philip McMichael (1990) suggests “incorporated comparison” as an alternative to purge holistic comparison of its mechanistic and self-validating logic. Here a kind of social whole is posited, but it does not preexist and regulate its parts, as does the world system. Instead, it is analytically “reconstituted” by conceptualizing different historical instances as interdependent moments, which, when cumulated and connected through time and space, “form” the whole as a general but empirically diverse historical process. Thus the very definition and selection of “cases” becomes

the object of theorizing and research and not its point of departure or merely the vehicle conveying data for analysis. Karl Polanyi’s ([1944] 1957) analysis of the emergence and decline of laissez-faire capitalism is a compelling example of this form of holistic comparison.

*Individualizing interpretive* analysts use historical comparison to demonstrate the historical and cultural particularity of individual cases. Analysts in this tradition often choose research questions and cases on the basis of their recurrent moral and historical significance rather than scientific representativeness or comprehensiveness. They also tend self-consciously to eschew general theory, quantification, and the methodology of causal analysis. Rather, interpretation is grounded in rich historical narration, in anthropological sensitivity to the import of cultural practice, in the culturally embedded meanings in social action, and in the use of persuasive concepts that crystallize or resonate with significant social values and themes (Bendix 1963; Geertz 1973; Bonnell 1980; Tilly 1984; Skocpol 1984b). Bendix’s ([1956] 1974) research on work and authority in industry is an exemplar of this strategy, and Daniel Goldhagen’s (1996) recent study, *Hitler’s Willing Executioners*, is another much more controversial example.

**Causal Interpretivism.** Historical comparisons rooted in *causal interpretivism* synthesize aspects of both analytical-formal and interpretive-historical comparisons. Inspired in part by Weber’s (1949) formulation of “causal interpretation” (and anthropologist Clifford Geertz’s [1980] similar idea of “interpretive explanation”), this strategy is stamped by explicit causal reasoning, the potential for explanatory generality across comparable instances, and the use of methodological procedures allowing for strict replication, on the one hand, and attention to historical narrative, cultural particularity, and subjective meaning, on the other.

Causal interpretivism has been most fruitfully developed and applied in the analysis of historical narratives (Sewell 1996; Griffin 1992, 1993; Quadagno and Knapp 1992; Somers 1998). Narratives are “sequential accounts” organizing information into “chronological order to tell stories about what happened. . .” and why it happened as it did, and, as such, are “tools for joining sequentiality, contingency, and generalizability” (Stryker 1996, p.305, 307). Narrativists conceptualize historical time so

that it is “eventful.” Any given moment in a historical chain of unfolding sequences is both a repository of past actions and a harbinger of future possibilities (Thompson 1978). Temporal order and the causal relationships among actions that make up events therefore become the major object of inquiry in comparative narrative analysis.

Narrative focuses squarely on the import of human agency in reproducing or changing societal arrangements, thereby complementing the attentiveness comparative-historical sociologists have traditionally given social structural constraints and opportunities (Stryker 1996). Channeling analytical attention in this fashion directs narrativists both to transformative historical happenings (the French Revolution, for instance [Sewell 1996]) and to the development and deployment of decidedly temporal concepts, such as path dependence, sequential unfolding, temporally cumulative causation, and the pace of social action (Abrams 1982; Aminzade 1992). The structure of action underpinning particular historical narratives, moreover, can also be strictly compared with an eye toward both showing general patterns across events and individual exceptions to those generalizations (Abbott 1992; Griffin 1993). Dietrich Rueschemeyer and John Stephens (1997), for example, demonstrate how “eventful time, including the order and sequence of key actions, can both be incorporated into causal generalizations for such large scale processes as democratization and capitalist economic development and [be] exploited to ascertain where those generalizations break down.

Comparative narrativists have developed a variety of procedures to enhance the theoretical payoff and replicability of their research: (1) event-structure analysis (Griffin 1993), which allows explicit codification of reasons for inferences and for cross-level generalization; (2) semantic grammar analysis, which uses linguistic rules to convert text, such as newspaper accounts of strikes in Italy, into numbers for subsequent statistical analysis (Franzosi 1995, 1998); and (3) systematizing electronic means to collect, store, code, and analyze diverse types of historical texts, from newspaper articles to court opinions and legislative hearings (Stryker 1996; Pedriana and Stryker 1997). Robin Stryker (1996) has also developed the useful notion of “strategic narrative” as a guide to case selection in comparative-historical analysis: Strategic narrative suggests

that some historical events and ways of constructing stories will promote theory-building more than will others, thus facilitating cumulation of knowledge.

Regardless of precisely how analytical formalism is incorporated, unpacking a narrative and reconstituting it as an explicit causal account—that is, comprehending and explaining the logical within the chronological, and the general within the particular—often requires that analysts imaginatively reconstruct participants’ cultural understandings (interpretation), as well as systematically harness theoretical abstractions, comparative generalizations, and replicable research methods. In effect, then, analysts explain because they are compelled to interpret, and they interpret because causal explanation is demanded (Beer 1963; Griffin 1993; Mahoney 1999).

#### METHODOLOGICAL ISSUES AND DIFFICULTIES

Though comparative-historical analysis is neither necessarily cross-national or cross-cultural, nor always confronting the same obstacles as narrative history, it does share many of the methodological problems (and solutions) and dilemmas traditionally associated with both types of research. These include, among others: (1) defining and selecting comparable analytical units; (2) case interdependence; (3) the nonrepresentativeness of cases; (4) determining conceptual equivalence and measurement reliability and validity across time and space; (5) the paucity of data, especially that which is quantitative, over long periods of time, and for newly emerging nations; (6) the selectivity and general unsoundness of the historical record; (7) the use of spatially and temporally aggregated data, and, more generally, the distance between what can be collected and systematically measured and what the theory or research question actually calls for; and (8) fruitful ways to wed historical narration and cultural specificity, on the one hand, to the development of general sociological theory and cumulative, replicable results, on the other. Some of these difficulties especially plague analytically formal comparisons (e.g., the artificiality and interdependence of cases; the paucity of systematic, quantifiable data at the proper theoretical level of analysis; and the excessive generality at the cost of important historical and cultural specificity);



others haunt with more force interpretive comparisons (e.g., nonrepresentative cases; analytical looseness and non replicability; and excessive particularity at the expense of generalization and theoretical development).

Granting their very real import, these difficulties are not normally intractable. As we have shown, methodologically self-conscious comparative-historical scholars, both qualitative and quantitative, have devised strategies to: construct and select the most relevant cases for analysis; exploit cross-level data (e.g., within and between nation-states) and make causal inferences across levels of analysis; guard against logically false or self-validating inferences; infuse the analysis of historical narrative with greater theoretical and analytical rigor; and increase the number and representativeness of their cases. They have also devised or imported useful ways to ensure maximum measurement equivalence across cultures; sample historical records; estimate data missing from statistical series and to correct for truncated data; capitalize on case interdependence with new statistical methods; historically ground time-series analysis; etc. (see Bloch [1928] 1969; Carr 1961; Przeworski and Teune 1970; Zelditch 1971; Tuma and Hannan 1984; Skocpol 1984b; Ragin 1987; Kohn 1989; Isaac and Griffin 1989; McMichael 1990; Kiser and Hechter 1991; Griffin 1993; Western 1994; Deane, Beck, and Tolnay 1998).

Some problems, admittedly, have no satisfactory solution, and nothing can compensate for the utter dearth of valid information on particular happenings or variables. This, though, is hardly different from any other form of social research. The litany of methodological difficulties discussed above, in fact, afflict all social research. Survey analysts, for example, deal with questionnaire responses of dubious accuracy and unknown meaning taken from geographically and temporally limited samples, and scholars working with institutional records of any sort must deal with information collected by institutionally self-interested actors.

It is important to recognize that although comparative-historical inquiry is no more "error-prone" or unreliable than are other forms of social research, it often adopts a very different underlying "working philosophy" from most theory-driven quantitative analysis. Usually in highly deductive and statistical analysis (including comparative

strategies), theory construction and theory testing are two clearly defined and quite separate phases of a research program (Kiser and Hechter 1991; Stryker 1996). Segmenting research practice in this way is thought necessary to ensure that hypothesis testing produces a fair empirical test, unbiased by prior analysis, or even intimate foreknowledge of the data. Given these rules and objectives, then, concepts, indicators, hypotheses, and theory are not usually changed in the middle of analysis; that would invalidate the statistical or logical test.

In contrast, comparative-historical inquiry that is qualitative and inductive, though often also analytically formal, establishes a continual dialogue between social theory and empirical evidence, so that conceptual abstraction and comparative history are mutually defined and constructed in reference to each other. In *The Protestant Ethic and the Spirit of Capitalism*, for example, Weber ([1904] 1958) used historical evidence on economic and religious practices and beliefs to refine his ideal types and his hypotheses and, simultaneously, conceptually frame that empirical data to render it intelligible and explicable. Many comparative-historical sociologists follow Weber's general strategy, that is, they start their inquiry with one set of concepts, indicators, or hypotheses, or all three, and end with another, modified set because they mutually adjust data and theory, ensuring a good "fit" between the two as analysis progresses. In their study of capitalist economic development and democratic political practice in many nations, for example, Rueschemeyer, John Stephens, and Evelyne Huber Stephens (1992) amended their theoretical definition of democracy, both broadening its historical scope and specifying more precisely its meaning, as they moved in their analysis from Europe to Latin America. Thus what may look like scientific "cheating" to a deductive or quantitative scholar is an appropriate, fruitful strategy for comparative-historical sociologists whose goals are concept formulation and reformulation, theory construction, and empirical understanding, rather than strict hypothesis testing.

Analytic rigor does not rest with the mechanistic application of technique per se, but, instead, with a systematic and methodologically self-disciplined stance toward both data collection and analysis and the use of evidence to advance inferences. Most comparative-historical scholars adhere

to explicit decision rules, permitting others to replicate, critically judge, and constructively critique the research. To help ensure the validity and reliability of measures in this sort of inquiry, and while also maintaining the needed flexibility to improve key concepts and measures during the analysis, for instance, Stryker (1996) has shown how researchers can adapt traditional content-analytic coding techniques to capture time and space as explanatory context. Context can be conceptualized and measured as a conjunction of “values” on general theoretical constructs, and, at the same time, researchers can develop what she terms “action coding” to construct more systematically the who, what, when, where, how, and why of action and event sequences. With enhanced systematization—that is, the use of precise concepts, measures, and coding techniques, and the clear communication of these—comes enhanced reliability and replicability (Stryker 1996).

#### THE FUTURE OF COMPARATIVE-HISTORICAL ANALYSIS

Jack Goldstone (1997, p. 119) has stated that comparative-historical inquiry has produced “many of the greatest achievements of macro-sociological scholarship.” At the end of the twentieth century, these achievements were severely tested by a series of interdependent, large-scale social transformations of every bit the historical magnitude and moral significance of those motivating sociology’s founders: economic and cultural globalization; the growth of transnational “super states” brought about by strategic military and economic alliances among sovereign nation-states; post-communist marketization and democratization in former Soviet-bloc nations; and armed nationalist struggles and resurgent ethnic conflicts. Broad interest in, and need for, illuminating comparative-historical analysis therefore is only likely to increase.

The utility of research in this tradition rests more on the creative use of good theory than it does on its methodology. But “creative use” of theory is itself largely a matter of how theory is combined with research strategy. On this score, the future of comparative-historical sociology would seem bright.

One encouraging trend is the increasing use of statistical techniques developed explicitly to

analyze temporal (Tuma and Hanna 1984; Isaac and Griffin 1989; Abbott 1992; McCammon 1998) and spatial processes (Deane et al. 1998). Time-series and spatial analyses are not always subject to the same limitations as are cross-national statistical analyses, and quantitative spatial techniques are particularly apt given heightened global interdependence. Spatial analysis, for instance, can uncover how historical processes move from place to place (country, region, etc.), and it can ascertain whether structures or practices in one country are diffused into other countries or deterred from entering them (Deane et al. 1998).

Another very positive trend that probably will become more widespread is combining and synthesizing different kinds of comparative logics. Both Jeffrey Paige (1975) and John Stephens (1980) combined cross-national quantitative analysis of a large number of nation-states with qualitative, case-oriented work on a small, but theoretically strategic subset of their cases. Paige combined his statistical generalizations with three parallel case studies, finding differences in the nature of agrarian revolts. Stephens, on the other hand, systematically compared a subset of four countries to show how the historical processes of welfare-state development differed from one nation to another. He therefore used both generalizing and individualizing comparative logics. Goldstone (1991) followed a similar strategy, combining “within-case” quantitative analyses with case-oriented comparisons to show how long-term population growth contributed to state breakdown in the early modern world. Combining diverse comparative-historical approaches in this way, and synthesizing analytical formalism and interpretation in causal interpretivism, yield insights that simply are not possible with any single methodological strategy.

Even if the differences among diverse approaches prove too great to overcome completely—a possibility suggested by the debate over the utility of highly deductive explanatory schemas such as rational choice theory in comparative-historical research (Quadagno and Knapp 1992; Kiser and Hechter 1991, 1998; Somers 1998)—diversity in goals, strategies, and procedures has spurred methodological innovation and the cross-fertilization of scholarship from divergent methodological strategies (Beer 1963; Rueschemeyer and Stephens 1997; Goldstone 1997). It thereby provides the continued promise of enhanced knowledge and richly

textured, multi-voiced portrayals and understandings of perennial questions of fundamental human significance.

## REFERENCES

- Abbott, Andrew 1992 "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods and Research* 20:428-55.
- Abrams, Philip 1982 *Historical Sociology*. Ithaca, N.Y.: Cornell University Press.
- Aminzade, Ron 1992 "Historical Sociology and Time." *Sociological Methods and Research* 20:456-480.
- Beer, Samuel 1963 "Causal Explanation and Imaginative Reenactment." *History and Theory* 3:6-29.
- Bendix, Reinhard (1956) 1974 *Work and Authority in Industry: Ideologies of Management in the Course of Industrialization*. Berkeley: University of California Press.
- 1963 "Concepts and Generalizations in Comparative Sociological Studies." *American Sociological Review* 28:532-539.
- Bloch, Marc (1928) 1969 "A Contribution Towards Comparative History of European Societies." In M. Bloch, ed., *Life and Work in Medieval Europe: Selected Papers of Marc Bloch*. New York: Harper Torchbooks.
- Bonnell, Victoria 1980 "On the Uses of Theory, Concepts and Comparison in Historical Sociology." *Comparative Studies in Society and History* 22:156-173.
- Burawoy, Michael 1989 "Two Methods in Search of Science: Skocpol versus Trotsky." *Theory and Society* 18:759-805.
- Carr, E. H. 1961. *What is History?* New York: Vintage.
- Chase-Dunn, Christopher 1979 "Comparative Research on World System Characteristics." *International Studies Quarterly* 23:601-625.
- Deane, Glenn, E.M. Beck, and Stewart E. Tolnay 1998 "Incorporating Space into Social Histories: How Spatial Processes Operate and How we Observe Them." *International Review of Social History* 43:57-80.
- Franzosi, Roberto 1998 "Narrative as Data: Linguistic and Statistical tools for the Quantitative Study of Historical Events." *International Journal of Social History* 43:81-104.
- Frendreis, James 1983 "Explanation of Variation and Detection of Covariation: The Purpose and Logic of Comparative Analysis." *Comparative Political Studies* 16:255-272.
- Geertz, Clifford 1980 "Blurred Genres: The Reconfiguration of Social Thought." *American Scholar* 49:165-179.
- 1973 *The Interpretation of Cultures*. New York: Basic Books.
- Goldhagen, Daniel 1996 *Hitler's Willing Executors: Ordinary Germans and the Holocaust*. New York: Knopf.
- Goldstone, Jack 1997 "Methodological Issues in Comparative Macrosociology." *Comparative Social Research* 16:107-120.
- 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- Griffin, Larry J. 1993 "Narrative, Event-Structure, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98:1094-1133.
- 1992 "Temporality, Events and Explanation in Historical Sociology: An Introduction." *Sociological Methods and Research* 20:403-27.
- Hibbs, Douglas 1978 "On the Political Economy of Long-run Trends in Strike Activity." *British Journal of Political Science* 8:51-75.
- Isaac, Larry, and Larry J. Griffin 1989 "Ahistoricism in Times Series Analysis of Historical Process: Critique, Redirection and Illustrations from U.S. Labor History." *American Sociological Review* 54:873-890.
- Jackman, Robert 1984 "Cross National Statistical Research and the Study of Comparative Politics." *American Journal of Political Science* 28:161-182.
- Joynt, Cary, and Nicholas Rescher 1960 "The Problem of Uniqueness in History." *History and Theory* 1:150-162.
- Kiser, Edgar, and Michael Hechter 1998 "The Debate on Historical Sociology: Rational Choice Theory and Its Critics." *American Journal of Sociology* 104:785-816.
- 1991 "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97:1-30.
- Kohn, Melvin (ed.) 1989 *Cross-National Research in Sociology*. Beverly Hills, Calif.: Sage.
- Korpi, Walter 1989 "Power, Politics and State Autonomy in the Development of Social Citizenship: Social Rights During Sickness in Eighteen OECD Countries Since 1930." *American Sociological Review* 54:309-328.
- Lijphart, Arend 1971 "Comparative Politics and the Comparative Method." *American Political Science Review* 65:682-93.
- Mahoney, James 1999 "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis" *American Journal of Sociology* 104:1154-1196.
- McCammon, Holly 1998 "Using Event History Analysis in Historical Research: With Illustrations from a Study of the Passage of Women's Protective Legislation." *International Review of Social History* 43:33-55.

- McMichael, Philip 1990 "Incorporating Comparison within a World-Historical Perspective: An Alternative Comparative Method." *American Sociological Review* 55:385-397.
- Mill, John Stuart (1843) 1967 *A System of Logic: Ratiocinative and Inductive*. Toronto: University of Toronto Press.
- Moore, Barrington 1966 *The Social Origins of Dictatorship and Democracy*. Boston: Beacon Press.
- Paige, Jeffrey 1975 *Agrarian Revolt: Social Movements and Export Agriculture in the Underdeveloped World*. New York: Free Press.
- Pampel, Fred, and John Williamson 1989 *Age, Class, Politics and the Welfare State*. Cambridge: Cambridge University Press.
- Pedriana, Nicholas, and Robin Stryker. 1997. "Political Culture Wars 1960s Style: Equal Employment Opportunity, Affirmative Action Law and the Philadelphia Plan." *American Journal of Sociology* 103:633-691.
- Polanyi, Karl (1944) 1957 *The Great Transformation: The Political and Economic Origins of Our Times*. Boston: Beacon Press.
- Przeworski, Adam, and Henry Teune 1970 *The Logic of Comparative Social Inquiry*. New York: Wiley.
- Quadagno, Jill, and Stan Knapp 1992 "Have Historical Sociologists Forsaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20:481-507.
- Ragin, Charles 1987 *The Comparative Method*. Berkeley: University of California Press.
- Rueschemeyer, Dietrich, and John Stephens 1997 "Comparing Historical Sequences: A Powerful Tool for Causal Analysis." *Comparative Social Research* 17:55-72.
- , and Evelyne Huber Stephens 1992 *Capitalist Development and Democracy*. Chicago: University of Chicago Press.
- Sewell, William, Jr. 1996. "Three Temporalities: Toward an Eventful Sociology." In, T. J. McDonald, ed., *The Historic Turn in the Human Sciences*. Ann Arbor: University of Michigan Press.
- 1967 "Marc Bloch and the Logic of Comparative History." *History and Society* 6:208-218.
- Skocpol, Theda (ed.) 1984a *Vision and Method in Historical Sociology*. New York: Cambridge University Press.
- 1984b "Emerging Agendas and Recurrent Strategies in Historical Sociology." In T. Skocpol, ed., *Vision and Method in Historical Sociology*, New York: Cambridge University Press.
- 1979 *States and Social Revolutions*. Cambridge: Cambridge University Press.
- , and Margaret Somers 1980 "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Study of Society and History* 22:174-197.
- Smelser, Neil 1976 *Comparative Methods in the Social Sciences*. Englewood Cliffs, N.J.: Prentice-Hall.
- Sombart, Werner (1906) 1976 *Why Is There No Socialism in the United States?* White Plains, N.Y.: M.E. Sharpe.
- Somers, Margaret 1998 "We're No Angels": Realism, Rationality Choice and Relationality in Social Science." *American Journal of Sociology* 104:722-784.
- Stephens, John 1980 *The Transition from Capitalism to Socialism*. Atlantic Highlands, N.J.: Humanities Press.
- Stryker, Robin 1996 "Beyond History vs. Theory: Strategic Narrative and Sociological Explanation." *Sociological Methods and Research* 24:304-352.
- Thompson, Edward P. 1978 *The Poverty of Theory and Other Essays*. London: Merlin.
- Tilly, Charles 1984 *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- , Louise Tilly, and Richard Tilly 1975 *The Rebelious Century 1830-1930*. Cambridge, Mass.: Harvard University Press.
- Tuma, Nancy, and Michael Hannan 1984 *Social Dynamics*. New York: Academic Press.
- Wallerstein, Immanuel 1974 *The Modern World System*. New York: Academic Press.
- Weber, Max (1904) 1958 *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribners.
- 1949 *The Methodology of the Social Sciences*. New York: Free Press.
- Western, Bruce 1994 "Unionization and Labor Market Institutions in Advanced Capitalism, 1950-85." *American Journal of Sociology* 99:1314-1341.
- Zelditch, Morris 1971 "Intelligible Comparisons." In I. Vallier, ed., *Comparative Methods in Sociology*, Berkeley: University of California Press.

LARRY J. GRIFFIN  
ROBIN STRYKER

## COMPLEX ORGANIZATIONS

Organizations, as a class, are socially constructed innovations, deliberately designed as solutions to problems. Although some forms of organization, such as churches and armies, have been around for centuries, only since the Industrial Revolution have complex organizations assumed the form people take for granted today. Because they are

shaped by the contexts or environments in which they are established, contemporary organizations reflect the impact of their historical origins in societies characterized by growing affluence and conflicts over the control and distribution of wealth. Organizations come in a bewildering variety of forms because they have been explicitly designed to deal with a wide range of problems and because they have emerged under widely varying environmental conditions.

### THE IMPORTANCE OF ORGANIZATIONS

Why are complex organizations important? Organizations, which produce goods, deliver services, maintain order, and challenge the established order are the fundamental building blocks of modern societies, and the basic vehicles through which collective action is undertaken. The prominence of organizations in contemporary society is apparent when we consider some consequences of their actions.

Organizations coordinate the actions of people in pursuit of activities too broad in scope to be accomplished by individuals alone. Railroads were the first large corporations in the United States, arising because small autonomous merchants and traders could not effectively coordinate the passage of long-distance shipments. In the twentieth century, the production of mass-market consumption goods, such as automobiles and electric appliances, was made possible by the rise of large, vertically integrated manufacturing firms. Similarly, in the public sector, the implementation of government social policies has necessitated the development of large government agencies that process thousands of cases on a universalistic, impersonal basis.

The concentration of power in organizations contributes not only to the attainment of large-scale goals in modern societies, but also to some major social problems. When organizations focus on attaining specific goals, such as making profits for shareholders, they may neglect the side effects or externalities of their actions. Examples include air and water pollution and the careless disposal of unwanted hazardous materials by large manufacturers. The manufacture of unsafe consumer products and instances of large-scale fraud and collusion also illustrate the capacity of organizations to do harm as well as good.

Increasingly, major tasks in society are addressed not by single organizations but instead by networks of interdependent organizations. Interorganizational arrangements pool the efforts of numerous specialized organizations in pursuit of a common end. For example, automobiles are produced in *vertical* production networks centered on assemblers that maintain ongoing relationships with multiple tiers of suppliers responsible for the manufacture and delivery of components. However, in the industrial districts of northern Italy, textile production occurs within *horizontal* production networks consisting of small and specialized firms that employ flexible machinery and temporarily cooperate with one another for particular projects. Cooperation enables them to adapt to changing market demands and to learn from one another. Likewise, services for the severely mentally ill are delivered by independent but interrelated agencies providing specific medical and social services. Many large-scale motion pictures are produced not by self-contained studios, but rather by combinations of producers, directors, cinematographers, and other personnel assembled only for the duration of a particular project. Ongoing arrangements involving hospitals, doctors, and university laboratories have been created by the National Cancer Institute to coordinate cancer research and treatment. As these examples demonstrate, interorganizational arrangements give organizations the flexibility to address the unique features of specific cases, while avoiding the rigidities sometimes found in very large organizations.

### DEFINITIONS OF ORGANIZATIONS

What are complex organizations? A simple definition is that organizations are goal-directed, boundary-maintaining, socially-constructed systems of human activity. Some definitions add other criteria, such as deliberateness of design, the existence of status structures, patterned understandings between participants, orientation to an environment, possession of a technical system for accomplishing tasks, and substitutability of personnel (Scott 1998).

Goal orientation and deliberate design of activity systems are distinctive features discriminating between organizations and other collectivities, such as families and small groups. Organizations are purposive systems in which members behave as if they are committed to the organization's

goals, although individual participants might personally feel indifferent toward those goals or even alienated from their organizations. Concerted collective action toward an apparent common purpose also distinguishes organizations from social units such as friendship circles, audiences, and mass publics. Because many organizational forms are now institutionalized in modern societies, people readily turn to them or construct them when a task or objective exceeds their own personal abilities and resources (Meyer and Rowan 1977; Zucker 1988).

Organizations have activity systems—or technologies—for accomplishing work, which can include processing raw materials, information, or people. Activity systems consist of bounded sets of interdependent role behaviors; the nature of the interdependencies is often contingent upon the techniques used.

Other key elements of organizations, such as socially constructed boundaries, are shared with other types of collectivities. The establishment of an “organization” implies a distinction between members and nonmembers, thus marking off organizations from their environments. Maintaining this distinction requires boundary-maintenance activity, because boundaries may be permeable, and thus some organizations establish an authoritative process to enforce membership distinctions. For example, businesses have human resource management departments that select, socialize, and monitor employees, and voluntary associations have membership committees that perform similar functions. Distinctive symbols of membership may include unique modes of dress and special vocabularies.

Within organizations, goal attainment and boundary maintenance manifest themselves as issues of coordination and control, as authorities construct arrangements for allocating resources or integrating workflows. These internal structures affect the perceived meaning and satisfaction of individual participants by, for example, differentially allocating power and affecting the characteristics of jobs. Control structures, which shape the way participants are directed, evaluated, and rewarded are constrained by participants’ multiple external social roles, some complementing, but others conflicting with organizational roles.

## ENVIRONMENTS

Organizations, with few exceptions, are incomplete social systems that are not self-sufficient and thus depend on interchanges with their environments. Therefore, goal setting by owners or leaders must take into account the sometimes contrary preferences of organizations and other actors, because activity systems are fueled by resources obtained from outsiders. For example, participants must be enticed or coerced into contributing to the organization’s activities: businesses pay people to work for them, while voluntary nonprofit organizations may offer more intangible benefits, such as sociable occasions.

Over the past few decades, organizational sociology has gradually expanded its scope to include more of the external elements associated with organizational life. Initially, theorists emphasized relatively tangible features: environments were seen as stocks of resources (such as raw materials, capital, or personnel) and information. To the degree that resources are scarce or concentrated in a few hands, organizations are more dependent on their environments and may be vulnerable to exploitation or external control by outsiders (Pfeffer and Salancik 1978). To the degree that information is dispersed, heterogeneous, or subject to change, organizations confront problems of uncertainty.

Much research has been devoted to understanding how organizations deal with the problems of dependence and uncertainty that environments pose. Some organizations respond internally, by developing specialized structures and processes. For example, boundary-spanning personnel are charged with dealing with actors outside the organization, acquiring resources, and disposing of products; they thereby enable others in the technical core to work under more certain conditions. Some organizations respond externally, by building bridges that stabilize their ties to environments. They use interorganizational devices such as long-term contracts and working agreements, joint programs or alliances, and interlocking directorates (Mizruchi and Galaskiewicz 1993).

More recent scholarship has stressed the socially constructed, *institutional* aspects of environments. Organizations must be attentive to these external aspects to acquire and maintain legitimacy and social standing. Scott (1995) distinguished

between regulative, normative, and cognitive-cultural understandings of institutions. Regulative institutions are external systems of rules, often established by states, to which organizations must conform. External rules may be prescriptive, such as when a law requires that publicly traded corporations have boards of directors with at least three members. Rules may also be proscriptive, such as when regulatory bodies or courts prohibit particular trade practices and ownership patterns.

Normative institutions are sets of beliefs and values shared by organizations and their participants about the morally appropriate way in which to organize activities. Professional associations are common sources of normative prescriptions, enforced through processes such as certification. Cognitive institutions are cultural frames or scripts that constitute the social forms seen as natural and taken for granted in a given society. In contemporary society, the idea of organizations itself is institutionalized; legitimate pursuit of an activity often requires the formation of an organization because of the widespread belief that this is an efficient, effective, and fair manner in which to proceed (Meyer and Rowan 1977).

### ORGANIZATIONAL DEMOGRAPHICS

The number of organizations in industrial societies is very large. Over five million businesses with at least one employee existed in the United States in 1998, and there were thousands of governmental agencies, nonprofit organizations, and voluntary associations (Aldrich 1999). Social scientists have developed various schemes to describe the diversity of organizations. For example, the North American Industry Classification System (Office of Management and the Budget 1997) classifies establishments by their products and the processes used to make them. Organizations can also be classified according to their goals or the social functions they serve (Parsons 1960), or on the basis of the prime beneficiaries of organizational actions: owners, members, clients, or the general public (Blau and Scott 1962). Another classification contrasts generalist and specialist organizations and hypothesizes that each type thrives in a different kind of environment (Hannan and Freeman 1989).

Industrial societies contain a large number of organizations, with three characteristics of special

note. First, the size distribution of businesses and nonprofit organizations is highly skewed, with a large number of very small organizations. In 1990, approximately 90 percent of the five million firms in the United States with employees employed fewer than twenty people. Over 98 percent employed fewer than one hundred workers (Small Business Administration 1994). Employment statistics were very similar in the European Union (ENSR 1993), with about 93 percent of firms employing fewer than ten workers, although the figure varied by nation. In 1994, most of the 4.3 million corporations in the United States had fewer than \$one hundred,000 in assets, and they accounted for less than 0.3 percent of all corporate assets. Thus, most organizations are fairly vulnerable to environmental forces and must adapt to them or disband.

Second, although the number of large organizations is small, they have achieved a dominant share of revenues and assets. Business organizations are highly stratified by size, and large firms have more resources with which to resist and counter environmental pressures. In 1994, the top 0.002 percent of corporations, each with a quarter-billion or more in assets, held about 83 percent of all corporate assets. In 1992, the fifty largest manufacturing firms in the United States accounted for about 24 percent of value added and about 13 percent of employment, and the two hundred largest firms accounted for 42 percent of value added and around 25 percent of employment (U.S. Bureau of the Census 1996).

Third, smaller organizations still employ a relatively large share of all workers. In the United States, about 39 percent of all employees work in firms that employ fewer than one hundred workers. In particular industries, such as agriculture and construction, substantially more than half the work force is employed in organizations with fewer than one hundred workers. In the European Union, firms with fewer than one hundred workers employ about 55 percent of the labor force, although variation across nations is substantial. For example, in Germany and the United Kingdom, firms larger than one hundred employ more than half the labor force, whereas in Greece and Italy, such firms employ only about 20 percent of the labor force (ESNR 1993).

Populations of organizations in modern societies are constantly undergoing processes of expansion, contraction, and change. Some organizations are founded in a flash of creative energy and then disband almost immediately, whereas others emerge slowly and then last for decades. Some organizations adapt readily to every environmental challenge, whereas others succumb to the first traumatic event they face. Sociologists have turned their attention to these vital events surrounding the reproduction and renewal of organizational populations, focusing on three processes: foundings, transformations, and disbandings (Aldrich 1999; Hannan and Freeman 1989).

New organizations are established fairly frequently, although systematic data on founding rates are available only for businesses. Studies in the United States and other Western industrialized nations show that about ten businesses are founded per year for every one hundred businesses active at the start of the year. In the United States, approximately 4 to 6 percent of the adult population in the mid-1990s has engaged in some activities with an intent to start a business, and about half of them actually succeeded in their initial efforts (Reynolds and White 1997). Explanations for variations in rates of organizational foundings have stressed the characteristics of opportunity structures, the organizing capacities of groups, and strategies adopted by entrepreneurs as they take account of opportunities and resources available to them (Aldrich and Wiedenmayer 1991).

Societal demands for special-purpose organizations increased with urbanization and with economic, political, and social differentiation, while the resources required to construct organizations grew more abundant with the development of a money economy and the spread of literacy (Stinchcombe 1965). The spread of facilitative legal, political, and other institutions also played a major role, by creating a stable, predictable context within which entrepreneurs could look forward to appropriating the gains from organizational foundings. Occasional periods of political upheaval and revolution stimulate foundings by freeing resources from previous uses, and thus massive changes occurred in the organizational populations of Eastern Europe in the 1990s. For example, centralized health care systems were disbanded with the demise of state socialist regimes,

and there is a good deal of flux in the organizations within this sector (Mechanic and Rochefort 1996).

Transformations occur when existing organizations adapt their structures to changing conditions. The issue of how frequently and under what conditions organizations change has provoked some of the most spirited debates in organizational sociology. Strategic choice-theorists have argued for managerial autonomy and adaptability, whereas ecological and institutional theorists have tended to stress organizational inertia and dependence. Research has moved away from polarizing debates and reframed the question of transformation, asking about the conditions under which organizations change and whether changes occur more frequently in core or peripheral features (Singh and Lumsden 1990).

In the 1980s and 1990s, growing awareness of techniques for performing dynamic analyses have produced some useful studies of transformation, such as those on diversification, top executive changes, and changes in corporate form (Fligstein 1991). These studies tell us that changes *do* occur, although they do not report whether rates of change go up or down over an organization's life cycle. Most of these studies are of the very largest business firms, for which data are publicly available, and not for representative samples.

If all newly founded organizations lived forever, the study of organizational change would be confined to issues of founding, adaptation, and inertia. Research has shown that organizations disband at a fairly high rate, however, and a sizable literature has grown up on organizational mortality (Baum 1996; Singh 1990). Organizations can cease to exist as separate entities in two ways: by completely dissolving—the process by which the vast majority of organizations disband—or by becoming part of a different entity through merger or acquisition. Throughout the 1990s, between two and six thousand mergers and acquisitions a year took place, amounting to less than 1 percent of the incorporated firm population in any given year (Aldrich 1999). For example, in 1994, mergers and acquisitions involved about one-tenth of one percent of all corporate assets. By contrast, between 1976 and 1984, the annual rate of dissolution in the business population of the United States was about 10 percent (Small Business Administration 1994).



Age and size are the strongest predictors of how long an organization will survive. Young organizations disband at a substantially higher rate than older ones; a conservative estimate is that only half of new organizations survive more than five years. Internally, new organizations depend upon the cooperation of strangers who must be taught new routines, some of which are unique to particular organizations (Stinchcombe 1965). Externally, new organizations must penetrate niches in potentially hostile environments, overcoming competitors and establishing their legitimacy with potential members, customers, suppliers, and others. Survival beyond infancy is easier when an organization adopts a form that has already been institutionalized and is widely regarded as legitimate and proper (Zucker 1988). Population ecologists have given special attention to the density of an organizational form—the number of organizations of a given kind—in their analyses of vital rates. Density is thought to be an indicator both of the acceptance or legitimacy of a form (which raises foundings and decreases disbandings) and of the number of competitors attempting to survive within a given environmental niche (which reduces foundings and increases disbandings). The life chances of an organization initially increase—though not smoothly—as density rises, and then decrease as competition intensifies.

#### **SOURCES OF INTERNAL DIVERSITY AMONG ORGANIZATIONS**

All models of organizations as coherent entities can be reduced to two basic views. The systemic view sees organizations as social systems, sustained by the roles allocated to their participants, whereas an associative perspective treats organizations as associations of self-interested parties, sustained by the rewards the participants derive from their association with the organization (Swanson 1971). These two views each have a venerable heritage in the social sciences. Despite subtle variations, all perspectives on organizations ultimately use one or both of these models.

Institutional, functionalist, and ecological perspectives rely on a systemic model, viewing organizations as relatively coherent, stable entities. Such models emphasize the activity systems in organizations that are deliberately designed to accomplish specific goals. Formal structures of organizations,

including a division of labor, authority relationships, and prescribed communication channels, are treated as fulfilling a purposeful design. For example, an institutional approach emphasizes member socialization and other processes that make the transmission of shared meanings easier. Ecological models usually treat organizations as units that are being selected for or against by their environments, and thus assume that organizations cohere as units.

Interpretive and more micro analytic views rely more on an associative model, leading to the expectation that organizations are constantly at risk of dissolution (March and Olsen 1976). For example, in the interpretive view, the reproduction of organizational structure depends on participants resubscribing to, or continually negotiating, a shared understanding of what they jointly are doing. Some cultural theories of organization emphasize the different, conflicting views that coexist within one organization (Martin 1992).

Views of organizations as marketplaces of incentives (Dow 1988), bundles of transactions (Williamson 1981), or arenas of class conflict (Clegg 1989), are in harmony with the associative view, insofar as they focus on actors' contributions to sustaining interaction. Indeed, organizational economics views organizations primarily as mechanisms for mediating exchanges among individuals, arguing that they arise only when market coordination has proven inadequate (Williamson 1981). One variant on this approach, agency theory, focuses on the relationships between self-interested principals (such as owners or stockholders) and the self-interested agents (such as workers or managers) engaged on behalf of principals. The self-interests of these actors may be divergent, and hence agency theory examines how systems of incentives, monitoring, and coordination can be designed to align the activities of agents with the objectives of principals, at minimal cost.

These complementary views of organizations—the associative and the systemic—highlight the sources of two fundamental problems of social organization, those of differentiation and integration. Some differentiation occurs through the division of labor among different roles and subunits; for example, employees may be divided into departments such as sales, finance, and manufacturing. Differentiation pressures also arise because

participants sometimes bring widely varying expectations to the same organization. Differentiation is thus a centrifugal force threatening the coherence of social units. Integration, by contrast, refers to procedures for maintaining coherence, as diverse roles are linked and activities coordinated to sustain an organization as a coherent entity. Examples of integrative processes include the holding of weekly departmental meetings or the circulation of inter office memos.

Differentiation increases organizational complexity because it increases the extent and nature of specialization (Blau 1972). Complexity increases with the number of different components, and may be horizontal (tasks spread over many roles or units), vertical (many levels in a hierarchy of authority), or spatial (many operating sites). Complexity also increases when tasks are grouped by product/market (soap, paper products, or foods) or by function (finance, production, or marketing).

Problems of coordination are present for any activity system, but especially for a complex one. Many concepts used to describe organization structure involve alternative processes used in attempts to achieve integration. One scheme, for example, identified five coordinating mechanisms: direct supervision, three forms of standardization, and mutual adjustment (Mintzberg 1983). With direct supervision, or simple control, decision making is highly centralized: persons at the top of a hierarchy make decisions that lower level personnel simply carry out. This coordination pattern was prevalent in preindustrial organizations, and today is especially likely within small organizations (Edwards 1979).

Coordination also can be attained through standardizing work processes, skills, or outputs. Work processes may be standardized through formalization, that is, the development of rules and procedures. Examples include rules for processing orders, for assembling and packaging products, or for conducting screening interviews for clients. A formalized organization may appear decentralized, since few explicit commands are given, and lower-level participants have freedom in making decisions within the rules. The rules may, however, be so restrictive as to leave little room for discretion. A variant on process standardization is technical control; here, rules and

procedures—and thereby coordination and control—are built into the design of machinery, as in an assembly line. Most worker discretion is eliminated by the structure of the technical system, and what remains is centralized in the upper echelons of the organization.

Standardization of skills involves considerable training and indoctrination of personnel, so that participants will carry out organizational policies with minimal oversight. Organizations employing large numbers of professionals are likely to rely on this coordination strategy (Von Glinow 1988). Professional participants enjoy considerable autonomy in making decisions, but their prior socialization sets most decision premises for them.

By producing products with standard properties, subunits of an organization are able to work independently of one another; if they use each other's outputs, the standards tell them what to anticipate. For example, large clothing firms produce massive runs of identical garments, and thus the various departments within firms know precisely what to expect from one another as they follow daily routines. Similar purposes are served by precise tolerances for parts used in manufacturing machinery and information systems designed to permit easy sharing of data within offices.

In small, young organizations, coordination is achieved by the simple mutual adjustment of personnel to one another. This form of coordination is also found in larger organizations working on complex, novel tasks, for which innovation is at a premium, such as consulting firms or research and development units. In such organizations, the activities of participants are reciprocally contingent, rather than sequentially ordered as in an assembly line. Formal structural devices have been advocated as ways of facilitating mutual adjustment under complex circumstances, such as integrating managers and project teams into matrix-management structures.

Standardization, bureaucratization, and formalization have done much to enable organizations to perform repetitive work efficiently. The same phenomena contribute to some well-known organizational pathologies; in particular, these coordination devices tend to reduce flexibility. In the 1990s, organizations struggled to develop less rigid structures for accomplishing work. Efforts to

achieve flexibility included reductions in size, expansion of the scope of employee responsibilities through cross-training and broadened job descriptions, reductions in layers of management, and maintenance of long-term relationships with suppliers as an alternative to vertical integration. Many of these devices rely on strengthening the internal social networks that link an organization's personnel as well as the external ones linking it to the environment. Such linkages enable organizations to remain adaptable, cope with market uncertainties, and take advantage of opportunities for learning.

### CONCLUSION

Complex organizations are the building blocks of industrial and post industrial societies, enabling people to accomplish tasks otherwise beyond their individual competencies. Organizations also help create and sustain structures of opportunities and constraints that affect nearly every aspect of societies. Sociologists and other social scientists are grappling with a variety of complementary and conflicting perspectives on organizations, ranging from microanalytic interpretive views to macroevolutionary institutional and ecological views. Regardless of perspective, research on organizations has shown its capacity to interest, inform, and provoke us, as sociologists investigate these social constructions that figure so prominently in our lives.

### REFERENCES

- Aldrich, Howard E. 1999 *Organizations Evolving*. London: Sage.
- , and Gabriele Wiedenmayer 1991 "From Traits to Rates: An Ecological Perspective on Organizational Foundings." In Jerome Katz and Robert Brockhaus, eds., *Advances in Entrepreneurship, Firm Emergence, and Growth*, vol. 1, 145–114 Greenwich, Conn.: JAI.
- Baum, Joel A.C. 1996 "Organizational Ecology." in Stewart R. Clegg, Cynthia Hardy, and Walter Nord, eds., *Handbook of Organization Studies* 77–114 London: Sage.
- Blau, Peter M. 1972 "Interdependence and Hierarchy in Organizations." *Social Science Research* 1:1–24.
- , and W. Richard Scott 1962 *Formal Organizations*. San Francisco: Chandler.
- Clegg, Stewart 1989 "Radical Revisions: Power, Discipline and Organizations." *Organization Studies* 10:97–115.
- Dow, Gregory K. 1988 "Configurational and Coactivational Views of Organizational Structure." *Academy of Management Review* 13:53–64.
- Edwards, Richard C. 1979 *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*. New York: Basic Books.
- ENSR 1993 *The European Observatory for SMEs: First Annual Report*. Zoetermeer, The Netherlands: European Observatory for SME Research and EIM Small Business Research and Consultancy.
- Fligstein, Neil 1991 *The Transformation of Corporate Control*. Cambridge, Mass.: Harvard University Press.
- Hannan, Michael T., and John H. Freeman 1989 *Organizational Ecology*. Cambridge, Mass: Harvard University Press.
- March, James G., and Johan P. Olsen 1976 *Ambiguity and Choice in Organizations*. Bergen, Norway: Universitetsforlaget.
- Martin, Joanne 1992 *Cultures in Organizations: Three Perspectives*. New York: Oxford University Press.
- Mechanic, David, and David A. Rochefort 1996 "Comparative Medical Systems." *Annual Review of Sociology* 22:239–270.
- Meyer, John W., and Brian Rowan 1977 "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83:340–363.
- Mintzberg, Henry 1983 *Structure in Fives: Designing Effective Organizations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Mizruchi, Mark S., and Joseph Galaskiewicz 1993 "Networks of Interorganizational Relations." *Sociological Methods & Research* 22:46–70.
- Office of Management and Budget, Statistical Policy Division 1997 *North American Industry Classification System*. Washington, D.C.: U.S. Government Printing Office.
- Parsons, Talcott 1960 *Structure and Process in Modern Societies*. New York: Free Press.
- Pfeffer, Jeffrey, and Gerald R. Salancik 1978 *The External Control of Organizations*. New York: Harper and Row.
- Reynolds, Paul D., and Sammis B. White 1997 *The Entrepreneurial Process: Economic Growth, Men, Women, and Minorities*. Westport, Conn.: Quorum Books.
- Scott, W. Richard 1998 *Organizations: Rational, Natural, and Open Systems*, 4th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- 1995 *Institutions and Organizations*. Thousand Oaks, Calif.: Sage.
- Singh, Jitendra (ed.) 1990 *Organizational Evolution*. Beverly Hills, Calif.: Sage.

- , and Charles Lumsden 1990 “Theory and Research in Organizational Ecology.” *Annual Review of Sociology* 16:161–195.
- Small Business Administration 1994 *Handbook of Small Business Data*. Washington, D.C.: U.S. Government Printing Office and Office of Advocacy, Small Business Administration.
- Stinchcombe, Arthur 1965 “Social Structure and Organizations.” In J. G. March, ed., *Handbook of Organizations*. 142–193 Chicago: Rand-McNally.
- Swanson, Guy E. 1971 “An Organizational Analysis of Collectivities.” *American Sociological Review* 36:607–614.
- U.S. Bureau of the Census 1996 Selected statistics from the 1992 Census of Manufactures report MC92-S-2 “Concentration Ratios in Manufacturing.” <http://www.census.gov/mcd/mancen/download/mc92cr.sum> (21 Jan. 1999)
- Von Glinow, Mary Ann 1988 *The New Professionals: Managing Today's High-Tech Employees*. Cambridge, Mass.: Ballinger.
- Williamson, Oliver 1981 “The Economics of Organization: The Transaction Cost Approach.” *American Journal of Sociology* 87:548–577.
- Zucker, Lynne G. (ed.) 1988 *Institutional Patterns and Organizations: Culture and Environment*. Cambridge, Mass.: Ballinger.

HOWARD E. ALDRICH  
PETER V. MARSDEN

## COMPLIANCE AND CONFORMITY

*Conformity* is a change in behavior or belief toward a group standard as a result of the group's influence on an individual. As this definition indicates, conformity is a type of social influence through which group members come to share similar beliefs and standards of behavior. It includes the processes by which group members converge on a given standard of belief or behavior as well as the pressures they exert on one another to uphold such standards. *Compliance* is behavioral conformity in order to achieve rewards or avoid punishments (Kelman 1958). Since one can behaviorally adhere to a group standard without personally believing in it, the term is often used to indicate conformity that is merely public rather than private as well. Compliance can also refer to behavioral conformity to the request or demand of another, especially an authority.

In an individualistic society such as the United States, conformity has a negative connotation (Markus and Kitiyama 1994). Yet conformity is a fundamental social process without which people would be unable to organize into groups and take effective action as a collectivity. For people to coordinate their behavior so that they can organize and work together as a group, they must develop and adhere to standards of behavior that make each other's actions mutually predictable. Simply driving down a street would be nearly impossible if most people did not conform to group norms that organize driving.

Conformity is also the process that establishes boundaries between groups. Through the conformity process, the members of one group become similar to one another and different from those of another group. This, in turn, creates a shared social identity for people as the members of a distinctive group. Given the pressure of ever-changing circumstances, social groups such as families, peer groups, business firms, and nations, only maintain their distinctive cultural beliefs and moderately stable social structures through the constant operation of conformity processes.

Perhaps because it is essential for social organization, conformity appears to be a universal human phenomenon. The level of conformity varies by culture, however. Collectivist cultures (e.g., Japan) that emphasize the interdependence of individuals show higher levels of conformity than individualistic cultures (e.g., the United States) that focus on the independence of individuals (Bond and Smith 1996).

Although essential, conformity always entails a conflict between a group standard and an alternative belief or behavior (Asch 1951; Moscovici 1985). For their physical and psychological survival, people need and want to belong to social groups. Yet to do so, they must curb the diversity and independence of their beliefs and behavior. Without even being aware of it, people usually willingly adopt the group position. Occasionally, however, individuals believe an alternative to be superior to the group standard and suffer painful conflict when pressured to conform.

Sometimes a nonconforming, deviate alternative is indeed superior to the group standard in that it offers a better response to the group circumstances. Innovation and change is as important to

a group's ability to adjust and survive as is conformity. In fact, a nonconforming member can influence the majority opinion even as the majority pressures the deviate to conform. As Irving Janis (1972) points out in his analysis of "groupthink," however, conformity pressures can grow so strong that they silence alternative opinions and strangle a group's ability to critically analyze and respond to the problems it faces. Thus, conformity is a double-edged sword. It enables people to unify for collective endeavors but it exacts a cost in potential innovation.

### CLASSIC EXPERIMENTS

The social scientific investigation of conformity began with the pioneering experiments of Muzafer Sherif (1936). They beautifully illustrate the easy, almost unconscious way people in groups influence one another to become similar. Sherif made use of the autokinetic effect, which is a visual illusion that makes a stationary pinpoint of light in a dark room appear to move. Sherif asked subjects in his experiments to estimate how far the light moved.

When individuals estimated the light alone, their estimates were often quite divergent. In one condition of the experiment, however, subjects viewed the light with two or three others, giving their estimates out loud, allowing them to hear each other's judgments. In this group setting, individuals gave initial estimates that were similar to one another and rapidly converged on a single group estimate. Different groups settled on very different estimates but all groups developed a consensus judgment that remained stable over time.

After three sessions together, group members were split up. When tested alone they continued to use their group standard to guide their personal estimates. This indicates that the group members had not merely induced one another to conform in outward behavior. They had influenced one another's very perception of the light so that they believed the group estimate to be the most accurate judgment of reality.

In another condition, Sherif first tested subjects alone so that they developed personal standards for their estimates. He then put together two or three people with widely divergent personal standards and tested them in a group setting. Over three group sessions, individual estimates merged

into a group standard. Thus, even when participants had well-established personal standards for judging, mere exposure to the differing judgments of others influenced them to gradually abandon their divergent points of view for a uniform group standard. This occurred despite a setting where the subjects, all strangers, had no power over one another and were only minimally organized as a group.

The Sherif experiment suggests that conformity pressures in groups are subtle and extremely powerful. But critics quickly noted that the extreme ambiguity of the autokinetic situation might be responsible for Sherif's results. In such an ambiguous situation, participants have little to base their personal judgments on, so perhaps it is not surprising that they turn to others to help them decide what to think. Do people conform when the task is clear and unambiguous? Will they yield to a group consensus if it is obvious that the consensus is wrong? These are the questions Solomon Asch (1951, 1956) addressed in his classic experiments.

To eliminate ambiguity, Asch employed clear-cut judgment tasks where subjects chose which of three comparison lines was the same length as a standard line. The correct answers were so obvious that individuals working alone reached 98 percent accuracy. Similar to the Sherif experiment, Asch's subjects gave their judgments in the presence of seven to nine of their peers (all participants were male college students). Unknown to the single naive subject in each group, all other group members were confederates of the experimenter. On seven of twelve trials, as the confederates announced their judgments one by one, they unanimously gave the wrong answer. It was arranged so that the naive subject always gave his judgment after the confederates.

The subject here was placed in a position of absolute conflict. Should he abide by what he knows to be true or go along with the unanimous opinion of others? A third of the time subjects violated the evidence of their own senses to agree with the group.

The Asch experiments clearly demonstrated that people feel pressure to conform to group standards even when they know the standards are wrong. It is striking that Asch, like Sherif, obtained these results with a minimal group situation. The

group members were strangers who meant little to one another. Yet they exerted substantial influence over one another simply by being in the same situation together. Because of the dramatic way it highlights the conflict inherent in conformity between individuals and groups, Asch's experimental design has become the paradigm for studying conformity.

### NORMATIVE AND INFORMATIONAL INFLUENCE

Sherif's and Asch's striking results stimulated an explosion of research to explain how conformity occurs (see Kiesler and Kiesler 1976, Cialdini and Trost 1998 for reviews). It is now clear that two analytically distinct influence processes are involved. Either or both can produce conformity in a given situation. Morton Deutsch and Harold Gerard (1955) labelled these *informational influence* and *normative influence*.

In informational influence, the group defines perceptual reality for the individual. Sherif's experiment is a good illustration of this. The best explanation derives from Leon Festinger's (1954) social comparison theory. According to the theory, people form judgments about ambiguous events by comparing their perceptions with those of similar others and constructing shared, socially validated definitions of the "reality" of the event. These consensual definitions constitute the social reality of the situation (Festinger 1950). Because people want the support of others to assure them of the validity of their beliefs, disagreeing with the majority is uncomfortable. People in such situations doubt their own judgment. They change to agree with the majority because they assume that the majority view is more likely to be accurate.

As this indicates, conformity as a result of informational influence is not unwilling compliance with the demands of others. Rather, the individual adopts the group standard as a matter of private belief as well as public behavior. Informational influence is especially powerful in regard to social beliefs, opinions, and situations since these are inherently ambiguous and socially constructed.

Normative influence occurs when people go along with the group majority in order to gain

rewards or avoid unpleasant costs. Thus it is normative influence that is behind compliance. People depend on others for many valued outcomes, such as inclusion in social relationships, a sense of shared identity, and social approval. Because of this dependency, even strangers have some power to reward and punish one another. Asch's results are a good example. Although a few of Asch's participants actually doubted their judgment (informational influence), most conformed in order to avoid the implicit rejection of being the odd person out. Studies show that fears of rejection for nonconformity are not unfounded (see Levine 1980 for a review). While nonconformists are sometimes admired, they are rarely liked. Furthermore, they are subject to intense persuasive pressure and criticism from the majority.

### FACTORS THAT INCREASE CONFORMITY

Anything that increases vulnerability to informational and normative influence increases conformity. Although there may be personality traits that incline people to conform, the evidence for this is conflicting (Crowne and Marlowe 1969; Moscovici 1985). Situational factors seem to be the most important determinants of conformity. Research indicates that conformity is increased by a) the ambiguity or difficulty of the task, b) the relative unimportance of the issue to the person, c) the necessity of making a public rather than private response, d) the similarity of group members, e) high interdependence among the group members, f) the attractiveness and cohesiveness of the group, and g) the unanimity of the majority (see Kiesler and Kiesler 1976; Cialdini and Trost 1998 for reviews).

When a task or situation is ambiguous or difficult, it is not easy to tell what the best response to it would be. As a result, much as in Sherif's experiments, group members rely heavily on each other's opinions to decide what is best, increasing their susceptibility to informational influence. When decision-making groups in government or business face complex, difficult decisions where the right choice is uncertain, informational influence increases the members' tendency to agree and can affect their critical analysis of the situation (Janis 1972). Tastes and beliefs about matters, such as clothing style or music, about which there are no objectively right choices are subject to

sudden fads or fashions for similar reasons. Powerful conformity processes take over as group standards define for the individual what the “right” clothes or music are.

The less people care about an issue, the more open they are to both informational and normative influence. Without the motivation to examine an issue personally, people usually accept the group standard about it, both because the agreement of others makes the standard seem right and because there are more rewards and fewer costs in going along with the group. Because of such rewards and costs, people are especially likely to go along when their response must be public rather than private.

Since people compare their perceptions and views most closely to those of people who are socially similar to them, similarity increases group members’ informational influence on one another. Similarity also increases liking and, when people like one another, they have more power to reward or punish each other, so normative influence increases as well. Because of the increased power of both informational and normative influence, conformity pressures are often especially strong in peer groups.

When members are highly dependent on one another for something they value, conformity pressure increases because the members have more power to reward or frustrate one another (normative influence). Similarly, when a group is very attractive to an individual, its members have more power to normatively influence the individual. Gangs, fraternities, and professional societies all use this principle to induce new members to adopt their groups’ distinctive standards. Also, when a group is very tight knit and cohesive, members’ commitment to the group gives it more power over their behavior, increasing the forces of conformity.

The unanimity of the majority in a group is an especially important factor in the conformity process. In his studies, Asch (1951) found that as long as it was unanimous, a majority of three was as effective in inducing conformity as one of sixteen. Subsequent research generally confirms that the size of a majority past three is not a crucial factor in conformity. It is unanimity that counts (see Allen 1975 for a review). When Asch (1951) had one confederate give the correct answer to the line

task, the naive subjects’ conformity to the majority dropped from a third to only 5 percent. One fellow dissenter shows an individual that nonconformity is possible and provides much needed social support for an alternate construction of social reality. Interestingly, a dissenter need not agree with an individual to encourage nonconformity. It is only necessary that the dissenter also break with the majority.

Another factor that affects conformity is the sex composition of the group. Although the results of studies are inconsistent, statistical summaries of them, called meta-analyses, indicate that there is an overall tendency for women to conform slightly more than men (Becker 1986; Eagly and Wood 1985). Sex differences in conformity are most likely when behavior is under the surveillance of others. The evidence suggests two explanations (see Eagly 1987 for a review). First, sex carries status value in interaction, which creates social expectations for women to be less competent and influential in the situation than men (Ridgeway 1993). Second, sex stereotypes pressure men to display independence when they are being observed.

#### THE INFLUENCE OF THE MINORITY ON THE MAJORITY

Conformity arises out of a social influence process between an individual and the group majority. The influence process is not always one way, however. As Serge Moscovici (1976) points out, a dissenting group member is not just a recipient of pressure from the majority, but also someone who, by breaking consensus, challenges the validity of the majority view, creating conflict, doubt, and the possibility of opinion change in the group. Dissenters sometimes modify the opinion of the majority in a process called *minority influence*. Research shows that for a minority opinion to affect the majority it must be presented consistently and clearly without wavering and it helps if there are two such dissenters in the group rather than one (see Moscovici 1985; Moscovici, Mucchi-Faina, and Maass 1994; Wood et al. 1994, for reviews). A dissenting minority increases divergent thinking among group members that can enhance the likelihood that they will arrive at creative solutions to the problems the group faces (Nemeth 1986).

### CONFORMITY AND STATUS

Research in the Asch and Sherif paradigms focuses on conformity pressures among peers. However, when group members differ in status, it affects the group's tolerance of their nonconformity. Higher status members receive fewer sanctions for nonconformity than lower status members (Gerson 1975). As long as they adhere to central group norms, high status members' nonconformity can actually increase their influence in the group (Berkowitz and Macauley 1961). Edwin Hollander (1958) argues that, because high status members are valued by the group, they are accorded "idiosyncrasy credits" that allow them to nonconform and innovate without penalty as long as they stay within certain bounds. It is middle status members who actually conform the most (Harvey and Consalvi 1960). They have fewer idiosyncrasy credits than high status members and more investment in the group than low status members.

Nonconformity can also affect the position of status and influence a person achieves in the group. Hollander (1958, 1960) proposed that individuals earn status and idiosyncrasy credits by initially conforming to group norms, but replications of his study do not support this conclusion (see Ridgeway 1981 for a review). Conformity tends to make a person "invisible" in a group and so does little to gain status. Nonconformity attracts attention and gives the appearance of confidence and competence which can enhance status. But it also appears self-interested, which detracts from status (Ridgeway 1981). Consequently, moderate levels of nonconformity are most likely to facilitate status attainment.

### COMPLIANCE WITH AUTHORITY

Reacting to the Nazi phenomenon of World War II, studies of compliance to authority have focused on explaining people's obedience even when ordered to engage in extreme or immoral behavior. Compliance in this situation is comparable to conformity in the Asch paradigm in that individuals must go against their own standards of conduct to obey. The power of a legitimate authority to compel obedience was dramatically demonstrated in the Milgram (1963, 1974) experiments. As part of an apparent learning study, a scientist-experimenter ordered subjects to give increasingly strong

electric shocks to another person. The shock generator used by the subject labelled increasing levels as "danger-severe shock" and "XXX" (at 450 volts). The victim (a confederate who received no actual shocks) protested, cried out, and complained of heart trouble. Despite this, 65 percent of the subjects complied with the scientist-experimenter and shocked the victim all the way to the 450 volt maximum. It is clear that most of the time, people do as they are told by legitimate authorities.

Uncertainty over their responsibilities in the situation (a definition of social reality issue) and concern for the authority's ability to punish or reward them seem to be the principle causes for people's compliance in such circumstances. Note the comparability of these factors to informational and normative influence. Situational factors that socially define the responsibility question as a duty to obey rather than to disobey increase compliance (Kelman and Hamilton 1989), as do factors that increase the authority's ability to sanction.

Research has demonstrated several such factors. Compliance is increased by the legitimacy of the authority figure and his or her surveillance of the individual's behavior (Milgram 1974; Zelditch and Walker 1984). When others in the situation obey or when the individual's position in the chain of command removes direct contact with the victim, compliance increases (Milgram 1974). On the other hand, when others present resist the authority, compliance drops dramatically. Stanley Milgram (1974) found that when two confederates working with the subject refused to obey the experimenter, only 10 percent of subjects complied fully themselves. As with a fellow dissenter from a unanimous majority, other resisters define disobedience as appropriate and provide support for resisting. In an analysis of "crimes of obedience" in many government and military settings, Herbert Kelman and Lee Hamilton (1989) show how such factors lead to compliance to illegal or immoral commands from authority.

Conformity and compliance are fundamental to the development of norms, social organization, group culture, and people's shared social identities. As a result, research on conformity and compliance continues to develop in several directions. Efforts are underway to develop broader models of the social influence process that can incorporate both conformity and compliance (see Cialdini



and Trost 1998). These efforts give added emphasis to people's dependence on social relationships and groups and address questions such as whether minority and majority influence work through different or similar processes. Also, new, more systematic cross-cultural research attempts to understand both what is universal and what is culturally variable about conformity and compliance (Markus and Kitayama 1994; Smith and Bond 1996).

## REFERENCES

- Allen, Vernon L. 1975 "Social Support for Nonconformity." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 8, 2-43. New York: Academic Press.
- Asch, Solomon E. 1951 "Effects of Group Pressure Upon the Modification and Distortion of Judgments." In H. Guetzkow, ed., *Groups, Leadership, and Men*. Pittsburgh: Carnegie Press.
- 1956 "Studies of Independence and Submission to Group Pressure: I. A Minority of One Against a Unanimous Majority." *Psychological Monographs* 70 (9, whole no. 416).
- Becker, B. J. 1986 "Influence Again: Another Look at Studies of Gender Differences in Social Influence." In Janet Shibley Hyde and Marci C. Lynn, eds., *The Psychology of Gender: Advances Through Meta-Analysis*. Baltimore: Johns Hopkins University Press.
- Berkowitz, Leonard, and J. R. Macaulay 1961 "Some Effects of Differences in Status Level and Status Stability." *Human Relations* 14:135-147.
- Bond, M. H. and P. B. Smith 1996 "Culture and Conformity: A Meta-Analysis of Studies Using Asch's Line Judgment Task." *Psychological Bulletin* 119:111-137.
- Cialdini, Robert B., and Melanie R. Trost 1998 "Social Influence: Social Norms, Conformity, and Compliance." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, 4th ed., vol. 2. New York: McGraw-Hill.
- Crowne, D. P., and D. Marlowe 1964 *The Approval Motive: Studies in Evaluative Dependence*. New York: Wiley.
- Eagly, Alice H. 1987 *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, N.J.: Earlbaum.
- , and Wendy Wood 1985 "Gender and Influenceability: Stereotype vs. Behavior." In V. E. O'Leary, R. K. Unger, and B. S. Watson, eds., *Women, Gender and Social Psychology*. Hillsdale, N.J.: Earlbaum.
- Festinger, Leon 1950 "Informal Social Communication." *Psychological Review* 57:217-282.
- 1954 "A Theory of Social Comparison Processes." *Human Relations* 7:117-140.
- Gerson, Lowell W. 1975 "Punishment and Position: The Sanctioning of Deviants in Small Groups." In P. V. Crosbie, ed., *Interaction in Small Groups*. New York: Macmillan.
- Harvey, O. J., and Conrad Consalvi 1960 "Status and Conformity to Pressures in Informal Groups." *Journal of Abnormal and Social Psychology* 60:182-187.
- Hollander, Edwin P. 1958 "Conformity, Status, and Idiosyncrasy Credit." *Psychological Review* 65:117-127.
- 1960 "Competence and Conformity in the Acceptance of Influence." *Journal of Abnormal and Social Psychology* 61:365-369.
- Janis, Irving L. 1972 *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*. Boston: Houghton-Mifflin.
- Kelman, Herbert C. 1958 "Compliance, Identification, and Internalization: Three Processes of Attitude Change." *Journal of Conflict Resolution* 2:51-60.
- , and V. Lee Hamilton 1989 *Crimes of Obedience*. New Haven, Conn.: Yale University Press.
- Kiesler, Charles A. and Sara B. Kiesler 1976 *Conformity*. 2d ed. Reading, Mass.: Addison-Wesley.
- Levine, John M. 1980 "Reaction to Opinion Deviance in Small Groups." In P. Paulus, ed., *Psychology of Group Influence*. Hillsdale, N.J.: Earlbaum.
- Markus, Hazel R., and S. Kitayama 1994 "A Collective Fear of the Collective: Implications for Selves and Theories of Selves." *Personality and Social Psychology Bulletin* 20:568-579.
- Milgram, Stanley 1963 "Behavioral Study of Obedience." *Journal of Abnormal and Social Psychology* 67:371-378.
- 1974 *Obedience to Authority: An Experimental View*. New York: Harper and Row.
- Moscovici, Serge 1976 *Social Influence and Social Change*. London: Academic Press.
- 1985 "Social Influence and Conformity." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*. 3d ed., vol. 2. New York: Random House.
- , A. Mucchi-Faina, and A. Maass 1994 *Minority Influence*. Chicago: Nelson-Hall.
- Nemeth, Charlan 1986 "Differential Contributions of Majority and Minority Influence." *Psychological Review* 93:23-32.
- Ridgeway, Cecilia L. 1981 "Nonconformity, Competence, and Influence in Groups: A Test of Two Theories." *American Sociological Review* 46:333-347.
- 1993 "Gender, Status, and the Social Psychology of Expectations." In P. England, ed., *Theory on Gender/Feminism on Theory*. New York: Aldine.

Sherif, Muzafer 1936 *The Psychology of Social Norms*. New York: Harper and Row.

Wood, Wendy, S. Lundgren, J. A. Ouelette, S. Busceme, and T. Blackstone 1994 "Minority Influence: A Meta-Analytic Review of Social Influence Processes." *Psychological Bulletin* 115:323-345.

Zelditch, Morris, Jr., and Henry A. Walker 1984 "Legitimacy and the Stability of Authority." In E. Lawler, ed., *Advances in Group Processes*, vol. 1. Greenwich, Conn.: JAI Press.

CECILIA L. RIDGEWAY

## COMPLEMENTARITY

See Mate Selection Theories.

## COMPUTER APPLICATIONS IN SOCIOLOGY

Most sociologists, both professionals and students, now have their own computer with direct access to a printer for writing and to the Internet for electronic mail (e-mail). Beyond the basic tasks of writing and e-mailing are a variety of other computer-supported research applications, both quantitative and qualitative. This article describes how sociologists and other social scientists use these applications and what resources are available.

The data and modeling requirements of social research have united sociologists with computers for over a hundred years. It was the 1890 U. S. census that inspired Herman Hollerith, a census researcher, to construct the first automated data processing machinery. Hollerith's punchcard system, while not a true computer by today's definitions, provided the foundation for contemporary computer-based data management.

In 1948 the U. S. Bureau of the Census, anticipating the voluminous tabulating requirements of the 1950 census, contracted for the building of Univac I, the first commercially produced electronic computer. The need to count, sort, and analyze the 1950 census data on this milestone computer led to the development of the first high-speed magnetic tape storage system, the first sort-merge software package, and the first statistical package, a set of matrix algebra programs.

Only a decade later many social scientists were exploring ways to use computers in their research. In the early 1960s, the first book devoted entirely to computer applications in social science research (Borko 1962) was published. Not only were social scientists writing about how to apply computers, they were designing and developing new software. Some of the most popular statistical software packages, e.g., SPSS (Nie, Bent, and Hull 1975), were developed by social scientists.

During the 1980s, universities and colleges began to acquire microcomputers, accepting the premise that all researchers needed their own desktop computing equipment. The American Council of Learned Societies (ACLS) survey in 1985 (Morton and Price 1986) reported that 50 percent of sociologists had a computer for exclusive use. A survey of academic departments supported by the American Sociological Association (Koppel, Dowdall, and Shostak 1985) found that slightly less than half of the sociology faculty reported to have immediate access to microcomputers. To put these findings into a more complete perspective, of the approximately 9,000 sociologists in 1985, about 4,500 had their own computers and about 5,200 reported routine computer use. Now, it is hard to find a sociologist's office without at least one computer. And in many countries most students in sociology have a computer for writing papers and accessing online resources.

**Sociology and the Web.** The Internet may be one of the largest and probably the most rapidly growing peaceful social movements in history. It is not just a technology, or a family of technologies, but a rapidly evolving socio-cultural phenomena often called "cyberspace" or "cyberculture." No matter how this phenomena is defined, it is changing the way sociologists conduct their work.

By the mid-1990s sociology, like most other academic disciplines, had come to depend upon e-mail. In addition, a rapidly growing number had begun to use the World Wide Web (WWW), commonly called the Web (Babbie 1996). Bainbridge (1995) claimed that the Web is "a significant medium of communication for sociologists, and extrapolation of present trends suggests it may swiftly become the essential fabric of sociology's existence." In January 1999, the author searched Web sites with the Alta Vista search engine for the word

“sociology” and found 750,000 instances. Two years earlier a search yielded only 250,000. Sociology is indeed rapidly building its presence on the Web.

The Internet and the Web are often described as a medium of communication because e-mail, electronic conferencing, online chats (synchronous discussion), groupware, and data exchange imply social interaction. Another major metaphor of the Web is that of a database for information search and retrieval. However, new roles are emerging for the Web. For the individual, the Web has become an important “presentation of self,” an opportunity for publishing personal and professional resumes and other such information. For the organization, the Web has become an opportunity for advertising, recruiting, communicating with the public, and conducting commerce itself. The rapid proliferation of personal Web sites in the form of various home pages suggests that many now see the Web primarily as a medium for personal and organizational impression management.

An appendix to this article provides a number of sources of sociological activity and information on the Internet, especially the Web. Not only does the appendix serve as a resource for information about sociology, but it also portrays the diverse, complex state of the institutionalization of sociology on the Web.

In the next few years the Web and its technologies will offer many new opportunities for conducting research. Already there are software packages developed for Web-assisted surveys and other types of data collection. The amount of data accumulating on the Web is enormous. And the graphical content of the Web is indeed a challenge to sociologists conducting computerized content analyses. The pace of these new developments will continue to challenge sociologists to not only to stay current with the new tools for research but to conduct research on the rapidly growing cultures of the Internet.

**Publications on Sociological Computing.** The primary source for articles on social science computer applications is the *Social Science Computer Review*, a quarterly publication of Sage Publications, Inc., which also regularly offers book and software reviews. Other relevant software reviews periodically appear in such journals as *Educational*

*and Psychological Measurement*, the *Journal of Marketing Research*, *The American Statistician*, and *Simulation and Games*. In addition, JAI Press publishes an occasional series volume on “Computers and the Social Sciences.”

In the late 1980s, the American Sociological Association (ASA) formed a “Section on Microcomputing.” Over 350 sociologists joined this new section in 1990, making it the fastest-growing new section in the history of the ASA. The section publishes a quarterly newsletter and organizes sessions at annual meetings. In 1993, the section name was officially changed to “Sociology and Computers.” A regular newsletter called *SCAN (Sociology and Computers: A Newsletter)* is published by members.

Another indicator of growing involvement in computing was the first annual conference, “Computing in the Social Sciences,” held in 1990 at Williamsburg, Virginia. From this conference emerged a professional association, the Social Science Computing Association. The official journal of this association is the *Social Science Computer Review*, and they still hold an annual conference.

**Computer Applications in Sociology.** The practice of computing in sociology has evolved rapidly. Computers have been applied to practically every research task, including such unlikely ones as field note-taking, interviewing, and hundreds of other tasks (Brent and Anderson 1990). The many diverse uses of computing technology in social research are difficult to categorize because applications overlap and evolve unpredictably. Nonetheless, it is necessary to discuss different categories of applications in order to describe the state of the art of computing in sociology. Since 1987, the Winter issue of the *Social Science Computer Review* has been devoted to an annual symposium on the “State of the Art of Social Science Computing.” The categorization of computer applications in this article reflects these discussions. First, some major types of applications are summarized in order of descending popularity. Then some of the challenges of computing for sociological research are noted.

**Writing and Publishing.** Once equated with the secretarial pool, word processing now is an

activity of nearly every graduate student and professional in sociology. It consists not only of writing but preparing tables, "typesetting" mathematical equations, and resizing objects, such as three-dimensional graphs embedded within text. Social researchers are using such capabilities and moving rapidly toward workstation environments that obscure the transition between data analysis and manuscript preparation (Steiger and Fouladi 1990). Not only do researchers use their computers for writing papers, but word processing software plays a central role in the refinement of data collection instruments, especially questionnaires and codebooks, which allows for rapid production of alternative forms and multiple drafts.

Trends in text production that blur traditional distinctions between writing and publishing (Lyman 1989) may in the long term have the most impact on what sociologists do. The growing body of articles and books in electronic-text form propel scholarship toward *hypertext*, which is a document system that provides for nonsequential reading of text using links that automatically access other documents. Contemporary word processors contain the capacity to easily produce documents in HTML (Hypertext Markup Language) that are ready for installation as sites on the Web. HTML can contain hypertext links, which with a single click of the mouse can bring up a totally different document from anywhere in the world, making it a truly new form of publishing.

There are several major forms of text entry that may also change the nature of writing and publishing. These forms include scanning for optical character recognition (OCR), voice recognition, and automated language translation. The technology for scanning text documents and producing computer text files has been in use for some years and requires only a scanning device and OCR software. This technology will continue to improve and its use will yield fewer errors and require considerably less effort in the future.

Likewise it is now possible to use voice recognition software to automatically transcribe dictation, interviews, and field notes into computer text files. One of the remaining problems in this approach is that all voice recognition software now requires considerable "training" time where the errors made in recognizing a speaker's word-sound pattern are corrected. The software is thus "taught"

to make refined guesses in translating vocalized sounds into electronic text. Even the best voice recognition software now makes a moderate number of errors, so it is not yet a panacea for manual transcription of either the spoken word or audio recordings.

It is possible now to find software that will translate text into many different languages. However, like voice recognition software, translation software still requires considerable time to manually check, decipher, and make judgments about the text produced by such programs. Future software may yield significantly improved results, automating nearly all of the voice recognition and translation process.

#### **Communicating Electronically (E-mail, etc.)**

Networks for computer-mediated communication (CMC) continue to expand internationally following the traditional logistic diffusion curve (Gurbaxani 1990). Electronic networks now supplement most other forms of social communication. E-mail, which is asynchronous or nonsimultaneous, is still the most common form of electronic interaction, but Internet-based, synchronous (simultaneous) "mailing lists" and "newsgroups" are also popular, as are "chat rooms" or "discussion groups." With improvements in transmitting digital audio and video files on the Internet, it is expected that some new forms of video conferencing will become commonplace. At the turn of the millennium, desktop video conferencing is available "off the shelf," but suffers from extraneous noise and rough motion. While individual sociologists vary greatly in how they utilize e-mail, nearly all sociologists in economically developed countries depend upon it for certain types of communication.

While e-mail messages are generally written in plain text, "attachments" to e-mail now make it possible for formatted documents, even those including graphics and multimedia, to be shared with others around the world in a matter of minutes. This remarkable technology makes co-authoring, and other forms of collaboration, far more feasible due to reduced time and cost.

As e-mail systems continue to expand, they offer social researchers new opportunities for conducting studies using electronic networks. For instance, Gaiser (1997) explored issues of running online focus groups. Online surveys have become

quite common in various forms: e-mail texts, e-mail attachments, entry forms on the Web, and as programs in external storage devices like diskettes and CD-ROMs. Sampling problems and low completion rates pose the greatest challenges. Ongoing methodological investigations will be necessary to determine the implications of this new mode of research.

**Statistics.** Hundreds of computer programs and articles have been written to address the needs of statistical computing in social research. Prior to the 1980s, all statistical work was performed on large or medium-size, mainframe computers. But advances in both hardware and software for microcomputers now make it possible to conduct the statistical data analysis of most small or moderate-size research studies on microcomputers. A large share of ongoing social data analysis, like analysis of massive census files, would never get done without computer technology. For example, one use of LISREL, a computer procedure which analyzes linear structural relationships by the method of maximum likelihood, would consume weeks or months without a computer.

Not only does statistical computing save time but it offers unique views of the patterns in one's data. Without the ability to quickly reorganize data and display it in a variety of forms, social researchers neglect important patterns and subtle relationships within complex data. Some patterns cannot be observed without special software tools. For example, Heise's (1988) computer program called Ethno gives the researcher a framework for conceptualizing, examining, and analyzing data containing event sequences. In addition, several general-purpose statistical packages offer powerful exploratory data analysis capabilities with bidirectionality through dynamic data links (Steiger and Fouladi 1990). One type of bidirectionality puts a graph in one window and frequency distributions in another, and when the user adjusts the data in one window, it automatically changes in the other.

Finding a statistical program tailored to a particular problem or technique is often challenging as the potential "user community" may be quite small. The best sources for such software are the notices and reviews in journals such as the *Social Science Computer Review*, *Educational and Psychological Measurement*, *Journal of Marketing Research*,

and *The American Statistician*. Another important source is the annual *Sociological Methodology* and its software list on the Web (<http://weber.u.washington.edu/~socmeth2/software.html>). Some of the software noted in these sources can be obtained at no cost or very low cost. No matter what the cost, one cannot assume that any complex program is free of errors. Thus is it important to run test data and to apply data to multiple programs in order to check for inaccuracies.

#### **Accessing, Retrieving and Managing Data.**

While years ago students and researchers had to use a library or similar institution to gain access to bibliographic data files, now such services are available from one's desktop using the Web or external storage units such as CD-ROM or DVD-ROM. Large bibliographic databases including Sociological Abstracts and Psychological Abstracts are available in these forms, as is a vast amount of data in the form of statistical tables and maps. Now that devices for "writing" onto CD-ROMs have become inexpensive, it is expected that data from even small research projects will be disseminated in this medium.

One major development is interactive access to data by means of the Web. A variety of models are used for interactive access to both preformatted text files and precoded data files. Among the systems are GSSDIRS from ICPSR at the University of Michigan (<http://www.ICPSR.umich.edu/gss>), IPUMS at the University of Minnesota (<http://www.hist.umn.edu/~ipums>), QSERVE from Queens College-CUNY (<http://www.soc.qc.edu/qserve>), and SDA Archive from the University of California at Berkeley (<http://csa.berkeley.edu:7502/archive.htm>).

The software technology for archiving and analyzing social data is less than a half-century old. But it is very plausible to expect many advancements in the next fifty years. Interactive data analysis sites of the Web hint about what these advancements might be. For instance, using the SDA Archive Web site (see [http](http://csa.berkeley.edu:7502/archive.htm) address above), one can get a large crosstabulation table for any three variables in the full General Social Survey of over 35,000 respondents in less time than it takes to type in the variable names.

The refinement of such systems faces issues such as how to balance functionality with ease of use, the plausibility of standardizing interfaces for

many numeric data files, and the use of gateways between the Web and third-party software such as statistical packages. It will take considerable research and development to sort out the feasibility of providing many different analytical facilities in the Web environment. A major challenge is determining the amount and type of data documentation necessary for typical users to get meaningful results in a reasonable amount of time. A considerable challenge is created when a variety of types of materials are necessary for retrieving useful data-related information. The main types are text, graphics, meta databases (data about databases), fielded text/data (such as bibliographic databases), and multimedia (audio and video) clips.

**Qualitative Computing.** Computer-based content analysis began with Stone (1966) and associates, and now plays an important role in the social sciences (Weber 1984; Kelle 1995). A survey of 110 qualitative-oriented researchers found three-fourths regularly used computers (Brent, Scott, and Spencer 1987). Ragin and Becker (1989, p. 54) persuasively claimed that qualitative research using these computing tools yields more systematic attention to diversity, for example, by encouraging a "more thorough examination of comparative contrasts among cases."

This type of computing became much more common as researchers combined content analysis with other tasks associated with qualitative analysis. Several general-purpose programs for qualitative analysis have been widely distributed (Tesch 1989; Fielding and Lee 1991). These tools make the analysis of large amounts of text more accurate and efficient, and potentially direct the focus of attention to analytic procedures. The general tasks of text entry, code assignment, counting, and data organization have been extended to include special routines for improving the quality of coding and code management (Carley 1988). Hesse-Biber, Dupuis, and Kinder (1997) technically extended this methodology to include the management and analysis of audio and video segments as well as text.

**Simulating and Modeling.** Early in the history of sociological computing, Coleman (1962) and McPhee and Glaser (1962) designed computer simulation models and showed how they could be used to identify elusive implications of different theoretical assumptions. Other social scientists

followed in their footsteps but the excitement of the pioneers was lost and few simulations and formal computer models were developed in the 1970s. With the emergence of artificial intelligence and other modeling methodologies, social researchers demonstrated renewed interest in formal computer-supported models of social processes (cf. Feinberg and Johnson 1995; Hanneman 1988; Markovsky, Lovaglia, and Thye 1997). New computer simulations for social policy analysis as well as pedagogy or instruction have emerged as well (Brent and Anderson 1990, pp. 188–210).

Neural networks combined with other techniques of artificial intelligence and expert systems have excited a number of social scientists (Garson 1990). Neural nets organize computer memory in ways that model human brain cells and their ability to process many things in parallel. Systems that use neural nets are especially good when pattern matching is required, however, the computations require high-performance computers.

**Computer-Assisted Data Collection.** CATI (Computer-Assisted Telephone Interviewing) is a computing system with online questionnaires or entry screens for telephone interviewers. It has become very common in sociological research, although its impact is not fully understood (Groves et al. 1988). It is used on free-standing PCs, networked PCs, or larger computers. These systems generally, but not always, have the following characteristics: centralized facilities for monitoring individual interviewer stations, instantaneous edit-checks with feedback for invalid responses, and automatic branching to different questions depending upon the respondents' answers. Other major forms of computer-supported data collection include (1) CAPI (Computer-Assisted Personal Interviewing), the acronym used in survey research to refer to face-to-face interviewing assisted with a laptop or hand-held computing device; (2) Computerized Self-Administered Questionnaires (CSAQ), online programs designed for direct input from respondents; and (3) data-entry programs to facilitate the entry of data collected manually at a prior time.

A related innovation is software built for designing online questionnaires. For example, the Questionnaire Programming Language (QPL), developed by Dooley (1989), allows the researcher to

draft a questionnaire with any word processor. From embedded branching commands within the questionnaire document, the QPL software package simultaneously produces two versions of the questionnaire: one for computer administration and the other for interviewer- or self-administration. An additional bonus of the program is that it automatically produces data definition commands for SPSS or SAS to use for data analysis.

**Visualization and Graphics.** Many social researchers have come to rely on computer graphic systems to produce maps, charts summarizing statistical data, network diagrams, and to retrieve data from GIS (Geographic Information Systems) databases. GIS data contain coordinates to associate a specific spatial point or area with any attribute, e.g., social density, associated with that point or collection of points. Because of the complexities of these data structures, the integration of these techniques onto sociologists' desktops has been slow. Another constraint is the paucity of techniques for analyzing such data. Work such as that of Cleveland (1993) for analyzing and visualizing data graphically may increase the utilization of such data by sociologists. In addition, techniques for storing and delivering interactive audio and video by means of the Web may stimulate sociologists to investigate multimedia data containing sounds and moving images.

**Teaching and Learning.** During the 1970s long before the microcomputer, a small group of social science instructors began to explore how to utilize computer technology in teaching (Bailey 1978). Now it has become quite common, and several sociologically-oriented instructional packages are widely used. The most popular have been Chipendale, designed by James Davis (1990), and MicroCase, developed by Roberts and Stark (cf. Roberts and Corbett 1996). These two packages have served as the basis for the exercises contained in at least a dozen published textbooks and workbooks. A variety of instructional approaches and software tools are described regularly in *Teaching Sociology* and the *Social Science Computer Review*.

### ISSUES AND CHALLENGES

The practice of computing in social research has evolved rapidly. Computers have been applied to

practically every research task, including such unlikely ones as interviewing, and hundreds of other tasks (Brent and Anderson 1990). This variety of computer applications will continue to evolve with the newer Internet-based technologies.

The application of computing to sociology is not without problems. Errors in data and software abound yet rarely do social scientists check their results by running more than one program on the same data. Data and software tend to be very costly, but there are many impediments to the sharing of these critical resources. Better software is needed but graduate students often are discouraged from programming new software. Nonetheless, new breakthroughs in computer technology will continue, and major new opportunities will emerge. Many of the advances in sociological computing during the next few years undoubtedly will follow the lines of progress already described: hypertext networks; integrated, high performance, graphic data analysis stations; software for computer-supported cooperative work; and neural networks for complex models of social systems.

Perhaps the most exciting challenge for the future involves a concert of these innovations directed at the problem of modeling and analyzing vast amounts of social data. One solution would incorporate three-dimensional, multicolored, dynamic graphical representations of complex social data structures. But new techniques for analyzing these data will require new models of dynamic social structures as well as parallel social processes. Computer representations of these models tend to require extremely fast processing. Access to such models on the Web, supplemented with audio and video displays, may evolve into an important part of the sociologist's tool kit of the future.

### REFERENCES

- Anderson, Ronald E. 1981 "Instructional computing in sociology: current status and future prospects." *Teaching Sociology*. 8,2 (Jan):171-195.
- 1997 "A Research Agenda for Computing and the Social Sciences." *Social Science Computer Review* 14, 1 (Summer):123-134.
- , and Edward Brent 1989 "Computing in Sociology: Promise and Practice." *Social Science Computer Review* 7,4 (Winter).
- 1991 "Sociological Computing: An Opportunity Missed?" *The American Sociologist* 22,11 (Spring):65-77.

- Babbie, Earl 1996 "We Am a Virtual Community." *The American Sociologist* 27,1 (Spring):65-71.
- Bailey, D. (ed.) 1978 *Computer Science in Social and Behavioral Science Education*. Englewood Cliffs, N.J.: Educational Technology Press.
- Bainbridge, William Sims 1995 "Sociology on the World Wide Web." *Social Science Computer Review* 13,4 (Winter):508-523.
- Borko, Harold (ed.) 1962 *Computer Applications in the Behavioral Sciences*. Santa Monica, Calif.: System Development Corporation.
- Brent, Edward, and Ronald E. Anderson 1990 *Computer Applications in the Social Sciences*. New York: McGraw-Hill.
- Brent, Edward, James Scott, and John Spencer 1987 "The use of computers by qualitative researchers." *Qualitative Sociology* 10(3):309-313.
- Carley, Kathleen 1988 "Formalizing the Social Expert's Knowledge." *Sociological Methods & Research* 17, 2:165-232.
- Cleveland, William S. 1993 *Visualizing Data*. Summit, N.J.: Hobart Press.
- Coleman, J. S. 1962 "Analysis of social structures and simulation of social processes with electronic computers." In H. Guetzkow, H., ed., *Simulation in Social Science: Readings*, 61-69 Englewood Cliffs, N.J.: Prentice-Hall.
- Davis, James 1990 CHIPendale. Zeda Data, 25 Haskins Rd. Hanover, NH.
- Dooley, Kevin 1989 QPL Data Collection Program. Washington D.C.: Human Resources Division, U. S. General Accounting Office.
- Fan, David 1996 Use of the Press to Predict Time Trends in Public Opinions and Behaviors. Paper presented at the 1996 Conference on Computing for the Social Sciences at the University of Minnesota.
- Feinberg, William E., and Norris R. Johnson 1995 "FIRESCAP: A Computer Simulation Model of Reaction to a Fire Alarm." *Journal of Mathematical Sociology* 20, 2-3:257-269.
- Fielding, Nigel G., and Raymond M. Lee (eds.) 1991 *Using Computers in Qualitative Research*. Newbury Park, Calif.: Sage Publications, Inc.
- Gaiser, Ted J. 1997 "Conducting On-Line Focus Groups" *Social Science Computer Review* 15,2 (Summer):135-144.
- Garson, G. David 1990 "Expert Systems: An Overview for Social Scientists." *Social Science Computer Review* 8,3 (Fall):387-410.
- Groves, Robert M., P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, and J. Waksberg (eds.) 1988 *Telephone Survey Methodology*. New York: John Wiley and Sons.
- Gurbaxani, Vijay 1990 "Diffusion in Computing Networks: BITNET." *Communications of the ACM* 33,12 (Dec):65-75.
- Hanneman, Robert A. 1988 *Computer-Assisted Theory Building*. Newbury Park, Calif.: Sage Publications, Inc.
- Heise, David 1988 "Computer analysis of cultural structures." *Social Science Computer Review* 6:183-196.
- Hesse-Biber, Sharlene, Paul Dupuis, and Scott Kinder 1997 "Anthropology: New Developments in Video Ethnography and Visual Sociology—Analyzing Multimedia Data Qualitatively." *Social Science Computer Review* 15, 1 (Spring):5-12.
- Kelle, Udo 1995 *Computer-Aided Qualitative Data Analysis*. Newbury Park, Calif.: Sage Publications, Inc.
- Kiesler, Sarah, and Lee S. Sproull 1986 "Response effects in the electronic survey." *Public Opinion Quarterly* 50:402-413.
- Koppel, R, G. Dowdall, and A. Shostak 1985 "Using computers to teach sociology: A departmental survey." *Footnotes* 13,8 (November):4.
- Lyman, Peter 1989 "The Future of Sociological Literature in an Age of Computerized Texts." In G. Blank, J. L. McCartney, and E. Brent, eds., *New Technology in Sociology: Practical Applications in Research and Work*, 17-32 New Brunswick, N.J.:
- Markovsky, Barry, Michael Lovaglia, and Shane Thye 1997 "Sociology: Computer-Aided Research at the Iowa Center for the Study of Group Processes." *Science Computer Review* 15, 1 (Spring):48-64.
- McPhee, W. N., and W. A. Glaser (eds.) 1962 *Public Opinion and Congressional Election*. New York: Free Press.
- Morton, H. C., and A. J. Price 1986 "The ACLS survey of scholars: views on publications, computers, libraries." *EDUCOM Bulletin* 21,3 (Winter):8-21.
- Nie, N., D. Bent, and W. Hull 1975 *Statistical Package for the Social Sciences*. New York: McGraw-Hill.
- Ragin, Charles C. and Howard S. Becker 1989 "How the Microcomputer is Changing our Analytic Habits." In G. Blank, J. L. McCartney, and E. Brent, eds., *New Technology in Sociology: Practical Applications in Research and Work*. New Brunswick, N.J.: Transaction.
- Roberts, Lynne and Michael Corbett 1996 *Social Research Using MicroCase*. Bellevue, Wash.: MicroCase Corporation.
- Steiger, James H., and Rachel T. Fouladi 1990 "Some Key Emerging Trends in Statistical and Graphical Software for the Social Scientist." *Social Science Computer Review* 8,4 (Winter):627-664.



Stone, Philip et al. 1966 *The General Inquirer*. Cambridge, Mass.: MIT Press.

Tesch, Renata 1989 "Computer Software and Qualitative Analysis: A Reassessment." In G. Blank, J. L. McCartney, and E. Brent, eds., *New Technology in Sociology: Practical Applications in Research and Work*. New Brunswick, N.J.: Transaction.

Weber, R. P. 1984 "Computer-Aided Content Analysis, A Short Primer." *Qualitative Sociology*, 7 (1/2):126-147.

Wellman, Barry 1996 "Computer Networks as Social Networks." *Annual Review of Sociology* 22.2/3:Annual Reviews, Inc.

**APPENDIX: SOCIOLOGY WEB/  
INTERNET SITES**

These lists of Web and other Internet sites include electronic addresses for some of sociology's professional associations and some special interest groups, several resource referral sites giving many additional links, several sites giving access to sociological data, and lists of Internet Mailing Lists on "listservs." The addresses given were all current as of January 1, 1999. While many of the addresses have been stable for several years, there is no guarantee that they will remain accessible. Nonetheless, these addresses profile sociology's presence or representation on the Internet, especially the Web, and serve as entry points to sociology-related information in "cyberspace."

**PROFESSIONAL ASSOCIATIONS OF  
SOCIOLOGISTS**

American Sociological Association (ASA)  
<http://www.asanet.org/>

International Sociological Association (ISA)  
<http://www.ucm.es/OTROS/isa/>

Midwest Sociological Society  
<http://www.drake.edu/MSS/>

Pacific Sociological Association  
<http://www.csus.edu/psa/>

Society for Applied Sociology  
<http://www.appliedsoc.org/>

Society for the Study of Symbolic Interaction  
<http://www.soci.niu.edu/~sssi/>

Society for the Study of Social Problems (SSSP)

<http://itc.utk.edu/sssp/>

**SOCIOLOGY INTEREST GROUPS**

ASA Section on Sociology and Computers  
<http://www.asanet.org/Sections/computer.html>  
(see also <http://www.princeton.edu/~socomp/links.html>)

ASA Section on Organizations,  
Occupations, and Work  
<http://www.northpark.edu/acad/soc/oow/>

International Network for Social Network  
Analysis  
<http://www.heinz.cmu.edu/project/INSNA/>

**ELECTRONIC JOURNALS ON SOCIOLOGY**

The Electronic Journal of Sociology  
<http://www.sociology.org/>

Sociological Research Online (A quarterly "refereed" electronic journal of sociology from the United Kingdom)  
<http://www.socresonline.org.uk/socresonline/>

Current Research in Social Psychology (CRISP)  
<http://www.uiowa.edu/~grpproc/crisp/crisp.html>

Journal of World-Systems Research  
<http://csf.colorado.edu/wsystems/jwsr.html>

**MAJOR SOCIOLOGY DATA SOURCES**

Inter-University Consortium for Political and  
Social Research (ICPSR)  
<http://www.icpsr.umich.edu/>

General Social Survey Data and Information  
Retrieval System (GSSDIRS)  
<http://www.icpsr.umich.edu/gss>

Council of European Social Sciences Data  
Archives (CESSDA)  
<http://www.nsd.uib.no/cessda/europe.html>

Integrated Public Use Microdata Series (IPUMS)  
<http://www.ipums.umn.edu>

Social Indicators of Development (World Bank's  
data on social effects of economic development)  
<http://www.ciesin.org/IC/wbank/sid-home.html>

**MAJOR REFERRAL SITES FOR SOCIOLOGY RESOURCES**

Demography and Population Studies WWW Virtual Library

<http://coombs.anu.edu.au/ResFacilities/DemographyPage.html>

WWW Virtual Library: Sociology—SOCNET—Courses, Resources

<http://www.mcmaster.ca/socscidocs/w3virtsoclib/socnet.htm>

SOCIOWEB A Sociology Resource Center

<http://www.socioweb.com/~markbl/socioweb/>

Research engines for the social sciences

<http://tile.net/~tile.net>

WeacTies (for Sociology of the Internet)

<http://www.princeton.edu/~soccomp/weacties/>

**SELECTED NONELECTRONIC PERIODICALS IN SOCIOLOGY**

Sociological Methodology: An Annual Volume

<http://weber.u.washington.edu/~socmeth2/software.html>

Social Science Computer Review

<http://sagepub.com>

The Information Society

<http://www.ics.uci.edu/~kling/tis.html>

**INTERNET MAILING LISTS (LIST SERVERS) ON SOCIOLOGICAL TOPICS**

This list of List Servers, or Listserv discussion groups, indicate the popularity of such communication vehicles. The format for each entry is: List-Name followed by a brief one-phrase description. The second line is the Internet subscription address. Unless otherwise noted, anyone can subscribe by sending a one-line message to these addresses. The one-line message should say “subscribe” followed by the List-Name.

ASASCAN Section on Sociology and Computing (ASA)

[listserv@vm.temple.edu](mailto:listserv@vm.temple.edu)

FAMILYSCI Family studies

[listserv@ukcc.uky.edu](mailto:listserv@ukcc.uky.edu)

IVSA International Visual Sociology Association  
[listserv@pdomain.uwindsor.ca](mailto:listserv@pdomain.uwindsor.ca)

MATHSOC Mathematical Sociology

[listserv@listserv.dartmouth.edu](mailto:listserv@listserv.dartmouth.edu)

RURSOCL Rural Sociology Discussion List

[listserv@lsv.uky.edu](mailto:listserv@lsv.uky.edu)

SOCIOLOGY-L Sociology Faculty Information

[listserv@american.edu](mailto:listserv@american.edu)

SOCNET Social Networks

[listserv@lists.ufl.edu](mailto:listserv@lists.ufl.edu)

**SOCIOLOGY USENET NEWSGROUPS**

In contrast to Internet Mailing Lists, UseNet Newsgroups do not broadcast a message to all those included on the list. Internet users must specifically access the postings (messages) for any such newsgroup. While on any given day, “sociology” will be mentioned in a variety of newsgroup discussions, the only newsgroup where academic sociological topics are regularly discussed is: alt.sci.sociology. Generally a few messages are added to this discussion area every day.

RONALD E. ANDERSON

**CONFIDENCE INTERVALS**

See Statistical Inference.

**CONFLICT THEORY**

*NOTE: Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

Conflict theory explains social structure and changes in it by arguing that actors pursue their interests in conflict with others and according to their resources for social organization. Conflict theory builds upon Marxist analysis of class conflicts, but it is detached from any ideological commitment to socialism. Max Weber generalized conflict to the arenas of power and status as well as economic class, and this multidimensional approach has become widespread since the 1950s.

### WHAT CONSTITUTES A CONFLICT GROUP?

For Marx and Engels, a society's conflicting interests derive from the division between owners and nonowners of property. Dahrendorf (1959) proposed that conflicts are based on power, dividing order-givers, who have an interest in maintaining the status quo, from order-takers, who have an interest in changing it. Property is only one of the bases of power conflict, and conflicts can be expected inside any type of organization, including socialist ones. In the Weberian model there are even more types of conflict, since every cultural group (such as ethnic, religious, or intellectual groups) can also struggle for advantage. In addition, economic conflict takes place in three different types of market relations, pitting employers against workers, producers against consumers, and lenders against borrowers (Wiley 1967). Gender stratification produces yet another dimension of conflict.

### THE PROCESS OF CONFLICT

Conflicting interests remain latent until a group becomes mobilized for active struggle. This occurs when its members are physically concentrated, have material resources for communicating among themselves, and share a similar culture. The higher social classes are typically more mobilized than lower classes, and most struggles over power take place among different factions of the higher classes. Lower classes tend to be fragmented into localized groups and are most easily mobilized when they are a homogeneous ethnic or religious group concentrated in a particular place. The better organized a conflict group is, the longer and more intensely it can struggle; such struggles become routinized, as in the case of entrenched labor unions or political parties. Less organized conflict groups that become temporarily mobilized are more likely to be violent but unable to sustain the conflict.

Overt conflict increases the solidarity of groups on both sides. Coser (1956), elaborating the theory of Georg Simmel, points out that conflict leads to a centralization of power within each group and motivates groups to seek allies. A conflict thus tends to polarize a society into two factions, or a world of warring states into two alliances. This

process is limited when there are cross-cutting memberships among groups, for instance, if class, ethnic, and religious categories overlap. In these cases, mobilization of one line of conflict (e.g., class conflict) puts a strain on other dimensions of conflict (e.g., ethnic identity). Thus, cross-cutting conflicts tend to neutralize each other. Conversely, when multiple lines of group membership are superimposed, conflicts are more extreme.

Conflicts escalate as each group retaliates against offenses received from the other. How long this process of escalation continues depends on how much resources a group can draw upon: its numbers of supporters, its weapons, and its economic goods. If one group has many more resources than the other, the conflict ends when the mobilizing capacity of the weaker side is exhausted. When both sides have further resources they have not yet mobilized, escalation continues. This is especially likely when one or both sides have sustained enough damage to outrage and mobilize their supporters but not great enough damage to destroy their organizational resources for struggle.

Deescalation of conflict occurs in two very different ways. If one side has overwhelming superiority over the other, it can destroy opposition by breaking the other group's organizational capacity to fight. The result is not harmony but an uneasy peace, in which the defeated party has been turned back into an unmobilized latent interest. If neither side is able to break up the other's organization, conflict eventually deescalates when resources are eaten up and the prospects of winning become dimmer. Although wars usually arouse popular solidarity at first, costs and casualties reduce enthusiasm and bring most wars to an end within a few years. Civilian uprisings, strikes, and other small-scale conflicts typically have fewer resources to sustain them; these conflicts deescalate more quickly. During a deescalation, the points of contention among the opponents modulate from extreme demands toward compromises and piecemeal negotiation of smaller issues (Kriesberg 1982). Very destructive levels of conflict tend to end more rapidly than moderate conflicts in which resources are continuously replenished.

### COERCIVE POWER AND REVOLUTION

In a highly coercive state, such as a traditional aristocracy or a military dictatorship, power is

organized as an enforcement coalition (Collins 1988; Schelling 1962). Members of the ruling organization monitor each other to ensure loyalty. A change in power is possible only when a majority of the enforcers disobey orders simultaneously. Revolts occur in a rapid “bandwagon effect,” during which most members scramble to become part of the winning coalition. The more coercive the state, the more extreme the swings between long periods of tyrannical stability and brief moments of political upheaval.

Since the state claims a monopoly on the instruments of violence, revolutionary changes in power occur through the reorganization of coercive coalitions. Revolts from below are almost always unsuccessful as long as the state’s military organization stays intact. For this reason, revolutions typically are preceded by a disintegration of the military, due to defeat in war, depletion of economic resources in previous conflicts, and splits within the ruling group (Skocpol 1979). These breakdowns of military power in turn are determined by geopolitical processes affecting the expansion or contraction of states in the surrounding world (Stinchcombe 1968; Collins 1986).

#### WHO WINS WHAT?

Conflict shapes the distribution of power, wealth, and prestige in a society. The victorious side is generally the group that is better mobilized to act in its collective interest. In many cases, the dominant group is well organized, while the opposing interest group remains latent. The result is a stable structure of stratification, in which overt conflict rarely occurs.

Lenski (1966) showed that concentration of wealth throughout world history is determined by the interaction of two factors. The higher the production of economic surplus (beyond what is necessary to keep people alive), the greater the potential for stratification. This surplus in turn is appropriated according to the distribution of power.

Turner (1984) theorizes that the concentration of power is unequal to the extent that there is external military threat to the society or there is a high level of internal conflict among social groups. Both external and internal conflict tend to centralize power, providing that the government wins

these conflicts; hence, another condition must also be present, that the society is relatively productive and organizationally well integrated. If the state has high resources relative to its enemies, conflict is the route by which it concentrates power in its own hands.

Prestige is determined by the concentration of power and wealth. Groups that have these resources can invest them in material possessions that make them impressive in social encounters. In addition, they can invest their resources in culture-producing organizations such as education, entertainment, and art, which give them cultural domination. According to Pierre Bourdieu’s research (1984), the realm of culture is stratified along the same lines as the stratification of the surrounding society.

#### EFFECTS OF CONFLICT GROUPS UPON INDIVIDUALS

The latent lines of conflict in a society divide people into distinctive styles of belief and emotion. Collins (1975) proposed that the differences among stratified groups are due to the microinteractions of daily experience, which can occur along the two dimensions of vertical power and horizontal solidarity. Persons who give orders take the initiative in the interaction rituals described by Goffman (1959). These persons who enact the rituals of power identify with their front-stage selves and with the official symbols of the organizations they control; whereas persons who take orders are alienated from official rituals and identify with their private, backstage selves. Individuals who belong to tightly enclosed, localized groups emphasize conformity to the group’s traditions; persons in such positions are suspicious of outsiders and react violently and emotionally against insiders who are disrespectful of the group’s symbols. Loosely organized networks have less solidarity and exert less pressure for conformity. Individuals build up emotional energy by microexperiences that give them power or solidarity, and they lose emotional energy when they are subordinated to power or lack experiences of solidarity (Collins 1988). Both emotions and beliefs reproduce the stratification of society in everyday life.

(SEE ALSO: *Coalitions; Game Theory and Strategic Interaction; Interpersonal Power*)

REFERENCES

Bailey, Kenneth D. 1997 "System and Conflict: Toward a Symbiotic Reconciliation." *Quality and Quantity* 31:425–442.

Bourdieu, Pierre 1984 *Distinction*. Cambridge, Mass.: Harvard University Press.

Briggs, E. Donald 1992 "New Directions in Conflict Theory: Conflict Resolution and Conflict Transformation." *Canadian Journal of Political Science / Revue canadienne de science politique* 25:430–431.

Chapin, Mark 1994 "Functional Conflict Theory, the Alcohol Beverage Industry, and the Alcoholism Treatment Industry." *Journal of Applied Social Sciences* 18:169–182.

Collins, Randall 1975 *Conflict Sociology*. New York: Academic Press.

— 1986 *Weberian Sociological Theory*. Cambridge: Cambridge University Press.

— 1988 *Theoretical Sociology*. San Diego: Harcourt Brace Jovanovich.

Collins, Randall 1993 "What Does Conflict Theory Predict about America's Future?" *Sociological Perspectives* 36:289–313.

Coser, Lewis A. 1956 *The Functions of Social Conflict*. New York: Free Press.

Dahrendorf, Ralf 1959 *Class and Class Conflict in Industrial Society*. Stanford, Calif.: Stanford University Press.

Gagnon, V. P., Jr. 1995 "Ethnic Conflict as an Intra-Group Phenomenon: A Preliminary Framework." *Revija za Sociologiju* 26:1–2.

Glaser, James M. 1994 "Back to the Black Belt: Racial Environment and White Racial Attitudes in the South." *Journal of Politics* 56:21–41.

Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. New York: Doubleday.

Hanneman, Robert A., Randall Collins, and Gabriele Mordt 1995 "Discovering Theory Dynamics by Computer Simulation: Experiments on State Legitimacy and Imperialist Capitalism." *Sociological Methodology* 25:1–46.

Haugaard, Mark 1997 "The Consensual Basis of Conflictual Power: A Critical Response to 'Using Power, Fighting Power' by Jane Mansbridge." *Constellations* 3:401–406.

Kriesberg, Louis 1982 *Social Conflicts*. Englewood Cliffs, N.J.: Prentice-Hall.

Lenski, Gerhard E. 1966 *Power and Privilege: A Theory of Stratification*. New York: McGraw-Hill.

Schelling, Thomas C. 1962 *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.

Skocpol, Theda 1979 *States and Social Revolutions*. New York: Cambridge University Press.

Stinchcombe, Arthur L. 1968 *Constructing Social Theories*. New York: Harcourt.

Turner, Jonathan H. 1984 *Societal Stratification: A Theoretical Analysis*. New York: Columbia University Press.

— 1993 *Classical Sociological Theory: A Positivist's Perspective*. Chicago, Ill.: Nelson-Hall Publishers.

Van-Huyssteen, Elsa F. 1994 "Interpretation of the South African Legal System in Terms of the Analytical Conflict Perspective." *Suid Afrikaanse Tydskrif vir Sosiologie / South African Journal of Sociology* 25:87–94.

Wiley, Norbert F. 1967 "America's Unique Class Politics: The Interplay of the Labor, Credit, and Commodity Markets." *American Sociological Review* 32:529–540.

RANDALL COLLINS

## CONSISTENCY THEORY

See Cognitive Consistency Theories.

## CONTENT ANALYSIS

"Content analysis" has evolved into an umbrella label that includes various procedures for making reliable, valid inferences from qualitative data, including text, speech, and images. These procedures have improved and expanded due to numerous developments in recent years since this encyclopedia's first edition.

Traditionally, "content analysis" has referred to systematic procedures for assigning prespecified codes to text, such as interviews, newspaper editorials, open-ended survey answers, or focus-group transcripts, and then analyzing patterns in the codings. Some projects will count each specific occurrence within a text, while others will have coders tally the number of column inches assigned a code. Either way, the procedure usually employs

a “top down” strategy, beginning with a theory and hypotheses to be tested, developing reliable coding categories, applying these to coding-specified bodies of text, and finally testing the hypotheses by statistically comparing code indexes across documents.

With the increasing popularity of qualitative sociology, content analysis has also come to refer to “grounded” inductive procedures for identifying patterns in various kinds of qualitative data including text, illustrations, and videos. For example, the data might include observers’ detailed notes of children’s behaviors under different forms of supervision, possibly supplemented with videotapes of those same behaviors. While traditional content analysis usually enlisted statistical analyses to test hypotheses, many of these researchers do not start with hypotheses, but carefully search for patterns in their data.

However, rather than just produce statistical analyses or search for patterns, investigators should also situate the results of a content analysis in terms of the contexts in which the documents were produced. A content-analysis comparison of letters to stockholders, for example, should take into consideration the particular business sectors covered and the prevailing economic climates in which they were written. An analysis of American presidential nomination acceptance speeches should consider that they changed dramatically in form once they started to be broadcast live on national radio. A content analysis may have reliable coding, but the inferences drawn from that coding may have little validity unless the researcher factors in such shaping forces.

Like any expanding domain, there has been a tendency for content analysis to segment into specialized topics. For example, Roberts (1997) focuses on drawing statistical inferences from text, including Carley’s networking strategies and Gottschalk’s clinical diagnostic tools. There also has been a stream of instructional books, including several series published by Sage, that focus on particular kinds of qualitative data such as focus-group transcripts. A technical literature also has developed addressing specialized computer software and video-analysis techniques. Nevertheless the common agenda is analyzing the content of qualitative data. Inasmuch as our lives are shaped

by different forms of media, and inasmuch as different analytic procedures can complement one another in uncovering important insights, it makes sense to strive toward an integration, rather than fragmentation.

Many advances in content-analysis procedures have been made possible by the convenience and power of desktop and laptop computers. In addition, an overwhelming proportion of text documents are now generated on computers, making their text files computer accessible for content analysis. And a revolution in hand-held analogue and digital video cameras, together with computer-based technology for editing and analyzing videotapes, makes new research procedures feasible.

Consider, for example, new possibilities for analyzing responses to open-ended questions in survey research. For years, survey researchers have been well aware that closed-ended questions require respondents to frame how they think about an issue in terms of a question’s multiple choices, even when the choice options had little to do with how a respondent views an issue. But the costs and time involved in analyzing open-ended responses resulted in such questions rarely being used. Even when they were included in a survey, the interviewers usually just recorded capsule summaries of the responses that omitted most nuances of what was said.

Contrast this then with survey research using today’s audio information-capturing technologies. Telephone survey interviewers are guided by instructions appearing on a computer screen. Whenever an open-ended question appears, the interviewer no longer needs to type short summaries of the responses. Instead, a computer digitally captures an audio recording of each open-ended response, labels it, and files it as a computer record. Any audio response can later be easily fetched and replayed, allowing a researcher, for example, to identify a “leaky voice,” that is, one indicative of the respondent’s underlying emotion or attitude. And the full audio responses are then available to be transcribed to text, including, if desired, notations indicating hesitations and voice inflections. Until computer voice recognition is completely reliable, transcribing usually remains a manual task. But with spreadsheet software (such as *Excel*) no longer restrictively limiting the amount of text

in any one cell, the text of each individual's entire response to a question can be placed in a spreadsheet cell, thus capturing both the closed-ended and open-ended data for a survey into a convenient, single spreadsheet for researchers to analyze.

With the survey data in this convenient form, researchers can then code open-ended responses manually, putting their assigned codes in additional spreadsheet columns. As a teaching exercise, it is instructive to assign students a task such as identifying gender differences among a thousand responses to a broad open-ended question, such as a question asking respondents' views about peace or family values. Students first might sort the spreadsheet by gender in order to read separately samples of male responses and female responses and obtain a sense of what possible gender differences exist. They then develop coding instructions that capture these differences and apply the codes to the entire set of responses. This coding, of course, is better done without knowledge of the respondents' genders, with responses in a random order, and on different respondents than those used to develop the codes. After coding several hundred responses, however, students usually begin to glaze over and soon the most ardent humanist student is asking whether the computer could possibly be of help in assigning codes.

For some kinds of coding, computer help is indeed available in the form of computer programs that assign codes. Such codings can be treated as advisory and then manually confirmed, augmented perhaps by also assigning a weight. Or they may be used as is after being spot-checked for accuracy. Not only can a computer complete huge amounts of tedious coding in minutes, possibly assigning many different types of codings to each text, but these codings may uncover statistically significant frequency differences that human coders would not uncover, if only because computer analysis is so even-handed and untiring. The static created by occasional miscodings may be more than offset by gains from a reliable consistency in making many codings.

Computer coding assignments are usually based on the occurrence of words, particular senses of words, or multiword idioms appearing in the text. For example, the word "father" in a text might be coded as "male," "family member," etc. as well as

possibly "authority-role." Computer content-analysis software may search the contexts of words in the text to ferret out and correctly code common word senses. For example, for a national study of people's perceptions of African-American young males on several open-ended questions, it was particularly important for the computer to identify correctly each respondent's usages of such multi-meaning words as "race," "white," "black," and "color."

In addition to developing their own coding categories, researchers may enlist existing computer-scored categories that are relevant to the task at hand. For example, it might be hypothesized that one group being studied is more optimistic and its responses will reflect more "positive thinking" while another is more negative or pessimistic. To code "positive thinking" a researcher may want the computer to apply an existing content-analysis category that includes over 1200 words, word roots, word senses, phrasal verbs, and idioms, thus essentially covering most expressions of "positive-thinking" that occur as infrequently as three times per million words of ordinary English text. A similar category exists for negative-thinking, allowing the investigator to check whether the groups being studied differ in their coded positive thinking, negative thinking, or both. And by enlisting such standard categories, the results obtained in one study can be readily compared with results found in other studies.

Once data has been captured in a convenient format for computer use, they can be repeatedly analyzed. For example, should our now glazed-over students have any energy left after analyzing the responses by gender, they could be given an additional assignment of identifying and coding rural-urban differences in these same open-ended responses. Given so many analyses that can be made, it makes sense to let the computer do what it can, saving manual labor for those types of codings that would be hard to have a computer assign. Even multimedia qualitative-analysis software such as *HyperResearch* includes some rudimentary tools for automatic assignments.

Moreover, desktop computer software has also become available that identifies patterns in text without having to develop coding categories. For example, SPSS's *TextSmart* uses an algorithm that

groups respondents into clusters based on word-use co-occurrences within responses and then maps these clusters into a two-dimensional grid that uses colors to represent each cluster group. If such an automated inductive procedure can produce additional valid insights that other techniques are likely to overlook, then why not use it too?

Pioneering work in inductive automatic categorizing, such as Iker's (1969), usually enlisted procedures based upon correlation matrixes, such as factor analysis. These procedures tended not to be particularly suited to analyzing text both because of the shape of word-usage frequency distributions as well as the limited number of words that a correlation matrix could feasibly handle. *TextSmart*, based upon a word-distance measure, provides much more suitable solutions. Further automatic categorizing procedures may be expected from artificial intelligence, as well as from categorizing techniques being developed for Internet search engines.

Content-analysis research strategies can thus now easily be multipronged, spanning from completely automatic inductive procedures to manual coding. But even manual coding these days is likely to utilize computer software to help coders manage information. Consider these changes in costs and convenience: Unlike mainframe computing of the 1960s and 1970s, when the cost of an hour of computer time was about the same as a coder's wage for several weeks, the marginal cost of using a desktop computer is essentially the electricity it uses. Today's desktop computer is likely to be more than fivefold faster at content-analysis coding than those mainframe computers ever were. They also can access much larger dictionaries and other information in their RAM than was ever feasible on a user partition of a mainframe computer, thus making their coding more accurate and comprehensive. Moreover, a single CD-ROM full of text to be analyzed is easily popped into a desktop computer, whereas in the days of mainframe computing, a comparable amount of text would have to be keypunched on over 3,000 boxes of IBM cards.

Given today's convenience and low cost of computer-based procedures, there is no reason to limit an analysis to one approach, especially if insights gained from one approach will differ and

often complement those gained from another. Instead of being limited by technology, the limits now may lie in the skills, proclivities, and comfort zones of the researchers. Research teams, rather than individual researchers, may prove the best solution, for only in a team made up of people with complementary strengths is one likely to find the full range of statistical, conceptual, intuitive, experiential, and perhaps clinical strengths needed to carry out penetrating, comprehensive content-analysis projects. Moreover, some researchers will prefer to learn from the main trends while others will learn more from studying outlying cases. Some will learn from bottom-line numbers while others will learn more from innovative graphics that highlight information patterns. Some will focus on current data while others will contextualize data historically by comparing them with data in archives. Data that has been gathered and assembled at considerable cost, especially data-gathering that imposed on many respondents, merit as thorough and comprehensive analyses as these various procedures collectively offer.

Unfortunately, however, an "either-or" assumption about how to do content analysis has continued to be supported both by books and computer software. Authors who do an excellent job of describing one approach to content analysis, such as Boyatzis (1998), give an impression that an either-or decision has to be made about which approach to use. Some software—especially that ported from mainframe computers or developed for early desktop computers—still may steer or even limit researchers who use it to just one approach. For example, some software packages create specialized data formats such as "classification trees" that then in effect constrain the user to analyses that can be readily derived from that format. Software reviews such as Lewis's (1998) excellent comparison of *ATLAS/ti* and *NUD-IST* software have been explicit about what assumptions a researcher buys into when utilizing each package.

Additional leverage in analyzing qualitative information has stemmed from computer-based tools, such as newer versions of *HyperResearch* and *ATLAS/ti*, that integrate the handling of multiple media (text, illustrations, and video). Especially as more software comes from countries where there are expert programming skills and programming



labor is relatively inexpensive, we can expect ambitious content-analysis software, of which *TextAnalyst* from Russia ([www.megaputer.com](http://www.megaputer.com)) may be a forerunner.

Given continuing content-analysis software developments, those who would like to learn what is currently available are advised search Internet sites rather than rely upon even recently published materials. One recommended starting point is the Georgia State University content-analysis site ([www.gsu.edu/~wwwcom/content.html](http://www.gsu.edu/~wwwcom/content.html)), which gives links to software web sites (including software mentioned in this article), indexes recent content-analysis publications, and has a mailing list of more than 700 members. Technical reviews of relevant language-analysis tools, such as Berleant's (1995), occasionally appear in computational linguistics journals and web sites. For training, the University of Essex Summer School in Social Science Data Analysis and Collection ([www.essex.ac.uk](http://www.essex.ac.uk)), as part of a program of a European consortium, has been offering a content-analysis module for years as part of its program.

Given the developments described here, some of the contributions that content analysis should be able to make to sociological research include:

1. A major shift from reliance upon closed-ended questions to an appropriate use of open-ended questions that lets people be heard in the ways they frame issues, as well as the way they think and feel about them, as discussed in detail by Stone (1997)
2. A better understanding of both print and television media and its impact on public opinion, both in setting agendas and in influencing opinion intensity, as laid out in Neuman (1989). This will involve research that compares the content of media with the content of opinions. Not only will survey research data be archived and accessible from Internet servers, but full-display media will be accessible from Lexis-Nexis and on-line editions supplied by media providers, as well as television news archives such as those at Vanderbilt University
3. Better use of historical qualitative data, including both text and graphic materials,

to address such issues as how economic cycles impact ideology, as examined by Namenwirth and Weber (1987), or to uncover cycles of creativity, as demonstrated by Martindale (1990)

4. Investigations, several of which are already underway, of both intranet communication patterns within organizations as well as Internet communications, including analyses of the content of communications over those networks.

There is also, however, good reason for caution. Never before in history has so much qualitative information been available electronically. High-volume image scanning will also further increase the amount of information that can be electronically accessed and content-analyzed. Quite understandably, those agencies responsible for limiting terrorist activities may look on content-analysis procedures as possibly providing early warnings that could save lives. But these procedures can also become tools for a "big-brother" monitoring society. Sociologists have an important role in anticipating these problems and helping resolve them.

#### REFERENCES

- Berleant, Daniel 1995 "Engineering 'Word Experts' for Word Disambiguation." *Natural Language Engineering* 1 (4):339-362.
- Boyatzis, Richard E. 1998 *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, Calif.: Sage.
- Iker, Howard, and Norman Harway 1969 "A Computer Systems Approach Toward the Recognition and Analysis of Content." In George Gerbner, Ole Hosti, Klaus Krippendorff, William Paisley, and Philip Stone, eds., *The Analysis of Communication Content*. New York: Wiley.
- Lewis, R. Berry 1998 "ATLAS/ti and NUD-IST: A Comparative Review of Two Leading Qualitative Data Analysis Packages." *Cultural Anthropology Methods* 10 (3):41-47.
- Martindale, Colin 1990 *The Clockword Muse: The Predictability of Artistic Change*. New York: Basic Books.
- Namenwirth, J. Zvi, and Robert Philip Weber 1987 *Dynamics of Culture*. Winchester, Mass.: Allen and Unwin.
- Neuman, Richard 1989 "Parallel Content Analysis: Old Paradigms and New Proposals." *Public Communication and Behavior* 2:205-289.

Roberts, Carl W. (ed.) 1997 *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Mahwah, N.J.: Lawrence Erlbaum.

Stone, Philip J. 1997 "Thematic Text Analysis: New Agendas for Analyzing Text Content." In Carl W. Roberts, ed., *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Mahwah, N.J.: Lawrence Erlbaum.

PHILIP STONE

## CONTINGENCY TABLES

See Tabular Analysis; Typologies.

## CONVERGENCE THEORIES

The idea that societies move toward a condition of similarity—that they converge in one or more respects—is a common feature of various theories of social change. The notion that differences among societies will decrease over time can be found in many works of eighteenth and nineteenth century social thinkers, from the prerevolutionary French *philosophes* and the Scottish moral philosophers through de Tocqueville, Toennies, Maine, Marx, Spencer, Weber, and Durkheim (Weinberg 1969; Baum 1974). More recently, the study of "postindustrial" society and the debate over "postmodernist" aspects of contemporary society also reflect to some degree the idea that there is a tendency for broadly similar conditions or attributes to emerge among otherwise distinct and dissimilar societies.

In sociological discourse since the 1960s, the term *convergence theory* has carried a more specific connotation, referring to the hypothesized link between economic development and concomitant changes in social organization, particularly work and industrial organization, class structure, demographic patterns, characteristics of the family, education, and the role of government in assuring basic social and economic security. The core notion of convergence theory is that as nations achieve similar levels of economic development they will become more alike in terms of these (and other) aspects of social life. In the 1950s and 1960s, predictions of societal convergence were most

closely associated with modernization theories, which generally held that developing societies will follow a path of economic development similar to that followed by developed societies of the West. Structural-functionalist theorists, such as Parsons (1951) and Davis (1948), while not actually employing the terminology of convergence theory, paved the way for its development and use in modernization studies through their efforts to develop a systematic statement of the functional prerequisites and structural imperatives of modern industrial society; these include an occupational structure based on achievement rather than ascription, and the common application of universalistic rather than particularistic evaluative criteria. Also, beginning in the 1960s, convergence theory was invoked to account for apparent similarities in industrial organization and patterns of stratification found in both capitalist and communist nations (Sorokin 1960; Goldthorpe 1964; Galbraith 1967).

### CONVERGENCE THEORY AND MODERNIZATION

The conventional and most controversial application of convergence theory has been in the study of modernization, where it is associated with the idea that the experience of developing nations will follow the path charted by Western industrialized nations. Related to this idea is the notion of a relatively fixed pattern of development through which developing nations must pass as they modernize (Rostow 1960). Inkeles (1966), Inkeles and Smith (1974), and Kahl (1968) pursued the idea of convergence at the level of individual attitudes, values, and beliefs, arguing that the emergence of a "modern" psychosocial orientation accompanies national modernization (see Armer and Schnaiberg 1972 for a critique).

Kerr and colleagues' *Industrialism and Industrial Man* (1960) offers the classic statement of the "logic of industrialism" thesis, which the authors proposed as a response to Marxian theory's equation of industrial society with capitalism. More specifically, Kerr et al. sought to identify the "inherent tendencies and implications of industrialization for the work place," hoping to construct from this a portrait of the "principal features of the new society" (p. 33). The features common to

industrial society, they argued, include rapid changes in science, technology, and methods of production; a high degree of occupational mobility, with continual training and retraining of the work force; increasing emphasis on formal education, particularly in the natural sciences, engineering, medicine, managerial training, and administrative law; a workforce highly differentiated in terms of occupational titles and job classifications; the increasing importance of urban areas as centers of economic activity; and the increasing role of government in providing expanded public services, orchestrating the varied activities of a large and complex economy, and administering the “web of rules” of industrial society. Importantly, Kerr et al. envisioned these developments as cutting across categories of political ideology and political systems.

Although the “logic of industrialism” argument is often cited as a prime example of convergence theory (see Form 1979; Moore 1979; Goldthorpe 1971), Kerr et al. never explicitly made this claim for their study. While mentioning convergence at various points in their study, the authors pay equal attention to important countercurrents leading toward diverse outcomes among industrial societies. The concluding chapter of *Industrialism and Industrial Man* is, in fact, entitled “Pluralistic Industrialism,” and addresses the sources of diversity as well as uniformity among industrial societies. Among sources of diversity identified are the persistence of existing national institutions, enduring cultural differences, variations in the timing of industrialization (late versus early), the nature of a nation’s dominant industry, and the size and density of population. Counterposed against these factors are various sources of uniformity, such as technological change, exposure to the industrial world, and a worldwide trend toward increased access to education leading to an attenuation of social and economic inequality.

The critique of convergence theory in the study of modernization recalls critiques of earlier theories of societal evolution advanced under the rubric of social Darwinism in the nineteenth century and structural functionalism in the mid-twentieth century. The use of convergence theory to analyze modernization has been attacked for its alleged assumptions of unilinearity and determinism (i.e., a single path of development that all societies must follow), its teleological or historicist

character (Goldthorpe 1971), its Western ideological bias (Portes 1973), and for ignoring the structurally dependent position of less-developed countries in the world economy (Wallerstein 1974). Yet a careful review of the literature suggests that many criticisms have often tended to caricature convergence theory rather than addressing its application in actual research studies. Since the 1960s few if any researchers have explicitly claimed convergence theory, at least in its unreconstructed form, as their own. For example, Moore (1979), an exponent of the “conventional” view of modernization, subtitled his book, *World Modernization, “the limits of convergence,”* and went to great pains to distance himself from the “model modernized society” position associated with early versions of convergence theory (see Moore 1979, pp. 26–28, 150–153). And Parsons (1966), whose name is virtually synonymous with structural functionalism, concluded one of his later writings on comparative sociology with the statement that “any linear theory of societal evolution” is “untenable” (p. 114). As Form (1979) observes, convergence theory passed through a cycle typical of social science theories: a burst of initial interest and enthusiasm, followed by intense criticism and controversy, finally giving way to neglect. The major challenge to those wishing to revive convergence theory and rescue it from its critics is to specify its theoretical underpinnings more precisely, to develop appropriate empirical studies, and finally account for variation as well as similarity among observed cases.

#### FORMS OF CONVERGENCE AND DIVERGENCE

In recent years Inkeles (1980, 1981; also Inkeles and Sirowy 1983) has made the most systematic attempt to reformulate convergence theory and respecify its core hypotheses and propositions. Inkeles (1981) argues that earlier versions of convergence theory failed to distinguish adequately between different elements of the social system, which is problematic because these elements not only change at different speeds, but may move in opposite directions. He proposes dividing the social system into a minimum of five elements for purposes of assessing convergence: modes of production and patterns of resource utilization; institutional arrays and institutional forms; structures

or patterns of social relationships; systems of popular attitudes, values, and behavior; and systems of political and economic control. Finally, he specifies the different forms convergence and divergence may take: (1) simple convergence involving the movement from diversity to uniformity; (2) convergence from different directions involving movement toward a common point by an increase for some cases and a decrease for others; (3) convergence via the crossing of thresholds rather than changes in absolute differences; (4) divergent paths toward convergence, where short-term fluctuations eventually fall into line or a “deviant” case that eventually defines the norm for other cases (for example, France’s move toward small family size in the late eighteenth century); and (5) convergence in the form of parallel change, where nations all moving in the same direction along some dimension of change continue to remain separated by a gap. Although parallel change of this sort does not represent true convergence, it is consistent with the key assumption of convergence theory, namely, that “insofar as they face comparable situations of action . . . nations and individuals will respond in broadly comparable ways” (p. 21).

Inkeles (1981) also describes various forms that divergence may take: (1) simple divergence, the mirror image of simple convergence, in which movement occurs away from a common point toward new points further apart than the original condition; (2) convergence with crossover, where lines intersect and then proceed to spread apart; and (3) convergent trends masking underlying diversity (for example, although the United States, Great Britain, and Sweden all experienced large increases in public assistance programs from 1950 to the early 1970s, the social groups receiving benefits were quite different among the three nations, as were the political dynamics associated with the spending increases within each nation). Finally, Inkeles (1981) notes the importance of selecting appropriate units of analysis, levels of analysis, and the time span for which convergence, divergence, or parallel change can be assessed. These comments echo earlier sentiments expressed by Weinberg (1969) and Baum (1974) about how to salvage the useful elements of standard convergence theory while avoiding the pitfalls of a simplistic functionalist-evolutionary approach. Common to these attempts to revive convergence theory is the exhortation to develop more and better

empirical research on specific institutional spheres and social processes. As the following sections demonstrate, a good deal of work along these lines is already being done across a wide range of substantive questions and topical concerns that can aptly be described in the plural as *convergence theories*, indicating their revisionist and more pluralistic approach.

### INDUSTRIAL SOCIOLOGY

Despite criticisms of Kerr and colleagues’ (1960) concept of the logic of industrialism, the question of convergent trends in industrial organization has remained the focus of active debate and much research. The large research literature related on this question, reviewed by Form (1979), has produced mixed evidence with respect to convergence. Studies by Shiba (1973, cited in Form 1979), Form (1976), and Form and Kyu Han (1988), covering a range of industrializing and advanced industrial societies found empirical support for convergence in workers’ adaptation to industrial and related social systems, while Gallie’s (1977, cited in Form 1979) study of oil refineries in Great Britain and France found consistent differences in workers’ attitudes toward systems of authority. On the question of sectoral and occupational shifts, Gibbs and Browning’s (1966) twelve-nation study of industrial and occupational division of labor found both similarities—consistent with the convergence hypothesis—as well as differences. Studies across nations varying in levels of industrial development revealed only “small and unsystematic differences” in worker commitment (Form 1979, p. 9), thus providing some support for the convergence hypothesis. Japan has been regarded as an exceptional case among industrialized nations because of its strong cultural traditions based on mutual obligation between employers and employees. These characteristics led Dore (1973), for example, to argue vigorously against the convergence hypothesis for Japan. A more recent study by Lincoln and Kalleberg (1990) “stands convergence on its head,” arguing that patterns of work organization in the United States are being impelled in the direction of the Japanese model. Finally, with respect to women in the labor force, the evidence of convergence is mixed. Some studies found no relationship between female labor-force participation and level of industrialization

(Ferber and Lowry 1977; Safilios-Rothchild 1971), though there is strong evidence of a trend toward increasing female participation in nonagricultural employment among advanced industrial societies (Paydarfar 1967; Wilensky 1968) along with the existence of dual labor markets stratified by sex, a pattern found in both communist and capitalist nations in the 1970s (Cooney 1975; Bibb and Form 1977; Lapidus 1976).

### STRATIFICATION

Closely related to the study of industrial organization is the question of converging patterns of stratification and mobility. The attempt to discover common features of the class structure across advanced industrial societies is a central concern for social theorists of many stripes. The question has inspired intense debate among both neo-Weberian and Marxist sociologists, although the latter, for obvious ideological reasons, tend to eschew the language of convergence theory. An early statement of the class convergence thesis was made by Lipset and Zetterberg (1959), to the effect that observed rates of mobility between social classes tend to be similar from one industrial society to another. Erikson et al. (1983) conducted a detailed test of the class mobility convergence hypothesis in England, France, and Sweden, and found little support for it. They conclude that the “process of industrialization is associated with very variable patterns . . . of the social division of labour” (p. 339).

A subcategory of comparative stratification research concerns the evidence of convergence in occupational prestige. A study published in 1956 by Inkeles and Rossi, based on data from six industrialized societies, concluded that the prestige hierarchy of occupations was “relatively invariable” and tended to support the hypothesis that modern industrial systems are “highly coherent. . . relatively impervious to the influence of traditional culture patterns” (p. 329). Although the authors did not specifically mention convergence, their conclusions were fully consistent with the idea of emergent similarities. A subsequent study by Treiman (1977) extended the comparison of occupational prestige to some sixty nations, ranging from the least developed to the most developed. The study found that occupational-prestige rankings were markedly similar across all

societies, raising the question of whether convergence theory or an explanation based on the functional imperatives of social structure of all complex societies, past or present, was most consistent with the empirical findings. The conclusion was that both explanations had some merit, since although all complex societies—whether developed, undeveloped, or developing—showed similar occupational-prestige rankings, there was also evidence that the more similar societies were in levels of industrialization, the more similar their patterns of occupational-prestige evaluation appeared to be.

### DEMOGRAPHIC PATTERNS

The theory of demographic transition provides one of the most straightforward examples of convergence. The essence of the theory is that fertility and mortality rates covary over time in a predictable and highly uniform manner. Moreover, these changes are directly linked to broad developmental patterns, such as the move from a rural, agriculturally based economy to an urban-industrial one, increases in per capita income, and adult literacy (Berelson 1978). In the first stage of the demographic transition, both fertility rates and death rates are high, with population remaining fairly constant. In the second stage, death rates drop (as a result of improvements in living conditions and medical care) while fertility rates remain high, and population levels increase rapidly. In the third stage, fertility rates begin to decline, with total size of the population leveling off or even decreasing. This simple model works remarkably well in accounting for demographic patterns observed among all industrialized (and many industrializing) societies during the post-World War II period. A large spread in fertility rates among nations at the beginning of the 1950s gave way to declining rates of fertility ending with a nearly uniform pattern of zero population growth in the 1970s.

The convergent tendencies predicted by the theory of demographic transition have not gone unchallenged, however. Freedman (1979), for example, suggests that cultural factors mediate the effects of social structural factors central to transition theory. Coale (1973) and Teitelbaum (1975) note that demographic transition theory has not provided much explanatory or predictive power

with regard to the timing of population changes or the regional variations observed within nations undergoing change.

### **FAMILY**

Inkeles (1980) explored the effects of putative convergent tendencies discussed above for family patterns. While he found evidence of convergence in some aspects of family life, other patterns continue "to be remarkably stable in the face of great variation in their surrounding socio-economic conditions" (p. 34). Aspects of family life that show clear convergent patterns include the trend toward falling fertility rates and a shift in relative power and resource control in the direction of increasing autonomy of women and declining authority of parents. Other aspects of the family, such as age at first marriage, appear to present a more complex picture, with short-term fluctuations obscuring long-term changes, and great variation from one culture to another. Still other characteristics of family life seem resistant to change; cited as examples in Inkeles (1980) are cultural patterns such as veneration of elders in many Asian societies, basic human needs for companionship and psychological support, and the role of husbands helping wives with housework. In all, Inkeles (1980) estimates that only about half the indicators of family life he examined showed any convergence, and even then not always of a linear nature.

### **EDUCATION**

Following Inkeles' (1981) reformulation of convergence theory, Inkeles and Sirowy (1983) studied the educational systems of seventy-three rich and poor nations. Among thirty different "patterns of change" in educational systems examined, they found evidence of marked convergence in fourteen, moderate convergence in four, considerable variability in nine, mixed results in two, and divergence in only one. Based on these findings, they conclude that the tendency toward convergence on common structures is "pervasive and deep. It is manifested at all levels of the educational system, and affects virtually every major aspect of that system" (p. 326). Also worthy of note is that while the authors take the conventional position

that convergence is a response to pressures arising from a complex, technologically advanced social and economic system, they also identify diffusion via integration of networks through which ideas, standards, and practices in education are shared. These networks operate largely through international organizations, such as UNESCO and the OECD; their role as mediating structures in a process leading toward cross-national similarities in education constitutes an important addition to convergence theory, with wide-ranging implications for convergence in other institutions.

### **THE WELFARE STATE**

The development of the welfare state has inspired active theoretical debate and empirical research on convergence theory, with researchers divided over the nature and extent of convergence found across nations. On the one hand, there is indisputable evidence that extensive social security, health care, and related benefit programs is restricted to nations that have reached a level of economic development where a sufficient surplus exists to support such efforts. Moreover, the development of programs of the welfare state appears to be empirically correlated with distinct bureaucratic and demographic patterns that are in turn grounded in economic development. For example, Wilensky (1975) found that among sixty nations studied the proportion of the population sixty-five years of age and older and the age of social security programs were the major determinants of levels of total welfare-state spending as a percent of gross national product. Since levels of economic development and growth of the elderly population both represent areas of convergence among advanced societies, it is reasonable to expect that patterns of welfare-state development will also tend to converge. Indeed, in such respects as the development of large and expensive pension and health-care programs, of which the elderly are the major clientele, this is the case (Coughlin and Armour 1982; Hage et al. 1989). Other empirical studies have found evidence of convergence in public attitudes toward constituent programs of the welfare state (Coughlin 1980), in egalitarian political movements affecting welfare effort across nations (Williamson and Weiss 1979), and in levels of spending (Pryor 1968), normative patterns (Mishra 1976), and social control functions of welfare-state

programs across capitalist and communist nations (Armour and Coughlin 1985).

Other researchers have challenged the idea of convergence in the welfare state. In a historical study of unemployment programs in thirteen Western European nations, Alber (1981) found no evidence that programs had become more alike in eligibility criteria, methods of financing, or generosity of benefits, although he did find some evidence of convergence in duration of unemployment benefits in nations with compulsory systems. A study conducted by O'Connor (1988) testing the convergence hypothesis with respect to trends in welfare spending from 1960 to 1980 concluded that "despite the adoption of apparently similar welfare programmes in economically developed countries there is not only diversity but divergence in welfare effort. Further, the level of divergence is increasing" (p. 295). A much broader challenge to the convergence hypothesis comes from studies focusing on variations in welfare-state development among western capitalist democracies. Hewitt (1977), Castles (1978, 1982), and Korpi (1983), to cite a few leading examples, argue that variations across nations in the strength and reformist character of labor unions and social democratic parties account for large differences in the levels of spending for and redistributive impact of welfare-state programs. However, the disagreement among these studies and scholars arguing for convergence may be simply a function of case selection. For example, in a study of nineteen rich nations, Wilensky (1976, 1981) linked cross-national diversity in the welfare state to differences in "democratic corporatism," and secondarily to the presence of Catholic political parties, thus rejecting the simplistic idea that the convergence observed across many nations at widely different levels of economic development extends to the often divergent policy developments in the relatively small number of advanced capitalist societies.

The debate over convergence in the welfare state is certain to continue. A major obstacle in resolving the question involves disagreement on the nations chosen for study, selection and construction of measures (see Uusitalo 1984), and judgments about the time frame appropriate for a definitive test of the convergence hypothesis. Wilensky et al. (1985, pp. 11–12) sum up the mixed status of current research on convergence in welfare-state development as follows:

*Convergence theorists are surely on solid ground when they assert that programs to protect against the seven or eight basic risks of industrial life are primarily responses to economic development . . . . However, showing that societies have adopted the same basic programs. . . is only a partial demonstration of convergence insofar as it does not demonstrate convergence in substantive features of the programs or in the amount of variation among affluent countries compared to poor countries.*

## GLOBALIZATION

Growing attention to a variety of large-scale changes in economic relations, technology, and cultural relations, broadly subsumed under the description "globalization," has inspired renewed interest in the ideas of convergence and modernity (see Robertson 1992 for a critical account). The literature on globalization has several threads. One approach focuses on the economic and cultural impact of transnational capitalist enterprises that are judged to be responsible for the spread of a pervasive ideology and culture of consumerism (Sklair 1995). Ritzer (1993) summarizes this phenomenon as the "McDonaldization of society"—a broad reference to the ubiquity and influence of the consumer brand names (and the large corporate interests behind them) that are instantly recognizable in virtually every country in the world today. The main implication of this perspective is that indigenous industries, habits, and culture are rapidly being driven aside or even into extinction by the "juggernaut" of the world capitalist economy dominated by a relatively few powerful interests.

Meyer et al. (1997) provide a different interpretation of globalization in their work on "world society." Although they argue that "many features of the contemporary nation-state derive from a worldwide model constructed and propagated through global cultural and associational processes" (pp. 144–145), the essence of their position is that nations are drawn toward a model that is "surprisingly consensual. . . in virtually all the domains of rationalized social life" (p. 145). Meyer et al. contend that various core principles, such as those legitimating human rights and favoring environmentalism, do not emerge spontaneously as an imperative of modernity, but rather diffuse

rapidly among nations worldwide through the agency of international organizations, networks of scientists and professionals, and other forms of association. Although not referring specifically to convergence theory, this world society and culture approach makes a strong case for the emergence of widely shared structural and cultural similarities, many of which hold out the promise of improvement, among otherwise diverse nation-states.

The rapid growth of telecommunications and computing technology, especially apparent in the emergence of the Internet as a major social and economic phenomenon of the 1990s, presents yet another aspect of globalization that holds profound implications for possible societal convergence. However important and wide-ranging, the precise patterns that will ultimately emerge from these technological innovations are not yet clear. While new computing and communication technologies compress the time and space dimensions of social interaction (Giddens 1990), and have the potential to undercut national identities and cultural differences along the lines envisioned by McLuhan's (1960) "global village," the same forces of advanced technology that can level traditional differences may ultimately reinforce the boundaries of nation, culture, and social class. For example, even as the computers and related communication technologies become more ever more widely disseminated, access to and benefits from the new technologies appear to be disproportionately concentrated among the "haves," leaving the "have nots" more and more excluded from participation (Wresch 1996). Over time, such disparities might well serve to widen differences both across and within nations, thus leading toward divergence rather than convergence.

Finally, interest in convergence has also been given a boost by various political developments in the 1990s. In particular, the twin developments of the collapse of communism in Eastern Europe and the Soviet Union and the progressive weakening of economic and political barriers in Europe are notable in this regard. The demise of state socialism has revived interest in the possibilities of global economic and political convergence among advanced industrial societies (see, for example, Lenski et al. 1991, p. 261; Fukuyama 1992). Although ongoing economic and political turmoil in the "transition to capitalism" in the former Soviet

Union during the 1990s may cast serious doubt on the long-term prospects for convergence, developments have clearly moved in that direction with astonishing speed.

The continuing movement toward unification in Europe associated with the European Union (EU, formerly the European Community) represents another significant case of political and economic convergence on a regional scale. The gradual abolition of restrictions on trade, the movement of labor, and travel among EU nations (and not least of all the establishment of a single currency in 1999), and the harmonization of social policies throughout the EU, all signal profound changes toward growing convergence in the region that promises to continue into the twenty-first century.

### CONCLUSION

The idea of convergence is both powerful and intuitively attractive to sociologists across a range of backgrounds and interests (Form 1979). It is difficult to conceive of an acceptable macro theory of social change that does not refer to the idea of convergence in one way or another. Despite the controversy over, and subsequent disillusionment with, early versions of convergence theory in the study of modernization, and the often mixed results of empirical studies discussed above, it is clear that the concept of societal convergence (and convergence theories that allow for the possibility of divergence and invariance) provides a useful and potentially powerful analytical framework within which to conduct cross-national studies across a broad range of social phenomena. Even where the convergence hypothesis ultimately ends up being rejected, the perspective offered by convergence theories can provide a useful point of departure for research. Appropriately reformulated, focused on elements of the social system amenable to empirical study, and stripped of the ideological baggage associated with its earlier versions, convergence theories hold promise to advance the understanding of the fundamental processes and regularities of social change.

### REFERENCES

- Alber, Jens 1981 "Government Responses to the Challenge of Unemployment: The Development of Unemployment Insurance in Western Europe. In Peter



- Flora and Arnold J. Heidenheimer, eds., *The Development of Welfare States in Europe and America*. New Brunswick, N.J.: Transaction Books.
- Armer, Michael, and Allan Schnaiberg 1972 "Measuring Individual Modernity: A Near Myth." *American Sociological Review* 37:301-316.
- Armour, Philip K., and Richard M. Coughlin 1985 "Social Control and Social Security: Theory and Research on Capitalist and Communist Nations." *Social Science Quarterly* 66:770-788.
- Aron, Raymond 1967 *The Industrial Society*. New York: Simon and Schuster.
- Baum, Rainer C. 1974 "Beyond Convergence: Toward Theoretical Relevance in Quantitative Modernization Research." *Sociological Inquiry* 44:225-240.
- Bendix, Reinhard 1967 "Tradition and Modernity Reconsidered." *Comparative Studies in Society and History* 9:292-346.
- Berelson, B. 1978 "Prospects and Programs for Fertility Reduction: What? Where?" *Population and Development Review* 4:579-616.
- Bibb, R. and W.H. Form 1977 "The Effects of Industrial, Occupational, and Sex Stratification in Blue-Collar Markets." *Social Forces* 55:974-996.
- Castles, Francis 1982 "The Impact of Parties on Public Expenditure." In Frances Castles, ed., *The Impact of Parties: Politics and Policies in Democratic Capitalist States*. Beverly Hills, Calif.: Sage.
- 1978 *The Social Democratic Image of Society*. London: Routledge and Kegan Paul.
- Coale, Ansley J. 1973 "The Demographic Transition." In *The International Population Conference, Liège*. Liège, Belgium: International Union for the Scientific Study of Population.
- Cooney, R.S. 1975 "Female Professional Work Opportunities: A Cross-National Study." *Demography* 12:107-120.
- Coughlin, Richard M. 1980 *Ideology, Public Opinion and Welfare Policy*. Berkeley: University of California, Institute of International Studies.
- , and Philip K. Armour 1982 "Sectoral Differentiation in Social Security Spending in the OECD Nations." *Comparative Social Research* 6:175-199.
- Davis, Kingsley 1948 *Human Society*. New York: Macmillan.
- Dore, Ronald 1973 *British Factory-Japanese Factory*. Berkeley: University of California Press.
- Erikson, Robert, John H. Goldthorpe, and Lucienne Portocarero 1983 "Intergenerational Class Mobility and the Convergence Thesis: England, France, and Sweden." *British Journal of Sociology* 34:303-343.
- Ferber, M., and H. Lowry 1977 "Women's Place: National Differences in the Occupational Mosaic." *Journal of Marketing* 41:23-30.
- Form, William 1979 "Comparative Industrial Sociology and the Convergence Hypothesis." *Annual Review of Sociology* 4:1-25.
- 1976 *Blue-Collar Stratification*. Princeton, N.J.: Princeton University Press.
- , and Bae Kyu Han 1988 "Convergence Theory and the Korean Connection." *Social Forces* 66:618-644.
- Fox, William S., and William W. Philliber 1980 "Class Convergence: An Empirical Test." *Sociology and Social Research* 64:236-248.
- Freedman, Ronald 1979 "Theories of Fertility Decline: A Reappraisal." *Social Forces* 58:1-17.
- Fukuyama, Francis 1992 *The End of History and the Last Man*. New York: Free Press.
- Galbraith, John Kenneth 1967 *The New Industrial State*. Houghton-Mifflin.
- Gibbs, J.P., and H.L. Browning 1966 "The Division of Labor, Technology, and the Organization of Production in Twelve Countries." *American Sociological Review* 31:81-92.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Stanford, Calif.: Stanford University Press.
- Goldthorpe, John H. 1971 "Theories of Industrial Society: Reflections on the Recrudescence of Historicism and the Future of Futurology." *Archives of European Sociology* 12:263-288.
- 1964 "Social Stratification in Industrial Society." In P. Halmos, ed., *The Development of Industrial Societies*, 97-123. Keele, Eng.: University of Keele.
- Hage, Jerald, Robert Hanneman, and Edward T. Gargan 1989 *State Responsiveness and State Activism*. London: Unwin Hyman.
- Hewitt, Christopher 1977 "The Effect of Political Democracy and Social Democracy on Equality in Industrial Societies: A Cross-National Comparison." *American Sociological Review* 42:450-464.
- Inkeles, Alex 1981 "Convergence and Divergence in Industrial Societies." In M.O. Attir, B.H. Holzner, and Z. Suda, eds., *Directions of Change: Modernization Theory, Research, and Realities*, 3-38. Boulder, Colo.: Westview Press.
- 1980 "Modernization and Family Patterns: A Test of Convergence Theory." In D.W. Hoover and J.T.A. Koumoulides, eds., *Conspectus of History*, 31-64. Cambridge, Eng.: Cambridge University Press.
- 1966 "The Modernization of Man." In M. Weiner, ed., *Modernization* New York: Basic Books.

- , and Larry Sirowy 1983 “Convergent and Divergent Trends in National Educational Systems.” *Social Forces* 62:303–333.
- Inkeles, Alex, and David E. Smith 1974 *Becoming Modern*. Cambridge, Mass.: Harvard University Press.
- Inkeles, Alex, and Peter H. Rossi 1956 “National Comparisons of Occupational Prestige.” *American Journal of Sociology* 61:329–339.
- Kahl, J.A. 1968 *The Measurement of Modernization: A Study of Values in Brazil and Mexico*. Austin: University of Texas Press.
- Kerr, Clark, John T. Dunlap, Frederick H. Harbison, and Charles A. Myers 1960 *Industrialism and Industrial Man*. Cambridge, Mass.: Harvard University Press.
- Korpi, Walter 1983 *The Democratic Class Struggle*. London: Routledge and Kegan Paul.
- Lapidus, Gail W. 1976 “Occupational Segregation and Public Policy: A Comparative Analysis of American and Soviet Patterns.” In M. Blaxell and B. Reagan, eds., *Women and the Workplace* Chicago: University of Chicago Press.
- Lenski, Gerhard, Jean Lenski, and Patrick Nolan 1991 *Human Societies*, 6th ed. New York: McGraw-Hill.
- Lincoln, James R., and Arne L. Kalleberg 1990 *Culture, Control, and Commitment: A Study of Work Organization and Work Attitudes in the United States and Japan*. Cambridge, Eng.: Cambridge University Press.
- Lipset, Seymour M., and Hans Zetterberg 1959 “Social Mobility in Industrial Societies.” In S.M. Lipset and R. Bendix, eds., *Social Mobility in Industrial Society*. London: Heinemann.
- McLuhan, Marshall 1960 *Exploration in Communications*. Boston: Beacon Press.
- Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez 1997 “World Society and the Nation-State.” *American Journal of Sociology* 103:144–181.
- Mishra, Ramesh 1976 “Convergence Theory and Social Change: The Development of Welfare in Britain and the Soviet Union.” *Comparative Studies in Society and History* 18:28–56.
- Moore, Wilbert E. 1979 *World Modernization: The Limits of Convergence*. New York: Elsevier.
- O’Connor, Julia S. 1988 “Convergence or Divergence?: Change in Welfare Effort in OECD Countries 1960–1980.” *European Journal of Political Research* 16:277–299.
- Parsons, Talcott 1966 *Societies: Evolutionary and Comparative Perspectives*. Englewood Cliffs, N.J.: Prentice Hall.
- Parsons, Talcott 1951 *The Social System*. Glencoe, Ill.: The Free Press.
- Paydarfar, A.A. 1967 “Modernization Process and Demographic Changes.” *Sociological Review* 15:141–153.
- Portes, Alejandro 1973 “Modernity and Development: A Critique.” *Comparative International Development* 8:247–279.
- Pryor, Frederick 1968 *Public Expenditure in Capitalist and Communist Nations*. Homewood, Ill.: Irwin.
- Ritzer, George 1993 *The McDonalidization of Society*. Newbury Park, Calif.: Sage.
- Robertson, Roland 1992 *Globalization: Social Theory and Global Culture*. London and Newbury Park, Calif.: Sage.
- Rostow, Walt W. 1960 *The Stages of Economic Growth*. Cambridge, Eng.: Cambridge University Press.
- Safilios-Rothchild, C. 1971 “A Cross-Cultural Examination of Women’s Marital, Educational and Occupational Aspirations.” *Acta Sociologica* 14:96–113.
- Sklair, Leslie 1995 *Sociology of the Global System*. Baltimore, Md.: Johns Hopkins University Press.
- Sorokin, Pitirim A. 1960 “Mutual Convergence of the United States and the U.S.S.R. to the Mixed Sociocultural Type.” *International Journal of Comparative Sociology*. 1:143–176.
- Teitelbaum, M.S. 1975 “Relevance of Demographic Transition for Developing Countries.” *Science* 188:420–425.
- Treiman, Donald J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
- Uusitalo, Hannu 1984 “Comparative Research on the Determinants of the Welfare State: The State of the Art.” *European Journal of Political Research* 12:403–442.
- Wallerstein, Immanuel 1974 “The Rise and Future Demise of the World Capitalist System: Concepts for Comparative Analysis.” *Comparative Studies in Social History* 16:287–415.
- Weed, Frank J. 1979 “Industrialization and Welfare Systems: A Critical Evaluation of the Convergence Hypothesis.” *International Journal of Comparative Sociology* 20:282–292.
- Weinberg, Ian 1969 “The Problem of the Convergence of Industrial Societies: A Critical Look at the State of the Theory.” *Comparative Studies in Society and History* 11:1–15.
- Wilensky, Harold L. 1981 “Leftism, Catholicism, and Democratic Corporatism: The Role of Political Parties in Recent Welfare State Development.” In Peter Flora and Arnold J. Heidenheimer, eds., *The Development of Welfare States in Europe and America*, 345–382. New Brunswick, N.J.: Transaction Books.
- 1976 *The “New Corporatism”: Centralization and the Welfare State*. Beverly Hills, Calif.: Sage.

- 1975 *The Welfare State and Equality*. Berkeley: University of California Press.
- 1968 “Women’s Work: Economic Growth, Ideology, Structure.” *Industrial Relations* 7:235–248.
- , Gregory M. Luebbert, Susan Reed Hahn, and Adrienne M. Jamieson 1985 *Comparative Social Policy: Theory, Methods, Findings*. Berkeley: University of California, Institute of International Studies.
- Williamson, John B., and Joseph W. Weiss 1979 “Egalitarian Political Movements, Social Welfare Effort and Convergence Theory: A Cross-National Analysis.” *Comparative Social Research* 2:289–302.
- Williamson, John B., and Jeanne J. Fleming 1977 “Convergence Theory and the Social Welfare Sector: A Cross-National Analysis.” *International Journal of Comparative Sociology* 18:242–253.
- Wresch, William 1996 *Disconnected: Haves and Have Nots in the Information Age*. New Brunswick, N.J.: Rutgers University Press.

RICHARD M. COUGHLIN

## CONVERSATION ANALYSIS

Conversation analysis has evolved over several decades as a distinct variant of ethnomethodology. Its beginnings can be traced to the mid-1960s, to the doctoral research and the unpublished but widely circulated lectures of Harvey Sacks. Sacks was a University of California sociologist who had studied with Harold Garfinkel, the founder of the ethnomethodological movement, as well as with Erving Goffman. While not an ethnomethodologist, Goffman’s proposal that face-to-face interaction could be an analytically independent domain of inquiry certainly helped inspire Sacks’s work. Two other key figures whose writings (separately and together with Sacks) contributed to the emergence of conversation analysis were Gail Jefferson, one of Sacks’s first students, and Emanuel A. Schegloff, another sociologist trained in the University of California system who was decisively influenced by Garfinkel and, in much the same manner as Sacks, by Goffman (Schegloff 1988).

Sacks, like Garfinkel, was preoccupied with discovering the methods or procedures by which humans coordinate and organize their activities, and thus with the procedures of practical, common-sense reasoning in and through which “social

order” is locally constituted (Garfinkel [1967] 1984). In addressing this problem, he devised a remarkably innovative approach. Working with tapes and transcripts of telephone calls to a suicide prevention center (and with recordings of other, somewhat more mundane sorts of conversations), Sacks began examining the talk as an object in its own right, as a fundamental type of social action, rather than primarily as a resource for documenting other social processes. In short, Sacks came to recognize that the talk itself was the action. It was in the details of the talk that we could discover just how what was getting done in the activity was accomplished, systematically and procedurally, then and there, by the coparticipants themselves. This appeared to be an especially fruitful way of investigating the local production of social order.

As Schegloff (1989, p. 199) later wrote in a memoir of these first years, Sacks’s strategy in his pioneering studies was to first take note of how members of society, in some actual occasion of interaction, achieved some interactional effect—for example, in the suicide center calls, how to exhibit (and have others appreciate) that one has reasonably, accountably, arrived at the finding “I have no one to turn to”—and then to ask: Was this outcome accomplished methodically? Can we describe it as the product of a method of conduct, such that we can find other enactments of that method that will yield the same outcome, the same recognizable effect? This approach provided, Sacks suggested, an opportunity to develop formal accounts of “members’ methods” for conducting social life.

In this way, Sacks sought to address the basic question of (as he put it in one of his early manuscripts) “what it is that sociology can aim to do, and . . . how it can proceed” (Sacks [1964–1968] 1984, p. 21). Sociology, he argued, could be a “natural observational science,” concerned with the methodic organization of naturally occurring events, rather than with behavior that was manipulated through experimental techniques or other interventions such as surveys, interviews, and the like. And it could be committed to direct observation of this organization *in situ*, rather than dependent upon analytic theorizing and a concomitant reliance on idealized models of action.

Naturalistic observation also met the ethnomethodological mandate that all evidence

for the use of members' methods, and for members' orientation to or tacit knowledge of them, was to be derived exclusively from the observed behavior of the coparticipants in an interactional event. As Schegloff and Sacks (1973, p. 290) subsequently summarized the logic of this stance, if the event, the recorded conversational encounter, exhibited a methodically achieved orderliness, it "did so not only to us [the observing analysts], indeed not in the first place for us, but for the co-participants who had produced" it. After all, the task was to discover members' methods for coordinating and ordering conversational events, and these could not in any way be determined by analysts' conceptual stipulations or deduced from inventive theories.

The contrast with other methods and approaches for studying interactional processes could not be sharper, particularly with those methods adopted by Bales (1950) and Homans (1961) and their many followers, with their commitment to theoretically derived and precise operational definitions of social phenomena as a prerequisite to any scientific investigation (see Sacks 1992, vol. 1, p. 28 and p. 105, for his thinking on Bales and Homans). But Sacks's methodological stance contrasts even with those traditions usually regarded as neighbors of conversation analysis, such as symbolic interactionism and Goffman's "micro-Durkheimian" approach. These approaches assume that without an analytically stipulated conceptual scheme, there is no orderliness (or the orderliness cannot be seen) in what Garfinkel (1991) terms "the plenum"—the plentitude of members' lived experience.

Sacks was also making a well-reasoned argument for the importance of studying mundane conversation, directly confronting the belief that sociology's overriding concern should be the study of "big issues"; that is, the belief that the search for social order should center on the analysis of large-scale, massive institutions. Social order, he insisted, can be found "at all points," and the close study of what from conventional sociology's point of view seemed like small (and trivial) phenomena—the details of conversation's organization—might actually give us an enormous understanding of the way humans do things and the kinds of methods they use to order their affairs (Sacks [1964–1968] 1984, p. 24).

This last proposition bears special emphasis, for Sacks felt that these details went unnoticed,

and perhaps could not even be imagined, by conventional analytic sociology. When had sociologists concerned themselves with the profoundly methodic character of things like how to avoid giving your name without refusing to give it, he argued, or with how to get help for suicidalness without requesting it? Or with members' methods for things like "doing describing" and "recognizing a description," methods that provide for hearing the first two sentences from a story told by a young child—"The baby cried. The mommy picked it up."—as saying: The mommy who picked up the baby is the baby's mommy, and she picked it up because it was crying. Though apparently mundane, this observation provided the basis for a series of investigations regarding members' categorization methods and eventually came to provide for an entirely different approach to studies of social institutions and phenomena like race and gender. Sacks noted that since every person can be categorized at any time in various ways (for example, in terms of age, gender, or stage of life), a person's use of, or reliance on, one category rather than another to guide his or her actions with others must be grounded in one of a multitude of discoverable systems of relevance, some of which are potentially applicable in any situation, such as age, race, and gender, and others that are more limited in their use, such as occupationally defined categories or those made relevant by the organization of conversation itself, such as speaker/hearer, caller/called, and the like.

As these last examples of categories suggest, "membership analysis"—making sense of who someone is for the purpose of appropriately designing some next action—is an unremitting problem for members of a society, one that has to be solved in real time, and for which there is no single solution. Furthermore, who a person relevantly "is" for the purposes of some next action can change from moment to moment. As a consequence, Sacks pointed out that sociologists can no longer innocently categorize populations whatever way they (and their theories) see fit. Instead, analysts must similarly demonstrate the relevance of a category *to the participants* in any scene, as well as its consequentiality for them in terms of how the action proceeds, in order to ground its use in any sociological investigation (see Silverman 1998 for a useful summary of this argument).

Such grounding would stand in contrast to the epidemiological uses of categorization that underpin standard social scientific research. Sacks's colleague, anthropologist Michael Moerman (1974, pp. 67–68), once noted that social scientists “have an apparent inability to distinguish between warm . . . human bodies and one kind of identification device which some of those bodies sometimes use.” Further, since Sacks was primarily concerned with categorization as a thoroughly practical, procedural activity for members he was not much interested in the content of categories, with drawing the cultural grids that preoccupied cognitive anthropologists and many social psychologists. Instead, Sacks believed that by starting with the close study of actual events—such as members' observable use of categories *in situ*—and showing that they happened in an endogenously, socially organized manner, a much sounder basis for studying and understanding social life could be established.

It should be evident, then, that the appellation *conversation analysis* does not really capture the enterprise's commitment to addressing the most basic problem for the social sciences: the underlying character and structure of social action. As the title of one of Sacks's first publications, “An Initial Investigation of the Usability of Conversational Data for Doing Sociology,” makes clear, the use of recorded conversational materials was more of an opportunistic research strategy than a commitment to studying talk *per se* (Sacks 1972). Tape recordings of conversations constituted a record of the details of actual, singular events that could be replayed and studied extensively, and would permit other researchers direct access to exactly these same details. Still, for these identical reasons, it is conversation's organization—its detectable, orderly properties—that has remained the concrete object of study for the enterprise.

During years since the “initial investigations,” conversation analysis has given rise to a substantial research literature. Pursuing the lines of analysis first identified in the early studies while simultaneously opening up many new avenues of inquiry, researchers working in this tradition have produced findings that are, in the words of one contemporary practitioner, “strikingly cumulative and interlocking” (Heritage 1987, p. 256). Important collections of papers include those of Sudnow (1972), Schenkein (1978), Atkinson and Heritage (1984), Button and Lee (1987), and ten Have and

Psathas (1995). Sacks's lectures have now been edited and published in complete form (Sacks 1992). Special issues of *Sociological Inquiry*, *Social Psychology Quarterly*, *Human Studies*, *Social Problems*, *Research on Language and Social Interaction*, *Text*, and the *Western Journal of Speech Communication* have also been devoted to ethnomethodological and conversation-analytic topics.

Three major domains in conversation's organization identified in this literature are the organization of sequences, of turn taking, and of repair. These organizations can be described as systems of naturally organized activity, systems known and used by members as courses of practical action and practical reasoning, and designed to resolve generic problems of coordination that confront any conversationalist (and perhaps members of all social species). A sketch of some research findings with respect to these organizations should serve to illustrate how they function in this fashion, as well as the interlocking nature of their domains.

Consider first the organization of sequences. Begin with the fact that even the most cursory inspection of conversational materials reveals that talk-in-interaction has a serial arrangement to it. For example, in a conversation between two parties, party A will talk first, then party B, then A, then B, and so forth. Accordingly, in two-party conversations, turns at talk constitute a series of alternately produced utterances: ABABAB. But overlaying this serial arrangement of utterances are distinctly characterizable conversational sequences, where turns at talk do not simply happen to occur one after the other but rather “belong together” as a socio-organizational unit, and where there is thus a methodic relationship between the various turns or parts.

This methodic, structurally linked relationship between sequence parts is central to how sequences work in resolving coordination problems in conversation. This point can be demonstrated by briefly focusing on one of the earliest studies of sequence organization by Schegloff (1968), an investigation into how the initiation of conversational interactions is coordinated. Schegloff directed attention to a frequently occurring initial exchange, which was called a “summons-answer sequence.” This sequence is composed, he discovered, of closely linked parts. The production of the

first turn in the sequence, the summons, projected a relevant next action, an answer, to be accomplished by the recipient of the summons in the very next turn. Moreover, the occurrence of the expected answer cannot properly be the final turn in the exchange. The summons–answer exchange is therefore nonterminal: Upon production of the answer, the summoner is then expected to speak again, to provide the reason for the summons. This provides for a coordinated entry into conversation, and for the possibility of an extended spate of talk.

Observe that a set of mutual obligations is established by the structural relationships between these sequence parts, with each current action projecting some “next.” In the strongest form of these obligations (sequence classes vary in this regard), the property of “conditional relevance” holds between the parts of a sequence unit. A “summons–answer” sequence is but one type of a large class of utterance units, known as “adjacency pairs,” that are characterized by this property. Examples here include “greeting–greeting,” “question–answer,” and “invitation–acceptance/declination.” In adjacency pairs, when one utterance or action is conditionally relevant on another, the production of the first provides for the occurrence of the second. It could be said, then, using the example above, that the issuance of a summons is an action that selects a particular next action, an answer, for its recipient. If this action does not occur, its nonoccurrence will be a noticeable event. That is to say, it is not only nonoccurring, it is notably, “officially” absent; accordingly, this would warrant various inferences and actions. For instance, the summoner might infer that a recipient “didn’t hear me,” which would provide for the relevance and grounds of a repetition of the summons.

The discovery that human activities like conversation were coordinated and organized in a very fundamental way by such methodic relationships between actions, with some current or “first” action projecting and providing for some appropriate “second,” led to investigations into the various methods by which the recipient of a first may accomplish a second, or recognizably hold its accomplishment in abeyance until issues relevant to its performance are clarified or resolved, or avoid its accomplishment altogether by undertaking some other activity. Researchers learned, for

example, that for some firsts, there was not a single appropriate second but rather a range of alternative seconds. Note that in the examples of adjacency pair structures listed just above, invitations project either an acceptance or a declination as a course of action available to the recipient. In this case, and in others like “request–granting/denial” and “compliment–acceptance/rejection,” it was found that the alternative second parts are not generally of equal status; rather, some second parts are preferred and others dispreferred, these properties being distinct from the desires or motivations of the coparticipants. “Preference” thus refers to a structural rather than dispositional relationship between alternative but nonequivalent courses of action. Evidence for this includes distributional data across a wide range of speakers and settings, and, more important, the fact that preferred and dispreferred alternatives are regularly performed in distinctively different ways. The preference status of an action is therefore exhibited in how it is done.

Related to this, conversation analytic researchers observed that the producers of a first action often dealt in systematic, methodic ways with these properties of preference organization. To take one example, the producer of a request can and often does analyze the recipient silence that follows as displaying or implicating a denial—a denial as-yet-unstated, but nevertheless projected—and seeks to preempt the occurrence of this dispreferred action by issuing a subsequent version of the request, before the recipient starts to speak. Subsequent versions attempt to make the request more acceptable and provide another opportunity for a favorable response (Davidson 1984).

Moreover, members were observed to orient to the properties of preference organization through their performance of actions plainly meant to be understood as specifically preliminary to some adjacency pair first action. Such “pre”-type actions are designed to explore the likelihood that producing that first part of some pair will not be responded to in a dispreferred way. For instance, an utterance like “Are you doing anything tonight?” provides, in a methodical way, an opportunity for its producer to determine, without yet having to actually issue the invitation, whether it would most likely be declined. Similarly, this provides an opportunity for the recipient of the “pre”

to indicate that a dispreferred action would be forthcoming without ever having to perform that action. Additionally, because “pre” actions themselves engender sequences by making some response to them a relevant next action, they constitute the first part of a “pre-sequence.” It follows that since these and other features of preference organization together maximize the likelihood of preferred actions and minimize the likelihood of dispreferred ones, they serve as important structural resources for maintaining social solidarity and “preserving face.”

These interrelated observations on the organization of sequences were generalized outward in conversation analytic research from the relatively simple adjacency pair organization by the recognition that virtually every utterance occurs at some sequentially relevant, structurally defined place in talk (see especially Atkinson and Heritage 1984, pp. 5–9). Moreover, it is this placement that provides the primary context for an utterance’s intelligibility and understanding. Put another way, utterances are in the first place contextually understood by reference to their placement and participation within sequences of action, and it is therefore sequences of action, rather than single utterances or actions, that have become the primary units of analysis for the conversation-analytic enterprise. Accordingly, researchers in this tradition have not restricted themselves to studying only especially “tight” sequence units, but have instead broadened their investigations to (mentioning just a few) the sequencing of laughter, disputes, story and joke telling, political oratory, and the initiation and closing of topics. In addition, the sequential organization of gaze and body movement in relation to turns at talk has been the focus of some truly pathbreaking research using video recordings (see, for example, Goodwin 1981, 1994; Heath 1986).

Now let us consider the organization of turn taking, surely a central feature of virtually all talk-in-interaction. Recall that the prior discussion on the organization of sequences frequently made reference to sequence parts as “turns,” implicitly trading on the understanding that talk in conversation is produced in and built for turns, with recurring speaker change and a consequent serial ordering of utterances. In conversation, this turn ordering, as well as the size and content of each

turn, is not predetermined or allocated in advance. Instead, it is locally determined, moment-by-moment, by the coparticipants in the talk. In fact, this completely local determination of who speaks when, how long they speak, and what they might say or do in their turn, is what provides for talk being hearable as a “conversation,” rather than as, say, a debate or a ceremony of some kind. But this does not tell us just how—*methodically*—speaker change is achieved such that, ordinarily, one party talks at a time and there is little or no silence (or “gap”) between turns. Clearly, this requires close coordination among coparticipants in any conversational encounter. The systematic practices by which this is accomplished are analyzed by Sacks, Schegloff, and Jefferson in a 1974 paper that remains one of the most important in the conversation-analysis literature.

Basic to the accomplishment of turn taking is the practice of changing speakers at possible utterance completion places, what Sacks, Schegloff, and Jefferson term *transition relevance places*. How are such places, where speaker change may relevantly occur but is in no way guaranteed or required, discernable by members? A key feature of the units by and through which turns are constructed offers one resource here: For an utterance to be usable as a turn constructional unit, it must have a recognizable completion, and that completion must be recognizable prior to its occurrence (Sacks, Schegloff, and Jefferson [1974] 1978, p. 12). That is to say, its completion is projectable, and a coparticipant in the conversation who wishes to speak next can therefore begin his or her turn just at the place where the current speaker projects completion.

Of course, this does not preclude this coparticipant, or any other, from starting to speak elsewhere in the course of a current speaker’s turn. (Indeed, what actually constitutes a “turn at talk” is as locally and mutually determined as any other aspect of conversation’s organization, even as the resources for doing so are general ones.) There are various interactional moves that could involve, as one way they might be accomplished, this sort of action. At the same time, however, research on turn-taking has revealed that turns beginning elsewhere may well be met with procedures systematically designed to enforce the practice of starting at possible completion places. Further, features of the turn taking system such as that

described just above account for a great deal of the overlapping speech that can occasionally be observed. For instance, a speaker might append a tag question like “you know?” to his or her turn, while a coparticipant, having no resources available to project such an action, starts to speak just prior to or at the beginning of that appended tag, at the place that was projectably the “first possible completion” of the turn. This would result in overlapping speech, with both parties talking simultaneously. This was just one example; studies of “more than one party at a time” speech have uncovered massive evidence that its occurrence and its resolution (the restoration of one party at a time), as well as the solution to the problem of which overlapping action should then be consequential for next action, is methodically organized.

Having described the function of turn constructional practices in turn taking, Sacks, Schegloff, and Jefferson still faced the issue of how coparticipants, at possible completion places, determine just who will be the “next speaker” (note in this regard that conversation can involve more than two parties) or even if there will be a next speaker, given that a current speaker might want to continue talking. They discovered that to deal with this problem, members have available a “turn allocational component” for the system. This component consists of a set of ordered rules that come into play at transition relevance places and which provide for the methodic allocation of the right to produce a next turn, or more accurately, a turn constructional unit. In related research, methods for securing the temporary suspension of turn-taking procedures (to tell an extended story, for example) and for coordinating exit from the system (to end the conversation) have been documented.

Finally there is the entire set of procedures by which any troubles in speaking, hearing, and understanding talk are systematically handled and “repaired.” As Schegloff (1979, p. 269) points out, insofar as “any of the systems and contingencies implicated in the production and reception of talk—articulatory, memory, sequential, syntactic, auditory, ambient noise, etc.—can fail,” any piece of talk is susceptible to, or can reveal, troubles in speaking, hearing, or understanding. As a consequence, members of society must have some systematically organized set of methods for managing such trouble when it arises. Further, in order for interaction to serve as a primary site for the

coordination of social activity, any such troubles must be located and dealt with as quickly as possible to avoid whole stretches of talk developing on a problematic basis. Finally, this set of methods must provide the opportunity to discover and display trouble in speaking, hearing, or understanding by any of the ratified coparticipants to the interaction, while simultaneously managing such trouble from the variety of quarters from which it might arise, whether the trouble is noticed or produced by the current speaker or her recipient, and whether its source is endogenous to the interaction or impinges on it from outside.

When Schegloff, Sacks, and Jefferson (1977) began examining the related set of practices through which speakers managed such troubles they discovered two important features. First, they noticed that participants in interaction treat the initiation of repair as a separate matter from the actual accomplishment of a solution. That is, they distinguish between the various practices for locating a trouble source and making it the focus of the interaction and the set of practices for implementing a solution. Second, they observed that these two activities were not distributed evenly among the parties: The organization of repair exhibited a preference for self-repair and a preference for self-initiation of repair. And they went on to show that this latter feature is primarily a product of the way that the organization of repair relies on, and is fitted to, the system for distributing turns.

The organization of repair initiation operates in a restricted “repair initiation opportunity space” that is organized around the trouble source or “repairable.” Within this repair initiation opportunity space each party to an interaction moves through a series of discrete opportunities to locate and indicate potential and actual troubles. In turn, these discrete opportunities to initiate repair shape where (relative to the trouble source) a repair is effected, and by whom. The current speaker has the first opportunity to initiate repair on any trouble source within his or her own turn while still in the midst of it, or just after it is complete but before a next speaker starts. If they do initiate repair during (or immediately following the possible completion of) their own turn, such speakers also have the first opportunity to effect repair as well.

Of course, as we noted above, conversation is characterized by the alternation between current



and next, thus once a current speaker completes her turn a next speaker begins, typically by addressing herself to that just-prior talk. Accordingly, if a next speaker has some trouble with the prior speaker's turn, the next turn is the place where she can initiate repair (using a variety of forms, including "what?" and "huh?" and other designs that vary in the degree to which they specify the exact source of trouble). By initiating repair using one of these methods, that speaker selects the speaker of the trouble source to speak next, and to offer a solution to the trouble indicated. If the next speaker has no trouble with the prior turn, and she uses it to move the action forward (instead of stopping it to initiate repair), her turn will display a variety of understandings regarding the talk it follows. In doing so, her turn may also reveal some type of misunderstanding (from the point of view of the speaker of the prior turn). If that occurs the speaker of the prior turn can then initiate repair in "turn after next" (or "third position") and offer a solution immediately. Perhaps the recurrent and recognizable format for this is "I don't mean *x*, I mean *y*."

Thus, the movement of talk through these three positions—current, next turn, and turn after next—systematically provides the various parties to the interaction the opportunity to detect any trouble in speaking, hearing, and understanding, whatever its source, and initiate repair on it. As a consequence almost all instances of repair are initiated in one of these adjacent locations. The localization of repair initiation opportunities, and the distribution of them over three turns, has several consequences for the organization of social life. First, the localization of repair within a finite, and relatively restricted, space ensures that trouble is dealt with swiftly. Second, and related to this, given the systematic relevance of repair, if speakers move through these three positions without any party initiating repair, a shared understanding of the talk is thereby confirmed *en passant*.

Finally, as with sequence organization, the issue of preference is best grasped as a structural property of the organization of talk-in-interaction (rather than being a product of concerns regarding the private desires of the parties). The two preferences observed by Schegloff, Sacks, and Jefferson are a product of the distribution of opportunities to initiate and effect repair that systematically favors the speaker of the trouble source over

others. As Sacks, Schegloff, and Jefferson ([1974] 1978, p. 40) put it, the organization of turn taking and the organization of repair "are thus 'made for each other' in a double sense." It is worth noting in this regard that insofar as interaction provides the primary site for the achievement of intersubjectivity, for what makes sociality possible, the organization of repair constitutes its last line of defense (Schegloff 1994).

Taken together, the operation of the turn-taking system and the practices involved in the organization of sequences and repair account for many of the detectable, orderly features of conversation. This orderliness was shown to be locally organized and managed, the product of members' methods. It will be useful to make note once again of the research strategy that enabled such findings. Because the data consisted of recordings of naturally occurring activity, a scientific account of the phenomenon under investigation could be empirically grounded in the details of actual occurrences. The investigation began with a set of observable outcomes of these occurrences—in the case of turn taking, for example, speaker change overwhelmingly recurred; overwhelmingly, one party talked at a time; turn order, size, and content were not fixed, but varied; and so on. It was then asked: Could these outcomes be described as products of certain social organized practices, of methods of conduct? At the same time, if members of society did in fact use such formal methods, how were they systematically employed to produce just those outcomes, in just those occurrences, in all their specificity? In addressing the problem in this way, then, conversation analysis was able to discover how cardinal forms of social order were locally constituted.

The research on turn taking in conversation has provided one starting point for more recent studies of interaction in "institutional" settings, such as news interviews, doctor-patient and other clinical consultations, courtrooms, plea bargaining sessions, job interviews, and citizen calls to emergency services. In many of these studies, researchers pursued Sacks, Schegloff, and Jefferson's ([1974] 1978, pp. 45–47) suggestion that the practices underlying the management of ordinary conversation are the "base" or primary ones (for an example, see Heritage and Maynard in press).

Other forms of interaction—in this case, so-called “institutional” forms—are in part constituted and recognizable through systematic variations from conversational turn taking, or through the narrowing and respecification of particular conversational practices involved in the organization of sequences, repair, and other activities.

Take the case of courtroom interaction. The turn-taking system operative in these encounters places restrictions on turn construction and allocation: Coparticipants ordinarily restrict themselves to producing turns that are at least minimally recognizable as “questions” and “answers,” and these turn types are pre-allocated to different parties rather than locally determined. The relatively restricted patterns of conduct observable in these settings is, in large part, the product of this form of turn taking. Accordingly, variation in turn taking in such settings has been shown to have a “pervasive influence both on the range and design of the interactional activities which the different parties routinely undertake and on the detailed management of such encounters” (Heritage 1987, p. 261; Atkinson and Drew 1979).

Note that throughout the above discussion, the term “institutional” has been presented with quotation marks around it. This was done to emphasize ethnomethodology’s preoccupation with the local production of social order. From this view, that some activity or encounter is recognizably either an “ordinary conversation” or more “institutional” in nature—for example, is recognizably a “cross-examination,” a “call to the police,” a “clinical consultation,” or whatever—is something that the coparticipants can and do realize, procedurally, at each and every moment of the encounter. The task for the analyst is to demonstrate how they actually do this; how, for example, they construct their conduct, turn by turn, so as to progressively constitute and thus jointly and collaboratively realize the occasion of their encounter, together with their own social roles in it, as having some distinctively institutional sense (Heritage and Greatbatch 1991). Conversation analytic research on “institutional” interaction has therefore undertaken, through its investigations into the methodic practices by which this gets done, a systematic study of a wide range of human activities.

This mode of research, with its commitment to understanding precisely how any activity becomes what it recognizably and accountably is—that is to say, how it acquires its social facticity—has tended to focus in the 1990s on work activities and settings, under the rubric of “workplace studies.” The scope of investigation has expanded to encompass all forms of “embodied action” (that is, not only the talk), with extensive use of video recordings and, influenced by Suchman’s (1987) pioneering study of human-machine interaction, with careful attention to how the machines, technologies, and other artifacts that saturate the modern work site are taken up and enter into the endogenous organization of work tasks (see, for example, Whalen (1995), and the papers collected in Luff, Hindmarsh, and Heath in press).

Research into conversation’s organization also continues to evolve. While there has been relatively little work that attempts to fundamentally deepen the original account of the turn-taking system developed by Sacks, Schegloff, and Jefferson (for a notable exception see Lerner 1996), there has been important research at the intersection of grammar and interaction—recognizing that talk-in-interaction is in fact the natural home of human language. This work demonstrates that the approach to language taken Chomsky’s Transformational Grammar could be supplanted by one based on naturalistic study of the “grammar for conversation.” Given that Chomsky’s approach has decisively shaped, both directly and indirectly, the understanding of language in cognitive science, psychology, computational linguistics, and the other disciplines that rely on a model of grammatical organization for their own research, these findings are plainly significant.

Conversation analytic work on grammar and interaction was launched by Schegloff’s (1979) paper on the “relevance of repair for a syntax for conversation.” This line of work has underscored the need for studies of language to draw on naturally occurring spates of talk. As Schegloff observed, while nearly every episode of ordinary talk contains instances of repair within the “sentences” (or sentential turn constructional units) out of which it is built, the entire view of language developed by linguists is based on imagined (or what might as well be imagined) instances of language

that are free of such repair. Schegloff went on to show that most instances of talk-in-interaction, at least in English, are organized by reference to the systematic relevance of repair, whether an instance of it actually occurs in the sentence or not.

Of course repair is not the only organization relevant for grammar, and so more recently scholars have begun to examine what more might be learned about language by studying it as produced in naturally occurring interaction. With respect to this problem, conversation analysts have argued that insofar as language most likely evolved in face-to-face encounters by members of our species, its structure and organization must have evolved, at least in part, to manage the basic exigencies confronted by speakers and hearers. Thus, in addition to the systematic relevance of repair, the structure and organization of grammar most likely evolved as resources that shape, and are shaped by, how opportunities to speak are distributed, what constraints are introduced by a current turn on subsequent ones, and how speakers' formulations of the events, persons, and objects are organized. Perhaps most developed are a series of findings that link the organization of grammar and the system for distributing turns at talk briefly described above.

As we stated earlier regarding turn-constructive units, one of their key features is that each sentence, or utterance, projects from its beginning roughly what it will take for it to be possibly complete. And over its course, each utterance projects in finer and finer detail the exact moment that a speaker may end her utterance. Thus, instead of expressing logical predicates or cognitive states, grammar may be best understood, in the first instance, as a sequentially sensitive resource that progressively projects the course and duration of turns at talk (Ford and Thompson 1996).

One of the most striking consequences of such a view of grammar is that the locus of its organization is transformed. While most approaches to grammar rely on the sentence as the basic unit of organization (with occasional nods to the organization of "discourse"), the grammatical units produced in interaction are fundamentally organized relative to their sequential environment, most proximally the just prior, current, and next turns. Thus, rather than the sentence, or even discourse, being the fundamental unit or environment of

analysis, interaction and sequences of turns appear to be that within which grammar is most proximally organized. This appears to be true even at levels beneath the turn whether a sentence, clause, or phrase. As Schegloff (1996) shows, turn beginnings and turn endings, as well as what happens in between, are sites of strategic manipulation. Through this manipulation, both grammatical and prosodic, speakers fit their utterances to prior talk, launch new actions, and shape when they will be heard as possibly complete. Any scientific analysis of language, then, must take into account this central function.

Thus, as the collection of papers assembled in Ochs, Schegloff, and Thompson (1996) suggest, rather than viewing grammar as an independent, clearly delineated, and internally coherent structure, it is best approached as one more of the interrelated set of resources through which interaction, and social life more broadly, is organized.

## REFERENCES

- Atkinson, J. Maxwell, and Paul Drew 1979 *Order in Court: The Organization of Verbal Interaction in Judicial Settings*. London: Macmillan.
- Atkinson, J. Maxwell, and John Heritage (eds.) 1984 *Structures of Social Action: Studies in Conversation Analysis*. Cambridge, Eng.: Cambridge University Press.
- 1984 "Introduction." In J. Maxwell Atkinson and John Heritage, eds., *Structures of Social Action: Studies in Conversation Analysis*. Cambridge, Eng.: Cambridge University Press.
- Bales, Robert F. 1950 *Interaction Process Analysis*. Cambridge, Mass.: Addison-Wesley.
- Button, Graham, and John R. E. Lee (eds.) 1987 *Talk and Social Organization*. Clevedon, Eng.: Multilingual Matters.
- Davidson, Judy 1984 "Subsequent Versions of Invitations, Offers, Requests, and Proposals Dealing with Potential or Actual Rejection." In J. Maxwell Atkinson and John Heritage, eds., *Structures of Social Action: Studies in Conversation Analysis*. Cambridge, Eng.: Cambridge University Press.
- Ford, Cecilia E., and Sandra A. Thompson 1996 "Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns." In Elinor Ochs, Emanuel A. Schegloff and Sandra A. Thompson, eds., *Interaction*

- and Grammar*. Cambridge, Eng.: Cambridge University Press.
- Garfinkel, Harold 1991 "Respecification: Evidence for Locally Produced, Naturally Accountable Phenomena of Order, Logic, Reason, Meaning, Method, etc. in and as of the Essential Haecceity of Immortal Ordinary Society (I)—An Announcement of Studies." In Graham Button, ed., *Ethnomethodology and the Human Sciences*. Cambridge, Eng.: Cambridge University Press.
- (1967) 1984 *Studies in Ethnomethodology*. Cambridge: Polity Press.
- Goodwin, Charles 1994 "Professional Vision." *American Anthropologist* 96:606–633.
- 1981 *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- Heath, Christian C. 1986 *Body Movement and Speech in Medical Interaction*. Cambridge, Eng.: Cambridge University Press.
- Heritage, John 1987 "Ethnomethodology." In Anthony Giddens and Jonathan H. Turner, eds., *Social Theory Today*. Stanford, Calif.: Stanford University Press.
- , and David Greatbatch 1991 "On the Institutional Character of Institutional Talk: The Case of News Interviews." In Deirdre Boden and Don H. Zimmerman, eds., *Talk and Social Structure*. Cambridge: Polity Press.
- Heritage, John, and Douglas W. Maynard (eds.) In press *Practicing Medicine: Talk and Action in Primary Care Encounters*. Cambridge, Eng.: Cambridge University Press.
- Homans, George C. 1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace.
- Lerner, Gene H. 1996 "On the 'Semi-Permeable' Character of Grammatical Units in Conversation: Conditional Entry into the Turn Space of Another Speaker." In Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, eds., *Interaction and Grammar*. Cambridge, Eng.: Cambridge University Press.
- Luff, Paul, John Hindmarsh, and Christian Heath (eds.) In press *Workplace Studies: Recovering Work Practice and Informing Systems Design*. Cambridge, Eng.: Cambridge University Press.
- Maynard, Douglas W. 1997 "The News Delivery Sequence: Bad News and Good News in Conversational Interaction." *Research on Language and Social Interaction* 30:93–130.
- Ochs, Elinor, Emanuel A. Schegloff, and Sandra A. Thompson (eds.) 1996 *Interaction and Grammar*. Cambridge, Eng.: Cambridge University Press.
- Sacks, Harvey 1992. *Lectures on Conversation*. 2 vols. Oxford: Blackwell.
- (1964–1968) 1984 "Notes on Methodology." In J. Maxwell Atkinson and John Heritage, eds., *Structures of Social Action: Studies in Conversation Analysis*. Cambridge, Eng.: Cambridge University Press.
- 1972 "An Initial Investigation into the Usability of Conversational Data for Doing Sociology." In David Sudnow, ed., *Studies in Social Interaction*. New York: Free Press.
- , Emanuel A. Schegloff, and Gail Jefferson (1974) 1978 "A Simplest Systematics for the Organization of Turn-Taking for Conversation." In Jim Schenkein, ed., *Studies in the Organization of Conversational Interaction*. New York: Academic Press.
- Schegloff, Emanuel A. 1996 "Turn Organization: One Intersection of Grammar and Interaction." In Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, eds., *Interaction and Grammar*. Cambridge, Eng.: Cambridge University Press.
- 1989 "Harvey Sacks—Lectures 1964–1965. An Introduction/Memoir." *Human Studies* 12:185–209.
- 1988 "Goffman and the Analysis of Conversation." In Paul Drew and Anthony Wootton, eds., *Erving Goffman: Exploring the Interaction Order*. Boston: Northeastern University Press.
- 1979 "The Relevance of Repair for Syntax-for-Conversation." In T. Givon, ed., *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press.
- 1968 "Sequencing in Conversational Openings." *American Anthropologist* 70:1075–1095.
- , Gail Jefferson, and Harvey Sacks 1977 "The Preference for Self-Correction in the Organization of Repair in Conversation." *Language* 53:361–382.
- Schegloff, Emanuel A., and Harvey Sacks 1973 "Opening Up Closings." *Semiotica* 7:289–327.
- Schenkein, Jim (ed.) 1978 *Studies in the Organization of Conversational Interaction*. New York: Academic Press.
- Silverman, David 1998 *Harvey Sacks: Social Science and Conversation Analysis*. New York: Oxford University Press.
- Suchman, Lucy A. 1987 *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge, Eng.: Cambridge University Press.
- Sudnow, David (ed.) 1972 *Studies in Social Interaction*. New York: Free Press.
- ten Have, Paul, and George Psathas 1995 *Situated Order: Studies in the Social Organization of Talk and Embodied Activities*. Washington, D.C.: University Press of America.
- Whalen, Jack 1995 "A Technology of Order Production: Computer-Aided Dispatch in 9-1-1 Communications." In Paul ten Have and George Psathas, eds.,

*Situated Order: Studies in the Social Organization of Talk and Embodied Activities.* Washington, D.C.: University Press of America.

JACK WHALEN  
GEOFF RAYMOND

## COOPERATION AND COMPETITION

See Small Groups.

## CORPORATE ORGANIZATIONS

**NOTE:** Although the following article has not been revised for this edition of the *Encyclopedia*, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

Societies carry out many of their activities through formal organizations. Organizations are units in which offices, or positions, have distinct but interdependent duties. Organizations—hospitals, schools, governments, business firms—share certain features. Usually, at least one of the offices serves as the linchpin: It coordinates the separate duties within the organization. The key office has ultimate authority in that the orders it issues constrain the actions of lower-level offices.

But organizations also differ from one another. In some, the assets belong to particular individuals. In others, ownership resides in a collectivity. The latter represents a corporate organization or *corporation*. Three features describe the modern corporation. First, it has certain legal rights and privileges. By law, a corporation can sue and be sued in the courts, make contracts, and purchase and receive property. Second, it usually exists in perpetuity: It outlasts the individuals who set it up. Ownership rests with stockholders, whose numbers and makeup can change from one time to another. Third, the owners have only a limited responsibility for the obligations the corporation makes.

These features distinguish the corporate organization from two other forms of ownership: the

*proprietorship* and the *partnership*. In a proprietorship a particular person owns the property of the organization; in a partnership, two or more persons share it. The right to handle the property and affairs of the organization rests with a designated proprietor or set of partners. Significantly, proprietors and partners bear personal responsibility for the debts of the organization.

The corporation constitutes a social invention. The form evolved to handle problems that arose within religious, political, and other kinds of communities. It holds a place of importance in contemporary Western societies. Because it is the product of social conditions and an influence on them, the corporation represents a topic of substantial interest in sociology.

At present, the corporation appears commonly within the world of business. But when the corporation began to take shape during the Middle Ages, the questions to be resolved lay outside that realm. One of these questions had to do with church ownership. In medieval Germany, landowners often set up churches on their estates and placed a priest in charge of them. As priests gained authority over their charges, they argued that the church and the land surrounding it no longer belonged to the donor. Deciding the true owner proved to be difficult. A given priest could die or be replaced; hence, any particular priest seemed to have no claim to ownership. One practice regarded the owner to be the saint whose name the church bore. Eventually, the idea developed that ownership inhered in the church, and that the church constituted a body independent of its current leaders or members (Coleman 1974; Stone 1975).

Thorny problems also arose as medieval settlements formed into towns. A town required someone to manage its affairs such as collecting tolls and transacting other business. But the laws that prevailed at the time applied only to individuals. Any actions individuals took obligated them personally. By this principle, managers would have to meet any commitments they made on behalf of the town. To eliminate the dilemmas that the principle situation posed, new laws made the town a corporate person. The corporate person would have all the rights and privileges of any human being. This action reduced the risks that public service might otherwise entail. For many of the same reasons

that the church and town became corporate persons, the university of the Middle Ages moved towards the corporate form.

The early corporations played rather passive roles. Essentially, they held property for a collective, whose members might change from time to time. Contrastingly, the corporations of the twentieth century constitute spirited forces. They hire multitudes of employees. They produce goods and services and mold ideals and tastes. The decisions their leaders make about where to locate often determine which locales will prosper and which will languish.

The influence that corporations have produces concerns about the control of them. Much of the work on corporations that sociologists have undertaken highlights these concerns. The work on control and corporations covers three topics: the means through which corporations control their employees; the allocation of control between owners and managers; and the extent to which societies control corporations. For all three topics, control implies command over the affairs of and operations within the corporate organization.

### **CONTROL OVER EMPLOYEES**

The corporate form has a long history, yet it did not typify the early factories that manufacturers established in the United States. Before the early 1900s, most factories operated as small operations under the control of a single entrepreneur. The entrepreneur hired an overseer who might in turn choose a foreman to hire, discipline, and fire workers. Through consolidation and merger, the economic landscape of the 1920s revealed far more large organizations than had the tableau of a half-century earlier.

More changed over the years of the late nineteenth and early twentieth centuries than just the size of organizations. The corporate form spread; the faceless corporation replaced the corporeal entrepreneur. Corporations moved towards professional management. Factories that businessmen once controlled personally now operated through abstract rules and procedures. The people whom the workers now contacted on a regular basis consisted of staff for the corporation and not the corporate owners themselves. Bureaucratic tenets took root.

A bureaucracy constitutes a particular mode that organizations can take. Consistent with all organizations, bureaucratic ones divide up duties. Two features separate a bureaucracy from other modes, however. First, a system of ranks or levels operates. Second, fixed rules and procedures govern actions. The rules define the tasks, responsibilities, and authority for each office and each level.

Few of the factories in nineteenth-century America operated as bureaucracies. Instead, the individuals who made the products decided how the work would be done. A minimum number of levels existed. Supervisors or foremen hired and fired workers, but workers made the rules on the work itself. The workers were craftsmen or artisans, and they contended that only those who possessed the skills that the work demands should decide how or if it should be divided. Gradually, machinery took over the skilled work. Machines and not workers controlled the pace. By the end of the 1920s, neither the laborers nor the machinery shaped the work. Professional managers did. These managers enforced rules and oversaw an organization where specialized tasks and graded authority prevailed (Nelson 1975; Clawson 1980; and Jacoby 1985).

The corporation of the late twentieth century continues to operate as a bureaucracy. Some sources argue that efficiency explains the adoption of the bureaucratic model (see especially Chandler 1980, 1984). Others challenge the emphasis on efficiency, charging it with being overly rational or too apolitical. The first challenge appears most notably in the work on organizations as institutions. This literature regards survival as the premier goal for any organization. The closer an organization approximates an institution—an element taken for granted in the society—the greater its chances for survival.

According to the institutional perspective, organizations adopt practices that appear to be reasonable. Myths develop about which patterns prove most useful and efficient, and any organization that does not adopt a pattern that the myth favors courts failure (Meyer and Rowan 1977; DiMaggio 1988; DiMaggio and Powell 1983; Tolbert and Zucker 1983; also see Scott 1987 for a review of the different branches of institutional theory).

A different argument maintains that the emphasis on efficiency fails to capture the politics of

corporations. This perspective treats corporations as systems in which the interests of owners clash with those of workers. Owners, it asserts, seek to reduce uncertainties and to eliminate the vagaries that can plague organizations. From this angle, bureaucracy serves the interests of owners primarily because it reduces the influence that workers exercise and thereby removes a source of uncertainty (Braverman 1974; Edwards 1979).

Workers need not have formal authority in order to affect outcomes within organizations. Studies document the creative ways in which employees enliven monotonous jobs and pursue their own ends (Roy 1952; Mechanic 1962; Burawoy 1979, 1985). Yet, officially, the higher levels have greater power than have the lower levels. This is the consequence of the bureaucratic nature of corporations, not of their pattern of ownership. The bureaucratic mode is not unique to corporations. Proprietorships and partnerships can display the traits of bureaucracy. The diffuseness of ownership that one finds in the corporation possibly makes formal control less obvious than obtains when ownership resides in identifiable persons.

#### **OWNER VERSUS MANAGERIAL CONTROL**

Managers occupy important places in the contemporary organization. One argument regards managers as more powerful than stockholders. Adolph Berle and Gardiner Means offered this argument in the 1930s. As Berle and Means saw the situation, stockholding had become too widely dispersed for any individual holder or even group of holders to command corporations. Managers, they contended, filled the void (Berle and Means 1932). Later discussions echoed the thesis that the expansion of the corporate form had raised the power of corporate managers (Berg and Zald 1978; also see Chandler 1962, 1977).

Critics contend that the thesis overstates the role and power of managers. They base their criticism on studies of the influence that corporate leaders wield. Maurice Zeitlin (1974) helped launch this line of research when he argued that few scholars had tested the Berle and Means thesis and that the handful of extant studies showed owners to be less fractious and fractionated than the thesis supposed. Michael Useem (1984), among others,

heeded the call from Zeitlin for research on the networks that link shareholders. Useem concluded from his study on contacts and networks among large shareholders that a corporate community operated, held together by an inner circle whose interests transcended company, region, and industry lines. Beth Mintz and Michael Schwartz (1985) examined the connections between financial institutions and other corporations and decided that control over corporate directions rested disproportionately in the world of finance. The work from the critics cautions us against the assumption that a multiplicity of owners implies control by managers.

#### **SOCIAL CONTROL OVER CORPORATIONS**

The corporate form constitutes a remarkable innovation. But as the corporation has become ever more active and entrenched, it has generated problems for society. Corporations have at times engaged in criminal behavior (Sutherland 1949; Clinard and Yeager 1980). At other times, their actions have violated no law but have put the well-being of the public at risk. Both situations often show the inadequacy of the mechanisms through which society attempts to control corporations.

Corporations are creatures of the state. Ostensibly, then, they operate only at the indulgence of the state. But myriad corporations now have greater resources than do the states that chartered them. Moreover, the laws that states have at their disposal often fit individuals better than they do corporations. Corporations can be sued for wrongdoing; but a fine that would bankrupt an individual might be a mere pittance for a large corporation. Both James Coleman (1974) and Christopher Stone (1975) have argued that the law can never be the sole means for controlling corporations; a sense of responsibility to the public must prevail within corporations.

Even if the law were shown to be effective in constraining corporations within a state, it might prove rather impotent in the case of multinational organizations. A multinational or transnational corporation holds a charter from one nation-state but transacts business in at least one other. The governmental entity that issues the charter cannot alter the policies the corporation pursues in its

other locales. In addition, the very size of many multinationals restricts the pressure that either the home or the host country can impose.

Through various actions corporations demonstrate that they are attentive to the societies they inhabit. Corporate leaders serve on the boards of social service agencies; corporate foundations provide funds for community programs; employees donate their time to local causes. The agenda of corporations long have included these and similar activities. Increasingly, the agenda organize such actions around the idea of corporate social responsibility. Acting responsibly means taking steps to promote the commonweal (Steckmest 1982).

Some corporations strive more consistently to advance social ends than do others. Differences in norms and values apparently explain the contrast. Norms, or maxims for behavior, indicate the culture of the organization (Deal and Kennedy 1982). The culture of some settings gives the highest priority to actions that protect the health, safety, and welfare of citizens and their heirs. Elsewhere, those are not what the culture emphasizes (Clinard 1983; Victor and Cullen 1988).

The large corporation had become such a dominant force by the 1980s that no one envisioned a return to an era of small, diffuse organizations. Yet, during that decade some sectors had started to move from growth to contraction. At times, the shift resulted from legislative action. When the Bell Telephone System divided in 1984, by order of the courts, the change marked a sharp reversal. For more than a century the system had glided toward integration and standardization (Barnett and Carroll 1987; Barnett 1990).

Whether through fiat or choice, corporations contract (Whetten 1987; Hambrick and D'Aveni 1988). Two perspectives associate the rise and fall in the fortunes of corporations to changes in the social context. The first perspective, resource dependence, centers on the idea that organizations must secure their resources from their environs (McCarthy and Zald 1977; Jenkins 1983). When those environs contain a wealth of resources—personnel in the numbers and with the qualifications the organization requires, funds to finance operations—the corporation can thrive. When hard times plague the environs, the corporation escapes that fate only with great difficulty.

The perspective known as population ecology likewise connects the destiny of organizations to conditions in their surroundings. Population ecologists think of organizations as members of a population. Changing social conditions can enrich or impoverish a population. Individual units within it can do little to offset the tide of events that threatens to envelop the entire population. (Hannan and Freeman 1988; Wholey and Brittain 1989; for a critique of the approach see Young 1988).

Neither resource dependency theory nor population ecology theory focuses explicitly on the corporate form. But just as analyses of corporations inform the discussions sociologists have undertaken on formal organizations, models drawn from studies of organizations have proved useful as scholars have tracked the progress of corporations.

The corporation clearly constitutes a power to be reckoned with. As with its precursors, the modern corporation serves needs that collectivities develop. In fact, the corporation rests on an assumption that is fundamental in sociology: A collectivity has an identity of its own. But the corporation of the twentieth century touches more than those persons who own its assets or produce its goods. This social instrument of the Middle Ages is now a social fixture.

(SEE ALSO: *Capitalism; Organizational Effectiveness; Organizational Structure; Transnational Corporations*)

#### REFERENCES

- Antai, Ariane-Berthoin 1992 *Corporate Social Performance: Rediscovering Actors in Their Organizational Contexts*. Boulder, Colo.: Westview Press.
- Barnett, William P. 1990 "The Organizational Ecology of a Technological System." *Administrative Science Quarterly* 35:31-60.
- Barnett, William P., and Glenn Carroll 1987 "Competition and Mutualism Among Early Telephone Companies." *Administrative Science Quarterly* 32:400-421.
- Berg, Ivar, and Mayer N. Zald 1978 "Business and Society." *Annual Review of Sociology* 4:115-143.
- Berle, Adolph A., and Gardiner C. Means 1932 *The Modern Corporation and Private Property*. New York: Macmillan.
- Braverman, Harry 1974 *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review.



- Breedlove, William L., and J. Michael Armer 1996 "Economic Disarticulation and Social Development in Less-Developed Nations: A Cross-National Study of Intervening Structures." *Sociological Focus* 29:359-378.
- Burawoy, Michael 1979 *Manufacturing Consent: Changes in the Labor Process under Monopoly Capitalism*. Chicago: University of Chicago Press.
- 1985 *The Politics of Production: Factory Regimes under Capitalism and Socialism*. London: Verso.
- Chandler, Alfred D., Jr. 1962 *Strategy and Structure*. Cambridge, Mass.: MIT Press.
- 1977 *The Visible Hand*. Cambridge, Mass.: Harvard University Press.
- 1980 "The United States: Seedbed of Managerial Capitalism." In A. D. Chandler, Jr., and Herman Daems, eds., *Managerial Hierarchies*. Cambridge, Mass.: Harvard University Press.
- 1984 "The Emergence of Managerial Capitalism." *Business History Review* 58:484-504.
- Clawson, Dan 1980 *Bureaucracy and the Labor Process*. New York: Monthly Labor Review.
- Clinard, Marshall B. 1983 *Corporate Ethics and Crime*. Beverly Hills, Calif.: Sage.
- , and Peter C. Yeager 1980 *Corporate Crime*. New York: Free Press.
- Coleman, James S. 1974 *Power and the Structure of Society*. New York: Norton.
- Coleman, James S. 1996 "The Asymmetric Society: Organizational Actors, Corporate Power, and the Irrelevance of Persons." In M. David Ermann, and Richard J. Lundman, eds., *Corporate and Governmental Deviance: Problems of Organizational Behavior in Contemporary Society*. New York: Oxford University Press.
- Cuthbert, Alexander R., and Keith G. McKinnell 1997 "Ambiguous Space, Ambiguous Rights-Corporate Power and Social Control in Hong Kong." *Cities* 14: 295-311.
- Davis, Gerald F. 1994 "The Corporate Elite and the Politics of Corporate Control." *Current Perspectives in Social Theory* 1: 215-238.
- Davis, Gerald F., Kristina A. Diekmann, and Catherine H. Tinsley. 1994. "The Decline and Fall of the Conglomerate Firm in the 1980s: The Deinstitutionalization of an Organizational Form." *American Sociological Review* 59: 547-570.
- Deal, Terrence, and Allen Kennedy 1982 *Corporate Culture*. Reading, Mass.: Addison-Wesley.
- DiMaggio, Paul 1988 "Interest and Agency in Institutional Theory." In Lynne G. Zucker, ed., *Research on Institutional Patterns: Environment and Culture*. Cambridge, Mass.: Ballinger.
- , and Walter Powell 1983 "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147-160.
- Donaldson, Thomas 1992 "Individual Rights and Multinational Corporate Responsibilities." *National Forum* 72: 7-9.
- Dunford, Louise, and Ann Ridley 1996 "No Soul to Be Damned, No Body to Be Kicked: Responsibility, Blame and Corporate Punishment." *International Journal of the Sociology of Law* 24:1-19.
- Dunworth, Terence, and Joel Rogers 1996 "Corporations in Court: Big Business Litigation in U.S. Federal Courts, 1971-1991." *Law and Social Inquiry* 21: 497-592.
- Edwards, Richard 1979 *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*. New York: Basic Books.
- Etzioni, Amitai 1993 "The U.S. Sentencing Commission on Corporate Crime: A Critique." *Annals of the American Academy of Political and Social Science* 525: 147-156.
- Hambrick, Donald C., and Richard A. D'Aveni 1988 "Large Corporate Failures as Downward Spirals." *Administrative Science Quarterly* 33:1-23.
- Hannan, Michael, and John Freeman 1988 *Organizational Ecology*. Cambridge, Mass.: Harvard University Press.
- Hawley, James P. 1995 "Political Voice, Fiduciary Activism, and the Institutional Ownership of U.S. Corporations: The Role of Public and Noncorporate Pension Funds." *Sociological Perspectives* 38:415-435.
- Ireland, Paddy 1996 "Corporate Governance, Stakeholding, and the Company: Towards a Less Degenerate Capitalism?" *Journal of Law and Society* 23:287-320.
- Jacoby, Sanford 1985 *Employing Bureaucracy: Managers, Unions and the Transformation of Work in American Industry, 1900-1945*. New York: Columbia University Press.
- Jenkins, J. Craig 1983 "Resource Mobilization Theory and the Study of Social Movements." *Annual Review of Sociology* 42:249-268.
- Jones, Marc T. 1996 "The Poverty of Corporate Social Responsibility." *Quarterly Journal of Ideology* 19:1-2.
- Lofquist, William S. 1993 "Legislating Organizational Probation: State Capacity, Business Power, and Corporate Crime Control." *Law and Society Review* 27: 741-783.
- Luchansky, Bill, and Gregory Hooks 1996 "Corporate Beneficiaries of the Mid-Century Wars: Respecifying

- Models of Corporate Growth, 1939–1959.” *Social Science Quarterly* 77:301–313.
- McCarthy, John D., and Mayer N. Zald 1977 “Resource Mobilization and Social Movements: A Partial Theory.” *American Journal of Sociology* 82:1212–1241.
- McKinlay, Alan, and Ken Starkey 1998 *Foucault, Management and Organizations Theory: From Panopticon to Technologies of Self*. London: Sage Publications.
- Mechanic, David 1962 “Sources of Power of Lower Participants in Complex Organizations.” *Administrative Science Quarterly* 7:349–364.
- Meyer, John, and Brian Rowan 1977 “Institutionalized Organizations: Formal Structure as Myth and Ceremony.” *American Sociological Review* 42:340–363.
- Mintz, Beth, and Michael Schwartz 1985 *The Power Structure of American Business*. Chicago: University of Chicago Press.
- Nelson, Daniel 1975 *Managers and Workers*. Madison: University of Wisconsin Press.
- Parker-Gwin, Rachel, and William G. Roy 1996 “Corporate Law and the Organization of Property in the United States: The Origin and Institutionalization of New Jersey Corporation Law, 1888–1903.” *Politics and Society* 24:111–135.
- Prechel, Harland 1994 “Economic Crisis and the Centralization of Control over the Managerial Process: Corporate Restructuring and Neo-Fordist Decision-Making.” *American Sociological Review* 59:723–745.
- Ridley, Ann, and Louise Dunford 1994 “Corporate Liability for Manslaughter: Reform and the Art of the Possible.” *International Journal of the Sociology of Law* 22:309–328.
- Rodman, Kenneth A. 1995 “Sanctions at Bay? Hegemonic Decline, Multinational Corporations, and U.S. Economic Sanctions since the Pipeline Case.” *International Organization* 49:105–137.
- Roy, Donald 1952 “Restriction of Output in a Piecework Machine Shop.” Ph.D. diss., University of Chicago, Chicago.
- Sanders, Joseph, and V. Lee Hamilton 1996 “Distributing Responsibility for Wrongdoing inside Corporate Hierarchies: Public Judgments in Three Societies.” *Law and Social Inquiry* 21:815–855.
- Scott, W. Richard 1987 “The Adolescence of Institutional Theory.” *Administrative Science Quarterly* 32:493–511.
- Steckmest, Francis W. 1982 *Corporate Performance: The Key to Public Trust*. New York: McGraw-Hill.
- Stone, Christopher D. 1975 *Where the Law Ends: The Social Control of Corporate Behavior*. New York: Harper.
- Sutherland, Edwin H. 1949 *White Collar Crime*. New York: Holt, Rinehart, and Winston.
- Tolbert, Pamela S., and Lynne G. Zucker 1983 “Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Service Reform, 1880–1935.” *Administrative Science Quarterly* 28:22–39.
- Useem, Michael 1984 *The Inner Circle: Large Corporations and the Rise of Political Activity in the U.S. and U.K.* New York: Oxford University Press.
- Victor, Bart, and John B. Cullen 1988 “The Organizational Bases of Ethical Work Climates.” *Administrative Science Quarterly* 33:101–125.
- Westphal, James D., and Edward J. Zajac 1997 “Defections from the Inner Circle: Social Exchange, Reciprocity, and the Diffusion of Board Independence in U.S. Corporations.” *Administrative Science Quarterly* 42:161–183.
- Whetten, David A. 1987 “Organizational Growth and Decline Processes.” *Annual Review of Sociology* 13:335–358.
- Wholey, Douglas, and Jack Brittain 1989 “Characterizing Environmental Variation.” *Academy of Management Journal* 32:867–882.
- Wilks, Stephen 1992 “Science, Technology and the Large Corporation.” *Government and Opposition* 27:190–212.
- Young, Ruth 1988 “Is Population Ecology a Useful Paradigm for the Study of Organizations?” *American Journal of Sociology* 94:1–24.
- Zeitlin, Maurice 1974 “Corporate Ownership and Control: The Large Corporation and the Capitalist Class.” *American Journal of Sociology* 79:1073–1119.
- Zey, Mary, and Brande Camp 1996 “The Transformation from Multidimensional Form to Corporate Groups of Subsidiaries in the 1980s: Capital Crisis Theory.” *Sociological Quarterly* 37:327–351.

CORA B. MARRETT

## CORRECTIONS SYSTEMS

See Penology; Criminal Sanctions; Criminology.

## CORRELATION AND REGRESSION ANALYSIS

In 1885, Francis Galton, a British biologist, published a paper in which he demonstrated with graphs and tables that the children of very tall

parents were, on average, shorter than their parents, while the children of very short parents tended to exceed their parents in height (cited in Walker 1929). Galton referred to this as “reversion” or the “law of regression” (i.e., regression to the average height of the species). Galton also saw in his graphs and tables a feature that he named the “co-relation” between variables. The stature of kinsmen are “co-related” variables, Galton stated, meaning, for example, that when the father was taller than average, his son was likely also to be taller than average. Although Galton devised a way of summarizing in a single figure the degree of “co-relation” between two variables, it was Galton’s associate Karl Pearson who developed the “coefficient of correlation,” as it is now applied. Galton’s original interest, the phenomenon of regression toward the mean, is no longer germane to contemporary correlation and regression analysis, but the term “regression” has been retained with a modified meaning.

Although Galton and Pearson originally focused their attention on bivariate (two variables) correlation and regression, in current applications more than two variables are typically incorporated into the analysis to yield partial correlation coefficients, multiple regression analysis, and several related techniques that facilitate the informed interpretation of the linkages between pairs of variables. This summary begins with two variables and then moves to the consideration of more than two variables.

Consider a very large sample of cases, with a measure of some variable,  $X$ , and another variable,  $Y$ , for each case. To make the illustration more concrete, consider a large number of adults and, for each, a measure of their education (years of school completed =  $X$ ) and their income (dollars earned over the past twelve months =  $Y$ ). Subdivide these adults by years of school completed, and for each such subset compute a mean income for a given level of education. Each such mean is called a *conditional mean* and is represented by  $\bar{Y}|X$ , that is, the mean of  $Y$  for a given value of  $X$ .

Imagine now an ordered arrangement of the subsets from left to right according to the years of school completed, with zero years of school on the left, followed by one year of school, and so on through the maximum number of years of school completed in this set of cases, as shown in Figure 1.

Assume that each of the  $\bar{Y}|X$  values (i.e., the mean income for each level of education) falls on a straight line, as in Figure 1. This straight line is the *regression line of  $Y$  on  $X$* . Thus the regression line of  $Y$  on  $X$  is the line that passes through the mean  $Y$  for each value of  $X$ —for example, the mean income for each educational level.

If this regression line is a straight line, as shown in Figure 1, then the income associated with each additional year of school completed is the same whether that additional year of school represents an increase, for example, from six to seven years of school completed or from twelve to thirteen years. While one can analyze curvilinear regression, a straight regression line greatly simplifies the analysis. Some (but not all) curvilinear regressions can be made into straight-line regressions by a relatively simple transformation of one of the variables (e.g., taking a logarithm). The common assumption that the regression line is a straight line is known as the *assumption of rectilinearity*, or more commonly (even if less precisely) as the *assumption of linearity*.

The slope of the regression line reflects one feature of the relationship between two variables. If the regression line slopes “uphill,” as in Figure 1, then  $Y$  increases as  $X$  increases, and the steeper the slope, the more  $Y$  increases for each unit increase in  $X$ . In contrast, if the regression line slopes “downhill” as one moves from left to right,  $Y$  *decreases* as  $X$  increases, and the steeper the slope, the more  $Y$  decreases for each unit increase in  $X$ . If the regression line doesn’t slope at all but is perfectly horizontal, then there is no relationship between the variables. But the slope does not tell how closely the two variables are “co-related” (i.e., how closely the values of  $Y$  cluster around the regression line).

A regression line may be represented by a simple mathematical formula for a straight line. Thus:

$$\bar{Y}|X = a_{yx} + b_{yx}X \quad (1)$$

where  $\bar{Y}|X$  = the mean  $Y$  for a given value of  $X$ , or the regression line values of  $Y$  given  $X$ ;  $a_{yx}$  = the  $Y$  intercept (i.e., the predicted value of  $\bar{Y}|X$  when  $X = 0$ ); and  $b_{yx}$  = the slope of the regression of  $Y$  on  $X$  (i.e., the amount by which  $\bar{Y}|X$  increases or decreases—depending on whether  $b$  is positive or negative—for each one-unit increase in  $X$ ).

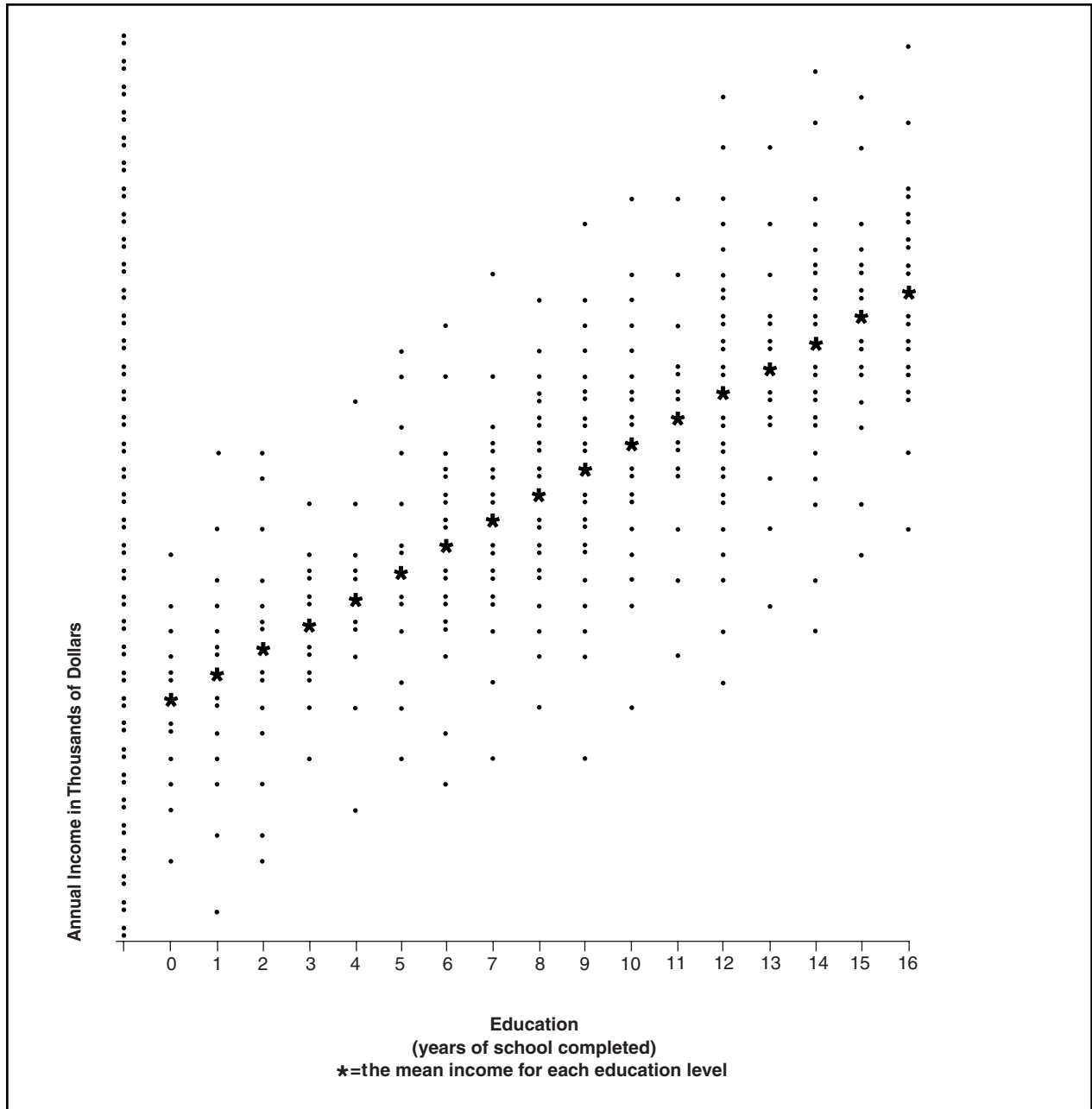


Figure 1. Hypothetical Regression of Income on Education

Equation 1 is commonly written in a slightly different form:

$$\hat{Y} = a_{yx} + b_{yx}X \quad (2)$$

where  $\hat{Y}$  = the regression prediction for Y for a given value of X, and  $a_{yx}$  and  $b_{yx}$  are as defined above, with  $\hat{Y}$  substituted for  $\bar{Y}|X$ .

Equations 1 and 2 are theoretically equivalent. Equation 1 highlights the fact that the points on the regression line are assumed to represent conditional means (i.e., the mean Y for a given X). Equation 2 highlights the fact that points on the regression line are not ordinarily found by computing a series of conditional means, but are found by alternative computational procedures.

Typically the number of cases is insufficient to yield a stable estimate of each of a series of conditional means, one for each level of X. Means based on a relatively small number of cases are inaccurate because of sampling variation, and a line connecting such unstable conditional means may not be straight even though the true regression line is. Hence, one assumes that the regression line is a straight line unless there are compelling reasons for assuming otherwise; one can then use the X and Y values for all cases together to estimate the Y intercept,  $a_{yx}$ , and the slope,  $b_{yx}$ , of the regression line that is best fit by the *criterion of least squares*. This criterion requires predicted values for Y that will minimize the sum of squared deviations between the predicted values and the observed values. Hence, a “least squares” regression line is the straight line that yields a lower sum of squared deviations between the predicted (regression line) values and the observed values than does any other straight line. One can find the parameters of the “least squares” regression line for a given set of X and Y values by computing

$$b_{yx} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2} \quad (3)$$

$$a_{yx} = \bar{Y} - b_{yx}\bar{X} \quad (4)$$

These parameters (substituted in equation 2) describe the straight regression line that best fits by the criterion of least squares. By substituting the X value for a given case into equation 2, one can then find  $\hat{Y}$  for that case. Otherwise stated, once  $a_{yx}$  and  $b_{yx}$  have been computed, equation 2 will yield a precise predicted income level ( $\hat{Y}$ ) for each education level.

These predicted values may be relatively good or relatively poor predictions, depending on whether the actual values of Y cluster closely around the predicted values on the regression line or spread themselves widely around that line. The *variance* of the Y values (income levels in this illustration) around the regression line will be relatively small if the Y values cluster closely around the predicted values (i.e., when the regression line provides relatively good predictions). On the other hand, the variance of Y values around the regression line will be relatively large if the Y values are spread widely around the predicted values (i.e., when the regression line provides relatively poor predictions). The

variance of the Y values around the regression predictions is defined as the mean of the squared deviations between them. The variances around each of the values along the regression line are assumed to be equal. This is known as the assumption of *homoscedasticity* (homogeneous scatter or variance). When the variances of the Y values around the regression predictions are larger for some values of X than for others (i.e., when homoscedasticity is not present), then X serves as a better predictor of Y in one part of its range than in another. The homoscedasticity assumption is usually at least approximately true.

The variance around the regression line is a measure of the accuracy of the regression predictions. But it is not an easily interpreted measure of the degree of correlation because it has not been “normed” to vary within a limited range. Two other measures, closely related to each other, provide such a normed measure. These measures, which are always between zero and one in absolute value (i.e., sign disregarded) are: (a) the correlation coefficient,  $r$ , which is the measure devised by Karl Pearson; and (b) the square of that coefficient,  $r^2$ , which, unlike  $r$ , can be interpreted as a percentage.

Pearson’s correlation coefficient,  $r$ , can be computed using the following formula:

$$r_{yx} = r_{xy} = \frac{\frac{\sum (X - \bar{X})(Y - \bar{Y})}{N}}{\sqrt{\left(\frac{\sum (X - \bar{X})^2}{N}\right)\left(\frac{\sum (Y - \bar{Y})^2}{N}\right)}} \quad (5)$$

The numerator in equation 5 is known as the *covariance* of X and Y. The denominator is the square root of the product of the variances of X and Y. Hence, equation 5 may be rewritten:

$$r_{yx} = r_{xy} = \frac{\text{Covariance (X, Y)}}{\sqrt{[\text{Variance (X)}][\text{Variance (Y)}]}} \quad (6)$$

While equation 5 may serve as a computing guide, neither equation 5 nor equation 6 tells why it describes the degree to which two variables covary. Such understanding may be enhanced by stating that  $r$  is the slope of the least squares regression line when both X and Y have been

transformed into “standard deviates” or “z measures.” Each value in a distribution may be transformed into a “z measure” by finding its deviation from the mean of the distribution and dividing by the standard deviation (the square root of the variance) of that distribution. Thus

$$Z_x = \frac{X - \bar{X}}{\sqrt{\frac{\sum (X - \bar{X})^2}{N}}} \quad (7)$$

When both the X and Y measures have been thus standardized,  $r_{yx} = r_{xy}$  is the slope of the regression of Y on X, and of X on Y. For standard deviates, the Y intercept is necessarily 0, and the following equation holds:

$$\hat{Z}_y = r_{yx} Z_x \quad (8)$$

where  $\hat{Z}_y$  = the regression prediction for the “Z measure” of Y, given X;  $Z_x$  = the standard deviates of X; and  $r_{yx} = r_{xy}$  = the Pearsonian correlation between X and Y.

Like the slope  $b_{yx}$ , for unstandardized measures, the slope for standardized measures,  $r$ , may be positive or negative. But unlike  $b_{yx}$ ,  $r$  is always between 0 and 1.0 in absolute value. The correlation coefficient,  $r$ , will be 0 when the standardized regression line is horizontal so that the two variables do not covary at all—and, incidentally, when the regression toward the mean, which was Galton’s original interest, is complete. On the other hand,  $r$  will be 1.0 or  $-1.0$  when all values of  $Z_y$  fall precisely on the regression line  $rZ_x$ . This means that when  $r = +1.0$ , for every case  $Z_x = Z_y$ —that is, each case deviates from the mean on X by exactly as much and in the same direction as it deviates from the mean on Y, when those deviations are measured in their respective standard deviation units. And when  $r = -1.0$ , the deviations from the mean measured in standard deviation units are exactly equal, but they are in opposite directions. (It is also true that when  $r = 1.0$ , there is no regression toward the mean, although this is very rarely of any interest in contemporary applications.) More commonly,  $r$  will be neither 0 nor 1.0 in absolute value but will fall between these extremes, closer to 1.0 in absolute value when the  $Z_y$  values cluster closely around the regression line, which, in this standardized form, implies that the

slope will be near 1.0, and closer to 0 when they scatter widely around the regression line.

But while  $r$  has a precise meaning—it is the slope of the regression line for standardized measures—that meaning is not intuitively understandable as a measure of the degree to which one variable can be accurately predicted from the other. The square of the correlation coefficient,  $r^2$ , does have such an intuitive meaning. Briefly stated,  $r^2$  indicates the percent of the possible reduction in prediction error (measured by the variance of actual values around predicted values) that is achieved by shifting from (a)  $\bar{Y}$  as the prediction, to (b) the regression line values as the prediction. Otherwise stated,

$$r^2 = \frac{\text{Variance of Y values around } \bar{Y} - \text{Variance of Y values around } \hat{Y}}{\text{Variance of Y values around } \bar{Y}} \quad (9)$$

The denominator of Equation 9 is called the *total variance* of Y. It is the sum of two components: (1) the variance of the Y values around  $\hat{Y}$ , and (2) the variance of the  $\hat{Y}$  around  $\bar{Y}$ . Hence the numerator of equation 9 is equal to the variance of the  $\hat{Y}$  values (regression values) around  $\bar{Y}$ . Therefore

$$r^2 = \frac{\text{Variance of } \hat{Y} \text{ values around } \bar{Y}}{\text{Variance of Y values around } \bar{Y}} \quad (10)$$

= proportion of variance explained

Even though it has become common to refer to  $r^2$  as the proportion of variance “explained,” such terminology should be used with caution. There are several possible reasons for two variables to be correlated, and some of these reasons are inconsistent with the connotations ordinarily attached to terms such as “explanation” or “explained.” One possible reason for the correlation between two variables is that X influences Y. This is presumably the reason for the positive correlation between education and income; higher education facilitates earning a higher income, and it is appropriate to refer to a part of the variation in income as being “explained” by variation in education. But there is also the possibility that two variables are correlated because both are measures of the same



dimension. For example, among twentieth-century nation-states, there is a high correlation between the energy consumption per capita and the gross national product per capita. These two variables are presumably correlated because both are indicators of the degree of industrial development. Hence, one variable does not “explain” variation in the other, if “explain” has any of its usual meanings. And two variables may be correlated because both are influenced by a common cause, in which case the two variables are “spuriously correlated.” For example, among elementary-school children, reading ability is positively correlated with shoe size. This correlation appears not because large feet facilitate learning, and not because both are measures of the same underlying dimension, but because both are influenced by age. As they grow older, schoolchildren learn to read better *and* their feet grow larger. Hence, shoe size and reading ability are “spuriously correlated” because of the dependence of both on age. It would therefore be misleading to conclude from the correlation between shoe size and reading ability that part of the variation in reading ability is “explained” by variation in shoe size, or vice versa.

In the attempt to discover the reasons for the correlation between two variables, it is often useful to include additional variables in the analysis. Several techniques are available for doing so.

### PARTIAL CORRELATION

One may wish to explore the correlation between two variables with a third variable “held constant.” The *partial correlation coefficient* may be used for this purpose. If the only reason for the correlation between shoe size and reading ability is because both are influenced by variation in age, then the correlation should disappear when the influence of variation in age is made nil—that is, when age is held constant. Given a sufficiently large number of cases, age could be held constant by considering each age grouping separately—that is, one could examine the correlation between shoe size and reading ability among children who are six years old, among children who are seven years old, eight years old, etc. (And one presumes that there would be no correlation between reading ability and shoe size among children who are homogeneous in age.) But such a procedure requires a relatively large number of children in each age grouping.

Lacking such a large sample, one may hold age constant by “statistical adjustment.”

To understand the underlying logic of partial correlation, one considers the *regression residuals* (i.e., for each case, the discrepancy between the regression line value and the observed value of the predicted variable). For example, the regression residual of reading ability on age for a given case is the discrepancy between the actual reading ability and the predicted reading ability based on age. Each residual will be either positive or negative (depending on whether the observed reading ability is higher or lower than the regression prediction). Each residual will also have a specific value, indicating how much higher or lower than the age-specific mean (i.e., regression line values) the reading ability is for each person. The complete set of these regression residuals, each being a deviation from the age-specific mean, describes the pattern of variation in reading abilities that would obtain if all of these schoolchildren were identical in age. Similarly, the regression residuals for shoe size on age describe the pattern of variation that would obtain if all of these schoolchildren were identical in age. Hence, the correlation between the two sets of residuals—(1) the regression residuals of shoe size on age and (2) the regression residuals of reading ability on age—is the correlation between shoe size and reading ability, with age “held constant.” In practice, it is not necessary to find each regression residual to compute the partial correlation, because shorter computational procedures have been developed. Hence,

$$r_{xy \cdot x} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}} \quad (11)$$

where  $r_{xy \cdot z}$  = the partial coefficient between X and Y, holding Z constant;  $r_{xy}$  = the bivariate correlation coefficient between X and Y;  $r_{xz}$  = the bivariate correlation coefficient between X and Z; and  $r_{yz}$  = the bivariate correlation coefficient between Y and Z.

It should be evident from equation 11 that if Z is unrelated to both X and Y, controlling for Z will yield a partial correlation that does not differ from the bivariate correlation. If all correlations are positive, each increase in the correlation between the control variable, Z, and each of the focal

variables, X and Y, will move the partial,  $r_{xy.z}$ , closer to 0, and in some circumstances a positive bivariate correlation may become negative after controlling for a third variable. When  $r_{xy}$  is positive and the algebraic sign of  $r_{yz}$  differs from the sign of  $r_{xz}$  (so that their product is negative), the partial will be *larger* than the bivariate correlation, indicating that Z is a *suppressor variable*—that is, a variable that diminishes the correlation between X and Y unless it is controlled. Further discussion of partial correlation and its interpretation will be found in Simon 1954; Mueller, Schuessler, and Costner 1977; and Blalock 1979.

Any correlation between two sets of regression residuals is called a *partial correlation coefficient*. The illustration immediately above is called a *first-order partial*, meaning that one and only one variable has been held constant. A *second-order partial* means that two variables have been held constant. More generally, an *nth-order partial* is one in which precisely n variables have been “controlled” or held constant by statistical adjustment.

When only one of the variables being correlated is a regression residual (e.g., X is correlated with the residuals of Y on Z), the correlation is called a *part correlation*. Although part correlations are rarely used, they are appropriate when it seems implausible to residualize one variable. Generally, part correlations are smaller in absolute value than the corresponding partial correlation.

## MULTIPLE REGRESSION

Earned income level is influenced not simply by one's education but also by work experience, skills developed outside of school and work, the prevailing compensation for the occupation or profession in which one works, the nature of the regional economy where one is employed, and numerous other factors. Hence it should not be surprising that education alone does not predict income with high accuracy. The deviations between actual income and income predicted on the basis of education are presumably due to the influence of all the other factors that have an effect, great or small, on one's income level. By including some of these other variables as additional predictors, the accuracy of prediction should be increased. Otherwise stated, one expects to predict Y better using both

$X_1$  and  $X_2$  (assuming both influence Y) than with either of these alone.

A regression equation including more than a single predictor of Y is called a *multiple regression equation*. For two predictors, the multiple regression equation is:

$$\hat{Y} = a_{y.12} + b_{y1.2}X_1 + b_{y2.1}X_2 \quad (12)$$

where  $\hat{Y}$  = the least squares prediction of Y based on  $X_1$  and  $X_2$ ;  $a_{y.12}$  = the Y intercept (i.e., the predicted value of Y when both  $X_1$  and  $X_2$  are 0);  $b_{y1.2}$  = the (unstandardized) regression slope of Y on  $X_1$ , holding  $X_2$  constant; and  $b_{y2.1}$  = the (unstandardized) regression slope of Y on  $X_2$ , holding  $X_1$  constant. In multiple regression analysis, the predicted variable (Y in equation 12) is commonly known as the *criterion variable*, and the X's are called predictors. As in a bivariate regression equation (equation 2), one assumes both rectilinearity and homoscedasticity, and one finds the Y intercept ( $a_{y.12}$  in equation 12) and the regression slopes (one for each predictor; they are  $b_{y1.2}$  and  $b_{y2.1}$  in equation 12) that best fit by the criterion of least squares. The b's or regression slopes are *partial regression coefficients*. The correlation between the resulting regression predictions ( $\hat{Y}$ ) and the observed values of Y is called the *multiple correlation coefficient*, symbolized by R.

In contemporary applications of multiple regression, the partial regression coefficients are typically the primary focus of attention. These coefficients describe the regression of the criterion variable on each predictor, holding constant all other predictors in the equation. The b's in equation 12 are *unstandardized* coefficients. The analogous multiple regression equation for all variables expressed in standardized form is

$$\hat{Z}_y = b^*_{y1.2}Z_1 + b^*_{y2.1}Z_2 \quad (13)$$

where  $\hat{Z}_y$  = the regression prediction for the “z measure” of Y, given  $X_1$  and  $X_2$ ;  $Z_1$  = the standard deviate of  $X_1$ ;  $Z_2$  = the standard deviate of  $X_2$ ;  $b^*_{y1.2}$  = the standardized slope of Y on  $X_1$ , holding  $X_2$  constant; and  $b^*_{y2.1}$  = the standardized slope of Y on  $X_2$ , holding  $X_1$  constant.

The standardized regression coefficients in an equation with two predictors may be calculated from the bivariate correlations as follows:



$$b_{y1.2}^* = \frac{r_{y1} - r_{y2}r_{y12}}{1 - r_{12}^2} \quad (14)$$

$$b_{y2.1}^* = \frac{r_{y2} - r_{y1}r_{12}}{1 - r_{12}^2} \quad (15)$$

where  $b_{y1.2}^*$  = the standardized partial regression coefficient of Y on  $X_1$ , controlling for  $X_2$ ; and  $b_{y2.1}^*$  = the standardized partial regression coefficient of Y on  $X_2$ , controlling for  $X_1$ .

Standardized partial regression coefficients, here symbolized by  $b^*$  (read “b star”), are frequently symbolized by the Greek letter beta, and they are commonly referred to as “betas,” “beta coefficients,” or “beta weights.” While this is common usage, it violates the established practice of using Greek letters to refer to population parameters instead of sample statistics.

A comparison of equation 14, describing the standardized partial regression coefficient,  $b_{y1.2}^*$ , with equation 11, describing the partial correlation coefficient,  $r_{y1.2}$ , will make it evident that these two coefficients are closely related. They have identical numerators but different denominators. The similarity can be succinctly expressed by

$$r_{y1.2}^2 = b_{y1.2}^* b_{1y.2}^* \quad (16)$$

If any one of the quantities in equation 16 is 0, all are 0, and if the partial correlation is 1.0 in absolute value, both of the standardized partial regression coefficients in equation 16 must also be 1.0 in absolute value. For absolute values between 0 and 1.0, the partial correlation coefficient and the standardized partial regression coefficient will have somewhat different values, although the general interpretation of two corresponding coefficients is the same in the sense that both coefficients represent the relationship between two variables, with one or more other variables held constant. The difference between them is rather subtle and rarely of major substantive import. Briefly stated, the partial correlation coefficient—e.g.,  $r_{y1.2}$ —is the regression of one standardized residual on another standardized residual. The corresponding standardized partial regression coefficient,  $b_{y1.2}^*$ , is the regression of one residual on another, but the residuals are standard measure discrepancies from standard measure predictions, rather than the

residuals themselves having been expressed in the form of standard deviates.

A standardized partial regression coefficient can be transformed into an unstandardized partial regression coefficient by

$$b_{y1.2} = b_{y1.2}^* \frac{s_y}{s_1} \quad (17)$$

$$b_{y2.1} = b_{y2.1}^* \frac{s_y}{s_2} \quad (18)$$

where  $b_{y1.2}$  = the unstandardized partial regression coefficient of Y on  $X_1$ , controlling for  $X_2$ ;  $b_{y2.1}$  = the unstandardized partial regression coefficient of Y on  $X_2$ , controlling for  $X_1$ ;  $b_{y1.2}^*$  and  $b_{y2.1}^*$  are standardized partial regression coefficients, as defined above;  $s_y$  = the standard deviation of Y;  $s_1$  = the standard deviation of  $X_1$ ; and  $s_2$  = the standard deviation of  $X_2$ .

Under all but exceptional circumstances, standardized partial regression coefficients fall between  $-1.0$  and  $+1.0$ . The relative magnitude of the *standardized* coefficients in a given regression equation indicates the relative magnitude of the relationship between the criterion variable and the predictor in question, after holding constant all the other predictors in that regression equation. Hence, the *standardized* partial regression coefficients in a given equation can be compared to infer which predictor has the strongest relationship to the criterion, after holding all other variables in that equation constant. The comparison of *unstandardized* partial regression coefficients for different predictors in the same equation does not ordinarily yield useful information because these coefficients are affected by the units of measure. On the other hand, it is frequently useful to compare *unstandardized* partial regression coefficients across equations. For example, in separate regression equations predicting income from education and work experience for the United States and Great Britain, if the unstandardized regression coefficient for education in the equation for Great Britain is greater than the unstandardized regression coefficient for education in the equation for the United States, the implication is that education has a greater influence on income in Great Britain than in the United States. It would be hazardous to draw any such conclusion from the comparison of *standardized* coefficients for Great Britain and the United

States because such coefficients are affected by the variances in the two populations.

The multiple correlation coefficient,  $R$ , is defined as the correlation between the observed values of  $Y$  and the values of  $Y$  predicted by the multiple regression equation. It would be unnecessarily tedious to calculate the multiple correlation coefficient in that way. The more convenient computational procedure is to compute  $R^2$  (for two predictors, and analogously for more than two predictors) by the following:

$$R^2 = b_{y1.2}^2 r_{y1} + b_{y2.1}^2 r_{y2} \quad (19)$$

Like  $r^2$ ,  $R^2$  varies from 0 to 1.0 and indicates the proportion of variance in the criterion that is “explained” by the predictors. Alternatively stated,  $R^2$  is the percent of the possible reduction in prediction error (measured by the variance of actual values around predicted values) that is achieved by shifting from (a)  $\bar{Y}$  as the prediction to (b) the multiple regression values,  $\hat{Y}$ , as the prediction.

#### VARIETIES OF MULTIPLE REGRESSION

The basic concept of multiple regression has been adapted to a variety of purposes other than those for which the technique was originally developed. The following paragraphs provide a brief summary of some of these adaptations.

**Dummy Variable Analysis.** As originally conceived, the correlation coefficient was designed to describe the relationship between continuous, normally distributed variables. Dichotomized predictors such as gender (male and female) were introduced early in bivariate regression and correlation, which led to the “point biserial correlation coefficient” (Walker and Lev 1953). For example, if one wishes to examine the correlation between gender and income, one may assign a “0” to each instance of male and a “1” to each instance of female to have numbers representing the two categories of the dichotomy. The unstandardized regression coefficient, computed as specified above in equation 3, is then the difference between the mean income for the two categories of the dichotomous predictor, and the computational formula for  $r$  (equation 5), will yield the point biserial correlation coefficient, which can be interpreted much like any other  $r$ . It was then only a small step to the

inclusion of dichotomies as predictors in multiple regression analysis, and then to the creation of a set of dichotomies from a categorical variable with more than two subdivisions—that is, to *dummy variable analysis* (Cohen 1968; Bohrnstedt and Knoke 1988; Hardy 1993).

Religious denomination—e.g., Protestant, Catholic, and Jewish—serves as an illustration. From these three categories, one forms two dichotomies, called “dummy variables.” In the first of these, for example, cases are classified as “1” if they are Catholic, and “0” otherwise (i.e., if Protestant or Jewish). In the second of the dichotomies, cases are classified as “1” if they are Jewish, and “0” otherwise (i.e., if Protestant or Catholic). In this illustration, Protestant is the “omitted” or “reference” category (but Protestants can be identified as those who are classified “0” on both of the other dichotomies). The resulting two dichotomized “dummy variables” can serve as the only predictors in a multiple regression equation, or they may be combined with other predictors. When the dummy variables mentioned are the only predictors, the unstandardized regression coefficient for the predictor in which Catholics are classified “1” is the difference between the mean  $Y$  for Catholics and Protestants (the “omitted” or “reference” category). Similarly, the unstandardized regression coefficient for the predictor in which Jews are classified “1” is the difference between the mean  $Y$  for Jews and Protestants. When the dummy variables are included with other predictors, the unstandardized regression coefficients are the same except that the difference of each mean from the mean of the “reference” category has been statistically adjusted to control for each of the other predictors in the regression equation.

The development of “dummy variable analysis” allowed multiple regression analysis to be linked to the experimental statistics developed by R. A. Fisher, including the analysis of variance and covariance. (See Cohen 1968.)

**Logistic Regression.** Early students of correlation anticipated the need for a measure of correlation when the predicted or dependent variable was dichotomous. Out of this came (a) the phi coefficient, which can be computed by applying the computational formula for  $r$  (equation 5) to two dichotomies, each coded “0” or “1,” and (b) the tetrachoric correlation coefficient, which uses

information in the form of two dichotomies to estimate the Pearsonian correlation for the corresponding continuous variables, assuming the dichotomies result from dividing two continuous and normally distributed variables by arbitrary cutting points (Kelley 1947; Walker and Lev 1953; Carroll 1961).

These early developments readily suggested use of a dichotomous predicted variable, coded "0" or "1," as the predicted variable in a multiple regression analysis. The predicted value is then the conditional proportion, which is the conditional mean for a dichotomized predicted variable. But this was not completely satisfactory in some circumstances because the regression predictions are, under some conditions, proportions greater than 1 or less than 0. *Logistic regression* (Retherford 1993; Kleinman 1994; Menard 1995) is responsive to this problem. After coding the predicted variable "0" or "1," the predicted variable is transformed to a logistic—that is, the logarithm of the "odds," which is to say the logarithm of the ratio of the number of 1's to the number of 0's. With the logistic as the predicted variable, impossible regression predictions do not result, but the unstandardized logistic regression coefficients, describing changes in the logarithm of the "odds," lack the intuitive meaning of ordinary regression coefficients. An additional computation is required to be able to describe the change in the predicted proportion for a given one-unit change in a predictor, with all other predictors in the equation held constant.

**Path Analysis.** The interpretation of multiple regression coefficients can be difficult or impossible when the predictors include an undifferentiated set of causes, consequences, or spurious correlates of the predicted variable. Path analysis was developed by Sewell Wright (1934) to facilitate the interpretation of multiple regression coefficients by making explicit assumptions about causal structure and including as predictors of a given variable only those variables that precede that given variable in the assumed causal structure. For example, if one assumes that  $Y$  is influenced by  $X_1$  and  $X_2$ , and  $X_1$  and  $X_2$  are, in turn, both influenced by  $Z_1$ ,  $Z_2$ , and  $Z_3$ , this specifies the assumed causal structure. One may then proceed to write multiple regression equations to predict  $X_1$ ,  $X_2$ , and  $Y$ , including in each equation only those variables that come prior in the assumed causal order. For

example, the  $Z$  variables are appropriate predictors in the equation predicting  $X_1$  because they are assumed causes of  $X_1$ . But  $X_2$  is not an appropriate predictor of  $X_1$  because it is assumed to be a spurious correlate of  $X_1$  (i.e.,  $X_1$  and  $X_2$  are presumed to be correlated only because they are both influenced by the  $Z$  variables, not because one influences the other). And  $Y$  is not an appropriate predictor of  $X_1$  because  $Y$  is assumed to be an effect of  $X_1$ , not one of its causes. When the assumptions about the causal structure linking a set of variables have been made explicit, the appropriate predictors for each variable have been identified from this assumed causal structure, and the resulting equations have been estimated by the techniques of regression analysis, the result is a path analysis, and each of the resulting coefficients is said to be a "path coefficient" (if expressed in standardized form) or a "path regression coefficient" (if expressed in unstandardized form).

If the assumed causal structure is correct, a path analysis allows one to "decompose" a correlation between two variables into "direct effects"; "indirect effects"; and, potentially, a "spurious component" as well (Land 1969; Bohrnstedt and Knoke 1988; McClendon 1994).

For example, we may consider the correlation between the occupational achievement of a set of fathers and the occupational achievement of their sons. Some of this correlation may occur because the father's occupational achievement influences the educational attainment of the son, and the son's educational attainment, in turn, influences his occupational achievement. This is an "indirect effect" of the father's occupational achievement on the son's occupational achievement "through" (or "mediated by") the son's education. A "direct effect," on the other hand, is an effect that is not mediated by any variable included in the analysis. Such mediating variables could probably be found, but if they have not been identified and included in this particular analysis, then the effects mediated through them are grouped together as the "direct effect"—that is, an effect not mediated by variables included in the analysis. If the father's occupational achievement and the son's occupational achievement are also correlated, in part, because both are influenced by a common cause (e.g., a common hereditary variable), then that

part of the correlation that is attributable to that common cause constitutes the “spurious component” of the correlation. If the variables responsible for the “spurious component” of a correlation have been included in the path analysis, the “spurious component” can be estimated; otherwise such a “spurious component” is merged into the “direct effect,” which, despite the connotations of the name, absorbs all omitted indirect effects and all omitted spurious components.

**“Stepwise” Regression Analysis.** The interpretation of regression results can sometimes be facilitated without specifying completely the presumed causal structure among a set of predictors. If the purpose of the analysis is to enhance understanding of the variation in a single dependent variable, and if the various predictors presumed to contribute to that variation can be grouped, for example, into proximate causes and distant causes, a *stepwise regression analysis* may be useful. Depending on one’s primary interest, one may proceed in two different ways. For example, one may begin by regressing the criterion variable on the distant causes, and then, in a second step, introduce the proximate causes into the regression equation. Comparison of the coefficients at each step will reveal the degree to which the effects of the distant causes are mediated by the proximate causes included in the analysis. Alternatively, one may begin by regressing the criterion variable on the proximate causes, and then introduce the distant causes into the regression equation in a second step. Comparing the coefficients at each step, one can infer the degree to which the first-step regression of the criterion variable on the proximate causes is spurious because of the dependence of both on the distant causes. A stepwise regression analysis may proceed with more than two stages if one wishes to distinguish more than two sets of predictors. One may think of a stepwise regression analysis of this kind as analogous to a path analysis but without a complete specification of the causal structure.

**Nonadditive Effects in Multiple Regression.** In the illustrative regression equations preceding this section, each predictor has appeared only once, and never in a “multiplicative” term. We now consider the following regression equation, which includes such a multiplicative term:

$$\hat{Y} = a_{y.12} + b_{y1.2}X_1 + b_{y2.1}X_2 + b_{y.12}X_1X_2 \quad (20)$$

In this equation,  $Y$  is said to be predicted, not simply by an additive combination of  $X_1$  and  $X_2$  but also by their product,  $X_1X_2$ . Although it may not be intuitively evident from the equation itself, the presence of a multiplicative effect (i.e., the regression coefficient for the multiplicative term,  $b_{y.12}$ , is not 0) implies that the effect of  $X_1$  on  $Y$  depends on the level of  $X_2$ , and vice versa. This is commonly called an *interaction effect* (Allison 1977; Blalock 1979; Jaccard, Turrisi and Wan 1990; Aiken and West 1991; McClendon 1994). The inclusion of multiplicative terms in a regression equation is especially appropriate when there are sound reasons for assuming that the effect of one variable differs for different levels of another variable. For example, if one assumes that the “return to education” (i.e., the annual income added by each additional year of schooling) will be greater for men than for women, this assumption can be explored by including all three predictors: education, gender, and the product of gender and education.

When product terms have been included in a regression equation, the interpretation of the resulting partial regression coefficients may become complex. For example, unless all predictors are “ratio variables” (i.e., variables measured in uniform units from an absolute 0), the inclusion of a product term in a regression equation renders the coefficients for the additive terms uninterpretable (see Allison 1977).

### SAMPLING VARIATION AND TESTS AGAINST THE NULL HYPOTHESIS

Descriptions based on incomplete information will be inaccurate because of “sampling variation.” Otherwise stated, different samplings of information will yield different results. This is true of sample regression and correlation coefficients, as it is for other descriptors. Assuming a random selection of observed information, the “shape” of the distribution of such sampling variation is often known by mathematical reasoning, and the magnitude of such variation can be estimated. For example, if the true correlation between  $X$  and  $Y$  is 0, a series of randomly selected observations will rarely yield a correlation that is precisely 0. Instead, the observed correlation will fluctuate around 0 in the

“shape” of a normal distribution, and the standard deviation of that normal sampling distribution—called the *standard error* of  $r$ —will be

$$\sigma_r = \frac{1}{\sqrt{N-1}} \quad (21)$$

where  $\sigma_r$  = the standard error of  $r$  (i.e. the standard deviation of the sampling distribution of  $r$ , given that the true correlation is 0); and  $N$  = the sample size (i.e., the number of randomly selected cases used in the calculation of  $r$ ). For example, if the true correlation were 0, a correlation coefficient based on a random selection of 400 cases will have a standard error of approximately  $1/20$  or  $.05$ . An observed correlation of  $.15$  or greater in absolute value would thus be at least three standard errors away from 0 and hence very unlikely to have appeared simply because of random fluctuations around a true value of 0. This kind of conclusion is commonly expressed by saying that the observed correlation is “significantly different from 0” at a given level of significance (in this instance, the level of significance cited could appropriately be  $.01$ ). Or the same conclusion may be more simply (but less precisely) stated by saying that the observed correlation is “significant.”

The standard error for an unstandardized bivariate regression coefficient, and for an unstandardized partial regression coefficient, may also be estimated (Cohen and Cohen 1983; Kleinman, Kupper and Muller 1988; Hamilton 1992; McClendon 1994; Fox 1997). Other things being equal, the standard error for the regression of the criterion on a given predictor will decrease as (1) the number of observations ( $N$ ) increases; (2) the variance of observed values around predicted values decreases; (3) the variance of the predictor increases; and (4) the correlation between the predictor and other predictors in the regression equation decreases.

### PROBLEMS IN REGRESSION ANALYSIS

Multiple regression is a special case of a very general and adaptable model of data analysis known as the *general linear model* (Cohen 1968; Fennessey 1968; Blalock 1979). Although the assumptions underlying multiple regression seem relatively demanding (see Berry 1993), the technique is remarkably “robust,” which is to say that the technique yields valid conclusions even when the

assumptions are met only approximately (Bohrnstedt and Carter 1971). Even so, restricted or biased samples may lead to conclusions that are misleading if they are inappropriately generalized. Furthermore, regression results may be misinterpreted if interpretation rests on an implicit causal model that is misspecified. For this reason it is advisable to make the causal model explicit, as in path analysis or structural equation modeling and to use regression equations that are appropriate for the model as specified. “Outliers” and “deviant cases” (i.e., cases extremely divergent from most) may have an excessive impact on regression coefficients, and hence may lead to erroneous conclusions. (See Berry and Feldman 1985; Fox 1991; Hamilton 1992.) A ubiquitous but still not widely recognized source of misleading results in regression analysis is measurement error (both random and non-random) in the variables (Stouffer 1936; Kahneman 1965; Gordon 1968; Bohrnstedt and Carter 1971; Fuller and Hidiroglou 1978; Berry and Feldman 1985). In *bivariate* correlation and regression, the effect of measurement error can be readily anticipated: on the average, random measurement error in the *predicted* variable attenuates (moves toward zero) the correlation coefficient (i.e., the standardized regression coefficient) but not the unstandardized regression coefficient, while random measurement error in the *predictor* variable will, on the average, attenuate both the standardized and the unstandardized coefficients. In multiple regression analysis, the effect of random measurement error is more complex. The unstandardized partial regression coefficient for a given predictor will be biased, not simply by random measurement error in that predictor, but also by other features. When random measurement error is entailed in a given predictor,  $X$ , that predictor is not completely controlled in a regression analysis. Consequently, the unstandardized partial regression coefficient for every other predictor that is correlated with  $X$  will be biased by random measurement error in  $X$ . Measurement error may be non-random as well as random, and anticipating the effect of non-random measurement error on regression results is even more challenging than anticipating the effect of random error. Non-random measurement errors may be correlated errors (i.e., errors that are correlated with other variables in the system being analyzed), and therefore they have the potential to distort greatly the estimates of both standardized and



unstandardized partial regression coefficients. Thus, if the measures used in regression analysis are relatively crude, lack high reliability, or include distortions (errors) likely to be correlated with other variables in the regression equation (or with measurement errors in those variables), regression analysis may yield misleading results. In these circumstances, the prudent investigator should interpret the results with considerable caution, or, preferably, shift from regression analysis to the analysis of structural equation models with multiple indicators. (See Herting 1985; Schumacker and Lomax 1996.) This alternative mode of analysis is well suited to correct for random measurement error and to help locate and correct for non-random measurement error.

## REFERENCES

- Agresti, Alan and Barbara Finlay 1997 *Statistical Methods for the Social Sciences*, 3rd ed. Upper Saddle River, N.J.: Prentice Hall.
- Aiken, Leona S., and Stephen G. West 1991 *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, Calif.: Sage.
- Allen, Michael Patrick 1997 *Understanding Regression Analysis*. New York: Plenum Press.
- Allison, Paul 1977 "Testing for Interaction in Multiple Regression." *American Journal of Sociology* 83:144-153.
- Berry, William D. 1993 *Understanding Regression Assumptions*. Newbury Park, Calif.: Sage.
- , and Stanley Feldman 1985 *Multiple Regression in Practice*. Beverly Hills, Calif.: Sage.
- Blalock, Hubert M., Jr. 1979 *Social Statistics*, 2nd ed. New York: McGraw-Hill.
- Bohrnstedt, George W., and T. Michael Carter 1971 "Robustness in Regression Analysis." In Herbert L. Costner, ed., *Sociological Methodology 1971*. San Francisco: Jossey-Bass.
- , and David Knoke 1988 *Statistics for Social Data Analysis*, 2nd ed. Itasca, Ill.: F. E. Peacock.
- Carroll, John B. 1961 "The Nature of the Data, or How to Choose a Correlation Coefficient." *Psychometrika* 26:347-372.
- Cohen, Jacob 1968 "Multiple Regression as a General Data Analytic System." *Psychological Bulletin* 70:426-443.
- , and Patricia Cohen 1983 *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Fennessey, James 1968 "The General Linear Model: A New Perspective on Some Familiar Topics." *American Journal of Sociology* 74:1-27.
- Fox, John 1991 *Regression Diagnostics*. Newbury Park, Calif.: Sage.
- , 1997 *Applied Regression Analysis, Linear Models and Related Methods*. Beverly Hills, Calif.: Sage.
- Fuller, W. A., and M. A. Hidiroglou 1978 "Regression Estimates after Correcting for Attenuation." *Journal of the American Statistical Association* 73:99-104.
- Gordon, Robert 1968 "Issues in Multiple Regression." *American Journal of Sociology* 73:592-616.
- Hamilton, Lawrence C. 1992 *Regression with Graphics: A Second Course in Applied Statistics*. Pacific Grove, Calif.: Brooks-Cole.
- Hardy, Melissa A. 1993 *Regression with Dummy Variables*. Newbury Park, Calif.: Sage.
- Heise, David R. 1969 "Problems in Path Analysis and Causal Inference." In Edgar Borgatta and George W. Bohrnstedt, eds., *Sociological Methodology 1969*. San Francisco: Jossey-Bass.
- Herting, Gerald 1985 "Multiple Indicator Models Using LISREL." In H. M. Blalock, Jr., ed., *Causal Models in the Social Sciences*. New York: Aldine.
- Jaccard, James, Robert Turrissi, and Choi K. Wan 1990 *Interaction Effects in Multiple Regression*. Newbury Park, Calif.: Sage.
- Kahneman, D. 1965 "Control of Spurious Association and the Reliability of the Controlled Variable." *Psychological Bulletin* 64:326-329.
- Kelley, Truman Lee 1947 *Fundamentals of Statistics*. Cambridge, Mass.: Harvard University Press.
- Kleinman, David G. 1994 *Logistic Regression: A Self-Learning Text*. New York: Springer.
- , Lawrence L. Kupper, and Keith E. Muller 1988 *Applied Regression Analysis and Other Multivariable Methods*. Boston: PWS-Kent.
- Land, Kenneth C. 1969 "Principles of Path Analysis." In Edgar Borgatta and George W. Bohrnstedt, eds., *Sociological Methodology 1969*. San Francisco: Jossey-Bass.
- McClendon, McKee J. 1994 *Multiple Regression and Causal Analysis*. Itasca, Ill.: F. E. Peacock.
- Menard, Scott 1995 *Applied Logistic Regression Analysis*. Beverly Hills, Calif.: Sage.
- Mueller, John H., Karl F. Schuessler, and Herbert L. Costner 1977 *Statistical Reasoning in Sociology*, 3rd ed. Boston: Houghton Mifflin.
- Retherford, Robert D. and Minju Kim Choe 1993 *Statistical Models for Causal Analysis*. New York: Wiley.

- Schroeder, Larry P., David L. Sjoquist and Paula E. Stephan 1986 *Understanding Regression Analysis: An Introductory Guide*. Beverly Hills, Calif.: Sage.
- Schumacker, Randall E., and Richard G. Lomax 1996 *A Beginner's Guide to Structural Equation Modeling*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Stouffer, Samuel A. 1936 "Evaluating the Effects of Inadequately Measured Variables in Partial Correlation Analysis." *Journal of the American Statistical Association* 31:348-360.
- Walker, Helen M. 1929 *Studies in the History of Statistical Method*. Baltimore: Williams and Wilkins.
- , and Joseph Lev 1953 *Statistical Inference*. New York: Holt.
- Weisberg, Sanford 1985 *Applied Linear Regression*. New York: Wiley.
- Wright, Sewell 1934 "The Method of Path Coefficients." *The Annals of Mathematical Statistics* 5:161-215.

HERBERT L. COSTNER

## COUNTERCULTURES

The enclaves in which people of the modern era live no longer resemble the small, integrated, and homogeneous communities of earlier times; rather, these have been replaced by large societies that are complex and diverse in their composition. The United States, a prime exemplar, is composed of multiple smaller groups holding characteristics, beliefs, customs, and interests that vary from the rest of society. While there are many cultural universals binding such groups to the mainstream, they also exhibit significant cultural diversity. Some of these groups display no clear boundaries demarcating them from the rest of society and fail to achieve any degree of permanence. Yet those that do, and that also share a distinctive set of norms, values, and behavior setting them off from the dominant culture, are considered *subcultures*. Subcultures can be organized around age, ethnicity, occupation, social class, religion, or lifestyle and usually contain specific knowledge, expressions, ways of dressing, and systems of stratification that serve and guide its members (Thornton 1997). Distinctive subcultures within the United States include jazz musicians, gangs, Chicanos, gay, college athletes, and drug dealers. While it was once hypothesized that these subcultures would merge together in a "melting pot," incorporating a mix

of the remnants of former subcultures (Irwin 1970), trends suggest that they resist total assimilation and retain their cultural diversity and distinct identity.

Some subcultures diverge from the dominant culture without morally rejecting the norms and values with which they differ. Others are more adamant in their condemnation, clearly conflicting with or opposing features of the larger society. Milton Yinger first proposed, in 1960, to call these *contracultures*, envisioning them as a subset of subcultures, specifically, those having an element of conflict with dominant norms, values, or both (Yinger 1960). Indeed, the feature he identified as most compelling about a contraculture is its specific organization in opposition to some cultural belief(s) or expression(s). Contracultures often arise, he noted, where there are conflicts of standards or values between subcultural groups and the larger society. Factors strengthening the conflict then strengthen the contracultural response. Contraculture members, especially from such groups as delinquent gangs, may be driven by their experiences of frustration, deprivation, or discrimination within society.

Yinger's conceptualization, although abstract and academic at first, came to enjoy widespread popularity with the advent of the 1960s and the student movement. Here was the kind of contraculture he had forecast, and his ideas were widely applied to the trends of the time, albeit under another label. Most analysts of contracultures preferred the term *counterculture*, and this soon overtook its predecessor as the predominant expression. In 1969 historian Theodore Roszak published his *The Making of a Counter Culture*, claiming that a large group of young people (ages fifteen to thirty) had arisen who adamantly rejected the technological and scientific outlook characteristic of Western industrialized culture, replacing this, instead, with a humanistic/mysticist alternative. In a more recent update, Roszak (1995) reflected back on that time, further locating the counterculture phenomenon as an historical aberration that arose out of the affluence of post-World War II America. Kenneth Keniston (1971) described this counterculture as composed of distinct subgroups (radicals, dropouts, hippies, drug users, communards, or those living in communes) rising from the most privileged children of the world's wealthiest nation. Jack Douglas (1970) also discussed the social,

political, and economic background to this movement and its roots in members' entrenchment in the welfare state and the existing youth and student cultures. While this movement was clearly political as well, Douglas outlined some of its social dimensions, including rejection of the workaday world and its idealization of leisure, feeling, openness, and antimaterialism. Richard Flacks (1971) and Fred Davis (1971) followed with descriptions of the counterculture's overarching lifestyles, values, political beliefs, and ideologies. Ralph Turner (1976), Nathan Adler (1972), and Erik Erikson (1968) explored the social psychological implications of this counterculture, positing, respectively, a transformation in the self from "institution" to "impulse," the rise of an antinomian personality, where individuals oppose the obligatoriness of the moral law, and the formation of the *negative identity*. John Rothchild and Susan Berns Wolf (1976) documented the vast extension of countercultural outposts around the country and their innovations in child rearing. Charles Reich (1970) emphatically stated that this counterculture, consisting mostly of students, was being reinforced by merging with nonstudent youth, educated labor, and the women's movement, already effecting a major transformation in Western laws, institutions, and social structure. There was strong belief that this movement would significantly and permanently alter both society and its consciousness (Wuthnow 1976). After researching one commune in depth, Berger (1981) later mused about the survival of the counterculture, acknowledging its failure to meet earlier expectations, yet examining how its ideals and values become incorporated into the mainstream culture (cf. Spates 1976).

Other subcultural analysts noted more broadly that these groups are typically popular among youth, who have the least investment in the existing culture, and that, lacking power within society, they are likely to feel the forces of social control swiftly moving against them, from the mass media to police action. Countercultures were further differentiated from subcultures by the fact that their particular norms and values, were not well integrated into the dominant culture, generally known among group members, and other mainstream subcultures.

Yinger reclaimed theoretical command of the counterculture concept with his reflective expansions on the term in a presidential address for the

American Sociological Association (Yinger 1977) and a book that serves as the definitive statement on the topic (1982). He asserted the fundamental import of studying these sharp contradictions to the dominant norms and values of a society as a means of gaining insight into social order. Countercultures, through their oppositional culture (polarity, reversal, inversion, and diametric opposition), attempt to reorganize drastically the normative bases of social order. These alternatives may range from rejecting a norm or value entirely to exaggerating its emphasis in their construction of countervalues. As a result, some countercultures fade rapidly while others become incorporated into the broader cultural value system. Examples of countercultural groups would include the 1960s student counterculture (in both its political and social dimensions); youth gangs (especially delinquent groups); motorcycle gangs (such as the Hell's Angels); revolutionary groups (the Weathermen of the Students for a Democratic Society, the Black Panthers, millennialists [Williams 1996]); terrorist organizations (such as the Symbionese Liberation Army); extremist racist groups (the Ku Klux Klan, skinheads [Baron 1997; Hamm 1995; Young and Craig 1997], the Aryan Nation); survivalists (Branch Davidians); punkers; bohemian Beats; "straight edgers;" Rainbow family; Earth First! (Lange 1990; Short 1991) and some extreme religious sects (such as the Amish and the Hare Krishnas [see Saliba 1996]).

#### VARIETIES OF COUNTERCULTURES

Yinger believed that countercultural groups could take several forms. The *radical activist* counterculture "preaches, creates, or demands new obligations" (Yinger 1977, p. 838). They are intimately involved with the larger culture in their attempts to transform it. Members of the *communitarian utopian* counterculture live as ascetics, withdrawing into an isolated community forged under the guidelines of their new values. *Mystical* countercultures search for the truth and for themselves, turning inward toward consciousness to realize their values. Theirs is more a disregard of society than an effort to change it. These three forms are not necessarily intended to describe particular groups. Rather, they are ideal types, offered to shed insights into characteristics or tendencies groups may combine or approximate in their formation. Hippie communities or bohemian Beat groups



combined the mystical and utopian features of countercultures in their withdrawal from conventional society and their search for a higher transcendence. Revolutionary youth gangs, such as the 1960s radicals, the Hell's Angels, and the punkers, fuse the mystical search for new experiences and insights (often through drug use) with an activist attack on the dominant culture and its institutional expressions. Survivalists, Amish, and Hare Krishnas fuse the radical critique of conventional values and lifestyles with a withdrawal into an isolated and protected community.

Countercultures can be differentiated by their primary breaks with the dominant culture. Some take odds with its *epistemology*, or the way society contends that it knows the truth. Hippies and other mystics, for example, have tended to seek insight in homespun wisdom, meditation, sensory deprivation, or drugs, rejecting the rationality of science and technology. Others assert an alternative system of *ethics*, or the values pursued in defining good or striving for the good life. Some, like skinheads or KKK members, may be quite conservative in their definition of the good life; others are libertarian, advocating for people to "do their own thing." Still other countercultures offer alternative *aesthetic standards* by which fashion, taste, and beauty are judged. Punk or acid rock musical movements, performance or postmodernist art movements, and bohemian or hippie fashion movements were all aesthetic statements that incorporated a radical rejection of the standards of conventional taste and its connection to conventional values. Thus, entire countercultural movements may be based on their advocacy of these competing beliefs.

### COUNTERCULTURES AND SOCIAL CHANGE

Due to their intense opposition to the dominant culture, countercultures are variously regarded as "engines of social change, symbols and effects of change, or mere faddist epiphenomena" (Yinger 1982, p. 285). Examining these in reverse order, countercultures are often considered *mutations* of the normative social order, encompassing such drastic lifestyle changes that they invoke deep ambivalence and persecution. Most major countercultural mutations appear in the form of

religious movements. Other countercultures arise out of underlying or developing societal stress: rapid political or economic change; demographic transformations in the population (age, gender, location); a swift influx of new ideas; drastic escalation or diminishment of hopes or aspirations; weakening of ties to primary support circles (families, neighborhoods, work groups); and the erosion of meaning in the deepest symbols and rituals of society. These factors are then augmented by communication among people sharing such experiences or beliefs, leading them to coalesce into normatively and ideologically integrated groups. Countercultures can also precipitate social change if the norms and values they champion are incorporated into the mainstream. In commenting on the 1960s student movement, Chief Justice Warren Burger of the United States Supreme Court stated that "the turbulent American youth, whose disorderly acts [I] once 'resented,' actually had pointed the way to higher spiritual values" (cited in Yinger 1977, p. 848). Lasting influence may not always result from major countercultural movements, as witnessed by the rapid erosion in influence of Mao's cultural revolution after his death, yet it is possible. This occurs through a cultural dialectic, wherein each existing system, containing antithetical, contradictory ideas, gives rise to the oppositional values of a counterculture. These are ultimately incorporated into a future new order.

### COUNTERCULTURE CASE STUDIES

While the student movement of the 1960s was undoubtedly the largest and most influential counterculture to arise in the United States, a review of three more contemporary American countercultures may yield further insight into the parameters and character of these movements. Let us focus on the Hare Krishnas, punks, and survivalists.

The Hare Krishna movement, also known as the International Society for Krishna Consciousness (ISKCON), is one of the religious movements that became popular in the United States during the great "cult" period of the 1970s (Judah 1974; Rochford 1985). Its rise after the decline of the 1960s student movement is not coincidental, for many people who were former hippies or who were influenced by or seeking the ideals and values of the 1960s turned toward new religions (Tipton

1982) in search of the same features of community, idealism, antimaterialism, mysticism, transcendence to a higher plane, and “a spiritual way of life, which stands outside the traditional institutions found in America” (Rochford 1985, p. 44). Its primary values conflicting with mainstream culture include the rejection of (1) material success through competitive labor; (2) education to promote that end; (3) possessions for sense gratification; (4) authority favoring the status quo; (5) imperialistic aggression; and (6) the hypocrisy of racial discrimination (Judah 1974, p. 16). After the death of its American spiritual master, Srila Prabhupada, in 1977, however, the movement peaked and became more commercialized, transferring its emphasis from self-expression and uniqueness perpetuation of the sect, thereby becoming more of a mass phenomenon.

In contrast to the religious and value components of the Hare Krishnas’ rejection of mainstream culture, the punk or punk rock counterculture of the late 1970s and early 1980s was more of a style movement (Hebdige 1981). As Fox (1987, p. 349) has noted, “The punks created a new aesthetic that revealed their lack of hope, cynicism, and rejection of societal norms.” This was expressed in both their appearance and their lifestyle. The punk belief system was antiestablishment and anarchistic, celebrating chaos, cynicism, and distrust of authority. Punks disdained the conventional system, with its bureaucracies, power structures, and competition for scarce goods (Fox 1987). Members lived outside the system, unemployed, in old abandoned houses or with friends, and engaged in heavy use of drugs such as heroin and glue. Hardcore commitment was usually associated with semipermanent alteration of members’ appearance through tattoos, shaven heads, or Mohawk hairstyles (Brake 1985). The musical scene associated with punks was contrary to established tastes as well and often involved self-abandonment characterized by “crash dancing” (Street 1986).

In contrast to the hippies, Krishnas, and punks, the survivalist counterculture was grounded in exaggeration of right-wing beliefs and values. While some of the former groups preached love, survivalists were characterized by hate. Formed out of extremist coalition splinter groups such as neo-Nazis, the *KKK*, the John Birch Society, fundamentalist Mormon Freemen, the White Aryan Resistance, and tax protesters from *Posse Comitatus*,

survivalists drew on long-standing convictions that an international conspiracy of Jews was taking over everything from banking, real estate, and the press to the Soviet Politburo, and that the white race was being “mongrelized” by civil-rights legislation. A cleansing nuclear war or act of God, with “secular” assistance, would soon bring the Armageddon, eradicating the “Beast” in their midst (Coates 1987). Members thus set about producing and distributing survivalist literature, stockpiling machine guns, fuel, food, and medical supplies on remote farms and in underground bunkers, joining survivalist retreat groups, and attending survivalist training courses (Peterson 1984). Within their retreat communities they rejected the rationalization, technologization, secularization, and commodification of society, creating an environment of creative self-expression where an individual could accomplish meaningful work with a few simple tools. In their withdrawn communities and “utopian” future scenario, men would reclaim their roles as heads of the family; women would regain mastery over crafts and nurturance. Theirs is thus a celebration of fantasy and irrationality (Mitchell, n.d.). Yet while they isolate themselves in countercultures composed of like-minded individuals, they try to influence mainstream society through activism in radical right-wing politics as well. Their actions and beliefs, although rejecting the directions and trends in contemporary society, arise out of and represent frustrations felt by embattled segments of the Moral Majority (mainly fundamentalist Christian, white groups).

Scholarly treatment of counterculture movements is not limited to the United States. In the field of new social movements research, many European scholars have looked at organizations that are designed to mobilize forces against nationalistic cultures. These studies, ranging in topics from nuclear weapons, ecology, squatters’ rights, gays, women, and other countercultural groups (i.e., Autonomes or terrorist organizations), have explored the common denominators inherent in all new social movements. Using quantitative data from protest events collected from newspaper sources in France, the Netherlands, Germany, and Switzerland, Hanspeter Kriesi and others (1995) outlined the “new cleavage” that exists in these Western European societies.

Countercultures thus stand on the periphery of culture, spawned by and spawning social trends

and changes [by their opposition to dominant culture]. As Yinger (1977, p. 850) noted, "Every society gets the countercultures it deserves, for they do not simply contradict, they also express the situation from which they emerge. . . . Countercultures borrow from the dominant culture even as they oppose it."

(SEE ALSO: *Alternative Lifestyles; Social Movements; Student Movements*)

## REFERENCES

- Adler, Nathan 1972 *The Underground Stream: New Life Styles and the Antinomian Personality*. New York: Harper.
- Baron, Stephen W. 1997 "Canadian Male Street Skinheads: Street Gang or Street Terrorists?" *Canadian Review of Sociology and Anthropology* 34:151-174.
- Berger, Bennett 1981 *The Survival of a Counterculture*. Berkeley, Calif.: University of California Press.
- Brake, Michael 1985 *Comparative Youth Culture: The Sociology of Youth Culture and Subcultures in America, Britain, and Canada*. London: Routledge.
- Coates, James 1987 *Armed and Dangerous*. New York: Hill and Wang.
- Davis, Fred 1971 *On Youth Subcultures: The Hippie Variant*. New York: General Learning Press.
- Douglas, Jack D. 1970 *Youth in Turmoil*. Chevy Chase, Md.: National Institute of Mental Health.
- Erikson, Erik 1968 *Identity: Youth and Crisis*. New York: W. W. Norton.
- Flacks, Richard 1971 *Youth and Social Change*. Chicago: Markham.
- Fox, Kathryn J. 1987 "Real Punks and Pretenders: The Social Organization of a Counterculture." *Journal of Contemporary Ethnography* 16:344-370.
- Hamm, Mark 1993 *American Skinheads*. Westport, Conn.: Praeger.
- Hebdige, Dick 1981 *Subcultures: The Meaning of Style*. New York: Methuen.
- Irwin, John 1970 "Notes on the Present Status of the Concept Subculture." In D.O. Arnold, ed., *The Sociology of Subcultures*. Berkeley, Calif.: Glendessary Press.
- Judah, J. Stillson 1974 *Hare Krishna and the Counterculture*. New York: Wiley.
- Keniston, Kenneth 1971 *Youth and Dissent: The Rise of a New Opposition*. New York: Harcourt Brace Jovanovich.
- Kriesi, Hanspeter, Ruud Koopmans, Jan Willem Dyvendak, and Marco G. Giugni 1995 *New Social Movements in Western Europe*. Minneapolis: University of Minnesota Press.
- Lange, Jonathan I. 1990 "Refusal to Compromise: The Case of Earth First!" *Western Journal of Communication* 54:473-494.
- Mitchell, Richard n.d. *Dancing at Armageddon: Domsday and Survivalists in America*. Unpublished paper, Oregon State University.
- 1991 "Secrecy and Disclosure in Fieldwork." In W. Shaffir and R. Stebbins, eds., *Experiencing Fieldwork*. Newbury Park, Calif.: Sage.
- Peterson, Richard G. 1984 "Preparing for Apocalypse: Survivalist Strategies." *Free Inquiry in Creative Sociology* 12:44-46.
- Reich, Charles A. 1970 *The Greening of America*. New York: Random House.
- Rochford, E. Burke 1985 *Hare Krishna in America*. New Brunswick, N.J.: Rutgers University Press.
- Roszak, Theodore 1969 *The Making of a Counter Culture*. New York: Doubleday.
- 1995 "Introduction to the 1995 Edition." In Theodore Roszak *The Making of a Counter Culture*. Berkeley, Calif.: University of California Press.
- Rothchild, John, and Susan Berns Wolf 1976 *The Children of the Counter-Culture*. New York: Doubleday.
- Saliba, John A. *Understanding New Religious Movements*. Grand Rapids, Mich.: William B. Eerdmans Press.
- Short, Brant 1991 "Earth First! And the Rhetorics of Moral Confrontation." *Communication Studies* 42:172-198.
- Spates, James L. 1976 "Counterculture and Dominant Culture Values: A Cross-National Analysis of the Underground Press and Dominant Culture Magazines." *American Sociological Review* 41:868-883.
- Street, J. 1986 *Rebel Rock: The Politics of Popular Music*. Oxford: Basil Blackwell.
- Thornton, Sarah 1997 "General Introduction." In Ken Gelder and Sarah Thornton, eds., *The Subcultures Reader*. London: Routledge.
- Tipton, Steven M. 1982 *Getting Saved from the Sixties*. Berkeley, Calif.: University of California Press.
- Turner, Ralph 1976 "The Real Self: From Institutions to Impulse." *American Journal of Sociology* 81:989-1016.
- Williams, Rhys H. "Politics, Religion and the Analysis of Culture." *Theory and Society* 25:883-900.
- Wuthnow, Robert 1976 *The Consciousness Revolution*. Berkeley, Calif.: University of California Press.
- Yinger, J. Milton 1960 "Contraculture and Subculture." *American Sociological Review* 25: 625-635.

— 1977 “Countercultures and Social Change.” *American Sociological Review* 42:833–853.

— 1982 *Countercultures*. New York: Free Press.

Young, Kevin, and Laura Craig 1997 “Beyond White Pride: Identity, Meaning, and Contradiction in the Canadian Skinhead Subculture.” *Canadian Review of Sociology and Anthropology* 34:175–206.

PATRICIA A. ADLER  
PETER ADLER

## COURT SYSTEMS AND LAW

Most sociological discussions of law begin with Weber’s definition in which a specific staff is charged with avenging norm violation or ensuring compliance (*Economy and Society* 1968, p. 34). Weber’s goal was to distinguish law from morality and convention, by which a whole community may act to impose sanctions. He also developed his now classic typology of formal legal systems (those limited only to legal as opposed to those legal systems he called “substantive,” based on religious, economic, or moral criteria) and rational (those legal systems based on rules as opposed to those involving use of oracles, oaths, and ordeals, for example). Although he was careful to call these distinctions “ideal-types,” that caution has not stopped persons from offering specific examples that are actually mixed types, as in speaking of “khadi justice” (a term, unfortunately, used by Weber himself) as a prime example of substantively irrational decision making in which a Moslem khadi sits under a palm tree and dispenses justice according to his personal feelings or inspiration. As Rosen (1989, ch. 1) shows, in actual cases he observed, the khadi does not exclude any evidence but relies on witnesses, notaries, documents, and any relevant evidence as well as testimony from interested parties. His goal, as in Islamic law generally, is to “put people back in the position of being able to negotiate their own permissible relationships . . .” (p. 17; see Nader 1969; Starr 1992). He follows a careful procedure, although the conduct of persons in his court may appear to Westerners to be more informal and more disorderly than that allowed in a typical Western court.

While anthropological studies of tribal societies have been shown to exhibit the whole panoply

of Weberian categories (cf. Gluckman 1954; Bohannan 1957, 1967; Howell 1954; Kuper and Kuper 1965), even the classic moot (eg. Gibbs 1963; Gulliver 1969), while classifiable as substantively irrational since it subjects disputes to discussion by a whole village of involved or even merely curious onlookers, still follows definite procedures and is guided by a mediator or authorized persons. See also Stone (1979) on miners’ meetings in the gold rush Yukon; MacLachland (1974) on the tribunal of the *Acordada* in eighteenth-century Mexico; and the People’s Courts in postrevolutionary Russia (Feifer 1964). The goal of the moot, as Gibbs notes, is not solely legal but is at least what he calls therapeutic since the end sought is restoration of relationships and harmony. It turns out that goal is quite consonant with what we find in formally rational systems in the Western world. In any case, bases of all legal systems include: first, a reasonable certainty and predictability on which persons can depend, which, in practice, means no retroactive laws (as Fuller 1969 notes); second, fairness, which comes down to treating like cases alike; and third, justice should not only be done but be seen to be done, which means that there are no secret decisions, and decisions are in accord with generally agreed values, assuming they exist. These similarities are more important than the many variations in detail that can be found in different legal systems (cf. Pospisil 1958).

In what follows, we shall focus on the two main systems of law, *common law* and *Civil law*, which are found widely in much of the world. In so doing, we can give only passing reference to such systems as Islamic, Jewish, or Tibetan law, or to the many subtle differences found in tribal law. To further complicate matters, many national legal systems are blends of other systems, especially as the result of conquest (as is the case in Japan, which took over a German code but then had common-law features of public law grafted onto it after its occupation by the United States in the 1940s). In modern society, legal systems have become identified with the nation-state so that Canadian law differs from American law, as does Scotch from French law. Nevertheless, many can be said to share one of the two traditions we are considering. Although the two systems have been converging in many respects, each is distinctive in outlook as well as in court organization and the kinds of legal careers likely to be found.

The *Civil law* is by far the older, going back at least to the *Corpus Juris Civilis* of the Emperor Justinian in the sixth century C. E., though some scholars would trace it back to the 12 Tables of Rome, said to have been put together in 450 B. C. E. This system is found throughout Western and now Eastern Europe, most of Central as well as South America, and in many other areas in Africa as well as Asia, plus the state of Louisiana (for an example see *Williams vs. Employers Liability Assurance Corporation, Limited*, 296 F.2d 569 (1961) U.S), the Canadian province of Quebec (Magnet 1980), and Scotland. The *common law* system is often dated from the Norman conquest of England in the eleventh century, when the Normans sought to impose a single system on the country, hence the name "common law." Common law is now found not only in Great Britain but in those countries that made up the British Empire, such as Canada, Australia, New Zealand, Ireland, India, as well as the United States, and is influential in many countries of Africa and Asia. During colonialism, the common law countries preserved legal ties with England through Privy Council appeal, a process largely abolished or greatly attenuated today. Comparable appeals or ties to a single place were largely unknown to the civil law system. Whether the European Community will establish some form of tie remains to be seen.

#### THE BASIC COURT PARADIGM: THE TRIADIC DILEMMA

When persons with a grievance decide to take action, they may: (1) act separately from each other as in direct attacks, seizure of property, etc. (see Black 1993, chs. 2, 5); (2) confront the other party; or (3) enlist the participation of a third party to help settle the dispute. Courts are preeminently concerned with this third option, what we call "the triadic dilemma." Although courts deal with matters other than dispute settlement, such as law making (when judges make a new policy in decision making) as well, of course, as social control, they all employ a variant of a triad. Two persons or collectivities who have been unable to settle their differences or who have accused others of public harm have in all societies sought the help or guidance of a third party. The hope of each party is that the third party will side with them. Although each may think it fairest if the third party is neutral, in practice each hopes for a decision in his or

her favor. As Simmel (1902; Wolff 1950; Caplow 1968) shows, parties of three are inherently unstable, being liable to break down when two of the three form a coalition to defeat the third, a phenomenon well-known to parents of small children.

The third party may be simply a go-between who offers his services unasked in the interest of preserving or restoring good relations. Or he or she may, as Shapiro (1986, ch. 1) notes, take a more active role as a mediator, though only with the consent of the two disputing parties. While mediators may seek to preserve their neutrality, in fact he may, as in the case of real estate agents, make proposals of their own and may shade the decision by how they handle the facts each shares with them. Less consent from the disputants is involved when an arbitrator appears who may be persons agreed to by the parties or persons imposed by the state or the terms of a labor contract, for example. Arbitrators are under no obligation of coming up with a solution agreeable to both, but they will usually try for such a solution. The most coercive situation of all is the appearance of a judge, either as required by law or by decision of a governmental body. Here the parties may have no say at all either in his appointment or on the body of rules he applies, which may be the terms of a contract or the rules of law.

The dilemma in all these triadic situations is, as noted, preventing the breakdown into two against one. Whatever the outcome, those who lose are likely to feel the other two have somehow combined against them or that they were defeated from the start. In common law systems, the myth of the neutral judge is much celebrated, although since judges are political figures who owe their position to their political activities and may even have been rewarded with the judgeship for their services, it is not surprising the loser feels he never had a chance. In Civil law systems, there is no pretence of neutrality. The judge is a member of the civil service and hence a part of government itself. With the prosecutor in criminal cases being also a member of the administration, the dice are loaded as two against one. On the other hand, this does not mean the civil law system is less fair or just.

Indeed, because of the many protections offered to the accused in common law systems, one author commented that if he were guilty, he would rather be tried in the United States but if innocent,

then he would prefer the Civil law system because, for reasons we shall explain, judges have far less power than they do in the United States, and the system is far less adversarial. The Civil law judge assumes a more neutral role in searching for “truth,” rather than winners or losers. Attempts are made in both systems to produce an outcome that each will feel is “just.” Disputants may be asked to submit voluntarily, and if they do, they will be felt to have offered at least a modicum of consent to the outcome. Alternatively, attempts are made to avoid an all-or-none outcome. For example, in auto accident cases, one party may be felt to be 70 percent at fault and the other 30 percent, with a division of property in that proportion. In criminal situations, as in the United States, a plea bargain may enable the accused to walk away with a much lesser penalty than he or she might otherwise have suffered.

Actually, although plea bargaining is thought of as peculiar to criminal proceeding, negotiation is equally common in what are called “civil situations” (involving actions between two or more persons) where, as in criminal proceedings, an estimated 90 percent of cases are settled without trial. Such resort to negotiation places a special strain on the triadic model, particularly where the judge may himself participate, at least to the point of interviewing the accused in case of crime, or the parties in civil law, to provide assurance, at least in his own mind, that the accused or the plaintiff is aware of what is happening and has voluntarily entered into the “deal.” (Klein 1976). What one needs to understand is that the ubiquity of negotiation in American law is an organizational model. Lawyers, both defending and prosecuting, as well as lawyers in other courtroom situations become “repeat players” (Galanter 1974), who deal routinely with one another as well as with judges, clerks, and others in the court. Gradually, their work becomes routinized, with all focusing on getting things moving and coming out with cases settled. The goal of the system, then, becomes one of efficiently moving cases through, with participants seeing themselves, however unwittingly, as agents of the system. As that takes place, the triad can be seen as disappearing altogether and being replaced by a work group consisting of three or more players who have an interest in the outcome of the game (see Eisenstein and Jacob 1977; Jacob 1983).

We proceed to a detailed examination of the two systems, beginning with the common law system as most familiar to American readers. But that very familiarity is likely to blind us to the assumptions of the system that contrast so sharply with the equally hidden assumptions of the civil law system.

### THE COMMON LAW SYSTEM

Although most persons are taken with the image of justice as a “blind lady” who acts on the basis of the facts and the inherent justice of the situation, as Jacob (1996) points out, courts in common law systems are ridden with policy assumptions, no more so in the United States than in other places. While courts go about their business of settling disputes and ensuring orderly procedures, their procedures send symbolic messages (see Nelken 1997; Sarat and Kearns 1988). This is especially the case for appellate courts where, in contrast to European courts, judges are fond of wrapping up their decisions in opinions that are often more widely cited and influential than the decision itself. While conservatives are usually at pains to insist that judges confine themselves to being “strict constructionists,” their opinions resonate with what the legal scholar Dworkin (1977) calls “principles.” In the classic case of *Riggs vs Palmer* (115 N.Y. 506, 22 N.E. 188, 190, 1889), a presumptive heir on becoming alarmed at the possibility that his grandfather might change his will proceeded to eliminate that possibility by murdering him. The grandson was properly tried, but defended his right to his inheritance. The court refused to award the inheritance to him, leading a dissenting judge in the case to ask for the court’s reasoning for this decision. After all, wrote the dissenter, the will was in order, was it not? There were the required witnesses, and there was no question that it was the intent of the testator that the young man should receive the bulk of the estate. With barely a reference to those issues, the majority awarded the bulk of the estate to the deceased’s daughters, who, under the will, were to receive only token amounts. In sum, the majority proceeded to rewrite the will in direct contravention of the deceased’s clearly expressed wishes. Although the majority hunted mightily for a source for their decision, turning variously to Aristotle, the Bible, an ancient case from Bologna, the Napoleonic Code, Roman law and, finally, to a rather desperate assumption that no specific law was really

needed, anyhow. In the end they asserted that there is a “fundamental maxim” found in all “civilized countries,” namely:

*No one shall be permitted to profit by his own fraud, or to take advantage of his own wrong, or to found any claim upon his own iniquity, or to acquire property by his own crime.* (Riggs vs. Palmer 1889)

But the dissenter asked: What is the legal source of this “maxim?” It could not be found in any case previously decided, nor in any statute (such is no longer the case in most jurisdictions). Although he agreed that the principle had intuitive appeal and might be found in most religious and moral systems, the law, at least Western secular law, is not simply a set of religious or moral principles. In the United States, the Constitution specifically erects a wall between church and state. Since that case, the principle, once announced by the court, has been cited and is indeed now a part of U. S. (and most other common law countries) law. But we must not go too far and declare that common law is simply a set of principles. Rather, the law is informed by policy assumptions, often hidden, but influential nonetheless. Although civil law is equally political, attempts are much stronger to hide the policy assumptions with an insistence that only a legislature can make law. All judges are supposed to do is apply it with as little innovation or interpretation as possible, let alone enunciating “principles.” The overall issue of how much politics and governmental policy is reflected in judicial decision making clearly affects the triadic system’s functioning to produce or not produce a sense of justice done and seen to be done. It is this feature that has been seized upon by Marxists and proponents of socialist law who see both common law and civil law as simply disguised systems whereby bourgeois ideologies are foisted on a powerless proletariat by capitalist exploiters, whether governmental or private. Their answer to the triadic dilemma is that there never can be neutral third parties, with the result that the only reality is endless strife between parties with temporary truces as victors seek to pick up the spoils.

In addition to taking policy positions, courts in common law systems can have massive effects on society by judicial review of legislative acts, a phenomenon severely limited in civil law systems as we shall note. Courts also provide a major

alternate route to persons who lack funds or who have been stymied by attempts to influence legislation. Persons can convert a private grievance into a public cause (Savat and Scheingold 1998) by, for example, deciding that severe burns from hot coffee served at McDonald’s requires court action to impose punitive damages as a warning to all companies serving the public that consumer safety must be a part of the design even in private firms. So too, persons suffering from lung cancer who have been unable to get protective legislation against cigarette companies have turned to the courts, not simply for financial awards, but for vindication of what they feel are violations of their rights, as citizens, to life. When added up, the damage awards in such suits as well as those even in routine automobile accidents, violations of privacy, and tort and contractual disputes have never been totaled but constitute a huge shift in resources comparable to that involved in taxation. Nor have we touched on the part that lawyers’ fees play in such cases.

Finally, unlike judges in civil law jurisdictions who are part of the civil service from the start (judges in France, for example, even go to special schools), judges especially in the United States, often come to their judgeships in mid-life, after serving in political positions, business or other areas of life. Jacob points out that:

*... between 1963 and 1992, between 58 and 73 percent of federal appeals judges had a record of party activism before their appointments; among federal district judges, between 49 and 61 percent had such a background.* (1996, p.19)

Yet such a background should not lead to cynicism that judges are necessarily biased or pro-business or pro-party. Their background in ordinary society helps ensure a commitment to the importance of the rule of law, of procedural fairness, and of individual rights as a legitimate expectation of ordinary citizens.

A special characteristic of common law court is often the source of surprise and some envy on the part of those schooled in civil law systems. American courts are often perceived as a jumble, with multiple urban, country, and state systems, so that what is actionable in Montana may not be actionable in Idaho and that if one is not happy with treatments in a state court, one can, with

some issues, move to a federal court. Such “forum shopping” seems a travesty to those accustomed to single systems. Yet the United States has more a single system than have many European countries. In those countries, there are often two sets of courts—ordinary courts that try the vast majority of civil cases, and a second set of administrative courts for those who have quarrels with the government or administration. In the United States most regular courts can handle any case that comes before them (though there are many specializations), and ultimately, the U. S. Supreme Court sits at the top as the final arbiter of law (considering constitutional questions especially important and above all). Nor, it should be added, are courts entirely separate systems. They are dependent for salary and other resources on what Congress and legislatures will provide, and they do not appoint their own numbers. The U. S. Supreme Court is indeed supreme on law, but it is in no sense the apex of a bureaucracy that can appoint and discipline members of lower courts. Further, it is basically passive, waiting for cases to be brought to it when there is a “case or controversy,” thus severely limiting its ability to act autonomously as a conscience of the nation. Often citizens gnash their teeth as the U. S. Supreme Court refuses to hear a case or decides it on the narrow issue that happens to have been presented to it. In that sense, the entire legal system is the property of the lawyers and what they choose to present to the courts, a situation vastly different from what we find in civil law systems.

The impact of lawyers in common law systems is greater than in civil law and indeed in any other system known to the world. Two major factors help account for lawyer dominance. One is the dependence on case law. Although legislation and statutes are basic sources of law, the necessity for interpretation and application to particular cases places enormous power in the hands of courts. If all legal systems require a strong sense of certainty, then that certainty is provided in common law systems by the majestic procession of cases, whether simply confirming one another, adding details, or overruling contradictory cases. To argue a case in an American court is to recite cases, and if there are enough of them, lawyers and judges feel the outcome can be considered settled. On the other hand, civil law systems depend on codes as enacted by revolutionary regimes, which codes are felt to

be the ultimate source of law, supplemented by legislation that is still felt to be a kind of supplement to the code. In a sense, the code is thought to settle the matter of certainty with little need for lawyers to interpret it. (This is less true in Germany where lawyers play a larger role.)

But in addition, especially in America, the model of the triad is felt to be basic. A court case consists of two adversaries who argue before a (one hopes, neutral and just) judge. The emphasis is on the adversarial process itself, a situation that produces not necessarily “truth” but rather a victory for one side. Lawyers play the key role, a matter that often leaves the outcome to the skill of the lawyers as much as to other features of the case. The judge is felt, if not to be neutral, at least to be passive, waiting for lawyers to present objections or evidence as they wish. If a lawyer chooses not to present a piece of evidence or simply sloppily forgets to do so, the judge cannot intervene to instruct the lawyer on what he has left out. In civil law systems, the judge, while a mere civil servant, has more power to direct the course of the trial, assuming what is often spoken of as an “inquisitorial style.” Given the major role that lawyers play in common law systems, it is important to give attention to how that role is played out. Since data are more complete, we shall use U.S. sources. However, comparable rates of increase for most categories are found in Canada and Great Britain (see Galanter 1992). This is not meant to deny the differences, especially cultural variations, in those countries (see Atiyah and Summers 1987).

#### **DOMINANCE OF LAWYERS IN COMMON LAW SYSTEMS**

In spite of their widespread influence and frequently very high income, lawyers in America are not a happy lot. They are not esteemed (a Gallup Poll found that 46 percent of respondents rated lawyers “low” or “very low” in honesty and ethical standards, just barely above used-car salesmen). A survey by the California Bar Association in 1992 reported that 70 percent of those polled said they would choose another career if they could. Even more—75 percent—confessed that they would not want their children to become lawyers. Other studies report that lawyer job satisfaction is dropping, along with much higher levels



of alcoholism, drug abuse, and symptoms of depression than those found in the general population.

In spite of such indices of self-destruction, the number of lawyers in the United States has been rising, especially in the 1990s. From 374,000 in 1975, the number of lawyers will soon top one million as 31,000+ new lawyers are admitted to the bar every year. The field is proving attractive to minorities and women. From a low of only about 3 percent in 1971, female lawyers now make up over one-quarter of the total of practicing lawyers, and nearly one-half of students in entering law school classes. Female lawyers, as a group, are younger, with only 7 percent being over 50 compared to 30 percent of their male colleagues.

The location of practice has, however, not changed significantly. Private practice is dominant, even increasing, so that by 1991, 73 percent of lawyers were in private practice, with only 8.8 percent found in private industry, and 8.2 percent in government. But private practice has been undergoing profound changes. Solo practitioners have become scarcer as lawyers move increasingly to firms. The increase is mainly in larger firms (those with at least eleven lawyers). In 1980, the very large firms (one hundred or more lawyers) accounted for only 7 percent of firm employment. By 1991, that percent jumped to twenty-three. The large firms are the more common locus for men, with women being more likely to be found in government, legal aid, and in public defender's offices. Some of these differences are declining as more women enter the profession and attain experience. On the other hand, more women, proportionally, are leaving the profession.

It is the large firms that attract more and more of the new lawyers who seek distinguished and lucrative careers. The largest—often called “mega firms”—range from the Washington, D. C., firm of Williams & Connolly with 127 lawyers (sixty-one partners) producing revenues of \$78 million, to true giants, such as the New York firm of Skadden, Arps, Slate, Meagher & Flom, with over 1,000 lawyers (236 partners) earning well over half-a-billion dollars in gross revenue. Some firms are even larger. They are not, of course, all under one roof but scattered in different cities as well as in foreign countries. Not only are these the places where the largest salaries are found, but they are also the platforms from which government and

other influential careers are launched, including leading positions on major committees and boards, as well as ambassadorships and presidencies. Twenty-five of the forty-one U. S. presidents have been lawyers, as well as half of U. S. senators and nearly half of all members of Congress. Lawyers are widely found in governorships and state legislatures as well. If these are not the most esteemed members of society, they certainly are among the most powerful and perhaps the most feared.

Although most persons in common law systems are aware of the presence of lawyers in those settings, it is not the image most have when they think of lawyers, and it is not the setting in which they see them in television drama. Instead, it is the lawyer, often solo or in a small firm, arguing for his or her client in a courtroom before a jury. The television image runs counter to the image of real lawyers in the news, leading to charges that the United States is a “litigious” society, in which large awards are given for burns suffered from spills of hot coffee, and there are suits against arrest for breast-feeding in public or for recovery of expenses on being stood up for a date. Though most such cases are thrown out immediately by disgusted judges or settled out of court for modest sums, critics person who may never hear of those outcomes continue to demand that we follow the lead of the nonlitigious Japanese, for example, who make do with very few lawyers. Actually, the number of lawyers in Japan is deliberately kept low by the governing elite to preserve a hierarchical social order. Nor are the Japanese devoid of a taste for litigation by any means (Haley 1978, 1991, ch. 5).

Nor has the United States had a “litigation explosion” nearly as great as some have claimed. There was only a moderate increase at the state level in the 1990s. It is true that there has been a large increase in the number of federal cases, not of the trivial sort noted above but rather of big businesses suing each other. The news of most of those ends up on the back pages of the *Wall Street Journal*. Other federal cases deal with asbestos and similar injuries as well as other suits by government (Galanter 1983; Galanter and Palay 1991). The reason it seems that the U.S. is experiencing a litigation explosion is that there is an increasing prominence of what have been called “mega” cases, in which large masses of lawyers and experts pursue a single case, sometimes for years on end. Although most involve business, a few involve

highly prominent individuals who have the resources and will to fight, intimidate, and otherwise bring up issues, subsidiary issues, and more of what Damaska (1978, p. 240) has called “companion litigation” where, along with the main case, separate suits are filed on discovery, on legal fees, on standing, and on other issues, which lengthen proceedings and often do little more than harass the other side or both sides into exhaustion. Such cases take the form of a “prolonged clinch and . . . settlement” (Galanter 1983, p. 163) while rarely ever ending up in court. Meanwhile, such cases contribute to the image of the United States as a “litigious” society.

Yet, even granting all such cases, very few lawyers are involved in doing such things. Only a small minority of all dealings of U. S. lawyers ever result in a contested court action. Most legal practice takes place in offices for the benefit of business firms, and only a small minority deal with individual clients at all. Even in those cases, lawyers spend much of their time persuading suit-eager clients not to go to court but to work out a settlement. Lawyers generally limit themselves to cases they think they can win. Often filing a case is symbolic of seriousness of intent, forcing a response from the other party; but the cases are settled, sometimes on the very eve of the court date or even as the trial, if there is one, is in process.

What all this amounts to is that the legal profession in the United States is “split.” Most lawyers quietly carry on the journeymen work of settling disputes and assisting persons to compromise so that they can carry on with their lives. A very, very few carry out the courtroom battles of the O. J. Simpson type that dominate the front pages of the nation’s newspapers. In such cases, lawyers are not seen as settling disputes but they are seen by many people as “getting people off,” leading to cynicism or to despising lawyers, even when or even because they win. Many cases take place outside the large law firm, but the reputation of lawyers created by the sensational cases affects the public image of all lawyers, wherever located.

One special feature of the split deserves attention; namely, income. Solo and small-firm lawyers carry out much of the work of helping persons set up partnerships, get a divorce or settlement from

an insurance company, draw up wills, and deal with persons who are in minor trouble with the law. On the other hand, work in a large law office is carried out by specialists who do the complex work of big business. Large businesses often make use of their own “in-house counsel” for the routine work of contracts, labor-management negotiation, and other repetitive legal activities. The company turns to outside law firms for the unusual, once-only activity, such as mergers and acquisitions, floating new securities, takeovers, and bankruptcies. Such activities call for the highest degree of expertise and knowledge, far beyond what a solo lawyer might be called upon to have. A major study of Chicago lawyers (Heinz and Laumann 1982) asked them to rank legal specialties in prestige. At the top were securities, tax, antitrust, patents, banking, and public utilities—the activities in which large law firms are involved. At the bottom were criminal defense and prosecution, personal injury, consumer debt, landlord-tenant, divorce, and family—the concerns of the solo and small-firm lawyer. The income differentials between the two clusters are equally impressive. A major study reported salaries from large firms in Indianapolis and New York to average around \$300,000 per partner for the year, but many make much more. A New York law firm reported that each of its 121 lawyers (sixty-one equity partners) earned over \$1 million. Many others were not far behind. Lawyers in solo and small firms are not poor but make a good deal less. A 1995 survey reported that those lawyers earn somewhere between \$75,000 and \$100,000 a year, assuming they work a full two-thousand billable hours, which some do not. Associates (that is, nonpartners) start out, according to a 1996 study, from lows of \$40,000 to as high as \$70,000, but then rise with each year in the larger firms to \$150,000 and up, plus bonuses. Stories of such incomes add little to offset the low esteem in which lawyers are held, especially since most persons who deal with lawyers find the lawyer wants money “up front” or on a retainer basis, unless a contingency arrangement is made, and often even then. (Note: The preceding section draws from the author’s paper, Gross 1998.)

### TOO MANY LAWYERS?

A final issue that troubles many observers both in common law and civil law countries is

**Judges, Lawyers, and Civil Litigation in Selected Countries**

Country	Judges		Lawyers		Civil Cases	
	Date	Number per Million	Date	Number per Million	Date	Number per Million
Australia	1977	41.6	1975	911.6	1975	62.06
Belgium	1975	105.7	1972	389.7	1969	28.31
Canada	1970	59.3	1972	890.1	1981-2	46.58
Denmark					1970	41.04
England/Wales	1973	50.9	1973	606.4	1973	41.1
France	1973	84.0	1973	206.4	1975	30.67
Italy	1973	100.8	1973	792.6	1973	9.66
Japan	1974	22.7	1973	91.2	1978	11.68
Netherlands	1975	39.8	1972	170.8	1970	8.25
New Zealand	1976	26.8	1975	1081.3	1978	53.32
Norway	1977	60.8	1977	450.0	1976	20.32
Spain	1970	31.0	1972	893.4	1970	3.45
Sweden	1973	99.6	1973	192.4	1973	35.0
United States	1980	94.9	1980	2348.7	1975	44.0
W. Germany	1973	213.4	1973	417.2	1977	23.35

*Table 1*

whether the United States, in particular, is “over-lawyered.” Some even see this question as helping account for the so-called litigiousness of American law. Whatever the numbers, an increasing number of civil cases are settled either during or after trial. Lawyers play their role in filing cases, but most of their work is done outside the court, which, of course, means lots of work for lawyers. Further, although cases may never reach court, as Mnookin and Kornhauser (1979) put it, much negotiation takes place in the “shadow of law.” That is, lawyers, well or poorly acquainted with actual court cases, call attention to what is “likely” to happen if they go to court, not to speak of the delay and expense. So law, or at least imagined law, plays a dominant role even when never specifically called into play (Ewick and Silbey 1998).

A useful table (table 1), if read with caution, is provided after careful research by Galanter (1983, p. 53).

Although we do not provide the sources, the results differ in dependability and the care with which they have been calculated. Still, the contrasts, however crude, are revealing. As to civil cases, the United States is seen to stand toward the middle, exceeded by Australia, New Zealand, and Canada, with many others being much lower. The United States is at the lower end in judges, being exceeded by W. Germany, Belgium, Italy, and

Sweden. But when we come to lawyers, the United States far exceeds other countries, though Australia, New Zealand, Canada, and Spain are also richly supplied. As we note later in the discussion of civil law countries, many others not called “lawyers” do what the United States would call “law.” Such persons include notaries, government officials of many kinds, law clerks in private firms, and, in Japan, the very high proportion of those who take the exam but are not allowed to practice (they do just about everything lawyers in the United States do except represent clients in court).

In making such international or intercultural comparisons, one should bear in mind differences in conceptions of what is worth disputing over and indeed what a dispute is in the first place. The studies by Felstiner, Abel, and Sarat (1980–81) suggest that a great many, perhaps most, injurious experiences are never perceived as such, but rather thought of as simply part of the risks of living. Some proportion of these are seen as violations of some right, but even many of those are simply “lumped,” that is, borne with equanimity or tolerated as not worth pursuing because of time or costs. A small proportion are, however, charged to a specific causal agent, and if a person or collectivity, then become what are called “grievances.” Some small proportion of those in turn, if voiced, turn into claims that, if rejected, become “disputes.” In turn, most (90 percent approximately)

disputes brought to lawyers are settled with varying degrees of satisfaction, lawyers often functioning to persuade the aggrieved that they should accept a settlement and move along with their lives. Whether persons are willing to pursue a grievance depends on technology and the ability to find a causal agent, as well as the existence of such legal devices as no-fault automobile insurance or divorce, which have the effect of diverting cases out of the legal system (see Kritzer, Bogart, and Vidmar 1991). Some countries, such as many on the European continent, have special labor tribunals and other systems that also divert cases out of what would, in the United States, be a legal case. The United States has fewer such alternative forums than is the case in the civil law world.

Jacob (1996, p. 52) speculates that the United States is, perhaps, more a nation of strangers, leading to a greater willingness to pursue disputes than is the case in countries with a stronger sense of community. In general, it may be said that the closer persons are, whether as family, neighbors or co-religionists, the less likely are they to sue one another. On the other hand, that does not mean there are fewer conflicts in such groups. Rather, there are internal mechanisms for settling them *within* such groups.

A final point about common law systems that contrasts with civil law systems is the widespread availability of appeal, particularly within the judicial systems. Although much of this is perfunctory and may involve little more than an attempt to satisfy clients with the appeals court routinely affirming the lower court, still appeal is possible, much more so than in civil law countries. Appellate judges often lack experience and are not required to have experience as trial judges. In some cases, judges have discretion on whether to hear an appeal, leading to selection of cases that may be controversial or present novel points of law. The extreme is presented by the U. S. Supreme Court which, in the 1990s, has elected to consider around 100 out of some 5,000 cases presented to it, usually reserving to hear constitutional cases and conflicts between the states or foreign governments. It is difficult to assess the impact of appeals on the civil or criminal process. Unlike trial judges, appeals judges do not merely decide cases but also give reasons. Such reasons are often examined by elite lawyers and are now

routinely discussed in the “legal” columns of popular magazines. It is not clear that the reasons affect policy in any obvious way. But the language of the court enters common discourse and affects thinking. School boards, church councils, Boy Scout boards, and even teachers’ decisions on classroom discipline become legalistic, with persons being given notice of charges, given chances to answer, and allowed to bring witnesses in their defense. In the United States, as in common law countries, the law seems to be everywhere (Galanter 1983), even if not formally invoked.

### THE CIVIL LAW TRADITION

In drawing comparisons between the common law and civil law traditions, it is important not to dismiss variations as due simply to “historical experience” or to even vaguer influence of “culture.” History and culture are, of course, operative at all times, but we seek not simply a description but a sociological explanation. We must begin with the recognition of what Zweigert and Kotz (1987) call “functionalism.” By that they mean that all legal systems deal with generally similar problems as, say, medical systems do. Whether the society employs witchcraft, herbs, appeals to the gods, leeches, hot baths, or Western-style x-rays and surgery, they all deal with illnesses of the body. So, too, legal systems concern themselves with trouble presented by the fact that humans live in society and must deal with each other. In Chiangmai, Thailand, for example, it is not surprising to find that three main classes of law suits appear: crimes as offenses against public order or the state, which are dealt with seriously by the courts; private wrongs, such as those arising from marital disputes, which are settled by negotiation; and those conflicts involving contracts and property rights, which are settled by careful examination of written and especially certified documents (Engel 1978). The details are indeed “cultural,” involving what Watson (1977) speaks of as “legal transplants,” by which a society adopts some procedure borrowed from another society because it is accessible, written in a language the elite can read (such as Latin for Roman law), or more commonly simply the law of a conquering power as with England or India. One wastes one’s time if one looks for rational reasons or tries to account for such transplants on grounds of “efficiency,” although often people come to

believe in the superior efficiency of the system of law they happen to use.

On the other hand, a close examination of legal procedures can have much to teach about the assumptions taken for granted in the culture (Ross 1993; Nelken 1997). In a report on a personal experience in Indonesia, Lev (1972) tells of an accident in a hotel in which a toilet tank, affixed high up on a wall, fell, nearly hitting a friend. To Lev's amazement, the hotel presented him and his friend with a bill for repairs. Lev refused, turning for help to a local judge, who was also a friend, for support for what Lev felt were his legal rights. The judge, while agreeing that those were indeed his legal rights, proposed that Lev make a token payment as evidence of "good will." With reluctance, Lev did so since compromise or peace was, the judge reminded him, after all, more important than vindication of rights. An even more obvious example is presented in a Mexican case in which the supreme court absolved a court for liability for the theft of money and jewels left in the care of the court pending settlement of the case. After all, said the court, Mexico is a poor country that cannot afford safe deposit boxes or secure storage places, but they do the best they can. The court then quotes what it clearly sees as a universal "principle of law," "*impossibile nulla obligatio est*," which the court translates simply as "No one is obligated to do the impossible" (quoted in Merryman, Clark, and Haley 1994, p. 684). These considerations are spoken of by Glendon (1987) as the "hortatory" function of law in civil law systems. She contrasts that with a view that American and British law usually involve a command backed up by punishment. Yet, whatever the system, laws, whether self-consciously doing so (as in civil law systems) or inadvertently (as in common law systems), always teach lessons as persons observe their operations.

Understanding of the civil law systems requires recognition that they come to us in two widely separated parts. The first is what is owed (and that is a great deal) to Roman law as codified in the sixth century under the Emperor Justinian as the *Corpus Juris Civilis*. This magnificent collection includes the law of persons, family, inheritance, property, contracts, and remedies, all of which the juriconsults (the legal experts) of the day saw as forming a unified body of law and which has been largely seen that way ever since. The influence was not simply on the Civil law system as such

but has had a strong effect on civil law (narrowly conceived) in common law countries as well. Basic principles the Roman jurists developed echo through the ages up to the present.

With the invasions of Rome that followed, much of this law fell into disuse or was united with the local laws of German tribes. However, *canon law*, as developed by the Catholic Church for its own uses, came to be widely adopted and grafted onto classic Roman law, influencing family law and civil procedure as well as much else, though not public law for the most part. Then, with the Renaissance, classic Justinian law was revived, especially in Bologna, where scholars gathered from all over Europe to study it, in Latin of course, and then spread it, where it came to be known as the *jus commune*. However, as it spread, it was inevitably influenced by local laws and customs that were often simply added to it in the interests of utility for solutions of local problems. A third development was *commercial law*, also from Italy, at about the time of Crusades, when there was much transport of goods and persons. The guilds and towns that developed this law were, for the most part, only tangentially influenced by Roman law since the focus was on rules merchants developed for their own use. Such rules spread even more widely than the *jus commune* up to the present day, where much of it can be found in admiralty law and related fields. It is quite clear, for example, that when two ships approach one another on the high seas there had better be a clear understanding as whether they both keep to the right or to the left, and that understanding must be clear whatever the differences in language, culture, or tradition. It is the nearest thing law offers to a truly international and intercultural system.

But there is more to modern civil law than a revival and Renaissance enrichment. The system was almost totally transformed by the ideas that gave birth to the political revolutions that began in the seventeenth century and are far from over at the present day. The revolutions were, of course, the American and French Revolutions, the military and ideological events associated with the unification of Italy and Germany, the coming new nations as the Turkish Empire disintegrated, the movements for freedom from Spanish and Portuguese domination in the Americas, as well as the chaos that followed the great wars of the nineteenth and twentieth centuries. While the word

“chaos” may be useful in a strictly descriptive sense, it is better to speak of these changes in the term employed by Schumpeter (1976) as winds of “creative destruction,” for they did not simply destroy but created what we speak of as the modern world. What they destroyed was basically feudalism and the concept of status fixed at birth as well as the conception of a divinely ordained, and hence unchangeable, social universe. In Weber’s classic phrase, we witnessed a “disenchantment of the world” which left humans, in the words of the existentialists, fated to create their own world and take responsibility for it. This meant that law was secular, torn loose from any religious basis, but focused instead on what came to be called “positive law,” that is law enacted by legislatures and parliaments. The transformation was not merely procedural but involved substantive changes in the assumptions of legal rights. These were the now-familiar rights stated in the American and French revolutionary documents—rights to liberty and property, the opportunity to change one’s status through one’s own efforts, the right to own land in one’s own person rather than merely, as in feudalism, as a serf or dependant of a feudal lord. Along with these changes came fundamental changes in loyalty and allegiance. Instead of fealty and subordination to lord or guild master, allegiance came to be narrowed to a single, overarching focus on the nation-state. The power of the church was similarly destroyed or greatly attenuated, as in England, and with that decline went the jurisdiction of ecclesiastical courts as well, though some of the traditions of canon law as, for example, in civil procedure (where courts make use of written records of proceedings and much less, than in common law systems, of oral testimony in trials), were retained.

Although the state and the law created by legislators came to make up the substance of law, it should be noted that in time it became evident that some controls were necessary on the state itself. In the United States, this control is institutionalized in the doctrine of the separation of powers, especially the ability of the courts to rule on the constitutionality of legislative enactments as well as the authority and legality of administrative acts and regulations. The civil law countries also elaborated a separation of powers but a very different one. The concern there was the enormous power judges had during the feudal period to act, usually in

support of the landed classes and the aristocracy. As Stone (1986) points out, the French *parlements* (panels of judges) had almost limitless power—they could arrest seditious persons, ban public gatherings, evaluate regulations of all kinds, supervise guilds and universities, and act as censors of public morals. Somewhat similar powers were enjoyed by the *audiencia* as representatives of royal power by the Spanish conquerors of Latin America. Such power led to their becoming wealthy and powerful, which led, in the case of France, ironically, to their own undoing. Although their vast powers might (and did occasionally) act as a break on royal powers, instead in a final act of defiance of the royal power, they threatened to resign on the very eve of the French Revolution. The Constituent Assembly voted to place them on indefinite vacation and then abolished them altogether. In a sense, their very arrogance and posturing led to a recognition that they would be a permanent obstacle to the new freedoms the revolutionaries wished to establish. A result was that there was a serious attempt to reduce the judge forever to a mechanical figure who would simply carry out the expressed will of the parliament in the name of the people. As such, the judge would have no inherent powers at all but would become a clerk or servant. He was not to presume even to interpret the will of the parliament. But how was that to be achieved?

The answer was to create a code that could answer all legal questions for the judge. In terms of the triadic model, two adversaries would argue before a neutral third who would simply delve into the code, find the answer, and impose the solution on them. In practice, as we can see from our vantage point, matters could never be so simple. As time went on, the concept of a legislative or code monopoly of law gave way to systems where by the judge could declare legislative or administrative acts unconstitutional, but the process involved much hedging by being careful, at first, at least to locate the places in which such review could take place outside the ordinary court systems in special constitutional courts (often not even called courts) as in France, Germany, Italy, and Spain, and in most Latin American countries, though in the latter, more influenced by U. S. practice, they were less reluctant to call them courts. There was little of this problem in England, which changed more slowly, retained more feudal practices, and, most important, did not go through a

bloody revolution to achieve the doctrine of parliamentary superiority (though the British did cut off at least one head and became, for a time, a republic).

Although one often speaks of common law systems as made up of cumulative, judge-dominated case systems, and Civil law countries as code systems, Merryman (1985, ch. v) reminds us that the contrast is overdrawn and misleading. All American law students are forced to master the Uniform Commercial Code, and many states routinely refer to their laws as “codes.” So too, there are code nations, such as Hungary, that actually did not enact a code until it became a socialist state, although it was a Civil law country before then. Instead, what is distinctive of codes in civil law systems is that they are unified documents that seek to express the spirit, ideology, and goals of the new state the revolution has created. Thus, the French code, the *Code Napoléon* of 1804, sought to express the ideology of the French Revolution—liberty, equality, and fraternity—in every clause. It was intended to be a blueprint for a utopia. Thus, every attempt was made to abolish or at least hide any earlier statutes or laws that were inconsistent with it and try to make a fresh start. Law would now begin with the *Code Napoléon*.

Further, in lines with the French *Declaration Rights of Man and of the Citizen*, the *Code Napoléon* must be one that the average Frenchman could read and interpret for himself without “humiliating” himself by going through clerks, officials, and other overlords to get to the courts. As such, lawyers would be unnecessary. For this to be possible, the code must be complete—without gaps. Everything would be covered. Although manifestly impossible, the Germans did make a valiant attempt to do so in the Prussian *Landrecht* of 1794, which laid out some 17,000 detailed “fact situations” that were felt to cover everything that could come up, thus eliminating any need for lawyers or interpreters. It failed, but it is a striking illustration of how persuasive was the ideology of the French Revolution, which created the belief that it could be done. In the *Code Napoléon* and others following it, the goal of completeness is achieved but only by broad statements that practically invite judicial interpretation. (For example, the Italian Civil Code of 1942 tells judges to follow the intention of the legislature, and if it is not entirely clear, then to reason by “analogy.”)

Germany, under Bismarck, did enact a code in the full sense but in what can only be seen as a very Germanic manner. The *Code Napoléon* began with certain assumptions about human nature (equality, liberty, etc.) and tried to produce a humanistic code that would, presumably, have universal application. Under the influence of Savigny, a major German historian, that approach was felt to be inappropriate. Instead, he insisted (in the face of heated controversy) that the German code (and he did agree that a code was necessary), since it was intended to represent the spirit of German society, must be based on the German *volkgeist* (folk spirit). But it was first necessary to decide what that was. To that end, and with help of German romantic writers, he felt it necessary to plumb German history for the basic elements of the *volkgeist* and build the code up from those elements. That code would then be not only historically oriented but also scientific (in being built up by logical and empirical deduction from basic principles) and professional. This code would not be revolutionary—quite the contrary—but would be a true code in being built up, paragraph by paragraph, from principles that could stand on their own as a legal document and manifesto of the new Germany. One difference from the French code was that with its complexity and its dependence on the many historical details that went into the *volkgeist*, it would require lawyers to explain and interpret it. Nevertheless, it was careful, like the French, to make sure judges would have little power, perhaps even less than was the case in France. For answers, the German litigant, with the help of his lawyers, was to go to the code and, above all, not to seek answers as American and common law lawyers do. That would just return power to the judges again.

Codes were also enacted in the many countries in Europe and elsewhere that followed the French or German systems (Japan tried doing both, with a dollop of the U. S. model thrown in as well (Haley 1991), though the German model eventually triumphed). When a code was shown to have gaps, scholars (in keeping with the tradition of drawing on the juriconsults in Roman law) would develop a new principle, as in an example provided by Watson (1981), wherein a doctrine similar to the British concept of *estoppel* was developed to cover cases where a person had acted contrary to his usual practice but others had come to rely on this new behavior. But when a new

interpretation (rather than a gap) was the problem, the court would draw on similar cases, not as precedents but to use as a basis for a new principle that would be held to govern the case at hand. In this manner, the spirit of a code based on permanent principles would be maintained.

The German approach to codification had a lasting effect in emphasizing the dominant role of the scholar in civil law systems generally. Although the scholars were everywhere evident, in the case of Germany, Savigny and his followers felt that in creating what they considered to be a code based on “scientific” principles they were creating a body of law that was indeed scientific in a sense not unlike that of the physical sciences. It was built up from empirical elements, could be found to be true or false, subject, as any science is, to modification as new facts came in. It came to be called “legal science,” which remains the dominant school of thought even up to the present, however much criticized. It had its own concepts, such as a “juridical act,” and was systematic in structure and therefore an infinite distance from such American schools of thought as “legal realism.” Above all, the Pandectists (as they came to be known, from the Latin word for Justinian’s *Digest–pandectae*) felt their great strength was their purity in being divorced from politics and everyday life.

Actually, as Merryman (1985, pp. 65) points out, the Pandectists were far from being value free. The doctrine was shot through and through with the basic assumptions of nineteenth-century European liberalism—private property, liberty of contract and, above all, individualism. They were most limited in their concept of law as a matter of transactions between private individuals, an assumption that was to collapse in the growth of giant collectivities, such as corporations and labor federations and, most important, the increased role of the state in managing economic and social life. Although civil law systems recognize the distinction between private and public law, even dividing up law in just that way, they hardly anticipated, nor could they, the merging of the two systems as states began to manage private life, and as private relations became imbued with public consequences as with pollution, the spilling over of populations across borders and, still later, the emergence of new national groups or even nations and new communities such as the European Community (cf. Gessner, Hoeland, and Varga 1996).

It is clear from the above that the Uniform Commercial Code in the United States, though called a code, has nothing in common with the civil law codes. It is not animated by any underlying utopian principles, it makes no claim to answer all questions, and it makes no attempt to supersede any laws. Instead, it is a collection that seeks to bring some order into the many elements of commercial law. States are free to ignore it (though few do so), and new laws can be tacked onto it at any time. It remains judge-made law, with judges being free to draw on it or not for precedent as they please.

Legal science did leave one imprint on American law, though a minor one. *Case law*, as taught in American law schools, was thought of as a kind of science, with cases as the raw materials. Conclusions from case accumulations might generate principles with wide applications. An attempt to state these principles took the form of what were called *restatements*, which are, from time to time, quoted by judges as they go about making law.

## JUDGES

As we have noted, the position of judges in civil law countries is vastly different from that in common law countries, especially the United States. Merryman (1985) points out that judges are not only respected in the United States but that some, such as Marshall, Holmes, Brandeis, and Cardozo, are culture heroes. The opinions of U. S. Supreme Court judges are studied carefully, as noted earlier, for hints of policy changes and guidance on how to proceed in deciding on difficult issues such as euthanasia, product safety, and whether schools may require bilingualism of its teachers. There seems no limit to areas into which judges may wander. Nor do judges hesitate in passing judgment on any law (if appropriate, of course) or even on the private life of the President of the United States. As has often been noted, American law is judge-made law built up from cases that lawyers have presented for decision (and, note, the judge must limit himself to such presentations, since he is very limited in his power to bring up issues on his own initiative). The supreme doctrine is that of *stare decisis*, whereby judges are required to follow decided cases; new cases must be compared to those already decided. If similar facts, then a similar decision. If the facts are significantly different, then there is a different decision. As noted by



Clark (Merryman 1991, p. 898), this need to do such case research, which falls on the shoulders of lawyers, helps explain (though only partly) the fact that the United States has more lawyers per capita than any other country for which we have reliable statistics.

Nothing could be further from that image among civil law judges. To begin with, the status of “judge” is usually much lower than is the case for common law judges. A civil law judge is a government employee, a civil servant, appointed to his position, whose career will follow that of other civil servants in rising by seniority and merit. His prestige is not necessarily low but reflects the prestige of civil servants at his level. He likely identifies with other civil servants, though more closely with judges, resulting in a certain insularity from the general public and its concerns. He is particularly isolated from any creative role in decision making. In line with the continuing suspicion of the dangers of judicial power going back to the *parlements* of pre-Revolutionary France (as well as similar excesses in other countries), he must not interpret the law or review legislation. As noted, he is a kind of expert in the application of the law to particular cases. More recently, constitutional review has begun to make its appearance in Austria, Spain, Italy, and Germany, but this goal is achieved not through giving judges in the ordinary courts new powers but rather through the creation of special constitutional review bodies that often are not called “courts” but that in time perform court functions.

This is not to say that appeal from judicial decisions was or is impossible. Quite the contrary, appeal is common but still dominated by attempts, at least in form, to restrict the power of judges. Thus, in France, appeal for what is claimed to be a misinterpretation of a law may be made to the Court (originally called a tribunal) of Cassation which could quash an incorrect interpretation by a lower court. It would indicate the correct interpretation but then remand it to the lower court to modify its ruling. The courts could ignore this decision but in practice would rarely do so. The process was time-consuming, and, in the case of Germany, the higher court would not only quash the lower decision but go ahead and revise the decision itself. It was also possible to appeal administrative rulings, though not, as in the American case, by declaring a law unconstitutional or

lacking in validity because of vagueness or for being over-broad. Instead, another body outside the court system was created. These might be what would amount to administrative courts, with a council of state at the top, a process seen not only in France but also in Italy and Belgium, as well as in Germany and Austria, where they were actually called administrative courts. Such attempts to counter state power were not necessary in a country such as England, where courts have the power of *quo warranto* (questioning the legality of an act by a public official) and *mandamus* (the ability to order a public official to perform as required by law). As noted, similar powers are possessed by ordinary American courts, though it is difficult for courts to use them.

The attempts to control the power and initiative of judges was tied up with another concern of civil law traditions; namely, the search for certainty. As long as judges had interpretive powers, the law was a tool in their hands that could be twisted to suit particular interests. Instead, the hope was that if the code plus legislation was clear and complete the judge would not need to exercise any initiative. Such a concern with certainty is not foreign to common law, either—persons need to know what the law says for law to be a guide to behavior. However, in civil law there is little that resembles the concept of equity at law. *Equity*—the power of a judge to limit the harshness of a law or to adapt the law to fit particular situations—gives him great powers. In England, equity reached its greatest development in the creation of chancery courts as a way of appealing to the king against what was felt to be an unjust rule of law. Civil law countries, though occasionally, and grudgingly, conceding a place for equity, preferred to confine it to the legislature, which might grant equitable powers to a court for a particular case or might make what amounts to an equitable grant of power to a court by telling it that, when the law is unclear, the court is to see to it that the parties acted “in good faith.” But the suspicion of judicial discretion remains and is not always a simple prejudice. Thus, the Nazi regime in Germany was able to make use of such discretion by using the courts to provide a patina of legality to its racist decrees, a process more difficult in Mussolini’s Italy, where discretion was more restricted.

One other important difference is that the English and American courts include in equitable

powers that of *contempt*. A litigant or witness who refuses to follow court processes or who refuses to carry out the will of the court may be subject to the contempt power, which can include fines and imprisonment. This degree of power is quite unknown in civil law countries where it is felt to give the judge what amounts to discretion to impose criminal penalties in civil cases. The civil law judge must, in comparable cases, limit himself to drawing on the person's property which may, of course, be felt by the person, as no less painful than a period in prison might be for others.

### THE LEGAL PROFESSIONS

In the United States, lawyers, as well as the general public, think of the legal profession in the singular, though specialties are recognized. This is especially the case, as noted earlier, since such a high proportion are in private practice (well over 70 percent) as compared, for example, to only 33 percent in Germany, 42 percent in Colombia and only 23 percent in Chile. The United States also has a very low proportion of the profession acting as judges (only 3 percent) compared to 17 percent in Germany, 23 percent in Chile, and 42 percent in Colombia. (Clark 1982, figures are for the 1960s and 1970s). In Germany, 70 percent of positions in general administration are filled by persons trained in law.

Apart from differences in distribution, the path to a legal career is very dissimilar in different cultures. In the United States, aspiring lawyers go to a graduate law school, then sit for the bar exam (often after an intensive cram course that prepares them for that exam), then after passing the bar (in most states, the success-rate percentage is over 60 percent, and higher for first-time exam takers) he or she enters directly into practice in a firm, in a small-firm partnership, or as a solo. There they learn as they go along. Persons may, and many do, shift around from service in a government department to a public prosecutor office to the corporate law office of a private firm, or elsewhere. They may run for political office and may end up being rewarded for service by appointment or election as judge of a lower-level court and, for a few, high judicial office in a circuit or even a supreme court.

In England, the distinction between solicitor and barrister, though less rigid than in the past, continues. The would-be barrister takes his pupillage

under a barrister in one of the Inns of Court where he may, under good conditions, receive an apprenticeship and possible appointment after passing an examination. Only barristers may argue cases in the higher courts. Solicitors maintain direct contact with clients, collect fees, and assist barristers in their work. Solicitors are now being allowed to argue cases in some lower courts. Some barristers may develop honored reputations, a few being chosen as judges as the culmination of a distinguished career. There is almost no shifting to other careers on the part of barristers and very little among solicitors. In that respect, they resemble the lawyers in civil law countries.

In civil law countries, a young lawyer must make an early choice as to whether he wishes to be a judge, a government lawyer, a regular lawyer in private practice, a public prosecutor, or a notary. If he finds later he made a mistake, exit to another legal career is difficult, and does his experience in one does not translate into credits in another. He spends his whole career in the one field, a process that often leads to rivalry and conflict between the fields. An attempt to deal with this problem is made in Germany (and some other countries) in the *referendarzeit*, where a lawyer spends two or so years of practical training as a government lawyer, a judge, and in private practice. If lawyers choose to be judge, they will begin their career in a low-level court but may move up as openings occur. Although this career process isolates the judges who take on guild-like characteristics as civil servants, it also means that judges are often better trained than is frequently the case in the United States (where it is not uncommon for judges to have had no judicial experience whatsoever). Further, the quality of judging may be higher since candidates are chosen from among the best law school graduates. At the top in constitutional courts, for example, the quality of decision making is the equal of that found anywhere.

Public prosecutor in Civil law countries are much like U. S. district attorneys, but they also are required to represent the public interest in proceedings between private persons in court situations. In Italy and France, the prosecutor is also a member of the judiciary, allowing some shifting back and forth from prosecutor to judge. Some degree of shifting back and forth also takes place in Germany. Those lawyers working for the government in administrative positions are career

bureaucrats. Closest to the U. S. lawyer in private practice is the French *avocat* (distinctions such as those between the *avoués*, who acted as solicitors in appeals courts, and the *conseil juridiques*, who give general advice and represent clients before commercial courts, are gradually being eliminated or reduced in significance), the German *anwalt* or Italian *avvocato* (Merryman et al., p. 918). Much different from that known in the United States is the notary who receives legal training and is an important person. He drafts important legal instruments, such as corporate charters, wills, and instruments transferring land, as well as contracts. Most important, he authenticates documents. Once he does so, the instrument is accepted in court without further question. They also have monopolistic control over assigned territories. In Germany the services of a notary are required to validate legal documents for purchase, sale, and mortgage of land, for official records of decisions of company meetings, and for sale of shares in a private company (Merryman et al., p 911). Academic lawyers are found in a law school where they carry on the tradition of the old Roman juriconsult. However, most academic lawyers work for a professor with little or no pay, and wait for a vacancy that may never come. In Latin American countries, such persons may hardly earn a living, having to take on regular work as a lawyer in private practice or in public office.

#### CIVIL AND CRIMINAL PROCEDURE

Something should be said, in brief, about variations in procedure between the two systems. Here, the word “civil” is used in contrast to criminal. There is no trial or jury in civil cases, as may often be the case in the United States, though not in England. The entire process is different. The presence of a jury in the United States forces an acceleration of the entire process because of the difficulty and expense of getting the jury assembled and empanelled. Once that is the case, the court proceeds immediately with the trial in an attempt to conclude as quickly as possible.

In Civil law jurisdictions, civil cases go much more slowly. There is a brief preliminary stage when pleadings are submitted to a hearing judge. Next follows an evidence-taking, where the hearing judge takes notes and prepares a written record. That is later submitted to the deciding judge

who receives briefs from the counsel and listens to arguments. All of this takes the form of a series of meetings as each issue is brought to the attention of the hearing judge. There is little surprise as each lawyer is notified of each issue as it comes up, and, without a jury, there is no cross-examination. Generally, questions are passed to the judge who may conduct the investigation. There are fewer of the rules of evidence familiar to American lawyers (such as the exclusionary rule whereby illegally gathered information is excluded from trial), though a number of rules are employed, such as excluding biased persons from testifying, as well as taking what is called a “decisory oath” in some countries. There is in many countries a “loser pays” rule, referring mostly to legal fees, though the amounts are usually limited by a court schedule. Contingent fees are usually considered illegal (France) or unethical (Germany) but are found in Japan, Indonesia, and Thailand (Merryman et al., p. 1026). Many foreign legal authorities are appalled by its prevalence in the United States, feeling that a lawyer should not be personally given a stake in the outcome of a case.

Although substantive criminal law is similar in both systems, civil law jurisdictions, in line with the revolutionary principle of limiting the power of judges, reject the American practice, which gives the judge power to award penal and general damages in criminal cases, not to speak of the contempt power (which is very rare), instead insisting the judge be limited to what is provided for in legislation. The contrast is often drawn (or overdrawn) between what is called the “accusatory” model in the United States and the “inquisitorial” model in civil law countries. Historically, the accusatory procedure is felt to have been a development from that of private vengeance in which the interested parties would be the main participants. Instead of settling their dispute by direct conflict or feud, a legal procedure would involve a neutral third party who would seek to secure a settlement. In such a triadic situation, as noted at the outset, the object was to secure an outcome that would settle the matter by leaving each party feeling justice had been done. In earlier times, when trial was by battle, ordeal, or other ways, although seeming to be a throwback to a time of “barbarous” cruelty, these methods, apart from their presumed psychological effects (the guilty party

feeling he would be caught and so offering confession at once), had the virtue that they brought the conflict to an end, and prevented further acts of vengeance that would disturb the peace of the community. With the accusatory practice, presumably conflict would also be terminated but might go on if each party felt justice had not been done.

The inquisitorial model introduced the state as an active participant in the trial. Now the judge, who is after all a representative of government, is in a coalition with the prosecution against the third party, the defendant. Although this biases the process, in practice, the introduction of the jury, the fact that proceedings are oral, as well as limitations on the power of the judge all combine to make the system fair, though excesses did and continue to exist.

The criminal trial is in three parts: the investigative phase, with the public prosecutor assuming an active role; the examining phase, presided over by a judge who assumes an active role in examining the evidence; and preparing a record and the trial. Judges may, if warranted, end the proceedings if they feel the evidence is not conclusive, or they may decide the case should go to trial. The accused is entitled to legal representation as well as the right to inspect the material the judge has collected. He can be questioned, but not sworn, and he or she may refuse to answer. Unlike the American system in which a defendant, if sworn, may then be cross-examined, no comparable procedure exists, though a refusal to answer by the accused may be taken into account by the jury. British judges assume a more active role in the trial process than is the case in the United States. However, this is not “inquisitorial” in a narrow sense but rather a reflection of the fact that in the English procedure the judge determines the relevancy of evidence, rather than the strict exclusionary and other rules emphasized in American courts. To do their jobs, English judges are much more willing to question witnesses or even raise issues (Glendon et al., 1982, ch.. 10).

Until the 1980s, plea bargaining was considered to be undesirable practice, only possible in America. But judicial scholars increasingly asked how could civil law jurisdictions possibly handle all the criminal cases they had to decide without some form of plea bargaining? After a spirited debate and careful research, it was finally concluded that

plea bargaining, though not exactly like that used in America, was in fact being used in European countries. In the case of Germany, Hermann (1991) reports that some kind of plea bargaining takes place in from 20 to 30 percent of all cases. While far from the American percentage of 90 percent, that is still considerable, particularly since it is accompanied by other forms of sentence reduction and mitigation of offenses. It is widely employed in complex cases, such as white-collar crimes, tax evasion, and drug offenses, which present nearly impossible evidentiary problems, as well as less serious crimes, which are settled with a fine. It is rare in cases involving violent crimes, however. Bargaining occurs at all stages of a criminal proceeding, often with the active participation of the judge, who may even take the initiative. A settlement may take the form of the accused agreeing to pay a sum of money to a charitable organization. Other alternatives include penal orders that are similar to *nolo contendere* pleas in which the accused agrees to a fine—usually for minor misdemeanor cases, such as traffic cases. Pleas also occur if the accused makes a confession. Normally, a confession does not lead to avoidance of a trial, as is the case in America. Instead, it usually leads to a reduction in the sentence. Other countries are also beginning to allow plea bargaining as many had been doing, but are now doing more openly. The most striking example is that of Italy, which even uses the term *patteggiamento*, the Italian word for bargain (Piazzi and Marafioti 1992; Merryman et al., pp. 1100 ff). However, there is no reduction of the charge, as in the American system, but there is a maximum reduction of one-third of the normal sentence, which may not exceed two years, which has the effect of limiting the practice to cases with shorter normal sentences. In Italy, as is the case in other civil law countries, the trial decision is made by a panel of judges, a practice which, however costly, is defended as superior to the American practice, which places what is felt to be excessive power in the hands of a single judge.

#### **WILL COMMON LAW AND CIVIL LAW SYSTEMS PERSIST?**

Although we have focused on differences between the two systems, and the differences are indeed substantial, there is considerable movement toward the convergence of common and Civil law systems. The attempts to severely limit the power

of judges have, over the years, been recognized as excessive and more a holdover from fears of arbitrary and class-based favoritism of judges. From the refusal to place limits on the power of legislatures, Civil law countries have developed constitutional limitations, even in the form of courts, though often hidden under other names. Historically, Civil law countries divided the field between private and public law, with most of the concern historically being with civil law (law of persons, marriage, contracts, torts, etc.). Public law was felt to be the concern of legislatures or the sovereign. However, the coming of the modern state has led to an enormous growth of administrative law, along with appropriate court systems, leading to situations not much different in essence from those found in developed common law countries.

Civil codes have receded in significance as state legislatures and parliaments enact more far-reaching laws never contemplated in earlier times, particularly those associated with large-scale industry, the welfare role of governments, complex bodies of labor law, and especially with the emergence of the government as an economic participant in national affairs. Perhaps of greatest significance is the emergence of new international entities, especially the European Community, before which national codes have been slowly giving way. Of course, that tendency may be reversed, but it seems strongly in process. On the other hand, common law countries have recognized the advantages of classic civil law procedures, such as the importance of certainty in legal decision making, though such certainty is sought not through a code but by a succession of cases. The power of judges to make decisions has been the object of attempts by the U. S. Congress to place caps on awards for injury, as well as the attempt of persons to bypass courts and seek justice through changes in legislation, as in cases of pollution, abortion, and other public issues.

Is one system better than the other? The question is unanswerable in that form. Even though civil law authorities continue to be suspicious of lawyers, they discover they must have them and grant them powers they would prefer not to. That is, of course, a well-known problem in the United States as well. Each system must be understood in the context of the culture and social structure of the country that employs it. Ultimately, any question of superiority must be answered in terms of

how well the system serves the legal needs of the country. Most countries, in fact, employ a mix of systems in which one finds bits and pieces of both common law and civil law systems.

## REFERENCES

- Atiyah, P.S., and Robert S. Summers 1987 *Form and Substance in Anglo-American Law*. Oxford: Clarendon Press.
- Black, Donald 1993 *The Social Structure of Right and Wrong*. San Diego, Calif.: Academic Press.
- Bohannon, Paul J. (ed.) 1967 *Law and Warfare: Studies in the Anthropology of Conflict*. Garden City, N.J.: Natural History Press.
- 1957 *Justice and Judgment Along the Tiv*, London: Oxford University Press.
- Boigeol, Ann 1998 "The French Bar: The Difficulties of Unifying a Divided Profession." In Richard L. Abel and Philip S. C. Lewis, eds., *Lawyers in Society: The Civil Law World*, vol. 2, 258 ff.
- Caplow, Theodore 1968 *Two Against One*, Englewood Cliffs, N.J.: Prentice-Hall.
- Clark, David S. 1975 "Judicial Protection of the Constitution in Latin America." *Hastings Constitution Law Quarterly* 2:405-442.
- 1982 "The Legal Profession in Comparative Perspective: Growth and Specialization." *American Journal of Comparative Law* 30, Supp.:163, 169.
- Damaska, M. 1978 "A Foreign Perspective on the American Judicial System." In T. J. Fetter, ed., *State Courts: A Blueprint for the Future*. Williamsburg, Va.: National Center for State Courts.
- Dworkin, Ronald 1977 *Taking Rights Seriously*. London: Duckworth.
- Eisenstein, James, and Herbert Jacob 1977 *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Littel, Brown.
- Engel, David M. 1978 *Code and Custom in a Thai Provincial Court: The Interaction of Formal and Informal Systems of Justice*. Tucson: University of Arizona Press.
- Ewick, Patricia, and Susan S. Silbey 1998 *The Common Place of Law*. Chicago: University of Chicago Press.
- Feifer, George 1964 *Justice in Moscow*. New York: Simon and Schuster.
- Felstiner, William, Richard L. Abel, and Austin Sarat 1980-1981 "The Emergence and Transformation of Disputes: Naming, Blaming and Claiming. . ." *Law and Society Review* 15:631-654.

- Fuller, Leon 1969 *The Morality of Law*. New Haven, Conn.: Yale University Press.
- Galanter, Marc 1992 "Law Abounding: Legalization Around the North Atlantic." *Modern Law Review* 55:1-24.
- 1983a "Mega-Law and Mega-Lawyering in the Contemporary United States." In Robert Dingwall and Philip Lewis, eds., *The Sociology of the Professions*. London: Macmillan.
- 1983b "Reading the Landscape of Disputes: What We Know and Don't Know (and Think We Know) About Our Allegedly Contentious and Litigious Society." *UCLA Law Review* 31:5-71.
- 1983c "The Radiating Effects of Courts." In Keith O. Boyum and Lynn Mather, eds., *Empirical Theories About Courts*. New York: Longman.
- 1974 "Why the Haves Come Out Ahead: Speculations on the Limits of Legal Change." *Law and Society Review* 9:95-125.
- , and Thomas Palay 1991 *Tournament of Lawyers*. Chicago: University of Chicago Press.
- Gessner, Volkmar, Armin Hoeland, and Csaba Varga (eds.) 1996 *European Legal Cultures*. Dartmouth: Aldershot.
- Gibbs, James L., Jr. 1963 "The Kpelle Moot: A Therapeutic Model for the Informal Settlement of Disputes." *Africa* 33 (Jan.):1-11.
- Glendon, Mary Ann 1987 *Abortion and Divorce in Western Law*. Cambridge, Mass.: Harvard University Press.
- , Michael Wallace Gordon, and Christopher Osakwe 1982 *Comparative Legal Traditions in a Nutshell*. St. Paul, Minn.: West Publishing Co.
- Gluckman, Max 1954 *The Judicial Process Among the Barotse of Northern Rhodesia*. Manchester, Eng.: Manchester University Press.
- Gross, Edward 1998 "Lawyers and Their Discontents." *Society* 36 (Nov./Dec.):26-31.
- Gulliver, P. H. 1969 "Dispute Settlement Without Courts: The Ndendeuli of Southern Tanzania." In Laura Nader, ed., *Law in Culture and Society*. Chicago: Aldine.
- Haley, John Owen 1991 *Authority without Power*. New York: Oxford University Press.
- 1978 "The Myth of the Reluctant Litigant." *Journal of Japanese Studies* 4:359-396.
- Heinz, John P., and Edward O. Laumann 1982 *Chicago Lawyers: The Social Structure of the Bar*. New York: Basic Books.
- Hermann, Joachim 1992 "Bargaining Justice—A Bargain for German Criminal Justice?" *University of Pittsburgh Law Review* 53:755-776.
- Howell, P. O. 1954 *A Handbook of Nuer Law*. London: Oxford University Press.
- Jacob, Herbert 1996 "Courts and Politics in the United States." In Herbert Jacob et al., ed., *Courts: Law & Politics in Comparative Perspective*. New Haven, Conn.: Yale University Press.
- 1983 "Courts as Organizations." In Keith O. Boyum and Lynn Mather, eds., *Empirical Theories About Courts*. New York: Longman.
- Klein, John F. 1976 *Let's Make a Deal*. Lexington, Mass.: Lexington Books.
- Kritzer, Herbert M., W. A. Bogart, and Neil Vidmar 1991 "The Aftermath of Injury: Cultural Factors in Compensation Seeking in Canada and the United States." *Law and Society Review* 25:499-543.
- Lev, Daniel S. 1972 "Judicial Institutions and Legal Culture in Indonesia." In Holt Claire, ed., *Culture and Politics in Indonesia*. Ithaca, N.Y.: Cornell University Press.
- Levasseur, Alain 1970 "On the Structure of a Civil Code." *Tulane Law Review* 44:693-719.
- MacLachland, Colin M. 1974 *Criminal Justice in Eighteenth Century Mexico: A Study of the Tribunal of the Acordada*. Berkeley: University of California Press.
- Magnet, S. Rodgers 1980 "The Right to Emergency Medical Assistance in the Province of Quebec." *Revue du Barreau*, Tome 40, Numero 3, Mai-Juin, 373-447.
- Merryman, John Henry 1985 *The Civil Law Tradition*. Stanford, Calif.: Stanford University Press.
- , Davis S. Clark, and John O. Haley 1994 *The Civil Law Tradition: Europe, Latin America and East Asia*. Charlottesville, Va.: Michie.
- Mnookin, Robert, and Lewis Kornhauser 1979 "Bargaining in the Shadow of Law: The Case of Divorce." *Yale Law Journal* 88:950-997.
- Nader, Laura 1969 "Styles of Court Procedure: To Make the Balance." In Laura Nader, ed., *Law in Culture and Society*. Chicago: Aldine.
- Nelken, David (ed.) 1997 *Comparing Legal Cultures*. Dartmouth: Aldershot.
- Pospisil, Leopold 1958 *Kapauka Papuans and Their Law*. New Haven, Conn.: Yale University Publications in Anthropology, No. 54.
- Rosen, Lawrence 1989 *The Anthropology of Justice*. New York: Cambridge University Press.
- Rosenn, Keith S. 1992 "A Comparison of the Protection of Individual Rights in the New Constitutions of Colombia and Brazil." *University of Miami Inter-American Law Review* 23:659-691.

- Ross Marc Howard 1993 *The Culture of Conflict*. New haven, Conn.: Yale University Press.
- G. Roth, and C. Wittig 1978 *Economy and Society*. Berkeley: University of California Press.
- Rouland, Norbert 1994 *Legal Anthropology*. Stanford, Calif.: Stanford University Press.
- Sarat, Austin, and Thomas R. Kearns (eds.) 1998 *Law in the Domains of Culture*. Ann Arbor: University of Michigan Press.
- , and Stuart Scheingold (eds.) 1998 *Cause Lawyering*. New York: Oxford University Press.
- Schumpeter, Joseph A. 1976 *Capitalism, Socialism and Democracy*. New York: Harper and Row.
- Shapiro, Martin 1986 *Courts: A Comparative and Political Analysis*. Chicago: University of Chicago Press.
- Simmel, Georg 1902 “The Number of Members as Determining the Sociological Form of the Group.” *American Journal of Sociology* 8:1–46, and 158–196.
- Starr, June 1992 *Law as Metaphor: From Islamic Courts to the Palace of Justice*. Albany, N.Y.: State University of New York Press.
- Stone, Bailey 1986 *The French Parlements and the Crisis of the Old Regime*. (excerpted from Merryman et al., eds., 436–437).
- Stone, Thomas 1979 “The Mounties as Vigilantes: Perceptions of Community and the Transformation of Law in the Yukon, 1885–1897.” *Law and Society Review* 14:83–114.
- Watson, Alan 1977 *Society and Legal Change*.
- Wolff, Kurt H. (trans, ed., and Introd.) 1950 *The Sociology of Georg Simmel*. New York: Glencoe Press.
- Zamudio, Hector Fix 1979 “A Brief Introduction to the Mexican Writ of Amparo.” *California Western International Law Journal* 9:306–348.
- Zweigert, Konrad, and Hein Kotz 1987 *Introduction to Comparative Law* (trans. Tony Weir). Oxford: Clarendon Press.

EDWARD GROSS

## COURTSHIP

Given the social centrality of the family institution and the role of courtship in the family formation process, it is not surprising that the study of courtship has received attention from several disciplines. Anthropologists have described practices

in primitive and other societies, historians have traced courtship patterns in America from colonial to contemporary times, psychologists and social psychologists have examined intra- and interpersonal components of relationships, and sociologists have developed research-based theories explaining the process of mate selection, and have investigated various courtships dynamics. Here, some attention will be given to each of these approaches, along the way selectively noting scholars who have made major contributions.

Historically, according to Rothman, the term *courtship* applied to situations where the intention to marry was explicit (if not formally—and mutually—stated). *Courting* was the broader term used to describe socializing between unmarried men and women” (Rothman 1984, p. 23, italics in original).

Scholars have disagreed as to whether dating—a twentieth-century term for a primarily recreational aspect of courting—should be considered a part of courtship since, according to Waller (1938) and others, dating may be merely thrill-seeking and exploitative, and not marriage oriented (but see Gordon 1981 for an opposing view). However, wooing (that is, seeking favor, affection, love, or any of these) may be integral to courtship and yet not result in marriage. For present purposes, then, courtship will be understood in its broadest sense—as a continuum from casual to serious. Thus, “the unattached flirt, the engaged college seniors, the eighth-grade ‘steadies,’ and the mismatched couple on a blind date are all engaging in courtship” (Bailey 1988, p. 6).

Queen, Habenstein, and Quadagno’s (1985) classic text provides much of the basis for the following brief and highly generalized overview of some mate-selection patterns unlike those found in contemporary America. Some of these systems involved little or no courtship. For example, among the ancient Chinese, Hebrews, and Romans, marriage was arranged by male heads of kin groups. Among the ancient Greeks and until recently among the Chinese, many brides and grooms did not meet until their wedding day. Around the turn of the century (1900), infant marriages were the rule among the Toda of south India, and the bride was deflowered at about age ten by a man who was not of her clan and not her husband. In medieval England, contrary to the literature of chivalry, love

had little to do with mate selection in any social class because marriages were arranged by lords or by parents with primary regard to the acquisition of property.

In societies where romantic love is not a basis for mate choice, such sentiments are seen as dangerous to the formation and stability of desirable marital unions—those that maintain stratification systems (see Goode 1969). Queen, Habenstein, and Quadagno (1985) describe still other mate-selection patterns that do involve some form of love, including the systems found on Israeli kibbutzim at midcentury, among ethnic immigrant groups in the United States, and among African-Americans during slavery (see also Ramu 1989). In the twentieth century, however, and especially since the 1920s, courtship in Western societies has been participant-run and based on ideas of romantic love. In the United States today, it is not uncommon for a couple to meet, woo, and wed almost without the knowledge of their respective kin. “Compared with other cultures, ours offers a wide range of choices and a minimum of control” (Queen, Habenstein, and Quadagno 1986, pp. 8–9).

In colonial America, practices differed somewhat between the North and South. In the North, mate choice was participant run, but a suitor’s father had control over the timing of marriage since he could delay the release of an adequate section of family land to his son while the son’s labor was still needed. Conjugal (but not romantic) love was thought to be the *sine qua non* of marriage, and couples came to know and trust one another during often lengthy courtships. In the South, a custom of chivalry developed, closely guarding the purity of (at least upper-class) women, but condoning promiscuity among men. Parental consent was required for the beginning of courtship and for marriage and open bargaining about property arrangements was commonplace. Unlike the colonial North, where marriage was considered a civil ceremony, in most parts of the South, Anglican church ministers were required to officiate at weddings. In both regions, banns were published prior to weddings.

During the 1800s, mate choice became more autonomous with the growth of cities and the spread of industrial employment. Choices were affected less by considerations of wealth than by

personal qualities—especially morality, spirituality, and “character.” Wooing was rather formal, with each participant carefully evaluating the qualities of the other. Courtship tended to be exclusive and directed toward marriage.

Then, from about 1900 to World War II, a system evolved in which there was much “playing the field” (casual dating), gradually more exclusive dating (“going steady”), engagement, and finally, wedding—a relatively fixed sequence. Following the war, stages of courtship were typically marked by symbols (e.g., wearing a fraternity pin, then an engagement ring), each stage implying increased commitment between the partners. By the 1950s, a separate youth culture had developed. Ages at first marriage declined dramatically, and dating started earlier than ever before. The sexual exploration that had previously been a part of the last stage of courtship now occurred earlier, even in very young couples.

During the 1960s, a time of “sexual revolution,” nonmarital cohabitation increased—not substituting for marriage but delaying it. In the post World War II. period and since, among the young especially, demands for both freedom and dependence (i.e., the right to sexual freedom without assuming responsibility for its multifaceted consequences) have been relatively widespread. Concurrently, rates of nonmarital pregnancy rose dramatically.

In general, every society attempts to control sexual activity among unmarried (and married) persons, but the forms of control (e.g., chaperonage) and the degree of enforcement have varied. Virginity, especially in women, is highly prized and guarded in some cultures but has no special value in others. Similarly, all societies attempt to limit the pool of those eligible to marry, but the precise constraints have differences across societies and from time to time. Typically, blood kin and relatives by marriage (and in some cases, baptismal relatives such as godparents) are delimited to greater or lesser degrees from the eligibility pool.

Where male elders have arranged marriages for their offspring (generally in ascription-based societies), the accumulation of family power and prestige has been of primary concern, with dowrys, bride prices, or both figuring prominently in prenuptial arrangements. In participant-run mate



selection (generally in achievement-based societies), the power and prestige of a dating partner (although defined in terms other than land, cattle, and the like) is still valuable. Good looks (however defined) in women, for instance, are a status symbol for men, and conspicuous consumption in men (cars, clothing, spending habits) provides status for women. Thus, within the field of eligibles is a smaller field of desirables. Unfortunately, the qualities that are valued in dates (from among whom a mate may be chosen) are not necessarily those one would want in a spouse.

Even in participant-run “free choice” systems, there is a tendency toward homogamy in the selection of partners, whether conscious or not. As a society becomes more varied in its mix of persons within residential, educational, religious, or work-related settings, the tendency toward heterogamy increases. That is, the field of eligibles and desirables broadens. Heterogeneity leads to a prediction of “universal availability” (Farber 1964) as the salience of social categories (such as race, age, religion and class) declines. For example, interracial relationships, once unthinkable (e.g., in the colonial American South), increased with urbanization, industrialization, and a general movement toward educational and income equality. Social class endogamy, however, is the general preference, although women are encouraged with varying degrees of subtlety to “marry up,” and a dating differential exists such that men tend to court women who are slightly younger, physically smaller, and somewhat less well educated or affluent than themselves.

Contemporary courtship, marked as it is by freedom of choice, has been likened to a market in which the buyer must be wary and in which there is no necessary truth in advertising. Persons compete, given their own assets, for the best marital “catch” or the most status-conferring date. Waller and Hill (1951) warned about the potential for exploitation in both casual and serious courtship and indeed, critics of conventional dating have decried it as a sexist bargaining arrangement in which men are exploited for money and women for sexual favors. The superficiality of dating, its commercialization, the deceit involved (given contradictory motives), and the high levels of anxiety provoked by fears of rejection (especially in men), are additional drawbacks. Since status differentials still characterize the sexes, dating may also be seen

as a contest in which a struggle for power and control between partners is part of the game. Thus, courtship’s emphasis on individualism, freedom, commercialism, competitive spirit, and success reflects the larger social system within which it functions. One may well ask whether such a system can prepare participants for marriage which, unlike courtship, requires cooperation and compromise for its successful survival.

Efforts to predict who marries whom and why, to delineate the courtship process itself, or both, have interested a number of scholars. Based on a large body of theoretical and empirical work, Adams (1979) developed a propositional theory to explain how courtship moves from initial acquaintance toward (or away from) marriage in an achievement-oriented society. The propositions, in slightly modified language, are as follows:

1. Proximity, which facilitates contact, is a precondition for courtship and marriage.
2. As time passes, marriage is increasingly more likely to be with a currently propinquitous than with a formerly propinquitous partner.
3. Propinquity increases the likelihood that one will meet, be attracted to, and marry someone of the same social categories as oneself.
4. Early attraction is a result of immediate stimuli such as physical attractiveness, valued surface behaviors, and similar interests.
5. The more favorable the reactions of significant others to an early relationship, the more likely the relationship will progress beyond the early attraction stage.
6. The more positive the reaction of the partners to self-disclosures, the better the rapport between them.
7. The better the rapport between the partners, the more likely the relationship will be perpetuated beyond the early attraction stage.
8. The greater the value compatibility—consensus between partners, the more likely that the relationship will progress to a deeper level of attraction.

9. The greater the similarity in physical attractiveness between the partners, the more likely that the relationship will progress to a deeper level of attraction.
10. The more the partners' personalities are similar, the more likely that the relationship will progress to a deeper level of attraction.
11. The more salient the categorical homogeneity of the partners, the more likely that the relationship will progress to a deeper level of attraction.
12. The more salient the categorial heterogeneity of the partners, the more likely that the relationship will terminate either before or after reaching a deeper level of attraction.
13. The greater the unfavorable parental intrusion, the more likely that the relationship will terminate either before or after reaching a deeper level of attraction.
14. An alternative attraction to the current partner may arise at any stage of a couple's relationship. The stronger that alternative attraction to either partner, the more likely that the original couple's relationship will terminate.
15. The greater the role compatibility of the partners, the more likely that the relationship will be perpetuated.
16. The greater the empathy between the partners, the more likely that the relationship will be perpetuated.
17. The more each partner defines the other as "right" or as "the best I can get," the less likely that the relationship will terminate short of marriage.
18. The more a relationship moves to the level of pair communality, the less likely it is that the relationship will terminate short of marriage.
19. The more a relationship moves through a series of formal and informal escalators, the less likely it is to terminate short of marriage (Adams 1979, pp. 260-267).

Adams (1979) also provides some warnings about these propositions. First, some factors (such as partner's good looks) have greater salience for

men than for women, while some (such as partner's empathic capacity) have greater salience for women than for men. Second, some factors such as parental interference may have different outcomes in the long run compared to the short run. Third, the timing of courtship may bring different considerations into play, e.g., courtship in later life such as following divorce or widowhood, or when children from previous marriages must be considered (see Bulcroft and O'Connor 1986). Finally, social class factors may affect the predictive value of the propositions. There is also a difference between traditional (male-dominated) and egalitarian relationships—the former more often found in the working class and among certain ethnic groups, the latter more likely to characterize the middle class. Thus, the kind of marriage one anticipates (traditional/egalitarian) may influence the mate-selection process. (See also Aronson 1972 for specifications of the conditions under which various interpersonal attraction predictors such as propinquity and similar interests operate).

Further, as courtship has moved away from the fixed-stage sequence of development, it may be viewed best from a circular-causal perspective (Stephen 1985) in which progress is strongly influenced by communication within the couple, leading to increased or decreased movement toward marriage.

The timing of marriage may be influenced by such factors as meaningful employment opportunities for women (which may diminish their motivation to marry), the increasing acceptability of nonmarital cohabitation and adult singlehood (see Stein 1981), and the effects of nonmarital pregnancy or of various intolerable conditions (such as violence) in the family of origin. Currently, a number of scholars are studying each of these topics. They affect not only the timing of marriage but also how we define courtship.

Regarding premarital factors that contribute to later marital adjustment, no scholar has presented evidence to refute Kirkpatrick's ([1955] 1963) conclusions: The happiness of parents' marriage; adequate length of courtship; adequate sex information in childhood; a happy childhood including a harmonious relationship with parents; approval of the courtship relationship by significant others; good premarital adjustment of the couple and strong motivation to marry; homogamy along age,

racial-ethnic, and religious lines; and, later age at marriage.

Murstein (1980) reviewed mate-selection scholarship from the 1970s and predicted that researchers would focus less on the “old standby” variables such as race, class, and religion and more on the dynamic aspects of courtship. He was correct. Some of the major themes that have interested scholars in recent years are identified below.

Studies of cohabitation included early efforts to identify its several types (both structural and motivational). Later studies focused on the effects of cohabitation on subsequent marital happiness, satisfaction, and stability. The general finding across such research is that living with someone prior to marriage has little or no positive effect. Instead, most studies show negative effects in terms of happiness, satisfaction, and stability. This research has been carried out in the United States, Canada, and other countries, and although the rates vary, they are quite uniform in showing that there is a greater tendency to divorce among those who have lived with someone (i.e., the future spouse or any other partner) prior to marriage than among those who did not previously cohabit. Most scholars point out that either or both of two factors are probably at work here; first, the less-than-full acceptance of cohabitation as a lifestyle (implying less or no social support for those who cohabit), and second, the kind of persons who choose a “deviant” lifestyle—persons who are risk-takers, and who are less commitment-oriented. (However, see Popenoe 1987 for a different view of cohabitation in a setting where it is more normative.)

As rates of sexual activity outside of marriage rose, and as sex was to some extent disengaged from procreation (since the arrival of the birth control pill in the 1960s), research and theoretical interest focused on changes in sexual behavior and values in courtship. (See Schur 1988 for a highly negative view of the “Americanization of sex.”) It should be noted that cohabitation appears to be a “sexier” arrangement than marriage (Call, Sprecher, and Schwartz 1955), which may account for why prior cohabitants’ marriages do not meet their expectations and thus, may be more divorce-prone.

Also on the negative side of the ledger, there is concern about the spread of sexually transmitted diseases, including AIDS, and on factors related to

the use or nonuse of “safe” sexual practices. Research continues to examine variations in premarital sexual activity rates and their effects. Frazier (1994) points out that the AIDS epidemic has not sidetracked the sexual revolution that began in the 1960s. This is because the forces that fueled the revolution are still in place, and some are intensifying—“mobility, democratization, urbanization, women in the workplace, birth control and other reproductive interventions, and media proliferation of sexual images, ideas, and variation” (p. 32). Moreover, cohabitation is increasing as are the single-person household and single parenthood. The pursuit of individuality and freedom continues. Many studies show that women are more sexual today than at any previous time in this century, says Frazier. On the positive side, a greater openness about sexuality-related information has occurred. The trend, as Frazier sees it, is away from the illusions of traditional ideas about romance and toward a more reality-based understanding between men and women. Also positive, and part of the same revolution, are expanded definitions of masculinity and femininity as the trend toward egalitarianism continues.

Unmarried households (i.e., single parenthood) have lost much of their past stigma, and increased numbers of women are choosing to remain single over the (potentially illusory) financial security of marriage, notes Frazier. This is largely a function of women’s increased earning capacities in an expanded set of labor market opportunities.

Along with the strong trend toward later marriages has come declining family size. The U. S. Department of Commerce (1992) tells us that the median age at marriage has been rising and in the 1990s was higher than it had been a century earlier. One outcome of this is, as noted earlier, a rise in nonmarital births. Related to this is a rise in the now considerable rate of child poverty, since women’s (i.e., single mothers) earnings are not as high as single fathers or of men in general.

Frazier, Arikian, Benson, Losoff, and Maurer (1996) report that, among unmarried singles over age thirty, reasons for remaining single have to do with barriers as well as choices and that men would like to marry more than would women. (This situation is reversed among younger adults, where women are more interested in marriage than are men.) Never-married adults want to marry more

than do divorced adults, and divorced women have the least desire for marriage. Again, this may have to do with the greater options (if not economic parity with men) open to women in recent years. Both men and women state their primary reasons for wanting to marry as love, the desire for a family, what they see as the “romantic” nature of marriage, a desire for economic security (which, as noted, may be illusory), and the opportunity for regular sexual activity. However, the desire to remain single—for both men and women—is linked to having unrestricted career opportunities, to the desire for an “exciting” lifestyle, and to having the freedom to change and experiment. Men also identify the restrictive nature of marriage and the limits on mobility and experiences as reasons to remain single, but women mention the desire to be self-sufficient and the possibility of poor communication in marriage as among their top reasons for choosing singlehood over marriage.

Around the world, childbearing by unwed women has increased, accounting for about one-third of all births in America and northern Europe in the mid-1990s. Despite the fact that there has been a recent decline in teen births in the United States, teen pregnancies are much higher in America than in other industrialized nations, for which our poor (or absent) sexuality education is often blamed (Ventura, Matthews, and Curtin 1998).

The rising age at marriage and its effects have interested scholars not only in developed countries (e.g., East Germany) but also in developing countries such as Sri Lanka, Java, and sub-Saharan Africa. One effect is the relatively large numbers of young adults still living in the parental home. Living arrangements and other family influences such as parental divorce or having had alcoholic parents have been studied for their effects on dating behavior, premarital pregnancy, violence in courtship, and on drinking behaviors of young adults.

Research shows that expectations of marriage among those of courting age are inflated, when compared to the expectations of persons with marital experience. Inflated expectations may be another of the causes of subsequent divorce. Moreover, scholars who study conduct on dates have uncovered the seeming paradox of egalitarian daters who behave traditionally during the earliest stages of courtship. Behavior that does not reflect

beliefs gives false impressions, and has obvious implications for the spontaneity and honesty (or lack of them) of the conditions under which courting partners get to know one another. In this era when dating/courtship has lost much of its coherence, we find advice books for young adults with names such as *Dating for Dummies* (1996) and *The Complete Idiot's Guide to Dating* (1996). The titles alone tell a story. Still other writers attempt to capitalize on the old notion of a war between the sexes, implying that in addition to knowing little about how to date/court, we know very little about one another since “men are from Mars and women are from Venus”—unless we study such guides as *Mars and Venus on a Date* (Gray 1988). Under the guise of assisting daters to communicate with one another, they present half-traditional, half-egalitarian versions of how to get along with persons who are seen as each others’ “opposites.” These popular books rest on the idea that by following their prescriptions, as in *The Rules* (Fein and Schneider 1996), our courtships will be successful and their outcomes happy. (*The Rules*, interestingly, is highly traditionalistic, and reads like a guide for 1950s “dating success.”)

In contrast to these for-profit offerings, scholars continue to study “close” or “intimate” premarital relationships as these have changed from stylized conventional dating to the more informal “hanging out” and “hooking up,” the latter an almost only just-for-sexual-purposes arrangement. These shifts follow in part from a weakened normative imperative to marry (Thornton 1989) and in part from the trend toward more egalitarian relationships between the sexes. However, in almost all research, male/female similarities and differences continue to form part of the data analysis. Recent studies have examined attraction in an effort to identify its bases, and, less broadly, have investigated “opening lines” used for meeting potential partners. Which lines work, which do not, and why are unsuccessful lines still in use? These are among the kinds of questions that are asked and answered by such investigations. Scholars working in criminal justice-related areas have provided information about the use of Rohypnol, a central nervous system depressant that is “abused throughout the United States by high school and college students, rave and nightclub attendees, and drug addicts and alcohol abusers.” Its use facilitates sexual assaults (Office of National Drug

Control Policy 1998). Other investigations examine dating among herpes- and HIV-infected persons. On the more positive and more conventional side, some of the standard variables such as age and education have been reexamined for their impact on mate choice patterns (Qian 1998).

Earlier, the trend toward expanded gender roles was noted. Considerable research interest has been devoted to identifying the components of conventional (traditional) masculinity and femininity and their effects, and on resistance to change in these stereotypes—for example, because of ongoing conventional socialization practices and, as communications experts have documented, because of the effects of various media portrayals supporting the *status quo ante*. Scholars have also noted the greater likelihood of relational success among androgynous than among conventionally masculine men and feminine women.

Research on courtship has extended to the study of “taboo” conversational topics, degrees and forms of honesty and deception, communication style differences between the sexes (one result of differential socialization), and methods of conflict resolution that enhance relationship survival or that presage relationship dissolution. Interest in failed relationships has attempted to identify factors at both individual and dyadic levels that might have predicted which pairings would last and which would not. In ongoing relationships, scholars have investigated the positive and negative effects of outside influences such as parental or peer pressure, and the parts played by same- and cross-sex friends. Other topics of interest have included barriers to the development of trust and the effect of its loss, the meanings of commitment, and the effects of self-disclosure, self-esteem, self-awareness, and jealousy on close relationships.

We have witnessed a virtual explosion in the study of love—attempts to identify its forms, its properties, and its distribution of types across women and men as well as the effects of all of these, especially in terms of romantic love. On the negative side, a number of studies of violence in courtship have shown relatively high rates of this kind of activity, especially between cohabitants—and also reveal that a sizable minority of those who have experienced violence in close relationships identify it with a loving motive. This seemingly odd

justification is explained by the practices of parents who, when using violence against their children, often indicate that they hit or spank “out of love.” Thus, the lesson (i.e., the rationalization) is learned early. Other negative aspects of courtship include the study of “mind games” and other facets of competition between partners, sexual aggression including date/acquaintance rape, and the effects of contrasts between idealized images and courtship realities.

As courtship itself has expanded, researchers have taken a interest in an expanded range of relationship types. For example, the romantic involvements between lesbian women and between gay men have been studied in their own right, and also for comparative purposes with heterosexual partnerships. A predictable area for future study is the legalization of marital relationships between same-sex partners.

Other innovations such as video- and computer-matching and personal advertisements in print media have also captured researchers’ attention, as has using the internet to make romantic contacts (sometimes called “cyberdating”). Scholars can be expected to pursue the study of the impact of prenuptial agreements on relationships. To a lesser extent, older lines of research have continued to probe the purported decline of, or changes in, the double standard, the “principle of least interest” (Waller 1938) as related to changes in gender roles, the dimensions of intimacy, and on desired traits in dating partners as these may differ between premarital partners and permanent mates.

As society grows more complex and the rate of change is increasingly rapid, confusion over the mate-selection process in all of its dimensions appears to be rife. A study of college student dating shows that these young adults have questions about virtually every aspect of the process and about the choices they make (Laner 1995). High divorce rates have produced a backlash of insecurity as the marriage decision approaches. This is reflected in the frequently asked question, “How can I be *sure* of making the right choice in a partner?” Sociologists and scholars in related disciplines continue to study a growing set of factors that shed light on the answer.

(SEE ALSO: *Alternative Lifestyles; Love; Mate-Selection Theories*)

## REFERENCES

- Adams, Bert N. 1979 "Mate Selection in the United States: A Theoretical Summarization." In W. R. Burr, R. Hill, F. I. Nye, and I. L. Reiss, eds., *Contemporary Theories About the Family*. New York: Free Press.
- Aronson, Elliot 1972 *The Social Animal*. New York: Viking.
- Bailey, Beth L. 1988 *From Front Porch to Back Seat: Courtship in Twentieth Century America*. Baltimore: Johns Hopkins University Press.
- Browne, Judy 1996 *Dating for Dummies*. Foster City, Calif.: IDG Books.
- Bulcroft, Kris, and Margaret O'Connor 1986 "The Importance of Dating Relationships on Quality of Life for Older Persons." *Family Relations* 35:397-401.
- Call, Vaughn, Susan Sprecher, and Pepper Schwartz 1995 "The Incidence and Frequency of Marital Sex in a National Sample." *Journal of Marriage and the Family* 57:639-652.
- Farber, Bernard 1964 *Family Organization and Interaction*. San Francisco: Chandler.
- Fein, Ellen, and Sherrie Schneider 1995 *The Rules: Time-Tested Secrets for Capturing the Heart of Mr. Right*. New York: Warner.
- Frazier, Patricia, Nancy Arikian, Sonja Benson, Ann Losoff, and Steven Maurer 1996 "Desire for Marriage and Life Satisfaction among Unmarried Heterosexual Adults." *Journal of Social and Personal Relationships* 13(2):225-239.
- Frazier, Shervert H. 1994 "Psychotrends: What Kind of People Are We Becoming?" *Psychology Today* January/February:32-37, 64, 66.
- Goode, William J. 1969 "The Theoretical Importance of Love." *American Sociological Review* 34:38-47.
- Gordon, Michael 1981 "Was Waller Ever Right? The Rating and Dating Complex Reconsidered." *Journal of Marriage and the Family* 43:67-76.
- Gray, John 1998 *Mars and Venus on a Date*. New York: Harper-Collins.
- Kirkpatrick, Clifford (1955) 1963 *The Family as Process and Institution*, 2nd ed. New York: Ronald.
- Kuriansky, Judy 1996 *The Complete Idiot's Guide to Dating*. New York: Alpha Books.
- Laner, Mary R. 1995 *Dating: Delights, Discontents, and Dilemmas*. Salem, Wisc.: Sheffield.
- Murstein, Bernard I. 1980 "Mate Selection in the 1970s." *Journal of Marriage and the Family* 42:777-792.
- Office of National Drug Control Policy 1998 "Fact Sheet: Rohypnol." Washington, D.C.: Drug Policy Information Clearing House.
- Popenoe, David 1987 "Beyond the Nuclear Family: A Statistical Portrait of the Changing Family in Sweden." *Journal of Marriage and the Family* 49:173-183.
- Qian, Zhenchao 1998 "Changes in Assortative Mating: The Impact of Age and Education, 1970-1990." *Demography* 35(3):279-292.
- Queen, Stuart A., Robert W. Habenstein, and Jill S. Quadagno 1985 *The Family in Various Cultures*, 5th ed. New York: Harper and Row.
- Ramu, G. N. 1989 "Patterns in Mate Selection." In K. Ishwaran, ed., *Family and Marriage: Cross-Cultural Perspectives*. Toronto: Wall and Thompson.
- Rothman, Ellen K. 1984 *Hands and Hearts: A History of Courtship in America*. New York: Basic Books.
- Schur, Edwin M. 1988 *The Americanization of Sex*. Philadelphia: Temple University Press.
- Stein, Peter J. (ed.) 1981 *Single Life: Unmarried Adults in Social Context*. New York: St. Martin's.
- Stephen, Timothy D. 1985 "Fixed-Sequence and Circular-Causal Models of Relationship Development: Divergent Views on the Role of Communication in Intimacy." *Journal of Marriage and the Family* 47:955-963.
- Thornton, Arland 1989 "Changing Attitudes toward Family Issues in the United States." *Journal of Marriage and Family* 51:873-893.
- U.S. Bureau of the Census 1992 Statistical Brief 92-13, *Family Life Today. . . and How It Has Changed*.
- Ventura, S. J., T. J. Matthews, and S. C. Curtin 1998 "Declines in Teenage Birth Rates, 1991-97; National and State Patterns." *National Vital Statistics Reports* 47(12). Hyattsville, Md: National Center for Health Statistics.
- Waller, Willard 1938 "The Rating and Dating Complex." *American Sociological Review* 2:727-734.
- , and Reuben Hill 1951 *The Family: A Dynamic Interpretation*, rev. ed. New York: Dryden.

MARY REIGE LANER

**CRIME DETERRENCE**

See Criminal Sanctions; Criminology; Social Control.

**CRIME RATES**

The interpretation of crime rates seems unproblematic, but those sociologists who study crime know that is not the case. Even simple counts of

crime raise difficult issues. To imagine the difficulties consider two men at a bar. One of them makes a rude comment to the other and the other person shoves him. The person shoved strikes the other person on the jaw before the fight is broken up. The person hit on the jaw calls the police to the scene and wants to press charges, but the person shoved claims self-defense and that the person who shoved him is guilty of an assault. The two men may decide that the matter is not worth pressing and the police may agree. On the other hand, the police may file a report that results in an arrest. If a grand jury thinks the evidence warrants it, the case may go to trial. Even if the case reaches trial, the jury may acquit the accused. Other possible steps in the process exist, but determining whether a crime occurred is often difficult. The process contains many decision points. Will the individuals involved report the incident to the police? Will the police determine that a crime occurred? Now imagine that an interviewer asks one of these men about their history of criminal offending or victimization. Would the individual report this incident to the interviewer? Would one of the individuals admit to an interviewer that he committed an assault? Crime counts constitute an important element in the calculation of crime rates, but they are only one factor.

Rather than report simple crime counts, social scientists often calculate crime rates. Depending on the context, they may do this because crime rates: (1) allow for the comparison of crime patterns across groupings of different sizes, (2) make possible the comparison of the relative frequency of crime over time, or (3) help in the assessment of the risk of victimization. In terms of comparability of patterns across groupings of different sizes, simply reporting the number of crimes in cities with 100,000 or more residents produces unsatisfactory results. Forty homicides may not represent a large number for New York City, but twenty homicides represent a large number in Eugene, Oregon. The calculation of homicide rates per 100,000 residents for New York City and Eugene makes comparisons between the two cities easier. Similar calculations of rates for different years facilitate comparisons over time. The number of homicides may have doubled in Los Angeles from 1920 to 1990 while the rate per 100,000 residents has decreased. In terms of the risk of motor vehicle theft, a rate based on the number of motor

vehicle thefts per 10,000 registered motor vehicles may be more helpful than the number of motor vehicles stolen.

Equation (1) represents the general formula for calculating a crime rate:

$$\text{Crime Rate Per Base} = \left( \frac{\text{Number of Incidents}}{\text{Relevant Population Size}} \right) \times \text{Base} \quad (1)$$

Each of the components in this formula: Base, Relevant Population Size, and Number of Incidents, represents an important decision point when calculating a crime rate. Determining the most appropriate measure for each component is not as straightforward as it might first appear.

*Selecting a Base:* If the base is 100,000, then the interpretation of the rate is the number of incidents per 100,000. With a base of 100 the interpretation is in terms of the number of incidents per 100 (a percent). Reports of crime rates often appear as rates per 100,000 or per 10,000. One reason for this is the relative infrequency of some crimes. For example, in 1997 the number of murders in the United States was 18,210 and the population of the United States was 268 million. If we choose 100 for the base, we find that the homicide rate per 100 is .0068 or .0068 percent. This figure challenges the intuition of most people. We should choose the base to make the rate easier to interpret. If we choose 100,000 for the base, we find that the homicide rate is 6.8 per 100,000; a much easier figure for most people to understand.

*Relevant Population Size:* The choice of the relevant population size depends upon the selection of an appropriate "relevant population." The Uniform Crime Reports (UCR) typically chooses the number of residents as the relevant population. This number is used to calculate crime rates for cities or for the United States as a whole in the UCR. The National Crime and Victimization Survey (NCVS) often uses the number of residents who are twelve years of age or older as the relevant population, since the survey only asks about crime incidents involving those in that age range. At other times households represent the relevant population as when victimization surveys calculate rates for households touched by crime.

---

**General Definitions of Part I Offenses in the Uniform Crime Reports**

<b>Criminal Homicide</b>	Murder and nonnegligent manslaughter: the willful (nonnegligent) killing of one human being by another. Deaths caused by negligence and justifiable homicide are excluded.
<b>Forcible Rape</b>	The carnal knowledge of a female forcibly and against her will. Included are rapes by force and attempts or assaults to rape. Statutory offenses (no force used and victim under age of consent) are excluded.
<b>Robbery</b>	The taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence or by putting the victim in fear.
<b>Aggravated Assault</b>	An unlawful attack by one person on another for the purpose of inflicting severe or aggravated bodily injury. This type of assault is usually accompanied by the use of a weapon or by means likely to produce death or great bodily harm. Simple assaults are excluded.
<b>Burglary</b>	Breaking or Entering: The unlawful entry of a structure to commit a felony or a theft. Attempted forcible entry is included.
<b>Larceny</b>	Theft (Except Motor Vehicle Theft): The unlawful taking, carrying, leading, or riding away of property from the possession or constructive possession of another. Attempted larcenies are included. Embezzlement, con games, forgery, worthless checks, etc., are excluded.
<b>Motor Vehicle Theft</b>	The theft or attempted theft of a motor vehicle. A motor vehicle is self-propelled and runs on the surface and not on rails. Specifically excluded from this category are motorboats, construction equipment, airplanes, and farming equipment.
<b>Arson</b>	Any willful or malicious burning or attempt to burn, with or without intent to defraud, a dwelling house, public building, motor vehicle or aircraft, personal property of another, etc.

Table 1

SOURCE: Federal Bureau of Investigation, 1998. *Crime in the United States—1997* (Appendix II).

Thus, the total number of people may not be the most appropriate “relevant population.” In the case of rape incidents, it might be more appropriate to compute rape rates using women as the relevant population. For motor vehicle theft, the relevant population might be the number of registered motor vehicles. For commercial burglaries, the relevant population might be the number of commercial establishments. Social scientists should choose these populations carefully.

*Number of Incidents:* Determining the value for the numerator of the crime rate formula raises particularly difficult issues. The UCR, for example, employs a complicated set of rules for counting the number of incidents. For example, the UCR “hierarchy rule” ensures that for Part I crimes (the most serious street crimes) only the *most* serious crime is recorded. Thus, if a person breaks into a store, is confronted by the owner and threatens the owner with a gun, and then kills the owner, the offense is recorded as a homicide, but not as a burglary or robbery or assault. The “hotel rule” is

invoked when a burglar breaks into several units in an apartment or hotel and only one burglary is counted. Researchers draw a distinction between an incident rate and a prevalence rate. This involves a distinction between the number of criminal incidents per some relevant population and the number of offenders per some relevant population. Using the number of offenders per some relevant population produces a prevalence rate.

Thus, many decisions help determine the most appropriate crime rate for a particular purpose. These decisions need to be considered when interpreting a particular rate or when attempting to compare one crime rate with another.

#### SOURCES OF CRIME DATA IN THE UNITED STATES

There are three major sources of crime data in the United States, and knowing the strengths and weaknesses of these data sources helps explain why crime rates derived from them often vary.



**Crimes and Crime Indexes For the United States, 1997**

<b>Offenses</b>	<b>Number</b>	<b>Rate per 100,000 Inhabitants</b>	<b>Percent Change in the Rate From 1988 to 1997</b>
Crime index total	13,175,100	4,922.7	-13.1
Violent crime index	1,634,770	610.8	-4.1
Homicide	18,210	6.8	-19.0
Forcible rape	96,120	35.9	-4.5
Robbery	497,950	186.1	-15.8
Aggravated assault	1,022,490	382.0	-3.2
Property crime index	11,540,300	4,311.9	-14.2
Burglary	2,461,100	919.6	-29.8
Larceny-theft	7,725,500	2,886.5	-7.9
Motor vehicle theft	1,353,700	505.8	-13.2

Table 2

SOURCE: Data are from *Crime in the United States-1997* (Federal Bureau of Investigation 1998, Table 1)

These three methods are similar to those used to collect crime data internationally: (1) official data collected by law enforcement agencies, (2) surveys of the victims of crimes, and (3) self-report studies based on offenders. Other sources also are available, but less commonly used (e.g., vital statistics, hospital records, and insurance records).

**The Uniform Crime Reports.** During the latter half of the 1920s the International Association of Chiefs of Police (IACP) developed uniform crime definitions and counting rules that would allow it to collect more comparable data from different law enforcement agencies. The IACP initiated the UCR in January of 1930. During the latter part of 1930 the Bureau of Investigation took over the UCR program, and it has remained a function of that bureau (renamed the Federal Bureau of Investigation, or FBI).

The general definitions of the UCR Part I offenses appear in Table 1, but more detailed definitions appear in the *Uniform Crime Reporting Handbook* (FBI 1985). As an example, the Uniform Crime Reporting System defines burglary as the unlawful entry of any fixed structure, vehicle, or vessel used for a regular residence, industry, or business, with or without force, with the intent to commit a felony, or larceny. Without uniform definitions, jurisdictions in which state law defines theft from a storage-shed as a burglary would report such an incident as a burglary to the FBI. In another jurisdiction, breaking and entering might be required for an incident to be classified as a

burglary. In such a jurisdiction an incident in which an offender entered a house through an open window in order to steal a stereo would be coded as unlawful entry with intent to commit a crime for the purposes of the state's law, but as a burglary for UCR purposes.

Participation in the UCR is voluntary. Local law-enforcement agencies send their data to the FBI or to state-level UCR programs that forward the data to the FBI. In the early years only the largest law enforcement agencies in the United States participated. But by the 1980s and 1990s, agencies representing about 97 percent of the population of the entire United States participated. The FBI reports two types of crime: Part I crimes and Part II crimes. For Part I crimes, data on crimes known to the police and on arrests are reported; for Part II crimes, only data on arrests are reported. Part I crimes include violent crimes (murder and non-negligent homicide, forcible rape, robbery, and aggravated assault), property crimes (burglary, larceny-theft, and motor vehicle theft), and arson. Each year the news media report crime rates derived from the FBI data and a summary of these data appear annually in *Crime in the United States*. This FBI publication provides crime rates for states, metropolitan statistical areas, counties, cities, other geographical subdivisions, and the nation as a whole.

Table 2 presents Part I crime data from the publication *Crime in the United States-1997* (FBI 1998). The second column presents the number of

incidents of a specific type, the third column reports the rate per 100,000 residents, and the final column indicates the change in the rate over a ten year period. In 1997, for example, 1,634,770 violent crimes were known to the police. The rate of violent crimes was 610.8 per 100,000 persons, which represents a decrease in the rate of violent crimes from 1988 to 1997 of 4.1 percent. Officially reported crime rates fell in all Part I categories from 1988 to 1997.

Crimes reported or discovered by the police and found through investigation to have occurred are labeled "crimes known to the police." A crime found not to have occurred is "unfounded," and does not appear as a crime known to the police. Although occasionally police discover a crime being committed (e.g., they happen upon a burglary in progress, or they arrest someone for resisting a police officer) they are highly dependent on citizen reports of criminal incidents.

The UCR also provides data on arrests for Part I and Part II crimes, and reports these crimes on the basis of the geographical divisions described above. Since these data are for arrests, the rates calculated from them should be labeled "arrest rates." Part II crimes include: other assaults (not including aggravated assaults), forgery and counterfeiting, fraud, embezzlement, stolen property (buying, receiving, and possessing), vandalism, weapons (carrying, possessing, etc.), prostitution and commercialized vice, sex offenses (except forcible rape and prostitution), drug abuse violations, gambling, offenses against family and children, driving under the influence, liquor law violations, drunkenness, disorderly conduct, vagrancy, curfew and loitering law violations, and runaways.

**National Crime and Victimization Survey (NCVS).** The NCVS data come from large surveys using probability samples of respondents. Since 1973 a national-level program has surveyed a probability sample of U.S. households. The U.S. Census Bureau conducts the survey, and during the 1990s the sample contained some 60,000 households and approximately 100,000 respondents. The results of this survey have been published annually since 1973 by The U.S. Department of Justice in *Criminal Victimization in the United States*.

The NCVS collects data on a much smaller number of crimes than the UCR. Since it is a

survey of the victims of crime, it does not record homicides, but it does record forcible rapes and sexual assaults, robberies, aggravated assaults, simple assaults, burglaries, larceny-thefts, and motor-vehicle thefts. It also provides details about the victims of crime not available in the UCR. These include the household income of the victim, any injuries sustained by the victim, insurance coverage, length of any hospitalization due to injuries, what protective measures the victim employed, and so on.

Careful field-testing preceded the initiation of data collection for NCVS in 1973 and before any changes made to its design since that time. An early decision involved the length of the reference period (the period the interviewers referred to when they asked whether victimizations occurred during the past period). Researchers chose a six-month reference period on the basis of the costs of doing more frequent surveys and studies that showed the effect of memory decay on remembering whether an event occurred. Respondents tend to telescope events into the reference period, and this can cause substantial increases in the number of victimizations reported during the past six months. To overcome this problem, each respondent remains in the sample for three years with interviews repeated every six months. Victimizations reported to the interviewer the first time a respondent is interviewed do not contribute to estimates of the crime rates. That first interview and subsequent interviews provide bounds for the six-month reference period. The interviewer asks about victimizations occurring since the last interview and has a list of incidents reported in the last interview to make sure that those incidents are not telescoped into the reference period.

To estimate rates for the entire nation from the NCVS data, researchers use the samples to estimate the number of victimizations occurring in the entire nation. They then use these estimates to calculate victimization rates. The derived estimates often contain large sampling errors, since they are based on only a "small" sample of the total population of the United States. These large sampling errors are greatest for the crimes least frequently reported by the respondents, for example, victimization rates for rapes involve more sampling error than rates for aggravated assault (O'Brien 1986).

**Self-Reports.** Unlike the UCR and NCVS, self-report studies refer to a collection of studies conducted by different researchers using a variety of methodologies. The basic methodology, however, involves sampling respondents and questioning them about criminal or delinquent acts. Although some isolated self-report studies occurred in the 1940s, the technique gained popularity with the work of Short and Nye (1957, 1958).

Perhaps the most ambitious self-report study is the National Youth Survey, or NYS (Elliott et al. 1983). This survey involves a national panel sample. When first instituted in 1976, the panel members were aged eleven to seventeen. Estimates of national-level crime rates for the age groups covered in this survey can be computed because the sample of respondents is a national one. These rates, unlike UCR- and NCVS-based rates, result from respondents answering questions about their offenses.

The NYS, and other self-report surveys, calculate two types of crime rates: "prevalence rates" and "incidence rates." To calculate these crime rates, the NYS uses the general formula represented in Equation (1). The prevalence rate uses the number of respondents who claimed they committed an offense for the numerator, the number of respondents for the relevant population size, and (we choose) 100 for the base. This yields the percentage of respondents who have committed a particular act. In the NYS in 1980, when the respondents were fifteen to twenty-one years old, the prevalence rate for robbery was 2. The incidence rate uses the number of times individuals claimed they committed a particular act as the numerator, the number of respondents as the relevant population size, and typically uses 100, 1000, or 10,000 for the base. We again use 100 for the base and find that in the 1980 NYS the incident rate for robbery was 10. Thus, there was an average of ten robberies per 100 respondents in this age group in 1980, with only two out of 100 respondents claiming involvement in a robbery.

As with the NCVS, self-report studies depend upon the accuracy of respondent reports. Therefore, issues such as telescoping, recall, and embarrassment are consequential for self-reports of criminal activities. In comparison to the UCR and NCVS, many of the crimes reported in self-report surveys are trivial crimes or not crimes at all (e.g., lying to

parents or cheating on a test). The sample sizes, even in the largest of such surveys, are not nearly as large as sample sizes in the NCVS. Therefore, given the relative infrequency of serious crimes, the rates for serious crimes are not very precisely measured by self-report surveys; that is, the estimates have large standard errors. Nonetheless, crime rates derived from self-report studies provide much of the data used to investigate the determinants of individual criminal offending.

### **SOME PROBLEMS WITH UCR, NCVS, AND SELF-REPORT DATA**

Using information from official data (UCR), victimization surveys (NCVS), and self-report surveys helps us to gauge the accuracy of crime rates. Crime rates vary depending upon whether they derive from UCR, NCVS, or self-report data. To understand why crime-rate estimates vary depending upon their source, we briefly outline some problematic aspects of these three data sources.

**UCR Data.** The most obvious problem with UCR data involves underreporting of crimes to the police. The NCVS interviewers ask respondents whether they reported a particular incident to the police. In 1994, 78 percent of the respondents who reported a motor vehicle theft to the interviewer said they reported that crime to the police, 50 percent of the burglaries were reported, 50 percent of the robberies, and only 32 percent of rapes and sexual assaults (U.S. Department of Justice 1997).

Once an incident comes to the attention of the police a number of factors influence whether the incident is recorded as a crime or whether an arrest is made. There may be organizational pressures to raise or lower the crime rate (e.g., Bell 1960; Chambliss 1984; DeFleur 1975; McCleary et al. 1982; Selke and Pepinsky 1982; Sheley and Hanlon 1978; Wheeler 1967). The professional styles of particular police departments may affect the recorded crime rates (Beattie 1960; Skogan 1976; Wilson 1967, 1978). The interactions between officers and offenders also help determine the classification of an incident as a crime and whether an arrest is made (Black 1970; Smith and Visher 1981; Visher 1983). Given these and other problems, it is not surprising that crime rates recorded by the UCR are known to contain substantial errors.

**NCVS Data.** The NCVS may well provide a more accurate picture of crime rate trends and comparative crime rates than the UCR. As Sellin's (1931, p. 346) dictum suggests, "the value of a crime for index purposes decreases as the distance from the crime itself in terms of procedure increases." Further, the NCVS standardizes the procedures it uses in compiling crime incidents. When the NCVS changes its procedures, careful methodological studies evaluate the changes.

Although the incidents reported to NCVS interviewers need not depend upon a respondent contacting the police, the police responding to the contact, and eventually police recording of the incident as founded, respondents do need to respond appropriately in an interview situation in order for an incident to be recorded as a victimization in the NCVS. In this context, note that the interview situation is a social interaction. The interviewers have little to offer respondents for their time. Respondents may be reluctant to share embarrassing information with the interviewer, such as a fight at a bar, an assault by a relative, or a sexual assault. As noted earlier, sometimes respondents may forget about incidents or telescope the incidents into or out of the reference period.

Turner (1972) used "reverse record" checks to investigate 206 cases of robbery, assault, and rape found in police records. Interviewers interviewed these "known victims" using an NCVS-like technique. Only 63.1 percent of these "known incidents" were reported to interviewers. The percentage reporting these incidents to the interviewer was strongly related to the relationship of the offender to the victim. Respondents reported 76.3 percent of the incidents involving a stranger; 56.9 percent of the incidents involving known offenders; and only 22.2 percent of the incidents involving a relative.

Not all of this underreporting of incidents to the interviewers is attributable to embarrassment. Turner (1972) found a strong relationship between the number of months between the interview and the incident and the respondents' reporting of the incident to the interviewer. Respondents recalled 69 percent of the incidents occurring one to three months before the interview, 50 percent of those occurring four to six months before, 46 percent of those occurring six to nine months

before, and only 30 percent of those occurring ten to twelve months before.

**Self-Report Data.** Like UCR and NCVS data, self-report data contain their own weaknesses. As with the NCVS data, the survey research situation raises questions of honesty, forgetting, bounding the reference period, and so on. Some problems are more severe than those encountered by the NCVS. One of these problems is sample size; even the best-financed surveys—like the National Youth Survey (Elliot et al. 1983)—have samples of less than 2,000. Such small samples almost guarantee large sampling errors associated with the resulting crime rates (especially for serious crimes). Small sample size further limits the usefulness of these data for conducting regional analyses (e.g., estimating the crime rates for states or cities). Similarly, it limits the accuracy of these estimates for comparing the crime rates for groups such as Hispanics or Asian females. Unlike the NCVS, many self-report surveys do not involve panels in which respondents are reinterviewed over an extended period of time (the NYS is an exception) so that the interviews cannot be bounded. Often the response rates are quite low. In the NYS (Elliot et al. 1983), a sample of 2,360 eligible youth originally were selected and of these 73 percent agreed to participate in the first wave of data collection in 1976. By 1980 the sample size dropped to 1,494 or 63 percent of the original sample. This contrasts with initial response rates well in excess of 90 percent for the NCVS.

Two other issues need mention. First, self-report studies concentrate on relatively trivial behaviors, such as lying to parents, defying authority, or cheating on tests. One reason for concentrating on such behaviors is that they are more commonly reported. Given the small sample sizes typical of self-report studies, these behaviors may be more reliably measured. Second, some studies do not accurately gauge the frequency of "delinquent behaviors." They use response sets such as "no," "once or twice," or "several times." Experienced researchers now ask about a wide range of behaviors and more specifically ask about the frequency of these behaviors. In the NYS, Elliot et al. (1983) ask about "arson," "prostitution," and "physical threat for sex" as well as "skipped class" and "didn't return change." They use response categories such as "2-3 times a day," "once a day," "2-3 time a week," "once a week," "once every 2-3

weeks,” and “once a month,” as well as asking the exact number of times respondents offended during the past year.

### COMPARISONS OF CRIME RATES GENERATED FROM UCR, NCVS, AND SELF- REPORTS

Fortunately, the problems outlined above do not mean that crime rates based on UCR, NCVS, and self-reports are not useful. Homicide rates, for example, are probably reasonably accurate in terms of absolute rates and in terms of homicide trends over the past sixty-five years. They also are appropriate for comparisons of homicide rates across large cities. When we judge the usefulness and accuracy of crime rates, the type of crime considered is important (e.g., homicide or rape). Further, even if a crime rate is inaccurate in terms of the absolute amount of crime it represents, it may be an accurate measure of the relative amount of that crime in different areas or across time. This would be the case if exactly half or some other proportion of the aggravated assaults were reported in different jurisdictions (or across time). Then we could perfectly compare the relative amount of crime across jurisdictions (or crime trends over time). If only approximately the same proportion of cases were reported across jurisdictions, then comparisons of the relative amount of crime would only be approximate.

**Measuring Absolute Crime Rates.** The most accurately measured UCR Part I crime almost certainly is homicide. The reasons for this include the seriousness of the crime and the physical evidence it produces (e.g., a body, weapon, or missing person), which make it unlikely that once detected this crime will be ignored. Cantor and Cohen (1980) compared homicide rates based on the U.S. Department of Health, Education, and Welfare's *Annual Vital Health Statistics Report* with those from the UCR for the years 1935–1975 and found a very close correspondence. Using a different approach, O'Brien (1996) found evidence that fluctuations in the homicide rate closely paralleled fluctuations in the rate of other violent crimes from 1973 to 1992.

Motor vehicle theft is the other Part I crime that appears to be well estimated in terms of

absolute rates. In 1994, for example, the NCVS estimated 1,764,000 motor vehicle thefts based on respondents' answers, while the UCR reported 1,539,100 motor vehicle thefts known to the police. Thus the NCVS estimate exceeded the UCR estimate by only 28 percent. When NCVS respondents report a motor vehicle theft they are asked if they reported the theft to the police. Based on their responses one would calculate that the police receive reports of only 1,379,448 motor vehicle thefts, which is fewer motor vehicle thefts than appear in the UCR records. This probably occurs because motor vehicle thefts from private citizens constitute only 80 to 85 percent of the total number of such thefts known to the police (Biderman and Lynch 1991). For the remaining Part I crimes covered by both the UCR and NCVS (personal robbery, aggravated assault, residential burglary, and rape), the NCVS and UCR data show greater differences in absolute rates. The smallest ratio of NCVS incidents to UCR incidents is 1.62 for rape and the largest 3.01 for residential burglary.

Difficulties arise when comparing the self-report rates to UCR and NCVS rates, because self-report studies typically only involve a sample of young people. The estimated crime rates for these young people, however, are so high that we may conclude that the rates generated by self-reports are far greater than those generated by either the NCVS or the UCR. The 1976 NYS produced an estimated incident rate for eleven to seventeen year-olds for aggravated assault of 170 per 1000 and for robbery a rate of 290 per 1000. This compares with rates for those twelve year-olds and over of 7.90 and 6.48 based on the NCVS and rates for all ages of 2.29 and 1.96 based on the UCR. Some part of this discrepancy results from differences in the age groups compared, but not all of it. In 1976 eleven to seventeen year-olds comprised 13.38 percent of the population. Thus, even if it were assumed that no assaults or robberies were committed by other age groups, the aggravated assault rate based on the NYS would be  $[\.1338 \times 170 =] 22.75$  and the rate for robbery would be  $[\.1338 \times 290 =] 38.80$ .

**Measuring Relative Crime Rates.** If absolute rates are well measured, then relative rates of crimes across cities or other units probably are well measured. This means that both homicide rates and motor vehicle theft rates should be good

measures of the relative rate of crime. In the mid-1970s the NCVS conducted victimization studies with reasonably large size samples; 10,000 or more households containing some 22,000 respondents, in each of twenty-six large U.S. cities. The same cities, of course, reported UCR crime rates for Part I crimes. This provided the opportunity to compare the consistency of these estimates in terms of the relative crime rates measured by the NCVS and the UCR by correlating the two crime-rate measures across the twenty-six cities. The closer these correlations are to 1.00, the greater the degree to which cities with relatively high rates (low rates) on one of the measures have relatively high rates (low rates) on the other. Across these cities, the correlation of UCR and NCVS motor vehicle theft rates was .90 or higher (Nelson 1978, 1979; O'Brien, Shichor, and Decker 1980). The correlations were close to zero for rape, negative for aggravated assault, and positive and moderately strong for burglary and robbery. This indicates that for burglary and robbery the relative rates of crime in these cities are similar whether the UCR or NCVS measures are used, but they are not similar for rape and aggravated assault.

Overall, the use of homicide rates and motor vehicle theft rates to measure the relative amount of these crimes across cities is supported by the above findings. The use of burglary and robbery rates receives weaker support, and the use of rape and aggravated assault rates to compare the relative amount of these crimes across cities does not receive support.

#### **Gender and Race Composition of Offenders.**

Hindelang (1978, 1979) addressed the issue of whether the UCR and victim reports (NCVS) produce similar percentages of offenders identified as African American or white or as male or female. O'Brien (1995, 2000) replicated these findings using more recent data. These comparisons indicate that for those crimes in which the victim and offender come into contact (which enables the victim to identify the race and sex of the offender), the percentages based on the UCR and NCVS are fairly similar. For example, O'Brien (2000) aggregates data for the years 1992 to 1994 and finds that the percentages of males involved in these crimes, according to the UCR and NCVS, never differ by more than 3 percent. For race of the offender, rape produces the largest difference with victims

stating that 33.2 percent of the offenders are African American and the UCR arrest data indicating that 42.7 percent of those arrested are African American. For robbery, aggravated assault, and simple assault the differences do not exceed 5 percent.

Self-report data at one time were considered to be inconsistent with UCR data, which indicated that males committed crimes at a much higher rate than females and that African Americans also committed crimes at a much higher rate than whites. These large differences between males and females and African Americans and whites, however, are not found when self-report studies ask about more serious crimes and ask more carefully about the frequency of offending (Elliot and Ageton 1980; Hindelang, Hirschi, and Weis 1979).

These comparisons only touch on the issue of the appropriate uses of UCR, NCVS, and self-report data. The appropriate uses depend upon the type of crime and the type of comparison. (For helpful information on which to judge the appropriate uses of these data see Biderman and Lynch 1991; Gove Hughes and Geerken 1985; O'Brien 1985, 2000).

### **INTERNATIONAL DATA**

Many nations collect data from law enforcement agencies, survey victims, and conduct self-report surveys. Difficulties arise, however, when comparing these data, since the definitions of crimes and the counting rules differ from nation to nation. In this section we note some of these problems and cautiously present crime-rate data for two nations. In addition, for a larger number of nations, we present homicide rates.

Perhaps the most comparable large national victimization surveys are the NCVS (conducted in the United States) and the British Crime Survey (conducted in England and Wales). But even in this case, there are important differences—the British Crime Survey (BCS) uses a twelve-month reference period and unbounded interviews, while the NCVS uses a six-month reference period and bounded interviews. The former difference should decrease the estimated victimization rates in the BCS due to “memory decay,” but the latter should increase these rates due to “telescoping.” The BCS does not ask about crimes involving victims

under the age of sixteen, while the NCVS does not ask about crimes involving victims under age twelve.

Further problems arise with the definitions of offenses. In the United States, for example, aggravated assaults include attempted murders and almost any assault in which a weapon is used or assaults for the purpose of inflicting severe or aggravated injury. In England attempted murders receive separate treatment while the assaults most similar to Part I aggravated assaults are "woundings." Woundings require some kind of cut or wound where the skin or a bone is broken, or medical attention is needed, whereas common assault occurs if the victim is punched, kicked or jostled, with negligible or no injury.

Since 1981 definitions of rape have changed substantially in England. Until 1981, to be classified as a rape an incident required a male offender and female victim and the penetration of the vagina by the penis. Husbands could not be convicted of raping their wives, and males under age fourteen could be convicted of a rape. By 1996 rapes could involve offenders as young as ten, spousal victims, male victims, and anal intercourse. In the United States, the UCR definition of rape includes male victims, male or female offenders, spousal victims, anal intercourse, and other sexual acts. With these nontrivial problems in mind, we compare crime rates (based both on surveys and police records) in England and the United States.

Table 3 presents victimization rates for robbery, assault, burglary, and motor vehicle theft in the United States and England (based on the NCVS and the BCS) for the years 1981 and 1995 (Langan and Farrington 1998). Perhaps the most surprising result is that the estimated victimization rates in England are higher than in the United States for these crimes in 1995. The surveys also indicate substantial increases in robbery, assault, burglary, and motor vehicle theft from 1981 to 1995 in England and decreases for all but motor vehicle theft in the United States. The cross-country comparisons of these absolute rates may be more problematic than over-time comparisons within the countries—given that the victimization survey methods and the definitions of these four crimes changed little in these countries over this period.

Table 4 presents crime rates based on police records (Langan and Farrington 1998). In this

**Victim Survey Crime Rates per 1,000  
in the United States and England:  
1981 and 1995**

	1981		1995	
	U.S.	England	U.S.	England
Robbery	7.4	4.2	5.3	7.6
Assault	12.0	13.1	8.8	20.0
Burglary	105.9	40.9	47.5	82.9
Motor Vehicle Theft	10.6	15.6	10.8	23.6

*Table 3*

SOURCE: Langan and Farrington 1998.

table, we report rates per 100,000 residents because of the relatively low rates for homicide and rape. The rates in Table 4 reinforce the patterns reported in Table 3, which showed a substantial increase in crime rates from 1981 to the mid-1990s in England for robbery, assault, burglary, and motor vehicle theft. Table 4 also shows higher rates for these crimes (except for robbery) in England than in the United States in 1996. Two new crimes appear in this table: Homicide, for which the rate in the United States is several times higher than in England, for both 1981 and 1996 (most criminologists agree that homicide rates in the United States are very high by international standards) and rape, which shows a dramatic increase from 4.19 to 21.77 per 100,00 in England (this results, at least in part, from changes in the legal definition of rape that occurred in England between 1981 and 1996).

Although we could attempt other comparisons, most such comparisons involve even greater difficulties than those between the United States and England. Some countries do not differentiate between simple and aggravated assaults; others do not record any but the most severe assaults when the assault occurs within the family. As noted, legal definitions of rape may differ widely (even over time within the same country), reports of burglaries may require breaking and entering into a business or residence, only entering, or even breaking into an outbuilding.

Homicide represents the crime with the most consistent definition across countries: the intentional taking of another person's life without legal justification. Even here some countries may not separate justifiable from unjustifiable homicides.



**Police Recorded Crime Rates per 100,000 in the United States and England: 1981 and 1996**

	1981		1996	
	U.S.	England	U.S.	England
Homicide	9.83	1.13	7.41	1.31
Rape	70.59	4.19	70.79	21.77
Robbery	258.75	40.86	202.44	142.35
Assault	289.73	197.49	388.19	439.60
Burglary	1649.47	1447.36	942.95	2239.15
Motor Vehicle Theft	474.72	670.09	525.93	948.83

Table 4

SOURCE: Langan and Farrington 1998.

The legal justifications for killing a spouse or even a person committing a crime may vary from country to country, as may the reporting and recording of homicides. Still, the data for this crime rate almost certainly are more accurate for international comparisons than those for rape, robbery, burglary, assault, theft, or arson.

Table 5 presents data for 1990 from the United Nations (United Nations Crime and Justice Network 1998) on intentional homicide rates per 100,000 for a selected set of countries. The rate for the United States was taken from the UCR (FBI 1991). Note the wide range of homicide rates, with Barbados and the United States with relatively high rates and Austria, Botswana, Denmark, Japan, Norway, and Spain with relatively low rates. There probably are real differences between these two sets of countries in terms of homicide rates. It is less clear that smaller differences across countries with rather different legal systems, levels of development, and cultural histories, represent differences in the unjustifiable intentional taking of another person's life. To reiterate the difficulty in making cross-national comparisons, we note that INTERPOL's international crime statistics come with the warning: "the information given is in no way intended for use as a basis for comparisons between different countries."

**CONCLUSIONS**

At the beginning of this article we noted that the computation of a crime rate might seem straightforward, but designating the most appropriate term for each component in equation (1) often is

**Intentional Homicide Rates per 100,000 in Selected Nations in 1990**

Austria	.52	Italy	3.11
Barbados	11.76	Japan	.50
Botswana	.46	Madagascar	1.31
Bulgaria	2.52	Norway	.99
Canada	2.22	Spain	.71
Denmark	.80	Sri Lanka	2.17
Egypt	1.07	Sweden	1.41
Fed. Rep. of Germany	2.13	Switzerland	1.64
India	4.24	Turkey	1.64
Israel	2.30	United States	9.4

Table 5

SOURCE: United Nations Crime and Justice Information Network 1998. <http://www.ifs.unvie.ac.at/uncjin/mosaic/at.inthom/txt>.

difficult. Designating the base represents the easiest decision. This choice does not greatly change the interpretation of the rate, although it may make the interpretation more or less intuitive for the reader. Choice of the appropriate relevant population size is more difficult; for example, basing rape rates on the number of females rather than the number of people or the motor vehicle theft rate on the number of registered motor vehicles rather than the number of people. Determining the number of incidents constitutes the most problematic component of calculating crime rates. Here, the differing definitions of crimes, varying counting rules, response rates, rates of reporting incidents to the police or interviewers, response categories, bounding of interviews, memory decay, police discretion in recording crimes, and so on, can greatly affect the estimated crime rate. These factors differentially affect estimates of the number of incidents based on self-reports, UCR, and NCVS data. Even so, within the United States for certain crimes and comparisons, data from these different sources lead to similar conclusions. Comparisons of crime rates across nations, with the possible exception of homicide, is extremely risky.

REFERENCES

Beattie, Ronald H. 1960 "Criminal Statistics in the United States—1960." *Journal of Criminal Law, Criminology, and Police Science* 51:49–65.

Bell, Daniel 1960 "The Myth of Crime Waves." In Daniel Bell, ed., *The End of Ideology: On the Exhaustion of Political Ideas in the Fifties*. New York: Free Press.



- Biderman, Albert D., and James P. Lynch 1991 *Understanding Crime Incidence Statistics: Why the UCR Diverges From the NCS*. New York: Springer-Verlag.
- Black, Donald J. 1970 "The Production of Crime Rates." *American Sociological Review* 35:733-748.
- Cantor, David, and Lawrence E. Cohen 1980 "Comparing Measures of Homicide Trends: Methodological and Substantive Differences in Vital Statistics and Uniform Crime Report Time Series (1935-1975)." *Social Science Research* 9:121-145.
- Chambliss, William J. 1984 "Crime Rates and Crime Myths." In William J. Chambliss, ed., *Criminal Law in Action*. New York: John Wiley.
- Defleur, Lois B. 1975 "Biasing Influences on Drug Arrest Records: Implications for Deviance Research." *American Sociological Review* 40:88-103.
- Elliott, Delbert S., and Suzanne S. Ageton 1980 "Reconciling Race and Class Differences in Self-Reported and Official Estimates of Delinquency." *American Sociological Review* 45:95-110.
- , David Huizinga, Brian A. Knowles, and Rachelle J. Canter 1983 *The Prevalence and Incidence of Delinquent Behavior: 1976-1980: National Estimates of Delinquent Behavior by Sex, Race, Social Class and Other Selected Variables*. Boulder, Colo.: Behavioral Research Institute.
- Federal Bureau of Investigation 1985 *Uniform Crime Reporting Handbook*. Washington, D.C.: U.S. Government Printing Office.
- 1991 *Crime in the United States-1990*. Washington, D.C.: U.S. Government Printing Office.
- 1998 *Crime in the United States-1997*. Washington, D.C.: U.S. Government Printing Office.
- Gove, Walter R., Michael Hughes, and Michael Geerken 1985 "Are Uniform Crime Reports a Valid Indicator of the Index Crimes? An Affirmative Answer with Minor Qualifications." *Criminology* 23:451-491.
- Hindelang, Michael J. 1978 "Race and Involvement in Common Law Personal Crimes." *American Sociological Review* 43:93-109.
- 1979 "Sex Differences in Criminal Activity." *Social Problems* 27:143-156.
- , Travis Hirschi, and Joseph G. Weis 1979 "Correlates of Delinquency: The Illusion of Discrepancy Between Self-Report and Official Measures." *American Sociological Review* 44:995-1014.
- Langan, Patrick A., and David Farrington 1998 *Crime and Justice in the United States and In England and Wales, 1981-96*. Washington, D.C.: U.S. Department of Justice.
- McCleary, Richard, Barbara C. Nienstedt, and James M. Erven 1982 "Uniform Crime Reports as Organizational Outcomes: Three Time Series Experiments." *Social Problems* 29:361-372.
- Nelson, James F. 1978 "Alternative Measures of Crime: A Comparison of the Uniform Crime Report and the National Crime Survey in 26 American Cities." Presented at the American Society of Criminology, Dallas, Tex.
- 1979 "Implications for the Ecological Study of Crime: A Research Note." In William H. Parsonage, ed., *Perspectives on Victimology*. Beverly Hills, Calif.: Sage.
- O'Brien, Robert M. 1985 *Crime and Victimization Data*. Beverly Hills, Calif.: Sage.
- 1986 "Rare Events, Sample Size, and Statistical Problems in the Analysis of NCS (National Crime Survey) City Surveys." *Journal of Criminal Justice* 14:441-448.
- 1996 "Police Productivity and Crime Rates: 1973-1992." *Criminology* 34:183-207.
- 2000 "Crime and Victimization Data." In Joseph F. Sheley, ed., *Criminology*. New York: Wadsworth.
- , David Shichor, and David L. Decker 1980 "An Empirical Comparison of the Validity of UCR and NCS Crime Rates." *Sociological Quarterly* 21:301-401.
- Selke, William L., and Harold E. Pepinsky 1982 "The Politics of Police Reporting in Indianapolis, 1948-1978." *Law and Human Behavior* 6:327-347.
- Sellin, Thorsten 1931 "The Bias of a Crime Index" *Journal of the Institute of Criminal Law and Criminology* 22:335-356.
- Sheley, Joseph F., and John J. Hanlon 1978 "Unintended Effects of Police Decisions to Actively Enforce Laws: Implications for Analysis of Crime Trends." *Contemporary Crises* 2:265-275.
- Short, James F., and F. Ivan Nye 1957 "Reported Behavior as a Criterion of Deviant Behavior." *Social Problems* 5:207-213.
- 1958 "Extent of Unrecorded Juvenile Delinquency: Tentative Conclusions." *Journal of Criminal Law and Criminology* 49:296-302.
- Skogan, Wesley G. 1976 "Crime and Crime Rates." In Wesley G. Skogan, ed., *Sample Surveys of the Victims of Crime*. Cambridge, Mass.: Ballinger.
- Smith, Douglas A., and Christy A. Visher 1981 "Street Level Justice: Situational Determinants of Police Arrest Decisions." *Social Problems* 29:167-177.
- Turner, Anthony G. 1972 *The San Jose Methods Test of Known Crime Victims*. Washington, D.C.: U.S. Government Printing Office. National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration.

- United Nations Crime and Justice Information Network 1998 <http://www.ifs.unvie.ac.at/uncjin/mosaic/at.inthom/txt>.
- U.S. Department of Justice 1997 *Criminal Victimization in the United States, 1994*. Washington, D.C.: U.S. Government Printing Office.
- Visher, Christy A. 1983 "Gender, Police, Arrest Decisions, and Notions of Chivalry." *Criminology* 21:5-28.
- Wheeler, Stanton 1967 "Criminal Statistics: A Reformulation of the Problem." *Journal of Criminal Law, Criminology, and Police Science* 58:317-324.
- Wilson, James Q. 1967 "The Police and Delinquents in Two Cities." In Stanton Wheeler and H. M. Hughes, eds., *Controlling Delinquents*. New York: John Wiley.
- 1978 *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, Mass.: Harvard University Press.

ROBERT M. O'BRIEN

## CRIME, THEORIES OF

Most accounts of the rise of criminological inquiry indicate that it had its beginnings in mid-nineteenth-century developments in Europe, including the work of Cesare Lombroso, an Italian prison physician, who argued that many criminals are atavists, that is, biological throwbacks to a human type, *homo delinquens*, that allegedly existed prior to the appearance of *homo sapiens*. Since the time of Lombroso and other early figures in criminology, the field has grown markedly, both in terms of the variety of scholars who have tried to uncover the causes of crime and also in terms of the diverse theories that have been produced by these persons (Gibbons 1994). Currently legal theorists, psychologists, economists, geographers, and representatives of other scholarly fields engage in criminological theorizing and research. There has also been renewed interest in sociobiological theorizing and investigation regarding criminality. Even so, the largest share of work has been and continues to be carried on by sociologists. Thus, criminology is frequently identified as a subfield of sociology (Gibbons 1979, 1994).

Although a few scholars have argued that crime should be defined as consisting of violations of basic human rights or for some other "social" conception, most criminologists opt for the legalistic view that crime and criminal behavior are

identified by the criminal laws of nations, states, and local jurisdictions. Acts that are not prohibited or required by the criminal law are not crimes, however much they may offend some members of the community. Also, the reach of the criminal law in modern societies is very broad, involving a wide range of behavioral acts that vary not only in form but in severity as well. The criminal laws of various states and nations prohibit morally repugnant acts such as murder or incest, but they also prohibit less serious offenses such as vandalism, petty theft, and myriad other acts. Parenthetically, there is considerable controversy in modern America, both among criminologists and among members of the general public, as to whether certain kinds of behavior, such as marijuana use, various consensual sex acts between adults, or abortion, ought to be expunged from or brought into the criminal codes.

Persons of all ages violate criminal laws, although a number of forms of criminality are most frequent among persons in their teens or early twenties. Except for "status offense" violations such as running away, truancy, and the like, which apply only to juveniles (usually defined as persons under eighteen years of age), juvenile delinquency and adult criminality are defined by the same body of criminal statutes. However, criminologists have often constructed theories about delinquency separate from explanations of adult criminality. Although many theories of delinquency closely resemble those dealing with adult crime, some of the former are not paralleled by theories of adult criminality. In the discussion to follow, most attention is upon explanatory arguments about adult lawbreaking, but some mention is also made of causal arguments about juvenile crime.

### CRIMINOLOGICAL QUESTIONS AND CAUSAL THEORIES

Given the broad compass of the criminal law, and given the variety of different perspectives from which the phenomenon of crime has been addressed, it is little wonder that there are many theories of crime. Most of these theories center on the explanation of crime patterns and crime rates, or what might be termed "crime in the aggregate," or are pitched at the individual level and endeavor to identify factors that account for the involvement of specific individuals in lawbreaking conduct (Cressey 1951; Gibbons 1992, pp. 35-39)

These are related but analytically separate questions about the causes of crime. As Donald Cressey (1951) argued many years ago, an adequate account of criminality should contain two distinct but consistent aspects: First, a statement that explains the statistical distribution of criminal behavior in time and space (epidemiology), and second, a statement that identifies the process or processes by which persons come to engage in criminal behavior.

Statistical distributions of criminal behavior in time and space are usually presented in the form of crime rates of one kind or another. One of the most familiar of these is the *index crime rate* reported annually for cities, states, and other jurisdictions by the Federal Bureau of Investigation. The index crime rate is comprised of the number of reported cases of murder, non-negligent manslaughter, forcible rape, aggravated assault, robbery, burglary, larceny, auto theft, and arson per jurisdiction, expressed as a rate per 100,000 persons in that jurisdiction's population.

Many crime rate patterns are well known, including relatively high rates of violence in the United States as compared to other nations, state-by-state variations in forcible rape rates, regional variations in homicide and other crimes within the United States, and so forth. However, criminological scholars continue to be hampered in their efforts to account for variations in crime across various nations in the world by the lack of detailed data about lawbreaking in nations and regions other than the United States (although see van Dijk, Mayhew, and Killias 1990).

Criminologists have developed a number of theories or explanations for many crime rate variations. One case in point is Larry Baron and Murray Straus's (1987) investigation of rape rates for the fifty American states, in which they hypothesized that state-to-state variations in gender inequality, social disorganization (high divorce rates, low church attendance, and the like), pornography readership, and "cultural spillover" (authorized paddling of school children, etc.) are major influences on forcible rape. Steven Messner and Richard Rosenfeld's (1994) *institutional anomie theory* is another example of theorizing that focuses on crime rate variations. They argued that in present-day America, cultural pressures to accumulate money and other forms of wealth are joined to

weak social controls arising from noneconomic elements of the social structure, principally the political system, along with religion, education, and family patterns. According to Messner and Rosenfeld, this pronounced emphasis on the accumulation of wealth and weak social restraints promotes high rates of instrumental criminal activity such as robbery, burglary, larceny, and auto theft.

Crime rates are important social indicators that reflect the quality of life in different regions, states, or areas. Additionally, theories that link various social factors to those rates provide considerable insight into the causes of lawbreaking. But, it is well to keep in mind that crime rates are the summary expression of illegal acts of individuals. Much of the time, the precise number of offenders who have carried out the reported offenses is unknown because individual law violators engage in varying numbers of crimes per year. Even so, crime rates summarize the illegal actions of individuals. Accordingly, theories of crime must ultimately deal with the processes by which these specific persons come to exhibit criminal behavior.

In practice, criminological theories that focus on crime rates and patterns often have had relatively little to say about the causes of individual behavior. For example, variations in income inequality from one place to another have been identified by criminologists as being related to rates of predatory property crime such as burglary, automobile theft, and larceny. Many of the studies that have reported this finding have had little to say about how income inequality, defined as the unequal distribution of income among an entire population of an area or locale, affects individuals. In short, explanations of crime rate variations often have failed to indicate how the explanatory variables they identify "get inside the heads of offenders," so to speak.

Although criminological theories about crime rates and crime patterns have often been developed independently of theories related to the processes by which specific persons come to exhibit criminal conduct, valid theories of these processes ought to have implications for the task of understanding the realities of individual criminal conduct. For example, if variations in gender inequality and levels of pornography are related to rates of forcible rape, it may be that males who carry out sexual assaults are also the individuals who most strongly

approve of discrimination against women and who are avid consumers of pornography. In the same way, if income inequality bears a consistent relationship to rates of predatory crime, it may be that individual predators express strong feelings of “relative deprivation,” that is, perceptions that they are economically disadvantaged and distressed about their situation. However, some additional factors may also have to be identified that determine which of the persons who oppose women’s rights or who feel relatively deprived become involved in illegal conduct and which do not.

### PERSPECTIVES, THEORIES, AND HYPOTHESES

A number of arguments about crime patterns and the processes through which individuals get involved in lawbreaking are examined below. Before moving to these specific theories, however, two other general observations are in order. First, in criminology, as in sociology more generally, there is considerable disagreement regarding the nature of perspectives, theories, and hypotheses (as well as paradigms, frameworks, and other theoretical constructions). Even so, *perspectives* are often identified as broad and relatively unsystematic arguments; while *theories* are often described as sets of concepts, along with interconnected propositions that link the concepts together into an “explanatory package”; and *hypotheses* are specific research propositions derived from theories. In practice, however, many causal explanations that have been described as theories have been incomplete and also conceptually imprecise. Jack Gibbs (1985) has labeled such “theories” as being in “the discursive mode” rather than as formal theories. Discursive arguments are stated in everyday language and their underlying logic is often difficult to identify. According to Gibbs, because many criminological theories are discursive, precise predictions cannot be deduced from them, nor is it possible to subject predictions to empirical test, that is, to validation through research.

Many criminological theories involve relatively vague concepts, faulty underlying logic, and other problems. At the same time, it is possible to identify a number of general theoretical perspectives in criminology and to differentiate these from relatively formalized and precise theories. For example, many criminologists contend that

American society is *criminogenic* because it involves social and economic features that appear to contribute heavily to criminality. However, this is a general perspective rather than a theory of crime in that it does not identify the full range of factors that contribute to lawbreaking, and it also lacks a set of explicit and interrelated propositions. By contrast, the income inequality argument more clearly qualifies as a causal theory, as does the formulation that links gender inequality, pornography readership, and certain other influences to forcible rape.

A few other comments are in order on theoretical perspectives in criminology. During most of the developmental history of criminology in the United States, from the early 1900s to the present, sociological criminologists voiced support for the criminogenic culture thesis that directs attention to social-structural factors thought to be responsible for criminality. Thus, this view might also be referred to as “mainstream criminology.” Most criminologists have linked lawbreaking to major “rents and tears” in societal structure at the same time that most of them have assumed that these crime-producing features can be remedied or lessened through social and economic reforms of one kind or another (Gibbons 1992, 1994; Currie 1985).

In the 1970s, a markedly different perspective competed for attention. Often referred to as “radical-Marxist” or “critical” criminology, it asserted that the causes of crime arise out of societal characteristics that are inherent in corporate capitalism (Gibbons 1992, pp.122–130; Chambliss 1975; Quinney 1974, 1977). According to radical-Marxist criminologists, criminal laws serve the interests of the capitalist ruling class. In turn, the system of corporate capitalism over which the ruling class presides depends for its survival on the exploitation of the resources and people of other countries and the economic oppression of citizens within capitalist nations. These conditions create economic strains for many persons, contribute to the deterioration of family life, and drive many individuals into desperate acts of lawbreaking.

The radical-Marxist perspective received considerable attention in the 1970s. Those who criticized it claimed that it presented a one-dimensional, oversimplified account of the social sources of criminality. For example, while some criminal laws favor the interests of the owners of capital, many

others serve broader social interests. Similarly, while some forms of crime may be related to economic problems, others are not.

A number of other alternative perspectives began to appear in criminology in the 1980s and 1990s, so that theorizing about crime and criminality has become even more diversified. These “new criminologies” (Gibbons 1994, pp. 151–175) include postmodernist viewpoints, feminist arguments, and a number of other strains of thought, all of which differ in a number of ways from “mainstream” criminology.

Although broad theorizing has continued to proliferate in criminology, another major trend in recent years has taken criminology in a different direction, toward relatively detailed theories specific to one or another form of crime and toward research investigations of those theories. Baron and Straus’s (1987) formulation that links gender inequality, pornography, and specific flaws in the social control system is a case in point, as is Kenneth Polk’s (1994) theorizing and research regarding the various “scenarios” of social interaction that culminate in lethal violence. Indeed, contemporary criminology has a rich accumulation of empirical evidence that can be drawn upon by those who seek to understand the nature and causes of criminality in modern societies.

#### **FORMS OF CRIME AND TYPES OF OFFENDERS**

The legal codes of the various states and of the federal government include hundreds of specific offenses, but the explanatory task is to develop a relatively small set of theories that make sense of this diverse collection of illegal activities.

In their response to this task, Michael Gottfredson and Travis Hirschi (1990) have argued that virtually all forms of criminal activity, and many kinds of deviant behavior as well, share certain features in common: they are spontaneous, unplanned actions requiring little or no skill for their commission. Further, Gottfredson and Hirschi have claimed that lawbreakers rarely specialize in specific acts of criminality. They concluded that virtually all of these varying criminal and deviant acts can be accounted for by a single, general theory that asserts they are the work of persons who are

characterized by low self-control. Accordingly, in their view, there is no need for schemes that classify types of crime or kinds of offenders or for separate theories to account for them.

However, many criminologists contend that there are relatively distinct forms of crime that differ from each other and also that the behavior of many criminals is relatively patterned. For example, some offenders concentrate their efforts upon larcenous acts while others of them are mainly involved in acts of violence.

A number of criminologists have tried to sort the diverse collection of illegal activities into a smaller number of sociologically meaningful groupings or crime forms (Farr and Gibbons 1990; Gibbons 1994). Some have singled out crude property crime, consisting of larceny, burglary, robbery, and kindred offenses, as one type of crime; others have placed homicide and assaultive acts into another crime type; while still others have treated forcible rape and other sexual offenses as yet another broad form of lawbreaking. Then, too, “white-collar” or organizational crime has often been singled out as a crime pattern (Sutherland 1949; Schrager and Short 1978; Coleman 1987), consisting in large part of criminal acts such as antitrust violations, financial fraud, and the like, carried on by corporations and other large organizations. “Organized crime” is still another type that has received a good deal of criminological attention. Some persons have also pointed to a collection of offenses that receive little visibility in the mass media and elsewhere and have termed these “folk crimes” (Ross 1960–1961, 1973) or “mundane crimes” (Gibbons 1983). Finally, “political crime” has been identified as a major pattern of lawbreaking (Turk 1982).

Although these groupings identify forms of lawbreaking that may differ from each other in important ways, it is also true that they are relatively crude in form in that the underlying dimensions or variables on which they are based have not been spelled out. Further, there is disagreement among criminologists as to the specific crimes that should be identified as instances of white-collar crime, mundane crime, or some other category.

Criminologists have also developed systems for sorting individual offenders into behavioral types (Gibbons 1965). Although related to crime

classification efforts, categorization of lawbreakers into types is a separate activity. While it may be possible to identify groupings such as predatory property crime, it may not be true that individual offenders specialize in that form of crime, hence it may be incorrect to speak of "predatory offenders" as a type of criminal. Most offender classification systems have been deficient in one respect or another (Gibbons 1985), but the most serious flaw is that they are oversimplified. Researchers have discovered that many offenders engage in a fairly diverse collection of offenses over their criminal "careers" rather than being crime specialists such as "burglars," "robbers," or "drug dealers" (Chaiken and Chaiken 1982).

### THEORIES OF CRIME

The number of theories regarding particular forms of crime is extensive, thus they cannot all be reviewed here (for a review of many of them, see Gibbons 1994). Additional to those theories mentioned previously, a sampling of the more important ones would include the *routine activities* explanation of predatory property crime. Lawrence Cohen and Marcus Felson (1979) contend that predatory property crime involves three major elements: the supply of motivated offenders, the supply of suitable targets, and the absence of capable guardians. In other words, these crimes are carried out by persons with criminal motives, but the incidence of such offenses also depends upon the number of opportunities to burglarize homes or to rob persons. Also, the number of burglaries from one community to another is influenced by the degree to which residents in local areas act as guardians by maintaining surveillance over homes in their neighborhoods or by taking other crime-control steps. This theory takes note of the fact that criminal opportunities have increased in the United States in recent decades at the same time that capable guardianship has declined, due principally to changes in employment patterns. In particular, the number of families in which both adult members work during the day has grown markedly, as has the number of employed, single-parent families. Research evidence lends considerable support to this theory (Cohen and Felson 1979).

Research evidence also indicates that income inequality is related to predatory property crime

(Braithwaite 1979; Carroll and Jackson 1983). Further, Leo Carroll and Pamela Jackson (1983) argue that the routine activities and income inequality arguments are interrelated. They suggest that the labor market trends identified in the former have led to increased crime opportunities, declines in guardianship, and heightened levels of income inequality.

### THEORIES OF CRIMINAL BEHAVIOR

While theories about crime patterns and rates have been developed principally by sociological criminologists, representatives of a number of disciplines have endeavored to identify factors and processes that explain the involvement or noninvolvement of specific individuals in lawbreaking. Three basic approaches can be noted: the biogenic-sociobiogenic, psychogenic, and sociogenic orientations. *Biogenic-sociobiogenic views* attribute the genesis or causes of lawbreaking, entirely or in part, to constitutional and hereditary factors, while *psychogenic perspectives* often contend that lawbreakers exhibit personality problems to which their illegal conduct is a response. By contrast, sociologists have most often advanced *sociological theories*, arguing that criminal behavior is learned in a socialization process by individuals who are neither biologically nor psychologically flawed. Also, some persons have constructed theories that combine or integrate elements of these three approaches, one case being James Wilson and Richard Herrnstein's (1985) argument that the behavior of criminals has genetic and constitutional roots and that offenders tend to be more mesomorphic in body build, less intelligent, and more burdened with personality defects than their noncriminal peers. Wilson and Herrnstein also contend that various social factors such as unemployment, community influences, and the like play some part in criminality.

Three generalizations can be made about biological theories: First, conclusive evidence supporting these arguments has not yet been produced; second, biological factors cannot be ruled out on the basis of the empirical evidence currently on hand; and third, if biological factors are involved in criminality, they are probably intertwined with social and psychological influences (Trasler 1987; Fishbein 1990).

In the first half of the twentieth century, psychological arguments about criminals centered on claims that these persons were feeble-minded, or somewhat later, that many of them were suffering from serious mental pathology of one sort or another. However, a number of reviews of the evidence, particularly that having to do with the alleged role of low intelligence or personality defects in criminality, turned up little or no support for such claims (Schuessler and Cressey 1950; Waldo and Dinitz 1967; Tennenbaum 1977).

Even so, there is a lingering suspicion among a number of criminologists that the criminal acts of at least some lawbreakers, including certain kinds of sexual offenders, can be attributed to faulty socialization and aberrant personality patterns (Gibbons 1994). Additionally, some psychologists have argued that even though the broad theory that criminality is due to marked personality defects on the part of lawbreakers lacks support, it is nonetheless true that individual differences in the form of personality patterns must be incorporated into criminological theories (Andrews and Wormith 1989; Blackburn 1993; Andrews and Bonta 1998). Moreover, in the opinion of a number of sociological criminologists, the argument that individual differences make a difference, both in accounting for criminality and for conformity, is persuasive (Gibbons 1989, 1994). Personality dynamics play a part in the behavior patterns that individuals exhibit, thus such concepts as role and status are often not entirely adequate to account for the behavior of individuals. Lawbreaking is quite probably related to the psychic needs of individuals as well as social and economic influences that play upon them. On this point, Jack Katz (1988) has explored the personal meanings of homicidal acts, shoplifting, and a number of other kinds of criminality to the persons who have engaged in these acts.

Sutherland's theory of *differential association* (Sutherland, Cressey, and Luckenbill 1992, pp.88–90) has been one of the most influential sociological theories about the processes through which persons come to engage in criminality. Sutherland maintained that criminal behavior, including techniques of committing crime and conduct definitions favorable to lawbreaking activity, is learned in association with other persons. Many of the associations of persons involve face-to-face contact, but conduct definitions favoring criminality

can be acquired indirectly from reference groups, that is, from persons who are important to individuals but with whom they do not directly associate. Sutherland also contended that associations vary in frequency, duration, priority, and intensity (the personal meaning or significance to individuals of particular social ties).

A very different theory, directed mainly at the explanation of juvenile delinquency, is that if, through faulty socialization, individuals fail to become bonded or connected to others (that is, if they do not develop positive attachments to adult persons such as parents or teachers), they will then be unlikely to refrain from misbehavior (Hirschi 1969). The emphasis in this argument is on the failure to acquire prosocial, nondelinquent sentiments rather than on the learning of antisocial ones. In this view, delinquency is the result of defective socialization rather than of socialization patterns through which criminal attitudes are learned. A more recent but related version of this argument, noted earlier in this essay, is that of Gottfredson and Hirschi (1990), who have claimed that criminality and other forms of deviance are most often engaged in by persons who are low on self-control.

### THEORETICAL INTEGRATION

Clearly, there is a wealth of differing arguments about the causes of crime and individual lawbreaking now in existence. Not surprisingly, then, a number of scholars have begun to ask whether it might be possible to amalgamate some or all of these varied lines of explanation into an integrated theory and thereby to develop a more powerful causal argument. Some criminologists have suggested that biological, psychological, and sociological contentions about crime all have some part to play in explaining crime and that, therefore, they should be integrated (Barak 1998). Others have proposed more limited forms of integration in which, for example, several sociological arguments might be merged into a single formulation (e.g., Tittle 1995; Braithwaite 1989) or in which psychological claims about lawbreaking might be linked or integrated with sociological ones. But to date, criminological investigators have not moved very far in the direction of sophisticated theoretical integrations. Further research on the interconnections between biological, psychological, and social factors in crime

and criminal conduct will probably be required if integrative efforts are to bear fruit.

(SEE ALSO: *Criminology*; *Juvenile Delinquency Theories*; *Social Control*)

## REFERENCES

- Andrews, D.A., and James Bonta 1998 *The Psychology of Criminal Conduct*, 2d ed. Cincinnati: Anderson.
- , and J. Stephen Wormith 1989 "Personality and Crime: Knowledge Destruction and Construction." *Justice Quarterly* 6:289-309.
- Barak, Gregg 1998 *Integrating Criminologies*. Boston: Allyn and Bacon.
- Baron, Larry, and Murray A. Straus 1987 "Four Theories of Rape: A Macrosociological Analysis." *Social Problems* 34:467-489.
- Blackburn, Ronald 1993 *The Psychology of Criminal Conduct*. New York: Wiley.
- Braithwaite, John 1979 *Inequality, Crime and Public Policy*. London: Routledge and Kegan Paul.
- 1989 *Crime, Shame, and Reintegration*. Cambridge, Eng.: Cambridge University Press.
- Carroll, Leo, and Pamela Irving Jackson 1983 "Inequality, Opportunity, and Crime Rates in Central Cities." *Criminology* 21:178-194.
- Chaiken, Jan, and Marcia Chaiken 1982 *Varieties of Criminal Behavior*. Santa Monica, Calif.: Rand.
- Chambliss, William 1975 "Toward a Political Economy of Crime." *Theory and Society* 2:148-155.
- Cohen, Lawrence E., and Marcus Felson 1979 "Social Change and Crime Rate Trends: A Routine Activities Approach." *American Sociological Review* 44:588-607.
- Coleman, James W. 1987 "Toward an Integrated Theory of White-Collar Crime." *American Journal of Sociology* 93:406-439.
- Cressey, Donald R. 1951 "Epidemiology and Individual Conduct: A Case From Criminology." *Pacific Sociological Review* 3:47-58.
- Currie, Elliott 1985 *Confronting Crime*. New York: Pantheon.
- Farr, Kathryn Ann, and Don C. Gibbons 1990 "Observations on the Development of Crime Categories." *International Journal of Offender Therapy and Comparative Criminology* 34:223-237.
- Fishbein, Diana H. 1990 "Biological Perspectives in Criminology." *Criminology* 28:27-72.
- Gibbons, Don C. 1965 *Changing the Lawbreaker*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1979 *The Criminological Enterprise*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1983 "Mundane Crime." *Crime and Delinquency* 29:213-227.
- 1985 "The Assumption of the Efficacy of Middle-Range Explanations: Typologies." In R.F. Meier, ed., *Theoretical Methods in Criminology*. Beverly Hills, Calif.: Sage.
- 1989 "Comment-Personality and Crime: Non-Issues, Real Issues, and a Theory and Research Agenda." *Justice Quarterly* 6:311-323.
- 1992 *Society, Crime, and Criminal Behavior*, 6th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- 1994 *Talking About Crime and Criminals*. Englewood Cliffs, N.J.: Prentice-Hall.
- Gibbs, Jack P. 1985 "The Methodology of Theory Construction in Criminology." In R.F. Meier, ed., *Theoretical Methods in Criminology*. Beverly Hills, Calif.: Sage.
- Gottfredson, Michael R., and Travis Hirschi 1990 *A General Theory of Crime*. Stanford, Calif.: Stanford University Press.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- Katz, Jack 1988 *Seductions of Crime*. New York: Basic Books.
- Messner, Steven F., and Richard Rosenfeld 1994 *Crime and the American Dream*. Belmont, Calif.: Wadsworth.
- Polk, Kenneth 1994 *When Men Kill*. Cambridge, Eng.: Cambridge University Press.
- Quinney, Richard 1974 *Critique of Legal Order*. Boston: Little, Brown.
- 1977 *Class, State, and Crime*. New York: McKay.
- Ross, H. Laurence 1960-61 "Traffic Law Violation: A Folk Crime." *Social Problems* 9:231-241.
- 1973 "Folk Crime Revisited." *Criminology* 11:41-85.
- Schrager, Laura Shill, and James F. Short, Jr. 1978 "Toward a Sociology of Organizational Crime." *Social Problems* 25:407-419.
- Schuessler, Karl F., and Donald R. Cressey 1950 "Personality Characteristics of Criminals." *American Journal of Sociology* 55:476-484.
- Sutherland, Edwin H. 1949 *White Collar Crime*. New York: Dryden.
- , Donald R. Cressey, and David F. Luckenbill 1992 *Principles of Criminology*, 11th ed. Dix Hills, N.Y.: General Hall.
- Tennenbaum, D. J. 1977 "Personality and Criminality: A Summary and Implications of the Literature." *Journal of Criminal Justice* 5:225-235.



- Tittle, Charles R. 1995 *Control Balance*. Boulder, Colo.: Westview.
- Trasler, Gordon 1987 "Biogenetic Factors." In Herbert C. Quay, ed., *Handbook of Juvenile Delinquency*. New York: Wiley.
- Turk, Austin T. 1982 *Political Criminality*. Beverly Hills, Calif.: Sage.
- van Dijk, Jan J., Pat Mayhew, and Martin Killias 1990 *Experiences of Crime Across the World*. Deventer, The Netherlands: Kluwer Law and Taxation Publishers.

DON C. GIBBONS

## CRIMINAL AND DELINQUENT SUBCULTURES

A subculture is derivative of, but different from, some larger referential culture. The term is used loosely to denote shared systems of norms, values, or interests that set apart some individuals, groups, or other aggregation of people from larger societies and from broader cultural systems. Common examples include youth subcultures, ethnic subcultures, regional subcultures, subcultures associated with particular occupations, and subcultures that develop among people who share special interests such as bird-watching, stamp collecting, or a criminal or delinquent behavior pattern.

Neither membership in a particular category (age, ethnicity, place of residence, occupation) nor behavior (bird-watching, stamp collecting, crime, or delinquency) is sufficient to define a subculture, however. The critical elements are, rather, (1) the degree to which values, norms, and identities associated with membership in a category or types of behaviors are shared, and (2) the nature of relationships, within some larger cultural system, between those who share these elements and those who do not.

In these terms, criminal or delinquent subcultures denote systems of norms, values, or interests that support criminal or delinquent behavior. The many behaviors specified in law as criminal or delinquent are associated with many criminal and delinquent subcultures. The norms, values, or interests of these subcultures may support particular criminal acts, a limited set of such acts (e.g., a subculture of pickpockets vs. a subculture of hustlers). "Professional criminals," for example, take pride in their craft, organize themselves for

the safe and efficient performance of the crimes in which they specialize, and generally avoid other types of criminal involvement that might bring them to the attention of the authorities (see Sutherland 1937; Cressey 1983). Not all criminal subcultures are this specific, however. Some are simply opportunistic, embracing several types of criminal behavior as opportunities arise. So it is with delinquent subcultures, where specialization is rare.

While delinquent subcultures typically are associated with a broad range of illegal behaviors, among delinquent groups and subcultures there is great variation in the nature and strength of group norms, values, and interests. Moreover, the extent to which delinquent behavior is attributable to these factors is problematic. Much delinquent behavior of highly delinquent gangs, for example, results from the operation of group processes rather than group norms per se (see Short 1997). The normative properties of groups vary greatly, but even the most delinquent gang devotes relatively little of its group life to the pursuit of delinquent behaviors. Further, when gangs do participate in delinquent episodes, some members of the gang typically do not become involved. This is so, in part, because subcultures typically consist of *collections of normative orders*—rules and practices related to a common value (Herbert 1998)—rather than norms oriented around a single value (such as being "macho," "cool," or exceptionally gifted in some way). In addition, individuals who are associated with a particular subculture tend also to be associated with other subcultures. Simply being associated with a subculture thus is unlikely to be a good predictor of the behavior of a particular individual.

For analytical purposes it is important to distinguish between subcultures and the particular individuals and groups who share the norms, values, and interests of the subculture. While members of a delinquent gang may be the sole carriers of a particular subculture in a particular location, some subcultures are shared by many gangs. Conflict subcultures, for example, are shared by rival fighting gangs among whom individual and group status involves values related to the defense of "turf" (territory) and "rep" (reputation) and norms supportive of these values. Subcultures oriented to theft and other forms of property crime vary in the extent to which they are associated with particular

groups. Some types of property crimes require organization and coordination of activities in order to be successful. Some also necessarily involve the efforts of others, such as “fences,” in addition to members of a criminally organized gang (see Klockars 1974). Others, such as mugging and other types of robbery, may be carried out by individual offenders, who nevertheless share a subculture supportive of such behavior. Most drug-using subcultures tend to be less oriented toward particular groups than are conflict subcultures because the subcultural orientation is toward drug consumption, and this orientation can be shared with other drug users in many types of group situations. To the extent that a subculture is oriented to experiences associated with a particular group, however, a drug-using subculture may also be unique to that group.

As this analysis suggests, cultures, subcultures, and the groups associated with them typically overlap, often in multiple and complex ways. To speak of youth culture, for example, is to denote a subculture of the larger adult-dominated and institutionally defined culture. Similarly, delinquent subcultures contain elements of both youth and adult cultures. Terry Williams’s (1989) lower class, minority, “cocaine kids,” for example, were entrepreneurial, worked long hours, and maintained self-discipline—all important elements in the achievement ideology of the American Dream (see Messner and Rosenfeld 1994; also Fagan 1996; Hagan, et al. 1998). Most saw their involvement in the drug trade as a way to get started in legitimate business or to pursue other conventional goals, and a few succeeded at least temporarily in doing so. The criminal subculture with which they identified shared a symbiotic relationship with their customers (including many middle- and upper-class persons), who shared subcultural values approving drug use but who participated in the subculture of drug distribution only as consumers. For the young drug dealers, selling drugs was a way to “be somebody,” to get ahead in life, and to acquire such things as jewelry, clothing, and cars—the symbols of wealth, power, and respect.

The nature of relationships between delinquent subcultures and larger cultural systems is further illustrated by Mercer Sullivan’s study of cliques of young men in Brooklyn, among whom the “cultural meaning of crime” was “constructed

in . . . interaction out of materials supplied from two sources: the local area in which they spend their time almost totally unsupervised and undirected by adults, and the consumerist youth culture promoted in the mass media” (1989, p. 249). In other words, crime becomes meaningful to young men when they interact with one another and when they participate in youth culture, with its highly commercialized messages. The research literature on criminal and delinquent subcultures is devoted largely to describing and accounting for these types of varied and complex relationships.

### THEORY AND RESEARCH

Despite efforts to define the theoretical construct, “subculture”—and related constructs—more precisely and to describe and account for the empirical reality they represent, no general theory of subcultures has emerged (Yinger 1960, 1977). Instead, research has continued to reveal enormous variation in subcultures, and theory has proceeded by illustration and analogy, with little progress in measurement or formal theoretical development. Despite this scientifically primitive situation, principles of subcultural formation have been identified, and knowledge of it has advanced.

It is a “fundamental law of sociology and anthropology,” noted Daniel Glaser, that “social separation produces cultural differentiation” (1971, p. 90). More formally, and more cautiously, social separation is a necessary but not sufficient condition for the formation of subcultures. To the extent that groups or categories of persons are socially separated from one another, subcultural formation is likely to occur.

Albert Cohen argued that a “crucial” (perhaps necessary) “condition for the emergence of new cultural forms is the existence, in effective interaction with one another, of a number of actors with similar problems of adjustment” (1955, p. 59). While the notion of “similar problems of adjustment” can be interpreted to include problems faced by quite conventional people with special interests who find themselves “in the same boat” with others who have these same interests (let us say, bird-watchers), this condition seems especially appropriate to subcultures that embrace vandalism, “hell raising,” and other types of nonutilitarian

delinquent behavior. Observing that this type of behavior occurs most frequently among working-class boys, Cohen hypothesized that this type of delinquent subculture was formed in reaction to status problems experienced by working-class boys in middle-class institutions such as schools. Many working-class boys are inadequately prepared for either the educational demands or the discipline of formal education, and they are evaluated poorly in terms of this "middle-class measuring rod." Working-class girls are less pressured in these terms, Cohen argued, because they are judged according to criteria associated with traditional female roles, and they are subject to closer controls in the family.

The solution to their status problems, as some working-class boys see it, according to Cohen's theory, is to reject the performance and status criteria of middle-class institutions, in effect turning middle-class values upside down. The theory thus seeks to account for the highly expressive and hedonistic quality of much delinquency and for the malicious and negativistic quality of vandalism.

Cohen did not attempt to account for the delinquent behavior of particular individuals or for the behavior of all working-class boys. Most of the latter do not become delinquent—at least not seriously so. They choose instead—or are channeled into—alternative adaptations such as the essentially nondelinquent "corner boys" or the high-achieving "college boys" described by William Foote Whyte (1943).

The forces propelling youngsters into alternative adaptations such as these are not completely understood. Clearly, however, working-class and lower-class boys and girls tend to be devalued in middle-class institutional contexts. Their marginality sets the stage for subcultural adaptations. Delinquents and criminals occupy even more marginal positions. This is particularly true of persistent delinquents and criminals who commit serious crimes, in contrast to those who only rarely transgress the law and with little consequence. When marginality is reinforced by labeling, stigmatization, or prejudicial treatment in schools and job markets, "problems of adjustment" magnify. The common ecological location of many delinquents, in the inner-city slums of large cities, and their coming together in schools, provides the setting for

"effective interaction." The result often is the formation of delinquent youth gangs, an increasingly common organizational form taken by delinquent subcultures (see Thrasher 1927; Klein, 1995; Short 1997).

There is no universally agreed-upon definition of youth gangs. For theoretical purposes, however, it is useful to define gangs as groups whose members meet together with some regularity over time and whose membership is group-selected, based on group-defined criteria. Similarly, organizational characteristics are group determined. Most importantly, gangs are not adult-sponsored groups. The nature of relationships between young people and conventional adults is the most critical difference between gangs and other youth groups (see Schwartz 1987). Gang members are less closely tied to conventional institutions and therefore less constrained by institutional controls than are nongang youth. Group processes of status achievement, allocation, and defense are more likely to result in delinquent behavior among gangs than is the case among adult-sponsored groups.

#### **HOW DO CRIMINAL AND DELINQUENT SUBCULTURES GET STARTED?**

Delinquent and criminal subcultures have a long history in industrialized societies (Cressey 1983). Herman and Julia Schwendinger (1985) trace the origins of adolescent subcultures, including delinquent varieties, to social changes that began in the sixteenth century. Traditional economic and social relationships were greatly altered with the advent of capitalism and the Industrial Revolution in Western Europe, leaving in their wake large numbers of unemployed persons and disrupting communities, families, and other primordial groups. Cut adrift from traditional crafts and communities, thousands roamed the countryside, subsisting as best they could off the land or by victimizing travelers. The "dangerous classes" eventually settled in cities, again to survive by whatever means were available, including crime. Criminal subcultures and organizational networks often developed under these circumstances.

The Schwendingers emphasize that criminal subcultures developed as a result of structural

changes associated with capitalist values and their accompanying norms and interests—individualism and competitiveness, acquisitiveness and exploitativeness—and the relationships that developed during this period between capitalists and emerging nation-states. While many of the facts upon which this Marxist interpretation is based are generally accepted, careful historical analysis of economic and political systems and their consequences cautions against any simple or straightforward interpretation (see Chirot 1985). The connection between global phenomena and crime and delinquency always is mediated by historical, cultural, and local circumstances—by the historically concrete (see Tilly 1981).

James Coleman and his associates (1974) identified more recent social changes that were associated with the rise and spread of youth culture throughout the United States: the Baby Boom following World War II and the increased affluence of young people associated with post-World War II economic prosperity combined to create a huge youth market with great economic power. At the same time, young people were spending more time in school and therefore delaying their entrance into the labor force; growing numbers of women entered the work force, further separating mothers from youth in homes and neighborhoods; adults increasingly were employed in large organizations where young people were not present; and mass media, catering more and more to the youth market, were greatly expanded. At the close of the twentieth century, each of these broad social changes was more pronounced—and their influence was more widespread throughout the world—than was the case when these observations were made.

In contrast to accounts of the origins of Western European youth cultures, and of youth culture in the United States, Ko-lin Chin (1996) traces the development of Chinese youth gangs in the United States to ancient secret society traditions, and to the more recent Triad societies that formed in the late seventeenth century in China, and their counterpart tongs in the United States. Formed as political groups representing disfavored Chinese officials and the alienated poor, these groups initially stressed patriotism, righteousness, and brotherhood as primary values. However, their secret nature was conducive to clandestine activities such as gambling, prostitution, and running opium dens.

Competition among Triad societies in these activities often led to violence. Failure to achieve political power led to their further transformation into criminal organizations involved in extortion, robbery, drug trafficking, and other serious crimes.

Street gangs comprised of Chinese adolescents did not form in the United States until the late 1950s, but their numbers increased dramatically during the following decade, when changed immigration laws permitted more Chinese to enter the country. Conflict between foreign- and American-born youths led to the emergence of gangs, some seriously delinquent. Chin attributes this development to alienating problems experienced by immigrant Chinese youth in their families, schools, and communities, in dramatic contrast to their American-born counterparts. From the beginning, many Chinese youth gangs have been associated in a variety of ways with established adult secret societies in the United States, Hong Kong, or Taiwan, or in all three places. The existence of “Triad-influenced” organizations has been critical to the types of gangs that have emerged in Chinese communities and to the nature of their criminal activities.

While the origins of delinquent subcultures may reside in antiquity, the formation and evolution of modern variations of them can be explained in terms of more immediate macro-level developments. Some of these developments relate primarily to the ongoing activities and interests of gang members rather than to racial or ethnic changes, or to sweeping social changes. The nature of these influences is illustrated by a drug-using group studied by James Short, Fred Strodbeck, and their associates (Short and Strodbeck 1965; see also Short 1997, 1998). This gang was observed as it developed its own unique subculture. The subculture of the “Pill Poppers,” as they became known to the research team, evolved from their relationship with a larger, conflict-oriented gang of which they had previously been a part. The Pill Poppers’ preoccupation with drugs and their refusal to participate in the more bellicose activities of the larger gang led to their withdrawal and increasing isolation, by mutual agreement. The researchers were able to observe the evolution of this subculture, which was characterized by normative approval of drug consumption, an elevated value on “getting high,” and mutual interest in the

“crazy” things that happened to them when they were under the influence of drugs. The latter, in particular, became legendary within the group, being told and retold with nostalgia and humor when members of the gang were together. The subculture of this gang contrasted sharply with that of other gangs that were participating in a well-developed conflict subculture.

#### LEVELS OF EXPLANATION OF CRIMINAL AND DELINQUENT SUBCULTURES

The theories of criminal and delinquent subcultures are macro-level theories (Short 1998). Their purpose is to identify what it is about political, economic, and other social systems that explain the emergence and the social distribution of these phenomena. Other theories at this level focus on the impact of local and broader community opportunities on delinquent subcultures and on youth subcultures generally. Walter Miller (1958) related lower-class culture to gang delinquency, while Richard Cloward and Lloyd Ohlin (1960) found that different delinquent subcultures were associated with the availability of legitimate and illegitimate economic opportunities in local communities. Gary Schwartz (1987) stressed the importance of local community youth-adult authority relationships in determining the nature of youth subcultures, including the extent and the nature of delinquent activities associated with them. Mercer Sullivan (1989) related gang adaptations to more global economic developments such as the transfer of manufacturing jobs from the United States to other countries and an increasingly segmented labor market, which has resulted in the concentration of low-wage and surplus labor in inner-city minority communities (see also Hagedorn 1987, 1998).

William Julius Wilson (1987) provided the most compelling theoretical argument relating economic changes to crime and delinquency among the “truly disadvantaged.” Wilson argued that a permanent underclass emerged in the United States during the 1960s and 1970s. Research on delinquent subcultures supports this argument and documents related changes in gang membership. Because fewer good jobs were available to poor, minority young men, more gang members continued their association with gangs as they entered

adulthood. In the past, gang members typically “grew out of” the gang to take jobs, get married, and often become associated with adult social clubs in stable, ethnically-based communities. These options have become less viable among the poor of all races, but minorities have increasingly become the truly disadvantaged. Gang organization has been affected by this change, as older members assume or continue leadership roles. The result often has been that gang involvement in criminal activities has become more sophisticated and instrumental, and younger members have been exploited in criminal enterprise. Relationships between young people and conventional adults also suffer, as older, stable role models and monitors of youthful behavior are replaced by young adult, often criminal, role models for the young (see Anderson 1990, 1999).

Sullivan’s study of groups and young males in three Brooklyn communities—black, predominantly Latino, and white—is particularly significant in this regard. The young men in Hamilton Park, the white group, were able to find better jobs than were the others at all ages. More important, because “they had become more familiar with the discipline of the workplace,” as they grew older they were able to secure better-quality jobs and to hold on to them, compared to the minority youth studied by Sullivan. Familiarity with the discipline of the workplace is a type of *human capital* that was made possible by an important type of *social capital*—the superior personal networks that the Hamilton Park youth shared with the adult community (Sullivan 1989, pp. 105, 226). The minority youth were disadvantaged, with respect to both human and social capital, in the family and in other ways (see Coleman 1988). Thus, while individual human and social capital are acquired through personal experience, communities and neighborhoods also vary in their stock of human and social capital, qualitatively and quantitatively—yet another way in which delinquent and criminal subcultures reflect aspects of their “parent” (larger) cultures.

All macro-level theories make certain assumptions about the individual level of explanation. The most prominent of these assumptions is that individuals learn subcultural norms, values, and the behaviors they encourage, by interacting with others in their environment. The general processes of learning have been well established by research and theory (see Bandura 1986; Eron 1987).

The child's most important early learning experiences take place in the family, but other influences quickly assert themselves, especially as children associate with age and gender peers. With the beginning of adolescence, the latter become especially powerful as young people experience important biological and social changes. It is at this point that youth subcultures become especially significant.

Subcultures are dynamic and ever changing, influenced by both external and internal forces and processes. Substantial knowledge gaps exist at each level of explanation; precisely how they relate to each other is not well understood, however. By documenting the ongoing interaction among gang members and between gang members and others, group processes help to inform the nature of the relationship in ways that are consistent with what is known at the individual and macro levels of explanation. Both intragang and intergang fighting often serve group purposes, for example, by demonstrating personal qualities that are highly valued by the gang or by reinforcing group solidarity (see Miller, Geertz, and Cutter 1961; Jansyn 1966). Gang conflict often occurs when a gang believes its "rep," its "turf," or its resources (for example, its share of a drug market) are threatened by another gang. Threats to individual or group status often result in violent or other types of criminal behavior (Short 1997).

### CONCLUSION

The rich research materials that have accumulated regarding criminal and delinquent subcultures suggest two conclusions: (1) there is a great variety of such adaptive phenomena, and (2) they are of an ever-changing nature. Because they both effect social change and adapt to it, subcultures—including criminal and delinquent subcultures—continue to be important theoretically, empirically, and practically (that is, as a matter of social policy). Robert Sampson and his colleagues (1997) find that willingness to intervene on behalf of the common good, together with collective efficacy (social cohesion among neighbors), is associated with lower rates of violent crime in Chicago neighborhoods. This suggests that developing ways to encourage identification of neighbors with each other and encouraging them to help one another in their common interests will enhance local social

control and help weaken the influence of subcultures that encourage criminal and delinquent behavior.

(SEE ALSO: *Crime, Theories of; Criminology; Juvenile Delinquency and Juvenile Crime; Juvenile Delinquency, Theories of*)

### REFERENCES

- Anderson, Elijah 1999 *The Code of the Streets*. New York: W.W. Norton.
- 1990 *Streetwise: Race, Class, and Change in an Urban Community*. Chicago: University of Chicago Press.
- Bandura, Albert 1986 *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, N.J.: Prentice-Hall.
- Chin, Ko-lin 1996 *Chinatown Gangs: Extortion, Enterprise, and Ethnicity*. New York: Oxford University Press.
- Chiro, Daniel 1985 "The Rise of the West." *American Sociological Review* 50:181–195.
- Cloward, Richard A., and Lloyd E. Ohlin 1960 *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press.
- Cohen, Albert K. 1955 *Delinquent Boys: The Culture of the Gang*. New York: Free Press.
- Coleman, James S. 1988 "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (Suppl.):S95–S120.
- , Robert H. Bremner, Burton R. Clark, John B. Davis, Doroty H. Eichorn, Zvi Griliches, Joseph F. Kett, Norman B. Ryder, Zahava Blum Doering, and John M. Mays 1974 *Youth Transition to Adulthood*. Report of the Panel of the President's Science Advisory Committee. Chicago: University of Chicago Press.
- Cressey, Donald R. 1983 "Delinquent and Criminal Subcultures." In S. E. Kadish, ed., *Encyclopedia of Crime and Justice*. New York: Free Press.
- Eron, Leonard D. 1987 "The Development of Aggressive Behavior from the Perspective of a Developing Behaviorism." *American Psychologist* 42:435–442.
- Fagan, Jeffrey 1996 "Gangs, Drugs, and Neighborhood Change." In C. Ronald Huff, ed., *Gangs in America*, 2d ed. Newbury Park, Calif.: Sage.
- Glaser, Daniel 1971 *Social Deviance*. Chicago: Markham.
- Hagan, John, Gerd Heffler, Gabriele Classen, Klaus Boehnke, and Hans Merckens 1998 "Subterranean Sources of Subcultural Delinquency Beyond the American Dream." *Criminology* 36:309–341.

- Hagedorn, John M. 1998 "Gang Violence in the Post-Industrial Era." In Michael Tonry and Mark H. Moore, eds., *Youth Violence*, vol. 24, *Crime and Justice*. Chicago: University of Chicago Press.
- , with Perry Macon 1987 *People and Folks: Gangs, Crime, and the Underclass in a Rustbelt City*. Chicago: Lake View Press.
- Herbert, Steve 1998 "Police Subculture Reconsidered." *Criminology* 36:343–369.
- Jansyn, Leon R. 1966 "Solidarity and Delinquency in a Street-Corner Group." *American Sociological Review* 31:600–614.
- Klein, Malcolm W. 1995 *The American Street Gang*. New York: Oxford University Press.
- Klockars, Carl B. 1974 *The Professional Fence*. New York: Free Press.
- Messner, Steven F., and Richard Rosenfeld 1994 *Crime and the American Dream*. Belmont, Calif.: Wadsworth.
- Miller, Walter B. 1958 "Lower Class Culture as a Generating Milieu of Gang Delinquency." *Journal of Social Issues* 14:5–19.
- , Hildred Geertz, Henry S. G. Cutter 1961 "Aggression in a Boys' Street-Corner Group." *Psychiatry* 24:283–298.
- Sampson, Robert J., S.W. Raudenbush, and F. Earls 1997 "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918–924.
- Schwartz, Gary 1987 *Beyond Conformity or Rebellion: Youth and Authority in America*. Chicago: University of Chicago Press.
- Schwendinger, Herman, and Julia Siegel Schwendinger 1985 *Adolescent Subcultures and Delinquency*. New York: Praeger.
- Short, James F., Jr. 1998 "The Level of Explanation Problem Revisited—The American Society of Criminology 1997 Presidential Address." *Criminology* 36:3–36.
- 1997 *Poverty, Ethnicity, and Violent Crime*. Boulder, Colo.: Westview.
- , and Fred L. Strodbeck 1965 *Group Process and Gang Delinquency*. Chicago: University of Chicago Press.
- Sullivan, Mercer L. 1989 *"Getting Paid": Youth Crime and Work in the Inner City*. Ithaca, N.Y.: Cornell University Press.
- Sutherland, Edwin H. 1937 *The Professional Thief*. Chicago: University of Chicago Press.
- Thrasher, Frederic M. 1927 *The Gang: A Study of 1,313 Gangs in Chicago*. Chicago: University of Chicago Press.
- Tilly, Charles 1981 *As Sociology Meets History*. New York: Academic Press.
- Whyte, William Foote 1943 *Street Corner Society*. Chicago: University of Chicago Press.
- Williams, Terry 1989 *The Cocaine Kids: The Inside Story of a Teenage Drug Ring*. Menlo Park, Calif.: Addison-Wesley.
- Wilson, William Julius 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Yinger, Milton 1960 "Contraculture and Subculture." *American Sociological Review* 23:625–635.
- 1977 "Countercultures and Social Change." *American Sociological Review* 42:833–853.

JAMES F. SHORT, JR.

## CRIMINAL SANCTIONS

The quality and quantity of normative sanctions have been viewed as a reflection of the nature of social solidarity (Durkheim 1964; Black 1976). In simple societies where the level of willing conformity is high, normative sanctions tend to be informal in nature, substantive in application, and limited in use. In complex societies where levels of willing conformity are lower, normative sanctions are more likely to be formal in nature, procedural in application, and frequent in use (Michalowski 1985). The increasing stratification, morphology, and bureaucracy of modern society have given rise to the predominance of formal justice in the form of criminal law and criminal sanctions (Black 1976). Consequently, the nature of crime has been transformed from an offense committed by one individual against another in the context of community to an offense committed against the society as a whole (Christie 1977). Behaviors considered harmful to the moral, political, economic, or social well-being of society are defined as criminal and thereby worthy of formal state sanctions (Walker 1980). Criminal behaviors include transgressions of both the prohibitions and obligations that define a particular society. Behaviors come to be defined as crimes through the process of criminalization, which includes the calculation of proportional sanctions for each crime.

Criminal sanctions include capital punishment, imprisonment, corporal punishment, banishment,

house arrest, community supervision, fines, restitution, and community service. The type and severity of criminal sanctions are prescribed by criminal law (Walker 1980). The quality and quantity of criminal penalties are determined by both the perceived seriousness of the offense and the underlying philosophy of punishment. Punishment by the state on behalf of society has traditionally been justified on either consequential or nonconsequential grounds (Garland 1994). Consequentialism justifies punishment as a means to the prevention of future crime. The utilitarian approach, for example, allows society to inflict harm (by punishment) in order to prevent greater harms that would be caused by future crimes. The utilitarian approach to criminal sanctions is governed by a set of limiting principles (Beccaria 1980). According to utilitarianism criminal sanctions should not be used to penalize behavior that does not harm, the severity of the penalty should only slightly outweigh the benefit derived from the criminal behavior, and alternatives to punishment should be utilized when they prove to be as effective (Bentham 1995).

Nonconsequentialism, on the other hand, justifies punishment as an intrinsically appropriate response to crime (Duff and Garland 1994). The retributive approach, for example, mandates punishment as a way of removing the advantage originally gained by the criminal behavior or as a means of restoring the moral balance that was lost as a result of a crime. According to the retributivist perspective crime separates the offender from the community and it is only through punishment that the separation can be repaired (Morris 1981). While the utilitarian and retributivist justifications have dominated the philosophical discussion of punishment, more recent justifications such as rehabilitation (Rotman 1990) and incapacitation (Morris 1982) have also been viewed in terms of their consequential and nonconsequential nature. From a sociological perspective these philosophies of punishment represent ideal types against which the actual practice of punishment must be examined.

#### THE SOCIOLOGY OF CRIMINAL SANCTIONS

Several lines of inquiry comprise the sociology of criminal sanctions including the examination of

the criminalizing process, the empirical assessment of the effectiveness of particular sanctions, and the analysis of the relationship between social structural changes and the evolution of criminal law.

**Critique of the Criminalization Process.** The question of which behaviors become defined as criminal and deemed worthy of punitive sanctions has been central to the study of crime and law. The *functional perspective* of criminal law suggests that the criminalization of a particular behavior is the result of a consensus among members of a society (Durkheim 1964). In the *consensus view* criminal law encompasses the behaviors that have been determined to be most threatening to the social structure of society and the well-being of its members (Wilson 1979).

According to the functionalist perspective, criminal law adapts to changes in the normative consensus of society. As society evolves, behaviors once considered criminal may be decriminalized while behaviors that had previously been acceptable may become criminalized. In this context criminal law and its accompanying sanctions are viewed as the collective moral will of society. Punishment is viewed as serving the essential function of creating a sense of moral superiority among the law-abiding members of society, thereby strengthening their social solidarity (Durkheim 1964). Therefore, crime is considered to be inevitable in the functionalist view because the pressure exerted against those who do not conform to normative expectations is necessary to reinforce the willing conformity among members of society. Solidarity results from the social forces directed against transgressors, with the most powerful of these forces being criminal sanctions (Vold 1958). Criminal sanctions enable societies to distinguish between behaviors that are simply considered unacceptable and those behaviors that are considered truly harmful.

Conversely, the conflict perspective suggests that the designation of behaviors as criminal is determined within a context of unequal power. Although various forms of the conflict perspective identify different sources for this power imbalance, they are in general agreement as to the results of the criminalizing process. According to the conflict view, society is made up of groups with conflicting needs, values, and interests that are



mediated by an organized state that represents the needs, values, and interests of groups that possess the power to control the state. Such control includes the capacity to determine which behaviors are considered to be criminal and which behaviors are not. As a result behaviors more likely to be committed by the less powerful are defined as criminal while behaviors more likely committed by the powerful are not defined as criminal (Bernard 1983; Reiman 1998). Furthermore, the conflict view suggests that the application of both criminal law and criminal sanctions are skewed in favor of those with power. Within the conflict perspective various theories are differentiated according to their identified source of group conflict. Among the major conflict theories are those that identify a conflict of cultural groups (Sellin 1938), those that identify a conflict of norms (Vold 1958), those that identify a conflict of socioeconomic interests (Marx 1964), and those that identify a conflict of bureaucratic interests (Turk 1969). Despite their different origins, conflict theories are united in their assessment that criminal law and criminal sanctions are utilized to support the average best interest of those with power. As a result the state is assumed to apply criminal sanctions in such a manner that they are not perceived as overly coercive while at the same time preserving the existing power arrangements thereby maintaining the legitimacy of criminal law (Turk 1969).

**An Empirical Context for Criminal Sanctions.**

Empirical assessments of criminal sanctions have been primarily carried out within the contexts of *penology* and the *sociology of punishment*. Penology is a practically oriented form of social science that traces and evaluates various practices of penal institutions and other punitively oriented institutions of the modern criminal justice system (Duff and Garland 1994). Penology began as an extension of the prison itself that its sole purpose was to evaluate the objectives of the institution and develop more efficient ways of achieving these institutional objectives. More recently penology has expanded its inquiry to include the assessment of the entire criminal sanctioning process (studying the prosecution, the court process, as well as alternative sanctions such as probation, fines, electric monitoring, community service, and restitution). Increasingly, penology has been guided by the major theories of crime causation thus becoming a recognized area of study within the discipline of

criminology. Contemporary, penological research serves as the empirical foundation for policy development (Bottomley 1989). Despite its more recent criminological grounding, penological research is still viewed as primarily evaluative rather than critical. Penology assumes an exclusively punitive perspective toward transgressions of criminal law, therefore, it is primarily concerned with the relative effectiveness of the various punitive responses to criminal transgressions. The critical dimension of penology is limited to identifying problems within existing institutions and suggesting ways to more efficiently achieve the goal of punishment. Because penology is limited to the study of punishment systems any questions concerning alternative models of crime control such as compensation or conciliation are addressed by the sociological study of punishment (Garland 1990).

Unlike penology, the sociology of punishment raises more fundamental questions concerning the relative effectiveness of both punitive and nonpunitive responses to normative transgressions. Within the context of the sociology of punishment special attention is directed to the way in which society organizes and deploys its power to respond to transgressions (Duff and Garland 1994). It is primarily concerned with the relationship between punishment and society. Punishment is examined as a social institution that reflects the nature of social life. The sociology of punishment is fundamentally interested in why particular types of societies employ particular types of criminal sanctions (Garland 1990). In addition, the sociological analysis of punishment examines the social conditions that necessitate certain styles of normative sanctioning such as the use of penal sanctions in social situations that are characterized by high levels of inequality, heterogeneity, and bureaucratic interaction (Black 1976). The sociology of punishment also includes the study of correlations between the application of various sanctions (e.g. imprisonment, probation, fines, etc.) and demographic variables such as class, race, gender, occupation, and education (Reiman 1998; Zimring and Hawkins 1990). Because of the dominance of imprisonment as the preferred sanction for serious criminal transgressions in Western societies much attention has been directed to the effect of penal confinement on prisoners. Sociological inquiries into the inner world of the prison have examined the social adaptation of prisoners to the unique

requirements and relationships that characterized "The total institution" (Goffman 1961; Sykes 1958). The sociological study of punishment suggests that the adoption and application of criminal sanctions are better understood as a reflection of social, political, and economic reality than as the product of moral consensus (Garland 1990).

**The Evolution of Criminal Sanctions.** Sociologists have paid considerable attention to the qualitative evolution of criminal sanctions. Special attention has been directed to the relationship between changes in the socioeconomic structure and the evolution of criminal law in modern Western societies (Melossi and Pavarini 1981). Within the more general discussion of law and society, the study of punishment and social structure emerged (Rusche and Kirchheimer 1939). The sociology of penal systems argues that the transformation of penal systems cannot primarily be explained by the changing needs of crime control. Instead, the sociological analysis of shifts in the systems of punishment and criminal law are best understood in terms of their relationship to the prevailing system of production (Sellen 1976). Each society generates criminal law practices and types of punishments that correspond to the nature of its productive relationships. Research on punishment and social structure has focused attention on the origin and evolution of penal systems, the use or avoidance of specific punishments, and the intensity of penal practices in terms of larger social forces—particularly economic and fiscal forces (Rusche and Kirchheimer 1939). For example, in slave economies where the supply of slaves was inadequate, penal slavery emerged; the emergence of the factory system decreased the demand for convict labor, which in turn led to the rise of reformatories and industrial prisons (Melossi and Pavarini 1981). Although the economic analysis of criminal sanctions dominated the early literature on the sociology of punishment, later analysis has focused more broadly on all five aspects of social life (i.e., stratification, morphology, bureaucracy, culture, and social control).

#### CRIMINAL SANCTIONS AND THE ASPECTS OF SOCIAL LIFE

The comparative study of law has identified criminal law and criminal sanctions as quantifiable variables (Black 1976). By quantifying criminal law

and criminal sanctions it is possible to compare both the type and the amount of law utilized by one society versus another. The quantification of law can then be correlated to the nature of social life. The type and amount of criminal law can either be correlated to individual aspects of social life or the collective integration of all the variables of social life. Social life consists of five variable aspects: *stratification* is the vertical aspect of social life (the hierarchy of people relative to their power), *morphology* is the horizontal aspect of social life (the distribution of people in relation to each other), *bureaucracy* is the corporate aspect of social life (the capacity for collective action), *culture* is the symbolic aspect of social life (the representation of ideas, beliefs, and values), and *social control* is the normative aspect of social life (the definition of deviance and the response to it). The first four aspects of social life collectively determine the style of social control.

Comparative legal studies have identified five styles of social control; *penal control*, *compensatory control*, *therapeutic control*, *educative control* and *conciliatory control* (Fogel 1975). The response to normative transgressions under each style is consistent with the general standard of behavior recognized by society at large. Transgressions of the prohibition standard of the penal style of control require a punitive solution in order to absolve the guilt of the transgressor. Violation of the obligation standard of the compensatory style requires payment in order to eliminate the debt incurred by the violation. Behaviors that fall outside the standard of normality that characterizes the therapeutic style require treatment to restore predictability to social interaction. Failure to meet the knowledge standard of the educative style requires reeducation in order to attain a full understanding of society's norms and underlying values. Disruption of the harmony standard of the conciliatory style requires resolution in order to reestablish the relationships among members of the community (Fogel 1975).

Criminal law and punitive sanctions are inversely correlated to other forms of social control. In addition to law, social control is found in many intermediate social institutions, including family, churches, schools, occupations, neighborhoods, and friendships. In societies where such intermediate institutions are dominant, the need for law

and formal criminal sanctions is limited. Conversely, in societies where intermediate institutions are less dominant informal social control is less effective, thereby necessitating the expansion of formal social control that typically takes the form of criminal law and punitive sanctions. The quantity of criminal law increases as the quantity of informal social control decreases (Black 1976). The primary measurement of the quantity of criminal law is the frequency and severity of criminal sanctions.

The relative quantity of criminal law and other forms of social control have been linked to the relative presence or absence of the other aspects of social life. At the extremes, societies characterized by significant stratification, morphological diversity, bureaucratic interaction, and cultural multiplicity will employ a more penal style of social control while societies characterized by general equality, homogeneity, face-to-face interaction, and cultural consensus will employ a more conciliatory style of social control. The therapeutic, educative, and compensatory styles of social control are to be found in societies somewhere between the extremes. In Western societies the increase in stratification, morphology, bureaucracy, and cultural multiplicity have combined to lower the level of willing conformity and diminish the effectiveness of informal mechanisms of social control (Garland 1990). The decline of informal social control has been traced to the less efficient use of shaming. Literature on the effectiveness of various normative sanctions has suggested that shame can be either reintegrative or disintegrative in nature depending on the strength of intermediate institutions and the level of willing conformity (Braithwaite 1994). Although the phenomenon of shame is found in both formal and informal systems of social control it is a more central feature of informal social control.

#### FORMAL VERSUS INFORMAL SANCTIONS

*Shaming* is the central deterrent element of most informal mechanisms of social control. Its use in formal mechanisms of social control has until the 1990s been limited due to the concerns related to labeling (Becker 1963; Braithwaite 1994). Deterrence research has shown a much stronger shaming effect for informal sanctions than for formal legal

sanctions (Paternoster and Iovanni 1986). It has been suggested that sanctions imposed by groups with emotional and social ties to the transgressor are more effective in deterring criminal behavior than sanctions that are imposed by a bureaucratic legal authority (Christie 1977). Although the empirical evidence comparing the deterrent effect of informal versus formal sanctions is limited, it has consistently shown that there is a greater concern among transgressors for the social impact of their arrest than for the punishment they may receive (Sherman and Berk 1984). Perceptions of the certainty and severity of formal criminal sanctions appear to have little effect on the considerations that lead to criminal behavior. Whatever effect formal criminal sanctions do have on deterring future criminal behavior is primarily dependent on the transgressor's perception of informal sanctions (Tittle 1980).

*Deterrence theory* is predicated on the assumption of fear (Zimring and Hawkes 1973). It has been assumed that deviants and especially criminal offenders fear the loss directly related to formal sanctions (e.g. loss of liberty, loss of material possessions, etc.). Research on formal sanctions that are coupled with informal sanctions suggests to the contrary that to the extent transgressors would be deterred by fear, the fear that is most relevant is that their transgression will result in a loss of respect or status among their family, friends, or associates (Tittle 1980). In the cost-benefit analysis that is central to the "rational actor" model of crime causation, loss of respect and status weighs much more heavily for most individuals than the direct loss of liberty or material.

Deterrence is considered to be irrelevant to the majority of the populace, therefore, most people are believed to comply with the law because of their internalization of the norms and values of society (Toby 1964). For the majority of the populace failure to abide by the norms and values would call into question their commitment to that society. For the well-socialized individual, moral conscience serves as the primary deterrent. The conscience delivers an anxiety response each time an individual transgresses the moral boundaries of society. The anxiety response fulfills the three requirements of deterrence: it is immediate, it is certain, and it is severe. Formal criminal sanctions on the other hand may be severe but they are not always certain nor swift (Braithwaite 1994).

Criminal sanctions represent a denial of confidence in the morality of the transgressor. In a criminal law system norm compliance is reduced to the calculus of cost versus benefit. Conversely, informal sanctions can be a reaffirmation of the morality of the transgressor by showing disappointment in the transgressor for acting out of character. The shaming associated with the disappointment is reintegrative because its ultimate goal is the restoration of the transgressor to full membership in the group, community, or society. Shaming in the informal context is reintegrative because of its view of the transgressor as a whole and valued member of society and its goal of social restoration. Criminal sanctions, on the other hand, define the transgressor strictly in terms of his or her transgression thus causing a disconnection between the transgressor and society (Garfinkel 1965). Criminal sanctions lack both the incentive and mechanism for reintegration. Without reintegrative potential, criminal sanctions stigmatize rather than shame. Stigmatization is exclusively concerned with labeling while reintegrative shaming is equally concerned with delabeling. In the context of most informal sanctions community disapproval is coupled with gestures of reacceptance. The reintegrative character of informal sanctions ensures that the deviant label be directed to the behavior rather than the person thus avoiding the assignment of a master status to the person (Braithwaite 1994).

In the study of criminal sanctions much attention has been directed to the effects of labeling on the sanctioned transgressor. *Labeling theory* argues that the deviant characteristic (e.g., drug addict, criminal, etc.) assigned to a person may become a status such that the 'label' will dominate all positive characteristics (e.g. parent, teacher, church member, etc.) that might otherwise characterize the person. Such a deviant master status is believed to limit and in some cases preclude the reintegration of the transgressor. For those prevented from reintegrating, the label becomes a self-fulfilling prophecy eventually leading to secondary deviance, namely deviance caused by the label itself (Becker 1963). The possibility of secondary deviance as a result of the stigmatization of the transgressor has been included by some labeling theorists in their assessment of the relative effectiveness of criminal sanctions (Gove 1980). However, it has been alternatively suggested that

stigmatization might enhance crime control because the thought of being a social outcast serves as a stronger sanction and more effective deterrent than being shamed and eventually reintegrated (Silberman 1976). The potential effectiveness of sanctions is dependent upon whether deterrence is more effectively achieved through the fear of social marginalization associated with formal sanctions or the fear of community disapproval associated with informal sanctions (Braithwaite 1994).

### CRIMINAL SANCTIONS AND DUE PROCESS

Despite the lack of conclusive evidence in support of the deterrent effect of criminal sanctions, the frequency and severity of such sanctions have significantly increased over the past quarter-century (Stafford and Warr 1993). The dramatic rise in both the prison population and persons under probationary supervision is a reflection of the increase in the age of the population most at risk to engage in criminal acts as well as a renewed commitment to the nonconsequential philosophy of punishment. A number of observers have pointed to the growing use of criminal sanctions as a sign of an increased emphasis on crime control; however, the increasing use of criminal sanctions has coincided with the expansion of legal protections for the accused. The limitations of criminal sanctions in the form of due process guarantees have produced a hybrid criminal process that legitimates the widespread use of punitive responses to normative transgressions by constraining the power of state-sponsored agents of control. The criminal process in the West includes both a strong crime-control component and a strong due process component (Packer 1968). The process is perceived as a contest between relatively equal actors, which helps to legitimate the pain associated with the imposition of criminal sanctions. The criminal process satisfies both society's demand for crime control and its desire to protect the liberty of the individual. Criminal sanctions are shaped by these competing demands of society. On the one hand the crime-control component is based on the assumption that the suppression of criminal conduct is by far the most important function of the criminal process, while the due process component is based on the assumption that ensuring that

only the guilty are punished is the most important function of the criminal process.

The crime-control model suggests that the failure to provide an adequate level of security for the general public will undermine the legitimacy of criminal law. In contrast, the due process model suggests that the wrongful punishment of an individual is a greater threat to the legitimacy of law (Turk 1969). Therefore, criminal sanctions must be severe enough to contribute to crime control and their application must be fair and consistent in order to satisfy due process requirements (Packer 1968). The additional constraint on criminal sanctions that they be applied only in cases where genuine intent can be conclusively determined has contributed to the rise of alternative systems of social controls.

#### SOCIAL CONTROL BEYOND CRIMINAL SANCTIONS

As the definition of what constitutes criminal behavior is narrowed, the social control of noncriminal normative transgression is transferred from the jurisdiction of criminal law to the jurisdictions of medicine and public health. Both public health and psychiatry have long been concerned with social control (Foucault 1965), but what is more significant is the dramatic expansion of their influence and control over the past half-century. Much of what was considered "badness" (and in some cases criminal) and thereby worthy of punishment has been redefined as sickness that requires a therapeutic response. Though the response is quite different, the objective of social control is fundamentally the same (Conrad and Schneider 1992). Some forms of normative transgressions have long been within the medical and public health domain (e.g., mental illness). Research has shown, however, that an increasing number of behaviors that had been punished in the past are now being treated within the medical jurisdiction (e.g., suicide, alcoholism, drug addiction, child abuse, violence, etc.). The most important implication in the shift of the locus of social control from law to medicine is the determination of individual responsibility. In the context of criminal law the full responsibility for normative transgressions is vested in the individual. In the medical context transgressors are not considered responsible for their

actions (or their responsibility is significantly mitigated), therefore, they cannot be punished for actions that are beyond their control. The societal response to behaviors so defined is therapeutic rather than punitive (Conrad and Schneider 1992).

The trend toward the decriminalization of normative transgressions can be traced to the emergence of *deterministic criminology*, which shifted the study of crime causation from rational action to biological, psychological, and sociological sources (Kittrie 1971). The deterministic perspective was initially applied to juvenile delinquency. The influence of deterministic criminology grew in proportion to the expansion of the juvenile justice system. As more questions were raised about the ability of criminal law and criminal sanctions to effectively deal with transgressors such as juvenile delinquents, alternative methods of control became more widely accepted. It has been suggested that the shift from simple to complex societies has necessitated an accompanying shift from punitive to therapeutic methods of social control (Kittrie 1971). For example, the strength of criminal sanctions is believed to be declining because of the increase in residential mobility coupled with the decrease in the influence of primary institutions such as family, church, and school (Braithwaite 1994). As the influence of primary institutions has waned, their ability to command conformity has declined. With the decline of informal social control and the perceived ineffectiveness of criminal sanctions society has increasingly turned to the promise of therapeutic social control as a means of responding to normative transgressions (Conrad and Schneider 1992).

#### REFERENCES

- Beccaria, Cesare (1764) 1980 *On Crimes and Punishment*. Indianapolis, Ind.: Bobbs-Merrill.
- Becker, Howard 1973 *Outsiders*. New York: The Free Press.
- Bentham, Jeremy (1823) 1995 "Punishment and Utility." In Jeffrie Murphy, ed., *Punishment and Rehabilitation*. Belmont, Calif.: Wadsworth.
- Bernard, Thomas 1983 *The Consensus–Conflict Debate: Form and Content in Social Theories*. New York: Columbia University Press.
- Black, Donald 1976 *The Behavior of Law*. New York: Academic Press.

- Bottomley, A. 1989 *Crime and Penal Politics: The Criminologist's Dilemma*. Hull, Eng.: Hull University Press.
- Braithwaite, John 1994 *Crime, Shame and Reintegration*. Cambridge, Eng.: Cambridge University Press.
- Christie, Nils 1977 "Conflicts as Property." *British Journal of Criminology* 17:1-15.
- Conrad, Peter, and Joseph Schneider 1992 *Deviance and Medicalization: From Badness to Sickness*. Philadelphia: Temple University Press.
- Duff, Anthony, and David Garland 1994 "Introduction: Thinking About Punishment." In R.A. Duff and D. Garland, eds., *A Reader on Punishment*. New York: Oxford University Press.
- Durkheim, Emile (1893) 1964 *The Division of Labor in Society*. New York: Free Press.
- Fogel, David 1975 *We Are Living Proof: The Justice Model of Corrections*. Cincinnati, Oh.: W.H. Anderson.
- Foucault, Michael 1965 *Madness and Civilization*. New York: Random House.
- Garfinkel, Harold 1965 "Conditions of Successful Degradation Ceremonies." *American Journal of Sociology* 61:420-424.
- Garland, David 1990 *Punishment and Modern Society: A Study in Social Theory*. Chicago: University of Chicago Press.
- Goffman, Erving 1961 *Asylums*. Garden City, N.Y.: Anchor Press.
- Gove, William 1980 *The Labeling of Deviance: Evaluating a Perspective*. Beverly Hills, Calif.: Sage.
- Kittrie, Nicholas 1971 *The Right to be Different: Deviance and Enforced Therapy*. Baltimore, Md.: Johns Hopkins University Press.
- Marx, Karl 1964 "Theories of Surplus Value." In T.B. Bottomore and M. Rubel, eds., *Karl Marx: Selected Writings in Sociology and Social Philosophy*. New York: McGraw-Hill.
- Melossi, Dario, and M. Pavarini 1981 *Prison and the Factory: Origins of the Penitentiary System*. Totowa, N.J.: Barnes & Noble Books.
- Michalowski, Raymond 1985 *Order, Law and Crime*. New York: Random House.
- Morris, Herbert 1981 "A Paternalistic Theory of Punishment." *American Philosophical Quarterly* 18:263-271.
- Morris, Norval 1982 "'Dangerousness and Incapacitation' Record of the Association of the Bar in the City of New York" 34:102-128. In A. Duff and D. Garland, eds., *1994 Reader on Punishment*. New York: Oxford University Press.
- Packer, Herbert 1968 *The Limits of the Criminal Sanction*. Palo Alto, Calif.: Stanford University Press.
- Paternoster, Ray, and Leeann Iovanni 1986 "The Deterrent Threat of Perceived Severity: A Reexamination." *Social Forces* 64:751-777.
- Reiman, Jeffrey 1998 *The Rich Get Richer and the Poor Get Prison*. New York: Wiley.
- Rotman, Edgardo 1990 *Beyond Punishment: A New View of the Rehabilitation of Offenders*. Westport, Conn.: Greenwood Press.
- Rusche, George, and Otto Kirchheimer 1939 *Punishment and Social Structure*. New York: Columbia University Press.
- Sellen, Thorsten 1976 *Slavery and the Penal System*. New York: Elsevier.
- 1938 *Culture, Conflict and Crime*. New York: Social Science Research Council.
- Sherman, Lawrence, and Richard Berk 1984 "The Special Deterrent Effect of Arrest for Domestic Assault." *American Sociological Review* 49:261-272.
- Silberman, Charles 1978 *Criminal Violence, Criminal Justice*. New York: Vintage.
- Stafford Mark, and Mark Warr 1993 "A Reconceptualization of General and Specific Deterrence." *Journal of Research on Crime and Delinquency* 30:123-135.
- Sykes, Gresham 1958 *The Society of Captives*. Princeton, N.J.: Princeton University Press.
- Tittle, Charles 1980 *Sanctions and Social Deviance*. New York: Praeger.
- Toby, Jackson 1964 "Is Punishment Necessary?" *Journal of Criminal Law, Criminology and Political Science* 55:332-337.
- Turk, Austin 1969 *Criminality and Legal Order*. Chicago: Rand McNally.
- Vold, George 1958 *Theoretical Criminology*. New York: Oxford University Press.
- Walker, Nigel 1980 *Punishment, Danger and Stigma: The Morality of Criminal Justice*. Oxford, Eng.: Basil Blackwell.
- Wilson, James Q. 1979 *Thinking About Crime*. New York: Basic Books.
- Zimring, Franklin, and Gordon Hawkins 1990 *The Scale of Imprisonment*. Chicago: University of Chicago Press.
- 1973 *Deterrence: The Legal Threat in Crime Control*. Chicago: University of Chicago Press.

PETER CORDELLA

## CRIMINALIZATION OF DEVIANCE

A 1996 household survey dealing with drug abuse revealed that 13 percent of persons eighteen to twenty-five, 6 percent of persons twenty-six to thirty-four, and 2 percent of persons thirty-five years or older had smoked marijuana within the last thirty days (Maguire and Pastore 1998, p. 246). The same survey showed that nearly a quarter of persons eighteen to twenty-five had smoked marijuana within the past year and about half of persons eighteen to thirty-four had smoked marijuana at least once during their lifetimes. (The prevalence of marijuana use is considerably greater among males than among females, so these statistics understate marijuana use among young adult males.) A different population survey, also conducted in 1996, revealed that the legalization of marijuana use enjoys considerable support among young adults: 38 percent among respondents eighteen to twenty, 30 percent among respondents twenty-one to twenty-nine, and 28 percent among respondents thirty to forty-nine (Maguire and Pastore 1998, p. 151). Males are much more supportive of legalization than females, and those who claim no religion are most supportive of all.

A reasonable interpretation of these data is that no clear consensus exists that smoking marijuana is wrong. Rather, American society contains a large group, probably a majority, that considers smoking marijuana wrong and a smaller group (but a substantial proportion of young adult males) that uses marijuana, considers it harmless, and is probably indignant at societal interference. Since a nonconformist is immensely strengthened by having even one ally, as the social psychologist Solomon Asch demonstrated in laboratory experiments (Asch 1955), such sizable support for marijuana use means that controlling its use is no easy task. Many other kinds of behavior in contemporary societies resemble marijuana use in that some people disapprove of the behaviors strongly and others are tolerant or supportive. In short, modern societies, being large and heterogeneous, are likely to provide allies even for behavior that the majority condemns.

Yet moral ambiguity does not characterize American society on every issue. Consensus exists

that persons who force others to participate unwillingly in sexual relations and persons with body odor are reprehensible. Body odor and rape seem an incongruous combination. What they have in common is that both are strongly disapproved of by the overwhelming majority of Americans. Where they differ is that only one (rape) is a statutory crime that can result in police arrest, a court trial, and a prison sentence. The other, smelling bad, although deviant, is not criminalized. When deviance is criminalized, the organized collectivity channels the indignant response of individuals into public condemnation and, possibly, punishment.

Some sociologists maintain that when sufficient consensus exists about the wrongfulness of an act, the act gets criminalized. This is usually the case but not always. Despite consensus that failing to bathe for three months is reprehensible, body odor has not been criminalized. And acts *have* become criminalized, for example, patent infringement and other white-collar crimes that do not arouse much public indignation. In short, the correlation between what is deviant and what is criminal, though positive, is not perfect.

For a clue to an explanation of how and why deviance gets criminalized, note how difficult it is for members of a society to know with certainty what is deviant. Conceptually, deviance refers to the purposive evasion or defiance of a normative consensus. Defiant deviance is fairly obvious. If Joe, a high school student, is asked a question in class by his English teacher pertaining to the lesson, and he replies, "I won't tell you, asshole," most Americans would probably agree that he is violating the role expectations for high school students in this society. Evasive deviance is less confrontational, albeit more common than defiant deviance, and therefore more ambiguous. If Joe never does the assigned homework or frequently comes late without a good explanation, many Americans would agree that he is not doing what he is supposed to do, although the point at which he steps over the line into outright deviance is fuzzy. Both evasive and defiant deviance require other members of the society to make a judgment that indignation is the appropriate response to the behavior in question. Such a judgment is difficult to make in a heterogeneous society because members of the society cannot be sure how closely other people share their values.

To be sure, from living in a society the individual has a pretty good idea what sorts of behavior will trigger indignant reactions, but not with great confidence. The issue is blurred by subgroup variation and because norms change with the passage of time. Bathing suits that were entirely proper in 1999 in the United States would have been scandalous in 1929. A survey might find that a large percentage of the population disapproves now of nudity on a public beach. The survey cannot reveal for how long the population will continue to disapprove of public nudity. Nor can the survey help much with the crucial problem of deciding how large a proportion of the population that disapproves strongly of a behavior is enough to justify categorizing public nudity as “deviant.”

In short, the scientific observer can decide that a normative consensus has been violated only after first establishing that a consensus exists on that issue at this moment of time, and that is not easy. To determine this, individual normative judgments must somehow be aggregated, say, by conducting a survey that would enable a representative sample of the population to express reactions to various kinds of behavior. Ideally, these responses differ only by degree, but in a large society there are often qualitatively different conceptions of right and wrong in different subgroups.

In practice, then, in modern societies neither the potential perpetrator nor the onlooker can be certain what is deviant. Consequently the social response to an act that is on the borderline between deviance and acceptability is unpredictable. This unpredictability may tempt the individual to engage in behavior he would not engage in if he knew that the response would be widespread disapproval (Toby 1998a). It may also restrain onlookers from taking action against the behavior—or at least expressing disapproval—against persons violating the informal rules.

#### WHAT CRIME INVOLVES: A COLLECTIVE RESPONSE

Crime is clearer. The ordinary citizen may not know precisely which acts are illegal in a particular jurisdiction. But a definite answer is possible. A lawyer familiar with the criminal code of the State of New Jersey can explain exactly what has to be

proved in order to convict a person of drunken driving in New Jersey. The codification of an act as criminal does not depend on its intrinsic danger to the society but on what societal leaders *perceive* as dangerous. For example, Cuba has the following provision in its criminal code:

*Article 108. (1) There will be a sanction of deprivation of freedom of from one to eight years imposed on anyone who: (a) incites against the social order, international solidarity or the socialist State by means of oral or written propaganda, or in any other form; (b) makes, distributes or possesses propaganda of the character mentioned in the preceding clause. (2) Anyone who spreads false news or malicious predictions liable to cause alarm or discontent in the population, or public disorder, is subject to a sanction of from three to four years imprisonment. (3) If the mass media are used for the execution of the actions described in the previous paragraphs, the sanction will be a deprivation of freedom from seven to fifteen years (Ripoll 1985, p. 20).*

In other words, mere possession of a mimeograph machine in Cuba is a very serious crime because Fidel Castro considers the dissemination of critical ideas a threat to his “socialist State,” and in Cuba Castro’s opinions are literally law. Hence, possession of a mimeograph machine is a punishable offense. Members of Jehovah’s Witnesses who used mimeograph machines to reproduce religious tracts have been given long prison sentences. On the other hand, reproducing religious tracts may not arouse indignation in the Cuban population. It is criminal but not necessarily deviant.

In California or New Jersey, as in Cuba, a crime is behavior punishable by the state. But the difference is that in the fifty states, as in all democratic societies, the legislators and judges who enact and interpret criminal laws do not simply codify their own moral sentiments; they criminalize behavior in response to influences brought to bear on them by members of their constituencies. True, women, children, members of ethnic and racial minorities, and the poorly educated may not have as much political input as affluent, middle-aged, white male professionals. But less influence does not mean they don’t count at all. In a dictatorship, on the other hand, the political process is closed;



few people count when it comes to deciding what is a crime.

### WHEN DEVIANCE IS CRIMINALIZED: POLITICIZATION

As was mentioned earlier, what is deviant is intrinsically ambiguous in a complex society whose norms are changing and whose ethnic mix has varied values. Criminalization solves the problem of predictability of response by transferring the obligation to respond to deviance from the individual members of society to agents of the state (the police). But criminalization means that some members of society are better able than others to persuade the state to enforce *their* moral sentiments. Criminalization implies the politicization of the social control of deviance. In every society, a political process occurs in the course of which deviant acts get criminalized. Generally, the political leadership of a society criminalizes an act when it becomes persuaded that without criminalization the deviant "contagion" will spread, thereby undermining social order (Toby 1996). The leaders may be wrong. Fidel Castro might be able to retain control of Cuban society even if Cubans were allowed access to mimeograph machines and word-processors. Nevertheless, leaders decide on crimes based on their perception of what is a threat to the collectivity. According to legal scholars (Packer 1968), the tendency in politically organized societies is to overcriminalize, that is, to involve the state excessively in the response to deviance. Political authorities, even in democracies, find it difficult to resist the temptation to perceive threats in what may only be harmless diversity and to attempt to stamp it out by state punishment.

Sociology's labeling perspective on deviance (Becker 1963; Lemert 1983) goes further; it suggests that overcriminalization may *increase* deviance by changing the self-concept of the stigmatized individual. Pinning the official label of "criminal" on someone stigmatizes him and thereby amplifies his criminal tendencies. Furthermore, an advantage of ignoring the deviance is that it may be ephemeral and will disappear on its own; thus in 1974 American society virtually ignored "streaking" instead of imprisoning streakers in large numbers for indecent exposure (Toby 1980),

and by 1975 streaking had become a historical curiosity. But whether deviance is self-limiting is an empirical question. The labeling perspective ignores the logical possibility that stigmatizing the deviant may be necessary to deter future deviance by bringing home to the offender (as well as potential offenders) the danger of antagonizing the community. In point of fact, the empirical evidence supports the deterrence possibility more than it does the amplification assumption (Gove 1980; Gibbs 1975). At its most extreme, the labeling perspective denies the desirability of any kind of criminalization:

*"The task [of radical reform] is to create a society in which the facts of human diversity, whether personal, organic, or social, are not subject to the power to criminalize"* (Taylor, Walton, and Young 1973, p. 282).

Thus, the labeling perspective flirts with philosophical anarchism. More reasonably, the issue is: which forms of deviance can be regarded as harmless diversity and which threaten societal cohesiveness sufficiently that they require criminalization in order to be contained within tolerable limits? Experience has taught us that the body odor of other people, though objectionable to most Americans, is tolerable. But what about consuming alcohol or cocaine to the point of chronic intoxication? What about sexual practices that shock most people such as sado-masochism or intercourse with a sheep? One way to finesse these thorny questions is to define such acts as the product of mental illness and therefore beyond the control of the individual. Instead of regarding drug abuse or alcoholism or bestiality as deviant choices in the face of temptations, society may choose to regard them as "addictions," that is to say, involuntary (Toby 1998b). Since the ill person is by definition unenviable, he is not a role model, and therefore the deviant contagion does not spread (Toby 1964). But suppose consensus does not exist that these acts are compulsive; suppose that many people feel that the perpetrators are perversely *choosing* to engage in these behaviors. Average citizens may become demoralized when they see their norms flouted or they may be tempted to engage in the deviant behavior themselves. This is why criminalization (and state-sponsored punishment) may be necessary. Punishment serves to deprive

the deviant of the benefits of his nonconformity, and therefore he becomes unenviable in the eyes of conformists.

Yes, society may stigmatize and perhaps imprison perpetrators, amid hope that imitators will be rare. But criminalization arouses opposition. Libertarians lean toward permitting almost any nonviolent behavior except the exploitation of children. Mental-health advocates perceive stealing to feed a passion for gambling as a symptom of illness; they may perceive even predatory violence as symptomatic of a sick personality for which the individual cannot be considered responsible. Pragmatists point out that when large numbers of people want to do something, such as gamble or use drugs, it is not practical to attempt to stop them by criminalizing the behavior. They argue that deviants cannot all be punished, certainly not by imprisonment; they corrupt police forces through bribes and deflect police efforts into tasks that cannot be accomplished instead of more feasible deviance-control activities; and the criminal organizations that emerge to cater to these forbidden pleasures promote crimes that would not have occurred in the absence of criminalization.

On the other side of the ledger is the tendency for the absence of criminalization to encourage individuals to engage in a behavior they would not have engaged in when faced with possible criminal sanctions. The Prohibition experiment of the 1920s, despite its perception as a failure, did succeed in reducing the incidence of alcoholism, as reflected in reduced incidence of alcoholic psychosis and of cirrhosis of the liver. But the social cost of criminalizing alcohol consumption was not only the proliferation of criminal enterprises to supply the demand of social drinkers; criminalization also prevented many people who wished to be social drinkers from having the freedom to do so. True, some of these would go on to become alcoholics, but most would not. Thus, the criminalization issue always involves a tradeoff between partially legitimating possibly harmful behavior, such as smoking cigarettes, or curtailing freedom by criminalizing the behavior. This judgment has to be made on a case by case basis. Most people who drink socially do not become alcoholics and most people who smoke do not get lung cancer; hence it is difficult to justify criminalizing drinking and

smoking despite the likelihood that more people become alcoholics and get lung cancer than would if smoking and drinking were criminalized. On the other hand, the tradeoff goes the other way with "hard" drugs; the main argument against decriminalizing the sale of heroin is that the health costs to the general population would be too great; decriminalization would inevitably increase experimentation with the drug and ultimately the number of addicts (Kaplan 1983).

### OTHER CONSEQUENCES OF CRIMINALIZATION

The absence of criminal law—and consequently of state-imposed sanctions for violations—is no threat to small primitive communities: Informal social controls can be counted on to prevent most deviance and to punish what deviance cannot be prevented. In heterogeneous modern societies, however, the lack of some criminalization would make moral unity difficult to achieve. When Emile Durkheim spoke of the collective conscience of a society, he was writing metaphorically; he knew that he was abstracting from the differing consciences of thousands of individuals. Nevertheless, the criminal law serves to resolve these differences and achieve a contrived—and indeed precarious—moral unity. In democratic societies, the unity is achieved by political compromise. In authoritarian or totalitarian societies the power wielders unify the society by imposing their own values on the population at large. In both cases law is a unifying force; large societies could not function without a legal system because universalistic rules, including the rules of the criminal law, meld in this way ethnic, regional, and class versions of what is deviant (Parsons 1977, pp. 138–139).

The unifying effect of the criminal law has unintended consequences. One major consequence is the development of a large bureaucracy devoted to enforcing criminal laws: police, judges, prosecutors, jailers, probation officers, parole officers, prison guards, and assorted professionals like psychologists and social workers who attempt to rehabilitate convicted offenders. Ideally, these employees of the state should perform their roles dispassionately, not favoring some accused persons or discriminating against others. In practice, however, members of the criminal justice bureaucracy bring to their jobs the parochial sentiments

of their social groups as well as a personal interest in financial gain or professional advancement. This helps to explain why police are often more enthusiastic about enforcing some criminal laws than they are about enforcing others.

Another consequence of criminalization is that the criminal law, being universal in its reach, cannot make allowances for subgroup variation in sentiments about what is right and what is wrong. Thus, some people are imprisoned for behavior that neither they nor members of their social group regard as reprehensible, as in Northern Ireland where members of the Irish Republican Army convicted of assassinating British soldiers considered themselves political prisoners. They went on hunger strikes—in some cases to the point of death—rather than wear the prison uniform of ordinary criminals.

### CONCLUSION

The more heterogeneous the culture and the more swiftly its norms are changing, the less consensus about right and wrong exists within the society. In the United States, moral values differ to some extent in various regions, occupations, religions, social classes, and ethnic groups. This sociocultural value pluralism means that it is difficult to identify behavior that everyone considers deviant. It is much easier to identify crime, which is codified in politically organized societies. The criminalization of deviance makes it clear when collective reprisals will be taken against those who violate rules.

Deviance exists in smaller social systems, too: in families, universities, and corporations. In addition to being subjected to the informal disapproval of other members of these collectivities, the deviant in the family, the university, or the work organization can be subjected to formally organized sanctioning procedures like a disciplinary hearing at a university. However, the worst sanction that these nonsocietal social systems can visit upon deviants is expulsion. A university cannot imprison a student who cheats on a final exam. Even in the larger society, however, not all deviance is criminalized, sometimes for cultural reasons as in the American refusal to criminalize the expression of political dissent, but also for pragmatic reasons as in the American failure to criminalize body odor, lying to one's friends, or smoking in church.

### REFERENCES

- Asch, Solomon E. 1955 "Opinions and Social Pressure." *Scientific American* 193, November.
- Becker, Howard S. 1963 *Outsiders. Studies in the Sociology of Deviance*. New York: Free Press.
- Gibbs, Jack P. 1975 *Crime, Punishment, and Deterrence*. New York: Elsevier.
- Gove, Walter R. 1980 *The Labelling of Deviance: Evaluation of a Perspective*, 2nd ed. Beverly Hills, Calif.: Sage.
- Kaplan, John 1983 *The Hardest Drug: Heroin and Public Policy*. Chicago: University of Chicago Press.
- Lemert, Edwin M. 1983 "Deviance." In Sanford H. Kadish, ed., *Encyclopedia of Crime and Justice*. New York: Free Press.
- Maguire, Kathleen, and Ann L. Pastore (eds.) *Sourcebook of Criminal Justice Statistics 1997*. Washington, D.C.: U.S. Government Printing Office. U. S. Department of Justice, Bureau of Justice Statistics.
- Packer, Herbert L. 1968 *The Limits of the Criminal Sanction*. Stanford, Calif.: Stanford University Press.
- Parsons, Talcott 1977 *The Evolution of Societies*. Englewood Cliffs, N.J.: Prentice-Hall.
- Ripoll, Carlos 1985 *Harnessing the Intellectual: Censoring Writers and Artists in Today's Cuba*. Washington, D.C.: Cuban American National Foundation.
- Taylor, Ian, Paul Walton, and Jock Young 1973 *The New Criminology*. London: Routledge and Kegan Paul.
- Toby, Jackson 1964 "Is Punishment Necessary?" *Journal of Criminal Law, Criminology, and Police Science*, 55, September.
- 1980 "Where are the Streakers Now?" In Hubert M. Blalock, ed., *Sociological Theory and Research*, pp. 304–313. New York: Free Press.
- 1996 "Reducing Crime: New York's Example," *Washington Post*, July 23.
- 1998a "Getting Serious about School Discipline," *The Public Interest*, 133, Fall.
- 1998b "Medicalizing Temptation." *The Public Interest*, 130, Fall.

JACKSON TOBY

### CRIMINOLOGY

The roots of modern criminology can be found in the writings of social philosophers, who addressed Hobbes's question: "How is society possible?" Locke and Rousseau believed that humans are endowed with free will and are self-interested. If

this is so, the very existence of society is problematic. If we are all free to maximize our own self-interest we cannot live together. Those who want more and are powerful can simply take from the less powerful. The question then, as now, focuses on how is it possible for us to live together. Criminologists are concerned with discovering answers to this basic question.

Locke and Rousseau, philosophers who are not considered criminologists, argued that society is possible because we all enter into a “social contract” in which we choose to give up some of our freedom to act in our own self-interest for the privilege of living in society. What happens though to those who do not make, or choose to break, this covenant? Societies enforce the contract by punishing those who violate it. Early societies punished violations of the social contract by removing the privilege of living in society through banishment or death. In the event of minor violations, sanctions such as ostracism or limited participation in the community for a time were administered. The history of sanctions clearly demonstrates the extreme and frequently arbitrary and capricious nature of sanctions (Foucault 1979).

The Classical School of criminology (Beccaria 1764; Bentham 1765) began as an attempt to bring order and reasonableness to the enforcement of the social contract. Beccaria in *On Crimes and Punishments* (1768) made an appeal for a system of “justice” that would define the appropriate amount of punishment for a violation as just that much that was needed to counter the pleasure and benefit from the wrong. In contemporary terms, this would shift the balance in a cost/benefit calculation, and would perhaps deter some crime. Bentham’s writings (1765) provided the philosophical foundation for the penitentiary movement that introduced a new and divisible form of sanction: incarceration. With the capacity to finally decide which punishment fits which crime, classical school criminologists believed that deterrence could be maximized and the cost to societal legitimacy of harsh, capricious, and excessive punishment could be avoided. In their tracts calling for reforms in how society sanctions rule-violators, we see the earliest attempts to explain two focal questions of criminology: Why do people commit crimes? How do societies try to control crime? The “classical school” of criminology’s answer to the first question is that individuals act rationally, and when the

benefits to violating the laws outweigh the cost then they are likely to choose to violate those laws. Their answer to the second question is deterrence. The use of sanctions was meant to discourage criminals from committing future crimes and at the same time send the message to noncriminals that crime does not pay. Beccaria and Bentham believed that a “just desserts” model of criminal justice would fix specific punishments for specific crimes.

In the mid-nineteenth century the early “scientific study” of human behavior turned to the question of why some people violate the law. The positivists, those who believed that the scientific means was the preeminent method of answering this and other questions, also believed that human behavior was not a product of choice nor individual free will. Instead they argued that human behavior was “determined behavior,” that is, the product of forces simply not in the control of the individual. The earliest positivistic criminologists believed that much crime could be traced to biological sources. Gall (Leek 1970), referred to by some as the “father of the bumps and grunts school of criminology,” studied convicts and concluded that observable physical features, such as cranial deformities and protuberances, could be used to identify “born criminals.” Lombroso (1876) and his students, Ferri and Garofalo, also embraced the notion that some were born with criminal constitutions, but they also advanced the idea that social forces were an additional source of criminal causation. These early positivists were critics of the Classical School. They did not go so far as to argue that punishment should not be used to respond to crime, but they did advance the notion that punishment was insufficient to prevent crime. Simply raising the cost of crime will not prevent violations if individuals are not freely choosing their behavior. The early positivists believed that effective crime control would have to confront the root causes of violations, be they biological or social in nature.

Around 1900, Ferri gave a series of lectures critiquing social control policies derived from classical and neo-classical theory. What is most remarkable about those lectures is that, considered from the vantage point of scholars at the end of the twentieth century, the arguments then were little different from public debates today about what are the most effective means of controlling crime.

Then, as now, the main alternatives were “get tough” deterrence strategies that assumed that potential criminals could be frightened into compliance with the law, versus strategies that would reduce the number of offenses by addressing the root causes of crime. We know far more about crime and criminals today than criminologists of the late nineteenth and early twentieth century knew, yet we continue the same debate, little changed from the one in which Ferri participated in.

The debates today pit those espousing *rational choice theories* of crime (control and deterrence theories being the most popular versions) against what still might best be called *positivistic theories*. To be sure, contemporary positivistic criminology is considerably different from the theories of Gall and Lombroso. Modern criminologists do not explain law-violating behavior using the shapes of heads and body forms. Yet there are still those who argue that biological traits can explain criminal behavior (Wilson and Herrenstein 1985; Mednick 1977), and still others who focus on psychological characteristics. But most modern criminologists are sociologists who focus on how social structures and culture explain criminal behavior. What all of these modern positivists have in common with their predecessors Gall, Lombroso, and company, is that they share a belief that human behavior, including crime, is not simply a consequence of individual choices. Behavior, they argue, is “determined” at least in part by biological, psychological, or social forces. The goal of modern positivistic criminologists is to unravel the combination of forces that make some people more likely than others to commit crimes.

Today the research of sociological criminologists focuses on three questions: What is the nature of crime? How do we explain crime? What are the effects of societies’ attempts to control crime? Approaches to answering these questions vary greatly, as do the answers offered by criminologists. For example the first question, what is the nature of crime, can be answered by detailing the characteristics of people who commit crimes. Alternatively, one can challenge the very definition of what crime, and consequently criminals, are. In an attempt to answer this question, some criminologists focus on how much crime there is. But of course, even this is a difficult question to answer because there are many ways to count crime, with each type

offering different and sometimes seemingly conflicting answers.

### WHAT IS THE NATURE OF CRIME?

Simply defining crime can be problematic. We can easily define crime legally: Crime is a violation of the criminal law. The simplicity of this definition is its virtue, but also its weakness. On the positive side, a legalistic definition clearly demarcates what will be counted as crime—those actions defined by the state as a violation. However, it is not as clear as to whom will be defined as criminals. Do we count as criminals those who violate the law and are not arrested? What about those who are arrested and not convicted? Some criminologists would argue that even an act that may appear to be criminal cannot be called “crime” until a response or evaluation has been made of that act. For example, how can we know if an offensive act is a crime if there has been no evaluation of the intent of the perpetrator? In Anglo-American law, without criminal intent, an offensive action is not considered a criminal action. Other critics of a legalistic definition of crime argue that it is an overly limited conception that too narrowly truncates criminological inquiry; a legalistic definition of crime accepts the state’s definition of “legal” and equates that with “legitimate.” Critical criminologists (Quinney 1974; Chambliss 1975; Taylor, Walton, and Young 1973) argue that we should ask questions about the creation of law such as whose interests are being served by classifying a particular behavior as “illegal.” Questions such as these tend to be ignored if we simply accept a legalistic definition. Indeed, the particular definition used influences the kind of questions criminologists ask. When a legalistic definition is used, criminologists tend to ask questions such as: “How much crime occurs?” “Who commits crime?” “Why are some people criminal?” If a broader conception of crime is the focus (i.e., one that addresses the rule-makers as well as the rule-breakers), then one might ask these same questions, but add others: “Where does the law come from?” “Why are some offensive acts considered crimes while others are not?” “Whose interests do the laws serve?” This is not a debate that will ever be resolved. Students of criminology should understand however, that the definition of crime they employ will have important implications for the kinds of questions they will ask, the

kinds of data they will use to study crime, and the kinds of explanations and theories they will apply to understand crime and criminal behavior.

### HOW IS CRIME STUDIED?

Having made choices about definitions, criminologists are then faced with an array of data and methodologies that can be brought to bear on criminological questions. The data and methodological approach used should be dictated by the definition of crime and the research question being asked. Some very interesting research problems require analysis of quantitative data while others require that the researcher use qualitative approaches to study crime or a criminal justice process. For example, if one is trying to describe the socio-demographic characteristics of criminals then one of several means of counting crimes and people who engage in them might be used. On the other hand, in his book *On The Take* (1978), because Chambliss was interested in the source and nature of organized crime, it was clearly more appropriate for him to spend time in the field observing and interviewing, rather than simply counting.

**Methods of Criminological Research.** For the most part, criminologists use the same types of research methods as do other sociologists. But a unique quality of crime is its “hiddenness.” The character of crime means that those who do it hide it. As a result, the criminologist must be a bit of a detective even while engaged in social science research. To do this criminologists use observational studies such as those conducted by Chambliss (1978) in studying organized crime, or Fisse and Braithwaite (1987) in studies of white-collar crime, or Sanchez-Jankowski (1992) in his studies of gang crime.

Those who use quantitative methods frequently use data generated by the criminal justice system, victimization surveys, or self-reports. All of these data collection procedures have strengths and weaknesses, and they are best used by criminologists who have an appreciation of both. The most widely used criminal justice data are produced by police departments and published by the Federal Bureau of Investigation each year in their *Uniform Crime Reports* (UCR). The UCR contains counts of the crimes reported to police, arrest data (including some characteristics of those arrested), police

manpower statistics, and other data potentially of interest to researchers. Victimization surveys may be conducted by individuals or teams of researchers. The National Crime Victimization Survey (NCVS) is conducted annually by the federal government and, as a consequence, is widely used. The virtue of victimization surveys is that they include criminal acts that are never reported to the police. As their name indicates, self-report studies simply ask a sample of people about their involvement in criminal activity. If done correctly, it is quite surprising what people will report to researchers.

**Characteristics of criminals.** It is always risky to speak of the characteristics of criminals because of the problems with crime data mentioned above. Also, one must take care to specify the types of crimes included in any description of those characteristics correlated with criminal involvement. For example, those engaging in white-collar crime have very different characteristics than perpetrators of what might be called “common crimes” or “street crime.” In order to engage in white-collar crime, one has to be old enough to have a white-collar job, and possess those characteristics (such as education) requisite for those jobs. Without these same requirements, the perpetrators of street crime can reasonably be expected to look different from white-collar criminals.

Criminologists frequently speak of “the big four correlates of crime”: age, sex, race, and social class. The first two are rather uncomplicated. Crime is, for the most part, a young person’s activity (Hirschi and Gottfredson 1983). Probable involvement escalates in the teen years reaching a peak at between ages 15 and 17 and then drops. Most people stop participating in criminal activity by their mid to late twenties, even if they have not been arrested, punished, or rehabilitated. The correlation between sex and crime is also quite straightforward. Males are more likely to engage in crime than females. Criminologists have not found a society where this pattern does not hold.

The correlation between race, or ethnicity, and crime is complex. Most Americans when asked state that they believe that minorities commit more crimes than whites. This oversimplification is not only inaccurate, but it obscures important patterns. First, some nonwhites in the United States are from groups with lower crime rates than whites

(e.g. Chinese and Japanese Americans), but even this pattern is more complex. Chinese Americans whose families have been in the United States for generations seem to have lower crime rates than white Americans, but more recent Chinese immigrants have higher rates than white Americans. The category "Asian Americans" is too diverse to make generalizations about the larger group's perpetration of crime. The same is true of Latinos and Latinas. African Americans have disproportionately high crime rates, but Africans and people of Caribbean island descent have lower crime rates than black Americans. To simply describe the correlations between broad racial categories and crime misses an important sociological point: The criminality associated with each group appears to be more a product of their experience in America than simply their membership in that racial/ethnic category. This point is buttressed by patterns of race, ethnicity, and crime in other countries. Visible minorities and immigrants are more likely to have higher crime rates and to be more frequently arrested in countries where they are subordinated (see Tonry 1997 for a collection of studies of race, ethnicity, crime, and criminal justice in several countries).

For decades, in fact probably for centuries, researchers assumed that people from the lower classes committed more crime than those of higher status. In fact, most sociological theories used to explain common crime are based on this assumption. In the mid-1970s several criminologists argued that there really is not a substantial correlation between social class and crime (Tittle, Villemez, and Smith 1978). Many criminologists today believe that if we are simply considering the likelihood of breaking the law then there probably is not much difference by social class. But if the criminal domain being studied is serious violent offenses, then there probably is a negative association between social class and crime. Still, the correlation between social class and violent crime is not as strong as most people would assume (Hindelang, Hirschi, and Weis 1981).

**Where crime occurs.** A very sociological way of describing crime is to examine the crime patterns that exist for different types of areas. The social "ecological" literature that does this usually focuses on states, metropolitan areas, cities, and

neighborhoods. When this occurs, we find that those areas with relatively large numbers of residents who are poor, African American, immigrants, young, and living in crowded conditions have higher crime rates. One must be very careful when interpreting these patterns. Because an area with relatively large numbers of people with these characteristics has a high crime rate, we cannot conclude that they are necessarily the people committing the crimes. For example, high poverty areas have high crime rates. The high crime rates found in poor neighborhoods however, could be produced by their victimization by the nonpoor. The interesting thing about these correlations, as well as patterns of individual crime correlations, are not the observed patterns themselves but rather the questions that arise from these observations.

**Victims of crime.** One of the more interesting things that we have learned from victimization studies is how similar the victims of crimes are to the offenders. Victims of crime tend to be young, male, and minority group members. In fact, the prototypical victim of violent crime in America is a young, African-American male. The poor and those with limited education are disproportionately the victims of crime as well. These patterns are obviously inconsistent with popular images of crime and victimization, but they are quite predictable when we examine what we know about crime.

### CRIMINOLOGY THEORIES

Three theoretical traditions in sociology dominated the study of crime from the early and mid-twentieth century. Contemporary versions of these theories continue to be used today. The "Chicago School" tradition, or *social disorganization theory*, was the earliest of the three. The other two are *differential association strain theories* or *anomie theory* (including subculture theories).

Sociologists working in the Chicago School tradition have focused on how rapid or dramatic social change causes increases in crime. Just as Durkheim, Marx, Toennies, and other European sociologists thought that the rapid changes produced by industrialization and urbanization produced crime and disorder, so too did the Chicago School theorists. The location of the University of Chicago provided an excellent opportunity for

Park, Burgess, and McKenzie to study the social ecology of the city. Shaw and McKay found (1931) that areas of the city characterized by high levels of social disorganization had higher rates of crime and delinquency.

In the 1920s and 1930s Chicago, like many American cities, experienced considerable immigration. Rapid population growth is a disorganizing influence, but growth resulting from in-migration of very different people is particularly disruptive. Chicago's in-migrants were both native-born whites and blacks from rural areas and small towns, and foreign immigrants. The heavy industry of cities like Chicago, Detroit, and Pittsburgh drew those seeking opportunities and new lives. Farmers and villagers from America's hinterland, like their European cousins of whom Durkheim wrote, moved in large numbers into cities. At the start of the twentieth century Americans were predominately a rural population, but by the century's mid-point most lived in urban areas. The social lives of these migrants, as well as those already living in the cities they moved to, were disrupted by the differences between urban and rural life. According to social disorganization theory, until the social ecology of the "new place" can adapt, this rapid change is a criminogenic influence. But most rural migrants, and even many of the foreign immigrants to the city, looked like and eventually spoke the same language as the natives of the cities into which they moved. These similarities allowed for more rapid social integration for these migrants than was the case for African Americans and most foreign immigrants.

In these same decades America experienced what has been called "the great migration": the massive movement of African Americans out of the rural South and into northern (and some southern) cities. The scale of this migration is one of the most dramatic in human history. These migrants, unlike their white counterparts, were not integrated into the cities they now called home. In fact, most American cities at the end of the twentieth century were characterized by high levels of racial residential segregation (Massey and Denton 1993). Failure to integrate these migrants, coupled with other forces of social disorganization such as crowding, poverty, and illness, caused crime rates to climb in the cities, particularly in the

segregated wards and neighborhoods where the migrants were forced to live.

Foreign immigrants during this period did not look as dramatically different from the rest of the population as blacks did, but the migrants from eastern and southern Europe who came to American cities did not speak English, and were frequently Catholic, while the native born were mostly Protestant. The combination of rapid population growth with the diversity of those moving into the cities created what the Chicago School sociologists called social disorganization. More specifically, the disorganized areas and neighborhoods where the unintegrated migrants lived were unable to exercise the social control that characterized organized, integrated communities. Here crime could flourish. Crime was not a consequence of who happened to live in a particular neighborhood, but rather of the character of the *social ecology* in which they lived. That is, the crime rate was a function of the area itself and not of the people who lived there. When members of an immigrant or ethnic group moved out of that area (usually in succeeding generations), that group's crime rate would go down. But, the old neighborhood, with its cheaper housing and disorganized conditions, would attract another, more recent group migrating to the city. Those groups settled there because that is where they could afford to live. When they became more integrated in American urban life, like those who came before them, they would move to better neighborhoods and as a result have lower crime rates. Chicago School sociologists called the process of one ethnic group moving out to be replaced by a newly arriving group (with the first group passing on to newcomers both the neighborhood and its high crime) "ethnic succession." This pattern seems to have worked for most ethnic groups except African Americans. For African Americans, residential segregation and racial discrimination prohibited the move to better, more organized neighborhoods.

Contemporary social disorganization theorists (Sampson and Groves 1989) are less concerned with the effects of ethnic migration. The disorganizing effects of urban poverty (Sampson and Wilson 1994), racial residential segregation (Massey and Denton 1993), and the social isolation of the urban poor (Wilson 1987), and the consequent effect on



crime is the focus of current research. This line of research has developed intriguing answers to the question of why some racial or ethnic groups have higher crime rates.

Both the anomie and differential association traditions grew out of critiques of the Chicago School version of social disorganization theory. The latter was developed by Sutherland (1924), himself a member of the University of Chicago faculty. Sutherland believed that a theory of crime should explain not only how bad behavior is produced in bad living conditions, but also how bad behavior arises from good living circumstances. This theory should also explain why most residents of disorganized neighborhoods do not become criminals or delinquents. Sutherland explained this by arguing that crime is more likely to occur when a person has a greater number of deviant associations, relative to non-deviant associations.

Differential association occurs when a person has internalized an excess of definitions favorable to violations of the law. Sutherland believed that we all experience a variety of forms of exposure to definitions favorable to violation of the law—"It is OK to steal this because the insurance company will pay for it"—and definitions unfavorable to violation of the law—"It is not OK to steal from stores, because all of us are hurt when prices go up to pay the stores' higher insurance premiums." Sutherland was very careful to point out what differential association was not. It is not a simple count of favorable and unfavorable definitions. Differential association theory is not a theory that focuses on who a person associates with. Indeed Sutherland argued that it is possible to receive definitions favorable to violation from the law-abiding. Of course those spending time with delinquent peers will be exposed to more criminal definitions, but the theory should not be reduced to simply a peer group theory of crime and delinquency (Sutherland and Cressey 1974). As Cressey has pointed out, if crime were simply a consequence of prolonged proximity with criminals, then prison guards would be the most criminal group in the population. Differential association theory continues to be of influence in contemporary sociology, but it does not occupy as central a role as it did in the 1940s and 1950s. Matsueda

(1988), the theory's major contemporary proponent, has emphasized the interactive quality (similar to labeling theory) of differential association theory. Matsueda has suggested new ways to operationalize the key concepts of differential association, and the theory is "evolving" in his writings to bring in aspects of both *interactionist* and *rational choice theory*.

Anomie theory's roots are in the work of Durkheim, who used the concept *anomie* to describe the disruption of regulating norms resulting from rapid social change. Strain theories, including anomie theory, focus on social structural strains, inequalities, and dislocations, which cause crime and delinquency. Durkheim stressed that societal norms that restrained the aspirations of individuals were important for preventing deviance. However, unrealistically high aspirations would be frustrated by a harsh social reality, leading to adaptations such as suicide, crime, and addiction. Merton adapted Durkheim's notion of anomie by combining this idea with the observation that not only do societal norms affect the likelihood of achieving aspirations, but they also determine to a large degree what we aspire to (1938). In other words, society helps determine the goals that we internalize by defining which of them are legitimate but it also defines the legitimate means of achieving these goals. American society, for example, defines material comfort as legitimate goals, and taking a well-paid job as a legitimate means of achieving them. Yet legitimate means such as these and economic success are not universally available. When a society is characterized by a disjunction between its legitimate goals and the legitimate means available to achieve them, crime is more likely.

This conceptualization led a number of criminologists to focus on these "opportunity structure" strains (Cloward and Ohlin 1960), as well as cultural strains (Cohen 1955), to explain why crime would be higher in lower social class neighborhoods. Contemporary strain theorists have moved in two directions. Some have proceeded in ways that are quite consistent with Merton's original conceptualization, while others have set forward a more social psychological conception of strain (Agnew 1992). What ties these contemporary versions of anomie theory to the earlier tradition is the notion that there are individual adaptations to

social strain, or structurally based unequal opportunity, and that some of these adaptations can result in crime.

Subcultural explanations of crime are similar to both differential association and anomie theories. Like differential association, *subculture theories* have an important learning component. Both types of theories emphasize that crime, the behavior, accompanying attitudes, justifications, etc., are learned by individuals within the context of the social environment in which they live. And, as anomie and other versions of strain theory emphasize, subcultural explanations of crime tend to focus on lower-class life as a generating milieu in which procriminal norms and values thrive. Cohen's book *Delinquent Boys* (1955) describes delinquency as a product of class-based social strains, which lead to a gang subculture conducive to delinquency. Miller (1958) argued that a prodelinquency value system springs from situations where young boys grow up in poor, female-headed households. He felt that adherence to these values, which he called "the focal concerns of the lower class," made it more likely that boys would join gangs and involve themselves in delinquency. This idea enjoys recurring popularity in explanations of behavior in poor, and especially urban, minority neighborhoods (Banfield 1968; Murray 1984) even though it is notoriously difficult to test empirically.

A different version of subculture theory has been championed by Wolfgang (Wolfgang and Ferracuti 1967). Wolfgang and his colleagues argued that people in some segments of communities internalized, carried, and intergenerationally transferred values that were proviolence. Accordingly, members of this subculture of violence would more frequently resort to violence in circumstances where others probably would not. Critics of this thesis have argued that it is difficult to assess who carries "subculture of violence values" except via the behavior that is being predicted.

Along with most other institutions and traditions, mainstream criminology was challenged in the 1960s. Critics raised questions about the theories, data, methods, and even the definitions of crime used in criminology. The early challenges came from *labeling* theorists. This group used a variant of symbolic interaction to argue that law-violating behavior was widespread in the general

population, and the official labeling of a selected subset of violators was more a consequence of who the person was than of what the person had done (Becker 1963). Consequently, crime should not be defined as behavior that violates the law, because many people violate the law and are never arrested, prosecuted, or convicted. Rather, crime is a behavior that is selectively sanctioned, depending on who is engaging in it. It follows then that when criminologists use data produced by the criminal justice system, we are not studying crime but the criminal justice system itself. These data only tell us about the people selected for sanctioning, not about all of those who break the law. And, labeling theorists argued, most of our theories have been trying to explain why lower-class people engage disproportionately in crime, but since they do not—they are simply disproportionately sanctioned—the theories are pointless. What should be explained, labeling theorists argued, is why some people are more likely to be labeled and sanctioned as criminals than are others who engage in the same or similar behavior. The answer they offered was that those with sufficient resources—enough money, the right racial or gender status, etc.—to fend off labeling by the criminal justice system are simply less likely to be arrested.

A further contribution by labeling theorists is the idea that the labeling process actually creates deviance. The act of labeling people as criminals sets in motion processes that marginalize them from the mainstream, create in them the self-identification as "criminal," which in turn affects their behavior. In other words, the act of sanctioning causes more of the behavior the criminal justice system wishes to extinguish.

*Marxist criminology* actually began in 1916 with the publication of Bonger's *Criminality and Economic Conditions*, but it became central to criminological discourse in the late 1960s and 1970s (Chambliss 1975; Platt 1969; Quinney 1974; Taylor, Walton, and Young 1973). The Marxist critique of mainstream criminology can be summarized by focusing on the argument that it tends to ask three types of questions: Who commits crimes? How much crime is there? Why do people break the law? The Marxist perspective argues that these may be important questions, but more important are those that do not reify the law itself. Marxists argue that in addition to the above questions, criminologists should ask: Where does the

law come from? Whose interest does the law serve? Why are the laws structured and enforced in particular ways? Their answers to these questions are based on a class-conflict analysis. Power in social systems is allocated according to social class, and the powerful use the law to protect their interests and the status quo that perpetuates their superordinate status. So the law and criminal justice system practices primarily serve the interest of elites; while they at times serve the interests of other classes, their *raison d'être* is the interests of the powerful. Law is structured to protect the current status arrangements.

Both the labeling and Marxist perspectives experienced broad popularity among students of criminology. Many scholars responded to their critiques not by joining them, but by taking some lessons from the debate and moving forward to develop theories consistent with traditional directions and research methods that were not as dependent on data generated by the criminal justice system. Victimization surveys and self-report studies of crime have become more widely used, in part as a consequence of these critiques.

All of the theories mentioned so far have had significant empirical challenges. Most of the initial statements have been falsified or their proponents have had to revise the perspective in the face of evidence that did not support it. Several of these explanations of crime, or how societies control their members, are used today in a modified form, while others have evolved into contemporary theories that are being tested by researchers. No doubt when someone writes about criminology in the future, some or all of the more recent theories will have joined those that have been falsified or modified as a result of empirical analyses. Contemporary criminological theories tend to be in the control theory, rational choice, or conflict traditions. However, these are not mutually exclusive (e.g. some control theories are very much rational choice theories).

*Control theorists* begin by saying that we ask the wrong question when we seek to understand why some people commit crimes. We should instead seek to explain why most people do *not* violate the rules. Control theorists reason that we do not need to explain why someone who is hungry or has less will steal from those who have what they want or

need. We do not need to explain why the frustrated and angry among us will express their feelings violently. The interesting question is: Why *don't* most people who have less or have grievances engage in property or violent crime? Other versions of control theory (Nye 1958; Reckless 1961) preceded his, but Hirschi's (1969) classic answer to this question is that those who are "socially bonded" to critical institutions such as families, schools, conventional norms, etc., are less likely to violate the law. The bonds give them a "stake in conformity," something they value and would risk losing should they violate important rules. Though not initially stated in rational choice terms of classical school ideas, one can see similarities between Hirschi's notions of "stakes in conformity" and the Classical School of criminology's concept of the "social contract."

Control theory has evolved in a number of directions. Gottfredson and Hirschi (1990) have argued that crime and deviance are a consequence of the failure of some to develop adequate self-control. Self-control is developed, if at all, early in life—by the age of eight or ten. Those who do not develop self-control will, they argue, exhibit various forms (depending on their age) of deviance throughout their lives. These people will commit crimes disproportionately during the crime-prone ages, between adolescence and the late twenties. Sampson and Laub's (1993) "life-course" perspective argues that social bonds change for people at different stages of their lives. Bonds to families are important early, to schools and law-abiding peers later, and eventually bonds to spouses, one's own children, and jobs and careers ultimately tie people to conventional norms and prevent crime and deviance. Tittle (1995) has offered a *control balance theory* of deviance. He believes that those who balance control exerted by self with control by others that they are subjected to are the least likely to commit acts of crime and deviance. Those whose "control ratio" is out of balance, with too much or too little regulation by the self or others, have a higher probability of breaking rules.

The contemporary rational choice perspective of crime has been most explicitly articulated by economists (Becker 1968; Ehrlich 1973). Becker described the choices people make in social behavior, including crime, as much like those made in economic behavior. Before a purchase we weigh

the cost of an item against the utility, or benefit to be gained by owning that item. Likewise, before engaging in crime, a person weighs the cost—prison, loss of prestige, relationships, or even life—against the benefits to crime—pleasure, expression of anger, or material gain. Becker argued that when the benefits outweigh the cost individuals would be more likely to commit criminal actions.

Control theory can also be thought of as a rational choice theory. The “bonded” among us have a stake in conformity, or something to weigh on the “to be lost” side of a cost-benefit analysis. To engage in crime is to risk the loss of valued bonds to family, teachers, or employers. The bond in control theory, thought of in this way, is an informal deterrent. *Deterrence theory* (Geerken and Gove 1975) is ordinarily written of in terms of a formal system of deterrence provided by the criminal justice system. As originally conceived by Classical School criminologists, police, prosecutors, and the courts endeavor to raise the cost of committing crime so that those costs outweigh the benefits from illegal behavior. In the case of those convicted, the courts attempt to do this by varying the sentence so that those convicted become convinced that they must avoid crime in the future. This deterrence aimed at those already in violation is referred to as “specific” or “special” deterrence. The noncriminal public is persuaded by general deterrence to avoid crime. They perceive that crime does not pay when they see others sanctioned. So, when considering criminal options, they weigh the utility of illegal action against the potential cost in the forms of sanctions. According to deterrence theory, when those positive utilities are outweighed by the perceived cost—effective deterrence—they choose not to engage in crime.

A number of contemporary *conflict theories* of criminology are legacies of 1960s and 1970s-era critical perspectives. Most prominent among these is *feminist criminological theory* (Simpson 1989). Contemporary conflict theories, like earlier Marxist theory, address important questions about interest groups. They begin with the position that to understand social life, including crime and responses to crime, the social, political, and economic interests of contesting parties and groups must be taken into account (Marxists, of course, focus on economic conflict). Feminist criminology focuses on gender conflict in society. These

criminologists argue that to understand crimes by women, against women, and the reaction to both, we must consider the subordinate status of women. If poverty is one of the root causes of crime, then we should recognize first that the largest impoverished group in American society, and in most western industrialized societies, is children. Increasingly, children are raised in female-headed households whose poverty can be traced to the lower earnings of women, a consequence of gender stratification.

Other conflict analyses of crime focus on income inequality (Blau and Blau 1982), racial inequality (Sampson and Wilson 1994), and labor-market stratification (Crutchfield and Pitchford 1997). What these approaches share is the idea that to understand why individuals engage in crime, or why some groups or geographic areas have higher crime rates, or why patterns of criminal justice are the way they are, consideration of important social conflicts and cleavages must be brought into the analysis. Blau and Blau (1982) illustrated how income inequality, especially inequality based on race, is correlated with higher rates of violent crime. Those metropolitan areas with relatively high levels of inequality tend also to have more violence. Others (Williams 1984; Messener 1989) have argued that it is not so much relative inequality that leads to higher violent crime rates, but rather high levels of poverty. Sampson and Wilson (1994) argue that racial stratification is linked to economic stratification and this complex association accounts for higher levels of crime participation among blacks. Crutchfield and Pitchford (1997), studying a particular form of institutional stratification that was produced by segmented labor markets, found that those marginalized in the work force under some circumstances are more likely to engage in crime.

Hagan, Gillis, and Simpson (1985) developed what they call a “power control theory of delinquency.” They combine elements of social control theory and conflict theory to explain class and gender patterns of delinquency. They argue that because of gender stratification, parents seek to control their daughters more and because of class stratification, those in the higher classes have more capacity (because they have the available resources) to monitor (control) their children. Yet, children of the upper classes are actually free to commit

more delinquency by virtue of their social-class standing. Consequently, upper-class girls will be more controlled than their lower-class counterparts, but their brothers will have minor delinquency rates resembling those of lower-class boys.

### SOCIETY'S ATTEMPT TO CONTROL CRIME

While the theories discussed above focus on the causes of crime, they are also important for describing how social systems control crime. Societies attempt to control the behavior of people living within their borders with a combination of formal and informal systems of control. Social disorganization theory and anomie theory are examples of how crime is produced when normative control breaks down. Differential association and subcultural explanations describe how informal social control is subverted by socialization that supports criminal behavior rather than compliant behavior. Control theories focus specifically on how weak systems of informal social control fail. The obvious exception to this last statement is that portion of control theories that are also deterrence theory. Deterrence theory does consider informal systems of control, but an important part of this thesis is aimed at explaining how formal systems, or the criminal justice system in western societies, attempt to control criminal behavior. Most criminologists believe that informal systems of control are considerably more efficient than formal systems. This makes sense if one remembers that police cannot regulate us as much as we ourselves can when we have internalized conventional norms, and similarly, police cannot watch our behavior nearly as much as our families, friends, teachers, and neighbors can. The criminal justice system then is reduced to supporting informal systems of control (thus the interest in block-watch programs by police departments), engaging in community policing and patrol patterns that discourage crime, or reacting after violations have occurred.

### CONCLUSION

Criminologists are interested in answering questions about how crime should be defined, why crime occurs, and how societies seek to control

crime. The history of modern criminology, which can be traced to the early nineteenth century, has not produced definitive answers to these questions. To some students that is a source of frustration. To many of us the resulting ambiguity is the source of continuing interesting debate. More importantly, the disagreement among criminologists captures the complexity of social life. Oversimplification to achieve artificial closure on these debates will not produce quality answers to these questions, nor will it, to the consternation of some politicians, lead to workable solutions to crime problems. Most criminologists recognize that the complex debates about the answers to these three seemingly simple questions will ultimately be a more productive route to understanding crime and to finding effective means to address crime problems.

### REFERENCES

- Agnew, Robert 1992 "Foundations for a General Strain Theory of Crime and Delinquency." *Criminology* 30:47-87.
- Banfield, Edward 1968 *The Unheavenly City*. Boston: Little, Brown.
- Beccaria, Cesare (1764) 1963 *On Crimes and Punishments*. Translated by Henry Paolucci. Indianapolis, Ind.: Bobbs-Merrill.
- Becker, Gary S. 1968 "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76:169-217.
- Becker, Howard 1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Bentham, Jeremy (1765) 1970 *An Introduction to the Principles of Morals and Legislation*. London: Athlone Press.
- Blau, Judith, and Peter Blau 1982 "The Cost of Inequality: Metropolitan Structure and Violent Crime." *American Sociological Review* 47:114-128.
- Bonger, William 1916 *Criminality and Economic Conditions*. Boston: Little, Brown.
- Chambliss, William J. 1978 *On the Take: From Petty Crooks to Presidents*. Bloomington: Indiana University Press.
- 1975 "Toward a Political Economy of Crime." *Theory and Society* 2:149-170.
- Cloward, Richard, and Lloyd Ohlin 1960 *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press of Glencoe.

- Cohen, Albert 1955 *Delinquent Boys*. New York: Free Press of Glencoe.
- Crutchfield, Robert D., and Susan R. Pitchford 1997 "Work and Crime: The Effects of Labor Stratification." *Social Forces* 76:93-118.
- Ehrlich, Isaac 1973 "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81:521-565.
- Fisse, Brent, and John Braithwaite 1987 "The Impact of Publicity on Corporate Offenders: Ford Motor Company and the Pinto Papers." In M. David Ermann and Richard J. Lundmann, eds., *Corporate and Governmental Deviance*. Oxford: Oxford University Press.
- Foucault, Michel 1979 *Discipline and Punish: The Birth of the Prison*. New York: Vintage Books.
- Geerken, Michael, and Walter Gove 1975 "Deterrence: Some Theoretical Considerations." *Law and Society Review* 9:497-514.
- Gottfredson, Michael, and Travis Hirschi 1990 *A General Theory of Crime*. Stanford, Calif.: Stanford University Press.
- Hagan, John, A. R. Gillis, and John Simpson 1985 "The Class Structure of Gender and Delinquency: Toward a Power-Control Theory of Common Delinquent Behavior." *American Journal of Sociology* 90:1151-1178.
- Hindelang, Michael, Travis Hirschi, and Joseph Weis 1981 *Measuring Delinquency*. Newbury Park, Calif.: Sage.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- , and Michael Gottfredson 1983 "Age and the Explanation of Crime." *American Journal of Sociology* 89:552-584.
- Leek, Sybill 1970 *Phrenology*. New York: Macmillan.
- Lombroso, Cesare 1876 *L'Uomo Delinquente*. Milan: Hoepli.
- Massey, Douglas, and Nancy Denton 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Matsueda, Ross L. 1988 "The Current State of Differential Association Theory." *Crime and Delinquency* 34:277-306.
- Mednick, Sarnoff A. 1977 "A Bio-Social Theory of Learning of Law-Abiding Behavior." In Sarnoff A. Mednick and Karl O. Christiansen, eds., *Biosocial Bases of Criminal Behavior*. New York: Gardner.
- Merton, Robert K. 1938 "Social Structure and Anomie." *American Sociological Review* 3:672-682.
- Messener, Steven 1989 "Economic Discrimination and Societal Homicide Rates: Further Evidence on the Cost of Inequality." *American Sociological Review* 54:597-611.
- Miller, Walter 1958 "Lower Class Culture as a Generating Milieu of Gang Delinquency." *Journal of Social Issues* 14:5-19.
- Murray, Charles 1984 *Losing Ground: American Social Policy 1950-1980*. New York: Basic Books.
- Nye, Ivan F. 1958 *Family Relationships and Delinquent Behavior*. New York: John Wiley.
- Platt, Anthony 1969 *The Child Savers: The Invention of Delinquency*. Chicago: University of Chicago Press.
- Quinney, Richard 1974 *Critique of the Legal Order: Crime Control in a Capitalist Society*. Boston: Little, Brown.
- Reckless, Walter C. 1961 "A New Theory of Delinquency and Crime." *Federal Probation* 25:42-46.
- Sampson, Robert, and W. Byron Groves 1989 "Community Structure and Crime: Testing Social Disorganization Theory of Crime." *American Journal of Sociology* 94:774-802.
- Sampson, Robert, and John Laub 1993 *Crime in the Making*. Cambridge, Mass.: Harvard University Press.
- Sampson, Robert, and William J. Wilson 1994 "Toward a Theory of Race, Crime and Urban Inequality." In John Hagan and Ruth Peterson, eds., *Crime and Inequality*. Stanford, Calif.: Stanford University Press.
- Sanchez-Jankowski, Martí 1992 *Islands in the Street*. Los Angeles: University of California Press.
- Shaw, Clifford, and Henry McKay 1931 *Social Factors in Juvenile Delinquency*. Washington, D.C.: National Commission of Law Observance and Enforcement.
- Simpson, Sally S. 1989 "Feminist Theory, Crime, and Justice." *Criminology* 27:705-631.
- Sutherland, Edwin 1924 *Criminology*. Philadelphia: Lippincott.
- , and Donald R. Cressey 1974 *Criminology*, 9th ed. Philadelphia: Lippincott.
- Taylor, Ian, Paul Walton, and Jock Young 1973 *The New Criminology: For a Social Theory of Deviance*. London: Routledge and Kegan Paul.
- Tittle, Charles 1995 *Control Balance: Toward a General Theory of Deviance*. Boulder, Colo.: Westview Press.
- , W. J. Vilemez, and Douglas Smith 1978 "The Myth of Social Class and Criminality: An Empirical Assessment of the Empirical Evidence." *American Sociological Review* 47:505-518.
- Tonry, Michael (ed.) 1997 *Ethnicity, Crime, and Immigration: Comparative and Cross-National Perspectives*. Chicago: University of Chicago Press.

- Williams, Kirk R. 1984 "Economic Sources of Homicide: Reestimating the Effects of Poverty and Inequality." *American Sociological Review* 20:545-572.
- Wilson, James Q., and Richard Herrnstein 1985 *Crime and Human Nature*. New York: Simon and Schuster.
- Wilson, William J. 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Wolfgang, Marvin E., and Franco Ferracuti 1967 *The Subculture of Violence*. London: Tavistock Publishers.

ROBERT D. CRUTCHFIELD  
CHARIS KUBRIN

## CRITICAL THEORY

The term *critical theory* was used originally by members of the Institute for Social Research in Frankfurt, Germany, after they emigrated to the United States in the late 1930s, following the rise of Hitler. The term served as a code word for their version of Marxist social theory and research (Kellner 1990a). The term now refers primarily to Marxist studies done or inspired by this so-called Frankfurt School and its contemporary representatives such as Jurgen Habermas. Critical sociologists working in this tradition share several common tenets including a rejection of sociological positivism and its separation of facts from values; a commitment to the emancipation of humanity from all forms of exploitation, domination, or oppression; and a stress on the importance of human agency in social relations.

### THE FRANKFURT SCHOOL OF CRITICAL THEORY

The Institute for Social Research was founded in 1923 as a center for Marxist studies and was loosely affiliated with the university at Frankfurt, Germany. It remained independent of political party ties. Max Horkheimer became its director in 1931. Theodor Adorno, Erich Fromm, Leo Lowenthal, Herbert Marcuse, and, more distantly, Karl Korsch and Walter Benjamin were among the prominent theorists and researchers associated with the institute (Jay 1973). Initially, institute scholars sought to update Marxist theory by studying new social developments such as the expanding role of the

state in social planning and control. The rise of fascism and the collapse of effective opposition by workers' parties, however, prompted them to investigate new sources and forms of authoritarianism in culture, ideology, and personality development and to search for new oppositional forces. By stressing the importance and semiautonomy of culture, consciousness, and activism, they developed an innovative, humanistic, and open-ended version of Marxist theory that avoided the determinism and class reductionism of much of the Marxist theory that characterized their era (Held 1980).

"Immanent critique," a method of description and evaluation derived from Karl Marx and Georg W. F. Hegel, formed the core of the Frankfurt School's interdisciplinary approach to social research (Antonio 1981). As Marxists, members of the Frankfurt School were committed to a revolutionary project of human emancipation. Rather than critique existing social arrangements in terms of a set of ethical values imposed from "outside," however, they sought to judge social institutions by those institutions' own internal (i.e., "immanent") values and self-espoused ideological claims. (An example of the practical application of such an approach is the southern civil rights movement of the 1960s, which judged the South's racial caste system in light of professed American values of democracy, equality, and justice.) Immanent critique thus provided members of the Frankfurt School with a nonarbitrary standpoint for the critical examination of social institutions while it sensitized them to contradictions between social appearances and the deeper levels of social reality.

Immanent critique, or what Adorno (1973) termed "non-identity thinking," is possible because, as Horkheimer (1972, p. 27) put it, there is always "an irreducible tension between concept and being." That is, in any social organization, contradictions inevitably exist between what social practices are called—for example, "democracy" or "freedom" or "workers' parties"—and what, in their full complexity, they really are. This gap between existence and essence or appearance and reality, according to Adorno (1973, p. 5), "indicates the untruth of identity, the fact that the concept does not exhaust the thing conceived." The point of immanent critique is thus to probe empirically whether a given social reality negates its own claims—as, for example, to represent a

“just” or “equal” situation—as well as to uncover internal tendencies with a potential for change including new sources of resistance and opposition to repressive institutions.

Frankfurt School theorists found a paradigmatic example of immanent critique in the works of Karl Marx, including both his early writings on alienation and his later analyses of industrial capitalism. Best articulated by Marcuse (1941), their reading of *Capital* interpreted Marx’s text as operating on two levels. On one level, *Capital* was read as a historical analysis of social institutions’ progressive evolution, which resulted from conflicts between “forces” (such as technology) and “relations” (such as class conflicts) in economic production. Scientistic readings of Marx, however—especially by the generation of Marxist theorists immediately after the death of Marx—essentialized this dimension into a dogma that tended to neglect the role of human agency and stressed economic determinism in social history. But the Frankfurt School also read *Capital* as a “negative” or “immanent” critique of an important form of ideology, the bourgeois pseudo-science of economics. Here, Marx showed that the essence of capitalism as the exploitation of wage slavery contradicts its ideological representation or appearance as being a free exchange among equal parties (e.g., laborers and employers).

Members of the Frankfurt School interpreted the efforts that Marx devoted to the critique of ideology as an indication of his belief that freeing the consciousness of social actors from ideological illusion is an important form of political practice that potentially contributes to the expansion of human agency. Thus, they interpreted Marx’s theory of the production and exploitation of economic values as an empirical effort to understand the historically specific “laws of motion” of market-driven, capitalist societies. At the same time, however, it was also interpreted as an effort—motivated by faith in the potential efficacy of active opposition—to see through capitalism’s objectified processes that made a humanly created social world appear to be the product of inevitable, autonomous, and “natural” forces and to call for forms of revolutionary activism to defeat such forces of “alienation.”

Members of the Frankfurt School attempted to honor both dimensions of the Marxian legacy.

On the one hand, they sought to understand diverse social phenomena holistically as parts of an innerconnected “totality” structured primarily by such capitalistic principles as the commodity form of exchange relations and bureaucratic rationality. On the other hand, they avoided reducing complex social factors to a predetermined existence as shadowlike reflections of these basic tendencies (Jay 1984). Thus, the methodology of immanent critique propelled a provisional, antifoundationalist, and inductive approach to “truth” that allowed for the open-endedness of social action and referred the ultimate verification of sociological insights to the efficacy of historical struggles rather than to the immediate observation of empirical facts (Horkheimer 1972). In effect, they were saying that social “facts” are never fixed once and for all, as in the world of nature, but rather are subject to constant revisions by both the conscious aims and unintended consequences of collective action.

In their concrete studies, members of the Frankfurt School concentrated on the sources of social conformism that, by the 1930s, had undermined the Left’s faith in the revolutionary potential of the working class. They were among the first Marxists to relate Freud’s insights into personality development to widespread changes in family and socialization patterns that they believed had weakened the ego boundary between self and society and reduced personal autonomy (Fromm 1941). After they emigrated to the United States, these studies culminated in a series of survey research efforts, directed by Adorno and carried out by social scientists at the University of California, that investigated the relation between prejudice, especially anti-Semitism, and “the authoritarian personality” (Adorno et al. 1950). Later, in a more radical interpretation of Freud, Marcuse (1955) questioned whether conflicts between social constraints and bodily needs and desires might provide an impetus for revolt against capitalist repression if such conflicts were mediated by progressively oriented politics.

Once in the United States, members of the Frankfurt School emphasized another important source of conformism, the mass media. Holding that the best of “authentic art” contains a critical dimension that negates the status quo by pointing in utopian directions, they argued that commercialized and popular culture, shaped predominantly by market and bureaucratic imperatives, is



merely “mimetic” or imitative of the surrounding world of appearances. Making no demands on its audience to think for itself, the highly standardized products of the “culture industry” reinforce conformism by presenting idealized and reified images of contemporary society as the best of all possible worlds (see Kellner 1984–1985).

The most important contribution of the Frankfurt School was its investigation of the “dialectic of enlightenment” (Horkheimer and Adorno [1947] 1972). During the European Enlightenment, scientific reason had played a partisan role in the advance of freedom by challenging religious dogmatism and political absolutism. But according to the Frankfurt School, a particular form of reason, the instrumental rationality of efficiency and technology, has become a source of unfreedom in both capitalist and socialist societies during the modern era. Science and technology no longer play a liberating role in the critique of social institutions but have become new forms of domination. Dogmatic ideologies of scientism and operationalism absolutize the status quo and treat the social world as a “second nature” composed of law-governed facts, subject to manipulation but not to revolutionary transformation. Thus, under the sway of positivism, social thought becomes increasingly “one-dimensional” (Marcuse 1964). Consequently, the dimension of critique, the rational reflection on societal values and directions, and the ability to see alternative possibilities and new sources of opposition are increasingly suppressed by the hegemony of an eviscerated form of thinking. One-dimensional thinking, as an instrument of the totally “administered society,” thus reinforces the conformist tendencies promoted by family socialization and the culture industry and threatens both to close off and absorb dissent.

The Frankfurt School’s interpretation of the domination of culture by instrumental reason was indebted to Georg Lukacs’s ([1923] 1971) theory of reification and to Max Weber’s theory of rationalization. In the case of Lukacs, “reification” was understood to be the principal manifestation of the “commodity form” of social life whereby human activities, such as labor, are bought and sold as objects. Under such circumstances, social actors come to view the world of their own making as an objectified entity beyond their control at the same time that they attribute human powers to things.

For Lukacs, however, this form of life was historically unique to the capitalist mode of production and would be abolished with socialism.

In the 1950s, as they grew more pessimistic about the prospects for change, Horkheimer and Adorno, especially, came to accept Weber’s belief that rationalization was more fundamental than capitalism as the primary source of human oppression. Thus, they located the roots of instrumental rationality in a drive to dominate nature that they traced back to the origins of Western thought in Greek and Hebrew myths. This historical drive toward destructive domination extended from nature to society and the self. At the same time, Horkheimer and Adorno moved closer to Weber’s pessimistic depiction of the modern world as one of no exit from the “iron cage” of rationalization. In the context of this totalizing view of the destructive tendencies of Western culture—the images for which were Auschwitz and Hiroshima—the only acts of defiance that seemed feasible were purely intellectual “negations,” or what Marcuse (1964) termed “the great refusal” of intellectuals to go along with the one-dimensional society. Consequently, their interest in empirical sociological investigations, along with their faith in the efficacy of mass political movements, withdrew to a distant horizon of their concerns. Marcuse, like Benjamin before him, remained somewhat optimistic. Marcuse continued to investigate and support sources of opposition in racial, sexual, and Third World liberation movements.

Benjamin’s *Passagen-Werk* (Arcades project), originally titled *Dialectical Fairy Scene*, was an unfinished project of the 1930s that culminated in a collection of notes on nineteenth-century industrial culture in Paris. The Paris Arcade was an early precursor to the modern department store, a structure of passages displaying commodities in window showcases; it reached its height in the world expositions (e.g., the Paris Exposition in 1900). Through an interpretation of Benjamin’s notes, Susan Buck-Morss (1995) has brought this unfinished project to life. Benjamin drew on allegory as a method for analyzing the content and form of cultural images. In contrast to Horkheimer and Adorno’s “iron cage” view of mass culture, his dialectical approach held out hope for the revolutionary potential of mass-produced culture. Anticipating aspects of Symbolic Interactionism and feminist theories of performativity, Benjamin

explored the relationship between mass production as form (e.g., montage as a form of film production) and political subversion. In contrast to Horkheimer and Adorno who lamented the loss of authority in art and the family, Benjamin welcomed the abolition of traditional sources of authority and hailed the rapidity of technological change in mass production as potentially positive. He interpreted mass production as a form of mimicry that reproduced existing relations of authority and domination while lending itself to potentially subversive reinterpretations and reenactments of existing social relations and social meanings (Buck-Morss 1995).

While retaining an analysis of instrumental reason as a source of domination, Benjamin's allegorical approach worked to unveil the forces of contradiction that were crystallized as promise, progress, and ruin in mythical modern images. These images bore the revolutionary potential of the new to fulfill collective wishes for an unrealized social utopia contained in a more distant past. At the same time, they represented progress as the unrealized potential of capitalism to satisfy material needs and desires. For Benjamin, images of ruin represented the transitoriness, fragility, and destructiveness of capitalism as well as the potential for reawakening and a critical retelling of history (Buck-Morss 1995).

Even though some of the most prominent founders of the Frankfurt School abandoned radical social research in favor of an immanent critique of philosophy (as in Adorno 1973), the legacy of their sociological thought has inspired a vigorous tradition of empirical research among contemporary American social scientists. In large measure, this trend can be seen as a result of the popularization of Frankfurt School themes in the 1960s, when the New Left stressed liberation and consciousness raising, themes that continue to influence sociological practice. Stanley Aronowitz (1973), for example, along with Richard Sennet and Jonathan Cobb (1973), have rekindled the Frankfurt School's original interests in working-class culture in the context of consumer society. Henry Braverman (1974) has directed attention to processes of reification in work settings by focusing on scientific management and the separation of conception from labor in modern industry. Penetrating analyses also have been made of the

impact of commodification and instrumental rationalization on the family and socialization (Lasch 1977), law (Balbus 1977), education (Giroux 1988), advertising culture (Haug 1986), and mass media (Kellner 1990b), as well as other institutional areas. Feminist theorists have contributed a "doubled vision" to critical theory by showing the "systematic connectedness" of gender, class, and race relations (Kelly 1979) and by criticizing critical theory itself for its neglect of gender as a fundamental category of social analysis (Benjamin 1978; Fraser 1989). Among the most far-reaching and innovative contemporary studies are those of the contemporary German sociologist and philosopher Jurgen Habermas.

### THE CRITICAL THEORY OF JURGEN HABERMAS

Perhaps no social theorist since Max Weber has combined as comprehensive an understanding of modern social life with as deeply reflective an approach to the implications of theory and methods as Jurgen Habermas. Habermas has attempted to further the emancipatory project of the Frankfurt School by steering critical theory away from the pessimism that characterized the closing decades of Frankfurt School thought. At the same time, he has resumed the dialogue between empirical social science and critical theory to the mutual benefit of both. Further, he has given critical theory a new ethical and empirical grounding by moving its focus away from the relationship between consciousness and society and toward the philosophical and sociological implications of a critical theory of communicative action.

In sharp contrast to the Frankfurt School's increasing pessimism about the "dialectic of enlightenment," Habermas has attempted to defend the liberative potential of reason in the continuing struggle for freedom. While agreeing with the Frankfurt School's assessment of the destruction caused by instrumental rationality's unbridled domination of social life, he nonetheless recognizes the potential benefits of modern science and technology. The solution he offers to one-dimensional thought is thus not to abandon the "project of modernity" but rather to expand rational discourse about the *ends* of modern society. In order to further this goal, he has tried to unite science and ethics (fact and values) by recovering the

inherently rational component in symbolic interaction as well as developing an empirical political sociology that helps to critique the political effects of positivism as well as to identify the progressive potential of contemporary social movements.

From the beginning, Habermas (1970) has agreed with the classical Frankfurt School's contention that science and technology have become legitimating rhetorics for domination in modern society. At the same time, he has argued that alternative ways of knowing are mutually legitimate by showing that they have complementary roles to play in human affairs, even though their forms of validity and realms of appropriate application are distinct. That is, plural forms of knowledge represent different but complementary "knowledge interests" (Habermas 1971).

"Instrumental knowledge," based on the ability to predict, represents an interest in the technical control or mastery of nature. "Hermeneutical knowledge" represents an interest in the clarification of intersubjective understanding. Finally, "emancipatory knowledge" is best typified in the self-clarification that occurs freely in the nondirective communicative context provided by psychoanalysis. In the context of a democratic "public sphere," such self-clarification would have a macro-social parallel in the form of ideology critique had this space not been severely eroded by elite domination and technocratic decision making (Habermas 1989). Emancipatory knowledge thus has an interest in overcoming the illusions of reification, whether in the form of neurosis at the level of psychology or ideology at the level of society. In contrast to testable empirical hypotheses about objectified processes, the validity of emancipatory knowledge can be determined only by its beneficiaries. Its validity rests on the extent to which its subjects find themselves increasingly free from compulsion. Thus, a central problem of modern society is the hegemony of instrumental knowledge that, though appropriate in the realm of nature, is used to objectify and manipulate social relations. Instrumental knowledge thus eclipses the interpretive and emancipatory forms of knowledge that are also essential for guiding social life.

When sufficient attention is paid to interpersonal communication, Habermas (1979) contends that every act of speech can be seen as implying a

universal demand that interpersonal understanding be based on the free exchange and clarification of meanings. In other words, an immanent critique of language performance (which Habermas terms "universal pragmatics") reveals the presumption that communication not be distorted by differences in power between speakers. Thus, human communication is implicitly a demand for freedom and equality. By this form of immanent critique—consistent with the methodological standards of the Frankfurt School—Habermas attempts to demonstrate the potential validity of emancipatory knowledge so that it can be seen as a compelling challenge to the hegemony of instrumental knowledge. The purpose of Habermas's communication theory is thus highly partisan. By showing that no forms of knowledge are "value free" but always "interested," and that human communication inherently demands to occur freely without distortions caused by social power differentials, Habermas seeks politically to delegitimize conventions that confine social science to investigations of the means rather than the rational ends of social life.

In his subsequent works, Habermas has tried to reformulate this philosophical position in terms of a political sociology. To do so, he has profoundly redirected "historical materialism," the Marxist project to which he remains committed (see Habermas 1979). Habermas contends that Marx gave insufficient attention to communicative action by restricting it to the social class relations of work. This restriction, he argues, inclined the Marxist tradition toward an uncritical attitude toward technological domination as well as toward forms of scientism that contribute to the suppression of critique in regimes legitimated by Marxist ideology. Habermas relates his immanent critique of language performance to historical materialism by showing that sociocultural evolution occurs not only through the increasing rationality of technical control over nature (as Marx recognized) but also through advances in communicative rationality, that is, nondistorted communication. Thus, instrumental rationality and communicative rationality are complementary forms of societal "learning mechanisms." The problem of modernity is not science and technology in and of themselves, because they promise increased control over the environment, but rather the fact that instrumental rationality has eclipsed communicative rationality in social life. In other words, in

advanced industrial society, technical forms of control are no longer guided by consensually derived societal values. Democratic decision making is diminished under circumstances in which technical experts manipulate an objectified world, in which citizens are displaced from political decision making, and in which “reason”—identified exclusively with the “value free” prediction of isolated “facts”—is disqualified from reflection about the ends of social life.

More recently, Habermas (1987) has restated this theory sociologically to describe an uneven process of institutional development governed by opposing principles of “system” and “lifeworld.” In this formulation, the cultural lifeworld—the source of cultural meanings, social solidarity, and personal identity—is increasingly subject to “colonization” by the objectivistic “steering mechanisms” of the marketplace (money) and bureaucracy (power). On the levels of culture, society, and personality, such colonization tends to produce political crises resulting from the loss of meaning, increase of anomie, and loss of motivation. At the same time, however, objectivistic steering mechanisms remain indispensable because large-scale social systems cannot be guided by the face-to-face interactions that characterize the lifeworld. Thus, the state becomes a battleground for struggles involving the balance between the structuring principles of systems and lifeworlds. Habermas contends that it is in response to such crises that the forces of conservatism and the “new social movements” such as feminism and ecology are embattled and that it is here that the struggle for human liberation at present is being contested most directly. As formulated by Habermas, a critical theory of society aims at clarifying such struggles in order to contribute to the progressive democratization of modern society.

#### **CURRENT DEBATES: CRITICAL THEORY AND POSTSTRUCTURALISM**

In the late 1980s and early 1990s, the heightened influence of poststructuralism sparked intense debate between critical theorists and poststructuralists. Theorists staked out positions that tended to collapse distinct theories into oppositional categories (critical theory *or* poststructuralism) yet they agreed on several points. Both critical and poststructural theorists critiqued the transcendental claims of

Enlightenment thought (e.g., that truth transcends the particular and exists “out there” in its universality), understood knowledge and consciousness to be shaped by culture and history, and attacked disciplinary boundaries by calling for supra-disciplinary approaches to knowledge construction. Polarization, nonetheless, worked to emphasize differences, underplay points of agreement, and restrict awareness of how these approaches might complement one another (Best and Kellner 1991; Fraser 1997).

Because critical theory aspires to understand semiautonomous social systems (e.g., capital, science and technology, the state, and the family) as interconnected in an overarching matrix of domination (Best and Kellner 1991, p. 220), poststructuralists charge that it is a “grand theory” still mired in Enlightenment traditions that seek to understand society as a totality. In viewing the path to emancipation as the recovery of reason through a critical analysis of instrumentalism, scientism, and late capitalism, critical theory is seen as promoting a centralized view of power as emitting from a macro-system of domination. That is, by promoting a view of social subjects as overdetermined by class, critical theory is said to reduce subjectivity to social relations of domination that hover in an orbit of capitalist imperatives. By theorizing that subjectivity is formed through social interaction (e.g., intersubjectivity), Habermas departs from Horkheimer and Adorno’s view of the social subject as ego centered—as a self-reflexive critical subject (Best and Kellner 1991). Nonetheless, poststructuralists contend that Habermas, like his predecessors, essentializes knowledge. In other words, the capacity to recover reason either through critical reflexivity (Horkheimer, Adorno, and Marcuse) or through a form of communicative action that appeals to a normative order (Habermas) promotes a false understanding of subjectivity as “quasi-transcendental.”

In rebuttal, critical theorists argue that poststructuralist views of power as decentralized and diffuse uncouple power from systems of domination (Best and Kellner 1991). Poststructuralists view social subjectivity as a cultural construction that is formed in, and through, multiple and diffuse webs of language and power. Critics charge that such a diffuse understanding of power promotes a vision of society as a “view from everywhere” (Bordo 1993). Social identities are seen as

indeterminate and social differences as differences of equivalency (Flax 1990). This perspective results in analyses that focus on identity to the exclusion of systemic forms of domination. Thus, for example, while feminists who adhere to critical theory tend to analyze gender as a system of patriarchal domination, poststructuralist feminists, by contrast, tend to focus on the cultural production of gendered subjects, that is, on representation and identity. Habermas ([1980] 1997) and others argue that the avoidance of analyses of systems in favor of more fragmentary micro-analyses of discrete institutions, discourses, or practices is an antimodern movement that obscures the emancipatory potential of modernity (Best and Kellner 1991).

Although this debate is still stirring, some scholars are moving away from oppositional positions in favor of more complex readings of both traditions in order to synthesize or forge alliances between approaches (Best and Kellner 1991; Kellner 1995; Fraser 1997). Thus poststructuralism may serve as a corrective to the totalizing tendencies in critical theory while the latter prevents the neglect of social systems and calls attention to the relationship between multiple systems of domination and social subjectivities. In other words, critical theory points to the need to understand systemic forms of domination while poststructuralism warns against the reduction of social subjectivity to macro-overarching systems of domination. Thus drawing on both traditions, Nancy Fraser (1997, p. 219) suggests that a more accurate picture of social complexity “might conceive subjectivity as endowed with critical capacities and as culturally constructed” while viewing “critique as simultaneously situated and amenable to self-reflection.”

Theoretical and empirical applications of such a “both/and approach” abound. For instance, in recognizing that all knowledge is partial, black feminist theorists such as Patricia Hill Collins (1990) articulate both critical theoretical tenets and poststructuralist sensibilities by conceptualizing identity as socially constructed, historically specific, and culturally located while stressing systemic forms of domination without reducing identities to single systems of oppression (also see Agger 1998). Postcolonial theories likewise draw on both traditions in order to understand the fluid relationships among culture, systems of domination, social subjectivity, the process of “othering,” and

identity formation (see Williams and Chrisman 1993). Douglas Kellner’s (1997) empirical work on media culture likewise employs a multiperspectival approach that combines insights from cultural studies and poststructuralism with critical theory in order to understand mass media as a source of both domination and resistance, and as a way to account for the formation and communicative positionality of social subjects constituted through multiple systems of race, class, and gender. Habermas’s (1996) current theorizing on procedural democracy reflects a move toward the poststructuralism of Ernesto Laclau and Chantal Mouffe’s (1985) theory of radical democracy that stresses the potential collaboration of diverse agents in progressive social movements that aim at defending and expanding citizen participation in public life.

(SEE ALSO: *Marxist Sociology*)

#### REFERENCES

- Adorno, Theodor 1973 *Negative Dialectics*. New York: Seabury Press.
- , E. Frenkel-Brunswick, D. Levinson, and R. N. Sanford 1950 *The Authoritarian Personality*. New York: Harper.
- Agger, Ben 1998 *Critical Social Theory: An Introduction*. Boulder, Colo.: Westview Press.
- Antonio, Robert J. 1981 “Immanent Critique as the Core of Critical Theory.” *British Journal of Sociology* 32:330–345.
- Aronowitz, Stanley 1973 *False Promises*. New York: McGraw-Hill.
- Balbus, Isaac 1977 “Commodity Form and Legal Form: An Essay on ‘Relative Autonomy.’” *Law and Society Review* 11:571–588.
- Benjamin, Jessica 1978 “Authority and the Family Revisited: Or, A World without Fathers?” *New German Critique* 13 (Winter):35–57.
- Best, Steven, and Douglas Kellner 1991 *Postmodern Theory: Critical Interrogations*. New York: The Guilford Press.
- Bordo, Susan 1993 *Unbearable Weight: Feminism, Western Culture, and the Body*. Berkeley: University of California Press.
- Braverman, Harry 1974 *Labor and Monopoly Capital*. New York: Monthly Review Press.
- Buck-Morss, Susan 1995 *The Dialectics of Seeing: Walter Benjamin and the Arcades Project*. Cambridge, Mass.: The MIT Press.

- Flax, Jane 1990 *Thinking Fragments: Psychoanalysis, Feminism, and Postmodernism in the Contemporary West*. Los Angeles: University of California Press.
- Fraser, Nancy 1997 *Justice Interruptus: Critical Reflections on the "Postsocialist" Condition*. New York: Routledge.
- 1989 *Unruly Practices: Power, Discourse, and Gender in Contemporary Social Theory*. Minneapolis: University of Minnesota Press.
- Fromm, Erich 1941 *Escape from Freedom*. New York: Avon.
- Giroux, Henry A. 1988 *Schooling and the Struggle for Public Life*. Minneapolis: University of Minnesota Press.
- Habermas, Jurgen 1996 *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, Mass.: The MIT Press.
- 1989 *The Structural Transformation of the Public Sphere*. Cambridge, Mass.: The MIT Press.
- 1987 *The Theory of Communicative Action II: Lifeworld and System—A Critique of Functionalist Reason*. Boston: Beacon Press.
- 1986 "Three Normative Models of Democracy." In Seyla Benhabib, ed., *Democracy and Difference*. Princeton, N.J.: Princeton University Press.
- (1980) 1997 "Modernity: An Unfinished Project." In d'Entregrave,ves, Maurizio Passerin and Seyla Benhabib, eds., *Habermas and the Unfinished Project of Modernity*. Cambridge, Mass.: The MIT Press.
- 1979 *Communication and the Evolution of Society*. Boston: Beacon Press.
- 1971 *Knowledge and Human Interests*. Boston: Beacon Press.
- 1970 *Toward a Rational Society*. Boston: Beacon Press.
- Haug, W. F. 1986 *Critique of Commodity Aesthetics*. Minneapolis: University of Minnesota Press.
- Held, David 1980 *Introduction to Critical Theory*. Berkeley: University of California Press.
- Hill Collins, Patricia 1990 *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge.
- Horkheimer, Max 1972 *Critical Theory: Selected Essays*. New York: Herder and Herder.
- , and Theodor Adorno (1947) 1972 *Dialectic of Enlightenment*. New York: Seabury Press.
- Jay, Martin 1984 *Marxism and Totality*. Berkeley: University of California Press.
- 1973 *The Dialectical Imagination*. Boston: Little, Brown.
- Kellner, Douglas 1995 *Media Culture: Cultural Studies, Identity, and Politics Between the Modern and the Postmodern*. New York: Routledge.
- 1984–85 "Critical Theory and the Culture Industries: A Reassessment." *Telos* 62:196–209.
- 1990a "Critical Theory and the Crisis of Social Theory." *Sociological Perspectives* 33:11–33.
- 1990b *Television and the Crisis of Democracy*. Boulder, Colo.: Westview Press.
- Kelly, Joan 1979 "The Doubled Vision of Feminist Theory." *Feminist Studies* 5:216–227.
- Laclau, Ernesto, and Chantal Mouffe 1985 *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. London: Verso.
- Lasch, Christopher 1977 *Haven in a Heartless World*. New York: Basic Books.
- Lukacs, Georg (1923) 1971 *History and Class Consciousness*. Cambridge, Mass.: The MIT Press.
- Marcuse, Herbert 1964 *One-Dimensional Man*. Boston: Beacon Press.
- 1955 *Eros and Civilization*. Boston: Beacon Press.
- 1941 *Reason and Revolution*. New York: Oxford University Press.
- Sennett, Richard, and Jonathan Cobb 1973 *The Hidden Injuries of Class*. New York: Vintage Books.
- Williams, Patrick, and Laura Chrisman 1993 *Colonial Discourse and Post Colonial Theory: A Reader*. New York: Columbia University Press.

DWIGHT B. BILLINGS  
PATRICIA JENNINGS

## CROSS-CULTURAL ANALYSIS

Cross-cultural research has a long history in sociology (Armer and Grimshaw 1973; Kohn 1989; Miller-Loessi 1995). It most generally involves social research across societies or ethnic and subcultural groups within a society. Because a discussion of macro-level comparative historical research appears in another chapter of this encyclopedia, the focus here is primarily on cross-cultural analysis of social psychological processes. These include communicative and interactive processes within social institutions and more generally the relation between the individual and society and its institutions.

Although all sociological research is seen as comparative in nature, comparisons across

subcultural or cultural groups have distinct advantages for generating and testing sociological theory. Specifically, cross-cultural research can help “distinguish between those regularities in social behavior that are system specific and those that are universal” (Grimshaw 1973, p. 5). In this way, sociologists can distinguish between generalizations that are true of all cultural groups and those that apply for one group at one point in time. The lack of cross-cultural research has often led to the inappropriate universal application of sociological concepts that imply an intermediate (one cultural group at one point in time) level (Bendix 1963).

In addition to documenting universal and system-specific patterns in social behavior, cross-cultural analysis can provide researchers with experimental treatments (independent variables) unavailable in their own culture. Thus, specific propositions can be investigated experimentally that would be impossible to establish in a laboratory in the researcher’s own country (Strodtbeck 1964). Finally, cross-cultural analysis is beneficial for theory building in at least two respects. First, the documentation of differences in processes across cultures is often the first step in the refinement of existing theory and the generation of novel theoretical models. Second, cross-cultural analysis can lead to the discovery of unknown facts (behavioral patterns or interactive processes) that suggest new research problems that are the basis for theory refinement and construction.

#### **INTERPRETIVE AND INFERENCEAL PROBLEMS**

Cross-cultural researchers face a number of challenging interpretive and inferential problems that are related to the methodological strategies they employ (Bollen, Entwisle, and Alderson 1993). For example, Charles Ragin (1989) argues that most cross-cultural research at the macro level involves either intensive studies of one or a small group of representative or theoretically decisive cases or the extensive analysis of a large number of cases. Not surprisingly, extensive studies tend to emphasize statistical regularities while intensive studies search for generalizations that are interpreted within a cultural or historical context. This same pattern also appears in most micro-level cross-cultural research, and it is clearly related to both theoretical orientation and methodological preferences.

Some scholars take the position that cultural regularities must always be interpreted in cultural and historical context, while others argue that what appear to be cross-cultural differences may really be explained by lawful regularities at a more general level of analysis. Those in the first group most often employ primarily qualitative research strategies (intensive ethnographic and historical analysis of a few cases), while those in the second usually rely on quantitative techniques (multivariate or other forms of statistical analysis of large data sets).

#### **METHODOLOGICAL TECHNIQUES AND AVAILABLE DATA SOURCES**

The wide variety of techniques employed in cross-cultural analysis reflect the training and disciplinary interests of their practitioners. We discuss the methods of anthropologists, psychologists, and sociologists in turn. Anthropologists generally rely on different types of ethnographic tools for data collection, analysis, and reporting. Ethnographic research has the dual task of cultural description and cultural interpretation. The first involves uncovering the “native’s point of view” or the criteria the people under study use “to discriminate among things and how they respond to them and assign them meaning, including everything in their physical, behavioral, and social environments” (Goodenough 1980, pp. 31–32); while the second involves “stating, as explicitly as we can manage, what the knowledge thus attained demonstrates about the society in which it is found and, beyond that, about social life as such” (Geertz 1973, p. 27).

Three types of ethnographic approaches and methodological tools are generally employed in cross-cultural research. The first involves long-term participant observation and the thick description of the culture under study in line with Clifford Geertz’s (1973) interpretive perspective of culture. From this perspective, culture is seen as “layered multiple networks of meaning carried by words, acts, conceptions and other symbolic forms” (Marcus and Fisher 1986, p. 29). Thus the metaphor of culture in the interpretive approach is that of a text to be discovered, described, and interpreted. The second involves methods of ethno-science including elicitation tasks and interviews with key informants that yield data amenable to logical and statistical analysis to generate the “organizing principles underlying behavior” (Tyler 1969; also see

Werner and Schoepfle 1987). Ethnoscience views these organizing principles as the “grammar of the culture” that is part of the mental competence of members. The final type of ethnographic research is more positivistic and comparative in orientation. In this approach cross-cultural analysis is specifically defined as the use of “data collected by anthropologists concerning the customs and characteristics of various peoples throughout the world to test hypotheses concerning human behavior” (Whiting 1968, p. 693).

All three types of ethnographic research generate data preserved in research monographs or data archives such as the Human Relations Area Files, the *Ethnographic Atlas*, and the ever-expanding World Cultures data set that has been constructed around George Murdock and Douglas White’s (1968) *Standard Cross Cultural Sample*. A number of scholars (Barry 1980; Lagacé 1977; Murdock 1967; Whiting 1968; Levinson and Malone 1980) have provided detailed discussions of the contents, coding schemes, and methodological strengths and weaknesses of these archives as well as data analysis strategies and overviews of the variety of studies utilizing such data.

Most cross-cultural research in psychology involves the use of quasi-experimental methods. These include classical experimentation, clinical tests and projective techniques, systematic observation, and unobtrusive methods (see Berry, Poortinga, and Pandey, 1996; Triandis and Berry 1980). However, a number of psychologists have recently turned to observational and ethnographic methods in what has been termed “cultural psychology” (Shweder 1990). Much of the recent research in this area focuses on culture and human development and spans disciplinary boundaries and involves a wide variety of interpretive research methods (Greenfield and Suzuki 1998; Shweder et al. 1998).

Sociologists have made good use of intensive interviewing (Bertaux 1990) and ethnography (Corsaro 1988, 1994; Corsaro and Heise 1992) in cross-cultural analysis. However, they more frequently rely on the survey method in cross-cultural research and have contributed to the development of a number of archives of survey data (Kohn 1987; Lane 1990). The growth of such international surveys in recent years has been impressive. The World Fertility Survey (WFS) is an early example.

In one of the first efforts of its kind, women in forty-two developing countries were interviewed between 1974 and 1982 about their fertility behavior, marital and work history, and other aspects of their background. The WFS spawned hundreds of comparative studies that have contributed greatly to the understanding of human fertility (see, for example, Bohgaarts and Watkins 1996; Kirk and Pillet 1998). The Demographic and Health Survey (DHS) largely took up where the WFS left off. In this ongoing project, begun in 1984, nationally representative samples of women aged 15–49 in forty-seven countries have been surveyed regarding lifetime reproduction, fertility preferences, family planning practices, and the health of their children. For some countries, detailed data are available for husbands also and, for a few countries, in-depth interview data are also available.

Less specialized is the international counterpart to the U.S. General Social Survey (GSS); the International Social Survey Program (ISSP). The ISSP got its start in 1984 as researchers in the United States, Germany, Britain, and Australia agreed to field common topical modules in the course of conducting their regular national surveys. Beginning 1985 with a survey of attitudes toward government in the four founding countries, the ISSP has expanded to include surveys on topics as diverse as social networks and social support and attitudes regarding family, religion, work, the environment, gender relations, and national identity. Some specific modules have been replicated and, overall, a large proportion of the items from earlier surveys are carried over to new modules, giving the ISSP both a cross-cultural and longitudinal dimension. At present, thirty-one nations are participating in the ISSP. A number of non-member nations have also replicated specific modules. An interesting offshoot of the ISSP is the International Survey of Economic Attitudes (ISEA). Building on an ISSP module concerning beliefs regarding social inequality, the ISEA collects a wide array of information on attitudes regarding income inequality, social class, and economic policy. The first round was carried out between 1991–1993 in three countries. A second round was carried out in five countries over the 1994–1997 period, and a third round is currently under way. Some of the important reports based on data from these surveys include Jones and Broyfield (1997), Kelley and Evans (1995), and Western (1994).



Also of general interest are the surveys undertaken under the auspices of the World Values Survey Group and Eurobarometer. The World Values Survey (WVS) (originally termed the "European Values Survey") began in 1981 as social scientists in nine Western European countries administered a common survey of social, political, moral, and religious values in their respective countries. Between 1981 and 1984, this survey was replicated in fourteen additional countries, including a number of non-European countries. In 1990–1993, a second wave of the World Values Survey was conducted in a broader group of forty-three nations and a third wave was undertaken in 1995–1996. In terms of content, the WVS is broadly organized around values and norms regarding work, family, the meaning and purpose of life, and topical social issues. Specifically, respondents are queried on everything from their views on good and evil to their general state of health, from their associational memberships to their opinions of the value of scientific discoveries (see Inglehart 1997; MacIntosh 1998). The Eurobarometer surveys began in 1974 as an extension of an earlier series of European Community surveys. Designed primarily to gauge public attitudes toward the Common Market and other EU institutions, the Eurobarometer surveys, carried out every Fall and Spring, have expanded to include a variety of special topics of interest to sociologists, ranging from attitudes regarding AIDS to beliefs about the role of women (see Pettigrew 1998; Quillian 1995).

Finally, there are two more specialized projects that are deserving of note for their scale and scope. Of interest to students of crime and deviance is the International Crime (Victim) Survey (IC(V)S). Begun in 1989 and carried out again in 1992 and 1996, the IC(V)S gathers reports of crime, in addition to surveying attitudes regarding the police and the criminal justice system, fear of crime, and crime prevention (see Alvazzi del Frate and Patrignani 1995; Zvekic 1996). At present, over fifty countries have participated in the IC(V)S. Scholars interested in cross-cultural dimensions of poverty and development have benefitted from the Living Standards Measurement Study (LSMS). In this World Bank-directed program of research, surveys have been conducted in over two dozen developing countries since 1980 with the aim of gauging the welfare of households, understanding household behavior, and assessing the impact of

government policy on standards of living. The central instrument is a household questionnaire that details patterns of consumption. Other modules, collecting information regarding the local community, pricing, and schools and health facilities, have also been administered in a number of cases (see Grosh and Glewwe 1998; Stecklov 1997).

### CHALLENGES AND PROBLEMS IN CROSS-CULTURAL RESEARCH

There are numerous methodological problems in cross-cultural research including: acquiring the needed linguistic and cultural skills and research funds; gaining access to field sites and data archives; defining and selecting comparable units; ensuring the representativeness of selected cases; and determining conceptual equivalence and measurement reliability and validity. These first two sets of problems are obvious, but not easily resolved. Cross-cultural analysis is costly in terms of time and money, and it usually demands at least a minimal level (and often much more) of education in the history, language, and culture of groups of people foreign to the researcher. The difficulties of gaining access to, and cooperation from, individuals and groups in cross-cultural research "are always experienced but rarely acknowledged by comparative researchers" (Armer 1973, pp. 58–59). Specific discussions of, and development of strategies for, gaining access are crucial because research can not begin without such access. Additionally, casual, insensitive, or ethnocentric presentation of self and research goals to foreign gatekeepers (officials, scholars, and those individuals directly studied) not only negatively affects the original study, but can also cause serious problems for others who plan future cross-cultural research (Form 1973; Portes 1973). Given the cultural isolation of many social scientists in the United States, it is not surprising that these practical problems have contributed to the lack of cross-cultural research in American sociology. However, the internationalization of the social sciences and the globalization of social and environmental issues are contributing to the gradual elimination of many of these practical problems (Sztompka 1988).

For the cross-cultural analysis of social psychological processes the unit of analysis is most often interactive events or individuals that are sampled

from whole cultures or subunits such as communities or institutions (e.g., family, school, or workplace). The appropriateness of individuals as the basic unit of analysis has been a hotly debated issue in sociology. The problem is even more acute in cross-cultural analysis, especially in cultures "that lack the individualistic, participatory characteristics of Western societies" (Armer 1983, p. 62). In addition to the special difficulties of representative, theoretical, or random sampling of cases (Elder 1973; Van Meter 1990), cross-cultural researchers must also deal with "Galton's problem." According to the British statistician, Sir Francis Galton, "valid comparison requires mutually independent and isolated cases, and therefore cultural diffusion, cultural contact, culture clash or outright conquest—with their consequent borrowing, imitation, migrations etc.—invalidates the results of comparative studies" (Sztompka 1988, p. 213). Although several researchers have presented strategies for dealing with Galton's problem for correlational studies of data archives (see Naroll, Michik, and Naroll 1980), the problem of cultural diffusion is often overlooked in many quantitative and qualitative cross-cultural studies.

Undoubtedly, ensuring conceptual equivalence and achieving valid measures are the most challenging methodological problems of cross-cultural research. Central to these problems is the wide variation in language and meaning systems across cultural groups. Anthropologists have attempted to address the problem of conceptual equivalence with the distinction between "emic" and "etic." Emics refer to local (single culture) meaning, function and structure, while etics are culture-free (or at least operate in more than one culture) aspects of the world (Pike 1966). A major problem in cross-cultural analysis is the use of emic concepts of one culture to explain characteristics of another culture. In fact, many cross-cultural studies involve the use of "imposed etics," that is Euro-American emics that are "imposed blindly and even ethnocentrically on a set of phenomena which occur in other cultural systems" (Berry 1980, p. 12). A number of procedures have been developed to ensure emic-etic distinctions and to estimate the validity of such measures (Brislin 1980; Naroll, Michik, and Naroll 1980).

Addressing conceptual relevance in cross-cultural research does not, of course, ensure valid measures. All forms of data collection and analysis

are dependent on implicit theories of language and communication (Cicourel 1964). As social scientists have come to learn more about communicative systems within and across cultures, there has been a growing awareness that problems related to language in cross-cultural analysis are not easily resolved. There is also a recognition that these problems go beyond the accurate translation of measurement instruments (Brislin 1970; Grimshaw 1973), to the incorporation of findings from studies on communicative competence across cultural groups into cross-cultural research (Briggs 1986; Gumperz 1982).

### THE FUTURE OF CROSS-CULTURAL ANALYSIS

There is a solid basis for optimism regarding the future of cross-cultural analysis. Over the last twenty years there has been remarkable growth in international organizations and cooperation among international scholars in the social sciences. These developments have not only resulted in an increase in cross-cultural research, but also have led to necessary debates about the theoretical and methodological state of cross-cultural analysis (Øyen 1990; Kohn 1989).

Cooperation among international scholars in cross-cultural analysis has also contributed to the breaking down of disciplinary boundaries. In the area of childhood socialization and the sociology of childhood, for example, there have been a number of cross-cultural contributions to what can be termed "development in sociocultural context" by anthropologists (Heath 1983), psychologists (Rogoff 1990), sociologists (Corsaro 1997), and linguists (Ochs 1988). Developing interest in children and childhood in sociology has resulted in the establishment of a new research committee ("Sociology of Childhood") in the International Sociological Association (ISA) and a new journal titled *Childhood: A Global Journal of Child Research*, as well as the publication of several international reports and edited volumes (see Qvortrup, Bardy, Sgritta, and Wintersberger 1994). Less interdisciplinary, perhaps, but no less impressive, has been the degree of international cooperation that has developed around a number of other research committees of the ISA. The ISA Research Committee on Stratification, for instance, has had a

long history of such collaboration. This has involved the development of a common agenda (official and unofficial) and the pursuit of a common program of research by scholars around the globe (Ganzeboom, Treiman, and Ultee 1991). In addition to these developments, the growth of international data sets and their ready availability due to the new technologies such as the Internet are also grounds for optimism regarding cross-cultural research.

Despite growing international and multi-discipline cooperation and recognition of the importance of comparative research, it is still fair to say that cross-cultural analysis remains at the periphery of American sociology and social psychology. Although there has been some reversal of the growing trend toward narrow specialization over the last ten years, such specialization is still apparent in the nature of publications and the training of graduate students in these disciplines. There is a clear need to instill a healthy skepticism regarding the cultural relativity of a great deal of theory and method in social psychology in future scholars. Only then will future sociologists and social psychologists come to appreciate fully the potential of cross-cultural analysis.

## REFERENCES

- Alvazzi del Frate, Anna and Angela Patrignani 1995 "Women's Victimization in Developing Countries." *Issues and Reports* (UNIRI) 5:1-16.
- Armer, Michael 1973 "Methodological Problems and Possibilities in Comparative Research." In M. Armer and A. Grimshaw, eds., *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- , and Allen Grimshaw, (eds.) 1973 *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- Barry, Herbert 1980 "Description and Uses of the Human Relations Area Files." In H. Triandis and J. Berry, eds., *Handbook of Cross Cultural Methodology, Volume 2 Methodology*. Boston: Allyn and Bacon.
- Bendix, 1963 "Concepts and Generalizations in Comparative Sociological Studies." *American Sociological Review* 28:532-539.
- Berry, J. W. 1980 "Introduction to Methodology." In H. Triandis and J. Berry, eds., *Handbook of Cross Cultural Methodology*. Boston: Allyn and Bacon.
- , Y. H. Poortinga, and J. Pandey, (eds.) 1996 *Handbook of Cross-Cultural Psychology: Theory and Method*, 2nd edition. Boston: Allyn and Bacon.
- Bertaux, Daniel 1990 "Oral History Approaches to an International Social Movement." In E. Øyen, ed., *Comparative Methodology: Theory and Practice in International Social Research*. Newbury Park, Calif.: Sage.
- Bollen, Kenneth A., Barbara Entwisle, and Arthur S. Alderson 1993 "Macrocomparative Research Methods." *Annual Review of Sociology* 19:321-351.
- Bongaarts, John, and Susan Cotts Watkins 1996 "Social Interactions and Contemporary Fertility Transitions." *Population and Development Review* 22:639-682.
- Briggs, Charles 1986 *Learning How to Ask*. New York: Cambridge.
- Brislin, Richard 1970 "Back-Translation for Cross-Cultural Research." *Journal of Cross-Cultural Psychology* 1:185-216.
- 1980 "Translation and Content Analysis of Oral and Written Materials." In H. Triandis and J. Berry, eds., *Handbook of Cross Cultural Methodology*. Boston: Allyn and Bacon.
- Cicourel, Aaron 1964 *Method and Measurement in Sociology*. New York: Free Press.
- Corsaro, William 1988 "Routines in the Peer Culture of American and Italian Nursery School Children." *Sociology of Education* 61:1-14.
- 1994 "Discussion, Debate, and Friendship: Peer Discourse in Nursery Schools in the U.S. and Italy." *Sociology of Education* 67:1-26.
- , and David Heise 1990 "Event Structure Models From Ethnographic Data." *Sociological Methodology* 20:1-57.
- Elder, Joseph 1973 "Problems of Cross-Cultural Methodology: Instrumentation and Interviewing in India." In M. Armer and A. Grimshaw, eds., *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- Form, William 1973 "Field Problems in Comparative Research: The Politics of Distrust." In M. Armer and A. Grimshaw, eds., *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- Ganzeboom, Harry B. G., Donald J. Treiman, and Wout C. Ultee 1991 "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277-302.
- Geertz, Clifford 1973 *The Interpretation of Cultures*. New York: Basic Books.
- Goodenough, Ward 1980 "Ethnographic Field Techniques." In H. Triandis and J. Berry, eds., *Handbook of Cross Cultural Methodology*. Boston: Allyn and Bacon.

- Greenfield, Patricia and Lalita Suzuki 1998 "Culture and Human Development: Implications for Parenting, Education, Pediatrics, and Mental Health." In W. Damon, I. Siegel, K. Anne Renninger and K. Ann Renninger, eds., *Handbook of Child Psychology: Volume 4, Child Psychology in Practice*. New York: Wiley.
- Grimshaw, Allen 1973 "Comparative Sociology: In What Ways Different from Other Sociologies?" In M. Armer and A. Grimshaw, eds., *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- Grosh, Margaret E., and Paul Glewwe 1998 "Data Watch: The World Bank's Living Standards Measurement Study Household Surveys." *Journal of Economic Perspectives* 12:187-196.
- Gumperz, John J. 1982 *Discourse Strategies*. New York: Cambridge University Press.
- Heath, Shirley 1989 "Oral and Literate Traditions Among Black Americans Living in Poverty." *American Psychologist* 44:45-56.
- Inglehart, Ronald 1997 "Modernization, Postmodernization and Changing Perceptions of Risk." *International Review of Sociology* 7:449-459.
- Jones, Rachel K., and April Brayfield 1997 "Life's Greatest Joy: European Attitudes Towards the Centrality of Children." *Social Forces* 75:1239-1270.
- Kelley, Jonathan and M. D. R. Evans 1995 "Class Conflict in Six Western Nations." *American Sociological Review* 60:157-178.
- Kirk, Dudley, and Bernard Pillet 1998 "Fertility Levels, Trends, and Differentials in Sub-Saharan Africa in the 1980s and 1990s." *Studies in Family Planning* 29:1-22.
- Kohn, Melvin 1987 "Cross-National Research as an Analytic Strategy." *American Sociological Review* 52:713-731.
- (ed.) 1989 *Cross-National Research in Sociology*. Newbury Park: Calif.: Sage.
- Lagacé, R. O. (ed.) 1977 *Sixty Cultures: A Guide to the HRAF Probability Sample Files*. New Haven, Conn.: HRAF.
- Lane, Jan-Erik 1990 "Data Archives as an Instrument for Comparative Research." In E. Øyen, ed., *Comparative Methodology: Theory and Practice in International Social Research*. Newbury Park, Calif.: Sage.
- Levinson, David and Martin J. Malone 1980 *Toward Explaining Human Culture: A Critical Review of the Findings of Worldwide Cross-Cultural Research*. New Haven, Conn.: HRAF Press.
- MacIntosh, Randall 1998 "Global Attitude Measurement: An Assessment of the World Values Survey Postmaterialism Scale." *American Sociological Review* 63:452-464.
- Miller-Loessi, Karen 1995 "Comparative Social Psychology: Cross-Cultural and Cross-National." In K. Cook, G. A. Fine, and J. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Murdock, G. P. 1967 *Ethnographic Atlas*. Pittsburgh.: University of Pittsburgh Press.
- , and Douglas White 1969 "Standard Cross-Cultural Sample." *Ethnography* 8:329-369.
- Naroll, Raoul, Gary Michik, and Frada Naroll 1980 "Holocultural Research." In H. Triandis and J. Berry, eds., *Handbook of Cross Cultural Methodology*. Boston: Allyn and Bacon.
- Ochs, Elinor 1988 *Culture and Language Development: Language Acquisition and Language Socialization in a Samoan Village*. New York: Cambridge University Press.
- Øyen, Else 1990 "The Imperfection of Comparisons." In E. Øyen, ed., *Comparative Methodology: Theory and Practice in International Social Research*. Newbury Park, Calif.: Sage.
- Pettigrew, Thomas 1998 "Reactions toward the New Minorities of Western Europe." *Annual Review of Sociology* 24:77-103.
- Pike, R. 1966 *Language in Relation to a United Theory of the Structure of Human Behavior*. The Hague: Mouton.
- Portes, Alejandro 1973 "Perception of the U.S. Sociologist and Its Impact on Cross-National Research." In M. Armer and A. Grimshaw, eds., *Comparative Social Research: Methodological Problems and Strategies*. New York: Wiley.
- Quillian, Lincoln 1995 "Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-Immigrant and Racial Prejudice in Europe." *American Sociological Review* 60:586-611.
- Qvortrup, Jens, Marjatta Bardy, Sgritta Giovanni, and Helmut Wintersberger (eds.) 1994 *Childhood Matters: Social Theory, Practice, and Politics*. Aldershot, England: Avebury.
- Ragin, Charles 1989 "New Directions in Comparative Research." In M. Kohn, ed., *Cross-National Research in Sociology*. Newbury Park, Calif.: Sage.
- Rogoff, Barbara 1990 *Apprenticeship in Thinking: Cognitive Development in Social Context*. New York: Oxford University Press.
- Shweder, Richard 1990 "Cultural Psychology: What Is It?" In J. Stigler, R. Shweder, and G. Herdt, eds., *Cultural Psychology: Essays on Comparative Human Development*. New York: Cambridge University Press.
- , J. Goodnow, G. Hatano, R. LeVine, H. Marcus, and P. Miller 1998 "The Cultural Psychology of Development: One Mine Many Mentalities." In W.

- Damon and R. Lerner, eds., *Handbook of Child Psychology: Volume 1, Theoretical Models of Human Development*. New York: Wiley.
- Stecklov, Guy 1997 "Intergenerational Resource Flows in Côte d'Ivoire: Empirical Analysis of Aggregate Flows." *Population and Development Review* 23:525-553.
- Strodtbeck, Fred 1964 "Considerations of Meta-Method in Cross-Cultural Studies." *American Anthropologist* 66:223-229.
- Sztompka, Piotr 1988 "Conceptual Frameworks in Comparative Inquiry: Divergent or Convergent?" *International Sociology* 3:207-219.
- Triandis, Harry, and John Berry 1980 *Handbook of Cross-Cultural Psychology Volume 2, Methodology*. Boston: Allyn and Bacon.
- Tyler, Stephen 1969 "Introduction." In S. Tyler, ed., *Cognitive Anthropology*. New York: Holt, Rinehart and Winston.
- Van Meter, Karl 1990 "Sampling and Cross-Classification Analysis in International Social Research." In E. Øyen, ed., *Comparative Methodology: Theory and Practice in International Social Research*. Newbury Park, Calif.: Sage.
- Werner, O, and G. M. Schoepfle 1987 *Systematic Fieldwork: Foundations of Ethnography and Interviewing*. Beverly Hills, Calif.: Sage.
- Western, Bruce 1994 "Institutional Mechanism for Unionization in Sixteen OECD Countries: An Analysis of Social Survey Data." *Social Forces* 73:497-520.
- Whiting, John 1968 "Methods and Problems in Cross-Cultural Research." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology, Volume 2*. Reading, Mass.: Addison-Wesley.
- Zvekic, Ugljesa 1996. "The International Crime (Victim) Survey: Issues of Comparative Advantages and Disadvantages." *International Criminal Justice Review* 6:1-21.

ARTHUR S. ALDERSON  
WILLIAM A. CORSARO

## CROWDS AND RIOTS

Crowds are a ubiquitous feature of everyday life. People have long assembled collectively to observe, to celebrate, and to protest various happenings in their everyday lives, be they natural events, such as a solar eclipse, or the result of human contrivance, such as the introduction of machinery into the production process. The historical

record is replete with examples of crowds functioning as important textual markers, helping to shape and define a particular event, as well as strategic precipitants and carriers of the events themselves. The storming of the Bastille and the sit-ins and marches associated with the civil rights movement are examples of crowds functioning as both important markers and carriers of some larger historical happening. Not all crowds function so significantly, of course. Most are mere sideshows to the flow of history. Nonetheless, the collective assemblages or gatherings called crowds are ongoing features of the social world and, as a consequence, have long been the object of theorizing and inquiry, ranging from the psychologistic renderings of Gustav LeBon (1895) and Sigmund Freud (1921) to the more sociological accounts of Neil Smelser (1963) and Ralph Turner and Lewis Killian (1987) to the highly systematic and empirically grounded observations of Clark McPhail and his associates (1983, 1991).

Crowds have traditionally been analyzed as a variant of the broader category of social phenomena called collective behavior. Broadly conceived, collective behavior refers to group problem solving behavior that encompasses crowds, mass phenomena, issue-specific publics, and social movements. More narrowly, collective behavior refers to "two or more persons engaged in one or more behaviors (e.g., orientation, locomotion, gesticulation, tactile manipulation, and/or vocalization) that can be judged common or convergent on one or more dimensions (e.g., direction, velocity tempo, and/or substantive content)" (McPhail and Wohlstein 1983, pp. 580-581). Implicit in both conceptions is a continuum on which collective behavior can vary in terms of the extent to which its participants are in close proximity or diffused in time and space. Instances of collective behavior in which individuals are in close physical proximity, such that they can monitor one another by being visible to, or within earshot of, one another, are constitutive of crowds. Examples include protest demonstrations, victory celebrations, riots, and the dispersal processes associated with flight from burning buildings. In contrast are forms of collective behavior that occur among individuals who are not physically proximate but who still share a common focus of attention and engage in some parallel or common behaviors without developing the debate characteristic of the public or the

organization of social movements, and who are linked together by social networks, the media, or both. Examples of this form of collective behavior, referred to as diffuse collective behavior (Turner and Killian 1987) or the mass (Lofland 1981), include fads and crazes, deviant epidemics, mass hysteria, and collective blaming. Although crowds and diffuse collective behavior are not mutually exclusive phenomena, they are analytically distinct and tend to generate somewhat different literatures—thus, the focus on crowds in this selection.

Understanding crowds and the kindred collective phenomenon called “riots” requires consideration of five questions: (1) How do these forms of collective behavior differ from the crowd forms typically associated with everyday behavior, such as audiences and queues? (2) What are the distinctive features of crowds as collective behavior? (3) What are the conditions underlying the emergence of crowds? (4) What accounts for the coordination of crowd behavior? and (5) What are the correlates and/or predictors of individual participation?

#### **DISTINGUISHING BETWEEN THE CROWDS OF COLLECTIVE BEHAVIOR AND EVERYDAY BEHAVIOR**

There has been increasing recognition of the continuity between collective behavior and everyday behavior, yet the existence of collective behavior as an area of sociological analysis rests in part on the assumption of significant differences between collective behavior and everyday institutionalized behavior. In the case of crowds, those commonly associated with everyday life, such as at sports events and holiday parades, tend to be highly conventionalized in at least two or three ways. Such gatherings are recurrent affairs that are scheduled for a definite place at a definite time; they are calendarized both temporally and spatially. Second, associated behaviors and emotional states are typically routinized in the sense that they are normatively regularized and anticipated. And third, they tend to be sponsored and orchestrated by the state, a community, or a societal institution, as in the case of most holiday parades and electoral political rallies. Accordingly, they are typically socially approved affairs that function to reaffirm rather than challenge some institutional arrangement or the larger social order itself.

In contrast, crowds commonly associated with collective behavior, such as protest demonstrations, victory celebrations, and riots, usually challenge or disrupt the existing order. This is due in part to the fact that these crowds are neither temporally nor spatially routinized. Instead, as David Snow, Louis Zurcher, and Robert Peters (1981) have noted, they are more likely to be unscheduled and staged in spatial areas (streets, parks, malls) or physical structures (office buildings, theaters, lunch counters) that were designed for institutionalized, everyday behavior rather than contentious or celebratory crowds. Such crowd activities are also extrainstitutional, and thus unconventional, in the sense that they are frequently based on normative guidelines that are emergent and ephemeral rather than enduring (Turner and Killian 1987), on the appropriation and redefinition of existing networks or social relationships (Weller and Quarantelli 1973), or on both.

Crowd behavior has long been described as “extraordinary” in the sense that its occurrence indicates that something unusual is happening. Precisely what it is that gives rise to the sense that something “outside the ordinary” is occurring is rarely specified unambiguously, however. John Lofland (1981) suggests that it is increased levels of emotional arousal, but such arousal is not peculiar to crowd episodes. The conceptualization offered here suggests several possibilities: It is the appropriation and use of spatial areas, physical structures, or social networks and relations for purposes other than those for which they were intended or designed that indicates something extraordinary is happening.

#### **THE CHARACTERISTIC FEATURES OF CROWDS**

Crowds have been portrayed historically and journalistically as if they are monolithic entities characterized by participant and behavioral homogeneity. Turner and Killian (1987, p. 26) called this image into question, referring to it as “the illusion of unanimity,” but not until the turn toward more systematic empirical examination of crowds was it firmly established that crowd behaviors are typically quite varied and highly differentiated, and that crowd participants are generally quite heterogeneous in terms of orientation and behavior.

**Variation in Crowd Behaviors and "Riots."**

Based on extensive field observation of crowds, Sam Wright (1978) differentiated between two broad categories of crowd behaviors: crowd activities and task activities. "Crowd activities" refer to the redundant behavior seemingly common to all incidents of crowd behavior, such as assemblage, milling, and divergence. In their overview of empirical research on behaviors within crowds, McPhail and Ronald Wohlstein (1983) include collective locomotion, collective orientation, collective gesticulation, and collective vocalization among the types of crowd behaviors "repeatedly observed across a variety of gatherings, demonstrations, and some riots" (p. 595).

Taking these observations together, one can identify the following "crowd activities" (Wright 1978) or "elementary forms" of crowd behavior (McPhail and Wohlstein 1983): assemblage/convergence; milling; collective orientation (e.g., common or convergent gaze, focus, or attention); collective locomotion (e.g., common or convergent movement or surges); collective gesticulation (e.g., common or convergent nonverbal signaling); collective vocalization (e.g., chanting, singing, booing, cheering); and divergence/dispersal. Given the recurrent and seemingly universal character of these basic crowd behaviors, it is clear that they do not distinguish between types of crowds, that is, between demonstrations, celebrations, and riots.

To get at the variation in types of crowds, attention must be turned to what Wright conceptualized as "task activities" (1978). These refer to joint activities that are particular to and necessary for the attainment of a specific goal or the resolution of a specific problem. It is these goal-directed and problem-oriented activities that constitute the primary object of attention and thus help give meaning to the larger collective episode. Examples of task activities include parading or mass marching, mass assembly with speechmaking, picketing, proselytizing, temporary occupation of premises, lynching, taunting and harassment, property destruction, looting, and sniping.

Several caveats should be kept in mind with respect to crowd task activities. First, any listing of task activities is unlikely to be exhaustive, because they vary historically and culturally. Charles Tilly's (1978) concept of "repertoires of collective action" underscores this variation. Tilly has stressed

that while there are innumerable ways in which people could pursue collective ends, alternatives are in fact limited by sociohistorical forces. His research suggests, for example, that the massed march, mass assembly, and temporary occupation of premises are all collective task activities specific to the twentieth century.

Second, crowd task activities are not mutually exclusive but are typically combined in an interactive fashion during the history of a crowd episode. The mass assembly, for example, is often preceded by the massed march, and property destruction and looting often occur together. Indeed, whether a crowd episode is constitutive of a protest demonstration, a celebration, or a riot depends, in part, on the particular configuration of task activities and, in part, on who or what is the object of protest, celebration, or violence. Both of these points can be illustrated with riots.

It is generally agreed that riots involve some level of collective violence against persons or property, but that not all incidents of collective violence are equally likely to be labeled riots. Collective violence against the state or its social control agents is more likely to be labeled riotous, for example, than violence perpetrated by the police against protesting demonstrators. Traditionally, what gets defined as a riot involves interpretive discretion, particularly by the state. But even when there is agreement that riots are constituted by some segment of a crowd or gathering engaging in violence against person(s) or property, distinctions are often made between types of riots, as evidenced by Morris Janowitz's (1979) distinction between "communal riots" and "commodity riots," Gary Marx's (1972) distinction between "protest riots" and "issueless riots," and the Walker Report's (1968) reference to "police riots." Communal riots involve religious, ethnic, and/or racial intergroup violence in which the members or property of one group are violently assaulted by members of another group, as occurred in the United States in Miami in 1980 (Porter and Dunn 1984) and in South Central Los Angeles in 1992 (Bergesen and Herman 1998). Commodity or property riots, in contrast, typically involve looting, arson, and vandalism against property, which has generally been posited as one of the defining features of the many urban riots that occurred across major U.S. cities

in the 1960s. But it has been argued that many of these riots, such as those that occurred following the assassination of Martin Luther King, Jr. in 1968, were also protest riots, or at least had elements of protest associated with them (Fogelson 1971). Even though there has been some effort to identify the conditions that lead to the designation of elements of crowd behavior as protest (Turner 1969), it is clear that communal, commodity, and protest riots are overlapping rather than mutually exclusive crowd phenomena and that distinguishing among them therefore involves some interpretive discretion (Turner 1994). The same is true, as well, with the category of issueless riots, such as the sporting victory celebrations that sometimes take on the flavor of property riots among some of the celebrants. Even in the case of police riots, which involve a loss of discipline and control among the ranks, as occurred during the 1968 Democratic National Convention in Chicago, there is often debate as to whether the assaultive behavior of the police is justified. That there is some level of ambiguity and debate associated with categorizing and distinguishing any crowd episode as not only a riot, but as particular kind of riot, is not surprising considering that all crowd episodes share various task activities or elementary forms even when they differ in terms of their defining task activities.

Following from these observations is a final caveat: The task activities associated with any given crowd episode vary in the degree to which they are the focus of attention. Not all are equally attended to by spectators, social control agents, or the media. Consequently, task activities can be classified according to the amount of attention they receive. One that is the major focus of attention and thus provides the phenomenal basis for defining the episode constitutes "the main task activity," whereas those subordinate to the main task activity are "subordinate or side activities." The main task activity is on center stage and typically is the focus of media attention, as illustrated by the extensive media coverage of property vandalism and looting associated with commodity riots. In contrast, the remaining task activities are sideshows, occasioned by and often parasitic to the focal task activity. Examples of subordinate task activity in the case of property or communal riots, or both, include spectating or observing, informal, unofficial attempts at social control, and even the work of the media.

**Variation in Participation Units.** Just as empirical research on crowds has discerned considerable heterogeneity in behavior, so there is corresponding variation in terms of participants. Some are engaged in various task activities, some are observing, and still others are involved in the containment and control of the other participants and their interactions. Indeed, most of the individuals who make up a crowd fall into one of three categories of actors: task performers, spectators or bystanders, and social control agents. Task performers include the individuals performing both main and subordinate tasks. In the case of an antiwar march, for example, the main task performers would include the protesting marchers, with counterdemonstrators, peace marshals, and the press or media constituting the subordinate task performers.

Spectators or bystanders, who constitute the second set of actors relevant to most instances of crowd behavior, have been differentiated into proximal and distal groupings according to proximity to the collective encounter and the nature of their response. Proximal spectators, who are physically co-present and can thus monitor firsthand the activities of task performers, have generally been treated as relatively passive and nonessential elements of crowd behavior. However, research on a series of victory celebrations shows that some spectators do not merely respond passively to the main task performance and accept the activity as given, but can actively influence the character of the activity as well (Snow, Zurcher, and Peters 1981). Accordingly, proximal spectators can vary in terms of whether they are passive or animated and aggressive. "Distal spectators" refer to individuals who take note of episodes of crowd behavior even though they are not physically present during the episodes themselves. Also referred to as "bystander publics" (Turner and Killian 1987), they indirectly monitor an instance of crowd behavior and respond to it, either favorably or unfavorably, by registering their respective views with the media and community officials. Although distal spectators may not affect the course of a single episode of crowd behavior, they can clearly have an impact on the career and character of a series of interconnected crowd episodes.

Social control agents, consisting primarily of police officers and military personnel, constitute



the final set of participants relevant to most instances of crowd behavior. Since they are the guardians of public order, their primary aim with respect to crowds is to maintain that order by controlling crowd behavior both spatially and temporally or by suppressing its occurrence. Given this aim, social control agents can clearly have a significant impact on the course and character of crowd behavior. This is evident in most protest demonstrations and victory celebrations, but it is particularly clear in the case of riots, which are often triggered by overzealous police activity and often involve considerable interpersonal violence perpetrated by the police. The urban riots of the 1960s in the United States illustrate both tendencies: police-citizen scuffles occasioned by traffic citation encounters or arrests sometimes functioned as a triggering event (Kerner 1968); and the vast majority of riot-associated deaths were attributed to social control agents (Bergesen 1980; Kerner 1968). This was not the case, however, with the riots in Miami in 1980 and in South Central Los Angeles in 1992, in which the clear majority of deaths were at the hands of civilians (McPhail 1994; Porter and Dunn 1984). Although these different sets of findings caution against premature generalization regarding the attribution of responsibility for riot-related deaths, they do not belie the important role of social control agents in affecting the course and character of crowd episodes (Della Porta and Reiter 1998).

Although there is no consensual taxonomy of crowd behaviors and interacting participation units, the foregoing observations indicate that behavioral and participant heterogeneity are characteristic features of most crowd episodes. In turn, the research on which these observations are based lays to rest the traditional image of crowds as monolithic entities composed of like-minded people engaged in undifferentiated behavior.

### CONDITIONS OF EMERGENCE

Under what conditions do individuals come together collectively to engage in crowd task activities constitutive of protest or celebration, and why do these occurrences sometimes turn violent or riotous? Three sets of interacting conditions are discernible in the literature: (1) conditions of conduciveness; (2) precipitating ambiguities or grievances; and (3) conditions for mobilization.

**Conditions of Conduciveness.** The concept of conduciveness directs attention to structural and cultural factors that make crowd behavior physically and socially possible (Smelser 1963). Conditions of conduciveness constitute the raw material for crowd behavior and include three sets of factors: ecological, technological, and social control. Ecological factors affect the arrangement and distribution of people in space so as to facilitate interaction and communication. One such factor found to be particularly conducive is the existence of large, spatially concentrated populations. The vast majority of campus protest demonstrations in the 1960s occurred on large universities, for example. Similarly, the urban riots of the 1960s typically occurred in densely populated residential areas, where there were large, easily mobilizable black populations. Seymour Spilerman's (1976) aggregate-level research on the occurrence of these riots found an almost linear relationship between the size of a city's black population and the likelihood and number of disorders experienced, thus suggesting that there was a threshold population size below which riots were unlikely. More recent analyses of the 1960 riots have found that "the propensity to riot was a function of far more than simply the number of Blacks available for rioting in a particular city" (Myers 1997, p. 108; Olzak, Shanahan, and McEneaney 1996), but such findings do not suggest that concentrated population density and the prospect of rioting are unrelated. Thus, these findings, in conjunction with the earlier observations, provide support for the hypothesis that, all other things being equal, the greater the population density, the greater the probability of crowd behavior.

The heightened prospect of interpersonal interaction and communication associated with population concentration can also be facilitated by the diffusion of communication technology, namely telephone, radio, television, and Internet access. But neither the diffusion of such technology nor population density guarantee the emergence of crowd behavior in the absence of a system of social control that allows for freedom of assembly and speech. It has been found repeatedly that incidence of public protest against the state diminishes considerably in political systems that prohibit and deal harshly with such crowd behavior, whereas the development of a more permissive attitude

toward public protest and a corresponding relaxation of measures of social control is frequently conducive to the development of protest behavior. Two concrete examples of this political-opportunity principle include the proliferation of public protest throughout Eastern Europe in 1989 and 1990 following the break-up of the Soviet Union, and prison riots, which research reveals are sparked in part by the erosion of prison security systems and the increased physical vulnerability of those systems (Useem and Kimball 1989).

**Precipitating Events and Conditions.** However conducive conditions might be for crowd behavior, it will not occur in the absence of some precipitating event or condition. Although the specific precipitants underlying the emergence of crowd behavior may be quite varied, most are variants of two generic conditions: (1) ambiguity; and (2) grievances against corporate entities, typically the state or some governmental, administrative unit, or against communal or status groups such as ethnic, racial, and religious groups.

Ambiguity is generated by the disruption or breakdown of everyday routines or expectancies, and has long been linked theoretically to the emergence of crowd behavior (Johnson and Feinberg 1990; Turner and Killian 1987). Evidence of its empirical linkage to the emergence of crowd behavior is abundant, as with the materialization of crowds of onlookers at the scene of accidents and fires; the celebrations that sometimes follow high-stakes, unanticipated athletic victories; the collective revelry sometimes associated with the disruption of interdependent networks of institutionalized roles, as in the case of power blackouts and police strikes; and prison riots that frequently follow on the heels of unanticipated change in administrative personnel, procedures, and control.

The existence of grievances against the state or some governmental, administrative unit, or against communal or status groups, can be equally facilitative of crowd behavior, particularly of the protest variety. Grievances against the state or other corporate actors are typically associated with the existence of economic and political inequities that are perceived as unjust or political decisions and policies that are seen as morally bankrupt or advantaging some interests to the exclusion of others. Examples of protest crowds triggered in

part by such grievances include the hostile gatherings of hungry citizens and displaced workers in industrializing Europe; the striking crowds of workers associated with the labor movement; and the mass demonstrations (marching, rallying, picketing, vigiling) associated with the civil rights, student, antiwar, and women's movements of the 1960s and 1970s.

Grievances against communal or status groups appear to occur most often in a context of competition and conflict between two or more ethnic or racial groups. A second generation of quantitative research on urban racial rioting in the United States has shown it to be associated with increasing intergroup competition sparked by patterns of hypersegregation of blacks (Olzak, Shanahan, and McEneaney 1996), heightened nonwhite unemployment (Myers 1997), and rapid ethnic succession (Bergesen and Herman 1998). Indeed, if there is a single structural-based source of grievance associated with intergroup rioting throughout much of modern history, it is probably intergroup competition triggered by ethnic/racial displacement and succession.

Crowd violence—"riotous" task activities such as property destruction, looting, fighting, and sniping—has been an occasional corollary of protest crowds, but it is not peculiar to such crowds. Moreover, the occurrence of crowd violence, whether in association with protest demonstrations or celebrations, is relatively infrequent in comparison to other crowd behaviors (Eisinger 1973; Gamson 1990; Lewis 1982). When it does occur, however, there are two discernible tendencies: interpersonal violence most often results from the dynamic interaction of protestors and police (Kritzer 1977; MacCannell 1973); and property violence, as in the case of riot looting, often tends to be more selective and semi-organized than random and chaotic (Berk and Aldrich 1972; Quarantelli and Dynes 1968; Tierney 1994).

**Conditions for Mobilization.** A precipitating condition coupled with a high degree of conduciveness is rarely sufficient to produce an instance of crowd behavior. In addition, people have to be assembled or brought into contact with one another, and attention must be focused on the accomplishment of some task. On some occasions in everyday life the condition of assemblage is already met, as in the case of the pedestrian crowd

and conventional audience. More often than not, however, protest crowds, large-scale victory celebrations, and riots do not grow out of conventional gatherings but require the rapid convergence of people in time and space. McPhail and David Miller (1973) found this assembling process to be contingent on the receipt of assembling instructions; ready access, either by foot or by other transportation, to the scene of the action; schedule congruence; and relatively free or discretionary time. It can also be facilitated by lifestyle circumstances and social networks. Again, the ghetto riots of the 1960s are a case in point. They typically occurred on weekday evenings or weekends in the summer, times when people were at home, were more readily available to receive instructions, and had ample discretionary time (Kerner 1968).

The focusing of attention typically occurs through some “keynoting” or “framing” process whereby the interpretive gesture or utterance of one or more individuals provides a resonant account or stimulus for action. It can occur spontaneously, as when someone yells “Cops!” or “Fire!”; it can be an unintended consequence of media broadcasts; or it can be the product of prior planning, thus implying the operation of a social movement.

### COORDINATION OF CROWD BEHAVIOR

Examination of protest demonstrations, celebratory crowds, and riots reveals in each case that the behaviors in question are patterned and collective rather than random and individualistic. Identification of the sources of coordination has thus been one of the central tasks confronting students of crowd behavior.

Earlier theorists attributed the coordination either to the rapid spread of emotional states and behavior in a contagion-like manner due to the presumably heightened suggestibility of crowd members (LeBon [1895] 1960; Blumer 1951) or to the convergence of individuals who are predisposed to behave in a similar fashion because of common dispositions or background characteristics (Allport 1924; Dollard et al. 1939). Both views are empirically off the mark. They assume a uniformity of action that glosses the existence of various categories of actors, variation in their behaviors, ongoing interaction among them, and the

role this interaction plays in determining the direction and character of crowd behavior. These oversights are primarily due to the perceptual trap of taking the behaviors of the most conspicuous element of the episode—the main task performers—as typifying all categories of actors, thus giving rise to the previously mentioned “illusion of unanimity” (Turner and Killian 1987).

A more modern variant of the convergence argument attributes coordination to a rational calculus in which individuals reach parallel assessments regarding the benefits of engaging in a particular task activity (Berk 1974; Granovetter 1978). Blending elements of this logic with strands of theorizing seemingly borrowed from LeBon and Freud, James Coleman (1990) argues that crowd behavior occurs when individuals make a unilateral transfer of control over their actions. Such accounts are no less troublesome than the earlier ones in that they remain highly individualistic and psychological, ignoring the extent to which crowd behavior is the product of collective decision making involving the “framing” and “reframing” of probable costs and benefits and the extent to which this collective decision making frequently has a history involving prior negotiation between various sets of crowd participants.

A sociologically more palatable view holds that crowd behavior is coordinated by definition of the situation that functions in normative fashion by encouraging behavior in accordance with the definition. The collective definition may be situationally emergent (Turner and Killian 1987) or preestablished by prior policing strategies or negotiation among the relevant sets of actors (Della Porta and Reiter 1998; Snow and Anderson 1985). When one or more sets of actors cease to adjust their behaviors to this normative understanding, violence is more likely, especially if the police seek to reestablish normative control, and the episode is likely to be labeled as riotous or mob-like.

Today it is generally conceded that most instances of crowd behavior are normatively regulated, but the dynamics underlying the emergence of such regulations are still not well understood empirically. Consequently, there is growing research interest in detailing the interactional dynamics underlying the process by which coordinating understandings emerge and change. Distinctive to

this research is the view that social interaction among relevant sets of actors, rather than the background characteristics and cognitive states of individuals, holds the key to understanding the course and character of crowd behavior (Snow and Anderson 1985; Turner 1994; Waddington, Jones, and Critcher 1989).

### THE CORRELATES AND PREDICTORS OF PARTICIPATION

Crowd phenomena associated with collective behavior have been distinguished from everyday, conventionalized crowds, the characteristic features of crowds have been elaborated, the major sets of conditions that facilitate the emergence or occurrence of crowds and riots have been identified, and the issue of behavioral coordination in crowd contexts has been explored. In addressing these orienting issues, only passing reference has been made to factors that make some individuals more likely than other individuals to participate in crowd episodes. For example, it is clear that the odds of participating in some crowd episodes are greater with increasing spatial proximity and access to those episodes, schedule congruence, and discretionary time (McPhail and Miller 1973). As well, individuals whose daily routines and expectancies have been rendered ambiguous or who share grievances that are linked to the occurrence of a crowd episode would appear to be more likely candidates for participation. But both of these sets of conditions typically hold for a far greater number of individuals than those who end up participating in a crowd episode in some capacity other than a social control agent or media representative. So what can be said about the personal and interpersonal correlates or predictors of participation?

There is no simple answer to this question or standard formula for predicting crowd participation. However, research on this question suggests at least four general, sensitizing observations, particularly with respect to participation in protest crowds and riots. The first general observation is that commonsensical psychological indicators of protest and riot participation, such as intense frustration or strong feelings of deprivation, have not been found to be valid or reliable predictors. For example, studies of individual riot participation,

which have been numerous, have failed to find consistently significant empirical correlations between measures of frustration or deprivation and participation (McPhail 1994). This is not to suggest that riot or protest participants may not sometimes feel deeply frustrated or deprived as compared to others, but that these psychological states do not reliably explain their participation or typically differentiate them from nonparticipants. Such findings are consistent with the second general observation: There are a diversity of motivations for participation in crowd episodes, ranging from curiosity to exploitation of the situation for personal gain (e.g., fun, material goods) to sympathy with the issue for which the episode is a marker or carrier to embracement of and identification with the cause from either a self-interested or altruistic standpoint (Turner and Killian 1987). That neither a distinctive psychological state or deficit nor a dominant motive have been found to be associated with crowd and riot participation does not mean that psychological or personality factors are without relevance to this issue. To the contrary, one such factor that appears to be consistently associated with participation as a main task performer in protest crowds and riots is the existence of a sense of "personal efficacy"—the belief that one's participation will make a difference, the confidence that one's efforts will contribute to the larger cause (Snow and Oliver 1995). This finding, which constitutes the third general observation regarding participation correlates, makes good sense when considered in conjunction with the fourth general observation: Participants in crowd episodes—whether they are victory celebrations, protest events, or riots—seldom participate alone. Instead, rather than being isolates or loners, they are typically in the company of friends or acquaintances; they are, in other words, part of a social network. Additionally, recruitment into many crowd episodes occurs through the very same social networks (Snow and Oliver 1995). Thus, participation in crowd episodes, particularly planned ones such as protest events, tends to be embedded in social networks, which also function to nurture a greater sense of both personal and collective efficacy. When these factors are coupled with the previously mentioned conditions for assemblage, and either ambiguity or target-specific grievances, participation becomes more likely and perhaps even more predictable.

## REFERENCES

- Allport, Floyd H. 1924 *Social Psychology*. Boston: Houghton Mifflin.
- Bergesen, Albert 1980 "Official Violence During the Watts, Newark, and Detroit Race Riots of the 1960s." In Pat Lauderdale, ed., *A Political Analysis of Deviance*. Minneapolis: University of Minnesota Press.
- , and Max Herman 1998 "Immigration, Race, and Riot: The 1992 Los Angeles Uprising." *American Sociological Review* 63:39–54.
- Berk, Richard 1974 "A Gaming Approach to Crowd Behavior." *American Sociological Review* 39:355–373.
- , and Howard Aldrich 1972 "Patterns of Vandalism During Civil Disorders as an Indicator of Selection of Targets." *American Sociological Review* 37:533–547.
- Blumer, Herbert 1951 "Collective Behavior." In Alfred McClung Lee, ed., *Principles of Sociology*. New York: Barnes and Noble.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Della Porta, Donatella, and Herbert Reiter (eds.) 1998 *Policing Protest: The Control of Mass Demonstrations in Western Democracies*. Minneapolis: University of Minnesota Press.
- Dollard, John, Leonard Doob, Neal Miller, Herbert Mowrer, and Robert Sears 1939 *Frustration and Aggression*. New Haven, Conn.: Yale University Press.
- Eisinger, Peter 1973 "The Conditions of Protest Behavior in American Cities." *American Political Science Review* 67:11–28.
- Fogelson, Robert M. 1971 *Violence as Protest: A Study of Riots and Ghettos*. New York: Doubleday.
- Freud, Sigmund 1921 *Group Psychology and Analysis of the Ego*. London: International Psychoanalytical Press.
- Gamson, William 1990 *The Strategy of Social Protest*, 2nd ed., rev. Homewood, Ill.: Dorsey.
- Granovetter, Mark 1978 "Threshold Models of Collective Behavior." *American Journal of Sociology* 83:1420–1443.
- Janowitz, Morris 1979 "Collective Racial Violence: A Contemporary History." In Hugh Davis Graham and Ted Gurr, eds., *Violence in America: Historical and Comparative Perspectives*, rev. Beverly Hills, Calif.: Sage.
- Johnson, Norris R., and William E. Feinberg 1990 "Ambiguity and Crowds: Results From A Computer Simulation Model." *Research in Social Movements, Conflict and Change* 12:35–66.
- Kerner, Otto 1968 *Report of the National Advisory Commission on Civil Disorders*. New York: E.P. Dutton.
- Kritzer, Herbert M. 1977 "Political Protest and Political Violence: A Nonrecursive Causal Model." *Social Forces* 55:630–640.
- LeBon, Gustave (1895) 1960 *The Crowd: A Study of the Popular Mind*. New York: Viking Press.
- Lewis, Jerry 1982 "Fan Violence: An American Social Problem." *Research on Social Problems and Public Policy* 2:175–206.
- Lofland, John 1981 "Collective Behavior: The Elementary Forms." In Morris Rosenberg and Ralph Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- MacCannell, Dean 1973 "Nonviolent Action as Theater." Nonviolent Action Research Project Series No. 10, Haverford College Center for Nonviolent Conflict Resolution, Haverford, Pa.
- Marx, Gary 1972 "Issueless Riots." In James F. Short and Marvin E. Wolfgang, eds., *Collective Violence*. Chicago: Aldine-Atherton.
- McPhail, Clark 1994 "The Dark Side of Purpose: Individual and Collective Violence in Riots." *Sociological Quarterly* 35:1–32.
- 1991 *The Myth of the Madding Crowd*. New York: Aldine de Gruyter.
- , and David L. Miller 1973 "The Assembling Process: A Theoretical and Empirical Examination." *American Sociological Review* 38:721–735.
- McPhail, Clark, and Ronald T. Wohlstein 1983 "Individual and Collective Behaviors Within Gatherings, Demonstrations, and Riots." *Annual Review of Sociology* 9:579–600.
- Myers, Daniel J. 1997 "Racial Rioting in the 1960s: An Event History Analysis of Local Conditions." *American Sociological Review* 62:94–112.
- Olzak, Susan, Suzanne Shanahan, and Elizabeth H. McEaney 1996 "Poverty, Segregation, and Race Riots: 1960–1993." *American Sociological Review* 61:590–613.
- Porter, Bruce, and Marvin Dunn 1984 *The Miami Riot of 1980: Crossing the Bounds*. Lexington, Mass.: Lexington Books.
- Ouarantelli, Enrico L., and Russell Dynes 1968 "Looting in Civil Disorders: An Index of Social Change." In Louis Masotti and Don R. Bowen, eds., *Riots and Rebellion*. Beverly Hills, Calif.: Sage.
- Smelser, Neil 1963 *Theory of Collective Behavior*. New York: Free Press.
- Snow, David A., and Leon Anderson 1985 "Field Methods and Conceptual Advances in Crowd Research." Paper presented at Conference on Research Methods in Collective Behavior and Social Movements,

Collective Behavior/Social Movements Section of the American Sociological Association, Bowling Green State University.

- Snow David A., and Pamela E. Oliver 1995 "Social Movements and Collective Behavior: Social Psychological Dimensions and Considerations." In Karen S. Cook, Gary Alan Fine, James S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Snow, David A., Louis A. Zurcher, and Robert Peters 1981 "Victory Celebrations as Theater: A Dramaturgical Approach to Crowd Behavior." *Symbolic Interaction* 4:21-42.
- Spilerman, Seymour 1976 "Structural Characteristics of Cities and the Severity of Racial Disorders." *American Sociological Review* 35:627-649.
- Tierney, Kathleen J. 1994 "Property Damage and Violence: A Collective Behavior Analysis. In Mark Baldassare, ed., *The Los Angeles Riots: Lessons for the Urban Future*. Boulder, Colo.: Westwood Press.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.
- Turner, Ralph H. 1994 "Race Riots Past and Present: A Culture-Collective Behavior Approach." *Symbolic Interaction* 17:309-324.
- 1969 "The Public Perception of Protest." *American Sociological Review* 34:815-831.
- , and Lewis Killian (1957, 1972) 1987 *Collective Behavior*, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Useem, Bert, and Peter Kimball 1989 *States of Siege: U.S. Prison Riots 1971-1986*. New York: Oxford University Press.
- Waddington, David, Karen Jones, and Chas Critcher 1989 *Flashpoints: Studies in Public Disorder*. London: Routledge.
- Walker, Daniel (ed.) 1968 *Rights in Conflict*. New York: Grossett and Dunlop.
- Weller, Jack, and Enrico L. Quarantelli 1973 "Neglected Characteristics of Collective Behavior." *American Journal of Sociology* 79:665-685.
- Wright, Sam 1978 *Crowds and Riots: A Study in Social Organization*. Beverly Hills, Calif.: Sage.

DAVID A. SNOW  
RONNELLE PAULSEN

## CULTS

See Religious Organizations; Religious Orientations; Religious Movements.

## CULTURE

To produce a definition of culture, one can examine the concept in the abstract, that is, explore the concept theoretically from a variety of standpoints and then justify the definition that emerges through deductive logic. Or one can explore how the concept is used in practice, that is, describe how sociologists, both individually and collectively, define culture in the research process and analyze how they inductively construct a shared definition. This essay takes the latter collective-inductive approach to defining culture. Such an approach is inherently sociological and does not presume to produce an independent definition for the field, rather it seeks to document how successful participants in the field have been in producing a shared definition for themselves. To produce such a "working" definition of culture, one starts by examining the social science roots that have helped determine the current status of the sociology of culture.

The focus on culture in sociology has flourished over the past twenty years, as evidenced by the fact that the Culture Section in the American Sociological Association has become one of the largest and is still one of the fastest-growing sections in the discipline. The growth of interest in culture is also nicely documented by the number of survey review articles and books written during this period (e.g., Denzin 1996; Crane 1994, 1992; Hall and Neitz 1993; Munch and Smelser 1992; Peterson 1990, 1989, 1979; Alexander and Seidman 1990; Wuthnow and Witten 1988; Blau 1988; Mukerji and Schudson 1986). As is clear from the reviews, interest in cultural analysis has grown significantly. The focus on culture in all spheres of research has increased tremendously; and culture is now readily accepted as a level of explanation in its own right. Even in traditionally materialist-oriented research arenas, such as stratification and Marxist studies, cultural activities and interests are not treated as subordinate to economic explanations in current research (e.g., Halle 1994; Nelson and Grossberg 1988; Bourdieu 1984; Williams 1981, 1977). Cultural studies and analysis have become one of the most fertile areas in sociology.

The rapid growth in the focus on culture and cultural explanation has produced some definitional boundary problems. The term *culture* has been

used in contemporary sociological research to describe everything from elite artistic activities (Becker 1982) to the values, styles, and ideology of day-to-day conduct (Swidler 1986). Along with art and everyday conduct, included among the “mixed bag” of research that takes place under the auspices of the sociology of culture is work in science (Latour 1987; Star 1989), religion (Neitz 1987), law (Katz 1988), media (Schudson 1978; Gitlin 1985; Tuchman 1978), popular culture (Peterson 1997; Weinstein 1991; Chambers 1986), and work organization (Fine 1996; Lincoln and Kalleberg 1990).

With such an extensive variety in the empirical focus of research in culture, the question for many participants in the field is how to translate this eclecticism into a coherent research field. This goal has not yet been reached, but while a coherent concept of culture is still evolving and the boundaries of the current field of sociology of culture are still fluid and expanding, it is possible to explore how different types of researchers in the social sciences, both currently and historically, have approached the concept of culture. In this inventory process, a better understanding of the concept of culture will emerge, that is, what different researchers believe the concept of culture includes, what the concept excludes, and how the distinction between categories has been made. This essay will provide a historical overview of the two major debates on the appropriate focus and limitations of the definition of culture, and then turn to the contemporary social context in an effort to clarify the issues underlying the current concept of culture.

### THE CULTURE–SOCIAL STRUCTURE DEBATE

From the turn of the century until the 1950s, the definition of culture was embroiled in a dialogue that sought to distinguish the concepts of culture and social structure. This distinction was a major bone of contention among social scientists, most noticeably among anthropologists divided between the cultural and social traditions of anthropology. Researchers in the cultural or ethnological tradition, such as Franz Boas (1896/1940), Bronislaw Malinowski (1927, 1931), Margaret Mead (1928, 1935), Alfred Kroeber (1923/1948, 1952), and Ruth Benedict (1934) believed culture was the central concept in social science. “Culturalists”

maintained that culture is primary in guiding all patterns of behavior, including who interacts with whom, and should therefore be given priority in theories about the organization of society. This position was countered by researchers in the structural tradition, such as A.R. Radcliffe-Brown ([1952] 1961) and E. E. Evans-Pritchard (1937, 1940) from the British school of social anthropology, and Claude Levi-Strauss ([1953] 1963) in French structuralism. “Structuralists” contended that social structure was the primary focus of social science and should be given priority in theories about society because social structure (e.g., kinship) determines patterns of social interaction and thought. Both schools had influential and large numbers of adherents.

The culturalists took a holistic approach to the concept of culture. Stemming from Edward Tylor’s classic definition, culture was “. . . that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits acquired by man as a member of society” ([1871] 1924, p.1). This definition leaves little out, but the orientation of the late nineteenth century intended the concept of culture to be as inclusive as possible. Culture is what distinguishes man as a species from other species. Therefore culture consists of all that is produced by human collectivities, that is, all of social life. The focus here stems from the “nature” vs. “nurture” disputes common during this period. Anything that differentiates man’s accomplishments from biological and evolutionary origins was relevant to the concept of culture. That includes religion as well as kinship structures, language as well as nation-states.

Following Boas, the study of culture was used to examine different types of society. All societies have cultures, and variations in cultural patterns helped further the argument that culture, not nature, played the most significant role in governing human behavior. In addition, the cultural variances observed in different societies helped break down the nineteenth-century anthropological notion of “the psychic unity of mankind, the unity of human history, and the unity of culture” (Singer 1968, p. 527). The pluralistic and relativistic approaches to culture that followed emphasized a more limited, localized conception. Culture was what produced a distinctive identity for a society,

socializing members for greater internal homogeneity and identifying outsiders. Culture is thus treated as differentiating concept, providing recognition factors for internal cohesion and external discrimination.

Although this tradition of ethnographic research on culture tended to be internal and localized, what is termed an “emic” approach in cognitive anthropology (Goodenough 1956), by the 1940s there emerged a strong desire among many anthropologists to develop a comparative “etic” approach to culture, that is, construct a generalized theory of cultural patterns. In the comparison of hundreds of ethnographies written in this period, A.L. Kroeber and Clyde Kluckhohn sought to build such a general definition of culture. They wrote,

*Culture consists of patterns, explicit and implicit, of and for behavior acquired and transmitted by symbols, constituting the distinctive achievement of human groups, including their embodiments in artifacts; the essential core of culture consists of traditional (i.e., historically derived and selected) ideas and especially their attached values; culture systems may, on the one hand, be considered as products of action, on the other as conditioning elements of further action ([1952] 1963, p. 181).*

Milton Singer (1968) characterized this “pattern theory” definition as a condensation of what most American anthropologists in the 1940s and 1950s called culture. It includes behavior, cultural objects, and cognitive predispositions as part of the concept, thus emphasizing that culture is both a product of social action and a process that guides future action. The pattern theory stated simply that behavior follows a relatively stable routine, from the simplest levels of custom in dress and diet to more complex levels of organization in political, economic, and religious life. The persistence of specific patterns is variable in different arenas and different societies, but larger configurations tend to be more stable, changing incrementally unless redirected by external forces. In addition, the theory emphasized that the culture from any given society can be formally described, that is, it can be placed in formal categories representing different spheres of social life to facilitate comparison between societies. As such, universal patterns of culture can be constructed.

In comparison, anthropological structuralists in this period conceive of culture less comprehensively. The structuralists’ concept of culture is made distinct through emphasis on a new concept of social structure. Largely through the efforts of Radcliffe-Brown, a theory emerged that argues social structure is more appropriately represented by a network or system of social relations than a set of norms. The structuralist argument is intended to clarify how actors in a society actively produce and are socially produced by their cultural context. By distinguishing the actors and interaction in a social system from the behavioral norms, structuralists seek to establish a referent for social structure that is analytically independent of the culture and artifacts produced in that system. The production of culture is thus grounded clearly in an international framework. Norms of interaction are also produced by interacting participants, but the question of causal primacy between culture and social structure can be considered separately. The initial effort here is simply not to reify the origins of culture.

The exact relationship of culture and social structure, however, becomes the central issue of the structuralist/culturalist debate. For example, how to identify the boundaries of a society one is researching is problematic when the society is not an isolate. Structuralists tend to give social relations, that is, the extent of a network, priority in identifying boundaries, while culturalists focus on the extent of particular types of cultural knowledge or practices. Since both elements are obviously operating interdependently, the efforts to disentangle these concepts make little headway. The arguments to establish causal priority for one concept vis-à-vis the other settle into a fairly predictable exchange. Structuralists base their priority claims on the fact that the interaction of actors in a society is empirically preliminary to the development and application of cultural elements. Culturalists respond that interaction itself is at least partially cultural phenomenon, and that in most complex societies cultural patterns have been well established prior to ongoing social relationships.

By the late 1950s, the concept of culture was becoming increasingly important to sociologists. To help resolve the now tired debate over cultural and structural foci and precedence, A.L. Kroeber and Talcott Parsons published a report in the



American Sociological Review titled "The Concepts of Culture and Social System" (1958), which seeks to establish some ground rules for differentiating the two concepts. At least for sociologists, many of whom identify explicitly with the structural-functional theories of the anthropological structuralists, acknowledgement of a separate social system component that delimits the scope of culture is not difficult. More difficult is ascertaining where the appropriate limits for the concept of culture lie within this domain. Kroeber and Parsons suggest restricting the usage of culture to, "transmitted and created content and patterns of values, ideas, and other symbolic-meaningful systems as factors in the shaping of human behavior and the artifacts produced behavior" (1958, p. 583). This definition emphasized the predispositional aspect of a cultural referent, limiting the scope of culture to a cognitive perspective, and concentrates on a carefully worded description of "symbolic-meaningful systems" as the appropriate referent for culture. While no longer the omnibus conception of a traditional, Tylor-derived approach, this type of cultural analysis is still potentially applicable to any realm of social activity.

#### THE HIGH-MASS CULTURE DEBATE

In the 1950s and early 1960s, the concept of culture became enmeshed in a new debate that like the previously documented dialogue has both influential and significant numbers of participants on each side of the dispute. Sociologists, however, are more central to the discussion, pitting those who support a broadly conceived, anthropological interpretation of culture that places both commonplace and elite activities in the same category, against a humanities oriented conception of culture that equates the identification of cultural activity with a value statement. This debate attempts to do two things: to classify different types of cultural activity, and to distinguish a purely descriptive approach to the concept of culture from an axiological approach that defines culture through an evaluative process.

That an axiological approach to culture can be considered legitimate by a "scientific" enterprise is perhaps surprising to contemporary sociologists entrenched in the positivistic interpretation of science, yet a central issue for many sociologists in

this period was how and whether to approach questions of moral values. For example, the critical theorist Leo Lowenthal (1950) characterized this period of social science as "applied ascetism" and stated that the moral or aesthetic evaluation of cultural products and activities is not only sociologically possible, but also should be a useful tool in the sociological analysis of cultural differentiation.

These evaluative questions certainly play a part in the analysis of "mass culture," a term that the critic Dwight McDonald explains is used to identify articles of culture that are produced for mass consumption, "like chewing gum" (McDonald 1953, p. 59). A number of commentators, including both sociologists and humanists, observe the growth of mass culture production in the post-World War II United States with a mixture of distaste and alarm. The concern of McDonald and critics like him is the decline of intrinsic value in cultural artifacts, a decline in quality that stems from, or is at least attributed to, a combination of economic and social factors associated with the growth of capitalism. For example, mass culture critics argue that the unchecked growth of capitalism in the production and distribution phases of culture industries leads to a "massification" of consumption patterns. Formerly localized, highly differentiated, and competitive markets become dominated by a single corporate actor who merges different sectors of the consumer landscape and monopolizes production resources and distribution outlets. Within these giant culture industry organizations the demand for greater efficiency and the vertical integration of production lead to a bureaucratically focused standardization of output. Both processes function to stamp out cultural differences and create greater homogeneity in moral and aesthetic values, all at the lowest common denominator.

Regardless of the causes of the mass culture phenomena, the critics of mass culture believe it to be a potentially revolutionary force that will transform the values of society. One critic states that "mass culture is a dynamic, revolutionary force, breaking down the old barriers of class, tradition, taste, and dissolving all cultural distinctions. It mixes and scrambles everything together, producing what might be called homogenized culture. . . It thus destroys all values, since value judgements imply discrimination" (McDonald 1953, p. 62).

In launching this attack, mass culture opponents see themselves as the saviors of a “true” or “high” culture (e.g., McDonald, Greenberg, Berelson, and Howe; see Rosenberg and White 1957). They argue that the consumption of mass culture undermines the very existence of legitimate high culture, that is, the elite arts and folk cultures. Without the ability to differentiate between increasingly blurred lines of cultural production, the average consumer turns toward mass culture due to its immediate accessibility. Further, simply through its creation, mass culture devalues elite art and folk cultures by borrowing the themes and devices of different cultural traditions and converting them into mechanical, formulaic systems (Greenberg 1946). Thus critics of mass culture argue that it is critical for the health of society to discriminate between types of culture.

Defenders of mass culture, or at least those who feel the attack on mass culture is too extreme, respond that mass culture critics seek to limit the production and appreciation of culture to an elitist minority. They contend that the elitist criticism of culture is ethnocentric and that not only is mass, popular, or public culture more diverse than given credit for (e.g., Lang 1957; Kracuer 1949), but also the benefits of mass cultural participation far outweigh the limitations of a mass media distribution system (White 1956; Seldes 1957). Post-World War II America experienced an economic boom that sent its citizens searching for a variety of new cultural outlets. The increase in cultural participation certainly included what some critics might call “vulgar” activities, but it also included a tremendous increase in audiences for the arts across the board. Essentially mass culture defenders assert that the argument over the legitimacy of mass culture comes down to a matter of ideology, one that positions the elitist minority against the growing democratization of culture.

To extricate themselves from this axiological conundrum, many sociologists of culture retreated from a morally evaluative stance to a normative one. As presented by Gertrude Jaeger and Philip Selznick (1964), the normative sociological approach to culture, while still evaluative, seeks to combine anthropological and humanist conceptions of culture through a diagnostic analysis of cultural experience. The emphasis here is on elaborating the nature of “symbolically meaningful” experience, the same focus for culture that Kroeber

and Parsons (1958) take in their differentiation of culture and social system. To do this, Jager and Selznick adopt a pragmatist perspective (Dewey 1958) that accords symbolic status to cultural objects or events through a social signification process. Interacting individuals create symbols through the communication of meaningful experience, using both denotative and connotative processes. By creating symbols, interacting individuals create culture. Thus the definition of culture becomes: “Culture consists of everything that is produced by, and is capable of sustaining, shared symbolic experience” (Jaeger and Selznick 1964, p. 663). In establishing this sociological definition of culture emphasizing the shared symbolic experience, Jaeger and Selznick also seek to maintain a humanist-oriented capability to distinguish between high and mass culture without marginalizing the focus on high culture. Following Dewey, they argue that the experience of art takes place on a continuum of cultural experience that differs in intensity from ordinary symbolic activities, but has essentially the same basis for the appreciation of meaning. Art or high culture is simply a more “effective” symbol, combining “economy of statement with richness of expression” (Jaeger and Selznick 1964, p. 664). As such, art, like all culture, is identified through the normative evaluation of experience.

In sum, the high culture-mass culture debate shifted the focus on the concept of culture from a question of appropriate scope to a question of appropriate values. From a functionalist point of view, the health of a society’s culture is not simply an issue of what type of values are advocated, but of how culture serves a moral and integrative function. Yet the mass culture critique was often unable to distinguish the cultural values of elite intellectuals from the effect of these values on society. To escape from this ethnocentric quagmire, contemporary sociologists have generally turned away from an evaluative position toward culture.

#### **THE CONTEMPORARY APPROACH TO CULTURE: MAPPING THE TERRAIN**

As mentioned at the beginning of this essay, the contemporary approach to culture is quite eclectic. Despite the elaborate historical lineage of the concept, there is no current, widely accepted, composite resolution for the definition of culture.

Instead, culture is still currently defined through an extensive variety of perspectives, sanctioning a broad, historically validated range of options. While the omnibus definition from the cultural anthropology tradition has been generally relegated to introductory texts, and the elitist attack on mass culture has been largely replaced by an antiethnocentric, relativist position open to a wide spectrum of symbolic arenas and perspectives, many of the elements of these old debates still appear in new cultural analyses.

For example, as categorized by Richard Peterson introducing a review of new studies in cultural analysis at the beginning of the 1990s, culture tends to be used two ways in sociological research; as a “code of conduct embedded in or constitutive of social life,” and as symbolic products of group activity” (Peterson 1990, p. 498). The first perspective is clearly indebted to the traditional cultural anthropology approach and indeed is used to analyze and characterize social units ranging from whole societies (e.g., Cerulo 1995; Bellah et al. 1985) to specific subcultures (e.g., Hebdige 1990, 1979; Willis 1977). Empirical applications using this perspective are also made to geographically dispersed social worlds that organize collective activities (e.g., Lofland 1993 on the peace movement; Fine 1987 on Little League baseball; Latour and Woolgar 1979 on scientific research in biology; Trawick 1988 on scientific research in physics). The second perspective takes the more concrete course of treating culture as specific socially constructed symbols and emphasizes the production and meaning of these specific forms of cultural expression. Most examples of this latter form of cultural research are conducted in substantive arenas collectively known as the “production of culture” (Peterson 1979; Crane 1992), however, the range of empirical focus for this perspective is considerable and includes research in such areas as the moral discourse on the abortion issue (Luker 1984), the politics and aesthetics of artistic evaluation and reception (DeNora 1995; Lang and Lang 1990; Griswold 1986), and the motivational and ideological context of organizational, professional, and work cultures (e.g., Fine 1996; Martin 1992; Katz 1999; Fantasia 1988; Harper 1987; Burawoy 1979).

From the array of activities mentioned above, it is clear that the contemporary concept of culture

in sociology does not exclude any particular empirical forms of activity, except perhaps through an emphasis on shared or collective practices, thus discounting purely individual foci. Since all collective social practices are potentially symbolic and therefore culturally expressive, any collective activity can be reasonably studied under the rubric of the sociology of culture. This “open borders” philosophy has at times made it difficult for participants in the sociology of culture to establish any kind of nomothetic perspective for cultural theory. The vast differentiation and sheer complexity of the expression of culture in various forms of social life resists ready categorization. Instead, participants in the sociology of culture have usually opted for the preliminary step of surveying and mapping the terrain of research in the sociology of culture with the goal of helping to define emerging theoretical perspectives in the field. Two particularly informative efforts are the contributions of John Hall and Mary Jo Neitz (1993) and Diana Crane (1992, 1994).

In *Culture: Sociological Perspectives* (1993), Hall and Neitz provide an excellent overview of the substantive and theoretical directions in which research in the sociology of culture has proliferated. They identify five “analytic frames” (p. 17) through which researchers can focus on particular aspects of culture and that emphasize associated processes of inquiry. The first frame is a focus on “institutional structures”: that is, research on culture specifically linked with social institutions and such issues as the construction of social and personal identity and conventional or moral conduct (e.g., Bellah et al. 1985; Gilligan 1982; Warner 1988). In the second analytic frame, Hall and Neitz describe “cultural history” and the influence of past cultural practices on the present. Research in this area includes a focus on the significance of rituals (e.g., Douglas 1973; Goffman 1968, 1971; Neitz 1987), the effects of rationalization on social processes and cultural consumption (e.g., Foucault 1965; Mukerji 1983; Born 1995), and the creation of mass culture (e.g., Ewen 1976; Schudson 1984). In the third analytic frame, Hall and Neitz focus on “the production and distribution of culture” with a special emphasis on stratification and power issues. Research in this area includes work on the socioeconomic differentiation of cultural strata (e.g., Gans 1974; Bourdieu 1984; Lamont 1992), gender and ethnic cultural differentiation and

their effect on inequality (e.g., Radway 1984; Lamont and Fournier 1992), and the production of culture (e.g., Becker 1982; Gilmore 1987; Hirsch 1972; Coser, Kadushin, and Powell 1982; Faulkner 1983; Crane 1987). The fourth analytic frame, "audience effects," looks at how cultural objects affect the people who consume them and the precise patterns of shared meaning and interpretive ideology that provide a compatible environment for the popular and critical success of particular cultural forms (e.g., Wuthnow 1987; Baxandall 1985; Long 1985). Finally the fifth analytic frame, "meaning and social action," refers to how actors in varied mainstream and subcultural settings use culture to guide behavior and establish social identity. In a range of ethnic, political, and ideological contexts, participants use visible expressive symbols and styles to assert cultural difference and communicate the social and personal significance of cultural objects (e.g., Rushing 1988; Ginsberg 1990; Schwartz 1991; Fine 1987).

These frames serve different purposes. For the nonsociologist or for sociologists from outside the field of culture, they provide a guide to current cultural research and a reasonably accurate descriptive picture of research segmentation within the field. For the sociologist of culture, however, these frames represent not only a "division of labor in sociohistorical inquiry, in the sense that any particular frame seems to generate boundaries. . . (within in the field), as Hall and Neitz claim (1993, p. 19), but a strategy to bring analytic coherence to a field that has experienced remarkable growth and empirical diffusion over a relatively short period. As such, in the future these frames may emerge through collective activity as problem areas within the field of culture that will guide empirical and methodological tendencies within particular research communities and influence theoretical interaction, that is, co-citation among researchers. The precise impact in the field, however, still remains to be seen.

A somewhat different mapping, primarily in terms of theoretical emphasis, is offered by Diana Crane in her book *The Production of Culture* (1992) and through her efforts as editor of *The Sociology of Culture: Emerging Theoretical Perspectives* (1994). Like Hall and Neitz, Crane seeks to help codify research segmentation in the field of culture, but she does not try to accomplish this daunting task simply by producing a comprehensive survey of

current research in the field. Instead, she attempts to give the reader a guide to theoretical issues in the sociology of culture, particularly the place of the concept of culture in the discipline of sociology as a whole, and how the centrality of culture as a variable in mainstream sociological models will determine the significance of future research in the field.

To start, Crane argues that culture has traditionally been regarded as "peripheral" to mainstream concerns in American sociology because of its relationship to classical theory (i.e., Marx, Weber, Durkheim). In comparison to the emphasis by these theorists on social structure, organization, and market forces, cultural elements have been consistently treated as secondary in their impact on peoples' behavior and attitudes, particularly surrounding significant life issues (e.g., economic considerations). One reason for this secondary status may be the difficulty classical and mainstream theorists have in conceptualizing and documenting everyday cultural practices. Crane states, "To American and some British structuralists, culture as a concept lacks a suitably rigorous definition" (Crane 1994, p. 2). And from Archer (1988, p. 1), "the notion of culture remains inordinately vague. . . In every way, 'culture' is the poor relation of 'structure.'" Thus culture, approached as the values, norms, beliefs, and attitudes of a population or subgroup, is treated as "an implicit feature of social life. . ." (Wuthnow and Witten 1988, p. 50-51), difficult to put one's finger on, and therefore difficult to document through specific empirical referents.

But Crane argues that culture in contemporary society is much more than implicit features. She states, "Culture today is expressed and negotiated almost entirely through culture as explicit social constructions or products, in other words, through *recorded culture*, culture that is recorded either in print, film, artifacts or, most recently, electronic media" (Crane 1994, p. 2). Further, contemporary sociologists of culture have tended to focus on this "recorded culture" as the principal empirical referent through which various types of contemporary culture are expressed and thus can easily be explored. Not surprisingly then, the primary direction through which the new sociology of culture has proliferated is in areas like art, science, popular culture, religion, media, technology, and other social worlds where recorded

forms of culture are readily accessible. These culture subfields have become the central substantive foci through which the field as a whole has undertaken to build theoretical coherence.

At the same time outside the boundaries of the field of culture per se, it is also clear from recent research in the 1990s that the concept of culture has gained significant relevance in many mainstream areas of the discipline that have traditionally been dominated by macrostructuralist approaches. For example, in both Ewa Morawska and Willfried Spohn's (1994), and Mabel Berezin's (1994) contributions to Crane's *The Sociology of Culture: Emerging Theoretical Perspectives*, the impact of cultural forces are discussed in a variety of macroinstitutional contexts. Morawska and Spohn's focus on examples from the historical perspective includes research on the effect of ideology in the macrostructural analysis of revolution and social change (e.g., Sewell 1985; Skocpol 1985; Goldstone 1991), issues of working-class consciousness and capitalist development (e.g., Aminzade 1981; Calhoun 1982), and the articulation of new forms of religious and ideological doctrines in a social-institutional context (e.g., Wuthnow 1989; Zaret 1985). Berezin's chapter examines the relationship of culture and politics in macromodels of political development and state formation (e.g., Greenfeld 1992; Mitchell 1991). Additional examples in organizational or economic contexts (e.g., Dobbin 1994; Granovetter 1985) only further emphasize the point, that the expanding application of cultural analysis to mainstream models means that for many sociologists, culture is more an explanatory perspective than a substantive area of study. As such, future limitations on the explanatory potential of cultural analysis in sociology will likely be conceptual, not empirical, and the above research suggests a broadly fertile spectrum of empirical possibilities.

Finally, a significant elaboration of the explanatory potential of cultural analysis has taken place in a field organized largely outside the discipline of sociology. "Cultural studies," identifying a loosely connected, interdisciplinary network of scholars from a wide spectrum of perspectives, including the humanities, the social sciences, the arts, and various status-specific programs (e.g., ethnic studies, feminist studies, gay and lesbian studies), has produced a tremendous number of

new kinds of cultural analyses that have implications for the sociology of culture. The approach to cultural analysis, however, is often radically different, both empirically and theoretically, than that conventionally used by sociologists. Cultural studies approaches range from a cultural text-based analysis that interprets meaning and sources of social influence directly from cultural objects (e.g., Hooks 1994; Giroux 1992; see Fiske 1994), to complex interpretative decodings of narratives around issues such as identity politics (e.g., Trinh 1989; Hall 1992) and postcolonial repression and resistance (e.g., Appadurai 1990; Grossberg et al. 1992). As a consequence, the history and emerging relationship of cultural studies to sociology is rather piecemeal. Indeed, Norman Denzin (1996) characterizes the potential association to be one of "colonization"; that is, "the attempt to locate and place cultural studies on the boundaries and margins of academic, cultural sociology" (Denzin 1996, p. XV). Others see the possibility of more reciprocal exchange with the possibility of a "revitalization" for sociological cultural perspectives (Seidman 1996). Whichever way the relationship develops, it is clear that efforts to rethink the concept of culture, the impact of cultural values, and approaches to cultural analysis that take place outside of sociology and even outside of academia will have an invigorating effect on the sociological conceptualization of culture. These battles (i.e., "culture wars") already have had important consequences for policy and resource allocation in education (e.g., Nolan 1996; Hunter 1991). There is no reason to think that sociology will or should be immune to these external influences.

In sum, there is a new appreciation of the salience of culture as an explanatory perspective in contemporary sociological research. Whether it involves the convention-setting influence of art worlds, the moral authority of organizational cultures, or the facilitation of class privileges through habitus, the concept of culture is used to explain behavior and social structure from a distinct and powerful perspective. The future elaboration of this perspective in sociology looks very promising.

## REFERENCES

- Alexander, Jeffrey C., and Steve Seidman (eds.) 1990 *Culture and Society*. Cambridge, Eng./New York: Cambridge University Press.

- Aminzade, Ronald 1981 *Class, Politics and Early Industrial Capitalism*. Albany: SUNY Press.
- Appadurai, A. 1990 "Disjuncture and Difference in the Global Cultural Economy." *Public Culture* 2:1-24.
- Archer, Margaret 1988 *Culture and Agency*. New York: Cambridge University Press.
- Baxandall, Michael 1985 *Patterns of Intention*. New Haven: Yale University Press.
- Becker, Howard S. 1982 *Art Worlds*. Berkeley: University of California Press.
- Bellah, Robert, Richard Madsen, William Sullivan, Ann Swidler, and Steven Tipton 1985 *Habits of the Heart*. Berkeley: University of California Press.
- Benedict, Ruth 1934 *Patterns of Culture*. Boston: Houghton Mifflin.
- Berezin, Mabel 1994 "Fissured Terrain: Methodological Approaches and Research Styles in Culture and Politics." In Diana Crane, ed., *The Sociology of Culture: Emerging Theoretical Perspectives*. Oxford: Basil Blackwell.
- Blau, Judith 1988 "Study of the Arts: A Reappraisal." *Annual Review of Sociology* 14:269-292.
- Boas, Franz (1896) 1940 "The Limitations of the Comparative Method of Anthropology." Reprinted in Boas, *Race, Language, and Culture*. New York: Macmillan.
- Born, Georgina 1995 *Rationalizing Culture*. Berkeley: University of California Press.
- Bourdieu, Pierre 1984 *Distinction*. Cambridge, Mass.: Harvard University Press.
- Burawoy, Michael 1979 *Manufacturing Consent*. Chicago: University of Chicago Press.
- Calhoun, Craig 1982 *The Question of Class Struggle*. Chicago: University of Chicago Press.
- Cerulo, Karen 1995 *Identity Designs: The Sights and Sounds of a Nation*. New Brunswick, N.J.: Rutgers University Press.
- Chambers, Iain 1986 *Popular Culture*. London: Methuen.
- Coser, Lewis, Charles Kadushin, and Walter Powell 1982. *Books: The Culture and Commerce of Publishing*. New York: Basic Books.
- Crane, Diana (ed.) 1994 *The Sociology of Culture: Emerging Theoretical Perspective*. Oxford: Basil Blackwell.
- 1992 *The Production of Culture*. Newbury Park, Calif.: Sage.
- 1987 *The Transformation of the Avant-Garde*. Chicago: University of Chicago Press.
- Denzin, Norman 1996 *Cultural Studies: A Research Volume*. Greenwich, Conn.: JAI Press.
- DeNora, Tia 1995 *Beethoven and the Construction of Genius*. Berkeley: University of California Press.
- Dewey, John 1958 *Art as Experience*. New York: Capricorn Books.
- Dobbin, Frank 1994 *Forging Industrial Policy: The United States, France, and Britain in the Railway Age*. New York: Cambridge University Press.
- Douglas, Mary 1973 *Natural Symbols: Explorations in Cosmology*. New York: Vintage Books.
- Dubin, Steve 1992 *Arresting Images: Impolitic Art and Uncivil Action*. New York: Routledge.
- Evans-Prichard, E. E. 1940 *The Nuer*. London: Oxford University Press.
- 1937 *Witchcraft, Oracles and Magic Among the Azande*. Oxford: Clarendon Press.
- Ewen, Stuart 1976 *Captains of Consciousness: Advertising and the Social Roots of the Consumer Culture*. New York: McGraw Hill.
- Fantasia, Rick 1988 *Cultures of Solidarity*. Berkeley: University of California Press.
- Faulkner, Robert 1983 *Music on Demand*. New Brunswick, NJ: Transaction.
- Fine, Gary Alan 1996 *Kitchens*. Berkeley: University of California Press.
- 1987 *With the Boys*. Chicago: University of Chicago Press.
- Fiske, John 1994 "Audiencing: Cultural Practices and Cultural Studies." In Norman Denzin and Yvonne Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Foucault, Michel 1965 *Madness and Civilization*. New York: Pantheon.
- Gans, Herbert J. 1974 *Popular Culture and High Culture*. New York: Basic Books.
- Gilligan, Carol 1982 *In a Different Voice*. Cambridge, Mass.: Harvard University Press.
- Gilmore, Samuel 1987 "Coordination and Convention: The Organization of the Concert Art World." *Symbolic Interaction* 10:209-227.
- Ginsberg, Faye 1990 *Contested Lives*. Berkeley: University of California Press.
- Giroux, Henry 1992 "Resisting Difference: Cultural Studies and the Discourse of Critical Pedagogy." In Larry Grossberg et al., eds., *Cultural Studies*. New York: Routledge.
- Gitlin, Todd 1985 *Inside Prime Time*. New York: Pantheon.
- Goffman, Erving 1971 *Relations in Public*. New York: Harper and Row.
- 1968 *Interaction Ritual*. Garden City, N.Y.: Anchor Books.

- Goldstone, Jack 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.
- Goodenough, Ward 1956 "Componential Analysis and the Study of Meaning." *Language* 32:22-37.
- Granovetter, Mark 1985 "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.
- Greenberg, Clement 1946 "Avant-Garde and Kitsch." *The Partisan Reader* 378-389.
- Greenfeld, Liah 1992 *Nationalism: Five Roads to Modernity*. Cambridge, Mass.: Harvard University Press.
- Griswold, Wendy 1986 *Renaissance Revivals*. Chicago: University of Chicago Press.
- Grossberg, Larry, et al. 1992. *Cultural Studies*. New York: Routledge.
- Hall, John, and Mary Jo Neitz 1993 *Culture: Sociological Perspectives*. Englewood Cliffs, N.J.: Prentice-Hall.
- Hall, Stuart 1992 "Cultural Studies and its Theoretical Legacies." In Larry Grossberg, et al., eds., *Cultural Studies*. New York: Routledge.
- Halle, David 1994 *Inside Culture: Class, Culture and Everyday Life in Modern America*. Chicago: University of Chicago Press.
- Harper, Douglas 1987 *Working Knowledge: Skill and Community in a Small Shop*. Chicago: University of Chicago Press.
- Hebdige, Dick 1990 *Cut 'N Mix: Culture, Identity, and Caribbean Music*. London: Routledge.
- 1979 *Subculture: The Meaning of Style*. London: Routledge.
- Hirsch, Paul 1972 "Processing Fads and Fashions: An Organization Set Analysis of the Cultural Industry System." *American Journal of Sociology* 77:639-659.
- Hooks, Bell 1994 *Outlaw Culture: Resisting Representations*. New York: Routledge.
- Hunter, James D. 1991 *Culture Wars: The Struggle to Define America*. New York: Basic Books.
- Jaeger, Gertrude, and Philip Selznick 1964 "A Normative Theory of Culture." *American Sociological Review* 29:653-669.
- Katz, Jack 1988 *Seductions of Crime: Moral and Sensual Attractions of Doing Evil*. New York: Basic Books.
- Katz, Pearl 1999 *The Scalpel's Edge*. Boston: Allyn and Bacon.
- Kracauer, Seigfried 1949 "National Types as Hollywood Presents Them." *Public Opinion Quarterly* 13:53-72.
- Kroeber, Alfred 1952 *The Nature of Culture*. Chicago: University of Chicago Press.
- (1923) 1948 *Anthropology*. New York: Harcourt.
- , and Clyde Kluckhohn (1952) 1963 *Culture: A Critical Review of Concepts and Definitions*. New York: Vintage Books.
- , and Talcott Parsons 1958 "The Concepts of Culture and Social System." *American Sociological Review* 23:582-583.
- Lamont, Michele 1992 *Money, Morals, and Manners*. Chicago: University of Chicago Press.
- , and Marcel Fournier (eds.) 1992 *Cultivating Differences: Symbolic Boundaries and the Making of Inequality*. Chicago: University of Chicago Press.
- Lang, Gladys Engel, and Kurt Lang 1990 *Etched in Memory: The Building and Survival of Artistic Reputation*. Chapel Hill/London: University of North Carolina Press.
- Lang, Kurt 1957 "Mass Appeal and Minority Tastes." In Bernard Rosenberg and David Manning White, eds., *Mass Culture*. New York: Free Press.
- Latour, Bruno 1987 *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, Mass.: Harvard University Press.
- , and Steve Woolgar 1979 *Laboratory Life*. Beverly Hills, Calif.: Sage.
- Levi-Strauss, Claude (1953) 1963 *Structural Anthropology*. New York: Basic Books.
- Lincoln, James, and Arne Kalleberg 1990 *Culture, Control and Commitment: A Study of Work Organization and Work Attitudes in the United States and Japan*. Cambridge, Eng.: Cambridge University Press.
- Lofland, John 1993 *Polite Protesters: The American Peace Movement of the 1980s*. Syracuse, N.Y.: Syracuse University Press.
- Long, Elizabeth 1985 *The American Dream and the Popular Novel*. Boston: Routledge and Kegan Paul.
- Lowenthal, Leo 1950 "Historical Perspectives of Popular Culture." *American Journal of Sociology* 55:323-332.
- Luker, Kristen 1984 *Abortion and the Politics of Motherhood*. Berkeley: University of California Press.
- Malinowski, Bronislaw 1931 "Culture." *Encyclopedia of the Social Sciences* 4:621-646.
- 1927 *Sex and Repression in Savage Society*. London: Routledge.
- Martin, Joanne 1992 *Cultures in Organizations: Three Perspectives*. New York: Oxford University Press.
- McDonald, Dwight 1953 "A Theory of Mass Culture." *Diogenes* 3:1-17.
- Mead, Margaret 1935 *Sex and Temperament in Three Primitive Societies*. New York: Morrow.

- 1928 *Coming of Age in Samoa*. New York: Morrow.
- Mitchell, Timothy 1991 *Colonizing Egypt*. Berkeley: University of California Press.
- Morawska, Ewa, and Willfried Spohn 1994 "'Cultural Pluralism' in Historical Sociology: Recent Theoretical Directions." In Diana Crane, ed., *The Sociology of Culture: Emerging Theoretical Perspectives*. Oxford: Basil Blackwell.
- Mukerji, Chandra 1983 *From Graven Images*. New York: Columbia University Press.
- , and Michael Schudson 1986 "Popular Culture." *Annual Review of Sociology* 12:47–66.
- Munch, Richard, and Neil J. Smelser (eds.) *The Theory of Culture*. Berkeley: University of California Press.
- Neitz, Mary Jo 1987 *Charisma and Community: A Study of Religious Commitment Within the Charismatic Renewal*. New Brunswick, N.J.: Transaction Books.
- Nelson, Cary, and Lawrence Grossberg (eds.) 1988 *Marxism and the Interpretation of Culture*. Urbana: University of Illinois Press.
- Nolan, James (ed.) 1996 *The American Culture Wars: Current Contests and Future Prospects*. Charlottesville: University of West Virginia Press.
- Peterson, Richard 1997 *Creating Country Music: Fabricating Authenticity*. Chicago: University of Chicago Press.
- 1990 "Symbols and Social Life: The Growth of Culture Studies." *Contemporary Sociology* 19:498–500.
- 1989 "La Sociologique de l'Art et de la Culture aux Etats-Unis." *L'Annuaire de sociologique* 39:153–179.
- 1979 "Revitalizing the Culture Concept." *Annual Review of Sociology* 5:137–166.
- Radcliffe-Brown, A. R. (1952) 1961 *Structure and Function in Primitive Society: Essays and Addresses*. New York: Free Press.
- Radway, Janice 1984 *Reading the Romance*. Chapel Hill: University of North Carolina Press.
- Rosenberg, Bernard, and David Manning White (eds.) 1957 *Mass Culture*. New York: Free Press.
- Rushing, Andrea Benton 1988 "Hair-Raising." *Feminist Studies* 14:325–336.
- Schudson, Michael 1984 *Advertising the Uneasy Persuasion*. New York: Basic Books.
- 1978 *Discovering the News: A Social History of American Newspapers*. New York: Basic Books.
- Schwartz, Barry 1991 "Social Change and Collective Memory: The Democratization of George Washington." *American Sociological Review* 56:221–236.
- Seidman, Steven 1996 "Revitalizing Sociology: The Challenge of Cultural Studies." In Norman Denzin, ed., *Cultural Studies*. Greenwich, Conn.: JAI Press.
- Seldes, Gilbert 1957 "The Public Arts." In Bernard Rosenberg and David Manning White, eds., *Mass Culture*. New York: Free Press.
- Sewell, William 1985 "Ideologies and Social Revolutions: Reflections on the French Case." *The Journal of Modern History* 57:57–85.
- Singer, Milton 1968 "Culture: The Concept of Culture." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- Skocpol, Theda 1985 "Cultural Idioms and Political Ideologies in the Revolutionary Reconstruction of State Power: A Rejoinder to Sewell." *The Journal of Modern History* 57:86–96.
- Star, Susan Leigh 1989 *Regions of the Mind: Brain Research and the Quest for Scientific Certainty*. Stanford, Calif.: Stanford University Press.
- Swidler, Ann 1986 "Culture in Action." *American Sociological Review* 51:273–286.



# D

---

## DATA BANKS AND DEPOSITORIES

### CONCEPT AND HISTORY OF DATA-BANKS

The tradition of thought that gave rise to the social sciences was based on the quest to understand the laws or regularities governing the emerging industrial societies with democratic political regimes. The political and economic revolutions that were shaking the world substituted the relative predictability of the traditional ways of handling power and production with the disconcerting uncertainties of political consensus and the new market of commodities.

Finding such laws was not easy. The model for scientific inquiry established by the successful endeavors in the various fields of physical and natural sciences was not applicable to social science research. For a period social scientists believed that the complexity of the social object and the immaturity of the field were responsible for the failure of the natural science model. Social scientists gradually became aware, however, that the epistemological foundations of the social sciences were different: in practical terms, because creating an experimental situation in social matters is extremely difficult; and in theoretical terms, because society is a moving target, readily reacting to changes in circumstances. In the end, mainstream social scientists learned to live with these difficulties.

The Newtonian challenge of formulating hypotheses, collecting the relevant data, and accepting only those hypotheses that fit observational

data, has been in one way or another the stimulus and the standard of advances in knowledge in the many fields that later composed the social sciences. The systematic collection of empirical observations has been the ballast that has kept modern social sciences from drowning in second-rate philosophy or outright ideology. Numerous social researchers and thinkers have stood up to the challenge of providing reliable observations on the social world. (Since every observer is also a member of society, it is not easy to stand aside and look at it from a fixed point of view.)

In retrospect it is understandable that different paths to the common goal of collecting systematic reliable data were tested with various methodological and technical tools, not always being understood as part of the same endeavor. In the latest bout of cultural dominance of Marxist theories, Karl Marx and his school were viewed as “grand theorists” squarely opposed to the “abstract empiricism” of contemporary sociology. But this was a misconception that completely overlooked the many years spent by Marx painstakingly collecting data on industrial society, and by desire to be a scientist like Charles Darwin—to whom he dedicated the book produced by his gigantic research effort, *Das Kapital*.

Despite the apparent confusion and turmoil, there was an underlying paradigm. Social sciences had to be empirical. No matter how radically critical and antipositivistic have been the epistemological conclusions of the various *Methodenstreiten*, the mainstream has resisted the idea of a data-free

social science. Altogether such theoretical orientation has provided this disciplinary area with a fairly resilient ballast against ideological nebulosities. And this can be said safely without underestimating the humongous and repetitive production of irrelevant trivia in tabular or graphic form produced over the years by the low-brows of “abstract empiricism”. Data, however, are a peculiar commodity insofar that they have to be produced. Which means most of the time it is costly, painstaking, and time consuming to produce data. By and large, data belong to one of two families. *Primary data* are collected by the researcher himself. *Secondary data* are collected by other agencies, mainly public and private large-scale organizations, in which case the researcher can only perform what is known as *secondary analysis*, and he obviously has no control over the collection of this data. In the latter case it is also useful to further distinguish between data collected for analytical purposes by agencies like the various census bureaus, and those collected for administrative purposes, such as data on health collected by hospitals, or data on education collected by educational institutions, or mortality data collected by governments and other agencies. These are *process-produced data*; data created as by-products of administrative activities.

There are disciplinary areas that have a high degree of institutional stability and control over their data such as physics, but also engineering, medicine or law. The latter two fields, of course, provided the original kernel for the development of universities and *studia generalia*, in Bologna, Padova, or Salerno. Other fields are less stable, usually in the area of the humanities and the social sciences. But in all cases the organized character of the type of knowledge provided by academic disciplines is predominant.

*Organized knowledge* is the practical rather than speculative knowledge accumulated by governments and their administrative apparatuses, by corporations, political parties, trade unions, churches, and other institutions. Thomas Kuhn has been the leading figure in analyzing innovation processes in the institutions of *organized knowledge* (Kuhn 1962). These types of secondary data are essential for social scientists, and in fact, as noticed long ago by Paul F. Lazarsfeld and Stein Rokkan, among others (Rokkan 1976), the work of sociologists,

especially in Europe, is largely based on interactions with the institutions in charge of this type of knowledge.

*Le Suicide* by Emile Durkheim is an excellent example of the advantages and drawbacks of secondary analysis of process-produced data. This masterpiece of sociological research gave empirical evidence of the theoretical tenet that a cogent collective agent can influence individual behavior even to the extreme gesture of annihilating the genetic commandment. It could not have been written without access and a hard sweated one at that, to secondary data. No individual scientific actor could collect data on events that are so numerous, and so highly dispersed in time and space. On the other hand, relying on data collected by other agencies means that the researcher relies on somebody else’s definition of events. One of the most damaging critiques to Durkheim’s work is that his sophisticated theoretical definition of suicides, is nullified by the fact that the cases of suicides in the data used were defined by officers or judges according to completely different, and uncontrollably variable, criteria.

The work of historians, too, would probably not be possible without access to the *organized knowledge* embodied in the archives of agencies of all kind. The early development of economics as a quantitative discipline was greatly favored by the availability of organized knowledge collected by public administrations. The same applies to demography, which could base itself also on data accumulated by the church, and in general to the whole field of statistics, which could today be defined as policy sciences (Cavalli 1972).

The Durkheim syndrome, the need to use data not collected by scholars, but by public employees, is a constant problem for social scientists using information coming from the realm of organizational knowledge. At the beginning of the nineteenth century Melchiorre Gioja expresses the irritation of a scholar dependent on low-quality public data. He took issue with “the many questions that various inept busybodies called secretaries send from the capital to the province. Questions that never produced other than the following three effects:

1. fear that the Government seeks the basis for some aggravation, and therefore opportunistically false answers;

2. ridicule resulting by the silliness, inconsistency, and imprecision of the questions, and thus answers biased by contempt;
3. heaps of paper uselessly encumbering archives, if the government mistrust them, or very serious mistakes if it uses them, not to speak of the time stolen from the municipal or provincial administrators who must prepare the answers." (Gioja 1852, p. 5)

A powerful answer to these problems came from the development of *survey research*, particularly from the 1940s on by American sociologists who developed the "art of asking why"; namely the inquiry into the rational motivations of individual behavior. It is not surprising that these new methods originally developed in two crucial areas of behavior: the choice of candidates in an electoral process, and the choice of a product in consumption behavior. There has been much criticism of mercantile attitudes in voting, and of course it is quite clear that some of the value motivations that are mobilized in choosing a candidate are evidently not the same as those that are mobilized in picking up one granola package rather than another from a shelf. But the critics miss the point. The two areas of behavior are similar, and the inquiry into behavioral motivations should not be diverted by undue considerations of political or ethical correctness. Little wonder that scholars had to develop new data collection tools on this type of behavior. And equally not surprising is that political and economic elites are willing to invest resources in the arduous enterprise of predicting the aggregate outcome of individual behavior. A prudent politician today constantly monitors the opinion of the electorate. And economists make use not only of large-scale models of the behavior of macroeconomic systems, but also of assessments of consumer behavior. Politicians, bureaucrats, entrepreneurs, and managers would have a hard time doing without the tools provided by the social sciences.

Survey data collection methods, however, presented the social scientist and his institutions with novel problems. Pollsters and survey people in general produce huge quantities of individually uninteresting questionnaires. Punched cards were developed to hold data; at first the cards were processed manually as the "McBee cards," but

soon after machine-readable cards were developed, such as those punched with the Hollerith code (universally known as IBM card). Cards were easier to store than questionnaires, but it was easy, in a routine research process, to lose the "codebook" of the research so that in many cases the cards were useless, even if they contained relevant information (Rokkan 1976). Thus the storing, handling, processing, and redistributing of punched cards required specific skills. Furthermore, the traditional institutional structure of universities and research centers is not well adapted to take on these tasks. Originally social scientists turned to libraries to store their data (Lucci and Rokkan 1957), but in the late 1950s and early 1960s libraries were not equipped to handle large masses of data requiring mainframe computers. The big machines were housed in separate structures and tended by IBM technicians. Social scientists had limited access to the mainframe computers and therefore, the data, until the scientists developed their own separate institutions on the model of the Inter University Consortium for Political Research at Ann Arbor, Michigan (ICPR, later turned into ICPSR when social research was added). In Europe the vision and farsightedness of scholars such as Philip Abrams in the United Kingdom, Stein Rokkan in Norway, and Erwin Scheuch in Germany, helped establish the first archives in Essex, the Social Science Research Center, in Bergen, the Norwegian Data Service, and in Koeln, the Zentralarchiv für Empirische Sozialforschung. Later on these archives developed their own organizations, first IFDO, International Federation of Data Organizations, and later on CESSDA, Council of European Social Science Data Archives. Archives developed in Italy at the Istituto Superiore di Sociologia of Milano (ADPSS), in Denmark (DDA), in the Netherlands (the Steinmetz archive), in Belgium (BASS), and in France (BDSP in Grenoble and in several other places).

### SOCIAL SCIENCE DATA ARCHIVES

Cultural and technological changes led to the creation of Social Science Data Archives (SSDA). SSDA are scientific institutions that retrieve, store, and distribute large amounts of data on social science. The oldest and the most important SSDA were established in the United States—where an emphasis on quantitative social research is deeply rooted—when there was growing attention by both

public administration and the scientific community to the use of social indicators as standards for the population welfare level. These social actors turned to social indicators for help in planning, applying, and the evaluating public-assistance programs when it became apparent that using economic indicators alone was insufficient and inappropriate.

The first major contribution of social indicators is the *Recent Social Trends* study, supervised by the sociologist William Ogburn and prescribed by the U.S. Government toward the end of the 1920s (Bauer 1966). In 1946 the *Employment Act* was published, which was a systematic collection of information on economy intended to affect policies and programs that would sustain employment rates. Most of the studies and undertakings of the 1960s described a "great society," which could overcome the widening economic and social gap.

The consciousness of social problems, together with the necessity of endowing a collection of social data, prompted the publication of *Towards a Social Report* at the end of the 1960s (Olson 1974). This publication was intended to be "a first step toward the evolution of a regular system in social reporting;" but still, like other similar and contemporary writings, data were used just as illustrations supplementing the text.

Technological developments that contributed to the establishment of SSDA include previously unseen data collection techniques and new quantitative methods to organize and analyze those same data. Throughout the 1960s improvements in computing technologies, specifically in data gathering and storing, allowed researchers to do previously unthinkable levels of analysis (Deutsch 1970).

The first SSDA were born autonomously, unrestricted by publicly administered archives or by the institutions traditionally related to the collection and promotion of data (libraries, museums, and data archives). The earliest SSDA were created in the United States, where in 1947 the Roper Public Opinion Research Center opened, followed by the Inter-University Consortium for Political and Social Research (ICPSR) in 1962 at the University of Michigan in Ann Arbor. In 1960, the University of Koln in Germany developed the Zentralarchiv für Empirische Sozialforschung (ZA). A few years later, in 1964, the Steinmetzarchief settled in at the

Amsterdam Arts and Sciences Real Academy. The Economic and Social Research Council Data Archive (ESRC-DA) was created at the University of Essex in 1967. These archives specialized in public opinion surveys and social research, but they also focused attention on the great amount of data compiled by statistical bureaus and public agencies (Herichsen 1989).

SSDA developed a culture of data sharing; data exchange and the repeated use of available data for new research projects intensified with the introduction of statistical packages for the social sciences and more compact media for data transfer. As data production and SSDA grew, more systematic acquisition policies were implemented, and transborder cooperation resulted in the exchange of data processing tools and of emerging archiving and service standards (Mochmann 1998).

In the early 1970s SSDA were created: in Norway in 1971, in Denmark in 1973, and in the United States (at the University of Wisconsin in Madison, U.C.L.A., and the University of North Carolina in Chapel Hill). In the 1980s SSDA were created in Sweden (1980), France (1981), Austria (1985), and then Canada, Hungary, Israel, Australia, New Zealand, and Switzerland (1993).

Currently, the most important SSDA is the ICPSR at the University of Michigan, which has over 40,000 data files and gathers information from more than 300 institutions. Besides data retrieval, processing, researching, and transferring, it publishes an annual catalog and a four-monthly bulletin, cooperates in great projects concerning data collection (e.g. the *General Social Surveys* and the *Panel Study on Income Dynamics*), conducts formative training (it organizes a summer school on methodology and statistics for social science), and offers educational activities. The ICPSR has been open to foreign institutions since the early 1970s.

SSDA differ on budget and staff size, functions, amount of data files in archives, and technical characteristics (e.g. type of hardware and software used, online data accessibility) (IFDO 1991). The SSDA network is tied, however, by international associations such as the Council of European Social Science Data Archives (CESSDA was instituted in 1976 to facilitate cooperation between the most important European SSDA), and

the International Federation of Data Organizations (IFDO was founded in 1977 to enhance the cooperation already started by CESSDA).

CESSDA promoted the acquisition, archiving, and distribution of data for social research throughout Europe, facilitated the exchange of data and technology among data organizations, and supported the development of standards for study description schemes to inform users about the archival holdings, classification schemes for access to variables by subject and continuity guides for coherent data collections. In the 1970s and 1980s, European membership grew continuously and CESSDA started to associate and cooperate with other international organizations sharing similar objectives. Today Europe has a good coverage of CESSDA member archives, while planning processes are under way in several more countries that still lack a social science infrastructure.

Recent trends reveal a type of limitation to the growth and proliferation of national SSDA. This occurs partly through the opening of decentralized bureaus and partly through more specialized data file collections—ones with more well-defined inquiry areas. Exemplary regarding this last issue is the *Rural Data Base* of the ESRC, or even the CESSDA internal agreement aiming at a study field subdivision between the different European SSDA (Tannenbaum 1986).

SSDA have been useful not only in their specific field (especially retrieving, storing, gathering, and making available data) but also have improved general quantitative research techniques, secondary analysis, statistical and administrative software packages, and have made available the best hardware in social research. Most of the SSDA publish their own bulletins, organize methodology schools for social science, hold updating seminars, and cooperate in research projects. In the second half of the 1980s SSDA were particularly active in defining the standards for the information system to be used among these institutions. More recently, SSDA have been working on international access practices on two levels: developing standard study description procedures (to define each data file) and promoting access to online archives.

## SOCIAL SURVEY INSTRUMENTS

Major social survey instruments are those institutionalized initiatives that have produced some of the most interesting results from the autonomous research applied to problems of general interest. The *General Social Surveys*, the *Continuous National Surveys*, and the *U.S. National Surveys* are the most interesting examples of longitudinal studies. Eurobarometers and the International Social Survey Program (ISSP) are discussed here because, although not strictly comparable to general social surveys, they are two of the most comprehensive and continuous academic survey programs. Although it is not discussed here, a student of the social sciences should examine the *European Household Panels*. This project integrated national household panels from Germany (since 1984), Sweden (since 1984), Luxembourg (since 1985), France (since 1985), Poland (since 1987), Great Britain (since 1991), and Belgium and Hungary (since 1992). These panels include variables on household composition, employment, earnings, occupational biographies, health, and satisfaction indicators. (In order to create an international comparative database for microdata from these projects, the Panel Comparability Project (PACO) was formed).

**General Social Surveys.** General Social Surveys (GSS) were developed to set out data on demographic, social, and economic characteristics of the population, as well as opinion data on social life (e.g. family, politics, institutions, relationships). This important survey instrument improved over time. In the mid 1940s the Survey Research Center of the University of Michigan conducted panel studies on a national level on specific issues, such as political behavior, socioeconomic status, and consumers' attitudes. In the early 1970s, the National Opinion Research Center (NORC) of the University of Chicago organized the first GSS, which was immediately followed by the Continuous National Surveys program.

Following the consolidation of the research method based on enlarged quantitative surveys on a national level, some institutions committed to the transfer of the results, both within the U.S. academic community and toward universities and research centers worldwide. One should note that data provided by the GSS are given a particular treatment inside the ICPSR, which, besides transferring the data, gathers the sociologists interested

in the study and the arrangement of social indicators. The importance of the GSS in U.S. social research is exemplified by the more than 1,000 publications related to GSS data analysis listed in the catalog prepared by NORC.

The GSS is done through interviews of a representative sample of U.S. residents who are over 18 years old and are not institutionalized. Representativeness is guaranteed by a complex sampling mechanism, which relies on a multilevel selection over metropolitan areas, municipalities, regions, and individuals. Sampling criteria have been modified throughout the course of different GSS, but to keep the results comparable between one edition and the following ones correctors have been added to allow for these adjustments.

The GSS questionnaire consists of standard questions that are asked each time, and sometimes groups of questions on specific themes are included. The topics treated are the ones most interesting in the study of society and its trends, with a particular focus on family, economic and social status, gender and ethnic group relationships, and moral questions. Political behavior and working activity are not included because the former is already studied in detail through surveys organized by the Institute for Social Research of the University of Michigan, and the latter is well described in the research on labor forces arranged by the U.S. Bureau of the Census.

Given that one of the fundamental aims of the GSS is to provide a general view of time trends, not only in population characteristics, but also and overall in opinions, evaluations, and behaviors over the most important topics describing the social scenery, these questions are included only from time to time, so the questionnaire is not weighted excessively. Some questions are included only two years in three, others every 10 years, and some only after particular and significant events. Over 100 sociologists worked on the first draft of the questionnaire (in 1972). These sociologists devised a final, definite version by voting for each single question. Every year the selection of questions is done by an ad hoc committee, selected by American Sociological Association members.

Some questions utilized in the GSS come from the national surveys run before 1972 that were promoted by commercial research institutes (e.g.

Gallup, Harris), university research institutes (especially ICPSR), and also federal commissions organized to study particular phenomena. Data comparability is assured by the question scheme, which is the same of the original study.

**Continuous National Surveys.** The *Continuous National Surveys* (CNS) are national studies (the first one dates back to 1973) conducted monthly, with the aim of supplying the various governmental agencies with the necessary data (e.g. welfare indicators) to schedule social programs. The sample plan consists of persons selected on the basis of their living groups. On this regard, the NORC has carefully prepared a master probability sample of households, that is, a multi-stage sample to collect on a first stage municipalities or else groups of municipalities. From them, all districts or block groups are selected, and then the proper cohabitational groups in which to choose the individuals to interview are selected.

**U.S. National Surveys.** U.S. national surveys on representative samples of U.S. population have been held since 1974 by the Survey Research Center of the University of Michigan. Some of these studies are held regularly (e.g. the *Surveys of Consumer Finances*, the *Survey of Consumer Attitudes and Behavior*, and the *Panel Studies of Income Dynamics*). In particular, the yearly *Panel Study of Income Dynamics* is one of the most interesting surveys on income trends, and, specifically, on the possible causes for changes in the economic status both of households and single individuals.

The representative sample initially extracted consisted of 2,930 households, to which has been added 1,872 households that were already survey subjects by the U.S. Bureau of the Census on the income topic in the two previous years. Each year these households are re-interviewed, and the sample has grown as members of the original sample established new households. For the first time ever, phone interviews were tested in these researches, and since then the telephone has become broadly used in social research and panel studies.

**Eurobarometers.** Eurobarometers are opinion studies that have been held twice a year since 1973 in what are now the European Union countries. These studies sample about 1,000 individuals for each country, which represents the population

over 15 years-old. The aim of these comparative surveys is to learn about the attitudes of European citizens on some broad interest topics. Questions regard public attitudes toward European integration, but sometimes also address specific problems of a single country or more generally economic, political, and social conditions. Two important functions of Eurobarometers are: being cross-national and easily comparable surveys, they helped integrate social research throughout Western Europe; and, they allowed (and still allow) analysis on social changes in Western Europe. As a matter of fact, there is no survey similar to eurobarometers in what concerns a regular check over time (more than 20 years to date) and space (every single European country). Eurobarometers are now at the disposal of the academic community thanks to the ICPSR and the ZA of Koln.

**The International Social Survey Program.** The International Social Survey Program (ISSP) combines a cross-national survey with a longitudinal time dimension by replicating particular question modules, ideally in five year intervals. The first survey on the “role of government” started in 1985 in four countries (the United States, Great Britain, West Germany, and Australia). Since then the ISSP has grown rapidly and now covers more than 30 countries around the world, including Bulgaria, the Czech Republic, Hungary, Latvia, Poland, Russia, and Slovenia. Topics of the ISSP have included social networks (1986), social inequality (1987, 1992), family and changing gender roles (1988, 1994), work orientations (1989, 1997), religion (1991), and environment (1993). Role of government was replicated in 1990 and 1996. The official data archive of the ISSP is the ZA, which makes the integrated data sets available via the archival network.

#### **DEPOSITED, ACCESS PROCEDURES AND USE**

The existence of SSDA in Europe and in the United States has had a positive effect on the scientific community because SSDA allow access to some data that is particularly useful in secondary analysis. In other words, researchers can collect and use data from different surveys (with particular hypothesis and conceptual frames) to support their own works.

SSDA vary on the type and volume of data that is held and delivered and on the exact services offered. The data differ within and between SSDA in terms of subject matter, time period, and geographical area covered. The data may be individual or aggregate, and the variables may be cross-sectional or time series and suitable for comparative or longitudinal analysis, or both. Charging policies differ from one SSDA to the next and depend upon the type of service being provided, the specific data set demanded, and the institutional affiliation of the requester.

Many data formats are used by SSDA to preserve and deliver the data that is increasingly required to be machine readable, and there is also an increasing demand for the archiving and disseminating of machine-readable metadata.

Despite the heterogeneity of SSDA, there are common goals and tasks. Although we know more about the SSDA in Western Europe, we hope to represent all SSDA in the following list of goals and tasks:

- To promote the acquisition, archiving, and distribution of electronic data for social science teaching and research and to exchange data and technology;
- To exploit the potential of the Internet by the expansion of web-based services and the improvement of dissemination to allow users more direct and immediate access to the data;
- To develop and use metadata standards for the management of data;
- To provide users access to comparative data and to multiple data sources across national boundaries;
- To allow users to receive all or subsets of data via download or on portable media in one of a number of formats;
- To extend the users' base beyond traditional boundaries;
- To better serve the needs of an increasing number of novice users, particularly in terms of data analysis software and improved searching aids through the development and use of social science thesauri and user-friendly interfaces.

The technological changes that contributed to the creation of SSDA have had far-reaching significance for social scientists. The changes throughout the 1990s were considerable: from simple writing (word processing) to data production (especially computer assisted interviewing technique); from archiving (from punch cards to floppy disks and CD-ROMs) to data analysis (the creation of various new software); from bibliographic reference (online access to bibliographic databases) to communication (e-mail) (Cooley and Ryan 1985). The boundaries of technological innovations are not yet clear. The compound archive of integrated data, text, images, and sounds (Garvel 1989), and the Geographical Information System, by which we can analyze data concerning space are examples of previously unimagined products that are possible because of computers (Unwin 1991). These instruments require high-level storing proceedings so that data can be a real resource in the development of empirical knowledge.

While social research in earlier decades lamented the lack of reliable data (in the 1960s the social sciences were considered data poor, and infrastructural support for social research was lacking) the situation has changed. Thousands and thousands of data sets, some of them dating back to the 1940s, are stored for secondary analysis in SSDA. These data sets represent a vast potential for comparative research on historical developments and social change.

### PROSPECTS FOR THE FUTURE

In 1994, the European Commission recognized Social Science Data Archives as legitimate institutions to be considered as applicants to the large-scale facility program for scientific research. This is tantamount to recognizing that the social sciences have infrastructure needs equal to those of the "hard" sciences. This goal was originally pursued through the work of a study panel created by the General Directorate for Science and Technology of the European Commission and particularly its Training and Mobility of Researchers program. The study panel met twice, and was entrusted with the tasks of: a) identifying the future priorities of European scientists concerning access to large installations in the social sciences, both in the

short and in the medium term (i.e. in the period 1994–2004); b) suggesting ways to meet these priorities—with an indication of their respective costs and benefits—in light of the present and expected availability of large installations, taking into account existing and planned future international cooperations, both within the community and outside as well as the particular needs of researchers working in regions where such installations do not exist; and c) consulting representatives from member states and reporting on what facilities, if any, might be available for possible support in the frame of successor programs. A large-scale facility has, to date, been defined as a large research installation that is rare or unparalleled in Europe and that is necessary for high-level good-quality research. Such facilities tend to have high initial costs of investment and comparatively high operating costs. The large-scale facilities program is intended to provide scientists with access to large installations within and especially outside their own countries, thereby promoting the mobility of researchers and encouraging the creation of a Europe-wide research community. And likewise, these installations are better utilized when their services, facilities, and knowledge are available to a wider community of users. Thus the beneficiaries of the large-scale facilities program are researchers who are provided with access to the facilities and the organizations that receive support for the use and improvement of their equipment. The level of support for such facilities has, to date, been based on the quality and unique features of the facility and the value—particularly in advanced training—to potential users.

As a sequel to the work and recommendations of the study panel, between 1994 and 1999 three Social Science Data Archives, the SSRC Essex archive, the German ZA, and the NSD in Bergen have been included in the large-scale facilities program together with other social science institutions. In addition the General Directorate for Science and Technology created a round table on large-scale facilities in the social sciences. This round table was constituted by the European Union commission to suggest actions needed to support a given field of scientific activity. The prospects of Social Science Data Archives are thus quite positive, and for the first time they have been given a tool capable of furthering their original aims.



## REFERENCES

- Bauer, R. A. (ed.) 1966 *Social Indicators*. Cambridge, Mass.: MIT Press.
- Cavalli, A. 1972 "La Sociologia e le Altre Scienze Sociali. Prospettive di Integrazione Interdisciplinare." In P. Rossi, ed., *Ricerca Sociologica e Ruolo del Sociologo*. Bologna: Il Mulino.
- Cooley, R. E., and N. S. Ryn 1985 "An Information Technology Strategy for Social Scientists." *University Computing* 7.
- Deutsch, K. 1970 "The Impact of Complex Data Bases on the Social Sciences." In R. L. Bisco, ed., *Data Bases, Computers and the Social Sciences*. New York: Wiley.
- Garvel, S. 1989 "National Archives and Electronic Records: Where Are We Going?" *IASSIST Quarterly* 3-4.
- Gioja, M. 1852 *Filosofia Della Statistica*. Torino.
- Henrichsen, B. 1989 "Data from the Central Bureau of Statistics to the Social Science Community: The Norwegian Experience." *IASSIST Quarterly* fall-winter.
- IFDO 1991 "IFDO Survey of Computerized Catalogues." *IFDO News* September.
- Kuhn, T. 1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lucci, T., and S. Rokkan (eds.) 1957 *A Library Center of Survey Research Data*. New York: Columbia University Press.
- Mochmann, E. 1998 "European Cooperation in Social Science Data Dissemination." Working paper, round table meeting on infrastructures for social economic research. Koln, 29-30 June.
- Rokkan, S. 1976 "Data Services in Western Europe." *American Behavioral Scientist* 19:443-454.
- Tannenbaum, E. 1989 "Sharing Information Begets Information." *IASSIST Quarterly*, Summer.
- Unwin, D. 1991 "Geographical Information Systems and the Social Science Research." *ESRC Data Archive Bulletin* 48.

GUIDO MARTINOTTI  
SONIA STEFANIZZI

## DEATH AND DYING

This essay asks three questions about death and dying: 1) Why should an entry on such phenomena, which are clearly of interdisciplinary interest, appear in an encyclopedia of sociology? 2) What

related topics have been studied by sociology? 3) What issues are currently pending that call for sociological attention?

### DEATH AND DYING AS A FIELD OF SOCIOLOGICAL INQUIRY

The answer to the first question is not readily found in the history of sociological thought, although Victor Marshall once bemoaned the fact that Georg Simmel in 1908 had identified but had not pursued the topic as suitable for sociological inquiry, and a half century later the topic was thought to be a neglected area for sociology (Faunce and Fulton 1958). On other hand, Fulton reminds us that "sociological interest in death is coexistent with the history of sociology" (Fulton and Bendiksen 1994; Fulton and Owen 1988). Both Marshall and Fulton are correct in that the "interest" has typically been peripheral. Herbert Spencer had noted that social progress depended on the separation of the world of the living from the world of the dead, but that was hardly his central theory. Emile Durkheim's *Suicide* depends on an elaborate theory of "anomie," not on any theory of death. Max Weber deals with the fact of death in that it interrupts the pursuit of one's calling—a basic observation later developed by Talcott Parsons. William Graham Sumner wrote widely about such death-related topics as fear of ghosts, mortuary rituals, widowhood, infanticide, war, and even the right to die, but all such references were illustrative of some more general point. In 1956 Herman Feifel chaired a conference on death for the American Psychological Association.

There may well be other precursors, but a bit of history suggests how and when sociology may have staked out its disciplinary claim in a field that had long been cultivated by medicine, theology, ethics, philosophy, law, and psychology.

In May 1967, as part of a new program on "Death Education" at the University of Minnesota, Robert Fulton, a sociologist, arranged perhaps the first interdisciplinary conference on death and dying in the United States. Fulton had just published his ground-breaking book *Death and Identity* (1965), the purpose of which was to help in "preserving rather than losing . . . personal identity . . ." when facing death. It was a time of broad and diverse interest in the subject. Examples had been popping up in many domains: Jay Lifton's notes

on the Hiroshima bombing; Eric Lindemann's report on the psychiatric effects of the disastrous Coconut Grove fire in Boston; Avery Weisman's clinical studies of dying patients; Lloyd Warner's interpretation of the meaning of ceremonial events that honor the dead; Herman Feifel's work on social taboos; Richard Kalish's early essays on teaching; Parsons' emerging theory of the relationship of social action to death, and so on. Fulton, clearly aware of these varying expressions of interest, was prompted to try to interpret the diversity and "get it all together." What Fulton did in *Death and Identity* (1965) was to piece together some three dozen edited excerpts from the works of a wide range of experts who had written on an equally wide range of death-related topics. He found American society to be essentially death denying.

**The Minnesota Conference, 1967.** At the Minnesota Conference of 1967, Alber Sullivan of the Minnesota Medical School and Jacques Choron of the New School of Social Research spoke to various medical and philosophical issues; Jeanne Quint from the University of California reported on the role of the nurse in dealing with terminal patients; Eric Lindemann of Harvard Medical School discussed the symptomatology of acute grief; Herman Feifel of the Veterans Administration, famous for his work on taboos, emphasized that death always carries many meanings; and Talcott Parsons from Harvard probed the topic in broad theoretical terms.

It was a heady agenda but there were some strange omissions. Elizabeth Kubler-Ross, then known to be working on the "stages" of dying (*On Death and Dying*, 1969) was scarcely mentioned, nor was much made of the equally influential work of Barney Glaser and Anselm Strauss (1965), whose book *Awareness of Dying* had been published two years earlier. The participants in the Minnesota Conference were intrigued by various sociological questions. They wrestled with Karl Mannheim's thought experiment of what society would be like if there were no death (Mannheim 1928, 1959). They wondered what mortality and fertility rates had been historically and how they are related (cf. Riley and Riley 1986). They debated Robert Blauner's thesis (Blauner 1966) that death in all known societies imposes imperatives (a corpse must be looked after, property must be reallocated, vacated roles must be reassigned, the solidarity of the

deceased's group must be reaffirmed). They attacked hospital regimens that depersonalized terminal patients, and they challenged the medical profession for treating death as "the enemy" and prolonging life at any cost.

The conference produced extravagant results in anticipating two critical issues: the norms and arrangements for dealing with dying persons were both confused and hazy; and the greater attention paid to caregivers than to dying persons. But the fact that the conference was consistent in insisting that dying always involves at least two persons turned out to be its most important message for the future agenda. Sociological interest in death and dying, of course, did not start with the Minnesota Conference, but what the conference did was to underscore the often overlooked sociological proposition that the dying process is essentially social in nature.

#### A REVIEW OF SOCIOLOGICAL INQUIRIES

The answer to the second question about related research is more straightforward. A bit of American history shows the range of topics that has received sociological attention. In the 1930s sociological interest in death and dying had focused mainly on the economic plight of the bereaved family (Eliot 1932). In the 1950s attention turned to the high cost of dying and the commercialization of funerals (Bowman 1959). Twenty years later, it shifted to a medley of popular topics of peripheral sociological interest, with many books written on various aspects of death and dying, such as *On the Side of Life*; *On Dying and Denying*; *After the Flowers Have Gone*; *Widow*; *Caretaker of the Dead*; *Death in the American Experience*; *Last Rights*; *Someone You Love is Dying*; *The Practice of Death*; *Grief and Mourning*; *No More Dying*; *Life After Life*; *The Way We Die*; *Death as a Fact of Life*; *The Immortality Factor*; *Facing Death*; *Death and Obscenity*; and *Living Your Dying*. One sociologist termed that burst of literature a "collective bustle," and characterized the "discovery" of death during the 1970s as "the happy death movement" (Lofland 1978). There can be no argument that the topic had become more open. Furthermore, the increasing use of life-sustaining technologies dictated that the circumstances of dying became more controllable and negotiable, even as increasing proportions of all deaths were occurring in the later years.

**Little Theoretical Work.** Interest in death and dying was varied and diverse during the 1970s and 1980s, and no widely accepted conceptual framework for its study emerged, except that Kathy Charmaz published a seminal book titled *The Social Reality of Death* in 1980. Sociologists had been critical both of the title and content of Kubler-Ross's widely read book *On Death and Dying* (1969) but they recognized the appeal of the subject matter (Riley 1968, 1983). In earlier decades death had been typically viewed as a social transition, as a "rite de passage," but new threads running through the literature were emerging. Formal "arrangements" were being negotiated prior to death, dying persons were generally more concerned about their survivors than they were about themselves, dying individuals were able to exercise a significant degree of control over the timing of their deaths, tensions typically existed between the requirements of formal care and the wishes of dying patients, and similar tensions almost always existed between formal and informal caregivers—between hospital bureaucracies and those significant others who were soon to be bereaved (Kalish 1985a, 1985b; Riley 1970, 1983).

Little systematic attention from sociologists, however, had emerged. *The Encyclopedia of the Social Sciences* (1968) contained but two entries, both on the social meanings of death. Similarly, there are only two indexed references in the 1988 *Handbook of Sociology* (Smelser 1988): one to poverty resulting from the death of breadwinners, the other to the role of death in popular religion. Sociologists had failed to generate any overarching theory. There have been, however, many attempts. Several kernels illustrate the broad range of these theoretical efforts. Parsons (1963) related the changing meanings of death to basic social values; Mannheim (1928, 1952) used mortality to explain social change; Renee Fox (1980, 1981) found that "life and death were coming to be viewed less as absolute . . . entities. . . and more as different points on a meta-spectrum..a new theodicy"; Dorothy and David Counts (1985) specified the role of death in the various social transformations from preliterate to modern societies; Paul Baker (1990), following Lloyd Warner (1959) and others (e.g. Kears and Rinaldi 1983) elaborated the long-recognized theory that images of the dead exert profound influences on the living, and Michael Kears wrote a more general statement in 1989. And

more recently, Fulton has published an essay on "Society and the Imperative of Death" (1994) in which he discusses the role of such customs and rituals as the Mardi Gras, the bullfight, the "Dani" of primitive societies, and other symbolic events in which either societal survival or individual salvation is at stake.

One exception to these various theoretical efforts is found in the sustained work of Marshall and collaborators. Starting in 1975 with a seminal article in *The American Journal of Sociology*, followed by his book *Last Chapters* (1980), he collaborated with Judith Levy in a review titled "Aging and Dying" (Marshall and Levy 1990). Marshall began his work with an empirical field study of socialization for impending death in a retirement village, followed by a compelling theoretical essay on age and awareness of finitude in developmental gerontology, and has been consistently engaged in such theoretical efforts. His basic postulate is that "awareness of finitude" operates as a trigger that permits socialization to death.

**Empirical Research Largely Topical.** In contrast to theoretical work, the empirical literature shows that sociological research on death and dying has been, and largely continues to be, essentially topical. Studies range widely, from the taboo on death to funerals and the social "causes" of death (Riley 1983; Marshall and Levy 1990). They include the following examples:

**Planning for Death.** A national survey conducted in the late 1960s showed that the great majority of Americans (85 percent) are quite realistic and consider it important to "try to make some plans about death," and to talk about it with those closest to them (Riley 1970). In addition, bereavement practices, once highly structured, are becoming increasingly varied and individually therapeutic; dying is feared primarily because it eliminates opportunities for self-fulfillment; and active adaptations to death tend to increase as one approaches the end of the life course (i.e. the making of wills, leaving instructions, negotiating conflicts).

**Death and Dying in a Hospital.** Among such studies, a detailed account of the "social organization" of death in a public hospital describes rules for dealing with the corpse (the body must be washed, catalogued, and ticketed). Dignity and

bureaucratic efficiency are typically found to be at odds (Sudnow 1967). A contrasting account of hospital rules governing disposition of the body in contemporary Ireland is even more sociological in its emphasis (Prior 1989). In another hospital study the “caring issue” has been seen as the main social problem. The selfhood of the dying person is found to be at risk since the hospital is essentially dedicated to efficiency (Kalish 1985b). However, studies suggest that an increasing proportion of deaths may now be occurring at home or under hospice care, which “mediates between the families and formal institutions that constitute the social organization of death and dying” (Marshall and Levy 1990; see also Bass 1985).

**The Funeral.** The funeral as a social institution has long been of sociological concern (cf. Habenstein 1968). For example, a massive cross-cultural study attests to its worldwide function in marking a major social transition (Habenstein and Lamers 1963; Howarth 1996). Durkheim had emphasized its ceremonial role in facilitating social regrouping. Later sociologists have shown that elaborate and extravagant funeral rites may be more reflective of commercial interests than of human grief or mourning (Parsons and Lidz 1967).

**The Bereaved Family.** The now classic study (Eliot 1932) of the economic consequences of death on the family stimulated a large literature that documents the general proposition that survivors—particularly significant others—require various types of social supports to “get through” the period of intense personal grief and the more publicly expressed mourning. In today’s societies, the time devoted to bereavement activities is generally shorter (Pratt 1981). This is consistent with Parsons’s (1963) position that in societies characterized by an “active” orientation, the bereaved are expected to carry out their grief work quickly and privately.

**Social Stressors as “Causes” of Death.** Sociologists and psychologists have investigated a range of individually experienced “social stressors” as causes of death, such as bereavement and retirement. The hypothesis that a bereaved spouse is at higher risk of death (the “broken heart” syndrome, or “death causes death”) has been widely investigated but with no conclusive results. Similarly, retirees in some longitudinal studies have been

shown to experience excess mortality, whereas other investigations have reported opposite results. Retirement is a complex process, not a simple or single event, and the mortality impact of retirement is moot.

In an era in which nursing homes play an important role in the lives of many older people, the mortality consequences of relocation have come under critical scrutiny. Several studies have reported that the “warehousing” of the frail elderly results in increased mortality while in other studies feelings of security in the new “home” are shown to enhance a sense of well-being resulting in lower mortality. Similar caveats apply to macro-level studies that attempt to relate such collectively experienced stressors as economic depressions, wars, and technological revolutions to trends in mortality. Advances in mathematical modelling and the increasing availability of large and relevant data sets make this problem an attractive area for continuing sociological research (see Riley 1983 for details and sources).

**Self-Motivated Death.** Durkheim’s studies of suicide spawned a wide, diverse, and sometimes confusing research literature. In most such studies social integration is the operative concept. If the theoretical relationship is believed to be unambiguous, the empirical relationship is far from tidy. The literature is vast and well beyond the reach of this review. Apart from suicide, it is a sociological truism that individuals are often socially motivated to influence the time of their own deaths. It has long been noted, for example, that both Thomas Jefferson and John Adams delayed their dying in order to participate in Independence Day celebrations. Several empirical studies have explored this so-called “anniversary effect” in which social events of significance are preceded by lower-than-expected mortality (Phillips and Feldman 1973). Such studies rest on Durkheim’s insight that if some people are so detached from society that they commit suicide, others may be so attached that they postpone their deaths in order to participate in social events of great significance (Phillips and Smith 1990). An example of the mortality impact of personal and local events is seen in studies of the “birthday dip.” One year-long study, in a test area, coded all obituaries for birthdates. The results were striking. Fewer than 10 percent of the deaths occurred during the three months prior to the birth date, whereas nearly half

were reported during the following three months. Along similar lines, several sociological investigations have explored the proposition that some people die socially before they die biologically. These studies center on the notion of "levels of awareness" of death (Glaser and Strauss 1965). When both the dying person and his or her significant others are cognizant of death as a soon-to-be-experienced event, the ensuing "open" awareness may enable them to negotiate various aspects of the final phase of life. Other research on "dying trajectories" involves certainties and uncertainties as to the time of death (Glaser and Strauss 1968).

**"The Right to Die."** As a final and critical example in this review of disparate empirical work, a basic and far-reaching question is being asked: Does the individual, in a society deeply committed to the preservation of life, have a "right" to die? This has become one of the most profound, complex, and pressing issues of our time (Glick 1992). It involves the "rights" and wishes of the dying person, the "rights" and responsibilities of his or her survivors, the "rights" and obligations of attending physicians, and the "rights" and constraints of the law. The human side of such issues is producing a tidal wave of expressions of public interest in television documentaries, opinion surveys, editorials, pamphleteering, and radio talk shows. The issue of euthanasia is openly debated in leading medical journals, an unthinkable topic only a few years ago. Hospital rules, in which do not resuscitate (DNR) orders were written on blackboards then quickly erased, are being changed. Certain aspects of the issue have reached the Supreme Court. A major book has proposed the rationing of medical resources (Callahan 1987). Radical movements have sprung up that advocate active euthanasia and offer recipes for self-deliverance. *Final Exit*, the Hemlock Society's handbook, was an instant best-seller (Humphry 1991). The costs of the last days of life have been dramatized, sometimes spoofed as a myth (Alliance for Aging Research 1996), and sometimes reported with great care (Congressional Research Service). Jack Kevorkian, often referred to as "Dr. Death," has become both hero and despised public enemy. In short, the problems and dilemmas inherent in the "management" of death have captured both popular and scientific attention (see various issues of the Hastings Center Reports). In both instances doctors and lawyers play ambiguous but critical

roles. It is, however, the "negotiation" that is of sociological interest. Norms designed to reduce the perplexities in wrenching decisions or to reassure the decision makers (including dying persons) are generally lacking (Wetle 1994). The need for relevant norms governing "the dying process" has been noted earlier (Riley and Riley 1986), and the main considerations have been specified (Logue 1989). The U. S. Office of Technology Assessment (1987) and the Hastings Center (1987) have issued medical and ethical guidelines, respectively, on the use of life-sustaining procedures. Many years ago sociologists developed research models for studying the social aspects of heroic operations and the treatment of nonsalvageable terminal patients (Fox and Swazey 1974; Crane 1975). Yet models necessary to the formation of norms capable of handling the "rights" and wishes of the various parties to the process of dying are still clearly needed. Furthermore, the conceptual problem of distinguishing between the two actors in the dying process, which the Minnesota Conference had emphasized, has not been resolved.

#### CURRENT ISSUES

With roots in these diverse studies, a set of three issue-laden topics cry out for more research and understanding: 1) dying individuals want a clearer voice in how their last days are to be treated; 2) policy questions are being raised that call attention to potential conflicts between the rights of individuals and the imperatives of society; and 3) programs and campaigns designed to reduce the difficulties of dying are demanding wide social action. These issues can be grouped under three shorthand labels: the living will; assisted suicide; and the quality of dying. While these issues are of great sociological interest, they are only now beginning to be framed in terms for sociological inquiry. The following discussions, consequently, rely largely on commission reports, conferences, public forums, social commentary in the media, brief reports in such journals as *Omega*, *Issues in Law and Medicine*, *Journal of the American Geriatrics Society*, *Hospice Journal*, and various unpublished materials.

**The Living Will.** As noted above, one of the most easily understood and practical developments in response to the dilemmas of dying in America is the "living will," the so-called durable power of

attorney, or some other form of advance directive. These are quasi-legal instruments, signed by the patient, that instruct attending physicians (or surrogates) as to the patient's preferred treatment at the end of life. Such directives are widely varied as to their specificity and the conditions of application, and it is currently impossible to know how many and what types of directives have been executed. There is, however, ample evidence that they are in widespread use. Simple do-it-yourself forms are available in stationery stores, and countless specialized directives have been developed to cover a wide variety of conditions and contingencies. But both the effectiveness and the ethics of such directives have become subject to wide debate: When and under what conditions is the withdrawal of food or fluids legally and medically permissible? When may guardians or surrogates act for incompetent patients? When does the constitutional right to privacy prevail? Under what conditions may the patient refuse treatment or take the initiative and disconnect respirators or tubes? When do oral directives, if ever, take precedence over written ones? The answers to such questions tend to be moot, although a broad legal doctrine has been promulgated that bears on the availability and use of advance directives. There must be "clear and convincing evidence" that the directive accurately reflects the patient's precise intentions—or would in cases of incompetency. This legal dictum, however, has proved to be both burdensome and murky. The well-known Karen Anne Quinlan case is illustrative. This young woman "existed" in a persistent vegetative state for ten years while the legal process could determine whether Karen's parents had met the "clear and convincing evidence" test (Karen's parents had testified that they knew their daughter well enough to be certain that she would not wish to live in such circumstances). In another widely cited case, Nancy Cruzan "lived" in a similar state for seven years while the intricacies of the law were being debated. These and other such cases point to the need for more useful and practical evidentiary tests. The Hastings Center published a special supplement titled "Advance Care Planning" in 1993.

In 1990, the Patient Self-Determination Act (PSDA) raised a set of new questions. The PSDA requires hospitals, nursing homes, and health providers to inform patients of their right to prepare a living will or some document of end-of-life

preferences. The significance of the PSDA was underscored by an occasionally distributed joint statement by the American Medical Association and the Harvard Medical School:

*Modern medicine can keep one alive long after any reasonable prospect of mental, spiritual, or emotional life is gone. The only way for a person to retain autonomy in such a situation is to record his or her preferences for medical care before they are needed. (Published in The Harvard Health Letter and elsewhere.)*

The force of the PSDA, however, has not been great and today it is generally believed that its lasting importance will be found in its power to enhance understanding of the still-developing and changing doctor-patient relationship. Indeed, the drama of that relationship has now been moved to a larger stage that involves both assisted suicide and the quality of dying.

**Assisted Suicide.** The role of law in cases where patients or their surrogates seek to control end-of-life decisions has always been debated. The early cases in the 1950s and 1960s had revolved around "informed consent." This rule surfaced when medical treatments resulted in unanticipated negative consequences, and when it could be shown that the patient had not been informed of the risks. Not surprisingly, this rule led to more complicated ones and, during the 1970s, the increasing demand for patient control resulted in an implied "right" to die by refusing treatment. Of sociological interest, it was popular experience—not statutory law or court decisions—that was bringing about social change (see M. W. Riley 1978 for a theoretical statement). Nor was it long before demands for the "right" to receive treatments specifically designed to hasten death were seriously being discussed and in some states actually outlawed. As these events unfolded they have been reported by major news services, and analyzed by Hastings Center reports beginning in 1995.

In 1996 the U.S. Court of Appeals for the Ninth Circuit struck down a Washington state statute that had been passed specifically to deny such a right. The presiding judge included this noteworthy statement:

*A competent, terminally ill adult, having lived nearly the full measure of his life, has a strong liberty interest in choosing a dignified and*

*humane death rather than being reduced at the end of his existence to a childlike state of helplessness—diapered, sedated, incompetent.*

The case was striking not only for its human interest but also because it invoked the guarantee of personal liberty in the Fourteenth Amendment to the U.S. Constitution. Shortly thereafter a New York statute was struck down by the U.S. Court of Appeals for the Second Circuit (Quill 1996), which had argued that if physicians were allowed to help people die it would put society on a slippery slope leading inevitably to abuse. The court pointed out, however, that physicians are not killers if they prescribe drugs to hasten death any more than they are killers when they discontinue life supports.

As expected, both cases were sent to the U.S. Supreme Court which has, at this writing, let stand state laws that prohibit any form of physician-assisted suicide. But the issue is far from settled. The Court's decision concluded with this surprising statement by Chief Justice Rehnquist:

*Throughout the nation, Americans are engaged in an earnest and profound debate about the morality, legality and practicality of physician-assisted suicide. Our holding permits this debate to continue.*

This statement was all the more remarkable since the chief justice, in his long opinion, rejected both the "liberty" and "due process" constitutional arguments, but came far short of putting the matter to rest. Indeed, continuing developments indicate that the question of physician-assisted suicide is not likely to go away soon. For example, the State of Oregon has passed two voter referenda, the most recent in 1998 with a 60 percent majority support which makes Oregon the only state (as of 1998) to permit, under strict conditions, physician-assisted suicide. The State of Michigan has rejected an Oregon-type statute and has finally convicted Dr. Kevorkian of second degree murder. Several other states are reported to be experimenting with alternatives that enhance the "right" of the individual to choose to die (for details of these developments see The Hastings Center reports). It may be that the law is overreaching its capacity to deal with such a basic and philosophical issue. The question is profound. Is there any logical (or sociological) difference between the right to refuse treatment designed to prolong

life and a parallel right to receive treatment designed to hasten death? Sociologists would ask: What is the distinction between the acceptance of death and its acceleration? Is the question so abstract that it defies empirical inquiry? Both the medical and legal answers to the issue are ambiguous and have been discussed in a book by a physician who publicly admitted helping a patient to die. Dr. Timothy Quill not only was the main plaintiff in the New York case, but he has become the leading medical voice on issues of assisted suicide. He writes with authority and sensitivity: "Death seems antithetical to modern medicine—no longer a natural and inevitable part of the life cycle, but a medical failure to be fought off, ignored, and minimized. The dark side of this desperate battle has patients spending their last days in the intensive care units of acute hospitals, tubes inserted into every body part, vainly trying to forestall death's inevitability. No one wants to die, but if we have to, there must be a better way" (for an account of the issues see Quill 1996).

**The Quality of Dying.** Sociological concern with assisted suicide has been paralleled with concern for how people die. It had been hoped that the PSDA not only would make advance directives more effective, but would also bring about better communications between doctors and terminal patients. The act, however, was deemed a failure even before it was formally put in place. An impressive experiment, of great sociological interest, was designed to solve these basic issues. Funded and launched by the Robert Wood Johnson Foundation, it was the advance directive problem cast in research terms. The "Study to Understand Prognoses and Preferences in Risks of Treatment" carried such an unwieldy title that it was quickly shortened to the acronym SUPPORT (for a detailed account see The Hastings Center special supplement that carries the subtitle "The Lessons of SUPPORT" 1995). Five teaching hospitals were invited to participate in this multimillion dollar project. Phase One called for baseline data on the end-of-life experiences of some 9,000 dying patients. When the data were analyzed, the researchers were not surprised to find much to criticize in various regimens of hospital care. Their main finding, however, was that doctors and attending nurses must attend not only to the physical comfort and pain management needs of patients but, more importantly, to their psychosocial needs.

Phase Two, consequently, called for experimentation. The 9,000 cases were randomized and an experimental intervention consisting of a protocol designed to sensitize doctors to attend more closely to psychosocial needs was administered to one-half of the cases and withheld from the other half, which had served as a control group. The experiment was continued for two years and expectations were high. Much to the chagrin of the study directors, when the two groups were compared, no differences were found! The experimental intervention showed no effects. This negative finding was so shocking that a number of evaluation panels were enlisted to reanalyze the data and scrutinize the research design. Their efforts, however, only served to corroborate the original analysis and Daniel Callahan (1995), then president of The Hastings Center, concluded with this statement, "This painstaking scrutiny into how people die only goes to show how difficult it is to make the process any better. . . . We thought that the care of dying patients could be set right by . . . some good talk between doctor and patient . . . we thought that we just needed reform . . . it is now obvious that we need a revolution" (Callahan 1995).

Callahan's dramatic statement had the effect of nourishing a spate of organizations and proposals that had sprung up to improve the care of dying patients. For example "Project Death in America" (PDIA), centered at Sloan Kettering Hospital in New York was funded by the billionaire George Soros; the "Center to Improve Care of the Dying" (CICD), based in the George Washington Medical School and originally funded by the Retirement Research Foundation in 1995, enjoys wide institutional support and in 1997 launched a program for individual advocacy, "Americans for Better Care of the Dying" (ABCD); the American Board of Internal Medicine (ABIM) began publishing educational materials on techniques to improve care of the dying in 1996; and the American Medical Association (AMA) began alerting its members to the most recent developments in end-of-life care. And finally, the Robert Wood Johnson Foundation, which has long served as the collective voice for issues on dying, launched "Last Acts" in 1997, which is designed to involve the public through such participating organizations as the Institute of Medicine, the American Hospital Association, the Health Care Financing Administration, the National Hospice Organization, the Veterans Health

Administration, the American College of Physicians, and a host of specialized associations such as the American Pain Society, the American Cancer Society, Choice in Dying, Partnership for Organ Donation, Memorial Societies of America, Wellness Councils of America, and on and on. Former First Lady Roslyn Carter, in a nationally broadcast speech, proclaimed that, "We need this coalition so that fewer people will die alone, in pain, and attached to machines, with the result that more people. . . can experience dying for what it ought to be. . . the last act in the journey of life." The President of the Robert Wood Johnson Foundation was both optimistic and enthusiastic:

*With all that is going on. . . we are seizing the moment, "Last Acts" will be much more than platitudes about a good death. . . it will undertake to improve care at the end of life. . . If the campaign succeeds . . . we will find a significant decrease in the number of people dying in pain, an increase in referrals to hospice, more people dying at home outside the hospital, and fewer requests for physician-assisted suicide (1997).*

If, however, the campaign does not live up to expectations, the foundation will work with the American Medical Association "on helping physicians to work with patients on advance care planning, and providing opportunities for physicians to increase their skills in palliative medicine and comfort care." It is too early to estimate the long-term effects of this blizzard of efforts to improve the way people die, but it is not too early to predict that any improvements in the quality of dying based only on comfort care are likely to be short term. But change is in the air. Dying persons are pressing for more participation in when and how they die, and caregivers are coming under increasing criticism of the limitations of their caring regimens.

### AN EVOLVING PERSPECTIVE

As indicated by this review of past research and current issues, the limitations of both legal and medical approaches to the problems of dying in contemporary society mean that the perspectives of sociology will be brought to bear. Perhaps future sociological attention will focus on the uniquely sociological principle that the dying process is not understandable in individual terms. This,



in turn, provokes a series of predictions: Sociologists will focus on how dying persons are defined by their survivors and caregivers as well as on the sociological hypothesis that dying persons are more concerned for others than they are for themselves. Sociologists will be required to disentangle the concatenation of forces that has produced today's caregiving regimens in which terminal patients tend to be treated more as objects than as persons. Sociologists will be asked to explain how the process of socialization has seemingly been reversed, with the individual being figuratively stripped of years of social experience and defined as a nonperson. In effect, sociologists will be asked to explain the historical alchemy whereby dying persons themselves have been conned into believing that all they needed was palliation and comfort care.

## REFERENCES

- Alliance for Aging Research 1996 *Uncovering the Facts about the Cost of the Last Year of Life*. Washington, D.C.: AAR.
- Baker, Paul M. 1990 "Socialization after Death: the Might of the Living Dead." In Beth Hess and Elizabeth Markson, eds., *Growing Old in America*. Rutgers: Transaction Books.
- Bass, David M. 1985 "The Hospice Ideology and Success of Hospice Care." *Research on Aging* 7:1.
- Blauner, Robert 1966 "Death and Social Structure." *Psychiatry* 29:378-94.
- Bowman, L. 1959 *The American Funeral: A Study in Guilt, Extravagance, and Sublimity*. Washington D.C.: Public Affairs Press.
- Callahan, D. 1995 "Dying Well in the Hospital." *Hastings Center Report*. Briarcliff Manor, N.Y.: Hastings Center.
- 1987 *Setting Limits; Medical Goals in an Aging Society*. New York: Simon and Schuster.
- Charmaz, K. C. 1980 *Social Reality of Death*. Reading, Mass.: Addison-Wesley.
- Counts, D.A., and D.C. Counts 1985 *Aging and Its Transformations: Moving Toward Death in Pacific Societies*. Lanham, Md: University Press of America.
- Crane, D. 1975 *The Sanctity of Social Life: Physicians' Treatment of Critically Ill Patients*. New York: Russell Sage Foundation.
- Eliot, T. 1932 "The Bereaved Family." Special issue, *The Annals* 160.
- Fox, R.C., and T.P. Swazey 1981 "The Sting of Death in American Society." *Social Service Review*. March:42-59.
- 1980 "The Social Meaning of Death." *The Annals* 447.
- 1974 *The Courage to Fail: A Social View of Organ Transplantation and Dialysis*. Chicago: University of Chicago Press.
- Fulton, R. 1994 "Society and the Imperative of Death." In I. Corless et al., eds., *Dying, Death and Bereavement*. Boston: Jones and Bartlett.
- 1965 *Death and Identity*. New York: John Wiley and Sons.
- Fulton, R., and R. Bendiksen 1994 *Death and Identity*, 3rd ed. Philadelphia: The Charles Press.
- Fulton, R., and G. Owen 1988 "Death and Society in Contemporary Society." *Omega* 18(4).
- Glaser, B.G., and A.L. Strauss 1965 *Awareness of Dying*. Chicago: Aldine.
- 1968 *Time for Dying*. Chicago: Aldine.
- Glick, H.R. 1992 *The Right to Die: Policy Innovation and its Consequences*. New York: Columbia University Press.
- Habenstein, R.W. 1968 "The Social Organization of Death." *The International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- , and M.W. Lamers 1963 *Funeral Customs the World Over*. Milwaukee: Bulfin.
- Hastings Center 1987 *Guidelines on the Termination of Life-Sustaining Treatment and the Care of the Dying*. Bloomington, Ind.: Indiana University Press.
- Howarth, G. 1996 *Last Rites: The Work of the Modern Funeral Director*. Amityville N.Y.: Baywood Publishing Co.
- Humphry, Derek 1991 *Final Exit: the Practicalities of Self-Deliverance and Assisted Suicide*. Eugene Oreg.: The Hemlock Society.
- Kalish, R.A. 1985a *Death, Grief and Caring Relationships*. Monterey, Calif.: Brooks/Cole.
- 1985b "The Social Context of Death and Dying." In R.H. Binstock and E. Shanas, eds., *Handbook of Aging and the Social Sciences*. New York: Van Nostrand Reinhold Co.
- Kearl, Michael C. 1989 *Endings: A Sociology of Death and Dying*. Oxford: Oxford University Press.
- , and A. Rinaldi 1983 "The Political Uses of the Dead as Symbols in Contemporary Civil Religions." *Social Forces* 61:3.
- Kubler-Ross, E. 1969 *On Death and Dying*. New York: Macmillan.
- Lofland, L. 1978 *The Craft of Dying: The Modern Face of Death*. Beverly Hills, Calif.: Sage.

- Logue, B. 1989 *Death Control and the Elderly: The Growing Acceptability of Euthanasia*. Providence: Population Studies and Training Center, Brown University.
- Mannheim, K. 1928, 1959 "The Problem of Generations." In P. Keeskemati, ed., *Essays in Sociological Knowledge*. London: Routledge and Kegan Paul.
- Marshall, V.W. 1980 *Last Chapters: A Sociology of Aging and Dying*. Belmont Calif.: Wordsworth.
- , and J. A. Levy 1990 "Aging and Dying." In R. H. Binstock and L. George, eds., *Handbook of Aging and the Social Sciences*, 3rd ed. San Diego: Academic Press.
- Parsons, T. 1963 "Death in American Society: A Brief Working Paper." *American Behavioral Scientist*.
- , and V. W. Lidz 1967 "Death in American Society." In E.S. Shneidman, ed., *Essays in Self Destruction*. New York: Science House.
- Phillips, D.P., and K.A. Feldman 1973 "A Dip in Deaths before Ceremonial Occasions: Some New Relationships between Integration and Mortality." *American Journal of Sociology* 87.
- , and D.G. Smith 1990 "Postponement of Death Until Symbolically Meaningful Occasions." *Journal of the American Medical Association*, vol. 203(14).
- Pratt, L.V. 1981 "Business Temporal Norms and Bereavement Behavior." *American Sociological Review* 4:317-333.
- Prior, Lindsay 1989 *The Social Organization of Death: Medical Discourses and Social Practices in Belfast*. New York: St.Martin's Press.
- Quill, T.E. 1996 *A Midwife through the Dying Process*. Baltimore, Md.: The Johns Hopkins University Press.
- Riley, J.W., Jr. 1983 "Dying and the Meanings of Death: Sociological Inquiries." *Annual Review of Sociology* 9:191-216.
- 1970 "What People Think about Death." In O. G. Brim, Jr. et al., eds., *The Dying Patient*. New York: Russell Sage Foundation.
- 1968 "Death and Bereavements." In D.L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- Riley, M. W., and J.W. Riley, Jr. 1986 "Longevity and Social Structure: the Added Years." *Daedalus* 115:51-75.
- Sudnow, D. 1967 *Passing On: The Social Organization of Dying*. Englewood Cliffs, N.J.: Prentice-Hall.
- U. S. Congress, Office of Technology Assessment 1987 *Life-Sustaining Technologies and the Elderly*. Washington D.C.: U.S. Government Printing Office.
- Warner, W.L. 1959 *The Living and the Dead*. New Haven, Conn.: Yale University Press.

- Wetle, T. 1994 "Individual Preferences and Advance Planning." *Hastings Center Report*. Briarcliff Manor, N.Y.: Hastings Center.

JOHN W. RILEY, JR.

## DECISION-MAKING THEORY AND RESEARCH

Decision making must be considered in any explanation of individual behavior, because behaviors are based on decisions or judgments people have made. Thus, decision-making theory and research is of interest in many fields that examine behavior, including cognitive psychology (e.g., Busemeyer, Medin, and Hastie 1995), social psychology (e.g., Ajzen 1996), industrial and organizational psychology (e.g., Stevenson, Busemeyer, and Naylor 1990), economics (e.g., Lopes 1994), management (e.g., Shapira 1995), and philosophy (e.g., Manktelow and Over 1993), as well as sociology. This section will be an overview of decision-making theory and research. Several excellent sources of further information include: Baron (1994), Dawes (1997), Gilovich (1993), and Hammond (1998).

### DECISION-MAKING THEORIES

Most decision-making theory has been developed in the twentieth century. The recency of this development is surprising considering that gambling has existed for millennia, so humans have a long history of making judgments of probabilistic events. Indeed, insurance, which is in effect a form of gambling (as it involves betting on the likelihood of an event happening, or, more often, not happening), was sold as early as the fifteenth and sixteenth centuries. Selling insurance prior to the development of probability theory, and in many early cases without any statistics for, or even frequencies of, the events being insured led to bankruptcy for many of the first insurance sellers (for more information about the history of probability and decision making see Hacking 1975, 1990; Gigerenzer et al. 1989).

**Bayes's Theorem.** One of the earliest theories about probability was *Bayes's Theorem* (1764/1958). This theorem was developed to relate the probability of one event to another; specifically, the

probability of one event occurring given the probability of another event occurring. These events are sometimes called the cause and effect, or the hypothesis and the data. Using H and D (hypothesis and data) as the two events, Bayes's Theorem is:  $P(H|D) = P(H)P(D|H)/P(D)$  [1] or  $P(H|D) = P(D|H)P(H)/[P(D|H)P(H)+P(D|-H)P(-H)]$  [2] That is, the probability of H given D has occurred is equal to [1] the probability of H occurring multiplied by the probability of D occurring given H has occurred divided by the probability of D occurring, or [2] the probability of D given H has occurred multiplied by the probability of H then divided by the probability of D given H has occurred multiplied by the probability of H plus the probability that D occurs given H has not occurred multiplied by the probability that H does not occur.

The cab problem (introduced in Kahneman and Tversky 1972) has been used in several studies as a measure of whether people's judgments are consistent with Bayes's Theorem. The problem is as follows: A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data: (a) 85 percent of the cabs in the city are Green and 15 percent are Blue. (b) a witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80 percent of the time and failed 20 percent of the time. What is the probability that the cab involved in the accident was Blue rather than Green? Using the provided information and formula [2] above,  $P(\text{Blue Cab}|\text{Witness says "Blue"}) = P(\text{Witness says "Blue"}|\text{Blue Cab})P(\text{Blue Cab})/[P(\text{Witness says "Blue"}|\text{Blue Cab})P(\text{Blue Cab}) + P(\text{Witness says "Blue"}|\text{Green Cab})P(\text{Green Cab})]$  or  $P(\text{Blue Cab}|\text{Witness says "Blue"}) = (.80)(.15)/[(.80)(.15)+(.20)(.85)] = (.12)/[(.12)+(.17)] = .41$

Thus, according to Bayes's Theorem, the probability that the cab involved in the accident was Blue, given the witness testifying it was Blue, is 0.41. So, despite the witness's testimony that the cab was Blue, it is more likely that the cab was Green (0.59 probability), because the probabilities for the base rates (85 percent of cabs are Green and 15 percent Blue) are more extreme than those for the witness's accuracy (80 percent accuracy).

Generally, people will rate the likelihood that the cab was Blue to be much higher than .41, and often the response will be .80—the witness's accuracy rate (Tversky and Kahneman 1982).

That finding has been used to argue that people often ignore base rate information (the proportions of each type of cab, in this case; Tversky and Kahneman 1982), which is irrational. However, other analyses of this situation are possible (cf. Birnbaum 1983; Gigerenzer and Hoffrage 1995), which suggest that people are not irrationally ignoring base rate information. The issue of rationality will be discussed further below.

**Expected Utility (EU) Theory.** Bayes's Theorem is useful, but often we are faced with decisions to choose one of several alternatives that have uncertain outcomes. The best choice would be the one that maximizes the outcome, as measured by utility (i.e., how useful something is to a person). Utility is not equal to money (high-priced goods may be less useful than lower-priced goods), although money may be used as a substitute measure of utility. EU Theory (von Neumann and Morganstern 1947) states that people should maximize their EU when choosing among a set of alternatives, as in:  $EU = \sum (U_i * P_i)$ ; where  $U_i$  is the utility for each alternative,  $i$ , and  $P_i$  is the probability associated with that alternative.

The earlier version of this theory (Expected Value Theory, or EV) used money to measure the worth of the alternatives, but utility recognizes people may use more than money to evaluate the alternatives. Regardless, in both EU and EV, the probabilities are regarded similarly, so people only need consider total EU/EV, not the probability involved in arriving at the total.

However, research suggests that people consider certain probabilities to be special, as their judgments involving these probabilities are often inconsistent with EU predictions. That is, people seem to consider events that have probabilities of 1.0 or 0.0 differently than events that are uncertain (probabilities other than 1.0 or 0.0). The special consideration given to certain probabilities is called the *certainty effect* (Kahneman and Tversky 1979). To illustrate this effect, which of these two options do you prefer? A. Winning \$50 with probability .5 B. Winning \$30 with probability .7 Now which of these next two options do you prefer? C. Winning

\$50 with probability .8 D. Winning \$30 with probability 1.0

Perhaps you preferred A and D, which many people do. However, according to EU, those choices are inconsistent, because D does not have a higher EU than C (for D,  $EU = \$30 = (1.0 * \$30)$ ; for C,  $EU = \$40 = (.8 * \$50)$ ). Recognize that A and B differ from C and D by a .3 increase in probability, and EU prescribes selecting the option with the highest EU. The certainty effect may also be seen in the following pair of options. E. Winning \$1,000,000 with probability 1.0 F. Winning \$2,000,000 with probability .5

According to EU, people should be indifferent between E and F, because they both have the same EU ( $\$1,000,000 * 1.0 = \$2,000,000 * .5$ ). However, people tend to prefer E to F. As the cliché goes, a bird in the hand is worth two in the bush. These results (choosing D and E) suggest that people are risk averse, because those are the certain options, and choosing them avoids risk or uncertainty. But risk aversion does not completely capture the issue. Consider this pair of options: G. Losing \$50 with probability .8 H. Losing \$30 with probability 1.0

If people were risk averse, then most would choose H, which has no risk; \$30 will be lost for sure. However, most people choose G, because they want to avoid a certain loss, even though it means risking a greater loss. In this case, people are risk seeking.

The tendency to treat certain probabilities differently from uncertain probabilities led to the development of decision-making theories that focused on explaining how people make choices, rather than how they should make choices.

**Prospect Theory and Rank-Dependent Theories.** Changing from EV to EU acknowledged that people do not simply assess the worth of alternatives on the basis of money. The certainty effect illustrates that people do not simply assess the likelihood of alternatives, so decision theories must take that into account. The first theory to do so was prospect theory (Kahneman and Tversky 1979).

*Prospect theory* proposes that people choose among prospects (alternatives) by assigning each prospect a subjective value and a decision weight (a value between 0.0 and 1.0), which may be functionally equal to monetary value and probability,

respectively, but need not be actually equal to them. The prospect with the highest value as calculated by multiplying the subjective value and the decision weight is chosen. Prospect theory assumes that losses have greater weight than gains, which explains why people tend to be risk seeking for losses but not for gains. Also, prospect theory assumes that people make judgments from a subjective reference point rather than an objective position of gaining or losing.

Prospect theory is similar to EU in that the decision weight is independent of the context. However, recent decision theories suggest weights are created within the context of the available alternatives based on a ranking of the alternatives (see Birnbaum, Coffey, Mellers, and Weiss 1992; Luce and Fishburn 1991; Tversky and Kahneman 1992). The need for a rank-dependent mechanism within decision theories is generally accepted (Mellers, Schwartz, and Cooke 1998), but the specifics of the mechanism are still debated (see Birnbaum and McIntosh 1996).

**Improper Linear Models.** Distinguishing between alternatives based on some factor (e.g., value, importance, etc.) and weighting the alternatives based on those distinctions has been suggested as a method for decision making (Dawes 1979). The idea is to create a linear model for the decision situation. Linear models are statistically derived weighted averages of the relevant predictors. For example:  $L(\text{lung cancer}) = w_1 * \text{age} + w_2 * \text{smoking} + w_3 * \text{family history}$ , where  $L(\text{lung cancer})$  is the likelihood of getting lung cancer, and  $w_x$  is the weight for each factor. Any number of factors could be included in the model, although only factors that are relevant to the decision should be included. Optimally, the weights for each factor should be constructed from examining relevant data for the decision.

However, Dawes (1979) has demonstrated that linear models using equal weighting are almost as good as models with optimal weights, although they require less work, because no weight calculations need be made; factors that make the event more likely are weighted +1, and those that make it less likely are weighted -1.

Furthermore, linear models are often better than a person's intuition, even when the person is an expert. Several studies of clinical judgment

(including medical doctors and clinical psychologists) have found that linear models always do as well as, if not better than, the clinical experts (see Einhorn 1972). Similarly, bank loan officers asked to judge which businesses will go bankrupt within three years of opening was about 75 percent accurate, but a statistical model was 82 percent accurate (Libby 1976).

Arkes, Dawes, and Christensen (1986) demonstrated this point with people knowledgeable about baseball. Participants were asked to identify “which of . . . three players won the MVP [most valuable player] award for that year.” Each player’s season statistics were provided. One of the three players was from the World Series winning team, and subjects were told that 70 percent of the time the MVP came from the World Series winning team, so if they were uncertain, they could use that decision rule.

Participants moderately knowledgeable of baseball did better than the participants highly knowledgeable, although the highly knowledgeable participants were more confident. The moderately knowledgeable group did better because they used the decision rule more. Yet, neither group did as well as they could have, if they had used the decision rule for every judgment. How a little knowledge can influence judgment will be further explained in the next section on how people make decisions.

### DECISION PROCESSING

Decision theories changed because studies revealed that people often do not make judgments that are consistent with how the theories said they should be making judgments. This section will describe evidence about how people make judgments. Specifically, several *heuristics* will be discussed, which are short cuts that people may use to process information when making a judgment (Kahneman, Slovic, and Tversky 1982, is a classic collection of papers on this topic).

**Availability.** Consider the following questions: What is the first digit that comes to mind?, What is the first one-digit number that comes to mind?, and What is the first digit, such as one, that comes to mind? Kubovy (1977) found that the second statement, which mentions “1” in passing, resulted in more “1” responses than either of the other

two statements. The explanation is that mentioning “1” made it more available in memory. People are using *availability*, when they make a judgment on the basis of what first comes to mind.

Interestingly, the third statement, which mentions “1” in an explicit manner, resulted in fewer “1” responses than the first statement, which does not mention “1” at all. Kubovy suggests there are fewer “1” responses, because people have an explanation of why they are thinking of “1” following that statement (it mentions “1”), so they do not choose it, because “1” has not been thought of at random. Thus, information must not only be available, but it must be perceived as relevant also.

**Mood.** Information that is available and may seem relevant to a judgment is a person’s present mood (for a review of mood and judgment research see Clore, Schwarz, and Conway 1994). Schwarz and Clore (1983) suggest that people will use their mood state in the judgment process, if it seems relevant to that judgment. They contacted people by phone on cloudy or sunny days, and predicted that people would be happier on the sunny days than on the cloudy days. That prediction was verified. However, if people were first asked what the weather was like, then there was no difference in people’s happiness on sunny and cloudy days. Schwarz and Clore suggested that asking people about the weather gave them a reason for their mood, so they did not use their mood in making the happiness judgment, because it did not seem relevant. Thus, according to Schwarz and Clore, people will use their mood as a heuristic for judgment, when the situation elicits actions from people as if they ask themselves, how do I feel about this?

**Quantity and Numerosity.** An effect that is also similar to availability was demonstrated in a series of experiments by Josephs, Giesler, and Silvera (1994). They found that subjects relied on observable quantity information when making personal performance judgments. They called this effect the *quantity principle*, because people seemed to use the size or quantity of material available to make their judgment. The experimenters had participants proofread text that was either attached to the source from which it came (e.g., a book or journal) or not, which was an unfamiliar task for the participants. They found that performance

estimations on an unfamiliar task were affected by the quantity of the task work completed. Participants whose work resulted in a large pile of material rated their performance higher than participants whose work resulted in small pile. However, the amount of actual work done was the same. Moreover, if the pile was not in sight at the time of judgment, then performance estimates did not differ.

Pelham, Sumarta, and Myaskovsky (1994) demonstrated an effect similar to the quantity principle, which they called the *numerosity heuristic*. This heuristic refers to the use of the number of units as a cue for judgment. For example, the researchers found that participants rated the area of a circle as larger when the circle was presented as pieces that were difficult to imagine as a whole circle than if the pieces were presented so that it was easy to imagine, and larger than the undivided circle. This result suggests that dividing a whole into pieces will lead to those pieces being thought of as larger than the whole, and this will be especially so when it is hard to imagine the pieces as the whole. That is, the numerosity of the stimulus to be rated can affect the rating of that stimulus.

**Anchoring.** *Anchoring* is similar to availability and quantity heuristics, because it involves making a judgment using some stimulus as a starting point, and adjusting from that point. For example, Sherif, Taub, and Hovland (1958) had people make estimates of weight on a six-point scale. Prior to each estimation, subjects held an “anchor” weight that was said to represent “6” on the scale. When the anchor’s weight was close to the other weights, then subjects’ judgments were also quite close to the anchor, with a modal response of 6 on the six-point scale. However, heavier anchors resulted in lower responses. The heaviest anchor produced a modal response of 2.

However, anchoring need not involve the direct experience of a stimulus. Simply mentioning a stimulus can lead to an anchoring effect. Kahneman and Tversky (1972) assigned people the number 10 or 65 by means of a seemingly random process (a wheel of fortune). These people then estimated the percentage of African countries in the United Nations. Those assigned the number 10 estimated 25 percent, and those assigned the number 65 estimated 45 percent (at the time the actual percentage was 35). This indicates the psychological

impact of one’s starting point (or anchor), when making a judgment.

An anecdotal example of someone trying to fight the anchoring phenomenon is a writer who tears up a draft of what she is working on, and throws it away, rather than trying to work with that draft. It may seem wasteful to have created something, only to throw it out. However, at times, writers may feel that what they have produced is holding them in a place that they do not want to be. Thus, it could be better to throw out that draft, rather than trying to rework it, which is akin to pulling up anchor.

**Endowment Effect.** Related to anchoring is the tendency people have to stay where they are. This tendency is known as the *endowment effect* (Thaler 1980) or the *status quo bias* (Samuelson and Zeckhauser 1988). This tendency results in people wanting more for something they already have than they would be willing to pay to acquire that same thing. In economic terms, this would be described as a discrepancy between what people are willing to pay (WTP) and what they are willing to accept (WTA). Kahneman, Knetsch, and Thaler (1990) found WTA amounts much higher for an item already possessed than WTP amounts to obtain that item. In their study, half of the people in a university class were given a university coffee mug, and half were not. A short time later, the people with the mugs were asked how much money they wanted for their mug (their WTA amount), and those without the mug were asked how much they would pay to get a mug (their WTP amount). The median WTA amount was about twice the median WTP amount: WTA = \$7.12, WTP = \$2.87. Thus, few trades were made, and this is consistent with the idea that people often seem to prefer what they have to what they could have.

**Representativeness.** People may make judgments using *representativeness*, which is the tendency to judge events as more likely if they represent the typical or expected features for that class of events. Thus, representativeness occurs when people judge an event using an impression of the event rather than a systematic analysis of it. Two examples of representativeness misleading people are the gambler’s fallacy and the conjunction fallacy.

The gambler’s fallacy is the confusion of independent and dependent events. Independent events are not causally related to each other (e.g., coin

flips, spins of a roulette wheel, etc.). Dependent events are causally related; what happened in the past has some bearing on what happens in the present (e.g., the amount of practice has some bearing on how well someone will perform in a competition). The confusion arises when people's expectations for independent events are violated. For example, if a roulette wheel comes up black eighteen times in a row, some people might think that red must be the result of the next spin. However, the likelihood of red on the next spin is the same as it is each and every spin, and the same as it would be if red, instead of black, had resulted on each of the previous eighteen spins. Each and every roulette wheel spin is an independent event.

The conjunction fallacy (Tversky and Kahneman 1983) occurs when people judge the conjunction of two events as more likely than (at least) one of the two events. The "Linda scenario" has been frequently studied: Linda is thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Following that description, subjects are asked to rank order in terms of probability several statements including the following: Linda is active in the feminist movement. [F] Linda is a bank teller. [B] Linda is a bank teller and is active in the feminist movement. [B&F]

The conjunction fallacy is committed if people rank the B&F statement higher (so more likely) than either the B or F statement alone, because that is logically impossible. The likelihood of B&F may be equal to B or F, but it cannot be greater than either of them, because B&F is a subset of the set of B events and the set of F events.

People find events with multiple parts (such as B&F) more plausible than separate events (such as B or F alone), but plausibility is not equal to likelihood. Making an event more plausible might make it a better story, which could be misleading and result in erroneous inferences (Markovits and Nantel 1989). Indeed, some have suggested that people act as if they are constructing stories in their minds and then make judgments based on the stories they construct (Pennington and Hastie 1993). But of course good stories are not always true, or even likely.

## INDIVIDUAL DIFFERENCES

Another question about decision processing is whether there are individual differences between people in their susceptibility to erroneous decision making. For example, do some people tend to inappropriately use the heuristics outlined above, and if so, is there a factor that accounts for that inappropriate use?

Stanovich and West (1998) had participants do several judgment tasks and related the performance on those tasks to assessments of cognitive ability and thinking styles. They found that cognitive capacity does account for some performance on some judgment tasks, which suggests that computational limitations could be a partial explanation of non-normative responding (i.e., judgment errors). Also, independent of cognitive ability, thinking styles accounted for some of the participants' performance on some judgment tasks.

A similar suggestion is that some erroneous judgments are the result of participants' conversational ability. For example, Slugoski and Wilson (1998) show that six errors in social judgment are related to people's conversational skills. They suggest that judgment errors may not be errors, because participants may be interpreting the information presented to them differently than the researcher intends (see also Hilton and Slugoski 1999).

Finally, experience affects decision-making ability. Nisbett, Krantz, Jepson, and Kunda (1983) found that participants with experience in the domain in question preferred explanations that reflected statistical inferences. Similarly, Fong, Krantz, and Nisbett (1986) found that statistical explanations were used more often by people with more statistical training. These results suggest that decision-making ability can improve through relevant domain experience, as well as through statistical training that is not domain specific.

## GROUP DECISION MAKING

**Social Dilemmas.** *Social dilemmas* occur when the goals of individuals conflict with the goals of their group; individuals face the dilemma of choosing between doing what is best for them personally and what is best for the group as a whole (Lopes 1994). Hardin (1965) was one of the first to write about these dilemmas in describing the "tragedy

of the commons.” The tragedy was that individuals tried to maximize what they could get from the common and, or the “commons,” which resulted in the commons being overused, thereby becoming depleted. If each individual had only used his allotted share of the commons, then it would have continued to be available to everyone.

**Prisoner’s Dilemma.** The best-known social dilemma is the prisoner’s dilemma (PD), which involves two individuals (most often, although formulations with more than two people are possible). The original PD involved two convicts’ decision whether or not to confess to a crime (Rapoport and Chammah 1965). But the following example is functionally equivalent.

Imagine you are selling an item to another person, but you cannot meet to make the exchange. You agree to make the exchange by post. You will send the other person the item, and receive the money in return. If you both do so, then you each get 3 units (arbitrary amount, but amounts received for each combination of choices is important). However, you imagine that you could simply not put the item in the post, yet still receive the money. Imagine doing so results in you getting 5 units and the other person  $-1$  units. However, the other person similarly thinks that not posting the money would result in getting the item for free, which would result in 5 units for the other person and  $-1$  for you. If you both do not put anything in the post, although you agreed to do so, you would be at the status quo (0 units each).

Do you post the item (i.e., cooperate) or not? Regardless of what the other person does, you will get more out of not cooperating (5 v. 3, when the other person cooperates, and 0 v.  $-1$ , when other does not). However, if you both do not cooperate, that produces an inferior group outcome, compared to cooperating (0 [0+0] v. 6 [3+3], respectively). Thus, the dilemma is that each individual has an incentive to not cooperate, but the best outcome for the group is obtained when each person cooperates. Can cooperation develop from such a situation?

Axelrod (1984; cf. Hofstadter 1985) investigated that question by soliciting people to participate

in a series of PD games (social dilemmas are often referred to as games). Each person submitted a strategy for choosing to cooperate or not over a series of interactions with the other strategies. Each interaction would result in points being awarded to each strategy, and the strategy that generated the most points won. The winning strategy was Tit for Tat. It was also the simplest strategy. The Tit for Tat strategy is to cooperate on the first turn, and then do whatever the other person just did (i.e., on turn  $x$ , Tit for Tat will do whatever its opponent did on turn  $x-1$ ).

Axelrod suggested that four qualities led to Tit for Tat’s success. First, it was a *nice* strategy, because it first cooperates, and Tit for Tat will cooperate as long as the other person cooperates. But when the other person does not cooperate, then it immediately retaliates. That is, it responds to noncooperation with noncooperation, which illustrates its second good quality. Tit for Tat is *provocable*, because it immediately reacts to noncooperation, rather than waiting to see what will happen next, or ignore the noncooperation. However, if the opponent goes back to cooperating, then Tit for Tat will also go back to cooperating. That is quality three: *forgiveness*. Tit for Tat will not continue punishing the other player for previous noncooperations. All that counts for Tit for Tat is what just happened, not the total amount of noncooperation that has happened. Finally, Tit for Tat has *clarity*, because it is simple to understand. A complex strategy can be confusing, so it may be misunderstood by opponents. If the opponent’s intentions are unclear, then noncooperation is best, because if or when a complex strategy is going to be cooperative cannot be predicted.

Thus, a cooperative strategy can be effective even when there are clear incentives for noncooperation. Furthermore, Axelrod did another computer simulation in which strategies were rewarded by reproducing themselves, rather than simply accumulating points. Thus, success meant that the strategy had more of its kind to interact with. Again, Tit for Tat was best, which further suggests that a cooperative strategy can be effective and can flourish in situations that seem to be designed for noncooperation.



**Bargaining and Fairness.** Bargaining and negotiation have received increasing attention in decision theory and research (e.g., Pruitt and Carnevale 1993), as has the issue of justice or fairness (e.g., Mellers and Baron 1993). These issues are involved in the “ultimatum game,” which involves two people and a resource (often a sum of money). The rules of the game are that one person proposes a division of the resource between them (a bargain), and the other person accepts or rejects the proposal. If the proposal is rejected, then both people get nothing, so the bargain is an ultimatum: this or nothing.

Expected utility (EU) theory suggests that the person dividing the resource should offer the other person just enough to get him to accept the bargain, but no more. Furthermore, EU suggests the person should accept any division, because any division will be more than zero, which is what the person will receive if the bargain is refused. However, typically the bargain is a fifty-fifty split (half of the resource to each person). Indeed, if people are offered anything less than a fifty-fifty split, they will often reject the offer, although that will mean they get nothing, rather than what they were offered, because that seems unfair.

In studies, people have evaluated these bargains in two ways. People can rate how attractive a bargain is (e.g., on a 1–7 scale). Possible divisions to be rated might be: \$40 for you, \$40 for the other person; \$50 for you, \$70 for the other person, and so on. Thus, bargains are presented in isolation, one after another, as if each was an individual case unrelated to anything else. This type of presentation is generally referred to as “absolute judgment” (Wever and Zener 1928).

Alternatively, people may evaluate bargains in pairs, and choose one. For example, do you prefer a bargain where you get \$40, and the other person gets \$40, or a bargain where you get \$50, and the other person gets \$70? Thus, the bargains are presented such that they can be compared, so people can see the relative outcomes. This type of presentation is generally referred to as “comparative judgment.”

Absolute and comparative judgment have different results in the “ultimate” bargaining game.

Blount and Bazerman (1996) gave pairs of participants \$10 to be divided between them. In absolute judgment (i.e., is this bargain acceptable?), the money holder accepted a minimum division of \$4 for him and \$6 for the other person. But asked in a comparative judgment format (i.e., do you prefer this bargain or nothing?), participants were willing to accept less (a minimum division of \$2.33 for the money holder and \$7.67 for the other person). This result suggests that considering situations involving the division or distribution of resources on a case-by-case basis (absolute judgment) may result in sub-optimal choices (relative to those resulting from comparative judgment) for each person involved, as well as the group as a whole.

Comparative and absolute judgment can be applied to social issues as adoption. There has been controversy about adoption, when the adopting parents have a different cultural heritage than the child being adopted. Some argue that a child should be adopted only by parents of the same cultural heritage as the child to preserve the child’s connection to his or her culture. That argument views this situation as an absolute judgment: should children be adopted by parents of a different cultural heritage or not?

However, there is an imbalance between the cultural heritages of the children to be adopted and those of the parents wanting to adopt. That imbalance creates the dilemma of what to do with children who would like to be adopted when there are no parents of the same cultural heritage wanting to adopt them. That dilemma suggests this comparative judgment: should children be adopted by parents of a different cultural heritage than their own, or should children be left unadopted (e.g., be brought up in a group home)?

The answers to these absolute and comparative judgments may differ, because the answer may be that a child should not be adopted by parents of a different cultural heritage as an absolute judgment, but as a comparative judgment the answer may be that a child should be adopted by such parents, despite the cultural differences, because having parents is better than not having parents. Thus, the best answer may differ depending on

how the situation is characterized. Such situations may involve more than one value (in this case, the values are providing parents for a child and preserving the cultural heritage that the child was born into). Typically, absolute judgements reflect an acceptance or rejection of one value, while comparative judgments reflect more than one value.

### GENERAL JUDGMENT AND DECISION-MAKING ISSUES

That the best solution for a situation can seem different if the situation is characterized differently is one of the most important issues in judgment and decision making. The theories mentioned above (Bayes's, EU, prospect, and rank-dependent) assume problem invariance. That is, they assume that people's judgments will not vary with how the problem is characterized. However, because the characterization of the problem affects how people frame the problem, people's decisions often do vary. (Tversky and Kahneman 1981).

An implication of this variability is that eliciting people's values becomes difficult (Baron 1997; 1998). However, different methods can produce contradictory results. For example, choice and matching tasks often reveal different values preferences. Choice tasks are comparative judgments: Do you prefer A or B? Matching tasks require participants to estimate one dimension of an alternative so that its attractiveness matches that of another alternative (e.g., program A will cure 60 percent of patients at a cost of \$5 million, what should B cost if it will cure 85 percent of patients?).

The difference produced by these tasks has been extensively examined in studies of *preference reversals* (Slovic and Lichtenstein 1983). Tversky, Sattath, and Slovic (1988) have suggested that the dimension of elicitation (e.g., probability or value) will be weighted most, so reversals can result from changing the elicitation dimension. Fischer and Hawkins (1993) suggested that preference reversals were the result of the compatibility between people's strategy for analyzing the problem and the elicitation mode. Preference reversals clearly occur, but the cause of them continues to be debated (cf. Payne, Bettman, and Johnson 1992).

Ideas about rationality have also been influenced by the variability of people's judgments. Generally, the idea of rationality originated with a theory and then examined whether people behaved in that way, rather than examining how people behave and then suggesting what is rational. That is, rationality has typically been examined based on a prescriptive theory, such as Bayes's or EU, about how people should make decisions rather than a descriptive theory based on how people actually process information. When studies resulted in judgments that were inconsistent with those prescriptive theories of decision making, researchers concluded that people often acted irrationally.

However, there is growing recognition that study participants may be thinking of situations differently than researchers have assumed (Chase, Hertwig, and Gigerenzer 1998), which has led several researchers to create theories about decision processing (e.g., Dougherty, Gettys, and Ogden in press; Gigerenzer, Hoffrage, and Kleinbolting 1991) and use those theories to address rationality issues rather than the reverse. Approaches that focus on processing have been present in decision theory for some time (cf. Brunswick 1952; Hammond 1955), but they have not been dominant in the field. The acknowledgement of multiple views of rationality coupled with the poor explanations of prescriptive theories for people's actual decision behavior may shift the emphasis to processing models.

Further consideration of the decision-making process has led to other questions that are garnering increased attention. For example, how do people make decisions within dynamic environments? Generally, people make decisions in a dynamic world (Brehmer 1990; Busemeyer and Townsend 1993; Diehl and Serman 1995), but many decision-making theories (such as those reviewed above) do not account for the dynamics of the environment. Also, how do people's emotions affect the decision-making process? Decisions can involve topics that evoke emotion or have emotional consequences (such as regret, Gilovich and Medvec 1995). Some decision theories have tried to include emotional considerations in decision

making (e.g., Bell 1982), but this topic deserves more attention. These questions, as well as the issues discussed above, will make decision theory and research an area of continued interest and relevance.

## REFERENCES

- Ajzen, I. 1996 "The Social Psychology of Decision Making." In E. T. Higgins and A. W. Kruglanski, *Social Psychology: Handbook of Basic Principles*. New York: Guilford.
- Arkes, H. R., R. Dawes, and C. Christensen 1986 "Factors Influencing the Use of a Decision Rule in a Probabilistic Task." *Organizational Behavior and Human Decision Making*, 37:93-110.
- Axelrod, R. 1984 *The Evolution of Cooperation*. New York: Basic Books.
- Baron, J. 1994 *Thinking And Deciding*, 2d ed. Cambridge, U.K.: Cambridge University Press.
- 1997 "Biases in the Quantitative Measurement of Values for Public Decisions." *Psychological Bulletin* 122:72-88.
- 1998 *Judgment Misguided: Intuition and Error in Public Decision Making*. Oxford: Oxford University Press.
- Bayes, T. (1764) 1958 "An Essay Towards Solving a Problem in the Doctrine of Chances." *Biometrika* 45:293-315.
- Bell, D. 1982 "Regret in Decision Making Under Uncertainty." *Operations Research* 30:961-981.
- Birnbaum, M. H. 1983 "Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem." *American Journal of Psychology* 96:85-94.
- Birnbaum, M. H., G. Coffey, B. A. Mellers, and R. Weiss 1992 "Utility Measurement: Configural-Weight Theory and the Judge's Point of View." *Journal of Experimental Psychology: Human Perception And Performance* 18:331-346.
- Birnbaum, M. H., and W. R. Mcintosh 1996 "Violations of Branch Independence in Choices Between Gambles." *Organizational Behavior And Human Decision Processes* 67:91-110.
- Blount, S., and M. H. Bazerman 1996 "The Inconsistent Evaluation of Absolute Versus Comparative Payoffs in Labor Supply and Bargaining." *Journal of Economic Behavior and Organization* 30:227-240.
- Brehmer, B. 1990 "Strategies in Real-Time, Dynamic Decision Making." In R. M. Hogarth, ed., *Insights in Decision Making: A Tribute To Hillel J. Einhorn* 262-279. Chicago: University of Chicago Press.
- Brunswick, E. 1952 *The Conceptual Framework of Psychology*. Chicago: University of Chicago Press.
- Busemeyer, J., D. L. Medin, and R. Hastie 1995 *Decision Making From a Cognitive Perspective*. San Diego: Academic.
- , and J. T. Townsend 1993 "Decision-Field Theory: A Dynamic-Cognitive Approach to Decision-Making in an Uncertain Environment." *Psychological Review* 100:432-459.
- Chase, V. M., R. Hertwig, and G. Gigerenzer 1998 "Visions of Rationality." *Trends in Cognitive Sciences*, 2:206-214.
- Clore, G. L., N. Schwarz, and M. Conway 1994 "Affective Causes and Consequences of Social Information Processing." In R. S. Wyer and T. K. Srull, eds., *Handbook of Social Cognition: 2nd Edition*. Hillsdale, N.J.: Erlbaum.
- Dawes, R. M. 1979 "The Robust Beauty of Improper Linear Models." *American Psychologist* 34:571-582.
- 1998 "Behavioral Decision Making and Judgment." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, 4th ed. Boston: McGraw-Hill.
- Diehl, E., and J. D. Sterman 1995 "Effects of Feedback Complexity on Dynamic Decision Making." *Organizational Behavior and Human Decision Processes* 62:198-215.
- Dougherty, M. R. P., C. F. Gettys, and E. E. Ogden (in press) "MINERVA-DM: A Memory Processes Model for Judgments of Likelihood." *Psychological Review* 106:180-209.
- Einhorn, H. J. 1972 "Expert Measurement and Mechanical Combination." *Organizational Behavior and Human Performance* 13:171-192.
- Fischer, G. W., and S. A. Hawkins 1993 "Strategy Compatibility, Scale Compatibility, and the Prominence Effect." *Journal of Experimental Psychology: Human Perception and Performance* 19:580-597.
- Fong, G. T., D. H. Krantz, and R. E. Nisbett 1986 "The Effects of Statistical Training on Thinking About Everyday Problems." *Cognitive Psychology* 18:253-292.
- Gigerenzer, G., and U. Hoffrage 1995 "How to Improve Bayesian Reasoning without Instruction: Frequency Formats." *Psychological Review* 102:684-704.

- , and H. Kleinbolting 1991 “Probabilistic Mental Models: A Brunswickian Theory of Confidence.” *Psychological Review* 98:506–528.
- Gigerenzer, G., Z. Swijtink, T. Porter, L. Daston, J. Beatty, and L. Kruger 1989 *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge, U.K.: Cambridge University Press.
- Gilovich, T., and V. H. Medvec 1995 “The Experience of Regret: What, Why, and When.” *Psychological Review* 102:379–395.
- Hacking, I. 1975 *The Emergence of Probability*. Cambridge, U.K.: Cambridge University Press.
- 1990 *The Taming of Chance*. Cambridge, U.K.: Cambridge University Press.
- Hammond, K. R. 1955 “Probabilistic Functioning and the Clinical Method.” *Psychological Review* 62:255–262.
- 1998 *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Justice*. New York: Oxford.
- Hardin, G. R. 1968 “The Tragedy of the Commons.” *Science* 162:1243–1248.
- Hastie, R., and N. Pennington 1991 “Cognitive and Social Processes in Decision Making.” In L. B. Resnick, J. M. Levine, and S. D. Teasley, eds., *Perspectives on Socially Shared Cognition*, 308–327. Washington, D.C.: APA.
- Hilton, D. J., and B. R. Slugoski 1999 “Judgment and Decision-Making in Social Context: Discourse Processes and Rational Inference.” In T. Connolly, ed., *Judgment and Decision-Making: An Interdisciplinary Reader* 2d ed. Cambridge, U.K.: Cambridge University Press.
- Hofstadter, D. R. 1985 “The Prisoner’s Dilemma Computer Tournaments and the Evolution of Cooperation.” In D. R. Hofstadter, *Metamagical Themas: Questing for the Essence of Mind and Pattern*, 715–734. New York: Basic Books.
- Josephs, R. A., R. B. Giesler, and D. H. Silvera 1994 “Judgment By Quantity.” *Journal of Experimental Psychology: General* 123:21–32.
- Kahneman, D., J. L. Knetsch, and R. H. Thaler 1990 “Experimental Tests of the Endowment Effect and the Coase Theorem.” *Journal of Political Economy* 98:1325–1348.
- , P. Slovic, and A. Tversky 1982 *Heuristics and Biases: Judgments Under Uncertainty*. Cambridge, U.K.: Cambridge University Press.
- , and A. Tversky 1972 “On Prediction and Judgment.” *ORI Research Monograph*, 12.
- 1979 “Prospect Theory: An Analysis of Decisions Under Risk.” *Econometrica* 47:263–291.
- Kubovy, M. 1977 “Response Availability and the Apparent Spontaneity of Numerical Choices.” *Journal of Experimental Psychology: Human Perception and Performance* 3:359–364.
- Libby, R. 1976 “Man Versus Model of Man: Some Conflicting Evidence.” *Organizational Behavior and Human Performance* 16:1–12.
- Lopes, L. L. 1994 “Psychology and Economics: Perspectives on Risk, Cooperation, and the Marketplace.” *Annual Review of Psychology* 45:197–227.
- Luce, R. D., and P. C. Fishburn 1991 “Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles.” *Journal of Risk and Uncertainty* 4:29–59.
- Manktelow, K. I., and D. E. Over 1993 *Rationality: Psychological and Philosophical Perspectives*. London: Routledge.
- Markovits, H., and G. Nantel 1989 “The Belief-Bias Effect in the Production and Evaluation of Logical Conclusions.” *Memory & Cognition* 17:11–17.
- Mellers, B. A., and J. Baron 1993 *Psychological Perspectives on Justice: Theory and Applications*. New York: Cambridge University Press.
- Mellers, B. A., A. Schwartz, and A. D. J. Cooke 1998 “Judgment and Decision Making.” *Annual Review of Psychology* 49:447–477.
- Nisbett, R. E., D. H. Krantz, D. Jepson, and Z. Kunda 1983 “The Use of Statistical Heuristics in Everyday Inductive Reasoning.” *Psychological Review* 90:339–363.
- Payne, J. W., J. R. Bettman, and E. J. Johnson 1992 “Behavioral Decision Research: A Constructive Processing Perspective.” *Annual Review of Psychology* 49:447–477.
- Pelham, B. W., T. T. Sumarta, and L. Myaskovsky 1994 “The Easy Path from Many to Much: The Numerosity Heuristic.” *Cognitive Psychology* 26:103–133.
- Pruitt, D. G., and P. J. Carnevale 1993 *Negotiation in Social Conflict*. Pacific Grove, Calif.: Brooks/Cole.
- Rapoport, A., and A. M. Chammah 1965 *Prisoner’s Dilemma: A Study in Conflict and Cooperation*. Ann Arbor, Mich.: University of Michigan Press.
- Samuelson, W., and R. Zeckhauser 1988 “Status-Quo Bias in Decision Making.” *Journal of Risk and Uncertainty* 1:7–59.

- Schwarz, N., and G. L. Clore 1983 "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States." *Journal of Personality and Social Psychology* 45:513-523.
- Shapira, Z. 1995 *Risk Taking: A Managerial Perspective*. New York: Russell Sage Foundation.
- Sherif, M., D. Taub, and C. I. Hovland 1958 "Assimilation and Contrast Effects of Anchoring Stimuli on Judgments." *Journal of Experimental Psychology* 55:150-155.
- Slovic, P., and S. Lichtenstein 1983 "Preference Reversals: A Broader Perspective." *American Economic Review* 73:596-605.
- Slugoski, B. R., and A. E. Wilson 1998 "Contribution of Conversation Skills to the Production of Judgmental Errors." *European Journal of Social Psychology* 28:575-601.
- Stanovich, K. E., and R. F. West 1998 "Individual Differences in Rational Thought." *Journal of Experimental Psychology: General* 127:161-188.
- Stevenson, M. K., J. R. Busemeyer, and J. C. Naylor 1990 "Judgment and Decision-Making Theory." In M. D. Dunette and L. M. Hough, eds., *Handbook of Industrial and Organizational Psychology*, 2d ed., vol. 1, 283-374. Palo Alto, Calif.: Consulting Psychologists Press.
- Tversky, A., and D. Kahneman 1981 "The Framing of Decisions and the Psychology of Choice." *Science* 211:453-458.
- 1982 "Evidential Impact of Base Rates." In D. Kahneman, P. Slovic, and A. Tversky, eds., *Heuristics and Biases: Judgments Under Uncertainty*. Cambridge, U.K.: Cambridge University Press.
- 1983 "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment." *Psychological Review* 90:293-315.
- 1992 "Advances in Prospect Theory: Cumulative Representations of Uncertainty." *Journal of Risk and Uncertainty* 5:297-323.
- Tversky, A., S. Sattath, and P. Slovic 1988 "Contingent Weighting in Judgment and Choice." *Psychological Review* 95:371-384.
- von Neumann, J., and O. Morganstern 1947 *Theory of Games and Economic Behavior*, 2d ed. Princeton, N.J.: Princeton University Press.
- Wever, E. G., and K. E. Zener 1928 "The Method of Absolute Judgment in Psychophysics." *Psychological Review* 35:466-493.

EVAN THACKERAY PRITCHARD

## DEDUCTION/INDUCTION

See Experiments; Quasi-Experimental Research Designs; Scientific Explanation; Statistical Inference.

## DEMOCRACY

Democracy is one of the most important subjects in the social sciences. From the work of de Tocqueville in the early nineteenth century through the work of the best contemporary scholars, democracy has been studied closely and debated widely (Tocqueville 1969). Democracy has drawn this attention primarily because, in spite of the fact that it is quite rare historically, it has come to have enormous legitimacy in the eyes of many individuals worldwide. This has not always been the case. Democracy has been severely criticized by those on both the political right and left. But few scholars today question whether democracy is a social good.

Democracy is also important because many historically undemocratic countries have adopted it as a system of government. Many such changes have occurred only in the years since the Cold War (Huntington 1991). By 1994, over half the countries in the world had some form of democratic governance, a doubling of the number of nation-states so organized within 25 years (Lipset 1994).

At the core of most discussions of democracy is a common understanding that democracy is a method of governance or decision making for organizations or societies in which the members of that organization or society participate, directly or indirectly, in the decision making of that group. Further, members affect decision making to such an extent that they can be thought of as actually governing that organization or society. In short, democracy is a system of governance in which members control group decision making.

Not all considerations on democracy have shared this understanding. For example, those working in the Marxist tradition saw any state, democratic or not, as the expression of a class struggle. As such, any state was inherently undemocratic, absent the creation of a classless, communist, and hence truly democratic, society (Held

1995). Subsequent to the collapse of the Soviet Union and its affiliated states, and the full exposure of the failures of those regimes, such an extreme view has been almost entirely rejected.

The common understanding of democracy as participation by members or citizens in the decision making of an organization or society still leaves considerable room for dispute. Two issues are central: First, who are, or should be, considered members of the society? Second, what does, or should, constitute a minimum level of control over decision making by members for a system to be thought of as democratic? In short, how much participation is necessary for a system to be democratic? These questions are not simply matters of empirical observation of the world, but also matters for moral and political philosophy.

Three additional factors add to the difficulty of approaching democracy as a field of sociological research. First, analysts of democracy all too often use different definitions of democracy, or fail to define democracy clearly. Democratic systems of governance can be characterized by many attributes—frequency of member participation, the form of member participation, and so forth. Establishing whether a particular system of governance is democratic, involves making decisions about which attributes are essential to a democratic system. Where there is no specific definition of democracy or where definitions conflict, evaluating research on democracy can be tasking (Macpherson 1972).

A second factor increasing the difficulty of this subject matter is that democracy is a system of governance found in many different kinds of collectivities, including states, formal organizations, and informal groups. It is thus necessary to be cautious in applying models, findings, and relationships across different types of collectivities. When general propositions about democracy are advanced, it is important to evaluate those propositions at multiple levels of analysis.

A third difficulty is that social scientists are interested in democracy not just for its own sake, but also because it is thought to be associated with other critical issues. Many important questions

involve considerations of democracy: What is the effect of democracy on the success of organizations and nation-states? Does democracy promote individual liberty? What is the effect of democracy on income inequality and social stratification? What is the relationship between democracy and civil society? Can market economies prosper in the absence of democracy? All these and more force sociologists to consider the consequences, as well as the causes, of democracy.

While the difficulties of studying democracy are daunting, much significant work has been done in this field. Democracy has been studied as an outcome and as a cause, and has been studied at both the level of the nation-state, and at the level of the organization.

To begin with, there is work in the sociology of democracy on the question of who is or should be members or citizens of a democratic polity. Most systems commonly thought to be democratic have throughout history excluded some portion of those subject to the will of the democracy from participation in the decision-making process. Such exclusion has occurred on the basis of race, sex, income, relationship to property, criminal status, mental health, religion, age, and other characteristics. While use of many of these categories as a justification for excluding individuals from participation has declined in recent times, others remain, and there is continuing disagreement about the moral and political bases for excluding or including specific groups or categories of individuals. Migrant workers in Western Europe, for example, are subject to the action of the state on a long-term basis and yet remain excluded from full political participation in those states (Brubaker 1989). At the level of the organization, there are individuals affected by the decisions of the organization who may have little or no say in the decisions that affect them. These can include individuals both outside the organization (“stakeholders”) and inside the organization. Regarding affected individuals outside the organization, there are occasional movements to increase the power of stakeholders over the decisions of the organization (Nader, Green, and Seligman 1976). While unsuccessful at a general level, there has been a shift toward permitting

stakeholders increased access to legal redress, for example, in class action lawsuits and in environmental litigation. And regarding those within the organization, one way to understand the management trends toward “total quality management,” participatory management, and economic democracy is that they are attempts to increase democratic participation of workers in decisions that affect them (Jarley, Fiorito, and Delaney 1997).

Sociologists have often focused on the forms of influence and participation that individuals have in decision making. *Representative democracy* is that form of governance in which members of the organization or polity exercise their control over the organization through the regular election of members of a decision-making body. It is traditional to view representative democracies as democratic because they provide for the expression of interests through the election of representatives. For example, in the United States, citizens directly elect senators and representatives to the U.S. Congress, which in turn makes political decisions about the actions of the federal government. Theorists since the Enlightenment have argued that representative democracy is an appropriate means for conveying participation in decision making in large-size organizations, where individuals are thought not to be able to participate in all decisions (Hobbes 1968; Locke 1980; Mill 1962; Rousseau 1977).

But this view has been attacked by numerous critics, many of them sociologists. One of the most scathing criticisms, building on work by Mosca, is the analysis of democracy by Michels (Mosca 1939). Michels’s argument is that in any large organization (and, by extension, in any nation-state), a democratic system of governance inevitably leads to the rise of an oligarchy, and worse, to an oligarchy whose leaders have interests that differ from those of the ordinary members or citizens (Michels 1949). Why is this inevitable? In every instance of large democratic organization, Michels argues, oligarchy arises as a result of the organization’s requirement for experienced, skilled leaders. Experience in leadership, however, tends to give leaders access to key organizational resources, such as mailing lists, publicity, and greater experience, that are significant resources that the leadership

can use to return themselves to office year after year. And as leaders remain in office over an extended time, their interests and attitudes are likely to diverge from those of members. The divergence of interests is a result of the changed work and social experiences that accrue to leaders. Hence Michels, while arguing that formal organization is necessary for social life, and especially for politics, also believes that democracy in such organizations is essentially impossible.

Michels’s analysis has been taken very seriously in the social sciences, and there is some supporting evidence for his propositions. Weber described, and Hecló and Wilson separately concede, that there is a tendency for the civil service and bureaucracy to become unresponsive to the wishes of the people, as their experiences and needs differ from the people (Hecló 1977; Weber 1978; Wilson 1989). Lincoln and Zeitz have shown that as unions tend to get more professional, there is less member participation in decision making (Lincoln and Zeitz 1980). And for both unions and social movement organizations, Michels’s critique is taken so seriously as to generate sometimes drastic proposals for counteracting the oligarchical tendencies in these organizations (Kochan 1980). Piven and Cloward argue that reform movements of the poor should waste few resources on creating long-lasting organizations, but should instead create massive and disruptive protests (Piven and Cloward 1977).

However, Michels is not without his critics. Nyden argues that democratic unions are possible (Nyden 1985). Weber himself, who was Michels’s teacher, was critical of his conclusions. Michels overstated the case, Weber argued, because he insisted on relying upon too pure or strict a definition of democracy. Having started with such an idealistic vision of democracy, Michels was bound to find that reality comes up short (Scaff 1981).

That too pure a definition of democracy can lead to a misplaced understanding of how democracy works, and a failure to appreciate its achievements, is the key assumption behind the most significant defense of democracy in the 1950s and 1960s—the *pluralist* account of democracy. Dahl’s

account defends democracy by admitting its weakness: voting in elections is not a terribly effective system for ensuring that the will of the people will be carried out. Instead, Dahl focuses attention on whether non-electoral forms of influence can yield democratic decision making in keeping with the wishes and interests of the public. Interest groups thus become not the bane, but the hope of democracy. Through lobbying in all its forms, interest groups are able to exert power and influence over decision making beyond elections; if they do this, then the system, with all its flaws, can be considered democratic (Dahl 1961).

A problem with the pluralist view is that not all groups in society may be able effectively to form interest groups to pursue their goals. Olson, in an early effort in what is now known as *rational choice theory*, argued that individuals must be assumed to be rational, and that rational individuals will not contribute to the formation of interest groups when they will obtain the benefits achieved by the interest group anyway. This is the *free rider* or *collective action* problem: if an interest group lobbies for clean air, and a person cannot be denied clean air because he or she does not belong to the interest group, why should that person contribute to the group? Only those interest groups with a particularly small constituency or those interest groups who are able to use “special” incentives—those available only to members of the group—to attract contributors will be able to form to lobby to advance their interests. Groups representing weak and powerless individuals may be unable to supply such special incentives (Olson 1971).

Olson’s pessimism about the chances for the disadvantaged to gain a voice in decision making has been the focus of much attention. Oliver and Marwell suggest that social movements are more likely to be formed as interest groups grow in size (Oliver and Marwell 1988). Knoke argues that the use of selective incentives may attract apathetic members, whereas a focus on the goal of lobbying may attract highly active members, thereby creating more effective organizations (Knoke 1988). Clemens points out that as interest groups come into existence, they are themselves models or templates for others to imitate. Those templates will

then increase the likelihood of the formation of more interest groups (Clemens 1997). These criticisms of Olson’s analysis of the collective action problem may actually serve to strengthen the pluralist account of democracy.

Yet many sociologists remain deeply critical of the pluralist account. Domhoff argues that pluralism is flawed not because the collective action problem retards the capacity of the disadvantaged to organize. Rather, pluralism is flawed because in the United States, and in other industrialized democracies, there is a governing class (Domhoff 1998). This governing class is composed of elites from business, the social upper class, and those in charge of organizations, both within and outside government, that are powerfully involved in the formation of public policy. While Domhoff admits that there is some conflict within these groups, he views them as cohesive in their opposition to the interests of the poor and the working class. Through their control of important organizations, through the strength of their social ties, and through the use of agenda-setting, the governing class achieves enormous power. And, Domhoff argues, the governing class is able to use that power consistently to defeat the interests of the majority.

Other critics of the pluralist account have drawn attention to the relationship between social class and voting. For some years, it appeared that class-based voting in the United States appeared to be declining (Clark, Lipset, and Rempel 1993; Manza, Hout, and Brooks 1995). Yet some scholars believe that class remained a significant factor in voting behavior (Burnham 1981). Piven and Cloward argued that the pluralist account failed because there was a systematic pattern to who voted and who did not. Because the poor and the working classes disproportionately failed to vote in elections, they were inadequately represented in the competition between interests; hence the poor were excluded from the pluralist democracy (Piven and Cloward 1988). It has been shown that class remains a powerful determinant of how people vote, even if the working class no longer votes for the Democratic party in the United States with as great a frequency as it did in the years immediately after WWII (Hout, Brooks, and Manza 1995).



Researchers in the 1970s and 1980s criticized pluralism for adopting a definition of democracy that was too satisfied with the status quo of only limited participation in decision making. These participatory democratic theories emphasized that as members participate in decision making, they learn more about the criteria that need to be used in effective decision making and become better at making decisions. But members also, it was argued, become better at evaluating the choice of candidates where representatives must be elected. Accordingly, empirical researchers began to investigate the causes and consequences of increased participation in decision making (Finley 1973; Pateman 1970).

In keeping with this view, researchers increasingly moved to consider cases of participation throughout society (Alford and Friedland 1975). Participation in decision making in unions, community organizations, municipalities, and protest groups has been analyzed and touted (Gans 1989; Cole 1975). In organization theory, research on different forms of worker participation in organizational decision making has gained increasing prominence. Total quality management, participatory management, and worker control all have been studied closely, not just for improvement in productivity and quality, but also for democracy. Much of this research emphasizes how democracy is consistent with both effectiveness and the improvement of the condition of workers (Jarley, Fiorito, and Delaney 1997). Some have even come to see the spread of democratic management as inevitable, although this is almost certainly exaggerated (Collins 1997). Yet many have worked to show that democratic systems of management are more broadly possible than has been thought, although the conditions under which such systems of decision making can be created and maintained remain under debate (Burawoy 1982; Kanter 1983).

Renewed definitions of democracy beg a central question in democratic research: Where do democracies come from? The question has been most closely studied for nation-states. Lipset identifies a set of central conditions that are associated with the rise of democracy in nation-states. The presence of a market economy appears to be a

necessary, although not a sufficient, condition for democracy. A minimum level of economic development is associated with democracy, although a key debate is the extent to which development leads to democracy (Bollen and Jackman 1995; Muller 1995). Also associated with democracy is a political culture in which the tolerance for the rights of others is recognized. Finally, countries with Protestant religious traditions have been more likely to be democratic, though the significance of that effect may be fading in the recent transition to democracy (Lipset 1994).

One of the most significant contributions to the account of the origins of democracy is that by Moore (Moore 1966). His analysis, standing in contrast to a Marxist emphasis on the role of the working class as a force in history, identified the relationship between peasant and lord prior to capitalism as the critical factor in determining whether a society became democratic or autocratic. In countries such as China, France, Japan, or Germany, repressive control over peasants by a dominant class led either to revolution or to continued autocracy. In China, revolution led to Communist autocracy; in France, because of the existence of a commercial class, revolution led, through fits and starts, to democracy. In Japan and Germany, the failure or absence of revolution led to continued dominance of repressive classes, leading to the rise of fascist regimes. Moore argued that in a country such as England, however, the greater status of labor, coupled with the nobility's increasing dependence on market-based agriculture, led to an eventual democratic solution to social conflict.

Moore's work has provoked considerable criticism and extension (Ross et al. 1998). Downing has shown conclusively that the nature of military conflicts affects the success of democracy in a country. How a nation fights its wars, and how often it must fight, is critically determinative of the need for repression in the mobilization of men and weapons to fight. England's peculiar move toward democracy is thus critically dependent on its position as an island nation, free from the necessity to fight long-term, massive land wars on the continent of Europe, and the necessity to

maintain a state and military administrative structure capable of that task (Downing 1992). Friedman has shown that former British colonies tend to be democratic, whereas countries ruled by Leninist parties tend to remain autocratic (Friedman 1998).

Students of social movements have tended to argue that social movements are significant sources of democracy (Giugni, McAdam, and Tilly 1998). There is evidence that social movements can push the transition to democracy faster (Hipsher 1998; Sandoval 1998). Certainly this is consistent with Tilly's theoretical model of democratization, in which he argues that social conflict, as embodied in social movements, and where mediated by third parties, can lead to the creation of rights essential for democracy (Tilly 1998). But surely the effects of social movements on democracy are contingent on many factors, and sociologists must be careful not to assume that the outcome of social protest will be democratization (Melucci and Lyyra 1998). Clearly this has not always been the case.

The sources of democracy in organizations is an understudied area. Two findings are worth noting. Knoke has argued that in the present-day United States, a minimum level of democratic procedure is just a part of the institutional building blocks from which organizations are constructed; in short, organizations such as unions may have democratic procedures simply because everybody expects them to have those procedures (Knoke 1990). And, returning to the tension between democracy and effectiveness identified by Michels, Jarley has found that the causes of democratic procedures in unions are independent of those which drive bureaucratization (Jarley, Fiorito, and Delaney 1997). Yet much more systematic work needs to be done in this area.

The other side of the coin in the study of democracy is the question of the relationship between democracy and other core subjects of sociological interest. The relationship between democracy and equality is a central issue and has been a focus of research since de Tocqueville (Tocqueville 1969). In recent years, research has centered around this specific question: Does democracy promote or retard income inequality in nation-states? There is

some evidence that democracy does not increase inequality, at least directly, and it might lead to increased equality (Bollen and Jackman 1985; Muller 1988).

Another question is whether democracies can be effective. The central issue, echoing Michels, is whether or not organizations, such as unions or parties or, for that matter, businesses, can be successful in competitive environments against organizations that are autocratically run. The evidence is conflicting. Some argue that democracy and effectiveness are in conflict in the context of unions (Lipset, Trow, and Coleman 1959; Piven and Cloward 1977). Others argue that democracy leads to effectiveness in achieving goals (Stepan-Norris and Zeitlin 1994). As yet, there seems to be no definitive answer, and the issue certainly merits more research.

Finally, the subject of democracy has been intimately tied in the 1990s to two related subjects.

The first is the subject of globalization. As the world has become more closely connected, as communications technologies radically change again, and as the world economy has grown larger, some have raised questions about the implications of this trend for democracy. Held, for example, has argued that decisions that affect citizens are increasingly being made at a level beyond that of the nation-state; in supra-national organizations and in the international economy. As a consequence of this globalization, the extent of democratic control over decisions is seen to have weakened (Held 1995). Others have criticized this argument, and the subject is still very much subject to research and debate (Hirst and Thompson 1996).

The second related subject is civil society. A focus of scholarly attention in part because of the demands for it from those who have emerged from socialist rule, civil society is commonly conceived as space for associational activity between the state and the individual (Gellner 1994). Many now see organizations and associations, independent of the state, as crucial to democracy, constituting a critical element of democratic society (Streeck and Schmitter 1985). Certainly, they are not the same: as Hall puts it, "Democracy can be

decidedly uncivil" (Hall 1995). But democracy depends on civil society (Somers 1993). This view echoes Tocqueville's assertion that the knowledge of how to combine is fundamental to democracy (Tocqueville 1969). Already attracting significant attention, much room remains to answer questions about the relationship between civil society and democracy.

In conclusion, it is well to remember that there are many forms of democracy: those with weaker or stronger civil liberties; those with weaker or stronger civil societies; those with weaker or stronger tolerance for diversity. Nor can it be assumed that these different forms are internally consistent: the rights of the community to choose what it wishes to be, and the rights of the individual to live as he or she wishes, are not easily reconciled.

It is also well to remember that democracy is not inevitable (Berger 1992). Neither, we should also recall, is democracy a simple outcome that, once achieved, is a permanent condition (Friedman 1998). Democracy can be strengthened; democracy can be weakened. And it can, as it has in the past, disappear. While democracy is today in the ascendant, the lessons of the French Revolution and of Weimar Germany should not be forgotten; although in both instances democracy was regained, it was not regained quickly or without cost. And in Germany, as in Japan, democracy was not regained from within, but imposed from without. History should teach us that we still have much to learn about democracy.

(SEE ALSO: *Capitalism, Development, Inequality, Civil Society, Individualism in Less Developed Countries, Political Sociology, Rational Choice Theory.*)

REFERENCES

Alford, Robert R., and Roger Friedland 1975 "Political Participation and Public Policy." *Annual Review of Sociology* 1:429-479.

Berger, Peter 1992 "The Uncertain Triumph of Democratic Capitalism." *Journal of Democracy* 3:7-17.

Bollen, Kenneth A., and Robert W. Jackman 1985 "Political Democracy." *American Sociological Review* 50:438-457.

— 1995 "Income Inequality and Democratization Revisited: Comment on Muller." *American Sociological Review* 60:983-989.

Brubaker, William Rogers 1989 *Immigration and the Politics of Citizenship in Europe and North America*. Lanham, Md.: University Press of America; German Marshall Fund of the United States.

Burawoy, Michael 1982 *Manufacturing Consent*. Chicago: University of Chicago Press.

— 1997 "Review Essay: The Soviet Descent into Capitalism." *American Journal of Sociology* 102:1430-1444.

Burke, Edmund 1955 *Reflections on the Revolution in France*. Indianapolis: Bobbs-Merrill.

Burnham, Walter Dean 1981 *The Current Crisis in American Politics*. New York: Oxford University Press.

Clark, Terry Nichols, Seymour Martin Lipset, and Michael Rempel 1993 "The Declining Political Significance of Class." *International Sociology* 8:323-333.

Clemens, Elisabeth S. 1997 *The People's Lobby: Organizational Innovation and the Rise of Interest Group Politics in the United States, 1890-1925*. Chicago: University of Chicago Press.

Cole, Richard L. 1975 "Citizen Participation in Municipal Politics." *American Journal of Political Science* 19:761-782.

Collins, Denis 1997 "The Ethical Superiority and Inevitability of Participatory Management as an Organizational System." *Organization Science*.

Dahl, Robert A. 1961 *Who Governs?* New Haven: Yale University Press.

Domhoff, G. William 1998 *Who Rules America? Power and Politics in the Year 2000*. Mountain View, Calif.: Mayfield Publishing Company.

Downing, Brian M. 1992 *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton, N.J.: Princeton University Press.

Finley, M. I. 1985 *Democracy: Ancient and Modern*. London: Hogarth Press.

Friedman, Edward 1998 "Development, Revolution, Democracy, and Dictatorship: China versus India?" In Theda Skocpol, ed., *Democracy, Revolution, and History*, 102-123. Ithaca, N.Y. and London: Cornell University Press.

Gans, Herbert J. 1989 "Comment on Varese's Review of Middle American Individualism: The Future of

- Liberal Democracy." *American Journal of Sociology* 95:188–190.
- Gellner, Ernest 1994 *Conditions of Liberty: Civil Society and its Rivals*. London: Penguin Books.
- Giugni, Marco G., Doug McAdam, and Charles Tilly 1998 "From Contention to Democracy." Lanham, Md.: Rowman and Littlefield.
- Hall, John A. 1995 "In Search of Civil Society." In John A. Hall, ed., *Civil Society: Theory, History, Comparison* 1–27. Cambridge, England: Polity Press.
- Hartz, Louis 1948 *Economic Policy and Democratic Thought: Pennsylvania 1776–1860*. Cambridge, Mass.: Harvard University Press.
- Hecllo, Hugh 1977 *A Government of Strangers*. Washington, D.C.: Brookings Institution.
- Held, David 1995 *Democracy and the Global Order: From the Modern State to Cosmopolitan Governance*. Stanford, Calif.: Stanford University Press.
- Hipsher, Patricia L. 1998 "Democratic Transitions and Social Movement Outcomes: The Chilean Shantytown Dwellers' Movement in Comparative Perspective." In Marco G. Giugni, Doug McAdam, and Charles Tilly, eds., *From Contention to Democracy* 149–168. Lanham, Md.: Rowman and Littlefield.
- Hirst, Paul, and Grahame Thompson 1996 *Globalization in Question: The International Economy and the Politics of Governance*. Cambridge, England: Polity Press.
- Hobbes, Thomas 1968 *Leviathan*. New York: Penguin Books.
- Hout, Michael, Clem Brooks, and Jeff Manza 1995 "The Democratic Class Struggle in the United States, 1948–1992." *American Sociological Review* 60:805–828.
- Huntington, Samuel P. 1991 *The Third Wave: Democratization in the Late Twentieth Century*. Norman, Okla.: University of Oklahoma.
- Jarley, Paul, Jack Fiorito, and John Thomas Delaney 1997. "A Structural Contingency Approach to Bureaucracy and Democracy in U.S. National Unions." *Academy of Management Journal* 40:831–861.
- Kanter, Rosabeth Moss 1983 *The Change Masters*. New York: Touchstone.
- Knoke, David 1988 "Organizational Incentives." *American Sociological Review* 53:311–329.
- 1990 *Organizing for Collective Action: The Political Economies of Associations*. New York: de Gruyter.
- Kochan, T. H. 1980 *Collective Bargaining and Industrial Relations: From Theory to Policy and Practice*. Homewood, Ill.: Irwin.
- Lincoln, J. R., and G. Zeitz 1980 "Organizational Properties from Aggregate Data: Separating Individual and Structural Effects." *American Sociological Review* 45:391–408.
- Lipset, Seymour Martin 1994 "The Social Requisites of Democracy Revisited: 1993 Presidential Address." *American Sociological Review* 59:1–22.
- , Martin A. Trow, and James S. Coleman 1959 *Union Democracy*. Glencoe, Ill.: Free Press.
- Locke, John 1980 *The Second Treatise of Government*. Indianapolis, Ind.: Hackett Publishers.
- Macpherson, C. B. 1972 *The Real World of Democracy*. New York: Oxford University Press.
- Manza, Jeff, Michael Hout, and Clem Brooks 1995 "Class Voting in Capitalist Democracies since WWII." *Annual Review of Sociology* 21:137–163.
- Melucci, Alberto, and Timo Lyyra 1998 "Collective Action, Change, and Democracy." In Marco G. Giugni, Doug McAdam, and Charles Tilly, eds., *From Contention to Democracy* 203–228. Lanham, Md.: Rowman and Littlefield.
- Michels, Robert 1949 *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy*. New York: Free Press.
- Mill, John Stuart 1962 *Considerations on Representative Government*. South Bend, Ind.: Gateway Editions.
- Moore, Barrington 1966 *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Mosca, Gaetano 1939 *The Ruling Class*. New York: McGraw-Hill.
- Muller, Edward N. 1988 "Democracy, Economic Development, and Income Inequality." *American Sociological Review* 53:50–68.
- 1995 "Economic Determinants of Democracy." *American Sociological Review* 60:966–982.
- Nader, Ralph, Mark Green, and E. J. Seligman 1976 *Constitutionalizing the Corporation: The Case for the Federal Chartering of Giant Corporations*. Washington, D.C.: Corporate Accountability Research Group.
- Nyden, Philip W. 1985 "Democratizing Organizations: A Case Study of a Union Reform Movement." *American Journal of Sociology* 90:1179–1203.
- Oliver, Pamela E., and Gerald Marwell 1988 "The Paradox of Group Size in Collective Action." *American Sociological Review* 53:1–8.

- Olson, Mancur 1971 *The Logic of Collective Action*. Cambridge, Mass.: Harvard University Press.
- Pateman, Carole 1970 *Participation and Democratic Theory*. Cambridge: Cambridge University Press.
- Piven, Frances Fox, and Richard Cloward 1977 *Poor People's Movements*. New York: Pantheon.
- 1988 *Why Americans Don't Vote*. New York: Pantheon.
- Ross, George, Theda Skocpol, Tony Smith, and Judith Eisenberg Vichniac 1998 "Barrington Moore's Social Origins and Beyond: Historical Social Analysis since the 1960s." In Theda Skocpol, ed., *Democracy, Revolution, and History* 1–21. Ithaca, N.Y. and London: Cornell University Press.
- Rousseau, Jean-Jacques 1977 *The Social Contract and Discourse on the Origin of Inequality*. New York: Washington Square Press.
- Sandoval, Salvador A. M. 1998 "Social Movements and Democratization: The Case of Brazil and the Latin Countries." In Marco G. Giugni, Doug McAdam, and Charles Tilly, eds., *From Contention to Democracy* 169–202. Lanham, Md.: Rowman and Littlefield.
- Scaff, Lawrence A. 1981 "Max Weber and Robert Michels." *American Journal of Sociology* 86:1269–1286.
- Schumpeter, Joseph A. 1976 *Capitalism, Socialism, and Democracy*. New York: Harper and Row.
- Somers, Margaret A. 1993 "Citizenship and the Place of the Public Sphere: Law, Community, and Political Culture in the Transition to Democracy." *American Sociological Review* 58:587–620.
- Stapan-Norris, Judith, and Maurice Zeitlin 1994 "Union Democracy, Radical Leadership, and the Hegemony of Capital." *American Sociological Review* 1995:829–850.
- Streeck, Wolfgang, and Philippe C. Schmitter 1985 "Community, Market, State—and Associations? The Prospective Contribution of Interest Governance to Social Order." In Wolfgang Streeck and Philippe C. Schmitter, eds., *Private Interest Government* 1–29. London: Sage Publications.
- Tilly, Charles 1998 "Where Do Rights Come From?" In Theda Skocpol, ed., *Democracy, Revolution, and History* 55–72. Ithaca, N.Y. and London: Cornell University Press.
- Tocqueville, Alexis de 1969 *Democracy in America*. Garden City, N.J.: Harper Perennial.
- Weber, Max 1978 "Parliament and Government in a Reconstructed Germany: A Contribution to the Political Critique of Officialdom and Party Politics." In

Guenther Roth and Claus Wittich, eds., *Economy and Society: An Outline of Interpretive Sociology* 1381–1469. Berkeley: University of California Press.

Wilson, James Q. 1989 *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.

ANDREW L. CREIGHTON

## DEMOGRAPHIC METHODS

Demographic methods are used to provide researchers and policymakers with useful information about the size and structure of human populations and the processes that govern population changes. A population, of course, may range in size from a small number of individuals surveyed locally to a large national population enumerated in periodic censuses to even larger aggregated entities.

We use demographic methods not only in purely demographic applications, but also in a variety of other fields, among them sociology, economics, anthropology, public health, and business. Demographers, like all researchers, must pay careful attention to the quality of data that enter into their analyses. Some circumstances under which we use these methods are more trying than others. In cases in which the data are viewed to be accurate and complete, the methods we use to analyze them are more straightforward than those that are applied to data of imperfect quality.

### DESCRIPTION OF DATA

First we must develop ways to describe our data in a fashion that allows the most important facts to leap out at us. As one example, let us examine a population's age structure or distribution. Simple descriptive statistics are doubtless helpful in summarizing aspects of population age structure, but demographers often use age "pyramids" to convey to an audience the youthfulness of a population, for example, or even to convey a rough sense of a nation's history.

We might note, for example, that 20 percent of Norway's population in 1997 was under the age of fifteen. In contrast, 39 percent of Mexico's population seven years earlier fell into this category. But it is perhaps more dramatic to create a visual display of these figures in the form of age pyramids.

An age pyramid is typically constructed as a bar graph, with horizontal bars—one representing each sex—emanating in both directions from a central vertical age axis. Age increases as one proceeds up the axis and the unit of the horizontal axis is either the proportion of the total population in each age group or the population size itself.

We see in panel A of figure 1 that Mexico's population is described by a very broad-based pyramid—actually, in two-dimensional space, a triangle, but by convention we refer to it as a pyramid—which reveals a remarkably large proportion of the population not yet having reached adulthood. The median age of this population is about twenty years. Such a distributional shape is common particularly among high fertility populations. In stark contrast we have the pyramid shown in panel B, representing the population of Norway. Rather than triangular, its age distribution is somewhat more rectangular in shape, which is typically seen among countries that have experienced an extended period of low fertility rates. It is easy to see that Norway's median age (about thirty-six years) is considerably higher than that of Mexico.

As mentioned above, not only can we examine the age structure of populations through the use of pyramids, but we can also gain much insight into a nation's history insofar as that history has either directly or indirectly influenced the size of successive birth cohorts. Note, for example, the age pyramid reflecting the population age structure of France on 1 January 1962 (figure 2).

In this figure, several notable events in France's history are apparent. We see (1) the military losses experienced in World War I by male birth cohorts of the mid to late 1890s, (2) the remarkable birth dearth brought about by that war (the cohort aged in their mid-forties or so in 1962), followed by (3) the return of prewar childbearing activity once the

war ended (the cohort who in 1962 were in their early forties), and a similar pattern revolving around World War II, during which (4) a substantial decline in births took place (the cohort around age twenty in 1962), succeeded by (5) a dramatic baby boom in the years immediately following (those around age fifteen in 1962).

### COMPARISON OF CRUDE RATES

Much of the work in which social scientists are engaged is comparative in nature. For example, we might seek to contrast the mortality levels of the populations of two countries. Surely, if the two countries in question have accurate vital registration and census data, this task would appear to be trivial. For each country, we would simply divide the total number of deaths (D) in a given year by the total population (P) at the midpoint of that year. Thus we would define the crude death rate (CDR) for country A as:

$$CDR^A = \frac{D^A[t,t+1]}{P^A[t+.5]} = \frac{\sum_{x=0,n}^{\omega-n} [{}_nM_x^A \cdot {}_n P_x^A]}{\sum_{x=0,n}^{\omega-n} {}_n P_x^A} \quad (1)$$

where  $t$  denotes the beginning of the calendar year,  $\omega$  is the oldest age attained in the population,  ${}_nM_x^A$  is the death rate of individuals aged  $x$  to  $x+n$ , and  ${}_n P_x^A$  is the number of individuals in that same age group.  ${}_nM_x^A$  and  ${}_n P_x^A$  are centered on the midpoint of the calendar year. The rightmost segment of this equation reminds us that the crude death rate is but the sum of the age-specific death rates weighted by the proportion of the population at each age.

Unfortunately, comparisons of crude death rates across populations can lead to misleading conclusions. This problem is in fact general to any sort of comparison based on crude rates that do not account for the confounding effects of factors that differentiate the two populations.

Let us examine the death rates of two very different countries, Mexico and Norway. The crude death rate (for both sexes combined) in Mexico in 1990, 5.2 per 1000, was approximately half that for

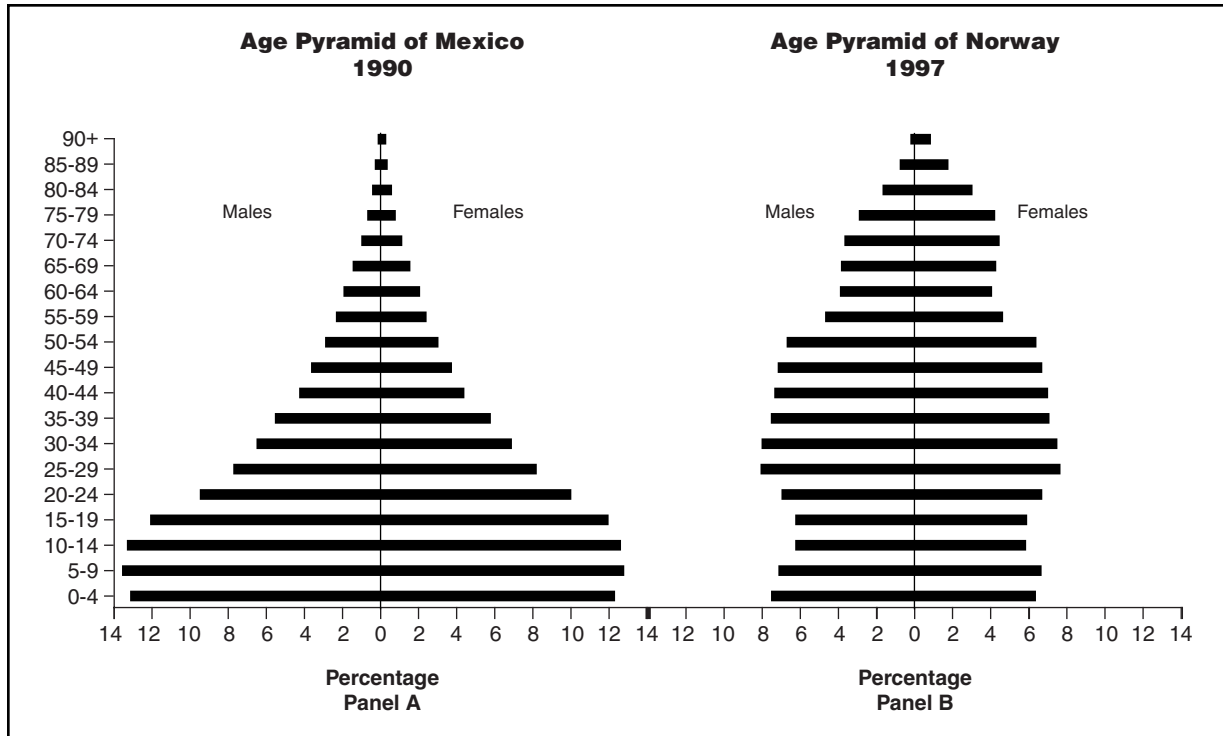


Figure 1

NOTE: Age pyramids for Mexico and Norway

Norway, 10.1, seven years later, in 1997. Knowing nothing else about these two countries, we might infer that the health of the Mexican population was considerably superior to that of the Norwegian population. However, once we standardize the crude death rates of the two countries on a common population age structure, we find just the reverse is true.

To accomplish this standardization, instead of applying the above equation to our data, we use the following:

$$ASCDR^A = \frac{\sum_{x=0,n}^{\omega-n} [{}_nM_x^A \cdot {}_nP_x^S]}{\sum_{x=0,n}^{\omega-n} {}_nP_x^S} \quad (2)$$

where  $ASCDR^A$  is the age-standardized crude death rate of country A, and  ${}_nP_x^S$  is the number of individuals in the standardized population of that same age group. Any age distribution may be chosen as the standard, however, it is common simply to use the average of the two proportionate

age distributions (i.e., each normalized to one in order to account for unequal population sizes).

Table 1 gives the death rates and number of persons for each country by five-year age groups (zero through four, five through nine, and so on through eighty and above). The first fact we glean from the table is that the mortality rates of the Norwegian population are substantially lower than those of their Mexican counterparts at virtually all ages. We see that nearly 40 percent of the Mexican population is concentrated in the three age groups having the lowest death rates (ages five through nineteen). Only half that proportion of the Norwegian population is found in the same age range. In contrast, only 6 percent of the Mexican population is above sixty years of age, an age range with which its highest level of mortality is associated. At the same time, about 20 percent of the Norwegian population is sixty or older. Compared with the Mexican crude death rate, then, the Norwegian rate is disproportionately weighted toward the relatively high age-specific death rates that exist in the older ages.

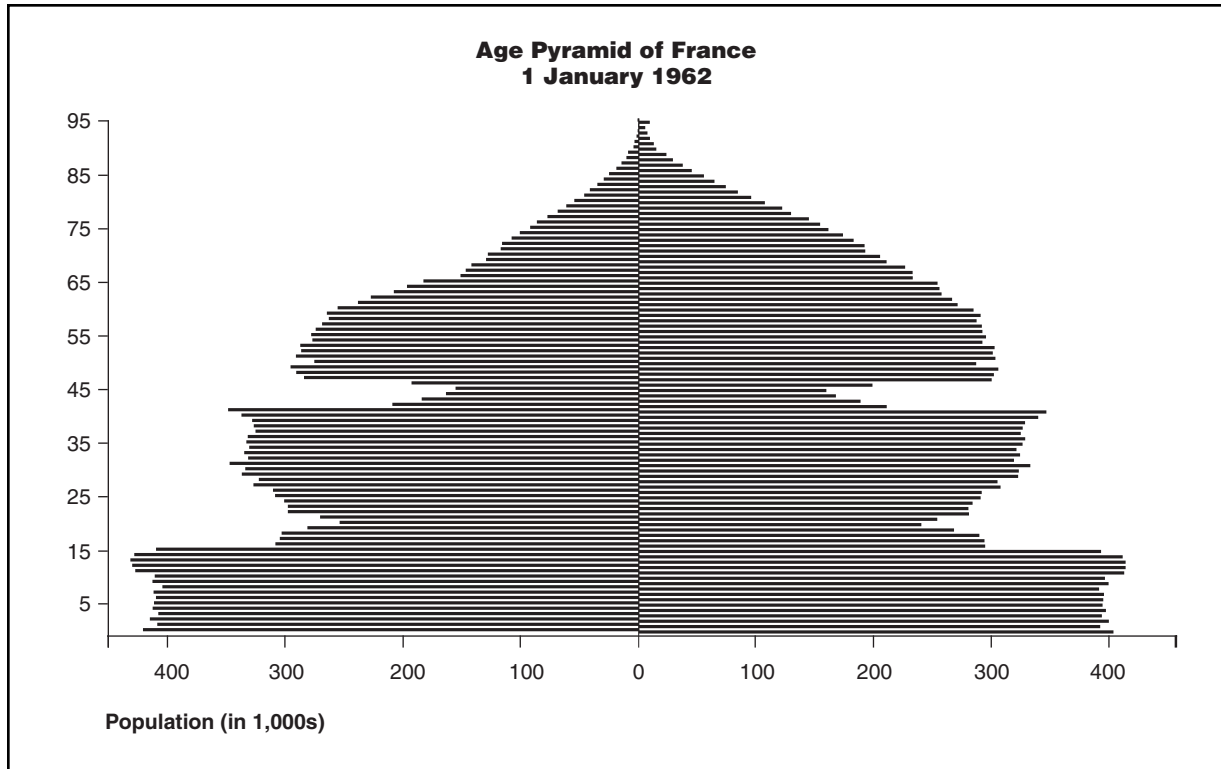


Figure 2

NOTE: Age pyramid France

Applying the standardization method described by the above equation to the average of Norway's and Mexico's proportionate age distributions, we find that the resulting crude death rates are consistent with what we would infer from the two series of age-specific mortality rates. Mexico's age-standardized crude death rate, 8.7 per 1000 population, is one-third greater than that of Norway, 6.6 per 1000.

### LIFE TABLES

The *life table* is a methodological device used by demographers and others to examine the life—either literal or figurative—of a particular duration-dependent phenomenon. The original application of the life table was to the mortality patterns of human populations. Today, the life table technique is applied to such diverse areas as contraceptive efficacy, marital formation and dissolution, and organizational failure. Thus it is a remarkably

general tool for examining the time-dependent survivorship in a given state.

To illustrate the use of the life table method, suppose we use as an example the mortality experience of the United States population in 1996. We might wish to derive the average number of years that an individual would live, subject to the series of age-specific death rates attributed to that population. To find the answer, we could construct a *complete life table*, as in table 2. It is complete in the sense that it is highly age-detailed, focusing on single years of age. This is in contrast to the *abridged life table*, which is usually constructed using five-year age groups. The abridged life table shown in table 3 refers also to the total population of the United States in 1996.

Most life tables that we see are called *period* or *current life tables*. They refer to a particular snapshot in time. Although they describe the mortality experience of an actual population, they do not



**Standardization using Death Rates and Population Distributions  
from Mexico (1990) and Norway (1997)**

Age Group	Mexico		STANDARD DISTRIBUTION (%)	Norway	
	Population in 1000s (% distribution)	Death Rate (per 1000)		Population in 1000s (% distribution)	Death Rate (per 1000)
0-4	10,257 (12.62)	8.40	9.75	303 (6.88)	1.06
5-9	10,627 (13.08)	0.61	9.96	301 (6.83)	0.19
10-14	10,452 (12.86)	0.52	9.43	264 (6.00)	0.14
15-19	9,723 (11.97)	0.99	9.00	265 (6.02)	0.48
20-24	7,877 (9.69)	1.50	8.24	299 (6.78)	0.54
25-29	6,444 (7.93)	1.88	7.86	343 (7.78)	0.68
30-34	5,420 (6.67)	2.21	7.17	338 (7.68)	0.77
35-39	4,607 (5.67)	2.88	6.46	319 (7.25)	0.94
40-44	3,519 (4.33)	3.80	5.72	313 (7.11)	1.55
45-49	2,990 (3.68)	5.30	5.28	303 (6.89)	2.45
50-54	2,408 (2.96)	7.30	4.73	286 (6.50)	3.88
55-59	1,906 (2.35)	11.32	3.48	203 (4.61)	6.19
60-64	1,621 (2.00)	15.17	2.97	174 (3.94)	10.31
65-69	1,191 (1.47)	23.57	2.74	177 (4.02)	16.55
70-74	832 (1.02)	33.27	2.53	177 (4.03)	28.66
75-79	594 (0.73)	54.53	2.13	156 (3.53)	57.11
80+	780 (0.96)	108.52	2.55	182 (4.14)	122.48

Table 1

 SOURCE: <http://www.ssb.no/www-open/english/yearbook/> and <http://www.census.gov/ipc/www/idbprint.html>

**Complete Life Table\* for the United States, 1996**

<i>Exact Age x</i>	${}_1q_x$	$l_x$	${}_1d_x$	${}_1L_x$	$T_x$	$e_x$
0	.00732	100,000	732	99,370	7,611,825	76.1
1	.00054	99,268	53	99,240	7,512,455	75.7
2	.00040	99,215	40	99,193	7,413,215	74.7
3	.00031	99,175	31	99,158	7,314,022	73.7
4	.00026	99,144	26	99,130	7,214,864	72.8
5	.00023	99,118	23	99,106	7,115,734	71.8
6	.00021	99,095	21	99,084	7,016,628	70.8
7	.00020	99,074	19	99,064	6,917,544	69.8
8	.00018	99,055	18	99,046	6,818,480	68.8
9	.00016	99,037	15	99,029	6,719,434	67.8
10	.00014	99,022	14	99,014	6,620,405	66.9
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
80	.05967	49,276	2,940	47,668	411,547	8.4
81	.06566	46,336	3,043	44,676	363,879	7.9
82	.07250	43,293	3,139	41,586	319,203	7.4
83	.08033	40,154	3,225	38,403	277,617	6.9
84	.08936	36,929	3,300	35,141	239,214	6.5
85	(1.00000)	33,629	33,629	204,073	204,073	6.1

Table 2

NOTE: \*This is technically an “interpolated life table” and not a complete life table based on single-year data.

SOURCE: This life table is available on the web at <http://www.cdc.gov/nchswww/datawh/statab/unpubd/mortabs/lewk2.htm>

describe the experience of an actual birth cohort—that is, a group of individuals who are born within a (narrowly) specified interval of time. If we wished to portray the mortality history of the birth cohort of 2000, for example, we would have to wait until the last individual of that cohort has died, or beyond the year 2110, before we would be able to calculate all of the values that comprise the life table. In such a life table, called a *generation* or *cohort life table*, we can explicitly obtain the probability of individuals surviving to a given age. As is intuitively clear, however, a generation life table is suitable primarily for historical analyses of cohorts now extinct. Any generation life table that we could calculate would be very much out of date and would in no way approximate the present mortality experience of a population. Thus, we realize the need for the period life table, which

treats a population at a given point in time as a *synthetic* or *hypothetical cohort*. The major drawback of the period life table is that it refers to no particular cohort of individuals. In an era of mortality rates declining at all ages, such a life table will underestimate true life expectancy for any cohort.

The most fundamental data that underlie the formation of a period life table are the number of deaths attributed to each age group in the population for a particular calendar year ( ${}_nD_x$ ), where  $x$  refers to the exact age at the beginning of the age interval and  $n$  is the width of that interval, and the number of individuals living at the midpoint of that year for each of those same age groups ( ${}_nP_x$ ).

To begin the life table’s construction, we take the ratio of these two sets of input data— ${}_nD_x$  and  ${}_nP_x$ —to form a series of age-specific death rates, or  ${}_nM_x$ :

**Abridged Life Table for the United States, 1996**

Exact Age $x$	${}_nD_x$	${}_nP_x$ (in 1,000s)	${}_nq_x$	$l_x$	${}_nd_x$	${}_nL_x$	$T_x$	$e_x$
0	28,487	3,769	.00732	100,000	732	99,370	7,611,825	76.1
1	5,948	16,516	.00151	99,268	150	396,721	7,512,455	75.7
5	3,780	19,441	.00097	99,118	96	495,329	7,115,734	71.8
10	4,550	18,981	.00118	99,022	117	494,883	6,620,405	66.9
15	14,663	18,662	.00390	98,905	386	493,650	6,125,522	61.9
20	17,780	17,560	.00506	98,519	499	491,372	5,631,872	57.2
25	20,730	19,007	.00544	98,020	533	488,766	5,140,500	52.4
30	30,417	21,361	.00710	97,487	692	485,746	4,651,734	47.7
35	42,499	22,577	.00944	96,795	914	481,820	4,165,988	43.0
40	53,534	20,816	.01283	95,881	1,230	476,549	3,684,168	38.4
45	67,032	18,436	.01801	94,651	1,705	469,305	3,207,619	33.9
50	77,297	13,934	.02733	92,946	2,540	458,779	2,738,314	29.5
55	96,726	11,362	.04177	90,406	3,776	443,132	2,279,535	25.2
60	136,999	9,999	.06649	86,630	5,760	419,530	1,836,403	21.2
65	200,045	9,892	.09663	80,870	7,814	385,659	1,416,873	17.5
70	273,849	8,778	.14556	73,056	10,634	339,620	1,031,214	14.1
75	321,223	6,873	.21060	62,422	13,146	280,047	691,594	11.1
80	342,067	4,557	.31754	49,276	15,647	207,474	411,547	8.4
85	576,541	3,762	1.00000	33,629	33,629	204,073	204,073	6.1

Table 3

SOURCE:  ${}_nD_x$  and  ${}_nP_x$  values are obtained from Peters, Kochanek, and Murphy 1998, and from the web site, <http://www.cdc.gov/nchswww/datawh/statab/unpubd/mortabs/pop6096.htm>, respectively.

$${}_nM_x = \frac{{}_nD_x}{{}_nP_x} \tag{3}$$

For each death rate, we compute the corresponding probability of dying within that age interval, given that one has survived to the beginning of the interval. This value, denoted by  ${}_nq_x$ , is computed using the following equation:

$${}_nq_x = \frac{n \cdot {}_nM_x}{1 + (n - {}_na_x) \cdot {}_nM_x} \tag{4}$$

where  ${}_na_x$  is the average number of years lived by those who die within the age interval  $x$  to  $x+n$ . (Except for the first year of life, it is typically assumed that deaths are uniformly distributed within an age interval, implying that  ${}_na_x = n/2$ .) Given the values of  $q$  and  $a$ , we are able to generate the entire life table.

The life table may be thought of as a tracking device, by which a cohort of individuals is followed from the moment of their birth until the last surviving individual dies. Under this interpretation, the various remaining columns are defined in the following manner:  $l_x$  equals the number of individuals in the life table surviving to exact age  $x$ . We arbitrarily set the number “born into” the life

table,  $l_0$ , which is otherwise known as the *radix*, to some value—most often, 100,000. We generate all subsequent  $l_x$  values by the following equation:

$$l_{x+n} = l_x \cdot [1 - {}_nq_x] \tag{5}$$

${}_nd_x$  equals the number of deaths experienced by the life table cohort within the age interval  $x$  to  $x+n$ . It is the product of the number of individuals alive at exact age  $x$  and the conditional probability of dying within the age interval:

$${}_nd_x = l_x \cdot {}_nq_x \tag{6}$$

The concept of “person-years” is critical to understanding life table construction. Each individual who survives from one birthday to the next contributes one additional person-year to those tallied by the cohort to which that person belongs. In the year in which the individual dies, the decedent contributes some fraction of a person-year to the overall number for that cohort.

${}_nL_x$  equals the total number of person-years experienced by a cohort in the age interval,  $x$  to  $x+n$ . It is the sum of person-years contributed by those who have survived to the end of the interval

and those contributed by individuals who die within that interval:

$${}_nL_x = [n'l_{x+n}] + [{}_na_x \cdot {}_nd_x] \quad (7)$$

$T_x$  equals the number of person-years lived beyond exact age  $x$ :

$$T_x = \sum_{a=x,n}^{\infty} {}_nL_a = T_{x+n} + {}_nL_x \quad (8)$$

$e_x$  equals the expected number of years of life remaining for an individual who has already survived to exact age  $x$ . It is the total number of person-years experienced by the cohort above that age divided by the number of individuals starting out at that age:

$$e_x = \frac{T_x}{l_x} \quad (9)$$

The  ${}_nL_x$  and  $T_x$  columns are generated from the oldest age to the youngest. If the last age category is, for example, eighty-five and above (it is typically “open-ended” in this way), we must have an initial value for  $T_{85}$  in order to begin the process. This value is derived in the following fashion: Since for this oldest age group,  $l_{85} = {}_{\infty}d_{85}$  (due to the fact that the number of individuals in a cohort who will die at age eighty-five or beyond is simply the number surviving to age eighty-five) and  $T_{85} = L_{85}$ , we have:

$$e_{85} = \frac{T_{85}}{l_{85}} = \frac{1}{l_{85} / T_{85}} = \frac{1}{{}_{\infty}d_{85} / L_{85}} \approx \frac{1}{{}_{\infty}M_{85}} \quad (10)$$

From the life table, we can obtain mortality information in a variety of ways. In table 2, we see, for example, that the expectation of life at birth,  $e_0$ , is 76.1 years. If an individual in this population survives to age eighty, then he or she might expect to live 8.4 years longer. We might also note that the probability of surviving from birth to one’s tenth birthday is  $l_{10}/l_0$ , or 0.99022. Given that one has already lived eighty years, the probability that one survives five additional years is  $l_{85}/l_{80}$ , or 33,629/49,276=0.68246.

**POPULATION PROJECTION**

The life table, in addition, is often used to project either total population size or the size of specific age groups. In so doing, we must invoke a different interpretation of the  ${}_nL_x$ ’s and the  $T_x$ ’s in the life table. We treat them as representing the age distribution of a *stationary population*—that is, a population having long been subject to zero growth. Thus,  ${}_5L_{20}$ , for example, represents the number of twenty- to twenty-four-year-olds in the life table “population,” into which  $l_0$ , or 100,000, individuals are born each year. (One will note by summing the  ${}_nd_x$  column that 100,000 die every year, thus giving rise to stationarity of the life table population.)

If we were to assume that the United States is a *closed population*—that is, a population whose net migration is zero—and, furthermore, that the mortality levels obtaining in 1996 were to remain constant for the following ten years, then we would be able to project the size of any U.S. cohort up to ten years into the future. Thus, if we wished to know the number of fifty- to fifty-four-year-olds in 2006, we would take advantage of the following relation that is assumed to hold approximately:

$$\frac{{}_nP_{x+t}^{\tau+t}}{{}_nP_x^{\tau}} \approx \frac{{}_nL_{x+t}^{\tau+t}}{{}_nL_x^{\tau}} \quad (11)$$

where  $\tau$  is the base year of the projection (e.g., 1996) and  $t$  is the number of years one is projecting the population forward. This equation implies that the fifty- to fifty-four-year-olds in 2006,  ${}_5P_{50}^{2006}$ , is simply the number of forty- to forty-four-year-olds ten years earlier,  ${}_5P_{40}^{1996}$ , multiplied by the proportion of forty- to forty-four-year-olds in the life table surviving ten years,  ${}_5L_{50}/{}_5L_{40}$ .

In practice, it is appropriate to use the above relation in population projection only if the width of the age interval under consideration,  $n$ , is sufficiently narrow. If the age interval is very broad—for example, in the extreme case in which we are attempting to project the number of people aged ten and above in 2006 from the number zero and above (i.e, the entire population) in 1996—we cannot be assured that the life table age distribution within that interval resembles closely enough the age distribution of the actual population. In other words, if the actual population’s age distribution within a broad age interval is significantly

different from that within the corresponding interval of the life table population, then implicitly by using this projection device we are improperly weighing the component parts of the broad interval with respect to survival probabilities.

Parenthetically, if we desired to determine the size of any component of the population under  $t$  years old—in this particular example, ten years old—we would have to draw upon fertility as well as mortality information, because at time  $\tau$  these individuals had not yet been born.

### HAZARDS MODELS

Suppose we were to examine the correlates of marital dissolution. In a life table analysis, the break-up of the marriage (as measured, e.g., by separation or divorce) would serve as the analogue to death, which is the means of exit in the standard life table analysis.

In the study of many duration-dependent phenomena, it is clear that several factors may affect whether an individual exits from a life table. Certainly, it is well-established that a large number of socioeconomic variables simultaneously impinge on the marital dissolution process. In many populations, whether one has given birth premaritally, cohabited premaritally, married at a young age, or had little in the way of formal education, among a whole host of other factors, have been found to be strongly associated with marital instability. In such studies, in which one attempts to disentangle the intricately related influences of several variables on survivorship in a given state, we invoke a hazards model approach. Such an approach may be thought of as a multivariate statistical extension of the simple life table analysis presented above (for theoretical underpinnings, see, e.g., Cox and Oakes 1984 and Allison 1984; for applications to marital stability, see, e.g., Menken, Trussell, Stempel, and Babakol 1981 and Bennett, Blanc, and Bloom 1988).

In the marital dissolution example, we would assume that there is a hazard, or risk, of dissolution at each marital duration,  $d$ , and we allow this duration-specific risk to depend on individual characteristics (such as age at marriage, education, etc.). In the *proportional hazards model*, a set of individual characteristics represented by a vector

of covariates shifts the hazard by the same proportional amount at all durations. Thus, for an individual  $i$  at duration  $d$ , with an observed set of characteristics represented by a vector of covariates,  $Z_i$ , the hazard function,  $\mu_i(d)$ , is given by:

$$\mu_i(d) = \exp[\lambda(d)] \exp[Z_i\beta] \quad (12)$$

where  $\beta$  is a vector of parameters and  $\lambda(d)$  is the underlying duration pattern of risk. In this model, then, the underlying risk of dissolution for an individual  $i$  with characteristics  $Z_i$  is multiplied by a factor equal to  $\exp[Z_i\beta]$ .

We may also implement a more general set of models to test for departures from some of the restrictive assumptions built into the proportional hazards framework. More specifically, we allow for time-varying covariates (for instance, in this example, the occurrence of a first marital birth) as well as allow for the effects of individual characteristics to vary with duration of first marriage. This model may be written as:

$$\mu_i(d) = \exp[\lambda(d)] \exp[Z_i(d)\beta(d)] \quad (13)$$

where  $\lambda(d)$  is defined as in the proportional hazards model,  $Z_i(d)$  is the vector of covariates, some of which may be time-varying, and  $\beta(d)$  represents a vector of parameters, some of which may give rise to nonproportional effects. The model parameters can be estimated using the method of maximum likelihood. The estimation procedure assumes that the hazard,  $\mu_i(d)$ , is constant within duration intervals. The interval width chosen by the analyst, of course, should be supported on both substantive and statistical grounds.

### INDIRECT DEMOGRAPHIC ESTIMATION

Unfortunately, many countries around the world have poor or nonexistent data pertaining to a wide array of demographic variables. In the industrialized nations, we typically have access to data from rigorous registration systems that collect data on mortality, marriage, fertility, and other demographic processes. However, when analyzing the demographic situation of less developed nations, we are often confronted with a paucity of available data on these fundamental processes. When such data are in fact collected, they are often sufficiently

inadequate to be significantly misleading. For example, in some countries we have learned that as few as half of all actual deaths are recorded. If we mistakenly assume the value of the actual number to be the registered number, then we will substantially overestimate life expectancy in these populations. In essence, we will incorrectly infer that people are dying at a slower rate than is truly the case.

**The Stable Population Model.** Much demographic estimation has relied on the notion of stability. A *stable population* is defined as one that is established by a long history of unchanging fertility and mortality patterns. This criterion gives rise to a fixed proportionate age distribution, constant birth and death rates, and a constant rate of population growth (see, e.g., Coale 1972). The basic stable population equation is:

$$c(a) = be^{-ra}p(a) \quad (14)$$

where  $c(a)$  is the proportion of the population exact age  $a$ ,  $b$  is the crude birth rate,  $r$  is the rate of population growth, and  $p(a)$  is the proportion of the population surviving to exact age  $a$ . Various mathematical relationships have been shown to obtain among the demographic variables in a stable population. This becomes clear when we multiply both sides of the equation by the total population size. Thus, we have:

$$N(a) = Be^{-ra}p(a) \quad (15)$$

where  $N(a)$  is the number of individuals in the population exact age  $a$  and  $B$  is the current annual number of births. We can see that the number of people aged  $a$  this year is simply the product of the number of births entering the population  $a$  years ago—namely, the current number of births times a growth rate factor, which discounts the births according to the constant population growth rate,  $r$  (which also applies to the growth of the number of births over time)—and the proportion of a birth cohort that survives to be aged  $a$  today. Note that the constancy over time of the mortality schedule,  $p(a)$ , and the growth rate,  $r$ , are crucial to the validity of this interpretation.

When we assume a population is stable, we are imposing structure upon the demographic relationships existing therein. In a country where data

are inadequate, indirect methods allow us—by drawing upon the known structure implied by stability—to piece together sometimes inaccurate information and ultimately derive sensible estimates of the population parameters. The essential strategy in indirect demographic estimation is to infer a value or set of values for a variable whose elements are either unobserved or inaccurate from the relationship among the remaining variables in the above equation (or an equation deriving from the one above). We find that these techniques are robust with respect to moderate departures from stability, as in the case of quasi-stable populations, in which only fertility has been constant and mortality has been gradually changing.

**The Nonstable Population Model.** Throughout much of the time span during which indirect estimation has evolved, there have been many countries where populations approximated stability. In recent decades, however, many countries have experienced rapidly declining mortality or declining or fluctuating fertility and, thus, have undergone a radical departure from stability. Consequently, previously successful indirect methods, grounded in stable population theory, are, with greater frequency, ill-suited to the task for which they were devised. As is often the case, necessity is the mother of invention and so demographers have sought to adapt their methodology to the changing world.

In the early 1980s, a methodology was developed that can be applied to populations that are far from stable (see, e.g., Bennett and Horiuchi 1981; and Preston and Coale 1982). Indeed, it is no longer necessary to invoke the assumption of stability, if we rely upon the following equation:

$$c(a) = b \cdot \exp \left[ -\int_0^a r(x) dx \right] \cdot p(a) \quad (16)$$

where  $r(x)$  is the growth rate of the population at exact age  $x$ . This equation holds true for any closed population, and, indeed, can be modified to accommodate populations open to migration.

The implied relationships among the age distribution of living persons and deaths, and rates of growth of different age groups, provide the basis for a wide range of indirect demographic methods that allow us to infer accurate estimates of basic



demographic parameters that ultimately can be used to better inform policy on a variety of issues. Two examples are as follows.

First, suppose we have the age distribution for a country at each of two points in time, in addition to the age distribution of deaths occurring during the intervening years. We may then estimate the completeness of death registration in that population using the following equation (Bennett and Horiuchi 1981):

$$\tilde{N}(a) = \int_a^{\infty} D(x) \exp \left[ \int_a^x r(u) du \right] dx \quad (17)$$

where  $\hat{N}(a)$  is the estimated number of people at exact age  $a$ ,  $D(x)$  is the number of deaths at exact age  $x$  and  $r(u)$  is the rate of the growth of the number of persons at exact age  $u$  between the two time points. By taking the ratio of the estimated number of persons with the enumerated population, we have an estimate of the completeness of death registration in the population relative to the completeness of the enumerated population. This relative completeness (in contrast to an “absolute” estimate of completeness) is all that is needed to determine the true, unobserved age-specific death rates, which in turn allows one to construct an unbiased life table.

A second example of the utility of the nonstable population framework is shown by the use of the following equation:

$$\frac{N(x)}{N(a)} = {}_{x-a}P_a \exp \left[ \int_a^x r(u) du \right] \quad (18)$$

where  $N(x)$  and  $N(a)$  are the number of people exact ages  $x$  and  $a$ , respectively, and  ${}_{x-a}P_a$  is the probability of surviving from age  $a$  to age  $x$  according to period mortality rates. By using variants of this equation, we can generate reliable population age distributions (e.g., in situations in which censuses are of poor quality) from a trustworthy life table (Bennett and Garson 1983).

### MORTALITY MODELING

The field of demography has a long tradition of developing models that are based upon empirical regularities. Typically in demographic modeling, as in all kinds of modeling, we try to adhere to the

principle of parsimony—that is, we want to be as efficient as possible with regard to the detail, and therefore the number of parameters, in a model.

Mortality schedules from around the world reveal that death rates follow a common pattern of relatively high rates of infant mortality, rates that decline through early childhood until they bottom out in the age range of five to fifteen or so, then rates that increase slowly through the young and middle adult years, and finally rising more rapidly during the older adult ages beyond the forties or fifties. Various mortality models exploit this regular pattern in the data. Countries differ with respect to the overall level of mortality, as reflected in the expectation of life at birth, and the precise relationship that exists among the different age components of the mortality curve.

Coale and Demeny (1983) examined 192 mortality schedules from different times and regions of the world and found that they could be categorized into four “families” of life tables. Although overall mortality levels might differ, within each family the relationships among the various age components of mortality were shown to be similar. For each family, Coale and Demeny constructed a “model life table” for females that was associated with each of twenty-five expectations of life at birth from twenty through eighty. A comparable set of tables was developed for males. In essence, a researcher can match bits of information that are known to be accurate in a population with the corresponding values in the model life tables, and ultimately derive a detailed life table for the population under study. In less developed countries, model life tables are often used to estimate basic mortality parameters, such as  $e_0$  or the crude death rate, from other mortality indicators that may be more easily observable.

Other mortality models have been developed, the most notable being that by Brass (1971). Brass noted that one mortality schedule could be related to another by means of a linear transformation of the logits of their respective survivorship probabilities (i.e., the vector of  $l_x$  values, given a radix of one). Thus, one may generate a life table by applying the logit system to a “standard” or “reference” life table, given an appropriate pair of parameters that reflect (1) the overall level of mortality in the population under study, and (2) the relationship between child and adult mortality.

**MARRIAGE, FERTILITY, AND  
MIGRATION MODELS**

Coale (1971) observed that age distributions of first marriages are structurally similar in different populations. These distributions tend to be smooth, unimodal, and skewed to the right, and to have a density close to zero below age fifteen and above age fifty. He also noted that the differences in age-at-marriage distributions across female populations are largely accounted for by differences in their means, standard deviations, and cumulative values at the older ages, for example, at age fifty. As a basis for the application of these observations, Coale constructed a standard schedule of age at first marriage using data from Sweden, covering the period 1865 through 1869. The model that is applied to marriage data is represented by the following equation:

$$g(a) = \frac{E}{\sigma} 1.2813 \exp\left\{-1.145 \left(\frac{a-\mu}{\sigma} + 0.805\right) - \exp[-1.896 \left(\frac{a-\mu}{\sigma} + 0.805\right)]\right\} \quad (20)$$

where  $g(a)$  is the proportion marrying at age  $a$  in the observed population and  $\mu$ ,  $\sigma$ , and  $E$  are, respectively, the mean and the standard deviation of age at first marriage (for those who ever marry), and the proportion ever marrying.

The model can be extended to allow for covariate effects by stipulating a functional relationship between the parameters of the model distribution and a set of covariates. This may be specified as follows:

$$\left. \begin{aligned} \mu_i &= \mathbf{X}_i' \alpha \quad , \\ \sigma_i &= \mathbf{Y}_i' \beta \quad , \\ E_i &= \mathbf{Z}_i' \gamma \quad , \end{aligned} \right\} \quad (20)$$

where  $\mathbf{X}_i$ ,  $\mathbf{Y}_i$ , and  $\mathbf{Z}_i$  are the vector values of characteristics of an individual that determine, respectively,  $\mu_i$ ,  $\sigma_i$ , and  $E_i$ , and  $\alpha$ ,  $\beta$ , and  $\gamma$  are the associated parameter vectors to be estimated.

Because the model is parametric, it can be applied to data referring to cohorts who have yet to complete their marriage experience. In this fashion, the model can be used for purposes of projection (see, e.g., Bloom and Bennett 1990). The model has also been found to replicate well

the first birth experience of cohorts (see, e.g., Bloom 1982).

Coale and Trussell (1974), recognizing the empirical regularities that exist among age profiles of fertility across time and space and extending the work of Louis Henry, developed a set of model fertility schedules. Their model is based in part on a reference distribution of age-specific marital fertility rates that describes the pattern of fertility in a *natural fertility population*—that is, one that exhibits no sign of controlling the extent of child-bearing activity. When fitted to an observed age pattern of fertility, the model's two parameters describe the overall level of fertility in the population and the degree to which their fertility within marriage is controlled by some means of contraception. Perhaps the greatest use of this model has been devoted to comparative analyses, which is facilitated by the two-parameter summary of any age pattern of fertility in question.

Although the application of indirect demographic estimation methods to migration analysis is not as mature as that to other demographic processes, strategies similar to those invoked by fertility and mortality researchers have been applied to the development of model migration schedules. Rogers and Castro (1981) found that similar age patterns of migration obtained among many different populations. They have summarized these regularities in a basic eleven-parameter model, and, using Brass and Coale logic, explore ways in which their model can be applied satisfactorily to data of imperfect quality.

The methods described above comprise only a small component of the methodological tools available to demographers and to social scientists, in general. Some of these methods are more readily applicable than others to fields outside of demography. It is clear, for example, how we may take advantage of the concept of standardization in a variety of disciplines. So, too, may we apply life table analysis and nonstable population analysis to problems outside the demographic domain. Any analogue to birth and death processes can be investigated productively using these central methods. Even the fundamental concept underlying the above mortality, fertility, marriage, and migration models—that is, exploiting the power to be found in empirical regularities—can be applied fruitfully to other research endeavors.



REFERENCES

- Allison, Paul D. 1984 *Event History Analysis*. Beverly Hills, Calif.: Sage Publications.
- Bennett, Neil G., Ann K. Blanc, and David E. Bloom 1988 "Commitment and the Modern Union: Assessing the Link between Premarital Cohabitation and Subsequent Marital Stability." *American Sociological Review* 53:127-138.
- Bennett, Neil G., and Lea Keil Garson 1983 "The Centenarian Question and Old-Age Mortality in the Soviet Union, 1959-1970." *Demography* 20:587-606.
- Bennett, Neil G., and Shiro Horiuchi 1981 "Estimating the Completeness of Death Registration in a Closed Population." *Population Index* 47:207-221.
- Bloom, David E. 1982 "What's Happening to the Age at First Birth in the United States? A Study of Recent Cohorts." *Demography* 19:351-370.
- , and Neil G. Bennett 1990 "Modeling American Marriage Patterns." *Journal of the American Statistical Association* 85 (December):1009-1017.
- Coale, Ansley J. 1972 *The Growth and Structure of Human Populations*. Princeton, N.J.: Princeton University Press.
- 1971 "Age Patterns of Marriage." *Population Studies* 25:193-214.
- , and Paul Demeny 1983 *Regional Model Life Tables and Stable Populations*, 2nd ed. New York: Academic Press.
- Coale, Ansley J., and James Trussell 1974 "Model Fertility Schedules: Variations in the Age Structure of Childbearing in Human Populations." *Population Index* 40:185-206.
- Cox, D. R., and D. Oakes 1984 *Analysis of Survival Data*. London: Chapman and Hall.
- Menken, Jane, James Trussell, Debra Stempel, and Ozer Babakol 1981 "Proportional Hazards Life Table Models: An Illustrative Analysis of Sociodemographic Influences on Marriage Dissolution in the United States." *Demography* 18:181-200.
- Peters, Kimberley D., Kenneth D. Kochanek, and Sherry L. Murphy 1998 "Deaths: Final Data for 1996." *National Vital Statistics Reports*, vol. 47, no. 9. Hyattsville, Md.: National Center for Health Statistics.
- Preston, Samuel H., and Ansley J. Coale 1982 "Age Structure, Growth, Attrition, and Accession: A New Synthesis." *Population Index* 48:217-259.
- Rogers, Andrei, and Luis J. Castro 1981 "Model Migration Schedules." (Research Report 81-30) Laxenburg,

Austria: International Institute for Applied Systems Analysis.

NEIL G. BENNETT

**DEMOGRAPHIC TRANSITION**

The human population has maintained relatively gradual growth throughout most of history by high, and nearly equal, rates of deaths and births. Since about 1800, however, this situation has changed dramatically, as most societies have undergone major declines in mortality, setting off high growth rates due to the imbalance between deaths and births. Some societies have eventually had fertility declines and emerged with a very gradual rate of growth as low levels of births matched low levels of mortality.

There are many versions of demographic transition theory (Mason 1997), but there is some consensus that each society has the potential to proceed sequentially through four general stages of variation in death and birth rates and population growth. Most societies in the world have passed through the first two stages, at different dates and speeds, and the contemporary world is primarily characterized by societies in the last two stages, although a few are still in the second stage.

Stage 1, presumably characterizing most of human history, involves high and relatively equal birth and death rates and little resulting population growth.

Stage 2 is characterized by a declining death rate, especially concentrated in the years of infancy and childhood. The fertility rate remains high, leading to at least moderate population growth.

Stage 3 involves further declines in mortality, usually to low levels, and initial sustained declines in fertility. Population growth may become quite high, as levels of fertility and mortality increasingly diverge.

Stage 4 is characterized by the achievement of low mortality and the rapid emergence of low fertility levels, usually near those of mortality. Population growth again becomes quite low or negligible.

While demographers argue about the details of demographic change in the past 200 years, clearcut declines in birth and death rates appeared on the European continent and in areas of overseas European settlement in the nineteenth century, especially in the last three decades. By 1900, life expectancies in “developed” societies such as the United States were probably in the mid-forties, having increased by a few years in the century (Preston and Haines 1991). By the end of the twentieth century, even more dramatic gains in mortality were evident, with life expectancies reaching into the mid- and high-seventies.

The European fertility transition of the late 1800s to the twentieth century involved a relatively continuous movement from average fertility levels of five or six children per couple to bare levels of replacement by the end of the 1930s. Fertility levels rose again after World War II, but then began another decline about 1960. Some countries now have levels of fertility that are well below long-term replacement levels.

With a few exceptions such as Japan, most other parts of the developing world did not experience striking declines in mortality and fertility until the midpoint of the twentieth century. Gains in life expectancy became quite common and very rapid in the post-World War II period throughout the developing world (often taking less than twenty years), although the amount of change was quite variable. Suddenly in the 1960s, fertility transitions emerged in a small number of societies, especially in the Caribbean and Southeast Asia, to be followed in the last part of the twentieth century by many other countries.

Clear variations in mortality characterize many parts of the world at the end of the twentieth century. Nevertheless, life expectancies in countries throughout the world are generally greater than those found in the most developed societies in 1900. A much greater range in fertility than mortality characterizes much of the world, but fertility declines seem to be spreading, including in “laggard” regions such as sub-Saharan Africa.

The speed with which the mortality transition was achieved among contemporary lesser-developed countries has had a profound effect on the magnitude of the population growth that has occurred during the past few decades. Sweden, a

model example of the nineteenth century European demographic transition, peaked at an annual rate of natural increase of 1.2 percent. In contrast, many developing countries have attained growth rates of over 3.0 percent. The world population grew at a rate of about 2 percent in the early 1970s but has now declined to about 1.4 percent, as fertility rates have become equal to the generally low mortality rates.

### CAUSES OF MORTALITY TRANSITIONS

The European mortality transition was gradual, associated with modernization and raised standards of living. While some dispute exists among demographers, historians, and others concerning the relative contribution of various causes (McKeown 1976; Razzell 1974), the key factors probably included increased agricultural productivity and improvements in transportation infrastructure which enabled more efficient food distribution and, therefore, greater nutrition to ward off disease. The European mortality transition was also probably influenced by improvements in medical knowledge, especially in the twentieth century, and by improvements in sanitation and personal hygiene. Infectious and environmental diseases especially have declined in importance relative to cancers and cardiovascular problems. Children and infants, most susceptible to infectious and environmental diseases, have showed the greatest gains in life expectancy.

The more recent and rapid mortality transitions in the rest of the world have mirrored the European change with a movement from infectious/environmental causes to cancers and cardiovascular problems. In addition, the greatest beneficiaries have been children and infants. These transitions result from many of the same factors as the European case, generally associated with economic development, but as Preston (1975) outlines, they have also been influenced by recent advances in medical technology and public health measures that have been imported from the highly-developed societies. For instance, relatively inexpensive vaccines are now available throughout the world for immunization against many infectious diseases. In addition, airborne sprays have been distributed at low cost to combat widespread

diseases such as malaria. Even relatively weak national governments have instituted major improvements in health conditions, although often only with the help of international agencies.

Nevertheless, mortality levels are still higher than those in many rich societies due to such factors as inadequate diets and living conditions, and inadequate development of health facilities such as hospitals and clinics. Preston (1976) observes that among non-Western lesser-developed countries, mortality from diarrheal diseases (e.g., cholera) has persisted despite control over other forms of infectious disease due to the close relationship between diarrheal diseases, poverty, and ignorance—and therefore a nation's level of socioeconomic development.

Scholars (Caldwell 1986; Palloni 1981) have warned that prospects for future success against high mortality may be tightly tied to aspects of social organization that are independent of simple measures of economic well-being: Governments may be more or less responsive to popular need for improved health; school system development may be important for educating citizens on how to care for themselves and their families; the equitable treatment of women may enhance life expectancy for the total population.

Recent worldwide mortality trends may be charted with the help of data on life expectancy at age zero that have been gathered, sometimes on the basis of estimates, by the Population Reference Bureau (PRB), a highly respected chronicler of world vital rates. For 165 countries with relatively stable borders over time, it is possible to relate estimated life expectancy in 1986 with the same figure for 1998. Of these countries, only 13.3 percent showed a decline in life expectancy during the time period. Some 80.0 percent had overall increasing life expectancy, but the gains were highly variable. Of all the countries, 29.7 percent actually had gains of at least five years or more, a sizable change given historical patterns of mortality.

An indication of the nature of change may be discerned by looking at Figure 1, which shows a graph of the life expectancy values for the 165 countries with stable borders. Each point represents a country and shows the level of life expectancy in 1986 and in 1998. Note the relatively high

levels of life expectancy by historical standards for most countries in both years. Not surprisingly, there is a strong tendency for life expectancy values to be correlated over time. A regression straight line, indicating average life expectancy in 1998 as a function of life expectancy in 1986 describes this relationship. As suggested above, the levels of life expectancy in 1998 tend to be slightly higher than the life expectancy in 1986. Since geography is highly associated with economic development, the points on the graph generally form a continuum from low to high life expectancy. African countries tend to have the lowest life expectancies, followed by Asia, Oceania, and the Americas. Europe has the highest life expectancies.

The African countries comprise virtually all the countries with declining life expectancies, probably a consequence of their struggles with acquired immune deficiency syndrome (AIDS), malnutrition, and civil disorder. Many of them have lost several years of life expectancy in a very short period of time. However, a number of the African countries also have sizeable increases in life expectancies.

Asian and American countries dominate the mid-levels of life expectancy, with the Asian countries showing a strong tendency to increase their life expectancies, consistent with high rates of economic development.

Unfortunately, Figure 1 does not include the republics of the former Soviet Union, since exactly comparable data are not available for both time points. Nevertheless, there is some consensus among experts that life expectancy has deteriorated in countries such as Russia that have made the transition from communism to economically-unstable capitalism.

### **WHAT DRIVES FERTILITY RATES?**

The analysis of fertility decline is somewhat more complicated analytically than mortality decline. One may presume that societies will try, if given resources and a choice, to minimize mortality levels, but it seems less necessarily so that societies have an inherent orientation toward low fertility, or, for that matter, any specific fertility level. In addition, fertility rates may vary quite widely across

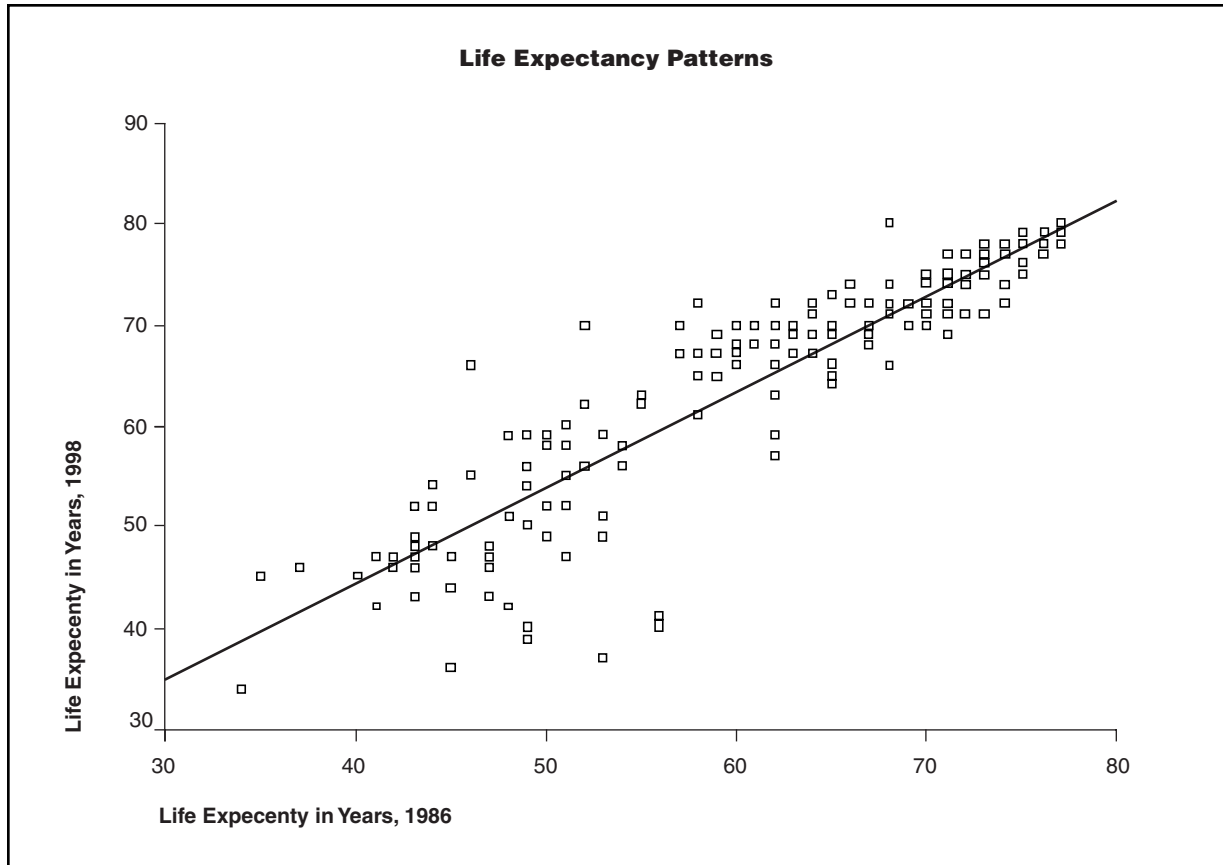


Figure 1

societies due to factors (Bongaarts 1975) that have little relationship to conscious desires such as prolonged breastfeeding which suppresses reproductive ovulation in women, the effectiveness of birth control methods, and the amount of involuntary foetal abortion. As a result of these analytic ambiguities, scholars seem to have less consensus on the social factors that might produce fertility than mortality decline (Hirschman 1994; Mason 1997).

Coale (1973), in an attempt to reconcile the diversity of circumstances under which fertility declines have been observed to occur, identified three major conditions for a major fall in fertility:

1. Fertility must be within the calculus of conscious choice. Parents must consider it an acceptable mode of thought and form of behavior to balance the advantages and disadvantages of having another child.

2. Reduced fertility must be viewed as socially or economically advantageous to couples.
3. Effective techniques of birth control must be available. Sexual partners must know these techniques and have a sustained will to use the them.

Beyond Coale's conditions, little consensus has emerged on the causes of fertility decline. There are, however, a number of major ideas about what causes fertility transitions that may be summarized in a few major hypotheses.

A major factor in causing fertility change may be the mortality transition itself. High-mortality societies depend on high fertility to ensure their survival. In such circumstances, individual couples will maximize their fertility to guarantee that at least a few of their children survive to adulthood, to perpetuate the family lineage and to care for

them in old age. The decline in mortality may also have other consequences for fertility rates. As mortality declines, couples may perceive that they can control the survival of family members by changing health and living practices such as cleanliness and diet. This sense of control may extend itself to the realm of fertility decisions, so that couples decide to calculate consciously the number of children they would prefer and then take steps to achieve that goal.

Another major factor may be the costs and benefits of children. High-mortality societies are often characterized by low technology in producing goods; in such a situation (as exemplified by many agricultural and mining societies), children may be economically useful to perform low-skilled work tasks. Parents have an incentive to bear children, or, at the minimum, they have little incentive not to bear children. However, high-technology societies place a greater premium on highly-skilled labor and often require extended periods of education. Children will have few economic benefits and may become quite costly as they are educated and fed for long periods of time.

Another major factor that may foster fertility decline is the transfer of functions from the family unit to the state. In low-technology societies, the family or kin group is often the fundamental unit, providing support for its members in times of economic distress and unemployment and for older members who can no longer contribute to the group through work activities. Children may be viewed as potential contributors to the unit, either in their youth or adulthood. In high-technology societies, some of the family functions are transferred to the state through unemployment insurance, welfare programs, and old age retirement systems. The family functions much more as a social or emotional unit where the economic benefits of membership are less tangible, thus decreasing the incentive to bear children.

Other major factors (Hirschman 1994; Mason 1997) in fertility declines may include urbanization and gender roles. Housing space is usually costly in cities, and the large family becomes untenable. In many high-technology societies, women develop alternatives to childbearing through

employment outside their homes, and increasingly assert their social and political rights to participate equally with men in the larger society. Because of socialization, men are generally unwilling to assume substantial child-raising responsibilities, leaving partners with little incentive to participate in sustained childbearing through their young adult lives.

No consensus exists on how to order these theories in relative importance. Indeed, each theory may have more explanatory power in some circumstances than others, and their relative importance may vary over time. For instance, declines in mortality may be crucial in starting fertility transitions, but significant alterations in the roles of children may be key for completing them. Even though it is difficult to pick the “best” theory, every country that has had a sustained mortality decline of at least thirty years has also had some evidence of a fertility decline. Many countries seem to have the fertility decline precondition of high life expectancy, but fewer have achieved the possible preconditions of high proportions of the population achieving a secondary education.

#### **EUROPEAN FERTILITY TRANSITION**

Much of what is known about the process of fertility transition is based upon research at Princeton University (known as the European Fertility Project) on the European fertility transition that took place primarily during the seventy-year period between 1870 and 1940. Researchers used aggregate government-collected data for the numerous “provinces” or districts of countries, typically comparing birth rates across time and provinces.

In that almost all births in nineteenth-century Europe occurred within marriage, the European model of fertility transition was defined to take place at the point marital fertility was observed to fall by more than 10 percent (Coale and Treadway 1986). Just as important, the Project scholars identified the existence of varying levels of natural fertility (birth rates when no deliberate fertility control is practiced) across Europe and throughout European history (Knodel 1977). Comparative use of natural fertility models and measures derived from these models have been of enormous use to demographers in identifying the initiation and progress of fertility transitions in more contemporary contexts.

Most scholars have concluded that European countries seemed to start fertility transitions from very different levels of natural fertility but moved at quite similar speeds to similar levels of controlled fertility on the eve of World War II (Coale and Treadway 1986). As the transition progressed, areal differences in fertility within and across countries declined, while the remaining differences were heavily between countries (Watkins 1991).

Although some consensus has emerged on descriptive aspects of the fertility transition, much less agreement exists on the social and economic factors that caused the long-term declines. Early theorists of fertility transitions (Notestein 1953) had posited a simple model driven by urban-industrial social structure, but this perspective clearly proved inadequate. For instance, the earliest declines did not occur in England, the most urban-industrial country of the time, but were in France, which maintained a strong rural culture. The similarity of the decline across provinces and countries of quite different social structures also seemed puzzling within the context of previous theorizing. Certainly, no one has demonstrated that variations in the fertility decline across countries, either in the timing or the speed, were related clearly to variations in crude levels of infant mortality, literacy rates, urbanization, and industrialization. Other findings from recent analysis of the European experience include the observation that in some instances, reductions in fertility preceded reductions in mortality (Cleland and Wilson 1987), a finding that is inconsistent with the four-stage theory of demographic transition.

The findings of the European Fertility Project have led some demographers (Knodel and van de Walle 1979) to reformulate ideas about why fertility declined. They suggest that European couples were interested in a small family well before the actual transition occurred. The transition itself was especially facilitated by the development of effective and cheap birth control devices such as the condom and diaphragm. Information about birth control rapidly and widely diffused through European society, producing transitions that seemed to occur independently of social structural factors such as mortality, urbanization, and educational attainment. In addition, these scholars argue that “cultural” factors were also important in the decline. This is based on the finding that provinces of some countries such as Belgium differed in their

fertility declines on the basis of areal religious composition (Lesthaeghe 1977) and that, in other countries such as Italy, areal variations in the nature of fertility decline were related to political factors such as the Socialist vote, probably reflecting anticlericalism (Livi-Bacci 1977). Others (Lesthaeghe 1983) have also argued for “cultural” causes of fertility transitions.

While the social causes of the European fertility transition may be more complex than originally thought, it may still be possible to rescue some of the traditional ideas. For instance, mortality data in Europe at the time of the fertility transition were often quite incomplete or unreliable, and most of the studies focused on infant (first year of life) mortality as possible causes of fertility decline. Matthiessen and McCann (1978) show that mortality data problems make some of the conclusions suspect and that infant mortality may sometimes be a weak indicator of child survivorship to adulthood. They argue that European countries with the earliest fertility declines may have been characterized by more impressive declines in post-infant (but childhood) mortality than infant mortality.

Conclusions about the effects of children’s roles on fertility decline have often been based on rates of simple literacy as an indicator of educational system development. However, basic literacy was achieved in many European societies well before the major fertility transitions, and the major costs of children would occur when secondary education was implemented on a large scale basis, which did not happen until near the end of the nineteenth century (Van de Walle 1980). In a time-series analysis of the United States fertility decline from 1870 to the early 1900s, Guest and Tolnay (1983) find a nearly perfect tendency for the fertility rate to fall as the educational system expanded in terms of student enrollments and length of the school year. Related research also shows that educational system development often occurred somewhat independently of urbanization and industrialization in parts of the United States (Guest 1981).

An important methodological issue in the study of the European transition (as in other transitions) is how one models the relationship between social structure and fertility. Many of the research reports from the European Fertility Project seem to assume that social structure and fertility had to be

closely related at all time points to support various theories about the causal importance of such factors as mortality and children's roles, but certain lags and superficial inconsistencies do not seem to prove fundamentally that fertility failed to respond as some of the above theories would suggest. The more basic question may be whether fertility eventually responded to changes in social structure such as mortality.

Even after admitting some problems with previous traditional interpretations of the European fertility transition, one cannot ignore the fact that the great decline in fertility occurred at almost the same time as the great decline in mortality and was associated (even if loosely) with a massive process of urbanization, industrialization, and the expansion of educational systems.

#### FERTILITY TRANSITIONS IN THE DEVELOPING WORLD

The great majority of countries in the developing world have undergone some fertility declines in the second half of the twentieth century. While the spectacularly rapid declines (Taiwan, South Korea) receive the most attention, a number are also very gradual (e.g. Guatemala, Haiti, Iraq, Cambodia), and a number are so incipient (especially in Africa) that their nature is difficult to discern.

The late twentieth century round of fertility transitions has occurred in a very different social context than the historical European pattern. In the past few decades, mortality has declined very rapidly. National governments have become very attuned to checking their unprecedented national growth rates through fertility control. Birth control technology has changed greatly through the development of inexpensive methods such as the intrauterine device (IUD). The world has become more economically and socially integrated through the expansion of transportation and developments in electronic communications, and "Western" products and cultural ideas have rapidly diffused throughout the world. Clearly, societies are not autonomous units that respond demographically as isolated social structures.

Leaders among developing countries in the process of demographic transition were found in East Asia and Latin America, and the Caribbean

(Coale 1983). The clear leaders among Asian nations, such as South Korea and Taiwan, generally had experienced substantial economic growth, rapid mortality decline, rising educational levels, and exposure to Western cultural influences (Freedman 1998). By 1998, South Korea and Taiwan had fertility rates that were below long-term replacement levels. China also experienced rapidly declining fertility, which cannot be said to have causes in either Westernization or more than moderate economic development, with a life expectancy estimated at seventy-one years and a rate of natural increase of 1.0 percent (PRB 1998).

Major Latin American nations that achieved substantial drops in fertility (exceeding 20 percent) in recent decades with life expectancies surpassing sixty years include Argentina, Brazil, Chile, Columbia, the Dominican Republic, Jamaica, Mexico, and Venezuela. All of these have also experienced substantial changes in mortality, education, or both, and economic development.

Unlike the European historical experience, fertility declines in the post-1960 period have not always sustained themselves until they reached near replacement levels. A number of countries have started declines but then leveled off with three or four children per reproductive age woman. For instance, Malaysia was considered a "miracle" case of fertility decline, along with South Korea and Taiwan, but in recent years its fertility level has stabilized somewhat above the replacement level.

Using the PRB data for 1986 and 1998, we can trace recent changes for 166 countries in estimated fertility as measured by the Total Fertility Rate (TFR), an indicator of the number of children typically born to a woman during her lifetime. Some 80.1 percent of the countries showed declines in fertility. Of all the countries, 37.3 percent had a decline of at least one child per woman, and 9.0 percent had a decline of at least two children per woman.

The region that encompasses countries having the highest rates of population growth is sub-Saharan Africa. Growth rates generally exceed 2 percent, with several countries having rates that clearly exceed 3 percent. This part of the world has been one of the latest to initiate fertility declines, but in the 1986-1998 period, Botswana, Kenya,

and Zimbabwe all sustained fertility declines of at least two children per woman, and some neighboring societies were also engaged in fertility transition. At the same time, many sub-Saharan countries are pre-transitional or only in the very early stages of a transition. Of the twenty-five countries that showed fertility increases in the PRB data, thirteen of them were sub-Saharan nations with TFRs of at least 5.0.

In general, countries of the Middle East and regions of Northern Africa populated by Moslems have also been slow to embark on the process of fertility transition. Some (Caldwell 1976) found this surprising since a number had experienced substantial economic advances and invited the benefits of Western medical technology in terms of mortality reduction. Their resistance to fertility transitions had been attributed partly to an alleged Moslem emphasis on the subordinate role of women to men, leading them to have limited alternatives to a homemaker role. However, the PRB data for 1986–1998 indicate that some of these countries (Algeria, Bangladesh, Jordan, Kuwait, Morocco, Syria, Turkey) are among the small number that achieved reduction of at least two children per woman.

The importance of the mortality transition in influencing the fertility transition is suggested by Figure 2. Each dot is a country, positioned in terms of graphical relationship in the PRB data between life expectancy in 1986 and the TFR in 1998. The relationship is quite striking. No country with a life expectancy less than fifty has a TFR below 3.0. Remember that before the twentieth century, virtually all countries had life expectancies below fifty years. In addition, the figure shows a very strong tendency for countries with life expectancies above seventy to have TFRs below 2.0.

For a number of years, experts on population policy were divided on the potential role of contraceptive programs in facilitating fertility declines (Davis 1967). Since contraceptive technology has become increasingly cheap and effective, some (Enke 1967) argue that modest international expenditures on these programs in high-fertility countries could have significant rapid impacts on reproduction rates. Others (Davis 1967) point out, however, that family planning programs would only permit couples to achieve their desires, which may not be compatible with societal replacement

level fertility. The primary implication was that family planning programs would not be effective without social structures that encouraged the small family. A recent consensus on the value of family planning programs relative to social structural change seems to have emerged. Namely, family planning programs may be quite useful for achieving low fertility where the social structure is consistent with a small family ideal (Mauldin and Berelson 1978).

While the outlook for further fertility declines in the world is good, it is difficult to say whether and when replacement-level fertility will be achieved. Many, many major social changes have occurred in societies throughout the world in the past half-century. These changes have generally been unprecedented in world history, and thus we have little historical experience from which to judge their impact on fertility, both levels and speed of change (Mason 1997).

Some caution should be exercised about future fertility declines in some of the societies that have been viewed as leaders in the developing world. For instance, in a number of Asian societies, a strong preference toward sons still exists, and couples are concerned as much about having an adequate number of sons survive to adulthood as they are about total sons and daughters. Since pre-birth gender control is still difficult, many couples have a number of girl babies before they are successful in bearing a son. If effective gender control is achieved, some of these societies will almost certainly attain replacement-level fertility.

In other parts of the world such as sub-Saharan Africa, the future of still-fragile fertility transitions may well depend on unknown changes in the organization of families. Caldwell (1976), in a widely respected theory of demographic transition that incorporates elements of both cultural innovation and recognition of the role of children in traditional societies in maintaining net flows of wealth to parents, has speculated that the traditional extended kinship family model now predominant in the region facilitates high fertility. Families often form economic units where children are important work resources. The extended structure of the household makes the cost of any additional member low relative to a nuclear family structure. Further declines in fertility will depend on the



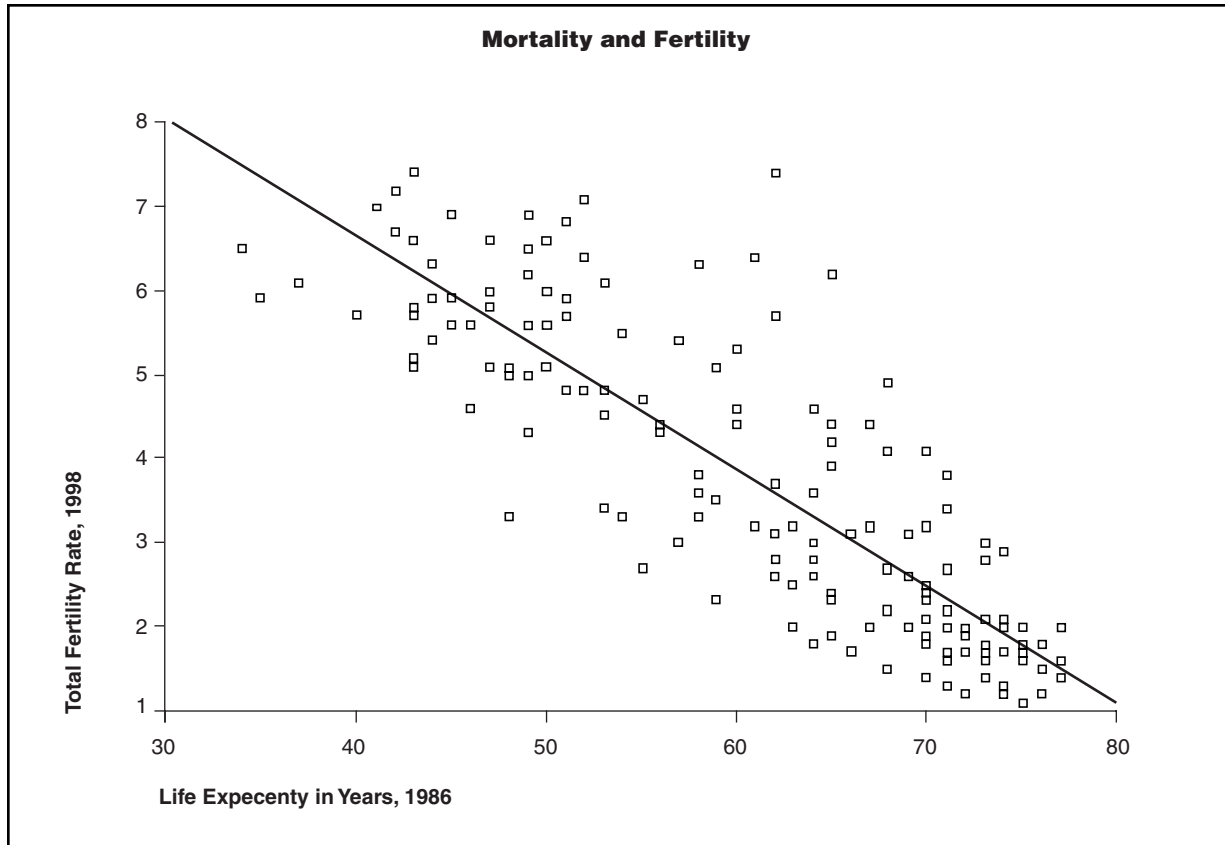


Figure 2

degree to which populations adopt the “Western” nuclear family, either through cultural diffusion or through autonomous changes in local social structure.

Taking the long view, the outlook for a completed state of demographic transition for the world population as a whole generally appears positive if not inevitable, although demographers are deeply divided on estimates of the size of world population at equilibrium, the timing of completed transition, the principal mechanisms at work, and the long-term ecological consequences. Certainly, the world population will continue to grow for some period of time, if only as a consequence of the previous momentum of high fertility relative to mortality. Most if not all demographers, however, subscribe to the view expressed by Coale (1974, p. 51) that the entire process of global demographic transition and the phase of phenomenal population growth that has accompanied

it will be a transitory (albeit spectacular) episode in human population history.

REFERENCES

Bongaarts, John 1975 “Why High Birth Rates are So Low.” *Population and Development Review* 1:289–296.

Caldwell, John 1976 “Toward a Restatement of Demographic Transition Theory.” *Population and Development Review* 2:321–366.

— 1986 “Routes to Low Mortality in Poor Countries.” *Population and Development Review* 12:171–220.

Cleland, John, and Christopher Wilson 1987 “Demand Theories of the Fertility Transition: An Iconoclastic View.” *Population Studies* 41:5–30.

Coale, Ansley 1973 “The Demographic Transition.” *International Population Conference, IUSSP*. Liege, Belgium.

— 1974 “The History of Human Population.” *Scientific American* 23:41–51.

- 1983 “Recent Trends in Fertility in Less Developed Countries.” *Science* 221:828–832.
- , and Roy Treadway 1986 “A Summary of the Changing Distribution of Overall Fertility, Marital Fertility, and the Proportion Married in the Provinces of Europe.” In Ansley Coale and Susan Watkins, eds., *The Decline of Fertility in Europe*. Princeton: Princeton University Press.
- Davis, Kingsley 1967 “Population Policy: Will Current Programs Succeed?” *Science* 158:730–739.
- Enke, Stephen 1967 “The Economic Case for Birth Control.” *Challenge/Magazine of Economic Affairs* 15:30–31.
- Freedman, Ronald 1998 “Observing Taiwan’s Demographic Transition: A Memoir.” Research Report 98–426. Ann Arbor: Population Studies Center, University of Michigan.
- Guest, Avery 1981 “Social Structure and U.S. Inter-State Fertility Differentials in 1900.” *Demography* 18:453–486.
- and Stewart Tolnay 1983 “Children’s Roles and Fertility: Late Nineteenth Century United States.” *Social Science History* 7:355–380.
- Hirschman, Charles 1994 “Why Fertility Changes.” *Annual Review of Sociology* 20:203–233.
- Knodel, John 1977 “Family Limitation and the Fertility Transition: Evidence from the Age Patterns of Fertility in Europe and Asia.” *Population Studies* 32:481–510.
- , and Etienne van de Walle 1979 “Lessons from the Past: Policy Implications of Historical Fertility Studies.” *Population and Development Review* 5:217–245.
- Lesthaeghe, Ron 1983 “A Century of Demographic Change in Western Europe: An Exploration of Underlying Dimensions.” *Population and Development Review* 9:411–436.
- Mason, Karen Oppenheim 1997 “Explaining Fertility Transitions.” *Demography* 34:443–454.
- Matthiessen, Paul C., and James C. McCann 1978 “The Role of Mortality in the European Fertility Transition: Aggregate-level Relations.” In Samuel H. Preston, ed., *The Effects of Infant and Child Mortality on Fertility*. New York: Academic Press.
- Mauldin, W. Parker, and Bernard Berelson 1978 “Conditions of Fertility Decline in Developing Countries.” *Studies in Family Planning* 9:89–148.
- McKeown, Thomas 1976 *The Modern Rise of Population*. New York: Academic Press.
- Notestein, Frank 1953 “Economic Problems of Population Change.” *Proceedings of the Eighth International Conference of Agricultural Economics*, 13–31. London: Oxford University Press.
- Palloni, Alberto 1981 “Mortality in Latin America: Emerging Patterns.” *Population and Development Review* 7:623–649.
- Population Reference Bureau 1986 *World Population Data Sheet*. Washington, D.C.
- 1998 *World Population Data Sheet*. Washington, D.C.
- Preston, Samuel 1975 “The Changing Relationship Between Mortality and Level of Economic Development.” *Population Studies* 29:231–248.
- 1976 *Mortality Patterns in National Populations*. New York: Academic Press.
- , and Michael Haines 1991 *Fatal Years: Child Mortality in Late Nineteenth-Century America*. Princeton: Princeton University Press.
- Razzell, P.E. 1974 “An Interpretation of the Modern Rise of Population in Europe - a Critique.” *Population Studies* 28:5–17.
- van de Walle, Francine 1980 “Education and the Demographic Transition in Switzerland.” *Population and Development Review* 6:463–472.
- Watkins, Susan 1991 *From Provinces into Nations: Demographic Integration in Western Europe, 1870–1960*. Princeton: Princeton University Press.

AVERY M. GUEST  
GUNNAR ALMGREN

## DEMOGRAPHY

Demography is the study of human populations. It is an important part of sociology and the other social sciences because all persisting social aggregates—societies, states, communities, racial or ethnic groups, professions, formal organizations, kinship groups, and so on—are also populations. The size of the population, its growth or decline, the location and spatial movement of its people, and their changing characteristics are important features of an aggregate whether one sees it as a culture, an economy, a polity, or a society. As a result some anthropologists, economists, historians, political scientists, and sociologists are also demographers, and most demographers are members of one of the traditional social science disciplines.

A central question for each of the social sciences is: How does the community, society, or whatever, seen as a culture, an economy, a polity, or whatever, produce and renew itself over the

years? Formal demography answers this question for aggregates seen as populations. This formal part of demography is fairly independent of the traditional social sciences and has a lengthy history in mathematics and statistics (Smith and Keyfitz 1977; Stone 1997; Desrosières 1998). It depends on a definition of age and on the relationship of age to fertility and mortality. Those relationships certainly are socially conditioned, but their major outlines are constrained by biology.

Beyond pursuing formal demography, the task of most social scientist-demographers is detailing the relationships between demographic change and other aspects of social change. Working with concepts, methods, and questions arising from the traditions of each of the social science disciplines as well as those of demography per se, scholars have investigated the relationship between demographic changes and such social changes as those in the nature of families (Davis 1985; Sweet and Bumpass 1987; Bumpass 1990; Waite 1995), levels of economic growth (Johnson and Lee 1987; Nerlove and Raut 1997), the development of colonialism (McNeill 1990), changes in kinship structures (Dyke and Morrill 1980), nationalism and interethnic strife (Tietelbaum and Winter 1998), and the development of the nation-state (Watkins 1991).

### FORMAL DEMOGRAPHY

At the heart of demography is a body of strong and useful mathematical theory about how populations renew themselves (Keyfitz 1968, 1985; Coale 1972; Bourgeois-Pichat 1994). The theory envisions a succession of female birth cohorts living out their lives subject to a schedule of age-specific mortality chances and age-specific chances of having a female baby. In the simplest form of the model the age-specific rates are presumed constant from year to year.

Each new annual birth cohort is created because the age-specific fertility rates affect women in earlier birth cohorts who have come to a specific age in the year in question. Thus, the mothers of a new cohort of babies are spread among previous cohorts. The size of the new cohort is a weighted average of the age-specific fertility rates. The sizes of preceding cohorts, survived to the year in question, are the weights.

As a cohort of women age through their fertile period, they die and have children in successive years according to the age-specific rates appropriate to those years. Thus, the children of a single birth cohort of women are spread over a sequence of succeeding birth cohorts.

The number of girls ever born to a birth cohort, taken as a ratio to the initial size of the cohort, is implicit in the age-specific fertility and mortality rates. This ratio, called the net reproduction rate, describes the growth rate over a generation that is implicit in the age-specific rates. The length of this generation is also implicit in the age-specific rates as the average age of mothers at the birth of the second-generation daughters. With a rate of increase over a generation and a length of the generation, it is clear that an annual rate of increase is intrinsic to the age-specific rates.

The distribution of the children of a birth cohort over a series of succeeding cohorts has an important effect. If an unusually small or large birth cohort is created, the effects of its largeness or smallness will be distributed among a number of succeeding cohorts. In each of those succeeding cohorts, the effect of the unusual cohort is averaged with that of other birth cohorts to create the new cohort's size. Those new cohorts' "inherited" smallness or largeness, now diminished by averaging, will also be spread over succeeding cohorts. In a few generations the smallness or largeness will have averaged out and no reflection of the initial disturbance will be apparent.

Thus, without regard to peculiarities in the initial age distribution, the eventual age distribution of a population experiencing fixed age-specific fertility and mortality will become proportionately constant. As this happens, the population will take on a fixed aggregate birth and death rate and, consequently, a fixed rate of increase. The population so created is called a stable population and its rates, called intrinsic rates, are those implicit in the net reproduction rate and the length of a generation. Such rates, as well as the net reproduction rates, are frequently calculated for the age-specific fertility and mortality rates occurring for a single year as a kind of descriptive, "what if" summary.

This theory is elaborated in a number of ways. In one variant, age-specific rates are not constant but change in a fixed way (Lopez 1961). In another

elaboration the population is divided into a number of states with fixed age-specific migration or mobility among the states (Land and Rogers 1982; Schoen 1988). States may be geographic regions, marital circumstances, educational levels, or whatever.

In part, the value of this theory is in the light it sheds on how populations work. For example, it explains how a population can outlive all of its contemporary members and yet retain its median age, percent in each race, and its regional distribution.

The fruit of the theory lies in its utility for estimation and forecasting. Using aspects of the theory, demographers are able to elaborate rather modest bits and pieces of information about a population to a fairly full description of its trajectory (Coale and Demeny 1983). Combined with this mathematical theory is a body of practical forecasting techniques, statistical estimation procedures, and data-collection wisdom that makes up a core area in demography that is sometimes called formal demography (United Nations 1983; Shryock and Siegel 1976; Pollard, Yusuf, and Pollard 1990; Namboodiri 1991; Hinde 1998).

#### DEMOGRAPHIC DATA

Generating the various rates and probabilities used in formal demography requires two different kinds of data. On the one hand are data that count the number of events occurring in the population in a given period of time. How many births, deaths, marriages, divorces, and so on have occurred in the past year? These kinds of data are usually collected through a vital registration system (National Research Council 1981; United Nations 1985, 1991). On the other hand are data that count the number of persons in a given circumstance at a given time. How many never-married women age twenty to twenty-four were there on July 1? These kinds of data are usually collected through a population census or large-scale demographic survey (United Nations 1992; Anderson 1988; Anderson and Fienberg 1999). From a vital registration system one gets, for example, the number of births to black women age twenty. From a census one collects the number of black women at age twenty. The division of the number of events by the population exposed to the risk of having the event occur to them yields the demographic rate, that is, the fertility rate for black women age twenty. These

two data collection systems—vital statistics and census—are remarkably different in their character. To be effective, a vital statistics system must be ever alert to see that an event is recorded promptly and accurately. A census is more of an emergency. Most countries conduct a census every ten years, trying to enumerate all of the population in a brief time.

If a vital registration system had existed for a long time, were very accurate, and there were no uncounted migrations, one could use past births and deaths to tally up the current population by age. To the degree that such a tallying up does not match a census, one or more of the data collections systems is faulty.

#### SOCIAL DEMOGRAPHY

One of the standard definitions of demography is that given by Hauser and Duncan: “Demography is the study of the size, territorial distribution and composition of populations, changes therein, and the components of such change, which may be identified as natality, mortality, territorial movement (migration) and social mobility (change of status)” (1959, p. 2). Each of these parts—size, territorial distribution, and composition—is a major arena in which the relationships between demographic change and social change are investigated by social scientist-demographers. Each part has a somewhat separate literature, tradition, method, and body of substantive theory.

*Population Size.* The issue of population size and change in size is dominated by the shadow of Malthus (Malthus 1959), who held that while food production can grow only arithmetically because of diminishing returns to investment in land and other resources, population can grow geometrically and will do so, given the chance. Writing in a time of limited information about birth control and considerable disapproval of its use, and holding little hope that many people would abstain from sexual relations, Malthus believed that populations would naturally grow to the point at which starvation and other deprivations would curtail future expansion. At that point, the average level of living would be barely above the starvation level. Transitory improvements in the supply of food would only lead to increased births and subsequent deaths as the population returned to its

equilibrium size. Any permanent improvement in food supplies due to technological advances would, in Malthus's theory, simply lead to a larger population surviving at the previous level of misery.

Although there is good evidence that Malthus understood his contemporary world quite well (Lee 1980; Wrigley 1983), he missed the beginnings of the birth control movement, which were contemporaneous (McLaren 1978). The ability to limit births, albeit at some cost, without limiting sexual activity requires important modifications to the Malthusian model.

Questions of the relationship among population growth, economic growth, and resources persist into the contemporary period. The bulk of the literature is in economics. A good summary of that literature can be found in T. P. Schultz (1981), in Rosensweig and Stark (1997), and in a National Research Council report on population policy (1985). A more polemical treatment, but one that may be more accessible to the noneconomist, is offered by the World Bank (1985).

Sharing the study of population size and its change with Malthusian issues is a body of substantive and empirical work on the *demographic revolution* or *transition* (Notestein 1945). The model for this transition is the course of fertility and mortality in Europe during the Industrial Revolution. The transition is thought to occur concurrently with "modernization" in many countries (Coale and Watkins 1986) and to be still in process in many less developed parts of the world (United Nations 1990). This transition is a change from (1) a condition of high and stable birthrates combined with high and fluctuating death rates, through (2) a period of initially lowering death rates and subsequently lowering birthrates, to (3) a period of low and fairly constant death rates combined with low and fluctuating birthrates. In the course of part (2) of this transition, the population grows very considerably because the rate of increase, absent migration, is the difference between the birthrate and the death rate.

In large measure because of anxiety that fertility might not fall rapidly enough in developing countries, a good deal of research has focused on the fertility part of the transition. One branch of this research has been a detailed historical investigation of what actually happened in Europe, since that is the base for the analogy about what is

thought likely to happen elsewhere (Coale and Watkins 1986). A second branch was the World Fertility Survey, perhaps the largest international social science research project ever undertaken. This project conducted carefully designed, comparable surveys with 341,300 women in seventy-one countries to investigate the circumstances of contemporary fertility decline (Cleland and Hobcraft 1985). In 1983, a continuation of this project was undertaken under the name Demographic and Health Surveys. To date, this project has provided technical assistance for more than 100 surveys in Africa, Asia, the Near East, Latin America, and the Caribbean. For more complete information about current activities see their web site at <http://www.macrpoint.com/dhs/>.

Scholars analyzing these projects come to a fairly similar assessment of the roots of historical and contemporary fertility decline as centering in an increased secular rationality and growing norms of individual responsibility.

Detailed investigation of the mortality side of the European demographic transition has primarily been conducted by historians. A particularly useful collection of papers is available in Schofield, Reher, and Bideau (1991). For the United States, a particularly useful book on child and infant mortality at the turn of the twentieth century is by Haines and Preston (1991).

The utility of the idea of a demographic transition to understand population change in the developing world has, of course, been a matter of considerable debate. A good summary of this literature is found in Jones et al. (1997).

*Territorial Distribution.* Research on the territorial distribution of populations is conducted in sociology, geography, and economics. The history of population distribution appears to be one of population dispersion at the macroscopic level of continents, nations, or regions, and one of population concentration at the more microscopic level of larger towns and cities.

The diffusion of the human population over the globe, begun perhaps as the ice shields retreated in the late Pleistocene, continues to the present (Barracough 1978; Bairoch 1988). More newly inhabited continents fill up and less habitable land becomes occupied as technological and social change makes it possible to live in previously

remote areas. Transportation lines, whether caravan, rail, or superhighway, extend across remote areas to connect distant population centers. Stops along the way become villages and towns specializing in servicing the travelers and the goods in transit. These places are no longer "remote." Exploitation of resources proximate to these places may now become viable because of the access to transportation as well as services newly available in the stopover towns.

As the increasing efficiency of agriculture has released larger and larger fractions of the population from the need to till the soil, it has become possible to sustain increasing numbers of urban people. There is a kind of urban transition more or less concomitant with the demographic transition during which a population's distribution by city size shifts to larger and larger sizes (Kelley and Williamson 1984; Wrigley 1987). Of course, the continuing urban transition presents continuing problems for the developing world (United Nations 1998).

*Population Composition.* The characteristics included as compositional ones in demographic work are not predefined theoretically, with the exception, perhaps, of age and sex. In general, compositional characteristics are those characteristics inquired about on censuses, demographic surveys, and vital registration forms. Such items vary over time as social, economic, and political concerns change. Nonetheless, it is possible to classify most compositional items into one of three classes. First are those items that are close to the reproductive core of a society. They include age, sex, family relationships, and household living arrangements. The second group of items are those characteristics that identify the major, usually fairly endogamous, social groups in the population. They include race, ethnicity, religion, and language. Finally there is a set of socioeconomic characteristics such as education, occupation, industry, earnings, and labor force participation.

Within the first category of characteristics, contemporary research interest has focused on families and households because the period since 1970 has seen such dramatic changes in developed countries (Van de Kaa 1987; Davis 1985; Farley 1996, ch. 2). Divorce, previously uncommon, has become a common event. Many couples now live together without record of a civil or religious

ceremony having occurred. Generally these unions are not initially for the purposes of procreation, although children are sometimes born into them. Sometimes they appear to be trial marriages and are succeeded by legal marriages. A small but increasing fraction of women in the more developed countries seem to feel that a husband is not a necessary ornament to motherhood. Because of these changes, older models of how marriage comes about and how marriage relates to fertility (Coale and McNeil 1972; Coale and Trussell 1974) are in need of review as demographers work toward a new demography of the family (Bongaarts 1983; Keyfitz 1987).

Those characteristics that indicate membership in one or another of the major social groupings within a population vary from place to place. Since such social groupings are the basis of inequality, political division, or cultural separation, their demography becomes of interest to social scientists and to policy makers (Harrison and Bennett 1995). To the degree that endogamy holds, it is often useful to analyze a social group as a separate population, as is done for the black population in the United States (Farley 1970). Fertility and mortality rates, as well as marital and family arrangements, for blacks in the United States are different from those of the majority population. The relationship between these facts, their implications, and the socioeconomic discrimination and residential segregation experienced by this population is a matter of historic and continuing scholarly work (Farley and Allen 1987; Lieberman 1980; Farley 1996, ch. 6).

For other groups, such as religious or ethnic groups in the United States, the issue of endogamy becomes central in determining the continuing importance of the characteristic for the social life of the larger population (Johnson 1980; Kalmijn 1991). Unlike the black population, ethnic and religious groups seem to be of decreasing appropriateness to analyze as separate populations within the United States, since membership may be a matter of changeable opinion.

The socioeconomic characteristics of a population are analyzed widely by sociologists, economists, and policy-oriented researchers. Other than being involved in the data production for much of this research, the uniquely demographic contributions come in two ways. First is the consideration

of the relationship between demographic change and change in these characteristics. The relationship between female labor force participation and changing fertility patterns has been a main topic of the "new home economics" within economic demography (Becker 1960; T. W. Schultz 1974; T. P. Schultz 1981; Bergstrom 1997). The relationship between cohort size and earnings is a topic treated by both economists and sociologists (Winsborough 1975; Welch 1979). A system of relationships between cohort size, economic well-being, and fertility has been proposed by Easterlin (1980) in an effort to explain both fertility cycles and long swings in the business cycle.

A second uniquely demographic contribution to the study of socioeconomic characteristics appears to be the notion of a cohort moving through its socioeconomic life course (Duncan and Hodge 1963; Hauser and Featherman 1977; Mare 1980). The process of leaving school, getting a first job, then subsequent jobs, each of which yields income, was initially modeled as a sequence of recursive equations and subsequently in more detailed ways. Early in the history of this project it was pointed out that the process could be modeled as a multistate population (Matras 1967), but early data collected in the project did not lend itself to such modeling and the idea was not pursued.

### POPULATION POLICY

The consideration and analysis of various population policies is often seen as a part of demography. Population policy has two important parts. First is policy related to the population of one's own nation. The United Nations routinely conducts inquiries about the population policies of member nations (United Nations 1995). Most responding governments claim to have official positions about a number of demographic issues, and many have policies to deal with them. It is interesting to note that the odds a developed country that states its fertility is too low has a policy to raise the rate is about seven to one, while the odds that a less developed country that states that its fertility is too high has a policy to lower the rate is about 4.6 to one. The prospect of declining population in the United States has already begun to generate policy proposals (Teitelbaum and Winter 1985).

The second part of population policy is the policy a nation has about the population of other

countries. For example, should a country insist on family-planning efforts in a developing country prior to providing economic aid? A selection of opinions and some recommendations for policy in the United States are provided in Menken (1986) and in the National Research Council (1985).

### IMPORTANT CURRENT RESEARCH AREAS

Three areas of demographic research are of especially current interest. Each is, in a sense, a sequela of the demographic transition. They are:

1. Institutional arrangements for the production of children,
2. Immigration, emigration, ethnicity, and nationalism,
3. Population aging, morbidity, and mortality.

**Institutional arrangements for the production of children.** The posttransition developed world has seen dramatic changes in the institutional arrangements for the production of children. Tracking these changes has become a major preoccupation of contemporary demography. Modern contraception has broken the biological link between sexual intercourse and pregnancy. Decisions about sexual relations can be made with less concern about pregnancy. The decision to have a child is less colored by the need for sexual gratification. People no longer must choose between the celibate single life or the succession of pregnancies and children of a married one. Women are employed before, during, and after marriage. Their earning power makes marriage more an option and a context for child raising than an alternative employment. Women marry later, are older when they have children, and are having fewer children than in the past. Divorce has moved from an uncommon event to a common one. Cohabitation is a new, and somewhat inchoate, lifestyle that influences the circumstances for childbearing and child rearing. Although only a modest number of children are currently born to cohabiting couples, most children will spend some time living with a parent in such a relationship before they are adults (Bumpass and Lu 1999). Changes such as these have occurred in most of Europe and North America and appear to be increasing in those parts of Asia where the demographic transitions occurred first. The changes have, of course, been met with

considerable resistance in many places. Will they continue to occur in other countries as the demographic transitions proceed around the world? How serious will the resistance be in countries more patriarchal than those of Europe? These issues await future demographic research.

**Immigration, emigration, ethnicity, and nationalism.** The difference between immigration and emigration is, of course, net migration as a component of population growth. In general, the direct impact of net migration on growth has been modest. Certainly emigration provided some outlet for population growth in Europe during the transition (Curtin 1989) and played a role in the population of the Americas, Australia, and New Zealand. Most of the effect of immigration to the New World was indirect, however. It was the children of immigrants who peopled the continents rather than the immigrants themselves. Contemporary interest in international migration has three sources. First is an abiding concern for refugees and displaced persons. Second is interest in a highly mobile international labor force in the construction industry. Third, and perhaps most important, is the migration of workers from less developed countries to developed ones in order to satisfy the demand for labor in a population in real or incipient decline (Teitelbaum and Winter 1998). In France, Germany, and England immigrants who come and stay change the ethnic mix of the population. In the United States, legal and illegal immigration from Mexico has received considerable attention (Chiswick and Sullivan 1995). The immigration of skilled Asians raises issues of another kind. In all of these countries, the permanent settlement of new immigrants arouses powerful nationalistic concerns. The situation provides ample opportunity for the investigation of the development of ethnicity and the place of endogamy in its maintenance. As with changes for child rearing, these changes seem consequences of the demographic transition. Below-replacement fertility seems to generate the need for imported workers and their consequent inclusion in the society. If this speculation is true, how will countries more recently passing through the transition react to this opportunity and challenge? This is another arena for future demographic investigation.

**Population aging, morbidity, and mortality.** Accompanying the reduction of fertility and of

population growth is the aging of the population (Treas and Torrecilha 1995). Much of the developed world is experiencing this aging and its concomitants. Countries that made a rapid demographic transition have especially dramatic imbalances in their age distribution. Concerns about retirement, health, and other programs for the elderly generate interest in demographic research in these areas. Will increasing life expectancy add years to the working life so that retirement can begin later? Or will the added years all be spent in nursing homes? These questions are currently of great policy interest and pose interesting and difficult problems for demographic research (Manton and Singer 1994). Traditional demographic modeling has assumed that frailty, the likelihood of death, varies according to measurable traits such as age, sex, race, education, wealth, and so on, but are constant within joint values of these variables. But what happens if frailty is a random variable? In survival models, randomness in frailty is not helpful, or at least benign, as it is in so many other statistical circumstances. Rather, it can have quite dramatic effects (Vaupel and Yashin 1985). How one models frailty appears to make a considerable amount of difference in the answers one gets to important policy questions surrounding population aging.

#### DEMOGRAPHY AS A PROFESSION

Most demographers in the United States are trained in sociology. Many others have their highest degree in economics, history, or public health. A few are anthropologists, statisticians, or political scientists. Graduate training of demographers in the United States and in much of the rest of the world now occurs primarily in centers. Demography centers are often quasi-departmental organizations that serve the research and training needs of scholars in several departments. In the United States there are about twenty such centers, twelve of which have National Institutes of Health grants. The Ford Foundation has supported similar demography centers at universities in the less developed parts of the world.

Today, then, most new demographers, without regard for their disciplinary leanings, are trained at a relatively few universities. Most will work as faculty or researchers within universities or at



demography centers. Another main source of employment for demographers is government agencies. Census bureaus and vital statistics agencies both provide much of the raw material for demographic work and employ many demographers around the world. There is a small but rapidly growing demand for demographers in the private sector in marketing and strategic planning.

Support for research and training in demography began in the United States in the 1920s with the interest of the Rockefeller Foundation in issues related to population problems. Its support led to the first demography center, the Office of Population Research at Princeton University. The Population Council in New York was established as a separate foundation by the Rockefeller brothers in the 1930s. Substantial additional foundation support for the field has come from the Ford Foundation, the Scripps Foundation, and, more recently, the Hewlett Foundation. Demography was the first of the social sciences to be supported by the newly founded National Science Foundation in the immediate post-World War II era. In the mid-1960s the National Institute of Child Health and Human Development undertook support of demographic research.

(SEE ALSO: *Birth and Death Rates; Census; Demographic Methods; Demographic Transition; Life Expectancy; Population*)

## REFERENCES

- Anderson, Margo 1988 *The American Census: A Social History*. New Haven: Yale University Press.
- , and Stephen Fienberg 1999 *Who Counts: The Politics of Census-Taking in Contemporary America*. New York: Russell Sage Foundation.
- Barracough, Geoffrey (ed.) 1978 *The Times Atlas of World History*. London: Times Books.
- Becker, G. S. 1960 "An Economic Analysis of Fertility." In Becker, ed., *Demographic and Economic Change in Developed Countries*. Princeton, N.J.: National Bureau of Economic Research.
- Bergstrom, Theodore 1997 "A Survey of Theories of the Family." In Mark Rosenzweig and Oded Stark, eds., *Handbook of Family and Population Economics*. Amsterdam: North-Holland.
- Bongaarts, John 1983 "The Formal Demography of Families and Households: An Overview." *IUSSP Newsletter* 17:27–42.
- Bourgeois-Pichat, Jean 1994 *La Dynamique Des Populations: Populations Stables, Semi-stable, Quasi-stables*. Paris: Presses Universitaires de France.
- Bumpass, Larry L. 1990 "What's Happening to the Family? Interactions Between Demographic and Institutional Change." *Demography* 27:483–498.
- , and H-H Lu 1999 "Trends in Cohabitation and Implications for Children's Family Contexts in the U.S." Forthcoming in *Population Studies*.
- Chiswick, Barry, and Teresa Sullivan 1995 "The New Immigrants." In Reynolds Farley, ed., *State of the Union: America in the 1990s, V2, Social Trends*. New York: Russell Sage Foundation.
- Cleland, John, and John Hobcraft 1985 *Reproductive Change in Developing Countries: Insights from the World Fertility Survey*. London: Oxford University Press.
- Coale, Ansley 1972 *The Growth and Structure of Human Populations: A Mathematical Investigation*. Princeton, N.J.: Princeton University Press.
- , and Paul Demeny 1983 *Regional Model Life Tables and Stable Populations*. New York: Academic Press.
- Coale, Ansley, and D. R. McNeil 1972 "Distribution by Age of Frequency of First Marriage in a Female Cohort." *Journal of the American Statistical Association* 67:743–749.
- , and James Trussell 1974 "Model Fertility Schedules: Measurement and Use in Fertility Models." *Population Index* 40:182–258.
- Coale, Ansley, and Susan C. Watkins (eds.) 1986 *The Decline of Fertility in Europe*. Princeton, N.J.: Princeton University Press.
- Curtin, Philip 1989 *Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century*. Cambridge, Mass.: Cambridge University Press.
- Davis, Kingsley (ed.) 1985 *Contemporary Marriage*. New York: Russell Sage Foundation.
- Desrosières, Alain 1998 *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Cambridge, Mass.: Harvard University Press.
- Duncan, O. D., and William Hodge 1963 "Education and Occupational Mobility." *American Journal of Sociology* 68:629–644.
- Dyke, Bennett, and Warren Morrill (eds.) 1980 *Genealogical Demography*. New York: Academic Press.
- Easterlin, Richard 1980 *Birth and Fortune: The Impact of Numbers on Personal Welfare*. New York: Basic Books.
- Farley, Reynolds 1970 *Growth of the Black Population: A Study of Demographic Trends*. Chicago: Markham.

- (ed.) 1995 *State of the Union: America in the 1990's*. Vol. 1, *Economic Trends*. Vol. 2, *Social Trends*. New York: Russell Sage Foundation.
- 1996 *The New American Reality: Who We Are, How We Got There, Where We Are Going*. New York: Russell Sage Foundation.
- , and Walter R. Allen 1987 *The Color Line and the Quality of Life in America*. New York: Russell Sage Foundation.
- Haines, M., and S. Preston 1991 *The Fatal Years: Child Mortality in Late Nineteenth-Century America*. Princeton, N.J.: Princeton University Press.
- Harrison, Roderick, and Claudette Bennett 1995 "Racial and Ethnic Diversity." In Farley, ed., *State of the Union: America in the 1990s*, V2, *Social Trends*. New York: Russell Sage Foundation.
- Hauser, Philip M., and Otis Dudley Duncan 1959 *The Study of Population: An Inventory and Appraisal*. Chicago: University of Chicago Press.
- Hauser, Robert M., and David L. Featherman 1977 *The Process of Stratification: Trends and Analyses*. New York: Academic Press.
- Hinde, Andrew 1998 *Demographic Methods*. London: Arnold.
- Johnson, D. Gale, and Ronald D. Lee (eds.) 1987 *Population Growth and Economic Development: Issues and Evidence*. Madison: University of Wisconsin Press.
- Johnson, Robert A. 1980 *Religious Assortative Marriage in the United States*. New York: Academic Press.
- Jones, Gavin et al. (eds.) 1997 *The Continuing Demographic Transition*. Oxford: Clarendon Press.
- Kalmijn, Matthijs 1991 "Shifting Boundaries: Trends in Religious and Educational Homogamy." *American Sociological Review* 56:786–800.
- Kelley, Allen, and Jeffrey Williamson 1984 "Modeling the Urban Transition." *Population and Development Review* 10:419–442.
- Keyfitz, Nathan 1968 *Introduction to the Mathematics of Populations*. Reading, Mass.: Addison-Wesley.
- 1985 *Applied Mathematical Demography*. New York: Springer-Verlag.
- 1987 "Form and Substance in Family Demography." In J. Bongaarts, T. Burch, and K. Wachter, eds., *Family Demography*. New York: Oxford University Press.
- Land, Kenneth C., and Andrei Rogers (eds.) 1982 *Multi-dimensional Mathematical Demography*. New York: Academic Press.
- Lee, Ronald 1980 "An Historical Perspective on Economic Aspects of the Population Explosion: The Case of Preindustrial England." In Richard Easterlin, ed., *Population and Economic Change in Developing Countries*. Chicago: University of Chicago Press.
- Lieberson, Stanley 1980 *A Piece of the Pie: Black and White Immigrants Since 1880*. Berkeley: University of California Press.
- Lopez, A. 1961 *Some Problems in Stable Population Theory*. Princeton, N.J.: Office of Population Research.
- Malthus, Thomas R. 1959 *Population: The First Essay*. Ann Arbor: University of Michigan Press.
- Mare, Robert D. 1985 "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75:295–305.
- Matras, Judah 1967 "Social Mobility and Social Structure: Some Insights from the Linear Model." *American Sociological Review* 33:608–614.
- McLaren, Angus 1978 *Birth Control in Nineteenth-Century England*. New York: Holmes & Meier.
- McNeill, William H. 1990 *Population and Politics Since 1750*. Charlottesville: University of Virginia Press.
- Menken, Jane (ed.) 1986 *World Population and U.S. Policy: The Choices Ahead*. New York: W. W. Norton.
- Namboodiri, Krishnan 1991 *Demographic Analysis: A Stochastic Approach*. San Diego: Academic Press.
- National Research Council 1981 *Collecting Data for Estimation of Fertility and Mortality*. Report no. 6, Committee on Population and Demography, Assembly of Behavioral and Social Sciences, National Research Council. Washington, D.C.: National Academy Press.
- Working Group on Population Growth and Economic Development, Committee on Population, Commission on Behavioral and Social Sciences and Education 1985 *Population Growth and Economic Development: Policy Questions*. Washington, D.C.: National Academy Press.
- Nerlove, Marc, and Lakshmi Raut 1997 "Growth Models with Endogenous Population: A General Framework." In Mark Rosenzweig and Oded Stark, eds., *Handbook of Family and Population Economics*. Amsterdam: North-Holland.
- Notestein, Frank W. 1945 "Population—The Long View." In Theodore W. Schultz, ed., *Food for the World*. Chicago: University of Chicago Press.
- Pollard, A. H., Farhat Yusuf, and G. N. Pollard 1990 *Demographic Techniques*. Sydney: Pergamon Press.
- Rosenzweig, Mark, and Oded Stark (eds.) 1997 *Handbook of Population and Family Economics*. Amsterdam: Elsevier.
- Schoen, Robert 1988 *Modeling Multigroup Populations*. New York: Plenum Press.

- Schofield, R., D. Reher, and A. Bideau (eds.) *The Decline of Mortality in Europe*. Oxford: Clarendon Press.
- Schultz, T. Paul 1981 *Economics of Population*. Reading, Mass.: Addison-Wesley.
- Schultz, T. W. (ed.) 1974 *Economics of the Family*. Chicago: University of Chicago Press.
- Shryock, Henry S., and Jacob S. Siegel 1976 *The Methods and Materials of Demography*. New York: Academic Press.
- Singer, B., and K. Manton 1994 "What's the Fuss About the Compression Morbidity?" *Chance*. Fall 1994.
- Smith, David P., and Nathan Keyfitz (eds.) 1977 *Mathematical Demography: Selected Essays*. Berlin and New York: Springer-Verlag.
- Stone, Richard 1997 *Some British Empiricists in the Social Sciences: 1650–1900*. Cambridge, Mass.: Cambridge University Press.
- Sweet, James A., and Larry L. Bumpass 1987 *American Families and Households*. New York: Russell Sage Foundation.
- Teitelbaum M., and J. Winter 1985 *The Fear of Population Decline*. Orlando, Fla.: Academic Press.
- 1998 *A Question of Numbers: High Migration, Low Fertility, and the Politics of National Identity*. New York: Wang and Hill.
- Treas, Judith, and Ramon Torrecilha 1995 "The Older Population." In Reynolds Farley, ed., *State of the Union: The 1990 Census Research Project*. New York: Russell Sage Foundation.
- United Nations, Department of International Economic and Social Affairs 1983 *Manual X: Indirect Techniques for Demographic Estimation*. Population Studies, no. 81, ST/ESA/SER.A/81.
- 1985 *Handbook of Vital Statistics Systems and Methods, Volume II: Review of National Practices*. ST/STAT/SER.F/35 V2.
- 1991 *Handbook of Vital Statistics Systems and Methods, Volume I: Legal, Organizational and Technical Aspects*. ST/STAT/SER.F/35V.1.
- 1992 *Handbook of Population and Housing Censuses*. ST/STAT/SER.F/54 V.1,2.
- 1995 *Results of the Seventh United Nations Population Inquiry Among Governments*. ST/ESA/SER.R/140.
- 1998 *World Urbanization Prospects: The 1996 Revision*. ST/ESA/SER.A/170.
- Van de Kaa, Dirk 1987 "Europe's Second Demographic Transition." *Population Bulletin* 42:1.
- Vaupel, James W., and Anatoli I.Yashin. 1985 "Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics." *American Statistician* 39:176–185.
- Waite, Linda 1995 "Does Marriage Matter?" *Demography* 32:483–507.
- Watkins, Susan C. 1991 *From Provinces into Nations: Demographic Integration in Western Europe, 1870–1960*. Princeton, N.J.: Princeton University Press.
- Welch, F. 1979 "Effect of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy* 2:565–598.
- Winsborough, H. 1975 "Age, Period, Cohort, and Education Effects of Earnings by Race—An Experiment with a Sequence of Cross-Sectional Surveys." In Kenneth Land and Seymour Spilerman, eds., *Social Indicator Models*. New York: Russell Sage Foundation.
- World Bank 1985 *Population Change and Economic Development*. New York: Oxford University Press.
- Wrigley, E. A. 1987 *People, Cities and Wealth: The Transformation of Traditional Society*. Oxford: Basil Blackwell.

HALLIMAN H. WINSBOROUGH

## DEPENDENCY THEORY

Conceived in the 1960s by analysts native to developing countries, dependency theory is an alternative to Eurocentric accounts of modernization as universalistic, unilinear evolution (Addo 1996). Instead, contemporary underdevelopment is seen as an outgrowth of asymmetrical contacts with capitalism. The thesis is straightforward: Following the first wave of modernization, less-developed countries are transformed by their interconnectedness with other nations, and the nature of their contacts, economies, and ideologies (Keith 1997). Interaction between social orders is never merely a benign manifestation of economic or cultural diffusion. Rather, both the direction and pace of change leads to internal restructuring designed to bolster the interests of the more powerful exchange partner without altering the worldwide distribution of affluence.

Widely utilized as both a heuristic and empirical template, dependency theory emerged from neo-Marxist critiques of the failure of significant capital infusion, through the United Nations Economic Commission for Latin America, to improve overall quality of life or provide significant returns

to countries in the region. Dependency theory quickly became an instrument for political commentary as well as an explanatory framework as it couched its arguments in terms of the consequences of substituting cash crops for subsistence farming and replacing local consumer goods with export commodities destined for developed markets. Frank (1967, 1969), an early proponent, established the premise for dependency theory; contemporary underdevelopment is a result of an international division of labor exploited by capitalist interests.

Frank's innovation was to incorporate the vantagepoint of underdeveloped countries experiencing capital infusion into the discussion of the dynamics of economic, social, and political change. In so doing he discounted Western and European models of modernization as self-serving, ethnocentric, and apolitical. He asserted that explanations of modernization have been disassociated from the colonialism that fueled industrial and economic developments characteristic of the modern era. However, much of the first wave of modernization might have been driven by intrinsic, internal factors; more recent development has taken place in light of change external to individual countries. Central to Frank's contention was a differentiation of undeveloped from underdeveloped countries. In the latter, a state of dependency exists as an outgrowth of a locality's colonial relationship with "advanced" areas. Frank referred to the "development of underdevelopment" and the domination of development efforts by advanced countries, using a "metropolis-satellite" analogy to denote the powerful center out of which innovations emerge and a dependent hinterland is held in its sway. He spoke of a chain of exploitation—or the flow and appropriation of capital through successive metropolis-satellite relationships, each participating in a perpetuation of relative inequalities even while experiencing some enrichment. Galtung (1972) characterized this same relationship in terms of "core and periphery," each with something to offer the other, thereby fostering a symbiotic but lopsided relationship. The effect is a sharp intersocietal precedence wherein the dominant core grows and becomes more complex, while the satellite is subordinated through the transfer of economic surpluses to the core in spite of any absolute economic changes that may occur (Hechter 1975).

Dos Santos (1971) extended Frank's attention to metropolis-satellite relationships. He maintained that whatever economic or social change that does take place occurs primarily for the benefit of the dominant core. This is not to say that the outlying areas are merely plundered or picked clean; to ensure their long-run usefulness, peripheral regions are allowed, indeed encouraged, to develop. Urbanization, industrialization, commercialization of agriculture, and more expedient social, legal, and political structures are induced. In this way the core not only guarantees a stable supply of raw materials, but a market for finished goods. With the bulk of economic surplus exported to the core, the less-developed region is powerless to disseminate change or innovation across multiple realms. Dos Santos distinguished colonial, financial-industrial, and technological-industrial forms of dependencies. Under colonialism there is outright control and expropriation of valued resources by absentee decision makers. The financial-industrial dimension is marked by a locally productive economic sector characterized by widespread specialization and focus ministering primarily to the export sector that coexists alongside an essential subsistence sector. The latter furnishes labor and resources but accrues little from economic gains made in the export sector. In effect, two separate economies exist side by side.

In the third form of dependency, technological-industrial change takes place in developing regions but is customarily channeled and mandated by external interests. As the core extends its catchment area, it maximizes its ascendancy by promoting a dispersed, regionalized division of labor to maximize its returns. International market considerations affect the types of activities local export sectors are permitted to engage in by restricting the infusion of capital for specified purposes only. So, for example, the World Bank makes loans for certain forms of productive activity while eschewing others, and, as a result, highly segmented labor markets occur as internal inequities proliferate. The international division of labor is reflected in local implementation of capital-intensive technology, improvements in the infrastructure—transportation, public facilities, communications—and, ultimately, even social programming occurring principally in central enclaves or along supply corridors that service export trade (Hoogvelt 1977; Jaffee 1985; So 1990). Dos Santos

maintained that the balance of payments is manipulated, *ex parte*, not only to bring about desired forms of change, but to ensure the outflow of capital accumulation to such a degree that decapitalization of the periphery is inescapable. Or, at the very least, that capitalist-based economic development leaves local economies entirely dependent on the vagaries of markets well beyond their control. International interests establish the price of local export products and also set purchase prices of those technological-industrial products essential to maintaining the infrastructure of development. There may be a partial diffusion of technology, since growth in the periphery also enhances profits for the core, but it also creates unequal pockets of surplus labor in secondary labor markets, thereby impeding growth of internal markets as well as deteriorating the quality of life for the general populace (Portes 1976; Jaffee 1985).

Because international capital is able to stipulate the terms of the exchange, external forces shape local political processes if only through disincentives for capitalization of activities not contributing to export trade (Cardoso and Faletto 1979). As Amin (1976) pointed out, nearly all development efforts are geared to enhancing productivity and value in the export sector, even as relative disadvantages accrue in other sectors and result in domestic policies aimed at ensuring stability in the export sector above all else. Internal inequalities are thereby exacerbated as those facets of the economy in contact with international concerns become more capital intensive and increasingly affluent while other sectors languish as they shoulder the transaction costs for the entire process. One consequence of the social relations of the new production arrangements standing side by side with traditional forms is a highly visible appearance of obsolescence as status is conferred by and derived from a "narrow primary production structure" (Hoogvelt 1977, p.96). DeJanvry (1981) spoke of the social disarticulation that results as legitimization and primacy are granted to those deemed necessary to maintain externally valued economic activities. Although overall economic growth may occur as measured by the value of exports relative to Gross Domestic Product (GDP), debt loads remain high and internal disparities bleak as few opportunities for redistribution occur even if local decision makers were so inclined in the face of crushing debt-service. As Ake

(1988) put it, even per capita income figures may be insufficient as they represent averages while GDP may be suspect in light of differences between economic growth and distributive development. Finally, although local economies may manifest substantial growth, profits are exported along with products so that capital accumulation is not at the point of production in the periphery but at the core. In its various guises, and despite substantial criticisms, dependency theory provides global-orienting principles for analyses of international capitalism, interlocking monetary policies, and their role in domestic policy in developing countries.

Baran's (1957) analysis of the relationship between India and Great Britain is frequently cited as an early effort to examine the aftermath of colonialism. The imperialism of Great Britain was said to exploit India, fostering a one-sided extraction of raw materials and the imposition of impediments to industrialization, except insofar as they were beneficial to Great Britain. Precolonial India was thought to be a locus of other-worldly philosophies and self-sustaining subsistence production. Postcolonial India came to be little more than a production satellite in which local elites relished their relative advantages as facilitators of British capitalism. In Baran's view, Indian politics, education, finance, and other institutional arrangements were restructured to secure maximal gain to British enterprise. With independence, sweeping changes were undertaken, including many exclusionary practices, as countermeasures to the yoke of colonial rule.

Latin American concerns gave rise to the dependency model, and many of the *dependistas*, as they were originally known, have concentrated on regional case studies to outline the nature of contact with international capitalism. For the most part they focused less on colonialism *per se* and more on Dos Santos's (1971) second and third types of dependency. Despite substantial resources, countries in the region found themselves burdened with inordinate trade deficits and international debt loads which had the potential to inundate them (Sweezy and Magdoff 1984). As a consequence, one debt restructuring followed another to ensure the preservation of gains already made, to maintain some modicum of political stability, and to ensure that interest payments continued or, in worst-case scenarios, bad debts

could be written down and tax obligations reduced. In many instances the World Bank and the International Monetary Fund (IMF) exercised control and supervision, one consequence of this action being promotion of political regimes unlikely to challenge the principle of the loans (So 1990). Chile proved an exception, but one with disastrous and disruptive consequences. In all cases, to default would be to undermine those emoluments and privileges accorded local elites likely to seek further rather than fewer contacts with external capital.

Dos Santos's third form of dependency has also been widely explored. Landsberg (1979) looked to Asia to find empirical support. Through an analysis of manufacturing relationships in Hong Kong, Singapore, Taiwan, and South Korea with the industrial West, he concluded that despite improvement in local circumstances, relative conditions remained little affected due to the domination of manufacturing and industrial production by multinational corporations. These corporations moved production "off-shore," to Asia or other less-developed regions, in order to limit capitalization and labor costs while selling "on-shore," thereby maximizing profits. Landsberg asserted that because external capital shapes the industrialization of developing nations, the latter becomes so specialized as to have little recourse when international monopolistic practices become unbearable.

Few more eloquent defenders of the broad dependency perspective have emerged than Brazil's Cardoso (1973, 1977). He labeled his rendering a *historical-structural model* to connote the manner in which local traditions, preexisting social patterns, and the time frame of contact serve to color the way in which generalized patterns of dependency play out. He also coined the phrase *associated-dependent development* to describe the nurturing of circumscribed internal prosperity in order to enhance profit realization on investments (Cardoso 1973). He recognized that outright exploitation may generate immediate profit but can only lead to stagnation over the long run. Indeed, the fact that many former colonies remain economic losers seems to suggest that global development has not been ubiquitous (Bertocchi and Conova 1996). Instead, foreign capital functions as a means to development, underwriting dynamic progress in those sectors likely to further exports

but able, as well, to absorb incoming consumer goods. In the process, internal inequalities are heightened in the face of wide-ranging economic dualities as the physical quality of life for those segments of the population not immediately necessary to export and production suffer as a consequence.

In his analysis of Brazil, Cardoso also broadened the discussion to political consequences of dependency spurred by capital penetration. His intent was not to imply that only a finite range of consequences may occur, but to suggest that local patterns of interaction, entitlements, domination, conflict, and so on have a reciprocal impact on the conditions of dependency. By looking at changes occurring under military rule in Brazil, Cardoso succeeded in demonstrating that foreign capital predominated in essential manufacturing and commercial arenas (foreign ownership of industries in Brazil's state of Rio Grande do Sul were so extensive that Brazilian ownership was notable for being an exception). At the same time, internal disparities were amplified as interests supportive of foreign capital gained advantage at the expense of any opposition. In the process, wages and other labor-related expenses tended not to keep pace with an expanding economy, thereby resulting in ever-larger profit margins. As the military and the bourgeoisie served at the behest of multinational corporations, they defined the interests of Brazil to be consonant with their own.

By the early 1980s the Brazilian economy had stagnated, and as the country entered the new millennium it appeared headed for recursive hard times. Evans (1983) examined how what he termed the "triple alliance"—the state, private, and international capital interests—combined to alter the Brazilian economic picture pretty drastically while managing to preserve their own interests. In an effort to continue an uninterrupted export of profits, international capitalists permitted some accumulation among a carefully circumscribed local elite, so that each shared in the largess of favorable political decisions and state-sponsored ventures. Still, incongruities abounded; per capita wages fell as GDP increased and consumer goods flourished as necessities became unattainable. At the same time, infant and female mortality remained high and few overall gains in life expectancy were experienced. An exacerbation of local

inequalities may have contributed to the downswing but so too did international financial shifts which eventually dictated that the majority of new loans were earmarked for servicing old debts.

In the face of these shifts, foreign capital gained concessions, subsidies, and favored-nation accommodations. The contradictions proliferated. Unable to follow through on promises to local capital, the Brazilian government had little choice but to rescind previous agreements, at the same time incurring the unintended consequence of reducing regional market-absorption capabilities. Without new orders and in the face of loan payments, local capital grew disillusioned and moderated its support for state initiatives despite a coercive bureaucracy that sent Brazil to the verge of economic failure. The value of Evans's work is that it highlights the entanglements imposed on local politics, policies, and capitalists by a dependent-development agenda lacking significant national autonomy. What became apparent was that governments legitimate themselves more in terms of multinational interests than in terms of local capital—and certainly more than in terms of local less-privileged groups seeking to influence government expenditures. In an examination of forty-five less-developed countries, Semyonov and Lewin-Epstein (1996) discovered that though external influences shape the growth of productive services, internal processes frequently remain capable of moderating the effect these changes have on other sectors.

In an analysis of Peru, Becker (1983) contended that internal alignments created close allegiances based on mutual interests and that hegemonic control of alliances in local decision making leads to the devaluing and disenfranchising of those who challenge business as usual or represent old arrangements. Bornschieer (1981) asserted that internal inequalities increase and the rate of economic growth decreases in inverse proportion to the degree of dependency and in light of narrow sectoral targeting of foreign capital's development dollars. A recurrent theme running through their findings and those of other researchers is that internal economic disparities grow unchecked as tertiary-sector employment eventually becomes the predominant form (Bornschieer 1981; Semyonov and Lewin-Epstein 1986; Delacroix and Ragin 1978; Chase-Dunn 1981; Boyce 1992).

Nearly all advocates of dependency models contend that many facets of less-developed countries, from structure of the labor force to mortality, public health, forms and types of services provided, and the role of the state in public welfare programs, are products of the penetration of external capital and the particulars of activities in the export sector. As capital-intensive production expands, surplus labor is relegated back to agrarian pursuits or to other tertiary and informal labor. It also fosters a personal-services industry in which marginal employees provide service to local elites but whose own well-being is dependent on the economic well-being of the elites. Distributional distortions, as embodied in state-sponsored social policies, are also thought to reflect the presence of external capitalism (Kohli et al. 1984; Clark and Phillipson 1991). Evans is unmistakable: the relationship of dependency and internal inequality “. . . is one of the most robust, quantitative, aggregate findings available” (1979, p.532).

Not everyone is convinced. As investigations of dependency theory proliferated, many investigators failed to find significant effects that could be predicted by the model (Dolan and Tomlin 1980). In fact, Gereffi's (1979) review of quantitative studies of “third world” development led him to maintain that there was little to support the belief that investment of foreign capital had any discernable effect on long-term economic gain. In fact, it is commonly claimed that most investigations rely on gross measures of the value of exports relative to GDP, thus treating all exports as contributing equally to economic growth (Talbot 1998). But when a surplus of primary commodities, “raw” extraction or agricultural products, is exported, prices become unstable, with the consequences being felt most explicitly in producing regions. When manufactured goods are exported, prices remain more stable, and local economies are less affected. Relying on covariant analysis of vertical trade (export of raw materials, import of manufactured goods), commodity concentration, and export processing, Jaffee (1985) maintained that when exports grow so too does overall economic viability. Yet he did note that consideration of economic vulnerabilities and export enclaves does yield “conditional effects” whereby economic growth is significantly reduced or even takes a negative turn.

In their research on what is sometimes termed the “resource curse” in resource-rich countries,

Auty (1993) and Khalil and Mansour (1993) suggest that competition between export sectors such as minerals, oil, and agriculture often impedes prosperity of the other two or one another. In countries where that occurs, governments tend to adopt lax economic policies that, for example, increase agricultural dependencies in favor of oil exports. Internal conflicts then become self-perpetuating and play out in the physical and economic well-being of the populace. Some investigators also point out that the internal dynamics of different types of export commodities will have differential impacts on the economy as a whole and on state bureaucracies as differing degrees of infrastructure are implicated (Talbot 1998).

A number of critics have asserted that dependency theory is flawed, fuzzy, and unable to withstand empirical scrutiny (Peckham 1992; Ake 1988; Becker 1983). Even Marxist theorists find fault with dependency theory for overemphasizing external, exploitative factors at the expense of attention to the role of local elite (Shannon 1996). Other critics have focused on the evidence mustered and suggested that only about one-third of the variance in inequality among nations is accounted for by penetration of multinational corporations or other forms of foreign investment (Kohli et al. 1984; Bornschier, Chase-Dunn, and Rubinson 1978). Interestingly, still others have concluded that economic development of the type being discussed here is a significant facilitator for political democracy (Bollen 1983).

Defenders react by challenging measures of operationalization, the way variables are defined, and whether the complex of concepts embraced by the multidimensionality of the notion of dependency can be assessed in customary ways or in the absence of a comparative framework juxtaposing developing nations with their industrial counterparts (Ragin 1983; Robinson and Holtzman 1982; Boyce 1993). Efforts to isolate commodity concentration and multinational corporate investments have not proven to be reliable indicators and, as noted, even per capita GDP has its detractors. While important questions on heterogeneity, dispersion, or heteroskedasticity can be addressed by slope differences and recognized estimation techniques (Delacroix and Ragin 1978), proponents of the model are adamant that contextualized historical analysis is not only appropriate but mandated by the logic of dependency itself (Bach

1977; Bertocchi and Canova 1996). In their investigation of former colonies in Africa, Bertocchi and Canova (1996) concluded that colonial status is central to explaining relatively poor economic performance in ensuing years.

Proponents have also turned to sophisticated statistical procedures to elucidate their claims. For example, Bertocchi and Canova (1996) ran regression models on forty-six former colonies and dependencies in Africa to test their contention that colonialization makes a difference in subsequent societal and economic well-being. In a separate examination of state size and debt size among African nations, Bradshaw and Tshandu (1990) concluded that international capital penetration in the face of a mounting debt crisis may precipitate antagonism and austerity measures as the IMF and foreign capital debt claims are pitted against local claimants such as governmental subsidies and wages. In discussing capital penetration and the debt crises facing many developing nations, Bradshaw and Huang (1991) attributed incidences of political turmoil to austerity measures imposed on domestic welfare programs by IMF conditions and transnational financial institutions. They went on to assert that dependency theorists must take into consideration international recessions and global monetary crises if they are to understand structural accommodations and shifts in the quality of life in developing nations. In light of interlocking monetary policies, a single country teetering on the brink of economic adversity portends consequences for not only its trading partners but many other countries as well.

Several dependency analysts have utilized advanced analytic techniques to examine whether income inequality within countries is related to status in the world economy (Rubinson 1976; London and Robinson 1989; Boyce 1993). Both Rubinson (1976) and London and Robinson (1989) looked at interlocking world economies and their affect on governmental bureaucracies and internal structural differentiation. London and Robinson (1989) noted that the extent of multinational corporate penetration, and indirectly, its affect on income inequality, is associated with political malaise. Walton and Ragin (1990) concurred, maintaining that the involvement of international economic interests in domestic political-economic policy combine with "overurbanization" and associated dependency to help pave the way to political protest.



Boswell and Dixon (1990) carried the analysis a step further. Using regression and path analysis, they examined both economic and military dependency, concluding that both forms contribute to political instability through their effects on domestic class and state structures. They asserted that corporate penetration impedes real growth while exacerbating inequalities and the type of class polarization that leads to political violence. In his analysis of the economic shambles created in the Philippines under Ferdinand Marcos, Boyce (1993) concluded that the development strategies adopted during the Marcos era were economically disastrous for the bulk of the populace as “imperfections” in world markets precluded even modest capital accumulation for all but a privileged few. Though not speaking strictly in terms of dependency theory, Milner and Keohane (1996) point out that domestic policies are inevitably affected by global economic currents insofar as new policy preferences and coalitions are created as a result, by triggering domestic political and economic crises or by the undermining of governmental autonomy and thereby control over local economic policy. Dependency theorists customarily incorporate attention to all three in addressing the role international markets play in local economic policy.

Alternative interpretations of underdevelopment began to gain strength in the early 1970s. The next step was a world systems perspective, which saw global unity accompanied by an international division of labor with corresponding political alignments. Wallerstein (1974), Chirot and Hall (1982), and others shifted the focus from spatial definitions of nation-states as the unit of analysis to corporate actors as the most significant players able to shape activities—including the export of capital—according to their own interests. Wallerstein suggested that the most powerful countries of the world constitute a *de facto* collective core that disperses productive activities so that dependent industrialization is an extension of what had previously been geographically localized divisions of labor. World systems analysts see multinational corporations rather than nation-states as the means by which articulation of global economic arrangements is maintained. So powerful have multinationals become that even the costs of corporate organization are borne by those countries in which the corporations do business, with costs calculated according to terms dictated by the multinational

corporations themselves. Yet state participation is undoubtedly necessary as a kind of subsidization of multinational corporate interests and as a means for providing local management that, in addition to facilitating political compliance and other functions, promotes capital concentration for more efficient marketing and the maintenance of demand for existing goods and services. Thus production, consumption, and political ideologies are transplanted globally, legitimated, and yield a thoroughgoing stratification that, while it cuts across national boundaries, always creates precedence at the local level.

Dependency theory and collateral notions have become dispositional concepts utilized by numerous investigations of the effect of dependent development on diverse dimensions of inequality. Using a liberal interpretation of the model, many investigators have sought to understand how political economy affects values, types of rationality, definitions of efficiency and so on, as well as evaluations of those who do not share those values, views, or competencies. The social organization of the marketplace is thus thought to exert suzerainty over many types of social relationships and will continue to drive analysis of internal disparities in underdeveloped regions (Zeitlin 1972; Hechter 1975; Ward 1990; Shen and Williamson 1997). It remains to be seen how the absence of Soviet influence, often underemphasized during the formative period of dependency theory, will play out in analyses of international economic development in the twenty-first century (Pai 1991).

Substituting a figurative, symbolic relationship for spatial criteria, a generalized dependency model alloyed ideas of internal colonialism and has been widely employed as an explanatory framework wherein social and psychological distance from the center of power is seen as a factor in shaping well-being and other aspects of quality of life such as school enrollment, labor force participation, social insurance programs, longevity and so on. Likely as not, the political economics of development will continue to inform analysis of ancillary spheres for some time to come.

Gamson's (1968) concept of “stable unrepresentation” helped emphasize how the politics of inequality are perpetuated by real or emblematic core complexes. Internal colonialism and

political economic variants have been widely adopted in examinations of many types of social problems. Blauner's (1970) analysis of American racial problems is illustrative of one such application. So too is Marshall's (1985) investigation of patterns of industrialization, investment debt, and export dependency on the status of women in sixty less-developed countries. While she was unable to draw firm conclusions relative to dependency per se, Marshall did assert that with thoughtful specification, gender patterns in employment and education may be found to be associated with dependency-based economic change. In a manner similar to Blauner, Townsend (1981) spoke of "the structured dependency of the elderly" as a consequence of economic utility in advanced industrial societies. Many analysts have advocated the use of dependency-driven approaches to examine various consequences of dependency and development such as fertility, mortality, differential life expectancy, health patterns, and education (Hendricks 1982, 1995; Neysmith and Edwardth 1984; London 1988; Ward 1990; Lena and London 1993; Shen and Williamson 1997). Such a perspective casts the situation of the elderly as a consequence of shifts in economic relationships and state policies designed to provide for their needs. For example, Neysmith (1991) maintained that as debts are refinanced to retain foreign capital, domestic policies are rewritten in such a way as to disenfranchise vulnerable populations within those countries in favor of debt service. To support her point, Neysmith cites a United Nations finding that human development programs tend to benefit males, households in urban areas, and middle- or higher-income people, while relatively fewer are targeted at women, rural residents, or low-income persons (United Nations Department of International Economic and Social Affairs 1988). Interestingly, according to a United Nations report issued two decades ago, women account for approximately half the world's adult population, one-third of the formal labor force, and two-thirds of all the recorded work hours, yet receive one-tenth of the earned income (cited in Tiano 1988; Ward 1990). In an examination of capital penetration in eighty-six countries, Shen and Williamson (1997) suggest that as penetration increases and sectorial inequalities are exacerbated between tertiary and informal labor markets vis-à-vis other sectors, there is a degradation of women's status in all economic activities. At the same time, fertility rates remain

high partly because child labor provides an integral component of household income.

Variation in the life experiences of subpopulations is one of the enduring themes of sociology. Despite wide disparities, a central focus has been the interconnections of societal arrangements and political, economic, and individual circumstances. It is through them that norms of reciprocity and distributive justice are fostered and shared. As contexts change, so too will norms of what is appropriate. The linkage between political and moral economies is nowhere more apparent than in dependency theory as it facilitates our understanding of the dynamic relationship between individuals and structural arrangements.

## REFERENCES

- Addo, Herb 1996 "Developmentalism: A Eurocentric Hoax, Delusion, and Chicanery." In Sing C. Chew and Robert A. Denmark, eds., *The Underdevelopment of Development*. Thousand Oaks, Calif.: Sage.
- Ake, Claude 1988 "The Political Economy of Development: Does it Have a Future?" *International Social Science Journal* 40:485-497.
- Amin, Samir 1976 *Unequal Development: An Essay on the Social Formation of Peripheral Capitalism*. New York: Monthly Review Press.
- Auty, Richard M. 1993 "Sustainable Development in Mineral Exporting Economies." *Resources-Policy* 19:14-29 (cited in Talbot).
- Bach, Robert 1977 "Methods of Analysis in the Study of the World Economy: A Comment on Rubinson." *American Sociological Review* 42:811-814.
- Baran, Paul 1957 *The Political Economy of Growth*. New York: Monthly Review Press.
- Becker, David G. 1983 *The New Bourgeoisie and the Limits of Dependency: Mining, Class, and Power in "Revolutionary" Peru*. Princeton: Princeton University Press.
- Bertocchi, Graziella, and Fabio Canova 1996 "Did Colonization Matter for Growth? An Empirical Exploration into the Historical Causes of Africa's Underdevelopment." Discussion Paper No. 1444, Center for Economic Policy Research, London.
- Blauner, Robert 1970 "Internal Colonialism and Ghetto Revolt." In J. H. Skolnick and E. Curries, eds., *Crisis in American Institutions*. Boston: Little, Brown.
- Bollen, Kenneth 1983 "World System Position, Dependency, and Democracy: The Cross-National Evidence." *American Sociological Review* 48:468-479.

- Bornschier, Volker 1981 "Dependent Industrialization in the World Economy." *Journal of Conflict Resolution* 25:371-400.
- , C. Chase-Dunn, and R. Rubinson 1978 "Cross-National Evidence of the Effects of Foreign Investment: A Study of Findings and Analysis." *American Journal of Sociology* 84:651-683.
- Boswell, Terry, and William J. Dixon 1990 "Dependency and Rebellion: A Cross-National Analysis." *American Sociological Review* 55:540-559.
- Boyce, James K. 1993 *The Philippines: The Political Economy of Growth and Impoverishment in the Marcos Era*. Honolulu: University Of Hawaii Press.
- Bradshaw, York W., and Zwelakhe Tshandu 1990 "Foreign Capital Penetration, State Intervention, and Development in Sub-Saharan Africa." *International Studies Quarterly* 34:229-251.
- , and Jie Huang 1991 "Intensifying Global Dependency: Foreign Debt, Structural Adjustment, and Third World Underdevelopment." *Sociological Quarterly* 32:321-342.
- Cardoso, Fernando H. 1973 "Associated-Dependent Development: Theoretical and Practical Implications." In Alfred Stephen, ed., *Authoritarian Brazil*. New Haven: Yale University Press.
- 1977 "The Consumption of Dependency Theory in the United States." *Latin American Research Review* 12:7-24.
- , and Enzo Faletto 1979 *Dependency and Development In Latin America*. Berkeley, Calif.: University of California Press.
- Chase-Dunn, Christopher 1981 "Instate System and Capitalist World-Economy: One Logic or Two?" *International Studies Quarterly* 25:19-42.
- Chirot, Daniel, and Thomas D. Hall 1982 "World System Theory." *Annual Review of Sociology* 8:81-106.
- Clark, Roger, and Rachel Filinson 1991 "Multinational Corporate Penetration, Industrialism, Region, and Social Security Expenditures: A Cross-National Analysis." *International Journal of Aging and Human Development* 32:143-159.
- DeJanvry, Alain 1981 *The Agrarian Question and Reformism in Latin America*. Baltimore, Md.: Johns Hopkins University Press.
- Delacroix, Jacques, and Charles Ragin 1978 "Modernizing Institutions, Mobilization, and Third World Development." *American Journal of Sociology* 84:123-149.
- Dolan, M., and B. W. Tomlin 1980 "First World-Third World Linkages: The Effects of External Relations upon Economic Growth, Imbalance, and Inequality in Developing Countries." *International Organization* 34:41-63.
- Dos Santos, Theotonio 1971 "The Structure of Dependence." In K. T. Kan and Donald C. Hodges, eds., *Readings in U.S. Imperialism*. Boston: Expanding Horizons.
- Evans, Peter 1979 *Dependent Development: The Alliance of Multinational State and Local Capital in Brazil*. Princeton: Princeton University Press.
- Frank, Andre Gunder 1967 *Capitalism and Underdevelopment in Latin America*. New York: Monthly Review Press.
- 1969 *Latin America: Underdevelopment or Revolution*. New York: Monthly Review Press.
- Galtung, Johan 1972 "A Structural Theory of Imperialism." *African Review* 1:93-138.
- Gamson, William A. 1968 "Stable Unrepresentation in American Society." *American Behavioral Scientist* 12:15-21.
- Gereffi, Gary 1979 "A Critical Evaluation of Quantitative, Cross-National Studies of Dependency." Paper presented to International Studies Association Annual Meeting. Toronto, March 1979.
- Hechter, Michael 1975 *Internal Colonialism: The Celtic Fringe in British National Development, 1536-1966*. Berkeley: University of California Press.
- Hendricks, Jon 1982 "The Elderly in Society: Beyond Modernization." *Social Science History* 6:321-345.
- 1995 "The Contributions of Older Women to Social and Economic Development." *Ageing International* (July): 4-9.
- Hoogvelt, A. M. M. 1977 *The Sociology of Developing Societies*. Atlantic Heights, N.J.: Humanities Press.
- Jaffee, David 1985 "Export Dependence and Economic Growth: A Reformulation and Respecification." *Social Forces* 64:102-118.
- Keith, Nelson W. 1997 *Reframing International Development: Globalism, Postmodernity and Difference*. Thousand Oaks, Calif.: Sage.
- Khalil, Mohammad, and Ali H. Mansour 1993 "Development Patterns of Oil Exporting Countries." *Journal of Economics and Finance* 17:43-49.
- Kohli, Atul, M. F. Altfeld, S. Lotfian, and R. Mardon 1984 "Inequality in The Third World." *Comparative Political Studies* 17:283-318.
- Landsberg, Martin 1979 "Export-Led Industrialization in the Third World: Manufacturing Imperialism." *Review of Radical Political Economics* 11:50-63.
- Lena, Hugh F., and Bruce London 1993 "The Political and Economic Determinants of Health Outcomes: A

- Cross-National Analysis." *International Journal of Health Services* 23:585-602.
- London, Bruce 1988 "Dependence, Distorted Development, and Fertility Trends in Noncore Nations: A Structural Analysis of Cross-National Data." *American Sociological Review* 53:606-618.
- , and Thomas D. Robinson 1989 "The Effects of International Dependence on Income Inequality and Political Violence." *American Sociological Review* 54:305-308.
- Marshall, Susan E. 1985 "Development, Dependence, and Gender Inequality in the Third World." *International Studies Quarterly* 29:217-240.
- Milner, Helen V., and Robert O. Keohane 1996 "Internationalization and Domestic Policies: A Conclusion." In Robert O. Keohane and Helen V. Milner, eds., *Internationalization and Domestic Politics*. Cambridge, Eng.: Cambridge University Press.
- Neysmith, Sheila M. 1991 "Dependency Among Third World Elderly: A Need for New Direction in the Nineties." In M. Minkler and C. L. Estes, eds., *Critical Perspectives on Aging: The Political and Moral Economy of Growing Old*. Amityville, N.Y.: Baywood Publishing.
- , and Joey Edwardth 1984 "Economic Dependence in the 1980s: Its Impact on Third World Elderly." *Ageing and Society* 5:21-44.
- O'Donnell, Guillermo 1988 *Bureaucratic Authoritarianism: Argentina, 1966-1973, in Comparative Perspective*. Berkeley: University of California Press.
- Packenham, Robert A. *The Dependency Movement: Scholarship and Politics in Development Studies*. Cambridge, Mass.: Harvard University Press.
- Pai, Sudha 1991 "Problems in the Comparative Analysis of Third World Societies." *Indian Journal of Social Science* 4:591-603.
- Portes, Alejandro 1976 "On the Sociology of National Development: Theories and Issues." *American Journal of Sociology* 29:339-357.
- Ragin, Charles 1983 "Theory and Method in the Study of Dependency and International Inequality." *International Journal of Comparative Sociology* 24:121-136.
- Rubinson, Richard 1976 "The World-Economy and the Distribution of Income within States: A Cross-National Study." *American Sociological Review* 41:638-659.
- , and Deborah Holtzman 1982 "Comparative Dependence and Economic Development." *International Journal of Comparative Sociology* 22:86-101.
- Semyonov, Moshe, and Noah Lewin-Epstein 1986 "Economic Development, Investment Dependence, and the Rise of Services in Less Developed Nations." *Social Forces* 64:582-598.
- Shannon, Thomas R. 1996 *An Introduction to the World-System Perspective*. Boulder, Colo.: Westview.
- Shen, Ce, and John B. Williamson 1997 "Child Mortality, Women's Status, Economic Dependency, and State Strength: A Cross-National Study of Less Developed Countries." *Social Forces* 76:667-687.
- So, Alvin Y. 1990 *Social Change and Development: Modernization, Dependency, and World-System Theories*. Newbury Park, Calif.: Sage.
- Sweezy, Paul, and Harry Magdoff 1984 "The Two Faces of Third World Debt: A Fragile Financial Environment and Debt Enslavement." *Monthly Review* 35:1-10.
- Talbot, John M. 1998 "Export Sectors and Economic Development: Do Sectors Shape the Development Prospects of States?" Paper presented to the annual meeting of the American Sociological Association. San Francisco, August 21-25, 1998.
- Tiano, Susan 1988 "Women's Work in the Public and Private Spheres: A Critique and Reformulation." In M. Francis Abraham and P. Subhadra Abraham, eds., *Women, Development and Change: The Third World Experience*. Bristol, Ind.: Wyndham Hall Press.
- Townsend, Peter 1981 "The Structured Dependency of the Elderly." *Ageing and Society* 1:5-28.
- United Nations Department of International Economic and Social Affairs 1988 *Human Resources Development: A Neglected Dimension of Development Strategy*. New York: United Nations.
- Wallerstein, Immanuel 1974 "The Rise and Future Demise of the World Capitalist System: Concepts for Comparative Analysis." *Comparative Studies in Society and History* 16:387-415.
- Walton, John, and Charles Ragin 1990 "Global and National Sources of Political Protest: Third World Responses to the Debt Crisis." *American Sociological Review* 55:876-890.
- Ward, Kathryn B. (ed.) 1990 *Women Workers and Global Restructuring*. Cornell, N.Y.: ILR Press.
- Zeitlin, I. M. 1972 *Capitalism and Imperialism*. Chicago: Markham.

JON HENDRICKS

## DEPRESSION

### INTRODUCTION

The term "depression" covers a wide range of thoughts, behaviors, and feelings. It is also one of the most commonly used terms to describe a wide range of negative moods. In fact there are many

types of depression, each of which vary in the number of symptoms, their severity, and persistence. The prevalence of depression is surprisingly high. Between 5 and 12 percent of men and 10 to 20 percent of women in the United States will suffer from a major depressive episode at some time in their lives. Approximately half of these individuals will become depressed more than once, and up to 10 percent will experience manic phases where they are elated and excited, in addition to depressive ones; an illness known as “manic-depressive” or bipolar disorder. Depression can involve the body, mood, thoughts, and many aspects of life. It affects the way people eat and sleep, the way they feel about themselves, and the way they think about things. Without treatment, symptoms can last for weeks, months, or years. Appropriate treatment, however, can help most people who suffer from depression.

#### TYPES OF DEPRESSION

Depression can be experienced for either a short period of time or can extend for years. It can range from causing only minor discomfort, to completely mentally and physically crippling the individual. The former case of short-term, mild depression is what is most commonly referred to as “the blues” or when people report “feeling low.” It is technically referred to as *dysphoric mood*. Feelings of depression tend to occur in almost all individuals at some point in their lives, and dysphoric mood has been associated with many key life events varying from minor to major life transitions (e.g., graduation, pregnancy, the death of a loved one). The feelings of separation and loss associated with leaving a town one has grown up in, moving to a new city for a job or school, or even leaving a work environment that one has grown accustomed to, can cause bouts of depression signaled by a loss of motivation and energy, and sadness. Other common features of dysphoric mood, include sighing, an empty feeling in the stomach, and muscular weakness, are also associated with changes such as the breakup of a dating relationship, divorce, or separation. In the case of bereavement, most survivors experience a dysphoric mood that is usually called grief (although some studies have shown that these feelings may be distinct from depression).

Many of these feelings are seen as representing the body’s short-term response to stress. Other

kinds of stressful events like losing a job; being rejected by a lover; being unable to pay the rent or having high debts; or losing everything in a fire, earthquake, or flood; may also bring on feelings of depression. Most of these feelings based on temporary situations are perfectly normal and tend to fade away.

A more severe type of depression than dysphoric mood, *dysthymia* (from the Greek word for defective or diseased mood), involves long-term, chronic symptoms that do not disable, but keep those individuals from functioning at their best or from feeling good. People with dysthymia tend to be depressed most of the day, more days than not, based on their own description or the description of others. Dysthymics have a least two of the following symptoms: eating problems, sleeping problems, tiredness and concentration problems, low opinions of themselves, and feelings of hopelessness. Unlike major depression, dysthymics can be of any age. Often people with dysthymia also experience major depressive episodes.

If “the blues” persist, it is more indicative of major depression, also referred to as *clinical depression* or a *depressive disorder*. Major depression is manifested by a combination of symptoms that interfere with the ability to work, sleep, eat, and enjoy once-pleasurable activities. These disabling episodes of depression can occur once, twice, or several times in a lifetime. Clinical practitioners (both clinical psychologists and psychiatrists) make use of a multiple-component classification system designed to summarize the diverse information relevant to an individual case rather than to just provide a single label. Using a specified set of criteria in the *Diagnostic and Statistical Manual of Mental Disorders* (referred to as *DSM-IV*, American Psychiatric Association 1994), a diagnosis of depressive disorder includes symptoms such as dissatisfaction and anxiety; changes in appetite, sleep, and psychological and motor functions; loss of interest and energy; feelings of guilt; thoughts of death; and diminished concentration. It is important to keep in mind that many of these symptoms are also reported by individuals who are not diagnosed with clinical depression. Only having many of these symptoms at any one time qualifies as depression. An individual is said to be experiencing a “major depressive episode” if he or she experiences a depressed mood or a loss of interest

or pleasure in almost all activities and exhibits at least four other symptoms from the following list: marked weight loss or gain when not dieting, constant sleeping problems, agitated or greatly slowed-down behavior, fatigue, inability to think clearly, feelings of worthlessness, and frequent thoughts about death or suicide. Anyone experiencing these symptoms for a prolonged period of time should see a doctor or psychiatrist immediately.

Another type of depression is *bipolar disorder*, formerly called manic-depressive illness. Not nearly as prevalent as other forms of depressive disorders, bipolar disorder involves interspersed periods of depression and elation or mania. Sometimes the mood switches are dramatic and rapid, but most often they are gradual. When in the depressed cycle, individuals have any or all of the symptoms of a depressive disorder. When in the manic cycle, individuals tend to show inappropriate elation, social behavior, and irritability; have disconnected and racing thoughts; experience severe insomnia and increased sexual appetite; talk uncontrollably; have grandiose notions; and demonstrate a marked increase in energy. Mania often affects thinking, judgment, and social behavior in ways that cause serious problems and embarrassment. For example, unwise business or financial decisions may be made when an individual is in a manic phase. Bipolar disorder is often a chronic condition. For more details on types of depression, including symptoms, see either a good textbook on abnormal psychology (e.g., Sarason and Sarason 1999), DSM-IV (American Psychiatric Association 1994) or use the search term “depression” on the web site for the National Institute of Health (<http://search.info.nih.gov/>).

### THEORIES OF DEPRESSION

Most theorists agree that depression can be best studied using what health psychologists refer to as a *biopsychosocial approach*. This holds that depression has a biological component (including genetic links and biochemical imbalances), a psychological component (including how people think, feel, and behave), and a social component (including family and societal pressures and cultural factors). Individual theories have tended to emphasize one or the other of these components. The main theories of depression are biological and

cognitive in nature, although there are also psychodynamic and behavioral explanations which are discussed below.

**Psychodynamic Theories of Depression.** The psychological study of depression was essentially begun by Sigmund Freud and Karl Abraham, a German physician. Both described depression as a complex reaction to the loss of a loved person or thing. This loss could be real or imagined, through death, separation, or rejection. For Abraham (1911/1968), individuals who are vulnerable to depression experience a marked ambivalence toward people, with positive and negative feelings alternating and blocking the expression of the other. These feelings were seen to be the result of early and repeated disappointments. Depression, or melancholy, as Freud called it, was grief out of control (Freud 1917/1957). Unlike those in mourning, however, depressed persons appeared to be more self-denigrating and lacking self-esteem. Freud theorized that the anger and disappointment that had previously been directed toward the lost person or thing was internalized, leading to a loss of self-esteem and a tendency to engage in self-criticism. Theorists who used a similar approach and modified Freud’s theories for depression were Sandor Rado and Melanie Klein, and most recently John Bowlby (1988).

**Behavioral Theories of Depression.** In contrast to a focus on early-childhood experiences and internal psychological processes, behavioral theories attempt to explain depression in terms of responses to stimuli and the overgeneralization of these responses. For example, loss of interest to a wide range of activities (food, sex, etc.) in response to a specific situation (e.g., loss of a job). The basic idea is that if a behavior is followed or accompanied by something good (a reward), the behavior will increase and persist. If the reward is taken away, lessened, or worse still, if the behavior is punished, the behavior will lessen or disappear. B. F. Skinner, a key figure in the behaviorist movement, postulated that depression was the result of a weakening of behavior due to the interruption of an established sequence of behavior that had been positively reinforced by the social environment. For example, the loss of a job would stop a lot of the activities that having a salary provides (e.g., dining out often, entertainment). Most behavioral theories extended this idea, focusing on specific

others as the sources of reinforcement (e.g., spouses, friends).

**Cognitive Theories of Depression.** Although it is probably indisputable that the final common pathways to clinical depression and even dysphoric mood involve biological changes in the brain, the most influential theories of depression today focus on the thoughts of the depressed individual. This cognitive perspective also recognizes that behavior and biochemistry are important components of depression, but it is more concerned with the quality, nature, and patterns of thought processes. Cognitive therapists believe that when depressive cognitions are changed, behavior and underlying biological responses change as well. Cognitive theories of depression differ from behavioral theories in two major ways (see Gotlieb and Hammen 1992 for a more detailed description). First, whereas behavioral theories focus on observable behaviors, cognitive theories emphasize the importance of intangible factors such as attitudes, self-statements, images, memories, and beliefs. Second, cognitive approaches to depression consider maladaptive, irrational, and in some cases, distorted thoughts to be the cause of the disorder and of its exacerbation and maintenance. Depressive behaviors, negative moods, lack of motivation, and physical symptoms that are seen to accompany depression are all seen as stemming from faulty thought patterns. There are three main cognitive theories of depression: *Beck's cognitive-distortion model*, *Seligman's learned helplessness model*, and the *hopelessness theory of depression*.

*Beck's cognitive-distortion model.* The most influential of these theories is Aaron Beck's cognitive-distortion model of depression (1967). Beck believes that depression is composed of three factors: negative thoughts about oneself, the situation, and the future. A depressed person misinterprets facts in a negative way, focuses on negative aspects of a situation, and has no hope for the future. Thus any problem or misfortune experienced, like the loss of a job, is completely assumed to be one's own fault. The depressed individual blames these events on his or her own personal defects. Awareness of these presumed defects becomes so intense that it overwhelms any positive aspects of the self and even ambiguous information is interpreted as evidence of the defect in lieu of positive explanations. A depressed person might focus on

a minor negative exchange within an entire conversation and interpret this as a sign of complete rejection. These types of thought patterns, also referred to as "automatic thoughts" when responses based on insufficient information are made, are persistent and act as negative filters for all of life's experiences. Excellent descriptions of the way these thoughts operate can be found in books by the psychologist Norman Endler (*Holiday of Darkness*, 1990) and the writer William Styron (*Darkness Visible: A Memoir of Madness*, 1982).

Together with the idea that depressed individuals mentally distort reality and engage in faulty processing of information, the most important part of Beck's cognitive model of depression is the notion of a "negative self-schema." A schema is a stored body of knowledge that affects how information is collected, processed, and used, and serves the function of efficiency and speed. In the context of depression, schemas are mental processes that represent a stable characteristic of the person, influencing him or her to evaluate and select information from the environment in a negative and pessimistic direction. Similar to psychoanalytical theories, negative self-schemas are theorized to develop from negative experiences in childhood. These schemas remain with the individual throughout life, functioning as a vulnerability factor for depression. Cognitive treatments of depression necessarily work to change these negative schemas and associated negative-automatic thought patterns.

*Seligman's learned helplessness model.* Based on work on animals (later replicated in humans), Martin Seligman's (1975) theory of learned helplessness and his model of depression holds that when individuals are exposed to uncontrollable stress they fail to respond to stimulation and show marked decrements in the ability to learn new behaviors. Because this theory did not sufficiently account for the self-esteem problems faced by depressed individuals, it was reformulated by Abramson who hypothesized that together with uncontrollable stress, people must also expect that future outcomes are uncontrollable. When they believe that these negative uncontrollable outcomes are their own doing (internal versus external), will be stable across time and will apply to everything they do (global), they feel helpless and depressed.

*Hopelessness theory of depression.* The most recent reformulation of the learned helplessness theory, referred to as the “hopelessness theory” of depression (Abramson, Seligman, and Alloy 1989) holds that depression is a result of expectations that highly undesired outcomes will occur and that one is powerless to change these outcomes. The hopelessness theory of depression is receiving a large amount of attention as it has been found to be particularly useful in predicting the likelihood of suicide among depressed people.

**Biological Theories of Depression.** The most compelling of the recent theories of depression rely heavily on the biological bases of behavior. Biological theories assume the cause of depression lies in some physiological problem, either in the genes themselves or in the way neurotransmitters (the chemicals that carry signals between nerve cells in our brains and around our bodies) are produced, released, transported, or recognized (see Honig and van Praag 1997 for a detailed review of biological theories of depression). Most of the work focuses on neurotransmitters, especially a category of chemicals in our bodies called the monoamines, the main examples of which are norepinephrine (also called noradrenalin), dopamine, and serotonin. These chemicals first attracted attention in the 1950s when physicians discovered that severe depression arose in a subset of people who were treated for hypertension with a drug (reserpine) that depleted monoamines. Simultaneously, researchers found that a drug that increased the monoamines, this time given to medicate tuberculosis, elevated mood in users who were depressed. Together these results suggested that low levels of monoamines in the brain cause depression. The most important monoamine seems to be norepinephrine although it is now acknowledged that changes in levels of this neurochemical do not influence moods in everyone. Nevertheless, this biochemical theory has received much experimental support.

Apart from the neurochemicals, there are also other physiological differences between depressed and nondepressed individuals. Hormones are chemical substances that circulate in the blood and enable communication between different systems of the body. Some hormones control the release of other hormones which then stimulate growth and help prepare the body to deal with,

and respond to, stress (e.g., adrenocorticotrophic hormone or ACTH). Depressed patients have repeatedly been demonstrated to show abnormal functioning of these hormones (see Nemeroff 1998 for a detailed review). Another difference is seen in one of the major systems of the body that affects how we respond to stress; the hypothalamic-pituitary-adrenal (HPA) axis. From the late 1960s and early 1970s, researchers have found increased activity in the HPA axis in unmedicated depressed patients as evidenced by increased levels of stress markers in bodily fluids. Now a large volume of studies confirm that substantial numbers of depressed patients display overactivity of the HPA axis. According to Charles Nemeroff (1998) and his colleagues, and based on studies on animals, all these biological factors including genetic inheritance of depression, neurotransmitter and hormonal levels, and HPA axis and related activity, could relate to early childhood abuse or neglect, although this theory has yet to be fully substantiated. The antecedents and consequences notwithstanding, it is well accepted that one of the major causes of depression is based in our biology.

### RISK FACTORS

Depression can have many different causes as indicated by the different theories that have been formulated to explain it. Accordingly, there are different factors that indicate a risk for depression. Some of the main risk factors for long-term depression include heredity, age, gender, and lack of social support.

Studies of twins and of families clearly suggest a strong genetic component to clinical depression, which increases with genetic closeness. There is a much greater risk of developing a major depression if one's identical twin has had it than if one's parent, brother, or sister developed it. Chances are even less if no close relatives have ever had it. Furthermore, the younger people are when they experience depression, the higher the chances that one of their relatives will also get severely depressed. Relatives of people who were over forty when they first had a major depression have little more than the normal risk for depression.

One of the most clear risk factors is gender. Women are at least twice as likely to experience



all types of depressed states than are men and this seems to occur from an early age. There are no gender differences in depression rates in prepubescent children, but after the age of fifteen, girls and women are about twice as likely to be depressed as boys and men. Many models have been advanced for how gender differences in depression might develop in early adolescence. For example, one model suggests that the causes of depression can be assumed to be the same for girls and boys, but these causes become more prevalent in girls than in boys in early adolescence. According to another model, there are different causes of depression in girls and boys, and the causes of girls' depression become more prevalent than the causes of boys' depression in early adolescence. The model that has received the most support suggests that girls are more likely than boys to carry risk factors for depression even before early adolescence, but these risk factors lead to depression only in the face of challenges that increase in early adolescence (Nolen-Hoeksema and Girgus 1994). For a review of the epidemiology of gender differences in depression including prominent theories for why women are more vulnerable to depression, cross-cultural studies of gender differences in depression, biological explanations for the gender difference in depression (including postpartum depression, premenstrual depression, pubescent depression), personality theories (relationship with others, assertiveness), and social factors for the gender difference (increases in sexual abuse in adolescent females) see Susan Nolen-Hoeksema (1995).

Age by itself is a major risk factor for depression, although as described this varies for each gender. For women, the risk for a first episode of depression is highest between the ages of twenty and twenty-nine. For men, the risk for a first episode is highest for those aged forty to forty-nine. A related risk concerns when a person was born. People born in recent decades have been found to have an increased risk for depression as compared to those born in earlier cohorts.

Another significant risk factor for depression is the availability and perception of social support. People who lack close supportive relationships are at added risk for depression. Additionally, the presence of supportive others may prevent depression in the face of severe life stressors. Support is

especially important in the context of short-term depression that can result from events like conflictual work or personal interactions, unemployment, the loss of a job, a relationship break-up, or the loss of a loved one.

### MEASUREMENT OF DEPRESSION

Most of the commonly used techniques to assess for depression come from clinical psychology and are heavily influenced by the cognitive theories of depression. For example, the work of Beck and other cognitive theorists has led to the development of many ways to measure the thoughts that depressed individuals may have. Most of these measures are completed by the individuals themselves, while some are administered in an interview format where the therapist asks a series of questions. Some interviews are delivered by trained clinical administrators (e.g., the Structured Clinical Interview for DSM-IV), while others are highly structured, can be computer scored to achieve diagnoses based on the DSM-IV, and can be administered by lay interviewers with minimal training (e.g., the Diagnostic Interview Schedule). Separate measures have also been designed for adults and children to compensate for differences in level of comprehension and sophistication, although measures of symptoms and diagnoses in children and adolescents are less-extensively studied. The methods used work well for children provided that information from both parent and child sources are included in the final decisions.

There are different types of self-report measures for depression. It can be assessed by having the patient fill out a questionnaire. Because our thought processes may operate at varying levels of consciousness, we may not always be able to access what they are to report on them. For this reason different cognitive measures of depression were designed to operate at various levels of consciousness. For example, the most direct measures ask about the frequency with which negative automatic thoughts have "popped" into a person's head in the past week (e.g., "no one understands me"). Another type of measure attempts to get at the cognitive and social cognitive mechanisms by which people formulate their beliefs and expectations. Because many negative thoughts take the form of comparing the self with others, these types of

scales try to understand the negative comparisons. An example of this type of measure is one where individuals are asked about the circumstances (interactions with people, idle thoughts, etc.), the dimensions (social skills, intelligence), the gender, and the type of relationship with the comparison target, and the individual's mood before and after the interaction. To get at the least-accessible level of thoughts, those that are believed to store, organize, and direct the processing of personally relevant information, researchers have used measures like the Stroop color-word task. Individuals are asked to name the color of the ink in which a word is printed but to ignore the meaning of the word itself. Slower response rates are thought to indicate greater effort to suppress words that are highly descriptive of the self. For example, depressed individuals take longer to name the color in which words like "sad" and "useless" are printed compared to the color for positive words.

In health, clinical, and counseling research and evaluation settings, the two most common measures of depression are the Beck Depression Inventory (BDI) and the Center for Epidemiological Studies Depression Scale (CESD). The BDI was designed to measure "symptom-attitude categories" associated with depression (Beck 1967). These include, among others, mood, pessimism, and sense of failure as well as somatic preoccupation. Many of the items reflect Beck's belief in the relevance of negative cognitions or self-evaluations in depression. Each item includes a group of statements that reflect increasing levels of one of these symptom-attitude categories. The test taker is asked to choose the statement within each item that reflects the way he or she has been feeling in the past week. The items are scored on a scale from 0–3, and reflect increasing levels of negativity. A sample item includes 0 = "I do not feel like a failure," to 3 = "I feel I am a complete failure as a person." The CESD is a twenty-item scale, is a widely used measure of depressive symptomatology, and has been shown to be valid and reliable in many samples. Participants are asked to best describe how often they felt or behaved during the previous week, in a variety of ways reflective of symptoms of depression, using a scale ranging from 0 (*Rarely or none of the time [less than 1 day]*) to 3 (*Most or all of the time [5–7 days]*). For example, participants are asked how often their sleep was restless or they felt that everything they did was an

effort. Other self-report measures include the Minnesota Multiphasic Personality Inventory Depression Scale (MMPI-D), the Zung Self-Rating Depression Scale (SDS), and the Depression Adjective Check List (DACL). Complete descriptions of these scales can be found in Constance Hammen (1997).

## TREATMENT

As can be expected, the type of treatment depends on the type of depression and to some extent the favored theory of the health-care provider (physician, psychologist, or therapist) one goes to for treatment.

**Biologically based treatments.** The most common treatment for depression that is thought to have a physiological basis is antidepressant medication. Based on biological theories suggesting that depression results from low levels of the monoamines serotonin, norepinephrine, and dopamine, antidepressant medications act to increase the levels of these chemicals in the bloodstream. These drugs work by either preventing the monoamines from being broken down and destroyed (referred to as monoamine oxidase (MOA) inhibitors and tricyclics) or by preventing them from being removed from where they work (referred to as selective serotonin inhibitors [SSRIs]). Elavil, Norpramin, and Tofranil are the trade names of some MAO inhibitors. Prozac, Paxil, Zoloft, and Luvox are examples of SSRIs.

Although these medications have been proven to be effective in reducing depression, they also have a variety of side effects and need to be taken only under medical supervision. For example, tricyclics also cause dry mouth, constipation, dizziness, irregular heartbeat, blurred vision, ringing in the ears, retention of urine, and excessive sweating. Some of the SSRIs were developed with an eye toward reducing side effects and are correspondingly more often prescribed. Unfortunately, they are most commonly associated with prescription drug overdoses resulting in many thousands of deaths a year.

Several medicinal herbs have antidepressant effects. The most powerful is St. John's wort, a natural MAO inhibitor. In addition, ginkgo and caffeine may also help. Although much more research remains to be done, studies to date support

the effectiveness of such alternative medicine. For example, a group of researchers in Texas, in collaboration with German scientists, surveyed studies including a total of 1,757 outpatients with mainly mild or moderately severe depressive disorders, and found that extracts of St. John's wort were more effective than placebos (i.e., inactive pills) and as effective as standard antidepressant medication in the treatment of depression. They also had fewer side effects than standard antidepressant drugs (Linde et al. 1996).

In extreme cases of depression when drugs have been tried and found to not have an effect, and when the patient does not have the time to wait for drugs to take an effect (sometimes up to two or three weeks), electroconvulsive therapy (ECT) is recommended. ECT involves passing a current of between 70 and 130 volts through the patient's head after the administration of an anesthetic and muscle relaxant (to prevent injury from the convulsion caused by the charge). ECT is effective in treating severe depression although the exact mechanisms by which it works have not been determined.

**Cognitive treatments.** Cognitive therapists focus on the thoughts of the depressed person and attempt to break the cycle of negative automatic thoughts and negative self-views. Therapy sessions are well structured and begin with a discussion of an agenda for the session, where a list of items is drawn up and then discussed one by one. The therapist then tries to identify, understand, and clarify the misinterpretations and unrealistic expectations held by the client. Therapists use several techniques to identify these thoughts including asking direct questions, asking the client to use imagery to evoke the thoughts, or role-playing. Identifying these thoughts is a critical part of cognitive therapy and clients are also asked to keep daily diaries to list automatic thoughts when they occur as they are often unnoticed by depressed individuals. The client is then asked to provide a written summary of the major conclusions from the session to solidify what has been achieved and finally, the therapist prescribes a "homework assignment" designed to help the client practice skills and behaviors worked on during the session. Behavioral therapy is closely related to cognitive therapy and involves training the client to have better social skills and behaviors that enable them to develop better relationships with others.

Which are more effective treatments: cognitive or biological? A large National Institute of Mental Health study suggests that there is little difference in the effectiveness of the two therapies although the two treatments seem to produce different effects over time. Patients who received cognitive therapy were less likely to have a return of depression over time as compared to patients with biological therapy, although the small sample size used in this study precludes a definite answer to this question. Both therapies have been found to be effective, and it is likely that one is better with some forms of depression than the other, depending on how long the person has been depressed and the exact nature of his or her symptoms. In general both treatments, whether cognitive or biological, are recommended to be continued for a short time after the depressed episode has ended in order to prevent relapse.

### CONCOMITANTS

A wide body of research has documented the links between depression and a wide variety of other factors. It is both a component of many other psychological disorders as well as something that follows many other disorders. In fact, some studies have shown that out of all the people at a given time with depression, only 44 percent of them display what can be called "pure" depression, whereas the others have depression and at least one other disorder or problem. The most common of these associated problems are anxiety, substance abuse, alcoholism, and eating disorders (see Hammen 1997 for more details). Given the symptoms of depression, individuals with the disorder also experience associated social problems including strained relationships with spouses, family, and friends, and in the workplace. Most alarming perhaps is that the children of depressed parents (especially mothers) are especially at risk for developing problems of their own.

Depression has also been linked to positive factors although not always with good results. For example, there is some evidence that depression is linked to creativity. Artists tend to suffer more than their share of depression according to psychiatrists at Harvard medical school, who charted the psychological histories of fifteen mid-twentieth-century artists. They found that at least half of

them, including artists like Jackson Pollock and Mark Rothko, suffered from varying degrees of depression (Schildkraut and Aurora 1996). Many of these artists eventually committed suicide, which is perhaps one of the most significant and dangerous results of depression. At least 15 percent of people with depression complete the act of suicide, but an even higher proportion will attempt it. Consequently, individuals with severe cases of depression may experience many suicide-related thoughts and sometimes need constant surveillance.

Depression is often seen in patients with chronic or terminal illnesses and in patients who are close to dying. For example, depression is a common experience of AIDS patients, and is related to a range of factors such as physical symptomatology, number of days spent in bed, and in the perceived sufficiency of social support. Depression has also been linked to factors that influence mortality and morbidity. Higher depressed mood has been significantly associated with immune parameters pertinent to HIV activity and progression: lower levels of CD4 T cells, immune activation, and a lower proliferative response to PHA (a natural biological reaction that is essential to good health). Depression is also a critical variable with respect to compliance with treatment, especially in HIV-positive women of low-socioeconomic status.

Depression is strongly related to the number and duration of stressors experienced, or chronic burden. Chronic burden, defined by Leonard Pearlin and Carmi Schooler (1978) as ongoing difficulties in major social roles, including difficulties in employment, marriage, finances, parenting, ethnic relations, and being single/separated/divorced contributes to depression and increases vulnerability to health problems by reducing the ability of the body to respond to a physiological challenge, such as mounting an immune response to a virus. Related to chronic burden, many aspects of depression are concomitants of low-socioeconomic status, traditionally measured by education, income, and occupation. Research showing clear social-class differences in depression also suggest the contribution of the stress of poverty, exposure to crime, and other chronic stressors that vary with social class. Jay Turner, Blair Wheaton, and David Lloyd (1995) found that individuals of low-socioeconomic status were exposed to more chronic strain in the form of life difficulties in

seven domains (e.g., parenting, relationships, and financial matters) than individuals of high-socioeconomic status, which could account for higher levels of depression.

The influence of culture is one factor that has not been sufficiently studied in the context of depression. To date, most clinical-disorder classification systems do not sufficiently acknowledge the role played by cultural factors in mental disorders. The experience of depression has very different meanings and forms of expression in different societies. Most cases of depression worldwide are experienced and expressed in bodily terms of aching backs, headaches, fatigue, and a wide assortment of symptoms that lead patients to regard this condition as a physical problem (Sarason and Sarason 1999). Only in contemporary Western societies is depression seen principally as an internal psychological experience. For example, many cultures tend to view their mental health problems in terms of physical bodily problems. That is, they tend to manifest their worries, guilt feelings, and strong negative emotions (such as depression) as physical complaints. This could be because bodily complaints do not carry the stigma or negative social consequences that psychological problems do, and are correspondingly easier to talk about.

Although not an essential part of aging, many people over age sixty-five develop clinical depression. Surveys suggest that only about 5 percent of healthy elderly people living independently suffer depression at any given moment, but more than 15 percent experience depression at some point during their elderly years, and the condition tends to be more chronic than in younger people. In addition, some 25 percent of elderly individuals experience periods of persistent sadness that last two weeks or longer, and more than 20 percent report persistent thoughts of death and dying. The likelihood of depression varies with the situation the person is in, and is more likely when the elderly person is away from his or her family in a novel setting. For example, some 20 percent of nursing home residents are depressed. Depression is also antagonized by serious medical conditions that elderly men and women may have. Correspondingly, depression is commonly associated with illnesses like cancer, heart attack, and stroke. Depression often goes undiagnosed and untreated in the elderly and is something that caregivers (spouse,

children, family, and friends) should be especially watchful for given the relationship between depression and suicide.

### CONCLUSIONS

Many people still carry the misperception that depression is either a character flaw, a problem that happens because of personal weaknesses, or is completely “in the head.” As described above, there are psychological, physiological, and societal components to depression. Most importantly, it is something that can and should be treated. There are too few people who see a doctor when they recognize symptoms of depression or think of getting medical treatment for it. Depression is so prevalent that it is often seen as a natural component of life events like pregnancy and old age, and depressed mothers and elderly men and women often do not get the attention they need. Today, much more is known about the causes and treatment of this mental-health problem, with the best form of treatment being a combination of medication and psychotherapy. Depression need not be “the end.”

### REFERENCES

- Abraham, Karl 1968 “Notes on the Psychoanalytic Investigation and Treatment of Manic-Depressive Insanity and Allied Conditions” (1911). In K. Abraham, ed., *Selected Papers of Karl Abraham*. New York: Basic Books.
- Abramson, Lauren Y., G. I. Metalsky, and L. B. Alloy 1989 “Hopelessness Depression: A Theory-Based Subtype of Depression.” *Psychological Review* 96:358–372.
- American Psychiatric Association 1994 *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*, 4th ed. Washington, D.C.: American Psychiatric Association.
- Beck, Aaron T. 1967 *Depression: Clinical, Experimental and Theoretical Aspects*. New York: Harper and Row.
- Beckham, E. E., and W. R. Leber (eds.) 1995 *Handbook of Depression: An Updated Review and Integration*, 2nd ed. New York: Guilford.
- Bowlby, John 1988 *A Secure Base: Parent-Child Attachment and Healthy Human Development*. New York: Basic Books.
- Brown, G. W., and T. O. Harris 1978 *Social Origins of Depression*. London: Free Press.
- Endler, Norman S. 1990 *Holiday of Darkness: A Psychologist's Personal Journey Out of His Depression*. New York: John Wiley and Sons.
- Freud, S. 1957 “Mourning and Melancholia” (1917). In J. Strachey, ed., *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol 14. London: Hogarth.
- Gotlib, Ian H., and Constance L. Hammen 1992 *Psychological Aspects of Depression: Toward a Cognitive-Interpersonal Integration*. New York: John Wiley and Sons.
- Hammen, Constance 1997 *Depression*. Hove, East Sussex, Eng.: Psychology Press.
- Honig, A., and H. M. van Praag (eds.) 1997 *Depression: Neurobiological, Psychopathological, and Therapeutic Advances*. New York: John Wiley and Sons.
- Linde K., G. Ramirez, C. D. Mulrow, A. Pauls, W. Weidenhammer, and D. Melchart 1996 “St. John's Wort for Depression: A Meta-Analysis of Randomized Clinical Trials.” *British Medical Journal*, 313, 253.
- Nemeroff, Charles B. 1998 “The Neurobiology of Depression.” *Scientific American* 278:42–49.
- Nolen-Hoeksema, Susan 1995 “Epidemiology and Theories of Gender Differences in Unipolar Depression.” In M. V. Seeman, ed., *Gender and Psychopathology*, 63–87. Washington, D.C.: American Psychiatric Press.
- , and J. S. Girus 1994 “The Emergence of Gender Differences in Depression During Adolescence.” *Psychological Bulletin* 115:424–443.
- Pearlin, L. I., and C. Schooler 1978 “The Structure of Coping.” *Journal of Health and Social Behavior* 19:2–21.
- Sarason, Irwin G., and Barbara R. Sarason 1999 *Abnormal Psychology: The Problem of Maladaptive Behavior*, 9th ed. Upper Saddle River, N.J.: Prentice-Hall.
- Schildkraut, J. J., and A. Otero (eds.) 1996 *Depression and the Spiritual in Modern Art: Homage to Miro*, 112–130. New York: John Wiley and Sons.
- Seligman, Martin E. P. 1975 *Helplessness: On Depression, Development and Death*. San Francisco: Freeman.
- Styron, William 1990 *Darkness Visible: A Memoir of Madness*. New York: Random House.
- Turner, R. J., B. Wheaton, and D. A. Lloyd 1995 “The Epidemiology of Social Stress.” *American Sociological Review* 60:104–125.

REGAN A. R. GURUNG, PH. D.

## DESCRIPTIVE STATISTICS

Descriptive statistics include data distribution and the summary information of the data. Researchers use descriptive statistics to organize and describe the data of a sample or population. The characteristics of the sample are statistics while those of the

population are parameters. Descriptive statistics are usually used to describe the characteristics of a sample. The procedure and methods to infer the statistics to parameters are the statistical inference. Descriptive statistics do not include statistical inference.

Though descriptive statistics are usually used to examine the distribution of single variables, they may also be used to measure the relationship of two or more variables. That is, descriptive statistics may refer to either univariate or bivariate relationship. Also, the level of the measurement of a variable, that is, nominal, ordinal, interval, and ratio level, can influence the method chosen.

### DATA DISTRIBUTION

To describe a set of data effectively, one should order the data and examine the distribution. An eyeball examination of the array of small data is often sufficient. For a set of large data, the aids of tables and graphs are necessary.

**Tabulation.** The table is expressed in counts or rates. The *frequency table* can display the distribution of one variable. It lists attributes, categories, or intervals with the number of observations reported. Data expressed in the frequency distribution are *grouped data*. To examine the central tendency and dispersion of large data, using grouped data is easier than using ungrouped data. Data usually are categorized into intervals that are mutually exclusive. One case or data point falls into one category only. Displaying frequency distribution of quantitative or continuous variables by intervals is especially efficient. For example, the frequency distribution of age in an imaginary sample can be seen in Table 1.

Here, age has been categorized into five intervals, i.e., 15 and below, 16–20, 21–25, 26–30, and 31–35, and they are mutually exclusive. Any age falls into one category only. This display is very efficient for understanding the age distribution in our imaginary sample. The distribution shows that twenty cases are aged fifteen or younger, twenty-five cases are sixteen to twenty years old, thirty-six cases are twenty-one to twenty-five years old, twenty cases are twenty-six to thirty years old, and nineteen cases are thirty-one to thirty-five years old. To compare categories or intervals and to

compare various samples or populations, the reporting percent or relative frequency of each category is important. The third column shows the percent of sample in each interval or category. For example, 30 percent of the sample falls into the range of twenty-one to twenty-five years old. The fourth column shows the proportion of observation for each interval or category. The proportion was called *relative frequency*. The cumulative frequency, the cumulative percent, and the cumulative relative frequency are other common elements in frequency tabulation. They are the sum of counts, percents, or proportions below or equal to the corresponding category or interval. For instance, the cumulative frequency of age thirty shows 101 persons or 84.2 percent of the sample age thirty or younger.

The frequency distribution displays one variable at a time. To study the *joint distribution* of two or more variables, we cross-tabulate them first. For example, the joint distribution of age and sex in the imaginary sample can be expressed in Table 2.

This table is a two-dimensional table: age is the column variable and sex is the row variable. We call this table a “two-by-five” table: two categories for sex and five categories for age. The marginal frequency can be seen as the frequency distribution of the corresponding variables. For example, there are fifty seven men in this sample. The marginal frequency for age is called *column frequency* and the marginal frequency for sex is called *row frequency*. The joint frequency of age and sex is *cell frequency*. For example, there are seventeen women twenty-one to twenty-five years old in this sample. The second number in each cell is *column percentage*; that is, the cell frequency divided by the column frequency and times 100 percent. For example, 47 percent in the group of twenty-one to twenty-five year olds are women. The third number in each cell is *row percentage*; that is, the cell frequency divided by the row frequency. For example, 27 percent of women are twenty-one to twenty-five years old. The *marginal frequency* can be seen as the frequency distribution of the corresponding variables. The row and column percentages are useful in examining the distribution of one variable conditioning on the other variable.

**Charts and Graphs.** Charts and graphs are efficient ways to show data distribution. Popular

<b>Age Distribution of an Imaginary Sample</b>					
Codes	Frequency	Percent	Relative Frequency	Cumulative Frequency	Cumulative Percent
15 and below	20	16.7	.17	20	16.7
16–20	25	20.8	.21	45	37.5
21–25	36	30.0	.30	81	67.5
26–30	20	16.7	.17	101	84.2
31–35	19	15.8	.16	120	100
<b>Total</b>	<b>120</b>	<b>100</b>	<b>1.0</b>		

Table 1

graphs for single variables are *bar graphs*, *histograms*, and *stem-and-leaf plots*. The bar graph shows the relative frequency distribution of discrete variables. A bar is drawn over each category with height of the bar representing the relative frequency of observations in that category. The histogram can be seen as a bar graph for the continuous variable. By connecting the midpoints of tops of all bars, a histogram becomes the *frequency polygon*. Histograms effectively show the shape of the distribution.

Stem-and-leaf plots represent each observation by its higher digit(s) and its lowest digit. The value of higher digits is the stem while the value of the final digit of each observation is the leaf. The stem-and-leaf plot conveys the same information as the bar graph or histogram. Additionally, it tells the exact value of each observation. Despite providing more information than bar graphs and histograms, stem-and-leaf plots are used mostly for small data.

Other frequently used graphs include *line graphs*, *ogives*, and *scatter plots*. Line graphs and ogives show the relationship between time and the variable. The line graph usually shows trends. The ogive is a form of a line graph for cumulative relative frequency or percentage. It is commonly used for survival data. The scatter plot shows the relationship between variables. In a two-dimensional scatter plot,  $x$  and  $y$  axes label values of the data. Conventionally, we use the horizontal axis ( $x$ -axis) for the explanatory variable and use the vertical axis ( $y$ -axis) for the outcome variable. The plain is naturally divided into four areas by two axes. For continuous variables, the value at the joint point of two axes is zero. When the  $x$ -axis

goes to the right or  $y$ -axis goes up, the value ascends; when the  $x$ -axis goes to the left or  $y$ -axis goes down, the value descends. The data points, determined by the joint attributes of the variables, are scattered in four areas or along the axes.

### SUMMARY STATISTICS

We may use measures of central tendency and dispersion to summarize the data. To measure the central tendency of a distribution is to measure its center or typicality. To measure the dispersion of a distribution is to measure its variation, heterogeneity, or deviation.

**Central Tendency.** Three popular measures of the central tendency are *mean*, *median*, and *mode*. The arithmetic mean or average is computed by taking the sum of the values and dividing by the number of the values. It is the balanced point of the sample or population weighted by values. Mean is an appropriate measure for continuous (ratio or interval) variables. However, the information might be misleading because the arithmetic mean is sensitive to the extreme value or outliers in a distribution. For example, the ages of five students are 21, 19, 20, 18, and 20. The ages of another five students are 53, 9, 12, 13, and 11. Though their distributions are very different, the mean age for both groups is 19.6.

Median is the value or attribute of the central case in an ordered distribution. If the number of cases is even, the median is the arithmetic average of the central two cases. In an ordered age distribution of thirty-five persons, the median is the age of the eighteenth person, while, in a distribution of thirty-six persons, the median is the average age of

Sex	Age					Total
	15 and below	16–20	21–25	26–30	31–25	
Male	9	12	19	9	8	57
	45.0%	48.0%	52.8%	45.0%	42.1%	47.5%
	15.8%	21.1%	33.3%	15.8%	14.0%	
Female	11	13	17	11	11	63
	55.0%	52.0%	47.2%	55.0%	57.9%	52.5%
	17.5%	20.6%	27.0%	17.5%	17.5%	
Total	20	25	36	20	19	120
	16.7%	20.8%	30.0%	16.7%	15.8%	100%

Table 2

the seventeenth and eighteenth persons. The median, like mean, can only tell the value of the physical center in an array of numbers, but cannot tell the dispersion. For example, the median of 21, 30, 45, and 100 is 27.5 and the median of 0, 27, 28, and 29 is also 27.5, but the two distributions are different. The mode is the most common value, category, or attribute in a distribution. Like the median, the mode has its limitations. For a set of values of 0, 2, 2, 4, 4, 4, 4, 5, and 10, the mode is four. For a set of values of 0, 0, 1, 4, 4, 4, 5, and 6, the mode is also four. One cannot tell one distribution from the other simply by examining the mode or median alone. The mode and median can be used to describe the central tendency of both continuous and discrete variables, and values of mode and median are less affected by the extreme value or the outlier than the mean.

One may also use upper and lower quartiles and percentiles to measure the central tendency. The  $n$  percentile is a number such that  $n$  percent of the distribution falls below it and  $(100-n)$  percent falls above it. The lower quartile is the twenty-fifth percentile, the upper quartile is the seventy-fifth percentile, and the median is the fiftieth percentile. For example, the lower quartile or the twenty-fifth percentile is two and the upper quartile or the seventy-fifth percentile is seven for a set of values of 1, 2, 3, 4, 5, 6, 7, and 8. Apparently, the upper and lower quartiles and the percentiles can provide more information about a distribution than the other measures of the central tendency.

**Dispersion.** The central tendency per se does not provide much information on the distribution. Yet the combination of measures of central tendency and dispersion becomes useful to study a

distribution. The most popular measures of dispersion are *range*, *standard deviation*, and *variance*. Range is the crude measure of a distribution from the highest value to the lowest value or the difference between the highest and the lowest values. For example, the range for a set of values of 1, 2, 3, 4, and 5 is one to five. The range is sensitive to the extreme value and may not provide sufficient information about the distribution. Alternatively, the dispersion can be measured by the distance between the mean and each value. The standard deviation is defined as the square root of the arithmetic mean of the squared deviation from the mean. For example, the standard deviation for a set of values of 1, 2, 3, 4, and 5 is 1.44. We take the square root of the squared deviation from the mean because the sum of the deviation from the mean is always zero. The variance is the square of the standard deviation. The variance is two in the previous array of numbers. The standard deviation is used as a standardized unit in statistical inference. Comparing with standard deviation, the unit of the variance is not substantively meaningful. It is, however, valuable to explain the relationship between variables. Mathematically, the variance defines the area of the normal curve while the standard deviation defines the average distance between the mean and each data point. Since they are derived from the distance from the mean, standard deviation and variance are sensitive to the extreme values.

The *interquartile range* (IQR) and *mean absolute deviation* (MAD) are also commonly used to measure the dispersion. The IQR is defined as the difference between the first and third quartiles. It is more stable than the range. MAD is the average



absolute values of the deviation of the observations from the mean. As standard deviation, MAD can avoid the problem that the sum of the deviation from the mean is zero, but it is not as useful in statistical inference as variance and standard deviation.

**Bivariate Relationship.** One may use the *covariance* and *correlation coefficients* to measure the direction and size of a relationship between two variables. The covariance is defined as the average product of the deviation from the mean between two variables. It also reports the extent to which the variables may vary together. On average, while one variable deviates one unit from the mean, the covariance tells the extent to which the corresponding value of the other variable may deviate from its own mean. A positive covariance suggests that, while the value of one variable increases, that of the other variable tends to increase. A negative covariance suggests that, while the value of one variable increases, that of the other variable tends to decrease. The correlation coefficient is defined as the ratio of the covariance to the product of the standard deviations of two variables. It can also be seen as a covariance rescaled by the standard deviation of both variables. The value of the correlation coefficient ranges from  $-1$  to  $1$ , where zero means no correlation,  $-1$  means perfectly negatively related, and  $1$  means perfectly positively related. The covariance and correlation are measures of the bivariate relationship between continuous variables. Many measures of association between categorical variables are calculated using cell frequencies or percentages in the cross-tabulation, for example, Yule's  $Q$ ,  $\phi$ , Goodman's  $\tau$ , Goodman's  $\gamma$ , and Somer's  $d$ . Though measures of association alone show the direction and size of a bivariate relationship, it is statistical inference to test the existence of such a relationship.

#### RELATIONSHIPS BETWEEN GRAPHS AND SUMMARY STATISTICS

The *box plot* is a useful tool to summarize the statistics and distribution. The box plot is consisted of a rectangular divided box and two extended lines attached to the ends of the box. The ends of the box define the upper and lower quartiles. The range of the distribution on each side is shown by an extended line attached to each quartile. A line

dividing the box shows the median. The plot can be placed vertically or horizontally. The box plot became popular because it can express the center and spread of the data simultaneously. Several boxes may be placed next to one another for comparison.

The order of mode, median, and mean is related to the shape of the distribution of a continuous variable. If mean, median, and mode are equal to each other, the shape of the histogram approximates a bell curve. However, a uniform distribution, in which all cases are equally distributed among all values and three measures of the central tendency are equal to each other, has a square shape with the width as the range and the height as the counts or relative frequency. In a bimodal distribution, two modes are placed in two ends of the distribution equally distanced from the center where the median and the mean are placed. We seldom see the true bell-curved, uniform, and bimodal distributions. Most of the distributions are more or less skewed to the left or to the right. If the mean is greater than median and the median is greater than mode, the shape is skewed to the right. If the mean is smaller than the median, and the median is smaller than the mode, the shape is skewed to the left. The outliers mainly lead the direction.

The shape and direction of the scatter plot can diagnose the relationship of two variables. When the distribution directs from the upper-right side to the lower-left side, the correlation coefficient is positive; when it directs from the upper-left side to the lower-right side, the correlation coefficient is negative. The correlation of a loosely scattered plot is weaker than that of a tightly scattered plot. A three-dimensional scatter plot can be used to show a bivariate relationship and its frequency distribution or a relationship of three variables. The former is commonly seen as a graph to examine a joint distribution.

Descriptive statistics is the first step in studying the data distribution. In omitting this step, one might misuse the advanced methods and thus be led to wrong estimates and conclusions. Some summary statistics such as standard deviation, variance, mean, correlation, and covariance, are also essential elements in statistical inference and advanced statistical methods.

## REFERENCES

- Agresti, Alan, and Barbara Finlay 1997 *Statistical Methods for the Social Sciences*, 3rd ed. New York: Simon & Schuster.
- Blalock, Jr., Hubert M. 1979 *Social Statistics*, rev. 2nd ed. New York: McGraw-Hill.
- Johnson, Allan G. 1988 *Statistics*. New York: Harcourt Brace Jovanovich.
- Wonnacott, Thomas H., and Ronald J. Wonnacott 1990 *Introductory Statistics*. 5th ed. New York: John Wiley & Sons, Inc.

DAPHNE KUO

**DEVIANCE**

*See* Alienation; Anomie; Criminalization of Deviance; Deviance Theories; Legislation of Morality.

**DEVIANCE THEORIES**

Since its inception as a discipline, sociology has studied the causes of deviant behavior, examining why some persons conform to social rules and expectations and why others do not. Typically, sociological theories of deviance reason that aspects of individuals' social relationships and the social areas in which they live and work assist in explaining the commission of deviant acts. This emphasis on social experiences, and how they contribute to deviant behavior, contrasts with the focus on the internal states of individuals taken by disciplines such as psychology and psychiatry.

Sociological theories are important in understanding the roots of social problems such as crime, violence, and mental illness and in explaining how these problems may be remedied. By specifying the causes of deviance, the theories reveal how aspects of the social environment influence the behavior of individuals and groups. Further, the theories suggest how changes in these influences may yield changes in levels of deviant behaviors. If a theory specifies that a particular set of factors cause deviant behavior, then it also implies that eliminating or altering those factors in the environment will change levels of deviance. By developing policies or measures that are informed

by sociological theories, government agencies or programs focused on problems like crime or violence are more likely to yield meaningful reductions in criminal or violent behavior.

Despite their importance, deviance theories disagree about the precise causes of deviant acts. Some look to the structure of society and groups or geographic areas within society, explaining deviance in terms of broad social conditions in which deviance is most likely to flourish. Others explain deviant behavior using the characteristics of individuals, focusing on those characteristics that are most highly associated with learning deviant acts. Other theories view deviance as a social status conferred by one group or person on others, a status that is imposed by persons or groups in power in order to protect their positions of power. These theories explain deviance in terms of differentials in power between individuals or groups.

This chapter reviews the major sociological theories of deviance. It offers an overview of each major theory, summarizing its explanation of deviant behavior. Before reviewing the theories, however, it may prove useful to describe two different dimensions of theory that will structure our discussion. The first of these, the *level of explanation*, refers to the scope of the theory and whether it focuses on the behavior and characteristics of individuals or on the characteristics of social aggregates such as neighborhoods, cities, or other social areas. *Micro-level* theories stress the individual, typically explaining deviant acts in terms of personal characteristics of individuals or the immediate social context in which deviant acts occur. In contrast, *macro-level* theories focus on social aggregates or groups, looking to the structural characteristics of areas in explaining the origins of deviance, particularly rates of deviance among those groups.

Theories of deviance also vary in relation to a second dimension, *causal focus*. This dimension divides theories into two groups, those that explain the social origins of norm violations and those explaining societal reactions to deviance. *Social origin* theories focus on the causes of norm violations. Typically, these theories identify aspects of the social environment that trigger norm violations; social conditions in which the violations are most likely to occur. In contrast, *social reaction* theories argue that deviance is often a matter of

social construction, a status imposed by one person or group on others and a status that ultimately may influence the subsequent behavior of the designated deviant. Social reaction theories argue that some individuals and groups may be designated or labeled as deviant and that the process of labeling may trap or engulf those individuals or groups in a deviant social role.

These two dimensions offer a four-fold scheme for classifying types of deviance theories. The first, *macro-level origin theories*, focus on the causes of norm violations associated with broad structural conditions in the society. These theories typically examine the influences of such structural characteristics of populations or communities like the concentration of poverty, levels of community integration, or the density and age distribution of the population on areal rates of deviance. The theories have clear implications for public policies to reduce levels of deviance. Most often, the theories highlight the need for altering structural characteristics of society, such as levels of poverty, that foster deviant behavior.

The second, *micro-level origin theories* focus on the characteristics of the deviant and his or her immediate social environment. These theories typically examine the relationship between a person's involvement in deviance and such characteristics as the influence of peers and significant others, persons' emotional stakes in conformity, their beliefs about the propriety of deviance and conformity, and their perceptions of the threat of punishments for deviant acts. In terms of their implications for public policy, micro-level origin theories emphasize the importance of assisting individuals in resisting negative peer influences while also increasing their attachment to conforming lifestyles and activities.

A third type of theories may be termed *micro-level reaction theories*. These accord importance to those aspects of interpersonal reactions that may seriously stigmatize or label the deviant and thereby reinforce her or his deviant social status. According to these theories, reactions to deviance may have the unintended effect of increasing the likelihood of subsequent deviant behavior. Because labeling may increase levels of deviance, micro-level reaction theories argue that agencies of social control (e.g. police, courts, correctional systems) should adopt policies of "nonintervention."

Finally, *macro-level reaction theories* emphasize broad structural conditions in society that are associated with the designation of entire groups or segments of the society as deviant. These theories tend to stress the importance of structural characteristics of populations, groups, or geographic areas, such as degrees of economic inequality or concentration of political power within communities or the larger society. According to macro-level reaction theories, powerful groups impose the status of deviant as a mechanism for controlling those groups that represent the greatest political, economic, or social threat to their position of power. The theories also imply that society can only achieve reduced levels of deviance by reducing the levels of economic and political inequality in society.

The rest of this article is divided into sections corresponding to each of these four "types" of deviance theory. The article concludes with a discussion of new directions for theory—the development of explanations that cut across and integrate different theory types and the elaboration of existing theories through greater specification of the conditions under which those theories apply.

#### MACRO-LEVEL ORIGINS OF DEVIANCE

Theories of the macro-level origins of deviance look to the broad, structural characteristics of society, and groups within society, to explain deviant behavior. Typically, these theories examine one of three aspects of social structure. The *first* is the pervasiveness and consequences of poverty in modern American society. Robert Merton's (1938) writing on American social structure and Richard Cloward and Lloyd Ohlin's (1960) subsequent work on urban gangs laid the theoretical foundation for this perspective. Reasoning that pervasive materialism in American culture creates unattainable aspirations for many segments of the population, Merton (1964) and others argue that there exists an environmental state of "strain" among the poor. The limited availability of legitimate opportunities for attaining material wealth forces the poor to adapt through deviance, either by achieving wealth through illegitimate means or by rejecting materialistic aspirations and withdrawing from society altogether.

According to this reasoning, deviance is a byproduct of poverty and a mechanism through

which the poor may attain wealth, albeit illegitimately. Thus, “strain” theories of deviance interpret behaviors such as illegal drug selling, prostitution, and armed robbery as innovative adaptations to blocked opportunities for legitimate economic or occupational success. Similarly, the theories interpret violent crimes in terms of the frustrations of poverty, as acts of aggression triggered by those frustrations (Blau and Blau 1982). Much of the current research in this tradition is examining the exact mechanisms by which poverty and economic inequality influence rates of deviant behavior.

Although once considered a leading theory of deviance, strain theory has come under criticism for its narrow focus on poverty as the primary cause of deviant behavior. Recent efforts have sought to revise and extend the basic principles of the theory by expanding and reformulating ideas about strain. Robert Agnew (1992) has made the most notable revisions to the theory. His reformulation emphasizes social psychological, rather than structural, sources of strain. Agnew also broadens the concept of strain, arguing that poverty may be a source of strain, but it is not the *only* source. Three sources of strain are important: failure to achieve positively valued goals, removal of positively valued stimuli, and confrontation with negative stimuli. The first type of strain, failure to achieve positively valued goals, may be the result of a failure to live up to one’s expectations or aspirations. Strain may also result if an individual feels that he or she is not being treated in a fair or just manner. The removal of a positively valued stimulus, such as the death of a family member or the loss of a boyfriend or girlfriend, can also result in strain. Finally, strain can also be produced by the presentation of negative stimuli, such as unpleasant school experiences. Thus, although this reformulation of strain theory retains the notion that deviance is often the result of strain, the concept of strain is broadened to include multiple sources of strain.

The second set of macro-level origin theories examine the role of culture in deviant behavior. Although not ignoring structural forces such as poverty in shaping deviance, this class of theories reasons that there may exist cultures within the larger culture that endorse or reinforce deviant values; deviant subcultures that produce higher rates of deviance among those segments of the population sharing subcultural values.

Subcultural explanations have their origin in two distinct sociological traditions. The first is writing on the properties of delinquent gangs that identifies a distinct lower-class culture of gang members that encourages aggression, thrill seeking, and antisocial behavior (e.g., Miller 1958). The second is writing on cultural conflict that recognizes that within complex societies there will occur contradictions between the conduct norms of different groups. Thorsten Sellin (1938) suggests that in heterogeneous societies several different subcultures may emerge, each with its own set of conduct norms. According to Sellin, the laws and norms applied to the entire society do not necessarily reflect cultural consensus but rather the values and beliefs of the dominant social groups.

Subcultural theories emerging from these two traditions argue that deviance is the product of a cohesive set of values and norms that favors deviant behavior and is endorsed by a segment of the general population. Perhaps most prominent among the theories is Marvin Wolfgang and Franco Ferracuti’s (1967) writing on subcultures of criminal violence. Wolfgang and Ferracuti reason that there may exist a distinct set of beliefs and expectations within the society; a subculture that promotes and encourages violent interactions. According to Wolfgang and Ferracuti, this violent subculture is pervasive among blacks in the United States and may explain extremely high rates of criminal homicide among young black males.

Although Wolfgang and Ferracuti offer little material specifying the subculture’s precise causes, or empirical evidence demonstrating the pervasiveness of subcultural beliefs, other writers have extended the theory by exploring the relationship between beliefs favoring violence and such factors as the structure of poverty in the United States (Curtis 1975; Messner 1983), the history of racial oppression of blacks (Silberman 1980), and ties to the rural South and a southern subculture of violence (Gastil 1971; Erlanger 1974). Even these writers, however, offer little empirical evidence of violent subcultures within U.S. society.

A third class of theories about the macro-level origins of deviance began with the work of sociologists at the University of Chicago in the 1920s. Unlike strain and subcultural theories, these stress the importance of the social integration of neighborhoods and communities—the degree to which

neighborhoods are stable and are characterized by a homogenous set of beliefs and values—as a force influencing rates of deviant behavior. As levels of integration increase, rates of deviance decrease. Based in the early work of sociologists such as Clifford Shaw and Henry McKay, the theories point to the structure of social controls in neighborhoods, arguing that neighborhoods lacking in social controls are “disorganized,” that is, areas in which there is a virtual vacuum of social norms. It is in this normative vacuum that deviance flourishes. Therefore, these theories view deviance as a property of areas or locations rather than specific groups of people.

Early writers in the “disorganization” tradition identified industrialization and urbanization as the causes of disorganized communities and neighborhoods. Witnessing immense growth in eastern cities such as Chicago, these writers argued that industrial and urban expansion create zones of disorganization within cities. Property owners move from the residential pockets on the edge of business and industrial areas and allow buildings to deteriorate in anticipation of the expansion of business and industry. This process of natural succession and change in cities disrupts traditional mechanisms of social control in neighborhoods. As property owners leave transitional areas, more mobile and diverse groups enter. But the added mobility and diversity of these groups translate into fewer primary relationships—families and extended kinship and friendship networks. And as the number of primary relationships decline, so will informal social controls in neighborhoods. Hence, rates of deviance will rise.

Recent writing from this perspective focuses on the mechanisms by which specific places in urban areas become the spawning grounds for deviant acts (Bursik and Webb 1982; Bursik 1984; and others). For example, Rodney Stark (1987) argues that high levels of population density are associated with particularly low levels of supervision of children. With little supervision, children perform poorly in school and are much less likely to develop “stakes in conformity”—that is, emotional and psychological investments in academic achievement and other conforming behaviors. Without such stakes, children and adolescents are much more likely to turn to deviant alternatives. Thus, according to Stark, rates of deviance will be high in densely populated areas because social

controls in the form of parental supervision are either weak or entirely absent.

Similarly, Robert Crutchfield (1989) argues that the structure of work opportunities in areas may have the same effect. Areas characterized primarily by secondary-sector work opportunities—low pay, few career opportunities, and high employee turnover—may tend to attract and retain persons with few stakes in conventional behavior—a “situation of company” in which deviance is likely to flourish.

Recent writing from the disorganization perspective has also taken the form of ethnographies; qualitative studies of urban areas and the deviance producing dynamics of communities. As Sullivan (1989, p. 9) states, ethnographies describe the community “as a locus of interaction, intermediate between the individual and the larger society, where the many constraints and opportunities of the total society are narrowed to a subset within which local individuals choose.” At the heart of Sullivan’s argument is the idea that social networks in neighborhoods are important in understanding whether individuals are capable of finding meaningful opportunities for work. For example, youth were less likely to turn to crime in those neighborhoods where they could take advantage of family and neighborhood connections to blue collar jobs. Because of the greater employment opportunities in these neighborhoods, even youth who become involved in crime were less likely to persist in high-risk criminal behaviors.

Similarly, Jay MacLeod (1995) attempts to explain how the aspirations of youth living in urban areas have been “leveled,” or reduced to the point where the youths have little hope for a better future. In an analysis of two urban gangs, MacLeod argues that the youths’ family and work experiences, along with their relationships with their peers, help explain why a predominantly white gang had lower aspirations and engaged in more delinquent and antisocial behavior than the other gang, predominantly comprised of African Americans. According to MacLeod, the parents of white youth were much less likely to discipline their children or to encourage them to achieve and do well in school. Also, white youth had more experience on the job market than the African American youth. This contributed to a more pessimistic outlook and a lowering of their future aspirations.

Finally, MacLeod argues that the white youths' immersion in a subculture, which emphasized rejecting the authority of the school, reinforced their negative attitudes to a much greater extent than the African American peer group.

In sum, theories of the macro-level origins of deviance argue that many of the causes of deviance may be found in the characteristics of groups within society, or in the characteristics of geographic areas and communities. They offer explanations of group and areal differences in deviance—for example, why some cities have relatively higher rates of crime than other cities or why blacks have higher rates of serious interpersonal violence than other ethnic groups. These theories make no attempt to explain the behavior of individuals or the occurrence of individual deviant acts. Indeed, they reason that deviance is best understood as a property of an area, community, or group, regardless of the individuals living in the area or community, or the individuals comprising the group.

The theories' implications for public policy focus on the characteristics of geographic areas and communities that lead to deviance. The impact of change on neighborhoods, for example, can be reduced if the boundaries of residential areas are preserved. By preserving such boundaries, communities are less likely to become transitional neighborhoods that foster deviance and crime. Also, by maintaining residential properties people become invested in their own community, which helps foster the mechanisms of informal social control that make deviance less likely. Strengthening schools and other stabilizing institutions in neighborhoods, such as churches and community centers, can also contribute to a reduction in deviance. Finally, establishing networks for jobs and job placement in disadvantaged areas may increase the opportunities of employment among youth. If they succeed in increasing employment, the networks should decrease the chances that youth will turn to careers in crime.

#### **MICRO-LEVEL ORIGINS OF DEVIANCE**

Many explanations of deviance argue that its causes are rooted in the background or personal circumstances of the individual. Micro-level origin theories have developed over the past fifty years, identifying mechanisms by which ordinarily conforming

individuals may become deviant. These theories assume the existence of a homogeneous, pervasive set of norms in society and proceed to explain why persons or entire groups of persons violate the norms. There exist two important traditions within this category of theories. The first tradition involves "social learning theories"—explanations that focus on the mechanisms through which people learn the techniques and attitudes favorable to committing deviant acts. The second tradition involves "social control theories"—explanations that emphasize factors in the social environment that regulate the behavior of individuals, thereby preventing the occurrence of deviant acts.

Edwin Sutherland's (1947) theory of differential association laid the foundation for learning theories. At the heart of this theory is the assumption that deviant behavior, like all other behaviors, is learned. Further, this learning occurs within intimate personal groups—networks of family members and close friends. Thus, according to these theories individuals learn deviance from persons closest to them. Sutherland specified a process of differential association, reasoning that persons become deviant in association with deviant others. Persons learn from others the techniques of committing deviant acts and attitudes favorable to the commission of those acts. Further, Sutherland reasoned that persons vary in their degree of association with deviant others; persons regularly exposed to close friends and family members who held beliefs favoring deviance and who committed deviant acts would be much more likely than others to develop those same beliefs and commit deviant acts.

Sutherland's ideas about learning processes have played a lasting role in micro-level deviance theories. Central to his perspective is the view that beliefs and values favoring deviance are a primary cause of deviant behavior. Robert Burgess and Ronald Akers (1966) and subsequently Akers (1985) extended Sutherland's ideas, integrating them with principles of operant conditioning. Reasoning that learning processes may best be understood in terms of the concrete rewards and punishments given for behavior, Burgess and Akers argue that deviance is learned through associations with others and through a system of rewards and punishments, imposed by close friends and relatives, for participation in deviant acts. Subsequent empirical studies offer compelling support for elements

of learning theory (Matsueda 1982; Akers et al. 1979; Matsueda and Heimer 1987).

Some examples may be useful at this point. According to the theory of differential association, juveniles develop beliefs favorable to the commission of delinquent acts and knowledge about the techniques of committing deviant acts from their closest friends, typically their peers. Thus, sufficient exposure to peers endorsing beliefs favoring deviance who also have knowledge about the commission of deviant acts will cause the otherwise conforming juvenile to commit deviant acts. Thus, if adolescent peer influences encourage smoking, drinking alcohol, and other forms of drug abuse—and exposure to these influences occurs frequently, over a long period of time, and involves relationships that are important to the conforming adolescent—then he or she is likely to develop beliefs and values favorable to committing these acts. Once those beliefs and values develop, he or she is likely to commit the acts.

The second class of micro-level origin theories, control theories, explores the causes of deviance from an altogether different perspective. Control theories take for granted the existence of a cohesive set of norms shared by most persons in the society and reason that most persons want to and will typically conform to these prevailing social norms. The emphasis in these theories, unlike learning theories, is on the factors that bond individuals to conforming lifestyles. The bonds act as social and psychological constraints on the individual, binding persons to normative conformity (Toby 1957; Hirschi 1969). People deviate from norms when these bonds to conventional lifestyles are weak, and hence, when they have little restraining influence over the individual. Among control theorists, Travis Hirschi (1969) has made the greatest contributions to our knowledge about bonding processes and deviant behavior. Writing on the causes of delinquency, he argued that four aspects of bonding are especially relevant to control theory: emotional attachments to conforming others, psychological commitments to conformity, involvements in conventional activities, and beliefs consistent with conformity to prevailing norms.

Among the most important of the bonding elements are emotional attachments individuals may have to conforming others and commitments

to conformity—psychological investments or stakes people hold in a conforming lifestyle. Those having weak attachments—that is, people who are insensitive to the opinions of conforming others—and who have few stakes in conformity, in the form of commitments to occupation or career and education, are more likely than others to deviate (see, e.g., Paternoster et al. 1983; Thornberry and Christenson 1984; Liska and Reed 1985). In effect, these individuals are “free” from the constraints that ordinarily bond people to normative conformity. Conversely, individuals concerned about the opinions of conforming others and who have heavy psychological investments in work or school will see the potential consequences of deviant acts—rejection by friends or loss of a job—as threatening or costly, and consequently will refrain from those acts.

A related concern is the role of sanctions in preventing deviant acts. Control theorists like Hirschi reason that most people are utilitarian in their judgments about deviant acts, and thus evaluate carefully the risks associated with each act. Control theories typically maintain that the threat of sanctions actually prevents deviant acts when the risks outweigh the gains. Much of the most recent writing on sanctions and their effects has stressed the importance of perceptual processes in decisions to commit deviant acts (Gibbs 1975, 1977; Tittle 1980; Paternoster et al. 1982, 1987; Piliavin et al. 1986; Matsueda, Piliavin, and Gartner 1988). At the heart of this perspective is the reasoning that individuals perceiving the threat of sanctions as high are much more likely to refrain from deviance than those perceiving the threat as low, regardless of the actual level of sanction threat.

Writing from the social control perspective attempts to build on and extend the basic assumptions and propositions of control theory. Michael Gottfredson, in conjunction with Hirschi, has developed a general theory of crime that identifies “low self-control,” as opposed to diminished social control, as the primary cause of deviant behavior (Hirschi and Gottfredson 1987; Gottfredson and Hirschi 1990). Arguing that all people are inherently self-interested, pursuing enhancement of personal pleasure and avoiding pain, Gottfredson and Hirschi suggest that most crimes, and for that matter most deviant acts, are the result of choices to maximize pleasure, minimize pain, or both.

Crimes occur when opportunities to maximize personal pleasure are high and when the certainty of painful consequences is low. Further, people who pursue short-term gratification with little consideration for the long-term consequences of their actions are most prone to criminal behavior. In terms of classical control theory, these are individuals who have weak bonds to conformity or who disregard or ignore the potentially painful consequences of their actions. They are “relatively unable or unwilling to delay gratification; they are indifferent to punishment and the interests of others” (Hirschi and Gottfredson 1987, pp. 959–960).

Building on traditional control theory, Charles Tittle (1995) reasons that it helps explain why individuals conform, but it also helps to explain why they engage in deviant behavior. Tittle (1995, p. 135) argues that “the amount of control to which an individual is subject, relative to the amount of control he or she can exercise, determines the probability of deviance occurring as well as the type of deviance likely to occur.” Conformity results when individuals are subjected to and exert roughly equal amounts of control—there is “control balance.” According to Tittle, however, individuals who are subjected to more control than they exert will be motivated to engage in deviance in order to escape being controlled by others.

Robert Sampson and John Laub (1993) have also expanded on the basic propositions of control theory. In their research, Sampson and Laub focus on stability and change in the antisocial behavior of individuals as they grow from juveniles to adults. Sampson and Laub argue that family, school, and peer relationships influence the likelihood of deviant behavior among juveniles. In particular, Sampson and Laub argue that the structure of the family (e.g., residential mobility, family size) affects family context or process (e.g., parental supervision, discipline), which, in turn, makes deviance among children more or less likely. Many adolescent delinquents grow up to become adult criminals because their juvenile delinquency makes the formation of adult social bonds to work and family less likely. Despite this continuity in antisocial behavior from adolescence to adulthood, however, Sampson and Laub argue that many juvenile delinquents do not commit deviant acts as adults because they develop adult social bonds, such as attachment to a spouse or commitment to a job.

In sum, micro-level origin theories look to those aspects of the individual’s social environment that influence her or his likelihood of deviance. Learning theories stress the importance of deviant peers and other significant individuals, and their impact on attitudes and behaviors favorable to the commission of deviant acts. These theories assume that the social environment acts as an agent of change, transforming otherwise conforming individuals into deviants through peer influences. People exposed to deviant others frequently and sufficiently, like persons exposed to a contagious disease who become ill, will become deviant themselves. Control theories avoid this “contagion” model, viewing the social environment as a composite of controls and restraints cementing the individual to a conforming lifestyle. Deviance occurs when elements of the bond— aspects of social control—are weak or broken, thereby freeing the individual to violate social norms. Sanctions and the threat of sanctions are particularly important to control theories, a central part of the calculus that rational actors use in choosing to commit or refrain from committing deviant acts.

The policy implications of micro-level origin theories are obvious. If, as learning theories argue, deviance is learned through association with deviant peers, then the way to eliminate deviance is to assist youths in resisting deviant peer influences and helping them to develop attitudes that disapprove of deviant behavior. Control theories, on the other hand, suggest that deviance can be reduced with programs that help families develop stronger bonds between parents and children. Control theory also implies that programs that help youths develop stronger commitments to conventional lines of activity and to evaluate the costs and benefits of deviant acts will also result in a reduction of problematic behavior.

#### **MICRO-LEVEL REACTIONS TO DEVIANCE**

Unlike micro-level origin theories, micro-level reaction theories make no assumptions about the existence of a homogeneous, pervasive set of norms in society. These theories take an altogether different approach to explaining deviant behavior, viewing deviance as a matter of definition; a social status imposed by individuals or groups on others. Most argue that there exists no single pervasive set



of norms in society and that deviant behavior may best be understood in terms of norms and their enforcement. These theories typically stress the importance of labeling processes—the mechanisms by which acts become defined or labeled as “deviant—and the consequences of labeling for the person so labeled. Many of these theories are concerned with the development of deviant lifestyles or careers; long-term commitments to deviant action.

One of the most important writers in this tradition is Howard Becker (1963). Becker argues that deviance is not a property inherent in any particular form of behavior but rather a property conferred on those behaviors by audiences witnessing them. Becker (1963, p. 9) notes that “. . . deviance is *not* a quality of the act the person commits, but rather a consequence of the application by others of rules and sanctions to an ‘offender.’ The deviant is one to whom that label has been successfully applied; deviant behavior is behavior that people so label.” Thus, Becker and others in this tradition orient the study of deviance on rules and sanctions, and the application of labels. Their primary concern is the social construction of deviance—that is, how some behaviors and classes of people come to be defined as “deviant” by others observing and judging the behavior.

Building on the idea that deviance is a property conferred on behavior that is witnessed by a social audience, Becker (1963) also developed a simple typology of deviant behavior. The dimensions upon which the typology is based are whether or not the individual is perceived as deviant and whether or not the behavior violates any rule. *Conforming* behavior is behavior that does not violate any rules and is not perceived as deviant. Individuals in the opposite scenario, in which the person both violates rules and is perceived by others as deviant, Becker labeled *pure deviants*. Some individuals, according to Becker, may be perceived as deviant, even though they have not violated any rules. Becker identified these individuals as the *falsely accused*. Finally, the *secret deviant* is one who has violated the rules, but, nonetheless, is not perceived by others as being deviant.

Equally important is the work of Edwin Lemert (1951). Stressing the importance of labeling to subsequent deviant behavior, he argues that repetitive deviance may arise from social reactions to

initial deviant acts. According to Lemert (1951, p. 287), deviance may often involve instances where “a person begins to employ his deviant behavior. . . as a means of defense, attack or adjustment to the. . . problems created by the consequent social reactions to him.” Therefore, a cause of deviant careers is negative social labeling; instances where reactions to initial deviant acts are harsh and reinforce a “deviant” self-definition. Such labeling forces the individual into a deviant social role, organizing his or her identity around a pattern of deviance that structures a way of life and perpetuates deviant behavior (Becker 1963; Schur 1971, 1985).

Perhaps the most significant developments in this tradition have contributed to knowledge about the causes of mental illness. Proponents of micro-level reaction theories argue that the label “mental illness” can be so stigmatizing to those labeled, especially when mental-health professionals impose the label, that they experience difficulty returning to nondeviant social roles. As a result, the labeling process may actually exacerbate mental disorders. Former mental patients may find themselves victims of discrimination at work, in personal relationships, or in other social spheres (Scheff 1966). This discrimination, and the widespread belief that others devalue and discriminate against mental patients, may lead to self-devaluation and fear of social rejection by others (Link 1982, 1987). In some instances, this devaluation and fear may be associated with demoralization of the patient, loss of employment and personal income, and the persistence of mental disorders following treatment (Link 1987).

Hence, micro-level reaction theories reason that deviant behavior is rooted in the process by which persons define and label the behavior of others as deviant. The theories offer explanations of individual differences in deviance, stressing the importance of audience reactions to initial deviant acts. However, these theories make no attempt to explain the origins of the initial acts (Scheff 1966). Rather, they are concerned primarily with the development and persistence of deviant careers.

Micro-level reaction theories have very different implications for public policy than macro- and micro-level origins theories. Micro-level reaction theories argue that unwarranted labeling can lead

to deviant careers. In effect, the reaction to deviance can cause deviant behavior to escalate. Thus, in order to reduce deviance, agencies of social control must adopt policies of nonintervention. Rather than being formally sanctioned and labeled as deviant, nonintervention policies must encourage diversion and deinstitutionalization. Formal sanctioning must be highly selective, focusing only on the most serious and threatening deviant acts.

### **MACRO-LEVEL REACTIONS TO DEVIANCE**

The final class of theories looks to the structure of economic and political power in society as a cause of deviant behavior. Macro-level reaction theories—either Marxist or other conflict theories—view deviance as a status imposed by dominant social classes to control and regulate populations that threaten political and economic hegemony. Like micro-level reaction theories, these theories view deviance as a social construction and accord greatest importance to the mechanisms by which society defines and controls entire classes of behavior and people as deviant in order to mediate the threat. However, these theories reason that the institutional control of deviants has integral ties to economic and political order in society.

Marxist theories stress the importance of the economic structure of society and begin with the assumption that the dominant norms in capitalist societies reflect the interests of the powerful economic class; the owners of business. But contemporary Marxist writers (Quinney 1970, 1974, 1980; Spitzer 1975; Young 1983) also argue that modern capitalist societies are characterized by large “problem populations”—people who have become displaced from the workforce and alienated from the society. Generally, the problem populations include racial and ethnic minorities, the chronically unemployed, and the extremely impoverished. They are a burden to the society and particularly to the capitalist class because they create a form of social expense that must be carefully controlled if the economic order is to be preserved.

Marxist theories reason that economic elites use institutions such as the legal, mental-health, and welfare systems to control and manage society’s problem populations. In effect, these institutions define and process society’s problem populations as deviant in order to ensure effective

management and control. In societies or communities characterized by rigid economic stratification, elites are likely to impose formal social control in order to preserve the prevailing economic order.

Conflict theories stress the importance of the political structure of society and focus on the degree of threat to the hegemony of political elites, arguing that elites employ formal social controls to regulate threats to political and social order (Turk 1976; Chambliss 1978; Chambliss and Mankoff 1976). According to these theories, threat varies in relation to the size of the problem population, with large problem populations substantially more threatening to political elites than small populations. Thus, elites in societies and communities in which those problem populations are large and perceived as especially threatening are more likely to process members of the problem populations as deviants than in areas where such problems are small.

Much of the writing in this tradition has addressed the differential processing of people defined as deviant. Typically, this writing has taken two forms. The first involves revisionist histories linking the development of prisons, mental asylums, and other institutions of social control to structural changes in U.S. and European societies. These histories demonstrate that those institutions often target the poor and chronically unemployed independent of their involvement in crime and other deviant acts, and thereby protect and serve the interests of dominant economic and political groups (Scull 1978; Rafter 1985).

A second and more extensive literature includes empirical studies of racial and ethnic disparities in criminal punishments. Among the most important of these studies is Martha Myers and Suzette Talarico’s (1987) analysis of the social and structural contexts that foster racial and ethnic disparities in the sentencing of criminal offenders. Myers and Talarico’s research, and other studies examining the linkages between community social structure and differential processing (Myers 1987, 1990; Peterson and Hagan 1984; Bridges, Crutchfield, and Simpson 1987; Bridges and Crutchfield 1988), demonstrate the vulnerability of minorities to differential processing during historical periods and in areas in which they are perceived by whites as serious threats to political and social order. In effect, minorities accused of crimes during these

periods and in these geographic areas are perceived as threats to white hegemony, and therefore become legitimate targets for social control.

In addition to studying the connections between community social structure and the differential processing of racial and ethnic minorities, researchers have also begun to examine how court officials' perceptions of offenders can influence disparities in punishments. Bridges and Steen (1998), for example, show how court officials' perceptions of white and minority youths differ, and how these different perceptions contribute to different recommendations for sentencing. Probation officers often attribute the offenses of minority youths to internal characteristics of the youths (i.e., aspects of their personality), while attributing the offenses of white youths to external characteristics (i.e., aspects of their environments). As a result of these differential attributions, minority youths are perceived as more threatening, more at risk for re-offending than whites and more likely to receive severe recommendations for sentences.

Thus, macro-level reaction theories view deviance as a by-product of inequality in modern society, a social status imposed by powerful groups on those who are less powerful. Unlike micro-level reaction theories, these theories focus on the forms of inequality in society and how entire groups within the society are managed and controlled as deviants by apparatuses of the state. Like those theories, however, macro-level reaction theories make little or no attempt to explain the origins of deviant acts, claiming instead that the status of "deviant" is, in large part, a social construction designed primarily to protect the interests of the most powerful social groups. The primary concern of these theories is explicating the linkages between inequality in society and inequality in the labeling and processing of deviants.

Since macro-level reaction theories view deviance as a status imposed by powerful groups on those with less power, the most immediate policy implication of these theories is that imbalances in power and inequality must be reduced in order to reduce levels of deviance and levels of inequality in the sanctioning of deviance. More effective monitoring of government agencies that are used to control problem populations, such as the criminal

justice system, can also help to reduce the disproportionate processing of less powerful groups, such as racial minorities, as deviant.

### NEW THEORETICAL DIRECTIONS

A recurring issue in the study of deviance is the contradictory nature of many deviance theories. The theories often begin with significantly different assumptions about the nature of human behavior and end with significantly different conclusions about the causes of deviant acts. Some scholars maintain that the oppositional nature of these theories—the theories are developed and based on systematic rejection of other theories (Hirschi 1989)—tends toward clarity and internal consistency in reasoning about the causes of deviance. However, other scholars argue that this oppositional nature is intellectually divisive—acceptance of one theory precludes acceptance of another—and "has made the field seem fragmented, if not in disarray" (Liska, Krohn, and Messner 1989, p. 1).

A related and equally troublesome problem is the contradictory nature of much of the scientific evidence supporting deviance theories. For each theory, there exists a literature of studies that supports and a literature that refutes major arguments of the theory. And although nearly every theory of deviance may receive empirical confirmation at some level, virtually no theory of deviance is sufficiently comprehensive to withstand empirical falsification at some other level. The difficult task for sociologists is discerning whether and under what circumstances negative findings should be treated as negating a particular theory (Walker and Cohen 1985).

In recent years, these two problems have renewed sociologists' interest in deviance theory and, at the same time, suggested new directions for the development of theory. The oppositional nature of theories has spawned interest in theoretical integration. Many scholars are dissatisfied with classical theories, arguing that their predictive power is exceedingly low (see Elliott 1985; Liska, Krohn, and Messner 1989). Limited to a few key explanatory variables, any one theory can explain only a limited range and amount of deviant behavior. And because most scholars reason that the causes of deviance are multiple and quite complex, most also contend that it may be "necessary to combine different theories to capture the

entire range of causal variables" (Liska, Krohn, and Messner 1989, p. 4).

Because it combines the elements of different theories, the new theory will have greater explanatory power than theories from which it was derived. However, meaningful integration of deviance theories will require much more than the simple combination of variables. Scholars must first reconcile the oppositional aspects of theories, including many of their underlying assumptions about society, the motivations of human behavior, and the causes of deviant acts. For example, learning theories focus heavily on the motivations for deviance, stressing the importance of beliefs and values that "turn" the individual to deviant acts. In contrast, control theories accord little importance to such motivations, examining instead those aspects of the social environment that constrain people from committing deviant acts. Reconciling such differences is never an easy task, and in some instances may be impossible (Hirschi 1979).

The problem of contradictory evidence suggests a related but different direction for deviance theory. Theories may vary significantly in the conditions—termed *scope conditions*—under which they apply (Walker and Cohen 1985; Tittle 1975; Tittle and Curran 1988). Under some scope conditions, theories may find extensive empirical support, and under others virtually none. For instance, macro-level origin theories concerned with the frustrating effects of poverty on deviance may have greater applicability to people living in densely populated urban areas than those living in rural areas. The frustration of urban poverty may be much more extreme than in rural areas, even though the actual levels of poverty may be the same. As a result, the frustrations of urban poverty may be more likely to cause deviant adaptations in the form of violent crime, drug abuse, and vice than those of rural poverty. In this instance, "urbanness" may constitute a condition that activates strain theories linking poverty to deviance. Obviously, the same theories simply may not apply in rural areas or under other conditions.

Effective development of deviance theory will require much greater attention to the specification of such scope conditions. Rather than combining causal variables from different theories as integrationists would recommend, this approach

to theory development encourages scholars to explore more fully the strengths and limitations of their own theories. This approach will require more complete elaboration of extant theory, explicitly specifying those circumstances under which each theory may be meaningfully tested and thus falsified. The result will be a greater specification of each theory's contribution to explanations of deviant behavior.

These two directions have clear and very different implications for the development of deviance theory. Theoretical integration offers overarching models of deviant behavior that cut across classical theories, combining different levels of explanation and causal focuses. If fundamental differences between theories can be reconciled, integration is promising. The specification of scope conditions offers greater clarification of existing theories, identifying those conditions under which each theory most effectively applies. Although this direction promises no general theories of deviance, it offers the hope of more meaningful and useful explanations of deviant behavior.

REFERENCES

- Agnew, Robert 1992 "Foundation for a General Strain Theory of Crime and Delinquency." *Criminology* 30:47-88.
- Akers, Ronald L. 1985 *Deviant Behavior: A Social Learning Approach*. Belmont, Calif.: Wadsworth.
- , Marvin D. Krohn, Lonn Lanza-Kaduce, and Marcia Radosevich 1979 "Social Learning and Deviant Behavior: A Specific Test of a General Theory." *American Sociological Review* 44:636-655.
- Becker, Howard 1963 *The Outsiders*. Chicago: University of Chicago Press.
- Blau, Judith, and Peter Blau 1982 "Metropolitan Structure and Violent Crime." *American Sociological Review* 47:114-128.
- Bridges, George S., and Robert D. Crutchfield 1988 "Law, Social Standing and Racial Disparities in Imprisonment." *Social Forces* 66:699-724.
- , and Edith Simpson 1987 "Crime, Social Structure and Criminal Punishment." *Social Problems* 34:344-361.
- , and Sara Steen 1998 "Racial Disparities in Official Assessments of Juvenile Offenders: Attributional Stereotypes as Mediating Mechanisms." *American Sociological Review* 63:554-570.

- Burgess, Robert L., and Ronald Akers 1966 "A Differential Association-Reinforcement Theory of Criminal Behavior." *Social Problems* 14:128-147.
- Bursik, Robert J. 1984 "Urban Dynamics and Ecological Studies of Delinquency." *Social Forces* 63:393-413.
- , and J. Webb 1982 "Community Change and Ecological Studies of Delinquency." *American Journal of Sociology* 88:24-42.
- Chambliss, William 1978 *On the Take*. Bloomington: Indiana University Press.
- , and Milton Mankoff 1976 *Whose Law, What Order?* New York: Wiley.
- Cloward, Richard 1959 "Illegitimate Means, Anomie and Deviant Behavior." *American Sociological Review* 24:164-176.
- , and Lloyd E. Ohlin 1960 *Delinquency and Opportunity*. New York: Free Press.
- Crutchfield, Robert D. 1989 "Labor Stratification and Violent Crime." *Social Forces* 68:489-513.
- Curtis, Lynn 1975 *Violence, Race and Culture*. Lexington, Mass.: Heath.
- Elliott, Delbert 1985 "The Assumption That Theories can be Combined with Increased Explanatory Power: Theoretical Integrations." In Robert F. Meier, ed., *Theoretical Methods in Criminology*. Beverly Hills, Calif.: Sage.
- Erlanger, Howard 1974 "The Empirical Status of the Subculture of Violence Thesis." *Social Problems* 22:280-292.
- Gastil, Raymond 1971 "Homicide and a Regional Culture of Violence." *American Sociological Review* 36:412-427.
- Gibbs, Jack P. 1975 *Crime, Punishment and Deterrence*. New York: Elsevier.
- 1977 "Social Control, Deterrence, and Perspectives on Social Order." *Social Forces* 62:359-374.
- Gottfredson, Michael R., and Travis Hirschi 1990 *A General Theory of Crime*. Palo Alto, Calif.: Stanford University Press.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- 1979 "Separate and Unequal is Better." *Journal of Research in Crime and Delinquency* 16:34-37.
- 1989 "Exploring Alternatives to Integrated Theory." In Steven F. Messner, Marvin D. Krohn, and Allen E. Liska, eds., *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*. Albany, N.Y.: SUNY Press.
- , and Michael R. Gottfredson 1987 "Causes of White Collar Crime." *Criminology* 25:949-974.
- Lemert, Edwin 1951 *Social Pathology*. New York: McGraw-Hill.
- Link, Bruce 1982 "Mental Patient Status, Work and Income: An Examination of the Effects of a Psychiatric Label." *American Sociological Review* 47:202-215.
- 1987 "Understanding Labeling Effects in the Area of Mental Disorders: An Assessment of the Effects of Expectations of Rejection." *American Sociological Review* 52:96-112.
- Liska, Allen E., Marvin D. Krohn, and Steven S. Messner 1989 "Strategies and Requisites for Theoretical Integration in the Study of Deviance and Crime." In Steven S. Messner, Marvin D. Krohn, and Allen E. Liska, eds., *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*. Albany, N.Y.: SUNY Press.
- MacLeod, Jay 1995 *Ain't No Makin' It*. Boulder, Colo.: Westview.
- Matsueda, Ross L. 1982 "Testing Control Theory and Differential Association: A Causal Modeling Approach." *American Sociological Review* 47:36-58.
- , and Karen Heimer 1987 "Race, Family Structure and Delinquency: A Test of Differential Association and Social Control Theories." *American Sociological Review* 52:826-840.
- , Irving Piliavin, and Rosemary Gartner 1988 "Ethical Assumptions Versus Empirical Research." *American Sociological Review* 53:305-307.
- Merton, Robert K. 1938 "Social Structure and Anomie." *American Sociological Review* 3:672-682.
- 1964 *Social Theory and Social Structure*. New York: Free Press.
- Messner, Steven F. 1983 "Regional and Racial Effects on the Urban Homicide Rate: The Subculture of Violence Revisited." *American Journal of Sociology* 88:997-1,007.
- Miller, Walter B. 1958 "Lower Class Culture as a Generating Milieu of Gang Delinquency." *Journal of Social Issues* 14:5-19.
- Myers, Martha 1987 "Economic Inequality and Discrimination in Sentencing." *Social Forces* 65:746-766.
- 1990 "Economic Threat and Racial Disparities in Incarceration: The Case of Postbellum Georgia." *Criminology* 28:627-656.
- , and Suzette Talarico 1987 *Social Contexts of Criminal Sentencing*. New York: Springer-Verlag.
- Paternoster, Raymond L., Linda Saltzman, Gordon P. Waldo, and Theodore Chiricos 1982 "Perceived Risk and Deterrence: Methodological Artifacts in Deterrence Research." *Journal of Criminal Law and Criminology* 73:1,243-1,255.

- 1983 “Perceived Risk and Social Control.” *Journal of Criminal Law and Criminology* 74:457–480.
- Peterson, Ruth D., and John Hagan 1984 “Changing Conceptions of Race: Toward an Account of Anomalous Findings of Sentencing Research.” *American Sociological Review* 49:56–70.
- Piliavin, Irving, Rosemary Gartner, Craig Thornton, and Ross L. Matusueda 1986 “Crime, Deterrence and Rational Choice.” *American Sociological Review* 51:101–120.
- Quinney, Richard 1970 *The Social Reality of Crime*. Boston: Little, Brown.
- 1974 *Critique of Legal Order*. Boston: Little, Brown.
- 1980 *Class, State and Crime*. New York: Longman.
- Rafter, Nicole Hahn 1985 *Partial Justice: Women in State Prisons, 1800–1935*. Boston: Northeastern University Press.
- Sampson, Robert J., and John H. Laub 1993 *Crime in the Making: Pathways and Turning Points Through Life*. Cambridge, Mass.: Harvard University Press.
- Scheff, Thomas 1966 *Being Mentally Ill*. Chicago: Aldine.
- Schur, Edwin 1971 *Labeling Deviant Behavior*. New York: Harper and Row.
- 1985 *Labeling Women Deviant: Gender, Stigma and Social Control*. New York: Harper and Row.
- Scull, Andrew T. 1978 *Museums of Madness*. London: Heinemann.
- Sellin, Thorsten 1938 *Culture, Conflict and Crime*. New York: Social Science Research Council.
- Shaw, Clifford 1930 *The Jack-Roller*. Chicago: University of Chicago Press.
- , and Henry McKay 1942 *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Silberman, Charles 1980 *Criminal Justice, Criminal Violence*. New York: Vintage.
- Spitzer, Steven 1975 “Toward a Marxian Theory of Deviance.” *Social Problems* 22:638–651.
- Stark, Rodney 1987 “Deviant Places: A Theory of the Ecology of Crime.” *Criminology* 25:893–910.
- Sullivan, Mercer L. 1989 *Getting Paid: Youth Crime and Work in the Inner City*. Ithaca, N.Y.: Cornell University Press.
- Sutherland, Edwin H. 1947 *Principles of Criminology*, 4th ed. Philadelphia: Lippincott.
- Thornberry, Terence P., and R. L. Christenson 1984 “Unemployment and Criminal Involvement: An Investigation of Reciprocal Causal Structures.” *American Sociological Review* 49:398–411.
- Tittle, Charles R. 1975 “Deterrents or Labeling?” *Social Forces* 53:399–410.
- 1980 *Sanctions and Social Deviance*. New York: Praeger.
- 1995 *Control Balance: Toward a General Theory of Deviance*. Boulder, Colo.: Westview Press.
- , and Barbara A. Curran 1988 “Contingencies for Dispositional Disparities in Juvenile Justice.” *Social Forces* 67:23–58.
- Toby, Jackson 1957 “Social Disorganization and Stakes in Conformity: Complimentary Factors in the Predatory Behavior of Hoodlums.” *Journal of Criminal Law and Criminology and Police Science* 48:12–17.
- Turk, Austin T. 1976 “Law as a Weapon in Social Conflict.” *Social Problems* 23:276–291.
- Walker, Howard, and Bernard P. Cohen 1985 “Scope Statements.” *American Sociological Review* 50:288–301.
- Wolfgang, Marvin E., and Franco Ferracuti 1967 *The Subculture of Violence: Towards an Integrated Theory in Criminology*. London: Tavistock.
- Young, Peter 1983 “Sociology, the State and Penal Relations.” In David Garland and Peter Young, eds., *The Power to Punish*. London: Heinemann.

GEORGE S. BRIDGES  
SCOTT A. DESMOND

## DIFFERENTIAL ASSOCIATION

See Crime, Theories of; Deviance Theories.

## DIFFUSION THEORIES

The concept of *diffusion* inherently focuses upon *process*. Diffusion refers to the dissemination of any physical element, idea, value, social practice, or attitude through and between populations. Diffusion is among the rare concepts used across the physical, natural, and social sciences, as well as in the arts. Diffusion is most closely associated with the social sciences, particularly rural sociology, anthropology, and communication. Diffusion thinking offers a logic through which to describe and perhaps explain myriad types of change that involve equally diverse foci, ranging from the adoption of internet technology (Adams 1997), to the spread of belief systems (Dean 1997).

Work connected to the concept of diffusion is arguably structured as theory. Certainly there is no

single, unified, deductively structured collection of propositions, widely regarded by social scientists as identifying the principal mechanisms of diffusion that could be employed across all substantive areas. There are however, many distinct collections of propositions, a few well-tested over decades, that describe different diffusion phenomena in different content areas. Indeed, in the area of innovation diffusion, Everett Rogers (1983) produced a formal theory that is broadly recognized, often tested, and that has been adapted to other content areas including disaster research and technology transfer. Part of the problem in formalizing diffusion theories is that the concept does not inherently specify content (rather a framework or process to structure thinking). However, it cannot be examined empirically without tying it to some substance. That is, studies of diffusion focus on *something* (a technology, idea, practice, attitude, etc.) that is being diffused. Consequently, the research on diffusion that would drive theory building has remained scattered in the literatures of different sciences, and as such it is not readily pulled together. These conditions do not facilitate ready assembly of information that would encourage creation of a general theory of diffusion. Thus, one must look to the growth of formal theory in sub-areas, although Rogers (in press) has begun to cross content spheres.

There are at least three traditions or theory families that can be historically discerned in the study of diffusion. Rogers (1983, p. 39) has pointed out that for many years these traditions remained largely distinct, with little overlap and cross-fertilization. Since the late 1970s the level of research and theoretical isolation has decreased, leading to an enhanced awareness among the perspectives and some integration of empirical findings into more general theoretical statements. The three theory families are: (1) cultural diffusion; (2) diffusion of innovations; and (3) collective behavior.

### CULTURAL DIFFUSION

The earliest social scientific use of the term *diffusion* is found in Edward Tylor's (1865) treatment of culture change. Anthropologists have long attempted to explain similarities and differences among cultures, especially those that were geographically adjacent. Tylor's work on culture change first

proposed the notion of diffusion as a means of explaining the appearance of similar culture elements in different groups and of understanding the progressive alteration of elements within the same group. As the twentieth century began, diffusion arose as an alternative to evolution as a basis for understanding cultural differences and change. Evolutionists argued that cultural similarities probably arose through independent invention. Those who embraced diffusion presented it as a more parsimonious explanation, emphasizing that traits and institutions could pass between groups by means of contact and interaction.

**Historical Development.** The English anthropologists W. J. Perry and Elliot Smith devised the most extreme position on cultural diffusion. These scholars held that human culture originated in Egypt and progressively diffused from that center over the remainder of the earth. In Germany, Fritz Graebner (1911) argued that critical aspects of cultures—toolmaking, for example—originated in a small number of geographically isolated societies. This hypothesis formed the basis for *culture circles* (“*kulturkreise*”), collections of societies sharing similar cultures. Unlike British diffusionists who emphasized tracking the movement of single culture elements, Graebner and others in his tradition focused on the dissemination of collections of elements or cultural complexes.

American anthropologists are credited with developing a social scientifically workable concept of diffusion. Franz Boas (1896) conceived of diffusion as a viable mechanism for culture exchanges among geographically adjacent areas. His view figured prominently in the intellectual move away from the deterministic view of diffusion proposed by early British anthropologists. Alfred Kroeber (1923, p. 126) and Robert Lowie (1937, p. 58)—students of Boas—subsequently developed a position called *moderate diffusionism*, which is currently widely accepted in anthropology. This position allowed for the coexistence of a variety of mechanisms of change and transfer—independent invention, acculturation, etc.—in addition to diffusion in accounting for culture change and differentiation. Clark Wissler (1929), a Kroeber contemporary, established an empirical basis for culture diffusion by identifying ten culture areas (regions with similar cultural inventories) in North and South America and the Caribbean.

**Theory in Cultural Diffusion.** In terms of theory development, *cultural diffusion* is the actual movement of a given social institution or physical implement, while *stimulus diffusion* is the exchange or movement of the *principle* upon which an institution or implement is based. In the cultural diffusion literature, scholars have enumerated assumptions, stated principles, and reviewed empirical work with the objective of identifying propositions tested repeatedly and not found to be false. Indeed, beginning with the work of early twentieth century anthropologists, one can identify at least five broadly accepted and empirically supported claims that form the core of what is called cultural diffusion theory. First, borrowed elements usually undergo some type of alteration or adaptation in the new host culture. Second, the act of borrowing depends on the extent to which the element can be integrated into the belief system of the new culture. Third, elements that are incompatible with the new culture's prevailing normative structure or religious belief system are likely to be rejected. Fourth, acceptance of an element depends upon its utility for the borrower. Finally, cultures with a history of past borrowing are more likely to borrow in the future. These claims constitute the "core propositions" of culture diffusion theory; over the years, each has been qualified and elaborated upon, and corollaries have been created (Stahl 1994).

Currently, diffusion is seen as a mechanism for culture change that typically accounts for a large proportion of any particular culture inventory. The deterministic, linear view of diffusion has been discredited by the empirical record. The concept of culture diffusion as a means of understanding cultural inventories is entrenched in the field of cultural anthropology. As the twenty-first century dawns, a principal controversy among cultural anthropologists centers on the definition of culture, rather than upon the acceptability of culture diffusion. Diffusion theory remains prominent in the archaeology literature, particularly as a means of tracing culture inventories for groups over time (Posnansky and DeCorse 1986).

Sociologists were initially involved in the use of cultural diffusion theory as a means of looking at cultural change (largely in terms of nonmaterial culture) in the United States. Initially theory drawn from anthropology was used (Chapin 1928), but

over time the sociological focus became identifying social psychological motivations and mechanisms supporting the diffusion process (Park and Burgess 1921, p. 20). Recently, sociological work directly on culture is bifurcated, with one group of scientists still emphasizing the social psychological issues in culture meaning (Wuthnow and Witten 1988), and another more concerned with structural (mathematical or statistical) models of culture processes themselves (Griswold 1987). Neither group has especially focused on diffusion theory as a mechanism to track or identify the content-outcomes of culture change.

### DIFFUSION OF INNOVATIONS

The diffusion of innovations has historically focused on the spread of an idea, procedure, or implement within a single social group or between multiple groups. For the most part, scholars of this tradition define diffusion as the process through which some innovation is communicated within a social system. Also important is the notion of a time dimension reflecting the *rate* of diffusion, and the importance of the individual adopter (or non-adopter) reflecting the role of social influence.

The study of innovation diffusion began rather narrowly, grew to dominate the field of rural sociology for a time, contracted in popularity for many years, and then spawned wide interest across several disciplines. Innovation diffusion study contains several groups: those who focus on content or the specific innovation being diffused; those who emphasize theoretical elaborations of generic principles of innovation diffusion; and those concerned with creating structural models to track diffusion. Particularly in the past decade, the literature has seen much cross-fertilization, although mathematical modelers tend to appear less often in the work of other diffusion scholars. Although the roots of innovation diffusion theory are seen to be largely in rural sociology, more recently the field has become distinctly interdisciplinary with major advancements made especially in the discipline of communication.

**Historical Development.** The definitive history of the diffusion of innovations as a paradigm was published by Thomas Valente and Everett Rogers (1995). The roots of innovation diffusion



are usually traced to Gabriel Tarde (1890) who didn't use the term *diffusion*, but was the first to address the notions of *adopters* and the role of social influence in adoption, as well as to identify the S-shaped curve associated with the rate of an innovation's adoption. The formative empirical work on innovation diffusion can be traced to Bryce Ryan's Iowa State University-based study of hybrid corn seeds (published with Gross in 1943), and Raymond Bowers's (1937) study of the acceptance and use of ham radio sets. For more than two decades following this pioneering work, the study of innovation diffusion and particularly theory development took place within the context of rural sociology. This circumstance was a function of a variety of forces, principal among which were the location of rural sociologists in land grant institutions charged with the dissemination of agricultural innovations to farmers (Hightower 1972) and the communication and stimulation accorded by the North Central Rural Sociology Committee's (a regional professional society) formation of a special subcommittee to deal with the issue of diffusion of agricultural innovations (Valente and Rogers 1995, p. 254).

Most scholars agree that contemporary views of innovation diffusion grew from hybrid corn seeds; specifically the research on adoption done by Bryce Ryan and Neal Gross (1943). These studies ultimately defined most of the issues that occupied diffusion researchers and builders of innovation diffusion theory for decades to come: the role of social influence, the timing of adoptions, the adoption process itself, and interactions among adopter characteristics and perceived characteristics of the innovation. From the middle 1940s through the 1950s, rural sociologists vigorously developed a body of empirical information on the diffusion of innovations. Most of these studies remained tied to agriculture and farming, and focused on the diffusion of new crop management systems, hybridizations, weed sprays, insect management strategies, chemical fertilizers, and machinery. A common criticism of the studies of this era is that many of the studies seem to be almost replications of the Ryan and Gross work, the main difference among them being the specific innovation studied. While it is true that these studies tend to share a common methodology and linear conception of diffusion, it is also true that they provide a strong foundation of empirical case studies.

Indeed, the replications that these studies represent substantially facilitated the later sophisticated theoretical work initiated in the early 1960s (Rogers 1962), and continued in the 1980s (Rogers 1983, 1988).

The 1960s marked the beginning of the decline of the central role of rural sociologists in innovation diffusion research. In large part this was due to changes in the field of rural sociology, but it also reflected the increasing involvement of researchers from other disciplines, changing the sheer proportion of rural sociologists working on innovation diffusion. After more than two decades of extensive research on the diffusion of agricultural innovations, rural sociologists—like other social scientists of the time—began to devote more time to the study of social problems and the consequences of technology. Indeed, Crane (1972) argued that around 1960 rural sociologists began to believe that the critical questions about innovation diffusion had already been answered. Although the late 1960s saw rural sociologists launch a series of diffusion studies on agricultural change in the international arena (particularly Latin America, Asia and Africa), by 1965 research on diffusion of innovations was no longer dominated by members of that field. Of course, innovation diffusion research by rural sociologists has continued, including studies of the impacts of technological innovation diffusion, and diffusion of conservation practices and other ecologically-based innovations (Fliegel 1993).

The infusion of researchers from many disciplines studying a variety of specific innovations initiated the process of expanding the empirical testing of innovation diffusion tenets. This began with studies in education addressing the diffusion of kindergartens and driver education classes in the 1950s, as well as Richard Carlson's (1965) study of the diffusion of modern math. Another major contribution came from the area of public health. Elihu Katz, Herbert Menzel, and James Coleman launched extensive studies of the diffusion of a new drug (the antibiotic tetracycline); first in a pilot study (Menzel and Katz 1955) and then in studies of four Illinois cities (Coleman, Menzel, and Katz 1957; Coleman, Katz, and Menzel 1966). This research greatly expanded knowledge of interpersonal diffusion networks, and in particular its influence in adoption. Interestingly, as Elihu Katz, M. L. Levine, and Harry Hamilton

(1963) indicated, the drug studies truly represented an independent replication of the principles of innovation diffusion developed by rural sociologists because the public health researchers were unaware of the agricultural diffusion research. Other studies in the public health arena focused on dissemination of new vaccines, family planning, and new medical technology.

Beginning in the late 1960s, there was a substantial increase in the amount of diffusion research in three disciplines: business marketing, communication, and transportation-technology transfer. Marketing research principally addressed the characteristics of adopters of new products and the role of opinion leaders in the adoption process (Howes 1996). This literature is based almost exclusively on commercial products, ranging from coffee brands and soap to touch-tone telephones, the personal computer, and internet services. The studies tend to be largely atheoretical, methodologically similar, and aimed simply at using knowledge of diffusion either to improve marketing and sales of the product or to describe product dissemination.

In sharp contrast, work done on innovation diffusion by scholars trained in communication has been considerably more theoretically oriented. Throughout the 1960s, universities in America began to establish separate departments of communication (Rogers 1994). Since diffusion of innovations was widely seen as one type of communication process, scholars in these new departments adopted this type of research as one staple of their work. Beginning with studies of the diffusion of news events (Deutschmann and Danielson 1960), this research tradition has branched out to study the dissemination of a wide variety of specific innovations (McQuail 1983, p. 194). Scholars working in this tradition have been principally responsible for the progressive refinements of formalized theory of innovation diffusion. Everett Rogers has consistently remained the leader in theory development in communication, revising and extending his 1962 book *Diffusion of Innovations* with help from co-author Floyd Shoemaker to produce *Communication of Innovations* in 1971. Subsequently, Rogers (1983, in press) restored the original title *Diffusion of Innovations*, broadened the theoretical base and incorporated diffusion studies and thinking from other disciplines. Generally, Rogers and other communication scholars

have studied the diffusion of many target material elements, phenomena, and other intangibles, but they have continued to produce theoretical statements dealing with communication channels, diffusion networks, interpersonal influence and the innovation-decision process. Finally, the technology dissemination and transfer issues have involved work by geographers, engineers, and others beginning in the 1970s. The primary focus of such studies has been the spread or dissemination of technology (Sahal 1981) and the development of network models of innovation diffusion (Valente 1995).

**Theory in Innovation Diffusion.** The theoretical work of Everett Rogers initially resulted in the collection of knowledge gained from the rural sociology tradition, then facilitated the transition to communication perspectives, and now has served as the mainstay of what is developing as a more cross-disciplinary focus on innovation diffusion. His contribution is twofold. First, he created inventories of findings from many disciplines and from many types of innovation. These inventories provided impetus for the development of a definition of innovation diffusion that was not bound by discipline. Second, Rogers assembled and refined theoretical structures aimed at explaining the principal features of innovation diffusion. The theoretical work has cemented a core of knowledge and principles that are widely identified (and used empirically) as the bases of the diffusion of innovations. Rogers's (1983) theory includes eighty-one generalizations (propositions) that have undergone empirical testing.

The theory of innovation diffusion may be understood as capturing the innovation-decision process, innovation characteristics, adopter characteristics, and opinion leadership. The innovation-decision process represents the framework on which diffusion research is built. It delineates the process through which a decision maker (representing any unit of analysis) chooses to adopt, reinvent (modify), or reject an innovation. This process consists of five stages. *Knowledge* is the initial stage when the decision maker detects the existence of the innovation and learns of its function. In the *persuasion* stage, the decision maker forms a positive or negative attitude toward the innovation. The third stage, *decision*, deals with the decision-maker's choice to accept or reject the

innovation. *Implementation*, the fourth stage, follows a decision to accept and involves putting the innovation into some use (in either its accepted form or some modified form). During the final stage of *confirmation*, decision makers assess an adopted innovation, gather information from significant others, and choose to continue to use the innovation as is, modify it (reinvention), or reject it. While some have criticized the stage model as too linear, Rogers (1983) has convincingly argued that existing formulations afford a degree of interpretative and predictive flexibility that averts historical problems with stage models in social science.

Different innovations have different probabilities of adoption and hence, different adoption rates. That is, they travel through the innovation-decision process at varying speeds. The literature demonstrates that five characteristics of innovations influence the adoption decision. *Compatibility* refers to the congruence between an innovation and the prevailing norms, values, and perceived needs of the potential adopter. Higher levels of compatibility are associated with greater likelihood of adoption. Innovation *complexity*, on the other hand, is negatively associated with adoption. The extent to which use of an innovation is visible to the social group—called *observability*—is positively related to adoption. *Relative advantage* refers to the extent to which an innovation is perceived to be “better” than the idea, practice, or element that it replaces. Higher relative advantage increases the probability of adoption. Finally, *trialability*—the extent to which an innovation may be experimented with—also increases the probability of adoption.

The third component of diffusion of innovation theory addresses adopter characteristics. Adopter categories are classifications of individuals by how readily they adopt an innovation. Rogers (1983, p. 260) identifies nine socioeconomic variables, twelve personality variables, and ten personal communication characteristics that have been demonstrated to bear upon adoption choices. In general, the literature holds that early adopters are more likely to be characterized by high socioeconomic status, high tolerance of uncertainty and change, low levels of fatalism and dogmatism, high integration into the social system, high exposure to mass media and interpersonal communication channels, and frequent engagement in information seeking.

Identifying the characteristics of people who adopt innovations raises the question of interpersonal influence. Three issues are addressed in the development of propositions about the role of interpersonal influence in the innovation decision process: information flow, opinion leadership, and diffusion networks. Over time, information flow has been seen as a “hypodermic needle” model, a two-step flow (to opinion leaders, then other adopters), and a multi-step flow. Currently, information flows are seen as multi-step in nature and are described in terms of *homophily* and *heterophily*—the degree to which pairs of interacting potential adopters are similar or dissimilar. Opinion leadership denotes the degree to which one member of a social system can influence the attitude and behavior of others. This concept is presently discussed relative to spheres of influence, wherein a given person may be a leader or follower depending upon the part of the diffusion network being referenced. The diffusion or communication network is the structural stage upon which social influence takes place. Considerable attention has been devoted to developing analysis strategies and tactics for such networks (Wigand 1988).

### COLLECTIVE BEHAVIOR

While diffusion is not a commonly used term in collective behavior, processes of diffusion are important in connection with understanding crowds, fashion, and some aspects of disaster behavior. In all cases, analytic concern centers on the dissemination of emotions, social practices, or physical elements through a collectivity. The study of human behavior in disasters is recent and multidisciplinary. In this field there has been a concern with diffusion in the classic sense of tracking ideas and practices through networks. The principle foci of research have been the adoption of protective measures and the dissemination of warning messages (Lindell and Perry 1992), with the aim of research being both the development of general theories of protective behaviors and more effective protection of endangered populations.

All three diffusion theory traditions converge in the study of crowd behavior. In proposing *imitation* as an explanatory mechanism for crowd actions, Gabriel Tarde (1890, p. 45) drew upon Edward Tylor’s concept of cultural diffusion. Subsequently, Gustave LeBon (1895) and Gabriel Tarde

(1901) approached crowd behavior in terms of *social contagion*: rapid dissemination of emotions among interacting people. Although Floyd Allport (1924) and Herbert Blumer (1939) extended and formalized the concept of contagion, it has been largely displaced as a theory of crowd behavior by convergence theory (Turner and Killian 1987, p.19).

Changes in dress have been conceptualized as diffusion processes. Alfred Kroeber (1919) studied *fashion cycles* which he believed diffused systematically through civilizations. Katz and Lazarsfeld (1955, p. 241) moved away from the initial concern with movement of fashion through networks to focus more on social influence. Herbert Blumer (1969) firmly established social psychological motivations as the basis for fashion behavior. Current theoretical work on fashion continues to emphasize social psychological approaches wherein fashion diffusion issues are peripheral (Davis 1985; Nagasawa, Hutton, and Kaiser 1991).

REFERENCES

- Adams, Tyrone 1997 "Using Diffusion of Innovations Theory to Enrich Virtual Organizations in Cyberspace." *Southern Communication Journal* 62:133-141.
- Allport, Floyd 1924 *Social Psychology*. Boston: Houghton-Mifflin.
- Blumer, Herbert 1939 "Collective Behavior." In Robert Park, ed., *Outline of the Principles of Sociology*. New York: Barnes and Noble.
- 1969 "Fashion." *Sociological Quarterly* 10:275-291.
- Boas, Franz 1896 *The Limitations of the Comparative Method of Anthropology*. New York: Alfred Knopf.
- Bowers, Raymond 1937 "The Direction of Intra-Societal Diffusion." *American Sociological Review* 2:826-836.
- Carlson, Richard 1965 *Adoption of Educational Innovations*. Eugene: University of Oregon Center for the Advanced Study of Educational Administration.
- Chapin, F. Stuart 1928 *Cultural Change*. New York: Century.
- Coleman, James, Herbert Menzel, and Elihu Katz 1957 "The Diffusion of an Innovation among Physicians." *Sociometry* 20:253-270.
- 1966 *Medical Innovation: A Diffusion Study*. New York: Bobbs-Merrill.
- Crane, Diana 1972 *Invisible Colleges*. Chicago: University of Chicago Press.
- Davis, Fred 1985 "Clothing and Fashion as Communication." *Symbolic Interaction* 5:111-126.
- Dean, John 1997 "The Diffusion of American Culture in Western Europe since World War Two." *Journal of American Culture* 20:11-32.
- Deutschmann, P. J. , and W. A. Danielson 1960 "Diffusion of Knowledge of the Major News Story." *Journalism Quarterly* 37:345-355.
- Fliegel, Frederick 1993 *Diffusion Research in Rural Sociology*. Westport, Conn.: Greenwood.
- Graebner, Fritz 1911 *Methode der Ethnologie*. Heidelberg, Germany: Carl Winter.
- Griswold, William 1987 "A Methodological Framework for the Sociology of Culture." *Sociological Methodology* 14:1-35.
- Hightower, James 1972 *Hard Tomatoes, Hard Times*. Cambridge, Mass.: Shenkman.
- Howes, David 1996 *Cross-Cultural Consumption*. New York: Routledge.
- Katz, Elihu, and Paul Lazarsfeld 1955 *Personal Influence*. New York: Free Press.
- Katz, Elihu, M. L. Levine, and H. Hamilton 1963 "Traditions of Research on Diffusion of Innovation." *American Sociological Review* 28:237-253.
- Kroeber, Alfred 1919 "On the Principle of Order in Civilization as Exemplified by Changes of Fashion." *American Anthropologist* 21:235-263.
- 1923 *Anthropology*. New York: Harcourt.
- LeBon, Gustave 1895 *Psychologie des foules*. Paris: Alcan.
- Lindell, Michael, and Ronald Perry 1992 *Behavioral Foundations of Community Emergency Planning*. Washington, D.C.: Hemisphere Publishers.
- Lowie, Robert 1937 *The History of Ethnological Theory*. New York: Rinehart and Company.
- McQuail, Denis 1983 *Mass Communication Theory*. Beverly Hills, Calif.: Sage.
- Menzel, Herbert, and Elihu Katz 1955 "Social Relations and Innovation in the Medical Profession." *Public Opinion Quarterly* 19:337-352.
- Nagasawa, Richard, Sandra Hutton, and Susan Kaiser 1991 "A Paradigm for the Study of the Social Meaning of Clothes." *Clothing and Textiles Research Journal* 10:53-62.
- Park, Robert, and Ernest Burgess 1921 *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.
- Posnansky, M., and C. R. DeCorse 1986 "Historical Archaeology in Sub-Saharan Africa." *Historical Archaeology* 20:1-14.
- Rogers, Everett 1962, 1995 *Diffusion of Innovations*. New York: Free Press.

- 1988 “The Intellectual Foundation and History of the Agricultural Extension Model.” *Knowledge* 9:492–510.
- 1994 *A History of Communication Study*. New York: Free Press.
- , and Floyd Shoemaker 1971 *Communication of Innovations*. New York: Free Press.
- Ryan, Bryce, and Neal Gross 1943 “The Diffusion of Hybrid Seed Corn in Two Iowa Communities.” *Rural Sociology* 8:15–24.
- Sahal, Devendra 1981 *Patterns of Technological Innovation*. Reading, Mass.: Addison-Wesley.
- Stahl, Ann 1994 “Innovation, Diffusion and Culture Contact.” *Journal of World Prehistory* 8:51–112.
- Tarde, Gabriel 1890 *Les Lois de l'imitation*. Paris: Alcan.
- 1901 *L'opinion et la foule*. Paris: Alcan.
- Turner, Ralph, and Lewis Killian 1987 *Collective Behavior*. Englewood Cliffs, N.J.: Prentice-Hall.
- Tylor, Edward 1865 *Early History of Mankind and the Development of Civilisation*. London: John Murray.
- Valente, Thomas 1995 *Network Models of the Diffusion of Innovations*. Cresskill, N.J.: Hampton.
- , and Everett Rogers 1995 “The Origins and Development of the Diffusion of Innovations Paradigm.” *Science Communication* 16:242–273.
- Wigand, Rolf 1988 “Communication Network Analysis.” In G. Goldhaber and G. Barnett, eds., *Handbook of Organizational Communication*. Norwood, N.J.: Ablex Publishing.
- Wissler, Clark 1929 *An Introduction to Social Anthropology*. New York: Henry Holt.
- Wuthnow, Robert, and Marsha Witten 1988 “New Directions in the Study of Culture.” *Annual Review of Sociology* 14:49–67.

RONALD W. PERRY

## DISASTER RESEARCH

### SOCIOHISTORY OF THE FIELD

Descriptions of calamities exist as far back as the earliest human writings, but systematic empirical studies and theoretical treatises on social features of disasters appeared only in the twentieth century. The first major publications in both instances were produced by sociologists. Samuel Prince (1920) wrote a doctoral dissertation in sociology at Columbia University that examined the social

change consequences of a munitions ship explosion in the harbor of Halifax, Canada. Pitirim Sorokin (1942) two decades later wrote *Man and Society in Calamity* that mostly speculated on how war, revolution, famine, and pestilence might affect the mental processes and behaviors, as well as the social organizations and the cultural aspects of impacted populations. However, there was no building on these pioneering efforts.

It was not until the early 1950s that disaster studies started to show any continuity and the accumulation of a knowledge base. Military interest in possible American civilian reactions to post-World War II threats from nuclear and biological warfare led to support of academic research on peacetime disasters. The National Opinion Research Center (NORC) undertook the key project at the University of Chicago between 1950 and 1954. This work, intended to be multidisciplinary, became dominated by sociologists as were other concurrent studies at the University of Oklahoma, Michigan State, and the University of Texas (Quarantelli 1987, 1994). The NORC study not only promoted field research as the major way for gathering data, but also brought sociological ideas from the literature on collective behavior and notions of organizational structure and functions into the thinking of disaster researchers (Dynes 1988; Dynes and Tierney 1994).

While the military interest quickly waned, research in the area obtained a strategic point of salience and support when the U.S. National Academy of Sciences in the late 1950s created the Disaster Research Group (DRG). Operationally run by sociologists using the NORC work as a prototype, the DRG supported field research of others as well as conducted its own studies (Fritz 1961). When the DRG was phased out in 1963, the Disaster Research Center (DRC) was established at the Ohio State University. DRC helped the field of study to become institutionalized by its continuous existence to the present day (having moved to the University of Delaware in 1985). In its thirty-six years of existence DRC has trained dozens of graduate students, built the largest specialized library in the world on social aspects of disasters, produced over six hundred publications and about three dozen Ph. D. dissertations (see <http://www.udel.edu/DRC/homepage.htm>), continually and consciously applied a sociological perspective to new disaster research topics, initiated an

interactive computer net of researchers in the area, and intentionally helped to create international networks and critical masses of disaster researchers.

DRC was joined in time in the United States by two other major social science research centers (both currently headed by sociologists). The Natural Hazards Center at the University of Colorado has as part of its prime mission the linking of disaster researchers and research-users in policy and operational areas. The Hazards Reduction and Recovery Center at Texas A & M University has a strong multidisciplinary orientation. The organization of these groups and others studying disasters partly reflects the fact that the sociological work in the area was joined in the late 1960s by geographers with interest in natural hazards (Cutter 1994), in the 1980s by risk analysts (including sociologists such as Perrow 1984; Short 1984) especially concerned with technological threats, and later by political scientists who initially were interested in political crises (Rosenthal and Kouzmin 1993). More important, in the 1980s disaster research spread around the world, which led to the development of a critical mass of researchers. This culminated in 1986 in the establishment within the International Sociological Association of the Research Committee on Disasters (# 39) (<http://sociweb.tamu.edu/ircd/>), with membership in over thirty countries; its own professional journal, *The International Journal of Mass Emergencies and Disasters* ([www.usc.edu/dept/sppd/ijmed](http://www.usc.edu/dept/sppd/ijmed)); and a newsletter, *Unscheduled Events*. At the 1998 World Congress of Sociology, this committee organized fourteen separate sessions with more than seventy-five papers from several dozen countries. The range of papers reflected that the initial focus on emergency time behavior has broadened to include studies on mitigation and prevention, as well as recovery and reconstruction.

### CONCEPTUALIZATION OF “DISASTER”

Conceptualizations of “disaster” have slowly evolved from employing everyday usages of the term, through a focus on social aspects, to attempts to set forth more sociological characterizations. The earliest definitions equated disasters with features of physical agents and made distinctions between “acts of God” and “technological” agents. This

view was followed by notions of disasters as phenomena that resulted in significant disruptions of social life, which, however, might not involve a physical agent of any kind (e.g., a false rumor might evoke the same kind of evacuation behavior that an actual threat would). Later, disasters came to be seen as crises resulting either from certain social constructions of reality, or from the application of politically driven definitions, rather than necessarily from one initial and actual social disruption of a social system. Other researchers equated disasters with occasions where the demand for emergency actions by community organizations exceeds their capabilities for response. By the late 1980s, disasters were being seen as overt manifestations of latent societal vulnerabilities, basically of weaknesses in social structures or systems (Schorr 1987; Kreps 1989).

Given these variants about the concept, it is not surprising that currently no one formulation is totally accepted within the disaster research community (see Quarantelli 1998 where it is noted that postmodernistic ideas are now also being applied). However, there would be considerable agreement that the following is what is involved in using the term “disaster” as a sensitizing concept: Disasters are relatively sudden occasions when, because of perceived threats, the routines of collective social units are seriously disrupted and when unplanned courses of action have to be undertaken to cope with the crisis.

The notion of “relatively sudden occasions” indicates that disasters have unexpected life histories that can be designated in social space and time. Disasters involve the perceptions of dangers and risks to valued social objects, especially people and property. The idea of disruption of routines indicates that everyday adjustive social mechanisms cannot cope with the perceived new threats. Disasters necessitate the emergence of new behaviors not available in the standard repertoire of the endangered collectivity, a community, which is usually the lowest social-level entity accepted by researchers as able to have a disaster. In the process of the refinement of the concept, sociologists have almost totally abandoned the distinction between “natural” and “technological” disasters, derived from earlier notions of “acts of God” and “man-made” happenings. Any disaster is seen as inherently social in nature in origin, manifestation, or consequences. However, there is lack of

consensus on whether social happenings involving intentional, deliberate human actions to produce social disruptions such as occur in riots, civil disturbances, terrorist attacks, product tampering or sabotage, or wars, should be conceptualized as disasters. The majority who oppose their inclusion argue that conflict situations are inherently different in their origins, careers, manifestations, and consequences. They note that in disaster occasions there are no conscious attempts to bring about negative effects as there are in conflict situations (Quarantelli 1993). Nevertheless, there is general agreement that both conflict- and consensus-type crises are part of a more general category of collective stress situations, as first suggested by Allan Barton (1969).

### MAJOR RESEARCH FINDINGS

While the research efforts have been uneven, much has been learned about the behavior of individuals and households, organizations, communities, and societies in the pre-, trans-, and postimpact time periods (Quarantelli and Dynes 1977; Kreps 1984; Drabek 1986). A separation of the disaster-planning cycle into mitigation, preparedness, response, and recovery phases has won partial acceptance at some policy and operational levels in the United States. However, international usage of the terms is far from total and there also is disagreement regarding what should be considered under mitigation. Therefore we will continue to discuss findings under the older "time" breakdown.

**Preimpact behavior.** *Individuals and households.* Most residents show little concern about disasters before they happen, even in risk-prone areas and where threats are recognized. Citizens tend to see disaster planning as primarily a moral even more than a legal responsibility of the government. Very few households ever plan in any concrete way for possible disasters. Exceptions to these passive attitudes are where there are many recurrent experiences of disasters as occur in some localities, where disaster subcultures (institutionalized expectations) have developed, and where potential disaster settings (such as at chemical complexes or nuclear plants) are the focus of activist citizen groups.

*Organizations.* Except for some disaster-oriented groups such as police and fire departments,

there usually is little organizational planning for disasters. Even agencies that plan tend to think of disasters as extensions of everyday emergencies and fail, according to researchers, to recognize the qualitative as well as quantitative differences between routine crises and disaster occasions. In disasters the responding organizations have to quickly relate to more and different groups than normal, adjust to losing part of their autonomy to overall coordinating groups, apply different performance standards and criteria, operate within a closer-than-usual public and private interface, and cannot function well when their own facilities and operations may be directly impacted by the disaster agent.

*Communities.* Low priority usually is given to preparing localities for disasters, and when there is some effort it is usually independent of general community development and planning. This reflects the reactive rather than proactive orientation of most politicians and bureaucrats and the fact that the issue of planning very seldom becomes a matter of broad community interest as would be indicated by mass media focus, discussions in the political arena, or the existence of advocacy groups. Efforts to initiate overall disaster preparedness often are hindered by prior organizational and community conflicts, cleavages, and disputes. Starting in the 1990s, major systematic efforts from the top down have been made in a few countries to push for the implementation of local mitigation measures. For programs to be implemented, however, people must accept the realistic criticism that while a disaster may be a high-impact, it is a very low-probability event.

*Societies.* Generally, disaster planning does not rank very high on the agenda of most societies. However, increasingly there are exceptions in developing countries when major recurrent disasters have had major impact on the gross national product and developmental programs. Also, in developed societies certain even distant catastrophes such as a Bhopal or Chernobyl can become symbolic occasions that lend impetus to instituting preparedness measures for specific disaster agents. Increasingly too, attention to national-level disaster planning has increased as citizens have come to expect their governments to provide more security in general for the population. Also, mitigation or prevention of disasters is being given higher priority than in the past.

**Transemergency Period Behavior.** *Individuals and households.* When disasters occur, individuals generally react very well. They are not paralyzed by a threat but actively seek relevant information and attempt to do what they can in the emergency. Victims while usually very frightened, not only act positively but also show little deviant behavior; they extremely seldom break in panic flight; they do not act irrationally especially from their perspective; and they very rarely engage in antisocial activities, although stories of such contrary behavior as looting may circulate very widely. Prosocial behavior especially comes to the fore, with the initial search and rescue being undertaken quickly and mostly by survivors in the immediate area. Most sudden needs, such as emergency housing, are met by kin and friends rather than official relief agencies. Family and household relationships are very important in affecting immediate reactions to disasters such as whether evacuation will occur or if warnings will be taken seriously, because mass media reports are filtered through primary ties.

*Organizations.* As a whole, organizations do not react as well to disasters as do individuals and households. But while there are many organizational problems in coping with the emergency time demands of a disaster, these difficulties are often not the expected ones. Often it is assumed that if there has been organizational disaster planning, there will be successful crisis or emergency management. But apart from the possibility of planning being poor in the first place, planning is not management and the former does not always translate well into the latter in community disasters. There typically are problems in intra- and interorganizational information flow, and in communication between and to organizations and the general public. Groups initially often have to struggle with major gaps in knowledge about the impacts of a disaster. There can be organizational problems in the exercise of authority and decision making. These can stem from losses of higher-echelon personnel because of overwork, conflict regarding authority over new disaster tasks, and clashes over organizational jurisdictional differences. Generally, there is much decentralization of organizational response which in most cases is highly functional. Organizations operating with a command and control model of response do not do well at emergency times. There often too are

problems associated with strained organizational relationships created by new disaster tasks and by the magnitude of a disaster impact.

*Communities.* Since disasters almost always cut across formal governmental boundaries, problems of coordination among different impacted political entities are all but inevitable. The greater the impact of a disaster, the more there will be the emergence of new and adaptive community structures and functions, especially emergent groups (that is, those without any preimpact existence). The greater the disaster also, the more organized improvisations of all kinds appear accompanied by pluralistic decision making. In addition, the mass convergence of outside but nonimpacted personnel and resources on impacted communities, while functional in some ways, creates major coordination problems.

*Societies.* Few societies ignore major disasters, but this sometimes occurs especially in the case of slow and diffuse occasions such as droughts and famines, especially if they primarily affect subgroups not in the mainstream of a developing country. In responding to domestic disasters, typically massive help is provided to impacted areas. Increasingly, most societies, including governmental officials at all levels, obtain their view of what is happening in their disasters from mass media accounts (what has been called the “CNN syndrome”); this also affects what is often remembered about the occasions. There is also a spreading belief, so far unsupported by research, that new technologies—especially computer-related ones—will allow major improvements in disaster planning and management.

**Postimpact Behavior.** *Individuals and households.* Overall, there is little personal learning as a result of undergoing a single disaster. While the experience of a major disaster is a memorable one from a social-psychological point of view, there are seldom lasting and widespread negative behavioral consequences. Disasters very seldom produce new psychoses or severe mental illnesses. They often, but not always, generate subclinical, short-lived and self-remitting surface reactions, such as loss of appetite, sleeplessness, and anxiety. More common are many problems in living that stem more from inefficient and ineffective relief and recovery efforts of helping organizations rather



than from the direct physical impacts of disasters. However, not all postimpact effects are negative; sometimes, the experience of undergoing a disaster results in positive self-images and closer social ties among victims.

*Organizations.* Organizational changes, whether for planning for disasters or for other purposes in the postimpact period, is not common and selective at best. Most modifications are simply accelerations of noncrisis-related changes already planned or underway. Postimpact discussion of how to improve disaster planning seldom gets translated into concrete actions (unlike civil disturbances which in American society in the 1960s led to many changes in organizations). However, overall, both in the United States and elsewhere, there have been in recent decades the growth of small, locally based, formal social groups primarily concerned with emergency time disaster planning and management. Partly because of seemingly constantly escalating economic losses, certain businesses in such sectors as banking and insurance have increasingly become interested in disaster preparedness and recovery.

*Communities.* There are selective longer-run outcomes and changes in communities that have been impacted by disasters. There can be acceleration of some ongoing and functional community trends (e.g., in local governmental arrangements and power structures), and generation of some limited new patterns (e.g., in providing local mental-health services or some mitigation measures such as floodproofing regulations). On the other hand, particularly as the result of rehousing and rebuilding, there can be magnifications of preimpact community conflicts as well as the generation of new ones; some of the latter is manifested in blame assignation, which, however, tends to deflect attention away from social structural flaws to mass-media-influenced search for individual scapegoats. It is also being recognized after disasters that changes in technology that create diffuse networks and systems, such as among lifeline organizations, are increasingly creating the need for regional rather than just community-based disaster planning.

*Societies.* In developed societies, there are few long-run negative consequences of disaster losses whether of people or property, since such effects are absorbed by the larger system. In developing

societies and very small countries, this is not necessarily true; a catastrophic disaster may reduce the gross national product five to ten percent as well as producing tens of thousands of casualties. Nevertheless, changes or improvement in national disaster planning often does not occur except in certain cases such as after the Mexico City earthquake where an unusual set of circumstances existed, including a "political will" to do something. But increasingly, in the aftermath of major disasters, to the extent that planning is instituted or improved, it is being linked to developmental planning, a move strongly supported by international agencies such as the World Bank.

#### UNIVERSALITY OF GENERALIZATIONS FROM AN UNEVEN RESEARCH BASE

Cross-societal and comparative research increased markedly in the 1990s. Studies have ranged from cooperative work on local mass-media reporting of community disasters in Japan and the United States (Mikami, Hiroi, Quarantelli, and Wenger 1992) and flood responses and crisis management in four Western European countries (Rosenthal and Hart 1998), to comparisons of perceptions of recurrent floods in Bangladesh by European engineers and local residents (Schmuck-Widmann 1996), and cross-national analyses of post-disaster political unrest in a dozen countries (Olson and Drury 1997), as well as methodological issues involved in cross-societal research in Italy, Mexico, Turkey, Peru, the United States, and Yugoslavia (Bates and Peacock 1993). However, this kind of comparative empirical research so far has been limited. Furthermore, although the bulk of disasters occur in developing countries, the majority of studies from which the generalizations advanced have been derived, have been done in developed societies. Thus, the question of the universality of disaster behaviors in different social systems has increasingly been raised. Some universals appear to have been found: Prosocial rather than antisocial behavior clearly predominates in responses everywhere; household members and significant others are crucial in validating warning messages, and the larger kin system is vital in providing emergency assistance; emergent groups always appear at the height of the crisis period; organizations have relatively more difficulty in adjusting to and coping with disasters than do individuals and

small groups; the disaster recovery period is fraught with problems at the household, organization, and community levels; mitigation measures are given little priority even in disaster-prone localities; and social change is seldom an outcome of most disasters.

Generalizations of a more limited nature also seem to exist. There are social-system-structure-specific behaviors. For example, there often is a major delay in the response to catastrophic disasters from centralized, compared to decentralized, governmental systems. There also may be culturally specific differences. For example, reflecting cultural values, individual volunteers in disasters very rarely appear in some societies such as Japan whereas they are typical in almost all American disasters.

### THE FUTURE

There is a dialectical process at work: There will be more and worse disasters at the same time that there will be more and better planning. Why more and worse disasters? Risks and threats to human beings and their societies are increasing. Traditional natural-disaster agents, such as earthquakes and floods, will simply have more to impact as the result of normal population growth and higher, denser concentration of inhabitants in risk-prone localities, such as floodplains or hurricane-vulnerable shorelines that otherwise are attractive for human occupancy. There is an escalating increase in certain kinds of technological accidents and mishaps in the chemical, nuclear, and hazardous-waste areas that were almost unknown before World War II. There are technological advances that create risks and complexities to old threats such as when fires are prevented in high-rise buildings by constructing them with highly toxic materials, or when the removal of hazardous substances from solid sewage waste generates products that contain dangerous viruses and gases. New versions of old threats are also appearing, such as the increasing probability of urban rather than rural droughts, or the potential large-scale collapse of the infrastructure of older metropolitan area life-line systems. Finally, there is the continual development of newer kinds of risks ranging from the biological threats that are inherent in genetic engineering, to the crises that will be generated as the

world increasingly becomes dependent on computers that are bound to fail somewhere at some key point, with drastic consequences for social systems. In addition, the newer threats are frequently dangerous at places and times distant from their initial source or origin as dramatized by the Chernobyl nuclear radiation fallout in European countries and smog pollution episodes such as forest fires in Indonesia which had negative effects in many Southeast Asian countries.

On the other hand, there is increasing concern and attention being paid to disaster planning of all kind. The future augers well for more and better planning. Citizens almost everywhere are coming to expect that their governments will take steps to protect them against disasters; this is often actualized in planning for emergency preparedness and response. Whereas two decades ago a number of societies had no preimpact disaster planning of any kind, this is no longer the case. A symbolic manifestation of this trend was the proclamation by the United Nations of the 1990s as The Decade for Natural-Disaster Reduction. This international attention accelerated efforts to prevent, prepare for, respond to, and recover from disasters. The effect was especially notable in the increased disaster planning in developed countries.

In developed societies, a focus on disaster planning and crisis management had started earlier, partly as a result of sociological and related research. By the 1980s, social scientists were increasingly influencing policies, political agenda settings, and operational matters regarding disasters. This can be seen in a variety of ways. Social scientists were represented on almost all national committees set up for the U. N. Decade, and contributed significantly to the reports prepared to mark the midpoint of the decade. The Board on Natural Disasters, established in 1992 in the U.S. National Academy of Sciences, always has had members from sociology and related disciplines. Sociologists have had major roles in national-disaster legislation in Greece and Italy. American social science disaster researchers typically testify before state and congressional committees considering disaster-related laws and policies. Many sociological disaster researchers provide both paid and unpaid consultant services to international, national, and local public and private groups involved in disaster-related activities. In places such

as the United States, Australia, New Zealand, and some European countries, the emergency-management and disaster-planning community has become more open to recognizing the practical implications of social science research.

However, a caveat is in order for the generally correct view that studies by sociologists and others have increasingly influenced policy decisions and operational activities in the disaster area. Research results and questions are sometimes more than counterbalanced by other factors: These include the vested interests of powerful professional and bureaucratic elites to maintain traditional stances, resistance to seriously questioning the viabilities and competencies of specific organizations, and an unwillingness to face up to false assumptions of some cultural beliefs and values. Research to the contrary, for example, has had little effect on the fad-like spread of the "Incident Command System" as a model for the emergency time operations of organizations, or an ever-spreading acceptance that victims are likely to suffer posttraumatic stress disorders, or the current common belief that mitigation measures are necessarily a better strategy for disaster planning than giving priority to improving resilience and response to crises.

### RELATIONSHIP TO SOCIOLOGY

Although not true everywhere, sociologists have been increasingly accepted as having an important contribution to make to disaster planning and management. In part this stems from the fact that in many countries such as Germany, Italy, Russia, and the United States, they have played the lead role among social scientists in undertaking disaster studies. While many reasons account for this, probably the crucial factor has been that much in general sociology can be used in doing research in the area.

There has been a close relationship between disaster studies and sociology from the earliest days of work in the area. In part this is because sociologists, being among the leading pioneers and researchers in the area, have tended to use what they could from their discipline. Thus, sociology has contributed to the research techniques used (e.g., field studies and open-ended interviewing), the research methodology utilized (e.g., the

"grounded theory" approach and the employment of inductive analytical models), the theoretical ideas used (e.g., the notion of emergence from collective-behavior thinking and the idea of informal and formal structures of organizations), and the general perspectives employed (e.g., that there can be latent as well as dysfunctional aspects of any behavior and that societies and communities have a social history that is not easily set aside). In the volume entitled *Sociology of Disasters: Contributions of Sociology to Disaster Research* (Dynes, De Marchi, and Pelanda 1987), these and other contributions to disaster theory, disaster research methods, disaster models, and disaster concepts are set forth in considerable detail.

The relationship has not been one-sided, since disaster research has also contributed to sociology. The field of collective behavior has been most influenced and this has been explicitly noted (Wenger 1987). Other significant contributions include the study of formal organizations, social roles, social problems, organizational and social change, mass communications, medical sociology, and the urban community (see Drabek 1986; Dynes, De Marchi, and Pelanda 1987; Dynes and Tierney 1994). A symposium on social structure and disaster, coattended by disaster researchers and prominent sociological theorists, examined how disaster studies not only are informed by but could also inform sociological theory; the proceedings were published in *Social Structure and Disaster* (Kreps 1989). It is also perhaps of interest that for several decades now, many introductory sociology textbooks have a section on disaster behavior, usually in the collective behavior chapter.

### REFERENCES

- Barton, Allan 1969 *Communities in Disaster: A Sociological Analysis*. Garden City, N.Y.: Anchor.
- Bates, F., and W. Peacock 1993 *Living Conditions, Disasters and Development: An Approach to Cross-Cultural Comparisons*. Athens, Ga.: University of Georgia Press.
- Cutter, Susan, ed. 1994 *Environmental Risks and Hazards*. Englewood Cliffs, N.J.: Prentice-Hall.
- Drabek, Thomas 1986 *Human System Responses to Disasters: An Inventory of Sociological Findings*. New York: Springer Verlag.
- Dynes, R. R., ed. 1988 "Disaster Classics Special Issue." *International Journal of Mass Emergencies and Disasters* 6:209-395.

- , and K. Tierney, eds. 1994 *Disasters, Collective Behavior, and Social Organization*. Newark, Del.: University of Delaware Press.
- Dynes, R. R., B. De Marchi, and C. Pelanda (eds.) 1987 *Sociology of Disasters: Contributions of Sociology to Disaster Research*. Milan: Franco Angeli.
- Fritz, C. 1961 "Disasters." In R. K. Merton and R. A. Nisbet, eds., *Contemporary Social Problems*. New York: Harcourt.
- Kreps, G. 1984 "Sociological Inquiry and Disaster Research." *Annual Review of Sociology* 10:309–330.
- 1989 *Social Structure and Disaster*. Newark, Del.: University of Delaware Press.
- Mikami, S., O. Hiroi, E. L. Quarantelli, and D. Wenger 1992 "A Cross-Cultural Comparative Study of Mass Communication in Disasters." *Bulletin of Faculty of Sociology, Toyo University* 29:59–202.
- Olson, Richard, and A. Cooper Drury 1997 "Un-Therapeutic Communities: A Cross-National Analysis of Post-Disaster Political Unrest." *International Journal of Mass Emergencies and Disasters* 15:221–238.
- Perrow, C. 1984 *Normal Accidents: Living With High-Risk Technologies*. New York: Basic Books.
- Prince, S. 1920 *Catastrophe and Social Change*. New York: Columbia University Press.
- Quarantelli, E. L. 1987 "Disaster Studies: An Analysis of the Social Historical Factors Affecting the Development of Research in the Area." *International Journal of Mass Emergencies and Disasters* 5:285–310.
- 1993 "Community Crises: An Exploratory Comparison of the Characteristics and Consequences of Disasters and Riots." *Journal of Contingencies and Crisis Management* 1:67–78.
- 1994 "Disaster Studies: The Consequences of the Historical Use of a Sociological Approach in the Development of Research." *International Journal of Mass Emergencies and Disasters* 12:25–49.
- , ed. 1998 *What is a Disaster? Perspectives on the Question*. London: Routledge.
- , and R. R. Dynes 1977 "Response to Social Crisis and Disaster." *Annual Review of Sociology* 3:23–49.
- Rosenthal, U., and P. Hart, eds. 1998 *Flood Response and Crisis Management in Western Europe: A Comparative Analysis*. New York: Springer Verlag.
- Rosenthal, U., and A. Kouzmin 1993 "Globalizing an Agenda for Contingencies and Crisis Management." *Journal of Contingencies and Crisis Management* 1:1–12.
- Schmuck-Widmann, Hanna 1996 *Living with the Floods: Survival Strategies of Char-Dwellers in Bangladesh*. Berlin: FDCL.
- Schorr, J. 1987 "Some Contributions of German Katastrophensoziologie to the Sociology of Disaster." *International Journal of Mass Emergencies and Disasters* 5:115–135.
- Short, James F. 1984 "The Social Fabric at Risk: Toward the Social Transformation of Risk Analysis." *American Sociological Review* 49:711–725.
- Sorokin, Pitirim 1942 *Man and Society in Calamity*. New York: Dutton.
- Wenger, Dennis 1987 "Collective Behavior and Disaster Research." In R. R. Dynes, B. De Marchi, and C. Pelanda, eds., *Sociology of Disasters: Contribution of Sociology to Disaster Research*, 213–238. Milan: Franco Angeli.

E. L. QUARANTELLI

## DISCRIMINATION

Discrimination, in its sociological meaning, involves highly complex social processes. The term derives from the Latin *discriminatio*, which means to perceive distinctions among phenomena or to be selective in one's judgment. Cognitive psychology retains the first of these meanings, popular usage the second. Individual behavior that limits the opportunities of a particular group is encompassed in many sociological considerations of discrimination. But exclusively individualistic approaches are too narrow for robust sociological treatment. Instead, sociologists understand discrimination not as isolated individual acts, but as a complex system of social relations that produces intergroup inequities in social outcomes.

This definitional expansion transforms "discrimination" into a truly sociological concept. But in its breadth, the sociological definition leaves room for ambiguity and controversy. Obstacles to consensus on what constitutes discrimination stem from two sources—one empirical, the other ideological and political. First, deficiencies in analysis and evidence limit our ability to trace thoroughly the dynamic web of effects produced by discrimination. Second, because social discrimination is contrary to professed national values and law, a judgment that unequal outcomes reflect discrimination is a call for costly remedies. Variable willingness to bear those social costs contributes to dissension about the extent of discrimination.

The broadest sociological definitions of discrimination assume that racial minorities, women, and other historical target groups have no inherent characteristics that warrant inferior social outcomes. Thus, all inequality is seen as a legacy of discrimination and a social injustice to be remedied. By contrast, political conservatives favor a far narrower definition that limits the concept's scope by including only actions *intended* to restrict a group's chances. For solid conceptual reasons, sociologists seldom follow suit (but see Burkey 1978, p. 79). First, an intentionality criterion returns the concept to the realm of psychology and deflects attention from restraining social structure. Second, the invisibility of intentions creates insuperable obstacles to documenting discrimination.

### DIRECT AND INDIRECT DISCRIMINATION

Sociologists' understanding of intricate societal patterns sensitizes them to the fact that disadvantage accruing from intentional discrimination typically cumulates, extends far beyond the original injury, and long outlives the deliberate perpetration. Many sociologists distinguish between *direct* and *indirect* discrimination (Pettigrew 1985). Direct discrimination occurs at points where inequality is generated, often intentionally. When decisions are based explicitly on race, discrimination is direct. Indirect discrimination is the perpetuation or magnification of the original injury. It occurs when the inequitable results of direct discrimination are used as a basis for later decisions ("past-in-present discrimination"), or decisions in linked institutions ("side-effect discrimination") (Feagin and Feagin 1986). Hence, discrimination is indirect when an ostensibly nonracial criterion serves as a *proxy* for race in determining social outcomes.

To illustrate with wages, direct discrimination exists when equally qualified blacks and whites or men and women are paid at different rates for the same work. Indirect discrimination exists when the two groups are paid unequally because *prior* discrimination in employment, education, or housing created apparent differences in qualifications or channeled the groups into better- and worse-paying jobs. This direct/indirect distinction resembles the legal distinction between disparate treatment and disparate impact. While intentional direct discrimination may have triggered the

causal chain, the original injury is often perpetuated and magnified by unwitting accomplices. Intentionality criteria deny that the continuing disadvantage is a legacy of discrimination.

Years ago, Williams outlined the concept differently: "Discrimination may be said to exist to the degree that individuals of a given group who are otherwise *formally qualified* are not treated in conformity with these nominally *universal institutionalized codes*." (Williams 1947, p. 39, italics added). For Antonovsky (1960, p. 81), discrimination involves "... injurious treatment of persons on grounds *rationally irrelevant* to the situation" (italics added). Economists use starker terms. Becker (1968, p. 81) held that economic discrimination occurs "... against members of a group whenever their earnings fall short of the amount '*warranted*' by their *abilities*'" (italics added).

Two problems arise with these definitions. First, the assessment of "abilities" and of what treatment is "rationally" relevant or "warranted" is no easy task. Critical examination of common practice uncovers many instances where formal qualifications and "nominally universal institutionalized codes" prove *not* to provide a logical basis for distinctions. Employment testing litigation demonstrates that when hiring criteria once legitimized by tradition or "logic" are put to scientific test, they often fail to predict job performance in the assumed fashion. Analogous fallacies have been identified in the conventional wisdom guiding admission to advanced education. Hence, nominally universalistic standards may provide an altogether illogical basis for decision making. If such misguided selection procedures also work to the disadvantage of historical victims of discrimination, these practices are not exempted from the charge of discrimination by their universalistic facade.

The second problem with these definitions is that they ignore another, prevalent form of indirect discrimination. Even where nominally universalistic standards do serve some legitimate social function, such as selecting competent workers, adverse impact of these standards on those who bear the cumulated disadvantage of historical discrimination cannot be disregarded.

The complexity of discrimination and unresolved issues about its definition impede easy application of social science methods to inform institutional policy. Apparently rigorous quantitative

analyses often only camouflage the crucial issues, as critical examination of wage differential decompositions reveals.

### THE DECOMPOSITION APPROACH

Assessments of discrimination produced by decomposing gross race or gender differences in wages or other social outcomes are common in sociology (e.g., Corcoran and Duncan 1978; Farley 1984, Rosenfeld and Kalleberg 1990) as well as in economics (e.g., Gill 1989). One segment of the gross intergroup differential is defined by its empirical linkage to “qualifications” and other factors deemed legitimate determinants of social rewards. The second, residual segment not demonstrably linked to “legitimate” determinants of the outcomes often is presented as the estimate of discrimination. However, in the absence of better information than usually available and greater agreement on what constitutes discrimination, no *unique* estimate is possible. Through their choice of control variables to index “legitimate” determinants of social outcomes and their interpretation of findings, researchers wittingly or unwittingly shape their answers. Any appearance of scientific certitude is an illusion. For example, estimated proportions of the gender earnings gap caused by discrimination in the United States range from Sanborn’s (1969) 10 percent or less to Blinder’s (1973) 100 percent. Predictably, each has been challenged (Bergman and Adelman 1973; Rosensweig and Morgan 1976).

**An Illustrative Decomposition Study.** An analysis of gender differentials in faculty salaries at a large university illustrates the difficulty of separating “legitimate” wage differentials from inequity (Taylor 1988). About 90 percent of a \$10,000 gender difference in faculty salaries was empirically linked to three factors widely considered legitimate determinants of faculty pay: academic rank, age, and discipline. Women tend to hold lower academic rank, to be younger, and more often to be affiliated with poorly-paid disciplines than men. Insofar as women’s lower salaries are linked to rank, age, and discipline, is the salary differential untainted by gender discrimination? Conventional wage differentials imply an unequivocal yes. If a simple answer is given, it should be no. But in truth, when policy makers ask for dollar estimates

of inequity that the institution is obliged to remedy, the answers are neither unequivocal nor simple.

If the university’s promotion system has operated fairly, a gender gap reflecting differences in rank may be warranted. If gender bias has existed in the university’s promotion system, depressing the average academic rank of women faculty, the resulting deficit in women’s salaries reflects indirect discrimination. Attention should then be directed to the offending promotion processes. However, direct salary adjustments may also be in order, because gender bias in promotions weakens the link between rank and merit. Inequitable depression of women’s ranks would not necessarily lessen their actual contributions to the faculty, just their status. Using rank as a universalistic determinate of salary would then undermine the goal this practice is claimed to promote—the matching of rewards to contributions.

Salary differences tied to the age differential of female and male faculty also raise troublesome questions. If the relative youth of women faculty reflects lower retention and higher turnover as a result of discriminatory review processes or generally inhospitable conditions, salary differentials tied to age differences are again examples of indirect discrimination. The evidence would signal a need for institutional efforts to improve the retention of women faculty. But here it is not clear that salary adjustments are warranted. Because faculty contributions may be a function of experience, application of the universalistic age criterion is arguably reasonable. Any gender gap in salary tied to age differentials could, then, be both a legacy of discrimination and a reasonable conditioning of rewards on contributions. Complicating the issue further is the fact that affirmative action efforts often meet with greatest success in recruiting junior candidates. Thus, without supplementary data, it is not even clear whether an age-linked gender gap in salary reflects continuing institutional discrimination or affirmative hiring.

Gender differences in salary associated with discipline present even more complicated interpretational problems. Women and men are distributed across academic disciplines in a fashion that mirrors the gender distribution across occupations. Disciplinary differences in average salary likewise mirror wage differentials across occupations. But are the differing occupational

distributions simply a matter of gender differences in preferences or abilities with no implication of discrimination? Or are women steered away from lucrative fields, so that gender differentials in salary linked to disciplinary affiliation represent indirect discrimination in training, recruitment, and hiring? Or does the pattern of occupational wage differentials mirrored in disciplinary differences represent direct discrimination, an influence of gender composition per se on occupational wage structures (England et al. 1988)? In the latter case, assignment of responsibility for remedy presents particular problems. The university, like any other single employer, is simultaneously vulnerable to competitive forces of the wider labor market and a constituent element of that market. Defiance of the market by a single organization is costly to that organization; adherence by all organizations to the broader occupational wage structure perpetuates gender inequity and carries broader costs.

This faculty salary study did not examine the role of scientific “productivity.” But the inclusion of productivity measures among the control variables raises further difficulties. On standard “productivity” measures, women faculty average lower scores than men (Fox 1990). Thus, an institutional study of salary differentials might find some segment of the male/female salary gap linked to productivity differences. Fox’s research demonstrates, however, that gender differences in scientific productivity reflect contrasting levels of resources that institutions provide to male and female faculty. Like age and rank, the gender difference in productivity may itself be a product of institutional discrimination. Thus, salary differentials based on male/female productivity differences also may represent indirect discrimination. The new ingredient here is that institutional shaping of productivity is subtle. Scientific productivity is ordinarily seen as an outgrowth of talent and effort, not potentially gender-biased institutional resource allocation. Solid documentation of this indirect discrimination process offers another challenge for researchers.

#### COMPLEXITIES OF DISCRIMINATION AND REMEDY

Critical reflection on this decomposition study highlights a set of interrelated points about the

complex nature of discrimination and unresolved issues of remedy.

1. In American society today, the injuries of indirect discrimination are often far more extensive than those of direct discrimination. This conclusion does not imply that direct discrimination no longer exists (Reskin 1998). The continued operation of direct forms of discrimination is indicated by employment complaint records. American women and minorities have filed almost 1.5 million job discrimination complaints since 1965 (Blumrosen 1996, p. 4). In 1994 alone, over 150,000 such complaints were filed; 91,000 to local and state agencies and 64,000 to the U.S. Equal Employment Opportunity Commission (Leonard 1994, p. 24). Of course, not all complaints reflect genuine discrimination; on the other hand, not all discrimination prompts formal complaints. Many major corporations were found guilty of direct race or gender discrimination in the 1990s. And employment audits using paired, equally qualified applicants reveal widespread direct discrimination (Reskin 1998, pp. 25–29).
2. Apparently reasonable universalistic principles may on closer examination be unnecessary or even dysfunctional. Scrutiny of employment criteria prompted by the Supreme Court’s 1971 *Griggs v. Duke Power Co.* decision has provided useful models for challenging nominally universalistic standards. Where it is possible to substitute standards that do as well or better at screening or evaluation without adversely affecting historical targets of discrimination, there are gains for all involved.
3. When gaps in actual qualifications are a legacy of discrimination, more extensive remedies are needed. Where training deficits impair employability, or inadequate preparation impedes admission to higher education, attention should be given to the earlier schooling processes that generated these deficiencies. This form of remedy aids future generations. In the meantime, compensatory training can

reduce the liabilities of those who have already fallen victim to inferior schools.

4. Microcosms cannot escape the discriminatory impact of the societal macrocosm. Just as salary differences across academic disciplines reflect general occupational wage structures, institutions are often both prey to and participant in broader social forces. Narrow, legalistic approaches to remedy are inadequate for addressing this dynamic of discrimination.
5. Empirical research on group discrimination must mirror the phenomenon in its variety and complexity. The regression decomposition approach has proven useful but has its limitations (see also Dempster 1988, and the ensuing commentary). Regression analyses could provide more pertinent information if based on more homogeneous job groups (Conway and Roberts 1994) and on structural equation models that test reciprocal causation. Most important, if the aim is to guide policy, a framework far more complex than the dichotomous discrimination-or-not approach is required. The sociological arsenal of methods offers other promising approaches. Research that traces the actual processes of institutional discrimination is essential (e.g., Braddock and McPartland 1987, 1989). Also needed is attention to victims' perceptions of discrimination (e.g., Feagin and Sikes 1994) and investigation of the changes generated by anti-discrimination efforts. Another approach involves cross-national comparative research, which we consider below.

#### EFFECTIVE REMEDIAL INTERVENTIONS

Direct racial and gender discrimination in the United States has declined in recent decades—more slowly in the 1980s and 1990s than in the 1960s and 1970s. But what caused this decline? Many factors were involved, but governmental intervention was an important impetus. For example, blacks in South Carolina made dramatic economic gains in manufacturing during the late 1960s. Heckman and Payner (1989) demonstrated that human capital, supply shifts, and tight labor

markets could not explain the sudden improvements. It was federal anti-discrimination programs, they concluded, that made a decisive contribution to the gains.

More general assessments also show that anti-discrimination legislation did reduce direct job discrimination nationally (Burstein 1985). It did not, however, eliminate the problem. Nor did such laws effectively attack indirect discrimination. For this more difficult problem, affirmative action programs were needed and have had some success (Reskin 1998). The resistance to such programs, however, underscores the difficulty of establishing effective remedies for the more subtle forms of discrimination.

#### DISCRIMINATION IN WESTERN EUROPE

Beyond racial and gender discrimination in the United States, the same basic concerns and principles arise for other nations and targets. Discrimination against Western Europe's new immigrant minorities is pervasive (Castles 1984; MacEwen 1995; Pettigrew 1998). Both direct and indirect discrimination are involved, though the indirect forms are largely unrecognized in Europe.

Investigators have repeatedly uncovered direct discrimination in England (Amin et al. 1988; Daniel 1968; Gordon and Klug 1984; Smith 1976). Controlled tests reveal the full litany of discriminatory forms involving employment, public accommodations, housing, the courts, insurance, banks, even car rentals. Employment discrimination poses the most serious problem. In every European Union nation, minorities have far higher unemployment rates than the majority group. In 1990 in the Netherlands, Moroccans and Turks had unemployment rates above 40 percent compared with the native Dutch rate of 13 percent (Pettigrew and Meertens 1996). During the 1974–1977 recession, West German manufacturing reduced its labor force by 765,000—42 percent of whom were foreign workers (Castles 1984, p. 148).

As in the United States, there are many reasons for minority unemployment disparities. The “last-in, first-out” principle selectively affects the younger minority workers. Typically less skilled, they are more affected by job upgrading. Minorities also are more likely to be in older, declining



industries. But these patterns are not accidental. Planners put minorities into these industries for cheaper labor precisely because of their decline. In addition, these multiple factors offer insufficient explanations for the greater unemployment of minorities. Veenman and Roelandt (1990) found that education, age, sex, region, and employment level explained only a small portion of the differential unemployment rates in the Netherlands.

Indirect discrimination arises when the inability to obtain citizenship restricts the opportunities of non-European Union minorities. It limits their ability to get suitable housing, employment, and schooling. A visa is necessary in order to travel to other European Union countries. In short, the lives of Europe's non-citizens are severely circumscribed (Wilpert 1993). Castles (1984) contends that the guest-worker system that brought many of the immigrants to Europe was itself a state-controlled system of institutional discrimination. It established the newcomers as a stigmatized "out-group" suitable for low-status jobs but not citizenship. Widespread indirect discrimination was inevitable for these victims of direct discrimination.

Anti-discrimination remediation has been largely ineffective in Europe. Basic rights in Germany are guaranteed only to citizens. So, the disadvantages of non-citizenship include limited means to combat discrimination. There is extensive German legislation to combat anti-Semitism and Nazi ideology, but these laws have proved difficult to apply to non-citizens. The German constitution explicitly forbids discrimination on the basis of origin, race, language, beliefs, or religion—but not citizenship. Indeed, the Federal Constitutional Court has ruled that differential treatment based on citizenship is constitutional if there is a reasonable basis for it and if it is not wholly arbitrary. Hence, a German court upheld higher taxes for foreign bar owners than German bar owners. And restaurants can refuse service to Turks and others on the grounds that their entry might lead to intergroup disturbances (Layton-Henry and Wilpert 1994).

Few means of combatting discrimination are available in France either. Commentators often view discrimination as "natural," as something universally triggered when a "threshold of tolerance" (*seuil de tolerance*) is surpassed (MacMaster

1991). Without supporting evidence, this rationalization supports quotas and dispersal policies that limit minority access to suitable housing.

The Netherlands, United Kingdom, and Sweden have enacted anti-discrimination legislation that specifically applies to the new immigrant minorities. And the Dutch have instituted modest affirmative action programs for women and minorities (De Vries and Pettigrew 1994). Not coincidentally, these countries make citizenship easier to obtain than Germany. Yet this legislation has been largely ineffective for two interrelated reasons. First, European legal systems do not allow class action suits—a forceful North American weapon to combat discrimination. Second, European efforts rely heavily on individual complaints rather than systemic remediation. Britain's 1976 Act gave the Commission for Racial Equality the power to cast a broad net, but individual complaints remain the chief tool (MacEwen 1995).

It is a sociological truism that individual efforts are unlikely to alter such systemic phenomena as discrimination. Mayhew (1968) showed how individual suits and complaints are largely non-strategic. Minorities bring few charges against the worst discriminators, because they avoid applying for jobs with them in the first place. Complaints about job promotion are common, but they are made against employers who hire minorities. Thus, effective anti-discrimination laws must provide broad powers to an enforcement agency to initiate strategic, institutionwide actions that uproot the structural foundations of discrimination. Mayhew's American analysis proves to be just as accurate in Western Europe.

## CONCLUSIONS

A comprehensive understanding of societal discrimination in both North America and Western Europe must encompass two propositions.

1. The long-lasting character of discrimination means that the effects typically outlive the initiators of discriminatory practices. Apart from its importance to the law, this feature of modern discrimination has critical implications for sociological theory. Discrimination is fundamentally normative; its structural web operates in large part independent of the dominant group's

present “tastes,” attitudes, or awareness. Hence, models based primarily on individual prejudice or “rationality,” whether psychological or economic, will uniformly understate and oversimplify the phenomenon.

2. Discrimination is typically cumulative and self-perpetuating. For example, an array of research on black Americans has demonstrated that neighborhood racial segregation leads to educational disadvantages, then to occupational disadvantage, and thus to income deficits (Pettigrew 1979, 1985). To be effective, structural remedies must reverse this “vicious circle” of discrimination. Affirmative action programs are one such remedy.

Seen in sociological perspective, then, discrimination is considerably more intricate and entrenched than commonly thought. The complexity of discrimination presents major challenges to social-scientific attempts to trace its impact. This complexity also precludes any one-to-one correspondence between perpetration and responsibility for remedy. Broad social programs will be necessary if the full legacy of direct and indirect discrimination is finally to be erased.

## REFERENCES

- Amin, K., M. Fernandes, and P. Gordon 1988 *Racism and Discrimination in Britain: A Select Bibliography, 1984–87*. London: Runnymede Trust.
- Antonovsky, A. 1960 “The Social Meaning of Discrimination.” *Phylon* 11:81–95.
- Becker, G. 1968 “Economic Discrimination.” In D. L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. IV. New York: Macmillan.
- Bergman, B. R., and I. Adelman 1973 “The 1973 Report of the President’s Council of Economic Advisors: The Economic Role of Women.” *American Economic Review* 63:509–514.
- Blinder, A. S. 1973 “Wage Discrimination: Reduced Form and Structural Estimates.” *Journal of Human Resources* 8:436–455.
- Blumrosen, A. W. 1996 *Declaration*. Statement submitted to the Supreme Court of California in response to Proposition 209 (Sept. 26, 1996).
- Braddock, J. H. II, and J. M. McPartland 1987 “How Minorities Continue to be Excluded from Equal Employment Opportunities: Research on Labor Market and Institutional Barriers.” *Journal of Social Issues* 43:5–39.
- 1989 “Social Psychological Processes that Perpetuate Racial Segregation: The Relationship Between School and Employment Desegregation.” *Journal of Black Studies* 19:267–289.
- Burkey, R. 1978 *Ethnic and Racial Groups: The Dynamics of Dominance*. Menlo Park, Calif.: Cummings.
- Burstein, P. 1985 *Discrimination, Jobs, and Politics*. Chicago: University of Chicago Press.
- Castles, S. 1984 *Here for Good: Western Europe’s New Ethnic Minorities*. London: Pluto Press.
- Conway, D. A., H.V. and Roberts 1994 “Analysis of Employment Discrimination through Homogeneous Job Groups.” *Journal of Econometrics* 61:103–131.
- Corcoran M., and G. J. Duncan 1978 “Work History, Labor Force Attachment, and Earning Differences Between Races and Sexes.” *Journal of Human Resources* 14:3–20.
- Daniel, W. W. 1968 *Racial Discrimination In England*. Harmondsworth, UK: Penguin.
- Dempster, A. P. 1988 “Employment Discrimination and Statistical Science.” *Statistical Science* 3:149–195.
- De Vries, S., and T. F. Pettigrew 1994 “A Comparative Perspective on Affirmative Action: *Positieve aktie* in the Netherlands.” *Basic and Applied Social Psychology* 15:179–199.
- England, P., G. Farkas, B. Kilbourne, and T. Dou 1988 “Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects.” *American Sociological Review* 53:544–558.
- Farley, R. 1984 *Blacks and Whites: Narrowing the Gap?* Cambridge, Mass.: Harvard University Press.
- Feagin, J. R., and C. B. Feagin 1986 *Discrimination American Style: Institutional Racism and Sexism*, 2d ed. Malabar, Fla.: Krieger.
- Feagin, J. R., and M. P. Sikes 1994 *Living with Racism: The Black Middle-Class Experience*. Boston, Mass.: Beacon Press.
- Fox, M. F. 1991 “Gender, Environmental Milieux, and Productivity in Science.” In J. Cole, H. Zuckerman, and J. Bruer, eds., *The Outer Circle: Women in the Scientific Community*. New York: Norton.
- Gill, A. M. 1989 “The Role of Discrimination in Determining Occupational Structure.” *Industrial and Labor Relations Review* 42:610–623.
- Gordon, P., and F. Klug 1984 *Racism and Discrimination in Britain: A Select Bibliography, 1970–83*. London: Runnymede Trust.

- Heckman, J. J., and B. S. Payner 1989 "Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina." *American Economic Review* 79:138-177.
- Layton-Henry, Z., and C. Wilpert 1994 *Discrimination, Racism and Citizenship: Inclusion and Exclusion in Britain and Germany*. London: Anglo-German Foundation for the Study of Industrial Society.
- Leonard, J. S. 1994 "Use of Enforcement Techniques in Eliminating Glass Ceiling Barriers." Report to the Glass Ceiling Commission. U. S. Dept. of Labor, Washington, D.C.
- MacEwen M. 1995 *Tackling Racism in Europe: An Examination of Anti-Discrimination Law in Practice*. Washington, D.C.: Berg.
- MacMaster N. 1991 "The 'Seuil de Tolerance': The Uses of a 'Scientific' Racist Concept." In M. Silverman, ed., *Race, Discourse and Power in France*. Aldershot, UK: Avebury.
- Mayhew, L. H. 1968 *Law and Equal Opportunity: A Study of the Massachusetts Commission Against Discrimination*. Cambridge, Mass.: Harvard University Press.
- Pettigrew, T. F. 1979 "Racial Change and Social Policy." *Annals of the American Association of Political and Social Science* 441:114-131.
- 1985 "New Black-White Patterns: How Best to Conceptualize Them?" *Annual Review of Sociology* 11:329-346.
- 1998 "Responses to the New Minorities of Western Europe." *Annual Review of Sociology* 24:77-103.
- , and R. W. Meertens 1996 "The Verzuiling Puzzle: Understanding Dutch Intergroup Relations." *Current Psychology* 15:3-13.
- Reskin, B. F. 1998 *The Realities of Affirmative Action in Employment*. Washington, D.C.: American Sociological Association.
- Rosenfeld, R. A., and A. L. Kalleberg 1990 "A Cross-national Comparison of the Gender Gap in Income." *American Journal of Sociology* 96:69-105.
- Rosensweig, M. R., and J. Morgan 1976 "Wage Discrimination: A Comment." *Journal of Human Resources* 11:3-7.
- Sanborn, H. 1969 "Pay Differences Between Men and Women." *Industrial and Labor Relations Review* 17:534-550.
- Smith, D. J. 1976 *The Facts of Racial Disadvantage: A National Survey*. London: PEP.
- Taylor, M. C. 1988 "Estimating Race and Sex Inequity in Wages: Substantive Implications of Methodological Choices." Paper presented at the 1989 Research

Conference of the Association for Public Policy Analysis and Management. Seattle, Wash.

- Veenman, J., and T. Roelandt 1990 "Allochtonen: Achterstand en Achterstelling." In J. J. Schippers, ed., *Arbeidsmarkt en Maatschappelijke Ongelijkheid*. Groningen, The Netherlands: Wolters-Noordhoff.
- Williams, R. M., Jr. 1947 *The Reduction of Intergroup Tensions*. New York: Social Science Research Council.
- Wilpert, C. 1993 "Ideological and Institutional Foundations of Racism in the Federal Republic of Germany." In J. Solomos and J. Wrench, eds., *Racism and Migration in Western Europe*. Oxford, UK: Berg.

THOMAS F. PETTIGREW  
MARYLEE C. TAYLOR

## DISENGAGEMENT THEORY

See Aging and the Life Course; Retirement.

## DISTRIBUTION-FREE STATISTICS

See Nonparametric Statistics.

## DISTRIBUTIVE JUSTICE

See Human Rights/Children's Rights; Social Justice.

## DIVISION OF LABOR

The division of labor is a multifaceted concept that applies to several levels of analysis: small groups, families, households, formal organizations, societies, and even the entire "world system" (Wallerstein 1979). Each level of analysis requires a slightly different focus; the division of labor may refer to the emergence of certain roles in groups, to the relative preponderance of industrial sectors (primary, secondary, tertiary) in an economy, to the distribution of societal roles by sex and age, or to variability among occupational groupings. Sociologists at the micro-level are concerned with "who does what," while macrosociologists focus on the larger structural issues of societal functions.

All of the major sociological theorists considered the division of labor to be a fundamental concept in understanding the development of modern society. The division of labor in society has been a focus of theoretical debate for more than a century, with some writers most concerned about hierarchical divisions, or the vertical dimension, and others emphasizing the heterogeneity or the horizontal dimension, of a social system. The analysis here is primarily theoretical, and so does not deal with debates about the technical problems of measurement (but see Baker 1981; Clemente 1972; Gibbs and Poston 1975; Land 1970; Rushing and Davies 1970).

### **THE DIVISION OF LABOR IN SMALL GROUPS**

Most small groups exhibit "role differentiation," as first described by Simmel (1890), Whyte (1943), and Bales and Slater (1955). The research on informal work groups stresses a hierarchical form of the division of labor, or the emergence of leadership; in particular, the emergence of both a "task leader" and a "social leader." It appears that defining group goals and enforcing norms is a type of activity incompatible with maintaining group cohesion, and therefore most groups have a shared leadership structure. Even in same-sex groups, this structure of coleadership is apparent; also, the degree of role differentiation appears to be greater in larger groups. The functional need for a social leader depends in part on the degree to which group members are task oriented, and on the degree to which the task leader is perceived as a legitimate authority (with the power to reward and punish other group members). Burke (1969) has also noted that a "scapegoat" role may emerge within task groups, reducing the need for a social leader to maintain harmony among the rest of the group members.

### **HOUSEHOLD AND FAMILY DIVISION OF LABOR**

When the family or household is the unit of analysis, issues concerning the division of tasks between spouses are of primary interest to sociologists. One major finding has been the persistence of disproportionately high levels of traditional housework by the wife, even when she is employed

outside the home. This reflects the "provider" vs. "homemaker" role distinction that formerly characterized most nuclear families in industrialized societies. Another major family role, occupied almost exclusively by the wife, is that of "kinkeeper." The maintenance of family traditions, the recording of important family anniversaries, and the coordination of visits between households are all vital elements of this role. The classic works by Bott (1957) and Blood and Wolfe (1960) documented the spousal division of labor as it was influenced by wider social networks and the relative power of the husband and wife. Kamo (1988) has noted that cultural ideology is a strong factor in the determination of spousal roles, although it is not entirely independent of the resources available to husband and wife.

Within a household, both age and sex structure the division of labor. Young children have few responsibilities, while parents have many, and males and females tend to do very different tasks. For example, White and Brinkerhoff (1981) studied the reported allocation of a variety of household tasks among a sample of Nebraska households, and found that the youngest cohort of children (aged two to five) performed relatively few household tasks and showed little differentiation by sex. Older children, however, diverged considerably in their roles (males doing more outdoor work, females more responsible for cooking or childcare). At the other end of the lifespan, the roles of the grandparent generation were strongly influenced by geographical proximity, the ages of the grandparents, and the ages of the children. Divorce among the parental generation can also substantially impact the roles adopted by grandparents (see Bengston and Robertson 1985).

### **FORMAL ORGANIZATIONS AND OCCUPATIONAL SPECIALIZATION**

Formal organizations are always structured by an explicit division of labor (an organization chart contains the names of functions, positions, or subunits). The horizontal differentiation in an organization, or task specialization, is normally based on functional units that are of roughly equivalent levels in the hierarchy of authority. The division of labor may also be constructed on a geographical basis, with the extreme examples being

transnational corporations that draw raw resources and labor from the developing nations, and management from the developed nations.

Labor is also organized hierarchically, into levels of authority. Max Weber was one of the first sociologists to analyze the emergence of modern bureaucracies, where the division of labor is fundamental. Weber's approach suggests a maximum possible level of specialization, so that each position can be filled by individuals who are experts in a narrow area of activity. An extreme form of the bureaucratic division of labor was advocated by Frederick Taylor's theory of "Scientific Management" (1911). By studying in minute detail the physical motions required to most efficiently operate any given piece of machinery, Taylor pioneered "time and motion studies." Along with Henry Ford, he also made the assembly line a standard industrial mode of production in modern society. However, this form of the division of labor can lead to isolation and alienation among workers, and so more recent organizational strategies, in some industrial sectors, emphasize the "craft" approach, in which a team of workers participate more or less equally in many aspects of the production process (see Blauner 1964; Hedley 1992).

### SOCIETAL DIVISION OF LABOR

Macrosociologists measure the societal division of labor in a number of ways, most commonly by considering the number of different occupational categories that appear in census statistics or other official documents (see Moore 1968). These lists are often implicitly or explicitly ranked, and so both horizontal and vertical division of labor can be analyzed. This occupational heterogeneity is dependent in part on the official definitions, but researchers on occupational prestige (the vertical dimension) have shown comparable levels in industrialized societies such as the United States, Canada, England, Japan, Sweden, Germany, and France (see Treiman 1977). The "world system" as described by Wallerstein (1979) is the upper limit of the analysis of the division of labor. Entire societies are characterized as "core" or "periphery" in Wallerstein's analysis of the global implications of postindustrialism.

### THEORETICAL APPROACHES

Adam Smith, in *The Wealth of Nations* (1776), produced the classic statement of the economic efficiencies of a complex division of labor. He observed that the manufacture of steel pins could be more than two hundred times more productive if each separate operation (and there were more than a dozen) were performed by a separate worker. The emergence of the systematic and intentional division of labor probably goes back to prehistoric societies, and is certainly in evidence in the ancient civilizations of China, Egypt, Greece, and Rome. The accumulation of an agricultural surplus and the establishment of markets both created and stimulated the differentiation of producers and consumers. Another phase in the transformation in Europe was the decline of the "craft guilds" that dominated from the twelfth to the sixteenth centuries. When labor became free from the constraints of the journeyman/apprentice system and when factories began to attract large numbers of laborers, the mechanization of modern work flourished. (For economic and political perspectives on the division of labor see Krause 1982; Putterman 1990).

Early sociologists such as Herbert Spencer considered the growth of societies as the primary determinant of the increased specialization and routinization of work; they also emphasized the positive impacts of this process. Spencer, like other functionalists, viewed human society as an organic system that became increasingly differentiated as it grew in size, much as a fertilized egg develops complex structures as it develops into a full-fledged embryo. In his *Principles of Sociology* (1884), Spencer considered the evolution of human society as a process of increasing differentiation of structure and function.

Karl Marx (1867) argued that the increasing division of labor in capitalist societies is a primary cause of alienation and class conflict, and therefore is a force in the eventual transformation to a socialist/communist society. In fact, a specific question asking for details of the division of labor appeared on one of the earliest questionnaire surveys in sociology done by Marx in 1880. Marx and his followers called for a new form of the division of labor, supported by an equalitarian ethos, in which individuals would be free to choose

their productive roles; labor would not be alienating because of the common ideology and sense of community.

Weber (1947) painted a darker picture when he documented the increasing “rationalization” of society, especially the ascendance of the bureaucratic division of labor with its coordinated system of roles, each highly specialized, with duties specified in writing and incumbents hired on the basis of their documented competence at specific tasks. The ideal-type bureaucracy was in actuality subject to the negative consequences of excessive specialization, however. Weber pointed out that the “iron cage” placed stifling limits on human freedom within the organization, and that decisions by bureaucrats often became so rule bound and inflexible that the clients were ill served.

Georg Simmel’s “differentiation and the principle of saving energy” (1976 [1890]) is a little-known essay that similarly describes the inevitable problems that offset the efficiencies gained by the division of labor; he called these “friction, indirectness, and superfluous coordination.” He also echoed Marx and Engels when he described the effects of high levels of differentiation upon the individual:

*... differentiation of the social group is evidently directly opposed to that of the individual. The former requires that the individual must be as specialized as possible, that some single task must absorb all his energies and that all his impulses, abilities and interests must be made compatible with this one task, because this specialization of the individual makes it both possible and necessary to the highest degree for him to be different from all other specialized individuals. Thus the economic setup of society forces the individual for life into the most monotonous work, the most extreme specialization, because in this way he will acquire the skill which makes possible the desired quality and cheapness of the product. (Simmel 1976, p. 130)*

While the foregoing theorists contributed substantially to the understanding of the division of labor, Emile Durkheim’s *The Division of Labor in Society* (1893) stands as the classic sociological statement of the causes and consequences of the historical shift from “mechanical solidarity” to

“organic solidarity.” The former is found in smaller, less-advanced societies where families and villages are mostly self-sufficient, independent, and united by similarities. The latter is found in larger, urbanized societies where specialization creates interdependence among social units.

Following Spencer’s lead, Durkheim noted that the specialization of functions always accompanies the growth of a society; he also observed that increasing population density—the urbanization of society that accompanies modernization—greatly increases the opportunities for further increases in the division of labor.

It should be noted that the shift to a modern division of labor could not have occurred without a preexisting solidarity; in his chapter on “organic and contractual solidarity” he departed from Spencer’s utilitarian explanation of social cohesion, and noted that the advanced division of labor can occur only among members of an existing society, where individuals and groups are united by preexisting similarities (of language, religion, etc.).

A sense of trust, obligation, and interdependence is essential for any large group in which there are many diverse roles; indirect exchanges occur; and individuals form smaller subgroupings based on occupational specialization. All of these changes create high levels of interdependence, but with increasing specialization, and different world views develop, along with different interests, values, and belief systems. This is the problem Durkheim saw in the shift from mechanical to organic solidarity; he feared the “anomie” or lack of cohesion that might result from a multiplicity of views, languages, and religions within a society (as in the France of his times, and even more so today). The problems of inequality in modern industrial society were not lost on Durkheim, either; he noted how the “pathological form of the division of labor” posed a threat to the full development of social solidarity (see Giddens 1971). Although many simplistic analyses of Durkheim’s approach suggest otherwise, he dealt at length with the problems of “the class war” and the need for justice and fraternity.

The division of labor is treated as a key element in Peter Blau’s book, *Inequality and Heterogeneity* (1977). This important work emphasizes the primacy of differentiation (division of labor) as an

influence on mobility, prejudice, conflict, affiliation, intermarriage, and inequality. In Blau's scheme, a society may be undifferentiated (all persons or positions are independent, self-sufficient, etc.) or strongly differentiated (a high degree of specialization and interdependence). This differentiation may be a matter of degree, on dimensions such as authority, power, prestige, etc.; this constitutes inequality. Or the differentiation may be a matter of kind, such as occupational differentiation—the division of labor.

Blau, like Durkheim, distinguished two major types of division of labor: *routinization* and *expert specialization*.

*The two major forms of division of labor are the subdivision of work into repetitive routines and its subdivision into expert specialties. . . when jobs are divided into repetitive routines, the training and skills needed to perform them are reduced, whereas when they are divided into fields of specialists, the narrower range of tasks permits greater expertness to be acquired and applied to the work, increasing the training and skills required to perform it.*  
(Blau 1977, p. 188)

Blau has shown that the division of labor always increases inequality in the organization. The managerial and technical experts coordinate the increasing number and diversity of routinized positions. As organizations and societies increase in size and population density, the division of labor increases. At the same time, the forces of industrialization and urbanization require and encourage further specialization, and in most cases, increase inequality, and indirectly, social integration. Some of the confounding boundaries include the degree of linguistic, ethnic, or cultural heterogeneity (all of which can inhibit integration), and social or geographic mobility (which can increase integration).

In a rare display of explicitly stated definitions and propositions, Blau created a landmark theory of social organization. Here are the most important of Blau's assumptions and theorems relating to the division of labor:

- The division of labor depends on opportunities for communication.
- Population density and urbanization increase the division of labor.

- Rising levels of education and qualifications promote an advanced division of labor.
- Large work organizations promote the division of labor in society.
- Linguistic heterogeneity impedes the division of labor.
- The more the division of labor is in the form of specialization rather than routinization, the higher are rates of associations among different occupations, which produces higher integration.
- The more the division of labor intersects with other nominal parameters (including kinship, language, religion, ethnicity, etc.) the greater is the probability that intergroup relations strengthen society's integration.
- The smaller an organization, the more its internal division of labor increases the probabilities of intergroup and interstratum associations, and therefore the higher the degree of integration.  
(Summarized from Blau 1977, pp. 214–215)

Blau concludes by noting that “Advances in the division of labor tend to be accompanied by decreases in various forms of inequality but by increases in inequality in power. Although the advancing division of labor does not generate the growing concentration of power, the two are likely to occur together, because the expansion of work organizations promotes both.” (p. 214).

## CONCLUSION

Ford's moving assembly lines began to produce the frames for Model T automobiles in 1913, at a rate of about one every two working days, but within months, refinements on the assembly process reduced this to four units per day. This eightfold increase in efficiency was accompanied by a decrease in the price of the cars and indirectly stimulated a very large industry. Now, automated factories, using robotics and highly specialized computer systems, have dramatically increased the efficiency of the automobile industry. The effects on morale and the environment, however, appear to be less salutary.

(SEE ALSO: *Bureaucracy, Complex Organizations, Convergence Theories, Family and Household Structure, Family Roles, Industrial Sociology, Industrialization, Parental Roles, Social Change, Social Structure, Technology and Society, and Work and Occupations.*)

REFERENCES

Baker, P.M. 1981 "The Division of Labor: Interdependence, Isolation, and Cohesion in Small Groups." *Small Group Behavior* 12(1):93-106.

Bales, R.F., and P.E. Slater 1955 "Role Differentiation in Small Decision-Making Groups." In T. Parsons et al., eds., *Family, Socialization, and Interaction Process*. Glencoe, Ill.: Free Press.

Bengston, V.L., and J.F. Robertson 1985 *Grandparenthood*. Beverly Hills, Calif.: Sage.

Blau, P.M. 1977 *Inequality and Heterogeneity*. New York: Free Press.

Blauner, R. 1964 *Alienation and Freedom*. Chicago: University of Chicago Press.

Blood, R., and D. Wolfe 1960 *Husbands and Wives: The Dynamics of Married Living*. New York: Free Press.

Bott, E. 1957 *Family and Social Network*. London: Tavistock Publications.

Burke, P.J. 1969 "Scapegoating: An Alternative to Role Differentiation." *Sociometry* 32(June):159-168.

Clemente, R. 1972 "The Measurement Problem in the Analysis of an Ecological Concept: The Division of Labor." *Pacific Sociological Review* 15:30-40.

Durkheim, E. (1893) 1933 *The Division of Labor in Society*. New York: Free Press.

Gibbs, J.P., and D.L. Poston 1975 "The Division of Labor: Conceptualization and Related Issues." *Social Forces* 53 (March) 468-476.

Giddens, A. 1971 *Capitalism and Modern Social Theory: Analysis of the Writings of Marx, Durkheim, and Max Weber*. Cambridge: Cambridge University Press.

Hedley, R.A. 1992 *Making a Living: Technology and Change*. New York: Harper Collins.

Kamo, Y. 1988 "Determinants of Household Labor: Resources, Power, and Ideology." *Journal of Family Issues* 9:177-200.

Krause, E.A. 1982 *The Division of Labor: A Political Perspective*. Westport, Conn.: Greenwood Press.

Land, K.C. 1970 "Mathematical Formalization of Durkheim's Theory of Division of Labor." E.F. Borgatta, ed., *Sociological Methodology 1970*, 257-282. San Francisco: Jossey-Bass.

Marx, K. (1867) 1977 *Capital: A Critique of Political Economy*. New York: Vintage Books.

Moore, W.E. 1968 *Economy and Society*. New York: Random House.

Putterman, L. 1990 *Division of Labor and Welfare: An Introduction to Economic Systems*. New York: Oxford University Press.

Rushing, W.A., and V. Davies 1970 "Note on the Mathematical Formalization of a Measure." *Social Forces* 48(March):394-396.

Simmel, G. (1890) 1976 "On Social Differentiation." In P. Lawrence, ed., *Georg Simmel, Sociologist and European*. New York: Nelson.

Smith, Adam (1776) 1985 *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York: Modern Library.

Smith, D., and R. Snow 1976 "The Division of Labor: Conceptual and Methodological Issues." *Social Forces* 55:520-528.

Spencer, H. 1884 *Principles of Sociology*. New York: Appleton-Century-Crofts.

Taylor, F. 1911 *Scientific Management*. New York: Harper.

Treiman, D. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

Wallerstein, I. 1979 *The Capitalist World-Economy*. New York: Cambridge University Press.

Weber, M. (1908) 1947 *The Theory of Social and Economic Organization*. New York: Free Press.

Whyte, W.F. 1943 *Street Corner Society*. Chicago: Chicago University Press.

PAUL MORGAN BAKER

**DIVORCE**

Divorce is of sociological significance for several reasons. To begin, divorce rates are often seen as indicators of the health of the institution of marriage. When divorce rates rise or fall, many sociologists view these changes as indicating something about the overall quality of marriages or, alternatively, the stability of social arrangements more generally. Viewed from another perspective, divorce interests sociologists as one of several important transitions in the life course of individuals. The adults and children who experience divorce have been studied to understand both the causes and consequences. From this perspective, a divorce is as much an event in the biography of



family members, as other life-course transitions (remarriage, childbirth, and retirement). The sociological interest in divorce also focuses on the social trends it is part of, figuring prominently in any sociological analysis of industrialization, poverty rates, educational attainment, strategies of conflict resolution, or law.

For sociologists, divorce may characterize an individual, a family, a region, a subgroup, a historical period, or an entire society. It may be studied as either the cause or consequence of other phenomena. Still, the overriding concern of almost all research on this topic has been the increase in divorce over time. Divorce is now almost as common as its absence in the lives of recently married couples. The National Center for Health Statistics estimates that 43 percent of marriages begun in the early 1990s will end in divorce (NCHS 1998), a significant decline from the estimates of 50 percent to 65 percent in the late 1980s (Martin and Bumpass 1989). The decline in divorce rates in the recent past is probably a result of the aging of the post-World War II Baby Boom generation who are no longer at high risk of divorce because of their age. It is also possible that American marriages are becoming somewhat more stable than they were a decade ago. Still, the fluctuations in divorce rates one decade to the next do not mask the more general trend for the past two centuries. Understanding the increase in divorce has been the larger sociological endeavor regardless of the particular perspective employed. A historical account of trends is necessary before considering contemporary issues associated with divorce.

#### A BRIEF HISTORICAL RECORD OF DIVORCE IN AMERICA

**The Colonial Period.** Divorce was not legal in any but the New England settlements. The Church of England allowed for legal separations (*a mensa et thoro*), but not for divorce. The New England Puritans who first landed at Plymouth in 1621, however, were disenchanted with this, as well as many other Anglican doctrines. Divorce was permitted on the grounds of adultery or seven-year desertion as early as 1639 in Plymouth. Other New England colonies followed similar guidelines. Divorce governed by rudimentary codified law was granted by legislative decree. Individual petitions for divorce were debated in colonial legislatures

and were effected by bills to dissolve a particular marriage. Still, though legal, divorce was very rare. During the seventeenth century, there were fifty-four petitions for divorce in Massachusetts, of which forty-four were successful (Phillips 1988, p. 138). The middle colonies provided annulments or divorces for serious matrimonial offenses such as prolonged absence or bigamy. The southern colonies afforded no provisions for divorce whatsoever.

**Post-Revolutionary War.** Immediately after the Revolutionary War, without British legal impediments to divorce, the states began discussion of laws to govern divorce. In New England and the middle states, divorce became the province of state courts while in the more restrictive southern states it was more often a legislative matter. By the turn of the nineteenth century, almost all states had enacted some form of divorce law. And by the middle of the century, even southern states were operating within a judicial divorce system.

The shift to judicial divorce is significant. By removing divorce deliberations from legislatures, states were forced to establish grounds that justified a divorce. Such clauses reflected the prevailing sentiments governing normative marriage—they indicated what was expected of marriage at the time. And by investing judges with the authority to interpret and adjudicate, such changes significantly liberalized the availability of divorce. Northern and southern states permitted divorces for specific offenses such as adultery, desertion, bigamy, and increasingly with time, cruelty. In the newer frontier western states, grounds resembled those of the East plus “any other cause for which the court shall deem it proper that the divorce shall be granted” (Phillips 1988, p. 453).

Throughout the nineteenth century, there was a gradual liberalization of divorce laws in the United States and a corresponding increase in divorce as well. Where divorces totaled a few hundred at the beginning of the nineteenth century, the numbers grew exponentially as the century wore on; 7,380 divorces in 1860, 10,962 in 1870, 19,663 in 1880, 33,461 in 1890, and 55,751 in 1900 (U.S. Bureau of the Census 1975). These figures assume greater significance when growth in population is removed from them. Whereas the divorce rate (number of divorces per 1,000 marriages) was but 1.2 in 1869, it had climbed to 4.0 by 1900. In short, the increase

in divorce outstripped the increase in population several times.

A number of factors have been identified as causes of such dramatic increases. In part, these can be described as social changes, which made marriage less essential. The growth of wage labor in the nineteenth century afforded women an alternative to economic dependence on a husband. In an economy dominated by individuals rather than families, marriage was simply less essential. Life as a single individual gradually lost its legal or social stigma (New England settlements had forbidden solitary dwelling while southern communities had taxed it heavily).

More important, however, were fundamental shifts in the meaning of marriage. Divorce codes reflected the growing belief that marriages should be imbued with heavy doses of affection and equality. Divorce grounds of cruelty or lack of support indicate that marriage was increasingly viewed as a partnership. Where a century earlier men had been granted greater discretion in their personal lives, latter nineteenth-century morality attacked such double standards. Men were not necessarily less culpable than women for their vices. Victorian morality stressed the highest standards of sexual behavior for both husbands and wives. Changing divorce codes coincided with the passage of laws restricting husbands' unilateral control over their wives' property. The passage of married women's property acts throughout the nation in the latter nineteenth century acknowledged married women's claims to property brought to or acquired in marriage. By 1887, thirty-three states and the District of Columbia gave married women control over their property and earnings (Degler 1980, p. 332).

Divorce codes including omnibus grounds such as "cruelty" (which could justify a divorce from a drunkard husband, for example) may be viewed as reflecting a Victorian American belief that women were morally sensitive and fragile, and in need of protection (Phillips 1988, p. 500). More particularly, the growing use of offenses against the intimate and emotional aspects of marriage reflected a growing belief that such things constituted matrimonial essentials. If a failure of intimacy could justify the dissolution of a marriage, then intimacy may be viewed as a core expectation of marriage.

**The Twentieth Century.** The first half of the twentieth century was a continuation of trends

established in the latter nineteenth century. Two world wars and the Great Depression interrupted gradually increasing divorce rates, however. During each war and during the Depression, divorce rates dropped. After each, rates soared before falling to levels somewhat higher than that which preceded these events. Sociological explanations for these trends focus on women's employment opportunities. Women's labor force participation permits the termination of intolerable unions. The separations, hastily timed marriages, and sexual misalliances characteristic of wartime were also undoubtedly factors in the post-war divorces rates. Further, the increases in divorce following these difficult times may be seen, in part, as a delayed reaction. Once the Depression or war was over, the reservoir of impending divorces broke. And finally, postwar optimism and affluence may have contributed to an unwillingness to sustain an unhappy marriage.

The second half of the century witnessed even more dramatic increases in divorce. With the exception of the peculiar 1950s (for an explanation of this anomaly, see Cherlin 1992), the trend for the second half of the 1900s was a regular and exponential growth in divorce until around 1980, at which point the increase stopped.

Though specific explanations for the increase in divorces during the twentieth century vary, several themes may be noted. First, marriage has lost much of its central economic and social significance—especially for women. For example, divorce was undoubtedly inhibited by the fact that prior to the twentieth century, custody of children was uniformly awarded to fathers (since they were legally responsible for financial support). With the acceptance of Freudian ideas of psychosexual development and similar ideas about intellectual and cognitive growth, the so-called Tender Years Doctrine became accepted practice in courts during the early 1900s which then awarded custody to mothers as regularly as they had once done to fathers. And as it became more commonplace, remarriage began to lose some of its stigma. All these changes made it possible for women to divorce their husbands if they wished. But why did so many wish to obtain divorces?

The simplest explanation is that more divorce is a consequence of higher expectations of marriage. More and more grounds for divorce are

developed as there are higher and higher expectations for what a marriage should be. In the nineteenth century, drunkenness, cruelty, and failure to provide were added to more traditional grounds of adultery and desertion. In the early twentieth century, cruelty was continually redefined to include not only physical, but mental cruelty as well.

The post-war surges in divorce created sufficient numbers of divorced persons so that the practice lost much of its stigma. The increase in divorce becomes more understandable when the loss of stigma is considered alongside the increase in women's employment since the mid 1960s. When women are employed, there is less constraint on them to remain in a marriage. But there is also less constraint on their husbands who will not be required to support their employed ex-wives after a divorce.

Since 1970, divorce has been fundamentally redefined. No-fault divorce laws passed since the early 1970s have defined as unacceptable those marriages in which couples are "incompatible," have "irreconcilable differences" or in which the marriage is "irretrievably broken." Prior to the no-fault regime, divorces required proof of a fault (crime) on the part of one spouse. The court decided whether to grant the divorce. Divorce proceedings were intentionally adversarial. Today, the non-adversarial grounds for divorce are almost entirely based on the failures of emotional essentials. Emotional marital breakdown may have been a feature of large numbers of marriages in earlier historical periods. Only now, however, is such a situation viewed as solely sufficient grounds for terminating the marriage.

### DIVORCE IN THE WEST

Any theory of divorce must be able to account for the broad similarities in historical (twentieth century) trends throughout the entire Western world. These similarities exist despite notable differences in national economies, forms of government, and the role of the church. The trends are well known. There was very little divorce until the end of the nineteenth century, a slow but constant growth in divorce rates through the first half of the twentieth century (interrupted by two world wars and an international economic depression), and significant increases in divorce rates since the 1960s. The

twentieth century, in short, is when most significant changes in divorce rates occurred. And the changes noted in America were seen in most other Western nations.

Between World Wars I and II, there were widespread changes in divorce laws that reflected changing beliefs about matrimony and its essentials. The strains of war and the associated problems that produced more divorces made the practice more conspicuous and consequently more acceptable. There is no doubt one cause of divorce is divorce. When obscure, the practice was stigmatized and there was little to counter stereotypes associated with its practice. When divorce became more commonplace, it lost some of its stigma.

Social changes pertaining to women's roles are a large part of the story of divorce during the postwar era. One sign of these changes was the growth, throughout the West, of women's labor force participation. But the most conspicuous symbol of the changing role of women was the passage of suffrage legislation throughout the Western world. Before 1914, women were permitted to vote only in New Zealand, Australia, Finland, Norway, and eleven western U.S. states. In the United States, women were enfranchised in 1920. In Britain, Sweden, Germany, and many other European countries, suffrage passed soon after World War I.

Divorce laws, similarly, were altered between the wars in accordance with changing views of marriage and the role of women. The British Parliament enacted divorce reform in 1937 by significantly extending the grounds for divorce (including cruelty) and granting women new options for filing for divorce. Scotland reformed its divorce laws in 1938 by extending grounds for divorce to include failures of emotional essentials—cruelty and habitual drunkenness, for example. In 1930, the Canadian Parliament for the first time empowered judicial magistrates to grant divorce rather than requiring legislative decrees. And the Spanish divorce law of 1932 was the most liberal in contemporary Europe—providing divorce by mutual consent (Phillips 1988, p. 539). Even Nazi Germany permitted no-fault divorce by 1938 (though divorce law was aimed at increasing the number of Aryan children born).

Following World War II, divorce rates throughout the Western world stabilized after an initial increase. The low divorce rates, high fertility, and

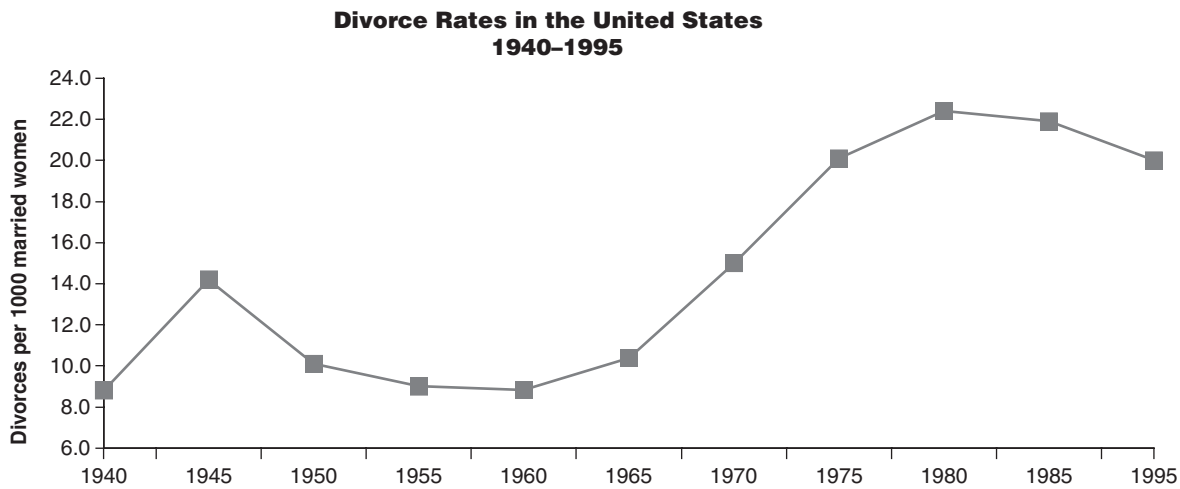


Figure 1

SOURCE: Vital Statistics of the United States

lower age at marriage that characterized all Western nations after World War II are trends that have not been adequately explained. Whether these trends reflected the consequences of war, the effects of having grown up during the worldwide depression, or a short-term rise in social conservatism is now debated. Regardless of the cause, the decade of the 1950s is universally regarded as a temporary aberration in otherwise long-term and continuous twentieth century trends. Not until the 1960s were there additional significant changes in divorce laws or divorce rates.

The 1960s were years of significant social change in almost all Western nations. The demographic consequences of high fertility during the 1950s became most apparent in the large and vocal youth movements challenging conventional sexual and marital norms, censorship, the war in Vietnam, and educational policies. Challenges to institutional authority were commonplace. Divorce laws were not immune to the general liberalization. "Between 1960 and 1986 divorce policy in almost all the countries of the West was either completely revised or substantially reformed" (Phillips 1988, p. 562). Most such reforms occurred in the late 1960s to the late 1970s. Unlike earlier divorce law reforms, those during the post-World War II era did not extend the grounds for divorce so much as they redefined the jurisdiction over it. The passage

of no-fault divorce laws signaled a profound shift in the way divorce was to be handled.

Most significantly, divorce became the prerogative of the married couple with little involvement of the state. No-fault divorce laws do not require either spouse to be guilty of an offense. Instead, they focus on the breakdown of the emotional relationship between the spouses. These statutes typically require a period of time during which the spouses do not live together. Beyond that, evidence must be adduced to substantiate one or both spouses' claim that the marriage is irretrievably broken. The significance of no-fault divorce lies entirely in the fact that decisions about divorce are no longer the prerogative of the state or church but rather of the married couple.

The passage of no-fault divorce laws in the West is properly viewed as a response to changing behaviors and attitudes. Indeed, social science research has shown that divorce rates began to increase significantly prior to passage of such laws and did not change any more dramatically afterward (Stetson and Wright 1975).

The changes in divorce law and actual divorce behaviors in the West are a reflection of the redefinition of marriage. More vulnerable and fragile emotional bonds have replaced the economic constraints that once held spouses together. The

availability of gainful employment for women makes marriage less essential and divorce more possible. Indeed, the significant changes in women's social positions and the corresponding changes in normative expectations (i.e., gender) have been the subject of significant sociological research. These changes are recognized as fundamentally altering almost all social institutions. Marriage is no exception.

The redefinition of marriage in the latter twentieth century throughout the West reflects the profound changes in relationships between men and women that have occurred. No longer an economic institution, marriage is now defined by its emotional significance. Love and companionship are not incidents of the institution. Rather, they are essentials. Meeting these high expectations may be difficult, but sustaining them is certainly more so.

Taken together, the changes in the second half of the twentieth century may be summarized as redefining the meaning of marriage. Children are not economic assets. Spouses are not economic necessities. Marriage is a conjugal arrangement where the primary emphasis is on the relationship between husband and wife. The reasons for divorce are direct consequences of the reasons for marriage. As one changes so does the other. Since it is more difficult to accomplish and sustain matrimonial essentials, it is easier to terminate the legal framework surrounding them. Divorce has become less costly (financially, legally, and reputationally) as marriage has become more so (in terms of the investments required to accomplish what is expected of it).

### CORRELATES OF DIVORCE

Sociologists have documented a number of demographic and personal characteristics that correlate with the probability of divorce. These include early age at marriage, premarital births, premarital cohabitation, divorce from a previous marriage, and low educational attainments. Social class is inversely related to divorce, yet wives' employment significantly increases divorce probabilities (see Huber and Spitze 1988 for a review).

Half of all recent marriages began with cohabitation (unmarried couples living together) (Bumpass and Sweet 1989). Repeated national studies have found that married couples who cohabited (either with each other, or with others) before marrying

have higher divorce rates than those who never cohabited (Nock 1995). The reason is still unclear. Research shows that cohabiting individuals are less committed to the idea of marriage or marital permanence. They are also less religious and tend to be drawn from lower social classes (both of which are associated with higher divorce rates) (Nock 1995). Cohabitation appears to foster (or reflect) a belief that problems in intimate relationships are solvable by ending the relationship. When such beliefs are carried into marriage, the result is higher divorce rates.

Race correlates with divorce—even after controls are imposed for socioeconomic correlates of race—with black individuals having divorce rates approximately twice those of whites. However, such large differences associated with race are recent in origin. Not until the late 1950s did significant differences in divorce, separation, and other marital statuses emerge between blacks and whites, even though a pattern of marginally higher marital disruption has been found among blacks for at least a century. Such findings suggest that the differences stem more from contemporary than historical circumstances. As Cherlin suggests, the recent changes in black Americans' family situations resemble those of other racial and ethnic groups, though they are more pronounced. The restructuring of the American economy, the decline in semi-skilled jobs, and the rise in service occupations has resulted in higher rates of black male unemployment or low wages, and better opportunities for black women. "Faced with difficult times economically, many blacks responded by drawing upon a model of social support that was in their cultural repertoire, a way of making it from day to day passed down by African Americans who came before them. This response relied heavily on extended kinship networks and de-emphasized marriage" (Cherlin 1992, p. 113).

### CONSEQUENCES OF DIVORCE

**For Children.** A central concern of much of the recent research on divorce is how children fare. Developmental psychologists describe five ways in which marital disruption may affect children's adjustment. First, some adults and some children are more vulnerable to the stress and strain of divorce. Personality characteristics, ethnicity, or

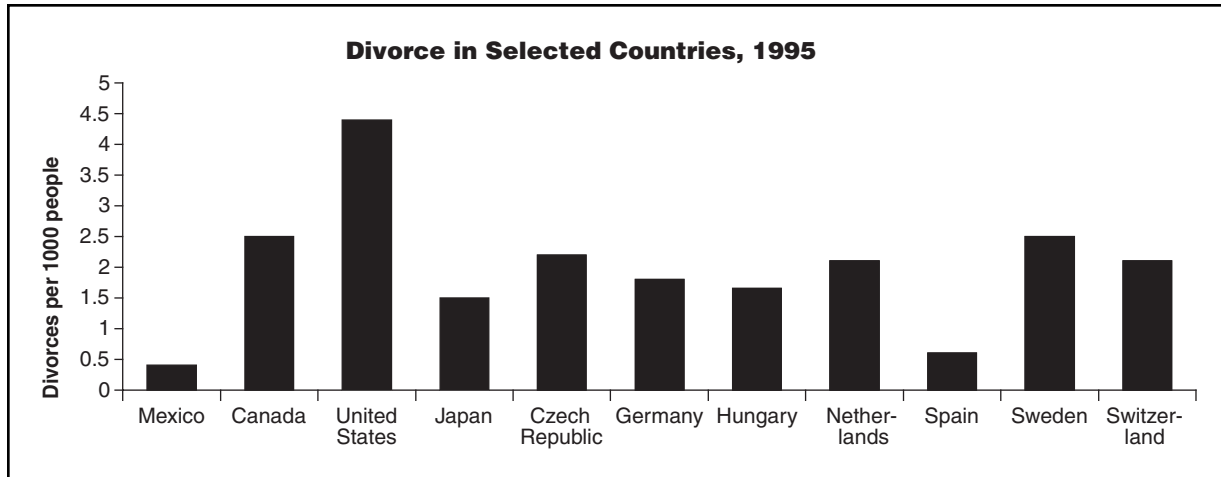


Figure 2

SOURCE: United Nations 1996 Demographic Yearbook

age for example, may make some individuals more susceptible to negative outcomes. Second, the absence of one parent, per se, may affect children's adjustment to divorce. Boys, in particular, appear to benefit from the presence of a male adult. Third, the loss of income creates many indirect problems for children, including changes in residence, school, neighborhood, and peer networks. Fourth, divorce often diminishes the custodial parent's ability to provide supportive and appropriate parenting, especially if depression follows marital disruption. And finally, negative, conflictual, and dysfunctional family relationships between parents, parents and children, and siblings are probably the most damaging consequence of divorce for children. (Hetherington, Bridges, and Insabella 1998).

Longitudinal research has shown that children who experience divorce differ from others before the disruption occurs. Cherlin showed that children whose parents were still married, but who would later divorce, showed more behavior problems and did less well in school than children whose parents would remain married (Cherlin et al. 1991).

Even after such predisruption differences are considered, divorce takes a toll in the lives of children who experience it. Divorce significantly increases the chances that young people will leave their homes due to friction with a parent, increases

the chances of premarital cohabitation, and increases the odds of premarital pregnancies or fatherhood (Cherlin, Kiernan, and Chase-Lansdale 1995). The effects of divorce in young adulthood include higher rates of unemployment and lower educational attainments. Divorce weakens young people's connections to their friends and neighbors due to higher rates of residential mobility (McLanahan and Sandefur 1994). Following divorce, many children are subjected to changes in residence, often to disadvantaged neighborhoods where peers have lower educational prospects. The lack of connections to others affects parents' ability to monitor their children. It also limits young people's knowledge about local employment opportunities.

The changed economic circumstances caused by divorce affects children in many indirect ways. The loss of available income may affect the quality of schools children attend if custodial parents move to poorer neighborhoods. The lack of income may limit children's opportunities for extracurricular activities (e.g., travel, or music lessons). The need for income often compels custodial parents to work more hours, reducing their ability to monitor children's after-school activities.

In their socioeconomic attainments, children who experienced their parents' divorce average one to two fewer years of educational attainment than children from intact homes (Krein and Beller

1988; Hetherington, Camara, and Featherman 1983). Such effects are found even after rigorous controls are imposed for such things as race, sex, years since the divorces, age at time of divorce, parental income, parental education, number of siblings, region of residence, educational materials in the home, or the number of years spent in the single-parent family. There are comparable effects of divorce on occupational prestige, income and earnings, and unemployment (Nock 1988).

White women who spent some childhood time in a single-parent family as a result of divorce are 53 percent more likely to have teenage marriages, 111 percent more likely to have teenage births, 164 percent more likely to have premarital births, and 92 percent more likely to experience marital disruptions than are daughters who grew up in two-parent families. The effects for black women are similar, though smaller. Controls for a wide range of background factors have little effect on the negative consequences of divorce. Further, remarriage does not remove these effects of divorce. And there is no difference between those who lived with their fathers and those who lived with their mothers after divorce. Experiencing parents' divorce has the same (statistical) consequences as being born to a never-married mother (McLanahan and Bumpass 1988; McLanahan and Sandefur 1994).

Such large and consistent negative effects have eluded simple explanation. Undoubtedly much of the divorce experience is associated with the altered family structure produced—in almost 90 percent of all cases a single-mother family—and the corresponding changes in family functioning. Such a structure is lacking in adult role models, in parental supervision, and in hierarchy. On this last dimension, research has shown that divorced women and their children are closer (less distinguished by generational distinctions) to one another than is true in intact families. Parent and child are drawn together more as peers, both struggling to keep the family going. The excessive demands on single parents force them to depend on their children in ways that parents in intact families do not, leading to a more reciprocal dependency relationship (Weiss 1975, 1976). Single mothers are “. . . likely to rely on their children for emotional support and assistance with the practical problems of daily life” (Hetherington, Camara,

and Featherman 1983, p. 218). In matters of discipline, single mothers have been found to rely on restrictive (authoritarian as opposed to authoritative) disciplinary methods—restricting children's freedom and relying on negative sanctions—a pattern psychologists believe reflects a lack of authority on the part of the parent (Hetherington 1972). Whatever else it implies, the lack of generational boundaries means a less hierarchical family and less authoritative generational distinctions.

The institutional contexts within which achievement occurs, however, are decidedly hierarchical in nature. Education, the economy, and occupations are typically bureaucratic structures in which an individual is categorically subordinate to a superior—an arrangement Goffman described as an “*eschelon authority structure*” (1961, p. 42). The nuclear family has been described as producing in children the skills and attitudes necessary for competition within such *eschelon authority relationships* in capitalist production and family childrearing. “The hierarchical division of labor (in the economy) is merely reflected in family life” (Bowles and Gintis 1976, p. 144–147). The relative absence of clear subordinate-superordinate relationships in single-parent families has been argued to inadequately socialize children, or place them in a disadvantageous position when and if they find themselves in hierarchical organizations (Nock 1988).

**For Adults.** A wide range of psychological problems has been noted among divorcing and recently divorced adults. A divorce occasions changes in most every aspect of adult life; residence, friendship networks, economic situation, and parental roles. Marriage in America makes significant contributions to individual well-being. Thus, regardless of the quality of the marriage that ends, emotional distress is a near-universal experience for those who divorce (Weiss 1979). Anxiety, anger, and fear are dominant psychological themes immediately before and after divorce. At least for a year or two after divorce, men and women report psychosomatic symptoms of headaches, loss of appetite, overeating, drinking too much, trembling, smoking more, sleeping problems, and nervousness (Group for the Advancement of Psychiatry 1980).

The emotional problems occasioned by divorce are accompanied by major changes in economic situations, as well—especially for women.

The vast majority of those involved in divorce experience a significant decline in their immediate standard of living. This problem is especially acute for women who—in almost 90 percent of cases—assume custody of children. Immediately after a divorce, women suffer an average 30 percent to 40 percent decline in their overall standards of living (Hoffman and Duncan 1988; Peterson 1989). Either in anticipation of or as a consequence of divorce, there is typically an increase in divorced women's labor force participation. Analyzing national longitudinal data, Peterson estimates that one year before the divorce decree (when most divorcing individuals are separated), women's average standard of living (total family income divided by the poverty threshold for a family of a particular size) is 70 percent of its level in the previous year. As a consequence of increased hours worked, the standard of living increases one year after divorce and by five or six years after divorce, "the standard of living of divorced women is about 85 percent of what it had been before separation" (1989, p. 48). Women who have not been employed during their marriages, however, are particularly hard-hit; the majority ending up in poverty.

Child support payments are not a solution to the economic problems created by divorce for two reasons. First, about one-quarter (24 percent) of women due child support receive none (39 percent of men awarded child support receive none). Another one-quarter receive less than the court-ordered amount. In 1991, the average amount of child support received by divorced mothers was \$3,011 per year (\$2,292 for men) (U.S. Bureau of the Census 1995). About 16.7 million, or 85 percent of the 19.8 million children in single-parent families in 1997 were living with the mother; 60 percent of whom were divorced (U.S. Bureau of the Census 1998a). Their median family income was \$22,999 compared with \$34,802 for those in single-father situations, and \$51,681 for children in households where both parents were present (U.S. Bureau of the Census 1998a). Families headed by single mothers are the most likely to be in poverty, and represent 55 percent of all poor families. In 1997, a third (31.6 percent) of all single-mother families were in poverty compared to 5.2 percent of two-parent families (U.S. Bureau of the Census 1998b). Analyzing national longitudinal data, Duncan concluded that changes in

family status—especially divorce and remarriage—are the most important cause of change in family economic well-being and poverty among women and children (1984).

Single-parent families in America have grown dramatically as a result of increasing divorce rates. And even though most divorced persons remarry, Bumpass has shown that the average duration of marital separation experienced by children under age 18 was 6.3 years and 7.5 years for whites and blacks respectively. In fact 38 percent of white and 73 percent of black children are still in a single-parent family 10 years after the marital disruption—a reflection of blacks' lower propensity to remarry and their longer intervals between divorce and remarriage (1984). The role of divorce in the formation of single-parent families differs by race. Among all single-parent white families, 25 percent are maintained by never-married mothers, 47 percent by divorced (or separated) mothers, 7 percent by never-married men, and 13 percent by divorced or separated men. Among black single-parent families, 59 percent are maintained by never-married women, 28 percent by divorced or separated women, 4 percent by never-married men, and only 3 percent by divorced or separated men. Divorce is the primary route to single-parent-hood for white mothers, whereas out-of-wedlock childbearing is for black mothers (U.S. Bureau of the Census 1998c, Table 11; 1998d).

Families headed by single women with children are the poorest of all major demographic groups regardless of how poverty is measured. Combined with frequent changes in residence and in employment following divorce, children and mothers in such households experience significant instabilities—a fact reflected in the higher rates of mental health problems among such women (Garfinkel and McLanahan 1986, pp. 11–17).

## CONCLUSION

High rates of remarriage following divorce clearly indicate that marital disruption does not signify a rejection of marriage. There is no evidence of widespread abandonment of conjugal life by Americans. Admittedly, marriage rates have dropped in recent years. However, such changes are best seen to be the result of higher educational attainments,



occupational commitments, and lower fertility expectations; not a rejection of marriage per se. Rather, increasing divorce rates reflect the fact that marriage is increasingly evaluated as an entirely emotional relationship between two persons. Marital breakdown, or the failure of marriage to fulfill emotional expectations, has come increasingly to be a cause for divorce. Since the 1970s, our laws have explicitly recognized this as justification for terminating a marriage—the best evidence we have that love and emotional closeness are the *sine qua non* of modern American marriage. Contemporary divorce rates thus signal a growing unwillingness to tolerate an unsatisfying emotional conjugal relationship.

The consequences of divorce for children are difficult to disentangle from the predictable changes in household structure. Whether the long-term consequences are produced by the single-parent situation typically experienced for five to ten years, or from the other circumstances surrounding divorce is not clear. It is quite apparent, however, that divorce occasions significant instabilities in children's and mothers' lives.

Our knowledge about the consequences of divorce for individuals is limited at this time by the absence of controlled studies that compare the divorced to the nondivorced. Virtually all research done to date follows the lives of divorced individuals without comparing them to a comparable group of individuals who have not divorced. A related concern is whether the consequences of divorce reflect the experience itself, or whether they reflect various selection effects. That is, are people who divorce different from others to begin with? Are their experiences the results of their divorce, or of antecedent factors?

When almost half of all marriages are predicted to end in divorce, it is clear that marital disruption is a conspicuous feature of our family and kinship system. Divorce creates new varieties of kin not traditionally incorporated in our dominant institutions. The rights and obligations attached to such kinship positions as spouse of the noncustodial father are ambiguous—itsself a source of problems. The social institution of the family is redefined as a consequence of divorce. Entering marriage, for example, is less commonly the beginning of adult responsibilities. Ending marriage is less commonly the consequence of death. Parents

are not necessarily co-residents with their children. And new categories of “quasi” kin are invented to accommodate the complex connections among previously married spouses and their new spouses and children. In many ways, divorce itself has become a dominant institution in American society. It is, however, significantly less structured by consensual normative beliefs than the family institutions to which it is allied.

## REFERENCES

- Bennett, Neil G., David E. Bloom, and Patricia H. Craig 1989 “The Divergence of Black and White Marriage Patterns.” *American Journal of Sociology* 95 (3):692–772.
- Bowles, Samuel, and Herbert Gintis 1976 *Schooling in Capitalist America*. New York: Basic Books.
- Bumpass, Larry L. 1984 “Children and Marital Disruption: A Replication and Update.” *Demography* 21:71–82.
- , and James A. Sweet 1989 “National Estimates of Cohabitation.” *Demography* 26:615–625.
- Cherlin, Andrew J. 1981 *Marriage, Divorce, and Remarriage*. Cambridge, Mass.: Harvard University Press.
- , F.F. Furstenberg, Jr., P.L. Chase-Lansdale, K.E. Kiernam, P.K. Robins, D.R. Morrison, and J.O. Teitler 1991 “Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States.” *Science* 252 (June 7):1386–1389.
- Cherlin, Andrew J., Kathleen E. Kiernam, and P. Lindsay Chase-Lansdale 1995 “Parental Divorce in Childhood and Demographic Outcomes in Young Adulthood.” *Demography* 32 (No. 3):299–318.
- Degler, Carl N. 1980 *At Odds: Women and the Family in America from the Revolution to the Present*. New York: Oxford University Press.
- Duncan, Greg J. 1984 *Years of Poverty, Years of Plenty*. Ann Arbor, Mich.: University of Michigan, Institute for Social Research.
- Garfinkel, Irwin, and Sara S. McLanahan 1986 *Single Mothers and their Children*. Washington, D.C.: The Urban Institute Press.
- Goffman, Irving 1961 *Asylums*. New York: Anchor.
- Group for the Advancement of Psychiatry 1980 *Divorce, Child Custody, and the Family*. San Francisco: Jossey-Bass.
- Hetherington, E. Mavis 1972 “Effects of Paternal Absence on Personality Development in Adolescent Daughters.” *Developmental Psychology* 7:313–326.
- , Margaret Bridges, and Glendessa M. Insabella 1998 “What Matters? What Does Not? Five Perspectives on the Association Between Marital Transitions

- and Children's Adjustment." *American Psychologist* 53 (No.2):167-184.
- Hetherington, E.M., K.A. Camara, and D.L. Featherman 1983 "Achievement and Intellectual Functioning of Children in One Parent Households." In J. T. Spence, ed., *Achievement and Achievement Motives*, 205-284 San Francisco: W. H. Freeman.
- Hoffman, Saul D., and Greg J. Duncan 1988 "What are the Economic Costs of Divorce?" *Demography* 25:641-645.
- Huber, Joan, and Glenna Spitze 1988 "Trends in Family Sociology." In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Krein, Sheila F., and Andrea H. Beller 1988 "Educational Attainment of Children from Single-Parent Families: Differences by Exposure, Gender, and Race." *Demography* 25:221-234.
- Martin, Teresa C., and Larry L. Bumpass 1989 "Recent Trends in Marital Disruption." *Demography* 26 (1):37-51.
- McLanahan, Sara S., and Larry Bumpass 1988 "Intergenerational Consequences of Family Disruption." *American Journal of Sociology* 94 (1):130-152.
- National Center for Health Statistics 1998 FASTATS: "Divorce." [www.cdc.gov/nchswww/fastats/divorce.htm](http://www.cdc.gov/nchswww/fastats/divorce.htm)
- Nock, Steven L. 1988 "The Family and Hierarchy." *Journal of Marriage and the Family* 50 (Nov):957-966.
- 1995 "A Comparison of Marriages and Cohabiting Relationships." *Journal of Family Issues* 16:53-76.
- Peterson, Richard R. 1989 *Women, Work, and Divorce*. Albany: State University of New York Press.
- Phillips, Roderick 1988 *Putting Asunder: A History of Divorce in Western Society*. New York: Cambridge University Press.
- Stetson, Dorothy M., and Gerald C. Wright, Jr. 1975 "The Effects of Law on Divorce in American States." *Journal of Marriage and the Family* (August):537-547.
- U.S. Bureau of the Census 1975 *Historical Statistics of the United States: Colonial Times to 1970. Part I*. Washington, D.C.: U.S. Government Printing Office.
- 1989 "Child Support and Alimony: 1985." *Current Population Reports*. P-23, No. 154.
- 1989 "Studies in Marriage and the Family." *Current Population Reports*. P-23, No. 162.
- 1995 "Who Receives Child Support?" *Statistical Brief*. [www.census.gov/socdemo/www/chldsupp.html](http://www.census.gov/socdemo/www/chldsupp.html)
- 1998a "Money Income in the United States: 1997." *Current Population Reports*. P-60, No. 200.
- 1998b "Poverty in the United States: 1997." *Current Population Reports*. P-60, No. 201.
- 1998c "Household and Family Characteristics: 1997." *Current Population Reports*. P-20, No. 515.
- 1998d "Marital Status and Living Arrangements: 1997." *Current Population Reports*. P-20, No. 506.
- Weiss, Robert 1975 *Marital Separation*. New York: Basic Books.
- 1976 "The Emotional Impact of Marital Separation." *Journal of Social Issues* 32:135-145.
- 1979 *Going it Alone: The Family Life and Social Situation of the Single Parent*. New York: Basic Books.

STEVEN L. NOCK  
ALISON BURKE

## DRAMATURGY

See Symbolic Interaction Theory.

## DRUG ABUSE

Drug abuse has been a major social problem in the United States for almost a century and we are now in the second decade of a continuing war on drugs. Drug abuse is a health and criminal justice problem that also has implications for nearly every facet of social life. It is a major element in the high cost of health care, a central reason for the United States's extraordinarily high rate of incarceration, and a focus of intensive education and treatment efforts. Substance abuse is an equal-opportunity problem that affects both high- and low-income persons, although its consequences are most often felt by those persons and communities that have the lowest social capital.

*Substance abuse*, with its connotations of disapproval or wrong or harmful or dysfunctional usage of mood-modifying substances, is a term that was developed in the United States. The more neutral term, *dependence*, is often used in other countries. *Addiction*, which formerly communicated the development of tolerance after use and a physical withdrawal reaction after a drug became unavailable, has assumed less explicit meanings. Whatever terminology is employed, there is intense societal concern about the use of psychoactive mood-altering substances that involve loss of control. This concern is manifest particularly for young people in the age group most likely to use such substances. Society is concerned that adolescents

and young adults, who should be preparing themselves for crucial educational, vocational, and other significant life choices, are instead diverted by the use of controlled substances.

The United States has the highest rate of drug abuse of any industrialized country and, not surprisingly, spends more public money than any other country to enforce laws that regulate the use of psychoactive drugs. Its efforts to control drug abuse reach out across its borders. The United States also plays a critical role in developing knowledge about substance abuse; more than 85 percent of the world's drug abuse research is supported by the National Institute on Drug Abuse.

### EPIDEMIOLOGY

Information on incidence and prevalence of drug use and abuse derives from a range of sources: surveys of samples of households and schools; hospital emergency room and coroners' reports; urine testing of samples of arrestees; treatment programs; and ethnographic studies. Such epidemiological information enables us to assess drug abuse programs and decide on allocation of resources (Winick 1997).

Since World War II, the peak years for illicit drug use were in the late 1970s, when approximately 25 million persons used a proscribed substance in any thirty-day period. Overall illicit drug use has been declining since 1985. The yearly National Household Survey on Drug Abuse, which is the most influential source of epidemiology data, reported that in 1997 marijuana was used by 11.1 million persons or 80 percent of illicit drug users (Office of Applied Studies 1999). Sixty percent only used marijuana but 20 percent used it along with another illicit substance. During the 1990s, the rate of marijuana initiation among youths aged twelve to seventeen reached a new high, of approximately 2.5 million per year. The level of current use of this age group (9.4 percent) is substantially less than the rate in 1979 (14.2 percent).

Twenty percent of illicit drug users in 1997, ingested a substance other than marijuana in the month preceding the interviews. Some 1.5 million Americans, down from 5.7 million in 1985, used cocaine in the same period; the number of crack users, approximately 600,000, has remained nearly constant for the last ten years. At least 408,000

individuals used heroin in 1997, with the estimated number of new users at the highest level in thirty years.

Data on incidence and prevalence of use must be interpreted in terms of social structure. Thus, one out of five of the American troops in Vietnam were addicted to heroin, but follow-up studies one year after veterans had returned to the United States found that only 1 percent were addicted (Robins, Helzer, Hesselbrook et al. 1980). In Vietnam, heroin use was typically found among enlisted men and not among officers. Knowing such aspects of social setting and role can help in understanding the trends and can contribute to understanding the use of other substances in other situations. In any setting, the frequency of substance use, the length of time over which it was taken, the manner of ingestion, whether it was used by itself or with other substances, its relationship to criminal activity and other user characteristics (e.g., mental illness), the degree to which its use was out of control, the setting, and whether it was part of a group activity are also important.

Rates of use by subgroup can vary greatly. Thus, for example, prevalence rates of drug use are higher among males than females and highest among males in their late teens through their twenties. Over half the users of illicit drugs work full time. About one-third of homeless persons and more than one-fourth of the mentally ill are physically or psychologically dependent on illicit drugs. The first survey of mothers delivering liveborns, in 1993, found that 5.5 percent had used illicit drugs at some time during their pregnancy. A survey of college students reported that in the previous year, 26.4 percent had used marijuana and 5.2 percent had used cocaine. National Household Survey data indicate that use of illicit drugs by persons over thirty-five, which was 10.3 percent in 1979, jumped to 29.4 percent by 1991 and was 33.5 percent in 1997.

Rates of cigarette smoking are of interest because of their possible relationship to the use of other psychoactive substances. Approximately one-eighth of cigarette smokers also use illicit drugs. In a typical month in 1997, 30 percent of Americans, or 64 million, had smoked cigarettes and one-fifth of youths between the ages of twelve and seventeen, were current smokers. Almost half of all American adults who ever smoked have stopped

smoking. Drug abusers may also be involved with alcohol.

### POLICY

A central contributor to current American policy toward mood-modifying drugs was the Harrison Act of 1914, which prevented physicians from dispensing narcotics to addicts (Musto 1987). The Marijuana Tax Act of 1937 and strict penalties for sale and possession of narcotics that were imposed by federal legislation in 1951 and 1966 expanded punitive strategies. An important change took place in 1971, when President Nixon—who had campaigned vigorously against drug use—established a national treatment network. Nixon was the only president to devote most of the federal drug budget to treatment; his successors have spent most of the budget on law enforcement.

In 1972, the Commission on Marijuana and Drug Abuse recommended a dual-focused policy that is both liberal and hard-line. The policy, which continues to the present, is liberal in that users who need help are encouraged to obtain treatment. But it is hard-line because it includes harsh criminal penalties for drug possession and sales. As a result, nearly two-thirds of the federal resources devoted to drug use are now spent by the criminal justice system to deter drug use and implement a zero-tolerance philosophy.

President Carter's 1977 unsuccessful attempt to decriminalize marijuana was the only effort by a national political leader to lessen harsh penalties for drug possession. Between 1981 and 1986, President Reagan doubled enforcement budgets to fight the "war on drugs." Politicians generally have felt that the traditional hard-line policy served their own and the country's best interests and there has been limited national support for legalization or decriminalization (Evans and Berent 1992).

Originating in several European countries, the policy of *harm reduction* has, during the last decade, generated growing interest in the United States as a politically viable alternative to legalization (Heather, Wodak, Nadelmann et al. 1998). It attempts to understand drug use nonevaluatively in the context of people's lives and to urge that the policies that regulate drug use should not lead to more harm than the use of the substance itself causes. A representative harm-reduction initiative

is the establishment of needle exchanges, for injecting users of heroin and other drugs, in order to minimize the possibility of HIV transmission resulting from the sharing of infected needles. The use of needles to inject illegal substances has been linked to one-third of the cumulative number of AIDS cases in the United States. In the United States, the use of federal government money for needle exchanges is prohibited, although there are approximately 1.3 million injecting drug users. Critics of these programs believe that such exchanges increase heroin use and send a latent message that it is acceptable to use drugs like heroin. Harm reductionists disagree and argue that needle exchanges lead to a decline in rates of HIV infection without encouraging use.

Another policy disagreement between America and other countries involves marijuana. In the United States many federal benefits, including student loans, are not available to those convicted of marijuana crimes. In contrast, marijuana has been decriminalized in a number of Western European countries, including Italy, Spain, and Holland. It is openly available in coffee houses in Holland, where officials believe that its use is relatively harmless and can deter young people from using heroin or cocaine. In America, marijuana is viewed by federal authorities as possibly hazardous and a potential "stepping stone" to heroin or cocaine use, and approximately 695,000 persons were arrested for its possession in 1997.

Other countries have experimented with ways to make drugs such as heroin legally available, albeit under control. Thus, in Switzerland, heroin addicts have been legally maintained. In England, methadone (a heroin substitute) can be obtained by prescription from a physician. In the United States, by contrast, an addict must enroll in a program to be able to receive methadone.

In the United States prevention of drug abuse has never been as important a policy dimension as treatment or law enforcement, in part because it requires legislators to commit resources in the present to solve a future problem. Prevention has, thus, accounted for less than one-seventh of the drug abuse budget. Because of the variety of prevention approaches and because of the American local approach to education, there are many viewpoints on how to conduct programs that will prevent young people from becoming drug users and

abusers. An information-didactic approach, often with the assistance of law enforcement personnel, has been traditional. A role-training, peer-oriented, values-clarification, alternatives, affective-education approach emerged in the 1970s, along with psychological inoculation. Addressing the social structure and family in which young people live, and targeted community action, attracted substantial support in the 1980s and 1990s.

National policy toward drug use is systematically promulgated by the Office of National Drug Control Policy (1999). The office has established the goal of reducing drug use and availability by 50 percent and reducing the rate of related crime and violence by 30 percent by 2007. It is proposed that these goals will be achieved by expanding current approaches.

Although drug abuse has been called "the American disease," physicians have had little impact on policy. Between 1912 and 1925, clinics in various states dispensed opiates to users. More recently, however, the federal government has opposed making marijuana available for medicinal purposes, even to treat persons with terminal or debilitating illnesses. Nevertheless, eleven states decriminalized marijuana possession in the 1970s and others, by referendum vote in the 1990s, have permitted physicians to recommend and patients to use marijuana medically.

### CONTROL

In the United States, programs to control the supply of mood-modifying substances are intended to interdict the importation of illicit materials, enforce the laws, and cooperate with other countries that are interested in minimizing the availability of controlled substances. In addition to illicit substances (such as heroin, that has no established medical use), prescription products can be abused. These include substances such as barbiturates, that are used without medical supervision in an inappropriate manner. The problem also includes over-the-counter drug products that are not used for the purpose for which they were manufactured. Some nondrug substances like airplane model glue and other inhalants that can provide a "high" and are difficult to regulate, are also considered part of the country's substance abuse burden.

Preventing illicit drugs from entering the United States is difficult because of heavily trafficked, long, porous borders. Large tax-free profits provide incentives for drug entrepreneurs to develop new ways to evade customs barriers, process the drug for the market, and sell it (Johnson, Goldstein, Preble et al. 1985). For example, approximately seven-eighths of the retail price represents profit after all costs of growing, smuggling, and processing cocaine for illegal sale in the United States. The increasing globalization of the world economy further facilitates the international trade in illicit substances.

A key component in efforts to reduce the supply of stimulants, depressants, and hallucinogens is the Comprehensive Drug Abuse and Control Act of 1970, which established a national system of schedules that differentiated the public health threat of various drugs of abuse. This law, which has been modified over the years, classifies controlled substances into five categories, based on their potential for abuse and dependency and their accepted medical use. Schedule I products, such as peyote, have no acceptable safe level of medical use. Schedule II products, such as morphine, have both medicinal value and high abuse potential. Schedule III substances, such as amphetamines, have medical uses but less abuse potential than categories I or II. Also acceptable medicinally, Schedule IV substances, such as phenobarbital have low abuse potential, although the potential is higher than Schedule V products, such as narcotics that are combined with non-narcotic active ingredients. Conviction for violation of federal law against possession or distribution of scheduled products can lead to imprisonment, fines, and asset forfeiture.

Ever since it assumed a major role in promoting the Hague Opium Convention of 1912, the United States has been a leader in the international regulation of drugs of abuse. The United States convened the 1961 Single Convention on Narcotic Drugs and the 1971 Convention on Psychotropic Substances. Some countries, like England and Holland, subscribe to the treaties but interpret them more liberally than does the United States. The United States has also provided technical assistance, financing, and encouragement to other countries to minimize the growth of drugs such as cocaine and marijuana. Programs have been conducted in Mexico and Turkey to eradicate these

crops and related programs have been encouraged in Peru and Columbia.

Since 1930, the U.S. Treasury Department has had responsibility for drug regulation in the United States. In 1973, the Drug Enforcement Administration of the Justice Department assumed the police and control function under federal law. Each state has laws that generally parallel the federal laws on possession and distribution of controlled substances and all states have a single state agency that coordinates other programs related to drug abuse.

Regular or frequent drug users, without outside income, are likely to engage in a range of criminal activities in order to buy controlled substances. They typically engage in six times more criminal activity when using than when they are not using drugs. Urine testing of arrestees, under the federal Arrestees Drug Abuse Monitoring program, indicates that some two-thirds of those arrested in urban communities had used an illicit substance prior to arrest. It is, thus, not surprising that the rates of street crimes tend to be positively correlated with the number of illegal drug users in a community.

During the last fifteen years, both state and federal prison populations have experienced a massive increase due to the number of people convicted and jailed for selling or using drugs. Other developments contributing to the surge in the prison population include aggressive enforcement, longer sentences, the decline of parole, and mandatory sentencing procedures that provide less latitude for judges. Thus, for example, federal penalties for possession of crack, a rock-like form of cocaine that became popular in the 1980s and sells for a low price on the street, are 100 times greater than for powdered cocaine. Sellers targeting crack to urban minorities represent one of several factors that have led to a disproportionate number of young blacks in federal and state prisons, for violation of possession laws. Ninety percent of prisoners in federal prisons for crack violations are black, although twice as many whites as blacks use it.

Survey and other data consistently report that the use of mood-modifying drugs is distributed among all the socioeconomic and ethnic groups in the United States; nevertheless, arrests, convictions, deaths, and other negative outcomes of

drug use are disproportionately concentrated in specific geographic areas and population subgroups. In state prisons, blacks make up some 60 percent of the drug-law violators although they represent 12 percent of the country's population and 15 percent of regular drug users. Selective enforcement of the laws might reasonably be considered a possible contributor to such statistics.

American attitudes toward drug use have historically reflected ethnic and class-related prejudices. Thus, earlier in the twentieth century, negative attitudes toward cocaine were associated with the hostility that Southern blacks, among whom cocaine use was thought to be widespread, were believed to harbor toward whites. The public's suspicion of Chinese immigrants was a reflection of their use of opium. A number of stereotypes about marijuana reflected beliefs about its use by Mexican immigrants and some occupations that had low status at the time, such as jazz musicians.

For members of both majority and minority groups sentenced to prison, recidivism rates are high and represent one reason that the United States has higher rates of incarceration (approaching two million) than any industrialized nation. Although treatment of former drug users in prison settings has produced some promising results, treatment opportunities in prison are scarce and have not kept pace with the growth in the population of incarcerated former users. Approximately one in eight state inmates and one in ten federal inmates have taken part in treatment since their admission to prison. On a limited basis, treatment is being offered in an effort to keep offenders from returning to prison.

## TREATMENT

The treatment of substance abuse has consistently been a lower priority than efforts to control drug abuse through interdiction and criminal sanctions, although cost-benefit studies have demonstrated that every dollar invested in treatment saves seven dollars in other costs. The federal government has usually spent more than two-thirds of its substance abuse budget (which now totals nearly \$20 billion) on such supply-reduction and criminal justice system strategies. Only a small minority of drug abusers have access to treatment, since health

insurers tend to discriminate against persons with substance abuse problems and there are inadequate treatment resources.

Current treatment for drug abuse, in addition to withdrawal, ranges from psychotherapeutic interventions (provided in both inpatient and outpatient settings), pharmacology agents, and various forms of milieu therapy. It frequently includes information on relapse prevention. Psychotherapy is often used in combination with other forms of treatment, and is provided both on an individual and group basis by therapists trained in medicine, psychology, social work, nursing and education. Pharmacological treatments include approaches, that substitute or block the effect of an abused substance, such as methadone maintenance for heroin (Ball and Ross 1991). Milieu therapies include a variety of residential programs where drug abusers can learn or relearn how to live substance free. Although some relatively short-term hospital-based programs exist (particularly for those with independent resources to pay for such services), the most common milieu consists of longer-term therapeutic communities such as Phoenix House, where drug users live in a setting in which they are closely monitored. The residents' progress through the several levels of the program's hierarchical social structure depends on their ability to implement the program's rules for "right living" (De Leon 1997).

Although often not considered a treatment, various fellowship groups deriving from the Alcoholics Anonymous model are widely used by drug abusers. Thus, for example, groups such as Narcotics Anonymous, Cocaine Anonymous, and parallel groups for spouses and parents of drug users exist in almost every community. Such groups, which have no professional staff and rely on the reinforcement of abstinence, support drug abusers in maintaining drug-free lives and help family members aid their drug-abusing relatives. Many treatment programs encourage their patients or clients to participate in such a twelve-step fellowship simultaneously with the treatment period or after treatment is completed.

Although substantial resources have been devoted to treatment outcome research, our knowledge of who does best in what treatment is limited. Particularly for cocaine, the use of which can lead to dependence in a short period of time, effective

pharmacological treatments are not available. A combination of strategies is often most effective especially if it recognizes that drug abuse is a chronic relapsing disorder that is likely to include multiple treatment failures on the way to an ultimately favorable outcome. Of the treatment approaches to drug abuse, milieu treatments have been among the most intensely studied. For those able to participate in such programs, they can have extremely high rates of relatively enduring positive outcomes. Whatever the treatment modality, it must include job readiness, habilitation and vocational rehabilitation, and other dimensions that will enable the former user to function effectively in the modern information-oriented community and economy.

### PREVENTION

In the late 1980s, the social problems associated with drug abuse, particularly in terms of the possession and sale of cocaine in urban areas, were perceived to have reached crisis dimensions and there was a marked increase in criminal justice efforts to control substance use. Another positive response was a renewed emphasis on the prevention of drug abuse and the collateral development of broad-based community strategies designed to reduce demand for illicit drugs. Currently, such demand-reduction efforts are undergoing systematic study in several long-term longitudinal public and private programs.

There has been a transformation in views of substance abuse as we have moved from a focus on individual pathology to programs that engage community institutions. Such efforts aim to change norms about substance use through the involvement of community members and the integration of the substance-abuse programs pursued by public and private agencies. The 1990s saw the expansion of community-based programs to include a broad range of institutions, including the police and courts, the voluntary sector, as well as the media (Falco 1994). Fostered by the government's Center for Substance Abuse Prevention (part of the Substance Abuse and Mental Health Administration) and efforts of the Robert Wood Johnson Foundation, the country's largest health foundation, hundreds of communities are engaged in broad-based efforts to change the culture within

which substance abuse takes place (Saxe, Reber, HalFors, et al. 1997).

The belief that substance abuse is sustained by community norms represents an ecological approach. Environmental conditions, whether they reflect physical conditions in the community, poverty, or available health care, are thus seen as risk factors for drug use and abuse. Supporters of this view believe that what is needed are coordinated, community-wide efforts to address drug abuse at multiple levels of social organization and the collaboration of many groups. The idea that multicomponent community-action efforts can prevent drug abuse derives from earlier studies of programs designed to cope with cardiovascular disease. It is consistent with efforts to promote a variety of other health issues, but substance abuse is now a primary focus of these efforts.

The largest of these comprehensive efforts is the federal government's Community Partnership Program, which has supported over 250 partnerships. The Community Partnership Program was initiated in 1990 after the Robert Wood Johnson Foundation had begun to develop a model and sponsor broad-based community efforts. Called "Fighting Back" programs, they now provide long-term support to more than a dozen communities to develop comprehensive demand-reduction interventions. The foundation has also provided support and technical assistance to hundreds of additional communities through groups such as Join Together.

A significant element of many such prevention programs is the presence of a strong media component. The Partnership for a Drug-Free America, for example, develops and places hundreds of millions of dollars of advertising each year, and communities are encouraged to leverage local media to present anti-drug messages directed at youth. Although there is limited direct evidence of the effectiveness of media campaigns, it is likely that they reinforce education and prevention messages being delivered to youth through other means.

Schools play a central role in prevention programs, under the assumption that drug abuse will be more easily prevented if programs are started early. The goal of these programs is to provide youth with the skills to become successful adults and to teach them the community's norms and

values. There is substantial evidence that positive school experiences are linked to lower levels of drug use and conversely, that drug use is related to delinquency and problems in school.

The role of school environments in affecting adolescent substance use has been validated by specific school-based trials. In both the Midwestern Prevention Project and Project Northland, significant reductions in the prevalence of substance use by adolescents were reported (Pentz, Dwyer, MacKinnon, et al. 1989; Perry, William, Veblen-Martenson 1996). Designed for students in grades six through eight, the programs include academic curricula, along with parental and community involvement. Often, a significant mass media component is part of the effort, with a focus on correcting misperceptions about the consequences of drug use and providing alternative positive behavior. The D.A.R.E. program (Drug Abuse Resistance Education) also has been a widely used school-based prevention strategy.

Schools are not the only public institution that affect youths' likelihood to abuse drugs. The police and justice agencies, as well as the network of health and social service agencies that serve a community, have a crucial influence and prevention activities typically involve such agencies. The ability of health and social service professionals to attend to drug use is clearly important, but their role is often reactive, providing treatment rather than prevention.

One of the most important programs that has contributed to attempts by law enforcement agencies to deal with drug abuse is community policing. It represents a shift from reactive policing where the goal is to arrest offenders, to an active strategy designed to identify crime problems and work with citizens—including offenders—to avoid further difficulties. The heart of the approach is that officers get to know citizens and help them deal with minor transgressions and, in so doing, avoid serious crime. A collateral approach, widely used in the 1990s ("fixing broken windows") is designed to improve morale and confidence and stem the physical and social deterioration of communities by prompt attention to small visible manifestations of community dysfunction or decay. There is evidence that such approaches are, at least partly, responsible for declines in violent



crime, which is closely related to substance-abuse problems.

A community's resources and social institutions have a critical impact on drug use, but the attitudes and behavior of peers and family may have an even more direct influence. The affluence of a community and the quality of its schools have a substantial effect on the initiation of drug use, but their impact is mediated by adolescents' peer relationships and their interactions with significant adults in their lives. Thus, peers and parents are perhaps the most vital elements of the community context—directing or guiding youngsters' needs and desires through the obstacles in their environment. Some of the most important programs designed to address community substance abuse focus on changing peer culture and addressing family attitudes and behavior.

Parents (or other adult “guides”) arguably have the greatest potential effect on how the youngster learns to negotiate the environment as it exists (good or bad), and they can also affect the influence that the youngsters' peers exert. The use of drugs by parents significantly increases the likelihood that their youngsters will also use drugs. Although this might seem to be a clear example of youngsters modeling the behavior of their parents, the influence of parents' own use of drugs is probably more complicated. Some research suggests that it is not merely that youngsters mimic parents' behavior, but instead such modeling interacts with what they see in their peers. If both peers and parents engage in substance use, there is far greater likelihood that young people will become regular users.

The influences of peers and parents may interact in complex ways and each community is different—its resources and institutions function differently. Communities can be directed to the key levers, but there is no simple formula available to determine which activities will be most important for a particular community. What is clear, however, is that to understand and develop strategies that reduce adolescent substance abuse, it is necessary to consider the social context in which a child lives. Only by identifying the resources available within a community, the roles played by the social institutions within that community, and the behaviors and values of the individuals (parents and

children) who live in that community, can the interactions among the multiple forms and levels of influence begin to be understood.

## RESEARCH

Social science research has played a critical role in the identification of the substance-abuse problem, its social consequences, and strategies to arrest the use of illicit drugs. There is now a long-standing tradition of surveys to identify drug use and attitudes toward the use of mood-modifying substances and their consequences. Surveys, such as the National Household Survey on Drug Abuse (which assesses the drug use of a random sample of U.S. residents over twelve years old) and Monitoring the Future (a school-based survey of junior and senior high school students), have each been conducted for more than two decades. Although there is considerable discussion about the validity of these surveys and how to ensure veridical data (Beveridge, Kadushin, Saxe, et al. forthcoming), there is no question that they have influenced social policy.

More recently, much of the focus of social research has shifted to assessing strategies to prevent drug use and to evaluate treatment programs. Under the auspices of the National Institute on Drug Abuse (a component of the National Institutes of Health), a variegated research program includes both biological and sociopsychological components. An emphasis of research is on assessment of programs such as D.A.R.E., the Community Partnership Program, and Fighting Back. Determining whether these programs achieve their goals of preventing substance abuse is a particularly difficult challenge. The programs are implemented differently across communities and the research design needs to separate the effects of race, socioeconomic status, and other factors from program implementation (Rindskopf and Saxe 1998).

It is also the case that antidrug programs develop loyal followings and their proponents develop a stake in showing that their efforts are successful. Thus, for example, there has been a major debate about the D.A.R.E. program and whether it is successful, with researchers claiming that the evidence suggests it is not effective. In other cases, such as the Community Partnership Program and Fighting Back, the issue has been the

availability of data that can show the effects of the program over time.

One development that will likely allow much better utilization of social research is the availability of sophisticated methodologies. Thus, for example, meta-analytic techniques are now available that permit the synthesis of data across multiple studies, allowing us to amalgamate multiple small-scale tests of programs. In addition, new analytic strategies are being developed to allow construction of multilevel statistical models. Such hierarchical linear modeling permits one to take account of the fact that programs are conducted in particular settings and facilitates the segregation of community effects from overall program effects. Qualitative ethnographic techniques have been used to track the life cycle of substance abuse and the structure of the illegal markets.

### FUTURE

The war on drugs is far from being “won,” but drug abuse appears to have stabilized, with use remaining nearly constant. Two trends, that could be counterreactions, have emerged and may help to shape future use of illicit drugs. The first is the call for legalization or decriminalization of the possession of drugs such as marijuana. Several national organizations have emerged to promote this goal and to urge a harm reduction approach. The second trend is the increased licit use of mood-altering prescription substances, such as Prozac and Ritalin. Such powerful psychotropic agents are being prescribed by physicians for depression, difficulties in concentration, and similar problems. As the medical options increase, misuse of prescription drugs will likely increase and it may be more difficult to control the sale of less powerful nonprescription drugs.

(SEE ALSO: *Alcohol*)

### REFERENCES

- Ball, J. C. , and A. Ross 1991 *The Effectiveness of Methadone Maintenance Treatment*. New York: Springer-Verlag.
- Beveridge, A., C. Kadushin, L. Saxe, D. Rindskopf, and D. Livert forthcoming “Survey Estimates of Drug Use Trends in Urban Communities: General Principles and Cautionary Examples.” *Substance Use and Misuse*.

- De Leon, G. (ed.) 1997 *Community As Method: Therapeutic Communities for Special Populations and Special Settings*. Westport, Conn.: Praeger.
- Evans, R., and I. Berent 1992 *Drug Legalization: For and Against*. La Salle, Ill.: Open Court Press.
- Falco, M. 1994 *The Making of A Drug Free America: Programs That Work*. New York: Times Books.
- Gerstein, D. R., and H. J. Harwood (eds.) 1990 *Treating Drug Problems*. Washington, D.C.: National Academy Press.
- Heather, N., A. Wodak, E. Nadelmann, et al. 1998 *Psychoactive Drugs and Harm Reduction*. London, U.K.: Whurr Publishers.
- Johnson, B. D., P. Goldstein, E. Preble, et al. 1985 *Taking Care of Business: The Economics of Crime by Heroin Abusers*. Lexington, Mass.: Lexington Books.
- Kleiman, M. A. R. 1992 *Against Excess: Drug Policy for Results*. New York: Basic Books.
- Musto, D. F. 1987 *The American Disease: Origins of Narcotic Control*. New York: Oxford University Press.
- Office of Applied Studies 1999 *National Household Survey on Drug Abuse Main Findings 1997*. Rockville, Md.: Substance Abuse and Mental Health Services Administration.
- Office of National Drug Control Policy 1999 *The National Drug Control Strategy 1999*. Washington D.C.: The White House.
- Pentz, M. A., J. H. Dwyer, D. P. MacKinnon, et al. 1989 “A Multi-Community Trial for Primary Prevention of Adolescent Drug Use: Effects on Drug Use Prevalence.” *Journal of the American Medical Association* 261:3259-3266.
- Perry, C. L., C. L. Williams, S. Veblen-Martenson, et al. 1996 “Project Northland: Outcome of A Community Wide Alcohol Use Prevention Program for Early Adolescence.” *American Journal of Public Health* 86:956-965.
- Rindskopf, D., and L. Saxe 1998 “Zero Effects in Substance Abuse Programs: Avoiding False Positives and False Negatives in the Evaluation of Community-based Programs.” *Evaluation Review* 22:78-94.
- Robins, L. N., J. E. Helzer, M. Hesselbrook, et al. 1980 “Vietnam Veterans Three Years After Vietnam.” In L. Brill and C. Winick, eds., *Yearbook of Substance Use and Abuse*, vol. 2, 213-232. New York: Human Sciences Press.
- Saxe, L., E. Reber, D. Halfors, C. Kadushin, D. Jones, D. Rindskopf, and A. Beveridge 1997 “Think locally, act globally: Assessing the impact of community-based substance abuse prevention.” *Evaluation and Program Planning* 20:357-366.

Winick, C. 1983 "Addicts and Alcoholics as Victimizers."  
In D.E.J. Mac Namara and A. Karmen, eds., *Deviants:  
Victims or Victimizers?* Beverly Hills, Calif.: Sage  
Publications.

—— "Epidemiology." In J.H. Lowinson, P. Ruiz, R.B.  
Millman, et al. eds., *Substance Abuse: A Comprehensive*

*Textbook* 3rd ed. 10–16. Baltimore, Md.: Williams and  
Wilkins.

——(ed.) 1974 *Sociological Aspects of Drug Dependence*.  
Cleveland, Oh: CRC Press.

LEONARD SAXE  
CHARLES WINICK

# Encyclopedia of Sociology

---

Second Edition

# Encyclopedia of Sociology

---

Second Edition

VOLUME 2

**Edgar F. Borgatta**

*Editor-in-Chief*

*University of Washington, Seattle*

**Rhonda J. V. Montgomery**

*Managing Editor*

*University of Kansas, Lawrence*

**Macmillan Reference USA**

*an imprint of the Gale Group*

*New York • Detroit • San Francisco • London • Boston • Woodbridge, CT*

# E

---

## ECOLOGY

See Demography; Environmental Equity; Environmental Sociology; Human Ecology and Environmental Analysis.

## ECONOMIC DETERMINISM

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

Economic determinism refers to a kind of causality in which an economic variable  $x$  causes a condition of behavior  $y$ . This statement of direct causality contains very little actual economic determinism. But in stating that economic condition  $x$  is the most determining factor in causing behavior  $y$ , we have a model for economic determinism that is quite common in economics and sociology. This model appears in Weber's (1979) *Economy and Society*; in his discussion of domination he states:

*Nor does domination utilize in every case economic power for its foundation and maintenance. But in the vast majority of cases, and indeed in the most important ones, this is just what happens in one way or another and often to such an extent that the mode of applying economic means for the purpose of maintaining domination, in turn exercises a determining influence on the structure of domination. (p. 942)*

The same model of economic determinism appears in Louis Althusser's (1970) *For Marx*, where the social formation has multiple determinants, but "the economy is determinant in the last instance" (p. 113). Neither model has the economic as a monocausal determinant of society, but economic categories, such as the economic market for Weber, clearly are part of a central structuring determination for society.

This model predates both Weber and Althusser and has its origins in eighteenth-century free market liberalism. In *The Federalist Papers*, James Madison assumes economic interests as the chief motivation of the people. In *The Wealth of Nations*, for Adam Smith it is "that in commercial society every man thus lives by exchanging, or becomes in some measure, a merchant" (1976, p. 26). For Smith people are buyers and sellers involved in production and consumption, and human behavior is an unending series of economic exchanges. Individuals pursue their own self-interest in rational ways, and their own self-interest consists of profit, which is regulated by competition in a predominantly self-regulating free market. In the pursuit of economic self-interest the individual promotes the social good, "led by an invisible hand to promote an end which was no part of his intention" (p. 477). In Smith's model an individual's pursuit of economic profit automatically structures the good for all of society.

Smith assumes an economic person who is always acting to optimize economic advantages. People here are determined by economic motives, that is, to increase profits and decrease losses.

Modern neoclassical economists, such as Milton Friedman and Gary Becker, have continued the tradition of Adam Smith. For them the only limitation on the free market model is the amount of information to which a rational actor has access. Since information is not perfect, mistakes can occur. But in the pursuit of profit a rational actor learns even from mistakes; mistakes therefore increase the amount of information that a rational actor acquires, thus increasing the chances of making correct choices in the future. Thus, the free market not only structures human behavior but determines the inevitability that these actors, in rationally pursuing profit, will acquire the information needed to produce profits continuously for themselves and to mold—even if unintentionally—the social good.

The model of economic determinism offers sociology an easily quantifiable object, a phenomenon that can be subjected to scientific procedures—observed, measured, tested, and verified. As a positivist social science, sociology requires a transhistorical and universal phenomenon such as the physical sciences have, and economic determinism provides it in economic concepts such as the market, money, the circulation of capital, and so forth. It allows sociologists to speak the language of science and to reduce all social phenomena to mathematical formulas.

All positivist social sciences use mathematical representations of social reality to understand society, and economic determinism is but one attempt at creating a scientific sociology. Still, the model has offered sociology a formula that possesses broad powers for explaining social factors and avoids problems of indeterminate multiple causes. This is what science traditionally means by lawfulness. These lawlike social categories are representations of social reality and provide a framework in which the aggregate behavior of individuals can be structured in terms of economic interests. This behavior is patterned and can be studied and subjected to scientific procedures. Thus, the model of economic determinism informs sociological analysis and research.

Exchange theory provides an example of a positivist methodology that is based on the model of economic determinism. In exchange theory, economic exchange is the determinant principle of behavior. George C. Homans, the originator of

this form of analysis, combines this economic model with a behaviorist psychology, but the economic is the determining structure. In *Social Behavior: Its Elementary Forms*, Homans states, “Human behavior as a function of its payoff: in amount and kind it depends on the amount and kind of reward and punishment it fetches” (1961, p. 13). Human relations have been reduced to the exchange and circulation of commodities. Homans assumes that a social actor is an economic person existing in a free market where individuals make rational choices to maximize profits and reduce costs. This is stated partly in the language of behavioral psychology, but the conditions of economic exchange are dominant. Thus, Homans sounds more like Adam Smith than like B. F. Skinner when he writes “we define psychic profits as rewards less costs, and we argue that no exchange continues unless both parties are making a profit” (1961, p. 61).

This model allows Homans to analyze individuals who live in a group, make numerous rational choices, and yet live a stable, patterned life. Individuals calculating their possibilities and making rational choices are led by something like Smith’s “invisible hand” to maintain group life in “practical equilibrium.”

Homans believed that his analysis was a value-neutral attempt to expand the scope of a scientifically valid sociology of human behavior. But, while exchanges are an important aspect of human interaction, the reduction of all human behavior to elementary exchanges is problematic. Homans has universalized the relations of individuals in the capitalist marketplace, relations that are historically specific and not the basis of a universal psychology of human behavior. Homans’s exchange theory is a conservative sociological theory in which pecuniary relations structure human behavior.

Another form of sociological analysis that uses the model of economic determinism is Marxist sociology. In Weber’s *The Protestant Ethic and the Spirit of Capitalism*, Marx’s analysis is called a “one-sided materialistic” interpretation (1958, p. 183). Marx provides evidence for Weber’s contention when he argues that the economic base determines the ideological superstructure. In the preface to *A Contribution to the Critique of Political Economy*, Marx writes that “the totality of these relations of production constitutes the economic

structure of society, the real foundation, on which arises a legal and political superstructure and to which correspond definite forms of social consciousness" (1970, p. 20). This famous passage, containing what is known as the base-superstructure model, became the hallmark statement for economic determinist Marxists from the Second International (Marxist Workers Congress) in 1889 to the present.

But Marx himself was not an economic determinist, even though many of his followers were. Marx theorized about a world in which human relations were subsumed under capitalist relations of production. His central concern was revolutionary change, which depended on the formation of a revolutionary class that could wage war against the dominant classes. But for Marx, class was determined not by economic conditions alone but by community and cultural conditions as well. Thus, he writes in *The Eighteenth Brumaire of Louis Bonaparte* on whether "small holding peasants" are or are not a class:

*In so far as millions of families live under economic conditions of existence that separate their mode of life, their interests and their culture from those of other classes and put them in hostile opposition to the latter, they form a class. In so far as there is merely interconnection among these small holding peasants, and the identity of their interests begets no community, no national bond and no political organization among them, they do not form a class. (Marx 1963, p. 124)*

This is not "economics in the last instance" but a multivalent fitting together of a series of necessary conditions. Economic conditions are insufficient, and class formation and class struggle do not occur unless community occurs as well.

Later Marxists differ from Marx on a number of issues. First, many understand the base-superstructure model as a determinant condition of class. That is, class depends exclusively on the economic base for its formation. For Marx, however, it was an insufficient condition and a "moment" in his analysis of capitalism. Furthermore, economic determinist Marxists view class struggle as a secondary phenomenon, even though it is important. Finally, many Marxists have critiqued Marx's concept of class for its inability to predict revolutionary change with scientific accuracy. Erik

Olin Wright, for instance, calls Marx's concept of class "vague" and "random"; it is much too relativistic for use as a neat, lawlike scientific formula. In his important book *Classes*, Wright preserves the Marxian tradition of the Second International and attempts to erect a positivist Marxist sociology whose propositions can be empirically verified. It is upon the base-superstructure model that Wright builds this science. As in exchange theory, Wright's model suggests that in capitalism rational actors make rational choices in pursuing their economic advantages. But Wright also attempts to build a scientific base for an analysis of class struggle, and in *Classes* he provides the causal link between capitalist exploitation and the actions of individuals in society. He does this by demonstrating that class structure is determined by property relations, a central determinant in modern society:

*Class structure is of pervasive importance in contemporary social life. The control over society's productive assets determines the fundamental material interests of actors and heavily shapes the capacities of both individuals and collectivities to pursue their interests. The fact that a substantial portion of the population may be relatively comfortable materially does not negate the fact that their capacities and interests remain bound up with property relations and the associated processes of exploitation. (Wright 1985, pp. 285–286)*

Economic factors are crucial for understanding the social world. But society is not reducible to economic determinants; it is too complex to be reduced to a set of economic propositions, or any single determinant set of propositions, that imply unilinear causality. Still, economic determinism is seductive because it offers a theory that has broad explanatory powers, is easily quantifiable, and can be reduced to simple mathematical formulas. Thus, it is not surprising to see it continue and expand, seemingly unaware of its limitations. For instance, George Gilder writes in *Wealth and Poverty*, "The man's earnings, unlike the woman's, will determine not only his standard of living but also his possibilities for marriage and children—whether he can be a sexual man" (1981, p. 109). The reduction of sex and marriage to economic determinants is highly problematic, both as science and as common sense. Yet generalizations about social life wholly based on economic causes persist, under the names of both science and myth. One can



restate Max Weber's warning against monocausal theories in the social sciences:

*But it is, of course, not my aim to substitute for a one-sided materialistic an equally one-sided spiritualistic causal interpretation of culture and of history. Each is equally possible, but each if it does not serve as the preparation, but as the conclusion of an investigation accomplishes equally little in the interest of historical truth.* (1958, p. 183)

(SEE ALSO: *Capitalism; Economic Sociology; Marxist Sociology*)

## REFERENCES

- Albritton, Robert 1995 "Political Economy in the Making: A Response." *Rethinking Marxism* 8:125–129.
- Aronowitz, Stanley 1988 *Science as Power: Discourse and Ideology in Modern Society*. Minneapolis: University of Minnesota Press.
- Dahms, Harry F. 1997 "Theory in Weberian Marxism: Patterns of Critical Social Theory in Lukacs and Habermas." *Sociological Theory* 15:181–214.
- Feenberg, Andrew 1992 "Subversive Rationalization: Technology, Power, and Democracy." *Inquiry* 35:3–4.
- Gilder, George 1982 *Wealth and Poverty*. New York: Bantam.
- Hodgson, Geoffrey M. 1995 "Varieties of Capitalism from the Perspectives of Veblen and Marx." *Journal of Economic Issues* 29:575–584.
- Kennedy, Paul M., Andrew Levine, Elliott Sober, and Erik-Olin Wright 1987 *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House.
- Homans, George C. 1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace, and World.
- Marx, Karl (1859) 1970 *A Contribution to the Critique of Political Economy*.
- (1852) 1963 *The Eighteenth Brumaire of Louis Bonaparte*. New York: International.
- Rammert, Werner 1997 "New Rules of Sociological Method: Rethinking Technology Studies." *British Journal of Sociology* 48:171–191.
- Resnick, Stephen A., and Richard D. Wolff 1987 *Knowledge and Class: A Marxian Critique of Political Economy*. Chicago: University of Chicago Press.
- Sherman, Howard J. 1998. "Critique of the Critique: Analysis of Hodgson on Marx on Evolution." *Review of Social Economy* 56:47–58.
- Smith, Adam 1976 *An Inquiry into the Nature and Causes of the Wealth of Nations*. Chicago: University of Chicago Press.
- Weber, Max (1947) 1979 *Economy and Society: An Outline of Interpretive Sociology*. Guenther Roth and Claus Wittich, eds. and trans. Berkeley: University of California Press.
- Wright Erik, O., Andrew Levine, and Elliott Sober 1992 *Reconstructing Marxism: Essays on Explanation and the Theory of History*. New York: Verso.
- 1994 "Historical Materialism: Theory and Methodology." *Science and Society* 58:53–60.
- (1904–1905) 1958 *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribners.
- Wright, Erik Olin 1985 *Classes*. London: Verso.

WILLIAM DiFAZIO

## ECONOMIC DEVELOPMENT

See Industrialization in Less-Developed Countries, Modernization Theory; Rural Sociology .

## ECONOMIC INSTITUTIONS

The analysis of economic institutions is central to the work of the classical figures of sociology—Marx, Weber, and Durkheim. These thinkers did not recognize a boundary line between sociological inquiry and economic inquiry; on the contrary, their efforts to make sense of the development of market capitalism led them to intensive analysis of market processes. Unfortunately, this thrust of sociological inquiry was largely abandoned by sociologists between World War I and the late 1960s. This was particularly true in the United States, where sociologists generally deferred to economists' claims of an exclusive mandate to study economic processes.

To be sure, there were a number of important intellectual figures during this period whose work integrated sociological and economic inquiry, but these individuals were rarely housed in sociology departments. Such economists as Thorstein Veblen, Joseph Schumpeter, and John Kenneth Galbraith have been retroactively recognized as sociologists. Similarly, the largely self-educated Hungarian scholar Karl Polanyi ([1944] 1957, 1971) is now acknowledged to have made seminal contributions

to the sociological analysis of economic institutions. Yet scholars working in the sociological mainstream either ignored economic topics or tended to incorporate the perspectives of neoclassical economics.

Since the late 1960s, however, the lines of inquiry pioneered by the classical writers have been revitalized. Sociologists working in a number of different intellectual traditions and on diverse empirical topics have developed sophisticated analyses of economic processes (Swedberg 1987). No longer content to defer to the expertise of professional economists, many of these writers have developed powerful critiques of the work of neoclassical economists (Zukin and DiMaggio 1990).

Although this body of work explores a range of different economic institutions, much of it can be understood through its analysis of the way that markets work. (Analyses of other economic institutions, such as the division of labor, money, and corporate organizations, are presented elsewhere in this encyclopedia.) In particular, an emergent economic sociological conception of market processes can usefully be contrasted with the conception of the market that is implicit in most economic writings.

### THE ORGANIZATION OF MARKETS

Neoclassical economists tend to assume an ideal market situation that allows changes in prices to equilibrate supply and demand. In this ideal market situation, there are multiple buyers and sellers whose transactions are fundamentally impersonal; information on the product and the price are the only relevant variables shaping the action of market participants. Contemporary economists recognize that this ideal market situation requires a basic symmetry in the information available to buyers and sellers. When there are significant differences in information, it is likely that the resulting price will diverge from the price that would effectively equilibrate supply and demand. Nevertheless, contemporary microeconomics rests on the assumption that most markets approximate the ideal situation, including information symmetry (Thurow 1983)

The sociological view of markets is fundamentally different. It stresses the embeddedness of behavior within markets, the central role of imitation in structuring markets, and the importance of blocked exchanges. The concept of embeddedness challenges the idea that impersonality is an important feature of actual market situations. While the individual actor in economic theory is a rational actor who is able to disregard his or her social ties in the market situation, the sociological actor is seen as embedded in a network of social relations at the time that he or she engages in market transactions. This embeddedness means that a wide range of social ties exert continuing influence over how the actor will both make and respond to price signals (Polanyi [1944] 1957; Granovetter 1985).

This view has two elements. First, the individual actor is decisively influenced by social ties. For example, a consumer might choose not to do business with retailers belonging to a stigmatized ethnic group, even when their prices are lower, because members of the consumer's ethnic group genuinely believe that the products of the other group will be inferior or that contact with the stigmatized group will jeopardize one's social position. The economists' argument that such an individual has a "taste for discrimination" does not adequately capture a reality in which discriminatory behavior often occurs with very little reflection because beliefs are deeply rooted. Second, the individual actor's dependence on social ties is necessary in order for him or her to accomplish a given economic goal. As Granovetter (1985) has pointed out, a purchasing officer at a corporation might well do business with a particular supplier regardless of price considerations because of longstanding ties to that individual. These ties provide assurance that the delivery will occur in a timely fashion and the merchandise will meet established quality standards. In other words, social bonds can provide protection against the uncertainties and risks that are always involved in transactions.

The standard economic view of markets tends to ignore these uncertainties; it is generally assumed that individuals will automatically obey the

rules that make transactions possible. As Oliver Williamson (1975) has noted, this represents a profound inconsistency in economic analysis. While individuals are assumed to be self-interested, they are also expected to avoid guile and deception. In the real world, however, every transaction involves the risk that one party is deliberately cheating the other, and the more “impersonal” the transaction—that is, the less one party knows about the other—the greater the risk becomes.

This is one of the reasons that actual markets tend to develop structures and rules designed to constrain and to embed individual behavior. Although commodity markets and stock markets most closely resemble the pure markets of economic theory, in which rapid price changes serve to balance supply and demand, these markets tend to evolve complex social structures. At one level, such markets appear to be completely impersonal, in that there is no contact between buyers and sellers. On another level, however, the actual transactions are handled by a community of brokers who are well known to each other and who are expected to follow a particular etiquette for managing transactions. This structure has evolved to provide protection against unknown brokers who might prove unreliable and to assure that known brokers will be discouraged from cheating their colleagues and clients (Adler and Adler 1984; Baker 1984a; Burk 1988).

The central point, however, is that the particular ways in which market behaviors are embedded have real and significant consequences. On the one hand, embeddedness allows market processes to go forward by diminishing the opportunities for guile and deceit. On the other hand, embeddedness assures that factors besides price will influence the behavior of market participants, so it can no longer be assumed that price will automatically equilibrate supply and demand.

One of the main implications of the concept of embeddedness is that actual markets will be characterized by imitative behavior. Economists assume that each economic actor will calculate his or her preference schedule independently of all other actors. In the sociological view, however,

actors are continually making their choices in reference to the behavior of others. In deciding the price to be asked for a particular commodity, a market participant will set it in comparison with the prices of competitors (White 1981). This is one of the key reasons that in actual markets there is often less price competition than would be suggested either by economic theory or by considerable differences across firms in the costs of production.

This role of imitation plays a particularly important role in financial panics and bubbles. Panics occur when many holders of a particular kind of asset rush to convert their holdings to cash, leading to precipitous price declines. Bubbles occur when enthusiasm for a particular asset drives prices far higher than can be justified by expected returns (Kindleberger 1978). Economists have trouble explaining the behavior of individuals in most panics or bubbles because rational actors with complete information would not engage in such irrational behavior; they would understand that the underlying value of an asset is not likely to change so dramatically in a short period of time. However, actual individuals have limited information, and they cope with uncertainty by observing the behavior of their friends and neighbors. Such observation makes them quite susceptible to the collective enthusiasms of panics and bubbles.

The threat of panics is another reason that financial markets develop institutional structures to embed individual behavior. The New York Stock Exchange, for example, has an elaborate system of specialist firms who have the responsibility to smooth out the market for particular stocks. Such firms are expected to be purchasers of last resort in situations where there are too many sellers of a stock (Baker 1984b). The idea is that such action should help to reduce the likelihood of panic. While such an institutional arrangement diverges from the economists’ conception of a self-regulating market, it makes sense in the context of imitative behavior.

Furthermore, in their emphasis on the virtues of markets, neoclassical economists often suggest a vision of society as a giant bazaar in which

everything is for sale. Sociologists, in contrast, are more likely to recognize that the viability of markets depends upon a wide variety of restrictions that block certain types of exchanges. Blocked exchanges are transactions that are in violation of the law or of widely shared ethical standards (Walzer 1983). Instances of blocked exchange include prohibitions on the resale of stolen merchandise, laws that prevent government officials from selling their services to the highest bidder, and the criminalization of prostitution and the purchase of certain drugs.

More subtle blocked exchanges play an important role in structuring the economy. Lawyers are not allowed to switch sides in the middle of a civil suit in response to an offer of a higher fee by the other side. Accountants are not supposed to produce a more favorable audit in response to a higher fee (Block 1990). Loan officers at a bank are prohibited from approving loans in exchange for side payments from the applicant. The members of a corporate board of directors are supposed to place their fiduciary responsibility to the shareholders ahead of their self-interest in contemplating offers to buy out the firm. In a word, there is a complicated and sometimes shifting boundary between legitimate and illegitimate transactions in contemporary economies.

The importance of blocked exchanges sheds further doubt on arguments that markets can regulate themselves without governmental interference (Polanyi [1944] 1957). The construction of any particular market involves rules as to what kinds of exchange are open and what kinds are blocked, but the incentives to violate these rules are often quite substantial. At the same time, the incentives for market participants to police each other are often lacking. Hence, there is often no alternative to a governmental role in policing the boundary between legitimate and illegitimate exchanges. Moreover, debates about social policy are often framed in a language of individual rights that is profoundly insensitive to the importance and pervasiveness of blocked exchanges. For example, when a woman serving as a surrogate mother in exchange for a fee decides she wants to keep the baby, the issue is often debated in terms of contract law. The more fundamental issue, however, is

whether the society believes that the rental of wombs is a legitimate or illegitimate transaction (Rothman 1989).

### VARIETIES OF MARKET SYSTEMS

A second important dimension of contemporary sociological work on economic institutions emphasizes the significant variations in the ways in which different market societies are institutionally organized. Against the common assumption that market societies will converge toward the same institutional arrangements, this work has used the comparative method to highlight significant and durable institutional differences. Out of this has come a rich body of literature on the wide variety of different "capitalisms" that can coexist at a particular moment in time. While there are a variety of typologies, this work has tended to identify an East Asian model, a Rhinish model that draws heavily on Germany and France, a social democratic model associated particularly with Sweden and Norway, and the Anglo-American model (Couch and Streeck 1997; Hollingsworth and Boyer 1997; Orru et al. 1997). While researchers have identified a wide variety of differences in economic institutions, much of this work has focused on examining differences in labor markets, capital markets, and in the markets for innovation.

The study of labor-market institutions encompasses variations in how the labor force acquires basic and advanced skills; variations in how social welfare provides citizens with income in the event of sickness, disability, unemployment, retirement, and other contingencies; variations in the rights and responsibilities that employees have in the workplace; and variations in the mechanisms for matching employees with job vacancies (Rogers and Streeck 1995; Wever and Turner 1995). The point is not simply that different countries are located on different points of a continuum in terms of variables such as union density or welfare generosity. It is rather that some of these variables tend to be grouped together, so that one can identify ideal typical complexes of institutions through which societies distribute some of the costs and benefits of economic growth. For example, studies by Dore (1997) and Streeck (1997)

have sought to show how Japanese and German systems of industrial relations both differ dramatically from the Anglo-American model and work in tandem with other specific institutional features of those societies to create significant economic advantages for Japanese and German firms.

Comparative work on capital markets has generally contrasted the stock market-centered system in the United States and the United Kingdom with the bank-centered systems that have prevailed in Japan and in Continental Europe (Mizruchi and Stearns 1994; Zysman 1983). In the former firms rely primarily on stock and bond issues in impersonal markets to raise the funds required for expansion, while in the latter firms are more reliant on long-term loans from banks. This difference has important implications for corporate governance, for the size of the debt burden on corporations, for the time horizons of corporate executives, and for the openness of the economy to new entrepreneurial initiatives.

Finally, studies of the market for innovation have identified distinct varieties of what some analysts have called national innovation systems (Amable et al. 1997). This encompasses the research efforts of scientists and engineers, initiatives by government agencies to foster innovation, and the entrepreneurial activities of both public and private firms. The ways in which these elements are brought together vary significantly both in terms of the division of labor between the public sector and the private sector and the way that interactions between them are organized. Since innovation clearly weighs heavily in international economic success, some analysts have begun to examine the specific institutional features of nations that have been most effective in fostering successful innovation (Evans 1995).

This effort to identify different institutional types of capitalism has occurred during the same period in which the countries of Eastern Europe and the former Soviet Union have been making a transition from state socialism to capitalism. While some Western advisers acted as though the Anglo-American model was the only relevant model for former socialist countries to emulate, the process

of transition had the effect of deepening the awareness of the institutional variations among capitalist societies (Stark and Bruszt 1998). Hungary, for example, had to make a set of decision as to how much weight to give the stock market in corporate governance and in capital allocation. This, in turn, generated increasing interest in studies showing the important differences in the relative role of stock markets and banks in different market societies.

### THE SCOPE OF MARKETS

A third important dimension of contemporary sociological work on economic institutions is a concern with the geographical scope of markets. Much of the sociological tradition has been oriented to the study of national societies or of subnational units. But an understanding of the scope of international markets calls that approach into question. Markets for raw materials, finished products, services, capital, and labor cross national boundaries and exert extraordinary influence over all aspects of social life (Swedberg 1987).

The most ambitious effort to date to chart the importance of these international markets has been the world-system theory of Immanuel Wallerstein (1974a, 1974b, 1980, 1989). Instead of using national societies as the basic unit of analysis, Wallerstein has sought to shift sociological analysis to the level of the capitalist world-system. For Wallerstein, this system is comprised of a world market and a competitive state system of divided sovereignties. Any analysis of patterns within a particular national society must begin by locating that society within the larger capitalist world-system. A nation's location at the core, the periphery, or the semiperiphery of the capitalist world-system can be expected to shape the nature of its economic and political institutions.

The existence of a single unified world market is central to Wallerstein's argument. Each nation that is part of the capitalist world-system must struggle for relative advantage in that market, and this has implications for both social relations within nations and for the political-military relations among nations. Moreover, it is a central part of

Wallerstein's project to detail the process by which various regions of the world have been incorporated into this unified world-system. The capitalist world-system began as a purely European phenomenon, but through colonization and the penetration of Western influence, this system became global. Regions that were once external to the system were progressively incorporated as peripheral areas that produced raw materials. Later-arriving nations such as the United States, Canada, and Japan moved to the semiperiphery and then to the core.

The contributions of Wallerstein and his followers have been extremely important in showing the systematic ways in which international markets shape developments within societies. The study of labor markets within the world-system, for example, has been particularly fruitful in making sense of international migrations—the massive movements of people across national boundaries—that loom so large in understanding both core and peripheral societies (Portes and Walton 1981).

Wallerstein's work, however, is vulnerable to criticism because it sometimes implies that there is a single unified world market for commodities, for labor, and for capital that operates quite similarly to the markets of economic theory. Critics have suggested that Wallerstein devotes too little attention to analyzing the specific institutional structures of global markets and that his work has not given sufficient weight to the institutional variations among capitalist societies. The response to the criticism is that the homogenizing logic of a global capitalist system will work itself out over time and will ultimately eliminate much of the variety in institutional forms.

Toward the end of the 1990s, debate over this precise issue has become increasingly important. The attention of a growing number of scholars is focused on the interactions between the varieties of market institutions at the national level and the dynamics of a global capitalist system. Particularly in discussions of "globalization" (see article in this encyclopedia), scholars ask whether there are—as Wallerstein suggests—powerful structural forces operating at the global level that are systematically

forcing all market societies to adopt Anglo-American economic institutions. One British scholar posed the question polemically by asking if there were a new type of Gresham's Law in which "bad capitalisms were driving out good" (Gray 1999).

One set of events that intensified interest in this question were the deep problems of the Japanese economy during the 1990s and the Asian economic crisis that began in Thailand in July of 1997 and spread to Malaysia, Indonesia, South Korea, and ultimately to Russia and Brazil. Enthusiasm for an East Asian model of capitalism was considerably dampened by the severity of Japan's long-term problems, by South Korea's overnight transition from success story to severe financial crisis, and by doubts that the model could facilitate sustainable growth in poorer countries such as Thailand and Indonesia. At the same time, the crisis heightened awareness that the institutions of global capitalism, especially the International Monetary Fund, were exerting powerful pressures on East Asian countries to align themselves more fully with Anglo-American ways of organizing their economic institutions. Most specifically, in the aftermath of the crisis, there were pressures on nations to shift from bank-centered finance to stock market-centered finance (Wade and Veneroso 1998).

Finally, the Asian crisis also drew attention to the power of what one observer has labeled the "electronic herd" (Friedman 1999)—the global network of traders in markets for bonds, equities, foreign exchange, and more exotic financial instruments. When significant numbers of these traders decide that a particular economy on any continent is being poorly managed, they have the capacity to precipitate a major financial crisis just by taking positions in the market that assume a depreciation of the value of the country's financial assets and currency (Block 1996). To be sure, this power is not yet universal; China and Vietnam have been able to insulate themselves from these pressures by maintaining controls on capital movements, and Hong Kong and Malaysia were successful in foiling the speculative strategies of the traders. Nevertheless, the power wielded by these traders is further evidence that the global market might be able to force steadily greater homogenization in economic institutions and practices across nations.

These same concerns extend to the more developed economies of Western Europe. Scholars have argued that the rapid movement of capital across national boundaries has fundamentally weakened the social democratic model as an alternative way of organizing market societies (Scharpf 1991). And others have begun to wonder aloud as to how long the German model can resist the pressures to abandon its particular institutions for organizing labor markets and capital markets (Gray 1999; Streeck 1997). Developments such as the cross-border merger between Daimler-Benz and Chrysler might portend German accommodation to the Anglo-American model.

The question of whether the global economy will increasingly converge to a single type of capitalism based on the Anglo-American model or whether there will be continuing diversity of the institutional forms of market societies is far from being resolved. The outcome will depend on economic and political developments over the next few decades. However, if there is convergence, the results are likely to be highly unstable unless there is significant progress in increasing the regulatory capacities of global economic institutions. If one starts from neoclassical economic premises, one can imagine that a self-regulating global economy with relatively weak international institutions could be viable. However, a sociological view of markets evokes deep skepticism about the stability of such a system (Polanyi [1944] 1957). Just as sustained economic activity within nations requires both embeddedness of markets to minimize guile and deceit and enforcement of blocked exchanges, so an increasingly integrated global economy has similar needs. Moreover, since globalization tends to weaken regulatory mechanisms at the national level, the need for effective enforcement at the global level becomes more acute.

The Asian economic crisis of 1997 and 1998 revealed that the International Monetary Fund lacks the resources and the capacity to contain international financial panics. Moreover, since the global economy lacks an effective "lender of last resort," a financial panic could lead to a cascade of failures by large financial institutions on an enormous scale. But even when the need to expand

global economic regulation is recognized, the political obstacles to the strengthening of international financial and regulatory institutions are enormous. Most national governments are extremely reluctant to cede sovereignty to international agencies. Since the same processes of globalization that increase the need for global regulation have already eroded the powers of national governments, political leaders are reluctant to limit their own authority even further. At the same time, while the United States has thrown its political and economic weight behind the forces of globalization, domestic political opposition in the United States to strengthened international agencies continues to be formidable. Hence, it appears unlikely that the globalized economy will soon have the regulatory institutions that it needs to function effectively.

(SEE ALSO: *Corporate Organizations*; *Division of Labor*; *Economic Sociology*; *Money*; *Transnational Corporations*)

## REFERENCES

- Adler, Patricia, and Peter Adler, eds. 1984 *The Social Dynamics of Financial Markets*. Greenwich, Conn.: JAI Press.
- Amable, Bruno, Rémi Barré, and Robert Boyer 1997 "Diversity, Coherence, and Transformations of Innovation Systems." In Rémi Barré, Michael Gibbons, Sir John Maddox, Ben Martin, Pierre Papon et al., eds., *Science in Tomorrow's Europe*. Paris: Economica International.
- Baker, Wayne 1984a "The Social Structure of a National Securities Market." *American Journal of Sociology* 89:775-811.
- 1984b "Floor Trading and Crowd Dynamics." In Patricia Adler and Peter Adler, eds., *The Social Dynamics of Financial Markets*. Greenwich, Conn.: JAI Press.
- Block, Fred 1990 *Postindustrial Possibilities: A Critique of Economic Discourse*. Berkeley: University of California Press.
- 1996 "Controlling Global Finance." *World Policy Journal* (Fall): 24-34.
- Burk, James 1988 *Values in the Marketplace: The American Stock Market Under Federal Securities Law*. New York: Walter de Gruyter.
- Couch, Colin, and Wolfgang Streeck, eds. 1997 *Political Economy of Modern Capitalism: Mapping Convergence and Diversity*. Thousand Oaks, Calif.: Sage.

- Dore, Ronald 1997 "The Distinctiveness of Japan" In Colin Crouch and Wolfgang Streeck, eds., *Political Economy of Modern Capitalism: Mapping Convergence and Diversity*. Thousand Oaks, Calif.: Sage.
- Evans, Peter 1995 *Embedded Autonomy: States and Industrial Transformation*. Princeton, N.J.: Princeton University Press.
- Friedman, Thomas 1999 *The Lexus and the Olive Tree*. New York: Farrar, Straus & Giroux.
- Granovetter, Mark 1985 "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.
- Gray, John 1999 *False Dawn: The Delusions of Global Capitalism*. New York: New Press.
- Hollingsworth, J. Rogers, and Robert Boyer 1997 *Contemporary Capitalism: The Embeddedness of Institutions*. New York: Cambridge University Press.
- Kindleberger, Charles P. 1978 *Manias, Panics, and Crashes*. New York: Basic Books.
- Mizuchi, Mark S., and Linda Brewster Stearns 1994 "Money, Banking, and Financial Markets." In Neil J. Smelser and Richard Swedberg, eds., *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press.
- Orru, Marco, Nicole Woolsey Biggart, and Gary G. Hamilton 1997 *The Economic Organization of East Asian Capitalism*. Thousand Oaks, Calif.: Sage.
- Polanyi, Karl (1944) 1957 *The Great Transformation*. Boston: Beacon Press.
- 1971 *Primitive, Archaic and Modern Economies*. Boston: Beacon Press.
- Portes, Alejandro, and John Walton 1981 *Labor, Class, and the International System*. New York: Academic Press.
- Rogers, Joel, and Wolfgang Streeck, eds. 1995 *Works Councils: Consultation, Representation and Cooperation in Industrial Relations*. Chicago: University of Chicago Press.
- Rothman, Barbara Katz 1989 *Recreating Motherhood: Ideology and Technology in a Patriarchal Society*. New York: Norton.
- Scharpf, Fritz 1991 *Crisis and Choice in European Social Democracy*, trans. Ruth Crowley and Fred Thompson. Ithaca, N.Y.: Cornell University Press.
- Stark, Davis and Laszlo Bruszt 1998 *Postsocialist Pathways: Transforming Politics and Property in East Central Europe*. Cambridge, U.K.: Cambridge University Press.
- Streeck, Wolfgang 1997 "German Capitalism: Does it Exist? Can it Survive?" In Colin Crouch and Wolfgang Streeck, eds., *Political Economy of Modern Capitalism: Mapping Convergence and Diversity*. Thousand Oaks, Calif.: Sage.
- Swedberg, Richard 1987 "Economic Sociology: Past and Present." *Current Sociology* 35:1-221.
- Thurow, Lester 1983 *Dangerous Currents: The State of Economics*. New York: Random House.
- Wade, Robert, and Frank Venoroso 1998 "The Asian Crisis: The High Debt Model Versus the Wall Street-Treasury-IMF Complex." *New Left Review* 228 (March-April): 3-25.
- Wallerstein, Immanuel 1974a *The Modern World-System, I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.
- 1974b "The Rise and Future Demise of the Capitalist World-System: Concepts for Comparative Analysis." *Comparative Studies in Society and History* 16:387-415.
- 1980 *The Modern World-System, II: Mercantilism and the Consolidation of the European World-Economy*. New York: Academic Press.
- 1989 *The Modern World-System, III: The Second Era of Great Expansion of the Capitalist World-Economy, 1730-1840s*. New York: Academic Press.
- Walzer, Michael 1983 *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.
- Wever, Kirschen S., and Lowell Turner, eds. 1995 *The Comparative Political Economy of Industrial Relations*. Madison, Wis.: Industrial Relations Research Association.
- White, Harrison 1981 "Where Do Markets Come from?" *American Journal of Sociology* 87:517-547.
- Williamson, Oliver 1975 *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Zukin, Sharon, and Paul DiMaggio 1990 *Structures of Capital: The Social Organization of the Economy*. New York: Cambridge University Press.
- Zysman, John 1983 *Governments, Markets, and Growth: Financial Systems and the Politics of Industrial Change*. Ithaca, N.Y.: Cornell University Press.

FRED BLOCK

## ECONOMIC SOCIOLOGY

Economic sociology constitutes its own distinct subfield in sociology and can be briefly defined as



*the sociological analysis of economic phenomena.* Economic sociology has a rich intellectual tradition and traces its roots to the founding fathers of sociology, especially to Max Weber and his *Economy and Society* (see Swedberg 1998). It should be noted that not only sociologists but also economists have made important contributions to economic sociology. This is particularly true for today's economic sociology, which is the result of works not only by sociologists (such as Mark Granovetter and Harrison White) but also by economists (such as Gary Becker and Oliver Williamson).

To define *economic sociology* as “the sociological analysis of economic phenomena” may seem bland and even tautological. It is therefore important to stress that it entails a definite conception of what topics may be studied by sociologists; that it implies a certain division of labor between economists and sociologists; and that it also has direct consequences for how the relationship between economic theory and sociology is conceived. That this is the case becomes very clear if we contrast this definition with two other ones that are commonly used: (1) that economic sociology primarily deals with a particular dimension of economic phenomena, namely their social dimension; and (2) that economic sociology is the study of social structures and organizations in the economy. That economic sociology deals with economic phenomena in general (our definition) means that it addresses issues not only at the periphery of the economy (such as, say, the influence of religious values on the economy or of ethnicity on entrepreneurship) but also at its core (such as the way markets operate or investment decisions are made). Sociological theory here emerges as either an alternative to economic theory or as a direct challenge to it. To look at the social dimension of economic phenomena (the first alternative definition) means, on the other hand, that sociologists only look at a limited number of economic issues, usually those that are left over once the economists have finished with their analyses. Economists may, for example, decide with the help of standard economic theory what salaries and prices are like in a certain industry, while sociologists, by looking at a factory or a work group as a social system, may

then add some additional information. Economic theory is not challenged by this type of economic sociology, since it only deals with those topics for which there is no economic theory. That economic sociology focuses on social structures or on organizations in the economy (the second alternative definition) means that a purely economic analysis may be regarded as economic sociology as long as it deals with certain topics. Why a firm rather than the market is used for a specific type of transaction may, for example, be explained by the fact that transaction costs are higher in this specific case in the market. This type of economic sociology is close to economic theory and basically dispenses with traditional sociology (although not necessarily with rational choice sociology; see, e.g., Coleman 1990).

These three ways of looking at economic sociology all have their followers. The one which emphasizes that the sociological perspective in principle can be applied to *all* types of economic phenomena is, however, the one that has been used most frequently throughout the history of economic sociology. That this is the case will become clear from the following brief overview of the field. That the two other definitions—economic sociology as the analysis of the social dimension of economic phenomena, and economic sociology as the study of social structures and organizations in the economy—also have their adherents will become obvious as well.

Since the mid-1980s economic sociology has been going through something of a renaissance in the sociological profession, not only in the United States but also in other countries. The advent of what is usually referred to as “new economic sociology” represents one of the most dynamic areas in contemporary sociology. Before the mid-1980s three separate attempts had been made to create a vigorous economic sociology, and something needs to be said about these. The first attempt was made in the early twentieth century by a group of German scholars of whom Max Weber is the most important. The second attempt was made during the same time period by Emile Durkheim and his followers in France. And the third attempt was made by some American sociologists, such as

Talcott Parsons and Neil Smelser, in the 1950s. A few words shall be said about each of these attempts before we discuss the contemporary situation.

### HISTORICAL ATTEMPTS AT ECONOMIC SOCIOLOGY

The first significant attempt to create a solid economic sociology was made in Germany during the period 1890–1930 by a group of scholars who were all trained in economics. The three key figures were Max Weber, Werner Sombart, and Joseph Schumpeter. A major reason that economic sociology developed so forcefully in German-speaking academia was probably its strong tradition of historical economics. There was also the fact that toward the end of the nineteenth century Gustav von Schmoller, the leader of the historical school of economics, became embroiled in a bitter academic fight with Carl Menger, one of the founders of marginal utility analysis. By the time Weber and Sombart became active, German economics had been polarized into two camps through the so-called battle of the methods, or the *Methodenstreit*: one that was overly theoretical and one that was overly historical. The idea of “economic sociology” was conceived by both Sombart and Weber as an attempt to get out of this dead end and to function as a kind of bridge between economic history and economic theory. Economic sociology should be analytical in nature, but historically grounded. While Sombart, however, wanted economic sociology to totally replace economic theory, Weber thought differently. In his mind, a healthy science of economics (Weber used the term *Sozialökonomie*, or “social economics”) should be broad and simultaneously draw on economic theory, economic history, and economic sociology (Weber 1949). Schumpeter basically shared Weber’s opinion, although economic theory would always rank higher in his mind than in Weber’s. The idea of such a broad-based social economics, however, never caught on.

Weber, Sombart, and Schumpeter all made a series of first-rate contributions to economic sociology. For one thing, all of them produced major studies of capitalism: Weber ([1921-22] 1978) in

*Economy and Society*; Sombart ([1902] 1987) in *Der moderne Kapitalismus*; and Schumpeter ([1942] 1976) in *Capitalism, Socialism and Democracy*. Weber emphasized that capitalism was becoming increasingly rationalized; Sombart was particularly interested in looking at the different historical stages of capitalism; and Schumpeter argued that modern capitalism was digging its own grave and was soon to be replaced by socialism. These visions of capitalism still dominate our thinking and are therefore of great interest. And so are many of the shorter studies by Weber, Sombart, and Schumpeter, such as Sombart’s ([1906] 1976) study of why there is no socialism in the United States, Weber’s ([1904-5] 1930) analysis of the relationship between Protestantism and the spirit of capitalism, and Schumpeter’s ([1919] 1954, [1918] 1971) two superb articles on imperialism and the tax state.

A special mention must also be made of Georg Simmel’s ([1907] 1990) *The Philosophy of Money*. This work contains an ingenious analysis of money that ranges from philosophy to sociology. No general theory of money is developed, but the author takes on a series of interesting topics, including credit, checks, and small change. Simmel should not only be credited with having made a serious attempt to develop a sociological approach of money; he was also the first sociologist to realize what an important role trust plays in economic life.

The only one to make a sustained effort to lay a theoretical foundation for economic sociology, however, was Max Weber. He did this in a chapter of *Economy and Society* (Weber [1921-22] 1978) entitled “Sociological Categories of Economic Action.” When Weber lectured on this chapter to his students, they found his analysis abstract and dry. He therefore decided to give a lecture course in economic history to supplement his theoretical ideas. This course became what is today known as *General Economic History* (Weber [1923] 1981), and it should be read together with *Economy and Society*. In the latter work Weber carefully constructs the various analytical categories that are needed in economic sociology. He starts with “the concept of economic action” and ends with macroeconomic phenomena, such as “market economies and planned economies.” He also defines and discusses such basic concepts as trade, money, and

the market—all from a sociological perspective. At various points in his discussion Weber carefully underlines when economic theory and economic sociology differ. It is, for example, imperative for economic sociologists to use the concept of economic power in their analyses, while this plays no role in marginal utility theory. In economic theory it is assumed that consumers are price givers, but economic sociology assumes that they are price takers. In economic theory it is usually assumed that prices are simply the result of demand and supply, while in economic sociology it is necessary to look at the strength of the various social groups in order to understand the unfolding of the “price struggle.” Finally, in economic sociology economic action must in principle be oriented to the behavior of others. Economists exclusively study rational economic action, Weber concludes, while sociologists have a much broader focus.

During about the same time that Weber, Sombart, and Schumpeter were active in Germany, a similar, though independent, effort to create an economic sociology was made in France. The key figures here are Emile Durkheim, Marcel Mauss, and François Simiand. All three felt that since economic theory is not a social theory (in the sense that it does not assign analytical priority to society as opposed to the individual), it should be replaced by a sociological approach to the economy or, more precisely, by economic sociology. In this they echoed Auguste Comte’s critique in the early 1800s of economic theorists for ignoring the fact that the economy is part of society and that, as a consequence, there is no need for a separate economic theory (Swedberg 1987). The two most important studies in the French school of economic sociology are *The Division of Labor in Society* by Durkheim ([1893] 1964) and *The Gift* by Marcel Mauss ([1925] 1969) (see also Simiand 1932). The latter work not only covers gift-giving but also contains a series of brilliant remarks on credit, interest, and consumption. In *The Division of Labor in Society* Durkheim raises the question of how to bring about solidarity in industrial society. His answer, which is further elaborated in other works (see especially Durkheim [1928] 1962, [1950] 1983), is that no society in which the economic element

predominates can survive. Economic life has to be restrained by a moral element; without a common morality, all persons would be at war with one another.

Both German and French economic sociology petered out in the 1930s. At around this time European sociology was exhausting itself, while U.S. sociology was in ascendancy. Among the multiple subfields that appeared at that time, several are of interest to economic sociology, such as industrial sociology, the sociology of professions, and stratification theory. None of these, however, dealt with core economic problems or with economic theory. Instead there was a firm division of labor in U.S. social science at this time between economists, who only studied economic topics, and sociologists, who only studied social topics. In the 1950s, however, some sociologists decided to challenge this division of labor, and their efforts have become known as the “economy and society approach,” so called both because two works with this title now appeared (Moore 1955; Parsons and Smelser 1956) and because a conscious effort was made to bring closer together two bodies of thought in the social sciences—economics and sociology—that most social scientists felt should be kept separate (see also Polanyi et al. 1957). Talcott Parsons and Neil Smelser (1956) argued, for example, that the economy is part of society or, in their terminology, “the economic sub-system” is part of “the social system.” In this sense they assigned a certain priority to society and implicitly to sociology. On the other hand, they also felt that economic theory was essentially correct—even if it needed to be complemented by a sociological approach. This dual position also informs the first textbook as well as the first reader in economic sociology—both produced by Smelser (1963, 1965).

### NEW ECONOMIC SOCIOLOGY

During the late 1960s and the 1970s little of interest happened in economic sociology. Since the mid-1980s, however, there has been a sharp increase of interest in this topic, and a new type of economic sociology has come into being (see Friedland and Robertson 1990; Granovetter 1990;

Zukin and DiMaggio 1990). Not only sociologists but also economists have contributed to this development. Since the mid-1970s mainstream economists have become increasingly interested in the role of social structures and organizations in the economy. This has led to a movement usually referred to as “new institutional economics” (e.g., Eggertsson 1990). Sources of inspiration for this new institutionalism include transaction cost economics, agency–principal theory, and game theory. Gary Becker (1976), for example, has convinced many economists that social phenomena can be analyzed with the help of the economist’s tools; Kenneth Arrow has written about the role of organizations in the economy; Thomas Schelling (1960) has used game theory to develop a science of “interdependent decision”; and Oliver Williamson (1975) has popularized the concept of transaction costs through his best-selling *Markets and Hierarchies* (see also Coase 1937; Swedberg 1990). Three Nobel Prizes have also been awarded to economists who in one way or another focus on the social aspects of the economy: R.H. Coase (1991), Gary Becker (1992), and Douglass North (1993). As a result of these and other events, mainstream economists today are interested not only in traditional issues relating to price formation but also in economic institutions and how these change. The last time this happened in the United States was in the early twentieth century, when American institutionalism was born (see, e.g., Commons 1924; Gruchy 1947; Veblen [1899] 1973). There exists, however, an important difference between the old form of institutionalism and new institutional economics. While Thorstein Veblen and his contemporaries tried to analyze economic institutions with the help of an approach that was very close to that of sociology, Becker and other current theorists claim that the reason economic institutions work the way they do can be analyzed with the help of the economist’s traditional tools (efficiency, rational choice, etc.). This approach has been severely criticized by some sociologists on the grounds that it simplifies and distorts the analysis (e.g., Etzioni 1988; Granovetter 1985).

Since the mid-1980s, as already mentioned, there has been a major revival of economic sociology, and what is usually referred to as “new

economic sociology” has come into being. The date of birth of this movement is usually set to 1985, since that year a highly influential article, which was to create much interest in economic sociology, was published. This was Mark Granovetter’s “Economic Action and Social Structure: The Problem of Embeddedness,” published in *The American Journal of Sociology*. The very same year, it can be added, Granovetter introduced the notion of “new economic sociology” in a brief paper at the annual meeting of the American Sociological Association. In his 1985 article on embeddedness, Granovetter sharply attacked the attempts by economists to explain the functioning of social institutions and accused them of simplicity. Just as economists have a tendency to ignore social relations through an “undersocialized concept of man,” Granovetter said, some sociologists view the individual as a reflex of the social structure, and they consequently have an “oversocialized concept of man.” The proper way to proceed, Granovetter suggested, is to tread a middle way between these two opposites, and this can best be done by assuming that individual actions are always “embedded” in social networks.

Granovetter’s article has been followed by a minor avalanche of writings in economic sociology, and there exist good reasons for arguing that new economic sociology today constitutes a minor school of its own. A large number of articles and quite a few monographs have been produced; a couple of introductory readers can be found on the market; and in the mid-1990s a huge *Handbook of Economic Sociology* was published (Smelser and Swedberg 1994). Other signs that a certain institutionalization of economic sociology has taken place is that a section in economic sociology has been organized at the American Sociological Association, which has also published a volume with course outlines and similar teaching materials (Green and Myhre 1996).

Before saying something about the concrete studies that have been produced since the mid-1980s, it should be pointed out that new economic sociology is primarily a creation of North American sociologists. In Europe and elsewhere in the world there also exists an interest in economic

sociology, but it tends to manifest itself in a less cohesive form than in the United States, and it is not held together through recurring conferences and the like. This is especially true for Europe. Most of the major European sociologists have written on economic topics in some work or another, but this is rarely perceived as an interest in economic sociology (e.g., Boltanski 1987; Bourdieu 1986; Luhmann 1988). There also exist articles and monographs by European sociologists who identify themselves as economic sociologists—but, to repeat, these tend not to be much noticed, since they are not held together by a strong and self-conscious tradition (e.g., Beckert 1997; Dodd 1994; Gislain and Steiner 1995). Finally, quite a bit of economic sociology has also been produced under the auspices of the section on economy and society within the International Sociological Association (e.g., Martinelli and Smelser 1990).

New economic sociology has advanced the understanding of economic phenomena in a number of ways, and it has especially been successful in analyzing the following three topics: (1) the role that networks play in the economy, (2) the way that culture and values influence the economy, and (3) what causes firms to be organized the way they are. Something will be said about each of these topics, but before doing so it should be noted that some interesting advances have also been made in many other areas, such as consumption, finance, and the role of gender in the economy (e.g., Abolafia 1996; Biggart 1989; Warde 1997). Finally, social capital is a topic that has attracted attention from sociologists as well as from political scientists and economists (e.g., Bourdieu 1986a; Coleman 1988; for an overview, see Woolcock 1998).

**Network Analyses.** Network analyses are often empirical in nature and sophisticated in their methodology, and this is also true for network studies in economic sociology. This latter type of studies made its first appearance in the 1970s, something which Granovetter's well-known *Getting a Job* (1974) is a reminder of. The same is true for studies of interlocks, that is, studies of the kind of links that emerge when some individual is a member of more than one corporate board. Interlock studies became popular with Marxist sociologists, who felt that they had found a way to document how the ruling class controls corporations

(e.g., Mintz and Schwartz 1985). A more subtle version of this argument can be found in Michael Useem's *The Inner Circle* (1984), based on interviews with chief executive officers (CEOs), whose main point is that CEOs who are members of several boards have a better overview of the economy, something that enables them to better defend their interests.

The simplistic type of interlock studies have been severely criticized, primarily on the grounds that it is unclear what the consequences are of the fact that two or more corporations are connected through interlocks. In one interesting study, it was also argued that if for some reason a link between two corporations was severed, it would have to be reconstituted relatively soon if this type of link indeed is as important as is often claimed. This study showed that only a minority of so-called broken ties were actually re-created (Palmer 1983; see also the discussion in Stearns and Mizruchi 1986). As of today, the opinion of many economic sociologists is that interlock studies can be quite valuable, but only on condition that they are complemented with other material, such as historical studies, interviews, and the like.

A few words must be said about Granovetter's *Getting a Job* (1974), since it represents a particularly fine example of what an empirically sophisticated and theoretically interesting study in economic sociology can look like. As Granovetter notes in the second edition of this work from 1995, his study has inspired quite a bit of research since its original publication in the 1970s. The main thrust of the study is to challenge the notion of mainstream economics that social relations can be abstracted from an analysis of how people get jobs. Through network data he had collected in a Boston suburb, Granovetter succeeded in showing that information about openings in the job market travels through social networks, and the more networks you belong to, the more likely you are to find this type of information. Having a few very close and helpful friends is not as effective in terms of getting information as being linked to many different networks ("the strength of weak ties"). A corollary of this thesis, Granovetter shows, is that people who have had several jobs are more likely

to find a new position when they become unemployed than those who have had only one employer.

Since the mid-1980s network studies have become very popular in economic sociology, and a number of advances have been made (see the studies cited in Powell and Smith-Doerr 1994). Several new topics have also been added to the repertoire, including industrial regions and ethnic entrepreneurship. A special mention should be made of Ronald Burt's *Structural Holes* (1992), in which competition and entrepreneurship are analyzed from a network perspective. Burt's study is centered around the argument that when an actor is the one and only link between two networks, he or she is in a good position to exploit this situation (*tertius gaudens*, or "the third who benefits," in Simmel's terminology). Granovetter (1994) has also suggested that the network approach can be used to study so-called business groups, that is, the kind of social formations that are made up of corporations that are bound together in some formal or informal way and that display a certain amount of solidarity. The applicability of the notion of business groups to the Korean *chaebol* or to the Japanese *keiretsu* is obvious, but it also appears that business groups exist in most Western countries.

**The Influence of Culture and Values.** A few economic sociologists have approached the study of the economy from a different perspective and emphasize the way that culture and values influence economic phenomena. The two most prominent contributors to this type of economic sociology are Paul DiMaggio (1994) and Viviana Zelizer (1979, 1985, 1994); the studies they have produced are of two kinds—general theoretical statements and empirical studies of a historical and qualitative character. Zelizer (1988) has sharply critiqued what she sees as an attempt in much of current economic sociology to eliminate values and to reduce everything to networks. Economic sociology, she argues, needs to introduce culture and values into the analysis, while simultaneously paying attention to the social structure.

Zelizer has also produced three empirical studies in which she attempts to show the impact of culture and values on economic phenomena. In

the first of these, Zelizer (1979) looks at the development of the life insurance industry in the United States, showing how difficult it was to get people to accept that an individual's life can be evaluated in purely monetary terms. In her second study, Zelizer (1985) looks at the same development but, so to speak, in reverse—namely, how something that had an economic value at one time in history can turn into something that has a sacred value at another. In the nineteenth century, as she shows, children were often seen as having an economic value, while today they have an exclusively emotional value. In her latest study, Zelizer (1994) looks at money, arguing that people usually distinguish between different types of money. Money—and this is the main point—is not some kind of homogeneous, asocial medium, as economists claim, but is social to its very core. Pin money, for example, differs from the kind of money that is set aside for ordinary expenses; and when money is given away as a gift, an effort is usually made to disguise its nature as money.

**Organization Theory.** For a number of reasons there exists a clear affinity between organization theory and economic sociology. One reason for this, no doubt, is that sociologists of organization often analyze economic organizations; another is that organization theory was to incorporate much of industrial sociology when this field disappeared in the 1970s. And, finally, roughly during the 1990s, business schools often hired sociologists to teach organization theory. Three schools or perspectives in organization theory have been of much importance to economic sociology: resource dependency, population ecology, and new institutionalism.

The basic idea of resource dependency is that an organization is dependent on resources in its environment to survive. This perspective, as especially Ronald Burt has shown, can be of some help in understanding how the economy works. At the center of Burt's work on resource dependency is his concept of structural autonomy, or the idea that a corporation has more room to maneuver the fewer competitors it has and the more suppliers and the more customers there are. That a corporation has more power if it is in a monopoly

position is clear; from this it follows that suppliers as well as customers are less powerful the more competitors they have. If Corporation A, for example, has only one supplier and one customer, both of these can wield quite a bit of power over Corporation A. Using a huge input-output data set for U.S. industry, Burt has also shown that the idea of structural autonomy has some support in empirical reality; in brief, the more structurally autonomous a corporation is, the more likely it is that profits will increase (Burt 1983).

Population ecology, as opposed to resource dependency, uses as its unit of analysis not the single corporation but whole populations of organizations. That these populations go through fairly distinct phases of growth and decline has been shown through a number of empirical studies, many of which are highly relevant to economic sociology since the organizations being studied are often economic organizations. Population ecology also looks at competition between organizations and the processes through which new organizational forms become accepted. The fact that population ecology typically looks at large populations of organizations means that relatively high-powered statistical methods are used. There is, however, little theoretical renewal going on in population ecology, and unless this changes, this perspective risks being exhausted in a few years.

A considerably higher degree of flexibility and creativity characterizes new institutionalism, or the kind of organization theory that has emerged around the work of John Meyer (e.g., Meyer and Rowan 1977; cf. DiMaggio and Powell 1991). A fundamental thesis in this approach is that rationality is often only a thin veneer and that organizations usually look the way they do for other than rational reasons. There also exist more or less distinct models for what a certain type of organization should look like, and these models are typically diffused through imitation. Since new institutionalism has such a flexible core, it can be used to analyze a variety of topics, in contrast to population ecology, which is considerably more limited in scope.

Two studies that illustrate this flexibility are Neil Fligstein's *The Transformation of Corporate Control* (1990) and Frank Dobbin's *Forging Industrial*

*Policy* (1994). The former is a study of the huge American corporation since the end of the nineteenth century that challenges several of Alfred Chandler's theses. According to Fligstein, U.S. corporations have created different concepts of control during different periods of time; by control, he means the general strategy that corporations follow for surviving and making money. While cartels, for example, represented a common strategy around the turn of the century in the United States, they were later replaced by vertical integration, the idea of conglomerates, and other concepts of control. Fligstein also shows not only that the famous multidivisional form was a response to the economic environment, as Chandler claimed, but also that it was diffused through imitation.

Fligstein, as opposed to Chandler, also points out that the state influences the way corporations operate and the way they decide on a certain concept of control. Dobbin makes a similar point in *Forging Industrial Policy* (1994), but the emphasis in this study is primarily on regulatory or industrial policy cultures. Drawing on empirical material of a historical character from France, England, and the United States in the nineteenth century, Dobbin shows how each of these countries developed different regulatory and industrial policy cultures, and in particular how they treated railroads in different ways. The state, for example, was actively involved in the railroad business in France but played a more passive role in England and the United States. Dobbin argues convincingly that there exists no single best way of doing things in the economy, as mainstream economists seem to think; what may seem natural and rational to do in one country does not seem so in another.

New economic sociology has also made some interesting progress in the analysis of the market. The reason this topic has attracted quite a bit of attention among sociologists is that the theory of the market constitutes the very heart of mainstream economics; and to challenge mainstream economics one first and foremost has to challenge its theory of the market. Of the empirical studies that sociologists have produced, the most innovative may well be Mitchell Abolafia's *Making Markets* (1996) (see also Uzzi 1996). Abolafia has investigated three important markets on Wall Street

(bonds, stocks, and futures markets) through participant observation; in particular, he has looked at the way that these are regulated. His major conclusion is that markets are social constructions and that regulation is related to "cycles of opportunism." When the existing regulation of a market is mild, opportunistic actors will take advantage of this fact, which will lead to a tightening of the rules; when regulation has been strong and effective for some time, demands are likely to be raised that milder rules should be introduced.

While most empirical studies of the market have focused on some aspect of the market rather than on its core, there do exist a few theoretical attempts by sociologists to explain the very nature of the market. Two of these are particularly interesting, namely, the analyses of Harrison White (1981) and of Neil Fligstein (1996). White's argument, which takes its departure in the typical production market with only a handful of actors, can be summarized in the following way: When a few actors produce similar products at similar prices, they may, by watching one another, come to realize that they make up a market and also behave according to this perception. More precisely, it is by watching the terms-of-trade schedule that this process takes place; and as long as the producers feel that they fit into this schedule, the market will continue to exist. By modeling his argument about the terms-of-trade schedule, White is also able to show under which theoretical conditions a market can come into being and when it will unravel.

While only a few attempts have been made to work directly with White's so-called W(y)-model, its general impact has been large in new economic sociology, especially through White's argument that a market comes into being when actors orient their behavior to one another in a rolelike manner. The most suggestive of the studies that have been influenced in a general way by the W(y)-model is Fligstein's theory of markets. Like White, Fligstein uses the typical production market as his point of departure, but the emphasis in his theory is quite different. Market actors, according to Fligstein, fear competition, since this makes it hard to predict what will happen, and they therefore attempt

to introduce stability into the market. This can be done in different ways, and for empirical illustration Fligstein draws on his study of the evolution of the huge American corporation (Fligstein 1990). In certain situations, competition can nonetheless be very strong, but this is usually accompanied by attempts to stabilize the market. As examples of this, Fligstein mentions the situation when a new market is coming into being, when a major innovation is introduced into an already existing market, and when some major social disturbance takes place.

### POSSIBLE FUTURE DIRECTIONS

If one were to summarize the situation in economic sociology at the end of the twentieth century, it could be said that economic sociology, which played such an important role in the classic works of sociology, has once again come alive. New and provocative studies have been produced, and a steadily growing number of sociologists are becoming interested in economic sociology. If one adds to this that mainstream economists are increasingly realizing the importance of institutions in the economy, it may well be the case that economic sociology will become one of the most interesting fields in sociology during the twenty-first century.

### REFERENCES

- Abolafia, Mitchell 1996 *Making Markets: Opportunism and Restraint on Wall Street*. Cambridge, Mass.: Harvard University Press.
- Arrow, Kenneth 1974 *The Limits of Organization*. New York: W.W. Norton.
- Becker, Gary 1976 *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- Beckert, Jens 1997 *Grenzen des Marktes*. Frankfurt: Campus Verlag.
- Biggart, Nicole Woolsey 1989 *Charismatic Capitalism: Direct Selling Organizations in America*. Chicago: University of Chicago Press.
- Boltanski, Luc 1987 *The Making of a Class: Cadres in French Society*. Cambridge, U.K.: Cambridge University Press.
- , and Laurent Thevenot 1987 *Les économies de la grandeur*. Paris: Presses Universitaires de France.



- Bourdieu, Pierre 1986a "The Forms of Capital." In John G. Richardson, ed., *Handbook of Theory and Research for the Sociology of Education*. Westport, Conn.: Greenwood Press.
- 1986b *The Distinction*. Cambridge, Mass.: Harvard University Press.
- Burt, Ronald 1983 *Corporate Profits and Cooptation: Networks of Market Constraints and Directorate Ties in the American Economy*. New York: Academic Press.
- 1992 *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- Coase, Ronald H. 1937 "The Nature of the Firm." *Economica N.S.* 4:385–405.
- Coleman, James 1988 "The Role of Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94:S95–S120.
- 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Commons, John R. 1924 *The Legal Foundations of Capitalism*. New York: Macmillan.
- DiMaggio, Paul 1994 "Culture and the Economy." In Neil Smelser and Richard Swedberg, eds., *Handbook of Economic Sociology*. Princeton, N.J., and New York: Princeton University Press and Russell Sage Foundation.
- , and Walter Powell, eds. 1991 *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Dobbin, Frank 1994 *Forging Industrial Policy: The United States, Britain and France in the Railway Age*. Cambridge, U.K.: Cambridge University Press.
- Dodd, Nigel 1994 *The Sociology of Money: Economics, Reason and Contemporary Society*. London: Polity Press.
- Durkheim, Emile (1928) 1962 *Socialism*. New York: Collier Books.
- (1893) 1964 *The Division of Labor in Society*. Glencoe, Ill.: Free Press.
- (1950) 1983 *Professional Ethics and Civic Morals*. Westport, Conn.: Greenwood Press.
- Eggertsson, Thráinn 1990 *Economic Behavior and Institutions*. Cambridge, U.K.: Cambridge University Press.
- Etzioni, Amitai 1988 *The Moral Dimension: Towards a New Economics*. New York: Free Press.
- Fligstein, Neil 1990 *The Transformation of Corporate Control*. Cambridge, Mass.: Harvard University Press.
- 1996 "Markets as Politics: A Political-Cultural Approach to Market Institutions," *American Sociological Review* 61:656–673.
- Friedland, Roger, and A. F. Robertson, eds. 1990 *Beyond the Marketplace: Rethinking Economy and Society*. New York: Aldine.
- Gislain, Jean-Jacques, and Philippe Steiner 1995 *La sociologie économique 1890–1920*. Paris: Presses Universitaires de France.
- Granovetter, Mark (1974) 1995 *Getting a Job: A Study in Contacts and Careers*. Cambridge, Mass.: Harvard University Press.
- 1985 "Economic Action and Social Structure: A Theory of Embeddedness." *American Journal of Sociology* 91:481–510.
- 1990 "The Old and the New Economic Sociology: A History and an Agenda." In Roger Friedland and A. F. Robertson, eds., *Beyond the Marketplace*. New York: Aldine.
- 1994 "Business Groups." In Neil Smelser and Richard Swedberg, eds., *Handbook of Economic Sociology*. Princeton, N.J. and New York: Princeton University Press and Russell Sage Foundation.
- Green, Gary P., and David Myhre 1996 *Economic Sociology: Syllabi and Instructional Materials*. Washington, D.C.: American Sociological Association Teaching Resources.
- Gruchy, Allan G. 1947 *Modern Economic Thought: The American Contribution*. New York: Prentice-Hall.
- Luhmann, Niklas 1988 *Die Wirtschaft der Gesellschaft*. Frankfurt: Suhrkamp Verlag.
- Martinelli, Alberto, and Neil Smelser, eds. 1990 *Economy and Society: Overviews in Economic Sociology*. London: Sage.
- Mauss, Marcel (1925) 1969 *The Gift: Forms and Functions of Exchange in Archaic Societies*. London: Cohen & West.
- Meyer, John, and Brian Rowan 1977 "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Sociological Review* 83:340–363.
- Mintz, Beth, and Michael Schwartz 1985 *The Power Structure of American Business*. Chicago: University of Chicago Press.
- Moore, Wilbert E. 1955 *Economy and Society*. New York: Doubleday.
- Palmer, Donald 1983 "Broken Ties: Interlocking Directorates, the Interorganizational Paradigm, and Intercorporate Coordination." *Administrative Science Quarterly* 28:40–55.
- Parsons, Talcott, and Neil Smelser 1956 *Economy and Society: A Study in the Integration of Economic and Social Theory*. London: Routledge and Kegan Paul.

- Polanyi, Karl, Conrad M. Arensberg, and Harry W. Pearson, eds. 1957 *Trade and Market in the Early Empires*. Glencoe, Ill.: Free Press.
- Powell, Walter, and Laurel Smith-Doerr 1994 "Networks and Economic Life." In Neil Smelser and Richard Swedberg, eds., *Handbook of Economic Sociology*. Princeton, N.J., and New York: Princeton University Press and Russell Sage Foundation.
- Schelling, Thomas 1960 *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.
- Schumpeter, Joseph A. (1918) 1954 "The Crisis of the Tax State." *International Economic Papers* 4:5–23.
- (1919) 1971 "The Sociology of Imperialisms." In *Imperialism and Social Classes*. New York: Meridian Books.
- (1942) 1976 *Capitalism, Socialism and Democracy*. New York: Harper & Row.
- Simiand, François 1932 *Le Salaire, l'Evolution Sociale et la Monnaie*. Paris: Alcan.
- Simmel, Georg (1907) 1990 *The Philosophy of Money*, 2nd enlarged ed. London: Routledge.
- Smelser, Neil 1963 *The Sociology of Economic Life*. Englewood Cliffs, N.J.: Prentice-Hall.
- , ed. 1965 *Readings on Economic Sociology*. Englewood Cliffs, N.J.: Prentice-Hall.
- , and Richard Swedberg, eds. 1994 *Handbook of Economic Sociology*. Princeton, N.J., and New York: Princeton University Press and Russell Sage Foundation.
- Sombart, Werner (1906) 1976 *Why Is There No Socialism in the United States?* New York: Sharpe.
- (1902) 1987 *Der moderne Kapitalismus*. Munich: Deutscher Taschenbuch Verlag.
- Stearns, Linda Brewster, and Mark Mizruchi 1986 "Broken-Tie Reconstitution and the Function of Interorganizational Interlocks: A Reexamination." *Administrative Science Quarterly* 31:522–538.
- Swedberg, Richard 1987 "Economic Sociology: Past and Present." *Current Sociology* 35(1):1–221.
- 1990 *Economics and Sociology: On Redefining Their Boundaries. Conversations with Economists and Sociologists*. Princeton, N.J.: Princeton University Press.
- 1998 *Max Weber and the Idea of Economic Sociology*. Princeton: Princeton University Press.
- Useem, Michael 1984 *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and U.K.* New York: Oxford University Press.
- Uzzi, Brian 1996 "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61:674–698.
- Veblen, Thorstein (1899) 1973 *The Theory of the Leisure Class*. New York: Macmillan.
- Warde, Alan 1997 *Consumption, Food and Taste*. London: Sage.
- Weber, Max (1904–05) 1930 *The Protestant Ethic and the Spirit of Capitalism*. London: Allen & Unwin.
- (1904) 1949 "'Objectivity' in Social Science and in Social Policy." In *The Methodology of the Social Sciences*. New York: Free Press.
- (1921–22) 1978 *Economy and Society: An Outline of Interpretive Sociology*. Berkeley: University of California Press.
- (1923) 1981 *General Economic History*. New Brunswick, N.J.: Transaction Books.
- White, Harrison C. 1981 "Where Do Markets Come From?" *American Journal of Sociology* 87:514–547.
- Williamson, Oliver 1975 *Markets and Hierarchies*. New York: Free Press.
- Woolcock, Michael 1998 "Social Capital and Economic Development." *Theory and Society* 27:151–208.
- Zelizer, Viviana 1979 *Morals and Markets: The Development of Life Insurance in the United States*. New York: Columbia University Press.
- 1985 *Pricing the Priceless Child: The Changing Social Value of Children*. New York: Basic Books.
- 1988 "Beyond the Polemics of the Market: Establishing a Theoretical and Empirical Agenda." *Sociological Forum* 3:614–634.
- 1994 *The Social Meaning of Money*. New York: Basic Books.
- Zukin, Sharon, and Paul DiMaggio, eds. 1990 *Structures of Capital: The Social Organization of the Economy*. Cambridge, U.K.: Cambridge University Press.

RICHARD SWEDBERG

## EDUCATION AND DEVELOPMENT

It is safe to say that the current living standard is the highest since the beginning of human history.

We have achieved unprecedented levels of life expectancy, income per capita, and educational attainment over the past few decades. This unprecedented prosperity and achievement would probably not have been attained without the continuous technological progress of the peaceful era after World War II. Most people would acknowledge the role of education in the advancement of our socioeconomic development. The value of education is widely studied. For example, it has been found that better-educated farmers are more responsive to new technical possibilities and that better-educated women are more effective at allocating resources within the family, including those that enhance child survival (Cleland and Van Ginneken 1988; Lockheed et al. 1980; Mensch et al. 1985; Schultz 1979). This article examines the empirical relationship between education and development during recent decades. Included are a brief description of the history of world education and socioeconomic development since the early 1960s as well as discussions of theoretical background, data sources, research methodology, and findings.

#### EDUCATIONAL DEVELOPMENT IN RECENT DECADES

During the past few decades, a rapid expansion of educational provision at primary, secondary, and tertiary levels in much of the world has been documented (Shavit and Blossfeld 1993; World Bank 1998). Column 3 of Table 1 shows the primary school gross enrollment ratio from 1960 through 1990. The gross enrollment ratio is the ratio of total primary school enrollment, regardless of age, to the population of the age group that officially corresponds to the "usual" primary education years (World Bank 1998). The ratio for the world as a whole has increased from 86 per 100 population in primary school age group in 1960 to a virtually universal rate in 1990. Impressive gains have been observed for many areas of the world. In 1960, for example, the primary school enrollment ratio was only 39 per 100 for sub-Saharan Africa; by 1990, it had increased to 73 percent.

Similarly, the secondary school gross enrollment ratio for the world as a whole increased from

27 percent in 1960 to 54 percent in 1990. Progress was especially impressive in the Middle East and North Africa, where the enrollment ratio surged from 12 percent in 1960 to 57 percent in 1990. This gross enrollment ratio is lowest in sub-Saharan Africa, where only 4 percent and 22 percent of the population were enrolled in secondary school in 1960 and 1990, respectively.

One important reason for such educational expansion has been the adoption of a compulsory education policy by many countries. Egalitarian values regarding education have also emerged with increasing of modernization. According to Lenski (1966), the Western industrial nations began to subscribe to an egalitarian-democratic ideology, in which equality of educational opportunity is highly valued, after industrialization took place. Presumably this egalitarian-democratic ideology became part of the philosophy of the United Nations through the influence of Western countries. The United Nations General Assembly's Universal Declaration of Human Rights, Article 26, which proclaimed education as a basic right and demanded that elementary education should be compulsory and free resulted from post-World War II expansion of the conception of education as a fundamental human right. In addition, the demand for more skilled workers in today's economy has also played a role in the expansion of education.

#### SOCIOECONOMIC DEVELOPMENT IN RECENT DECADES

As discussed earlier, the overall progress in socioeconomic development during the past few decades has pushed our living standard to the highest level ever. The World Bank defines development as follows:

*Development is about people and their well-being—about people developing their capabilities to provide for their families, to act as stewards of the environment, to form civil societies that are just and orderly. (World Bank 1998, p. 35)*

At the national level, development is generally divided into two dimensions: social and economic. Indicators of social development include

## EDUCATION AND DEVELOPMENT

## World Education and Socioeconomic Development, 1960–1996

Region	Year	% Primary Enroll- ment	% Secondary Enroll- ment	Level of Indust- rialization	Level of Urban- ization	% with Radio	GNP per Capita (Constant 1987 US \$)	Infant Mortality	Life Expec- tancy	Population Growth Rate	Total Fertility Rate
<b>World</b>											
	1960	86	27	39	33	.	.	.	.	.	.
	1970	83	31	45	37	9	\$2,574	98	58.7	.	4.8
	1980	97	49	48	39	29	\$3,037	80	62.7	1.7	3.7
	1990	103	54	51	43	36	\$3,374	61	65.5	1.7	3.1
	1996	.	.	.	46	.	\$3,502	54	66.7	1.4	2.8
<b>Sub-Saharan Africa</b>											
	1960	39	4	18	15	.	.	.	.	.	.
	1970	50	7	21	19	4	\$508	137	44.2	2.7	6.6
	1980	78	14	28	23	9	\$556	115	47.6	3.1	6.6
	1990	73	22	32	28	16	\$488	100	50.7	2.9	6.0
	1996	.	.	.	32	.	\$475	91	52.2	2.8	5.6
<b>South Asia</b>											
	1960	56	18	25	17	.	.	.	.	.	.
	1970	67	25	29	19	2	\$233	139	48.8	2.4	6.0
	1980	76	27	31	22	4	\$250	120	53.8	2.4	5.3
	1990	91	39	36	25	8	\$348	89	59.2	2.2	4.0
	1996	.	.	.	27	.	\$419	73	62.1	1.8	3.4
<b>Middle East and North Africa</b>											
	1960	54	12	41	33	.	.	.	.	.	.
	1970	68	24	50	41	11	.	134	52.8	2.7	6.8
	1980	87	42	53	48	17	\$2,653	96	58.5	3.2	6.1
	1990	97	57	65	54	26	\$2,016	61	64.5	2.8	4.9
	1996	.	.	.	57	.	.	50	67.0	1.9	4.0
<b>East Asia and Pacific</b>											
	1960	101	19	18	17	.	.	.	.	.	.
	1970	88	24	24	19	1	\$147	79	59.2	2.7	5.8
	1980	111	43	28	21	7	\$225	56	64.5	1.5	3.1
	1990	122	47	31	28	17	\$391	42	67.3	1.6	2.4
	1996	.	69	.	32	.	\$636	39	68.2	1.2	2.2
<b>Latin America and Caribbean</b>											
	1960	89	15	52	49	.	.	.	.	.	.
	1970	.	28	59	57	13	\$1,386	84	60.6	2.6	5.2
	1980	106	42	66	65	26	\$1,867	59	64.8	2.2	4.1
	1990	106	48	74	71	35	\$1,675	42	68.1	1.9	3.2
	1996	.	.	.	74	.	\$1,877	33	69.5	1.6	2.8
<b>Europe and Central Asia</b>											
	1960	.	.	54	45	.	.	.	.	.	.
	1970	.	.	67	52	.	.	.	.	.	2.6
	1980	97	84	73	58	.	\$1,865	41	67.8	1.0	2.5
	1990	98	85	77	63	.	\$2,154	27	69.3	0.7	2.3
	1996	.	.	.	66	.	\$1,548	24	68.3	0.1	1.8
<b>North America<sup>#</sup></b>											
	1960	113	66	90	69	.	.	27	.	.	3.7
	1970	101*	65*	94	75	.	\$12,084	19	71.6	1.3	2.4
	1980	99	90	95	75	136	\$15,079	12	74.2	1.1	1.8
	1990	103	97	97	76	155	\$17,543	8	76.2	1.3	2.0
	1996	.	.	.	77	.	\$18,265	7	77.9	1.0	1.9

Table 1

SOURCE: 1998 World Development Indicator, World Bank CD-ROM.

*Table 1 - notes*

NOTE: Primary enrollment ratio: The ratio of total primary school enrollment, regardless of age, to the population of the age group that officially corresponds to the "usual" primary education/years.

NOTE: Secondary enrollment ratio: The ratio of total secondary school enrollment, regardless of age, to the population of the age group that officially corresponds to the "usual" secondary education/years.

NOTE: Level of industrialization: The proportion of the total labor force recorded as not working in agriculture, hunting, forestry, or fishing.

NOTE: eization: The percentage of the total population living in urban areas.

NOTE: GNP per capita (constant 1987 US\$): Gross national product divided by midyear population.

NOTE: \*The average of U.S.A. and Canadian data.

NOTE: †Canadian data only.

life expectancy, infant mortality, and educational attainment. The most commonly used indicator for economic development is GNP per capita (the gross national product divided by midyear population).

Table 1 shows world education and socioeconomic development from 1960 through 1996. The selected indicators include level of industrialization, level of urbanization, percent of population with radio, GNP per capita, infant mortality, life expectancy, population growth rate, and total fertility rate. The results are summarized below.

**Level of Industrialization.** Level of industrialization is defined as the percent of the total labor force employed in areas other than agriculture, hunting, forestry, and fishing. In 1960, 39 percent of the world labor force was so employed; by 1990, this figure had risen to 51 percent. Similar gains were observed for all world regions. The Middle East and North Africa posted the largest change in the level of industrialization. For example, 41 percent of the labor force in the Middle East and North Africa was employed in areas other than agriculture, hunting, forestry, and fishing in 1960; by 1990, the figure had risen to 65 percent. On the other hand, the gain in North America was low: The level of industrialization increased from 90 percent to 97 percent during the same period. It should be noted that the level of industrialization in North America started at a very high level, This accounts for the low level of gains in North America.

**Level of Urbanization.** Urbanization is the process whereby the proportion of people in a population who live in urban places increases. As our world moves toward being a more industrial

one, more people migrate from rural to urban areas to pursue better economic opportunities. The push-pull theory of migration is often cited to account for rural-to-urban migration (Ravenstein 1898; Lee 1886). The push factors are the unfavorable internal and external conditions in the places of origin that push individuals to leave their jobs/residences. Unfavorable internal employment conditions include lack of economic opportunities, low pay, low prospect for upward mobility, poor interpersonal relations, lack of challenge in the job, poor working environments, and so forth. Individuals in such circumstances are more likely to be pushed out of their jobs. Adverse external conditions that push individuals away from their residences/jobs, include such unfavorable structural conditions as high crime rate, pollution, and traffic congestion.

On the other hand, pull factors are favorable conditions in the new place of employment that attract individuals to migrate there. Facing the pressures of population growth and deteriorating economic opportunities, rural residents are being pushed out of their villages and attracted to urban areas, where they find a variety of economic opportunities to raise their living standard. According to Table 1, only 33 percent of the world's population lived in urban areas in 1960; by 1996, this number had increased to 46 percent. All the world regions have experienced similar gains, with the Middle East/North Africa and Latin America/Caribbean countries posting the largest gain (about 25 percentage points). North America ranked first in level of urbanization in 1996, with 77 percent of the population living in urban areas.

**Percent of Population with a Radio.** Another indicator of development is the percent of population with a radio. This indicator is an indirect

measure of exposure to modern values and ideas. In 1970, only 9 percent of the world's population had a radio; this figure had risen to 36 percent in 1990. South Asia has the lowest level of radio possession; in 1990, only 8 percent of the population in South Asia had a radio. North America has the highest level of radio possession; specifically, there were more than 1,500 radios per 1,000 population in 1990.

**GNP Per Capita.** GNP per capita is defined as the gross national product divided by midyear population. Based on Table 1, the GNP per capita (inconstant 1987 US\$) increased for most world regions from 1970 through 1996. For example, the GNP per capita in the world increased from \$2,574 in 1970 to \$3,502 in 1996; in North America, it increased from \$12,084 in 1970 to \$18,265 in 1996; in South Asia, it increased from \$233 in 1970 to \$419 in 1996.

**Life Expectancy.** Life expectancy at birth is the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life (World Bank 1998). According to Table 1, the life expectancy at birth for the world increased from 58.7 years in 1970 to 66.7 years in 1996. The gain is found for all world regions. The Middle East and North Africa posted the largest increase in life expectancy (from 52.8 in 1970 to 67 in 1996); the progress in South Asia has also been very impressive (from about 48.8 in 1970 to 62.1 in 1996). The observed increase in life expectancy from 1970 through 1996 is a strong indication of world socioeconomic development that enables newborns to live a longer life.

**Infant Mortality Rate.** Infant mortality rate (IMR) is defined as the number of deaths during the first year of life per 1,000 live births. The negative relationship between infant mortality and the level of economic development is often used as a barometer for economic development (United Nations 1982). Young (1993) found strong support for this relationship in developed countries. Krikshnan (1975) and Rodgers (1979) also reported a negative relationship between infant mortality and level of economic development for developing countries. Similarly, Berg (1973) and Gaise

(1979) also maintain that as a country's GNP increases, the standard of living improves, which leads to an improvement in nutrition and health services. Preston (1976), however, notes that one should not expect to find a direct relationship between mortality and per capita income, because per capita income is an average measure and does not take the distribution of income into account. The effects of income on mortality are likely to be greater at the lower end of the income distribution.

Table 1 shows that the world's IMR decreased from 98 in 1970 to 54 in 1996. All world regions experience a decreased IMR during the same period; for example, the IMR decreased from 137 to 91 for sub-Saharan Africa. The progress in reducing IMR was especially prominent in the Middle East and North Africa, where the IMR decreased from 134 in 1970 to 50 in 1996. North America had the lowest IMR—7—in 1996.

**Total Fertility Rate.** Total fertility rate is an estimate of the average number of children that would be born to a woman if the current age-specific birthrates remained constant. The reproductive revolution or the transition from high to low fertility is one of the dimensions of socioeconomic development. According to demographic transition theory, socioeconomic development facilitates fertility decline through the following mechanisms: 1) reducing infant/child mortality rate, 2) raising the status of women (including an increased level of education for women and an increased proportion of women employed in the nonagricultural sectors), 3) raising the marriage age and celibacy rate, 4) increasing the costs of raising children, and 5) reducing the economic value of children. Caldwell (1982) also argues that modernization creates reversed intergenerational wealth flows from parents to children. Such flow, unlike traditional wealth flow from children to parents, discourages couples to have high fertility. These changes coupled with accessible contraceptives, a higher value placed on smaller families, a latent demand for smaller families, and governmental family-planning policies are commonly cited factors that account for fertility decline. Moreover, diffusion/interaction theory (Bongaarts and Watkins 1886; Rosero-Boxby 1883; Casterline 1985)

and ideational theory (Lesthaeghe 1883) also provide significant theoretical insights on fertility decline.

In 1970, the total fertility rate (TFR) for the world was 4.8 children per woman; it had dropped to 2.8 in 1996. Similar patterns of fertility decline are found for all world regions. The worldwide reduction in fertility is as predicted by the demographic transition theory. In most of the more developed countries and some East Asian countries (e.g., Taiwan, Singapore, South Korea, and Hongkong), the total fertility rate has reached the replacement level (when TFR=2.1) or even below replacement level. Hirschman and Young (1999) also found that the total fertility rate had dropped to below 2.0 for Thailand in 1990. (Impressively the reproductive revolution in Thailand occurred when its GNP per capita was only \$1,470 in 1990 (World Bank 1994).)

In sum, the latest World Bank data show that we have made tremendous progresses in education and socioeconomic development over the past few decades. The central question, however, remains to be answered: What is the impact of education on socioeconomic development?

### THEORETICAL BACKGROUND

The title "Education and Development" does not imply a straight, unidirectional causal effect of education on development. Actually, the relationship between education and development can well be covariational.

To identify the exact cause and effect between the two is a difficult task. For example, education can facilitate development by providing the better-educated human resources that are essential for socioeconomic development. Education also reduces the fertility rate and thus population growth rate. It also transforms the labor-force structure and promotes rural-to-urban migration. On the other hand, the level of socioeconomic development in a country is likely to influence the level of education for that country. As development progresses, countries would have more resources to invest in education (thus development affects education). According to the functionalist theorists, the

rapidly changing technology of the twentieth century has generated a demand for a better-educated labor force. The expansion of schooling can be viewed as a direct response to these technological changes. Moreover, the need for a more skilled labor force would encourage government to invest more in education in order to keep the economy competitive in today's world economy.

Thus, the relationship between education and development can be best viewed as covariational. Here I assume that education and development are related to each other in the initial stage of development. My goal is to investigate the net impact of education in an early stage on later stages of socioeconomic development after controlling for early-stage development and other important intervening variables. There are three main theories (modernization, human capital, and world-system) that address the impact of education on development.

**Modernization Theory.** Modernization is a transformation of social and economic structures. The *International Encyclopedia of the Social Sciences* (1868) defines modernization as "the process of social change in which development is the economic component" (p. 387). From a comparative perspective, modernization can be viewed as the process of social change whereby less developed countries acquire characteristics common to more developed countries. Lasswell (1965) argued that modernization not only shapes economic factors but also reshapes all social values such as power, respect, rectitude, affection, well-being, skill, and enlightenment. Common characteristics of modernity include: (1) a degree of self-sustaining growth in the economy; (2) a measure of public participation in the polity; (3) a diffusion of secular-rational norms in the culture; (4) an increment of mobility in the society; and (5) a corresponding transformation in the modal personality that equips individuals to function effectively in a social order that operates according to the foregoing characteristics.

Proponents of modernization theory argue that the transformations of socioeconomic structures—such as the mechanization of agriculture,

urbanization, a mass communication network, demographic transition, the expansion and integration of a national market, and an increase in political participation—are necessary preconditions for sustained economic growth (Apter 1965; Rostow 1960). Obviously, changes in these social forces create new opportunities, incentives, and normative influences that can affect (1) an individual's view on the world and (2) his or her behavior. Another thread of the modernization theory stresses that exposure to modern values leads to socio-economic development. Inkeles and Smith (1975) maintain that modern people, as opposed to traditional people, are prepared to act on their world rather than fatalistically accept it; have a cosmopolitan rather than a local orientation; see the sense of deferring gratification; welcome rather than distrust change; are not constrained by irrational religious or cultural forms; and recognize the value of education. According to modernization theory, education plays a crucial role in making the route to modernization possible.

Macroeconomic studies have shown that education is positively correlated with overall economic growth, with one year of additional schooling of the labor force possibly leading to as much as a 9 percent increase in gross domestic product (GDP) for the first three years of schooling and a yearly 4 percent increase for the next three years (Summers 1994). Increased education has also been found to result in greater agricultural productivity, even in developing countries (Jamison and Lau 1982; Lockheed et al. 1980). The National Research Council (1986) also reported that “urbanization plays a beneficial role in the development process, providing an increasing share of population with access to relatively high-wage employment, education, health care, and other modern public services” (p. 76).

However, in a case study of Egypt, Faksh (1977) found that “educational expansion in modern Egypt thus far has not been conducive to development in the general configuration of the Egyptian polity” (p. 238).

**Human Capital Theory.** Human capital theory argues that education leads to development by

increasing the efficiency and productivity of workers. Investment in human capital is a key element in achieving long-term sustainable economic growth. According to this perspective, “the main contribution of education to economic growth was to increase the level of cognitive skills possessed by the work force and consequently to improve their marginal productivity” (Benavot 1989, p. 15). According to the theory, the provision of education is not a form of consumption but a productive investment in society's “stock” of human capital. Investment in human capital is at least as profitable as investment in physical capital. At the national level, increasing the overall level of education will raise the stock of the human capital, which will have a positive impact on national productivity and economic growth. At the individual level, education level provides some indication of the ability of a person to perform certain duties and adapt to other work situations.

**World-System/Dependency Theory.** World-system and dependency theories suggest that the specialization of some countries in the export of raw materials and lightly processed goods is an important cause of their underdevelopment. Moreover, the world-system/dependency theories argue that the needs and interests of Western capitalism determine the pattern of education in developing countries. Education is seen as part of the process whereby peripheral countries are kept underdeveloped. The prevalence of foreign investment capital, the presence of multinational corporations, the concentration on exporting primary products, and the dependence on imported technologies and manufactured goods constrain long-term economic development (Bornschieer and Chase-Dunn 1985; Delacroix and Ragin 1978). According to the theory, “education, far from a key component in development, modernization, self-sufficiency, and so on, is in fact yet another instrument of enslavement, a way of tightening, rather than loosening, the dependency bond.” (Dale 1982, p. 412).

Each of the above theory examines the relationship between education and development from different angles. Each theory delineates part of the dynamics of the relationship. Due to the limitation



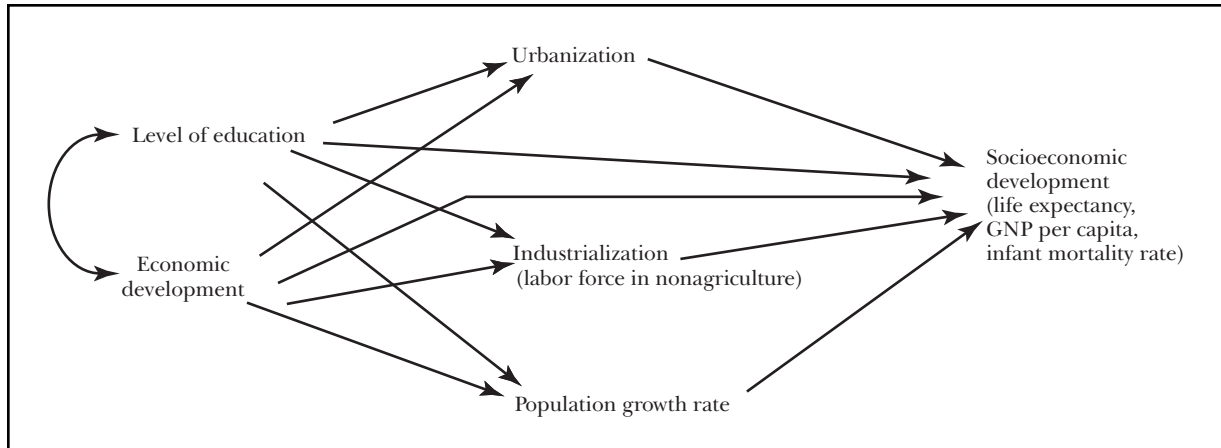


Figure 1. Analytical Model of Education and Development

of data, this article will examine only the validity of modernization theory.

**DATA SOURCE, ANALYTICAL MODEL, MEASUREMENT, AND RESEARCH METHODS**

**Data Source.** The World Bank compiles national data on education, demographics, and socioeconomic development from various years and sources. The data used for this research is from the World Bank’s 1998 World Development Indicator CD-ROM.

**Analytical Model.** In this article, education and economic development are assumed to be related to each other in the initial stage of analysis (e.g., measures for both variables obtained in 1970). Both education and development are the independent variables. From there, I analyze how both 1970 factors affected the intervening variables (1980 level of industrialization, urbanization, and population growth rate). Finally, I estimate the net effect of 1970 education on 1990 development, after controlling for 1970 development and other 1980 intervening variables. The effect of education on development is also estimated for the 1980–1996 period. This model is shown in Figure 1.

**Measurement.** Education is defined as the formal schooling, although education can also be defined as an alternative to family education, an

instrument of state social policy, a site of civic reform, or a form of humanistic progress. Socioeconomic development is broadly defined as the progress in the areas of GNP per capita, infant mortality, and life expectancy.

There are three sets of variables—the independent, the intervening, and the dependent variables. The independent variables include the level of education and economic development observed at a earlier time period. The intervening variables, including urbanization, industrialization, and population growth rate, are controlled to mediate the effects of the independent variables on the dependent variable. Finally, the dependent variable is the level of socioeconomic development observed at a later period of time. Three indicators are used: GNP per capita, life expectancy, and infant mortality.

*Level of education* is defined as the secondary school gross enrollment ratio, defined in the previous section on “Educational Development in Recent Decades.” I use this as the main independent variable rather than the primary school enrollment ratio because for many countries compulsory education is limited to the primary level. Lack of variation in primary school enrollment ratio could pose a threat to the validity of the study.

*Socioeconomic development* include three indicators: life expectancy, GNP per capita, and infant mortality rate. Lift expectancy at birth measures

**Descriptive Statistics for All Variables in Panel Analyses**

Variables	Mean	SD	Minimum	Maximum	N
<b>1970–1990</b>					
<b>Independent variables</b>					
Percent of secondary enrollment, 1970	31.6	26.4	1.0	102.0	128
Logged per capita GNP, 1970	7.2	1.5	4.5	11.0	119
<b>Intervening variables</b>					
Industrialization, 1980	57.1	28.0	5.6	98.7	173
Urbanization, 1980	45.9	24.1	3.9	100.0	195
Population growth rate, 1980	2.0	1.5	-2.4	9.7	194
<b>Dependent variables</b>					
Logged per capita GNP, 1990	7.4	1.4	4.7	10.3	159
Life expectancy, 1990	64.9	10.3	35.3	78.8	192
Infant mortality rate, 1990	48.0	41.8	4.6	189.0	194
<b>1980–1996</b>					
<b>Independent variables</b>					
Education, 1980	49.37	31.56	3	114	147
Logged per capita GNP, 1980	7.42	1.47	4.8	10.72	136
<b>Intervening variables</b>					
Industrialization, 1990	62.36	28.28	5.88	99.63	173
Urbanization, 1990	50.41	23.73	5.2	100	194
Population growth rate, 1990	1.9	1.68	-4.87	14.12	199
<b>Dependent variables</b>					
Logged per capita GNP, 1996	7.23	1.44	4.55	10.32	138
Life expectancy, 1996	66.08	10.22	36.88	79.76	194
Infant mortality rate, 1996	42.14	38.75	3.5	174.2	195

Table 2

the overall quality of life. It is defined as “the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life” (World Bank 1998: 19). The GNP per capita measures the economic aspect of progress. Infant death is the final biological expression of a process that is determined basically by the economic and social structure of a country or region. These conditions influence the occurrence and spread of disease as well as quality and availability of health care facilities, all of which are crucial to survival probabilities. The structural determinants are mediated at the family level, because the child’s growth and development are heavily dependent on the living conditions of the family.

**Intervening Variables.** The intervening variables include urbanization of population in urban areas, industrialization of labor force in nonagricultural activities, and annual population growth rate. Time-lag path analysis is used to investigate

the direct and indirect effects of education at an earlier time (e.g., 1970) on socioeconomic development at the later time (e.g., 1990), after controlling for urbanization, population growth rate, and industrialization. Specifically, two models will be examined. The first model uses 1970 data for the independent variables, 1980 data for the intervening variables, and 1990 data for the dependent variables. The second model examines the periods between 1980 and 1996. That is, I use 1980 data for the independent variables, 1990 data for the intervening variables, and the 1996 data for the dependent variables. For both analyses, the unit of analysis is country.

**Research Method.** Path analysis will be used to study the proposed model. The path coefficient represents the standardized regression coefficient. The standardized regression coefficient ( $\beta$ ) represents the change in the standard deviation of the dependent variable associated with one standard deviation of change in the independent variable, when all other variables are controlled for.

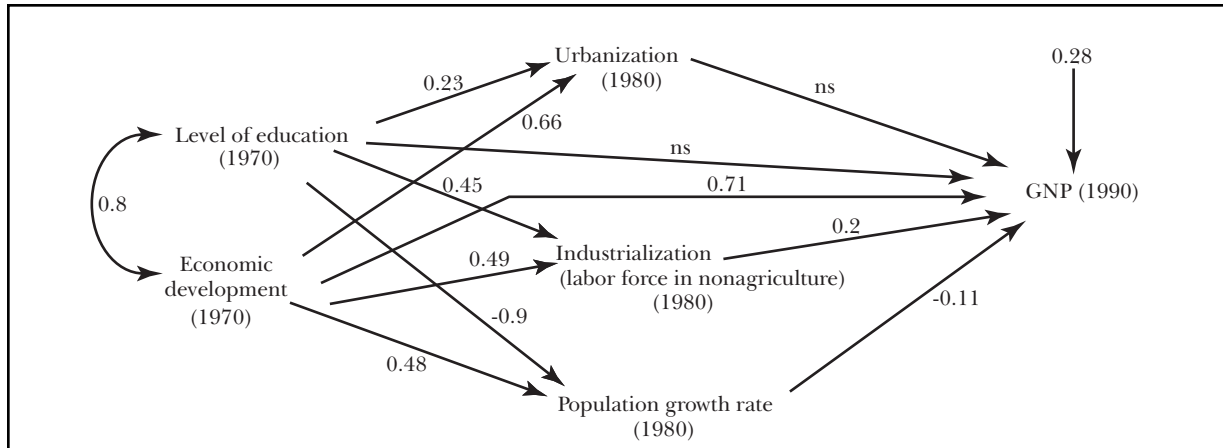


Figure 2. Education and Economic Development, 1970–1990.

The coefficient of alienation (defined as the square root of  $1-R^2$ ) is also provided to show how well each development indicator is predicted by all the independent variables. A larger coefficient of alienation indicates that the development model has a smaller  $R^2$  or coefficient of determination. On the other hand, if a development model has a smaller coefficient of alienation, then the independent variables in that model explain more variation in that development indicator.

**FINDINGS AND DISCUSSION**

**Descriptive Statistics.** Table 2 presents means, standard deviations, the minimum values, and the

maximum values for all variables used in the analysis. The upper panel shows the data for 1970–1990 period and the lower panel for 1980–1996 period. There are two independent variables, percent of secondary school enrollment and logged GNP per capita. GNP per capita has an extremely skewed distribution. In regression analysis, normal distributions for all variables are expected. To correct this problem, I take the natural log of GNP per capita. The mean logged GNP per capita in 1970 was 7.2. The 1990 logged GNP per capita was 7.4.

The intervening variables include industrialization, urbanization, and population growth rate. Finally, the dependent variables include economic

**Correlation Matrix of Variables Used in Panel Path Analyses**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1. Education, 1970	1.00															
2. GNP per capita, 1970	0.80	1.00														
3. Industrialization, 1980	0.81	0.84	1.00													
4. Urbanization, 1980	0.74	0.82	0.89	1.00												
5. Population growth rate, 1980	-0.52	-0.32	-0.36	-0.24	1.00											
6. GNP per capita, 1990	0.81	0.95	0.86	0.81	-0.32	1.00										
7. Life expectancy, 1990	0.79	0.73	0.88	0.73	-0.43	0.81	1.00									
8. Infant mortality rate, 1990	-0.76	-0.70	-0.85	-0.69	0.43	-0.80	-0.96	1.00								
9. Education, 1980	0.92	0.78	0.81	0.72	-0.48	0.72	0.78	-0.77	1.00							
10. GNP per capita, 1980	0.76	0.98	0.87	0.83	-0.19	0.98	0.78	-0.75	0.78	1.00						
11. Industrialization, 1990	0.79	0.82	0.99	0.88	-0.35	0.85	0.88	-0.85	0.80	0.86	1.00					
12. Urbanization, 1990	0.70	0.80	0.87	0.98	-0.17	0.80	0.70	-0.66	0.68	0.82	0.87	1.00				
13. Population growth rate, 1990	-0.48	-0.17	-0.32	-0.16	0.75	-0.24	-0.37	0.38	-0.41	-0.12	-0.32	-0.13	1.00			
14. GNP per capita, 1996	0.84	0.94	0.81	0.78	-0.61	0.98	0.79	-0.77	0.69	0.97	0.80	0.76	-0.58	1.00		
15. Life expectancy, 1996	0.79	0.72	0.87	0.73	-0.41	0.80	0.99	-0.94	0.75	0.77	0.88	0.70	-0.35	0.79	1.00	
16. Infant mortality rate, 1996	-0.75	-0.68	-0.84	-0.68	0.41	-0.78	-0.96	0.99	-0.75	-0.73	-0.85	-0.66	0.36	-0.76	-0.95	1.00

Table 3

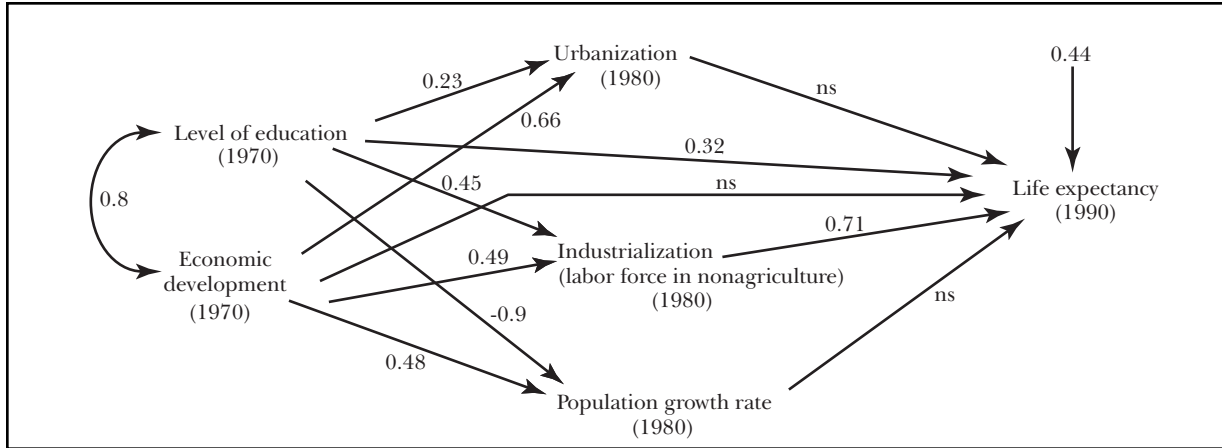


Figure 3. Education and Life Expectancy, 1970–1990

development (logged GNP per capita at a later year), life expectancy, and infant mortality.

**Corelation Matrix.** Table 3 shows the correlation matrix for all variables used in the two-panel analyses. All the correlation coefficients are significant at .01 level. Furthermore, the strengths of all correlations are substantially strong. Moreover, the direction of relationship is as expected by modernization theory. For example, there is a high correlation between 1970 education and 1990 socioeconomic development. Specifically, .81, .79, and  $-.76$  are the correlations between 1970 education and 1990 logged GNP per capita, 1990 life expectancy, and 1990 infant mortality rate, respectively.

Similarly, there is also a high correlation between 1980 education and 1996 socioeconomic development. Specifically, the correlations between 1980 education and 1996 logged GNP per capita, 1996 life expectancy, and 1996 infant mortality rate are .69, .75, and  $-.75$ , respectively. Since the bivariate correlation between the independent variable and the dependent variables does not control for other causal mechanisms, I will be controlling for intervening variables in path analyses. The results of path analyses are reported below.

**Path Models.** Figure 2 shows the relationship between education and economic development during the 1970–1990 period. The path model shows that the level of education, as measured by secondary school enrollment rate in 1970, has no

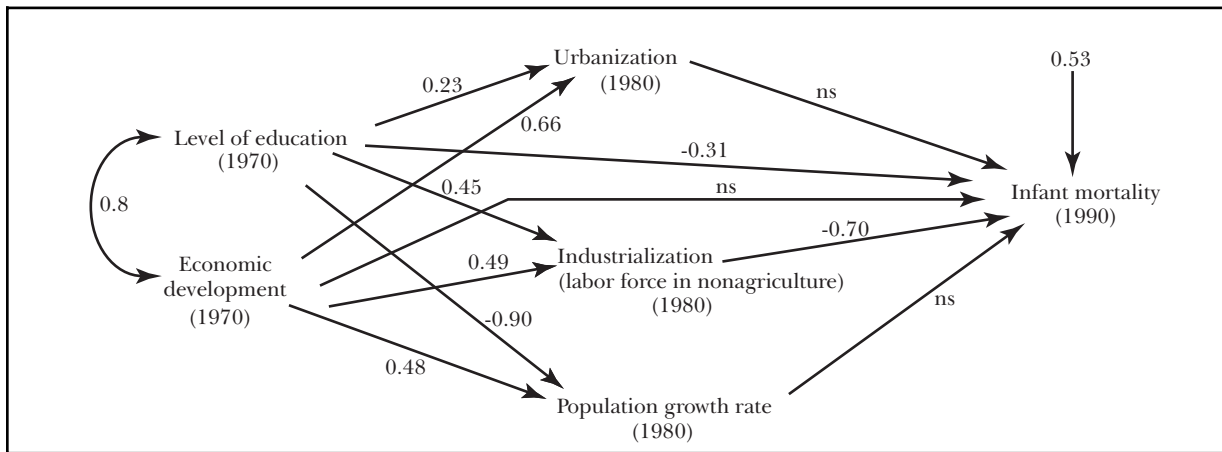


Figure 4. Education and Infant Mortality, 1970–1990

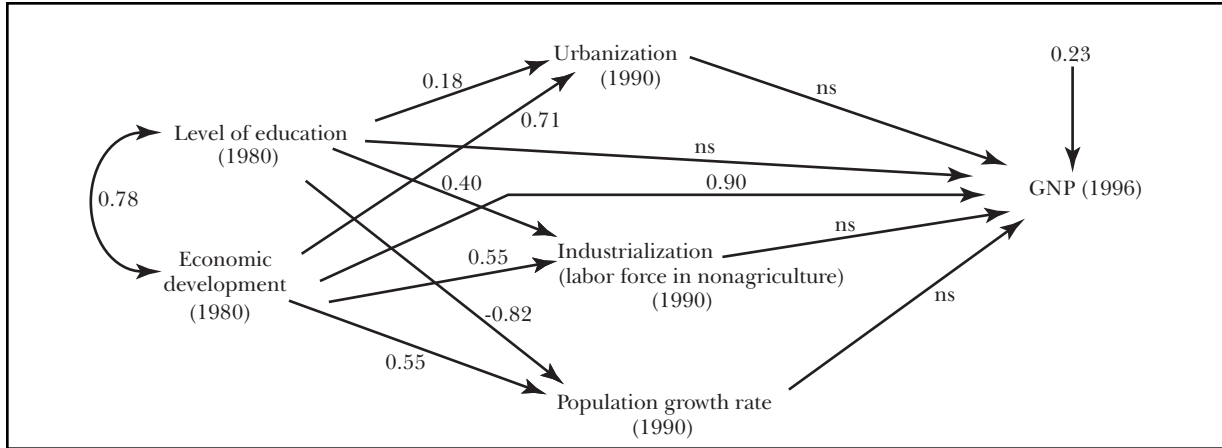


Figure 5. Education and Economic Development, 1980–1996

direct effect on economic development, as measured by GNP per capita in 1990, after the 1980 urbanization, 1980 industrialization, 1980 population growth rate, and 1970 economic development are held constant. Similarly, urbanization has no effect on economic development after other variables are controlled for, although cross-national studies suggest that urbanization is related to the level of economic development as measured by per capita income or GNP (Chenery and Syrquin 1975).

Nevertheless, education has indirect effects on economic development. The first indirect effect of .09 ( $.45 \times .02 = .09$ ) from education to economic development is through its effect on industrialization (the path coefficient is .45). High

enrollment rate in secondary school is found to have a moderate and positive effect on level of industrialization. Industrialization is found to have a positive direct effect on economic development (the path coefficient is .20). The second indirect path from education to economic development is through population growth rate. High secondary school enrollment rate lowers a country's population growth rate (path coefficient is  $-.90$ ). A country's population growth rate is found to have a weak and negative effect on its economic development (the path coefficient is  $-.11$ ). The second indirect effect of education on economic development is  $.10$  ( $-.99 \times -.11 = .10$ ). The total indirect effect of education on economic development is  $.19$  ( $.09 + .10 = .19$ ).

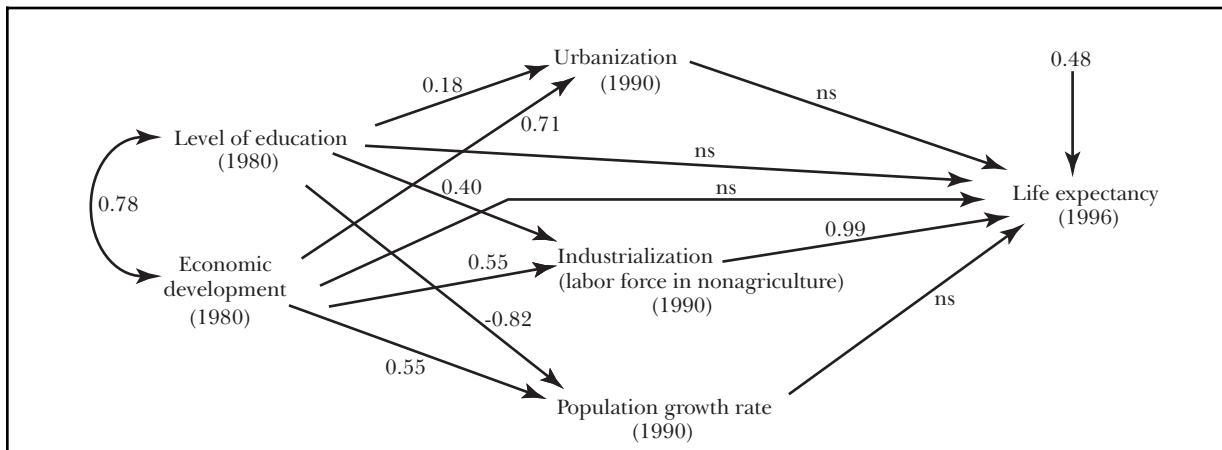


Figure 6. Education and Life Expectancy, 1980–1996

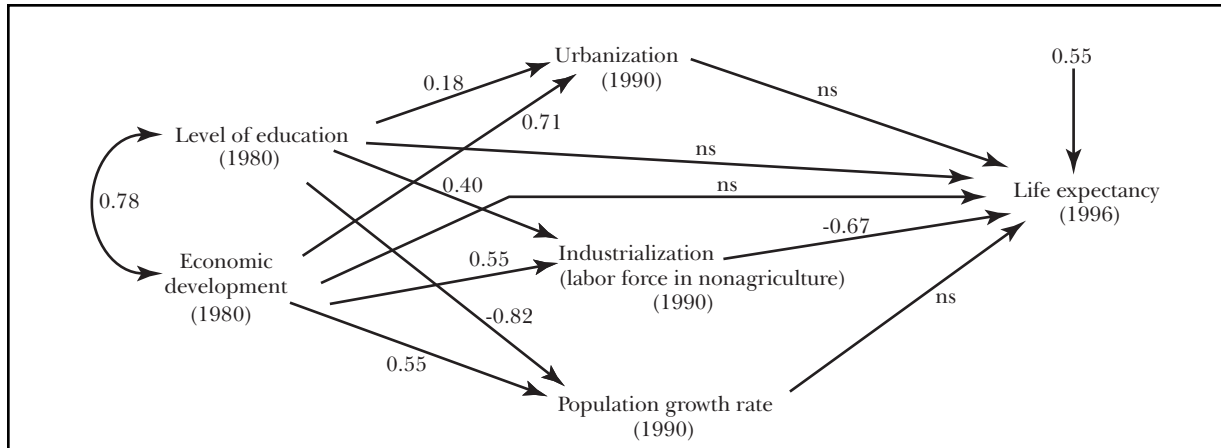


Figure 7. Education and Infant Mortality, 1980–1996

Figure 3 shows the relationship between education and life expectancy during the 1970–1990 period. The path model shows that the level of education in 1970, as measured by secondary school enrollment rate, has a positive and moderate direct effect on 1990 life expectancy (the path coefficient is .32), after 1980 urbanization, 1980 industrialization, 1980 population growth, and 1970 GNP per capita are controlled for. Countries that invest their resources in education can directly increase their population’s life expectancy. Urbanization and population growth rate are found to have no effect on life expectancy. In addition to its direct effect on life expectancy, education also has an indirect effect on life expectancy. The indirect effect of .32 (.45×.71=.32) from education to life expectancy is through its effect on industrialization (the path coefficient is .45). High level of industrialization is found to have a very strong positive effect on life expectancy (path coefficient is .71). High proportion of population engaged in nonagricultural occupations increases a country’s life expectancy. The total effect of education on life expectancy is .64 (the sum of direct effect, .32, + indirect effect, .32).

The last model for the relationship between education and development between 1970 and 1990, as measured by infant mortality, is shown in Figure 4. According to Figure 4, level of education in 1970 reduces the infant mortality rate in 1990. This finding is consistent with other studies based on individual-level analysis. Moreover, 1980 industrialization also has a direct and negative effect on

infant mortality. Countries with a low proportion of the labor force in nonagricultural activities are more likely to have higher infant mortality rate. The indirect effect of education on infant mortality through industrialization is  $-.32$  (.45 ×  $-.70 = -.32$ ). The total effect on education on infant mortality is  $-.63$  ( $-.32 + -.31 = -.63$ ).

Figures 5, 6, and 7 examine the relationship between education and economic development, life expectancy, and infant mortality rate during 1980–1996. The intervening variables—urbanization, industrialization, and population growth rate—are based on 1990 data. Surprisingly, the results from Figure 5 shows that there was no direct or indirect effect of education on economic development during 1980–1996. Only 1980 GNP per capita had a direct impact on the 1996 GNP per capita.

Figure 6 shows that the level of education in 1980 had no direct effect on the 1996 life expectancy. However, it had an indirect effect on life expectancy through its effect on industrialization. The indirect effect of education on life expectancy is .31 (.40 - .77 = .31).

Similar to what was reported in Figure 6, Figure 7 shows that the level of education in 1980 had no direct effect on 1996 infant mortality. However, it had an indirect effect on infant mortality through its effect on industrialization. The 1980 industrialization has a substantially strong negative and direct effect on 1996 infant mortality rate (path coefficient =  $-.67$ ). The indirect effect of

education on infant mortality is  $-.27 (.40 \times -.67 = -.27)$  for the 1980–1996 period. The level of economic development in 1960 is also found to have an indirect and negative impact on 1996 infant mortality rate. The indirect effect of economic development on infant mortality is  $-.37 (.55 \times -.67 = -.37)$ , which is stronger than the impact of education during this period.

There were some weaknesses in the study. The variables used in the study are period data. Period data are collected in a given year when their values are very much influenced by macro socioeconomic conditions. The time intervals between the independent and the dependent variables of 20 years (or 16 years for second model) may seem long. Finally, secondary school enrollment was used as the independent variable. Further studies may consider tertiary education as an indicator of education.

#### SUMMARY AND CONCLUSION

Modernization theory maintains that education promotes development. For many developing countries, education is a prominent means of attempting to narrow the knowledge gap between the highly industrialized countries and the developing countries.

This article reported the world's impressive education and socioeconomic developments over the past few decades. It also examined a diverse set of mechanisms through which education affects socioeconomic development for two periods: 1970–1990 and 1980–1996. The overall findings suggest that education has a positive effect on life expectancy for both time periods examined. Moreover, education was found to have had negative relationship to infant mortality for the 1970–1990 and 1980–1996 periods. However, the effect of education on economic development is more complicated. Education is found to have had a positive effect on economic development for only the 1970–1990 period, not for 1980–1996 period. Another interesting finding is that the effect of education on development was lower during 1980–1996 than the 1970–1990 period.

The study shows that education, industrialization, and development are inextricably interrelated. To achieve a higher level of socioeconomic

development, policy makers would need to consider the complex relationship between education and development. Several countries have taken steps to improve their educational systems in order to bolster their economies and improve conditions in their nations. In recent history, for example, China and Taiwan have made attempts to modernize and strengthen their economies by encouraging people to further their education, especially in science and technology. It is clear from this study that to improve the welfare of the billions of people in the developing world, governments in developing countries need to continue and expand their investments in education for their population.

The sociology of education and the sociology of development have become very important areas in sociological research. As more advanced data become available, we will be able to do better research in this area by conducting more sophisticated and comprehensive studies. Examining modernization and other theories in the area of education and development will enable sociologists to provide solid knowledge about the mechanisms whereby education affects development, which can be important information for policy makers when they implement policies related to education and development.

#### REFERENCES

- Apter, D. 1965 *The Politics of Modernization*. Chicago: University of Chicago Press.
- Benavot, Aaron 1989 "Education, Gender, and Economic Development: A Cross-National Study." *Sociology of Education* 62(1):14–32.
- Berg, B. 1973 "Nutrition and National Development." In Alan Berg, Neven S. Scrimshaw, and David L. Call, eds., *Nutrition, National Development and Planning*. Cambridge, Mass.: MIT Press.
- Bornschier, Y., and C. Chase-Dunn 1985 *Transnational Corporations and Underdevelopment*. New York: Praeger Publishers.
- Bongaarts, John, and Susan C. Watkins 1996 "Social Interactions and Contemporary Fertility Transitions." *Population and Development Review* 22:639–682.
- Caldwell, John C. 1982 *Theory of Fertility Decline*. London: Academic Press.
- Casterline, John B. ed. 1985 *The Collection and Analysis of Community Data*. Voorburg, Netherlands: International Statistical Institute.



- Chenery, H., and M. Syrquin 1975 *Patterns of Development, 1950–1970*. New York: Oxford University Press.
- Dale, Roger 1982 “Learning to Be . . . What? Shaping Education in Developing Societies.” In Hamza Alavi and Teodor Shanin, eds., *Introduction to the Sociology of “Developing Societies.”* New York: Monthly Review Press.
- Delacroix, J., and C. Ragin 1978 “Modernizing Institutions, Mobilization, and Third World Development: A Cross-National Study.” *American Journal of Sociology* 84:123–149.
- Diebolt, Claude 1997 “Education and Economic Growth in Germany Before the Second World War.” *Historical-Social-Research*.
- Faksh, Mahmud A. 1977 “The Chimera of Education for Development in Egypt: The Socio-Economic Roles of University Graduates.” *Middle Eastern Studies* 13:229–240.
- Gaise, S. K. 1979 “Some Aspects of Socioeconomic Determinants of Mortality in Tropical Africa” (WHO doc: DSI/SE WP/79:13).
- Hirschman, Charles, and Yih-Jin Young 1999 “The Decline of Fertility in Thailand, Malaysia, Indonesia and the Philippines.” In Ronald D. Lee and C. Y. Cyrus Chu, eds., *Economic Aspects of Demographic Transition in Asia-Pacific Countries. A supplement to Population and Development Review* (Forthcoming).
- 1998 “Fertility Transition in Southeast Asia.” Paper presented at the Conference on Economic Aspects of Demographic Transition: The Experience of Asian-Pacific Countries, Taipei, Taiwan, June 19–20, 1998.
- Inkeles, Alex, and David Smith 1975 *Becoming Modern*. London: Heinemann.
- Jamison, Dean, and Lawrence J. Lau 1982 *Farmer Education and Farm Efficiency*. Published for the World Bank. Baltimore: The Johns Hopkins University Press.
- Lasswell, Harold D. 1965 “The Policy Sciences of Development.” *World Politics* 17:286–309.
- Lee, Everett 1966 “A Theory of Migration.” *Demography* 3:47–57.
- Lesthaeghe, Ron 1983 “A Century of Demographic and Cultural Change in Western Europe: An Exploration of Underlying Dimensions.” *Population and Development Review* 9:411–435.
- Lockheed, Marlane E., Dean T. Jamison, and Lawrence Lau 1980 “Farmer Education and Farm Efficiency: A Survey.” *Economic Development and Cultural Change* 29(1):37–76.
- Mensch, B., H. Lentzner, and S. Preston 1985 *Child Mortality Differentials in Developing Countries*. New York: United Nations.
- National Research Council 1986 *Population Growth and Economic Development: Policy Questions*. Washington, D.C.: National Academy Press.
- Preston, Samuel H. 1976 *Mortality Patterns in National Populations: With Special Reference to Recorded Causes of Death*. New York: Academic Press.
- Ravenstein, E. 1898 “The Laws of Migration.” *The Journal of the Royal Statistical Society* 52:241–301.
- Rodgers, G. B. 1979 “Income and Inequality as Determinants of Mortality: An International Cross-Section Analysis.” *Population Studies* 33(2):343–351.
- Rosero-Bixby, Luis, and John B. Casterline 1993 “Modelling Diffusion Effects in Fertility Transition.” *Population Studies* 47:147–167.
- Rostow, W. W. 1960 *The Stages of Economic Growth*. Cambridge, U.K.: Cambridge University Press.
- Schultz, T. W. 1979 “Investment in Population Quality in Low Income Countries.” In P. M. Haieser, ed., *World Population and Development: Challenges and Prospects*. Syracuse, N.Y.: Syracuse University Press.
- Shavit, Yossi, and Hans-Peter Blossfeld, eds. 1993 *Persistent Inequality, Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview Press.
- Summers, Lawrence H. 1994 “Investing in All the People: Educating Women in Developing Countries” (EDI Seminar Paper No. 45), Washington, D.C.: Economic Development Institute of the World Bank.
- United Nations 1982 “Levels and Trends of Mortality Since 1950: A Joint Study by the UN and WHO.” New York: United Nations.
- World Bank 1994 *World Tables 1994*. Baltimore: John Hopkins University Press.
- 1998 *World Development Indicators 1998*. Washington, D.C.: The World Bank.
- Young, Yih-Jin 1993 “A Macro-Comparative Analysis on Socioeconomic Development and Infant Mortality.” Paper presented at the Canadian Association for the Study of International Development Ninth Annual Conference, Ottawa, Canada, June 1993.

YIH-JIN YOUNG

## EDUCATION AND MOBILITY

One of the main reasons education is valued so highly in modern societies is the role it plays in relation to social mobility and reproduction. This role has long been debated between those who emphasize its contribution to social mobility and



those who focus on its contribution to social reproduction. In order to understand this debate, it is useful to review the key concepts and theoretical perspectives before considering the empirical evidence and then offering a resolution.

*Social stratification* refers to institutionalized inequality, that is, to hierarchically structured social positions (strata) and to the inequality in social rewards received by people who belong to different strata. Social stratification is based mainly on class or status, although other forms of stratification exist (for elaboration, see Grusky and Takata 1992; Haller 1992). *Class* is the term preferred by theorists who view the social order as consisting of distinctive economic groupings struggling to maximize their interests vis-à-vis each other, while *status* is preferred by theorists who perceive a continuing distribution of socioeconomic variation without clear-cut divisions and conflict.

*Social mobility* is the movement from one class or status to another. The emphasis here, as with most studies of social mobility, is on *intergenerational mobility*, which refers to the change in class or status from parents to their adult children. An example of intergenerational mobility is when the daughter or the son of peasants becomes a doctor. In contrast, when the child of peasants ends up being a peasant, it is an example of *social reproduction*.

The class or status positions that individuals occupy in society are usually attributed to both *ascriptive* and *achievement* processes. These are generally viewed as opposite or contradictory processes involving either ascribed characteristics based on biological factors and family of origin or achieved characteristics based on individual traits and behaviors. Stratification systems that emphasize ascriptive characteristics for class or status placement are defined as “closed” and lead to status inheritance or class reproduction. Those stratification systems that emphasize achieved characteristics are defined as “open” and are expected to lead to social mobility.

The opposing positions are formalized in the functionalist and conflict theories of social stratification. With respect to the role of education in producing social mobility, functionalists argue that different social roles require different skills and abilities and that, if society is to function effectively, they must be filled by individuals possessing the appropriate skills and abilities (Davis and Moore

1945). The positions most valued by society are usually the most critical for societal functioning and the most demanding of individual skills and ability. In order to encourage individuals to invest the time and effort for training and to attract the best-qualified individuals, these positions have to be accompanied by higher social and economic rewards. Education is widely viewed as both developing and reflecting individual skills and abilities, and it is therefore used as a means of social selection. Thus, education enhances social mobility by providing for social selection based on achieved rather than ascribed characteristics of individuals.

Conflict theorists start with the premise that society consists of different groups with conflicting interest, and they argue that stratification exists because groups that acquired power, wealth, and prestige want to maintain and enhance their position at the expense of less privileged groups. In respect to education, most conflict theorists agree that schools help to reproduce and legitimize the stratification system by portraying attainment as an achieved individual characteristic, while in fact they select and process individuals on the basis of ascriptive characteristics (Bowles and Gintis 1976; Bourdieu 1977; Willis 1977).

Empirical research on the role of education in the process of social mobility or reproduction has produced conflicting evidence. The argument of mobility through education as suggested by functional theories depends on the validity of two general conditions: (1) Educational attainment must be used as a criterion of eventual class or status position, and (2) the level of educational attainment of individuals must not be influenced by the level of their family’s class or status. Boudon (1976) calls these two conditions necessary for social mobility “meritocracy” and “equality of educational opportunity” respectively. It is important to note that social mobility exists only if both conditions are met and that each of them alone is a necessary but insufficient condition for social mobility.

These conditions necessitate a distinction, as far as the role of education is concerned, that is very rarely made between class and status. The role of education differs considerably in class and status mobility. Education plays a very weak role in class mobility or reproduction (Katsillis and Armer

1992). More specifically, the meritocracy condition is not satisfied, since education is almost never used as a criterion for class. Following Marx, most class theorists see two major classes, capitalists and workers, although the perception of the exact number of classes and their composition varies (see Poulantzas 1974; Wright 1978, 1985; for a discussion of the different views, see Grusky and Takata 1992). Education would play an important role in class reproduction only if it were a major criterion for becoming a capitalist or a worker and the vehicle through which the class of the parents is transferred to their children. But it is neither. In fact, if education were a principle determinant of class, one would expect most Ph.D.'s to be capitalists and a significant number of children from bourgeois families to become workers because of educational difficulties (Katsillis and Armer 1992).

A detailed discussion of the process through which class is reproduced is beyond the scope of this article. It suffices to say here that other institutions, such as the legal system and, more specifically, the inheritance and property transfer laws, are much more reliable and effective for the transfer of class from parents to their children. Whether these children have a high school or higher diploma may be important for other reasons, but it is clearly not essential for the reproduction of their class (Katsillis and Armer 1992).

In addition, educational attainment does not seem to be influenced by family class. Indeed, there is no reliable and consistent empirical evidence that supports such a relationship. When class is measured as determined by the relation of production and not as a status position, it has no effect on educational attainment, especially if the status position is held constant (see Katsillis and Armer 1992; Katsillis and Rubinson 1990; Katsillis and Spinthourakis 1997). At first glance, this would indicate that the equality of educational opportunity hypothesis is satisfied in relation to class. However, taking into account the absence of meritocracy in class determination, the lack of influence of class is more an indication of the weak role education plays in class reproduction or mobility than of its equalizing potential.

In contrast to class, education plays an important role in status allocation. Many studies have empirically tested the meritocracy hypothesis, and

almost all have found a significant relationship between educational and later socioeconomic attainments (Blau and Duncan 1967; Duncan et al. 1972; Jencks et al. 1972; Sewell et al. 1969). Indeed, most studies in the United States have found educational attainment to be among the most important determinants of occupational and status attainment, although the findings regarding its relationship to income are not as conclusive (Jencks et al. 1972). In short, the meritocracy condition is well supported by empirical evidence.

However, other studies of social mobility have found that employing meritocracy in the allocation of occupational and social status does result in substantial increases in social mobility (Boudon 1974; Collins 1979). Boudon, using data from Western industrialized countries, and Collins, analyzing data from the United States, found that the tremendous expansion of education in the nineteenth and twentieth century left the opportunities for social mobility essentially unchanged. It did expand the educational attainment of many social groups, but, as the educational attainment of individuals from lower socioeconomic strata increased, individuals from higher strata acquired even more education, thus shifting the overall educational attainment of the population upward but keeping intact the stratification of educational attainment (Boudon 1974; Collins 1979). Given that meritocracy in the allocation of social positions exists, these findings suggest the lack of equality of educational opportunity.

In relation to the latter, the functionalist position is that schools do provide equality of opportunity. For empirical support, they point to the numerous empirical studies suggesting that the process of educational attainment is an achievement process. The best-known model of educational and status attainment in the United States, known as the Wisconsin model, describes the process as one whereby individual and background characteristics are translated into differential status attainment only after they have been transformed into individual performance and psychological variables (Sewell et al. 1969; 1970). Although this model has been criticized for excluding social constraints and related structural variables (see Kerckhoff 1976), its explanatory power remains strong, and it has withstood a number of replications (Alexander, Eckland, and Griffin 1978; Jencks et al. 1983). Indeed, most of the research on the

social selection process that followed the Wisconsin model has shown that schools select, process, and reward students based on individual traits and achievements, such as aspirations and ability, and that educational attainment in turn is the major determinant of occupational status attainment.

Conflict theorists and researchers, by contrast, have not been very successful in describing and explaining a social selection process that leads to social reproduction. The explanations and the evidence as to why individuals from higher social strata acquire consistently better education have not been able to dispute or account for the fact that the educational selection process is ostensibly an “achievement” process. In general, the argument from a conflict perspective is that structural limitations imposed on the schooling of some groups restrict their educational success, thus helping to reproduce the educational and social hierarchies.

Some structural limitations on both the quality and quantity of educational opportunity of children from low socioeconomic strata do indeed exist. Differential quality and quantity of schooling may have been especially influential in the past, but it still exists today. Some of the best schools at all educational levels in the United States are private, with high tuition, and obviously not all social groups have equal access to or success in these institutions. Also, fewer institutions, especially at the postprimary levels, are available in rural or low-income areas. Nonetheless, the impact of variation in quality and quantity of schooling has been reduced over the years, and evidence does not indicate it as a major determinant of educational attainment. For example, the well-publicized report *Equality of Educational Opportunity* (Coleman et al. 1966) found that differences between public schools had no significant effect on student performance. In general, even though some relevant quality differences between schools may still exist today, this structural variable is a relatively weak factor in explaining differential educational attainment.

Other structural variables, such as curriculum tracking (Alexander, Cook, and McDill 1978) and differential treatment by teachers and counselors (Karabel 1972; Rist 1970), also have been found to exert significant influence on educational attainment. In addition, researchers have found that

cultural differences linked to differential social origin are also responsible for the unequal educational attainment of students from different social groups (Bourdieu 1977; Bourdieu and Passeron 1977; DiMaggio 1982). Overall, however, structural limitations and cultural deficiencies account for only a small amount of attainment differences as compared to the individual achievement variables.

Summarizing the findings on equality of educational opportunity and meritocracy presents a paradoxical picture of social stratification and leaves the issue of social mobility through education largely unresolved: Status attainment research has shown that educational and status attainment is a meritocratic process based on individual achievement variables, but it has not explained the relatively low social mobility rates. Critical research, on the other hand, has shown reproduction of social status, but it has not been able to unseat the equality of opportunity thesis resting on the association of individual achievement with educational status attainment.

This apparent paradox may be due in part to the fact that research on educational and social stratification in the last few decades has been dominated by the ascription–achievement controversy without necessarily examining the relationship between this controversy and the broader mobility–reproduction debate. The underlying assumption of this focus seems to have been that achievement leads to mobility and ascription leads to reproduction. But however important achievement and ascription may be, they do not address the same issues as mobility and reproduction. A way out of this impasse may be to challenge the assumed correspondence between what we traditionally consider individual achievements on the one hand and social mobility on the other. There is no reason to assume that an empirical finding of schooling as an achievement process is necessarily incompatible with a theoretical argument of schooling as a process of reproduction. As long as individual qualities and achievements are determined by the social origin of the student, educational systems can promote not only social reproduction and individual achievement but also social reproduction *through* achievement (Katsillis and Rubinson 1990).

One of the most consistent findings of the research on educational and status attainment is

that the socioeconomic status of the family influences the whole educational process, including many, if not all, individual student achievements and abilities that lead to socioeconomic status attainment. Thus, once the assumption that achievement implies social status or class mobility is abandoned, there is no contradiction between the findings of status attainment research that indicate an achievement-oriented educational selection system and the findings of critical research that schools reproduce social status or class inequalities.

This, of course, poses some interesting questions, especially in relation to the meaning and the role of equality of educational opportunity as we understand it today: Does a process that transforms family status inequality into differential individual achievements or qualities and, subsequently, into unequal educational attainment constitute equality of educational opportunity? If it does not, what constitutes equality of educational opportunity, and how is it attained? We may have to rethink the whole process of equality of opportunity before we are able to provide satisfactory answers to the question of education and social mobility.

(SEE ALSO: *Equality of Opportunity*; *Social Mobility*; *Social Stratification*; *Sociology of Education*; *Status Attainment*)

REFERENCES

Alexander, K. L., M. A. Cook, and E. L. McDill 1978 "Curriculum Tracking and Educational Stratification: Some Further Evidence." *American Sociological Review* 43:47-66.

Alexander, K. L., B. K. Eckland, and L. J. Griffin 1978 "The Wisconsin Model of Socioeconomic Attainment: A Replication." *American Journal of Sociology* 81:324-342.

Blau, P. M., and O. D. Duncan 1967 *The American Occupational Structure*. New York: Wiley.

Boudon, R. 1974 *Education, Opportunity and Social Inequality*. New York: Wiley.

Bourdieu, P. 1977 "Cultural Reproduction and Social Reproduction." In J. Karabel and A. H. Halsey, eds., *Power and Ideology in Education*. New York: Oxford University Press.

—, and J. C. Passeron 1977 *Reproduction in Education, Society, and Culture*. Beverly Hills, Calif.: Sage Publications.

Bowles, S., and H. Gintis 1976 *Schooling in Capitalist America*. New York: Basic Books.

Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York 1966 *The Equality of Educational Opportunity*. Washington, D.C.: U.S. Department of Health, Education, and Welfare.

Davis, K., and W. Moore 1945 "Some Principles of Stratification." *American Sociological Review* 10:242-249.

DiMaggio, P. 1982 "Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students." *American Sociological Review* 47:189-201.

Duncan, O. D., D. L. Featherman, and B. Duncan 1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.

Grusky, D. B., and A. A. Takata 1992 "Social Stratification." In E. F. Borgatta and M. L. Borgatta, eds., *Encyclopedia of Sociology*. New York: Macmillan.

Jencks, C., J. Crouse, and P. Mueser 1983 "The Wisconsin Model of Status Attainment: A National Replication with Improved Measures of Ability and Aspirations." *Sociology of Education* 56:3-19.

Jencks, C., M. Smith, H. Acland, J. M. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson 1972 *Inequality: A Reassessment of Family and Schooling in America*. New York: Basic Books.

Karabel, J. 1972 "Community Colleges and Social Stratification." *Harvard Educational Review* 42:521-562.

—, and A. H. Halsey, eds. 1977 *Power and Ideology in Education*. New York: Oxford University Press.

Katsillis, J., and J. M. Armer 1992 "Class Reproduction and Status Competition: Incompatible or Complementary Processes?" Paper presented at the annual meeting of the Southern Sociological Society, New Orleans, LA.

— and R. Rubinson 1990 "Cultural Capital, Student Achievement, and Educational Reproduction: The Case of Greece." *American Sociological Review* 55:270-279.

— and J. A. Spithourakis 1997 "Social Determination of Language Learning in Greek High Schools." Paper presented at the Twelfth International Congress of the World Association for Educational Research (WAER), Rethymno, Crete.

Katsillis, J. M., C. Alexopoulos, P. Moustairas, and P. Skartsilas 1997 "Equality of Educational Opportunity and Social Equality: Concepts, Theory and Practice." In T. Moyne, ed. *Sociology of Greek Education: Review, New Research and Prospects*. Pp. 47-60, Patras, Greece: Achaikes Ekklisis. In Greek.

- Kerckhoff, A. C. 1976 "The Status Attainment Process: Socialization or Allocation?" *Social Forces* 52:368-381.
- Poulantzas, N. 1974 *Classes in Contemporary Capitalism*. London: Verso.
- Rist, R. C. 1970 "Student Social Class and Teacher Expectations: The Self-Fulfilling Prophecy in Ghetto Education." *Harvard Educational Review* 40:411-451.
- Sewell, W. H., A. O. Haller, and G. W. Ohlendorf 1970 "The Educational and Early Occupational Status Attainment Process: Replications and Revisions." *American Sociological Review* 34:1014-1027.
- , and A. Portes 1969 "The Educational and Early Occupational Status Attainment Process." *American Sociological Review* 34:82-92.
- Willis, P. 1977 *Learning to Labor*. New York: Columbia University Press.
- Wright, E. O. 1978 *Class Structure and Income Determination*. New York: Academic Press.
- 1985 *Classes*. London: Verso.

JOHN KATSILLIS  
J. MICHAEL ARMER

## EDUCATIONAL ORGANIZATION

Education and schooling are not synonymous. Education is the more encompassing concept, referring to the general process by which a social group—whether an entire society, a family, or a corporation—transmits attitudes, beliefs, behaviors, and skills to its members. Within these broad boundaries, we can distinguish three general types of education—informal, nonformal, and formal—according to the location of instruction, the characteristics of the teachers, the methods of instruction, and what is learned.

Informal education takes place in the context of everyday life, and the educators include family members, peers, workmates, and the mass media. Formal education or schooling, meanwhile, takes place outside the family in institutions that specialize in education, is conducted by teachers who are not students' intimates and whose principal occupation is education, and stresses learning more through verbal and written description and guided inquiry than through observation and imitation. Finally, nonformal education—which takes

such forms as on-the-job training, agricultural extension programs, and family-planning outreach programs—is more organized than informal education but has aims that are more specific and short term than those of formal education.

Virtually all societies utilize all three forms of education, but they differ in the relative predominance of these forms. In nonindustrialized societies, informal education dominates, with formal and nonformal education only marginally present. But in industrialized societies such as the United States, formal education rivals, if not exceeds, nonformal and informal education in importance and the use of society's resources. However, this is not to say that such agencies of informal education as the mass media do not have very profound effects. The ubiquity of the modern mass media and the fact that they are now held in relatively few hands allow them to widely and deeply shape many of our beliefs, attitudes, and behaviors (Bagdikian 1990; Herman and Chomsky 1988).

## THE STRUCTURE OF NATIONAL SCHOOL SYSTEMS

School systems across the world are converging more and more in structure and content (Meyer, Kamens, and Benavot 1992; Meyer, Ramirez, and Soysal 1992). Yet school systems still differ considerably, even among countries comparable in economic development. One key axis of variation is relative size. Nations greatly differ in the proportion of their total population, especially the young, enrolled in school. For example, in 1994 the proportion of youth of secondary school age enrolled in school averaged 94 percent across twenty-one advanced industrial societies (sixteen European countries, the United States, Canada, Japan, Korea, and Australia). But the average percentage was 55 percent for eighteen less developed Asian countries (excluding Japan, Korea, Taiwan, Hong Kong, and Singapore) and 32 percent for fifteen African countries (U.S. National Center for Education Statistics 1997).

Clearly, differences in wealth and degree of industrialization explain a major part of this variation. But even when we control for these factors, we still find enormous differences among societies in the structure of their school systems.

**Economically Advanced Countries.** Even economically advanced societies differ greatly in how their school systems are governed and how their students' school careers are structured. Nations differ greatly in how much control the national government exercises over how schools are financed and operated. There are several countries that lodge governance primarily at the subnational or provincial level, such as the United States, Canada, and Germany. But most advanced societies vest control in a national central educational authority, usually a national Ministry of Education. For example, the Japanese Ministry of Education provides most of the funding for schooling, determines national curriculum requirements (the subjects to be taught and the depth in which they are to be covered), selects lists of acceptable text books, sets standards for teacher training and certification, and administers the 166 or so public universities. To be sure, local prefectural boards establish or close schools, hire and supervise teachers, and plan the curriculum. But they do all this within parameters set by the national ministry, which can veto their decisions (Kanaya 1994).

Nations with strongly centralized school governance leave much less room for local control and therefore for local variation in the content and structure of schooling. But the flip side of the coin is that such nations also suffer from much less inequality in school spending across localities and—because of class and racial segregation in housing—across social classes and races.

School systems in advanced industrial societies vary also in the structure of students' careers; that is, the timing of career selection and therefore curricular differentiation; the proportion of secondary students specializing in vocational studies; the strength of the tie between vocational training in secondary school and labor-market outcomes; the proportion of students entering and graduating from higher education; and the degree of differentiation within higher education, including whether there is an elite sector with privileged ties to elite public and private jobs (Brint 1998; Hopper 1977). The United States and Germany are nearly polar opposite on virtually all these dimensions.

The United States puts off occupational selection until very late. The main branching point comes after high school, when a student decides

whether to go to college, which college to enter, and later what field to major in. As a result, U.S. high schools have a weak connection to the labor market. Because of this, many current educational reforms—such as school-to-work partnerships between schools and employers—are directed toward enhancing the connection between secondary school curricula and labor market opportunities.

Meanwhile, the proportion of students entering and graduating from higher education is huge. In 1996, 65 percent of high school graduates (or about 58 percent of all college-age youth, given a dropout rate of about 11 percent) entered higher education. According to the High School and Beyond Survey, about one-quarter of college entrants eventually receive a baccalaureate degree or higher and another one-fifth receive a one- or two-year certificate or degree. These figures for college entrance and graduation are about double those for Germany (Organization for Economic Cooperation and Development 1998; U.S. National Center for Education Statistics 1997). Because so many U.S. students go on to college and because student selection occurs to a great extent within college, U.S. higher education institutions are quite varied in curriculum, prestige, and student-body composition (see below).

Germany, meanwhile, has a very different school system. Student selection occurs at age ten, when students are divided between academic high school (Gymnasium) and two types of vocational secondary schooling (Realschule and Hauptschule). At age fifteen students graduate from the vocational high schools into either more advanced vocational schools or apprenticeship programs combining on-the-job and classroom training. Both are strongly connected to specific employment. Meanwhile, graduates of the Gymnasium go on to take the Abitur exam, which determines if they will be allowed into university. All told, only one-quarter of German students enter the university and only 15 percent get university degrees (Brint 1998; Organization for Economic Cooperation and Development 1998).

These differences in student careers fundamentally shape educational outcomes. A tightly coupled school career system, in which test results largely determine admission to the best schools and graduation from them in turn strongly shapes

job placement, will tend to produce students who work hard at their schooling and their exam performance, as is the case in Japan (Brint 1998; Rohlen 1983). But the effects of school structure reach further. School systems that have small, highly selective higher educational sectors with little or no distinctions made among universities, as in the case of Germany, will tend to generate greater class consciousness and solidarity. In contrast, schools systems with large, internally heterogeneous higher education sectors, such as in the United States, foster weaker class consciousness (Brint 1998).

**Economically Less Developed Countries.** Economically less developed countries (LDCs) vary greatly as well in the size and structure of their school systems. For example, in Africa, the ratio of secondary school students to the secondary-age population ranges from 7 percent in Mozambique to 77 percent in Egypt (U.S. National Center for Education Statistics 1997). In addition, LDCs vary greatly in how socially exclusive their higher education systems are; for example, in 1985 the proportion of postsecondary students who are female ranged from 24 percent in sub-Saharan Africa to 52 percent in the Caribbean (Ramirez and Riddle 1991).

A major source of this diversity in size and structure is, of course, differences in degree and form of economic development even among less developed societies. But other factors also play an important role in causing this variation. Though most LDCs were at some point colonies or protectorates of one of the European powers or the United States, this colonial inheritance was not homogeneous. For example, the British and French colonial heritages were quite different, rooted in the different educational and political systems of those two countries. British colonies typically had higher rates of college attendance and lower rates of grade repetition than French colonies, echoing the differences between their colonial masters' own school systems (Brint 1998). Furthermore, the nature of the political elites—whether enthusiastic modernizers as in Turkey or Iraq or conservatives as in Saudi Arabia—has made a difference in how much emphasis they put on expanding the school system (Brint 1998).

Despite these variations, educational systems in economically less developed countries (LDCs)

do exhibit considerable homogeneity in structure. A lack of resources has tended generally to force a lower level of educational provision (Brint 1998). In addition, many LDCs share a common colonial inheritance; for example, across the former British colonies in Africa and the Caribbean, secondary education remains dominated by the British “O-level” and “A-level” examinations (Brint 1998). Moreover, modernizing movements of quite various ideological stripes have seen education as a way of creating loyalty to and solidarity with their new ideas (Brint 1998; Meyer, et al. 1992). Finally, the World Bank has been playing a homogenizing role by strongly urging particular reforms (such as emphasizing primary over tertiary education and deemphasizing vocational education) on nations applying for loans (Brint 1998).

#### A CLOSER LOOK AT THE STRUCTURE OF THE U.S. SCHOOL SYSTEM

As noted above, the U.S. school system is quite unlike that of most other advanced industrial societies. The United States is virtually unique among advanced societies in that education is not mentioned in the national constitution and educational governance is not lodged with the national government (Ramirez and Boli-Bennett 1982). Instead, schooling in the United States is a state and local responsibility. Consequently, the United States has more than fifty separate sovereign educational authorities.

The United States has no national universities (except for the military academies and a few other specialized institutions). There is no required national exam for university entrance. While the Scholastic Assessment Test (SAT) and the American College Testing (ACT) exam are widely used, they are privately operated and individual colleges decide whether and how their results will be used.

The connection between educational credentials and workplace opportunities is comparatively weak in the United States (Collins 1979). Of course, for some occupations, the connection is quite strong, with a standardized curriculum preparing graduates for licensing examinations. But for most college graduates, the connection between their college major and their work careers is tenuous at best. Significant labor-market advantages go

to those who attend and graduate from college, but the school system has relatively weak connections to most occupational sectors. Recent reforms, such as the 1994 federal School-to-Work Act, aim to tighten the links between secondary and postsecondary training and the labor market (Van Horn 1995). But it will take many years of such efforts before the United States even approximates Germany or Japan in the closeness of linkage between school and work.

These structural features have created an educational system in the United States that is wide open and characterized by very high enrollments and great student and institutional diversity. We make available a seat in some college somewhere for virtually everyone who wants to attend. Consequently, our secondary education system is less decisive than in most other countries, as “second-chance” opportunities abound. Secondary school students do not have to make hard decisions about their educational futures until quite late, often in college.

In order to better understand these unusual features of the American system, let us examine the structure of U.S. education in greater detail.

**Elementary and Secondary Education.** All elementary and secondary (K–12) school districts operate within the confines of the relevant state education law, which specifies requirements for graduation, certification of teachers, and so forth. State governments also provide on average about 47 percent of public school funding, with most of the rest coming from local taxes (U.S. National Center for Education Statistics 1997).

Private schools, too, must conform to state education law, but they are less restricted than are public schools. For example, in most states, the regulations governing teacher certification are less strict if one teaches in a private school than a public school. It is almost entirely up to the private school and its sponsors to generate financial support. No tax-derived funds may be used to support private K–12 schooling unless special conditions are met (for example, private schools may receive public aid if they enroll handicapped students). Interestingly, public aid flows much more easily to private colleges. They can receive student financial aid, grants to build academic facilities, and grants

and contracts to conduct research and run academic programs.

The operation of public education at the elementary and secondary levels largely rests with the local school district. In 1995–1996 there were about 16,000 separate school districts in the nation, each with its own school board, superintendent, and schools (U.S. National Center for Education Statistics 1997). Although smaller school districts have often been consolidated into larger ones, many states still have hundreds of separate districts. School district boundaries are usually coterminous with local political boundaries, but elected school boards are rarely identified with a political party. Localities provide about 46 percent of public school funding (U.S. National Center for Education Statistics 1997).

This reliance on local revenues derived from property taxes generates great disparities in per-pupil spending across property-rich and property-poor districts. Though states have increased their share of educational expenses, largely due to legal challenges to relying on local property taxes for funding, spending disparities have decreased only a little. The poorest districts do have more money to spend, but rich districts have increased their tax levels in order to maintain their spending lead (Ballantine 1997).

An important consequence of the U.S. pattern of considerable local control is that local concerns are more likely to be reflected in school policies and practices than is common in countries with more centralized educational systems. Citizens elect local school boards and frequently vote on budgets, property tax rates, and bond issues. Moreover, parents exercise considerable informal political power through parent–teacher associations, informal conferences with school teachers and administrators, and decisions about whether to send their children to a particular school or not. (See the section below on modes of influence over schools.)

Despite the absence of strong national control, U.S. elementary and secondary schools do share many similarities across the country. One reason is that the federal government does exercise a homogenizing influence through its policy recommendations and funding for particular programs. In addition, national professional associations of educators and regional nongovernmental



accrediting agencies provide common definitions across states and localities of what constitutes good educational practice. Also, college admissions requirements, though they vary across colleges, are similar enough to influence the course offerings of secondary schools. Moreover, textbook writers and publishers, who provide instructional material for schools nationwide, influence what is taught and often how it is taught by marketing the same instructional materials nationwide (Apple 1986). Finally, the high geographic mobility of students and teachers has helped reduce the isolation and consequent diversity among schools.

Across the United States, formal public schooling generally begins at age six, but enrollment in preschool and kindergarten is widespread and growing. In 1995 about 61 percent of three- to five-year-olds were enrolled part time or full time in nursery schools or kindergartens (U.S. National Center for Educational Statistics 1997).

Elementary schools are smaller than secondary schools. They are also less differentiated internally, in that all students are exposed to essentially the same subject matter by their "home room" teacher. However, within the home room, teachers often do divide students into different groups that are supposed to learn the same material at different speeds. This within-class differentiation is often termed "ability grouping," but in actuality test scores are often only a weak predictor of group assignment. Nonetheless, this grouping by putative aptitude is an important source of later class and, less so, racial differences in achievement (Dougherty 1996).

While elementary schools are generally alike in organization and curriculum, they differ widely in student composition. Because they draw from neighborhoods differing in class and racial composition, they end up differing from each other in student composition.

Secondary schooling begins around age thirteen. For the most part, U.S. secondary schools are "comprehensive"; that is, intended as much for those who will not attend college as for those who will. The comprehensive high school provides college preparation, vocational education, and general secondary education under one roof (Clark 1985; Krug 1964, 1972).

The comprehensive nature of U.S. secondary schools is fairly atypical, for the usual pattern abroad is to have separate academic and vocational high schools. For example, in Germany, academic and vocational training is assigned to separate secondary schools, with nearly half of all students entering the latter (Brint 1998).

The United States is also atypical among industrialized societies in awarding secondary school diplomas qualifying their holders for college entrance solely on the basis of the number and kinds of courses taken. Most other countries require passage of a national exam to receive a degree that qualifies one for university entrance (Brint 1998). To be sure, seventeen U.S. states do use minimum-competency examinations for awarding high school degrees (Airasian 1987; U.S. National Center for Education Statistics 1997), but the country still does not have a European-style national examination that alone determines university entrance.

The exceptional devotion of the United States to comprehensive schooling is traceable to two factors. The strong local role in educational governance in the U.S. system makes it more likely that the demands of non-college-goers will be listened to. Moreover, the heterogeneity of the U.S. population has made social integration a more pressing concern than in most European societies. This is evident in the words of the highly influential 1918 report, *Cardinal Principles of Secondary Education*: "The comprehensive school is the prototype of a democracy in which various groups have a degree of self-consciousness as groups and yet are federated into a larger whole through the recognition of common interests and ideals. Life in such a school is a natural and valuable preparation for life in a democracy" (National Education Association 1918, p. 26).

While U.S. comprehensive schooling may have been successful in its aim of social integration, its diffuse character has also been widely criticized. As many recent critics have noted, the variety of curricular goals and educational and social purposes served by U.S. secondary schools blurs their academic mission. When large proportions of students are not particularly academically inclined, the rigor and sense of purpose necessary to motivate student effort are missing. Moreover, the huge size of many U.S. schools makes them impersonal and hard put to maintain the involvement

and commitment of students (Cusick 1983; Goodlad 1984; Powell et al. 1985; Sizer 1985).

While the U.S. school system is much less differentiated than is typical abroad, U.S. secondary schooling is by no means entirely undifferentiated. To begin with, there is an extensive private sector. Ten percent of all U.S. K-12 students attend private schools. These schools vary enormously, from individual Montessori schools, Christian academies, and elite private schools to city-wide systems of Catholic parochial schools. While small in numbers and enrollments, the elite private schools, which are variously termed "prep" or "boarding" or "country day" schools, carry great prestige and importance. Most areas of the United States have elite schools, but the most famous are the boarding schools of New England, such as Phillips Exeter, Choate/Rosemary Hall, Groton, Hotchkiss, and St. Paul's. These schools have enrolled such famous Americans as Franklin D. Roosevelt, John F. Kennedy, Adlai Stevenson, George Bush, and Nelson Rockefeller. The importance of these schools lies not only in the fact that they heavily enroll the sons and daughters of the upper and upper-middle classes, particularly those of long-established wealth and prominence, but also in the fact that they provide their students with privileged access to the top universities and, in turn, corporate and governmental leadership (Cookson and Persell 1985; Hammack and Cookson 1980).

But it is not just the public/private divide that provides differentiation within the U.S. system. Even the comprehensive public high schools provide alternatives within their walls in the form of different curricular groupings (college prep, vocational, and general) and courses at different levels of rigor. However, this phenomenon, typically termed "tracking," has been criticized as a significant source of class and racial inequality in educational attainment (Dougherty 1996). Consequently, a movement has developed to "detrack" schools by eliminating grouping and instead relying on "cooperative learning" within mixed-ability classrooms (Oakes and Lipton 1992; Wells and Oakes 1996).

In addition to different tracks, most U.S. urban school districts maintain specialized vocational and academic secondary schools. For example,

Boston, New York, Philadelphia, and San Francisco have maintained old and distinguished academic high schools such as Boston Latin.

Since about 1970, most urban school systems have introduced a wide variety of programs in order to meet demands for more choice, retain middle-class white students, and better motivate students (Dougherty and Sostre 1992). New York City provides a good example of how highly differentiated some urban school systems have become. The city does have a large number of general "academic/comprehensive neighborhood schools." However, within many of these schools, there are "academies" or other magnet programs, which are operationally independent and have some freedom to select their students from wider attendance areas. In fact, there are about fifty small alternative high schools that offer special curricular emphases and that are open to students citywide. New York City also has a variety of vocational schools. Special "educational opportunity" high schools are organized around vocational themes (such as health, business, or aviation) and have the right to select their students. On the other hand, there are also ordinary vocational schools that have no particular focus and have open admissions. Finally, the city boasts four very well-known college preparatory schools, such as the Bronx High School of Science, that grant entry solely on the basis of an examination or audition (Board of Education of the City of New York 1997).

In recent years, a new form of differentiation has arisen within the public schools: "charter schools." Since 1991, more than thirty-five states have passed legislation making these schools possible, and perhaps more than 1,000 of them are now operating. Once granted a "charter" by the state or other designated authority, these schools operate independently of many existing school regulations but are financed by funds that would otherwise go to the districts. Charter schools are accountable to the chartering authority, and the renewal of their charter depends on meeting the goals set forth in their mission statement. The idea behind these schools is to free public school parents, teachers, and administrators to create schools that "break the mold" of existing schools and, by competing with existing public schools for students, force them to improve their performance and attractiveness. Moreover, many charter advocates have seen charter schools as a way of meeting

the growing parental demand for choice—among not only affluent white parents but also working-class minority parents—but keeping it from tipping into a demand for vouchers to allow student to attend private schools. Beyond these commonalities, charter schools are very diverse in size, mission, student composition, and sponsorship. And because of this variation and their youth, it is unclear what impact charter schools will ultimately have. The jury is still out on whether they will enroll more than a fraction of public school students, successfully “routinize the charisma” of their founders after those founders move on, significantly enhance the performance of their students, and effectively stimulate regular public schools to improve. Many fear that charter schools may simply cream off the most advantaged students and leave the regular public schools more segregated and academically impoverished than ever (Cobb and Glass 1999; Manno et al. 1998; Wells et al. 1999).

**Higher Education.** As with the K–12 system, the U.S. higher education system is also quite unusual. It is much larger than, and organized very differently from, most other nations’ systems. As of 1995, the United States had 3,706 institutions of higher education enrolling 14.3 million students in credit-bearing courses, which corresponded to about 35 percent of the population age 18–21 (keeping in mind that many college students are older than twenty-one). In addition, there were some 6,300 noncollegiate postsecondary institutions enrolling 850,000 students (U.S. National Center for Education Statistics 1997). All these numbers are much larger than those for comparable advanced industrial societies.

Given their number, it is not surprising that American colleges are quite varied. This variation can be usefully categorized along three axes: control, degrees and programs offered, and student-body composition.

U.S. colleges are legally owned by a wide variety of bodies. Some 1,700 colleges are public, owned by local, state, and federal governmental bodies. They account for 45 percent of all colleges but 78 percent of all college enrollments. Meanwhile, about 2,000 colleges are private, owned either by religious groups, profit-making corporations, or nonsectarian, non-profit-making boards. The nonsectarian, nonprofit private institutions

include both many of the most prestigious doctorate-granting universities in the world and many small, undistinguished liberal arts colleges (U.S. National Center for Education Statistics 1997).

The U.S. Department of Education distinguishes five kinds of colleges according to the degrees and programs that they offer. The first group is the 171 “doctoral” institutions that offer doctoral and professional programs and produce a large number of graduates with either Ph.D.’s or medical and dental degrees. “Comprehensive” institutions, numbering about 420, make up the second category. They offer graduate programs but graduate few people with doctoral or medical degrees. Rather, they specialize in undergraduate, master’s, and law programs. Quite often these institutions are former teacher-training colleges that broadened into general liberal arts schools and added graduate programs. The nearly 700 “general baccalaureate” or “liberal arts” colleges emphasize undergraduate education and have very few, if any, graduate programs. “Specialized” colleges, which number about 600, emphasize one field, such as engineering or the arts, and offer either a baccalaureate or postbaccalaureate training. Finally, the nearly 1,500 “two-year colleges” specialize in two-year associate’s degrees, one-year vocational certificates, and noncredit training. They enroll not only college-age students seeking academic or vocational training but also older adults seeking job retraining, skills upgrading, or avocational knowledge (Cohen and Brawer 1996; Dougherty 1994).

U.S. higher educational institutions also differ in their student-body composition. As one moves from universities to four-year colleges to two-year colleges, the proportion of students who are male, white, upper-class, or academically high-performing drops. In addition, some colleges serve distinct student populations; for example, nearly 200 colleges are single-sex and nearly 100 are all-black (U.S. National Center for Education Statistics 1997).

#### MODES OF INFLUENCE OVER SCHOOLS

The governance of education has been a repeated motif in our discussion above. We would now like to discuss it in greater detail. So far, in our discussion of control we have focused on political authority, whether exercised by national, state, or

local governments or the citizens that elect them. But political authority is only one of several, often contradictory, mechanisms of influence over U.S. schooling. Also operative are market competition, bureaucratic power, professional authority, and ideological formation (Weiss 1990).

**Political Authority.** Political authority is vested in the various elected bodies of government and ultimately in the citizenry. State governors, legislatures, and boards of education control the schools through state funding (which amounts to nearly half of all public school revenues) and through laws specifying minimum curriculum and graduation requirements, the minimum length of the school day and year, required facilities, standards for teacher education and certification, standards for school plant, school district lines, and so forth (Campbell et al. 1990; Wirt and Kirst 1992).

However, state governments delegate political authority over the day-to-day operation of schools to local schools boards elected by local citizens. These local boards in turn have the power to hire and supervise district superintendents and school principals. The boards also vote on the school budget, the local tax rate (though usually subject to voter referendum), curriculum, teaching, facilities standards beyond state minimal, and the rules for hiring and supervising teachers (Campbell et al. 1990; Wirt and Kirst 1992).

The federal government, meanwhile, only contributes about 7 percent of K-12 public school revenues, mostly in the form of categorical aid (discussed below under "Market Competition"). However, through the federal courts, the federal government has had a profound effect on school policies involving the treatment of pupils, particularly women, racial and linguistic minorities, and the handicapped. Moreover, as will be discussed below, the federal government has also exercised great ideological power (Campbell et al. 1990; Wirt and Kirst 1992).

Citizens, finally, exert political authority. Very frequently they vote on who will represent them on a school board or in state office. They also exercise direct democratic control by voting on tax rates and bond issues through local and state referenda and initiative elections. And in states such as California and Washington, it has become commonplace for voters to vote on school policies

such as affirmative action in student admissions and teacher hiring (Wirt and Kirst 1992).

However legitimate and powerful political authority is, it can also be ineffective, particularly in a highly decentralized political system such as that of the United States. When a host of different government bodies impose multiple, often conflicting, mandates on schools, the effectiveness and authority of any one given political body is undermined (Weiss 1990).

**Market Competition.** Market control is less coercive than political authority. Schools can refuse to act in the way a market actor wishes, but that actor achieves compliance by supplying or denying resources that the school values and that the school cannot easily acquire from alternative sources (Pfeffer and Salancik 1978). These resources include funds, students, teachers, and jobs. Market control is particularly obvious in the case of private schools, which usually do not have a guaranteed clientele or funds and must recruit new students every year. But public schools also face market competition. In fact, market control over public schooling is steadily rising with the current vogue for school choice, charter schools, performance-based funding, and other means of making public schools more "accountable."

Students and their parents exercise market control over schools through their decisions about which schools to attend (Spicer and Hill 1990; Weiss 1990). If many middle-class students desert a school district, it loses state funding, which is largely enrollment-driven. Moreover, local tax revenues may also decline. Real estate values are strongly affected by perceptions of the quality of local schools, and these perceptions are in turn shaped by how many middle-class and white students attend a school. But even if school revenues are not affected, the desertion of middle-class students can still affect schools by impinging on teachers' perceived quality of work. Teachers usually clamor for better-prepared students, and a loss of middle-class students can lead the better teachers to themselves desert a particular school.

In order to retain students, particularly middle-class white ones, school districts adopt a variety of expedients. They create "gifted" programs or "magnet" schools that attract such students not only by offering superior academic resources but

also by largely segregating them from working-class and nonwhite students (Metz 1986; Wells 1993).

School competition for students has risen in the 1990s. It used to be that the main choices parents had available were between sending their children to public or private schools or between living in one school district versus another. But competition for students has increased with the advent of greater choice within the public schools in the form of magnet schools, charter schools, and interdistrict choice plans (Cookson 1994; Metz 1986; Wells 1993).

Students and their parents exert market power not only over schools overall but also over the classrooms within them. Student decisions about whether to take one or another course or whether or not to actively participate in class deeply shape the character of teaching and learning within classrooms. In order to attract students and then motivate them to work hard and actively participate, teachers often resort to such devices as giving students more choice over course selection or course content, making course content less abstruse and technical (more "relevant"), reducing academic demands, and grading less stringently (Powell et al. 1985; Weiss 1990).

But teachers themselves are also market actors. Teachers can choose whether to go to work for one or another district or, if they have enough seniority, work at one school versus another in the same district. Hence, school districts compete to hire and retain teachers, particularly if they are in fields such as math and science, where qualified teachers are scarce. Moreover, teachers have shaped the schools through their collective capacity to withhold their labor through unions. With the rapid growth of teacher unions since the early 1960s, teachers have been able to secure considerably higher salaries and greater voice in how schools run than they had before (Campbell et al. 1990; Kerchner et al. 1993).

Business also shapes schools through market control. For one thing, business controls jobs. Schools compete to place their students in good jobs because a good placement record can be used in turn to attract students. In order to get their students placed in good jobs, schools inculcate the kinds of skills, attitudes, and behaviors that business is looking for in new workers (Brint and

Karabel 1989). In fact, business's influence based on its role as future employer of students has been institutionalized in the form of a myriad of business/school or school-to-work "compacts" or "partnerships" in which formal links are established and schools receive resources and job placements in return for greater responsiveness to business opinions about the desirable content of education (Gelberg 1997; Van Horn 1995).

In addition, business along with foundations and government influence schools through discretionary funding. Almost all business and foundation aid and most federal aid to schools takes the form of categorical grants. These funds will flow to a school only if it successfully competes with other schools to demonstrate that it is willing and able to engage in actions that the funder wishes to encourage. Moreover, this avenue of market control is increasing, as state governments establish not only more categorical grant programs but also performance funding, in which a certain portion of state formula aid is conditioned on meeting certain performance targets.

Market competition can be a very powerful control device, but it is also less effective than its evangelists believe. For example, greater student choice may not cause the deserted schools to change. The schools may not know why students are leaving, and the loss of funds, good students, and good teachers may impede its capacity to improve. Moreover, even if schools do react, they may get the wrong cues because students and their parents make bad choices due to lack of good information (Weiss 1990). Similarly, categorical aid often fails to accomplish its purpose. What schools do to secure the aid may bear little resemblance to how the aid is actually used. This has been a perennial problem with federal Title I funding for high-poverty schools. It often does not end up benefiting the students to which it is ostensibly targeted (Somini 1999).

**Bureaucratic Power.** Schools are bureaucratic organizations. They have explicit goals. Their work is done through a division of labor involving specialized jobs. There is a chain of command, with explicit differences in the authority of members according to their place in the organizational hierarchy. The members' actions are largely governed by formal rules and a norm of professionalism (impersonality). Organizational decisions are

recorded through explicit and voluminous records. And personnel decisions are supposedly governed by merit (Bidwell 1965).

Within this bureaucratic structure, administrators—such as district superintendents or college presidents, school principals or deans—exercise great power. They create jobs and define their responsibilities, establish organizational rules, allocate scarce resources (money, space, staff, etc.), order specific actions, referee conflicts among subordinates, and hire and supervise subordinates (Campbell et al.; Weiss 1990).

Historically, teachers have been objects of administrative power. But increasingly, they themselves are participating in the exercise of administrative duties. The movement for school-based management has given teachers the potential to exercise greater power over how schools are run, though it is still not clear to what extent this has become a reality. In numerous communities, school councils have been set up that include teacher members. These councils have the authority to exercise considerable voice over such things as a school's budget, teacher hiring (what areas to hire in and whom to recommend to the district), and student discipline rules (Kerchner et al. 1993; Mohrman et al. 1994). However, this authority is often not exercised in practice. Between principal resistance to sharing authority and teacher reluctance to assume it, school councils often end up exerting much less authority than authorized.

Bureaucratic control had become perhaps the dominant form of school control by the end of the Progressive era. But despite its power, bureaucratic control does not handle localized, specific situations well. The general orientation of bureaucratic rulemaking is toward general prescriptions because the aim is to circumscribe the discretion of organizational staff (Weiss 1990). However, student-centered education—particularly in a highly diverse, politically decentralized society such as the United States—often does not fit easily within bureaucratic universalism.

**Professional Authority.** However bureaucratic schools are, they are also professional organizations because teachers make up such a large portion of the labor force and administrators are invariably former teachers. The main fount of professional authority lies in the fact that effective teaching requires the exercise of discretion—how

teachers are to interact with students cannot be prescribed—and teachers largely monopolize the knowledge necessary to correctly exercise that discretion. Teachers use their professional authority to strongly shape curriculum, student evaluation, student discipline, proper classroom practices, and teacher training (Weiss 1990).

Nonetheless, teacher professional authority has always been uncertain and contested. This authority is at its apex in the classroom and fades as one goes up the bureaucratic hierarchy (Metz 1978; Weiss 1990). The weakness of teachers' claims for professional power and autonomy results from several factors. The majority of teachers are women, giving teaching less status than more male-dominated professions. Also, because of the unique complexity of the teacher–student relationship, teachers are less able to deliver consistent results than members of professions such as medicine and engineering. Finally, teachers must coexist with a powerful and numerous body of competitors for influence over students; namely, parents. (We discuss teacher–parent struggles for control below.) In the face of these limits to professional authority, teachers have increasingly resorted to market control, in the form of unionism, in order to bolster their influence over the schools.

**Ideological Formation.** Various actors can shape schools by the power of their ideas; that is, by their successful socialization of educational policy makers to certain values and beliefs (Weiss 1990). This ideological power has been strongly used by the federal government. Repeatedly, it has stimulated schools to take action by focusing attention on certain problems or offering exemplary solutions. One of the most notable examples has been the educational “excellence” movement of the 1980s and 1990s, which was strongly accelerated—though not really sparked—by the report of the National Commission on Excellence in Education, *A Nation at Risk* (1983). Within a year of its publication, many states and localities established commissions similar to the National Commission and passed laws implementing its recommendations. On a more global level, the power of ideological persuasion can be seen in the unusual homogeneity across the world in how nations have pursued the expansion and centralization of their educational systems. In large part this commonality of action is rooted in widespread support for a model of societal modernization that emphasizes

national unification and development by means of the mass mobilization of citizens through a unified school system (Ramirez and Boli 1982).

Once established, ideological control can be extremely powerful and durable. Its main limitation is that it usually takes a long time to establish. And unlike the other forms of power, it is particularly dependent on the willing acquiescence of those who would be influenced (Weiss 1990).

#### **Conflict Between Various Modes of Control.**

Many actors attempting to influence the schools utilize—wittingly or unwittingly—several of these modes of control. For example, state governments use political authority, market competition (through categorical aid), and ideological persuasion to get school personnel to act in certain ways. These different modes of influence can often yield great power if they are effectively meshed. But quite often they contradict each other.

Assertions of bureaucratic authority have been met by counterclaims by teachers in the name of professional authority or market control. Teachers have resisted state and local expansions of bureaucratic authority by mobilizing professional associations (for example, the National Council of Teachers of Mathematics) to shape the content of state curriculum standards (Massell 1994; Ravitch 1995). But because assertions of professional authority are often resisted, teachers have resorted as well to market control, in the form of teacher unionism.

But these assertions of power by teachers and by local and state school bureaucrats have in turn provoked democratic counterclaims by groups representing conservative parents concerned about parental prerogatives over education. These groups have strongly criticized teacher unionism and various curricular and pedagogical innovations advocated by teacher professional associations at the national, state, and local levels. These controversial innovations have included not only sex education and values clarification but also state content, performance, and evaluation standards. In California and Pennsylvania, parents associated with such New Right groups as Citizens for Excellence in Education, Focus on the Family, and Eagle Forum have vociferously attacked statewide goals and standards, performance-based assessment, whole-language instruction, and conceptually oriented math education. These groups reject such

curricular and pedagogical reforms as ineffective and unwarranted educational experimentation on children that undermines parental prerogatives to determine the content of their children's education (Boyd et al. 1996; Kirst and Mazzeo 1995).

Interestingly, there has been little conflict between market control and democratic authority. For example, inadequate critical attention has been devoted to business/school partnerships and the question of how compatible are business desires and public interests in schooling. This absence of scrutiny may be due to the weakness of the socialist tradition in the United States. Because of this weakness, democracy and the market are seen in the popular mind as largely compatible. Both voting and buying tend to be seen similarly, as decisions by atomized actors operating on the basis of narrow self-interest. Public discussion and the public interest tend to be seen as no more relevant to voting than to buying. As a result, efforts to increase market competition within schooling through such devices as vouchers and charter schools are often portrayed in the U.S. as democratic innovations—because they “empower” individuals—whereas in Europe there is much more hesitation to equate consumer choice and citizen sovereignty (Whitty et al. 1998).

## **CONCLUSION**

The sociology of U.S. schooling can benefit enormously from keeping in mind several features of that system. The educational system goes well beyond the schools to include such other institutions as families, the mass media, employers, and churches. Even when we focus on the schools, it is important to keep in mind that the U.S. school system is highly unusual compared to those in other advanced industrial societies. And when we turn to control of the U.S. system, we need to look beyond political authority to also consider other, often contradictory, mechanisms of influence over the schools: market competition, bureaucratic decision making, professional authority, and ideological formation.

#### REFERENCES

- Airasian, Peter W. 1987 “State Mandated Testing and Educational Reform: Context and Consequences.” *American Journal of Education* 95 (May):393-412.

- Apple, Michael 1986 *Teachers and Texts*. New York: Routledge, Chapman and Hall.
- Bagdikian, Ben 1990 *The Media Monopoly*, 3rd ed. Boston: Beacon Press.
- Ballantine, Jeanne H. 1997 *The Sociology of Education: A Systematic Analysis*. Upper Saddle River, N.J.: Prentice Hall.
- Bidwell, Charles 1965 "The School as a Formal Organization." In James G. March, ed., *Handbook of Organizations*. Chicago: Rand McNally.
- Board of Education of the City of New York 1997 *Directory of the Public High Schools, 1997-1998*. New York: Author.
- Boyd, William Lowe, Catherine A. Lugg, and Gerald L. Zahorchak 1996 "Social Traditionalists, Religious Conservatives, and the Politics of Outcomes-Based Education." *Education and Urban Society* 28 (May):347-365.
- Brint, Steven 1998 *Schools and Society*. Thousand Oaks, Calif.: Pine Forge Press.
- Brint, Steven, and Jerome Karabel 1989 *The Diverted Dream*. New York: Oxford University Press.
- Campbell, Roald F., Luvern L. Cunningham, Raphael O. Nystrand, and Michael D. Usdan 1990 *Organization and Control of American Schools*, 6th ed. Columbus, Ohio: Merrill.
- Clark, Burton R. 1985 "The High School and the University: What Went Wrong in America" (2 parts). *Phi Delta Kappan* (February and March):391-397; 472-475.
- Cobb, Casey D., and Gene V. Glass 1999 "Ethnic Segregation in Arizona Charter Schools." *Educational Policy Analysis Archives* 7 (January). <http://epaa.asu.edu/epaa/v7n1z>.
- Cohen, Arthur M., and Florence Brawer 1996 *The American Community College*, 3rd ed. San Francisco: Jossey-Bass.
- Collins, Randall 1979 *The Credential Society*. New York: Academic Press.
- Cookson, Peter W., Jr. 1994 *School Choice: The Struggle for the Soul of American Education*. New Haven, Conn.: Yale University Press.
- , and Caroline Hodges Persell 1985 *Preparing for Power: America's Elite Boarding Schools*. New York: Basic Books.
- Cusick, Philip 1983 *The Egalitarian Ideal and the American High School*. New York: Longman.
- Dougherty, Kevin J. 1994 *The Contradictory College: The Conflicting Origins, Outcomes, and Futures of the Community College*. Albany: State University of New York Press.
- 1996 "Opportunity to Learn Standards: A Sociological Critique." *Sociology of Education* 69(5):40-65.
- , and Lizabeth Sostre 1992 "Minerva and the Market: The Sources of the Movement for School Choice." *Educational Policy* 6 (June):160-179.
- Gelberg, Denise 1997 *The "Business" of Reforming American Schools*. Albany: State University of New York Press.
- Goodlad, John 1984 *A Place Called School: Prospects for the Future*. New York: McGraw-Hill.
- Hammack, Floyd Morgan, and Peter W. Cookson, Jr. 1980 "Colleges Attended by Graduates of Elite Secondary Schools." *The Educational Forum* 44 (May):483-490.
- Herman, Edward S., and Noam Chomsky 1988 *Manufacturing Consent*. New York: Pantheon.
- Hopper, Earl 1977 "A Typology for the Classification of Educational Systems." In Jerome Karabel and A.H. Halsey, eds., *Power and Ideology in Education*. New York: Oxford University Press.
- Kanaya, T. 1994 "Japan: System of Education." Pp. 3078-3086. In T. Husen and N. Postlethwaite, eds., *The International Encyclopedia of Education*. New York: Pergamon.
- Kerchner, Charles Taylor, Julia Koppich, and William Ayers 1993 *A Union of Professionals: Labor Relations and Educational Reform*. New York: Teachers College Press.
- Kirst, Michael, and Christopher Mazzeo 1995 "The Rise, Fall, and Rise of State Assessment in California." *Phi Delta Kappan* (December):319-325.
- Krug, Edward 1964 *The Shaping of the American High School*, vol. 1. New York: Harper and Row.
- 1972 *The Shaping of the American High School*, vol. 2. Madison, Wisc.: University of Wisconsin Press.
- Manno, Bruce V., Chester E. Finn, Jr., Louann Bierlein, and Gregg Vanourek 1998 "Charter Schools: Accomplishments and Dilemmas." *Teachers College Record* 99 (Spring):537-558.
- Massell, Diane 1994 "Setting Standards in Mathematics and Social Studies." *Education and Urban Society* 26 (February):118-140.
- Metz, Mary Haywood 1978 *Classrooms and Corridors*. Berkeley: University of California Press.
- 1986 *Different by Design: Politics, Purpose, and Practice in Three Magnet Schools*. Boston: Routledge and Chapman Hall.
- Meyer, John W., David H. Kamens, and Aaron Benavot 1992 *School Knowledge for the Masses*. London: Falmer.



- Meyer, John W., Francisco Ramirez, and Yasmin N. Soysal 1992 "World Expansion of Mass Education, 1870-1980." *Sociology of Education* 65:128-149.
- Mohrman, Susan Albers, and Priscilla Wohlstetter 1994 *School Based Management*. San Francisco: Jossey-Bass.
- National Commission on Excellence in Education 1983 *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C.: U.S. Government Printing Office.
- National Educational Association 1918 *Cardinal Principles of Secondary Education*. Washington, D.C.: U.S. Department of the Interior, Bureau of Education.
- Oakes, Jeannie, and Martin Lipton 1992 "Detracking Schools: Early Lessons from the Field." *Phi Delta Kappan* (February):448-454.
- Organization for Economic Cooperation and Development 1998 *Education at a Glance*. Paris, France: Author.
- Pfeffer, Jeffrey, and Gerald R. Salancik 1978 *The External Control of Organizations*. New York: Harper and Row.
- Powell, Arthur, Eleanor Farrar, and David Cohen 1985 *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Boston: Houghton Mifflin.
- Ramirez, Francisco O., and John Boli-Bennett 1982 "Global Patterns of Educational Institutionalization." In Philip G. Altbach, Robert F. Arno, and Gail P. Kelly, eds., *Comparative Education*. New York: Macmillan.
- Ramirez, Francisco O., and Phyllis Riddle 1991 "The Expansion of Higher Education." In Philip G. Altbach, ed., *International Higher Education: An Encyclopedia*, vol. 1. New York: Garland.
- Ravitch, Diane 1995 *National Standards in American Education*. Washington, D.C.: Brookings Institution.
- Rohlen, Thomas P. 1983 *Japan's High Schools*. Berkeley: University of California Press.
- Sizer, Theodore 1985 *Horace's Compromise: The Dilemma of the American High School*. Boston: Houghton Mifflin.
- Somini, Sengupta 1999 "Integration Tool Bypasses Racially Isolated Schools in Queens." *New York Times* (February 23):B1, B5.
- Spicer, Michael W., and Edward W. Hill 1990 "Evaluating Parental Choice in Public Education." *American Journal of Education* 98 (February):97-113.
- U.S. National Center for Education Statistics 1997 *Digest of Education Statistics, 1997*. Washington, D.C.: U.S. Government Printing Office.
- Van Horn, Carl E. 1995 *Enhancing the Connection Between Higher Education and the Workplace: A Survey of Employers*. Denver, Colo.: State Higher Education Executive Officers and Education Commission of the States.
- Weiss, Janet A. 1990 "Control in School Organizations: Theoretical Perspectives." In William H. Clune and John F. Witte, eds., *Choice and Control in American Education*, vol. 1. Bristol, Pa.: Falmer.
- Wells, Amy Stuart 1993 *Time to Choose*. New York: Hill and Wang.
- , and Jeannie Oakes 1996 "Potential Pitfalls of Systemic Reform: Early Lessons from Detracking Research." *Sociology of Education* 56(5):135-143.
- , Cynthia Grutzick, Sibyll Carnochan, Julie Slayton, and Ash Vasudeva 1999 "Underlying Policy Assumptions of Charter School Reform: The Multiple Meanings of a Movement." *Teachers College Record* 100:513-535.
- Whitty, Geoff, Sally Power, and David Halpin 1998 *Devolution and Choice in Education: The School, the State, and the Market*. Bristol, Pa.: Open University Press.
- Wirt, Frederick M., and Michael W. Kirst 1992 *Schools in Conflict*. San Francisco: McCutchan.

KEVIN J. DOUGHERTY  
FLOYD M. HAMMACK

## ELITES

See Intellectuals; Social and Political Elites.

## EMOTIONS

Emotions arise in individual experience, frequently with noticeable physiological signs, such as a racing heart, flushed or pallid face, tense gut, cold hands, and so forth, and thus may seem an unsuitable topic for sociological examination. This attitude, however, reckons without the incontrovertible facts that *most human emotions result from real, imagined, recollected, or anticipated outcomes of social interaction* and that interaction is the fundamental stuff of sociological analysis. Thus emotions are empirically linked to the social by virtue of their being a consequence of involvement in interaction. But emotions are also precursors of the social, by virtue of their mobilizing energy and motivation for the accomplishment of major social tasks, not the least of which is social solidarity itself.

Although emotions are thus important features of social life, they have had a varying place in the history of sociology. Early recognized as important by the great founders of the field—Marx, Durkheim, Weber, and Simmel—emotions declined in importance after the 1920s, as behavioral and cognitive approaches came to the fore in the social sciences. (Although these two approaches are thoroughly antagonistic to each other, they have agreed that emotions are irrelevant or unamenable to sociological analysis.) Resurrected after a long hiatus, emotions have once again come into their own and are now recognized both as exerting causal effects as well as serving as important results of social endeavors.

This article covers both early and present-day approaches to what has come to be known as the sociology of emotions and examines the role of emotions as both independent and dependent variable.

#### THE PLACE OF EMOTIONS IN NINETEENTH-CENTURY SOCIOLOGY

In the great effort to topple the entrenched structures of monarchy and clergy in eighteenth-century Europe, a signal weapon was *reason*. By virtue of the application of reason to traditional modes of thought and conduct, philosophers and political theorists concluded that the dominion of kings and priests was falsely premised and that all men were created equal. That reason alone could lead to such a startling conclusion for that day was the result of the conquest by reason in another domain, namely natural science. In astronomy, physics, chemistry and biology, the work of reason had accomplished transformations of thought and understanding of a revolutionary nature. The success of reason here gave social thinkers confidence that reason applied to social life would produce results equally significant. And, indeed, it did in form of the upheavals that culminated in the American and French Revolutions.

What had been only thought was put into practice and the tradition of a thousand years swept away. In the ensuing years, reason was enshrined, a new god replacing the old ones that had been discarded. But as this occurred, a counterrevolutionary ideological process was getting under way. In England, Thomas Burke, at first

partial to the aims of the French Revolution, recoiled in horror at its excesses—reason had run amok, with a degree of passion and emotion that was terrifying. Burke argued that society was not amenable to quick constructions or reconstructions. Rather, a slow accretion of time-tested ways that ultimately refined the social rules and customs was the only reliable method for attaining a safe, stable, and satisfactory social order. Opposing the worship of reason, Burke posed rather the validity of an emotional basis for social solidarity. Longstanding mores garnered emotional support and, by virtue of this, violent social change, with its attendant upheaval, was averted.

Burke's conservative message, which rejected the application of reason to social forms, was to lie fallow for a century, until it was resurrected in a more systematic sociological way by Émile Durkheim. In the interim, Karl Marx, also an exponent of reason in the reorganization of society, had propounded his comprehensive historical theory of the growth and decay of societal organization. Yet in developing his understanding of how society, particularly capitalist society, works, Marx ([1842–1844] 1971, [1867] 1967; Engels [1846] 1947) also produced a paradigm for how to examine emotions sociologically. His approach remains important today, even if the specific political and ideological interests of Marxism are no longer attached to it.

Marx propounded the view that forms of social organization—feudalism, capitalism, socialism—were products of such factors as technology and the division of labor, which he termed the *mode of production*, and certain forms of authority and property rights, which he termed the *relations of production*. A given historical mode of production gave rise, for Marx, to a necessary set of relations of production. In feudal times, the relatively primitive technology required a great deal of hand labor, as exemplified in craft production, where producers owned their tools, their raw materials, and the products of their labor. Property rights and social relations were in line with the division of labor made possible by the technology of the time.

After the Industrial Revolution, with its relatively advanced technology, a new form of division of labor and property rights emerged. The factory

system entailed labor sold by its owners, the workers, as if it were any other commodity. Factory workers owned neither the tools, the raw materials, nor the products of their labor. Thus a new set of property and authority relations came into existence because of the more advanced technology and the division of labor this enabled.

Here Marx looked closely at the horrible conditions of early entrepreneurial capitalism. The principal, although not the only, defect of capitalism was that workers were reduced to a state of what Marx called “immiseration,” a condition fostered by subsistence wages. Poor housing, poor food, poor health, and impoverishment of spirit were the common lot of factory workers. Beyond this was the state of “alienation,” a concept Marx adopted from the work of philosopher Georg Frederick Hegel. Hegel had reflected on the sad fate that befell human effort, which, once crystallized in material objects, exists separately and apart from the individual who exerted it and, in this sense, was alienated from its producer. Marx used this concept to forge a theory of emotional consequences for workers under capitalism.

The primary form of workers’ alienation was the fundamental experience of chagrin, bitterness, and resentment because of the loss, or alienation, of the product of their labor to the capitalist, who did not labor or produce anything. According to Marx, this was the main emotional result of a system in which those who produced things were not the same ones who owned them.

A second form of alienation resulted from the boring and mentally numbing tasks of factory-based production. Instead of the inherent satisfactions of earlier craft forms of work, workers in modern industry, with its extensive division of labor, obtained little pleasure from their largely repetitive tasks.

A third form of alienation was the emotional isolation and competitive envy of others: both owners and fellow workers. This led to the breakdown of solidarity and community, the comfortable sense of belonging to a group in which one was an accepted member. Melvin Seeman (1959) translated Marx’s forms of alienation into five elements: (1) powerlessness, or a feeling that one had no control over one’s fate; (2) meaninglessness,

or a sense of confusion about the value and significance of one’s efforts; (3) self-estrangement, or a feeling of distance between what one felt oneself to be and what one was required to do at work; (4) isolation, or the longing for a sense of connection with others; and (5) normlessness, or the feeling that one’s efforts lacked aim or goal.

Even if Marx’s critique of capitalism is ignored, it is important to recognize his innovation in sociological analysis, what today we would call social psychology; that is, how social organization affects individual variables. In Marx’s case view, how the social patterns by which labor is organized and its benefits distributed affects emotions. This mode of analysis, designated by the term *social structure*, is one of the two main types of approach to the analysis of emotions in current sociology. The other, designated as *culture*, was fostered by the work of Max Weber.

Although he recognized the power and utility of Marx’s social structural analysis of capitalism and its emotions, Weber cast a different light on the process and examined different emotions. In his most famous work, *The Protestant Ethic and the Spirit of Capitalism*, Weber ([1904–1905] 1958) engaged Marx in an intellectual battle over who better understood the historical processes that led to capitalism. According to Marx, the economic infrastructure of society—the mode and relations of production—gave rise to the superstructures of society; namely, the way in which other institutions, such as the family, religion, politics, art, and so on, were organized. Further, Marx contended that the infrastructure even determined ideas. And here Weber demurred. Instead, Weber sought to show that capitalism was the product of a distinct set of religious ideas, along with the emotions that those ideas fostered.

For Weber, the crucial ideational matrix for capitalism was the Calvinist doctrine of predestination. This is the view that salvation in the religious sense is determined for everyone prior to birth. Some, the elect, are destined to be saved and the remainder, doomed to eternal damnation. According to this doctrine, there is nothing one can do to change one’s fate—no action, no amount of devotion can change one’s predestined fate. For many today, this seems like an odd doctrine, and many wonder how it could have held sway. But in

Geneva and other places where Calvinism reigned for a while, the belief was of the utmost importance, since it spelled out the possible eternal fate of one's soul.

Weber conjectured that such a doctrine must have been accompanied by some powerful emotions, mainly a terrible anxiety. At this point, Weber argues, a reaction formation of necessary optimism intervened to forestall despair. Calvinists came to believe that somehow, contrary to doctrine, they might change a dire fate by the most faithful and dedicated service to God's commandments. One of the most important ways in which to do this was in one's God-given *Beruf*, or vocation; that is, one's life's work, whether it be as a farmer, artisan, merchant, and so forth. Pursuing one's daily work in so dedicated a fashion would perhaps avert the severity of a predestined fate.

But such religiously and emotionally driven attention to one's occupation brought an unsought consequence: It led to worldly success. This in itself, despite doctrinal objections, seemed like a divine signal. Those who succeeded in worldly pursuits began to see this as a hopeful sign that they were among the elect. And, if this were so, then even greater worldly success would ensure that prognosis. Given that Calvinism also inveighed against material display, it led the successful to plow their profits back into their work. In other words, said Weber, it led to the routine practice of the capitalist method of doing business, thus giving rise to that historic form of entrepreneurship. In Weber's view, anxiety and a form of coping with it were central to the emergence of capitalism.

In the terms of the present-day sociology of emotions, Marx originated an approach based on *social structure*, while Weber fostered an approach based on *culture*. The effects of this distinction will be elaborated below.

Émile Durkheim also entered the debate against Marx, but on a different front from Weber. Marx had viewed religion as an opiate, offering hope to and dulling the pain felt by workers who suffered from the extremes of dislocation and poverty, especially during the early stages of capitalism. Marx supposed that when socialism conquered, it would obviate the need for religious belief and practice. Durkheim ([1915] 1965) viewed this as a

form of ultrarationalism and sought to uncover the universal wellsprings of religion, which, he believed, would flourish regardless of what form society took. In order to investigate religious phenomena at their simplest—and hence most likely to prove universal—Durkheim examined the religion of Australian Aborigines, a group that appeared to be about as primitive as one could find according to nineteenth-century understanding.

First Durkheim proposed a distinction between religious behavior and everyday behavior, the sacred and the profane. He saw the two as segregated by time and place as well as by behavior. Among the Aborigines, the sacred emerged during clan gatherings, with the worship of totems—animals that stood for the clan and with whom clan members identified. Worship was a matter of ritual practice involving highly emotional activities—rhythmic chanting, frenetic dancing, sexual intercourse—with individuals giving over emotional control to the impulses of the moment. Durkheim reasoned that in these moments of heightened emotional arousal, individuals felt themselves to be in the grip of forces greater than themselves, providing a sense of the presence of something greater than any single individual—a spirit, a god. But Durkheim argued that the only supra-individual entity present was the group itself; that is, the co-presence of the members of the group, giving themselves over to strong emotions, provided the sense of a superior force. In worshipping that force, group members were, unwittingly, worshipping the group itself.

Durkheim generalized this point, arguing that ritual conduct, with its element of emotional arousal, leads to a sense of solidarity among those who engage in the ritual. Thus, religious ritual, far from representing the dead hand of the past, as Marx would have it, enabled group members to cleave to each other and make group continuity possible. This is true, said Durkheim, even when the shared emotion is sadness. Indeed, in examining funerary rites, he showed that coming together on the death of a group member stirred up emotions that transcended the sadness itself. Remarkably, a feeling of strength and renewed vitality emerged from the co-presence of other mourners. Emotions, thus, are not merely individual phenomena; they are crucial to the existence of human groups.

Society itself is possible only because its members periodically share emotions, whether elation over great victory or sadness over great loss.

### EMOTIONS IN MODERN PERSPECTIVE

Although present-day sociologists of emotions treat them differently in many ways than did the founders of the field, they nonetheless hue close to the main insight of the founders; namely, that emotions are socially constituted. This means several things.

First, emotions emerge from episodes of interaction in which valued outcomes are at stake. Second, emotions have a socially normative component. Thus, when one's emotions deviate from the normative emotional prescription, there is some constraint to adjust one's emotions to what is normatively specified. Third, over time, as social conditions change, emotional requirements change as well. New conditions mean a somewhat new recipe for emotional life.

In the present-day sociology of emotions, two major sociological traditions converge on the social constitution of emotions, but they do so in different ways. The *social structural* approach examines emotions as direct products of social interaction and its outcomes. For example, insult leads to anger; deflated ego leads to shame; threat leads to fear; and so on. In each case, the end result of the social interaction is experienced as an emotion.

By contrast, the *cultural* approach looks mainly at the regulation of emotional expression by social rules. For example, we are enjoined to be happy at weddings, sad at funerals, angry at injustice, and so on. In each case, the individual who is emotionally out of line with what is expected in the situation confronts the possibility of emotional deviance.

The social structural and the cultural approaches form the main body of work in the modern sociology of emotions, and we discuss each in turn.

**Social Structural Approaches.** The earliest modern sociologist of emotion was Erving Goffman (1959, 1967, 1981), a transitional figure between the works of the founders, especially Durkheim, and later approaches. He focused powerfully on

ordinary conversation as one of the main settings of social interaction. In the social interaction we call conversation, he found high emotional drama. Although conversation most often proceeds without conscious reflection about it by the participants, a conversation has only to break down to receive considerable and sometimes perplexed attention. Goffman saw that conversation actually comprises a miniritual, much in the manner Durkheim described, and that a successful conversation is like a successful ritual: Everyone plays a necessary prescribed role, even if the exact content of the role is not prescribed. In a successful conversation, participants experience a degree of self-realization that produces the kind of satisfaction that ensues from a whole-hearted and authentic participation in a religious ritual. Indeed, there is a similar sacred aspect to conversation, with the most important sacred object on display being the participant's self. In conversation, that self is accorded a due reverence by other participants or is demeaned by them.

Goffman proposed that in a conversation every actor offers a "line" about him- or herself, one which the other participants are expected to take at face value. For example, one has just been promoted to a higher position, or one has sacrificed dearly for a loved one, or one has learned how to play a Bach partita, and so on. If one carries off the line in the conversation such that others confirm one's projected persona, one feels confident and self-assured. But if one stumbles conversationally and the projected self becomes noncredible, then embarrassment ensues. Unless the actor has what Goffman calls "poise," which is the ability to mask one's embarrassment, the conversation falters.

Since conversation by its nature is interactional, other participants must cooperate in sustaining the viability of the conversational interaction. Except in unusual circumstances, Goffman saw in ordinary talk a precarious social order that is supported by the emotions of the participants, requiring their cooperation and emotional sensibility to succeed. The social order of ordinary talk is an instance of social order in general. Although Goffman did not pursue the larger questions of social order in society, his point about the emotional underpinnings of the micro order was extended to the social macro order, especially in the

formal precedence orders of social power and status. Randall Collins and Arlie Russell Hochschild examined this question in the domain of stratification.

Stratification is the division of the social order into a hierarchy of power, status, and benefits. Some individuals have more and some have less of these desirable attributes and commodities. Since so many life chances are tied up in the stratification order, it is understandable that emotions are focal here, too.

According to Collins (1975, 1981, 1990), social systems, whether large or small, are arenas of conflict in which power, status, and benefits are sought by the participants. In modern societies, the main field of conflict is in organizations devoted to work. Here there are order-givers and order-takers, whose daily interactions are fraught with emotional consequences. Relative to order-takers, order-givers are secure, confident, and charged with emotional energy. This is because they are at the head of organizational units that support their initiatives and legitimate their orders. Order-takers, on the other hand, are relatively resentful, indifferent, and alienated. They are under constraint from the organizational coalitions headed by the order-givers; they will be punished if they do not obey.

Following Durkheim's model, Collins views the daily interactions of order-givers and order-takers as a ritual in which the former display their social value by providing information, evaluating the performances of others, and mobilizing the morale and energies of group members. The order-takers are the objects of the ritual, which is intended to arouse them to the same pitch of organizational commitment experienced by the order-givers. But since the order-givers, as the main beneficiaries of the system of organizational stratification, are most aroused in the ritual, they are most likely to cleave loyally to the symbols of the organizations they represent. As beneficiaries of far less, order-takers are also aroused—but with different emotions that are less conducive to good feelings, high energy, and loyal commitment.

Hochschild (1979, 1983) examined some of the emotional accompaniments of labor in service organizations. In an exemplary study of airline flight attendants, she found these workers to be

often conflicted in their emotions. Although they were trained specifically to do what Hochschild called emotional labor—namely, to cater to passengers' expectations in situations of high time and space constraints, as well as passengers' sometimes impolite or obstreperous behavior—flight attendants' own emotions had to be suppressed. Insulted by a passenger, they were required to suppress anger; disgusted by a passenger's conduct, they were required to overlook their own emotion; and so on—all in the service of not alienating customers. According to Hochschild, this left many of the flight attendants emotionally numb from the overpractice of emotion suppression. Hochschild generalized these findings to the service economy at large, where most of the occupations that require emotional labor are staffed disproportionately by women.

Taking a broader perspective on social structure, Theodore Kemper (1978, 1987, 1989, 1990) proposed that most emotions could be examined as resultants of interaction outcomes in two dimensions; namely, power and status. Power is a relationship in which one actor compels another actor to do something the latter does not want to do. This entails the employment or threat of employment of a variety of noxious stimuli—from physical and verbal violence, through withholding of deserved rewards, to subtle manipulations such as lying and deception. Power is exercised pandemically, whether in informal social interactions or in organizations that carry out the broad work of society.

Status (or status-accord), on the other hand, is a social relationship in which actors voluntarily provide rewards and gratification to other actors. Coercion plays no part here. Status-giving is the basis of true social solidarity, and individuals who voluntarily share membership with each other in a social group are normally motivated to provide each other with the benefits of status, the ultimate degree of which is love.

Kemper proposed that outcomes of interaction in power and status terms give rise directly to emotions. For a sense of how this works, it is important to see that a relatively simple model of social relational outcomes suffices for the analysis. Every relational exchange takes place along the dimensions of power and status. For each actor,

the outcomes can be: an increase in power and/or status, a decline, or no change. This produces 12 possible outcomes (2 actors  $\times$  2 dimensions  $\times$  3 results), but only 4 of these will actually occur, namely a power and status outcome for each actor A. Emotions will flow from these.

In the power dimension, the available evidence suggests that actor A loses power or actor B gains it, the emotional outcome for actor A is some degree of fear or anxiety. If actor A gains power and/or actor B loses it, the emotional outcome for actor is likely to be a sense of security. In the status domain, the outcomes are a bit more complex. If actor A gains status, he is likely to feel contented, satisfied, and happy. If actor A loses status, the emotional outcome depends on the felt sense of agency: Who was ultimately responsible for the status loss? If actor B is held culpable, the emotional result is anger. If the self is held culpable, the emotional result is shame or, more seriously, depression, if actor A feels that the situation is irremediable.

When the situation is one in which actor A gives status to actor B, we may expect satisfaction and contentment on the part of actor A. When actor A withholds or withdraws status from actor B, several emotional outcomes are possible. If actor B is deemed responsible (“deserving” this treatment), the result is likely to be satisfaction of a self-righteous kind on the part of actor A. If actor B was not responsible (not “deserving” the treatment), then two emotions are possible, either singly or jointly. If actor A deems that he or she did not live up to his or her character standing—that is, the amount of status with which he or she is generally credited by others as deserving—he or she will feel shame. On the other hand, if actor A concentrates on the harm or hurt he or she did to actor B, guilt is likely to ensue. In the instance of shame, actor A recognizes that he or she has acted in such a way as not to deserve the status he or she has been accorded. In the case of guilt, actor A recognizes that he or she has used an excess of power against actor B.

Some aspects of emotion also follow from anticipations that are either confirmed or disconfirmed by subsequent interaction. Thus, the actor who anticipated losing power but did not should feel especially secure. The actor who did

not anticipate receiving status but received it is likely to feel especially pleased. In general, disconfirmations of expectations tend to exert a multiplier effect on the emotions that would ordinarily be felt from the power–status outcome in the situation.

Kemper’s approach to emotions through power and status analysis also allows a social relational perspective on love and on its frequently confused near-relation, liking. Kemper defines a love relationship as one in which at least one actor gives or is prepared to give extreme amounts of status (for terminological purposes, status is equated with affection) to another actor. The definition includes nothing about power, but since power is a feature of all relationships, including it allows for seven different ideal-typical love relationships: (1) *Adulation*, in which one actor gives or is prepared to give extreme affection to another who may not even know of the other’s existence. Neither actor has or uses power. (2) *Ideal love*, in which each actor gives or is prepared to give extreme amounts of affection to the other. Neither actor has or uses power. (3) *Transference* or *mentor love*, in which each actor gives or is prepared to give extreme amounts of affection to the other, but one of the actors also has a great deal of power over the other. This pattern is prevalent in ideal teacher–student, therapist–client, or mentor–mentee relationships, in which the former in each case is the one with power over the latter. (4) *Romantic love*, in which each actor gives or is prepared to give extreme amounts of love to the other, but each actor also has a great deal of power over the other. This is the kind of relationship in which each can suffer intensely from the real or imagined withdrawal of affection by the other. (5) *Unfaithful love*, in which only one actor gives or is prepared to give extreme amounts of affection to the other, but both actors have a great amount of power. Infidelity is a case of this kind of relationship; the betrayer has withdrawn affection from the betrayed, although the betrayed still has a great deal of power, which is why most infidelities are kept secret. (6) *Infatuation* or *unrequited love*, in which one actor gives or is prepared to give extreme amounts of affection to the other but has no power, while the other actor gives no affection but has a great deal of power. (7) *Parent–Infant Love*, in which the parent gives extreme amounts of affection to the infant and the

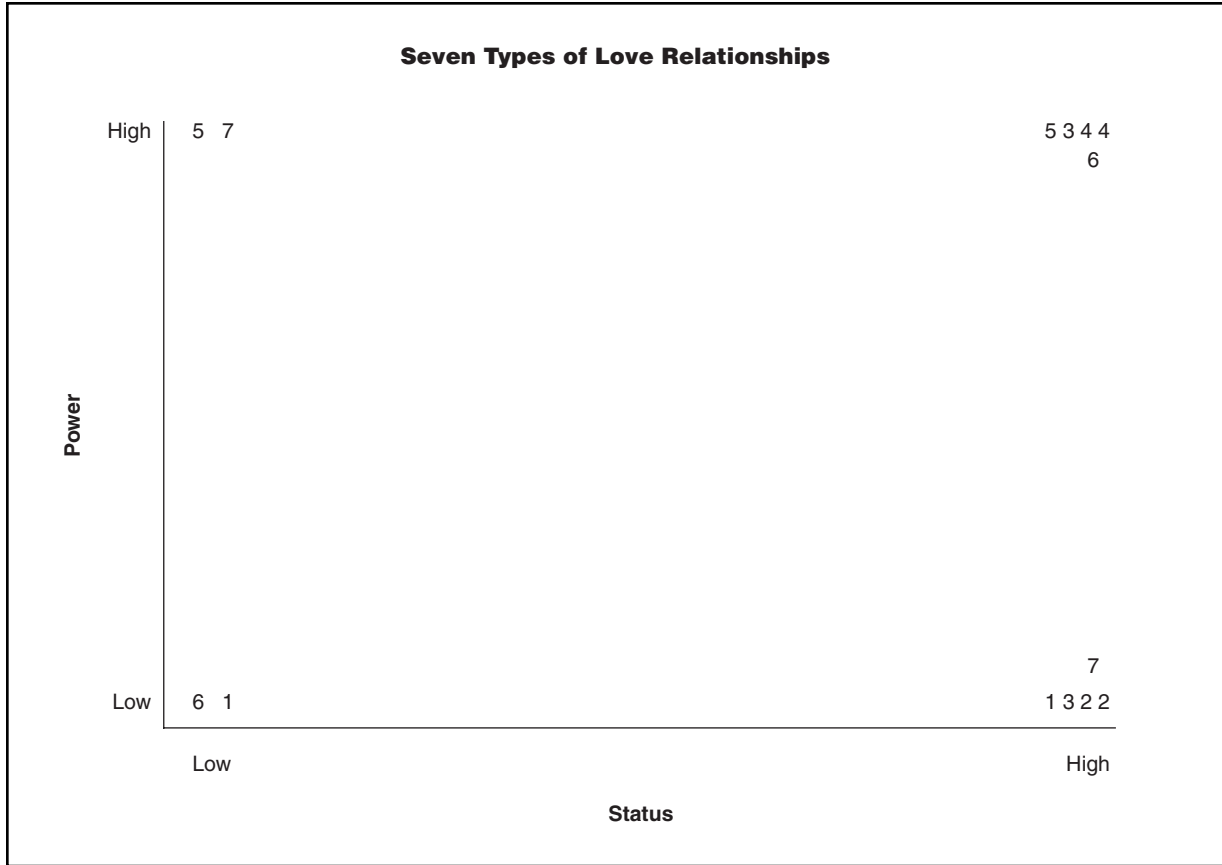


Figure 1

NOTE: 1-1 Adulation

NOTE: 2-2 Ideal Love

NOTE: 3-3 Transference or Mentor Love

NOTE: 4-4 Romantic Love

NOTE: 5-5 Unfaithful Love

NOTE: 6-6 Inatuation or Unrequited Love

NOTE: 7-7 Parent-Infant Love

infant gives nothing in return, while the parent has a great deal of power and the infant has none. These seven relationships can also be represented visually in a two-dimensional space according to their power and status (affection) locations (See figure 1).

Kemper (1989; and Reid 1997) distinguishes love from liking as follows: one feels love for another when the other's qualities match one's standards to an extremely high degree. It does not matter what the standards are; they may be trivial (e.g., he dresses well; she dances well) or they may be profound (e.g., he is a noble man who would

risk his life to preserve others; she is compassionate woman whose concern is with the well-being of others). In either case, if one has standards that are met by these qualities in the other, then it is likely that feelings of love will be induced. Ordinarily, more than one standard must be met by the qualities of the other in order for love to bloom. No small standard in Western culture is appearance, which often dominates all others, at least among the young. Love thus has to do with the qualities of the other and how they match our standards.

Liking, on the other hand, according to Kemper, is the pleasant feeling that arises when the other



gives us status or affection. Liking requires that the other act well toward us; love has no such requirement, as shown by such types (above) as adulation, unfaithful love, infatuation or unrequited love, and parent-child love. In each of these, one party loves another who gives no affection in return. In love this is possible; in liking, it is not.

Kemper's use of power and status is matched by Heise (1979) and his colleagues (Smith-Lovin and Heise 1988; MacKinnon 1994) in another structural approach to emotions known as affect control theory (ACT). Heise's model of interaction and emotions stems from a linguistic paradigm, most notably the semantic differential (SD), which purports to get at fundamental categories of meaning in the use of language. Supported by many cross-cultural studies, the proponents of the SD have found three fundamental dimensions: potency (power), evaluation (status), and activity. Kemper does not include activity in his approach to emotions, although it has a place in a more general social organizational framework as standing for technical activity in the division of labor.

Working with the potency, evaluation, and activity dimensions, Heise and his colleagues have developed a mathematically sophisticated set of formulas to predict a variety of outcomes, including emotions. First, common language terms—nouns, verbs, adjectives—are rated by samples of respondents for their potency, evaluation, and activity standing. For example, the term *father* may stand relatively high on each of these; on the other hand, the term *criminal* may stand high on potency and activity but low on evaluation. Heise has compiled a large dictionary of such terms whose potency, evaluation, and activity values can be entered into complex regression equations to predict emotions.

The basic notion in ACT is that individuals behave so as to maintain their fundamental identity. If something has occurred—whether action by oneself or another—that questions the validity of that identity, an emotion results. For example, a parent may usually be loving to his or her child, but if the parent acts out of character by behaving cruelly or indifferently to the child, there is a felt need to reequilibrate the relationship and reclaim the identity of loving parent. The emotion, whether it is shame or guilt, provides some of the

motivational energy to repair the relationship and maintain the fundamental identity.

In some examples of how this works, Heise and his colleagues found that when a father serves a son, the father's emotions were predicted to be pleased, contented, and relieved, while the son's emotions (the emotions of the recipient of interaction can also be predicted) were predicted to be amused, light-hearted, and euphoric. If a judge sentences a gangster, the judge's emotions were predicted to be contented, relieved, and proud, while the gangster's emotions were predicted to be uneasy and awestruck. Several emotions are usually generated by the ACT equations because they have approximately equal value in potency, evaluation, and activity terms for the given situation.

Heise has examined complex instances in which an individual is observed to perform an act and reveal a number of emotions. How others will judge the actor is predicted to be based on the emotions the focal actor reveals. For example, when a man kisses a woman, if he is cheerful, he is identified as a gentleman, pal, or mate, identities that gain significantly in evaluation, with some loss in the potency dimension. On the other hand, if he kisses a woman and manifests disgust, he is identified at a much lower evaluation level, but with his potency remaining unchanged. If he displays nervousness, he is identified as having lower standing in both evaluation and potency.

Heise's method also allows for a distinction between emotions and moods. For example, if a father ignores a son, the father is predicted to feel unhappy. ACT now predicts that the father will act to reinstate his identity as a father who does not ignore his son. In contrast to the father who became unhappy over his act, there is the unhappy father, a fundamental identity combining a social position with an emotion. Heise identifies this as a mood. According to ACT, moods give rise to consistent behavior; for example, the unhappy father might neglect or attack the son.

In all, the structural theorists of emotion are concerned with predicting emotions from the social locations and relationships of actor. This assumes that there is a natural (or pancultural) universal reaction to certain kinds of social relational outcomes. For example, insult evokes anger. One

of the most important differences between the social structural and cultural approaches (to be discussed next) hinges on whether this assumption is correct.

**Cultural Approaches to Emotions.** In contrast with the social structural view of emotions as direct results of social relations between actors, the cultural approach inserts an intervening stage; namely, the normative definition of situations and the specification of what emotions are appropriate in them. Hochschild proposes that these “feeling rules” define and regulate the expression of emotion. Examining emotions from this perspective leads to a strong emphasis on the study of how emotions are managed so that they conform with the normative requirements of given situations. Culturally oriented sociologists of emotion are also concerned with how emotions contribute to social order. Guilt and shame are significant emotions for this purpose.

The cultural approach to emotions is partial to cognitive and idealist models, since these concern themselves with mental processes that come to determine emotions. The fundamental source of the cultural approach is *symbolic interactionism*. In this school, whose prime practitioner and exemplar was George Herbert Mead (1934), the fundamental notion is that virtually nothing, not even mind or self, precedes social interaction. Thus social interaction actually constitutes or constructs these fundamental categories.

Mead proposed that after some cooperative interactions with another person, we have the capacity to call up in ourself the probable reaction of the other to any proactive behavior of our own. This is based on the recollection of the pattern of interactions with the other when one or another behavior on our part elicited one or another response on the part of the other. The ability to recapitulate all this in our head prior to any actual behavior is what Mead termed *mind* and the process, *thinking*.

Having derived mind from prior social interaction, Mead went further and located the origin of the self in the same kinds of interactive encounters with others. In Mead’s terms, we become capable of putting ourself in the place of the other and looking at ourself as if we were an object. This

mental operation provides us with a sense of self—our identity—as derived from the perspective of another. The actual self of the individual is some composite of the many selves that are available when one takes the perspective of the many others with whom one interacts.

The most widely known approach to emotions from the symbolic interactionist perspective is offered by Hochschild (1979, 1983). She posits that emotions arise in a somewhat natural way in situations or frames. But the emotion is then subjected to examination as more or less appropriate from the perspective of the normative borders of the situation. Expressed as “feeling rules,” the norms specify the required emotions in given frames: happy at birthday parties, sad at funerals, and so on. Because most people react emotionally to situations in more or less the ways that the rules require, they are rarely conscious of the rules. But should there be a discrepancy between the emotion and the rule, there is a sense of discomfort and a felt pressure to adapt. Hochschild offers a number of likely strategies that are used to manage one’s emotions in such emotionally deviant situations. Principally, one may engage in either surface or deep acting. In the former, one puts on the manifest signs of the emotion even if one does not authentically feel it; for example, smiling at the host of a party, despite the fact that one despises him. In deep acting, the individual actually tries to evoke the prescribed emotion.

For Hochschild, emotion essentially results from a discrepancy between what we perceive and what we expected. The inchoate feeling is labeled by cultural fiat as anger, fear, shame, and so on, and this provides a reservoir of cultural associations with the significance, meaning, implications, and so on of having such an emotion. These may ramify into modifying or seeking to change the emotion. Thus culture, the aggregate of normative understandings derived from others, intervenes early in the emotion process, leading to the judgment that emotion is a social construction. For example, shame is constructed from five perceptions: motive (I want to do right); possession (I have done wrong); value (I disapprove); agency (I am the cause of the event); and self-agent relations (the audience for my act is better than I am).

Hochschild’s (1983) widely cited study of emotion management among airline flight attendants

has led to a significant body of research that has focused on the *emotional* effects of managing emotions. These studies are reviewed by Morris and Feldman (1996) and Gibson (1997).

Peggy Thoits (1990) has taken Hochschild's ideas on emotion management to another level. In her view, emotions are comprised of four elements: situational cues, physiological changes, expressive gestures, and an emotion label. These are so connected in memory and behavior pattern that the elicitation of one evokes the others. Thoits proposes that when emotions and feeling rules are discrepant, the actor can manage this through either cognitive or behavioral manipulation of the four elements of emotion. For example, one can withdraw when a deviant emotion is felt (behavior-situational), or one can exercise or take drugs to change the physiological base of the emotion (behavior-physiological), or one can redefine the situation so that its implication for emotion changes (cognitive-situational), and so on.

Thoits also expands Hochschild's notion of emotional deviance through proposing four situations that might dispose toward it: multiple role occupancy, subcultural marginality, role transition, and rigid rules governing ongoing roles or ceremonial occasions. Thoits also proposes that a deviant emotion not only violates the feeling rules but also includes emotions that are too prolonged, too intense, or directed at the wrong target. Yet even deviant emotions may become legitimate if they are widely shared, thus leading to a change in social norms. An example is the change in national sentiment brought about by the protests of the antiwar movement during the Vietnam period.

In another fundamentally symbolic interactionist approach to emotions, Clark (1997) examines sympathy, an emotion treated importantly by Adam Smith ([1759] 1853) in his *Theory of Moral Sentiments*. Clark sees sympathy as a pervasive emotion, making society possible, for without it there would be no "social glue." Indeed, since sympathy is so important, "sympathy entrepreneurs," emerge to facilitate the evocation and display of this emotion. These are voluntary organizations designed to mobilize sympathy for specific victims (for example, Mothers Against Drunk Driving [MADD]) or commercial organizations, such as greeting card companies, that facilitate expressions of sympathy by large publics. Clark also sees what she refers to

as "sympathy margins" as regular features of social relationships. These are earned credits, so to speak, that enable individuals to call on the understanding, sympathy, and forgiveness of others when they are caught out, or caught short, or have otherwise become hapless victims.

Steven Gordon (1981) is another exponent of the social constructionist view of emotions. Although he acknowledges that emotions per se are elemental and biological, he contends that shortly after childhood they are culturally transformed into what he calls "sentiments." For example, the elemental emotion of anger is converted into such sentiments as resentment, righteous indignation, moral outrage, and so on. This is presumed to be the fate of all emotions. Since sentiments are socially formed, they can be invented or even abandoned.

Gordon stands within the tradition of Norbert Elias (1978a, 1978b), who offered an insightful historical perspective on the emergence of the emotion of shame as an important feature of social relations between different social classes. Elias examined especially the relations between the rising bourgeoisie of the fifteenth and sixteenth centuries and the aristocrats who disdained them. In many instances the nobility, seeking to insulate itself from the incursions of lower-status merchants and traders, refined their practices in the execution of common daily chores involving dining, self-cleansing, toilet functions, and so on. Those who could not match the newly defined limits of gentility were exposed to ridicule and shaming. For a long historical period, the aristocracy were the arbiters of manners, that is, what passed for acceptable social conduct. Those who were ill trained in the ultrarefinements thus properly felt shame.

Although symbolic interactionism has been used to refute the idea of fixity in the domain of emotions, Susan Shott (1979) and Thomas Scheff (1979, 1988, 1994, 1997) have also employed it to show how emotions underlie social order and stability. The fact that pattern and predictability exist in much of social life, rather than chaos and randomness, poses one of the longstanding problems in sociological analysis. It is generally acknowledged that the reverse side of social order is social control, indicating that society manages somehow to instill in individuals a propensity to comply with required social forms and that deviance from

these forms, though greater in some periods than in others, is actually quite limited. How is this accomplished?

There are two main answers to the question, and both of them turn on emotion. First, social order may be imposed by dominant and powerful groups. Their tactic is to evoke fear for nonconforming behavior. Although it has been argued by some, such as Talcott Parsons, that such regimes cannot be stable in the long run, in the short run they can have remarkable sticking power. Decades or even centuries may elapse before a feared government is overthrown.

The second ground for social order entails acceptance of the existing pattern of things. And here sociologists have split on what is accepted. On the one hand, social order can flow from belief in the validity of the social norms. One pays one's taxes, serves in the armed forces, does not steal even when there is an opportunity to do so undetected, and so on because the rules are deemed valid and it feels morally right to abide by them. A second view, which has come to prominence, is that underlying social order is an emotional order. Without an emotional basis, social order would not be possible.

Attacking the question from a symbolic interactionist perspective, Shott proposes a set of "role-taking emotions" that are central to social control: guilt, shame, embarrassment, pride, and vanity. Each of these involves the central symbolic interactionist mechanism of putting oneself in the place, or taking the role, of the other person and thereby evoking his or her perspective. The result of such role-taking can be an emotion directed toward the self, because it evokes in the self the judgment that others are making about the self. Guilt involves the self- and (presumed) other-judgment of "moral inadequacy." Shame entails the self- and (presumed) other-rejection of an "idealized self-image." Embarrassment arises from the realization that others view one's self-presentation as "inept." Pride comes from placing oneself in the position of others and regarding oneself with approbation; vanity is a reduced form of this, in that one is not sure of other's approval.

These emotions (with perhaps the exception of vanity) operate homeostatically. Individuals are moved to reduce the incidence of such unpleasant emotions as guilt, shame, and embarrassment and

increase the incidence of pleasant emotions such as pride. How is this done? Obviously, one avoids the unpleasant emotions by avoiding conduct that would earn the disapprobation of others. One gains pleasant emotions by engaging in conduct of which others approve. In general, the emotions fit one into the moral requirements of others, who are themselves governed by the same set of pre- and proscriptions for social conduct.

Shott suggests that the role-taking emotions are an inexpensive way for society to obtain social order, since they make each person his or her own guardian in ensuring that the emotional tone of one's life remains, on balance, more pleasant than painful. Where such self-control fails, the prospects of social order are not entirely dim, for the emotions that ensue—guilt, shame, or embarrassment—motivate reparative action to reequilibrate the social order as well as others—therefore, one's own—opinion of oneself. Thus, guilt and shame have been shown to increase compensatory altruism toward others, and embarrassment has been shown to evoke compensatory supererogation or attainment as a way of reequilibrating the judgments of others about the self to return to a positive balance.

Finally, Shott proposes a role-taking emotion that is not reflexive, in that it does not pertain to a judgment of the self. This is empathy, which allows one to feel what the other person in the situation is feeling, or what one would likely feel if one were in the place of the other person. Empathy makes any emotion vicariously accessible. Where the emotion reveals the other to be in a socially vulnerable place, one has the embodied sense of the need that other has for social rescue, and the likelihood of engaging in that rescue is enhanced. Thus empathy allows for the evocation of solidarity with others and the preservation of social order through protective behaviors that take up the slack when others are unable to act suitably in their own behalf.

Scheff is also concerned with social order, but he focuses on shame and pride as the ne plus ultra emotions in this regard. He takes a cue from Charles Horton Cooley's ([1902] 1922) famous looking-glass metaphor: "Each to each a looking glass, reflects the other that doth pass" (p. 184). Cooley asserted that pride and shame were the emotional engines for getting individuals to conform to the requirements of their fellow members

in society. Scheff proposes that individuals are continuously in a state of either pride or shame. But then the question arises as to why, if these emotions are so important, there is so little evidence of them.

Scheff proposes that shame is a recursive emotion. That is, once present it has a tendency to evoke more shame, or even anger, over the fact that one is ashamed. This can lead to a spiral of emotion about emotion about emotion . . . that leads to both an inability to escape the emotion and a tendency to hide it from others—a frequent response when coping with shame. This hiddenness, proposed as a defining feature of shame, ties in with the work of Helen Block Lewis (1971), whose intensive analysis of psychotherapy protocols revealed two types of “unacknowledged” shame: *overt, undifferentiated shame* and *bypassed shame*. The former is manifested by painful feelings and self-derogation (“I am stupid, foolish, feckless, incompetent,” and the like). Often this is accompanied by stammering, unnecessary word repetition, averted gaze, and declining audibility of speech. Both the verbal, paralinguistic and proxemic forms are means of hiding the self from the evaluating gaze of others.

Bypassed shame, on the other hand, leads to covert such symptoms as obsessive focusing on the episode that evoked the inadequate response, as if the replay could retrieve the lost status. Thought and speech are hyperactive, actually preventing one from participating with others in the natural rhythm of conversational flow. Both types of shame share the common characteristic of low visibility, thus demonstrating the power of shame as the emotional foundation of social control. Only those with sufficient self-esteem can acknowledge their shame and thus discharge it. But self-esteem itself, as proposed by Mead, derives from the good opinion of others, which itself arises when one conforms to their normative requirements; that is, when there has been social control.

This formulation, implying catharsis in the discharge of shame, ties into Scheff’s work on the problem of undischarged emotions. Scheff proposes that catharsis of these residual emotions can only occur in properly “distanced” settings. Here Scheff adopts the concept of “aesthetic distance,” employed by Bullough in his analysis of drama. According to Bullough (1912), an audience can

experience a dramatic presentation in various ways according to its emotional distance from what appears on stage. Too little distance involves the audience so deeply that it forgets it is merely watching a play and wants to mount the stage in defense of the hero. Too much distance leaves the audience uninvolved, indifferent to whatever murder or mayhem may be happening on stage. Optimum, or aesthetic, distance, like the last of Goldilocks’s porridge bowls and beds, is “just right,” providing a comfortable level of emotional arousal that leads to the “purgation of pity and terror,” which according to Aristotle was the aim of drama.

In a similar vein, Scheff proposes that troublesome residual emotions may also be purged in social settings where there is optimum distance. To do so, he suggests, requires that the expressive emotional content be retrieved (for example, crying, trembling, sweating), but in a context in which the individual can be both participant and observer of his or her own emotional display. When these conditions for emotional aesthetic distance are met, catharsis occurs. This is signaled by an anomalous emotional outcome: even though the residual emotion may be unpleasant, discharging it is not unpleasant, and there is a succeeding state of clarity of thought, relaxation, renewed energy or exhilaration. Although the catharsis paradigm is somewhat different from the social control paradigm, the two are joined in that unacknowledged shame, which is not discharged, often leads to spirals of emotion (e.g., anger over shame over fear) that incapacitate individuals in their social interactions, leading sometimes to violent outbursts that break through all the bonds of social control.

Robert Thamm (1992) builds a theory of emotions on the foundations of Talcott Parsons’s and Edward Shils’s scheme for a general theory of action. In their formulation, social actors are linked in reciprocal forms of action and response through expectations and sanctions. In social settings, individuals have expectations of each other, and in light of those expectations they reward or fail to reward each other’s behavior. From each actor’s perspective, according to Thamm, this leads to four questions: (1) Is the self meeting expectations? (2) Is the self receiving rewards? (3) Is the other meeting expectations? (4) Is the other receiving rewards? These constitute a social matrix for

the production of emotions. As the answers to these questions vary from yes (+) to no (-) to don't know (0), different emotions result. A given state of the system of self and other's expectations and sanctions can be coded by a pertinent series of pluses, minuses, and zeros. For example, if the answer to all four questions is yes, the coding is [++++]; if the answer to all four is no, the coding is [----]; if the answer to the first two is yes and the last two is no, the coding is [++--].

Based on the permutations of the many possible states of the expectations-sanctions system, Thamm hypothesizes a variety of emotional resultants. For example, when the self meets expectations [+000], the self feels pleased with itself. When the self does not meet expectation [-000], the self feels disappointed with itself. When the other does not meet expectations [00-0], the self feels disappointed in the other. When the self meets expectations but is not rewarded [+00], the self feels powerless. When the other does not meet expectations, so that the self is not rewarded [0--0], the self feels anger at the other. Many additional hypotheses follow from the variations along the spectrum of expectations-sanctions possibilities.

The reliance of the cultural approach to emotions on mental structions susceptible to socialization and variable according to historical conditions of change, directly conflicts with at least some elements of the social structural position. The latter places more emphasis on universal situational determinants of emotion and on some biological mechanisms that articulate with the situation-emotion nexus. Kemper (1987) has attempted to reconcile some of the opposing views through a syncretic analysis, focusing on the issue of primary and secondary emotions.

Kemper proposes that a large body of cross-cultural, phylogenetic, autonomic, social relational, and classificatory evidence leads to a model of four primary emotions: fear, anger, sadness, and joy (or nominal variants of these). Since there are additional emotions, the question is: What is their source? Kemper proposes that emotions beyond the primary ones may arise from a specific pattern of socialization in which a social definition and label are applied to a situation in which one of the primary emotions is being felt. For example, pride may be derived from socialization to the idea of self-regard for accomplishment in a context of joy.

Shame may result from socialization to the idea of self-rejection in a context of anger. And guilt may derive from socialization to the idea of self-rejection for what is defined as a morally wrong action in the context of fear. The primary emotion contexts are important because they provide the autonomic, therefore specifically emotional, underpinnings of the secondary emotions. The cultural components, such as situational definitions and emotional labels, are important because they help the person differentiate and ascribe the feeling to particular social and behavior contexts.

Gibson (1997) offers a different approach to the reconciliation of the structural and cultural approaches to emotions by incorporating aspects of both in his model for feeling and expression of emotions in organizations. Both structural conditions and display rules operate to instigate and control emotions in organizational contexts.

### REASON AND EMOTION

One of the longest-standing problems in the study of emotions is the relationship between emotions and reason. This question has engaged two millennia of philosophers and psychologists, including Aristotle, Aquinas, Spinoza, Hume, and Freud. Sociologists are latecomers here, but have substantially and persuasively claimed that reason is not an isolated domain of human action, but is imbued with emotion (Kemper 1993), a point that a sociological approach makes particularly clear.

Max Weber ([1922] 1947) set the stage by distinguishing between *Zweckrational*, or expedient action, and at least two types of emotional action: *Affectuel*, or impulsive action, and *Wertrational*, or value-oriented action. But he did little to develop the relationship between the two emotional types and the strictly expedient type. The latter is the prototypical action of economic theory, where means are examined and selected to attain the best possible outcome. In the economic version, money is the usual yardstick for the efficacy of the decision based on expediency. Other considerations are treated more or less as "noise," disturbing the adequacy of the model. The sociological version of this economic approach is rational choice theory.

In a radical confrontation with rational choice theory, Collins (1993) throws down the gauntlet to

economically based theories. He argues that the main preference order is based on emotional currency, namely emotional energy (EE) which is acquired in successful interactions with others (see earlier description of Collins' work). The good feelings—confidence and enthusiasm—that individuals derive from participation in interaction are the *summum bonum*, and this is what individuals are attempting to maximize in their so-called rational choices, regardless of what the currency may appear to be—money, status, practical or aesthetic enjoyment of material goods, and so on.

Lawler and Thye (1999) approach the rational choice issue through the window of social exchange theory, in which self-interested actors are trying to obtain something of value from other self-interested actors. They examine the *context* of exchange, which ordinarily will have a certain emotional tone and emotional requirements, the *processes* of exchange, which may make actors feel satisfied, excited, or otherwise emotionally aroused, and the *outcome* of exchange at which point actors may feel gratified or angry, prideful or crestfallen. In a useful schematic formulation, Lawler and Thye organize the context, process, and outcome features of exchange according to six different mainly sociological approaches to the study of emotions. For context, the *cultural-normative* approach (Hochschild 1979) and the *structural-relation* approach (Collins 1975, Kemper 1978); for processes, a psychologically oriented *social-cognitive* approach (Bower 1991) and a *sensory-informational* approach (Heise 1979); and for outcomes, a *social attribution* approach (Weiner 1986) and a *social formations* approach (Collins 1981 and Lawler and Yoon 1998). This schematic formulation enables analysts to move easily into the examination of motions in social exchange situations.

### EMOTIONS AND MACROPROCESSES

Most sociological examinations of emotion are social-psychological; that is, social structures, processes, or outcomes of these are seen to produce emotions in the individuals involved, with emotions differing according to where in the structure, process, or outcome the individual stands. Jack Barbalet (1998) provides an important exception to this social-psychological approach, conceiving of emotion as integral to social relations and social

processes themselves. Emotion is felt by individuals—this cannot be escaped—but as an aspect of societal patterns of social organization in terms of class, gender, race, and the like. This leads to another perspectival difference: Most sociological approaches to emotion examine social processes and social relations as the independent variables—they cause or produce emotions. Barbalet reverses this and examines how emotions cause or produce social processes and social relations. Furthermore, this is conceived at the macro level, engaging societal, as opposed to interpersonal, processes.

For example, working-class individuals might be expected to harbor social resentment against those who are better off, but such resentment is scant in the United States and has led to no effective political movements. Following Bensman and Vidich (1962), Barbalet tries to explain this in part by locating different sectors of the working class in different places in normal trade cycles in capitalist societies. A dynamic economy contains both expanding (e.g., computers) and contracting (e.g., textiles) industries, and workers in the different industries cannot be expected to experience the same emotions, thus vitiating any theory or program that views workers in a monolithic way.

In another venture into the macrosociology of emotions, Barbalet examines the emotion, mood, or feeling of confidence as an important feature of social process. Particularly in the business community, confidence is a necessary condition of investment. Indeed, in Barbalet's view, confidence dominates even rational calculation. This is because rational business planning is limited by the fact that it is future-oriented and the information that rational assessment requires is unavailable—it can only unfold in the future. Therefore, to undertake action under conditions of limited rationality, the business community must rely on its intuition that investment will be profitable. Put otherwise, it must have confidence. Government is an important constituent of the situation, sometimes enhancing and sometimes depressing business confidence. Barbalet proposes that what differentiates these effects of government policy is whether they reflect "acceptance and recognition" of the business community. For example, government spending on infrastructure or the bailout of the savings and loan industry reflect such appreciation. On the other hand, business interests feel slighted when the government proposes strict policies to

reduce global warming, and business confidence falls accordingly. In both examples—those of the working class and business—the emotions are aggregated products of many individuals that then act as a discrete force in society.

In an unusually strong entry of a macrosociological approach into the domain of emotions, Jasper (1998) and Goodwin and Jasper (1999) have argued for the overlooked importance of emotions in the understanding of large scale social movements. Social movements, they argue, are awash in emotions. Anger, fear, envy, guilt, pity, shame, awe, passion and other feelings play a part either in the formation of social movements, in their relations with their targets who are either antagonists or possible collaborators, and in the lives of potential recruits and members. Without the emotions engaged in movement environments, dynamics, and structure it would be hard to explain how social movements arise, amass critical levels of support, maintain such support in long enduring campaigns in the face of often intense opposition, and provide means for recruiting and sustaining supporters, both as active members and as favorably disposed publics and bystanders. Understanding the dynamics of emotions thus clarifies social movement dynamics.

### CONCLUSION

The sociology of emotions has a long history but only a short recent life. It is a diverse speciality, reflecting many of the axial divisions that currently rend the field, but one that lends itself to the illumination of a large number of problem areas from the micro to the macro level. The main requirement for the present-day sociology of emotions is, as Peggy Thoits (1989) has argued, to pursue empirical support for its many theories. Only in this way will it become clear in which direction theory can most fruitfully go.

(SEE ALSO: *Affect Control Theory and Impression Formation; Rational Choice; Social Exchange Theory*)

### REFERENCES

Barbalet, Jack 1998 *Emotion, Social Theory and Social Structure: A Macrosociological Analysis*. Cambridge, U.K.: Cambridge University Press.

- Bensman, Joseph, and Arthur Vidich 1962 "Business Cycles, Class and Personality." *Psychoanalysis and the Psychoanalytic Review* 49:30–52.
- Bower, G. H. 1991 "Mood Congruity of Social Judgments." In J. Forgas, ed., *Emotion and Social Judgments*. Oxford: Pergamon.
- Bullough, Edward 1912 "Psychic Distance as a Factor in Art and as an Aesthetic Principle." *British Journal of Psychology* 5: 669–679.
- Clark, Candace 1997 *Misery and Company: Sympathy in Everyday Life*. Chicago: University of Chicago Press.
- Collins, Randall 1975 *Conflict Sociology*. New York: Academic Press.
- 1981 "On the Microfoundations of Macrosociology." *American Journal of Sociology* 86: 984–1014.
- 1990 "Stratification, Emotional Energy, and the Transient Emotions." In T. D. Kemper, ed., *Research Agendas in the Sociology of Emotions*. Albany: State University of New York Press.
- 1993 "Emotional Energy as the Common Denominator or Rational Action." *Rationality and Society* 5: 203–230.
- Cooley, Charles H. (1902) 1922 *Human Nature and the Social Order*, Rev. ed. New York: Scribner's.
- Durkheim, Émile (1915) 1965 *The Elementary Forms of the Religious Life*, trans. J. W. Swain. New York: Free Press.
- Elias, Norbert 1978a *The Civilizing Process*. New York: Urizen Books.
- 1978b *The History of Manners*. New York: Urizen Books.
- Gibson, Donald E. 1997 "The Struggle for Reason: The Sociology of Emotions in Organizations." In R. J. Erickson and B. Cuthbertson-Johnson, eds., *Social Perspectives on Emotion*, Vol. 4. Greenwich, Conn.: JAI Press.
- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, N.Y.: Doubleday.
- 1967 *Interaction Ritual*. Garden City, N.Y.: Doubleday.
- 1981 *Forms of Talk*. Philadelphia, Pa.: University of Pennsylvania Press.
- Goodwin, Jeff, and James M. Jasper 1999 "Caught in a Winding, Snarling Vine: The Structural Bias of Political Process Theory." *Sociological Forum* 14: 27–54.
- Gordon, Steven L. 1981 "The Sociology of Sentiments and Emotions." In M. Rosenberg and R. S. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.



- Heise, David R. 1979 *Understanding Events: Affect and the Construction of Social Action*. New York: Cambridge University Press.
- Hochschild, Arlie Russell 1979 "Emotion Work, Feeling Rules, and Social Structure." *American Journal of Sociology* 85:551-575.
- 1983 *The Managed Heart: Commercialization of Human Feelings*. Berkeley: University of California Press.
- Jasper, James M. 1998 "The Emotions of Protest: Affective and Reactive Emotions in and Around Social Movement." *Sociological Forum* 13: 397-424.
- Kemper, Theodore D. 1978 *A Social Interactional Theory of Emotions*. New York: Wiley.
- 1987 "How Many Emotions Are There? Wedding the Social and the Autonomic Components." *American Journal of Sociology* 93:263-289.
- 1989 "Love and Like and Love and *Love*." In D. D. Franks and E. D. McCarthy, eds., *The Sociology of Emotions: Original Essays and Research Papers*. Greenwich, Conn.: JAI Press.
- 1990 "Social Relations and Emotions: A Structural Approach." In T. D. Kemper, ed., *Research Agendas in the Sociology of Emotions*. Albany: State University of New York Press.
- 1993 "Reason in Emotions or Emotions in Reason." *Rationality and Society* 5: 275-282.
- and Muriel T. Reid 1997 "Love and Liking in the Attraction and Maintenance Phases of Long-Term Relationships." In R. Erickson and B. Cuthbertson-Johnson, eds., *Social Perspectives on Emotion*, Vol. 4. Greenwich, Conn.: JAI Press.
- Lawler, Edward J., and Jeongkoo Yoon 1998 "Network Structure and Emotion in Exchange Relations." *American Sociological Review* 63: 871-894.
- and Shane R. Thye 1999 "Bringing Emotions into Social Exchange Theory." In K. Cook and J. Hagan, eds., *Annual Review of Sociology*, Vol. 25. Palo Alto, Calif.: Annual Reviews.
- Lewis, Helen B. 1971 *Shame and Guilt in Neurosis*. New York: Wiley.
- MacKinnon, Neil J. 1994 *Symbolic Interaction as Affect Control*. Albany: State University of New York Press.
- Marx, Karl (1842-1844) 1971 *The Early Texts*, Ed. D. McLellan, Oxford: Blackwell.
- (1867) 1967 *Capital: A Critique of Political Economy*. New York: International Publishers.
- and Friedrich Engels (1846) 1947 *The German Ideology*. New York: International Publishers.
- Mead, George H. 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Morris, J. Andrew, and Daniel C. Feldman 1996 "The Dimensions, Antecedents, and Consequences of Emotional Labor." *Academy of Management Review* 21: 986-1010.
- 1979 *Catharsis in Healing, Ritual, and Drama*. Berkeley: University of California Press.
- 1988 "Shame and Conformity: The Deference-Emotion System." *American Sociological Review* 53:395-406.
- Scheff, Thomas J. 1994 *Bloody Revenge: Emotions, Nationalism, and War*. Boulder, Colo.: Westview Press.
- 1997 *Emotions, the Social Bond, and Human Reality: Part/Whole Analysis*. New York: Cambridge University Press.
- Seeman, Melvin 1959 "On the Meaning of Alienation." *American Sociological Review* 24: 783-791
- Shott, Susan 1979 "Emotion and Social Life: A Symbolic Interactionist Analysis." *American Journal of Sociology* 84:1317-1334.
- Smith, Adam (1759) 1853 *The Theory of Moral Sentiments*. Rev. ed. London: G. Bell and Sons.
- Smith-Lovin, Lynn, and David R. Heise 1988 *Analyzing Social Interaction: Advances in Affect Control Theory*. New York: Gordon and Breach.
- Thamm, Robert 1992 "Social Structure and Emotion." *Sociological Perspectives* 35:649-671.
- Thoits, Peggy 1989 "The Sociology of Emotions." *Annual Review of Sociology* 15:317-342,
- 1990 "Emotional Deviance: Research Agendas." In T. D. Kemper, ed., *Research Agendas in the Sociology of Emotions*. Albany: State University of New York Press.
- Weber, Max (1904-1905) 1958 *The Protestant Ethic and the Spirit of Capitalism*, trans. T. Parsons. New York: Scribner's.
- Weiner, Bernard 1986 *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.

THEODORE D. KEMPER

## ENVIRONMENTAL EQUITY

The scientific study of the social distribution of environmental impacts has quickly become an important area of inquiry within environmental sociology. Scholarly interest in this topic—which has been referred to as environmental justice,

environmental equity, or environmental racism—does not derive from formal theories of differential environmental impacts, but rather has been inspired by the rapidly evolving environmental justice movement. Sociologists have studied both the social movement itself and the claims of environmental inequity made by the movement's proponents. The study of environmental justice has important implications for other areas of sociology, including stratification, race relations, sociology of health, and the study of social movements.

### THE ENVIRONMENTAL JUSTICE MOVEMENT

**Origins of the Movement.** In the summer of 1978, the state of New York finally acknowledged what the residents of Love Canal had suspected for some time—that some of the 21,000 tons of chemical waste that had been dumped into an abandoned canal between 1942 and 1952 were now leaking into the basements of nearby homes and percolating into the playgrounds of the elementary school that had been built directly above the filled canal (Levine 1982). The situation presented “great and imminent peril to the health of the general public residing at or near the site,” according to the New York state commissioner of health (as quoted in Levine 1982, p. 7). This announcement galvanized the concerns and suspicions of many residents of this working-class community, who for years had endured bad smells, mysterious sinkholes, skin irritations, respiratory problems, spontaneous abortions, and birth defects. Soon after this announcement, the private concerns of individual people began to develop into an organized social protest that pressured the state and federal governments to take action to protect the health, safety, and property of the citizens.

The events at Love Canal continued to make national headlines for the next two years. In 1979, before the media or the public could forget about environmental health risks, an accident occurred at the Three Mile Island nuclear power plant near Harrisburg, Pennsylvania (Walsh 1981). Large quantities of radioactive gas and water were released, and over 100,000 residents within a 15-mile radius of the plant were evacuated. Overnight, public support rallied behind local anti-nuclear groups that had previously been viewed as radical.

Soon after, grassroots anti-toxics movements began to proliferate in the United States. In addition to the media attention given to Love Canal and Three Mile Island, a number of other factors set the stage for the emergence of these movements. The growing popularity of the mainstream environmental movement and books such as Rachel Carson's *Silent Spring* (1962) had made much of the public aware that there were health hazards surrounding man-made chemicals, that citizens could be exposed to them unknowingly, and that technical and political solutions to such problems were not readily available (Cable and Shriver 1995; Freudenberg and Steinsapir 1992; Szasz 1994). In addition, problems associated with the unsafe disposal practices used during the post-World War II industrial boom began to surface at this time, causing increasing discovery of new toxic contamination problems.

Racial issues quickly became central to the environmental justice movement. The first environmental justice protest organized by African Americans took place in 1982 in Warren County, North Carolina, where the state proposed to build a landfill for the disposal of hazardous polychlorinated biphenyls (PCBs) that had been illegally dumped in 14 different North Carolina counties (Bullard 1994). This was a poor, mostly black, rural community. Why had this community been chosen to receive the landfill? Local residents believed that race was a deciding factor. They called upon black civil rights leaders for support, and a protest was organized. Although the landfill was eventually constructed, the Warren County protest was a landmark in the history of the environmental justice movement. Since then numerous other communities of color, primarily African American, Hispanic, and Native American, have become active in the environmental justice movement (for examples, see Bullard 1993 or Hofrichter 1993).

The formation of local-level anti-toxics movements has tended to follow a general pattern (Cable and Benson 1993; Cable and Shriver 1995). The first stage is the recognition that there is a local toxics hazard. Often, this recognition is a result of health problems among members of the community. Thus it is frequently women, particularly housewives, who are the first to become involved with toxics issues, since they are most likely to take responsibility for the health of the family (Krauss 1993; Brown and Masterson-Allen

1994). Once a toxics problem is perceived by community members, they typically turn to official administrative and regulatory bodies for correction of the problem. The results are usually disappointing, and residents often come to believe that these institutions actually serve the interests of polluters more than the interests of residents (Bullard 1994; Cable and Benson 1993; Krauss 1989; Levine 1982; Molotch 1970). A sense of injustice, and often a deep sense of betrayal, develops among local activists, and they begin to pursue alternative means of social control through social movement organizations.

Most anti-toxics movements emerge in poor, working class, or minority communities (Brown and Masterson-Allen 1994; Freudenberg and Steinsapir 1992). Many activists and analysts alike argue that this is due to the disproportionate share of toxic pollution borne by these communities (Austin and Schill 1991; Bullard 1994). More affluent communities have more resources to draw on—including adequate funding, an educated membership, access to professional advice, and social networks that overlap with local power structures—and thus they are better able to control pollution problems through the political and regulatory processes (Hurley 1995; Molotch 1970).

The notion that less powerful communities have been targeted for polluting facilities was further reinforced by the widespread circulation among the grassroots community of a report written by the consulting firm Cerrell Associates, Inc. for the California waste industry. The report recommended that waste incinerators would most easily be sited in communities least likely to generate political resistance (i.e., communities that have populations that are older, more conservative, less educated, and have lower average incomes) (Cerrell Associates, Inc. 1984; Cable and Shriver 1995).

**Relationship to Mainstream Environmental Groups.** The environmental justice movement differs significantly from the “mainstream” environmental movement in terms of its membership, values, and tactics. The mainstream environmental movement has focused primarily on issues of conservation of natural resources and preservation of natural beauty. Membership in mainstream environmental organizations grew rapidly in the United States from 1960 to 1990, with three groups—the Sierra Club, the National Wildlife

Federation, and the National Audubon Society—each claiming more than 500,000 members (Mitchell et al. 1991). Mainstream environmental groups have professional staffs, generate funds through large foundation grants and membership dues, and work primarily within the system by using lobbying and litigation to influence government action.

Mainstream environmentalism has often been characterized as an elite movement. Early studies supported this general view when they found the greatest support for environmental causes among people who are young, white, well-off, urban, well-educated, and politically liberal (Buttel 1979; Buttel and Flinn 1974; Kreger 1973; Morrison et al. 1972). Some more recent research has confirmed these findings (Jones and Dunlap, 1992). However, one longitudinal study suggests the well-educated and well-off were simply the first to embrace environmentalism (Kanagy et al. 1994).

There has been some debate about the role of race in mainstream environmentalism. It has been common for social scientists to assume that African Americans will be less supportive of environmentalism than whites, but most evidence suggests that when it comes to environmental concern, there is no significant difference between blacks and whites (Jones and Carter 1994). Blacks *are* less likely, however, to be involved in mainstream environmental action (Taylor 1989; Mohai 1990). One explanation for this lag in environmental action is that blacks are more likely to experience a variety of stressors which detract from their ability to act upon environmental concerns. As Robert Bullard put it: “Decent and affordable housing, for example, is a top environmental problem for inner-city blacks” (1994, pp. 10–11).

In contrast to mainstream environmentalism, environmental justice movements tend to develop in poor, working class, or minority communities (Brown and Masterson-Allen 1994; Freudenberg and Steinsapir 1992). The two movements also have different goals—concerns about local health threats are central to the environmental justice movement. Protests often focus on pollution-generating facilities, such as factories, landfills, toxic waste dumps, incinerators, or nuclear power plants. Participants in these movements are generally people who believe they are personally at risk from the

facility. Thus, at least initially, these protests are generally grassroots-generated, and locally focused.

This focus on local issues has prompted some members of the mainstream environmental movement to charge that environmental justice activists are mere “NIMBY” (Not In My BackYard) protesters (Freudenberg and Steinsapir 1992). They accuse environmental justice activists of lacking concern for the environment in general and of acting out of pure self-interest. In turn, environmental justice activists have accused the mainstream environmental movement of representing elite interests and being content to let disadvantaged groups bear the toxic burden of modern affluence. As relative newcomers to environmental activism, members of the environmental justice movement have generally been disappointed with the lack of assistance given their cause from the larger, more established environmental groups (Bullard and Wright 1992; Taylor 1992).

Grassroots environmental protests have tended to use different tactics than the mainstream environmental groups. They typically lack the funds or the political clout to lobby and negotiate with elites. Thus they are more likely to use disruptive and attention-getting tactics, such as public protests, sit-ins, letter writing campaigns, and attendance at public hearings (Bullard 1994). In one extreme example, a group of residents in Love Canal took an EPA official hostage for several hours in an effort to generate media attention (Levine 1982). Environmental justice groups have also made use of the courts, primarily through tort actions that seek reparations for damage to health and property.

These differences between the two movements have had conflicting implications for social movement theory. The elite membership of mainstream environmental groups has often been viewed as support for *resource mobilization theory*. This theory proposes that the emergence of social movements has little to do with the existence or extent of grievances experienced by movement participants. Grievances are a given, everyone has them, and the development of active social movements depends upon the availability of resources available to particular groups of people to advance their causes (McCarthy and Zald 1977). In contrast, the environmental justice movement is based on grievances and is comprised of members with relatively

little resources. Current social movement theory appears inadequate to explain the environmental justice movement (Brown and Masterson-Allen 1994; Masterson-Allen and Brown 1990; Walsh 1981; Walsh et al. 1993).

**Expansion of the Environmental Justice Movement.** The environmental justice movement has evolved beyond its origin in the late 1970s as a collection of isolated local protests. Regional coalitions, such as the Southern Organizing Committee and the Southwest Network for Environmental and Economic Justice, have formed to facilitate the sharing of ideas, resources, and networks among geographically dispersed groups. A number of national support organizations have also been formed, most notably the Citizen’s Clearing House for Hazardous Waste (led by Lois Gibbs of Love Canal) and the National Toxics Campaign Environmental Justice Project (Moore and Head 1993). These organizations provide resources and information to smaller groups, and have conducted their own organizing and lobbying campaigns at the national level. Activists have also connected with the academic community, and a number of research and support centers for environmental justice have been established. These include the Environmental Justice Resource Center at Clark Atlanta University in Atlanta and the Deep South Center for Environmental Justice at Xavier University in New Orleans, among others (Wright 1995). Members of the environmental justice movement have increasingly sponsored meetings and panels that bring together activists, scholars, and the media.

Institutional responses to the environmental justice movement have also increased at the national level. In 1992 the Environmental Protection Agency established an Office of Environmental Equity (Cutter 1995). In 1994, President Clinton signed Executive Order 12898, which requires every federal agency to adhere to principles of environmental justice in its operations (Clinton 1994). The issue of racial disparity has allowed activists to file administrative complaints under Title VI of the 1964 Federal Civil Rights Act.

**A New Worldview—Ecological Democracy.** Many scholars have noted the profound change in worldview that often results from individual participation in the environmental justice movement. As noted earlier, environmental justice activists

often become disillusioned with the ability of the government to protect them from toxic contamination. This change is often conceived as a reframing of the toxics problem: what was initially perceived as an isolated technical or regulatory problem is now seen as a larger social issue involving the relative power of corporations and citizens, the role of the state, and conceptions of social justice (Capek 1993; Brown and Masterson-Allen 1994; Cable and Shriver 1995). Capek (1993) argues that, especially for black communities, the civil rights movement provided a “master frame” which legitimated the quest for environmental equality and permitted critical analysis of the social structure that created the inequity.

This new, broader perspective now dominates discourse on the causes of environmental inequity. As Heiman (1990) puts it, the Not In My BackYard perspective has become the Not In Anybody’s BackYard critique. When challenged as to where hazardous facilities should be located, rather than try to locate them in someone else’s backyard, environmental justice activists have responded by questioning the assumption that such facilities are a public good that must be sited somewhere. “The not-in-anybody’s-backyard stand forces the debate away from the suitability of specific waste treatment facilities of locations, and toward a more fundamental reassessment of the propriety of a production system under private control where, in the quest for profit, the public is exposed to known risks” (Heiman 1990, p. 361).

This new perspective was expressed in the *Principles of Environmental Justice*, a document created in 1991 by a group of activists at the People of Color Environmental Leadership Summit. Among the 17 principles adopted are these:

*1. Environmental Justice affirms the sacredness of Mother Earth, ecological unity and the interdependence of all species, and the right to be free from ecological destruction. 3. Environmental justice mandates the right to ethical, balanced and responsible uses of land and renewable resources in the interest of a sustainable planet for humans and other living things. 5. Environmental justice affirms the fundamental right to political, economic, cultural and environmental self-determination of all peoples. 6. Environmental justice demands the cessation of the production of all toxins,*

*hazardous wastes, and radioactive materials, and that all past and current producers be held strictly accountable to the people for detoxification and the containment at the point of production. 7. Environmental justice demands the right to participate as equal partners at every level of decision-making including needs assessment, planning, implementation, enforcement and evaluation. (Reprinted in Hofrichter, 1993, pp.237–239)*

In searching for the causes of environmental inequity, many analysts—academics and activists alike—have come to see issues of environmental justice as a result of an inherent contradiction within the liberal democratic capitalist state (Cable and Benson 1993; Hamilton 1993; Heiman 1990; Krauss 1989; see also O’Connor 1973). The essence of this argument is that the state is charged with two conflicting goals: to further the accumulation of capital and profits for private industry; and to protect the interests of individual citizens and communities. Particularly in terms of environmental regulation, these two goals are at odds because environmental regulations protect citizens from industrial externalities at the expense of profits.

This line of reasoning is compatible with that of social ecologists and eco-Marxists. For example, Faber (1998) states:

*The roots of America’s ecological problems and injustices are grounded in the expansionary dynamics of the global capitalists system . . . . To sustain the process of capital accumulation and higher profits in the new global economy, American capital is increasingly relying on ecologically and socially unsustainable forms of production. . . . In so doing, America’s corporate ruling class, the 1 percent of the population that owns 60 percent of all corporate stock and business assets, is serving its own narrow material interests at the expense of the environment, communities, and the health of working people. The reason is that corporate expenses related to human health and environmental quality do not typically increase labor productivity (hence potential profits) sufficiently to outweigh such expenditures. . . . It costs capital and the state much less to displace environmental health problems onto people who lack health care*

*insurance, possess lower incomes and property values, and as unskilled or semiskilled laborers are more easily replaced if they become sick or die. In this sense, environmental inequalities in all forms, whether they be class, race, gender, or geographically based, are socially constructed features grounded in the systemic logic of capitalist accumulation. (pp. 2–5)*

The emergence of this theoretical model of environmental justice has inspired some writers to call for a restructuring of society and a return to basic principles of democracy. They reject the false contradiction between individuals rights and social goods. In particular, they are challenging corporate power and control over production decisions (Pullido 1994; O'Connor 1993; Hamilton 1993). They favor a new social contract with the state to create an economic democracy in which decision making is decentralized. For example, O'Connor (1993) states:

*For the people to regain their sovereignty and restore their environment, fundamental changes in the structure of government are required that permit effective discussion, debate, and decision making. Every citizen, regardless of wealth, should be given an equal opportunity to bring ideas to stop environmental decline directly to the American people. In addition, each citizen will need new rights, in relation to polluting or poisonous industries, in order to allow citizens an opportunity to protect themselves and their communities from the everyday abuses of economic power, against which no level of regulation or conventional governance can fully protect them. (p. 52)*

**Global Issues in Environmental Justice.** The global dimensions of environmental justice issues have been receiving increased attention. Third World environmental problems, particularly natural resource depletion, have generally been viewed by established international authorities (particularly the United Nations) as the result of insufficient development and a lack of advanced technology (World Commission on Environment and Development, 1987). Alternatively, however, critical scholars see Third World environmental problems as the legacy of colonialism and continuing exploitation of the periphery by core nations (Esteva 1992; Harvey 1997; Redclift 1987; Sachs 1997; Weissman 1993). Plans to promote development

in the Third World often involve large-scale construction projects (often funded by the World Bank) that cause large-scale destruction of natural resources such as forests, rivers, farmlands, or native species. Agyeman and Evans (1996) explain the situation this way:

*The well-entrenched belief endemic to Northern societies is that the extant economic and social conditions of many Southern societies legitimately permit commercial behaviour that would not be accepted in the North, often justified in terms of bringing employment to areas of no work. However, the implicit, unstated position is fundamentally racist, in that such commercial behaviour is deemed as appropriate for black people whereas it is inappropriate for white people. It is far away from the economic power centres of the North, and any local representations concerning pollution or environmental degradation are likely to be muted for fear of unemployment. (pp. 74–75)*

A case in point is the Union Carbide pesticide factory near Bhopal, India that released a cloud of poisonous gasses in 1984, killing an estimated 6,600 people and injuring as many as 600,000. Such a massive facility probably could never have been sited in a Northern country, and it would have required much more expensive safety measures.

The involvement of U.S. environmental groups in Third World environmental issues has been strongest in the protection of forests and natural habitat, and has generally involved mainstream environmental groups and other organizations dedicated to global justice and equity (Keifer and Benjamin, 1993). One popular intervention is the debt-for-nature swap, in which a Northern environmental group pays a portion of a Third World country's debt on the condition that the recipient nation will establish a nature preserve or other conservation program. These types of programs are increasingly resented by the nations of the South, who view them as paternalistic extensions of the North's colonial control that interfere with local self-determination. One alternative that has been developed is the debt-for-Indian-stewardship swap, in which indigenous people gain control over parcels of land in exchange for debt relief (Alston and Brown 1993).

The export of toxic waste to Third World nations has also become a major political issue. As the cost of disposal rises, facilities in the United States and Western Europe are increasingly looking for new locations to dispose of their wastes. For example, a ton of waste that costs \$200 to store in the United States will cost \$40 to dispose in Benin (Mpanya 1992). About half of all African nations have been approached by Western interests to serve as dump sites. At least 25 African nations have agreed, motivated by payments of \$10 to \$25 million that can be used to make payments on international debt. There have also been several rather widely publicized cases (and probably many undiscovered cases) where toxic waste has been imported into Third World countries under false pretenses, often labeled as fertilizer or construction material. For example, 4,000 tons of toxic incinerator ash from Philadelphia were unloaded on a beach in Haiti, falsely labeled as fertilizer. The attempt to repatriate the toxic ash has dragged on for years (Bruno, 1998).

The economic logic of these practices was revealed in a leaked internal World Bank memo, written in 1991 by then chief economist Lawrence Summers, which was subsequently published in *The Economist* (September 8, 1992). It reads in part:

*Just between you and me, shouldn't the World Bank be encouraging more migration of the dirty industries to the LDC's (Lesser-Developed-Countries)? I can think of three reasons: 1) The measurement of the costs of health-impairing pollution depends on the foregone earnings from increased morbidity and mortality. From this point of view a given amount of health-impairing pollution should be done in the country with the lowest cost, which will be the country with the lowest wages. I think the economic logic behind dumping a load of toxic waste in the lowest-wage country is impeccable and we should face up to that. 2) The costs of pollution are likely to be non-linear as the initial increments of pollution probably have very low cost. I've always thought that underpopulated countries in Africa are vastly under-polluted; their (air pollution) is probably vastly inefficiently low compared to Los Angeles or Mexico City. Only the lamentable facts that so much pollution is generated by non-tradable industries (transport, electric generation) and that the unit transportation costs*

*of solid waste are so high prevent world welfare-enhancing trade in air pollution and waste. 3) The demand for a clean environment for aesthetic and health reasons is likely to have very high income elasticity. The concern over an agent that causes a one in a million change in the odds of prostate cancer is obviously going to be much higher in a country where people survive to get prostate cancer than in a country where under-5 mortality is 200 per thousand. Also, much of the concern over industrial atmosphere discharge is about visibility of particulates. These discharges may have little direct health impact. Clearly trade in goods that embody aesthetic pollution concerns could be welfare enhancing.*

In an attempt to address the problem of increasing toxic exports, the United Nations organized a convention held in Basel, Switzerland regarding the international movement of toxic waste. In 1989, 118 nations signed the Basel Convention on the Control of Transboundary Movements of Hazardous Waste and their Disposal. The Basel Convention prohibits the transfer of toxic wastes from member nations of the Organization for Economic Cooperation Development (OECD—an organization comprising the most industrialized and developed nations in the world) to non-OECD countries. However, many nations, including the United States, have yet to ratify the treaty and make it the law of the land. Furthermore, other international agreements, particularly international trade agreements such as the North American Free Trade Agreement (NAFTA) and the General Agreement on Tariffs and Trade (GATT) actually make it more difficult for individual nations to restrict the flow of toxic substances across their borders, since such restrictions can present impediments to free trade.

#### EMPIRICAL RESEARCH ON ENVIRONMENTAL EQUITY

**Research Findings.** For many environmental justice activists, the existence of inequitable environmental impacts along race and class lines is a reality. For sociologists and other scholars, however, the nature and extent of these inequities needs to be established through systematic and objective research. The protest against the Warren County landfill (described above) launched the current

interest in empirical investigation into claims of disproportionate siting of environmentally hazardous facilities in disadvantaged communities. One of the activists involved in the Warren County protest was Congressman Walter E. Fauntroy, who initiated a study by the General Accounting Office (GAO) of hazardous landfill siting in the South (GAO 1983). The GAO study, though limited in scope and analysis, found a relationship between the siting of landfills and the racial and economic characteristics of residents in surrounding communities. A more comprehensive study, also in response to the Warren County protest, was conducted by the Commission for Racial Justice of the United Church of Christ (United Church of Christ 1987). This study analyzed the presence of hazardous waste treatment, storage, and disposal facilities (TSDFs) in all zip codes in the United States, and concluded that racial composition was the strongest predictor of the presence of a TDSF. This study was circulated widely within the civil rights community and is considered to be the seminal investigation into racial bias in environmental risk. Interestingly, as Szasz and Mueser (1997) point out, the United Church of Christ analysis was not the first study of environmental justice. A number of investigations conducted in the 1970s, which focused more on social class than on race, had already found disparities in exposures to environmental risks, but these studies were largely ignored by the social science community. It took the politicizing force of the charge of racism to get scholars to take up serious investigation of the distribution of environmental hazards.

Since the United Church of Christ report, research on this topic has grown rapidly. Most studies, with some significant exceptions, have reported some type of social disparity in the distribution of environmental hazards. (See Szasz and Meuser 1997 for a comprehensive review of environmental justice research.) Higher levels of environmental hazards in places with lower socioeconomic status have been found by the United Church of Christ (1987), Burke (1993), Pollack and Vittas (1995), Been (1997), Boer et al. (1997), Brooks and Sethi (1997), Ringquist (1997), Kreig (1998), and Daniels and Friedman (1999). Similar findings for minority communities have been reported by the United Church of Christ (1987), Burke (1993), Zimmerman (1993), Perlin et al. (1995), Pollack

and Vittas (1995), Been (1997), Boer et al. (1997), Brooks and Sethi (1997), Ringquist (1997), Kreig (1998), Stretesky and Hogan (1998), and Daniels and Friedman (1999). Research by Anderton et al. (1994; 1997), Bowen et al. (1995), and Cutter et al. (1996), however, found no significant evidence of race or class disparity in the distribution of environmental hazards.

**Methodological Issues.** It is important to note that the studies mentioned above have used a variety of different research designs, and thus are not entirely comparable. For example, researchers have studied a wide range of geographic locations, ranging in size from the entire United States to a single metropolitan area, and results from one location may have little bearing on the situation in others. Furthermore, some of these differences in design point to important methodological issues that need to be considered in evaluating environmental equity.

One of the most important methodological issues involves the conceptualization and measurement of pollution. What constitutes a hazard? A wide variety of environmental risks have been discussed in the literature, including lead poisoning, pesticide exposure, toxic fish consumption, occupational hazards, nuclear facilities, municipal landfills, hazardous waste sites, and industrial emissions. Empirical research to date has most often used data on three types of hazards: hazardous waste treatment storage and disposal facilities (TSDFs) (United Church of Christ 1987; Anderton et al. 1994; Cutter et al. 1996; Been 1997; Boer et al. 1997); Superfund sites or other sites known to be contaminated with hazardous wastes (Anderton et al. 1997; Cutter et al. 1996; Kreig 1998; Stretesky and Hogan 1998; Zimmerman 1993); and data on industrial toxic releases from the EPA's Toxics Release Inventory (TRI) (Burke 1993; Bowen et al. 1995; Brooks and Sethi 1997; Daniels and Friedman 1999; Perlin et al. 1995; Pollack and Vittas 1995; Ringquist 1997). The focus on these particular environmental hazards stems partially from the seriousness of their potential impacts on human health, but it is also due to the simple fact that there are publicly available data for these hazards.

The extent of the risk posed by these forms of pollution is an important variable that sociologists have yet to fully address. Some researchers have



tried to take this issue into account by measuring the distance of resident populations from the hazard (Pollack and Vittas 1995), and others have included data on the toxicity of released chemicals in their analyses (Bowen et al. 1995). However, we still have relatively little information about exposure pathways and the risks to health that result from the environmental hazards under study. Epidemiologists, statisticians, and others have reported associations between the proximity of various environmental hazards and heightened morbidity and mortality (e.g., Geschwind et al. 1992; Kelsall 1997), but there are currently many more questions than answers about the mechanisms by which pollutants cause disease, the dose-response relationships involved, and potential synergistic effects (National Research Council 1991). These issues, although outside the domain of sociological inquiry, are highly relevant to the study of environmental equity, since adverse health outcomes lie at the core of citizen concerns about environmental contamination.

Another methodological issue involves a controversy in the literature concerning the appropriate geographic unit for the study of environmental equity. Some analysts have declared that, due to their smaller size, census tracts or block groups are more appropriate than zip codes, counties, or regions (Anderton et al. 1994; Bowen et al. 1995; Cutter et al. 1996). They suggest that findings of racial inequity such as that found in the United Church of Christ study are the result of an ecological fallacy resulting from the use of too large a geographic unit. Large units can mask significant heterogeneity in the distribution of residents, hazards, or both. Whether the ecological fallacy has played a role in findings of environmental inequity is still unresolved, as attempts to address the issue have been inconclusive. Bowen et al. (1995) claimed to compare two geographic levels in their study of environmental equity in Ohio, however, they analyzed first all counties in Ohio and then census tracts only in Cuyahoga County. Thus it is impossible to know if differences between the two levels were due to the unit of analysis or the area of study. Cutter et al. (1996) studied environmental justice in South Carolina at the county, tract, and block group levels. While they did observe that bivariate relationships changed significantly across geographic units, they found little evidence of

environmental inequity in South Carolina at any level, and thus no evidence of the ecological fallacy.

**Future Directions.** Evidence is mounting that disparities do exist in the distribution of environmental hazards according to race and class. But there is still much to know with regard to which groups of people are exposed to which hazards. As noted above, there are many types of environmental hazards and they may have very different distributions and impacts.

Perhaps the most significant question raised by this line of research is the question of process. How have disadvantaged communities come to be associated with greater environmental risks? Which came first—are hazardous facilities sited in already disadvantaged communities, or do poor and minority residents tend to settle near already existing facilities? This question is particularly significant because the answers may have important implications for policies intended to remedy inequities. Szasz and Meuser (1997) suggest six possible scenarios, ranging from overt discrimination in the siting process to market rationality. Most of the empirical research has been cross-sectional in design, and thus cannot address this important issue. One major barrier has been the lack of good historical data on environmental hazards. A few longitudinal studies have been conducted, but the results have been inconclusive (Been 1994; Been 1997; Oakes et al. 1996). Evidence about process from case studies of polluted communities, while perhaps not generalizable, suggests the operation of complex reciprocal processes (Hersch 1995; Hurley 1995). Traditional heavy industry in Pittsburgh, Pennsylvania and Gary, Indiana, for example, appears to have been located in places near raw materials and transportation. Initially, the lack of transportation meant that more-affluent white workers lived closer to the factories. Over time, a number of processes brought African Americans and poor whites into closer proximity to pollution. Improved transportation and affluence, combined with housing segregation, allowed middle-class whites to move away from the factories into new housing on the suburban fringe, while blacks moved into downtown locations. As the cities grew and environmental regulation increased, it became necessary to locate sites for the disposal of wastes. Affluent communities with more political clout were able to keep such facilities outside their

borders. Less advantaged communities were more vulnerable to the siting of waste facilities due to a lack of effective zoning regulations, a reluctance to protest anything that might bring jobs, and a lack of information about the types of facilities that were being proposed for construction.

This question of process points to the need to establish links between existing sociological theories and the issues raised by the environmental justice movement regarding the distribution of environmental impacts. What larger social forces have had a role in producing these outcomes? As noted above, some scholars have begun to link environmental inequities to general theories of capitalist production. In order to demonstrate the relevance of these ideas (or other theories, for that matter) to environmental justice, more specific mechanisms of inequity must be identified. Several factors have been suggested, including housing discrimination, the market dynamics of land values, occupational segregation, and procedural inequities in environmental regulation. Hopefully, future research will begin to address these challenging issues.

## REFERENCES

- Agyeman, Julian, and Bob Evans 1996 "Black on Green: Race, Ethnicity and the Environment." In Susan Buckingham-Hatfield and Bob Evans, eds., *Environmental Planning and Sustainability*. New York: John Wiley and Sons.
- Alston, Dana, and Nicole Brown 1993 "Global Threats to People of Color." In Robert D. Bullard, ed., *Confronting Environmental Racism: Voices from the Grassroots*. Boston: South End Press.
- Anderton, Douglas L., Andy B. Anderson, John Michael Oakes, and Michael R. Fraser 1994 "Environmental Equity: The Demographics of Dumping." *Demography* 31:2 (May) 229-248.
- , John Michael Oakes, and Karla L. Egan 1997 "Demographics of the Discovery and Prioritization of Abandoned Toxic Sites." *Evaluation Review* 21(1):3-26.
- Austin, Regina, and Michael Schill 1991 "Black, Brown, Poor, and Poisoned: Minority Grassroots Environmentalism and the Quest for Eco-Justice." *The Kansas Journal of Law & Public Policy* 1(1):69-82.
- Been, Vicki 1997 "Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims." *Ecology Law Quarterly* 24(1):1-56.
- 1994 "Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics?" *The Yale Law Journal* 103:6 (April):1383-1422.
- Boer, J. Tom, Manuel Pastor, Jr., James L. Sadd, and Lori D. Snyder 1997 "Is There Environmental Racism? The Demographics of Hazardous Waste in Los Angeles County." *Social Science Quarterly* 78(4):793-810.
- Bowen, William M., Mark J. Salling, Kingsley Haynes, and Ellen J. Cyran 1995 "Toward Environmental Justice: Spatial Equity in Ohio and Cleveland." *Annals of the Association of American Geographers* 85(4):641-663.
- Brooks, Nancy, and Rajiv Sethi 1997 "The Distribution of Pollution: Community Characteristics and Exposure to Air Toxics." *Journal of Environmental Economics and Management* 32:233-250.
- Brown, Phil, and Susan Masterson-Allen 1994 "The Toxic Waste Movement: A New Type of Activism." *Society and Natural Resources* 7:269-287.
- Bruno, Kenny 1998 "Philly Waste Go Home." *Multinational Monitor* (February).
- Bullard, Robert D., ed. 1993 *Confronting Environmental Racism: Voices from the Grassroots*. Boston: South End Press.
- 1994 *Dumping in Dixie: Race, Class, and Environmental Quality*, 2nd ed. San Francisco: Westview Press.
- , and Beverly H. Wright 1992 "The Quest for Environmental Equity: Mobilizing the African-American Community for Social Change." In Riley E. Dunlap and Angela G. Mertig, eds., *American Environmentalism: The U.S. Environmental Movement*. New York: Taylor and Francis.
- Burke, Laretta M. 1993 "Race and Environmental Equity: A Geographic Analysis in Los Angeles." *Geo Info Systems* (October):44-50.
- Buttel, Frederick H. 1979 "Age and Environmental Concern: A Multivariate Analysis." *Youth and Society* 10(3):237-256.
- Buttel, F. H., and W. L. Flinn 1974 "The Structure of Support for the Environmental Movement, 1968-1970." *Rural Sociology* 39:56-69.
- Cable, Sherry, and Michael Benson 1993 "Acting Locally: Environmental Injustice and the Emergence of Grass-roots Environmental Organizations." *Social Problems* 40(4):464-477.
- , and Thomas Shriver 1995 "Production and Extrapolation of Meaning in the Environmental Justice Movement." *Sociological Spectrum* 15(4):419-442.

- Capek, Stella M. 1993 "The 'Environmental Justice' Frame: A Conceptual Discussion and an Application." *Social Problems* 40(1):5-24.
- Carson, Rachel 1962 *Silent Spring*. New York: Fawcett.
- Cerrell Associates, Inc. 1984 *Political Difficulties Facing Waste-to-Energy Conversion Plant Siting*. Los Angeles: Cerrell Associates.
- Clinton, William J. 1994 "Executive Order 12898 of February 11, 1994: Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations." *The Federal Register* 59:32 (February 16):7629-7633.
- Cutter, Susan L. 1995 "Race, Class and Environmental Justice." *Progress in Human Geography* 19(1):111-122.
- , Danika Holm, and Lloyd Clark 1996 "The Role of Geographic Scale in Monitoring Environmental Justice." *Risk Analysis* 16(4):517-526.
- Daniels, Glynis, and Samantha R. Friedman 1999 "Spatial Inequality and the Distribution of Industrial Toxic Releases: Evidence from the 1990 TRI." *Social Science Quarterly* 80(2):244-262.
- Esteva, Gustavo 1992 "Development." In Wolfgang Sachs, ed., *The Development Dictionary: A Guide to Knowledge as Power*. London: Zed Books.
- Faber, Daniel 1998 "The Struggle for Ecological Democracy and Environmental Justice." In Daniel Faber, ed., *The Struggle for Ecological Democracy: Environmental Justice Movements in the United States*. New York: Guildford Press.
- Freudenberg, Nicholas, and Carol Steinsapir 1992 "Not in Our Backyards: The Grassroots Environmental Movement." In Riley E. Dunlap and Angela G. Mertig, eds., *American Environmentalism*. New York: Taylor and Francis.
- General Accounting Office 1983 *Siting of Hazardous Waste Landfills and Their Correlation with Racial and Economic Status of Surrounding Communities*. (June 1) Washington, D.C.: U.S. General Accounting Office.
- Geschwind, Sandra A., Jan A. J. Stolwijk, Michael Bracken, Edward Fitzgerald, Alice Stark, Carolyn Olsen, and James Melius 1992 "Risk of Congenital Malformations Associated with Proximity to Hazardous Waste Sites." *American Journal of Epidemiology* 135(11):1197-1207.
- Hamilton, Cynthia 1993 "Environmental Consequences of Urban Growth and Blight." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- Harvey, David 1997 "The Environment of Justice." In Andy Merrifield and Erik Swyngedouw, eds., *The Urbanization of Injustice*. New York: New York University Press.
- Heiman, Michael 1990 "From 'Not in My Backyard!' to 'Not in Anybody's Backyard!' Grassroots Challenge to Hazardous Waste Facility Siting." *Journal of the American Planning Association* 56:359-362.
- Hersch, Robert 1995 "Race and Industrial Hazards: An Historical Geography of the Pittsburgh Region, 1900-1990." Discussion Paper 95-18. Washington, D.C.: Resources for the Future.
- Hofrichter, Richard, ed. 1993 *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- Hurley, Andrew 1995 *Environmental Inequalities: Class, Race, and Industrial Pollution in Gary, Indiana, 1945-1980*. Chapel Hill, N.C.: University of North Carolina Press.
- Jones, Robert Emmet, and Riley E. Dunlap 1992 "The Social Bases of Environmental Concern: Have They Changed Over Time?" *Rural Sociology* 57(1):28-47.
- Kanagy, Conrad L., Craig R. Humphrey, and Glenn Firebaugh 1994 "Surging Environmentalism: Changing Public Opinion or Changing Publics?" *Social Science Quarterly* 75(4):804-819.
- Kelsall, J. E., J. M. Samet, S. L. Zeger, and J. Xu 1998 "Air Pollution and Mortality in Philadelphia, 1974-1988." *American Journal of Epidemiology* 146(9):750-762.
- Kiefer, Chris, and Medea Benjamin 1993 "Solidarity with the Third World: Building an International Environmental-Justice Movement." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- Krauss, Celene 1993 "Blue Collar Women and Toxic Waste Protests." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- 1989 "Community Struggles and the Shaping of Democratic Consciousness." *Sociological Forum* 4(2):227-239.
- Kreger, Janet 1973 "Ecology and Black Student Opinion." *Journal of Environmental Education*. 4(3):30-34.
- Kreig, Eric J. 1998 "The Two Faces of Toxic Waste: Trends in the Spread of Environmental Hazards." *Sociological Forum* 13(1):3-23.
- Levine, Adeline Gordon 1982 *Love Canal: Science, Politics, and People*. Lexington, Mass.: Lexington Books.
- Masterson-Allen, Susan, and Phil Brown 1990 "Public Reaction to Toxic Waste Contamination: Analysis of

- a Social Movement." *International Journal of Health Services* 20(3):485-500.
- McCarthy, John D., and Mayer N. Zald 1977. "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82:1212-1241.
- Mitchell, Robert Cameron, Angela G. Mertig, and Riley E. Dunlap 1991 "Twenty Years of Environmental Mobilization: Trends Among National Environmental Organizations." *Society and Natural Resources* 4(3):219-234.
- Mohai, Paul 1990 "Black Environmentalism." *Social Science Quarterly* 71:745-765.
- Molotch, Harvey 1970 "Oil in Santa Barbara and Power in America." *Sociological Inquiry* 40 (Winter):131-144.
- Moore, Richard, and Louis Head 1993 "Acknowledging the Past, Confronting the Present: Environmental Justice in the 1990s." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- Morrison, D. E., K. E. Hornback, and W. K. Warner 1972 "The Environmental Movement: Some Preliminary Observations and Predictions." In W. R. Burch, Jr., N. H. Cheek Jr., and L. Taylor, eds., *Social Behavior, Natural Resources, and the Environment*. New York: Harper and Row.
- Mpanya, Mutombo 1992 "The Dumping of Toxic Waste in African Countries: A Case of Poverty and Racism." In Bunyan Bryant and Paul Mohai, eds., *Race and the Incidence of Environmental Hazards: A Time for Discourse*. San Francisco: Westview Press.
- National Research Council 1991 *Environmental Epidemiology: Public Health and Hazardous Wastes*. Washington, D.C.: National Academy Press.
- Oakes, John Michael, Douglas L. Anderton, and Andy B. Anderson 1996 "A Longitudinal Analysis of Environmental Equity in Communities with Hazardous Waste Facilities." *Social Science Research* 25:125-148.
- O'Connor, James R. 1973 *The Fiscal Crisis of the State*. New York: St. Martin's Press.
- O'Connor, John 1993 "The Promise of Environmental Democracy." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.
- Perlin, Susan A., R. Woodrow Setzer, John Creason, and Ken Sexton 1995 "Distribution of Industrial Air Emissions by Income and Race in the United States: An Approach Using the Toxic Release Inventory." *Environmental Science & Technology* 29(1):69-80.
- Pollock, Phillip H., III, and M. Elliot Vittas 1995 "Who Bears the Burdens of Environmental Pollution? Race, Ethnicity, and Environmental Equity in Florida." *Social Science Quarterly* 76(2):294-310.
- Pullido, Laura 1994 "Restructuring and the Contraction and Expansion of Environmental Rights in the United States." *Environment and Planning A* 26:915-936.
- Redclift, Michael 1987 *Sustainable Development: Exploring the Contradictions*. New York: Routledge.
- Ringquist, Evan J. 1997 "Equity and the Distribution of Environmental Risk: The Case of TRI Facilities." *Social Science Quarterly* 78(4):811-829.
- Sachs, Wolfgang 1997 "Sustainable Development." In Michael Redclift and Graham Woodgate, eds., *The International Handbook of Environmental Sociology*. Northampton, Mass.: Edward Elgar Publishing, Inc.
- Stretesky, Paul, and Michael J. Hogan 1998 "Environmental Justice: An Analysis of Superfund Sites in Florida." *Social Problems* 45(2):268-287.
- Szasz, Andrew 1994 *Ecopopulism: Toxic Waste and the Movement for Environmental Justice*. Minneapolis: University of Minnesota Press.
- , and Michael Meuser 1997 "Environmental Inequalities: Literature Review and Proposals for New Directions in Research and Theory." *Current Sociology* 45(3):99-120.
- Taylor, Dorceta 1992 "Can the Environmental Movement Attract and Maintain the Support of Minorities?" In Bunyan Bryant and Paul Mohai, eds., *Race and the Incidence of Environmental Hazards: A Time for Discourse*. San Francisco: Westview Press.
- Taylor, Dorceta E. 1989 "Blacks and the Environment: Toward an Explanation of the Concern and Action Gap Between Blacks and Whites." *Environment and Behavior* 21:175-205.
- United Church of Christ (Commission for Racial Justice) 1987 *Toxic Wastes and Race in the United States: A National Report on the Racial and Socioeconomic Characteristics of Communities with Hazardous Waste Sites*. New York: United Church of Christ.
- Walsh, Edward J. 1981 "Resource Mobilization and Citizen Protest in Communities Around Three Mile Island." *Social Problems* 29(1):1-21.
- , Rex Warland, and D. Clayton Smith 1993 "Backyards, NIMBYs, and Incinerator Sitings: Implications for Social Movement Theory." *Social Problems* 40(1):25-38.
- Weissman, Robert 1993 "Corporate Plundering of Third World Resources." In Richard Hofrichter, ed., *Toxic Struggles: The Theory and Practice of Environmental Justice*. Philadelphia: New Society Publishers.

World Commission on Environment and Development 1987 *Our Common Future*. Oxford: Oxford University Press.

Wright, Beverly 1995 "Environmental Equity Justice Centers: A Response to Inequity." In Bunyan Bryant, ed., *Environmental Justice: Issues, Policies, and Solutions*. Washington, D.C.: Island Press.

Zimmerman, Rae 1993 "Social Equity and Environmental Risk." *Risk Analysis* 13(6):649-666.

GLYNIS DANIELS

## ENVIRONMENTAL SOCIOLOGY

Environmental sociology is a relatively new area of inquiry that emerged largely in response to increased societal recognition of the seriousness of environmental problems. Many areas of sociology have similarly arisen as a result of societal attention to problematic conditions, including poverty and inequality, racial and gender discrimination, and crime and delinquency. Environmental sociology is unique, however, in that sociological attention to environmental problems had to overcome strong disciplinary traditions that discouraged giving attention to nonsocial conditions such as environmental quality. Consequently, the growth of sociological work on environmental issues has been accompanied by a critique and reassessment of core sociological assumptions and practices, with the result that environmental sociology has a somewhat ambivalent stance toward its parent discipline.

We begin with a brief examination of the nature and evolution of environmental problems, in order to clarify the kinds of issues that are of concern to environmental sociologists. Then we describe the emergence of societal attention to environmental problems, highlighting sociological work on environmental activism and related topics. Next we describe sociology's response to the increased salience of environmental problems, including the development of environmental sociology as an area of inquiry as well as its critique of mainstream sociology's neglect of environmental issues. Then we review some important emphases of the field, including analyses of the causes of environmental problems, examinations of the social impacts of these problems, and analyses of

solutions to such problems. We end with a brief overview of recent trends and debates in the field.

### SOCIETAL-ENVIRONMENTAL INTERACTIONS AND THE EVOLUTION OF ENVIRONMENTAL PROBLEMS

*Environmental sociology* is typically defined as the study of relations between human societies and their physical environments or, more simply, "societal-environmental interactions" (Dunlap and Catton 1979). Such interactions include the ways in which humans influence the environment as well as the ways in which environmental conditions (often modified by human action) influence human affairs, plus the manner in which such interactions are socially construed and acted upon. The relevance of these interactions to sociology stems from the fact that human populations depend upon the biophysical environment for survival, and this in turn necessitates a closer look at the functions that the environment serves for human beings.

**Three Functions of the Environment.** The biophysical environment serves many essential functions for human populations, as it does for all other species (Daily 1997), but three basic types can be singled out. First, the environment provides us with the resources that are necessary for life, ranging from air and water to food to materials needed for shelter, transportation, and the vast range of economic goods we produce. Human ecologists thus view the environment as providing the "sustenance base" for human societies, and we can also think of it as a "supply depot." Some resources, such as forests, are potentially renewable while others, like fossil fuels, are nonrenewable or finite. When we use resources faster than the environment can supply them, even if they are potentially renewable (such as clean water), we create resource shortages or scarcities (Catton 1980).

Second, in the process of consuming resources humans, like all species, produce "waste" products; indeed, humans produce a far greater quantity and variety of waste products than do other species. The environment must serve as a "sink" or "waste repository" for these wastes, either absorbing or recycling them into useful or at least harmless substances (as when trees absorb carbon dioxide and return oxygen to the air). When land

was sparsely populated and utilization of resources was minimal, this was seldom a problem. Modern and/or densely populated societies generate more waste than the environment can process, however, and the result is the various forms of “pollution” that are so prevalent worldwide.

Humans, like other species, must also have a place to exist, and the environment provides our home—where we live, work, play, travel, and spend our lives. In the most general case, the planet Earth provides the home for our species. Thus, the third function of the environment is to provide a “living space” or habitat for human populations. When too many people try to live in a given space, the result is overcrowding, a common occurrence in many urban areas (especially in poorer nations). Some analysts suggest that the entire planet is now overpopulated by human beings, although efforts to determine the number of people the Earth can support has proven to be difficult and contentious (Cohen 1995).

When humans overuse an environment’s ability to fulfill these three functions, “environmental problems” in the form of pollution, resource scarcities, and overcrowding and/or overpopulation are the result. However, not only must the environment serve all three functions for humans, but when a given environment is used for one function its ability to fulfill the other two is often impaired. Such conditions of functional competition often yield newer, more complex environmental problems.

Competition among environmental functions is especially obvious in conflicts between the living-space and waste-repository functions, as using an area for a waste site typically makes it unsuitable for living space. When an area is used as a garbage landfill or hazardous waste site, for example, people don’t even want to live near it, much less on it (Freudenburg 1997). Likewise, if hazardous materials escape from a waste repository and contaminate the soil, water, or air, the area can no longer serve as a supply depot for drinking water or for growing agricultural products. Finally, converting farmland or forests into housing subdivisions creates more living space for people, but it means that the land can no longer function as a supply depot for food or timber (or as habitat for wildlife).

**The Evolution of Environmental Problems.** Understanding these three functions played by the environment provides insight into the evolution of

environmental problems, or the problematic conditions created by human overuse of the environment. In the 1960s and early 1970s when awareness of environmental problems was growing rapidly in the United States, primary attention was given to air and water pollution and to litter—problems stemming from the environment’s inability to absorb human waste products—as well as to the importance of protecting areas of natural beauty. The “energy crisis” of 1973 highlighted the dependence of modern industrialized nations on fossil fuels and raised the specter of resource scarcity in general. The living-space function came to the forefront in the late 1970s when it was discovered that a neighborhood in Niagara Falls, New York, was built on an abandoned chemical waste site that had begun to leak toxic materials. Love Canal came to symbolize the growing problems of using an area as both waste repository and living space.

New environmental problems continually emerge, the result of humans trying to make incompatible uses of given environments. Global warming is an excellent example. It is primarily a consequence of a rapid increase in carbon dioxide in the Earth’s atmosphere produced by a wide range of human activities—especially burning fossil fuels (coal, gas, and oil), wood, and forest lands. This buildup of carbon dioxide (CO<sub>2</sub>) traps more of the Sun’s heat, thus raising the temperature of the Earth’s atmosphere. While global warming results from overuse of the Earth’s atmosphere as a waste site, the resulting warming may in turn produce changes that make our planet less suitable as a living space (not only for humans, but especially for other forms of life). A warmer climate may also affect the Earth’s ability to continue producing natural resources, especially food supplies (Stern et al. 1992).

These examples of how human activities are harming the ability of the environment to serve as our supply depot, living space, and waste repository involve focusing on specific aspects of particular environments (e.g., a given river’s ability to absorb wastes without becoming polluted). However, it is increasingly recognized that the health of entire ecosystems is being jeopardized as a result of growing human demands being placed on them. An *ecosystem* is an interacting set of living organisms (animals and plants) and their nonliving environment (air, land, water) that are bound together

by a flow of energy and nutrients (e.g., food chains); it can range in size from a small pond, to a large region such the Brazilian rainforest, to the entire biosphere—the Earth’s global ecosystem (Freese 1997). Technically, it is not “the environment” but “ecosystems” that serve the three functions for humans—and for all other living species.

Exceeding the capacity of a given ecosystem to fulfill one of the three functions may disrupt not only its ability to fulfill the other two but its ability to continue to function at all. As a recent U.S. Environmental Protection Agency (EPA) report notes, “Ecological systems like the atmosphere, oceans, and wetlands have a limited capacity for absorbing the environmental degradation caused by human activities. After that capacity is exceeded, it is only a matter of time before those ecosystems begin to deteriorate and human health and welfare begin to suffer” (EPA Science Advisory Board 1990, p. 17). Human overuse of ecosystems thus creates “ecological disruptions” that become “ecological problems” for humans. As more and more people require places to live, use resources, and produce wastes, it is likely that ecological problems will worsen and that new ones will continue to emerge.

The notion that human societies face “limits to growth” was originally based on the assumption that we would run out of natural resources such as oil, but nowadays it is recognized that the ability of ecosystems to fulfill any of the three necessary functions can be exceeded. Ozone depletion, for example, stems from exceeding the atmosphere’s limited ability to absorb chlorofluorocarbons (CFCs) and other pollutants. Thus, it is not the supply of natural resources per se, but the finite ability of the global ecosystem to provide us with resources, absorb our wastes, and still offer suitable living space (all of which, as we have seen, are interrelated) that constrains human societies. The emergence of problems such as ozone depletion, climate change, species extinction, and rainforest destruction are indications that modern societies may be taxing the limits of the global ecosystem (see, e.g., Daily 1997).

This brief sketch of the nature and evolution of environmental problems reveals the rich subject matter studied by environmental sociologists. While broadly construed as the study of relations

between human societies and their physical environments, the field focuses primary attention on the ways in which modern societies are altering their environments and the ways in which such alterations eventually create problematic conditions for our societies—as well as the ambiguities and controversies involved in assessing and responding to these alterations and impacts (Hannigan 1995). Despite improvements in specific environmental problems, such as urban air quality and the quality of many streams and lakes, in the last quarter of the twentieth century, newer and often more serious environmental problems—often affecting wider geographic areas—have continued to emerge as greater effort is made to monitor the quality of the environment. Consequently, it seems safe to assume that environmental sociology will have no shortage of subject matter in the foreseeable future.

#### **SOCIETAL RESPONSE TO ENVIRONMENTAL PROBLEMS**

Environmental concerns first attained considerable prominence in the United States with the rise of the progressive conservation movement in the late 1800s and early 1900s in response to reckless exploitation of the nation’s resources. This movement included both “utilitarians” like Gifford Pinchot, who sought to manage natural resources such as forests wisely to ensure their continued availability, and “preservationists” like John Muir, who sought to preserve areas of natural beauty for their own sake. Although these two factions eventually came into conflict, their joint efforts led to legislation establishing national parks and agencies such as the U.S. Forest Service while also spawning organizations such as the Sierra Club and National Audubon Society.

These “conservation” organizations and agencies continued to exist and grow periodically throughout the first half of the twentieth century, but they did not become highly visible until the 1950s and 1960s, when preservationist organizations such as the Sierra Club achieved renewed visibility by fighting for the protection of areas of natural beauty such as the Grand Canyon (which had been threatened by damming). The older preservationist concern with natural areas gradually coalesced with concerns about issues such as

pesticide contamination—publicized by Rachel Carson's *Silent Spring*—and air and water pollution and other urban-based problems (Taylor 1997). These newer “environmental” problems tended to be more complex in origin (often stemming from new technologies), and had delayed, complex, and difficult-to-detect effects that were consequential for human (as well as nonhuman) health and welfare. Encompassing pollution and loss of recreational and aesthetic resources, and ultimately the consequences of overusing all three functions of the environment, such problems were increasingly viewed as threats not only to “environmental quality” but to our “quality of life” (Dunlap and Mertig 1992).

By the late 1960s the older conservation movement had evolved into a modern “environmental movement,” as traditional conservation organizations joined with newer, multi-issue organizations, such as the Natural Resources Defense Council and Environmental Defense Fund, to pursue a wide range of environmental goals. The transformation from conservationism to environmentalism—symbolized by celebration of the first “Earth Day” in 1970—is reflected by the growth of an “environmental” discourse that largely supplanted the older conservation/preservation discourses (Brulle 1996) as well as by the explosive growth of both local and national environmental organizations concerned with a wide range of issues in the 1970s (McLaughlin and Khawaja 1999). With a reported 20 million participants, “E-Day” not only launched the contemporary environmental movement but mobilized a far greater base of support for environmentalism than had ever been achieved by the earlier conservation movement (Dunlap and Mertig 1992).

Environmental sociologists quickly tried to explain the emergence of the modern environmental movement. Besides highlighting the crucial roles played by older conservation organizations (such as the Sierra Club) in mobilizing public support, they pointed to the following factors: growth of scientific knowledge about problems such as pesticides and smog; intense media coverage devoted to incidents such as the 1969 Santa Barbara oil spill; rapid increase in outdoor recreation that brought more and more people into direct contact with threatened environments; widespread post-World War II affluence that stimulated increasing concern with quality of life over

more materialistic pursuits; and, finally, the general climate of social activism that grew out of the civil rights, anti-Vietnam War and student-power movements (see, e.g., Morrison et al. 1972). More recent analyses have emphasized the decline of class politics and materialist concerns and the subsequent emergence of new social movements devoted to quality-of-life goals as the broad socio-cultural context in which environmentalism developed (e.g., Buttel 1992).

The last three decades of the twentieth century have seen enormous changes in the nature of environmentalism, often chronicled by sociologists. The large national organizations, like the Sierra Club, constituting the core of the “environmental lobby” have remained strong, but their limited ability to produce effective results and the continual emergence of new issues that escape their purview have resulted in the emergence of numerous alternative strands of environmental activism. The result is that environmentalism is far more diverse than in the older days of the utilitarian-preservationist bifurcation (Dunlap and Mertig 1992). A major divide is that between the national organizations and the burgeoning number of local, grassroots groups that are typically concerned with hazardous conditions in their communities. A particularly potent form of grassroots activism has arisen in response to evidence of environmental racism, or the disproportionate location of hazardous facilities in minority communities, and the “environmental justice” movement spearheaded by people of color is the result (Taylor 1997). A variety of other splinter groups, including radical environmentalism (exemplified by the direct-action tactics pursued by Earth First!), deep ecology (a biocentric philosophy urging the equality of all forms of life), and ecofeminism (which links environmental degradation to the exploitation of women), along with the increasing internationalization of environmental organizations, make the contemporary environmental movement a highly complex entity (Brulle 1996).

Regardless of its historical causes, the establishment and continued existence of a viable environmental movement has ensured that environmental issues remain on the nation's policy agenda. Several pieces of landmark legislation aimed at protecting and improving environmental quality were passed in the late 1960s and early 1970s, including the National Environmental Policy Act



requiring environmental impact assessments and establishment of the Environmental Protection Agency. Despite some ups and downs, environmental quality has remained a major national goal for the last three decades of the twentieth century, with new concerns continually emerging—often in response to specific “crises” such as the *Exxon Valdez* oil spill in Alaska.

The environmental movement has been widely criticized for failing to halt environmental degradation in our nation (and the world). Nonetheless, compared to other social movements, environmentalism has clearly been one of the most influential movements of the last half of the twentieth century, ranking with the women’s movement, civil rights movement, and peace/antiwar movements as having changed the contours of contemporary life. Environmental concerns are now institutionalized throughout our society, not only in government laws and agencies, but in the form of environmental education (K–12) and college-level environmental studies, environmental reporters and news beats among major media, environmental affairs offices within major corporations, the growing involvement of mainstream religions with environmental problems, and a reasonably well funded and institutionalized discipline of environmental science. To these institutional indicators one can add cultural changes such as the emergence of an environmental discourse throughout society, including advertising; the ubiquity (if not effectiveness) of recycling and the gradual growth of green consumer behaviors; and the normative disapproval of outright assaults on environmental quality whether by governments or industry. Despite the intermittent emergence of open opposition to environmentalism—as represented by the current “wise-use” movement that seeks to lift governmental restrictions on use of natural resources (Switzer 1997)—the overall trends toward “environmentalization” validate claims that we are witnessing a “greening” of society at both the institutional and cultural levels (Buttel 1992).

#### **DISCIPLINARY RESPONSE: THE BIRTH OF ENVIRONMENTAL SOCIOLOGY**

Although there was scattered sociological attention to natural resource issues prior to the 1970s, environmental sociology developed in that decade

as sociology’s own response to the emergence of environmental problems. At first sociologists tended to pay more attention to societal response to environmental problems than to the problems themselves. As noted earlier, analyses of the environmental movement were popular, as were studies of public attitudes toward environmental issues. Prime topics included identifying the social sectors from which environmental activists were drawn and the social bases of pro-environmental attitudes among the general public (Dunlap and Catton 1979). Broader analyses of the ways in which “environment” was being constructed as a social problem, and the vital roles played by both activists and the media in this process, also received attention (Albrecht 1975). In addition, rural sociologists conducted a growing number of studies of natural resource agencies, while other sociologists examined environmental politics and policy making. In general, sociological work on environmental issues typically employed perspectives from the larger discipline to shed light on societal awareness of and response to environmental problems. In today’s parlance, these initial efforts largely involved analyses of aspects of the “social construction of environmental problems” and represented what was termed a “sociology of environmental issues” (Dunlap and Catton 1979).

As sociologists paid more attention to environmental issues, a few began to look beyond societal attention to environmental problems to the underlying relationships between modern, industrialized societies and the physical environments they inhabit. Concern with the causes of environmental pollution was supplemented by a focus on the social impacts of pollution and resource constraints. In some cases there was explicit attention to the reciprocal relationships between societies and their environments, or to the “ecosystem-dependence” of modern societies (Dunlap and Catton 1994). These concerns were bolstered by the 1973–1974 “energy crisis,” as the interrupted flow of oil from Arab nations generated dramatic and widespread impacts and vividly demonstrated the vulnerability of modern industrial societies to an interruption of their fossil fuel supplies and—by extension—to natural resources in general (Rosa et al. 1988). Sociologists were quick to respond with numerous studies of the impacts, particularly the inequitable distribution of negative ones, of energy shortages (Schnaiberg 1975).

Sociological interest in the impacts of energy and other resource scarcities accelerated the emergence of environmental sociology as a distinct area of inquiry by heightening awareness that “environment” was more than just another social problem, and that environmental conditions could indeed have societal consequences. Studies of the societal impacts of energy shortages thus facilitated a transition from the early “sociology of environmental issues” to a self-conscious “environmental sociology” focused explicitly on societal–environmental relations. In retrospect, it is apparent that this concern also contributed to a rather one-sided view of such interactions, however, as the effects of resource constraints on society received far more emphasis than did the impacts of society on the environment (something that has been rectified in more recent research on the causes of environmental degradation).

The nascent environmental sociology of the 1970s was quickly institutionalized via formation of interest groups within the national sociological associations. These groups provided an organizational base for the emergence of environmental sociology as a thriving area of specialization and attracted scholars interested in all aspects of the physical environment—from environmental activism to energy and other natural resources, natural hazards and disasters, social impact assessment, and housing and the built environment (Dunlap and Catton 1983). The late 1970s was a vibrant era of growth for American environmental sociology, but momentum proved difficult to sustain during the 1980s, as the Reagan era was a troublesome period for the field and social science more generally. Ironically, however, sociological interest in environmental issues was beginning to spread internationally, and by the late 1980s and the 1990s environmental sociology was not only reinvigorated in the United States but was being institutionalized in countries around the world and within the International Sociological Association (Dunlap and Catton 1994).

The resurgence of environmental sociology in the United States and its emergence internationally benefited from key societal events. Publicity surrounding Love Canal and other local environmental hazards stimulated interest in the impacts of such hazards on local communities, while major accidents at Three Mile Island, Bophal (India), and

Chernobyl dramatized the importance of technological hazards and helped generate sociological interest in the environmental and technological risks facing modern societies (Short 1984). More recently, growing awareness of global environmental problems such as ozone depletion, global climate change, and tropical deforestation have served to enhance sociological interest in environmental problems—particularly at the global level (Dunlap and Catton 1994).

### ENVIRONMENTAL SOCIOLOGY AND THE LARGER DISCIPLINE

As previously noted, early sociological work on environmental issues typically involved application of standard sociological perspectives drawn from social movements, social psychology, social problems, and so forth to empirical work focusing on societal response to environmental issues. Efforts to theorize about environmental matters were rare, and they tended to involve demonstrations of the utility of established theoretical perspectives, such as Parsonian theory, for viewing environmental issues rather than asking whether such perspectives were adequate for understanding the relations between modern societies and their biophysical environments (Klausner 1971). Since sociologists interested in theorizing about the relations between modern societies and their environments found little guidance from the larger discipline, they drew heavily upon other disciplines, such as general ecology (Catton 1976), or combined ecological and sociological insights in order to develop new theoretical perspectives (Schnaiberg 1975).

Unlike the larger society, in the 1970s mainstream sociology was remarkably oblivious to the relevance of environmental matters. This disciplinary blindness stemmed from a long period of neglect of such matters stimulated by both societal developments and disciplinary traditions. The Durkheimian emphasis on explaining social phenomena only in terms of other “social facts,” plus an aversion to earlier excesses of biological and geological “determinisms,” had led sociologists to ignore the physical world. These disciplinary traditions were further strengthened by sociology’s emergence during an era of unprecedented growth and prosperity, when limits to resource abundance and technological progress were unimaginable, and increased urbanization, which reduced

direct contact with the natural environment. With modern, industrialized societies appearing to be increasingly independent of the biophysical world, sociology came to assume that the exceptional features of *Homo sapiens*—language, technology, science, and culture more generally—made these societies “exempt” from the constraints of nature. Thus, the core task of sociology was to examine the uniquely social determinants of contemporary human life (Dunlap and Catton 1979, 1983). In short, mainstream sociology offered infertile ground for planting sustained interest in the relations between societies and their biophysical environments.

It is not surprising, therefore, that efforts to establish environmental sociology as a legitimate and important area of inquiry included criticism of the larger discipline’s blindness to environmental matters. Indeed, efforts to define and codify the field of environmental sociology were accompanied by explication and critique of the “human exemptionalism paradigm” (HEP) on which contemporary sociology was premised. While not denying that human beings are obviously an exceptional species, environmental sociologists argued that our special skills and capabilities nonetheless failed to exempt us from the constraints of the natural environment. Consequently, it was argued that the HEP needed to be replaced by a more ecologically sound perspective, a new ecological paradigm (NEP), that acknowledges the ecosystem-dependence of human societies (Catton and Dunlap 1978, 1980). It was further argued that much of environmental sociology, particularly examinations of the relations between social and environmental factors (as opposed to analyses of the social construction of environmental issues), entailed at least implicit rejection of the HEP with its assumed irrelevance of nonsocial phenomena to modern societies (Dunlap and Catton 1979, 1983).

The call for revision of mainstream sociology’s dominant paradigm, and particularly the urging of adoption of an ecological perspective, has been a controversial feature of environmental sociology. While regarded as a core element of the field’s commitment to ensuring that the material bases of modern societies are no longer neglected by sociology, the argument has been criticized for deflecting efforts to utilize classical and mainstream theoretical perspectives in environmental sociology (Buttel 1996, 1997). Fortunately, debate about the

need for an ecological perspective versus the relevance of mainstream sociological theories has taken a new turn in recent years, as several environmental sociologists have independently begun to develop ecologically informed versions of classical theoretical perspectives. Efforts to develop “green” versions of Durkheimian, Weberian, and especially Marxian macro-sociologies as well as micro-level perspectives, such as symbolic interactionism, bear the fruit of integrating an ecological paradigm with classical theoretical traditions (see Foster 1999 and references in Dunlap 1997).

Increasing awareness of the societal significance of ecological conditions has not only stimulated efforts to develop greener sociological theories but also opened the floodgates of empirical research on societal–environmental relations. Studies such as those conducted by Freudenburg and Gramling (1994) on oil development in coastal waters, which convincingly show how development depends on *both* environmental conditions and social forces, and by Couch and Kroll-Smith (1985) on community hazards, which demonstrates the differing impacts of natural versus human-made disasters, clearly violate the disciplinary tradition of ignoring all but the social causes of social facts (as do many of the studies to be reviewed later).

While the empirical thrust of environmental sociology thus represents at least implicit rejection of mainstream sociology’s “exemptionalist” orientation by continually demonstrating the relevance of environmental factors in modern, industrialized societies, the situation regarding adoption of an ecological paradigm or perspective is less clear. Some environmental sociologists follow Catton’s (1980) lead in applying ecological theory and concepts to human societies (e.g., Fischer-Kowalski 1997), and others employ an ecological perspective as an “orienting strategy” that encourages them to raise questions about issues such as the long-term sustainability of current consumption patterns in the wealthy nations (Redclift 1996). However, some environmental sociologists express caution regarding the utility of ecological *theory* as a guiding framework for environmental sociology (Buttel 1997) or disavow its utility altogether (Macnaghten and Urry 1998). These differing orientations stem from the inherent ambiguities involved in applying concepts and findings from general ecology to human societies (Freese 1997)

as well as differences of opinion about the roles of paradigms versus theories in empirical inquiry.

### SOME CURRENT RESEARCH EMPHASES

Over the last three decades of the twentieth century, sociological research on environmental issues has focused on a wide range of topics, and we cannot begin to do justice to the full body of work by environmental sociologists. Fortunately, several existing reviews and compilations provide fairly comprehensive overviews of this research, covering the 1970s (Dunlap and Catton 1979), 1980s (Buttel 1987; Rosa et al. 1988), and 1990s (Buttel 1996; Dunlap and Catton 1994; Redclift and Woodgate 1997). Consequently, we limit our attention to sociological work on three particularly important theoretical and policy-relevant topics: the causes of environmental problems, the impacts of such problems, and the solutions to these problems.

**Causes of Environmental Problems.** Given that environmental sociology arose largely in response to increased recognition of environmental problems, it is not surprising that a good deal of work in the area has been devoted to trying to explain the origins of environmental degradation. Much of the early work, however, was devoted more to analyses and critiques of the rather simplistic views of the causes of environmental degradation that predominated in the literature rather than original studies. The need for such analyses stemmed from the fact that the predominant conceptions of the origins of environmental problems tended to emphasize the importance of single factors such as population growth (emphasized by Paul Ehrlich) or technological development (stressed by Barry Commoner), rather than recognizing the multiplicity of factors involved, and also ignored or simplified the distinctively social causes of environmental degradation. In this context, environmental sociologists tended to explicate the competing range of explanations (Dunlap and Catton 1983) and to criticize the most widely accepted ones for their shortcomings (Schnaiberg 1980).

The most powerful sociological critique of common conceptions of the origins of environmental problems in general, and those by biologists such as Ehrlich and Commoner in particular, was provided by Schnaiberg (1980). Schnaiberg

criticized Ehrlich's view by noting the enormous variation in environmental impact between populations of rich and poor nations as well as between the wealthy and poor sectors within individual nations, emphasized that population growth is interrelated with factors such as poverty—which induces poor people to have more children for work-force and security reasons. Similarly, Commoner's perspective was criticized for viewing technology as an autonomous force, ignoring the degree to which technological developments are driven by political and especially economic forces—particularly the need for profit and capital accumulation.

Besides demonstrating the oversimplification involved in attributing environmental degradation to either population or technology, Schnaiberg also critiqued a third factor widely mentioned as a cause—the wasteful lifestyles of consumers. In particular, Schnaiberg distinguished between the production and consumption spheres of society, arguing that the former is the more crucial contributor to environmental degradation. Attributing environmental degradation to the affluence of consumers ignores the fact that decisions made in the production realm (e.g., as to what types of transportation will be available to consumers) are far more significant than are the purchasing behaviors of individual consumers. Consequently, Schnaiberg emphasized the “treadmill of production,” or the inherent need of market-based economic systems to grow and the powerful coalition of capital, state, and labor supporting such growth, as the most fundamental contributor to environmental degradation.

While Schnaiberg's analysis, which he has continued to update and refine (see, e.g., Schnaiberg and Gould 1994), has become highly influential within environmental sociology (Buttel 1987, 1997), it has proven difficult to translate into concrete empirical research beyond local case studies of organized opposition to treadmill processes (Gould et al. 1996). Consequently, a new generation of sociological analysts, while cognizant of Schnaiberg's insights, have adopted a broader framework for investigating the causes of crucial environmental problems, particularly pressing global problems.

Ironically, one line of this new work has involved revisiting the Ehrlich–Commoner debate over the relative importance of population and

technological factors in generating environmental degradation. As their debate progressed, both sides realized that they could not totally ignore the other's preferred cause, or more distinctively social factors, and the debate became encapsulated in differing interpretations of a simple formulation known as the "IPAT" equation. Both Ehrlich and Commoner came to agree that environmental impact = population  $\times$  affluence  $\times$  technology, although debate continued over which factor on the right side of the equation had the most impact on environmental degradation (for more on this debate, see Dietz and Rosa 1994; Dunlap et al. 1994).

In recent years environmental sociologists have begun to reassess the IPAT model's utility, particularly as a means of examining the crucial causal forces generating global-level environmental problems such as tropical deforestation and global climate change. Taking into account earlier critiques of the IPAT model, Dietz and Rosa (1994; see also Rosa and Dietz 1998) have proposed a major revision that they label "STIRPAT," for "stochastic impacts by regression on population, affluence and technology." Whereas IPAT is an accounting equation that assumes direct proportionality between impact and each of the other three factors, such that a 10-percent increase in population (or technology or affluence) is assumed to produce a 10-percent increase in impact, STIRPAT treats such linkages as hypotheses to be tested. In addition, it allows both for decomposition of the individual factors (e.g., population size versus growth rate) in the model and the consideration of cultural, institutional, and political factors as additional sources of environmental impact.

While the STIRPAT model can be applied to any environmental impact, the initial application has been to global climate change, where the basic model was used to estimate CO<sub>2</sub> emissions (Dietz and Rosa 1997). Two interesting results emerged: First, diseconomies of scale apparently exist at the largest population sizes, as countries with the largest populations (e.g., China and India) have a disproportionate impact on CO<sub>2</sub> loads. Second, the results also replicated recurrent findings by economists that suggest the existence of an environmental Kuznets curve. Economists argue in these studies that as affluence increases, environmental impact per unit of affluence decreases, producing an inverted U curve, or a Kuznets curve, named as the relationship between development

and inequality posited by Simon Kuznets (see Dietz and Rosa 1997 for citations).

A related sociological study, involving a longitudinal analysis of national CO<sub>2</sub> emissions, challenges the economists' faith that economic development will lessen environmental impacts because of the Kuznets curve. Roberts and Grimes (1997) conclude that since the energy crisis of the 1970s, affluent nations have become more carbon-efficient, producing more gross national product per unit of CO<sub>2</sub> emissions, but that the carbon efficiency of middle-income nations has gone down slightly and that of the poor nations has dropped substantially. Consequently, the poorest nations are locked into a pattern of high and even increasing environmental impact per unit of affluence, while wealthy nations may indeed follow the patterns proposed by development theorists. If true, this challenge to development theory suggests that the continued development of poor but populous nations like China and India will lead to increasing levels of CO<sub>2</sub> and other pollutants rather than the pattern observed in wealthy nations.

Tropical deforestation is another global-level problem that has attracted increasing attention by environmental sociologists seeking to understand its causes. A series of studies by Rudel and colleagues (see Rudel and Roper 1997) employ a variety of theoretical models encompassing population, technology, and affluence (the components of the IPAT and STIRPAT models) as well as a wide range of both national and international social-structural variables such as level of inequality, urbanization, and international debt. These studies typically find a significant impact of population growth in general, and patterns of rural versus urban growth in particular, on deforestation rates, but these effects are moderated by a wide range of environmental conditions and social-structural factors. Importantly, but not surprisingly, Rudel finds that deforestation is partly a function of the size of nations' forests and the topography of forest land as well as political and economic decisions to invest capital in forestry, levels of economic development, and international indebtedness.

Sociological studies of global-level problems such as deforestation and emissions of greenhouse gases, both of which directly contribute to global climate change, are yielding important findings as well as conceptual and methodological strategies

for developing a better understanding of the driving forces producing global environmental change and other environmental problems (Rosa and Dietz 1998). As such, they represent an important supplement to natural-science research programs on these topics.

**Impacts of Environmental Problems.** As noted earlier, environmental sociology was just emerging at the time of the 1973–1974 energy crisis, and it is not surprising that a good deal of effort was made to identify real as well as potential social impacts of energy and other natural resources in this early period of the field. While diverse impacts, from regional migration to consumer lifestyles, were investigated, heavy emphasis was placed on investigating the “equity” impacts of both energy shortages and the policies designed to ameliorate them (Rosa et al. 1988). A general finding was that both the problems and policies often had regressive impacts, with the lower socioeconomic strata bearing a disproportionate cost due, for example, to rising energy costs (Schnaiberg 1975).

Equity has been a persistent concern in environmental sociology, and researchers gradually shifted their attention to the distribution of exposure to environmental hazards (ranging from air and water pollution to hazardous wastes). Again, a persistent finding has been that exposure to environmental hazards is generally negatively correlated with socioeconomic status. A growing number of studies have also found that minority populations are disproportionately exposed to environmental hazards, in part because of their lower-than-average socioeconomic levels but perhaps also because of conscious decisions to locate hazardous sites in minority communities. Such findings, which a few recent studies have challenged, have led to charges of “environmental racism” and efforts to achieve “environmental justice.” At a broader level, international equity is attracting the attention of environmental sociologists, who are investigating the export of polluting industries from wealthy to poor nations, the disproportionate contribution of wealthy nations to many global-level problems, and the consequent hurdles these phenomena pose for international cooperation to solve environmental problems (Redclift and Sage 1998).

Sociologists have not limited themselves to investigating the equity impacts of environmental

problems, and studies of communities exposed to technological or human-made hazards (such as Love Canal) offer particularly rich portrayals of the diverse impacts caused by discovery of community hazards. Whereas natural hazards—such as floods, hurricanes, and earthquakes—have been found to result in a therapeutic response in which communities unite in efforts to help victims, repair damage, and reestablish life as it was before the disaster struck, technological disasters have been found to have very different impacts (Freudenburg 1997). Although a putative hazard often appears obvious to those who feel affected by it, the ambiguities involved in detecting and assessing such hazards often generate a pattern of intense community conflict. Unlike those affected by natural hazards, the “victims” often find themselves at odds not only with public officials but also with other residents who fail to acknowledge the seriousness of the hazard (for fear of economic loss in terms of property values, jobs, etc.). In many cases, such conflicts have resulted in a long-term erosion of community life as well as exacerbation of the victims’ personal traumas stemming from their exposure to the hazards (Couch and Kroll-Smith 1985).

**Solutions to Environmental Problems.** As was true for the causes of environmental problems, early work by environmental sociologists interested in solutions to these problems often involved explications and critiques of predominant approaches. Early on Heberlein (1974) noted the predilection of the United States for solving environmental problems via a “technological fix,” or developing and applying new technologies to solve problems such as air and water pollution. Understandably popular in a nation with a history of technological progress, such a solution is appealing because it avoids mandating behavioral and institutional change. Unfortunately, solving problems with new technologies sometimes creates even more problems, as illustrated by attempts to solve energy shortages with nuclear power. Consequently, as the seriousness and pervasiveness of environmental problems became more obvious, attention was given to a variety of “social fixes,” or efforts to change individual and institutional behaviors.

Expanding on Heberlein’s analysis, other sociologists (e.g., Dunlap et al. 1994) have identified three broad types of social fixes, or implicit policy

types: (1) the cognitive (or knowledge) fix, which assumes that information and persuasion will suffice to produce the necessary changes in behavior, illustrated by campaigns encouraging energy conservation and recycling; (2) the structural fix, which relies on laws and regulations that mandate behavioral change, reflected in highway speed limits or enforced water conservation; and (3) the intermediary behavioral fix, which employs incentives and disincentives to encourage changes in behavior, as illustrated by pollution taxes (penalties) and tax credits (rewards) for installing pollution-abatement technology (see Gardner and Stern 1996 for a more refined typology of policy approaches and detailed examples of each).

Environmental sociologists, in conjunction with other behavioral scientists, have conducted a range of studies that bear on the efficacy of these differing strategies for solving environmental problems, ranging from field experiments to test the effectiveness of information campaigns in inducing energy and water conservation to evaluations of alternative strategies for generating participation in recycling programs (see Gardner and Stern 1996 for a good summary). A noteworthy sociological study was Derksen and Gartrell's (1993) investigation of recycling in Edmonton, Alberta, which found that individuals' level of environmental concern (and, by implication, knowledge about the importance of recycling) was not as important in predicting recycling behavior as was ready access to a curbside recycling program. While sociologists have conducted numerous field experiments and evaluations of community environmental programs, typically investigating the efficacy of one or more of the above-noted "fixes," they have generally left examinations of national and international environmental policy making to political scientists and economists. However, sociologists have begun to pay attention to efforts to negotiate international agreements to achieve reduction of greenhouse gases (Redclift and Sage, 1998), and we expect more sociological work along these lines.

#### CURRENT TRENDS AND DEBATES

As the foregoing illustrates, environmental sociology not only emerged in response to societal attention to environmental problems but has focused much of its energy on understanding these problems, especially their causes, impacts, and

solutions. The field has proved to be more than a passing fad, becoming well institutionalized and also increasingly internationalized. But in the process, fundamental assumptions that once served to unify the field—agreement over the reality of environmental degradation; diagnoses of such degradation as inherent to modern, industrialized societies; and the sense that mainstream sociology was largely blind to the significance of environmental matters—have become matters of debate (Buttel 1996, 1997).

The emergence of environmental problems provided the *raison d'être* for environmental sociology, and the seriousness of such problems was seldom challenged. While environmental sociologists from the outset paid attention to ways in which claims about environmental conditions are socially constructed and become the subject of societal conflict (e.g., Albrecht 1975), such efforts seldom questioned the objective existence of environmental problems. In recent years, however, environmental sociology has felt the effects of the larger discipline's turn toward more cultural and interpretative orientations. A growing number of scholars, particularly in Europe, have not only highlighted the contested nature of claims about environmental problems but—in the postmodern tradition—concluded that there is no reason for privileging the claims of any parties to these debates, including those of environmental scientists as well as activists (Macnaghten and Urry 1998).

Such work has led to the emergence of a strong constructivist/interpretative orientation in environmental sociology that challenges the objectivist/realist perspective that has traditionally been dominant. Whereas the realist orientation assumes that the environment is a biophysical entity existing independent of humans, thereby providing the setting for study of human-environment interactions as the core of environmental sociology, the constructivist orientation leads its adherents to adopt an agnostic view of such interactions, preferring instead to examine knowledge claims—and the social forces they reflect—about these interactions (Macnaghten and Urry 1998). These competing perspectives can be readily seen in sociological work on global environmental change (GEC), where those in the realist camp have sought to complement natural-science research by examining, for example, societal processes leading to tropical deforestation

and greenhouse gas emissions (noted previously), while constructivists have highlighted the uncertainties in scientific evidence for GEC and the social, political, and historical forces that have made GEC a central topic of scientific and policy-making interest (see Rosa and Dietz 1998). These differing orientations have led to debate among environmental sociologists, with realists pointing to shortcomings of the constructivist approach (Dickens 1996; Murphy 1997) and constructivists demonstrating the utility of their perspective (Hannigan 1995). Fortunately promising syntheses of constructivist and realist perspectives are beginning to emerge (Rosa 1998).

Another source of debate is the inevitability of continued environmental degradation, particularly on the part of advanced, industrialized nations. Whereas environmental sociologists have traditionally seen the drive toward capital accumulation inherent in such societies as making environmental degradation inevitable (as epitomized by Schnaiberg's "treadmill of production" argument), European scholars have increasingly suggested that this may not be the case. Obvious successes in environmental amelioration within advanced European nations have led them to build upon economic models of "industrial ecology," which suggest that modernization of industrial processes can permit production with ever-decreasing levels of material input and pollution output, heralding a new era of "ecological modernization" (Spaargaren and Mol 1992).

This perspective not only adopts a more sanguine image of the future of industrialized societies but, as Buttel (1996, 1997) notes, involves a shift in focus for environmental sociology: from a preoccupation with the origins of environmental degradation to efforts to explain the institutionalization of environmental amelioration (via technological innovation, policy incentives, pressures from citizens' groups, etc.). While representing an important complement to traditional perspectives in the field and building on larger sociological debates about the future of "modernity," ecological modernization is vulnerable to criticism. First, its development in northern Europe leads to concerns that ecological modernization is not applicable to less wealthy and technologically disadvantaged nations. Second, although nations such as the Netherlands have made considerable advances

in protecting their own environments, their import of natural resources and export of pollution creates a large "ecological footprint" well beyond their borders (Wackernagel and Rees 1996). Finally, even if continual progress is made in creating cleaner, more efficient production processes, these gains may be offset by continued economic growth and consumption and consequent increased demand for materials and energy (Bunker 1996).

The trends toward adoption of more constructivist/interpretive frameworks and models of ecological modernization are related to a third trend in environmental sociology—the ongoing reassessment of its relationship to the larger discipline. As noted earlier, the emergence of environmental sociology was marked by criticism of mainstream sociology's neglect of the ecosystem-dependence of modern, industrialized societies and consequent inattention to the challenge posed by environmental problems. This critical orientation led many environmental sociologists to look to other disciplines (such as ecology and environmental science) for guidance and probably contributed to a somewhat insular perspective vis-à-vis mainstream sociology.

But in the 1990s environmental problems, particularly global-level threats like climate change, have caught the attention of growing numbers of eminent sociologists, such as Giddens (1990), who have recognized that such problems cannot be ignored in analyses of the future course of industrial societies. Greater interaction between environmental and mainstream sociology has resulted, and the penetration of currently fashionable perspectives, such as cultural/interpretive frameworks and theories of the nature of late modernity into environmental sociology has followed—as exemplified by the popularity of the constructivist and ecological modernization perspectives. This is resulting in considerable debate and self-reflection among environmental sociologists, something that a maturing field can afford. Hopefully, environmental sociology will emerge with renewed relevance to the larger discipline of sociology as well as continued relevance to societal efforts to ensure an ecologically sustainable future for humankind.

(SEE ALSO: *Environmental Equity*)



## REFERENCES

- Albrecht, Stan L. 1975 "The Environment as a Social Problem." In A. L. Mauss, ed., *Social Problems as Social Movements*. Philadelphia: J. P. Lippincott.
- Bulle, Robert J. 1996 "Environmental Discourse and Social Movement Organizations." *Sociological Inquiry* 66:58-83.
- Bunker, Stephen G. 1996 "Raw Material and the Global Economy: Oversights and Distortions in Industrial Ecology." *Society and Natural Resources* 9:419-430.
- Buttel, Frederick H. 1987 "New Directions in Environmental Sociology." *Annual Review of Sociology* 13:465-488.
- 1992 "Environmentalization: Origins, Processes, and Implications for Rural Social Change." *Rural Sociology* 57:1-27.
- 1996 "Environmental and Resource Sociology: Theoretical Issues and Opportunities for Synthesis." *Rural Sociology* 61:56-76.
- 1997 "Social Institutions and Environmental Change." In M. Redclift and G. Woodgate, eds., *The International Handbook of Environmental Sociology*. Cheltenham, U.K.: Edward Elgar.
- Catton, William R., Jr. 1976 "Why the Future Isn't What It Used to Be." *Social Science Quarterly* 57:276-291.
- 1980 *Overshoot: The Ecological Basis of Revolutionary Change*. Urbana: University of Illinois Press.
- and Riley E. Dunlap 1978 "Environmental Sociology: A New Paradigm." *The American Sociologist* 13:41-49.
- , and Riley E. Dunlap 1980 "A New Ecological Paradigm for Post-Exuberant Sociology." *American Behavioral Scientist* 24:15-47.
- Cohen, Joel E. 1995 *How Many People Can the Earth Support?* New York: Norton.
- Couch, Stephen R., and J. Stephen Kroll-Smith 1985 "The Chronic Technical Disaster: Toward a Social Scientific Perspective." *Social Science Quarterly* 66:564-575.
- Daily, Gretchen C., ed. 1997. *Nature's Services: Social Dependence on Natural Ecosystems*. Washington, D.C.: Island Press.
- Derksen, Linda, and John Gartrell 1993 "The Social Context of Recycling." *American Sociological Review* 58:434-442.
- Dickens, Peter 1996 *Reconstructing Nature: Alienation, Emancipation and the Division of Labour*. London: Routledge.
- Dietz, Thomas, and Eugene A. Rosa 1994 "Rethinking the Environmental Impacts of Population, Affluence and Technology." *Human Ecology Review* 1:277-300.
- 1997. "Effects of Population and Affluence on CO<sub>2</sub> Emissions." *Proceedings of the National Academy of Sciences* 94:175-179.
- Dunlap, Riley E. 1997 "The Evolution of Environmental Sociology." In M. Redclift and G. Woodgate, eds., *The International Handbook of Environmental Sociology*. Cheltenham, U.K.: Edward Elgar.
- , and William R. Catton, Jr. 1979 "Environmental Sociology." *Annual Review of Sociology* 5:243-273.
- , and William R. Catton, Jr. 1983 "What Environmental Sociologists Have in Common (Whether Concerned with 'Built' or 'Natural' Environments)." *Sociological Inquiry* 53:113-135.
- , and William R. Catton, Jr. 1994. "Struggling with Human Exemptionalism: The Rise, Decline and Revitalization of Environmental Sociology." *The American Sociologist* 25:5-30.
- , and Angela G. Mertig, eds. 1992 *American Environmentalism: The U.S. Environmental Movement: 1970-1990*. Bristol, Pa.: Taylor and Francis.
- , Loren A. Lutzenhiser, and Eugene A. Rosa 1994 "Understanding Environmental Problems: A Sociological Perspective." In B. Burgenmeier, ed., *Economy, Environment, and Technology*. Armonk, N.Y.: M.E. Sharpe.
- EPA Science Advisory Board 1990 *Reducing Risk: Setting Priorities and Strategies for Environmental Protection*. Washington, D.C.: U.S. Environmental Protection Agency.
- Fischer-Kowalski, Marina 1997 "Society's Metabolism." In M. Redclift and G. Woodgate, eds., *The International Handbook of Environmental Sociology*. Cheltenham, U.K.: Edward Elgar.
- Foster, John Bellamy 1999 "Marx's Theory of Social-Ecological Metabolism and the Critique of Capitalist Society: Classical Foundations for Environmental Sociology." *American Journal of Sociology*: 105:366-405.
- Freese, Lee 1997 *Environmental Connections*. Supplement I (Part B) to *Advances in Human Ecology*. Greenwich, Conn.: JAI Press.
- Freudenburg, William R. 1997 "Contamination, Corrosion and the Social Order: An Overview." *Current Sociology* 45:41-57.
- , and Robert Gramling 1994 *Oil in Troubled Waters: Perceptions, Politics, and the Battle over Offshore Drilling*. Albany: State University of New York Press.
- Gardner, Gerald T., and Paul C. Stern 1996 *Environmental Problems and Human Behavior*. Boston: Allyn and Bacon.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Stanford, Calif. Stanford University Press.

- Gould, Kenneth A., Allan Schnaiberg, and Adam S. Weinberg 1996 *Local Environmental Struggles: Citizen Activism in the Treadmill of Production*. New York: Cambridge University Press.
- Hannigan, John A. 1995 *Environmental Sociology: A Social Constructionist Perspective*. London: Routledge.
- Heberlein, Thomas 1974 "The Three Fixes: Technological, Cognitive and Structural." In D. R. Field, J. C. Barron, and B. F. Long, eds., *Water and Community Development: Social and Economic Perspectives*. Ann Arbor, Mich.: Ann Arbor Science.
- Klausner, Samuel A. 1971 *On Man in His Environment*. San Francisco, Jossey-Bass.
- Macnaghten, Phil, and John Urry 1998 *Contested Natures*. London: Sage Publications.
- McLaughlin, and Marwan Khawaja 1999 "The Organizational Dynamics of the U.S. Environmental Movement." *Rural Sociology*: In press.
- Morrison, Denton E., Kenneth E. Hornback, and W. Keith Warner 1972 "The Environmental Movement: Some Preliminary Observations and Predictions." In W. R. Burch, Jr., N. H. Cheek, Jr., and L. Taylor, eds., *Social Behavior, Natural Resources, and the Environment*. New York: Harper and Row.
- Murphy, Raymond 1997 *Sociology and Nature: Social Action in Context*. Boulder, Colo.: Westview Press.
- Redclift, Michael 1996 *Wasted: Counting the Costs of Global Consumption*. London: Earthscan.
- , and Colin Sage 1998 "Global Environmental Change and Global Inequality: North/South Perspectives." *International Sociology* 13: 499–516.
- , and Graham Woodgate, eds. 1997 *The International Handbook of Environmental Sociology*. Cheltenham, U.K.: Edward Elgar.
- Roberts, J. Timmons, and Peter E. Grimes 1997 "Carbon Intensity and Economic Development 1962–1991: A Brief Exploration of the Environmental Kuznets Curve." *World Development* 25:191–198.
- Rosa, Eugene A. 1998 "Metatheoretical Foundations for Post-Normal Risk." *Risk Analysis* 1:15–44.
- , and Thomas Dietz 1998 "Climate Change and Society: Speculation, Construction, and Scientific Investigation." *International Sociology* 13:421–455.
- , Gary E. Machlis, and Kenneth M. Keating 1988 "Energy and Society." *Annual Review of Sociology* 14:149–172.
- Rudel, Thomas K., and Jill Roper 1997 "The Paths to Rain Forest Destruction: Crossnational Patterns of Tropical Deforestation, 1975–90" *World Development* 25:53–65.
- Schnaiberg, Allan 1975 "Social Syntheses of the Social–Environmental Dialectic: The Role of Distributional Impacts." *Social Science Quarterly* 56:5–20.
- 1980 *The Environment: From Surplus to Scarcity*. New York: Oxford University Press.
- , and Kenneth Alan Gould 1994 *Environment and Society: The Enduring Conflict*. New York: St. Martin's Press.
- Short, James F., Jr. 1984 "The Social Fabric at Risk: Toward the Social Transformation of Risk Analysis." *American Sociological Review* 49:711–725.
- Spaargaren, Gert, and Arthur P.J. Mol 1992 "Sociology, Environment, and Modernity: Ecological Modernization as a Theory of Social Change." *Society and Natural Resources* 5:323–344.
- Stern, Paul C., Oran R. Young, and Daniel Druckman, eds. 1992 *Global Environmental Change: Understanding the Human Dimensions*. Washington, D.C.: National Academy Press.
- Switzer, Jacqueline Vaughn 1997 *Green Backlash*. Boulder, Colo.: Lynne Rienner.
- Taylor, Dorceta E. 1997 "American Environmentalism: The Role of Race, Class and Gender in Shaping Activism, 1820–1995." *Race, Gender and Class* 5:16–62.
- Wackernagel, Mathis, and William Rees 1996 *Our Ecological Footprint: Reducing Human Impact on the Earth*. Philadelphia: New Society Publishers.

RILEY E. DUNLAP  
EUGENE A. ROSA

## EPIDEMIOLOGY

*Epidemiology* is the study of the distribution of disease and its determinants in human populations. Epidemiology usually takes place in an applied public health context. It focuses on the occurrence of disease by time, place, and person and seeks to identify and control outbreaks of disease through identification of etiological factors. Its approach is to identify associated risk factors and then work back to causes.

Historically, the focus of epidemiology was on large outbreaks, usually of infectious disease. The substance and methods of epidemiology have been applied to most forms of acute and chronic disease and many other physical and mental health conditions. Along with the broadening of the subject matter of epidemiology has come more focus on its methods. Along with the specialization of

epidemiological methods has come the professional identification of persons as epidemiologists. In this entry, we address the origins, methodology, current topics, and professional issues related to epidemiology.

### PROFESSIONAL ROLE

Epidemiology is the professional identification of an increasing number of persons who have received specialized training in departments of epidemiology in schools of public health. Many complete approved courses of study at the master's or doctoral level leading to an M.P.H., Dr. P.H., or Ph.D. with epidemiology as an area of concentration. Some schools offer the M.P.H. to physicians after an abbreviated course of study. Like most other developing professions, graduate epidemiologists are protective of their professional identification and would seek to differentiate their professional credentials from those who have found their ways to epidemiological roles from other educational backgrounds. The problem in this developing professionalism is how to define a unique intellectual content or theory of epidemiology, when the applied nature of the discipline demands that the latest theories and methods of investigating disease be incorporated into any ongoing investigation.

### ORIGINS

Langmuir (Roeché 1967, p. xiii) traces the origins of epidemiology as far back as Hypocrites' report of an outbreak of mumps among Greek athletes, but the written history of mankind is full of major outbreaks of disease, which the practitioners of the time attempted to control with methods ranging from folk remedies and the available medicine, to religion and witchcraft, and even to civil and sanitary engineering. It was the careful and scientific observation of these outbreaks that led to the development of modern epidemiology.

Perhaps the most famous early epidemiological inquiry was John Snow's careful observation of the house-by-house locations of incident cases of cholera during epidemics in London in the 1840s and 1850s. Snow's correlation of new cases with some water supply systems, but not others, changed the conventional medical wisdom about how cholera was spread and did so before the microbial nature

of the disease was discovered. Roeché's *Annals of Epidemiology* provides a number of examples of the application of epidemiological methods in relation to local outbreaks in the earlier part of the twentieth century. More recent well-known investigations have included those relating to Legionnaires' disease and AIDS.

Epidemiology as a discipline has grown along with the many associated disciplines of public health and medicine. Microbiology has identified the organisms and modes of transmission for many infectious diseases. Biochemistry, physiology, virology, and related basic medical sciences have provided the scientific background to support the development of the epidemiology of infectious disease. Other sciences ranging from genetics to sociology have contributed to the various specializations within epidemiology, which are broadly classified as infectious disease, chronic disease, cancer, cardiovascular disease, genetic, perinatal, reproductive, oral, occupational, environmental, air pollution, respiratory, nutritional, injury, substance abuse, psychiatric, social, and health care epidemiology as well as pharmaco-epidemiology.

### A CONCEPTUAL PARADIGM

A typical paradigm in epidemiology focuses on the interaction of: host, agent, and environment. The host is typically a person but is sometimes another species or organism which provides a reservoir of infection that is subsequently transmitted to humans. In the presence of an epidemic, one tries to understand the characteristics of the host that provide susceptibility to the epidemic condition. Characteristics that are frequently considered are: (1) genetic characteristics; (2) biological characteristics, such as immunology, physiology, and anatomy; (3) demographic characteristics, such as age, sex, race, ethnicity, place of birth, and place of residence; (4) social and economic factors, such as socioeconomic status, education, and occupation; and (5) personal behaviors, such as diet, exercise, substance use, and use of health services.

The agents in this paradigm have classically been rats, lice, and insects. They may be biological, such as viruses or bacteria, as is the case in most infectious disease epidemiology. Typhus, cholera, smallpox, the Ebola virus, Legionnaires' disease, and AIDS follow this model. Chemical agents have also contributed to significant epidemics. Some

have been physiological poisons such as the mercury once used in hatmaking, lead in old paint, water pipes, or moonshine liquor; others have been carcinogens such as tobacco or PCBs found as a contaminant in oil spread on roads. Recent focus has been on alcohol and other drugs of abuse. Other physical agents have included asbestos, coal dust, and radioactive fallout. Guns, automobiles, and industrial equipment are also agents of injury and mortality. Sometimes the “agent” has been a nutritive excess or deficiency, demonstrated in Goldberger’s 1915 discovery that a deficiency of niacin, part of the vitamin B complex, causes pellagra. Similarly, stress and other elements of lifestyle have been investigated as contributors to cardiovascular disease. An element in the consideration of agents is the degree of exposure and whether the agent alone is a necessary or sufficient cause of the disorder.

The third element of this paradigm is the environment, which may promote the presence of an agent or increase host susceptibility to the agent. Many elements of the environment may be relevant to the disease process, including (1) the physical environment, such as climate, housing, and degree of crowding; (2) the biological environment, including plant and animal populations, especially humans; and (3) the demographic and socioeconomic environments of the host.

At the core of this simple paradigm is the potential for multiple paths of causation for an epidemic. Rarely is a single cause sufficient to ensure the onset of a disease or condition. Thus the approach considers alternative modes of transmission, such as the influence of a common agent versus transmission from host to host. For many agents there is an incubation period, with delayed onset after “infection” or contact with the agent. Further, exposure may lead to a spectrum of disease with variation in the type and severity of the response to the agent. An important element to observe in any outbreak is *who* is not affected and whether that is due to immunity or some other protective factor.

## METHODOLOGY

**Field Methods.** The historical method of epidemiology began with the observation that an epidemic was present, with the initial response

beginning with “shoe leather” investigation of who was affected and how. The initial approach focused on identified cases and their distribution over time and place, leading to a methodology that was often able to discern the risk factors for a disease even before a particular pathogen could be identified. The description of cases in terms of geography and environment as well as various demographic characteristics and exposures remains at the core of epidemiology.

Essential complements to epidemiologic field methods are careful clinical observation, measurement, and classification of the disorder, as well as laboratory identification of any pathogens and identification of potential risk factors for the observed disorder. This methodology is drawn from the methods of associated fields such as pathology, bacteriology, virology, immunology, and molecular genetics. Efforts at classification of pathogens have been augmented by the collection of libraries of reference specimens from past outbreaks.

A problem with the investigation of known outbreaks is that many types of epidemic may develop unobserved until a significant proportion of the population is affected. This has led to the development of surveillance strategies. Among these have been the reporting of multiple causes on death certificates and the mandatory reporting to the health department of many communicable diseases, such as tuberculosis and sexually transmitted diseases. The early detection and reporting of outbreaks make various public health interventions possible. Statistical analysis of mortality and morbidity has become a major component of the public health systems of most countries.

**Case Registers.** An extension to this approach involves case registers, in which identified mortality or morbidity is investigated and compiled in a statistical database, with subsequent investigation of the detailed context for each case. Case registers for particular types of diseases typically identify a case and then collect additional information through review of medical charts and direct field investigations and interviews. Although not as detailed as most case registers, large databases of treated disorders are gathered by various funders or providers of health services. These include the Health Care Finance Administration (Medicare and Medicaid) and various private insurance companies.

The systematic analysis of large databases, whether case registers or those of administrative agencies, makes it possible to monitor the prevalence of various conditions and to detect increases in prevalence or outbreaks long before they would be detected by individual clinicians. Analyses of mortality and morbidity are presented by most public health agencies. Typical methods include the presentation of rates of disorder, such as the number of cases per 100,000 population for geographic and demographic subpopulations. Analyses might ask whether rates are higher in one place compared to another, for a particular birth cohort, or for persons of a particular socioeconomic status. Such comparisons may be crude or adjusted for factors such as age and sex, which may bias such comparisons. Sometimes the analyses provide detailed age- and sex-specific comparisons or breakouts by other risk factors. Yet this approach is dependent on people with diseases being identified through the health system.

**Prevalence Surveys.** In order to discover the true prevalence of various conditions and risk factors, it is possible to conduct sample surveys of the population of a country or other geographic unit. This approach avoids potential reporting bias from the health care system and can identify conditions that might not otherwise be recognized or reported. Typically a representative sample of persons would be interviewed about their health, and sometimes various examinations or tests would be administered. Large comprehensive surveys of health and nutrition are conducted regularly by the National Center for Health Statistics. More specialized surveys of particular conditions such as blood pressure, substance use, and mental health have been conducted by federal, state, and private agencies. Because participation in most such surveys is voluntary, investigative procedures are usually limited to interviews and test procedures with little discomfort or risk. The limitation of prevalence surveys is that they are expensive and may have respondent selection biases. On the other hand, they have great potential for collecting detailed information about disorders and potential risk factors.

**Case-Control Studies.** The case-control method is used extensively in epidemiology. This approach typically starts with a sample of cases of a particular disorder and identifies one or more samples of persons who appear similar but who do

not have the disorder. Careful comparison of the groups has potential for identifying risk factors associated with having the disorder or, alternatively, factors that are protective. The most important issue in designing case-control studies is the designation of an appropriate control group in such a way that the process selecting controls neither hides the real risk factors nor pinpoints apparent but false ones. Control groups are frequently designated by geographic or ecological variables and are often matched on individual characteristics such as age and sex. These methods are somewhat similar to those used in natural experiments, in which two groups differ on one or more risk factors and the rates of disorder are compared, but the case-control method starts with identified cases.

**Cohort Studies.** Unlike the prevalence survey and most case-control studies, which are cross-sectional or retrospective in nature, a cohort study endeavors to identify groups of persons who do not have the disorder or disorders in question and then follows them prospectively through the occurrence of the disease, with ascertainment of factors likely to contribute to the etiology of the disease. Such studies may start with a general sample of the population or may select groups based on the presence or absence of hypothesized risk factors prior to onset of the disorder. They then follow the samples for incident disorder. Unlike cross-sectional studies, cohort studies provide the possibility of determining causal order for risk factors and thus avoid the possibility that apparent risk factors are simply consequences of the disease. The main problem with true cohort studies is that they must last as long as it takes for the disease to develop, which may even take longer than the working life of the investigator. Sometimes, however, cohorts can be found that have been identified and assessed historically and thus can be compared in the present. Examples may be insurance groups, occupational groups, or even residents of certain areas. Then the task is to gather the information on exposure and relate it to health outcomes.

**Experimental and Quasi-Experimental Methods.** A true experiment is dependent on random assignment of persons or groups to conditions, but one would hardly assign a person to a condition known to produce a serious disease. Nonetheless, there are research designs that approximate

true experiments. The most direct is an intervention study in which an intervention that cannot be provided to everyone is provided to one group but not another—or provided to the alternative group at a later time. A quasi-experimental design would identify a group that has been exposed to a risk factor and compare it to another that has not. If the factors that govern the initial exposure appear to be truly accidental, then the quasi-experiment may be nearly as random as a true experiment. In such circumstances or in true randomization, one can determine the effect of the risk factor on outcomes with little risk of the outcome determining the status of the risk factor.

**Alternative Methods.** There are a number of research designs found in epidemiology. Those presented above are typical but certainly not exhaustive of the alternatives. At the core of nearly all of these methods, however, is the identification of factors associated with the incidence or prevalence of particular types of disorder.

### CURRENT ISSUES

In recent epidemiological literature there has been some concern about the future development of epidemiology. A part of this discussion focuses on the basic paradigm being used and whether it has shifted from an ecological approach consistent with the spread of infectious disease and related pathogens to individual risk factors that are more closely related to the current emphasis on chronic diseases. The growing influence of molecular and genetic methods increases this tendency. It has been suggested that this focus on individual and proximate causes blames the individual rather than fixing the prior cause. Some authors have suggested that epidemiology should regain its public health orientation by focusing more on ecological issues affecting the risk status of groups rather than on individual-level factors, and one suggested that the primary focus should be on fighting poverty. A more integrative approach has been suggested by Susser and Susser (1996a, 1996b), who criticize the field for continuing to rely on a multiple risk factor “black box” approach, which is of declining utility, in favor of an approach that encompasses multiple levels of organization and causes from the molecular to the societal.

A second area of concern, identified by Bracken (1998), is the relationship of epidemiology to corporate litigation when various rare disorders are related to sources of risk such as cellular telephones, breast implants, and the like. The issue is the extent to which epidemiologists can provide risk information to journalists and the public without undue burden from unbridled discovery and from the inconsistencies of findings of rare exposures related to rare diseases.

A third issue appears to be the search for professional identity in a discipline that has applicability across the full range of health professions, human and otherwise. The professional positions identified in a 1999 Internet search indicated that there is a continuing market for persons with epidemiological training but that these positions are distributed widely in departments supporting specific disciplines or located in various public health settings. Thus the job titles are often designated in a hyphenated form, such as psychiatric-epidemiologist, cancer-epidemiologist, and genetic-epidemiologist. For a more complete introduction to the principles and methods of epidemiology see Kleinbaum, Kupper, and Morgenstern (1982), Littlefeld and Stolley (1994), or MacMahon and Trichopoulos (1996).

### REFERENCES

- Ahlbom, Anders, and Staffan Norell 1990 *Introduction to Modern Epidemiology*. Newton Lower Falls, Mass.: Epidemiology Resources, Inc.
- Bracken, M. B. 1998 “Alarums False, Alarums Real: Challenges and Threats to the Future of Epidemiology.” *Annals of Epidemiology* 8(2):79–82.
- Friedman, Gary D. 1980 *Primer of Epidemiology*, 2nd ed. New York: McGraw-Hill.
- Lilienfeld, David E., and Paul D. Stolley 1994 *Foundations of Epidemiology*. New York: Oxford University Press.
- MacMahon, Brian, and Dimitrios Trichopoulos *Epidemiology: Principles and Methods* 2nd ed. Boston: Little, Brown.
- Roueché, Berton 1967 *Annals of Epidemiology*. Boston: Little, Brown.
- Susser M., and E. Susser 1996a “Choosing a Future for Epidemiology: I. Eras and Paradigms.” *American Journal of Public Health* 86(5):630–632.

— 1996b “Choosing a Future for Epidemiology: II. From Black Box to Chinese Boxes and Eco-Epidemiology.” *American Journal of Public Health* 86(5):674–677.

CHARLES E. HOLZER, III

## EPISTEMOLOGY

The term *epistemology* is used with two separate meanings according to different cultural traditions. In English-speaking countries, epistemology denotes the philosophical theory of knowledge in general: in this sense, it includes themes and problems such as the question of the possibility of valid knowledge, the analysis of the nature of such validity, the foundation of knowledge on reason or on experience and the senses, the analysis of different types of knowledge, and the limits of knowledge. In continental Europe, the above issues are considered part of the field of the more general discipline of *gnoseology* (*gnoséologie* in French, *gnoseologia* in Italian and Spanish) or *theory of knowledge* (*Erkenntnistheorie* in German), whereas the aim of epistemological inquiry is restricted to scientific knowledge.

In this second sense, with particular respect to social sciences (and sociology, in particular) the fundamental epistemological question becomes: “Is it possible to acquire any valid knowledge of human social reality? And, if so, by what means?” As these questions show, epistemological issues are inescapably interconnected with methodological problems; however, they cannot be reduced to simple technical procedures and their validity, as a long empiricist tradition among sociologists has tried to do. A full epistemological awareness, from a sociological point of view, should cope with at least four main issues:

1. Is the nature of the object of social sciences (i.e., social reality) fundamentally different from that of the object of natural sciences (i.e., natural reality)?
2. Consequently, what is the most appropriategnoseological procedure with which to study and understand social reality?
3. Are we sure that the particular knowledge we get by studying a particular social reality can be generalized?

4. What kind of causality can we postulate between social events, if any?

Historically, the various sociological traditions or schools have answered these four questions differently. We shall trace them by following the three main epistemological debates that took place between three pairs of opposing schools of sociological and epistemological thinking: positivism versus historicism, logical empiricism versus dialectical theory, and realism versus constructivism.

### POSITIVISM VERSUS HISTORICISM

At the beginning of the nineteenth century, the discipline of sociology originated as the positivistic science of human society and aligned itself against metaphysical and philosophical speculation about social life. This, according to Auguste Comte, the founder of sociology, implied the straight application of the scientific method used in physics and the other natural sciences to the analysis of human society. His *social physics* (Comte 1830–42) is the expression of the new culture of the Industrial Revolution and of its practical ambition of a totally planned social life. Comte’s positivistic doctrine is organic and progressive and holds a naive faith in the possibility of automatically translating scientific knowledge of the laws of society into a new harmonious social order.

The uncritical assumptions of Comte as to the reliability and universal applicability of scientific method were at least partially tempered by John Stuart Mill, who considered illiberal and doctrinaire Comte’s idea that, once sociological laws are established by the same scientific method of natural science, they can no longer be questioned. Mill objected that scientific knowledge never provides absolute certainties, and he organically settled the British empiricist tradition in his *A System of Logic, Ratiocinative and Inductive* (1834). Starting from the epistemological positions of David Hume, Mill considered knowledge as fundamentally grounded on human experience, and the induction as its proper method. The possibility of inductive generalization is based on the idea of regularity and uniformity of human nature, ordained by laws; and even social events are correlated in the form of empirical laws, which we can understand using the scientific method, considered as a formal logical structure independent from human subjectivity.

The individualistic and liberal features (based on utilitarian ethics) of British positivism were particularly evident in Herbert Spencer, who opposed to Comte's organicistic view of society the greater importance of the particular with respect to the organic whole. His central focus on the concept of evolution supplied him with a synthetic perspective with which to study reality as a whole, using the analogy of biological and inorganic evolution to consider social evolution (1876–1896). However, the most complete development of a positivistic epistemology in sociological thought is certainly contained in Emile Durkheim's work *Les règles de la méthode sociologique* (1895). What he terms *faits sociaux* (social facts) can be characterized according to four criteria: two of the criteria are related to the subjects observed, whereas the other two are concerned with the sociologist observing them.

The first criterion considers the social facts as a reality external to the individual conscience: social institutions have their own life independent from individual life. The second criterion is that this external, objective nature of social facts consequently confers upon them a normative, coercive power over the individual: these social facts impose themselves on him, even without his will. Morals, public opinion, law, customs are all examples of this. These two features of social facts require, on the part of the social scientist observing them, the observance of two further qualifying criteria. Firstly, they should be considered as "things," that is, they should be studied as an external, objective reality, separate from the individual consciousness. Secondly, consistent with their nature, social facts can be explained only by other social facts: the nature of social causality is specific, and can be reduced neither to psychological causes of individual behavior, nor to biological causality, as Spencer's evolutionism suggested. The emerging nature of social phenomena from the level of psychological and biological phenomena grounds their autonomous, distinct reality: this, in turn, constitutes the specific field of sociology and of the other social sciences. To explain social phenomena, sociology should adopt a unilinear concept of causation: the same effect always corresponds to the same cause. Any plurality of causes, according to Durkheim, involves the impossibility of sorting a scientific principle of causality.

In sum, we can say that the classic positivistic paradigm of sociology is characterized by the recognition of the specific nature of social facts as emerging from the other spheres of reality. However, this implies an idea of sociology as a naturalistic science of society, using the same methods already applied with success in natural sciences (*methodological monism*), and grounded on experience as perceived by the senses and generalized on the basis of universal human nature (*induction*). Finally, a proper explanation of social facts should take into account only one cause for each effect (*mono-causal determinism*).

A serious challenge to this epistemological view of sociology came from the German Historical School: the debate about the method (*methodenstreit*) which took place in Germany during the last decades of the nineteenth century is the first example of epistemological debate in sociological history. It started with Wilhelm Dilthey, the major representative of German historicism, who argued against Comte and Mill's proposal of introducing the methods of physics into the "moral sciences." What he termed "spiritual sciences" (*Geisteswissenschaften*) were of a radically different type from the natural sciences and therefore they could not share the same method (Dilthey 1833). In fact, their objects were states of mind, spiritual experiences that could be apprehended only by means of an "empathic understanding" (*Verstehen*): "we explain nature, we understand psychic life" (*ibid.*). In the first phase of Dilthey's thought, this understanding, on which the sociohistorical nature of the spiritual sciences is grounded, was not considered to be mediated by sense perception, but rather as producing a direct and immediate intuitive knowledge-by-acquaintance. In the second phase of his thought (1905), however, he believed that psychic life was not immediately understandable, but would require the hermeneutic interpretation of its objective displays in cultural life. In any case, his distinction between the aim of natural sciences (to explain *erklären* data by means of external senses) and the scope of human sciences (to understand *verstehen* through an intrapsychic experience) was posited as a long-lasting distinction in the social sciences.

Wizhelm Windelband and Heinrich Rickert, two other adherents of German historicism, criticized Dilthey's distinction between nature and



spirit and the related sciences, since they considered human knowledge as always a spiritual activity, without regard to its *object*. In its place, they proposed a distinction based on the *form* and on the different type of *methods* used: whereas the *idiographic* method is a description of singular events, the *nomothetic* method is concerned with the inquiry of regularities and general laws (Windelband 1894). The first type of procedure is typical of—though not exclusive to—the *sciences of the particular*, the sciences of culture that, being historical, had to interpret and understand the individual character of the historical event; the second type of method is more typical of the *sciences of the general*, natural sciences that are aimed at establishing general laws. Knowledge is always a simplification of reality, which is a heterogeneous continuum (*eterogenes Kontinuum*); knowledge can either proceed by making a homogeneous *Kontinuum*, as the natural sciences do, or by sectionalizing a portion of heterogeneous *Diskretum*, such as the sciences of culture do. Rickert (1896) further developed the necessity of referring to values as a foundation for sociohistorical knowledge: the sciences of culture (*Kulturwissenschaft*) are not idiographic disciplines only, they should also refer to *value relevances* (*Wertbeziehung*) if they want to understand the actual meaning of sociohistorical events.

The most significant attempt to reconcile the two antitheses (understanding/explanation and idiographic/nomothetic) has been carried out by Max Weber, the founder of German sociology, who was profoundly influenced by German historicism, although critical of its idealistic orientation. In his methodological essays (1904–17) and in subsequent works, Weber undertakes an inquiry into the nature and validity of methods used in the social sciences and presents a general epistemological framework for his *interpretive sociology* (*Verstehende Soziologie*). He argued in particular for the necessity of when investigating social actions, 1) resorting to an *interpretive understanding*, which is not separated by *causal explanation*, 2) while avoiding *value judgements*, establishing *value relevances* (*Wertbeziehungen*) of the subject matter as criteria for its cultural importance and scientific pertinence, and 3) preserving the social and cultural uniqueness of the historical event, using the *ideal-type* methodology. These three points constitute

Weber's main elaboration of the epistemological tenets of German historicism.

With regard to the first point, Weber criticized Wilhelm Wundt and George Simmel (who were influenced by Dilthey), who reduced socio-historical knowledge to psychological understanding, since this position cannot support an objective knowledge: to properly understand a socio-historical event requires establishing some relations of cause and effect in order to test our interpretive comprehension. Therefore, Weber conceived an idea of sociology as both a generalizing, nomothetic and interpretive, idiographic science, arguing for the complementarity of interpretive understanding and causal explanation: that is, a researcher's personal understanding should be balanced by empirical and statistically established regularities of scientific explanation (1913).

Secondly, Weber distinguishes *value judgements*, based on personal faith and belief, and *fact judgements*, based on science. Scientific knowledge cannot be neutral, in the sense that the values of the researcher inevitably influence his choices of relevance—the specific selection of problems and focus. However, it can be *nonevaluative* and *objective* in the sense that, once its priorities have been selected according to the researcher's personal values, the researcher should proceed by testing empirical evidence supporting his hypotheses, thus expressing fact judgements only.

Finally, the ideal-type concept was employed by Weber to explain certain unique historical events using a model of *plural causality*: a singular event is always the effect of a multiplicity of historical connections. The problem becomes, in the infinite flow of events, to sort out the proper model of factors that under certain conditions and at a certain time can *probably* explain that effect. This also implies a *probabilistic* idea of causation, which gets beyond the deterministic approach typical of positivistic unilinear causation by a model of plural probabilistic-causal connections.

#### LOGICAL EMPIRICISM VERSUS DIALECTICAL THEORY

The epistemological legacy of German historicism—with its *dualistic* approach to the relationship between social and natural sciences, its preferential bias for a gnoseological model based on

*interpretive understanding*, and its *probabilistic* model of *multiple causality*—has undoubtedly influenced a significant part of contemporary twentieth century sociology, such as the School of Chicago (Park's contacts with Windelband and Simmel are well known) and, more strongly, the later interpretive sociology (through Weber and Simmel) and the phenomenological tradition in sociology (through Schutz). Its impact is manifest even in the antipositivistic methodological choices of many sociologists as different as Pitrim Aliksandrovič Sorokin (1959) and Charles Wright Mills (1959). Altogether, these different currents answer negatively the epistemological question about the possibility of acquiring knowledge of human social reality by means of empirical data alone. They reject the methodological unity of the natural and social sciences and do not accept the straight application of the scientific method to sociological analysis. When researchers undertake empirical research, they resort to methodological perspectives that result from the researcher's epistemological choices: they tend to reject technical terminologies and statistical quantification, to privilege common-sense conceptualization and language, and to take the point of view of the social actor instead of that of the researcher and of the scientific community. Qualitative methodologies of this kind are also favored by those microsociologies—such as symbolic interactionism, ethnomethodology, and sociologies of daily life of different varieties—that have not been directly influenced by the idealistic and antipositivistic tenets of the German historicism. Although seldom explicitly stated, their epistemological foundations, which tacitly direct their choice of methods and tools of social research, are mostly consonant with those of interpretive and phenomenological sociologies.

The large majority of sociologists, however, followed an empirical-quantitative approach and accepted the scientific method borrowed from natural sciences as the appropriate and valid method of social sciences. The positivistic epistemological foundations of this *mainstream* sociology are quite evident in its most significant representatives Lazarsfeld and Lundberg, and in their utilization of the *social survey* as the main research tool of sociological inquiry. Lazarsfeld, in particular, although reluctant during his work at Columbia University to directly tackle epistemological issues

while preferring practical empirical research, was well known for his neo-positivistic convictions (Gallino 1973:27), which were rooted in his mathematical and psychological background within the Vienna philosophical movements of the post-World War I period.

In fact, during the twenties and thirties of the twentieth century, a new group of positivists, the so-called “Vienna Circle,” arose in the Austrian capital, profoundly influenced by the work of the physicist Ernst Mach, the pioneers of mathematical logic Gottlob Frege, Bertrand Russell, Alfred North Whitehead, and Ludwig Wittgenstein, and by the French *conventionalism* of Pierre Duhem and Jules-Henri Poincaré. Mach considered science not a fact-finding but a fact-related activity: it cannot claim to discover the absolute truth about reality, since its laws are not absolute but, at best, “limitations of possibilities.” Russell, among the pioneers of mathematical logic, attempted a logical foundation of mathematics that considers the element of necessity present in both disciplines. This can be considered as the basis of the logical structure of thinking. Finally, Duhem and Poincaré suggested an idea of scientific theories as being purely hypothetical and conventional constructions of the human mind.

On these premises, the main representatives of the “Vienna Circle”—Moritz Schlick, Rudolf Carnap, and Otto Neurath—preferred the label of *logical empiricism* or *neopositivism* to distinguish themselves from classical positivism, since they discarded its sensist view—the idea that knowledge originates from experience gained only through the senses, and from the observation and verification of that experience. They rejected direct observation of experience as the only means of hypothesis testing, and they considered verification as “testability in principle.” In this way, they accepted theoretical constructs without direct empirical referents as meaningful. They renewed empiricism on the new basis of the logical analysis of language: in fact, they argued that linguistic statements can be shown to be true or false by appeal to both logic and experience, and that both means, being factual, can be considered meaningful.

Logical empiricism still maintained some fundamental epistemological tenets of classical positivism, however, such as the idea that only empirically verifiable knowledge was meaningful, that

science was a cumulative process based on induction, that the method of physics was the method of all sciences including social sciences (*methodological monism*), and that the discovery of natural and general laws was the fundamental aim of any science. The criticism of Karl Popper was aimed at most of these foundations and reversed them in a sort of *negativism* (Cipolla 1990). Since his first major publication (Popper 1934), the Austrian philosopher formulated his tenets in opposition to the main assumptions of the neopositivist: first and foremost, the *induction* as an appropriate method to infer general laws from the observation of a singular case. He argued that one case is enough to demonstrate the falsity of induction and that verification always depends on observational theories that are often derived from substantially the same theory from which hypotheses under test have been deducted, making their test logically inconclusive. Popper dropped the notion of *verification* itself in favor of that of *falsification*, and proposed a *hypothetic-deductive* method, based on the assumption that a hypothesis can be definitely considered false if it fails an adequate test, while if it is congruent with data this does not necessarily mean it is true. In any case, Popper argued that a theory is maintained as true only provisionally, and it always remains a conjectural hypothesis that can be confuted in the future. In this way, scientific knowledge is never a closed, completed system, but always remains open to new possibilities. And even though Popper believed in the *methodological monism*—the unified theory of method—this is not absolute but conjectural, critical, and subject to falsification (1963).

A substantial refusal of *methodological monism* is represented by the Frankfurt School with its “critical theory of society” that advocates the necessity of the dialectical method in sociology. Its epistemological tenets are derived partly from Weber’s thought on Western reason and rationalization and partly from the young Hegel-influenced Marx of the *Philosophical-Economical Manuscripts* of 1844. It was also strongly influenced by the two critical Marxists Lukacs and Korsch and by the psychoanalysis of Freud. Dialectical epistemology implies:

1. The *holistic* approach, that is, the necessity of considering the totality in order to understand the parts, which are mediated by totality.

2. The denial of the separation of history and sociological theory in order to grasp the dialectical process of change.
3. A demystifying attitude in undertaking critical analysis of society and of the irrationality of Enlightenment reason once it has become purely instrumental and separated from its goals, that is, when it’s become a pure means of domination over men and society.

In 1961 the second most significant debate in the history of sociological epistemology took place at Tübingen between Theodor Adorno and Jurgen Habermas (two of the most prominent representatives of the Frankfurt School) on one side and Karl Popper and Hans Albert on the other (Adorno et al. 1969). The neopositivistic epistemology of the latter two was strongly questioned by the former two, who criticized its dialectic nature and who considered it a sort of logical formalism without any connection with its contents and that was incapable of understanding the social reality. The aim of sociology is to go beyond apparent phenomena, to grasp social contradictions and conflicts by interpreting society as a totality. This implies a refusal of the individualistic approach of positivism, of its *monism* with regard to the scientific method, of its measurement and *quantification* of social reality. Beyond Popper and Albert, the actual target of the Frankfurt scholars was the American school of sociology, whose positivistic analytical framework (theoretical categories and their translation into research tools) was considered an ideological reflection of the domination structures in the late capitalistic society.

### REALISM VERSUS CONSTRUCTIVISM

Since the 1960s, logical empiricism, under criticism from many sources, has dissolved in a plurality of *postpositivistic* approaches, whose common denominator is a reformulation of positivistic tradition. This allows them to be grouped under the label of *scientific realism*, which asserts the absolute or relative independence of the reality under scientific scrutiny from the human researcher studying it. This, in turn, is based on the fundamental distinction between the objectivity of the reality observed and the subjectivity of the scientific observer studying it. The common tenet of realisms

of all types is that the observer does not belong to the reality he observes; by means of his techniques of scientific inquiry, he should avoid any involvement and influence on reality, maintaining a neutral position. In this way, the subjectivity of the observer is limited to his “discovery” of the objective reality.

One of the first attacks on the neopositivistic legacy was by Willard Van Ormand Quine (1952), who proposed a holistic view of knowledge, considered as a field of forces whose limiting conditions are constituted by experience. In turn, Mary Hesse (1974) considered scientific language as a dynamic system whose continuing growth is due to metaphorical extension of the natural language. Thomas Kuhn (1962) greatly contributed to the growing consciousness of the reality of scientific change by analyzing its actual historical development, challenging the common-sense concept of science as a purely rational enterprise. Taking scientific practice into account, he showed how it is usually ruled by a *paradigm*, a world view, legitimized by the scientific community, that remains dominant until a new paradigm replaces it.

Paul K. Feyerabend, with his “methodological anarchism” (1975), proposed a paradoxical and extreme view of science. In Feyerabend’s view, the scientific method remains the only rational procedure for deciding and agreeing upon which theory is more adequate to describe and explain a state of affairs in the natural as well as in the social worlds. Of course, rationality depends on common premises and procedures: the advantage of accepting them is great because it is through them that intersubjectivity is achieved. Scientific objectivity resides neither in the object of knowledge, as classical positivism maintained, nor in the subject, as idealists tend to believe, but rather in the intersubjectivity that results when researchers adopt the same procedures and accept the premises on which those procedures are based. It is the reproducibility of the methods, techniques, and tools of scientific research that secures the replicability of results. And their reproducibility is due largely to their being public procedures, easily scrutinized and reapplied. The results of the research in the natural sciences may be more objective than those in the social sciences, but this is due to the standardization and publicity of procedures in the natural sciences. There is no ground for supposing that

natural sciences possess a special objective attitude, while social scientists possess a value-oriented attitude. The problem is that many research procedures in the social sciences are not reproducible, because they often reflect a private state of mind communicated through linguistic expression full of connotative meanings, often not shared by all in the research community.

The *critical realism* of Roy Bhaskar (1975) proposes an ontology based on the distinction of three spheres: the *real*, the *actual*, and the *empiric*, arguing in favor of the existence of structures or hidden mechanisms that can work independently from our knowledge, but whose power can be empirically investigated both in closed and open systems. This realistic view of science is supported also by Rom Harré (1986), who stresses the role of models in the development of theory. The new realists also suggest a *relational paradigm* for sociology in order to overcome the traditional antinomy between structure and action by a theory of *structuralization* (Giddens 1984) and by a *transformative model of social activity* (Bhaskar 1989). According to this new paradigm, the social structure should be considered as both the omnipresent condition and the continuously reproduced outcome of the intentional human action. This view is shared by the *applied rationalism* of Pierre Bourdieu, who systematically applies relational concepts and methodically compares his theoretical models with the empirical material that is the result of different research methodologies (Bourdieu et al. 1973). The critical view of science taken by the new realists, joined with their transformative conception of social reality, has produced a *critical naturalism* that considers both the positivist and the historicist traditions as dependent on the same positivistic conception of the natural sciences. This critical naturalism holds that, in actuality, both the human social life and the natural life are liable to scientific explanation, although of a different kind (Bhaskar 1989).

This realistic view has been seriously challenged in the 1980s and 1990s by the *constructivistic movement*, giving rise to the third, and still ongoing, epistemological debate in the history of sociology. Constructivism proposes the unification of the objective reality observed with the subjective reality of the observer. The rationale is that the reality investigated and the science used to investigate it have an equally subjective origin, which

implies that the reality observed “depends” on the observer. This does not mean that constructivism denies the actual existence of an autonomous reality from the observer, but rather that an “objective” representation of this external reality by scientific knowledge is not possible. All that can be said about reality is inevitably a “construction” of the observer (von Foerster 1984), who, in turn, being part of this reality, is a “black box” whose internal components are unobservable. Therefore, even for the constructivists, reality is an elusive objectivity, with its autonomous existence which cannot be reduced to a simple subjective and arbitrary experience. What is “constructed” is the knowledge and not the reality (Von Glasersfeld 1987).

There are different types of constructivism. The most radical (Von Glasersfeld 1995) maintains that nothing can be said even by the observer, apart from the trivial verification that he observes. An unknowable observer faces an unknowable reality. A more moderated type of constructivism considers still possible a theory of the observer, even though it denies the same possibility for the reality observed (Maturana 1988). And finally, there is a constructivism which accepts both a theory of the observer and of the observed, considering them as mental systems (Luhmann 1990).

#### TOWARDS AN INTEGRATED AND PLURALISTIC PARADIGM

As the current debate between realism and constructivism shows, the positions only appear distant. Especially when they move from the more abstract theory to the empirical research, the respective positions become more ambiguous, vague, and overlapping. Moreover, there are clear indications that the old cleavages between positivistic and interpretive sociologies may slowly fade away. Even though quantitative research techniques appear more precise and rigorous—with their standardization and reproducibility features—than qualitative methods, it is difficult to justify using quantification for all types of sociological data. Many sociologists believe that not all aspects of social phenomena could be subject to the rules of quantification. And most of them start out thinking that quantitative and qualitative methods and techniques are complementary. Yet, such an integrated and pluralistic methodology needs to be justified by an explicit epistemology. Therefore,

the times seem ripe for a new epistemological foundation of an emerging integrated and pluralistic methodology, even though a few attempts have been made towards this goal.

The first attempt was the *ecological paradigm* of Gregory Bateson (1972). Strongly influenced by his participation in the 1940s in the interdisciplinary research group on cybernetics composed of Von Neumann, Shannon, Von Foerster, and Wiener, among others, he then organized all his subsequent work in different fields around a few central cybernetics concepts, including schismogenesis, circular communication, and feedback. On the ground of this new theoretical perspective, he built up an epistemology aimed at getting beyond the boundaries between the internal and the external components of the observer, the “mind,” conceived as a network of founding relationships. In this process, the identity of the observing subject is dissolved into its ecological environment, in the relationship between subjective meanings, action, and action objective. Even though Bateson’s epistemology precedes the debate between realism and constructivism, it has clearly inspired most of the constructivistic approach.

A second attempt, historically, can be traced in Edgar Morin’s attempt to set up a new general method (Morin 1977, 1980, 1986), which can be applied in every field of knowledge according to a methodological *plurivers*. In this way, he tries to overcome the shortcomings of the traditional scientific paradigm, which he considers *disjointed*, *reductive*, and *simplified*, since it does not take into account the complexity of reality. The *multidimensional* character of reality is precisely the starting point of his very comprehensive attempt to found a new epistemology based on the *auto-eco-organizing* principle. An attempt, however, which for the most part is affected by a biologicistic bias, seems quite inappropriate for the social sciences.

Finally, Costantino Cipolla has proposed a new *correlational* paradigm based on an *epistemology of tolerance* (Cipolla 1997) which prefers an “inter-” and “co-” perspective linking together and integrating the traditionally opposite epistemological poles in an attempt to reassemble them into a new pluralistic approach. Closer to the sociological tradition of the first two attempts, this new correlational paradigm takes into account the reasons of both the realists and the constructivists

and focuses its attention on the connection, the relation, and the link between the objective reality and the subjective observer in order to avoid any absolutism of either of the poles. It is starting from this bivalent conception of reality—which considers reality as both autonomous in itself and “constructed” by the subject, who is in turn the outcome of the real social forces, so that each pole is considered only partially autonomous and strictly interconnected with the other—that it should be possible to re-found the sociological epistemology on the ground of the methodological research questions of the discipline, in recognition of the fundamental ambivalence of reality and of the need of an *adductive* procedure as a “double movement” between induction and deduction, particular and general, theoretical hypotheses and social reality beyond any self-contained reductivistic monism.

## REFERENCES

- Adorno, T. W. et al. 1969 *Der Positivismusstreit in der deutschen Soziologie*. Berlin: Luchterhand.
- Bateson, G. 1972 *Steps to an Ecology of Mind*. San Francisco: Chandler.
- Bhaskar, R. (1975) 1978 *A Realist Theory of Science*. Hemel Hempstead: Harvester Wheatsheaf.
- 1989 *The Possibility of Naturalism*. Hemel Hempstead: Harvester Wheatsheaf.
- Bourdieu, P. et al. (1973) 1991 *The Craft of Sociology: Epistemological Preliminaries*. Berlin and New York: De Gruyter.
- Cipolla, C. 1990 *Dopo Popper*. Roma: Borla.
- 1997 *Epistemologia della tolleranza*. vol. 5 Milano: Angeli.
- Comte, A. 1830–42 *Cours de philosophie positive*. Paris: Bachelier.
- Dilthey, W. (1883) 1966 *Einleitung in die Geisteswissenschaften*. Vol 1 of *Gesammelte Schriften*. Stuttgart: Teubner.
- 1905 *Kritik der historischen Vernunft*. Stuttgart: Teubner.
- Durkheim, E. 1895 *Les règles de la méthode sociologique*. Paris: Alcan.
- Feyerabend, P. K. 1975 *Against Method: Outline of an Anarchist Theory of Knowledge*. London: NBL.
- Gallino, L. 1973 “Metodologia neopositivistica e teoria sociologica.” In *Quaderni di Sociologia*. XXII, 1.
- Giddens, A. 1984 *The Constitution of Society*. Cambridge: Polity Press.
- Harré, R. 1986 *Varieties of Realism*. Oxford: Blackwell.
- Hesse, M. 1974 *The Structure of Scientific Inference*. London: Macmillan.
- Kuhn, T. S. 1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Luhmann, N. 1990 *Die Wissenschaft der Gesellschaft*. Frankfurt am Mein: Suhrkamp.
- Maturana, H. 1988 “Reality: the Search for Objectivity and the Quest for a Compelling Argument.” *The Irish Journal of Psychology*, Special Issue, 9,1.
- Morin, E. 1977 *La methode 1. La nature de la nature*. Paris: Seuil.
- 1980 *La methode 2. La vie de la vie*. Paris: Seuil.
- 1986 *La methode 3. La connaissance de la connaissance*. Paris: Seuil.
- Mill, J. Stuart (1834) 1947 *A System of Logic, Ratiocinative and Inductive*. New York: Longman.
- Mills, C. Wright 1959 *The Sociological Imagination*. New York: Oxford University Press.
- Popper, K. 1939 *Logik der Forschung*. Wien: Springer.
- 1963 *Conjectures and Refutations*. London: Routledge and Kegan Paul.
- Quine, W. V. O. (1952) 1963 *From a Logical Point of View*. New York: Harper and Row.
- Rickert, H. (1896) 1913 *Die Grenzen der naturwissenschaftlichen Begriffsbildung*. Tübingen: Mohr.
- Sorokin, P. A. 1959 *Fads and Foibles in Modern Sociology and Related Sciences*. New York: Henry Regnery.
- Spencer, H. 1876–96 *The Principles of Sociology*. London: Williams and Norgate.
- Von Foerster, H. 1984 “Erkenntnistheorien und Selbstorganisation.” *Delphin* 4:6–19.
- Von Glasersfeld, E. 1987 *The Construction of Knowledge*. Seaside, Calif.: Intersystems Publishers.
- 1995 *Radical Constructivism. A Way of Knowing and Learning*. London: Falmer Press.
- Weber, M. (1904–17) 1949 *The Methodology of the Social Sciences*. New York: Free Press.
- (1913) 1968 “Über einige Kategorien der verstehenden Soziologie.” In *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: Mohr.
- Windleband, W. (1894) 1914 “Geschichte und Naturwissenschaft.” In W. Windleband, ed., *PrÄtudien*. Tübingen: Mohr.

COSTANTINO CIPOLLA  
GUIDO GIARELLI

## EQUALITY OF OPPORTUNITY

*Equality of opportunity* refers to the fairness of processes through which individuals with different backgrounds or from different social groups reach particular outcomes, such as educational or occupational goals. Sociologists have developed several alternative approaches to defining and assessing equality of opportunity in each outcome domain, including trends in demographic gaps, residual differences after relevant qualifications are taken into account, process differences in the variables linking individual attributes to outcomes, and structural differences in the barriers encountered in preparing for, learning about, or obtaining particular educational or occupational achievements. Each approach has advantages and disadvantages for particular scientific, policy, and practical purposes.

### TRENDS IN DEMOGRAPHIC GAPS

Equality of opportunity is usually judged with reference to major demographic groupings, such as race, sex, or socioeconomic status. When there are significant changes over an extended period of time in educational or occupational outcome gaps for major subgroups, inferences may be made that changes in opportunity structures underlie the trends. This argument is best made when confidence is high on the accuracy of the outcome trend data, the timing of changes in outcomes and specific opportunity processes can be matched, and other subgroup differential changes in personal resources can be discounted as nonexistent or following a different time sequence.

In education, changes in racial/ethnic gaps in achievement scores and college attendance rates have been subjected to trend analyses to pinpoint opportunity processes. The test score gap showing higher average achievement in basic skills by white students compared to African-American students has shown a significant narrowing since the 1970s, although the convergence may have stalled or begun to reverse on some tests by the end of the century (Jencks and Phillips 1998). Since the largest gains during this period occurred for African-American students who entered school between the late 1960s and late 1970s, particularly in the South, some have credited the antipoverty legislation and school desegregation enforcements of

the time aimed at increasing educational opportunities for minorities. Trends in college entry among gender and racial/ethnic groups show interesting patterns in recent decades that invite explanations of changing opportunities for selected groups. After a long-term upward trend from the 1940s to the mid 1970s, when college entry increased more for African Americans than whites to narrow the gap in educational attainments, college entry actually declined among African-American high school graduates from the mid 1970s to the mid 1980s, with some signs of recovery in black initial college enrollments in the 1990s that still did not match the progress of whites during the same period. Careful analyses to match various contemporaneous changes to the college entry trends through the 1980s ruled out most changes in personal resources, such as family income and academic achievement, although modifications in college financial aid policies from grants to loans in the face of rapidly rising college costs may account for minority college entry declines (Hauser 1993).

Racial/ethnic gaps in college degree attainments, after gradually closing for successive decades, still remain very large and recently also have begun to grow even larger. The number of African-American recipients of either bachelor or advanced degrees has actually declined since the beginning of the 1980s, which can be tied to numerous inequalities of access and support in the American education system across the grades and the court-inspired decline of racial considerations in the admissions policies of some colleges (Miller 1995; Orfield and Miller 1998).

Economic changes over recent decades in the distribution of income have also been subjected to trend analyses. Income inequality as measured by the gaps in the percentage of annual income held by different social groups, such as the top fifth versus bottom fifth of the population, showed moderate improvements, with only small periods of slight reversals over the years following World War II through the end of the 1970s. But the period since 1980 has been one in which income inequality increased sharply, sometimes called the shrinking middle class, as the rich got richer, acquiring a higher percent of total income, while productivity growth eliminated jobs or decreased earnings for many less educated workers (Levy,

1995). Numerous factors have been associated with recent trends, including reduced access to employment opportunities due to the movement of jobs from many urban locations where poor minority workers live (Wilson 1996).

### RESIDUAL DIFFERENCES IN EQUALITY OF OPPORTUNITY

Even when major outcome gaps are observed, the issue remains whether individuals from major population subgroups have had the same chances to achieve educational or occupational success, assuming that they possess the same distributions of personal attributes to qualify for success. Because any initial average outcome gaps between subgroups can be due to unequal possession of relevant qualifications, as well as to unfair access to the opportunities that link qualifications to achievements, it is necessary to take into account differences in personal qualifications before deciding that unequal opportunities exist.

Researchers have frequently tested for inequalities of opportunity by estimating the residual gap between the educational or occupational success of selected race, sex, or social-class groups after individual differences in relevant credentials or competencies and educational or labor market locations have been statistically controlled for. The usual methodology is to estimate a prediction equation or to use other methods of standardization for selected individual resource variables that permit a researcher to compare the actual group difference in an educational or occupational outcome with the residual gap that would be expected if one group's productivity resources were replaced by the average resources of the other group (Farley and Allen 1987, Chap. 11). For example, the actual average difference in annual earnings of African-American and white workers in the North would be compared against the residual earnings gap when one assumes that African-American workers' resources (such as education and labor-market experience) deliver the same rate of return in earnings as that experienced by white workers. Some problems are inherent in this approach, including the risk of overestimating the residual gap if some important qualification variables are omitted or poorly measured and the chance of underestimating the residual gap when some groups

are deprived of relevant qualifications due to earlier unequal opportunities not reflected in the estimation methodology. Nevertheless, several important residual race, sex, or social-class gaps have been identified for various important educational and occupational outcomes. However, these gaps are often associated with some subgroups but not others, and some gaps have been changing more rapidly than others in recent years.

Numerous national and regional studies have been conducted since the 1960s to estimate the inequality of job opportunities, including research that examines residual subgroup differences in unemployment rates, occupational distributions, and dollar returns from holding a job. Studies of race, sex, and social-class residual gaps in earnings and income of employed workers have been particularly noteworthy, with more than twenty-five major national studies having been published since 1965 (Farley and Allen 1987).

The research on earnings gaps that estimates the "cost of being black" due to inequality of job opportunities has contrasted the experiences of male and female workers and reported the continuing but declining significance of race. After taking into account differences in educational attainment, age or years of potential labor-market experience, hours of work, and regional location, large residual gaps in earnings are found between male African-American and white workers, with African Americans earning 10 to 20 percent less than comparable whites in various regions and at various educational levels. Women continue to earn much less than men of the same race with similar educational credentials, but the residual race gap for women is no longer the same as reported for men. In 1960, African-American women earned less than white women at all educational levels except college graduate, but this gap had been eliminated or reversed by 1980, when college-educated African-American women actually reported greater earnings, largely because of greater hours of employment. The residual race gap in earnings for employed workers also appeared to grow somewhat smaller, for men between 1960 and the 1990s, but it still remains between 10 and 15 percent at all educational levels.

At the same time, evidence is mounting that racial gaps in rates of unemployment are significant and have been growing worse since the 1960s



for African-American men in most age and education categories; they are especially severe for unmarried young African-American men in the North who have limited educational attainments (Farley and Allen 1987, chaps. 8–11; Jaynes and Williams 1989, Chap. 6). While the economic boom period in the 1990s benefited the average employment prospects and median incomes of all race/ethnic and gender groups, the lower levels of the income distribution and the poorly educated did not keep pace and actually fell behind in some of these years.

Race inequalities of accumulated wealth have also been investigated after statistically taking into account factors that affect how individuals encounter financial opportunity structures. Measured by either the net worth of a household total assets less any debts or as net financial assets that exclude equity accrued in a home or vehicle that is more difficult to convert into other resources, very large racial differences are found in wealth that have grown even larger in recent years. The average black family held \$3,779 in mean net worth in 1967, which rose to \$19,736 in 1984 and \$23,818 in 1988, in comparison to the average white family's mean net worth, which stood at \$20,153, \$76,267, and \$95,667 for the same time periods, for a race gap that widened by \$40,000 during these years and reached \$71,849 by 1988. After controlling for differences in age, annual household income, household composition, and professional and self-employed status, nearly three-quarters of the racial gap is left unexplained. Institutional and policy factors that may account for the residual race differentials in wealth include mortgage loan and interest rate practices, evaluation of different neighborhoods, and various inheritance mechanisms whereby status and resources can be transmitted across generations (Oliver and Shapiro 1995).

Inequalities of educational opportunity have been examined by estimating residual race, sex, or social-class gaps in outcomes net of initial resources, especially for college enrollment and completion rates. Among the earliest evidence of a social-class gap in college attendance net of academic ability is data from the 1960s showing that even after controlling on standardized test performance, students from lower categories of socioeconomic status are much less likely to enter college

within five years of high school graduation. The talent loss due to unequal social-class background was estimated to be 50 percent of top-ability students who do not enter college from the lowest socioeconomic quartile, compared with a loss of only 5 percent of high-ability students from the highest socioeconomic quartile (U.S. Department of Health, Education and Welfare 1969). The importance of social-class factors for educational equity was reinforced by extensive research on Wisconsin high school students that included measures of race as well as student achievement on standardized tests. Social-class disparities in educational attainment net of academic ability were again in evidence, as it was reported that top-ability students were only half as likely to attend college or to graduate from college if they came from the lowest quarter, rather than the highest quarter, in socioeconomic status (Sewell and Hauser 1980). These studies also estimated that observed African-American–white differences in years of educational attainment can largely be accounted for by social-class differences between the racial groups.

However, race differences in students' achievement test performance are not so well explained by socioeconomic status alone. A study of seven national probability samples of adolescents from 1965 to 1996 described trends in black–white test score gaps after adjustments had been made for social class, family structure, and community variables. About a third of the test score gap is accounted for by racial differences in social class, so a major race gap remains in adjusted test scores; this gap has narrowed somewhat since 1965, although the rate of closure seems to have decreased or reversed since 1972. Group differences at the extremes of the distribution reveal contrasting gaps and the importance of social-class factors. Social-class-adjusted differences at the bottom of the distribution are closing more rapidly, especially in reading, but differences at the top of the achievement distribution are large and they are neither improving over time nor due to relative changes in social class (Hedges and Nowell 1999). This supports that argument that gaps in test scores are due to factors other than social class and family structure, such as discrimination, residential segregation, and the quality of schooling available to African Americans (Jaynes and Williams 1989).

**PROCESS DIFFERENCES IN OPPORTUNITIES**

Another approach to assessing equality of opportunity is to compare the attainment processes that link personal resources or investments to educational or occupational achievements for different social groups. Opportunities can be defined as unequal when the major avenues to advancement used by one group are not as effective for another. Researchers have frequently reported attainment process differences in the degree to which various population subgroups have been able to capitalize on advantages of family background or have experienced a high rate of return on investments in building relevant competencies or credentials. Some of this work has been criticized for possible shortcomings of methodology and interpretation.

Studies of social-group differences in an attainment process are important because they help to estimate the long-run prospects for closing existing achievement gaps (Featherman and Hauser 1978, Chap. 6). The prospects are positive if each subgroup has access to an attainment process that will translate improvements of personal and family resources into achievement outcomes, especially when programs and policies are available for investments in upgrading resources of groups that at present are weak. But if some groups are lagging in relevant skills and credentials, and are exposed only to attainment processes that provide poor returns in comparison with other groups, then the prospects are dim for closing existing gaps.

Studies of general social mobility processes have identified the special problems of African American males in translating any advantages from the family of origin into attainments in their own adult lives. For the white male population in this country, clear intergenerational processes have been evidenced in which sons can build upon a middle- or upper-class family background, as shown by the strong relationship between father's and son's occupational status for whites over many recent decades. In contrast, through the 1960s, African-American males have not been as able to capitalize on any family advantages in building their occupational careers, as shown by the weak relationship for intergenerational mobility and the frequency with which substantial proportions of African-American males from nonmanual or white-collar households are downwardly mobile

and unable to benefit from their family advantages. There is some indication that since the 1970s race differences for males in the opportunities to benefit from any inheritance of family social-class advantages have closed (Farley and Allen 1987; Featherman and Hauser 1978.)

Race differences in the processes of school effects on achievement have been reported in two national studies by the sociologist James S. Coleman and his research coworkers. In a 1966 national study of public schools, differences in school resources and learning environments were found to have larger average effects on African-American students' achievement than on white students' achievement (Coleman et al. 1966). The result was interpreted as a differential sensitivity of disadvantaged students to school improvements, because these students from poor families relied more on good schools for their development of academic skills. A similar race difference in educational processes was found in the 1980s with national data from public and private high schools (Coleman and Hoffer 1987). African Americans, Latinos, and students from low socioeconomic backgrounds were found to do better in Catholic schools than in public high schools, in terms of both higher test scores and lower dropout rates. It was argued that these students especially benefited from the greater academic demands that can be enforced by the sense of community established by Catholic schools, which compensates for family disadvantages of many of these students. Again, minority and disadvantaged students were found to be more responsive to changes in school environments that have effects on high school students' achievement and completion rates. Other researchers have questioned the recent results on the grounds that key student/family self-selection variables were not controlled in the analyses of public-Catholic school differences and that the sizes of the race interaction effects were not impressive by conventional statistical standards (Alexander and Pallas 1985; McPartland and McDill 1982).

Research has indicated that race inequalities are currently much less evident in educational attainment processes than in occupational attainment processes. Analyses using appropriate statistical tests of the processes that yield important educational achievements—such as additional years of schooling and scholastic outcomes, including

grades and test scores—have found great similarities between African Americans and whites (Gottfredson 1981; Wolfle 1985). Thus, not only have African-American-white differences in the frequency of high school graduation and college education been diminishing, the processes that link social background and school input variations to educational achievements have become very similar for African Americans and whites. At the same time, race gaps in school test scores have been closing more slowly, and serious disparities persist in the level of financing and concentration of single-race and disadvantaged student bodies in schools attended by racial minorities, even though education attainment processes would translate improvements of such inputs into attainments for African Americans (Jaynes and Williams 1989).

However, major race and sex differences continue in the occupational domain regarding both the processes of attainment and the gaps in achievement. Labor-market disparities by race and sex are much more apparent than differences in educational opportunities, but the disparities are exhibited in complex patterns or processes according to individuals' social-class position, labor-market location, career stage, and other factors (Farley and Allen 1987; Featherman and Hauser 1978; Jaynes and Williams 1989; Wilson 1987). The chances are equally good for African Americans and whites of each sex who are highly educated to gain entry to good jobs, but advancement opportunities to higher positions at later career stages are more likely to be missed by African Americans. At the same time, African-American male workers with less advanced credentials are much more likely to have periods of unemployment or reduced hours, and to be paid less when employed, than white males with equivalent years of schooling. The greatest race discrepancies are observed for poorly educated young African-American males, who are much more likely than comparable whites to be unemployed, to have dropped out of the labor force, or to report no annual earnings. William J. Wilson (1980) has developed a theory of the "declining significance of race" that considers the growing social-class gaps within the African-American population in occupational success, as well as the special difficulties faced by poorly educated African-American males in urban racial ghettos, whom he views as "the truly disadvantaged" (Wilson 1987).

### STRUCTURAL BARRIERS TO EDUCATIONAL OPPORTUNITIES

While careful studies of residual differences and attainment process differences can document the existence of unequal opportunities, other research is required on specific interactions and practices in schools or labor markets to understand the actual barriers that unfairly inhibit individuals because of their sex, race, or social-class position. For education, research on differential access to specific components of schooling, studies of tracking and grouping policies in elementary and secondary schools, and examinations of financial aid practices in higher education have identified some specific structural barriers in educational opportunities.

A landmark study was conducted in response to a congressional request under the 1964 Civil Rights Act and published in 1966 with the title *Equality of Educational Opportunity*. Also known as the "Coleman Report," after the sociologist James S. Coleman who directed the research, it was both influential and controversial for the way it examined educational opportunities and for its major findings (Coleman et al. 1966). Based on a large national survey of students and schools at both elementary and secondary levels, the Coleman Report collected the most comprehensive data available at that time on equity issues in education. It was not satisfied to compare only the average school input resources experienced by different race and ethnic groups—such as textbooks, libraries and laboratories, per pupil expenditures, teacher qualifications, or class size. The Coleman Report also considered race and ethnic differences on student outcomes as measured by standardized tests in major subjects, and asked how different school components contributed to student learning, in order to weigh inequalities of school inputs by their importance for student outcomes. The simultaneous examination of school inputs, student outcomes, and their relationships to one another had not been attempted before in assessing equity issues, and the published results have been a continuing source of reanalysis and reinterpretation.

The Coleman Report did find large differences in test scores between white and most racial and ethnic minority groups that existed from the time students began school and were not reduced,

on the average, as students moved from grade 1 to grade 12. These differences in student outcomes could not be explained by variations in the school input factors measured by the Coleman Report surveys, because within each region no great disparities of school inputs appeared for different racial and ethnic groups, and these factors did not relate strongly to student outcomes in any case after family background and social-class factors were statistically controlled. In fact, when school factors were combined into three clusters for analysis—(1) instructional materials and resources, (2) teacher and staff characteristics, and (3) student body composition—the most important component in accounting for variations in student test scores net of family background was the attributes of fellow students. Thus, the large observed group differences in student outcomes were not found to be accounted for by existing variations in conventional school and teacher components, although attending a school with fellow students who were college-bound did seem to make a positive contribution to the learning environment.

Subsequent investigations have shown that improvements in school factors, such as smaller class size, better-qualified teachers, and well-directed extra resources, can actually make a significant difference for student learning (Burtless 1996; Jencks and Phillips 1998). But the general picture drawn by the Coleman Report has been confirmed of an educational system that does little to reduce the large racial and ethnic differences in academic test scores with which students begin elementary grades (Jencks et al. 1972; Puma et al. 1997).

The Coleman Report data did not measure within-school differences in educational resources and learning environments, and consequently was unable to analyze major barriers to equal opportunities from specific internal school practices, such as tracking and ability grouping. Other research has shown that when students are tracked into separate programs or separate courses according to their earlier test scores or grades, those in the lower-level groups are likely to encounter serious barriers to their educational growth and progress. Lower tracks and lower-level courses have been shown to offer weaker educational resources, such as fewer expert teachers and poorer educational climates with lower academic expectations, that can lead to lower average student achievement test scores and decreased probabilities of completing

high school and continuing education in college (Gamoran 1986; Hallinan 1988; Oakes 1985). Tracking is now seen as a major barrier to equal educational opportunities because tracking and ability grouping are very common practices in American schools, and minorities and socioeconomically disadvantaged students are much more likely to be assigned to the lower-level programs and courses within their schools (Oakes 1985).

Moreover, the educational resources available at the school level are thought to be more unequal for minorities and disadvantaged students at the present time than they were found to be in the 1966 Coleman Report assessments (Smith and O'Day 1991). Since the 1960s, demographic trends have created greater concentrations of poverty in large urban schools, and changes in funding support for public education in central city districts have reduced those districts' relative ability to purchase adequate classroom supplies and materials, and to recruit and retain highly qualified teachers. In addition, trends in school desegregation, which produced increasing numbers of racially mixed schools with improved learning environments for minorities because of court decisions from 1954 through the 1970s, were reversed in the late 1980s and early 1990s, when school segregation grew in most regions and in the nation as a whole (Orfield and Eaton, 1996). Consequently, school-level barriers to equal educational opportunities have worsened since the 1960s, because the changing urban demographics and negative fiscal trends have dramatically altered the student body composition and the quality of the teaching staff that the Coleman Report found to be the most important factors of a good school.

Barriers have been identified in college educational opportunities, which also may have gotten worse, especially for African-American males, in recent years. Minorities have long been underrepresented as students at four-year colleges, in scientific major fields, and in obtaining advanced degrees (Trent and Braddock 1987). Some of these gaps had been closing through the 1970s, but since that time, uniform progress is no longer evident and some actual downturns in minority enrollments and attainments have been recorded (Jaynes and Williams 1989; Miller 1995; Wirt et al. 1998). African-American and Latino students often encounter special problems in pursuing college

programs because of insufficient social and academic support on campus or inadequate prior educational experiences (Green 1989). Recent reversals in minority enrollments have been explained by increasing tensions related to race and ethnicity on some college campuses and to changes from grants to loans in many financial assistance programs which poor students are less likely to receive or use (Blackwell 1990).

### **STRUCTURAL BARRIERS TO OCCUPATIONAL OPPORTUNITIES**

To help account for residual sex or race gaps in job success and in the career attainment process, research has identified specific structural barriers to sex equality and to racial equality in occupational opportunities.

Studies of the large average earnings differences between men and women workers show that very large gaps remain after statistically controlling on individual differences in input variables such as education and experience, but these gaps are substantially reduced by adding measures of each person's occupation or occupational group. This result indicates that sex gaps in earnings have much of their source in the extreme job segregation by sex in the American labor market—many occupations are primarily filled by women or primarily filled by men—and the wage levels are much lower for “female” occupations (Treiman and Hartmann 1981). Since fully two-thirds of men and women would have to change jobs to achieve similar representation of each sex across occupations, full enforcement of antidiscrimination laws against unequal pay for men and women in the same occupation can achieve only modest improvements in wage differentials by sex. Other suggested approaches to reducing sex segregation of jobs and associated wage gaps—such as enriching the socialization experiences toward a wider range of career exposures for children and youth of both sexes, or incorporating policies of “comparable worth” that establish wage rates by job features, irrespective of sex or race of incumbents (Hartmann 1985; Marini 1989)—have not yet made large inroads.

To specify how occupational opportunities continue to be unequal for racial or ethnic minorities, research has identified structural barriers at each stage of the occupational career process.

Barriers can appear at the job candidate stage, when employers are recruiting the pool of candidates for job openings; at the job entry stage, when an individual is actually selected to fill a vacancy; and at the job promotion stage, when transfers are made within a firm to fill spots at higher levels (Braddock and McPartland 1987; Feagin and Feagin 1978; Marini 1989).

At the job candidate stage, qualified minorities of either sex may fail to learn about many desirable job openings because they are excluded from useful social networks that provide others with information about and contacts for particular employment opportunities. Employers find job candidates more frequently from walk-ins and friends of current employees (the result of informal social networks) than any other recruitment means for lower- and middle-level jobs. The social contacts used by many minorities are racially segregated networks that on the average are not as well tied to good job information as the social networks available to whites. This barrier to equal opportunities at the job candidate stage is partially kept in place by the continued racial segregation of the schools and neighborhoods that create many social networks and by the underrepresentation of minorities in the upper levels of firms, where informal information for friends and relatives about job openings is often best acquired (Crain 1970; Rossi et al. 1974).

At the job entry stage, otherwise qualified minorities are often not selected because of barriers of statistical discrimination and information bias. Employers who do not wish to invest much to obtain extensive information about job applicants will often use a group identifier, such as sex or race, in hiring decisions when they believe that traits on which subgroups may differ statistically predict job performance. For example, such “statistical discrimination” can occur when an employer selects a white over a minority applicant for a job requiring good academic skills, based on a belief in average racial group differences on academic test scores rather than on actual individual candidates' differences in academic skills shown on tests administered or obtained by the employer during the screening process (Bielby and Baron 1986; Braddock et al. 1986; Thurow 1975).

Even when qualification data from individuals are relied upon in hiring decisions, other barriers

to equal opportunities occur due to “information bias” of data on minority candidates. References and recommendations from school or employment officials for African-American applicants may be viewed as less credible by white employers who are less familiar with an African-American school, a member of the African-American clergy, or an African-American firm, or who may be more wary of information provided by minority sponsors due to stigmas or stereotypes attached to these sources. Similarly, minority job applicants who grow up in communities that have high youth unemployment rates will be less able to satisfy prospective employers’ interests in previous employment experiences and references (Braddock and McPartland 1987).

At the job promotion stage, minorities may face unfair barriers due to internal recruitment methods or because they are poorly positioned within internal labor markets. However, findings from a national study indicate the potential benefits to minorities of seeking internal promotions: The average pay differential between African-American and white workers is less for jobs filled from inside a firm than for jobs filled from outside for individuals of the same sex and education level, suggesting that unfair selection is reduced when employers process information on applicants’ actual job performance within their firm. On the other hand, the same study showed that unless an internal vacancy is widely advertised within a firm, whites are more likely to be sought out for available promotions (Braddock and McPartland 1987). Moreover, research has shown that minorities are less likely to have entered a firm on a career ladder that ordinarily leads to promotion opportunities, so they may never be eligible to compete for advancement through an internal labor market (Rosenfield 1980).

### **POLICIES AND PRACTICES**

Governments and courts have established policies and practices in recent decades that are intended to eliminate race and sex discrimination and to ensure equality of opportunity. These range from the 1954 Supreme Court decision against segregated schools to the civil rights legislation of the 1960s and the executive orders to establish affirmative action guidelines in employment (Burstein 1985; Jaynes and Williams 1989). However, in the

final decades of the twentieth century, major court decisions have stepped back from considerations of race in school attendance patterns, college admissions policies, and employment selection practices (Orfield and Eaton 1996; Orfield and Miller 1998). At the same time, the expectation of high performance in elementary and secondary schools by students regardless of race or ethnicity has become widespread national and state policy, as a common-core academic curriculum is being mandated for all learners. It has yet to be determined how testing of students will combine with higher standards to influence access to educational opportunities and the gaps in dropout rates and achievement scores.

Although it is difficult to distinguish the effects of one governmental action from those of another in improving the life chances of women and minorities, clear advances have been made that can be attributed to the combined impacts of various public policies for equal rights. For example, from 1970 to 1990 the race gaps in academic test scores of schoolchildren decreased between 25 and 50 percent for different age groups (Smith and O’Day 1991). Reductions in the race gaps in terms of years of school completed have been dramatic, especially among female students. Greater equity is also evident in some labor-market behaviors, including the distribution of occupations by race within sex groups. On the other hand, inequalities in the distribution of income have grown in recent years, with a much higher share going to the top earners; average racial improvements are not evident in employment rates and income levels of adult males; some stagnation or reversals have occurred in upward trends of minority test scores and college attainment rates; and extensive racial segregation of housing and schooling remains a dominant feature of American life. The increasing diversity of minority groups in this country will amplify issues of equality of opportunity that concern language and cultural background differences in the population (Harrison and Bennett 1995).

Controversy continues to accompany further efforts to sustain current policies and to institute new practices for equal opportunities. The differences are most evident on whether outcome-based policies are required to overcome systemic barriers—for instance, affirmative action programs that use guidelines and timetables—or whether efforts

should concentrate only on intentional discrimination or on specific aspects of the processes that inhibit equal rights (Levinger 1987).

(SEE ALSO: *Affirmative Action*; *Discrimination*; *Education and Mobility*; *Ethnicity*; *Equity Theory*; *Social Mobility*; *Race*)

REFERENCES

Alexander, Karl A., and Aaron M. Pallas 1985 "School Sector and Cognitive Performance." *Sociology of Education* 58:115-127.

Bielby, William T., and James N. Baron 1986 "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91:759-799.

Blackwell, James E. 1990 "Current Issues Affecting Blacks and Hispanics in the Educational Pipeline." In Gail E. Thomas, ed., *U.S. Race Relations in the 1980s and 1990s*. New York: Hemisphere.

Braddock, Jomills H., II, and James M. McPartland 1987 "How Minorities Continue to Be Excluded from Equal Employment Opportunities: Research on Labor Market and Institutional Barriers." *Journal of Social Issues* 43(1):5-39.

—, et al. 1986 "Applicant Race and Job Placement Decisions: A National Survey Experiment." *International Journal of Sociology and Social Policy* 6:3-24.

Burstein, Paul 1985 *Discrimination, Jobs and Politics: The Struggle for Equal Employment Opportunity in the United States Since the New Deal*. Chicago: University of Chicago Press.

Burtless, Gary, ed. 1996 *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution Press.

—, and Thomas Hoffer 1987 *Public and Private High Schools*. New York: Basic Books.

— et al. 1966 *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.

Crain, Robert L. 1970 "School Integration and Occupational Achievement of Negroes." *American Journal of Sociology* 75:593-606.

Farley, Reynolds, and Walter R. Allen 1987 *The Color Line and the Quality of Life in America*. New York: Russell Sage Foundation.

Feagin, J. R., and C. B. Feagin 1978 *Discrimination American Style: Institutional Racism and Sexism*. Englewood Cliffs, N.J.: Prentice-Hall.

Featherman, David L., and Robert M. Hauser 1978 *Opportunity and Change*. New York: Academic Press.

Gamoran, Adam 1986 "Instructional and Institutional Effects of Ability Grouping." *Sociology of Education* 59:185-198.

Gottfredson, Denise C. 1981 "Black-White Differences in the Educational Attainment Process: What Have We Learned?" *American Sociological Review* 46:542-557.

Granovetter, Mark S. 1974 *Getting a Job: A Study of Contacts and Careers*. Cambridge, Mass.: Harvard University Press.

Green, Madeleine F. 1989 *Minorities on Campus: A Handbook for Enhancing Diversity*. Washington, D.C.: American Council on Education.

Hallinan, Maureen T. 1988 "Equality of Educational Opportunity." *Annual Review of Sociology* 14:249-268.

Harrison, Roderick J., and Claudette E. Bennett 1995 "Race and Ethnic Diversity." In Reynolds Farley, ed., *State of the Union: America in the 1990s: vol. 2. Social Trends*. New York: Russell Sage Foundation.

Hartmann, Heidi I., ed. 1985 *Comparable Worth: New Directions for Research*. Washington, D.C.: National Academy Press.

Hauser, Robert M. 1993 "The Decline in College Entry Among African Americans: Findings in Search of Explanations." In Paul M. Snideman, Philip E. Tetlock, and Edward G. Carmines, eds., *Prejudice, Politics, and the American Dilemma*. Stanford, Calif.: Stanford University Press.

Hedges, Larry V., and Amy Nowell 1999 "Changes in the Black-White Gap in Achievement Test Scores." *Sociology of Education* 72:111-135.

Jaynes, Gerald David, and Robin M. Williams, Jr., eds. 1989 *A Common Destiny: Blacks and American Society*. Washington, D.C.: National Academy Press.

Jencks, Christopher, and Meredith Phillips, eds. 1998 *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press.

Jencks, Christopher, et al. 1972 *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

Levinger, George (ed.) 1987 "Black Employment Opportunities: Macro and Micro Perspectives." *Journal of Social Issues* 43:1-156.

Levy, Frank 1995 "Incomes and Income Inequality." In Reynolds Farley, ed., *State of the Union: America in the 1990s: vol. 1. Economic Trends*. New York: Russell Sage Foundation.

Marini, Margaret Mooney 1989 "Sex Differences in Earnings in the United States." *Annual Review of Sociology* 15:343-380.

- McPartland, James M., and Edward L. McDill 1982 "Control and Differentiation in the Structure of American Education." *Sociology of Education* 55:65-76.
- Miller, L. Scott 1995 *An American Imperative: Accelerating Minority Educational Advancement*. New Haven, Conn.: Yale University Press.
- Oakes, Jeannie 1985 *Keeping Track*. New Haven, Conn.: Yale University Press.
- Oliver, Melvin L., and Thomas M. Shapiro 1995 *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge.
- Orfield, Gary, and Susan E. Eaton 1996 *Dismantling Desegregation*. New York: New Press.
- Orfield, Gary, and Edward Miller, eds. 1998 *Chilling Admissions: The Affirmative Action Crisis and the Search for Alternatives*. Cambridge, Mass.: Harvard Education Publishing Group.
- Puma, M.J., N. Karweit, Price C., Ricciuti, A., Thompson, W., and Vaden-Kiernan, M. 1997. *Prospects: Final report on student outcomes*. Bethesda, MD: Abt Associates, Inc.
- Rosenfield, Rachel A. 1980 "Race and Sex Differences in Career Dynamics." *American Sociological Review* 45:583-609.
- Rossi, Peter H., Richard A. Berk, and Betty K. Eidson 1974 *The Roots of Urban Discontent: Public Policy, Municipal Institutions and the Ghetto*. New York: Wiley.
- Sewell, William H., and Robert M. Hauser 1980 "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." In Alan C. Kerckhoff, ed., *Research in Sociology of Education and Socialization*, vol. I. Greenwich, Conn.: JAI Press.
- Smith, Marshall S., and Jennifer O'Day 1991 "Educational Equality: 1966 and Now." In Deborah Versteegen, ed., *Spheres of Justice in American Schools*. New York: Harper Business.
- Thurow, Lester 1975 *Generating Inequality*. New York: Basic Books.
- Treiman, Donald J., and Heidi I. Hartmann, eds. 1981 *Women, Work and Wages: Equal Pay for Jobs of Equal Value*. Washington, D.C.: National Academy Press.
- Trent, William T., and Jomills Henry Braddock II 1987 "Trends in Black Enrollment and Degree Attainment." In John B. Williams, ed., *Title VI Regulation of Higher Education: Problems and Progress*. New York: Teachers College Press.
- U.S. Department of Health, Education and Welfare 1969 *Toward a Social Report*. Washington, D.C.: U.S. Government Printing Office.
- Wilson, William Julius 1980 *The Declining Significance of Race: Blacks and Changing American Institutions*, 2nd ed. Chicago: University of Chicago Press.
- 1987 *The Truly Disadvantaged: The Inner City, the Underclass and Public Policy*. Chicago: University of Chicago Press.
- 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.
- Wirt, John, Tom Snyder, Jennifer Sable, Susan P. Choy, Yupin Baes, Janis Stennett, Allison Gruner, and Marianne Perie 1998. *The Condition of Education 1998*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Wofle, Lee M. 1985 "Postsecondary Educational Attainment Among Whites and Blacks." *American Educational Research Journal* 22:501-525.

JOMILLS HENRY BRADDOCK II  
JAMES M. MCPARTLAND

## EQUILIBRIUM THEORY

See Cognitive Consistency Theories; Social Dynamics.

## EQUITY THEORY

See Social Justice.

## ETHICS IN SOCIAL RESEARCH

The immediacy of subject matter in social science underscores the importance of ethical issues in research by social scientists. This is particularly true in sociology. A rather small percentage of sociologists use historical documents or cultural products as data. The majority rely upon interviews with actively cooperating subjects, records relating to persons still living or recently alive, unobtrusive observation of live actors, or participant studies within interacting groups. Sociological research typically focuses on relatively large study populations and poses questions relevant to many dimensions of individual and social life. Both the process and application of sociological inquiry may conceivably affect large numbers of subjects in an adverse manner. Thus, the question of "right" and "wrong" in research has been a continual (though not always powerful or explicit) concern within the profession.



Ethics may be conceptualized as a special case of norms governing individual or social action. In any individual act or interpersonal exchange, ethics connotes principles of obligation to serve values over and above benefits to the people who are directly involved. Examination of ethical standards in any collectivity provides insights into its fundamental values; identification of ethical issue provides clues to its basic conflicts. This is as true of sociology as a profession as it is of other social systems.

The most abstract and general statements about ethics in sociological literature reflect broad agreement about the values that social inquiry should serve. Bellah (1983) writes that ethics constitutes an important, though typically implicit, topic in the thinking of sociology's founders (such as Durkheim and Weber) and leading modern practitioners (such as Shils and Janowitz). Even while consciously striving to distinguish their emerging discipline as a science free of values and moralizing, the early sociologists appeared to have a distinct ethical focus. The discipline's founders implied and sometimes stated that sociology necessarily involved ethical ends, such as identification of emerging social consensus or the development of guidelines for assessing social good. Modern sociologists have emphasized improvement of society's understanding of itself as the discipline's principal ethical end, as opposed to determining a specific direction or developing technology for social change. In the broadest sense, contemporary sociologists seem to consider the raising of consciousness as quintessentially ethical activity and social engineering by private or parochial interests as ethically most objectionable. In the phraseology of Edward Shils, this means contributing to "the self-understanding of society rather than its manipulated improvement" (Shils 1980, p. 76).

Dedication to advancement of society's understanding of itself through diverse scientific approaches may comprise the fundamental ethic of sociology. A *Code of Ethics* published by the American Sociological Association (ASA) in 1989 (American Sociological Association 1989) gave concrete expression to this ethic. Concentrating primarily on research, the *Code of Ethics* emphasized three specific areas of concern: (1) full disclosure of motivations for and background of research; (2) avoidance of material harm to research subjects, with special emphasis on issues of confidentiality;

and (3) qualifications to the technical expertise of sociology.

The first area appeared concerned primarily with a fear among sociologists that agencies of social control (such as military or criminal justice units) may seek intelligence under the guise of social research. Thus, the code advised sociologists not to "misuse their positions as professional social scientists for fraudulent purposes or as a pretext for gathering intelligence for any organization or government." The mandate for disclosure has implications involving relations among professionals, between professionals and research subjects, and between professionals and the public. Another provision of the code read, "Sociologists must report fully all sources of financial support in their publications and must note any special relation to any sponsor." (p. 1)

The second area of concern in the code placed special emphasis on assurance of confidentiality to research subjects. It stressed the need for extraordinary caution in making and adhering to commitments. As if to recognize the absence of legal protection for confidentiality in the research relationship and to mandate its protection nevertheless, the code stated: "Sociologists should not make any guarantees to respondents, individuals, groups, or organizations—unless there is full intention and ability to honor such commitments. All such guarantees, once made, must be honored" (p. 2).

As a subject of professional ethics, the third area is extraordinary. Provisions mandating disclosure of purpose and assurance of confidentiality might appear in the code of ethics of any profession dealing regularly with human clients or subjects. But it is surprising to find, as a provision in the 1989 ASA *Code of Ethics*, the mandate that sociologists explicitly state the shortcomings of methodologies and the openness of findings to varying interpretation. The following quote illustrates provisions of this nature:

*Since individual sociologists vary in their research modes, skills, and experience, sociologists should always set forth ex ante the limits of their knowledge and the disciplinary and personal limitations that condition the validity of findings. To the best of their ability, sociologists should . . . disclose details of their*

*theories, methods and research designs that might bear upon interpretation of research findings. Sociologists should take particular care to state all significant qualifications on the findings and interpretations of their research.* (p. 2)

Themes in the 1989 *Code of Ethics* dealing with disclosure and confidentiality reflect widely shared values and beliefs in the profession. Historically, sociology has stood out among the learned professions as critical of the authority of established institutions such as governments and large business firms. But propositions about the limitations of theories and methodologies and the openness of findings to varying interpretation suggest conflict. In the late twentieth century, sociological methodologies encompassed both highly sophisticated mathematical modeling of quantitative data and observation and theory building based entirely on qualitative techniques. Acknowledgment of the legitimacy of these differences in an *ethical* principle reflects a strenuous attempt by sociology as a social system to accommodate subgroups whose basic approaches to the discipline are inconsistent with each other in important respects.

A more recent formulation of the ASA *Code of Ethics*, published in 1997 (American Sociological Association 1997), restates the basic principals of serving the public good through scientific inquiry and avoiding harm to individuals or groups studied. But a shift in emphasis appears to have occurred. The 1989 *Code* explicitly cited the danger of governmental or corporate exploitation of the sociologist's expertise. The 1997 *Code*, though, stresses ethical challenges originating primarily from the researcher's personal objectives and decisions.

The 1997 *Code of Ethics*, for example, contains a major section on conflict of interest. According to this section, "conflicts of interest arise when sociologists' *personal or financial* interests prevent them from performing their professional work in an unbiased manner" (p. 6; emphasis added). A brief item on "disclosure" asserts an obligation by sociologists to make known "relevant sources of financial support and relevant personal or professional relationships" that may result in conflicts of interest vis-a-vis to employers, clients, and the public (p. 7).

The two most extensive sections in the 1997 *Code* are those on confidentiality and informed consent. The directives addressing confidentiality place extraordinary responsibility on the individual sociologist. Pertinent language states that "confidential information provided by research participants, students, employees, clients, or others is treated as such by sociologists *even if there is no legal protection or privilege to do so*" (emphasis added). The *Code* further instructs sociologists to "inform themselves fully about all laws and rules which may limit or alter guarantees of confidentiality" and to discuss "relevant limitations on confidentiality" and "foreseeable uses of the information generated" with research subjects (p. 9). It is recommended that information of this kind be provided at the "outset of the relationship." Sociologists are neither absolutely enjoined from disclosing information obtained under assurances of confidentiality nor given clear guidance about resolving pertinent conflicts. The *Code of Ethics* states:

*Sociologists may confront unanticipated circumstances where they become aware of information that is clearly health- or life-threatening to research participants, students, employees, clients, or others. In these cases, sociologists balance the importance of guarantees of confidentiality with other priorities in [the] Code of Ethics, standards of conduct, and applicable law.* (p. 9)

The section on informed consent, the most extensive in the 1997 *Code of Ethics*, reflects a frequent dilemma among sociologists. The basic tenets of informed consent as stated here approximate those in all fields of science. Obtaining true consent requires eliminating any element of undue pressure (as might occur in the use of students as research subjects) or deception regarding the nature of the research or risks and benefits associated with participation. In social research, however, statement of the objectives of an investigation may affect attitudes and behavior among research subjects in a manner that undermines the validity of the research design. Recognizing this possibility, the *Code* acknowledges instances when deceptive techniques may be acceptable. These include cases where the use of deception "will not be harmful to research participants," is "justified by the study's prospective scientific, educational, or applied value," and cannot be substituted for by alternative procedures (p. 12).

A review of historical developments, events, and controversies of special importance to sociologists in the decades preceding the 1989 and 1997 *Codes of Ethics* promotes a further appreciation of the concerns they embody. Perhaps the most far-reaching development in this era was the introduction of government funding into new areas of the sociological enterprise. In sociology, as in many areas of science, government funding provided opportunities to expand the scope and sophistication of research, but it created new ethical dilemmas and accentuated old ones.

Increased government funding created interrelated problems of independence for the sociological researcher and anonymity for the research subject. A report by Trend (1980) on work done under contract with the U.S. Department of Housing and Urban Development (HUD) illustrates one aspect of this problem. Possessing a legal right to audit HUD's operations, the General Accounting Office (GAO) could have examined raw data complete with individual identifiers despite written assurances of confidentiality to the subjects by the research team. Sensitivity on the part of the GAO and creativity by the sociologists averted an involuntary though real ethical transgression in this instance. But the case illustrates both the importance of honoring commitments to subjects and the possibility that ethical responsibilities may clash with legal obligations.

Legal provisions designed explicitly to protect human subjects emerged in the 1970s. Regulations developed by the U.S. Department of Health and Human Services (DHHS) require that universities, laboratories, and other organizations requesting funds establish institutional review boards (IRBs) for protection of human subjects. The 1997 ASA *Code of Ethics* makes frequent reference to these boards as a resource for resolution of ethical dilemmas.

Sociologists, however, have not always expressed confidence in the contributions of IRBs. One commentary (Hessler and Freerks 1995) argues that IRBs are subject to great variability in protecting the rights of research subjects at the local level. Others contend that deliberations of these boards take place in the absence of appropriate standards or methods of analysis. The expertise and concerns of IRBs may not apply well to

actual risks posed by sociological research methods. Biomedical research, the primary business of most IRBs, potentially poses risks of physical injury or death to the research subject. Except in extraordinary circumstances, sociological techniques expose subjects at worst to risks of embarrassment or transient emotional disturbance. IRB requirements often seem inappropriate or irrelevant to sociology. In the words of one commentator, the requirement by IRBs that researchers predict adverse consequences of proposed studies encourages sociologists to engage in exercises of "futility, creativity, or mendacity" (Wax and Cassell 1981, p. 226).

Several instances of highly controversial research have helped frame discussion of ethics among sociologists. Perhaps most famous is the work of Stanley Milgram (1963), who led subjects to believe (erroneously) that they were inflicting severe pain on others in a laboratory situation. This experiment, which revealed much about the individual's susceptibility to direction by authority figures, was said by some to present risk of emotional trauma to subjects. Milgram's procedure itself seemed to duplicate the manipulative techniques of authoritarian dictators. Distaste among sociologists for Milgram's procedure helped crystallize sentiment in favor of public and professional scrutiny of research ethics.

The Vietnam era saw increasing suspicion among sociologists that government might use its expertise to manipulate populations both at home and abroad. A seminal event during this period was the controversy over a U.S. Army-funded research effort known as Project Camelot. According to one commentator, Project Camelot aimed at ascertaining "the conditions that might lead to armed insurrections in . . . developing countries so as to enable United States authorities to help friendly governments eliminate the causes of such insurrections or to deal with them should they occur" (Davison 1967, p. 397). Critical scrutiny by scholars, diplomats, and congressional committees led to cancellation of the project. But provisions in the 1989 *Code of Ethics* on disclosure and possible impacts of research clearly reflect its influence.

The end of the Cold War and increasing litigiousness among Americans may help explain the shift in emphasis between the 1989 and 1997

ASA *Codes of Ethics*. As noted above, the later *Code* appears to emphasize ethical issues facing sociologists as individuals rather than as potential tools of government and big business. Many sociologists have stories to tell about actual or potential encounters with the legal system over the confidentiality of data obtained from research subjects. The visibility and frequency of such encounters may have helped shape the 1997 *Code's* section on confidentiality.

The most celebrated confrontation of a sociologist with the law involved Rik Scarce, who was incarcerated for 159 days for refusing to testify before a grand jury investigating his research subjects. Scarce's case is described by Erikson (1995):

*Scarce found himself in an awful predicament. He was engaged in research that rested on interviews with environmental activists, among them members of the Animal Liberation Front. One of his research subjects came under investigation in connection with a raid on a local campus, and Scarce was ordered to appear before a grand jury investigation. He refused to answer questions put to him, was found to be in contempt, and was jailed for more than five months.*

Some evidence suggests that the institutional structure surrounding social research has proven an uncertain asset in personal resolution of ethical issues such as Scarce's. The 1997 ASA *Code of Ethics* advises sociologists confronting dilemmas regarding informed consent to seek advice and approval from institutional review boards or other "authoritative [bodies] with expertise in the ethics of research." But IRBs typically serve as reviewers of research plans rather than consultative bodies regarding issues encountered in execution of research; the phrase "authoritative [bodies] with expertise in the ethics of research" has a vague ring. Lee Clark's (1995) description of his search for guidance in responding to a law firm's request for his research notes illustrates the limitations of IRBs and related individuals and agencies:

*. . . I talked with first amendment attorneys, who said academic researchers don't enjoy journalists' protections. . . . I was told that if I destroyed the documents, when there was reason to expect a subpoena, then I would be held in contempt of court. I talked with ASA officials and the chair of ASA's Ethics Committee, all*

*sympathetic but unable to promise money for an attorney. They were equally certain of my obligations according to the Ethics Code. . . . I talked with lawyers from Stony Brook [where Clark had performed his research], who told me that the institution would not help. Lawyers for Rutgers, where I was . . . employed, said they wouldn't help either.*

In all human activity, individuals ultimately face ethical issues capable of resolution only through personal choice among alternatives. But increasingly, sociologists seem to face these choices unaided by distinct guidelines from their profession. This default to personal responsibility derives in part from the ambiguity in two philosophical principles widely encountered in sociological discourse, utilitarianism and moral relativism.

As an ethical principle, utilitarianism seems to provide a convenient rule of making decisions. The prevailing morality among modern cosmopolitans, utilitarianism applies the principle of the greatest net gain to society in deciding questions of research ethics. This perspective places emphasis on degrees of risk or magnitude of harm that might result from a given research effort. Under this perspective, Project Camelot (cited above) may have deserved a more favorable reception. Davison (1967) suggests that completion of the project would probably not have caused appreciable harm. He comments:

*If past experience is any guide, it would have contributed to our knowledge about developing societies, it would have enriched the literature, but its effects on this country's international relations would probably have been tangential and indirect. (p. 399)*

Several well-known and ethically controversial studies may be justified on utilitarian grounds. Among the best known is Laud Humphreys's study of impersonal sex in public places (1975). Humphreys gained access to the secret world of male homosexuals seeking contacts in public restrooms by volunteering his services as a lookout. Despite its obvious deception, Humphreys's work received the support of several homophile organizations (Warwick 1973, p. 57), in part because it illustrated the prevalence of sexual preferences widely considered abnormal. In his study of mental institutions, Rosenhan (1973) placed normal (i.e.,

nonpsychotic) observers in psychiatric wards without the knowledge or consent of most staff. His study generated highly useful information on the imperfections of care in these institutions, but the deception and manipulation of his subjects (hospital staff) are undeniable.

As a rule for making decisions, though, utilitarianism presents both practical and conceptual problems. Bok (1978) points out the difficulty in estimating risks of harm (as well as benefits) from any research activity. The subtle and uncertain impacts of sociological research techniques (as well as associated findings) make prospective assessment of utilitarian trade-offs extremely problematical. Many traditional ethical constructs, moreover, contradict utilitarianism, implying that acts must be assessed on the basis of accountability to abstract principles and values (e.g., religious ones) rather than the practical consequences of the acts themselves.

Moral relativism provides some direction to the uncertainty implicit in utilitarianism. This principle assumes that "there are no hard and fast rules about what is right and what is wrong in all settings and situations" (Leo 1995). Under this principle, ethical judgment applies to ends as well as means. Moral relativism might provide ethical justification for a sociologist who, believing that the public requires greater knowledge of clandestine police practices, misrepresents his personal beliefs or interests in order to observe these practices. The very relativism of this principal, however, invites controversy.

The 1997 ASA *Code of Ethics* restates the profession's fundamental ethic as striving to "contribute to the public good" and to "respect the rights, dignity, and worth of all people" (p. 4). Regarding research activity, the *Code* places primary emphasis on informed consent, protection of subjects from harm, confidentiality, and disclosure of conflicts of interests. But the *Code*, the institutional milieu of sociology, and the practical conditions under which sociological research takes place preclude strong direction for individuals in the ethical dilemmas they encounter.

## REFERENCES

American Sociological Association 1989 *Code of Ethics*. Washington, D.C.: Author.

— 1997 *Code of Ethics*. Washington, D.C.: Author.

Bellah, R. N. 1983 "The Ethical Aims of Sociological Inquiry." In N. Haan, R. N. Bellah, P. Rabinow, and E. M. Sullivan, eds., *Social Science as Moral Inquiry*. New York: Columbia University Press.

Bok, S. 1978 "Freedom and Risk." *Daedalus* 107 (Spring):115-127.

Clark, L. 1995 "An Unethical Ethics Code?" *The American Sociologist* 26(2):12-21.

Davison, W. P. 1967 "Foreign Policy." In P. F. Lazarsfeld, W. H. Sewell, and H. L. Wilensky, eds., *The Uses of Sociology*. New York: Basic Books.

Erikson, K. 1995 "Commentary." *The American Sociologist* 26(2):4-11.

Hessler, R. M., and K. Freerks 1995 "Privacy Ethics in the Age of Disclosure: Sweden and America Compared." *The American Sociologist* 26(2):35-53.

Humphreys, L. 1975 *Tearoom Trade: Impersonal Sex in Public Places*. Chicago: Aldine.

Leo, R. A. 1995 "The Ethics of Deceptive Research Roles Reconsidered: A Response to Kai Erikson." *The American Sociologist* 27(1):122-128.

Milgram, S. 1963 "Behavioral Study of Obedience." *Journal of Abnormal and Social Psychology* 67:371-378.

Rosenham, D. L. 1973 "On Being Sane in Insane Places." *Science* 179 (January 1973):250-258.

Shils, E. 1980 *The Calling of Sociology: Essays on the Pursuit of Learning*. Chicago: University of Chicago Press.

Trend, M. G. 1980 "Applied Social Research and the Government: Notes on the Limits of Confidentiality." *Social Problems* 27:343-349.

Warwick, D. P. 1973 "Tearoom Trade: Ends and Means in Social Research." *The Hastings Center Studies* 1:27-38.

Wax, M. L., and J. Casell 1981 "From Regulation to Reflection: Ethics in Social Research." *The American Sociologist* 16:224-229.

HOWARD P. GREENWALD

## ETHNIC CLEANSING

See Genocide.

## ETHNICITY

Ethnicity is a salient feature of numerous societies throughout the world. Few societies are ethnically

homogeneous, even when they proclaim themselves to be. Consequently, ethnicity has been a preoccupation of sociologists since the early days of the discipline (although more so in the United States than elsewhere).

Yet there is not complete agreement on how the subject should be defined. In the past, it was common to highlight cultural difference as an essential feature of ethnic distinctiveness (see, e.g., van den Berghe 1967). Recently, this has lost favor on the grounds that cultural differences may vary from one setting to another and from one historical period to another. Following an approach attributed to Frederik Barth (1969), recent definitions have therefore focused on the existence of a recognized social boundary. But still among the most useful definitions is the classic one of Max Weber ([1922] 1968): An ethnic group is one whose members “entertain a subjective belief in their common descent because of similarities of physical type or of customs or both, or because of memories of colonization and migration” Weber adds insightfully, “It does not matter whether or not an objective blood relationship exists” (p. 389).

Despite definitional disagreements, there is general recognition that a number of characteristics appear as hallmarks of ethnicity; not all of them will be present in every case, but many will be. They include features shared by group members, such as the same or similar geographic origin, language, religion, foods, traditions, folklore, music, and residential patterns. Also typical are special political concerns, particularly with regard to a homeland; institutions (e.g., social clubs) to serve the group; and a consciousness of kind, or sense of distinctiveness from others (for the full listing, see Thernstrom et al. 1980, p. vi).

There is controversy over whether race should be viewed as a form of ethnicity. In this context, “race” should not be understood as a bundle of genetically determined traits that of themselves generate social differences—a view that has been repudiated by the vast majority of social scientists—but as a kind of social classification used by members of a society. Many scholars distinguish between ethnicity and race. For example, Pierre van den Berghe (1967) defines race as a social classification based on putative physical traits and ethnicity as a classification based on cultural ones

(see also Omi and Winant 1994). The contrast between the two can also be formulated in terms of volition vs. external constraint, with racial categories seen as more imposed by outsiders and ethnic ones as more claimed by the group members themselves (see, e.g., Waters 1990). But equally commonly, race is seen as a variant of ethnicity: A racial group is, then, an ethnic group whose members are believed, by others if not also by themselves, to be physiologically distinctive. This is the approach adopted in this article for several reasons. Not only do racial groups typically have the characteristics of ethnic groups (e.g., cultural distinctiveness), but many seemingly nonracial ethnic groups may also be believed to possess some distinctive physical features (e.g., the olive skin tone of Italians). The distinction between the two types of groups is therefore not hard and fast, a conclusion that is underscored by the historical transmutation of some racial groups into nonracial ones, for example, the Irish (Ignatiev 1995; Roediger 1991).

Sociologists recognize that the imprint of history on the contemporary ethnic relations of any society is deep, and this gives rise to another distinction that is potentially central to any discussion of ethnicity. It pertains to the mode of entry of a group into a society, and it has been formulated by Stanley Lieberson (1961; see also Blauner 1972) in terms of the situation that obtains just after contact between an indigenous group and one migrating into an area. One possibility is that the migrant group dominates, typically through conquest (often aided by the introduction of new diseases)—this is exemplified in the contacts between indigenes and European settlers in Australia and the United States. The other is that the indigenous group dominates, as occurred during the century of mass immigration (1830–1930) to the United States. The crux of the matter here is whether a group is incorporated into a society through force or through more or less voluntary migration. Lieberson argues that a group’s mode of entry is fateful for its trajectory of development in a society, and this is amply borne out in the literature on ethnicity.

Stated in very broad terms, three approaches have dominated the sociological study of ethnicity; a fourth and much newer one appears to hold the promise of eventual parity with them. Of the older

approaches, one, the *assimilation* perspective, focuses on social processes and outcomes that tend to dissolve ethnic distinctions, leading to the assimilation of one ethnic group by another or by the larger society. The second approach could be labeled as *stratification*. As the name implies, it addresses the origins and consequences of inequalities of various kinds among ethnic groups. The third approach focuses on *ethnic-group resources*, and encompasses processes, such as mobilization and solidarity, by which the members of ethnic groups attempt to use their ethnicity to compete successfully with others. The newest approach is a *social constructionist* one. It stems from the recognition that ethnic boundaries are malleable and is concerned with the ways by which such boundaries are created, maintained, and transformed.

No one of these approaches could be described as preeminent; each is a major presence in contemporary research on ethnicity and has shown that it has something to contribute. Other approaches are possible but are not as theoretically and empirically developed as these four. One other approach seeks a basis for ethnicity in sociobiology, viewing it as a form of genetic nepotism, a generalization of the universal tendency among animals to favor kin. van den Berghe (1981) has been an exponent of such an approach, but as yet no body of evidence has been developed to distinguish it from more sociological approaches; other sociologists have not followed his lead. Ethnicity has also been viewed as “primordial,” deriving from deeply seated human impulses and needs that are not eradicated by modernization (Isaacs 1975). But this viewpoint has not led to sociologically interesting research, and it has lacked exponents in recent decades. Another attempt, stemming from “rational choice theory” (Banton 1983; Hechter 1987), seeks to explain ethnic phenomena in terms of the efforts of individuals to maximize their advantages (or, in technical language, utilities). Research using rational choice theory is still too immature to draw up a meaningful balance sheet on it.

The assimilation approach has deep roots in classical social theory as well as in American sociology, where it is often traced to Robert E. Park’s 1926 formulation of a race relations cycle of “contacts, competition, accommodation, and eventual assimilation” (quoted from Park 1950, p. 150; cf. Lal 1990). The canonical statement of the assimilation approach is by Milton Gordon (1964; for an

updated revision, see Alba and Nee 1997). Although Gordon was addressing the role of ethnicity in the United States, his formulation is so general that it can be readily applied to other societies. At the heart of his contribution is the recognition that assimilation is a multidimensional concept. He distinguished, in fact, among seven types of assimilation, but the critical distinction lies between two of them: acculturation and structural (or social) assimilation. *Acculturation* means the adaptation by an ethnic group of the cultural patterns of the dominant, or majority, group. Such acculturation encompasses not only external cultural traits, such as dress and language, but also internal ones, such as beliefs and values. Gordon (1964) theorized that acculturation is typically the first of the types of assimilation to occur and that the stage of “acculturation only” may continue indefinitely” (p. 77)—hence the importance of the second assimilation type, structural assimilation. *Structural assimilation* is defined by Gordon to mean the entry of an ethnic group’s members into close, or primary, relationships with members of the dominant group (or, at least, with ethnic outsiders). The cardinal hypothesis in Gordon’s (1964) scheme is that structural assimilation is the key that unlocks all other types: “Once structural assimilation has occurred . . . all of the other types of assimilation will naturally follow” (Gordon 1964, p. 81). Once structural assimilation occurs, the way is open to widespread intermarriage, an abating of prejudice and discrimination, and the full participation of ethnic-group members in the life of a society.

Gordon discussed certain models, or theories, of the assimilation process (they might also be described as ideologies because of their value-laden character). Although these were, again, developed for the U.S. context, Gordon’s discussion is so lucid that the models have passed into more general application. One is labeled as “Anglo-conformity” by Gordon, and it describes an assimilation that is limited to acculturation to the behavior and values of the core ethnic group, taken to be Protestants with ancestry from the British Isles in the American context. A second model is that of the “melting pot.” It envisions an assimilation process that operates on cultural and structural planes. One outcome is a culture that contains contributions from numerous ethnic

groups and is adopted by their members. A parallel outcome on a structural plane is a pattern of widespread marriage across ethnic lines, in which the members of all ethnic groups participate and which leads ultimately to population made up of individuals of quite intermixed ancestry. The melting-pot idea corresponds with some popular notions about U.S. society, but so does the last model explicated by Gordon—namely, “cultural pluralism.” Cultural pluralism corresponds with a situation in which ethnic groups remain socially differentiated, often with their own institutions and high rates of in-group marriage, and retain some culturally distinctive features. It is, in fact, an apt description of many societies throughout the world.

Urban ecology, dating back to the origins of the Chicago School of American sociology, is quite compatible with the assimilation approach. A core tenet of this tradition is that the spatial separation of one group from another mirrors the social distance between them and changes as this does. This ultimately implies a model of *spatial assimilation* (Massey 1985), according to which residential mobility follows from the acculturation and socioeconomic mobility of ethnic-group members and is often an intermediate step on the way to more complete—that is, structural—assimilation. The model envisions an early stage of residential segregation, as the members of ethnic groups—typically, immigrants and their children—are concentrated in urban enclaves, which frequently results in the displacement of other groups. But as the members of an ethnic group acculturate and establish themselves in the labor markets of the host society, they attempt to leave behind less successful co-ethnics and to convert socioeconomic and assimilation progress into residential gain by “purchasing” residence in places with greater advantages and amenities. This process implies, on the one hand, a tendency toward dispersion of an ethnic group, opening the way for increased contact with members of the ethnic majority, and, on the other hand, greater resemblance in terms of residential characteristics between successful ethnic-group members and their peers from the majority group.

The assimilation perspective has been successfully applied to American ethnic groups derived from European immigration. In a review of the evidence, Charles Hirschman (1983) documents the abating of ethnic differences in the white

population in terms of socioeconomic achievement, residential location, and intermarriage. To cite some representative research findings, Stanley Lieberson and Mary Waters (1988), comparing the occupations of men of European ancestry in 1900 and 1980, find a marked decline in occupational concentrations, although these still show traces of the patterns of the past. These authors and Richard Alba (1990) also demonstrate the great extent to which interethnic marriage now takes place within the white population: Three of every four marriages in this group involve some degree of ethnic boundary crossing. The assimilation perspective as applied to European Americans has not been without its critics; Andrew Greeley (1971) has done the most to assemble evidence of persisting ethnic differences.

There can be no doubt that African Americans do not exemplify the patterns expected under the assimilation perspective (e.g., Massey and Denton 1993). A question that now motivates much debate and research is the degree to which the assimilation patterns will be found among contemporary immigrants and their descendants (see Alba and Nee 1997; Portes and Rumbaut 1996). It is too early to answer the question, but reflecting on the potential differences between past and contemporary immigrations, Portes and Zhou (1993) have added a new concept to the ethnicity arsenal—namely, *segmented assimilation*, which acknowledges that the strata of American society into which individuals assimilate may vary considerably. In their view, the children of dark-skinned, working-class immigrants who grow up in the inner city are at great risk of assimilating into the indigenous lower class and thus experiencing assimilation with little or no upward mobility.

Much of the evidence on assimilation and ethnic change is derived from cross-sectional studies rather than those done over time; the latter are difficult to conduct because of the limited availability of comparable data from different time points. Cross-sectional analyses involve some dissection of an ethnic group into parts expected to display a trajectory of change. One basis for such a dissection is generational groups. *Generation* here refers to distance in descent from the point of entry into a society. (By convention, generations are numbered with immigrants as the “first,” so that their children are the “second,” their grandchildren are



the “third,” etc.) Generally speaking, later generations are expected to be more assimilated than earlier ones. Another basis for dissection is *birth cohorts*, defined as groups born during the same period. Cohort differences can provide insight into historical changes in a group’s position. Both kinds of differences have been used to study ethnic changes in the United States (for an application of the generational method, see Neidert and Farley 1985; for the cohort method, see Alba 1988).

The second major approach to the study of ethnicity and race, labeled above as “stratification,” is considerably less unified than the assimilation approach, encompassing quite diverse theoretical underpinnings and research findings. Yet there are some common threads throughout. One is an assumption that ethnic groups generally are hierarchically ordered: There is typically a dominant, or *superordinate*, group, which is often described as the *majority* group (even though in some societies, such as South Africa, it may be a numerical minority of the population). There are also *subordinate* groups, often called *minorities* (although they may be numerical majorities). Second, these groups are assumed to be in conflict over scarce resources, which may relate to power, favorable occupational position, educational opportunity, and so forth. In this conflict, the dominant group employs a variety of strategies to defend or enhance its position, while minority groups seek to challenge it. Often, the focus of the stratification approach is on the mechanisms that help preserve ethnic inequalities, although there has also been some attention to the means that enable minorities successfully to challenge entrenched inequality.

One tradition in ethnic stratification research has looked to mechanisms of inequality that are rooted in ideologies and outlooks that are then manifested in the behavior of individuals. This is, in fact, a common meaning for the word *racism*. A longstanding research concern has been with *prejudice*, which is generally defined as a fixed set of opinions, attitudes, and feelings, usually unfavorable, about the members of a group (Allport 1954). Prejudice is frequently an outgrowth of *ethnocentrism*, the tendency to value positively one’s own group and denigrate others. It can lead to *discrimination*, which is a behavior: the denial of equal treatment to a group’s members, exemplified by the refusal to sell homes in certain neighborhoods to minority-group members. The investigation of

prejudice was one of the early testing grounds for survey research. In the United States, this research uncovered a dimension of *social distance*, expressing the specific gradations of social intimacy the majority is willing to tolerate with the members of various ethnic groups (Bogardus 1928). Recent research has revealed a paradoxical set of changes: on the one hand, a secular decline in traditional prejudice, most notably the prejudiced attitudes and beliefs held by whites about blacks; on the other, little increase in support for government policies that implement principles of racial equality (Schuman et al. 1998). This divergence has led many scholars to theorize about the emergence of modern forms of prejudice, exemplified by the concept of *symbolic racism* (Kinder and Sears 1981).

However persuasive as explanatory factors prejudice and discrimination may appear to the layperson, sociologists have in recent decades more and more neglected them in favor of *institutional*, or *structural*, mechanisms of inequality (see, e.g., Bonilla-Silva 1997). One reason for this shift has been skepticism that prejudice and individual-level discrimination by themselves are adequate to account for the depth and durability of racial and ethnic cleavages in industrial societies, especially since these purported explanatory factors have seemed to decline in tandem with rising educational levels. (However, the emphasis on structural mechanisms can itself be faulted for neglecting the ideological component in racism.)

One expression of the focus on structural factors has been the notion of *institutional racism* (Blauner 1972). According to it, inequality among racial and ethnic groups depends not so much on individual acts of discrimination as it does on the workings of such institutions as the schools and the police, which process and sort individuals according to their racial and ethnic origins and ultimately impose very different outcomes on them. An assumption of this approach is that this sort of discrimination can occur on a wide scale without equally widespread prejudice. Indeed, it may even be possible without any discriminatory intent on the part of individuals in authority; an example would be educational tracking systems, which sort students according to racial background based on culturally and socially biased cues that are presumed by teachers and administrators to be related to intellectual ability. Studies deriving from the notion of institutional racism have in fact provided

some compelling analyses of the perpetuation of inequalities (on education, see Persell 1977), although they also can easily descend into controversy, as when any unequal outcome is declared to indicate the operation of racism.

A crucial arena in which both institutional and individual forms of racism operate to produce inequality is that of residence. Residential segregation is probably the most prominent indicator of the persisting importance of race in the United States; and it is critical to many other inequalities because life chances, especially for children, vary sharply across neighborhoods, which differ in many ways, from the qualities of the housing and schools they contain to the risks that their residents will be victims of crime. Decades of research with highly refined data and measures, such as the well-known Index of Dissimilarity, have shown that levels of neighborhood segregation by race are quite high and, at best, moderating very slowly (Farley and Frey 1994; Massey and Denton 1993). The pattern in many American cities can be described as one of *hypersegregation*, in the sense that the overwhelming majority of African Americans reside in large, consolidated ghettos containing virtually no whites and few members of other groups (Massey and Denton 1993). Residential segregation is not much explained by the economic inequalities between whites and blacks. So-called audit studies, involving matched pairs of white and black housing applicants, reveal considerable outright discrimination in the housing market. While whites are now more willing than in the past to accept blacks as neighbors, it appears that their tolerance is usually limited to small numbers. Contemporary segregation is also the consequence of government policies, past and present. The policies of the Federal Housing Administration, which effectively led in the 1930s to the modern mortgage instrument but were biased against mortgage lending in areas with many black residents, have had an enduring impact of American residential patterns, reflected in the division between largely white suburbs and largely nonwhite cities. (For a thorough discussion of the mechanisms behind segregation, see Massey and Denton 1993.)

A major theme in the stratification approach is the often complicated relationship, or interaction, between ethnicity and social class. One viewpoint is that ethnicity is, to some degree at least, a

manifestation of deeply rooted class dynamics. This has led to analyses that emphasize the economic and material foundations of what appear superficially to be cultural and ethnic distinctions. Analyses of this type have sometimes been inspired by Marxism, but they are hardly limited to Marxists. For example, Herbert Gans ([1962] 1982), in an influential analysis of second-generation Italians in a Boston neighborhood, argued that many of their distinctive traits could be understood as a function of their working-class position, which was not greatly changed from the situation of their southern Italian ancestors. In a related vein, Stephen Steinberg (1989) argues that cultural explanations of ethnic inequalities, which impute “undesirable” characteristics to some groups and “desirable” ones to others, are often rationalizations of economic privilege.

It is sometimes argued that inequalities that once rested on an ethnic basis now rest primarily on one of class. An important, if controversial, instance is William J. Wilson’s (1978, 1987) claim of a “declining significance of race” for American blacks. One part of Wilson’s argument focuses on an increasing socioeconomic division within the black population. This is held to result from the increasing opportunities available to young, well-educated African Americans since the 1960s. However, while improvements have been registered for a minority of the group, the lot of the black poor has not improved—it has even worsened. Wilson describes their situation as one of an *underclass*, which he defines in terms of isolation from the mainstream economy and society. His explanations for the emergence of the underclass are structural, not individualistic, and include the spatial concentration of the black poor in run-down urban neighborhoods, which have been stripped of their institutional fabric and middle-class residents, and the exodus of suitable job opportunities from central cities to suburbs and Sunbelt areas. In opposition to Wilson, others have argued that the emergence of underclass ghettos is better understood as a consequence of racism, as exemplified in residential segregation (e.g., Massey and Denton 1993).

An economic approach has also been used to explain ethnic conflict, which is seen as an outgrowth of the conflicting material interests of different ethnic groups. An exemplar is provided

by the theory of the ethnically *split labor market* (Bonacich 1972). Such a labor market develops when two ethnically different groups of workers compete (or could compete) for the same jobs at *different costs to employers*. It is typical in such situations for the higher-priced group of workers to have the same ethnic origins as employers, and therefore for the lower-priced group to be ethnically different from both. Nevertheless, it is in the interest of employers to substitute lower-priced workers for higher-priced ones wherever possible, despite the ethnic ties they share with the latter. Intense ethnic conflict can therefore develop between the two groups of workers, as the higher-priced group seeks to eliminate the threat to its interests. Two strategies may be employed: exclusion of the lower-priced group (for example, through legal restrictions on the immigration of its members) or creation of a caste system, that is, the limitation of the lower-priced group to a separate sphere of undesirable jobs. Split-labor-market theory has been applied to black–white relations in South Africa and the United States.

Yet, even in terms of a strictly economic approach, the precise genesis of the conflict between different ethnic groups of workers is open to question, and the theory of *segmented labor markets* draws another picture (Piore 1979). This theory divides the economies and labor markets of advanced capitalist societies into a primary sector, which contains relatively secure, well-paid jobs with decent working conditions and the opportunity for advancement, and a secondary sector, made up of insecure, dead-end jobs at low wages. Regardless of their class position, workers from the dominant group prefer to avoid jobs in the secondary sector, and usually they can manage to do so. Even unemployment may not be sufficient to force them into the secondary sector, since the benefits and resources available to most members of the dominant group, such as relatively generous unemployment compensation and seniority rights, enable them to wait out periodic unemployment. Hence, there is a need for another supply of workers, typically drawn from minorities and immigrants, who have no alternative but to accept employment in the secondary sector. Immigrants, in fact, are often willing to take these jobs because, as sojourners, they find the social stigma attached to the work to be less meaningful than do the

native-born. In contrast to the theory of the split labor market, which takes the existence of an ethnic difference among workers as a given, segmented-labor-market theory can explain why ethnic differences, especially between natives and immigrants, are so prevalent and persistent in the industrial societies of the West.

An economic explanation of ethnic differences is sometimes placed in a context of worldwide colonialism and capitalist exploitation (Rex 1981). Indeed, ethnic inequalities within a society are sometimes seen as the consequence of international relations between colonizers and the colonized. The notion that subordinate groups form economically exploited *internal colonies* in Western societies is an expression of this view (Blauner 1972). This notion is compatible with a hypothesis of a *cultural division of labor*, according to which positions in the socioeconomic order are assigned on the basis of cultural markers and hence ethnic origin (Hechter 1975).

The stratification approach need not focus exclusively on socioeconomic differences. Some scholars, in fact, prefer to see inequalities of power as more fundamental (Horowitz 1985; Stone 1985). This is a very general perspective on ethnic stratification, which is quite compatible with such fundamental notions as dominant and subordinate groups. According to it, social-class relations are but one instance, no matter how important, of the institutionalized inequalities between ethnic groups. Equally important, ethnic dominance cannot be reduced to, or explained solely in terms of, social-class mechanisms. (An implication is that class analysis of ethnic relations can be reductionist, an attempt to explain away ethnicity's causal independence.) Thus, the antagonism and sectarian violence between Catholics and Protestants in Northern Ireland is not comprehensible solely in terms of a social-class analysis, even though aggregate class differences between the groups exist as a result of centuries-long Protestant domination. This domination, the legacy of the colonial treatment of Ireland by the British, is manifest in a number of areas—in separate residential neighborhoods and schools, in social relations between members of the two groups, and in the political system. In short, domination encompasses much more than social-class privilege and gives even

working-class members of the Protestant group a sense of status and superiority.

Distinguishing empirically between ethnic stratification based on power and that rooted in economic structure has proven difficult. In one attempt, Hubert Blalock (1967) formulated a *power threat hypothesis*, to be contrasted with one derived from economic competition between groups. These two hypotheses can be tested in the relationship between discrimination and the size of a minority group. In particular, threats to the power of the dominant group are expected to result in discrimination that rises sharply with increases in the size of a minority; the same is not true for economic competition. So far, this test has been applied mainly to the American South (Tolnay and Beck 1995).

Theories concerning power differentials among ethnic groups border on the third major approach to the study of ethnicity, with its focus on ethnic-group resources. This approach, like the preceding one, takes its point of departure from the inequalities among groups. However, its vision is less one of the domination of some groups over others than it is of a more balanced competition that is affected by characteristics of the groups, such as their numbers, their solidarity, and their ability to form separate ethnic subeconomies. Such characteristics can give the group and its members relative advantages, or disadvantages, in this competition. Insofar as advantages are conferred, there may be incentives for individuals to maintain their attachments to a group rather than assimilate. In a sense, theories of ethnic-group resources can be seen as counterarguments to assimilation theories.

This is certainly clear in Nathan Glazer and Daniel Patrick Moynihan's ([1963] 1970) politically based explanation for the continuing importance of ethnicity in the United States. These authors acknowledge that immigrant cultures fade quickly under the impact of the assimilation process; assimilation is accomplished to this degree. However, ethnicity comes to coincide with differences in American circumstances, such as residential and occupational concentrations, which are similarly affected by government policies and actions. Hence, ethnicity takes on importance in the political sphere: Ethnic groups become "interest groups," reflecting the interests of many similarly

situated individuals. This role breathes new life into what might otherwise languish as an Old World social form. Glazer and Moynihan give many examples of the working of such interest groups in New York City.

Others have argued that ethnicity has become "politicized" in many contemporary societies, including many industrialized ones, and this leads to an unanticipated ethnic "resurgence." Daniel Bell (1975) states one basis for this point of view, claiming that politics is increasingly replacing the market as the chief instrument of distribution and that politics recognizes only group claims, thus enhancing ethnicity's political import. (The direction of change seems to have shifted since then, however.) In a more general fashion, Francois Nielsen (1985) contends that ethnicity offers a wider basis for political recruitment than the chief alternative, social class, and therefore ethnic-based movements have a greater chance for success. Addressing the situation in Third World societies, Donald Horowitz (1985) sees the ethnic political conflict that troubles many of them as originating in part in colonial policies and then being intensified by the anxieties of groups over their status in the postcolonial order.

Students of ethnic politicization have focused especially on the phenomenon of ethnic mobilization, which is epitomized in separatist movements in modern states, as in the Congo, Quebec, and Tibet (Olzak 1983). Mobilization can be regarded as one manifestation of the state of ethnic solidarity, a core concept in the literature on ethnicity. Ethnic groups marked by *solidarity* can be defined as self-conscious communities whose members interact with each other to achieve common purposes, and *mobilization* occurs when members take some collective action to advance these purposes. Recent research on ethnic movements appears to demonstrate that they are not generally interpretable in modern polities as the vestiges of traditional loyalties that have yet to be submerged by the modernization processes attendant upon development; rather, such movements can be outcomes of these processes and thus increase as economic development proceeds. The specific causes of this linkage are disputed, however (Olzak 1983).

Culture is another domain in which the search for group resources has been carried out. The

group-resources approach is compatible with the cultural-pluralist description of society, described earlier. More commonly in the past than today, the relative success of ethnic groups has been explained in terms of cultural traits. Quite often, the advantages these give have been analyzed in social-psychological terms. A well-known attempt along these lines was that of Bernard Rosen (1959), who matched American ethnic groups against the profile of the "achievement syndrome," a configuration of values that was presumed to predispose individuals to success. Included was an orientation to the future rather than the past and a downplaying of fatalism. In Rosen's analysis, the presence or absence of these traits in the culture of a group was explained according to the group's history and experience, and frequently in terms of the culture of the society from which it came. This sort of analysis, presuming stable cultural traits and rooting socioeconomic success in social-psychological prerequisites, has fallen into disfavor of late. In fact, it is often seen more as popular myth than as social science (Steinberg 1989). Cultural explanations, however, are not limited to the social-psychological realm. As one example, Ivan Light (1972) has devised an intriguing partial explanation for the entrepreneurial proclivities of different groups—in terms of the extent to which their cultures sponsor mechanisms that generate capital for the start-up of small businesses. Light argues that the business success of some Asian groups can be understood in part as an outcome of the rotating-credit association, a traditional social form imported from their home societies. Nevertheless, sociologists recently have stressed the malleability of culture and have tended to view it more as an adaptation to, and hence outcome of, socioeconomic position than as a cause of it. Consequently, cultural interpretations currently play only a minor role in the study of ethnicity. This may be a neglect engendered by cyclical intellectual fashion—in the future, they may loom larger, especially in the analysis of the immigrant groups proliferating in many societies (see, e.g., Zhou and Bankston 1998).

A focus of intense interest in recent years has been on the economic resources some groups are able to attain and thus on the advantages and opportunities adhering to ethnicity for their members. This interest is expressed in somewhat divergent research streams, running along distinctive

conceptual channels. One has been carved out around the linkage that appears in many societies between minority-group status and entrepreneurial activity. This linkage makes sense in terms of the disadvantages borne by many minority groups in mainstream economies, where they may suffer from various forms of discrimination and be channeled into low-status positions. Entrepreneurial activity, then, represents an attempt to evade these economic disadvantages. But, as many observers have pointed out, this hypothesis alone cannot account for the wide variation in entrepreneurial levels and forms among groups (Light 1984). An early attempt at a general theory is that of middleman minorities, which begins from the observation that a few groups, such as the Chinese and Jews, have occupied entrepreneurial niches in numerous societies (Bonacich 1973). The theory views entrepreneurialism, especially in commercial forms, as consistent with the sojourner status and interstitial position of these groups, between elites and masses. Yet this theory is too specific to account for contemporary immigrant entrepreneurialism, which frequently, for instance, involves nonsojourner groups. Therefore, Ivan Light (1984) has turned toward broader conceptual accounting schemes, seeing a range of ethnic and class resources as behind entrepreneurialism. Very intriguing here is the notion of ethnic resources, which include the established ethnic networks that train newly arrived immigrants and set them up in business. This notion helps to account for the clustering that is so obvious in ethnic small business, such as the concentration of New York's Koreans in greengroceries and dry cleaning.

A second stream of research builds partly on the immigrant proclivity toward entrepreneurialism. The theory of the ethnic enclave, developed by Alejandro Portes and various collaborators, holds that some groups are able to form ethnic subeconomies, which shelter not only entrepreneurs but also workers from the disadvantages they would face in the secondary sectors of the mainstream economy (Portes and Bach 1985). Enclave theory is particularly interesting because it provides a material motive for resisting assimilation. As exemplified in the case of the Cubans of Miami, enclave-forming groups typically contain a high-status stratum composed of individuals with professional occupations, capital, or both, along

with lower-status strata of workers seeking employment. A key to a full-fledged enclave is the establishment of networks of firms in interrelated economic sectors. The success of these firms is predicated on ethnic loyalties to an important degree. Workers and bosses may find mutual advantage where they share the same ethnicity. Workers may be willing to work longer hours or for lower wages, thus enhancing the profitability of a business, because they are able to work in a culturally familiar environment (usually speaking their mother tongue, for example). Workers may also have the opportunity to learn about running a business, and some eventually graduate to become entrepreneurs themselves. Enclave economies may be sustained in part by servicing the needs of their own ethnic communities, but if they are to be truly robust, they must also plug into the mainstream economy. This is the case, for instance, with the many ethnic firms in the American garment industry.

Despite its attractiveness on theoretical grounds, the implications of the enclave economy are disputed. One criticism is that such an economy offers few benefits for workers; the economic gains accrue to ethnic entrepreneurs (Sanders and Nee 1987). In addition, there is growing recognition that fully formed enclave economies are uncommon (Logan et al. 1994). These problems have led investigators to reinvigorate a broader notion, that of the *ethnic niche*, which refers to any economic position where a group is sufficiently concentrated to draw advantage from it, typically by being able to steer its own members into openings (Model 1997; Waldinger 1996). Thus, one can speak of an entrepreneurial niche, such as the Koreans have established in various lines of small business, or of an occupational niche, exemplified by the former dominance of the Irish in municipal employment in numerous cities. The ethnic-niche idea implicates other notions, such as the operation of ethnic networks that assist group members in finding positions in a niche. The niche idea is also frequently linked to that of an ethnic queue, an hierarchical ordering among minority groups, and to demographic and socioeconomic shifts. The latter open up niches, as exemplified by the withdrawal of white ethnics from some lines of small business in American cities, and the former determine which other groups are best positioned to take advantage of the openings (Waldinger 1996).

Oddly, a relatively neglected dimension of the ethnic-resources perspective concerns ethnic communities themselves, despite an almost universal recognition among scholars that some ethnics prefer to live in communities where their fellow ethnics are a numerous, if not the predominant, element of the population. Ethnic neighborhoods continue to be a salient aspect of metropolitan life in the United States and connected to inequalities that affect the well-being and life chances of their residents; and now they are emerging on a substantial scale in suburban settings (Horton 1995). A longstanding idea is that the residents of "institutionally complete" communities are less likely to assimilate (Breton 1964). Given the burgeoning of immigrant communities in the United States and elsewhere, it seems certain that this scholarly neglect will soon be repaired.

The final orientation, the social constructionist one, is so new that it cannot yet be said to have accumulated a substantial corpus of findings (for general reviews, see Nagel 1994; Omi and Winant 1996). This is not to deny its roots in the classical conception of ethnicity, as exemplified by the Weberian definition cited earlier, with its emphasis on the subjective view and, implicitly, on the relation of ethnic boundaries to social closure. As its name implies, this perspective stresses that ethnic boundaries and meanings are not preordained but malleable and thus that an understanding of ethnic distinctions must generally be sought in contemporary, rather than historical, circumstances. This perspective is especially attuned to the possible emergence of new forms of ethnicity, as some see, for instance, in the cultural hybridity associated with the rising frequency of mixed racial ancestry in the United States. There is also an emphasis on heightened ethnic fluidity, even the potential for revolutionary shifts, in the present because of the rise of the global economy and associated phenomena, such as large-scale migrations across borders, cultural diffusions, and transnational networks (Sassen 1988).

The most striking findings to emerge from this perspective concern historical shifts in racial boundaries, confounding the lay view that racial distinctions are determined by impossible-to-overlook physical differences among humans. Yet, historically, such boundaries have shifted in the United States within relatively brief intervals. During their immigrations, Irish Catholics, Italians, and

Eastern European Jews were all perceived as racially different from native-born whites, as nineteenth-century cartoon depictions of the Irish demonstrate. Today, the racial element in the perception of ethnic differences among whites has disappeared. The limited historical studies have not yet fully explicated the decades-long process by which this happened but make clear that it was not benign, as immigrants used violence and stereotyping, among other means, to create social distance between themselves and African Americans and thereby make plausible their candidacy for “whiteness” (see Ignatiev 1995; Roediger 1991). Then, their acceptability to other white Americans grew as they climbed the social ladder and mixed occupationally and residentially with other whites.

In conclusion, one must acknowledge that the literature on ethnicity remains unsettled in its theoretical core. The persistence—perhaps even the resurgence—of ethnic difference and conflict in societies throughout the world has attracted much attention from sociologists and other social scientists. But the paradoxes associated with ethnicity, evidenced in the United States by the assimilation of some groups and the continued separateness and even subordination of others, have yet to be resolved. They remain fruitful for sociology, nevertheless: The study of ethnicity has produced some of the discipline’s most striking findings and, no doubt, will continue to do so.

## REFERENCES

- Alba, Richard 1988 “Cohorts and the Dynamics of Ethnic Change.” In Matilda White Riley, Bettina J. Huber, and Beth B. Hess, eds., *Social Structures and Human Lives*. Newbury Park, Calif.: Sage.
- 1990 *Ethnic Identity: The Transformation of White America*. New Haven, Conn.: Yale University Press.
- , and Victor Nee 1997 “Rethinking Assimilation for a New Era of Immigration.” *International Migration Review* 31:826–874.
- Allport, Gordon 1954 *The Nature of Prejudice*. New York: Addison-Wesley.
- Banton, Michael 1983 *Racial and Ethnic Competition*. Cambridge, U.K.: Cambridge University Press.
- Barth, Frederik 1969 “Introduction.” In Frederik Barth, ed., *Ethnic Groups and Boundaries*. Boston, Mass.: Little, Brown.
- Bell, Daniel 1975 “Ethnicity and Social Change.” In Nathan Glazer and Daniel Patrick Moynihan, eds., *Ethnicity: Theory and Experience*. Cambridge, Mass.: Harvard University Press.
- Blalock, Hubert 1967 *Toward a Theory of Minority-Group Relations*. New York: Capricorn.
- Blauner, Robert 1972 *Racial Oppression in America*. New York: Harper and Row.
- Bogardus, Emory 1928 *Immigration and Race Attitudes*. Boston: D.C. Heath.
- Bonacich, Edna 1972 “A Theory of Ethnic Antagonism: The Split Labor Market.” *American Sociological Review* 37:547–559.
- 1973 “A Theory of Middleman Minorities.” *American Sociological Review* 38:583–594.
- Bonilla-Silva, Eduardo 1997 “Rethinking Racism: Toward a Structural Interpretation.” *American Sociological Review* 62:465–480.
- Breton, Raymond 1964 “Institutional Completeness of Ethnic Communities and the Personal Relations of Immigrants.” *American Journal of Sociology* 70:193–205.
- Farley, Reynolds, and William Frey 1994 “Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps Towards a More Integrated Society.” *American Sociological Review* 59:23–45.
- Gans, Herbert (1962) 1982 *The Urban Villagers: Group and Class in the Life of Italian Americans*. New York: Free Press.
- Glazer, Nathan, and Daniel Patrick Moynihan (1963) 1970 *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City*. Cambridge, Mass.: MIT Press.
- Gordon, Milton 1964 *Assimilation in American Life*. New York: Oxford University Press.
- Greeley, Andrew 1971 *Why Can't They Be Like Us?* New York: Dutton.
- Hechter, Michael 1975 *Internal Colonialism: The Celtic Fringe and British National Development*. Berkeley: University of California Press.
- 1987 *Principles of Group Solidarity*. Berkeley: University of California Press.
- Hirschman, Charles 1983 “America’s Melting Pot Reconsidered.” *Annual Review of Sociology* 9:397–423.
- Horowitz, Donald 1985 *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Horton, John 1995 *The Politics of Diversity: Immigration, Resistance, and Change in Monterey Park, California*. Philadelphia: Temple University Press.
- Ignatiev, Noel 1995 *How the Irish Became White*. New York: Routledge.

- Isaacs, Harold 1975 "Basic Group Identity: The Idols of the Tribe." In Nathan Glazer and Daniel Patrick Moynihan, eds., *Ethnicity: Theory and Experience*. Cambridge, Mass.: Harvard University Press.
- Kinder, Donald, and David Sears 1981 "Prejudice and Politics: Symbolic Racism Versus Threats to the Good Life." *Journal of Personality and Social Psychology* 40:414-431.
- Lal, Barbara Ballis 1990 *The Romance of Culture in an Urban Civilization: Robert E. Park on Race and Ethnic Relations in Cities*. London: Routledge.
- Lieberson, Stanley 1961 "A Societal Theory of Race and Ethnic Relations." *American Sociological Review* 26:902-910.
- , and Mary Waters 1988 *From Many Strands: Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage Foundation.
- Light, Ivan 1972 *Ethnic Enterprise in America: Business and Welfare Among Chinese, Japanese, and Blacks*. Berkeley: University of California Press.
- 1984 "Immigrant and Ethnic Enterprise in North America." *Ethnic and Racial Studies* 7:195-216.
- Logan, John, Richard Alba, and Thomas McNulty 1994 "Ethnic Economies in Metropolitan Regions: Miami and Beyond." *Social Forces* 72:691-724.
- Massey, Douglas 1985 "Ethnic Residential Segregation: A Theoretical Synthesis and Empirical Review." *Sociology and Social Research* 69:315-350.
- , and Nancy Denton 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Model, Suzanne 1997 "Ethnic Economy and Industry in Mid-Twentieth Century Gotham." *Social Problems* 44:445-463.
- Nagel, Joane 1994 "Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture." *Social Problems* 41:152-176.
- Neidert, Lisa, and Reynolds Farley 1985 "Assimilation in the United States: An Analysis of Ethnic and Generation Differences in Status and Achievement." *American Sociological Review* 50:840-850.
- Nielsen, Francois 1985 "Toward a Theory of Ethnic Solidarity in Modern Societies." *American Sociological Review* 50:133-149.
- Olzak, Susan 1983 "Contemporary Ethnic Mobilization." *Annual Review of Sociology* 9:355-374.
- Omi, Michael, and Howard Winant 1994 *Racial Formation in the United States: From the 1960s to the 1990s*, 2nd ed. New York: Routledge and Kegan Paul.
- Park, Robert 1950 *Race and Culture*. New York: Free Press.
- Persell, Caroline 1977 *Education and Inequality: The Roots of Stratification in America's Schools*. New York: Free Press.
- Piore, Michael 1979 *Birds of Passage: Migrant Labor and Industrial Societies*. Cambridge, U.K.: Cambridge University Press.
- Portes, Alejandro, and Robert Bach 1985 *Latin Journey: Cuban and Mexican Immigrants in the United States*. Berkeley: University of California Press.
- Portes, Alejandro, and Ruben Rumbaut 1996 *Immigrant America: A Portrait*, 2nd ed. Berkeley: University of California Press.
- Portes, Alejandro, and Min Zhou 1993 "The New Second Generation: Segmented Assimilation and Its Variants." *The Annals* 530:74-96.
- Rex, John 1981 "A Working Paradigm for Race Relations Research." *Ethnic and Racial Studies* 4:1-25.
- Roediger, David 1991 *The Wages of Whiteness: Race and the Making of the American Working Class*. New York: Verso.
- Rosen, Bernard 1959 "Race, Ethnicity, and the Achievement Syndrome." *American Sociological Review* 24:47-60.
- Sanders, Jimmy, and Victor Nee 1987 "Limits of Ethnic Solidarity in the Ethnic Enclave." *American Sociological Review* 52:745-767.
- Sassen, Saskia 1988 *The Mobility of Capital and Labor*. Cambridge, U.K.: Cambridge University Press.
- Schuman, Howard, Charlotte Steeh, Lawrence Bobo, and Maria Krysan 1998 *Racial Attitudes in America: Trends and Interpretations*, rev. ed. Cambridge, Mass.: Harvard University Press.
- Steinberg, Stephen 1989 *The Ethnic Myth: Race, Ethnicity, and Class in America*. Boston: Beacon.
- Stone, John 1985 *Racial Conflict in Contemporary Society*. London: Fontana Press/Collins.
- Thernstrom, Stephan, Ann Orlov, and Oscar Handlin 1980 *Harvard Encyclopedia of American Ethnic Groups*. Cambridge, Mass.: Harvard University Press.
- Tolnay, Stewart, and E. M. Beck 1995 *A Festival of Violence: An Analysis of Southern Lynchings, 1882-1930*. Urbana: University of Illinois Press.
- van den Berghe, Pierre 1967 *Race and Racism: A Comparative Perspective*. New York: Wiley.
- 1981 *The Ethnic Phenomenon*. New York: Elsevier.
- Waldinger, Roger 1996 *Still the Promised City? African-Americans and New Immigrants in Postindustrial New York*. Cambridge, Mass.: Harvard University Press.
- Waters, Mary 1990 *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.



Weber, Max (1922) 1968 *Economy and Society*. New York: Bedminster Press.

Wilson, William J. 1978 *The Declining Significance of Race: Blacks and Changing American Institutions*. Chicago: University of Chicago Press.

— 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Zhou, Min, and Carl Bankston 1998 *Growing Up American: How Vietnamese Children Adapt to Life in the United States*. New York: Russell Sage Foundation.

RICHARD D. ALBA

## ETHNOCENTRISM

See Ethnicity; Nationalism.

## ETHNOGRAPHY

Ethnographic research (also referred to as *field research* or *participant observation*) is a qualitative social science method that involves the observation of the interactions of everyday life, whether in public parks, business organizations, or mental health clinics. It is social constructionist, exploring intersubjective cultural meanings rather than positivist “social facts,” or laws. The theoretical intent of ethnography is inductive, generating concepts and theories from the data. However, in sociology, ethnography is also associated with micro-level theories such as symbolic interactionism, ethnomethodology, labeling theory, and dramaturgy. In particular, the theories of everyday life developed by Georg Simmel (1950) and Erving Goffman (1959) form part of the disciplinary repertoire from which workers in the field draw their concepts. Some field researchers also link their ethnography to critical, macro-theoretical frameworks emphasizing power relations within hierarchies of gender, class, race, age, nationality, or sexual orientation (Diamond, 1993).

The goal of field research is on the development of analytic descriptions or grounded theories (Glaser and Strauss, 1967) of the social world, usually written but sometimes audio- or videotaped. As Clifford Geertz (1988) notes, “thick description” is the foundation of ethnography.

But equally important is analysis: the generation of concepts, patterns, or typologies from thick description, and their linkage to concepts, theories, and literatures already established in the discipline. Although some academic fieldworkers may have other goals (such as helping the people in the setting), most seek to publish their ethnographic work as books or articles. There are also some field researchers who have turned to alternative forms of inscription, such as poems or plays, to try to bring to life the social worlds they study or (as in autoethnography) inhabit.

## HISTORIES OF ETHNOGRAPHY

The contemporary origins of sociological ethnography are traced to the “Chicago School” of 1915–1940 (Bulmer 1984), and to nineteenth century sociological theory, reformist endeavors, and anthropological exploration. “Origin myths” seek the roots of ethnography in the writings of ancient and medieval travelers who sought first-hand knowledge of other cultures. Herodotus, for example, has been called the father of both history and sociology. In the fifth century B.C.E., Herodotus traveled to distant lands and recorded comments about the peoples and customs he found there; throughout the centuries into modern times, European, Persian, and Chinese explorers, traders, and missionaries followed suit. But it was in the nineteenth century, with the development of sociology as a discipline, that first-hand observation became part of the modern methodological repertoire.

Although the canonical work of the nineteenth century European theorists such as Karl Marx, Max Weber, and Émile Durkheim was macro-level and comparative-historical, the works are read for their ethnographic warrant. In particular, Weber’s concept of *verstehen*, read as an interpretive immersion into the subjective worlds of respondents, is cited as an epistemological foundation of contemporary ethnography. Feminist ethnographers of the late twentieth century have added the names of prominent nineteenth and early twentieth century women to the histories of field research, citing Harriett Martineau’s interest in the everyday lives of women and children, and Charlotte Perkins Gilman’s “autoethnographic” account of her moral treatment for neurasthenia (Bailey 1996).

Anglo-American reform movements also played a part in the development of nineteenth and early twentieth century observational methods. With an Enlightenment focus on reason and the scientific method, reform efforts came to be accompanied by observational studies of the problems seen as needing remediation, from conditions in mental hospitals to poverty in the streets of London or New York. The work of reformist scholars such as Beatrice Webb and Charles Booth in Victorian London among the poor involved the use of survey and field research methods. During the same era, Beatrix Potter took a position in a London sweatshop in order to document working conditions and attempt to change them (Bailey 1996).

British and American anthropology, following Bronislaw Malinowski's lead, instituted several of the key elements of field research: the time commitment, immersion in the geographical space of the respondents, the use of cultural insiders as key informants, and the inscription of observations and thoughts. Other aspects of ethnographic research are traced to the Chicago School, with its eclectic mix of pragmatist philosophy, symbolic interaction, journalistic interests, and reformist tendencies (Bulmer 1984; Becker 1999). Robert Park's injunction to his Chicago students has been taught to subsequent generations of ethnographers:

*go and sit in the lounges of the luxury hotels and on the doorsteps of the flophouses; sit on the Gold Coast settees and on the slum shakedowns; sit in the Orchestra Hall and in the Star and Garter Burlesk. In short, gentlemen, go get the seat of your pants dirty in real research (Bulmer 1984; 97).*

The Chicago School sociologists pioneered the case study method, which included oral histories, documents, statistical surveys, and some fieldwork. And it was from Chicago that the themes of 1950s to 1970s fieldwork emerged: the focus on studies of everyday life exemplified by the work of Goffman (1959), and on labeling and deviance by Howard Becker (1963) and others. The methods learned by contemporary sociological ethnographers are grounded in the Chicago, neo-Chicago (Fine 1995) and "California" schools (Adler and Adler 1987), while there has also been convergence with other disciplines practicing ethnography, such as anthropology, education, and business.

## DOING ETHNOGRAPHIC RESEARCH

Introductory texts on ethnographic research (for example, Bailey 1996; Lofland and Lofland 1995) emphasize the nonlinear aspect of the method: the reflexive interplay of data collection, writing, and analysis. However, two distinct sets of activities are involved in doing ethnography: interacting within the "field" and inscription. Inscription occurs in the form of field notes (and perhaps diaries, memos, analytic notes, and other writings) and drafts of the ethnographic paper or monograph. Interaction within the field poses the usual everyday life issues of roles and relationships, together with more particular concerns such as human subjects regulations, research ethics, and entrée. Thus, the methodology of field research poses intertwined issues of representation and interaction.

Before initiating research in the field, the ethnographer selects a setting, and makes her preparations for attempting entrée. These preparations may be as simple as walking across the road to the local park, or as complex as negotiating with the governments of Cuba and the United States for permission to travel and stay in Havana. For many studies, especially funded team ethnography, it is necessary for the ethnographer to obtain permission to do the study from Institutional Review Boards (IRBs) at her University, and perhaps at the organization she wishes to study, for example, a prison or mental hospital. These Institutional Review Boards ensure that the fieldworker's plans are in compliance with the basic tenets of human subjects protections: informed consent (which is often not possible in non-interview based field research) and protection of confidentiality (in practice, not naming respondents, nor providing identifying details in published work).

All fieldwork in public places, and some fieldwork in other places, is exempt from IRB scrutiny, but no fieldwork is exempt from ethical considerations. Ethical issues that have preoccupied ethnographers for several decades are those associated with the deception of respondents. In covert fieldwork, where the ethnographer pretends to be a member of the setting (for example, Alcoholics Anonymous), the respondents are deceived about the researcher's identity and purposes. This level of deception is probably less common than it was up until the early 1980s. More common, however, is fieldwork in settings where the researcher

is already a member or participant, such as a bar where he is employed, or a bisexual or gay organization of which he is a member. In this type of “nonstranger” research, the fieldworker’s identity is known, but she may or may not share her new research purposes with the other members.

**Entrée and Incorporation.** In a setting where she is initially a stranger, the ethnographer must negotiate entrée into the setting. If this is a public park, she simply goes and sits on a park bench. If this is an organization, she will have already negotiated initial entrée with an administrative “gatekeeper,” and now must thread her way through the geographical and relational space of the setting, from the first to the last time she does fieldwork there. It is a truism of field research that entrée is not a one-time event: both physical and relational entrée must continually be negotiated and renegotiated throughout the field study. Initial setting gatekeepers may be followed by new gatekeepers of other parts of the setting, while “freeze outs” bar the door to some occasions or locales.

Once in the door, hoping to be allowed to stay and move about (and perhaps conduct formal or informal interviews), the researcher is incorporated into the setting: assigned a physical and conceptual space within the field by each individual respondent, and by groups of people within the setting. The language of incorporation, as against that of role playing, highlights the reciprocal nature of the interaction between the respondents and the ethnographer, in which the researcher cannot just “decide” what kind of role to play, but, rather, is assigned a place by the people in the setting. The ethnographer’s personal and status characteristics—gender, age, nationality, appearance—are an important aspect of incorporation, and determine the place that will be afforded him within the setting. These characteristics may, indeed, preclude or smooth access to particular persons, events, or parts of the field. Much has been written about gender and ethnography (Warren and Hackney 2000, since gender and sexuality are focal elements of hierarchy, interactions, and relationships within many settings).

Incorporation is a powerful emotional force field for many ethnographers, drawing them into the web of relationships in the field. Old hands at fieldwork warn of the necessity—if an analytic

description is to be written—of balancing immersion with distance, touching cognitive and emotional base with both disciplinary concepts and out-of-setting roles in order to avoid engulfment in the field (Pollner and Emerson 1988). Researchers who begin as members of the setting they are studying sometimes report the opposite, becoming more detached as they move from immersed membership to analytic observer: the field, once warmly familiar, can become a place of distance and even alienation.

In the classic language of the “research bargain,” ethnographers may find that they are approached and sometimes used as economic, medical, status, labor, or even sexual resources in the organizations or communities where they conduct their fieldwork. Fieldwork “warnings and advice,” although they avoid the positivist language of bias, note that such research bargains may take a toll on researchers in the form of time, emotions or funds. And at times, the incorporation of the ethnographer into the setting makes leaving the field emotionally difficult, although practically necessary. Some ethnographers, anthropologists in particular, revisit “their” fields at intervals of years or months in order to document change over time. But, once out of the field, the researcher is faced with the disciplinary requisite of starting or finishing the ethnographic writing of the inscription already begun in the field.

**Ethnographic Representation.** The endpoint of fieldwork is ethnographic representation: an article or book about the field, published in a scholarly journal, or as a monograph or trade book. Two major publication outlets for sociological ethnographies are the *Journal of Contemporary Ethnography* and *Qualitative Sociology*; fieldworkers may also publish in general, specialty (for example, gerontology or education), or anthropology journals. Many published ethnographies are “realist tales” (Van Maanen 1988), written in the traditional style of the analytic description: “I am here and this is what is going on here.” Some are “confessional tales” which focus on the writer rather than on the respondents, while others experiment with alternative representational forms such as poetry or plays. There is also a well-established tradition of visual sociology which includes photography, film, video, and the internet together with more contemporary genres such as performance ethnography.

Written ethnographic representations begin with fieldnotes: the inscription of the field (Emerson et al. 1995). Fieldworkers are taught and exhorted to write down everything they can, including verbatim conversations, in the service of thick description. This exhortation approaches feasibility in bounded settings such as a self-help group meeting that occurs once a week for fifty minutes. It becomes more problematic when the ethnographer is faced with an entire village, or even perhaps a large urban area. In practice, fieldworkers often take note of, and inscribe, aspects of events and persons to which they habitually attend; for example, one person studying a restaurant takes note of spatial patterns, while another inscribes social types by appearance and clothing, and still another notes the interaction between servers and cooks.

The inscription of fieldnotes is vital to the traditional ethnographic enterprise, since the identification of analytic patterns is difficult without an adequate amount of field notes (what is “adequate” varying with the size and complexity of the field). Alongside fieldnotes (separated by parentheses, or on separate pieces of paper), fieldworkers write memos, diaries, and analytic notes which record their own feelings and thoughts, analytic ideas and possibilities, and sometimes events which they do not want recorded in their formal fieldnotes. Although most fieldnotes have no audience but the self, fieldnotes written for graduate seminars or funded team research may be read by instructors, supervisors, or co-workers. In inscribing the Other, fieldnotes also inscribe the self (Warren 2000).

### POSTMODERN ETHNOGRAPHY AND BEYOND

Ethnography, like most corners of the social sciences, has been affected by postmodernism. As representations, both fieldnotes and published ethnographies are open to the postmodern critiques of social scientific representation, critiques which began in sociology with the work of Joseph Gusfield (1976) on the literary rhetoric of sociological writing. One result of this postmodern critique is the deconstruction of realist or traditional ethnography, which, in turn, has resulted in a withdrawal from the field (the Other) into the text or the self. This movement away from traditional fieldwork has given rise to an ironic reversal

of earlier exhortations to “get out of the armchairs and into the streets.” There is a joke in ethnographic circles today that goes something like this: Respondent to Ethnographer: “Well, enough about you. Why don’t we talk a bit about me?”

While the postmodern ethnography of the 1990s sometimes takes place in the armchair and away from the streets, dwelling upon the self rather than upon the Other, it is likely that the new century and millenium will see a renewed concern with the activities, as well as the representations, of ethnography. Those sociologists concerned with relations of power and hierarchy—of gender, race, class, and nation—are particularly concerned that a focus on textual discourse and polyvocality (attending equally to the voices of the ethnographer and those of all the respondents) does not erase the critical function of ethnographic research. Jaber Gubrium and James Holstein (1997) suggest that contemporary ethnographers attend to both the “what” (the field), and the “how” (the representations of fieldwork), becoming “self-consciously attentive to both the world researched and the researcher” (1997: 212). This is the reflexivity toward which our ethnographic practice strives at the beginning of the twenty-first century.

#### REFERENCES

- Adler, Patricia A., and Peter Adler 1987 *Membership Roles in Field Research*. Newbury Park: Sage.
- Bailey, Carol A. 1996 *A Guide to Field Research*. Thousand Oaks, Calif.: Pine Forge Press.
- Becker, Howard S. 1999 “The Chicago School, So-Called.” *Qualitative Sociology* 22 (Spring): 3–12.
- Becker, Howard S. 1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Bulmer, Martin 1984 *The Chicago School of Sociology*. Chicago: University of Chicago Press.
- Diamond, Tim 1993 *Making Grey Gold: Narratives of Nursing Home Care*. Chicago: University of Chicago Press.
- Emerson, Robert M., Rachel I. Fretz, and Linda L. Shaw 1995 *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press.
- Fine, Gary Alan, ed. 1995 *A Second Chicago School?* Chicago: University of Chicago Press.
- Geertz, Clifford 1988 “Toward an Interpretive Theory of Culture.” In Robert M. Emerson, ed., *Contemporary Field Research*. Prospect Heights, Ill.: Waveland Press.

- Glaser, Barney and Anselm L. Strauss 1967 *The Discovery of Grounded Theory*. Chicago: Aldine.
- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, NY: Anchor Books.
- Gubrium, Jaber F. and James A. Holstein 1997 *The New Language of Qualitative Research*. New York: Oxford University Press.
- Gusfield, Joseph 1976 "The Literary Rhetoric of Science: Comedy and Pathos in Drinking Driver Research." *American Sociological Review* 41 (February): 16–34.
- Lofland, John and Lyn H. Lofland 1995 *Analyzing Social Settings*. third; ed. Belmont: Wadsworth.
- Pollner, Melvin and Robert M. Emerson 1988 "The Dynamics of Inclusion and Distance in Field Research." In Robert M. Emerson, ed., *Contemporary Field Research*. Prospect Heights, Ill.: Waveland Press.
- Simmel, Georg 1950 *The Sociology of Georg Simmel*, trans./ed. Kurt H. Wolff. New York: Free Press.
- Van Maanen, John 1988 *Tales of the Field: on Writing Ethnography*. Chicago: University of Chicago Press.
- Warren, Carol A.B. 2000 "Writing the Other, Inscribing the Self." *Qualitative Sociology*. Forthcoming.
- , and Jennifer K. Hackney 2000 *Gender Issues in Ethnography*. 2nd ed. Thousand Oaks, Calif.: Sage.
- Wolff, Kurt H., ed. 1950 *The Sociology of Georg Simmel*. New York: Free Press.

CAROL A.B. WARREN

## ETHNOMETHODOLOGY

Ethnomethodology is a field of sociology that studies the commonsense resources, procedures, and practices through which the members of a culture produce and recognize mutually intelligible objects, events, and courses of action. The field emerged in the late 1960s in reaction to a range of sociological perspectives, most prominently structural functionalism, which treated conduct as causally determined by social structural factors. In contrast, ethnomethodology stressed that social actions and social organization are produced by knowledgeable agents who guide their actions by the use of situated commonsense reasoning. Rather than treating the achievement of social organization as a given from which the analysis of social structure could proceed, ethnomethodological research was directed at the hidden social processes

underlying that achievement. The resulting research focus on the properties of commonsense knowledge and reasoning represents one strand of what has been termed the "cognitive revolution" in the social sciences. As a sociological perspective however, ethnomethodology deals with the socially shared and publicly accountable nature of commonsense reasoning rather than with psychological aspects of cognitive processes. Its primary research stance has been descriptive and naturalistic rather than explanatory or experimental.

### BACKGROUND AND DEVELOPMENT

The basic outlook of ethnomethodology was developed by Harold Garfinkel (1967a) during a twenty-year period spanning graduate research at Harvard under the supervision of Talcott Parsons through an extensive number of empirical investigations at UCLA. Garfinkel's starting point was the vestigial treatment in the sociological analyses of the 1950s of how actors employ knowledge to understand and act in ordinary social contexts (Heritage 1984). With respect to the prevailing treatment of internalized norms as motivational "drivers" of behavior, Garfinkel noted that the achievement of goals requires actions based on knowledge of real circumstances and that where coordinated action is necessary, that knowledge must be socially shared. What is the character of this knowledge? How is it implemented and updated? By what means are shared and dynamically changing knowledge and understanding concerning actions and events sustained? Merely to raise these questions was to point to fundamental deficiencies in the theory of action.

In developing answers to these questions, Garfinkel drew on the theoretical writings of the phenomenological sociologist Alfred Schutz (Schutz 1962–1966). Schutz observed that each actor approaches the social world with a "stock of knowledge at hand" made up of commonsense constructs and categories that are primarily social in origin. The actor's grasp of the real world is achieved through the use of these constructs, which, Schutz stressed, are employed presuppositionally, dynamically, and in a taken-for-granted fashion. Schutz also observed that these constructs are held in typified form, that they are approximate and revisable, that actions are guided by a patchwork of "recipe knowledge," and that

intersubjective understanding between actors who employ these constructs is a constructive achievement that is sustained on a moment-to-moment basis. Ethnomethodology took shape from Garfinkel's efforts to develop these theoretical observations into a program of empirical research.

A major component of these efforts took the form of the famous "breaching experiments," which were inspired by the earlier "incongruity experiments" pioneered by Solomon Asch and Jerome Bruner. The breaching experiments employed a variety of techniques to engineer drastic departures from ordinary expectations and understandings about social behavior. By "making trouble" in ordinary social situations, Garfinkel was able to demonstrate the centrality of taken-for-granted background understandings and contextual knowledge in persons' shared recognition of social events and in their management of coordinated social action. He concluded that understanding actions and events involves a circular process of reasoning in which part and whole, foreground and background, are dynamically adjusted to one another. Following Karl Mannheim, he termed this process "the documentary method of interpretation." In this process, basic presuppositions and inferential procedures are employed to assemble linkages between an action or an event and aspects of its real worldly and normative context. The character of the action is thus grasped as a "gestalt contexture" (Gurwitsch 1966) that is inferentially and procedurally created through the interlacing of action and context. Here temporal aspects of actions and events assume a central significance (Garfinkel 1967a), not least because background and context have to be construed as dynamic in character. Within this analysis, presuppositions, tacit background knowledge, and contextual detail are the inescapable resources through which a grasp of events is achieved.

Garfinkel (1967a) also showed that the recognition, description, or coding of actions and events is an inherently approximate affair. The particulars of objects and events do not have a "one-to-one" fit with their less specific representations in descriptions or codings. The fitting process thus inevitably involves a range of approximating activities that Garfinkel terms "*ad hoc* practices" (Garfinkel 1967a). This finding is, of course, the inverse of his well-known observation that descriptions, actions, and so forth have *indexical* properties: Their sense

is elaborated and particularized by their contextual location. An important consequence of these observations is that shared understandings cannot be engendered by a "common culture" through a simple matching of shared words or concepts but rather can only be achieved constructively in a dynamic social process (Garfinkel 1967a). Similar conclusions apply to the social functioning of rules and norms.

In summary: Garfinkel's researches indicate that every aspect of shared understandings of the social world depends on a multiplicity of tacit *methods of reasoning*. These methods are procedural in character, they are socially shared, and they are ceaselessly used during every waking moment to recognize ordinary social objects and events. A shared social world, with its immense variegation of social objects and events, is jointly constructed and recognized through, and thus ultimately rests on, a shared base of procedures of practical reasoning that operationalize and particularize a body of inexact knowledge.

In addition to functioning as a base for understanding actions, these procedures also function as a resource for the production of actions. Actors tacitly draw on them so as to produce actions that will be *accountable*—that is, recognizable and describable—in context. Thus, shared methods of reasoning are publicly available on the surface of social life because the results of their application are inscribed in social action and interaction. As Garfinkel (1967a) put it: "The activities whereby members produce and manage the settings of organized everyday affairs are identical with members' procedures for making these settings *accountable*" (p. 1).

While the results of Garfinkel's experiments showed that the application of joint methods of reasoning is central to the production and understanding of social action, they also showed that the application of these methods is strongly "trusted" (Garfinkel 1963, 1967a). This "trust" has a normative background and is insisted upon through a powerful moral rhetoric. Those whose actions could not be interpreted by means of this reasoning were met with anger and demands that they explain themselves. Garfinkel's experiments thus showed the underlying *morality* of practical reasoning and that the procedural basis of action and understanding is a part—perhaps the deepest part—of

the moral order. Such a finding is consistent with the view that this procedural base is foundational to organized social life and that departures from it represent a primordial threat to the possibility of sociality itself.

### CONTEMPORARY RESEARCH INITIATIVES

Garfinkel's writings have stimulated a wide range of commentary, theoretical reaction, and empirical initiatives. In what follows, only the latter will be described. Empirical research in ethnomethodology will be discussed under three headings: (1) social structures as normal environments, (2) the creation and maintenance of social worlds, and (3) studies of work. A fourth, and possibly the most conspicuous, initiative—conversation analysis—is discussed elsewhere in this Encyclopedia.

#### **Social Structures as Normal Environments.**

In his theoretical writings, Schutz (1962) argued that human consciousness is inherently typifying and that language is the central medium for the transmission of socially standardized typifications. In a number of his empirical studies, Garfinkel developed this idea in relation to social process, noting the ways in which commonsense reasoning is used—often within a moral idiom—to typify and normalize persons and events. A number of influential ethnomethodological studies have taken up this theme and focused on the ways in which participants may be actively or tacitly engaged in creating or reproducing a texture of normality in their everyday affairs.

Much of this work was focused in the fields of deviance and bureaucratic record keeping. This focus was far from accidental. In both fields, the participants are concerned with the administration of socially consequential categories and in both—with their indigenous preoccupation with classification and definition—normalizing processes were close to the surface of organizational life and were somewhat easier to track. Pioneering studies in this area included Sudnow's (1965) analysis of "normal crimes," in which he showed that California lawyers employed models of "typical" offenders and offences in plea-bargaining procedures that departed substantially from the provisions of the California criminal code. Zimmerman's (1969) work on record keeping in a public welfare agency showed that bureaucratic records employed

typifications of clients that could only be interpreted by reference to detailed background knowledge of the organization's procedures (see also Garfinkel 1967a, pp. 186–207). Wieder's (1974) work on a halfway house for paroled narcotics offenders showed that a "convict code" profoundly shaped how staff and inmates perceived events inside the institution—with disastrous consequences for its success.

Related works on deviance—by Cicourel (1968) on the policing of juveniles and by Atkinson (1978) on suicide—crystallized points of friction between ethnomethodology and more traditional approaches to the study of deviant behavior. Both studies examined the social processes underlying the classification of deviants. Each of them detailed a complex of commonsense considerations that enter into the determination of the nature of a deviant act and (in the case of juvenile offenders) the treatment of its perpetrator.

Cicourel's study showed that police treatment of juveniles was informed by a lay theory that posited a connection between juvenile offenses and the home background of the offender. Offenses by juveniles from "broken homes" were treated more seriously than offenses by those from other social backgrounds. In consequence, offenses by juveniles from broken homes were more likely to be officially reported, were more commonly the object of court proceedings, and had a greater tendency to result in custodial sentences. Police records, Cicourel showed, embodied a related process of idealization and typification in which case records, as they were developed through the system, became increasingly concise, selective, and consistent with the assumptions, objectives, and dispositions of the legal agencies. At the core of Cicourel's argument was the claim that the processing of juvenile offenders exhibits a circular process. Basic assumptions about the causal factors associated with juvenile crime were being used to normalize offenders and were built into the differential treatment of juveniles. From this point, these assumptions became built into police records and statistics and, finally, into social scientific treatments of the statistics that "recovered" the initial assumptions as valid explanations of juvenile crime.

Similar conclusions were reached by Atkinson (1978) in relation to the treatment of suicide.

Drawing on the work of Garfinkel (1967G) and Douglas (1967), Atkinson argued that police conceptions of “typical suicides” profoundly influenced how particular cases of sudden death were investigated and treated. These conceptions not only influenced individual verdicts but, through the accumulation of verdicts, the official statistics on suicide. Atkinson concluded that sociological studies of suicide based on official statistics are unavoidably engaged in decoding the commonsense typifications of suicide that were constitutive in the recognition of, and verdicts on, individual cases and that accumulate in the statistical record.

In sum, ethnomethodological studies of typification in relation to deviance and organizational records have had both “constructive” and “deconstructive” moments. New and important social processes that inform the categorial activities of public agencies of various kinds have been uncovered. At the same time, these discoveries have challenged traditional sociological treatments of official statistics. The “deconstructive” conclusion that official statistics of social phenomena may be largely artifactual and, in many cases, can tell us only about the kinds of assumptions and practices that animate the relevant officials has provoked debates in the discipline that are unresolved to the present date.

**The Creation and Maintenance of Social Institutions and Social Worlds.** An important aspect of ethnomethodological theorizing is the notion that social institutions are sustained as real entities through vocabularies of accounts (or accounting frameworks) through which the events of the social world are recognized and acted upon. Although this idea can be traced back to C. Wright Mills (1940), it found vivid expression in Garfinkel’s (1967a) analysis of a transsexual individual, which he used as an occasion to investigate the nature of gender as (1) the achievement of a particular individual that (2) was made possible by the person’s grasp of, and subscription to, appropriately “gendered” practices and accounting frameworks that are generally hidden or taken for granted by normally sexed persons.

A number of subsequent studies have developed this preoccupation with the role of accounting frameworks that are employed in the taken-for-granted production and reproduction of social

institutions and social realities. An early and influential work was Wieder’s (1974) study of the role of the “convict code” in a halfway house for paroled narcotics offenders. Wieder showed the ways in which the code—which prescribed a range of activities hostile to staff members—functioned both as a fundamental way of seeing “what was going on” in the halfway house and as a resource that could be invoked in accounting for noncompliant conduct in interaction with staff members. Of particular interest is Weider’s finding that the “code” functioned in these ways among both offenders and staff despite their distinctive and conflicting perspectives on the activities of the halfway house. In transcending the formal power structure of the institution, the code was the predominant medium through which events were defined and acted upon by all participants and, for this reason, served as a source of power and control for the offenders.

At a still more general level, Pollner (1974; 1987) has explored the ways in which a version of reality is socially sustained within a collectivity. Our sense of reality, he argues, is a social institution that is sustained by particular socially organized practices, which he labels “mundane reason.” Within this framework of practices, we start from the presumption that real-world objects and events are intersubjectively available as determinate, definite, noncontradictory, and self-identical. That this presumption is actively sustained, he shows, emerges in environments—ranging from everyday events to more specialized contexts such as the law courts, mental hospitals, and research science—where witnesses disagree in their depiction of objects and events. In such contexts, Pollner observes, a range of procedures are invoked that discount one version of events and privilege the other. The invocation of such procedures sustains (by restoring or repairing) persons’ belief in a single noncontradictory reality. “Explaining away” a witness’s testimony—by, for example, arguing that he couldn’t have seen what he saw, was not competent to do so, or was lying or even insane—functions in the way described by Evans-Pritchard (1937) as the “secondary elaboration of belief.” Such secondary elaborations, Pollner observes, are inevitably used in defense of the factual status of all versions of the world. They are the universal repair kit of the real. Pollner’s work has stimulated a range of empirical



studies in a variety of settings, ranging from mental hospitals (Coulter, 1975) to a research community of biochemists (Gilbert and Mulkay 1984), and has strongly influenced the recent “discourse analysis” movement in social psychology.

**Workplace Studies.** In recent years, ethnomethodological research has been most prominently visible in social studies of science (Lynch 1995) and in studies of knowledge and action in the workplace (Button 1993). These studies are continuous with earlier investigations in that they examine the practical context and practical achievement of scientific and workplace activities. However they also radicalize earlier studies in their demonstrations of the practical contingencies that underly the production and use of “hard” scientific and technological findings just as much as they do more ephemeral cultural objects such as sketch maps (Garfinkel 1996, in press).

The workplace studies began with a focus on embodied courses of practical reasoning and action that are involved in the technically competent performance of work tasks that range from playing jazz (Sudnow 1978) to proving a mathematical theorem (Livingston 1986), laboratory work in brain science (Lynch 1985), or astronomical discovery (Lynch et al. 1983). Work in these fields was studied for the detailed texture of the shop-talk and workbench practices that comprise the recognizably competent performance of work tasks. A further focus on workplace studies has emerged from Lucy Suchman’s (1987) pioneering study of human-machine communication. A very large body of ethnomethodological work has focused on a variety of aspects of machine-mediated observation and communication (for representative studies from a large corpus, see Button 1993, Luff et al. in press), as well as fundamental aspects of the relationship among computers, human action, and communication (Agre 1997; Button et al. 1995). Another set of studies addresses disciplinarily specific processes of representation in science (Lynch and Woolgar 1990), police work in courtroom testimony (Goodwin 1994), and historical events in the political context of congressional hearings (Lynch and Bogen 1996). Finally, a new body of work by Lynch and his collaborators (see Jordan and Lynch 1998) examines the nature and application of scientific techniques concerned with DNA sequencing in a variety of contexts, ranging from research and industrial to forensic contexts. This

research explores the contingencies that support the stability and objectivity of these techniques in some contexts, while subverting them in others.

## CONCLUSION

Since its emergence in the 1960s, ethnomethodology has developed as a complex set of research initiatives that have raised topic areas, problems, and issues for analysis where none were previously seen to exist or perceived to be relevant. A number of these initiatives have had a pronounced “deconstructive” dimension that has sometimes appeared iconoclastic or even nihilistic (Pollner 1991). Yet in its oscillation between constructive and deconstructive tendencies, ethnomethodology has been a significant site of theoretical and empirical innovation within sociology. It is exerting a continuing impact on the sensibility of the discipline. Finally, it has also had a wide ranging influence on a range of adjacent disciplines that are concerned with knowledge systems, communication, action, and practical reasoning

(SEE ALSO: *Conversation Analysis*)

## REFERENCES

- Agre, P. 1997. *Computation and Human Experience*. Cambridge, U.K.: Cambridge University Press.
- Atkinson, J.M. 1978 *Discovering Suicide: Studies in the Social Organization of Sudden Death*. London: Macmillan.
- Button, G., ed. 1993 *Technology in Working Order: Studies of Work, Interaction and Technology*. London: Routledge.
- , Jeff Coulter, John Lee, and Wesley Sharrock 1995 *Computers, Minds and Conduct*. Oxford: Blackwell.
- Cicourel, A.V. 1968 *The Social Organization of Juvenile Justice*. New York: Wiley.
- Coulter, J. 1975 “Perceptual Accounts and Interpretive Asymmetries.” *Sociology* 9:385–396.
- Douglas, J. 1967 *The Social Meanings of Suicide*. Princeton, N.J.: Princeton University Press.
- Evans-Pritchard, E. E. 1937 *Witchcraft, Oracles and Magic Among the Azande*. Oxford: Oxford University Press.
- Garfinkel, H. 1963 “A Conception of, and Experiments with, ‘Trust’ as a Condition of Stable Concerted Actions.” In O. J. Harvey, ed., *Motivation and Social Interaction*. New York: Ronald Press.
- 1967a *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice Hall.

- 1967b “Practical Sociological Reasoning: Some Features of the Work of the Los Angeles Suicide Prevention Center.” In E. S. Schneidman, ed., *Essays in Self-Destruction*. New York: International Science Press.
- 1996. “Ethnomethodology’s Program.” *Social Psychology Quarterly* 59(1):5–21.
- in press. *A Catalogue of Ethnomethodological Investigations*, ed. Anne Warfield Rawls. Boston: Basil Blackwell.
- , M. Lynch, and E. Livingston 1981 “The Work of a Discovering Science Construed with Materials from the Optically Discovered Pulsar.” *Philosophy of the Social Sciences* 11:131–158.
- Gilbert, G. N., and M. Mulkay, 1984 *Opening Pandora’s Box: A Sociological Analysis of Scientists’ Discourse*. Cambridge, U.K.: Cambridge University Press.
- Goodwin, C. 1994 “Professional Vision.” *American Anthropologist* 96(3):606–633.
- Gurwitsch, A. 1966 *Studies in Phenomenology and Psychology*. Evanston, IL: Northwestern University Press.
- Heritage, J. 1984 *Garfinkel and Ethnomethodology*. Cambridge, Mass.: Polity Press.
- Jordan, K. and M. Lynch 1998 “The Dissemination, Standardization and Routinization of a Molecular Biological Technique.” *Social Studies of Science*, 28:773–800.
- Livingston, E. 1986 *Ethnomethodological Foundations of Mathematics*. London: Routledge.
- Luff, P., J. Hindmarsh, and C. Heath, eds., in press. *Workplace Studies: Recovering Work Practice and Informing System Design*. Cambridge, U.K.: Cambridge University Press.
- Lynch, M. 1985 *Art and Artifact in Laboratory Science*. London: Routledge.
- 1995 *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge, U.K.: Cambridge University Press.
- , and D. Bogen 1996 *The Spectacle of History: Speech, Text and Memory at the Iran–Contra Hearings*. Durham, N.C.: Duke University Press.
- , E. Livingston, and H. Garfinkel 1983 “Temporal Order in Laboratory Work.” In K. Knorr-Cetina and M. Mulkay, eds., *Science Observed*. London: Sage.
- , and S. Woolgar, eds. 1990 *Representation in Scientific Practice*. Cambridge, Mass.: MIT Press.
- Mills, C. W. 1940 “Situated Actions and Vocabularies of Motive.” *American Sociological Review* 5:904–913.
- Pollner, M. 1974 “Mundane Reasoning.” *Philosophy of the Social Sciences* 4:35–54.
- 1987 *Mundane Reason: Reality in Everyday and Sociological Discourse*. Cambridge, U.K.: Cambridge University Press.
- 1991 “Left of Ethnomethodology,” *American Sociological Review* 56:370–380.
- Schutz, A. 1962 “Commonsense and Scientific Interpretation of Human Action.” In A. Schutz, *Collected Papers: Vol. 1. The Problem of Social Reality*, ed. by M. Natanson. The Hague: Martinus Nijhoff.
- 1962–1966 *Collected Papers*, 3 vol. The Hague: Martinus Nijhoff.
- Suchman, L. 1987 *Plans and Situated Action*. Cambridge, U.K.: Cambridge University Press.
- Sudnow, D. 1965 “Normal Crimes.” *Social Problems* 12:255–276.
- 1978 *Ways of the Hand*. Cambridge, Mass.: Harvard University Press.
- Wieder, D. L. 1974 *Language and Social Reality*. The Hague: Mouton.
- Zimmerman, D. 1969 “Record Keeping and the Intake Process in a Public Welfare Agency.” In S. Wheeler, ed., *On Record: Files and Dossiers in American Life*. Beverly Hills, Calif.: Sage.

JOHN HERITAGE

## EUTHANASIA

See Death and Dying.

## EVALUATION RESEARCH

### BRIEF HISTORY

There is no uniformly accepted definition of what constitutes evaluation research. At perhaps its narrowest point, the field of *evaluation research* can be defined as “the use of scientific methods to measure the implementation and outcomes of programs for decision-making purposes” (Rutman 1984, p. 10). A broader, and more widely accepted, definition is “the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs” (Rossi and Freeman 1993, p. 5). A much broader definition is offered by Scriven (1991), who suggests that evaluation is “the process of determining the merit, worth and value of things” (p. 1). In the latter definition, the notion of what can be evaluated is

not limited to a social program or specific type of intervention but encompasses, quite literally, everything.

Any description of the history of evaluation research depends on how the term is defined. Certainly, individuals have been making pronouncements about the relative worth of things since time immemorial. In the case of social programs, proficiency requirements to guide the selection of public officials using formal tests were recorded as early as 2200 B.C. in China (Guba and Lincoln 1981). Most observers, however, date the rise of evaluation research to the twentieth century. For example, programs in the 1930s established by the New Deal were viewed as great opportunities to implement social science methods to aid social planning by providing an accounting of program effects (Stephan, 1935). Modern evaluation research, however, underwent explosive growth in the 1960s as a result of several factors (Shadish et al. 1991). First, the total amount of social programming increased tremendously under the administrations of Presidents Kennedy, Johnson, and Nixon. New programs were directed toward social issues such as education, housing, health, crime, and income maintenance. Second, along with these huge financial investments came the concern by Congress about whether these programs were achieving their intended effect. As a result, Congress began mandating evaluations. Third, program managers were concerned whether programs were being implemented in the manner intended, and consequently data were required to monitor program operations. In addition, there were intellectual issues about how best to implement programs and the relative effectiveness of various approaches to offsetting various social ills. Outcome data were needed to compare competing approaches. The result was a burgeoning demand for trained evaluators; and the large number of scientists involved in the common enterprise of evaluation became sufficient to support the development of evaluation research as a scientific specialty area.

The field of evaluation research is no longer expanding at the rate it was in the 1960s and 1970s (Freeman 1992). By the 1980s, there was a substantial decline in the funding for evaluation activities that was motivated, in part, by the budget cuts of the Reagan administration. By then, however, the field of evaluation research had been established.

It continues to thrive for several reasons (Desautels 1997). First, difficult decisions are always required by public administrators and, in the face of continuing budget constraints, these decisions are often based on accountability for results. Second, an increasingly important aspect of service provision by both public and private program managers is service quality. Monitoring quality requires information about program practices and outcomes. Third, there is growing public demand for accountability in government, a view increasingly echoed by government representatives. Meeting these demands requires measurement of results and a management system that uses evaluation for strategic planning and tactical decision making.

Early in its history, evaluation was seen primarily as a tool of the political left (Freeman 1992). Clearly, that is no longer the case. Evaluation activities have demonstrated their utility to both conservatives and liberals. Although the programs of today may be different from those launched in the 1960s, evaluation studies are more pervasive than ever. As long as difficult decisions need to be made by administrators serving a public that is demanding ever-increasing levels of quality and accountability, there will be a growing market for evaluation research.

### PURPOSES OF EVALUATION RESEARCH

A wide variety of activities are subsumed under the broad rubric of "evaluation research." This diversity proceeds from the multiplicity of purposes underlying evaluation activities. Chelimsky (1997) identifies three different purposes of evaluation: evaluation for accountability, evaluation for development, and evaluation for knowledge.

**Accountability.** From the perspective of auditors and funding agencies, evaluations are necessary to establish accountability. Evaluations of this type frequently attempt to answer the question of whether the program or policy "worked" or whether anything changed as a result. The conceptual distinction between program and policy evaluations is a subtle but important one (Sonnad and Borgatta 1992). Programs are usually characterized by specific descriptions of what is to be done, how it is to be done, and what is to be accomplished. Policies are broader statements of objectives than programs, with greater latitude in how they are implemented and with potentially more

diverse outcomes. Questions addressed by either program or policy evaluations from an accountability standpoint are usually cause-and-effect questions requiring research methodology appropriate to such questions (e.g., experiments or quasi-experiments). Studies of this type are often referred to as summative evaluations (Scriven 1991) or impact assessments (Rossi and Freeman 1993). Although the term “outcome” evaluation is frequently used when the focus of the evaluation is on accountability, this term is less precise, since all evaluations, whether conducted for reasons of accountability, development, or knowledge, yield outcomes of some kind (Scriven 1991).

**Development.** Evaluation for development is usually conducted to improve institutional performance. Developmental evaluations received heightened importance as a result of public pressure during the 1980s and early 1990s for public management reforms based on notions such as “total quality management” and “reinventing government” (e.g., see Gore 1993). Developmental evaluations often address questions such as: How can management performance or organizational performance be improved? What data systems are necessary to monitor program accomplishment? What are appropriate indicators of program success and what are appropriate organizational goals? Studies designed primarily to improve programs or the delivery of a product or service are sometimes referred to as formative or process evaluations (Scriven 1991). In such studies, the focus is on the treatment rather than its outcomes. Depending on the specific question being addressed, methodology may include experiments, quasi-experiments, or case studies. Data may be quantitative or qualitative. Formative or process evaluations may be sufficient by themselves if a strong relationship is known to exist between the treatment and its outcomes. In other cases, they may be accompanied by summative evaluations as well.

**Knowledge.** In evaluation for knowledge, the focus of the research is on improving our understanding of the etiology of social problems and on detailing the logic of how specific programs or policies can ameliorate them. Just as evaluation for accountability is of greatest interest to funding or oversight agencies, and evaluation for performance is most useful to program administrators, evaluation for knowledge is frequently of greatest interest to researchers, program designers, and

evaluators themselves. Questions might include such things as the causes of crime, homelessness, or voter apathy. Since these are largely cause-and-effect questions, rigorous research designs appropriate to such questions are generally required.

## CONTEMPORARY ISSUES IN EVALUATION

**Utilization of Findings.** Implicit in the enterprise of evaluation research is the belief that the findings from evaluation studies will be utilized by policy makers to shape their decisions. Indeed, such a view was espoused explicitly by Campbell (1969), who argued that social reforms should be regarded as social experiments and that the findings concerning program effectiveness should determine which programs to retain and which to discard. This process of rational decision making, however, has not been consistently embraced by policy makers and has been a source of concern and disillusionment for many evaluators. Rossi (1994) sums up the situation by noting:

*Although some of us may have entertained hopes that in the “experimenting society” the experimenter was going to be king, that delusion, however grand, did not last for long. It often seemed that programs had robust lives of their own, appearing, continuing, and disappearing following some unknown processes that did not appear responsive to evaluations and their outcomes. (p. 26)*

One source of the utilization problem, as Weiss (1975, 1987) has noted, is the fact that evaluations take place in a political context. Although accomplishing its stated objectives is important to program success, it may not be the only—or even the most important—measure of program success. From this perspective, it is not that administrators and policy makers are irrational—they simply use a different model of rationality than do evaluators. Indeed, the view of policy makers and program administrators may be more “rational” than that of evaluators because it has been shown repeatedly that programs can and do survive negative evaluations. Programs are less likely, however, to survive a hostile congressional committee, negative press, or lack of public support. There are generally multiple stakeholders, often with competing interests, associated with any large program. Negative findings are of very little use to individuals whose reputations and jobs are dependent on program

success. Thus, rather than bemoaning a lack of utilization of findings, evaluators need to recognize that evaluation findings represent only one piece of a complex political process.

Evaluators concerned with utilization frequently make a distinction between the immediate or instrumental use of findings to make direct policy decisions versus the conceptual use of findings, which serves primarily to enlighten decision makers and perhaps influence later decision making (Leviton and Hughes 1981). In a related vein, Scriven (1993) makes an important distinction between “lack of implementation” and “lack of utilization.” Lack of implementation merely refers to a failure to implement recommendations. In contrast, utilization is more ambiguous. It is often not clear what outcomes or actions actually constitute a utilization of findings. Evaluation findings can have great utility but may not necessarily lead to a particular behavior. For example, a consumer can read an evaluation of a product in a publication such as *Consumer Reports* and then decide not to buy the product. Although the evaluation did not lead to a particular behavior (i.e., purchasing the product), it was nonetheless extremely useful to the consumer, and the information can be said to have been utilized. Some observers have noted that the concern about underutilization of evaluation findings belies what is actually happening in the field of evaluation research. Chelimsky and Shadish (1997) provide numerous examples of how evaluation findings have had substantial impacts on policy and decision making, not only in government but also in the private sector, and not only in the United States but internationally as well.

**Quantitative Versus Qualitative Research.** The rise of evaluation research in the 1960s began with a decidedly quantitative stance. In an early, influential book, Suchman (1967) unambiguously defined *evaluation research* as “the utilization of scientific research methods and techniques” (p. 7) and cited a recent book by Campbell and Stanley (1963) on experimental and quasi-experimental designs as providing instruction on the appropriate methodology. It was not long, however, before the dominance of quantitative methods in evaluation research came under attack. Cook (1997) identifies two reasons. First, there has been a longstanding debate, especially in sociology, over the merits of qualitative research and the limits of quantitative methods. Sociologists brought the debate with

them when they entered the field of evaluation. Second, evaluation researchers, even those trained primarily in quantitative methods, began to recognize the epistemological limitations of the quantitative approach (e.g., Guba and Lincoln 1981). There were also practical reasons to turn toward qualitative methods. For example, Weiss (1987) noted that quantitative outcome measures are frequently too insensitive to detect program effects. Also, the expected time lag between treatment implementation and any observed outcomes is frequently unknown, with program effects often taking years to emerge. Moreover, due to limited budgets, time constraints, program attrition, multiple outcomes, multiple program sites, and other difficulties associated with applied research, quantitative field studies rarely achieved the potential they exuded on the drawing board. As a result, Weiss recommended supplementing quantitative with qualitative methods.

Focus on the quantitative–qualitative debate in evaluation research was sharpened when successive presidents of the American Evaluation Association expressed differing views on the matter. On the qualitative side, it was suggested that the focus on rigor associated with quantitative evaluations may have blinded evaluators to “artistic aspects” of the evaluation process that have traditionally been unrecognized or simply ignored. The time had come “to move beyond cost benefit analyses and objective achievement measures to interpretive realms” in the conduct of evaluation studies (Lincoln 1991, p. 6). From the quantitative perspective, it was acknowledged that while it is true that evaluations have frequently failed to produce strong empirical support for many attractive programs, to blame that failure on quantitative evaluations is akin to shooting the messenger. Moreover, at a time when research and statistical methods (e.g., regression discontinuity designs, structural equations with latent variables, etc.) were finally catching up to the complexities of contemporary research questions, it would be a shame to abandon the quantitative approach (Sechrest 1992). The ensuing controversy only served to polarize the two camps further.

The debate over which approach is best, quantitative or qualitative, is presently unresolved and, most likely, will remain so. Each paradigm has different strengths and weaknesses. As Cook (1997) points out, quantitative methods are good for

generalizing and describing causal relationships. In contrast, qualitative methods are well suited for exploring program processes. Ironically, it is the very differences between the two approaches that may ultimately resolve the issue because, to the extent that their limitations differ, the two methods used jointly will generally be better than either used singly (Reichardt and Rallis 1994).

**Research Synthesis.** Evaluation research, as it was practiced in the 1960s and 1970s, drew heavily on the experimental model. The work of Donald Campbell was very influential in this regard. Although he is very well known for his explication of quasi-experimental research designs (Campbell and Stanley 1963; Cook and Campbell 1979), much of his work actually de-emphasized quasi-experimentation in favor of experiments (Shadish et al. 1991). Campbell pointed out that quasi-experiments frequently lead to ambiguous causal inferences, sometimes with dire consequences (Campbell and Erlebacher 1970). In addition, he noted that experiments have wide applicability, even in applied settings where random assignment may not initially seem feasible (Campbell and Boruch 1975). Campbell also advocated implementing such rigorous methods in the evaluation of social programs (Campbell 1969). As a result, Campbell is frequently credited with proposing a rational model of social reform in which a program is first evaluated using rigorous social science methods, such as experiments, when possible, and then a report is issued to a decision maker who acts on the findings.

Whatever its source, it was not long before the rational model was criticized as being too narrow to serve as a template for evaluation research. In particular, Cronbach and colleagues (Cronbach et al. 1980) argued that evaluation is as much a political process as a scientific one, that decisions are rarely made but more likely emerge, that there is rarely a single decision maker, and that programs are often amorphous undertakings with no single outcome. From Cronbach's perspective, the notion that the outcome of a single study could influence the existence of a program is inconsistent with the political realities of most programs.

Understanding the ensuing controversy requires an understanding of the notion of validity. Campbell distinguished between two types of validity: internal and external (Campbell 1957; Campbell and Stanley 1963). Internal validity refers to

whether the innovation or treatment has an effect. In contrast, external validity addresses the issue of generalizability of effects; specifically, "To what populations, settings, treatment variables, and measurement variables can this effect be generalized" (Campbell and Stanley 1963, p. 5). Campbell clearly assigned greater importance to internal validity than to external validity. Of what use is it, he asked, to generalize experimental outcomes to some population if one has doubts about the very existence of the relationship that one seeks to generalize (Shadish et al. 1991)? Campbell's emphasis on internal validity was clearly consistent with his focus on experiments, since the latter are particularly useful in examining causal relationships.

In contrast, Cronbach (1982) opposed the emphasis on internal validity that had so profoundly shaped the approach to evaluation research throughout the 1960s and 1970s. Although experiments have high internal validity, they tend to be weak in external validity; and, according to Cronbach, it is external validity that is of greatest utility in evaluation studies. That is, decision makers are rarely interested in the impact of a particular treatment on a unique set of subjects in a highly specific experimental setting. Instead, they want to know whether a program or treatment, which may not always be administered in exactly the same way from agency to agency, will have an effect if it is administered on other individuals, and in other settings, from those studied in the experimental situation. From Cronbach's perspective, the rational model of evaluation research based on rigorous social research procedures is a flawed model because there are no reliable methods for generalizing beyond the factors that have been studied in the first place and it is the generalized rather than the specific findings in which evaluators are interested. As a result, Cronbach viewed evaluation as more of an art than a scientific enterprise.

The debate over which has priority in evaluation research, internal or external validity, seems to have been resolved in the increasing popularity of research syntheses. Evaluation syntheses represent a meta-analytic technique in which research results from numerous independent evaluation studies are first converted to a common metric and then aggregated using a variety of statistical techniques. The product is a meaningful summary of the collective results of many individual studies. Research synthesis based on meta-analysis has

helped to resolve the debate over the priority of internal versus external validity in that, if studies with rigorous designs are used, results will be internally valid. Moreover, by drawing on findings from many different samples, in many different settings, using many different outcome measures, the robustness of findings and generalizability can be evaluated as well.

Although meta-analysis has many strengths, including increased power relative to individual studies to detect treatment effects, the results are obviously limited by the quality of the original studies. The major drawback to meta-analysis, then, deals with repeating or failing to compensate for the limitations inherent in the original research on which the syntheses are based (Figueredo 1993). Since many evaluations use nonexperimental designs, these methodological limitations can be considerable, although they potentially exist in experiments as well (e.g., a large proportion of experiments suffer from low external validity).

An emerging theory underlying research syntheses of experimental and nonexperimental studies, referred to as critical multiplism (Shadish 1993) and based on Campbell and Fiske's (1959) notion of multiple operationalism, addresses these issues directly. "Multiplism" refers to the fact that there are multiple ways of proceeding in any research endeavor, with no single way being uniformly superior to all others. That is, every study will involve specific operationalizations of causes and effects that necessarily underrepresent the potential range of relevant components in the presumed causal process while introducing irrelevancies unique to the particular study (Cook 1993). For example, a persuasive communication may be intended to change attitudes about an issue. In a study to evaluate this resumed cause-and-effect relationship, the communication may be presented via television and attitudes may be assessed using paper-and-pencil inventory. Clearly, the medium used underrepresents the range of potential persuasive techniques (e.g., radio or newspapers might have been used) and the paper-and-pencil task introduces irrelevancies that, from a measurement perspective, constitute sources of error. The term "critical" refers to the attempt to identify biases in the research approach chosen. The logic, then, of critical multiplism is to synthesize the results of studies that are heterogeneous with

respect to sources of bias and to avoid any constant biases. In this manner, meta-analytic techniques can be used to implement critical multiplism ideas, thereby increasing our confidence in the generalizability of evaluation findings.

The increasing use of research syntheses represents one of the most important changes in the field of evaluation during the past twenty-five years (Cook 1997). Research synthesis functions in the service of increasing both internal and external validity. Although it may seem that the use of research syntheses is a far cry from Campbell's notion of an experimenting society, in reality Campbell never really suggested that a single study might resolve an important social issue. In "Reforms as Experiments" (1969) Campbell states:

*Too many social scientists expect single experiments to settle issues once and for all. . . . Because we social scientists have less ability to achieve "experimental isolation," because we have good reason to expect our treatment effects to interact significantly with a wide variety of social factors many of which we have not yet mapped, we have much greater needs for replication experiments than do the physical sciences. (pp. 427-428)*

Ironically, perhaps, the increasing use of research syntheses in evaluation research is perfectly consistent with Campbell's original vision of an experimenting society.

#### DIRECTIONS FOR THE FUTURE

The field of evaluation research has undergone a professionalization since the early 1970s. Today, the field of evaluation research is characterized by its own national organization (the American Evaluation Association), journals, and professional standards. The field continues to evolve as practitioners continue the debate over exactly what constitutes evaluation research, how it should be conducted, and who should do it. In this regard, Shadish and colleagues (1991) make a compelling argument that the integration of the field will ultimately depend on the continued development of comprehensive theories that are capable of integrating the diverse activities and procedures traditionally subsumed under the broad rubric of evaluation research. In particular, they identify a number of basic issues that any theory of evaluation must

address in order to integrate the practice of evaluation research. These remaining issues include knowledge construction, the nature of social programming and knowledge use, the role of values, and the practice of evaluation.

**Knowledge Construction.** A persisting issue in the field of evaluation concerns the nature of the knowledge that should emerge as a product from program evaluations. Issues of epistemology and research methods are particularly germane in this regard. For example, the controversy over whether quantitative approaches to the generation of knowledge are superior to qualitative methods, or whether any method can be consistently superior to another regardless of the purpose of the evaluation, is really an issue of knowledge construction. Other examples include whether knowledge about program outcomes is more important than knowledge concerning program processes, or whether knowledge about how programs effects occur is more important than describing and documenting those effects. Future theories of evaluation must address questions such as which types of knowledge have priority in evaluation research, under what conditions various knowledge-generation strategies (e.g., experiments, quasi-experiments, case studies, or participatory evaluation) might be used, and who should decide (e.g., evaluators or stakeholders). By so doing, the field will become more unified, characterized by common purpose rather than by competing methodologies and philosophies.

**Social Programming and Knowledge Use.** The ostensible purpose of evaluation lies in the belief that problems can be ameliorated by improving the programs or strategies designed to address those problems. Thus, a social problem might be remediated by improving an existing program or by getting rid of an ineffective program and replacing it with a different one. The history of evaluation research, however, has demonstrated repeatedly how difficult it is to impact social programming. Early evaluators from academia were, perhaps, naive in this regard. Social programs are highly resist to change processes because there are generally multiple stakeholders, each with a vested interest in the program and with their own constituencies to support. Complicating the matter is the fact that knowledge is used in different ways in different circumstances. Several important distinctions concerning knowledge use

can be made: (1) use in the short term versus use in the long term, (2) information for instrumental use in making direct decisions versus information intended for enlightenment or persuasion, and (3) lack of implementation of findings versus lack of utilization of findings. These different types of use progress at different rates and in different ways. Consequently, any resulting program changes are likely to appear slow and sporadic. But the extent to which such change processes should represent a source of disappointment and frustration for evaluators requires further clarification. Specifically, theories of evaluation are needed that take into account the complexities of social programming in modern societies, that delineate appropriate strategies for change in differing contexts, and that elucidate the relevance of evaluation findings for decision makers and change agents.

**Values.** Some evaluators, especially in the early history of the field, believed that evaluation should be conducted as a value-free process. The value-free doctrine was imported from the social sciences by early evaluators who brought it along as a by-product of their methodological training. This view proved to be problematic because evaluation is an intrinsically value-laden process in which the ultimate goal is to make a pronouncement about the value of something. As Scriven (1993) has cogently argued, the values-free model of evaluation is also wrong. As proof, he notes that statements such as "evaluative conclusions cannot be established by any legitimate scientific process" are clearly self-refuting because they are themselves evaluative statements. If evaluators cling to a values-free philosophy, then the inevitable and necessary application of values in evaluation research can only be done indirectly, by incorporating the values of other persons who might be connected with the programs, such as program administrators, program users, or other stakeholders (Scriven 1991). Obviously, evaluators will do a better job if they are able to consider explicitly values-laden questions such as: On what social values is this intervention based? What values does it foster? What values does it harm? How should merit be judged? Who decides? As Shadish and colleagues (1991) point out, evaluations are often controversial and explosive enterprises in the first place and debates about values only make them more so. Perhaps that is why values theory has gotten short shrift in the past. Clearly, however,



future theory needs to address the issue of values, acknowledging and clarifying their central role in evaluation research.

**The Practice of Evaluation.** Evaluation research is an extremely applied activity. In the end, evaluation theory has relevance only to the extent that it influences the actual practice of evaluation research. Any theory of evaluation practice must necessarily draw on all the aforementioned issues (i.e., knowledge construction, social programming and information use, and values), since they all have direct implications for practice. In addition, there are pragmatic issues that directly affect the conduct of evaluation research. One important contemporary issue examines the relationship between the evaluator and individuals associated with the program. For example, participatory evaluation is a controversial approach to evaluation research that favors collaboration between evaluation researchers and individuals who have some stake in the program under evaluation. The core assumption of participatory evaluation is that, by involving stakeholders, ownership of the evaluation will be shared, the findings will be more relevant to interested parties, and the outcomes are then more likely to be utilized (Cousins and Whitmore 1998). From an opposing perspective, participatory evaluation is inconsistent with the notion that the investigator should remain detached from the object of investigation in order to remain objective and impartial. Not surprisingly, the appropriateness of participatory evaluation is still being debated.

Other aspects of practice are equally controversial and require clarification as well. For example: Who is qualified to conduct an evaluation? How should professional evaluators be trained and by whom? Should evaluators be licensed? Without doubt, the field of evaluation research has reached a level of maturity where such questions warrant serious consideration and their answers will ultimately determine the future course of the field.

## REFERENCES

- Campbell, Donald 1957 "Factors Relevant to the Validity of Experiments in Social Settings." *Psychological Bulletin* 54: 297-312.
- 1969 "Reforms as Experiments." *American Psychologist* 24:409-429.

- , and Robert Boruch 1975 "Making the Case for Randomized Assignment to Treatments by Considering the Alternatives: Six Ways in Which Quasi-Experimental Evaluations in Compensatory Education Tend to Underestimate Effects." In C. Bennett and A. Lumsdaine, eds., *Evaluation and Experiments: Some Critical Issues in Assessing Social Programs*. New York: Academic Press.
- , and Albert Erlebacher 1970 "How Regression Artifacts Can Mistakenly Make Compensatory Education Programs Look Harmful." In J. Hellmuth, ed., *The Disadvantaged Child: vol. 3. Compensatory Education: A National Debate*. New York: Brunner/Mazel.
- , and Donald Fiske 1959 "Convergent and Discriminant Validity by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81-105.
- , and Julian Stanley 1963 *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Chelimsky, Eleanor 1997 "The Coming Transformations in Evaluation." In E. Chelimsky and W. Shadish, eds., *Evaluation for the Twenty-first Century*. Thousand Oaks, Calif.: Sage.
- , and William Shadish 1997 *Evaluation for the Twenty-first Century*. Thousand Oaks, Calif.: Sage.
- Cook, Thomas 1993 "A Quasi-Sampling Theory of the Generalization of Causal Relationships." In L. Sechrest and A. Scott, eds., *Understanding Causes and Generalizing About Them* (New Directions for Program Evaluation, No. 57). San Francisco: Jossey-Bass.
- 1997 "Lessons Learned in Evaluation over the Past 25 Years." In Eleanor Chelimsky and William Shadish, eds., *Evaluation for the Twenty-first Century*. Thousand Oaks, Calif.: Sage.
- , and Donald Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Cousins, J. Bradley, and Elizabeth Whitmore 1998 "Framing Participatory Evaluation." In Elizabeth Whitmore, ed., *Understanding and Practicing Participatory Evaluation* (New Directions for Evaluation, No. 80). San Francisco: Jossey-Bass.
- Cronbach, Lee 1982 *Designing Evaluations of Educational and Social Programs*. San Francisco: Jossey-Bass.
- , Sueann Ambron, Sanford Dornbusch, Robert Hess, Robert Hornik, D. C. Phillips, Decker Walker, and Stephen Weiner 1980 *Toward Reform of Program Evaluation*. San Francisco: Jossey-Bass.
- Desautels, L. Denis 1997 "Evaluation as an Essential Component of 'Value-for-Money.'" In Eleanor

- Clemimsky and William R. Shadish, eds., *Evaluation for the Twenty-first Century*. Thousand Oaks, Calif.: Sage.
- Figueredo, Aurelio 1993 "Critical Multiplism, Meta-Analysis, and Generalization: An Integrative Commentary." In L. Sechrest, ed., *Program Evaluation: A Pluralistic Enterprise* (New Directions for Program Evaluation, No. 60). San Francisco: Jossey-Bass.
- Freeman, Howard 1992 "Evaluation Research." In E. Borgatta and M. Borgatta, eds., *Encyclopedia of Sociology*. New York: Macmillan.
- Gore, Albert 1993 *From Red Tape to Results: Creating a Government That Works Better and Costs Less*. New York: Plume/Penguin.
- Guba, Egon, and Yvonna Lincoln 1981 *Effective Evaluation*. San Francisco: Jossey-Bass.
- Leviton, Laura, and Edward Hughes 1981 "Research on the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review* 5:525-548.
- Lincoln, Yvonna 1991 "The Arts and Sciences of Program Evaluation." *Evaluation Practice* 12:1-7.
- Reichardt, Charles, and Sharon Rallis 1994 "The Relationship Between the Qualitative and Quantitative Research Traditions." In Charles Reichardt and Sharon Rallis, eds., (New Directions for Program Evaluation, No. 61). San Francisco: Jossey-Bass.
- Rossi, Peter 1994 The War Between the Quals and the Quants. Is a Lasting Peace Possible? In Charles Reichardt and Sharon Rallis, eds., *The Qualitative-Quantitative Debate: New Perspectives* (New Directions for Program Evaluation, No. 61). San Francisco: Jossey-Bass.
- , and Howard E. Freeman 1993 *Evaluation*, fifth ed. Newbury Park, Calif.: Sage.
- Rutman, Leonard 1984 *Evaluation Research Methods*. Newbury Park, Calif.: Sage.
- Scriven, Michael 1991 *Evaluation Thesaurus*, fourth ed. Newbury Park, Calif.: Sage.
- 1993 *Hard-Won Lessons in Program Evaluation* (New Directions for Program Evaluation, No. 58). San Francisco: Jossey-Bass.
- Sechrest, Lee 1992 "Roots: Back to Our First Generations." *Evaluation Practice* 13:1-7.
- Shadish, William 1993 "Critical Multiplism: A Research Strategy and Its Attendant Tactics." In L. Sechrest, ed., *Program Evaluation: A Pluralistic Enterprise* (New Directions for Program Evaluation, No. 60). San Francisco: Jossey-Bass.
- , Thomas Cook, and Laura Leviton 1991 *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, Calif.: Sage.
- Sonnad, Subhash, and Edgar Borgatta 1992 "Evaluation Research and Social Gerontology." *Research on Aging* 14:267-280.
- Stephan, A. Stephen 1935 "Prospects and Possibilities: The New Deal and the New Social Research." *Social Forces* 13:515-521.
- Suchman, Edward 1967 *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. New York: Russell Sage Foundation.
- Weiss, Carol 1975 "Evaluation Research in the Political Context." In Marcia Guttentag and Elmer Struening, eds., *Handbook of Evaluation Research*, vol. 1. Beverly Hills, Calif.: Sage.
- 1987 "Evaluating Social Programs: What Have We Learned?" *Society* 25:40-45.

KARL KOSLOSKI

## EVENT HISTORY ANALYSIS

Event history analysis is a collection of statistical methods for the analysis of longitudinal data on the occurrence and timing of events. As used in sociology, event history analysis is very similar to linear or logistic regression analysis, except that the dependent variable is a measure of the likelihood or speed of event occurrence. As with other regression methods, event history analysis is often used to develop causal or predictive models for the occurrence of events. Event history analysis has become quite popular in sociology since the mid 1980s, with applications to such diverse events as divorces (Bennett et al. 1988), births (Kallan and Udry 1986), deaths (Moore and Hayward 1990), job changes (Carroll and Mayer 1986), organizational foundings (Halliday et al. 1987), migrations (Baydar et al. 1990), and friendship choices (Hallinan and Williams 1987).

Although event history methods have been developed and utilized by statistical practitioners in a variety of disciplines, the term *event history analysis* is primarily used in sociology and closely allied disciplines. Elsewhere the methodology is known as survival analysis (biology and medicine), failure-time analysis (engineering), or duration analysis (economics). Introductory treatments for social scientists can be found in Teachman (1983), Allison (1984, 1995), Tuma and Hannan (1984), Kiefer (1988), and Blossfeld and Rohwer (1995).

For a biostatistical point of view, see Collett (1994), Kleinbaum (1996), or Klein and Moeschberger (1997).

### EVENT HISTORY DATA

The first requirement for an event history analysis is event history data. An event history is simply a longitudinal record of when events occurred for an individual or a sample of individuals. For example, an event history might be constructed by asking a sample of people to report the dates of any past changes in marital status. If the goal is a causal analysis, the event history should also include information on explanatory variables. Some of these, such as race and gender, will be constant over time while others, such as income, will vary. If the timing of each event is known with considerable precision (as with exact dates of marriages), the data are called continuous-time data. Frequently, however, events are only known to have occurred within some relatively large interval of time, for example, the year of a marriage. Such data are referred to as discrete-time data or grouped data.

Event history data are often contrasted with panel data, in which the individual's status is known only at a set of regular, fixed points in time. For example, employment status and other variables may be measured in annual interviews. Panel data collected at frequent intervals can often be treated as discrete-time event history data. But if the intervals between data collections are long, one of the major attractions of event history analysis can be lost—the ability to disentangle causal ordering. While this capability is by no means unequivocal, the combination of event history data and event history analysis is perhaps the best available nonexperimental methodology for studying causal relationships.

### PROBLEMS WITH CONVENTIONAL METHODS

Despite the attractiveness of event history data, they typically possess two characteristics that make conventional statistical methods highly unsuitable. Censoring is the most common problem. Suppose, for example, that the aim is to study the causes of divorce. The sample might consist of a

number of couples who married in 1990 and who are followed for the next five years. For the couples who get divorced, the length of the marriage is the principal variable of interest. But a large fraction of the couples will not divorce during the five-year interval. Marriages that are still in progress when the study ends are said to be *censored*. The problem is to combine the data on timing with the data on occurrence in a statistically consistent fashion. Ad hoc methods, such as excluding the censored cases or assigning the maximum length of time observed, can lead to substantial biases or loss of precision.

The second problem is time-varying explanatory variables (also known as time-dependent covariates). Suppose, in our divorce example, that the researcher wants to include number of children as a predictor of divorce. But number of children may change over the marriage, and it is not obvious how such a variable should be included in a regression model. If there were no censored cases, one might be tempted to regress the length of the marriage on the number of children at the end of the marriage. But longer marriages are likely to have produced more children simply because more time is available to have them. This would produce a spurious positive relationship between number of children and the length of the marriage.

One method for dealing with the censoring problem has been around since the seventeenth century and is still widely used—the *life table*. The life table is one example of a variety of methods that are primarily concerned with estimating the *distribution* of event times without regard for the effects of explanatory variables. For a comprehensive survey of such methods, see Elandt-Johnson and Johnson (1980). The remainder of this article focuses on regression methods that estimate the effects of explanatory variables on the occurrence and timing of events.

### ACCELERATED FAILURE-TIME MODELS

Suppose the goal is to estimate a model predicting the timing of first marriages, and the sample consists of women who are interviewed at age twenty-five. For each woman ( $i=1, \dots, n$ ), we learn her age in days at the time of the marriage, denoted by  $T_i$ .

For women who still were not married at age twenty-five (the censored cases),  $T_i^*$  is their age in days at the time of the interview. We also have data on a set of explanatory variables  $x_{i1}, \dots, x_{i11}$ . For the moment, let us suppose that these are variables that do not change over time, such as race, parents' education, and eighth-grade test scores.

One class of models that is appropriate for data such as these is the accelerated failure-time (AFT) model. The general formulation is

$$\log T_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_2 x_{i2} + \varepsilon_i \quad (1)$$

By taking the logarithm on the left-hand side, we ensure that  $T_i$  is always greater than 0, regardless of the values of the  $x$  variables. Specific cases of this general model are obtained by choosing particular distributions for the random disturbance  $\varepsilon_i$ . The most common distributions are normal, extreme-value, logistic, and log-gamma. These imply that  $T_i$  has distributions that are, respectively, lognormal, Weibull, log-logistic, and gamma, which are the names usually given to these models. The disturbance term  $\varepsilon_i$  is assumed to be independent of the  $x$ 's and to have constant variance.

If there are no censored data, these models can be easily estimated by ordinary least-squares regression of  $\log T$  on the  $x$ 's. The resulting coefficients are best linear unbiased estimators. But the presence of censored data requires a different method. The standard approach is maximum likelihood, which combines the censored and uncensored data in an optimal fashion. Maximum likelihood estimation for these models is now widely available in several statistical packages (e.g., BMDP, LIMDEP, SAS, SYSTAT, Stata).

Here's an example from criminology. In the early 1970s, 432 inmates from Maryland state prisons were followed for one year after their release (Rossi et al. 1980). The event of interest is the first arrest that occurred to each person during the one-year observation period. Only 26 percent of the released inmates were arrested. We'll use the following variables:

ARREST 1 if arrested, otherwise 0

WEEK Week of first arrest for those who were arrested (ranges 1 to 52); for those not arrested, week 52

FIN 1 if they received financial aid after release, otherwise 0

AGE Age in years at the time of release

RACE 1 if black, otherwise 0

MAR 1 if married at the time of release, otherwise 0

PRIO Number of prior convictions

Using these data, I estimated a Weibull version of the accelerated failure time model by maximum likelihood with the SAS® statistical package. Results are shown in Table 1. Looking first at the  $p$ -values, we see that race and marital status do not have a significant impact on the timing of arrests. On the other hand, we see highly significant effects of age and number of prior convictions, and a just barely significant effect of financial aid.

The negative coefficient for PRIO tells us that having more prior convictions is associated with *shorter* times to arrest. The positive coefficient for AGE tells us that inmates who were older when they released have *longer* times to arrest. Similarly, those who got financial aid have longer times to arrest. We can interpret the magnitudes of the coefficients by applying the transformation  $100[\exp(\beta)-1]$ , which gives the percentage change in time to event for a 1-unit increase in a particular independent variable. For PRIO we get  $100[\exp(-.071)-1]=-6.8$ , which tells us that each additional conviction lowers the time to arrest by 6.8 percent, controlling for other variables in the model. For FIN we get  $100[\exp(.268)-1]=31$ . Those who got financial aid have times to arrest that are 31 percent longer than those who did not get financial aid.

In addition to the Weibull model, I also estimated gamma, lognormal, and log-logistic models. Results were very similar across the different models.

### PROPORTIONAL HAZARDS MODELS

A second class of regression models for continuous-time data is the proportional hazards model. To explain this model, it is first necessary to define the hazard function, denoted by  $h(t)$ , which is the fundamental dependent variable. Let  $P(t+\Delta t)$  be the conditional probability that an event occurs in the time interval  $(t+\Delta t)$ , given that it has not already

**Results from Weibull Regression Model Predicting the Time of First Arrest**

Variable	Coefficient	Standard Error	Chi-Square	p-value
FIN	.268	.137	3.79	.05
AGE	.043	.015	8.00	.004
RACE	-.226	.220	1.06	.30
MAR	.353	.269	1.73	.19
PRIO	-.071	.020	12.85	.0003
Intercept	4.037	.402		

Table 1

occurred prior to  $t$ . To get the hazard function, we divide this probability by the length of the interval  $\Delta t$ , and take the limit as  $\Delta t$  goes to 0:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t, t + \Delta t)}{\Delta t} \tag{2}$$

Other common symbols for the hazard function are  $r(t)$  and  $\lambda(t)$ . The hazard may be regarded as the instantaneous likelihood that an event will occur at exactly time  $t$ . It is not a probability, however, since it may be greater than 1.0 (although never less than 0).

If  $h(t)$  has a constant value  $c$ , it can be interpreted as the expected number of events in a 1-unit interval of time. Alternatively,  $1/c$  is the expected length of time until the next event. Suppose, for example, that the events are residence changes, time is measured in years, and the estimated hazard of a residence change is .20. That would imply that, for a given individual, the expected number of changes in a year is .20 and the expected length of time between changes is  $1/.20 = 5$  years.

Like a probability (from which it is derived), the hazard is never directly observed. Nevertheless, it governs both the occurrence and timing of events, and models formulated in terms of the hazard may be estimated from observed data.

The general proportional hazards (PH) model is given by

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i1} + \dots + \beta_k x_{ik} \tag{3}$$

where  $\alpha(t)$  may be any function of time. It is called the proportional hazards model because the ratio of the hazards for any two individuals is a constant over time. Notice that, unlike the AFT model,

there is no disturbance term in this equation. That does not mean that the model is deterministic, however, because there is random variation in the relationship between  $h(t)$  and the observed occurrence and timing of events.

Different versions of the PH model are obtained by choosing specific forms for  $\alpha(t)$ . For example, the Gompertz model sets  $\alpha(t) = \alpha_1 + \alpha_2 t$ , which says that the hazard is an increasing (or decreasing) function of time. Similarly, the Weibull model has  $\alpha(t) = \alpha_1 + \alpha_2 \log t$ . (The Weibull model is the only model that is a member of both the AFT class and the PH class.) The exponential model—a special case of both the Weibull and the Gompertz models—sets  $\alpha(t) = \alpha$ , a constant over time. For any specific member of the PH class, maximum likelihood is the standard approach to estimation.

In a path-breaking paper, the British statistician David Cox (1972) showed how the PH model could be estimated without choosing a specific functional form for  $\alpha(t)$ , using a method called partial likelihood. This method is very much like maximum likelihood, except that only a part of the likelihood function is maximized. Specifically, partial likelihood takes account only of the ordering of events, not their exact timing. The combination of partial likelihood and the proportional hazards model has come to be known as Cox regression. The method has become extremely popular because, although some precision is sacrificed, the resulting estimates are much more robust. Computer programs that implement this method are now available in most full-featured statistical packages (SPSS, SAS, LIMDEP, BMDP, S-Plus, Stata, SYSTAT).

As an example, we'll estimate a proportional hazards model for the recidivism data discussed

**Results from Cox Regression Model Predicting the Hazard of First Arrest**

Variable	Coefficient	Standard Error	Chi-Square	p-value
FIN	-.373	.191	3.82	.05
AGE	-.061	.021	8.47	.004
RACE	.317	.308	1.06	.30
MAR	-.493	.375	1.73	.19
PRIO	.099	.027	13.39	.0003
Intercept	4.037	.402		

Table 2

earlier. The chi-squares and  $p$ -values shown in Table 2 for the five variables are remarkably similar to those in Table 1 that were obtained with maximum likelihood estimation of a Weibull model. On the other hand, the coefficients are noticeably different in magnitude and even have signs that are reversed from those in Table 1. The sign reversal is a quite general phenomenon that stems from the fact that the dependent variable is the time of the event in the AFT model and the hazard of the event in the PH model. People with high hazards are very likely to have events at any point in time, so their times to events tend to be short. By contrast, people with low hazards tend to have long times until event occurrence.

To interpret the magnitudes of the coefficients, we can use the same transformation used for the AFT models. Specifically,  $100[\exp(\beta)-1]$  gives the percentage change in the hazard of an event for a 1-unit increase in a particular independent variable. Thus, for FIN we have  $100[\exp(-.373)-1] = -31$ , which says that those who got financial aid have hazards of arrest that are 31 percent lower than those who did not get aid. For AGE we have  $100[\exp(-.061)-1] = -6$ . Each additional year of age at release yields a 6 percent reduction in the hazard of arrest. Finally, each additional conviction is associated with  $100[\exp(.099)-1] = 10.4$  percent increase in the hazard of an arrest.

The partial likelihood method also allows one to easily introduce time-varying explanatory variables. For example, suppose that the hazard for arrest is thought to depend on both financial aid ( $x_1$ ) and employment status ( $x_2$ , coded 1 for employed, 0 for unemployed). A suitable PH model might be

$$\log h(t) = \alpha(t) + \beta_1 x_1 + \beta_2 x_2(t) \quad (4)$$

which says that the hazard at time  $t$  depends on financial aid, on employment status at time  $t$ , and on time itself. If longitudinal data on income are available, models such as this can be estimated in a straightforward fashion with the partial likelihood method. How to do this is shown in Chapter 5 of Allison (1995).

**MULTIPLE KINDS OF EVENTS**

To this point, it has been assumed that all events can be treated alike. In many applications, however, there is a compelling need to distinguish among two or more types of events. For example, if the events of interest are job terminations, one might expect that explanatory variables would have vastly different effects on voluntary and involuntary terminations. For recidivism studies, it might be desirable to distinguish arrests for crimes against persons and crimes against property. The statistical analysis should take these distinctions into account.

All of the methods already discussed can be easily applied to multiple kinds of events. In essence, a separate model is estimated for each kind of event. In doing an analysis for one kind of event, one simply treats other kinds of events as though the individual were censored at the time when the event occurred, a method known as "competing risks." Thus, no new methodology is required to handle this situation.

An alternative approach is to estimate a single event history model for the timing of events, without distinguishing different event types. Then,



after eliminating all the individuals who did not have events (the censored cases), one estimates a logistic regression model for the determinants of the type of event. This method of analysis is most appropriate when the different kinds of events are functionally alternative ways of achieving a single objective. For example, the event might be purchase of a computer and the two different types might be a Windows-based computer versus a Macintosh computer.

### REPEATED EVENTS

The discussion so far has presumed that each individual experiences no more than one event. Obviously, however, such events as childbirths, job changes, arrests, and car purchases can occur many times over the life of an individual. The methods already described have been routinely applied to cases of repeated events, taking one of two alternative approaches. One approach is to do a separate analysis for each successive event. For example, one event history model is estimated for the birth of the first child, a second model is estimated for the birth of the second child, and so on. The alternative approach is to break each individual's event history into a set of intervals between events, treat each of these intervals as a distinct observation, and then pool all the intervals into a single analysis.

Neither of these alternatives is entirely satisfactory. The sequential analysis is rather tedious, wastes information if the process is invariant across the sequence, and is prone to selection biases for later events in the sequence. The pooled analysis, on the other hand, makes the rather questionable assumption that the multiple intervals for a single individual are independent. This can lead to standard errors that are biased downward and test statistics that are biased upward.

Several methods are available for dealing with the lack of independence. One is to estimate standard errors and test statistics using the robust method developed by White (1982). These robust standard errors have been incorporated into some Cox regression programs (e.g., Stata, S-Plus). Another is to do a "fixed-effects" Cox regression that stratifies on the individual (Allison 1996; Yamaguchi 1986). These models can be estimated with most Cox regression software, and they have the advantage of automatically controlling for all stable

characteristics of the individual. On the other hand, fixed-effects models cannot produce coefficient estimates for stable characteristics such as sex or race. Finally, there are "random-effects" or "frailty" models that explicitly build the dependence into the model as an additional random disturbance. Unfortunately, there is little commercial software available to estimate random-effects event history models.

### DISCRETE-TIME METHODS

When event times are measured coarsely, the continuous-time methods already discussed may yield somewhat biased estimates. In such cases, methods specifically designed for discrete-time data are more appropriate (Allison 1982). Moreover, such methods are easily employed and are particularly attractive for handling large numbers of time-varying explanatory variables.

Suppose that the time scale is divided into a set of equal intervals, indexed by  $t = 1, 2, 3, \dots$ . The discrete-time analog of the hazard function, denoted by  $P_t$ , is the conditional probability that an event occurs in interval  $t$ , given that it has not occurred prior to  $t$ . A popular model for expressing the dependence of  $P_t$  on explanatory variables is the logit model

$$\log\left(\frac{P_{it}}{1-P_{it}}\right) = \alpha_t + \beta_1 x_{i1} + \dots + \beta_2 x_{i2} \quad (5)$$

where the subscript on  $\alpha_t$  indicates that the intercept may differ for each interval of time. Similarly the explanatory variables may take on different values at each interval of time. This model can be estimated by the method of maximum likelihood, using the following computational strategy:

1. Break each individual's event history into a set of discrete time units, for example, person-years.
2. Create a dependent variable that has a value of 1 for time units in which events occurred; otherwise use 0. Explanatory variables are assigned whatever values they had at the beginning of the time unit.
3. Pool all these time units, and estimate a logistic regression model using a standard maximum likelihood logit program.

Other models and computational methods are also available for the discrete-time case. These methods can be easily extended to allow for multiple kinds of events and repeated events.

## REFERENCES

- Allison, Paul D. 1982 "Discrete Time Methods for the Analysis of Event Histories." In Samuel Leinhardt, ed., *Sociological Methodology 1982*. San Francisco: Jossey-Bass.
- 1984 *Event History Analysis*. Beverly Hills, Calif.: Sage.
- 1995 *Survival Analysis Using the SAS® System: A Practical Guide*. Cary, N.C.: SAS Institute.
- 1996 "Fixed Effects Partial Likelihood for Repeated Events." *Sociological Methods and Research* 25:207–222.
- Baydar, Nazli, Michael J. White, Charles Simkins, and Ozer Babakol 1990 "Effects of Agricultural Development Policies on Migration in Peninsular Malaysia." *Demography* 27:97–110.
- Bennett, Neil G., Ann Klimas Blanc, and David E. Bloom 1988 "Commitment and the Modern Union: Assessing the Link Between Premarital Cohabitation and Subsequent Marital Stability." *American Sociological Review* 53:127–138.
- Blossfeld, Hans-Peter, and Götz Rohwer 1995 *Techniques of Event History Modeling: New Approaches to Causal Modeling*. Mahwah, N.J.: Lawrence Earlbaum.
- Carroll, Glenn R., and Karl Ulrich Mayer 1986 "Job-Shift Patterns in the Federal Republic of Germany: The Effects of Social Class, Industrial Sector, and Organizational Size." *American Sociological Review* 51:323–341.
- Collett, D. 1994 *Modelling Survival Data in Medical Research*. London: Chapman and Hill.
- Cox, David R. 1972 "Regression Models and Life Tables." *Journal of the Royal Statistical Society, Series B*, 34:187–202.
- Elandt-Johnson, R. C., and N. L. Johnson 1980 *Survival Models and Data Analysis*. New York: Wiley.
- Halliday, Terence C., Michael J. Powell, and Mark W. Granfors 1987 "Minimalist Organizations: Vital Events in State Bar Associations, 1870–1930." *American Sociological Review* 52:456–471.
- Hallinan, Maureen, and Richard A. Williams 1987 "The Stability of Students' Interracial Friendships." *American Sociological Review* 52:653–664.
- Kallan, Jeffrey, and J. R. Udry 1986 "The Determinants of Effective Fecundability Based on the First Birth Interval." *Demography* 23:53–66.
- Kiefer, Nicholas M. 1988 "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26:646–679.
- Klein, John P., and Melvin Moeschberger 1997 *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer-Verlag.
- Kleinbaum, David G. 1996 *Survival Analysis: A Self-Learning Text*. New York: Springer-Verlag.
- Moore, David E., and Mark D. Hayward 1990 "Occupational Careers and Mortality of Elderly Men." *Demography* 27:31–53.
- Rossi, P. H., R. A. Berk, and K. J. Lenihan 1980 *Money, Work and Crime*. New York: Academic Press.
- Teachman, Jay D. 1983 "Analyzing Social Processes: Life Tables and Proportional Hazards Models." *Social Science Research* 12:263–301.
- Tuma, Nancy Brandon, and Michael T. Hannan 1984 *Social Dynamics: Models and Methods*. Orlando, Fla.: Academic Press.
- White, Halbert 1982 "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50:1–25.
- Yamaguchi, Kazuo 1986 "Alternative Approaches to Unobserved Heterogeneity in the Analysis of Repeated Events." In Nancy Brandon Tuma, ed., *Sociological Methodology 1986*. Washington, D.C.: American Sociological Association.

PAUL D. ALLISON

## EVOLUTION: BIOLOGICAL, SOCIAL, CULTURAL

The diverse forms of life on earth have emerged probably from a common source, through a process of evolution that has the following characteristics:

1. The course of evolution does not always proceed along a straight path (for example, from simple to complex forms). Instead, it can meander like a stream, directed largely by environmental circumstances, although also affected by limitations on the capacities of organisms to respond to environmental challenges, and by unpredictable "chance" factors. Occasionally it reverses direction in certain respects, as when our own distant monkey-like ancestors became adapted to life in



the trees, and our more recent ancestors readapted to living on the ground. When the course of evolution does reverse direction, it generally does so with respect to comparatively few features only, not all features. Thus, we humans retain various characteristics evolved earlier in connection with life in the trees: stereoscopic vision, visual acuity, reduced sense of smell, and hands adapted for grasping.

2. Different groups of organisms sometimes evolve in similar directions in certain respects when exposed to similar environmental conditions. Thus whales—descendants of mammals that lived on land—acquired fishlike shapes when they adapted to life in the water. However, parallelism in evolutionary development generally remains limited: different evolutionary lines do not come close to “merging.” Whales, even though living in the ocean and acquiring fishlike shapes, retain basic anatomical and physiological features of mammals, quite different from those of fish.

The life cycle of a human being involves fixed stages (infancy, childhood, adolescence, adulthood, old age) that represent an unfolding of innate potentialities and that culminate in inevitable death. The evolution of a new biological species, by contrast, is a unique historical development involving movement in directions shaped primarily by environmental pressures, without any inevitable “death” or other pre-determined end state.

A “Lamarckian” evolutionary process (named after Jean Baptiste Lamarck, 1744–1829), would involve inheritance of acquired characteristics. Thus, if evolution followed the Lamarckian pattern, giraffes might have acquired their long necks because in each generation necks were stretched (perhaps to obtain food high up in trees, or to detect approaching enemies), and because the effect of each generation’s stretching tended to be inherited by the generation that followed. However, Lamarckian ideas were found to be invalid long ago.

A different approach developed by Charles Darwin (1859) and others does not assume that

acquired characteristics (such as effects of neck-stretching) are inherited. Rather, it assumes (1) random or randomlike variation among the offspring in each generation (with some giraffes happening by chance to have longer necks than others); (2) natural selection, involving tendencies for certain variants (longer-necked giraffes) to survive and reproduce more than others; and (3) preservation through inheritance of the effects of natural selection (longer-necked giraffes tending to have similarly long-necked offspring, although—as suggested above—there would still be some random variation among these offspring with respect to neck length).

Darwinian ideas challenged traditional Christian religious beliefs by suggesting (1) that we are descended from ape-like creatures and, ultimately, from elementary life forms; (2) that our evolution was basically an unplanned outcome of diverse environmental pressures rather than something planned in advance; and (3) that the earth is old enough for evolutionary processes to have had time to produce the variety and complexity of life forms that we actually find. Although today still resisted by many on religious grounds, Darwinian theory ultimately came to be generally accepted by biologists, with important modifications and certain disagreements about details, and in combination with new knowledge in other areas of biology and other scientific disciplines that was not available to Darwin (see Gould 1982; Stebbins and Ayala 1985; Mayr and Provine 1998).

An inherited trait that has evolved over a long period of time (like the giraffe’s long neck) has very likely evolved because it makes some contribution to the survival and reproduction of the organisms possessing it, and because it has consequently emerged and persisted through the Darwinian mechanisms outlined above. However, some traits may appear and persist that are neutral in their implications for survival and reproduction, and even traits that have negative implications may appear and persist if they happen to be linked with traits whose implications are positive.

Bioevolutionary explanations of human social phenomena, sometimes under the label of “sociobiology” (see Wilson 1975), pose a challenge to sociology, as the following example suggests. Stepparents tend to abuse their stepchildren

more than biological parents tend to abuse their own children. Sociologists would ordinarily seek to explain this in a way that utilizes sociological and social-psychological concepts only, avoiding assumptions about inherited and biologically evolved traits. A sociobiologist, by contrast, would be more likely to view it as a human manifestation or extension of a biologically evolved tendency prevailing widely among nonhuman mammals: a tendency for individual mammals in stepparent-like positions to kill their “stepchildren,” thus increasing their own prospects for producing survivable offspring and perpetuating their own genes (see Beckstrom 1993, pp. 23–29).

An evolutionary explanation of an aspect of human life or society may present that aspect as having emerged (1) as a part of, or a continuation of, biological evolution (as illustrated in the discussion of stepparental child abuse above), or (2) through a process analogous to biological evolution but nevertheless distinct from it, that is, a process of social or cultural evolution. Concepts analogous to those of biological evolution are especially applicable to aspects of culture such as science (Hull 1988) and technology (Basalla 1988) in which the cumulative character of culture is most strongly manifested. Competing scientific hypotheses are the “randomlike variations” in science as an evolutionary process. Research results that evaluate such hypotheses select some of them for survival and others for extinction, and the knowledge that constitutes the outcome of this process in any given generation of scientists is “inherited” by subsequent generations through textbooks, teaching, and research publications. In technological evolution, positive selection (i.e., acceptance) of variations (innovations) depends not only on research results (i.e., on how well they “work”) but also on costs, competitive pressures, compatibility with prevailing culture, and other factors.

The evolutionary model is not appropriately applied to social or cultural changes that are cyclical, easily reversed, primarily planned, or repetitive. It may be usefully applied when a complex social or cultural transformation that would be hard to repeat or to reverse occurs gradually without being planned, through environmentally determined selections from among divergent alternative directions of change, and with enough

stability to preserve the results of past environmental influence (see Campbell 1965; Richter 1982, pp. 19–34).

Not every process that is called “evolutionary” seems to deserve that label. Some theories pertaining to major transformations of society illustrate this point.

The idea that human society evolves from simple beginnings through comparatively fixed stages came to be commonly accepted in the nineteenth century, although different theorists had different conceptions of what these stages were. Thus we find analyses of transitions from theological to metaphysical to “positivistic” thought-styles (Comte 1875); from savagery to barbarism to civilization (Morgan 1877); from tribalism to slavery, feudalism, capitalism, and then communism (Marx and Engels [1846] 1947); and from simple to compounded, doubly compounded and then triply compounded societies (Spencer [1885–1886] 1967).

When societies, cultures, or civilizations are said to pass from childhood to adulthood, and then to old age, or are said to grow and then decline, we have an analogy not with the evolution of a species but rather with the human life cycle, as illustrated in the works of Oswald Spengler and Arnold J. Toynbee. Nineteenth century “social-evolutionary” theories departed from the life-cycle model in that they did not involve the idea of decline or old age followed by death as the end point of a cycle. Quite the opposite: the trends they described tended to culminate in triumphant achievements—“positivism” (Comte), “civilization” (Morgan), “communism” (Marx and Engels), and societies “compounded” many times (Spencer). In this one respect, but only in this respect, these theories resembled the bioevolutionary model: biological species do not have to become extinct—even though many of them actually do—in the same sense that individuals have to die. However, these theories nevertheless reflect a life-cycle model, insofar as they involve one society after another following essentially the same pattern of development, just as one infant after another follows the same general route to adulthood. And they do not appear to involve, in any major way, the evolutionary mechanisms of random variation and selection. In fact, they are primarily theories of progress, not of evolution in any sense that biologists

would recognize, regardless of the “evolutionary” label commonly attached to them (Nisbet 1969).

The social-evolutionary idea fell into disfavor among sociologists around the turn of the twentieth century, but was revived several decades later with, for example, Talcott Parsons’ analysis of primitive, intermediate, and modern societies (1977), and Lenski and Lenski’s (1982) analysis of transitions from hunting-and-gathering to horticultural, agrarian, and industrial societies. Although based on more accurate and more extensive facts than were available in the nineteenth century, these newer schemes pertaining to the evolution of whole societies are nevertheless basically similar to their nineteenth century predecessors in that they present various societies as passing through specific stages in a predominantly unidirectional pattern, which, as noted above, represents a divergence from the bioevolutionary model. This is not necessarily a defect: the development of societies of new types may simply not be a process to which concepts of a bioevolutionary sort are fully applicable.

Evolutionary concepts have been used not only in attempts to understand some social and cultural phenomena, as discussed above, but also in attempts to formulate and justify certain social policies. In the late nineteenth and early twentieth centuries, evolutionary concepts and associated slogans such as “survival of the fittest” appeared to provide rationales for what have been labelled “Social Darwinist” policies. Focusing on the selective (rather than on the random-variation) phase of the evolutionary process, various business, political, ideological, and military leaders in several countries, along with some scholars in several disciplines (including the sociologists Herbert Spencer and William Graham Sumner), emphasized the importance of the struggle for survival in maintaining a hardy population and a vigorous society, and opposed social welfare measures that they thought would mitigate that struggle. Supporters of this approach differed in several ways: some emphasized individual struggle, others, group (e.g., racial or national) struggle; some emphasized nonviolent means (e.g., economic competition), and others, armed conflict (see Hofstadter 1955).

A statement by Darwin himself illustrates the basic idea underlying “Social Darwinism.” Although Darwin did not oppose smallpox vaccination, he

worried about its effects: “There is reason to believe that vaccination has preserved thousands, who from a weak constitution would formerly have succumbed to smallpox. Thus the weak members of civilized societies propagate their kind. No one who has attended to the breeding of domestic animals will doubt that this must be highly injurious to the race of man” (Darwin [1871] 1897, p. 134).

Darwin did not believe that war led to the selective survival of the fittest people, as he thought smallpox tended to do. Quite the opposite: he saw war as destroying “the finest” young men while those with “poor constitutions” would avoid fighting and thus survive and reproduce.

But evolutionary concepts can be applied at diverse levels. Nineteenth-century war might have led to the selective survival of societies or governments that were “fittest” in some sense, even if not to the survival of the “fittest” young men within particular societies. Some Social Darwinists adopted that perspective. Thus: “Storm purifies the air and destroys the frail trees, leaving the sturdy oaks standing. War is the test of a nation’s political, physical and intellectual worth. The State in which there is much that is rotten may vegetate for a while in peace, but in war its weakness is revealed. . . . It is better to spend money on armaments and battleships than luxury, motormanias and other sensual living” (Baron Karl von Stengel, cited by Angell 1911, p. 168).

There is a problem here: what an exceedingly powerful storm is most likely to leave standing are not sturdy oaks but blades of grass. And this leads to a question: what does “survival of the fittest” actually mean? Defining the fittest as “those who survive” would make the phrase tautological, but any other definition would very likely make it untrue except under a limited range of conditions. In any case, “fitness” is not necessarily equivalent to “desirability” as we might define that, or to what the people of a society actually desire.

The problem of defining “fitness” may be illustrated by examining what Darwin said about smallpox. Suppose people who would have succumbed to smallpox if not vaccinated, but who were actually saved by vaccination, have descendants who inherit their potential vulnerability to that disease. In this respect, such descendants might be considered relatively “unfit.” But if these descendants are all protected from smallpox by

vaccination, or if smallpox itself is permanently eliminated, then “potential vulnerability to smallpox” might have no practical implications and might not entail “unfitness” in any practically relevant sense.

If the term “Social Darwinism” is to be applied to attempts to improve society through competition and struggle that eliminate some variants while perpetuating others, then attempts to improve society by improving people within society (through education, health care, and so on) might reasonably be labelled “Social Lamarckism.” The fact that Lamarckian ideas have been empirically disconfirmed and rejected in explanations of biological evolution has no bearing on the question of the reasonableness of analogous ideas in the socio-cultural realm. And, in both biological and socio-cultural contexts, Lamarckian and Darwinian ideas can logically coexist and have done so. Darwin himself accepted the validity of Lamarckian ideas in explaining biological evolution, even while he was developing different ideas of his own; the decisive rejection of Lamarckism by evolutionary biologists came after Darwin had died. And Herbert Spencer, who coined the Darwinian phrase “survival of the fittest” and has traditionally been considered a major Social Darwinist, has also been called a Social Lamarckist (Bowler 1995, p. 113) on the ground that he was more committed to struggle as a (Lamarckian) means for invigorating and improving the individuals who participate in it, than to struggle as a (Darwinian) means for eliminating the “unfit.”

Partial artificial control over reproduction and hence over the evolutionary process is illustrated by methods used to develop new variants of farm animals, work animals, laboratory animals, and pets. Improvements in these methods have ultimately led to “cloning,” which involves total and exact replication of individuals’ inherited characteristics without mating. Control over human reproduction was sought on a limited basis by a “eugenics” movement that emerged in the late nineteenth century, founded by Darwin’s cousin Francis Galton. This movement achieved some successes, including arrangements for sterilization of some criminals and mentally ill people in various jurisdictions. However, it also faced problems, including: lack of consensus about the traits to be targeted; complexities in the process of hereditary transmission (for example, the fact that parents

may transmit to their offspring a defect that they do not manifest themselves); political and legal obstacles to compulsory eugenic measures in societies with well-established traditions of democracy and personal freedom; and, the difficulty of distinguishing between traits transmitted through heredity and traits transmitted through socialization. Supporters of eugenics did not always care about this last point: they sometimes assumed that preventing people with undesirable traits from having children would reduce the prevalence of people bearing these traits regardless of how the traits themselves were transmitted from one generation to the next.

The eugenics movement developed in a tragically racist direction. Some major leaders of the movement in the United States thought that white people of northern European ancestry constituted a superior race and regarded as genetically inferior the southern and eastern Europeans who were immigrating to the country in large numbers. The movement came to be discredited after its ideas came to be used in Hitler’s Germany to support genocidal killings of Jews and members of various other groups.

## REFERENCES

- Angell, Norman 1911 *The Great Illusion*. New York: Putnam.
- Basalla, George 1988 *The Evolution of Technology*. New York: Cambridge University Press.
- Beckstrom, John H. 1993 *Darwinism Applied: Evolutionary Paths to Social Goals*. Westport, Conn.: Praeger.
- Bowler, Peter J. 1995 “Social Metaphors in Evolutionary Biology, 1870–1930.” In Sabine Maasen, Everett Mendelsohn, and Peter Weingart, eds., *Biology as Society, Society as Biology: Metaphors*. Dordrecht: Kluwer Academic Publishers.
- Campbell, Donald T. 1965 “Variation and Selective Retention in Socio-Cultural Evolution.” In Herbert R. Barringer, George I. Blanksten, and Raymond W. Mack, eds., *Social Change in Developing Areas*. Cambridge, Mass.: Schenkman.
- Comte, Auguste 1875 *The Positive Philosophy*, 2nd ed. Harriet Martineau, trans. and cond. London: Trubner.
- Darwin, Charles 1859 *On the Origin of Species*. London: Murray.
- Darwin, Charles 1897 *The Descent of Man*. New York: D. Appleton.
- Gould, Stephen Jay 1982 “Darwinism and the Expansion of Evolutionary Theory.” *Science* 216:380–387.

- Hofstadter, Richard 1955 *Social Darwinism in American Thought*. Boston: Beacon Press.
- Hull, David L. 1988 *Science As a Process*. Chicago: University of Chicago Press.
- Lenski, Gerhard, and Jean Lenski 1982 *Human Societies: An Introduction to Macrosociology*. New York: McGraw-Hill.
- Marx, Karl, and Friedrich Engels 1947 *The German Ideology*. New York: International Publishers.
- Mayr, Ernst, and William B. Provine, eds., 1998 *The Evolutionary Synthesis*. Cambridge, Mass.: Harvard University Press.
- Morgan, Lewis H. 1877 *Ancient Society*. Chicago: Charles H. Kerr and Co.
- Nisbet, Robert A. 1969 *Social Change and History*. London: Oxford University Press.
- Parsons, Talcott 1977 *The Evolution of Societies*. Jackson Toby, ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Richter, Maurice N., Jr. 1982 *Technology and Social Complexity*. Albany, N.Y.: State University of New York Press.
- Spencer, Herbert 1967 *The Evolution of Society*. Robert L. Carneiro, ed. Chicago: University of Chicago Press. Excerpted from *Principles of Sociology*, vols. 1 and 2, rev. ed. London: Williams and Margate, 1885–1886.
- Stebbins, G. Ledyard, and Francisco J. Ayala 1985 “The Evolution of Darwinism.” *Scientific American* 253:72–82.
- Wilson, Edward O. 1975 *Sociobiology: The New Synthesis*. Cambridge, Mass.: The Belknap Press of Harvard University Press.

MAURICE N. RICHTER, JR.

## EXCHANGE THEORY

See Social Exchange Theory.

## EXPECTATION STATES THEORY

Expectation states theory is a well-established, ongoing research program investigating various aspects of group interaction. The focus is on small, task-oriented groups; the central interest is in both the processes through which group members assign levels of task competence to each other and the consequences this assignment has for their interaction. Originating as a single theory developed by Joseph Berger (1958), expectation states

theory has grown to include various branches sharing a core of basic concepts, definitions, and propositions about group interaction processes, as well as methodological and metatheoretical assumptions. Thus, the program in fact contains not just one theory but several. (Unless otherwise specified, the expression “expectation states theory” refers here to the entire program rather than to any particular theory within it.) Expectation states theory has received strong support from extensive empirical research.

Two key concepts in the program are “status characteristics” and “performance expectations.” A status characteristic is any valued attribute implying task competence. Such characteristics are viewed as having at least two levels (e.g., being either high or low in mechanical ability, being either male or female), one carrying a more positive evaluation than the other. They are also defined as varying from specific to diffuse, depending on the range of their perceived applicability. For instance, mechanical ability is usually considered to be relatively specific, or associated with well-defined performance expectations. Gender, on the other hand, is frequently treated as diffuse, or carrying both limited and general performance expectations. The “diffuseness” refers to the fact that since there is no explicitly set limit to the expectations, the characteristic is viewed as relevant to a large, indeterminate number of different tasks. Other attributes commonly treated as diffuse status characteristics are ethnicity, skin color, socio-economic class, level of education, organizational rank, age, and physical attractiveness.

Performance expectations link status characteristics to observable behavior. Thus, levels of these characteristics are associated with levels of competence and corresponding expectations that, in turn, determine what is known as “the power and prestige order of the group.” This concept refers to a set of interrelated behaviors, namely, the unequal distribution in the offer and acceptance of opportunities to perform, the type of evaluations received for each unit of performance, and the rates of influence exerted among group members. Note, then, that performance expectations are distinguished from evaluations of units of performance: while the latter are evaluations of a single act, the former refer to the level of competence that a person is predicted to exhibit over a

number of such performances. Once formed, expectations tend to be stable, since the behaviors that make up the power and prestige order of a group operate in a way that reinforces the status quo.

Research in expectation states theory follows a situational approach. Propositions are formulated from the point of view of an individual (also referred to as “self” or “focal actor”) who performs a given task with a partner (“other”). Performance expectations are therefore relative to a specific pair of actors, and an individual is not seen as holding high (or low, or medium) expectations for self but, rather, an “expectation state” that is defined in terms of self *and* other. In other words, the focal actor is said to hold expectations for self that are either higher than, lower than, or equal to, the expectations he or she holds for the partner. Such expectations are also conceptualized as relative to a particular situation. Thus, for example, with a different partner or in the context of a different task, a person’s expectation state could vary. It should also be noted that an expectation state is a theoretical construct, not an observable phenomenon. In particular, although expectations are seen as reflecting a person’s beliefs about the distribution of task competence in the group, they are *not* assumed to be self’s conscious calculations of advantage (or disadvantage, or equality) in this respect. Rather, what is proposed are models to be used to predict the focal actor’s behavior, and this person is to be seen only *as if* he or she performed the operations specified in the models.

Propositions in expectation states theory also have been formulated within well defined scope conditions, or clauses specifying the limits within which the propositions apply. Since scope conditions are part of the theory, they are expressed in abstract terms (as are the key concepts of “status characteristic,” “performance expectations,” and “power and prestige order”). Three important scope conditions are that the focal actor is assumed to: (1) value the task in question (he or she views it as having two or more possible outcomes, none due exclusively to chance, and considers one of these to be success and the other failure); (2) be task-oriented (motivated to do the task well); and (3) be collectively oriented (prepared to accept the partner’s ideas if they are thought to contribute to the task solution). In other words, the third scope condition specifies that solving the task must be

more important to the focal actor than any other considerations (such as having his or her individual contributions accepted, or being liked by the partner).

The situational approach and the use of scope conditions are part of an overall theory-construction strategy whereby the initial focus is on the simplest contexts, which are studied with the aid of minimum assumptions. Complexities are then added gradually as knowledge accumulates. More recent expectation states research has, for example, extended the investigation from dyads to larger groups. Similarly, the scope condition of collective orientation has been relaxed in some of these studies, with results showing that status still affects expectations under such circumstances.

The core of the theory investigates the performance expectations an individual forms about self and other in settings where the two are engaged in the joint solution of a task. As indicated previously, the focal actor is assumed to value the task and to be both task-oriented and collectively oriented. A valued task implies that competence at the task is more desirable than incompetence, and therefore that levels of task ability constitute a status characteristic for that actor. Basically, the assignment of levels of this characteristic and corresponding performance expectations may occur in one of two ways: directly, from actual evaluations of task performance, or indirectly, on the basis of other status characteristics of the performers that self perceives to be relevant to the task. For simplicity, I first consider situations where expectations are formed exclusively in one fashion or the other. These two ways comprise the major branches of expectation states theory, the former focusing on the effects of performance evaluations and the latter on the effects of other status information. These branches and their development are described below. (Because of space limitations, the emphasis in this article is on the basic ideas, and specific references can only be given for a sample of the theoretical and empirical work done in the program. This sample consists of key earlier pieces and a selection of the more recent research. It should be noted, however, that expectation states theory has been a collaborative effort from its beginnings and that a considerable number of researchers have worked in it. For references to their individual contributions at the various stages of the program, see the reviews and edited

collections, as well as the examples of research, mentioned at the end of this article.)

Initial work in the first branch of the theory investigated the formation of performance expectations in situations where two persons begin their interaction as status equals (i.e., they have no information from either inside or outside the group that would enable them to assign different levels of task competence to each other). In this case, the assignment of levels of competence that eventually occurs is the result of a generalization from evaluations of units of performance. These evaluations may be made by the performers themselves as they resolve disagreements between them regarding the correct solution to the task (Berger and Conner 1969; Berger and Snell 1961) or by a "source" or third party with the right to evaluate the performers (Crundall and Foddy 1981; Webster 1969; Webster and Sobieszek 1974). Moreover, the evaluations may or may not be made through the use of objective criteria. Various other topics were later investigated within this branch. These include the effects on expectations of the type of evaluations received (e.g., either positive or negative; either highly or moderately positive), the consistency of these evaluations (e.g., across performances or across sources), and the types of attributes that confer evaluative competence on a source (such as level of task competence or vicarious exposure to the task). Other recent work has dealt with the effects of applying either strict or lenient standards for competence (as well as for lack of competence) to the processing of evaluations. The study of second-order expectations (those based on self's perceptions of the expectations held by the partner) also has received attention in this program, and is of direct relevance to this branch.

The second branch of the theory is primarily concerned with situations where the actors differ in status. It is in this branch, which has come to be known as "status characteristics theory," that a large proportion of expectation states research has been conducted. This work started with the investigation of situations in which two persons are differentiated with respect to a single diffuse status characteristic (Berger et al. [1966] 1972; Moore 1968). Furthermore, this difference constitutes the only information that the focal actor has about self and other, and he or she treats the attribute as a diffuse status characteristic (i.e., attaches different evaluations to its various levels

and holds different performance expectations, specific as well as general, for each level).

The basic proposition in this branch states that unless the focal actor believes the characteristic to be irrelevant to the task at hand, he or she will use the status difference between self and other to organize their interaction. In other words, performance expectations for the task will be the result of importing status information from outside the group. This process is known as "status generalization." Let us consider gender as an example (with "man" carrying a more positive valuation than "woman") and assume that the focal actor is a man and that his partner is a woman. The theory predicts different expectations, depending on what the individual believes the sex linkage of the task to be. If the task is perceived as masculine (one at which, in general, men do better than women), self will use this information to infer that he is the better of the two. If he has no information about its sex linkage (perceives the task as neither associated with nor dissociated from sex differences), the "burden-of-proof principle" will apply: those of lower status will be considered to have inferior task competence, unless they demonstrate the opposite. Accordingly, also in this situation, the man will expect to be more competent at the task than his female partner. On the other hand, the theory predicts that he will *not* form expectations of his own superiority if the task has been explicitly dissociated from gender. Finally, if the valued task is perceived to be one at which women generally excel, the prediction is that the male actor will consider himself inferior to his female partner.

It is important to emphasize that the above predictions assume that gender is a diffuse status characteristic for the focal actor (and that, in this example, "man" carries a higher value than "woman"). If gender is not such a characteristic, the propositions do not apply. Because of the inclusion of this scope condition, the theory is able to incorporate cultural, historical, and individual differences. Thus, through this condition, it incorporates the fact that although gender is a diffuse status characteristic in most societies, it is more so in some than in others. The extent to which this is the case also may vary within a given society and from one historical period to another. Furthermore, the theory reflects the fact that even in

strongly sexist societies there are likely to be individuals who are less sexist than the majority, and perhaps some who are not sexist at all. (This may be due, for example, to variations in socialization practices within a given culture, and/or to the different degrees of success of these practices). Note that the theory assumes that women, *as well as* men, may treat gender as a diffuse status characteristic and devalue women's performances. (For a theoretical account of the social construction of status beliefs, see Ridgeway 1991).

Let us then reanalyze the earlier example from the point of view of the female actor, and assume that she considers gender to be a diffuse status characteristic to the same degree as her male partner does. The theory predicts that she will then form expectations that are exactly complementary to those of her partner. Thus, if she either views the task as masculine or relates it to sex differences through a burden-of-proof process, she will consider herself the less competent of the two performers. (She will not form such expectations if she sees the task as dissociated from gender, and will believe herself to be the better performer if she views the task as feminine.) However, as with most status systems, those persons who benefit from the existing order can be expected to be more supportive of it than those who do not, socialization practices notwithstanding. Thus it could be that men tend to meet this scope condition more than women do. This is not an issue about the theory itself but an empirical matter regarding the particular instances to which it applies. Note, also, that while the preceding discussion has been limited to gender for the sake of providing an example, the comments are meant to apply to any diffuse status characteristic. In general, the strength of the link between status and expectations is a function of the extent to which the attribute is a status characteristic for the focal actor.

Performance expectations formed through status generalization may, of course, be affected by other factors. Accordingly, the basic ideas of the initial formulation of status characteristics theory have been expanded and elaborated in several directions. In particular, these include the study of situations where actors possess more than one status characteristic. These may be either specific or diffuse, or either consistent or inconsistent with each other in terms of the levels of competence

they imply. Furthermore, as a set, they may either equate or differentiate the performers to various degrees (Berger et al. 1977). This extended formulation specifies the conditions under which the available status items will become salient (or "activated"), as well as the rules by which they will be processed. It is proposed that all (and only) activated items will be used and that they will be combined according to the following principles: inconsistent information will have more impact than consistent information, and each additional item of consistent information will have less weight than it would by itself. Furthermore, this version includes referent actors (nonparticipating actors who serve as objects of comparison for the performers) and their role in the formation of expectations.

The two main branches of the program have not developed independently of each other. In fact, the branches touch and intertwine at various points, as expectations are formed on the basis of *both* performance evaluation and status information. For example, the extent to which the focal actor accepts the evaluations from a source is often affected by the diffuse status characteristics (and their levels) of everyone involved in the situation. Of particular interest is the resilience of status-based expectations when these are contradicted by actual performance evaluations. This resistance shows itself in, for example, the higher level of performance that lower-status actors have to achieve in order to escape the effects of status generalization. The inconsistency between performance evaluation and status information is of course a special case of the extension of status characteristics theory to more than one attribute. A recent theoretical proposal in this area links the resilience to the use of double (or even multiple) standards for both competence and incompetence that protect higher-status actors and penalize lower-status ones.

Other expectation states research, also of relevance to both branches, has included a rich variety of topics. For example, theoretical and empirical work has investigated the relationship between performance expectations and reward expectations, the status-value view of distributive justice, the transfer of expectations across actors and situations, and the legitimation of status positions. The latter work examines how and why individuals come to view the status hierarchies in which they



are involved as right and proper. Also of special interest is the analysis of the way in which task settings generate emotional reactions and how these, in turn, affect performance expectations. Other important areas have been: The extension of status characteristics theory to include self-fulfilling effects of status on performance; the study of the role that status cues (such as accent, dress, demeanor) have in the assignment of task competence; and the work on how the attribution of personality characteristics (such as friendly, rigid, outgoing) and moral characteristics (such as honest, fair, selfish) affects this assignment and is, in turn, affected by it.

No review of expectation states theory would be complete without a discussion of three features that, together with those mentioned earlier, serve to characterize this program. *First*, a sizable portion of the research has used a standardized experimental situation involving two subjects (Webster and Sobieszek 1974, Appendices 1 and 2). The utilization of an experimental approach clearly is well suited to the program's analytical strategy of investigating the effects of a few variables at a time, and the standardized setting has contributed to cumulativeness in the research findings by allowing comparability across studies. In this setting, the task consists of a series of trials, each involving a binary choice regarding visual stimuli, with subjects allowed to interact only indirectly (as opposed to face-to-face) with each other. The situation consists of two phases. In the first phase, performance expectations are formed on the basis of information controlled by the experimenter. This may include evaluations of each person's task performance, cues regarding the participants' other status characteristics, or a combination of the two. In the second phase, each trial consists of an individual initial choice, the experimentally controlled communication of the partner's choice, and a final individual choice. In most trials, the communication indicates that the partner disagrees with the subject. The dependent variable of central interest is the amount of influence (one of the components of the power and prestige order of the group) that a partner exerts on the subject in the resolution of those disagreements. Several versions of this setting, including computerized ones, have now been developed. (Additional bases of empirical support for the theory include direct evidence from research conducted in discussion

groups, classrooms, and various other settings—such as those mentioned below—as well as indirect evidence from work originating in several other theoretical traditions. These studies corroborate and extend the knowledge gained through the use of the standardized setting. For a review of the different types of support for the theory, see Berger et al. 1989).

*Second*, the program has been characterized by a strategy of stating propositions as part of deductive systems. These systems have been formalized to various extents, and different techniques (such as stochastic models, a Bayesian approach, graph theory, fuzzy set theory) have been used. The most comprehensive of these formulations is the graph theoretic model of the extended version of status characteristic theory (Berger, Fisek, Norman, and Zelditch 1977). This model has also been used to formalize other areas of the program, such as the formation of expectations in groups where members initially are status equals, and the legitimation and delegitimation of status hierarchies.

*Third*, from its beginnings, researchers in the program have shown a definite interest in applications and interventions designed to understand and alleviate specific social problems. A large portion of this work has included the development of techniques to reduce the effects of status generalization, particularly those based on skin color and ethnic background, in classroom settings (Cohen 1972, 1982; Cohen and Lotan 1997; Entwisle and Webster 1972). It is important to note that this research has progressed in close association with theoretical work; thus applications and interventions have both stimulated and benefited from the development of abstract formulations.

Expectation states theory continues to generate strong interest among researchers in group processes and related areas, and to develop in several directions. For example, the refinement, expansion, and integration of various aspects of the program are currently under way. There are also some exciting new areas of research, such as the study of the relationship between status and various other factors in the assignment of task competence, namely affect (emotions and sentiments), formal authority, power, dominance, and social identity. Furthermore, the notion of an "expectation state" has been generalized from its

original reference to status-based processes, to incorporate those concerning affect and control. Finally, the research methodology has been extended to comprise a wide range of settings and tasks, including face-to-face contexts, larger groups, interactions where selected aspects of formal organizations have been re-created, vignette studies, computerized tasks, and settings where actors are not required to be collectively oriented.

For reviews and assessments of the expectation states program at various stages, see Berger (1988); Berger, Rosenholtz, and Zelditch (1980); Berger, Wagner, and Zelditch (1985, 1989); Meeker (1981); Ridgeway and Walker (1995); Wagner and Berger (1993); Webster and Foschi (1988a). For edited collections of work on various topics, see Berger, Conner, and Fisek (1974); Berger and Zelditch (1985; 1998b); Szmataka, Skvoretz, and Berger (1997); Webster and Foschi (1988b). For examples of the more recent research, see Balkwell, Berger, Webster, Nelson-Kilger, and Cashen (1992); Cohen and Zhou (1991); Foschi (1996); Foschi, Lai, and Sigerson (1994); Gerber (1996); Lovaglia and Houser (1996); Lovaglia, Lucas, Houser, Thye, and Markovsky (1998); Riches and Foddy (1989); Ridgeway (1989); Ridgeway, Boyle, Kuipers, and Robinson (1998); Ridgeway and Johnson (1990); Shelly and Webster (1997) Stewart and Moore (1992); Troyer and Younts (1997); Wagner, Ford, and Ford (1986); Webster and Whitmeyer (1999). For discussions of the program's methodological and metatheoretical assumptions, see Berger and Zelditch (1998a); Berger, Zelditch, and Anderson (1972); Cohen (1989, esp. chap. 6); Zelditch (1969).

REFERENCES

Balkwell, James W., Joseph Berger, Murray Webster, Jr., Max Nelson-Kilger, and Jacqueline Cashen 1992 "Processing Status Information." *Advances in Group Processes* 9:1-20.

Berger, Joseph 1958 "Relations Between Performance, Rewards, and Action-Opportunities in Small Groups." Ph.D. diss., Harvard University.

—1988 "Directions in Expectation States Research." In Murray Webster, Jr., and Martha Foschi, eds., *Status Generalization: New Theory and Research*. Stanford, Calif.: Stanford University Press.

—, Bernard P. Cohen, and Morris Zelditch, Jr. 1966 "Status Characteristics and Expectation States." In

Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress*, vol. 1. Boston: Houghton Mifflin.

—, Bernard P. Cohen, and Morris Zelditch, Jr. 1972 "Status Characteristics and Social Interaction." *American Sociological Review* 37:241-255.

—, and Thomas L. Conner 1969 "Performance Expectations and Behavior in Small Groups." *Acta Sociologica* 12:186-198.

—, Thomas L. Conner, and M. Hamit Fisek, eds., 1974 *Expectation States Theory: A Theoretical Research Program*. Cambridge, Mass.: Winthrop.

—, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch, Jr. 1977 *Status Characteristics and Social Interaction: An Expectation-States Approach*. New York: Elsevier.

—, Susan J. Rosenholtz, and Morris Zelditch, Jr. 1980 "Status Organizing Processes." In Alex Inkeles, Neil J. Smelser, and Ralph H. Turner, eds., *Annual Review of Sociology*, vol. 6. Palo Alto, Calif.: Annual Reviews.

—, and J. Laurie Snell 1961 "A Stochastic Theory for Self-Other Expectations." Technical Report no. 1. Stanford, Calif.: Laboratory for Social Research, Stanford University.

—, David G. Wagner, and Morris Zelditch, Jr. 1985 "Expectation States Theory: Review and Assessment." In Joseph Berger and Morris Zelditch, Jr., eds., *Status, Rewards and Influence: How Expectations Organize Behavior*. San Francisco: Jossey-Bass.

—, David G. Wagner, and Morris Zelditch, Jr. 1989 "Theory Growth, Social Processes, and Metatheory." In Jonathan H. Turner, ed., *Theory Building in Sociology: Assessing Theoretical Cumulation*. Newbury Park, Calif.: Sage.

—, and Morris Zelditch, Jr., eds., 1985 *Status, Rewards, and Influence: How Expectations Organize Behavior*. San Francisco: Jossey-Bass.

—, and Morris Zelditch, Jr. 1998a "Strategies of Theory Construction." In Joseph Berger and Morris Zelditch, eds., *Status, Power and Legitimacy: Strategies and Theories*. New Brunswick, N.J.: Transaction.

—, and Morris Zelditch, Jr., eds., 1998b *Status, Power and Legitimacy: Strategies and Theories*. New Brunswick, N.J.: Transaction.

—, Morris Zelditch, Jr., and Bo Anderson 1972 "Introduction." In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.

Cohen, Bernard P. 1989 *Developing Sociological Knowledge: Theory and Method*, 2nd ed. Chicago: Nelson-Hall.

- , and Xueguang Zhou 1991 “Status Processes in Enduring Work Groups.” *American Sociological Review* 56:179–188.
- Cohen, Elizabeth G. 1972 “Interracial Interaction Disability.” *Human Relations* 37:648–655.
- 1982 “Expectation States and Interracial Interaction in School Settings.” In Ralph H. Turner and James F. Short, Jr., *Annual Review of Sociology*, vol. 8. Palo Alto, Calif.: Annual Reviews.
- , and Rachel A. Lotan, eds., 1997 *Working for Equity in Heterogeneous Classrooms: Sociological Theory in Practice*. New York: Teachers College Press.
- Crundall, Ian, and Margaret Foddy 1981 “Vicarious Exposure to a Task as a Basis of Evaluative Competence.” *Social Psychology Quarterly* 44:331–338.
- Entwisle, Doris R., and M. Webster, Jr. 1972 “Raising Children’s Performance Expectations: A Classroom Demonstration.” *Social Science Research* 1:147–158.
- Foschi, Martha 1996 “Double Standards in the Evaluation of Men and Women.” *Social Psychology Quarterly* 59:237–254.
- , Larissa Lai, and Kirsten Sigerson 1994 “Gender and Double Standards in the Assessment of Job Applicants.” *Social Psychology Quarterly* 57:326–339.
- Gerber, Gwendolyn L. 1996 “Status in Same-Gender and Mixed-Gender Police Dyads: Effects on Personality Attributions.” *Social Psychology Quarterly* 59:350–363.
- Lovaglia, Michael J., and Jeffrey A. Houser 1996 “Emotional Reactions and Status in Groups.” *American Sociological Review* 61:867–883.
- , Jeffrey W. Lucas, Jeffrey A. Houser, Shane R. Thye, and Barry Markovsky 1998 “Status Processes and Mental Ability Test Scores.” *American Journal of Sociology* 104:195–228.
- Meeker, Barbara F. 1981 “Expectation States and Interpersonal Behavior.” In Morris Rosenberg and Ralph H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Moore, James C., Jr. 1968 “Status and Influence in Small Group Interactions.” *Sociometry* 31:47–63.
- Riches, Phoebe, and Margaret Foddy 1989 “Ethic Accent as a Status Cue.” *Social Psychology Quarterly* 52:197–206.
- Ridgeway, Cecilia L. 1989 “Understanding Legitimation in Informal Status Orders.” In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress: New Formulations*. Newbury Park, Calif.: Sage.
- 1991 “The Social Construction of Status Value: Gender and Other Nominal Characteristics.” *Social Forces* 70:367–386.
- , and Cathryn Johnson 1990 “What Is the Relationship Between Socioemotional Behavior and Status in Task Groups?” *American Journal of Sociology* 95:1189–1212.
- , Elizabeth H. Boyle, Kathy J. Kuipers, and Dawn T. Robinson 1998 “How Do Status Beliefs Develop? The Role of Resources and Interactional Experience.” *American Sociological Review* 63:331–350.
- , and Henry A. Walker 1995 “Status Structures.” In Karen S. Cook, Gary A. Fine, and James S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Shelly, Robert K., and Murray Webster, Jr. 1997 “How Formal Status, Liking, and Ability Status Structure Interaction: Three Theoretical Principles and a Test.” *Sociological Perspectives* 40:81–107.
- Stewart, Penni A., and James C. Moore, Jr. 1992 “Wage Disparities and Performance Expectations.” *Social Psychology Quarterly* 55:78–85.
- Szmatka, Jacek, John Skvoretz, and Joseph Berger, eds., 1997 *Status, Network, and Structure: Theory Development in Group Processes*. Stanford, Calif.: Stanford University Press.
- Troyer, Lisa, and C. Wesley Younts 1997 “Whose Expectations Matter? The Relative Power of First- and Second-Order Expectations in Determining Social Influence.” *American Journal of Sociology* 103:692–732.
- Wagner, David G., and J. Berger 1993 “Status Characteristics Theory: The Growth of a Program.” In Joseph Berger and Morris Zelditch, Jr., eds., *Theoretical Research Programs: Studies in the Growth of Theory*. Stanford, Calif.: Stanford University Press.
- , Rebecca S. Ford, and Thomas W. Ford 1986 “Can Gender Inequalities Be Reduced?” *American Sociological Review* 51:47–61.
- Webster, Murray, Jr. 1969 “Source of Evaluations and Expectations for Performance.” *Sociometry* 32:243–258.
- , and Martha Foschi 1988a “Overview of Status Generalization.” In Murray Webster, Jr. and Martha Foschi, eds., *Status Generalization: New Theory and Research*. Stanford, Calif.: Stanford University Press.
- , and Martha Foschi, eds., 1988b *Status Generalization: New Theory and Research*. Stanford, Calif.: Stanford University Press.
- , and Barbara Sobieszek 1974 *Sources of Self-Evaluation: A Formal Theory of Significant Others and Social Influence*. New York: Wiley.
- , and Joseph M. Whitmeyer 1999 “A Theory of Second-Order Expectations and Behavior.” *Social Psychology Quarterly* 62:17–31.

Zelditch, Morris, Jr. 1969 "Can You Really Study an Army in the Laboratory?" In Amitai Etzioni, ed., *A Sociological Reader on Complex Organizations*, 2nd ed. New York: Holt, Rinehart and Winston.

MARTHA FOSCHI

## EXPERIMENTS

Sociologists usually reserve the term *experiment* for studies in which the researcher manipulates one or more key independent variables; that is, for studies in which the experimenter controls the decision as to which subjects are exposed to what level of the independent variable. Although experiments comprise a minority of sociological research studies, they are still quite common. Laboratory versions are traditional in much of sociological social psychology (examples are Bonacich 1990; Molm 1990). Field experiments are prominent in program evaluations, particularly in applied areas such as education (e.g., Slavin and Karweit, 1985), criminal justice (e.g. Rossi et al. 1980; see also Zeisel, 1982), and even large-scale social policy efforts such as studies of a proposed guaranteed minimum income (Hannon et al. 1977). Questionnaire experiments (such as Schuman and Presser 1981) test various methodological questions, while "vignette" experiments (such as Rossi and Nock 1982) have become popular approaches to the study of attitudes.

The primary attraction of the experimental method is undoubtedly that it is more persuasive than other methods in its fit with causal arguments (Kish 1987). Campbell and Stanley, whose 1963 book is the most influential discussion of experiments for contemporary sociologists, call this "internal validity" (for other historically important discussions of experiments, see especially Cochran and Cox 1957; Fisher 1935). To them, "true" experiments are studies in which subjects are randomly allocated into "experimental" and "control" groups. The former receive a "treatment," such as an educational program, while the latter do not. Randomization allows the researcher to assume the similarity of groups at the beginning of the treatment, with known statistical chances of error, and to avoid a variety of "threats" to the validity of the conclusion that the treatment "caused" any found differences between the groups in post-treatment behavior or other outcome. In

contrast, "quasi-experiments" (most other forms of research), which do not use randomization, require a variety of additional, often heroic, assumptions in order to make causal inferences (see Lieberman 1985 for an argument that these assumptions may be generally unacceptable).

## A MODEL

Consider an experiment to test the proposition that increasing the sensitivity to Hispanic cultures of non-Hispanic second-grade teachers will improve the performance of Hispanic children in their classes, particularly students with poor English-language skills. Teachers are randomly assigned to attend one of three seminars on Hispanic cultures that vary in intensity (experimental treatments) or a seminar on Asian culture (control). Following the approach of Alwin and Tessler (1974), Figure 1 presents a model of this experiment. Our proposition is tested by estimating the relationship between two "unobservable" variables—student performance in second grade ( $Y$ ) and the sensitivity of their teachers ( $X$ ). The variables are "unobservable" in the sense that we cannot unambiguously and directly measure the rather complex concepts we have. For example, true "student performance" probably includes facility with mathematics, language acquisition, and many other components, each of which we can measure only imperfectly. The model also contains two other variables that may be unobservable—the "Hispanic intensity" of the course ( $T$ ), and the English-language skill of the student ( $L$ )—and might include a fifth unobservable variable (as indicated by the broken line), which is the student's performance in first grade ( $P$ ). Including  $P$  would allow us to test for the effect of taking the seminar on *change* in student performance.

As is conventional in such models, each unobservable variable is measured by indicators ( $x_1, x_2, t_1$ , etc.) and is affected by a stochastic term ( $e_x, e_T$ , etc.) representing our inability to measure the concept perfectly with our indicators as well. It is obvious that student performance ( $Y$  and  $P$ ) would be best measured by a series of indicators such as teacher's ratings, grades in various topics, and performances on standardized tests. Since similar or identical measures of  $Y$  and  $P$  might be used, errors could be correlated, as shown by the curved line between  $e_{y2}$  and  $e_{p2}$ , thereby biasing

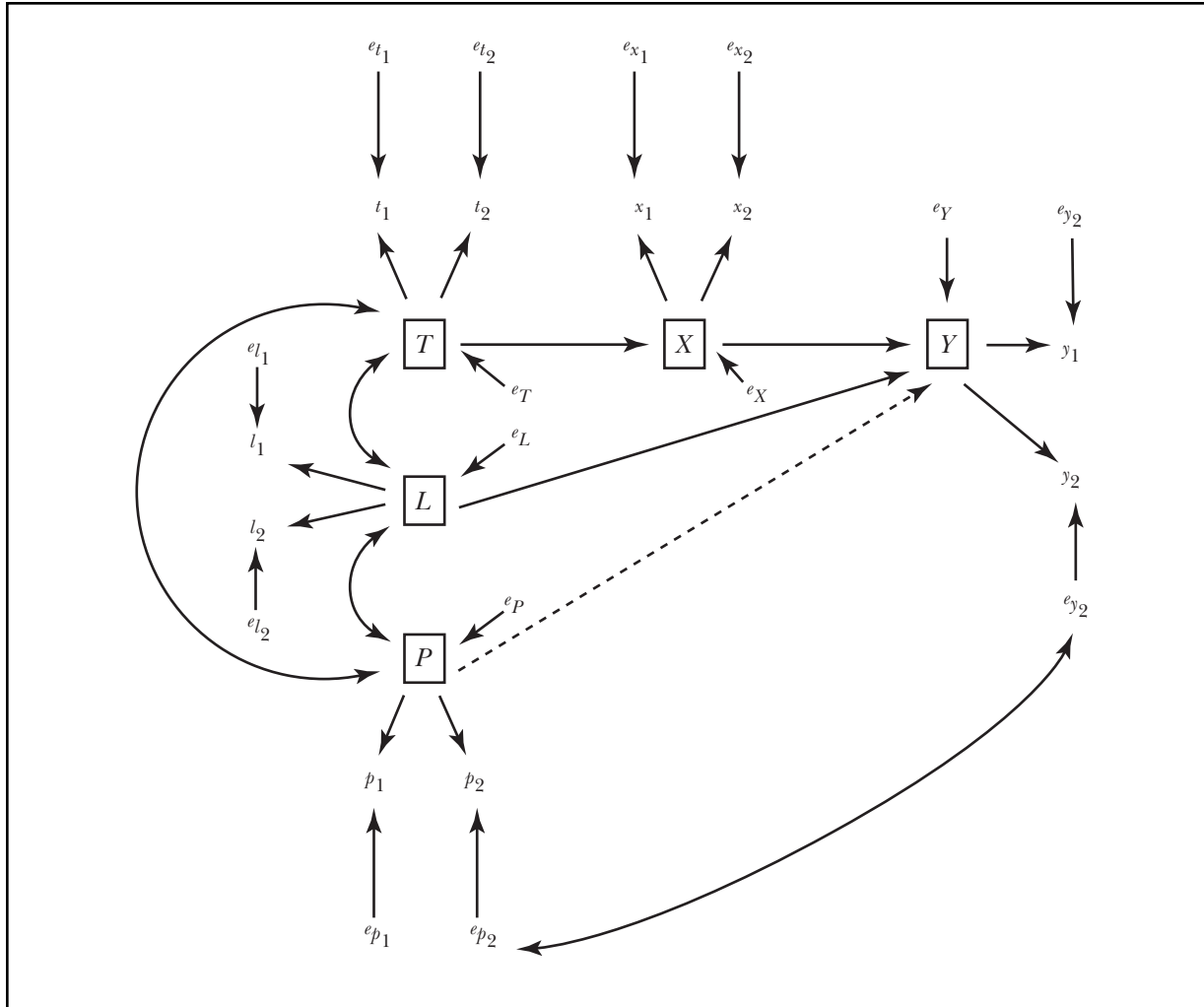


Figure 1

NOTE: Model based on Alwin and Tessler

results. This is why Campbell and Stanley (1963) argue in favor of generally using “post-test only” designs.

Alwin and Tessler call  $x_1$  and  $x_2$  “manipulation checks.” Hispanic culture seminars ( $T$ ) may or may not change a given teacher’s sensitivity toward Hispanic students ( $X$ ). It is even possible that Asian culture seminars make teachers more sensitive to all minorities, including Hispanics. Heckman and Smith (1995) also point out that subjects who do not get the “treatment” in field experiments can sometimes find “substitutes.” For example, the teachers assigned to Asian culture seminars might feel they also need sensitivity to Hispanic

culture and take courses in the subject voluntarily. All these possibilities point out that if we simply assume that  $T$  successfully manipulated  $X$ , we might be making a grave error. There may actually be little difference in Hispanic sensitivity ( $X$ ) between experimentals and controls. We will have to check, which means that we will have to measure  $X$ .

### VARIATIONS IN TREATMENT

Experimenters rarely realize that they often need to measure  $T$  as well as  $Y$ , even if  $T$  is “under their control.” If the treatment is complicated, assigning a subject to be treated does not guarantee that

he or she receives the treatment that is intended. The instructor in one of the culture seminars might do a poor job of presenting the materials. A teacher might be ill and miss one of the sessions. Most experimenters assume that the relationship between  $T$  and  $X$  is 1.0, but that is frequently, perhaps mostly, not the case. To estimate the actual effect of  $T$  on  $X$ , various “treatment checks” ( $t_1, t_2$ ) would be required.

Note that a sophisticated experimental design may often require many levels of the “treatment” and even of the “control.” Or there may be no “control” at all, just comparison among different treatments. If  $X$  is conceptualized as a continuous variable, such as the intensity of the seminar (where  $T$  may be number of sessions), there is no reason to have only three levels of  $T$ . Instead, we might randomly assign each member of the treatment group to a different number of sessions from one to a whole semester (if the seminar was being offered at the university anyway). This is particularly important in something like a guaranteed income experiment, where a wide range of levels of guaranteed income are possible. We might even wish to have every family receive a different level of income support, so as to well represent the possible range and thus help us choose the best level of support for an actual program.

Furthermore, note that in our example the control group is not comprised of teachers who get no seminars at all. This is to avoid what has been commonly termed the “Hawthorne effect,” in which experimental subjects respond to the attention they get by being sent to a seminar rather than to the content (Hispanic culture) of the seminar itself. Designs that do not attend to the effects of just being treated are always subject to errors of interpretation. Of course, an even more sophisticated design might include a second control group of teachers who were not even informed that others were attending seminars (and perhaps other control groups at various levels of “intensity” of the Asian culture seminar, as well). The “no seminar” group would provide data that allow us to estimate the size of the difference in class performance between students whose teachers had attended sensitivity seminars and students in a school system with no sensitivity program at all.

The Hawthorne effect is actually only one example of a general class of problems that might

arise in interpreting found relationships (or nonrelationships) between  $T$  and  $Y$ . Smith (1990) argues that even if  $T$  does manipulate  $X$  as designed, it might also affect some additional unmeasured and perhaps untheorized variable ( $A$ ) that in turn has an important effect on  $Y$ . This is particularly a problem for “real-world” experiments. For example, teachers who take the Hispanic culture seminars might be treated with more respect by their Hispanic principal, who feels closer to them. These teachers may have no greater sensitivity, but may work harder and teach better in response to the principal’s behavior. If  $A$  is not in the causal model, the effects of  $A$  on  $Y$  will be misattributed to  $X$ . Thus, some of the difficulties with nonexperimental research designs can affect complex experimental research as well.

### BLOCKING

In most studies, a number of variables other than  $X$  may seem to the researcher to be probable causes of  $Y$ . Unlike  $A$ , above, these variables are both theorized and measurable. However, these variables cannot be manipulated, either because such manipulation is impossible or because it is otherwise undesirable. For example, English-language facility ( $L$ ) would probably affect the performance of most second-grade students in the United States. The experimenter cannot manipulate this characteristic of the subjects. Because of randomization we expect  $T$ , and hence  $X$ , to be uncorrelated with  $L$  (see Figure 1) and with all other variables. However, sampling variation means that we might be unlucky, so that  $T$  and  $L$  are empirically correlated among the students we are actually studying. Under these conditions our estimate of the relationship between  $X$  and  $Y$  may be unbiased, but it is also less efficient than when  $L$  is controlled, because we add sampling error to other sources of error.

We can increase the efficiency of the estimate by using  $L$  as a “blocking variable,” perhaps by using their ranks on  $L$  to create matched pairs of students. We then randomly assign one of the two highest-rated students to the experimental group and the other to the controls, one of the two second-highest-rated students to the experimental group and the other to the controls, and so forth, until the lowest  $L$  pair is assigned. We thus guarantee a correlation of 0 between  $T$  and  $L$ . In our

analysis, we can statistically estimate the effects of blocked variables, such as  $L$  on  $Y$ , and reduce the variance that we attribute to “error.”

### COST

Efficiency is particularly important because experiments are often very expensive. When “real life” manipulates the independent variables—for example making one person male and the other female, one born to rich parents another to poor—the researcher does not have to pay for the “treatment.” But experiments usually require the researcher to get the subject to do something, which can often be done only through pay, and then to do something to the subject, which also usually costs money. In our example, the teachers will have to be paid for their time spent going to the seminars, and so will the instructors of those seminars. These costs generally mean that experimenters very much need to keep the number of subjects they study small, which threatens the internal (and external, see below) validity of the research.

### GENERALIZATION

The most frequent criticism of experiments is that their results are often difficult to generalize, a problem Campbell and Stanley (1963) call “external invalidity.” Because treatments tend to require special arrangements, the sample studied is usually restricted to some relatively “special” group, which is therefore unrepresentative (Kish 1987) of the population to whom we might like the results to apply. For example, getting students from more than one or two school systems to participate in our study is hard to imagine. Certainly a random sample from all the Hispanic students in the whole country is impossible. Yet the school system to which we have access, and within which we do the study, might be very atypical. For example, it may be a suburb whose Hispanic community is comparatively well-off. A strong relationship between  $X$  and  $Y$  might exist in a poorer community, but our study might estimate the relationship as 0. It is possible to do representative samples in experiments (see Marwell and Ames 1979), particularly those involving such things as question wording (e.g., Schuman and Presser 1981). However, such opportunities are rare.

Experiments also usually prescribe strong “control” over various extraneous conditions. To assure that they all get the identical treatment, the teachers might be brought to a tightly run seminar at the university. However, if the program were actually implemented as a regular policy by a number of school systems, many of the seminars would be taught at the schools, at various times, by various leaders, and so forth. Can one generalize from the experiment to everyday practice? As Kish (1987) points out, experiments are often necessarily weak on the *reality* of the conditions they impose on subjects.

Campbell and Stanley spend much of their time discussing the possibility that the very process of measuring the variables of interest may make the experiment unrepresentative of other situations. For example, if we measure performance in both the second ( $Y$ ) and first ( $P$ ) grades with the same standardized test, some students might have had their learning during the second grade affected by taking the test in the first grade. They might have become sensitized to the material asked in the early test and “learned” it more than would students in the “real world” who had not been tested. Again, the issue is whether what appear to be the effects of  $X$  on  $Y$  might not be the effects of some interaction between  $X$  and  $P$ —and therefore not be generalizable to the usual situation in the school. This is one reason Campbell and Stanley recommend using “post-test only” designs (i.e., without  $P$ ) where feasible, relying on randomization to make groups equivalent.

In fact, interactions with nonmanipulated independent variables are a particular problem for experiments. For example, the sensitivity program may work only with Hispanic students who are at least second-generation Americans. First-generation Hispanics have such language problems that “sensitivity” matters much less than other skills. If our community contains only first-generation students, we may falsely assume that sensitivity training was generally worthless (Smith 1990).

### CONCLUSIONS

Although recognizing the advantages of experimental studies for making causal inferences, sociologists have tended to associate experiments with

laboratory-based studies and to denigrate the possibility of generalizing from their results. A broader view of experiments sees them as serving a variety of purposes in a variety of settings, while retaining the association with causal models that makes them attractive.

One of the less recognized of these advantages is that experiments tend to force the researcher to clarify the causal model underlying his or her argument, since such theoretical clarity is necessary for the development of an effective design.

The future of experiments in sociology seems particularly uncertain at this time. On the one hand, the growth of observational and historical studies, along with the continuing influence of econometric methods and large data sets, all appear to be reserving for experiments a smaller and smaller part of the research agenda. The study of group structure and group behavior, perhaps the sociological area most suffused with experimental research, appears to be shifting out of sociological centers and journals and into business schools and publications. Technically skilled economists and sociologists such as Heckman and Smith have been rethinking the problems of experiments and have been making clearer their shortcomings, even for causal assertions.

On the other hand, work on sociological topics done in business schools, sometimes by trained sociologists, need not be interpreted as a decline in interest or activity by the field. Within sociology there has been a growth of attention to social networks, and some of the work in this area has used experiments to good effect.

The reconsideration of experiments by economists has followed a surge of interest in experimental methods by a growing number of economists, with the institutionalization of a journal called *Experimental Economics*, a book with same title (Davis and Holt 1993), and even a *Handbook of Experimental Economics* (Kagel and Roth 1991). The growth of interest by economists is largely (although not entirely) driven by the increasing willingness of government and other large funders to undertake experimental evaluations of social programs in areas such as welfare reform, education, and criminal justice. Sociologists have played a continuing and central role in these efforts, and the end of such investments does not seem in sight.

## REFERENCES

- Alwin, Duane F., and Richard C. Tessler 1974 "Causal Models, Unobservable Variables, and Experimental Data." *American Journal of Sociology* 80:58-86.
- Bonacich, Philip 1990 "Communication Dilemmas in Social Networks: An Experimental Study." *American Sociological Review* 55:427-447.
- Campbell, Donald T., and Julian C. Stanley 1963 *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNalley.
- Cochran, W. G., and G. Cox 1957 *Experimental Designs*. New York: John Wiley.
- Davis, Douglas D., and Charles A. Holt 1993 *Experimental Economics*. Princeton, N.J.: Princeton University Press.
- Fisher, R. A. 1935 *The Design of Experiments*. London: Oliver and Boyd.
- Hannan, Michael, Nancy Brandon Tuma, and Lyle Groenveld 1977 "Income and Marital Events." *American Journal of Sociology* 82:1186-1211.
- Heckman, James J., and Jeffrey A. Smith 1995 "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9:85-110.
- Kagel, J., and A. Roth, eds. 1991 *Handbook of Experimental Economics*. Princeton, N.J.: Princeton University Press.
- Kish, Leslie 1987 *Statistical Design for Research*. New York: John Wiley.
- Lieberson, Stanley 1985 *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California Press.
- Marwell, Gerald, and Ruth Ames 1979 "Experiments on the Provision of Public Goods, I: Resources, Interest, Group Size and the Free Rider Problem." *American Journal of Sociology* 84:1335-1360.
- Molm, Linda 1990 "Structure, Action and Outcomes: The Dynamics of Power in Social Exchange." *American Sociological Review* 55:427-447.
- Rossi, Peter H., Richard A. Berk, and Kenneth J. Lenihan 1980 *Money, Work, and Crime: Experimental Evidence*. New York: Academic Press.
- Rossi, Peter H., and Steven L. Nock 1982 *Measuring Social Judgments: The Factorial Survey Approach*. Beverly Hills, Calif.: Sage.
- Schuman, Howard, and Stanley Presser 1981 *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. Orlando, Fla.: Academic Press.



- Slavin, Robert E., and Nancy C. Karweit 1985 "Effects of Whole Class, Ability Grouped and Individual Instruction on Mathematics Achievement." *American Educational Review Journal* 22:351-367.
- Smith, Herbert L. 1990 "Specification Problems in Experimental and Nonexperimental Social Research." *Sociological Methodology* 20:59-92.
- Zeisel, Hans 1982 "Disagreement over the Evaluation of a Controlled Experiment." *American Journal of Sociology* 88:378-389.

GERALD MARWELL

## **EXTREME INFLUENCE – THOUGHT REFORM, HIGH CONTROL GROUPS, INTERROGATION AND RECOVERED MEMORY PSYCHOTHERAPY**

That social influence is ever present and pervasive is one of the most fundamental observations that may be made about social life. While we are both guided and constrained by the influence factors present at every moment in our lives, we typically fail to fully appreciate how much of an effect cultural, community, and interpersonal influences have on our values, beliefs, and choices. This is in part due to the fact that we rarely find ourselves in situations in which essentially all of the influence forces to which we are exposed are strongly organized and directed in support of a particular ideology, perception, and/or set of actions—rather we usually find ourselves in social situations in which the factors pressuring us to act in one way are offset somewhat by opposing pressures directing us toward different value positions, understandings of the world, and/or different actions. The variability and diversity of the influences to which we are exposed allow us to maintain a sense that we are the source of our choices and actions and contribute to our sense of personal autonomy, freedom of choice, and individuality.

The substantial power of community and interpersonal influence to shape perceptions and actions can most clearly be appreciated through the study of social environments in which influence factors are hyper-organized rather than relatively loosely knit together. It is through the study

of influence environments that have been consciously designed to elicit conformity, promote radical change in a person's values and beliefs, and influence an individual's choices that it is possible to gauge the extent to which our sense of personal autonomy and individuality is rooted in social organization. Four examples of environments constructed for the purpose of inducing radical and dramatic shifts in major components of persons' values or perceptions, or both, or environments designed to induce a shift on a single, but nevertheless important decision will be briefly reviewed herein—programs of thought reform or coercive persuasion (such as programs of political reeducation attempted in China under Mao); programs of indoctrination carried out by ideologically focused high control groups (groups labeled as cults and organizations that market experiences that are alleged to psychologically transform an individual); modern police interrogation tactics that, through psychological means, elicit false confessions from innocent suspects; and quack psychotherapy treatments in which patients are led to mislabel hypnotic fantasies, dreams, and hunches as "recovered memories" of lengthy histories of sexual abuse that were entirely unknown to them prior to entering treatment.

### **REEDUCATION PROGRAMS**

*Coercive persuasion* and *thought reform* are alternate names for programs of social influence capable of producing substantial behavior and attitude change through the use of coercive tactics, persuasion, and/or interpersonal and group-based influence manipulations (Schein 1961; Lifton 1961). Such programs have also been labeled "brainwashing" (Hunter 1951), a term more often used in the media than in scientific literature. However identified, these programs are distinguishable from other elaborate attempts to influence behavior and attitudes, to socialize, and to accomplish social control. Their distinguishing features are their totalistic qualities (Lifton 1961), the types of influence procedures they employ, and the organization of these procedures into three distinctive subphases of the overall process (Schein 1961; Ofshe and Singer 1986). The key factors that distinguish coercive persuasion from other training and socialization schemes are (1) the reliance on intense interpersonal and psychological attack

to destabilize an individual's sense of self to promote compliance, (2) the use of an organized peer group, (3) applying interpersonal pressure to promote conformity, and (4) the manipulation of the totality of the person's social environment to stabilize behavior once modified.

Thought-reform programs can vary considerably in their construction. The first systems studied ranged from those in which confinement and physical assault were employed (Schein 1956; Lifton 1954; Lifton 1961, pp. 19–85) to applications that were carried out under nonconfined conditions, in which nonphysical coercion substituted for assault (Lifton 1961, pp. 242–273; Schein 1961, pp. 290–298). The individuals to whom these influence programs were applied were in some cases unwilling subjects (prisoner populations) and in other cases volunteers who sought to participate in what they believed might be a career-beneficial, educational experience (Lifton 1961, p. 248).

Significant differences existed between the social environments and the control mechanisms employed in the two types of programs initially studied. Their similarities, however, are of more importance in understanding their ability to influence behavior and beliefs than are their differences. They shared the utilization of coercive persuasion's key effective-influence mechanisms: a focused attack on the stability of a person's sense of self; reliance on peer group interaction; the development of interpersonal bonds between targets and their controllers and peers; and an ability to control communication among participants. Edgar Schein captured the essential similarity between the types of programs in his definition of the coercive-persuasion phenomenon. Schein noted that even for prisoners, what happened was a subjection to "unusually intense and prolonged persuasion" that they could not avoid; thus, "they were coerced into allowing themselves to be persuaded" (Schein 1961, p. 18).

Programs of both types (confined/assaultive and nonconfined/nonassaultive) cause a range of cognitive and behavioral responses. The reported cognitive responses vary from apparently rare instances, classifiable as internalized belief change (enduring change), to a frequently observed transient alteration in beliefs that appears to be *situationally adaptive* and, finally, to reactions of nothing less than firm intellectual resistance and

hostility (Lifton 1961, pp. 117–151, 399–415; Schein 1961, pp. 157–166).

The phrase *situationally adaptive belief change* refers to attitude change that is *not* stable and *is environment dependent*. This type of response to the influence pressures of coercive-persuasion programs is perhaps the most surprising of the responses that have been observed. The combination of psychological assault on the self, interpersonal pressure, and the social organization of the environment creates a situation that can only be coped with by adapting and acting so as to present oneself to others in terms of the ideology supported in the environment (see below for discussion). Eliciting the desired verbal and interactive behavior sets up conditions likely to stimulate the development of attitudes consistent with and that function to rationalize new behavior in which the individual is engaging. Models of attitude change, such as the theory of Cognitive Dissonance (Festinger 1957) or Self-Perception Theory (Bem 1972), explain the tendency for consistent attitudes to develop as a consequence of behavior.

The surprising aspect of the situationally adaptive response is that the attitudes that develop are unstable. They tend to change dramatically once the person is removed from an environment that has totalistic properties and is organized to support the adaptive attitudes. Once removed from such an environment, the person is able to interact with others who permit and encourage the expression of criticisms and doubts, which were previously stifled because of the normative rules of the reform environment (Schein 1961, p. 163; Lifton 1961, pp. 87–116, 399–415; Ofshe and Singer 1986). This pattern of change, first in one direction and then the other, dramatically highlights the profound importance of social support in the explanation of attitude change and stability. This relationship has for decades been one of the principal interests in the field of social psychology.

Statements supportive of the proffered ideology that indicate adaptive attitude change during the period of the target's involvement in the reform environment and immediately following separation should not be taken as mere playacting in reaction to necessity. Targets tend to become genuinely involved in the interaction. The reform experience focuses on genuine vulnerabilities as

the method for undermining self-concept: manipulating genuine feelings of guilt about past conduct; inducing the target to make public denunciations of his or her prior life as being unworthy; and carrying this forward through interaction with peers for whom the target develops strong bonds. Involvement developed in these ways prevents the target from maintaining both psychological distance or emotional independence from the experience.

The reaction pattern of persons who display adaptive attitude-change responses is not one of an immediate and easy rejection of the proffered ideology. This response would be expected if they had been faking their reactions as a conscious strategy to defend against the pressures to which they were exposed. Rather, they appear to be conflicted about the sentiments they developed and their reevaluation of these sentiments. This response has been observed in persons reformed under both confined/assaultive and nonconfined/nonassaultive conditions (Schein 1962, pp. 163–165; Lifton 1961, pp. 86–116, 400–401).

Self-concept and belief-related attitude change in response to closely controlled social environments have been observed in other organizational settings that, like reform programs, can be classified as total institutions (Goffman 1957). Thought-reform reactions also appear to be related to, but are far more extreme than, responses to the typically less-identity-assaultive and less-totalistic socialization programs carried out by organizations with central commitments to specifiable ideologies, and which undertake the training of social roles (e.g., in military academies and religious-indoctrination settings (Dornbush 1955; Hulme 1956).

The relatively rare instances in which belief changes are internalized and endure have been analyzed as attributable to the degree to which the acquired belief system and imposed peer relations function to fully resolve the identity crisis that is routinely precipitated during the first phase of the reform process (Schein 1961, p. 164; Lifton 1961, pp. 131–132, 400). Whatever the explanation for why some persons internalize the proffered ideology in response to the reform procedures, this extreme reaction should be recognized as both atypical and probably attributable to an interaction between long-standing personality traits and

the mechanisms of influence utilized during the reform process.

Much of the attention to reform programs was stimulated because it was suspected that a predictable and highly effective method for profoundly changing beliefs had been designed, implemented, and was in operation. These suspicions are not supported by fact. Programs identified as thought reforming are not very effective at actually changing people's beliefs in any fashion that endures apart from an elaborate supporting social context. Evaluated only on the criterion of their ability to genuinely change beliefs, the programs have to be judged abject failures and massive wastes of effort.

The programs are, however, impressive in their ability to prepare targets for integration into and long-term participation in the organizations that operate them. Rather than assuming that individual belief change is the major goal of these programs, it is perhaps more productive to view the programs as elaborate role-training regimes. That is, as resocialization programs in which targets are being prepared to conduct themselves in a fashion appropriate for the social roles they are expected to occupy following conclusion of the training process.

If identified as training programs, it is clear that the goals of such programs are to reshape behavior and that they are organized around issues of social control important to the organizations that operate the programs. Their objectives then appear to be behavioral training of the target, which result in an ability to present self, values, aspirations, and past history in a style appropriate to the ideology of the controlling organization; to train an ability to reason in terms of the ideology; and to train a willingness to accept direction from those in authority with minimum apparent resistance. Belief changes that follow from successfully coercing or inducing the person to behave in the prescribed manner can be thought of as by-products of the training experience. As attitude-change models would predict, they arise "naturally" as a result of efforts to reshape behavior (Festinger 1957; Bem 1972).

The tactical dimension most clearly distinguishing reform processes from other sorts of training programs is the reliance on psychological coercion: procedures that generate pressure to

comply as a means of escaping a punishing experience (e.g., public humiliation, sleep deprivation, guilt manipulation, etc.). Coercion differs from other influencing factors also present in thought reform, such as content-based persuasive attempts (e.g., presentation of new information, reference to authorities, etc.) or reliance on influence variables operative in all interaction (status relations, demeanor, normal assertiveness differentials, etc.). Coercion is principally utilized to gain behavioral compliance at key points and to ensure participation in activities likely to have influencing effects; that is, to engage the person in the role-training activities and in procedures likely to lead to strong emotional responses, to cognitive confusion, or to attributions to self as the source of beliefs promoted during the process.

Robert Lifton labeled the extraordinarily high degree of social control characteristic of organizations that operate reform programs as their totalistic quality (Lifton 1961). This concept refers to the mobilization of the entirety of the person's social, and often physical, environment in support of the manipulative effort. Lifton identified eight themes or properties of reform environments that contribute to their totalistic quality: (1) control of communication, (2) emotional and behavioral manipulation, (3) demands for absolute conformity to behavior prescriptions derived from the ideology, (4) obsessive demands for confession, (5) agreement that the ideology is faultless, (6) manipulation of language in which clichés substitute for analytic thought, (7) reinterpretation of human experience and emotion in terms of doctrine, and (8) classification of those not sharing the ideology as inferior and not worthy of respect (Lifton 1961, pp. 419–437, 1987).

Schein's analysis of the behavioral sequence underlying coercive persuasion separated the process into three subphases: *unfreezing*, *change*, and *refreezing* (Schein 1961, pp. 111–139). Phases differ in their principal goals and their admixtures of persuasive, influencing, and coercive tactics. Although others have described the process differently, their analyses are not inconsistent with Schein's three-phase breakdown (Lifton 1961; Farber, Harlow, and West 1956; Meerloo 1956; Sargent 1957; Ofshe and Singer 1986). Although Schein's terminology is adopted here, the descriptions of phase activities have been broadened to reflect later research.

*Unfreezing* is the first step in eliciting behavior and developing a belief system that facilitates the long-term management of a person. It consists of attempting to undercut a person's psychological basis for resisting demands for behavioral compliance to the routines and rituals of the reform program. The goals of unfreezing are to destabilize a person's sense of identity (i.e., to precipitate an identity crisis), to diminish confidence in prior social judgments, and to foster a sense of powerlessness, if not hopelessness. Successful destabilization induces a negative shift in global self-evaluations and increases uncertainty about one's values and position in society. It thereby reduces resistance to the new demands for compliance while increasing suggestibility.

Destabilization of identity is accomplished by bringing into play varying sets of manipulative techniques. The first programs to be studied utilized techniques such as repeatedly demonstrating the person's inability to control his or her own fate, the use of degradation ceremonies, attempts to induce reevaluation of the adequacy and/or propriety of prior conduct, and techniques designed to encourage the reemergence of latent feelings of guilt and emotional turmoil (Hinkle and Wolfe 1956; Lifton 1954, 1961; Schein 1956, 1961; Schein, Cooley, and Singer 1960). Contemporary programs have been observed to utilize far more psychologically sophisticated procedures to accomplish destabilization. These techniques are often adapted from the traditions of psychiatry, psychotherapy, hypnotherapy, and the human-potential movement, as well as from religious practice (Ofshe and Singer 1986; Lifton 1987; Singer and Lalich 1995).

The *change* phase allows the individual an opportunity to escape punishing destabilization procedures by demonstrating that he or she has learned the proffered ideology, can demonstrate an ability to interpret reality in its terms, and is willing to participate in competition with peers to demonstrate zeal through displays of commitment. In addition to study and/or formal instruction, the techniques used to facilitate learning and the skill basis that can lead to opinion change include scheduling events that have predictable influencing consequences, rewarding certain conduct, and manipulating emotions to create punishing experiences. Some of the practices designed to promote influence might include requiring the target

to assume responsibility for the progress of less-advanced “students,” to become the responsibility of those further along in the program, to assume the role of a teacher of the ideology, or to develop ever more refined and detailed confession statements that recast the person’s former life in terms of the required ideological position. Group structure is often manipulated by making rewards or punishments for an entire peer group contingent on the performance of the weakest person, requiring the group to utilize a vocabulary appropriate to the ideology, making status and privilege changes commensurate with behavioral compliance, subjecting the target to strong criticism and humiliation from peers for lack of progress, and peer monitoring for expressions of reservations or dissent. If progress is unsatisfactory, the individual can again be subjected to the punishing destabilization procedures used during unfreezing to undermine identity, to humiliate, and to provoke feelings of shame and guilt.

*Refreezing* denotes an attempt to promote and reinforce behavior acceptable to the controlling organization. Satisfactory performance is rewarded with social approval, status gains, and small privileges. Part of the social structure of the environment is the norm of interpreting the target’s display of the desired conduct as demonstrating the person’s progress in understanding the errors of his or her former life. The combination of reinforcing approved behavior and interpreting its symbolic meaning as demonstrating the emergence of a new individual fosters the development of an environment-specific, supposedly reborn social identity. The person is encouraged to claim this identity and is rewarded for doing so.

Lengthy participation in an appropriately constructed and managed environment fosters peer relations, an interaction history, and other behavior consistent with a public identity that incorporates approved values and opinions. Promoting the development of an interaction history in which persons engage in cooperative activity with peers that is not blatantly coerced and in which they are encouraged but not forced to make verbal claims to “truly understanding the ideology and having been transformed,” will tend to lead them to conclude that they hold beliefs consistent with their actions (i.e., to make attributions to self as the source of their behaviors). These reinforcement procedures can result in a significant degree

of cognitive confusion and an alteration in what the person takes to be his or her beliefs and attitudes while involved in the controlled environment (Bem 1972; Ofshe et al. 1974).

Continuous use of *refreezing* procedures can sustain the expression of what appears to be significant attitude change for long periods of time. Maintaining compliance with a requirement that the person display behavior signifying unreserved acceptance of an imposed ideology and gaining other forms of long-term behavioral control requires continuous effort. The person must be carefully managed, monitored, and manipulated through peer pressure, the threat or use of punishment (material, social, and emotional) and through the normative rules of the community (e.g., expectations prohibiting careers independent of the organization, prohibiting formation of independent nuclear families, prohibiting accumulation of significant personal economic resources, etc.) (Whyte 1976; Ofshe 1980; Ofshe and Singer 1986).

The rate at which a once-attained level of attitude change deteriorates depends on the type of social support the person receives over time (Schein 1961 pp. 158–166; Lifton pp. 399–415). In keeping with the *refreezing* metaphor, even when the reform process is to some degree successful at shaping behavior and attitudes, the new shape tends to be maintained only as long as temperature is appropriately controlled.

One of the essential components of the reform process in general and of long-term *refreezing* in particular is monitoring and controlling the contents of communication among persons in the managed group (Lifton 1961; Schein 1960; Ofshe et al. 1974). If successfully accomplished, communication control eliminates a person’s ability to safely express criticisms or to share private doubts and reservations. The result is to confer on the community the quality of being a spy system of the whole, upon the whole.

The typically observed complex of communication-controlling rules requires people to self-report critical thoughts to authorities or to make doubts known only in approved and readily managed settings (e.g., small groups or private counseling sessions). Admitting “negativity” leads to punishment or reindoctrination through procedures

sometimes euphemistically termed “education” or “therapy.” Individual social isolation is furthered by rules requiring peers to “help” colleagues to progress, by reporting their expressions of doubt. If it is discovered, failure to make a report is punishable, because it reflects on the low level of commitment of the person who did not “help” a colleague to make progress.

Controlling communication effectively blocks individuals from testing the appropriateness of privately held critical perceptions against the views of even their families and most-valued associates. Community norms encourage doubters to interpret lingering reservations as signs of a personal failure to comprehend the truth of the ideology; if involved with religious organizations, to interpret doubt as evidence of sinfulness or the result of demonic influences; if involved with an organization delivering a supposed psychological or medical therapy, as evidence of continuing illness and/or failure to progress in treatment.

The significance of communication control is illustrated by the collapse of a large psychotherapy organization in immediate reaction to the leadership’s loss of effective control over interpersonal communication. At a meeting of several hundred of the members of this “therapeutic community” clients were allowed to openly voice privately held reservations about their treatment and exploitation. They had been subjected to abusive practices which included assault, sexual and economic exploitation, extremes of public humiliation, and others. When members discovered the extent to which their sentiments about these practices were shared by their peers they rebelled (Ayalla 1998).

Two widespread myths have developed from misreading the early studies of thought-reforming influence systems (Zablocki 1991). These studies dealt in part with their use to elicit false confessions in the Soviet Union after the 1917 revolution; from American and United Nations forces held as POWs during the Korean War; and from their application to Western missionaries held in China following Mao’s revolution.

The first myth concerns the necessity and effectiveness of physical abuse in the reform process. The myth is that physical abuse is not only necessary but is the prime cause of apparent belief

change. Reports about the treatment of POWs and foreign prisoners in China documented that physical abuse was present. Studies of the role of assault in the promotion of attitude change and in eliciting false confessions even from U.S. servicemen revealed, however, that it was ineffective. Belief change and compliance was more likely when physical abuse was minimal or absent (Biderman 1960). Both Schein (1961) and Lifton (1961) reported that physical abuse was a minor element in the understanding of even prison reform programs in China.

In the main, efforts at resocializing China’s nationals were conducted under nonconfined/nonassaultive conditions. Millions of China’s citizens underwent reform in schools, special-training centers, factories, and neighborhood groups in which physical assault was not used as a coercive technique. One such setting for which many participants actively sought admission, the “Revolutionary University,” was classified by Lifton as the “hard core of the entire Chinese thought reform movement” (Lifton 1961, p. 248).

Attribution theories would predict that if there were differences between the power of reform programs to promote belief change in settings that were relatively more or less blatantly coercive and physically threatening, the effect would be greatest in less-coercive programs. Consistent with this expectation, Lifton concluded that reform efforts directed against Chinese citizens were “much more successful” than efforts directed against Westerners (Lifton 1961, p. 400).

A second myth concerns the purported effects of brainwashing. Media reports about thought reform’s effects far exceed the findings of scientific studies—which show coercive persuasion’s upper limit of impact to be that of inducing personal confusion and significant, but typically transitory, attitude change. Brainwashing was promoted as capable of stripping victims of their *capacity* to assert their wills, thereby rendering them unable to resist the orders of their controllers. People subjected to “brainwashing” were not merely influenced to adopt new attitudes but, according to the myth, suffered essentially an alteration in their psychiatric status from *normal* to *pathological*, while losing their capacity to decide to comply with or resist orders.

This lurid promotion of the power of thought-reforming influence techniques to change a person's capacity to resist direction is entirely without basis in fact: No evidence, scientific or otherwise, supports this proposition. No known mental disorder produces the loss of will that is alleged to be the result of brainwashing. Whatever behavior and attitude changes result from exposure to the process, they are most reasonably classified as the responses of normal individuals to a complex program of influence.

The U.S. Central Intelligence Agency seems to have taken seriously the myth about brainwashing's power to destroy the will. Due, perhaps, to concern that an enemy had perfected a method for dependably overcoming will—or perhaps in hope of being the first to develop such a method—the Agency embarked on a research program, code-named MKULTRA. It was a pathetic and tragic failure. On the one hand, it funded some innocuous and uncontroversial research projects; on the other, it funded or supervised the execution of several far-fetched, unethical, and dangerous experiments that failed completely (Marks 1979; Thomas 1989).

Although no evidence suggests that thought reform is a process capable of stripping a person of their will to resist, a relationship does exist between thought reform and changes in psychiatric status. The stress and pressure of the reform process cause some percentage of *psychological casualties*. To reduce resistance and to motivate behavior change, thought-reform procedures rely on psychological stressors, induction of high degrees of emotional distress, and on other intrinsically dangerous influence techniques (Heide and Borkovec 1983). The process has a potential to cause psychiatric injury, which is sometimes realized. The major early studies (Hinkle and Wolfe 1961; Lifton 1961; Schein 1961) reported that during the *unfreezing* phase individuals were intentionally stressed to a point at which some persons displayed symptoms of being on the brink of psychosis. Managers attempted to reduce psychological pressure when this happened, to avoid serious psychological injury to those obviously near the breaking point.

Contemporary programs speed up the reform process through the use of more psychologically

sophisticated and dangerous procedures to accomplish destabilization. In contemporary programs the process is sometimes carried forward on a large group basis, which reduces the ability of managers to detect symptoms of impending psychiatric emergencies. In addition, in some of the “therapeutic” ideologies espoused by thought-reforming organizations, extreme emotional distress is valued positively, as a sign of progress. Studies of contemporary programs have reported on a variety of psychological injuries related to the reform process. Injuries include psychosis, major depressions, manic episodes, and debilitating anxiety (Glass, Kirsch, and Parris 1977, Haaken and Adams 1983, Heide and Borkovec 1983; Higget and Murray 1983; Kirsch and Glass 1977; Yalom and Lieberman 1971; Lieberman 1987; Singer and Ofshe 1990; Singer and Lalich 1995, 1996).

#### **HIGH CONTROL GROUPS AND LARGE GROUP AWARENESS TRAINING**

Political reeducation in China was backed by the power of the state to request, if not to compel participation. Contemporary examples of the use of extreme influence that rise to the level of attempting to induce a major shift in belief typically lack the power to coerce someone into remaining involved long enough for the persuasive procedures of the program to unfreeze, change, and refreeze a person's values, perceptions, and preferences. Rather, participants must be attracted to the group and participate long enough for the tactics utilized during the unfreezing phase to destabilize the person's sense of identity so that he or she will be motivated to adopt the new world view or the new perceptions promoted by the program's managers, will desire to maintain association with committed believers in the new perspective, and will thereby obtain the long-term social support necessary to restabilize his or her identity.

Contemporary thought-reform programs are generally far more sophisticated in their selection of both destabilization and influence techniques than were the programs studied during the 1950s (Ofshe and Singer 1986; Singer and Lalich 1995, 1996). For example, hypnosis was entirely absent from the first programs studied but is often observed in modern programs. In most modern

examples in which hypnosis is present, it functions as a remarkably powerful technique for manipulating subjective experience and for intensifying emotional response. It provides a method for influencing people to imagine impossible events such as those that supposedly occurred in their "past lives," the future, or during visits to other planets. If persons so manipulated misidentify the hypnotically induced fantasies, and classify them as previously unavailable memories, their confidence in the content of a particular ideology can be increased (Bainbridge and Stark 1980).

Hypnosis can also be used to lead people to allow themselves to relive actual traumatic life events (e.g., rape, childhood sexual abuse, near-death experiences, etc.) or to fantasize the existence of such events and, thereby, stimulate the experience of extreme emotional distress. When imbedded in a reform program, repeatedly leading the person to experience such events can function simply as punishment, useful for coercing compliance.

Accounts of contemporary programs also describe the use of sophisticated techniques intended to strip away psychological defenses, to induce regression to primitive levels of coping, and to flood targets with powerful emotion (Ayalla 1998; Haaken and Adams 1983; Hockman 1984; Temerlin and Temerlin 1982; Singer and Lalich 1996). In some instances stress and fatigue have been used to promote hallucinatory experiences that are defined as therapeutic (Gerstel 1982). Drugs have been used to facilitate disinhibition and heightened suggestibility (Watkins 1980). Thought-reform subjects have been punished for disobedience by being ordered to self-inflict severe pain, justified by the claim that the result will be therapeutic (Bellack et al. *v.* Murietta Foundation et al.).

Programs attempting thought reform appear in various forms in contemporary society. They depend on the voluntary initial participation of targets. This is usually accomplished because the target assumes that there is a common goal that unites him or her with the organization or that involvement will confer some benefit (e.g., relief of symptoms, personal growth, spiritual development, etc.). Apparently some programs were developed based on the assumption that they could be used to facilitate desirable changes (e.g., certain

rehabilitation or psychotherapy programs). Some religious organizations and social movements utilize them for recruitment purposes. Some commercial organizations utilize them as methods for promoting sales. In some instances, reform programs appear to have been operated for the sole purpose of gaining a high degree of control over individuals to facilitate their exploitation (Ofshe 1986; McGuire and Norton 1988; Watkins 1980).

Virtually any acknowledged expertise or authority can serve as a power base to develop the social structure necessary to carry out thought reform. In the course of developing a new form of rehabilitation, psychotherapy, religious organization, utopian community, school, or sales organization, it is not difficult to justify the introduction of thought-reform procedures.

Perhaps the most famous example of a thought-reforming program developed for the ostensible purpose of rehabilitation was Synanon, a drug-treatment program (Sarbin and Adler 1970, Yablonsky 1965; Ofshe et al. 1974). The Synanon environment possessed all of Lifton's eight themes. It used as its principle coercive procedure a highly aggressive encounter/therapy group interaction. In form it resembled "struggle groups" observed in China (Whyte 1976), but it differed in content. Individuals were vilified and humiliated not for past political behavior but for current conduct as well as far more psychologically intimate subjects, such as early childhood experiences, sexual experiences, degrading experiences as adults, etc. The coercive power of the group experience to affect behavior was substantial as was its ability to induce psychological injury (Lieberman, Yalom, and Miles 1973; Ofshe et al. 1974).

Allegedly started as a drug-rehabilitation program, Synanon failed to accomplish significant long-term rehabilitation. Eventually, Synanon's leader, Charles Diederich, promoted the idea that any degree of drug abuse was incurable and that persons so afflicted needed to spend their lives in the Synanon community. Synanon's influence program was successful in convincing many that this was so. Under Diederich's direction, Synanon evolved from an organization that espoused non-violence into one that was violent. Its soldiers were dispatched to assault and attempt to murder persons identified by Diederich as Synanon's enemies (Mitchell, Mitchell, and Ofshe 1981).



The manipulative techniques of self-styled messiahs, such as People's Temple leader Jim Jones (Reiterman 1982), and influence programs operated by religious organizations, such as the Unification Church (Taylor 1978) and Scientology (Wallis 1977; Bainbridge and Stark 1980), can be analyzed as thought-reform programs. The most controversial recruitment system operated by a religious organization in recent American history was that of the Northern California branch of the Unification Church (Reverend Moon's organization). The influence program was built directly from procedures of psychological manipulation that were commonplace in the human-potential movement (Bromley and Shupe 1981). The procedures involved various group-based exercises as well as events designed to elicit from participants information about their emotional needs and vulnerabilities. Blended into this program was content intended to slowly introduce the newcomer to the group's ideology. Typically, the program's connection with the Unification Church or any religious mission was denied during the early stages of the reform process. The target was monitored around the clock and prevented from communicating with peers who might reinforce doubt and support a desire to leave. The physical setting was an isolated rural facility far from public transportation.

Initial focus on personal failures, guilt-laden memories, and unfulfilled aspirations shifted to the opportunity to realize infantile desires and idealistic goals, by affiliating with the group and its mission to save the world. The person was encouraged to develop strong affective bonds with current members. They showed unflinching interest, affection, and concern, sometimes to the point of spoon-feeding the person's meals and accompanying the individual everywhere, including to the toilet. If the *unfreezing* and *change* phases of the program succeeded, the individual was told of the group's affiliation with the Unification Church and assigned to another unit of the organization within which *refreezing* procedures could be carried forward.

#### **POLICE INTERROGATION AND FALSE CONFESSIONS**

Police interrogation in America was transformed in the twentieth century from a method of gaining

compliance relying largely on physical coercion into one that relies almost exclusively on psychological means (Leo 1992). Psychological methods of interrogation were developed to influence persons who know they are guilty of a crime to change their decision to deny guilt, to admit responsibility for the crime, and to confess fully to their role in the crime (see Hilgendorf and Irving 1981; Ofshe and Leo 1997a,b). Compared to the other extreme influence environments described in this entry, a police interrogation is of relatively short duration and is entirely focused on the single issue of eliciting a confession. It is nevertheless an environment in which a person can be influenced to make a dramatic shift in position—from denial of guilt to confession.

The techniques and tactics that lead a guilty suspect to admit guilt constitute an impressive display of the power of influence to change a person's decision even when the consequence of the shift is obviously disadvantageous. If the procedures of interrogation are misused, modern interrogation methods can have an even more impressive result. If the influence procedures and techniques of modern interrogation methods are directed at innocent persons, some false confessions will result (Bedau and Radelet 1987; Leo and Ofshe 1998). In most instances when an innocent person is led to give a false confession the cause is coercion—the use of a threat of severe punishment if the person maintains that he is innocent and an offer of leniency if he complies with the interrogator's demand to confess. Most persons who decide to comply and offer a false confession in response to coercion remain certain of their innocence and know that they are falsely confessing in order to avoid the most severe possible punishment.

Under some circumstances, however, interrogation tactics can cause an innocent person to give what is called a persuaded false confession—a false confession that is believed to be true when it is given. Influence procedures now commonly used during modern police interrogation can sometimes inadvertently manipulate innocent persons' beliefs about their own innocence and, thereby, cause them to falsely confess. Confessions resulting from accomplishing the *unfreezing* and *change* phases of thought reform are classified as persuaded false confessions (Kassin and Wrightsman 1985;

Gudjonsson and MacKeith 1988; Ofshe and Leo 1997a). Although they rarely come together simultaneously, the ingredients necessary to elicit a temporarily believed false confession are: erroneous police suspicion, the use of *certain* commonly employed interrogation procedures, and some degree of psychological vulnerability in the suspect. Philip Zimbardo (1971) has reviewed the coercive factors generally present in modern interrogation settings. Richard Ofshe and Richard Leo (1989, 1997a) have identified those influence procedures that if present in a suspect's interrogation contribute to causing *unfreezing* and *change*.

Techniques that contribute to *unfreezing* include falsely telling a suspect that the police have evidence proving the person's guilt (e.g., fingerprints, eyewitness testimony, etc.). Suspects may be given a polygraph examination and then falsely told (due either to error or design) that they failed and the test reveals their unconscious knowledge of guilt. Suspects may be told that their lack of memory of the crime was caused by an alcohol- or drug-induced blackout, was repressed, or is explained because the individual is a multiple personality.

The techniques listed above regularly appear in modern American police interrogations. They are used to lead persons who know that they have committed the crime at issue to decide that the police have sufficient evidence to convict them or to counter typical objections to admitting guilt (e.g., "I can't remember having done that."). In conjunction with the other disorienting and distressing elements of a modern accusatory interrogation, these tactics can sometimes lead innocent suspects to doubt themselves and question their lack of knowledge of the crime. If innocent persons subjected to these sorts of influence techniques do not reject the false evidence and realize that the interrogators are lying to them, they have no choice but to doubt their memories. If the interrogator supplies an explanation for why the suspect's memory is untrustworthy, the person may reason that "I must have committed this crime."

Tactics used to *change* the suspect's position and elicit a confession include maneuvers designed to intensify feelings of guilt and emotional distress following from the suspect's assumption of guilt.

Suspects may be offered an escape from the emotional distress through confession. It may also be suggested that confession will provide evidence of remorse that will benefit the suspect in court.

### RECOVERED MEMORY PSYCHOTHERAPY

The shifts in belief and conduct promoted during political reeducation concern an intellectual analysis of society and one's role in it; in the context of high control groups the beliefs are about theology and one's role in a community; and in police interrogation the newly created beliefs concern a crime the suspect has no knowledge of having committed. The progression in these illustrations is from influence efforts directed at an intellectual or philosophical assumption to that of an influence effort directed at changing beliefs about a single fact—Did the suspect commit a crime and not know of it due to a memory defect? The final example of extreme influence and belief change is arguably the most dramatic example of the power of interpersonal influence and particular influence techniques to change a person's beliefs about the historical truth relating to a major dimension of his or her life history, such as whether or not he or she had been viciously raped and horribly brutalized by a parent (or parents, or siblings, or teachers, or neighbors, etc.) for periods of as long as two decades.

Psychotherapy directed at causing a patient to retrieve allegedly repressed and therefore supposedly unavailable memories (e.g., of sexual abuse, and/or having spent one's life suffering from an unrecognized multiple personality disorder, and/or having spent one's childhood and teenage years as a member of a murderous satanic cult) is one of the most stunning examples of psychological/psychiatric quackery of the twentieth century. It is also perhaps the most potent example of the power of social influence to predictably create beliefs (in this case beliefs that are utterly false) and thereby alter a person's choices and conduct (Loftus 1994; Perdergrast 1995; Ofshe and Watters 1994; Spanos 1996).

The elements of a thought-reforming process are visible in the steps through which a patient undergoing recovered memory therapy is manipulated. The patient's identity is destabilized by the therapist's insistence that the reason for the patient's distress or mental illness is that she or he

suffered a series of sexual traumas in childhood that he or she don't know about because he or she has *repressed* them. Recovered memory therapists organize their treatment programs on the presumption that repression exists despite the fact that the notion of repression was never more than a fanciful, unsubstantiated speculation that was long ago rejected as useless by the scientific community that studies human memory (Loftus 1994; Crews 1998; Watters and Ofshe 1999). Recovered memory therapists rely on assertions of authority, claims of expertise, and outright trickery (e.g., inducing the patient to experience hypnotic fantasies of sexual abuse that they lead patients to misclassify as memories and interpreting the patient's dreams as proof that they suffer from repressed memories) to convince the patient of the existence of repression. Recovered memory therapists rely on the alleged "repression mechanism" to undermine patients' confidence in their normal memories of their lives. If patients can be successfully convinced that repression exists and that the hypnotic fantasies that the therapist suggests to them are memories of events that happened during their lives, they can no longer trust that they know even the broad outlines of their life histories.

Once the patient's identity has been destabilized, the therapist guides the patients to build a new identity centered on his or her status as victim of sexual abuse. This victim role may include requiring the patient to learn to act as if she or he suffers from multiple personality disorder, or suggesting that the patient publicly denounce, sue, or file criminal charges against the persons who supposedly abused them.

### CONCLUSION

Extreme influence environments are not easy to study. The history of research on extreme influence has been one in which most of the basic descriptive work has been conducted through post-hoc interviewing of persons exposed to the influence procedures. The second-most frequently employed method has been that of participant observation. In connection with work being done on police interrogation methods, it has been possible to analyze contemporaneous recordings of interrogation sessions in which targets' beliefs are actually made to undergo radical change. All this work

has contributed to the development of an understanding in several ways.

Studying these environments demonstrates that the extremes of influence are no more or less difficult to understand than any other complex social event. The characteristics that distinguish extreme influence environments from other examples of social settings are influence in the order in which the influence procedures are assembled and the degree to which the target's environment is manipulated in the service of social control. These are at most unusual arrangements of commonplace bits and pieces.

As it is with all complex, real-world social phenomena that cannot be studied experimentally, understanding information about the thought-reform process proceeds through the application of theories that have been independently developed. Explaining data that describe the type and organization of the influence procedures that constitute an extreme influence environment depends on applying established social-psychological theories about the manipulation of behavior and attitude change. Assessing reports about the impact of the experiences on the persons subjected to intense influence procedures depends on the application of current theories of personality formation and change. Understanding instances in which the thought-reform experience appears related to psychiatric injury requires proceeding as one would ordinarily in evaluating any case history of a stress-related or other type of psychological injury.

(SEE ALSO: *Attitudes; Persuasion; Social Control*)

### REFERENCES

- Ayalla, Marybeth 1998 *Insane Therapy*. Philadelphia: Temple University Press.
- Bainbridge, William S., and Rodney Stark 1980 "Scientology, to Be Perfectly Clear." *Sociological Analysis* 41:128-136.
- Bedau, Hugo, and Michael Radelet 1987 "Miscarriages of Justice in Potentially Capital Cases." *Stanford University Law Review* 40:21-197.
- Bellack, Catherine et al. *v.* Murietta Foundation et al. United States District Court, Central District of California. Civil No. 87-08597.
- Bem, Darryl 1972 "Self-Perception Theory." In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, vol 6. New York: Academic.

- Biderman, Albert D. 1960 "Social-Psychological Needs and Involuntary Behavior as Illustrated by Compliance in Interrogation." *Sociometry* 23:120-147.
- Bromley, David G., and Anson D. Shupe, Jr. 1981 *Strange Gods*. Boston: Beacon Press.
- Dornbush, Sanford M. 1955 "The Military Academy as an Assimilating Institution." *Social Forces* 33:316-321.
- Farber, I. E., Harry F. Harlow, and Louis J. West 1956 "Brainwashing, Conditioning and DDD: Debility, Dependency and Dread." *Sociometry* 20:271-285.
- Festinger, Leon 1957 *A Theory of Cognitive Dissonance*. Evanston, Ill.: Row Peterson.
- Gerstel, David 1982 *Paradise Incorporated: Synanon*. Novato, Calif.: Presidio.
- Glass, Leonard L., Michael A. Kirsch, and Frederick A. Parris 1977 "Psychiatric Disturbances Associated with Erhard Seminars Training: I. A Report of Cases." *American Journal of Psychiatry* 134:245-247.
- Goffman, Erving 1957 "On the Characteristics of Total Institutions." Proceedings of the Symposium on Preventive and Social Psychiatry. Washington, D.C.: Walter Reed Army Institute of Research.
- Gudjonsson, Gisli, and Bent Lebegue 1989 "Psychological and Psychiatric Aspects of a Coerced-Internalized False Confession." *Journal of Forensic Science* 29:261-269.
- Gudjonsson, Gisli H., and James A. MacKeith 1988 "Retracted Confessions: Legal, Psychological and Psychiatric Aspects." *Medical Science and Law* 28: 187-194.
- Haaken, Janice, and Richard Adams 1983 "Pathology as 'Personal Growth': A Participant Observation Study of Lifespring Training." *Psychiatry* 46:270-280.
- Heide, F. J., and T. D. Borkovec 1984 "Relaxation Induced Anxiety: Mechanism and Theoretical Implications." *Behavior Research and Therapy* 22:1-12.
- Higget, Anna C., and Robin M. Murray 1983 "A Psychotic Episode Following Erhard Seminars Training." *Acta Psychiatrica Scandinavica* 67:436-439.
- Hilgendorf, Edward, and Barrie Irving 1981 "A Decision-Making Model of Confession." In S.M.A. Loyd-Bostock, ed., *Psychology in Legal Contexts*. London: Macmillan.
- Hinkle, L. E., and Harold G. Wolfe 1956 "Communist Interrogation and Indoctrination of Enemies of the State." *Archives of Neurology and Psychiatry* 20:271-285.
- Hochman, John A. 1984 "Iatrogenic Symptoms Associated with a Therapy Cult: Examination of an Extinct 'New Therapy' with Respect to Psychiatric Deterioration and 'Brainwashing'." *Psychiatry* 47:366-377.
- Hulme, Kathryn 1956 *The Nun's Story*. Boston: Little, Brown.
- Hunter, Edward 1951 *Brain-washing in China*. New York: Vanguard.
- Kassin, Saul, and Lawrence Wrightsman 1985 "Confession Evidence." In Samuel Kassin and Lawrence Wrightsman, eds., *The Psychology of Evidence and Trial Procedure*. London: Sage.
- Kirsch, Michael A., and Leonard L. Glass 1977 "Psychiatric Disturbances Associated with Erhard Seminars Training: II. Additional Cases and Theoretical Considerations." *American Journal of Psychiatry* 134: 1254-1258.
- Leo, Richard 1992 "From Coercion to Deception: The Changing Nature of Police Interrogation in America." *Journal of Criminal Law and Social Change* 88, 27-52.
- and Richard Ofshe 1998 "The Consequences of False Confessions." *Journal of Criminal Law and Criminology* 88, 2:429-496.
- Lieberman, Morton A. 1987 "Effect of Large Group Awareness Training on Participants' Psychiatric Status." *American Journal of Psychiatry*, 144:460-464.
- , Irvin D. Yalom, and M. B. Miles 1973 *Encounter Groups: First Facts*. New York: Basic Books.
- Lifton, Robert J. 1954 "Home by Ship: Reaction Patterns of American Prisoners of War Repatriated from North Korea." *American Journal of Psychiatry* 110:732-739.
- 1961 *Thought Reform and the Psychology of Totalism*. New York: Norton.
- 1986 *The Nazi Doctors*. New York: Basic Books.
- 1987 "Cults: Totalism and Civil Liberties." In Robert J. Lifton, ed., *The Future of Immortality and Other Essays for a Nuclear Age*. New York: Basic Books.
- Loftus, Elizabeth, and Katherine Ketcham 1994 *The Myth of Repressed Memory*. New York; St. Martins.
- Marks, John 1979 *The Search for the Manchurian Candidate*. New York: Dell.
- McGuire, Christine, and Carla Norton 1988 *Perfect Victim*. New York: Arbor House.
- Meerloo, Jorst A. 1956 *The Rape of the Mind: The Psychology of Thought Control, Menticide and Brainwashing*. Cleveland, Ohio: World Publishing.
- Mitchell, David V., Catherine Mitchell, and Richard J. Ofshe 1981 *The Light on Synanon*. New York: Seaview.
- Ofshe, Richard J. 1980 "The Social Development of the Synanon Cult: The Managerial Strategy of Organizational Transformation." *Sociological Analysis* 41: 109-127.
- 1986 "The Rabbi and the Sex Cult." *Cultic Studies Journal* 3:173-189.

- 1989 “Coerced Confessions: The Logic of Seemingly Irrational Action.” *Cultic Studies Journal* 6:1–15.
- 1992 “Inadvertant Hypnosis During Interrogation: False Confession Due to Dissociative State; Mis-Identified Multiple Personality and the Satanic Cult Hypothesis.” *International Journal of Clinical and Experimental Hypnosis* 40, (3):125–156.
- , and Richard Leo 1997a “The Social Psychology of Police Interrogation: The Theory and Classification of True and False Confessions.” *Studies in Law, Politics and Society* 16: 190–251.
- , 1997b “The Decision to Confess Falsely: Rational Choice and Irrational Action.” *Denver University Law Review* 74, 4:979–1122.
- Ofshe, Richard and Margaret T. Singer 1986 “Attacks on Peripheral Versus Central Elements of Self and the Impact of Thought Reforming Techniques.” *Cultic Studies Journal* 3:3–24.
- , Nancy Eisenberg, Richard Coughlin, and Gregory Dolinajec 1974 “Social Structure and Social Control in Synanon.” *Voluntary Action Research* 3:67–76.
- Ofshe, Richard, and Ethan Watters 1994 *Making Monsters*. New York: Scribners.
- Perdergrast, Mark 1995 *Victims of Memory*. Hinesberg, Vt.: Upper Access.
- Reiterman, Timothy, and Dan Jacobs 1982 *The Raven*. New York: Dutton.
- Sarbin, Theodore R., and Nathan Adler 1970 “Self-Reconstitution Processes: A Preliminary Report.” *Psychoanalytic Review* 4:599–616.
- Sargent, William 1957 *Battle for the Mind: How Evangelists, Psychiatrists, and Medicine Men Can Change Your Beliefs and Behavior*. Garden City, N.Y.: Doubleday.
- Schein, Edgar W. 1961 *Coercive Persuasion*. New York: Norton.
- , W. E. Cooley, and Margaret T. Singer 1960 *A Psychological Follow-up of Former Prisoners of the Chinese Communists, Part I, Results of Interview Study*. Cambridge: MIT.
- Shurmann, Franz 1968 *Ideology and Organization in Communist China*. Berkeley: University of California Press.
- Singer, Margaret, and Janja Lalich 1995 *Cults in Our Midst*. San Francisco: Jossey-Bass.
- 1996 *Crazy Therapies*. San Francisco: Jossey-Bass.
- Singer, Margaret T., and Richard J. Ofshe 1990 “Thought Reform Programs and the Production of Psychiatric Casualties.” *Psychiatric Annals* 20:188–193.
- Taylor, David 1978 “Social Organization and Recruitment in the Unification Church.” Master’s diss., University of Montana.
- Temerlin, Maurice K., and Jane W. Temerlin 1982 “Psychotherapy Cults: An Iatrogenic Phenomenon.” *Psychotherapy Theory, Research Practice* 19:131–41.
- Thomas, Gordon 1989 *Journey Into Madness*. New York: Bantam.
- Wallis, Roy 1977 *The Road to Total Freedom*. New York: Columbia University Press.
- Watkins, Paul 1980 *My Life With Charles Manson*. New York: Bantam.
- Whyte, Martin K. 1976 *Small Groups and Political Behavior in China*. Berkeley: University of California Press.
- Wright, Stewart 1987 *Leaving Cults: The Dynamics of Defection*. Society of the Scientific Study of Religion, Monograph no. 7, Washington, D.C.
- Yablonski, Louis 1965 *The Tunnel Back: Synanon*. New York: Macmillan.
- Yalom, Irvin D., and Morton Lieberman 1971 “A Study of Encounter Group Casualties.” *Archives of General Psychiatry* 25:16–30.
- Zablocki, Benjamin 1991 *The Scientific Investigation of the Brainwashing Conjecture*. Washington D.C.: American Association for the Advancement of Science.
- Zimbardo, Philip G. 1971 “Coercion and Compliance.” In Charles Perruci and Mark Pilisuk, eds., *The Triple Revolution*. Boston: Little, Brown.

RICHARD J. OFSHE

# F

---

## FACTOR ANALYSIS

Factor analysis is a mathematical and statistical technique for analyzing differences among units of analysis and the structure of relationships among variables assessing those units. The units of analysis may be persons, groups, organizations, ecological units, or any other justifiable basis of aggregation although persons are most often the focus of analysis. The chief purpose of the method is the attainment of scientific parsimony, which is achieved by positing a set of latent common factors that underlie the data. The factor model was developed by Charles Spearman (1904a, 1927) to be used to describe economically the correlations among mental test scores observed for persons. Spearman's famous bi-factor model of intelligence held that measures of mental abilities had two major sources: a factor common to all measures of ability, which he called the *g*-factor (factor of general ability), and a specific component of variation (an *s*-factor) unique to the test. For example, a test of numerical ability may be affected in part by a general factor of intelligence as well as a factor specific to numerical aptitude. This model, although never the predominant psychological theory of mental tests, has persisted in the culture in the sense that people often believe there is a general factor of intelligence underlying performance across different domains (see Gould 1981 for a critique of this view).

Although Spearman's work did not go very far beyond such a simple model, his approach to model construction and theory testing using tetrad

differences has provided the basis for much further work (see, e.g., Glymour et al. 1987). Many contemporaries of Spearman—Cyril Burt, Karl Pearson, Godfrey Thomson, J. C. Maxwell Garnett, and others—working in the fields of human abilities and statistics also contributed to the development of factor analysis. Several worked to modify Spearman's bi-factor model to include multiple factors of intelligence. But the most radical departure from the *g*-factor view of human intelligence came with Thurstone's (1938) publication of *Primary Mental Abilities*, in which he demonstrated empirically through the application of *multiple factor analysis* that several common factors were necessary to explain correlations among measures of human abilities. While Thurstone (1947) is usually credited with the popularization of this more general technique, the concept of multiple factor analysis first arose in the work of Garnett (1919–1920; see Harmon 1976).

*Multiple factor analysis* proved to be a major advance over the Spearman model, which was later to be seen as a special case (the one-factor case). Multiple factor analysis permitted a general solution to the possibility of positing multiple factors (*k*) in a set of variables (*p*). Within this framework, two competing research strategies emerged, each resting on distinct principles. One was based on Pearson's principle of *principal axes*, which was later developed by Hotelling (1933) as the method of *principal components*. This approach emphasized the objective of "extracting" a maximum of variance from the set of *p* variables so that the *k* factors explained as much of the variance in

the variables as they could. This tradition still exists in approaches to factor analysis that rely on principal components analysis, and, although many researchers use the technique, they are likely to be unaware of the objectives underlying the approach (see Harman 1976).

In contrast to the strategy of maximizing the variance explained in the variables, the other basic strategy—more squarely in the tradition of Spearman—emphasized the objective of reproducing the observed correlations among the variables. These two objectives—one emphasizing the extraction of maximum variance in the variables and the other emphasizing the fit to the correlations among the variables—eventually became the object of serious debate. However, with time there has emerged a consensus that the “debate” between these approaches rested on a misconception. The method of principal axes, which is the basis of principal components analysis—involving the analysis of a correlation matrix with unities in the diagonal—is now better understood as a computational method and not a model, as the factor analysis approach is now considered (see Maxwell 1977).

The early developments in the field of factor analysis and related techniques were carried out primarily by psychometricians. These early developments were followed by many important contributions to estimation, computation, and model construction during the post-World War II period. Some of the most important contributions to the method during the 1950s were made by a sociologist, Louis Guttman. Guttman made important contributions to the resolution of the issue of deciding upon the best number of latent factors (Guttman 1954), the problem of “factor indeterminacy” (Guttman 1955), and the problem of estimating communalities (Guttman 1956), among many others. Guttman (1960) also invented yet a third model, called *image analysis*, which has a certain elegance but is rarely used (see Harris 1964; Kaiser 1963).

Research workers in many fields made contributions to the problem of deciding how best to represent a particular factor model in a theoretical/geometrical space, via the transformation or rotation of factors. Methods of rotation included the quartimax (Neuhaus and Wrigley 1954), varimax (Kaiser 1958), and oblimax (Harman 1976), among

others. Several contributions were made during the early development of factor analysis with respect to the most useful strategies for estimating factor scores (see reviews by Harris 1967 and McDonald and Burr 1967) and for dealing with the problem of assessing factorial invariance (e.g., Meredith 1964a, 1964b; Mulaik 1972). Beginning in the mid-1960s the advances in the field of factor analysis have focused on the development of maximum-likelihood estimation techniques (Lawley and Maxwell 1971; Jöreskog 1967; Jöreskog and Lawley 1968); alternative distribution-free techniques (Bentler 1983, 1989; Bentler and Weeks 1980; Browne 1974, 1984; Browne and Shapiro 1988); the development of confirmatory factor analysis, which permits the setting of specific model constraints on the data to be analyzed (Bentler 1989; Jöreskog 1966, 1967, 1970, 1971a, 1973; Jöreskog and Sörbom 1986); and the development of factor analysis strategies for categorical variables (Christofferson 1975; Jöreskog and Sörbom 1988; Muthén 1983, 1988).

Factor analysis is used extensively by sociologists as a research tool. It is used in at least four related ways. *First*, it is frequently used as a data reduction or item analysis technique in index construction. *Second*, it is used as an exploratory device for examining the dimensional structure of content within a well-specified domain. *Third*, it is used as a confirmatory, hypothesis-testing tool aimed at testing prior hypotheses about the dimensional structure of a set of variables. And *fourth*, it is used to conceptualize the relationships of multiple indicators of latent variables in a causal modeling framework in which a factor model is assumed for the relationships between latent variables and their indicators. After a brief introduction to each of these four ways in which factor analytic tools are used in sociological research, this discussion covers the basic factor model and issues that arise in its application, either in the *exploratory* or *confirmatory* frameworks of analysis.

#### DATA REDUCTION APPROACHES

When a researcher wishes to build a composite score from a set of empirical variables, factor analysis and related techniques are often useful. Indeed, it is perhaps in this area of “index construction” that factor analysis is most often used by sociologists. There are various related data

reduction approaches that fall under the heading of “dimensional analysis” or “cluster analysis,” but the basic goal of all these techniques is to perform some decomposition of the data into sets of variables, each of which is relatively independent and homogeneous in content. When factor analysis is used in this way the researcher is essentially interested in determining the sets of linear dependence among a set of variables that are intended to measure the same general domain of content. The factor analysis of such variables may proceed in a number of different ways, but the basic goal is to determine the number of clusters of homogeneous content and the extent of relationship among various clusters or factors. Relationships among factors may be conceptualized either in terms of uncorrelated (or orthogonal) sets of factors or in terms of correlated (or oblique) factors. Such analyses are often supplemented with information on how to build “factor scores,” with item-analysis information, such as item-to-total score correlations, and with information estimating the “internal consistency” or “reliability” of such composite scores (see Greene and Carmines 1979; Maxwell 1971).

When using factor analysis and related techniques as a basis for index construction, one of two situations is typically the case. Either the investigator has some a priori basis for expecting that items in a set have a common factor (or common factors) underlying them, and therefore the investigator has certain well-founded expectations that the items can be combined into a scale, *or* the investigator has no a priori set of hypotheses for what clusters will be found and is willing to let the inherent properties of the data themselves determine the set of clusters. In the first case confirmatory factor models are appropriate, whereas in the second case exploratory methods are mandated. In either case the use of factor analysis as a data reduction tool is aimed at the development and construction of a set of “scores,” based on factor analysis, that can then be introduced as variables in research.

**Exploratory Factor Analysis.** As noted above, in situations where the researcher has no a priori expectations of the number of factors or the nature of the pattern of loadings of variables on factors, we normally refer to applications of factor analysis as exploratory. In the case of exploratory factor analysis the goal is to find a set of  $k$  latent

dimensions that will best reproduce the correlations among the set of  $p$  observed variables. It is usually desirable that  $k$  be considerably less than  $p$  and as small as possible. In exploratory factor analysis one typically does not have a clear idea of the number of factors but instead begins with uncertainty about what the data will reveal. The most common practice is to find  $k$  orthogonal (uncorrelated) dimensions that will best reproduce the correlations among the variables, but there is nothing intrinsic to the factor analytic model that restricts the conceptual domain to several orthogonal dimensions.

**Confirmatory Factor Analysis.** Confirmatory factor analysis, in contrast, refers to situations in which the investigator wishes to test some hypotheses regarding the structure of relationships in the presence of a strong set of assumptions about the number of factors, the values of the factor pattern coefficients, the presence or absence of correlations of factors, or other aspects of the model. In confirmatory factor analysis it is essential that one begin with a theory that contains enough detailed specification regarding constraints that should be imposed on the data in order to provide such a test, whereas in exploratory factor analysis there is no requirement that one specify the number of factors and expected relationships to be predicted in the data. Confirmatory approaches are thus more theory-driven, whereas exploratory approaches are more data-driven (see Alwin 1990). However, much of the so-called confirmatory factor analysis that is carried out in modern social and behavioral science is in fact exploratory, and much current research would be more realistically appraised if such confusion did not exist. Often, there is considerable tinkering with “confirmatory” models in order to improve their fit to the data, either by removing variables or by loosening up (or “freeing”) certain parameters. It is also often the case that the “confirmatory” factor analyses are actually preceded by an exploratory analysis, and then a confirmatory model based on these results is fit to the same data. Although very common, this approach “capitalizes” on chance and gives an illusory sense that one has confirmed (or verified) a particular model. Placed in the proper perspective, there is nothing in principal wrong with the approach, as long as the “test” of the model is cross-validated in other data.



**FACTOR ANALYSIS AND MULTIPLE INDICATOR CAUSAL MODELS**

In the 1960s and 1970s, with the introduction of causal modeling strategies in social science (see Blalock 1964; Duncan 1966, 1975; Heise 1968), a fundamental shift occurred in the nature and uses of the common factor models by sociologists. Methods and the logic of causal modeling with nonexperimental statistical designs had been around for a long time. Due largely to the influence of Lazarsfeld (1968), causal inference strategies had been prevalent especially among analysts of sample survey data since the 1940s, but the research strategies were based on tabular presentation of data and the calculation of percentage differences. In the early 1960s there was a general infusion of techniques of causal modeling in sociology and other social science disciplines. Path analysis, principal of these newly adopted techniques, of course, was invented before 1920 by the great geneticist Sewall Wright, but his contributions were not appreciated by social and behavioral scientists, including the psychometricians responsible for the development of factor analysis. Wright (1921) developed path models as *deductive* systems for deriving correlations of genetic traits among relatives of stated degree. He also used the method inductively to model complex economic and social processes using correlational data (Wright 1925).

Psychometricians, like Spearman, had been dealing with models that could be thought of as “causal models,” which could be understood in Wright’s path analysis framework—common factors were viewed as the causes underlying the observed variables—but Spearman and others who developed common factor models were unfamiliar with Wright’s work. None of the early psychometricians apparently recognized the possibility of causal relationships among the latent variables of their models, or for that matter among their indicators. However, with the publication of work by Jöreskog (1970) and others working in the “new” field of *structural equation models* (see Goldberger 1971, 1972; Goldberger and Duncan 1973; Hauser and Goldberger 1971); the convergence and integration of linear models in the path analysis tradition and those in the factor analysis tradition provided a basic “breakthrough” in one of the major analytic paradigms most prevalent in social science. These developments were assisted by the interest in conceptualizing measurement

errors within a causal analysis framework. A number of researchers began to incorporate conceptions of measurement error into their causal analyses (Alwin 1973a, 1974; Blalock 1965, 1969, 1970; Costner 1969; Duncan 1972; Heise 1969; Siegel and Hodge 1968), ushering in a new approach that essentially combined factor models and path models.

At about this same time, Karl Jöreskog and his colleagues were developing efficient procedures for estimating the parameters of such models—called LISREL models, named after Jöreskog and his colleagues’ computer program, LISREL—and this provided a major impetus for the widespread use of confirmatory approaches to the estimation of structural equation models. Jöreskog’s (1967) early contributions to maximum-likelihood factor analysis became readily applied by even the most novice of analysts. Unfortunately, the widespread availability of these techniques to researchers who do not understand them has led to serious risks of abuse. This can be true of any technique, including the techniques of exploratory factor analysis. In any event, the proper use and interpretation of the results of LISREL-type model estimation is a significant challenge to the present generation of data analysts.

**THE COMMON FACTOR MODEL**

The formal mathematical properties of the common factor model are well known and can be found in many of the accompanying references. It is useful for purposes of exposition briefly to review its salient features. Although the model can be most compactly represented using vector and matrix notation, it is normally best to begin with a scalar representation for the data, such as in equation 1.

$$\begin{aligned}
 z_1 &= a_{11}f_1 + a_{12}f_2 + \dots + a_{1k}f_k + u_1 \\
 z_2 &= a_{21}f_1 + a_{22}f_2 + \dots + a_{2k}f_k + u_2 \\
 &\dots \\
 z_p &= a_{p1}f_1 + a_{p2}f_2 + \dots + a_{pk}f_k + u_p
 \end{aligned}
 \tag{1}$$

Here the  $z$  variables are empirical quantities observed in a sample of units of observation (e.g., persons, groups, ecological units), and for present purposes the variables are standardized to have a mean of zero and standard deviation of unity. This

scaling is not a requirement of the model. In fact, in confirmatory factor models, it is often desirable to leave the variables in their original metric, especially when comparing the properties of these models across populations or subpopulations (see Alwin 1988b).

According to the common factor model, each observed  $z$  variable is, then, a linear function of a set of  $k$  latent or unobserved variables and a residual variable,  $ui$ , (also unobserved), which contains variation specific to that particular variable and random measurement error. The  $a$  coefficients in equation 1 are “factor loadings.” They reflect the linkage between the “observed” variables and the “unobserved” factors. In the case of uncorrelated factor these loadings equal the correlations of the variables with the factors. The loadings thus provide a basis for interpreting the factors in the model; factors obtain their meaning from the variables to which they are linked and vice versa. Thus, in many investigations the primary objective is to estimate the magnitudes of these factor loadings in order to obtain a meaningful understanding of the nature of the data.

The  $k$  latent variables are called *common* factors because they represent common sources of variation in the observed variables. As such, these common factors are thought to be responsible for covariation among the variables. The unique parts of the variables, by contrast, contribute to lack of covariation among the variables. Covariation among the variables is greater when they measure the same factors, whereas covariation is less when the unique parts of the variables dominate. Indeed, this is the essence of the model–variables correlate because they measure the same things. This was the basis of Spearman’s original reasoning about the correlations among tests of mental ability. Those tests correlated *because* they measured the same general factor. In the general case those variables that correlate do so because of their multiple sources of common variation.

Common variation in the aggregate is referred to as *communality*. More precisely, a variable’s communality is the proportion of its total variation that is due to its common sources of variation. The communality of variable  $i$  is denoted  $h^2_i$ . A variable’s *uniqueness*, denoted  $u^2_i$ , is the complement of the communality; that is,  $u^2_i = 1.0 - h^2_i$ . The

uniqueness is thought of as being composed of two independent parts, one representing *specific* variation and one representing *random measurement error* variation; that is,  $u^2_i = s^2_i + e^2_i$ . (This notation follows traditional psychometric factor analysis notation. Each of these quantities is a variance, and, thus, in the covariance modeling or structural equation modeling tradition these quantities would be represented as variances, such that  $\sigma^2_{ui} = \sigma^2_{ii} + \sigma^2_{ei}$  [see below].) Specific variance is *reliable* variance, and thus the reliability of variable  $i$  can be expressed as  $r^2_i = h^2_i + s^2_i$ . Unfortunately, because of the presence of specific variance in most variables, it is virtually impossible to use the traditional form of the common factor model as a basis for reliability estimation (see Alwin 1989; Alwin and Jackson 1979). The problem is that the common factor model typically does not permit the partitioning of  $u^2_i$  into its components,  $s^2_i$  and  $e^2_i$ . In the absence of specific variance, classical reliability models may be viewed as a special case of the common factor model, but in general it is risky to assume that  $e^2_i = u^2_i$ . Alwin and Jackson (1979) discuss this issue in detail. Some attempts have been made to augment the traditional latent “trait” model inherent in the common factor model by adding “method” factors based on the *multitrait–multimethod* design of measurement within the framework of confirmatory factor models (see Alwin 1974; Alwin and Jackson 1979; Werts and Linn 1970). This provides a partitioning of the specific variance due to method, but it does not provide a general solution to the problem of handling specific variance.

Returning to the above example (in equation 1), in Spearman’s case (the one-factor case) each variable contains a common factor and a specific factor, as shown in equation 2.

$$\begin{aligned} z_1 &= a_{11}f_1 + s_1 \\ z_2 &= a_{21}f_1 + s_2 \\ &\dots \\ z_p &= a_{p1}f_1 + s_p \end{aligned} \tag{2}$$

In this case  $h^2_i = a^2_i$  and  $u^2_i = s^2_i$ . Spearman’s (1927) theory in essence assumes perfect measurement, not unlike most common factor models. However, unlike researchers of today, Spearman was very

concerned about measurement errors, and he went to great lengths to correct his observed correlations for imperfections due to random errors of measurement (Spearman 1904b). Thus, when applied to such corrected correlational data, these assumptions may be appropriate.

As can be seen from the equations for Spearman's model (equation 2), the correlations among variables  $z_i$  and  $z_j$ ,  $r_{ij} = E[z_i z_j]$  (the expected value of the cross-products of the  $z$  scores for the two variables) may be written as  $r_{ij} = a_i a_j$ . For example, if  $p = 3$ , the correlations among the variables can be written as  $r_{12} = a_1 a_2$ ,  $r_{13} = a_1 a_3$ , and  $r_{23} = a_2 a_3$ . In vector notation (introduced in greater detail below), the common parts of the correlations among the variables of the model are composed of the matrix product  $AA'$ . In the case where  $p = 3$ , the matrix  $A$  is written as in equation 3,

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{3}$$

and the product  $AA'$  is written as in equation 4.

$$\begin{bmatrix} a_1 a_1 & a_1 a_2 & a_1 a_3 \\ a_2 a_1 & a_2 a_2 & a_2 a_3 \\ a_3 a_1 & a_3 a_2 & a_3 a_3 \end{bmatrix} \tag{4}$$

The variances of the variables are also affected by the common factors, but, as indicated in the foregoing, there is a residual portion of variance containing specific and unreliable variance. In Spearman's model the variance of variables  $i$  is as shown in equation 5.

$$\begin{aligned} \sigma_{z_i}^2 &= 1.0 = a_i^2 + \sigma_{u_i}^2 \\ &= a_i^2 + \sigma_{s_i}^2 + \sigma_{e_i}^2 \end{aligned} \tag{5}$$

Then it can be seen that the correlation matrix is equal to  $R = AA' + U^2$ , where the matrix  $U^2$  for the  $p = 3$  case is written in matrix form as in equation 6.

$$\begin{bmatrix} \sigma_{u_1}^2 & 0 & 0 \\ 0 & \sigma_{u_2}^2 & 0 \\ 0 & 0 & \sigma_{u_3}^2 \end{bmatrix} \tag{6}$$

These results have general applicability, as will be seen below.

### ESTIMATION AND TESTING OF THE FACTOR MODEL

Before proceeding to the more general uses of the model, it is important to review the logic behind Spearman's approach. In the general Spearman case, the correlation of two variables is equal to the product of their loadings on the general factor; that is,  $r_{ij} = a_i a_j$ . Recall that under this model the  $a$  coefficients represent the correlations of the variables with the factor. Spearman reasoned therefore that if the model were true (that is, if a single unobserved common factor could account for the correlations among the observed variables), then certain things had to hold in the empirical data.

Spearman reasoned that, if the single factor model holds, the partial correlation between any two variables, holding constant the underlying common factor,  $r_{ij \cdot f}$  should be zero. This stems from the fact that the numerator of this partial correlation,  $r_{ij} - r_{if} r_{jf}$ , is zero, because *under the model*  $r_{ij} - a_i a_j = a_i a_j - a_i a_j = 0$ . Of course, it is not possible to calculate such a partial correlation from the data because the factor score,  $f$ , does not exist except in the theory. Spearman, however, recognized a specific pattern to the components of the correlations *under the model*. He noted that, if the single factor model held for  $p = 4$ , the intercorrelations of the variables had to satisfy two independent conditions, referred to by Spearman (1927) as *vanishing tetrads*, shown in equation 7.

$$\begin{aligned} r_{12} r_{34} - r_{14} r_{23} &= 0 \\ r_{13} r_{24} - r_{14} r_{23} &= 0 \end{aligned} \tag{7}$$

Note that the case of  $p = 3$  is a trivial case, since a one-factor model can always be used to describe the intercorrelations among three variables. For  $p = 5$  there are  $[p(p - 1)(p - 2)(p - 3)]/8$  different tetrads (see Harman 1976), which equals fifteen. Not all of the possible tetrad differences formed from these fifteen are independent, and for one factor to explain the correlations, there are  $p(p - 3)/2$  independent tetrad differences. Thus, in the case of five variables there are five tetrad differences that must vanish, and for six there are nine, and so forth.

Although in recent years there has been a revival of interest in Spearman's vanishing tetrads for sets of four variables (Glymour et al. 1987), at

the time he developed this logic there was little that could be done computationally with very large problems. Thurstone (1947) developed the *centroid* method as an approximation to the principal axes approach involved in Spearman's early work, which was in common use during the 1940s and 1950s, but with the development of the high-speed computer, principal axes methods became (and remain) quite common in many applications of the model.

In exploratory factor analysis, where the number of factors of the model is not known beforehand, estimation is carried out by way of an eigenvalue/eigen-vector decomposition of some matrix, either  $R$  or some estimate of  $R - U^2$ . There is a wide variety of types of factor analyses that can be done—principal component factor analysis (which analyzes the  $p$  first nonzero components of  $R$ ), communality-based factor analysis (which analyzes  $R$  with a communality estimate in the diagonal), alpha factor analysis, canonical factor analysis, or image analysis (see Harris 1963, 1964). Few developments have been made in these approaches since the 1960s, although there continues to be considerable debate about the desirable properties of these various approaches (e.g., see Widaman 1991).

Perhaps the most important development affecting exploratory factor analysis since the 1960s has been the development of maximum-likelihood factor analysis. Maximum-likelihood estimation, however, requires the prior estimate of the number of factors. These methods are most often discussed in connection with confirmatory factor analysis, although the approach to exploratory factor analysis discussed by Lawley and Maxwell (1971) illustrates how a form of traditional exploratory factor analysis can be done by setting minimal constraints on the model and testing successive hypotheses about the number of factors. A discussion of these models occurs in a subsequent section on confirmatory factor analysis. Before this, a more formal presentation of the factor model in matrix form is given, along with a discussion of several of the longstanding problems that dominate the literature on factor analysis, specifically the problem of estimating communalities, the problem of estimating factor scores, and the problem of determining factorial invariance.

### THE FACTOR MODEL IN MATRIX NOTATION

We can generalize the model given above for the case of multiple factors,  $k$ , in matrix notation. And again, the factor model can be easily represented in terms of the data matrix at hand. (The model can also be written compactly in vector notation for populations of interest. This is the approach taken in the subsequent discussion of confirmatory factor analysis.) The data matrix in this case can be represented as a  $p$  by  $n$  array of variable scores. Let  $Z'$  symbolize this  $p$  by  $n$  data matrix. Using this notation, write the common factor model for a set of  $p$  variables as  $Z' = AF' + UW'$  where  $Z'$  is as defined above,  $A$  is a  $p$  by  $k$  *factor pattern matrix* (in the case of uncorrelated factors  $A$  is called the *factor structure matrix*),  $F'$  is a  $k$  by  $n$  matrix of hypothetical factor scores,  $U$  is a  $p$  by  $p$  diagonal matrix of unique-score standard deviations (defined such that the element  $u_i$  is the square root of the unique variances,  $\sigma^2 u_i$ ), and  $W'$  is a  $p$  by  $n$  matrix of hypothetical unique scores. Note that the factors (both common and unique) are never observed—they exist purely at the hypothetical level. Note also that because we have standardized the variables (the  $z$ 's) to be centered about the mean and to have standard deviations of unity, the factor scores in this model are theoretically standardized in the same fashion. In other words,  $E(F'F) = I_k$ , and the variances of the unique scores are equal to  $E(W'W) = U^2$ , assumed to be a diagonal matrix.

Traditionally, the factor model assumed that the factors of this model were uncorrelated, that the unique parts of the data (the  $W$ ) are uncorrelated with the common parts (the  $F$ ), and that the unique variation in variable  $i$  is uncorrelated with the unique variation in variable  $j$ , for all  $i$  and  $j$ . In matrix notation, the factor model assumes that  $E(F'F) = I_k$ ,  $E(W'W) = U^2$ , and  $E(F'W) = E(W'F) = 0$ . In other words, the factors of the model are uncorrelated with one another and have variances of unity. Also, the common factors and unique factors are uncorrelated, and the unique factors are uncorrelated among themselves.

This type of notation helps clarify the fact that factor analysis is in effect interested in the "reduced" data matrix,  $Z' - UW'$ , rather than  $Z'$ . Consequently, the factor model is concerned with

the decomposition of the matrix  $R - U^2$  (the correlation matrix with communalities in the diagonal) rather than  $R$  (the correlation matrix with unities in the diagonal), since in equation 8 we demonstrate the following:

$$\begin{aligned} R &= E(Z'Z) = E(AF' + UW')(AF' + UW')' \\ &= AA' + U^2 \end{aligned} \tag{8}$$

and  $R - U^2 = AA'$ .

This demonstrates an often misunderstood fact—namely, that factor analysis focuses on the reduced-correlation matrix,  $R - U^2$ , rather than on the correlation matrix (with 1s in the diagonal). As will be clarified below, this is the fact that differentiates factor analysis from principal components analysis—the latter operates on the correlation matrix. In factor analysis, then, one must begin with some estimate of  $I - U^2$  or  $H^2$ , the matrix of communalities, and then work on the decomposition of  $R - U^2$ . This poses a dilemma, since neither the common nor unique factors are observed, and it is therefore not possible to know  $U^2$  and  $H^2$  beforehand. The objective is to come up with an estimate of  $H^2$  that retains a positive semidefinite property to  $R - U^2$ . At the same time,  $H^2$  is one aspect of what one wants to discover from the analysis, and yet in order to estimate the model one must know this matrix beforehand. The solution to this problem is to begin with an “estimate” of the communalities of the variables, and then through an iterative procedure new estimates are obtained and the solution is reached through convergence to some criterion of fit.

### COMMUNALITY ESTIMATION AND THE NUMBER OF FACTORS

Determining the number of factors in exploratory factor analysis is one of the fundamental problems involved in arriving at a solution to the parameters of the factor model. The problem essentially involves determining the rank,  $k$ , of the matrix  $R - U^2$ , where these matrices are as defined above. Technically, we want to find a matrix  $U^2$  that will retain the property of positive semidefiniteness in  $R - U^2$  with the smallest possible rank (Guttman 1954). The rank of this matrix in this case is the minimum number of factors necessary to reproduce the off-diagonal elements of  $R$ . Thus, the problem of

determining  $k$  is closely related to the communality estimation problem, that is, determining an estimate for the diagonal of  $R - U^2$ , that is,  $H^2$ .

Guttman (1954) outlined the problem of deciding on the number of factors and compared three principles for estimating  $k$  via solutions to the communality estimation problem. He described a “weak lower bound” on the number of factors,  $k_1$ , as the nonnegative roots (eigen values) of the matrix  $R - I$ . This is equivalent to the number of roots of  $R$  greater or equal to unity, since  $R$  and  $R - I$  differ only by  $I$ , and their roots differ therefore only by unity. Guttman shows that  $k_1$  is a lower bound to  $k$ , that is,  $k \geq k_1$ . A second principle, one that also implies another approach to estimating communality, is based on the matrix  $R - D$ , where  $D$  is a diagonal matrix whose elements,  $1 - r_j^2$  ( $j = 1, p$ ), are the result of unity minus the square of the largest correlation of variable  $j$  with any of the  $p - 1$  other variables. Guttman shows that  $k_2$  is also a lower bound to  $k$ , such that  $k \geq k_2$ . A third and perhaps the most common approach to estimating communalities is based on the idea that the squared multiple correlation for each variable predicted on the basis of all the other variables in the model is the upper limit on what the factors of the model might reasonably explain. If we define the matrix  $R - C^2$ , where  $C^2$  is a diagonal matrix whose elements  $C_j^2$  ( $1, p$ ) are equal to  $1 - r_j^2$ , where  $r_j^2$  is the squared multiple correlation of variable  $j$  with the remaining  $p - 1$  variables. Guttman shows that  $k_3$  is also a lower bound to  $k$ . This third lower bound is often referred to as Guttman’s *strong* lower bound since he showed the following relationships among the lower bounds:  $k \geq k_3 \geq k_2 \geq k_1$ . In practice,  $k_1$  may be adequate but it could be wrong, and, if wrong, it is likely to be too small. The use of  $k_1$  is probably questionable in the general case. The use of  $k_2$  is obsolete and not practicable. It estimates communality in the way of the Thurstone centroid method, which is only a rough approximation to a least-squares approach. Perhaps the best solution is the choice of  $k_3$ . It is less likely to overlook common factors, as  $k_1$  might, since  $1 - s_j^2$  is a lower bound to  $h_j^2$ . It should be pointed out that the lower bounds  $k_1$  and  $k_3$  are what distinguish the two main approaches to factor analysis, namely an incomplete principal components decomposition (referred to as principal components factor analysis) and the principal factor method of analysis.



**FACTOR ANALYSIS VERSUS PRINCIPAL COMPONENTS ANALYSIS**

It was mentioned above that it is not understood well enough that factor analysis is concerned mainly with the matrix  $R - U^2$  rather than with  $R$ . This is in fact one of the things that distinguishes factor analysis from principal components analysis. However, the differences between the two are more fundamental. Factor analysis is based on a *model*, a particular theoretical view (hypothesis, if you like) about the covariance (or correlational) structure of the variables. This model states (as given above) that the correlation matrix for a set of variables can be partitioned into two parts—one representing the common parts of the data and one representing uniqueness—that is,  $R = A'A + U^2$ . The factor model states, first, that the off-diagonal elements of  $A'A$  equal the off-diagonal elements of  $R$  and, second, that the elements of  $U^2$  (a diagonal matrix) when added to the diagonal elements of  $A'A$  give the diagonal elements of  $R$ . Thus, the factor model posits a set of  $k$  hypothetical variables ( $k < p$ ) that can account for the interrelationships (or correlations) of the variables but not for their total variances.

In contrast to this, principal components is not a model in the same sense—it is best viewed as a method. It is one method for obtaining an initial approximation to the common factor model (see Guttman's weak lower bound, discussed above), but it is extremely important to distinguish such an "incomplete" principal components solution (one associated with the roots of  $R$  that are equal to or greater than unity) from the full-rank principal components decomposition of  $R$  (see Maxwell 1977).

Any square symmetric nonsingular matrix, for example,  $R = R'$ , can be written in the form  $R = QD^2Q'$ , where  $D^2$  is a diagonal matrix of order  $p$  containing eigen values ordered according to decreasing magnitude, and  $Q$  is a matrix of unit-length eigen vectors (as columns) associated with the eigen values.  $Q$  is an orthonormal matrix,  $Q'Q = I = QQ'$ . This model is referred to as the principal components decomposition of  $R$ . Typically, one either analyzes a correlation matrix with 1s in the diagonal, or a covariance matrix with variances in the diagonal, in the application of this decomposition. In this model the correlation matrix,  $R$ , is formed by a centered or deviation-score data matrix scaled so that the variables have variances of

unity. Let  $Z'$  be the  $p \times n$  data matrix, as above. Note that the expected value of  $Z'Z = QD^2Q'$ , since  $Z' = QDY'$ .

If the correlation matrix is of full rank, then there will be  $p$  columns in  $Q$ . This means that in this case the principal components model involves a transformation of  $p$  variables into a set of  $p$  orthogonal components. When the correlation matrix is singular, meaning that the rank of the matrix is less than  $p$ , the principal components decomposition is said to be incomplete, but from the point of view of factor analysis this is often irrelevant since it is the matrix  $R - U^2$  that is of interest to the factor analyst.

If  $P$  is an  $r \times p$  components matrix ( $P = QD$ ), and  $r = p$ , then it is well known that  $Y' = (P'P)^{-1}P'Z' = P^1Z'$ , where  $Y'$  is a set of  $r$  component scores,  $P$  is as defined above, and  $Z'$  is a  $p \times n$  data matrix involving  $p$  variables and  $n$  units of observation (e.g., persons, cities, social organizations). In other words, component scores (in contrast to factor scores) are directly calculable.

**THE ROTATION PROBLEM—CORRELATED VERSUS UNCORRELATED FACTORS**

Principal components are by definition *uncorrelated* with one another. The basic objective of the method is to obtain a set of  $p$  orthogonal (uncorrelated) new variables via a linear transformation of the  $p$  original variables. Factors are different. Factors may be uncorrelated, and in classical exploratory factor analysis one always begins with a set of uncorrelated factors, but in general this is not a requirement. Indeed, in exploratory factor analysis the factors one obtains are uncorrelated because of the nature of the methods used, but normally one performs a transformation or rotation of these factors to achieve a more pleasing representation for interpretation purposes.

Two types of rotations are available—those that preserve the uncorrelated nature of the factors, such as the varimax and quartimax rotations (see Kaiser 1958; Neuhaus and Wrigley 1954), and those that allow the factors to be correlated. The latter are called "oblique" rotations because they move the factors out of the orthogonal reference into a vector space that reduces the geometric

angles between them. Using either of these approaches, the basic goal of rotation is to achieve what Thurstone called *simple structure*, the principle that variables should simultaneously load highly on one factor and low on all other factors. These rotational approaches are relatively straightforward and discussed in all of the textbook descriptions of factor analysis.

### FACTORIAL INVARIANCE

Following Thurstone's (1938, 1947) discussions of factor analysis, students of the method have frequently been concerned with the problem of the correspondence between factors identified in separate studies or in subgroups of the same study. Using Thurstone's terminology, a concern with the correspondence of factors refers to the *invariance* of factors. The concern with factorial invariance has generated an array of methods for comparing factors (see Mulaik 1972). The most common approach to the problem involves the computation of an index of *factor similarity* for corresponding factors given estimates of a factor model using the same variables in two or more samples. The details of various strategies for estimating factor similarity will not be covered here, as descriptions can be found in a variety of factor analysis textbooks.

These approaches were developed primarily for results obtained from exploratory factor analysis, and it can be argued that the issues of factorial invariance can be more fruitfully addressed using the methods of confirmatory factor analysis (see Jöreskog 1971b; Lawley and Maxwell 1971). The technical aspects of these methods will not be reviewed here, as they have been explicated elsewhere (see Alwin and Jackson 1979, 1981). Suffice it to say that issues of factorial invariance can be phrased, not only with respect to the correspondence of the factor pattern coefficients (the *A* matrix) across populations, but also with respect to other parameters of the model as well, particularly the matrix of factor interrelationships (correlations and covariances) and the matrix of disturbance covariances.

It is perhaps useful in this context to raise a more general question regarding the nature of factorial invariance that is sought in the analysis of the factorial content of measures. In general there

is no consensus regarding whether stronger or weaker forms of invariance are necessary for comparisons across populations or subpopulations. Horn and associates (1983), for example, suggest that rather than the "equivalence" of factor structures across populations, weaker "configurational" forms of invariance are "more interesting" and "more accurate representations of the true complexity of nature." By contrast, Schaie and Hertzog (1985, pp. 83-85) argue that the condition of factorial invariance, that is, "the equivalence of unstandardized factor loadings across multiple groups," is critical to the analysis of differences among groups and developmental changes within groups.

Clearly, these represent extremes along a continuum of what is meant by the question of factorial invariance. On the one hand, for strict comparison of content across groups, it is necessary to have the same units of measurement, that is, invariance of metric. This requires the same variables measured across populations, and some would also argue that such invariance of metric requires that the relationships of the variables and the factors be equivalent across populations (see Jöreskog 1971a). On the other hand, if the same pattern of loadings seems to exist, it may be an example of misplaced precision to require equivalence in the strictest sense. Of course, the resolution of these issues has implications for other uses to which factor analysis is typically put, especially the construction of factor scores and the use of causal modeling strategies to compare substantive processes across groups.

### THE PROBLEM OF FACTOR SCORE ESTIMATION

Researchers using the common factor model as a data reduction tool typically engage in the estimation of such models in order to obtain scores based on the factors of the model, which can then be used to represent those factors in further research. As will be shown here, *factor scores* can never be computed directly (as in the case of *component scores*). Factor scores are always estimated, and, due to the nature of the factor model, "estimated factor scores" never correlate perfectly with the underlying factors of the model. An important

alternative to factor scores is what have come to be called “factor-based” scores, which are scores derived from the results of the factor analysis, using unit versus zero weightings for the variables instead of the factor score weights derived from one or another method of estimating factor scores. Factor-based scores, which are frequently more easy to justify and much more practical, typically correlate so highly with factor score estimates as to make one skeptical of the need for factor score estimates at all (see, e.g., Alwin 1973b).

However, it is important that factor analysts understand the nature of the factor score estimation problem, regardless of whether factor score estimates become any more appealing than the simpler and more stable factor-based scores. The factor score estimation problem can best be seen in terms of an interest in solving for the matrix  $F'$  in the above matrix representation of the common factor model,  $Z' = AF' + UW'$ . This can be done analytically, but, as will be seen, it is not possible to do so empirically because of the nature of the model. To solve for  $F'$  in this model we arrive at the following representation (without going through all of the necessary steps):  $F' = (A'A)^{-1} A' [Z' - UW']$ . The calculations implied by this expression cannot actually be carried out because one never knows  $W'$ . This is known as the “factor measurement problem,” which results in the fact that factor scores cannot be computed directly and must therefore be estimated.

The question, then, becomes whether it is possible to estimate factor scores in a manner that is useful, given what is known— $Z$ ,  $A$ , and  $U^2$ . Several approaches have been set forth for estimating the factors, all of which involve some transformation of the data matrix  $Z'$  into a set of  $k$  scores that vary in their properties (see Harris 1967; McDonald and Burr 1967). Most of these methods bear some resemblance to the analytic solution for  $F'$  above, but there are some technical differences. One of the most commonly misunderstood facts involved in the estimation of factor scores is that the factor pattern coefficient matrix,  $A$ , cannot be applied directly to the estimation of factors; that is,  $F'$  cannot be estimated by  $A'Z'$ . This, of course, should be clear from the above representation, but it is often used, probably due to ignorance of the more “correct” factor score estimation strategies.

There are four recognized strategies for estimating scores representing the common factors of the model, given  $Z$ ,  $A$ , and  $U^2$  (Alwin 1973b). All of these approaches are typically discussed for a model such as that discussed above, namely a set of uncorrelated factors scaled to have 0 means and standard deviations of 1. It is not the purpose of the present discussion to evaluate the properties of these various approaches to factor score estimation, but a brief summary can perhaps provide a guide to the technical literature on this topic. It is important to emphasize that *none* of these approaches produces factor score “estimates” that are perfectly correlated with the underlying factors of the model. Some of these approaches produce *univocal* score estimates, meaning that each factor score correlates *only* with the factors they are intended to measure and not with factors they are not intended to measure. Only one of the approaches produces a set of factor score estimates that reflect the property of uncorrelated factors with unit standard deviations. But it is difficult in practice to evaluate the desirability of any of the properties of factor score estimates.

#### METHODS OF CONFIRMATORY FACTOR ANALYSIS

Confirmatory factor analysis, unlike the methods of exploratory factor analysis, begins with prior knowledge regarding the number of factors and something about the nature of their relationships to the observed variables. In the absence of such knowledge, confirmatory factor analysis is not appropriate. In the typical situation of confirmatory factor analysis, then, the investigator begins with some specific theoretical hypotheses involving a model that can be tested against the data. Naturally, there is an extremely large number of such possible models, so it should be obvious that the techniques cannot easily be used to “search” for the best possible set of restrictions involving  $k$  factors (but see Glymour et al. 1987).

The bulk of this review has been devoted to the use of exploratory techniques of factor analysis. This imbalance is perhaps justified, given what is known within sociology about the common factors in our data. Exploratory factor analysis techniques are likely to be much more useful,



especially at a stage where less knowledge has been developed. And within a field like sociology, where there is a broad variety of competing concepts and paradigms, exploration of data may often be the most salutary strategy. There are, however, clear-cut instances where the use of confirmatory factor analysis techniques is in order, and the remainder of this discussion focuses on these situations.

Consider the following model for a set of  $p$  variables:  $y = v + \lambda \eta + \varepsilon$ , where  $v$  is a vector of location parameters or constants representing the origins of measurement of the  $p$  observed variables,  $\eta$  is a vector of  $k$  latent variables or factors, and  $\varepsilon$  is a vector of random disturbances for the  $p$  observed variables. The covariance properties associated with  $\eta$  and  $\varepsilon$  are basically the same as those discussed in the section on exploratory factor analysis for  $F$  and  $W$ , except that in general these are not required to be uncorrelated within the common factor set. And, of course, there is no restriction on the metric of the variables; that is, the  $p$  variables and  $k$  factors are not necessarily standardized to have 0 means and standard deviations of unity. The coefficient matrix,  $\lambda$ , is a matrix of regression coefficients relating the  $p$  observed variables to the  $k$  latent factors. In the case of a single population, one can easily consider the  $p$  variables centered (which would remove the vector of location constants) and scaled to have unit variance, but in the situation where one wants to compare populations neither of these constraints is probably desirable.

The use of these models requires the distinction between constrained and unconstrained parameters. Typically, one refers to parameters of the model as *fixed* if they are constrained to have a particular value, such as a factor loading of 0 or, in the case of the variance of a latent factor, a variance of unity. By contrast, the unknown parameters of the model, for example the  $\lambda$ s, are referred to as *free* parameters, which means that they are estimated in the model under the constraints specified for *fixed* parameters. Thus, one speaks of fixed or constrained parameters on the one hand and free or estimable parameters on the other. The major breakthrough in the use of this type of model was the development of computer programs that allow one to fix certain parameters of the model to known quantities while estimating

the free parameters under these constraints. The general approach also allows one to specify causal connections among the latent factors of the model, and it allows one to specify correlations among the errors in the variables and the errors in the equations connecting the latent factors.

Consider the factor model for the situation where there are  $p = 4$  variables and  $k = 2$  factors, with the pattern of factor pattern coefficients shown in equation 9, where the first two variables are believed to measure  $\eta_1$  and the third and fourth variables are said to measure  $\eta_2$ . This is the kind of situation described by Costner (1969), who developed an approach to confirmatory factor analysis using Spearman's tetrad differences. Of course, there are more efficient estimation strategies than those proposed by Costner. In any event, the point of this example is that the investigator begins not only with a specific number of factors in mind but also with a specific set of assumptions about the pattern of loadings.

$$\Lambda = \begin{bmatrix} \lambda_{11} & 0.0 \\ \lambda_{21} & 0.0 \\ 0.0 & \lambda_{32} \\ 0.0 & \lambda_{42} \end{bmatrix} \tag{9}$$

In the general case the covariances and correlations of the common factors of the model,  $E(\eta'\eta)$ , can be symbolized by  $\Psi$  (sometimes this matrix is denoted as  $\Phi$ , but there are many ways in which to symbolize these quantities), and the covariances of the disturbances (or errors) on the variables can be symbolized by  $\theta_c$ . Neither of these two matrices,  $\Psi$  and  $\theta_c$ , is required by the model to be diagonal, although in the simplest form of the confirmatory model,  $\theta_c$  is often assumed to represent a set of uncorrelated disturbances. In more complicated forms of the model, within constraints placed by the identification problem, both of these matrices can be nondiagonal. In either case, the model here is written with the assumption that the investigator has prior theoretical knowledge regarding the number of sources of common variation and that the  $\eta$  vector exhausts those sources.

Any application of this model requires that it be *identified*, which essentially means that there must be enough independent information within the covariance and correlation structure being

analyzed sufficient to solve for the unknown parameters of the model. In general, there need to be  $k^2$  constraints on a particular common factor model, that is, in the parameters in  $\lambda$  and  $\psi$ . Other constraints, of course, are possible. Space does not permit the discussion of these matters, but a detailed treatment of these issues can be found elsewhere (see Alwin 1988a; Alwin and Jackson 1979).

**The True-Score Models.** It can be shown that a well-known class of measurement models that form the basis for *classical test theory* (Lord and Novick 1968) can be specified as a special case of confirmatory factor analysis (see Alwin and Jackson 1979; Jöreskog 1971a). In brief, by placing particular constraints on the  $\lambda$  and  $\theta$ , matrices of the model, one can estimate the parameters of models that assume the measures are *parallel*, *tau-equivalent*, or *congeneric*. Of course, as indicated earlier, in order to use the common factor model in such a fashion, one must be reasonably sure that there is little or no specific variance in the measures. Otherwise, one runs the risk of confusing reliable specific variance with measurement error variance.

**Multitrait–Multimethod Models.** In addition to the application of confirmatory factor analysis to the estimation of classical true-score models, several attempts have been made to augment the traditional latent “trait” model, inherent in the classical model, by the addition of “method” factors based on the *multitrait–multimethod* design of measurement within the framework of confirmatory factor models (see Alwin 1974; Alwin and Jackson 1979; Werts and Linn 1970). This provides a partitioning of the specific variance due to method, but it does not provide a general solution to the problem of handling specific variance. While these models can be very useful for partitioning item-level variance into components due to trait, method, and error, they place relatively high demands on the measurement design. And while the required designs are relatively rare in practice, these models help sensitize the researcher to problems of correlated method error (see e.g., Costner 1969). Recent work in this area has shown that the multitrait-multimethod model can be fruitfully applied to questions of survey measurement quality, assessing the extent to which correlated measurement errors account for covariation among survey measures (see Alwin, 1997; Scherpenzeel, 1995).

**Multiple-Indicator, Multiple-Cause Models.**

One of the simplest forms of causal models involving latent common factor models is one in which a single latent endogenous variable having several indicators is determined by several perfectly having measured exogenous variables. Jöreskog (1974) and Jöreskog and Goldberger (1975) refer to this as a *multiple-indicator, multiple-cause* (MIMC) model. This kind of model has certain similarities to the canonical correlation problem (see Hauser and Goldberger 1971).

**Analysis of Change–Simplex Models.** One type of model that can be viewed as a confirmatory factor model, and is useful in analyzing change with respect to the latent common factors over time, falls under the rubric of *simplex* models (Jöreskog 1974). Such models are characterized by a series of measures of the same variables separated in time, positing a Markovian (lag-1) process to describe change and stability in the underlying latent variable. This model can be used in situations where there is a single variable measured over time (see Heise 1969) or in situations where there are multiple measures of the latent variable at each time (see Wheaton et al. 1977).

These models have proven to be valuable in the study of human development and change, especially applied to panel studies of individual lives over time (Alwin 1994, 1995a; Alwin and Krosnick 1991; Alwin et al. 1991; Asendorf 1992.); A related set of models in this area are latent growth curves, in which levels and trajectories of the growth in latent factors can be conceptualized and estimated (e.g., Karney and Bradbury 1995; McArdle 1991; McArdle and Anderson 1990; Willett and Sayer 1994). In such applications the focus is on the nature of growth processes and the correlates/predictors of individual change in latent factors over time. This is a natural extension of traditional methods of confirmatory factor analysis and causal modeling strategies within the context of longitudinal data.

**Causal Modeling of Factors.** If one obtains multiple measures of factors, for which common factor models are believed to hold, and the factors can be specified to be causally related, then it is possible to use confirmatory techniques to estimate the causal influences of factors on one another. Of course, one must be able to justify these

models strongly in terms of theoretical considerations, and there must be considerable prior knowledge (as in the use of confirmatory factor analysis) that justifies the specification of such measurement models. The logic involved in dealing with the linkage between observed and unobserved variables is essentially that involved in confirmatory factor analysis, while the logic applied in dealing with the causal linkages among factors is that involved in path analysis and structural equation modeling. The main point is that the parameters of models that essentially contain two parts—a measurement part specifying a model linking observed and latent variables and a structural model linking the latent variables—can be estimated within the framework of LISREL-type models. The measurement part can typically be viewed within the framework of confirmatory factor analysis, although in some cases an “induced variable” model is more appropriate (Alwin, 1988b).

### CONCLUSION

This review is designed to provide an overview of the major issues involved in the use of factor analysis as a research tool, including both exploratory and confirmatory techniques. There are several useful textbook discussions of factor analysis that will aid those who desire further study. Among these, the texts by Gorsuch (1984), Harman (1976), Mulaik (1972), McDonald (1985), and Lawley and Maxwell (1971) are some of the best sources on exploratory factor analysis. There are some recent texts offering instruction on the conceptual understanding and practical guidance in the use of these confirmatory factor analysis and causal modeling strategies (e.g., Bollen 1989; Loehlin 1992). The newest developments in the area involve advances in the analysis of categorical and ordinal data, statistical estimation in the presence of incomplete data, the provision of graphic interfaces for ease of specification of causal models. There are now several major competitors in the area of software packages that can be used to estimate many of the confirmatory factor models discussed here. Although these several packages offer largely comparable material, each offers a somewhat special approach. The LISREL approach to the analysis of covariance structures was first made available in the early 1970s. It is now in its

eighth version and offers many improvements to its use (Jöreskog and Sörbom 1996a). The LISREL8 program is now supplemented by PRELIS (Jöreskog and Sörbom 1996b), which provides data transformation capabilities, and by SIMPLIS (Jöreskog and Sörbom 1993), which provides a simple command language for formulating LISREL-type models. There are several alternatives to the LISREL approach, including EQS (Bentler 1992), AMOS (Arbuckle 1997), and Mplus (Muthén and Muthén 1998), among others, which provide many attractive features (see West, 1997).

(SEE ALSO: *Causal Inference Models; Measurement; Multiple Indicator Models; Validity*).

### REFERENCES

- Alwin, D. F. 1973a “Making Inferences from Attitude–Behavior Correlations.” *Sociometry* 36:253–278.
- 1973b “The Use of Factor Analysis in the Construction of Linear Composites in Social Research.” *Sociological Methods and Research* 2:191–214.
- 1974 “Approaches to the Interpretation of Relationships in the Multitrait-Multimethod Matrix.” In H. L. Costner, ed., *Sociological Methodology* 1973–74. San Francisco: Jossey-Bass.
- 1988a “Structural Equation Models in Research on Human Development and Aging.” In K. Warner Schaie, Richard T. Campbell, William Meredith, and Samuel C. Rawlings, eds., *Methodological Issues in Aging Research*. New York: Springer.
- 1988b “Measurement and the Interpretation of Coefficients in Structural Equation Models.” In J. S. Long, ed., *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research*. Beverly Hills, Calif.: Sage.
- 1989 “Problems in the Estimation and Interpretation of the Reliability of Survey Data.” *Quality and Quantity* 23:277–331.
- 1990 “From Causal Theory to Causal Modeling: Conceptualization and Measurement in Social Science.” In J. J. Hox and J. De-Jong. Gierveld, eds., *Operationalization and Research Strategy*. Amsterdam: Swets and Zeitlinger.
- 1994 “Aging, Personality and Social Change.” In D. L. Featherman, R. M. Lerner, and M. Perlmutter, eds., *Life-Span Development and Behavior*, vol. 12. Hillsdale, N.J.: Lawrence Erlbaum.
- 1995a “Taking Time Seriously: Social Change, Social Character Structure, and Human Lives.” In P.

- Moen, G. H. Elder, Jr., and K. Lüscher, eds., *Examining Lives in Context: Perspectives on the Ecology of Human Development*. Washington, D.C.: American Psychological Association.
- 1995b “Quantitative Methods in Social Psychology.” In K. Cook, G. Fine, and J. House, eds., *Sociological Perspectives on Social Psychology*. New York: Allyn and Bacon.
- 1997 “Feeling Thermometers Versus 7-Point Scales—Which Are Better?” *Sociological Methods and Research* 25:318–340.
- , R.L. Cohen, and T.M. Newcomb 1991 *Political Attitudes Over the Life-Span: The Bennington Women After 50 Years*. Madison: University of Wisconsin Press.
- , and D. J. Jackson 1979 “Measurement Models for Response Errors in Surveys: Issues and Applications.” In K. F. Schuessler, ed., *Sociological Methodology 1980*. San Francisco: Jossey-Bass.
- , 1981 “Applications of Simultaneous Factor Analysis to Issues of Factorial Invariance.” In D. J. Jackson and E. F. Borgatta, eds., *Factor Analysis and Measurement in Sociological Research*. Beverly Hills, Calif.: Sage.
- , and D. J. Jackson 1982 “Adult Values for Children: An Application of Factor Analysis to Ranked Preference Data.” In R. M. Hauser et al. eds., *Social Structure and Behavior*. New York: Academic Press.
- , and J. A. Krosnick. 1991 “Aging, Cohorts and the Stability of Socio-Political Orientations Over the Life-Span.” *American Journal of Sociology* 97:169–195.
- , and R. C. Tessler. 1974 “Causal Models, Unobserved Variables, and Experimental Data.” *American Journal of Sociology* 80:58–86.
- Arbuckle, J. L. 1997 *Amos Users’ Guide*. Version 3.6. Chicago: SmallWaters Corporation.
- Asendorpf J. B. 1992 “Continuity and Stability of Personality Traits and Personality Patterns.” In J. B. Asendorpf and J. Valsiner, eds., *Stability and Change in Development*. Newbury Park, Calif.: Sage.
- Bentler, P. M. 1983 “Some Contributions to Efficient Statistics for Structural Models: Specification and Estimation of Moment Structures.” *Psychometrika* 48:493–517.
- 1989 *EQS Structural Equations Program Manual, Version 3.0*. Los Angeles: BMDP Statistical Software.
- , and D. G. Weeks 1980 “Linear Structural Equations with Latent Variables.” *Psychometrika* 45:289–308.
- Blalock, H. M., Jr. 1964 *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- 1965 “Some Implications of Random Measurement Error for Causal Inferences.” *American Journal of Sociology* 71:37–47.
- 1969 “Multiple Indicators and the Causal Approach to Measurement Error.” *American Journal of Sociology* 75:264–272.
- 1970 “Estimating Measurement Error Using Multiple Indicators and Several Points in Time.” *American Sociological Review* 35:101–111.
- Bollen, K. A. 1989 *Structural Equations with Latent Variables*. New York: John Wiley and Sons.
- Browne, M. W. 1974 “Generalized Least Squares Estimators in the Analysis of Covariance Structure.” *South African Statistical Journal* 8:1–24.
- 1984 “Asymptotically Distribution-Free Methods for the Analysis of Covariance Structures.” *British Journal of Mathematical and Statistical Psychology* 37:62–83.
- , and A. Shapiro 1988 “Robustness of Normal Theory Methods in the Analysis of Linear Latent Variate Models.” *British Journal of Mathematical and Statistical Psychology* 41:193–208.
- Christofferson, A. 1975 “Factor Analysis of Dichotomized Variables.” *Psychometrika* 40:5–32.
- Costner, H. L. 1969 “Theory, Deduction, and Rules of Correspondence.” *American Journal of Sociology* 75:245–263.
- Duncan, O. D. 1966 “Path Analysis: Sociological Examples.” *American Journal of Sociology* 72: 1–16.
- 1972 “Unmeasured Variables in Linear Models for Panel Analysis.” In H. L. Costner, ed., *Sociological Methodology 1972*. San Francisco: Jossey-Bass.
- 1975 *Introduction to Structural Equation Models*. New York: Academic Press.
- Garnett, J. C. M. 1919–20 “On Certain Independent Factors in Mental Measurement.” *Proceedings of the Royal Society of London A*. 96:91–111.
- Glymour, C., R. Scheines, P. Spines, and K. Kelly 1987 *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modeling*. New York: Academic Press.
- Goldberger, A. S. 1971 “Economics and Psychometrics: A Survey of Communalities.” *Psychometrika* 36:83–107.
- 1972 “Structural Equation Models in the Social Sciences.” *Econometrika* 40:979–999.
- , and O. D. Duncan 1973 *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
- Gorsuch, R. L. 1984 *Factor Analysis*, 2d ed. Hillsdale, N.J.: Lawrence Erlbaum.

- Gould, S. J. 1981 *The Mismeasure of Man*. New York: Norton.
- Greene, V. L., and E. G. Carmines 1979 "Assessing the Reliability of Linear Composites." In K. F. Schuessler, ed., *Sociological Methodology 1980*. San Francisco: Jossey-Bass.
- Guttman, L. 1954 "Some Necessary Conditions for Common Factor Analysis." *Psychometrika* 19:149-161.
- 1955 "The Determinacy of Factor Score Matrices with Implications for Five Other Basic Problems of Common-Factor Theory." *British Journal of Statistical Psychology* 8:65-81.
- 1956 "'Best Possible' Systematic Estimates of Communalities." *Psychometrika* 21:273-285.
- 1960 "The Matrices of Linear Least-Squares Image Analysis." *British Journal of Statistical Psychology* 13:109-118.
- Harman, H. H. 1976 *Modern Factor Analysis*. Chicago: University of Chicago Press.
- Harris, C. W. 1963 "Canonical Factor Models for the Description of Change." in C. W. Harris, ed., *Problems in Measuring Change*. Madison: University of Wisconsin Press.
- 1964 "Some Recent Developments in Factor Analysis." *Educational and Psychological Measurement* 24:193-206.
- 1967 "On Factors and Factor Scores." *Psychometrika* 32:193-379.
- Hauser, R. M. 1972 "Disaggregating a Social Psychological Model of Educational Attainment." *Social Science Research* 1:159-188.
- , and A. S. Goldberger 1971 "The Treatment of Unobservable Variables in Path Analysis." In H. L. Costner, ed., *Sociological Methodology 1971*. San Francisco: Jossey-Bass.
- Heise, D. R. 1968 "Problems in Path Analysis and Causal Inference." In E. F. Borgatta, ed., *Sociological Methodology 1969*. San Francisco: Jossey-Bass.
- 1969 "Separating, Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34:93-101.
- 1972 "Employing Nominal Variables, Induced Variables, and Block Variables in Path Analysis." *Sociological Methods and Research* 1:147-174.
- , and G. W. Bohrnstedt 1970 "The Validity, Invalidity, and Reliability of a Composite Score. In E. F. Borgatta and G. W. Bohrnstedt, eds., *Sociological Methodology 1970*. San Francisco: Jossey-Bass.
- Horn, J. L., J. J. McArdle, and R. Mason 1983 "When Is Invariance Not Invariant?: A Practical Scientist's Look at the Ethereal Concept of Factor Invariance." *The Southern Psychologist* 1:179-188.
- Hotelling, H. 1933 "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of Educational Psychology* 24:417-441, 498-520.
- Jöreskog, K. G. 1966 "Testing a Simple Structure Hypothesis in Factor Analysis." *Psychometrika* 31: 165-178.
- 1967 "Some Contributions to Maximum-Likelihood Factor Analysis." *Psychometrika* 32:443-482.
- 1970 "A General Method for Analysis of Covariance Structures." *Biometrika* 56:239-251.
- 1971a "Simultaneous Factor Analysis in Several Populations." *Psychometrika* 36:409-426.
- 1971b "Statistical Analysis of Sets of Congeneric Tests." *Psychometrika* 36:109-133.
- 1973 "A General Method for Estimating a Linear Structural Equation System." In A. S. Goldberger and O. D. Duncan, eds., *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
- 1974 "Analyzing Psychological Data by Structural Analysis of Covariance Matrices." In D. H. Krantz, et al., eds., *Measurement, Psychophysics, and Neural Information Processing*. San Francisco: Freeman.
- , and A. S. Goldberger 1975 "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable." *Journal of the American Statistical Association* 70:631-639.
- , and D. N. Lawley 1968 "New Methods in Maximum-Likelihood Factor Analysis." *British Journal of Mathematical and Statistical Psychology* 21:85-96.
- , and D. Sörbom 1986 *LISREL: Analysis of Linear Structural Relationships by the Method of Maximum-Likelihood*. User's Guide. Version 6. Chicago: Scientific Software.
- , and D. Sörbom 1988 *PRELIS: A Program for Multivariate Data Screening and Data Summarization (A Preprocessor for LISREL)*. Version 1.8, 2d ed. Chicago: Scientific Software.
- , and D. Sörbom 1996a *LISREL8: User's Reference Guide*. Chicago: Scientific Software International.
- , and D. Sörbom 1996b *PRELIS2: User's Reference Guide*. Chicago: Scientific Software International, Inc. 7383 N. Lincoln Avenue. Chicago IL 60646-1704.
- , and D. Sörbom 1993. *LISREL8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International.
- Kaiser, H. F. 1958 "The Varimax Criterion for Analytic Rotation in Factor Analysis." *Psychometrika* 23:187-200.

- 1963 “Image Analysis.” In C. W. Harris, ed., *Problems in Measuring Change*. Madison: University of Wisconsin Press.
- Karney, B. R., and T. N. Bradbury 1995 “Assessing Longitudinal Change in Marriage: An Introduction to the Analysis of Growth Curves.” *Journal of Marriage and the Family* 57:1091–1108.
- Lawley, D. N., and A. E. Maxwell 1971 *Factor Analysis as a Statistical Method*. London: Butterworths.
- Lazarsfeld, P. F. 1968 “The Analysis of Attribute Data.” In D. L. Sills, ed., *International Encyclopedia of the Social Science*, vol. 15. New York: Macmillan and Free Press.
- Loehlin, J. C. 1992 *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*. Hillsdale, N.J.: Lawrence Erlbaum.
- Lord, F. M., and M. L. Novick 1968 *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- Maxwell, A. E. 1971 “Estimating True Scores and Their Reliabilities in the Case of Composite Psychological Tests.” *British Journal of Mathematical and Statistical Psychology* 24:195–204.
- 1977 *Multivariate Analysis in Behavioral Research*. London: Chapman and Hall.
- McArdle, J.J. 1991 “Structural Models of Developmental Theory in Psychology.” *Annals of Theoretical Psychology* 7:139–160.
- McArdle, J.J., and E. Anderson. 1990 “Latent Variable Growth Models for Research on Aging.” In J.E. Birren and K.W. Schaie, eds., *Handbook of the Psychology of Aging*, 3rd ed. New York: Academic Press.
- Meredith, W. 1964a “Notes on Factorial Invariance.” *Psychometrika* 29:177–185.
- 1964b “Rotation to Achieve Factorial Invariance.” *Psychometrika* 29:187–206.
- Mulaik, S. A. 1972 *The Foundations of Factor Analysis*. New York: McGraw-Hill.
- 1975 “Confirmatory Factor Analysis.” In D. J. Amick and H.J. Walberg, eds., *Introductory Multivariate Analysis for Educational, Psychological, and Social Research*. Berkeley: McCutchan Publishing Corp.
- Muthén, B. 1978 “Contributions to Factor Analysis of Dichotomous Variables.” *Psychometrika* 43:551–560.
- 1983 “Latent Variable Structural Equation Modeling with Categorical Data.” *Journal of Econometrics* 22:43–65.
- 1988 *LISCOMP—Analysis of Linear Structural Equations with a Comprehensive Measurement Model: A Program for Advanced Research*. Version 1.1. Chicago: Scientific Software.
- , and B.O. Muthén 1998 *Mplus—The Comprehensive Modeling Program for Applied Researchers*. User’s Guide Version 1.0. Los Angeles: Muthen and Muthen.
- Neuhaus, J. O., and C. Wrigley 1954 “The Quartimax Method: An Analytical Approach to Orthogonal Simple Structure.” *British Journal of Statistical Psychology* 7:81–91.
- Schaie, K. W., and C. Hertzog 1985 “Measurement in the Psychology of Adulthood and Aging.” In J. E. Birren and K. W. Schaie, eds., *Handbook of the Psychology of Aging*. New York: Van Nostrand.
- Scherpenzeel, A. 1995 “A Question of Quality: Evaluating Survey Questions By Multitrait–Multimethod Studies.” Unpublished doctoral dissertation, University of Amsterdam, The Netherlands.
- Siegel, P. M., and R. W. Hodge 1968 “A Causal Approach to Measurement Error.” In H. M. Blalock, Jr. and A. B. Blalock, eds., *Methodology in Social Research*. New York: McGraw-Hill.
- Spearman, C. 1904a “‘General Intelligence,’ Objectively Determined and Measured.” *American Journal of Psychology* 15:201–293.
- 1904b “The Proof and Measurement of Association between Two Things.” *American Journal of Psychology* 15:88–103.
- 1927 *The Abilities of Man*. New York: Macmillan.
- Thurstone, L. L. 1931 “Multiple Factor Analysis.” *Psychological Review* 38:406–427.
- 1938 *Primary Mental Abilities*. Psychometric Monographs, no. 1. Chicago: University of Chicago Press.
- 1947 *Multiple Factor Analysis*. Chicago: University of Chicago Press.
- Werts, C. E., and R. L. Linn 1970 “Path Analysis: Psychological Examples.” *Psychological Bulletin* 74:194–212.
- Wheaton, B., et al. 1977 “Assessing Reliability and Stability in Panel Models.” In D. R. Heise, ed., *Sociological Methodology 1977*. San Francisco: Jossey-Bass.
- West, Joel 1997 “Structural Equation Software.” <http://www.gsm.uci.edu/~joelwest/SEM/Software.htm>.
- Widaman, K. 1991 “Common Factor Analysis vs. Component Analysis: Differential Bias in Representing Model Parameters.” Department of Psychology, University of California, Riverside. Typescript.
- Wiley, D. E., and J. A. Wiley 1970 “The Estimation of Measurement Error in Panel Data.” *American Sociological Review* 35:112–117.
- Willett, J. B., and A. G. Sayer 1994 “Using Covariance Structure Analysis to Detect Correlates and Predictors of Individual Change Over Time.” *Psychological Bulletin* 116:363–381.

West, Joel 1997 "Structural Equation Software." <http://www.gsm.uci.edu/~joelwest/SEM/Software.htm>.

Wright, S. 1921 "Systems of Mating." *Genetics* 6:111-178.

—1925 *Corn and Hog Correlations*. U.S. Dept. of Agriculture, Bulletin no. 1300L1-60. Washington, D.C.: U.S. Government Printing Office.

DUANE F. ALWIN

## FAMILISM

See Filial Responsibility; Family and Household Structure; Intergenerational Relations; Intergenerational Resource Transfers.

## FAMILY AND HOUSEHOLD STRUCTURE

The family system of the United States is often characterized as consisting of nuclear-family households—that is, households consisting of no more than the parent(s) and dependent children, if any (Lee 1999). This is certainly true of the great majority of family households. In fact, there has never been a point in American history in which extended-family households predominated statistically (Ruggles 1994a; Seward 1978). In 1997 only about 4.1 percent of all families in the United States were "related subfamilies"—a married couple or single parent with children living with a related householder (U.S. Bureau of the Census 1998, Table 69). However, an analysis of census data from 1970 through 1990 by Glick and colleagues (1997) showed that the percentage of all households containing nonnuclear kin increased from 9.9 percent in 1980 to 12.2 percent in 1990, reversing a nearly century-long pattern of decline. In 1910 about 20 percent of the households of white families and 24 percent of those of black families contained nonnuclear kin (Ruggles 1994b). Apparently we have seen a long-term decline in the prevalence of extended-family households, very slightly counterbalanced by an increase in the 1980s; what happened in the 1990s is not yet known.

Not all of the of the households that do not contain extended families consist of the stereotypical

nuclear family of two parents and their dependent children, however. There is great diversity among American families and households, and this diversity is increasing. Table 1 presents data on the composition of American households and families from 1960 to 1998. Even over this relatively brief period, substantial changes are apparent. The average size of both households and families decreased dramatically from 1960 to 1990, although they have both been stable in the 1990s. Many fewer households contain families and married couples in the late 1990s than in 1960, while the proportion of nonfamily households has more than doubled and the proportion of single-person households has nearly doubled. Female householders have increased substantially as a proportion of both all households and all families.

There are many factors responsible for these changes. To understand them, changes in marriage rates and age at marriage, divorce and remarriage rates, rates of nonmarital cohabitation, the departure of children from their parents' homes, and the predilection of unmarried persons to live alone will be briefly examined. Each of these factors has affected family and household structure.

Marriage rates have declined considerably since 1960 (see Table 2). This is not readily apparent from the "crude" marriage rate (the number of marriages per 1,000 population) because this rate does not take the marital status or age distributions of the population into account. The crude marriage rate was artificially low in 1960 because, as a result of the postwar baby boom, a large proportion of the population consisted of children too young to marry. The rates per 1,000 unmarried women (for both ages 15 and over and ages 15 to 44) show the frequency of occurrence of marriage for persons exposed to the risk of marriage, and here there is clear evidence of decline. Some of this, however, is attributable to increases in the median age at first marriage, which declined throughout the twentieth century until about 1960, but has been increasing rapidly since 1970. As age at marriage increases, more and more people temporarily remain unmarried each year, thus driving the marriage rate down. The best evidence (Oppenheimer et al. 1997) indicates that a major cause of delayed marriage is the deteriorating economic circumstances of young men since the

FAMILY AND HOUSEHOLD STRUCTURE

**Changes in U.S. Households and Families Since 1960**

	1960	1970	1980	1990	1998
Number of households (in 1,000s)	52,799	63,401	80,776	93,347	102,528
Average size	3.35	3.14	2.76	2.63	2.62
Family households (%)	85.04	81.15	73.72	70.80	69.13
Married-couple households (%)	74.34	70.54	60.79	56.04	52.97
Female householder (%)	8.38	8.67	10.77	11.66	12.34
Nonfamily households (%)	14.95	18.84	26.27	29.19	30.86
Single-person households (%)	13.10	17.11	22.65	24.63	25.67
Number of families (in 1,000s)	45,111	51,586	59,550	66,090	70,880
Average size of families	3.67	3.58	3.29	3.17	3.18
Married-couple families (%)	87.18	86.75	82.47	79.16	76.63
Female householder (%)	9.99	10.83	14.61	16.47	17.81

Table 1

SOURCE: U.S. Bureau of the Census (1998; Table 69), and Current Population Reports, P20-509 and earlier reports.

1970s. Perhaps the improving economy of the later 1990s will eventually produce some change in this trend.

The rising divorce rate has also contributed greatly to the declining proportion of married-couple households and the increases in female householders and single-person households. The crude divorce rate rose from 2.2 per 1,000 in 1960 to 5.2 in 1980 (reaching peaks of 5.3 in both 1979 and 1981) but has declined modestly since then to 4.3 in 1996. The rate of divorce per 1,000 married women 15 and older followed a similar pattern, reaching a high of 22.6 in 1980 and declining to 19.5 in 1996. Some of this decline is illusory, because the large baby boom cohorts are aging out of the most divorce-prone years (Martin and Bumpass 1989). However, although the divorce rate remains high, it has not been increasing since 1980.

Sweeney (1997) notes that, for recent cohorts, about half of all marriages have involved at least one previously married partner. However, rates of remarriage after divorce have been declining steadily. Annual remarriage rates were 204.5 per 1,000 divorced men and 123.3 per 1,000 divorced women in 1970; by 1990 they had decreased to 105.9 for men and 76.2 for women (U.S. Bureau of the Census 1998).

Decreasing marriage and remarriage rates and increasing divorce rates have combined to produce increases in single-person and single-parent households. This trend is mitigated, however, by

the increasing prevalence of nonmarital heterosexual cohabitation. Most such unions (those without children) appear as nonfamily households in Table 1. Evidence from the National Survey of Families and Households (Bumpass 1994; Waite 1995) shows that, in the early 1990s, nearly one-quarter of all unmarried adults aged 25 to 29 were cohabiting. This percentage declines with age, but still exceeded 20 percent for those in their late thirties. The National Survey of Family Growth found that, in 1995, more than 41 percent of all women aged 15 to 44 had cohabited or were currently cohabiting (National Center for Health Statistics 1997). Of course many of the women who had not cohabited at the time of the survey will do so in the future. The best estimates suggest that more than half of all couples who marry now cohabit prior to marriage; further, about 60 percent of all cohabiting unions eventuate in marriage (Bumpass 1994; Bumpass et al. 1991).

To a considerable extent the increase in cohabitation has offset the decline in marriage. This is particularly the case among blacks, for whom the decrease in marriage rates over the past several decades has been much more precipitous than it has been for whites (Raley 1996; Waite 1995). Although cohabiting unions are less stable than marriages, ignoring cohabitation results in substantial underestimates of the prevalence of heterosexual unions in the United States.

In spite of the increase in cohabitation, changes in marriage and divorce behavior have had substantial effects on household and family structure



FAMILY AND HOUSEHOLD STRUCTURE

**Changes in U.S. Marriage and Divorce Since 1960**

	1960	1970	1980	1990	1996
Number of marriages (in 1,000s) <sup>1</sup>	1,523	2,159	2,390	2,443	2,344
Rate per 1,000 population	8.5	10.6	10.6	9.8	8.8
Rate per 1,000 unmarried women, age 15+	73.5	76.5	61.4	54.5	49.7
Rate per 1,000 unmarried women, age 15–44	148.0	140.2	102.6	91.3	81.5
Median age at first marriage <sup>2</sup>					
Males	22.8	23.2	24.7	26.1	27.1
Females	20.3	20.8	22.0	23.9	24.8
Divorce and annulments (in 1,000s) <sup>3</sup>	393	708	1,189	1,182	1,150
Rate per 1,000 population	2.2	3.5	5.2	4.7	4.3
Rate per 1,000 married women 15+	9.2	14.9	22.6	20.9	19.5

Table 2

NOTE: 1. Monthly Vital Statistics Report, Vol. 43, No. 12 (S), July 14, 1995 for 1960–90; and U.S. Bureau of the Census (1988): Table 156 for 1996.

2. U.S. Bureau of the Census, Current Population Reports, P20–514 (March 1998) and earlier reports for 1960–90; and U.S. Bureau of the Census web site for 1996.

3. Monthly Vital Statistics Report, Vol. 43, No. 9 (S), March 22, 1995 for 1960–90; and U.S. Bureau of the Census (1998): Table 156 for 1996.

in the United States over the past four decades. Fewer people are marrying, those who marry are doing so at later ages, more married people are divorcing, and fewer divorced people are remarrying. This means that Americans are living in smaller households than they did in 1960, but there are more of them. The rate of growth in the number of households has substantially exceeded the rate of growth in the number of families. Referring back to Table 1, from 1960 to 1998 the number of households increased by more than 94 percent, while the increase in the number of families was only about 57 percent. Over the same time period, the total population of the United States increased by just under 50 percent (U.S. Bureau of the Census 1998, Table 2). Our population, therefore, is distributed in a larger number of smaller households than was the case in 1960.

One cause of the decline in household size shown in Table 1 is decreased fertility. The fertility rate (number of births per 1,000 women aged 15 to 44) was 118.0 in 1960; by 1997 it had decreased to 65.0, although most of the decrease occurred prior to 1980 (National Center for Health Statistics 1999). The trend toward smaller households and families is reflective to some extent of decreases in the number of children per family.

A larger cause of the decrease in household size, however, is the proliferation of single-person

households, as shown in Table 1. Single-person households consist of three types of persons: the never-married, who are primarily young adults; the divorced and separated without coresident children, who are primarily young and middle-aged; and the widowed who live alone, who are primarily elderly. Each of these types has increased, but for somewhat different reasons. Each must therefore be examined separately.

As shown in Table 2, average ages at marriage have risen markedly since 1960, and the percentage of young adults who have never married has increased proportionately (Waite 1995). This has been accompanied by a long-term decline (since prior to World War II) in the average age of leaving the parental home (Goldscheider 1997). Prior to 1970 most of this decline was driven by decreasing ages at marriage, but since then it has reflected an increasing gap between leaving the family of orientation and beginning the family of procreation. More young adults are living independently of both parents and spouses. Some of them are cohabiting, of course, but increasing numbers are residing in either single-person or other nonfamily households (Goldscheider 1997; White 1994).

Since about 1970 there has been some increase in the proportion of young adults who live with their parents. This marks the reversal of a

long-term decline in age at leaving home (White 1994). This is, in part, a by-product of increasing age at marriage. However, decreases in exits from parental homes to marriage have been largely offset by increases in exits to independent living, so this recent increase in young adults living with parents is actually very small (Goldscheider 1997). On the other hand, there is increasing evidence that the process of launching children has become much more complex than in previous years. Goldscheider (1997) also shows that the proportion of young adults who return to their parents' homes after an initial exit has more than doubled from the 1930s to the 1990s; increases have been particularly striking since the early 1960s. This is a response, in part, to the rising divorce rate, but also an indication that it has gotten increasingly difficult for young adults, particularly young men, to make a living (Oppenheimer et al. 1997). Nonetheless, the proportion of young adults living independently of both parents and spouses continues to increase, contributing to the prevalence of nonfamily households.

The increase in divorce and decrease in remarriage have contributed to the rise in single-person households, as formerly married persons establish their own residences and, increasingly, maintain them for longer periods of time. They have also contributed to the rise in family households that do not contain married couples. As shown in Table 1, families headed by females (without husband present) increased from 10 percent of all families in 1960 to nearly 18 percent in 1998. Families headed by males (without wife present) also increased, from 2.8 percent of all families in 1960 to 5.5 percent in 1998. Among families with children under 18 in 1998, 20 percent were headed by women without spouses and 5 percent by men without spouses (U.S. Bureau of the Census 1998, Table 70).

As a consequence of these changes plus the rise in nonmarital childbearing, the proportion of children under 18 living with both parents decreased from 88 percent in 1960 to 68 percent in 1997 (U.S. Bureau of the Census 1998, Table 84). In addition, there is a large race difference in the living arrangements of children. Only 35 percent of black children lived with both parents in 1997, compared to 75 percent of white children. More than half (52 percent) of all black children lived

with their mothers only, as did 18 percent of white children. Further, 8 percent of black children and 3 percent of white children lived with neither parent. Some of these children are living with, and being cared for by, their grandparents (Pebbley and Rudkin 1999). This raises the issue of the living arrangements of older persons.

A somewhat longer perspective is necessary to observe changes in the living arrangements of older persons. Ruggles (1994a) has shown that, in 1880, nearly 65 percent of all elderly whites and more than 57 percent of all elderly nonwhites lived with a child. Since about 18 percent of all older persons had no living children, Ruggles estimates that about 78 percent of whites and 70 percent of nonwhites who had children lived with a child. By 1980 the percentages living with children had decreased to 16 for whites and 29 for nonwhites. There is little evidence of major changes in the proportion living with children since 1980. Further, Ruggles (1996) found that only 6 percent of all elderly women and 3 percent of elderly men lived alone in 1880. By 1997 the percentages living alone had increased to 41 for women and 17 for men (U.S. Bureau of the Census 1998, Table 50). The growth of single-person households among older people has been particularly rapid since about 1940.

Two sets of factors appear to be primarily responsible for the "migration" of older people from typically sharing households with their children in the late nineteenth century to living alone (or with their spouses only) in the late twentieth century. First, the family life cycle was quite different in 1900 than today. People married a bit later (and markedly later than in the 1960s and early 1970s), had more children, and had children later in life. Consequently, a significant proportion of people in their sixties had unmarried children who simply had not yet left the parental home. Ruggles (1994a) shows that, in 1880, about 32 percent of all unmarried elders and 57 percent of the married resided with a never-married child. Of course many of these children may have remained home precisely in order to care for their aging parents. Unmarried elders were more likely to live with married children.

Second, economic factors played a major role. Social Security did not exist until 1940. In 1900, 85

percent of all men between the ages of 65 and 69 were in the labor force, as were 49 percent of all men 85 and over (Smith 1979). However, this option was much less available to women; the comparable proportions in the labor force were 12 and 6 percent. Many older persons, particularly women, had no means of support other than their children. Rates of coresidence of aging parents with their adult children have decreased as the prosperity of the elderly has increased; more can now afford to live independently.

However, Ruggles (1994a) found that wealthier older people were more likely to share a household with children than were poorer elders in the nineteenth century, and the majority of multigenerational families lived in households headed by the elderly parent(s). These facts suggest that adult children benefited economically from coresidence and that the possibility of inheriting a farm or business from aging parents may have motivated many adults to coreside with parents. Today coresidence is more common among poorer than wealthier people (Ruggles 1994a, 1996).

As of March 1998, 41 percent of all women aged 65 and older lived alone, as did 17 percent of all older men. These percentages increase to 53 percent and 22 percent for women and men, respectively, for those age 75 and over (U.S. Bureau of the Census Web site). The reason for this large gender difference, of course, is the difference in marital status between men and women. Among men 75 and over, nearly two-thirds are married and less than one-quarter are widowed; among women these figures are almost exactly reversed. According to 1980 census data, the proportion of all elderly persons living alone increases from 22 percent among those 65 to 69 to more than 41 percent in the 85–89 age category, then drops to 33 percent for those 90 and over (Coward et al. 1989), after which the modal category becomes living with children. Older persons who have lost their spouses through death are clearly exhibiting a tendency to live alone as long as possible, which for many of them extends into the latest years of life.

Older persons now constitute nearly 13 percent of the total population of the United States, compared to about 4 percent in 1900. With so many of them maintaining their own residences,

either with their spouses or alone following widowhood, their contribution to the proliferation of small and single-person households is substantial.

If so many older persons lived with their children in the late nineteenth century, why were there so few extended-family households? Ruggles (1994a) shows that just under 20 percent of the households of whites contained extended families in both 1880 and 1900; this compares to less than 7 percent in 1980, but it was still very much a minority statistical pattern. There were three primary reasons. First, because of more limited life expectancies and relatively high fertility rates, there were proportionally few older people in the population, so where they lived made less difference to the nation's household structure. Second, as noted above, many older persons lived with an unmarried child; unless other relatives are present, this arrangement constitutes a nuclear-family household regardless of the age of the parent. Third, while these cohorts of older persons typically had many children (an average of 5.4 per woman in 1880), these children did not live together as adults, so older persons could live with only one; their remaining children lived in nuclear families. Ruggles (1994a) estimates that more than 70 percent of all elders who could have lived with a child actually did so in 1880; the comparable percentage in 1980 was 16. In comparison to the last century, older persons today are much less likely to live with children and much more likely to live alone, contributing to the proliferation of small and single-person households.

To this point, factors that have contributed to long-term decreases in household and family size, and consequent increases in the numbers of households and families, have been elucidated. There is evidence of changes in these directions in all age segments of the population. These trends do not mean, however, that more complex family households are not part of the contemporary American experience.

As noted at the beginning of this entry, the United States has never been characterized by a statistical predominance of extended-family households, although it appears that the preference was for intergenerational coresidence in the form of stem families (families containing an older parent or parents and one of their married children) until

the early years of the twentieth century. But extended family households do occur today. At any single point in time, they constitute less than 10 percent of all households (Glick et al. 1997; Ruggles 1994a). However, a dynamic perspective presents a somewhat different picture.

Beck and Beck (1989) analyzed the household compositions of a large sample of middle-aged women who were followed from 1969 to 1984. The presence of nonnuclear kin in their households was noted for specific years and was also calculated for the entire fifteen-year period. In 1984, when these women were between the ages of 47 and 61, only 8 percent of white married women and 20 percent of white unmarried women lived in households containing their parents, grandchildren, or other nonnuclear kin. The proportions were higher for comparable black women: 27 percent of the married and 34 percent of the unmarried. However, over the fifteen years covered by the survey, about one-third of all white women and fully two-thirds of the black women lived in a household containing extended kin at some point.

These and other data (Ruggles 1994a, 1994b) show that today blacks are more likely than whites to live in extended-family households. This was not the case until about 1940. What has happened is that the decrease in intergenerational coresidence since the late nineteenth century has been much steeper for whites than for blacks. This is probably connected to much lower rates of marriage among blacks; living in multigenerational households is much more common for unmarried than for married persons. It may also reflect the shift in the distribution of extended families from the wealthier to the poorer segments of the economic structure. Rather than serving as a means of ensuring inheritance and keeping wealth in the family, extended family living today is more likely to be motivated by a need to share and conserve resources.

The family and household structure of the United States has changed dramatically over the past century, in spite of the fact that our family system has remained nuclear in at least the statistical sense. More and more Americans are living in single-person households before, between, and after marriages. More are living in single-parent households. Collectively Americans are spending smaller proportions of their lives in families of any description than they did in the past (Watkins et al.

1987). However, they are more likely than ever before to live in nonmarital heterosexual unions, and many of them live in households that contain nonnuclear kin at some point in their lives. In fact, there is evidence (Glick et al. 1997) that the proportion of extended-family households increased between 1980 and 1990.

The growth of small and single-person households is in many ways indicative of the fact that more Americans can now afford to remain unmarried, leave unhappy marriages, and maintain their own residences in later life. The proliferation of households represents the proliferation of choices. The consequences of these choices remain to be seen.

## REFERENCES

- Beck, Rubye W., and Scott H. Beck 1989 "The Incidence of Extended Households Among Middle-Aged Black and White Women: Estimates from a 15-Year Panel Study." *Journal of Family Issues* 10:147-168.
- Bumpass, Larry L. 1994. "The Declining Significance of Marriage: Changing Family Life in the United States." Paper presented at the Potsdam International Conference, "Changing Families and Childhood."
- , James A. Sweet, and Andrew J. Cherlin 1991 "The Role of Cohabitation in Declining Rates of Marriage." *Journal of Marriage and the Family* 53:913-927.
- Coward, Raymond T., Stephen Cutler, and Frederick Schmidt 1989 "Differences in the Household Composition of Elders by Age, Gender, and Area of Residence." *The Gerontologist* 29:814-821.
- Glick, Jennifer E., Frank D. Bean, and Jennifer V. W. Van Hook 1997 "Immigration and Changing Patterns of Extended Family Household Structure in the United States: 1970-1990." *Journal of Marriage and the Family* 59:177-191.
- Goldscheider, Frances 1997 "Recent Changes in U.S. Young Adult Living Arrangements in Comparative Perspective." *Journal of Family Issues* 18:708-724.
- Lee, Gary R. 1999 "Comparative Perspectives." In Marvin B. Sussman, Suzanne K. Steinmetz, and Gary W. Peterson, eds., *Handbook of Marriage and the Family*, 2nd ed. New York: Plenum.
- Martin, Teresa Castro, and Larry L. Bumpass 1989 "Recent Trends in Marital Disruption." *Demography* 26:37-51.
- National Center for Health Statistics 1997 "Fertility, Family Planning, and Women's Health: New Data from the 1995 National Survey of Family Growth."

*Vital and Health Statistics*, Series 23, No. 19. Hyattsville, Md.: Public Health Service.

- 1999 “Births: Final Data for 1997.” *National Vital Statistics Reports*, series 47, no. 18. Hyattsville, Md.: National Center for Health Statistics.
- Oppenheimer, Valerie K., Matthijs Kalmijn, and Nelson Lim 1997 “Men’s Career Development and Marriage Timing During a Period of Rising Inequality.” *Demography* 34:311–330.
- Pebley, Anne R., and Laura L. Rudkin 1999 “Grandparents Caring for Grandchildren: What Do We Know?” *Journal of Family Issues* 20:218–242.
- Raley, R. Kelly 1996 “A Shortage of Marriageable Men? A Note on the Role of Cohabitation in Black–White Differences in Marriage Rates.” *American Sociological Review* 61:973–983.
- Ruggles, Steven 1994a “The Transformation of American Family Structure.” *American Historical Review* 99:103–128.
- 1994b “The Origins of African American Family Structure.” *American Sociological Review* 59:136–151.
- 1996 “Living Arrangements of the Elderly in the United States.” In Tamara K. Hareven, ed., *Aging and Intergenerational Relations: Historical and Cross-Cultural Perspectives*. Berlin, Germany: de Gruyter.
- Seward, Rudy R. 1978 *The American Family: A Demographic History*. Thousand Oaks, Calif.: Sage.
- Smith, Daniel Scott 1979 “Life Course, Norms, and the Family System of Older Americans in 1900.” *Journal of Family History* 4:285–298.
- Sweeney, Megan M. 1997 “Remarriage of Women and Men After Divorce.” *Journal of Family Issues* 18:479–502.
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States*, 118th ed. Washington, D.C.: U.S. Government Printing Office.
- 1998 “Marital Status and Living Arrangements: March 1998.” <http://www.gov/population/www/socdemo/ms-la.html>.
- Waite, Linda J. 1995 “Does Marriage Matter?” *Demography* 32:483–507.
- Watkins, Susan Cotts, Jane A. Menken, and Jon Bongaarts 1987 “Demographic Foundations of Family Change.” *American Sociological Review* 52:346–358.
- White, Lynn 1994 “Coresidence and Leaving Home: Young Adults and Their Parents.” *Annual Review of Sociology* 20:81–102.

## FAMILY AND POPULATION POLICY IN LESS DEVELOPED COUNTRIES

In 1798, Thomas Malthus penned his famous dictum that agricultural productivity increases arithmetically while populations increase geometrically. Ever since that time, scholars have concerned themselves with the relationship between population dynamics and economic development. With the rise of the modern nation-state, interest grew in how governmental family and population policies might affect the health of national economies. At the turn of twenty-first century comes, accelerating economic globalization, the decline of communism, and the growing influence of the United Nations and other international organizations provide an increasingly complex context within which to understand family and population policy in less developed countries.

The designation of countries as “less developed” or “more developed” (or “developing” versus “developed”, “newly industrializing” versus “industrialized”) relies on a multidimensional construct of development, ranging from the conditions of education and health to agricultural surplus generation, industrialization, and capital accumulation (Staudt 1997). Perhaps the best single measure of development is a country’s real gross domestic product (GDP) per capita (Crenshaw et al. 1997). However, while meager per capita productivity is characteristic of less developed countries, these nations fall along a productivity continuum of considerable range. Less developed countries also vary greatly in size, geographic location, and quality of natural resources. A study of population dynamics in seventy-five developing nations included countries as diverse as Mexico, Brazil, Sierra Leone, Rwanda, Egypt, Pakistan, Malaysia, and the Philippines (Crenshaw et al. 1997).

Like more developed nations, some of these countries have highly centralized governments; others are weak states with little coordination of productive or distributive activities. All, however, are perceived as promoting some form of family policy. In many of these countries, policies that attempt to regulate the family per se as a social institution are not explicitly articulated. Rather, policies regarding the family often operate indirectly through measures directed toward population control (e.g., family-planning programs and

GARY R. LEE  
PADMA SRINIVASAN

restrictions on internal migration) or economic development (e.g., agrarian reform and import-substitution policies). In some parts of the developing world (e.g., China, India, the Islamic states), extensive legal and political measures attempt to govern the formation, functioning, and dissolution of families more directly through, for example, laws on inheritance, dowry, divorce, and filiation. These more explicit family policies often bear the traces of traditional religious interests, but they may also reflect attempts to “modernize” countries by weakening traditional patriarchal control over the disposition of property and the labor of family members. Thus, family policy in less developed countries, whether indirect or explicit, is inextricably bound to questions of economic development, including those concerning the relationship between population and resources.

As a result, few social scientists today treat comparative family policy as an arena of inquiry separate or distinct from comparative studies of population or economic development policy. In fact, attempts to define “family policy” as a distinct object of study have been conspicuously absent from the scholarly literature since the late 1970s (for an example of one such attempt, see Kamerman and Kahn 1978). In part, waning efforts to define and examine family policy per se reflect heated debates about whether a uniform definition of the “family” is useful or appropriate. Feminist critiques in particular have challenged uniform definitions of the family as overly monolithic, static, and undifferentiated, claiming that these definitions ignore the diversity of family patterns not only in less developed nations but also in “modern” countries such as the United States (Beneria and Roldan 1987; Collins 1991; Baca Zinn 1990).

These definitional disputes have given rise to different concepts of what the family is. For example, some scholars argue that variations in household and kinship relations exist worldwide; they tend to prefer the term “household” to “family,” defining the former as a residential unit wherein people live in close proximity and organize activities to meet daily needs (Chow and Berheide 1994). Others, however, argue that adopting the household as the unit of analysis, with an attendant focus on how households respond “strategically” to the external economic and political environment, obscures conflicts of interest that arise within households (Bourque and Warren 1987; Wolf

1992). These critics tend to conceptualize the family or household not as a unit or entity but as a site that offers a cultural and moral context for the interaction of personalities, for the production and consumption of goods, and for the reproduction of populations (Acker 1988). This view of the family tends to emphasize how different interests within families are shaped by gender and age—interests that are themselves related to competition and conflict between larger systems of authority and of control over labor (e.g., capitalism, socialism, patriarchy) within national settings.

Although the study of family policy in less developed nations has largely been overtaken by more fundamental debates about the “proper” unit of analysis, some of these arguments provoke new ways of thinking about *why* developing nations have an interest in the regulation of families. The question of what constitutes a “family” for the purposes of analyzing policy might be found in answers to the question: What is it that families do, in addition to their contributions to population dynamics and economic development, that captures the interest of governments? That is, why do states have a “compelling interest” in families?

Part of an answer may be found in Cherlin’s (1999) definition of the “public family,” a definition that attempts to locate the contribution of families to the public welfare in their caring for dependent members of the population, such as children, the sick, and the elderly. Adapting Cherlin’s definition to a global context, public families may be defined as one adult, or two or more adults who are related by marriage, kin ties, shared parenthood, or coresidence, who is/are taking care of dependents and the dependents themselves. In nearly every society, most caretaking of dependents is performed within a small web of relations like those defined above. Although this work is largely unpaid, it is of great social value, being necessary for the reproduction of a society from one generation to the next. Hence the state’s interest in how well families manage the care of dependents. But government interest extends beyond merely ensuring the replenishment of a population. As Folbre (1994) argues, the caretaking that goes on within families can result in the production of valuable public goods—most notably through the rearing of children who will grow up to be productive adult members of a society. Thus, governments have an interest in promoting

caretaking within families that gives rise to positive collective benefits (productive members of the next generation, including those who will continue to provide caretaking within families).

Policies that provide explicit incentives for family caretaking in the form of tax breaks or government subsidies for day care, parental leave, or child allowances tend to emerge in countries with a highly developed welfare-state apparatus, which presumes the existence of large, stable (if not expanding) public revenues. That is, these policies tend to presume an already-high level of economic development that ensures stable public revenues. The European social democracies—in particular, Sweden—are examples in this regard (Acker 1994). However, some developing countries, such as China, have instituted policies that support caretaking in the form of parenting education programs, child allowances, maternity leave, and planning for social security for the aged (Chow and Chen 1994). The difference in the Chinese case, as in that of other less developed nations, is that these policies are tethered to population policies and goals. The Chinese government, for example, makes eligibility for forms of caretaking support contingent on compliance with its one-child policy, levying penalties in the form of fines, wage deductions, possible job demotion, and forfeiture of other social benefits on parents who have a second child (Chow and Chen 1994). Thus, caretaking policies in the developing world are often put into place as a mechanism for achieving family planning or development goals that aim, over the long term, to ensure national prosperity; they less often derive from such prosperity to begin with, as in the West. The extent to which developing countries manage to implement such “conditional” caretaking policies depends on the strength of the state—in particular, on whether this strength relies on autocratic forms of rule (e.g., China). In developing countries with weaker governments or less autocratic systems of power, support for caretaking more often arises from the adoption of programs or initiatives that are funded by international organizations committed to population control and/or globally organized economic growth. These programs often take the form of children’s health initiatives, such as immunization and nutritional-support projects sponsored by the World Health Organization and the World Bank.

A second aspect of governmental interest in the family among less developed countries concerns the role of family and kin networks as the matrix for traditional authority relations. In many emerging nation-states, political power remains locally controlled by clans or confederations of kin that resist attempts by central governments to consolidate power. Loyalties to religious or ethnically based claims to authority also tend to be coupled closely with loyalties to family and larger kin groups. Thus, the emergence of explicit family policies in less developed countries may reflect attempts on the part of central governments to wrest political control from clans or lineages (such as in sub-Saharan Africa; see Migdal 1988). In some parts of the developing world, changes in family law and policy are a function of competition between bureaucratic state interests and traditional religious claims to authority. For example, in the mid-1950s the government of Tunisia outlawed “repudiation,” or the husband’s Quranic, unilateral prerogative to terminate a marriage at will without judicial involvement. Prior to this time, family matters had been regulated by traditional Islamic doctrine, with allegiance to Islamic law being regionally based and deeply rooted in kin loyalties (Charrad 1994). Changes in the Tunisian civil code to abolish repudiation and make judicial intervention mandatory for divorce had the effect of unifying and strengthening the judicial system by creating a national network of courts. In terms of actual implementation of the new code, judges found themselves shifting between the values of traditional Islamic doctrine and the principles of reform. Nonetheless, the new policy has had a transformative effect on the structure of Tunisian rule, strengthening and consolidating the power of civil law. The case of the Philippines offers another example of how government policy toward families reflects competition between religious and more secular authority claims; the 1986 Philippine constitutional statement about the family reflects an uneasy compromise between Roman Catholic groups and more liberal elements of Filipino society (“Proceedings of the Constitutional Commission” 1986).

In short, despite widespread disagreement in the literature over definitions of “family” and the scope of “family policy,” government policies in less developed nations reveal interests in the caretaking of dependents and in the consolidation

and centralization of authority that might be usefully thought of as targeting “the family.” In contrast to the field of family policy, the terrain of population policy is much more clearly defined, though no less marked by controversy. Population policies differ from family policies because the former are designed to meet specific demographic goals. The most common population policies found in less developed countries are those that attempt to ameliorate population growth by reducing birthrates, for example, through programs in family-planning education and in the distribution of contraceptives (International Institute for Sustainable Development 1994). Other population policies attempt to control migration flows from, for example, rural to urban economic sectors in an attempt to control the mobility of labor and patterns of economic development. These migration policies are especially prevalent in sub-Saharan African countries, where increasing urbanization and the influx of wartime refugees from neighboring areas have placed severe pressures on cities and depleted the supply of labor for rural agriculture. For example, in the 1980s Zambia took steps to curb rural-to-urban migration and promote the resettlement of rural areas through such measures as removing subsidies for commodities in urban areas, a “back to the land” policy, and youth training programs in agriculture and farm management, carpentry, and other rural community-building skills (Mijere and Chilivumbo 1994).

Several demographic problems are a source of shared concern among less developed nations. Concern over rapid population growth has motivated a variety of family-oriented laws and programs, including raising the minimum legal age for marriage in order to shorten the reproductive span during which women are exposed to regular sexual activity (Piepmeier and Hellyer 1977), providing free family-planning services, and, as in China, offering incentives and disincentives for childbearing that involve the provision or withholding of housing, paid maternity leave, and medical or educational services (Greenhalgh 1990; Quah 1990).

Another demographic problem faced by many less developed countries is the lengthening of childhood dependency through adolescence and the concomitant rise of an independent youth culture brought about by the earlier onset of puberty and rising ages at marriage (United Nations

1989). A number of governments in Asia have established sex education or “family life education” policies and programs that target burgeoning adolescent populations, partly as a way of maintaining government family-planning programs in the face of declining levels of marital fertility (Xenos 1990). Other governments have increasingly differentiated youth from adults in their constitutions (Boli-Bennett and Meyer 1978), and programs that explicitly target adolescents and youth are an increasingly important element of social policy in less developed nations (see Central Committee on Youth 1988; Paxman 1984).

Another emerging demographic problem affecting many less developed countries is the growth of the aged population, a situation that typically arises whenever birthrates fall. While less developed nations may have a clear and compelling interest in the production of children as “public goods,” the question of how to care for a growing population of the aged amid declining birthrates and the emergence of smaller families raises thorny issues. In some countries, policies are being considered to keep this burden within families rather than making it a government responsibility, reflecting the state’s reluctance to absorb the costs of dependency among those who are no longer members of the productive work force.

All three of the above concerns—about rapid population growth, an expanding youth culture, and care for the aged—mark the recent history of policy in China. The Chinese one-child policy was adopted in 1979 to ameliorate population pressure on China’s food supply and to reduce state expenditures. Later, it developed into a policy promoting modernization and economic development, with the Chinese government offering several rationales for planned fertility: It would foster better health care for children and mothers, establish better social conditions for the rearing of future generations, increase work efficiency and political awareness, and promote equality of the sexes (Huang 1982). The government created a number of assistance programs, such as fertility education, free birth control and abortion, and planning for social security for the aged; it also offered numerous economic incentives for compliance, such as extra maternity leave, housing or



farmland, free doctor visits, day care, and wage bonuses.

Overall, these provisions appear to have relieved China's population crisis: As of 1994, population growth rates were the lowest since 1949, and the percentages of women marrying at an early age and bearing more than one child have declined (Beijing Review 1995). But because enforcement of the one-child policy is left to local jurisdictions, implementation and outcomes vary by location. The policy appears to have gained greater acceptance in the wealthier, more urban provinces, where the incentives are more abundant. Even in these areas, however, the policy has had an important, unanticipated consequence: The preciousness of the single child has led parents, especially mothers, to spend more time on child care and housework than parents in multiple-child households. Instead of promoting women's greater participation in employment, the policy has led to a reassertion of women's traditional roles as homemakers and nurturers. The policy's weakening of the traditional Chinese system of old-age dependence on adult children (primarily sons) was more readily anticipated, but an alternative system of old-age support has yet to emerge.

As the aims (though not the consequences) of the Chinese policy illustrate, another area of concern involves global social movements, including those supporting equal rights for women and other oppressed groups (Bandarage 1997; Chow and Berheide 1994; Staudt 1997). A second global trend is that of environmentalism (McMichael 1996). A variety of "green" movements have emerged that question the assumptions behind unrestricted economic growth and have redefined the debate about development in Third World countries, focusing on the need for agricultural sustainability, the protection of ecologically vulnerable habitats (particularly in the tropical zones), and the maintenance of biodiversity. Concerns about the environmental degradation sometimes wrought by development projects are often coupled with concerns about the rights and well-being of oppressed groups—particularly in Africa, where mining and rural development policies have differentially affected female-headed households

and, as in the case of South Africa, magnified racial inequalities (Schonfield 1994).

Emerging global viewpoints and the sanction given to them by international organizations represent another trend that is likely to influence future policy concerning families in developing countries. In the 1990s alone, various arms of the United Nations offered explicit recommendations on a range of policies, including those on education, family health, women's labor-force participation, and the family (International Institute for Sustainable Development 1994, 1995). Often these recommendations arose from the participation of nongovernmental organizations (NGOs) in planning and development activities. Such recommendations have been closely tied to the system of international aid that includes donors such as the World Bank and the U.S. Agency for International Development. However, the growing visibility of NGOs at U.N. conferences such as the 1994 Cairo Conference on Population and Development and the 1995 Beijing Conference on Women in part reflects the NGOs' lack of efficacy at the national level: These conferences provide forums for NGOs that have encountered resistance from national governments or have failed to mobilize support from the local populace (Baden and Goetz 1997).

Local and national resistance to NGO or multilateral initiatives regarding families sometimes stems from the belief that these initiatives have been "co-opted" by large international aid agencies that neglect the needs of local populations and their families (Mukherjee 1993). For example, donor-influenced policies in Malawi have disadvantaged female-headed households, favoring households headed by men, even though the traditional culture recognizes the mother-child pair as the most important family unit (Rodgers 1980; Spring 1986). Donor-influenced structural adjustment policies (SAPs) have become another source of conflict about policy. In response to recent economic stagnation and the Third World debt crisis, these adjustments have involved reducing levels of government employment, "floating" national currencies, and cutting tariffs and subsidies, while encouraging export-driven economic growth through labor-intensive manufacturing. In

many countries, women have responded by increasing their participation in home-based industries, often to provide themselves with security should a spouse lose a formal-sector or government job as a result of structural adjustment (Osirim 1994).

The case of Zimbabwe provides an example of how these SAPs can come into conflict with the aims of government policies concerning families. In the early 1980s, Zimbabwe passed laws that ensured women's claims to property in the event of divorce and allowed married women to obtain loans and property in their own names to establish businesses. These laws, supplemented by government loan and training programs to assist women entrepreneurs, were designed to encourage the greater participation of women in national development and to advance their civil rights. However, the advent of Zimbabwe's SAP in 1990 signaled a shift in priorities toward "shrinking the state." The resulting increases in transportation, wholesale, and licensing costs for businesses were particularly hard on women entrepreneurs, some of whom had become the sole wage earners of their families after their husbands had lost a government job (Osirim 1994).

The increasing involvement of transnational social movements, international donor agencies, and nongovernmental organizations in policy development raises an important issue: the strain in many less developed countries between a Western vision of the family that may be supported by the government, NGOs, or international aid agencies, and the traditional family forms supported at the local, grassroots level. In many less developed countries, the government vision of the family may result from the colonial experience, the exposure of elites to Western society, or the perception that a more Westernized family system fits better with other governmental goals, such as economic development or undermining the traditional authority of kin, religious, or ethnic groups. Wherever political struggle focuses on family and population policies, however, the underlying issue often at stake—to which debates in the literature attest—is competition between alternative definitions of the family.

## REFERENCES

- Acker, Joan 1998 "Class, Gender, and the Relations of Distribution." *Signs* 13:473-497.
- 1994 "Women, Families, and Public Policy in Sweden." In Esther Ngan-Ling Chow and Catherine White Berheide, eds., *Women, The Family, and Policy* Albany, N.Y.: State University of New York Press.
- Baca Zirr, Maxine 1990 "Family, Feminism, and Race in America." *Gender and Society* 4:68-82.
- Baden, Sally, and Anne Marie Goetz 1997 "Who Needs [Sex] When You Can Have [Gender]: Conflicting Discourses on Gender at Beijing." In Kathleen Staudt, ed., *Women, International Development, and Politics: The Bureaucratic Mire*. Philadelphia: Temple University Press.
- Bandarage, Asoka 1997 *Women, Population, and Global Crisis: A Political-Economic Analysis*. London: Zed Books.
- Beijing Review 1995. *State Statistics Bureau* 38:26. Beijing: Beijing Review.
- Beneria, Lourdes, and Martha Roldan 1987 *The Crossroads of Class and Gender*. Chicago: University of Chicago Press.
- Boli-Bennett, John, and John W. Meyer 1978 "The Ideology of Childhood and the State: Rules Distinguishing Children in National Constitutions, 1870-1970." *American Sociological Review* 43:797-812.
- Bourque, Susan, and Kay Warren 1987 "Technology, Gender, and Development." *Daedalus* 116:173-197.
- Central Committee on Youth 1988 *Report on Youth Policy*. Hong Kong: Author.
- Charrad, Mounira 1994 "Repudiation versus Divorce: Responses to State Policy in Tunisia." In Esther Ngan-ling Chow and Catherine White Berheide, eds., *Women, The Family, and Policy*. Albany, N.Y.: State University of New York Press.
- Cherlin, Andrew 1999 *Public and Private Families: An Introduction*. Boston: McGraw-Hill.
- Chow, Esther Ngan-ling, and Catherine White Berheide 1994 "Studying Women, Families, and Policies Globally." In Esther Ngan-ling Chow and Catherine White Berheide, eds., *Women, The Family and Policy*. Albany, N.Y.: State University of New York Press.
- , and Kevin Chen 1994 "The Impact of One-Child Policy on Women and the Patriarchal Family in the People's Republic of China." In Esther Ngan-ling Chow and Catherine White Berheide, eds., *Women, The Family, and Policy*. Albany, N.Y.: State University of New York Press.

- Collins, Patricia Hill 1991 *Black Feminist Thought: Knowledge, Consciousness, and The Politics of Empowerment*. New York: Routledge.
- Crenshaw, Edward M., Ansari Z. Ameen, and Matthew Christenson 1997 "Population Dynamics and Economic Development: Age-Specific Population Growth Rates and Economic Growth in Developing Countries, 1965 to 1990." *American Sociological Review* 62:974-984.
- Folbre, Nancy 1994 "Children as Public Goods." *American Economic Review* 84:86-90.
- Greenhalgh, Susan 1990 *State-Society Links: Political Dimensions of Population Policies and Programs, with Special Reference to China*. (Research Division Working Papers No. 18). New York: Population Council.
- Huang, Lucy Jen 1982 "Planned Fertility of One-Couple/One-Child Policy and Gender Equality in the People's Republic of China." *Journal of Marriage and the Family* 44:775-784.
- International Institute for Sustainable Development 1994 "International Conference on Population and Development." *Earth Negotiations Bulletin* 6 (40). Author.
- 1995 "Summary of the Fourth World Conference on Women." *Earth Negotiations Bulletin* 14 (21). New York Author.
- Kammerman, Sheila B., and Alfred J. Kahn, eds. 1978 *Family Policy: Government and Families in Fourteen Countries*. New York: Columbia University Press.
- McMichael, Philip 1996 *Development and Social Change: A Global Perspective*. Thousand Oaks, Calif.: Pine Forge Press.
- Migdal, Joel S. 1988 *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton, N.J.: Princeton University Press.
- Mijere, Nsolo, and Alifeyo Chilivumbo 1994 "Rural-Urban Migration and Urbanization in Zambia during the Colonial and Postcolonial Periods." In Ezekiel Kalipeni, ed., *Population, Growth and Environmental Degradation in Southern Africa*. Boulder, Colo.: Lynne Rienner Publishers.
- Mukherjee, Vanita Nayak 1993 *Shaping a Better Future: Women's Perspectives on Alternative Economic Framework and Population and Reproductive Rights*. Kuala Lumpur, Malaysia: Asian and Pacific Development Centre.
- Osirim, Mary 1994 "Women, Work, and Public Policy: Structural Adjustment and the Informal Sector in Zimbabwe." In Ezekiel Kalipeni, ed., *Population Growth and Environmental Degradation in Southern Africa*. Boulder, Colo.: Lynne Rienner.
- Paxman, John M. 1984 *Law, Policy, and Adolescent Fertility: An International Overview*. London: International Planned Parenthood Federation.
- Piepmeire, Katherine B., and Elizabeth Hellyer 1977 "Minimum Age at Marriage: 20 Years of Legal Reform." *People* 4 (3):
- "Proceedings of the Constitutional Commission on the Family, 1986 (Part 1)" *1986 Philippine Population Journal* 2(1-4):135-156.
- Quah, Stella R. 1990 "The Social Significance of Marriage and Parenthood in Singapore: Policy and Trends." In S. R. Quah, ed., *The Family as an Asset: An International Perspective on Marriage, Parenthood, and Social Policy*. Singapore: Times Academic Press.
- Rodgers, Barbara 1980 *The Domestication of Women: Discrimination in Developing Societies*. New York: St. Martin's Press.
- Schonfield, Anne 1994 "Securing the Future: Environmental Issues as a Priority in South Africa." In Ezekiel Kalipeni, ed., *Population Growth and Environmental Degradation in Southern Africa*. Boulder, Colo.: Lynne Rienner Publishers.
- Spring, Anita 1986 "Men and Women Smallholder Participants in a Stall-Feeder Livestock Program in Malawi." *Human Organization* 45(2):154-161.
- Staudt, Kathleen 1997 "Gender Politics in Bureaucracy: Theoretical Issues in Comparative Perspective." In Kathleen Staudt, ed., *Women, International Development, and Politics*. Philadelphia: Temple University Press.
- United Nations 1989 *Adolescent Reproductive Behaviour: Evidence from Developing Countries*, vol. 2. New York: Author.
- Wolf, Diane 1992 *Factory Daughters*. Berkeley: University of California Press.
- Xenos, Peter 1990 "Youth, Sexuality and Public Policy in Asia: A Research Perspective." In S. R. Quah, ed., *The Family as Asset: An International Perspective on Marriage, Parenthood, and Social Policy*. Singapore: Times Academic Press.

JULIE BRINES

## FAMILY AND RELIGION

Social scientific notions of the disappearance or vestigialization of religion and family are deeply

rooted in our theoretical conceptions of the social processes that created the modern world and that now are transforming that modernity into postindustrial, postmodern society. Theories of modernization envision social change as entailing the rationalization of all spheres of existence. In a statement characterizing the classic modernization approach, Moore (1963, p. 79) says, "A major feature of the modern world . . . is that the rational orientation is pervasive and a major basis for deliberate change in virtually every aspect of man's concerns." There is little room for the seemingly irrational and unscientific impulses of religion, primary emotions, and familial concerns.

With this approach, the secularization of religion is a given. Moore (1963, p. 80) states, "Even with regard to the role of religion in human affairs, the 'rational spirit' takes the form of *secularization*, the substitution of nonreligious beliefs and practices for religious ones." Though religion survives, it addresses "personal misfortune and bereavement" above all else in modern society (Moore 1963, p. 104).

Furthermore, "economic modernization" tends to have "negative consequences for extended kinship systems" and leads to "extensive 'family disorganization'" accompanying the "breakdown of traditional patterns and the incomplete establishment of new institutions" (Moore 1963, p. 102). For modernization theorists, although families remain significant as consumption units, the "decline" of the family (Popenoe 1988) is, at minimum, a metaphor for its consignment to a peripheral societal role. The analogue of the notion of linear secularization of religion is the idea of the loss of family functions (Vago 1989, pp. 150-157). Shaped in the eighteenth and nineteenth centuries, modernization views have continued to dominate public opinion and much of social scientific discourse. In general, according to modernization theories, both family and religion are relegated to the "private" sphere, set apart from the broader social processes and, thus, less significant than those broader processes.

Despite this widespread orientation, a revolution in the social sciences has been gaining momentum over the last twenty years or so. The

message of this revolution is that the modernization perspective is no longer an adequate vision for understanding the dynamics of modernity or the potentials of postmodernity for religion and family. In the sociology of religion, the paradigm shift moves social science away from a focus on religion as a disconnected phenomenon to a much more complex view of the nature of religious inter-institutional relations. Reflecting this shift, numerous scholars have begun to examine religion as an influence on and as an effect of varied social, political, and economic variables (e.g., Carter 1996; Cousineau 1998; Hammond 1985; Misztal and Shupe 1992; Roberts 1995; Rubenstein 1987; Shupe and Misztal 1998; Swatos 1992; Witte 1993). A market model of religion, based in rational choice propositions, has become the most strongly debated version of the new way of looking at the religious institution (Hadden 1995; Warner 1993; Young 1997).

Similarly for the family, there are many who now argue that, in spite of its changing forms and functions, the family as an institution remains crucially central to social processes and to the patterns of change determining the future of human societies (cf. Cherlin 1996). If not taken into consideration, family processes themselves are liable to torpedo efforts at planned social change and to deflect the vectors of unplanned change in unexpected directions (Settles 1996).

Attuned to the inter-institutional perspective, this article examines the linkage between family and religion. Then, after discussing social change processes, we make the point that the religion and family linkage today is important not only in the burgeoning private sphere but in the public realm as well. In spite of its importance, sociologists have given relatively little attention to family and religion together (Thomas and Cornwall 1990). A few journal articles and four collected volumes have addressed this institutional linkage. D'Antonio and Aldous (1983) and Thomas (1988) edited general volumes, and a collection by Ammerman and Roof (1995) focused on family and religion in relation to work. As of 1999, only one book (Houseknecht and Pankhurst 2000) had taken an international comparative approach to the topic. (An expanded

discussion of some of the themes presented here can be found in the introduction of the latter volume.)

### THE RELIGION AND FAMILY LINKAGE

In order to understand the significance of religion and family for both the private and public spheres, we must see clearly their unique characteristics and how they interrelate. The inter-institutional relations between family and religion are strong and qualitatively different from other institutional relationships. Berger (1967) noted that in premodern societies kinship was permeated with religious meaning, and in modern societies religion remains closely connected to the family. Hargrove (1979), in her systematization of the sociology of religion, argued that religion and family have had a close relationship throughout history in both Western and non-Western societies. D'Antonio and colleagues (1982) also stressed the significance of the connections between these two institutions.

Both the familial and religious institutions are characterized by what MacIver (1970, p. 45) called cultural rather than secondary interests. In other words, associations within the religious and familial spheres pursue interests for their own sake, because they bring direct satisfaction, not because they are means to other interests, as in the case of economy and polity. Both family and religion are devoted to organizing primary group relations. They stand out as the only two institutions that deal with the person as a whole rather than just segmented aspects of individual lives. These various similarities that religion and family share serve to strengthen the inter-institutional ties between them.

The interrelations between the institutions of religion and family are reciprocal (Thomas and Henry 1985; Thornton 1985). Religion provides the symbolic legitimation for family patterns (cf. Berger 1969), and the family is a requisite for a vigorous religious system because it produces members and instills them with religious values. In fact, numerous familial events are marked in religious contexts (e.g., weddings and funerals), and many religious observances take place within the familial setting (e.g., prayers at meals and bedtime).

The special affinity of religion and family as institutions takes varied forms. Almost everywhere, religion provides ritual support for family and kinship structures. This is the case even in a highly secularized society such as Sweden (Trost and Palm 2000). In some societies, this ritual support may be seen in ancestor rites or memorialization, for example, in Cameroonian Kedjom rites (Diduk and Maynard 2000), Japan's core religion (Smith 2000), Taiwan's folk religion (Yang et al. 2000), and French Mormonism (Jarvis 2000). Such practices support family life and, at the same time, fulfill a central function for religion itself (cf. Berger 1969, p. 62). In fact, Jarvis argues that the familism of Mormonism, expressed both ritually and in church values, is its greatest asset in the eroding environment for traditional families in France. Moreover, utopian experiments and new religious movements often take the family as their essential focus. According to Christiano (2000), in the Unification Church ("Moonies"), the family serves as more than the organizing metaphor for the group: It provides a basic model for the church's self-conception.

### FAMILY, RELIGION, AND SOCIAL CHANGE

The institutions of family and religion, in their interactions with each other and with the rest of society, can be considered in terms of two broad social change patterns—institutional differentiation and institutional dominance. Secularization can be seen as a special instance of institutional differentiation, and we discuss the concept and its use below, indicating the value of religious economy models over the conventional secularization approach.

**Institutional Differentiation.** Underlying all the dimensions of social change is the notion of institutional differentiation. This phenomenon implies greater specialization. Although the paths and extent of institutional differentiation vary across societies, when differentiation does proceed, we see fundamental changes in inter-institutional relations (cf. Alexander and Colomy 1990; Beckford 1989). Institutional differentiation affects all institutions, and we argue more broadly that one cannot assume inter-institutional isolation of religion

and family in the private sphere, even in highly differentiated societies. The effects of these institutions are always felt across inter-institutional divisions in some measure. In this section, we examine the issue of differentiation of religion and family conceptually.

Our previous description of the religion and family linkage as involving special affinity and reciprocity was not to say that religion and family always and everywhere are, or must be, equally intimately entangled. We can see a continuum in the level of differentiation of these institutions. On the one end of the continuum, Islam in Egypt (Houseknecht 2000) displays a lack of differentiation, an elaborate interweaving of the two institutions that makes each strongly dependent on the other. And research on the Cameroonian Kedjom funerary rites (Diduk and Maynard 2000) provides an example of religious practices that are hardly differentiated from the kinship context; they are precisely an affirmation of the kinship patterns of Kedjom society.

On the opposite end of the spectrum, Sweden, a country in which the Lutheran Church is officially established and a large majority of the population are nominal members, is highly secular. Developments in the family there have widely diverged from the traditional model that Protestant Christianity had advocated. In Sweden, the two institutions are intertwined only in limited ways (Trost and Palm 2000). It is primarily in regard to life-cycle rites that the two intersect. The individualized faith of many Americans also accompanies a highly differentiated system of institutional relationships (Christiano 2000).

**Secularization Theories.** Secularization is a special process of differentiation in which that which was previously under the “sacred canopy” (Berger 1969) of religion is removed from that realm and placed in a nonreligious institutional context. Allegedly, education, the acceptance of science, urbanization, industrialized work-life, and the like take away the mystery of religion and strip it of its “plausibility” in many areas of concern. Thus, cure of disease, protection from misfortune, explanation of the universe, and so forth are made

rational and thus not subject to religious intervention. In this approach, religion remains relevant only for very personal spiritual quests and solace in the private sphere, and most of its social institutional ramifications become, first, empty shells, and eventually vanish.

The notion of a unilinear process of secularization has long troubled many sociologists (cf. Hadden and Shupe 1986; Hadden 1987; Hammond 1985). Some have developed variants that see secularization as a cyclical process with long historical waves. Nisbet (1970), for example, reminds us of the rationality of the eras of classical Greece and of the Renaissance and Age of Reason, with the period from first century Rome to the Renaissance having Christianity “virtually eliminating secular rationalism from the European continent for more than a thousand years” (p. 391). Though we are now in a rationalizing or secularizing age, “To argue permanence for this age would be, on the testimony of history, absurd” (p. 391). Recently, as we shall see, sociologists of religion have focused upon shorter waves of secularization and have viewed the process as self-limiting. The most prominent versions of this approach apply economic models to religious markets, putting aside the notion that religion must be irrational or otherwise antimodern. Some of these postsecularization theorists argue for a rational choice approach to religion (Young 1997), an approach that is largely alien to the mode of thought underlying secularization theory.

In the long debate about the validity of notions of secularization within the sociology of religion, it has become clear that secularization cannot be understood in a simplistic way, if one wants to keep the concept at all. While certain evidence of secularization seems apparent, there are counter-movements suggesting that religion is truly vital in the modern world (cf. Marty and Appleby 1995). The spread of Liberation Theology throughout Latin America in the 1970s and 1980s (Berryman 1987; Smith 1991), followed by the more recent explosion of evangelical, fundamentalist, and pentecostal Protestantism in the same world region (Martin 1990), suggest the power of the human concern regarding spiritual or nonempirical matters. Similarly, the tenacious attachment of Americans to belief in God (Greeley 1989, 1992; cf. Wald

1986), their high levels of religious activity, and the numerical growth and public voice of evangelical Protestantism indicate that religious sentiment of some sort is compatible even with a society that is highly developed socioeconomically. Going beyond the simple Marxist assertion that religion is illusion, even if religious claims are often masks for the interests of power or wealth, religion must be understood as a very real and consequential part of sociocultural life.

Casanova (1994) carried out one of the most extensive recent cross-cultural analyses of religion, a close examination of the conditions of evangelical Protestantism and Catholicism in the United States and of Catholicism in Brazil, Spain, and Poland. He argued that the social scientific literature depicts secularization as having three correlated dimensions, but his research challenges this idea by convincingly showing that the three dimensions are not always present together. First, Casanova accepts the validity of claims that secularization entails a structural differentiation of the religious institution from other institutions as societal modernization takes place. This differentiation means, in particular, the “emancipation of the secular spheres from religious institutions and norms” (Casanova 1994, p. 6). However, the second dynamic often subsumed under secularization—the decline in religious beliefs and practices—cannot be taken for granted, and it does not necessarily follow from the first. The third dynamic, which is the core of the privatization thesis, is that religion will sequester itself in the private sphere under modernity, and, according to some analysts, will be marginalized there. However, this process, too, cannot be assumed to be associated with secularizing institutional differentiation. The second and third dynamics are unwarranted correlates of the first. While Casanova would have us accept the first as the true essence of secularization, he argues that the second and third are not supported by empirical evidence and should not be seen as part of the secularization process.

While there is no question that institutional differentiation is a sort of master process of the modern era, we cannot assume that it has progressed equally far everywhere. As already noted,

cases like modern Egypt and Cameroon, though both experiencing significant pressures toward greater differentiation, evidence far less differentiation between family and religion, and between these two and other institutions, than does, for example, the United States or Sweden. Furthermore, even if there is great differentiation, one cannot presume that the influences of the religious and familial institutions end. As the debates on abortion policy illustrate, even in a highly differentiated society like the United States, there is plenty of room for religious assertions beyond the alleged parameters of secularized religion and into political life. This circumstance indicates that we cannot take for granted notions of the irrelevance of religion for social policy, as secularist analysts are prone to do.

In addition, it also is possible for there to be de-differentiation in a highly differentiated society. Some of this has happened in Belarus during the post-Soviet period (Vardomatskii and Pankhurst 2000). Better known to the U.S. public is the recent passage of laws in Russia favoring the Russian Orthodox Church (Pankhurst 1998) after many decades of antireligious, extremely secularist communist control under the Soviets and after a brief period of strict legal disestablishment of all religions between 1991 and 1997. Here are instances of seeming de-differentiation, where the gap between politics and religion is narrowed. Similarly, the debate in Indonesia over divorce law indicates efforts to reassert religious authority over a legal arena that had been under state control for several decades (Cammack et al. 2000).

**Religious and Familial Markets.** The notion of unilinear secularization seems untenable, but there are certainly processes of decline and growth in religious phenomena that must be explained. The general inter-institutional perspective focuses attention on relationships that are important for these variations in the strength and character of religion. In addition, sociologists of religion have found market model approaches useful, within the general inter-institutional perspective, for generating testable hypotheses about several aspects of religion and family in various societies.

Social differentiation approaches, including secularization theories, start from observations of

society from the top down. An alternative approach is to look from the bottom up, moving from the level of individual social action toward the institutional structures that the aggregation of such action creates. Such an approach is found in rational choice theories, which begin by analyzing patterns of individual behavior according to the economic logic of consumer choice. In rational choice approaches to religion, churches and other religious organizations are seen as firms offering a variety of goods and services to consumer-believers and consumer-parishioners. When patterns of individual consumer choice are aggregated, market structures become apparent. Such markets provide the context within which patterns of supply and demand are worked out. They distribute goods and services to consumers, as well as "market share" to religious firms. Seen from the market model viewpoint, the issue of the strength of religion boils down to the likelihood on the part of potential consumer-believers to buy into a given religious belief or practice, or to affiliate with a given religious firm such as a temple or a missionary organization. Just as the level of economic purchases rises and falls over time, so does the level of various religious "products," like church membership or belief in God, rise and fall. What governs variation in purchases or adoptions (or church membership or belief in God) is the logic of rational choice among the options in the marketplace that are available to the consumer-believers or the consumer-members. Religious organizational leaders, like business executives, proffer a variety of products to the consuming public and vary the price of such products in order to attract consumer-believers. The leaders seek market share in the religious market.

This emerging "new paradigm" for understanding religious change relies upon economic models of the religious market to understand differing levels of religious group affiliation and participation (Hadden 1995; Warner 1993; Young 1997). Stark and Bainbridge (1987) provide an elaborate deductive theory of religion based upon rational choice principles, and this theory provides the backdrop for a series of more recent studies by these authors and collaborators (Hadden

1996). Perhaps most prominent among the studies developed in connection with this theoretical approach is Finke and Stark's (1992) analysis of American religion, which shows that, over the last three centuries, it has grown, rather than declined, in number of participants and proportion of the U.S. population, contrary to what secularization theory would predict. The authors argue that the growth is the result of competition in a pluralistic market. The economist Laurence Iannaccone (1995, 1997, and works cited in these sources) has elaborated several propositions in line with the theory and expanded the application of economic modeling.

The approach seems to hold greater promise for explaining and predicting the dynamism of religious phenomena than do other approaches that are primarily based in traditional functionalist secularization theorizing. For present purposes, its primary wisdom is that secularization processes are self-limiting. That is, when "the processes that erode commitments to a particular set of supernatural assumptions" (Stark and Bainbridge 1987, p. 311) advance far enough, the religious market will be open to new options. These options, according to Stark and Bainbridge (1987), take the form of either schisms from established groups, sects, or cults.

One of the great assets of the cultural market approach is that it is inter-institutional while at the same time giving individuals agency. That is, market-based institutions are structured by the individual patterns of choice that are aggregated in any society. Simultaneously, institution building is regulated by relationships with other institutions in an interactive process. Clearly, there are many avenues for development of other theoretical approaches to these issues. In particular, conflict theory, social movement theory, and Weberian theory have valuable traditions of analysis that apply to family and religion. Theoretically and conceptually, there are similarities between the market approach and these other approaches. But the market approach, which has assimilated numerous insights from other approaches, establishes a greater balance between micro and macro elements.



The economic model in which the market approach is rooted certainly has its own limitations. Perhaps its biggest problem for many readers is the use of a language that has its own implications that do not and should not apply to religious and familial processes. From a strict theoretical perspective, perhaps the most daunting criticism of the theory argues that religious choice by its nature cannot be strictly rational (cf., Chaves 1995; Demerath 1995; Neitz and Mueser 1997). It is this point of view that has discredited the theory most strongly in the religious studies field, where its advocates are relatively few. In his review of Finke and Stark's (1992) revisionist study of American religious history, Martin Marty (1993, p. 88) writes that their "world contains no God or religion or spirituality, no issue of truth or beauty or goodness, no faith or hope or love, no justice or mercy; only winning and losing in the churching game matters." The work, Marty says, is reductionistic, oversimplifying complex issues. Further, James Spickard (1998, p. 110), though expressing considerable sympathy for the rational choice model, has argued that "a rational-choice [*sic*] model can duplicate the overall structure of a religious marketplace, but it cannot demonstrate that individual people think in market terms."

Some sociologists (Ammerman 1997; Ellison 1995; Sherkat 1997) have contended that the notion of rational choice does not adequately take into account the structuring of individual preferences by a host of contextual, cultural, or environmental variables. Several have argued that the value of the approach in analyzing the open, pluralistic market of the United States may not extend to other societies (cf. Warner 1997). The approach also has been charged with being androcentric (Carroll 1996) and with ignoring gender as a variable (Neitz and Mueser 1997). These criticisms are important, but they do not, in our eyes, undermine the approach so much as indicate an agenda for research and further conceptual development.

At this juncture, it should be noted that much of the logic that has just been applied in the discussion of religion could presumably be extended to families. Among many social scientists, there is an overgeneralization from patterns of

differentiation related to families that is similar to that which Casanova identified in secularization thinking about religion (cf. Hargrove 1985). There is a widespread notion that changing patterns in the family (related to institutional differentiation) mean that the family is not centrally important in the modern age. From this premise, we are led to focus on the "private" family to the neglect of the "public" functions that the family retains (Cherlin 1996). It is as if the many "problems" of families remove the family from useful consideration. While scholars often take families for granted, there is clearly no more powerful socializer of the young and no other viable means for the reproduction of society through the birthing of new members than the family, whatever form it takes. If we assume the phenomenon can be identified by its function, that is, by what it does, rather than by its form, then we cannot imagine a society without families and without the effects that families by definition have.

Family sociologists and economists have long examined the economic side of families in its own right, looking at the effects on families of work, the household division of labor, the patterns of money and time budgets, and the like. There also have been significant studies using the market analogy to understand strictly familial phenomena. The economist Gary Becker (1973, 1974, 1976, 1981; Becker et al. 1977) has been the most deliberate in applying analytical tools from economics to family matters. (His impact on the religious economy approach is acknowledged by Iannaccone, 1997.) Others using economic analytical tools in studies of aspects of families that are not formally economic include Grossbard (1978), Huber and Spitze (1980), and Johnson (1980).

In an open "family market"—a notion that deserves much more elaboration and evaluation than we can give it here—the adoption of one family form over another (say, single parenting over dual parenting, isolated nuclear households over interconnected extended kin networks, or formal marriage over nonmarital cohabitation) would relate to the evaluation by participants of the costs and rewards associated with the adoption of the given form. Costs and rewards would be assessed by the individual taking into account the

surrounding culture and relevant subcultures, which would presumably restrict options for “choice” in numerous ways. Trost and Palm (2000) have outlined the conditions for a fairly open familial market system in Sweden.

As some work in the economic approaches to religion shows, one can include consideration of contextual variables in developing criteria of “choice” in such matters (see, for example, Ellison 1995; Iannaccone 1997; Sherkat 1997). Among significant influences on choice would be individual religious beliefs or the values of the religious affiliation a person may have, both of which are the product, to some degree, of the family socialization and community experience of the person. We can extend this line of thought by including a range of cultural variables that allow the research to be fully comparative, potentially applicable to any society or subsocietal unit. In short, conceptually exploiting the analogy with economic markets has potential for the analysis not only of religion but of family matters as well. In the end, such an approach would identify the conditions under which particular choices in the realm of families can be seen as signs of the strength of adaptive families rather than compromises of weak and ineffectual families.

Finally, it is interesting and important to note that the economic approach to religion relies on an understanding of religious choices as based in the household as much as in the individual. Religious “goods” are “household commodities” that the household invests other goods, labor, and skills in producing, according to Iannaccone (1997). In this imagery of the *household* producing religion, then, we again emphasize the intimate connection between religion and family. From this point, we need to establish its place in the larger sociocultural context.

**Institutional Dominance in Modern Societies.** Modern structures of institutional dominance—a concept first articulated by Williams in ([1955] 1970)—do not negate the importance of the religious and familial institutions. In many societies today, family and religion do not rank high in the relative dominance of major societal

institutions. On the contrary, there usually is some other institution that dominates the entire social system, most often the economy (as in Sweden and the United States, for example). Although this means that the economy is much more likely to be the instigator of social change in the less powerful institutions than vice versa, it does not follow that family and religion cannot initiate change as well. They can and they do, although such occurrences are much less frequent than in the case of the dominant institution(s).

Sometimes the question is asked, Do family and religion facilitate or hinder social change? But this question is an inappropriate one to ask. It is one that never is asked about the economy or the polity. Perhaps it is raised in the case of religion and family because these two institutions are viewed as conservative—as part of the past that is slowing down forward motion. This view, however, is untenable. All societal institutions—including family, religion, polity, economy, education, health care, welfare, and so forth—both facilitate and hinder social change. The direction of their effects depends on what best meshes with their interests at a particular point in time. Research on Brazil and Mexico illustrate facilitating effects of religion for change in gender roles. In fact, according to Rosado Nunes (2000), Catholicism in Brazil did more than facilitate—it orchestrated dramatic changes in women’s roles, restricting women to a very limited range of options within the household by the nineteenth century. Fortuny Loret de Mola (2000) showed how the mobilization of Protestantism in Mexico has provided for certain aspects of authority and legitimacy for women that the dominant Catholic culture did not support, that is, religion has fostered change in gender roles in a direction opposite to that of Brazilian Catholicism. While these examples evidence the change-oriented qualities of religion and family, work on Egypt (Houseknecht 2000) provides a good contrasting example in that it describes the hindering influences that religion and family can have with regard to social change in other institutions. Finally, in Godsell’s (2000) study of South Africa, we see examples of both facilitation and hindrance of social change by the religious and familial institutions. While the East Asian Hindu and Muslim

networks promote entrepreneurship, the black African Christian ones tend to inhibit the development of entrepreneurial activities.

In the modern world, the intersection of institutions is largely in the realm of economic issues—because of the dominance of the economic institution. This is true even in the interactions of noneconomic institutions, since the values and norms of the dominant institution “permeate a great many areas of life and enter into the operation of other institutions” (Gouldner and Gouldner 1963, p. 496).

### RELIGION AND FAMILY IN THE PRIVATE AND PUBLIC SPHERES

Although analysts have tended to see family and religion as institutions of the private sphere, there is no question but that they are both found in the public sphere as well. In the following sections, we elaborate on the dynamics that are found within and between these spheres.

**The Private Sphere.** Some theorists have argued that “privatization” (Luckmann 1967; Berger 1967) characterizes religion in modern societies and that the private sphere is shared by the family (Berger 1969). With modernization, these two institutions tend to become more specialized and to take on new forms as they structure and give substance to the private sphere. A high degree of differentiation, though, does not put an end to the family and religion connection, even though it weakens the relationships that they have with each other and with other societal institutions. The private sphere may seem to be, in a macro sense, peripheral in the modern world, but it nevertheless is where we find a bedrock of mutually reinforcing relations between family and religion. As noted earlier, both of these institutions focus on primary relations, and, in the past, primary relations were much more encompassing of all inter-institutional relations. This meant that the points of intersection that these two institutions had with each other and with other societal institutions were many. In modernized societies, though, the relevance of primary relations has come to be

limited more to the private sphere than in the past. And it is in the private sphere where we see what is really unique about the family and religion linkage. It is here that the connection is cut to the bare bones, and it is here that we see the affinity (although reduced) that persists despite differentiation.

With modernization and postmodernity, the private sphere has emerged as a unique social phenomenon. Although family and religion come to be dominated by other societal institutions in the modern setting, the private sphere that they constitute is of growing importance. It is a buffer zone in which individuals receive support that helps them absorb the stresses and strains brought on by their public activities in other institutional spheres. Because the public and private domains become less and less well integrated, the need for retreat to privacy grows. Not only is the private sphere an essential retreat, but it is also a place where people can devote more and more of the bounty of economic development—increased leisure time, less constant concern with mere survival, and greater financial and other material assets. In short, the private sphere is expanding in the modern world both because of the social psychological need for it and because of the availability of greater resources that can be devoted to it.

The growth of the private sphere signifies one way that the importance of family and religion is growing. Roof (1993) and Wuthnow (1994a) have argued that a primary form of contemporary American religion is found in the self-help or support group (which is frequently located within the context of the church, though it need not be). Aimed at solving problems of individual adaptation, interpersonal relationships, and local community issues, such groups would seem to represent a therapy technique for private troubles. As the private sphere grows, this function grows to match. However, both Roof and Wuthnow also claim that these groups are, in fact, linking mechanisms which bring the private sphere and the public sphere into connection. They are means of overcoming one-dimensional individualism and of connecting the individual with the broader community.

**The Public Sphere.** The recognition that the private functions of religion and family are vital

and even expanding in modern societies is not inconsistent with the fact that the public side of these two institutions retains significance (cf. Beyer 1994). Both family and religion provide important public functions that, in the end, demand our attention. They are public functions that are in crisis in many ways in modern and globalizing societies, and understanding them clearly should aid those who seek to overcome the crises. Following the work of Cherlin (1996) on families and of Casanova (1994), Cochran (1990), and Wuthnow (1994b) on religion, we note that both these institutions provide a range of significant public goods. Such goods are general benefits for the society; they cannot be denied to those who did not participate in producing them, and they often are in short supply due to the tendency for nonproducers to “free-ride” on the efforts of those voluntary actors who take part in the production.

Families provide the public good of children—they give birth to them and raise them to be contributing members of society. These children, then, by being productive, paying taxes, and paying into Social Security and other pension systems, help support all persons in the society, but particularly persons who have aged past the productive years. In addition to reproduction, Cherlin (1996) asserts that, in bearing the burden of dependence created mostly by care for children and care for the frail elderly, families are providing public goods. These public functions are fewer than the public functions of families in the past (Demos 1970), but they are nevertheless extremely important in the modern world.

In a similar vein, religion provides important public goods as well. The first of these is moral values. In modern societies, many citizens do not nurture moral values through church participation and support, and yet the Ten Commandments are nearly universally honored as moral values. Although derived from religious sources and cultivated and rehearsed in religious settings on a regular basis, they are not perceived as sectarian or limited in applicability. They are the concrete form of broad societal values that are not, for the most part, in dispute. Still, it is only the religious groups—aided and abetted by families, of

course—who spend the time reminding us of the importance of these fundamentals. Even atheists profit from the order and social stability that such an emphasis nurtures (cf. Cochran 1990). While Durkheim (1973) thought that adequate progress would lead to the usurpation of this moral production role of religion by education, with schools sustaining the essential values for the modern society, this situation has not yet arrived. The difficulty in designing a moral program that can be taught in the schools seems only more and more dependent upon the interpretation of relevant faith traditions. The schools, rather than displacing religious groups in this task, seem to be calling upon them for clarification and support (cf. Wuthnow 1994b).

Cochran (1990) stresses that religion is also important as a forum for public participation, that is, a place where people come together and discuss, evaluate, lobby, and generally keep informed about public issues and problems. In fact, while the private side of religion focuses upon individual salvation, there is also the supremely public side of “prophetic” religion, to use Weber’s ([1922] 1963) term for the kind of religion that challenges and calls for reform in society. As a forum for participation, the church, synagogue, mosque, or temple provides a venue in which individuals and families work out their positions vis-à-vis the politics and economics of their communities.

Casanova (1994) argues that religions today are more and more asserting their influence and enunciating their interest in secular affairs, mostly rising against the presumptions of the state and the market to prevent their incursion into religious matters. There is a new era, he asserts, of “public religion.” The core of Casanova’s argument is that, instead of a retreat by religion into a segregated and marginalized private sphere, there is today a “deprivatization” of religion in the world as one after another faith reasserts its claims to an active role in the broader society, especially in politics and economics. The mobilization of religious groups around the abortion debate in the United States is one such example, with the well-known engagement of American Catholic and conservative Protestant groups in unexpected political contestation with the government (cf. Pankhurst

and Houseknecht 1983). Another form of this deprivatization is the demand that the church not ally itself with the political and economic elites when the latter create and enforce policies or patterns of control that disenfranchise segments of the population, such as the poor or minorities. From within the faith, there has arisen the counterdemand that the poor be given a chance to improve and to escape the bonds of poverty. We can see these counterdemands dramatically stated in the Liberation Theology of the 1980s in Latin America. They also are advocated by a variety of Christian groups in the United States (Hart [1992] 1996).

In the end, such processes indicate the reinvigoration of public religion around the world. For Casanova (1994), even Islamic movements can be seen in this light and should not be consigned to a peripheralized "fundamentalist reaction" to modernization. Whether progressive or conservative, public religions' calls for a new vision of politics and economic life are an essential characteristic of the present era. We cannot, then, think of religion as circumscribed by the private sphere, Casanova argues, and must not neglect the public roles that religions play if we are to perceive the modern world clearly.

In sum, religion and family, both public and private, are essential components of modern life. If anything, their future roles seem to be growing rather than shrinking. Biomedical advances promise to play out most dramatically the questions of life and death, spirit and body in the public church and the public family, which will have to adapt to longer life and, perhaps, longer periods of dependency for the elderly. The need to prepare children for appropriate careers in the new fields being created by technological advances will have to be addressed in the personal councils around the kitchen table that always so strongly influence the occupational choices that young people make. Those occupational choices will rely, in some indirect but profound ways, on the economic ethic that is espoused in the religious community that the family adheres to. In the Third World, development hinges, to some degree, on the fertility decisions made in the family under the influence

of the faith that defines appropriate sexual and reproductive behavior. In the process of economic and political change that so many societies experience as disruption and disorganization, family and religion provide important sources of stability and order, even as they adapt to the changing circumstances they find themselves in. It is there that one should go to find the processes working out the morality for the new age and the lifestyles for the new era.

## REFERENCES

- Alexander, Jeffrey C., and Paul Colomy, eds. 1990 *Differentiation Theory and Social Change: Comparative and Historical Perspectives*. New York: Columbia University Press.
- Ammerman, Nancy Tatom 1997 "Religious Choice and Religious Vitality: The Market and Beyond." In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Ammerman, Nancy Tatom, and Wade Clark Roof, eds. 1995 *Work, Family, and Religion in Contemporary Society*. New York: Routledge.
- Becker, Gary S. 1973 "A Theory of Marriage: Part I." *Journal of Political Economy* 81(4):413-446.
- 1974 "A Theory of Marriage: Part II." *Journal of Political Economy* 82(2):511-526.
- 1976 *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- 1981 *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.
- Becker, Gary S., Elisabeth M. Landes, and Robert T. Michael 1977 "An Economic Analysis of Marital Instability." *Journal of Political Economy* 85(6):1141-1187.
- Beckford, James A. 1989 *Religion and Advanced Industrial Society*. London: Unwin Hyman.
- Berger, Peter L. 1967 "Religious Institutions." In Neil J. Smelser, ed., *Sociology: An Introduction*. New York: John Wiley and Sons.
- 1969 *The Sacred Canopy: Elements of a Sociological Theory of Religion*. Garden City, N.Y.: Doubleday Anchor Books.
- Berryman, Phillip 1987 *Liberation Theology: The Essential Facts About the Revolutionary Movement in Latin America*. New York: Pantheon Books.
- Beyer, Peter 1994 *Religion and Globalization*. London: Sage.

- Cammack, Mark, Lawrence A. Young, and Tim B. Heaton 2000 "The State, Religion, and the Family in Indonesia: The Case of Divorce Reform." In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Carroll, Michael P. 1996 "Stark Realities and Androcentric/Eurocentric Bias in the the Sociology of Religion." *Sociology of Religion* 57(3):225-239.
- Carter, Lewis F., ed. 1996 *The Issue of Authenticity in the Study of Religions* (Religion and the Social Order Series, Volume 6). Greenwich, Conn.: JAI Press.
- Casanova, Josè 1994 *Public Religions in the Modern World*. Chicago: University of Chicago Press.
- Chaves, Mark 1995 "On the Rational Choice Approach to Religion." *Journal for the Scientific Study of Religion* 34(1):98-104.
- Cherlin, Andrew 1996 *Public and Private Families: An Introduction*. New York: McGraw-Hill.
- Christiano, Kevin J. 2000 "Religion and Family in Modern American Culture." In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Cochran, Clarke E. 1990 *Religion in Public and Private Life*. New York: Routledge.
- Cousineau, Madeleine, ed. 1998 *Religion in a Changing World: Comparative Studies in Sociology*. Westport, Conn.: Praeger.
- D'Antonio, William V., and Joan Aldous, eds. 1983 *Families and Religions: Conflict and Change in Modern Society*. Beverly Hills, Calif.: Sage.
- D'Antonio, William V., William A. Newman, and Stuart A. Wright 1982 "Religion and Family Life: How Social Scientists View the Relationship." *Journal for the Scientific Study of Religion* 21(3):218-225.
- Demerath, N. J., III 1995 "Rational Paradigms, A-Rational Religion, and the Debate over Secularization." *Journal for the Scientific Study of Religion* 34(1):105-112.
- Demos, John 1970 *A Little Commonwealth: Family Life in Plymouth Colony*. Oxford: Oxford University Press.
- Diduk, Susan E., and Kent Maynard 2000 "A Woman's Pillow and the Political Economy of Kedjom Family Life in Cameroon." In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Durkheim, Emile 1973 *Emile Durkheim: On Morality and Society; Selected Writings* Chicago: University of Chicago Press.
- Ellison, Christopher G. 1995 "Rational Choice Explanations of Individual Religious Behavior: Notes on the Problem of Social Embeddedness." *Journal for the Scientific Study of Religion* 34(1):89-97.
- Finke, Roger, and Rodney Stark 1992 *The Churching of America, 1776-1990: Winners and Losers in Our Religious Economy*. New Brunswick, N.J.: Rutgers University Press.
- Fortuny Loret de Mola, Patricia 2000 "Converted Women Redefining Their Family Roles In Mexico." In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Godsell, Gillian 1999 "Religious and Familial Networks as Entrepreneurial Resources in South Africa." In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Gouldner, Alvin W., and Helen P. Gouldner 1963 *Modern Sociology: An Introduction to the Study of Human Interaction*. New York: Harcourt, Brace and World.
- Greeley, Andrew M. 1989 *Religious Change in America*. Cambridge, Mass.: Harvard University Press.
- 1992 "Religion in Britain, Ireland and the USA." In R. Jowell, L. Brook, G. Prior, and B. Taylor, eds., *British Social Attitudes: The 9th Report*. Social and Community Planning Research. Brookfield, Vt.: Dartmouth Publishing.
- Grossbard, Amyra 1978 "Towards a Marriage Between Economics and Anthropology and a General Theory of Marriage." *American Economic Review* 68(2):33-37.
- Hadden, Jeffrey K. 1987 "Toward Desacralizing Secularization Theory." *Social Forces* 65:587-611.
- 1995 "Religion and the Quest for Meaning and Order: Old Paradigms, New Realities." *Sociological Focus* 28:83-100.
- 1996 "Foreword" to the paperback edition. In Rodney Stark and William Sims Bainbridge *A Theory of Religion*. New Brunswick, N.J.: Rutgers University Press.
- Hadden, Jeffrey K., and Anson Shupe 1986 "Introduction." In Jeffrey K. Hadden and Anson Shupe, eds., *Prophetic Religions and Politics: Religion and the Political Order*, vol. 1. New York: Paragon House.

- Hammond, Phillip E. 1985 *The Sacred in a Secular Age: Toward Revision in the Scientific Study of Religion*. Berkeley: University of California Press.
- Hargrove, Barbara 1979 *The Sociology of Religion: Classical and Contemporary Approaches*. Arlington Heights, Ill.: AHM Publishing.
- 1985 “Gender, the Family, and the Sacred.” In Phillip E. Hammond, ed., *The Sacred in a Secular Age: Toward Revision in the Scientific Study of Religion*. Berkeley: University of California Press.
- Hart, Stephen (1992) 1996 *What Does the Lord Require: How American Christians Think about Economic Justice*. New Brunswick, N.J.: Rutgers University Press.
- Houseknecht, Sharon K. 2000 “Social Change in Egypt: The Roles of Religion and Family.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Houseknecht, Sharon K., and Jerry G. Pankhurst, eds. 2000 *Family, Religion, and Social Change in Diverse Societies*. New York: Oxford University Press.
- Huber, Joan, and Glenna Spitze 1980 “Considering Divorce: An Expansion of Becker’s Theory of Marital Instability.” *American Journal of Sociology* 86:75–89.
- Iannaccone, Laurence R. 1995 “Voodoo Economics? Reviewing the Rational Choice Approach to Religion.” *Journal for the Scientific Study of Religion* 34(1):76–89.
- 1997 “Rational Choice: Framework for the Scientific Study of Religion.” In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Jarvis, John 2000 “Mormonism in France: The Family as a Universal Value in Globalizing Religion.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Johnson, Robert Alan 1980 *Religious Assortative Marriage in the United States*. New York: Academic Press.
- Luckmann, Thomas 1967 *The Invisible Religion*. New York: Macmillan.
- MacIver, Robert M. 1970 *On Community, Society and Power*. Chicago: University of Chicago Press.
- Martin, David 1990 *Tongues of Fire: The Explosion of Protestantism in Latin America*. Oxford: Basil Blackwell.
- Marty, Martin 1993 Review of *The Churching of America, 1776–1990: Winners and Losers in Our Religious Economy* (Rutgers University Press, 1992). *Christian Century*, January 27, p. 88.
- Marty, Martin E., and R. Scott Appleby 1995 *The Fundamentalism Project*. Chicago: University of Chicago Press.
- Misztal, Bronislaw, and Anson Shupe, eds. 1992 *Religion and Politics in Comparative Perspective: Revival of Religious Fundamentalism in East and West*. Westport, Conn.: Praeger.
- Moore, Wilbert E. 1963 *Social Change*. Englewood Cliffs, N.J.: Prentice-Hall.
- Neitz, Mary Jo, and Peter R. Mueser 1997 “Economic Man and the Sociology of Religion: A Critique of the Rational Choice Approach.” In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Nisbet, Robert A. 1970 *The Social Bond: An Introduction to the Study of Society*. New York: Alfred A. Knopf.
- Pankhurst, Jerry G. 1998 “Russia’s Religious Market: Struggling with the Heritage of Russian Orthodox Monopoly.” In Madeleine Cousineau, ed., *Religion in a Changing World: Comparative Studies in Sociology*. Westport, Conn.: Praeger.
- Pankhurst, Jerry G., and Sharon K. Houseknecht 1983 “The Family, Politics and Religion in the 1980s: In Feat of the New Individualism.” *Journal of Family Issues* 4:5–34.
- Popenoe, David 1988 *Disturbing the Nest: Family Change and Decline in Modern Societies*. New York: Aldine de Gruyter.
- Roberts, Richard H., ed. 1995 *Religion and the Transformations of Capitalism: Comparative Approaches*. London: Routledge.
- Roof, Wade Clark 1993 *A Generation of Seekers: The Spiritual Journey of the Baby Boom Generation*. New York: HarperCollins.
- Rosado Nunes, Maria Jose Fontelas 2000 “Women, Family and Catholicism in Brazil: The Issue of Power.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Rubenstein, Richard L., ed. 1987 *Spirit Matters: The Worldwide Impact of Religion on Contemporary Politics*. New York: Paragon House.
- Settles, Barbara H. 1996 “The International Study of Families and Rapid Social Change: Issues for Family Sociologists and their Professional Organizations.” Paper presented at the annual meeting of the American Sociological Association, August 16–20, New York.
- Sherkat, Darren E. 1997 “Embedding Religious Choices: Integrating Preferences and Social Constraints into

- Rational Choice Theories of Religious Behavior.” In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Shupe, Anson, and Bronislaw Misztal, eds. 1998 *Religion, Mobilization, and Social Action*. Westport, Conn.: Praeger.
- Smith, Christian 1991 *The Emergence of Liberation Theology: Radical Religion and Social Movement Theory*. Chicago: University of Chicago Press.
- Smith, Stephen R. 2000 “Land of the Rising Son? Domestic Organization, Ancestor Worship, and Economic Change in Japan.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Spickard, James V. 1998 “Rethinking Religious Social Action: What is ‘Rational’ About Rational Choice Theory?” *Sociology of Religion* 59(2):99–115.
- Stark, Rodney, and William Sims Bainbridge 1987 *A Theory of Religion*. New York: Peter Lang.
- Swatos, William H., ed. 1992 *Twentieth-Century World Religious Movements in Neo-Weberian Perspective*. Lewiston, N.Y.: Edwin Mellen Press.
- Thomas, Darwin L., ed. 1988 *The Religion and Family Connection: Social Science Perspectives*. Provo, Utah: Religious Studies Center, Brigham Young University.
- Thomas, Darwin L., and Marie Cornwall 1990 “Religion and Family in the 1980s: Discovery and Development.” *Journal of Marriage and the Family* 52:983–992.
- Thomas, Darwin L., and Gwendolyn C. Henry 1985 “The Religion and Family Connection: Increasing Dialogue in the Social Sciences.” *Journal of Marriage and the Family* 47(2):369–379.
- Thornton, Arland 1985 “Reciprocal Influences of Family and Religion in a Changing World.” *Journal of Marriage and the Family* 47(2):381–394.
- Trost, Jan, and Irving Palm 2000 “Family and Religion in Sweden.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Vago, Steven 1989 *Social Change*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Vardomatskii, Andrei, and Jerry G. Pankhurst 2000 “Belarus on the Cusp of Change: The Relationship Between Religion and Family in a Newly Open Religious Market.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Wald, Kenneth 1986 *Religion and Politics in the United States*. New York: St. Martin’s Press.
- Warner, Stephen R. 1993 “Work in Progress Toward a New Paradigm for the Sociological Study of Religion in the United States.” *American Journal of Sociology* 98(5):1044–1093.
- 1997 “Convergence Toward the New Paradigm: A Case of Induction.” In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Weber, Max (1922) 1963 *The Sociology of Religion*. Boston: Beacon Press.
- Williams, Robin M., Jr. (1955) 1970 *American Society: A Sociological Interpretation*, 3rd ed. New York: Alfred A. Knopf.
- Witte, John, Jr., ed. 1993 *Christianity and Democracy in Global Context*. Boulder, Colo.: Westview Press.
- Wuthnow, Robert 1994a *Sharing the Journey: Support Groups and America’s New Quest for Community*. New York: Free Press.
- 1994b *Producing the Sacred: An Essay on Public Religion*. Urbana: University of Illinois Press.
- Yang, Li-shou, Arland Thornton, and Thomas Fricke 2000 “Religion and Family Formation in Taiwan: The Decline of Ancestral Authority.” In Sharon K. Houseknecht and Jerry G. Pankhurst, eds., *Family, Religion and Social Change in Diverse Societies*. New York: Oxford University Press.
- Young, Lawrence A., ed. 1997 *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.

JERRY G. PANKHURST  
SHARON K. HOUSEKNECHT

## FAMILY LAW

Family law is that body of law having to do with creating, ordering, and dissolving marital and family groups. Although the exact scope of family law is given differently by different authors, at its core family law is concerned with such issues and events as marriage, separation, divorce, alimony, custody, child support, and adoption, as well as the more arcane topics of annulment, paternity, legitimacy, artificial insemination, and surrogate parenting.

This entry on family law in the United States should be read with two important caveats in



mind. First, it is somewhat misleading to write of “United States” family law. Because the power to regulate domestic life is not one of the powers delegated to the federal government by the Constitution, in the United States most family law has been “a virtually exclusive province of the states” (*Sosna v. Iowa*, 419 U.S. 393 [1975]). Despite considerable variation in state law, however, certain general trends can be identified. Moreover, researchers have identified similar trends in a number of European countries (Glendon 1989).

Second, many kinds of law that affect the family cannot be discussed here. These include laws that are not ordinarily listed under the rubric of “family law” but that have significant effects on family life in this country. These range from the laws of inheritance to zoning regulations and regulations about social welfare programs. While the impact on the family of a diversity of laws seems to become increasingly significant, this is not a uniquely modern phenomenon. For example, 200 years ago poor laws affected family life in ways that anticipated the impact of modern welfare laws (tenBroek 1964).

As was typical of much of early law in this country, most American family law was received from English law; but family law was atypical in that much of it was not derived from secular or “temporal” English law. In England, from the late twelfth century (Pollock and Maitland 1898, Vol. 2, p. 367) until the passage of the Matrimonial Causes Act of 1857, issues pertaining to marriage and divorce were governed by canon law, and most family matters were thus subject to the jurisdiction of ecclesiastical courts. While the American colonies had no ecclesiastical courts, English canonical rules concerning family relations were incorporated—either by statute or by common law tradition—into the laws of the colonies and, later, the states (Clark 1980).

Notwithstanding its religious heritage, family law in this country was completely secular. Although marriages were frequently performed by members of the clergy, the authority to solemnize marriages was vested in them by the state, not the church. In legal theory, at the basis of the family

was a marriage that was a civil contract and not a religious sacrament.

This contractual view of marriage had some interesting consequences. For example, it led to official recognition of informal as well as formal marriage. This informal union, the so-called common law marriage, was effected by the simple express agreement of a man and a woman to be married, followed by their cohabitation. (Contrary to popular myth, common law marriages did not require a specific number of years to go into effect.) While today they are recognized only in fourteen states, until the twentieth century common law marriages were as valid as formal marriages in nearly every state (Wardle et al. 1988, § 3:17). Recognition of common law marriage meant that settlers on the geographic fringes of society, without access to officials, could enjoy the same protection of their property rights and their children’s legitimacy as was afforded in formal marriages. In 1833, Chief Justice Gibson of Pennsylvania ruled that rigid marriage laws were “ill adapted to the habits and customs of society as it now exists.” Not recognizing common law marriage, or so Gibson suggested, would “bastardize the vast majority of children which have been born within the state for half a century” (*Rodenbaugh v. Sanks*, 2 Watts 9).

After the Civil War, there was a movement to strengthen state regulation of marriage. Most states already required marriage licenses, but in antebellum America, courts had treated these licenses as a means “to register, not to restrict marriage” (Grossberg 1985, p. 78). By the end of the nineteenth century, however, marriage licenses had clearly become a means of social control. Because the process of acquiring a marriage license brought the couple under scrutiny of some official, licensing requirements helped states prevent marriages of people who were too young or too closely related, either by blood (consanguinity) or marriage (affinity). Official scrutiny of those seeking to wed also helped to enforce laws against bigamy and polygamy.

But legislators, encouraged by eugenicists who believed that crime, mental illness, and other social ills could be traced to hereditary biological

factors, also enacted laws enumerating other kinds of forbidden marriages. For example, marriage was prohibited to those not mentally capable of contracting owing to conditions variously labeled as insanity, lunacy, idiocy, feeble-mindedness, imbecility, or unsound mind (Clark 1968, pp. 95–96). Marriage was also prohibited to those physically incapable of performing the “marriage essentials.” Generally, this latter criterion involved only the capacity to have “normal” or “successful” sexual intercourse, not necessarily the ability to procreate. As one author explained it, “Copula, not fruitfulness, is the test” (Tiffany 1921, p. 29).

Eugenics also justified, scientifically, laws that prohibited people with certain diseases (e.g., epilepsy, tuberculosis, and venereal disease) and statuses (e.g., habitual criminal, rapist) from marrying. In most cases, such obstacles could be overcome only if the person consented to sterilization. Many believed such statutes were necessary to “prevent the demise of civilized-society” (Linn and Bowers 1978, p. 629). Even some of the most respected legal thinkers joined the eugenicists. Justice Oliver Wendell Holmes of the United States Supreme Court, for example, wrote that it would be “better for all the world, if instead of waiting to execute degenerate offspring for crimes, or to let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind” (*Buck v. Bell*, 274 U.S. 200 [1927]).

The most notorious marriage impediment was race. By 1930, thirty states had enacted statutes prohibiting interracial marriages (Clark 1968, p. 91). For the most part, these antimiscegenation laws forbade marriages between whites and blacks, but in several cases the prohibition was extended to, for example, white and Malays, whites and Mongolians, whites and Native Americans, and blacks and Native Americans (Kennedy 1959, pp. 59–69).

Divorce was even more strictly regulated than marriage. However, the absence of ecclesiastical restrictions made divorces much easier to obtain in the United States than in England. This was especially true in the northern states. A few states even allowed divorce simply where the cause

seemed “just and reasonable.” Connecticut, for example, permitted divorce for conduct that “permanently destroys the happiness of the petitioner and defeats the purpose of the marriage relation” (Clark 1968, p. 283). During the latter part of the nineteenth century such generous statutes were repealed, and divorce was allowed only in response to specific types of fault—usually adultery, desertion, cruelty, or long-term imprisonment.

Despite stringent regulation of entrance to and exit from marriage, husbands and wives in intact marriages were generally protected from legal scrutiny. Indeed, traditionally, the principle of nonintervention was so strong that neither husbands nor wives could invoke the law to resolve marital disputes even when they wished to. In one case, for example, the wife of a well-to-do but stingy husband asked the Nebraska courts to require him to pay for indoor plumbing and to provide a reasonable allowance to her. The court agreed that, given his “wealth and circumstances,” the husband’s attitude “leaves little to be said in his behalf.” But, said the court, “the living standards of a family are a matter of concern to the household and not for the courts to determine” (*McGuire v. McGuire*, 157 Neb. 226, 59 N.W.2d 336 [1953]). Similarly, the courts preferred a hands-off approach to parent–child relationships. As the United States Supreme Court ruled in 1944, “the custody, care, and nurture of the child reside first in the parents, whose primary function and freedom include preparation for obligations the state can neither supply nor hinder. . . . It is in recognition of this that [earlier] decisions have respected the private realm of family life which the state cannot enter” (*Prince v. Massachusetts*, 321 U.S. 158 [1944]).

The extent of the courts’ reluctance to intervene in family matters or, as it was sometimes put, to “disrupt family harmony,” was shown in the rule that spouses could not sue one another for personal torts or injuries. If, for example, a husband assaulted or battered his wife, she was enjoined from taking legal action against him in civil court (Keeton et al. 1984, pp. 901–902). In theory, the husband could be prosecuted in criminal court, but police and criminal courts too were reluctant to interfere in domestic matters (Pleck 1987, p. 187).

The practice of nonintervention was carried a step further at the turn of the century when the courts invented the doctrine of “parental immunity.” Owing to reasons of “sound public policy, designed to subserve the repose of families and the best interests of society” (*Hewellette v. George*, 68 Miss. 703, 9 So. 885 [1891]), an unemancipated minor was barred from suing his or her parents for negligent or intentional wrongdoing.

Owing to the state’s reluctance to intervene, the family has had a great deal of autonomy in this country, even to the extent that some have referred to the family as a “minisovereignty” (O’Donnell and Jones 1982, p. 7). In recent times, this autonomy has been justified on the basis of privacy rights. Speaking of the married couple’s right to make decisions about the use of contraception, the United States Supreme Court said in 1965, for example, “we deal with a right of privacy older than the Bill of Rights” (*Griswold v. Connecticut*, 381 U.S. 479 [1965]).

But things began to change in the late twentieth century. First, beginning in the 1960s, strict regulation of entrance to and exit from marriage began to unravel. In the 1967 case of *Loving v. Virginia*, the United States Supreme Court ruled unconstitutional all antimiscegenation laws, saying that the states had no right to “prevent marriages between persons solely on the basis of racial classification.” “Marriage,” said the Court, “is one of the ‘basic civil rights of man,’ fundamental to our very existence and survival” (388 U.S. 1; quoting *Skinner v. Oklahoma*, 316 U.S. 535 [1942]). Since *Loving*, many other marriage restrictions have been repealed or eased. Age requirements in many states have been lowered; the mental ability needed to contract marriage has been ruled to be less than that required for other sorts of contracts; and the necessary mental competency is presumed to be present unless there is “clear and definite” proof to the contrary. Moreover, “there is a trend in modern times to abolish affinity restrictions” (Wardle et al. 1988, § 2:09); only one state (Missouri) still prohibits epileptics to marry (Wardle et al. 1988, § 2:47), and in many states, even prison inmates are deemed to have a right to marry (*In re Carrafa*, 77 Cal. App.3d 788 [1978]).

These changes reflect the courts’ willingness to protect the rights of individuals to make their own choices about marriage and related matters. The decision to marry, according to the Supreme Court, is among “the personal decisions protected by the right to privacy” (*Zablocki v. Redhail*, 434 U.S. 374 [1978]).

Presumably, much the same can be said about the decision to divorce; recent changes in divorce laws have, if anything, been even more dramatic than changes in marriage laws. Implicitly accepting the principle that there is a right to divorce, the Supreme Court ruled in 1971 that welfare recipients could not be denied access to divorce courts because they could not afford to pay court costs and fees (*Boddie v. Connecticut*, 101 U.S. 371 [1971]). By the mid-1980s, every state had either replaced fault-based divorce laws with no-fault laws or added no-fault grounds to existing laws (Freed and Walker 1986, p. 444). No longer, then, must there be a “guilty” and an “innocent” party in a divorce. Instead, one spouse simply needs to assert that the couple is no longer getting along or has been living apart for a certain amount of time.

While regulations governing entrance to and exit from marriage and family life have decreased, there has been a corresponding increase in regulations affecting relations in ongoing families. Spousal immunity has been abolished in most states. Moreover, in many states the law recognizes the crime of “marital rape.” Similarly, children now have more rights that can be asserted against their parents. For example, minors have the right to obtain information about and to use birth control without a parent’s consent (*Carey v. Population Services International*, 431 U.S. 678 [1977]); to receive psychiatric care (*In re Alyne E.*, 113 Misc. 2d 307, 448 N.Y.S.2d 984 [1982]); and perhaps even to separate from their parents should the parents and children prove “incompatible” (*In re Snyder*, 85 Wash. 2d 182, 532 P.2d 278 [1975]). At base, says the Supreme Court, children “are ‘persons’ under the Constitution” and have rights that should be protected by the state (*Tinker v. Des Moines Independent School District*, 393 U.S. 503 [1969]).

Both the easing of marriage and divorce restrictions and the loss of family autonomy can be

traced to the growth of individual rights that began in the 1960s. The idea of family autonomy and privacy and, hence, the policy of nonintervention were traditionally based on “paternal” authority; the authority of the family patriarch. This pattern can be traced back to the Roman idea of *patria potestas*—or the right of the father to exert absolute control over his family, including the power of life or death.

Family autonomy and privacy that is based on paternal power is viable only when other members of the family are unable to invoke the power of the state against the father. It was for this reason, then, that traditionally the woman’s power to invoke the law was suspended from the moment of her marriage (Blackstone [1769] 1979, Vol. 1, p. 430). Children, likewise, had no legal standing until they reached the age of majority.

Things are much different today: While children still have many “legal disabilities,” they can no longer be regarded as chattel. Women have achieved at least technical legal equality (though whether this has served to their advantage in divorce law is still subject to debate—compare Weitzman 1985 and Jacob 1988). Although the courts still speak of “family privacy,” it is becoming clear that such privacy is based on family members’ individual rights and exists only as long as family members are not in serious conflict about how they wish to assert those rights.

Some mourn the loss of near total family autonomy; the family, they say, has lost its integrity (Peirce 1988). There is no doubt that the notion of family autonomy or privacy served an important value: It has been “a convenient way for dealing with a problem . . . [that is] especially acute in the United States—that of devising family law that is suited to the needs and desires of persons with different ethnic and religious backgrounds, different social status, and different standards of living” (Glendon 1989, p. 95). In many instances, however, nonintervention created private Hobbesian jungles in which the strong ruled and the weak could not call upon the law for help.

As we move into the twenty-first century, families will continue to play an important role in

society, and there can be little doubt that family relations will continue to be regarded as legally different from other relations and worthy of special legal protection. The question is, To whom is the law’s protection to be extended in domestic matters as the United States embarks upon the twenty-first century? Traditionally, lawmakers have extended this protection to a limited variety of relations—the father–mother–children household. If present trends continue, however, the traditional ideal-typical nuclear family will be something that is achieved (and perhaps aspired to) by only a small fraction of Americans (Difonzo 1997; Estlund and Nussbaum 1998; McIntyre and Sussman 1995; Reagin 1999).

As we move through the first decades of this new century, new and more complex family legal issues will emerge as people construct new communal arrangements, call them family, and seek the protections accorded by the law to more traditional arrangements (Dolgin 1999; Edwards 1999; Minow 1993). The most pressing question facing lawmakers is this: Will the law continue to afford its protections only to those domestic arrangements that mirror traditional family *forms*, or will it embrace and protect domestic arrangements insofar as they fulfill traditional family *functions*?

(SEE ALSO: *Family and Household Structure*)

#### REFERENCES

- Blackstone, William (1769) 1979 *Commentaries on the Laws of England*, 4 vols. Chicago: University of Chicago Press.
- Clark, Homer H., Jr. 1968 *Law of Domestic Relations*. St. Paul, Minn.: West.
- 1980 *Cases and Problems on Domestic Relations*. 3rd ed. St. Paul, Minn.: West.
- Difonzo, J. Herbie 1997 *Beneath the Fault Line: The Popular and Legal Culture of Divorce in the Twentieth Century*. Charlottesville: University Press of Virginia.
- Dolgin, Janet L. 1999 *Defining the Family: Law, Technology and Reproduction in an Uneasy Age*. New York: New York University Press.
- Edwards, Jeanette 1999 *Technologies of Procreation: Kinship in the Age of Assisted Conception*. New York: Routledge.

- Estlund, David M. and Martha C. Nussbaum 1998 *Sex, Preference, and Family: Essays on Law and Nature*. New York: Oxford University Press.
- Freed, Doris J., and Timothy B. Walker 1986 "Family Law in the Fifty States: An Overview." *Family Law Quarterly* 20:439-587.
- Glendon, Mary A. 1989 *The Transformation of Family Law, State, Law, and Family in the United States and Western Europe*. Chicago: University of Chicago Press.
- Grossberg, Michael 1985 *Governing the Hearth: Law and Family in Nineteenth-Century America*. Chapel Hill: University of North Carolina Press.
- Jacob, Herbert 1988 *Silent Revolution*. Chicago: University of Chicago Press.
- Keeton, W. Page, Dan B. Dobbs, Robert E. Keeton, and David G. Owen 1984 *Prosser and Keeton on the Law of Torts*, 5th ed. St. Paul, Minn.: West.
- Kennedy, Stetson 1959 *The Jim Crow Guide to the USA*. London: Lawrence and Wishart.
- Linn, Brian J., and Lesly A. Bowers 1978. "The Historical Fallacies Behind Legal Prohibitions of Marriages Involving Mentally Retarded Persons: The Eternal Child Grows Up." *Gonzaga Law Review* 13:625-690.
- Long, Joseph R. 1905 *A Treatise on the Law of Domestic Relations*. St. Paul, Minn.: Keefe-Davidson.
- McIntyre, Lisa J. and Marvin Sussman 1995 *Families and Law*. New York: Haworth Press.
- Minow, Martha 1993 *Family Matters: Readings on Family Lives and the Law*. New York: New Press.
- O'Donnell, William J., and David A. Jones 1982 *The Law of Marriage and Marital Alternatives*. Lexington, Mass.: D. C. Heath.
- Peirce, Dorothy S. 1988 "BRI v. Leonard: The Role of the Courts in Preserving Family Integrity." *New England Law Review* 23:185-219.
- Pleck, Elizabeth 1987 *Domestic Violence: The Making of American Social Policy Against Family Violence from Colonial Times to the Present*. New York: Oxford University Press.
- Pollock, Frederick, and Frederic W. Maitland 1898 *The History of English Law Before the Time of Edward I*, 2 vols. Cambridge: Cambridge University Press.
- Reagin, Milton C. *Alone Together: Law and the Meanings of Marriage*. New York: Oxford University Press.
- tenBroek, Jacobus 1964 *Family Law and the Poor*. Westport, Conn.: Greenwood.

- Tiffany, Walter C. 1921 *Handbook on the Law of Personal and Domestic Relations*. 3rd ed. St. Paul, Minn.: West.
- Wardle, Lynn D., Christopher L. Blakesley, and Jacqueline Y. Parker 1988 *Contemporary Family Law: Principles, Policy, and Practice*, 4 vols. Deerfield, Ill.: Callaghan.
- Weitzman, Lenore 1985 *The Divorce Revolution*. New York: Free Press.

LISA J. MCINTYRE

## FAMILY PLANNING

The ability of couples to plan the size of their family and the timing of births has important sociological implications for both individual families and society as a whole. Women's roles and labor-force participation, the socialization of children, social and economic development, and ultimately the ability of the earth to sustain human life are all affected in one way or another by the ability of couples to practice family planning and the success with which they do so. In the United States, women expect to complete their childbearing with an average of 2.2 children per woman (Abma et al. 1997), and, on average, women have 2.0 births over their lifetime (Ventura et al.). Throughout the world, the average number of children desired varies from about two in most industrialized nations to between six and eight in some African nations (Alan Guttmacher Institute 1995). In order to limit lifetime births to the number desired, couples must abstain from intercourse, have high levels of contraceptive use, or resort to abortion. Indeed, sexually active women would average eighteen births over their lifetime if they used no contraception and no induced abortion (Harlap et al. 1991). This article summarizes information regarding sexual activity; the risk and occurrence of unplanned pregnancy; contraceptive use and failure; and the provision of family planning-related information, education, and services in the United States. For comparison, worldwide variation in the planning status of pregnancies and births and the use and availability of contraception are also presented.

### EXPOSURE TO THE RISK OF PREGNANCY

Most Americans begin to have intercourse during their late adolescence and continue to be sexually

active throughout their reproductive lives. In 1995, 55 percent of all men aged 15–19 in the United States had had intercourse (Sonenstein et al. 1998). Similarly, about half of all women aged 15–19 report ever having had sex (Table 1). These data, collected in 1995, indicate a leveling off in the trend toward earlier ages of sexual initiation. Whereas the percentage of adolescents who reported being sexually experienced rose steadily throughout most of the 1980s, the percentage of adolescent females who had ever had sex did not change significantly between 1988 and 1995 and the percentage of adolescent males who were sexually experienced actually fell during that period (Singh and Darroch 1999).

Once sexually active, most women become at risk for an unintended pregnancy. Table 1 shows information on the percentage of all U.S. women aged 15–44 who are at risk for becoming pregnant by age and union-status groups (currently married, cohabiting, formerly married or never married). The proportion who are not at risk of an unintended pregnancy because they have never had intercourse decreases quickly from 50 percent of teenagers 15–19 to only 1 percent of all women aged 35–44. Five to seven percent of women in all age groups have had intercourse but are not currently in a sexual relationship (i.e., they have not had sex within the last three months). Some 5 percent of women are infertile, or noncontraceptively sterile, because of illness, surgery (that was not for contraceptive purposes), menopause, or some other reason. The proportion that is infertile increases steadily with age, from 1 to 2 percent of women under 30 to about 13 percent of those aged 40–44. Some women, especially those in their twenties and early thirties, are not at risk of an *unintended pregnancy* because they are already pregnant, postpartum, or seeking pregnancy. Eleven to fifteen percent of women aged 20–34 are in this category.

Women who are at risk for an unintended pregnancy account for more than two-thirds of all women ages 15–44 at any point in time. Women at risk are those who are currently in a sexual relationship, are fertile, and wish to avoid becoming pregnant. The proportion of women at risk of

unintended pregnancy increases from less than 40 percent of teenagers to about three-quarters of all women ages 25–44. Women who are currently married or cohabiting are most likely to be at risk for unintended pregnancy—81 to 83 percent of them are at risk, compared with 72 percent of formerly married women and 49 percent of never-married women. The most common reason some married or cohabiting women are not at risk of unintended pregnancy is that they are pregnant, postpartum, or trying to become pregnant. Among women who have never been married, never having had intercourse or no recent intercourse are the most common reasons.

### OCCURRENCE OF UNINTENDED PREGNANCY

Nearly one-half of all pregnancies (49 percent) in the United States are unintended (Henshaw 1998), that is, they occur to women who want to have a baby later but not now (generally called “mistimed”) or to women who did not want to have any (more) children at all (called “unwanted”) (Table 2). The proportion of pregnancies that are unintended is highest among adolescents—78 percent—and varies considerably by age. The percentage of pregnancies that are unintended is lowest among women aged 30–34 (33 percent) and rises again among older women to 51 percent among women aged 40 and older. Unintended pregnancies are also relatively more likely to occur among never-married women (78 percent), black women (72 percent), and low-income women (61 percent for women under 100 percent of the federal poverty level).

The percentage of pregnancies that are unintended has declined in recent years—from 57 percent in 1987 to 49 percent in 1994 (see Table 2). These declines have occurred across all age categories but have been more significant among older women. These declines have also been more significant among low-income women. In 1987, 75 percent of all pregnancies to women with family incomes under 100 percent of the poverty level were unintended. This dropped to 61 percent in 1994. In comparison, the percentage of unintended pregnancies to women with incomes 200 percent or more of the federal poverty level fell from 45 percent in 1987 to 41 percent in 1994.

**Percentage Distribution of Women Aged 15–44 According to Exposure to the Risk of Unintended Pregnancy, by Age and Union Status, National Survey of Family Growth, 1995**

	Total	AGE						UNION STATUS			
		15–19	20–24	25–29	30–34	35–39	40–44	Currently Married	Cohabiting	Formerly Married	Never Married
Total	100	100	100	100	100	100	100	100	100	100	100
<b>At risk of unintended pregnancy</b>	<b>69</b>	<b>37</b>	<b>70</b>	<b>74</b>	<b>77</b>	<b>77</b>	<b>77</b>	<b>81</b>	<b>83</b>	<b>72</b>	<b>49</b>
Using contraception	64	30	64	69	73	73	71	76	78	67	43
Not using contraception	5	7	6	5	4	4	5	4	5	6	7
<b>Not at risk</b>	<b>31</b>	<b>63</b>	<b>31</b>	<b>26</b>	<b>23</b>	<b>23</b>	<b>24</b>	<b>19</b>	<b>18</b>	<b>28</b>	<b>51</b>
Infertile	5	1	1	2	4	9	13	6	5	8	2
Pregnant/postpartum seeking pregnancy	9	5	11	15	12	7	3	13	12	4	3
No recent intercourse*	6	7	7	6	5	6	7	1	1	16	13
Never had intercourse	11	50	12	4	3	1	1	0	0	0	33

Table 1

SOURCE: Alan Guttmacher Institute tabulations of the 1995 National Survey of Family Growth (Cycle V).

NOTE: \*Have not had intercourse in the past three months.

Among all unintended pregnancies, more than half (54 percent) end in an abortion while 46 percent result in an unintended birth. This relationship differs for adolescents, who, in recent years, have been more likely to resolve unintended pregnancies with a birth. More than one-half of all unintended pregnancies to adolescents result in an unintended birth (55 percent), while 45 percent are resolved with an abortion. These percentages represent a significant change in the resolution of unintended pregnancies among adolescents. Throughout the 1980s, adolescents who were pregnant unintentionally were more likely to obtain an abortion (55–53 percent) than to carry the pregnancy to term.

Nearly half (48 percent) of all women aged 15–44 have had at least one unintended pregnancy at some time in their lives; 28 percent have had one or more unplanned births, 30 percent have had one or more abortions, and 11 percent have had both. Given current rates of pregnancy and abortion, by the time they are 45 years old, the typical woman in the United States will have experienced 1.42 unintended pregnancies and 43 percent will have had an abortion.

Women who are using no contraceptive method account for about 8 percent of all women at risk

of unintended pregnancy, but, because they are more likely to become pregnant than are those using a method, they account for nearly one-half of all unplanned pregnancies, an estimated 47 percent. Significant reductions in unintended pregnancy and abortion could occur with increased contraceptive use, with more effective use of existing methods, and with the development and marketing of additional methods.

**CONTRACEPTIVE USE**

Women and men in the United States rely on a variety of contraceptive methods to plan the timing and number of children they bear and to avoid unintended pregnancies. Surgical contraceptive sterilization is available to both men and women. Oral contraceptives, Depo Provera injectibles, Norplant implants, the IUD, and female barrier methods such as the diaphragm and the cervical cap are available from physicians and clinic providers. Other methods—condoms and spermicidal foam, cream, jelly, and film—can be purchased over the counter in pharmacies or other stores. Instruction in periodic abstinence is available from physicians and other family planning providers as well as through classes where only that method is taught.

More than nine in ten women aged 15–44 in 1995 who were at risk of unintended pregnancy

**Percentage of All Pregnancies (Excluding Miscarriages) That are Unintended by Women's Age, Marital Status, Race, Ethnicity and Poverty Status, 1987 and 1995**

WOMEN'S CHARACTERISTICS	% OF ALL PREGNANCIES THAT ARE UNINTENDED	
	1987*	1994**
All women	57.3%	49.2%
Age		
15-19	81.7%	78.0%
20-24	60.6%	58.5%
25-29	45.2%	39.7%
30-34	42.1%	33.1%
35-39	55.9%	40.8%
40-44	76.9%	50.7%
Marital status		
Currently married	40.1%	30.7%
Formerly married	68.5%	62.5%
Never married	88.2%	77.7%
Race		
White	NA	42.9%
Black	NA	72.3%
Other	NA	50.0%
Ethnicity		
Hispanic	NA	48.6%
Non-Hispanic	NA	49.3%
Poverty status		
<100% poverty	75.4%	61.4%
100-199% poverty	64.0%	53.2%
200+% poverty	45.0%	41.2%

Table 2

SOURCE: \*Forrest (1994); \*\*Henshaw (1998).

NOTE: na=not available.

were using a contraceptive method, as shown in Table 3. Thirty-six percent relied on contraceptive sterilization of themselves or their partner, 52 percent used reversible medical methods, 5 percent used nonmedical methods such as withdrawal and periodic abstinence, and 7.5 percent were currently using no contraceptive, even though they were at risk of unintended pregnancy.

Patterns of contraceptive use differ by age. Younger women at risk of unintended pregnancy are more likely than older women to use no method of contraception. Nearly one in five teenage women at risk use no method, compared to 6 to 7 percent of women at risk aged 25 and older. The proportion using reversible medical methods declines steeply with age—from more than four out of five women aged 20-24 to less than one-quarter

of those ages 40-44. Oral contraceptives are the most commonly used method among women under 30, used by 35 to 48 percent of these women. Condoms are second in popularity among this age group, used by 23 to 30 percent of women. Although fewer than 3 percent of women at risk use Depo Provera injectible contraceptives, this method has grown in popularity since its introduction into the United States, particularly among young women. Eight percent of teenagers at risk used this method. As women become older and complete their families, male and female contraceptive sterilization become increasingly common, rising steeply from 5 percent of women at risk aged 20-24 to one in five women in their late twenties and to two out of three women aged 40-44. Among women in their twenties, female sterilization is about four times more common than vasectomy. The margin narrows among older women to between two and two and a half times more common.

The proportion of women at risk of unintended pregnancy who use no contraceptive method is highest among never-married women, 14 percent as compared to 5 percent of those who are currently married or cohabiting and 8 percent of formerly married women. Sterilization is the most frequently used method among women who are currently married (46 percent) as well as formerly married women (50 percent). The pill is the most commonly used method among never-married women (38 percent) and cohabiting women (34 percent). Condoms are most likely to be used by never-married women (28 percent).

Although poor women and minority women at risk of unintended pregnancy have, in the past, been more likely than higher-income and non-Hispanic white women to be using no contraceptive method, these differences have lessened. Compared to the 1980s, in 1995 there were no significant race/ethnicity or poverty differences in the percentages of women at risk of unintended pregnancy who used no method of contraception. However, there is some variation in the types of methods used among these subgroups. Low-income women are less likely to rely on reversible methods and more likely to rely on sterilization than higher income women. Forty percent of women at risk of unintended pregnancy who are under



FAMILY PLANNING

**Percentage Distribution of Women at Risk of Unintended Pregnancy by Contraceptive Method Use and Age, Union Status, Race/Ethnicity and Poverty, National Survey of Family Growth, 1995**

Contraceptive Method Used	Total at Risk of Unintended Pregnancy	AGE							Currently Married	UNION STATUS			RACE/ETHNICITY			POVERTY STATUS		
		15-19	20-24	25-29	30-34	35-39	40-44	Cohab- iting		Formerly Married	Never Married	NonHispanic White	Black	Other	His-panic	0-149%	150-299%	300%+
<b>Total</b>	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<b>Sterilization</b>	36	0	5	20	38	56	66	46	27	50	8	36	38	29	37	40	37	32
Female	26	0	4	16	28	39	47	30	23	47	8	23	36	21	34	37	27	19
Male	10	0	1	4	10	18	19	16	4	3	0	13	2	8	4	3	10	14
<b>Reversible medical methods</b>	52	76	83	68	51	33	23	43	63	39	74	52	49	57	49	48	50	55
Oral contraceptives	25	35	48	37	27	11	6	19	34	19	38	27	21	18	21	22	25	26
Male condom	19	30	24	23	17	16	12	17	18	14	28	18	18	34	19	16	18	21
Depo Provera injectible	3	8	6	4	2	1	0	2	4	2	5	2	5	2	4	5	3	2
Barrier methods*	2	0	1	1	3	3	3	3	2	1	1	2	1	1	1	1	2	3
Norplant (implant)	1	2	3	2	1	0	0	1	2	1	2	1	2	2	2	3	1	1
Spermicides	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1
IUD	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
<b>Nonmedical methods</b>	5	5	4	5	6	6	5	6	5	4	4	5	3	9	5	4	5	6
Withdrawal	3	3	3	4	3	3	2	3	3	3	3	3	1	4	3	3	3	3
Periodic abstinence	2	1	1	1	3	3	2	3	2	1	1	2	1	5	2	1	2	3
<b>No method</b>	8	19	9	6	6	6	7	5	5	8	14	7	10	5	9	8	8	7

Table 3

SOURCE: Alan Guttmacher Institute tabulations of the 1995 National Survey of Family Growth (Cycle V).

NOTE: \*Female barrier methods include the diaphragm, cervical cap, sponge, and female condom.

150 percent of the poverty level use sterilization compared to 32 percent of women at 300 percent of the poverty level and above. Poor women relying on sterilization are much more likely than higher-income women to have been sterilized themselves rather than have a partner who has had a vasectomy. Female sterilization accounts for 92 percent of all contraceptive sterilization among poor women, compared with 58 percent among those with higher incomes.

**CONTRACEPTIVE EFFECTIVENESS**

Pregnancies occur to couples using contraceptive methods for two reasons—because of the inadequacy of the method itself or because it was not used correctly or consistently. Estimates have been made (either theoretically or empirically during clinical trials) regarding the efficacy of each contraceptive method given perfect use (Trussell 1998). In addition, estimates are made that measure the typical use effectiveness of each method, which relates to the experience of an actual group of users. The most recent estimates of typical use contraceptive effectiveness by method have been

made using the 1995 National Survey of Family Growth, corrected for abortion underreporting and standardized for variation in the proportions of women from different subgroups using certain methods (Fu et al. 1999). Failure rates differ by method, with some methods consistently showing higher effectiveness than other methods. Rates also differ by sociodemographic subgroup within study populations.

Table 4 provides estimates of method-specific failure rates given both perfect use and typical use for the most commonly used reversible contraceptive methods. For each contraceptive method, the typical failure rates observed among women are substantially higher than the estimated rates given perfect use, and the rates differ widely among marital status, age, and poverty of women subgroups. The lowest failure rates are achieved with long-acting hormonal contraceptives that require little user compliance. Among methods that women must use daily or per coital episode, oral contraceptives are most effective, while spermicides, withdrawal, and periodic abstinence have the highest failure rates. In general, women who are young,

**Estimated Percentage of Women Who Would Experience a Contraceptive Failure During the First Twelve Months of Perfect Method Use\* and the Corrected Use-failure Rates Given Typical Method Use\*\* for all Users and for Age, Poverty, and Marital Status Subgroups of Women Experiencing the Lowest or Highest Use-failure Rates**

METHOD	Perfect Use*	Typical Use**
<b>Total</b>		<b>13.1</b>
Norplant implants	0.05	2.0
Depo Provera injectible	0.3	3.5
Oral contraceptives	.1 to .5	8.5
Diaphragm/cervical cap	6/9 to 26	13.2
Male condom	3	14.9
Spermicides	6	28.2
Withdrawal	4	26.0
Periodic abstinence	1 to 9	21.8

*Table 4*

SOURCES: \*Trussell (1998), Table 31-1, p. 800. The ranges presented correspond to different formulations of the method, except for the cervical cap, where the range is due to woman's parity status. \*\*Fu et al., (1999), (From corrected Table 1, available on <http://www.agi-usa.org/pubs/journals/3105699.html>).

unmarried or cohabiting, and poor have higher failure rates. The differences in failure rates between methods and between subgroups of women are much greater than what any difference in method effectiveness or in the biology of women would cause and are assumed primarily to reflect differences in the correctness and consistency of method use (although reporting errors may also play a role).

#### **FAMILY PLANNING INFORMATION AND EDUCATION**

Rising public concern over the occurrence of unintended pregnancy and, particularly, of unintended, nonmarital adolescent pregnancy and childbearing in the United States has drawn attention to the manner in which young people are educated about sexuality, contraception, and how to avoid pregnancy and other negative consequences of

sexual activity. Parents and other adults have long played a key role in controlling the sexual behavior of adolescents and in providing basic information about sex and pregnancy avoidance. During the past twenty-five years, there has been a proliferation of organized efforts to augment the information, education, and support traditionally provided by families. Beginning with programs and services for young pregnant women, these efforts have expanded to include legislative mandates regarding the teaching of sexuality or family life education in schools, development and distribution of a variety of sexuality-education curricula, as well as integrated community interventions and media involvement. Organized efforts to implement sexuality education and related activities have also been influenced by growing public concern and awareness of HIV/AIDs and the need to provide young people with the information and means to avoid infection.

Increasingly, policies and programs to encourage abstinence among unmarried teenagers have become popular. Some of these programs attempt to accomplish this objective by giving young people encouragement, offering moral support and teaching interpersonal skills to resist pressures to become sexually active. Others, which seek to convince teenagers that sex before marriage is immoral, emphasize the negative consequences of sexual intercourse, occasionally withhold or distort information about the availability and effectiveness of contraception (Alan Guttmacher Institute 1994a). In fact, although most public schools provide some sort of sexuality education to middle or junior and senior high school students, the education provided is often too little, too late.

On a broader scale, community and service organizations have implemented interventions aimed at increasing the life options of disadvantaged young people through, for example, role models and mentoring, community service projects, job training, and activities aimed at reducing risky behaviors. Such interventions are expected indirectly to reduce levels of unintended teenage pregnancy, childbearing, and sexually transmitted infections, based on the belief that teenagers who are more positive about their futures are less likely

to participate in risk-taking behaviors, including risky sexual practices.

Other policies or programs implemented with the hope of reducing unprotected teenage sexual behavior include (1) comprehensive school-based sexuality-education curricula that include discussion of abstinence but also include information about contraceptive methods and services; (2) programs that address the social pressures faced by teenagers to have sex and that provide modeling and practice of communication, negotiation, and refusal skills; (3) condom availability programs in schools; and (4) multicomponent programs that include communitywide activities—such as media involvement, social marketing, and links between school-based activities and contraceptive service providers (Alan Guttmacher Institute 1994a).

Evaluations of a variety of programs and approaches aimed at affecting teenage sexual and reproductive behavior, although still somewhat inconclusive, have shown that some programs have had a positive effect on the behavior of youth. In addition, results of multiple studies indicate that the provision of contraceptive information and access does not encourage young people to become sexually active at younger ages. Reviews of the evaluation research point to the need for integrated approaches that both address the antecedents of sexual risk taking (e.g., poverty, violence, social disorganization) and provide young people (who will soon become adults) with the information, skills, and resources to make responsible decisions about sexual behavior and the avoidance of unintended outcomes (e.g., Kirby 1997).

### CONTRACEPTIVE SERVICE PROVISION

In the United States, women can receive contraceptive services from private practice general and family practitioners and obstetrician-gynecologists, as well as from publicly supported clinics run by hospitals, health departments, community health centers, and Planned Parenthood affiliates or independent clinic providers. In addition, some teenage and young adult women receive contraceptive services from school-based clinics and college or university health centers.

Private practice physicians are the most numerous providers in the United States that are available to women seeking contraceptive information and services. More than 40,000 family practice doctors and nearly 30,000 obstetrician-gynecologists provide office-based outpatient services (Alan Guttmacher Institute 1997). About seven in ten women seeking family planning services report going to a private practitioner or health maintenance organization (HMO) for their care (Aloma et al. 1997).

Annually, some 6.5 million U.S. women receive contraceptive services, supplies, and information from more than 7,000 publicly supported family planning clinics, located in 85 percent of all U.S. counties (Frost 1996). Family planning clinics, using a combination of federal, state, and local funds, provide care to those who cannot afford services from private physicians or who cannot use private physicians for other reasons. In most clinics, fees are based on the client's ability to pay, confidential services to teenagers are assured, and a full range of contraceptive methods are offered. As a result, family planning clinic clients are primarily low-income (57 percent are below 100 percent of the federal poverty level, and 33 percent are between 100 percent and 249 percent of the federal poverty level) and young (20 percent are under age 20; 50 percent are aged 20–29). Although a majority of clinic clients are non-Hispanic whites, nearly 40 percent are minority women (19 percent are black, 14 percent are Hispanic, and 7 percent are Asian or other races) (Frost and Bolzan). Lower-income women go to clinics primarily because they cannot afford physicians' fees, because the clinic is more conveniently located, or because the clinic will accept Medicaid payment. Adolescents often go to clinics because of the free or low-cost services and because they are afraid a private physician will tell their parents about their contraceptive use. In addition, some women, especially teenagers who have never been to a physician on their own, go to clinics because they do not know a physician who would serve them. Clinic clients usually shift to private physicians when their incomes rise and as they become older.

Sixty percent of all publicly supported clinics receive federal Title X support and must therefore

follow federal standard of care guidelines. These guidelines provide medical protocols as well as mandates regarding confidentiality and key areas of outreach that clinics should seek to address. As a result, many publicly supported clinics provide outreach and information or education in local schools or in other community locations. These clinics often seek to reach out to women (and men) in need of contraceptive care who have special needs or risk factors for unintended pregnancy (e.g., because of homelessness, drug or alcohol abuse, domestic violence, or other reasons).

The provision of contraceptive services, like all areas of health care, has been affected by changes in the structure of health care financing and the rise of managed care. In the past, most privately insured women had employer-based indemnity health insurance plans that rarely covered either routine gynecological checkups or reversible contraceptive services and supplies. However, such plans often covered sterilization services. Today, most privately insured women are enrolled in managed care plans. These plans are more likely to cover preventative care, including routine gynecological checkups and some reversible contraceptive services and supplies. However, not all managed care plans cover all or even most methods, and often the process of obtaining contraceptive services within managed care plans places additional burdens on women seeking contraceptive care. These burdens include prior authorization requirements that may cause some women to delay care or forgo sensitive care that a woman may not want to disclose to her primary care physician (Alan Guttmacher Institute 1994b, 1996).

### INTERNATIONAL COMPARISONS

Women in the United States have both similarities and differences with women throughout the world in their efforts to plan the number and timing of children. In adolescence, American women are somewhat less successful in their attempts to prevent unplanned pregnancies than are young women in most other industrialized countries. Table 5 presents recent birthrates among women aged 15–19 for selected European and North American

countries. Although there is no evidence that young women in the United States are more (or less) sexually active than young women in many other industrialized countries, the United States has adolescent birthrates that are nine to ten times higher than the rates for the Netherlands and Sweden and more than twice as high as the rates for Canada, England, and Wales. The factors responsible for these differences are not entirely clear; however, it is likely that they are due in part to differences in the levels of disadvantage among countries, to variation in the family planning education and services provided to youth and to greater or lesser openness regarding sexuality among countries.

In developing countries, young women marry earlier and have children at younger ages than in the United States. More than half of the young women in sub-Saharan Africa bear a child during the teenage years and about one-third of the young women in Asia, North Africa, the Middle East, and Latin America bear children as teens, compared to just one-fifth of teenage women in the United States. Although the situation is improving and the proportion of adolescents using contraception is increasing in many developing countries, few married adolescents use contraceptive methods and high percentages of sexually active unmarried adolescents rely on traditional nonmedical methods, such as withdrawal and periodic abstinence. Greater use of reliable contraceptive methods by adolescents worldwide is often hampered by inadequate or inaccurate information, poor access to services, and community expectations that value early childbearing within marriage and punish sexual activity outside of marriage. As a result, a substantial proportion of births to adolescents are unplanned—40 to 60 percent in several Latin American and sub-Saharan African countries, and 20 percent or more even in countries where almost all births are to married adolescents. By comparison, 66 percent of adolescent births in the United States are unplanned (Alan Guttmacher Institute 1998).

The patterns of pregnancy, childbearing, and contraceptive use among women of all ages vary from region to region throughout the world. Table 6 presents information on pregnancy rates

**Adolescent Birthrates for Selected European and North American Countries, Mid-1990s**

COUNTRY	BIRTHS PER 1,000 WOMEN AGED 15-19
Switzerland	5.7
Netherlands	5.8
Italy	6.9
Sweden	7.7
Spain	7.8
Denmark	8.3
Belgium	9.1
Slovenia	9.3
Finland	9.8
France	10.0
Greece	13.0
Germany	13.2
Norway	13.5
Austria	15.6
Czech Republic	20.1
Portugal	20.9
Poland	21.1
Iceland	22.1
Canada	24.2
England and Wales	28.4
Hungary	29.5
Slovak Republic	32.3
Estonia	33.4
Belarus	39.0
Bulgaria	49.6
Russian Federation	45.6
United States	54.4

Table 5

SOURCE: Singh et al. (2000).

(per 1,000 women aged 15-44), planning status of pregnancy, abortion, and contraceptive use for women in different regions of the world. Women in Africa experience the highest rates of pregnancy, the lowest percentage of pregnancies aborted, and the highest percentage of married women

using no contraceptive method. Although women in the United States have somewhat higher pregnancy rates than women in most of Europe and other North American countries, they have significantly lower pregnancy rates than most of the women in the rest of the world. The percentage of pregnancies that are planned varies considerably among the different regions of the world, from only 25 percent of pregnancies in eastern Europe being planned to 54 percent of African pregnancies being planned. Similarly, there is variation in the percentage of pregnancies that are aborted and in the patterns of contraceptive use. Compared to other regions of the world, the percentage of U.S. pregnancies that are aborted is similar to that in most of Europe and Latin America but less than half the percentage in eastern Europe. In terms of contraceptive use, married women in the United States are more likely than women in most other regions of the world (except for East Asia) to choose sterilization as their method of contraception. Married women in Africa and in Europe (both eastern and western Europe) rarely choose sterilization, and high percentages of Europeans rely on nonmedical methods, such as withdrawal.

Worldwide, only 47 percent of all pregnancies (including miscarriages) are planned and, at the same time, 42 percent of married women are using no method of contraception. In most regions of the world, one-third or more of married women use no method. Ensuring greater information, education, and access to family planning services worldwide has the potential to greatly reduce the level of unplanned pregnancy and abortion.

(SEE ALSO: *Birth and Death Rates; Family Size; Fertility Determinants; Pregnancy and Pregnancy Termination*)

REFERENCES

Abma, J. C., et al. 1997 "Fertility, Family Planning, and Women's Health: New Data from the 1995 Survey of Family Growth." *Vital and Health Statistics*, Series 23, No. 19, Table 1.

Alan Guttmacher Institute 1994 *Sex and America's Teenagers*. New York: Alan Guttmacher Institute.

— 1994b *Uneven and Unequal: Insurance Coverage and Reproductive Health Services*. New York: Alan Guttmacher Institute.

FAMILY PLANNING

**Total Number of Pregnancies per 1,000 Women Aged 15–44, Percentage of Pregnancies Planned and Aborted, and Contraceptive Use Among Married Women, Major Regions of the World, 1990s**

	TOTAL PREGNANCY RATE*	% OF PREGNANCIES PLANNED**	% OF PREGNANCIES ABORTED**	Contraceptive Use by Married Women Aged 15–49				
				Sterilization	Reversible Methods	Nonmedical Methods	No Method	Total
World	160	47%	22%	23	27	8	42	100
Africa	262	54%	12%	2	14	4	80	100
East Asia	123	47%	30%	44	38	1	17	100
Rest of Asia	182	51%	17%	17	17	8	58	100
Latin America	159	33%	23%	30	28	8	34	100
Eastern Europe	157	25%	57%	2	29	38	31	100
Rest of Europe	81	52%	21%	8	48	18	26	100
North America	100	41%	23%	27	34	6	33	100
United States	107	43%	23%	37	35	5	24	100

Table 6

SOURCES: columns 1–4: Alan Guttmacher Institute (1999); columns 5–9: United Nations (1999).

NOTE: \*Per 1,000 women aged 15–44.

NOTE: \*\*Pregnancies include miscarriages in this table.

— 1995 *Hopes and Realities: Closing the Gap Between Women's Aspirations and Their Reproductive Experiences*. New York: Alan Guttmacher Institute.

— 1996 *Improving the Fit: Reproductive Health Services in Managed Care Settings*. New York: Alan Guttmacher Institute.

— 1997 *Contraceptive Needs and Services*. New York: Alan Guttmacher Institute.

— 1998 *Into a New World: Young Women's Sexual and Reproductive Lives*. New York: Alan Guttmacher Institute.

— 1999 *Sharing Responsibility: Women, Society and Abortion Worldwide*. New York: Alan Guttmacher Institute.

Forrest, J. D. "Epidemiology of Unintended Pregnancy and Contraceptive Use." *American Journal of Obstetrics and Gynecology* 170(5):1485–1489.

Frost, J. 1996 "Family Planning Clinic Services in the United States, 1994." *Family Planning Perspectives* 28(3):92–100.

—, and M. Bolzan 1997 "The Provision of Public Sector Services by Family Planning Agencies in 1995." *Family Planning Perspectives* 29(1):6–14.

Fu, H., J. E. Darroch, T. Haas, and N. Ranjit 1999 "Contraceptive Failure Rates: New Estimates from

the 1995 National Survey of Family Growth." *Family Planning Perspectives* 13(2):56–63.

Harlap, S., K. Kost, and J. D. Forrest 1991 *Preventing Pregnancy, Protecting Health: A New Look at Birth Control Choices in the United States*. New York: Alan Guttmacher Institute.

Henshaw, S. K. 1998 "Unintended Pregnancy in the United States." *Family Planning Perspectives* 30(1):24–29, 46.

Kirby, D. 1997 "No Easy Answers: Research Findings on Programs to Reduce Teen Pregnancy." (Research review commissioned by The National Campaign to Prevent Teen Pregnancy Task Force on Effective Programs and Research, Washington, D.C.)

Singh, S., and J. E. Darroch 1999 "Trends in Sexual Activity Among Adolescent American Women: 1982–1995." *Family Planning Perspectives* 31(5):212–219.

—, and J. E. Darroch, Forthcoming "Adolescent Pregnancy and Childbearing: Levels of Trends in Developed Countries." *Family Planning Perspectives*.

Sonenstein, F. L., et al. 1998 "Changes in Sexual Behavior and Condom Use Among Teenaged Men: 1988 to 1995." *American Journal of Public Health* 88(6):956–959.

Trussell, J. 1998 "Contraceptive Efficacy." In R. A. Hatcher et al., eds., *Contraceptive Technology*, 17th rev. ed. New York: Ardent Media.

United Nations 1999 *Levels and Trends of Contraceptive Use as Assessed in 1998*. New York: United Nations.

Ventura, S. J., et al. 1998 "Births and Deaths: Preliminary Data for 1997." *National Vital Statistics Reports* 47(4):4.

JENNIFER J. FROST  
JACQUELINE E. DARROCH

## FAMILY POLICY IN WESTERN SOCIETIES

The dramatic growth in the industrial capacities and economic power of the United States and the countries of Western Europe since World War II has demonstrably improved the per capita income and quality of life of the average citizen in these countries. Based on such indicators as available health services, declining death rates, unemployment protection, and retirement benefits, it is reasonable to infer that, more than at any other time in human history, the majority of the people in these countries are assured that they can obtain the basic elements necessary for their survival and the survival of their families. Moreover, their quality of life, as measured by education, availability and quality of housing, ownership of automobiles and other durable goods, and available leisure time, suggests a lifestyle heretofore reserved only for the rich.

It is ironic that this unparalleled affluence has been accompanied by fundamental changes in family structure and family relationships—changes that challenge our basic conceptions of how families are organized and function. Marital stability and the size of the birthrate are common indicators of family health (Kittrie 1997). Over the early 1970s, virtually all the highly industrialized countries have experienced accelerating increases in divorce, decreased family size, increases in pregnancies among unmarried women, and an increase in childless marriages. Unlike many Third World countries, where population is growing at a rapid rate, birthrates in Western countries are below the level of population replacement (see Van de Kaa 1987 for Western Europe; Sweet and Bumpass 1987 for the United States). According to Census Bureau data, the American family reached its smallest size in 1990.

There is general agreement among family scholars that these changes are due, at least in part, to the expanding number of married women in the labor force. In 1920, about 9 percent of the families in the United States included wives who were employed outside the home (Hayghe 1990, p. 14). By 1987 the number of dual-earner families outnumbered families in which the husband was the only breadwinner by two to one (Lerner 1994). This situation does not change appreciably when children enter the family. Labor-force participation for married women with children under the age of six increased from 18.6 percent of all married women in 1960 to 63.6 percent in 1997. For married women with children between the ages of six and thirteen, the rate peaked in 1997 at 76.5 percent. Similar data exist for countries in Western Europe. Haavio-Mannila and Kauppinen (1990) note that full-time homemaking has almost disappeared in the Nordic countries. These changes have had a direct impact on the division of labor in the family and an indirect effect on marital-role relationships. Closely linked to women's growing labor-market activity is the fact that they are making better use of the educational opportunities available to them. There is evidence that educational achievement is converging for males and females in most of the Western world (Haavio-Mannila 1988).

As both educational and occupational career opportunities increase for women, there is a tendency to delay marriage. In the United States, the median age for first marriage peaked in 1996 for men (27.1 years) and in 1997 for women (25.0 years). Moreover, with growing educational equality and an increased ability to compete effectively for jobs in markets that value the skills obtained through education, the time such women spend raising children becomes more costly, both economically and psychologically (Sweet and Bumpass 1987). The result is fewer children per family and less time spent in parent-child interaction in those families with offspring (Lerner 1994).

These changes have altered marital and parental roles in families. Sweet and Bumpass (1987), describing the situation for U.S. families, claim that although husbands have made only a modest contribution to sharing household duties, they can no longer claim the role of "breadwinner." The conception of the prototypical family as consisting



of a legally married husband and wife living together for a lifetime, raising children in a household managed by the wife and mother and financially supported by the husband and father, is anachronistic. Indeed, such family structures constitute less than 7 percent of families in the United States (Lerner 1994).

Many scholars regard this decline in traditional family structures with considerable alarm. The decline is associated with values that emphasize individualistic and self-oriented goals over family well-being and, by extension, concern for others in the community. It is argued that these values lead to social isolation, growing levels of interpersonal distrust, and the the involvement by individuals in an endless and empty search for gratification (Elshtain 1997).

Other scholars, though not sanguine about the human problems created by these changes, tend to consider them as inevitable consequences of increased individual opportunity and freedom. Such freedom can only occur when developed societies establish equality between the sexes and underwrite the financial and social risks of illness, job loss, and aging. To the extent that greater individual autonomy and freedom of opportunity are gained at the cost of stable family relations, that cost should be borne (Myrdal 1967). Most important, those who look positively upon these changes point with approval to the changing role of women. The increased participation of women in the labor force in all industrialized countries has increased women's freedom to choose their own lifestyles, both in the world of work and in the family (Lerner 1994). If changes in family structure contribute to greater equality and greater opportunities for women to achieve their full potential, then, it is argued, society as a whole benefits (Lanca 1997).

Interestingly, people on both sides of the debate tend to define themselves as pro-family. Virtually everyone seems to acknowledge the social importance of the family in caring for and socializing children. There is also a perception of the family as a sanctuary that provides its members with protection, affection, and succor in what is often a hostile and, generally, impersonal environment. For the most part, advocates of both perspectives accept as problematic the fact that, increasingly, families fail to carry out these functions.

The high divorce rate, single-parent households, teenage pregnancy, and domestic violence are disturbing components in modern family life, and there is general agreement that something should be done about these problems. There is a consensus in all the advanced industrialized countries that the family is a proper concern of public policy. There are, however, fundamental differences in the conceptions of what that policy should be, what institutions or groups should be responsible for implementing policy, and what types of families or individuals should be the targets of such policy.

A conservative movement seeks to reverse the trends discussed above by shoring up the traditional family. They advocate reinforcing the importance of legal marriages, making divorce more difficult, emphasizing parental responsibility for the care and protection of children, and, more generally, strengthening the family by protecting it from government intrusions. Governments, from this perspective, should be restrained from taking over responsibilities that "rightfully" belong to the family. In this view, only limited indirect support in the form of tax relief or tax credits to help families care for their dependent members and to make homeownership easier are considered appropriate forms of government involvement (Carlson 1988). Advocates of restricting government involvement maintain that families, if left alone, would function quite well. Some critics of government involvement in family life argue for the privatization of most social services geared toward helping families. The essential claim is that privately organized and funded services to families would be more efficient and would protect the right of the family to make choices for itself (Brodkin and Young 1989).

Related to this position is the argument that a weakening of the family can be attributed to constitutional neglect—a failure to define and protect the status of the family as a publicly recognized institution (Kittrie 1997). Additionally, it is argued that contemporary social policy emphasizes individualistic interests rather than family concerns, thus reducing the family to a contractual arrangement rather than a socially recognized institution.

At the other extreme there are those who believe that the appropriate role of government is to assist the family in maximizing individual potential for each family member. They maintain that, in



democratic societies, government should mediate among various interest groups, each seeking its own advantages. Only the government, responsible to the electorate, is in a position to protect the public interest. Far from weakening the family, the government, it is argued, is the only institution in a position to use its resources to support and strengthen families and their members. This is especially true for families with children living under current conditions of rapid social change. Thus, it is argued that “social and demographic changes have combined to diminish the likelihood that families can assure their children’s healthy growth and development without help from outside the family” (Schorr et al. 1986).

Some have maintained that the question of government intervention is a moot point, since governments (as political entities) serve the purposes of their most influential constituencies. For example, they point to government tax incentives to assist businesses, the building of highways in the 1950s to benefit the automobile industry, and the government’s readiness to come to the aid of large corporations, banks, and savings and loan associations during times of crisis. In brief, for many observers, the notion of “minimalist government”—the ideal advocated in the rhetoric of the Reagan administration in the 1980s—is not a reality in any modern country (Kahn and Kamerman 1975). The real question is not whether government should come to the aid of its citizens, but which citizens will be the beneficiaries of such aid.

William J. Wilson’s (1987) answer to the question is that government services should be as universal as possible. Rather than limiting services to crisis intervention and emergencies or providing aid just to the disadvantaged, public services should be expanded to cover all citizens. Wilson maintains that only when public programs serve all segments of a society will it be possible to prevent system breakdowns and integrate all elements as productive members of the society.

The disagreements about what is appropriate family policy are so basic that it is hard to identify “objective” investigations or investigators. Policy implies advocacy, and advocacy implies bias. As a consequence, debates circle around such fundamentals as what is meant by the phrase *family policy*; indeed, there is disagreement on what is meant by *family*. Consider the following three

definitions: (1) “The family is a society limited in members but nonetheless a true society, anterior to every state or nation, with rights and duties of its own, wholly independent of the commonwealth” (Pope Leo XIII, quoted in Strong and DeVault 1989, p. 6). (2) “The family consists of two or more persons living together and related by blood, marriage or adoption” (U.S. Bureau of the Census 1990). (3) “One or more adults related by blood, marriage, or affiliation who cooperate economically, share a common dwelling place, and may rear children” (Strong and DeVault 1996, p. 15). Note that definition (1) places the family as independent and free from the state. By implication, what goes on in the family is none of the government’s business. Definition (2) restricts the family to shared residence and formal kin ties determined by blood or legal marriage, whereas definition (3), by introducing the vague term *affiliation*, is intended to include under the family label such “family” forms as cohabiting families and gay and lesbian families.

Still other definitions point to the fact that families need not always reside in the same household to maintain familial bonds and obligations, suggesting that emotional ties may be more salient than biological ties when defining family. A recent public opinion poll found that only 22 percent of respondents defined the American family in terms of blood, marriage, or adoption (Footlick 1990).

Those who define the family as the Census Bureau does are more likely to advocate policies that support only households whose members are linked through legally defined ties of kinship. Those who adhere to definition (3), on the other hand, are more likely to advocate government recognition of and assistance to different kinds of heterosexual and same-sex unions formed only by the consent of the participants without any state recognition or sanction. Adherents of definitions (1) or (2) are more likely to maintain that governmental aid to some of the groups that fit definition (3) is not aid to families but a major contribution toward the declining significance of the family in modern life (Kittrie 1997).

Despite disagreements concerning an adequate definition, some shared understanding of what is meant by “family” is necessary to provide a referent for the discussion that follows. The working

definition provided below is designed to incorporate essential elements that make up most definitions. It is, however, not likely to satisfy staunch advocates of any particular policy orientation. The family is defined here as *any group of individuals who are bound together by publicly acknowledged and socially sanctioned kinship ties*. The key to this definition is in the concept of *kinship*. Kinship is a special type of social relationship determined by ancestry, marriage, and adoption. In every society certain kin have designated duties toward and responsibilities for other members of the kin group. These responsibilities are obligatory; the only way an individual can free him- or herself from such obligations is by ceasing to be kin, a difficult, if not impossible, task. These obligations tend to be lifelong. For key kin the obligations can pertain to virtually all physical and emotional needs of family members. For example, in modern industrial societies parents are expected not only to care for their children's physical needs but to provide them with love and affection and to prepare them to function effectively in the real world. Failure to do so can result in severe sanctions against the parent. Thus all Western industrialized countries have laws designed to require that parents provide financial support and care for their children, even if the husband and wife are separated. Each country also has laws designed to sanction parents for physically abusing or neglecting their minor children. Similar obligations accrue as a consequence of marriage. Formal marriages are legal entities sanctioned by the state (and often the church), and violations of the marital agreement are monitored and adjudicated by reference to secular laws, sacred laws, or both.

Whereas the care, nurturance, and socialization of children continue to be the responsibility of parents, the care of aging parents by adult children is becoming less common. Social Security and retirement pensions, national health programs, and public support for housing for the elderly have been introduced in all highly industrialized countries. As these programs increase, family obligations tend to decline.

The tendency toward declining family responsibility for its members extends beyond care for the elderly. Over the years, as modes of production have changed, the family has gradually relinquished responsibility for providing either employment or occupational training for its members.

The family is no longer directly responsible for formal and religious education, and, more recently, responsibility for child care has tended to be relegated to nonfamily institutions. As the family gives up these functions to the workplace, schools, and churches, its relative power and influence on the economic and political life of the society diminishes. The issue thus becomes whether public policy, established and administered by agents of the governmental or economic segments of society, should be used to assist this weakened institution. More specifically, the question arises whether public services in the form of health care, housing, unemployment insurance, support for education, maternity benefits, support for child care, and so forth help to strengthen the family in its remaining responsibility for the care, nurturance, and socialization of its members or further weaken it by making family members dependent on nonfamily sources for benefits and social support. This is the key question surrounding the debate over what is a proper family policy in the Western world.

The term *policy*, like *family*, is burdened with multiple meanings and definitions. It has been variously described as the means for focusing on fundamental problems of individuals in relation to societies (Lasswell 1968), a set of decisions designed to support an agreed-upon course of action (Zimmerman 1988), or as governmental goals (Dumon and Aldous 1979). Thus the different definitions have covered the gamut, from means to process to goals. Kahn's (1969) conception of policy incorporates these elements within a context that gives it specific meaning. He refers to policy as "the general guide to action, the cluster of overall decisions relevant to the achievement of the goal, the guiding principles, the standing plan" (p. 131). Thus, a definition that seems to have some consensual meaning would hold that policy is a commitment to action that utilizes a consciously designated strategy designed to attain specified goals.

Although some writers insist that the term *policy* is applicable only to governmental action, the more general view is that any actor, whether an individual or a large collectivity, can make and implement policy (Zimmerman 1988). What is important, and the source of considerable debate, is the fit between the policy goals and the actor's ability to implement those goals. For example, individual families of moderate means are not

likely, by themselves, to be able to implement a policy for attaining high-standard, low-cost day care for their children; conversely, neither national governments nor national corporations are likely to be able effectively to develop policies for resolving neighborhood conflicts between ethnic groups. The difficulty is not only in matching appropriate levels of scale (e.g., community problems are best solved by community organizations); it also lies in the actor's ability to bring the right resources to bear on a given problem. This requires access to relevant information, the ability to recruit appropriate personnel, and control over the appropriate resources. The debate about family policy generally focuses on three primary issues: goals, the strategies designed to attain those goals, and the appropriate agencies for implementing these strategies. The sharpest differences concerning these issues are apparent at the national level.

The differences among policy goals, strategies, and designated agents are most evident when we compare the approach to family policy in the United States with the industrialized nations of Western Europe. As noted earlier, the United States shares many of the same family problems with these countries. The solutions offered in the form of public policy, however, differ considerably. The United States, often referred to as the "reluctant welfare state," has tended to be less forthcoming than the European countries in providing funds or services for its families or individuals. Moreover, there has been a strong inclination to have agencies other than government administer and implement social programs. America's revolutionary history has left its citizens with a pervasive distrust of government; rather than viewing government as a source of aid and assistance to its people, they often consider it a threat capable of usurping individual freedom and autonomy (Schorr 1979). Where the family is considered, this fear of government is enhanced by the tradition in Anglo-Saxon jurisprudence of noninterference in family life (Glendon 1989).

Differences in approach to family policy in the United States and Western Europe are also affected by the different cultural, racial, and ethnic compositions of the various countries. For example, despite the recent immigration of foreign workers to European countries such as West Germany, Switzerland, and Sweden, no European country approximates the United States in its ethnic

and racial heterogeneity. Whereas minorities make up between 6 and 7 percent of the population of countries like West Germany and Sweden, the nonwhite and Hispanic minorities in the United States represent 20 percent of the population (*Statistical Abstract of Sweden 1990*; *Statistisches Jahrbuch für die Bundesrepublik Deutschland 1989*; U.S. Bureau of the Census 1990).

Also important are the differences between the countries in the extent to which governments are centralized and can make policy at the national level. With the exception of Switzerland, the Western European countries tend to be more centralized. This makes it possible to have a meaningful national debate about the pros and cons of a national family policy. Such a debate in the United States is less conclusive, principally because decision-making power on such an issue is distributed among federal, state, and local governments (Dumon and Aldous 1979).

These factors tend to contribute to one incontrovertible fact: The United States has less legislation directly concerned with assisting families and provides less financial support designated to aid families than any other advanced Western industrial society (Kahn and Kamerman 1983; Kittrie 1997).

Americans have been more reluctant than Europeans to pass legislation that might be interpreted as violating the sanctity and the privacy of the family. Legislators in most European countries seem willing to write laws designed to protect the child's safety, health, and psychological well-being without great concern that such legislation might interfere with family prerogatives (Glendon 1989). Although the reluctance of the U.S. courts to venture into the family domain is gradually changing, especially in the areas of child-support enforcement and prevention of domestic violence, they have moved much more slowly and reluctantly in this area than their counterparts in the Scandinavian countries, France, the Netherlands, and West Germany. It would be unheard of, for example, for courts or legislatures in the United States to follow the Swedish example and forbid parents to spank their children (Davis 1997).

Although divorce is a common condition in all the Western developed countries, the United States and England have been hesitant to interfere in the financial arrangements resulting from divorce on

behalf of the children involved. This is not the case in the Nordic countries or in continental Europe, where, according to Glendon (1989), there is "genuine judicial supervision of the spouses' financial arrangements for children: mechanisms to ensure that child support is fixed at realistic and fair levels; highly efficient collection systems; 'maintenance advance systems' in which the state not only collects unpaid child support, but partially absorbs the risk of non-payment of advancing support up to a fixed amount in cases of default" (pp. 236-237).

The available data indicate that, with the possible exception of Social Security benefits for the elderly, the United States is not as generous as most of the European countries with regard to funding programs designed to assist families (Kahn and Kamerman 1983). This is most apparent when we consider families with children. The United States has tended to limit its assistance to such programs as Aid to Families with Dependent Children, which was geared toward providing financial aid to single-parent families living below the poverty line. Some additional assistance is provided to poor families through food stamps (unique to the United States) and, in the short run, unemployment insurance. Financial assistance to all families with children generally takes the form of a standard income tax deduction of \$2,700 for each dependent child. For families whose income is below \$10,000 an earned income credit is provided.

Most of the countries of Western Europe, on the other hand, provide direct cash payments to families for each child. These payments are tax-free and serve as family income supplements. The per-child payments tend to increase with increased family size. Funds are also provided to assist families during the childbearing phase in the form of direct grants to cover income loss and pre- and postnatal medical care (Kamerman and Kahn 1981).

The Nordic states and Western continental Europe also provide direct housing assistance for both rental and homeownership payments. Housing allowances generally decrease as income rises, thus providing an income leveling mechanism in these countries (Herrstrom 1986). Unlike the U.S. policy of tax deductions for interest paid on mortgages, which benefit better-off families, the European housing allowances tend to favor lower-income families (Kahn and Kamerman 1983; Herrstrom 1986).

In addition to these direct income-transfer programs, many of the countries in Western Europe provide a variety of benefits designed to assist families with the child-rearing challenges associated with both parents' working outside the home. Sweden, for example, provides a guaranteed pregnancy leave as well as a nine-month parental leave at 90 percent of salary and an additional three months at a flat rate (Lerner 1994). Austria, Germany, Norway, and the Netherlands have policies of at least three months of paid leave at 100 percent salary. Denmark and France have policies of at least sixteen weeks at 90 percent of salary, and Finland has a policy of eleven months at 80 percent of salary (Lerner 1994). Efforts are underway in Finland to extend this type of parental leave to three years. In the Nordic countries parents during the first three years of a child's life can choose to obtain state support for one parent to remain home with the child or to use the funds for day-care services. By comparison, the Family Support Act, passed in 1988 in the United States, renewed the requirement that a proportion of mothers receiving public assistance find employment (Michel 1998). At the same time, efforts to pass universal child-care legislation failed. Until January 1993, the United States was the only industrialized nation without a national parental leave policy (Lerner 1994).

Regardless of the merits of the American and European approaches to family policy, one significant current fact differentiates family conditions in the United States from those in the developed countries of continental Europe: In 1993, 40.1 percent of children under age six living in the United States were poor (U.S. Bureau of the Census 1994). This high level of poverty in the midst of affluence has produced a growing sense of despair and hopelessness concerning the life chances of these children. Miller (1990) summarizes the available data as follows: "Many studies have documented the much poorer health status of poor and minority children, the greater rates of child abuse, the higher incidence of injuries and death attributable to violence and crime, high rates of illiteracy, teenage pregnancy, educational failure and school dropouts." He goes on to note that the communities in which these children live are more dangerous, the schools are less adequate, and the rates of drug abuse are higher than in the general population. A sizable proportion of these children, but by

no means all, are African American Hispanic, or Native American. The majority live in female-headed, single-parent families, a phenomenon increasingly referred to as the “feminization of poverty” in the United States.

Thus, the essential issue concerning family policy for Americans is inextricably linked with the questions of poverty and what to do about it. The fact that this situation does not exist in the advanced countries of Europe (with the possible exception of the United Kingdom) raises the inevitable question whether the family policies of these “welfare states” have anything to offer Americans in their search for solutions to the problem of family poverty.

It will come as no surprise that expert opinion on this issue is greatly divided. Differences exist even among those who usually share the same ideological perspective. Some who are concerned with providing greater equality in our society advocate programs such as those in Western Europe. Wilson (1987, p. 149) refers to such programs as “universal programs of reform,” that is, programs that provide services for all citizens rather than earmarking aid for just minorities or the poor. The advantages of such programs are that since all citizens benefit, there is no stigma attached to being a recipient and no resentment among nonrecipients that taxes are being spent to benefit others.

This position is not universally shared even among those committed to egalitarian solutions. Some feel strongly that massive resources must be put into play to assist only those families at the greatest risk of being trapped in an abyss of poverty and hopelessness (Schorr 1988).

Predictably, those more concerned with issues of freedom than equality have argued that the problem of poverty in the United States is not the result of inadequate public services but rather is a consequence of such services. According to Carlson (1988), the “matriarchal welfare state has produced family disruptive results” and the “poverty crisis, the ageism crisis, the teen pregnancy crisis, the overpopulation crisis, the juvenile delinquency crisis, the eugenics crisis, the child abuse crisis, the youth suicide crisis” have all contributed to expanding the power of the state at the expense of

the family (p. 273). Charles Murray (1984) compiled an impressive amount of data that purports to illustrate that public services designed to reduce poverty in the United States were ineffectual. Murray’s data have been challenged by a number of investigators, and the ensuing debate reaffirms the difficulty in interpreting data relevant to this issue in ways that are free of bias. The difficulty is that socioeconomic conditions change so rapidly that it is not likely that any single factor can or should be justified as constituting the cause for poverty.

Whatever the causes, poverty is a reality in the United States, and, increasingly, the discrepancies in the distribution of wealth in this country demand the attention of policy makers (Phillips 1990). It also seems apparent that, given the ubiquity of the family in social life, any policy regarding poverty must also be a policy pertaining to families. Whether that policy will be concerned with all families or just the poor remains to be seen; whether it will be a national policy or policies instigated and implemented at the state levels, and whether its funding will be in the form of direct grants or tax benefits and incentives, is also hard to predict. It is likely, however, given American traditions and the deep differences that exist within the country about family policy, that the eventual solutions and compromises reached will be different from those that have evolved in Europe. Nevertheless, in this world of instant communication and visibility, the state of the family in Europe and the resources available to European families must have an impact on the policy decisions made in the United States.

(SEE ALSO: *Alternative Life styles; American Families; Family Law; Family Planning; Family Size; Fertility Determinants; Marriage and Divorce Rates*)

#### REFERENCES

- Brodtkin, Evelyn Z., and Dennis Young 1989 “Making Sense of Privatization: What Can We Learn from Economic and Political Analysis?” In Sheila B. Kamerman and Alfred J. Kahn, eds., *Privatization and the Welfare State*. Princeton, N.J.: Princeton University Press.
- Carlson, Allan C. 1988 *Family Questions: Reflections on the American Social Crisis*. New Brunswick, N.J.: Transaction Books.

- Davis, Phillip W. 1997 "The Changing Meanings of Spanking." In Arlene S. Skolnick and Jerome S. Skolnick, eds., *Family in Transition*, 9th ed. New York: Longman.
- Dumon, Wilfried, and Joan Aldous 1979 "European and United States Political Contexts for Family Policy Research." In Gerald W. McDonald and F. Ivan Nye, eds., *Family Policy*. Minneapolis: National Council on Family Relations.
- Elshtain, Jean Bethke 1997. "Reflections on the Family at Century's End." In Gordon L. Anderson, ed., *The Family in Global Transition*. St. Paul, Minn. PWPA/Paragon House.
- Footlick, Jerold 1990 "What Happened to the American Family?" *Newsweek* (December 1990): 18.
- Glendon, Mary Ann 1989 *The Transformation of Family Law: State, Law, and Family in the United States and Western Europe*. Chicago: University of Chicago Press.
- Haavio-Mannila, Elina 1988 "Converging Tendencies in Gender Roles." Paper presented at symposium, Can America Continue to Learn from Sweden and Finland?, University of Delaware, Newark, Del. October.
- Kaisa Kauppinen 1990 "Women's Lives and Women's Work in the Nordic Countries." In Hilda Kehore and Janet Giele, eds., *Women's Lives and Women's Work in Modernizing and Industrial Countries*. Boulder, Colo.: Westview Press.
- Hayghe, Howard V. 1990 "Family Members in the Work Force." *Monthly Labor Review* 113(3):14–19.
- Herrstrom, Staffan 1986 "Swedish Family Policy." *Current Sweden* 348 (September).
- Kahn, Alfred J. 1969 *Theory and Practice in Social Planning*. New York: Russell Sage Foundation.
- , and Sheila B. Kammerman 1983 *Income Transfers for Families with Children*. Philadelphia: Temple University Press.
- 1981 *Child Care, Family Benefits, and Working Parents: A Comparative Study*. New York: Columbia University Press.
- Kittrie, Nicholas N. 1997 "Policy, Law and Family Stability: From Antiquity to the Dawn of the Third Millennium." In Gordon L. Anderson, ed., *The Family in Global Transition*. St. Paul, Minn.: Paragon House.
- LANCA, Patricia 1997 "Feminism and the Family." In Gordon L. Anderson, ed., *The Family in Global Transition*. St. Paul, Minn.: Paragon House.
- Laswell, Harold 1968 "The Policy Orientation." In Daniel Lerner and Harold Lasswell, eds., *Policy Sciences*. Stanford, Calif.: Stanford University Press.
- Michel, Sonya 1998 "The Politics of Child Care in America's Public/Private Welfare State." In Karen V. Hansen and Anita Iltta Garey, eds., *Families in the U.S.: Kinship and Domestic Politics*. Philadelphia: Temple University Press.
- Miller, Gary J. 1990 "A Contextual Approach to Child Development" (Working Paper No. 9) Minneapolis: The City Inc.
- Myrdal, Alva 1967 "Forward." In Edmund Dahlstrom, ed., *The Changing Roles of Men and Women*. London: Gerald Duckworth.
- Murray, Charles 1984 *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books.
- Myrdal, Alva 1967 "Forward." In Edmund Dahlstrom, ed., *The Changing Roles of Men and Women*. London: Gerald Duckworth.
- Phillips, Kevin 1990 *The Politics of Rich and Poor: Wealth and the American Electorate in the Reagan Aftermath*. New York: Random House.
- Schorr, Alvin L. 1979 "Views of Family Policy." In Gerald McDonald and F. Ivan Nye, eds., *Family Policy*. Minneapolis: National Council on Family Relations.
- Schorr, Lisbeth B. 1988 *Within Our Reach*. New York: Anchor Books.
- , C. Arden Miller, and Amy Fine 1986 "The Social Policy Context for Families Today." In Michael W. Yogman and T. Berry Brazeiton, eds., *In Support of Families*. Cambridge, Mass.: Harvard University Press.
- Statistical Abstract of Sweden* 1990 Stockholm: Statistiska Centralbyran.
- Statistisches Jahrbuch für die Bundesrepublik Deutschland* 1989 Bonn: W. Kohlhammer.
- Sweet, James A., and L. L. Bumpass 1987 *American Families and Households*. New York: Russell Sage Foundation.
- U.S. Bureau of the Census 1990 *Statistical Abstract of the United States*. Washington, D.C.: U.S. Government Printing Office.
- Van de Kaa, Kirk J. 1987 "Europe's Second Demographic Transition." *Population Bulletin* 42:1–57.
- Wilson, William Julius 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- 1988 *Understanding Family Policy*. Newbury Park, Calif.: Sage Publications.

IRVING TALLMAN  
GINNA M. BABCOCK

## FAMILY ROLES

See Alternative Life Styles; American Families; Parental Roles.

## FAMILY SIZE

Family size may be considered from two perspectives. At the individual (micro) level, it defines one aspect of an individual's family background or environment. As such, it represents a potential influence on the development and accomplishments of family members. At the societal (macro) level, family size is an indicator of societal structure that may vary over time, with concomitant implications for individual development and social relations in different cohorts. In this essay, consideration is given to both aspects of family size, as it is reflected in sociological theory and research.

While the term *family size* is sometimes used to represent the total number of individuals comprising a family unit, Treas (1981) argues convincingly for decomposing the concept into two components: numbers of children and numbers of adults in the household. This distinction is important, as observed patterns of change in overall family size may be attributable to one component or the other, as may effects of overall family size. In the present discussion, family size is defined in terms of the number of children in the household.

A further distinction is made between family size in the parental and filial households, sometimes referred to as the family of origin (or orientation) and the family of procreation. Some use the term *sibship size* to refer to the number of children in an individual's parental family (Blake 1989; Ryder 1986). However, the two are not directly comparable: Mean family size takes into account those families which have no children, while mean sibship size is necessarily restricted to families with children.

Family size can also be differentiated from fertility, which reflects the aggregate numbers of births relative to the numbers of women in the population, without regard for the distribution of those births across family units. Fertility and family size are both important characteristics of cohorts; however, for assessing relationships at the

individual level, family size or sibship size is the more meaningful construct (Ryder 1986).

The subsequent sections address the following aspects of family size: demographic trends in family size, antecedents and correlates of family size, and implications of sibship size and family size for child and adult members of the family.

### DEMOGRAPHIC TRENDS

The twentieth century has witnessed substantial change in both fertility and family size (as indicated by the number of children in the household), with the overall trend being toward smaller families. Such trends can be examined through comparisons of fertility rates and mean family size, and also through investigation of parity distributions—that is, the numbers of families with zero, one, two (and so on) children.

Drawing on fertility tables compiled by the National Center for Health Statistics, Ryder (1986) presents time-series data for successive cohorts of women in the United States born between 1867 and 1955 (and who would thus be bearing children between approximately 1885 and 1975) that show the following general trends in fertility and family size:

1. Total fertility declined by 52 percent in the period being considered, from 4.00 for women born in 1867–1870 to 1.92 for women born in 1951–1955. A similar rate of decline occurred in marital fertility.
2. This decline was punctuated by a temporary upsurge in fertility for women born in 1916–1940, who were bearing children during the two decades following World War II (the “baby boom” years).
3. Variation in fertility rates increased for cohorts through 1910 and since then has consistently decreased, suggesting that in recent years there have been fewer women bearing no children or large numbers of children and an increasing concentration of families of moderate size.
4. Family size (the mean number of children in the family) decreased by 61 percent from a high of 7.3 for women born in 1867–1870 to 2.8 for women born in 1951–1955.

It thus appears that during the period under consideration, mean family size decreased at an even faster rate than fertility. Further, the increased fertility during the baby boom years appears to have been offset by reduced variation in fertility for those cohorts of women, with the result that mean family size held relatively constant during that period, then continued its pattern of decline.

Treas (1981) examined changes in family size between 1955 and 1978 for whites and for nonwhites, using data from the March Current Population Surveys. Throughout the period, nonwhites consistently had larger families than did whites: In 1955 the mean number of children was 1.26 in white families and 1.80 in nonwhite families; in 1978 the corresponding figures were 1.04 and 1.56. During this period Treas found similar patterns of increases in family size through the 1960s, followed by decreasing family size in the 1970s, for both groups. However, the shifts were considerably more pronounced among nonwhite families.

Data obtained from the U.S. Census on the distribution of family sizes (parity distributions) provide further insight on the trend toward smaller families. During the years between 1970 and 1988 the proportion of families with no children under eighteen increased substantially, from 44 percent to 51 percent, while the proportion of families with one child or two children increased only slightly (from 18 percent to 21 percent and from 17 percent to 18 percent, respectively). However, the proportion of families with three or more children decreased markedly, from 20 percent to 10 percent during this period. Among black and Hispanic families, the increase in families with no children was not as pronounced as among white families, but the increases in families with one or two children were greater, as were the decreases in families with three or more children (U.S. Bureau of the Census 1990, p. 51).

Further insight into the decline in family size is provided by investigations of parity progression, or the probability of having (or intending to have) an additional child at each parity level. Decomposing his time-series data into parity progressions, Ryder (1986) reports that the baby boom was the result of an increase in progression from parities one and two, but that progression from parities

three and higher have shown consistent declines. Similarly, data on intended parities show that the proportions intending progression from parity one have increased over time, while the intended progression ratios for parity three and higher have declined.

Other data on ideal, or normative, family sizes support this pattern of increasing concentration of smaller families. West and Morgan (1987) cite historical data showing that fertility norms have fluctuated in parallel with fertility rates and family sizes: During the 1930s and early 1940s two- and three-child families were preferred. During the post-World War II era three- and four-child families became the ideal, but in the late 1960s preferences reverted to the two- or three-child family. They further report that, among a sample of contemporary adults, a significant majority (64.8 percent) view the two-child family as ideal; that belief was surprisingly consistent across various subgroups defined by current family size, marital status, race, and religion.

At the same time that families have tended to become smaller on average, there has been increased variability in the timing of childbearing. One trend that has been widely noted has been the increase in childbearing among teenagers, particularly among those who are of lower socioeconomic statuses (SES), nonwhite, and less academically able youth (Card and Wise 1978). At the same time, there has been an increase in the proportion of women who delay childbearing until their early and mid-thirties or who remain childless (Bloom and Trussell 1984). As will be discussed below, the timing of the first birth has implications for the eventual family size and thus for the development and accomplishment of family members.

In sum, in the United States there appears to have been a strong shift toward smaller families, with the ideal being a two- or three-child family. A similar trend toward smaller families is found in other developed countries, while in developing countries families are more likely to be larger (Lopreato and Yu 1988). One exception to this generalization concerns countries, such as the People's Republic of China, that are trying to implement a policy of restricting families to one child. However, while the policy appears to have led to lower mean family sizes, numerous families have



continued to have two or more children, and a preferred family size of two continues to be the mode (Whyte and Gu 1987).

### ANTECEDENTS AND CORRELATES OF FAMILY SIZE

Determinants of family size have been investigated at both the societal and the individual level. At the societal level, researchers have sought to account for differences in fertility and family size over time or between societies. Easterlin (1980) advanced the theory that changes in fertility and family size over time are a function of individuals' economic resources and aspirations. He attributes the baby boom surge in fertility and family size to the generation of young men following World War II who experienced high wages, as a result of the expanding economy, and had relatively low material aspirations, as a result of being raised during the Depression. Conversely, the baby boom generation confronted increased competition for jobs, which, combined with higher aspirations, led to the "baby bust" of the 1970s and 1980s. One implication of Easterlin's theory is that smaller birth cohorts are likely to experience more favorable labor markets, resulting in higher fertility.

A variation of this theory is espoused by Devaney (1983), who argues that the decline in fertility observed during the 1960s and 1970s can be attributed to increases in female wages and female employment, which in turn served to depress fertility, rather than to conscious decisions to limit fertility in the face of disadvantageous economic conditions. Her analyses, based on national fertility data and data on female labor-force participation rates and male and female earnings, suggest (1) that female labor-force participation and fertility are highly and negatively correlated and (2) that female wage rates are the dominant factor in explaining recent variations in fertility and female employment. While this model differs from Easterlin's in terms of the process by which economic factors are thought to influence fertility, they are similar in viewing fertility as a response to economic market conditions.

Studies of developing countries have focused on several sociocultural as well as socioeconomic factors associated with fertility and family size:

modernization (Levy 1985); contraceptive use and family-planning programs (Koenig et al. 1987); and cultural attitudes and values, such as the perceived old-age security value of children (Rani 1986) or the view of children as risk insurance (Robinson 1986).

At the individual level, researchers have examined the extent to which fertility and family size may vary depending on individuals' family backgrounds, social and psychological characteristics, or economic status. Inverse relationships between social class and family size have been documented in a number of data sets: Individuals from larger families tend to have less-well-educated fathers who have lower-status occupations. Also, farm background is associated with larger family sizes (Blake 1989).

Parents' sibship size (the number of siblings that each parent had) is a second major determinant of family size: Women and men from larger families are more likely to have larger families (Ben-Porath 1975; Thornton 1980). This gives rise to an apparent paradox: While there is an overall trend toward small families, a high proportion of children come from larger families (Blake 1989). This paradox arises from the distinction noted above between cohort fertility rates, which are based on all women or all families, and children's sibship sizes, which are necessarily limited to women or families who have had children.

Retherford and Sewell (1988) investigated the relationship between intelligence and family size in their analysis of data from the Wisconsin Longitudinal Study of the High School Class of 1957, finding that the overall relationship between IQ and family size was negative for both sexes. However, the relationship proved to be much stronger for females, who showed consistent declines in family size as IQ increased. Among men the relationship was less consistent. Retherford and Sewell also reviewed the results of other, earlier studies, noting that the negative relationship between IQ and family size appears to have become more pronounced in the post-baby boom cohorts.

Additional factors associated with family size pertain primarily to family and achievement-related characteristics of the mother: More education, later age at marriage, longer interval between marriage and the birth of the first child, and

employment status are all associated with smaller families—that is, fewer children (Wagner et al. 1985). Family configuration has also been found to be associated with increased family size, with the probability of having an additional child being higher in families with all children of the same sex (Gualtieri and Hicks 1986). Also, only children are disproportionately likely to come from broken families (Blake 1989).

The interaction between wives' employment and childbearing has been a topic of much study, as women have increasingly entered or remained in the work force, but the results obtained are inconsistent. Waite and Stolzenberg (1976) found a significant negative relationship between wife's work and family size. However, based on analyses of longitudinal data that allowed for the study of recursive processes as well as inclusion of several additional measures, Bagozzi and Van Loo (1988) found no causal relationships between wife's employment and family size; they suggested that both labor-force participation and family size are codetermined by the wife's achievement motivation, sex-role norms, and perceived value of children.

Oropesa (1985) used data from the National Opinion Research Center (NORC) General Social Surveys to test the hypotheses represented in Easterlin's model at the micro level, using relative affluence as the predictor and expected family size as the outcome of interest. He found that relative affluence is more likely to be associated with expected births for women than for men, and that the effects are stronger with regard to expected births in the short term than with total expected family size.

The research cited above focuses on static determinants of childbearing and family size. However, some investigators have examined fertility and childbearing decisions as a dynamic process, influenced by life situation and life events, that may change over time, as well as by relatively fixed individual characteristics. One line of investigation has focused on timing of first birth as a determinant of eventual family size. Card and Wise (1978) and Hofferth and Moore (1979) demonstrated that early first births are associated with larger families; Bloom and Trussell (1984) similarly demonstrated that delayed childbearing is associated with smaller average family sizes, as well as with childlessness.

A second line of research has investigated the relationships between parity level and fertility decisions. Udry (1983) examined the relative influence of initial fertility plans and intervening life events (such as births during the interval, change in household income, change in education, female work status, change in marital satisfaction) on couple's fertility decisions at different parity levels. He found that including intervening events in the analyses improved the prediction of both fertility plans and, especially, actual fertility behavior, providing support for a sequential model of fertility decision making. White and Kim (1987) investigated whether the determinants of fertility choices vary by parity; they found a nonlinear relationship between fertility determinants and childbearing, especially with regard to factors related to women's roles. Both sex-role traditionalism and achievement in nonfamily roles were associated with a higher probability of having a child at parity zero or one, but a lower probability of having a child among women at higher parities. These findings are somewhat contrary to those based on cross-sectional analyses of family size, suggesting the importance of taking parity level into account in such investigations.

#### IMPLICATIONS OF SIBSHIP AND FAMILY SIZE

The effects of sibship/family size and family composition on children and on adults has long been a topic of popular interest and in recent years has become the focus of a considerable body of sociological and psychological inquiry. In particular, attention has been directed to effects of sibship size on children's cognitive development, physical and social-psychological development, educational attainment, and socioeconomic attainment and mobility. Consideration is also given to effects of family size on parents and on family well-being.

**Cognitive Development.** Interest in the relationship between sibship size and intelligence dates back to Anne Anastasi's (1956) review, which found an inverse relationship between the two. Subsequent empirical studies, in the United States as well as in Europe, using various measures of ability and controlling for family background characteristics, have confirmed this finding (Belmont and Marolla 1973; Breland 1974; Claudy et al. 1974). Blake (1989) provides a comprehensive review of

this literature, including a discussion of limitations and weaknesses in the prior studies.

Only children present a special case. Numerous studies have reported that only children do not perform as well on intelligence measures as do children from two-child families. Indeed, in the Belmont and Marolla study (1973), only children were found to be lower in intelligence than first-borns in families up to size four, and lower than second-borns in families up to size three. Claudy and associates (1974) obtained similar results after controlling for differences in SES. However, when differences in family composition were taken into account by restricting the sample to only children in two-parent families, the differences between only children and first-born children in larger families became nonsignificant (Claudy et al. 1979).

In an effort to account for the observed relationships between sibship size and intellectual ability, Zajonc (1976) introduced the "confluence model," which postulates that the intellectual environment in the home, defined by the combined intellectual levels of the parents and children, accounts for the observed relationships. According to his theory, the intellectual level is at its peak in families with two adults and no children; as the number of children in the home increases, the intellectual environment afforded to any individual child is effectively diluted. There are two implications of the "confluence model": Children from smaller families should show higher intelligence, and children born earlier in families should show higher intelligence. While the former hypothesis has been supported by a number of empirical studies, the latter did not account for the findings pertaining to only children. In response, Zajonc expanded the confluence model, postulating that younger siblings provide an opportunity for teaching, thus enriching the intellectual experience of older children; the lower intellectual performance of only children is attributed to the fact that they cannot avail themselves of this opportunity. While the confluence model has generated considerable discussion and debate, particularly regarding possible interactions between family size and birth order, and with family SES (for example, see Steelman 1985; Zajonc 1986), a systematic test of the model remains to be conducted.

Blake (1989) identifies two limitations in the previous work: lack of differentiation of various

kinds of intellectual ability (such as verbal and nonverbal) and potential interactions with SES. She finds that the inverse relationship between sibship size and intelligence holds for measures of verbal skill, but not for measures of nonverbal ability, and that the verbal ability deficits observed among children in large families are not limited to those from more disadvantaged backgrounds.

**Physical and Social-Psychological Development.** Compared with other outcome measures, relatively little attention has been given to the study of sibship-size effects on children's physical and social-psychological development. Mednick and associates (1985) and Wagner and associates (1985) provide brief reviews of this literature. Family size has been found to be inversely related to children's height and weight; it is also positively correlated with morbidity and mortality. With regard to social-psychological development, children from larger families have been found to have poorer self-concepts, to value conformity and self-control rather than independence and self-expression, and to show a greater tendency toward anti-social behavior. They are also less likely to be interested in white-collar occupations.

Blake (1989) investigated the relationship between sibship size and educational expectations, using data from three different cohorts of youth, and found that young people from smaller families, as well as from higher-status families, tend to have higher educational goals. These effects, however, are mediated through ability and grades as well as through parents' expectations.

**Educational Attainment.** Blake's (1989) book *Family Size and Achievement* provides the most comprehensive assessment to date of this area. Two sets of questions are addressed: First, does sibship size affect educational expectations and attainment, and if so, where in the educational process? Second, what is the relative importance of sibship size, relative to other measures of family background?

With regard to the first question, sibship size does appear to have a substantial effect on educational attainment. Individuals from small families had approximately two additional years of schooling, relative to their peers from larger families—net of differences attributable to parental characteristics. The greatest impact on education occurred at the high school level, with individuals

from larger families more likely to drop out of high school.

With regard to the second question, relative to other background variables in the analysis, sibship size was consistently second in importance for years of schooling, behind father's education. However, the negative effects of large families were somewhat mitigated by high parental SES and by membership in certain religious or ethnic groups. Similarly, the effects of parental SES were somewhat mitigated for youth in small families.

Some have argued that sibship size is simply a proxy for otherwise unmeasured characteristics of parents' family background and does not exert any independent effect on education in its own right. To address this concern, Blake (1989) examined the extent to which children from different-sized families have different home environments that might, in turn, influence educational attainment. In particular, attention was given to characteristics of the home setting (such as time spent reading newspapers, reading books, watching television) and to parental behaviors directed toward the child (such as encouragement, correction, goal setting). Children from smaller families were more likely to spend time in intellectual and cultural pursuits, to spend time playing alone, to have been read to as children, and to have had music or dance lessons. However, no significant differences were found in parental values for their children or in parenting style after parents' education and SES were taken into account. Thus, while there appear to be differences in the home environments afforded to children in smaller versus larger families, these differences do not appear to be attributable to differences in parental values or parenting style.

**Socioeconomic Attainment and Mobility.** A long tradition of research has addressed the question of how family background conditions or constrains individuals' socioeconomic attainment and social mobility. While primary consideration has been given to the impact of family social resources (father's education and occupation) on children's attainment, sibship size also was found to be related to occupational attainment (Blau and Duncan 1967). Among both women and men, those from larger families were more likely to have lower-status jobs and lower earnings, even after adjusting for differences in fathers' SES and educational attainment, both of which are correlated

with family size. Among women, the effect of sibship size on earnings was stronger than the effect of father's occupation (Featherman and Hauser 1976). Using path analysis to model both indirect and direct relationships, however, Duncan and associates (1972) found that the negative effect of sibship size on men's occupational status could be accounted for primarily by the effect of sibship size on educational attainment. This finding lends some support to arguments that larger families result in a dilution of family economic resources, thus constraining the opportunities available to children.

**Parents' Economic Well-Being.** Duncan and associates (1972) examined the impact of family size (as contrasted with sibship size) as a contingency in men's socioeconomic attainment, finding a slight and negative effect on occupational status but a positive effect on earnings, net of other background variables. Studies that included women found evidence of reciprocal relationships between family size and labor-force participation, which in turn affected women's career attainment (Waite and Stolzenberg 1976). However, as noted previously, Bagozzi and Van Loo (1988) suggested that women's work and family size are not causally related but are mutually dependent on other, achievement-related characteristics of the wife.

Relationships have been reported between the timing of childbearing and subsequent economic well-being. Card and Wise (1978) found that teenage parents of both sexes tended to have less education, lower job prestige, and lower earnings, relative to later childbearers, net of differences in background characteristics. Investigating this relationship in greater depth, Hofferth and Moore (1979) found that the effects of early childbearing on women's subsequent earnings were primarily attributable to the larger family sizes of these women and to the consequent implications for (less) work experience. However, they also found that early childbearing was less of a handicap for black women, due to weaker relationships between early childbearing and subsequent education and employment. Hofferth (1984) found that among women aged 60 or over, the number of children per se was not related to measures of economic well-being, but that the timing of childbearing was: Those who delayed the first birth until after age thirty had higher family incomes and higher standards of living than did women

whose first child was born before age thirty. This relationship was most pronounced among delayed childbearers who had small families, suggesting an interaction between timing of childbearing and family size.

Massagli (1987) has argued for a life-cycle model of the process of stratification that incorporates information on family size in both the parental and the filial generations. He hypothesizes that sibship size does not affect socioeconomic attainment directly but, rather, is related to the timing of early life-cycle transitions and to marital fertility; the observed negative effects of sibship size on attainment are attributed to the product of the relationship with life-cycle transitions and marital fertility and the negative effect of marital fertility on attainment.

**Parental Attitudes and Well-Being.** Wagner and associates (1985) review a number of studies of effects of family size on parental attitudes and parental health. They find that parental attitudes and treatment of children vary with family size: Larger families are more family centered, with a greater role played by fathers; at the same time, parents in larger families tend to be more authoritarian and more inclined to treat all children alike. Parents in larger families have also been found to have poorer marital relations. Finally, men and women who have many children are at greater risk of hypertension and other physical ailments.

In sum, sibship size and family size both appear to exert significant influence on the children and on the parents. Sibship size is closely related to family socioeconomic background, however, which is also a major influence on children's development and attainment. As a result, care must be taken to differentiate between effects of sibship size per se and effects of socioeconomic background. Similarly, family size among adults (the number of children they have) is highly correlated with socioeconomic status, intelligence, and other characteristics; again, it is important to consider the effects of family size net of these other factors. In many instances, the effects of sibship size and family size appear to be indirect. For example, sibship size is highly correlated with educational attainment and thus with subsequent occupational attainment. Similarly, among adults, family size is correlated with employment and thus with socioeconomic attainment. Finally, family size is often

closely related to other characteristics of the family: Among children, it may be related to birth order, and among parents, it may be related to the timing of childbearing. Understanding these indirect as well as direct relationships yields a better understanding of the ways in which, and the extent to which, sibship size and family size may affect the lives of children and adults.

### AT THE TURN OF THE CENTURY

The United States—as well as other developed and developing countries—has witnessed significant changes in fertility patterns and in family structure, which together combine to impact family size. This closing section reviews the more salient of these developments and examines how they have been reflected in recent sociological and demographic research.

**Family Size and Fertility.** Because family size is inextricably linked to fertility, it has been impacted by the fertility transition (i.e., the change from higher to lower rates of fertility) that has been well documented in the United States and is now being seen in both developed and developing societies elsewhere in the world. In the United States fertility has remained relatively constant since the early 1980s, ranging from 1.7 to 1.9 births per woman. However, this apparent stability masks a dramatic shift toward having children at later ages, especially among white women (Chen and Morgan 1991) and more highly educated women (Rindfuss et al. 1996).

The stability seen in the United States is in sharp contrast to Europe, where most countries have experienced significant declines in fertility during this period, and to many developing countries, which are also now evidencing fertility declines (Rutenberg and Diamond 1993; Thomas and Muvandi 1994). Global fertility projections for the twenty-first century (released by the United Nations in 1992) range from 1.7 to 2.5 births per woman (Cohen 1996). These declines have been linked to three factors:

1. Widespread changes in the *social and economic roles and opportunities* available to women, including the increased availability of child care (Hirschman and Guest 1990; Mason and Kuhlthau 1992; Rindfuss et al. 1996)

2. Increased availability of *contraception and family-planning services*, as well as in some developing countries policies supporting fertility limitation in some developing countries (Axinn 1992; DeGraff 1991; Lavelly and Freedman 1990; Njogu 1991)
3. Changes in the *social norms governing childbearing and child rearing*, including the emerging concept of “numeracy” about children—that is, the idea of having a particular family size as a goal—in developing countries (van de Walle 1992)

Corresponding decreases in actual and expected family size are also seen for this period. The average family size in 1993 was 3.16, down from 3.29 in 1980; similarly, the proportion of family households with three or more children had fallen by half since 1970 (Dortch 1993). On the 1994 General Social Survey 55 percent of Americans reported that they preferred two-child families—up from 41 percent in 1972—while the percentage of preferring substantially larger families declined commensurately. By 1988, the proportion of women expecting to remain childless had increased to 9 percent (National Center for Health Statistics 1996).

**Family Size and Family Structure.** Family size is also closely linked to family structure and to changes in patterns of family formation. Two somewhat related changes in particular have significantly impacted the size of family units: increased rates of marital dissolution and increased rates of out-of-wedlock births, both of which have contributed to a dramatic increase in single-parent family units.

Rates of *marital dissolution* have increased dramatically, both in the United States and elsewhere. In the United States more than half of all marriages are now expected to end in divorce; in less developed countries, approximately 25 percent of first marriages, on average, have dissolved as a result of death, divorce, or separation (Bruce et al. 1995). Not only does marital dissolution contribute directly to smaller family size (Lillard and Waite 1993); it also has an indirect effect—maternal divorce not followed by remarriage substantially reduces children’s preferred family size (Axinn and Thornton 1996).

Beginning in the 1980s, women were increasingly likely to have *children out of wedlock*, signaling

a significant change in the norms governing childbearing. By the early 1990s, 2 out of 3 black children and nearly 1 of 4 white children were born to unmarried mothers (Smith et al. 1996). One-fourth of these out-of-wedlock births were to cohabiting couples (Bumpass 1990). Thus, while fewer women were marrying and staying married, alternative family structures involving children were emerging.

The number of *single-parent families* in the United States grew dramatically from 1960 (10.5 percent) to 1990 (23.3 percent) (Garasky and Meyer, 1996). As a result, it is estimated that half of today’s young children will spend some time in single-parent family (Bumpass 1990). While the majority of single-parent family units are headed by the mother, the number of father-only families has grown at nearly twice the rate as the number of mother-only families. Nor is this phenomenon limited to the United States: In the former Soviet Union, the proportion of households headed by a single parent doubled in the fifteen-year period from 1980 to 1995 to 20 percent; in developing countries, the incidence of female-headed households as of 1995 ranged from 11 percent in the Philippines, to 13 percent in Mexico, to 19 percent in Cameroon, to more than 25 percent in Hong Kong (Bruce et al. 1995).

#### **Implications for the Study of Family Size.**

Returning to the framework initially presented in this essay, what are the implications of these trends and developments for the conceptualization of the “family” and “family size,” and for research on the correlates and implications of family size?

*Conceptualization of the “family” and “family size.”* The decreasing variance in family size is being offset by increasing complexity in family structure. In addition to the growing interest in single-parent families—and within that category, differentiation of mother-only and father-only families—researchers also identify nonmarital cohabitation (Bumpass 1990), parent-stepparent and blended families (Astone and McLanahan 1991; Dortch 1993; Wojtkiewicz 1993), and intergenerational households (Macunovich and Easterlin 1990). This evolving conceptualization of the family and—in particular—family structure is of interest not only in its own right but also for its implications for models of intergenerational transmission of status, resources, and values (Smith et al. 1996).

*Correlates of changes in family size and structure.* Increasingly research is directed toward linking social change at a macro level to individual-level fertility behavior. Structural factors, including increased labor-force participation of women (Rindfuss et al 1996), availability of contraceptive technology (Lavelly and Freedman 1990), and availability of child care (Mason and Kuhlthau 1992; Rindfuss et al. 1996), continue to be a subject of study in both developed and developing countries. Of equal interest is the social context surrounding child-bearing decisions, including the husband's and wife's own values regarding desired family size (Thomson 1997; Thomson et al. 1990), their parents' preferences and behavior (Axinn et al. 1994; Axinn and Thornton 1996), and societal norms (van de Walle 1992). As increasing attention is given to fertility transitions occurring in other countries, attention is also being given to identifying cultural factors that can potentially bias data and findings, such as nonresponse or qualitative responses to questions about expected or desired family size (Hermalin and Liu 1990; Riley et al. 1993).

*Implications of changes in family size and structure.* Considerable attention continues to be devoted to studying the impact of family size and structure on children's achievement. The inverse relationship between family size and children's attainment that has been widely documented in the United States is also observed in a number of developing countries, including Thailand (Knodel and Wonsith 1991), Vietnam (Anh et al. 1998), Ghana (Lloyd and Gage-Brandon 1995), and Israel (Shavit and Pierce 1991). Research is increasingly focusing on delineating the processes underlying these relationships, including the greater availability of parental economic and interpersonal resources in smaller families (Downey 1995; Macunovich and Easterlin 1990; Powell and Steelman 1993). Similarly, studies of the negative impact of marital disruption on children's achievement also explore how social factors such as reductions in parental expectations and involvement mediate this relationship (Astone and McLanahan 1991; Wojtkiewicz 1993).

These changes in family size and structure have significant implications for policy as well as for research. Domestically, Dortch (1993) raises the question of how the trend toward smaller

families will impact caring and support relationships for older family members, especially as the number of older American increases over the next few decades. In developing countries, where many governments are proactively working to foster economic development and social well-being, policies supporting lower fertility and smaller families may have both direct and indirect benefits: As the number of children coming from smaller families increases, so too should their prospects for educational and economic attainment (Knodel and Wonsith 1991).

(SEE ALSO: *Family Planning; Family Policy in Western Societies; Fertility Determinants*)

## REFERENCES

- Anastasi, A. 1956 "Intelligence and Family Size." *Psychological Bulletin* 53:187-209.
- Anh, Truong Si, John Knodel, David Lam, and Jed Friedman 1998 "Family Size and Children's Education in Vietnam." *Demography* 35(1):57-70.
- Astone, Nan M., and Sara S. McLanahan 1991 "Family Structure, Parental Practices, and High School Completion." *American Sociological Review*, 56 (June):309-320.
- Axinn, William G. 1992 "Family Organization and Fertility Limitation in Nepal." *Demography* 27(4):503-521.
- , Marin E. Clarkberg, and Arland Thornton 1994 "Family Influences on Family Size Preferences." *Demography* 31(1):65-79.
- , and Arland Thornton 1996 "The Influence of Parents' Marital Dissolutions on Children's Attitudes Toward Family Formation." *Demography* 33(1):66-81.
- Bagozzi, Richard P., and M. Frances Van Loo 1988 "An Investigation of the Relationship Between Work and Family Size Decisions over Time." *Multivariate Behavioral Research* 23:3-34.
- Belmont, L., and F. A. Marolla 1973 "Birth Order, Family Size, and Intelligence." *Science* 182:1096-1101.
- Ben-Porath, Y. 1975 "First-Generation Effects on Second-Generation Fertility." *Demography* 12:397-405.
- Blake, Judith 1986 "Number of Siblings, Family Background, and the Process of Educational Attainment." *Social Biology* 33:5-21.
- 1989 *Family Size and Achievement*. Berkeley: University of California Press.
- Blau, Peter M., and Otis D. Duncan 1967 *The American Occupational Structure*. New York: Free Press.

- Bloom, David E., and James Trussell 1984 "What Are the Determinants of Delayed Childbearing and Permanent Childlessness in the United States?" *Demography* 21:591-611.
- Breland, H. M. 1974 "Birth Order, Family Configuration, and Verbal Achievement." *Child Development* 45:1011-1019.
- Bruce, Judith, Cynthia B. Lloyd, and Ann Leonard 1995 *Families in Focus: New Perspectives on Mothers, Fathers, and Children*. New York: Population Council.
- Bumpass, Larry L. 1990 "What's Happening to the Family? Interactions Between Demographic and Institutional Change." *Demography* 27(4):483-498.
- Card, Josefina J., and Lauress L. Wise 1978 "Teenage Mothers and Teenage Fathers: The Impact of Early Childbearing on the Parents' Personal and Professional Lives." *Family Planning Perspectives* 10:199-205.
- Chen, Renbao, and S. Philip Morgan 1991 "Recent Trends in the Timing of First Births in the United States." *Demography* 28(4):513-533.
- Claudy, John G., William S. Farrell, Jr., and Charles W. Dayton 1979 "The Consequences of Being an Only Child: An Analysis of Project TALENT Data." Final report. Palo Alto, Calif.: American Institutes for Research.
- Claudy, John G., David E. Gross, and Rebecca D. Strause 1974 "Two Population Studies: I. Family Size, Birth Order, and Characteristics of Young Adults, and II. A Study of Married Couples in Knox County, Tennessee." Final report. Palo Alto, Calif.: American Institutes for Research.
- Cohen, Joel E. 1996 "How Many Can the Earth Support?" *American Demographics*.
- DeGraff, Deborah S. 1991 "Increasing Contraceptive Use in Bangladesh: The Role of Demand and Supply Factors." *Demography* 28(1):65-81.
- Devaney, Barbara 1983 "An Analysis of Variations in U.S. Fertility and Female Labor Force Participation Trends." *Demography* 20:147-161.
- Dortch, Shannon 1993 "The Future of Kinship." *American Demographics*.
- Downey, Douglas B. 1995 "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60 (October):746-761.
- Duncan, Otis D., David L. Featherman, and Beverly Duncan 1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Easterlin, Richard A. 1980 *Birth and Fortune*. New York: Basic Books.
- Featherman, David L., and Robert M. Hauser 1976 "Sexual Inequalities and Socioeconomic Achievement." *American Sociological Review* 41:462-483.
- Garasky, Steven, and Daniel R. Meyer 1996 "Reconsidering the Increase in Father-Only Families." *Demography* 33(3):385-393.
- Gualtieri, C. Thomas, and Robert E. Hicks 1986 "Family Configuration and Family Size." *Social Biology* 33:146-147.
- Hermalin, Albert I., and Xian Liu 1990 "Gauging the Validity of Responses to Questions on Family Size Preferences in China." *Population and Development Review* 16(2):337-354.
- Hirschman, Charles, and Philip Guest 1990 "Multilevel Models of Fertility Determination in Four Southeast Asian Countries: 1970 and 1980." *Demography* 27(3):369-396.
- Hofferth, Sandra L. 1984 "Long-Term Economic Consequences for Women of Delayed Childbearing and Reduced Family Size." *Demography* 21:141-155.
- , and Kristin A. Moore 1979 "Early Childbearing and Later Economic Well Being." *American Sociological Review* 44:784-815.
- Knodel, John, and Malinee Wonsith 1991 "Family Size and Children's Education in Thailand: Evidence from a National Sample." *Demography* 28(1):119-131.
- Koenig, Michael A., James F. Phillips, Ruth S. Simmons, and Mehrab A. Khan 1987 "Trends in Family Size Preferences and Contraceptive Use in Matlab, Bangladesh." *Studies in Family Planning* 18:117-127.
- Lavelly, William, and Ronald Freedman 1990 "The Origins of the Chinese Fertility Decline." *Demography* 27(3):357-367.
- Levy, Victor 1985 "Cropping Pattern, Mechanization, Child Labor, and Fertility Behavior in a Farming Economy: Rural Egypt." *Economic Development and Cultural Change* 33:777-791.
- Lillard, Lee A., and Linda J. Waite 1993 "A Joint Model of Marital Childbearing and Marital Disruption." *Demography* 30(4):653-681.
- Lloyd, Cynthia B., and Anastasia J. Gage-Brandon. 1995 *Does Sibsize Matter? The Implications of Family Size for Children's Education in Ghana* (Working Paper no. 45). New York: Population Council.
- Lopreato, Joseph, and Mai-yu Yu 1988 "Human Fertility and Fitness Optimization." *Ethology and Sociobiology* 9:269-289.
- Macunovich, Diane J., and Richard A. Easterlin 1990 "Life Cycle Demographic Decisions and the Economic Status of Young Children." *Population and Development Review* 16(2):301-325.



- Mason, Karen O., and Karen Kuhlthau 1992 "The Perceived Impact of Child Care Costs on Women's Labor Supply and Fertility." *Demography* 29(4): 523-543.
- Massagli, Michael P. 1987 "Effects of Family Size on the Process of Stratification: A Structural Equation Model for White Couples in the U.S. in 1962 and 1973." In Robert V. Robinson, ed. *Research in Social Stratification and Mobility*, vol. 6. Greenwich, Conn.: JAI Press.
- Mednick, Birgitte R., Robert L. Baker, and Dennis Hocevar 1985 "Family Size and Birth Order Correlates of Intellectual, Psychosocial, and Physical Growth." *Merrill-Palmer Quarterly* 31:67-84.
- National Center for Health Statistics 1996 "Birth Expectations of Women in the United States, 1973-88."
- Njogu, Wamucii 1991 "Contraceptive Use in Kenya: Trends and Determinants." *Demography* 28(1): 83-99.
- Oropesa, R. S. 1985 "Subject Relative Affluence and Expected Family Size." *Sociology and Social Research* 69:501-515.
- Powell, Brian, and Lala Carr Steelman 1993 "The Educational Benefits of Being Spaced Out: Sibship Density and Educational Progress." *American Sociological Review* 58 (June):367-381.
- Rani, Usha D. 1986 "Old Age Security Value of Children and Fertility in Relation to Social Policy." Paper presented at the annual meeting of the International Sociological Association.
- Retherford, Robert D., and William H. Sewell 1988 "Intelligence and Family Size Reconsidered." *Social Biology* 35:1-40.
- Riley, Ann P., Albert I. Hermalin, and Luis Rosero-Bixby 1993 "A New Look at the Determinants of Non-Numeric Response to Desired Family Size: The Case of Costa Rica." *Demography* 30(2):159-174.
- Rindfuss, Ronald R., S. Philip Morgan, and Kate Offutt 1996 "Education and the Changing Age Pattern of American Fertility: 1963-1989." *Demography* 33(3):277-290.
- Robinson, W. C. 1986 "High Fertility as Risk-Insurance." *Population Studies* 40:289-298.
- Rutenberg, Naomi, and Ian Diamond 1993 "Fertility in Botswana: The Recent Decline and Future Prospects." *Demography* 30(2):143-157.
- Ryder, Norman B. 1986 "Observations on the History of Cohort Fertility in the United States." *Population and Development Review* 12:617-643.
- Shavit, Yossi, and Jennifer L. Pierce 1991 "Sibship Size and Educational Attainment in Nuclear and Extended Families: Arabs and Jews in Israel." *American Sociological Review* 56 (June):321-330.
- Smith, Herbert L., S. Philip Morgan, and Tanya Koropecykj-Cox 1996 "A Decomposition of Trends in the Nonmarital Fertility Ratios of Blacks and Whites in the United States, 1960-1992." *Demography* 33(2):141-151.
- Steelman, Lala C. 1985 "A Tale of Two Variables: A Review of the Intellectual Consequences of Sibship Size and Birth Order." *Review of Educational Research* 55:353-386.
- Thomas, Duncan, and Ityai Muvandi 1994 "The Demographic Transition in Southern Africa: Another Look at the Evidence from Botswana and Zimbabwe." *Demography* 31 (2):185-207.
- Thomson, Elizabeth 1997 "Couple Childbearing Desires, Intentions, and Births." *Demography* 34(3):343-354.
- , Elaine McDonald, and Larry L. Bumpass 1990 "Fertility Desires and Fertility: Hers, His, and Theirs." *Demography* 27(4):579-600.
- Thornton, A. 1980 "The Influence of First Generation Fertility and Economic Status on Second Generation Fertility." *Population and Environment* 3:51-72.
- Treas, Judith 1981 "Postwar Trends in Family Size." *Demography* 18:321-334.
- Udry, J. Richard 1983 "Do Couples Make Fertility Plans One Birth at a Time?" *Demography* 20:117-128.
- U.S. Bureau of the Census 1990 *Statistical Abstract of the United States: 1990*. Washington, D.C.: U.S. Government Printing Office.
- van de Walle, Etienne 1992 "Fertility Transition, Conscious Choice, and Numeracy." *Demography* 29(4):487-502.
- Wagner, Mazie E., Herman J. P. Schubert, and Daniel S. P. Schubert 1985 "Family Size Effects: A Review." *Journal of Genetic Psychology* 146:65-78.
- Waite, Linda J., and Ross M. Stolzenberg 1976 "Intended Childbearing and Labor Force Participation of Young Women: Insights from Nonrecursive Models." *American Sociological Review* 41:235-252.
- West, Kirsten K., and Leslie A. Morgan 1987 "Public Perceptions of the Ideal Number of Children for Contemporary Families." *Population and Environment* 9:160-171.
- White, Lynn K., and Hyunju Kim 1987 "The Family-Building Process: Childbearing Choices by Parity." *Journal of Marriage and the Family* 49:271-279.
- Whyte, Martin K., and S. Z. Gu 1987 "Popular Response to China's Fertility Transition." *Population and Development Review* 13:471-494.
- Wojtkiewicz, Roger A. 1993 "Simplicity and Complexity in the Effects of Parental Structure on High School Graduation." *Demography* 30(4):701-717.

Zajonc, Robert B. 1976 "Family Configuration and Intelligence." *Science* 192:227-236.

—1986 "Family Factors and Intellectual Test Performance: A Reply to Steelman." *Review of Educational Research* 56:365-371.

LAURI STEEL

## FAMILY VIOLENCE

Physical violence of all types, from slaps to murder, probably occurs more frequently in the family than in any other setting or group except the armed services or police in time of war or riot. This article summarizes the prevalence rates and examines reasons for the high rates, with emphasis on the characteristics of the family as a social institution and on social inequality.

Physical violence is defined as an act carried out with the intention or perceived intention of causing physical pain or injury to another person (Gelles and Straus 1979). For certain purposes, the term "assault" is preferable because much intrafamily violence is a statutory crime. However, not all violence is criminal. Hitting a misbehaving child is legal and expected in all but a few countries. Corporal punishment of an "errant wife" was legal under common law in the United States until the 1870s (Calvert 1974).

Child abuse was not regarded as a widespread social problem by sociologists, family therapists, or the public until the 1960s (Nelson 1984; Pfohl 1977), and wife beating not until the women's movement made it a national issue in the mid-1970s. The subsequent emergence of public concern about and research on these and other aspects of family violence reflects major social changes, including the following:

1. The social activism of the 1960s, which sought to aid oppressed groups of all types, was extended to this aspect of the oppression of children and women.
2. The rising homicide and assault rates, violent political and social protest and assassinations, terrorist activity, and the Vietnam War sensitized people to violence.
3. Disenchantment with the family in the 1960s and early 1970s facilitated the

perception of negative features of family life, including violence.

4. The growth of paid employment by married women provided the economic means for them no longer to tolerate the abuse that had long been the lot of women.
5. The reemerged women's movement made battering a central issue in the mid-1970s and gave it wide publicity.
6. The creation by the women's movement of a new social institution—shelters for battered women—did more than provide material assistance. Shelters were ideologically important because they concretized and publicized a phenomenon that had previously been ignored.
7. Changes in theoretical perspectives in sociology put the consensus model of society under attack by conflict theory. The inevitability of conflict in all human groups, including the family, was recognized, along with the possibility of violent conflict.

### PREVALENCE OF FAMILY VIOLENCE

**Homicide.** In the United States, about one-quarter of all murders involve family members (Straus 1986). In other industrialized countries the percentage is much higher, for example, 40 percent in Canada and 67 percent in Denmark (Straus 1987). These high percentages occur because Canada has a low homicide rate and Denmark an even lower one: The few family homicides that occur are a large proportion of the low overall rate. This suggests that when homicide has been almost eliminated in a society, such as in Denmark, the family is the setting in which it is most likely to persist. Homicides of domestic partners have been decreasing in the United States since the mid 1970s (Greenfeld et al. 1998) and in Canada (Fedorowycz 1999).

### OFFICIAL STATISTICS ON CHILD ABUSE AND SPOUSE ABUSE

National statistics on child abuse in the United States have been published since 1976. These statistics vastly underestimate the actual extent of

child abuse. Many times more children are severely beaten each year but do not come to public attention. Officially reported cases grew from about 600,000 in 1976 to about 3 million annually in the mid-1990s. However, rather than being a 400 percent increase, the growth in cases reported to child protective services reflects social changes such as mandatory child abuse reporting laws, hot lines, child abuse education campaigns, an increasingly educated population, and a growth in professionals concerned with aiding and protecting children. These changes led the public and professionals to report cases that previously would have been ignored. This is consistent with historical and survey evidence suggesting that the true incidence of physical child abuse has been slowly decreasing since the late seventeenth century (Radbill 1987; Straus and Gelles 1986).

There are no official statistics for the United States on violence between spouses because the Uniform Crime Reporting System used by almost all police departments does not record the relationship between victim and offender. A new "incident-based reporting system" includes that information and is now being used in twelve states. However, because only about 7 percent of domestic assaults come to the attention of the police (Kaufman Kantor and Straus 1990), the new system will uncover only a small fraction of the cases. A similar problem makes the U.S. National Crime Survey (Gaquin 1977-1978; U.S. Department of Justice 1980) vastly underestimate the incidence of wife beating (Straus 1999). The public tends to consider assault by a spouse as a "family problem" rather than a "crime" and rarely informs the survey interviewer of such events.

**The National Family Violence Surveys.** National surveys of U.S. families were conducted in 1975 (Straus et al. 1980) and 1985 (Gelles and Straus 1988), and national surveys focused on specific aspects of family violence were conducted in 1992 (Kaufman Kantor et al. 1994) and 1995 (Straus et al. 1998). These studies provide a better estimate of the prevalence of family violence than is possible from police statistics or crime studies. They were made possible by the development of the "Conflict Tactics Scales" to measure family violence (Straus 1990; Straus et al. 1996). The resulting rates, which are based on the 6,002 households surveyed in 1985, are many times greater

than rates based on cases known to child welfare professionals, the police, shelters, or the National Crime Survey, but they are still believed to be lower-bound estimates.

Sixteen percent of the couples surveyed reported one or more incidents involving physical violence during the previous twelve months. Attacks by husbands on wives that were serious enough to warrant the term "wife beating" (because they involved punching, biting, kicking, choking, etc.) were reported for 3.2 percent of wives, resulting in a lower-bound estimate of 1.7 million beaten women. The National Family Violence Surveys, and all other studies of marital violence that do not use samples selected from the clientele of shelters and similar agencies, find that women assault their husbands at about the same rate as men assault their wives (Straus 1999); however, women are injured at seven times the rate of injury to men (Stets and Straus 1990; Straus 1990).

The most violent role within the family is that of parent, because almost all parents use corporal punishment (Straus and Stewart 1999). More than a third of the parents of infants in the 1995 survey reported hitting their child that year. Ninety-four percent of parents of three- to five-year-old children used corporal punishment. The percentage decreased steadily from age five on, but one-third of parents of children in their early teens reported hitting the child that year.

*Child abuse* is more difficult to operationalize than *corporal punishment* because the line differentiating *abuse* from *physical punishment* is to a considerable extent a matter of social norms. If one includes hitting a child with an object such as a belt, paddle, or hairbrush, the 1995 national survey data found a rate of 4.9 percent (Straus et al. 1998), which is twelve times higher than the rate of cases reported to protective service agencies in 1995.

Intrafamily relationships between children are also extremely violent. But, like the violence of parents, it is not perceived as such because there is an implicit normative tolerance. Almost all young children hit a sibling, and more than a third hit a parent. Even in their late teens (age fifteen to seventeen), the rate of violence between siblings is enormous: More than two-thirds of that age group hit a sibling during the year of the survey.

## EXPLANATIONS OF FAMILY VIOLENCE

Numerous family characteristics affect the level of family violence, several of which are discussed below.

**High Level of Family Conflict.** One characteristic of the family that helps account for the high rate of violence is its inherently high level of conflict. One reason for high conflict is that, as in other primary groups, family members are concerned with “the whole person.” Consequently, there are more issues over which conflict can occur than in nonprimary relationships. Moreover, when conflict does occur, the deep commitment makes arguments emotionally charged. A disagreement about music with colleagues at work is unlikely to have the same emotional intensity as when children favor rock and parents favor Bach. The likelihood of conflicts is further multiplied because families usually consist of both males and females and parents and children, thus juxtaposing differences in the orientations and interests of different genders and generations. The family is the prime locus of the “battle of the sexes” and the “generation gap.”

**Norms Tolerating or Requiring Violence.** Although conflict is endemic in families, it is not the only group or institution with a high level of conflict. Conflict is also high in academic departments and congressional committees, yet physical violence is practically nonexistent in those groups. Additional factors are needed to explain why violence is so much more frequent in the family than in other groups. One of these is the existence of cultural norms that tolerate or require violence. The clearest example is the right of parents to use corporal punishment to correct and control a child. At least two-thirds of Americans believe that “it is sometimes necessary to discipline a child with a good hard spanking” (Straus and Mathur 1996). These norms contrast with those prevailing within other institutions. Even prison authorities are no longer permitted to use corporal punishment.

Similar norms apply to husband–wife relations. However, they are implicit and taken for granted, and therefore largely unrecognized. Just as parenthood gives the right to hit, so the marriage license is also an implicit hitting license (Gelles and Straus 1988; Greenblat 1983). As with other licenses, rules limit its use. Slapping a spouse,

for example, is tolerable if the spouse is perceived to be persisting in a serious wrong and “won’t listen to reason.” Many of the men and women interviewed by Gelles (1974) expressed this normative principle with such phrases as “I asked for it” or “She needed to be brought to her senses” (p. 58).

The common law right of a husband to use corporal punishment on an “errant wife” was recognized by U.S. courts until the late nineteenth century (Calvert 1974). Informally, it lived on in the behavior of the public, the police, and the courts, and it continues to do so. Under pressure from the women’s movement, this is changing, but slowly. There have been major reforms of police and legal procedures, but the general public and many police officers continue to believe that “it’s their own business” if spouses are violent to each other, provided the blow is not severe enough to cause an injury that requires medical treatment, whereas they would not tolerate a similar pattern of assault in an office, factory, or church. Only a very small percent of men believe that a legal sanction would be likely if they assaulted their wife (Carmody and Williams 1987). In one study, of the more than 600 women who were assaulted by their husbands, the police were involved in only 6.7 percent of the incidents and an arrest was made in only five cases (Kaufman Kantor and Straus 1990). The probability of legal sanction for assaulting a wife is even less than the .008 indicated by those five cases, because two-thirds of the 600 women were assaulted more than once during the year of the survey.

**Family Socialization in Violence.** In a certain sense it begs the question to attribute the high rate of family violence to norms that tolerate, permit, or require violence because it does not explain why the norms for families are different from those for other social groups or institutions. There are a number of reasons, but one of the most fundamental is that the family is the setting in which physical violence is first experienced and in which the normative legitimacy of violence is learned. As noted above, corporal punishment is experienced by at least 94 percent of American children (Straus and Stewart 1999). Corporal punishment is used to teach that certain types of behavior are not condoned, but simultaneously,

social learning processes teach the legitimacy of and behavioral script for violence.

The example of corporal punishment also links love with violence. Since corporal punishment begins in infancy, parents are the first, and usually the only, ones to hit an infant. From the earliest level of psychosocial development, children learn that those to whom they are most closely bonded are also those who hit. Second, since corporal punishment is used to train the child in morally correct behavior or to teach about danger to be avoided, it establishes the normative legitimacy of hitting other family members. Third, corporal punishment teaches the cultural script for use of violence. For example, parents often refrain from hitting until their anger or frustration reaches a certain point. The child therefore learns that anger and frustration justify the use of physical force.

As a result of these social learning processes, use of violence becomes internalized and generalized to other social relationships, especially such intimate relationships as husband and wife and parent and child. The National Family Violence Surveys found that the more corporal punishment experienced as a child, the higher the probability of hitting a spouse (Straus et al. 1980; Straus and Yodanis 1996). Many children do not even need to extrapolate from corporal punishment of children to other relationships because they directly observe role models of physical violence between their parents.

**Gender Inequality.** Despite egalitarian rhetoric and the trend toward a more egalitarian family structure, male dominance in the family and in other spheres remains an important cause of family violence (Straus 1976). Most Americans continue to think of the husband as the “head of the family,” and many believe that status gives him the right to have the final say. This sets the stage for violence, because force is ultimately necessary to back up the right to have the final say (Goode 1974).

Numerous structural patterns sustain the system of male dominance: The income of women employed full time is about a third lower than the income of men, and money is a source of power. Men tend to marry women who are younger, shorter, and less well educated; and age, physical

size, and education form a basis for exercising power. Thus, the typical marriage begins with an advantage to the man. If the initial advantage changes or is challenged, many men feel morally justified in using their greater size and strength to maintain the right to have the final say, which they perceive to have been agreed on at the time of the marriage (LaRossa 1980). As a result, male-dominant marriages have been found to have the highest rate of wife beating (Coleman and Straus 1986; Straus et al. 1980), and societies in which male-dominant marriages prevail have higher rates of marital violence than more egalitarian societies (Levinson 1989; Straus 1994).

The privileged economic position of men also helps to explain why beaten wives so often stay with an assaulting husband (Kalmuss and Straus 1983). Women with full-time jobs earn only about 65 percent of what men earn (U.S. Bureau of the Census 1992). When marriages end, children stay with the mother in about 90 percent of the cases. Child support payments are typically inadequate and often defaulted on after a year or two. No-fault divorce has worked to the economic disadvantage of women (Weitzmen 1986). Consequently, many women stay in violent marriages because the alternative is bringing up their children in poverty.

**Other Factors.** Many other factors contribute to the high rate of intrafamily violence in the United States, even though they do not explain why the family is, on the average, more violent than other groups. Space permits only some of these to be identified briefly.

The empirical evidence shows that the greater the number of stressful events experienced by a family, the higher the rate of marital violence and child abuse (Makepeace 1983; Straus 1980; Straus and Kaufman Kantor 1987). In addition to specific stressful events that impinge on families, chronic stresses, such as marital conflict and poverty, are also strongly associated with child abuse and spouse abuse.

Almost all studies find a strong association between drinking and family violence (Coleman and Straus 1983; Kaufman Kantor and Straus 1987). However, even though heavy drinkers have two to three times the violence rate of abstainers, most heavy drinkers do *not* engage in spouse abuse or child abuse (Kaufman Kantor and Straus 1987).

The higher the level of nonfamily violence in a society, the higher the rate of child abuse and spouse abuse (Levinson 1989; Straus 1977). The nonfamily violence can be in the form of violent crime or socially legitimate violence such as warfare. The carryover of violent behavior from one sphere of life to another may be strongest when the societal violence is "legitimate violence" rather than "criminal violence," because most individual acts of violence are carried out to correct some perceived wrong. Archer and Gartner (1984) and Huggins and Straus (1980) (1980) found that war is associated with an increase in interpersonal violence. Straus constructed an index to measure differences between the states of the United States in the extent to which violence was used or supported for socially legitimate purposes, such as corporal punishment in the schools or expenditure per capita on the National Guard (Baron and Straus 1989). The higher the score of a state on this Legitimate Violence Index, the higher the rate of *criminal* violence such as homicide (Baron and Straus 1988) and rape (Baron and Straus 1989).

Family violence occurs at all social levels, but it is more prevalent at the lowest socioeconomic level and among disadvantaged minorities. Socioeconomic group differences in corporal punishment of children or slapping of spouses are relatively small, but the more severe the violence, the greater the socioeconomic difference. Thus, punching, biting, choking, attacking with weapons, and killing of family members occur much more often among the most disadvantaged sectors of society (Linsky et al. 1995; Straus et al. 1980).

**The Overall Pattern.** No single factor, such as male dominance or growing up in a violent family, has been shown to account for more than a small percentage of the incidence of child abuse or spouse abuse. However, a study of the potential effect of twenty-five such "risk factors" found that in families where only one or two of the factors existed, there were no incidents of wife beating during the year studied. On the other hand, wife beating occurred in 70 percent of the families with twelve or more of the twenty-five factors (Straus et al. 1980). Similar results were found for child abuse. Thus, the key to unraveling the paradox of family violence appears to lie in understanding the interplay of numerous causal factors.

## THE FUTURE

During the period 1965 to 1985, the age-old phenomena of child abuse and wife beating underwent an evolution from "private trouble" to "social problem" and, in the case of wife beating, to a statutory crime. Every state in the United States now employs large numbers of "child protective service" workers, and there are national and local voluntary groups devoted to prevention and treatment of child abuse. There are more than 1,000 shelters for battered women, whereas none existed in 1973. There are growing numbers of treatment programs for batterers and of family dispute mediation programs. Criminal prosecution of violent husbands, although still the exception, has become frequent, often with mandated participation in a treatment program as an alternative to fines or incarceration (Sherman et al. 1992). After lagging behind the states for more than a decade, in 1994 Congress passed the Violence Against Women act, which provides funds for services, education, and research. In the 1990s there was also an exponential growth of family therapy by psychologists and social workers, and psychology replaced sociology as the discipline most active in research on family violence.

Comparison of the 1975, 1985, and 1992 National Family Violence Surveys found a substantial reduction in the rates of child abuse and wife beating (Straus 1995; Straus and Gelles 1986; Straus et al. 1997). However, domestic assaults *by* women remained about the same, perhaps because there has been no national effort to confront this aspect of family violence and perhaps because it is a perverse aspect of the movement toward gender equality. These decreases in family violence are an acceleration of a centuries-long trend. The acceleration probably results from a combination of the educational programs, services, and legal changes discussed in this article; as well as from changes in characteristics of the family and the society that lie at the root of family violence. These include increases in the educational level of the population, later marriages, and fewer children, all of which reduce family stress; parent education programs and media, which help parents manage children without corporal punishment or more severe violence; greater equality between men and women, which reduces some of the power struggles that lead to violence; and wives in paid jobs and a greater acceptability of divorce, which (along with

shelters for battered women) enables more women to escape from violent marriages. Although child abuse and spouse abuse rates have declined, they are still extremely high. American society still has a long way to go before a typical citizen is as safe in his or her own home as on the streets or in the workplace.

## REFERENCES

- Archer, D., and R. Gartner 1984 *Violence and Crime in Cross-National Perspective*. New Haven, Conn.: Yale University Press.
- Baron, L., and M. A. Straus 1988 "Cultural and Economic Sources of Homicide in the United States." *The Sociological Quarterly* 29(3):371-190.
- 1989 *Four Theories of Rape in American Society: A State-Level Analysis*. New Haven, Conn.: Yale University Press.
- Calvert, R. 1974 "Criminal and Civil Liability in Husband-Wife Assaults." In S. K. Steinmetz and M. A. Straus, eds., *Violence in the Family*. New York: Harper and Row.
- Carmody, D. C., and K. R. Williams 1987 "Wife Assault and Perceptions of Sanctions." *Violence and Victims* 2(1):25-38.
- Coleman, D. H., and M. A. Straus 1983 "Alcohol Abuse and Family Violence." In E. Gottheil, K. A. Druley, T. E. Skolada, and H. M. Waxman, eds., *Alcohol, Drug Abuse and Aggression*. Springfield, Ill.: Thomas.
- 1986 "Marital Power, Conflict, and Violence in a Nationally Representative Sample of American Couples." *Violence and Victims* 1(2):141-157.
- Fedorowycz, O. 1999 "Homicide in Canada-1998." *Juristat: Canadian Centre for Justice Statistics*, 19(10).
- Gaquin, D. A. 1977-1978 "Spouse Abuse: Data from the National Crime Survey." *Victimology* 2:632-643.
- Gelles, R. J. 1974 *The Violent Home: A Study of Physical Aggression Between Husbands and Wives*. Beverly Hills, Calif.: Sage.
- , M. A. Straus 1979 "Determinants of Violence in the Family: Toward a Theoretical Integration." In W. R. Burr, F. Rueben Hill, I. Nye, and I. L. Reiss, eds., *Contemporary Theories About the Family*. New York: Free Press.
- , and M. A. Straus 1988 *Intimate Violence*. New York: Simon and Schuster.
- Goode, W. J. 1974 "Force and Violence in the Family." In S. K. Steinmetz and M. A. Straus, eds., *Violence in the Family*. New York: Harper and Row.
- Greenblat, C. S. 1983 "A Hit Is a Hit Is a Hit . . . Or Is It? Approval and Tolerance of the Use of Physical Force by Spouses." In D. Finkelhor, R. J. Gelles, G. T. Hotaling, and M. A. Straus, eds., *The Dark Side of Families*. Beverly Hills, Calif.: Sage.
- Greenfeld, L. A., M. R. Rand, D. Craven, P. A. Klaus, C. A. Perkins, C. Ringel, G. Warchol, C. Maston, and J. A. Fox 1998 *Violence by Intimates: Analysis of Data on Crimes by Current or Former Spouses, Boyfriends, and Girlfriends* (Bureau of Justice Statistics Factbook: No. NCJ-167237 ). Washington, D.C.: U.S. Department of Justice.
- Huggins, M. D., and M. A. Straus 1980 "Violence and the Social Structure as Reflected in Children's Books from 1850 to 1970." In M. A. Straus and G. T. Hotaling, eds., *The Social Causes of Husband-Wife Violence*. Minneapolis: University of Minnesota Press.
- Kalmuss, D. S., and M. A. Straus 1983 "Feminist, Political, and Economic Determinants of Wife Abuse Services in American States." In D. Finkelhor, R. J. Gelles, G. T. Hotaling, and M. A. Straus, eds., *The Dark Side of Families: Current Family Violence Research*. Newbury Park, Calif.: Sage.
- Kaufman Kantor, G., J. L. Jasinski, and E. Aldarondo 1994 "Sociocultural Status and Incidence of Marital Violence in Hispanic Families." *Violence and Victims* 9:207-222.
- Kaufman Kantor, G., and M. A. Straus 1987 "The 'Drunken Bum' Theory of Wife Beating." *Social Problems* 34(3):213-230.
- 1990 "Response of Victims and the Police to Assaults on Wives." In M. A. Straus and R. J. Gelles, eds., *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*. New Brunswick, N.J.: Transaction Publishers.
- Larossa, R. 1980 "And We Haven't Had Any Problems Since: Conjugal Violence and the Politics of Marriage." In M. A. Straus and G. T. Hotaling, eds., *The Social Causes of Husband-Wife Violence*. Minneapolis: University of Minnesota Press.
- Levinson, D. 1989 *Family Violence in Cross-Cultural Perspective*, vol. 1. Newbury Park, Calif.: Sage.
- Linsky, A. S., R. Bachman, and M. A. Straus 1995 *Stress Culture and Aggression*. New Haven, Conn.: Yale University Press.
- Makepeace, J. M. 1983 "Life Events Stress and Courtship Violence." *Family Relations* 32(January):101-109.
- Nelson, B. J. 1984 *Making an Issue of Child Abuse: Political Agenda Setting for Social Problems*. Chicago: University of Chicago Press.
- Pfohl, S. J. 1977 "The Discovery of Child Abuse." *Social Problems* 24(February):310-323.

- Radbill, S. X. 1987 "Children in a World of Violence: A History of Child Abuse and Infanticide." In R. E. Helfer and C. H. Kempe, eds., *The Battered Child*, 4th ed. Chicago: University of Chicago Press.
- Sherman, L. W., D. A. Smith, J. D. Schmidt, and D. P. Rogan 1992 "Crime, Punishment, and Stake in Conformity: Legal and Informal Control of Domestic Violence." *American Sociological Review* 57:680-690.
- Stets, J. E., and M. A. Straus 1990 "Gender Differences in Reporting of Marital Violence and its Medical and Psychological Consequences." In M. A. Straus and R. J. Gelles, eds., *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*. New Brunswick, N.J.: Transaction Publishers.
- Straus, M. A. 1976 "Sexual Inequality, Cultural Norms, and Wife-Beating." In E. C. Viano, ed., *Victims and Society*. Washington, D.C.: Visage Press.
- 1977 "Societal Morphogenesis and Intrafamily Violence in Cross-Cultural Perspective." *Annals of the New York Academy of Sciences* 285:718-730.
- 1980 "Social Stress and Marital Violence in a National Sample of American Families." *Annals of the New York Academy of Sciences* 347:229-250.
- 1986 "Domestic Violence and Homicide Antecedents." *Bulletin of the New York Academy of Medicine* 62:446-465.
- 1987 "Primary Group Characteristics and Intra-Family Homicide." Paper Presented At the Third National Conference For Family Violence Researchers, Family Research Laboratory, Durham, NH.
- 1990 "The Conflict Tactics Scales and its Critics: An Evaluation and New Data on Validity and Reliability." In M. A. Straus and R. J. Gelles, eds., *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*. New Brunswick, N.J.: Transaction Publishers.
- 1994 "State-to-State Differences in Social Inequality and Social Bonds in Relation to Assaults on Wives in the United States." *Journal of Comparative Family Studies* 25(1):7-24.
- 1995 "Trends in Cultural Norms and Rates of Partner Violence: An Update to 1992." In S. Stith and M. A. Straus, eds., *Understanding Partner Violence: Prevalence, Causes, Consequences, and Solutions*. Minneapolis, Minn.: National Council on Family Relations.
- 1999 "The Controversy over Domestic Violence by Women: A Methodological, Theoretical, and Sociology of Science Analysis." In X. Arriaga and S. Oskamp, eds., *Violence in Intimate Relationships*. Thousand Oaks, Calif.: Sage.
- , and R. J. Gelles 1986 "Societal Change and Change in Family Violence from 1975 to 1985 as Revealed by Two National Surveys." *Journal of Marriage and the Family* 48:465-479.
- , R. J. Gelles, and S. K. Steinmetz 1980b *Behind Closed Doors: Violence in the American Family*. New York: Doubleday/Anchor.
- , S. L. Hamby, S. Boney-McCoy, and D. B. Sugarman 1996 "The Revised Conflict Tactics Scales (CTS2): Development and Preliminary Psychometric Data." *Journal of Family Issues* 17(3):283-316.
- , S. L. Hamby, D. Finkelhor, D. W. Moore, and D. Runyan 1998 "Identification of Child Maltreatment with the Parent-Child Conflict Tactics Scales: Development and Psychometric Data for a National Sample of American Parents." *Child Abuse and Neglect* 22:249-270.
- , and G. Kaufman Kantor 1987 "Stress and Child Abuse." In R. E. Helfer and R. S. Kempe, eds., *The Battered Child*, 4th ed. Chicago: University of Chicago Press.
- , G. Kaufman Kantor, and D. W. Moore 1997 "Change in Cultural Norms Approving Marital Violence: From 1968 to 1994." In G. Kaufman Kantor and J. L. Jasinski, eds., *Out of the Darkness: Contemporary Perspectives on Family Violence*. Thousand Oaks, Calif.: Sage.
- , and A. K. Mathur 1996 "Social Change and Change in Approval of Corporal Punishment by Parents from 1968 to 1994." In D. Frehsee, W. Horn, and K.-D. Bussmann, eds., *Family Violence Against Children: A Challenge for Society*. New York: Walter deGruyter.
- , and J. H. Stewart 1999 "Corporal Punishment by American Parents: National Data on Prevalence, Chronicity, Severity, and Duration, in Relation to Child and Family Characteristics." *Clinical Child and Family Psychology Review* 2(2):55-70.
- , and C. L. Yodanis 1996 "Corporal Punishment in Adolescence and Physical Assaults on Spouses Later in Life: What Accounts for the Link?" *Journal of Marriage and the Family* 58(4):825-841.
- U.S. Bureau of the Census 1992 *Statistical Abstract of the United States*. Washington, D.C.: Author.
- U.S. Department of Justice 1980 *Intimate Victims: A Study of Violence Among Friends and Relatives* (A National Crime Survey Report). Washington, D.C.: Bureau of Justice Statistics.
- Weitzmen, L. J. 1986 *The Divorce Revolution*. New York: Free Press.

MURRAY A. STRAUS



## FEMINIST THEORY

The term *feminist theory* is an invention of the academic branch of the mid- and late twentieth-century feminist movement. It refers to generating systematic ideas that define women's place in society and culture, including the depiction of women—large questions, indeed. The task of feminist theorists is necessarily monumental. It requires the wisdom, courage, and perseverance that Penelope displayed as she wove and unwove her tapestry to trick the suitors who sought to appropriate her kingdom and so steal her child's birthright.

For many reasons the task of feminist theorists is difficult. First, it is interdisciplinary. Literary critics, art historians, musicologists, historians, and philosophers—to name some specialists associated with the humanities—have all offered powerful and sometimes conflicting ideas about women in society and culture. So have sociologists, anthropologists, economists, psychologists, and psychoanalysts. Although the biological and physical sciences do not usually make fruitful contributions to contemporary debates about social and cultural issues, feminist scientists have posed questions that challenge the presuppositions of their own fields. They, too, have augmented the scope of feminist theory. Indeed, specialists in so many disciplines have offered apt ideas that no one essay or writer can even pretend to outline the scope of contemporary feminist theory. This presentation will concentrate on ideas developed in the social sciences.

Second, because feminist theory has its basis in the current women's movement, it is necessarily infused with the political concerns of the contemporary era. (Any system of arranging facts, including the writing of history, is necessarily influenced by the dominant concerns and ideologies of its times.) The most important of these is the relationship among race (ethnicity), gender, and class, both cross-culturally and historically. For even as Americans and Europeans have sought to confront institutionalized racism and sexism, as well as the unequal distribution of income and wealth in their own nations, women in developing nations have posed issues regarding the application of generalizations based on those experiences to their own situations. So, too, historical research has

raised the challenge of process—namely, the problem that any particular historical outcome is not predetermined, so that the development of relationships among gender, class, and race (or ethnicity) may vary greatly. Such variations make the act of generalizing hazardous, if not foolhardy.

Third, feminist theory has not merely existed in a sociopolitical context but has been informed by it. This means that many theorists realize that their ideas have been influenced by their own material conditions and cultures. Thus, they have had to confront epistemological issues, including the meaning of objectivity and the way male dominance has shaped notions important to all branches of human inquiry. Put somewhat differently, theorists have broached two issues: (1) the inextricable association between ideas and methods of inquiry, and (2) how both dominant ideas and methods have been influenced by the male hegemony over academic and scientific discourse. As is true in most contemporary fields of study, each of these issues is controversial.

Given these challenges, one might well wonder why anyone would try to generate feminist theory. But feminist academics felt that they could make a significant contribution by using their training first to document and later to analyze women's place in society. When the feminists of academe began to debate their understandings of women's place in society and culture, no sure path seemed available. With the exception of Simone de Beauvoir's *The Second Sex* (1952), men had penned the two canonized (nineteenth-century) texts most familiar to these academics, namely John Stuart Mill's *The Subjection of Women* and Frederick Engels's *The Origins of Private Property, the Family, and the State*. Although Western women had debated their own situation since at least 1400, when, as Joan Kelly (1982) notes, Christine de Pisan “sparked . . . the four-century-long debate . . . known as the *querelle des femmes*,” twentieth-century academics were largely ignorant of that polemical tradition. Instead, they had been schooled in the thought of great men—whose writings were included in the first anthologies of feminist thought, such as Miriam Schneir's *Feminism: The Essential Historical Writings* (1972) and Alice Rossi's *The Feminist Papers* (1973). Such anthologies also introduced American academics to great women outside their own fields.

Nonhistorians learned of the generative ideas of participants in the Seneca Falls convention of 1848, as well as about Susan B. Anthony, Emma Goldman, and the British activist Emily Pankhurst; noneconomists met Charlotte Perkins Gilman; non-literary critics met Virginia Woolf. Although feminist intellectuals might find strength in the knowledge that other women had provided trail markers to guide their way, antifeminists were not convinced that gender inequality still existed.

Thus, the first task confronting feminist theory was to document both past and present inequalities. Many of the early writings addressing this project discussed women as either “other” or “victim.” These characterizations ran through writings that might be classed as either liberal (the belief that women have the same capabilities as men and should receive equal treatment); socialist Marxist (variations of the notion that capitalism created or augmented gender inequality); or radical feminist (versions of the idea that women are inextricably different from men and at least equal, and possibly superior, to them).

#### WOMAN AS OTHER AND AS VICTIM

To some extent the notions of “other” and “victim” are implicit in any mid-twentieth-century demonstration of inequality. Those who are maltreated for unacceptable reasons appear to be victims, as implicit in the late 1960s’ and early 1970s’ political slogan “Don’t blame the victim.” They seem to be “others” because of the historical and cross-cultural tendency of dominant groups to justify discriminatory actions by arguing that members of subdominant groups are “alien,” not fully human, or simply “not like us.” (In the American case, blacks were deemed “not fully human” when procedures for counting the male population were defined in the early years of the Republic.) This dichotomy between “subject” and “object,” “self” and “other,” has also been crucial to modern European thought, including the philosophic basis for de Beauvoir’s *The Second Sex*, which was Jean-Paul Sartre’s concept of existentialism. Such notions seemed to provide a conceptual framework with which to document *sexism*, a term introduced by those members of the American women’s movement who had participated in the civil rights movement and who wished for a term that

reverberated with connotations of despicable conduct implicit in the more familiar term *racism*. That is, academic feminists could view themselves as demonstrating how specific practices or institutions viewed women as “others,” maltreated them, and so transformed them into “victims” not responsible for their “despised” status. Once those processes were identified, feminist activists could seek to reform or to revolutionize the relevant institutions.

Social scientists provided confirmation of victimization by gathering data comparing women and men. Men were more likely to dominate professions (even such “female work” as grade school teaching and librarianship, in which men were likely to be principals and department heads), earn more money, receive higher education, be awarded scholarships and fellowships, earn advanced degrees, hold positions of political leadership, be granted credit cards, be treated as legally responsible for their actions, and be permitted to make decisions about their own bodies. (Both theorists and activists hotly discussed the “body issues” such as abortion, incest, rape, and sexual harassment and wife battering.) Psychologists and psychiatrists pointed out that their colleagues had equated mental health with supposedly male characteristics. Humanists demonstrated that in art, music, and literature, men had inscribed themselves in the “cultural canon.” Not only did the canon identify men’s accomplishments as the “most important” Western works, but the so-called Western cultural tradition also represented history, literature, art, and philosophy from a male point of view. Sometimes these great works dwelled on the dichotomy between the concepts of madonna and whore; sometimes on the secular objectification of women’s sexuality (as seen in renditions of the female nude). Whether religious or secular, both the cultural canon and academic knowledge were discovered to ignore or belittle women in various ways, as well as to devalue their contributions to civic and cultural life throughout the centuries.

Feminists confronted the dilemma of what to do about this devaluation. Liberal feminists seemed to echo one theme implicit in de Beauvoir’s *The Second Sex* and Mill’s *The Subjection of Women*: Become more like men—that is, remove the barriers preventing women from having the same opportunities as men, and, in the future, women will accomplish as much as men. Because, as some

psychologists argued, there is no innate difference between women and men, equal treatment and equal opportunity will result in equal accomplishment.

The solution offered by Marxist and socialist feminists was not all that different from the ideas of liberals. They, too, believed that the eradication of obstacles would liberate women. But Marxist and socialist feminists were haunted by the “problem of the hyphen.” That is, for them, the barriers confronting women were not simply posed by what all feminists termed “patriarchy” (shorthand for “male dominance”). Rather, as they saw it, patriarchy was itself inextricably related to capitalism. So, disentangling that relationship was a complex task. If capitalism had been a primary cause of women’s inferior position, then women should have found greater equality in noncapitalist nations. In actuality, women had *not* prospered under the brands of communism found in the Soviet Union, Eastern Europe, or China. In those nations, too, women were clustered in jobs that paid less than those filled by men of equivalent education.

As feminists learned, communist lands had treated women favorably while still in their revolutionary stages. For example, in the Soviet Union of the 1920s, “changes in property relationships and inheritance laws weakened the family as an economic unit and reduced the dominance of the male household head, while new family codes undermined the legal and religious basis of marriage and removed restriction on divorce” (Lapidus 1978, p. 60). By undercutting the power of both the church and the traditional family, these measures strengthened the state. Once the Communist party had institutionalized its power and declining birthrates challenged economic growth, however, it redefined the family as “the bulwark of the social system, a microcosm of the new socialist society . . . [supposed] to serve above all as a model of social order” (Lapidus 1978, p. 112). Divorce became difficult; motherhood was defined as a contradiction, simultaneously a joy and the “supreme obligation of Soviet women.” As in capitalist countries, the Soviet Union then began to glorify women’s role in the family (what feminist theory identified as the private sphere).

The discrepancy that arose between communist practice and socialist theory created a theoretical dilemma. One might insist that the so-called communist countries had radically departed from

the theoretical ideal, and so the impact of socialism on women’s lives had yet to receive a valid test. One might point to the relatively enlightened laws of the Scandinavian societies, where social policies assisted women who tried to combine work and family life. But even in these nations, women assumed more of the responsibility for children than did men. Although Scandinavian laws enabled father or mother to take parental leave after the birth or adoption of a child, few men exercised that legal right. Thus, another option seemed necessary: One might seek to reconceptualize the link between private property and patriarchy.

Anthropologists and historians were among the first feminists to attempt that (re)vision. Three of their solutions were particularly influential. First, drawing on de Beauvoir, some anthropologists (and at least one sociologist) returned to the idea of woman as other (see Rosaldo and Lamphere 1974). They suggested that extant societies embodied an analogy: Woman is to nature as man is to culture. That is, traditional and industrialized societies assumed that woman is closer to nature than man is. Men had supposedly thrust themselves upon nature and transformed it.

Second, some theorists incorporated Marxist notions by pointing out that men had defined women as private property. Rubin (1975) offered the most influential argument about what she termed “the sex/gender system.” Assuming that women and men are more like than unlike one another, she asked how societies create “difference” or transform sex into gender. Answering her question, she retained her anthropologist’s conviction that kinship relations are at the basis of society while she drew on her own “freely interpretive” readings of Claude Levi-Strauss and Sigmund Freud. In essence, Rubin argued, men exchange women to create and to cement their own social relationships. This exchange “does imply a distinction between gift and giver. If women are the gifts, then men are the exchange partners. And it is the partners, not the presents, upon whom reciprocal exchange confers its quasi-mystical powers of social linkage. The relations of such a system are such that women are in no position to realize the benefits of their own circulation. As long as the relations specify that men exchange women, it is the men who are the beneficiaries of the product of such exchanges—social organization” (Rubin 1975, p. 174).

Prohibitions on incest keep this exchange system going because they intend to ensure the availability of women to be exchanged. Yet the women must be willing to be “gifts”; that is, they must have internalized the appropriate societal norms. Supposedly, what Freud describes as the Oedipal complex provides that internalization. According to Rubin, the Oedipal complex is a record of “how [contemporary] phallic culture domesticates women” (1975, p. 198). Furthermore, psychoanalytic findings about women’s inferiority to men is a palimpsest “of the effects in women of their domestication” (p. 198). Thus, any society based on the exchange of women has molded the inequalities of the sex/gender system into its very essence. According to Rubin, this generalization is as applicable to today’s industrialized societies as to nonindustrial societies. What feminists term *patriarchy* is actually the operation of the sex/gender system. By implication, to achieve equality, feminists had to challenge the sex/gender system.

Third, some historians and anthropologists responded to the conflation of capitalism and patriarchy by seeking nonindustrial models where women held power. Mainly they sought examples of women as a force in the public sphere. Eleanor Leacock wrote about one classic case, the Iroquois. In the Iroquois Confederation, women elected the (male) chiefs and were also empowered to remove them from office. In medieval society, Joan Kelly maintained, (aristocratic) women had power over the education of their daughters. During the Renaissance, men absconded with that power, devalued the knowledge women had shared, deprived women of the right to educate one another, and also denied women equal access to the then newly discovered “classics.” Kelly concluded her article by suggesting that the Renaissance had a different meaning for women and men; and so, by challenging historians’ periodization, another conclusion was also possible. One might infer that during the Renaissance, men transformed women into victims.

By viewing women as either other or victim, all these theorists were implicitly accepting the male assumption that the public sphere is more important than the private. Even the search for examples of women who had once collectively held significant political power can be viewed as an affirmation of the dominant (male) view that the public sphere is more important than the private (home).

However, the third approach—the search for examples of institutionalized female power—also foreshadowed a new phase of feminist theory: “the (re)vision of public and private spheres.” As introduced by the poet Adrienne Rich, the term *(re)vision* is a deliberate pun referring to both a reconsideration of past ideas and a new vision of women’s role in society.

### THE (RE)VISION OF PUBLIC AND PRIVATE SPHERES

In the late 1970s, the feminist movement was maturing; many middle- and upper-class women (including nonfeminists) began to flock to the male-dominated professional schools from which they had once been excluded. Feminists in several fields began to reassess the value of activity in the private sphere—the world of the home in which most women were ensconced. Could the private sphere serve as a launchpad for social change? Had it ever done so? Are these spheres indeed separate, or does the persistence of this dichotomy conflate the errors of nineteenth-century thought? (A reconsideration of the relationship between the public and private spheres is implicit in the titles of such books as *Beyond Separate Spheres* and *Private Woman, Public Stage*. In sociology, Marxist and socialist feminists use different language to discuss women’s and men’s spheres.)

Nineteenth-century social theories had implied that the private sphere was of equal importance to the public world of work. Although the early nineteenth-century “cult of domesticity” banished women to the home, where they were to serve as models of religiosity and virtue, they were also enshrined as “mothers of civilization”—a role that might imply power. But from the vantage of the late twentieth century, the role of “mother of civilization” did not seem so vital. If the private sphere was so important, why was the role of “parent of civilization” not available to men, who historically seemed almost to have monopolized positions of power? If the roles of women and men (wives and husbands) were of equal importance, as Talcott Parsons had implied, why were American women more likely than men to complain of the sorts of physical and mental ailments associated with an inferior social position?

Yet, in the late 1970s, historians, anthropologists, and some sociologists began to find positive

aspects of women's role in the private sphere. According to historians, women had used their roles to initiate social reforms. They had been especially active in trying to ameliorate some of the social problems resulting from the transformation of an agricultural society into an industrial one. For example, through voluntary associations, middle- and upper-class women in New York tried to decrease the destructive impact of poverty on the poor, especially on poor women. In Oneida County, they sought to reform the behaviors of the many single men who had moved to the city from the farms, lived in boarding houses, and sometimes disrupted the civil order. That the activities of the volunteers resulted in the enshrinement of women in the home once their activities had been successful is a historical irony (Ryan 1981). But that outcome is irrelevant to the main theoretical point offered by feminist theorists: Activity that nineteenth-century women had viewed as an extension of their roles in the private sphere had indeed influenced the public sphere. Put somewhat differently, the domestic and public spheres are not necessarily dichotomous.

Yet problems remain. First, variations suggest that generalization is premature. Second, rejection of the dichotomy between domestic and public spheres challenges the residues of nineteenth-century thought remaining in twentieth-century theories but does not necessarily lead to new theoretical formulation. Indeed, neither historical nor anthropological discoveries of variations on the common pattern—male dominance—*necessarily* facilitate theorizing. Rather, they might and did lead some feminists to search for the origins of male dominance (as in the article by Gayle Rubin) and to champion causal explanations.

Yet the (re)vision of public and private spheres did enable some feminist theorists to ask new questions. The anthropologist M. Z. Rosaldo (1980) explains:

*Sexual asymmetry can be discovered in all human social groups, just as can kinship systems, marriages, and mothers. But asking "Why?" or "How did it begin?" appears inevitably to turn our thoughts from an account of the significance of gender for the organization of all human institutional forms (and reciprocally, of the significance of all social facts to gender) toward dichotomous*

*assumptions that link the roles of men and women to the different things that they, as individuals, are apt to do.*

Rosaldo (1980) continues:

*What traditional social scientists have failed to grasp is not that sexual asymmetries exist but that they are as fully social as the hunter's or the capitalist's role, and that they figure in the very facts, like racism or social class, that social science claims to understand. A crucial task for feminist scholars emerges, then, not as the relatively limited one of documenting pervasive sexism as a social fact—or showing how we can now hope to change or have in the past been able to survive it. Instead, it seems that we are challenged to provide new ways of linking the particulars of women's lives, activities, and goals to inequalities wherever they exist.*

To advance beyond naivete, feminist theories must grasp *how* meanings of gender are constructed, not why they exist.

Sociologist Nancy Chodorow provided one such demonstration in her now-classic but still controversial *The Reproduction of Mothering*. Using psychoanalytic object-relations theory and some elements of Marxist thought, she argued that in contemporary capitalist societies the roles of women and men within the family (re)produce the roles women and men are expected to fill in Western societies. Because women are responsible for the care of small children, young girls and boys initially identify with their mothers. Girls are encouraged to continue this relational identification with their mothers; boys are not. For girls, the omnipresence of women in early childhood leads to a problem of boundaries—of knowing where their mothers end and they begin. As adults, this lack of boundaries may be advantageous: Out of their ability to see the world as others do, they may have a richer emotional life than men and also be more emphatic than they are. For boys, the omnipresence of women means that men learn the meaning of masculinity through the eventual demand that they separate from their mothers and identify with a role (male gender). Theirs is a positional identification. Ultimately, Chodorow argues, these scenarios play themselves out so that women and men try to reproduce the sorts of

modern families in which they were reared. Additionally, Chodorow (1978) claims,

*An increasingly father-absent, mother-involved family produces in men a personality that both corresponds to masculinity and male dominance as these are currently constituted in the sex-gender system, and fits appropriately with participation in capitalist relations of production. Men continue to enforce the sexual division of spheres as a defense against powerlessness in the [capitalist] labor market. Male denial of dependence and of attachment to women helps to guarantee both masculinity and performance in the world of work.*

She continues,

*The relative unavailability of the father and the overavailability of the mother create negative definitions of masculinity and men's fear and resentment of women, as well as the lack of inner autonomy in men that enables, depending on the particular family constellation and class origin, either rule-following or the easy internalization of the values of the organization.*

Chodorow's theory is controversial. Some liberals object to the inference that women are more empathetic than men. They claim that psychological studies show that women and men have the same innate emotional capabilities. Nevertheless, radical feminists—people who believe that male dominance is the primary cause of women's subjugation—sometimes celebrate women's alleged superiority to men.

Others challenge Chodorow's theory by questioning whether any theory about the modern American family can possibly be applied to other historical epochs or cultures. Yet this objection misses the point in two ways. First, Chodorow discusses a specific time and place—contemporary America. Her contrast between how women and men learn their (relational or positional) roles should be even more important today than it was in 1978 when *The Reproduction of Mothering* was first published. Since 1978 the percentage of female-headed households has increased. Concomitantly, more young boys have even less contact with men; they must form the positional identifications that, Chodorow claims, prepare them to uphold the orientations to work and

family required to maintain postindustrial capitalism. Second, Chodorow anticipated Rosaldo's call to understand how gender is socially constructed to articulate with other roles. She did not ask the origin of all sex/gender systems.

Chodorow's argument is important in a third way: She transforms "normal" understandings of men's and women's roles. She turns one current interpretation of Freud's thought, object-relations theory, on its head. Woman-centered (written from the perspective of a feminist) yet comparative (examining both women and men), Chodorow's book anticipated the challenge that feminist scholarship currently offers other theoretical projects.

### CHALLENGING CONVENTIONAL INTERPRETATIONS

The theme of difference is key to Carol Gilligan's *In a Different Voice* (1981). Extending and (re)vising Lawrence Kohlberg's work on the development of moral judgments among men, Gilligan argues that Kohlberg's scale of moral maturity slights women. Kohlberg argues that the highest stages of maturity involve the application of "rules to universal principles of justice" (Gilligan 1981, p. 18). In Kohlberg's model, women seem stuck at an intermediate stage because of their lack of opportunity to enter the public sphere and so to master and to apply those universal principles. Gilligan contends that Kohlberg has erred in his assessment of women (and other subordinated groups). In the case of women, he has not understood how their concepts of morality are based on their socialization. On the basis of her empirical studies and Chodorow's argument, Gilligan concludes that women do not make inferior moral judgments but different ones. They employ a relational ethic that stresses interpersonal caring, including a responsibility for self and others.

Gilligan's theory is as controversial as Chodorow's—perhaps even more so. Radical feminists cite this theory to argue again that women are morally superior to men. Some liberal feminists, who are interested in pushing for women's entry to upper levels of corporate management, use Gilligan's theory to claim that women are better equipped than men to develop innovative management styles that bind teams to the corporation. Other liberals (e.g., Epstein 1989) insist that there

are no *innate* psychological distinctions between women and men.

However, both Chodorow and Gilligan raise an additional issue, potentially more controversial than the debates about “innate” gender differences. They try at one and the same time to follow the practices of their respective disciplines and to view social life from a woman’s perspective. Since science, social science, and the literary canon are based on male perspectives (as is true of Kohlberg’s studies), how are we to forge an epistemology (and hence a methodology) that can simultaneously claim veracity and be true to women’s experience of the social world? Can any theory that is either androcentric (man-centered) or gynocentric (woman-centered) be valid?

### FEMINIST THEORY AND THE PROBLEM OF KNOWLEDGE

Some scientists or social scientists might believe that the question of validity applies only to explicitly interpretive work. After all, many decades have passed since Heisenberg enunciated his famous principle that the technologies (and by extension the theories and methods) used to view a phenomenon necessarily influence what is viewed. Supposedly, scientists and social scientists have been taking this principle into account as they explain how their generalizations apply “all other things being equal.” But feminist theorists, like postmodern theorists, have challenged the very basis of the deductive methods at the heart of contemporary science and social science. They “question the very foundation of post-Enlightenment science and social science: the ‘objectivist illusion’ (Keller 1982) that observation can be separated from explanation, the knower from the known, theory from practice, the public from the private, culture from nature, and other dualisms that undergird systems of social stratification” (Hess 1990, p. 77). The feminist challenge differs from other deliberations about epistemology because, as Hess observes, it is “more political.”

The work of scientist Evelyn Fox Keller is exemplary of this controversy. Keller has written rather abstract philosophic essays on such issues as Plato’s epistemology and Bacon’s notions of mastery and obedience to argue that modern science is infused with “male” notions. Rather than attempt to present these complex arguments briefly,

let us concentrate on a more concrete work, Keller’s biography of biologist Barbara McClintock, *A Feeling for the Organism* (1982).

In this biography, Keller claims that McClintock accomplished her Nobel Prize-winning research on the transposition of genes in part because McClintock’s research style radically departed from the dominant male model, molecular genetics. According to Keller, the dominant model presumed a hierarchical structure of genetic organization that resembles organizational charts of corporate structure and assumes a unidirectional flow of information. That assumption permitted the quick payoffs in research on the structure of specific bacteria that facilitates significant scientific findings (and so careers). The hierarchical model, though, is also supposedly a “male” model.

McClintock, however, believed in a more complex and less hierarchical “old-fashioned” model: “To McClintock, as to many other biologists, mechanism and structure have never been adequate answers to the question ‘How do genes work?’ To her an adequate understanding would, by definition, have to include an account about how they function in relation to the rest of the cell, and, of course, to the organism as a whole” (Keller 1985, p. 168). In this view, even a genome is an organism, and it, too, must be considered in relation to its environment.

Keller does not claim that McClintock’s model was female: McClintock was trained by men. Keller does argue that many biologists have missed the essence of McClintock’s vision. Keller includes among them biologists who are trying to incorporate McClintock’s work on transposition into the hierarchical model. And, Keller believes,

*The matter of gender never does drop away. . . . The radical core of McClintock’s stance can be located right here. Because she is not a man, in a world of men, her commitment to a gender-free science has been binding; because concepts of gender have deeply influenced the basic categories of science, that commitment has been transformative. In short, the relevance of McClintock’s gender in this story is to be found not in its role in her personal socialization but precisely in the role of gender in the construction of science. (Keller 1982, p. 174)*

For, Keller explains, contemporary science names the object of its inquiry (nature) "as female and the parallel naming of subject (mind) as male" (1982 pp. 174). Thus, women scientists are faced with a necessary contradiction between their roles in the world and their role as scientist. Even as the social structure of science tends to place women on the periphery of the invisible colleges that constitute the scientific world, that contradiction may limit the creative scientific imagination of both women and men.

However, as Keller readily admits, not all men believe in the hierarchical model. Nor, to paraphrase Keller, have they all "embraced" the notion of science as a female to be put on the rack and tortured to reveal her secrets. But, Keller insists, both the naming of subject and object and the hierarchical model are androcentric and limit the possibilities of scientific inquiry.

Keller's work has been challenged on a number of grounds. Stephen Jay Gould has launched the most telling attack. His field, paleontology, is also dominated by men but rejects the hierarchical "male" model. Unlike nineteenth-century evolutionary thought, contemporary theories do not view human beings as the proud culmination of the past. Other lifeforms have been more successful. Furthermore, to reconstruct the past, paleontologists must grasp "wholes." Cynthia Fuchs Epstein adds that one cannot even argue that, as social scientists, women have been more empathic or observant than men. Epstein's counterexample is Erving Goffman, whose ability to "see" is almost legendary among sociologists.

For the purposes of this article, it is irrelevant whether Keller, Gould, or Epstein is correct. What matters is that developments in feminist theory have led feminists to participate in the postmodern debate about the nature of knowledge. During the late 1980s, this debate was at the center of controversies in the social sciences and the humanities. It will probably remain important for some years to come. The debate infuses interpretations of literary works, reconstructions of the past, and understandings of social scientific models. It seems to transcend schools of thought. Marxists are divided about postmodernism, as are liberals and, within the feminist community, radicals. Feminist theory has followed the course of all contemporary theories because feminists, too, are members of the

societies about which they write and which they are trying to change.

Throughout this essay, terms such as *liberal*, *Marxist* or *socialist*, and *radical* feminists have been intended to distinguish different political orientations and experiences that shaped each group of feminists, as so well documented in histories and sociologies of the mid-twentieth-century feminist movement. By 1990, these very different groups of feminists had become self-consciously aware; that is, many understood that even as feminism had sought social change, it, too, had shaped and been shaped by its environment. For instance, Stacey (1990) argued that in the late 1960s and early 1970s, as feminism stressed female participation in the public sphere and denigrated motherhood, it contributed to rising divorce rates and the devaluation of feminine tasks.

In the 1990s, feminist theory has maintained its own questions and has continued to be divided into camps that emphasize either the similarities between women and men or the differences between them. However, as is axiomatic, all enterprises are embedded in and shaped by their structural contexts. Feminist theory has not been an exception to this well-known rule.

In the 1990s, feminist theories have been influenced by general tendencies in other academic specialties. Although it is still rare (and perhaps because it is still rare) for someone to receive a Ph. D. from an academic department called "women's studies," "gender studies," or "feminism," feminist theory has become increasingly influenced by movements toward specialization within the social sciences. These movements have included the ever more narrow delineation of topics of inquiry, such that a business school course on women in non-profit organizations may share few if any readings with a sociology course on women and the professions and a course on men and masculinity taught in a school of family studies may not overlap an anthropology course on cross-cultural notions of masculinity. Additionally, within the social sciences, academic disciplines are so beset by specialized theories that researchers may encourage a thousand flowers to bloom by ignoring one another's ideas rather than by addressing the basic and competing philosophic premises on which their theoretical systems are built. Within sociology, for instance, the articles of feminist theorists who



pledge allegiance to rational choice theory rarely cite the works of feminist theorists who use key ideas of the symbolic interactionists—and vice versa. Indeed, they may not even agree to call themselves feminist theorists, but rather speak of themselves as rational choice theorists interested in the problems of gender or as symbolic interactionists exploring theories about how women experience their bodies under the parameters set by a national medical system.

Nonetheless, there have been some general ideas about which feminist theorists have come to agree. Perhaps the most important of these is a problem central to all social science: to wit, the theoretical tension between approaches that emphasize individuals as the agents of social stability and social change and approaches that stress how either macro- or microstructures influence and shape both individuals and their decisions. Indeed, as more social scientists have come to agree that nature and nurture interact, the theoretical problems posed by agency and structure have become more salient.

One basic way that feminist theorists have dealt with these problems is to redefine gender itself. In the past, gender has been thought of as a cultural phenomenon—either normative patterns of behaviors, roles, or scripts. Today gender itself is being viewed as a structural phenomenon. This means that many feminist theorists now define gender as an institution, as basic to social structure as the economy, religion, education, and the state—and perhaps more basic than the family (Connell 1995; Lorber 1994). Such theorists emphasize that in some societies there are more than two (dichotomous) genders and that socially and historically embedded notions of gender influence the operation of all institutions, even as they are influenced by other sociohistoric institutions. Further, these theorists continue, identification of gender as an institution enables one to recognize that gender is stratified and that it is also processual (part of unfolding social processes). Thus, this redefinition of gender enables theorists to link structure and agency by discussing how individuals' gendered behavior (re)produces gendered social structures, much as Giddens's (1984) structuration theory seeks to marry structural and agentic approaches. Furthermore, the recognition of gender as an institution entails such a broad paradigmatic shift that, at one and the same time, it refocuses

how theorists view gender and facilitates work within different theoretical traditions. Adherents of symbolic interactionism, neofunctionalism, or rational choice theory could all absorb this notion into their work and test its implications with their method of choice (Ferree et al. 1999).

(SEE ALSO: *Comparable Worth; Gender; Social Movements*)

## REFERENCES

- Chodorow, Nancy 1978 *The Reproduction of Mothering*. Berkeley: University of California Press.
- Connell, R. W. 1995 *Masculinities*. Berkeley: University of California Press.
- de Beauvoir, Simone 1952 *The Second Sex*. New York: Alfred A. Knopf.
- Ferree, Myra Marx, Judith Lorber, and Beth Hess, eds. 1999 *Revisioning Gender*. Thousand Oaks, Calif.: Sage.
- Giddens, Anthony 1984 *The Constitution of Society*. Berkeley: University of California Press.
- Gilligan, Carol 1981 *In a Different Voice*. Cambridge, Mass.: Harvard University Press.
- Hess, Beth 1990 "Beyond Dichotomies: Drawing Distinctions and Embracing Differences." *Sociological Forum* 5:75–94.
- Keller, Evelyn Fox 1982 *A Feeling for the Organism*. New York: W. H. Freeman.
- 1985 *Reflections on Gender and Science*. New Haven, Conn.: Yale University Press.
- Kelly, Joan 1982 "Early Feminist Theory and the *Querelle des femmes*, 1400–1789." *Signs* 8:4–28.
- Lapidus, Gail 1978 *Women in Soviet Society*. Berkeley: University of California Press.
- Lorber, Judith 1994 *Paradoxes of Gender*. New Haven, Conn: Yale University Press.
- Rosaldo, M. Z. 1980 "The Uses and Abuses of Anthropology." *Signs* 5:389–417.
- and Louise Lamphere, (eds.) 1974 *Women, Culture, and Society*. Stanford, Calif.: Stanford University Press.
- Rossi, Alice S., ed. 1973 *The Feminist Papers*. New York: Columbia University Press.
- Rubin, Gayle 1975 "The Traffic in Women: Notes on the Political Economy of Sex." In Rayna Rapp Reiter, ed., *Toward an Anthropology of Women*. New York: Monthly Review Press.
- Ryan, Mary P. 1981 *Cradle of the Middle Class*. New York: Cambridge University Press.

Schneir, Miriam, ed. 1972 *Feminism: The Essential Historical Writings*. New York: Random House.

Stacey, Judith 1990 *Brave New Families*. Berkeley: University of California Press.

GAYE TUCHMAN

## FEMININITY/MASCULINITY

Femininity and masculinity, or one's *gender identity* (Burke et al. 1988; Spence 1985), refer to the degree to which persons see themselves as masculine or feminine given what it means to be a man or woman in society. Femininity and masculinity are rooted in the social (one's gender) rather than the biological (one's sex). Societal members decide what being male or female means (e.g., dominant or passive, brave or emotional), and males will generally respond by defining themselves as masculine while females will generally define themselves as feminine. Because these are social definitions, however, it is possible for a person to be female and see herself as masculine or male and see himself as feminine.

It is important to distinguish *gender identity*, as presented above, from other gender-related concepts such as *gender roles*, which are shared expectations of behavior given one's gender. For example, gender roles might include women investing in the domestic role and men investing in the worker role (Eagly 1987). The concept of gender identity is also different from *gender stereotypes*, which are shared views of personality traits often tied to one's gender, such as instrumentality in men and expressiveness in women (Spence and Helmreich 1978). And gender identity is different from *gender attitudes*, which are the views of others or situations commonly associated with one's gender, such as men thinking in terms of justice and women thinking in terms of care (Gilligan 1982). Although gender roles, gender stereotypes, and gender attitudes influence one's gender identity, they are not the same as gender identity (Katz 1986; Spence and Sawin 1985).

From a sociological perspective, gender identity involves all the meanings that are applied to oneself on the basis of one's gender identification. In turn, these self-meanings are a source of motivation for gender-related behavior (Burke 1980). A person with a more masculine identity should act

more masculine, that is, engage in behaviors whose meanings are more masculine such as behaving in a more dominant, competitive, and autonomous manner (Ashmore et al. 1986). It is not the behaviors themselves that are important, but the meanings implied by those behaviors.

Beginning at birth, self-meanings regarding one's gender are formed in social situations, stemming from ongoing interaction with significant others such as parents, peers, and educators (Katz 1986). Although individuals draw upon the shared cultural conceptions of what it means to be male or female that are transmitted through institutions such as religion or the educational system, they may come to see themselves as departing from the masculine or feminine cultural model. A person may label herself female, but instead of seeing herself in a stereotypical female manner, such as being expressive, warm, and submissive (Ashmore et al. 1986), she may view herself in a somewhat stereotypically masculine manner, such as being somewhat instrumental, rational, and dominant. The point is that people view themselves along a feminine-masculine continuum, some seeing themselves more feminine, some as more masculine, and some as a mixture of the two. It is this self-perception along the feminine-masculine continuum that is their gender identity, and it is this that guides their behavior.

## THE ROOTS OF FEMININITY/ MASCULINITY

In Western culture, men are stereotypically seen as being aggressive, competitive, and instrumentally oriented while women are seen as being passive, cooperative, and expressive. Early thinking often assumed that this division was based on underlying innate differences in traits, characteristics, and temperaments of males and females. In this older context, measures of femininity/masculinity were often used to diagnose what were understood as problems of basic gender identification, for example, feminine males or masculine females (cf. Terman and Miles 1936).

We now understand that femininity and masculinity are not innate but are based upon social and cultural conditions. Anthropologist Margaret Mead addressed the issue of differences in temperament of males and females in *Sex and Temperament in Three Primitive Societies* (1935). This early

study concluded that there are no necessary differences in traits or temperaments between the sexes. Observed differences in temperament between men and women are not a function of their biological differences. Rather, they result from differences in socialization and the cultural expectations held for each sex.

Mead came to this conclusion because the three societies showed patterns of temperament that varied greatly with our own. Among the Arapesh, both males and females displayed what we would consider a “feminine” temperament (passive, cooperative, and expressive). Among the Mundugamor, both males and females displayed what we would consider a “masculine” temperament (active, competitive, and instrumental). Finally, among the Tchambuli, men and women displayed temperaments that were different from each other, but opposite to our own pattern. In that society, men were emotional and expressive while women were active and instrumental.

Mead’s study caused people to rethink the nature of femininity/masculinity. Different gender-related traits, temperaments, roles, and identities could no longer be inextricably tied to biological sex. Since Mead’s study, the nature–nurture issue has been examined extensively, and with much controversy, but no firm conclusions are yet clear (Maccoby and Jacklin 1974). While there may be small sex differences in temperament at birth (and the evidence on this is not consistent), there is far more variability within each sex group (Spence and Helmreich 1978). Further, the pressures of socialization and learning far outweigh the impact of possible innate sex differences in temperament. We examine this next.

### **THE DEVELOPMENT OF FEMININITY AND MASCULINITY**

There are at least three major theories that explain the development of femininity and masculinity: psychoanalytic theory (Freud 1927), cognitive-developmental theory (Kohlberg 1966), and learning theories that emphasize direct reinforcement (Weitzman 1979) and modeling (Mischel 1966). In all of these theories, a two-part process is involved. In the first part, the child comes to know that she or he is female or male. In the second part, the child comes to know what being female or male means in terms of femininity or masculinity.

According to psychoanalytic theory, one’s gender identity develops through identification with the same-sex parent. This identification emerges out of the conflict inherent in the oedipal stage of psychosexual development. By about age 3, a child develops a strong sexual attachment to the opposite-sex parent. Simultaneously, negative feelings emerge for the same-sex parent that are rooted in resentment and jealousy. By age 6, the child resolves the psychic conflict by relinquishing desires for the opposite-sex parent and identifying with the same-sex parent. Thus, boys come to learn masculinity from their fathers and girls to learn femininity from their mothers.

A later formulation of psychoanalytic theory suggests that mothers play an important role in gender-identity development. According to Chodorow (1978), mothers are more likely to relate to their sons as different and separate because they are not of the same sex. At the same time, they experience a sense of oneness and continuity with their daughters because they are of the same sex. As a consequence, mothers will bond with their daughters, thereby fostering femininity in girls. Simultaneously, mothers distance themselves from their sons, who respond by shifting their attention away from their mother and toward their father. Through identification with their father, boys learn masculinity.

Cognitive-developmental theory is another psychological theory of gender-identity development (Kohlberg 1966). Like psychoanalytic theory, it suggests that certain critical events have a lasting effect on gender-identity development, but these events are seen as cognitive rather than psychosexual in origin. Unlike psychoanalytic theory and learning theory (which is discussed next), cognitive-developmental theory sees the development of a gender identity as preceding rather than following from identification with the same-sex parent. Once a child’s gender identity becomes established, the self is then motivated to display gender-congruent attitudes and behaviors, well before same-sex modeling takes hold. Same-sex modeling simply moves the process along.

Kohlberg identifies two crucial stages of gender identity development: (1) acquiring a fixed gender identity and (2) establishing gender-identity constancy. The first stage begins with the child’s identification as male or female when hearing the

labels “boy” or “girl” applied to the self. By about age 3, the child can apply the appropriate gender label to the self. This is when gender identity becomes fixed. By about age 4, these gender labels are appropriately applied to others. Within another year or two, the child reaches the second critical phase, that of gender constancy. This is the child’s recognition that his or her gender will not change despite changes in age or outward appearance.

The most social of the theories of gender-identity development are the learning theories. In these theories, it is the social environment of the child, such as parents and teachers, that shapes the child’s gender identity. Here, parents and teachers instruct the child on femininity and masculinity, either directly through rewards and punishments or indirectly through acting as models to be imitated. Direct rewards or punishments are often given for outward appearance, such as what to wear (girls in dresses and boys in pants); object choice, such as toy preferences (dolls for girl and trucks for boys); and behavior (passivity and dependence in girls and aggressiveness and independence in boys). Through rewards and punishments, children learn appropriate appearance and behavior. Indirect learning of one’s gender identity emerges from modeling same-sex parents, teachers, peers, or models in the media. Children imitate rewarded models’ thoughts, feelings, or behavior because they anticipate that they will receive the same rewards that the models received.

#### MEASURING MASCULINITY AND FEMININITY: A PSYCHOLOGICAL VIEW

Conceptualizing masculinity and femininity and measuring these orientations in men and women originated in the work of Lewis Terman and Catherine Cox Miles (1936), who created a 455-item test to detect masculinity and femininity. They called it the Attitude Interest Analysis Test (AIAT) to conceal its purpose from subjects. The test included such things as word associations, inkblot associations, interest items, and introversion–extroversion items. For example, on the interest items, persons got femininity points for liking (and masculinity points for disliking) “nursing,” “babies,” and “charades.” Individuals received masculinity points for liking (and femininity points for disliking) “people with loud voices” and “hunting.” On the introversion–extroversion items, persons got

femininity points for agreeing (and masculinity points for disagreeing) that they “always prefer someone else to take the lead” and that they are “often afraid of the dark.” And they got masculinity points for agreeing (and femininity points for denying) that “as children [they were] extremely disobedient” and that they can “stand as much pain as others.” The responses did discriminate between the sexes, with men displaying higher masculinity and women displaying higher femininity.

Terman and Miles’s masculinity–femininity (M–F) scale became a model for M–F scales for more than three decades (see Morawski 1985 for a review). The M–F scales that followed shared four assumptions with the scale created by Terman and Miles. These included the assumptions that masculinity and femininity were: (1) deep-seated, enduring characteristics of people, (2) not readily apparent in overt behavior, (3) linked to mental health (an incongruence in sex and masculinity and femininity signaled problems in psychological adjustment), and (4) opposite ends of a continuum (Morawski 1987).

By the 1970s, researchers had become disenchanted with M–F scales, a timing that coincided with the reemergence of the women’s movement. Three criticisms had developed: (1) The early M–F scales fostered research that exaggerated the differences between men and women, (2) the feminine characteristics in M–F scales often carried negative connotations, and (3) the bipolar conception of masculinity–femininity was seen as problematic, that is, one could be masculine or feminine but not both (Morawski 1987). From the third criticism arose the concept of *androgyny* and scales to assess it (see Morawski 1987 for a review of the problems surrounding the concept of androgyny).

Androgyny is a combination or balance of masculinity and femininity, allowing for the possibility that individuals can express both. Instead of conceptualizing masculinity and femininity as opposite ends of a continuum, where masculinity on one end precludes one from being feminine on the other end, in androgyny, masculinity and femininity are separate dimensions that can be combined. People can be masculine, feminine, or both (androgynous). Two of the more famous inventories that emerged from the impetus to measure masculinity and femininity on separate, independent dimensions were in psychology: the Bem Sex

Role Inventory (BSRI) (Bem 1974) and the Personal Attributes Questionnaire (PAQ) (Spence and Helmreich 1978).

Both the BSRI and PAQ listed attributes that are positively valued for both sexes but are more normative for either males or females to endorse. These are known as the masculine scale and feminine scale, respectively. In the BSRI, respondents indicate the degree to which a series of descriptions are true about them. Examples of descriptions for the masculine scale include “acts as a leader,” “makes decisions easily,” and “willing to take risks.” Examples of descriptions for the feminine scale include being “affectionate,” “gentle,” and “sensitive to the needs of others.”

In the PAQ, respondents rate themselves on a series of bipolar items. For the masculine scale, the items range from masculine to not masculine, while the items for the feminine scale range from feminine to not feminine. Examples of items from the masculine scale include “very independent” (vs. “not at all independent”) and “can make decisions easily” (vs. “has difficulty making decisions”). The feminine scale includes bipolar items such as “very emotional” (vs. “not at all emotional”) and “very helpful to others” (vs. “not at all helpful to others”). In addition to the masculinity and femininity scales, the PAQ has a third scale, labeled masculinity–femininity, that is in keeping with the bipolar M–F measurement tradition. The bipolar items for this scale are culturally appropriate for males on one end, and culturally appropriate for females on the other. Typical items include “very dominant” (vs. “very submissive”) and “feelings not easily hurt” (vs. “feelings easily hurt”).

With separate measures of masculinity and femininity, it is possible to ask about the relationship between the measure of masculinity and the measure of femininity. When this relationship was examined, it was found that the two scales were not strongly negatively related, as would be expected if masculinity were the opposite of femininity. Instead, using either the BSRI or the PAQ, the two ratings were relatively unrelated; a person’s score on one scale did not predict that person’s score on the other scale (Bem 1974; Spence and Helmreich 1978). People had all combinations of scores. Initially, people who combined high scores on masculinity with high scores on femininity were labeled

as androgynous (Bem 1977; Spence and Helmreich 1978). Later, androgyny was indicated by a small difference between masculinity and femininity scores, representing balanced levels of these two characteristics. The other classifications were masculine (high M and low F scores), feminine (high F and low M scores), and undifferentiated (low F and low M scores).

The BSRI and the PAQ are embedded in very different theories about how gender-related characteristics are organized. For Bem (1981, 1993), scores on the BSRI measure not only the different dimensions of masculinity and femininity; more importantly, the scores measure an underlying unidimensional construct known as *gender schematization*. Gender schematization is an internalized tendency to see the world in gendered terms. One who is gender-schematic classifies stimuli as male or female rather than other according to other dimensions that could equally be used. Those who score high on masculinity or high on femininity are gender-schematic because they tend to organize information along gender lines. Androgynous people are gender-aschematic.

Spence (1984, 1993), on the other hand, suggests that gender phenomena are multifactorial. In this view, there are numerous attributes, attitudes, and behaviors that culturally distinguish between men and women, but these are not bound together as a single underlying property such as gender schematization. For Spence (1985; Spence and Sawin 1985) the important underlying construct is *gender identity*, or one’s sense of being masculine or feminine. Culturally defined personality traits, physical attributes, abilities, and occupational preferences, among other things, all contribute to one’s gender identity in unique and individualized combinations. Individuals draw upon these gender characteristics and choose those qualities that are compatible for them as they define themselves as masculine or feminine and ignore other gender qualities. Thus, while societal members may agree on the representation of masculinity and femininity, one’s own masculinity and femininity tends to be more variable and idiosyncratic in nature.

Rather than conceptualizing the items on the PAQ and the BSRI as referring to the broad categories of masculinity and femininity, Spence

maintains that these items tap into socially desirable instrumental and expressive traits in men and women, respectively (Spence and Helmreich 1978, 1980). While these traits are related to masculinity and femininity, they do not define one's overall gender identity. They are simply one of the set of contributors to one's gender-based self-image. This is supported by the fact that scores on the PAQ and BSRI are not strongly related to scores on other measures of gender attitudes, attributes, and behaviors (Spence 1993; Spence and Sawin 1985).

### MEASURING MASCULINITY AND FEMININITY: A SOCIOLOGICAL VIEW

In sociology, the symbolic interactionist view of masculinity-femininity (Burke 1989; Burke et al. 1988; Burke and Tully 1977) shares much in common with the view held by the psychologist Spence and her colleagues. For symbolic interactionists, gender identity is understood in the context of a body of research known as identity theory (Stryker 1980). According to identity theory, the self is an organized collection of hierarchically arranged identities (self-meanings) that serve as a source of motivation for our behavior (Burke 1980). Identities are organized as control systems that act to maintain congruency between the internalized self-meanings (one's identity standard) and perceptions of the meaning of the self in ongoing social situations (Burke 1991). The key in this is one's self-meanings (Osgood et al. 1957).

Individuals' gender identity as masculine or feminine is based on the meanings they have internalized from their association with the role of male or female, respectively, in society. Since these are self-meanings, they cannot be directly observed; they must be inferred from behaviors and expressions in which people engage. Gender identity is one of many role identities people hold. In sociology, we assume that roles do not stand in isolation but presuppose and are related to counterroles (Lindesmith and Strauss 1956). For example, the role of teacher takes on meaning in connection with the role of student, the role of mother takes on meaning in relation to the role of child, and so on. The same is true of identities.

Just as the meaning of student (the student identity) is understood in relation to that of teacher (the teacher identity), so too is the meaning of

male (masculinity) relative to that of female (femininity). The meanings of masculine and feminine are necessarily contrastive. To be male (masculine) is to be not female (feminine) and vice versa (Storms 1979). Gender meanings thus relate to one another as opposite ends of a single continuum, returning to the bipolar conceptualization of masculinity and femininity. Indeed, masculinity and femininity are negatively related when individuals are asked to judge themselves based on the self-descriptors "masculine" and "feminine" (Spence 1993). Interestingly, young children initially do not see masculine and feminine characteristics as opposites, but as they get older, their views of the genders become increasingly bipolar (Biernat 1991). This contrasting of masculinity and femininity in self-meanings does not necessarily hold for behaviors, since one can engage in both masculine and feminine behaviors.

The procedure that symbolic interactionists use to measure gender identity is based on the method devised by Burke and Tully to measure self-meanings (1977). In an analysis of middle school children's gender identities, Burke and Tully first collected sets of adjectives from children that the children themselves used to describe the images of boys and girls. These adjectives, together with their opposites, were used as adjective pairs to form a semantic differential scale to measure the meanings of the male and female roles (Osgood et al. 1957). The stem for the semantic differential was "Usually [boys, girls] are . . . "Then, through a statistical procedure known as discriminant function analysis, the researchers selected those items that best discriminated between the meanings of boys and girls. Examples of items that best distinguished "girl" meanings from "boy" meanings for these children included "soft" (vs. "hard"), "weak" (vs. "strong"), and "emotional" (vs. "not emotional"). After selecting the most discriminating items, children's self-descriptions ("As a [boy, girl] I usually am . . .") were then summed to form a scale of gender identity.

A gender identity scale constructed along the lines described above has certain properties. First, the scale evolves out of the meanings of maleness in relation to femaleness that actually are held in the population from which the sample is drawn. This procedure contrasts with much research that

uses attributes that are assumed to carry meanings of masculinity and femininity with no attempt to check whether they have these meanings for respondents. Second, the measure outlined above incorporates the assumption that meaning is contrastive. The meaning of female is in contrast to the meaning of male and vice versa. Third, by focusing upon self-meanings, we separate issues of who one is (gender identity) from what one does (gender roles) or what one believes (gender attitudes and stereotypes). From this perspective, androgyny may be thought of not as combining masculine and feminine meanings, but as being flexible in the kinds of behaviors in which one engages (sometimes more masculine in meaning, sometimes more feminine in meaning). We now review some of the important research on gender identity that has emerged out of the Burke-Tully method, for which the symbolic interactionist perspective has served as the backdrop.

#### RESEARCH ON FEMININITY AND MASCULINITY

The symbolic interactionist perspective suggests that the self is defined through interactions with others. Burke and Tully's (1977) work found that children with cross-sex identities (boys who thought of themselves in ways similar to the way most girls thought of themselves and vice versa) were more likely than children with gender-appropriate identities: (1) to have engaged in gender-inappropriate behavior, (2) to have been warned about engaging in gender-inappropriate behavior, and (3) to have been called names like "tomboy," "sissy," or "homo." Not surprisingly, boys and girls with cross-sex gender identities were more likely to have low self-esteem.

Another symbolic interactionist tenet is that people will choose behaviors that are similar in meaning to the meanings of their identities (Burke and Reitzes 1981). Burke (1989) found that among middle school children, boys and girls with a more feminine gender identity earned higher grades than those with a more masculine gender identity. This was true independent of the child's sex, race, or grade, the subject area, or the sex of the teacher. Since the early years of schooling are more likely to be "feminized" because there are more female

than male teachers (Lipman-Blumen 1984), children with a more feminine identity will likely perform better in a "feminine" institution. Among college students, research has shown that males and females with a more feminine gender identity are more likely to inflict and sustain both physical and sexual abuse in dating relationships (Burke et al. 1988). People with more feminine gender identities are likely to be more emotionally expressive and relationship-oriented. Aggression may be used as a last resort to attain a closer relationship.

Within the symbolic interactionist tradition, research demonstrates that the meanings that people attribute to themselves as masculine or feminine (their gender identity) are sometimes more important in predicting how they will behave than is their gender (male or female). For example, early research on conversational behavior reported that males were more likely than females to use more dominant and assertive speech patterns in interaction, such as interrupting (West and Zimmerman 1983) and talking more (Aries 1976). However, a later review of the many empirical studies on interruptions and time spent talking showed that gender had inconsistent effects (James and Clarke 1993; James and Drakich 1993). This inconsistency might be explained through an analysis of gender identity rather than gender. For example, research shows that persons with a more masculine gender identity, irrespective of their gender as male or female, are more likely to use overlaps, interruptions, and challenging statements in a conversation (Drass 1986; Spencer and Drass 1989).

While gender identity may sometimes be more important than gender in determining outcomes, it is also possible for one's gender (male or female) and one's gender identity (masculine or feminine) to each result in different displays of behavior. For example, in an analysis of problem-solving discussions between newly married spouses, females and those with a more masculine gender identity were more likely to express negative, oppositional, and dominating behavior, such as complaining, criticizing, or putting down their spouse (Stets and Burke 1996). While masculinity more than femininity should increase dominating behavior (as discussed above), it was surprising that females engaged in more dominating behavior than males.

It was discovered that this dominating behavior emerged especially from females who were viewed by their spouses as being in a weak, subordinate position in our society. These women were apparently using coercive communication to counteract the (subordinate) view of them and gain some control. The problem is that in compensating for their weaker status by behaving in a dominant fashion, women may unwittingly be reminding men of their weak position.

While one's gender identity is generally stable over time, it sometimes changes in response to different experiences. To examine stability and change in gender identity, Burke and Cast (1997) examined the gender identities of newly married couples over the first three years of marriage, finding that the year-to-year stability in gender identity was moderately high. This means that while the gender identities of the respondents did change over this period of time, they did not change markedly. Looking from month to month or week to week, there was almost no observable change.

Identity theory suggests that identities are most likely to change in the face of persistent changes in the environment. The birth of a first child represents a dramatic and persistent change in the environment that confers femininity on women and masculinity on men. Burke and Cast showed that when a couple had their first child, women's gender identities became more feminine and men's gender identities become more masculine. Social psychological processes may also modify one's gender identity. Burke and Cast also found that the more a spouse took the perspective of the other in the marriage, the more the spouse shifted his or her gender identity in the direction of the other's gender identity. The other's gender identity was thus verified and supported by the spouse, which may act to minimize marital conflict.

### THE FUTURE

We mention below several avenues of possible future work in femininity/masculinity. Many more avenues could be identified, for this is an area rich for continuing investigation, but our space is limited. First, we are only beginning to understand issues of stability and change in one's gender

identity. How might gender identities be modified through participation in societal institutions such as the economy, religion, and politics? For example, to what extent and in what ways might employers socialize employees into particular views of being masculine or feminine in order to maintain a smooth flow of work and profit? Are some people more resistant to this socialization than others? Is this issue tied to how relevant gender identity is to individuals?

Related to the above is a second avenue of research, that is, the salience of gender identity across individuals, groups, even cultures. Salience refers to the probability that a particular identity will be invoked in a situation (Stryker 1980). This will vary by situation, but it also varies across individuals. For some, gender is not very relevant, and for others gender is almost always relevant. This returns us to Bem's notion of gender schematization, or the tendency to see the world in gendered terms. What makes gender identity more or less salient for people, and what are the consequences of that?

Third, we know very little about subcultural, cultural, and cross-cultural differences in the meanings that are attached to femininity and masculinity. Most of what we know concerns Western cultures, yet as Margaret Mead discovered long ago, these patterns are not universal. We need to investigate the variation in the meanings of being masculine and feminine. Such studies may help us understand a society's division of labor, its differential power and status structure, and how its privileges and responsibilities are allocated. To modify the social system may mean first modifying individual beliefs about masculinity and femininity.

### REFERENCES

- Aries, Elizabeth 1976 "Interaction Patterns and Themes of Male, Female, and Mixed Groups." *Small Group Behavior*. 7:7-18.
- Ashmore, Richard D., Frances K. Del Boca, and Arthur J. Wohlrs 1986 "Gender Stereotypes." In Richard D. Ashmore and Frances K. Del Boca, eds., *The Social Psychology of Female-Male Relations: A Critical Analysis of Central Concepts*. New York: Academic Press.
- Bem, Sandra, L. 1974 "The Measurement of Psychological Androgyny." *Journal of Consulting and Clinical Psychology* 42:155-162.



- 1977 “On the Utility of Alternative Procedures for Assessing Psychological Androgyny.” *Journal of Consulting and Clinical Psychology* 45:196–205.
- 1981 “Gender Schema Theory: A Cognitive Account of Sex Typing.” *Psychological Review* 88:354–364.
- 1993 *The Lenses of Gender: Transforming the Debate on Sexual Inequality*. New Haven, Conn.: Yale University Press.
- Biernat, Monica 1991 “Gender Stereotypes and the Relationship Between Masculinity and Femininity: A Developmental Analysis.” *Journal of Personality and Social Psychology* 61:351–365.
- Burke, Peter J. 1980 “The Self: Measurement Implications from a Symbolic Interactionist Perspective.” *Social Psychology Quarterly* 43:18–29.
- 1989 “Gender Identity, Sex, and School Performance.” *Social Psychology Quarterly* 52:159–169.
- 1991 “Identity Processes and Social Stress.” *American Sociological Review* 56:836–849.
- , and Alicia D. Cast 1997 “Stability and Change in the Gender Identities of Newly Married Couples.” *Social Psychology Quarterly* 60:277–290.
- , and Donald C. Reitzes 1981 “The Link Between Identity and Role Performance.” *Social Psychology Quarterly* 44:83–92.
- , Jan E. Stets, and Maureen A. Pirog-Good 1988 “Gender Identity, Self-Esteem, and Physical and Sexual Abuse in Dating Relationships.” *Social Psychology Quarterly* 51:272–285.
- , and Judy Tully 1977 “The Measurement of Role/Identity.” *Social Forces* 55:880–897.
- Chodorow, Nancy 1978 *The Reproduction of Mothering: Psychoanalysis and the Sociology of Gender*. Berkeley: University of California Press.
- Drass, Kriss A. 1986 “The Effect of Gender Identity on Conversation.” *Social Psychology Quarterly* 49:294–301.
- Eagly, Alice H. 1987 *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, N.J.: Lawrence Erlbaum.
- Freud, Sigmund 1927 “Some Psychological Consequences of the Anatomical Distinction Between the Sexes.” *International Journal of Psychoanalysis* 8:133–142.
- Gilligan, Carol 1982 *In a Different Voice: Psychological Theory and Women’s Development*. Cambridge, Mass.: Harvard University Press.
- James, Deborah, and Sandra Clarke 1993 “Women, Men and Interruptions: A Critical Review.” In Deborah Tannen, ed., *Gender and Conversational Interaction*. New York: Oxford University Press.
- , and Janice Drakich 1993 “Understanding Gender Differences in Amount of Talk: A Critical Review of Research.” In Deborah Tannen, ed., *Gender and Conversational Interaction*. New York: Oxford University Press.
- Katz, Phyllis A. 1986 “Gender Identity: Development and Consequences.” In Richard D. Ashmore and Frances K. Del Boca, eds., *The Social Psychology of Female-Male Relations: A Critical Analysis of Central Concepts*. New York: Academic Press.
- Kohlberg, Lawrence 1966 “A Cognitive-Developmental Analysis of Children’s Sex-Role Concepts and Attitudes.” In Eleanor E. Maccoby, ed., *The Development of Sex Differences*. Stanford, Calif.: Stanford University Press.
- Lindesmith, Alfred R., and Anselm L. Strauss 1956 *Social Psychology*. New York: Holt, Rinehart and Winston.
- Lipman-Blumen, Jean 1984 *Gender Roles and Power*. Englewood Cliffs, N.J.: Prentice-Hall.
- Maccoby, Eleanor E., and Carol N. Jacklin 1974 *The Psychology of Sex Differences*. Stanford, Calif.: Stanford University Press.
- Mead, Margaret 1935 *Sex and Temperament in Three Primitive Societies*. New York: Dell.
- Mischel, Walter 1966 “A Social-Learning View of Sex Differences in Behavior.” In Eleanor E. Maccoby, ed., *The Development of Sex Differences*. Stanford, Calif.: Stanford University Press.
- Morawski, J. G. 1985 “The Measurement of Masculinity and Femininity: Engendering Categorical Realities.” In Abigail J. Stewart and M. Brinton Lykes, eds., *Gender and Personality: Current Perspectives on Theory and Research*. Durham, N.C.: Duke University Press.
- Morawski, J. G. 1987 “The Troubled Quest for Masculinity, Femininity, and Androgyny.” In Phillip Shaver and Clyde Hendrick, eds., *Sex and Gender*. Newbury Park, Calif.: Sage.
- Osgood, Charles E., George J. Succi, and Percy H. Tannenbaum 1957 *The Measurement of Meaning*. Urbana: University of Illinois Press.
- Spence, Janet T. 1984 “Masculinity, Femininity, and Gender-Related Traits: A Conceptual Analysis and Critique of Current Research.” In B. A. Maher and W. Maher, eds., *Progress in Experimental Research*. San Diego: Academic Press.
- 1985 “Gender Identity and Implications for Concepts of Masculinity and Femininity.” In T. B.

- Sonderegger, ed., *Nebraska Symposium on Motivation: Psychology and Gender*. Lincoln: University of Nebraska Press.
- 1993 “Gender-Related Traits and Gender Ideology: Evidence for a Multifactorial Theory.” *Journal of Personality and Social Psychology* 64:624–635.
- , and Robert L. Helmreich 1978 *Masculinity and Femininity: Their Psychological Dimensions, Correlates, and Antecedents*. Austin: University of Texas Press.
- , and Robert L. Helmreich 1980 “Masculine Instrumentality and Feminine Expressiveness: Their Relationships with Sex Role Attitudes and Behaviors.” *Psychology of Women Quarterly* 5:147–163.
- , and Linda L. Sawin 1985 “Images of Masculinity and Femininity: A Reconceptualization.” In Virginia E. O’Leary, Rhoda Kesler Unger, and Barbara Strudler Wallston, eds., *Women, Gender, and Social Psychology*. Hillsdale, N.J.: Lawrence Erlbaum.
- Spencer, J. William, and Kriss A. Drass 1989 “The Transformation of Gender into Conversational Advantage: A Symbolic Interactionist Approach.” *The Sociological Quarterly* 30:363–383.
- Stets, Jan E., and Peter J. Burke 1996 “Gender, Control, and Interaction.” *Social Psychology Quarterly* 59:193–220.
- Storms, Michael D. 1979 “Sex Role Identity and its Relationships to Sex Role Attitudes and Sex Role Stereotypes.” *Journal of Personality and Social Psychology* 37:1779–1789.
- Stryker, Sheldon 1980 *Symbolic Interactionism*. Menlo Park, Calif.: Benjamin/Cummings.
- Terman, Lewis M., and Catherine C. Miles 1936 *Sex and Personality*. New York: McGraw-Hill.
- Weitzman, Lenore J. 1979 *Sex Role Socialization*. Palo Alto, Calif.: Mayfield.
- West, Candace, and Don H. Zimmerman 1983 “Small Insults: A Study of Interruptions in Cross-Sex Conversations Between Unacquainted Persons.” In Barrie Thorn, Chris Kramarae, and Nancy Heuley, eds., *Language, Gender, and Society*. Rowley, Mass.: Newbury House.

JAN E. STETS  
PETER J. BURKE

## FERTILITY DETERMINANTS

There have been two periods of intense interest in the determinants of fertility by demographers.

The first period, which encompasses the late nineteenth and early twentieth centuries, was dominated by a concern about differential fertility within Western countries; in this period, leaders of the eugenics movement enlisted the services of demographers to learn how these differentials could be reduced, either by increasing the fertility of some groups or lowering the fertility of others. The second period, encompassing the 1940s to the late 1980s, was dominated by concern about differences in fertility (and thus in population growth rates) between Western countries and those countries known variously as, “less developed,” “developing,” “Third World,” and, most recently, “Southern.” In this period, there was little effort to raise fertility in low-fertility societies but a great deal of effort by an international population movement composed of neo-Malthusians and family planners to lower fertility in high-fertility societies by promoting the use of modern contraception. Again, policy makers turned to demographers to learn about the determinants of fertility. At the end of the twentieth century, fertility appears to have a lower place on the agenda of demographers. In some Western countries where national population growth rates are below replacement, there is increasing concern about fertility that is considered to be too low; groups that have fertility considered to be too high are largely adolescents and, in some countries, immigrants. Many Southern countries and some Western policy makers, however, remain concerned about levels of fertility, particularly in poor countries where fertility declines are just beginning. I begin by describing fertility determinants in pretransition societies (those characterized by an absence of deliberate attempts to limit family size) and then describe fertility transitions in Europe and the Third World.

## DEFINITIONS

Fertility refers to the actual childbearing performance of individuals, couples, groups, or populations. Fertility transitions are best defined not in terms of a change in level (from high to low) but rather in terms of reproductive practices. Pretransition societies are those in which *married* couples do not effectively stop childbearing once

they reach the number of children they desire; correspondingly, the onset of a fertility transition is marked by the adoption of practices to stop childbearing before the woman's physiological capacity to reproduce is exhausted. More precisely, in pretransition societies behavior is not parity-specific; that is, it does not depend on the number of children the couple has already borne (Henry 1961). At some point in time, presumably all societies were pretransition; currently, there are few societies where fertility remains high (most are in sub-Saharan Africa) and even fewer where there are no signs that a fertility transition may have begun (Bongaarts and Watkins 1996).

#### FERTILITY DETERMINANTS IN PRETRANSITION SOCIETIES

Despite the absence of parity-specific control in pretransition societies, fertility varies across individuals, couples, and groups: Observed total fertility rates for a population are as high as twelve (married Hutterites in the 1920s) and as low as four to five (the Kung hunters and gatherers in the 1970s and rural Chinese farmers around 1930). Important advances in understanding the sources of this variation followed a distinction between the "proximate" determinants of fertility and the "true" determinants of fertility, a distinction that owed much to an earlier systematic classification of influences on fertility made by Kingsley Davis and Judith Blake (1956). The proximate determinants are *direct* determinants of fertility, the combination of biological and behavioral channels through which the "true" determinants—the social, economic, psychological, and environmental factors—affect fertility (Bongaarts 1978).

In pretransition societies, the two most important of the proximate determinants of the overall level of fertility are marriage patterns and breast-feeding patterns. The other proximate determinants—fecundability (the monthly probability of conceiving among women who menstruate regularly but do not practice contraception), the use of contraception, the risk of spontaneous intrauterine mortality, and induced abortion—play a lesser role in accounting for variations in fertility (Bongaarts

and Menken 1983). Marriage patterns are important, since in most societies childbearing occurs within marriage. Thus, the proportion of women who are currently married at the ages at which reproduction is physiologically possible has a major influence on fertility. There is a striking difference between the marriage patterns of Western Europe and countries of European settlement, such as the United States, Australia, and New Zealand, and societies in other parts of the world, particularly Asia and Africa. In the former, marriage has been relatively late at least since the fourteenth century, and substantial proportions of women remain lifelong spinsters. In Africa and Asia, the typical pattern has been that marriage for women is early (usually as soon as a woman is able to bear a child she marries), and virtually all women marry. Obviously, the reproductive span in Western countries has been on average shorter than that in Asian and African societies. The peculiar Western European marriage pattern is associated with a nuclear family household ideal. In Western countries, couples typically set up their own households after marriage; in Asia and Africa, in contrast, the ideal has usually been that sons bring wives into their parents' household, while daughters go to live in the parental households of their husbands (Laslett 1972). Since the ability to support one's own household is associated with age, it is not surprising that the age of marriage has been later in Western countries than elsewhere.

Within marriage, the major determinants of variations in fertility across groups in populations in which little or no contraception is practiced is breast-feeding, since nursing inhibits the return of ovulation. In addition, in some societies a period of abstinence from sexual intercourse is prescribed by local custom, often because it is believed that intercourse during this period will harm the mother or the child. Post-partum abstinence taboos have been found to vary from a few weeks to as much as a year; if these are followed, they can account for substantial variations in marital fertility. Little research has been done on the determinants of post-partum abstinence. They are clearly cultural, in the sense that they characterize societies of interacting individuals, but they may also be linked to other factors.

Although differences in marriage patterns and breast-feeding patterns account for much of the observed differences in fertility across groups in pretransition societies, it is unlikely that their variation reflects variation in desired family size, for either the individual or the couple. Whether marriage was early or late, or whether breast-feeding was short or long, seems largely the outcome of other social concerns (Kreager 1982). It is probable that these patterns were determined by community norms or social structures, rather than individual preferences: Communities seem to have differed more in these respects than did individuals within these communities (Watkins 1991).

### FERTILITY TRANSITIONS

It is conventional to distinguish between fertility transitions in Western societies and those in other parts of the world. There are evident differences not only in geography but also in timing, with Western fertility transitions occurring earlier. In addition, fertility transitions in non-Western countries were promoted by an international population movement that played a major role in making modern contraception accessible in the Third World, whereas fertility declines in the West occurred without such efforts by governments or social movements. There have recently been a number of reviews of the determinants of fertility change; taken together, they show little consensus on similarities or differences in other determinants of fertility change (Hirschman 1994; Kirk 1996; Mason 1997), although most include mortality decline, the perception that large numbers of children are increasingly unaffordable, the attitudes and moralities concerning family life, and the costs of birth control (Casterline in press).

The earliest sustained fertility transitions at the national level occurred in France, where fertility decline began around the time of the French Revolution, and in the United States, where fertility control was evident in a number of New England communities by the end of the first quarter of the nineteenth century and widespread among women who married on the eve of the Civil War (David and Sanderson 1987). Other fertility transitions spread throughout Europe between 1870

and 1930 (Coale and Watkins 1986), with similar timing in Australia. These changes began in the core countries of northwest Europe and occurred later in the periphery of Central Europe and the Mediterranean countries. In contrast, most analysts agree that there was little evidence of efforts to deliberately stop childbearing or a decline in marital fertility anywhere in the developing world before 1960, except for Argentina, Uruguay, and Chile (largely populated by settlers from Western Europe).

The observation was nearly universal that industrialization was causing this decline, with those who were in more industrial settings (i.e., cities) having lower fertility than rural populations. Industrialization was thought to produce a rising standard of living, an increasingly complex division of labor, an open class system, a competitive social milieu, and individualism. These changes, most thought, induced a desire for smaller families (Hodgson and Watkins 1997). What most concerned observers of fertility changes was the differential fertility by class and ethnicity; the latter was particularly important in late nineteenth- and early twentieth-century America, a period of massive immigration from Eastern, Central, and Southern Europe (Watkins 1994). Although some early observers saw the declining fertility of the wealthy and the urban as fostering prosperity, by the last quarter of the nineteenth century most emphasized the consequences of differential fertility for the composition of the population. For example, U.S. President Theodore Roosevelt decried the "race suicide" of upper-class women who were avoiding marriage or having small families: "The greatest problem of civilization is to be found in the fact that the well-to-do families tend to die out; there results, in consequence, a tendency to the elimination instead of the survival of the fittest" (1907, p. 550). At the time, eugenicists worried that the "prudent and thoughtful" would be the ones to practice birth control, while knowledge of birth control was unlikely to affect the fertility of the "reckless" lower classes (Hodgson and Watkins 1997, pp. 473–474). Women, according to most commentators, were the instigators of fertility decline, and many linked their turn to

abortion and contraception to their reassessment of the value of motherhood. The problem was seen to be particularly acute among elite women. When initial attempts to persuade elite women to bear more children failed, attention was turned to persuading others to have fewer, and access to contraception was gradually liberalized. In Western countries the concerns of demographers and policy makers with domestic population composition faded with the widespread low fertility of the 1930s, and the eugenics movement was dealt a serious blow by its association with Nazi Germany in the 1930s and 1940s.

In the 1950s, Western attention turned to population growth rates in developing countries, many of which had until recently been colonies of Western countries. Mortality was declining, but until the 1960s, fertility in most developing countries was relatively high and apparently stable, aside from brief fluctuations associated with wars, famines, and other upheavals. This stimulated the formation of an international population movement, an alliance of neo-Malthusians, who emphasized the problems consequent on rapid population growth, and birth-controllers, who emphasized the importance of providing women with the means to control their reproduction. Subsequently, the previous pattern of stable reproduction came to an abrupt halt with the onset of rapid fertility transitions in a majority of countries. Between the early 1960s (1960–1965) and the late 1980s (1985–1990), the total fertility rate of the developing world as a whole declined by an estimated 36 percent—from 6.0 to 3.8 births per woman (United Nations 1995). Declines have been most rapid in Asia and Latin America (–42 and –43 percent respectively), less rapid but still substantial in the Middle East and North Africa (–25 percent), and almost nonexistent in sub-Saharan Africa. These averages conceal wide variations among countries in the timing of the onset of transitions and their subsequent pace. At one end of the spectrum of experience are a few countries (e.g., Hong Kong and Singapore) where a fertility transition started around 1960, followed by swift further reductions to the replacement level. At the other extreme are other countries, mostly in sub-Saharan Africa, that have not yet entered the transition.

These remarkable trends in reproductive behavior have been extensively documented in censuses and surveys, and the empirical record is not in dispute (Bongaarts and Watkins 1996). The causes of these trends, however, are the subject of often-contentious debate. Conventional theories of fertility decline, from the modernization versions dominant in the 1950s and 1960s to neoclassical economic and rational actor versions of recent decades, assume the fundamental importance of socioeconomic change, much as did nineteenth-century theories about industrialization and fertility decline. Socioeconomic development results in shifts in the costs and benefits of children and hence in the demand for them. As desired family size declines, fertility reduction soon follows with the widespread adoption of birth control, especially when governments make contraceptive services available through family-planning programs. While this broad explanation is widely accepted, analysts vigorously debate the precise variables and processes involved in this chain of causation. These disagreements have been stimulating and fruitful, producing a wide variety of increasingly refined and detailed views that have guided empirical investigations.

Although rises in female marriage age have contributed to the decline in fertility, this decline is largely due to the adoption of new behavior in marriage: More precisely, it is due to the adoption of parity-specific control using modern contraceptives. In Europe, fertility declines were accomplished initially by the use of abortion, withdrawal, and/or abstinence by married couples to stop childbearing, and only later by the use of modern contraceptives to space children as well as to limit their number. In the Third World, fertility decline was closely associated with the use of modern contraceptives.

Why did fertility decline? Why did couples start to deliberately limit the number of children they bore? What are the “true” determinants of fertility? While a comprehensive theory of fertility would account for both the shift from high to low fertility and variations in fertility at each stage of the fertility transition, most of the attempts to understand the social, economic, and cultural influences on fertility have focused on attempts to

understand the onset of the fertility transition. Almost anything that distinguishes traditional from modern societies has been considered relevant to the explanation of the fertility decline (Cleland 1985; see also reviews of fertility determinants by Hirschman 1994; Kirk 1996; Mason 1997).

The most influential theories that have guided demographic research into the determinants of fertility over the last four decades have been those that assume the fundamental importance of economic factor. Predominant in the 1960s and 1970s was the theory of the demographic transition (classic statements are Davis 1963; Freedman 1961–1962; Notestein 1953; Thompson 1929). Demographic transition theory is based on the assumption that the means of fertility control used in the early stages of the Western fertility transition were always known. Hence, fertility declines can be attributed to changes in the motivations of individuals or couples, changes thought to be related to “modernization,” especially increasing literacy, urbanization, the shift to paid, nonagricultural labor, and declines in infant and child mortality. Neoclassical economic theory, and in particular the New Home Economics associated with Gary Becker (1991), provides a translation from macro-level structural changes to the micro-level calculus of parents (for a more thorough review, see Pollak and Watkins 1993).

Empirical examinations driven by these theories gave them some support. It is now generally acknowledged that economic factors—often described in terms of the “costs” and “benefits” of children—are important determinants of fertility decline. It is, however, also acknowledged that economic factors do not provide a complete explanation. Currently, interesting research focuses on several additions to classical demographic transition theory and to neoclassical economic approaches.

Much attention has been devoted to evaluating the role of family-planning programs in the fertility decline in the Third World, where it seems that the methods used initially in the West were either not known or considered too costly in personal terms (Knodel et al. 1984). In the 1950s, it became evident that population growth rates in

Third World countries were high because of declining mortality but stable fertility. This was believed to have substantial consequences, ranging from changes in the composition of the world’s population (an increasing proportion of which was projected to come from Third World countries) to effects on Third World countries themselves, including famine, political instability, and the constraints that population growth was expected to place on the ability of these countries to develop economically and to modernize more generally. This led to concerted efforts by international agencies, Western governments, and Third World countries themselves to reduce fertility by making modern contraception desirable and accessible in the Third World (Hodgson and Watkins 1997). There has been considerable debate about the effectiveness of these efforts, with some according them little importance (e.g., Pritchett 1994) and others giving them more weight (e.g., Bongaarts 1997). There was a significant impact on fertility levels in the late 1980s, but whether this program effect operates primarily by affecting the timing of the onset of the transition or by the pace of fertility decline cannot be determined with available data (Bongaarts and Watkins 1996).

There has also been considerable interest in institutional determinants of fertility change. These are typically social institutions (e.g., systems of landholding) but occasionally are emergent properties of the collective behavior of individuals (Smith 1989). Therefore, in understanding the frequent association between education and fertility decline, for example, it may be more relevant to ask what proportion of the community has attended school than to ask whether a particular individual has. Similarly, both class relations and gender relations are aspects of the community rather than the individual, and both are likely to be associated with fertility change.

Another perspective emphasizes ideational change. Ideational changes are sometimes broadly, sometimes more narrowly, defined. Among the former is a shift in ideational systems toward individualism, which offered justification for challenging traditional authorities and practices, including those that concerned reproduction

(Lesthaeghe 1983). In a similar vein, John Caldwell argues that much of the fertility decline in developing countries can be explained in terms of the introduction of images of the egalitarian Western family into the more patriarchal family systems of the developing world. It was not so much that the relative balance of costs and benefits of children changed, but that the moral economy shifted: It came to be seen as inappropriate to derive economic benefit from one's children (Caldwell 1982). Among the narrower ideational changes are re-evaluations of the acceptability of controlling births within marriage (Watkins in press) and changes in the acceptability of modern family planning (Cleland and Wilson 1987). Explanations for fertility declines in terms of ideational change are often linked to a focus on diffusion as an important mechanism of change, where diffusion can be postulated as stemming from a central source such as the media (e.g., Westoff and Rodríguez 1995) and/or from person to person in global, national, and/or local networks of social interaction (Bongaarts and Watkins 1996). It is likely that personal networks influence fertility through social learning and the exercise of social influence (Montgomery and Casterline 1996). Intensive efforts to examine local networks of social interaction and their relation to fertility are currently underway in several countries (Agyeman et al 1995; Behrman et al. 1999; Entwistle et al. 1997).

What will happen to fertility in the future? In Western countries, fertility is close to replacement level, although in some (e.g., Italy, Spain) it is well below. There is less variation in fertility than there was either in pretransition societies or during the transition (Watkins 1991). Most couples desire only a few children (rarely more than two), and most use effective means to achieve their desires. Accordingly, analysts have concentrated on the determinants of fertility in specific subgroups of the population, such as teenagers or ethnic or racial minorities. In doing so, they have drawn on much the same combination of socioeconomic characteristics, institutional factors, and ideational change. For example, the higher fertility of teenagers is usually explained in terms both of their differing socioeconomic characteristics and

of their lesser access to effective contraception, as well as to an unwillingness to use it.

Few expect fertility to rise, and there seems to be a consensus that fertility will either remain around replacement level or decline further (see, e.g., Lesthaeghe and Willems 1999). The predictions of even lower fertility in the West are based on a combination of proximate and true determinants. Since the 1960s, marriage age has risen sharply in most developed countries, as have divorce rates; if these trends continue—and as long as most children continue to be born within marriage—lower fertility will follow. There has also been some increase in the proportion who are unable to bear children, in part because some couples postpone marriage and childbearing so long that they are unable to have the children they want, and in part because it is likely that involuntary sterility associated with sexually transmitted diseases may have increased at least slightly (Menken 1985). But the major predictions of lower fertility in the future emphasize the characteristics of modern societies that make childbearing less rewarding compared to the other opportunities available to women and the continued inroads into the family that individualism is making (Keyfitz 1986; Preston 1986; Schoen et al. 1997).

Similarly, fertility is likely to begin to decline in Third World countries that have not yet begun a fertility transition and to continue to decline in those where this process has begun. The course of fertility decline in countries where fertility is now low suggests that once the process of the fertility transition has started, fertility levels decline monotonically until very low levels are reached. Moreover, there is no turning back: The new reproductive behavior is not abandoned. Thus, it is likely that past differences in fertility both within and across countries will diminish.

(SEE ALSO: *Demographic Transition; Family and Population Policy in Less Developed Nations; Family Planning; Family Policy in Western Societies; Family Size*)

REFERENCES

- Agyeman, D., P. Aglobitse, C. Fayorse, J. Casterline, and M. Montgomery 1995 "Social Networks and Fertility

- Control in Ghana." Paper presented at the annual meeting of the Population Association of America, San Francisco, April 6-8.
- Becker, Gary S. 1991 *Treatise on the Family*. Cambridge, Mass.: Harvard University Press.
- Behrman, Jere R., Peter Kohler, and Susan C. Watkins 1999 "Family Planning Programs and Social Interaction: Exploration of Two Dimensions of Specification." Manuscript Submitted for Publication.
- Bongaarts, John 1978 "A Framework for Analyzing the Proximate Determinants of Fertility." *Population and Development Review* 4(1):105-132.
- 1997 "The Role of Family Planning Programmes in Contemporary Fertility Transitions." In G. W. Jones, R. M. Douglas, J. C. Caldwell, and R. M. D'Souza, eds., *The Continuing Demographic Transition* Oxford: Clarendon Press.
- , and Jane Menken 1983 "The Supply of Children: A Critical Essay." In R. Bulatao and R. Lee, eds., *Determinants of Fertility in Developing Countries*, vol. 1. New York: Academic Press.
- , and Susan C. Watkins 1996 "Social Interactions and Contemporary Fertility Transitions." *Population and Development Review* 22(4):639-682.
- Caldwell, John C. 1982 *Theory of Fertility Decline*. New York: Academic Press.
- Casterline, John B. in press "The Onset and Pace of Fertility Transition." *Population and Development Review*.
- Cleland, John 1985 "Marital Fertility Decline in Developing Countries: Theories and the Evidence." In J. Cleland and J. Hobcraft, eds., *Reproductive Change in Developing Countries: Insights from the World Fertility Survey*. London: Oxford University Press.
- , and Chris Wilson 1987 "Demand Theories of the Fertility Decline: An Iconoclastic View." *Population Studies* 41:5-30.
- Coale, Ansley J., and Susan C. Watkins, eds. 1986 *The Decline of Fertility in Europe*. Princeton, N.J.: Princeton University Press.
- David, Paul A., and Warren C. Sanderson 1987 "The Emergence of a Two-Child Norm Among American Birth-Controllers." *Population and Development Review* 13(1):1-41.
- Davis, Kingsley 1963 "The Theory of Change and Response in Modern Demographic History." *Population Index* 29(4):345-366.
- , and Judith Blake 1956 "Social Structure and Fertility: An Analytic Framework." *Economic Development and Cultural Change* 4(3):211-235.
- Entwisle, Barbara, Ronald R. Rindfuss, Stephen J. Walsh, Tom P. Evans, and Sara R. Curran 1997 "Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice." *Demography* 34(2):171-187.
- Freedman, Ronald 1961-1962 "The Sociology of Human Fertility." *Current Sociology* 10/11(2):35-119.
- Henry, Louis 1961 "Some Data on Natural Fertility." *Eugenics Quarterly* 8(2):81-91.
- Hirschman, Charles 1994 "Why Fertility Changes." *Annual Review of Sociology* 20:203-233.
- Hodgson, Dennis, and Susan Cotts Watkins 1997 "Feminists and Neo-Malthusians: Past and Present Alliances." *Population and Development Review* 23(3):469-523.
- Keyfitz, Nathan 1986 "The Family That Does Not Reproduce Itself." *Population and Development Review* 12(suppl.):139-154.
- Kirk, Dudley 1996 "Demographic Transition Theory." *Population Studies* 50(3):361-387.
- Knodel, John, Havanon Napaporn, and Anthony Pramualratana 1984 "Fertility Transition in Thailand: A Qualitative Analysis." *Population and Development Review* 10(2):297-328.
- Kreager, Philip 1982 "Demography in Situ." *Population and Development Review* 8(2):237-266.
- Laslett, Peter 1972 "Introduction: The History of the Family." In P. Laslett and R. Wall, eds., *Household and Family in Past Time*. Cambridge, U.K.: Cambridge University Press.
- Lesthaeghe, Ron 1983 "A Century of Demographic and Cultural Change in Western Europe." *Population and Development Review* 9(3):411-435.
- , and Paul Willems 1999 "Is Low Fertility a Temporary Phenomenon in the European Union?" *Population and Development Review* 25(2):211-228.
- Mason, Karen O. 1997 "Explaining Fertility Transitions." *Demography* 34:443-454.
- Menken, Jane A. 1985 "Age and Fertility: How Late Can You Wait?" *Demography* 22(4):469-483.
- Montgomery, Mark R., and John B. Casterline 1996 "Social Learning, Social Influence and New Models of Fertility." *Population and Development Review* 22(suppl.):151-175.
- Notestein, Frank 1953 "Economic Problems of Population Change." In *Proceedings of the Eighth International*



*Conference of Agricultural Economists*. London: Oxford University Press.

- Pollak, Robert A., and Susan C. Watkins 1993 "Cultural and Economic Approaches to Fertility: A Proper Marriage or a Mésalliance?" *Population and Development Review* 19(3):467–496.
- Preston, Samuel H. 1986 "Changing Values and Falling Birth Rates." *Population and Development Review* 12(suppl.):176–195.
- Pritchett, Lant H. 1994 "Desired Fertility and the Impact of Population Policies." *Population and Development Review* 20(1):1–56.
- Roosevelt, Theodore 1907 "A letter from President Roosevelt on race suicide." *American Monthly Review of Reviews* 35(5):550–551.
- Schoen, Robert, Young J. Kim, Constance A. Nathanson, Jason Fields, and Nan M. Astone 1997 "Why Do Americans Want Children?" *Population and Development Review* 23(2):333–358.
- Smith, Herbert L. 1989 "Integrating Theory and Research on the Determinants of Fertility." *Demography* 26(2):171–184.
- Thompson, Warren S. 1929 "Population." *American Journal of Sociology* 34:959–975.
- United Nations 1995 *World Population Prospects: The 1994 Revision*. New York: United Nations.
- Watkins, Susan C. 1991 *From Provinces into Nations: Demographic Integration in Western Europe, 1870–1960*. Princeton, N.J.: Princeton University Press.
- 1994 *After Ellis Island: Newcomers and Natives in the 1910 Census*. New York: Russell Sage Foundation.
- in press "Local and Foreign Models of Reproduction in Nyanza Province, Kenya." *Population and Development Review*.
- Westoff, Charles F., and Germán Rodríguez 1995 "The Mass Media and Family Planning in Kenya." *International Family Planning Perspectives* 21(1):26–31, 36.

SUSAN COTTIS WATKINS

## FIELD RESEARCH METHODS

See Case Studies; Ethnography; Ethnomethodology; Sociocultural Anthropology; Qualitative Methods.

## FIELD THEORY

It was perhaps only the youthful optimism of a new science that allowed Kurt Lewin and his colleagues to believe that they had within their grasp the key elements of a "field theory of the social sciences." Social psychology made great strides in 1930s and 1940s. Lewin and Lippitt (1938) seemed to have demonstrated in the laboratory the inherent superiority of democracy over autocracy. Lewin (1948) provided a theoretical framework for resolving social conflicts, and after his death his colleagues quickly shaped his legacy into a social scientific theory (Lewin, 1951). How could they resist? The physicists had just announced developments that shook the foundations of Newtonian physics. It was expected that Einstein, safely ensconced in the Institute for Advanced Studies, would any day announce the "unified field theory" that would once again make the physical world an orderly place. Could they ask less of the social sciences?

By 1968 one of Lewin's former students, Morton Deutsch, would declare field theory—and all other grand theories of social psychology—moribund. A few years later, Nicholas Mullins (1973) would eulogize the entire field of small group research as "the light that failed," a victim of the untimely death of its only real intellectual leader, Kurt Lewin. Mullins's borrowing of the Kipling title is compelling not only because it suggests that small group research promised much and failed to deliver but also because it suggests that field theory extended its reach beyond its grasp. In recent years there has been a revival of interest in field theory, though, for better or worse, much of the youthful optimism has faded.

### LEWIN AND THE ORIGINS OF FIELD THEORY

Lewin's (1935) *A Dynamic Theory of Personality* called for a shift in psychology from the Aristotelian to the Galilean mode of thought. Epitomized by the now-classic formulation "behavior is a function of personality and environment," or  $B=f(PE)$ , the new perspective placed social psychology squarely at the intersection of psychology and sociology.

It required abandoning the hope that social behavior could be explained by reference to personality variables and seeking explanations in the dynamic relationships among actors and situations. In this book, Lewin defined the building blocks of the field theory that was to come: force (a vector directed at a point of application), valence (the push or pull of the force), and conflict (the opposition of roughly equivalent forces). At this point, he clearly had in mind a metric space of social life, the concepts of vector and direction having limited meaning in topological (or nonmetric) space.

The young field of topological mathematics soon freed Lewin from the necessity of defining a metric space of social life. In *Principles of Topological Psychology*, he defined the new nonmetric space: "By this term is meant that we are dealing with mathematical relationships which can be characterized without measurement. No distances are defined in topological space" (Lewin 1936, p. 53). The concept was a failure; his presentation to the mathematicians at MIT made it clear that he had overreached. He had ventured into the murky realm of topology when, in fact, he always intended to return to metric space. Two additional difficulties also appear in this volume. First, Lewin insisted that the new topological psychology deal with the entire life-space of the individual. Much as Simmel (1917) conceived of the individual as lying at the intersection of various "social circles," Lewin saw the individual life-space as made up of the totality of available social relations. For practical purposes, this made the specification of a single life-space almost impossibly complex. If one then tried to understand even a small group of actors by merging life-spaces, the problem became overwhelming. Second, Lewin often seemed to think of life-space in terms of physical space. Thus, locomotion almost literally meant moving from one physical location to another. This confusion of metaphor with reality prevented Lewin from proposing a consistent conceptual space of the sort suggested by Borgatta (1963), Bales (1985), and others.

A collection of Lewin's (1948) more applied American papers, *Resolving Social Conflicts*, appeared the year he died. Lewin translated concrete conflicts into abstract life-space "capsules." They were

designed to illustrate the concept of "range of free movement" as "a topological region encircled by other regions that are inaccessible" (Lewin 1948, p. 5). The researchers who followed him have since used his graphic depiction of life-space as capsules only rarely. In the same volume, Lewin defined two useful characteristics of the boundaries between sectors of the life-space: sharpness (the clarity of boundaries) and rigidity (the ease with which boundaries may shift). *Field Theory in the Social Sciences*, a collection of Lewin's more theoretical writings, appeared in 1951. In these papers, Lewin introduced the most crucial concepts. Conflict is defined as "the overlapping of two force fields," force as "the tendency toward locomotion," and position as "a spatial relation of regions" (1951, pp. 39-40). The example of a conflict between husband and wife illustrates Lewin's concept of "subjective" and "objective" social fields. The subjectively defined life-spaces of two people differ, and so a single interpersonal act may have very different meanings for the two actors. Repeated reality testing is necessary to bring the individuals to a consensually defined "objective" social life-space.

Lewin never really succeeded in developing a predictive theory of group dynamics: "The clarification of the problems of past and future has been much delayed by the fact that the psychological field which exists at a given time contains the views of that individual about his future and past" (Lewin 1951, p. 53). The field is still struggling with this problem, lacking an adequate theory of even state-to-state transition.

In the years following Lewin's death, the focus shifted from the theoretical to the applied. Much of the work done at the University of Michigan's Research Center for Group Dynamics and at the National Training Laboratory's facilities in Bethel, Maine, was only very loosely tied to the concepts of field theory. While Lewin is often credited with founding the field of organizational development (Weisbord 1987), his careful work on the nature of the social field is often ignored in favor of deceptively simple quotes such as "There's nothing so practical as a good theory." While Lewin believed that he could not be sure that he had fully understood a social situation unless he

could change it, he never assumed that the ability to bring about change implied understanding. Only in the past several years has attention returned to the difficult legacy of an incomplete field theory.

### THE CURRENT STATE OF FIELD THEORY

The Society for the Advancement of Field Theory was founded at a Temple University conference in 1984. Stivers and Wheelan (1986) have since published the proceedings of the conference as *The Lewin Legacy*. Later, papers from the two subsequent biennial conferences have been published as *Advances in Field Theory* (Wheelan et al. 1990).

*The Lewin Legacy* includes historical essays and applications of field theory to therapy, education, organizational development, and community psychology. A brief set of papers at the end calls for a revitalization of Lewinian thought, particularly within the tradition of action research. It is clear, however, that the authors are responding more to Lewin's research philosophy than to the theoretical constructs of his field theory.

*Advances in Field Theory* continues to focus on application, but with more explicit reference to theory. Gold's (1990) paper titled "Two Field Theories" addresses the issue of whether field theory is a "real theory" or simply an "approach." It distinguishes between the theory as described above and the approach—or "metatheory"—that has served as the guide for the generations of scholars that followed Lewin. Pointing out that much of social psychology has drifted away from the consideration of social life-space toward the understanding of internal cognitive processes, the author calls for a revitalization of the Galilean mode of thought. The remainder of the volume is once again directed toward the solution of practical problems. The authors deal with families, psychiatry, human development, education, conflict, organizations, and cross-cultural concerns. Virtually all of this work seems to draw on Lewin's "approach," or "metatheory," rather than his "specific field theory."

All of this is not to deny the tremendous impact that Lewin and his students have had on

both academia and the resolution of social problems. What impressed his students most was Lewin's commitment to democracy and fairness. This commitment started in the research community, where he unselfishly gave of his time and intellect to help his students and colleagues expand their understanding of society. It extended to society at large. He and his students worked on behalf of the war effort during World War II and actively combated prejudice at home. As Weisbord (1987) has pointed out, the field of organizational development (OD) probably owes its existence and current shape to Lewin and his students. Organizational development is as much an ideology as a theory of change; its democratic value orientation owes much to the ideas of Kurt Lewin. Citing Marrow's (1969) excellent biography of Lewin, Weisbord draws parallels between the lives of Frederick Taylor and Kurt Lewin. Lewin's "humanization of the Taylor system" can be thought of as a blueprint for achieving the central goal of organizational development: increasing organizational effectiveness through the application of social science knowledge. In the seminal work in the field, *The Human Side of Enterprise*, Douglas McGregor (1960) blends the social science of Kurt Lewin with the humanistic psychology of Abraham Maslow (1954). The origins of participatory management, teambuilding, feedback, process consultation, and third-party intervention lie in both the theoretical and empirical work of Lewin and his colleagues: Ronald and Gordon Lippett, Ralph White, Kurt Back, Kenneth Beene, Dorwin Carwright, Alvin Zander, and others.

### REKINDLING THE LIGHT: ATTEMPTS AT SYNTHESIS

The group process school, with its origins in Parsonian functionalism, has recently moved toward integration with the group dynamics school. Bales (1985), tracing the origins of field theory to Dewey's (1896) "reflex arc," suggested that his three-dimensional conceptual space of social interaction made possible a new field theory. He argued for the universality of the dimensions but stopped short of offering the integration: "The new field theory of social psychology is the needed

framework, I believe, for the long-desired integration of social psychology. But to explore that thesis is a major undertaking, and here we must be content with a tentative case for the major dimensions of the framework" (Bales 1985, p. 17). Bales's colleagues (Hare et al. 1996) have since announced the existence of this "new field theory," but it has yet to be published in full and the academic community seems not to be eager to accept it as a major theoretical advance. Whether or not opinion changes as a result of the forthcoming publication of Bales's *Social Interaction Systems: Theory and Measurement* (1999) remains to be seen, but Bales's past efforts have been seen as much stronger in measurement than in theory. To the extent that Bales's new field theory has succeeded, it has been an operationalization of Lewin's concept of subjective life-space. His current theories of polarization are organized around positive and negative images in the minds of individual participants, not around consensual social reality.

Polley (1989) validated an updated set of dimensions and offered a series of explicit operational definitions for the basic concepts of Lewinian field theory. Figure 1 presents a "field diagram" that illustrates some of the basic principles. Two dimensions of interpersonal behavior (friendly-unfriendly and conventional-unconventional) define the plane of the diagram, and the third (dominant-submissive) is represented by circle size. Larger circles represent more dominant members, while smaller circles represent less dominant members. Members close together in the plane of the field diagram are drawn closer by vectors of positive valence, while distinct subgroups repel one another by vectors of negative valence. Members positioned at right angles to the central conflict between two subgroups tend to serve as mediators if they lie toward the "friendly" side of the space and as scapegoats if they lie toward the "unfriendly" side of the space. Both have the potential to draw opposing subgroups together, reducing the severity of the conflict. These analyses are based on "group average" perceptions and so are an attempt to operationalize Lewin's "objective" or consensual life-space. In 1994, Polley and Eid attempted an integration of the three major fields of

small group theory: Lewin's group dynamics, Bales's group process, and Moreno's sociometry. But again, while the methodology and theory are in use at a number of research centers, there has been no consensus among the academic community that either the synthesis or the operationalization of field theory have succeeded.

Two additional lines of research have drawn heavily on field theory concepts. Both developed observation systems based for the measurement of temporal patterns in groups. McGrath's (1991) research group developed a theory known as time, interaction, and performance (TIP) and an observation system known as TEMPO. Wheelan's (1999) research group is seeking to develop a model of group development based on a synthesis of Lewinian group dynamics and Bion's psychodynamic theory of groups. (For more information on these two lines of research, see Observation Systems.)

#### THE FUTURE OF FIELD THEORY

It now appears that Deutsch's announcement of the death of field theory was essentially correct, though there are a few signs of life, as indicated above. However, neither *Psychological Abstracts* nor *Sociological Abstracts* have listed "field theory" as a research topic in recent years. A search of the years 1985–1999 using *PsyINFO* revealed 177 books and articles under the key phrase "field theory." Of those, 65 dealt not with social field theory but with unrelated concepts that simply happened to share the same name. Of the remaining 112 citations, 31 were historical in nature and 63 used field theory as what Gold (1990) refers to as a "general approach" rather than as a specific theory capable of generating testable hypotheses. The remaining 18 articles use field theory as a specific theory. The applications of specific field theory are varied; the articles summarized below are not inclusive but should serve to give the flavor of the impact of Lewinian field theory.

Houston and associates (1988) and Dube and associates (1991) have applied field theoretic concepts to the study of customer evaluation of quality of service. Their model integrates concepts from marketing with those from Lewinian field theory

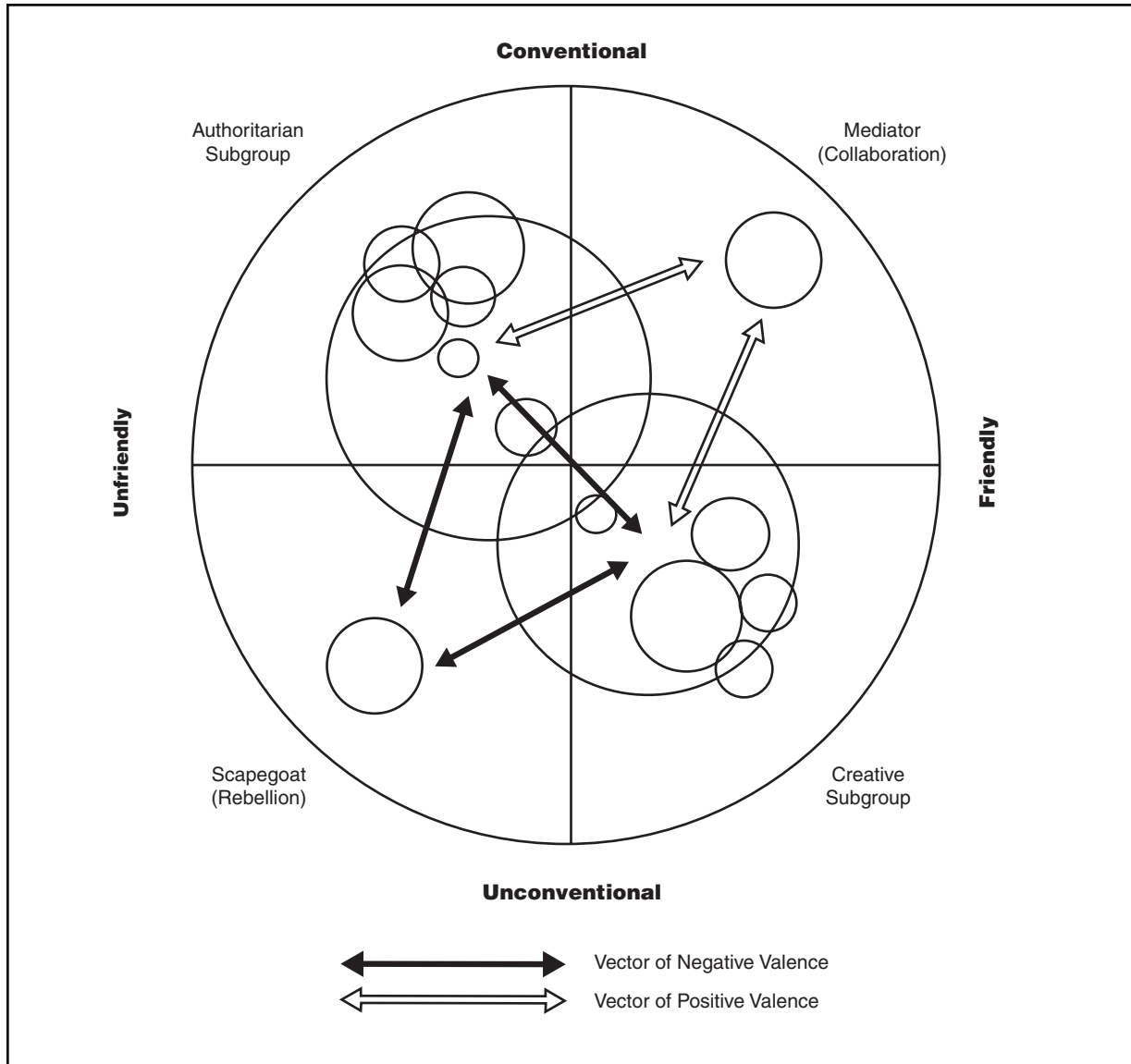


Figure 1. A Three-Dimensional Field Diagram, Showing Two Subgroups, a Mediator, and a Scapegoat.

to develop a deeper understanding of customer satisfaction. Granberg and Holberg (1986) compared behavioristic theory with field theory in attempting to explain voting behavior in Sweden and the United States. They conclude that field theory provides greater explanatory power than behavioristic psychology. In a related series of studies, Dillbeck (1990) and Assimakis and Dillbeck (1995) have employed field theory concepts to explain the effect of a Transcendental Meditation program on social change and perceptions of

quality of life. Viser (1994) pitted field theory against Lazarsfeld's sociological model of voting behavior and contends that field theory provides a better explanation of voter choice. Greenberg (1988) employed field theory concepts of competing force fields to develop a model of employee theft behavior. Diederich (1997) used field theory as the basis for her multiattribute dynamic decision model (MADD), which she uses to predict decision making under time constraints. In the field of communications, Hample (1997) and

Greene (1997) have based theories of “message production” on the Lewinian concepts of life-space and planes of reality and unreality. In the realm of therapy, Barber (1996) has employed field theory to the understanding of the educational community as an agent of change. Dube and Schmitt (1996) tested field theory-based predictions of perceived time judgments. They demonstrated that “unfilled intervals” were perceived to be longer if they occurred during a social process rather than before or after the process. Smith and Smith (1996) identified the concept of the social field as necessary and sufficient for the explanation of social behavior. Finally, Diamond (1992) contrasted field theory and rational choice as explanations for social policy choices and concluded that field theory explains many effects that are dismissed as “irrational” by the rational choice model. This brief tour of empirical research is not intended as comprehensive, but it does suggest that field theory as a specific theory continues to be taken seriously in a variety of disciplines.

There may never be a widely accepted “grand theory” of social psychology. Certainly, the attempts that we have seen so far have not received widespread support from sociologists and psychologists. Still, Kurt Lewin’s legacy is alive and well—as a philosophical orientation in a number of applied fields, such as organizational development, social work, conflict management, and therapy; as a general orientation for researchers in sociology and psychology; and even as a specific theory that continues to generate testable hypotheses. His all-too-brief career has served, and will likely continue to serve, as a guiding light to generations of researchers, practitioners, and social activists.

## REFERENCES

- Assimakis, Panayotis D., and Michael Dillbeck 1995 “Time Series Analysis of Improved Quality of Life in Canada: Social Change, Collective Consciousness, and the TM-Sidhi Program.” *Psychological Reports* 76:1171–1193.
- Bales, Robert F. 1985 “The New Field Theory in Social Psychology.” *International Journal of Small Group Research* 1:1–18.
- 1999 *Social Interaction Systems: Theory and Measurement*. Somerset, N.J.: Transaction Publishers.
- Barber, Paul 1996 “The Therapeutic ‘Educational’ Community as an Agent of Change: Towards a Lewinian Model of Peer Learning.” *Therapeutic Communities: The International Journal for Therapeutic & Supportive Organizations* 12:241–252.
- Borgatta, Edgar 1963 “A New Systematic Interaction Observation System.” *Journal of Psychological Studies* 14:24–44.
- Deutsch, Morton 1968 “Field Theory in Social Psychology.” In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*. Reading, Mass.: Addison-Wesley.
- Dewey, John 1896 “The Reflex Arc Concept in Social Psychology.” *Psychological Review* 3:357–370.
- Diamond, Gregory A. 1992 “Field Theory and Rational Choice: A Lewinian Approach to Modeling Motivation.” *Journal of Social Issues* 48:79–94.
- Diederich, Adele 1997 “Dynamic Stochastic Models for Decision Making Under Time Constraints.” *Journal of Mathematical Psychology* 41:260–274.
- Dillbeck, Michael 1990 “Test of a Field Theory of Consciousness and Social Change.” *Social Indicators Research* 22:399–418.
- Dube, Laurette, and Bernd H. Schmitt 1996 “The Temporal Dimension of Social Episodes: Position Effect in Time Judgments of Unfilled Intervals.” *Journal of Applied Social Psychology* 26:1816–1826.
- Dube, Laurette, Bernd H. Schmitt, and France Leclerc 1991 “Consumers’ Affective Response to Delays at Different Phases of a Service Delivery.” *Journal of Applied Social Psychology* 21:810–820.
- Gold, Martin 1990 “Two Field Theories.” In S. Wheelan, E. A. Pepitone, and V. Apt, eds., *Advances in Field Theory*. Newbury Park, Calif.: Sage Publications.
- Granberg, Donald, and Soeren Holberg 1986 “Prior Behavior, Recalled Behavior, and the Prediction of Subsequent Voting Behavior in Sweden and the U.S.” *Human Relations* 39:135–148.
- Greenberg, Jerald 1998 “The Cognitive Geometry of Employee Theft: Negotiating ‘The Line’ Between Taking and Stealing.” In Ricky Griffing and Anne O’Leary-Kelly, eds., *Dysfunctional Behavior in Organizations*, vol. 23. Stamford, Conn.: JAI Press.
- Greene, John O., ed. 1997 *Message Production: Advances in Communication Theory*. Mahwah, N.J.: Lawrence Erlbaum.
- Hample, Dale 1997 “Framing Message Production Research with Field Theory.” In John O. Greene, ed., *Message Production: Advances in Communication Theory*. Mahwah, N.J.: Lawrence Erlbaum.

- Hare, A. Paul, Sharon E. Hare, and Robert J. Koenigs 1996 "Implicit Personality Theory, Social Desirability, and Reflected Appraisal of Self in the Context of the New Field Theory (SYMLOG)." *Small Group Research* 27:504-531.
- Houston, Mark B. Lance A. Bettencourt, and Sutha Wenger 1998 "The Relationship Between Waiting in a Service Queue and Evaluations of Service Quality: A Field Theory Perspective." *Psychology and Marketing* 15:735-753.
- Lewin, Kurt 1935 *A Dynamic Theory of Personality*. New York: McGraw-Hill.
- 1936 *Principles of Topological Psychology*. New York: McGraw-Hill.
- 1948 *Resolving Social Conflicts*. New York: Harper.
- 1951 *Field Theory in the Social Sciences*. New York: Harper.
- , and Ronald Lippitt 1938 "An Experimental Approach to the Study of Autocracy and Democracy: A Preliminary Note." *Sociometry* 1:292-300.
- Marrow, Alfred 1969 *The Practical Theorist*. New York: Basic Books.
- Maslow, Abraham 1954 *Motivation and Personality*. New York: Harper and Row.
- McGrath, Joseph E. 1991 "Time, Interaction, and Performance (TIP): A Theory of Groups." *Small Group Research* 22:147-174.
- McGregor, Douglas 1960 *The Human Side of Enterprise*. New York: McGraw-Hill.
- Mullins, Nicholas C. 1973 *Theories and Theory Groups in Contemporary American Sociology*. New York: Harper and Row.
- Polley, Richard B. 1989. "Operationalizing Lewinian Field Theory." In E. J. Lawler and B. Markovsky, eds., *Advances in Group Processes: Theory and Method*. Greenwich, Conn.: JAI Press.
- , and J. Eid 1994 "First Among Equals: Leaders, Peers, and Choice." *Journal of Group Psychotherapy, Psychodrama and Sociometry* 47:59-76.
- Simmel, Georg 1955 *Conflict and the Web of Group Affiliation* (K. H. Wolf and R. Bendix, trans.). New York: Free Press.
- Smith, Noel W., and Lance L. Smith 1996 "Field Theory in Science: Its Role as a Necessary and Sufficient Condition in Psychology." *Psychological Record* 46:3-19.
- Stivers, Eugene, and Susan Wheelan 1986 *The Lewin Legacy*. New York: Springer-Verlag.
- Viser, Max 1994 "The Psychology of Voting Action." *Journal of the History of the Behavioral Sciences* 30:43-52.

- Weisbord, Marvin R. 1987 *Productive Workplaces*. San Francisco: Jossey-Bass.
- Wheelan, Susan A., ed. 1999 "Group Development." *Small Group Research* (special issue) 30:3-129.
- , Emmy A. Pepitone, and Vicki Apt, eds. 1990 *Advances in Field Theory*. Newbury Park, Calif.: Sage Publications.

RICHARD B. KETTNER-POLLEY

## FILIAL RESPONSIBILITY

The term *filial responsibility* denotes the "responsibility for parents exercised by children. The term emphasizes duty rather than satisfaction and is usually connected with protection, care, or financial support" (Schorr, 1980, p. 1). Although it is popularly believed that the obligation of children to care for their parents has ancient origins based on widely held moral beliefs, both historical and sociological evidence suggests that neither element of this belief is true (Finch 1989; Schorr 1980).

### LEGAL MANDATES FOR FINANCIAL SUPPORT

One source of the persistent belief in both the historical existence of filial responsibility and its moral derivation stems from an equally persistent belief that there was a time in the past when "the family" had a stronger sense of responsibility for looking after its aging members (Finch 1989). However, a literature on the history of families and demographic trends has emerged that consistently refutes this belief (Laslett 1972). Central to this literature is an understanding of the effects of the demographic transitions, medical advancements, and changes in social and health practices that have resulted in a significant increase in the numbers of persons living into old age. Together these factors have resulted in dramatic changes in the population structure of modern societies wherein the aging segments account for a significantly higher proportion of the population than ever before (Himes 1994). Put simply, until recently only an extremely small segment of the population survived into an extended and dependent old age (Anderson 1980).

Furthermore, due to the nature of the preindustrial economy, those persons who did

survive into old age were not dependent upon the goodwill of or a sense of obligation on the part of their children to assure their financial support. In preindustrial economies both land and businesses, which served as the family's source of sustenance, were owned by the oldest generation and passed down to the younger generation only at death. Hence, it was not the oldest generation that was dependent upon the younger generation, but rather the younger generation that was dependent on the older generation for housing and income. What care may have been provided to parents was, therefore, not necessarily based on a sense of moral obligation but instead stemmed from economic necessity (Haber and Gratton 1994; Schorr 1980).

As long as this economic basis for support existed, there was no need for filial responsibility laws (Schorr 1980). Only with the introduction of an industrial economy that provided a means for children to leave the parental home and obtain an income was there a need for the community to regulate filial responsibility. As Finch (1989, p. 84) notes:

*Much of the historical evidence. . . suggests that under the harsh conditions of poverty which prevailed for most people in the early industrial period, family relationships necessarily were highly instrumental, with support being offered only if there was. . . hope of mutual benefit precisely because anything else would have been an unaffordable luxury (p. 84).*

The shift to an industrial economy changed the financial base of the family and introduced the possibility of independent access to financial resources by children. This shift effectively emphasized the possibility of elders becoming destitute (Bulcroft et al. 1989).

#### **LEGAL MANDATES FOR FINANCIAL SUPPORT**

From their first inception in Roman and English codes, filial responsibility laws were enacted as a means to protect the public purse (and minimize the financial responsibility of the wealthy class) (Jacobson 1995). The aim of such laws was to enable public authorities to recover the costs of

maintaining paupers by making claims on those relatives assumed to be naturally bound by normal obligations of kinship to offer support. The poor laws established for the first time in history that the community would help an indigent parent only after the means of his child had been exhausted (Schorr 1980). These laws were not simply about poverty; they were laws intended to govern the lower class in order to keep down public expenditure. This motivation continues to underlie current laws and practices concerning the care of the elderly in both England and the United States. Despite the moral rhetoric in which filial responsibility laws are often couched, these laws emerged not as a result of moral values but in response to the threat of the growing financial burden that dependent elders posed to an industrial society (Jacobson 1995; Montgomery 1999).

Perhaps this absence of a moral basis for filial responsibility laws accounts for the inconsistency among states in the statement of filial responsibility. Filial responsibility laws, which in 1999 existed in some form in twenty-two states, are variously located in domestic statutes, poor laws, penal codes, and human resource laws. The nature of the fiscal responsibility incorporated into these statutes varies widely and tends to be quite vague. There are variations in the specificity as to which members of the family are responsible and the conditions under which they are responsible. Moreover, these laws are inconsistently enforced (Narayanan 1996), and many questions have been raised about their constitutionality and practicality (Jacobson 1995). For the most part, filial responsibility as incorporated in laws is limited to financial support; and these laws have generally gone unenforced. In fact, several court challenges to Medicaid laws have affirmed the illegality of states' requiring adult children to pay for care provided under Medicaid regulations (Jacobson 1995; Kline 1992).

#### **ATTITUDES TOWARD FILIAL RESPONSIBILITY**

Just as the discrepancy between the existence of filial responsibility laws and the lack of enforcement suggests an equivocal attitude toward children's responsibilities, results of attitude surveys about filial responsibility have been inconsistent.



While there is evidence that both parents and children acknowledge the existence of filial responsibility as a social expectation, individuals' reports of filial responsibility have been found to differ as a function of the type of support (Ganong et al. 1998), geographic distance from parents (Finley et al. 1988), income (Lee, Netzer and Coward 1994), and ethnicity (Lee et al. 1998; Stein et al. 1998). Findings regarding gender differences in attitudes toward responsibility have been mixed (Ganong et al. 1998).

Surveys generally yield respectable percentages of responses favoring filial responsibility as long as the question is limited to ethical and general terms (Schorr 1980). However, a majority of aged persons oppose filial responsibility when the question is reframed to introduce an individual's personal responsibility or to force a choice between the child and governmental/ community sources as the primary source of support (Schorr 1980). Recent studies have revealed a general trend in both Western and Asian cultures toward acceptance of the government or insurance programs as the source most appropriately responsible for financial support of the elderly (Sung 1990). The introduction of Social Security and Supplemental Security Income (SSI) in the United States and the introduction of national health care programs for the elderly in most other developed countries exemplify this trend. In a similar trend, adult children and their parents express a preference for independent living arrangements for the two generations even in Asian countries, where filial piety has long been a central value (Bleiszner and Hamon 1992).

Current practices of filial responsibility in the form of financial support correspond to these expressed preferences. For most families, the flow of financial support between generations is primarily from the oldest generation to the younger generation; and in most developed countries the two generations most commonly reside in separate housing units (Goldscheider and Lawton 1998). When parents and adult children do live together, it is usually due to the needs of the younger generation or the mutual benefits for both generations. The situation of a parent moving into the home of an adult child is relatively rare and is usually associated with a need for personal care rather than a need for housing.

### FILIAL RESPONSIBILITY AS DIRECT PARENT CARE

The assumption by large numbers of adult children of responsibilities for the welfare and direct care of their parents has been the most significant change in family roles and filial responsibility practices in recent history. Although the motives that prompt adult children to care for their frail parents are not well understood, the predominance of adult children among informal caregivers is undisputed when all types and levels of assistance are considered. The prevalence of children as sources of emotional support, assistance with transportation and banking matters, and help with household chores and activities of daily living has been widely documented over the past two decades (e.g., Brody 1985; Merrill 1997).

**Gender Differences.** There is, however, a significant difference between sons and daughters in their caregiving activities and experience. Almost uniformly, studies have shown that greater numbers of daughters than sons assist their parents with a wider range of tasks. Daughters' predominance is especially strong with respect to direct personal assistance to their impaired parents (Merrill 1997). As a rule, daughters are more likely than sons to help elders with household chores, especially food preparation and laundry chores, as well as with personal care tasks that require hands-on care and daily assistance (Montgomery 1992). In contrast, sons tend to concentrate their efforts on tasks that are more circumscribed and sporadic, such as occasional shopping trips and annual yard and house maintenance activities.

The difference in the types of tasks that sons and daughters assume is related to the types of caregiving roles that each group tends to assume. Daughters are not only more likely to engage in caregiving tasks; they are also more likely to assume the role of primary caregiver (Merrill 1997). As such, they are more likely to provide "routine" care over longer periods of time (Matthews and Rosner 1988; Stoller 1990). Sons, in contrast, assume supportive roles that require commitments over shorter periods of time; they tend to be peripheral helpers within a caregiving network rather than the central actors (Matthews 1995).

There is evidence that gender differences in patterns of care persist across cultures and classes. However, these patterns are mediated by the level

of dependency of the elder and the size and gender composition of the family network (Coward and Dwyer 1990; Matthews 1995). Fewer gender differences are observed in patterns of care for parents with low levels of disability. Also, sons from family networks that do not include daughters tend to provide care in a style that is more similar to that of daughters.

Numerous explanations have been advanced to account for the observed divergence of caregiving behaviors between sons and daughters. Most of these explanations center on differences between male and female *roles* and gender differences in power within the family and within society in general (see Finley 1989). However, there is little evidence to support any of the hypothesized explanations, and none of them account for the persistence of divergent parent care activities.

**Class Variations.** Patterns of parent care have also been found to differ within the United States by class and income. Lower-class and working-class caregivers are more likely to live with the elder and to provide direct care. In contrast, higher economic status has been found to be associated with larger financial gifts and help with procuring services (Merrill 1997). This tendency for wealthy families to purchase care for the elderly has also been noted in a wide range of countries at different stages of development throughout the world (Kosberg 1992).

**Cultural Variations.** Cultural differences in informal care have been the focus of a number of studies. However, evidence of cultural differences in parent care within the United States remains equivocal. Greater levels of kin support among African Americans and Hispanics than among whites have been documented in a few studies (Lee et al. 1998; Tennstedt and Chang 1998); but other researchers have reported no differences in informal support (Belgrave and Bradsher 1994) or have documented greater use of informal supports by whites (Silverstein and Waite 1993). These inconsistencies in findings are likely due to differences in sampling designs and the confounding of ethnicity with socioeconomic status.

#### AFFECTION AND OBLIGATION

The persistence of parent care activities, despite the absence of legal or economic imperatives,

has often led researchers and policy makers to focus on affection as the primary motivation for these filial responsibility practices (Jarrett 1985). Numerous studies have noted the relationship between affection and the felt obligation to provide for parent care (Blieszner and Hamon 1992) as well as the importance of attitudes of obligation as correlates of contact with and assistance to parents (Walker et al. 1990). However, there is a growing literature that questions the importance of "affection" as the primary force underlying filial responsibility and/or the performance of caregiving tasks. Repeatedly, it has been shown that there can be emotional closeness between parent and child without contact or aid being given (Walker et al. 1989). At the same time, it has been shown that children who do not feel a great amount of affection for their parents are still able and willing to provide needed assistance (Lee and Sung 1997; Walker et al. 1990). Furthermore, there is growing evidence that caregiving is governed by a cluster of motives that encompass both affection and obligation (Belgrave and Bradsher 1994). For many children, affection may influence the way in which responsibilities are experienced, but these children frequently provide care simply because parents need care (Litwin 1994) or because they perceive few alternatives (Guberman et al. 1992).

#### PARENT CARE AS A CONSEQUENCE SOCIAL POLICY

It has been suggested that the lack of alternatives for parent care is due to existing social policy, which incorporates assumptions about filial responsibility although it does not necessarily explicitly legislate this responsibility (Finch 1989; Montgomery 1999). Specifically, as implemented and practiced, policies and programs concerned with services for the elderly in both developed and developing countries reflect the belief that there is a latent willingness, presumably based upon a residual sense of responsibility that can be activated. The decline in economic liability of children for their parents, which has resulted from the introduction of old-age pensions and health care programs for the elderly, has not been accompanied by a decline in expectations that children will care for and support their parents. In fact, the general trend in long-term care policy, which has emphasized the benefits of community-based care,

has been accompanied by a greater expectation of family care (Finch 1989; Hendricks and Hatch 1993). This expectation for relatives to provide direct care, however, remains more implicit in current long term-care policies than the fiscal liability of relatives, which was explicitly written into the original English poor laws and later adopted by many states (Doty 1995; Kapp 1995). For example, changes in Medicare coverage that are related to the prospective payment system have resulted in shorter hospital stays for elders, who are now being discharged with greater needs for assistance (Estes et al. 1993). The assumption appears to be that family members are, or should be, available to provide such care. Another example of the way in which current policy delegates the direct care responsibilities to family members is found in eligibility guidelines for (SSI), which reduce benefits for beneficiaries who live with other persons or who receive in-kind contributions from them (Hendricks and Hatch 1993). Built into this provision is the expectation that family members will provide the beneficiary with needed care at no cost. Although these regulations assure minimal financial commitment for states that otherwise would be responsible for nursing home care, they differ from fiscal responsibility laws in that the assignment of such responsibility is more diffuse (Doty 1995).

Most often, the delegation of responsibility for direct care to family is implicit rather than explicit. Expectations for direct care are often conveyed directly and indirectly to adult children by case-workers and physicians who serve as gatekeepers for community services (Doty 1995). Even when elders are eligible for assistance through formal providers, there is often pressure placed on family members to perform care (Guberman et al. 1992). When resources are scarce and must be rationed, the living arrangement of the elder as well as the gender and location of relatives become criteria for the distribution of services. The effect of these family characteristics on distribution of home- and community-based services is contingent upon the elder's level of functioning. For elders with low- to mid-level disability, priority for in home care services is given to those who live alone, with the expectation that family members will provide the necessary care for elders who live with them (Greene and Coleman 1995). In contrast, elders with high levels of disability who live with family members

are more likely to receive home care services than those who live alone because discharge planners and case managers view such supports as supplemental but not sufficient for an elder living alone to remain in the community (Doty 1995).

Regardless of the elder's level of functioning, when practitioners convey expectations for care to family members, they are often influenced by cultural and gender norms regarding the division of labor within households (Walker 1983). Cumulative results overwhelmingly show that long-term care, like housework and child care, has been institutionalized as women's work (Keith 1995). Practitioners, who serve as gatekeepers for community resources, tend to offer greater supports to male caregivers and convey lower expectations for sons to provide direct care services (Doty 1995). The impact of these practices has been observed by researchers, who report care-sharing patterns that are associated with gender and note that sons tend to receive more assistance from formal service providers than do daughters (Merrill 1997).

The clear line between long-term care policies and family care patterns provides evidence that care patterns are not simply a result of choices made by family members to adhere to some "inherent" moral order. Rather, these patterns of care emerge as a consequence of laws and practices that limit the alternatives available to family members (Neysmith 1993). Taken together, the evidence regarding filial responsibility suggests trends in laws and practice that reflect a decline in expectations for financial support of the elderly but a simultaneous growth in expectations for children, especially daughters, to support parents in more direct ways. There is also an indication that filial responsibility, as practiced in terms of both financial and direct care, stems largely from necessity and is created by social policies and practices that tend to invoke a questionable or mythical moral basis for such responsibility.

#### REFERENCES

- Anderson, M. 1980 *Approaches to the History of the Western Family 1500-1914*. London, U.K.: Macmillan.
- Belgrave, Linda, and Julia Bradsher 1994 "Health as a Factor in Institutionalization: Disparities between African Americans and Whites." *Research on Aging* 16:115-141.

- Blieszner, R., and R. R. Hamon 1992 "Filial Responsibility: Attitudes, Motivators and Behaviors." In J. W. Dwyer and R. T. Coward, eds., *Gender, Families and Elder Care*. Newbury Park, Calif.: Sage.
- Brody, E. M. 1985 "Parent Care as a Normative Family Stress." *The Gerontologist* 25:19-29.
- Bulcroft, K., J. Van Leynseele, and E.F. Borgatta 1989 "Filial Responsibility Laws: Issues and State Statutes." *Research on Aging* 11:374-393.
- Coward, R. T., and J. W. Dwyer 1990 "The Association of Gender, Sibling Network Composition, and Patterns of Parent Care by Adult Children." *Research on Aging* 12:158-181.
- Doty, P. 1995 "Family Caregiving and Access to Publicly Funded Home Care." In R. A. Kane and J. D. Penrod, eds., *Family Caregiving in an Aging Society*. Thousand Oaks, Calif.: Sage Publications.
- Estes, C. L., J. H Swan, J. H and Associates 1993 *The Long Term Care Crisis*. Newbury Park, Calif.: Sage Publications.
- Finch, J. 1989 *Family Obligations and Social Change*. Cambridge, U.K.: Polity Press and Basil Blackwell.
- Finley, N. J. 1989 "Theories of Family Labor as Applied to Gender Differences in Caregiving for Elderly Parents." *Journal of Marriage and the Family* 51:79-86.
- Finley, N. J., M. D. Roberts, and B. F. Banahan, III 1988 "Motivators and Inhibitors of Attitudes of Filial Obligation Toward Aging Parents." *The Gerontologist* 28:73-78.
- Ganong, L., M. Coleman, K. McDaniel, and T. Killian 1998 "Attitudes Regarding Obligations to Assist an Older Parent or Stepparent Following Later-Life Remarriage." *Journal of Marriage and the Family* 60:595-610.
- Goldscheider, F. K., and L. Lawton 1998 "Family Experiences and the Erosion of Support for Intergenerational Coresidence." *Journal of Marriage and the Family* 60:623-632.
- Greene, V., and P. D. Coleman 1995 "Direct Services for Family Caregivers." In R. A. Kane and J. D. Penrod, eds., *Family Caregiving in an Aging Society: Policy Perspectives*. Thousand Oaks, Calif.: Sage Publications.
- Guberman, N., P. Maheu, and C. Maille 1992 "Women as Family Caregivers: Why Do They Care?" *The Gerontologist* 32:607-617.
- Haber, C., and B. Gratton 1994 *Old Age and the Search for Security*. Bloomington: Indiana University Press.
- Hendricks, J., and L. R. Hatch 1993 "Federal Policy and Family Life of Older Americans." In J. Hendricks and C. J. Rosenthal, eds., *The Remainder of Their Days: Domestic Policy and Older Families in the United States and Canada*. New York: Garland Publishers.
- Himes, Christine 1994 "Parental Caregiving by Adult Women: A Demographic Perspective." *Research on Aging* 16:191-211.
- Jacobson, R. 1995 "Americana Healthcare Center v. Randall: The Renaissance of Filial Responsibility." *South Dakota Law Review* 40:518-545.
- Kapp, M. B. 1995 "Legal and Ethical Issues in Family Caregiving and the Role of Public Policy." In R. A. Kane and J. D. Penrod, eds., *Family Caregiving in an Aging Society*. Thousand Oaks, Calif.: Sage Publications.
- Kline, T. 1992 "The Rational Role for Filial Responsibility Laws in Modern Society." *Family Law Quarterly* 26(3):193-210.
- Keith, C. 1995 "Family Caregiving Systems: Models, Resources and Values." *Journal of Marriage and the Family* 57:179-189.
- Kosberg, Jordan 1992 *Family Care of the Elderly: Social and Cultural Changes*. Thousand Oaks, Calif.: Sage Publications.
- Laslett, P. 1972 "Introduction: The History of Family." In P. Laslett and R. Wall, eds., *Household and Family in the Past Time*. Cambridge, U.K.: Cambridge University Press.
- Lee, Gary R., J. K Netzer, and Raymond T. Coward 1994 "Filial Responsibility Expectations and Patterns of Intergenerational Assistance." *Journal of Marriage and the Family* 56:559-565.
- Lee, Gary R, Chuck Peek, and Raymond T. Coward 1998 "Race Differences in Filial Responsibility Expectations Among Older Parents." *Journal of Marriage and the Family* 60:404-412.
- Lee, Yoon-Ro, and Kyu-Taik Sung 1997 "Cultural Differences in Caregiving Motivations for Demented Parents: Korean Caregivers Versus American Caregivers." *International Journal of Aging and Human Development* 44:115-117.
- Litwin, Howard 1994 "Filial Responsibility and Informal Support Among Family Caregivers of the Elderly in Jerusalem: A Path Analysis." *International Journal of Aging and Human Development* 38:137-151.
- Matthews, S. 1995 "Gender and the Division of Filial Responsibility Between Lone Sisters and Their Brothers." *Journals of Gerontology: Social and Behavioral Sciences* 50B:S312-S320.
- and T. T. Rosner 1988 "Shared Filial Responsibility: The Family as the Primary Caregiver." *Journal of Marriage and the Family* 50:185-195.

- Merrill, D. M. 1997 *Caring for Elderly Parents*. Westport, Conn.: Auburn House.
- Montgomery, R. J. V. 1992 "Gender Difference in Patterns of Child-Parent Caregiving Relationships." In J. Dwyer and R. Coward, eds., *Gender, Families, and Elder Care*. Newbury Park, Calif.: Sage.
- 1999 "The Family Role in the Context of Long-Term Care." *Journal of Aging and Health* 11(3):401–434.
- Narayanan, U. 1996 "The Government's Role in Fostering the Relationship Between Adult Children and Their Elder Parents: From Filial Responsibility Laws to . . . What?, A Cross-Cultural Perspective." *The Elder Law Journal* 4(2):369–406.
- Neysmith, S. M. 1993 "Developing a Home Care System to Meet the Needs of Aging Canadians and Their Families." In J. Hendricks and C. J. Rosenthal, eds., *The Remainder of Their Days: Domestic Policy and Older Families in the United States and Canada*. New York: Garland Publishers.
- Schorr, Alvin L. 1980 *Thy Father and Thy Mother. . . A Second Look at Filial Responsibility and Family Policy*. Washington D.C.: Government Printing Office.
- Silverstein, M., and L. J. Waite 1993 "Are Blacks More Likely Than Whites to Receive and Provide Social Support in Middle and Old Age? Yes, No, and Maybe So." *Journals of Gerontology: Social Sciences* 48:S212–S222.
- Stein, Catherine, Virginia A. Wemmerus, Marcia Ward, Michelle E. Gaines, Andrew L. Freeberg, and C. Jewell Thomas 1998 "Because They're My Parents: An Intergenerational Study of Felt Obligation and Parental Caregiving." *Journal of Marriage and the Family* 60:611–622.
- Stoller, E. P. 1990 "Males as Helpers: The Role of Sons, Relatives, and Friends." *The Gerontologist* 30:228–235.
- Sung, K. 1990 "A New Look at Filial Piety: Ideals and Practices of Family-Centered Parent Care in Korea." *The Gerontologist* 30:610–617.
- Tennstedt, S., and B. Chang 1998 "The Relative Contribution of Ethnicity Versus Socioeconomic Status in Explaining Differences in Disability and Receipt of Informal Care." *Journals of Gerontology: Social Sciences* 53B(2):S61–S70.
- Walker, A. 1983 "Care for Elderly People: A Conflict Between Women and the State." In J. Finch and D. Groves, eds., *A Labour of Love: Women, Work and Caring*. Boston: Routledge and Kegan Paul.
- C. C. Pratt, H.-Y. Shin, and L. L. Jones 1990 "Motives for Parental Caregiving and Relationship Quality." *Family Relations* 39:51–56.

RHONDA J. V. MONTGOMERY

## THE FRENCH SCHOOL OF SOCIOLOGY

The French School of Sociology was formed during the last decade of the nineteenth century and the first quarter of the twentieth century. The nucleus of the school was created by Emile Durkheim (1858–1917), to whose work was joined the crystalizing efforts in the new science that were performed by the team of *L'Année sociologique*, which was founded in 1898. In recent times, scholars have undertaken the examination of the effects of this major contribution to the field by studying the vicissitudes of Durkheim's legacy from the period between the two world wars and onward. We have also concentrated on clarifying the methods that permitted the exploitation and application of this legacy. In this regard, from 1979 to 1982, P. Besnard has fully informed us on the establishment and functioning of assistance strategies in the university and publishing fields. As important as the stakes of power may be, they are much less so than the thematic orientations that Durkheim and his disciples assigned to the science that, in 1838, Comte publicly baptized "sociology." Thus, in this entry, we will first apply ourselves to the task of recalling the path that was forged by the French School of Sociology; we will then examine schematically how it was charted and, finally, discuss the new directions in which sociological research is currently headed.

### PRINCIPAL BRANCHES OF THE FRENCH SCHOOL OF SOCIOLOGY

Problems of periodization are particularly acute in the field of the history of ideas. They are associated with problems related to affiliations, to influences, to rupture and continuity in thought, and to collaboration. Of what, precisely, does "a group," "a school," or "a generation" in a given discipline consist? In his amply documented and unjustly maligned *Manuel de sociologie* (first published in 1950), A. Cuvillier resolved the question a priori: In the history of sociology there exists a "before" and "after" Comte, a series of national schools as well as an ensemble of disciplinary cross-currents. Today, one would more likely delineate a "before" and "after" Durkheim and to regard differently those who are situated upstream of him. At present, it is more precisely posited that "sociological

tradition," formed in the nineteenth century, begins to have meaning with and after the appearance of four Durkheim masterpieces—*De la division du travail social* (1993), *Les règles de la méthode sociologique* (1995), *Le Suicide* (1997), *Les Formes élémentaires de la vie religieuse* (1912)—and that this tradition finds its origins in the works of Montesquieu and Rousseau, who are considered to be “the precursors of sociology.” As a rule, innovators tend to undervalue the works of those who preceded them. As the founder of scientific sociology, Durkheim does not escape this rule. He is even considered an obstacle according to a whole current of opinion, the perpetual ignorance of which has even given rise to the belief in a “blank” between Comte and Durkheim in the development of the discipline (cf. Yamashita 1995).

The concept of a scientifically based sociology was thus imposed. This concept affirms the specificity of social context; it pays close attention to the morphological substratum. While accenting collective tendencies, “forces which are just as real as cosmic forces,” it shows that “social life is essentially made up of representations.” From the introduction of the concept of anomie in the dissertation of 1893 to the analysis of the sacred and of beliefs that was developed in the 1912 work, the principal themes of a general sociology and those of specialized sociologies are freed from prejudices and preconceptions by a rigorously codified approach. Almost entirely, the domains in which they were implemented were “covered” with Durkheimians: judicial sociology by, notably, H. Lévy-Bruhl, economic sociology by F. Simiand (1873–1935), moral and political sociology by P. Fauçonnet and especially C. Bouglé (1870–1940), and religious sociology being particularly well by H. Hubert and M. Mauss (1872–1950). Again, the latter does not limit himself to this one sector, any more than M. Halbwachs (1877–1945) does to social morphology, or L. Lévy-Bruhl (1857–1939) to the writing of a masterpiece, *La morale et la science des mœurs*, published in 1903. Many other specialists—G. Bourgin, L. Gernet, M. Granet, C. Lalo, and so forth—produced works that, combined with those of their master, had a profound influence abroad. It was thus understood that every new development in the discipline had to rest upon the base that Durkheimism had furnished.

The fact that this vigorously formulated concept of sociology provoked the marginalization of

other trends is not at all surprising. Such was the case with the School of Social Reform, founded by F. Le Play (1806–1888), with the dissident school of E. Emolins, and with all the more or less obedient followers of Le Play whose allegiance inspired the family of monographs gathered together in *Les Ouvriers européens* (1855–1879) and *Les Ouvriers des deux mondes* (1857–1885): P. Bureau, E. Cheysson, P. Descamps, P. du Maroussem, P. de Rousiers, and the abbot of Tourville, all of whom were forgotten innovators, but all of whom were recently discovered by Kalaora and Savoya (1989). One will simply note that the collections of *La Réforme sociale* (from 1881 onward), those of *La Science sociale*, directed by Demolins after 1886, and those of *Le Musée social* (which became *Cahiers du Musée social* after 1945) comprise a documentary corpus parallel to that of *L'Année sociologique*. The same may be said for the *Revue internationale de sociologie*, published from 1893 onward by René Worms (1867–1926), who was the director of the International Sociological Library from 1896, the author of an important work, *Philosophie des sciences sociales* (3 vol., 1903–1907), and, most notably, the founder in 1894 of the International Institute of Sociology and in 1895 of the Society of Sociology of Paris.

One may, however, extend to sociology the relation that R. Aron extended for the field of history expanded into historiography: Every society produces the social science that it needs. Durkheim's brand of sociology, strongly tinged with morality, was necessary to the Third Republic, which had been badly shaken by successive crises (Panama, boulangisme, the Dreyfus affair). Scholars have often stressed the concerns that are raised in Durkheim's work about political thought, which is itself centered on social integration and cohesion; these concerns are evidenced by the course he taught in 1902–1903 entitled *L'Éducation Morale*. With regard to these concerns, what could be worth provoking objections raised in the name of old-fashioned individualism, and what interest could this “intermental” psychology present—that is, a collective conceptualized by G. Tarde (1843–1904), author of *Lois de l'imitation* (1890) and of *Logique sociale* (1894)—against which Durkheim polemicizes so effectively and so unjustly? A number of ruptures were deliberately intentional on Durkheim's part; it was against psychology that he intended to build sociology, while

refusing moreover to consider history to be a science and simultaneously reducing economics to nothing more than an incarnation of metaphysics. He was so effective that, whereas in Germany sociology could be developed at the crossroads of history, economics, and psychology, in France all the conditions were right for it to be defined in opposition to these disciplines.

Perhaps this sociological tradition was ordered at a later date. One has only to refer to the *Eléments de sociologie* by C. Bougle and J. Raffault (first edition, 1926; second edition, 1930) to measure the reshaping that occurred between the two world wars. The range of this anthology is wide open, from Bonald to Jaurès, and passing by way of Constant, Tocqueville, and Guizot. Texts by Le Play and Tarde figure alongside those of Durkheim and some of the Durkheimians who, of course, carve out the lion's share. Spencer, Frazer, Jhering, and Mommsen are among the others who are cited, while excerpts of *Principes historiques du droit* by the Russian, P. Vinogradoff, are presented. Yet, more than because of the authors they assemble, these selected pieces interest us because of whom they exclude: Tönnies, Weber, Simmel, Michels, Mosca, Pareto, and Comte, who is only cited a single time! Shortly thereafter, Aron introduced *Sociologie allemande contemporaine* (1935) to French researchers, but for a long while he remained closed to the sociology of Pareto. Other perspectives finally emerged on the eve of World War II: a major contribution to this process was made by Stoetzel, who was the founder of the IFOP in 1938 and of the review *Sondages* in 1939 upon his return from the United States, where he became acquainted with Gallup's works; he would later devote his doctoral thesis to a *Esquisse d'une théorie des opinions* (1943). With Stoetzel was born in France an electoral psychosociology during the electoral ecology inaugurated in 1913 by A. Siegfried's *Tableau politique de la France de l'Quest*. Other important contributors to the broadening of perspectives were G. Bataille, R. Caillois, and M. Leiris, who were reunited within the College of Sociology between 1938 and 1939; as brief as the existence of this institution was, it was the framework for interdisciplinary contributions on myth, the sacred, the imaginary, and the problems of the age (democracy and totalitarianism) that are still striking today for their modernity.

## FRENCH SOCIOLOGIES OF TODAY

Rupture or continuity? For M. Verret, co-author with H. Mendras of a collection of studies entitled *Les Champs de la sociologie française* Mendras and Verret (1988), there is no doubt about the answer: "French contemporary sociology cannot be understood without taking into consideration the great rupture between the two world wars; 1940 was a terrible critical test for French society. . . . There is nothing surprising about the fact that during this disaster, French sociology was also on the rocks. This includes not only Durkheimian sociology which constituted the latest face of it, but all the tradition from which it proceeded." The rereading of the fundamental chapters of *Traité de sociologie générale* (2 vols., 1958–1960), begun by G. Gurvitch (1897–1965) a dozen years after the end of World War II, leads one to nuance this assessment, which links devastation and reconstruction. A number of their authors began their careers before the rupture of 1940: This is the case of the taskmaster himself, whose *L'Idée de droit social* was published in 1932; of G. Friedman, who, beginning in 1936, inaugurated in France research in industrial sociology; and of J. Stoetzel. Others began during the Occupation: for example, A. Girard, working within the framework of the Alexis Carrel Foundation, which was transformed upon the Liberation into the National Institute of Demographic Studies. The documentation upon which they rely owes a great deal to Durkheim and his disciples.

May one say that these authors, senior and junior—F. Bourricaud, J. Cazeneuve, H. Mendras, and others—approach the study of social phenomena in a radically new spirit by breaking the connections established in the twentieth century between knowledge and power over society, by abandoning the research of the great consonance targeted by Durkheimism, and by stripping the discipline of its conquering aspects? It would seem not. A degree of optimism characterized the works of sociologists during the two and a half decades that followed the end of World War II. This is indicated by a flourishing of publications, all marked by a certain voluntarism, like the essays gathered together through the initiative of the French Society of Sociology under the title *Tendances et volontés de la société française* (1966). In a number of areas, people trusted this young science: One expected from its application a decisive improvement in the

government of mankind and in the management of things: various adjustments in the highly industrialized societies and in the developing countries such that they might contribute to the realization of the chosen model of growth.

However, many changes took place behind this permanence that had found its incarnation in a Durkheimian orthodoxy. G. Davy (1883–1973), whose *Eléments de sociologie* (1932) was republished in 1950, beginning in 1931, gathered together “sociologists of yesterday and sociologists of today” who were in the pursuit of the same goals. Those which are easiest to identify are of an institutional nature: the creation, under the auspices of the National Center of Scientific Research, of the Center for Sociological Studies, organizer of important “sociological weeks” that treated various issues such as *Industrialisme et technocratie* (1948), *Villes et campagnes* (1951), and *La Famille contemporaine* (1954); the constitution of numerous research groups (such as the Social Ethnology group run by P.-H. Chombart de Lauwe) and of associations like the one for the Development of the Sociology of Work; and the launching of new reviews, for example, the *Cahiers Internationaux de Sociologie*, founded in 1946 by G. Gurvitch, then, ten years later, the *Archives de Sociologie des Religions*, an organ of the sociology of religions group over which G. Le Bras presided, and two more in 1960. *La Revue Française de Sociologie* by J. Stoetzel, and *Sociologie du travail* by G. Friedman. On the academic front, sociology integrated with the study of philosophy became an entirely separate discipline with the creation of a B.A. in Sociology in 1958; this was the beginning of full academic recognition, which would lead to the institution of an *agrégation* in Social Science in 1976.

To these organizational transformations are added conflicts of paradigms. They are linked to the affirmation of, to use the expression of D. Lerner (1959), the “American concept,” that is, to the diffusion of a concept and a practice of sociology that had come from the far side of the Atlantic. Without oversimplifying, one can, in fact, say that during the period under consideration, the study of social facts was pursued on the one hand within the framework of functionalism, and on the other, on the bases of historical and dialectical materialism. It occurred on a basis of debate, called upon in the 1960s to take a polemical and ideological turn, when one wrongly and sterilely opposed the

descriptive and the explicative, the qualitative and the quantitative, empirical research and theoretical speculation, the exploitation of investigation and of social criticism.

Against the sociological scientism with which one assimilated the new American referent (whether it concern the sociometry of J.-L. Moreno or the functionalism of T. Parsons, the works of R. K. Merton or those of P. Lazarsfeld), the tenets of the dialectic hark back to C. W. Mills, who denounced the invasion of bureaucratic techniques into the social sciences (in *The Sociological Imagination*, 1957), and to P. Sorokin, vigorous critic of the compulsion to enumerate which then reigned, notably in electoral sociology. Imported into France, the German quarrel in the social sciences which, in the beginning of the 1960s, pitted T. Adorno, one of the principal representatives of the Frankfurt School, against K. Popper, the great epistemologist of Vienna, underlined the rift between the protagonists of both camps. Caught up in the turbulence of the late 1960s, sociology then entered a period of crisis (cf. R. Boudon 1971). The gulf widened between the functionalist sociologists, who adhered to systemic analysis, which could be denounced (and not without reason) for its schematism, its optimism, and its conservatism, and their adversaries who countered with a critical rhetoric that often sounded hollow, unsubstantiated as it was by precise data concerning controversial problems such as social change.

The major event of this period would remain, for French sociology, the coming of structuralism, with, in the background, the antagonistic conceptual orientations of G. Gurvitch, who was a proponent of the study of global societies, and of R. Aron, who shared the open perspective of Tocqueville. With the replacement of the Marxist explicative schema by the theorized structural analysis of C. Lévi-Strauss, a mutation of knowledge occurred: A paradigm centered on the idea of conflict was replaced by another that was formed of the stable elements of structure; new modes of conceptualization appeared and with them, new intellectual modes. The history of ideas shows how, in the succession of such modes, a model that has been dominant is shaken, and how it loses or even exhausts the merit that has been attributed to it. The reemergence of structuralism, which one must take care not to confuse with structural analysis, and whose range and limits have been



stressed, permitted the subsistence of four theoretical orientations that P. Ansart (1990) has clearly identified. These have certain affinities with the four schools distinguished by A. Touraine (1988): genetic structuralism, dynamic sociology, the functionalist and strategic approach, and, finally, methodological individualism.

These theoretical models reflect different visions of the world. Thus, P. Bourdieu's analyses lead to a highlighting of the division of society into classes; studies of the frequenting of museums or of the *grandes écoles* reveal practices that differ widely according to origin and to class. Associated with the study of determining structures, which is the goal of genetic structuralism, are the analyses of the author of *La Distinction* (1979) and those of the research team he conducts, who publish their works in *Actes de la recherche en sciences sociales*. The epistemology they display can be qualified as "poststructuralist" because of the introduction of the concept of *habitus* and because of the importance of the developments that support this conceptual proposition. Associated with the study of social dynamism are the names G. Balandier and A. Touraine. The former, beginning with African societies, conceptualized economic and social change; the latter, working from a base of the sociology of work, theorized collective action and social movement. Individual behavior and the problems posed by their aggregation occupies a central place in the work of R. Boudon, notably in *Effets pervers et ordre social* (1977), in *La Logique du social* (1979), and in *La Place du désordre* (1984).

One will note the cleavage that separates the latter two theoretical constructions (M. Crozier, R. Boudon) from the preceding ones. These latter constructions reveal the same structural social model of social determinations: The former prioritize the strategy of the participant and value the individual; nevertheless, their openness to social psychology remains distinct. Thus, M. Crozier reapproaches several interactionist models for treating interaction as a generalizable element, as though interaction revealed the totality of the regulations of a system of action. One is aware, on the other hand, of the close relationship between individualism and social interactionism.

It is not only the radical position that exists between the principles applied by genetic structuralism and those which prevail in methodological

individualism that must here be noted, the "epistemological distance" being perfectly measured by the debate on the inequality of chance. A nodal point of divergence appears, constituted by the conception of this *subject*. What is, in fact, the subject in generic structuralism and in a strategic conception of individual actions? Whether it be a question of the analysis of conflicts and of the representation of social relationships that it implies, a question of symbolic systems which contemporary sociologists are working to rethink, or finally a question of the place, the role and the function of the sociologist in the city, it clearly appears, as P. Ansart (1990) notes, that "the works of R. Boudin mark a critical position with regard to other sociological paradigms . . . , and, like every explicitly critical position, this position clearly marks the divergences, the points of disagreement which can be taken to be, on the epistemological level, insurmountable" (p. 285).

To conclude, these four theoretical orientations prolong and renew tendencies that were charted in the nineteenth century. In a certain manner, P. Bourdieu carries on the ambition of the first sociologists of the nineteenth century, which was picked up by Durkheim: that of constituting a science of social phenomena, with the certitude that social reality is indeed a reality and that it is ponderable by means of strictly codified rules. One can also say that dynamic sociology retains a vision of the world that has its origins in the works of Saint-Simon, Comte, Marx, and Spencer, insofar as these thinkers posed the problem of social change and proposed explicative models for it. One also recognizes without difficulty that in the background of methodological individualism may be distinguished the influence of Tocqueville, Weber, the Austrian marginalist school, and the debates of the Vienna Circle. As for the functionalist and strategic approach, one may consider it to be the inheritor of an administrative science that, with M. Vivien and M. Block, was put into place in the second half of the nineteenth century.

## CONCLUSION

What, all in all, is sociology in France at the end of the twentieth century? The statement made at the beginning of the 1990s by J.-M. Berthelot in *La*

*Construction de la sociologie* (1991) is still valid. The programmatic crossovers that have been identified and the new uncertainties that have been discerned have been both confirmed and emphasized. To these, one adds a transformation of the very objective of sociology by reason of a notable subjectivization of social material. Often denounced as a fetishistic idea, society tends to be thought of as an ensemble of forms and networks of *sociability*, in the same way that *individual*, a simple statistical element passively submitted to diverse sorting operations, tends to be substituted by a complex *individuality*, more or less free of contradictions, the subject of a statement that claims to be “personal” and, in all cases rebellious to classification, recalcitrant toward enumeration and cut off from the “masses”. To weigh the new relationships maintained by the individual and society, one has recourse to sociologies that, unlike those of Durkheim and his followers, remained engaged in the psychological givens: Such is indeed the direction of interest aroused by the recent translations of G. Simmel’s works, in which, in the analysis of phenomena under study, the content is less important than the form and the part more significant than the whole.

One sees to what degree, in this perspective, the rereading of “sociological tradition” is imposed, as well as the reintegration into the discipline’s history of contributions that had more or less been excluded from it, and of reflections on the epistemology of the social sciences. Concurrently being pursued is a *conceptual clarification*, including the admirable *Dictionnaire critique de la sociologie*, continually reedited since its appearance in 1982 by R. Boudon and F. Bourricaud, which is its point of departure and canonical example. At the same time, also being pursued is the analysis of *social problems* related to the evolution of mores (the family, the young, the relations between generations, etc.) and to economic transformations (unemployment, exclusion, social justice, etc.). Finally, parallel to these pursuits is the research being conducted on the improvement of descriptive and explicative models of the changes that society is witnessing today. One points out in this regard the exemplary character of the investigation being conducted by Louis Dirn (anagram for *lundi soir*), that is, the group of sociologists who meet at the beginning of each week in the company of H. Mendras and M. Forsé. Thus, for the most

part freed from past dogmatism, French sociology today asserts its vitality by better adjusting its procedures for examining social objects and phenomena to which it is working to give a sense.

#### Bibliography

- Ansart, P. 1990 *Les Sociologies contemporaines*. Paris: Seuil.
- Berthelot, J.-M. 1991 *La Construction de la sociologie*. Paris: P.U.F.
- Besnard, P. dir. 1979 “The Durkheimians”, *Revue française de sociologie*, January–March, XX-1.
- 1981 “Sociologies françaises au tournant du siècle.” *Revue française de sociologie*, July–September, 1981, XXII-3.
- 1985 “La Sociologie française dans l’entre-deux-guerres.” *Revue française de sociologie*, April–June, 1985, XXVI-2.
- Boudon, R. 1971 *La Crise de sociologie*. Geneva: Droz.
- Hollier, D. 1979 *Le Collège de sociologie*. Paris: Gallimard.
- Kalaora, B., and Savoye, A. 1989 *Les Inventeurs oubliés: Le Play et ses continuateurs*. Paris: Champ Vallon.
- Lerner, D. 1959 “De la tradition européenne au concept américain. *Esprit* (January): pp. 3–17.
- Mendras, H. and Verret, M. 1988 *Les Champs de la sociologie française*. Paris: A. Colin.
- Touraine, A. 1988 “Sociologies et sociologues.” In *L’état des sciences sociales*. Paris: La Découverte.
- Valade, B. 1989 “La Sociologie depuis la Seconde Guerre mondiale.” In the *Encyclopédie philosophique universelle*, vol. 1. Paris: P. U. F.
- with R. Felliéule 1996 *Introduction aux sciences sociales*. Paris: P.U.F.
- Yamashita, M. 1995 La Sociologie française entre Auguste Comte et Emile Durkheim.” in *L’Année sociologique*, vol. 45, 1995, I.

BERNARD VALADE

## FUNCTIONALISM AND STRUCTURALISM

Sociology’s first theoretical orientation was functionalism. In trying to legitimate the new discipline of sociology, Auguste Comte (1830–1842,

1851–1854) revived analogies made by the Greeks and, closer to his time, by Hobbes and Rousseau that society is a kind of organism. In so doing, Comte effectively linked sociology with the prestige of biological science. For functional theory, then, society is like a biological organism that grows, and as a consequence, its parts can be examined with respect to how they operate (or function) to maintain the viability of the body social as it grows and develops. As Comte emphasized (1851–1854, p. 239), there is a “true correspondence between Statical Analysis of the Social Organism in Sociology, and that of the Individual Organism in Biology” (1851–1854, p. 239). Moreover, Comte went so far as to “decompose structure anatomically into *elements, tissues, and organs*” (1851–1854, p. 240) and to “treat the Social Organism as definitely composed of the Families which are the true elements or cells, next the Classes or Castes which are its proper tissues, and lastly, of the cities and Communes which are its real organs” (pp. 211–212). Yet, since these analogies were not systematically pursued by Comte, his main contribution was to give sociology its name and to reintroduce organismic reasoning into the new science of society.

It was Herbert Spencer who used the organismic analogy to create an explicit form of functional analysis. Drawing upon materials from his monumental *The Principles of Biology* (1864–1867), Spencer’s *The Principles of Sociology* (1874–1896) is filled with analogies between organisms and society as well as between ecological processes (variation, competition, and selection) and societal evolution (which he saw as driven by war). Spencer did not see society as an actual organism; rather, he conceptualized “superorganic systems” (organization of organisms) as revealing certain similarities in their “principles of arrangement” to biological organisms (1874–1896, part 2, pp. 451–462). In so doing, he introduced the notion of “functional requisites” or “needs,” thereby creating functionalism. For Spencer, there were three basic requisites of superorganic systems: (1) the need to secure and circulate resources, (2) the need to produce usable substances, and (3) the need to regulate, control, and administer system activities (1874–1896, part 2, p. 477). Thus, any pattern of social organization reveals these three classes of functional requisites, and the goal of sociological analysis is to see how these needs are met in empirical social systems.

Later functionalists produced somewhat different lists of requisites. Émile Durkheim argued that sociological explanations “must seek separately the efficient cause [of a phenomenon]—and the function it fulfills” (1895, p. 96), but, in contrast to Spencer, he posited only one functional requisite: the need for social integration. For Durkheim, then, sociological analysis would involve assessment of the causes of phenomena and their consequences or functions for meeting the needs of social structures for integration.

Were it not for the activities of theoretically oriented anthropologists, functionalism probably would have died with Durkheim, especially since Spencer’s star had faded by World War I (Turner and Turner 1990). As the traditional societies studied by early anthropologists were generally without a written history, anthropologists were confronted with the problem of explaining the existence of activities and structures in these societies. The explanatory problem became particularly acute in the post-World War I period with the demise of evolutionism and diffusionism as deciphering tools (Turner and Maryanski 1979). Functional analysis provided a novel alternative: Analyze structures such as kinship or activities such as rituals in terms of their functions for maintaining the society. It was A. R. Radcliffe-Brown ([1914] 1922, 1924, 1935, 1952) who sustained the Durkheimian tradition by emphasizing the importance of integrative needs and then analyzing how structures—most notably kinship systems—operate to meet such integrative requisites. In contrast, Bronislaw Malinowski (1913, 1944) extended functional analysis in a more Spencerian direction, emphasizing that there are distinct system levels (biological, social, and cultural), each of which reveals its own distinctive requisites. Extending Spencer and anticipating Talcott Parsons, Malinowski (1944) posited four basic requisites at the social system level: (1) production and distribution, (2) social control and regulation, (3) education and socialization, and (4) organization and integration.

Thus functionalism was carried to the midpoint of the twentieth century by anthropological work. Then during the 1930s, a group of Harvard sociologists—led by a graduate student, Robert Merton (1949)—began thinking about functional analysis, especially as it had been carried forth by

Radcliffe-Brown and Malinowski (Turner and Maryanski 1979). As a result, many of Merton's fellow students became the leading figures in the revival of functionalism in sociology. This revival began with Kingsley Davis and Wilbert Moore's classic article on "Some Principles of Stratification" (1945), followed by Davis's basic text, *Human Societies* (1948), and a variety of articles and books by others listing the "functional requisites" of societies (e.g., Levy 1952). But it was Talcott Parsons, a young Harvard instructor during Merton's tenure as a graduate student, who was to become the archfunctionalist of modern times (Parsons 1951; Parsons et al. 1953). For Parsons, the social universe was conceptualized in terms of four distinct types and levels of "action systems" (culture, social, personality, and organismic/behavioral), with each system having to meet the same four functional needs: (1) adaptation (securing and distributing environmental resources), (2) goal attainment (mobilizing resources to goals or ends), (3) integration (coordinating system parts), and (4) latency (managing tensions within parts and generating new parts). The operation and interchanges of structures and processes within and between system levels were then analyzed with respect to these basic requisites.

As functional theorizing became dominant in American theory in the 1950s and 1960s, criticism escalated. Opposition came from several different quarters and took a number of distinctive lines of attack. From interactionist theorizing came criticism about functionalism's failure to conceptualize adequately the nature of actors and the process of interaction (Blumer 1969); from Marxist-inspired theory, which was just emerging from the academic closets in the post-McCarthy era, came attacks on the conservative and static nature of analysis that emphasized the functions of phenomena for maintaining the status quo (e.g., Coser 1956; Dahrendorf 1958; Mills 1959); from theory construction advocates came questions about the utility of excessively classificatory or typological theories that pigeonholed phenomena in terms of their functions (e.g., Merton 1957, pp. 44–61); and from philosophers and logicians came questions about tautology and illegitimate teleology in explanations that saw phenomena as meeting needs and needs as generating phenomena (e.g., Dore 1961). The result was the decline of functional theorizing in the early 1970s.

Functionalism has never fully recovered from these criticisms, although there persist vibrant modes of functional analysis in many disciplines, both within and outside of the social sciences. Within sociology, functionalism adapted to a hostile intellectual environment in several ways. One was to downplay functional requisites in a "neofunctionalism" (Alexander 1985; Alexander and Colomy 1985) and, instead, emphasize the dynamics that had always been at the center of functional analysis: differentiation and social change. As with Parsons, neofunctionalism maintains the analytical distinctions among culture, social system, and personality; and as with Durkheim and Radcliffe-Browne, integration among (1) cultural symbols, (2) differentiating social systems, and (3) persons is given emphasis without, however, the baggage of functional requisites. Another strategy was to maintain Parsons's functional requisites, expanding and elaborating his scheme, but at the same time generating testable propositions (Münch, 1982). In this way, the criticism that Parsons's theory was, in reality, a category system could be mitigated because a system of categories could be used to generate propositions and hypotheses. Still another approach was to downplay the notion of functional requisites, as was done by neofunctionalists, and analyze specific institutional systems in terms of their consequences for overcoming the problem of system complexity generated by differentiation (Luhmann 1982). Yet another effort emphasized cultural processes and the functions of rituals, ideology, and values for integrating social structures (Wuthnow 1987). A final strategy has been to translate the notion of functional requisites into a more ecological perspective, with the notion of requisites being translated into selection pressures in a way that corresponds to the notion of functions in the biological sciences (Turner 1995). That is, function becomes a shorthand way to summarize selection processes as they operated in the past to create a new structure. When this logic is applied to sociocultural processes, function is simply a name given to selection forces; and the analysis shifts from listing requisites to assessing how selection pressures increases or decreases for certain types of sociocultural formations.

Aside from changing under the impact of critiques leveled against it, functionalism has also

spawned a number of other theoretical traditions. The most obvious is structuralism, which will be examined in more detail shortly, but there is another that is equally important in sociology, namely, human ecology. Human ecology is a perspective that comes from Spencer's and Durkheim's sociology. Both Spencer and Durkheim recognized that any social structure or system of symbols exists in a resource environment and that there is often competition for resources among alternative structures and symbols when social density increases. Thus, religious or political ideologies must compete for adherents, or businesses or governmental agencies must compete in their respective resource niche. Ecological reasoning came initially via the Chicago School's interest in how areas of cities were utilized, and like Durkheim and Spencer before them, members of the Chicago School visualized the competition for space created by dynamic real estate markets in Darwinian terms (e.g., Park 1936; Park et al. 1925; Hoyt 1939). Later others (Hannan and Freeman 1977) applied the idea of niche, niche density, competition, and selection to populations of organizations to examine their rates of success and failure. Thus, one of the more visible theoretical perspectives in sociology comes from early functionalists, who not only borrowed an image of society as being like an organism but also developed ideas that paralleled those of Darwin in examining the mechanisms behind sociocultural differentiation (or "social speciation"). For them, differentiation is both the context of niche specialization and the result of previous competitions, as sociocultural forms compete with each other, with those unable to compete in one niche dying, moving to a new niche, or creating a new niche (and, thereby, increasing the level of differentiation).

The other major offspring of functionalism were various lines of structuralism that were inspired primarily by Durkheim's and various collaborators' work. Durkheim (1893, 1895) implicitly borrowed Comte's distinction between statics and dynamics, although Durkheim conceptualized social statics in Montesquieu's terms as social "morphology." Static or morphological analysis was seen to involve an assessment of the "nature," "number," "interrelations," and "arrangement" of parts in a systemic whole (Durkheim 1895, p. 85). For Durkheim, sociological explanation still

sought to discover "cause" and "function," but the basic structural units of sociological analysis—that is, the "things" that are caused and functioning—are to be classified by "the nature and number of the component elements and their mode of combination" (Durkheim 1895, p. 81).

### FRENCH STRUCTURALISM

French structuralism stood Durkheim on his head through Claude Lévi-Strauss's (1945a, [1949] 1969, 1963) adoption of ideas in less sociologically prominent works by Durkheim and his nephew, Marcel Mauss. In Durkheim's "Incest: The Nature and Origin of the Taboo" (1898) as well as his and Mauss's *Primitive Classification* ([1903] 1963), emphasis shifts to the origins and functions of rules of exogamy and to human classification systems and modes of symbolic thought. For Durkheim and Mauss, the way that humans cognitively perceive and classify the world reflects the morphological or material structure of society (nature, number, arrangement, and combination of parts). In developing this argument, which Durkheim repeats in less extreme form in *Elementary Forms of the Religious Life* (1912), Durkheim and Mauss posit all the basic elements of Lévi-Strauss's (1963) structuralism. First, although societies differ in their evolutionary development, they are all fundamentally similar because they are based upon the same "underlying principles" (Durkheim and Mauss [1903] 1963, p. 74), and these principles provide individuals with a basis for classifying and constructing their universe. Second, "mythology" is a universal method of classification. Third, such classification systems represent "relations of things" to each other and "are thus intended, above all, to connect ideas, to unify knowledge" (p. 81). Fourth, classification systems are, in essence, created by oppositions in the material social world—sacred—profane, pure—impure, friends—enemies, favorable—unfavorable.

This last idea, which Lévi-Strauss was to conceptualize as "binary oppositions," is carried further by Mauss (in collaboration with Henri Beuchat) in *Seasonal Variations of the Eskimo: A Study in Social Morphology* ([1904–05] 1979), where the sharp dualisms of Eskimo life are outlined as they are created by the seasonal nature of Eskimo activities. Later, in *The Gift* ([1925] 1941), Mauss reasserts Durkheim's earlier conclusion that activity, such as

the famous Kwakiutl potlatch, is a surface exchange reflecting a “deeper” and “more complex” underlying structure in which gifts symbolize and assure the continuation of social relations among diverse groups.

Lévi-Strauss regarded *The Gift* as the inauguration of a new era in the social sciences and saw Mauss as a new Moses “conducting his people all the way to a promised land whose splendour he would never behold” (Lévi-Strauss [1950] 1987, pp. 41–45). Following directly in Mauss’s footsteps, Lévi-Strauss ennobled the concept of exchange as a “total social phenomena” by arguing in his first major work, *The Elementary Structures of Kinship*, that a “principle of reciprocity” is the most general and universal property of society, with the exchange of women as the most fundamental expression of this principle (Lévi Strauss [1949] 1969, pp. 60–62).

Yet it is doubtful that these ideas alone would have produced structuralism. Two other scholars are critical to Lévi-Strauss’s transmutation of Durkheim and Mauss. One is Robert Hertz, a young member of Durkheim’s “Année School.” Before his death in World War I, Hertz produced a number of essays, the two best-known ones being published in *Death and the Right Hand* ([1909] 1960). In this work, Hertz continues the Durkheim–Mauss theme of the duality in the structure of society and documents how this is reflected in ideas about society (myths, classifications, and other “representations”), but the imagery is much more like modern structuralism in that the goal of inquiry is (1) to show the meaning of observed facts in their interrelations and (2) to uncover the underlying structural principles beneath the surface of such observed phenomena.

The final and perhaps most critical influence on Lévi-Strauss’s reversal of Durkheim is the Swiss linguist Ferdinand de Saussure ([1915] 1966), whose lectures, posthumously published under the title *Course in General Linguistics*, serve as the pioneering and authoritative work for modern-day structural linguistics. Saussure appears to have influenced Lévi-Strauss indirectly through Nikolai Trubetzkoi ([1949] 1964, 1968) and Roman Jakobson (1962, 1971), but they simply extend Saussure’s key insight that language is a system whose units—whether sounds or morphemes—are only points in an overall structure. Moreover, while speech

(*parole*) can be directly observed, it reflects an underlying system or language structure (*langue*), and, thus, it is critical to use *parole* to discover the *langue*.

In Lévi-Strauss’s hands, this combined legacy was to produce French structuralism. At times, Lévi-Strauss is critical of Durkheim, but he concludes in the end that Durkheim’s work “is an inspiration” (Lévi-Strauss 1945b, p. 522) and that “the entire purpose of the French school lies in an attempt to break up the categories of the layman, and to group the data into a deeper, sounder classification” (Lévi-Strauss 1945b, pp. 524–525). In early work such as *The Elementary Structures of Kinship* (Lévi-Strauss [1949] 1969), there is a clear debt to Durkheim and Mauss as Lévi-Strauss seeks to uncover the underlying principles (such as “reciprocity”) of bridal exchanges, but he soon turns to the methods of structural linguistics and sees surface social structures (Durkheim’s “morphology”) as reflecting a more fundamental reality, lodged in the human unconscious and the biochemistry of the human brain and involving “rules and principles” to organize “binary oppositions” that are used to create empirically observable patterns of social reality (Lévi-Strauss 1953). But these empirical patterns (e.g., myths, beliefs, kinship structures) are not the “really real” structure; they are like *parole* is to *langue*, a reflection of a deeper structural reality of which existing phenomena are only the surface realizations.

Thus, by the middle of the twentieth century, Lévi-Strauss (1953) had created a structuralism concerned with understanding cultural and social patterns as reflections of universal mental processes rooted in the biochemistry of the human brain. His basic argument can be summarized as follows: (1) The empirically observable must be viewed as a system of relations, whether these be a kinship system, elements of myth or folk tales, or any sociocultural pattern. (2) Statistical models of these observable systems need to be developed in order to summarize the empirically observable relations among components. (3) These models are, however, only surface manifestations of more fundamental reality and must, therefore, be used to construct mechanical models in which the organization of binary oppositions is organized by basic rules lodged presumably in the human brain. (4) These mechanical models are the more “real” since they

get at the underlying processes by which sociocultural systems are generated by the biology and neurology of the brain.

Marshall Sahlins (1966) once sarcastically remarked of Lévi-Strauss's scheme that "what is apparent is false and what is hidden from perception and contradicts it is true." But Levi-Strauss was not alone in this kind of neurological reasoning. Within linguistics, Noam Chomsky (1980) was making a similar argument, asserting that the brain is wired to produce generative grammars that are less the result of the sociocultural environment of humans than of the structure of their brains. We might term these kinds of structuralisms "biostructuralism" because they all argue that the social world as it is perceived and structured is being systematically generated by rules of organization that are part of the neurobiology of the human brain.

Not all structuralists went this far. Most were content to argue that symbol systems and social structures are organized by rules that need to be discovered. Since the rules are not obvious when examining symbols and text, the task of structural analysis is to uncover the underlying rules, programs, principles, and other generative forces that have produced a particular system of symbols or social structure. The presumption that these rules are built into the neurology of the brain was rarely invoked in most structuralist analysis, whether in the social sciences or the humanities.

Within sociology proper, structuralism exerted a rather vague influence. For a time in the 1970s and 1980s, it was quite faddish to search for the underlying structure of texts, social structures, cultural systems, and virtually any phenomena studied by sociologists. Jargon from structuralism appeared in the works of prominent theorists, such as Anthony Giddens's (1984) use of the notion of "structural principles" as part of the explanation for how the "rules and resources" are implicated in "structuration." Others such as Richard Harvey Brown (1987) proclaimed that we should conceptualize "society as text" in which methods of structural linguistics can get at the underlying forms that organize societies. A variety of texts and edited books on structuralism began to appear (e.g., Rossi 1982), but in the end this kind of analysis did not take a firm hold in American

sociology. It had a much greater influence in continental Europe, but here, too, structuralist sociology did not last very long (Giddens 1987).

But arising out of structuralism came "poststructuralism," which was even more vague than structuralism. Still, poststructuralism did take hold and evolved into or merged with what is now called "postmodernism." Ironically, postmodernism is generally anti-science and highly relativistic, emphasizing that there are no grand narratives. All is text, and none should enjoy a privileged voice. By the time poststructuralism had become part of a broad postmodern movement, then, it bore little resemblance to the scientific stance of structuralists, who were not only decidedly pro-science but also were committed to discovering the universal principles by which texts and all other human systems are generated, giving such a discovery a highly privileged voice.

Despite its origins in structural linguistics, French structuralism became ever more vague. In contrast, structuralism took a very different turn in England and the United States, evolving into one of the most precise ways of conceptualizing structures. While French-inspired structuralism got much of the press, British and American versions of network analysis have endured. These approaches also come out of Durkheim, particularly the British line but also the American line, although today the two approaches are the same.

### **BRITISH STRUCTURALISM**

Radcliffe-Brown's early works were decidedly functional in the Durkheimian tradition of analyzing sociocultural patterns in terms of their functions for the larger social whole. Radcliffe-Brown also began to develop a conception of structure that was very close to Durkheim's emphasis on the number, nature, and relations among parts. And increasingly, Radcliffe-Brown emphasized structural analysis over functional analysis, stressing the importance of examining relations among entities. Radcliffe-Brown's disagreement with Lévi-Strauss is clear in a letter: "I use the term 'social structure' in a sense so different from yours as to make discussion so different as to be unlikely to be profitable. While for you, social structure has nothing to do with reality but with models that are built

up, I regard social structure as a reality" (quoted in Murdock 1953). A key British figure in the transition from Durkheim's emphasis on relations among entities was S. F. Nadel (1957), a social anthropologist who separated structural from functional analysis and emphasized that structural analysis must concentrate on the properties of relations that are invariant and always occur rather than the nature of the actors in these relations. For Nadel, structure is to be viewed as clusters of networks of relations among positions or actors. This point of view was developed into network analysis by J. Clyde Mitchell (1974) and John A. Barnes (1972), who took Nadel's imagery and converted it into a more pure form of network analysis.

### AMERICAN STRUCTURALISM

At the same time that French-inspired structuralism was making inroads into American intellectual life, social psychologists in the United States were developing an alternative way to examine structure. Jacob Moreno (1953) was critical here, as he developed the methodology of sociograms in which a matrix of relations among a set of actors was plotted and then converted into a line drawing of relations. Studies by Alex Bavelas (1948) and Harold Leavitt (1951) on communication in groups were the first to assess how the network structure of groups influenced patterns of communication. At about the same time, gestalt and balance theory approaches by Fritz Heider (1946), Theodore Newcomb (1953), and Dorwin Cartwright and Frank Harray (1956) used network graphs to examine patterns of balance in groups. Concurrently within mathematics, a more formal system of notation for networks was being developed (Luce and Perry, 1949), and it is from this system that modern-day network analysis emerged.

### THE MERGER OF BRITISH AND AMERICAN STRUCTURALISM INTO NETWORK ANALYSIS.

The British and American lines of what became network analysis are now merged in a view of structure that emphasizes (1) nodes or positions in a system, (2) mapping the relations among these nodes, and (3) examining generic patterns and forms of ties along such dimensions as the nature,

number, direction, reciprocity, transitivity, centrality, equivalence, density, and strength of ties. A fairly standard vocabulary and set of analytic techniques have emerged from network analysis, and the approach is now used in many other areas of sociological analysis. Thus, Durkheim's original vision of structure has been more fully realized in contemporary network analysis.

### FUNCTIONALISM AND STRUCTURALISM TODAY

Functionalism and structuralism share common roots, especially the works of Durkheim and Mauss, and hence it should not be surprising that these two theoretical traditions reveal certain affinities. First, both emphasize that the subject matter of sociology is to be relations among parts. Second, both stress part-whole analysis—the functions of parts for sustaining the whole in functionalism, and the view that parts are understood only as elements of a more inclusive whole, revealing underlying properties, in structuralism. Beyond these affinities, however, functionalism and structuralism diverge. French structuralism moves toward a highly vague view of underlying structures below visible phenomena whose generative processes constitute reality, whereas the merger of British and American network approaches moves toward analysis of the form of relations among interrelated units. In contrast, functionalism sustains an emphasis on how units meet the needs of larger social wholes, or it moves toward models of system differentiation during historical change and, alternatively, toward ecological models of population differentiation. It is safe to say that, today, all forms of functionalism and structuralism, even the network branch of structuralism, have little in common besides part-whole analysis. But even here, this kind of part-whole approach is conducted in very different ways by functionalists and structuralists.

Moreover, neither functionalism nor structuralism represented coherent approaches as we enter the twenty-first century. Functionalism was plagued by its emphasis on functional requisites, whereas French structuralism was always handicapped by a vagueness about what structure is. And so, few sociologists today would proclaim themselves to be either functionalists or structuralists;



these labels have negative connotations, and most seek to avoid such stigmatizing labels. If they perform elements of functional or structuralist analysis, they do so by another name.

## REFERENCES

- Alexander, Jeffrey C., ed. 1985 *Neofunctionalism*. Beverly Hills, Calif.: Sage.
- , and Paul Colomy 1985 "Toward Neofunctionalism." *Sociological Theory* 3:11–32.
- Barnes, John A. 1972 "Social Networks." Menlo Park, Calif.: Addison-Wesley Module No. 26.
- Bavelas, Alex 1948 "A Mathematical Model for Group Structures." *Applied Anthropology* 7:16–30.
- Blumer, Herbert 1969 *Symbolic Interaction: Perspective and Method*. Englewood Cliffs, N.J.: Prentice-Hall.
- Brown, Richard Harvey 1987 *Society as Text: Essays on Rhetoric, Reason, and Reality*. Chicago: University of Chicago Press.
- Cartwright, Dorin, and Frank Harary 1956 "Structural Balance: A Generalization of Heider's Theory." *Psychological Review* 63:277–293.
- Chomsky, Noam 1980 *Rules and Representations*. New York: Columbia University Press.
- Comte, Auguste 1830–1842 *Cours de Philosophie Positive*, 5 vols., Paris: Bachelier.
- 1851–1854 *Système de Politique: ou, Traité de Sociologie, Instituant la Religion de L'Humanité*. Paris: L. Mathias.
- Coser, Lewis A. 1956 *The Functions of Social Conflict*. London: Free Press.
- Dahrendorf, Ralf 1958 "Out of Utopia: Toward a Reorientation of Sociological Analysis." *American Journal of Sociology* 74:115–127.
- Davis, Kingsley 1948 *Human Societies*. New York: Macmillan.
- , and Wilbert Moore 1945 "Some Principles of Stratification." *American Sociological Review* 4:431–442.
- Dore, Phillip Ronald 1961 "Function and Cause." *American Sociological Review*. 26:843–853.
- Durkheim, Émile 1893 *The Division of Labor in Society*. New York: Free Press.
- 1895 *The Rules of the Sociological Method*. New York: Free Press.
- 1898 "Incest: The Nature and Origin of the Taboo." *Année Sociologique* 1:1–70.
- 1912 *Elementary Forms of the Religious Life*. New York: Macmillan.
- , and Marcell Mauss (1903) 1963 *Primitive Classification*. London: Cohen and West.
- Giddens, Anthony 1984 *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley: University of California Press.
- 1987 "Structuralism, Post-structuralism and the Production of Culture." In A. Giddens and J. Turner, eds., *Social Theory Today*. Cambridge, U.K.: Polity Press.
- Hannan, Michael, and John Freeman 1977 "The Population Ecology of Organizations." *American Journal of Sociology* 82:929–964.
- Heider, Fritz 1946 "Attitudes and Cognitive Orientation." *Journal of Psychology* 2:107–112.
- Hertz, Robert (1909) 1960 *Death and the Right Hand*. London: Cohen and West.
- Hoyt, Homer 1939 *The Structure and Growth of Residential Neighborhoods in American Cities*. Washington, D.C.: Federal Housing Authority.
- Jakobson, Roman 1962 *Selected Writings I: Phonological Studies*. The Hague: Mouton.
- 1971 *Selected Writings II: Word and Language*. The Hague: Mouton.
- Leavitt, Harold J. 1951 "Some Effects of Certain Communication Patterns on Group Performance." *Journal of Abnormal and Social Psychology* 36:38–50.
- Lévi-Strauss, Claude 1945a "The Analysis of Structure in Linguistics and in Anthropology." *Word* 1:1–21.
- 1945b "French Sociology." In Georges Gurwitsch and Wilbert E. Moore, eds., *Twentieth Century Sociology*. New York: Philosophical Library.
- (1949) 1969 *The Elementary Structures of Kinship*. Boston: Beacon Press.
- (1950) 1987 *Introduction to the Work of Marcel Mauss*. London: Routledge and Kegan Paul.
- 1953 "Social Structure." In A. Kroeber, ed., *Anthropology Today*. Chicago: University of Chicago Press.
- 1963 *Structural Anthropology*. New York: Basic Books.
- Levy, Marion J. 1952 *The Structure of Society*. New Haven, Conn.: Yale University Press.
- Luce, R. Duncan, and A. D. Perry 1949 "A Method of Matrix Analysis of Group Structure." *Psychometrika* 14:94–116.
- Luhmann, Niklas 1982 *The Differentiation of Society*. New York: Columbia University Press.
- Malinowski, Bronislaw 1913 *The Family Among the Australian Aborigines*. New York: Schocken.

- 1944 *A Scientific Theory of Culture*. Chapel Hill: University of North Carolina Press.
- Mauss, Marcel, with Henry Beuchat (1904–1905) 1979 *Seasonal Variations of the Eskimo: A Study in Social Morphology*. London: Routledge and Kegan Paul.
- (1925) 1941 *The Gift, Forms and Functions of Exchange in Archaic Societies*. New York: Free Press.
- Merton, Robert K. 1949 “Manifest and Latent Functions.” In *Social Theory and Social Structure*. New York: Free Press.
- 1957 *Social Theory and Social Structure*, rev. ed. New York: Free Press.
- Mills, C. Wright 1959 *The Sociological Imagination*. New York: Oxford University Press.
- Mitchell, J. Clyde 1974 “Social Networks.” *Annual Review of Anthropology* 3:279–299.
- Moreno, Jacob 1953 *Who Shall Survive?*, rev. ed. New York: Beacon House.
- Münch, Richard 1982 *Theory of Action: Reconstructing the Contribution of Talcott Parsons, Emile Durkheim and Max Weber*, 2 vols. Frankfurt: Suhrkamp.
- Murdock, George P. 1953 “Social Structure.” In Sol Tax, Loren Eiseley, Irving Rouse, and Carl Voegelin, eds., *An Appraisal of Anthropology Today*. Chicago: University of Chicago Press.
- Nadel, S. F. 1957 *The Theory of Social Structure*. New York: Free Press.
- Newcomb, Theodore M. 1953 “An Approach to the Study of Communicative Acts.” *Psychological Review* 60:393–409.
- Park, Robert 1936 “Human Ecology.” *American Journal of Sociology* 42:1–15.
- , Ernest Burgess, and Roderick D. McKenzie 1925 *The City*. Chicago: University of Chicago Press.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- , Robert F. Bales, and Edward Shils 1953 *Working Papers in the Theory of Action*. New York: Free Press.
- Radcliffe-Brown, A. R. (1914) 1922 *The Andaman Islanders*. Cambridge, U.K.: Cambridge University Press.
- 1924 “The Mother’s Brother in South Africa.” *South African Journal of Science* 21:42–63.
- 1935 “On the Concept of Function in Social Science” (reply to Lesser 1935). *American Anthropologist* 37:394–402.
- 1952 *Structure and Function in Primitive Society*. London: Cohen and West.
- Rossi, Ino, ed. 1982 *Structural Sociology*. New York: Columbia University Press.
- Sahlins, Marshall D. 1966 “On the Delphic Writings of Claude Lévi-Strauss.” *Scientific American* 214:134.
- Saussure, Ferdinand de (1915) 1966 *Course in General Linguistics*. New York: McGraw-Hill.
- Spencer, Herbert 1864–1867. *The Principles of Biology*, 3 vols. New York: D. Appleton.
- 1874–1896 *The Principles of Sociology*, 3 vols. New York: D. Appleton.
- Trubetzkoi, Nikolai (1949) 1964 *Principes de Phonologie*. Paris: Klincksieck.
- 1968 *Introduction to the Principles of Phonological Description*. The Hague: Martinus Nijhoff.
- Turner, Jonathan H. 1995 *Macrodynamics: Toward a General Theory on the Organization of Human Populations*. New Brunswick, N.J.: Rutgers University Press.
- , and Alexandra Maryanski 1979 *Functionalism*. Menlo Park, Calif: Benjamin Cummings.
- Turner, Stephan Park, and Jonathan H. Turner 1990 *The Impossible Science: An Institutional Analysis of Sociology*. Newbury Park, Calif: Sage.
- Wuthnow, Robert 1987 *Meaning and Moral Order: Explorations in Cultural Analysis*. Berkeley: University of California Press.

ALEXANDRA MARYANSKI  
JONATHAN H. TURNER

## FUTURES STUDIES AS HUMAN AND SOCIAL ACTIVITY

Future thinking has always been part of human history, for humans become human when they think about the future, as John McHale (1969) wrote. Some form of future thinking exists in almost all known societies; it is a universal phenomenon, as Wendell Bell (1997) writes, that ranges from divination, which involves the unveiling of the unknown, to the recent development of futures studies. Divination may be related to decisions taken by whole tribes and nations, just as the aim of scientifically rigorous futures studies is to support decision making.

In this particular context, it is important to trace the development of future thinking and the formalization of futures studies after World War II.

## TERMINOLOGY

As in most disciplines, issues of terminology in futures studies are related as much to the word itself as to the underlying concept expressed by it.

From the 1940s to the 1970s, *forecasting* was the most important concept of futures studies. In 1967, Eric Jantsch defined *forecasting* “as a probabilistic statement, on a relatively high level of confidence about the future” (p. 15). In this rigorous definition, the use of the adverb “relatively” indicates the element of uncertainty that is always present in futures studies in general.

In France and in Spanish-speaking countries, however, they tend to prefer the term *prospective*. The term and the concept were actually born in France in the 1950s when Gaston Berger (1958) described *prospective* as “a way of focusing and concentrating on the future by imagining it full-blown rather than by drawing deductions from the present” (p. 3). In *prospective*, the main element is a project for the future. The idea is that by understanding the past (which cannot be changed but only interpreted), one can, in the present, choose from among the many possible and probable futures the one that is more desirable. According to Bertrand de Jouvenel (1967), the present cannot be changed either, as it is but a fleeting moment. Hence, the future is the only space we have to influence. Clearly, the prospective approach is completely different from the forecasting approach, as Michel Godet (1979) has correctly underlined.

The term *prevision*, as indicated by Barbieri Masini (1993) is not used in English or in French. In English it is thought to have the specific meaning of *pre-vision*, or seeing ahead (something not possible, as the future does not exist); in French it is not used because of its theological associations. In Italian the term *previsione* is used because a term equivalent to prospective does not exist. *Prospettiva* has a different meaning and projection is closer to forecast, which, as already mentioned here and discussed by Godet (1979), does not have the implications of prospective. Consequently when the term *previsione* is used in Italian it is necessary to qualify it. For example, *previsione umana* and *previsione sociale*, which has in itself the capacity to build a set of futures that are different from the past and the present (as with prospective) but with the addition of some other qualifications, as indicated by Barbieri Masini and Medina (1993). In

this meaning, to build the future is seen simultaneously as a need, a choice, and a way of life that are related to the human being and to a given society. The need is due in part to the rapidity of change in modern times but also to an ethical requirement to create a future that is humane, as Josef Fuchs (1977) says, and at the same time social, as related to social choices. According to Barbieri Masini (1993), embedded in *previsione umana e sociale* is the need to build the future, but with a choice in terms of deciding to propose one or many alternative futures. It is a way of life, as it goes beyond a discipline and becomes a distinguishing feature of a group of people. Hence, *previsione umana e sociale* has the following basic functions: It is geared to a project of the future; it clarifies the goals of a given person or social group; and it is educational, in terms of helping to create responsibility for the future.

*Prognosis* is a term that is very much used in German-speaking countries and in Central and Eastern Europe. It is very close to forecast and forecasting.

Among the many other terms, it is important to mention *futuribles*, which was used by Bertrand de Jouvenel (1967) to indicate the many possible, probable and desirable futures that exist. In the debate about *futuribles*, Jouvenel’s thinking is reminiscent of the thinking of many in the Middle Ages, especially the emphasis put on freedom by the Jesuit Molina (in 1588), whose writings were published in 1876 in France. Jouvenel, in the conceptual discussion of terms, especially “*futuribles*,” refers to the discussion of the symmetry between the past and the future that took place in the eighteenth century and goes back to 1588 to the Jesuit Molina who had favored the position of human freedom of choice in relation to the interpretation of the past and the choice of the future.

An interesting more recent term is *foresight*. Actually it was first used by D.H. Wells in 1932 in a speech for the BBC advocating the need for experts in foresight. As described by Irvin and Martin (1984), nowadays the term is mostly used in the sense of outlook, meaning a cluster of systematic attempts to look ahead and to choose more effectively. Foresight takes into account the fact that there is not one future but many—only one of which, however, will occur. Foresight is also used at the national level as a support for national

planning, to identify the possible technological futures and their environmental and social impact.

As regards the actual discipline, in addition to *futures studies*—a very broad term that contains all the different ways of looking at the future, from projection to utopia—there is the term *futurology*, which was defined by Ossip Flechtheim (1966) as the search for the logic of the future in parallel to the search for the logic of the past in history. Unfortunately, this concept has lost its correct meaning and is now used casually to refer to any fantasy about the future.

*Futuristics* is another term that is often used, especially in North America. Its meaning is similar to that of futures studies.

### A BRIEF HISTORY OF FUTURES STUDIES

It is hard to establish a definite date for the beginning of futures studies. Many indicate the period immediately after World War II with the development in the United States of what was called *technological forecasting* at RAND (Research and Development) Corporation. According to Wendell Bell (1997), this can be linked to the *operational research* conducted in the United Kingdom to predict and project the moves of the German bombers.

More or less in the same period, the Europeans also contributed to what may be considered the “beginnings” of futures studies—in France, Bertrand de Jouvenel with *futuribles* and Gaston Berger with *prospective*, followed by Pierre Massé and many others; Pater in the Netherlands, Fred Polak with his important books *The Image of the Future* and *Prognosis: A Science in the Making*.

The three major organizations dealing with futures studies were established at the end of the 1960s: the World Futures Studies Federation, which started in embryo in Oslo in 1967 and was formally founded in Paris at UNESCO in 1973; the World Future Society, essentially a North American organization founded by Edward Cornish in 1967; and the Club of Rome, founded in Rome by Aurelio Peccei and Alexander King in 1968. The latter in particular had a major impact on decision making with the project “The Limits to Growth,” which is still relevant both in relation to the limits

of the earth and of the need to look at the future in the long term, and in relation to a cluster of human and social issues (the global problematic).

The 1970s posed a serious challenge to futures studies. The severe economic crisis of that period showed that in looking at the future it is necessary to address many aspects and paces of change, given that economic and technological changes are much more rapid than social and cultural changes, for example. The 1970s also challenged the concept of the future as being linear development of the present and the past. There was a re-appraisal of futures studies in the 1980s, especially in Europe. In countries such as the Netherlands, Sweden, and Finland, futures studies acquired political importance. In the 1990s developing countries also started to make use of futures studies; with the exception of India, China, and certain Latin American countries, most developing countries had had little interest in futures studies up until then. In the same period, futures studies reemerged very strongly in the United States with the launching of such major projects as the Millennium Project, supported by the United Nations University and the Smithsonian Institute. The Millennium Project, directed by Jerome C. Glenn and Theodore J. Gordon, studies the major threats and challenges of the future and produces the “state of the future” report. The debate on methodologies was also revived, sparking the publication of many basic texts that became prominent in the world debate. Schools of futures studies emerged in Finland and Hungary, and there was increased interest in Australia.

Numerous sociologists have made important contributions to futures studies: William F. Ogburn was appointed by President Herbert Hoover to chair the Research Committee on Social Trends as early as 1929. Harold Lasswell, who laid the foundations for the science of political choices also wrote many less well-known articles on futures perspectives. Daniel Bell wrote the famous report of the Commission Toward the Year 2000 for the American Academy of Arts and Sciences in 1967. John McHale wrote many books and contributed to the founding of the World Futures Studies Federation in the 1960s and 1970s. Elise Boulding greatly contributed with her work on peace in the future and the role of women in the future. One

could mention many more names. The main point to stress is that futures studies and sociology have not always been considered two separate and distinct disciplines.

### CHARACTERISTICS OF FUTURES STUDIES

Many characteristics of futures studies relate to the linking of present action with future results. Barbieri Masini (1993) identifies seven characteristics in particular:

1. *Transdisciplinary*. Futures studies go beyond the interdisciplinary nature of social science. Futures studies needs the parallel approaches of different disciplines as well as the combined effort of many approaches in addressing the complexity of present problems in a rapidly changing society. It needs disciplinary approaches that identify common assumptions and use common methodologies. In terms of methodologies, examples of such approaches are the Delphi and the scenario-building methods, which function with the support of mathematics, statistics, economics, sociology, history, and psychology. To quote Fred Polak: "all kinds of separate, fragmented portions of the jigsaw puzzle are of little avail, unless they are fitted together in the best possible way, to form an image of the future depicting a number of main areas of development" (1971, p. 261).

2. *Complex*. This represents the other side of the transdisciplinary nature of futures studies. As a concept, complexity was much debated in futures studies, even before it became a clear issue in social sciences, because it is related to uncertainty: The greater the complexity of a given situation, the greater the level of uncertainty; the greater the number of variables needed to describe a social situation, the greater the uncertainty. Because the future is so uncertain, futures studies has to devise methods for lowering the level of uncertainty: One such method is alternative scenario building.

3. *Global*. Its global nature is probably the best-known characteristic of futures studies. It is now generally accepted that the future is determined by a continuous interplay between local issues and global issues. Some issues emerge as local and then become global, such as an economic crisis in a given country or part of the world. Some issues are

born global and subsequently affect the local level such as the greenhouse effect.

4. *Normative*. The impact of the normative is much greater in futures studies than is usually the case in social sciences. Frameworks of values and systems of values can never be completely eliminated in futures studies. Many methods try to lower the level of impact of the normative, but it is always present: Thinking about the future, making decisions about the future, is always related to some hope or fear.

5. *Scientific*. This is the most debated characteristic of futures studies, precisely because the normative is always present. Great effort has been made to invent devices and methods for increasing the scientific level of futures studies. If it is not possible to define futures studies as scientific, then it must be absolutely rigorous in terms of application and methods. It must be used according to strict scientific rules. Yehezkel Dror (1974) stresses the importance of combining a clinical and a human approach.

6. *Dynamic*. Futures studies is very dynamic: It is constantly in search of stronger foundations, of better approaches, and of more effective methods for facing the rapidity of change.

7. *Participatory*. There is less consensus on this characteristic of futures studies than on the previous ones. It is obvious that futures studies methods should be developed with the participation of those who are responsible for choosing the future, that is, decision makers at every level. This may not be possible at every step in the application of methods, but it is certainly something to strive for. Participation is definitely an essential characteristic of scenario-building and Delphi applications.

### LIMITS OF FUTURES STUDIES

Like all disciplines futures studies has limits. Barbieri Masini (1993) has singled out six such limits.

1. *Self-altering*. The moment a forecast, or the result of a futures study, becomes public, it produces consequences that alter the reality in which it operates. In other words, a self-realizing or a self-defeating effect may invalidate the value of the forecast. This danger was underlined long ago by Robert Merton ([1938] 1973) and is still valid today.

2. *Psychological aspects.* These are crucial in futures studies, as inevitably in looking ahead one is influenced by hopes and fears. Fear of the unknown is ever present and can have a negative impact on the need to think in alternative terms as the future requires. It is also easy to underestimate the changes that will occur in the future.

3. *Irrational aspects.* In future processes and events there is always an irrational element that cannot be quantified or evaluated; for example, the whims of a given head of state or the religious reaction of a given population.

4. *Implicit hypotheses.* These are present in any futures exercise. This aspect is of course related to the normative characteristic mentioned earlier, however, in using futures studies, it is essential to detect the implicit hypotheses.

5. *A posteriori verification.* A definite limit of futures studies is that it is only possible to verify the validity of a given study after the events forecast have occurred. This is often done and is an extremely useful way of learning more about the application of methods.

6. *Availability of reliable data.* This clearly is a crucial aspect: Any future-oriented study must be able to rely on good quantitative and qualitative data. There may, for example, be a lack of historical data on which to base a forecast or a lack of relevant ad hoc surveys. This is one of the most important limits of futures studies, for if forecasts are to be reliable, data must be rigorously evaluated.

## METHODS

Before entering the debate on methods, it is important to underline the accepted typology of futures studies methods. Futures studies can be *extrapolative* (opportunity-oriented) or *normative* (mission-oriented). The definitions in parenthesis are those of Eric Jantsch (1967). In other words, to extrapolate from the past to the present and into the future constitutes one group of methods, of which projections are one. Another is to start from what is needed in the future—an image, a goal, an objective—and work backward, searching in the present for possible and probable ways of realizing it. Nowadays futures studies are never wholly

extrapolative or normative; they usually emphasize one or the other aspect, given the normative characteristic of futures studies.

## SCENARIOS

The term and method were introduced by Herman Kahn during the 1960s. Nowadays *scenario* is a general term that is used to refer to different approaches: those of the Stanford Research Institute (SRI), of the Battelle Institute, of the studies by Michel Godet and others. A scenario can be defined as an ensemble created to describe a future situation and the sequence of events leading from the present situation to a future one.

Theoretically speaking, scenarios are a synthesis of several different hypothetical routes (events, actors, strategies) that lead to different possible futures. Practically speaking, scenarios often describe specific sets of events and variables that have been put together with the aim of focusing on causal processes and related decision-making. According to Herman Kahn (1967), scenarios answer the following basic questions: How does a situation evolve step by step from the present to the future? What are the possible alternatives that different actors in the different moments of decision making can use in order to anticipate, change, or facilitate the process?

From an epistemological point of view, scenarios are analytical and empirical constructions. They are hypotheses that do not “predict” the future but rather indicate a series of options and probable situations. No scenario will ever precisely anticipate the occurrence of a given event; rather it will suggest alternatives with the aim of sensitizing decision makers to what might happen. Hypotheses are never invented but are founded on a rational, consistent diagnosis of the forces that may model events. According to Michel Godet (1995), the scenario hypothesis should respond to the following five conditions. They must be: pertinent, coherent, realistic, important, and transparent.

Scenarios are instruments for better decision making. Their aim is to lower the level of uncertainty and increase the level of understanding of the consequences of actions in the present.

In looking at scenarios, one presumes an awareness on the part of the author of the rapidity and

interrelatedness of change, which will affect both decisions and the understanding of consequences, which can change a preexistent situation, either completely or in part. Scenarios are both synoptic and simultaneous: They can be used to analyze many variables at once and identify the respective turning points in terms of decision-making in steps of five, ten, or twenty years. Timing depends on the area of interest. For example, in the economic area the time span considered will be much shorter than in the educational area.

In recent years, use of the scenario technique has increased greatly, as have its applications. The scenario method seems to be particularly useful in the following situations: (1) to detect long-term trends that can help to formulate alternative within a given context; (2) to identify potential discontinuities and situations and alert organizations, countries, or regions to foresee them and thus prepare contingency plans; (3) to alert organizations, regions, or countries of possible interrelated changes as a reference for planning; (4) to provide a basis for analyzing risks as interactions between socio-economic areas that may lead to risks and are not understandable if seen only in a specific and isolated area; (5) to evaluate the results of different strategies in different areas that may not have been developed in the awareness of interrelations.

In the public sector the best-known applications include the following. In the 1960s, Herman Kahn used scenarios connected to military and strategic studies that were developed at the RAND corporation. In the 1960s and 1970s, territorial planning by DATAR was used in France. Scenarios derived from systems analysis and global models, such as *The Limits to Growth* by Meadows and Meadows (1972), analyzed the global consequences of the interrelated growth of population, agricultural production, industrial development, environmental pollution, and use of natural resources. Finally, Jacques Lesourne directed an important exercise of Interfutures for the Organization for Economic Cooperation and development (OECD), which identified the alternatives of the relationships between the North and the South regions of the world.

In scenario building the private sector intervened, perhaps late, but in a sophisticated manner. Shell International Petroleum Company (Royal

Dutch/Shell Group) used scenarios before the energy crises of 1973 and before the debacle of the Soviet Union, its major competitor in the petroleum area. Other large oil companies followed Shell's example and made extensive use of scenarios—ARGO in the late 1970s and, more recently, Pacific Gas and Electric.

In the private sector scenarios are used by enterprises in many sectors. Financial services have used scenarios to understand competitors and regulate uncertainty. Insurance companies, such as the Allied Irish Bank, have used scenarios to support strategic planning within a constantly changing context.

## REFERENCES

- Barbieri Masini, Eleonora 1993 *Why Futures Studies?* London: Grey Seal.
- , and Javier Medina Vasquez 2000 "Scenarios as Seen From a Human and Social Perspective." In *Technological Forecasting and Social Change*. Amsterdam: Elsevier.
- Bell, Wendell 1997 *Foundations of Futures Studies* 2 vols. News Brunswick, N.J.: Transaction Publishers.
- Berger, Gaston 1958 *L'Attitude*. Paris: Presses Universitaires de France.
- Dror, Yehezkel 1974 "Futures Studies—Quo Vadis?" In *Human Futures, Needs—Societies—Technologies*. London: IPC Science and Technology Press.
- Flechteim, Ossip 1966 *History and Futurology* Verlag Anton Main Meisenheim, Arn Glam.
- Fuchs, Josef 1977 "Morale come progettazione del futuro dell'uomo" [Ethics as Building the Human Future]. In P. C. Beltrao, ed., *Pensare il futuro (Thinking the Future)*. Rome: Edizioni Paoline.
- Glenn, Jerome C., Gordon Theodore, Jr. 1999 *State of the Future, Challenges We Face at the Millenium*. American Council of United Nations Universes.
- Godet, Michel 1995 "From Anticipation to Action." *A Book of Strategic Prospectice, Future Oriented Studies*. Paris: UNESCO Publishing.
- 1995 "Global Scenarios: Morphological and Probability Analysis." In *Scenario Building, Convergence, and Differences* Profutures Workshop, Institute for Prospective Technological Studies (IPTS). European Commission Joint Research Council, 17–30.
- 1995 "From Anticipation to Action: A Book of Strategic Prospective." Paris: UNESCO Publishing.
- 1979 *The Crisis in Forecasting and the Emergence of the Prospective Approach*, UNITAR. New York: Pergamon.

- Hatem, Fabrice 1993 "Introduction à la prospective." Paris: *Gestion-Economica*. 86–89.
- 1996 "La prospective. Pratiques et méthodes." *Gestion-Economica*. 222–232.
- Irvin, J., and B. Martin, B. 1984 *Foresight in Science, Picking the Winners*. London: Pinter Publishers.
- Jantsch, Eric 1967 *Technological Forecasting in Perspective*. Paris: Ocse.
- Jouvenel, Bertrand de 1967 *The Art of Conjecture*. New York: Basic Books.
- Kahn, Herman, and Anthony Wiener 1967 *The Year 2000—A Framework for Speculation in the Next 33 Years*. London: Macmillan.
- Malaska, Pentti 1995 "Survey of the Use of the Multiple Scenario Approach in Big European Companies Since 1973." In *Scenario Building, Convergence, and Differences*. Profutures Workshop, Institute for Prospective Technological Studies (IPTS), European Commission Joint Research Council, 42–46.
- McHale, John 1969 *The Future of the Future*. New York: George Braziller.
- Merton, Robert K. (1938) 1973 *The Sociology of Science*. Chicago: University of Chicago Press.
- Norse, David 1979 "Scenario Analysis in Interfutures." *Futures*. (October):412–428.
- Ozbekahn, Hasan 1970 "Verso una teoria generale della pianificazione." (Towards a General Theory of Planning). *Futuribili* (October):25–128.
- Polak, Fred 1971 *Prognosis: A Science in the Making*. Amsterdam: Elsevier.
- 1973 *The Image of the Future*. (ed. and abridged by Elise Boulding). New York: Elsevier.
- Schwartz, Peter 1996 "The Art of the Long View." *Planning for the Future in an Uncertain World*. New York: Doubleday.
- , and Kees Van der Heijden 1996 "Culture d'entreprise et planification par scénarios: une relation de coévolution." In Jacques Lesourne and Christian Stoffaës, eds., *La prospective stratégique d'entreprise*. Paris: Interditions.
- Wack, Peter 1985 "Scenarios: Shooting the Rapids." *Harvard Business Review* (November/December):139–150.
- 1985 "Scenarios: Uncharted Waters Ahead." *Harvard Business Review* (September/October):73–89.
- Wilson, Ian 1995 "Linking Intuition and Structure: an Integrated Approach to Scenario Development." In *Scenario Building, Convergence, and Differences*, Profutures Workshop, Institute for Prospective Technological Studies (IPTS), European Commission, Joint Research Council, 31–41.
- 1978 "Scenarios." In Jib Fowles, ed., *Handbook of Futures Research*. Westport, Conn.: Greenwood Press.

ELEONORA BARBIERI MASINI



# G

---

## GAME THEORY AND STRATEGIC INTERACTION

A *game* is a situation that involves two or more decision makers (called *players*), where (1) each player faces a choice between at least two behavioral options, (2) each player strives to maximize utility (i.e., to achieve the greatest payoff possible), and (3) the payoff obtained by a given player depends not only on the option that he or she chooses but also on the option(s) chosen by the other player(s). In virtually all games, some or all of the players have fully or partially opposing interests; this causes the behavior of players to be proactive and strategic.

The *theory of games* is a branch of applied mathematics that rigorously treats the topic of optimal behavior in two-person and  $n$ -person games. Its origins go back at least to 1710, when the German mathematician-philosopher Leibniz foresaw the need for a theory of games of strategy. Soon afterward, James Waldegrave (in Montmort 1713; 1980) formulated the concept of maximin, a decision criterion important to game theory. In his book *Mathematical Psychics*, Edgeworth (1881; 1995) made explicit the similarity between economic processes and games of strategy. Later, theorists such as Zermelo (1913) stated specialized propositions for certain games (e.g., chess). Not until the work of Borel ([1921–1927] 1953) and von Neumann (1928), however, did the foundations of a true theory of games appear. A landmark of the modern era, von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1944), extended game theory to problems involving more

than two players. Luce and Raiffa (1957) published the first widely used textbook in game theory. For more details regarding the early history of game theory, see Dimand and Dimand (1996) and Weintraub (1992).

Game theory has continued to develop substantially in recent years. Many introductory presentations of the modern mathematical theory are available. Among these are Friedman (1990), Jones (1980), Myerson (1991), Owen (1982), Romp (1997), and Szep and Forgo (1985).

Beyond its status as a branch of applied mathematics, game theory serves social scientists as a tool for studying situations and institutions with multiple decision makers. Some of these investigations are empirical, while others are primarily analytic in character. The dependent variables of central concern in games include allocation of payoffs (i.e., who receives what rewards or bears what costs) and formation of coalitions (i.e., which of various possible alliances among players occur in a game.) Other concerns include whether outcomes of a game are stable or not, whether outcomes are collectively efficient or not, and whether outcomes are fair or not in some specific sense.

### GAME-THEORETIC CONCEPTS

Mathematical game theory provides three main tools that assist in the analysis of multiperson decision problems. These include a descriptive framework, a typology of games, and a variety of solution concepts.

**Descriptive Framework.** At base, a description of any game requires a list of all players, the strategies available to each player, the logically possible outcomes in the game, and the payoff of each outcome to each player. In some instances, a game's description also includes a specification of the dynamic sequence of play and of the (possibly limited or incomplete) information sets available to players. Payoffs in a game are expressed in terms of *utility*; this provides a standard means of comparing otherwise diverse outcomes.

An analyst can model or represent a game in various forms. *Extensive form* depicts all possible strategies of players in a tree format. It is especially useful for modeling games in which play occurs in stages or over time. *Strategic form* (also called normal form or a "payoff matrix") shows payoffs to players as a function of all strategy combinations. *Characteristic function form* lists the minimum payoffs assured for each of the coalitions in a game. Whereas extensive and normal forms pertain to virtually all types of games, characteristic function form pertains only to cooperative games (i.e., games that permit coalitions).

**Typology of Games.** The second tool from game theory is a general typology of games. This provides a means of codifying or classifying games vis-à-vis one another. At base, there are four major types of games. Games can be either *static* (i.e., single time period) or *dynamic* (multiple time periods), and they can involve either *complete information* (all relevant information is shared and held in common) or *incomplete information* (some information is private and held only by some players). Much of classic game theory was formulated with reference to static games involving complete information; more recent developments have extended the theory to dynamic games and also to games involving incomplete information.

Games can be *two-person* or *n-person* (more than two players), and they can be further classified as *cooperative* or *noncooperative*. Cooperative games permit players to communicate before reaching decisions and include some mechanism that enables players to make binding agreements regarding coordination of strategies. Noncooperative games do not permit players to communicate or to form binding agreements prior to play. In other words, cooperative games enable players to form coalitions whereas noncooperative games do not.

Among cooperative games, some are *sidepayment games* while others are *nonsidepayment games*. Sidepayment games permit players to transfer payoffs (utility) within coalitions; nonsidepayment games do not. A further distinction applicable to cooperative sidepayment games is that between *simple games* and *nonsimple games*. Simple games are those in which the characteristic function assumes only two values, whereas nonsimple games are those in which the characteristic function has more than two values. Analysts use simple games primarily to model social processes with binary outcomes (e.g., win-lose, succeed-fail, etc.)

**Solution Concepts.** The third set of tools provided by game theory is a variety of solution concepts. A *solution concept* is theory of equilibrium that predicts (behaviorally) or prescribes (normatively) the allocation of payoffs to players in games. In other words, a solution concept specifies how a game will turn out when played. For this reason, solution concepts are among the most important contributions of game theory.

Game theorists have developed numerous solution concepts. These differ not only in the underlying assumptions but also in the predictions they make. For static noncooperative games, the most prominent solution is the Nash equilibrium (Nash 1951); there are many extensions of this concept (summarized in van Damme 1987). Other approaches to the solution of noncooperative games are those of Harsanyi and Selten (1988) and Fraser and Hipel (1984).

For static cooperative games, there are several classes of solution concepts. One prominent class consists of solutions that predict outcomes which are collectively rational (i.e., imputations). Included in this class are the core (Aumann 1961; Gillies 1959), the Shapley value (Shapley 1953), and the nucleolus (Schmeidler 1969). Other solutions in this class are the disruption nucleolus (Gately 1974; Littlechild and Vaidya 1976), the disruption value (Charnes et al. 1978), the p-center solution (Spinetto 1974), and the aspiration solution (Bennett 1983). Another class of solutions for cooperative games includes concepts that make payoff predictions contingent upon the coalition structures that form during play; these payoff allocations are usually coalitionally rational. Included are the  $M_1(i)$  bargaining set (Aumann and Dreze 1974; Aumann and Maschler 1964), the competitive bargaining

set (Horowitz 1973), the kernel (Davis and Maschler 1965), the Myerson–Shapley solution (Aumann and Myerson 1988; Myerson 1977), the equal division kernel (Crott and Albers 1981), and the alpha-power solution (Rapoport and Kahan 1982). Recently, a third class of solutions has emerged for cooperative games. Solutions in this class attempt not only to determine endogenously which coalition structure(s) will emerge but also to specify the associated payoffs to players. One solution in this class is the central-union theory (Michener and Au 1994; Michener and Myers 1998), which predicts coalition formation probabilistically. Another solution in this class is the viable proposals theory (Sengupta and Sengupta 1994).

### EXPERIMENTAL STUDIES OF GAMES

Laboratory experimentation on two-person and  $n$ -person games commenced in the early 1950s (e.g., Flood 1952) and it continues to the present. Some gaming studies are primarily descriptive in nature, whereas others investigate the predictive accuracy of various solution concepts.

**Experiments of Two-Person Games.** Investigators have conducted literally thousands of experiments on two-person games. Most of these treat noncooperative games, although some do treat cooperative games in various forms. Some studies investigate constant-sum games, whereas others treat non-constant-sum games (primarily such archetypal games as the prisoner’s dilemma, chicken, battle of the sexes, etc.).

The major dependent variables in the two-person studies are the strategies used by players (particularly the frequency of cooperative choices) and the payoffs received by players. Independent variables include the type of game, strategy of the confederate, information set, interpersonal attitudes of players, sex of players, motivational orientation of players, and magnitude and form of payoffs.

Some of this research seeks to understand how differences in game matrices affect play (Harris 1972; Rapoport et al. 1976). Another portion describes how players’ strategies vary as a function of the confederate’s strategy (i.e., partner’s history of play over time); this is reviewed in Oskamp (1971). Another portion of this work investigates

the extent to which predictions from the minimax theorem approximate observed payoffs in constant-sum games; Colman (1982, ch. 5) reviewed these findings. Still other work covers cooperative bargaining models; Roth (1995) reviewed research on bargaining experiments. Some experimentation on two-person games has addressed the impact of players’ value orientation on cooperation (McClintock and Liebrand 1988; Van Lange and Liebrand 1991). General reviews of experimental research on two-person games appear in Colman (1982), Komorita and Parks (1995), and Pruitt and Kimmel (1977).

**Experiments of  $n$ -Person Noncooperative Games.** There are several lines of experimentation on  $n$ -person noncooperative games. One line investigates multiperson compound games derived from  $2 \times 2$  matrices (e.g.,  $n$ -person chicken,  $n$ -person battle of the sexes, etc.). Important among these is the  $n$ -person dilemma (NPD) game, wherein individually rational strategies produce outcomes that are not collectively rational. The NPD serves as an abstract model of many phenomena, including conservation of scarce natural resources, voluntary wage restraint, and situations involving the tragedy of the commons (Hardin 1968; Hartwick and Yeung 1997; Moulin and Watts 1997). The literature contains many experimental studies of the NPD and other social dilemmas (e.g., Liebrand et al. 1992; Rapoport 1988). In addition to varying the payoff matrix itself, studies of this type investigate the effects of such factors as group identity, self-efficacy, perceptions of other players, value orientation, uncertainty, and players’ expectations of cooperation. Reviews of research on NPDs and similar games appear in Dawes (1980), Kollock (1998), Komorita and Parks (1999), Liebrand and colleagues (1992), Liebrand and Messick (1996), and Messick and Brewer (1983).

A related line of research is that by experimental economists on markets and auctions (Smith 1982). This work investigates market structures (such as competitive exchange, oligopoly, and auction bidding) in laboratory settings (Friedman and Hoggatt 1980; Plott and Sunder 1982). Many of these structures can be viewed as noncooperative games. Plott (1982) provides a review of studies investigating equilibrium solutions of markets—the competitive equilibrium, the Cournot model, and the monopoly (joint maximization) model.

There is an increasingly large experimental literature on auctions, some of which is game-theoretic in character. Since auctions usually entail incomplete information (buyers have private information about their willingness to pay and ability to pay), these studies investigate the effects on bidding behavior of such variables as differential information, asymmetric beliefs, and risk aversion. They also investigate different institutional forms, such as English auctions, Dutch auctions, double auctions, and sealed bid-offer auctions (e.g., Cox et al. 1984; Smith et al. 1982). Reviews of the theoretical literature on auctions appear in Engelbrecht-Wiggans (1980), Laffont (1997), and McAfee and McMillan (1987). Kagel (1995) provides a survey of experimental research on auctions.

**Experiments of  $n$ -Person Weighted Majority Games.** Weighted majority games are an important subclass of cooperative, sidepayment, simple games. They serve as models of legislative or voting systems. Theorists have developed many special solution concepts for these games. Early theories applicable to weighted majority games are the minimum power theory and minimum resource theory (Ganson 1961). Riker's size principle predicts the formation of minimal winning coalitions in these games. Other theories for weighted majority games include the bargaining theory (Komorita and Chertkoff 1973; Kravitz 1986) and the equal excess model (Komorita 1979). The bargaining theory posits that players in a coalition will divide payoffs in a manner midway between equality and proportionality to resources (votes) contributed. The equal excess model is similar but uses the equal excess norm instead of proportionality.

Numerous experiments on coalition bargaining in weighted majority games have tested these and related theories (e.g., Cole et al. 1995; Komorita et al. 1989; Miller and Komorita 1986). Results of these studies generally support the bargaining theory and the equal excess model over the others, although all have deficiencies. Reviews of some experiments in this line appear in Komorita (1984) and Komorita and Kravitz (1983).

**Experiments of Other  $n$ -Person Cooperative Games.** Beyond NPD and weighted majority games, investigators have studied a wide variety of  $n$ -person cooperative games in other forms. The primary objective of the work is to discover which

game-theoretic solution concepts predict most accurately the outcomes of these games.

Numerous studies have investigated cooperative sidepayment games in characteristic function form (e.g., Michener et al. 1986; Murnighan and Roth 1980; Rappaport 1990). Other studies have investigated similar games in strategic form. This work shows that in games with empty core, solution concepts such as the nucleolus and the kernel predict fairly well; in games with a nonempty core, however, the Shapley value is often more accurate. Reviews of parts of this research appear in Kahan and Rapoport (1984), Michener and Potter (1981), and Murnighan (1978).

Other studies have investigated cooperative nonsidepayment games. Some of this research pertains to bargaining models in sequential games of status (Friend et al. 1977). Other research tests various solution concepts (such as the core and the lambda transfer value) in nonsidepayment games in strategic form (McKelvey and Ordeshook 1982; Michener et al. 1985; Michener and Salzer 1989).

Another line of experimentation on cooperative nonsidepayment games is that conducted by political scientists interested in committee games or spatial voting games. These are  $n$ -person voting games in which policies are represented as positions in multidimensional space. For the most part, this research attempts to test predictions from alternative solution concepts (Ferejohn et al. 1980; Ordeshook and Winer 1980). Some of this work has led to new theories, such as the competitive solution (McKelvey and Ordeshook 1983; McKelvey et al. 1978), and to further developments regarding established ones, such as methods for computing the Copeland winner (Grofman et al. 1987). Experimental research on spatial games is reviewed in McKelvey and Ordeshook (1990).

## DYNAMIC GAMES

Although early developments in game theory centered primarily on static games (i.e., games in which interaction among players is single-period or single-play in nature), many subsequent developments have addressed *dynamic games* occurring over time. In a dynamic game, time (or stage) is an important consideration in strategy, and the choices and actions of players at any stage are conditional on the history of prior choices in the game. There

is a growing theoretical literature on various classes of dynamic games, including repeated games, differential games, and evolutionary games. Introductions to the topic of dynamic games appear in Friedman (1990), Fudenberg and Tirole (1991), Owen (1982), and Thomas (1984). The empirical literature on dynamic games is still small relative to that on static games, although experimental studies of repeated games appear increasingly often.

The term *supergame* refers to a sequence of (ordinary) games played by a fixed set of players. One important type of supergame is the *repeated game*, wherein the same constituent game is played at each stage in the sequence. For instance, if some players play a prisoner's dilemma game again and again, they are engaging in a repeated game. At this point in historical time, the dominant paradigm for the study of dynamic strategic behavior is that of repeated games. Certain repeated games are of interest because they allow collectively rational outcomes to result from noncooperative equilibrium strategies. Axelrod (1984) has analyzed the development of cooperation in repeated games. Selten and Stoecker (1983) have used a learning theory approach to model end-game behavior of players in repeated prisoner's dilemma games. Aumann and Maschler (1995) have studied repeated games with incomplete information. A survey of literature on repeated games appears in Mertens and colleagues (1994a,b,c).

Theorists have developed various solution concepts applicable to repeated games and multistage games. Among these are the backward induction process, the subgame perfect equilibrium (Selten 1975), and the Pareto perfect equilibrium (Bernheim et al. 1987). Cronshaw (1997) describes computational techniques for finding all equilibria in infinitely repeated games with discounting and perfect monitoring.

Another class of dynamic games is the *differential game*, played in continuous time. Much of the literature on differential games focuses on the two-person zero-sum case. Some applications of differential games are military, such as pursuit games, where the goal of, say, a pursuing aircraft is to minimize time or distance required to catch an evading aircraft (Hajek 1975). The classic works on differential games include Friedman (1971) and Isaacs (1965). Models of differential games with more than two players are discussed in Leitman

(1974). Other useful works on differential games include Basar and Bernard (1989) and Lewin (1994). The vector-valued maximin for these games is discussed in Zhukovskiy and Salukvadze (1994).

Biologists and economists have used game theoretic concepts to study *evolutionary games*, which are dynamic models of social evolution that explain why certain inherited traits (i.e., behavioral patterns) arise in a human or animal population and remain stable over time. In some evolutionary games (especially those with animal populations), the individuals are modeled as having neither rationality, nor conscience, nor expectations, so strategy selection and equilibrium derive from behavioral phenotypes rather than from rational thought processes. Models of this type often incorporate such phenomena as mutation, acquisition (learning), and the consequences of random perturbations. Theorists have advanced various concepts of evolutionary stability and evolutionarily stable strategies (Amir and Berninghaus 1998; Bomze and Potscher 1989; Gardner et al. 1987; Maynard Smith 1982). Summaries and extensions of work on evolutionary games appear in Bomze (1996), Friedman (1991, 1998), Samuelson (1997), and Weibull (1995).

### INSTITUTIONAL ANALYSIS VIA GAME THEORY

Economists and political scientists have long used game theory in the analysis of social institutions. In work of this type, an analyst specifies an institution (such as a Cournot oligopoly or an approval voting system) in hypothetical or ideal-typical terms and then applies game-theoretic solution concepts to see which payoff allocation(s) may result at equilibrium. Through this approach, an analyst can compare the outcomes of alternative institutional forms with respect to stability, efficiency, and fairness. Broad discussions and reviews of this literature appear in Schotter (1981, 1994), Schotter and Schwodiauer (1980), and Shubik (1982, 1984).

**Economic Institutions.** Von Neumann and Morgenstern (1944) were among the first to explore the role of  $n$ -person game theory in economic analysis. Since that time, economists have analyzed a variety of institutions in game-theoretic terms, including *oligopoly and other imperfect markets*. Markets in which there are only a few sellers

(oligopoly), two sellers (duopoly, a type of oligopoly), one buyer and one seller (bilateral monopoly), and so on, lend themselves to game-theoretic analysis because the payoffs to each player depend on the strategies of the other players.

Economists have modeled oligopolies both as noncooperative games and as cooperative games. Several analyses of oligopolies as noncooperative games show that the standard Chamberlin price-setting strategy is equivalent to a Nash equilibrium in pure strategies (Telser 1972). Beyond that, various analyses have treated oligopoly as a noncooperative multistage game. These analyses have produced generalizations concerning the effects of adjustment speed, demand and cost functions, and incomplete information regarding demand on the dynamic stability of the traditional Cournot solution and other equilibria (Friedman 1977; Radner 1980).

Analyses of collusion among oligopolists usually view this as a cooperative game. These treatments analyze outcomes via such solution concepts as the core or the bargaining set (Kaneko 1978). Reviews and discussions of game-theoretic models of oligopoly appear in deFraja and Delbono (1990), Friedman (1983), Kurz (1985), and Shubik (1984). Game theoretic literature on collusive equilibria is surveyed in Rees (1993).

A second topic of interest to game-theoretic economists is general equilibrium in a *multilateral exchange economy*. An early paper (Arrow and Debreu 1954) modeled a general competitive exchange economy (involving production, exchange, and consumption) as a noncooperative game, and then showed that a generalized Nash equilibrium existed for the model. Other theorists have modeled a multilateral exchange economy as a cooperative  $n$ -person game. Shubik (1959) showed that the core of such a game is identical to Edgeworth's traditional contract curve. More generally, Debreu and Scarf (1963) showed that the core converges to the Walrasian competitive equilibrium. Telser (1996) discussed the core as a tool for finding market-clearing prices in economies with many inputs and outputs. Roth and Sotomayor (1990) discussed the matching problem in labor and marriage markets.

By representing an economy as an  $n$ -person cooperative game, analysts can investigate general equilibrium even in markets that do not fulfill the

neoclassical regularity assumptions (e.g., convexity of preferences, divisibility of commodities, absence of externalities) (Rosenthal 1971; Telser 1972). Beyond that, some economists model general equilibrium in game-theoretic terms because they can introduce and analyze a variety of alternative institutional arrangements. This includes, for instance, models with or without trade, with or without production, with various types of money (commodity money, fiat money, bank money, accounting money), and with various types of financial institutions (shares, central banks, bankruptcy rules, and the like) (Dagan et al. 1997; Dubey and Shubik 1977; Karatzas et al. 1997). Depending on the models used, such cooperative solutions as the core, Shapley value, and nucleolus play an important role in these analyses, as does the Nash noncooperative solution.

A third concern of game-theoretically oriented economists is the analysis of *public goods and services* (e.g., bridges, roads, dams, harbors, libraries, police services, public health services, and the like). Of special relevance is the pricing and cross-subsidization of public utilities; a central issue is how different classes of customers should divide the costs of providing public utilities.

Theorists can model such problems by letting the cost function of a public utility determine the characteristic function of a (cooperative) cost-sharing game. Games of this type are amenable to analysis via such solution concepts as the Shapley value (Loehman and Whinston 1974), the nucleolus (Littlechild and Vaidya 1976; Nakayama and Suzuki 1977), and the core (Littlechild and Thompson 1977), each of which represents alternative cost-sharing criteria. Moulin (1996) compared alternative cost-sharing mechanisms under increasing returns. Other work has used game-theoretic concepts to model negotiation with respect to provision of public goods (Dearden 1998; Schofield 1984a). A survey of experimental research on public goods appears in Ledyard (1995).

**Political Institutions.** Like economists, political scientists have analyzed a variety of institutions in game-theoretic terms. One broad line of work—termed the study of *social choice*—investigates various methods for aggregating individual preferences into collective decisions (Moulin 1994). Of special concern is the stability of outcomes produced by alternative *voting systems* (Hinich and Munger

1997; Nurmi 1987; Ordeshook 1986). Analysts have studied many different voting systems (e.g., majority voting, plurality voting, weighted voting, approval voting, and so on.) Strikingly, this work has demonstrated that in majority voting systems where voters choose among more than two alternatives, the conditions for equilibria (i.e., the conditions that assure a decisive winner) are so restrictive as to render equilibria virtually nonexistent (Fishburn 1973; Riker 1982).

Topics in voting include the detailed analysis of cyclic majority phenomena (generalized Condorcet paradox situations), the analysis of equilibria in novel voting systems such as weighted voting (Banzhaf 1965) and approval voting (i.e., a method of voting wherein voters may endorse as many candidates as they like in multicandidate elections) (Brams and Fishburn 1983), and the development of predictive solution concepts for voting systems that possess no stable equilibrium (Ferejohn and Grether 1982).

A second line of work by game theorists concerns the strategic manipulation of political institutions to gain favorable outcomes. One topic here is the consequences of manipulative *agenda control* in committees (Banks 1990; Plott and Levine 1978). Another topic is the effects of *strategic voting*—that is, voting in which players strive to manipulate the decision by voting for candidates or motions other than their real preferences (Cox and Shugart 1996; Feddersen and Pesendorfer 1998; Niemi and Frank 1982). Analyses by Gibbard (1973) and Satterthwaite (1975) showed that no voting procedure can be completely strategy-proof (in the sense of offering voters no incentive to vote strategically) without violating some more fundamental condition of democratic acceptability. In particular, any voting mechanism that is strategy-proof is also necessarily dictatorial. An important issue here is how to design voting systems with at least some desirable properties that encourage sincere revelation of preferences.

A third broad line of work concerns the *indexing of players' power* in political systems (Owen 1982, ch. 10). Frequently this entails assessing differences in a priori voting strength of members of committees or legislatures. For example, a classic study applied the Shapley–Shubik index of power to the U.S. legislative system and assessed the relative power of the president, senators, and

representatives (Shapley and Shubik 1954). Other measures of power include the Banzhaf–Coleman index (Banzhaf 1965; Dubey and Shapley 1979) and Straffin's probabilistic indices (Straffin 1978).

A fourth game-theoretic topic is *cabinet coalition formation*, especially in the context of European governments. This topic is of interest because political fragmentation can produce instability of cabinet coalitions, which in turn can lead to collapse of entire governments. Work on this problem ties in with that on spatial voting games and weighted majority games, discussed above. Some cabinet coalition models stress policy (or ideological) alignment among members, while other models stress the transfer of value (payoffs) among members. Important theoretical models include DeSwaan (1973), Grofman (1982), Laver and Shepsle (1996), and Schofield (1984b).

#### REFERENCES

- Amir, M., and S. Berninghaus 1998 "Scale Functions in Equilibrium Selection Games." *Journal of Evolutionary Economics* 8(1):1–13.
- Arrow, K. J., and G. Debreu 1954 "Existence of an Equilibrium for a Competitive Economy." *Econometrica* 22:265–290.
- Aumann, R. J. 1961 "The Core of a Cooperative Game Without Sidepayments." *Transactions of the American Mathematical Society* 98:539–552.
- , and J. H. Dreze 1974 "Cooperative Games with Coalition Structures." *International Journal of Game Theory* 3:217–237.
- , and M. Maschler 1964 "The Bargaining Set for Cooperative Games." In M. Dresher, L. S. Shapley, and A. W. Tucker, eds., *Advances in Game Theory* (Annals of Mathematics Studies, No. 52), Princeton, N.J.: Princeton University Press.
- , and M. B. Maschler, with R. E. Stearns 1995 *Repeated Games with Incomplete Information*. Cambridge, Mass.: MIT Press.
- , and R. B. Myerson 1988 "Endogenous Formation of Links Between Players and of Coalitions: An Application of the Shapley Value." In A. E. Roth, ed., *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. New York: Cambridge University Press.
- Axelrod, R. 1984 *The Evolution of Cooperation*. New York: Basic Books.
- Banks, J. S. 1990 "Monopoly Agenda Control and Asymmetric Information." *Quarterly Journal of Economics* 105(2):445–464.

- Banzhaf, J. F., III 1965 "Weighted Voting Doesn't Work: A Mathematical Analysis." *Rutgers Law Review* 19:317-343.
- Basar, T. S., and P. Bernard, (eds.) 1989 *Differential Games and Applications*. New York: Springer-Verlag.
- Bennett, E. 1983 "The Aspiration Approach to Predicting Coalition Formation and Payoff Distribution in Sidepayment Games." *International Journal of Game Theory* 12:1-28.
- Bernheim, B. D., B. Peleg, and M. Whinston 1987 "Coalition-Proof Nash Equilibria. I: Concepts." *Journal of Economic Theory* 42:1-12.
- Bomze, I. M. 1996 "Evolutionary Stability Is Not a Foolish Game." *Central European Journal for Operations Research and Economics* 4(1):25-56.
- , and B. M. Potscher 1989 *Game Theoretic Foundations of Evolutionary Stability*. Berlin: Springer-Verlag.
- Borel, E. (1921-1927) 1953 "'The Theory of Play and Integral Equations with Skew Symmetric Kernels,' 'On Games That Involve Chance and the Skill of the Players,' and 'On Systems of Linear Forms of Skew Symmetric Determinant and the General Theory of Play,'" trans. L. J. Savage. *Econometrica* 21:97-117.
- Brams, S. J., and P. C. Fishburn 1983 *Approval Voting*. Cambridge, Mass.: Birkhauser.
- Charnes, A., J. Rousseau, and L. Seiford 1978 "Complements, Mollifiers, and the Propensity to Disrupt." *International Journal of Game Theory* 7:37-50.
- Cole, S. G., P. R. Nail, and M. Pugh 1995 "Coalition Preference as a Function of Expected Values in a Tetradic Weighted-Majority Game." *Basic and Applied Social Psychology* 16:109-120.
- Colman, A. M. 1982 *Game Theory and Experimental Games*. Oxford: Pergamon Press.
- Cox, J. C., V. L. Smith, and J. M. Walker 1984 "Theory and Behavior of Multiple Unit Discriminative Auctions." *Journal of Finance* 39:983-1010.
- Cox, G. W., and M. S. Shugart 1996 "Strategic Voting under Proportional Representation." *Journal of Law, Economics, and Organization* 12(2):299-324.
- Cronshaw, M. B. 1997 "Algorithms for Finding Repeated Game Equilibria." *Computational Economics* 10(2):139-168.
- Crott, H. W., and W. Albers 1981 "The Equal Division Kernel: An Equity Approach to Coalition Formation and Payoff Distribution in  $n$ -Person Games." *European Journal of Social Psychology* 11(3):285-305.
- Dagan, N., R. Serrano, and O. Volij 1997 "A Noncooperative View of Consistent Bankruptcy Rules." *Games and Economic Behavior* 18(1):55-72.
- Davis, M., and M. Maschler 1965 "The Kernel of a Cooperative Game." *Naval Research Logistics Quarterly* 12:223-259.
- Dawes, R. M. 1980 "Social Dilemmas." *Annual Review of Psychology* 31:169-193.
- Dearden, J. A. 1998 "Serial Cost Sharing of Excludable Public Goods: General Cost Functions." *Economic Theory* 12(1):189-198.
- Debreu, G., and H. E. Scarf 1963 "A Limit Theorem on the Core of an Economy." *International Economic Review* 4:235-246.
- deFraja, G., and F. Delbono 1990 "Game Theoretic Models of Mixed Oligopoly." *Journal of Economic Surveys* 4(1):1-17.
- DeSwaan, A. 1973 *Coalition Theories and Cabinet Formation*. San Francisco: Jossey-Bass.
- Dimand, M. A., and R. W. Dimand 1996 *A History of Game Theory*. New York: Routledge.
- Dubey, P., and L. S. Shapley 1979 "Mathematical Properties of the Banzhaf Power Index." *Mathematics of Operations Research* 4:99-130.
- Dubey, P., and M. Shubik 1977 "Trade and Prices in a Closed Economy with Exogenous Uncertainty, Different Levels of Information, Money and Compound Future Markets." *Econometrica* 45:1657-1680.
- Edgeworth, F. Y. 1881 *Mathematical Psychics*. London: Kegan Paul.
- 1995 *Foundations of Mathematical Economics Series*, vol. 3. San Diego, Calif.: James and Gordon.
- Engelbrecht-Wiggans, R. 1980 "Auctions and Bidding Models: A Survey." *Management Science* 26:119-142.
- Feddersen, T., and W. Pesendorfer 1998 "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts Under Strategic Voting." *American Political Science Review* 92(1):23-35.
- Ferejohn, J. A., M. P. Fiorina, and E. W. Packel 1980 "Nonequilibrium Solutions for Legislative Systems." *Behavioral Science* 25:140-148.
- Ferejohn, J. A., and D. M. Grether 1982 "On the Properties of Stable Decision Procedures." In P. C. Ordeshook and K. A. Shepsle, eds., *Political Equilibrium*. Boston: Kluwer-Nijhoff Publishing.
- Fishburn, P. C. 1973 *The Theory of Social Choice*. Princeton, N.J.: Princeton University Press.
- Flood, M. M. 1952 "Some Experimental Games." Research Memorandum RM-789. Santa Monica, Calif.: The Rand Corporation.
- Fraser, N. M., and K. W. Hipel 1984 *Conflict Analysis: Models and Resolutions*. New York: Elsevier Science.
- Friedman, A. 1971 *Differential Games*. New York: John Wiley.



- Friedman, D. 1991 "Evolutionary Games in Economics." *Econometrica* 59(3):637-666.
- 1998 "On Economic Applications of Evolutionary Game Theory." *Journal of Evolutionary Economics* 8(1):15-43.
- Friedman, J. W. 1977 *Oligopoly and the Theory of Games*. Amsterdam: North-Holland.
- 1983 *Oligopoly Theory*. New York: Cambridge University Press.
- 1990 *Game Theory with Applications to Economics*, 2nd ed. New York: Oxford University Press.
- , and A. C. Hoggatt 1980 "An Experiment in Noncooperative Oligopoly." *Research in Experimental Economics*, vol. 1, supplement 1. Greenwich, Conn.: JAI Press.
- Friend, K. E., J. D. Laing, and R. J. Morrison 1977 "Game-Theoretic Analyses of Coalition Behavior." *Theory and Decision* 8:127-157.
- Fudenberg, D., and J. Tirole 1991 *Game Theory*. Cambridge, Mass.: MIT Press.
- Gamson, W. A. 1961 "A Theory of Coalition Formation." *American Sociological Review* 26:373-382.
- Gardner, R., M. R. Morris, and C. E. Nelson 1987 "Conditional Evolutionarily Stable Strategies." *Animal Behaviour* 35(2):507-517.
- Gately, D. 1974 "Sharing the Gains From Regional Cooperation: A Game Theoretic Application to Planning Investment in Electric Power." *International Economic Review* 15:195-208.
- Gibbard, A. 1973 "Manipulation of Voting Schemes: A General Result." *Econometrica* 41:587-601.
- Gillies, D. B. 1959 "Solutions to General Non-Zero-Sum Games." *Annals of Mathematics Studies* 40:47-85.
- Grofman, B. 1982 "A Dynamic Model of Proto-coalition Formation in Ideological  $N$ -Space." *Behavioral Science* 27:77-90.
- , G. Owen, N. Noviello, and A. Glazer 1987 "Stability and Centrality in Legislative Choice in the Spatial Context." *American Political Science Review* 81:539-552.
- Hajek, O. 1975 *Pursuit Games*. New York: Academic Press.
- Hardin, G. 1968 "The Tragedy of the Commons." *Science* 162:1243-1248.
- Harris, R. J. 1972 "An Interval-Scale Classification System for All  $2 \times 2$  Games." *Behavioral Science* 17:371-383.
- Harsanyi, J. C., and R. Selten 1988 *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.
- Hartwick, J. M., and D. W. K. Yeung 1997 "The Tragedy of the Commons Revisited." *Pacific Economic Review* 2(1):45-62.
- Hinich, M. J., and M. C. Munger 1997 *Analytical Politics*. New York: Cambridge University Press.
- Horowitz, A. D. 1973 "The Competitive Bargaining Set for Cooperative  $n$ -Person Games." *Journal of Mathematical Psychology* 10:265-289. (Erratum: 1974, 11:161).
- Isaacs, R. 1965 *Differential Games*. New York: John Wiley.
- Jones, A. J. 1980 *Game Theory: Mathematical Models of Conflict*. Chichester, U.K.: Ellis Horwood.
- Kagel, J. H. 1995 "Auctions: A Survey of Experimental Research." In J. H. Kagel and A. E. Roth, eds., *Handbook of Experimental Economics*. Princeton, N.J.: Princeton University Press.
- Kahan, J. P., and Am. Rapoport 1984 *Theories of Coalition Formation*. Hillsdale, N.J.: Lawrence Erlbaum.
- Kaneko, M. 1978 "Price Oligopoly as a Cooperative Game." *International Journal of Game Theory* 7:137-150.
- Karatzas, J., M. Shubik, and W. D. Sudderth 1997 "A Strategic Market Game with Secured Lending." *Journal of Mathematical Economics* 28(2):207-247.
- Kollock, P. 1998 "Social Dilemmas: The Anatomy of Cooperation." *Annual Review of Sociology* 24:183-214.
- Komorita, S. S. 1979 "An Equal Excess Model of Coalition Formation." *Behavioral Science* 24:369-381.
- 1984 "Coalition Bargaining." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 18. New York: Academic Press.
- , K. F. Aquino, and A. L. Ellis 1989 "Coalition Bargaining: A Comparison of Theories Based on Allocation Norms and Theories Based on Bargaining Strength." *Social Psychology Quarterly* 52(3):183-196.
- , and J. M. Chertkoff 1973 "A Bargaining Theory of Coalition Formation." *Psychological Review* 80:149-162.
- , and D. A. Kravitz 1983 "Coalition Formation: A Social Psychological Approach." In P. B. Paulus, ed., *Basic Group Processes*. New York: Springer-Verlag.
- , and C. D. Parks 1995 "Interpersonal Relations: Mixed-Motive Interaction." *Annual Review of Psychology* 46:183-207.
- , and C. D. Parks 1999 "Reciprocity and Cooperation in Social Dilemmas: Review and Future Directions." In D. V. Budescu, I. Erev, and R. Zwick, eds., *Games and Human Behavior: Essays in Honor of Amnon Rapoport*. Mahwah, N.J.: Lawrence Erlbaum.
- Kravitz, D. A. 1986 "Extensions of Komorita and Chertkoff's (1973) Bargaining Theory to Simple

- Weighted Majority Games in Which Players Lack Alternative Minimal Winning Coalitions." In R. W. Scholz, ed., *Current Issues in West German Decision Research: Vol. 4, Psychology of Decisions and Conflict*. Frankfurt, Federal Republic of Germany: Peter Lang Publishing.
- Kurz, M. 1985 "Cooperative Oligopoly Equilibrium." *European Economic Review* 27(1):3-24.
- Laffont, J. J. 1997 "Game Theory and Empirical Economics: The Case of Auction Data." *European Economic Review*, 41(1):1-35.
- Laver, M., and K. A. Shepsle 1996 *Making and Breaking Governments: Cabinets and Legislatures in Parliamentary Democracies*. New York: Cambridge University Press.
- Ledyard, J. O. 1995 "Public Goods: A Survey of Experimental Research." In J. H. Kagel and A. E. Roth, eds., *Handbook of Experimental Economics*. Princeton, N.J.: Princeton University Press.
- Leitmann, G. 1974 *Cooperative and Non-Cooperative Many Player Differential Games*. New York: Springer-Verlag.
- Lewin, J. 1994 *Differential Games: Theory and Methods for Solving Game Problems with Singular Surfaces*. New York: Springer-Verlag.
- Liebrand, W. B. G., and D. M. Messick, (eds.) 1996 *Frontiers in Social Dilemmas Research*. New York: Springer.
- , and H. A. M. Wilke, (eds.) 1992 *Social Dilemmas: Theoretical Issues and Research Findings*. Oxford: Pergamon Press.
- Littlechild, S. C., and G. E. Thompson 1977 "Aircraft Landing Fees: A Game Theory Approach." *Bell Journal of Economics* 8:186-207.
- Littlechild, S. C., and K. G. Vaidya 1976 "The Propensity to Disrupt and the Disruption Nucleolus of a Characteristic Function Game." *International Journal of Game Theory* 5:151-161.
- Loehman, E., and A. Whinston 1974 "An Axiomatic Approach to Cost Allocations for Public Investments." *Public Finance Quarterly* 2:236-251.
- Luce, R. D., and H. Raiffa 1957 *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley.
- Maynard Smith, J. 1982 *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridge University Press.
- McAfee, R. P., and J. McMillan 1987. "Auctions and Bidding." *Journal of Economic Literature* 25:699-738.
- McClintock, C.G., and W. B. G. Liebrand 1988. "Role of Interdependence Structure, Individual Value Orientation, and Another's Strategy in Decision Making: A Transformational Analysis." *Journal of Personality and Social Psychology* 55:396-409.
- McKelvey, R. D., and P. C. Ordeshook 1982 "An Experimental Test of Solution Theories for Cooperative Games in Normal Form." In P. C. Ordeshook and K. A. Shepsle, eds., *Political Equilibrium*. Boston: Kluwer-Nijhoff Publishing.
- 1983 "Some Experimental Results That Fail to Support the Competitive Solution." *Public Choice* 40(3):281-291.
- 1990 "A Decade of Experimental Research on Spatial Models of Elections and Committees." In J. M. Enelow and M. J. Hinich, eds., *Advances in the Spatial Theory of Voting*. New York: Cambridge University Press.
- , and M. D. Winer 1978 "The Competitive Solution for  $n$ -Person Games without Transferable Utility, with an Application to Committee Games." *American Political Science Review* 72:599-615.
- Mertens, J.-F., S. Sorin, and S. Zamir 1994a "Repeated Games Part A: Background Material" (CORE Discussion Paper No. 9420). Universite Catholique de Louvain.
- 1994b "Repeated Games Part B: The Central Results" (CORE Discussion Paper No. 9421). Universite Catholique de Louvain.
- 1994c "Repeated Games Part C: Further Development" (CORE Discussion Paper No. 9423). Universite Catholique de Louvain.
- Messick, D. M., and M. D. Brewer 1983 "Solving Social Dilemmas: A Review." In L. Wheeler and P. Shaver, eds., *Review of Personality and Social Psychology*, vol. 4. Beverly Hills, Calif.: Sage.
- Michener, H. A., and W. T. Au. 1994 "A Probabilistic Theory of Coalition Formation in  $n$ -Person Sidepayment Games." *Journal of Mathematical Sociology* 19:165-188.
- Michener, H. A., D. C. Dettman, J. M. Ekman, and Y. C. Choi 1985 "A Comparison of the Alpha- and Beta-Characteristic Functions in Cooperative Non-Sidepayment  $n$ -Person Games." *Journal of Mathematical Sociology* 11:307-330.
- Michener, H. A, G. B. Macheel, C. G. Depies, and C. A. Bowen 1986 "Mollifier Representation in Non-Constant-Sum Games: An Experimental Test." *Journal of Conflict Resolution* 30:361-382.
- Michener, H. A., and D. J. Myers 1998 "Probabilistic Coalition Structure Theories: An Empirical Comparison in Four-Person Superadditive Sidepayment Games." *Journal of Conflict Resolution* 42(6):853-883.
- Michener, H. A., and K. Potter 1981 "Generalizability of Tests in  $n$ -Person Sidepayment Games." *Journal of Conflict Resolution* 25:733-749.

- Michener, H. A., and M. S. Salzer 1989 "Comparative Accuracy of Value Solutions in Non-Sidepayment Games with Empty Core." *Theory and Decision* 26:205–233.
- Miller, C. E., and S. S. Komorita 1986 "Changes in Outcomes in Coalition Bargaining." *Journal of Personality and Social Psychology* 51(4):721–729.
- Montmort, P. R. de 1713 *Essay d'Analyse sur les Jeux de Hazard*, 2nd ed. Paris: Quillau.
- Moulin, H. 1994 "Social Choice." In R. J. Aumann and S. Hart, eds., *Handbook of Game Theory with Economic Applications*, vol. 2. New York: Elsevier, North-Holland.
- 1996 "Cost Sharing under Increasing Returns: A Comparison of Simple Mechanisms." *Games and Economic Behavior* 13(2):225–251.
- , and A. Watts 1997 "Two Versions of the Tragedy of the Commons." *Economic Design* 2(4):399–421.
- Murnighan, J. K. 1978 "Models of Coalition Behavior: Game Theoretic, Social Psychological, and Political Perspectives." *Psychological Bulletin* 85:1130–1153.
- , and A. E. Roth 1980 "Effects of Group Size and Communication Availability on Coalition Bargaining in a Veto Game." *Journal of Personality and Social Psychology* 39:92–103.
- Myerson, R. B. 1977 "Graphs and Cooperation in Games." *Mathematics of Operations Research* 2:225–229.
- 1991 *Game Theory: Analysis of Conflict*. Cambridge, Mass.: Harvard University Press.
- Nakayama, M., and M. Suzuki 1977 "The Cost Assignment of Cooperative Water Resource Development: A Game Theoretical Approach." In R. Henn and O. Moeschlin, eds., *Mathematical Economics and Game Theory*. Berlin: Springer-Verlag.
- Nash, J. F., Jr. 1951 "Non-Cooperative Games." *Annals of Mathematics* 54:286–295.
- Niemi, R. G., and A. Q. Frank 1982 "Sophisticated Voting Under the Plurality Procedure." In P. C. Ordeshook and K. A. Shepsle, eds., *Political Equilibrium*. Boston: Kluwer-Nijhoff Publishing.
- Nurmi, H. 1987 *Comparing Voting Systems*. Dordrecht: D. Reidel.
- Ordeshook, P. C. 1986 *Game Theory and Political Theory*. New York: Cambridge University Press.
- , and M. Winer 1980 "Coalitions and Spatial Policy Outcomes in Parliamentary Systems: Some Experimental Results." *American Journal of Political Science* 24:730–752.
- Oskamp, S. 1971 "Effects of Programmed Strategies on Cooperation in the Prisoner's Dilemma and Other Mixed Motive Games." *Journal of Conflict Resolution* 15:225–259.
- Owen, G. 1982 *Game Theory*, 2nd ed. New York: Academic Press.
- Plott, C. R. 1982 "Industrial Organization Theory and Experimental Economics." *Journal of Economic Literature* 20:1485–1527.
- , and M. E. Levine 1978 "A Model of Agenda Influence on Committee Decisions." *American Economic Review* 68:146–160.
- , and S. Sunder 1982 "Efficiency of Experimental Securities Markets with Insider Information." *Journal of Political Economy* 90:663–698.
- Pruitt, D. G., and M. J. Kimmel 1977 "Twenty Years of Experimental Gaming: Critique, Synthesis, and Suggestions for the Future." *Annual Review of Psychology* 28:363–392.
- Radner, R. 1980 "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives." *Journal of Economic Theory* 22:136–154.
- Rapoport, Amnon 1988 "Experiments with N-Person Social Traps II: Tragedy of the Commons." *Journal of Conflict Resolution* 32(3):473–488.
- 1990 *Experimental Studies of Interactive Decisions*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- , and J. P. Kahan 1982 "The Power of a Coalition and Payoff Disbursement in Three-Person Negotiable Conflicts." *Journal of Mathematical Sociology* 8:193–224.
- , M. Guyer, and D. G. Gordon 1976 *The 2 × 2 Game*. Ann Arbor: University of Michigan Press.
- Rees, R. 1993 "Tacit Collusion." *Oxford Review of Economic Policy* 9(2):27–40.
- Riker, W. 1982 "Implication from the Disequilibrium of Majority Rule for the Study of Institutions." In P. C. Ordeshook and K. A. Shepsle, eds., *Political Equilibrium*. Boston: Kluwer-Nijhoff Publishing.
- Romp, G. 1997 *Game Theory: Introduction and Applications*. New York: Oxford University Press.
- Rosenthal, R. W. 1971 "External Economies and Cores." *Journal of Economic Theory* 3:182–188.
- Roth, A. E. 1995 "Bargaining Experiments." In J. H. Kagel and A. E. Roth, eds., *Handbook of Experimental Economics*. Princeton, N.J.: Princeton University Press.
- , and M. A. O. Sotomayor 1990 *Two-Sided Matching*. New York: Cambridge University Press.
- Samuelson, L. 1997 *Evolutionary Games and Equilibrium Selection*. Cambridge, Mass.: MIT Press.

- Satterthwaite, M. A. 1975 "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10:187-217.
- Schmeidler, D. 1969 "The Nucleolus of a Characteristic Function Game." *SIAM Journal of Applied Mathematics* 17:1163-1170.
- Schofield, N. 1984a "Bargaining over Public Goods." In M. J. Holler, ed., *Coalitions and Collective Action*. Wurzburg, Germany: Physica-Verlag.
- 1984b "Political Fragmentation and the Stability of Coalition Governments in Western Europe." In M. J. Holler, ed., *Coalitions and Collective Action*. Wurzburg, Germany: Physica-Verlag.
- Schotter, A. 1981 *The Economic Theory of Social Institutions*. New York: Cambridge University Press.
- 1994 *Microeconomics: A Modern Approach*. New York: HarperCollins.
- , and G. Schwodiauer 1980 "Economics and the Theory of Games: A Survey." *Journal of Economic Literature* 18:479-527.
- Selten, R. 1975 "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *International Journal of Game Theory* 4:25-55.
- , and R. Stoecker 1983 *End Behavior In Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach*. Bielefeld, Germany: Institut für Mathematische Wirtschaftsforschung an der Universität Bielefeld.
- Sengupta, A., and K. Sengupta 1994 "Viable Proposals." *International Economic Review* 35(2):347-359.
- Shapley, L. S. 1953 "A Value for  $n$ -Person Games." In H. W. Kuhn and A. W. Tucker, eds., *Contributions to the Theory of Games, II* (Annals of Mathematics Studies, No. 28). Princeton, N.J.: Princeton University Press.
- , and M. Shubik 1954 "A Method of Evaluating the Distribution of Power in a Committee System." *American Political Science Review* 48:787-792.
- Shubik, M. 1959 "Edgeworth Market Games." In A. W. Tucker and R. D. Luce, eds., *Contributions to the Theory of Games*, vol. 4 (Annals of Mathematics Studies, No. 40). Princeton, N.J.: Princeton University Press.
- 1982 *Game Theory in the Social Sciences*. Cambridge, Mass.: MIT Press.
- 1984 *A Game-Theoretic Approach to Political Economy*. Cambridge, Mass.: MIT Press.
- Smith, V. L. 1982 "Microeconomic Systems as an Experimental Science." *American Economic Review* 72:923-955.
- , A. W. Williams, W. K. Bratton, and M. G. Vannoni 1982 "Competitive Market Institutions: Double Auctions vs. Sealed Bid-Offer Auctions." *American Economic Review* 72:58-77.
- Spinetto, R. 1974 "The Geometry of Solution Concepts for  $n$ -Person Cooperative Games." *Management Science* 20:1292-1299.
- Straffin, P. D., Jr. 1978 "Probability Models for Power Indices." In P. C. Ordeshook, ed., *Game Theory and Political Science*. New York: New York University Press.
- Szep, J., and F. Forgo 1985 *Introduction to the Theory of Games*. Dordrecht, Holland: D. Reidel.
- Telser, L. 1972 *Competition, Collusion, and Game Theory*. Chicago: Aldine-Atherton.
- 1996 "Competition and the Core." *Journal of Political Economy* 104(1):85-107.
- Thomas, L. C. 1984 *Games, Theory and Applications*. Chichester, U.K.: Halsted Press.
- van Damme, E. 1987 *Stability and Perfection of Nash Equilibria*. Berlin: Springer-Verlag.
- Van Lange, P. A., and W. B. Liebrand 1991 "The Influence of Other's Morality and Own Social Value Orientation on Cooperation in the Netherlands and the U.S.A." *International Journal of Psychology* 26(4):429-499.
- von Neumann, J. 1928 "Zur Theorie der Gesellschaftsspiele." *Mathematische Annalen* 100:295-320.
- , and O. Morgenstern 1944 *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.
- Weibull, J. W. 1995 *Evolutionary Game Theory*. Cambridge, Mass.: MIT Press.
- Weintraub, E. R., (ed.) 1992 *Toward a History of Game Theory*. Durham, N.C.: Duke University Press.
- Zermelo, E. 1913 "Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels." In E. W. Hobson and A. E. H. Love, eds., *Proceedings of the Fifth International Congress of Mathematicians*, vol. 2. Cambridge, U.K.: Cambridge University Press.
- Zhukovskiy, V. I., and M. E. Salukvadze 1994 *The Vector-Valued Maximin*. Boston: Academic Press.

H. ANDREW MICHENER

## GANGS

See Criminal and Delinquent Subcultures; Juvenile Delinquency.

## GENDER

Gender, race, ethnicity, and social class are the most commonly used categories in sociology. They represent the major social statuses that determine the life chances of individuals in heterogeneous societies, and together they form a hierarchy of access to property, power, and prestige.

Gender is the division of people into two categories, “men” and “women.” Through interaction with caretakers, socialization in childhood, peer pressure in adolescence, and gendered work and family roles, women and men are socially constructed to be different in behavior, attitudes, and emotions. The gendered social order is based on and maintains these differences.

In sociology, the main ordering principles of social life are called *institutions*. Gender is a social institution as encompassing as the four main institutions of traditional sociology—family, economy, religion, and symbolic language. Like these institutions, gender structures social life, patterns social roles, and provides individuals with identities and values. And just as the institutions of family, economy, religion, and language are intertwined and affect each other reciprocally, as a social institution, gender pervades kinship and family life, work roles and organizations, the rules of most religions, and the symbolism and meanings of language and other cultural representations of human life. The outcome is a gendered social order.

The source of gendered social orders lies in the evolution of human societies and their diversity in history. The gendered division of work has shifted with changing means of producing food and other goods, which in turn modifies patterns of child care and family structures. Gendered power imbalances, which are usually based on the ability to amass and distribute material resources, change with rules about property ownership and inheritance. Men’s domination of women has not been the same throughout time and place; rather, it varies with political, economic, and family structures, and is differently justified by religions and laws. As an underlying principle of how people are categorized and valued, gender is differently constructed throughout the world and has been throughout history. In societies with other major social divisions, such as race, ethnicity, religion,

and social class, gender is intricately intertwined with these other statuses.

As pervasive as gender is, it is important to remember that it is constructed and maintained through daily interaction and therefore can be resisted, reformed, and even rebelled against. The social construction perspective argues that people create their social realities and identities, including their gender, through their actions with others—their families, friends, colleagues. It also argues that their actions are hemmed in by the general rules of social life, by their culture’s expectations, their workplace’s and occupation’s norms, and their government’s laws. These social restraints are amenable to change—but not easily.

Gender is deeply rooted in every aspect of social life and social organization in Western-influenced societies. The Western world is a very gendered world, consisting of only two legal categories—“men” and “women.” Despite the variety of playful and serious attempts at blurring gender boundaries with androgynous dress and desegregating gender-typed jobs, third genders and gender neutrality are rare in Western societies. Those who cross gender boundaries by passing as a member of the opposite gender, or by sex-change surgery, want to be taken as a “normal” man or woman.

Although it is almost impossible to be anything but a “woman” or a “man,” a “girl” or a “boy” in Western societies, this does not mean that one cannot have three, four, five, or more socially recognized genders—there are societies that have at least three. Not all societies base gender categories on male and female bodies—Native Americans, for example, have biological males whose gender status is that of women. Some African societies have females with the gender status of sons or husbands. Others use age categories as organizing principles, not gender statuses. Even in Western societies, where there are only two genders, we can think about restructuring families and workplaces so they are not as rigidly gendered as they are today.

### WHY GENDER AND NOT SEX

Gender was first conceptualized as distinct from sex in order to highlight the social and cultural processes that constructed different social roles

for females and males and that prescribed sex-appropriate behavior, demeanor, personality characteristics, and dress. However, *sex* and *gender* were often conflated and interchanged, to the extent that this early usage was called “sex roles” theory. More recently, gender has been conceptually separated from sex and also from sexuality.

Understanding gender practices and structures is easier if what is usually conflated as sex/gender or sex/sexuality/gender is split into three conceptually distinct categories—sex (or biology, physiology), sexuality (desire, sexual preference, sexual orientation, sexual behavior), and gender (social status, position in the social order, personal identity). Each is socially constructed but in different ways. Gender is an overarching category—a major social status that organizes almost all areas of social life. Therefore bodies and sexuality are gendered—biology and sexuality, in contrast, do not add up to gender.

Conceptually separating sex and gender makes it easier to explain how female and male bodies are socially constructed to be feminine and masculine through sports and in popular culture. In medicine, separating sex from gender helps to pinpoint how much of the differences in longevity and propensity to different illnesses is due to biology and how much to socially induced behavior, such as alcohol and drug abuse, which is higher among men than women. The outcome is a greater number of recorded illnesses but longer life expectancy for women of all races, ethnicities, and social classes when compared to men with the same social characteristics. This phenomenon is known in social epidemiology as “women get sicker, but men die quicker.”

### SOCIAL CONSTRUCTION OF GENDER

As with any other aspect of social life, gender is shaped by an individual’s genetic heritage, physical body, and physiological development. Socially, however, gendering begins as soon as the sex of the fetus is identified. At birth, infants are placed in one of two sex categories, based on the appearance of the genitalia. In cases of ambiguity, since Western societies do not have a third gender for hermaphrodites as some cultures do, the genitalia are now “clarified” surgically, so that the child can be categorized as a boy or a girl. Gendering then

takes place through interaction with parents and other family members, teachers, and peers (“significant others”). Through socialization and gendered personality development, the child develops a gendered identity that, in most cases, reproduces the values, attitudes, and behavior that the child’s social milieu deems appropriate for a girl or a boy.

Borrowing from Freudian psychoanalytic theory, Nancy Chodorow developed an influential argument for the gendering of personalities in the two-parent, heterogendered nuclear family. Because women are the primary parents, infants bond with them. Boys have to separate from their mothers and identify with their fathers in order to establish their masculinity. They thus develop strong ego boundaries and a capacity for the independent action, objectivity, and rational thinking so valued in Western culture. Women are a threat to their independence and masculine sexuality because they remind men of their dependence on their mothers. However, men need women for the emotional sustenance and intimacy they rarely give each other. Their ambivalence toward women comes out in heterosexual love–hate relationships and in misogynistic depictions of women in popular culture and in novels, plays, and operas.

Girls continue to identify with their mothers, and so they grow up with fluid ego boundaries that make them sensitive, empathic, and emotional. It is these qualities that make them potentially good mothers and keep them open to men’s emotional needs. But because the men in their lives have developed personalities that make them emotionally guarded, women want to have children to bond with. Thus, psychological gendering of children is continually reproduced.

To develop nurturing capabilities in men and to break the cycle of the reproduction of gendered personality structures would, according to this theory, take fully shared parenting. There is little data on whether the same psychic processes produce similarly gendered personalities in single-parent families, in households where both parents are the same gender, or in differently structured families in non-Western cultures.

Children are also gendered at school, in the classroom, where boys and girls are often treated differently by teachers. Boys are encouraged to develop their math abilities and science interests; girls are steered toward the humanities and social

sciences. The result is that women students in the United States outnumber men students in college, but only in the liberal arts; in science programs, men still outnumber women. Men also predominate in enrollment in the elite colleges, which prepare for high-level careers in finance, the professions, and government.

This data on gender imbalance, however, when broken down by race, ethnicity, and social class, is more complex. In the United States, upper- and middle-class boys are pushed ahead of girls in school and do better on standardized tests, although girls of all social statuses get better grades than boys. African-American, Hispanic, and white working-class and poor boys do particularly badly, partly because of a peer culture that denigrates “book learning” and rewards defiance and risk taking. In the context of an unresponsive educational structure, discouraged teachers, and crowded, poorly maintained school buildings, the pedagogical needs of marginal students do not get enough attention; they are also much more likely to be treated as discipline problems.

Children’s peer culture is another site of gender construction. On playgrounds, girls and boys divide up into separate groups whose borders are defended against opposite-gender intruders. Within the group, girls tend to be more cooperative and play people-based games. Boys tend to play rule-based games that are competitive. These tendencies have been observed in Western societies; anthropological data about children’s socialization shows different patterns of gendering. Everywhere, children’s gender socialization is closely attuned to expected adult behavior.

### GENDER AND WORK

Whether they have fully internalized Western society’s gender binarism and social construction of gender differences or rebelled against gender typing, adults encounter a gendered work world. The workplaces in industrialized societies are either gender-segregated or composed of all one gender. During the 1970s and 1980s, decades in which women were thought to have made inroads into many occupations previously dominated by men in the United States, about 6 percent of occupations saw an increase of women workers that was significantly greater than the increase of women

overall in the paid labor force during that period. Rather than desegregating occupations, most of the new women workers went into occupations where most of the employees were women, and those who went into occupations where the employees were predominantly men soon found that their coworkers became predominantly women. Some U.S. occupations that went from having a predominance of workers who were men to mostly women workers were personnel, training, and labor relations specialists; computer operators; and insurance adjusters, examiners, and investigators. These occupations had *resegregated*. When women and men work in nontraditional occupations, gender typing is often maintained symbolically, as when policewomen view their work as social work and men nurses emphasize the technical and physical strength aspects of what they do.

The processes that sort women and men of different racial and ethnic groups into different types of work include a matching of ranked workers and jobs, or *queues* of workers and jobs. Workers are ranked by employers from their first picks to their last. Jobs are ranked by workers similarly. Lower-ranked workers get the chance to move into better jobs than they have held in the past when these jobs are abandoned by favored workers or there are too few of these workers to go around, such as in wartime. The process works the other way, too; when there are too few of the best jobs for the preferred workers, as in a recession, only the best qualified or experienced among them will be hired; those with fewer credentials and less seniority move down the queue, bumping out lower-ranked workers. When workers are moving up, the most preferred on the basis of race, ethnicity, and gender usually get the better jobs. If gender segregation is so rigid that men will not apply or be hired for “women’s work,” when manufacturing jobs decline or are taken elsewhere, women in service, sales, and clerical work may continue to work as men’s unemployment rates rise.

Workers rank jobs on the basis of payoff for education and experience in salaries and also in fringe benefits, prestige, autonomy, security, and chances of promotion. For some workers, having any job may be an improvement over economic dependency. Employers’ preferences for workers, however, are not so uniform. Some will rank gender and race above qualifications; others will choose the most highly qualified of the preferred race and

gender and then go down the line, looking for the most qualified each time. Another variable employers factor in is the going pay scale for the workers they want; they may have to settle for less preferred workers to see more of a profit or sacrifice some profits to avoid protests from highly paid entrenched workers.

Although worker demographics, industry growth, and employer preferences produce changes in occupational gender composition, the main factor that redistributes workers of different races and genders is change in the structure of the work process and in the quality of particular jobs within occupations, which can be manipulated by employers. That is, jobs can be automated and deskilled or made part time or home based to justify reducing labor costs, with a few better-paid workers retained in supervisory positions.

During shifts of labor queues up and down the job ladder, the potential for conflict between women and men as well as between members of dominant and subordinated racial and ethnic groups is high. Dominant men want to perpetuate the work conditions that justify their high pay; employers who want to reduce their labor costs degrade the work process so they can hire cheaper labor, and then these new workers are accused of depressing the job's qualifications and skills. Gender segregation of jobs is historically the way employers have kept their men workers satisfied, while expanding the number of cheaper women workers. Such job divisions undercut unions that want to organize women and demand the same pay for them as similarly situated men workers. In a growing or stable job market, dominant men are much less resistant to incoming new types of workers, since they do not see them as competition. In those cases, the job may come closer to being integrated along lines of gender and race.

Occupational gender segregation does not result in separate but equal jobs. Rather, women's work tends to be lower in pay, prestige, and fringe benefits, such as health insurance. Workers themselves rate jobs where most of the employees are women as inferior to jobs where most of the employees are men. The criteria are number and flexibility of hours, earnings, educational requirements, on-the-job training, having a union contract, extent of supervision and place in the hierarchy, repetitiveness, risk of job loss, and being a

government employee. Workers rate women's jobs as better in vacation days and not getting dirty at work, even though changing bedpans is as "dirty" as changing oil pans. If wages were used to compensate for unattractive nonmonetary job characteristics, women's jobs would have to pay four times as much as men's jobs for workers to rate them equally. Nor is there a trade-off of pay for compatibility with child care—most full-time jobs held by mothers are incompatible with parenting demands; flexibility of schedules and control and timing of work-related tasks are the prerogatives of men managers, not their women secretaries.

The best-paid jobs are shaped on an ideal, dominant man's career—long-term, continuous work in the same occupation, with steady pay raises and a pension at retirement. Men's gender status is an advantage to them as workers; they are expected to earn more money when they marry and when they have children, so employers tend to view them as better workers than women. Women workers are felt to be entitled only to supplementary wages, whether they are married or single, because they are not considered legitimate workers but primarily wives and mothers. In actuality, research has shown that married women with children work harder and are more productive than married men with children.

The structured patterns of opportunities and access far override most individual employers' tastes or individual workers' motivations, ambitions, personal desires, and material needs. By the 1970s in the United States, adolescent girls were considerably less likely than in previous years to plan on entering an occupation in which most of the workers were women, especially if they lived in a woman-headed household. But they continued to value working with people, helping others, using their abilities, and being creative; boys wanted jobs with status, high earnings, freedom from supervision, and leadership potential. The jobs women are likely to end up in are more gender typed and less fulfilling than their occupational aspirations, but ambitious and hard-working men can often reach their early goals.

Women of all educational levels and men disadvantaged because of race, immigrant status, lack of education, or outmoded job skills are profitable workers because they tend to receive low wages; they also get promoted less frequently and



therefore receive fewer raises. Many work part time and get no benefits. They can be paid little because the pool of such workers is larger than that of privileged men workers. The size and social characteristics of the pool of low-waged workers are affected by state policies encouraging or discouraging the employment of women, the influx of immigrants, and the flight of capital from one area of a country to another or offshore.

Other processes that segregate and stratify occupations are *segmentation* and *ghettoization*. Segmented occupations are horizontally or vertically divided into sectors with different educational or credential requirements for hiring, different promotion ladders, different work assignments, and different pay scales. Typically, these segments are gendered and frequently also exhibit racial and ethnic clustering. However, occupations in which almost all the workers are of one gender can also be segmented. For instance, in the United States, doctors and nurses are gender-segregated segments in hospitals. Physicians are segmented between those in primary care and those who are hospital-based specialists, who have more prestige and power and higher incomes. Women physicians are often found in primary care. Nursing is also segmented into registered nurses, licensed practical nurses, nurses' aides, and home health workers. Nurses are virtually all women, but the segments are racially differentiated: The majority of registered nurses tend to be white or Asian-American; most of the lower-paid health workers tend to be black or Hispanic. Men who go into nursing tend to specialize in the more lucrative specializations and become administrators.

Segmentation is legitimized by bureaucratic rules or legal requirements for qualifying credentials, but ghettoization separates the lower-paid "women's" jobs from the better-paid "men's" jobs within an occupation through informal gender typing. What is dubbed "women's work" or "men's work" has a sense of normality and naturalness, an almost moral quality, even though the justification for such typing is usually an after-the-fact rationalization. The assumption is that the skills, competence, strength, and other qualities needed to do a job are tied up with masculinity and femininity, but gendered identities as workers are constructed in the gendered organization of the workplace and reinforced in training and organizational sociability, such as company golf games and sports teams.

Within gender-typed occupations, jobs or specialties may be gender typed in the opposite direction. For example, the majority of physicians are men in the United States and women in Russia, but the same specialties are seen as appropriate for one gender—pediatrics for women and neurosurgery for men. In both countries, neurosurgery pays better and has more prestige than pediatrics.

Both structural segmentation and gender typing that puts some jobs into a low-wage ghetto have the same results. They limit the extent of competition for the better positions, make it easier for privileged workers to justify their advantageous salary scales, and create a group of workers whose lack of credentials or requisite skills legitimate their lower pay. Credentials and skills, however, as well as experience, are manipulated or circumvented to favor workers with certain social characteristics, as when men with less lower-rank experience in women's jobs are hired as supervisors. In addition, femaleness and maleness are stereotypically linked to certain capabilities, such as finger dexterity and physical strength; gender then becomes the discriminant criterion for hiring, not what potential employees can actually do with their hands, backs, and heads.

Promotion ladders are also gender segregated. Women and men who are not of the dominant racial or ethnic group tend not to rise to the top in their work organization, unless practically all the workers are women or men of the same racial or ethnic group. White men tend to dominate positions of authority whether or not they are numerically predominant. This pattern is known as the *glass ceiling*—the lid on women's rise to the top of their work organizations. In occupations where the majority of the workers are women, positions of authority tend to be held by men—elementary school teachers are predominantly women in the United States, but principals and superintendents are predominantly men. That is, token men in a woman's occupation tend to be promoted faster than the women workers. This parallel phenomenon has been dubbed the *glass escalator*.

These pervasive patterns of occupational segregation and stratification are the result of deliberate actions and also inaction on the part of governments, owners and managers, and organized groups of workers—and change has to come from the same sources. Since gender segregation involves

occupations and professions, job titles, and specific work sites, integration has to involve more than simply increasing the numbers of women. True occupational gender equality would mean that women and men would have the same opportunities to obtain professional credentials and occupational training, and would be distributed in the same proportions as they are in the paid work force across workplaces, job titles, occupations, and hierarchical positions. Instead, in most industrialized countries, women are overrepresented in clerical and service jobs, low-prestige professional and technical work, and sales. In developing countries, and in areas of industrialized countries where there are concentrations of poor people and recent immigrants, women tend to be concentrated in labor-intensive factory work, agriculture, and the informal (off-the-books) economy.

This gendered organization of paid labor dovetails with the gendered organization of domestic labor. Low pay, uninteresting jobs, and the glass ceiling encourage single women to marry and married women to devote energy and attention to child rearing and domestic work. The job market encourages women to be a reserve army of labor—available for full-time work in times of scarce labor, but fired or put on part-time schedules when there is less work. Better job opportunities are offered to men of the dominant racial and ethnic groups to encourage them to give their all to the job. Employers (mostly men) benefit from women's cheap labor and men's need to earn more to support a family; men who live with women benefit from women's unpaid labor at home. Highly educated and professional women are caught in the structural conflicts of these two forms of labor; in order to compete with the men of their status, they have to hire "wives"—other women to do their domestic work. This pool of paid domestic labor historically is made up of the least advantaged women—native poor and recent immigrants of a variety of racial and ethnic groups.

### GENDER INEQUALITY

Gender inequality takes many different forms, depending on the economic structure and social organization of a particular society and on the culture of any particular group within that society. Although we speak of *gender* inequality, it is usually women who are disadvantaged when compared to

similarly situated men. In the job market, women often receive lower pay for the same or comparable work and are frequently blocked in their chances for advancement, especially to top positions. There is usually an imbalance in the amount of housework and child care a wife does compared to her husband, even when both spend the same amount of time in waged work outside the home. When women professionals are matched with men of comparable productiveness, men get greater recognition for their work and move up career ladders faster. On an overall basis, work most often done by women, such as teaching small children and nursing, is paid less than work most often done by men, such as computer programming and engineering. Gender inequality also takes the form of girls getting less education than boys of the same social class. It often means an unequal distribution of health care services between women and men, and research priorities that focus on diseases men are more likely to get than women.

Gender inequality takes even more oppressive and exploitative forms. Throughout the world, women are vulnerable to beatings, rape, and murder—often by their husbands or boyfriends, and especially when they try to leave an abusive relationship. The bodies of girls and women are used in sex work—pornography and prostitution. They undergo cosmetic surgery and are on display in movies, television, and advertising in Western cultures. In other cultures, their genitals are mutilated and their bodies are covered from head to toe in the name of chastity. They may be forced to bear children they do not want or have abortions or be sterilized against their will. In countries with overpopulation, infant girls are much more often abandoned in orphanages than infant boys. In cultural groups that value boys over girls, if the sex of the fetus can be determined, it is girls who are aborted.

Gender inequality can also disadvantage men. In many countries, only men serve in the armed forces, and in most countries, only men are sent into direct combat. It is mostly men who do the more dangerous work, such as firefighting and policing. Although women have fought in wars and are entering police forces and fire departments, the gender arrangements of most societies assume that women will do the work of bearing and caring for children, while men do the work of protecting them and supporting them economically.

Most women in industrial and postindustrial societies do not spend their lives having and caring for babies, and most women throughout the world do paid and unpaid work to supply their families with food, clothing, and shelter, even while they are taking care of children. The modern forms of gender inequality are not a complementary exchange of responsibilities, but a social system within which women are exploitable. In a succinct summary of gender inequality, it was estimated by a United Nations report in 1980 that women do two-thirds of the world's work, receive 10 percent of the world's income, and own 1 percent of the world's property.

The major social and cultural institutions support this system of gender inequality. Religions legitimate the social arrangements that produce it, justifying them as right and proper. Laws support the status quo and also often make it impossible to redress the outcomes—to prosecute husbands for beating their wives, or boyfriends for raping their girlfriends. In the arts, women's productions are so often ignored that they are virtually invisible, which led Virginia Woolf to conclude that *Anon* must have been a woman. Much scientific research assumes that differences between women and men are genetic or hormonal and looks for data to support these beliefs, ignoring findings that show gender overlaps or input from the social environment. In the social sciences, gender is entered into research designs only as a binary, erasing the effects of racial, social class, and ethnic variations.

Except for the Scandinavian countries, which have the greatest participation of women in government and the most gender-equal laws and state policies, most governments are run by socially dominant men, and their policies reflect their interests. In every period of change, including those of revolutionary upheaval, men's interests, not women's, have prevailed, and many men, but few women, have benefited from progressive social policies. Equality and justice for all usually means for men only. Women have never had their revolution because the structure of gender as a social institution has never been seriously challenged. Therefore, all men benefit from the "patriarchal dividend"—women's unpaid work maintaining homes and bringing up children; women's low-paid work servicing hospitals, schools, and myriad other workplaces.

Gender inequality is deeply ingrained in the structure of Western, industrialized societies. It is built into the organization of marriage and families, work and the economy, politics, religion, sports, the arts and other cultural productions, and the very language we speak. Making women and men equal, therefore, necessitates social, not individual, solutions.

### CHANGING GENDER

Changing a gendered society entails structural and institutional change. Attitudes and values must change, too, but these are often altered when social policies and practices shift. Which changes have occurred since the beginning of the feminist movement of the early 1970s and which have not? What kind of programs target institutions and social structures? What is still needed for gender equality?

Affirmative action and comparable worth pay scales were two efforts to effect structural change—one to desegregate occupations and the other to distribute economic rewards for work on a gender-neutral basis. Affirmative action (hiring women in occupations dominated by men and men for work usually done by women) was widely implemented in the United States and did desegregate some occupations, but without continuous effort, gender segregation reestablishes itself as jobs and work organizations change. Another effort to establish gender-neutral work policies was comparable worth pay scales—assessing the characteristics of the job and paying on the basis of type of work done, not on who does the work. These programs were not widely implemented; women's work continues to be paid less than men's even when a man does the work. Women have entered the professions, especially medicine and the law, in large numbers and have moved up career ladders, but in most large-scale corporations and professional organizations, the top positions of authority are still held by men.

Sexual harassment guidelines have been another effective effort at changing thinking about acceptable behavior, and again, while the results are imperfect (and the guidelines are being used in ways that do not empower women), thinking about what was "normal" behavior in the workplace did change drastically.

In Europe, but not in the United States, subsidized parental leave for either parent and child care for every mother has changed mothering from a full-time occupation to something that can be combined with paid work out of the home without a constant struggle. The Scandinavian countries provide “daddy days”—leave time in the first year of a newborn’s life that the father must take or it is forfeited.

A radical effort at restructuring government has taken place in France—a proposed program for mandating equal numbers of women and men representatives at the national level of government. Parity is not likely to become a widespread policy, but even redressing gender imbalance would give women and men a more equal opportunity to make laws and influence social policy. Gender differences in voting patterns in the United States indicate that women do have a different perspective on many issues. For women to be elected in greater numbers, powerful men now in politics would have to encourage young women to consider a career in politics, foster their advancement through mentoring, nominate them for national offices, and campaign with them and raise money for them when they run for office. Paradoxically, it has seemed easier for women to become heads of state than for these same states to vote in an equal number of women and men in their governing bodies. It has also been easier for women to become heads of parliamentary governments, where a party chooses the prime minister. The appointment of women to high positions, such as Madeleine Albright as U.S. Secretary of State, has also been welcomed. Yet when it comes to directly putting women into leadership positions of great authority, whether in government or in major corporations, there is still a public reluctance to grant women as much power as men.

On a more personal level, some people have structured their families to be gender-equal on every level—domestic work, child care, and financial contribution to the household economy. However, as long as work is structured for a married-man-with-wife career pattern, and men’s work is paid better than women’s work, gender-equal families will be very hard to attain by the majority of people. Other heterosexual couples have reversed roles—the woman is the breadwinner and the man cares for the children and keeps house. Here, the problem is that the domestic world is so gendered

that male househusbands suffer from ostracism and isolation, as well as from a suspicion of homosexuality. Oddly, lesbian and gay couples who have reared children in a variety of family arrangements have blended more easily into hetero-coupled social worlds, at least in some communities. Corporate and government policies that offer health insurance and other benefits to any couple in a long-term household arrangement have also helped to restructure family life in ways that do not assume heterosexuality and marriage. Note, however, that communal domestic households have waned in popularity in Western countries, although they are the norm in polygamous cultures.

Least amenable to change have been the gendered divisions of work in the global economy. Financed by capital from developed countries, work organizations around the world exploit the labor of young, unmarried women under sweatshop-like conditions, while reserving better-paid jobs and support for entrepreneurship to men. The policies of the International Monetary Fund and other financial restructuring agencies do not include among their goals gender desegregation or encouraging women’s education and access to health resources. In many developing countries, violence and sexual exploitation, as well as the heterosexual spread of AIDS, seriously undermine efforts to upgrade the lives of women and girls.

In sum, to change gendered social orders to be more equal (or, alternatively, less gendered) will take individual effort and modification of gender-stereotyped attitudes and values, but most of all, a restructuring of work and family through the policies and practices of large-scale corporations and the governments of dominant nations.

## REFERENCES

- Acker, J. 1989 *Doing Comparable Worth: Gender, Class, and Pay Equity*. Philadelphia: Temple University Press.
- Barrett, M. 1988 *Women’s Oppression Today*, rev. ed. London: Verso.
- Bart, P. B., and E. G. Moran (eds.) 1993 *Violence Against Women*. Thousand Oaks, Calif.: Sage.
- Berk, S. F. 1985 *The Gender Factory: The Apportionment of Work in American Households*. New York: Plenum.
- Boserup, E. 1987 *Women’s Role in Economic Development*, 2nd ed. New York: St. Martin’s Press.

- Butler, J. 1990 *Gender Trouble*. New York and London: Routledge.
- Chafetz, J. S. 1990 *Gender Equity*. Thousand Oaks, Calif.: Sage.
- Chodorow, N. 1978 *The Reproduction of Mothering*. Berkeley: University of California Press.
- 1994 *Femininities, Masculinities, Sexualities*. Lexington: University Press of Kentucky.
- Collins, P. H. 1990 *Black Feminist Thought*. Boston: Unwin Hyman.
- Connell, R. W. 1987 *Gender and Power*. Stanford, Calif.: Stanford University Press.
- 1995 *Masculinities*. Berkeley: University of California Press.
- Coontz, S., and P. Henderson 1986 *Women's Work, Men's Property: The Origins of Gender and Class*. London: Verso.
- Epstein, C. F. 1988 *Deceptive Distinctions: Sex, Gender and the Social Order*. New Haven, Conn.: Yale University Press.
- Ferree, M. M., J. Lorber, and B. B. Hess, (eds.) 1998 *Revisioning Gender*. Thousand Oaks, Calif.: Sage.
- Herdt, G., (ed.) 1994 *Third Sex Third Gender: Beyond Sexual Dimorphism in Culture and History*. New York: Zone Books.
- Hochschild, A. 1989 *The Second Shift*. New York: Viking.
- hooks, b. (1984) 1990 *Yearning: Race, Gender, and Cultural Politics*. Boston, Mass.: South End Press.
- Kanter, R. M. 1977 *Men and Women of the Corporation*. New York: Basic Books.
- Keller, E. F. 1985 *Reflections on Gender and Science*. New Haven, Conn.: Yale University Press.
- Kessler, S. J. 1998 *Lessons from the Intersexed*. New Brunswick, N.J.: Rutgers University Press.
- Kessler, S. J., and W. McKenna 1978 *Gender: An Ethnomethodological Approach*. Chicago: University of Chicago Press.
- Laqueur, T. 1990 *Making Sex: Body and Gender from the Greeks to Freud*. Cambridge, Mass.: Harvard University Press.
- Laslett, B., S. G. Kohlstedt, H. Longino, and E. Hammonds, (eds.) 1996 *Gender and Scientific Authority*. Chicago: University of Chicago Press.
- Lorber, J. 1994 *Paradoxes of Gender*. New Haven, Conn.: Yale University Press.
- 1998 *Gender Inequality: Feminist Theories and Politics*. Los Angeles: Roxbury.
- MacKinnon, C. A. 1989 *Toward a Feminist Theory of the State*. Cambridge, Mass.: Harvard University Press.
- Melhuus, M., and K. A. Stolen, (eds.) 1997 *Machos, Mistresses, Madonnas: Contesting the Power of Latin American Gender Imagery*. New York and London: Verso.
- Messner, M. A., and D. F. Sabo, (eds.) 1990 *Sport, Men, and the Gender Order*. Champaign, Ill.: Human Kinetics.
- Oyewúmi, O. 1997 *The Invention of Women: Making an African Sense of Western Gender Discourses*. Minneapolis: University of Minnesota Press.
- Peterson, V. S., (ed.) 1992 *Gendered States: Feminist (Re)visions of International Relations Theory*. Boulder, Colo.: Lynne Rienner.
- Potuchek, J. L. 1997 *Who Supports the Family? Gender and Breadwinning in Dual-Earner Marriages*. Stanford, Calif.: Stanford University Press.
- Redclift, N., and M. T. Stewart, (eds.) 1991 *Working Women: International Perspectives on Women and Gender Ideology*. New York and London: Routledge.
- Reskin, B. F., (ed.) 1984 *Sex Segregation in the Workplace: Trends, Explanations, Remedies*. Washington, D.C.: National Academy Press.
- Reskin, B. F., and P. A. Roos 1990 *Job Queues, Gender Queues*. Philadelphia: Temple University Press.
- Risman, B. 1997 *Gender Vertigo: Toward a Post-Gender Family*. New Haven, Conn.: Yale University Press.
- Rothman, B. K. 1989 *Recreating Motherhood*. New York: Norton.
- Sainsbury, D., (ed.) 1994 *Gendering Welfare States*. Thousand Oaks, Calif.: Sage.
- Segal, M. T., and V. Demos, (eds.) 1996–98 *Advances in Gender Research*, vols. 1–3. Greenwich, Conn.: JAI.
- Smith, D. E. 1987 *The Everyday World as Problematic*. Toronto: University of Toronto Press.
- 1990 *The Conceptual Practices of Power*. Toronto: University of Toronto Press.
- 1990 *Texts, Facts, and Femininity*. New York and London: Routledge.
- Staples, R. 1982 *Black Masculinity*. San Francisco: Black Scholar Press.
- Thorne, B. 1993 *Gender Play: Girls and Boys at School*. New Brunswick, N.J.: Rutgers University Press.
- Tilley, L. A., and P. Gurin, (eds.) 1990 *Women, Politics, and Change*. New York: Russell Sage.
- Van den Wijngaard, M. 1997 *Reinventing the Sexes: The Biomedical Construction of Femininity and Masculinity*. Bloomington: Indiana University Press.
- Walby, S. 1997 *Gender Transformations*. New York and London: Routledge.

West, C., and D. Zimmerman 1987 "Doing Gender." *Gender and Society* 1:125–151.

Williams, C. L. 1995 *Still a Man's World: Men Who Do Women's Work*. Berkeley: University of California Press.

JUDITH LORBER

## GENERAL LINEAR MODEL

See Analysis of Variance and Covariance; Causal and Inference Models; Correlation and Regression Analysis.

## GENOCIDE

When, in 1881, the German anti-Semite Dühring urged a "final settling of accounts" with the Jews, he spoke, rather obliquely, of the need for a "Carthaginian" solution (p. 113f.). This was not merely a euphemism, however, as it might now seem. On the contrary, it was well known that Rome had utterly destroyed Carthage in the last Punic War. Hence Dühring's meaning was clear. But his phrasing remained cryptic because the vocabulary of mass death had not yet attained technical precision. Not until World War II, in fact, were the terms "genocide" and "crimes against humanity" introduced. Thus, in December 1948, when the United Nations declared the willful destruction of an entire people to be a breach of international law—a principle enshrined in the U.N. Convention for the Prevention and Punishment of the Crime of Genocide—the very word for mass murder was still novel.

### CONCEPTUAL ROOTS

The word "genocide" was coined by the jurist Raphaël Lemkin in his book *Axis Rule in Occupied Europe*. "New conceptions require new terms," Lemkin wrote. "By 'genocide,' we mean the destruction of a nation or . . . ethnic group"—that is, "murder, though on a vastly greater scale" (1944, pp. 79, 91).

Mass murder, of course, is hardly new. But only since the French Revolution has there been widespread concern with what the philosopher Hegel called "the fanaticism of destruction," which

aims, Hegel said, to eliminate all enemies, real or imagined ([1821] 1991, p. 38). An early expression of this concern came from Gracchus Babeuf, the leading socialist in the French Revolution, who decried the "exterminating wheel" of Robespierre's Terror. Robespierre's "general system of government," Babeuf contended, is also a "general system of extermination" ([1794] 1987, pp. 96–98). In a similar spirit, the liberal poet August von Platen deplored the Russian repression of the Polish revolution of 1831 as an instance of "*Volksmorder*" (Jonassohn 1998, p. 140). Others later made similar charges. But it was not until the Nazi Holocaust that the international community felt the need to adopt a special term for the attempted murder of a people.

In 1945, soon after Lemkin's book appeared, the defeated Nazi rulers were charged with "deliberate and systematic genocide." A year later the United Nations ratified a working definition of genocide as the "denial of the right of existence of entire human groups" (Kuper 1981, pp. 22–23). Lemkin's new coinage had won swift acceptance. Since then, it has become universally popular—and equally controversial.

### GLOBAL VIOLENCE

The popularity of the idea of genocide reflects its continuing relevance in today's world. Globally, there were at least forty-four state-organized mass slaughters in the period from 1945 to 1989, many of which were quasi-genocidal by almost any definition. These massacres resulted in an average of 1.6 million to 3.9 million deaths per annum—significantly more, that is, than the total number of fatalities caused by all wars and natural disasters in this period. And in every decade since 1945 another 1.85 million people have died in wars and civil wars (Gurr and Harff 1989). Nor has the picture improved since 1989. In 1993, for example, there were twenty-three wars, 25 million refugees, and a record number of armed U.N. interventions in conflicts around the world (Gurr and Harff 1994). In 1994, a genocidal massacre in Rwanda caused roughly 850,000 deaths (Smith 1998). Elsewhere—in Burundi, east Timor, and myriad other places—similar if smaller tragedies have unfolded.

Many groups, in other words, have been victims of systematic violence in the recent past, and

allegations of genocide abound. Yet the concept itself remains elusive—partly, indeed, thanks to its own success. So many people now equate genocide with evil that nearly every conflict is marked by charges and countercharges of genocide. Abortion, birth control, forced sterilization, intermarriage, and desegregation are just a few of the countless practices that have been labeled genocidal. Almost anything controversial, it seems, can be called genocide.

For social scientists, plainly, the matter is more complex. This was clear from the outset in Lemkin's work, which remains paradigmatic in many ways.

### TOTAL WAR

Sociologically, Lemkin believed that the Holocaust was not a contingent feature of German Nazism but rather an expression of intrinsic Nazi tendencies. As an absolutist form of nationalism, Nazism claimed absolute sovereignty over non-Germans on the ground of purported German racial-national superiority. This, in turn, led to a Nazi conception of war as Total War—the conviction, that is, that wars are not merely struggles between states but rather inevitable racial contests between peoples, with total victory or defeat as the ultimate outcomes. And the Nazi repudiation of universalist ethics inspired the belief that, in Total War, enemy peoples may be totally destroyed without ethical qualms.

Genocide was not, however, the entire Nazi agenda. On the contrary, in Lemkin's view, genocide was a single part of a many-sided strategy. The ultimate Nazi goal was to build a German empire, and in pursuing this goal only certain types of enemies were slated for annihilation: Jews, who were seen as the poisonous racial antithesis of everything German; so-called Gypsies, who were regarded as racially inferior pests; and a variety of others (including democratic intellectuals, Marxists, and resistance fighters) whom the Nazis saw as irreconcilable political foes. Enemies of other kinds, however, were to be ruled rather than massacred. Some, indeed, were to be incorporated into the Nazi empire as lesser "Germanic" peoples (including the Dutch, Flemings, and Scandinavians). Others, including Russians, Serbs, and Poles, were to be enslaved and reduced in numbers (by limits on

birthrates, food rationing, etc.) but were not to be liquidated *per se*.

### AN AGE OF ABSOLUTES

For Lemkin, that is, genocide was a strand in the fabric of Nazi war aims. And though the genocidal aspect of Nazi policy had historical precedents, it was also uniquely modern in many ways. Neither of the two pillars of the Holocaust ("master-race mythology and aggressive technology") could be grasped apart from specifically modern forces, including science and industry, nationalism, racism, and statism. Nor could modern German militarism be equated with the warrior ethics of times past. Other peoples had periodically committed mass murder, but German militarism had risen to new heights of virulence, driven, Lemkin concluded, by an acute "national and racial emotionalism" that was given a uniquely destructive force by modern technology (1944, p. xiv).

The consequence, as the sociologist Jessie Bernard noted shortly after the war ended (1949, p. 652), is that "ethnic conflict in our day has taken an unexpectedly brutal turn," even reaching the level of "exterminating whole peoples—genocide." Previous conflicts, however brutal, now seemed relatively small and unsystematic by comparison.

### VARIETIES OF GENOCIDE

This is not to say, however, that genocide is entirely unique to modernity. In recent decades, sociologists have tended to agree that the willful destruction of entire peoples is, in fact, an ancient practice, common to many cultures. But it is also widely accepted that mass murder has assumed new forms in the twentieth century (as Lemkin said). This change is not merely quantitative but qualitative, and it is often framed in terms of a contrast between two ideal-types: (1) *instrumental genocide*, or mass murder to achieve pragmatic goals; and (2) *ideological genocide*, or mass murder as an end in itself.

Save for a relatively small proportion of cases that prefigure contemporary trends (e.g., wars of religion like the Crusades), most premodern genocides were essentially instrumental in nature. Modern genocides, in contrast, are often said to be largely if not primarily ideological—though instrumental motives remain powerful as well.

### THE ORIGIN OF MASS MURDER

Many scholars, like Kurt Jonassohn, believe that genocides have occurred “throughout history in all parts of the world” (Jonassohn 1988, in Chalk and Jonassohn 1990, p. 415). Leo Kuper, in a classic study, called genocide “an odious scourge which has inflicted great losses on humanity in all periods of history” (1981, p. 11). Sociobiologists, meanwhile, have been even bolder, claiming that genocide is not only universal for humanity but for our closest evolutionary relatives as well (including chimpanzees and our evolutionary ancestors; cf. Diamond 1991, p. 264). E. O. Wilson, the founder of sociobiology, has speculated that genocide may be a “primitive cultural capacity,” rooted in “the possession of certain genes,” that confers a selective advantage on predatory groups that wish to spread their genes at the expense of their neighbors (1980, p. 298).

In reality, however, there is very little evidence of anything like genocidal violence in the evolutionary or historical record until comparatively recently. Sociobiologists who defend the idea that people and chimps share “a continuous, 5-million-year habit of lethal aggression” have offered, as their best evidence, the testimony of witnesses who report that, in more than 100 years of observation at four sites, seven chimps were assaulted by groups from neighboring chimp communities; and though none of the victims were killed on the spot, several later died or disappeared (Wrangham and Peterson 1996, pp. 16–21, 63).

This, it seems plain, is slight evidence for the large claim that people and their relatives and ancestors are genocidal by nature. Little else in the ethological or evolutionary record seems to lend credibility to this claim. Nor does history before the discovery of agriculture show significant evidence of mass murder.

To date, the earliest known evidence of collective violence is a gravesite in the Nile Valley containing fifty-nine men, women, and children, most of whom were killed by weapons with projectile points. This massacre, dating back 12,000–14,000 years, was “almost certainly” a consequence of disruptions produced by a crisis in the proto-agricultural system that was evolving in the region (Reader 1998, p. 146). The only other authenticated evidence of mass violence in remote antiquity is

a similar but smaller burial ground in Bavaria dating back 7,700 years (Frayner 1997).

### PLOUGHSHARES INTO SWORDS

In general, archaeological evidence makes it clear that wars, massacres, and even murders were comparatively rare in pre-agricultural days (Ferguson 1997). And the ethnographic evidence shows that, while hunter-gatherers are seldom entirely non-violent, they have significantly fewer wars, internal conflicts, and interpersonal violence than agrarian peoples (Ember and Ember 1997; Ross 1993). Even the famously “fierce” Yamomami of northern Brazil and Venezuela seem to have been driven to violence as much by the influence of outside forces as by internal pressures (Ferguson 1992). “It seems unlikely,” as the genocide scholars Chalk and Jonassohn conclude, “that early man engaged in genocide during the hunting and gathering stage” (1990, p. 32).

For some peoples, this stage lasted until the recent past. For humanity as a whole, hunting and foraging remained universal until roughly 11,000 years ago. Shortly afterwards, agriculture arose in the Middle East; and then, in succession, in China 11,000 years ago in Papua, Mexico, and the Andes 10,000 years ago in West Africa 6,000 years ago and elsewhere.

Once agriculture had appeared, mass violence assumed new dimensions.

### EMPIRE AND TERROR

The first weapons of war, and the earliest signs of mass warfare, appeared at the dawn of the agrarian era. In the ensuing millenia the first true wars were fought (Ferrill 1997, pp. 18–19). The intent of these wars was usually instrumental—to win land and labor, to keep enemies at bay, to build states and empires. This was the context in which mass violence became familiar. Genocide in particular was almost always linked to empire building in this period.

Historically, empire builders generally seek tribute or servile labor from the peoples they conquer, which leads them to spare the lives of noncombatants, especially peasants. This was the



spirit in which Sumerian, Hittite, and Babylonian epics of remote antiquity warned of the folly of genocide (Cohn 1996). But as fortified cities emerged as obstacles in the path of empire, conquest sometimes took more brutal forms. And peasant resistance often provoked vengeful repression. Hence, by the third millennium B.C.E., genocidal conquests had become comparatively common in the early Egyptian empire (Jonassohn 1998). Elsewhere, genocidal massacres seem to have become common by the first millennium B.C.E.

The Assyrians in particular became legendary for their merciless assaults on fortified cities, including Babylon, which Sennacherib decimated in 689 B.C.E. The goal of such assaults was usually to destroy an imperial or commercial rival. The preferred method was to lay siege to a city and starve its citizens into submission by burning crops and destroying livestock. Once a campaign of this kind triumphed, the conquerors would often raze the defeated city, salt the earth, and slaughter the citizens. The legalistic Romans dubbed this practice “devastation.”

Early “devastations” included the destruction of Melos by Athens in 416 B.C.E. (for refusing to take sides in the Peloponnesian Wars); the destruction of two Sicilian cities by Carthage in 409 B.C.E.; and the annihilation of Carthage by Rome in 146 B.C.E. In the latter case, as many as 150,000 people are said to have died.

#### THE CONTINUATION OF WAR BY OTHER MEANS

Sociologists have identified a range of motives for instrumental genocide. Chalk and Jonassohn, for example, say that most genocides of this type (which they call “utilitarian”) have been prompted by the wish to acquire wealth, to terrorize real or potential enemies, or to avert threats (1990, pp. 34–36). Until comparatively recently, motives of this kind have been the main effective spurs to genocide both in the circum-Mediterranean world and elsewhere. In brief, it seems that whenever empires fight, great opportunities and equally great dangers are at stake. In such cases, genocidal massacres can become the extraordinary means of attaining goals that, in other contexts, would be pursued by ordinary statecraft.

Examples abound. Canton was gutted in 879 C.E. and its population was decimated. In the twelfth century C.E., Afghan conquerors from Ghur destroyed the city of Ghazni and all its men. In the next century the Mongols laid waste to Transoxania and Khurasan, evidently killing hundreds of thousands of people. Not long afterwards, all 12,000 of the Meos people (south of Delhi) were killed in the wake of a failed uprising. Many events in the sixteenth-century Spanish conquest of the Americas were plainly genocidal. And many other examples of this kind can be cited as well (Chalk and Jonassohn 1990; Jonassohn, 1998).

A related form of genocide accompanied the rise of modern colonialism. In the Americas, for example, the English Puritans liquidated the Pequot in 1637, the French destroyed the Natchez in 1731, and Argentine settlers virtually eliminated the Araucanians in 1878–1885. In Africa, many peoples have been devastated by colonial conquests; some (including the Herero of Namibia and the Matumbi of Tanzania) have suffered mass destruction. Many indigenous peoples in Australia have been destroyed, and in Tasmania the entire aboriginal population was extinguished by 1876.

These are just a few of the many cases in which genocide has accompanied instrumental conquest.

#### IDEOLOGICAL GENOCIDE

In some cases, however, genocide is not an instrumental means to an end, but rather an end in itself. Genocides of this type, in which the destruction of entire peoples is sought for intrinsic reasons, are generally classified as “ideological” (Fein 1993; Smith 1987). These massacres, which are typically fueled by religious or ethnoracial fanaticism, are often among the deadliest of genocides, since they seek to destroy targeted peoples for reasons of “principle,” not simply expediency. And they are often instrumentally irrational in an almost chemically pure sense. In 1994, for example, the rulers of Rwanda were so preoccupied by the wish to annihilate an unarmed ethnic minority that they left themselves virtually defenseless against an army of invading rebels from Uganda (Des Forges 1999). Between 1975 and 1979, the Pol Pot regime wrought havoc across the length and breadth of Cambodia in a vain effort to purge Cambodia of every trace of foreign, urban, and intellectual influence (Kiernan

1996). In the Nazi Holocaust, the German leaders diverted vital resources from the Russian front to fuel the militarily useless operations of the death camps (Postone 1980). Earlier, during World War I, Turkey failed to finish the strategically crucial Berlin–Bagdad railway largely because the Turkish rulers were so intent on destroying the Armenians that they liquidated large numbers of Armenian construction workers (Jonassohn 1998).

Genocides of this type are thus doubly destructive—first, because they are prosecuted with systematic ruthlessness; and second, because they often hurt the perpetrators as well as the victims.

### BEFORE IDEOLOGY?

Only a few instances of premodern mass violence are clearly ideological in inspiration—including, for example, the Crusades and the persecution of Cathar heretics in thirteenth-century Languedoc. But many other early cases also reveal a noninstrumental side.

Beginning in 1634, for example, the Iroquois fought a series of “mourning wars” against the Hurons, which resulted in large-scale massacres of Huron men and women. The main object, in this case, was to “requicken,” in the bodies of Huron children, the spirits of Iroquois who had died in epidemics. This was perhaps not an “ideological” motive in the modern sense, but it was clearly inspired by religious belief (Richter 1992). Similarly, the genocidal massacres that accompanied the Mongol conquests were fueled, in part, by apocalyptic faith in Mongol destiny (Jagchid and Hyer 1979). And several Indian cases recorded by Jonassohn (1998) were largely religious in character—for example, the massacre of Jains by the Hindu king of Pandya in the eleventh century C.E., the destruction of the Buddhist majority of Bihar by Muslim conquerors in the twelfth century C.E., and the persistent efforts of Bahmani sultans to exterminate or convert Hindus in the Deccan region in the period 1347–1482 C.E.

Broadly defined, in other words, ideological genocide may have been more common in the past than scholars have tended to say. But it is nonetheless plain that the twentieth century marked a new departure. Only in this century did mass murder routinely followed paths defined by doctrine.

### A NEW SELF-CONSCIOUSNESS

Raphaël Lemkin had been well aware that genocidal practices predated the twentieth century (Fein 1993). But he believed that the absolutist racism and nationalism of the modern era gave contemporary mass murder decisive new features. Above all, genocide was now pursued with heightened self-consciousness. In the characteristic genocides of the twentieth century—most notably the liquidation of Armenians by Turkey in World War I, the Nazi Holocaust during World War II, and the postwar massacres of minority groups by governments in Indonesia, Pakistan, Cambodia, Burundi, and Rwanda—the intent was genocidal from the start. Entire groups were sentenced to death for ideological as well as geopolitical reasons. And these sentences were pronounced with full awareness of their antihumanistic implications, since they defied the humanitarian trend, rooted in Puritanism and bourgeois constitutionalism (Tergel 1998; Troeltsch [1912] 1958), which had swept the Euro-Atlantic world in the nineteenth century.

What explains this genocidal intent? What gives “ideological” genocide its compelling force? For Lemkin the answer lies, in part, in social psychology. He sees evidence, in particular, of such psychological tendencies as cynicism, denial, “contempt for the victim,” and “exaggerated pride,” among others (1992, pp. 189–235). This line of reasoning is richly developed in the research literature on the personality traits most relevant to genocide: aggressiveness, submissiveness, and punitiveness. It was first shown, in a flawed but classic study (Adorno et al. 1950), that these three traits systematically covary in structured and intelligible ways. Later research has significantly refined this original discovery.

It is now well understood that certain kinds of harsh personal experience tend to make people (1) *submissive* to leaders and (2) *aggressive* toward—outsiders. And people imbued with this dually obedient and violent tendency are often exceptionally willing to persecute or fight those who are authoritatively defined as their “enemies” (Altemeyer 1996; Hopf 1993; Lederer and Schmidt 1995). The significance of this fact for the study of genocide was revealed by Ross, who showed that conflict intensity is strongly related to personality variables of just this type. While objective social divisions tend to determine *who fights whom*—class versus

class, nation versus nation, and so forth—personality-forming experiences play an exceptionally large role in determining the destructiveness of any given conflict. Ross found, after systematically studying the interplay of forty-one variables in ninety nonindustrial societies, that “the more affectionate and warm and the less harsh the socialization in a society, the lower the level of political conflict and violence” (1993, p. 99); and indeed, no other single factor has a greater impact.

### DIVIDED SOCIETIES

Patterns of socialization are thus highly relevant to the study of genocide. No less relevant are structural determinants—few of which, however, have yet received systematic attention. This is largely due to the fact that the sociology of genocide is still just a few decades old. While studies of anti-Semitism, ethnocentrism, and the Holocaust have been common throughout the postwar era, the study of genocide as a phenomenon *sui generis* was first pursued in the 1970s by a small cohort of sociologists—Dadrian, Kuper, Fein—who have remained central to the enterprise ever since.

Leo Kuper, the “doyen of genocide scholars” (Charny 1994, p. 64), set the stage for much of the research that was to follow. Focusing entirely on the twentieth century, Kuper noted that genocide in this century has been distinctive in several ways. In addition to the rising influence of ideological motives, Kuper observed that contemporary genocides are also exceptionally likely to occur within divided societies, as autocratic states seek to defend their monopoly of power against the challenges of domestic rivals. Genocidal massacres are still common in the conflicts between societies as well—especially, Kuper says, when imperialist motives are at work—but internal conflicts are now especially acute. This is particularly true in postcolonial societies, many of which were profoundly destabilized by the experience of colonial rule. Such societies are often pure artifacts of the colonial imagination, rigidly stratified and ethnically divided in new and untenable ways. When dissent arises in these “new nations,” the rulers—*sans* consent—resort to coercion, often on a grand scale. Hence the state-sponsored massacres in such postcolonial realms as Indonesia, the former “East Pakistan” (now Bangladesh), Rwanda, and Burundi. And communal divisions in these nations run so

deep that genocidal massacres often unfold entirely within civil society, unbidden by the state (Kuper 1981).

Many of Kuper’s themes were adapted from classical texts by Sartre (1968) and Cohn (1967), and many later scholars have pursued variations on these themes (see Andreopolous 1994; Totten et al. 1995). Except for Jonassohn, who delves into remoter history, most social scientists continue to focus mainly on twentieth-century cases (Dadrian 1995; Mazian 1990; Melson 1992; Wallimann and Dobkowski 1987). Much has been learned about these cases. Yet in many ways genocide remains as elusive as ever. Part of the problem is definitional.

### DUELLING DEFINITIONS

As the study of genocide proves, definitions are deeply political. This was apparent from the beginning of the United Nations debate over genocide in 1948. For the Russian delegates, “genocide” was a logical effect and extension of fascist doctrine, revolving around the victimization of racial and ethnic groups. Others argued that social classes and political groups can also be the victims of genocide. When the final vote was taken, religious as well as national and ethnic groups were identified as possible targets of genocide, but political groups were excluded. And the United Nations further limited the applicability of its definition by specifying that genocide entails an “intent to destroy” some group, “in whole or in part.” Since proving intent is notoriously difficult, some delegates questioned whether the Genocide Convention, so restricted, could ever be enforced. Similar doubts were voiced about the idea of “destroying” a group. Lemkin, in his original formulation, had said that genocide can take cultural as well as physical forms. But does this mean that efforts to sabotage the culture, language, or religion of a group are just as much “acts of genocide” as attempts to physically destroy the group?

Questions of this kind quickly proved to have direct political relevance. In 1951, W. E. B. Du Bois, Paul Robeson, and others addressed a book-length petition to the United Nations accusing the U.S. government of systematically genocidal actions against African Americans (Patterson et al. 1951). Reluctance to entertain charges of this kind

led many U. S. senators to oppose the Genocide Convention; and, indeed, the United States did not endorse the Convention until 1986 (LeBlanc 1991). Nor did the United Nations prove eager to enforce the Convention. Not until the Rwandan massacres of 1994, in fact, did the United Nations classify a postwar event as “genocide”—and even this came after the fact. It has remained extremely difficult, in short, for policy makers to operationally define or apply the notion of genocide.

Scholars, meanwhile, have spent considerable time on definitional issues. Kuper offered an influential distinction between genocide and “genocidal massacre” as a means of distinguishing the attempted murder of entire peoples from large-scale massacres that are not prompted by the wish to destroy groups *in toto*. He did not, however, establish clear criteria for this distinction, which limits its utility. Nor has any other approach won consensus. Thus, debate rages over the classification of many events: “the rape of Nanking,” Stalin’s purges, the nuclear bombing of Hiroshima, the carpet bombing of Vietnam, “human-made” famine in Mao’s China, and similar events. Other debates pivot on questions of principle. Are there unintended genocides? Is genocide distinct from the destruction of a culture? Does genocide belong on a continuum with other forms of destructiveness, or is it qualitatively unique?

#### THE LIMITS OF CLASSIFICATION

Many answers have been given to these questions, but it remains far from clear that “genocide” is, indeed, a unitary phenomenon or amenable to general explanation. Most scholars have tended to focus on relatively clear cases of mass murder inspired by modern racial-national ideologies, but even these vary in significant respects. Another approach, proposed by the psychologist Charny (1994), is to subsume the full range of mass-destructive practices into a wider matrix; but it seems unlikely that the criterion of mass destructiveness alone is sufficient to justify treating ecological catastrophe, ethnic slaughter, and cultural repression as elements in a unified field of events.

For fuller, more nuanced understanding, comparative research is needed on a widening scale. Whether the general notion of “genocide” will continue to prove compelling remains to be seen.

#### REFERENCES

- Adorno, Theodor W. Else Frenkel-Brunswik, Daniel Levinson, and R. Nevitt Sanford 1950 *The Authoritarian Personality*. New York: Harper and Bros.
- Altemeyer, Bob 1996 *The Authoritarian Specter*. Cambridge, Mass.: Harvard University Press.
- Andreopoulos, George J., (ed.) 1994 *Genocide: Conceptual and Historical Dimensions*. Philadelphia: University of Pennsylvania Press.
- Babeuf, Gracchus (1794) 1987 *La guerre de la Vendée et le système de dépopulation*. Paris: Tallandier.
- Bernard, Jessie 1949 *American Community Behavior*. New York: Dryden.
- Chalk, Frank, and Kurt Jonassohn 1990 *The History and Sociology of Genocide*. New Haven, Conn.: Yale University Press.
- Charny, Israel W. 1994 “Toward a Generic Definition of Genocide.” In George J. Andreopoulos ed., *Genocide*. Philadelphia: University of Pennsylvania Press.
- Cohn, Norman 1967 *Warrant for Genocide*. New York: Harper and Row.
- 1996 *Noah’s Flood*. New Haven, Conn.: Yale University Press.
- Dadrian, Vahakn N. 1995 *The History of the Armenian Genocide*. Providence, R.I.: Bergahn.
- Des Forges, Alison 1999 *Leave None to Tell the Story: Genocide in Rwanda*. New York: Human Rights Watch.
- Diamond, Jared 1991 *The Rise and Fall of the Third Chimpanzee*. London: Radius.
- Dühring, Eugen 1881 *Die Judenfrage als Racen-, Sitten- und Kulturfrage*, 2nd ed. Leipzig and Karlsruhe: H. Reuther.
- Ember, Carol R., and Melvin Ember 1997 “Violence in the Ethnographic Record.” In Debra L. Martin and David W. Frayer, eds. *Troubled Times*. Amsterdam: Gordon and Breach.
- Fein, Helen 1993 *Genocide: A Sociological Perspective*. Newbury Park, Calif.: Sage.
- Ferguson, R. Brian 1992 “A Savage Encounter.” In R. Brian Ferguson and Neil L. Whitehead, eds. *War in the Tribal Zone*. Santa Fe, N. Mex.: School of American Research Press.
- 1997 “Violence and War in Prehistory.” In Debra L. Martin and David W. Frayer, eds. *Troubled Times*. Amsterdam: Gordon and Breach.
- Ferrill, Arther 1997 *The Origins of War*. Boulder, Colo.: Westview.
- Frayer, David W. 1997 “Ofnet.” In Debra L. Martin and David W. Frayer, eds. *Troubled Times*. Amsterdam: Gordon and Breach.

- Gurr, Ted Robert, and Barbara Harff 1994 *Ethnic Conflict in World Politics*. Boulder, Colo.: Westview.
- 1989 “Victims of the State.” *International Review of Victimology* 1(1):23–41.
- Hegel, G. W. F. (1821) 1991 *Elements of the Philosophy of Right*. New York: Cambridge University Press.
- Hopf, Christel 1993 “Authoritarians and Their Families.” In William F. Stone, Gerda Lederer, and Richard Christie, eds., *Strength and Weakness: The Authoritarian Personality Today*. New York: Springer-Verlag.
- Jagchid, Sechin, and Paul Hyer 1979 *Mongolia’s Culture and Society*. Boulder, Colo.: Westview.
- Jonassohn, Kurt with Karin Solveig Björnson 1998 *Genocide and Gross Human Rights Violations in Comparative Perspective*. New Brunswick, N.J.: Transaction.
- Kiernan, Ben 1996 *The Pol Pot Regime: Race, Power, and Genocide in Cambodia Under the Khmer Rouge, 1975–79*. New Haven, Conn.: Yale University Press.
- Kuper, Leo 1981 *Genocide: Its Political Use in the Twentieth Century*. New Haven, Conn.: Yale University Press.
- LeBlanc, Lawrence J. 1991 *The United States and the Genocide Convention*. Durham, N.C.: Duke University Press.
- Lederer, Gerda, and Peter Schmidt, (eds.) 1995 *Autoritarismus und Gesellschaft*. Opladen, Germany: Leske + Budrich.
- Lemkin, Raphaël 1944 *Axis Rule in Occupied Europe*. Washington D.C.: Carnegie Endowment for International Peace.
- 1992 *Not Guilty? Raphael Lemkin’s Thoughts on Nazi Genocide*. In Steven L. Jacobs, ed. Lewiston, U.K.: Edwin Mellen.
- Mazian, Florence 1990 *Why Genocide? The Armenian and Jewish Experiences in Perspective*. Ames: Iowa State University Press.
- Melson, Robert 1992 *Revolution and Genocide*. Chicago: University of Chicago Press.
- Patterson, William, W. E. B. Du Bois, et al. 1951 *We Charge Genocide: The Historic Petition to the United Nations for Relief from a Crime of the United States Government Against the Negro People*. New York: Civil Rights Congress.
- Postone, Moishe 1980 “Anti-Semitism and National Socialism.” *New German Critique* 19:97–116.
- Reader, John 1998 *Africa*. New York: Knopf.
- Richter, Daniel K. 1992 *The Ordeal of the Long House*. Chapel Hill: University of North Carolina Press.
- Ross, Marc Howard 1993 *The Culture of Conflict*. New Haven, Conn.: Yale University Press.
- Sartre, Jean-Paul 1968 *On Genocide*. Boston: Beacon Press.
- Smith, David Norman 1998 “The Psychocultural Roots of Genocide: Legitimacy and Crisis in Rwanda.” *The American Psychologist* 53(7): 743–753.
- Smith, Roger W. 1987 “Human Destructiveness and Politics.” In Isidor Wallimann and Michael N. Dobkowski, eds., *Genocide and the Modern Age*. New York: Greenwood.
- Tergel, Alf 1998 *Human Rights in Cultural and Religious Traditions*. Uppsala: Acta Universitatis Upsaliensis.
- Totten, Samuel, William S. Parsons, and Israel W. Charny, (eds.) 1995 *Genocide in the Twentieth Century: Critical Essays and Eyewitness Accounts*. New York: Garland.
- Troeltsch, Ernst (1912) 1958 *Protestantism and Progress*. Boston: Beacon Press.
- Wallimann, Isidor, and Michael N. Dobkowski, (eds.) 1987 *Genocide and the Modern Age*. New York: Greenwood.
- Wilson, E. O. 1980 *Sociobiology*, 2nd ed. Cambridge, Mass.: Harvard University Press.
- Wrangham, Richard, and Dale Peterson 1996 *Demonic Males: Apes and the Origins of Human Violence*. Boston: Houghton Mifflin.

DAVID NORMAN SMITH

## GERONTOLOGY

See Aging and the Life Course; Cohort Perspectives; Filial Responsibility; Intergenerational Relations; Intergenerational Resource Transfer, Long Term Care, Long Term Care Facilities; Retirement; Widowhood.

## GERMAN SOCIOLOGY

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

“German Sociology” has two specific traits: It is part of the general humanities tradition of German culture—that is, it has a philosophical orientation—and it emphasizes epistemological reflection, favoring the understanding of human action through *verstehen* (intuitive oneness with the explanandum). This is the way in which Raymond Aron (1935) characterized sociology in Germany,

studying it at the time of the Nazi regime, when it was mainly a memory and not a living field of knowledge or profession.

Since then a large number of books and essays in the United States have treated sociology as practiced in Germany, at least some of which broaden the image (Nisbet 1966; Fletcher 1971; Oberschall 1965, 1972; Schad 1972; Freund 1978). The dominant meaning of the intellectual commodity called German sociology—as continued in the works of Salomon (1945), Barnes (1938), Coser (1977), Zeitlin (1981), and especially Meja, Misgeld, and Stehr (1987)—is that of grand theory with a teleological meaning. This “German Sociology” is only a small part of sociology in Germany, but it is the aspect to which the intellectual community at large reacts; and it is reinforced by selectivity in translation and citation. For the social sciences in the United States, those parts of sociology in other countries which are least like American sociology are searched out as welcome completions.

### THE CLASSICAL PERIOD

By the turn of the twentieth century, sociology had become an exciting intellectual concern in the United States, France, England, and Germany. Since an educated person at that time was expected to know English, German, and French, there was intensive direct interaction. Of special importance were the exchanges among Albert Schäffle (1884), Ferdinand Tönnies (1887), Emile Durkheim (e.g., 1887), and Georg Simmel (1890). Durkheim first made a name for himself in France by reviewing German social science literature by authors such as Wilhelm Wundt and, especially, Ferdinand Tönnies, and was therefore attacked as a Germanophile. Tönnies’s distinction *Gemeinschaft* (community) and *Gesellschaft* (society) influenced Durkheim’s dichotomy mechanical and organic solidarity, and Tönnies in turn commented on Durkheim’s notion of the division of labor as the central element in social evolution (Gephardt 1982). Initially the builders of large “systems,” such as Ludwig Gumplowitz (1883), Gustav Ratzenhofer (1893), and Theodore Abel (1929), dominated German sociological literature.

In a review of the period up to 1914, Leopold von Wiese (1959) lists Ferdinand Tönnies, Max Weber, and Werner Sombart as the most important sociologists; the list should also include Alfred

Vierkandt, Franz Oppenheimer, Alfred Weber, Roberto Michels, and Hermann Kantorowicz (Käsler 1984). This inner circle of sociologists of the “classical period” is still influential today.

In the closing days of the German empire, sociology was established in an elitist academy with Tönnies as president, Sombart as vice president, and von Wiese as the eminence gris. Simmel and Weber were considered to be the leading scholars within this academy, even though both had become merely observers. In 1912, after numerous editions, Tönnies’s *Gemeinschaft und Gesellschaft*, first published in 1887, at last achieved wide recognition among the educated public, paralleling the impact of Edward A. Ross’s *Social Control* in the United States.

### THE WEIMAR PERIOD

Upon the reestablishment of sociology in 1919, von Wiese was able to retain bureaucratic control over the policy of the academy and largely over its conventions. Not until the end of the 1920s was the first chair of sociology instituted in Frankfurt. However, during the fourteen years of the Weimar Republic, forty professorships were created that combined sociology with another discipline, such as economics, philosophy, or law. Eight periodicals had *Sociology* in their titles, and another eight regularly published sociological contributions. Von Wiese was instrumental in founding the first permanent research institute at the University of Cologne, *Forschungsinstitut für Sozial- und Verwaltungswissenschaften*, in 1919 (Alemann 1978), followed in 1923 by the *Institut für Sozialforschung* in Frankfurt (Jay 1973). In the early 1930s nearly all universities in Germany regularly offered courses in sociology, and by the mid-1920s the *Deutsche Gesellschaft für Soziologie* had begun to question the wisdom of opposing a degree curriculum in the universities.

The 1920s were a time of abrasive partisanship in German intellectual life. René König groups the various positions along a dividing line between left and right Hegelians, the Kantian tradition having paled in the humanities (König 1987). Left Hegelian translates into Marxism, but that itself was very heterogeneous. When Eisenstadt and Curelaru (1976, p. 122) and Zeitlin (1981, p. v) maintain that “a critical reexamination of Marxism” is a main

focus of German sociology, this is an error; Marxists and non-Marxists mostly tended to ignore each other, although not in the 1920s.

“In the 1920s there was no dominant figure in sociology, which evolved in a number of milieus with little common direction. Even within its local centers in the Weimar Republic there was practically no paradigmatic unity” (Lepsius 1987, p. 40). These local centers were Frankfurt, Cologne, Berlin, and (later) Leipzig. In this characterization of the Weimar turmoil, there are two omissions: Max Weber’s influence is not mentioned, and the Frankfurt School is bypassed. In both cases it is done for the same reason: at that time they were not very important for sociology. Shortly after accepting a professorship in Munich, Weber died in 1920. Some of his most important works on religion had appeared during the war, and his magnum opus, *Wirtschaft und Gesellschaft* (Economy and Society) was published posthumously by his wife Marianne Weber. Other works by Weber were not readily available until after about 1925.

The Institut für Sozialforschung in Frankfurt did not see itself as an institute for sociologists. Felix J. Weill, who obtained the funds for the institute from his grain merchant father, would have preferred to call it Institute for Marxism, but it was then judged prudent to choose the neutral title Social Research. All of the members shared an aesthetic disgust with bourgeois society, though they themselves were from well-off-to-very-rich families, and they wanted to convert fellow intellectuals to this view. The most important effect of the institute until the 1940s, however, was to give younger social scientists a chance to develop their talents: Max Horkheimer, Karl August Wittfogel, Franz Borkenau, Leo Löwenthal, Herbert Marcuse, and Theodor Adorno. “Although without much impact in Weimar, and with even less during the period of exile that followed, the Frankfurt School was to become a major force in the revitalization of Western European Marxism in the postwar years” (Jay 1973, p. 4).

### AN END TO SOCIOLOGY

It is easier to name sociologists who did not emigrate as the Nazi regime came to power than to list the émigrés. Of all the sociologists with a reputation, only the social philosopher Hans Freyer welcomed the new regime. Werner Sombart, who

during his lifetime took just about every political stance available, was in an anticapitalist anti-Semitic phase around 1933. It must have protected him when, in 1938, he ridiculed racism and the glorification of the people. Franz Oppenheimer and Eric Vögelin were eventually forced to emigrate, even though they tried to remain. Othmar Spann lost his professorship in 1938 and was imprisoned. Ferdinand Tönnies, whose show of opposition bordered on the suicidal, was ostracized. Alfred Weber was dismissed. Alfred Vierkandt had to cease lecturing. Alfred von Martin resigned. The Deutsche Gesellschaft für Soziologie was suspended in 1934 by von Wiese in order to avoid a takeover. Von Wiese stopped publishing the *Kölner Zeitschrift* in the same year, and from then on, lectured only on the history of economic thought. René König, a young candidate for a professorial career who had to emigrate in 1935, attributes to the Nazis a complete stoppage of sociology worthy of its name (König 1978, p. 14).

### A SECOND BEGINNING

Sociology after 1945 could not have been a continuation of a tradition; and even if it had been possible, it would not have been a discipline in which basic issues had been settled. This is especially true for the issue of professionalization, which translates into the question “What public are the sociologists addressing?”

Emigrants returning to Germany (René König, Max Horkheimer, Theodor W. Adorno, Siegfried Landshut, Helmut Plessner, Arnold Bergsträsser, Emmerich K. Francis, and, much later, Norbert Elias) often had been influenced by developments in America. American sociology was most influential, since books from the United States often were the only ones available. The “Young Turks” of sociology born between 1926 and 1930—Karl Martin Bolte, Rainer M. Lepsius, Burkhard Lutz, Renate Mayntz, Erwin K. Scheuch, and Friedrich Tenbruck—studied the subject in American universities.

Three of the five research centers of the period of reconstruction were financed by foreign sources: the Sozialforschungsstelle Dortmund with funds from the Rockefeller Foundation; the UNESCO Institute for Social Research in Cologne, through the initiative of Alva Myrdal; and the Institut für Sozialwissenschaftliche Forschung in Darmstadt,

with American government money. The trade unions established sociology as a “democratic discipline” in Hamburg at the Akademie für Gemeinschaft. Von Wiese reopened the institute in Cologne that later was integrated into the university. With generous financing from American sources, such as the government-sponsored Voice of America, the Institut für Sozialforschung in Frankfurt resumed operations. And in Göttingen, Hans Paul Bahrdt took up the tradition of industrial sociology of the late Weimar times and founded the Soziologisches Forschungsinstitut, Göttingen—the SOFI Institut. Thus, from the resumption of sociology in Germany, there was an infrastructure for empirical social research well connected to the international community. With the exception of von Wiese’s institute, however, it later proved to be impossible to integrate these institutes for empirical research into the universities, and therefore the UNESCO Institute, the Sozialforschungsstelle Dortmund, and the Darmstadt group ceased to exist.

Empirical social research at that time was largely understood to be an import from the United States. At the UNESCO Institute in Cologne and in Darmstadt, important community studies inspired by American studies of the 1930s and 1940s were carried out. The reports by Renate Mayntz and Erich Reigrotzki are still very much worth reading. Parallel to this, survey research was developed as a commercial service; here, too, the American standards of the time were immediately imported. Of great influence were the survey units that the American and British governments had begun as their troops moved into Germany. It was in these survey units that the core personnel of the later German institutes learned their techniques (Scheuch 1990b).

The 1950s were characterized by a coexistence of professors who had learned sociology largely on their own before 1945 and a larger number of sociologists born between 1926 and 1930 who were virtually identical in skill and outlook with their American contemporaries. For this generation Weimar sociology was forgotten, and the classics were read with an American selectivity and perspective.

By the end of the 1950s sociology—in terms of chairs, curricula, students, and volume of empirical research—had surpassed the level of 1933. Much against the wishes of von Wiese and

Horkheimer, the German Sociological Association had been transformed from an elitist academy into a professional association with increasingly important sections or research committees in which the Young Turks were able to attract followers. The sections for industrial sociology (for the more theoretically minded) and for methodology (for the mainly empirically oriented) offered an alternative to the plenary meetings that were still dominated by professors who were the last academic mandarins of German tradition. In research nearly everyone who later influenced the discipline worked on social stratification, which as a central topic succeeded the community studies of the first half of the 1950s. Among the Young Turks and their following a structural functional approach was the common paradigm, and Talcott Parsons was held to be the great theorist of the time.

Sociology was a deeply divided discipline. One dividing line pitted those who had emigrated against sociologists who had supported the Nazi regime. Hans Freyer, Arnold Gehlen, and Helmut Schelsky were accused of collaborating with the regime; Gunter Ipsen did more than that; and Karl-Heinz Pfeffer and Karl Valentin Müller were justly denounced as racists. In 1960 the German Sociological Association nearly split along the dividing line between collaborators with the Nazi regime and the majority led by former immigrants.

However, it was not possible simply to dismiss Gehlen and Schelsky as former Nazi sympathizers. While they had been exactly that, they now did important scholarly work. Gehlen developed an anthropology that included perhaps the best approach to the analysis of social institutions, and Schelsky had initiated many important studies on youth, the family, and social stratification. The two men published a textbook in sociology that saw four editions within three years (Gehlen and Schelsky 1955).

The second dividing line ran between the Frankfurt Institute and all others. Horkheimer, Adorno, and Pollack had been successful in the United States, the home of modern sociology. By now their Marxism had toned down to a variant of left Hegelianism that they christened “critical theory.” Initially a political camouflage, it now characterized the retained commitment to cultural criticism of bourgeois society. The combination of



Marxism with cultural criticism proved to be a winning message with the cultural establishment.

The opposition to Schelsky/Gehlen and Adorno/Horkheimer crystallized around René König, whom von Wiese had invited to Cologne. Originally a pure humanist—he wrote what is considered the standard monograph on Niccolò Machiavelli (1941)—in exile he had identified with the post-Durkheim school in France. Upon returning to Germany in 1950, he recognized the need to familiarize the young generation of sociologists and sociology students with mainstream American sociology. Although himself not a quantitatively minded scholar, he encouraged quantitative social research as an antidote to the temptations of speculative grand theory. Knowing his personal preference for cultural anthropology and French culture, it is ironic that he was the key figure in thoroughly Americanizing the larger part of an academic generation in sociology.

König's *Soziologie heute* (Sociology Today, 1949), a sociological rendering of existentialist philosophy, was the first postwar best-seller in sociology. The dictionary of sociology that he edited (1958) brought together most of the Young Turks and became the largest selling sociology book ever in Germany, with more than twenty editions and almost half a million copies. Most important for the profession, however, was the two-volume handbook of empirical social research (1962) that he edited, conceived along the lines of Gardner-Lindzey's *Handbook of Social Psychology* (König 1962). The handbook is still a standard for empirical sociology and has gone through several editions. All of these works were translated into English, French, Spanish, and Italian, and some into Japanese as well.

The different positions in the first phase of sociology after 1945 were sorted into three "schools" that were to constitute sociology in Germany: the Frankfurt School, the Cologne School, and the Schelsky school. Within the discipline the Cologne school set standards for curricula and research, the *Kölner Zeitschrift* being stronger than the Dortmund-based *Soziale Welt*. The readers of these two largest social science journals do not overlap much. *Kölner Zeitschrift* is the journal for the discipline, while *Soziale Welt* is for practitioners of social science in bureaucracies and the helping professions. Thus, the publics of the three "schools"

were different and have remained so: the Frankfurt school for the cultural intelligentsia; the Cologne school for social scientists; the Schelsky school for practical applications in welfare and bureaucracy. It was largely the pressure from the respective audiences more than the preferences of the professors that for a long time caused sociology in Germany to be divided into these three camps.

### THE TIME OF EXPANSION

The second most formative decade was the 1960s. It marked a return to important subjects of the classics and to some authors of that time. The Americanization of the discipline had peaked.

After the vituperous quarrels in the Deutsche Gesellschaft in the 1950s it was agreed to meet in closed session at Tübingen in 1961, to discuss basis issues in a purely scholarly atmosphere. Unexpectedly this led to a controversy between two radically different views of social science, the chief protagonists of which were Theodor W. Adorno and Karl Popper (Adorno 1962), between science as a vehicle for emancipation and a scientific view of science. Popper's work is in the philosophy of science rather conventional sociology. For Adorno empiricism means supremacy of instrumental reasoning, which subjugates reason to the rule of facts; empiricism is tantamount to "treating facts as fetishes" (Adorno 1969, p. 14). This dispute was continued by two nonsociologists—the economist Hans Albert and the philosopher Jürgen Habermas—as the "positivism controversy," Habermas representing the Frankfurt school of critical theory and Albert the Cologne school of neopositivism (Adorno 1975). This was a revival of the methodological controversies in the Verein für Socialpolitik in the decade prior to 1914.

The controversy between Habermas und Albert occurred during the 1964 sociology convention that was dedicated to the rediscovery of Max Weber. The topic that sent shock waves throughout the cultural establishment (Stammer 1965) was the revival of the debate on value judgments as part of science. The "radical" sociologists who began to appear at this time took as their central argument the charge that positivism was blind to the forces its research served.

The repercussions were much larger: the arguments of critical theory against positivism became

a credo among the cultural intelligentsia. It is important for an understanding of sociology in Germany that this old controversy did not spring from developments within the discipline; rather, it became an unavoidable topic because the intellectual environment forced it onto sociology.

Meanwhile, a seeming alternative to the Cologne and Frankfurt Schools, though closely related to the latter in its concerns, became the major intellectual success: Ralf Dahrendorf with his *Gesellschaft und Demokratie in Deutschland* (1965). How was the Nazi regime possible? Could it reappear in a different guise? This topic keeps emerging in Germany as a central focus in intellectual attempts at self-analysis. In the 1980s the intellectual public expected answers from a controversy among historians; in the 1960s the expectations were focused on sociology and Dahrendorf.

Dahrendorf, trained as a philosopher, is a self-taught sociologist. By temperament he is a moralist, as is evident in his dissertation exploring the idea of justice in Marx's thought (1954). He was subsequently among the Young Turks who concentrated on industrial sociology, and published an immensely successful book on the subject (Dahrendorf 1956) that overshadowed a more original monograph by Heinrich Popitz, Hans Paul Bahrdt, Ernst-August Jüres and Hanno Kesting (1957) that is yet to be recognized internationally as a classic. Among younger sociologists Dahrendorf became a central focus of controversy as a result of his long essay on role theory, in which he characterizes modern society as a cage of obligations that prevent individuality from asserting itself (Dahrendorf 1959b). While the profession largely rejected this notion as a misperception of role theory and economics, the book *Homo Sociologicus* remained for decades the most popular sociological treatise among students (Dahrendorf, 1959b). An essay on chances for complete social equality was Dahrendorf's contribution to the topic of social stratification (Dahrendorf 1961), then dominant among the Young Turks. A much more important publication for an understanding of his concerns is his treatment of the United States as the one case where presumably there was an attempt to construct a society in the spirit of the French Enlightenment (Dahrendorf 1963). Characteristically, Dahrendorf takes as his point of departure a source of general intellectual importance, in this case Alexis de Tocqueville's report on the United States in the 1830s.

As Adorno had introduced the concerns of the cultural public into the profession, so Dahrendorf did the same for the intelligentsia concerned with public affairs. In the English-speaking world he is often included in sociological curricula as a conflict theoretician, but there is no conflict theory in his writings, except for an attempt at a taxonomy of conflicts. The label *conflict theoretician* was affixed because of his criticism of Talcott Parsons, whose structural functionalism he accused of glorifying social harmony. Dahrendorf's central theme is to explore the chances for the values of British liberalism to become the guiding principle in society. In practical life this orientation met with mixed success.

In the 1970s Dahrendorf returned to England to become the first foreign president of the London School of Economics. He later became a professor at Oxford and remains a respected commentator on public affairs with a continuing interest in sociology as an intellectual endeavor. Although frequently in error in his statements that have empirical content, he is invariably sensitive in choosing topics that relate the discipline to an intellectual public.

The mood of the times was moving away from the liberal creed of Dahrendorf. The controversies in the United States, stimulated by the Vietnam intervention, were taken up by politicized students in Germany. Marxism was believed to provide answers that the liberal promise of the postwar Western world apparently could not. The varieties of Marxism were studied with religious fervor.

At the 1968 convention of the German Sociological Association demands for an alternative to mainstream sociology, especially the "Cologne Americanism," exploded. Adorno had chosen as the congress's theme, "Late Capitalism or Industrial Society"—alternative ways of conceptualizing the same reality. The term "late capitalism" was favored by Neo-marxists as implying that the demise of bourgeois society was imminent. It also implied that the notion of a basic sameness of all industrial societies—that the United States and the Soviet Union were structurally related as industrial societies—was wrong. The Neo-Marxist conceptualization "late capitalism" and, by contrast, critical theory implied that the social orders of the Soviet Union and the United States were

dedicated to a different telos, and that this difference was what mattered most (Adorno 1969).

With the May 1968 uprising in Paris, the student movement in Germany turned into a crusade for leftism as the mandatory creed on campus, and specifically for sociologists. For a moment the adherents of the Frankfurt school could see the student movement as the realization of their hopes. Max Horkheimer had always rejected any methodological constraints on philosophizing, and on any scholarship. Methodology would restrict science to the factual, while for Horkheimer, Adorno, and Bloch the prime goal of scholarship was creative utopianism. According to Horkheimer, it was now up to science to provide the answers given earlier by religion (Horkheimer 1968). The student movement, however, accused the Frankfurt school of running away from reality, and critical theory of having only negative messages and no direction for positive action. Adorno had always maintained this view (Adorno 1969). Deposing a ruling class would not change the basic character of relations that were permeated with instrumental reasoning; there was no longer a revolutionary subject to realize the cause of emancipation. After some very ugly clashes, Adorno died in 1969. Dahrendorf turned against the student movement, as did such well-known sociologists as Helmut Schelsky and Helmut Schoeck. Erwin Scheuch's *Anabaptists of the "Affluent Society"* (1968) became a best-seller. Even Habermas, irritated by the uncompromising nature of the New Left, severed his connection by diagnosing its "left fascism" (Habermas 1969).

In the turbulent years that followed 1968, sociologists were more concerned with reacting to the New Left than with developments of their own. Sociology had been utterly surprised by a student movement that it neither predicted nor understood. Rainer Lepsius contrasts this situation with the immediate success of empirical sociology in analyzing the right-wing protest NPD party (National-Demokratische Partei Deutschlands) of the late 1960s (Lepsius 1976, p. 7).

The public, however, identified sociology with the New Left. An explosion of leftist literature, using sociological terminology, buried the output of sociologists. The public image of sociology became one of a haven for radicals opposed to bourgeois society. For about ten years after 1968,

sociology became the quarry for cultural discourse. Academic fields that were based on shared beliefs in civil society and cultural values—such as pedagogy, art appreciation, literary criticism, and political education—had their criteria damaged, even destroyed. Sociology was used to provide an alternative rationale: service to the cause of emancipation from bourgeois society. Tenbruck (1984) has charged that sociology had "colonized" the humanities, and even life in general. It actually was the other way around: as the humanities lost their belief in civil society, they raided sociology for arguments in the name of society. Professional sociology completely lost control over the use of its vocabulary—with disastrous results for the self-selection of students of sociology and the standing of the field in the world of scholarship.

The student protest movement, and various alternative cultures loosely associated with it, became a regular part of public life in Germany, as it had in the United States. It became fashionable to call a great number of protest forms *sociology* while denouncing the profession carrying out "normal science" (Kuhn 1970) as "bourgeois" sociology. Sociologists were asked to react to fashionable topics of the day, such as permissive education, Third World dependency, mind-expanding drugs, feminism, gay power, autonomous living. Many of the newly appointed professors went along, contributing to the erroneous public impression that sociology was spearheading the cultural revolution in the name of anticapitalism. The opposite was true: sociologists often caved in to demands from the protest movements.

Even in a less turbulent intellectual environment, the period between 1968 and 1973 would have been most unsettling, because it was a period of unique expansion. By the mid-1960s social democratic state governments had been convinced by proponents of the Frankfurt school that sociology should be included in the curricula of secondary education. The profession was divided on this, and as late as 1959 Dahrendorf counseled against the inclusion of sociology in degree curricula even at universities. With the youth, however, sociology was a huge success; many teachers were needed, and that meant many more professors. In 1968 there were fifty-five tenured positions as professor of sociology in the Federal Republic of Germany; in 1973 there were 190 (Lamnek 1991). From that

year, the rate of expansion levelled, and by 1980 there were 252 “chairs” (Sahner 1982, p. 79).

Lepsius modified this feeling of an avalanche of sociology overwhelming academe (Lepsius 1976, p. 12). At the beginning of the 1970s, sociology had 1.3 percent of all the positions in university budgets, no larger a share than ten years earlier; 1.2 percent of the students took a major in sociology, also the percentage before the expansion; public money for sociological research had always hovered around 1 to 2 percent of all grants. The universities had expanded with explosive rapidity, and in this general explosion sociology merely kept its former share.

As a discipline, sociology, however, was among those least prepared to keep its place during an expansion that within a few years quadrupled the number of students in the Federal Republic. Consequently a great many sociologists were appointed to tenured positions who in other times would not have been. In this process the discipline lost cohesion and common scholarly standards.

### THE CONSOLIDATION

After 1968 the Deutsche Gesellschaft für Soziologie was taken over by the Young Turks; from their ranks Dahrendorf, Scheuch, and Lepsius were successively president. The leadership of the association was tired of conventions serving as forums for all kinds of protests, thus further damaging the reputation of the field. Consequently, biannual sociological conventions were suspended until 1974, when a meeting was held in Kassel. While the field had a Babel of views, there was now agreement to coexist.

At the Kassel convention of 1974, a custom was started called *Theorievergleich* (comparison of theory). This is in reality a juxtaposition of sociological “denominations,” not a weighing of alternative theoretical propositions (Lepsius 1976, sec. II). At that time four such denominations were defined: (a) behaviorism, the chief proponent of which in Germany was Karl-Dieter Opp; (b) action theory as represented by Hans-Joachim Hummel; (c) functionalism and systems theory, with Niklas Luhman as its prominent representative; and (d) historical materialism, which translates into Marxism. Somewhat later, parts of behaviorism and action theory amalgamated to

become rational choice; two German-speaking Dutch sociologists, Reinhard Wippler and Siegwart Lindenberg, are its best-known proponents. Phenomenological sociology has been successful especially among young sociologists, with Jürgen Helle as the chief representative.

The lines between these denominations keep shifting as new variants emerge. In general, though, this approach to theory—to choose a topic such as evolution and then listen to what each denomination has to say about it—appears to have spent itself.

Concern with theory in Germany again means reacting to theory builders, a most dissimilar pair of whom dominated the scene in the 1970s and early 1980s: Jürgen Habermas and Niklas Luhmann. Luhmann is a self-taught sociologist who studied law and became a career administrator. His first specialty was the sociology of organization, which showed a strong influence of Chester I. Barnard. Subsequently Luhmann analyzed phenomena of the *Lebenswelt* (the key term of phenomenology for the world of immediate, unreflected experience—in contrast to the world that science portrayed) at that time influenced by the phenomenologist Edmund Husserl, as in Luhmann’s monograph on trust as the basis for social cohesion. Luhmann later met Talcott Parsons, whose work he understood from the perspective of a systems theorist. From then on, he focused on pure theory, going so far as to reject the application of theoretical statements to the empirical world, declaring empirical evidence to be irrelevant for his theoretizing (Luhmann 1987). Since Luhmann writes prolifically in English, his views should be well known outside Germany.

It will be helpful to note his shift in emphasis over the years. The central notion was first functional differentiation as the guiding principle in evolution. This results in an increasing competence of the system if the functionally differentiated areas are allowed to develop their area-specific rationality (*Eigenlogik*). The economy is seen as a prime example of an *Eigendynamik*, provided it is not shackled by attempts at political or ethical guidance. Differentiated systems, according to Luhman, need constant feedback to permit a creative reaction. Luhman conceptualizes these feedbacks as *selbstreferenzielle Prozesse* (self-referential processes). More recently he has become interested in

chaos theory and has given self-organization—which he calls *Autopoiesis*—a central place in explaining system functioning. It appears that the interest in Luhman has increased as his writings have become more abstract and his terms more outlandish—something he does self-consciously, since he can be a lucid writer.

This is less true for Jürgen Habermas, a philosopher who is self-taught in sociology. He calls his approach *critical sociology*, a choice that expresses his initial indebtedness to left Hegelianism of the Frankfurt school. While he has included Marxist terminology in his writings, he is probably best understood as a sociological disciple of the idealistic philosopher Fichte. In this perspective society is a problem for man's true calling: emancipation. He understands emancipation in the spirit of the French Enlightenment. The characteristic element of Habermas's "critical sociology" is *doppelte Reflexivität* (double feedback): the sociologist reflects on the context of discovery, and again on the context of utilization of his findings. With this attitude Habermas approaches the problematic relationship between theory and praxis (a Hegelian variant of practice) at a time when human existence is determined by a technological *Eigendynamik*. Habermas's writings, available in translation, are internationally known. Viewed over a period of around thirty years, it seems that he is sliding into a position of Great Cultural Theory, of the kind like Pitirim Sorokin's. And in becoming a synthesizer, he becomes more empirically minded, as Luhman ascends into the more abstract.

### SOCIOLOGY AS "NORMAL SCIENCE"

While practically all German universities offer degrees in sociology up to the doctoral level, there are some centers in terms of number of students, research facilities, and number of teaching personnel. In terms of faculty size, these centers are Bielefeld, Berlin, and Frankfurt, Munich, Cologne, and Mannheim. If one includes among the criteria the number of degrees granted, then Bochum, Hamburg, and Göttingen must be added as centers.

There is now a very developed infrastructure for empirical research. In addition to many public service research institutes both inside and outside universities, there are some 165 commercial institutes for market and social research, the largest employing 600 academics and having a business

volume of 90 million dollars. There is an academic network of three service institutions with a yearly volume of ten million dollars from the budgets of state and federal governments: the Informationszentrum, in Bonn, providing on-line information on research projects and literature abstracts; ZUMA, Zentrum für Umfragen, Methoden und Analysen, in Mannheim, doing some of the work of NORC, National Opinion Research Center, Chicago, and helping anyone in need of methodological support; and the Zentralarchiv in Cologne, which provides data for secondary analysis, as the Roper Center in the United States does. All three together form GESIS, Gesellschaft sozialwissenschaftliche Infrastruktureinrichtungen, Bonn, a package that can be used by the mostly small research institutes of German universities. There are also larger institutes, the biggest in terms of number of academics employed being the Deutsche Jugendinstitut (German Youth Institute) and the largest in terms of finance (eleven million dollars) being the Wissenschaftszentrum Berlin (Social Science Center in Berlin). All have regular budgets from tax money.

There is a yearly survey on work in progress, and between 1978 and 1988, 29,000 research projects were reported. The Zentralarchiv keeps a count on quantitative research, and limiting its attention only to such work where data are in machine readable format, the yearly academic production in quantitative research is around 900 projects. The preferred method is the personal interview, which during the years had a share of 50 percent for all projects where quantitative data were collected. The doorstep interview is being replaced by the telephone interview, although to a lesser extent than in the United States.

Among the more than 1,000 members of the German Sociological Association the attention given to applied fields is dominant. A content analysis of journal articles shows that only 15.4 percent of all published manuscripts—the rejection rate is around 80 percent of all articles submitted—deal with theory. The most important applied fields are (in order of frequency) industrial sociology, social psychology, methods of quantitative research, sociology of politics, and sociology of the family. All these have shares in publications above five percent. Twenty other fields of sociology contribute altogether 50 percent to journal publications (Sahner 1982).

Heinz Sahner's (1991) content analysis of the theoretical paradigms used in journal articles between 1970 and 1987 found the expected correlation with the intellectual climate. Marxism declined rapidly from the late 1970s, and it never dominated in German professional journals. Structural functional sociology is still the single most important theoretical paradigm, with phenomenological approaches gaining steadily. Even more so than earlier, sociology in Germany is now a multiparadigm field, including as a prominent part, "German Sociology."

## REFERENCES

- Abel, Theodore 1929 *Systematic Sociology in Germany*. PhD diss. Columbia University, New York.
- Adair-Toteff, Christopher 1994 "Kant's Philosophical Influence on Classical German Sociology: Simmel's 'Exkurs über das Problem: Wie ist Gesellschaft möglich?'" *Simmel Newsletter* 4:308.
- Adorno, Theodor W. 1962 "Zur Logik der Sozialwissenschaften." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 14:249-263.
- 1975 *Der Positivismusstreit in der deutschen Soziologie*, 4th ed. Neuwied: Luchterband.
- , (ed.) 1969 *Spätkapitalismus oder Industriegesellschaft. Verhandlungen des 16. Deutschen Soziologentages*. Stuttgart: Entee.
- Alemann, Heine von 1978 "Geschichte und Arbeitsweise des Forschungsinstituts für Sozialwissenschaften in Köln." In Erwin K. Scheuch and Heine von Alemann, eds., *Das Forschungsinstitut*. Erlangen: Universität, Institut für Gesellschaft und Wissenschaft.
- Aron, Raymond 1935 *La sociologie allemande contemporaine*. Paris: pr. univer. de France. Translated into English as *German Sociology*. New York, 1964.
- Barnes, Harry E., and Howard Becker 1938 *Social Thought from Love to Science*, vol. 2. New York: D. C. Heath.
- Bernhard Schafers 1995 *Soziologie in Deutschland. Entwicklung, Institutionalisierung und Berufsfelder. Theoretische Kontroversen (Sociology in Germany. Development, Institutionalization and Professional Fields. Theoretical Controversies)*. Opladen, Germany: Leske and Budrich.
- Best, Heinrich and Renate Ohly 1994 "From Paradigms to Eclecticism: Thematic Profiles of German Language Core Sociology Journals 1984-1991." *BMS, Bulletin de Methodologie Sociologique* 42:95-113.
- Coser, Lewis A. 1977 *Masters of Sociological Thought*, 2nd ed. New York: Harcourt Brace Jovanovich.
- Dahrendorf, Ralf 1954 *Marx in Perspektive*. Hannover: Dietz.
- 1956 *Industrie- und Betriebssoziologie*. Berlin: De Gruyter.
- 1959a "Betrachtungen zu einigen Aspekten der gegenwärtigen deutschen Soziologie." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 11:132-153.
- 1959b *Homo sociologicus*. Opladen: Westdeutscher Verlag.
- 1961 *Über den Ursprung der Ungleichheit unter den Menschen*. Tübingen: Mohr.
- 1963 *Die angewandte Aufklärung*. Munich: Piper.
- 1965 *Gesellschaft und Demokratie in Deutschland*. Munich: Piper.
- Durkheim, Emile 1887 "La philosophie dans les universités allemandes." In *Revue internationale de l'enseignement*, Bd. 13.
- Eisenstadt, Shmuel Noah, and Miriam Curelaru 1976 *The Form of Sociology*. New York: Wiley.
- Fletcher, Ronald 1971 *The Making of Sociology*, 2 vols. London: Nelson.
- Freund, Julien 1978 "German Sociology in the Time of Max Weber." In Tom Bottomore and Robert Nisbet, eds., *A History of Sociological Analysis*. New York: Basic Books.
- Gehlen, Arnold, and Helmut Schelsky 1955 *Sociologie: Ein Lehr- und Handbuch*. Düsseldorf Köln: Diederich Verlag.
- Gephardt, Werner 1982 "Soziologie im Aufbruch." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 34:1-24.
- Gumplovitz, Ludwig 1883 *Der Rassenkampf: Sociologische Untersuchungen*. Innsbruck: Wagner.
- Habermas, Jürgen 1969 *Protestbewegung und Hochschulreform*. Frankfurt a.M.: Suhrkamp.
- , and William Rehg 1996 *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, Mass.: MIT Press.
- Hagemann-White, Carol 1997 "Current Debates in German Social Science." *Contemporary Sociology* 26:556-559.
- Hartmann, Jürgen 1992 "Portrait. In Memoriam Rene König (1905-1992)." *International Sociology* 7:481-483.
- Holub, Robert C. 1991 *Jürgen Habermas: Critic in the Public Sphere*. New York: Routledge.
- Horkheimer, Max 1968 *Kritische Theorie*. In Alfred Schmidt, ed., *Eine Dokumentation*. Frankfurt: S. Fischer. a.M.
- Hradil, Stefan 1993 "New German Social Structure Analysis." *Schweizerische Zeitschrift für Soziologie / Revue Suisse de sociologie* 19:663-688.

- Jay, Martin 1973 *The Dialectical Imagination*. London: Heinemann.
- Käsler, Dirk 1984 *Die frühe deutsche Soziologie 1909–1934*. Opladen: Westdeutscher Verlag.
- 1985 *Soziologische Abenteuer*. Opladen: Westdeutscher Verlag.
- Kettler, David, and Volker Meja 1994 “That Typically German Kind of Sociology which Verges towards Philosophy: The Dispute about Ideology and Utopia in the United States.” *Sociological Theory* 12:279–303.
- King, Michael, and Anton Schutz 1994 “The Ambitious Modesty of Niklas Luhmann.” *Journal of Law and Society* 21:261–287.
- König, René 1941 *Niccolo Machiavelli: Zur Krisenanalyse einer Zeitenwende*. Erlench-Zürich: Rentsch.
- 1949 *Soziologie heute*. Zürich: Regio-Verlag.
- 1958 *Sociologie*. Frankfurt: Fischer.
- 1962 *Handbuch der empirischen Sozialforschung*. Stuttgart: Enke.
- 1987 *Soziologie in Deutschland*. Munich: Hauser-Verlag.
- (ed.) 1978 *Soziologie*, 11th ed. Frankfurt: Fischer.
- Kuhn, Thomas Samuel 1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lamnek, Siegfried 1991 “Gesellschaftliche Interessen und Geschichte der Ausbildung.” In Harald Kerber, and Arnold Schmieder, eds., *Soziologie Theorien, Arbeitsfelder, Ausbildung: Ein Grundkurs*. Reinbek and Hamburg: Rowolt.
- Lepenes, Wolfgang (ed.) 1981 *Geschichte der Soziologie*, 4 vols. Frankfurt: Suhrkamp.
- Lepsius, Mario Rainer (ed.) 1976 *Zwischenbilanz der Soziologie*. Stuttgart: Enke.
- 1987 “Sociology in the Interwar Period.” In Volker Meja, Dieter Misgeld, and Nico Stehr, eds., *Modern German Sociology*. New York: Columbia University Press.
- Luhmann, Niklas 1987 “Soziologische Aufklärung.” In vol. 4, *Beiträge zur funktionalen Differenzierung der Gesellschaft*. Opladen: Westdeutscher Verlag.
- 1995 *Social Systems*. Translated by John Bednarz, Jr., with Dick Baecker. Stanford, Calif.: Stanford University Press.
- Meja, Volker, Dieter Misgeld, and Nico Stehl (eds.) 1987 *Modern German Sociology*. New York: Publisher.
- Muller, Hans-Peter 1992 “German Sociology at the Beginning of the 90s.” *Schweizerische Zeitschrift für Soziologie / Revue Suisse de sociologie* 18:751–762.
- 1995 “Fateful Encounter? German Sociology Meets Culture.” *Schweizerische Zeitschrift für Soziologie / Revue Suisse de sociologie* 21:189–198.
- Muller, Harro 1994 “Luhmann’s Systems Theory as a Theory of Modernity.” *New German Critique* 61:39–54.
- Nisbet, Robert A. 1966 *The Sociological Tradition*. London: Heinemann.
- Oberschall, Anthony 1965 *Empirical Social Research in Germany, 1848–1914*. Paris: Mouton.
- 1972 *The Establishment of Empirical Sociology*. New York: Harper and Row.
- Paterson, John 1997 “An Introduction to Luhmann.” *Theory, Culture and Society* 14:37–39.
- Popitz, Heinrich, et al. 1957 *Das Gesellschaftsbild des Arbeiters*. Tübingen: Mohr.
- Ratzenhofer, Gustav 1893 *Wesen und Zweck der Politik*. Leipzig: F. A. Brockhaus.
- Sahner, Heinz 1982 *Theorie und Forschung*. Opladen: Westdeutscher Verlag.
- 1991 “Paradigms Gained—Paradigms Lost: Die Entwicklung der Nachkriegssoziologie im Spiegel der Fachzeitschriften.” In *Soziale Welt*, Index for vols. 1–40.
- Salomon, Albert 1945 “German Sociology.” In Georges Gurvitch and Wilbert E. Moore, eds., *Twentieth Century Sociology*. New York: Basic Books.
- Salter, Michael 1997 “Habermas’s New Contribution to Legal Scholarship.” *Journal of Law and Society* 24:285–305.
- Schad, Susanna Petra 1972 *Empirical Social Research in Weimar Germany*. Paris: Mouton.
- Schäffle, Albert 1885 *Gesammelte Aufsätze*. Tübingen: J. C. B. Mohr.
- Scheuch, Erwin K. 1968 *Anabaptists of the “Affluent Society.”*
- 1990a “Von der deutschen Soziologie zur Soziologie in Deutschland.” *Österreichische Zeitschrift für Soziologie* 15:30–50.
- 1990b “Von der Pioniertät zur Institution.” In Dieter Franke and Joachim Scharioth, eds., *40 Jahre Markt- und Sozialforschung in der Bundesrepublik Deutschland*. Munich: Oldenbourg Verlag.
- Simmel, Georg 1890 *Über Soziale Differenzierung: Soziologie und psychologie Untersuchungen*. Leipzig: Duncker and Humblod.
- Stammer, Otto (ed.) 1965 *Max Weber und die Soziologie heute: Verhandlungen des 15. Deutschen Sociologentages*. Tübingen: Mohr.
- Tenbruck, Friedrich H. 1984 *Die unbewältigten Sozialwissenschaften oder die Abschaffung des Menschen*. Graz: Styria.

- Tönnies, Ferdinand (1887) 1912 *Gemeinschaft und Gesellschaft*. 2nd ed. Berlin: Carl Curtius.
- Vanberg, Viktor 1975 *Die Zwei Soziologien: Individualismus und Kollektivismus in der Sozialtheorie*.
- Vierckandt, Alfred 1931 *Handwörterbuch der Soziologie*. Stuttgart: Entze.
- Wagner, Gerhard 1997 "The End of Luhmann's Social Systems Theory." *Philosophy of the Social Sciences* 27:387-409.
- Wiese, Leopold von 1926 *Soziologie: Geschichte und Hauptprobleme*. Berlin: de Gruyter.
- 1959a "Die Deutsche Gesellschaft für Soziologie." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 11:11-20.
- 1959b "Deutsche Gesellschaft für Soziologie." In *Handwörterbuch für Sozialwissenschaften*, vol. 2. Stuttgart: Fischer.
- Zeitlin, Irving M. 1981 *Ideology and the Development of Sociological Theory*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.

ERWIN K. SCHEUCH

## GLOBAL SYSTEMS ANALYSIS

See Globalization and Global Systems Analysis.

## GLOBALIZATION AND GLOBAL SYSTEMS ANALYSIS

Since the early 1960s, mounting empirical pressure has forced sociology to abandon the assumption that national societies could be understood without looking beyond their borders. The nation-state remains a crucial unit of analysis, but it must be analyzed as intertwined with the operations of a larger global social system. Global flows of culture, technology, people, goods, and capital determines to an ever larger degree how societies change at the national level. The dynamics of transnational social, economic and political structures have become a focus of study in their own right.

The value of international trade has grown more rapidly than the value of goods and services produced and sold within national boundaries.

Manufacturing means assembly of components from around the globe. Electronic youth culture in the United States revolves around Japanese cartoon characters, while a generation of Third World television viewers take their cultural cues from North American TV serials. Capital markets operate around the globe and around the clock as trading moves from Tokyo to London to New York over the course of each day. Sociologists cannot yet claim to understand the working of this global system, but a number of fruitful avenues of analysis have emerged since about 1960.

The approaches favored by sociologists have differed from the study of "international relations" as it has been traditionally defined. International relations approaches have seen the global system as structured primarily by interactions in which nations are unitary actors (see Waltz 1979). The global system is presumed to be structured largely by the distribution of national power among the advanced nations. Sociologists, on the other hand, have been very interested in how the global system is structured by flows of resources, people, ideas, and attitudes across geographic boundaries, flows that often occur without, or even in spite of, national actors.

Within sociology itself, "globalization" has been transformed from a relatively narrow field of study, largely within the political economy of development, to a central theme across a wide range of subfields within the discipline. The topic of globalization has also often drawn sociologists into interdisciplinary debates with political scientists, anthropologists, geographers, and even occasionally economists, among others. As recently as the late 1980s, a number of central approaches to global systems analysis could be identified—Marxism, modernization theory, dependency theory, and world-system theory. However, although these research traditions continue to be elaborated, studies of globalization have proliferated and research perspectives have become fragmented. This article reviews the early and continuing contributions of these perspectives and outlines the main themes that current global systems analyses are exploring.

### MARXISM AND MODERNIZATION THEORY

In the years after World War II, Marxism and modernization theory debated vigorously the problems and consequences of capitalist development.



However, despite their ideological opposition, they shared some basic assumptions regarding the character of this development. In particular, they shared the assumption that the global system is dominated by processes of diffusion. In the words of the *Communist Manifesto*, the expansive character of capitalist production “draws all, even the most barbarian, nations into civilization.” Warren (1980) argues from a Marxist perspective that it is only by the expansion of economic ties to core countries, principally in the form of importing more foreign capital, that poor nations are likely to be able to reduce the gap that separates them from rich nations. Sklair (1991) argues that a world-system organized through nation states is being superseded by a “global system” dominated by transnational economic, political, and social structures. He points to the emergence of a transnational capitalist class that organizes the world economy to its own benefit, in contrast to the world systems view of contending national capitalist classes (for an empirical study, see Bottomore and Brym 1989). Robinson (1998) argues that sociology must move beyond nation-state based analytical approaches to make “transnational social structure” its proper object of study, since individuals are predominantly defined by their position in such transnational economic, political, and cultural structures rather than by national or local characteristics—transnational structures dominated by capital. Other Marxist-oriented theorists emphasize the irrationality of capitalist competition on a global scale, since competition between national capitalisms creates a crisis of profitability for capital due to global overcapacity, which in turn leads to declining conditions for labor (R. Brenner 1998; Walker 1999). Nonetheless, the emphasis remains the dilemmas of the diffusion of capitalism—the fact of that diffusion is presumed.

Investigating the prevalence of that complex of ideas and attitudes associated with “modernity” provided an important impetus for modernization theorists to take a global look at the diffusion of culture and social structures. In a classic study, Inkeles and Smith (1974) analyzed the attitudes of citizens in six Third World countries spread around the globe. Modern attitudes were found in social contexts more characteristic of advanced industrial countries. The more time respondents had spent living in cities or working in factories, the more

their attitudes resembled those associated with the culture of advanced industrial countries. The findings suggested a global system structured largely by processes of diffusion. The implicit model was one of gradual convergence around a similar set of “modern” values and attitudes. The spread of modern social institutions helps inculcate the values and attitudes, and the values and attitudes reinforce the institutions.

The work of Meyer and associates (Meyer 1980; Meyer et al. 1997a) shows how ideas and institutions originating in the advanced industrial countries become embodied in a global culture, which in turn shapes local institutions in all countries. Beginning from the surprising homogeneity of national political and economic institutional forms around the world, Meyer and associates (1997a) argue that contemporary actors, including nation-states, organize and legitimate themselves in terms of highly rationalistic, universalistic, world-cultural models. These models largely conform to the prescriptions of modernization theory, with core values such as citizenship, rationality, socioeconomic development, and so on. These models are diffused through the world system largely through the organizations and associations of “world society”—intergovernmental and nongovernmental transnational organizations which act to promote a “shared modernity” rooted in scientific rationalism and in which scientists and professionals are particularly dominant (Boli and Thomas 1999; Meyer et al. 1997a). Nation-states are not destroyed by globalization but, in contrast, the nation-state cultural form is diffused around the world through the institutions of world society.

An extensive body of empirical work has been developed in support of these theoretical insights. Boli-Bennett (1979), for example, examines the way in which national constitutions reflect global legal norms rather than local conditions. Ramirez and Boli (1982) look at the ways in which schools take on similar shapes around the globe as conceptions of what constitutes an effective educational institution come to be shared with surprising speed across geographical boundaries. Meyer and associates (1992) argue that world-society connections diffused mass education to former colonies, rather than education systems developing simply as a function of level of economic development. Meyer

and associates (1997b) examine how a world environmental regime emerged in recent decades through world-society associational processes, facilitated by the statelessness of this world society.

The interaction of these world-cultural forms with local contexts is unclear, however, as Meyer and associates (1997a) argue that there may be significant “decoupling” of these symbolic forms from the social practices in each context. Escobar (1995) argues that development discourses have relied exclusively on Western knowledge systems to the extent that non-Western knowledge systems are inherently marginalized by these “modernization” discourses. Ferguson (1990) shows that in Lesotho, the effect of modernization-inspired development policies was not to produce development but to depoliticize poverty and the further entrenchment of the state and Western modernizing influences. Other empirical work suggests that there are in fact a number of contesting models of the global economy, implying a more conflictual process than that depicted by Meyer and associates (Wade 1998).

Other social theorists take a more critical view of modernization while still placing the condition of modernity at the center of their analyses. Beck and associates (1994) have argued that advanced industrial societies are undergoing a new “reflexive modernization” that is closely tied to globalization. These authors see modernization bringing increased control over nature and society, as orthodox modernization theory predicts, but further rounds of modernization generating problems based precisely on its own success. The advance of rationalism and scientism remains central to their theories, but they are more skeptical about the ability of the experts to anticipate or control the side effects of their innovations. Pollution and other ecological disasters are critical political issues but also examples of how technical and social change can result in unanticipated system-level crises.

Giddens (1991) argues that modernity is inherently globalizing. Both are linked by the process of “disembedding”—the lifting out of social relations from local contexts of interaction. There are two main mechanisms of disembedding: symbolic tokens (universal media such as money) and expert systems (bodies of technical knowledge

that can be applied across a range of different contexts). Each produces a homogenization of social life across different contexts and relies on trust in absent repositories of expertise. With this trust, therefore, comes a strong element of risk, a risk that is an inherent element of the globalization process.

Beck (1992) places risk squarely at the center of his approach. For Beck, the populations of the advanced capitalist countries are living in a post-scarcity age where the major dilemmas are the side effects of success. Risks such as pollution are also inherently global in that they are not limited to the local contexts in which they are produced and also in that they have a tendency to affect all members of society relatively equally. We are also becoming increasingly conscious of these risks and therefore of our global interdependence with other societies. There is a new politics of risk—a politics that is concerned with the distribution not of “goods” but of “bads,” a politics that is not limited by territory and in which multiple groups of experts struggle to legitimate their expertise as the potential solution to risk management.

The globalization of modernist worldviews and their primary agents—scientists and professionals—is once again placed at the center of the analysis. However, Giddens and Beck add an understanding of the dilemmas associated with such globalized rationalist institutions. Lash and Urry (1994) point up the postmodernist emphasis on aesthetic symbols and signs, in contrast to the modernist emphasis on rationalism and the power of scientists and professionals. They direct our attention toward cultural consumption on a global scale rather than toward the diffusion of rationalist worldviews and their effects.

Sociologists concerned with modernity and modernization emphasize the diffusion of rationalistic organizational principles (such as the nation-state form) on a global scale rather than the spread of global capitalism. Individuals are culturally constituted by the global system as rationalized, individualized actors rather than as workers or members of classes, while capitalists take second place to experts as the prime movers of the historical process of globalization. Nonetheless, theorists of modernization and modernity share with Marxists the view of an increasingly rationalized world, for better or worse.

### DEPENDENCY THEORY

In contrast, sociologists who analyzed the global system from a dependency perspective emphasized the extent to which Third World political economies evolve differently from those of First World countries because they confront a world dominated by already industrialized countries. Deriving their initial inspiration from economists such as Prebisch (1950) and Baran (1957), the dependency approach saw the global system as consisting of a “core” of advanced industrial countries connected both economically and politically to a larger “periphery” of poor nations, rather than as consisting simply of a set of nations that can be ranked along various continua according to individual characteristics such as size and wealth. The structure of the global system was conceptualized primarily in terms of trade and capital flows reinforced by political domination. The principal concern was with the consequences of these ties for social, political, and economic change in the countries of the periphery.

Cardoso and Faletto’s *Dependency and Development in Latin America* (1979 [1969]) is still the classic exemplar of this tradition. Their analysis shows that the way in which economic elites are connected to the global economy shapes not only their strategies of investment but also their willingness to make political alliances with other groups and classes. For example, Latin American countries whose primary ties to the core were formed by mineral exports under the control of foreign capital experienced a different political history from countries who relied on agricultural exports controlled by local elites.

The expectations of the dependency approach with regard to changes in the structure of the global system over time stand in sharp contrast to those of both traditional Marxist and modernization approaches. The dependency perspective emphasizes the ways in which economic elites and their allies in the periphery have an interest in preventing the full diffusion of economic capacities from core to periphery. For example, those who have an interest in the system of trading agrarian exports for core-country manufactures may consider nascent local industrialists to be competitors for both labor and political power.

Empirical work generated by the dependency perspective suggested that the consequences of

core–periphery capital flows, rather than being the most important stimulus to the growth of the periphery as Marxists like Warren suggested, might in fact create obstacles to growth. Korzeniewicz and Moran (1997) show that overall world income inequality increased between 1965 and 1992, particularly in the 1980s—the most significant component of that inequality being between-country inequality. Initial cross-national quantitative analyses discovered that the buildup of stocks of foreign capital had negative, rather than the expected positive, consequences for growth (Bornschiefer et al. 1978; Chase-Dunn 1975). Firebaugh (1992, 1996), however, shows that the negative effects of foreign investment are in part a result of methodological problems with the analyses and that, while foreign investment is less beneficial than domestic investment, it has a generally positive short-term effect on growth. Dixon and Boswell (1996a, 1996b) reformulate the dependency argument to emphasize the negative externalities associated with foreign investment that dampen the productivity of domestic investment. Kentor (1998) finds that dependency on foreign capital has particularly harmful effects on growth over the long term (thirty years) through the distortion and disarticulation of the domestic political economy. Crowley and associates (1998), in a review of cross-national quantitative studies in economics and sociology, find that foreign investment has a short-term positive effect on growth but plays a more negative role over the long term. However, the dynamics through which foreign investment generates negative externalities within a political economy are not well understood (Dixon and Boswell 1996b). While the findings of these studies are still contested, they clearly demonstrate that the results of transnational capital flows are not those predicted by a simple diffusionist model.

The impact of trade liberalization has been just as controversial as that of foreign investment. O’Hearn’s (1990) longitudinal analysis of the Republic of Ireland shows that while foreign investment did not have a direct negative effect, free trade (which was inextricably tied to the foreign-investment policy) had substantial negative effects on growth and inequality. Crowley and associates (1998) review studies of the impact of trade on growth and find that while export reliance is generally positively related to growth, this positive effect is reduced or reversed under conditions that

approximate to those theorized as “unequal exchange” (Emmanuel 1972). Rodrik (1997), reviewing the economics literature, finds that unskilled workers are likely to experience increased labor-market insecurity in the face of free trade, that nations may have legitimate reasons to limit trade, and that national systems of social security are indeed threatened by trade liberalization.

Again, quantitative cross-national studies played a valuable role in specifying the consequences of industrializing in an already industrialized world. Chase Dunn (1975) found that a greater role for foreign capital was associated with high levels of inequality, and his findings were confirmed by a variety of subsequent studies (e.g., Evans and Timberlake 1980; Rubinson 1976). These studies also confirm the political consequences of global economic ties. Delacroix and Ragin (1981) show that peripheral status in terms of trade relations is associated with weak state apparatuses. Bornschieer and Ballmer-Cao (1979) argue on the basis of cross-national data that core-peripheral capital flows strengthen the political position of traditional power holders at the expense of labor and middle-class groups.

One of the criticisms that can be leveled against the dependency approach is that it has focused too much on capital and not enough on labor. Whereas dependency theorists have been concerned with the consequences of transnational capital flows for labor in both periphery and core (e.g., Frobel et al. 1981), the tradition contains no series of cross-national quantitative analyses of either wage levels or the structure of international labor flows studies comparable to the literature on international capital flows. With some notable exceptions (e.g., Portes and Walton 1981) the global structure of labor flows and their consequences is understudied. The 1990s saw a series of valuable case studies of work under globalization, although the consequences of work organization patterns in shaping patterns of dependency has not been explored in detail (Bonacich et al. 1994; Hodson 1998; Salzinger 1997).

Even within a global economy that is becoming more polarized, there have been significant cases of upward mobility. Spain and Portugal, which once formed the core, later moved to the semiperiphery, where they are joined by Taiwan

and Korea, which have moved up from the periphery (Korzeniewicz and Moran 1997). Although dependency theory has concentrated on the obstacles to mobility facing peripheral economies, it has also stimulated a wide range of research on the causes of mobility within the world-system. Analysts of “dependent development” (Cardoso 1974; Evans 1979) argue that intensification of ties with core countries is a dynamic element in reshaping the political economies of Third World countries. Dependent development shares with classic Marxist approaches the assumption that capital flows between core and periphery can play a significant role in generating industrialization in the Third World. It emphasizes, however, that both the economic character of this industrialization and its social and political concomitants are likely to be different from the experience of core countries.

The question of domestic dynamics is particularly important in relation to explanations of the mobility of individual nations from one position to another. While such mobility depends in part on changes in the system as a whole, it also depends on the outcome of domestic political struggles. In his work on Chile, for example, Zeitlin (1984) argues that the outcome of political struggles among economic elites in the nineteenth century was determinative of Chile’s retaining its role as an exporter of raw materials rather than moving in the direction of trying to transform its mineral resources into more processed exports. Zeitlin’s analysis of Chile shows how the outcome of domestic political contests can perpetuate peripheral status.

Brenner’s (1976, 1977) interpretation of England’s rise to the core provides another example of the domestic roots of systemic mobility. In Brenner’s view, the differences between the agrarian strategies of England on the one hand and Spain and Portugal on the other were not simply the result of trade possibilities generated by changes at the level of the world-system. They depended crucially on interactions between peasant communities and agrarian elites at the local level, which in turn were rooted in longstanding historical characteristics of the peasant communities themselves.

Likewise, Taiwan and Korea are used as examples of the way in which internal dynamics may allow construction of more effective state apparatuses, which in turn enable a country to improve its

position in the global system (Evans 1987). In the East Asian developmental states, and especially in Japan and Korea, the state sat at the center of an alliance between the large business groups, domestic banks, and certain key state agencies—prodding and poking firms and banks in more developmental directions, organizing political alliances, and mobilizing social resources (Amsden 1989; Evans 1995; Wade 1990). Research stimulated by the dependency perspective has been both a powerful corrective to views of globalization as an inevitable, relatively even process of diffusion from the core to the periphery. Recognizing the obstacles to diffusion has also problematized the cases where mobility has occurred and has focused analysts' attention on the forces shaping a nation's ability to overcome these obstacles.

### WORLD-SYSTEM THEORY

The dependency approach's contribution to our understanding of the international system has, however, been limited by the fact that it does not focus directly on the structure of the global system itself. The "world-system" approach, launched by Immanuel Wallerstein (1974) and others (see Chase-Dunn 1989) at the beginning of the 1970s, took the overall structure of the system as its starting point. Wallerstein's contribution lay not only in directing attention to analysis of the global system itself, but also in setting the contemporary capitalist world-system in the context of previous systems spanning more than one society.

In Wallerstein's world-system, the hierarchical structure is postulated as essential for its survival. The geographic expansion of northwestern Europe's economic and political influence beginning at the end of the "long" fourteenth century was, in Wallerstein's view, essential to the transformation of productive organization in that region. Subsequent interchange among regions with different modes of extraction has been central to sustaining the process of accumulation in the system as a whole. Wallerstein envisions at least three structural positions within the system. Defined in terms of the nature of their exchange relations with other regions, they are (1) the "core," which exports goods produced by processes more intensive in their use of capital and new technology; (2) the

periphery, which relies on the production of labor and resource-intensive goods; and (3) the semiperiphery, which "trades both ways". Arrighi (1994) emphasizes the role of finance in the world-system as a complement to the Wallersteinian focus on trade and the division of labor. He traces the process of capital accumulation across four systemic cycles of accumulation in the world-system. Each one is associated with a different hegemonic (state) power that coordinates the securing of the conditions for capital accumulation on a world scale.

Quantitative cross-national analysis provides support for the idea that countries can be categorized according to their interactions with other nations and that countries in the same position experience shared benefits (or costs). Snyder and Kick (1979), using "block modeling" techniques, found a block of nations whose characteristics corresponded roughly to those of the core, several blocks that corresponded to the periphery, and a set of nations with the intermediary properties attributed to the semiperiphery. Smith and White (1992) used a similar methodology but improved on Snyder and Kick by examining longitudinal changes in the structure of the system. They were able to demonstrate changes in the structure of relations within and between blocks of nations and chart the mobility of individual nations within the system. Van Rossem (1996), however, using an analysis of the role structure of the world economy, argues that world-system position has little direct effect on growth but is better seen as a framework within which nations can act rather than as a determining force in national economic development.

Some of the most interesting work stimulated by world-system thinking has involved analysis of systems that antedated the emergence of the contemporary one (e.g., Abu-Lughod 1989). Wallerstein argues that systems prior to the current one were primarily of two types. "Minisystems" extended across boundaries defined by unified political control and ethnic solidarity but did not come close to being global in scope. These minisystems have been the subject of a number of recent studies, with interesting comparisons across different world-systems (Chase-Dunn and Hall 1997; Chase-Dunn and Mann 1998). "World empires" joined various ethnic and social groups in a single division of

labor by extending political control over a broader geographic area. Only the contemporary capitalist world-system, however, unites such a broad geographic region (essentially the entire globe) in a single division of labor without unified political control of corresponding scope. Frank (1998) has challenged the Eurocentrism of world-systems theory, arguing that the Western European capitalist heyday is but an interlude in the dominance of Asia in the world economy. Arrighi and Silver (1999) compare the contemporary period to the two major hegemonic transitions—from Dutch to British power in the 1700s and from British to U.S. power in the early twentieth century. They argue that the financial expansion of the last two decades of the twentieth century, is not so much a sign of deepening globalization as of systemic crisis, with an uncertain outcome which depending largely on the response of the declining hegemonic power, the United States. The comparative perspective across geographic scales and historical time periods that world-systems theory can provide gives valuable insight into the novelty or otherwise of contemporary globalization processes.

Another interesting trend in current sociological work on the global system is the increasing concern of sociologists with the politics of international relations among states. The third volume of Wallerstein's (1989) epic analysis of the modern world-system makes it clear that the world-system perspective, often accused in the past of being excessively "economistic," is now focusing much more on the logic of interstate politics, an emphasis that is also central to Arrighi's work. At the same time, other major sociological figures have turned their attention to the political and military aspects of the international system (e.g., Giddens 1985; Mann 1988; Tilly 1992). There has also been some attempt within world-systems theory to incorporate the dynamics of households (Smith and Wallerstein 1992) and of social movements. The emphasis within the tradition remains firmly on structural, macro-level processes, despite an increasing attention to cultural constructions (Wallerstein 1990). World-systems theory does analyze how the core countries are affected by the development of the world-system—unlike in the previous perspectives. However, this analysis is concentrated on the struggle among core nations for hegemonic power within the world system.

There is very little analysis of the impact of peripheral development, or even of globalization processes more generally, on the core within any of the perspectives discussed so far.

A quite different systems perspective on globalization has been advanced by Robertson (1992). Building on a concern to link a Parsonian concept of social systems to an analysis of international relations, Robertson posits that globalization involves the structuration of a social system at the global level. He argues that this process of globalization has been underway since the early fifteenth century and has resulted in ever-increasing global complexity within the social system. He shares with world-systems theory a "long view" of the development of globalization tendencies. However, he rejects the "reductionism" and "economism" of world-systems theory in favor of an analysis of the cultural and social conditions of social order within a global system. Robertson argues that there is an intensification of global consciousness in the sense that individuals increasingly orient toward the world as a whole. This rise in global consciousness is sustained both by the greater material interdependence of people around the world and by the interplay of four components of the global social system: the individual self, the national society, the world-system of societies, and humankind (1992, p. 27). The self-definition of each element can no longer be sustained in isolation from the other elements—under conditions of globalization each component is constituted relative to the others, providing the defining feature of globalization.

#### **CURRENT GLOBAL SYSTEMS ANALYSIS**

Each of the "classical" perspectives on globalization has therefore generated insights into the process and continues to sustain a research tradition of its own. However, the theme of globalization has spread beyond the confines of these traditions to become a central organizing concept across a wide range of sociological perspectives and research. Since approximately the early 1980s, globalization has had an increasingly visible impact on the core countries themselves, shaping the concerns of sociologists located in those countries. Sociology itself has been at least partially globalized—both by the emergence of "indigenous" sociological

communities around the world and by the growing international communication between these communities (Albrow 1996). Previous eras of sociological research into the global system had been shaped by a concern for the impact of the core on the periphery. However, these new patterns of globalization and new locations and perspectives of sociologists have generated an interest in how specific components of that system are being rearticulated and in particular how the global, the local, and the national are being reconstituted (Waters 1995). This diverse body of research typically examines globalization by generating “theories of the middle range,” often as a reaction against what was seen as overly abstract or deterministic systems-level theories (Portes and Walton 1981). The diffusion of the concept of globalization through sociology and this concern for revealing the mechanisms of globalization were reflected in a fragmentation of perspectives in the 1990s. In the rest of this article, we will review some of this recent research, focusing on the issues of the globalization of production, the constitution of a “global culture” and the future of the state in a global economy.

Over the course of the 1970s and 1980s, a growing body of work emerged that took the influence of the global system seriously in its analysis of change at the domestic level while at the same time using analysis of domestic political struggles and their economic consequences to explain changing ties with the global system (see Evans and Stephens 1988). These attempts to integrate international and comparative analysis are not unique to sociology. A parallel trend can be observed in political science as well (see Gourevitch 1986; Putnam 1988).

In the 1990s, work on global and local production systems, while often informed by dependency theory, also took a somewhat different approach. Reich (1991) argues that “national champion” corporations are rapidly being transformed into “global webs,” or “virtual corporations,” that coordinate knowledge inputs from around the globe. Castells (1997) argues that the structural logic of society is increasingly expressed through a “space of flows” that stretches across the globe. Flows of capital, information technology, organizational interaction, images, sounds, and symbols are the expression of the dominant processes of global society. They are based on an electronic infrastructure,

supported through key nodes or locations, and directed by a dominant, managerial elite. Most people continue to live in the “space of places” where local context dominates their experience. However, the space of flows tends to dominate the space of places as function and power are increasingly organized in flows while experience remains rooted for the most part in places (Castells 1997, p. 428).

Organizational theory suggests, however, that the model of the virtual corporation is likely to be problematic as home and host country effects, corporate cultures, and sectoral characteristics still shape corporate strategies in significant ways. Gereffi advances the concept of “global commodity chains” (GCCs) as the organizational structure that ties together these various social and institutional contexts in a global production network (Gereffi 1994; Gereffi and Korzeniewicz 1994; Harrison 1994). These GCCs are dominated either by key producers or by dominant brand name marketing firms, creating “producer-driven” and “buyer-driven” GCCs. The overall organization of the GCCs and location within the GCC shape workers’ wages, conditions, and power in the workplace (Bonacich et al. 1994).

This perspective therefore goes well beyond perspectives that simply warn about capital mobility and virtual corporations. It also allows us to provide more nuanced accounts of the dynamics involved in territories’ dependency on foreign capital, although the analysis of the political contexts of GCCs has not yet been fully developed. This more nuanced analysis of global production structures also holds the potential for improved integration of the role of domestic dynamics in determining the position of individual nations within the system and, by extension, shaping the structure of the system itself.

Increased attention is also being paid to how, even in an era of globalization, long-established patterns of interaction and cultural forms can profoundly shape a nation’s ability to compete globally. The success of Japan, Korea, and Taiwan can, for example, be argued to be due in part to the contribution of communitarian, patrimonial, and patrilineal cultural logics to creating the organizational networks that are at the heart of these countries’ economic success (Orrú et al. 1997). A

country's historically shaped 'logic of social organization' can make it better suited to competition in particular industries—for example, a more decentralized, less hierarchical culture such as that of the United States may be more suitable for industries, such as software, that are based on knowledge sharing (Biggart and Guillen 1999; Guillen 1994).

Theorists of "industrial districts" have made a somewhat different point regarding the importance of local social and cultural practices in shaping global competitiveness. They argue that in an era of post-Fordist production, where information processing is critical and the need for "flexible specialization" central to competitiveness, local face-to-face relationships are necessary to build up the trust that supports such competitiveness (Piore and Sabel 1984). Initial versions of this approach pointed to long-established civic traditions, craft traditions, and kin relations as underpinning such organizational forms, in Italy in particular (Piore and Sabel 1984; Putnam 1993). Later research has explored the emergence of such cooperation and trust in "new" regions, such as Silicon Valley (Saxenian 1994), Hollywood (Storper 1997), Southern California (Scott 1993), and numerous other regions identified in a huge range of studies (Castells and Hall 1994). Sabel has explored at a theoretical level how such "constitutional orders" can develop or even be promoted by policy (Sabel 1996a).

Sassen explicitly attempts to integrate these local and global perspectives through an analysis of how labor flows are shaped by capital flows (Sassen 1988, 1998), how "global cities" arise as centralized nodes of control over decentralized production systems (Sassen 1990), how domestic structures of inequality are created by the demand for high-wage professionals and low-wage service workers (Sassen 1988, 1990), and how the state comes to play a critical role in globalization, even as it is transformed by the process (Sassen 1996, 1998). She extends this analysis to begin to analyze the incorporation of women into the global economy and the opportunities and threats posed for them by this process (Sassen 1998). Sassen demonstrates the interaction between transnational corporate networks and labor flows and territorial entities such as global cities and local and national states in shaping global capitalism.

This creative tension between local and global is also visible in research on "global culture." For many authors, the global reach of markets has brought with it the cultural construction of actors around the world as individualized consumers, often in association with the more general spread of U.S. cultural norms internationally (Barber 1995; Featherstone 1990; Ritzer 1993; Sklair 1991). Others see a reaction against this process in a return to "tribalism," in the increase in ethnic and nationalist conflicts and atrocities, and in the general struggle between "locals" caught in the space of places and "cosmopolitans" operating in the space of flows (Barber 1995; Castells 1997).

Each of these implies a vision of global culture as a homogenizing force, with local identities surfacing as a reaction against this domination or, in certain cases, exclusion. However, other authors emphasize how global cultural forms become transformed in the local culture and how local cultures form part of a more varied and heterogeneous global culture. Appadurai (1996) argues that global cultural flows are shaped by the multiplicity of perspectives generated by flows of people, money, technologies, ideologies, and media technologies and symbols. Working through these varied cultural landscapes, local cultures work to incorporate global symbols but in ways specific to the local context. There is no pure local culture that is untainted by global culture but rather a variety of local cultures that are increasingly interpenetrated and constantly remade out of elements of global cultural flows (Appadurai 1996; Hall 1991).

Research on migration and transnational communities provides an opportunity to examine the interaction of local and global forces in shaping "global culture." Hodagneu-Sotelo (1994) shows how migration can present opportunities for women to play a more central role in their communities and in mediating relations between the community and the state. Transnational communities of migrants emerge that may have huge economic and social impacts on the home and host countries by creating new identities and structuring flows of social and financial resources (Massey and Parrata 1998; Portes 1996). Another social group that has received much research attention, but rarely under the rubric of globalization, is professionals, who, as we have seen, are central to the process of globalization and are most likely to be integrated into global networks in the workplace (Reich 1991)



or in their migration patterns (Castells 1997). Finally, research into virtual communities promises to provide insights into the nature of community and shared culture in an ostensibly placeless cyberspace (Turkle 1995; Wellman et al. 1996).

Finally, we turn to the state, which is often claimed to be a doomed social actor in the face of globalization as mobile corporations, finance, migrants, and symbols undermine its economic, political, and cultural authority (Esping-Andersen 1996). Furthermore, globalization is one factor shifting the focus of politics away from the state and established interests and toward “new social movements” organized around new identities forged from “codes” from around the world (Melucci 1996). A variety of research, mainly in political economy, has aimed to show the continuing national diversity in socioeconomic organization and the persistent power of the state to shape economic outcomes (Boyer and Drache 1996; Hirst and Thompson 1996; Wade 1996). Indeed, globalization has in many ways been produced by states. Sassen (1998) argues that the state has played a key role in creating the international and local conditions for global production and capital accumulation. States have negotiated new legal regimes that secure the rights of capital on a global scale and have supported the strategic sites through which the global economy is organized. Economic sociology’s improved understanding of the social and political construction of markets can be applied very usefully to the continuing national diversity in state–market relations (Berger and Dore 1996). Fligstein (1990), for example, shows the domestic conditions for the increased U.S. corporate emphasis on finance, an emphasis that, at the end of the 1990s, threatened the East Asian economies (Wade 1998). The particular property rights and exchange rules underpinning the construction of a single market within the European Union have also been profoundly shaped by interstate negotiations (Fligstein and Mara-Drita 1996).

Neil Brenner (1998, 1999) argues that state structures and strategies are being reconstituted in order to mediate the global and local processes discussed above. The “glocal state” attempts to promote the global competitiveness of its major urban regions through increased ties to both local and international actors. Sassen (1999) points out that this remaking of certain parts of the state

apparatus to promote global competitiveness begins to shape the character of the state as a whole, as even those sections of the state still predominantly concerned with national social policy become accountable to the new priorities and as the national legal and institutional structure is changed to accommodate international capital. These processes are not inevitable, however, as the international hegemony of Anglo-American economic ideology contributes to the creation of a “leaner, meaner” state that retains the institutional capability to support capital even as its capacity to provide social services or to support civil society weakens (Evans 1997). This excessive global liberalization and integration of financial markets is likely to have a corrosive effect on domestic state capacities and social welfare (Block 1996).

The state is therefore remade by globalization, even as it plays a critical role in constituting the global economy. Sociologists are beginning to explore the implications of this reconstitution of the state for economic sovereignty and citizenship. Castells (1997) argues that a “network state” is emerging, a state in which sovereignty is pooled between increasingly intertwined local, national, and international levels of governance, such as in the European Union. Debates persist as to whether this network state means the eclipse of local or national institutions as European Union institutions increasingly come to predominate (Streeck and Schmitter 1991) or even as a “postnational society” is created in Europe (Habermas 1998). There are persistent examples within Europe, however, that local and national institutions continue to play a central role in generating economic and social progress. Recent examples include national corporatist institutions in the Netherlands (Visser and Hemerijck 1996), local “micro-corporatism” in Italy (Locke 1995; Regini 1995), and a combination of the two in the Republic of Ireland (Sabel 1996b).

States are clearly being transformed by both the localization and globalization of economic life. Boyer and Hollingsworth (1996) argue that economic action is best conceived of as “nested” within a combination of local, national, and international institutions, rather than as “embedded” within national regimes. Indeed, formal sovereignty is shifting away from the state in many cases—toward both privatized transnational legal

regimes (such as international commercial arbitration) and increasingly legitimate international human rights codes and instruments. These developments are also transforming the character of citizenship—on the one hand weakening certain provisions around economic and social citizenship (Sassen 1996) while on the other universalizing certain human rights (Soysal 1994). In short, then, the state remains a vital actor within the global system but one whose role, purpose, and structure are being transformed in new and unexpected ways.

Sociology itself is being transformed and globalized, if only partially up to this point. In the process, global systems analysis has moved from the specialization of a number of key theories to perhaps becoming a defining element of sociology in the twenty-first century. New methodologies are being developed to counter the abstraction of systems analysis and to provide insights into the process of globalization as revealed in local contexts and as practiced by social actors (Burawoy et al. forthcoming). What new theoretical and methodological traditions will emerge from this period of change remains unclear. Overall, our understanding of the global system must still be considered a project “under construction” rather than a finished set of tools easily applied to specific problems.

This is all the more true given the rapid changes in the global system itself. Some things are clear nonetheless. We know that trajectories of change in national societies cannot be analyzed without reference to the global system in which they are embedded any more than the analysis of change in individual communities can be attempted without awareness of the national society in which they are embedded; but we also know that the character of relations between an individual state and the larger system is shaped not just only by the evolution of the global system but also by the outcome of political struggles at the local level. We know that the nations located at the bottom of this structure are disadvantaged economically as well as politically, but we also know that mobility is possible. We know that diffusion of ideas and norms throughout the global system has a powerful influence on how social institutions are structured in individual nations, but we also know that this diffusion takes place within a system that has a very hierarchical structure. Increasingly, we understand that local

societies incorporate global cultural forms in complex and contested ways.

We know that the contemporary global system is an invention of the last half-dozen centuries; predicting how long it will endure is another question. For some, the future seems gloomy as the market threatens to undermine both state and society and lead to barbarism, as Polanyi (1944) suggested. For others, the future holds the possibility of the re-creation of civilizing institutions on a more comprehensive global scale. Still others seek to identify spaces for action within globalization processes, hoping to shape alternatives to the pessimistic Polanyian vision of creeping barbarism.

## REFERENCES

- Abu-Lughod, J. 1989 *Before European Hegemony*. New York: Oxford University Press.
- Albrow, M. 1996 *The Global Age*. Stanford, Calif.: Stanford University Press.
- Amsden, A. H. 1989 *Asia's Next Giant: South Korea and Late Industrialization*. Oxford: Oxford University Press.
- Appadurai, A. 1996 *Modernity at Large*. Minneapolis: University of Minnesota Press.
- Arrighi, G. 1994 *The Long Twentieth Century*. London: Verso.
- , and B. Silver 1999 *Chaos and Governance in the Modern World System*. Minneapolis: University of Minnesota Press.
- Baran, P. 1957 *The Political Economy of Growth*. New York: Monthly Review Press.
- Barber, B. R. 1995 *Jihad vs. McWorld*. New York: Times Books.
- Beck, U. 1992 *Risk Society: Towards a New Modernity*. London: Sage.
- , A. Giddens, and S. Lash 1994 *Reflexive Modernization*. Stanford, Calif.: Stanford University Press.
- Berger, S., and R. Dore, eds. 1996 *National Diversity and Global Capitalism*. Ithaca, N.Y.: Cornell University Press.
- Biggart, N., and M. Guillen 1999 “Developing Difference: Social Organization and the Rise of the Automotive Industries in South Korea, Taiwan, Spain, and Argentina.” *American Sociological Review*, forthcoming.
- Block, F. 1996 *The Vampire State: And Other Myths and Fallacies About the US Economy*. New York: The New Press.

- Boli, J., and G. M. Thomas 1999 *Constructing World Culture: International Nongovernmental Organizations Since 1875*. Stanford, Calif.: Stanford University Press.
- Boli-Bennett, J. 1979 "The Ideology of Expanding State Authority in National Constitutions, 1870-1970." In J. Meyer and M. Hannan, eds., *National Development and the World System*. Chicago: University of Chicago Press.
- Bonacich, E., et al., eds. 1994 *Global Production: The Apparel Industry in the Pacific Rim*. Philadelphia: Temple University Press.
- Bornschieer, V., and T.-H. Ballmer-Cao 1979 "Income Inequality: A Cross-National Study of the Relationships Between MNC-Penetration, Dimensions of Power-Structure and Income Distribution." *American Sociological Review* 44:487-506.
- Bornschieer, V., C. Chase-Dunn, and R. Rubinson 1978 "Cross-National Evidence of the Effects of Foreign Investment and Aid on Economic Growth and Inequality: A Survey of Findings and a Re-analysis." *American Journal of Sociology* 84:651-683.
- Bottomore, T., and R. J. Brym, eds. 1989 *The Capitalist Class: An International Study*. New York: New York University Press.
- Boyer, R., and D. Drache 1996 *States Against Markets: The Limits of Globalization*. London: Routledge.
- Boyer, R., and J. R. Hollingsworth 1996 "From National Embeddedness to Spatial and Institutional Nestedness" In R. Boyer and J. R. Hollingsworth, eds., *Contemporary Capitalism*. Cambridge, U.K.: Cambridge University Press.
- Brenner, N. 1998 "Global Cities, Global States: Global City Formation and State Territorial Restructuring in Contemporary Europe." *Review of International Political Economy* 5(1):1-37.
- 1999 "Beyond State-Centrism? Space, Territoriality and Geographical Scale in Globalization Studies." *Theory and Society* 28(1):39-78.
- Brenner, R. 1976 "Agrarian Class Structure and Economic Development in Pre-Industrial Europe." *Past and Present* 70:30-75.
- 1977 "The Origins of Capitalist Development: A Critique of Neo-Smithian Marxism." *New Left Review* 104:24-92.
- 1998 "The Economics of Global Turbulence." *New Left Review* 229:1-264.
- Burawoy, M., et al. forthcoming *Global Ethnography*. Berkeley: University of California Press.
- Cardoso, F. H. 1974 "As tradicoes de desenvolvimento-associado." *Estudos Cebrap* 8:41-75.
- , and Enzo Faletto 1979 (1969) *Dependency and Development in Latin America*. Berkeley: University of California Press.
- Castells, M. 1997 *The Information Age*, 3 vols. Oxford: Blackwell.
- , and P. Hall 1994 *Technopoles of the World*. London: Routledge.
- Chase-Dunn, Christopher 1975 "The Effect of International Dependence on Development and Inequality: A Cross-National Study." *American Sociological Review* 40(6):720-738.
- 1989 *Global Formation: Structures of the World Economy*. New York: Blackwell.
- , and T. D. Hall 1997 *Rise and Demise: Comparing World-Systems*. Boulder: Westview Press.
- , and K. M. Mann 1998 *The Wintu and their Neighbors: A Very Small World-System in Northern California*. Tucson: University of Arizona Press.
- Crowly, A. M., J. Rauch, S. Seagrove, and D. A. Smith 1998 "Quantitative Cross-National Studies of Economic Development: A Comparison of the Economics and Sociology Literatures." *Studies in Comparative International Development* 33(2):30-57.
- Delacroix, J., and C. Ragin 1981 "Structural Blockage: A Cross-National Study of Economic Dependence, State Efficacy and Underdevelopment." *American Journal of Sociology* 86:1311-1347.
- Dixon, W. J., and T. Boswell 1996a "Dependency, Disarticulation and Denominator Effects: Another Look at Foreign Capital Penetration." *American Journal of Sociology* 102(2):543-562.
- 1996b "Differential Productivity, Negative Externalities, and Foreign Capital Dependency: Reply to Firebaugh." *American Journal of Sociology* 102(2):576-584.
- Emmanuel, A. 1972 *Unequal Exchange: A Study of the Imperialism of Trade*. New York: Monthly Review Press.
- Escobar, A. 1995 *Encountering Development: The Making and Unmaking of the Third World*. Princeton, N.J.: Princeton University Press.
- Esping-Andersen, G., ed. 1996 *Welfare States in Transition: National Adaptations in Global Economies*. Thousand Oaks, Calif.: Sage.
- Evans, P. B. 1979 *Dependent Development: The Alliance of Multinational State and Local Capital in Brazil*. Princeton, N.J.: Princeton University Press.
- 1987 "Class, State, and Dependence in East Asia: Some Lessons for Latin Americanists." In F. Deyo, ed., *The Political Economy of the New Asian Industrialism*. Ithaca, N.Y.: Cornell University Press.

- 1995 *Embedded Autonomy*. Princeton, N.J.: Princeton University Press.
- 1997 “The Eclipse of the State? Reflections on Stateness in an Era of Globalization.” *World Politics* 50(1):62–87.
- , and John Stephens 1988 “Development and the World Economy.” In Neil Smelser, ed., *The Handbook of Sociology*. Beverly Hills, Calif.: Sage Publications.
- , and Michael Timberlake 1980 “Dependence, Inequality and the Growth of the Tertiary: A Comparative Analysis of Less Developed Countries.” *American Sociological Review* 45:531–53.
- Featherstone, M., ed. 1990 *Global Culture*. London: Sage.
- Ferguson, J. 1990 *The Anti-Politics Machine: ‘Development’, Depoliticization and Bureaucratic Power in Lesotho*. Cambridge, U.K.: Cambridge University Press.
- Firebaugh, G. 1992 “Growth Effects of Foreign and Domestic Investment.” *American Journal of Sociology* 98(1):105–130.
- 1996 “Does Foreign Capital Harm Poor Nations? New Estimates Based on Dixon and Boswell’s Measures of Capital Penetration.” *American Journal of Sociology* 102(2):563–575.
- Fligstein, N. 1990 *The Transformation of Corporate Control*. Cambridge, Mass.: Harvard University Press.
- , and I. Mara-Drita 1996 “How to Make a Market: Reflections on the Attempt to Create a Single Market in the European Union.” *American Journal of Sociology* 102(1):1–33.
- Frank, A. G. 1998 *Re-Orient: Global Economy in the Asian Age*. Berkeley: University of California Press.
- Frobel, F., J. Heinrich, and O. Kreye 1981 *The New International Division of Labor: Structural Unemployment in Industrialized Countries and Industrialization in Developing Countries*. Cambridge, U.K.: Cambridge University Press.
- Gereffi, G. 1994 “The International Economy.” In N. Smelser and R. Swedberg, eds., *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press/Russell Sage Foundation.
- , and M. Korzeniewicz 1994 *Commodity Chains and Global Capitalism*. Westport, Conn.: Greenwood Press.
- Giddens, A. 1985 *The Nation-State and Violence*. Berkeley: University of California Press.
- 1991 *The Consequences of Modernity*. Cambridge, U.K.: Cambridge University Press.
- Gourevitch, P. 1986 *The Politics of Hard Times*. Ithaca, N.Y.: Cornell University Press.
- Guillen, M. F. 1994 *Models of Management: Work, Authority, and Organization in a Comparative Perspective*. Chicago: University of Chicago Press.
- Habermas, J. 1998 *The Inclusion of the Other*. Cambridge, Mass.: MIT Press.
- Hall, S. 1991 “The Local and the Global: Globalization and Ethnicity” In A. D. King, ed., *Culture, Globalization and the World-System*. Binghamton: Department of Art and Art History, State University of New York-Binghamton.
- Harrison, B. 1994 *Lean and Mean*. New York: Basic Books.
- Hirst, P., and G. Thompson 1996 *Globalization in Question*. Cambridge, Mass.: Blackwell.
- Hodagneu-Sotelo, P. 1994 *Gendered Transitions*. Berkeley: University of California Press.
- Hodson, R., ed. 1997 *The Globalization of Work*. Greenwich, Conn.: JAI Press.
- Inkeles, A., and D. Smith 1974 *Becoming Modern: Individual Change in Six Developing Countries*. Cambridge, Mass.: Harvard University Press.
- Kentor, J. 1998 “The Long-Term Effects of Foreign Investment Dependence on Economic Growth, 1940–1990.” *American Journal of Sociology* 103(4):1024–1046.
- Korzeniewicz, R. P., and T. P. Moran 1997 “World-Economic Trends in the Distribution of Income, 1965–1992.” *American Journal of Sociology* 102(4):1000–1039.
- Lash, S., and J. Urry 1994 *Economies of Signs and Space*. London: Sage.
- Locke, R. 1995 *Remaking the Italian Economy*. Ithaca, N.Y.: Cornell University Press.
- Mann, M. 1988 *States, War and Capitalism*. Oxford: Blackwell.
- Massey, D. S., and E. A. Parrado 1998 “International Migration and Business Formation in Mexico.” *Social Science Quarterly* 79(1):1–20.
- Melucci, A. 1996 *Challenging Codes: Collective Action in the Information Age*. Cambridge, U.K.: Cambridge University Press.
- Meyer, J. W. 1980 “The World Polity and the Authority of the Welfare State.” In A. Bergesen, ed., *Studies of the Modern World-System*. New York: Academic Press.
- J. Boli, G. M. Thomas, and F. O. Ramirez 1997 “World Society and the Nation-State.” *American Journal of Sociology* 103(1):144–181.
- , D. Frank, A. Hironaka, E. Schofer, N. B. Tuma 1997 “The Structuring of a World Environmental Regime, 1870–1990.” *International Organization* 51(4):623–651.

- , F. O. Ramirez, and Y. N. Soysal 1992 “World Expansion of Mass Education, 1870–1980.” *Sociology of Education* 65(2):128–149.
- O’Hearn, D. 1990 “TNCs, Intervening Mechanisms and Economic Growth in Ireland: A Longitudinal Test and Extension of the Bornschier Model.” *World Development* 18(3):417–429.
- Orrù, M., N. Biggart, and G. Hamilton 1997 *The Economic Organization of East Asian Capitalism*. Thousand Oaks, Calif.: Sage.
- Piore, M., and C. Sabel 1984 *The Second Industrial Divide*. New York: Basic Books.
- Portes, A. 1996 “Global Villagers: The Rise of Transnational Communities” *The American Prospect* (March–April):74–77.
- , and J. Walton 1981 *Labor, Class and the International System*. New York: Academic Press.
- Prebisch, R. 1950 “The Economic Development of Latin America and its Principal Problems.” *Economic Bulletin for Latin America* 7(February):1–22.
- Putnam, R. 1988 “Diplomacy and Domestic Politics: The Logic of Two-Level Games.” *International Organization* 42(Summer):427–460.
- 1993 *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, N.J.: Princeton University Press.
- Ramirez, F., and J. Boli 1982 “Global Patterns of Educational Institutionalization.” In P. Albach et al., eds., *Comparative Education*. New York: MacMillan.
- Regini, M. 1995 *Uncertain Boundaries: The Social and Political Construction of European Economies*. Cambridge, U.K.: Cambridge University Press.
- Reich, R. 1991 *The Work of Nations*. New York: Vintage Books.
- Ritzer, G. 1993 *The McDonaldisation of Society*. Thousand Oaks, Calif.: Pine Forge.
- Robertson, R. 1992 *Globalization*. London: Sage
- Robinson, W. I. 1998 “Beyond Nation-State Paradigms: Globalization, Sociology, and the Challenge of Transnational Studies.” *Sociological Forum* 13(4):561–594.
- Rodrik, D. 1997 *Has Globalization Gone Too Far?* Washington, D.C.: Institute for International Economics.
- Rubinson, R. 1976 “The World Economy and the Distribution of Income within States.” *American Sociological Review* 41:638–659.
- Sabel, C. 1996a “Constitutional Orders: Trust Building and Response to Change” In R. Boyer and J. R. Hollingsworth, eds., *Contemporary Capitalism*. Cambridge, U.K.: Cambridge University Press.
- 1996 *Ireland: Local Partnerships and Social Innovation*. Paris: Organization for Economic Cooperation and Development, Paris.
- Salzinger, L. 1997 “From High Heels to Swathed Bodies: Gendered Meanings Under Production in Mexico’s Export-Processing Industry.” *Feminist Studies* 23(3):549–574.
- Sassen, S. 1988 *The Mobility of Labour and Capital*. Cambridge, U.K.: Cambridge University Press.
- 1990 *The Global City*. Princeton, N.J.: Princeton University Press.
- 1996 *Losing Control? Sovereignty in an Age of Globalization*. New York: Columbia University Press.
- 1998 *Globalization and its Discontents*. New York: The New Press.
- 1999 “Cracked Casings: Notes Towards an Analytics for Studying Transnational Processes.” Mimeo, Department of Sociology, University of Chicago.
- Saxenian, A. 1994 *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, Mass.: Harvard University Press.
- Scott, A. J. 1993 *Technopolis*. Berkeley: University of California Press.
- Sklair, L. 1991 *Sociology of the Global System*. Hemel Hempstead, U.K.: Harvester Wheatsheaf.
- Smith D. A., and D. R. White 1991 “Structure and Dynamics of the Global Economy: Network Analysis of International Trade, 1965–1980.” *Social Forces* 70(4):857–893.
- Smith, J., and I. Wallerstein 1992 *Creating and Transforming Households: The Constraints of the World-Economy*. Cambridge, U.K., and Paris: Cambridge University Press and Maison des Sciences de l’Homme.
- Snidal, D. 1985 “The Limits of Hegemonic Stability Theory.” *International Organization* 39(4):580–614.
- Snyder, D., and E. Kick 1979 “Structural Position in the World System and Economic Growth, 1955–1970: A Multiple Network Approach.” *American Journal of Sociology* 84(5):1096–1126.
- Storper, M. 1997 *The Regional World: Territorial Development in a Global Economy*. London: Guilford Press.
- Streeck, W., and P. Schmitter 1991 “From National Corporatism to Transnational Pluralism: Organized Interests in the Single European Market.” *Politics and Society* 19(2):133–164.
- Soysal, Y. 1994 *Limits of Citizenship*. Chicago: University of Chicago Press.
- Tilly, C. 1992 *Coercion, Capital and European States, AD 990–1990*. Cambridge, Mass.: Blackwell.

- Turkle, S. 1995 *Life on the Screen*. New York: Touchstone/Simon and Schuster.
- Van Rossem, R. 1996 "The World System Paradigm as General Theory of Development—A Cross-National Test." *American Sociological Review* 61(3):508–527.
- Visser, J., and A. Hemerijck 1996 *A Dutch Miracle*. Amsterdam: Amsterdam University Press.
- Wade, R. 1990 *Governing the Market*. Princeton, N.J.: Princeton University Press.
- 1996 "The Limits of Globalization." In S. Berger and R. Dore, eds., *National Diversity and Global Capitalism*. Ithaca, N.Y.: Cornell University Press.
- 1998 "The Coming Fight Over Capital Flows." *Foreign Policy* 113:41–54.
- Walker, R. A. 1999 "Putting Capital in Its Place: Globalization and the Prospects for Labor." *Geoforum*, forthcoming.
- Wallerstein, I. 1974 *The Modern World-System I: Capitalist Agriculture and the Origins of the European World Economy*. New York: Academic Press.
- 1989 *The Modern World-System, III: The Second Era of Great Expansion of the Capitalist World Economy, 1730's–1840's*. New York: Academic Press.
- 1990 "Culture as the Ideological Battleground of the Modern World-System" In M. Featherstone, ed., *Global Culture*. London: Sage.
- Waltz, K. 1979 *Theory of International Relations*. Reading, Mass.: Addison-Wesley.
- Warren, B. 1980 *Imperialism: Pioneer of Capitalism*. London: New Left Books.
- Waters, M. 1995 *Globalization*. London: Routledge.
- Wellman, B., et al. 1996 "Computer Networks as Social Networks." *Annual Review of Sociology* 22:213–238.
- Zeitlin, M. 1984 *The Civil Wars in Chile (or the Bourgeois Revolutions That Never Were)*. Princeton, N.J.: Princeton University Press.

SEÁN Ó RIAIN  
PETER B. EVANS

## GOVERNMENT REGULATION

Government regulation is part of two larger areas of study, one encompassing all *state* policy making and administration, whether regulatory or not, the other encompassing all *regulatory* and *deregulatory*

activity, whether by the state or by some other institution. Viewed either way, the subject remains an interdisciplinary growth industry, with contributions made by political scientists, economists, legal scholars, historians, and sociologists. Scholarly emphasis in the 1990s on economic globalization and its consequences has added to an already rich literature on government regulation, deregulation, and re-regulation. Now attention is focused on the supranational as well as the national level. Cross-national, comparative studies of government regulation complement a large literature focused on the United States.

### WHAT IS GOVERNMENT REGULATION?

There is no uniformly agreed-upon concept of regulation that separates it from other kinds of government activity. Mitnick (1980, pp. 3–19) offers a good overview of concepts of regulation. On the one hand, narrow definitions typically focus on government action affecting private business by policing market entry and exit, rate or price, and profit structures and competitive environment. Some narrow definitions confine regulatory activity to that undertaken by administrative agencies (see also Majone 1994). If courts are the exclusive site for state rule making and enforcement, it is not considered government regulation. On the other hand, the broadest definitions conceive of regulation as government action affecting private businesses or citizens. Government regulation then becomes virtually coterminous with all government policy making and administration, whether by legislatures, administrative agencies, courts, or some combination.

Mitnick (1980) shows that American scholarship has provided for much variation in the conceptualizing of government regulatory activity. However, Majone (1994) suggests that in the past, American concepts typically were narrower than those adopted explicitly or implicitly by European scholars. For example, self-labeled *regulation theory* is a "quasi-Marxist theory [in which] the notion of regulation . . . refers to institutions and norms that permit the reproduction of conflictual or contradictory social relations" (see Steinmetz 1996, p. 346). In this predominantly European tradition, modes of regulation are broad political-economic and cultural governance forms. They revolve

around capital accumulation and involve state action, including macroeconomic, social, and labor-market policies, but also involve systems of interest intermediation in the workplace, economic rule making by banks and other nongovernmental institutions, and cultural schemata followed and taught in families and schools. Lange and Regini (1989) and Regini (1995) reject such an all-encompassing definition of regulation in favor of a somewhat narrower one. But they also call attention to how regulatory action structures and reconciles conflicts and allocates resources, as well as coordinates interaction and relationships in production and distribution.

Lange and Regini (1989) argue that regulatory principles and regulatory institutions must be separated analytically. Both economic and political institutions regulate or engage in governance. We tend to equate market exchange principles with economic institutions and legal rulings and administrative decrees with the state. However, as Lange and Regini (1989) demonstrate, economic institutions also employ command and control logic, while the state may employ the logic of exchange. Most recently, European scholars have moved away from equating regulation with the realm of *all* institutional governance or of *all* government legislation and social control. Many now distinguish “the regulation issue” both from other modes of institutional governance and from other modes of state action, including nationalization and government planning (Majone 1994, p. 77).

Sabatier (1975) has offered a useful definition of *government* regulation in between the broad and narrow extremes. His definition is based on the goals and content of government policy, not on the means of enforcement. It highlights the distinction between government policing of behavior and government allocation of goods and services. Distributive (e.g., defense contracts) and redistributive policies (e.g., the income tax, social welfare policies) allocate goods and services. Government policing is self-regulatory if it polices behavior to the benefit of the group whose behavior is policed. It is regulatory if it “seek(s) to change the behavior of some actors in order to benefit others” (Sabatier 1975, p. 307). Pollution control, antidiscrimination, consumer protection, occupational safety and health, employment relations, and antitrust are examples of regulatory policies.

Sociologists often distinguish between economic and social regulation. Where economic regulation controls market activities, such as entry and exit or price controls, social regulation controls aspects of production, such as occupational safety and health standards and pollution control (e.g., Szasz 1986). The term *social regulation* is also used to signal regulation that directly affects people rather, or more than, markets (Mitnick 1980, p. 15). In the 1990s literature on European economic integration, a distinction has been made between regulation (governance oriented to making markets) and reregulation (governance oriented to constraining markets) (e.g., Streeck 1998). But the term *reregulation* is also used more broadly, to signal regulatory reform that both liberalizes markets and institutes new rules to police them (Vogel 1996).

Regulation is dynamic. It is “an ongoing process or relation” between regulator and regulated parties (Mitnick 1980, p. 6). Because of the nature of the legal system in the United States, regulation U.S.-style tends to involve issuing and applying legal rules (Sabatier 1975, p. 307). For example, Congress has legislated federal statutes to promote competitive markets, to prevent race and gender discrimination in employment, and to increase workplace safety. These laws have been interpreted and enforced by the appropriate federal administrative agencies and by the federal courts. Federal regulatory agencies include the Interstate Commerce Commission (ICC), Federal Trade Commission (FTC), Federal Communications Commission (FCC), Securities and Exchange Commission (SEC), Equal Employment Opportunity Commission (EEOC), National Labor Relations Board (NLRB), Environmental Protection Agency (EPA), Food and Drug Administration (FDA), and Occupational Safety and Health Administration (OSHA). (For case studies of many of these agencies, see Derthick and Quirk 1985; Wilson 1980b.)

Consistent with the U.S. emphasis on legal rules as implementing mechanisms, the institutional forms used to reach regulatory goals are varied. Breyer (1982) provides an overview of the ideal-typical workings of various government regulatory forms, including cost-of-service rate making (e.g., public utility regulation), standard setting (e.g., administrative rule making and enforcement

by the EPA and OSHA), and individualized screening (e.g., the FDA regulations pursuant to which food additives can be marketed). Mitnick (1980) also provides an overview of government regulatory forms and contrasts regulation by directive (e.g., administrative and adjudicative rule making) with regulation by incentive (e.g., tax incentives, effluent charges, and subsidies).

It is no accident that European scholars in the 1990s are devoting heightened attention to government regulation and are also beginning to conceive of it more similarly to their U.S. counterparts (see, e.g., Majone 1994; Scharpf 1997a; Vogel 1996). European economic integration has been accompanied by concern that national governments would compete to lessen business costs in part by lowering standards for environmental, health and safety, financial, and other regulations. European integration has also involved a cumulative process of European Court of Justice rule making geared to constructing and policing the integrated market (see e.g., Leibfried and Pierson 1995; Scharpf 1997a). Thus, European integration has created a situation in which courts and judicial review are more important than they were in the past.

According to Majone (1994, p. 77), “regulation has become the new border between the state and the economy [in Europe] and the battleground for ideas on how the economy should be run.” In addition, since national courts in ordinary administrative and civil proceedings apply the market-making and market-policing rules formulated by the supranational European Court, regulatory *law* is beginning to become visible to ordinary citizens of European countries as it has been for some time to citizens of the United States. These developments do *not* mean that we can assume a future convergence of either the concept or the reality of the “regulatory state” in Europe and the United States. They *do* mean that there is increasing potential for the cross-fertilization of scholarly concepts, theories, and empirical work from both sides of the Atlantic. These developments also provide new opportunities for informative comparative studies of government regulation.

### THEORIES OF REGULATION

There are various general theoretical approaches to government regulation. Most are concerned

with regulatory origins or processes, but often they also address questions of impact, at least implicitly. Because regulation is not just an object of scholarly inquiry but also an ongoing political process, it is easy to confuse normative perspectives on regulation with explanations for the empirical phenomenon. Here, I focus on the latter, that is, on positive as opposed to normative theories. Mitnick (1980) and Moe (1987) provide detailed exposition and evaluation of a large range of these positive theories. All are theories of “interest.” Majone (1994) reviews the predominant normative perspective. The latter holds that corrective government action to improve economic efficiency is justified by such diverse types of market failure as natural monopoly, imperfect information and negative externalities (see also Breyer 1982).

Bernstein’s classic life-cycle theory argues that regulatory agencies designed in the public interest become captured by the powerful private interests they are designed to regulate (see Mitnick 1980, pp. 45–50). The diffuse majority favoring government regulation loses interest once the initial statute is legislated. This leaves the regulatory agency with few political resources to confront strong, well-organized regulatory parties with a large stake in agency outcomes.

Arguing that regulatory agencies are not simply captured by private interests but are designed from the beginning to do their bidding, Stigler (1971) and others have developed the economic theory of regulation. This theory assumes that all actors behave rationally in their own self-interest and so try to use government to achieve their own ends. But economic interest does not necessarily result in effective mobilization of resources. Because “there is a mobilization bias in favor of small groups, particularly those having one or more members with sizable individual stakes in political outcomes,” concentrated business interests have great advantages over diffuse groups in mobilizing for regulatory legislation (Moe 1987, pp. 274–275). Regulatory capture results when the costs of regulation fall upon a concentrated group (e.g., a particular industry such as railroads or airlines) and the benefits of regulation fall upon a diffuse group (e.g., consumers). Similarly, when benefits fall upon a concentrated group and costs on a diffuse one, regulation will be designed to benefit regulated parties.



The economic theory of regulation does not always predict capture. Generally, regulatory policies result from a chain of control running from economic groups to politicians to bureaucrats. These policies reflect the underlying balance of power among economic groups, whatever that balance may be. Considering different distributions of regulatory costs relative to regulatory benefits, Wilson (1980a, pp. 364–374) sketches four different scenarios for the origins of regulation. Exemplified by the origin and operation of the Civil Aeronautics Board, “client politics” is consistent with Stigler’s prediction that regulation reflects the regulated industry’s desires. Client politics result when costs are widely distributed and benefits are concentrated. When both costs and benefits are narrowly concentrated, both sides have strong incentives to organize and exert influence, so “interest group politics” results. Wilson views passage of the Commerce Act in 1886 as a product of conflict over rate regulation, in which interest group participants included railroads, farmers, and shippers.

According to Wilson, when both costs and benefits are widely distributed, interest groups have little incentive to form around regulatory issues because none can expect to capture most of the benefits or to avoid most of the costs. “Majoritarian politics,” in which the mobilization of popular opinion is likely to play an important role, governs passage of such legislation. Finally, “entrepreneurial politics” characterizes the dynamics of mobilization around policies that offer widely distributed benefits but narrowly concentrated costs. Here, although policy opponents benefit from the mobilization bias of small numbers and have strong incentives to organize, a “policy entrepreneur” can “mobilize latent public sentiment . . . [and] put opponents” on the defensive (Wilson 1980a, p. 370). For Wilson, pollution-control laws enforced by the EPA exemplify entrepreneurial politics. Although the traditional economic theory of regulation predicts ultimate capture of agencies created by entrepreneurial politics, Sabatier (1975) argues that such agencies can avoid capture by concentrated business interests if they actively develop a supportive constituency able to monitor regulatory policy effectively.

Economic theories of government regulation have much to say about the political dynamics of social groups seeking and resisting regulation, but

they do not attend to political and administrative institutions. In contrast, the positive theory of institutions “traces the congressional and bureaucratic linkages by which interests are translated into public policy” (Moe 1987, p. 279). This theory is one of a large group of more specific theories falling under the burgeoning “new institutionalism” in the social sciences (Eisner 1991; Powell and DiMaggio 1991). Like all variants of institutionalism, the positive theory of institutions argues that political institutions and rules of the game matter. Although actors try to create rules that lead to outcomes they favor, institutionalized rules may well be out of sync with underlying economic interests. Whether the regulatory policies of the U.S. Congress reflect any given economic interest depends on the distribution of that interest across congressional districts, the location of members of Congress who support that interest on particular committees with particular prerogatives and jurisdictions, and the rules of the congressional game.

The positive theory of institutions ordinarily begins with and focuses on the self-interest of actors in Congress and the regulatory agencies rather than that of actors outside these legislative and administrative institutions. It argues that legislative choice of regulatory forms as well as of regulatory content can be modeled as a function of the costs and benefits to legislators of selecting particular regulatory strategies (see, e.g., Fiorina 1982). These costs and benefits are a function of the distribution of economic interests across districts and the political-institutional rules of the game. In general, electoral incentives prevent members of Congress from placing high priority on controlling administrative agencies. The principal-agent models of control employed by the positive theory of institutions “suggest . . . that even when legislators do have incentives to control agencies toward specific ends” they probably will fail “owing to . . . conflicts of interest, information asymmetries, and opportunities for bureaucratic ‘shirking’” (Moe 1987, p. 281). Game-theoretic models of regulatory enforcement developed in this theory indicate ample opportunity for the capture of the regulators by regulated parties (Ayres and Braithwaite 1989). However, where some forms of capture are economically undesirable, others are economically (Pareto) efficient.

Other theoretical perspectives used by sociologists to study regulation include various forms

of neo-marxist political economy or class theory (see Levine 1988; Steinmetz 1997; Yeager 1990) and the political-institutionalist view developed by Theda Skocpol and others (Skocpol 1992; Weir et al. 1988). Where the former parallels the economic theory of regulation in focusing on the organization and mobilization of nongovernmental actors—specifically classes and segments of classes—in support of their interests, the latter parallels the positive theory of institutions in stressing the import of political structures and rules of the game. But in contrast to economic and positive theories, which largely model comparative statics (Moe 1987), class and political-institutional theories ordinarily focus on historical dynamics.

Political institutionalists stress, for example, the importance of feedback from prior to current regulatory policies and of political learning by government actors (see Pedriana and Stryker 1997). Feedback and political learning can help account for deregulation as well as for regulation (see Majone 1994). Class theorists stress how regulatory enforcement and cycles of regulation and deregulation evolve over time in response both to the structural constraints of a capitalist economy and to active struggles over regulation by classes and class segments. For example, Yeager (1990) argues that because government in a capitalist society depends on tax revenues from the private accumulation of capital, it tends to resolve conflict conservatively over such negative consequences of production as air or water pollution, so as not to threaten economic growth. Many aspects of U.S. regulatory processes make it likely that laws passed against powerful economic actors will be limited in impact or will have unintended effects that exacerbate the problems that initially caused regulation.

The effectiveness of regulatory statutes may be limited by implementation decisions relying on cost-benefit considerations because ordinarily costs are more easily determined than benefits and because cost-benefit analyses assert the primacy of private production. Moreover, government relies upon signals from private business to gauge when regulation is preventing adequate economic growth. Limited effectiveness of regulation also results from enforcement procedures tilted in favor of regulated parties that have the technical and financial resources needed to negotiate with agency officials. Corporate organizational forms encourage leniency and negotiations about compliance.

Corporate officials seldom are prosecuted for criminal violations because the corporate form makes it hard to locate individual culpability. Because courts emphasize proper legal reasoning when reviewing agency decisions, regulatory agencies may focus on procedure rather than substance. Ambiguous statutes are likely to heighten a procedural approach to regulatory enforcement (see Edelman 1992). In turn, focus on procedures over substance will tilt enforcement toward the interests of regulated parties. Finally, because no unit of government has complete control over any given policy from legislation through funding and implementation, parties bearing the cost of regulation need thwart regulation at only one point in the process, while supporters of regulation must promote it effectively at all points. In implementation, advocates of tough enforcement are likely to lose to more resource-rich segments of business seeking to limit regulation (Yeager 1990).

Notwithstanding forces that load regulatory processes in favor of the regulated business community and particularly the larger, more powerful corporations at the expense of smaller firms, consumers, environmentalists, and labor, class theorists also see limits on regulatory leniency. For example, Yeager (1990) argues that pollution-control enforcement biased toward large corporations dominating the U.S. economy will reproduce both the dominance of this business segment and of large-scale pollution. Regulatory ineffectiveness may lead to a loss of legitimacy for government as the public responds to higher risk and to perceived governmental failure by pressuring for additional pollution-control efforts.

Finally, although the concept of interest is central to theories of regulation, sociologists studying regulation are sensitive to the causal role of cultural schemata, norms, ideas, values, and beliefs as well as of economic and political interests and political institutions. For example, elaborating on Swidler's (1986) notion of culture as a tool kit, Pedriana and Stryker (1997) examine the diverse cultural strategies involved in the symbolic framing of regulatory enforcement efforts in U.S. equal employment law. They show that these frames are cultural resources developed by social and institutional actors in variable ways as a function of their variable political-economic, political-cultural, and legal circumstances. They also show how

actors' mobilization of cultural resources affects the subsequent path of regulatory policy making. Construction of cultural resources, then, is one key mechanism through which policy feedbacks occur and political learning is given effect. Meidinger (1987), too, highlights the role of culture, focusing on the way understandings—including understandings about costs, benefits, and appropriate trade-offs—are negotiated and enacted by actors in regulatory arenas. Because statutes are indeterminate, regulators always possess some discretion. In addition, the mutual interdependence among regulated parties and regulators calls attention to the formation of regulatory communities in which shared cognitive and normative orientations develop, forming the basis for ongoing regulatory cultures.

Ayres and Braithwaite (1989) harness the notion of regulatory culture to their search for economically efficient regulatory schemes. They approach the problem of regulatory capture through a synthesis of economic interest and socialization mechanisms. Seeking a social framework to facilitate economically efficient forms of capture while deterring inefficient capture, they point to benefits obtainable if all participants in regulatory processes that empower public interest groups adhere to a culture of regulatory reasonableness. For example, social and self-disapproval sanctions in a regulatory ethic that is firm but reasonable will inhibit regulators from capitulating to law evasion by industry and from punitive enforcement when industry is complying with regulatory law. Yeager (1990) has a somewhat different view of regulatory reasonableness. He views limits on regulatory laws controlling pollution as a function of prevailing cultural belief systems as well as of class and group relations. Notions of regulatory responsiveness and reasonableness are negotiated in enforcement interactions between regulators and regulated parties within an overall cultural framework attributing moral ambivalence rather than unqualified harm to regulated conduct. This facilitates adoption of a technical orientation to solving "noncompliance" problems rather than of a more punitive approach. Because the regulation of business has to be justified constantly within highly market-oriented cultures like the United States, administering market-constraining regulation itself becomes morally ambivalent and contributes to less aggressive enforcement.

### FROM REGULATION TO DEREGULATION AND REREGULATION

Discussions of dynamism and change, whether through structural contradiction and class conflict as stressed by neo-Marxist perspectives, or through policy feedback and political learning as stressed by political-institutionalists, lead naturally toward explicit theorization and empirical research on periods or cycles of regulation and deregulation or reregulation. As Majone (1994) points out, deregulatory ideologies and politics in the United States were preceded by decades of scholarship on the economics, politics, and law of government regulatory processes. In Europe, by contrast, the term "deregulation" gained much more "sudden currency" (Majone 1994, p. 98). Unsurprisingly, on both sides of the Atlantic, the concepts and perspectives used to study deregulation parallel the alternative economic interest and political interest/political-institutional foci of theories of regulation themselves.

On the one hand, for example, Szasz (1986) analyzes deregulatory social movements in the United States through the lens of presumed accumulation and legitimation functions of the capitalist state. He suggests that changing economic circumstances provided political opportunity for the deregulatory movement in occupational safety and health. In the 1970s, a substantially worsening economy altered the balance of class forces and changed the political situation confronting the state. Deteriorating economic conditions weakened the economic and political power of organized labor, a major supporter of occupational safety and health legislation. These same conditions encouraged big business to join the already existing but to this point unsuccessful small business attacks on the Occupational Safety and Health Administration. Where small business argued for the complete elimination of OSHA, big business relied on cost-benefit analyses to argue that sound economics required reforming the implementation process. Though economic conditions made deregulation possible, the success it achieved and the form it took required business interests to mount a conscious, ideological campaign to mold favorable public opinion.

On the other hand, Derthick and Quirk (1985), examining deregulatory processes in the realm of *economic*, as opposed to *social* regulation, criticize

*non*-state-centered analyses of deregulation. They argue that, at least in the United States, regulated industries with a putative stake in deregulation did not ask to be deregulated. Nor were consumers' movements a major force. Instead, the deregulatory push emanated predominantly from within state regulatory agencies and courts, with commissioners and judges acting as policy entrepreneurs. Deregulatory politics and deregulation itself were only later and often quite reluctantly accepted by regulated industries such as airlines, trucking, and communications.

The foci of Derthick and Quirk (1985) and Szasz (1986) converge to highlight the role played by academic and policy think-tank experts in paving the way for and promoting pro-competitive regulatory reform. Szasz shows how U.S. economists and political scientists built a rationale for deregulation in the 1970s (see Breyer 1982 for a sophisticated but very readable overview of economic justifications and analyses of regulation and of economic justifications for deregulation). Derthick and Quirk (1985) push the role played by these experts further back in time, albeit noting that the earliest promoters of regulatory reform would never have anticipated the successful political movement for which they helped paved the way. Nonetheless, U.S. administrative law and public administration experts long had found fault with government regulatory structures and procedures. By the 1960s, economists had joined the chorus, attacking economic regulation for fostering costly inefficiencies and for shielding industries from competition. Economists also attacked economic and social regulation for producing costs in excess of benefits.

As Majone (1994) points out, where the United States tended to create regulated industries, allowing critics to catalogue subsequent *regulatory* failures, Europe traditionally tended toward public ownership, with its own set of corresponding failures to interpret and experience. When deregulatory ideologies were produced in Europe or diffused from the United States, privatization became the rallying point. But neither privatization nor the search for "less restrictive" or "less rigid" government intervention necessarily means the retreat of the state (Majone 1994, p. 80).

Deregulation is most precisely conceptualized as *reduction in the level of government regulation*. It

involves eliminating or reducing government rules or lessening their strictness (Vogel 1996). Strictly speaking, deregulation moves institutional governance toward *self*-governed markets. But according to Vogel (1996), much scholarship is remiss in equating deregulation with *any* kind of liberalization or pro-competitive regulatory reform (see, e.g., Derthick and Quirk 1985). Equating deregulation with market liberalization is undesirable because it forecloses by definitional fiat the question of whether and how liberalization may involve *more* government rule making rather than less. Liberalization may involve *changing* government rules rather than eliminating them (Vogel 1996).

Indeed, Vogel (1996) argues that across capitalist democracies the trends are toward what he terms reregulation rather than deregulation. Rather than reduce their levels of regulation of the private sector, governments have reorganized their control over it. When governments privatize previously nationalized industries and when they liberalize regulated markets to introduce more competition, ordinarily this involves both the reformulation of old rules and the creation of new ones. As the title of Vogel's book suggests, then, the price of "freer markets" is "more rules" (Vogel 1996; see also, e.g., Majone 1994; Streeck 1998). Increased conceptual precision helps Vogel solve what otherwise appears as a puzzle and paradox: that, as noted by Derthick and Quirk (1985), state actors themselves promote a great deal of deregulatory activity. To the question of why governments would take action apparently against their own interests, Vogel answers "they don't." Instead, as political-institutional perspectives on regulation would suggest, governments initiate regulatory *reform* and shape *reregulation* in their own interests. Pro-competitive regulatory reform represents neither "the triumph of markets over government" nor "the triumph of [economic] interests over government" (Vogel 1996, pp. 10, 13).

Vogel rejects exclusively economic theories of deregulation that argue either that increasingly integrated global markets force governments to deregulate or that interest groups, especially regulated industries, orchestrate reform. Instead, he provides a synthesis of sorts between economic and political-institutional views. He hypothesizes that, on the one hand, governments of advanced capitalist democracies *do* face a common set of economic and cultural pressures. Technologically

induced global market changes in particularly dynamic sectors like telecommunications and financial services compel governments to respond in *some* way, but without setting the terms of the response. Diffusion of market and deregulatory ideologies from the United States also exerts pressures—albeit somewhat less strong—for a response. In addition, this ideological diffusion helps explain why governments across the advanced capitalist world adopt similar reform *rhetorics*. It does *not* explain when and why they undertake reform action or the form their reregulation takes. Finally, governments do face a common politics of economic slowdown, in which they find that “the growth in demand for government services outpaces the growth of government resources for meeting this demand (Vogel 1996, p. 40). This creates political opportunity. But it does not explain why conservative and even left political parties take that opportunity in some countries, while neither left nor even conservative parties do so in others.

In short, according to Vogel’s theory of deregulation, there are a set of common forces for change—some stronger, some weaker, some broader, some narrower—that set the stage for specific national responses. However, “states themselves, even more than private interest groups, have driven the reform process” (Vogel 1996, p. 4). Governments in the advanced industrial world cannot ignore private groups’ interests and demands, but they take the initiative in shaping reform and constructing politically acceptable compromises. In this, governments do *not* converge in a common deregulatory trend. Instead, they adopt particular types and distinctive styles of reregulation as they achieve liberalized markets to different degrees. At its core, so-called deregulation is about “finding new ways to raise government revenue and designing new mechanisms of policy implementation” (Vogel 1996, p. 19). These goals typically concern states more than private interests, so it becomes no surprise that state actors actively mobilize to shape regulatory reform.

Vogel categorizes diverse reregulatory styles and processes in terms of two dimensions: whether the emphasis is more on liberalization or more on reregulation, and whether the reregulation undermines or enhances government control over industry. In turn, the diverse reregulatory styles and processes emerge as a function of variation

across countries in political-institutional regulatory regimes, developed over time as a function of each country’s own unique history, especially its history of industrialization. Regulatory regimes are “comprised of specific constellations of ideas and institutions” (Vogel 1996, p. 20). The ideas, or *regime orientation*, involve “state actors’ beliefs about the proper scope, goals and methods of government intervention in the economy and about how this intervention affects economic performance” (Vogel 1996, p. 20). Diverse regime orientations cause government officials to define the public interest in varied ways, to interpret common economic and ideological pressures and trends differently, and to conceive of different kinds of responses to such pressures and trends as appropriate. *Regime organization* involves how state regulators concerned with a given industry are structured internally and how they are linked to the private sector. In contemplating reform, government actors will assess how diverse alternatives are likely to affect existing institutions and arrangements. Differences in regime organization affect especially *who*—whether political parties, bureaucrats, and so forth—will control reform processes, *whether* government officials will try to refrain state capabilities, and *what* capabilities government officials will try to retain or develop for themselves in the reform process.

Vogel’s (1996) framework fundamentally reorients scholars to distinguish concepts of deregulation and reregulation and to approach both in terms of an overarching perspective that considers regulation, deregulation, and reregulation as part of the broader study of regulatory change. (Parallel efforts to integrate explanations of welfare development and retrenchment into a broader theory of change in social policy are equally underway [see, e.g., discussions in Steinmetz 1997; Stryker 1998]). Vogel’s framework is conducive to investigating the interaction of international pressures and domestic politics, as well as the interaction of governments and private actors. It is likewise conducive to investigating how institutional and cultural boundaries between public and private have been variably articulated across countries and over time, and to investigating how globalization shapes opportunities for and constraints on national-level government regulation and on the development of supranational regulatory institutions.

### EMPIRICAL STUDIES

Empirical research on regulation includes studies of regulatory origins (e.g., Majone 1994; Sanders 1981, 1986; Steinberg 1982), processes (e.g., Edelman 1992; Eisner 1991; Moe 1987; Yeager 1990), and impact (e.g., Beller 1982; Donahue and Heckman 1991; Mendelhoff 1979). It also includes studies of deregulation and reregulation (e.g., Derthick and Quirk 1985; Streeck 1998; Szasz 1986; Vogel 1996). Studies of processes look at the evolution of regulatory forms (e.g., Majone 1994; Stryker 1989, 1990) as well as at the substance of regulatory rules (e.g., McCammon 1990; Melnick 1983; Vogel 1996). Researchers employ a variety of methodologies. These include quantitative assessment of causes and consequences of regulation (e.g., Donahue and Heckman 1991; Mendeloff 1979; Steinberg 1982) and quantitative models of regulatory processes (e.g., Edelman 1992; Edelman et al. 1999; Yeager 1990). They also include qualitative, case-oriented legal, historical, or comparative accounts of regulatory, deregulatory, and reregulatory evolution (e.g., Majone 1994; Melnick 1983; Sanders 1981; Stryker 1990; Szasz 1986; Vogel 1996). It is hard to generalize about findings from empirical studies of regulation. A few things, however, are reasonably clear.

First, no general theory or perspective on regulation enjoys unqualified support when stacked up against the variety and complexity of regulatory experiences. Second, all extant theories have something to offer the empirical analyst. Third, in response to the first and second points, the field seems to be moving away from accounts that focus on either economic interests or political-institutional rules to more integrative or synthetic accounts that encompass a role for both. Fourth, European Union integration has increased interest in empirical research on supranational regulatory bodies, as a key part of the broader study of multitiered governance structures. Fifth, empirical building blocks are being constructed for overarching concepts and theories that account for variation in regulatory regimes and for regulatory change, whether toward increased or decreased regulation or from one institutional principle (e.g., command and control) to another (e.g., market incentives). The rest of this article elaborates on these points.

Empirical studies suggest that economic interests and resources are a major factor but not the sole one, in the dynamics of political struggles over regulatory origins and administration (Moe, 1987; Sanders 1986; Stryker 1989, 1990; Szasz 1986; Yeager 1990). Political structures and rules of the game matter because they are the mechanisms through which economic and social actors must translate their interests into regulatory policy (Moe 1987). But for legislative, administrative, and judicial participants in policy processes, these institutional mechanisms also create independent interests in, and resources for, regulatory policy making. Sanders (1981) shows that the regulation of natural gas in the United States has been a function of four sets of regionally based economic interests, including gas producer regions of the United States and gas consumer regions, as well as of electoral rules and structures. Regulatory outcomes have resulted from a dynamic relationship among political actors who reflect the changing market positions of their constituents. "The potential for sectional conflict is exacerbated by the territorial basis of elections, the weakness of the party system, and a federal structure that not only encloses different political cultures and legal systems, but also supports fifty sets of elected officials sensitive to encroachments on their respective turfs" (Sanders 1981, p. 196).

Current regulatory structures and policies *do* have feedback effects constraining and providing opportunities for subsequent regulatory policies as well as for subsequent action by parties with interests at stake in regulation (Sanders 1981; Steinberg 1982; Stryker 1990). Feedbacks occur through cultural as well as political-institutional mechanisms and political learning (e.g., Pedriana and Stryker 1997; Vogel 1996). In this regard, Vogel's (1996) comparative study of deregulation and regulation of telecommunications and financial services in the United States, the United Kingdom, France, Germany, and Japan highlights the mediating role of nationally specific regime orientations. State actors interpret situations and conceive of responses through the lens of regime orientation. Their cognitive and normative interpretive work then shapes the form and content of regulatory reform. Pedriana and Stryker (1997) demonstrate that both general equal opportunity values and the specific language in which

they are expressed provide raw materials for construction of symbolic resources by actors struggling over the enforcement of equal employment and affirmative action law in the United States. The work of Majone (1994) and Boyer (1996), among others, suggests that political learning occurs through the experience and interpretation of regulatory failures as well as of market failures.

In addition, the legal structures and culture through which most regulation is administered in the United States significantly shape regulatory processes and outcomes. For example, Melnick (1983) shows how the narrow, highly structured, reactive, and adversarial legal processes through which pollution control takes place in the United States have led to court decisions that simultaneously extend the scope of EPA programs and lessen agency resources for achieving pollution control goals. Appellate judges tend to promote stringent antipollution standards because they are removed from local concerns and are likely to be inspired by broad public goals. In a different institutional location, trial judges observe the impact on local businesses and citizens of imposing strict regulation. Their flexibility in response to the perceived harm of strict regulation generates an equity-balancing enforcement that counteracts what is accomplished in standard setting.

Likewise, because legal mandates are not self-executing and many are ambiguous, the response of regulated parties is an important mediator of regulatory impact. This response includes actions taken by organizations to demonstrate their compliance with law. For example, Edelman (1992) and Edelman and colleagues (1999) show that organizations respond to federal equal employment law in the United States by creating equal employment opportunity policies, organizational units, and grievance procedures. These both promote symbolism over substance and shape later court constructions of what constitutes compliance and what will insulate organizations from liability. More obviously, the response of regulated parties also includes whether and how public and private institutions and individuals invoke regulatory law on behalf of aggrieved parties (e.g., Burstein 1991). Additionally, it includes how public and private actors mobilize the values and language encapsulated in the law as political-cultural and legal resources to *change* the law (e.g., Pedriana and Stryker 1997).

Yet another insight from empirical studies is that regulatory implementation is influenced by internal agency politics as well as by the agency's external environment. Likewise, technical experts play an important role in shaping regulatory evolution. Stryker (1989, 1990) has shown how, in conjunction with class and political institutional factors, intra-NLRB conflict between agency economists and lawyers over the proper administrative use of social science caused Congress to abolish the NLRB's economic research unit. Katzmann (1980) and Eisner (1991) have shown how internal jockeying by economists within the FTC changed enforcement priorities and outcomes over time. Even more generally, empirical studies of regulation and deregulation point to the justificatory and mobilizing import of diverse kinds of scientific and technical expertise (e.g., Derthick and Quirk 1985; Eisner 1991; Szasz 1986). While heavily relied upon to promote deregulation and pro-competitive regulatory reform, economic analysis also can be mobilized to promote more stringent regulation and diverse types of reregulation (e.g., Rose-Ackerman 1992; Stryker 1989).

Empirical studies of regulation also show that regulation often has unintended effects. Yeager (1990) shows how EPA sanctioning decisions and processes, while rational in the face of economic, political, and legal constraints on the agency, reproduce private sector inequality by favoring large corporations that have financial and technical resources. Large companies have greater access to agency proceedings than do small companies. Agency proceedings often change pollution-control requirements in favor of regulated firms, so that ultimately large corporations have fewer pollution violations. In decisions to apply the harshest sanctions—criminal and civil prosecutions—the EPA may well avoid tangling with the most resource-rich firms for fear of losing in court. Melnick (1983, p. 354) indicates a similar dynamic. Ostensibly neutral procedures, then, create inequitable law enforcement and may also help reproduce the problems that led to the initial pollution-control legislation.

Yet another important message emphasized by empirical studies of regulation in the 1990s is the need to consider the growth of supranational mechanisms of governance and how these interrelate with national government regulation. Majone

(1994), for example, shows that with minimal explicit legal mandate and with very limited resources, there has nonetheless been continuous growth in the final three decades of the twentieth century in regulation by the European Community (EC, now the European Union, or EU). Economic and social regulation is “the core” of EU policy making (Majone 1994, p. 77). It is undertaken by the European Commission (Commission) in tandem with the European Court of Justice (ECJ), now supplemented with a court of first instance (see Leibfried and Pierson 1995). Between 1967–1987, for example, even *before* the Single European Act recognized EC authority to legislate to protect the environment, there were close to “200 environmental directives, regulations and decisions made by the European Commission” (Majone 1994, p. 85). Consumer product safety, banking and financial services, and medical drug testing also have been areas of high-volume Commission regulatory activity. Regulation has provided a way for the Commission to expand its role in spite of tight EC budgets and the serious political-institutional constraints embedded in the EC’s legal framework, at the same time as EC member states have been willing to delegate to a supranational authority because agreements among the EC national governments had low credibility (Majone 1994).

In the regulatory arena, the ECJ has been as important as, or even more important than, the Commission (see, e.g., Leibfried and Pierson 1995). Ostner and Lewis (1995), for example, stress the *inter-relationship* of the Commission and the ECJ. Even before the Single European Act in 1987, “gender policies . . . evolved through the intricate interplay between these two supranational bodies, within the range of outcomes tolerated by member states. By the late 1980s the Court’s interpretations of article 199 [of the Treaty of Rome], Commission-fostered directives that [gave] the article concrete form and extend[ed] it, and the Court’s subsequent rulings about the meaning of the directives yielded a body of gender-related policies of substantial scope” (Ostner and Lewis 1995, p. 159). No wonder scholars have characterized the EU as a “state of courts and technocrats” (Leibfried 1992, p. 249) and have highlighted “the rise of the regulatory state in Europe” (Majone 1994, p. 77). In turn, European scholars’ awareness of the import of Commission and ECJ regulatory activity

has fueled their growing research interest in American-style regulation (Majone 1994; see also Leibfried and Pierson 1995).

Finally, although capture of government regulators by regulated parties can and does occur (see Sabatier 1975; Sanders 1981), it need not. Enactment of regulatory legislation can also lead to cycles of aggressive enforcement alternating with periods of capture or, similarly, to enforcement that oscillates between or among the interests at stake in regulation or between periods of regulation and deregulation or reregulation. For example, over time, FTC enforcement has alternated between favoring big or small business and core or peripheral economic regions of the United States (Stryker 1990). Sanders’s (1981) study of natural gas regulation in the United States shows that the initial federal legislation mixed goals of consumer protection and of industry promotion. Federal Power Commission interaction with its environment did not result in stable capture by gas producers but rather in oscillation between capture by gas consumers and capture by gas producers. Clearly, consumers, labor, and other subordinate groups can be, and have been, benefited by regulation (see, e.g., Sanders 1981; Steinberg 1982; Stryker 1989). But the political economy of capitalism also sets structural and cultural limits to these benefits (McCammon 1990; Szasz 1986; Yeager 1990).

In this regard, economic globalization and European economic integration enhance the political and economic resources of business groups at the expense of labor, providing pressures and opportunities for governments to undertake market-liberalizing regulatory reforms (Streeck 1995, 1998; Stryker 1998; Vogel 1996). However, these same processes also may generate counterpressures and counteropportunities. As Streeck (1998) shows, European integration has been a process of economic liberalization by international means. In this process, national-level regulations are exposed to competitive market pressures, including the threat of “regulatory arbitrage”—business corporations moving capital or firms from countries with less favorable regulations to countries with a more favorable regulatory climate. Competitive market pressures then further advance liberalization. But liberalization likewise “calls forth demands” from individuals and communities for market-constraining reregulation, so that they can “cope with the uncertainties of free markets and stabilize



their social existence in dynamically changing economic conditions” (Streeck 1998, p. 432).

A major challenge to theories and empirical research on government regulation in the future is to model and explain the historical and comparative dynamics of both economic and social regulation at intersecting subnational, national, and supranational levels. Further work should continue to address diversity and change over time and place in regulatory scope, levels, institutional forms, and cultural justifications. Ideally, further juxtaposition of abstract theory and concrete historical and comparative research, both qualitative and quantitative, can lead to integrated theories of regulatory origins, processes, and impact. Ideally, as well, these theories can explain not just regulation but also deregulation and reregulation. This is a tall order, but the seeds have been planted in scholarship like that of Vogel (1996), which is equally sensitive to economic and organizational interests and resources, to political structures and rules, and to regulatory cultures (see also the empirically informed analytic frameworks offered in, e.g., Scharpf 1997b; Stryker 1996). Seeds also have been planted in research programs, like Vogel’s (1996), that are sensitive to periods or cycles in which different economic and other institutional arrangements, incentives, and constraints operate, and to feedback effects from past to future regulatory policies and processes (see also Boyer 1996).

Whether or not such an integrative and synthetic theory is achieved, a combination of unfolding social processes, including globalization, court-led European integration, and democratization and marketization in eastern Europe and elsewhere all will continue to enhance interest in the study of government regulation. All these processes simultaneously promote economic liberalization and the regulatory state. Whatever else these current political-economic changes bring, they certainly should enhance scholarly dialogue and also synergy across national borders in the study of regulation.

## REFERENCES

Ayres, Ian, and John Braithwaite 1989 “Tripartism, Empowerment and Game-Theoretic Notions of Regulatory Capture” (American Bar Foundation Working Paper No. 8902). Chicago: American Bar Foundation.

- Beller, Andrea 1982 “Occupational Segregation by Sex: Determinants and Changes.” *Journal of Human Resources* 17:371–392.
- Breyer, Stephen 1982 *Regulation and Its Reform*. Cambridge, Mass.: Harvard University Press.
- Boyer, Robert 1996 “State and Market: A New Engagement for the Twenty-First Century?” In R. Boyer and D. Drache, eds., *States Against Markets: The Limits of Globalization*. London: Routledge.
- Burstein, Paul 1991 “Legal Mobilization as a Social Movement Tactic: The Struggle for Equal Employment Opportunity.” *American Journal of Sociology* 96:1201–1225.
- Derthick, Martha, and Paul Quirk 1985 *The Politics of Deregulation*. Washington D.C.: Brookings Institution.
- Donahue, John J., III, and James Heckman 1991 “Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks.” *Journal of Economic Literature* 29:1603–1643.
- Edelman, Lauren 1992 “Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law.” *American Journal of Sociology* 97:1531–1576.
- , Christopher Uggen, and Howard S. Erlanger 1999 “The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth.” *American Journal of Sociology* 105:406–454.
- Eisner, Mark Allen 1991 *Antitrust and the Triumph of Economics: Institutions, Expertise and Policy Change*. Chapel Hill: University of North Carolina Press.
- Fiorina, Morris P. 1982 “Legislative Choice of Regulatory Forms: Legal Process or Administrative Process?” *Public Choice* 30:33–66.
- Katzmann, Robert A. 1980 *Regulatory Bureaucracy: The Federal Trade Commission and Antitrust Policy*. Cambridge, Mass.: MIT Press.
- Lange, Peter, and Marino Regini 1989 “Interests and Institutions: Forms of Social Regulation and Public Policy Making.” In P. Lange and M. Regini, eds., *State, Market and Social Regulation: New Perspectives on Italy*. Cambridge, U.K.: Cambridge University Press.
- Leibfried, Stephan 1992 “Towards a European Welfare State? On Integrating Poverty Regimes into the European Community.” In Z. Ferge and J. E. Kolberg, eds., *Social Policy in a Changing Europe*. Boulder, Colo.: Westview Press.
- , and Paul Pierson, eds. 1995 *European Social Policy: Between Fragmentation and Integration*. Washington D.C.: Brookings.
- Levine, Rhonda 1988 *Class Structure and the New Deal: Industrial Labor, Industrial Capital and the State*. Lawrence: University of Kansas Press.

- Majone, Giandomenico 1994 "The Rise of the Regulatory State in Europe." *West European Politics* 17:77-101.
- McCammon, Holly J. 1990 "Legal Limits on Labor Militancy: U.S. Labor Law and the Right to Strike Since the New Deal." *Social Problems* 37:206-229.
- Meidinger, Errol 1987 "Regulatory Culture: A Theoretical Outline." *Law & Policy* 9:355-385.
- Melnick, R. Shep 1983 *Regulation and the Courts: The Case of the Clean Air Act*. Washington, D.C.: Brookings Institution.
- Mendeloff, John 1979 *Regulating Safety: An Economic and Political Analysis of Occupational Safety and Health Policy*. Cambridge, Mass.: MIT Press.
- Mitnick, Barry M. 1980 *The Political Philosophy of Regulation: Creating, Designing and Removing Regulatory Forms*. New York: Columbia University Press.
- Moe, Terry 1987 "Interests, Institutions and Positive Theory: The Politics of the NLRB." *Studies in American Political Development* 2:236-299.
- Ostner, Ilona, and Jane Lewis 1995 "Gender and the Evolution of European Social Policies." In S. Leibfried and P. Pierson, eds., *European Social Policy: Between Fragmentation and Integration*. Washington, D.C.: Brookings Institution.
- Pedriana, Nicholas, and Robin Stryker 1997 "Political Culture Wars 1960s Style: Equal Opportunity-Affirmative Action Law and the Philadelphia Plan." *American Journal of Sociology* 103:633-691.
- Powell, Walter W., and Paul J. DiMaggio, eds. 1991 *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Regini, Marino 1995 *Uncertain Boundaries: The Social and Political Construction of European Economies*. Cambridge, U.K.: Cambridge University Press.
- Rose-Ackerman, Susan 1992 *Rethinking the Progressive Agenda: The Reform of the American Regulatory State*. New York and Toronto: The Free Press.
- Sabatier, Paul 1975 "Social Movements and Regulatory Agencies: Toward a More Adequate—and Less Pessimistic—Theory of 'Clientele Capture'." *Policy Sciences* 6:301-342.
- Sanders, M. Elizabeth 1981 *The Regulation of Natural Gas: Policy and Politics, 1938-1978*. Philadelphia: Temple University Press.
- 1986 "Industrial Concentration, Sectional Competition, and Antitrust Politics in America, 1880-1980." *Studies in American Political Development* 1:142-214.
- Scharpf, Fritz 1997a "Economic Integration, Democracy and the Welfare State." *Journal of European Public Policy* 4 (March):18-36.
- 1997b *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*. Boulder, Colo.: Westview Press.
- Skocpol, Theda 1992 *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, Mass.: Belknap.
- Steinberg, Ronnie 1982 *Wages and Hours: Labor and Reform in Twentieth Century America*. New Brunswick, N.J.: Rutgers University Press.
- Steinmetz, George 1997 "Social Class and the Reemergence of the Radical Right in Contemporary Germany." In John R. Hall, ed., *Reworking Class*. Ithaca, N.Y.: Cornell University Press.
- Stigler, George 1971 "The Theory of Economic Regulation." *Bell Journal of Economic and Management Science* 2:3-21.
- Streeck, Wolfgang 1995 "From Market Making to State Building? Reflections on the Political Economy of European Social Policy." In S. Leibfried and P. Pierson, eds., *European Social Policy: Between Fragmentation and Integration*. Washington D.C.: Brookings Institution.
- 1998 "The Internationalization of Industrial Relations in Europe: Prospects and Problems." *Politics & Society* 26:429-459.
- Stryker, Robin 1989 "Limits on Technocratization of the Law: The Elimination of the National Labor Relations Board's Division of Economic Research." *American Sociological Review* 54:341-358.
- 1990 "A Tale of Two Agencies: Class, Political-Institutional and Organizational Factors Affecting State Reliance on Social Science." *Politics & Society* 18:101-141.
- 1996 "Beyond History vs. Theory: Strategic Narrative and Sociological Explanation." *Sociological Methods and Research* 24:304-352.
- 1998 "Globalization and the Welfare State." *International Journal of Sociology and Social Policy* 18:1-49.
- Swidler, Ann 1986 "Culture in Action." *American Sociological Review* 51:273-286.
- Szasz, Andrew 1986 "The Reversal of Federal Policy Toward Worker Safety and Health." *Science and Society* 50:25-51.
- Vogel, Steven K. 1996 *Freer Markets, More Rules: Regulatory Reform in Advanced Industrial Countries*. Ithaca, N.Y.: Cornell University Press.
- Weir, Margaret, Ann Shola Orloff, and Theda Skocpol 1988 "Understanding American Social Politics." In M. Weir, A. Orloff, and T. Skocpol, eds., *The Politics of Social Policy in the United States*. Princeton, N.J.: Princeton University Press.

Wilson, James Q. 1980a "The Politics of Regulation." In J. A. Wilson, ed., *The Politics of Regulation*. New York: Basic Books.

— ed. 1980b *The Politics of Regulation*. New York: Basic Books.

Yeager, Peter C. 1990 *The Limits of Law: The Public Regulation of Private Pollution*. Cambridge, U.K.: Cambridge University Press.

ROBIN STRYKER

## GROUP COHESIVENESS

See Interpersonal Attraction; Small Groups.

## GROUP CONFLICT RESOLUTION

To understand conflict resolution among groups it is helpful first to consider the role of conflict in and among groups. Conflict analysis of groups is divided between the study of *intragroup* conflict, that which happens within a group among its members, and *intergroup* conflict, that which happens between one or more groups where the conflict is viewed as involving the group as a whole. The study of these phenomena is closely related to the study of both group dynamics and cross-cultural relations. A further distinction is made according to the group level being studied. Here the separation is generally between group conflicts and international conflicts. Group conflicts include both communal group conflicts and workplace conflicts. Communal group conflicts can involve just about any group that provides people with social identity. Social identity is itself a group-level concept, since it is defined as the identity we gain from being part of a collectivity, a group. Since most of the hours of our days are organized around group membership, the study of groups and their interactions is central to the work of sociologists.

When groups are in conflict, the very presence of the group intensifies and changes the way conflict between individuals is perceived. Membership in the group affords the individual two perceptions that impact on the conflict: (1) that the individual is right and justified to engage in and

attempt to win the conflict and (2) that the individual will be evaluated and either further embraced or rejected based on his or her performance in a conflict situation. Within groups, conflict often results in patterns of splintering or perhaps a coup d'état and expulsion. Between groups, conflict can encourage deindividuation, ethnocentrism, and diabolical imaging of the enemy. Groups often provide individuals with a way to rationalize their involvement in a conflict and perhaps to take actions they might otherwise avoid. Conflict resolution is likewise changed in a group situation.

Defining exactly what a conflict is has also been an important part of the analysis of conflict and its resolution. There are generally two accepted ways of defining "conflict": (1) realistic and (2) perceived. The former involves tangible, verifiable competing interests. The latter refers to situations where it is believed by one or both parties that the other stands in the way of achieving what is desired. Conflicts occur over resources, power distribution, and values. They are classified as latent (yet to be noticed) or manifest. A final method of defining "conflict" is as destructive or constructive. Although viewed by most people as risky and something to be avoided, conflict is often seen differently by conflict analysts. Most who study conflict and its resolution argue that conflict holds the potential to be constructive, to create positive social change. These theorists and practitioners argue that how one copes with conflict is the important distinction. If the proper structures exist (such as skilled trainers or participants), conflict can actually have creative and constructive outcomes.

## CONFLICT RESOLUTION

The area of conflict resolution is itself an interdisciplinary field of research. Despite the great diversity of scholars involved in this work, they are brought together by a common interest in understanding how and when conflicts are resolved. What to call the work varies, including such terms as "conflict management," "conflict resolution," and "conflict transformation." A variable here is to what extent and in what way one might get involved to eliminate a conflict. Scholars in this area favor searching for nonviolent and noncontentious methods for ending conflicts, perhaps because these methods are felt to be more

permanent than the cycle of escalation likely to emerge with aggressive measures for “ending” conflicts. Avoidance and repression are also considered to be negative responses to conflict.

Here we focus not generally on the field of conflict resolution but specifically on it in relation to groups. While there is much application of theory and research across the levels of study, there is not much agreement on the transferability of ideas from the interpersonal level to the group or international level. Conflict resolution among groups remains an important but understudied social process. Great strides have been taken in developing various models of conflict resolution at the interpersonal level, but among groups the process remains less easily identifiable. At the international level, only small inroads have been made in developing and testing possible methods. War or military threats remain the most common responses to conflict among nations.

Another variable is when and how conflict might be a positive social phenomenon. Functionalists may wonder about the purpose of the conflict in the larger social construct. Although much more writing exists on the negative impact of conflicts, there are many who argue for the creative potential of conflict. The positive view generally takes one of two tacks. Most common are those scholars who argue that conflict leads to greater understanding, relaying of information, and potentially new ideas and stronger relationships. For them, what matters is how conflict is handled. With proper training and skills, one can learn to turn conflict into a creative interaction. A second group argues that conflict in the form of competition can improve work quality, increase production, and improve unity. They are quick to point out that task conflict especially has this effect; interpersonal conflict rarely yields this positive effect.

### **INTERPERSONAL VERSUS GROUP CONFLICT**

Conflicts happen at all levels of society. The major classifications of levels of conflict include interpersonal, intergroup, and international. The primary distinction between intergroup and international conflicts is that nationalism provides a particular framework for and gives an intensity to conflicts that distinguish them from group conflicts in general. Intergroup conflicts may be organization based

or community based, or they may stem from racial, ethnic, or class differences among groups not recognized at the nation-state level. Fisher (1990) argues that the study of group and international conflict is not nearly as well developed as the research on conflict at the interpersonal level. He cautions against assuming that analysis derived from interpersonal experiences can be generalized to other levels. He points to basic information about how groups affect conflict and suggests that these differences necessitate different strategies for addressing and reducing conflict among groups and nations.

### **IMPACT OF GROUPS ON CONFLICT**

Group conflict is typically divided into two categories: (1) intragroup and (2) intergroup. Conflict within groups is viewed as potentially harmful since it can lead to competition for leadership, loss of focus on group purpose, and membership loss. Intragroup conflicts are typically about power and control over content, purpose, and goals. It can also emerge out of recognition of intragroup inequalities that are found unacceptable by those with less power.

The primary means of limiting intragroup conflict is maintaining a strong normative structure with effective member socialization. When conflicts arise, avoidance often leads the group to dissolve. Methods for solving conflicts within groups include facilitation or problem solving led by an external professional or perhaps a skilled insider, capitulation by one party, subdivision of the group, or re-prioritization of group goals. The latter often occurs in the presence of an outside threat. In fact, a useful response to internal conflict is the identification of a common fear or enemy. External threat increases the sense of homogeneity within the group, encourages commitment to ensuring group survival, and rationalizes stern measures to limit internal challenges to the status quo.

Intergroup conflicts are typically about control and distribution of resources (objective conflicts) or threats to identity or values (perceptual, or subjective, conflicts). Realistic group conflict theory (RCT) focuses on rational mechanisms for understanding group conflict. Social identity theory (SIT) focuses on the idea that the mere perception of belonging to a discrete group is sufficient to produce intergroup discrimination. According

to SIT, groups are the major basis for the formation of an individual's sense of identity. And groups produce a positive social identity by being favorably compared to other groups (Fisher 1990). As suggested above, intergroup conflicts can also emerge when there is a desire to quell intragroup conflict through the creation of an enemy.

Usually intergroup conflict is considered more dynamic, more likely to be intractable, quicker to escalate, and more in need of external intervention than interpersonal conflict. According to White, escalation of intergroup conflict is fed by (1) a diabolical image of the enemy, (2) a virile self-image, and (3) a moral self-image (cited in Fisher 1990). Misperceptions and stereotypes of those outside the group are confirmed by group members in an effort to solidify group identity and establish the boundaries of membership. According to balance theory, groups tend to minimize intragroup differences and exaggerate intergroup differences.

Social psychologists suggest that the magnification process stems from an audience effect. In a group, one has confirmation and support for one's position as right and justified. In a group, a clear leader often emerges as external threats grow. Because groups in conflict tend to prefer more aggressive leaders, leaders of groups find themselves pushed to take stances more extreme than they would as individuals. They are expected to succeed (win) or lose power or membership in the group. Leaders often find it difficult to negotiate on behalf of a group with diverse opinions about priorities. Members of the group are more inclined to follow orders in the face of an external threat and to be convinced to commit acts (sometimes heinous) to ensure group survival. This may help individuals' involvement in genocide, torture, and other crimes they would not consider participating in for individual gain.

### RESPONSES TO CONFLICT

Scholars in the field point to five possible responses to conflict. The actual names of these responses vary, but they are well represented by the following terms: contending, yielding, avoiding, compromising, and problem solving (integrative solutions). These categories are appropriate for all

levels of conflict—interpersonal, intergroup, and international. The first three (contending, yielding, and avoiding) tend to escalate conflict. The last two fall clearly within the purview of creative conflict resolution.

*Contending* behavior is when a party to a conflict seeks to solve the conflict at the other party's expense. Rubin and colleagues (1994) suggest that contentious tactics range from light to heavy, where "light" tactics result in acceptable or neutral consequences for the other party and "heavy" tactics impose or threaten severe consequences. Contending behavior is a key part of the upward spiral of escalation, since it usually evokes a defensive response from the other party. As long as a party believes it can win (benefits outweigh costs), it is likely to practice contending behavior.

*Yielding* responses are sometimes the partner to contending behavior. A party will yield when it is more important that the conflict be over than that the party wins. Yielding may be an initial strategy to lay the basis for getting something else that is a priority at a later time. Or perhaps the concern is just not an issue for the party that yields.

*Avoiding* conflict is practiced by a party that does not want to encourage the conflict to become manifest. Usually the hope is that if avoided, the conflict will become unimportant over time. While sometimes avoiding does allow space for conflict to fade, it rarely works to actually resolve those issues underlying the conflict. Therefore there is commonly a resurgence of conflict in another way or at a later time. Avoiding conflict can also be perceived by the initiating party as an act of escalation. As such, it can cause the upward spiral of conflict to begin.

*Compromising* about a conflict is the outcome that most people associate with conflict resolution processes. For some analysts, compromise is a sufficient method to diffuse conflict; others argue that it still falls short of the goal of a "win-win" solution. In a compromise, parties to the conflict are expected to give up some part of their needs or goals if they stand in the way of the other party's reaching important needs or goals. Each party is partially successful in achieving what it sets out to gain or protect at the beginning of the conflict. Neither one completely loses and neither one completely wins.

*Problem-solving* behavior entails engaging in a process with the goal of achieving a win-win solution to the conflict. A win-win solution is also called an “integrated solution” when the parties’ needs and goals are different but a resolution is found that meets the concerns of both. The process of problem solving is designed to encourage conflicting parties to think creatively to establish possible joint needs and goals and to work cooperatively to solve their problems. There are various methods, such as log-rolling, expanding the pie, and bridging, that can assist in achieving a win-win solution. Log-rolling encourages parties to prioritize their needs and preferences so that lower priority requirements can be negotiated as trade offs in a win-win solution. Expanding the pie requires the infusion of additional resources—often by a party external to the conflict. This makes it possible to satisfy competing needs. In bridging, parties discover mutual definitions of their needs and a single solution leave both satisfied.

#### ACHIEVING CONFLICT RESOLUTION

A central question in the study of conflict and its resolution involves the conditions that support constructive outcomes to conflict. According to Deutsch (1973), a cooperative environment breeds cooperation and a competitive environment encourages groups to compete. He also suggests that there are three conditions that encourage problem-solving behavior: (1) the arousal of the desire to solve the conflict; (2) conditions that permit reformulation of the issues of the conflict; and (3) the availability of diverse and flexible ideas. Rubin and colleagues (1994) suggest (as do many other scholars) that to turn from contentious behavior to problem solving, the parties must believe that the costs outweigh the benefits if they continue to fight. This may happen because resources are scarce, because the parties fear a permanent loss of their relationship, or perhaps because another party intervenes and provides an incentive (or threat) that encourages seeking a resolution. Sometimes parties in conflict are moved to problem solving by the introduction of a superordinate goal or by the emergence of a common or shared threat.

Establishing communication between the parties is an essential part of turning competitive conflict into cooperative problem solving. The

dynamics of groups can make this particularly challenging to accomplish. Leaders will not want to admit that they exaggerated stereotypes and negative associations to encourage diabolical imaging of the enemy. And leaders that emerged during the escalation of conflict will worry about damaging their credibility by agreeing to enter negotiations. For most people, negotiations mean compromise, which means appearing to have lost, at least partially. Communication and understanding can also be challenging across cultures. Lederach (1995) suggests that while the general progression of a conflict and its resolution remains fairly similar from one culture to the next, the specific processes and techniques for solving problems are dramatically different from one cultural group to the next.

Although it might prove challenging to establish, in any conflict communication is essential to building trust, eliciting interests, and defining the problems and possible solutions. Trust is also built through reciprocity of positive initiatives. As Kriesberg (1998) suggests, changes at three levels—within the group, between the groups, and in the larger social context—all contribute to the initiation of deescalation. Ongoing case studies help us understand how seemingly intractable conflicts can be softened through the transformation of interparty relationships and the intraparty absorption of those changes.

#### APPLICATIONS OF CONFLICT RESOLUTION FOR GROUPS

**Contact Theory.** Much effort has gone into identifying conditions that encourage the deescalation of conflict. Organized and well-structured contact between group members is regarded as essential for breaking down stereotypes and assumptions that prevent initiating discussion of the conflict(s). It is thought that this contact needs to be not simply positive but genuinely cooperative, so that new methods of interacting can be learned. Also, the parties need to enter into the contract as equals. Two studies of groups—by Sherif and by Blake and Mouton—suggest that building trust between competing groups requires the creation of superordinate goals if the groups are to move beyond stereotypes and negative assumptions (cited in Fisher 1990). Furthermore, many researchers (Fisher 1990; Rubin et al. 1994) point to the

positive function of superordinate goals or shared tasks in reducing intergroup tensions, especially bias or ethnocentrism. The cooperative nature of this structured contact can decrease perceived threat from the other group, thus supporting the reduction of bias and ethnocentrism. Much of the work in cross-cultural conflict resolution and sensitivity or diversity training operates from this belief that structured contact can be used to challenge people's assumptions about other groups.

**Third-Party Intervention (Intermediaries).**

Here, an intermediary intercedes with the goal of influencing or facilitating the settlement of a conflict. Historically force has been the common form of intervention. However, Princen (1992) argues that the intermediary role is becoming much more diverse in the contemporary international system. As an intermediary, a party does not impose a solution or have a direct stake in the outcome. However, it is possible that an intermediary has an indirect interest that the problem be solved. Thus Princen (1992) distinguishes those with an indirect stake (principal mediator) from those with no stake at all (neutral mediator). He suggests that the motivation to intervene shapes the form and content of the intervention made by a third party.

A typology of the forms of third-party intervention in conflict resolution would certainly include the following:

- **Peacekeeping.** Placement of military forces from a third party within another country on behalf of the third party or an international organization (e.g., the United Nations) to enforce or maintain a cease-fire and/or a peace agreement. While soldiers are typically armed, they limit their use of weapons to self-defense and defense of the peace agreement.
- **Arbitration.** An individual or a board of arbitrators is appointed (usually by a court or someone in a position of higher power) to listen to the concerns of the parties and elicit their suggested solutions. The arbitrator crafts an agreement to address these concerns. Parties who go to an arbitrator agree to abide by the decision, regardless of whether the outcome is considered favorable to one's concerns (binding arbitration).
- **Mediation.** Parties in conflict meet with an intermediary who facilitates a discussion about the conflict and each party's concerns. The mediator is expected to be neutral on the outcome of the issue and typically is a stranger to each party. There are exceptions in some international situations, where it is often a preference of the parties to know and trust the mediator prior to the meeting. And in some cases the international mediator may have an indirect interest in the outcome of the conflict (principal mediator). Regardless, the philosophy of mediation is that the parties solve their own conflict while the mediator encourages the process, especially the initiation of communication early in the discussions.
- **Conciliation.** Sometimes also known as "shuttle diplomacy" or "track-two diplomacy," this is mediation work that happens without bringing the parties face-to-face. Often this is the first step in developing enough trust for the parties to meet (especially at the international level). These intermediaries must establish high levels of trust and legitimacy with the parties in order to be successful. Often they are known to the parties, or are powerful international players, or are generally recognized for their success in international mediation.
- **Consultation.** In this approach, parties are brought together by a conflict resolution specialist in a workshop format to learn problem-solving skills. Usually the entire group (or significant leadership in the group) is involved. The workshop leader works with the groups to encourage a positive motivation for solving the conflict, to improve communication between the members of the groups, to help the groups define the conflict, and to regulate interaction as the groups discuss their problems. Whereas in much arbitration and mediation the concern is for agreement, here the concern is for building new, positive relationships between the parties by changing their interactions in a controlled setting (Fisher 1990).

These workshops are commonly organized around issues such as communication skills, group decision making, and cross-cultural understanding. They are commonly used with community groups or in work settings where a long-term cooperative relationship is either necessary or desirable.

**Unilateral Mechanisms.** Even without the intervention by a third party, sometimes a player in a conflict will initiate the deescalation process. This most likely occurs when there is apparently little to be gained (and much to be lost) by continuing the conflict. Unilateral moves typically begin with small confidence-building measures. These actions stand on their own and are done without the requirement of reciprocation (although that may be the hope). The goal is to break the upward spiral of conflict by opening a channel of communication or establishing a foundation for trust.

- **GRIT.** This is the acronym for “graduated and reciprocated initiatives in tension-reduction,” a strategy designed by Charles Osgood in the early 1960s that is intended to streamline unilateral measures (cited in Rubin et al. 1994). Some of his guidelines include announcing the series of unilateral initiatives ahead of time with a clear time line and continuing the announced steps even if the other party does not reciprocate. At the same time, the initiating party maintains its ability to retaliate in case the opposition responds with contentious behavior. The initiatives should reward the other party for cooperating and thus become increasingly favorable to the other party (Fisher 1990; Kreisberg 1998; Rubin et al. 1994).
- **Dialogue groups.** This long-term approach to building understanding between groups entails bringing members of groups together on a regular basis for general discussion. Often these meetings bring together people not directly engaged in the conflict. Hubbard (1997) reports the use of these groups in many communities in the United States to bring together Israelis and Palestinians. The groups typically also include members unrelated to either of the conflict groups. The groups

discuss issues generally or specifically, depending on local, national, or international events. Dialogue groups have also been used in international conflicts to bring together the “average, everyday” people who must learn once again to live side-by-side once the hostilities have ceased.

### **CURRENT STATUS OF THE FIELD AND RECOMMENDATIONS FOR FURTHER RESEARCH**

The field of conflict resolution is clearly interdisciplinary. With the end of the Cold War and the changing nature of conflict internationally, the field has become increasingly influential. Sociologists and social psychologists have contributed to theory building and the development of practical applications at all levels (interpersonal, group, international). The particular concern that sociologists have for group relations and for the alleviation of social inequalities (a common source of conflict) has made a natural fit between the areas of research. Conflict analysts are heirs to the work of Marx, Weber, Simmel, and many other influential sociologists.

A primary interest in both research and in the application of the field is in the area of ethnic conflict. Sociologists are working on these issues both in urban communities and at the international level. Another important area of interest is teaching conflict resolution theory and skills to young people. Mediation became well established both in the schools (elementary through university levels) and in the judicial system in the 1990s. There is also interest in an alternative to the adjudication system, an approach called “restorative justice.”

A challenge to researchers and practitioners is that each conflict has its unique elements. While we have been able to isolate basic concepts and “rules” of the escalation and deescalation of conflict, the nature of human beings resists a single approach to the science of conflict resolution. Most scholars acknowledge that there is more need for study of individual cases of conflicts and for the comparison of specific conflict and conflict resolution processes. This would certainly provide



more resources for the teachers and learners of conflict resolution techniques. At the same time, there is strong resistance to developing a monolithic model, since cultural, historical, and individual characteristics must be taken into account with each conflict. Much of the success of the application of the knowledge gained in the field is dependent upon the skill of the conflict resolution specialist and his or her willingness to respond to case-specific information. There is also a concern that while research is necessary to increase our understanding, the most significant work is done by those who work as practitioners of the conflict resolution process.

## REFERENCES

- Deutsch, Morton 1973 *The Resolution of Conflict*. New Haven, Conn.: Yale University Press.
- Fisher, Ronald 1990 *The Social Psychology of Intergroup and International Conflict Resolution*. New York: Springer-Verlag.
- Hubbard, Amy 1997 "Face-to-Face at Arm's Length: Conflict Norms and Extra-Group Relations in Grassroots Dialogue Groups." *Human Organization* 56(3):265-274.
- Kriesberg, Louis 1998 *Constructive Conflicts*. Lanham, Md.: Rowman and Littlefield.
- Lederach, John Paul 1995 *Preparing for Peace*. Syracuse, N.Y.: Syracuse University Press.
- Osgood, Charles 1962 *An Alternative to War or Surrender*. Urbana, Ill.: University of Illinois Press.
- Princen, Tom 1992 *Intermediaries in International Conflict*. Princeton, N.J.: Princeton University Press.
- Rubin, Jeffrey, Dean Pruitt, and Sung Hee Kim 1994 *Social Conflict: Escalation, Stalemate, and Settlement*. New York: McGraw-Hill.
- Simmel, Georg 1955 *Conflict and the Web of Group Affiliations*. New York: Free Press.

LYNNE WOEHRL

**GROUP PROBLEM SOLVING**

See Decision Making Theory and Research; Group Conflict Resolution; Interpersonal Conflict Resolution; Small Groups; Social Psychology.

**GROUP SIZE**

Some of the earliest and most basic ideas about groups in sociology concern group size. Cooley (1922) described how people universally are members of primary groups, which are small in size, face-to-face, highly intimate, cooperative, and enduring. Prototypes include families and groups of friends. Although people also participate in secondary groups, which are typically larger, less personal, more formally organized, and more limited in duration and purpose, primary group participation appears necessary for teaching children the requisites for participating in society, including language, basic social skills, values, and identities. Simmel (1950) described the profound effects on interpersonal relations when the smallest group, the dyad (pair), expands to a triad. The dyad is unique because it can be destroyed by the loss of a single member, and this feature often leads to a degree of intimacy and closeness not found in larger groups. Adding another person, a third party, to form a triad dramatically transforms the character of the relations. Simmel noted three different roles served by a third party. One role is as a nonpartisan or mediator, which can serve to draw the members of the pair closer together. For example, the birth of a child may enhance the emotional bond between the parents, or a third person may try to settle a dispute between two friends. Another role is as the *tertius gaudens* ("the third party who enjoys"). The third party benefits when two members in conflict try to win the third party's support. For example, parents who are in conflict may compete to win the affection of their child by using gifts and favors. A third role is to divide and conquer. Here the third party benefits by actively encouraging conflict in the pair. For example, a parent may gain greater control over two children by having them compete for the parent's favors and affection. Simmel thus drew attention to the fact that the triad (and larger groups) brings into play processes that are impossible in the dyad. A systematic account of the effects of group size, however, awaited the emergence of an experimental research tradition.

In sociology the term *group* has been used very loosely, referring at one extreme to small aggregations of people whose members are mutually aware of each other and can potentially interact (McGrath

1984) and at the other extreme to enormous aggregations whose members simply share some characteristic, as in the case of an "ethnic group." In studies of group processes, the more restricted definition has been used. The groups studied have generally been created explicitly for the purpose of research, rather than being ones that existed naturally for other purposes.

Group size has been studied in groups facing two very different circumstances. The first concerns groups that are formed to achieve some common purpose or goal. Interest has focused on how people behave in such groups and conditions that enhance or impede their effectiveness in achieving the goal. The second concerns groups facing collective dilemmas, where each member chooses between two actions with contrasting consequences. One action maximizes each member's own interest, whereas the other action maximizes collective interests, that is, the interests of the group as a whole. The dilemma is that each member profits more from the selfish choice, but if all members make that selfish choice each is worse off than if all chose to maximize collective interests (Komorita and Parks 1994).

### GOAL-ORIENTED GROUPS

As Steiner (1972) notes, purposes for which groups are formed can vary on a continuum from task performance at one extreme (e.g., solving a problem, creating some product) to sociability at the other (e.g., simply enjoying one another's company). Groups whose purpose is primarily sociability have been little studied. Assuming that sociability is defined by spontaneous, reciprocal conversation, enjoyment should be highest in groups with no more than five or six members, the largest group size in which this kind of interaction occurs easily. This rule may be qualified in larger groups, however, if the situation is arranged so that people interact in smaller subgroups.

Most experimental research has focused on groups that are more task oriented. Here the most consequential feature is the nature of the task, or *task structure*. The first major program studying task-oriented groups experimentally was conducted by Bales and associates (Bales 1950). Bales and other early researchers focused on groups attempting to solve decision or judgmental tasks, problems for which there is no correct answer and

where differences of opinion are likely regarding the resolution (e.g., a personnel issue facing an administrator in an organization). Groups were told to discuss the assigned problem for a few minutes and then arrive at a group recommendation. Bales's method of interaction process analysis (IPA) coded verbal and nonverbal communication among group members as they addressed two basic ongoing concerns: the task and the relationships among the members. Research by Bales and associates showed the feasibility of measuring ongoing behavior in discussion groups and was instrumental in the emergence of small-group research as a distinctive field in the social sciences.

Group size is a fundamental consideration in discussion groups because the number of possible symmetrical relations between pairs of members increases much more rapidly than the number of members added to the group. The formula is where  $x$ =the number of symmetrical relationships and  $n$ =number of members (Bossard 1945). In addition to the pair relations, there are also relations between each group member and the group as a whole. If group members interact for a fixed amount of time, adding group members forces changes in the nature of existing relations and limits the number and nature of new relations (e.g., time available with each member, degree to which conversation is reciprocal).

As group size increases, groups tend to adopt a more direct and organized approach in soliciting information, and a task leader is more likely to emerge or be chosen. The top communicator initiates a greater proportion of the group's communications, and the differences in amount of communication by other members diminish (Bales et al. 1951). An increasing proportion of communication is directed toward the group as a whole, and the number of members who participate minimally by simply listening and giving emotional expressions increases (Bales and Borgatta 1955). In addition, the larger the group, the greater the conflict, the less likely members are to reach agreement on controversial issues, the less the conformity to group norms is, and the less the members are satisfied with the group and its activities (Levine and Moreland 1998; Thomas and Fink 1961).

Latané and L'Herrou (1996) suggest that group size is important in determining whether or not

subgroups holding a minority opinion are likely to form and persist over time. As the number of members grows larger, the larger the number of members who may hold minority views, and hence the greater the opportunity for them to interact with each other and avoid exposure to majority views. The emergence of minority subgroups in groups is also affected by whether the group has an odd or even number of members. When opinion is divided, odd-numbered groups (e.g. 3, 5, or 7 members) are more likely to break into majorities and minorities, avoiding the possible even splits of even-numbered groups. With this pattern, odd-numbered groups should have fewer deadlocks and should reach decisions more quickly than even-numbered ones (Bales and Borgatta 1955).

The ideal size of discussion groups is often considered to be four to six members, a number large enough to ensure some diversity in member resources but small enough so that everyone can participate. Groups naturally occurring in public are rarely larger than five or six people, and more often contain two or three (Moreland et al. 1996).

Juries are discussion groups entrusted by the state to make decisions of the utmost consequence for the accused. Because legislation now permits juries smaller than twelve members, jury size has emerged as a variable of interest in studies of actual and mock (experimenter-created) juries. Although twelve-person juries should be less likely to reach unanimous verdicts (or verdicts based on a two-thirds majority) than smaller juries (e.g., six or eight members), the predicted difference is small and it has not been found in the studies themselves (Davis 1989).

Group size affects performance for other types of tasks as well. Steiner (1972) describes the effects for tasks having products or solutions whose adequacy can be judged objectively. Tasks are *unitary* when they cannot be readily divided into subtasks and where mutual assistance is impractical (e.g., working one arithmetic problem, hammering a nail). Tasks are *divisible* when they can be divided into subtasks, allowing different group members to work on the parts before the final product is assembled (e.g., working ten arithmetic problems, assembly-line production). Steiner identified several types of unitary tasks. A task is *disjunctive* when any single member can supply the group's product (e.g., the answer to a question). Here potential

group productivity depends on the resources of the most competent member. The larger the group, the greater the likelihood that at least one member will have the needed resources. As the group grows larger, however, each additional person adds a smaller increment to the group's potential. Productivity gains are thus a decelerating function of group size. A task is *conjunctive* when the resources of the least competent member determine group productivity (e.g., mountain climbers trying to reach the summit while roped together). Here the larger the group, the greater the likelihood that at least one member will be low in competence. As the group grows larger, productivity decreases, with each additional person adding a smaller decrement to the group's potential. Productivity losses are thus a decelerating function of group size. A task is *additive* when individual resources are added to obtain the group product (e.g., stacking firewood, offering ideas to improve a product). Here, the larger the group, the greater the group productivity.

For divisible tasks, variation can occur both in the nature of the subtasks and in the way that subtasks are combined to form the final product. A subtask can be disjunctive, conjunctive, or additive if two or more group members are working on it, and the process of assembling the subtasks can be disjunctive, conjunctive, or additive, thus creating a large number of possibilities. Where the rule for combining subtasks is additive or disjunctive, group productivity should be a positive, decelerating function of group size. The larger the group, the more likely it is that the subgroups can be staffed optimally, but communication, coordination, and matching members to the subtasks most appropriate for their interests and skills can become increasing problems. An increase in group size has the opposite effect, however, where the rule for combining subtasks is conjunctive. For tasks where the labor is divided and it is critical that each member make the correct response (e.g., an airline crew, a team of surgeons), the larger the group, the greater the chance that one of the critical members (or subgroups) will fail to perform adequately, and hence the group will perform poorly.

Steiner notes that actual task productivity often fails to reach potential productivity because groups use faulty or limiting processes to address

the task and may also fail to motivate members sufficiently. Both problems typically increase with group size. With divisible tasks, for example, assignments given to members may not provide a good match between member skills and the sub-tasks they are to perform. In face-to-face groups where each member contributes orally (e.g., giving novel ideas on some topic), larger groups derive decreasing benefits as group size increases simply because only one person can talk at a time. The best-studied example of motivational losses involves additive tasks where individual performances are simultaneous and anonymous. For example, where group members each make some response on cue (e.g., pulling on a rope, shouting), individual contributions are smaller the larger the group, an effect that is termed “social loafing.”

In addition to a task structure, most groups also have a *reward structure*—the arrangement of payoffs or rewards that motivate members to work on the task. The type and nature of the reward structure may be imposed by a third party (e.g., a supervisor, leader), be part of the group’s history, or be chosen by the group itself. In most experimental studies of task-oriented groups, the reward structure is *cooperative*: The payoff or reward is a result of the group’s efforts in meeting some criterion and is shared by all members (although not necessarily equally). The payoff may be intangible (e.g., the satisfaction of solving the problem or completing the product) or tangible (e.g., money or prizes). Reward structures can also be *competitive*, where rewards are distributed unequally to members based on *relative* individual performance. Cooperative and competitive contingencies are often compared with a third alternative, *individual* contingency, where a member receives a reward when he or she meets an individual performance criterion; this is the case in most work-for-pay groups in industry.

These reward structures are not equally appropriate for all tasks. In particular, cooperative rewards are effective across a range of tasks (Johnson et al. 1981; Qin-Zhining et al. 1995) and are uniquely appropriate when the task requires collaborative activities such as response coordination, task subdivision, or information sharing. Collaboration is rewarded with cooperative rewards, increasing the likelihood that the criterion for the group reward will be met. By contrast, competitive

rewards are effective only where task responses can be made independently by each person. With competitive rewards, blocking another’s responses, not collaboration, is likely to lead to winning. When competition is appropriate, though, it is often more cost-effective (more responses made per unit of reward), easier to implement, and capable of producing short-term performance rates that are higher than those of the other conditions (Schmitt 1987). Competitive motivation can also arise outside of the formal reward structure. Under cooperative conditions, members may work harder simply to be the best performer in the group.

A group member’s motivation to perform a task can vary greatly depending on reward structure, and group size affects important aspects of that structure. With a cooperative structure, all members are rewarded, but inequities may exist in the size of the rewards received (e.g., some people contributing to a task may get more money than others). In general, people working on tasks expect their rewards to be proportional to their contributions (Homans 1974; Walster et al. 1978). Thus, if person A and person B have similar task skills, A will be upset if A and B contributed equally to the task but A received a reward half the size of B’s. Person A will not be upset, however, if A made a contribution half the size of B’s to the task. People who are inequitably underpaid relative to their contributions often work less hard on future tasks and may choose to leave the group (Marwell and Schmitt 1975). In the 1990s professional sports revealed a number of cases where athletes earning millions of dollars per year refused to play for their teams (or played less energetically) because comparable performers on their own or other teams earned more. In cases where the total amount of the cooperative reward for completing a task is proportional to the number of group members (e.g., \$50 for a five-member group and \$100 for a ten-member group), the larger the group, the greater the potential for larger inequities, that is, several of the members receiving a large share of the reward. If large inequities are present, productivity gains are likely to be a decelerating function of group size (and member discontent may provoke a change in distribution).

In arranging for competition, the reward structure is defined by the unequal rewards distributed to winners and losers at the end of the contest. In

most contests the distribution is fixed in advance and is known to the competitors. As with inequitable cooperative structures, the variety of competitive structures depends on the number of competitors in the group, assuming that the total contest amount is proportional to the number of group members (i.e., the larger the group, the larger the contest amount). Two properties of the distribution are relevant. One is the proportion of competitors receiving rewards in each contest. At one extreme, only one competitor receives a reward; at the other extreme, all competitors are rewarded, but in varying amounts. With a single winner, the larger the size of the group, the larger the competitive reward. As has been shown with large lottery prizes, the larger the competitive reward, the greater the motivation of group members to compete, at least in a single contest. However, if there is a series of contests and a difference in competitive skills causes some members to lose continually, the lack of earnings will lead to their withdraw from the group, thus lowering group productivity (Schmitt 1998). Maximizing the proportion of competitors rewarded should encourage poorer performers to remain in contests regardless of group size. When more than one competitor is rewarded, variation can occur in a second property—reward differential or spread, that is, the difference between the highest and lowest reward amounts in each contest. Again, the larger the size of the group, the larger the total competitive reward, hence the larger the possible differential. Maximizing reward differential more highly motivates those who have a chance of winning but gives those who lose frequently less incentive to continue over a series of contests. In sum, group size is a factor when competitive reward structures specify that few are rewarded or have reward differentials that are extreme. Although increasing group size (and the reward pool) is likely to increase member motivation and productivity in initial contests, over a series of contests these gains are likely to be a decelerating function of group size, as those who earn little contribute less or quit.

With individual reward structures, the potential effects of group size are similar to those for cooperative structures. Again assuming that the total reward amount for the group is proportional to the number of members, the larger the group, the greater the size of the reward inequities that are possible in arranging the individual rewards.

With large inequities, productivity gains are likely to be a decelerating function of group size.

### SOCIAL DILEMMAS

Several types of social dilemmas have been investigated. The best-known type is prisoners' dilemma. Originally conceived as a two-person game, an  $n$ -party prisoners' dilemma in which the number playing the game can be varied has been used to study the effects of group size (Komorita and Parks 1994). Each group member has two choices: to cooperate (C), which maximizes payoffs for the group as a whole, or to defect (D), which maximizes the individual's own payoff. The actual payoffs for each member depend both on own and others' choices. Payoffs for each choice increase with the proportion of members who make the C choice, but the D always produces the higher individual payoff. Finally, the payoff if everyone chooses C is greater than the payoff if everyone chooses D, the selfish choice. Each group member chooses repeatedly over a number of trials. This basic dilemma has numerous everyday counterparts, as when commuters each prefer to use a private automobile instead of a bus, but if each does so the resulting chaos leaves everyone dissatisfied.

Other types of social dilemmas involve a pool of sources to which the group members have access (Pruitt 1998). One is the commons dilemma, based on the Tragedy of the Commons (Hardin 1968), in which a village's common land is overgrazed because of the selfish actions of the individual herdsmen. As studied experimentally, subjects take turns removing resources (e.g., money) from a pool that is replenished periodically based on the amount remaining. The pool can be productive indefinitely, provided that the subjects don't destroy it by taking all the resources. Another is the public goods dilemma, in which the resource pool is built up through individual contributions, as when people contribute money to support public television or some charity. Here the temptation is for individuals to "free ride" and let others make the contribution. As studied experimentally, subjects take turns contributing resources to a pool that is later enhanced by the experimenter and then divided equally.

Studies have found that the larger the group in social dilemmas, the less the cooperation and the greater the selfish behavior, although there is

little change in groups larger than eight members (Pruitt 1998). Various explanations of this relation have been proposed (Komorita and Parks 1994; Pruitt 1998). One possibility is that any defection breaches the trust required for cooperation, and that if one person defects, others view cooperation as unlikely and follow suit. The possibility of one person defecting increases as the group grows larger. In addition, as the group grows larger, each selfish individual response becomes less identifiable and may be seen as less responsible for the lower group payoffs. By the same logic, a cooperative response may also be seen as contributing less to the group product as the group grows larger. Finally, any opportunities by the group to communicate the group's cooperative interests or to sanction defectors are more difficult to carry out in larger groups.

### CONTEXT

Any social interaction among group members must take place in some context, for example, people standing, seated at a table, at stations on an assembly line, or in separate offices linked by a computer network. Physical contexts are very important because they affect both the nature of the interaction among members and the ease with which various members can interact. The nature of interaction among group members is determined in part by the number and kind of stimuli presented by those who interact. Stimuli may be verbal in the form of oral or written communication, or they may be nonverbal—information transmitted without using language (e.g., facial expressions, gestures, posture, appearance, voice quality, rate and temporal patterning of speaking).

The context in which a group functions determines whether verbal and nonverbal stimuli are transmitted among group members. If the context includes face-to-face contact, speakers can deliver messages using both verbal and nonverbal stimuli. Nonverbal stimuli add elements that can be both enriching and distracting. For example, a speaker's facial expressions and gestures can reveal the strength with which a position is held, the truthfulness of a message, or the degree of intimacy sought by the speaker. Aspects of the speaker's demeanor or appearance can also distract listeners from paying close attention to the message.

People differ, however, in how they use such stimuli and hence the meaning given to a message. This variability is greatly reduced if the messages are written. In addition, the quality of interaction may differ, depending on the presence of the verbal and nonverbal stimuli. For example, people are more likely to harm others (e.g., deliver punishment or bad news) when communication is restricted and nonverbal behavior is absent compared with face-to-face contact (where people are more fully personalized). In general, restricted communication does not lead to the full development of interpersonal relations, and it minimizes cues that reveal differences in status, power, and prestige among group members. Hence, differences in interaction and influence among group members tend to be less extreme than those in face-to-face groups (McGrath, 1984). Restricted communication is more likely as groups become large and more formally organized.

Another aspect of context concerns the timing of information transmitted among group members. When people interact, verbal and nonverbal stimuli may be presented synchronously (i.e., immediately and in real time), or asynchronously (i.e., at the member's own time, place, and pace). Face-to-face interaction, telephone conversations, and video conferences are synchronous. Letters, memos, videotapes, electronic mail, and answering machine messages are asynchronous. Synchronous interaction, whether in person or electronically mediated, constrains interaction because only one person normally talks at a time. With an increase in group size, a few members typically lead or dominate the conversation. Asynchronous interaction, by contrast, typically permits any number of messages to any or all members, and the fact that the messages are necessarily recorded means that they are normally available for comparison and review later.

Asynchronous communication has increased enormously in popularity in recent years as computer-mediated electronic mail has become available in all organizations and many households. Computer-mediated communication has unique facilitating features: the ability to link anyone who has a network connection and the opportunity to send messages instantaneously to any number of people at very low logistical and social cost. For

some tasks, the use of asynchronous communication changes the effects of group size on productivity, compared with traditional face-to-face synchronous communication. For example, Valacich and associates (1995) investigated groups formed to generate new ideas, that is, to “brainstorm,” where different group members have the potential for generating different ideas. In face-to-face interaction, larger groups derived decreasing benefits as group size increased from five to ten members. In computer-mediated groups, where members communicated via typed messages, larger groups derived increasing benefits as group size increased. Computer-mediated groups appear to use the advantage of reviewing the ideas of others to avoid redundant ideas and build new ones. The study of computer-mediated groups with other tasks promises to reveal further distinctions between synchronous and asynchronous interaction (see Kiesler 1997 for discussion of a number of computer- and Internet-related issues).

An intriguing and potentially distinguishing feature of computer-mediated compared with face-to-face groups is that group size itself may become difficult to define and detect (Sproull and Faraj 1997). People in face-to-face groups take up space, and their physical presence and nonverbal reactions can affect others’ behavior even if a group member says nothing. By contrast, the readers of computer-mediated messages are typically invisible, and speakers may have little notion of who has seen or read a message.

In both industry and education, computers have been used to create humanlike partners who interact with people in various ways. Kiesler and associates (1996) investigated subjects’ responses in a prisoners’ dilemma game where the “partner” was known by the subject to be either a person or computer based. In both conditions, the subject and partner discussed options on each trial, and the partner asked the subject for commitments. For both types of partners, discussion and agreements with the subjects facilitated cooperation, although the effect was stronger with the human partner. Thus, nonhuman partners can produce “social” responses in people provided that the partners make humanlike use of social stimuli and responses.

Much of the early research on context concerned constraints on who could communicate

with whom in small groups. These constraints determine the group’s communication network (for a summary, see Shaw 1981). For example, two contrasting networks in groups of three or more members are the circle, where each member can communicate with two adjacent members, and the wheel, where one member occupies a central position and others can communicate only with that central person (the “hub”). Wheels and circles are examples of highly centralized and decentralized networks, respectively. Groups studied in such networks have typically been small (e.g., three to five members), and communication has usually occurred via written notes. Comparisons of group problem solving using various networks have found that when tasks are relatively simple and require that members collect their information, centralized networks are more efficient than decentralized ones (with fewer messages and errors). Where tasks are more complex and require that members perform additional operations on their information, decentralized networks are more efficient than centralized ones (Shaw 1981). With complex tasks the communication overload (termed “saturation”) experienced by the member in the central position slows the attainment of a solution. Saturation is likely to be a problem in all networks as the group increases in size, and it should emerge more rapidly when interaction is synchronous.

Communication networks arrange the contacts among group members in a decisive manner, but even where members are in face-to-face contact, aspects of the setting can make interaction between some members more likely than others, thus producing networklike effects. Seating arrangement is an example. Groups discussing problems are frequently seated at tables. One of the earliest findings was that people tend to communicate with others across the table and facing them instead of with those seated alongside them (Steinzor 1950). As groups grow larger, people are frequently seated at rectangular tables. Studies have found that people at the table’s end positions tend to participate more, are seen as having more influence, and are more likely to be chosen as leaders than those on the sides (for a summary, see Shaw 1981). Networklike arrangements can also be created if group members are instructed to follow a particular communication pattern in a discussion (e.g., “Talk only with the leader.”). Instructions can also be used to limit style of expression (e.g.,

“Give ideas for solving the problem but don’t criticize others’ ideas.”).

In conclusion, differences in group size pose both opportunities and problems for the members. Because the nature of the consequences depends on the type of group task, reward structure, context, and the skills of the group members, little can be said about a group if only its size is known. Under certain circumstances, however, group size is highly consequential for the group’s performance and stability. It should be noted that for groups that form in everyday life, unlike those created for experimental purposes, problems posed by size do not necessarily condemn a group to subpar performance. Everyday groups usually have both pasts and futures, and are often skillful in identifying problems and in making changes that help mitigate them.

## REFERENCES

- Bales, R. F. 1950 *Interaction Process Analysis: A Method for the Study of Small Groups*. Cambridge, Mass.: Addison-Wesley.
- Bales, R. F., and E. F. Borgatta 1955 “Size of Group as a Factor in the Interaction Profile.” In A. P. Hare, E. F. Borgatta, and R. F. Bales, eds., *Small Groups: Studies in Social Interaction*. New York: Knopf.
- Bales, R. F., F. L. Strodtbeck, T. M. Mills, and M. E. Rosenborough 1951 “Channels of Communication in Small Groups.” *American Sociological Review* 16:461–468.
- Bossard, J. J. S. 1945 “The Law of Family Interaction.” *American Journal of Sociology* 50:292–294.
- Cooley, C. H. 1922 *Human Nature and the Social Order*. New York: Chas. Scribner’s Sons.
- Davis, J. 1989 “Psychology and the Law: The Last Fifteen Years.” *Journal of Applied Social Psychology* 19:199–230.
- Hardin, G. 1968 “The Tragedy of the Commons.” *Science* 162:1243–1248.
- Homans, G. 1974 *Social Behavior: Its Elementary Forms*, 2nd ed. New York: Harcourt Brace Jovanovich.
- Johnson, D. W., G. Maruyama, R. Johnson, D. Nelson, and L. Skon 1981 “Effects of Cooperative, Competitive, and Individualistic Goal Structures on Achievement: A Meta-Analysis.” *Psychological Bulletin* 89:47–62.
- Kiesler, S. (ed.) 1997 *Culture of the Internet*. Mahwah, N.J.: Erlbaum.
- , L. Sproull, and K. Waters 1996 “A Prisoner’s Dilemma Experiment on Cooperation with People and Human-Like Computers.” *Journal of Personality and Social Psychology* 70:47–65.
- Komorita, S. S., and C. D. Parks 1994 *Social Dilemmas*. Dubuque, Iowa: Brown and Benchmark.
- Latané, B., and T. L’Herrou 1996 “Spatial Clustering in the Conformity Game: Dynamic Social Impact in Electronic Groups.” *Journal of Personality and Social Psychology* 70:1218–1230.
- Levine, J. M., and R. L. Moreland 1998 “Small Groups.” In D. T. Gilbert, S. T. Fiske, and G. Lindsley, eds., *The Handbook of Social Psychology*, vol. II, 4th Ed. Boston: McGraw-Hill.
- Marwell, G., and D. R. Schmitt 1975 *Cooperation: An Experimental Analysis*. New York: Academic Press.
- McGrath, J. E. 1984 *Groups: Interaction and Performance*. Englewood Cliff, N.J.: Prentice-Hall.
- Moreland, R. L., J. M. Levine, and M. L. Wingert 1996 “Creating the Ideal Group: Composition Effects at Work.” In E. H. Witte and J. H. Davis, eds., *Understanding Group Behavior*, vol. II. Mahwah, N.J.: Erlbaum.
- Pruitt, D. G. 1998 “Social Conflict.” In D. T. Gilbert, S. T. Fiske, and G. Lindsley, eds., *The Handbook of Social Psychology*, vol. II, 4th ed. Boston: McGraw-Hill.
- Qin-Zhining, D. W. Johnson, and R. T. Johnson 1995 “Cooperative Versus Competitive Efforts and Problem Solving.” *Review of Educational Research* 65:129–143.
- Schmitt, D. R. 1987 “Interpersonal Contingencies: Performance Differences and Cost-Effectiveness.” *Journal of the Experimental Analysis of Behavior* 48:221–234.
- Schmitt, D. R. 1998 “Effects of Reward Distribution and Performance Feedback on Competitive Responding.” *Journal of the Experimental Analysis of Behavior* 69:263–273.
- Shaw, M. E. 1981 *Group Dynamics*, 3rd ed. New York: McGraw-Hill.
- Simmel, G. 1950 *The Sociology of George Simmel*; Translated, edited, and with an introduction by Kurt H. Wolff. Glencoe, Ill.: Free Press.
- Sproull, L., and S. Faraj 1997 “Atheism, Sex, and Databases, The Net as a Social Technology.” In S. Kiesler, ed., *Culture of the Internet*. Mahwah, N.J.: Erlbaum.
- Steiner, I. D. 1972 *Group Process and Productivity*. New York: Academic Press.
- Steinzor, B. 1950 “The Spatial Factor in Face-to-Face Discussion Groups.” *Journal of Abnormal and Applied Psychology* 45:552–555.
- Thomas, E. J., and C. F. Fink 1961 “Effects of Group Size.” *Psychological Bulletin* 60:371–385.
- Valacich, J. S., B. C. Wheeler, B. E. Mennecke, and R. Wachter 1995 “The Effects of Numerical and Logical



Group Size on Computer-Mediated Idea Generation." *Organizational Behavior and Human Decision Processes* 62:318-329.

Walster, E., G. W. Walster, and E. Berscheid 1978  
*Equity: Theory and Research*. Boston: Allyn and Bacon.

DAVID R. SCHMITT

# H

---

## **HAWTHORNE EFFECT**

*See* Industrial Sociology; Quasi-Experimental Research Designs.

## **HEALTH AND ILLNESS BEHAVIOR**

Health and illness behaviors are associated with level of disability, quality of life, patterns of illness, and risk of death. It is tempting to view such health-related outcomes solely through the lenses provided by the biomedical sciences; however, the behaviors that importantly shape individuals' experiences of sickness or wellness, and life or death, are more completely understood from a sociological perspective. The confluence of individuals' life histories, their personality characteristics and social experiences, and their social positions influences health and illness behaviors and tells us much about how to enhance health and well-being, and mitigate disability and sickness. An examination of health and illness behaviors, therefore, has important public health implications.

### **HEALTH BEHAVIORS**

Health behavior usually refers to preventive orientations and the positive steps people take to enhance their physical well-being and vitality. Traditionally, work in health behavior has focused on the use of preventive services such as immunizations, medical checkups, hypertension screening, and

prophylactic dentistry (Becker 1974). It also includes research on such behaviors as cigarette smoking, seat-belt use, medication adherence, substance abuse, nutritional practices, and exercise (Janz and Becker 1994).

The conventional approach to health behavior has been limited, focusing on the origins of particular behaviors damaging to health and strategies to modify them. The most widely used general model—the health belief model—conceptualizes preventive health action within a psychological cost-benefit analysis (Rosenstock 1974). The health belief model conceptualizes decisions to take positive health actions as motivated by perceived threat (either susceptibility to a particularly condition or perceptions that the condition is severe) and judgments about the barriers and benefits associated with specific changes in behavior. Behavior change is seen as following motives that are salient, in situations where people have conflicting motives, following those that are perceived as yielding valuable benefits. An important component of the model involves cues to action, since an activating stimulus often appears to be necessary in the initiation of a new behavioral sequence. Both internal (e.g., feelings of symptoms) and external (e.g., suggestions from doctors, peers, or the media) stimuli may act as cues motivating change. Over the years, this model has been expanded (Becker and Maiman 1983), but it serves more as an organizing framework for the study of preventive health behavior than as a successful predictive model. An analysis of studies that have used the health belief model to explain a variety of health

behaviors indicates that the predictive value of the model is, at best, modest; the average variance in health behaviors explained is approximately 20 percent (Harrison, Mullen and Green 1992).

A second commonly used model to explain health decisions is the theory of planned behavior, originally developed as the theory of reasoned action (Ajzen 1991; Ajzen and Fishbein 1977). Like the health belief model, the theory of planned behavior conceptualizes changes in behavior as products of the saliency of individuals' beliefs about the potential costs and benefits associated with an outcome or action. The theory of planned behavior, however, places greater weight on persons' intentions, arguing that behavior is centrally motivated by intentions that are shaped by normative beliefs, feelings of control, and judgments about the barriers and benefits associated with potential change. Again, however, the model has had only modest success in predicting an array of health behavior; the association between intentions and behavior is typically about .40 (Conner and Norman 1994).

There are many other models and theories proposed to predict health behaviors, and much has been written comparing the relative efficacy of each for predicting health behaviors (Conner and Norman 1994; Mullen et al. 1987; Weinstein 1993). It appears that efforts to develop a general theory are limited by the fact that behavior conducive to health derives from diverse and sometimes conflicting motives. Consistently, research indicates that health behavior, or a healthy lifestyle, is not a unitary construct (Johnson et al. 1998; Sobal et al. 1992). One study of health behaviors among a nationally representative sample of adult Americans examined the clustering of four health behaviors: diet quality, alcohol use, tobacco use, and physical exercise (Patterson et al. 1994). The results (based on data from the late 1980s) suggest that approximately 10 percent of Americans live a "healthy" lifestyle, defined by a good diet, low use of tobacco and alcohol, and engaging in regular physical exercise. In comparison, about 2 percent of the population practice unhealthy behavior on all four of these dimensions. Thus, most Americans fall somewhere between the two extremes, practicing some positive health behaviors while neglecting others. Research that has attempted to

establish specific clusters of related health behaviors has proved inconclusive. Some studies have identified many apparently interrelated clusters of behaviors—for example, smoking/drinking and exercise/diet—and others have identified few (cf. Johnson et al. 1998; Sobal et al. 1992). In short, the research indicates that there is no simple identifiable positive health orientation that can serve as a basis for promoting risk aversion and health maintenance.

The lack of such a general orientation results because most behaviors with important implications for health arise from motives not related to health and are significantly programmed into the daily patterns and institutional life of communities and families (Mechanic 1990). Health-protective behaviors that are consequences of accepted, everyday, conventional activities require neither conscious motivation nor special efforts to be sustained. The favorable health experience of Mormons, for example, is a product of their belief systems and patterns of activity reinforced by the way of life of this cultural community (Mechanic 1990). To the extent that health behaviors are more the result of habits than cognitive decisions, we might expect that past health behaviors are robust predictors of current health behaviors. Yet neither the theory of reasoned action nor the health belief model adequately incorporates past behavior in its model, an omission that might partially explain their lack of predictive power (Conner and Norman 1994).

Promoting health may be more a matter of changing culture and social structure than of modifying personal motives or intentions. Patterns of behavior that depend on sustained conscious motivation are less stable than those that are a natural consequence of the accepted norms and understandings within a community. Expectations not only affect the prevalence of varying behaviors but also establish constraints on the acquired behaviors of children and adolescents. Changes in the social constraints on smoking, and the growing unacceptability of smoking in varying social contexts, may have more significance than any program to change personal behavior for explaining the dramatic decline from about 42 percent of the U.S. adult population being current smokers in 1965 to 25 percent in 1995 (National Center for Health Statistics 1998).

Although there is no evidence for a unitary health orientation, some social factors, particularly socioeconomic status (SES), predict good outcomes across a wide range of health indicators (Bunker et al. 1989; Marmot 1998; Ross and Wu 1995). Occupational status, income, and education each reflect some part of SES, and all are associated with health behaviors, whether one is comparing the health behaviors of populations or of individuals. Generally, wealthy nations show the highest rates of preventative health practices, such as child immunization, routine dental care, and the use of mammography, compared with less wealthy nations. But the importance of SES may be indirect, through social conditions. Indeed, as Caldwell (1986, 1993) has argued, mothers' educational attainment appears to be particularly important, influencing health outcomes net of its relationship to per capita income. Caldwell suggests that maternal education increases women's autonomy, enhances their ability to interact efficaciously with available health services and technology (even when such technology is not advanced), and enables women to increasingly control their own health and that of family members.

Within nations, there are also important SES differences in health behaviors. Table 1 presents some examples of SES differences in the U.S. population; the behaviors presented are meant to illustrate the gap and are not an exhaustive list of possible health behaviors. As shown in Table 1, Americans with lower SES are more likely to engage in health-risk behaviors and less likely to engage in health-promoting behaviors. The precise ways in which SES affects these outcomes are not fully understood, but the consistent findings point to an explanation of health behaviors that goes beyond personal responsibility and free choice. Research that examines the impact of SES over the life course illustrates the fallacy of relying on "choice" explanations to account for SES differences in health behaviors. In a sample of Finnish men, Lynch and his colleagues (1997) examined the effects of SES during childhood (parents' occupation), adolescence (education), and adulthood (occupation) on psychosocial characteristics, such as hostility and hopelessness, that are important to health and on health behaviors, including smoking and alcohol use. They found that lower SES in childhood and adolescence is associated with greater health-risk behaviors in adulthood, in addition

to greater feelings of hopelessness and hostility. Given that childhood SES is not a matter of choice, their findings support an explanation of SES differences in health behavior rooted in persistent structural disadvantages and the accompanying differential opportunities and constraints. Others have also demonstrated that higher SES provides not only obvious economic advantages and related opportunities but also enhanced personal autonomy, increased sense of control, and greater social participation (Marmot 1998; Ross and Wu 1995, 1996), all of which also influence health outcomes.

A variety of behaviors noxious to health (smoking, drug use, and drinking) develop or increase during adolescence and young adulthood. However, young people who have a good relationship with their parents and who are attuned to parent-oriented values—as measured by school performance, attendance at religious services, and participation in meals with parents—do relatively well across a variety of health measures (Hansell and Mechanic 1990). In contrast, high engagement with peer-oriented social activities is associated with increases in behavior associated with health risk. In addition, children model their parents' health behaviors, an effect that persists at least into young adulthood (Lau et al. 1990).

Although it is apparent that adolescence is a time of life where there are likely to be changes in health behaviors, we know very little about other stages of the life course or the life transitions that may be especially important (Prohaska and Clark 1997). Prohaska and his colleagues propose a "stages of change" model that recognizes important transitions in the life course as explanations for changes in health practices (Prohaska and Clark 1997; Prohaska et al. 1994). They argue that individuals go through a number of steps, from not thinking there is a need for change to maintaining the new behavior after change. Public health efforts, therefore, could benefit from understanding what motivates or hinders a person's progression through the steps. For instance, life transitions such as motherhood may motivate progression toward positive health practices, while transitions such as death of a spouse may make it difficult to maintain health practices and, therefore, may explain deterioration of positive health behaviors among persons recently widowed. Moreover, according to this model, persons may be differently prepared to progress through the steps depending

**Socioeconomic Status and Health Behaviors in the United States**

Behavior	Education		Year
	Less Than 12 Years %	16 or More Years %	
Cigarette smoking (person 25 years and older)	35.7	13.6	1995
Dental visit last year (persons 25 years and older)*	38.0	73.8	1993
Prenatal care in first trimester (mothers 20 years and older)	68.0	93.9	1996
Heavy alcohol use (men 25–49 years)	16.3	6.1	1994/96
Overweight (women 25–74 years)	45.8	26.3	1988/94

Behavior	Income/Poverty		Year
	Below Poverty %	At or Above Poverty %	
Mammography use past two years (women 40 years and older)	44.4	64.8	1994
Dental visit past year (persons 25 years and older)	35.9	64.3	1993
No physician visit in last year (children under 6 years)**	11.2	5.2	1994/1995
Vaccination (children 19–35 months)	69.0	80.0	1996

*Table 1*

SOURCE: National Center for Health Statistics (1998).

NOTE: \*Comparison is less than 12 years versus 13+ years.

\*\*Comparison is poor versus nonpoor.

on their stage in the life course; for example, older persons may be better prepared to contemplate the health risks of smoking or heavy alcohol use, while adolescents are not. The important point for our purposes is that understanding how to modify health practices requires appreciation of how stages in the life course and transitions may influence when people are willing or able to make changes.

**ILLNESS BEHAVIOR**

The study of illness behavior, in contrast to health behavior, is concerned with the way people monitor their bodies, define and interpret bodily indications, make decisions about needed treatment, and use informal and formal sources of care (Mechanic 1986, 1995). Like other behavior, illness behavior is learned through socialization in families and peer groups and through exposure to the mass media and education. There is great diversity of attitudes, beliefs, knowledge, and behavior, all of which affect the definitions of problematic symptoms, the meanings and causal attributions that

explain them, socially anticipated responses, and the definition of appropriate remedies and sources of care. Motivation and learning affect the initial recognition of symptoms, reactions to pain, the extent of stoicism and hypochondriasis, and the readiness to seek release from work, school, and other obligations and to seek help (Mechanic 1978).

Illness behavior begins prior to the use of services with the recognition of illness or sickness. While a complex array of variables might explain variations in interpretation of sickness, they can be summarized in ten general categories: (1) the visibility, recognizability, or perceptual salience of deviant signs and symptoms; (2) the extent to which the person perceives the symptoms as serious (that is, the person's estimate of the present and future probabilities of danger); (3) the extent to which symptoms disrupt family, work, and other social activities; (4) the frequency of the appearance of deviant signs and symptoms, or their persistence, or their frequency of recurrence; (5) the tolerance threshold of those who are exposed

to and evaluate the deviance signs of symptoms; (6) the information available to, the knowledge of, and the cultural assumptions and understanding of the evaluator; (7) the degree to which autistic psychological processes (perceptual processes that distort reality) are present; (8) the presence of needs that conflict with the recognition of illness or the assumptions of the sick role; (9) the possibility that competing interpretations can be assigned to the symptoms once they are recognized; and (10) the availability of treatment resources, their physical proximity, and the psychological and monetary costs of taking action (including not only physical distance and costs of time, money, and effort, but also stigmatization, resulting social distance, and feelings of humiliation resulting from a particular illness decision) (Mechanic 1978).

In short, illness appraisal is a two-step process (Mechanic 1972). In the initial step, persons monitor their bodies to assess the location, duration, intensity, and persistence of discomfort. In the second stage, which may occur almost concurrently, they seek explanations for perceived changes. If an obvious explanation is not available, or is disconfirmed by further checking, individuals look to their environment for new cues and explanations. These interpretations, in light of knowledge and other beliefs, may then play a role in the formal initiation of care.

Individuals' appraisals of symptoms vary importantly. Processes of symptom appraisal are influenced by the manner in which symptoms occur and their characteristics, by knowledge, and by past experiences with illness (Leventhal 1986; Mechanic 1972). Some symptoms are so painful and incapacitating that they inevitably lead to intervention without significant inquiry. Others are so familiar and generally understood as self-limited that they also are dealt with routinely. Many symptoms, however, are neither familiar nor easily understood, resulting in a process of interpretation within the context of personality, situational cues and stressors, and environmental influences. Only a small proportion of symptoms lead to formal consultation or care. The vast majority are denied, normalized, or evaluated as having little significance

Much research has focused on how persons come to make judgments of their own health status. One approach is to study health appraisals

among children and adolescents who have little serious illness. In a prospective study of adolescents, those who were more competent and more engaged in age-related activities, as measured by school performance and participation in sports and other exercise, rated their health more highly (Mechanic and Hansell 1987). Adolescents' health assessments are shaped by their overall sense of functioning, and they do not seem to differentiate among physical and psychological aspects of well-being in making general assessments of how they feel.

This finding is consistent with the body of research examining the impact of global judgments of health on mortality and disability. Several longitudinal studies of the elderly have found that subjective self-assessments of health predict future mortality after taking account of known risk factors and sociodemographic measures (Idler and Angel 1990; Idler and Benyamini 1997; Idler and Kasl 1995; Mossey and Shapiro 1982). Likewise, some research suggests that global assessments of health predict future level of functioning or disability (Farmer and Ferraro 1997; Idler and Kasl 1995). It is remarkable that simple self-assessments of health prospectively predict longevity and disability better than physicians' assessments or known health-risk factors. Many studies suggest that judgments of health and well-being and interpretations of sickness are shaped by factors beyond those traditionally captured by biomedical conceptions of illness. Individuals' appraisals of their health appear to depend as much on their global sense of well-being as they do on specific patterns of illness. The identification of the exact psychological and social factors responsible for self-assessed health, however, remains elusive. Some have suggested that individuals take into account important social and psychological resources, such as social support, feelings of control, and optimism, when making judgments of their own health—and that these psychosocial resources provide protection against morbidity and mortality (Idler and Benyamini 1997; Kaplan and Camacho 1983).

One of the most consistent research findings has been that perceptions of physical illness depend on the person's psychological well-being or level of psychological distress (Aneshensel et al. 1984; Farmer and Ferraro 1997; Tessler and Mechanic 1978). Individuals tend to assess their health

holistically in terms of vitality and capacity to perform their social activities and roles. Psychological distress often diminishes vitality and functioning as much as serious medical conditions. The RAND Medical Outcomes Study found that depressive symptoms were more disabling than many chronic physical conditions that physicians view as extremely serious (Wells et al. 1989).

Persons seeking medical care commonly express their distress and lowered sense of well-being through many diffuse physical complaints, such as fatigue, insomnia, and aches and pains in different bodily systems, a process referred to as somatization (Kleinman 1986). Although much of the existing literature focuses on somatization as a problematic process, it is by far the predominant pattern for expressing distress. Until the last forty or fifty years, it was uncommon to conceptualize distress in psychological terms, and even now such expressions are used primarily among well-educated populations receptive to psychological interpretations. General distress has both physical and psychological concomitants; the language that people use to characterize distress depends on the cultural context, the perceived appropriateness of psychological complaints, and the stigma attached to emotional disorder.

There is controversy as to whether psychological idioms are inaccessible to many people due to cultural factors or limited schooling, or whether somatization represents a choice among alternative idioms because such presentations are seen as more consistent with the medical care context. Rates of reported depression in Chinese cultures, for example, are extremely low, although "neurasthenia" is a common diagnosis in Chinese medical care settings (Kleinman 1986). Psychiatrists in China routinely view neurasthenia as a "disorder of brain function involving asthenia of cerebral cortical activity," but the symptoms reported are strikingly similar to the physical manifestations of depressive disorders more commonly seen in Western countries. It remains unclear whether these diagnoses characterize the same underlying disorders that are expressed differently in varying cultural groups or whether they are fundamentally different. Kleinman (1986) treated Chinese patients diagnosed as neurasthenic who met the American diagnostic criteria for major

depression with antidepressant drugs but found on follow-up that while a majority showed significant improvements in clinical psychiatric symptoms, they continued to be impaired, to function badly, and to seek help for their condition. They also remained skeptical of the drug treatment they received. Kleinman links these responses to the patients' needs for the medical legitimization of their "illness" to explain past failures and to justify continuing difficulties in meeting social expectations.

Similar examples are provided within the American context when "new" disorders become part of our popular understanding of what is (and is not) physical illness. For example, in recent years we have seen much debate around the legitimacy of chronic fatigue syndrome and Gulf War illness as physical disorders (Abbey and Garfinkel 1991; NIH 1994; Presidential Advisory Committee 1996). The case definitions associated with these conditions are not well specified, and the etiology of each remains undetermined. Those suffering from the symptoms of each condition have benefited greatly from the efforts of interest groups who have lobbied to have the conditions treated as unique clinical disorders. The status of disorder often brings access to insurance coverage and other entitlement programs. Equally important, recognition of a group of symptoms as a specific disorder or distinct illness acts to reduce stigma for those suffering and provides some legitimacy to persons disabled by the symptoms. Studies of patients diagnosed with chronic fatigue syndrome highlight the important role of attribution in the experience of illness. Some research indicates that patients who view their illness as essentially physical have less favorable outcomes than those who attribute the cause of their illness to social or psychological factors (Joyce et al. 1997).

In situations where there are no obvious explanations for the occurrence of symptoms, individuals seek meanings for changes in their feeling states. The commonsense theories they apply may either be idiosyncratic or drawn from socially prevalent conventional explanations such as stress, lack of sleep, overwork, and overeating. These lay explanations are influential on subsequent behavior such as care seeking and use of medication (Kleinman 1980; Leventhal et al. 1980; Leventhal et al. 1985). For example, it is commonly believed that stress

increases blood pressure and that relaxation reduces it. Most individuals, however, cannot assess whether their blood pressure is high or low on the basis of available cues, yet many believe they can. Persons with hypertension, an asymptomatic condition, commonly use self-assessments of their stress levels or relaxation as an indicator of their blood pressure levels and adjust their medication accordingly, despite medical advice to the contrary (Leventhal et al. 1985). Similarly, many patients with limited understanding of the biological processes through which drugs such as antibiotics or antidepressants act increase or decrease medication in relation to changes in how they feel and environmental cues.

The decision to seek medical care in response to symptoms also depends on prevailing norms about the conditions that are within the purview of the medical field. Such norms differ across cultures and time. In recent years, the role of the popular media in influencing help-seeking appears to have increased. In the late 1980s, Prozac was widely advertised and discussed as a treatment for a wide range of symptoms ranging from clinical depression to shyness, self-criticism, low self-esteem, or just feeling blue (Barondes 1994; Kramer 1993). It became the most quickly accepted drug in U.S. history (Kramer 1993). In April 1998 Viagra was introduced to the market, and the popular press was inundated with stories of the potential of the drug to increase sexual pleasure. It replaced Prozac as the fastest-selling pharmaceutical in history, and by the end of the year approximately 7 million prescriptions had been written for some 3 million patients in the United States (Pfizer Inc. 1999). Unmet need partially explains the rapid acceptance of both drugs. However, it is also likely that variations in human feelings or behavior that would previously have been normalized were more likely to be viewed as medical conditions requiring treatment.

Sickness is an accepted role in society, bringing sympathetic attention and legitimate release from expected performance (Parsons 1951). Determinations of what is (and what is not) illness may involve intense negotiations about individuals' claims that, when legitimized, may justify failure to meet expectations or allow escape from onerous obligations (Mechanic 1978). In some

situations, the sick role becomes a point of tension and conflict between the claimant, who seeks legitimation of sickness with its special privileges, and other interested parties, including families, employers, and welfare administrators, who may seek to limit release from social obligations or diminish special privileges granted to the sick and disabled (Field 1957). Most illness situations are neither problematic nor sources of conflict, but the contested cases make evident the social assumptions and expectations around which illness is organized.

Certification of illness becomes a public issue when physicians, government bureaucrats, or managed care companies have moral and legal authority to define illness and disability and to sanction the sick role. Such influence is found in certifying justified absenteeism for employers, in litigation, and in decisions on eligibility in insurance and disability entitlement programs. Efforts are often made to maintain the illusion that these are objective decisions based solely on medical expertise and clinical experience, but judgments often depend on whom the physician represents. The state or other formal organizations may thus attempt to control physicians by limiting their discretion, as happened in the Soviet Union when physicians were viewed as allowing people excuses to escape work too easily (Field 1957).

American examples of the processes involved in legitimation of illness come from observations of legislative changes affecting definitions of what is (and is not) disability. For example, in 1996 Congress passed legislation that removed substance use as a disability qualifying for Supplemental Insurance Income (SSI). Approximately 140,000 people lost their official status as "disabled" under the Social Security Act and concomitantly lost the right to income and medical benefits (Gresenz et al. 1998). Similarly, substance-use disorders associated with illegal drugs are explicitly excluded from the classes of disorders covered by the Americans with Disability Act (Mechanic 1998a). The removal or exclusion of a class of illness from the definition of disability highlights the importance of societal norms and values in sanctioning which groups of persons may legitimately occupy the sick role. Moreover, much of the debate surrounding implementation of both acts concerned judgments of



personal responsibility and “badness” and “goodness,” rather than the clinical significance of substance-use disorders as disabling conditions. In contrast, veterans of the Persian Gulf War have been awarded medical and disability entitlements based on Gulf War illness, despite the weight of medical evidence suggesting that there is no such unique clinical entity (NIH 1994; Presidential Advisory Committee 1996). Both examples illustrate the point that the certification of persons as legitimately sick (and thus deserving of public benefits) is not always the result of clinical judgments. Definition of illness and disability is the “rope” in a tug of war in which competing parties seek determinations in their own interest.

The study of illness behavior has many applications in research, clinical care, public health, and social policy (Mechanic 1998b). Such patterns of behavior substantially affect pathways into care and selectively shape the samples studied in clinical contexts. Failure to understand these selection effects and how they operate leads to erroneous conclusions about the nature of basic disease processes. A common error is to attribute to the etiology of the illness influences triggering use of services, a problem that plagues much research on stress and illness. At the clinical level, awareness of how people construe illness, present symptoms, and respond to care can improve understanding and communication and help health professionals to guide the treatment regimen more effectively.

Health and illness behavior studies make clear that the forces affecting health and treatment outcomes transcend medical care and the transactions that take place between doctor and patient. In recent decades there have been increasing tendencies toward the medicalization of social problems and a failure to address the complicated attitudinal needs of patients with serious chronic conditions and disabilities. Moreover, the problems in managing the frailties of old age, characterized by a combination of medical, instrumental, and social needs, increasingly challenge assumptions of treatment focused on narrow definitions of disease and care. Studies of health and illness behavior teach the importance of moving beyond initial complaints and narrow definitions of problems and toward examining the broad context of individuals' lives and the factors that affect social

functioning and quality of life. They point to the diverse adaptations among persons with comparable social and physical debility and potential. They reinforce the need to take account of the environmental and social context of people's lives, their potential assets, and their disease. A medical care system responsive to these broad concerns would be better prepared for the impending health care challenges of the new millennium.

## REFERENCES

- Abbey, S. E., and P. E. Garfinkel 1991 “Neurasthenia and Chronic Disease Syndrome: The Role of Culture in the Making of a Diagnosis.” *American Journal of Psychiatry* 148:1638–1646.
- Ajzen, I. 1991 “The Theory of Planned Behavior.” *Organizational Behavior and Human Decision Processes* 50:179–211.
- , and M. Fishbein 1977 “Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research.” *Psychological Bulletin* 84:888–918.
- Aneshensel, Carol S, Ralph R. Frerichs, and George Huba 1984 “Depression and Physical Illness: A Multiwave, Nonrecursive Causal Model.” *Journal of Health and Social Behavior* 25:350–371.
- Barondes, Samuel H. 1994 “Thinking About Prozac.” *Science* 263:1102–1103.
- Becker, Marshall, (ed.) 1974 *The Health Belief Model and Personal Health Behavior*. Thorofare, N.J.: Slack.
- , and Lois Maiman 1983 “Models of Health Related Behavior.” In D. Mechanic, ed., *Handbook of Health, Health Care and the Health Professions*. New York: Free Press.
- Bunker, John, Deanna Gomby, and Barbara Kehrler, (eds.) 1989 *Pathways to Health: The Role of Social Factors*. Menlo Park, Calif.: Kaiser Family Foundation.
- Caldwell, John 1986 “Routes to Low Mortality in Poor Countries.” *Population and Development Review* 12:171–220.
- 1993 “Health Transition: The Cultural, Social, and Behavioural Determinants of Health in The Third World.” *Social Science and Medicine* 36:125–135.
- Conner, Mark, and Paul Norman 1994 “Comparing the Health Belief Model and the Theory of Planned Behavior in Health Screening.” In D.R. Rutter and Lyn Quine, eds., *Social Psychology and Health: European Perspectives*. Brookfield, Vt.: Avebury.
- Farmer, Melissa M., and Kenneth F. Ferraro 1997 “Distress and Perceived Health: Mechanisms of Health Decline.” *Journal of Health and Social Behavior* 39:298–311.

- Field, Mark G. 1957 *Doctor and Patient in Soviet Russia*. Cambridge, Mass.: Harvard University Press.
- Gresenz, Carole Roan, Katherine Watkins, and Deborah Podus 1998 "Supplemental Security Income (SSI), Disability Insurance (DI) and Substance Abusers." *Community Mental Health Journal* 34(4): 337-346.
- Hansell, Stephen, and David Mechanic 1990 "Parent and Peer Effects on Adolescent Health Behavior." In Klaus Hurrelman and Friedrich Losel, eds., *Health Hazards in Adolescence*. New York: DeGruyter.
- Harrison, J. A., P. D. Mullen, and L. W. Green 1992 "A Meta-Analysis of Studies of the Health Belief Model with Adults." *Health Education Research* 7:107-116.
- Idler, Ellen L., and Ronald J. Angel 1990 "Self-Rated Health and Mortality in the NHANES-I Epidemiologic Follow-Up Study." *American Journal of Public Health* 80:446-452.
- Idler, Ellen L., and Yael Benyamini 1997 "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38:21-37.
- Idler, Ellen L., and Stanislav V. Kasl 1995 "Self-Ratings of Health: Do They Also Predict Change in Functional Ability?" *Journal of Gerontology: Social Sciences* 50B:S344-S353.
- Janz, Nancy K, and Marshall H. Becker 1984 "The Health Belief Model: A Decade Later." *Health Education Quarterly* 11:1-47.
- Johnson, Marilyn F., Jeanne F. Nichols, James F. Sallis, Karen J. Calfas, and Melbourne F. Hovell 1998 "Interrelationships Between Physical Activity and Other Health Behaviors Among University Women and Men." *Preventative Medicine* 27:536-544.
- Joyce, J., M. Hotopf, and S. Wessely 1997 "The Prognosis of Chronic Fatigue and Chronic Fatigue Syndrome." *Quarterly Journal of Medicine* 90:223-233.
- Kaplan, George, and Terry Camacho 1983 "Perceived Health and Mortality: A Nine-Year Follow-Up of the Human Population Laboratory Cohort." *American Journal of Epidemiology* 117:292-504.
- Kleinman, Arthur 1980 *Patients and Healers in the Context of Culture*. Berkeley: University of California Press.
- , 1986 *Social Origins of Distress and Disease: Depression, Neurasthenia, and Pain in Modern China*. New Haven, Conn.: Yale University Press.
- Kramer, Peter D. 1993 *Listening to Prozac*. New York: Viking.
- Lau, Richard R., Marilyn Jacobs Quadrel, and Karen A. Hartman 1990 "Development and Change of Young Adults' Preventative Health Beliefs: Influences from Parents and Peer." *Journal of Health and Social Behavior* 31:240-259.
- Leventhal, Howard 1986 "Symptom Reporting: A Focus on Process." In Sean McHugh and T. Michael Vallis, eds., *Illness Behavior: A Multidisciplinary Model*. New York: Plenum.
- , Daniel Meyer, and David Nerenz 1980 "The Common-Sense Representation of Illness Danger." In Stanley Rachman, ed., *Contributions to Medical Psychology*. New York: Pergamon.
- , Thomas Prohaska, and Robert Hirschman 1985 "Preventive Health Behavior Across the Life Span." In James Rosen and Laura Solomon, eds., *Prevention in Health Psychology*. Hanover, N.H.: University Press of New England.
- Lynch, J. W., G. A. Kaplan, and J. T. Salonen 1997 "Why Do Poor People Behave Poorly? Variation in Adult Health Behaviours and Psychosocial Characteristics by Stages of the Socioeconomic Lifecourse." *Social Science and Medicine* 44(6):809-819.
- Marmot, Michael G. 1998 "Contributions of Psychosocial Factors to Socioeconomic Differences in Health." *The Milbank Quarterly* 76:403-448.
- Mechanic, David 1972 "Social Psychological Factors Affecting the Presentation of Bodily Complaints." *The New England Journal of Medicine* 286:1132-1139.
- 1978 *Medical Sociology*, 2nd ed. New York: Free Press.
- 1986 "Illness Behavior: An Overview." In S. McHugh and T. M. Vallis, eds., *Illness Behavior: A Multidisciplinary Model*. New York: Plenum.
- 1990 "Promoting Health." *Society* 27(2): 16-22.
- 1995 "Sociological Dimensions of Illness Behavior." *Social Science and Medicine* 34:1345-1350.
- 1998a *Mental Health and Social Policy: The Emergence of Managed Care*, 4th ed. Allyn and Bacon.
- 1988b "Cultural and Organizational Aspects of Application of the Americans with Disabilities Act to Persons with Psychiatric Disabilities." *The Milbank Quarterly* 76:5-23.
- , and Stephen Hansell 1987 "Adolescent Competence, Psychological Well-Being and Self-Assessed Physical Health." *Journal of Health and Social Behavior* 28:364-374.
- Mossey, Jana M, and Evelyn Shapiro 1982 "Self-Rated Health: A Predictor of Mortality Among the Elderly." *American Journal of Public Health* 72:800-808.
- Mullen, Patricia D., James C. Hersey, and Donald C. Iverson 1987 "Health Behavior Models Compared." *Social Science and Medicine* 11:974-981.

- National Center for Health Statistics 1998. *Health, United States, 1998, with Socioeconomic Status Health Chartbook*. Hyattsville, Md.: Author.
- NIH Technology Assessment Workshop Panel 1994 "The Persian Gulf Experience and Health." *Journal of the American Medical Association* 274:391-396.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- Patterson, Ruth E., Pamela S. Haines, and Barry M. Popkin 1994 "Health Lifestyle Patterns of U.S. Adults." *Preventative Medicine* 23:453-460.
- Pfizer Inc 1999 "Pfizer Inc Fourth-Quarter Total Revenues Increased 26 Percent; Net Income from Continuing Operations, Excluding Unusual Items, Increased 42 Percent." Company Press Release, January 19.
- Presidential Advisory Committee on Gulf War Veteran's Illness 1996 *Presidential Advisory Committee on Gulf War Veteran's Illness: Final Report*. Washington, D.C.: U.S. Government Printing Office.
- Prohaska, Thomas R., and Melissa A. Clark 1997 "Health Behavior and the Human Life Cycle." In David S. Gochman, ed., *Handbook of Health Behavior Research. Vol. III. Demography, Development, and Diversity* New York: Plenum.
- Prohaska, Thomas R., W. Velicer, J. Rossi, M. Goldstein, B. Marcus, W. Rakowski, C. Fiore, L. Harlow, C. Redding, D. Rosenbloom, and S. Rossi 1994 "Stages of Change and Decisional Balance for 12 Problem Behaviors." *Health Psychology* 13:39-46.
- Rosenstock, Irvin 1974 "The Health Belief Model and Preventive Health Behavior." In Marshall Becker, ed., *The Health Belief Model and Personal Health Behavior*. Thorofare, N.J.: Slack.
- Ross, Catherine E., and Chia-ling Wu 1995 "The Links Between Education and Health." *American Sociological Review* 60:719-745.
- 1996 "Education, Age and the Cumulative Advantage in Health." *Journal of Health and Social Behavior* 37:104-120.
- Sobal, Jeffery, Dennis Revicki, and Bruce R. DeForge 1992 "Patterns of Interrelationships Among Health-Promotion Behaviors." *American Journal of Preventative Medicine* 8:351-359.
- Tessler, Richard, and David Mechanic 1978 "Psychological Distress and Perceived Health Status." *Journal of Health and Social Behavior* 19:254-262.
- Weinstein, Neil D. 1993 "Testing Four Competing Theories of Health-Protective Behavior." *Health Psychology* 12:324-333.

- Wells, Kenneth, Anita Stewart, Ron D. Hays, Audrey Burnam, William Rogers, Marcia Daniels, Sandra Berry, Sheldon Greenfield, and John Ware 1989 "The Functioning and Well-Being of Depressed Patients: Results from the Medical Outcomes Study." *Journal of the American Medical Association* 262:914-919.

DAVID MECHANIC  
DONNA D. MCALPINE

## HEALTH AND THE LIFE COURSE

*Health* and *the life course* are two broad concepts of interest to sociologists. Each of these concepts must be nominally defined.

### CONCEPTIONS OF HEALTH

Health can be conceptualized in three major ways: the medical model (or physical definition); the functional model (or social definition); and the psychological model (or the subjective evaluation of health: Liang 1986). In the medical model, health is defined as the absence of disease. The presence of any disease condition is determined by reports from the patient, observations by health practitioners, or medical tests. The social definition of health is derived from Parsons's (1951) work and refers to an individual's ability to perform roles, that is, to function socially. Illness or impairment is a function of reduced capacity to perform expected roles, commonly measured in terms of activities of daily living (ADLs—eating, dressing, bathing, walking, grooming, etc). The psychological model, or the subjective evaluation of health, is often based on the response to a single question asking one to rate one's health on a scale from poor to excellent. The definition of health used by the World Health Organization since 1946 reflects this multidimensional perspective: "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity."

It has been suggested (e.g., Schroots 1988) that a distinction be made between *disease* and *illness*. It is argued that disease refers to an objective diagnosis of a disorder, while illness refers to the presence of a disease plus the individual's

perception of and response to the disease. Thus, one may have a disease, but as long as one does not acknowledge it and behave accordingly (e.g., take medicine), one will perceive oneself as healthy (Birren and Zarit 1985).

A distinction should also be made between *acute* and *chronic* conditions. These two types of health conditions are differentially related to older and younger age groups (discussed more below). That is, there is a morbidity shift from acute to chronic diseases as an individual ages. In addition, Western societies experienced a dramatic shift from infectious diseases (a form of acute condition) to chronic, degenerative diseases in the late nineteenth century and the first half of the twentieth century.

### CONCEPTIONS OF THE LIFE COURSE

The life course is a progression through time (Clausen 1986), in particular, social time. Social time is a set of norms governing life transitions for particular social groups. These transitions may vary from one group to another (e.g., working class versus middle class) and from one historical period to another. The life-course approach focuses on “age related transitions that are *socially created, socially recognized, and shared*” (Hagestad and Neugarten 1985, p. 35). Historical time plays a key role in life-course analysis because of the emphasis on social time and social transitions (Elder 1977; Hareven 1978). Changes that take place in society lead to a restructuring of individual life courses. Thus, life courses will vary from one cohort (generation) to the next.

The life-course perspective should be differentiated from the life-span perspective or other developmental models of psychology. In these latter approaches the focus is on the individual, especially on personality, cognition, and other intrapsychic phenomena (George 1982). In these developmental approaches, change results from within the individual, and this change is universal—it is a function of human nature. Typically, developmental changes are linked to chronological age, with little or no reference to the social context or the sociohistorical or individual-historical context. The life-course perspective, in contrast, focuses on transitions when the “social persona” (Hagestad and Neugarten 1985, p. 35) undergoes change.

### CONCEPTIONS OF AGING

In order to understand health and the life course it is also important to understand the aging process. Aging is best understood in a life-course perspective. Persons do not suddenly become old at age sixty or sixty-five or at retirement. Aging is the result of a lifetime of social, behavioral, and biological processes interacting with one another. While genetics may play a part in predisposing individuals to certain diseases or impairments, length and quality of life have been found to be highly dependent on behaviors, lifestyles, and health-related attitudes (e.g., Haug and Ory 1987).

A distinction is often made between primary and secondary aging (see Schroots 1988). Primary aging, or normal aging, refers to the steady declines in functioning in the absence of disease or despite good health. Secondary aging, or pathological aging, refers to the declines that are due to illnesses associated with age but not to aging itself. This suggests that secondary aging can be reversed, at least in principle (Kohn 1985).

### VARIATIONS IN HEALTH AND LIFE EXPECTANCY

The largest cause of death in America for people under age forty-five is accidents and adverse effects (National Center for Health Statistics [NCHS] 1999). For people five to fourteen years of age and twenty-five to forty-four years of age, malignant neoplasms (tumors) rank second as a cause of death. For persons fifteen to twenty-four years of age, homicide, followed closely by suicide, are the next leading causes of death.

For adults ages sixty-five and over the causes of death are quite different. Cardiovascular disease, malignant neoplasms, cerebrovascular disease, and chronic obstructive pulmonary disease are the most common causes of death (NCHS, 1999). Older persons, too, are more likely to suffer from chronic, and often limiting, conditions. Most common among these are arthritis, hypertension, hearing impairments, heart conditions, chronic sinusitis, visual impairments, and orthopedic impairments (e.g., back). Interestingly, these same conditions are among the most commonly mentioned by persons ages forty-five to sixty-four,

though their prevalence is generally considerably less than among persons sixty-five and older.

At the turn of the century, life expectancy was about 48 years. By 1950, life expectancy was 68 years (66 years for males and 71 years for females). By 1997, life expectancy had increased to 76.5 years (74 years for males and 79 for females). Many of the improvements in life expectancy came about before large-scale immunization programs. These programs largely affected the health of those born during the 1940s and 1950s. These programs have, however, reduced infant mortality and reduced the likelihood of certain debilitating diseases (e.g., polio).

The chance of surviving to old age with few functional disabilities is strongly related to socioeconomic position, educational level, and race (Berkman 1988). People in lower classes and with less education have higher mortality risk and have higher incidence and prevalence of diseases and injuries. They have more hospitalizations, disability days, and functional limitations.

Life expectancy also varies by social class. At age twenty-five life expectancy, for those with four or fewer years of education, is forty-four years for men and almost forty-seven years for women. For men and women with some college education, life expectancy is forty-seven years and fifty-six years respectively. After age sixty-five, however, this relationship becomes less clear-cut, suggesting that for older cohorts a different set of factors is involved.

Another area where health and the life course intersect is that of health inequalities in the life course. Increasing evidence indicates that many illnesses in middle and later life have their beginnings in childhood or prior to birth (Wadsworth 1997). For example, low birthweight indicates poor prenatal growth, and both are associated with higher risk of respiratory problems in adult life. Further, lower birthweight is associated with poorer health practices of expectant mothers, suggesting that these babies will be born into family/household/social environments that do not facilitate optimal health.

Life course trajectories vary by key social characteristics such as age, gender, race, and socioeconomic status (SES) (Bartley et al. 1997). Further,

these life-course trajectories are related to disease risk. For example, an individual born into a poorer or working-class family will have a different trajectory of disease over the life course than an individual born into a family with better financial and social means. The former individual accumulates risks or disadvantages over the life course that begin to show up in adulthood.

## ISSUES AND IMPLICATIONS

Differences in health conditions by age raise at least two issues regarding the analysis and understanding of health. First, it has been suggested that in trying to understand the health and health behavior of the elderly, especially as our models become more complex, the *individual* is the critical unit of analysis (Wolinsky and Arnold 1988). That is, we must focus on individual differentiation over the life course. Aging is a highly individual process, resulting from large inter- and intraindividual differences in health and functioning.

The second issue concerns the extent to which many processes thought to be life-course processes may in fact be cohort differences (see Dannefer 1988). An assumption is often made that the heterogeneity within older cohorts is an intracohort, life-course process: Age peers become increasingly dissimilar as they grow older. This conclusion is, however, often based on cross-sectional data and may lead to a life-course fallacy. Age differences may reduce to cohort differences. If each succeeding cohort becomes more homogeneous, older cohorts will display greater heterogeneity compared to younger cohorts. Evidence suggests that for several cohort characteristics this may be the case. For example, there has been increasing standardization of years of education, age of labor-force entry and exit, age at first marriage, number of children, and so on. Thus, younger age groups would exhibit less diversity than older cohorts.

Not all health deterioration is a normal process of aging. Some of it appears to be the result of an accumulation of life experiences and behaviors. Many of the experiences and behaviors are different for older and younger cohorts, suggesting that an understanding of factors affecting health for older cohorts may not hold for younger cohorts as they age.

Two possible scenarios exist. One is that older people in the future will experience less morbidity than today's elderly, even though later life will be longer. That is, they will be older longer, sick for a very short period of time, and then die. An alternative situation is one in which elderly live longer and are sick or impaired for many of those years. That is, they will be sick for an extensive period of their later life. Given the fact that more people are living longer, the general expectation is that the demand for health care by the elderly will be greater in the future. The extent of the demand will depend, in part, on which of these two scenarios is closer to the truth. The conservative approach, and the one generally adopted, is that estimates of tomorrow's needs for care are based on data from today's elderly. However, a life-course perspective might yield quite a different picture because the life experiences, behaviors, and health attitudes of today's elderly may be quite different from those of younger cohorts, tomorrow's elderly.

Health is a lifelong experience of an individual, composed of accumulations of risk. Logically, reducing health problems in middle and later years will be more successful the earlier attempts are made to reduce these risks (Wadsworth 1997). Similarly, health inequalities exist, and these are associated with social and biological characteristics of individuals (e.g., gender, race, SES). Reducing health inequalities will require reducing risk factors for these different groups of individuals and doing so early in life.

(SEE ALSO: *Aging and the Life Course; Health and Illness Behavior; Life Course; Life Expectancy*)

## REFERENCES

- Bartley, Mel, David Blane, and Scott Montgomery 1997 "Health and the Life Course: Why Safety Nets Matter." *British Medical Journal* 314:1194-1196.
- Berkman, Lisa F. 1988 "The Changing and Heterogeneous Nature of Aging and Longevity: A Social and Biomedical Perspective." In G. L. Maddox and M. P. Lawton, eds., *Varieties of Aging*, vol. 8 of *Annual Review of Gerontology and Geriatrics*. New York: Springer.
- Birren, James E., and Judy M. Zarit 1985 "Concepts of Health, Behavior, and Aging." In J. E. Birren and J. Livingston, eds., *Cognition, Stress and Aging*. Englewood Cliffs, N.J.: Prentice-Hall.
- Clausen, John A. 1986 *The Life Course: A Sociological Perspective*. Englewood Cliffs, N.J.: Prentice-Hall.
- Dannefer, Dale 1988 "Differential Gerontology and the Stratified Life Course: Conceptual and Methodological Issues." In G. L. Maddox and M. P. Lawton, eds., *Varieties of Aging*, vol. 8 of *Annual Review of Gerontology and Geriatrics*. New York: Springer.
- Elder, Glen H., Jr. 1977 "Family History and the Life Course." *Journal of Family History* 2:279-304.
- George, Linda K. 1982 "Models of Transitions in Middle and Later Life." *Annals of the Academy of Political and Social Science* 464:22-37.
- Hagestad, Gunhild O., and Bernice L. Neugarten 1985 "Age and the Life Course." In R. H. Binstock and E. Shanas, eds., *Handbook of Aging and the Social Sciences*. New York: Van Nostrand Reinhold.
- Hareven, T. K. 1978 *Transitions: The Family and the Life Course in Historical Perspective*. New York: Academic Press.
- Haug, Marie R., and Marcia G. Ory 1987 "Issues in Elderly Patient-Provider Interactions." *Research on Aging* 9:3-44.
- Kohn, R. R. 1985 "Aging and Age-Related Diseases: Normal Processes." In H. A. Johnson, ed., *Relations Between Normal Aging and Disease*. New York: Raven Press.
- Liang, Jersey 1986 "Self-Reported Physical Health among Aged Adults." *Journal of Gerontology* 41:248-260.
- National Center for Health Statistics (NCHS) 1999 "Deaths: Final Data for 1997." *National Vital Statistics Report*. Vol. 47 No. 19. Hyattsville, Md.: Public Health Service.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- Schroots, Johannes J. F. 1988 "Current Perspectives on Aging, Health, and Behavior." In J. J. F. Schroots, J. E. Birren, and A. Svanborg, eds., *Health and Aging: Perspectives and Prospects*. New York: Springer.
- Wadsworth, M. E. J. 1997 "Health Inequalities in the Life Course." *Social Science and Medicine* 44:859-869.
- Wolinsky, Fredric D., and Connie Lea Arnold 1988 "A Different Perspective on Health Services Utilization." In G. L. Maddox and M. P. Lawton, eds., *Varieties of Aging*, vol. 8 of *Annual Review of Gerontology and Geriatrics*. New York: Springer.

DONALD E. STULL

## HEALTH CARE FINANCING

*See* Health Care Utilization and Expenditures; Health Policy Analysis; Medical Sociology.

## HEALTH CARE UTILIZATION AND EXPENDITURES

The U.S. health care system is unique among industrialized nations because it lacks a national health insurance program. The United States relies instead on private health insurance that individuals or companies must purchase. The public insurance system is limited to those who are aged or disabled (Medicare) and to some individuals who are poor (Medicaid). The delivery system for health care in the United States is almost entirely private, with only a small sector of government providers who primarily target the poor and uninsured population. Moreover, the U.S. health care system is largely unplanned and has limited regulation, even though the government is a large payer of services. In contrast, other industrialized nations have national health systems that provide coverage for their populations that are generally independent of employment (Blendon et al. 1995). These systems are more comprehensive and less expensive than the U.S. system, which excludes 43 million people (U.S. Bureau of the Census 1998b). This article examines the factors that have contributed to a system that severely limits access to millions of individuals in the United States.

We first examine health care utilization and expenditure patterns at the macro level. This approach is used by public policy makers at the federal, state, and local levels and by private insurers in decision making about the allocation and control of public and private resources for health care services. The second approach is a micro-level focus on the patterns of use and expenditures by individuals or groups of people in society and the factors that limit (or improve) access to services.

### GROWTH IN EXPENDITURES

The most significant feature of health care in the United States during the 1980s and 1990s was its overall high growth rates and its total expenditures. Total national health expenditures increased by 283 percent between 1980 and 1990 (Smith et al.

1998). Although the annual growth rate in spending slowed from 11 percent in 1990 to 4.8 percent in 1997 (Levit et al. 1998), expenditures were over \$1 trillion in 1996 and are projected to exceed \$2 trillion by the year 2007 (Smith et al. 1998). Health care expenditures in 1997 amounted to \$4,000 per capita (Table 1).

Health expenditures increased from 12.2 percent of the gross domestic product (GDP) to 13.5 percent in 1997 (Levit et al. 1998), and they are expected to reach 16.6 percent of GDP in the year 2007 (Smith et al. 1998). The devotion of a large percentage of the total GDP to health costs is a concern because such dollars are then not available for other needs.

**Factors Increasing Expenditure Growth.** Levit and colleagues (1998) have examined three basic components in the growth of health spending. The first component is the economy-wide increase in prices and inflation (about half of the total increases in the 1995–1997 period). The second is the change in volume and intensity of services. Levit and colleagues (1998) found that the volume and intensity of service use were about the same in the 1990s (about one-third of the total expenditure increases in 1997). Finally, there is medical price inflation over and above these other factors. In the 1990–1995 period, excess medical prices were important factors—almost one-third of the total increases. Between 1995 and 1997, this factor was only about one-sixth of the total increase (Levit et al. 1998). Even so, excess medical prices, or the “greed factors,” are an important focus for policy makers because they represent the area where savings may be targeted.

Although health economists rarely focus on administrative costs, some scholars are now paying attention to the role of high administrative costs in the United States. For example, Hellander and colleagues (1994) estimated that administrative overhead accounted for 25 percent of total U.S. health system expenditures in 1993. This included the costs of insurance overhead, hospital administration, nursing home administration, and physicians’ billing and overhead. There is some evidence that such costs are increasing. The excess bureaucracy is related to the complex payment and billing systems that have been created in the United States. Consolidating the 1,500 insurers in the United States could reduce overhead costs.

**National Health Expenditures by Category of Funds for 1990, 1997  
and Projected for 2007**

(in millions)

<i>Source of Funds</i>	<b>1990*</b>	<b>Year 1997*</b>	<b>2007**</b>
Total health expenditures	\$699.4	\$1,092.4	\$2,133.3
<b>Total health services and supplies</b>	674.8	1,057.5	2,085.3
Hospital care	256.4	371.1	649.4
Physician services	146.3	217.6	427.3
Other professional services	40.7	160.0	229.7
Nursing home care	50.9	82.8	148.3
Home health care	13.3	32.3	66.1
Prescription drugs	37.7	78.9	171.1
Other personal care services	69.4	26.3	167.3
Other supplies and services	60.1	88.5	226.1
<b>Other expenditures</b>			
Research and construction	24.5	34.9	48.0
<b>Other expenditures not included above</b>			
Program administration and net cost of private health insurance	40.5	50.0	151.3
Government public health	19.6	38.5	74.9

*Table 1*

SOURCE: Adapted from \*Levit et al. (1998); \*\*Smith et al. (1998).

Kenkel (1993) found that health maintenance organizations (HMOs) reported that 18 percent of their revenues were used for overhead and profits, compared with traditional insurance plans (14 percent) and Medicare (only 3.5 percent).

Economic analysts argue that the slower spending during the latter part of the 1990s reflected several factors. First, where employers sponsored health plans, there was a shift of many workers from indemnity health plans into managed care (Levit et al. 1998), where premiums may have been lower. Second, managed care was successful in negotiating discounts with providers because there was an excess capacity of providers. Growth in managed care enrollment was found to have only a small effect on the volume and intensity of services. Finally, the general and medical-specific inflation rates declined (Levit et al. 1998).

Ginsburg and Gabel (1998) argue that spending was down in the 1990s because of a sharp decline in provider revenues and payrolls. They show that hospital and physician spending increases in 1997 were less than the growth rate in the GDP (which was 2.8 percent in 1997). Although there was a slight increase in payroll costs between 1996 and 1997 in hospitals, most of the growth was

due to increases in hours worked, and not in number of employees. At the same time, drug costs had increased by 11.5 percent, primarily because of increases in the use of drugs rather than the price (Ginsburg and Gabel 1998). Premiums for health insurance increased by only 3.3 percent between 1997 and 1998 (Ginsburg and Gabel 1998). The premium increases showed substantial profit margins in 1992–1994, but lagged behind expenditures while profit margins in health plans dropped in 1996–1997 (Ginsburg and Gabel 1998).

Health care spending increases are expected to continue because of the growing demand for medical services, continued economy-wide inflation, and medical price inflation. See Table 1. Most of the expenditure growth is expected to take place in the private sector, with lower spending in the public sector (Smith et al. 1998).

The high growth rate for health expenditures has had several consequences. First, the growth rate has dominated the attention of policy makers at all levels as they have attempted to struggle with cost containment. Most policy efforts have been directed toward reducing program utilization and expenditures. This focus on cost containment has



set the national political agenda, relegating issues of access to health care and the quality of care to a second level of consideration. This agenda is of great concern because the number of uninsured is increasing and the quality of health care urgently needs to be improved (Chassin et al. 1998). Moreover, the high expenditure rates continue to be used as a rationale by some policy makers for why the expansion of public programs and a national health insurance policy are not financially feasible in the United States.

**Growth Through Consolidations, Mergers, and Acquisitions.** Overall health expenditures are increasing in part because of the changing structure of health care financing and delivery systems. Kassirer (1996) documented the rapid trends in consolidation of the health care industry, creating larger corporations during the 1990s. Srinivasan and colleagues (1998) reported that, in 1997 alone, there were 483 mergers and acquisitions involving health services companies totaling \$27 billion and 33 mergers and acquisitions involving HMOs totaling \$13 billion. The big winners in mergers and acquisitions are the corporate executives of these large organizations, the stockholders, the lawyers who broker the consolidations, and the health care consultants. The losers are the physicians and nurses who may lose income and even jobs. It is not clear that these mergers and acquisitions are resulting in any overall savings to the system or in improvements in the quality of care. Rather, they appear to be increasing profits and market power (Kassirer 1996).

Srinivasan and colleagues (1998) reported that the growth rate in for-profit health service companies and HMOs has increased rapidly in recent years. Between 1981 and 1997, for-profit HMOs grew from 12 to 62 percent of total HMO enrollees and from 18 to 75 percent of all health plans. Between 1987 and 1997, there were 210 offerings for health services companies valued at \$6.7 billion and 23 HMO public offerings for \$1.4 billion. Health services capitalization was \$113 billion and HMO capitalization was \$39 billion in 1997, showing dramatic growth rates. Stock prices for health care industry firms generally outperformed the overall stock market over the decade. As these changes occurred, there were also many conversions of nonprofit health care organizations to for-profit status in order to obtain capital and survive in a highly competitive environment. In 1997, there

were 81 conversion foundations with assets of \$9 billion as a result of the conversions to for-profit status (Srinivasan et al. 1998).

Every year *Forbes* magazine lists the largest health care corporations in the United States and details their stock performance. Gallagher (1999) reported \$10 billion in mergers and acquisitions in 1998, including many in the HMO and long term care sectors. The managed care stocks were reportedly healthy in 1998, climbing 15 percent over 1997 (Gallagher 1999). HMO mergers and acquisitions were slower in 1998 than in the previous year. By comparison, hospital stocks were off 22 percent on average and long-term care providers were off 56 percent (Gallagher 1999).

What scholars and policy makers have generally thus far ignored are both the economic and the social costs of these changes in the structure of the health care system. The costs associated with mergers, consolidations, and profits are not identified in the analyses of health care costs (Levit et al. 1997). As long as the health industry is profitable, we can expect growth in the private profit-making sector.

#### PRIVATE PAYERS

The payers of health services dominate the policy decision-making process in the United States. At the same time, the payers of health care have shifted over the past decade.

*Private health insurance* corporations dominate the health system because they have the lion's share of the resources and power in the health industry. For instance, private health insurance firms earned \$348 billion in premium income in 1997 in the United States (Levit et al. 1998). See Table 2. Private health insurance companies include both the traditional indemnity insurance plans as well as managed care plans. The private health insurance industry—including managed care organizations—spent \$34.5 billion (or 10 percent) of their premium income in 1997 on administration and profits.

Light (1992) has documented many of the problems with private indemnity health insurance plans. These plans have practice medical underwriting that excludes individuals with medical problems (e.g., chronic diseases) or raises the premiums for such policies to levels where individuals

**National Health Expenditures by Source of Funds for 1990, 1997  
and Projected for 2007**

(in millions)

<i>Source of Funds</i>	<i>Year</i>		
	<b>1990*</b>	<b>1997*</b>	<b>2007**</b>
Total health expenditures	\$699.5	\$1,092.4	\$2,133.3
<b>Private funds</b>			
Private health insurance	239.6	348.0	754.4
Out of pocket	145.0	187.6	310.7
Other private funds	31.6	49.7	80.8
<b>Public funds</b>			
Medicare	111.5	214.6	415.6
Medicaid	75.4	59.9	337.0
Other public	96.3	132.7	234.8

*Table 2*

SOURCE: Adapted from \*Levit et al. (1998); \*\*Smith et al. (1998).

cannot afford to purchase them. There are indirect approaches to reducing risk, such as waiting periods, copayments, and payment ceilings, along with the exclusion of certain procedures, tests, or drugs. Light (1992) also documented the problems with inaccurate or manipulative premiums and discrimination, which includes policy churning (switching policies before the waiting period ends) within group underwriting for those individuals at high risk, renewal underwriting (increasing the costs for those with new medical conditions), and selective marketing. There are many techniques to deny or delay payment for services. All these practices have negative consequences for individuals and groups seeking insurance. These problems are the result of a health care system geared toward profit and cost containment and of ineffective regulation of health insurance companies at the state level, although states do vary in their regulatory activities.

In 1996, Congress passed the Health Insurance Portability and Accountability Act (HIPAA). This legislation, sponsored by Senators Edward Kennedy and Nancy Kassebaum, was designed to address some of the problems with health insurance by making it easier for persons who have lost employer-provided insurance to qualify for other coverage. However, Kuttner (1999) concludes that HIPAA is of little help because there are no regulatory controls on insurance premiums and no subsidies for those persons who need the insurance because most have lost their jobs and the ability to

pay for the insurance. Similarly, the 1985 Consolidated Omnibus Budget Reconciliation Act (COBRA) allowed people leaving employment to pay insurance premiums out of pocket for up to eighteen months to retain their coverage, but again, this is of little help for those without the resources to pay for the insurance (Kuttner 1999).

*Managed care organizations* (MCOs) had the most dramatic growth in health care in the 1990s. As one study reported, "Managed care isn't coming; it has arrived" (Jensen et al. 1997). In 1995, 73 percent of insured U.S. workers received their coverage through an HMO, a preferred provider organization (PPO), or a point-of-service plan (POS) (Jensen et al. 1997, p. 126). See Table 3. HMOs are health plans that provide care from an established panel of providers to an enrolled population on a prepaid basis. PPOs are health plan arrangements where networks of providers agree to accept plan payment rates and utilization controls. POS plans are HMOs where members may self-refer outside the established network but are required to pay deductibles and other costs. This means that conventional, or fee-for-service (FFS), plans have experienced a precipitous decline in recent years. Between 1993 and 1995, the percentage of workers in conventional health plans dropped from 49 to 27 percent (22 percent), while significant growth occurred in HMOs (6 percent), PPOs (5 percent), and POSs (11 percent) (Jensen et al. 1997, p. 126). Eighty-five percent of the U.S. work force was covered by some type of managed care plan in 1997 (Levit et al. 1997).

**Percentage of Workers in Various Types of Plans, 1993-1995**

Plan Type	Year	
	1993	1995
Conventional	48.9%	27.4%
Health maintenance organization (HMO)	22.4%	27.5%
Preferred provider organization (PPO)	19.6%	25.0%
Point-of-service plan (POS)	9.1%	20.1%

Table 3

SOURCE: Jenson et al. (1997).

There have been numerous studies of HMOs and their performance (Miller 1998; Miller and Luft, 1994, 1997). Davis and colleagues (1995) surveyed individuals in FFS plans in comparison to those in managed care. They found that many individuals (54 percent) had changed plans within the past three years. Of those who changed plans, 73 percent had done so involuntarily, either because of employment coverage changes or changes in jobs or moves. Of those persons in managed care who had changed plans, 41 percent also reporting having to change physicians involuntarily, compared with only 12 percent of enrollees in fee-for-service coverage.

Although most respondents were reasonably satisfied with their health insurance, those in managed care were more likely to rate their plan as "fair" or "poor" and were less like to rate their plan as "excellent." Those individuals in managed care plans reported less access to specialty care, less availability of emergency care, more waiting times for appointments, and less overall plan satisfaction (Davis et al. 1995).

Other research on the quality of managed care services points to concerns for populations with particular health needs. Miller (1998) found mixed results for access to care in HMOs among enrollees with chronic conditions and diseases. He found enough negative results to raise serious concerns about access. Nelson and colleagues (1997) found less satisfaction with the amount of services received among enrollees who used HMO services. For populations with a variety of needs, the concerns are also mounting. For example, studies of elderly ill persons revealed that HMO enrollees report poor quality of care and lower utilization

than FFS enrollees (Clement et al. 1994; Shaughnessy et al. 1994). The problems that the elderly confront reveal spaces where government programs (i.e., Medicare) and private institutional practices (HMOs) need expanded coverage and improved coordination.

Between 1989 and 1995, the enrollment of Medicare beneficiaries in HMOs grew rapidly, so that HMOs covered about 10 percent of the total beneficiaries and the remaining persons received traditional FFS care (Welch, 1996). The Balanced Budget Act of 1997 made a number of changes in the Medicare program, called Medicare+Choice, that were intended to expand Medicare beneficiaries' options (Christensen 1998). These included: (1) PPOs, where networks of providers agree to accept plan payment rates and utilization controls; (2) provider-sponsored organizations (PSOs), which are plans organized by affiliated health care providers; and (3) medical saving accounts (MSAs) established by enrollees to receive the Medicare amount that would be paid to an HMO. Enrollees must pay all costs out of the MSA that would be covered by Medicare as well as out-of-pocket expenses (Christensen 1998). Congress expected that the Balanced Budget Act would accelerate the growth in HMO enrollment to nearly 34 percent by 2005. In spite of these changes to implement this new program, growth may be slower than projected because some HMOs are unwilling to accept Medicare members because they consider Medicare rates to be too low. In 1998, HMOs announced that they were dropping approximately 400,000 of their 6 million enrollees because they consider the rates too low (Pear 1998).

There are two advantages of HMO enrollment for Medicare beneficiaries. The first is that premium costs have generally been minimal and the second is that some HMOs include prescription drug coverage and other benefits not covered by Medicare. Recently, however, drug benefits for Medicare beneficiaries have been reduced or capped by a number of HMOs (Kuttner 1999).

HMOs have also dramatically captured the Medicaid enrollment in many states. President Bill Clinton streamlined the process that allowed states to have waivers of the federal regulations that provide mandatory enrollment of Medicaid beneficiaries in HMOs (Iglehart 1995). The expectation

was that HMOs would control Medicaid costs and encourage more coordinated forms of care by increasing the likelihood of having a primary care physician. The Kaiser Commission on the Future of Medicaid (1995) summarized 139 articles, books, and reports on Medicaid managed care and found mixed results. They found that the use of specialty services and emergency room care declined, but there was not a consistent increase or decrease in use of physicians' services or hospital care. The issue of cost savings was mixed, and quality and satisfaction rates were considered to be comparable. On the other hand, Iglehart (1995) documented serious problems with the implementation of Medicaid managed care in some states, including Florida and Tennessee.

Pauly (1998) argues that for-profit health plans enjoy monopsony power in the health care market. And while a monopsony can result in lower provider premiums, it may also reduce benefits as the providers reduce their spending on care. Pauly concludes that traditional antitrust policies have not addressed the problem of monopsony or monopoly power.

Woolhandler and Himmelstein (1995) view the growth in HMOs as a transformation that not only places physicians at financial risk because of cost-containment practices but also pressures physicians to withhold needed care. These financial arrangements encourage physicians to collaborate in risk selection by seeking to avoid patients who may have high care needs. Mechanic and Schlesinger (1996) have argued that the impact of managed care has had a negative impact on patients' trust of medical care and their physicians. The insurance industry and the managed care organizations hold great power and authority over the care provided to individuals and groups. They also have a great deal of influence over physicians, other health professionals, and health care workers. The current trends can have serious negative consequences for access and the health status of the population.

### PUBLIC PAYERS

The two major public payers are Medicare and Medicaid, programs established in 1965. Medicare pays for those individuals who are aged or disabled under Title 18 of the Social Security Act. Medicaid

is generally provided to persons age 65 and over who are entitled to Social Security and to individuals who are disabled for a period of at least 24 months. There were 38 million people eligible in 1997. Medicaid, under Title 19 of the Social Security Act, is funded jointly by federal, state, and local governments. It provides coverage to those persons who are eligible for Supplemental Security Income, for low-income individuals who are aged, blind, or disabled. It also pays for recipients and their children under the Aid to Families with Dependent Children, which is now the Temporary Assistance for Needy Families program. The program provided assistance to 36 million persons in 1996 (Waid 1998). There are other public payers, such as the Veterans' Administration and local governments, that are not discussed here in detail.

**Medicare.** In 1997, Medicare paid \$214.6 billion for its 38.4 million aged and disabled enrollees (Levit et al. 1998). Medicare is the single largest public program, accounting for 20 percent of the nation's health expenditures. See Table 2. The growth rates in Medicare spending declined from 12.2 percent in 1994 to 7.2 percent in 1997. This decline was the result of slower price inflation and policy changes. Physician payments were limited, and fraud and abuse detection programs were increased. Limits on payments and aggressive abuse detection programs are credited with decreasing home health care spending (Levit et al. 1998). Future declines in nursing home spending are expected because of the adoption of a prospective payment system that was implemented in 1998.

In the 1980s and 1990s, Medicare spending growth had been much lower than private health insurance annual growth rates. But in 1997, Medicare spending growth rates were 5.8 percent compared with 3.8 percent for private health insurance (Levit et al. 1998). Unfortunately, this growth in spending has contributed to public policy attention on reducing spending for the Medicare program.

And although spending for Medicare has increased, the program pays for only a limited amount of health insurance for those who are aged and disabled. Data from the 1997 Current Population Survey indicate that Medicare beneficiaries paid 67 percent of their total expenditures. Private insurance paid for only 10 percent of the bills,

even though a large proportion of beneficiaries have private supplemental insurance. The remainder was paid out of pocket by beneficiaries (Carrasquillo et al. 1999).

When the Medicare, Medicaid, and other public programs are combined, their expenditures represent 46 percent of total national expenditures. Because this is such a large percentage of the total, the large increases in public spending per year contribute to the public policy focus on cost containment.

**Medicaid.** Medicaid is a program designed for the poor as a safety net. Medicaid spending was \$159.9 billion in 1997, or a 3.8 percent increase over the previous year (Levit et al. 1998). Other governmental expenditures for health care (federal, state, and local) totaled \$132.7 billion. These public programs represented 27 percent of total U.S. expenditures. See Table 2. There was a rapid growth in Medicaid expenditures in the 1988–1994 period, related to the growth in (1) the number of enrollees, (2) the spending per enrollee, and (3) increases in spending for hospitals with disproportionate shares of Medicaid patients. The spending slowed in the 1994–1997 period because the number of Medicaid enrollees dropped during this period. The states also were allowed to have waivers to implement mandatory enrollment in managed care plans.

Congress passed welfare reform legislation in 1996 that replaced the Aid to Families With Dependent Children (AFDC) with Temporary Assistance for Needy Families (TANF). This legislation broke the link between welfare and Medicaid eligibility, but states must continue to provide benefits to persons who met the eligibility requirements for Medicaid prior to the legislation. The law changed the definition of childhood disability under the Supplemental Security Income (SSI) program, curtailed benefits to legal immigrants, and banned Medicaid eligibility for qualified aliens admitted to the United States after the legislation. Some benefits were restored to legal immigrants under the Balanced Budget Act of 1997 (Levit et al. 1998).

The passage of TANF means that welfare limits will gradually become effective in most states during 1999, and millions of women will lose their eligibility for welfare benefits. Although states may

continue to offer Medicaid eligibility to those individuals who go off welfare and are working in low-income jobs, many former welfare recipients are not informed of their eligibility for Medicaid services and thus have not retained their benefits. Those former welfare recipients who successfully obtain employment will generally be placed in relatively low-wage occupations that will probably not offer health insurance or job security. Overall, the legislation will result in a loss of health insurance for many women and children (Kuttner 1999).

In response to the health needs of children, Congress passed the Children's Health Insurance Program (CHIP) as a part of the 1997 Balanced Budget Act. This program provided states with \$24 billion over five years to expand coverage of Medicaid for children or to establish a new children's health insurance program. The new insurance program is complex and difficult for people to understand and gain access to. In 1998, it was reported that the program was dramatically undersubscribed for a number of reasons across the country. California, for example, had enrolled only about 4 percent of the 580,000 children eligible for the program (Kuttner 1999). Some immigrants were reported to be concerned that enrollment of their children could affect their immigration status.

Several barriers restrict access to this program, particularly for immigrant populations. First, a large state bureaucracy and complicated application process have limited access to services. Second, eligibility coverage varies for children of different ages, so that one child may be covered while another is not in a single family. Many immigrant parents declined to participate in the program because they felt that providing health care for one child while not obtaining coverage for another was immoral (Altman 1998). Finally, eligibility for this program was initially tied to citizenship, which created access problems because, although their children were often documented and legal residents, undocumented immigrant parents declined to apply to the program for fear of deportation by immigration authorities (Korenbrodt et al. 1999).

The Congressional Budget Office estimated that the Children's Health Insurance Program and the Medicaid expansions could extend coverage to

2 million of the 10.6 million uninsured children in 1997 (Kuttner 1999). It is uncertain if more than a fraction of this total will ever receive coverage.

### CONSUMER POWER

Consumers have seen a drop in their out-of-pocket spending for medical care over the past two decades. In 1980, out-of-pocket spending was 24 percent of the total expenditures in contrast with 17 percent of spending in 1997 (Levit et al. 1998). The total out-of-pocket spending was \$187.6 billion in 1997. Out-of-pocket costs include premiums for insurance, copayments, and deductibles for services use, and direct costs for services not covered by insurance, such as outpatient drugs and long-term care. The reason that out-of-pocket spending is lower is that some consumers are also paying a lower proportion of their income on premiums for health insurance. Managed care plans generally had more limited copayments and deductibles than indemnity insurance, and more individuals were enrolled in managed care (Ginsburg and Gabel 1998). Out-of-pocket costs did, however, go up in 1997 because some HMOs were reportedly increasing their copayment and deductible charges (Levit et al. 1998). Consumer spending for drugs increased as well.

Some studies have reported that 61 percent of the population has employment related coverage (AHCPR 1998; Levit et al. 1998). These estimates group individuals who have public and private employment together and tend to count anyone with private insurance even if that insurance is not the primary payer of health care. New data from the 1997 Current Population Survey showed that only 43 percent of the population had their health insurance paid for by private-sector employers, 34 percent have publicly funded insurance, 7 percent purchased their own insurance, and 16 percent were uninsured (Carrasquillo et al. 1999). Employers also contribute to the Medicare Hospital Insurance Trust Fund through Social Security payroll taxes, but these account for only a small amount of the total expenditures.

Carrasquillo and colleagues (1999) point out that the role of employers in paying for insurance has been exaggerated and, as a consequence, employers have played a dominant role in public policy discussions related to health insurance. For

example, President Clinton's national health insurance plan proposed to link health insurance to that paid for by private employers, even though employers are not the major payers of insurance. Employers played a role, along with many other interest groups, in defeating the Clinton plan because of fears that their costs would be increased and they would lose control over their work force (Navarro 1995). At the same time, employers are given an estimated \$100 billion in tax subsidies in order to cover the costs of employer-sponsored health insurance (Reinhardt 1997). The \$100 billion in subsidies could be eliminated if a national health insurance plan were paid directly by the government.

### TYPES OF UTILIZATION AND EXPENDITURES

National expenditures for health care reveal how health resources are utilized and allocated to provide care in the United States. The patterns are not necessarily the way the public would allocate the resources. Instead, the expenditure patterns reflect historical decisions about how public and private dollars should be spent.

The patterns of expenditures changed rapidly in the latter half of the 1990s because of the growth in managed care. Managed care plans have instituted their own cost controls on the providers that they pay. The health industry has also been consolidating through a large number of mergers and acquisitions. These changes have resulted in fewer larger health care provider organizations.

**Hospital Care.** Hospital care has historically been and continues to be the largest component of health spending, accounting for 38 percent of total spending (\$371 billion) in 1997. See Table 1. The spending for inpatient care dropped considerably and marked increases occurred in outpatient services. This shift represents a restructuring to outpatient care, which is less costly. In the 1990–1997 period, there was a 6-percent reduction in hospital admissions per capita and a 16-percent decline in inpatient days in community hospitals (American Healthcare Infosource 1998).

As result of the restructuring and downsizing, U.S. hospitals closed 88,000 beds (10 percent) and occupancy rates fell from 64.5 percent in 1990 to 59.6 percent in 1997 (Levit et al. 1998). Although

utilization fell and price increases were controlled, hospitals were able to keep profit margins high, since almost all their revenue is from third-party insurance (consumers paid directly for only 3 percent of hospital costs).

**Physician Services.** Costs of physicians' services were \$218 billion in 1997, or 21 percent of total health services and supplies. Spending increases for physicians were associated with the increasing dominance of managed care organizations. The American Medical Association (AMA) reported that 92 percent of all physicians were in managed care, and these contracts accounted for 49 percent of their income (Levit et al. 1998). Physicians are also expected to pay for an increasing share of ancillary services under the capitated managed care contracts. Physicians have reacted to managed care cost controls by expanding the sizes of their practices to large group practices. These practices allow for greater leverage in negotiating contracts, economies of scale, and increases in capital. More research on the effects of the changing structure of medical practices on health care access, costs, and quality is needed.

**Drugs and Medical Nondurables.** Drugs and medical nondurables accounted for \$108.9 billion in 1997, or 10 percent of total national expenditures (Levit et al. 1998). Prescription drugs as a subset grew faster than other types of health care. These costs grew at 14.1 percent in 1997 compared to 4.8 percent for health spending in 1989. Out-of-pocket drug costs were 51 percent of total payments in 1989, later dropping to only 29 percent in 1997, and the rest were paid by third parties. The switch to managed care has increased the amount of covered outpatient drugs with relatively low copayments (Levit et al. 1998). Thus, the growth in drug costs was associated with increases in the number of prescriptions—not price increases, as had been the case in the 1980s (Levit et al. 1998).

The demand for drugs continues to increase. The Food and Drug Administration (FDA) approved fifty-three new drugs in 1996 and thirty-nine in 1997—both record highs. These approvals, along with increased advertising efforts by pharmaceutical firms, may lead to increases in demand. Drug makers reported spending \$21 billion on research and development in 1998, compared with \$8 billion in 1990 (Herrera 1999). They also increased spending on advertising to \$1.3 billion

in 1998, having great successes with Viagra (for impotence), Claritin (for allergies), and Propecia (for baldness) (Herrera 1999).

**Other Services.** Nursing homes represented about 8 percent of total health service and supply expenditures in 1997, and this is projected to decline to 7 percent in 2007. Other professional services were 4.6 percent of expenditures, home health care was 3 percent, dental services were 4.6 percent, and vision products and other durable medical equipment were only 1.4 percent in 1996. These expenditures were expected to remain about the same percentage of the total over the next ten-year period (Smith et al. 1998). Governmental public health activities were estimated to be about \$35.5 billion in 1996, representing 3.4 percent of total U.S. health expenditures. Research expenditures were 1.6 percent of the total in 1996, and construction was 1.4 percent of the total (Smith et al. 1998); these were also projected to remain the same over the next ten-year period.

#### LONG-TERM CARE IN THE UNITED STATES

The only segment of the U.S. population whose cost of long-term care is fully covered is made up of those individuals below the poverty threshold who are enrolled in the state-run, federally supported Medicaid plans (Harrington et al. 1991). Many persons of moderate incomes needing long-term care are unable to afford the costs of long-term care services, which can be as much as \$50,000 per year for nursing home care. If individuals “spend down” to the poverty threshold, they can become Medicaid eligible as a last resort (Wiener 1996). This not only constitutes a hardship to the patient but creates dependence on federal and state assistance, which would be unnecessary if the entire population were insured, with premiums derived from sources other than the government.

In contrast, the nonpoor elderly enrolled in Medicare are entitled only to a limited number of skilled nursing care days (up to 100 days) if medically required following hospitalization. With some exceptions, the rest of the population must either pay for care out of pocket or purchase private long-term care insurance (Levit et al. 1997).

In 1996, national estimates for long-term care spending were \$125.5 billion (Levit et al. 1997). Of

the total expenditures, 30 percent (\$38 billion) was for home health care (including hospitals and freestanding agencies) and 70 percent was for nursing home care (\$87.5 billion; including hospital and freestanding facilities) (Levit et al. 1997). Home health expenditures are expected to double between 1996 and 2007, while nursing homes are expected to grow by 50 percent during the same period (Smith et al. 1998).

Medicaid paid for 48 percent of all nursing home care and 14 percent of all home health care in 1996 (Levit et al. 1997). Medicare paid for 45 percent of home health care and 11 percent of nursing home care. Overall, the government paid 59 percent of home health costs and 61.5 percent of nursing home costs (Levit et al. 1997). Most of the burden for government spending is from general taxes used to pay for Medicaid and a combination of general taxes and payroll taxes that support the Medicare program.

Private health insurance paid only an estimated 5 percent of nursing home care and 10.6 percent of home health care expenditures in 1996 (Levit et al. 1997). The remainder of the expenditures was paid directly out of pocket by those needing long-term care (31.5 percent for nursing home care and 19.5 percent for home health care) (Levit et al. 1997).

Private, voluntary long-term care insurance is not a viable approach to financing long-term care (Wiener 1994). Although private long-term care insurance has been available since the late 1980s, only 4.5 million long-term care insurance policies had been sold by 1994 (and not all of these policies were still in effect) (Cohen and Kumar 1997). Sold on an individual basis, private long-term care insurance is primarily attractive to persons whose health condition places them at high risk and makes them likely candidates for long-term care. Only about 10 to 20 percent of the elderly can afford to purchase long-term care insurance (Wiener 1994). Premiums for two policies purchased at age 65 were estimated to cost an average of \$3,500 per year, which would be about 13 percent of a median elderly couple's income (Consumer Reports 1997, p. 46). These models determined that private long-term care insurance is unlikely to have a significant impact on public spending for long-term care, even though such insurance is expected

to increase (Wiener and Illston 1994; Wiener et al. 1994).

A mandatory social insurance program for long-term care would have many advantages (Harrington et al. 1991). If everyone paid into the system, then individuals would have access to coverage when they are chronically ill or disabled without the humiliation of having to become poor (i.e., to "spend down") to receive services. The program might be more appealing if enrollees could have the advantage of paying in advance (prefunding) so that services would be available when needed. No stigma would be attached to receiving services, and such a program should have wide public support. The program could also reduce the access problems that are currently experienced by those who are in the Medicaid program. The financial risk would be spread across the entire population so that individual premium costs or taxes would be relatively low, in comparison to the costs of insurance purchased when individuals are older and at risk of needing long-term care.

Although many have argued that the United States cannot afford to adopt a public social insurance program, Germany mandated a social insurance program for long-term care in 1995 (Geraedts et al. 1999). This program was funded through a combination of public taxes and payroll taxes paid by employers and employees (.08 percent of wages for each or a total of 1.6 percent) (Geraedts et al. 1999). During the first three years of this program, it maintained its financial solvency and expanded long-term care to the entire population who have disabilities. This example demonstrates that such public insurance programs, if paid for by the entire population, can ensure coverage at a reasonable cost. This approach avoids restricted access, the stigma, and the stress placed on those individuals needing long term care and their families who do not have sufficient funds to pay for them in the United States (Harrington et al. 1999).

#### **THE UNINSURED AND THEIR HEALTH: MICRO-LEVEL ISSUES**

An estimated 43 million individuals in the United States have no private health care insurance (Kuttner 1999; U.S. Bureau of the Census 1998b). These individuals have no Medicare, Medicaid, or other public insurance coverage. The percent of



the population without insurance increased from 14.2 in 1995, to 15.3 percent in 1996, and to 16.1 percent in 1997 (Kuttner 1999). This represents a steady increase in the number of uninsured since the 1980s, when the Reagan administration first began cutbacks in public health expenditures. The U.S. Bureau of the Census (1998b) also estimated that about 71.5 million individuals lacked insurance for at least part of the year in 1996.

The uninsured are primarily those who are poor and members of minority groups. Of those with incomes of less than \$25,000, 24 percent had no health insurance compared with only 8 percent of those with incomes over \$75,000 (Kuttner 1999). For those from minority groups, 50 percent of Hispanics and 37 percent of blacks had at least one month without insurance coverage compared to only 25 percent of whites (Kuttner 1999). Moreover, approximately one in three children had no health insurance for part of the year during 1995 and 1996 (Families USA 1997). The rate of children lacking insurance rose between 1989 and 1996 (U.S. Bureau of the Census 1998a). Lack of insurance for children was closely correlated with income. The group with the highest uninsured rate was young adults (38 percent, or twice the rate for other Americans), because many of them are dependent on their parents, who are working in low-wage jobs (AHCPR, 1998).

#### **REASONS FOR THE GROWING NUMBERS OF UNINSURED**

There are many reasons why individuals do not have health insurance, but the primary one is that private health insurance rates are too high relative to the incomes of those who are poor. The costs of insurance would represent about 26 to 40 percent of incomes of those who are poor, so many of these individuals are unable to afford coverage (Kuttner 1999). Another major problem is that 44 percent of those individuals who lost their jobs in the 1993–1995 period also lost their health insurance coverage.

The Medicaid program does not provide health insurance to all those individuals living below the U.S. poverty rate. State Medicaid programs are allowed to establish their own eligibility standards for those on welfare (TANF) in order to limit the number who can be served. The Census Bureau reported that 49 percent of people who were fully

employed but living below the poverty line had no Medicaid or private insurance, and this number increased to 52 percent in 1996 (Kuttner 1999). Due to reductions in Medicaid coverage, the rate of poor children with insurance fell from 16.5 million to 15.5 million between 1995 and 1996 (U.S. Bureau of the Census 1998a). A U.S. General Accounting Office study (1996) found that nearly 3 million children who were eligible for Medicaid were not enrolled in the program because of inadequate outreach, fears by immigrants about immigration problems, and other barriers. Additionally, the combination of immigration reform and welfare reforms has produced a “chilling effect” on many immigrants and their health care providers (Korenbrodt et al. 1999).

Another reason that health insurance coverage is eroding is the rising cost of health insurance premiums, especially for people who purchase individual premiums privately rather than through the job (Kuttner 1999). Kuttner (1999) also cites the trends toward temporary and part-time employment, where most workers do not have insurance. The rising costs of Medigap premiums for the elderly who are on Medicare is also a major problem. Finally, the trend away from community rating of insurance to individual and group ratings makes the costs substantially higher for individuals and some groups that have higher injury or illness rates. Employers are also reducing their supplemental health coverage for retirees (from 60 percent to only 40 percent of retirees in 1995) (Kuttner 1999).

There has been a reduction in benefit coverage, particularly for pharmaceutical benefits. Many health plans are capping their outpatient drug benefits, although prescription drugs are the largest category of out-of-pocket costs for the elderly and the costs are increasing rapidly (Kuttner 1999). As noted earlier, drugs costs have been increasing rapidly (Levit et al. 1997) and many cannot afford to pay these costs. Many HMOs and insurance companies have been reducing drug benefits because of these cost increases. One study found that 84 percent of Medigap policies for the elderly had no drug coverage (McCormack et al. 1996). The result is that many elderly individuals are unable to afford drugs that are prescribed for them.

Another problem has been limited health insurance coverage for persons with mental health

problems and mental illness. Mechanic and Rochefort (1992) document the deinstitutionalization of the mentally ill from public mental hospitals and the problems associated with inadequate mental health services. Inadequate housing and community health services, as well as the limits on mental health insurance coverage have all contributed to a general decline in mental health services.

### **CONSEQUENCES OF LIMITED ACCESS TO HEALTH CARE**

The United States and South Africa continue to hold the unenviable distinction as the only two existing industrialized countries without national health insurance. Thus, when the high U.S. expenditures per capita are compared with those of other industrialized countries, the majority of these nations have their populations covered by insurance, while the United States is excluding 43 million individuals. The lack of health insurance leads to a number of serious problems in access to health care services.

There are many other health-related problems that plague the U.S. population in comparison to those of other nations. Among the top twenty-four industrialized nations, the United States ranks sixteenth in life expectancy for women and seventeenth for men, while it ranks twenty-first in infant mortality (Andrews 1995, p. 38). Although there were substantial reductions in childhood mortality in the United States between 1950 and 1993 (Singh and Yu 1996), the United States is well behind many other industrialized countries in childhood mortality (two to four times higher than Japan and Sweden) not only because of higher mortality from medical causes (e.g., heart disease) but also because of injuries and violence. Moreover, there are large differences across groups, with male children experiencing higher mortality rates and African-American children experiencing rates of more than twice that of white children (Singh and Yu 1996).

A number of studies have been conducted on the barriers to access related to the inability to pay for services (Berk et al. 1995; Braveman et al. 1989; Berk et al. 1995; Hafner-Eaton 1994; Mueller et al. 1998; Weissman et al. 1991). Mueller and colleagues (1998) confirmed that lack of health insurance—regardless of race or ethnicity, or living in a

rural environment—is the major determinant of the utilization of health care services.

Approximately 12 percent of all American families (12.8 million) experienced barriers to receiving needed health care services in 1996. Sixty percent of these families reported that the barrier to care was their inability to afford the care, and 20 percent cited insurance-related problems (AHCPR 1997). The barriers to health care included difficulties finding care, delays, or not receiving the care that was needed.

Nearly 46 million Americans were without a regular source of health care such as a doctor's office, clinic, or health center in 1996 (AHCPR 1997). Those persons without insurance were two to three times more likely to have no regular source of health care and two to three times more likely to have encountered barriers in receiving needed health care than those persons with health insurance (AHCPR 1997). Hispanics (30 percent), the uninsured under age 65 (38 percent), and young adults (34 percent) were more likely to lack a stable source of health care. Those individuals with a history of serious medical illness were twice as likely to be unable to obtain care. African Americans and Hispanics were also more likely to be unable to obtain care than whites and others (Himmelstein and Woolhandler 1995).

### **OTHER BARRIERS TO HEALTH CARE ACCESS**

There are many other barriers to access to care. One is the lack of providers in areas close to people who need care. This is a problem for those persons living in rural areas as well as for those in central- or inner-city areas (Clarke et al. 1995). Weisgrau (1995) reviewed the myriad problems rural Americans confront in obtaining health care and mental health services due to the shortages of health care providers and the continued closures of rural hospitals. The stigma of obtaining mental health services in rural areas was also considered to be a problem.

Studies by Wennberg and colleagues (1989) have documented the wide geographic variations in health outcomes. These outcomes are considered to be, in part, related to variations in provider practices. A recent study by the Department of

Veterans' Affairs found that even though the system predominantly serves low-income men, they found substantial geographic variation in service use for different diagnoses in hospitals and clinic use across the United States which they attribute to different practice styles (Ashton et al. 1999).

Another problem is cultural barriers to care. These include language differences between patients and providers and other communication problems. Escarce and colleagues (1993), Gornick and colleagues (1996), Friedman (1994), and Korenbrot and colleagues (1999) have identified these problems. Many studies have found persistent disparities across racial and ethnic groups in access to care (Escarce et al. 1993). Racial and ethnic minorities and rural residents were less likely to use physician services. This pattern was even stronger for rural Latinos and Asians (Mueller et al. 1998), particularly refugee populations. Mortality rates for black Americans are about 50 percent higher than for white Americans, and Native Americans also show very high death rates, especially among the young (Nickens 1995). In part, the high mortality rates are related to lower socioeconomic status (SES) measured by occupation, income, and education attainment, which are generally lower for minorities. Mortality rates were more strongly related to SES in 1986 than in 1960. The substantial black and white differentials in infant mortality have been well documented (NCHS, 1994). High infant mortality rates for minorities are a function not only of poverty and low SES, but also of racial segregation in housing markets that are in close proximity to industrial pollution generators (incinerators, factories, etc.). These sources of pollution produce poor air quality in many communities of color and have a direct effect on infant mortality rates and adult morbidity (Hurley 1995).

In addition to poverty, low SES, and poor environmental quality, infant mortality rates of minorities are also influenced by continuing discrimination in society. The discrimination accumulates over time and results in perceived powerlessness, frustration, and negative self-images that in turn contribute to higher mortality rates. Discrimination against minority group members by health care providers may be a contributing factor that has not been eliminated. Using simulated video interviews of patients and physicians, a recent study documented that the race and sex of patients

have an independent influence on physician decisions regarding how to manage care for chest pain. Women, blacks, and younger patients were less likely to be referred for diagnostic tests (cardiac catheterization), and the interaction effects were also significant (Schulman et al. 1999).

### **SOCIOLOGICAL MODELS OF ACCESS TO SERVICE USE**

Much of the study of health care utilization and access has been developed using the theoretical model adopted by Andersen, Aday, and colleagues at the University of Chicago (Aday and Andersen 1974, 1991; Andersen 1995; Andersen and Aday 1978).

Andersen's (1995) original model was designed to predict or explain the use of health services. Three major components were developed for the model. First, predisposing characteristics were considered to include those that were endogenous to use including (1) demographics, such as age, gender, and race; (2) social structure, such as education, occupation, and ethnicity; and (3) health beliefs, including attitudes, values, and knowledge about health. Various critics have suggested (and Andersen 1995 agreed) that social networks, social interactions, and culture are all concepts that can be incorporated into the social structural factors. Genetic makeup might also be added to the list of predisposing factors. Second, enabling factors were considered to be those community and personal resources that may enable an individual or family to use health services. These include the supply of health personnel and organizations or facilities to provide care. Enabling factors also include income, health insurance, regular source of care, travel and waiting times, and other such factors. Social relationships, or "social support," are also important enabling factors according to Andersen (1995).

Finally, the original model included the need for health care services. "Need" included both the perceived need and the need as evaluated by health professionals. In the original model, the outcomes included the amount of physician, hospital, and other health services used. This model is one of the most widely used approaches by scholars, with hundreds of articles presenting the results of testing these factors. In Andersen's original model,

access was described as potential access and realized access (actual use) of services. Access could be found to be equitable or inequitable, depending on value judgments. Inequitable access occurred when predisposing factors (such as race, ethnicity, or gender) or enabling factors (e.g., health insurance) determine who receives services rather than services determined by need. In later models, the health care system was incorporated, including national health policy, resources, and organizations. Outcomes of health services were revised to include consumer satisfaction about convenience, availability, financing, providers, and quality (Aday and Andersen 1974). The model was expanded further to include ethical criteria to measure equity, such as freedom of choice, equal treatment, and decent basic minimum care. Subsequently, the model has included understandings of effective and efficient access to services and improved health status as outcomes (Andersen 1995). Finally, Andersen sees the model as interactive and dynamic, with utilization and outcomes influencing predisposing, enabling, and need factors.

As a test of the importance of the behavioral model, a study in the late 1990s examined its use in the social science literature between 1975 and 1995. The study found 139 empirical research articles that used the model; of these articles, 45 percent included environmental variables and 51 percent included provider-related variables (Phillips et al. 1998). Extensive work has been done to examine health status as measures of health service outcomes. These include mortality, morbidity, well-being, and functioning. Outcomes can also include equity of service access. More elaborate conceptual models have been developed by the Institute of Medicine (Gold 1998; Millman 1993).

#### **ISSUES RAISED BY TRENDS IN HEALTH CARE UTILIZATION AND EXPENDITURES**

The macro-level and micro-level patterns of health care utilization and expenditure suggest that consumers, workers, the poor, people of color, the elderly, nurses, and even physicians are losing ground with respect to access to care and decision making within the health care system. Each year—indeed, every month—thousands more individuals join the ranks of the uninsured and thousands more remain *underinsured* (Kuttner 1999). At the same time, expenditures for the nation's health

care system are expected to rise steeply in the first decade of the twenty-first century. While most policy makers and health policy analysts regard the American health care system as characterized by "private insurance," less than 50 percent of health insurance is in fact paid for by private employers—and these payments are subsidized by the government at a total of \$100 billion (Carrasquillo et al. 1999). Thus, the role of employers in paying for insurance has been overstated and employers have played a prominent role in health policy discussions. And while the number of employers offering their employees health insurance has either remained unchanged or even decreased in recent years, the cost of these plans has been prohibitive for many workers, who then go without coverage. The result is a nationwide decline in persons with employer-sponsored health coverage (NCH Statistics 1994).

These trends raise questions about the role of the U.S. health care system in providing care to the nation's population. What becomes clear is that while the system is heavily supported by public funds, the benefits are accruing more and more to private corporations. Thus, we suggest that in order to understand changing patterns in health care utilization and expenditure, a political economy perspective is useful. A political economy perspective recognizes the enduring underlying conflicts among different interest groups (workers, purchasers, providers, doctors, insurers, nurses, the elderly) in gaining access to the myriad benefits a health care system provides. For example, many scholars argue that while the U.S. medical system is officially designed to provide the best possible care to any person who needs it and can pay for services, health care is actually only a secondary or tertiary function of this system (Navarro 1976; Relman 1980; Waitzkin 1983). Profit, above all else, is the "driving force" behind the health care system (Woolhandler and Himmelstein 1995).

The benefits and costs of health care in the United States are hotly contested, and stakeholders are engaged in a continuous struggle to gain access to the former and to externalize the latter. Doctors and nurses have organized unions; consumers have demanded a "patient's bill of rights" and a long list of other legislative reforms, producing a "managed care backlash"; health care providers

and insurers have consolidated in a flurry of mergers and acquisitions in the 1990s; pharmaceutical and other health care firms have sought out markets around the globe; and managers of health care systems have slashed wages and jobs in an effort to contain costs (Andrews 1995). The results have been: (1) less regulation of the health care industry; (2) increasing profits for industry; (3) declining consumer satisfaction and less access to care and insurance under managed care; (4) loss of control and autonomy by doctors and nurses; (5) rising health care expenditures; and (6) no discernible improvement in the quality of health care. Our health care system is transforming rapidly without a coherent national approach to either the structure of the delivery of care or to the problem of rising expenditures. It is incumbent upon sociologists studying the health care system of this nation and others to provide a counterbalance to the currently dominant economic cost-benefit analytical approach to this topic. A sociological perspective offers a balanced understanding of the social basis on which the health care system operates and can, therefore, lead the way toward improvements in both theory and policy making in the areas of health and health care.

## REFERENCES

- Aday, L. A., and R. M. Andersen 1974 "A Framework for the Study of Access to Medical Care." *Health Services Research* 9(2):208-220.
- 1991 "Equity to Access to Medical Care." *Medical Care* 19(12):4-27.
- Agency for Health Care Policy and Research (AHCPR) 1997 *Access to Health Care in America, 1996*. Medical Expenditure Panel Survey, May (3):1-3.
- 1998 *Health Insurance Coverage In America - 1996*. Medical Expenditure Panel Survey, May (4):1-3.
- Altman, D. 1998 "The Kaiser Family Foundation: Our Mission." Paper presented to the Robert Wood Johnson Foundation Scholars in Health Policy Research Program, University of California, Berkeley.
- American Healthcare InfoSource, Inc. 1998 *Hospital Statistics*. Chicago: American Hospital Association.
- Andersen, R. M. 1995 "Revisiting the Behavioral Model and Access to Medical Care: Does it Matter?" *Journal of Health and Social Behavior* 36:1-10.
- , and L. A. Aday 1978 "Access to Medical Care in the U.S.: Realized and Potential." *Medical Care* 16(7):533-546.
- Andrews, C. 1995 *Profit Fever: The Drive to Corporatize Health Care and How to Stop it*. Monroe, Maine: Common Courage Press.
- Ashton, C. M., N. J. Petersen, J. Soucek, et al. 1999 "Geographic Variations in Utilization Rates in Veterans Affairs Hospitals and Clinics." *New England Journal of Medicine*. 340(1):32-39.
- Berk, M. L., C. L. Schur, and J. C. Cantor 1995 "Ability to Obtain Health Care: Recent Estimates from the Robert Wood Johnson Foundation National Access to Care Survey." *Health Affairs* 14(3):139-146.
- Blendon, R. J., J. Benson, K. Donelan, R. Leitman, H. Taylor, C. Koeck, and D. Gitterman 1995 "Who Has the Best Health Care System? A Second Look." *Health Affairs* 14(4):220-230.
- Braveman, P., G. Oliva, M. A. Miller, R. Reiter, and S. Egerter 1989 "Adverse Outcomes and Lack of Health Insurance Among Newborns in an Eight-County Area of California, 1982 to 1986." *New England Journal of Medicine* 321(8):508-513.
- Carrasquillo, O., D. U. Himmelstein, S. Woolhandler, and D. H. Bor 1999 "A Reappraisal of Private Employers' Role in Providing Health Insurance." *New England Journal of Medicine* 340(2):109-114.
- Chassin, M. R., R. W. Galvin, and the National Roundtable on Health Care Quality 1998 "The Urgent Need to Improve Health Care Quality." *Journal of the American Medical Association* 280(11):1000-1005.
- Christensen, S. 1998 "Medicare+Choice Provisions in the Balanced Budget Act of 1997." *Health Affairs* 17(4):224-231.
- Clarke, L. L., C. A. Bono, M. K. Miller, and S. C. Malone 1995 "Prenatal Care Use in Nonmetropolitan and Metropolitan America: Racial/Ethnic Differences." *Journal of Health Care for the Poor and Underserved* 6(4):410-433.
- Clement, D. G., S. M. Retchin, R. S. Borwn, and M. H. Stegall 1994 "Access and Outcomes of Elderly Patients Enrolled in Managed Care." *Journal of the American Medical Association* 271(19):1487-1492.
- Cohen, M. A., and A. K. N. Kumar 1997 "The Changing Face of Long-Term Care Insurance in 1994: Profiles and Innovations in a Dynamic Market." *Inquiry* 34:50-61.
- Consumer Reports 1997 "Long Term-Care Insurance: Special Report." *Consumer Reports* (October):36-50.
- Davis, K., K. S. Collins, C. Schoen, and C. Morris 1995 "Choice Matters: Enrollees' Views of Their Health Plans." *Health Affairs* :99-112.
- Escarce, J. J., K. R. Epstein, D. C. Colby, and J. S. Schwartz 1993 "Racial Differences in the Elderly's

- Use of Medical Procedures and Diagnostic Tests.” *American Journal of Public Health* 83 (July):948–954.
- Families USA 1997 *One Out of Three: Kids Without Health Insurance, 1995–1996*. Washington, D.C.: Families USA.
- Friedman, E. 1994 “Money Isn’t Everything: Nonfinancial Barriers to Access.” *Journal of the American Medical Association*. 271(19):1535–1538.
- Gallagher, L. 1999 “Health Care Services: The Big Money is in Small Towns.” *Forbes* 163(1):182–183.
- Geraedts, M., G. V. Heller, and C. Harrington 1999 “Germany’s Long Term Care Insurance: Putting A Social Insurance Model into Practice.” Manuscript submitted for publication.
- Ginsburg, P. B., and J. R. Gabel 1998 “Tracking Health Care Costs: What’s New in 1998.” *Health Affairs* 17(5):141–146.
- Gold, M. 1998 “Beyond Coverage and Supply: Measuring Access to Healthcare in Today’s Market.” *Health Services Research* 33(3):625–652.
- Gornick, M. E., P. W. Eggers, T. W. Reilly, R. M. Mentnech, L. K. Fitterman, L. E. Kucken, and B. C. Vladeck 1996 “Effects of Race and Income on Mortality and Use of Services Among Medicare Beneficiaries.” *The New England Journal of Medicine* 335(11):791–799.
- Hafner-Eaton, C. 1994 “Patterns of Hospital and Physician Utilization Among the Uninsured.” *Journal of Health Care for the Poor and Underserved* 5(4):297–315.
- Harrington, C., C. Cassel, C. L. Estes, et al. 1991 “A National Long-Term Care Program for the United States.” *Journal of the American Medical Association* 266(21):3023.
- , G. V. Heller, and M. Geraedts 1999 “Commentary on Germany’s Long Term Care Insurance Model: Lessons for the United States.” Manuscript submitted for publication.
- Hellander, I., D. U. Himmelstein, S. Woolhandler, and S. Wolfe 1994 “Health Care Paper Chase, 1993: The Cost to the Nation, the States, and the District of Columbia.” *International Journal of Health Services* 24(1):1–9.
- Herrera, S. 1999 “Health Care Products.” *Forbes* 163(1):180–181.
- Himmelstein, D. U., and S. Woolhandler 1995 “Care Denied: US Residents Who Are Unable to Obtain Needed Medical Services.” *American Journal of Public Health* 85(3):341–344.
- 1986 “Cost Without Benefit: Administrative Waste in the U.S.” *New England Journal of Medicine* 314:440–441.
- Hurley, A. 1995 *Environmental Inequalities: Race, Class, and Industrial Decline in Gary, Indiana 1945–1980*. Chapel Hill: University of North Carolina Press.
- Iglehart, J. K. 1995 “Medicaid and Managed Care.” *The New England Journal of Medicine* 332(25):1727–1731.
- Jensen, G. A., M. A. Morrisey, S. Gaffney, and D. Liston 1997 “The New Dominance of Managed Care: Insurance Trends in the 1990s.” *Health Affairs* 16:125–136.
- Kaiser Commission on the Future of Medicaid 1995 *Medicaid and Managed Care: Lessons from the Literature*. Menlo Park, Calif.: Henry J. Kaiser Family Foundation.
- Kassirer, J. P. 1996 “Mergers and Acquisitions—Who Benefits? Who Loses?” *The New England Journal of Medicine* 334(11):722–723.
- Kenkel, P. J. 1993 “Provider Based Managed-Care Plans Continue Growth Trend.” *Modern Healthcare* (May 19):26–34.
- Korenbrodt, C., L. S. Park, R. Sarnoff, J. Greene, and C. Bender 1999 *Impact of Recent Welfare and Immigration Reforms on Low-Income Pregnant Women and Their Health Care Providers*. Unpublished report to the California Endowment. Institute for Health Policy Studies, University of California at San Francisco.
- Kuttner, R. 1999 “The American Health Care System: Health Insurance Coverage.” *The New England Journal of Medicine* 340(2):163–168.
- Levit, K. R., C. Cowan, B. R. Braden, B.R., et al. 1998 “National Health Expenditures, 1997: More Slow Growth.” *Health Affairs* 17(5):99–110.
- , H. C. Lazenby, B. R. Braden, et al. 1997 “National Health Expenditures, 1996.” *Health Care Financing Review* 19(1):161–200.
- Light, D. W. 1992 “The Practice and Ethics of Risk-Rated Health Insurance.” *Journal of the American Medical Association* 267(18):2503–2508.
- McCormack, L. A., P. D. Fox, T. Rice, and M. L. Graham 1996 “Medigap Reform Legislation of 1990: Have the Objectives Been Met?” *Health Care Financing Review* 18(1):157–174.
- Mechanic, D., and D. A. Rochefort 1992 “A Policy of Inclusion for the Mentally Ill.” *Health Affairs* 11(1):128–150.
- , and M. Schlesinger 1996 “The Impact of Managed Care on Patients’ Trust in Medical Care and Their Physicians.” *Journal of the American Medical Association* 275(21):1693–1697.
- Miller, R. H. 1998 “Healthcare Organizational Change: Implications for Access to Care and Its Measurement.” *Health Services Research* 33(3):653–680.

- , and H. S. Luft 1994 “Managed Care Plan Performance Since 1980: A Literature Analysis.” *Journal of the American Medical Association* 271(19):1512–1519.
- , and H. S. Luft 1997 “Does Managed Care Lead to Better or Worse Quality of Care?” *New England Journal of Medicine* 16(5):1–18.
- Millman, M. (ed.) 1993 *Access to Health Care in America. Institute of Medicine Committee on Monitoring Access to Personal Health Services*. Washington, D.C.: National Academy Press.
- Mueller, K. J., K. Patil, and E. Boilesen 1998 “The Role of Uninsurance and Race in Healthcare Utilization by Rural Minorities.” *Health Services Research* 33(3):597–610.
- National Center for Health Statistics (NCHS) 1994 *Health, United States, 1993*. Hyattsville, Md.: U.S. Public Health Service.
- Navarro, V. 1995 “Why Congress Did Not Enact Health Care Reform.” *Journal of Health Politics, Policy and Law* 20:455–461.
- 1976 *Medicine Under Capitalism*. New York: Prodist.
- Nelson, L., R. Brown, M. Gold, A. B. Ciemnecki, and E. Docteur 1997 “Access to Care in Medicare HMOs, 1996.” *Health Affairs* 16(2):148–156.
- Nickens, H. W. 1995 “The Role of Race/Ethnicity and Social Class in Minority Health Status.” *Health Services Research*. 30(31):151–162.
- Pauly, M. V. 1998 “Managed Care, Market Power, and Monopsony.” *Health Services Research* 33(5):1439–1460.
- Pear, R. 1998 “HMOs are Retreating from Medicare, Citing High Costs.” *New York Times* (October 2):A20.
- Phillips, K. A., K. R. Morrison, R. Andersen, and L. A. Aday 1998 “Understanding the Context of Healthcare Utilization: Assessing Environmental and Provider-Related Variables in the Behavioral Model of Utilization.” *Health Services Research* 33(3):571–596.
- Reinhardt, U. E. 1997 “Wanted: A Clearly Articulated Social Ethic for American Health Care.” *Journal of the American Medical Association* 278:1446–1447.
- Relman, A. S. 1980 “The New Medical-Industrial Complex.” *New England Journal of Medicine* 303:963–970.
- Schulman, K. A., J. A. Berlin, W. Harless, J. F. Kerner, et al. 1999 “The Effect of Race and Sex on Physicians’ Recommendations for Cardiac Catheterization.” *The New England Journal of Medicine* 340(8):618–626.
- Shaughnessy, P. W., R. E. Schlenker, and D. F. Hittle 1994 “Home Health Care Outcomes Under Capitated and Fee-For-Service Payment.” *Health Care Financing Review* 16(1):187–221.
- Singh, G. K., and S. M. Yu 1996 “US Childhood Mortality, 1950 Through 1993: Trends and Socioeconomic Differentials.” *American Journal of Public Health* 86(4):505–512.
- Smith, S., M. Freeland, S. Heffler, D. McKusick, and the Health Expenditures Projection Team 1998 “The Next Ten Years of Health Spending: What Does the Future Hold?” *Health Affairs* 17(5):128–140.
- Srinivasan, S., L. Levitt, and J. Lundy 1998 “Wall Street’s Love Affair with Health Care.” *Health Affairs* 17(4):158–164.
- U.S. Bureau of the Census 1998a *Census Brief: Children Without Health Insurance*. Washington, D.C.: U.S. Government Printing Office.
- 1998b *Health Insurance Coverage, 1997*. Washington, D.C.: U.S. Government Printing Office.
- U.S. General Accounting Office 1996 *Health Insurance for Children: Private Insurance Coverage Continues to Deteriorate*. GAO/HEHS 96-129. Washington, D.C.: U.S. General Accounting.
- Waid, M. O., U.S. Health Care Financing Administration (USHCFA) 1998 Overview of the Medicare and Medicaid Programs. *Health Care Financing Review* 1998 Statistical Supplement.
- Waitzkin, H. 1983 *The Second Sickness: Contradictions of Capitalist Health Care*. New York: Free Press.
- Weisgrau, S. 1995 “Issues in Rural Health: Access, Hospitals, and Reform.” *Health Care Financing Review* 17(1):1–14.
- Weissman, J. S., R. Stern, S. L. Fielding, and A. M. Epstein 1991 “Delayed Access to Health Care: Risk Factors, Reasons, and Consequences.” *Annals of Internal Medicine* 114(4):325–331.
- Welch, W. P. 1996 “Growth in HMO Share of the Medicare Market, 1989–1995.” *Health Affairs* 15(3):201–214.
- Wennberg, J. E., J. L. Freeman, R. M. Shelton, and T. A. Bubolz 1989 “Hospital Use and Mortality Among Medicare Beneficiaries in Boston and New Haven.” *New England Journal of Medicine* 321:1168–1173.
- Wiener, J. M. 1996 “Can Medicaid Long-Term Care Expenditures for the Elderly Be Reduced?” *The Gerontologist* 36(6):800–811.
- , and L. H. Illston 1994 “How to Share the Burden: Long-Term Care Reform in the 1990s.” *The Brookings Review* 12(2):17–21.
- , and R. J. Hanley 1994 *Sharing the Burden: Strategies for Public and Private Long-Term Care Insurance*. Washington, D.C.: Brookings Institution.
- Woolhandler, S., and D. U. Himmelstein 1995 “Extreme Risk—The New Corporate Proposition for

Physicians." *New England Journal of Medicine* 333(25):1706-1708.

CHARLENE HARRINGTON  
DAVID N. PELLOW

## HEALTH POLICY ANALYSIS

Health policy analysis is of increasing interest to sociologists in the areas of medical sociology and health services research. Health policy analysis draws on perspectives from across the social science disciplines: from anthropology and economics to political science and sociology, as well as law, medical ethics, and the applied fields of public health, public administration, and public policy. Leading sources of policy analysis are scholars in twenty to thirty university-based health policy and health services research centers and institutes and the myriad and growing number of private sector "think tanks" such as the Brookings Institution, the Urban Institute, RAND, the National Bureau of Economic Research, Project Hope, and the American Enterprise Institute. An early indicator of advances in the field of policy studies was the publication of the *Policy Studies Review Annual*, which commenced in 1977 (Nagel 1977) and continues to cover the field with an editorial advisory board made up of distinguished social scientists.

Major federal agencies that both sponsor and conduct health policy analysis include the National Center for Health Statistics (NCHS); the Health Care Financing Administration (HCFA); the Agency for Health Care Policy and Research (AHCPR); the Alcohol, Drug Abuse, and Mental Health Administration (ADMHA); the Social Security Administration; the National Institute on Aging (NIA); and the Office of the Assistant Secretary for Planning and Evaluation. Federal research funding is the mother's milk of health policy analysis; although limited, it has assured the slow but gradual accumulation of health services research knowledge.

Several journals are sources for the latest developments in health policy analysis: *Health Affairs*, *Health Care Financing Review*, *Health Services Research*, *Journal of Health and Social Behavior*, *Journal of the American Medical Association*, *Milbank Quarterly*, *International Journal of Health Services*, *New England Journal of Medicine*, and *Journal of Health Politics, Policy and Law*. A recent, comprehensive

text, *Understanding Health Policy* (Bodenheimer and Grumbach 1998), though written by physicians and with a clinical orientation, is nonetheless critical of chronic systemic tensions and inequalities in U.S. health care delivery. The authors integrate social science literature throughout, one indication of a growing consensus regarding salient problems among scholars, practitioners, and "patients" alike, in a nation shaken by unprecedented corporate intrusions into health and medical encounters.

There are multiple paradigms in and approaches to health policy analysis in schools of public policy, public health, public administration, and social work. The same diversity is present in sociology and other social science disciplines. However, we detect and discuss important areas of convergence between current controversies in U.S. health policy and perspectives and methods that are well established in sociology. We believe these areas of convergence are likely to enhance the stature and usefulness of the discipline in the analysis of health policy, in public as well as in academic life.

The various disciplines, substantive specializations, and methodologies represented in such work have contributed an array of perspectives to the definition of health policy analysis, how it is conducted, and how professional training is oriented and organized. As the number of programs offering health and related policy training has increased, the academic respectability of such work has grown apace. In sociology, vestiges of an invidious distinction between "basic" and "applied" research are still with us, and policy research is both less visible and less valued than is warranted, given its potential public impact. Nonetheless, an expanded topical definition of health policy analysis, following from the recent political and cultural tumult over changes in health care, is conducive to research in several vibrant research genres in sociology, including political economy (see, e.g., *International Journal of Health Services*), constructionist approaches to medical encounters and social problems (Brown 1995; Spector and Kitsuse 1987), phenomenology of illness and medical practice (Benner 1994), community-based studies (Israel et al. 1998), and comparative sociocultural studies of health systems (Mechanic 1996; Kleinman 1980). The latter, bordering medical anthropology, encompasses conventional treatment regimen, as



well as “self-help” and various nonbiomedical, “alternative” health practices, be they traditional or sacred (Baer 1995). Such topical breadth is also evident in “mainstream,” medically oriented outlets. In recent years, the *Journal of the American Medical Association (JAMA)* has devoted sustained attention to public health issues such as gun violence (Sinauer et al. 1996), domestic violence, motor vehicle accidents, and terrorism.

For sociologists, the primary point of entry into health policy analysis has been medical sociology, which has long been sustained by its applied relevance to and sponsorship by agencies in government and medicine (Cockerham 1988). Other contributing subfields include aging/social gerontology, political sociology, gender studies, and social stratification. Despite productive cross-pollination between these related fields of scholarship, the number of sociologists working in health policy analysis is small relative to those involved overall in studies of health care and of social policy, broadly conceived. Though medical sociologists continue to comprise one of the largest sections of the American Sociological Association (ASA)—with more than 1,000 members—their presence in the smaller Association for Health Services Research (AHSR), a major professional association, is modest: only 5 percent of its 1,400 members report primary disciplinary training in sociology, compared to 20 percent from medicine and 14 percent from economics (other members were trained in other social sciences, the allied health fields, and business). Health policy analysis is not, however, confined to conventional research roles and careers; many working in health policy analysis hold master’s degrees, are employed on the staffs of governmental and private agencies, and are not oriented toward academic theory or publication (Luft 1999).

Sociologists’ limited involvement in health policy analysis reflects the sources of, and agendas driving, health services research funding. Many problems in the planning and administration of large, complex programs favor orientations and methodological skills others can best provide, primarily those in economics and business. It reflects as well an unfortunate trend in which “the division of intellectual labor in our discipline tends to replicate program divisions. Experts on aging study

Social Security; experts on health care study Medicare; experts on poverty study Aid to Families with Dependent Children” (Quadagno 1999, p. 8). More generally, the American health care system is itself increasingly governed by business principles of cost control and administrative efficiency, under corporate managed care. The dramatic growth of for-profit health maintenance organizations (HMOs), “now accounting for 75 percent of all HMOs and enrolling over 50 percent of all subscribers” (Fein 1998, p. 10), has intensified public debate over quality, access, and humanity in health care.

The products of health policy analysis range from journalistic and descriptive accounts to sophisticated quantitative analyses and projections. But over the last decade or more, health policy analysis has reflected a societal struggle to come to terms with a secular change in the organization and financing of medical care, away from solo, fee-for-service practice toward corporate managed care. Given the current emphasis on cost savings and efficiencies, and on mechanisms for achieving them such as capitation, risk adjustment, and “utilization management” or “practice guides” for physicians’ clinical discretion, economic models and analyses have been paramount in health policy analyses funded by government agencies and large corporate entities. The justification advanced for these competitive efforts has typically been a need to check inflationary costs and “excessive” demands for medical services by consumers, ostensibly free to operate in a “market” for such services.

Consequently, traditional foci of sociological interest—including professional status and autonomy, access to and stratification of health care services, and continuity Federal safety-net policies rooted in the postwar *social contract* (Quadagno 1999; Rubin 1996)—have been pushed to the margins of public and policy debates. However, sociological perspectives are both rejuvenated and needed at this time. One important line of critique has been to challenge attributions of market choice and consumer autonomy in the face of corporate managed care (Freidson 1994; Freund and McGuire 1999). Another is to reject the very notion of “system” in relation to health care and medical coverage in the United States and instead to document, as does Diamond (1995), the collective vulnerability and implications arising from the arbitrary and confusing patchwork quilt that is American

health policy—a paradox of “excess and deprivation” (Bodenheimer and Grumbach 1998). Yet another is to demonstrate how health care professions may act to mediate between users of health services and their often remote provider organizations. In an important analysis of how doctors are implicated in this process, Freidson (1994) argues for a rebirth of *professionalism*, based on client service and trust, as bases for health care reform. Although one may question the likelihood of this scenario, given the mistrust of doctors at the heart of the consumer backlash of recent decades, the answers are sure to be significant both for social theory and policy. These are but a few examples of the distinctive contributions sociologists are making to health policy analysis, broadly defined.

Consistent with this public spotlight, sociological research on health policy and other segments of the welfare state is gaining momentum. The 1998 president of the American Sociological Association devoted her address to a historical analysis of changes in welfare policy provision, including social security and medicare, as central to understanding the erosion of the postwar social contract in the United States (Quadagno 1999). Furthermore, the demographic aging of America, along with a dramatic increase since mid-century in women’s labor-force participation (with resultant strains in traditional sources of familial care), are propelling the neglected problems of chronic illness, community-based care, and allocation of resources—that is, between capital-intensive hospital treatments and more equitable provision of basic health care—to the forefront of the national and research agendas.

In recent years, then, the inventory and scope of topics subsumed under the heading *health policy analysis* have expanded in ways that energize and demand the attention of sociologists. The legislative failure of the Clinton administration’s national health plan demonstrated the necessity for a coherent set of principles—moral and political, as well as technocratic—in order to implement large-scale policy reform; resistance to “environmental racism” by those unduly exposed to hazardous jobs and industrial toxins has assumed global dimensions; and such widely publicized conflicts as those over public versus corporate liability for the expense of tobacco-related illness (Glantz et al.

1996) and firearms and other forms of violence (Prothrow-Stith 1998; Sinauer et al. 1996)—all these have underscored the political, economic, and cultural forces that shape the health problems, as well as the spectrum of policy options, that analysts address. Indeed, health policy analysts have periodically been buffeted directly by political currents. During the Reagan administration, conservative forces in Congress sought to curtail sharply the collection of health-related data at the federal level; and spokespersons for the failed Clinton plan were attacked as proponents of a federal “takeover” of health care. This attack reflected and accelerated the devolution of federal discretion and responsibility for health care and other policies to state and local governments. Thus, health policy analysis, like health policy itself, has become increasingly politicized.

Research in health policy analysis necessarily concerns itself most directly with timeliness, pragmatism, and specificity in an effort to improve health and health care delivery. Research and analysis are conceived to inform social policy by (1) illuminating features of social organization and social action that are relevant to health policy planning, (2) identifying the social and health problems that require formulation in attempts to develop health policy, and (3) organizing and interpreting data that monitor the effects and outcomes of health policy decisions and the relative impact of programmatic alternatives.

In response to this mission, health services research contributes two major types of knowledge: *engineering* and *enlightenment* knowledge (Weiss 1978). In turn, these models imply distinctly different roles for analysts in the policy process (Marris 1990). In the *engineering* model, researchers seek to provide instrumental knowledge for practical assessment of alternatives and problem solving, accepting the values and goals inherent in existing policies largely as givens. Many influential health policy analyses first appear as fugitive documents directed to internal governmental audiences, addressing particularistic needs and interests of government agencies and actors, and are based on reports designed with an evaluative purpose. Policy analysis of this kind is, again, primarily funded and supported by government, with a lesser role played by such private foundations as the Robert Wood Johnson Foundation.

In the *enlightenment* model, researchers critically—even irreverently—scrutinize the implicit empirical, moral, and political assumptions embedded either in discrete policies or in broader debates (e.g., about the “right to die” or national health insurance). Rather than dealing with how policies work in a technical or engineering sense, enlightenment research contributes to the root understanding of how, by whom, and with what unintended consequences problems in health policy are socially constructed. Often, enlightenment research promotes shifts in what Thomas Kuhn (1970) calls “paradigms,” that is, fundamental ways of looking at problems. The enlightenment model is rooted as well in a critical, Weberian tradition in which the *formal rationality* of internal program functioning is juxtaposed with the *substantive rationality* of such programs, as they affect individual freedom and social equity.

As Marris (1990) shows, the engineering model is most effective and appropriate when policies have clear goals, enjoy broad consensual support, and can be linked directly to social outcomes. At the macro level especially, such conditions have rarely obtained regarding health policy in the United States. Moreover, experienced observers have concluded that however well conceived and conducted, research has had a limited direct role on the adoption and implementation of health policy (e.g., Lee 1998; Mechanic 1974). Important, though less often discussed, is that analysts in the engineering model are dependent on access to reliable, comprehensive, and timely data sets. Such a research infrastructure is difficult to develop and maintain, even where data collection is mandated at state or federal levels of government (Mechanic 1974). Given the present trend of privatization in the management and delivery of health services, sources, collection, and linkage of data are correspondingly more varied and less subject to public oversight. For example, while public health departments have a responsibility to serve the population at large, HMOs, however carefully they document utilization of services among their thousands of subscribers, have no such obligation to the public. This poses serious questions regarding the coordination of public and private health entities (Goldberg 1998).

Among other fertile research questions being posed in the expanding, multidisciplinary field of

health policy analysis are the following: How is the global resurgence in infectious disease—termed the *third epidemiologic transition* (Barret et al. 1998)—linked to our more global economy and consumer culture, and what strains is it likely to impose on outdated public health networks? To what extent is globalization leading to convergence in the organization of health care systems internationally (Mechanic 1996)? How are the successes of the American health care system in increasing human longevity creating new conceptions of and practices in medical ethics? Inasmuch as chronic illnesses are often peripheral to direct treatment by doctors, what roles are nurses and other medical practitioners playing in the revision of medical ethics and practice (Thomasma 1994)? What is the place and role of communities in our increasingly corporate health care system? And how might we rethink research practices to better conceptualize and tap community-level perspectives and dynamics (Israel et al. 1998)? Many contemporary problems in health care—from mechanisms by which AIDS and other diseases are transmitted, to discrimination against minority groups seeking care—would seem to rest on understanding community-level dynamics.

Sociology has a long tradition of reformism and interest in finding solutions to applied problems. Robert Lynd’s *Knowledge for What?* (1986) called sociologists to the task, and a long line of American sociologists have worked within the applied tradition. Particular examples are from the Chicago School (Bulmer 1984; Deegan 1988; Deegan and Burger 1981; Park 1952) and Columbia University, where Lazarsfeld and his colleagues advanced the field of applied research after World War II. These efforts were followed by work on the uses of sociology (Lazarsfeld et al. 1967) and a burgeoning of critical scholarship in the wake of the “counterculture” of the 1960s. These forerunners laid the foundation for what has become an increasingly exciting enterprise: the study of health policy. Freeman’s (1978) observation on the nature of health policy analysis as a scientific enterprise remains applicable: that policy studies are rather specialized and “content limited,” demonstrating few attempts to develop overriding conclusions about the policy process; hence, “there is practically no effort at ‘grand theory’ and little at ‘middle-range theory’ either” (Freeman 1978, p.

11). Nevertheless, narrow, highly specialized studies are not policy studies if they have no use beyond the most limited and specialized areas of concern. "Policy studies . . . need to be broad in implications, insightful to those beyond the narrow band of experts in a particular field, and intermeshed with work in related areas" (Freeman 1978, p. 12). The stimulation of and funding for policy analyses has been driven largely by the immediacy of existing (rather than emerging) problems that catch the attention of policy makers. Therefore, there is tension between the need to conduct carefully controlled definitive studies and the need to enlarge the focus of such research to contribute broader application and significance.

The growth of health policy analysis was shaped by the social problem definitions of health care from the 1960s to the 1980s (Rist 1985), and these, in turn, have been shaped by the political and economic exigencies of these periods. Health care was defined in the 1960s by the crisis of access, in the 1970s by the crisis of fragmentation and lack of comprehensive planning, and in the 1980s by the crisis of cost and the resurgence of market forces in health care. The widening reverberations of these forces throughout the 1990s presents sociology with an urgent and relevant research agenda. The cost of medical care continues to rise at two or three times the rate of inflation; the costs to business, government, and individuals skyrocket; more and more Americans are uninsured each year; the annual expenditure on the medical-industrial complex climbs above \$600 billion; and the population is aging. In the wake of these dramatic developments, the health care system and the policies creating it have been increasingly exposed to criticism and investigation. The key health policy issues for the new century are the cost, quality, and outcomes of care; the organization, financing, and delivery of acute and long-term care services; and expanding access to care.

(SEE ALSO: *Health-Care Utilization and Expenditures*; *Health Promotion and Health Status*; *Medical-Industrial Complex*; *Medical Sociology*)

## REFERENCES

Baer, Hans A. 1995 "Medical Pluralism in the United States: A Review." *Medical Anthropology Quarterly* 9(4):493-502.

- Barrett, Ronald, et al. 1998 "Emerging and Re-Emerging Infectious Diseases: The Third Epidemiologic Transition." *Annual Review of Anthropology* 27:247-271.
- Benner, Patricia, ed. 1994 *Interpretive Phenomenology: Embodiment, Caring, and Ethics in Health and Illness*. Thousand Oaks, Calif.: Sage.
- Bodenheimer, Thomas, and Kevin Grumbach 1998 *Understanding Health Policy: A Clinical Approach*, 2nd ed. Stamford, Conn.: Appleton and Lange.
- Brown, Phil 1995 "Naming and Framing: The Social Construction of Diagnosis and Illness." *Journal of Health and Social Behavior* 33:267-281.
- Bulmer, Martin 1984 *The Chicago School of Sociology: Institutionalization, Diversity and the Rise of Sociological Research*. Chicago: University of Chicago Press.
- Cockerham, William C. 1988 "Medical Sociology." In Neil J. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- Deegan, Mary Jo 1988 *Jane Addams and the Men of the Chicago School, 1892-1918*. New Brunswick, N.J.: Transaction Books.
- , and John S. Burger 1981 "W. I. Thomas and Social Reform: His Work and Writings." *Journal of the History of Behavioral Sciences* 17:114-125.
- Diamond, Timothy 1995 "Breaking the Dichotomy Between Theory and Research: The Method of Institutional Ethnography." Paper presented at the annual meeting of the American Sociological Association, Washington, D.C.
- Fein, Rashi 1998 "The HMO Revolution" (Part 1). *Dissent*, 45 (2):29-36.
- Freeman, Howard E., ed. 1978 *Policy Studies Review Annual*, vol. 2. Beverly Hills, Calif.: Sage.
- Freidson, Eliot 1994 *Professionalism Reborn*. Chicago: University of Chicago Press.
- Freund, Peter E. S., and Meredith B. McGuire 1999 *Health, Illness, and the Social Body: A Critical Sociology*, 3rd ed. Upper Saddle River, N.J.: Prentice-Hall.
- Glantz, Stanton A., et al. 1996 *The Cigarette Papers*. Berkeley: University of California Press.
- Goldberg, Bruce W. 1998 "Managed Care and Public Health Departments: Who Is Responsible for the Health of the Population?" *Annual Review of Public Health* 19:527-537.
- Israel, Barbara A., et al. 1998 "Review of Community-Based Research: Assessing Partnership Approaches to Improve Public Health." *Annual Review of Public Health* 19:173-202.

- Kleinman, Arthur 1980 *Patients and Healers in the Context of Culture*. Berkeley: University of California Press.
- Kuhn, Thomas 1970 *Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- Lazarsfeld, Paul F., William H. Sewell, and Howard L. Wilensky 1967 *The Uses of Sociology*. New York: Basic Books.
- Lee, Philip R. 1998 "Health Policy: A Fifty-Year Perspective." Unpublished manuscript, Institute for Health Policy Studies, University of California San Francisco.
- Luft, Harold S. 1999 (Director, Institute for Health Policy Studies, University of California San Francisco) Personal communication.
- Lynd, Robert S. 1986 *Knowledge for What The Place of Social Science in American Culture*. Middletown, Conn.: Wesleyan University Press.
- Marris, Peter 1990 "Witnesses, Engineers, or Storytellers? Roles for Sociologists in Social Policy." In Herbert J. Gans, ed., *Sociology in America*. Newbury Park, Calif.: Sage.
- Mechanic, David 1974 *Politics, Medicine, and Social Science*. New York: John Wiley & Sons.
- 1996 "Comparative Medical Systems." *Annual Review of Sociology* 22:239–270.
- Nagel, Stuart S., ed. 1977 *Policy Studies Review Annual*, vol. 1. Beverly Hills, Calif.: Sage.
- Park, Robert Ezra 1952 *Human Communities*. Glencoe, Ill.: Free Press.
- Prothrow-Stith, Deborah 1998 *Peace by Piece: A Violence Prevention Guide for Communities*. Boston: Harvard School of Public Health.
- Quadagno, Jill 1999 "Creating a Capital Investment Welfare State." *American Sociological Review* 64(1):1–11.
- Rist, Ray C. 1985 "Introduction." In Ray C. Rist, ed., *Policy Studies Review Annual*, vol. 7. New Brunswick, N.J.: Transaction Books.
- Rubin, Beth A. 1996 *Shifts in the Social Contract*. Thousand Oaks, Calif.: Pine Forge.
- Sinauer, N., J. L. Annet, and J. A. Mercy 1996 "Unintentional, Nonfatal Firearm-Related Injuries: A Preventable Public Health Burden." *Journal of the American Medical Association* 275(22):1740–1743.
- Spector, Malcolm, and John I. Kitsuse 1987 *Constructing Social Problems*. New York: Aldine de Gruyter.
- Thomasma, David C. 1994 "Toward a New Medical Ethics: Implications for Ethics in Nursing." In Patricia Benner, ed., *Interpretive Phenomenology*. Thousand Oaks, Calif.: Sage.

- Weiss, Carol H. 1978 "Research for Policy's Sake: The Enlightenment Function of Social Research." In Howard Freeman, ed., *Policy Studies Review Annual*, vol. 2. Beverly Hills, Calif.: Sage.

CARROLL L. ESTES  
CHRISTOPHER WELLIN

## HEALTH PROMOTION AND HEALTH STATUS

*Health promotion*, a general term, refers to a wide range of health-enhancing activities that seek to maintain health and functional ability, increase longevity, and reduce the prevalence and consequences of disease. These diverse activities include distributing free needles to substance abusers, identifying and modifying genes such as those linked to the development of Alzheimer's disease, proposing laws that seek to deter cigarette smoking, and offering blood pressure and cholesterol screening. They also include personal health practices whereby individuals engage in healthy lifestyles, consume vitamins, and the like. Therefore, these wide-ranging activities are conducted by persons trained in the field of public health, traditional health care workers, practitioners of complementary medicine, basic science researchers, politicians, health policy experts, and the individuals who practice health-enhancing behavior, to name but a few.

Long ignored in favor of medical, surgical, and pharmaceutical treatments that seek to cure or arrest health problems, health promotion information is now a prominent feature of popular magazines and nightly newscasts. Indicative of its rising importance, the United States Public Health Service has expanded the official title of the Centers for Disease Control to the Centers for Disease Control and Prevention.

The significance of health promotion activity is well known to the medical community. It offers the best, if not cheapest, method of reducing the burden of life-threatening conditions and maintaining a healthy, well-functioning, long-lived population (Pope and Tarlov 1991). As has been stated, "many of the most serious disorders. . . can be prevented or postponed by immunizations,

chemoprophylaxis, and health life-styles. To an unprecedented extent, clinicians now have the opportunities, skills, and resources to prevent disease and promote health, as well as cure disease” (Office of Disease Prevention and Health Promotion 1994). Thus, health promotion is the ounce of prevention to avoid the pound of disease/disability/decline.

### THE RANGE OF HEALTH PROMOTION ACTIVITIES

Activities that constitute health promotion are conducted on macro and micro levels. They address many spheres of society including: (1) the physical environment, (2) political/economic institutions, (3) health and medical care systems, (4) the social environment, (5) the fields of human biology, molecular medicine, and genetics, and (6) human behavior.

Macro-level efforts range from the international arena (e.g., drafting treaties about water safety), to the national level (e.g., policies about smoking and air pollution standards), and further to the local level (e.g., offering services of the city or county public health agency). They also include the private sector. Some of these activities address the training of medical personnel, the financing and delivery of medical care at private and public sites, the provision of health screening services, the development of active surveillance systems about disease threats, the provision of immunization programs, the offering of health education programs, and the scientific discovery of pathways to disease.

Micro-level health promotion involves individuals and small units. A single person’s actions to maintain a healthy body, forestall the development of disease, achieve a longer life, or reduce emotional stress are micro-level health-promoting activities. Also included in this category are behaviors of the patient–health provider dyad. Such behavior often pertains to establishing a set of health-promoting practices and complying with this plan.

In a narrower approach to health promotion, disease control is the focus. This has resulted in distinguishing health promotion activities according to their potential for primary, secondary, and

tertiary prevention of disease. The terms are used in relationship to the stage of a disease or condition; thus primary prevention includes the pre-disease phase and applies to the period prior to onset or diagnosis of a disease. In contrast, secondary and tertiary prevention occur after a disease has manifested itself. Efforts to detect the disease and keep it controlled, contained, and manageable are included (secondary prevention) as well as medical treatment activity addressed to preventing disability, improving life quality, and delaying death during the course of well-established disease.

Primary prevention can begin at any time during the life course. Directed toward individuals without disease manifestations, it seeks to help them maintain well-functioning, disease-free bodies. One of the oldest primary prevention practices is handwashing to prevent spread of communicable disease. One of the newest is testing women for the BRCA1 gene for breast cancer. If it is detected in an individual, she may be prescribed a drug that has been observed to have cancer-prevention potential. This type of primary prevention activity reflects cutting-edge medical technology from the fields of molecular medicine and genetic therapy. It is growing exponentially as a health promotion strategy. Indeed, \$25 million of macro-level prevention efforts were recently announced to prevent or curb the spread of infectious diseases by using state-of-the-art technology (*New York Times*, December 27, 1998).

Avoiding exposure to agents that cause disease, injury, or defects is clearly the focus of much of the primary prevention activity. Public health agencies achieve this by offering immunization programs to children and adults alike. Health policy and health law personnel propose laws to prohibit cigarette advertisements. Physicians offer general and condition-specific health education, such as teaching patients about the value of weight control. Laboratory personnel seek to develop new vaccines or analyze the genetic characteristics of bacteria linked to certain infectious diseases.

Personal behavior is often the target of primary prevention efforts. Individuals with a family history of heart disease may be advised about dietary change from fatty foods or the need to engage in exercise, stress reduction, smoking cessation, and/or weight reduction programs. Given

recent data, these activities appear to be quite important in preventing the development of a heart attack (myocardial infarction) and blockage of the coronary arteries. Even reducing one of the modifiable risk factors for heart disease (smoking, high blood pressure, obesity, physical deconditioning, high lipid count, impaired glucose tolerance) significantly reduces the risk of developing this disease (Herd et al. 1987; Kannel et al. 1987). Since more than half of the men and women over age 65 currently die of heart disease, such activity could have profound effects. Indeed, the development of a major heart condition may take fifty years, beginning in childhood with unhealthy food and exercise habits (Fries and Crapo 1986). Therefore, the individual is responsible for this type of primary prevention, which can actually begin with parental efforts to offer low-fat diets to their children. The idea is that the person would then carry on healthy living behavior throughout the life cycle.

The importance of personal behavior in health promotion and disease prevention was established in a recent effort to reconceptualize the causative factors for death (mortality). Previously, mortality data listed specific diseases as primary or secondary causes of death. The new approach designates specific health-risk behaviors (e.g., smoking and exercise behavior, food habits, alcohol use) as actual causes of death (McGinnis and Foege 1993). These behaviors accounted for half of all U.S. deaths in 1990. The three leading risk behaviors were tobacco use (400,000 deaths) and diet and activity patterns (300,000 deaths). By now considering this health-risk behavior to be the cause of death, the prevailing disease-centered approach is diminishing in importance. Reducing the prevalence of these behaviors is becoming a goal of interventive efforts and can facilitate the shift to primary prevention efforts.

Secondary prevention is undertaken after a disease has been detected. This may occur a half century after the causal agent or offending behavior initiated the disease process (Fries and Crapo 1986). For example, a heart attack at age 70 may have as its etiology poor dietary habits beginning at age 15, smoking behavior that started at 20, and at ages 30, 40, and 45, respectively, physical deconditioning, job stress, and the development of hypertension.

In contrast to primary prevention, secondary prevention involves control of a disease or condition. Aims are to ensure early detection, follow this with prompt and effective treatment, and educate the individual about risk-reduction behavior. Such efforts seek to halt, slow, or possibly reverse the progression of a condition and to prevent secondary effects or complications. Therefore, the heart patient would not only receive medication that might improve functional capacity of the heart but also receive referral to a smoking-cessation program.

Secondary prevention efforts are addressed to all acute and chronic conditions, but five chronic diseases account for most of the deaths, hospital care, and disability of the U.S. population. In order of prevalence they are arthritis, high blood pressure (hypertension), heart disease, chronic bronchitis, and diabetes. These chronic diseases are readily amenable to preventive action (Pope and Tarlov 1991). Physicians can treat these conditions medically, pharmaceutically, and surgically to prevent complications or arrest disease progression. The individuals suffering from these conditions can engage in a series of health behaviors and modify their lifestyles. For example, persons with diabetes can be asked to visit their doctor regularly, have frequent tests of blood and urine to detect disease progression, and visit specialists for control of diabetes complications involving eyes, kidneys, nervous system, skin, and so forth. Simultaneously, they can be taught to monitor and control their conditions, engage in dietary change, and actively pursue weight-control and exercise programs.

One of the most important tools for secondary prevention is health screening. This procedure seeks to detect disease, abnormal body states, and sensory loss. It comprises activities as diverse as mammography; Pap smears; and blood pressure, glaucoma, blood sugar, prostate specific antigen (PSA), and hearing tests. These secondary prevention services may be offered in physician's offices, at community sites (such as health fairs or churches on Sunday), or at work sites.

Secondary prevention procedures are immensely valuable, especially if they allow a disease to be detected in its early stages. Through early detection, medical and personal care can begin before the disease has progressed. Complications may be

avoided, dangerous clinical thresholds may be averted, and the downward trajectory to disability and death may be prevented. For example, mammography that detects a small breast tumor may be responsible for saving a woman's life, preventing disfigurement from mastectomy (seldom performed in the early stages of cancer), averting the need for chemotherapy and its severe side effects, and enabling the patient to have hope for complete remission. In contrast, if a tumor is detected by clinical breast examination by a physician or breast self-examination, it will be larger than one identified through mammography. A tiny tumor or precancerous breast tissue change will not be apparent, through clinical breast examination and the cancer detected may be larger or have metastasized.

Physicians and nurses are the predominant practitioners of secondary prevention. Physician assistants, pharmacists, nutritionists, health educators, and physical therapists are among others who form the secondary prevention medical team. Unfortunately, there is much less use of this health promotion strategy than is optimal, or even desirable. Studies have shown that fewer than half of all physicians indicated they schedule proctoscopic examinations and chest x-rays for asymptomatic patients with no personal history of cancer (American Cancer Society 1990). This study also showed that specialists in internal medicine offer their patients the most cancer screening relative to other medical specialties (Pap tests; mammograms; stool blood determinations; chest x-rays; and breast, digital rectal, and proctoscopic examination).

Tertiary prevention is the last type of disease-focused prevention. It seeks to provide good health care to persons with diseases that have progressed beyond their initial stages. Therefore, since it is too late to prevent illness or arrest its progression to a more serious phase, tertiary prevention includes medical and surgical interventions that can prevent functional decline, improve life quality, or delay death.

The U.S. medical care system has focused on tertiary prevention and uses an increasingly large medical and surgical armamentarium. Breakthrough technology is at the basis of these efforts. It features replacement of organs with mechanical, animal, or other human parts; using genetically engineered products to alter disease agents; and cloning

or otherwise duplicating disease-free cells to replace unhealthy tissue. These efforts are unprecedented and unique to the U.S. health care system. No other nation offers a comparable level of advanced medical care. Unfortunately, these tertiary prevention services have made our health care system the most costly in the world on a per capita basis but have not improved our life expectancy to the level of most of the industrialized world. In 1994 life expectancy at birth in the United States was 75.9 years, compared with 80.1 years in Hong Kong, 79.3 in Japan, 78.2 in France, and 78.1 in our neighbor, Canada (U.S. Bureau of Census 1994). Men, in particular, have failed to experience major gains in life expectancy. Women of all racial groups have outlived men since 1900, and white females currently have a seven-year advantage in length of life (U.S. Bureaus of Census 1994).

The fact that the gender gap in survival is occurring simultaneously with a major decline in mortality from conditions such as heart disease may speak to the limitations of tertiary prevention. Some attribute lower mortality from heart disease to procedures such as replacing valves and arteries, heart transplantation, removing blockages in coronary arteries and introducing stents, or offering pharmacological treatment in the form of clot-reducing drugs. Yet others suggest that medical measures have been less effective than those that involve change in health-risk behavior. It is proposed that the reduction in heart disease mortality has been achieved mainly through primary and secondary prevention practices. Since males are less likely to have good dietary practices, and more likely to ingest foods that contribute to artery blockages, use tobacco and have poor exercise habits, this may be affecting their mortality rates. Men may not be living as long as women because they fail to engage in health-promoting behavior, relying instead on medical treatment for an existing condition.

Tertiary prevention is clearly important. The vast majority of individuals suffering from cancer, severely debilitating disease, or life-threatening heart problems seek, or await the development of, advanced medical care. Medical journal articles about life-saving, and sometimes life-enhancing, treatments, are reported regularly on evening news programs. Only a small number of people refuse



to have organ transplants or kidney dialysis, and many wait for new AIDS or cancer drugs to be offered for clinical trials. Canadian citizens routinely cross the U.S. border to avail themselves of medical, surgical, and technological procedures that are unavailable in their country or subject to waiting lists.

As stated in a classic medical sociology article, medical care in this country has not been responsible for a decline in mortality rates; rather, the primary causes of the reduction are public health measures that lessen the risk of acquiring disease (McKinlay and McKinlay 1977). We need to refocus our efforts toward heavier emphasis on primary and secondary prevention, and reduce our reliance on costly and technology-heavy tertiary interventions.

### **THE ROLE OF THE INDIVIDUAL IN HEALTH PROMOTION**

As presented above, the responsibility for health promotion rests with several agents; the national and international health care systems, health law and health policy experts, public and private health facilities, health professionals, and individuals. However, the individual is often the center of attention as the burden of health promotion is shifted to the person.

This approach rests on the premise that a person's actions account for whether he or she remains healthy and does not experience progression of or complications from disease. The corollary to this is that persons who fail to exert control over health behaviors are contributing to disease, disability, and death. Movements to empower people and help them take charge of their medical conditions reflect this emphasis on the person's responsibility for health.

Indicative of this trend is that patient noncompliance with the medical regimen is considered to be a cause of medical treatment failure. Clearly, many people fail to adhere to the treatment plan prescribed by their physician. They may fail to have their prescriptions filled, take the amount of medication prescribed, keep medical follow-up appointments, or adhere to a dietary program necessary for a controlling a condition. Only 57

percent of persons with elevated cholesterol levels actually visited a physician after being notified of their state. Gender is an important part of compliance. Women visit physicians more than men, especially for preventive care (Verbrugge 1990). They also act as gatekeepers to the medical system for children and husbands, offering suggestions and advice, and sometimes making medical appointments without the consent of the latter.

This line of reasoning—patients are largely responsible for poor health outcomes because they do not follow the doctor's advice—ignores the fact that health professionals may create situations that foster noncompliance. Presenting an ultimatum about smoking cessation to heart patients, informing people with diabetes they must avoid sugar for the rest of their lives, or telling people diagnosed with hypertension they must fill a costly prescription for a condition that does not cause them physical distress seldom results in compliance. People do not like to be given lists of do's and don'ts. If reasons for the recommended procedures, medications, or behavioral changes are not given, compliance may fail to materialize. Asymmetric models of the patient-physician encounter (Bloom 1974) pose physician authority on the one hand and childlike response on the other. These models have lost their impact, especially among educated middle-class patients, and compliant behavior as an automatic response by the patient is becoming much less widespread.

Another reason that health promotion should not be considered the sole responsibility of the individual is that knowledge is the foundation for action. However, individuals may have little knowledge of appropriate health-promoting behavior. They are consumers of health information. As such, they must either depend on others for information or take the initiative for self study. While knowledgeable, enlightened patients can certainly help to control and/or contain their medical problems, only medically trained persons familiar with that patient's health problems can offer appropriate recommendations! To illustrate, the heart patient's physician is responsible for educating the patient about specific exercises to reduce the risk of a second heart attack because it is the physician who is intricately familiar with the functional and/or anatomical state of that person's heart. Physicians

must educate their patients and, in doing so, carefully explain the regimen. It is only when patients receive adequate information and careful explanation about a particular regimen that they can be held accountable.

### **BARRIERS TO HEALTH PROMOTION**

Reducing barriers to health service use helps considerably to promote health and reduce disease prevalence (Orlandi 1987). Analysis of the major barriers to health service utilization shows that they are both societal and individual. Societal impediments are known as structural barriers and refer to variables that originate in the economic, political, and medical organizational spheres of society. Individual, or personal, barriers refer to behavioral variables that seek to avoid, delay, or underutilize health care.

Structural barriers limit access to health promotion programs, as they do to medical service utilization. They can act at several levels. At the economic and health insurance level, absence of such resources make it unlikely that patients will receive some recommended screening or early detection tests. Even for those persons with health insurance coverage, policies may not cover the recommended tests. Constraints on health promotion services are also due to policies or factors that reflect how health care services are organized. To illustrate, political decisions that reduce the availability of public transportation in turn reduce access to sites offering health promotion services. Another example would be health policies that encourage medical sites with state-of-the-art equipment to proliferate in suburban areas. If the newest mammography equipment is available only to women in affluent areas, while inner-city sites have machines that may be old and of poor quality, clearly mammography screening will be less likely in these latter areas.

Personal barriers to health promotion activity include several factors that relate to an individual's perceptions, beliefs about cause and cure, and/or attitudes toward use of formal health care services. These barriers influence whether people will seek health care, follow the advice of medical professionals, and comply with a health promotion regimen and are reflected in the health belief model

(Rosenstock 1974). This model of health services utilization proposes that people seek medical services according to: (1) perception of the threat posed by a health problem, and beliefs about their susceptibility to it (e.g., "heart attacks are serious but no one in my family died of heart disease"); (2) the possible inconvenience of the health-related activity versus potential rewards ("I can't manage without a morning cigarette, and besides my heart can't be in such bad shape since I only smoke half a pack a day"); and (3) response cues ("Did you have your annual prostate cancer checkup?").

Cultural factors affect the likelihood that persons will engage in recommended health promotion activity. Beliefs and practices handed down through the generations prescribe alternative health behaviors and nutrition habits. Some are harmless, others may be beneficial (chicken soup for colds, for example), and still others may exacerbate illness or even cause death (refusing blood transfusions on religious grounds). Beliefs also affect primary prevention; for example, some cultural traditions associate obesity with beauty or strength, not risk for chronic disease.

Denial is a particularly important personal barrier that can be added to this model, since denial of illness is directly related to avoidance of medical care. Many women fail to have a mammogram or delay the procedure because they seek to deny the possible threat of cancer. Some of these women have intense fear of the disease, others believe it is incurable, still others fear the surgical or chemotherapeutic treatment involved (Young 1998). In either case, they avoid thinking about it and feel they are better off not knowing whether they have a breast tumor. Delay in scheduling mammograms is a current interest of people seeking to improve rates of mammography compliance (Rimer et al. 1996).

Belief in one's own ability to control one's life also relates to health promotion. Health locus of control measures indicate whether individuals are internally controlled or whether they believe control is due to chance or the actions of powerful others (Wallston et al. 1978). Thus, individuals with internal health locus of control are the best candidates for health promotion programs because they believe in exerting control over their health status. They represent, on a conceptual level, the action stage of Prochaska and DiClemente's

(1984) transtheoretical model of health behavior. Using this same framework, persons who are not engaging in risk-reduction behavior might be considered to be in precontemplation or contemplation stages and may never move to action.

### HEALTH PROMOTION FOR OLDER PEOPLE

Health promotion and disease prevention programs generally target working-age people rather than older adults (Young 1994). This is quite unfortunate, since approximately 85 percent of people aged 65 and over suffer from chronic disease and the three most prevalent conditions are arthritis, high blood pressure, and heart disease (National Center for Health Statistics 1996). When considered simultaneously with the three leading causes of death in this population (heart disease, cancer, and stroke), the need for preventive care is quite apparent. These five conditions are largely amenable to risk-factor reduction practices. Indeed, readily available and well-known health promotion and disease prevention practices can alter the course of most, if not all, chronic and killer diseases of older people. There is great potential for improving the health of older adults by including them in health promotion programs (Pope and Tarlov 1991).

Health promotion efforts for older people should certainly involve risk-reduction behavior. As they are taught about behavioral and lifestyle changes, these efforts can result in a sharp drop in the medical consequences of chronic disease. Heart disease provides an important example. Rates of coronary heart disease are at least ten times as high in persons aged 65 and over as among their counterparts under age 45. Women as well as men show a dramatic rise in heart disease rates after they reach age 65. It has clearly been shown that much heart disease can be prevented by diet, exercise, smoking cessation, and similar healthy living practices. Furthermore, some of these same practices can reduce the likelihood of death, development of secondary complications of heart disease, and occurrence of a second heart attack (Kannel et al. 1987). Such behavioral change also brings psychological and social benefits to heart patients and results in gains in several areas of their lives. Yet

older people recovering from heart attacks have been found to be significantly less likely to receive preventive behavior advice than their younger counterparts (Young et al. 1987). Particularly absent for persons aged 60 and over is advice to cease smoking and enroll in a cardiac rehabilitation exercise program. The question then emerges, "How can older people engage in health-promoting behavior that may prevent another heart attack, if not given the proper information from their physicians?"

Another way health promotion can be achieved is through following recommendations of major authorities for periodic administration of specific tests, examinations, and immunizations. The guidelines for adult preventive care offered by the U.S. Preventive Service Task Force (1989) include specific age-related tests. The set of recommendations for people aged 65 and over includes annual influenza shots; breast, thyroid, mouth, skin, ovarian, testicular, lymph node, rectal, and prostate cancer examinations; and dental examinations. Tests to determine blood pressure and visual ability should be conducted every two years. There are also recommendations for periodic tests of urine, hearing, estrogen levels in women, and cholesterol. Of course all of these recommendations apply to persons without specific health problems (those who are asymptomatic and of normal risk). Cholesterol and blood pressure tests for older people with heart disease would need to be performed frequently.

Thus, health promotion for older people can be achieved if (1) physicians conduct specific examinations and tests at defined periods of time and (2) if older people engage in risk-reduction behavior. For either to be successful, the older person must follow a recommended plan for medical visits and for individual behavior. This plan must be carefully explained to the older patient. However, studies show that patient-physician interaction in later life is poor (Coe 1987), and physicians may communicate poorly with their older patients and fail to offer a recommended illness-management plan (Young et al. 1987). Clearly, if older people are to change poor dietary habits, they must first be informed about good food choices. Since this information is condition-specific, it needs to be provided by the physicians

who treat the individuals or their medical agents, such as nurses or physician assistants.

Older individuals also bear responsibility for behavioral change. They must be willing to engage in health behaviors that prevent illness or modify high-risk profiles. Even when these behaviors represent change in lifelong habits, they must be willing to pursue them. However, for effective risk-reduction practices to be targeted to older people, patient-physician partnerships must be formed (Hess 1991). Including the older patient in the health promotion program is essential.

### HEALTH PROMOTION FOR MINORITIES

Significant advances in understanding, managing, and treating chronic disease have failed to eliminate the excess death and disability found in older minorities. African Americans, for example, have a five-year deficit in life expectancy (U.S. Bureau of the Census 1994). They have higher mortality rates for cancer and heart disease (Polednak 1989) and twice the risk of severe complications from diabetes, such as blindness, neurological decline, and illnesses that require dialysis or amputation (Lieberman 1988). Other ethnic minorities such as Hispanic Americans, Native Americans, and some Asian Americans also have a major gap in mortality and morbidity, compared with whites; death rates for heart disease and stroke, two leading causes of death, are twice as high for older minorities as for whites (Polednak 1989). Indeed, the health gap of minorities in the United States is wide and shows no signs of being bridged.

This bleak prognosis is certainly a call for massive efforts at treatment, on the one hand, and extensive health promotion efforts, on the other. Yet despite the fact that health needs vary across populations, health promotion efforts have usually been generic (Young and Olson 1993). Often they focus on the needs of white, middle-class adults (Gottlieb and Green 1987).

Programs to target ethnic minorities must be culturally sensitive; they must also be specific to a particular population (Tseng and Ellyne 1990). One-size-fits-all programs will not fit ethnic minorities as a whole, or individual ethnic groups. The cultural background of the particular target

population must be taken into account, along with educational level, health beliefs, and attitudes toward health providers (Young and Olson 1993).

The need to modify generic health promotion program for minority groups is increasingly recognized. Without targeted programs, it is doubtful that health promotion efforts can achieve high success. Knowledge about and appreciation of specific cultural norms, values, and beliefs should be a major component of health promotion programs for minorities (Leavitt 1990).

An example of a health promotion program that targeted a particular ethnic minority was conducted among African Americans with coexistent diabetes and hypertension (Waller et al. 1994). Sixty subjects were randomly assigned to an educational/self-care intervention group or an observation group. The former were instructed by a health educator, using culturally relevant materials. Diet and food preparation information were specific to the food habits of lower-income African Americans and focused on culturally acceptable food and cooking changes. Materials prepared reflected the average educational level of the target population. After one year, high-risk behavior among the intervention group was significantly less than among the observational group. Risk for complications of diabetes, heart disease, and stroke all declined. Furthermore, the diseases were considered to be controlled.

The determinants of success of health promotion programs are clearly their effect on (1) health-risk behavior and (2) medical outcome (control over the condition, laboratory test values within acceptable ranges, etc.). Programs that can change behavior and also effect better outcome for ethnic minorities cannot be generic. They must be culturally sensitive and appropriate to the population they seek to serve. Otherwise, there will be little effect on the mortality and morbidity status of the minority group.

### HEALTH STATUS ASSESSMENT

Health promotion goals are often measured in terms of reduction in the prevalence of specific diseases or decline in death rates for a particular condition. The sciences of epidemiology and

biostatistics and the field of medical sociology propose several ways to determine health status. They use measurements and statistics for populations, subgroups, and individual cases.

Health status data are collected by epidemiologists, agencies, and medical professionals. Among multiple strategies used are disease surveillance, population surveys, and conducting probability and nonprobability studies of the health of subpopulations (Friedman 1988). Surveys that assess health-risk behavior (e.g., Behavioral Risk Factor Surveys from the Centers for Disease Control and Prevention) are well accepted means of collecting health status data.

Measurements often used to assess health status include rates, ratios, counts, proportions, distributions, ranges, and quantiles. These measurements may indicate prevalence of a particular disease (number of persons with the disease in proportion to the total number in the group); incidence (number of new cases in a given period); or number of persons of a particular age, gender, or ethnic group with the disease. They can also indicate ranking systems and represent change over time in disease or death rates.

Rates are often calculated as the number of cases per 1,000, 10,000, or 100,000 persons. They are also standardized to take factors such as age into consideration. Therefore, since age increases risk of lung cancer, epidemiologists can eliminate this potential bias by using appropriate statistical techniques. They may use age adjustment or age standardization procedures to calculate the expected rate of lung cancer (Friedman 1988). This enables them to assume nonsmokers are the same age composition as smokers.

After rate comparisons are made, the relative and attributable risks of a disease can be determined. In these calculations, biostatisticians evaluate prevalence or incidence of a disease in a particular population and may then compare risks of two different populations. Other analyses may be of change in risk over time, or the lifetime risk of acquiring a disease such as cancer.

A frequently encountered health indicator is the proportion of cases that fall into a particular quantile. This generic term refers to several tiers,

each representing a proportion of the population. To illustrate, people may be categorized into four quartiles, representing cholesterol levels. Each 25 percent segment of the population may then be compared. In these analyses, the association of an outcome variable, such as a second heart attack, with membership in a high-cholesterol-level group (quartile 1, for example) is clearly shown.

From a public health perspective, important indicators of health status are infant mortality and life expectancy rates. Low infant mortality and high life expectancy are considered to represent good health in a society. The premise is that societies that can reduce unnecessary maternal and child deaths and increase the average life expectancy at birth have established good medical care procedures and may have health practices that reduce the burden of illness and premature death.

In the United States, infant mortality rates have declined over the past century. In 1991 the overall infant death rate in the United States was 8.9/1,000 births (National Center for Health Statistics 1996). This represented a decline from 20/1,000 in 1970. Among certain subpopulations infant mortality was much higher than for the population at large. Blacks, for example, had infant death rates of 17.6, versus 7.3 for whites.

Life expectancy at birth is approximately 76 years in the United States. However, it is close to 80 years for white women and 73 years for white men. These figures represent a rise in life expectancy for men and women alike, and although there is a major racial gap, nonwhites have also shown an increase over the past 90 years.

In comparison with industrialized countries that offer all or most of their citizens publicly funded health care, U.S. men and women show up to a four-year gap in life expectancy (U.S. Bureau of Census 1994). U.S. infants are also more likely to die than those in several industrialized nations.

Sociologists use a different approach to assess health status than epidemiologists or biostatisticians. They measure functional ability, perceived health status, emotional and psychological health, and quality of life, to name a few. Medical measures are used infrequently, and statistics are seldom expressed in rates. Rather, sociologists assess health

status by tabulating scale scores of health-specific indices and/or calculating frequency distributions of individual health status questions. In the case of the latter, the investigator might determine presence/absence of a specific health problem (e.g., Do you have arthritis?) and measure central tendencies and variance in the data.

Functional ability is considered to an important indicator of health status in a sociological context. While the level of functioning can be measured as physical, mental, or social impairment, most of the interest is in physical functioning. This approach conceptualizes functional ability as ability to conduct activities of daily living (ADLs) or instrumental activities of daily living (IADLs) without assistance from others. Instruments used include measures suitable for older people (e.g., the OARS instrument of Duke University 1978), disease-specific measures (e.g., the arthritis functioning measure of Patrick and Deyo 1989), and instruments that assess cognitive functioning as a component of physical functioning (e.g., Keller et al. 1993).

The individual's perception of his or her health is often measured. In these investigations, the concern is with how the person assesses general health status at the present time or in reference to other people and other times. Thus, they may be asked "Is your health generally excellent, good, fair, poor, bad?" or "How does your health compare with that of people your own age?" or "Is your health better or worse than one year ago?" They may also be asked about perceived functional ability (Duke University 1978; Lawton et al. 1982).

Emotional and psychological health are also measured. Well-validated instruments such as the CES-D of the Center for Epidemiological Studies or the Zung Depression Measure are used for depression (DeForge and Sobal 1988). Investigators also seek to determine mood (Profile of Mood States of McNair et al. 1971), positive and negative affect (Bradburn 1969), morale (Lawton et al. 1982), and subjective well-being (Dupuy 1984).

Quality of life indicators are among the least well validated instruments used to assess health status. There are many different approaches to quality of life that represent medical, psychological, and social models of illness. Since sociologists

prefer a multidimensional view, Levine and Croog (1984) presented five components of quality of life: social-role performance, physiologic state, emotional state, intellectual function, and general satisfaction or feeling of well-being. Also considered to measure quality of life are some subscales of the Sickness Impact Profile (Bergner 1984). The full 134-item instrument assesses physical, social, psychological, and interactional aspects of illness, but some scales are specific to pain or impairment level and tend to reflect a medical, rather than psychosocial, view of quality of life. Indeed, the medical approach may concentrate exclusively on disease-specific or treatment-specific variables. It may measure, for example inability to eat or excessive fatigue among cancer patients undergoing chemotherapy, or frequency of urination among individuals with hypertension who are prescribed medications to expel fluids from the body. Still other medically focused approaches to quality of life may focus on the experience of pain, as in the previously mentioned Sickness Impact Profile.

Health status assessment is clearly a broad field. It includes measurement of disease patterns in a population, self-reports by individuals of generalized health status, and middle-level health measurements. These latter measurements are neither macro level, like population-based mortality and morbidity statistics, nor micro level, like individual reports. Rather, they include validated scales, indices, or series of questions that may be widely used among general or specific populations. Many, if not most, measures have been found to be reliable and valid indicators of health status. The medical sociologist or related researcher thus has a wide range of instruments to assess health status.

## CONCLUSION

Health promotion includes a wide-ranging set of activities that (1) enhance health status, (2) prevent disease, (3) seek to control the spread of chronic or infectious disease, and (4) attempt to arrest or delay deterioration that occurs as the result of these conditions. Health-promoting activities occur at the societal and individual levels and include a long list of agents. Essentially, health promotion represents the principle that maintaining health, preventing disease, and avoiding decline or complications of progressive illness are all

achievable goals. For any society to have a healthy, vital citizenry, it must reduce the financial, social, and medical burdens of illness. All these are accomplished with health-promoting practices.

## REFERENCES

- American Cancer Society 1990 "1989 Survey of Physicians' Attitudes and Practices in Early Cancer Detection." *Ca-A Cancer Journal for Clinicians* 40:77-101.
- Bergner, M. 1984 "The Sickness Impact Profile (SIP)." In N.K. Wenger, M. Mattson, C. Furberg, and J. Elinson, eds., *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*. New York: Le-Jacq.
- Bloom, S. W. 1974 *The Doctor and His Patient*. New York: Free Press.
- Bradburn, N. 1969 *The Structure of Psychological Well-Being*. Chicago: Aldine.
- Coe, R. M. 1987 "Communication and Medical Care Outcomes: Analysis of Conversations Between Doctors and Elderly Patients." In R. Ward, and S. Tobin, eds., *Health and Aging: Socioissues and Policy Directions*. New York: Springer.
- DeForge, B.R., and J. Sobal, 1988 "Self-Report Depression Scales in the Elderly: The Relationship Between the CES-D and Zung." *International Journal of Psychiatry in Medicine* 18:325-338.
- Duke University Center for the Study of Aging and Human Development 1978 *Multidimensional Functional Assessment: The OARS Methodology*. Durham, N.C.: Duke University Medical Center.
- Dupuy, H. J. 1984 "The Psychological General Well-Being (PGWB) Index." In N. Wenger, M. Mattson, C. Furberg, and J. Elinson, eds., *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*. New York: Le-Jacq.
- Friedman, G.D. 1988 *Primer of Epidemiology*, 3rd ed. New York: McGraw-Hill.
- Fries, J. F., and L. M. Crapo, 1986 "The Elimination of Premature Disease." In Ken Dychtwald, ed., *Wellness and Health Promotion for the Elderly*. Rockville, Md.: Aspen.
- Gottlieb, N. H., and L.W. Green, 1987 "Ethnicity and Lifestyle Health Risk: Some Possible Mechanisms." *American Journal of Health Promotion* 2(1):37-51.
- Herd, J. A., J. J. W. Alastair, J. Blumenthal, J. E. Daugherty, and R. Harris, 1987 "Medical Therapy in the Elderly." *Journal of the American College of Cardiology* 10:29-34.
- Hess, J. W. 1991 "Health Promotion and Risk Reduction for Later Life." In R. Young and E. Olson, eds., *Health, Illness, and Disability in Later Life: Practice Issues and Interventions*. Newbury Park: Sage.
- Kannel, W. B., J. T. Doyle, R. J. Shepard, et al. 1987 "Prevention of Cardiovascular Disease in the Elderly." *Journal of the American College of Cardiology* 10A:25-28.
- Keller, D. M., M. G. Kovar, J. B. Jobe, and L. G. Branch, 1993 "Problems Eliciting Elder's Reports of Functional Status." *Journal of Aging and Health* 5:306-318.
- Lawton, M. P., M. Moss, M. Fulcomer, and M.H. Kleban, 1982 "A Research and Service-Oriented Multilevel Assessment Instrument." *Journal of Gerontology* 37:91-99.
- Leavitt, R. 1990 "The Appreciation of Cultural Diversity. How to Integrate Content into Curriculum." In G. Price and P. Fitz, eds., *Issues in Aging: Cultural Diversity and the Allied Health Curriculum*. Hartford, Conn.: University of Connecticut.
- Levine, S., and S. H. Croog 1984 "What Constitutes Quality of Life? A Conceptualization of the Dimensions of Life Quality in Healthy Populations and Patients With Cardiovascular Disease." In N. Wenger, M. Mattson, C. Furberg, and J. Elinson, eds., *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*. New York: Le-Jacq.
- Lieberman, L. S. 1988 "Diabetes and Obesity in Elderly Black Americans." In J. S. Jackson, ed., *The Black Elderly: Research on Physical and Psychosocial Health*. New York: Springer.
- McGinnis J. M., and W. H. Foege, 1993 "Actual Causes of Death in the United States." *Journal of the American Medical Association* 270(18): 2207-2212.
- McKinlay, J.B., and S.M. McKinlay, 1977 "The Questionable Contribution of Medical Measures to the Decline of Mortality in the United States in the Twentieth Century." *Health and Society* 55(3):405-426.
- McNair, D., M. Lorr, and L. Dopleman, 1971 *Manual for the Profile of Mood States*. San Diego: Educational and Industrial Testing Service.
- National Center for Health Statistics 1996 *Vital Statistics of the United States 1991*, vol. 2. Hyattsville, Md.: U.S. Government Printing Office.
- New York Times. "Clinton Plans 125 Million Initiative on Infectious Disease." December 27, 1998.
- Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services Public Health Service 1994 *Clinician's Handbook of Preventive Services*. Washington D.C.: U. S. Government Printing Office.
- Orlandi, M. A. 1987 "Clinical Perspectives. Promoting Health and Preventing Disease in Health Care Settings: An Analysis of Barriers." *Preventive Medicine* 16:119-130.

- Patrick, D. L., and R. A. Deyo 1989 "Generic and Disease-Specific Measures in Assessing Health Status and Quality of Life." *Medical Care* 27 (Suppl. 3):S217-S232.
- Polednak, A.P. 1989 *Racial and Ethnic Differences in Disease*. New York: Oxford University Press.
- Pope, A. M., and A. Tarlov 1991 *Institute of Medicine Disability in America*. Washington, D.C.: National Academy Press.
- Prochaska, J. O., and C. C. DiClemente 1984 *The Transtheoretical Approach: Crossing the Traditional Boundaries of Therapy*. Chicago: Dow Jones/Irwin.
- Rimer, B. K., J. M. Schildkraut, C. Lerman, T. H. Lin, and J. Audrain, 1996 "Participation in a Women's Breast Cancer Risk Counseling Trial: Who Participates? Who Declines?" *Journal of American Cancer Society* 77(11):2348-2355.
- Rosenstock, I. M. 1974 "Historical Origins of the Health Belief Model." *Health Education Monographs* 2:344.
- Tseng, R. and D. Ellyne, 1990 "Perspectives on Developing a Course in Multicultural Health Promotion." *Issues in Aging*, 44-48.
- U.S. Bureau of Census 1994 *Current Population Reports*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Preventive Services Task Force 1989: *Guide to Clinical Preventive Services*. Baltimore: Williams and Wilkins.
- Verbrugge, L. M. 1990 "The Twain Meet: Empirical Explanations of Sex Differences in Health and Mortality." In M. G. Ory and H. R. Warner, eds., *Gender, Health, and Longevity*. New York: Springer Publishing Company.
- Waller, J. B., R. Young, and J. R. Sowers, 1994 *Frail Elderly* (Report to National Institutes of Health, NIA No. 10428). Washington, D.C.: U.S. Government Printing Office.
- Wallston, K. A., B. S. Wallston, and R. DeVellis, 1978 "Development of the Multi-Dimensional Health Locus of Control MHLC Scales." *Journal of Health Education Monographs*. 6(2):160-170.
- Young, R. F. 1994 "Older People as Consumers of Health Promotion Recommendations." *Generations* 18(1):69-73.
- 1998 "Delay in Breast Cancer Screening." Report to Karmonos Cancer Institute.
- , E. Kahana, and M. Rubenfire, 1987 "Preventive Health Behavior Advice: A Study with Older Myocardial Infarction Patients." *Evaluation and the Health Professions*. 10:4394-4407.
- Young R. F., and E. A. Olson, 1993 "Health Promotion Among Minority Aged: Challenges for the Health

Professions." *Journal of Continuing Education in the Health Professions* 13:235-242.

ROSALIE F. YOUNG

## HEALTH STATUS MEASUREMENT

See Health and Illness Behavior; Health Promotion and Health Status; Quality of Life; Medical Sociology.

## HETEROSEXUAL BEHAVIOR PATTERNS

See Courtship; Sexual Behavior in Marriage and Close Relationships; Sexual Behavior Patterns; Sexual Orientation; Sexual Violence and Exploitation.

## HIERARCHICAL LINEAR MODELS

Hierarchical linear models are applicable in situations where data have been collected from two (or more) different levels. Sociology's initial interest in such multilevel relationships can be traced back to Durkheim's research into the impact of community on suicide (Durkheim [1898] 1951). More recently, these models have been related to the topic of contextual analysis (Boyd and Iversen 1979), where researchers are interested in investigating linkages between micro-level and macro-level variables. Sociological theories have been classified into three groups according to the degree to which they incorporate multilevel variables (Coleman 1986). In one group, variation in a dependent variable is explained through independent variables obtained from the same social level (e.g., country, community, individual). In a second group, attempts are made to account for differences in a dependent variable at one level by examining variation in an independent variable at a higher level; and in a third group, variations in a dependent variable are explained by variations in an independent variable at a lower level. Theories



that fall into either the second or third group are multilevel theories and can be explored using hierarchical linear models.

**SPECIFICATION OF THE HIERARCHICAL LINEAR MODEL**

A wide variety of hierarchical models can be specified. However, in order to outline the basic features of such models, a simple example will be developed. Assume that a researcher is interested in modeling the length of hospital stay (LOS) for a specific individual ( $Y_i$ ) as a function of the severity of that individual's illness ( $X_i$ ) and the bed occupancy rate for the institution in which that individual is hospitalized ( $G_j$ ). In this hypothetical model we have one criterion (or dependent) variable,  $Y_i$ , at the micro level, one micro-level predictor (or independent) variable,  $X_i$ , and one macro predictor (or independent) variable,  $G_j$ . This produces a two-level hierarchical model. The technique is quite flexible and can be expanded to include multiple predictor variables at either (or both) the micro- and macro-levels and additional levels. In the given example, an index of individual comorbidity could be included as an additional micro-level predictor, type of hospital (e.g., public vs. private) could be included as an additional macro-level predictor, and an additional level of the gross national product (GNP) of the country in which the hospital is located could be added to create a three-level model.

The first step in developing hierarchical models is to specify a model for the micro-level variables that is identical for all contexts. In the present example a linear model relating LOS as a function of severity of illness is specified for each of the hospitals.

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \varepsilon_{ij} \tag{1}$$

Where  $j=1, 2, \dots, j$  denotes the macro-level contexts (e.g., the hospitals) and  $i=1, 2, \dots, n_j$  denotes micro-level observations within contexts (e.g., individuals within hospitals). The intercepts from Equation 1 ( $\beta_{0j}$ ) provide estimates of the expected LOS for individual  $i$  in hospital  $j$  whose severity of illness is zero, whereas the slopes ( $\beta_{1j}$ ) provide estimates for the effect of a unit change in the severity of the illness for individual  $i$  in hospital

$j$ . Finally, the  $\beta_{ij}$ 's represent random errors or residuals. It is assumed that these errors are normally distributed within each context with a mean of zero and a constant variance  $\sigma^2$ . This is a standard linear model with the exception that the coefficients (i.e., the  $\beta_j$ 's) are allowed to vary across contexts (hospitals).

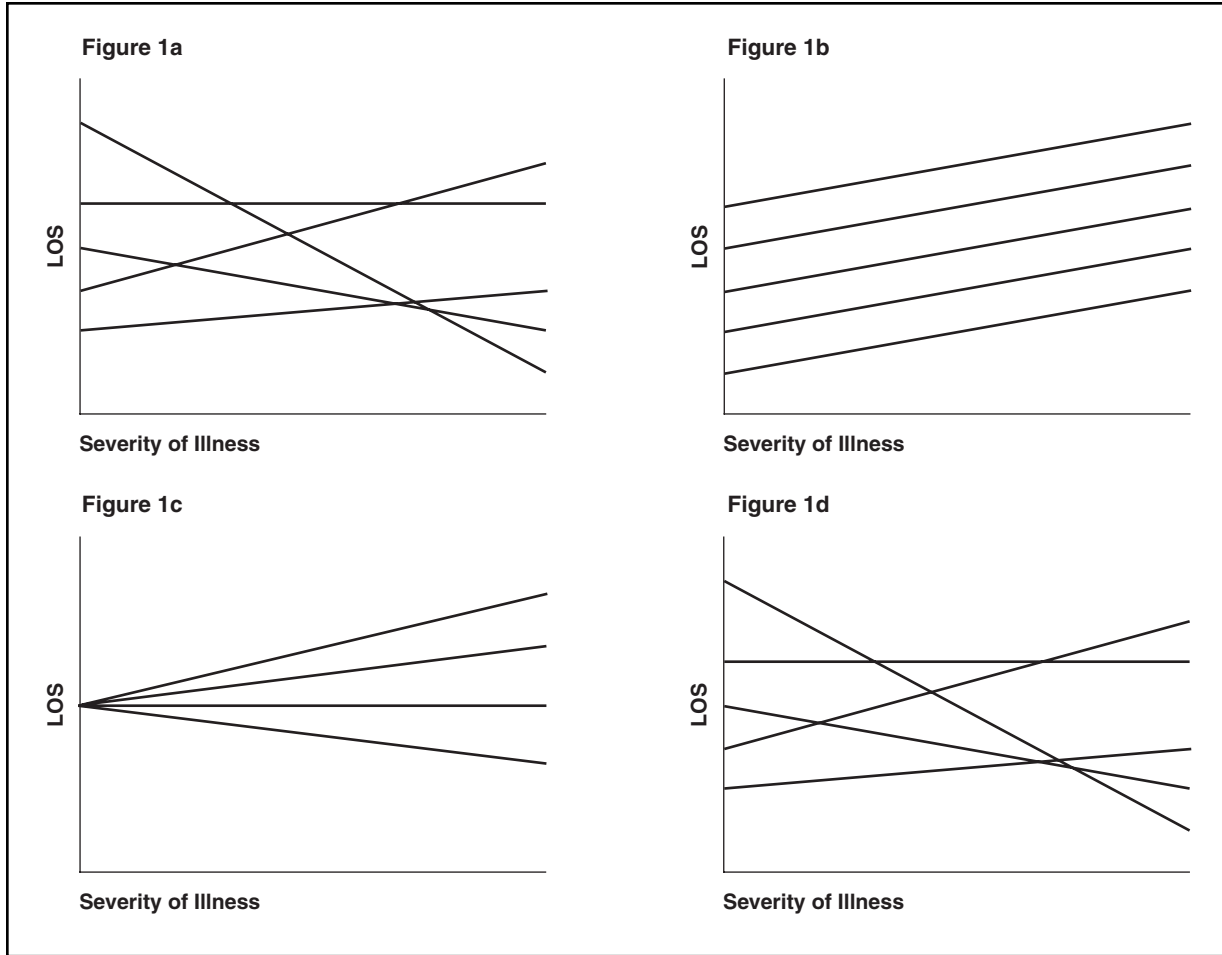
In situations where separate regression equations are estimated for various contexts, four different patterns can emerge. These patterns are depicted in Figures 1a, 1b, 1c, and 1d. In Figure 1a, the functional relationship between the micro-level variables is identical for all the contexts, and thus the intercepts and slopes are the same for all contexts. In Figure 1b, the degree of linear relationship between the micro-level variables is equivalent across contexts; however, the initial "location" (i.e., the intercept) of this relationship varies across contexts. In Figure 1c, the degree of linear relationship between the micro-level variables varies as a function of context, although the initial "location" is consistent across contexts. Finally, in Figure 1d, both the initial location and the relationship between the micro-level variables vary significantly across contexts.

Systematic differences across contexts are reflected in three of the figures (viz., Figures 1b, 1c, and 1d). The presence of these differences leads to questions of whether there are contextual or macro-level variables that could be associated with the varying micro-level coefficients (i.e., the slopes and/or intercepts). Questions of this type are addressed by specifying a second-level model. For example, if there is significant variation among the micro-level coefficients, then this variation could be modeled as a function of contextual or macro-level variables as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}G_j + U_{0j} \tag{2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}G_j + U_{1j} \tag{3}$$

where  $G_j$  is a contextual (or macro-level) variable,  $\gamma_{00}$  and  $\gamma_{10}$  are the intercepts from the second-level models,  $\gamma_{01}$  and  $\gamma_{11}$  are the slopes from the second-level model, and  $U_{0j}$  and  $U_{1j}$  are the second-level residuals. It is assumed that the residuals are distributed multivariate normal with mean vector  $\mathbf{0}$  and variance-covariance matrix  $\mathbf{T}$ . In the present example, Equation 2 would be used to model differences across hospitals among the intercepts



of the micro-level equations (cf. Figures 1b and 1d), whereas Equation 3 would be used to model differences across hospitals in the slopes of the micro-level equations (cf. Figures 1c and 1d).

Depending on the actual variability of the micro-level coefficients (i.e., the  $\beta_j$ 's), different second-level models would be justified. For example, in situations where there is no variation in the slopes across contexts (see Figure 1b), the inclusion of  $G_j$  in Equation 3 would not be meaningful given that  $\beta_{1j}$  is the same across all contexts. Similarly, in situations where there is no variation in the intercepts across contexts (see Figure 1c), the inclusion of  $G_j$  in Equation 2 would not be meaningful given that  $\beta_{0j}$  is the same across all contexts.

By substituting Equations 2 and 3 into Equation 1, we can obtain a single equation form of the hierarchical model as follows:

$$Y_{ij} = \gamma_{00} + \gamma_{01} G_j + \gamma_{10} X_{1ij} + \gamma_{11} G_j X_{1ij} + (U_{0j} + U_{1j} X_{1ij} + \varepsilon_{ij}) \tag{4}$$

The model represented by Equation 4 is a mixed model with both fixed coefficients (viz., the  $\gamma$ 's) and random coefficients (viz., the  $U$ 's and the  $\varepsilon$ 's). Further, since the random coefficients are allowed to covary across contexts, it can be called a variance component model.

The approach to investigating relationships occurring across hierarchical levels represented by the equations above is not new. Burstein and colleagues (1978) discussed a similar approach under the conceptualization of "slopes as outcomes." Conceptually, this is an accurate description, given that the regression coefficients estimated within each context at the micro level are used

as criterion (or dependent) measures in the macro-level (or second-level) model (cf. Equations 2 and 3). However, while this conceptualization of the relationship between micro- and macro-level variables has been understood for a number of years, concerns about the adequacy of estimating such models using traditional statistical techniques (viz., ordinary least squares, OLS) have been expressed. However, separate statistical advances throughout the 1980s improved the estimation procedures for these models (for reviews see Burstein and colleagues 1989; Raudenbush 1988), with the advances resulting in several different software packages being developed specifically for the estimation of hierarchical linear models (e.g., GENMOD, HLM, ML3, and VARCL).

### ESTIMATION OF HIERARCHICAL LINEAR MODELS

In estimating the various components of the hierarchical linear model, a distinction is made among fixed effects, random effects, and variance components. Specifically, fixed effects are those parameter estimates that are assumed to be constant across contexts (e.g., the  $\gamma$ 's from Equations 2 and 3), whereas random effects are parameter estimates that are free to vary across contexts (e.g.,  $\beta_{0j}$  and  $\beta_{1j}$  from Equation 1). Hierarchical linear models also allow for the estimation of the variance components of the model. These include (1) the variance of the residuals from the micro-level model (i.e., the variance of the  $\epsilon_{ij}$ 's identified as  $\sigma^2$  above); (2) the variance of the second-level residuals (i.e.,  $U_{0j}$  and  $U_{1j}$ ); and (3) the covariance of the second-level residuals (i.e., the covariance of  $U_{0j}$  and  $U_{1j}$ ). The variance-covariance matrix of the second-level residuals was previously defined as  $\mathbf{T}$ .

**Estimation of Fixed Effects.** One approach that could be used to estimate the  $\gamma$ 's from Equations 2 and 3 is traditional OLS regression. However, because the precision of estimation of these parameters will vary as a function of contexts, the usual OLS assumption of equal error variances (i.e., homoscedasticity) will be violated. In order to deal with this violation the second-level regression coefficients (the  $\gamma$ 's) are estimated using a more sophisticated procedure, generalized least squares (GLS). GLS techniques provide weighted estimates

of the second-level regression coefficients such that the contexts that have more precise estimation of the micro-level parameters receive more weight in the estimation. That is, those contexts in which there is greater precision in estimating the parameters (the slopes and the intercepts) receive more weight in estimating the second-level regressions.

**Estimation of Variance-Covariance Components.** The components of the variance-covariance matrix  $\mathbf{T}$  include the variance of the micro-level residuals, and the variance and covariance of the second-level residuals. These components are used in the GLS estimation of the fixed effects of the second-level model. However, the values of the components of this matrix are typically not known and must be estimated. The best methods for doing this are iterative methods that alternatively estimate the parameters of the models and then estimate the variance-covariance matrix  $\mathbf{T}$  until a convergence is reached. Hierarchical linear models adopt the EM algorithm (Dempster et al. 1977) that produces maximum likelihood estimates for the variance-covariance components of  $\mathbf{T}$ .

**Estimation of Random Effects.** The simplest way of estimating the coefficients for the micro-level model (i.e., Equation 1) is to compute an OLS regression for a specific context. In the present example, this would involve obtaining a regression equation relating expected LOS to severity of illness for all individuals within a specific hospital. If there are reasonably large sample sizes within each context, this analysis would provide relatively precise estimates of the coefficients of interest. These estimates will not be stable, however, if sample sizes are smaller. Further, inspection of the second-level models reveals that there is a second estimate of the coefficients from the micro-level models. Thus, for any particular observational unit there are two separate estimates of the micro-level regression coefficients: one from the micro-level regressions themselves and the other from the second-level regression model. The question that this leaves is which of these provides a more accurate estimate of the population parameters for the particular observational unit.

Rather than forcing a choice between one of these two estimates, hierarchical linear models use empirical Bayes estimation procedures (Morris

1983) to compute an optimally weighted combination of the two estimates. The empirical Bayes estimates are a weighted composite of the two estimates discussed above. The micro-level regression coefficients (the  $\beta_j$ 's) estimated by OLS are weighted according to the precision with which they are estimated (i.e., their reliability). In cases where the OLS estimates are not very reliable (e.g., due to small sample size), the empirical Bayes procedure allots greater weight to the second-level estimates. Essentially, then, the weighted composite "shrinks" the micro-level estimate toward the second-level estimate, with the level of shrinkage being determined by the reliability of the micro-level estimate. It has been demonstrated that, in general, the empirical Bayes estimates have smaller mean squared errors than OLS estimates.

**Statistical Tests.** A variety of statistical tests for hypothesis testing are provided by the various computer programs used to estimate hierarchical linear model. For example, HLM (Bryk et al. 1994) computes a *t*-test to evaluate the hypothesis that the second-level regression parameters depart significantly from zero. In addition, chi-square tests are provided for tests of whether or not there is significant variation in the second-level residuals. These latter tests allow the researcher to determine the model that best fits the observed data. For example, it might be that there is no significant variation in the slopes across contexts; however, there might be significant variation in the intercepts (as in Figure 1b).

**FURTHER ISSUES WITH HIERARCHICAL LINEAR MODELS**

**Centering.** Often, as in the present example, interpretation of the intercepts is not straightforward, since a value of zero for the independent variable (in the present case, severity of illness) is not meaningful. In situations like this, it is possible to "center" the independent variable as a deviation from the mean level of that variable in the sample as follows.

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{1j}) + \varepsilon_{ij} \quad (5)$$

With this specification, the intercepts now represent estimates of the expected length of stay for individuals in a specific hospital whose severity of

illness is at the mean. The interpretation of the other parameters remain unaltered.

**Longitudinal Data.** Hierarchical linear models can also be used to analyze longitudinal data collected in order to examine questions regarding the assessment of change (Bryk and Raudenbush 1987). Under this approach, there are repeated observations within an observational unit and there is a sample of different units. This allows for a two-level conceptualization of development such that change in the individual units is modeled as a function of time and differences in the patterns of change across individual units can be modeled as a function of measurable characteristics of the individual units. Under this conceptualization, interest is in between-individual (unit) differences in within-individual (unit) change.

**Statistical Software.** As previously noted, a number of different software programs have been specifically developed in order to estimate hierarchical linear models. Kreft and colleagues (1994) reviewed five of the then-available packages. While they recommended ML3 (Prosser et al. 1991) for the "serious" user, they concluded that HLM's main advantage is its ease of use. Since that time, both programs have been updated and now versions for Windows '95 are available (viz., HLM 4 and MlwiN. Information on the latest version of these programs is available from the following Web sites: go to <http://www.ssicentral.com/hlm/mainhlm.htm> for information on HLM 4; go to <http://www.ioe.ac.uk/mlwin> for information on MlwiN.).

**CONCLUSION**

Hierarchical linear models provide statistically sophisticated ways for dealing with analyses in which data are obtained from multiple levels. Such data are common in sociological research, especially if the investigation deals with contextual effects or longitudinal designs. For more detailed discussions of hierarchical linear models, the interested reader is directed to the following sources that provide more in-depth coverage: Bryk and Raudenbush (1992); Goldstein (1995); or Hox (1995). (At the time of publication, a complete, online version of this text was available from: <http://ioe.ac.uk/multilevel.what-new.html>.)

## REFERENCES

- Boyd, L. H., and G. R. Iversen 1979 *Contextual Analysis: Concepts and Statistical Techniques*. Belmont, Calif.: Wadsworth.
- Bryk, A. S., and S. W. Raudenbush 1987 "Application of Hierarchical Linear Models to Assessing Change." *Psychological Bulletin* 101:147-158.
- 1992 *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage.
- , and R. J. Congdon 1994 *Hierarchical Linear Modeling with the HLM/2L and HLM/3L Programs*. Chicago: Scientific Software International.
- Burstein, L., K. S. Kim, and G. Delandshere 1989 "Multilevel Investigations of Systematically Varying Slopes: Issues, Alternatives, and Consequences." In R. D. Bock, ed., *Multilevel Analysis of Educational Data*. New York: Academic Press.
- Burstein, L., R. L. Linn, and F. J. Capell 1978 "Analyzing Multilevel Data in the Presence of Heterogeneous Within-Class Regressions." *Journal of Educational Statistics* 3:347-383.
- Coleman, J. S. 1986 "Social Theory, Social Research, and a Theory of Action." *American Journal of Sociology* 91:1309-1336.
- Dempster, A. P., N. M. Laird, and D. B. Rubin 1977 "Maximum Likelihood from Incomplete Data Via the EM Algorithm." *Journal of the Royal Statistical Society, Series B* 39:1-8.
- Durkheim, E. (1898) 1951 *Suicide*. Glencoe Ill.: Free Press.
- Goldstein, H. 1995 *Multilevel Statistical Models*. New York: Halstead.
- Hox, J. J. 1995 *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.
- Kreft, I., J. de Leeuw, and R. van der Leeden 1994 "Review of Five Multilevel Analysis Programs: BMDP-5V, GENMOD, HLM, ML3, VARCL." *The American Statistician* 48:324-335.
- Morris, C. 1983 "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association* 78:47-65.
- Prosser, R., J. Rasbash, and H. Goldstein 1991 *ML3: Software for Three-Level Analysis*. London: Institute of Education, University of London.
- Raudenbush, S. W. 1988 "Educational Applications of Hierarchical Linear Models: A Review." *Journal of Educational Statistics* 13:85-116.

GEORGE ALDER

## HIGHER EDUCATION

Colleges and universities seem to defy the maxim that only highly rationalized institutions can succeed in the modern world. Only the Catholic Church has a longer continuous existence among Western institutions. Higher education has done more than survive; it is in many ways a pivot of key developments in the social structure and culture. It is central for the generation of research and technological innovations. It is also central in the selection, training, and credentialing of young men and women for higher-level positions in the occupational structure.

Among the most important sociological questions surrounding higher education are the following: (1) To what extent have advanced industrial societies become based on a "knowledge economy" closely related to university research and training? Related to this question is another: To what extent do we see the rise of a "new class" of "knowledge workers" with advanced training—differing in interest and outlook from both business elites and earlier aristocracies of labor? (2) To what extent do institutions of higher education reproduce social inequalities by certifying the cultural advantages of children from the upper classes, or reshuffle the social hierarchy by rewarding intellect and ability independent of students' social-class background? (3) Do institutions of higher education, with their traditions of collegial control and tenure, represent an alternative model to corporate forms of organization? These issues can be addressed only after examining the historical development, the existing organizational structures, and the contemporary pressures on higher education.

First, it is necessary to define the dimensions of higher education. Formal educational systems are conventionally divided between primary (the first six years), secondary (the next four to six years), and postsecondary education. Some postsecondary schools offer courses of study that are narrowly vocational and very short in duration. These institutions (including secretarial, business, and vocational-technical colleges) are not usually considered to be part of higher education. Institutions must award degrees that are recognized by baccalaureate-granting colleges and universities to be considered part of higher education. By this criterion, community colleges in the United States,

*Fachhochschulen* in Germany, and the “further education” colleges in the United Kingdom represent bottom tiers of their respective national higher education systems. These are all short-cycle, vocationally oriented institutions, but some of their degrees are transferable to higher-level colleges and universities. Above these lower-tier institutions are a vast array of colleges, universities, and specialized institutions (for example, seminaries and art schools) that constitute the core of the higher education sector in all contemporary societies. Levels in this institutional hierarchy are structured, most fundamentally, by the type of credentials offered. In the United States, for example, levels are marked by movement from the associate to the baccalaureate to the master’s to the doctoral degree.

### HISTORICAL DEVELOPMENT

The distant relatives of today’s institutions of higher education go back in the West to the Greek academies of the fourth and fifth centuries B.C.E. In these academies, young men from the governing classes studied rhetoric and philosophy (and lesser subjects) as training for public life (Marrou [1948] 1982). In the East, the roots of higher education go back to the training of future government bureaucrats at the feet of masters of Confucian philosophy, poetry, and calligraphy. In both East and West, a close relationship existed among social class, high culture, and preparation for public life.

However, modern institutions of higher education trace a more direct lineage from the medieval *studium generale*. In the first European universities of the twelfth and early thirteenth centuries (notably Salerno, Bologna, and Paris), students and masters came together to pore over the new knowledge discovered in ancient texts and developed by the Arab scholars of Spain. These gatherings of students and teachers were a product of the revival of scholarly inquiry in what has been called the “twelfth-century Renaissance.” The medieval universities were similar to modern higher education in that they were permanent institutions of learning with at least a rudimentary formal organization. Courses of study were formally organized, lectures and examinations were given at scheduled times, administrative officials presided, graduation ceremonies were held, and students lived in

lodgings near the university buildings. The *studium generale*, or leading universities, were recognized as such because they housed at least one of the “higher faculties” in law, medicine, or theology in addition to faculties of the arts. Courses in the arts, typically with an emphasis on logic and philosophy, were common preparation for study in the three learned professions. Thus, from the beginning, a certain vocational emphasis is evident in the university. Degrees awarded on the completion of professional studies certified accomplishments that made their recipients worthy of entry into professional life. Nevertheless, the spirit of inquiry was equally important in the medieval universities; these were places renowned for famous teachers, such as Abelard in Paris and Irnerius in Bologna. Civic competition led to a proliferation of universities. By the end of the Middle Ages, eighty had been founded in different parts of Europe (Rashdall [1985] 1936).

In the seventeenth and eighteenth centuries, the fortunes of colleges and universities waned. The causes for decline are numerous, including the attractiveness of commercial over scholarly careers, the interference (in some places) of religious and political authorities, and the insularity of faculty who jealously guarded their guild privileges but resisted new currents of thought. During this period, colleges and universities became places concerned with the transmission of ancient texts rather than the further advance of knowledge. Professional training moved out of the universities: into Inns of Court, medical colleges, and seminaries. New elites interested in technical and scientific progress established entirely new institutions rather than allying with the colleges and universities. Napoleon, for example, founded elite professional training institutions, the *grandes ecoles*, and the early investigators in the natural sciences created separate elite societies to encourage research and discussion.

The revived university is the product of nineteenth-century European reform movements, led in the beginning by intellectually oriented aristocrats and eminent philosophers and theologians. The University of Berlin, founded in 1810, was the first reformed university, and others shortly followed in its wake. The new university was founded on the “Humboldtian principles” of the unity of teaching and research (meaning that both functions were performed by the professoriate) and

the freedom to teach and learn without fear of outside interference. The development of new academic components, such as the research seminar and the specialized lecture, created an environment in which path-breaking researchers, such as Leopold Ranke in history and Justus von Liebig in chemistry, emerged (McClelland 1980). By mid-century, the German research universities had become a model for reformers throughout Europe and from as far away as the United States and Japan. The first research university in the United States, Johns Hopkins University, founded in 1876, was explicitly modeled on the German research university.

Higher education's current emphasis on training for a wide range of applied fields has an equally important history. Here the United States, rather than Germany, was the decisive innovator. The Morrill Acts (passed in 1862 and 1890) provided funds for states to establish "land grant" universities to provide both general education and practical training in agricultural and mechanical arts for all qualified applicants. Such institutions encouraged both the democratization of American higher education and a closer connection between universities and emerging markets for educated labor. The American university's role in society was further enhanced by its willingness to work collaboratively with government, professional associations, and (somewhat later) business and community organizations. The "Wisconsin Idea" encouraged close connection between university experts and government officials during the period before World War I. Universities also cooperated closely with professional associations to raise educational training standards. Connections between university and state were extended, particularly in the sciences, during World War II and the Cold War, when government grants for university-based scientific research became a very large source of support. These developments encouraged a new view of higher education. In the 1960s, Clark Kerr (1963) coined the term "multiversity" to describe institutions, like his own University of California, as service-based enterprises specializing in training, research, and advice for all major sectors of society. Junior colleges, first established just after the turn of the twentieth century, were by the 1960s even more systematically tied than universities to local and regional markets for semiprofessional and technical labor (Brint and Karabel

1989). In terms of growth, these two-year colleges are the great success story of twentieth-century higher education, and their influence is now evident even in four-year institutions. The utilitarian approach of American educators was resisted for some time in Europe and Asia, where access to higher education was strictly limited to those students who passed rigorous examinations and where higher degrees had long served as important badges of social status linked to cultural refinement. However, by the last quarter of the twentieth century, the entrepreneurial multiversity had become an important model throughout the developed world (Clark 1998).

Institutions of higher education rarely shed their earlier identities completely; instead, they incorporate new emphases through reorganizing and adding new components and new role expectations. Today, all major historical stages of university development remain very much in evidence. Much of the nomenclature, hierarchy, and ritual of the medieval university remains and is in full display at graduation ceremonies. Although the major fields of study have changed dramatically, the underlying liberal arts emphasis of the ancient academies has remained central in the first two years of undergraduate study (the lower division). The nineteenth-century emphasis on specialization is evident in the second two years of undergraduate study (the upper division) and in graduate and professional programs. The nineteenth-century emphasis on research remains an absorbing occupation of faculty and graduate students. The twentieth-century emphases on ancillary training, service, and advisory activities are organized in separate components (as in the case of university extension programs, agricultural experiment stations, university-based hospitals, and collegiate sports teams) or performed by research faculty in their capacity as consultants and lecturers in the community.

### ACADEMIC ORGANIZATION

Contemporary institutions of higher education are organized horizontally by divisions among fields of knowledge and vertically by ranks of authority. The dual hierarchy of professors and administrators is a structural feature of academic organization with particularly important consequences.

Knowledge is divided among discipline-based departments and interdisciplinary programs and research centers. Interdisciplinary programs have become increasingly important in the organization of research activity. However, academic appointments remain based in departments. For this reason, departments must be considered fundamental to academic organization. In chair systems, characteristic of continental European universities, one or two senior professors hold chairs and organize research programs, while the other professors serve in subsidiary roles under the direction of the chair. In American universities, departmental faculty operate independently, pursuing their own research programs, only occasionally in collaboration.

The larger structures of knowledge-based organization are the colleges and professional schools. A college of humanities will, for example, typically include all departments in the fine arts (such as music and theatre) and the humanities disciplines (such as philosophy and English). Colleges and professional schools are administrative units. The number of colleges and professional schools varies by the size of the campus. A very large campus will have separate divisions for the arts, humanities, social sciences, biological sciences, and physical sciences. It may also have half a dozen or more professional schools. A small campus may have only a single college of arts and sciences.

Colleges and universities are under no obligation to represent all fields of basic and applied knowledge, and most do not. (The term “university” does not, as many believe, refer to the universe of all fields of knowledge. Originally, it meant simply “an aggregate of persons.”) New disciplines must fight for a place in the university, and old disciplines sometimes fragment or disappear altogether. Sociology and psychology, for example, both broke away from philosophy, while the nineteenth-century discipline of political economy eventually divided into political science and economics. Today the fate of disciplines in particular colleges and universities depends on a number of factors, including, most notably, student demand for courses in the department, the strength of the departmental faculty in national ratings, the ability of faculty to bring in grants and contracts, the exodus or retirement of “star” faculty, and the effectiveness of the department in making its case for new hiring. These factors have little to do with

any purely objective intellectual principles of value. Over the long run, disciplines with a strong profile in the labor market and those with access to large research grants have been least vulnerable to retrenchments. One might even argue that the modern university is moving away from a liberal arts core in the direction of a “practical arts core” composed of departments closely tied to technological and economic advance or to national security (such as economics, molecular biology, physics, and international affairs) and professional schools providing training for the highest-income occupations (such as medicine, law, and finance). Strength in this practical arts core does not necessarily come at the expense of strength in traditional liberal arts disciplines, however. In the larger universities, powerful disciplines help to subsidize less powerful ones, which, in turn, may teach a disproportionate share of students.

Modern institutions of higher education are far from *collegiua* in their authority structure, but they also do not fit an ideal-type corporate model of top-down control. Instead, decision-making practices are based, at least in principle, on divided spheres of power and ongoing consultation among the major “branches” of institutional governances. In this dual structure, both administrative and knowledge-based authority are represented. The authority structure of knowledge is constituted by the departments and, within the departments, by the professorial ranks. Advancement in the professorial hierarchy is based in principle on the quality of a faculty member’s professional accomplishments (typically involving assessments of research, teaching, and service). Differences in rank are associated with both deference and income. This hierarchy moves from the temporary ranks of lecturer and instructor to the regular ranks of assistant, associate, and full professor. Highly visible full professors may be appointed to named chairs that provide both additional symbolic recognition and a separate budget for research and travel.

The top level of the administrative hierarchy is composed of a president or chancellor, who is responsible for fund-raising and interaction with important resource providers as well as overall supervision; a provost or executive vice chancellor, who is responsible for internal academic matters; and the deans of the colleges and schools. Top administrators are usually drawn from members of the faculty, although an increasing number



of lower-tier institutions now hire professional managers at the presidential level. Top administrators make the ultimate decisions about budget allocations, hiring and promotion, and planning for the future. However, the faculty, at the leading authorities in their specialized domains of knowledge, retain a decisive say at least in the better institutions, over all decisions involving curricular organization and instruction. They also retain the predominant say in hiring and promotion decisions within the academic departments, expecting only very rare overrules by administrators. The faculty typically play a significant advisory role in the development of new programs and centers and in discussions of institutional priorities. Top universities depend for prestige and resources on the accomplishments of their faculty; as a general rule, the less distinguished the faculty, the more powerful the administration (Blan 1973). Faculty in nonelite institutions have, consequently, sometimes chosen to organize in collective bargaining units to control administrative discretion through contractual means (Rhoades 1998).

The unique institution of tenure greatly enhances the influence of faculty. After a six-year probationary period, assistant professors come up for a decision on promotion to tenure and accompanying advancement in rank. Tenure, a conventional rather than a legal status, guarantees lifetime employment for those who continue to hold classes and act within broad bounds of moral acceptability. Together, dual authority and tenure guarantee opposition to any administrative efforts to abandon existing programs or to downgrade the work conditions and privileges of faculty.

The primary funding for colleges and universities varies by national circumstances. Most institutions of higher education in Europe and in the developing world are state-supported. Modest fees are sometimes charged students enrolling in expensive or high-demand fields. However, the idea of tuition is only now developing. In the United States, public colleges and universities are primarily supported by state appropriations, but they also charge tuition and fees. Private colleges and universities, lacking state appropriations, charge substantially higher tuitions and, therefore, attract a larger proportion of students from the higher social classes. They also use interest from their investments to support the operating budget. In

both public and private universities, research contracts and grants are another important source of funds.

### COMPARISON OF NATIONAL SYSTEMS

Sociologists frequently use the term “system” to describe national patterns of higher education. This term should be used advisedly, since many national “systems” are not in fact highly coordinated. Societies with strong traditions of state planning have relatively centralized systems. The Russian, French, and Swedish systems remain among the most centralized today. But even in these countries, some private institutions operate independently of the centrally organized public system. Societies with weak traditions of state planning and strong traditions of voluntarism have decentralized and highly diverse systems. American higher education is a clear example of this pattern. Colleges and universities have been organized by religious bodies, secular elites, state legislators, and individual entrepreneurs. The result is a system of some 4,000 largely independent institutions. Institutions emulate and compete with one another in a complex ecological setting whose major dimensions are defined by level of selectivity, by institutional identity (for example, denominational or nondenominational, residential or commuter), and, perhaps most of all, by geography. One of the few forms of regulation is the requirement that curricular programs meet accreditation standards.

It is possible to classify national systems in many ways. Clark (1965) proposed dividing them by the primary influence on the coordination of the system. He placed the former Soviet Union near the pole of state-based coordination and the United States near the pole of market-based coordination. He classified Italy as the clearest example of coordination by an “academic oligarchy.” Here powerful academics were the decisive influence in the development of rules and policies for the system as a whole. Clark argued further that the dominant mode of coordination has important consequences for the ethos and structure of the system. State-based systems place a strong and focused emphasis on manpower planning objectives and scientific and technical development. Market-coordinated systems place consumer interests first and compete to satisfy simultaneously

the social, vocational, and educational interests of students. In "oligarchical" systems, faculty traditions and privileges are jealously preserved and institutions do not change easily.

Clark's framework remains useful. Depending on the question of interest, however, other dimensions of comparison may be equally important. National systems, for example, can also be characterized in relation to their (1) size and openness, (2) institutional diversity, and (3) interinstitutional stratification structure. Countries vary significantly across these three dimensions. The United States represents an unusually large, diverse, and stratified system. Two-thirds of secondary school students enter higher education, but they enter a very heterogeneous set of institutions that are highly stratified by acceptance rates. Germany, by contrast, represents a still relatively small, homogeneous, and unstratified structure. About one-third of German high school students enter higher education. Four-year institutions are designed to be quite similar to one another, and there is no clear ranking system among them. In the United States, therefore, life fates are determined within the system; in Germany, they are determined to a greater degree by inclusion in or exclusion from the system. Some systems in the industrialized world remain relatively small but nonetheless include also a highly differentiated elite track. This is true, for example, in France, where the *grandes écoles* represent a clearly defined upper tier reserved for the very best students. It is also true in Japan, where an institution such as the University of Tokyo retains very close linkages to elite positions in the Japanese state and private economy. Differences across these dimensions have important implications for student consciousness. The highly educated are, for example, more likely to be seen as a separate status group in societies in which access to higher education is relatively restricted. By contrast, opportunity consciousness tends to replace class consciousness in more open systems.

#### SOURCES AND CONSEQUENCES OF GROWTH

Since the 1960s, the trend in the industrialized world has been in the direction of the American model, with an increasing proportion of students entering higher education but with stratification

among institutions and major subjects also increasing. In most countries of Europe, for example, access to higher education is now possible from all secondary school tracks (including vocational tracks) and once-rigorous secondary school-leaving examinations have been relaxed to allow a larger flow of students into higher education. Nevertheless, both attendance and graduation rates in most of the industrialized world remain at about half that in the United States. Thus, higher education in Europe and East Asia is no longer class education, but it has not reached the level of mass education found in the United States.

Theorists of postindustrial society have suggested that the growth of the knowledge sector in the economy is behind this expansion of higher education. Estimates vary on the rate of growth of the "knowledge sector," depending on the definition used. Industries employing high proportions of professionals are growing faster, by and large, than other industries, but some estimates show them slowing down over time (Rubin and Huber 1986). Every estimate shows that they do not as yet contribute a dominant share of the gross national product or even a dominant share of the most dynamic export industries.

The growth of the knowledge sector is undoubtedly an important factor in the expansion of graduate and professional education. Its importance at the undergraduate level is more debatable. In relation to undergraduate enrollments, at least three other sources of growth must be given proper emphasis. One is the interest of states in expanding educational opportunities for their citizens. Another is the interest of students, given these opportunities, to differentiate themselves in the labor market. As secondary school completion approaches universality and higher education attendance becomes more feasible, more students have a motive to differentiate themselves by pursuing higher degrees (Meyer et al. 1979). Finally, and perhaps most important, is the increasing role played by educational credentials as a means of access to desirable jobs in the economy. Credentials are not simply (or in many cases primarily) a guarantee of technical skills. They also signal that their holders are likely to have cultural and personality characteristics sought by employers. These characteristics include middle-class manners, a competitive outlook, literacy and communication skills,

and persistence. Colleges both reward and inculcate these qualities (Collins 1979).

Sociologists agree that restrictive systems of higher education tend to reproduce the social inequalities of the larger society, because the cultural information, motivation, and academic skills needed to pass rigorous examinations are highly correlated with social class. Social-class advantages do not disappear in more open systems, but these systems do generally allow a higher proportion of academically able students from the lower classes to advance. The sheer size of a system does not, however, guarantee decreasing inequality (Blossfeld and Shavit 1993). Much depends on the circumstances of students in the system and the levels of stratification within in the system. Since 1980, the number of college graduates in the United States has continued to grow, but this growth has occurred almost exclusively from among students whose families are in the top quartile of household income. Students from families in the bottom quartiles are entering at higher rates, but they have not graduated at higher rates. The reasons are clear: These students are often less prepared and less motivated to succeed, more likely to feel the press of work and family responsibilities, and more likely to struggle financially with the high cost of four years of college. They are also more likely to enter two-year institutions emphasizing job-related training.

### CONTEMPORARY PRESSURES

Colleges and universities are increasingly costly operations. In state-organized systems, growth is subject to fiscal circumstances and state priorities. In market-organized systems, developments are pushed to a considerable degree by the value of college degrees in the labor market and by competition among colleges and universities. To finance the growth that allows for development of new fields without sharp cutbacks in older fields, colleges and universities compete vigorously for research funds, private gifts, and preeminence in markets for educational services. They also compete vigorously for top faculty and students, the foundations for an institution's reputation.

The high costs of operation and the increasingly competitive environment have led to several

important developments. In the broad field of institutions of higher education, two quite separate market segments tend to develop: one for largely well-to-do students who can afford an expensive four-year residential experience and another for largely moderate- to lower-income students who desire convenience and flexibility as they juggle school, family, and work. In the former, the liberal arts tradition remains strong at the undergraduate level. In the latter, the emphasis is on practical, "consumer-friendly" job-relevant training. As a result of this bifurcation of market segments, the lower tier of liberal arts colleges has begun to disappear in the United States. In most cases, these institutions have transformed themselves into comprehensive colleges with large undergraduate professional programs in areas such as business, engineering, technology, and education (Breneman 1994). The same general trend toward practical, job-relevant training is evident at all but the most selective public four-year colleges and universities.

The size of operations and the increasing competition among institutions have strengthened the influence of top administrators. Managers have started to think strategically about areas of comparative advantage, a striking departure from the model of the past, which emphasized representation of all major fields of study. As a result of this strategic thinking, most departments can no longer depend on automatic replacements for departing faculty, even at the senior level. Administrators have also added resources to student services and development offices to strengthen their relations with key resource providers. For the first time in the postwar period, close partnerships have been developed at some institutions with private firms, which can provide new sources of research funding (Cohen et al. 1998). The ability to attract top students and sizable research grants has improved the position of some departments and schools while weakening the relative position of others. Within institutions, power and influence has continued to shift in the direction of the major professional school faculties and faculties in scientific disciplines with access to large federal research grants. Although still far from completely rationalized along market and bureaucratic-hierarchical lines, institutions of higher education moved significantly in this direction in the last decades of the

twentieth century. Far from reinforcing the position of universities as an alternative to corporate organizations, this movement suggests a tendency for academic institutions to become more like corporations, with power concentrated in the hands of managers, who are conscious, above all, of the markets for their organization's services. Nevertheless, as long as subject-area experts remain central to research and instruction, dual authority will be necessary for academic organization—and academic and corporate forms of organization will never completely converge.

These organizational developments help to explain weaknesses of “new-class” theories. “Knowledge workers” (including professors) do not represent a stratum with social and political interests distinct from those of business elites and nonprofessional workers. Instead, the interests of the knowledge workers are decisively influenced by their particular occupational, organizational, and market circumstances (Brint 1994). This is also true within universities. Those faculty located in professional programs are usually closely allied with top administrators, as are “star” faculty, while those in traditional liberal arts are more likely to express an independent, and somewhat critical, outlook.

All segments of the faculty do, however, share certain guild-like interests in maintaining control over recruitment, employment, and working conditions. The development of new electronic technologies of learning (such as distance learning, “virtual universities,” and Web-based courses) may pose a more significant long-run threat to these guild interests than any of the recent managerial efforts to rationalize campus operations. Studies thus far have not shown consistently significant differences in learning between students taking courses off-site in technologically mediated settings and those taking conventional, on-site courses. This tends to raise questions about the most powerful faculty rationale for the current campus-based organization of academic work. In the future, campuses will undoubtedly continue to exist for elite students, because of the importance of face-to-face contact for building networks that carry over into adult life. They will also be necessary for coursework and research requiring laboratory equipment. It is possible, however, that the number of campuses serving nonelite students will

shrink over the long run, with campus-based instruction shifting toward still more convenient and flexible computer-mediated technologies.

## REFERENCES

- Blan, Peter M. 1973 *The Organization of Academic Work*. New York: John Wiley and Sons.
- Blossfeld, Hans-Peter, and Yossi Shavit 1993 “Persisting Barriers: Changes in Educational Opportunities in 13 Countries.” In Yossi Shavit and Hans-Peter Blossfeld, eds., *Persistent Inequalities*. Boulder, Colo.: Westview Press.
- Breneman, David W. 1994 *Liberal Arts Colleges—Thriving, Surviving or Endangered?* Washington, D.C.: Brookings Institution Press.
- Brint, Steven 1994 *In an Age of Experts: The Changing Role of Professionals in Politics and Public Life*. Princeton, N.J.: Princeton University Press.
- , and Jerome Karabel 1989 *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity, 1900–1985*. New York: Oxford University Press.
- Clark, Burton R. 1965 *The Higher Education System*. Berkeley: University of California Press.
- 1998 *Creating Entrepreneurial Universities: Organizational Pathways of Transformation*. New York: Pergamon.
- Cohen, Wesley M., et al. 1998 “Industry and the Academy: Uneasy Partners in the Cause of Technological Advance.” In Roger G. Noll, ed., *Challenges to Research Universities*. Washington, D.C.: Brookings Institution Press.
- Collins, Randall 1979 *The Credential Society*. New York: Academic Press.
- Kerr, Clark 1963 *The Uses of the University*. New York: Harper and Row.
- Marrou, Henri [1948] 1982 *A History of Education in Antiquity*. Madison: University of Wisconsin Press.
- McClelland, Charles E. 1980 *State, Society, and Univeristy in Germany, 1700–1914*. Cambridge, U.K.: Cambridge University Press.
- Meyer, John W., Francisco O. Ramirez, Richard Rubinson, and John Boli-Bennett 1979 “The World Education Revolution, 1950–1970.” In John W. Meyer and Michael T. Hannan, eds., *National Development and the World System*. Chicago: University of Chicago Press.
- Rashdall, Hastings [1895] 1936 *The Universities of Europe in the Middle Ages*, ed. F. M. Powicke and A. B. Emden, vol. 1. Oxford: Oxford University Press.

Rhoades, Gary 1998 *Managed Professionals: Unionized Faculty and Restructuring of Academic Labor*. Albany, N.Y.: SUNY Press.

Rubin, Michael R., and Mary Taylor Huber 1986 *The Knowledge Industry in the United States, 1960–1980*. Princeton, N.J.: Princeton University Press.

STEVEN BRINT

## HISPANIC AMERICANS

Despite their common linguistic heritage, Hispanic Americans are a heterogeneous and rapidly growing population that includes no less than twenty-three distinct national identities and combines recent legal and undocumented immigrants with groups whose ancestors predate the formation of the United States as we know it today. The label *Hispanic* is derived from *Hispania*, the Latin word for Iberia. In 1973 the U.S. Department of Health, Education and Welfare adopted the term “Hispanic” at the recommendation of the Task Force on Racial/Ethnic Categories to designate U.S. residents who trace their origins to a Spanish-speaking country. Following suit, the U.S. Census Bureau adopted this label as a statistical shorthand for the Hispanic national-origin groups (del Pinal and Singer 1997; Haverluk 1997). Originating in the western United States, the term “Latino” has been adopted as an alternative by groups that view “Hispanic” as a conservative pan-ethnic label imposed by the government that ignores their political and economic struggles for equality and representation. These distinctions notwithstanding, both labels serve as umbrellas for a highly diverse segment of the U.S. population.

Hispanics are one of the fastest growing segments of the U.S. population. High levels of immigration combined with high fertility rates yield a growth rate for Hispanics that is seven times that of the non-Hispanic population (U.S. Department of Commerce 1993). In 1990 the U.S. Census Bureau enumerated 22.4 million Hispanics, representing 9 percent of the aggregate population, but the 1997 population estimate reached 29.7 million, accounting for 11 percent of the national total (U.S. Department of Commerce 1998). Hispanics accounted for about one-third of national population growth during the 1980s, and their

contribution to aggregate demographic growth is expected to increase in the future. Census Bureau projections made in 1995 predicted that by the year 2000 the Hispanic population would reach 31.4 million, but this estimate is conservative because annual estimates since that time have consistently been exceeded (U.S. Department of Commerce 1996). Hispanics are projected to surpass blacks as the largest minority by 2003—perhaps sooner, depending on the volume of legal and undocumented immigration from Central and South America and the Spanish-speaking Caribbean. Already in 1998, Hispanic children outnumbered black children.

Although immigration has figured prominently in the growth of the Hispanic population since 1941, its influence on demographic growth, ethnic diversification and renewal, and population replenishment has been especially pronounced during the 1980s and 1990s. Immigration was responsible for approximately one-third of the phenomenal growth of the Hispanic population in the 1980s and 1990s. At the end of the 1990s, two-thirds of the population were immigrants or children of immigrants (del Pinal and Singer 1997), and trends in fertility and immigration suggested higher growth of the Hispanic population well into the twenty-first century. By the year 2020, the U.S. Hispanic population is projected to reach 52.6 million, representing approximately 16 percent of the national total (U.S. Department of Commerce 1996).

Nearly two-thirds of all U.S. Hispanics (64 percent) are of Mexican origin, while 11 percent trace their origins to Puerto Rico, 4 percent to Cuba, and 14 percent to other Central and South American nations. An additional 7 percent of Hispanics are of unspecified national origin, which includes mixed Spanish-speaking nationalities, Spaniards, and “Hispanos,” the descendants of the original Spanish settlers in what came to be known as Colorado and New Mexico. This national-origin profile of the Hispanic population has evolved since 1970 because of the differential growth of selected groups. In particular, since 1970 the Mexican, Central American, and South American population shares have increased, while the Cuban, Puerto Rican, and other Hispanic shares have declined (Bean and Tienda 1987; del Pinal and Singer 1997). Differential growth rates derive from rising immigration flows combined with high fertility among the foreign-born.

Defining features of the U.S. Hispanic population, in addition to vigorous growth, include increasing diversity, segmented integration, and rising political influence. Among the more salient aspects of diversity are immigrant and generational status, national origin, residential distribution, socioeconomic status, political participation, and reproductive behavior. Segmented integration is evident in the emergence of a solid middle class coupled with rising poverty, especially among children and immigrants, and greater geographic dispersion coupled with persisting residential concentration—nationally in a few large metropolitan areas and locally in ethnic neighborhoods within major cities (Haverluk 1997). As the number of Hispanics increases, politics also becomes increasingly important in determining their social and economic destiny. Greater political influence is already evident in the growing presence of Hispanics among elected officials. However, Hispanic voter turnout remains exceedingly low, suggesting that the potential political impacts of population growth have not fully unfolded.

The continued rapid growth of U.S. Hispanic population coupled with increased diversification raises concerns about long-term prospects for their social integration, particularly recent arrivals from Mexico, Central and South America, and the Spanish-speaking Caribbean. Among casual observers who note the emergence of ethnic neighborhoods where Spanish is spoken publicly as well as privately, the growing Hispanic presence has raised fears about the potential “Latinization” or “Hispanicization” of the United States. However, larger numbers are not likely to be the most decisive force shaping the Hispanic imprint on the U.S. social and economic landscape, for size automatically confers neither power nor status. Therefore, to evaluate the recent past and future imprint of Hispanics in the United States, what follows summarizes several themes that have developed in the sociological literature on Hispanic Americans. These include: (1) the origin of current ethnic labels, (2) the roots of diversification, (3) the changing social, demographic, and economic composition of the population, and (4) the implications for societal integration of recent social and economic trends. A concluding section summarizes key lessons from existing studies and identifies areas for future investigation.

## EVOLVING LABELS

Ethnic labels are partly imposed by the host society and partly chosen by immigrant groups who wish to preserve their national identity. The labels “Spanish origin” and “Hispanic” originally were coined as terms of convenience for official reporting purposes. Before the mid-1960s, Hispanics were unfamiliar to most observers outside the Southwest, where persons of Mexican ancestry were well represented, and the Northeast, where Puerto Rican communities began to flourish after World War II. Therefore, until 1960 the “Spanish surname” concept was adequate for identifying persons of Mexican origin residing in the Southwest, and “Puerto Rican stock” was used to identify persons who resided in the Northeast (predominantly New York) and who were born in or whose parents were born in Puerto Rico. However, with increasing intermarriage, residential dispersion, and generational succession, these concepts became progressively less viable to identify persons from Mexico and Puerto Rico. Furthermore, the influx of immigrants from Central and South America and the Spanish-speaking Caribbean into areas traditionally inhabited by Mexicans and Puerto Ricans necessitated labels that could more adequately represent the growing diversity along national-origin lines.

In recognition of the growing residential, marital, and generational heterogeneity of the Spanish-speaking population, in 1970 the U.S. Census Bureau adopted the “Spanish origin” concept, which was based on self-identification and could be administered to the U.S. population on a national level (Bean and Tienda 1987). Symbolically, this decision, which was also an important political gesture, recognized Hispanics as a national minority group rather than as regionally distributed subgroups. And in 1980 the term “Hispanic” accompanied the “Spanish origin” item on the census schedule to identify persons from Latin America, Spain, and the Spanish-speaking Caribbean. However, through the 1980s the term “Latino” came into popular use as an alternative to “Hispanic.” As a symbol of self-determination and self-definition, “Latino” is the label preferred by many ethnic scholars.

Despite their popular use and administrative legitimacy, pan-ethnic terms such as “Latino” or “Hispanic” are less desirable than specific national-origin designations, such as Puerto Rican,

Venezuelan, Cuban, or Mexican, which better reveal the appreciable socioeconomic and generational diversity of Hispanic Americans. Not surprisingly, diversity along national-origin lines became a major intellectual theme in the scholarly writings about Hispanics during the mid- to late 1980s and remained so at the turn of the century as the powerful forces of international and internal migration continued to diversify the composition of this rapidly growing population.

### ROOTS OF DIVERSIFICATION

Social science interest in Hispanic Americans has increased greatly since 1960, and the scope of topics investigated has expanded accordingly. Whereas studies conducted during the 1960s and through the mid-1970s tended to focus on regionally localized populations, the 1980s witnessed a proliferation of designs that compare group experiences. This shift in the research agenda, which coincided with the designation of Hispanics as a national population, also brought into focus the theme of diversity and inequality among national-origin groups. More recent studies have focused on economic opportunities afforded by geographic dispersal and the changing fortunes of the burgeoning second generation—especially the children of recent immigrants.

The Hispanic presence in the United States predates the formation of the nation as we currently know it. Over a decade before the Pilgrims landed at Plymouth Rock, Spaniards had already settled in present-day New Mexico. Several subsequent events shaped the Hispanic imprint on the United States. These include (1) the Louisiana Purchase, which provided a powerful impetus for westward expansion during the 1800s; (2) the Treaty of Guadalupe Hidalgo, which ended the Mexican-American War and ceded all Mexican territories north of the Rio Grande to the United States; (3) the Mexican Revolution, which propelled thousands of Mexicans north to seek refuge from the bloody conflict; (4) the Cuban Revolution, which provided the impetus for several unprecedented refugee flows; (5) the political instability in Central America during the 1980s, which has augmented legal and undocumented migrant streams from El Salvador, Guatemala, and Nicaragua; and (6) the economic crisis of the 1980s, which encouraged higher levels of documented

and undocumented migration from South America, especially Colombia, Peru, and Ecuador (U.S. Department of Commerce 1993). International events in Latin America continue to influence the U.S. Hispanic population by altering immigration streams and their reception as refugees or labor migrants.

The uneven integration experiences of the various national-origin groups are as powerful in diversifying the U.S. Hispanic population as are migration streams of largely unskilled workers from Latin America. The cultural and socioeconomic diversity of Hispanic Americans can be traced partly to the diverse modes of their incorporation into the United States and partly to the changing opportunities to become Americanized. Nelson and Tienda (1985) proposed a framework for conceptualizing the emergence, consolidation, and persistence of distinct Hispanic ethnic groups. They identified three domains of immigrant incorporation that are pertinent for understanding the socioeconomic stratification of Hispanics: (1) the mode of entry, namely, the conditions of migration as voluntary labor or as political migrants; (2) the mode of integration, that is, the climate of reception at the time of mass entrance to the host society; and (3) the circumstances that precipitate reaffirmation of national origin. The latter emphasizes the distinction between the cultural or symbolic content of Hispanic origin and the economic consequences of ethnicity that result in the formation of minority groups. This distinction between the economic and the cultural underpinnings of Hispanic ethnicity is pertinent for theorizing about the long-term integration prospects of specific nationality groups.

Along these three domains of ethnic incorporation, the major national-origin groups exhibit considerable diversity. For example, the origin of the Mexican and Puerto Rican communities can be traced to annexation, although the timing and particulars of the two cases were quite distinct. The annexation of Mexican territory resulted from a political settlement subsequent to military struggle and was followed by massive and voluntary wage-labor migration throughout the twentieth century, but particularly after 1960. The Puerto Rican annexation, which was formalized at the culmination of the Spanish-American War, will remain an incomplete process until statehood or

independence is achieved. However, like the Mexican experience, the Puerto Rican presence on the U.S. mainland was characterized by a massive wage-labor flow after World War II that has ebbed and flowed according to economic conditions on both the island and the mainland.

Mexico has been the leading source country for Latin American immigrants since 1820, but the Mexican flow began in earnest during the time of the Mexican Revolution. The Bracero Program, which was a binational agreement that allowed entry of temporary agricultural workers, institutionalized migrant streams that persisted long after the formal agreement was terminated in 1964. Also, Mexico is the leading source of undocumented migrants to the United States. Although 2.3 million Mexicans obtained legal status under the provisions of the Immigration Reform and Control Act (IRCA) of 1986, undocumented migrant workers continue to enter the United States and are currently settling in such nontraditional areas as North Carolina, Atlanta, New Jersey, and New York.

These distinct modes of incorporation are sharpened by the experience of Cubans, whose socioeconomic success is as striking as the limited socioeconomic achievements of Puerto Ricans and Mexicans. Although the U.S. Cuban community was established before Fidel Castro's rise to power, Cuba's internal politics are largely responsible for the dramatic growth of the Cuban population in the United States since 1960. The Cuban Revolution created a wave of U.S.-bound political refugees who were themselves differentiated by social classes. The so-called golden exile cohort, which virtually gutted the Cuban middle class, was followed by the exodus of skilled and semiskilled workers who made up the vast majority of Cuban emigrés. Although the distinction between political and economic migrants is murky, political refugees, unlike wage-labor migrants, usually command immediate acceptance from the host society. The Cuban experience, however, stands in sharp contrast to that of later political refugees from Central America (primarily El Salvador, Nicaragua, and Guatemala), whose refugee story is one of clandestine entry and extended legal and political struggles for recognition. Probably this difference reflects the fact that Cubans were fleeing a communist regime while this was not the case for Central American refugees.

Immigration from Central and South America and the Spanish-speaking Caribbean has continued to diversify not only the national-origin composition of the Hispanic population but also the socioeconomic position of the various groups. The Dominican Republic, a country of only 8 million inhabitants, was the fourth-largest source of U.S. immigrants in 1995, including many who entered without documentation (del Pinal and Singer 1997). Although internal political problems precipitated Dominican emigration after 1960, declining economic conditions fueled the migrant stream during the 1970s. Likewise, the Central American wage labor-flow has been propelled by poor economic conditions coupled with political upheavals, especially during the 1980s and 1990s. That more than two-thirds of Central Americans residing in the United States in 1995 entered after 1980 testifies to the recency and intensity of this wage-labor flow. South American immigration to the United States derives mainly from Colombia, but there is a growing presence of Ecuadorians and Argentineans among this stream. Compared to Mexican, Central American, and Dominican flows, undocumented migrants are less common among South American immigrants.

The future role of immigration in stratifying the Hispanic population is highly uncertain, as it depends both on changes in U.S. foreign policy toward Central and South America and on revisions in immigration policy concerning the disposition of undocumented aliens and quotas on admissions of close family members of legal residents. Equally important is the role of expanded social networks in drawing new migrants to U.S. shores. Trends in the 1990s indicate that these flows are likely to continue into the foreseeable future. A comparison of wage-labor migration histories of Hispanics clearly illustrates how diverse modes of entry and integration have fueled the diversification of Hispanic Americans.

Ethnic reaffirmation and consolidation can best be understood as immigrant minority communities define themselves vis-à-vis the host society. For groups that were relatively successful in adapting to the host society, such as Cubans, national heritage acquired a highly symbolic character that is used for economic relationships when expedient and downplayed otherwise. Alternatively, class position and national origin become inextricably linked when immigrants are destined for



the lower ranks of the social hierarchy, as seems to have occurred for Mexicans, Puerto Ricans, and Central Americans, or when illegal status forces many underground. Thus, the distinction between symbolic ethnicity and minority status revolves around the degree of choice groups have in controlling their socioeconomic destiny (Vincent 1974).

### **SOCIAL, DEMOGRAPHIC, AND ECONOMIC TRENDS**

Until the 1960s, the U.S. Hispanic population was overwhelmingly Mexican and almost exclusively located in the Southwest (Haverluk 1997). But migration from Puerto Rico to the Northeast during the 1950s and the arrival of thousands of Cubans in south Florida and the Northeast following the 1959 Cuban Revolution led to the establishment of distinct Hispanic communities in Miami and New York City. Heavy migration from Mexico and other Latin American countries since the 1970s has altered the regional landscape and created a more geographically and socially heterogeneous Hispanic population whose imprint reaches well beyond the traditional states of residence.

Although Hispanics reside in all states, historically the population has been concentrated in nine states, where 85 percent of Latinos reside. These include the five southwestern states of California, Texas, Arizona, New Mexico, and Colorado, as well as New York, Florida, Illinois, and New Jersey. Two states—California and Texas—were home to more than half of all Hispanics in 1995, with Mexicans and Central Americans disproportionately concentrated in these states. Despite residential dispersion of Hispanics during the 1980s and 1990s, the visibility of distinct Hispanic communities is reinforced by persisting regional concentration along national origin lines. Puerto Ricans remain concentrated in the Northeast—predominantly in major cities of New York and New Jersey—while Cubans have become bimodally distributed in south Florida and large northeastern cities (Bean and Tienda 1987). Mexicans remain disproportionately concentrated in the major cities of the Southwest and Chicago.

Cities with large Hispanic populations witnessed the greatest growth during the 1990s (Frey 1998). The ten metro areas with the largest Hispanic populations experienced the largest population gains. These include Miami, New York City,

and Chicago as well as others close to the Mexican border. In San Antonio, Miami, and El Paso, Hispanics are the majority population. In 1996, more than 6 million Hispanics resided in Los Angeles, making it the second largest Latino city in the world (behind Mexico City). Residential segregation further compounds geographic concentration by spatially isolating Hispanics of low socioeconomic status and recent immigrants from non-Hispanic whites. This segregation reinforces the cultural distinctiveness of Hispanics in regions where visible communities have been established even as increased residential dispersion concurrently fosters national integration. It is in this sense—increasing geographic dispersion combined with persisting concentration—that Hispanics experience segmented residential integration.

The demography of the Hispanic population helps in understanding other characteristics, such as educational standing and economic well-being. For instance, the high immigrant composition of the population means that many have not had the opportunity to acquire a U.S. education, or much education at all. The youthful age structure also means that Hispanic school enrollment rates are higher and retirement rates lower compared to other population groups.

A social and economic profile of the Hispanic population since 1960 provides signals of both optimism and pessimism for the long-term economic prospects of distinct nationality groups. Because education is key for economic and social mobility, as well as for integration of new arrivals, trends in educational attainment are quite revealing about the diverse futures facing Hispanics. On the one hand, from 1960 to 1996 Hispanics witnessed an unmistakable improvement in their educational attainment. Whereas only 30 percent of Hispanics aged 25 and over had completed a high school education or more in 1970, by 1996 more than 53 percent had done so (del Pinal and Singer 1997; U.S. Department of Commerce 1993). However, there have been very modest gains in educational progress since 1980. More disturbing is evidence that educational gaps between Hispanics and non-Hispanics have widened because gains of the latter have been faster. These widened gaps are evident at all levels of educational attainment, beginning in preschool. Because Hispanics average less pre-school experience, they begin elementary school with fewer social skills. Furthermore,

high school noncompletion rates remain distressingly high for Mexicans and Puerto Ricans—the two groups with the longest history in the United States.

That Cubans outperform Mexicans and Puerto Ricans in educational attainment undermines simplistic explanations that immigration and language are the reasons for the continued educational underachievement of the latter. The continued arrival of new immigrants with very low levels of education also contributes to the slow educational progress of Hispanics. However, this is only a partial explanation because even among the native-born, Hispanics lag behind African Americans and non-Hispanic whites in educational attainment. Low socioeconomic status and especially the preponderance of parents with low educational attainment are major forces driving the persisting educational underachievement of Mexicans and Puerto Ricans (Bean and Tienda 1987). Most research shows that Hispanics with family backgrounds comparable to those of non-Hispanic whites are less likely to withdraw from high school and as likely to continue on to college. Unfortunately, the current educational distribution implies intergenerational perpetuation of educational disadvantages well into the future.

These educational disadvantages carry over into the labor market, yet Hispanics have higher age-specific labor-force participation rates than most groups. This labor-market advantage exists largely because immigrants have very high rates of labor-force participation and they comprise a large and growing share of the working-age Hispanic population (Bean and Tienda 1987). However, relatively low skills and limited proficiency in English channel Hispanic immigrants to low-wage jobs with few benefits. Yet because Hispanic immigrants themselves are highly diverse along national origin lines, group experiences in the U.S. labor market are equally varied.

The 1960s and 1970s witnessed a deterioration of the labor-market standing of Puerto Rican men and women, while Cubans became virtually indistinguishable from non-Hispanic whites in terms of participation rates, unemployment rates, and occupational profiles. Mexicans stand somewhere between these extremes, with greater success than Puerto Ricans in securing employment, but usually at low wages owing to their low levels of

human capital. Therefore, Mexicans are more highly represented among the working poor than are non-Hispanic whites, but Puerto Ricans are more likely than other Hispanic groups to join the ranks of the nonworking poor. Puerto Rican men witnessed unusually high rates of labor-force withdrawal during the 1970s, and their labor-force behavior converged with that of economically disadvantaged blacks (Tienda, 1989).

Numerous explanations have been provided for the deteriorating labor-market position of Puerto Ricans. These include the unusually sharp decline of manufacturing jobs in the Northeast; the decline of union jobs in which Puerto Ricans traditionally concentrated; increased labor-market competition with Colombian and Dominican immigrants; and the placement of Puerto Rican workers at the bottom of a labor queue. Direct empirical tests of these working hypotheses—all with merit—have not been forthcoming. However, the dramatic improvement in the labor-market standing of Puerto Ricans during the 1990s as a result of tight labor markets and vigorous economic growth suggest that Puerto Ricans are, like blacks, located at the bottom of a hiring queue and hence more vulnerable to dislocation in accordance with the business cycle.

Less debatable than the causes of Hispanics' labor-market status are the associated economic consequences. The median family incomes of Hispanic families place them below white and above African-American families (Bean and Tienda 1987). Although this socioeconomic ranking by groups has been in place since the 1970s, in 1995 Hispanic median family income actually fell below that of African Americans (del Pinal and Singer 1997). However, the emergence of a solid middle class supports the view of segmented economic integration wherein some groups experience upward mobility while others do not.

Trends in poverty rates also indicate that Hispanics are losing ground relative to African Americans. Not only have Hispanic poverty rates risen over time, but they appear resistant to improvement when the economy rebounds. Although black poverty rates have historically exceeded Hispanic poverty rates, this situation changed during the 1990s, when black poverty dropped faster than Hispanic poverty. Thus, by 1995 black and Hispanic family poverty rates had converged, and for

some population subgroups (e.g., single-mother families and children), Hispanic poverty rates exceeded those of blacks for the first time.

Poverty rates are higher among immigrants compared to native-born Hispanics; and, among specific national origin groups, Mexicans and Puerto Ricans experience the highest risk of poverty. In 1995 more than one in three Puerto Rican families had incomes below the poverty level, compared to 22 to 28 percent for the other Hispanic groups. Even Cubans, the most highly educated of the Hispanic groups, experienced poverty rates double those of non-Hispanic whites in 1995—16 versus 6 percent, respectively (del Pinal and Singer 1997). Puerto Rican families headed by single women have the highest poverty rates, as nearly two of every three such families were poor in 1995 compared to only 30 percent of Cuban and 50 percent of Mexican mother-only families. In part, these differences among national origin groups reflect the groups' nativity composition. That is, because recent arrivals have relatively low skills, even their high rates of labor-force participation cannot shield them from poverty wages. A working poverty explanation also is consistent with the failure of Hispanic poverty to rebound during periods of economic recovery. Puerto Ricans are a possible exception, inasmuch as their poverty derives more from nonwork than that of Mexicans or Central/South Americans.

### INTEGRATION PROSPECTS

The changing demography of Hispanics has direct implications for the integration of Hispanic immigrant minority groups and raises questions about assimilation prospects. The rapid growth and residential concentration of the groups, coupled with the rising salience of immigration as a component of demographic growth, revitalizes Hispanic cultures and fosters maintenance of Spanish. That Hispanics have some of the lowest naturalization rates of all immigrant groups hinders their political participation and, for some observers, raises questions about their commitment to becoming American.

Puerto Ricans are U.S. citizens by birth and enjoy most of the privileges that citizenship confers, while Mexican and other Latin American immigrants who enter legally must wait at least five

years to apply for citizenship. Those who enter illegally have an uncertain status in the United States as long as they remain undocumented. That Puerto Ricans have been less successful economically than Mexicans raises questions about the significance of citizenship as a requisite for socioeconomic integration. The newest wage-labor migrants—Colombians, Dominicans, and Central Americans—appear to fare better than Puerto Ricans, posing yet another challenge to conventional understandings about immigrant assimilation processes. However, it is too early to determine their long-term prospects, which likely will depend more on changes in labor market opportunities and the educational achievement of the second generation than on processes of discrimination and social exclusion. Furthermore, the growing number of undocumented migrants among these groups could undermine their social and political leverage over the short to medium term.

Although it is difficult to predict the long-term integration prospects of any group, the diversity of the Hispanic experience complicates this task further because of the uncertain future of immigration and the economy, and because the political participation of Hispanics has traditionally been low. While the spatial concentration of Hispanics allows native languages to flourish and under some circumstances promotes ethnic cohesion, some observers interpret the rise and proliferation of ethnic neighborhoods as evidence of limited integration prospects, irrespective of whether the segregation is voluntary or involuntary. An alternative interpretation of high levels of residential segregation among Hispanics, particularly recent arrivals, is that of ethnic resilience. This perspective maintains that in the face of inter-ethnic tensions and economic adversity, individuals rely on their ethnic compatriots for social supports and hence promote solidarity along ethnic lines. One implication of this view is that ethnic resilience is the *consequence* rather than the *cause* of unequal integration prospects—that Hispanics' tendency toward elaborate ethnic ties reflects their tentative acceptance by the dominant society. As such, ethnic traits become enduring rather than transitional features of Hispanic neighborhoods.

A similar debate over the integration prospects of Hispanics clouds the issue of Spanish-language retention, which is politically significant

because it provides a ready target for policies designed to assimilate linguistically diverse populations and because Spanish-language retention is a ready scapegoat for the poor educational achievement of Hispanics. Spanish is the second most common language spoken at home in the United States, but it is also a limiting attribute because two of five Hispanics reported that they did not speak English well or at all in 1990 (del Pinal and Singer 1997). Although this circumstance could prohibit labor-market integration, this is not necessarily the case in ethnic neighborhoods where virtually all business transactions can be conducted in Spanish. According to Lopez (1996), in Los Angeles, Spanish has become the *lingua franca* of the city's working class. As such, the public use of Spanish symbolizes and reinforces the ethnic stratification of the Hispanic population, particularly in areas of high residential concentration. Public use of Spanish by the poorer segments of the Hispanic population not only reinforces stereotypes about poverty and immigration but also reinforces beliefs about the limited integration prospects of Hispanics.

Furthermore, considerable controversy about the socioeconomic consequences of bilingualism exists, yet the preponderance of research shows that Spanish retention and bilingualism per se are not the sources of Hispanic underachievement. Rather, the failure to acquire proficiency in English is the principal culprit (Bean and Tienda 1987). In support of the distinction between bilingualism and lack of proficiency in English, there is some evidence that bilingualism may be an asset, albeit only among the middle classes who are able to convert this skill into social and financial resources (Tienda 1982). The failure of many Hispanics to achieve proficiency in English certainly limits their economic opportunities, but it is facile to equate lack of proficiency in English with bilingualism, which does not preclude proficiency in English. Although Spanish maintenance among Hispanics is more pervasive than is the maintenance of Asian languages among respective national origin groups, language shifts clearly indicate linguistic assimilation toward English dominance (Lopez 1996). Bilingualism is pervasive among second-generation Hispanics, but by the third generation bilingualism is less common than English monolingualism (Lopez 1997b). In short, language trends indicate that Hispanics are becoming assimilated to U.S. society and culture.

Finally, the long-term integration prospects of Hispanics depend on whether and how groups from specific localities or national-origin groups mobilize themselves to serve ethnic interests. Despite the convenience of the pan-ethnic label for statistical reporting purposes, Hispanics do not represent a unified political force nationally, and the fragility of their coalitions is easily undermined by demographic, economic, and social diversification. Hispanic populations of all ages and citizenship statuses affect routine electoral politics and contribute to the electoral power of registered Hispanic voters by the weight of numbers, which is the basis of redistricting. However, greater numbers are insufficient to guarantee increased representation, particularly for a population with a large number of persons under voting age, low rates of naturalization and voter registration, and, among registered voters, a dismal record of turning out at the polls. Most important is the low adult citizenship rate of 54 percent, compared to 95 percent for non-Hispanic whites and 92 percent for blacks (Rosenfeld 1998). That voter turnout also varies by national origin further divides the fragile Hispanic political alliances, such as the Hispanic congressional caucus, by splitting the Hispanic vote along class and party lines.

In the 1992 U.S. elections, Mexican Americans cast only about 16 votes per 100 persons compared to 50 for every 100 non-Hispanic whites. Cubans cast 29 votes for every 100 persons in the 1992 election; and despite their high citizenship rate (93 percent), Puerto Ricans cast only 24 votes for every 100 residents on the U.S. mainland. Other Hispanics, with an adult citizenship rate of 43 percent, cast 17 votes for every 100 residents in 1992. Despite low rates of voting relative to population shares, Hispanics have increased their representation at all levels of political office. Rosenfeld (1998) analyzed changes in Hispanic elected officials in the five southwest states from 1960 to 1995, a period that captures the entire trajectory of Hispanic electoralism in all states except New Mexico. The analysis considered the universe of elected officials in each state, including statewide and legislative offices as well as U.S. House and Senate seats relative to the voting-age population. Empirical trends indicate the emergence of Hispanic electoralism, which he characterizes as the achievement of a steady and significant level of

Hispanic representation throughout the political system.

Thus, political behavior also points to greater political integration, albeit slow and with some uncertainty for the future. This is because most of the gains in political representation and Hispanic electoralism occurred prior to 1980, and the rapidly changing demographic and socioeconomic profile is less predictable. The political status of Hispanics underwent rapid changes during the 1990s, and the eventual outcomes are far from certain. Low rates of naturalization are beginning to reverse. For Mexicans, this has been facilitated by changes in Mexican law that permit dual citizenship in Mexico and the United States. Furthermore, legislative changes that tie access to welfare benefits to citizenship status have also encouraged many immigrants to naturalize. Finally, the higher priority given to citizens in sponsoring relatives for admission to the United States also intensified demand for naturalization. However, there persist large differentials by national origin (del Pinal and Singer 1997), with Mexicans lagging far behind Cubans and Central/South Americans.

### CONCLUSIONS

Several lessons can be culled from the scholarly literature and public discourse about the social and economic future of Hispanic Americans. One is that generic labels, such as "Hispanic" and "Latino," are not useful for portraying the heterogeneous socioeconomic integration experiences of specific national-origin groups. A second major lesson is that the evolving differentials in economic standing of Hispanic national origin groups are rooted in the distinct *modes of incorporation* of each group, which in turn have profound implications for short- and long-term integration prospects. A third major lesson, which is related to the second, is that the socioeconomic imprint of Hispanic Americans will be as varied as the population itself. Changes in immigrant composition and residential segregation will play decisive parts in determining how Hispanics shape the ethnic landscape of the United States into the twenty-first century.

The nagging question is: Why does there persist a close association between Hispanic national origin and low social standing? On this matter there is much debate, but both sides accord great emphasis to the role of immigration in deciding

the socioeconomic destiny of Hispanic Americans. One interpretation—the replenishment argument—emphasizes how immigration continues to diversify the composition of the population by introducing new arrivals on the lower steps of the social escalator, even as earlier arrivals experience gradual improvements in their economic and social statuses. The recency of immigration of a large share of the Hispanic population has direct implications for other demographic, social, and economic characteristics. They are younger, and less proficient in English, and many cannot vote because they are not citizens. Immigrants are also more likely than native-born Hispanics to be working class and to have lower education levels. Consistent with the predictions of classical assimilation theory, the replenishment hypothesis implies that observed differences in socioeconomic standing among Hispanics will disappear with time, irrespective of country of origin or period of arrival. Lending support to this prediction is a growing body of evidence showing that later arrivals fare better in the U.S. labor market and social institutions than do earlier arrivals.

Despite some compelling aspects of the replenishment argument, it falls short of accounting for the limited social mobility of Mexicans and Puerto Ricans, the two groups with the longest exposure to U.S. institutions and traditions. These experiences challenge the immigrant-replenishment hypothesis and place greater emphasis on the complex set of circumstances that define distinct modes of incorporation for the major national-origin groups and subsequent entry cohorts. The structural interpretation emphasizes the role of unique historical circumstances in which each national origin group established its presence in the United States and acknowledges that social opportunities depend greatly on the state of the economy and public receptiveness toward new arrivals.

Unfortunately, it is too early to evaluate the relative merits of the two hypotheses, especially in the absence of the longitudinal data required to trace socioeconomic trajectories of successive generations. That the future of immigration (its volume, source countries, and composition) is highly indeterminate further aggravates the difficulties of assessing these hypotheses. Of course, the greatest uncertainty about immigration concerns the volume and sources of illegal migrants. Finally, the socioeconomic fate of Hispanic Americans as a

whole and as separate national origin groups will also depend on the extent to which Hispanic elected and appointed officials use ethnicity as a criterion for defining their political agendas. The early decades of the twenty-first century will be pivotal in resolving these uncertainties.

(SEE ALSO: *Discrimination; Ethnicity*)

REFERENCES

- Bean, Frank D., and Marta Tienda 1987 *The Hispanic Population of the United States*. New York: Russell Sage Foundation.
- del Pinal, Jorge, and Audrey Singer 1997 "Generations of Diversity: Latinos in the United States." *Population Bulletin* 52(3):2-48.
- Frey, William H. 1998 "The Diversity Myth." *American Demographics* (June):39-43.
- Haverluk, Terrence 1997 "The Changing Geography of U.S. Hispanics, 1850-1990." *Journal of Geography* 96(3):134-145.
- Lopez, David E. 1996 "Language: Diversity and Assimilation." In Roger Waldinger and Mehdi Bozorgmehr, eds., *Ethnic Los Angeles*. New York: Russell Sage Foundation.
- Nelson, Candace, and Marta Tienda 1985 "The Structuring of Hispanic Ethnicity: Historical and Contemporary Perspectives." *Ethnic and Racial Studies* 8:49-74.
- Rosenfeld, Michael 1998 "Mexican Immigrants and Mexican American Political Assimilation." In *Migration Between Mexico and the United States: Binational Study*, vol. 3. Washington, D.C.: U.S. Commission on Immigration Reform.
- Tienda, Marta 1982 "Sex, Ethnicity, and Chicano Status Attainment." *International Migration Review* 16:435-472.
- 1989 "Puerto Ricans and the Underclass Debate." *Annals of the American Academy of Political and Social Sciences* 501:105-119.
- U.S. Department of Commerce 1993 *We the American Hispanics*. Washington, D.C.: Bureau of the Census, Economics and Statistics Administration.
- 1996 "Population Projections of the United States by Age, Sex, Race and Hispanic Origin: 1995-2050." *Current Population Reports*, P25-1130. Washington, D.C.: Bureau of the Census, Economics and Statistics Administration.
- 1998 *The Hispanic Population in the United States: March, 1997. Detailed Tables for Current Population Reports P20-511*. Washington D.C.: Bureau of the Census, Ethnic and Hispanic Statistics Branch.

Vincent, Joan 1974 "The Structuring of Ethnicity." *Human Organization* 33:375-379.

MARTA TIENDA

**HISTOGRAMS**

See Statistical Graphics.

**HISTORICAL ANALYSIS**

See Comparative-Historical Analysis; Event-History Analysis; Historical Sociology.

**HISTORICAL SOCIOLOGY**

Used as a category to identify social scientific research that constructs or illustrates theory by careful attention to culturally, geographically, and temporally located facts, historical sociology (or, from the historian's vantage point, sociological history) exists as a self-conscious research orientation within both of its parent disciplines. Popular assertions in the 1950s and 1960s that history involved the study of particular facts while sociology involved the formulation of general hypotheses (see Franzosi and Mohr 1997, pp. 133-139 [for the historians' view]; Lipset 1968, pp. 22-23; McDonald 1996; cf. Mills 1959, pp. 143-164) had turned around by the early 1980s, at which time Abrams claimed that "sociological explanation is necessarily historical" (Abrams 1982, p. 2; see Burke 1980, p. 28). While this integrative interpretation of the two disciplines reflects the attitudes of many researchers in the respective disciplines, a few persons remain extremely cautious about the merits of integrating them (see Ferrarotti 1997, pp. 12-13).

Abrams's claim is true for much of sociological theorizing, and sociologists realize that seminal research in their discipline has been informed by careful attention to historical information. Nonetheless, fundamental differences exist between history and sociology regarding the choice of research strategies and methodologies. Historical research emphasizes the sociocultural context of events and actors within the broad range of human culture, and when examining events that occurred in early periods of the human record it borrows from archaeology and cultural anthropology, two

companion disciplines. Historians, therefore, who examine premodern material often borrow anthropological rather than sociological insights to elucidate areas where the historical record is weak, under the assumption that preindustrialized societies share basic similarities that sociological theory rarely addresses (Thomas 1963, 1971; cf. Thompson 1972). Historians are likely to choose research topics that are culturally and temporally delimited and that emerge “from the logic of events of a given place and period” (Smelser 1968, p. 35; see Bonnell 1980, p. 159). They tend to supplement secondary sources with primary texts or archival data (see Tilly 1981, p. 12).

In contrast, sociological research stresses theory construction and development, and its heavy emphasis on quantification limits most of its research to issues that affect societies after they begin to modernize or industrialize (and hence develop accurate record keeping that researchers can translate into data [see Burke 1980, p. 22]). Many of their methodological techniques—including surveys, interviews, qualitative fieldwork, questionnaires, various statistical procedures, and social-psychological experimentation—have little if any applicability to historians (Wilson 1971, p. 106). (Interestingly, however, a number of historical sociologists are utilizing narratives to elucidate sociologically traditional issues like capital–state relations and urban development [Gotham and Staples 1996; see Isaac 1997; Kiser 1996.]) Given their orientation toward theory, sociologists are likely to choose research topics that are “rooted in and generated by some conceptual apparatus” (Smelser 1968, p. 35; see Bonnell 1980, p. 159). Their data sources infrequently involve archival searches (see Schwartz 1987, p. 12) or heavy dependence on primary texts. Sociologists seem more willing than historians both to “undertake comparative analysis across national and temporal boundaries” and to present generalizations that relate to either a number of cases or universal phenomena (as opposed to a single case [Bonnell 1980, p. 159]). Reflecting these basic differences, social history, which developed out of the historical discipline, concentrates on speaking about lived experiences, while traditional historical sociologists concentrate on analysing structural transformations (Skocpol 1987, p. 28).

Although some historical sociological studies attempt either to refine concepts or rigorously to

test existing theoretical explanations, more often they attempt “to develop new theories capable of providing more convincing and comprehensive explanations for historical patterns and structures” (Bonnell 1980, p. 161). When testing existing theories, historical sociologists argue deductively (by attempting to locate evidence that supports or refutes theoretical propositions), case-comparatively (by juxtaposing examples from equivalent units), or case-illustratively (by comparing cases to a single theory or concept [Bonnell 1980, pp. 162–167]). Both case comparisons and case illustrations can show either that cases share a common set of “hypothesized causal factors” that adequately explain similar historical outcomes or that they contain crucial differences that lead to divergent historical results (Skocpol 1984, pp. 378–379). In any case, historical sociologists occasionally need reminding that, in both their quantitative and qualitative research, they must avoid the “common tendency . . . to play fast and loose with time” (Jensen 1997, p. 50), since the temporal dimensions of events have profound implications for the validity of comparisons.

Sociology’s founding figures—Marx and Engels, Weber, Tocqueville, and, to a limited degree, Durkheim—utilized history in the formulation of concepts and research agendas that still influence the discipline. In various works Marx and Engels demonstrated adroit sociohistorical skills, particularly in *The Eighteenth Brumaire of Louis Bonaparte* (Marx [1885] 1963) and *The Peasant War in Germany* (Engels [1870] 1926). In these studies they moved deftly among analyses of “short-term, day-to-day phenomena” of sociopolitical life, the underlying structure of that life, and “the level of the social structure as a whole” (Abrams 1982, p. 59) to provide powerful examples of historically grounded materialist analysis (see Abrams 1982, p. 63; Sztompka 1986, p. 325).

Weber, who was steeped in the ancient and medieval history of both the East and West, believed that, through the use of heuristically useful ideal types, researchers could “understand on the one hand the relationships and cultural significance of individual events in their contemporary manifestations and on the other the causes of their being historically *so* and not *otherwise*” (Weber 1949, p. 72). However much contemporary researchers have faulted his *Protestant Ethic and the Spirit Of Capitalism* (1930) for misunderstanding

Puritan religious traditions (Kent 1990; MacKinnon 1988a, 1988b) and interpreting them through preexisting philosophical categories (Kent 1983, 1985), it remains the quintessential example of his historically informed sociological studies (see Marshall 1982). Weber's extensive contribution to many aspects of historical sociology (including the linkage between "agency and structure," multicausality, ideal types, and various methodology and substantive issues) receives extensive attention in Kalberg (1994).

Variouly assessed as a historian and a political scientist, Tocqueville also contributed to historical sociology with books that examined two processes—democratization (in the United States) and political centralization (in France)—that remain standard topics of historical sociological research (Tocqueville [1835] 1969, [1858] 1955; see Poggi 1972; Sztompka 1986, p. 325). Similar praise for historical sensitivity, however, has not always gone to a fellow Frenchman of a later era, Emile Durkheim, whose concepts were scornfully called by one historical sociologist an "early form of ahistoricism" (Sztompka 1986, p. 324). Bellah, nevertheless, asserted that "history was always of central importance in Durkheim's sociological work" (Bellah 1959, p. 153) and even argued that, "at several points Durkheim went so far as to question whether or not sociology and history could in fact be considered two separate disciplines" (Bellah 1959, p. 154). Indeed, recent research has identified Durkheim's cross-disciplinary sensitivity in his analysis of the history of French education (Emirbayer 1996a) and "the structures and processes of civil society" (Emirbayer 1996b, p. 112). Nevertheless, Abrams's compromise interpretation may be most accurate concerning most of Durkheim's work. He acknowledges that Durkheim identified the broad process of the Western transition to industrialization, even though his "extremely general framework" demands specific historical elaboration (Abrams 1982, p. 32).

Despite the prominence of history within major studies by sociology's founding figures, subsequent sociologists produced few historically informed works until the late 1950s (cf. Merton 1938). Also during this period (in 1958) the historian Sylvia Thrupp founded the *Journal of Comparative Studies in Society and History*, and since then other journals have followed that are sympathetic

to historical sociology (including *Journal of Family History*, *Journal of Interdisciplinary History*, *Journal of Social History*, *Labor History*, *Past and Present*, *Social History*, and *Social Science History* [Bart and Frankel 1986, pp. 114–116]). The output of interdisciplinary books continued growing throughout the 1960s, and by the 1970s "the sociological study of history achieved full status within the discipline" (Bonnell 1980, p. 157, see p. 156; see Smith, 1991 [for a clear historical overview]).

Studies of capitalist expansion examine such topics as the emergence and consequences of the Industrial Revolution, the rise of the working class, population growth, and the developmental operations of the modern world system. Exemplary studies include Smelser's *Social Change in the Industrial Revolution* (1959) and Wallerstein's *The Modern World System* (1974). Basing his middle-range model on Parsons's general theory of action, Smelser deduced a supposedly universal sequence through which all changes move that involve structural differentiation in industrializing societies (see Bonnell 1980, p. 162; Skocpol 1984, p. 363). He illustrated the applicability of his framework by drawing examples from the economic changes within the British cotton industry during the nineteenth century, followed by additional examples of changes to the lives and activities of workers in that industry. These historical facts, however, were secondary to the model itself.

Wallerstein borrowed from Marxism and functionalism to devise a "world system" theory of the global economy that purports to be universal in its interpretive and explanatory power. He argued that after the late fifteenth and early sixteenth centuries, a "world economy" developed in which economically advantaged and politically strong areas, called "core states," dominated other economically nondiversified and politically weak "peripheral areas." Through "semi-peripheral areas" that serve as "middle trading groups in an empire," resources flow out of the peripheral areas and into the core states for capitalist development, consumption, and often export back to their areas of origin (Wallerstein 1974, pp. 348–350). Within this model Wallerstein mustered a phalanx of historical facts in order to demonstrate the emergence of the world economy above the limited events in various nation-states, and in doing so he "has promoted serious historical work within sociology" (Tilly 1981, p. 42).



E.P. Thompson's exemplary study (1963) took the sociological concept of "class" and presented its historical unfolding in England between 1780 and passage of the parliamentary Reform Bill in 1832 (Thompson 1963, p. 11). He argued that "the finest-meshed sociological net cannot give us a pure specimen of class . . . . The relationship [of class] must always be embodied in real people and in a real context" (p. 9). By the end of the era that he examined, "a more clearly-defined class consciousness, in the customary Marxist sense, was maturing, in which working people were aware of continuing both old and new battles on their own" (p. 712). His study stands among the finest examples of historically careful development of a sociological concept.

In contrast to the economic focus on issues involving capitalist expansion, studies of the growth of national states and systems of states examine political topics such as state bureaucratization, the democratization of politics, revolutions, and the interaction of nations in the international arena. Three heralded historical sociology studies in this genre are Eisenstadt (1963), Moore (1966), and Skocpol (1979). Eisenstadt studied twenty-two preindustrial states that had centralized, impersonal, bureaucratic empires through which political power operated. After a tightly woven and systematic analysis of comparative social, political, and bureaucratic patterns, he concluded that "in any of the historical bureaucratic societies, their continued prominence was dependent upon the nature of the political process that developed in the society: first, on the policies of the rulers; second, on the orientations, goals, and political activities of the principal strata; and third, on the interrelations between these two" (Eisenstadt 1963, p. 362).

Moore's case studies of revolutions in England, France, the United States, China, Japan, and India sought "to understand the role of the landed upper classes and peasants in the bourgeois revolutions leading to capitalist democracy, the abortive bourgeois revolutions leading to fascism, and the peasant revolutions leading to communism" (Moore 1966, p. xvii). His own sympathies, however, lay in the development of political and social systems that fostered freedom, and he realized the importance of "a violent past" in the development of English, French, and American democracies

(pp. 39, 108, 153). He concluded "that an independent nobility is an essential ingredient in the growth of democracy" (p. 417) yet realized that a nobility's efforts to free itself from royal controls "is highly unfavorable to the Western version of democracy," unless these efforts occur in the context of a bourgeois revolution (p. 418).

Skocpol scrutinized the "causes and processes" of social revolutions in France, Russia, and China "from a nonvoluntarist, structural perspective, attending to international and world-historical, as well as intranational, structures and processes." While doing so she moved "states—understood as potentially autonomous organizations located at the interface of class structures and international situations—to the very center of attention" (Skocpol 1979, p. 33). She concluded that "revolutionary political crises, culminating in administrative and military breakdowns, emerged because the imperial states became caught in cross-pressures between intensified military competition or intrusions from abroad and constraints imposed on monarchical responses by the existing agrarian class structures and political institutions" (p. 285).

Underappreciated by theorists of historical sociology is the growing number of studies that apply sociological categories and concepts to the emergence and development of historically significant religious traditions (see Swatos 1977). By doing so, these scholars have surpassed the traditional sociological and historical colleagues who limit their efforts primarily to political and structural issues, especially ones arising during the late eighteenth to twentieth centuries. Swanson (1960), for example, coded material on fifty hunting-and-gathering societies in an effort to connect religion and magic to social structure, and subsequently analyzed relationships between constitutional structures and religious beliefs around the period of the Protestant Reformation (Swanson 1967). The emergence and early development of major religious traditions has received considerable sociological attention. They include, for example, analyses of early Christianity as a social movement (Blasi 1988), as a millenarian movement (Gager 1975; see Lang 1989, p. 339; Meeks 1983, pp. 173–180), and as a dramatically expansive young religion during the first centuries of the Common Era (Stark 1996). Concepts from sociological studies of modern sectarianism have informed historical studies of Mahayana Buddhism (Kent 1982) and

Valentinian Gnosticism (Green 1982). Weberian examinations continue to influence historically grounded studies of numerous world religions, including ancient Judaism (Zeitlin 1984), Islam (Turner 1974), and other religious traditions from around the world (see Swatos 1990). On a grand scale, religion figures prominently in the work of globalization theorist Roland Robertson (e.g., Robertson 1985; see Robertson 1992).

The historically grounded research in the sociology of religion, along with the works of Eisenstadt and others, suggests that future historical sociological studies will continue pushing beyond the confines of modern, macrosociological topics and into a wide range of premodern historical areas. Likewise, historical issues likely will become more consciously developed in microsociological studies (see Abrams 1982, pp. 227–266), and there will appear more sociologically informed historical examinations of cultural development (still exemplified by Elias 1978). Nonetheless, considerable macrosociological research still needs to be performed on historical issues affecting preindustrializing and Third World countries as well as on recent international realignments between forms of capitalism and communism.

Sometimes coupled with communism's collapse is the reappearance of nationalism and ethnic violence in various parts of the world. The deterioration of the Balkans in what (under Soviet influence) was Yugoslavia, for example, prompted Veljko Vujacic (1996) to develop "a comparative examination of contemporary Russian and Serbian nationalism" by attempting "to demonstrate the usefulness of some of Weber's key theoretical ideas on nations, nationalism, and imperialism" (p. 789). Specifically, he tried "to show how long-term historical and institutional legacies, shared memories, and defining political experiences, played themselves out in the contemporary period, influencing the different availability of mass constituencies in Russia and Serbia for nationalist mobilization under the auspices of new 'empire-saving coalitions'" (p. 789). Despite being faced with the difficult task of utilizing Weberian theory in this comparative context, Vujacic warned, "there is no need to gloss over the frequently bloody and unpredictable consequences of [people's] struggles with unduly abstract sociological generalizations. Instead, we should theorize our narrative, while giving contingency its place" (p. 790). In essence,

narrative accuracy must not be compromised by the demands of theory building.

Innovative methodological techniques continue to enter the historical sociology realm. As outlined by Roberto Franzosi and John Mohr, two such methodologies are network analysis (mapping "the connectivities among people, groups, and organizations" [Franzosi and Mohr 1997, pp. 145–154]) and content analysis ("the use of formal methods for extracting meanings from textual materials" [Franzosi and Mohr 1997, p. 149; see Tilly 1997, pp. 219ff.]). These and other methodologies contribute to the production of respected historical sociology studies on such topics as capitalist expansion, "the growth of national states and systems of states," collective action (Tilly 1981, p. 44; and see Tarrow 1996 for a review of Tilly's work), and the sociology of religious development. (Not to be overlooked, however, are studies of both capital-state relations and urban development and poverty [Gotham and Staples 1996].) Discussion will continue over the relative merits of globalization theory, world systems theory, and modernization theories (see Mandaliou 1996), and globalization theory likely will expand and modify to include greater emphasis on regional world systems, particularly in East-Southeast Asia (Ikeda 1996). The extent to which Michel Foucault's historically grounded, poststructuralist work on the relationships among cultural meanings, self-identities, and social control extend into the historical sociology corpus (for example, Foucault 1979, 1980; Dean 1994; see Seidman 1994, pp. 212–233) remains to be seen.

As older sociological methodologies disseminate within history and newer ones enter the discipline, we will continue to appreciate "the importance of historical sociology in enlarging our understanding of the historical variations in class, gender, revolutions, state formation, religion, and cultural identity" (Somers 1996, p. 53). On a practical level, the profile of historical sociology in its respective disciplines will be enhanced as universities extend cross-appointments of faculty, cross-list courses in methodologies and other topics, share theses and dissertation supervisory responsibilities among history and sociology faculties; encourage interdepartmental co-sponsorship of visiting speakers, and generally enhance the learning and training environments for graduate

students and staff. On a professional level, the American Sociological Association continues to sponsor a section in its annual conference entitled "Comparative and Historical Society."

(SEE ALSO: *Comparative-Historical Sociology*; *Event History Analysis*)

## REFERENCES

- Abrams, Philip 1982 *Historical Sociology*. Shepton Mallet, U.K.: Open House.
- Bart, Pauline, and Linda Frankel 1986 *The Student Sociologist's Handbook*, 4th ed. New York: Random House.
- Bellah, Robert N. 1959 "Durkheim and History." *American Sociological Review* 24:447-461. Reprinted in Robert A. Nisbet, ed., *Emile Durkheim*, Englewood Cliffs, N.J.: Prentice-Hall, 1965.
- Blasi, Anthony J. 1988 *Early Christianity as a Social Movement*. New York: Peter Lang.
- Bonnell, Victoria E. 1980 "The Uses of Theory, Concepts, and Comparison in Historical Sociology." *Comparative Studies in Society and History* 22:156-173.
- Burke, Peter 1980 *Sociology and History*. London: George Allen and Unwin.
- Dean, Mitchell 1994 *Critical and Effective Histories: Foucault's Methods and Historical Sociology*. London: Routledge.
- Eisenstadt, S. N. 1963 *The Political System of Empires*. New York: Free Press.
- Elias, Norbert 1978 *The Civilizing Process: vol. 1. The History of Manners*, trans. Edmund Jephcott. New York: Urizen Books.
- Emirbayer, Mustafa 1996a "Durkheim's Contribution to the Sociological Analysis of History." *Sociological Forum* 11:263-284.
- 1996b "Useful Durkheim." *Sociological Theory* 14:109-130.
- Engels, Frederick (1870) 1926 *The Peasant War in Germany*, 2nd ed., trans. Moissaye J. Olgin. New York: International Publishers.
- Ferrarotti, Franco 1997 "The Relation Between History and Sociology: Synthesis or Conflict?" *International Journal of Contemporary Sociology* 34:1-16.
- Foucault, Michele 1979 *Discipline and Punish: The Birth of the Prison*. New York: Vintage.
- 1980 *The History of Sexuality: vol. 1. An Introduction*. New York: Vintage.
- Franzosi, Roberto, and John W. Mohr 1997 "New Directions in Formalization and Historical Analysis." *Theory and Society* 26:133-160.
- Gager, John G. 1975 *Kingdom and Community: The Social World of Early Christianity*. Englewood Cliffs, N.J.: Prentice-Hall.
- Gotham, Kevin Fox, and William G. Staples 1996 "Narrative Analysis and the New Historical Sociology." *The Sociological Quarterly* 37:481-501.
- Green, Henry A. 1982 "Ritual in Valentinian Gnosticism: A Sociological Interpretation." *Journal of Religious History* 12:109-124.
- Ikeda, Satoshi 1996 "The History of the Capitalist World-System vs. the History of East-Southeast Asia." *Review* (Fernand Braudel Center) 19:49-77.
- Isaac, Larry W. 1997 "Transforming Localities: Reflections on Time, Causality, and Narrative in Contemporary Historical Sociology." *Historical Methods* 30:4-12.
- Jensen, Gary F. 1997 "Time and Social History." *Historical Methods* 30:46-57.
- Kalberg, Stephen 1994 *Max Weber's Comparative-Historical Sociology*. Oxford: Polity Press.
- Kent, Stephen A. 1982 "A Sectarian Interpretation of the Rise of Mahayana." *Religion* 12:311-332.
- 1983 "Weber, Goethe, and the Nietzschean Allusion: Capturing the Source of the 'Iron Cage' Metaphor." *Sociological Analysis* 44:297-319.
- 1985 "Weber, Goethe, and William Penn: Themes of Marital Love." *Sociological Analysis* 46:315-320.
- 1990 "The Quaker Ethic and the Fixed Price Policy: Max Weber and Beyond." In William H. Swatos, ed., *Time, Place, and Circumstance: Neo-Weberian Studies in Comparative Religious History*. Westport, Conn.: Greenwood Press.
- Kiser, Edgar 1996 "The Revival of Narrative in Historical Sociology: What Rational Choice Theory Can Contribute." *Politics and Society* 24:249-271.
- Lang, Graeme 1989 "Oppression and Revolt in Ancient Palestine: The Evidence in Jewish Literature from the Prophets to Josephus." *Sociological Analysis* 49:325-342.
- Lipset, Seymour Martin 1968 "History and Sociology: Some Methodological Considerations." In Seymour Martin Lipset and Richard Hofstadter, eds., *Sociology and History: Methods*. New York: Basic Books.
- MacKinnon, Malcolm H. 1988a "Calvinism and the Infallible Assurance of Grace: The Weber Thesis Reconsidered." *British Journal of Sociology* 39:143-177.
- 1988b "Weber's Exploration of Calvinism." *British Journal of Sociology* 39:178-210.

- Mandalios, John 1996 "Historical Sociology." In Bryan S. Turner, ed., *The Blackwell Companion to Social Theory*. Oxford: Blackwell.
- Marshall, Gordon 1982 *In Search of the Spirit of Capitalism: An Essay on Max Weber's Protestant Ethic Thesis*. New York: Columbia University Press.
- Marx, Karl (1885) 1963 *The Eighteenth Brumaire of Louis Bonaparte*, 3rd ed. New York: International Publishers.
- McDonald, Terrence J. 1996 "What We Talk About When We Talk about History: The Conversation of History and Sociology." In Terrence J. McDonald, ed., *The Historic Turn in the Human Sciences*. Ann Arbor: University of Michigan Press.
- Meeks, Wayne A. 1983 *The First Urban Christians: The Social World of the Apostle Paul*. New Haven, Conn.: Yale University Press.
- Merton, Robert 1938 "Science, Technology, and Society in Seventeenth-Century England." *Osiris* 4:360–632. Reprinted with a new preface, New York: Harper Torchbooks, 1970.
- Mills, C. Wright 1959 *The Sociological Imagination*. New York: Oxford University Press.
- Moore, Barrington, Jr. 1966 *Social Origins of Dictatorship and Democracy*. Boston: Beacon Press.
- Poggi, Gianfranco 1972 *Images of Society: Essays on the Sociological Theories of Tocqueville, Marx, and Durkheim*. Stanford, Calif.: Stanford University Press.
- Robertson, Roland 1985 "The Sacred and the World System." In Phillip E. Hammond, ed., *The Sacred in a Secular Age*. Berkeley: University of California Press.
- 1992 *Globalization: Social Theory and Global Culture*. London: Sage.
- Schwartz, Mildred A. 1987 "Historical Sociology in the History of American Sociology." *Social Science History* 11:1–16.
- Seidman, Steven 1994 *Contested Knowledge: Social Theory in the Postmodern Era*. Oxford: Routledge.
- Skocpol, Theda 1979 *States and Social Revolutions*. London: Cambridge University Press.
- 1984 "Emerging Agendas and Recurrent Strategies in Historical Sociology." In Theda Skocpol, ed., *Vision and Method in Historical Sociology*. Cambridge, U.K.: Cambridge University Press.
- 1987 "Social History and Historical Sociology: Contrasts and Complementarities." *Social Science History* 11:17–30.
- Smelser, Neil, J. 1959 *Social Change in the Industrial Revolution: An Application to of Theory to the British Cotton Industry*. Chicago: University of Chicago Press.
- 1968 *Essays in Sociological Explanation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Smith, Dennis 1991 *The Rise of Historical Sociology*. Oxford: Polity Press.
- Somers, Margaret R. 1996 "Where Is Sociology After the Historic Turn? Knowledge Cultures, Narrativity, and Historical Epistemologies." In Terrence J. McDonald, ed., *The Historic Turn in the Human Sciences*. Ann Arbor: University of Michigan Press.
- Stark, Rodney 1996 *The Rise of Christianity: A Sociologist Reconsiders History*. Princeton, N.J.: Princeton University Press.
- Swanson, Guy 1960 *The Birth of the Gods: The Origin of Primitive Beliefs*. Ann Arbor: University of Michigan Press.
- 1967 *Religion and Regime: A Sociological Account of the Reformation*. Ann Arbor: University of Michigan Press.
- Swatos, William H., Jr. 1977 "The Comparative Method and the Special Vocation of the Sociology of Religion." *Sociological Analysis* 38:106–114.
- ed. 1990 *Time, Place and Circumstance: Neo-Weberian Studies in Comparative Religious History*. Westport, Conn.: Greenwood Press.
- Sztompka, Piotr 1986 "The Renaissance of Historical Orientation in Sociology." *International Sociology* 1:321–337.
- Tarrow, Sidney 1996 "The People's Two Rhythms: Charles Tilly and the Study of Contentious Politics. A Review Article." *Comparative Studies in Society and History* Vol. 38, 3:586–600.
- Thomas, Keith 1963 "History and Anthropology." *Past and Present* 24:3–24.
- 1971 *Religion and the Decline of Magic*. New York: Scribner's.
- Thompson, E. P. 1972 "Anthropology and the Discipline of Historical Context." *Midland History* 1:41–55.
- 1963 *The Making of the English Working Class*, 1966 ed., New York: Vintage Books.
- Tilly, Charles 1981 *As Sociology Meets History*. New York: Academic Press.
- 1997 *Roads from Past to Future*. Oxford: Rowman and Littlefield.
- Tocqueville, Alexis de (1835) 1969 *Democracy in America*, 2 vols., trans. George Lawrence. London: Collins.
- (1858) 1955 *The Old Regime and the French Revolution*, 4th ed., trans. Stuart Gilbert. Garden City, N. Y.: Doubleday.

- Turner, Bryan S. 1974 *Weber and Islam: A Critical Study*. London: Routledge and Kegan Paul.
- Vujacic, Veljko 1996 "Historical Legacies, Nationalist Mobilization, and Political Outcomes in Russia and Serbia: A Weberian View." *Theory and Society* 25:763–891.
- Wallerstein, Immanuel 1974 *The Modern World System*. New York: Academic Press.
- Weber, Max 1930 *The Protestant Ethic and the Spirit of Capitalism*, trans. and ed. Talcott Parsons. New York: Scribners. 1976 Reprint. (originally published in German in 1920).
- 1949 *The Methodology of the Social Sciences*, trans. and eds. E. Shils and F. Finch. New York: Free Press.
- Wilson, B. R. 1971 "Sociological Methods in the Study of History." *Transactions of the Royal Historical Society* (5th ser.) 21:101–118.
- Zeitlin, Irving 1984 *Ancient Judaism: Biblical Criticism from Max Weber to the Present*. New York: Polity Press.

STEPHEN A. KENT

## HOMELESSNESS

*Literal* homelessness—lacking permanent housing of one's own—is a condition that has been present throughout human history. It has always been dangerous as well, given the necessity of shelter for survival. Nevertheless, the routine occurrence of homelessness in the past probably prevented the problem from generating any extraordinary degree of collective concern. Members of premodern societies often experienced losses or disruptions of residence as a result of food scarcity, natural disaster, epidemic disease, warfare, and other environmental and self-inflicted circumstances. Such forces contributed to the likelihood, if not the expectation, that most people would be homeless at some point in the life cycle.

Ironically, now that homelessness is relatively rare in Western societies, it has achieved a special notoriety. When shelter security becomes the norm, the significance of housing evolves beyond the purely functional. Homes, like jobs, constitute master statuses, anchoring their occupants in the stratification system. Hence, being without a home portends a more general and threatening *disaffiliation*, defined as "a detachment from society characterized by the absence or attenuation of the

affiliative bonds that link settled persons to a network of interconnected social structures" (Caplow et al. 1968, p. 494). This is the broadest meaning associated with the concept of homelessness, at the opposite end of the continuum from its literal definition.

Homelessness, broadly construed, first appeared on the American scene during the early stages of colonial settlement, with paupers, indentured servants, petty criminals, unemployed seamen, and the mentally impaired forming a pool of individuals at risk of vagrancy. It began to assume major proportions as a social problem in the United States near the end of the nineteenth century. Over the several preceding decades, urban homeless populations had emerged in response to a series of events at the national level, including Civil War displacement; the arrival of impoverished European immigrants; seasonal employment patterns in agriculture, construction, and the extractive industries; and severe economic setbacks in the early 1870s and 1890s (Rooney 1970).

As a makeshift remedy, downtown warehouses and old hotels were converted into inexpensive, dormitory-style lodging facilities. The proximity of the lodging facilities to one another, along with the distinctive mix of commercial and recreational establishments growing up around them, helped to concentrate the homeless physically in areas that came to be known as *skid rows* (supposedly named for a "skid road" in Seattle used to slide logs downhill). At the turn of the century, these areas were less burdened by the seedy images later evoked by the term "skid row." Instead, they were vibrant neighborhoods offering a temporary resting place and a range of services to thousands of tramps, the mobile workers who laid the foundation for the U.S. industrial economy.

The manpower needs created by World War I subsequently drained skid row districts, but a pool of footloose veterans replenished them at war's end. An even greater surge in homelessness—one extending well beyond the boundaries of skid row—was soon sparked by the Great Depression. The widespread hardship of the period forced previously domiciled individuals into a migrant lifestyle, and shantytowns (dubbed "Hoovervilles") sprang up in urban and rural settings alike. These new manifestations of homelessness in turn stimulated the first generation of sustained research on

the subject among sociologists. Anderson (1940), Sutherland and Locke (1936), and other scholars conducted studies of different segments of the homeless population as part of the Depression-era relief effort.

A second generation of research started in the 1950s. Large-scale single-city surveys—many of which were funded by urban renewal agencies—formed the debate over what to do about deteriorating skid row areas (Bahr and Caplow 1974; Bogue 1963). Demographic data obtained during the surveys showed homeless respondents to be predominantly male, white, single, older, and of local origins. The surveys also lent credibility to the popular view of the homeless as deviant “outsiders.” Depending on the city under examination, between one-fourth and one-half reportedly were problem drinkers, a higher percentage had spent time in jail or prison, most were unable or unwilling to hold down steady employment, many suffered from poor health, and few were enmeshed in supportive social networks. This negative profile based on the survey findings was countered by a parallel body of ethnographic evidence. Field observers like Wallace (1965) portrayed the homeless of skid row in subcultural terms, as a cohesive group with their own language, norms, and status hierarchy. Participation in the subculture was believed to help members cope with a problem more serious than their presumed deviance: extreme poverty.

In the 1970s, almost a century after skid row became a recognizable entity in the American city, its demise seemed imminent. Urban renewal and redevelopment projects had eliminated much of the infrastructure of skid row, while slackening demand for short-term unskilled labor was eroding one of the few legitimate economic roles the area could claim to play. Consequently, several investigators predicted skid row’s disappearance and, by implication, the decline of the U.S. homeless population (Lee 1980). Yet within a decade of such forecasts, homelessness had resurfaced as an important national issue. During the 1980s media coverage of the so-called *new homeless* increased dramatically, and federal legislation (most notably the McKinney Act) was formulated to address their plight. The amount of social scientific inquiry rose as well. Indeed, the recent outpouring of scholarly monographs on the topic has surpassed that of any prior generation of research.

Despite this renewed interest, what is known about contemporary homelessness remains limited, for several reasons. Unlike most groups surveyed by sociologists, the homeless are not easily reached at residential addresses or telephone numbers. The demolition of skid row districts in general and of single-room-occupancy (SRO) hotels in particular, accompanied by social control measures designed to reduce the public visibility of drunks, panhandlers, and other “undesirables,” has intensified the difficulties involved in finding homeless people, pushing a higher percentage of them into more dispersed, obscure locations. Those referred to as the *doubled up*, who stay with settled relatives or friends, are virtually inaccessible to investigators. Also poorly captured by surveys are the many individuals for whom homelessness is of brief duration or episodic in character. Even among the homeless who can be found, participation rates fall far short of perfect. The prospect of further stigma and humiliation keeps some from admitting their condition, thus excluding them from sample membership, while others are too suspicious or incoherent to take part in an interview.

Finally, the political context surrounding the latest wave of research magnifies the significance of each methodological obstacle just identified. Because the homelessness issue has been transformed into a referendum on the ability of the state to meet its citizens’ needs, liberals and conservatives both use the slightest technical shortcoming as ammunition with which to attack any study unfavorable to their own position. Similarly, members of both camps—not to mention the media, advocacy groups, government agencies, and other actors—selectively draw on research results to frame the homelessness problem in a way that attracts (or diverts) public attention. Thus, apparently straightforward “facts” about homelessness—and there are few of these to begin with—become matters open to debate.

“Snapshot,” or single-point-in-time, data on the size of the national homeless population illustrate the uncertain nature of the existing knowledge base. According to an early assertion by advocates, the number of homeless in the United States as of 1982 stood at 2.2 million, or approximately 1 percent of the total population of the country (Hombs and Snyder 1982). However, only two years later the U.S. Department of Housing and Urban Development (HUD) (1984) compiled

a series of estimates, extrapolated from street counts and surveys of informants and shelter operators, that yielded a "most reliable" range of 250,000 to 350,000. A 1987 Urban Institute study arrived at a figure—500,000 to 600,000 homeless nationwide on a single day—that fell between the advocate and HUD extremes (Burt and Cohen 1989). More recently, the Census Bureau enumerated 240,000 homeless people in the course of its massive yet heavily criticized 1990 S-night (street and shelter) operation (U.S. Bureau of the Census 1992).

Although most experts now dismiss the advocate-generated 2 million figure as groundless, the remaining estimates vary considerably. One explanation for this variation is that all are *point estimates*, depicting the size of the homeless population at a specific moment (e.g., a particular day or week). To the extent that the number of homeless changed during the 1980s, studies conducted on different dates should produce different results. Indeed, the trend retrospectively uncovered by several investigators (Jencks 1994)—slow growth early in the decade, rapid increase in the middle, and decline after 1987–1988—seems consistent with the magnitudes of the HUD, Urban Institute, and Census estimates. Others believe that the homeless population grew rapidly throughout the decade, by as much as 25 percent annually in some places. That growth rate could be inflated, though, given the relative stability documented in one of the few large cities (Nashville, Tennessee) for which longitudinal observations are available (Lee 1989).

It is hard to know whether the most credible point estimates accurately reflect the true scope of homelessness. If the homeless population is marked by high turnover, with many people entering and exiting quickly, the total number who experience homelessness over a longer period will be grossly underestimated by a point estimate. Two recent *period-prevalence* studies illustrate this dynamic. In the first study, counts of unduplicated shelter users in New York and Philadelphia suggest that roughly 1 percent of the residents of both cities are homeless each year, and the figure rises to 3 percent for a three-to-five-year interval (Culhane et al. 1994). In the second, 15 percent of the respondents to a nationally representative telephone survey reported that they had been literally homeless or had doubled up with someone else during their lifetimes (Link et al. 1995).

While definitional and methodological differences underlie much of the disagreement over the magnitude of the homeless population, generalizations about its composition have been complicated by (1) the selective emphasis of many inquiries on atypical "slices" of the whole (homeless veterans, the mentally ill, etc.) and (2) real variation in the characteristics of the homeless across communities. Contrary to media reports and popular perceptions, the modal homeless individual is still an unattached male with local roots, similar in fundamental ways to his skid row counterpart of the 1950s or 1960s. Yet there clearly have been significant compositional shifts during the intervening period. Blacks and other minorities, rarely found on skid row, are now overrepresented among the homeless, and women, children, young adults, and high school graduates constitute larger segments of the population both absolutely and proportionally than they once did (Burt 1992; Rossi 1989). Family groupings, usually headed by the mother alone, have become more common as well. Taking these elements of demographic continuity and change together, perhaps the safest conclusion to be drawn is that a trend toward greater diversity distinguishes the new homelessness from the old.

The same conclusion applies with respect to deviant characteristics. Alcoholism, which previously constituted the most noticeable form of deviance among the homeless, is now rivaled by other kinds of substance abuse, and mental illness has surpassed physical illness as an object of public concern. Beyond a rough consensus regarding the greater variety of such problems in the current homeless population, little of a definitive nature is known about them. For example, a review of nine studies cited mental illness prevalence rates that run from a low of one-tenth to a high of one-half of all homeless (U.S. General Accounting Office 1988), and occasional reports suggest that as many as 90 percent are at least mildly clinically impaired. This wide range leaves room for opposing arguments: on the one hand, that pervasive mental illness is the principal cause of contemporary homelessness; on the other, that its presumed causal role represents a stereotypic myth created by the visibility of a small minority of disturbed folk.

Even if the extent of mental illness has been exaggerated, there can be no doubt that the general well-being of the homeless remains low. This is

hardly surprising in light of the stresses that accompany life on the street. The absence of shelter exposes homeless persons to the weather, violence, and other threatening conditions. They have trouble fulfilling basic needs that most Americans take for granted, such as finding a job, obtaining nutritionally adequate meals, getting around town, washing clothes, storing belongings, and locating toilet and bathing facilities. To cope with these difficulties, homeless people draw on a repertoire of subsistence strategies (Snow and Anderson 1993). One of the most common is temporary low-wage employment, often as a day laborer. For some, *shadow work*—engaging in unconventional activities outside the formal economy (scavenging, panhandling, selling blood, trading junk)—offers a means of survival. Others resort to crime, especially petty theft, prostitution, and drug dealing, or become dependent on service providers.

While frequently creative, such strategies heighten the physical health risks to which the homeless are subjected. Compared to the settled population, a larger percentage of homeless individuals suffer from chronic disorders, and rates for most infectious diseases are at least five to six times greater (Wright et al. 1998). The collective consequence of these conditions is a drastically shortened life expectancy. However, to identify homelessness as the direct cause of higher morbidity and mortality would be an oversimplification. Preexisting health problems can reduce a person's employability, prompting a downward "drift" toward homelessness and lessening the chances of returning to a normal life. Homelessness can also be a complicating factor in the provision of health care. In part because of their circumstances (e.g., lacking transportation, distrusting authorities, being unable to store medicine), many homeless miss appointments and do not follow through with their prescribed treatment. They are, in short, less than ideal patients from the perspective of health professionals, whose goal is to insure continuity of care.

Poor health and other disadvantages associated with homelessness tend to worsen as the length of time on the streets increases. Some people still experience the longer-duration bouts common in the skid row era; close to 10 percent may be homeless for five continuous years or more. These "chronics," by virtue of their visibility, disproportionately influence public perceptions of who the

homeless are, but they now constitute the exception rather than the rule. Results from most surveys indicate that the median episode of homelessness lasts between two months and one year (Burt 1992; Link et al. 1995). Of the persons who fall into this "temporary" category, some are homeless only once in their lives. Many, though, exhibit a more complex pattern marked by frequent exits from and returns to homelessness (Piliavin et al. 1996). For such individuals, being without shelter is just one manifestation of prolonged residential instability.

Whether temporarily or chronically homeless, few prefer to be in that state. But if preference can be ruled out, what forces account for the new homelessness? Among the numerous answers elicited by this question so far, two general classes are discernible. *Structural* explanations treat homelessness as a function of large-scale trends that constrain people's chances for success and that are beyond their immediate control. Scholars point in particular to (1) the decreasing availability of affordable housing; (2) the growth of the poverty population; (3) changes in the economy (e.g., deindustrialization and the expansion of the service sector) resulting in fewer decent-paying, limited-skill jobs; (4) intensifying competition among members of the baby boom cohort during their adult years; (5) the declining appeal of marriage (and the heightened vulnerability of unmarried women and men); (6) the deinstitutionalization movement in mental health care policy; and (7) wider access to controlled substances, dramatically illustrated by the crack cocaine "epidemic" (Burt 1992; Jencks 1994; Wright et al. 1998). The rise of the new homeless is typically attributed to the convergence of two or more of these trends in the 1980s.

The availability of affordable housing has probably received the most attention of any structural factor, in part because all the other trends are thought to operate in conjunction with this one to produce homelessness. The thrust of the housing thesis is that government action, a supply-demand imbalance, inner-city revitalization, and related events have not only priced many low-income households out of rental status but have also eliminated a key fallback option historically open to them: SRO units in downtown residential hotels (Hoch and Slayton 1989; Ringheim 1990). With the depletion of the SRO stock, displacement



from other sectors of the housing market may lead directly to a homeless outcome.

In contrast to the structural approach, *individualistic* explanations posit traits, orientations, or experiences specific to the person as the main causes of homelessness. Few researchers have found much evidence that lacking permanent residence is a freely chosen lifestyle. Nevertheless, the enlargement of the emergency shelter system in recent years has made it easier for poor people who are exposed to domestic conflict or doubled up in a crowded unit to voluntarily head for a shelter as a way of coping with their untenable housing situations. In similar fashion, older thinking about the inherent immorality and wanderlust of skid row denizens has given way to revisionist claims that the primary antecedents of homelessness are deficits in talent or motivation or the debilitating effects of mental illness or substance abuse. Traumatic life events, either in childhood (e.g., sexual violence, placement in foster care or an orphanage) or adulthood (divorce, job loss, a serious health problem), can increase the likelihood of homelessness as well.

Interestingly, many experts who subscribe to some version of the individualistic view have had to invoke associated structural trends—deinstitutionalization in the case of mental illness, for example—in order to explain the size and compositional changes that have occurred in the homeless population in recent years. The tendency to draw on both individualistic and structural perspectives has grown more pronounced with the realization that a theory of homelessness, like that of any social phenomenon, can never be fully satisfying when cast in exclusively micro- or macro-level terms. To date, the work of Rossi (1989) offers the most compelling cross-level synthesis. He contends that structural changes have put everyone in extreme poverty at higher risk of becoming homeless, especially those poor people who exhibit an “accumulation of disabilities,” such as drug abuse, bad health, unemployment, and a criminal record. Being “disabled” forces one to rely on a network of friends and family for support, often over prolonged periods. If the strain placed on this support network is too great and it collapses, homelessness is the likely result.

Though Rossi’s central idea—that structural factors and individual problems combine to make

certain segments of the poor more vulnerable to homelessness than others—seems reasonable to social scientists, it could prove less acceptable to members of the general public. Based on previous research into public beliefs about the causes of poverty, most Americans might be expected to hold the homeless responsible for their lot. However, the small amount of evidence that bears directly on this expectation contradicts rather than confirms it. Findings from a local survey, supplemented with data from a national opinion poll, indicate that (1) more people blame homelessness on structural variables and bad luck than on individualistic causes and (2) many hold a mixture of structural and individualistic beliefs, consistent with the complex roots of the condition (Lee et al. 1990).

The relative frequency of the two types of beliefs is a matter of substantial political significance, since the study just cited shows that each type implies a distinctive set of policy attitudes. As a rule, members of the public who believe in structural causes consider homelessness a very important problem, feel that the response to it has been inadequate, and endorse a variety of corrective proposals, including a tax increase and government-subsidized housing. This policy orientation stands at odds with that for individualistic believers, who tend to devalue homelessness as an issue and favor restrictive measures (vagrancy enforcement, access limitation, etc.) over service provision. Regardless of which orientation ultimately registers the greatest impact on policy making, the sharp contrast between them says much about how homelessness has managed to stay near the top of the U.S. domestic social agenda for the past two decades.

Homelessness also persists as an issue because it is so hard to solve. Out of frustration, many communities have resorted to controlling the homeless. Historically, mechanisms to achieve this goal have included expulsion, spatial containment, and institutionalization (the latter for the sake of monitoring, rehabilitation, or punishment). Other efforts have been aimed at amelioration. For example, the Stewart B. McKinney Homeless Assistance Act recognizes the responsibility of the federal government to meet the basic needs of the “down and out.” Since the McKinney Act was signed into law in 1987, it has authorized billions of dollars for

food, shelter, health care, and other forms of aid (Foscarinis 1996). However, most of the funding is used to support emergency relief programs. The act designates only modest amounts for reintegration initiatives (for example, moving people into transitional or permanent housing) and virtually nothing for prevention. In short, it treats symptoms but not causes.

The effectiveness of legislation like the McKinney Act is further hindered by our federal system of government, which requires an unrealistically high degree of coordination among units at different levels to insure successful and equitable implementation. The current political climate does not bode well for such legislation, either. Repeated challenges have been made to the McKinney Act; critics want to reduce the size of the federal commitment, redirect the homeless toward existing social services (although many do not appear eligible to receive benefits), and give local government more flexibility. In practice, this is likely to mean a continuation of the piecemeal approach already taken in many places, with an assortment of non-profit organizations, religious groups, advocates, volunteers, and state and municipal agencies attempting to do their part. As long as communities lack specific intervention strategies for keeping at-risk residents from losing their housing and for equipping them with essential skills, there will be little change in the status quo.

The United Nations' designation of 1987 as the "Year of Shelter for the Homeless" attests that homelessness has been an international as well as an American concern. The situation in Europe resembles that in the United States in several respects. Although rates of homelessness appear to be slightly lower across the dozen or so European countries for which data are available, the number of people affected annually (at least 2.5 million) is large (Wright et al. 1998). The compositional profile of the European homeless population looks familiar: Its members are disproportionately single, male, from disadvantaged socioeconomic backgrounds, and in poor physical and mental health. Like their American counterparts, they have become homeless as a result of both structural pressures (e.g., a poverty and affordable housing "squeeze") and individual experiences (family breakup, substance abuse, etc.). Remedies for these causes are no easier to come by in Europe than

they are in the United States. The primary response thus far has been to offer emergency relief, with the burden of service provision falling on the private sector.

Homelessness takes a different, more acute form in the developing countries of the Third World, where rapid population growth outstrips the expansion of the housing stock and the economy by a wide margin. Compounding the growth-housing mismatch are prevailing patterns of spatial redistribution: Rural-to-urban migration streams have created huge pools of homeless people in tenements, in squatter communities, and on the streets of many large cities. Besides such demographic trends, periodic events of the kind that once created literal homelessness in premodern societies—drought, earthquakes, food shortages, and the like—still contribute to the problem today outside the West. War (including "ethnic cleansing") and political instability add to the toll.

What is perhaps most striking about homelessness in the Third World context is its youthful face. Visitors to developing nations cannot help but notice the ubiquitous street children; UNICEF estimates that there may be as many as 100 million globally (Glasser 1994). A majority are "on the street" during the daytime, typically performing some sort of economic activity (begging, vending, etc.), but they have a family and dwelling to return to at night. Perhaps 10 percent qualify as literally homeless or "of the street." The children belonging to this group may have run away from difficult family circumstances, been discarded as "surplus kids" by parents unable or unwilling to care for them, or been discharged from an orphanage or other institutional setting. Because children of the street must hustle to survive, they are occasionally romanticized as savvy and resilient. But they also lack adequate diets, are susceptible to criminal victimization, and engage in behaviors such as drug use and prostitution that jeopardize their health.

Sadly, the prospects for weaving a safety net to catch homeless children and adults—let alone for targeting the sources of the problem—must be judged slim in the face of the financial debts, service demands, and other burdens under which Third World governments operate. Possibly because these burdens are so overwhelming,

homelessness—while important—has yet to achieve dominant-issue standing. As one informed observer put the matter, “neither the resources to address the plight of the homeless nor the degree of aroused public sympathy present in the United States are in evidence in the developing world” (Knight 1987, p. 268). However, that is the sector of the world in which a vast majority of all homeless persons will continue to live for the foreseeable future.

## REFERENCES

- Anderson, Nels 1940 *Men on the Move*. Chicago: University of Chicago Press.
- Bahr, Howard M., and Theodore Caplow 1974 *Old Men Drunk and Sober*. New York: New York University Press.
- Bogue, Donald J. 1963 *Skid Row in American Cities*. Chicago: Community and Family Study Center, University of Chicago.
- Burt, Martha R. 1992 *Over the Edge: The Growth of Homelessness in the 1980s*. New York: Russell Sage Foundation.
- , and Barbara E. Cohen 1989 *America's Homeless: Numbers, Characteristics, and the Programs That Serve Them*. Washington, D.C.: Urban Institute.
- Caplow, Theodore, Howard M. Bahr, and David Sternberg 1968 “Homelessness.” In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 6. New York: Macmillan.
- Culhane, Dennis P., Edmund F. Dejowski, Julie Ibanez, Elizabeth Needham, and Irene Macchia 1994 “Public Shelter Admission Rates in Philadelphia and New York City: The Implications of Turnover for Sheltered Population Counts.” *Housing Policy Debate* 5:107–140.
- Foscarinis, Maria 1996 “The Federal Response: The Stewart B. McKinney Homeless Assistance Act.” In Jim Baumohl, ed., *Homelessness in America*. Phoenix, Ariz.: Oryx.
- Glasser, Irene 1994 *Homelessness in Global Perspective*. New York: G.K. Hall.
- Hoch, Charles, and Robert A. Slayton 1989 *New Homeless and Old: Community and the Skid Row Hotel*. Philadelphia: Temple University Press.
- Hombs, Mary E., and Mitch Snyder 1982 *Homelessness in America: A Forced March to Nowhere*. Washington, D.C.: Community for Creative Non-Violence.
- Jencks, Christopher 1994 *The Homeless*. Cambridge, Mass.: Harvard University Press.
- Knight, Rudolph H. 1987 “Homelessness: An American Problem?” In Richard D. Bingham, Roy E. Green, and Sammis B. White, eds., *The Homeless in Contemporary Society*. Newbury Park, Calif.: Sage.
- Lee, Barrett A. 1980 “The Disappearance of Skid Row: Some Ecological Evidence.” *Urban Affairs Quarterly* 16:81–107.
- 1989 “Stability and Change in an Urban Homeless Population.” *Demography* 26:323–334.
- , Sue Hinze Jones, and David W. Lewis 1990 “Public Beliefs About the Causes of Homelessness.” *Social Forces* 69:253–265.
- Link, Bruce, Jo Phelan, Michaeline Bresnahan, Ann Stueve, Robert Moore, and Ezra Susser 1995 “Lifetime and Five-Year Prevalence of Homelessness in the United States: New Evidence on an Old Debate.” *American Journal of Orthopsychiatry* 65:347–354.
- Piliavin, Irving, Bradley R. Entner Wright, Robert D. Mare, and Alex H. Westerfelt 1996 “Exits from and Returns to Homelessness.” *Social Service Review* 70:33–57.
- Ringheim, Karin 1990 *At Risk of Homelessness: The Roles of Income and Rent*. New York: Praeger.
- Rooney, James F. 1970 “Societal Forces and the Unattached Male: An Historical Review.” In Howard M. Bahr, ed., *Disaffiliated Man: Essays and Bibliography on Skid Row, Vagrancy, and Outsiders*. Toronto: University of Toronto Press.
- Rossi, Peter H. 1989 *Down and Out in America: The Origins of Homelessness*. Chicago: University of Chicago Press.
- Snow, David A., and Leon Anderson 1993 *Down on Their Luck: A Study of Homeless Street People*. Berkeley: University of California Press.
- Sutherland, Edwin H., and Harvey J. Locke 1936 *Twenty Thousand Homeless Men*. Chicago: J.B. Lippincott.
- U.S. Bureau of the Census 1992 *Statistical Abstract of the United States: 1992* (112th ed.). Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Housing and Urban Development 1984 *A Report to the Secretary on the Homeless and Emergency Shelters*. Washington, D.C.: Office of Policy Development and Research, U.S. Department of Housing and Urban Development.
- U.S. General Accounting Office 1988 *Homeless Mentally Ill: Problems and Options in Estimating Numbers and Trends*. Washington, D.C.: Program Evaluation and Methodology Division, U.S. General Accounting Office.
- Wallace, Samuel E. 1965 *Skid Row as a Way of Life*. Totowa, N.J.: Bedminster.

Wright, James D., Beth A. Rubin, and Joel A. Devine  
1998 *Beside the Golden Door: Policy, Politics, and the Homeless*. New York: Aldine de Gruyter.

BARRETT A. LEE

## HOMICIDE

See Crime Rates; Criminology.

## HOMOSEXUALITY

See Alternative Life Styles; Sexual Orientation;  
Sexually Transmitted Diseases.

## HUMAN ECOLOGY AND ENVIRONMENTAL ANALYSIS

With the growing awareness of the critical environmental problems facing the world today, ecology, the scientific study of the complex web of interdependent relationships in ecosystems, has moved to the center stage of academic and public discourse. The term *ecology* comes from the Greek word *oikos* (“house”) and, significantly, has the same Greek root as the word *economics*, from *oikonomos* (“household manager”). Ernst Haeckel, the German biologist who coined the word *ecology* in 1868, viewed ecology as a body of knowledge concerning the economy of nature, highlighting its roots in economics and evolutionary theory. He defined *ecology* as the study of all those complex interrelations referred to by Darwin as the conditions of the struggle for existence.

Ecologists like to look at the environment as an ecosystem of interlocking relationships and exchanges that constitute the web of life. Populations of organisms occupying the same environment (habitat) are said to constitute a *community*. Together, the communities and their abiotic environments constitute an *ecosystem*. The various ecosystems taken together constitute the *ecosphere*, the largest ecological unit. Living organisms exist in the much narrower range of the *biosphere*, which extends a few hundred feet above the land or under the sea. On its fragile film of air, water, and soil, all life depends. For the sociologist, the most important ecological concepts are diversity and

dominance, competition and cooperation, succession and adaptation, evolution and expansion, and carrying capacity and the balance of nature. Over the years, the human ecological, the neo-Malthusian, and the political economy approaches and their variants have come to characterize the field of human ecology.

## CLASSICAL HUMAN ECOLOGY

The Chicago sociologists Louis Wirth, Robert Ezra Park, Ernest W. Burgess, and Roderick McKenzie are recognized as the founders of the human ecological approach in sociology. In the early decades of the twentieth century, American cities were passing through a period of great turbulence due to the effects of rapid industrialization and urbanization. The urban commercial world, with its fierce competition for territory and survival, appeared to mirror the very life-world studied by plant ecologists. In their search for the principles of order, human ecologists turned to the fundamental process of *cooperative competition* and its two dependent ecological principles of *dominance* and *succession*. For classical human ecologists, such as Park (1936), these processes determine the distribution of population and the location and limits of residential areas. City development is then understood in terms of succession—an orderly series of invasion—resistance—succession cycles in the course of which a biotic community moves from a relatively unstable (primary) to a more or less permanent (climax) stage. If resistance fails and the local population withdraws, the neighborhood eventually turns over and the local group is succeeded by the invading social, economic, or racial population. Each individual and every community thus finds its appropriate *niche* in the physical and living environment. In the hands of the classical human ecologists, human ecology became synonymous with the ecology of space. Park and Burgess identified the natural areas of land use, which come into existence without a preconceived design. Quite influential and popular for a while was the “Burgess hypothesis” regarding the spatial order of the city as a series of concentric zones emanating from the central business district. However, Hawley (1984) has pointed out that with urban characteristics now diffused throughout society, one in effect deals with a system of cities in which the urban hierarchy is cast in terms of functional rather than spatial relations.

Since competition among humans is regulated by culture, Park (1936) made a distinction between the biotic and cultural levels of society: above the symbiotic society based upon competition stands the cultural society based upon communication and consensus. Park identified the problematic of human ecology as the investigation of the processes by which biotic balance and social equilibrium are maintained by the interaction of the three factors constituting what he termed the *social complex* (population, technological culture [artifacts], and nonmaterial culture [customs and beliefs]), to which he also added a fourth, the natural resources of the habitat. However, while human ecology is here defined as the study of how the interaction of these elements helps maintain or disrupts the biotic balance and the social equilibrium, human agency and the cultural level are left out of consideration by Park and other human ecologists.

#### NEOCLASSICAL HUMAN ECOLOGY

Essentially the same factors reappear as the four POET variables (population, organization, environment, and technology) in Otis Dudley Duncan's (1964) *ecological complex*, indicating its point of contact with the early human ecological approach. In any case, it was McKenzie who, by shifting attention from spatial relations to the analysis of sustenance relations, provided the thread of continuity between the classical and the neoclassical approaches. His student Amos Hawley, who has been the "exemplar" of neoclassical human ecology since the 1940s, defines human ecology as the attempt to deal holistically with the phenomenon of organization.

Hawley (1986) views the *ecosystem* as the adaptive mechanism that emerges out of the interaction of population, organization, and the environment. *Organization* is the adaptive form that enables a population to act as a unit. The process of system *adaptation* involves members in relations of interdependence in order to secure sustenance from the environment. *Growth* is the development of the system's inner potential to the maximum size and complexity afforded by the existing technology for transportation and communication. *Evolution* is the creation of greater potential for resumption of system development through the incorporation of new information that enhances the capacity for

the movement of people, materials, and messages. In this manner, the system moves from simple to more complex forms.

Hawley (1984) has identified the following propositions, which affirm the interdependence of the demographic and structural factors, as constituting the core of the human ecological paradigm:

1. Adaptation to environment proceeds through the formation of a system of interdependences among the members of a population,
2. system development continues, *ceteris paribus*, to the maximum size and complexity afforded by the existing facilities for transportation and communication,
3. system development is resumed with the introduction of new information that increases the capacity for movement of materials, people, and messages and continues until that capacity is fully utilized. (p. 905)

The four ecological principles of *interdependence*, *key function*, *differentiation*, and *dominance* define the processes of system functioning and change. A system is viewed as made up of functioning parts that are related to one another. Adaptation to the environment involves the development of interdependence among members, which increases their collective capacity for action. Differentiation then allows human populations to restore the balance between population and environment that has been upset by competition or improvements in communication and transportation technologies. As adaptation proceeds through a differentiation of environmental relationships, one or a few functions come to mediate environmental inputs to all other functions. Since power follows function in Hawley's view, dominance attaches to those units that control the flow of sustenance into the ecosystem. The productivity of the key function, which controls the flow of sustenance, determines the extent of functional differentiation. As a result, the dominant units in the system are likely to be economic rather than political.

Since the environment is always in a state of flux, every social system is continuously subject to change. Change alters the life condition of all participants, an alteration to which they must adapt in order to remain in the system. One of the most

significant nonrecurrent alterations is *cumulative change*, involving both endogenous and exogenous changes as complementary phases of a single process. While evolution implies a movement from simple to complex, proceeding through variation and natural selection, cumulative change refers to an increase in scale and complexity as a result of increases in population and territory. Whether the process leads to growth or evolution depends on the concurrent nature of the advances in scale and complexity.

Generalizing the process of cumulative change as a principle of *expansion*, Hawley (1979, 1986) applies this framework to account for growth phases that intervene between stages of development. When scale and complexity advance together, the normal conditions for growth or expansion arise from the colonization process itself. Expansion, driven by increases in population and in knowledge, involves the growth of a center of activity from which dominance is exercised. The evolution of the system takes place when its scale and complexity do not go hand in hand. Change is resumed as the system acquires new items of information, especially those that reduce the costs of movement away from its environment. Thus an imbalance between population and the carrying capacity of the environment may create external pressures for branching off into colonies and establishing niches in a new environment. Since efficiencies in transportation and communication determine the size of a population, the scope of territorial access, and the opportunity for participation in information flows, Hawley (1979) identifies the technology of movement as the most critical variable. In addition to governing accessibility and, therefore, the spread of settlements and the creation of interaction networks among them, it determines the changes in hierarchy and division of labor. In general, the above process can work on any scale and is limited only by the level of development of the technologies of communication and transportation.

Hawley (1986, pp. 104–106) points out how with the growth of a new regional and international division of labor, states now draw sustenance from a single biophysical environment and are converted to subsystems in a more inclusive world system by the expansive process. In this way, free trade and resocialization of cultures create a far more efficient and cost-effective global reach. The result is a global system thoroughly interlinked by

transportation and communication networks. The key positions in this international network are occupied by the technologically advanced nations with their monopoly of information and rich resource bases. However, as larger portions of system territories are brought under their jurisdiction, the management of scale becomes highly problematic. In the absence of a supranational polity, a multipolar international pecking order is then subject to increasing instability, challenge, and change. With mounting costs of administration, the system again tends to return to scale, resulting in some degree of decentralization and local autonomy, but new information and improvements in the technologies of movement put the system back into gear and start the growth process all over again. In the modern period of “ecological transition,” a large portion of the biophysical environment has progressively come under the control of the social system. Hawley, therefore, believes that the growth of social systems has now reached a point at which the evolutionary model has lost its usefulness in explaining cumulative change.

Hawley points out that while expansionism in the past relied on political domination, its modern variant aims at structural convergence along economic and cultural axes to obviate the need for direct rule by the center. The process of modernization and the activities of multinational corporations are a prime example of this type of system expansion, which undermines traditional modes of life and results in the loss of autonomy and sovereignty by individual states. Convergence of divergent patterns of urbanization is brought about by increased economic interdependence among nations and the development of compatible organizational forms and institutional arrangements. This approach, as Wilson (1984, p. 300) points out, is based on the assumption that convergence is mainly a result of market forces that allow countries to compete in the world on an equal basis. He cites evidence that shows how the subordinate status of non-Western nations has hindered their socioeconomic development, sharpened inequalities, increased rural-to-urban migration and rural-urban disparities, and led to the expansion of squatter settlements.

Human ecological theory accounts for the existence of an international hierarchy in terms of functional differences and the operation of its

universal principles of ecosystem domination and expansion. Quite understandably, underdevelopment is defined by Hawley simply in terms of inferiority in this network. Since not all can enjoy equal position on scales of size, resources, and centrality with respect to information flows, Hawley believes that the resulting “inequality among polities might well be an unavoidable condition of an international division of labor, whether built on private or state capitalist principles” (1986, pp. 106, 119).

As the process moves toward a world system, all the limiting conditions of cumulative change are reasserted at a higher level. On the one hand, a single world order with only a small tolerance for errors harbors the seeds of totalitarianism (Giddens 1990). On the other, there is also the grave danger that a fatal error may destroy the whole system. Human ecologists, however, rely on further expansion as the sure remedy for the problems created by expansion. To restore ecological balance, they put their faith in the creation of value consensus, rational planning, trickle-downs, market mechanisms, technological fixes and breakthroughs, native “know-how,” and sheer luck.

The real irony of this relentless global expansion elaborated by Hawley lies, however, in the coexistence of the extreme opulence and affluence of the few with the stark poverty and misery of the majority at home and abroad. The large metropolitan centers provide a very poor quality of life. The very scale of urban decay underscores the huge problems facing the city—congestion, polluted air, untreated sewage, high crime rates, dilapidated housing, domestic violence, and broken lives. One therefore needs to ask: What prospect does this scale and level of complexity hold for the future?

#### HUMAN ECOLOGY AND THE PROBLEMATICS OF “CHAOS”

Chaos theory is the latest attempt to unravel the hidden structure in apparently random systems and to handle chaos within and between systems. In this view, order and disorder (chaos) are seen as two dimensions of the same process: Order generates chaos and chaos generates order (Baker 1993, p. 123). At the heart of both lies a dynamic element, an “attractor,” that creates the turbulence

as well as re-creates the order. In the human–social realm, Baker has identified center–periphery, or *centriphery*, as the attractor. Baker, however, uses the concepts of center and periphery more broadly to cover not only their application in the dependency approach (which views the exploitation and impoverishment of the non-Western peripheral societies as basic to the rise of the dominant Western capitalist center), but also to carry the connotation of humans as “world-constructors.” Centriphery is, then, the universal dynamic process that creates both order and disorder, as well as accounts for the pattern of human social evolution. The center has an entropic effect on the periphery, causing increasing randomness and denuding it of its resources. But as the entropic effects mount, they are fed back to the center. Beyond a certain point, the costs of controlling the periphery become prohibitive. Should the center fail to come up with new centering strategies, it may split off into subcenters or be absorbed by another more powerful center. Baker is thus led to conclude that “although the effect of feedback is unpredictable, the iteration of a pattern leads to turbulence. The mechanism for change and evolution are endemic to the centriphery process” (Baker 1993, p. 141).

Several things need to be noted about this approach. For one, since these eruptive episodes are random, “the emergence of repeated patterns . . . must be seen as random . . . not as mechanically predictable occurrences. Among other things, the precise character of the emergent pattern cannot be predicted, even though we would no longer be surprised to find a new thing emerging” (Francis 1993, p. 239). For another, Baker’s centriphery theory is essentially Hawley’s human ecological theory recast in the language of chaos theory, with the important difference that a specific reference is now made both to the role of *agency* as “world-constructors” involved in “categorizing, controlling, dominating, manipulating, absorbing, transforming, and so on,” and to their devastating impact on the peripheralized “others,” the victims of progress, who suffer maximum entropy, exploitation, impoverishment, death, and devastation. Even so, Baker’s is the latest, though undoubtedly unintended, attempt to generalize and rationalize Western expansionism and its “chaotic (unpredictable) . . . devastating, and now increasingly well known, impact on native peoples” (Baker

1993, p. 137). As such, the centripetal process, said to explain both order (stability) and disorder (change), is presented not only as evolutionary and irreversible, but also as natural and universal: "Thus, the Western world became a center through the peripheralization of the non-Western world. And within the Western world, particularly in North America, the city, which peripheralized the rural hinterland, became the megapolis whose peripheralizing effects were simultaneously wider and greater." (p. 136)]

Not only the recurrent iteration of this pattern but even its "unpredictable" outcomes (new strategies of control, splitting off into new subcenters, absorption into a larger center, etc.), are also prefigured in Hawley (1986). Its process is expansion, and its "attractor" is none other than the old master principle of sociology: domination or control (Gibbs 1989). While Friedmann and Wolff (1982) characterize world cities as the material manifestation of the control functions of transnational capital in its attempt to organize the world for the efficient extraction of surplus, Lechner (1985) leaves little doubt that Western "[materialism] and the emphasis on man's relation to nature are not simply analytical or philosophical devices, but are logically part of an effort to restore world-mastery" (p. 182).

#### **"ECOLOGICAL DEMOGRAPHY"?**

Since the study of organizational dynamics as well as the structure and dynamics of population are at the core of sociology, Namboodiri (1988) claims that rather than being peripheral to sociology, human ecology and demography constitute its core. As a result, he contends that the hybrid "ecological demography" promises a more systematic and comprehensive handling of a common core of sociological problems—such as the analysis of power relations, conflict processes, social stratification, societal evolution, and the like—than any other competing sociological paradigm. However, although human ecologists recognize the possibility of other pairwise interactions in addition to competition, and even highlight the points of convergence between the human ecological and the Marxist point of view (Hawley 1984), human ecology as such does not directly focus on conflict in a central way. In this connection, Namboodiri (1988) points out how the very

expansionist imperative of human and social systems, identified as a central postulate by human ecologists, generates the possibility of conflict between the haves and the have-nots far more in a milieu of frustrated expectations, felt injustice, and a growing awareness of entitlements, which includes claims to their own resources by nations and to a higher standard of living by deprived populations. How these factors affect the development of and distribution of resources and the relationships among populations by sex, race, ethnicity, and other stratifiers should obviously be of concern to a socially responsible human ecology, one that moreover should be responsive to Borgatta's call for a "proactive sociology" (1989, 1996).

The general neglect of cultural factors and the role of norms and agency in human and organizational interaction has also been a cause for concern to many sociologists. While some latitude is provided for incorporating social norms in specific analyses (e.g., in the relationship between group membership and fertility behavior), their macro-orientation and focus on whole populations compels human ecologists and demographers to ignore the role of the subjective values and purposes of individual actors in ecological and demographic processes (Namboodiri 1988, pp. 625–627).

#### **THE HUMAN ECOLOGICAL APPROACH: AN EVALUATION**

While the human ecological approach has strong theoretical underpinnings and proven heuristic value in describing Western expansionism and the colonizing process in supposedly objective terms, its central problem is one of ideology. Like structural functionalism, it is a theory of the status quo that supports existing institutions and arrangements by explaining them as the outcome of invariant principles: "Its concerns are the concerns of the dominant groups in society—it talks about maximizing efficiency but has nothing to say about increasing accountability, it talks of maintaining equilibrium through gradual change and readjustment and rules out even the possibility of fundamental restructuring" (Saunders 1986, pp. 80–81). Not surprisingly, human ecologists downplay the role of social class by subsuming it under the abstract concept of a "categorical unit." They also fail to analyze the role of the state and of the interlocking power of the political, military, and



economic establishment, which are centrally involved in the process of expansion and colonization of peoples and cultures. These omissions account for their total lack of concern for the fate of the "excluded others" and the "dark side of expansionism": colonial exploitation, war, genocide, poverty, pollution, environmental degradation, and ecological destruction. Hutchinson (1993) blames the human ecologists for neglecting or downplaying the role of socioeconomic practices and government policies in creating rental, economic resource, and other differentials. He claims that their analyses tend to be descriptive because they take for granted the existence of phenomena such as socioeconomic or racial and ethnic segregation rather than looking at them in terms of spatial processes that result from the competition between capital and labor.

#### **ENVIRONMENTAL SOCIOLOGY AND THE NEW HUMAN ECOLOGY**

The mounting public concern since the 1970s about fuel shortages, oil spills, nuclear power plant accidents, acid rain, dying lakes, urban smog, famine and death in the Sahel, rainforest destruction, and the like has made social scientists realize that overexploitation of the ecosystem may destroy the very basis of our planetary survival. Many environmentalists have blamed the voracious appetite of industrial societies and their obsession with growth for the destruction of the fragile balance among the components of the ecological complex.

Having encountered a seemingly unlimited frontier and an expanding economy, the West has come to believe that expansion is in the nature of things. A major reason for the neglect of the physical environment by American sociologists has, therefore, been the anti-ecological worldview of the dominant social paradigm that has been shaped by this belief. At the same time, the exaggerated emphasis by human ecologists on culture, science, and technology as "exceptional" human achievements has led to the illusion that humans are "exempt" from bioecological constraints to which all species are subject. This awareness has led Catton and Dunlap (1978) to develop the fields of new human ecology (Buttel 1987) and environmental sociology to deal with the reciprocal interaction between human activities and the physical

environment. They believe that the POET model, broadened to include the role of human agency and culture, provides a useful analytical framework for grounding environmental sociology in the ecological perspective.

In a comprehensive review of the new field, Buttel (1987) has pressed for shifting the focus of environmental sociology from the imbalance of population and resources, emphasized by Catton, to the reality of the unequal distribution of these resources. Allan Schnaiberg's idea of the "treadmill of production" (1980), which emerges from a dialectical relationship between economic growth and ecological structures, points to the need to focus on production institutions as the primary determinants of economic expansion and to incorporate a conflict dimension in environmental analysis. Buttel's own work in environmental sociology draws on the "political economy tradition" of the neo-Marxists and the neo-Weberians. Catton's major contributions, on the other hand, are in the neo-Malthusian tradition. While the problem of order created by the harsh realities of industrial life and expansionism had earlier defined the central problematic of sociology and human ecology, the problem of survival now defines the central problematic of environmental sociology and the neo-Malthusian new human ecology: to the earlier question of how social order is possible is now added the more urgent concern with survival itself.

#### **THE POLITICAL ECONOMY APPROACH AND THE NEW URBAN SOCIOLOGY**

The conservative nature of the classical and neoclassical human ecology paradigms has also come under attack from theorists who focus on the internal contradictions and the global reach of capitalism to understand urban phenomena.

Smith (1995) has argued that a new urban sociology paradigm, which draws on neo-Marxist sociological theory, urban political economy, dependency theory, world-system analysis, and critical theory, has now become dominant and largely supplanted human ecology and the old urban sociology approach to the study of urban phenomena. The conflict between the two approaches is an aspect of the old conflict between the structural-functional and the neo-Marxist (conflict) perspectives in the field of sociology generally. Whereas

human ecology's main concern is with how technological change enables population aggregates to adapt to their environment through changes in social and spatial organization, the new perspective underplays the role of technological determinants or functional imperatives in shaping the urban landscape. It focuses instead on social inequality and conflict, and highlights the role of economic and political elites, states and other institutional actors, and powerful global forces in order to analyze the problematic "underside" of modern city life: urban poverty, housing segregation by race and social class, urban fiscal crises, deindustrialization, structured inequality in the built environment, and the massive level of human misery associated with the rapid growth of megacities in the Third World (Smith 1995, p. 432.) The new approach looks at urban growth within the context of the international division of labor engendered by the global reach of the expansionary logic of competitive capitalism. This process, which translates aspects of competitive capitalism into geographic space, involves "the creation and destruction of land and built environments we term 'cities.' [Moreover,] this leads to concentrations and locational shifts of human populations, infrastructure, and buildings within the urban landscape (resulting in suburbs, neighborhoods, slums, etc.)" (Feagin 1988, p. 23, quoted in Smith 1995, p. 438).

A "new urban paradigm" in the political economy tradition has been put forward by Gottdiener and Feagin (1988) as an alternative to the human ecological and the expansionist paradigms discussed earlier. Rather than treating societies as mere population aggregates or as unified biotic communities, the new urban paradigm treats them as specified by their mode of production. In this view, crisis tendencies and profit generation constitute the core of societal development, which is seen as dominated by the process of capital accumulation. Thus, to take one example, conventional human ecologists like to regard central-city restructuring as a consequence of adaptation to increasing population size and the growing complexity of social organization. They then relate these changes to the size of the metropolitan hinterland. The new urban paradigm, on the other hand, emphasizes the impact of the global economy, multinational corporations, the shift to functional specialization in world-system financial and

administrative activities, the constant subsidization by the state, the efforts of pro-growth coalitions, and changes in labor-force requirements leading to some renovation and central-city gentrification. It tends to focus on power and inequality, the production and reproduction of capital accumulation, crisis adjustment in sociospatial organizations, and such other processes. The following are some of the basic questions that the new urban sociology paradigm seeks to answer: What is the character of power and inequality? How do they relate to "ecological" patterns? How do production and reproduction processes of capital accumulation, as well as the processes of crisis adjustment, manifest themselves in sociospatial organization?

### THE CRISIS OF THE NEW URBAN SOCIOLOGY

However, having apparently supplanted human ecology, the new urban sociology itself appears to be in a state of deep crisis (Hutchinson 1993). Among other things, many of its practitioners are now claiming that the new urban sociology lacks a paradigm equivalent to that of human ecology; that its contribution is critical rather than substantive; that its viewpoint is far more ideological than empirical; and that it lacks a unifying focus, there being as many new urban sociologies as there are its practitioners (La Gory 1993, p. 113). At the same time, while asserting that "what is most salient about the new approach is . . . its direct challenge to the theory and method of ecology," Gottdiener and Feagin (1988) deride the attempt "to pick and choose from some of the new literature . . . areas of compatibility, thereby turning the new approach into a mere footnote of the old" (p. 167). However, in view of the inadequacy of both approaches, La Gory (1993) suggests the use of network analysis as the preferred strategy for devising a revised urban paradigm that is informed by both the strengths and shortcomings of these two perspectives. And noting the considerable conceptual convergence between the two approaches, Smith (1995) argues for a *synthesized* urban theory that will require the fleshing out of Hawley's theory of social organization, technology, and population distributions by incorporating the contributions of the new urban theory regarding the nature and content of the global competitive capitalist system. Thus, Smith claims that while the new

urban sociology can provide human ecology with a better understanding of power and dominance and how class interests play a central role in shaping urbanization, human ecology can help the new urban theory pay more attention to the demographic processes and variables in order to develop a theory of demographic change under global capitalism.

### **ECO-CATASTROPHE AND ENVIRONMENTAL COLLAPSE**

Industrial and industrializing nations are now beset with more or less the same devastating problems of air, land, and water pollution and environmental destruction. Large numbers of lakes and rivers that were not naturally eutrophic have now become so as a result of pollution and chemical runoffs. In the United States, Love Canal and Times Beach, Missouri, made headlines in the 1980s as much as Chernobyl did in 1986 in the Soviet Union. Sulfur dioxide emissions from industrial and power plants cause acid rain that inflicts irreparable damage on buildings, monuments, marine life, trees, and plants. More than 60,000 synthetic chemicals are now on the market, of which a sizable number contaminate the environment and pose health hazards. Over half a million tons of toxic wastes are produced each year in the United States, while the five-year cost of cleaning nuclear waste, which remains dangerously radioactive for thousands of years, may well exceed \$30 billion. The soil, lake water, and groundwater near nuclear power and weapons plants are heavily contaminated with such toxins as mercury, arsenic, and many types of solvents, as well as with deadly radioactive materials such as plutonium, tritium, and strontium-90. The contamination is so bad in eight states that huge tracts of land are said to be totally unfit for human habitation and pose serious health hazards for the surrounding communities. The siting of dump sites in minority communities and the international shipment of hazardous waste to non-Western nations raise serious issues of environmental justice. With an annual production of solid waste that doubled between 1960 and the late 1990s to nearly 225 million tons, the United States is producing more garbage than any other nation in the world and will soon be facing a huge problem of disposal as its 2,300 landfills run out of room and their leachates pose serious threats of toxicity.

The environmental destruction is far more serious and widespread in eastern Europe and the republics of the former Soviet Union. These countries are the sites of some of the world's worst pollution. Lakes and rivers are dead or dying. Water is so contaminated in some areas that it is undrinkable. Chemical runoff and sewage and wastewater dumping have created serious groundwater contamination. Lignite (brown coal), the major source of energy for industry and homes in some of these nations, is responsible for the heavy concentration of sulfur dioxide and dust in the air that has caused serious respiratory problems and additional health damage. The haze-covered cities are an environmental disaster. According to Worldwatch estimates, the former Soviet Union alone accounted for a fifth each of global carbon dioxide and sulfur dioxide emissions—the former are implicated in global warming; the latter are the principal ingredient of acid rain. These environmental problems thus not only span transboundaries, they also cut across ideological labels.

In non-Western nations, a million people suffer acute poisoning and 20,000 persons die every year from pesticides. Pesticides are a major source of environmental and health problems in the United States as well. But the United States alone exports more than half a billion pounds of pesticides that are restricted in or banned from domestic use. The ecology, natural environment, and resources of these non-Western nations are being destroyed and contaminated at a frightening rate. Irreversible damage is being done by large-scale destruction of rainforests and the intensive use of marginal lands, as well as by the imbalances that result from population pressures and the practices of multinational firms and national elites. Desertification now threatens a third of the earth's land surface. Poverty, hunger, starvation, famine, and death are endemic throughout much of the world.

### **RELATION BETWEEN POPULATION AND THE ENVIRONMENT**

What lends urgency to the current population-resource crisis for the West is the fact that while human numbers are declining or standing still at most in industrial societies, they are increasing disproportionately in the rest of the world, a

world divided today not only economically and sociopolitically, but also demographically. The technological mastery of the world has resulted in a higher material standard of living in the West, but it has also spelled economic polarization, ecological ruin, and environmental disasters worldwide. At the same time, hunger, famine, poverty, and overpopulation in the rest of the world have raised critical issues of equity, justice, security, and human survival. While the close link among poverty, population growth, and environmental degradation is invariably highlighted by the neo-Malthusians and the media, the impact of unsustainable patterns of consumption and production on the environment does not receive equal emphasis. Much more disconcerting is the fact that the use of the population argument tends to divert attention away from the role of exploitative and oppressive social institutions and arrangements. Schnaiberg and Gould (1994) find the lack of control over industrial production systems rather than population growth to be the main factor contributing to the underdevelopment of Southern societies. Without minimizing the danger of overpopulation, they find clear historical evidence that the worldwide environmental disruption has been caused not by population growth but by the enormous expansion of production, profits, and surplus in the past century. And based on available evidence, Humphrey and Buttel (1982) have been led to conclude that “[one] of the most important findings to come from the study of the relationship between population size and the environment is the misplaced importance given to world population size as cause of natural resource scarcities and pollution . . . [We do not] imply that world population growth should be . . . neglected as a cause of environmental problems, [but] a fixation on it as the major reason for pollution and energy crises would be sociologically misguided” (p. 60).

Depending on their consensus or conflict orientation, we find that the dominant perspectives on the population–resource dynamic place differential emphasis on the alternative modes of resolving competition over scarce resources. In this respect, the modern division of labor, highlighted by Durkheim, is but one of several modes of resolving competition over scarce resources. Schnore (1965, pp. 12–13) offers a number of alternative survival strategies, which may involve one or a combination of the following: (1) demographic

changes resulting in the elimination of excess numbers (increase in the death rate, decrease in the birthrate, and migration); (2) technological changes that allow for the expansion of the resource base (the exploitation of unused or existing resources, availability of new areas and new resources, resource substitutions, etc.); and (3) organizational changes that allow for the support of larger numbers (occupational and territorial differentiation, revolutionary changes that redistribute the surplus among the many, and reduction in the general level of living to support increased numbers).

For human ecology, the most salient aspect of the population–environment relationship is the way it affects human survival and the quality of human life. Under the impact of the interlocking crises of overpopulation, resource depletion, and environmental degradation, issues of sustainability and survival have come to occupy center stage. Corresponding to the main approaches in human ecology, three broad positions may be identified for discussion: the pro-growth (expansionist), the neo-Malthusian, and the political economy perspectives. Our discussion of these positions is followed by a consideration of the Brundtland Report, issued by the World Commission on Environment and Development, and of the traditional-Gandhian view of the ecological crisis. Extended treatment of the issues involved may be found in Catton (1980), de la Court (1990), Humphrey and Buttel (1982), Mellos (1988), and Schnaiberg and Gould (1994).

#### **The Pro-Growth (Expansionist) Perspective.**

To explain the growth patterns of modern society, this approach builds on the foundations of “the new synthetic theory developed in the biological sciences in the last forty years, . . . mixing in elements of neo-Malthusian theory, Marx’s historical materialism, and modern systems theory” (Lenski 1979, p. 14). It seems quite likely, however, that the basic elements of expansionism, now presented as a natural universal process, were derived from the fundamental American experience of abundance and an open frontier conceived as a process. As Avery O. Craven (1937) put it more than sixty years ago, the basic idea was “that American history . . . presents a series of recurring social evolutions in diverse geographical areas as a people advance to colonize a continent. The chief characteristic is

expansion; the chief peculiarity of institutions, constant readjustment . . . Into . . . raw and differing areas men and institutions and ideas poured from older basins, there to return to a more or less primitive state and then to climb slowly back toward complexity . . . . The process was similar in each case, with some common results but always with ‘essential differences’ due to time and place” (quoted in Potter 1954, p. 145–146.)

In expansionist thinking, scale, complexity, and acceleration—that is, the constant broadening of the limits of the maximum permitted by prevailing circumstances—mark the human–environment encounter. Hawley’s version of human ecology, with its focus on population growth and differentiation as significant processes of continuous change, provides a concise exposition of the pro-growth or expansionist view on the population–resource problematic. Hawley believes that industrial systems have no known upper limits on either the number of specializations or the size of the populations that can be supported. Similar pro-growth sentiments are expressed by other expansionist thinkers. Asserting that resource supplies are finite but unbounded, Hawley (1986, pp. 110–111) questions the neo-Malthusian assumption that overpopulation and overuse will soon exhaust a declining supply of fixed resources. While acknowledging the threat of overpopulation and pollution to the quality of the environment, he points to the inherently expansive nature of populations, technology, and organization that has resulted in a long history of resource expansion through more efficient extraction and use of new and existing resources, new resource development, and resource substitutions. With regard to global food-producing resources, he presents evidence to show that the size of arable land, its productivity, and its agricultural output can be increased beyond the rate of population growth. He blames poverty and the structural conditions that generate it for the chronic food shortages in parts of the world and points to the indispensability of further increments of growth and the creation of central organizations capable of tackling these and other environmental problems. Contrary to the view of the Malthusians, he holds that the expansive power of populations by itself does not cause war, resource depletion, or environmental degradation; it does so only under specific organizational circumstances. Hawley (1986, p. 26) views

these outcomes as the result of the maladaptation or malfunctioning of organization, with disequilibrium opening the possibility for evolutionary change through a movement to a higher level of complexity.

While Colin Clark directly links population numbers to power, Herman Kahn (1974) views population increase as a necessary stimulus to economic growth and believes the earth can easily support 15 billion people at \$20,000 per capita for a millennium. In fact, he believes that the wider the gap between the rich and the poor, the more the riches will percolate downward. In any case, he is unconvinced that the rich would agree to part with their income to ensure a more equitable distribution of wealth. Roger Revelle (1974) believes the earth can actually support nearly thirty times the present population in terms of food supplies and that it would take almost 150 years to hit that mark. While economic development is necessary to provide people with the basis to control their fertility, Revelle is certain the world would drown in its own filth if most of the people in the world were to live at Western standards. Finally, the postindustrial sociologist Daniel Bell (1977) is convinced that economic growth is necessary to reduce the gap between the rich and poor nations. He has little doubt that the “super-productivity society,” with less than 4 percent of its labor force devoted to agriculture, could feed the whole population of the United States, and most of the world as well. In his opinion, pollution exists because the market principle has never been applied to the use of collective goods. Actually, Bell suggests that the government itself could utilize the market to demand a public accounting from all parties on issues of public interest, levy effluent charges for pollutants, and bring effective compliance through the price mechanism.

However, while corporations have shown greater sensitivity and self-regulation, there is evidence that attempts to enforce the “polluter-pays” principle are likely to be resisted or the costs passed on to the public. The negative impact of governmental policies that alleviate energy and resource scarcities is more likely to be felt at the lower socioeconomic levels (Morrison 1976). Dunlap (1979) presents evidence to show that the effects of pollution and the costs of cleaning the environment are borne disproportionately by the poor and may actually serve to reinforce class inequalities.

**The Neo-Malthusian Perspective.** Within the context of actual and perceived resource scarcities worldwide, neo-Malthusianism has gained ascendance since the 1970s over the earlier theories of demographic transition and neoclassical human ecology (expansionism), which were dominant through the 1960s. Based on the Malthusian notion that population invariably outruns food supply because of a lag between the simple arithmetical increases in resources and the exponential rates of population growth, the neo-Malthusians bring in the notion of *carrying capacity* to identify overpopulation as the main threat to human planetary survival. However, in spite of the fact that there is no exact or objective formula for determining the *optimum population*, the neo-Malthusians tie in the notion of carrying capacity—the optimum population that a given environmental resource base can support at a given time—with the idea of an acceptable quality of life that one *insists* on living. Sometimes the theory of demographic transition, discussed below, is also invoked to explain why Western societies have been able to avoid the Malthusian apocalypse by joining declining death rates and birthrates with increasing standards of living, while non-Western societies cannot, given the least likelihood of their ever achieving Western levels of industrial and economic development, and the sheer impossibility of the whole world living at U.S. standards within the constraints imposed by the finite nature of the earth's resources (Daly 1979).

Compounding the environmental effects of the poverty-stricken and “food-hungry” populations of the world are the impacts of massive consumption and pollution by the “energy hungry” nations (Miller 1972, p. 117). The latter rise sharply with even a slight growth in the population of Western nations, where one-quarter of the world's population is responsible for more than 85 percent of worldwide consumption of natural resources and the environmental sinks. Within the United States, a bare 6 percent of the world's population consumes more than half of the world's nonrenewable resources and more than a third of all the raw materials produced. G. Tyler Miller, Jr. (1972, p. 122) believes that the real threat to our life-support system, therefore, comes not from the poor but from the affluent megaconsumers and megapolluters who occupy more space, consume more natural resources, disturb the ecology more,

and directly and indirectly pollute the land, air, and water with ever-increasing amounts of thermal, chemical, and radioactive wastes. While the Club of Rome (Meadows et al. 1972) and the other neo-Malthusians gave a grace period of thirty or so years, Catton (1980) believes we have already overshot the maximum carrying capacity and are now on a catastrophic downward crash course. In any case, he is convinced that our best bet would be to act as if a crash were imminent and to take advance measures to minimize its impact.

However, the basic premises of Malthusian theory have not stood the test of time. With each technological breakthrough, the Western nations have so far been able to raise their carrying capacity through extending their territorial and environmental reach, which now reaches to the ends of the globe. The social and economic forces unleashed by the Industrial Revolution not only telescoped the doubling of human population within a shorter time span, they also brought about ever-rising material standards of living due to astronomical increases in the scale and speed of agricultural and industrial production in the advanced nations. While the Malthusian theory predicts the fall in growth rates of population as a result of rising death rates, this prediction failed to apply during the period of industrial growth. The theory of *demographic transition* was proposed to cover the anomalous results. The theory specifies declining fertility as a consequence of modernization and economic development. However, in the West itself, smaller families became a clear option only after the newly affluent had suffered a major setback in higher living standards during the Great Depression. On the other hand, the downward transition of fertility worldwide appears to be the result of a complex of factors, including the declining role of tradition and religion, rising levels of income, the increasing role of women's education and outside employment, urban residence and industrialization, and the awareness and availability of fertility-control measures (Weinstein 1976, p. 85). Many of the generalizations based on the demographic-transition theory have thus proved to be culturally and historically specific.

At the same time, the “development” of poor nations has created a new set of claimants for the resources needed to maintain the high material standard of living of affluent nations. As the poor nations begin to assert control over their own

resources, try to set terms of their exchange, or resist outside pressures to transform them into “environmental preserves” or the “global commons,” the prospects of conflict, particularly over critical water, mineral, forest, and energy resources, are greatly magnified. Amartya Sen (1981) has looked at the famine situation as essentially a “crisis of entitlement,” not so much because there is lack of food but because many are denied any claims to it because of the very nature of the market economy. In the West, the entitlement revolution has entailed huge welfare expenditures, which could be financed either by economic growth or by direct redistribution of income (Bell 1976, p. 20). For Bell and the neo-Malthusians, the latter is out of the question.

To restore the population–resource balance—with global economic development, equitable distribution of resources, and the perfectibility of man and society now largely ruled out—the neo-Malthusians rely on sophisticated computer models to predict the end to development and limits to economic and demographic growth for the non-Western nations; others favor “sustained environment development”. Still others despair of the effort to avert disaster through population control or the *preventive checks* of moral restraint proposed by Malthus. Instead, they invoke the operation of the Malthusian *positive checks* (wars, famines, pestilence, and natural disasters) and raise the specter of massive famines and die-offs to justify triage, war, secessionist movements, adding sterilants to drinking water, forced sterilization, violent and coercive contraceptive technologies for women, even genocide. In a piece published in 1969 in the Stanford Alumni Almanac, and appropriately titled “The Immorality of Being Softhearted,” Garrett Hardin is quite clear that food would be the worst thing to send to the poor. Nothing short of the final solution will do. “Atomic bombs would be kinder. For a few moments the misery would be acute, but it would soon come to an end for most of the people, leaving a very few survivors to suffer thereafter.” These solutions, which would bring about the decimation of entire populations, have been called *ecofascist*. Such sentiments are by no means uncommon among the neo-Malthusians.

To revert to the neo-Malthusian argument: The tragedy of numbers is compounded by the “free rider,” who derives personal benefits from

the collective efforts of others, and the more serious “tragedy of the commons” (Hardin 1966), where each herdsman will add cattle without limit, ignoring the costs imposed on the others and degrading the land held in common. The “tragedy of the commons” is really the tragedy of individualism run amuck, an individualism from which all constraints of private and common morality have been removed. However, others have been quick to point to the equal or far greater extent of environment pollution and ecological destruction in socialist countries as one more evidence of the inevitable convergence of capitalism and socialism!

Many environmental problems are clearly transideological and transnational. Acid rain, oil spills, the destruction of the ozone layer, the threat of global warming—all call for common responsibility and joint regulation. Ironically, it appears that the expressed concern about the destruction of the global commons through overpopulation or industrial pollution is seldom matched by a parallel commitment by powerful nations to preserve or clean up the environment or provide support for international population-control efforts. Instead, one witnesses a mad scramble to divide up the remaining oceanic and other planetary resources without regard to equity, ecology, or environment. As a result, the air and the oceans, as well as the forests and lands of other nations, are being overexploited or used as garbage and toxic dumps with impunity.

Of no small consequence globally is the environmental impact of waste, widespread corruption at all levels, hoarding and price-fixing, and poor storage, distribution, and transportation networks. Not surprisingly, “formidable mafias based on a triangular alliance between the corrupt bureaucrat, the corrupt politician and the corrupt businessman emerged in all [Indian] States and became a most powerful threat to the conservation of the country’s tree cover” (Vohra 1985, p. 50). When one adds to this list the role of political and economic elites and multinational corporations, and of huge debts, huge dams, and huge arms stockpiles, it becomes clear that poverty, hunger, malnutrition, and starvation may have far more to do with political, cultural, and socioeconomic components of food shortages than with sheer numbers alone. This is not to underestimate the immensity of the population problem or to minimize the difficulty of its solution.

**The Political Economy Perspective.** With the rise of the “consumer society” and the “welfare state” in the West, and the coupling of the “revolution of rising expectations” and the “entitlement revolution” with the impossibility of generalizing Western levels of consumption worldwide, the problematic nature of the relationship between consumption and production has again come to the fore. On the one hand, food aid and food supply have become powerful political weapons globally; on the other, welfare programs constitute a potentially deprivational means of control (Gibbs 1989, p. 453). Saunders points out that today the basic class divisions centered in the relations of production are increasingly being overlaid by “a division, which cuts right across the class structure, between those with access to individual forms of consumption and those who are reliant on collective provision” (1986, p. 232).

Arguing that the issues of collective consumption must be separated from issues of production and class struggle and that urban struggles develop around the question of social consumption, Saunders contends that by recognizing that consumption may generate its own effects, Castells (1985) opens up the possibility of identifying nonclass bases of power and popular mobilization, as well as nonclass forms of popular aspiration and identity (Saunders 1986, p. 226). It is precisely for this reason that the anticommunism of the counterculture during the 1960s was both hailed as a true “revolution without Marx or Jesus” and also seen as a threat to the very existence of a mass-production society. On the other hand, the current neo-Malthusian demands centered on the impossibility of generalizing Western standards of living (level of consumption) to the rest of the world, or on stabilizing consumption at some “optimum” level for achieving a steady state within western societies, also rest on the possibility of regulating production by manipulating mass-consumption patterns.

Barry Commoner (1974) faults socialist as much as capitalist economic theories for neglecting the biosphere as a major factor of production but regards both poverty and population growth as outcomes of colonial exploitation. The world, he believes, has enough food and resources to support nearly twice its present population. The problem, in his view, is a result of gross distributive imbalances between the rich and the poor, and

requires a massive redistribution of wealth and resources to abolish poverty and raise standards of living in order to wipe out the root cause of overpopulation. The alternative to this humane solution is the unsavory one of genocide or natural destruction. A study of environmental destruction in southern Honduras by Susan Stonich (1989) illustrates the power of a perspective that combines the concerns of political economy, ecology, and demography. Her conclusion is that environmental degradation arises from fundamental social structures and is intricately connected to problems of land tenure, unemployment, poverty, and demography.

Stonich identifies political and economic factors and export-promotion policies of international lending institutions and aid agencies as the key elements of a development policy for the whole of Central America that is likely to lead to destruction of the remaining tropical forests, worsen poverty and malnutrition, and increase inequality and conflicts within and between nations. Government policies encourage commercial agriculture in order to earn foreign exchange in the face of mounting external debt, which rose by 170 percent in just seven years to equal three-fourths of the 1986 gross national product. The expansion of export-oriented agriculture and the integration of resource-poor rural households into the capitalist sector, often by ruthless and violent means, concentrates the highest population densities in the most marginal highland areas and encourages intensive land-use and adaptive strategies that accelerate ecological decline. Between 1952 and 1974, as a result of changes in land-use patterns, forest land declined by more than two-fifths and the area lying fallow by three-fifths. In the same period, food-crop production was reduced drastically while pasture area rose by more than half regionally and by more than 150 percent in the highlands, where the number of cattle rose by about 70 percent. By 1974, a third of all rural families were landless; two-fifths were below the subsistence level in 1979.

The result has been the evolution of a class of rich peasants raising export-oriented cattle and cash crops, a class of land-poor and landless peasants and wage laborers, and a class of middlemen operatives who serve as transportation links in an expanding regional and national network. The whole socioeconomic structure has an extremely deleterious effect on the regional ecology and



environment. These patterns are being repeated all over Africa and Asia. Even the “green revolution,” which uses the model of American agribusiness to commercialize agriculture in non-Western countries, provides only a temporary respite. Its recurrent and increasingly high capital requirements for seed, fertilizer, insecticide, water, land, and machinery wipe out the small farmers and landless laborers. It destroys peasant agriculture, exposes the monocultures to destruction by disease and pests, magnifies inequality, and sows the seeds of social instability and rural strife. To those who subscribe to the political economy perspective, the bioecological explanation thus appears to be too simplistic. It overlooks the social context of development and land distribution within which worldwide destruction of traditional agriculture and the rainforests is now occurring.

In sum, these considerations bring out the fact that debates surrounding resource distribution and the control of population and consumption patterns are neither entirely scientific nor purely ecologically inspired. As Barry Commoner (1974) points out, they are political value positions. Will the changes come voluntarily, or will they involve totalitarian nightmares? “Sustainable development” and “traditional ecology” hold out two contrasting possibilities for the future.

#### THE BRUNDTLAND REPORT AND THE PROMISE OF SUSTAINABLE DEVELOPMENT

At the heart of the Brundtland Report of the World Commission on Environment and Development (1987) is the idea of *sustainable development* that has become the rallying point for diverse agendas linking poverty, underdevelopment, and overpopulation to environmental degradation and “environmental security.” The report defines sustainable development as “development which meets the needs of the present without compromising the ability of future generations to meet their own needs.” Its popularity lies in its ability to accommodate the opposing idea of limits to growth within the context of economic expansion, but with a new twist. As pointed out by Gro Brundtland (1989), the “central pivot” of the notion of sustainable development remains “progress, growth, the generation of wealth, and the use of resources.” The imposition of limits on consumption is then

justified in order to protect the resource base of the environment both locally and globally. At the same time, continuous economic growth is considered essential to meeting the needs of the world’s neediest. In fact, the Brundtland Report indicates that “a five-to-tenfold increase in world industrial output can be anticipated by the time world population stabilizes sometime in the next century” (World Commission on Environment and Development 1997).

Sustainable development is also seen as a strategy to enhance global security by reducing the threat posed by conflict and violence in an inequitable and resource-hungry world. To this end, it promotes a commitment to multilateralism, with a call for strong international institutions to ward off the new threats to security and for the collective management of global interdependence (Brundtland 1989, p. 14). As a result, the interests of economic growth and the environment are seen as mutually reinforcing rather than contradictory. The World Commission Report (1987) duly notes that ecology and economy “are becoming ever more interwoven—locally, regionally, nationally, and globally—into a seamless net of causes and effects.” And the 1990 Worldwatch Institute Report predicts that the world will have a sustainable society by the year 2030 (Brown et al. 1990, p. 175). Meanwhile, the challenge, as Arnold (1989, p. 22) states, is to ensure that the sustainability vision “is not trivialized or, worse, used as one more way to legitimize the exploitation of the weak and vulnerable in the name of global interest and solidarity.”

To its credit, the Brundtland Report singled out some forms of economic growth that destroy resources and the environment. The present \$1 trillion expenditure on armaments, for example, constitutes “more than the total income of the poorest half of humanity.” According to 1990 United Nations estimates, military expenditures in developing countries, which account for 75 percent of the arms trade, had multiplied by seven times since 1965 to almost \$200 billion, compared with a doubling by the industrialized countries. In addition, burgeoning debt, adverse trade policies, and internal instability constitute the overwhelming obstacles to sustained development. The fifty-two poorest nations of the world are now burdened with nearly a \$400 billion debt. With Africa’s total debt approaching \$200 billion (half of its overall gross national product and three to four

times its annual income from exports), its average debt repayments amount to more than half the export income. The debt burden forces the African nations to concentrate on monocrop export agriculture to the detriment of food-crop production and pushes hungry and landless farmers and nomads to marginal lands that they overgraze and overexploit in order to survive. However, with respect to fixing the responsibility for deforestation, the Brundtland Report appears to be of two minds (de la Court 1990). In asserting that to most farmers, especially the poor ones, "wood is a 'free good' until the last available tree is cut down," the report partly sides with the "tragedy of the commons" argument, accusing the poor farmers of being "both victims and agents of destruction." On the other hand, it also points to a different cause: "The fuelwood crisis and deforestation—although related—are not the same problems. Wood fuels destined for urban and industrial consumers do tend to come from the forests. But only a small proportion of that used by the rural poor comes from forests. Even in these cases, villagers rarely chop down trees; most collect dead branches or cut them from trees" (quoted in de la Court 1990, p. 68). In any case, the Brundtland Report became the focal point for global environmental efforts in the 1990s, even though in the United States it remained "America's best-kept secret." It undoubtedly played a crucial role in the United Nations conference on the global environment held in Brazil in 1992.

#### **SOCIAL CHAOS AND THE PROBLEMATICS OF SOCIAL AND ECOLOGICAL INTEGRATION**

Dennis Wrong (1994, pp. 295–296) has pointed out how the globalization of the world economy and the accompanying spread of communication and transportation systems to embrace more of the world represent an unprecedented increase in the dissociation of system from social integration, and how the fear of disorder has become more acute in a greatly interdependent world suffering from persistent economic scarcity and a limited capacity for human sympathy with others.

To Scott Greer (1979), the rapid growth of American society, made possible by the increase in societal scale, has led to the following Durkheimian

dilemma: "Increasing interdependence requires more cultural integration than we can manage; growth itself has undermined the cultural support system. While bureaucratization may increase order within a segment of the society, what is to guarantee order among segments?" In the absence of the spirit of consensus generated by war, economic disaster, or a universalized humanity, Greer feels that symbiosis rather than cultural integration may best remedy the fragmentation accompanying the discontinuities of societal growth. Such an approach would not only emphasize trading partners, controlled markets, and formal and informal co-optation, but "given the increasing number of role players who do not 'know their place,' from white working-class men to black college-educated women, such a system will take an awful lot of work by leaders, middlemen, and fixers, as well as some luck" (Greer 1979, pp. 315–316).

Daniel Bell (1976) has made the critical point that while the dominant nineteenth-century view of society as an interrelated web, a structured whole unified by some inner principle, still rules Marxist and functionalist thought, it is no longer applicable. On the contrary, society today is composed of three distinct realms—the technoeconomic structure, the polity, and the culture—each obedient to a different axial (value) principle, each having different rhythms of change, and each following different norms that legitimate different and even contradictory types of behavior. The discordances between these realms are responsible for various contradictions within society. Bell has proposed the creation of a "public household" to overcome the disjunctions among the family, the economy, and the state through the use of modified market mechanisms to further social goals. And given his conviction that the crisis is really a spiritual one of belief and meaning, he recommends the return in Western society of some conception of religion to restore the continuity of generations and provide a ground for humility and care for others. This presents a formidable challenge, however, for Bell is painfully aware that one can neither manufacture such a continuity nor engineer a cultural revolution.

It is doubtful that the problems of order created by the "normal" but dangerous nonrelation between the life-sustaining (ecology–economy) and

order-maintaining (sociopolitical) systems of contemporary society can be corrected, as Bell believes, by the creation of a miraculous hybrid—a “public household”—protected and nurtured by both the polity and the household to serve the interests of the technoeconomic structure and by the side-door entry of the “religious” to provide for the integrative and “higher-order” needs of a socially disjointed and spiritually vacuous society. Even the frantic use of a “holistic” ecological approach is bound to fail if its actual goal is somehow to dominate or desperately hold on to a sundered reality in which everything is so hopelessly *unrelated* to everything else. Since a dependent part cannot grow infinitely at the expense of the others, or usurp the whole for its own purposes without undermining the very conditions of its own existence, the high-powered technoeconomic structure, driven by the insatiable demand for energy, resources, and markets, is inherently disorder-producing and anti-ecological. Its immensity of scale and utilitarian thrust not only destroy traditional structures and sociocultural diversity but also set in motion irreversible and ecologically damaging global processes whose attempted solutions greatly magnify the problems.

That the philosophy and ideology of progress may promote activities inconsistent with sound ecological management and the prospects of human survival is also increasingly being recognized (Peck 1987). But the fact that “‘history’ is not on our side, has no teleology, and supplies us with no guarantees” does not mean to Giddens (1990) “that we should, or that we can, give up in our attempts to steer the juggernaut . . . For we can envisage alternative futures whose very propagation might help them be realised. What is needed is the creation of models of utopian realism” (p. 154). Unwilling and unable to abandon the world-constructionist project, he offers a “post-scarcity” system as perhaps the only possibility but is also bothered by the “dark side of modernity,” the creation of totalitarian power, based on his original insight that totalitarianism and modernity are not just contingently but inherently connected (Giddens 1990, p. 172).

In the stark asymmetry between the disorder-producing and the order-creating powers of the centrophery lies the real “nightmare of reason” and the roots of the current crisis and worldwide chaos. The process is not only incomprehensible

(“complex”), but totally out of control. In this state of affairs, “what is there to love or preserve in a universe of chaos? How are people supposed to behave in such a universe? If that is the kind of place we inhabit, why not go ahead with all our private ambitions, free of any fear that we may be doing special damage” (Worster 1994, p. A3).

### TRADITIONAL ECOLOGY AND THE ENVIRONMENT

Patterns of human social organization and technology use reflect the vision a people have of themselves and of their place in the universe. According to Karl Polanyi ([1947] 1974), the question of how to organize human life in a machine society confronts us with a new urgency:

*Behind the fading fabric of competitive capitalism there looms the portent of an industrial civilization with its paralyzing division of labor, standardization of life, supremacy of mechanism over organism, and organization over spontaneity. Science itself is haunted by insanity. This is the abiding concern.*” (pp. 213–214)

A. K. Saran (1978) does not doubt in the least that the ecological crisis is a self-inflicted one, because an entropic environmental system and an infinitely expanding economy and technology are mutually incompatible. His main argument is that the modern system does not provide a coherent worldview or the proper regulative principle to satisfy the needs of the different orders in a unitive way. In such a system, only a technological solution to the problem of order in the sociopolitical realm can be contemplated, and a piecemeal approach will be relied on to deal with the consequences of a discordant and disharmonious order. In addition to generating tremendous violence, universal disorder, and planetary destruction in the desperate attempt to hold the parts together under its hegemony, such an approach is bound to fail. Since the symbolic is not an integral part of the modern literal consciousness, the attempt to appropriate Mother Earth or other symbols, such as that by the proponents of Gaia, may be ideologically seductive but is both scientifically irrelevant and spiritually vacuous. Neo-Malthusian disclaimers notwithstanding, since evolution has been the master concept to organize and rearrange the world in human terms, the ontology of modern

science is necessarily anthropocentric. Saran's conclusion, therefore, is that there can be no ecological science unless it is grounded in traditional cosmology.

In a study of the Tukano Indians of the northwest Amazon, G. Reichel-Dolmatoff (1977) shows how aboriginal cosmologies, myths, and rituals

*represent in all respects a set of ecological principles . . . that formulate a system of social and economic rules that have a highly adaptive value in the continuous endeavor to maintain a viable equilibrium between the resources of the environment and the demands of society. The cosmological myths which express the Tukano world-view do not describe Man's Place in Nature in terms of dominion, or mastery over a subordinate environment, nor do they in any way express the notion of what some of us might call a sense of "harmony with nature". Nature, in their view, is not a physical entity apart from man and, therefore, he cannot confront it or oppose it or harmonize with it as a separate entity. Occasionally man can unbalance it by his personal malfunctioning as a component, but he never stands apart from it. Man is taken to be a part of a set of supra-individual systems which—be they biological or cultural—transcend our individual lives and within which survival and maintenance of a certain quality of life are possible only if all other life forms too are allowed to evolve according to their specific needs, as stated in cosmological myths and traditions. . . . This cosmological model . . . constitutes a religious proposition which is ultimately connected with the social and economic organization of the group. In this way, the general balance of energy flow becomes a religious objective in which native ecological concepts play a dominant organizational role. To understand the structure and functioning of the ecosystem becomes therefore a vital task to the Tukano. (pp. 9–11)*

However, modernity in its essence has been totally destructive of the traditional vision of human nature, our proper place in the "web of life," and our conception of the ultimate good. Polanyi ([1947] 1974) points out how with the modern separation of "economy" as the realm of hunger and gain,

*[our] animal dependence upon food has been bared and the naked fear of starvation permitted to run loose. Our humiliating enslavement to the "material", which all human culture is designed to mitigate, was deliberately made more rigorous. This is the root of the "sickness" of an acquisitive society that Tawney warned of . . . [The task of] adapting life in such surroundings to the requirements of human existence must be resolved if man is to continue on earth. (p. 219)*

The post-World War II creation of the global economy through the idea of "development" is the other half of the story. As pointed out by Wolfgang Sachs (1990), and in line with Hawley's observation, the concept of development provided the United States with the vision of a new global order in which the former colonies were held together not through political domination but through economic interdependence. But to "define the economic exploitation of the land and its treasures as 'development' was a heritage of the productivist arrogance of the 19th century. Through the trick of a biological metaphor, a simple economic activity turns into a natural and evolutionary process. [Soon] traditions, hierarchies, mental habits—the whole texture of societies—were all dissolved in the planner's mechanistic models . . . patterned on the American way of life" (p. 42).

However, even after nearly two decades of development work, the results were far from heartening. Instead of declining, inequality, poverty, unemployment, hunger, and squalor actually increased manifold in all "developing" countries. To summarize: While the expansionist vision tries to tie ecology, economy, and polity together, and the neo-Malthusians add biology to the list, it is in *sustainable development* that all these orders are firmly knit together—but at a price. The paradoxical nature of the term *sustainable development* arises from the fact that it attempts to combine the contradictory notions of limits to growth and active growth promotion. However, if the key to maintaining ecological integrity is economic self-sufficiency and production for use, then the problem today is surely one of the inhuman scale of enterprise based on the "techniques of degradation" (Marcel 1952), which serve nothing higher than human self-interest, and of the concept of

man as having an economically rather than a spiritually determined nature (Coomaraswamy 1946, p. 2).

Roy Rappaport (1976) has documented how the Maring of New Guinea support as many as 200 people per square mile by cultivating nearly forty-five acres of cleared forest at a time without damaging the environment. But then they look at the world and its "resources" through very different eyes!

### A PROACTIVE ENVIRONMENTAL SOCIOLOGY

Heightened concern with ecology and the environment has now moved into the mainstream of public life as a major priority at the national and international levels. This provides important opportunities for environmental sociologists to contribute to the understanding and solution of these problems. Constance Holden (1989) has highlighted the report of the National Academy of Sciences that outlines an agenda for both micro and macro social scientific studies of "anthropogenic" stresses on the resources and the environment in the north circumpolar region and that has general application. The fragile Arctic region has a great wealth of natural resources. It comprises one-tenth of the earth's area and has a population of 8 million people, of whom a quarter are natives. The report placed major emphasis on interdisciplinary studies, particularly those linking the social and physical sciences and basic and applied research. It emphasized the need for drawing on native knowledge and put urgent priority on issues such as cultural survival and the allocation of scarce resources. It also asked the social scientists to come up with models generalizable to other areas.

An interesting insight concerns how each of the several identities of the Arctic (e.g., as homeland for the natives, as a "colony" exploited for its natural resources, and as the last wilderness) results in a distinctive approach to human-environment relationships. "These approaches have come increasingly into conflict as subsistence hunters and commercial interests vie for limited stocks of fish and game; communities are shaken by boom and bust cycles in scrambles for mineral resources; and rapid modernization has inflicted trauma on native cultures" (Holden 1989, p. 883). The committee identified three areas of interest to the social scientist. In the area of human-environment

relationships, there is a need for studies on conflict resolution to strike a balance between commercial needs and the interests of subsistence hunters, sportsmen, and conservation. The second area pertains to community viability, for which a systematic approach is needed to help develop a physical and social services infrastructure to meet the special climatic needs of the region. A final area pertains to the study of the impact of rapid social change (single-industry cash economy, the snowmobile revolution), which is exacting a heavy price from the local inhabitants in terms of higher rates of alcoholism, suicide, stress, loneliness, accidents, and violence.

The shift from expansionist to neo-Malthusian thinking seemingly implies an attempt to come to terms with the finitude of the total ecosphere and changed global realities, but the underlying assumptions and contradictions are again not made explicit. The overriding concern has been with the protection of industrial and commercial interests, even where these interests clearly come in conflict with the interests of individuals, communities, and their environment. The commitment to protect growth or a certain way of life at any cost has led human ecologists and neo-Malthusians to disregard the minimum well-being or sheer survival of the rest of humankind. In fact, Hawley (1986) admits that while "competition," resulting from demand exceeding the carrying capacity, may account for the exclusion of some contestants from access to their share of a limited resource, it does not shed "any light on what happens to the excluded members of a population after their exclusion" (p. 127). This serious neglect of the concern for the underdog and the undermined is matched by the self admitted tendency of the human ecologists "not [to] confront policy matters directly" (Hawley 1986, p. 127). Given its reliance on large-scale macro forces to explain other macro-level phenomena, the human ecological approach does not readily lend itself to policy considerations. Even otherwise, since it views spatial patterns as the natural outcome of ecological processes rather than the result of power relations, it becomes a conservative force in policy applications (La Gory 1993, p. 112).

Edgar Borgatta (1989, 1996) has sought to develop an important field called "proactive sociology," with a view to closing the wide gap between sociological theory and practice and to save

sociology from sheer irrelevance. Sociological approaches, even when application-oriented, have been largely timid, inactive, or merely reactive. They have restricted their focus to studying how changes in social policy may alter behavior and social situations, but they have refrained from making policy recommendations for designing social structures to serve accepted values. Proactive sociologists, on the other hand, would have the task not only of clarifying values and specifying their meaning but also of assuming the responsibility for making policy recommendations based on an understanding of appropriate models of change. In stark contrast to the pretended value-neutral and value-free stance of mainstream sociology, Borgatta (1996) would, therefore, push to include the consideration of values, as well as the examination of possible structures to implement preferred values, among the central tasks for proactive sociologists.

In human ecology, for instance, it is with reference to the population-consumption problematic that questions of value and their interpretation become evident. Rather than waiting to study only the aftereffects of “all in the path”—the Three Mile Island radiation leak or the Exxon Valdez oil spill—or stepping in at the end of the “issue-attention” cycle, when the problem is historically interesting but socially irrelevant, a proactive sociology would concern itself with the dynamics out of which problems arise, anticipating potential problem areas and their alternative solutions as the means to translate desired values into effective policy. This would involve identifying possible futures and the consequences of action or inaction for their attainment—a policy dimension ignored by sociologists, despite their belief that this may make all the difference in a fast-changing and turbulent world in which the ability to handle and manage change requires the ability to anticipate change and to adapt social structures to changing requirements. To this end, the sociologist would need to ask whether what he or she was doing would make an impact and be useful to society. The fundamental assumption here is that if we know something about the impact of social structure on behavior, we should be able to propose models for changes in social structures that will effectively implement values that have priority status in society (Borgatta 1989, p. 15). Sociologists would then be obligated to “address societal

values more directly by providing alternate models of potential changes and exploring the consequences these changes may produce if identified values are implemented” (Borgatta and Hatch 1988, p. 354).

Following this lead, a “proactive environmental sociology” would have particular application to the “sociology of environmental issues,” which is concerned with the study of environmental movements, resource management problems, and the like. It would broaden the scope of the sociology of environmental issues by focusing specifically on the changes that are required to effectively implement stated values such as equity, environmental justice, rights of the deprived and of “future generations,” resource conservation, equitable sharing of the global commons, the right to clean and healthy environment, sustainable lifestyles and consumption patterns, and the like, and by exploring the possible consequences of these changes. Thus David Mahar, an adviser to the World Bank, has argued that blaming peasant colonists for deforestation is “tantamount to blaming the victim” for “misguided public policies” that promote road building, official colonization of the forest, and extensive live-stock development, and that “purposely or inadvertently encourage rapid depletion of the forest.” This *definition of the situation* led Mahar to propose an “alternative development model” that would put government action on hold so that, based on land-use surveys, lands “found to have limited agricultural potential—virtually whole of terra firme [*sic*] of Amazonia—would be held in perpetuity as forest reserves.” These and other unconventional conclusions are stated by Mahar as his own and carry the disclaimer that they do not necessarily represent the views and policies of the World Bank itself (cited in Hildyard 1989).

This example also brings out Borgatta’s point that a proactive stance may involve the espousing of unpopular positions. It may lead a proactive environmental sociologist to examine the role of established institutions and values (crass individualism and the impact of “anthropocentric,” “cowboy,” “superpower,” and “commercialized conservation” approaches to the use of finite resources and a fragile environment; wasteful consumption patterns; draconian measures of population control; corporate nonaccountability and the global impact of multinationals, the state, and the like) in

order to facilitate the creation of ecologically sustainable social structures that implement positive environmental values. On this basis, a systematic concern with morality and the application of knowledge would lead to a "proactive environmental sociology" that would prompt the sociologist to formulate alternative policies with respect to the set of environmental values or goals that are to be implemented (Borgatta and Cook 1988, p. 17). This approach would also ensure that the applied aspects of "environmental sociology" would flourish within the discipline and not become detached from sociology, as has been the fate of industrial sociology and many other areas in the past (Borgatta 1989).

### OVERVIEW

Environmental sociologists have complained of the lack of a unifying focus within the field and have noted its specialized, fragmented, and dualistic tendencies, which hinder concept and theory development (Buttel 1987, p. 466). This should be a cause for serious concern insofar as the new human ecology is supposed to provide a holistic, integrated understanding of human-environment interactions. In addition to the problems surrounding functionalist as well as Marxist categories and assumptions is the difficulty of adapting bioecological concepts to the human context. Notions such as ecosystem, niche, succession, climax communities, balance of nature, even evolution—none have clear social referents and all pose formidable problems of inappropriate or illegitimate transferal of concepts. Thus, while one finds constant reference to urban or social "ecosystems" in the literature, the wide-ranging, even global, energy-exchange patterns make the boundaries so diffuse that it becomes impossible to locate an urban ecosystem in time and space, at least in biological terms (Young 1983, p. 195). Or, if humans are defined as niche dwellers, the term *niche*, "if adopted directly from biology would produce only one worldwide niche for the entire global species, a result that would render the concept useless. How can the species problem be overcome in adapting such a concept to human ecology?" (Young 1983, p. 795).

Terms such as *the environment* are not easy to define or conceptualize; nor are ecological chain reactions, multiple causal paths, and feedback

mechanisms in complex ecosystems easy to delineate. In recognition of the substantial difference between human and bioecological orders, some human ecologists, such as Hawley, have moved away from bioecological models. Thus, Hawley (1986) is highly critical of the neo-Malthusian application of the "carrying-capacity" notion, on the ground that "while the argument may be suitable for plants and animals, its transfer to the human species is highly questionable" (p. 53). While still shying away from assigning a critical role to human agency, or even a policy-making role to the human ecologist, Hawley has nonetheless broadened the scope of his theory by incorporating culture and norms as ecosystem variables. However, as a commentary on Western-style development and its total disregard for limits, Rappaport has raised more basic objections: To treat the components of the environment as if they were mere resources is to view them exclusively in economic terms and invite "the use and abuse of biological systems of all classes and the neglect of moral and aesthetic considerations in general. Whatever may be meant by the phrase 'quality of life,' exploitation does not enhance it" (1978, pp. 266-267).

Human ecologists, in general, have not dealt adequately with such concerns, nor with the problems of power, domination, and the role of the state and of "values" in human-environmental relationships. At a minimum, one needs to know the role of the state in the regulation, maintenance, expansion, suppression, and "resocialization" of peoples and societies. If ecosystems are constituted of interdependent parts, one needs to know the nature of the reciprocal relationships among the parts and among the parts and the ecosystem. Rappaport (1978) has drawn attention to the maladaptive tendency of subsystems to become increasingly powerful and to dominate and use the larger system for their own benefit, to the detriment of the general interest and the adaptive flexibility of the system. He mentions the dominating positions occupied by huge corporations and the "military industrial complex" as examples. More broadly, Rappaport ties pollution, "resource" depletion, and the diminution of the quality of life and the destruction of its meanings to the scale of modern societies and the voracious appetites of their industrial metabolisms. Thus, while he does

not deny that population increase may have a negative impact on the quality of life, he has little doubt that the real cause of ecosystem destruction and the deterioration of the quality of life is to be found in the way societies are organized, not in their population trends. If that is so, what alternative do humans have?

Within the human ecological perspectives, environmental problems are seen as arising either from the unplanned nature of growth and expansionism, from its attendant externalities and "commons" tragedies, from growth and market restrictions (the pro-growth, expansionist perspective), or from the excess of population over the carrying capacity of the environment (the neo-Malthusian perspective). To restore ecological balance and environmental health, human ecologists place their faith in value consensus, rational planning, systems theory, computer models, economic growth, trickle-downs, market mechanisms, and technological fixes (the pro-growth perspective) or in limits to growth, sustainable development, sticks and carrots, benign neglect, triage, die-offs, outright compulsion, and even genocide (the neo-Malthusian perspective). Within the political economy perspective, on the other hand, the emphasis is on internal contradictions, uneven development, center-periphery relations, capitalist exploitation, the role of multinationals and the state, trade imbalances, and the treadmill of production. To ensure environmental and ecological protection, equitable distribution, and social justice, the proposals from a political economy standpoint range from social revolution to conflict and confrontation, from redistribution to social welfarism and mixed economies.

The political economy perspective is critical of the basic assumption of the Chicago ecologists that changes in population, organization, and the technologies of movement explain expansionary movements and territorial arrangements. By allowing planners to alter spatial forms to dissipate class conflict and social unrest, Smith (1979, p. 255) believes, the perspective becomes a powerful depoliticizing weapon in their hands. He favors "client-centered" planning, which does not assume that "physical structure determines social structure," but holds that both are shaped by the economic and political structures of society, which

provide selective access to opportunities and further discriminatory patterns of land use and investment. Smith therefore offers "conflictual planning" on behalf of the poor and the powerless in order to call attention to the hidden social costs of development and to increase the political costs of pursuing repressive policies disguised as rationally planned allocational, locational, and investment choices. This approach poses three basic questions: "Whose values, interests, and social actions will determine the purpose, pace, and direction of historical change? Can the costs and benefits of historical change be distributed fairly? Can the changes that do occur further the cause of social justice?" (Smith 1979, p. 288).

Schnaiberg (1980) has identified three responses to the contradiction between production expansion and ecological limits: (1) the expansionist, which will be temporary, increasingly unequal, environmentally stressful, and authoritarian; (2) the business-as-usual, which will be unstable, socially regressive, unequal, and of limited environmental value; and (3) the ecological, involving appropriate technology, reduced consumption, and reduced inequality, which will be the most durable but also the most socially disruptive and the least desirable. Schnaiberg's own preference, short of a social revolution, is for a mixed social democratic system like Sweden's, with some production expansion and improved welfare distribution under close state supervision.

However, this solution does not quite address the critical concerns of the environment or the needs of three-fourths of humanity. It presents to the world the anti-ecological model of the "treadmill of production" under a more benign form. It ignores the reality of global inequity, environmental injustice, and global-resource wars—the fact that behind every environmental struggle of today lies a struggle over the expansion of the treadmill of production. It is worth noting that at least in the case of sustainable-yield forestry, Schnaiberg and Gould (1994) no longer look to Sweden as the model. With more than 90 percent of Sweden's native forest now extinct, and its native species replaced by North American tree species, the authors decry that "native plants and animals have been exterminated, sacrificed for an economically-sustainable industry. The Swedish treadmill is



sustained at the expense of the Swedish environment" (1994, p. 210). The real thrust, therefore, is not sustainable development at all but what Vandana Shiva (1990) has called *commercialized conservation*, which puts a dollar value on biodiversity and "justifies conservation in terms of present or future commercial returns" (p. 14). Smith (1979) believes that the environmental problem, like the problem of poverty, has arisen and remains insoluble because of our commitment to existing economic arrangements and institutions and because wealth and power are valued over persons and human need. In fact, the global treadmill operates in such a way that the poor countries often end up financing the development of the rich ones. Thus, during the period from 1982 to 1990, Foster (1995) reports how Third World nations became a net exporter of hard currency to developed countries to the tune of about \$30 billion per year, while also remitting almost \$12.5 billion in monthly payments on debt alone to their creditors in wealthy nations.

While the West is busy presiding over a general reorganization of the global economy and the ways of living throughout the world, the common problem, as Hilary French (1990) points out, is that of finding the proper balance between sufficiency and excess, which she says will be as difficult for the former socialist nations as it has been for the West. In this context, she points out how Czechoslovakia's president, Vaclav Havel, has identified the omnipresent dictatorship of consumption, production, advertising, commerce, and consumer culture as the common enemy.

The uncontrolled greed of the global treadmill has been blamed for the frightening global environmental degradation and for overpowering our sense of responsibility to future generations. "How much is enough" for living a good life (Durning 1992) has become the critical issue today. This concern with sufficiency should lead us to question the equation of consumption with meaningful existence and of the good life with the material standard of living. But such an equation cannot be avoided in a society of abundance, which has to follow the imperative of consumption if its expanding productive capacities are to be put to use. However, since redistribution or any real systemic changes are ruled out, Daly (1979) recommends a control from within based on obedience

to objective value, warning that "if interior restraints on will and appetite diminish, then exterior restraints, coercive police powers, and Malthusian positive checks must increase." (p. 53)

Gandhi, aware both of the fatal attraction and the destructive potential of wanton materialism, saw it as constituting the gravest threat to human freedom, survival, and environmental security. He therefore opted for a simple, nonexploitative, and ecologically sustainable social order. Such a decentralized social order, based on truth and nonviolence, was to be governed by the metaphysically determined optimum levels of wants, technology, and resource use fitted to the requirements of the human scale. In the interim, he demanded that the rich become trustees of the poor in order to serve justice, to mitigate the negative impacts of the differentials of wealth and power, and to avoid class conflicts. His radical vision of a normal social order, nowhere realized as yet, provides a useful yardstick for measuring how ecologically sound and environmentally sustainable a society is in its actual operation. Noting that the world has enough for everyone's needs but not for everyone's greed, Gandhi was convinced that such a social order would come about

*only if the means of production of the elementary necessities of life remain in the control of the masses. These should be freely available to all as God's air and water are or ought to be; they should not be made a vehicle of traffic for the exploitation of others. Their monopolization by any country, nation or group of persons would be unjust. The neglect of this simple principle is the cause of the destitution that we witness today not only in this unhappy land but in other parts of the world too.* (quoted in Sinha 1976, p. 81.)

From this point of view, while there is little disagreement that overpopulation aggravates environmental and other problems, the attempts to eradicate the root causes of social instability, inequality, and poverty are bound to be far more effective in the long run than the impressive but partially effective approaches to population control. Brian Tokar (1988) has pointed out that, historically, rapid increases in population occur when people become dislocated from their traditional land base and become less secure about their personal and family survival. On the other

hand, populations become stable when the future is secure, the infant mortality rate is low, social choices for women are expanding, and parents are not worried about who will support them in their old age.

How to effect the radical changes required to restore the proper ecological balance and preserve the biocultural integrity and diversity of the global "household," but "without the most fantastic 'bust' of all time" (Ehrlich 1968, p. 169), is the formidable challenge and the urgent task facing humankind. This will involve a redirection of the vast, creative human energies away from a self-defeating and ecodestructive expansionist and wasteful orientation, and their rechanneling into life-giving and life-promoting forms of human action and human social organization.

## REFERENCES

- Arnold, Steven H. 1989 "Sustainable Development: A Solution to the Development Puzzle." *Development, Journal of SID* 2/3:21-24.
- Baker, Patrick L. 1993 "Chaos, Order, and Sociological Theory." *Sociological Inquiry* 63:123-149.
- Bell, Daniel 1976 *The Cultural Contradictions of Capitalism*. New York: Basic Books.
- 1977 "Are There 'Social Limits' to Growth?" In Kenneth D. Wilson, ed., *Prospects for Growth: Changing Expectations for the Future*. New York: Praeger.
- Borgatta, Edgar F. 1989 "Towards a Proactive Sociology." Paper presented at the 29th International Congress of the International Institute of Sociology, Rome.
- 1996 "Proactive Sociology: The Need for Relevance in Sociology." *The Annals of the International Institute of Sociology* (New Series, vol. 5). Trieste, Italy: University of Trieste.
- , and Karen S. Cook 1988 "Sociology and Its Future." In Edgar F. Borgatta and Karen S. Cook, eds., *The Future of Sociology*. Newbury Park, Calif.: Sage.
- , and Laurie R. Hatch 1988 "Social Stratification." In Edgar F. Borgatta and Karen S. Cook, eds., *The Future of Sociology*. Newbury Park, Calif.: Sage.
- Brown, Lester R., Christopher Flavin, and Sandra Postel 1990 "Picturing a Sustainable Society." In Lester R. Brown and associates, *State of the World 1990: A Worldwatch Institute Report on Progress Toward a Sustainable Society*. New York: W. W. Norton.
- Brundtland, Gro Harlem 1989 "Sustainable Development: An Overview." *Development, Journal of SID* 2/3:13-14.
- Buttel, Frederick H. 1987 "New Directions in Environmental Sociology." *Annual Review of Sociology* 13:465-488.
- Castells, Manuel 1985 "From the Urban Question to the City and the Grassroots." *Urban and Regional Studies* (Working Paper No. 47). University of Sussex, U.K.
- Catton, William R., Jr. 1980 *Overshoot*. Urbana: University of Illinois Press.
- , and Riley E. Dunlap 1978 "Environmental Sociology: A New Paradigm." *American Sociologist* 13:41-49.
- Commoner, Barry 1974 "Interview on Growth." In Willem L. Oltmans, ed., *On Growth: The Crisis of Exploding Population and Resource Depletion*. New York: Capricorn Books.
- Coomaraswamy, A. K. 1946 *The Religious Basis of the Forms of Indian Society*. New York: Orientalia.
- Craven, Avery O. 1937 "Frederick Jackson Turner." In William T. Hutchinson, ed., *Marcus W. Jernegan Essays in American Historiography*. Chicago: University of Chicago Press.
- Daly, Herman E. 1979 "Ethical Implications of Limits to Global Development." In William M. Finnin, Jr. and Gerald A. Smith, eds., *The Morality of Scarcity: Limited Resources and Social Policy*. Baton Rouge: Louisiana State University Press.
- de la Court, Thijs 1990 *Beyond Brundtland: Green Developments in the 1990s*. New York: New Horizon Press.
- Duncan, Otis Dudley 1964 "Social Organization and the Ecosystem." In Robert E. L. Faris, ed., *Handbook of Modern Sociology*. Chicago: Rand McNally.
- Dunlap, Riley E. 1979 "Environmental Sociology." *Annual Review of Sociology* 5:243-273.
- Durning, Alan 1992 *How Much Is Enough*. New York: W.W. Norton.
- Ehrlich, Paul R. 1968 *The Population Bomb*. New York: Ballantine Books.
- Feagin, Joe R. 1988 *The Free Enterprise City: Houston in Political-Economic Perspective*. New Brunswick, N.J.: Rutgers University Press.
- Foster, John Bellamy 1995 "Global Ecology and the Common Good." *Monthly Review* 46(9):1-10.
- Francis, Roy 1993 "Chaos, Order, and Sociological Theory: A Comment." *Sociological Inquiry* 63:239-241.
- French, Hilary F. 1990 *Green Revolution: Environmental Reconstruction in Eastern Europe and the Soviet Union* (Worldwatch Paper No. 99). Washington, D.C.: Worldwatch Institute.
- Friedmann, J., and G. Wolff 1982 "World City Formation: An Agenda for Research and Action." *International Journal of Urban and Regional Research* 6:309-344.

- Gibbs, Jack P. 1989 *Control: Sociology's Central Notion*. Urbana: University of Illinois Press.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Stanford, Calif.: Stanford University Press.
- Gottdiener, M., and Joe Feagin 1988 "The Paradigm Shift in Urban Sociology." *Urban Affairs Quarterly* 24:163-187.
- Greer, Scott 1979 "Discontinuities and Fragmentation in Societal Growth." In Amos H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- Hardin, Garrett 1968 "The Tragedy of the Commons." *Science* 162:1243-1248.
- Hawley, Amos H. 1979 "Cumulative Change in Theory and in History." In Amos H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- 1984 "Sociological Human Ecology: Past, Present, and Future." In Michael Micklin and Harvey M. Choldin, eds., *Sociological Human Ecology: Contemporary Issues and Applications*. Boulder, Colo.: Westview.
- 1986 *Human Ecology: A Theoretical Essay*. Chicago: University of Chicago Press.
- Hildyard, Nicholas B. 1989 "Adios Amazonia? A Report from Altimira Gathering." *The Ecologist* 19:53-67.
- Holden, Constance 1989 "Environment, Culture, and Change in the Arctic." *Science* 243:883.
- Humphrey, Craig R., and Frederick H. Buttel 1982 *Environment, Energy, and Society*. Belmont, Calif.: Wadsworth.
- Hutchinson, Ray 1993 "The Crisis of Urban Sociology." In Ray Hutchinson, ed., *Urban Sociology in Transition: Research in Urban Sociology* 3. Greenwich, Conn.: JAI Press.
- Kahn, Herman 1974 "Interview on Growth." In Willem L. Oltmans, ed., *On Growth: The Crisis of Exploding Population and Resource Depletion*. New York: Capricorn Books.
- La Gory, Mark 1993 "Spatial Structure and the Urban Experience: Ecology and the New Urban Sociology." In Ray Hutchinson, ed., *Urban Sociology in Transition: Research in Urban Sociology* 3. Greenwich, Conn.: JAI Press.
- Lechner, Frank, J. 1985 "Modernity and its Discontents." In Jeffrey C. Alexander, ed., *Neofunctionalism*. Beverly Hills, Calif.: Sage.
- Lenski, Gerhard 1979 "Directions and Continuities in Social Growth." In Amos H. Hawley, ed., *Societal Growth: Process and Implications*. New York: Free Press.
- Marcel, Gabriel 1952 *Men Against Humanity*. London: Harvill Press.
- Meadows, Donella, Dennis Meadows, Jorgen Randers, and William Behrens 1972 *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. New York: Universe Books.
- Mellos, Koula 1988 *Perspectives on Ecology: A Critical Essay*. New York: St. Martin's Press.
- Miller, G. Tyler, Jr. 1972 *Replenish the Earth: A Primer in Human Ecology*. Belmont, Calif.: Wadsworth.
- Morrison, Denton E. 1976 "Growth, Environment, Equity and Scarcity." *Social Science Quarterly* 57:292-306.
- Namboodiri, Krishnan 1988 "Ecological Demography: Its Place in Sociology." *American Sociological Review* 53:619-633.
- Park, Robert E. 1936 "Human Ecology." *American Journal of Sociology* 42:1-15.
- Peck, Dennis L. 1987 "Introduction to Science and Technology: Critical Assessments of Progress." *Quarterly Journal of Ideology* 11(2):i-iii.
- Polanyi, Karl (1947) 1974 "Our Obsolete Market Mentality." *The Ecologist* 4:213-220.
- Potter, David M. 1954 *People of Abundance: Economic Abundance and the American Character*. Chicago: Phoenix Books.
- Rappaport, Roy A. 1976 "Forests and Man." *The Ecologist* 6:240-246.
- 1978 "Biology, Meaning, and the Quality of Life." In J. Milton Yinger and Stephen J. Cutler, eds., *Major Social Issues: A Multidisciplinary View*. New York: Free Press.
- Reichel-Dolmatoff, G. 1977 "Cosmology as Ecological Analysis: A View from the Rain Forest." *The Ecologist* 7:4-11.
- Revelle, Roger 1974 "Interview on Growth." In Willem L. Oltmans, ed., *On Growth: The Crisis of Exploding Population and Resource Depletion*. New York: Capricorn Books.
- Sachs, Wolfgang 1990 "On the Archeology of the Development Idea." *The Ecologist* 20:42-43.
- Saran, A. K. 1978 "The Traditional Vision of Man." Paper presented at UNESCO Seminar, Hyderabad, India. Mimeo.
- Saunders, P. 1986 *Social Theory and the Urban Question*. New York: Holmes and Meier.
- Schnaiberg, Allan 1980 *The Environment: From Surplus to Scarcity*. New York: Oxford University Press.
- , and Kenneth A. Gould 1994 *Environment and Society: The Enduring Conflict*. New York: St. Martin's Press.

- Schnore, Leo F. 1965 *The Urban Scene: Human Ecology and Demography*. New York: The Free Press.
- Sen, Amartya K. 1981 *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Clarendon Press.
- Shiva, Vandana 1990 "Biodiversity, Biotechnology and Profit: The Need for a Peoples' Plan to Protect Biological Diversity." *The Ecologist* 20/2 (March/April).
- Sinha, Radha 1976 *Food and Poverty: The Political Economy of Confrontation*. London: Croom Helm.
- Smith, David A. 1995 "The New Urban Sociology Meets the Old: Rereading Some Classical Human Ecology." *Urban Affairs Review* 30(3):432-458.
- Smith, Michael P. 1979 *The City and Social Theory*. New York: St. Martin's Press.
- Stonich, Susan C. 1989 "The Dynamics of Social Processes and Environmental Destruction: A Central American Case Study." *Population and Development Review* 15:269-296.
- Tokar, Brian 1988 "Social Ecology, Deep Ecology and the Future of Green Political Thought." *The Ecologist* 18:132-141.
- Vohra, B. B. 1985 "Why India's Forests Have Been Cut Down." *The Ecologist* 15:50-51.
- Weinstein, Jay A. 1976 *Demographic Transition and Social Change*. Morristown, N.J.: General Learning Press.
- World Commission on Environment and Development 1987 *Our Common Future* (The Brundtland Report). New York: Oxford University Press.
- Worster, Dennis 1994 "Chaos Theory Seeps Into Ecology Debate, Stirring Up a Tempest." In Dennis Farney, Great Divides: Scenes From the Politics of American Culture. *The Wall Street Journal* July 11.
- Young, Gerald, ed. 1983 *Origins of Human Ecology*. Stroudsburg, Pa.: Hutchinson Ross.

LAKSHMI K. BHARADWAJ

## HUMAN NATURE

Debates over the nature of human nature have characterized social theory since it emerged in the Renaissance. As Thomas Sowell has argued, these debates generally take two forms: the optimistic and the pessimistic (Sowell 1987). The former position, associated with Rousseau and anarchists such as William Goodwin, holds that humans are essentially good but are turned bad by the institutions of their society. The latter position is rooted

in the assumption that humans are fundamentally egoistic and selfish, thereby requiring either a strong state to regulate them or, in a less pessimistic account, an institution like the market to guide their affairs toward an optimal result.

For sociologists, neither position became the dominant way of thinking about human nature; instead, the plasticity of human experience was emphasized. Durkheim (1973) wrote the most important defense of a pluralistic approach to the subject, one that remains unsurpassed to this day in its clarity of presentation. Human nature is dualistic, he argued, speaking to the needs of both body and soul, the sacred and the profane, the emotional and the cognitive, and other such dualities. We are, in short, what we make ourselves. This version of a flexible approach to human nature would come to characterize contemporary theorists such as Parsons, who spoke of "much discussed 'plasticity' of the human organism, its capacity to learn any one of a large number of alternative patterns of behavior instead of being bound by its generic constitution to a very limited range of alternatives" (1951, p. 32).

In current sociological debates, the plasticity of human nature is emphasized by the general term *social construction*. If one argues that we ought to speak of gender roles rather than sex roles—the former is recognized to be the product of how people arrange their cultural rules, whereas the latter is understood to be fixed biologically—one is making a case for plasticity (Epstein 1988). Indeed, given the importance of feminism in contemporary theory, which tends to argue that "nothing about the body, including women's reproductive organs, determines univocally how social divisions will be shaped" (Scott 1988, p. 2), the strength of a plasticity approach to human nature is probably stronger than ever before.

Current research in many areas of sociology is premised on a social construction approach. Work stimulated by ethnomethodology is one clear case. In contrast to a Chomskian understanding of language as originating in rules hard-wired in the brain, the tradition of conversation analysis examines how human beings in real conversation twist and shape their utterances to account for context and nuance (Schegloff and Sacks 1979; Scheff 1986). Moreover, since the language we use is a

reflection of the way we think, it is possible to argue that the mind itself is socially constructed, that the a priori nature of the way we think is relatively minimal (Coulter 1979). Accounts of sociological practice based on the assumption of plasticity do not end there. It has been argued that homosexuality is not driven by biological destiny but is a socially constructed phenomenon (Greenberg 1988). Morality, as well, can be understood as socially constructed (Wolfe 1989). Underlying a wide variety of approaches to sociology—from symbolic interactionism to social problems—is an underlying premise that human nature is not driven by any one thing.

The only dissent from a general consensus about human nature's plasticity is rational choice theory. At least among economists who believe that economic methodologies can be used to study social institutions such as the family, there is a belief that "human behavior is not compartmentalized, sometimes based on maximizing, sometimes not, sometimes motivated by stable preferences, sometimes by volatile ones, sometimes resulting in an optimal accumulation, sometimes not" (Becker 1976, p. 14). Yet there are many versions of rational choice theory; at least one of them, that associated with Jon Elster, is committed to methodological individualism but is also willing to concede the existence of a "multiple self" (Elster 1986). It is far more common in contemporary sociology to speak of egoism and altruism as existing in some kind of unstable combination rather than giving the priority totally to one or the other (Etzioni 1988).

Arguments about human nature, in turn, are related to the philosophical anthropology that shaped so much social theory. It has been a consistent theme of the sociological enterprise to argue that humans are different from other species. From the emphasis on *homo faber* in Marx and Engels, through Weber's notions about the advantages of culture, to Mead's account of why dogs and other animals are incapable of exchanging significant symbols, humans have been understood to possess unique characteristics that determine the organization of their society. Twentieth-century theorists such as Arnold Gehlen or Helmuth Plessner carried forward this tradition and are increasingly translated and read (for an overview, see Honneth and Joas 1988). Even Niklas Luhmann,

whose work is heavily influenced by biology and cybernetics, can still claim that "the decisive advantage of human interaction over animal interaction stems from this elemental achievement of language" (1982, p. 72).

The most important shift in philosophical anthropology in recent years is a shift from an essentially materialist understanding of human capacities to an essentially mental one. Powers of interpretation and narrative, it has been argued, constitute the essential features of the human self (Taylor 1989). Just as an argument about the plasticity of human nature enables sociology to avoid reduction into psychological categories, an emphasis on the interpretive powers of humans prevents a reduction of sociology to sociobiology and other basically algorithmic ways of thinking about evolution.

As with the issue of plasticity, not all sociologists agree either that there are specific human characteristics or that, if there are, they ought to be understood as primarily mental and interpretive. Sociobiologists argue not only that humans are driven by their genetic structure more than they would like to believe but also that other animals also possess cultural skills. There is therefore no fundamental difference between human and nonhuman species, they are merely points along a continuum (Lumsden and Wilson 1981). Both sociologists and anthropologists, consequently, have argued for the use of sociobiological approaches in the social sciences (Lopreato 1984; Rindos 1986; Wozniak 1984), although there are also critics who question such an enterprise (Blute 1987).

In the 1990s, sociobiology, now often called evolutionary psychology, was something of a growth industry. Edward O. Wilson, who did so much to originate the field, sees the possibility of a unified approach to knowledge, one in which the laws of human interaction could eventually be deduced from the physical and the biological sciences (Wilson 1998). In the meantime, others influenced by evolutionary psychology have argued that cultural products such as language and mind can be understood through the laws of evolution (Blackmore 1999; Lynch 1996). Even such specific cultural products as novels and works of art are formed by processes of cultural selection, it has been argued (Taylor 1996). While still something of a minority point of view, trends such as these are premised on

the idea that the natural sciences, especially biology, offer a better model for understanding human societies than the sociological tradition as derived from Durkheim, Weber, and Mead.

Another challenge to the anthropocentric view that social scientists have taken toward humans has arisen with cognitive science and artificial intelligence. Whereas classical sociological theory compared humans to other animal species, we can now compare them to machines. Computers, after all, process information just as human brains do, use language to communicate, and reason, and can, especially in new approaches to artificial intelligence called connectionist, learn from their mistakes. There are, consequently, some efforts to apply artificial intelligence to sociology just as there are efforts to use the insights of sociobiology (Gilbert and Heath 1986), although here, again, there are strong critical voices (Wolfe 1991; Woolger 1985). In the more recent work of Niklas Luhmann, as well as in the writings of some other theorists, emphasis is placed on information science, systems theory, even thermodynamics—all of which are approaches based on a denial that human systems require special ways of understanding that are different from other systems (Bailey 1990; Beniger 1986; Luhmann 1989).

In spite of efforts to develop sociological theory on the basis of algorithmic self-reproducing systems, it is unlikely that assumptions about the unique, interpretative, meaning-producing capacities of humans will be seriously challenged. It is the capacity to recognize the contexts in which messages are transmitted and thereby to interpret those messages that make human mental capacities distinct from any organism, whether natural or artificial, that is preprogrammed to follow explicit instructions. One reason humans are able to recognize contexts is precisely the plasticity of their mental capacities. The plastic theory of human nature, in short, overlaps with an emphasis on philosophical anthropology to produce an understanding of human behavior that does not so much follow already-existing rules so much as it alters and bends rules as it goes along.

Both understandings of human nature and accounts of specifically human capacities will be relevant to future efforts in sociological theory to reconcile micro and macro approaches. Although there has been a good deal of effort to establish a

micro-macro link (Alexander et al. 1987), the more interesting question may turn out not to be not whether it can be done but whether (and how) it ought to be done. Systems theory and the information sciences provide a relatively easy way to make a link between parts and wholes: Each part is understood to have as little autonomy as possible, so that the system as a whole can function autonomously with respect to other systems. The micro, like a bit of information in a computer program, would be structured to be as dumb as possible so that the macro system itself can be intelligent. Nonhuman enterprises—computers on the one hand and the structure of DNA in other animal species on the other—show that there is a major bridge between the macro and the micro. But the cost of constructing that bridge is the denial of the autonomy of the parts, a high cost for humans to pay.

But the conception of human beings as preprogrammed rule followers is not the only way to conceptualize micro sociological processes. The traditions of ethnomethodology and symbolic interactionism, which are more compatible with notions emphasizing the plasticity of human nature, imagine the human parts of any social system as engaged in a constant process of renegotiating the rules that govern the system. When the micro is understood as plastic, the macro can be understood as capable of existing even in imperfect, entropy-producing states of disorder. Indeed, for human systems, as opposed to those of machines and other species, disorder is the norm, integration the exception. If there is going to be a micro-macro link in sociology, it may well come about not by denying human plasticity and uniqueness, but rather by accounting for the particular and special property humans possess of having no fixed nature but rather a repertoire of social practices that in turn make human society different from any other kind of system.

(SEE ALSO: *Evolution: Biological, Social, Cultural; Intelligence; Sex Differences*)

#### REFERENCES

- Alexander, Jeffrey C., Bernhard Giesen, Richard Munch, and Neil J. Smelser 1987 *The Macro-Micro Link*. Berkeley: University of California Press.

- Bailey, Kenneth D. 1990 *Social Entropy Theory*. Albany: State University of New York Press.
- Becker, Gary 1976 *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- Beniger, James R. 1986 *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, Mass.: Harvard University Press.
- Blackmore, Susan 1999 *The Meme Machine*. Oxford: Oxford University Press.
- Blute, Marion 1987 "Biologists on Sociocultural Evolution: A Critical Analysis." *Sociological Theory* 5:185–193.
- Coulter, Jeff 1979 *The Social Construction of Mind*. London: Macmillan.
- Durkheim, Emile 1973 "The Dualism of Human Nature and Its Social Conditions." In Robert N. Bellah, ed., *Emile Durkheim on Morality and Society: Selected Writings*. Chicago: University of Chicago Press.
- Elster, Jon, ed. 1986 *The Multiple Self*. Cambridge, U.K.: Cambridge University Press.
- Epstein, Cynthia 1988 *Deceptive Distinctions: Sex, Gender, and the Social Order*. New Haven, Conn.: Yale University Press.
- Etzioni, Amitai 1988 *The Moral Dimension*. New York: Free Press.
- Gilbert, C. Nigel, and Christian Heath, eds. 1986 *Social Action and Artificial Intelligence: Surrey Conferences on Sociological Theory and Method 3*. Aldershot, U.K.: Gower.
- Greenberg, David 1988 *The Construction of Homosexuality*. Chicago: University of Chicago Press.
- Honneth, Axel, and Hans Joas 1988 *Social Action and Human Nature*. Cambridge, U.K.: Cambridge University Press.
- Lopreato, Joseph 1984 *Human Nature and Biocultural Evolution*. Boston: Allen and Unwin.
- Luhmann, Niklas 1982 *The Differentiation of Society*, trans. Stephen Holmes and Charles Larmore. New York: Columbia University Press.
- 1989 *Ecological Communication*, trans. John Bednarz. Chicago: University of Chicago Press.
- Lumsden, Charles J., and Edward O. Wilson 1981 *Genes, Mind, and Culture: The Coevolutionary Process*. Cambridge, Mass.: Harvard University Press.
- Lynch, Aaron 1996 *Thought Contagion: How Beliefs Spread Throughout Society*. New York: Basic Books.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- Rindos, David 1986 "The Evolution of the Capacity for Culture: Sociobiology, Structuralism, and Cultural Selectiveness." *Current Anthropology* 27:315–332.
- Scheff, Thomas J. 1986 "Micro-Linguistics and Social Structure: A Theory of Social Action." *Sociological Theory* 4:71–83.
- Schegloff, Emmanuel, and Harvey Sacks 1979 "Opening Up Closings." In Ray Turner, ed., *Ethnomethodology: Selected Readings*. Baltimore: Penguin.
- Scott, Joan Wallach 1988 *Gender and the Politics of History*. New York: Columbia University Press.
- Sowell, Thomas 1987 *A Conflict of Visions*. New York: William Morrow.
- Taylor, Charles 1989 *Sources of the Self: The Making of the Modern Identity*. Cambridge, Mass.: Harvard University Press.
- Taylor, Gary 1996 *Cultural Selection: Why Some Achievements Survive the Test of Time—and Others Don't*. New York: Basic Books.
- Wilson, Edward O. 1998 *Consilience: The Unity of Knowledge*. New York: Knopf.
- Wolfe, Alan 1989 *Whose Keeper?: Social Science and Moral Obligation*. Berkeley: University of California Press.
- 1991 "Mind, Self, Society, and Computer: Artificial Intelligence and the Sociology of Mind." *American Journal of Sociology* 96:1073–1096.
- Woolger, Steve 1985 "Why Not a Sociology of Machines?: The Case of Sociology and Artificial Intelligence." *Sociology* 19:557–572.
- Wozniak, Paul R. 1984 "Making Sociobiological Sense out of Sociology." *Sociological Quarterly* 25:191–204.

ALAN WOLFE

## HUMAN RIGHTS, CHILDREN'S RIGHTS, AND DEMOCRACY

Sociology has existed in universities and as a research and scholarly profession for more than a century, but it has serious problems in justifying its existence in terms of demonstrating that the knowledge it produces is useful. In its early stages of development, sociology was often coupled with social work or with other fields such as anthropology, political science, and economics. However, as universities grew, the social and behavioral science fields became differentiated, and in the post-World War II period sociology not only grew but flourished and developed a major identity in the universities. However, more important, within the field applied interests were often disparaged. In

large part, fields such as social work, family studies, industrial relations, systems analysis, administration, criminology, penology, and others became independent major entities as sociologists denigrated the notion of being involved in applied research and practice. In fact, in the last decades of the twentieth century sociologists appear more often to be seen as possible peripheral members in social research rather than the organizers of major social research, which is frequently motivated by practical, applied considerations.

Still, it is occasionally noted that sociology has been useful, but it is frequently and necessarily a quite modest statement. For example, it is often pointed out that sociology led much of the development and diffusion of statistical and other analytic procedures in the social sciences. Sociology also led the way in developing procedures for systematic data collection, which it has not only shared with other academic disciplines but which has become a part of common daily life, such as polling, investigative reporting, and so on. And the interest in social problems, consistent but peripheral for sociology over the years, has led to some involvement in race and ethnic relations, women's rights, and human rights more generally. Incidentally, while sociology has failed to move in the direction of applied science, psychology has virtually exploded in that direction, leading to concern over the years that that discipline might detach its academic and research interests from the applied professional and clinical aspects. Indeed, the field of psychology has moved into areas of application neglected by sociology, such as family relations, group counseling and therapy, and environmental and ecological studies; other sister disciplines, such as anthropology, economics, and political science have made a similar move in order to increase the scope of development of applied interests.

#### THE RELEVANCE OF SOCIOLOGY

It is not unreasonable to raise the question: If sociology is presumed to be adding to the knowledge of the world, how can it be useful? This question has been raised throughout the history of the discipline, but that does not mean it has gotten serious attention. As a recent positive example, note the work of Turner (1998) and the subsequent comments on the topic. Another recent

consideration of "The Value of Sociology" that provides perspective is the article by Snow (1999), but to some extent much of the discussion about the value of sociology is abstract, albeit broader than some earlier particularistic orientations, which took the view that if sociology was not valuable, it would not be in the college or university curriculum. Two sentences from the abstract of Turner's article (p.243) provide an orientation to answering the question: "It is argued that sociological theory and its applications to real world problems should constitute the core of the discipline . . . . Sociology should redefine and reorient its practice to create an engineering discipline where abstract theoretical principles are boiled down to rules of thumb and used to build or tear down social structures" (p. 243). This is a more direct statement of the idea that sociology should be useful.

Part of the reason for the hesitancy of sociologists to become interested in applied research is the way the field was defined—as "the science of society." Scientists study and analyze their subject matter, but they do not determine what is right and proper for society. The task of determining laws, custom, and mores is not seen by them as being their task, and indeed there is commonly great concern that the values of social scientists should not influence their research and scholarly work. Thus, applied research in sociology has often been defined as examining what is going on in a social situation, in order to determine the underlying social forces at work. If a social policy is involved, the aim of the research may be to see if it is working; with the introduction of a new social policy, research is most typically aimed at seeing whether stated objectives can be confirmed. Much of the interest in applied research in the social sciences became associated with the concept of "evaluation research," particularly in the post-World War II period, with stimulus from sociologists such as Donald Young and Leonard S. Cottrell, Jr., at Russell Sage Foundation. However, stimulus for evaluation research occurred in other ways as well, such as through laboratories and research centers in educational psychology, social psychology, and other behavioral and social sciences, with research often being supported by such government agencies as the National Institute of Mental Health. These thrusts led to two important developments for the social and behavioral sciences.



The first was the growth of methodological sophistication, often identified with the classic codifying work on quasi-experimental research design by Campbell and Stanley (1963) and Cook and Campbell (1979). The second was the notion that evaluation research could be of two types: conventional research, in which the disinterested researcher merely seeks to describe what happens when policy changes are introduced, and "formative" research, wherein the social scientist is not only involved in the evaluation of the consequences of policy changes but is a consultant in the process of identifying the changes needed to accomplish the goals of proposed policy, often with sequential evaluations and changes when the desired results do not materialize with the initial changes.

Throughout this history of social science involvement in applied research, evaluations have often produced negative (i.e., no change) findings, and in some cases even retrogression or reversal on the intended change direction and/or unintended consequences. But, independently of all these results, it can be asked whether the presumed knowledge of sociology and the social sciences is being used effectively, and the answer has been that this does not appear to be the case, at least in the context of one type of criticism. This critique has been advanced many times in different ways, and in the context of this article it will be phrased as not satisfying options of "proactive sociology" (see Borgatta 1991, 1994, 1996). In particular, do sociologists evaluate the social and behavioral structures that exist in society and put them in a comparative perspective? Do sociologists examine the values that exist and describe them objectively, presumably correctly interpreting what they are? Do sociologists examine stated values and see how well social structures that exist implement the identified values? Before proceeding to further consideration of these questions, consider one relevant case.

### DEMOCRACY AND INDIVIDUAL RIGHTS

While no form of government can effectively be defined as maximizing the provision of human rights and children's rights, there is a strong basis for arguing that democratic forms are more likely to do so in the long run than other governmental forms. There are a fair number of nations that at this juncture appear to be reasonably well oriented

to protecting human and children's rights, but attention to a little bit of history should keep one from becoming complaisant about the prospect. For example, in the United States slavery existed until the end of the Civil War, less than a century and a half ago. Until the amendments that followed the Civil War, if one read the Constitution of the United States one might have thought women were entitled to the same rights as men, but they were not. The assumption that women were not entitled to the same rights as men was so ingrained in the culture that it was not even necessary that it be made explicit. It was a world in which men governed; women did not even get the right to vote until half a century after the end of the Civil War. And the history of the expansion of the United States and its relationship with Mexico and with the American Indians may give pause to those who think that simply having a government that is labeled a democracy is one in which "social justice" writ small and large is to be expected.

In a democracy the *de facto* rules do not necessarily provide equal protection under the law for all persons, and the majority or even a smaller segment of society may define the rules. In more recent times, there have been examples of restrictive abuse in the United States, as in the passing of the constitutional amendment that established "prohibition" of alcohol. Aside from attempting to restrict the choice behavior of citizens, the amendment created the era of gangsters and racketeers. Although that constitutional amendment was repealed, there remain numerous laws prohibiting access to certain substances among the most obvious of which are the laws restricting "drugs." More than the outlawing of alcohol during prohibition, the drug laws have created the circumstances leading to the incarceration of the overwhelming majority of the roughly 1.5 million people in the U.S. prison system today. But more generally, in a democracy, all kinds of laws may be passed that restrict individual rights. Germany's progression from a democratic government to Hitler's dictatorship, although an extreme case, should never be forgotten. Similarly, the cases in which liberal or "socialist" regimes have been overthrown with the help of the great American Democracy, only to be followed by totalitarian regimes, may be seen as disturbing.

To continue with the example of the United States, the McCarthy period in the 1950s is an

example of rampant political witch hunts. But consider even the more recent partisan political situation, also termed a witch hunt by some, involving a person as powerful as the president of the United States. A Republican majority in the House of Representatives voted to impeach the president, sending two articles of impeachment to the Senate, which was to "try" the president to determine his guilt or innocence, with a guilty finding by two-thirds of the Senate requiring removal from office. Here we will not discuss the constitutional issues involved, or the moral and legal failings of the president in the matter, but only pay analytical attention to the strange articles of impeachment to show how subtly the power of a majority can create a situation that counters "fairness" and fails to protect even the rights of a president (see *Congressional Record*, December 19, 1998).

Two articles of impeachment were passed in the House of Representatives by an essentially partisan vote of the Republican majority; here *attention will be given only to the question of what is minimally necessary to arrive at a guilty verdict by two-thirds, or 67, of the 100 senators*. Article I states in part: "Contrary to that oath, William Jefferson Clinton willfully provided perjurious, false and misleading testimony to the grand jury concerning one or more of the following . . ." Then four alleged "testimonies" were noted. Now, any sociologist or statistician should be able to analyze the charge and ask the question, "What would the minimum guilt attribution be to create a majority of 67 senators voting guilty?" Simple, since only one alleged testimony is required to judge the president guilty, if 16 senators voted guilty on the first testimony and 17 different (nonoverlapping) senators voted guilty on each of the three other testimonies, the majority of 67 would be created to judge the president guilty of Article I. Thus, the extraordinary situation exists that *theoretically the president could be judged guilty by no more than 17 senators on any of the four testimonies, but he could be judged guilty by 67 senators because of the way the question was formulated!* What a strange way for a democracy to operate. Of course, the situation is even more ridiculous with regard to Article II, which is also defined in the "one or more" fashion, but involves seven "acts," leading to the situation that theoretically the minimum number of senators needed (nonoverlapping) is 9 for three "acts"

and 10 for each of the other four to get a cumulative majority of 67 senators voting guilty. Thus, it can be seen how the Republican congressmen were able to structure a vote in any way they wished, since they constituted the majority in the House of Representatives. It has been said many times that an effective democracy requires constant attention to how power is used and how the rights of all individuals, even of the president, are protected.

Democracy is a common topic of study, and topics relevant to issues of human rights are implicit in many presentations, as, for example, those in a special issue of *International Sociology* ("Democratic Culture," 1999).

### HUMAN RIGHTS AND CHILDREN'S RIGHTS

With regard to human rights and children's rights, attention must be given to providing, as much as possible, for the protection for all persons, and presumably this can happen most dependably in a democracy. However, some things need to be emphasized. Special rights can be legislated for particular individuals or groups, and because they are often accepted as reasonable, this may lead to some abuses and strange situations. For example, *it has been repeatedly stated that with regard to human rights, the attribution of status on an inherited basis does not correspond to a concept of equal opportunity and equal treatment*. If this is the case, then how do equal opportunity and equal treatment operate with regard to other aspects of inheritance? If a parent is rich, owning \$70 billion of assets, the child could inherit roughly \$30 billion in the United States under current law. If the parent is poor, the child may not inherit anything. Further, in being raised in the rich family, the child may have extraordinary forms of support and access to resources that the poor child does not have. The capitalist system may provide the possibility of upward mobility in a democracy, but considering the advantages of the wealthy, the playing fields for the rich and the poor are not equal. Thus, even if the state mandates a strong support system for all children, human rights and children's rights are still subject to a relative notion of the importance of inherited status. *Question: How much attention do sociologists give to this kind of analysis and identification of the values involved?* The sociologist, of course,

should be able to note that the general system even in a democracy can provide a very strong support of inherited status. The acceptance of inherited status is still rampant in much of the world, beginning with any nation in which there is a notion of royalty, which of course includes Great Britain and some other developed nations that are classed as democracies.

This article began by noting that sociologists and other social scientists should become more concerned with the possibility of proactive involvement in social change. As regards the above example, a major concern of the sociologist would be not only to consider and analyze the consequences of existing social policy, but also to *realistically represent* the value system that presumably underlies the social policy. As has been implicit thus far, social policy often does not correspond to the values that are espoused. There is reason for concern in being involved in this proactive orientation, since often values are not consistently aligned and may conflict. *Most dramatically, a traditional view of the family may conflict with the notion that children are due equal opportunity and access to resources in society, the provision of which necessarily becomes a community or state function. The tradition of parental "ownership" of children simply may conflict with providing children with equal access to the resources that exist in the society.* Thus, the analysis of values becomes a critical area for sociological research. Beyond this, if there is a thrust for moving toward "universal" values regarding human and children's rights, then in the proactive orientation sociologists and other social scientists who are ostensibly knowledgeable about how social structures operate should be able to model and propose social structures in order to implement these values.

### HUMAN RIGHTS

Human rights, or the rights and privileges that every person should expect and receive in a just or "good" society, have been a concern of philosophers through the ages, and commonly this interest has been reassociated with such concepts as "social justice." Although this philosophical topic has also been addressed in religious and political codes and documents, the first really broad and broadly supported international statement on the topic is

to be found in the United Nations Universal Declaration of Human Rights (1949), which was approved by the U.N. General Assembly on 10 December 1948. Sociologists might give this document critical attention. Are the rights that are noted consistent with each other; if not, what are the inconsistencies and how might they be resolved? Thus, the study of implicit values becomes a vital area for sociologists. Questions that become obvious include questions about the priorities of the values. The Universal Declaration begins with a preamble that states: "Whereas recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world . . ." It goes on to elaborate on these rights and then states: "The General Assembly proclaims this Universal Declaration of Human Rights as a common standard of achievement for all peoples and all nations, to the end that every individual and every organ of society, keeping this Declaration constantly in mind, shall strive by teaching and education to promote respect for these rights and freedoms and by progressive measures, national and international, to secure their universal and effective recognition and observance, both among the peoples of Member States themselves and among the peoples of territories under their jurisdiction." There followed thirty articles that can be summarized as follows:

1. "All human beings are born free and equal in dignity and rights . . ."
2. All are entitled ". . . without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status . . ."
3. "Everyone has the right to life, liberty and the security of person."
4. "No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms."
5. "No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment."
6. "Everyone has the right to recognition everywhere as a person before the law."
7. "All are equal before the law and . . . equal protection before the law . . ."

- 
- |   |   |
|---|---|
| <p>8. "Everyone has the right to . . . national tribunals . . ."</p> <p>9. "No one shall be subjected to arbitrary arrest, detention or exile."</p> <p>10. Elaboration is provided for entitlement to ". . . fair and public hearing . . ."</p> <p>11. "Everyone charged with a penal offence has the right to be presumed innocent until proved guilty . . . No one shall be held guilty of any penal offence on account of any act or omission which did not constitute a penal offence . . ."</p> <p>12. "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation . . ."</p> <p>13. "Everyone has the right to freedom of movement and residence within the borders of each state. Everyone has the right to leave any country, including his own, and to return to his country."</p> <p>14. "Everyone has the right to seek and to enjoy in other countries asylum from persecution."</p> <p>15. "Everyone has the right to a nationality . . ."</p> <p>16. "Men and women of full age, without any limitation due to race, nationality, or religion, have the right to marry and to found a family. They are entitled to equal rights . . . entered into only with the free and full consent of the intending spouses. The family is the natural and fundamental group unit of society and is entitled to protection by society and the State."</p> <p>17. "Everyone has the right to own property alone as well as in association with others . . ."</p> <p>18. "Everyone has the right to freedom of thought, conscience and religion . . ."</p> <p>19. "Everyone has the right to freedom of opinion and expression . . ."</p> <p>20. "Everyone has the right to freedom of peaceful assembly and association . . ."</p> <p>21. "Everyone has the right to take part in the government . . ."</p> | <p>22. "Everyone . . . has the right to social security . . ."</p> <p>23. "Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment . . ."</p> <p>24. "Everyone has the right to rest and leisure . . ."</p> <p>25. "Everyone has the right to a standard of living adequate for health and well-being . . ."</p> <p>26. "Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit. Education shall be directed to the full development of the human personality and to the strengthening of respect for human right and fundamental freedoms . . . Parents have a prior right to choose the kind of education that shall be given to their children."</p> <p>27. "Everyone has the right freely to participate in the cultural life . . . Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author."</p> |
|---|---|
- Articles 28 to 30 deal with (1) the rights to an international community in which these rights are realized; (2) respect for the rights and freedoms of others; and (3) maintenance of morality, public order, and the general welfare of a democratic society, in a manner that is not contrary to the purposes and principles of the United Nations.
- It is not until Article 29 that one realizes that the Universal Declaration of Human Rights implicitly assumes a democratic form of government in all nations. Therefore, the approval of the Universal Declaration in 1948 by a vote of 48 to 0 with 8 abstentions—from the U.S.S.R., Ukraine, Byelorussia, Poland, Czechoslovakia, Yugoslavia, Saudi Arabia, and the Union of South Africa—is not ununderstandable. However, what is important for sociologists is that this is an attempt at identifying of universals, which presumably can be the
-

basis for examining existing social structures and dealing with explorations of how social structures can be made to support the values involved. But before getting into such an enterprise, it is necessary to reiterate the idea that the values incorporated into the Universal Declaration themselves need to be made explicit.

Before going on to subsequent developments in the area of human rights, a few remarks about the Universal Declaration are appropriate. First, human rights in this document are allocated specifically to individuals. This is in keeping with the notion that only individuals have status before the law. Of course, throughout history, particularly with the beginning of industrialization, the law has given status to entities other than individuals, most obviously corporate structures, including religious organizations, schools and foundations, and unions. Provision of rights before the law to such entities has consequences for individual human rights. Second—and this is a consideration that will become more prominent in subsequent comments—the rights provided are ambiguously allocated to adults, but the Universal Declaration provides no definition of adult status. Thus the definition of a vital concept is omitted—but quite a few other important concepts are also named but not defined. Implicitly, the concepts are subject to definition in the state where the individual resides. To give an example of the ambiguity, let us look at Article 16: “Men and women of full age, without limitation due to race, nationality or religion, have the right to marry and to found a family.” It is not clear what full age means. Is it twelve years of age, fifteen, sixteen, eighteen, or twenty-one? Men and women have a right to marry, but does that mean one of each, or any combination of men and women? In terms of history and ethnic/religious variations, this is not a trivial question. “The family is the natural and fundamental group unit of society and is entitled to protection by society and the State.” What is a family? There is so much variation in any anthropological view of the family as to make this a very vague concept, and the definition is continuing to change in modern industrial societies, what with the arrival of birth control and divorce and sequential multiple marriage. Sociologists and other social scientists obviously can provide some bases for clarification of these statements of rights and values. Other issues

that arise with United Nations documents are emphasized in subsequent sections.

At this point it is appropriate to note that when the Universal Declaration of Human Rights was voted on, it was emphasized that it did not have the status of a treaty. A treaty was in preparation, however, and the International Covenant on Civil and Political Rights (1966), which incorporated much of the Universal Declaration, was adopted by the U.N. General Assembly, in 1966. By 1985, 80 of the 159 members had ratified the treaty. The treaty changed some of the orientation, however, as can be seen by Article 1, which states: “All people have the right to self-determination. By virtue of that right they freely determine their political status and freely pursue their economic, social and cultural development. The States . . . shall promote the realization of the right of self-determination . . .” The treaty begins with an ambiguous statement on how states are to operate, not on individual human rights. Unfortunately, it also contains language that responds to political entities, and one radical interpretation of the language suggests that any political entity effectively has the right to become independent. It needs to be emphasized that human rights—and closely topics related such as genocide, women’s rights, and children’s rights (to be considered below)—have been the subject of many actions in the U.N. General Assembly.

The literature in the field of human rights does not appear to have a major sociological component, but some sociological studies do exist (Buergethal 1997; Gros Espeill 1998; Magnarella 1995; Pace 1998).

### CHILDREN'S RIGHTS

The issues associated with children’s rights are anticipated in part in the movement from the Universal Declaration to the International Covenant. Article 26 of the Universal Declaration states that everyone has the right to education but that parents have a prior right to choose the kind of education to be given to their children. In the International Covenant, more is said about the parental role. Article 18, for example, states: “The States . . . undertake to have respect for the liberty of parents . . . to ensure the religious and moral education of their children in conformity with their own convictions.” Further, Article 23 states:

"Every child shall have . . . the right to such measures of protection required by his status as a minor, on the part of his family, the society and the State." This elaboration on the involvement of the society and the State raises many questions about the rights of children. Let us look at this in terms of education issues.

First, some attention must be given to the question of what constitutes education. A primary concept obviously is that there should be access to literacy—the ability to read and write. But other skills can also be important, as is exposure to what is classed as knowledge. Using a simplistic index, such as the one used in "Why Family Planning Matters" (1999), it is estimated that, of 4.8 billion people in developing countries, 45 percent of females and 54 percent of males of the relevant age group are in secondary school, suggesting that education is still relatively primitive in most of the world. Possibly more striking is that in twenty-six of the thirty-nine sub-Saharan African nations for which data are reported, less than 20 percent of females of the relevant age group are in secondary school. So, somehow much of the world is not yet in tune with the values in the U.N. treaties with regard to education. For sociologists, even these facts can be a challenge. Assuming that the value of education is seen as an important "universal" value, what is necessary to implement it in developing countries?

The question of the education of children goes beyond such components as literacy, and basic knowledge bases to other values that need attention to be clarified. For example, consider the allocation of rights ". . . of parents . . . to ensure the religious and moral education of the their children in conformity with their own convictions." This is indoctrination, and it is directly inconsistent with the notion that the children should be educated. To state the matter crassly, children are subject to indoctrination into specific belief systems, with no opportunity for choice. Children are not exposed to competing belief systems. Beyond indoctrination into belief systems, children are subject to behavior restrictions and physical practices. For example, circumcision of male children is a common form of irreversible mutilation that is imposed by parents and often by the dominant culture. Health-based rationale are sometimes advanced for the practice, without apparent support, and in addition it is suggested that the

practice does little harm. The same cannot be said about the "circumcision" of female children, in which the mutilation involves the removal of sensitive sexual tissue. While there has been substantial negative reaction to female circumcision, it is prevalent in a number of less developed countries. It has to be noted that the "community" and the governing bodies in these countries often support the indoctrination of children. This is not surprising, since many of these countries tend to be either formally or informally theocratic and children are indoctrinated not only through their parents but also through other community support systems. By contrast, in developed countries, where education is more widespread and substantial, belief systems may persist through parents and some community support systems, but modifications are obvious. For example, in Italy, where most of the population are Roman Catholic, the birthrate is the lowest among all developed countries, in spite of the papal condemnation of birth control methods other than abstention or rhythm. It would be naive to believe that Italians are not using modern techniques of birth control.

The treaty basis for children's rights, the Convention on the Rights of the Child (1989), was adopted by the U.N. General Assembly on 20 November 1989. The document contains fifty-four articles, but the greater detail does not alter the problem of value clarification, and in some ways some of the statements of rights create greater ambiguity. For example, Article 14 states: "States . . . shall respect the right of the child to freedom of thought, conscience and religion." From where will the child have acquired thought, conscience and religion?" Note also the role of the parents in Article 18: "Parents . . . have the primary responsibilities for upbringing and development of the child." Much of the document seems to consist of platitudinous statements that the parents will be good and the state will help them. However, sociologists and other social scientists may find it challenging to try to assess structures that might facilitate this. Situations in different nations may be quite different because of underlying differences in the social and economic circumstances. For example, much current political rhetoric in the United States focuses on strengthening the family so that it can better serve children. How is this to be done? Presumably, sociologists would examine the current dichotomy between reality

and rhetoric, noting the shift over the course of the twentieth century from large families with a single working head to small families, sometimes with two working parents, sometimes with single parent. How could the family be strengthened? The conventional/conservative notion of the mother staying home might not be the right answer. One possible answer that needs to be looked at is discarding school patterns that have their roots in the rural past and replacing them with patterns that meet current needs, such as full-day, year-long schooling, which recognizes that parents are not available until the end of the workday and do not have three-month summer vacations. Another possible answer would be to provide supervision and instruction at a level that corresponds to what research says is needed: rather than acquiescing to the economics of minimum support and unrealistic expectations about what schools as now constituted can produce. When something like the latter is suggested, too often the response is that it is not realistic because of cost and convention. On the contrary, and this is an important point: *If that is what the analysis of social scientists finds, there should be insistence that it is realistic.*

While there has been a substantial amount written about the rights of the child, one particularly accessible source is especially recommended—the 1996 issue of the *American Psychologist* that is devoted in large part to the topic (Limber and Wilcox 1996; Melton 1996; Murphy-Berman and Weisz 1996; Saks 1996) With regard to the Convention on the Rights of the Child, it should be noted that at this writing the treaty has been ratified by all member nations of the United Nations except two: Somalia and the United States.

### CONCLUSION

Human rights and children's rights are topics that provide rich potential for research and applied involvement for sociologists and other social scientists. The United Nations has given considerable attention to issues of human rights, and this brief consideration has barely scratched the surface. Sociologists have given attention to some aspects of human rights in detail, including issues of women's rights and racial and other discrimination, but even in these topics the focus has often been narrow considering the varieties of problems and issues in the world.

Special circumstances bring issues of human rights to prominence, and at the end of the twentieth century the issue of genocide has been given substantial attention. Even this issue should be carefully examined by sociologists and social scientists. For example, in a syndicated column Alexander Cockburn wrote: "In 1996, [Madeleine] Albright was asked the following question on CBS-TV's '60 Minutes' by Lesley Stahl: 'We have heard that half a million children have died [in Iraq]. I mean, that's more children than died in Hiroshima. And you know, is the price worth it?' Albright infamously replied, 'I think this is a very hard choice, but the price—we think the price is worth it.' So, back in Nuremberg time, Albright would certainly have been condemned and maybe hanged, if the standards applied to Seyss-Inquart had been leveled against her and if she had been on the losing side. So would her commanding officer, Bill Clinton" (*Seattle Times*, June 3, 1999). The winning side has always defined what is acceptable, but its actions should be subject to objective analysis. Cockburn went on to note: "The protocols of the Geneva Convention of 1949 prohibit bombing not justified by clear military necessity. If there is any likelihood the target has a civilian function, then bombing is forbidden. NATO's bombers have damaged and often destroyed hospitals and health-care centers, public housing, infrastructure vital to the well-being of civilians, refineries, warehouses, agricultural facilities, schools, road and railways. If Slobodan Milosevic goes on trial before before the International Criminal Court, Clinton, Albright and Defense Secretary William Cohen should have their place on the court's calendar, too." It has also been pointed out that the NATO bombings in Serbia were carried out with the defined expectation that there would be no NATO military casualties, and there were none. The Serb military was to be reduced in effectiveness, meaning the destruction not only of military physical resources but also of personnel in barracks and persons in government centers and other related circumstances; there was also the expectation that there would be "collateral damage," that is, civilian casualties. Of course, the Serbs killed Kosovo Albanians during the forced expulsion, but the irony was that after the collapse of the Serbs, the process was reversed. The agenda for social scientists in such matters is one of requiring objectivity in the analysis of the values and the documents and in the reporting of the events. While Cockburn's critique may be

rejected, it should be done so on the basis of the objective facts and the relevance of defined values, not on the basis of the interpretation of the winning side. History is full of uncomfortable facts, and in retrospectively looking at questions of human rights, we have to remember such things as that victory over Japan in World War II came by means of nuclear destruction of two Japanese cities, not two Japanese armies.

Sociologists have not focused sufficiently on human and children's rights with regard to clarifying the "universal" values that the United Nations has promulgated and in which the international community has concurred. Some issues are very complex, loaded with traditional and religious values that directly conflict with those in the Universal Declaration, which tend to be stated in broad and/or general humanistic terms. For example, in a nation that is dominated by a single religion—one that is virtually actually a theocracy—dogmas and traditions supported by the community and in families, may dictate highly restricted status for behavior by women. In a modern industrial nation, this may be viewed as depriving women of their human rights.

The conflict between values that emphasize the good of the broader community and those presumed to be individual human rights is also a critical area for attention by sociologists. One such conflict exists in the values involved in family planning and fertility control. From the community point of view, which may be phrased as "for the good of the society at large," there may be a need for fertility control. For example, although the world (or a nation) may be able to support a tremendous increase in population, such population growth also carries the potential for catastrophic famines and epidemics. The potential problems may result in a general value being placed on the importance of family planning and fertility control by both developed and less developed nations. How is population growth to be attenuated? "Educational" programs have had some success in less developed nations, especially those with a potential for modernization and industrialization. In China, with a tradition of valuing large families, especially male children, the government decided that a strict policy of one child only was necessary to stay population growth. When this policy was implemented, the population was about

90 percent rural, and it currently is still predominantly so. Is such a policy reasonable, and how can it be implemented? The implementation was possible through the form of government that existed, but not without some abuses and also some relaxations of the policy. Here the sociologist can be asked the practical question. How could the policy be implemented without some problems, that is, what would the sociologist instruct the policy makers to do to get the same effect as was accomplished? How are people to be made to adhere to a "one-child-family" policy? Many different types of suggestions can be advanced, including questioning the premise that such a policy is necessary or even appropriate. The argument may shift to empirical questions and estimates, and the like, while the world population continues to increase from 6 billion at the beginning of the twenty-first century to estimates as high as 15 billion by the year 2050. Then the question of quality of life comes into play. At this point, roughly 1.5 billion people live in more or less developed nations and 4.5 billion in less developed nations. What would happen if those in less developed nations used resources at the same rate as those in developed nations? What would happen if, in fifty years, 15 billion people used resources at the level they are currently used in the United States today? The critical analyst might point out that population control is more essential in developed nations, since one additional person there might use twenty (or as much as one hundred) times the resources as one person in a less developed nation. Sociologists obviously have fertile ground for research in examining how social structures might support values that are advanced as "universals" for human and children's rights.

## REFERENCES

- Borgatta, Edgar F. 1991 "Towards a Proactive Sociology." In *Annals of the International Institute of Sociology*. New Series, Vol. 1. Rome, Italy: University of Rome.
- Borgatta, Edgar F. 1994 "Sociology and the Reality of the Press on Environmental Resources." In William V. D'Antonio, Sasaki Masamichi, and Yoshio Yonebayashi, eds., *Ecology, Society, and the Quality of Social Life*. New Brunswick, N.J.: Transaction Publishers.
- Borgatta, Edgar F. 1996 "The Relevance of Sociology in Coping with Societal Problems." In *Annals de L'Institute Internationale de Sociologie*. Vol. 5, pp. 125 - 136. Trieste, Italy: University of Trieste.



- Buergenthal, Thomas 1997 "The Normative and Institutional Evolution of International Human Rights." *Human Rights Quarterly* 19:703-723.
- Campbell, Donald T., and Julian C. Stanley 1963 "Experimental and Quasi-Experimental Designs for Research on Teaching." In N. L. Gage, ed., *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Congressional Record. December 19, 1998, 105th Congress, 2nd Session, H.Res. 611 in the Senate of the United States. Washington, D.C.: U.S. Government Printing Office. (<http://thomas.loc.gov/cgi-bin/query/D?c105:3:./temp/~c105uc7qTc::>).
- "Convention on the Rights of the Child" 1989 *United Nations Document A/Res/44/25*. New York: United Nations.
- Cook, Thomas D., and Donald T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- "Democratic Culture: Ethnos and Demos in Global Perspective" 1999 *International Sociology* 14(3).
- Gros Espiell, Hector 1998 "Universality of Human Rights and Cultural Diversity." *International Social Science Journal* 50:523-534.
- "International Covenant on Civil and Political Rights" 1966 *Department of State, Selected Documents, No. 5. (revised)* (1978). Washington, D.C.: U.S. Government Printing Office.
- Limber, S. P., and B. L. Wilcox 1996 "Application of the U.N. Convention on the Rights of the Child to the United States." *American Psychologist* 51:1246-1250.
- Magnarella, Paul J. 1995 "Universal Jurisdiction and Universal Human Rights: A Global Progression." *Journal of Third World Studies* 12:159-171.
- Melton, G. B. 1996 "The Child's Right to a Family Environment." *American Psychologist* 51:1234-1238.
- Murphy-Berman, V., and V. Weisz 1996 "Convention on the Rights of the Child." *American Psychologist* 51:1231-1233.
- Pace, John P. 1998 "The Development of Human Rights Law in the United Nations, Its Control and Monitoring Machinery." *International Social Science Journal* 50:499-511.
- Saks, M. J. 1996 "The Role of Research in Implementing the U. N. Convention on the Rights of the Child." *American Psychologist* 51:1262-1266.
- Snow, David A. 1999 "The Value of Sociology." *Sociological Perspectives* 42(1):1-22.
- Turner, Jonathan H. 1998 "Must Sociological Theory and Sociological Practice Be So Far Apart?: A Polemical Answer." *Sociological Perspectives* 41(2):243-277.

"Universal Declaration of Human Rights" 1949 *Department of State Publication 3381*. Washington, D.C.: U.S. Government Printing Office. (Note: Copies of U.N. publications and documents can be located on the Internet, both at U.N. sites and university sites, e.g., <http://www.un.org/overview/rights.html>.)

"Why Family Planning Matters" 1999 *Population Reports*, Series J, Number 49, Vol. 27.

EDGAR F. BORGATTA

## HUMANISM

Humanism in its broadest sense can be traced to the philosophical movement that originated in Italy in the second half of the fourteenth century and that affirmed the dignity of the human being. Although over the centuries there have been numerous varieties of humanism, both religious and nonreligious, all have been in agreement on the basic tenet that every human being has dignity and worth and therefore should be the measure of all things.

Humanism, as practiced in sociology, starts from two fundamental assumptions. The first of these is that sociology should be a moral enterprise, one whose fundamental purpose is to challenge the views and conditions that restrain human potential in a given society. The second is that sociology should not be defined as a scientific discipline that embraces "positivism"—the position that facts exist independently of the observer and that the observer should be a value-neutral compiler of facts.

Sociologists operating in the humanist tradition hold that the study of society begins with the premise that human beings are free to create their social world and that whatever impinges on that freedom is ultimately negative and destructive. They argue that the use of one of the traditional methodological tools of science—dispassionate observation—has not only taken sociology away from its Enlightenment origins in moral philosophy but is based on a faulty epistemology.

Although diverse theoretical frameworks, such as Marxism, conflict theory, phenomenology, symbolic interaction, and feminist sociology, can all be said to have some form of a humanistic orientation

as a part of their overall framework's, humanism in sociology is most readily identified with those sociologists who in their teaching, research, and activism gravitate around the Association for Humanist Sociology (AHS), which was founded in 1976 by Alfred McClung Lee, Elizabeth McClung Lee, and Charles Flynn.

The fundamental underpinnings of sociological humanism can be traced back to two traditions that came out of the Enlightenment: moral philosophy and empiricism. Although Modern sociologists see these traditions as separate, to the Enlightenment French and Scottish philosophers (collectively known as the *philosophes*) they were intertwined and interdependent. The *philosophes* called for a fusion of morals and science, for a social science that sought to liberate the human spirit and ensure the fullest development of the person. It is this emphasis on moral philosophy and empiricism, as modified by German idealism and more recently by the American philosophical tradition of pragmatism, that constitutes the foundations of humanism in sociology, today.

#### THE ENLIGHTENMENT AND THE LEGACY OF SOCIOLOGICAL HUMANISM

Although the Enlightenment *philosophes* initiated the enterprise of modern sociology through their call for the application of scientific principles to the study of human behavior (Rossides 1998), humanist sociologists stress that the *philosophes* were first and foremost moral philosophers. Science and morality were to be fused, not separated; the "is" and the "ought" were to be merged into a moral science, a science for the betterment of humankind. It was Jean Jacques Rousseau, with his arguments against inequality and for the dignity of the person, who best represents this tradition of moral science tradition Rousseau (1755–1985) started with the fundamental assumption that all people are created equal and from this formulated a radical system of politics. Rousseau and the *philosophes* were wedded to the idea that individual liberty and freedom prospered only under conditions of minimal external constraint that had to be consensually based. In the eighteenth century, the *philosophes* articulated their doctrine of individual liberty and freedom chiefly in the idiom of natural rights (Seidman 1983).

The *philosophes* held that the most important value was the freedom of the individual in a humane society that ensured this freedom. Not having any developed psychology of the individual, of the subjective side of human behavior, or of how institutions are formed, they could not go beyond this modest beginning. They could not fashion a full-blown vision of the free individual within a society based on the principle of human freedom.

This tradition of a moral science is overlooked by contemporary sociologists, who instead focus on the empiricism of the *philosophes*; but although empiricism without doubt played the greatest role in the rise of social science, it is only one part of what the *philosophes* advocated. In their dismissal of the moral science tradition and in their virtually unquestioning embrace of the positivism of Comte, Spenser, Durkheim, and the other early founders of sociology as a discipline, contemporary sociologists overlook the *philosophes*' concern that there was an epistemological dilemma inherent in the new empirical science they envisioned. If a social science was to arise out of the Enlightenment, it had to have a new conception of knowledge, one that rejected Greek and medieval Christian epistemology. The Aristotelian view held that a definite entity resided within the human body, an entity that passively observed what was going on in the world, just as a spectator does. The observer sees a picture of the world, and it is this passive observation that constitutes *experience*. Science, in the Aristotelian model, was the process of observing objects as they were thought to be conceived in the human mind. Following Newton, the world was to be understood in terms of mathematical equations by means of axioms that were put in the minds of humans by God and that enabled the mind to picture reality. John Locke's *Essay Concerning Human Understanding* ([1690] 1894) represented an early attempt to show that the extreme rationalist notion that the world precisely followed mathematical axioms was in error. Locke argued that first principles did not exist a priori but came from the facts of experience. Locke, however, became caught up in the epistemological dilemma that experience was mental, and not physical, and therefore still had to be located in the "unscientific" concept of mind. This led Locke, like David Hume (1711–1776), to conclude that an exact science of human behavior was unattainable (Randall 1976). Only probabilistic knowledge could be arrived at,

and this could only modestly be used to guide humankind.

Although the epistemological dilemma posed by Locke and other Enlightenment thinkers was real to them, the development of sociology in France, England, and later in the United States discarded these concerns and embraced positivism as the cornerstones of the discipline. Sociology, however, developed differently in Germany, and it is through German social science that the tradition of humanism in sociology was kept alive.

### GERMAN IDEALISM

German social science, unlike its English, French, and later American counterparts, was much more influenced by idealism than by empiricism. This influence is due to two giants of philosophy: Immanuel Kant and Georg William Freidrich Hegel.

**Immanuel Kant (1724–1804).** Kant ([1781] 1965) was interested in answering the basic question of how autonomy and free will were possible in a deterministic Newtonian universe. His answer led him back to Locke's epistemological dilemma. According to Kant, there is a *phenomenal* and a *neumenological* self. Kant called the world as experienced by the individual *phenomena* and the thing in itself *neumena*. Since science is concerned with experience, Kant relegated it to the study of phenomena. The neumenon is beyond the scientist's realm of inquiry, because Kant wanted to claim the neumenon for the moral philosopher. For Kant, the basis of moral philosophy was to be found in the human mind; moral law located a priori in the mind and can be deduced rationally. Kant, like Locke before him, was faced with the dilemma of how the mind works.

Kant's explanation was that objects of scientific investigation are not simply discovered in the world but are constituted and synthesized a priori in the human mind. The external world that human beings experience is not a copy of reality, but something that can only be experienced and understood in light of a priori forms and categories. According to Kant, these forms and categories determine the form but not the content of external reality. Causation is a product of the mind and is a necessary precondition for the conception of an orderly universe.

Kant believed that he had solved the problem of knowledge through the forms and that he could do the same for ethics. Morally right action, too, is located in the mind. Going back to Rousseau and before him to the fourteenth-century humanists, Kant ([1788] 1949) focused on the dignity of the human being. His notion of the *categorical imperative*, that each person be treated as an end and never as a means, solidifies the importance of the person as the cornerstone of philosophical inquiry and of humanism. Natural rights are part of the neumenal world, part of the moral self. Kant thus began to look to the mind, to the self, as the primary origin of society. Moral values come from human consciousness: but, lacking a viable theory of consciousness, Kant could go only so far. It was Hegel who took up the challenge and subsequently made further progress toward the development of a humanistic orientation in sociology.

**Georg Wilhelm Freidrich Hegel (1770–1831).** Hegel was well versed in the social and moral philosophy of his day, and was particularly steeped in the work of Kant, who was the dominant figure in Germany philosophy at the time. Although Hegel ([1821] 1967) held that Kant's epistemology was successful in explaining how scientific knowledge was possible, he differed with Kant by rejecting Kant's belief that the categories were innate and therefore ahistorical. For Hegel, the human mind has to be understood in the context of human history. Human reasons is the product of collective action and as such is constantly evolving toward an ultimate understanding of its own consciousness. There are adumbrations of the sociology of knowledge in Hegel's view, specifically in his arguments that the Kantian categories, which are used to make sense of the world, change as the political and social climate changes. Hegel is very close to modern sociology in other aspects of his thought, and it is extremely unfortunate that he is so often dismissed because of his ultimate reliance on the metaphysical assumption that total understanding would only come with the realization of the absolute spirit in human history. When Hegel's contributions are mentioned, it is usually only as having had an influence on Marx, and even then it is inevitably pointed out that Marx turned Hegel "upside down." These interpretations overlook the fact that Hegel was the first modern theorist to develop an antipositivistic critical approach to society. Hegel rejected positivism because of its

overreliance on empiricism, which forces the individual to find sense impressions meaningful. As was Kant's philosophy, Hegel's philosophy was humanist at its core.

Also overlooked is the fact that Hegel not only offered an active epistemology but a social one as well. This socially based epistemology (the categories are conditioned by social and political factors) also led him to conceive of a socially based moral philosophy. Whereas Kant held that the concept of freedom was based in the mind of the individual, Hegel, like Rousseau, believed that freedom could only be expressed in terms of a supportive community.

Perhaps Hegel's most important contribution to modern social science is that he was among the first theorists to look at the social development of self, something that makes him a forerunner of humanist sociology. For Hegel ([1807] 1967), the self must be understood as a process, not as a static reality. The self develops as the mind negotiates intersubjectivity. We experience ourselves as both an intending subject and as an object of experience. The mind develops and strives for ultimate truth in this context, which, to Hegel, is freedom. The essence of being is, therefore, a self-reflexive struggle for freedom. Hegel's idealism led him to conclude that objective analysis is always mediated by subjective factors and points toward freedom. In Hegel, there is the outline of a critical, humanistic sociology. He offered an active, antipositivistic, socially conditioned epistemology, with an emphasis upon freedom through the seeking of self knowledge, along with a critique of any non-morally based society.

### PRAGMATISM AND HUMANISM

The importance of pragmatism for a humanistic orientation in sociology lies in its assumption of an active epistemology that undergirds an active theory of the mind, thereby challenging the positivistic behaviorism of the time made popular by the likes of John B. Watson. For the pragmatists, how the mind comes to know cannot be separated from how the mind actually develops.

George Herbert Mead ([1934] 1974) exemplifies the pragmatists' view concerning the development of mind. Consciousness and will arise from problems. Individuals ascertain the intentions of

others and then respond on the basis of their interpretations. If there were no interactions with others, there would be no development of the mind. Individuals possess the ability to modify their own behavior; they are subjects who construct their acts rather than simply responding in predetermined ways. Human beings are capable of reflexive behavior: that is, they can turn back and think about their experiences. The individual is not a passive agent who merely reacts to external constraints, but someone who actively chooses among alternative courses of action. Individuals interpret data furnished to them in social situation. Choices of potential solutions are only limited by the given facts of the individual's presence in the larger network of society. This ability to choose among alternatives makes individuals both determined and determiners (Meltzer et al. 1977).

What Mead and the pragmatists stressed was the important notion that the determination of ideas—in particular, the impact of social structure on the mind of an individual—is a social-psychological process. Thinking follows the pattern of language. Language is the mechanism through which humans develop a self and mind, and language is social because words assume meaning only when they are interpreted by social behavior. Social patterns establish meanings. Language sets the basis for reason, logic, and by extension all scientific and moral endeavors. One is logical when one is in agreement with one's universe of discourse; one is moral when one is in agreement with one's community. Language is a mediator of social behavior in that with a language come values and norms. Value judgments and collective patterns exist behind words; meaning is socially bestowed.

Although Mead was the most important pragmatist for understanding the development of self, the epistemology of pragmatism was most precisely formulated by John Dewey (1929, 1931). Dewey's epistemology represented a final break with the notion that the mind knows because it is a spectator to reality. For Dewey, thought is spatiotemporal. Eternal truths, universals, *a priori* systems are all suspect. Experience is the experience of the environment—an environment that is physical, biological, and cultural. Ideas are not Platonic essences but rather are functional to the experience of the individual (Dewey 1931). This position is antipositivistic in that the mind deals

only with ideas and, therefore, does not ponder reality, but only ideas about reality. Truth is not absolute but is simply what is consistent with experience.

The individual is engaged in an active confrontation with the world; mind and self develop in a social process. The pragmatists provided an epistemological justification for freedom (the basic tenet of humanism). The mind develops in a social context and comes to know as it comes into being. Any restriction on the freedom of the mind to inquire and know implies a restriction on the mind to fully develop. The pragmatists rounded out Hegel's ([1807] 1967) view that ultimate truth is freedom by showing that the mind needs freedom to develop in a social context. Epistemology and freedom are inseparable.

Pragmatism, by joining epistemology and freedom via the social development of mind, also provides a solution for the seeming incompatibility between an instrumental and an intrinsic approach to values. The value of freedom is instrumental in that it is created in action (the action of the developing mind); but it is also intrinsic in that the mind cannot fully develop without the creation of an environment that ensures freedom. This integrated epistemological framework provides the basis for a humanistic methodology for sociology.

#### PRAGMATISM, METHODOLOGY, AND HUMANISM

Dewey and Mead developed a methodology that gave social scientists a different frame of reference from that of the "traditional scientific methodology." Flexibility was the main characteristic of their pragmatic methodology—it did not offer specific forms or languages to which social problems had to be adapted. Instead, the form and language of the method grew out of the problem itself. The social scientist, thus, fashions his or her own methodology depending upon the problem studied. New concepts and methodologies arise from efforts to overcome obstacles to successful research. Techniques are developed that enable the researcher to be both a participant in and observer of social structures. There is an instrumentalist linkage between theory and practice as it is incorporated into the humanist sociologist's life. This is what Alfred McClung Lee (1978), a leading humanist sociologist, meant when he wrote: "Sociologists cannot

be persons apart from the human condition they presumably seek to understand" (p. 35). This is what C. Wright Mills (1959) meant when he wrote: "The most admirable thinkers within the scholarly community . . . do not split their work from their lives." (p. 195).

For the humanist sociologist, the main purpose in amassing a body of knowledge is to serve human needs; knowledge must be useful. By accepting this dictum, humanist sociologists extend the analysis of what *is* to the analysis of what *ought* to be. Knowledge should provide answers for bringing about a desired future state of affairs, a plan that can be achieved through the methodological insights of pragmatism whereby the researcher is both participant and observer.

The dilemma of which values to choose is answered by opting for the pragmatist's emphasis upon responsibility as a moral standard which assumes that a fundamental quality of human beings is their potentiality for ethical autonomy. People not only *are* but *ought* to be in charge of their own destiny within the limits permitted by their environment. Individual character development takes place to the extent that persons can and do decide on alternative courses of action (Dewey 1939).

Pragmatism is grounded upon an assumption of freedom of choice. However, choice among alternatives is always limited. It is in pointing out these limitations in the form of power relations and vested interests behind social structures that humanist sociology builds upon pragmatism and thereby confronts the basic sociological criticism of pragmatism—that it lacks a viable notion of social structure. Humanist sociology seeks to fashion a full-blown vision of the free individual within a society based on the principle of human freedom.

#### HUMANIST SOCIOLOGY TODAY

Humanist sociology has moved beyond pragmatism via its attempt to spell out the social structural conditions for the maximization of freedom. Humanist sociology is based on moral precepts, the foremost of which is that of *freedom*—"the maximization of alternatives" (Scimecca 1995, p.1). This is assumed to be *the* most desirable state for human beings—and the goal of sociology is to work toward the realization of conditions that insure this

freedom. Given its Meadian theory of self (an active theory of self that chooses between alternatives), humanist sociology is concerned with how this is best realized within a community. Humanist sociology begins with the fundamental assumption that all varieties of humanism hold—that individuals are the measure of all things. Using a nonpositivistic epistemological foundation, humanist sociologists employ their methods of research to answer the most important question that can be asked by a humanist sociologist about human behavior, the one originally raised by the Enlightenment *philosophes*: How can social science help to fashion a humane society in which freedom can best be realized?

## REFERENCES

- Dewey, John 1929 *The Quest for Certainty*. New York: Minton, Balch and Company.
- 1931 *Context and Thought*. Berkeley: University of California Press.
- 1939 *Freedom and Culture*. New York: G.P. Putnam's Sons.
- Flynn, Charles 1976 *Association for Humanist Sociology Newsletter* 1(1) pp 1-2.
- Hegel, Georg Wilhelm Freidrich (1807) 1967 *The Phenomenology of Mind*. New York: Harper Colophon Books.
- (1821) 1967. *The Philosophy of Right*. Oxford: Clarendon Press.
- Kant, Immanuel (1788) 1949 *The Critique of Practical Reason*. Chicago: University of Chicago Press.
- (1781) 1965 *The Critique of Pure Reason*. New York: St. Martin's Press.
- Lee, Alfred McClung 1978 *Sociology for Whom?* New York: Oxford University Press.
- Locke, John (1690) 1894 *Essay Concerning Human Understanding*. Oxford: Oxford University Press.
- Mead, George Herbert (1934) 1974 *Mind, Self and Society*. Chicago: University of Chicago Press.
- Meltzer, Bernard N., John W. Petras, and Larry T. Reynolds 1977 *Symbolic Interactionism: Genesis, Varieties, and Criticisms*. Boston: Routledge and Kegan Paul.
- Mills, C. Wright 1959 *The Sociological Imagination*. New York: Oxford University Press.
- Randall, John Herman, Jr. 1976 *The Making of the Modern Mind*. New York: Columbia University Press.
- Rossides, Daniel W. 1998 *Social Theory: Its Origins, History, and Contemporary Relevance*. Dix Hills, N.J.: General Hall.
- Rousseau, Jean Jacques (1755) 1985 *Discourse on Human Inequality*. New York: Penguin Books.
- Scimecca, Joseph A. 1995 *Society and Freedom*, 2nd ed. Chicago: Nelson-Hall.
- Seidman, Steven 1983 *Liberalism and the Origins of European Social Thought*. Berkeley: University of California Press.

JOSEPH A. SCIMECCA

## HYPOTHESIS TESTING

See Scientific Explanation; Statistical Inference.

# I

---

## IDEAL TYPES

*See* Typologies.

## IDENTITY THEORY

Identity theory, in the present context, has its referent in a specific and delimited literature that seeks to develop and empirically examine a theoretical explanation, derived from what has been called a structural symbolic interactionist perspective (Stryker 1980), of role choice behavior. It is only one of a large number of formulations—social scientific, therapeutic, humanistic—in which the concept of identity plays a central role, formulations having their roots in a variety of disciplines ranging from theology through philosophy to political science, psychology, social psychology, and sociology. Further mention of these diverse formulations will be forgone in order to focus on identity theory as specified above; those who desire leads into the literature of sociology and social psychology to which identity theory most closely relates will find them in McCall and Simmons (1978); Stryker (1980); Weigert (1983); Stryker and Statham (1985); Hewitt (1997a, 1997b); MacKinnon (1994); and Burke and Gecas (1995).

The prototypical question addressed by identity theory, phrased illustratively, is: Why is it that one person, given a free weekend afternoon, chooses to take his or her children to the zoo while another person opts to spend that time on the golf course with friends? The language of this prototypical question implies a scope limitation of the

theory that is important to recognize at the outset of the discussion. The theory is intended to apply to situations where alternative courses of action are reasonably, and reasonably equivalently, open to the actor. A defining assumption of the symbolic interactionist theoretical framework is that human beings are actors, not merely reactors. Identity theory shares this assumption, which recognizes the possibility of choice as a ubiquitous feature of human existence. At the same time, however, identity theory recognizes the sociological truth that social structure and social interaction are equally ubiquitous in constraining—not in a strict sense “determining”—human action. That constraint is variable. It may be true in an abstract and philosophical sense that people are “free” to act in any way they choose in any situation in which they may find themselves, including choosing to endure great punishment or even death rather than to behave in ways demanded by others; but surely it is entirely reasonable to presume that jailed prisoners have no viable options with respect to many—likely most—facets of life and in any event have fewer viable options than persons who are not jailed. Identity theory has more to say on those the latter persons than on the former, and more to say on those—perhaps few—aspects of life about which the former do have reasonable choice than on those many aspects of prisoner life where options, as a practical matter, do not exist.

As a derivative of a symbolic interactionist theoretical framework, identity theory shares a number of the assumptions or premises of

interactionist thought in general. One, that human beings are actors as well as reactors, has already been suggested. A second is that human action and interaction are critically shaped by definitions or interpretations of the situations of action and interaction, which definitions and interpretations are based on shared meanings developed in the course of interaction with others. A third premise is that the meanings which persons attribute to themselves, their self-conceptions, are especially critical to the process producing their action and interaction. And a fourth premise is that self-conceptions, like other meanings, are shaped in the course of interaction with others and are, at least in the initial instance and at least largely, the outcomes of others' responses to the person.

The fourth premise has sometimes been phrased as "self reflects society." Taken in conjunction with the third premise, it gives rise to the basic theoretical proposition or formula of symbolic interactionism: Society shapes self, which shapes social behavior. That formula, it must be noted, admits of and, indeed, insists upon the possibility of reciprocity among its components—social behavior impacts self and society, and self can impact society. Identity theory builds upon refinements of the traditional symbolic interactionist framework and specifications of its basic formula.

The refinements essentially have to do with three facets of the traditional symbolic interactionist framework as it evolved from Mead ([1934] 1962), Cooley ([1902] 1983), Blumer (1969), and others: the conceptualization of society, the conceptualization of self, and the relative weight to be accorded social structure versus interpretive processes in accounts of human behavior. The traditional framework tends to view "society" as unitary, as a relatively undifferentiated and unorganized phenomenon with few, if any, internal barriers to the evolution of universally shared meanings. It also tends to a view "society" as an unstable and ephemeral reflection, even reification, of relatively transient, ever-shifting patterns of interaction. In this view of society, social structures, as these are typically conceived of by sociologists, have little place in accounts of persons' behaviors. These accounts tend to be innocent of a coherent sense of extant social constraints on those behaviors, and there are few means of linking the dynamics of social

interaction in reasonably precise ways to the broader social settings that serve as context for persons' action and interaction.

Further, and enlarging this theme of an inadequate conceptualization and consequent neglect of social structure, this view of society tends to dissolve social structure in the universal solvent of subjective definitions and interpretations, thus missing the obdurate reality of social forms whose impact on behavior is undeniable. To say this does not deny the import for social life of the definitional and interpretative processes central to interactionist thinking and explanation. It is, however, to say that seeing these processes as in large degree unanchored and without bounds, as open to any possibility whatsoever without recognizing that some are much more probable than others, results in visualizing social life as less a product of external constraints and more a product of persons' phenomenology than is likely warranted. Finally, on the premise that self reflects society, this view of society leads directly to a view of self as unitary, as equivalently internally undifferentiated, unorganized, unstable, and ephemeral.

Contemporary sociology's image of society is considerably different from that contained in traditional symbolic interactionism, and it is the contemporary sociological conceptualization of society that is incorporated into the structural symbolic interactionist frame from which identity theory derives. This contemporary conceptualization emphasizes the durability of the patterned interactions and relationships that are at the heart of sociology's sense of social structure. It emphasizes social structure's resistance to change and its tendency to reproduce itself. The contemporary image differs as well by visualizing societies as highly differentiated yet organized systems of interactions and relationships; as complex mosaics of groups, communities, organizations, institutions; and as encompassing a wide variety of crosscutting lines of social demarcation based upon social class, age, gender, ethnicity, religion, and more. This vast diversity of parts is seen as organized in multiple and overlapping ways—interactionally, functionally, and hierarchically. At the same time, the diverse parts of society are taken to be sometimes highly interdependent and sometimes relatively independent of one another, sometimes implicated in close and cooperative interaction and sometimes conflicting.



The symbolic interactionist premise that self reflects society now requires a very different conceptualization of self, one that mirrors the altered conception of society. Self must be seen as multifaceted, as comprised of a variety of parts that are sometimes interdependent and sometimes independent of other parts, sometimes mutually reinforcing and sometimes conflicting, and that are organized in multiple ways. It requires a sense of self in keeping with James's ([1890] 1990) view that persons have as many selves as there are other persons who react to them, or at least as many as there are groups of others who do so.

Equally important, viewing both society and self as complex and multifaceted as well as organized opens the way to escaping the overly general, almost banal, and essentially untestable qualities of the basic symbolic interactionist formula by permitting theorization of the relations between particular parts of society and particular parts of self, and by permitting reasonable operationalizations of those parts.

In identity theory, this theorization proceeds by specifying the terms of the basic symbolic interactionist formula, doing so by focusing on particulars hypothesized as especially likely to be important in impacting role choice. That is, first of all, the general category of social behavior is specified by taking role choice—opting to pursue action meeting the expectations contained in one role rather than another—as the object of explanation. Role choice is hypothesized to be a consequence of identity salience, a specification of the general category of self, and identity salience is hypothesized to be a consequence of commitment, a specification of society. Identity theory's fundamental proposition, then, is: Commitment impacts identity salience impacts role choice.

The concept of identity salience develops from the multifaceted view of self articulated above. Self is conceptualized as comprised of a set of discrete identities, or internalized role designations, with persons potentially having as many identities as there are organized systems of role relationships in which they participate. Identities require both that persons be placed as social objects by having others assign a positional designation to them and that the persons accept that designation (Stone 1962; Stryker 1968). By this usage, identities are

self-cognitions tied to roles and, through roles, to positions in organized social relationships; one may speak of the identities of mother, husband, child, doctor, salesman, employee, senator, candidate, priest, tennis player, churchgoer, and so on. By this usage, too, identities are cognitive schemas (Markus 1977), structures of cognitive associations, with the capacity of such schemas to impact ongoing cognitive and perceptual processes (Stryker and Serpe 1994).

Self is not only multifaceted; it is also postulated to be organized. Identity theory takes hierarchy as a principal mode of organization of identities; in particular, it assumes that identities, given their properties as cognitive schemas, will vary in their salience, and that self is a structure of identities organized in a salience hierarchy. Identity salience is defined as the probability that a given identity will be invoked, or called into play, in a variety of situations; alternatively, it can be defined as the differential probability, across persons, that a given identity will be invoked in a given situation. Identity theory's fundamental proposition hypothesizes that choice between or among behaviors expressive of particular roles will reflect the relative location in the identity salience hierarchy of the identities associated with those roles.

The concept of commitment has its basic referent in the networks of social relationships in which persons participate; as such, commitment is a social structural term. Associated with the "complex mosaic of differentiated parts" image of society is the recognition that persons conduct their lives not in the context of society as a whole but, rather, in the many contexts of relatively small and specialized social networks, networks made up of persons to whom they relate by virtue of occupancy of particular social positions and the playing of the associated roles. To say that persons are committed to some social network is to say that their relationships to the other members of that network depend on their playing particular roles and having particular identities: To the degree that one's relationships to specific others depend on being a particular kind of person, one is committed to being that kind of person. Thus, commitment is measured by the costs of giving up meaningful relations to others should an alternative course of action be pursued. Commitment, so

defined and measured, is hypothesized by identity theory to be the source of the salience attached to given identities (Stryker 1968, 1980, 1987a).

Two analytically distinct and possibly independent dimensions or forms of commitment have been discerned: interactional and affective (Serpe 1987; Stryker 1968). The former has its referent in the number of relationships entered by virtue of having a given identity and by the ties across various networks of relationships (for instance, one may relate as husband not only to one's spouse, her friends, and her relatives but also to members of a couples' bridge club and other such groups). The latter has its referent in the depth of emotional attachment to particular sets of others.

Reciprocity among the three terms of the identity theory formula is again recognized; but the dominant thrust of the process is hypothesized to be as stated by the proposition, on the grounds that identity, as a strictly cognitive phenomenon, can change more readily than can commitment, whose conceptual core is interaction rather than cognition.

The empirical evidence brought to bear on the hypotheses contained in the fundamental identity theory formula has been supportive. Stryker and Serpe (1982) demonstrate that both time spent acting out a religious role and preferred distribution of time to that role are tied to the salience of the identity associated with the role; they demonstrate as well that the salience of the religious identity is tied to commitment (in this case, the measure of commitment combines interactional and affective commitment) to others known through religious activities. Burke and various associates (Burke and Hoelter 1988; Burke and Reitzes 1981; Burke and Tully 1977) show the link between identity and gender, academic attainment and aspirations, and occupational aspirations, finding evidence that the linkage reflects the commonality of meaning of identity and behavior. Lee (1998) finds that the correspondence of meanings of students' personal identities and meanings they attach to those occupying positions in scientific disciplines predicts interest in science as well as appreciably accounting for gender differences in intention to become scientists. Serpe and Stryker (1987), using data on student-related identities obtained at three points in time from students

entering a residential college, provide evidence that the salience of these identities is reasonably stable over time; that in a situation in which earlier commitments have been attenuated by a move to a residential university, high identity salience leads to efforts to reconstruct social relationships that permit playing the role associated with the salient identity, efforts taking the form of joining appropriate organizations; and that when such efforts are not successful, the level of salience of the identity subsequently drops and self-structure is altered. Callero (1985), Callero and associates (1987), and Charng and associates (1988) show that commitment and identity salience add appreciably to the ability to account for the behavior of repeated blood donors. Sparks and Shepherd (1992) find, to their considerable surprise, that identity theory-based predictions stand up well in accounting for behavioral intentions with regard to green consumerism, the predicted relationships holding when examined in the context of the variables of a theory of planned behavior. Serpe (1987) shows that over time there is indeed a reciprocal relationship between commitment to various student role relationships and the salience of identities associated with those roles, and that the identity theory hypothesis arguing the greater impact over time of commitment on salience than vice versa is correct.

The success of identity theory attested to in this brief and incomplete review of empirical evidence notwithstanding, however, there is reason to believe that the theory requires development and extension beyond the basic proposition that has been the major focus of attention to this point. Indeed, such work has begun; and it has a variety of thrusts. How to incorporate varying degrees of situational constraint into the theory—the impact that variations in “choice” have on the ways in which the relationships among commitment, identity salience, and role performances play themselves out—is one such thrust (Serpe 1987). Another seeks to explore mechanisms underlying the linkages among commitment and identity salience, and identity salience and behavior; to this end, work (especially by Burke and Reitzes 1981, 1991) exploits the basic symbolic interactionist idea that it is commonality of meaning which makes social life possible. Some attention has been given to extending the applicability and predictive power

of identity theory by incorporating into it other than strictly role-based identities; in particular, the concern has been with what have been termed “master statuses” (such as age, gender, and class) and personal traits (such as aggressiveness and honesty), and the suggestion is that master statuses and traits may affect identity processes by modifying the meaning of the roles from which identities derive (Stryker 1987a). Stryker and Serpe (1994) treat the conceptual and measurement confusion of the importance ranking of identities and identity salience, in the process showing that both importance and salience respond to commitment, that importance and salience are related, and that both contribute to the prediction of role-related behavior.

An effort is being made to correct the almost totally cognitive focus of identity theory (as well as its parent and grandparent interactionist frameworks) by recognizing the importance of affect and emotion to the processes with which the theory is concerned. The earliest statement of the theory (Stryker 1968) posited a cathectic modality of self that parallels the cognitive modality from which the emphasis on identity flows; however, subsequent work on the theory has not pursued that idea. Stryker (1987b) has attempted to integrate emotion into the theory by arguing, with Hochschild (1979), that emotional expressions carry important messages from self and, beyond Hochschild, that the experiences of emotions are messages to self informing those who experience those emotions of the strength of commitments and the salience of identities.

Finally—and here work has barely begun—it is time to make good on the promise to provide more adequate conceptualization of the linkages between identity theory processes and the wider social structures within which these processes are embedded. From the point of view of structural symbolic interactionism, structures of class, ethnicity, age, gender, and so on operate as social boundaries making it more or less probable that particular persons will form interactional networks; in this way, such social structures enter identity theory directly through their impact on commitments. However, the relation of such structures to identity processes clearly goes beyond this direct impact; they affect not only the probabilities of

interaction but also the content (meanings) of the roles entailed in interaction and, thus, the meanings of identities, the symbolic and material resources available to those who enter interaction with others, and the objectives or ends to which interactions are oriented. Explication of these impacts, both direct and indirect, of social structure on the processes that relate commitment, identity salience, and role performance remains to be accomplished.

## REFERENCES

- Blumer, Herbert 1969 *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, N.J.: Prentice-Hall.
- Burke, Peter J., and Viktor Gecas 1995 “Self and Identity.” In Karen Cook, Gary A. Fine, and James House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Burke, Peter J., and Jon W. Hoelter 1988 “Identity and Sex–Race Differences in Educational and Occupational Aspirations Formation.” *Social Science Research* 17:29–47.
- Burke, Peter J., and Donald C. Reitzes 1981 “The Link Between Identity and Role Performance.” *Social Psychology Quarterly* 44:83–92.
- 1991 “An Identity Theory Approach to Commitment.” *Social Psychology Quarterly* 54:239–251.
- Burke, Peter J., and Judy Tully 1977 “The Measurement of Role/Identity.” *Social Forces* 55:880–897.
- Callero, Peter L. 1985 “Role-Identity Salience.” *Social Psychology Quarterly* 48:203–215.
- , Judith A. Howard, and Jane A. Piliavin 1987 “Helping Behavior as Role Behavior: Disclosing Social Structure and History in the Analysis of Pro-Social Action.” *Social Psychology Quarterly* 50:247–256.
- Charng, Hong-Wen, June Allyn Piliavin, and Peter L. Callero 1988 “Role-Identity and Reasoned Action in the Prediction of Repeated Behavior.” *Social Psychology Quarterly* 51:303–317.
- Cooley, Charles H. (1902) 1983 *Human Nature and Social Order*. New Brunswick, N.J.: Transaction Books.
- Hewitt, John P. 1997a *Dilemmas of the American Self*. Philadelphia: Temple University Press.
- 1997b *Self and Society*, 7th ed. Needham Heights, Mass.: Allyn and Bacon.
- Hochschild, Arlie R. 1979 “Emotion Work, Feeling Rules, and Social Structure.” *American Journal of Sociology* 85:551–575.

- James, William (1890) 1990 *Principles of Psychology*. Chicago: Encyclopedia Britannica, Inc.
- Lee, James D. 1998 "What Kids Can 'Become' Scientists? The Effects of Gender, Self-Concepts, and Perceptions of Scientists." *Social Psychology Quarterly* 61:199-219.
- MacKinnon, Neil J. 1994 *Symbolic Interaction as Affect Control*. Albany: State University of New York Press.
- Markus, Hazel 1977 "Self-Schemas and Processing Information About the Self." *Journal of Personality and Social Psychology* 35:63-78.
- McCall, George, and J. S. Simmons 1978 *Identities and Interaction*, rev. ed. New York: Free Press.
- Mead, George H. (1934) 1963 *Mind, Self and Society*. Chicago: University of Chicago Press.
- Serpe, Richard T. 1987 "Stability and Change in Self: A Structural Symbolic Interactionist Explanation." *Social Psychology Quarterly* 50:44-55.
- , and Sheldon Stryker 1987 "The Construction of Self and the Reconstruction of Social Relationships." In Edward J. Lawler and Barry Markovsky, eds., *Advances in Group Processes*, vol. 4. Greenwich, Conn.: JAI Press.
- Sparks, Paul, and Richard Shepherd 1992 "Self-Identity and the Theory of Planned Behavior: Assessing the Role of Identification with 'Green Consumerism.'" *Social Psychology Quarterly* 55:388-399.
- Stone, Gregory P. 1962 "Appearance and the Self." In Arnold M. Rose, ed., *Human Behavior and the Social Process*. Boston: Houghton Mifflin.
- Stryker, Sheldon 1968 "Identity Salience and Role Performance." *Journal of Marriage and the Family* 30:558-564.
- 1980 *Symbolic Interactionism: A Social Structural Version*. Menlo Park, Calif.: Benjamin/Cummings.
- 1987a "Identity Theory: Developments and Extensions." In Krysia Yardley and Terry Honess, eds., *Self and Society: Psychosocial Perspectives*. New York: Wiley.
- 1987b "The Interplay of Affect and Identity: Exploring the Relationships of Social Structure, Social Interaction, Self and Emotion." Paper presented at annual meeting of the American Sociological Association, Chicago.
- , and Richard T. Serpe 1982 "Commitment, Identity Salience, and Role Behavior: Theory and Research Example." In William Ickes and Eric S. Knowles, eds., *Personality, Roles, and Social Behavior*. New York: Springer-Verlag.
- , and Richard T. Serpe 1994 "Identity Salience and Psychological Centrality: Equivalent, Overlapping, or Complementary Concepts?" *Social Psychology Quarterly* 57:16-35.

—, and Ann Statham 1985 "Symbolic Interactionism and Role Theory." In Gardner Lindzey and Eliot Aronson, eds., *The Handbook of Social Psychology*, 3rd ed. New York: Random House.

Weigert, Andrew J. 1983 "Identity: Its Emergence Within Sociological Psychology." *Symbolic Interaction* 6:183-206.

SHELDON STRYKER

## ILLEGAL ALIENS/ UNDOCUMENTED IMMIGRANTS

See International Migration.

## ILLEGITIMACY

Until the 1960s, it was widely assumed that marriage was a universal or nearly universal institution for licensing parenthood. Marriage assigned paternity rights to fathers (and their families) and guaranteed social recognition and economic support to mothers and their offspring. According to Malinowski (1930), who first articulated "the principle of legitimacy," and to Davis (1939, 1949), who extended Malinowski's theory into sociology, marriage provides the added benefit to children of connecting them to a wider network of adults who have a stake in their long-term development.

This functional explanation for the universality of marriage as a mechanism for legitimating parenthood became a source of intense debate in anthropology and sociology in the 1960s. Evidence accumulated from cross-cultural investigations showed considerable variation in marriage forms and differing levels of commitment to the norm of legitimacy (Bell and Vogel 1968; Blake 1961; Coser 1964; Goode 1961). More recently, historical evidence indicates that the institution of marriage was not firmly in place in parts of Western Europe until the end of the Middle Ages (Glendon 1989; Goody 1983).

The accumulation of contradictory data led Goode (1960, 1971) to modify Malinowski's theory to take account of high rates of informal unions

and nonmarital childbearing in many New World nations and among dispossessed cultural minorities. Goode (1971) argued that the norm of legitimacy was likely to be enforced only when fathers possessed wealth and property or when their potential economic investment in child rearing was high. Therefore, he predicted that when "giving a name" to children offers few material, social, or cultural benefits, the norms upholding marriage will become attenuated.

So vast have been the changes in the perceived benefits of marriage since the 1960s in the United States and most Western nations that even Goode's modification of Malinowski's theory of legitimacy now seems to be in doubt (Cherlin 1992; Davis 1985). Indeed, the term "illegitimacy" has fallen into disfavor precisely because it implies inferior status to children born out of wedlock. Both legal and feminist scholars have been critical of the notion that the presence of a father confers status on the child (Burns and Scott 1994; Mason et al. 1998). The nuclear unit (biological parents and their offspring)—once regarded as the cornerstone of our kinship system—remains the model family form, but it no longer represents the exclusive cultural ideal, as was the case in the mid-1960s. The incentives for marriage in the event of premarital pregnancy have declined, and the sanctions against remaining single have diminished (Cherlin 1999; McLanahan and Casper 1995). In the 1990s, considerable scholarly attention and public policy debate has been devoted to ways of restoring and reinvesting in the institution of marriage (Furstenberg 1996; McLanahan and Sandefur 1994; Popenoe et al. 1996).

### TRENDS IN NONMARITAL CHILDBEARING

Premarital pregnancy has never been rare in the United States or in most Western European nations (Burns and Scott 1994; Goode 1961; Smith 1978; Vinovskis 1988). Apparently, the tolerance for pregnancy before marriage has varied over time and varies geographically at any given time. Throughout the first half of the twentieth century, premarital pregnancy almost always led to hasty marriages rather than out-of-wedlock births—even for very young women (O'Connell and Moore 1981; Vincent 1961). In 1940, illegitimacy was

uncommon in the United States, at least among whites. Nonmarital births were estimated at about 3.6 per 1,000 unmarried white women, while the comparable rate for nonwhites was 35.6. For all age groups, among whites and nonwhites alike, a spectacular rise occurred over the next five decades (Clague and Ventura 1968; Cutright 1972; McLanahan and Casper 1995).

In the 1960s and the 1970s, nonmarital childbearing rates continued to increase for younger women, albeit at a slower pace, while for women in their late twenties and thirties rates temporarily declined. Then, in the late 1970s, nonmarital childbearing rates rose again for all age groups and among both whites and African Americans. This rise continued until the mid-1990s, when levels of nonmarital childbearing stabilized or even declined (Ventura et al. 1996). Since the early 1970s, rates of marriage and marital childbearing have fallen precipitously. Thus, the ratio of total births to single women has climbed continuously (Smith et al. 1996). Nearly a third of all births (32.4 percent) in 1996 occurred out of wedlock, more than seven times the proportion in 1955 (4.5 percent) and more than twice that in 1975 (14.3 percent). The declining connection between marriage and parenthood is evident among all age groups but is especially pronounced among women in their teens and early twenties. Three out of four births to teens and nearly half of all births to women ages twenty to twenty-four occurred out of wedlock. Virtually all younger blacks who had children in 1995 (more than 95 percent) were unmarried, while two-thirds of white teens and more than a third of white women twenty to twenty-four were single when they gave birth.

Nonmarital childbearing was initially defined as a problem among teenagers and black women (Furstenberg 1991). But these recent trends strongly suggest that disintegration of the norm of legitimacy has spread to all segments of the population. First the link between marriage and sexual initiation dissolved, and now the link between marriage and parenthood has become weak. Whether this trend is temporary or a more permanent feature of the Western family system is not known. But public opinion data suggest that a high proportion of the population finds single parenthood acceptable. A Roper study ("Virginia Slims American

Women's Opinion Poll" 1985) revealed that 49 percent of women agreed that "There is no reason why single women should not have children and raise them if they want to."

Citing similar attitudinal evidence from the National Survey of Families and Households in 1987–1988, Bumpass (1990) concludes that there has been an "erosion of norms" proscribing nonmarital childbearing. He concludes that this behavior is not so much motivated by the desire to have children out of wedlock as it is by the reduced commitment to marriage and the limited sanctions forbidding nonmarital childbearing. Bumpass argues that much of the nonmarital childbearing is unplanned and ill timed.

### THE CONSEQUENCES OF NONMARITAL CHILDBEARING

Although extensive research exists on the economic, social, and psychological sequelae of single parenthood for adults and children, relatively little of this research has distinguished between the consequences of marital disruption and nonmarriage (Furstenberg 1989; Furstenberg and Cherlin 1991; Garfinkel and McLanahan 1986; Maynard 1997). A substantial literature exists on the consequences of nonmarital childbearing, but it is almost entirely restricted to teenage childbearers (Chilman 1983; Hofferth and Hayes 1987; Institute of Medicine 1995; Miller and Moore 1990; Moore et al. 1986). It is difficult, then, to sort out the separate effects of premature parenthood, marital disruption, and out-of-wedlock childbearing on parents and their offspring.

Nonmarital childbearing most certainly places mothers and their children at risk of long-term economic disadvantage (Institute of Medicine 1995; Maynard 1997; McLanahan and Booth 1989). Out-of-wedlock childbearing increases the odds of going on welfare and of long-term welfare dependency (Duncan and Hoffman 1990). The link between nonmarital childbearing and poverty can probably be traced to two separate sources. The first is "selective recruitment," that is, women who bear children out of wedlock have poor economic prospects before they become pregnant, and their willingness to bear a child out of wedlock may reflect the bleak future prospects of many unmarried pregnant women, especially younger women

(Furstenberg 1990; Geronimus 1987; Hayes 1987; Hogan and Kitagawa 1985; Maynard 1997). But is also likely that out-of-wedlock childbearing—particularly when it occurs early in life—directly contributes to economic vulnerability because it reduces educational attainment and may limit a young woman's prospects of entering a stable union (Furstenberg 1991; Hofferth and Hayes 1987; Hoffman et al. 1993; Trussell 1988).

If nonmarital childbearing increases the risk of lengthy periods of poverty for women and their children, it is also likely that it restricts the opportunities for intra- and intergenerational mobility of families formed as single-parent units. Growing up in poverty restricts access to health, high-quality schools, and community resources that may promote success in later life (Ellwood 1988; Wilson 1987). Apart from the risks associated with poverty, some studies have shown that growing up in a single-parent family may put children at greater risk because they receive less parental supervision and support (Amato and Booth 1997; Dornbush 1989; McLanahan and Booth 1989; McLanahan and Sandefur 1994). As yet, however, researchers have not carefully distinguished between the separate sources of disadvantage that may be tied to nonmarital childbearing: economic disadvantage (which could restrict social opportunities or increase social isolation) and psychological disadvantage (which could foster poor parenting practices or limit family support).

Even though nonmarital childbearing may put children at risk of long-term disadvantage, it is also possible that over time the advantages conferred by marriage may be decreasing in those segments of the population that experience high rates of marital disruption (Bumpass 1990; Edin 1998; Furstenberg 1995). Moreover, the social and legal stigmata once associated with nonmarital childbearing have all but disappeared in the United States and many other Western nations (Glendon 1989). Over time, then, the hazards associated with nonmarital childbearing (compared with ill-timed marital childbearing) for women and their children may have declined. Whatever the reasons for these trends, it appears that nonmarital childbearing may have peaked by the mid-1990s. Whether the leveling off in rates of nonmarital childbearing signals a shift in family formation patterns or is

merely a response to the robust economy of the 1990s remains to be seen.

### NONMARITAL CHILDBEARING AND PUBLIC POLICY

Growing rates of nonmarital childbearing in the United States and many other Western nations suggest the possibility that the pattern of childbearing before marriage or between marriages may be spreading upward into the middle class. In Scandinavia, where marriage has declined most dramatically, it is difficult to discern whether formal matrimony is being replaced by a *de facto* system of informal marriage (Hoem and Hoem 1988). If this were to happen, the impact on the kinship system or the circumstances of children might not be as dramatic as some have speculated. But if the institution of marriage is in serious decline, then we may be in the midst of a major transformation in the Western family.

The weakening of marriage has created confusion and dispute over parenting rights and responsibilities. A growing body of evidence indicates that most nonresidential biological fathers, especially those who never marry, typically become disengaged from their children (Arendell 1995; King and Heard 1999; Seltzer 1991; Teachman 1990). Most are unwilling or unable to pay regular child support, and relatively few have constant relationships with their children. Instead, the costs of child rearing have been largely assumed by mothers and their families, aided by public assistance. A minority of fathers do manage to fulfill economic and social obligations, and some argue that many others would do so if they had the means and social support for continuing a relationship with their children (Marsiglio 1998; Smollar and Ooms 1987).

The uncertain relationship between biological fathers and their children has created a demand for public policies to shore up the family system (Garfinkel et al. 1996; Popenoe 1996). Widespread disagreement exists over specific policies for addressing current problems. Advocates who accept the current reality of high levels of nonmarriage and marital instability propose more generous economic allowances and extensive social support to women and their children to offset the limited

economic role of men in disadvantaged families (Ellwood 1988). Critics of this approach contend that such policies may further erode the marriage system (Vinovskis and Chase-Lansdale 1987). Yet few realistic measures have been advanced for strengthening the institution of marriage (Furstenberg and Cherline 1991).

Enforcement of child support has attracted broad public support. A series of legislative initiatives culminating in the Family Support Act of 1988 have increased the role of federal and state governments in collecting child support from absent parents (typically fathers) and standardizing levels of child support. There has been a steady but modest improvement in the collection of child support in the 1990s. It is much less clear whether the strengthening of child support has worked to the benefit of children (Furstenberg et al. 1992; Garfinkel et al. 1996). It is too early to tell whether these sweeping measures will succeed in strengthening the economic contributions of fathers who live apart from their children. And, if it does, will greater economic support by absent parents reinforce social and psychological bonds to their children (Furstenberg 1989; Garfinkel and McLanahan 1990)? The other great experiment of the 1990s was the Welfare Reform Act in 1996, which replaced the longstanding entitlements to public assistance with temporary provisional support. It is much too early to tell what, if any, the effects of this policy will be on marriage and fertility practices. Advocates of welfare reform claimed that it would reduce out-of-wedlock childbearing and help restore marriage (Murray 1984). But the link between welfare payments and marriage patterns has never been strong (Moffitt 1998). Still, it may be possible to devise a test of the consequences of the different policies given the large state variations in program implementation.

As for the future of marriage, few, if any, sociologists and demographers are predicting a return to the status quo or a restoration of the norm of legitimacy. Short of a strong ideological swing favoring marriage and condemning nonmarital sexual activity and childbearing, it is difficult to foresee a sharp reversal in present trends (Blankenhorn et al. 1990). Predicting the future, however, has never been a strong point of demographic and sociological research.

(SEE ALSO: *Deviance Theories; Law and Society*)

REFERENCES

- Amato, P.R., and A. Booth 1997 *A Generation at Risk: Growing Up in an Era of Family Upheaval*. Cambridge, Mass.: Harvard University Press.
- Arendell, T. 1995 *Fathers and Divorce*. Thousand Oaks, Calif.: Sage Publications.
- Bell, N. W., and E. F. Vogel (eds.) 1968 *A Modern Introduction to the Family*. New York: Free Press.
- Blake, J. 1961 *Family Structure in Jamaica*. New York: Free Press.
- Blakenhorn, D., S. Bayne, and J. B. Elshtain 1990 *Rebuilding the Nest*. Milwaukee, Wis.: Family Service America.
- Bumpass, L. 1990 "What's Happening to the Family? Interactions Between Demographic and Institutional Change." *Demography* 27:483-498.
- Burns, A., and C. Scott 1994 *Mother-Headed Families and Why They Have Increased*. Hillsdale, N.J.: Lawrence Erlbaum.
- Cherlin, A. J. 1992 *Marriage, Divorce, Remarriage*, rev. and enl. ed. Cambridge, Mass.: Harvard University Press.
- 1999 *Public and Private Families*, 2nd ed. Boston: McGraw-Hill.
- Chilman, C. S. 1983 *Adolescent Sexuality in a Changing American Society: Social and Psychological Perspectives for the Human Services Professions*. New York: Wiley.
- Clague, A. J., and S. J. Ventura 1968 *Trends in Illegitimacy: United States 1940-1965*. (Vital and Health Statistics, Public Health Service Publication no. 1000, ser. 21, no. 15.) Washington, D.C.: U.S. Government Printing Office.
- Coser, R. L., (ed.) 1964 *The Family: Its Structures and Functions*. New York: St. Martin's Press.
- Cutright, P. 1972 "Illegitimacy in the United States, 1920-1968." In C. Westoff and R. Parks, eds., *Demographic and Social Aspects of Population Growth*. Washington D.C.: U.S. Government Printing Office.
- Davis, K. 1939 "Illegitimacy and the Social Structure." *American Journal of Sociology* 45:215-233.
- 1949 *Human Society*. New York: Macmillan.
- 1985 *Contemporary Marriage: Comparative Perspectives on a Changing Institution*. New York: Russell Sage Foundation.
- Dornbush, S. M. 1989 "The Sociology of Adolescence." *The Annual Review of Sociology* 15: 233-259.
- Duncan, G. J., and S. D. Hoffman 1990 "Teenage Welfare Receipt and Subsequent Dependence Among Black Adolescent Mothers." *Family Planning Perspectives*. 22:219-223.
- Edin, K. 1998 "Why Don't Poor Single Mothers Get Married (or Remarried)?" Paper presented at the Russell Sage Foundation, New York, May.
- Ellwood, D. T. 1988 *Poor Support*. New York: Basic Books.
- Furstenberg, F. F., Jr. 1989 "Supporting Fathers: Implications of the Family Support Act for Men." Paper presented at the forum on the Family Support Act sponsored by the Foundation for Child Development, Washington, D.C., November.
- 1990 "Coming of Age in a Changing Family System." In F. Feldman and G. Elliott, eds., *At the Threshold: The Developing Adolescent*. Cambridge, Mass.: Harvard University Press.
- 1991 "As the Pendulum Swings: Teenage Childbearing and Social Concern." *Family Relations* 40:127-138.
- 1995 "Fathering in the Inner-City: Paternal Participation and Public Policy." In W. Marsiglio, ed., *Fatherhood: Contemporary Theory, Research, and Social Policy*. Thousand Oaks, Calif.: Sage Publications.
- 1996 "The Future of Marriage." *American Demographics* (June): 34-40.
- , and A. Cherlin 1991 *Divided Families: What Happens to Children When Their Parents Part*. Cambridge, Mass.: Harvard University Press.
- K. E. Sherwood, and M. L. Sullivan 1992 *Caring and Paying: What Fathers and Mothers Say About Child Support*. New York: Manpower Demonstration Research Corporation.
- Garfinkel, I., J. L. Hochschild, and S. McLanahan (eds.) 1996 *Social Policies for Children*. Washington, D.C.: Brookings Institution.
- Garfinkel, I., and S. McLanahan 1986 *Single Mothers and Their Children*. Washington, D.C.: Urban Institute Press.
- 1990 "The Effects of the Child Support Provisions of the Family Support Act of 1988 on Child Well-Being." *Population Research and Policy Review* 9:205-234.
- Geronimus, A. T. 1987 "On Teenage Childbearing in the United States." *Population and Developmental Review* 13:245-279.
- Glendon, M. A. 1989 *The Transformation of Family Law*. Chicago: University of Chicago Press.
- Goode, W. J. 1960 "A Deviant Case: Illegitimacy in the Caribbean." *American Sociological Review* 25:21-30.
- 1961 "Illegitimacy, Anomie, and Cultural Penetration." *American Sociological Review* 26:319-325.



- 1971 "Family Disorganization." In R. K. Merton and R. Nisbet, eds., *Contemporary Social Problems*, 3rd ed. New York: Harcourt Brace Jovanovitch.
- Goody, J. 1983 *The Development of the Family and Marriage in Europe*. Cambridge, U.K.: Cambridge University Press.
- Hayes, C. D. 1987 *Risking the Future*, vol. 1. Washington, D.C.: National Academy Press.
- Hoem, B., and J. M. Hoem 1988 "The Swedish Family: Aspects of Contemporary Development." *Journal of Family Issues* 9:397-424.
- Hofferth, S. L., and C. D. Hayes 1987 *Risking the Future*, vol. 2. Washington, D.C.: National Academy Press.
- Hoffman, S. D., E. M. Foster, and F. F. Furstenberg, Jr. 1993 "Re-Evaluating the Costs of Teenage Childbearing." *Demography* 30:1-13.
- Hogan, D. P., and E. Kitagawa 1985 "The Impact of Social Status, Family Structure, and Neighborhood on the Family Structure of Black Adolescents." *American Journal of Sociology* 90:825-855.
- Institute of Medicine 1995 *The Best Intentions: Unintended Pregnancy and the Well-Being of Children and Families*. Washington, D.C.: National Academy Press.
- King, V., and Heard, H. E. 1999 "Nonresident Father Visitation, Parental Conflict, and Mother's Satisfaction: What's Best for Child Well-Being?" *Journal of Marriage and the Family* 61:385-396.
- Malinowski, B. 1930 "Parenthood, the Basis of Social Structure." In R. L. Coser, ed., *The Family: Its Structure and Functions*. New York: St. Martin's Press.
- Marsiglio, W. 1998 *Procreative Man*. New York: New York University Press.
- Mason, M. A., A. Skolnick, and S. D. Sugarman (eds.) 1998 *All Our Families: New Policies for a New Century*. New York: Oxford University Press.
- Maynard, R. A. (ed.) 1997 *Kids Having Kids: Economic Costs and Social Consequences of Teen Pregnancy*. Washington, D.C.: Urban Institute Press.
- McLanahan, S., and K. Booth 1989 "Mother-Only Families: Problems, Prospects, and Politics." *Journal of Marriage and the Family* 51:577-580.
- McLanahan, S., and L. Casper 1995 "Growing Diversity and Inequality in the American Family." In R. Farley, ed., *State of the Union: America in the 1990s*, vol. 2. New York: Russell Sage Foundation.
- McLanahan, S., and G. Sandefur 1994 *Growing up with a Single Parent: What Hurts, What Helps*. Cambridge, Mass.: Harvard University Press.
- Miller, B. C., and K. A. Moore 1990 "Adolescent Sexual Behavior, Pregnancy, and Parenting." *Journal of Marriage and the Family* 524:1025-1044.
- Moffitt, R. A. (ed.) 1998 *Welfare, the Family and Reproductive Behavior*. Washington, D.C.: National Academy Press.
- Moore, K. A., M. C. Simms, and C. L. Betsey 1986 *Choice and Circumstance: Racial Differences in Adolescent Sexuality and Fertility*. New Brunswick, N.J.: Transaction Books.
- Murray, C. 1984 *Losing Ground*. New York: Basic Books.
- O'Connell, M., and M. J. Moore 1981 "The Legitimacy Status of First Births to U.S. Women Aged 15-24, 1939-1978." In F. F. Furstenberg, Jr., R. Lincoln, and J. Menken, eds., *Teenage Sexuality, Pregnancy, and Childbearing*. Philadelphia: University of Pennsylvania Press.
- Popenoe, D. 1996 *Life Without Father*. New York: Free Press.
- , J. B. Elshtain, and D. Blankenhorn (eds.) 1996 *Promises to Keep: Decline and Renewal of Marriage in America*. Lanham, Md.: Rowman & Littlefield.
- Seltzer, J. A. 1991 "Relationships Between Fathers and Children Who Live Apart: The Father's Role after Separation." *Journal of Marriage and the Family* 53:79-101.
- Smith, H. L., S. P. Morgan, and T. Koropeckyz-Cox 1996 "A Decomposition of Trends in the Nonmarital Fertility Ratios of Blacks and Whites in the United States, 1960-1992." *Demography* 33:141-151.
- Smollar, J., and T. Ooms 1987 *Young Unwed Fathers: Research Review, Policy Dilemmas, and Options*. Washington, D.C.: Maximus.
- Teachman, J. D. 1990 "Still Fathers? The Reorganization of Parental Obligations Following Divorce." Paper presented at the Albany Conference of Demographic Perspectives on the American Family: Patterns and Prospects, April.
- Trussel, J. 1988 "Teenage Pregnancy in the United States." *Family Planning Perspectives* 20:262-272.
- Ventura, S. J., J. A. Martin, S. C. Curtin, and T. J. Matthews 1996 "Report of Final Natality: Statistics." *Monthly Vital Statistics Report* 45 (11) (supp. 2) Hyattsville, Md.: National Center for Health Statistics.
- Vincent, C. E. 1961 *Unmarried Mothers*. New York: Free Press.
- Vinovskis, M. 1988 *An "Epidemic" of Adolescent Pregnancy? Some Historical and Policy Considerations*. New York: Oxford University Press.
- , and P. L. Chase-Lansdale 1987 "Should We Discourage Teenage Marriage?" *Public Interest* 98:23-37.
- "Virginia Slims American Women's Public Opinion Poll" 1985 A study conducted by the Roper Center for Public Opinion.

Wilson, W. J. 1987 *The Truly Disadvantaged*. Chicago: University of Chicago Press.

FRANK F. FURSTENBERG JR.

## IMPERIALISM, COLONIALISM, AND DECOLONIZATION

The colonial expansion of European states into the Americas, Asia, Africa, and the Pacific, followed by the collapse of these empires and their replacement by sovereign nation-states, is a double movement of great historical importance. The following briefly reviews the larger contours of this history and outlines some central arguments about its causes and consequences.

### IMPERIALISM AND COLONIALISM

The term “imperialism” was first used in the 1830s to recall Napoleonic ambitions. It gained its core contemporary meaning around the turn of the century as a description of the feverish colonial expansion of Britain, France, Germany, Russia, the United States, and Italy. But the term is not confined to formal colonial expansion; in particular, the continuing dependence of much of the Third World on Western states and multinational corporations is often understood as neocolonialism or neoimperialism (Magdoff 1969; Nkrumah 1966).

Contemporary efforts to distill these diverse usages generally define imperialism as the construction and maintenance of relationships of domination between political communities. Such relations are often seen as explicitly political, either in the narrow sense of direct administrative control or more broadly as formal or informal control over state policy. Economic conceptions of imperialism sometimes develop an analogue to these notions, where relations of economic control or exploitation replace political domination.

“Imperialism” has also been appropriated in a much narrower sense to support particular lines of argument. Lenin’s statement that “imperialism is the monopoly stage of capitalism” ([1917] 1939, p. 88) may be understood as a move of this sort. Arrighi, among others, argues that Lenin is better understood as formulating a substantive proposition; he suggests the interpretation “imperialism,

or the tendency to war between capitalist countries, is a necessary consequence of the transformation of capitalism into monopoly or finance capital” (1978, p. 14).

Even when imperialism is equated with the establishment and maintenance of political domination, an awkward relationship between *imperialism* and *empire* persists. Classically, “empire” refers to the great agrarian bureaucracies that dominated antiquity, from the Aztec to the Chinese, from ancient Sumer to Imperial Rome. It is not clear how much these structures have in common with the overseas colonial empires of Western states, much less with contemporary structures of dependence on foreign investment. Agrarian bureaucracies involved ethnic divisions that separated classes (most importantly separating warriors and peasant producers) rather than entire communities or nations (Gellner 1983).

A second historical use of “empire” is the medieval image of a temporal parallel to the Roman Church (Folz 1969; Guenee 1985). Rather than an alien and illegitimate structure, empire was seen as a political order unifying the Christian world. Revived by Charlemagne, the notion of a universal polity lived on, in an increasingly ghostly fashion, through the Holy Roman Empire. It receded into the background as a real political force with the construction of absolutist states and was lost as a compelling image with the rise of the nation-state.

In contrast to these historical understandings of empire, modern conceptions of imperialism rest on the notion that popular sovereignty forms the basis of political community. Only with the notion of popular sovereignty does domination refer to relationships between rather than within communities. If the criteria that the United Nations uses today to identify colonialism were to be applied before 1700, all territories would be parts of empires and all peoples would be dependent subjects. It is thus no accident that the notion of imperialism arose with the nation-state; it connotes the expansionary drive of a community that is internally organized around (the myth of) popular sovereignty.

**European Political Expansion.** European overseas expansion can be described crudely as occurring in two stages, the colonial and the imperial. In the fifteenth and sixteenth centuries, seagoing

powers constructed networks of colonial enclaves along the route to the East Indies. Less than half a century after the voyages of Columbus, the conquistadores had laid waste to the Incan and Aztec empires and were sending gold and silver back to Spain. In the following two centuries, Spain, Portugal, Great Britain, France, and the Netherlands colonized virtually the whole of the Caribbean, Central and South America, and the North Atlantic seaboard. The colonial period per se came to a close with the wars of national independence in the Americas between 1776 and 1830, leaving European states in control of little more than trading posts and exhausted sugar plantations.

The second period of expansion, one of imperial rather than colonial expansion, began after an interregnum marked by British naval hegemony. In the three decades after 1880, a scramble for territory partitioned Africa, Southeast Asia, and the Pacific among Great Britain, France, Germany, Belgium, and Portugal, while the United States annexed the remains of the Spanish Empire. None of this expansion involved much metropolitan emigration; colonial officials, traders, planters, and missionaries formed a thin veneer on indigenous societies.

The sources of the political structures of imperial rule lie in both metropolitan and indigenous traditions. Colonies tended to be formally organized along metropolitan lines (Fieldhouse 1966). Settler colonies mirrored domestic political structures quite directly (Lang 1975), while nonsettler colonies recall metropolitan structures in a more abstract fashion. For example, the British tried to fashion systems of local rule (Lugard [1922] 1965), while the French strove to create a unified, centralized administration. But the superficiality of most imperial rule led to great variation in actual administrative arrangements. Even empires whose guiding rationale was assimilation (the French and the Portuguese, especially) depended heavily on indigenous authorities and traditions.

Overseas colonies also varied in the strength and character of their economic relationship to the metropolis. Only a few colonies were the source of great riches for the metropolitan economy: most prominently, the American settler and plantation colonies, British India, and the Dutch East Indies. Others had a largely strategic value; much

of the British Empire, for instance, was acquired in the effort to maintain lines of communication to India. The great majority of colonies acquired after 1880 had rather little importance for the metropolis, either as markets for imperial products or as sources of raw materials (Fieldhouse 1973).

**Theories of Imperialism.** The starting point for modern theories of imperialism is John Hobson's *Imperialism: A Study* ([1902] 1965). A liberal critic of the Boer War, Hobson saw imperial expansion as a search for new outlets for investment. He found the source of this search in the surplus capital amassed by increasingly monopolistic corporate trusts. Hobson viewed imperial expansion as costly for the nation as a whole and sought to expose the special interests promoting imperialism. He also contended that capital surpluses could be consumed domestically by equalizing the distribution of income.

Lenin's *Imperialism: The Highest Stage of Capitalism* ([1917] 1939) provides the most influential statement of an economic analysis of imperialism. Lenin agreed with Hobson that imperialism flowed from the need to invest outside the domestic economy, drawing explicitly on Hilferding's ([1910] 1981) analysis of finance capital as a stage of capitalism. He was concerned to show that imperialism was a necessary consequence of the dynamics of capitalism (in contrast to Hobson's anticipation of Keynes) and that the expansionary impulse could not be globally coordinated (versus Kautsky's notion of an ultra-imperialism). Lenin argued that the unevenness of development makes imperialist war inevitable, as "late starters" demand their own place in the sun.

More contemporary writers like Baran and Sweezy (1966), Frank (1967), and Wallerstein (1974) drew upon both the Marxist tradition and Latin American theories of *dependencia* to suggest an alternative economic analysis of imperialism. They argued that international economic relations involve a net transfer of capital from the "periphery" to the "core" of the economic system and point to the continuities in this process from early colonial expansion to contemporary neoimperialism. This is in sharp contrast with the Leninist tradition, which argues that colonial forays bring noncapitalist societies into the world economy.

Other writers consider political ambitions or relationships to be the taproot of imperialism. In

perhaps the most interesting account of this sort, Schumpeter (1951) turned the Marxist perspective on its head. He noted that the characteristic motif of the ancient empires is military expansion for its own sake. Schumpeter argued that imperialism appears as an atavistic trait in the landed aristocracy of modern societies, stressing the mismatch between the social psychology of the warrior and the industrious, calculating spirit of the entrepreneur. A more political perspective treats imperial activity as flowing from the anarchical structure of the Western state system (Cohen 1973; Waltz 1979). In the absence of an enforceable legal order, states are motivated to expand when possible or endure decline relative to more aggressive states. This perspective explains European imperialism in the nineteenth century as the product of increasing levels of international competition and conflict.

Whether economic or political, most analyses of imperialism find its sources in the logic of the West, ignoring indigenous peoples in the process. John Gallagher and Ronald Robinson led a historiographic revision aiming to redress this imbalance. Their seminal essay "The Imperialism of Free Trade" (1953) emphasized the continuity in British policy between the informal imperialism of the mid-eighteenth century and the rush for colonies after 1880. In *Africa and the Victorians*, Robinson and colleagues (1961) argued that it was increasing indigenous resistance to European influence that led Britain to replace informal domination with formal empire. In later work, Robinson (1972) emphasized the other side of the coin—the extent to which Western empires as political systems were dependent on local collaboration.

### DECOLONIZATION

"Decolonization" refers to a polity's movement from a status of political dependence or subordination to a status of formal autonomy or sovereignty. In modern usage, it is generally assumed that the imperial or metropolitan center is physically separated from the dependency and that the two societies are ethnically distinct. The term refers specifically to the disintegration of Western overseas empires and their replacement by sovereign states in the Americas, Asia, and Africa.

There are several routes by which decolonization can take place. Most frequently, the dependency

becomes a new sovereign state, a political entity recognized in the international arena as independent of other states and as possessing final jurisdiction over a defined territory and population. Less often, decolonization may occur through the dependency's full incorporation into an existing polity, such that it is no longer separate and subordinate.

It is often unclear when (or whether) decolonization has occurred. Puerto Rico's relation to the United States can be described as one of colonial dependency or as free association. In the 1960s, Portugal claimed to have no colonies, only overseas territories formally incorporated into a unitary Portuguese state (Nogueira 1963). And where political relations are not contested, the absence of overt conflict makes it difficult to know when sovereignty has been achieved. For example, arguments can be made for dating Canadian independence at 1867, 1919, 1926, or 1931.

**European Political Contraction.** Virtually all of the decolonization of Western overseas empires occurred in two historical eras (Bergesen and Schoenberg 1980). The major American colonies became independent during the late eighteenth and early nineteenth centuries. The mid-twentieth century witnessed a more rapid and complete wave of decolonization worldwide. The types of colonies in existence in each period and the nature of the decolonization process varied greatly across the two periods (Fieldhouse 1966; Strang 1991a).

The first wave of decolonization began with the independence of Britain's thirteen continental colonies as the United States of America. The French Revolution touched off a slave uprising that led ultimately to the independence of the French colony of Saint Domingue as Haiti. Portuguese Brazil and Spanish Central and South America became independent after the Napoleonic Wars, which had cut Latin America off from the Iberian peninsula.

While the first period of decolonization was limited to the Americas, twentieth-century decolonization was global in scope. It included the independence of most of the Indian subcontinent, Southeast Asia and Australasia, the Middle East, Africa, and the Caribbean. Between the world wars, some of Britain's settler colonies and a number of loosely held protectorates became fully sovereign. Soon after World War II, the major

Asian colonies—India, Indonesia, Indochina, and the Philippines—achieved independence. The pace of change rapidly accelerated during the 1960s, which saw the decolonization of nearly all of Africa. By the 1980s, nearly all Western colonies had become independent or had been fully incorporated into sovereign states.

One fundamental difference between the two eras of decolonization has to do with who sought independence. Early American decolonizations were creole revolutions, as the descendants of European settlers sought political autonomy from the “mother country.” The American Revolution and the Spanish Wars for Independence were political rather than social revolutions. Slave revolt in Haiti provided the sole exception, to the horror of creole nationalists as well as loyalists elsewhere.

By contrast, twentieth-century decolonization was rooted in indigenous rather than creole movements for independence, as decolonization came to mean freedom from racially alien rule. After World War II, settler minorities opposed decolonization, since national independence spelled an end to their privileged political, economic, and social position. Only in South Africa did a racist minority regime survive decolonization.

The first and second waves of decolonization also differed importantly in the amount of violence involved. Early decolonization in the Americas was won through military combat between settler and imperial forces. Wars for independence raged in Britain’s thirteen continental colonies, in Spanish Central and South America, and in Haiti. Only in Portuguese Brazil was independence achieved without a fight, largely because Brazil was several times richer and more populous than Portugal.

During the twentieth century, protracted wars for independence were fought in Indochina, Indonesia, Algeria, and Angola. But these were the exceptions to the rule. Most colonies became independent with little or no organized violence between the imperial state and colonial nationalists. In much of Africa, imperial powers virtually abandoned colonies at the first sign of popular opposition to the colonial regime. By the mid-1960s, decolonization had become a rather routine activity for many imperial powers, often achieved through institutionalized expressions of popular will such as plebiscites.

**Causes of Decolonization.** A variety of arguments have been developed about factors contributing to decolonization. While most treatments have dealt with a single dependency or empire, there have been a number of efforts to develop explicitly comparative analyses (see Albertini 1982; Anderson 1983; Bergesen and Schoenberg 1980; Boswell 1989; Emerson 1960; Grimal 1978; Lang 1975; Smith 1978; Strang 1990).

Decolonization is often seen as the result of structural change in the dependency itself. Settler colonies are thought to undergo a natural process of maturation, well expressed in the physiocratic maxim that colonies are like fruit that fall from the tree when they are ripe. Indigenous populations are also importantly affected by contact with Western economic and political structures.

In both kinds of colonies, the specific condition that seems to precipitate decolonization is the emergence of peripheral nationalism. Settler colonies generally began as economic corporations chartered by European states. Non-Western peoples were generally tribal or segmental societies prior to colonization, and imperial structures were fundamentally dependent on the collaboration of indigenous elites (Robinson 1972). Decolonization required a new vision of the colonial dependency as a national society (Anderson 1983; Diamond 1958).

Colonial powers contributed unintentionally to the formulation of a national vision. They did so partly by spurring the rise of new social groups—indigenous bourgeois, landless workers, civil servants, teachers—who proved to be the carriers of colonial nationalism and independence. Contact with the colonial power exposed these groups to the notions and institutions of the Western nation-state while simultaneously denying them participation rights. Settlers, of course, carried notions of these rights with them (see Bailyn 1967 and Greene 1986 on the ideological origins of the American Revolution). Under these conditions, nationalism was a weapon easily turned on its creators (Emerson 1960; Strang 1992).

While pressures for decolonization invariably stemmed from the dependency itself, the response of the metropolitan power played a crucial role in the outcome. The classic contrast in imperial policy is between British “association” and French “assimilation” (though parallel contrasts between

the imperial policies of the United States and Portugal are even more striking). The British empire was administratively structured around indirect rule and local autonomy, which permitted considerable flexibility in the imperial reaction to pressures for decolonization. By contrast, the French aimed at the assimilation of their colonies into a unitary republic, which led to firmer resistance to decolonization.

Some decolonization occurred because the balance of military capacity shifted against metropolitan states. This was crucial to the decolonization of the Americas, as increasingly vigorous creole societies could not be controlled from across the Atlantic. After World War II, metropolitan weakness was a contributing factor as European colonial states found themselves reduced to the position of second-rate powers. At critical junctures after 1945, metropolitan governments were unable to project sufficient military power abroad to control events: the British in Suez, the French in Indochina and Algeria, the Dutch in Indonesia.

Finally, several systemic factors or processes may be linked to decolonization. One is the presence of a state that is economically and politically dominant on the world stage. It has been argued that "hegemonic" states seek to construct a global free-market system that provides them access to markets more cheaply than does formal control (Bergesen and Schoenberg 1980; Boswell 1989; Krasner 1976). This argument has much in common with Gallagher and Robinson's (1953) notion of an "imperialism of free trade" that metropolitan states preferred to the overhead costs of empire. Both Britain and the United States supported the collapse of rival empires and avoided empire building themselves during the periods of their global hegemony.

A second systemic factor is the contagiousness of decolonization. The American Revolution served as a model for insurrection in both Haiti and Latin America. After World War II, the independence of India, Indochina, and Indonesia had a substantial impact on colonized peoples everywhere. The contagiousness of independence is most apparent in the rapidity with which decolonization swept across Africa, where thirty-three colonies became independent between 1957 and 1966.

Finally, global political understandings and discourse play a critical role in decolonization

(Strang 1990). After 1945, the two superpowers were ideologically opposed to colonialism, though each accused the other of imperialism. Even the major imperial powers (Great Britain, France, the Netherlands) found it difficult to reconcile colonial possessions with the political ideas and institutions of the evolving national polity. As a result, the rationale for colonialism crumbled under pressure, with imperialism being formally denounced by the United Nations in 1960.

The social construction of a society or administrative unit as a "colonial dependency" is thus crucially important to mobilization around national independence, particularly given contemporary political understandings. A sharp distinction is generally drawn between the overseas empires of European states and "internal colonies" such as Wales, Armenia, or Eritrea. While the empirical basis for such a distinction is sometimes shaky, it is real in its consequences. Identification as a colonial dependency greatly increases the chances of mobilizing internal and external support for indigenous nationalism; it also vastly reduces the compulsion that the metropolitan state can legitimately bring to bear.

**Consequences of Decolonization.** The implications of decolonization for more general notions of international domination or exploitation are strongly contested. Dependency and world systems theorists view decolonization as producing change in the mere form, but not the content, of core-periphery relations (Chase-Dunn and Rubinson 1979). The argument is that contact between more and less developed economies tends generally to reinforce the differential between them, even in the absence of explicit political controls. Dependence on foreign capital has been argued to slow long-term economic growth (Bornschieer et al. 1978) and more generally to shape the political and economic structure of the dependent society (Amin 1973; Cardoso and Faletto 1979).

Despite these concerns, it is clear that decolonization involves a fundamental shift in the structures regulating international exchange, especially in the post-World War II era. Contemporary states are armed with widely accepted rights to control economic activity within their boundaries, including rights to nationalize foreign-owned

industries and renegotiate contracts with multinational corporations (Krasner 1978; Lipson 1985). Third World nations mobilize around these rights (Krasner 1985), and the negative impact of economic dependency seems to decrease when the peripheral state is strong (Delacroix and Ragin 1981).

While notions of dependency and neocolonialism are the subject of vigorous debate, a straightforward consequence of decolonization is the worldwide expansion of an originally European state society. Most of the present members of the United Nations became sovereign through decolonization (McNeely 1995). And historically, the political units emerging from decolonization have been strikingly stable. Few ex-dependencies have been recolonized or annexed or have merged or dissolved (Jackson and Rosberg 1982; Strang 1991b). Against much expectation, decolonization has produced states that figure as core players in the contemporary international economic and political order.

## REFERENCES

- Albertini, Rudolf von 1982 *Decolonization: The Administration and Future of the Colonies 1919–1960*. New York: Holmes and Meier.
- Amin, Samir 1973 *Neo-Colonialism in West Africa*. London: Monthly Review Press.
- Anderson, Benedict 1983 *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. London: Verso.
- Arrighi, Giovanni 1978 *The Geometry of Imperialism*. London: New Left Books.
- Bailyn, Bernard 1967 *The Ideological Origins of the American Revolution*. Cambridge, Mass.: Harvard University Press.
- Baran, Paul, and Paul Sweezy 1966 *Monopoly Capital*. New York: Monthly Review Press.
- Bergesen, Albert, and Ronald Schoenberg 1980 “The Long Waves of Colonial Expansion and Contraction 1415–1970.” In A. Bergesen, ed., *Studies of the Modern World System*. New York: Academic Press.
- Bornschiefer, Volker, Christopher Chase-Dunn, and Richard Rubinson 1978 “Cross-National Evidence of the Effects of Foreign Investment and Aid on Economic Growth and Inequality: A Survey of Findings and a Reanalysis.” *American Journal of Sociology* 84:651–683.
- Boswell, Terry 1989 “Colonial Empires and the Capitalist World-Economy: A Time Series Analysis of Colonization, 1640–1960.” *American Sociological Review* 54:180–196.
- Cardoso, Fernando H., and Enzo Faletto 1979 *Dependency and Development in Latin America*. Berkeley: University of California Press.
- Chase-Dunn, Christopher, and Richard Rubinson 1979 “Toward a Structural Perspective on the World-System.” *Politics and Society* 7:453–476.
- Cohen, Benjamin J. 1973 *The Question of Imperialism*. New York: Basic Books.
- Delacroix, Jacques, and Charles C. Ragin 1981 “Structural Blockage: A Cross-National Study of Economic Dependency, State Efficacy, and Underdevelopment.” *American Journal of Sociology* 86:1311–1347.
- Diamond, Sigmund 1958 “From Organization to Society.” *American Journal of Sociology* 63:457–475.
- Emerson, Rupert 1960 *From Empire to Nation*. Cambridge, Mass.: Harvard University Press.
- Fieldhouse, David K. 1966 *The Colonial Empires*. New York: Dell.
- 1973 *Economics and Empire, 1830–1914*. Ithaca, N.Y.: Cornell University Press.
- Folz, Robert 1969 *The Concept of Empire in Western Europe from the Fifth to the Fourteenth Century*. London: Arnold.
- Frank, Andre Gunder 1967 *Capitalism and Underdevelopment in Latin America*. New York: Monthly Review Press.
- Gallagher, John, and Ronald Robinson 1953 “The Imperialism of Free Trade.” *Economic History Review* S5:1–15.
- Gellner, Ernest 1983 *Nations and Nationalism*. Oxford: Basil Blackwell.
- Greene, Jack P. 1986 *Peripheries and Center: Constitutional Development in the Extended Politics of the British Empire and the United States 1607–1788*. Athens: University of Georgia Press.
- Grimal, Henri 1978 *Decolonization*. London: Routledge and Kegan Paul.
- Guenee, Bernard 1985 *States and Rulers in Later Medieval Europe*. Oxford: Basil Blackwell.
- Hancock, William K. 1950. *The Wealth of Colonies*. Cambridge: Cambridge University Press.
- Hilferding, Rudolf (1910) 1981 *Finance Capital*. London: Routledge and Kegan Paul.
- Hobson, John A. (1902) 1965 *Imperialism: A Study*. New York: Allen and Unwin.
- Holland, R. F. 1985 *European Decolonization 1918–1981*. New York: St. Martin’s Press.
- Jackson, Robert, and Carl Rosberg 1982 “Why Africa’s Weak States Persist: The Empirical and the Juridical in Statehood.” *World Politics* 35:124.

- Krasner, Stephen D. 1976 "State Power and the Structure of International Trade." *World Politics* 28:317-343.
- 1978 *Defending the National Interest*. Princeton, N.J.: Princeton University Press.
- 1985 *Structural Conflict*. Berkeley: University of California Press.
- Lang, James 1975 *Conquest and Commerce: Spain and England in the Americas*. New York: Academic Press.
- Lenin, Vladimir Ilich (1917) 1939 *Imperialism: The Highest Stage of Capitalism*. London: International Publishers.
- Lipson, Charles 1985 *Standing Guard: Protecting Foreign Capital in the Nineteenth and Twentieth Centuries*. Berkeley: University of California Press.
- Lugard, Frederick J. D. (1922) 1965 *The Dual Mandate in British Tropical Africa*, 5th ed. London: Cass.
- Magdoff, Harry 1969 *Age of Imperialism*. New York: Monthly Review Press.
- McNeely, Connie L. 1995 *Constructing the Nation-State: International Organizations and Prescriptive Action*. Westport, Conn.: Greenwood.
- Nkrumah, Kwame 1966 *Neo-Colonialism: The Last Stage of Imperialism*. London: International Publishers.
- Nogueira, Franco 1963 *The United Nations and Portugal: A Study of Anti-Colonialism*. London: Sidgwick and Jackson.
- Robinson, Ronald 1972 "Non-European Foundations of European Imperialism: Sketch for a Theory of Collaboration." In Roger Owen and Bob Sutcliffe, eds., *Studies in the Theory of Imperialism*. Bristol, U.K.: G. B. Longman.
- , John Gallagher, and Alice Denny 1961 *Africa and the Victorians: The Official Mind of Imperialism*. London: St. Martin's Press.
- Schumpeter, Joseph 1951 "The Sociology of Imperialism." In *Imperialism and Social Classes*. New York: Kelley.
- Smith, Tony 1978 "A Comparative Study of French and British Decolonization." *Comparative Studies in Society and History* 20:70-102.
- Strang, David 1990 "From Dependency to Sovereignty: An Event History Analysis of Decolonization." *American Sociological Review* 55:846-860.
- 1991a "Global Patterns of Decolonization 1500-1987." *International Studies Quarterly* 35:429-454.
- 1991b "Anomaly and Commonplace in European Political Expansion: Realist and Institutional Accounts." *International Organization* 45:142-161.
- 1992 "The Inner Incompatibility of Empire and Nation: Popular Sovereignty and Decolonization." *Sociological Perspectives* 35:367-384.

Wallerstein, Immanuel 1974, 1980, 1989 *The Modern World-System*, 3 vols. Cambridge: Cambridge University Press.

Waltz, Kenneth 1979 *Theory of International Politics*. Reading, Mass.: Addison-Wesley.

Weber, Max 1958 *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribners.

DAVID STRANG

## IMPRESSION FORMATION

See Affect Control Theory and Impression Formation; Social Psychology.

## INCEST

Incest is illicit sex or marriage between persons socially or legally defined as related too closely to one another. All societies have rules regarding incest. Incest is conceptualized in four ways: as a proscribed or prescribed marriage form, as a taboo, as prohibited coitus, and as child abuse. The first three conceptualizations are most closely related to early scholars (mid-1800s to mid-1900s), who tended to overlap them. The last conceptualization has become prominent more recently.

Incest-as-marriage rules are usually proscriptive ("Thou shalt not"). Prescriptive ("Thou shalt") incestuous marriage rules have been documented for royalty in Old Iran and ancient Egypt and for Mormons in the United States (Lester 1972). Some groups historically encouraged brother-sister incest. Cases in point were the ruling families of Egypt and Polynesia, where preservation of family resources and ethnic identity took precedence over political alliances with other groups (Firth 1936, 1994).

That some groups proscribe while others prescribe incestuous marriages has caused some to be skeptical of many theories about incest, especially theories that assume that incest avoidance is natural, that close inbreeding is genetically disadvantageous, or that there is an incest taboo. John F. McLennan ([1865] 1876), a lawyer, coined the terms *endogamy* (within-the-group marriage) and



*exogamy* (outside-the-group marriage). He defined incest as endogamy. Based on his analyses of marriage in Ireland, Australia, ancient Greece, and other societies, he concluded that rules proscribing endogamy evolved as a group survival mechanism. He reasoned that as members of one tribe or group married into others, “blood ties” emerged. These blood ties encouraged reciprocity between groups and cooperation and harmony within groups.

Anthropologist Lewis Henry Morgan ([1877] 1964) generally agreed with McLennan’s definition of incest as endogamy and with McLennan’s assumption that proscribing incest promoted both exogamy and group survival. Morgan assumed, however, that incest originally became prohibited due to the presumed deleterious effects of inbreeding. That is, Morgan assumed that at some point, humans developed the capacity to see that exogamy increased the “vigor of the stock” (p. 65). Moreover, in agreement with sociologist Herbert Spencer’s ([1876] 1898, pp. 623–642) earlier critique of McLennan’s application of the terms “endogamy” and “exogamy,” Morgan ([1877] 1964) contended that given the very contradictory nature by which kinship is recognized and organized among humans, McLennan’s “terms and his conclusions are of little value” (p. 432). Finally, Morgan contended that practices such as brother-sister and father-daughter marriage among Hawaiians, for instance, spoke to earlier, savage stages of human existence. To Morgan, human groups evolve through three basic epochs: savagery, barbarism, and civilization. The first two epochs each have two stages, the last one only one stage. In Morgan’s schema of the evolution of humankind, at the very first stage there is promiscuous intercourse and the marriage of brothers and sisters. In the second stage, brother-sister marriage is prohibited. The third stage of evolution sees prominence given to the Gentile pattern of social organization, which includes monogamy. The patriarchal family characterizes the fourth stage. The fifth, and final, stage in the schema, arising with the concept of property, is hallmarked by the monogamian family and the superiority of the “Aryan, Semitic, and Uralian system of consanguinity and affinity” (Morgan [1877] 1964, pp. 421–422). To Morgan, the Aryan family represented the highest level of human evolution, and thus the epitome of civilized humans. This high civilization was attributed to the “providence of God” (p.

468). To Morgan, in this civilized arrangement, cousin marriage is prohibited. In reading Morgan’s work, one gets the impression that he saw this last stage as the best and the highest, and seemed to personally identify with it. Thus, it is somewhat curious that he married his cousin, Mary Elizabeth Steele.

Tylor (1889) elaborated on McLennan’s notion of exchange and reciprocity among exogamous groups. He noted that men made political alliances with men in other groups by exchanging women in marriage. Spencer ([1876] 1898) argued that endogamy was probably the original practice among early peoples, especially among peaceful tribes. Exogamy probably began through wife stealing. He assumed that some early groups probably stole wives from other groups because they faced a scarcity of women. Other groups forcefully captured wives in the process of war. Thus, Spencer positioned exogamy as originally speaking to extremely brutal treatment of women. He assumed that as time passed and circumstances changed, exogamy would be seen as advantageous to groups who could make the linkages between sexual intercourse and procreation. That is, these groups might come to see exogamy as leading to a healthier stock of people. In addition, he assumed that, over time, such exchanges also would be seen as creating bonds among groups who otherwise might compete for scarce resources or else annihilate each other through warfare.

The common thread among these theories is the focus on incest rules as social organizational principles. A shared weakness is the reliance on analyses of primitive or premodern groups. These inclinations are also found in varying degrees in writings by Sir James Frazer (1927), Brenda Seligman (1929, 1932), Robert Briffault (1930), Bronislaw Malinowski (1922, 1927, 1929), George Murdock (1949), and Claude Levi-Strauss ([1949] 1969). Many of these and other writers also fail to differentiate between incest rules and exogamy rules. What one group calls incestuous marriage may not biologically be such. Likewise, some groups’ exogamy rules permit biologically incestuous marriage. Recognizing that incest rules are socially defined, Sumner (1906) presaged today’s sociobiologists. Sumner argued that incest rules should be modified as researchers gathered genetic evidence that

dispelled fallacious beliefs that all incestuous matings are deleterious.

Sociologist Emile Durkheim's *Incest: The Nature and Origin of the Taboo* ([1898] 1963) is an infrequently cited magnum opus. Based on ethnographic research in Australia, this book emphasized and illustrated the social and moral origins of incest taboos. Durkheim noted the ways in which prohibitions against incest and penalties for rule violations organize social groups internally. Cooperation and alliances with other groups via exogamy are consequently prompted. Durkheim contended that the incest taboo has a religious origin. It is derived from the clan's sentiments surrounding blood, specifically menstrual blood. While blood is taboo in a general way, contact with the blood of the clan is taboo in specific ways. Durkheim reasoned that menstrual blood represents a flowing away of the clan's lifeblood. This renders women taboo for and inferior to men of the same clan. The taboo relates to intercourse in general and to marriage in particular. Blood and incest are presumed to be related such that a man violating the taboo is seen as a murderer. Thus, in a real sense, Durkheim posited women as having a dual nature, being at once both sacred and profane. That women bear children, thus rendering a clan immortal, made them sacred. However, since they spilt blood, which itself is sacred, they had a profane nature to them. To marry a clan's women to another group kept the clan sacred, while at the same making it possible for other clans to perpetuate themselves.

Durkheim contended that the origin of the taboo is lost to the consciousness of clans over time but that the taboo itself is replaced by a generalized repugnance of incest. This repugnance prompts men to exchange women with other groups, thus facilitating political alliances between men of different groups. In addition, in his *The Division of Labor in Society* ([1893] 1963), Durkheim posited that when incest loses its religiously based criminal status, prohibitions against it are or will be codified into law.

While brother-sister marriages are negatively sanctioned in many parts of the world, the United States contrasts sharply with European and other nations in having civil laws that prohibit cousin marriage as well (Ottenheimer 1990). Roughly 60

percent of U.S. states prohibit cousin marriage. Many of the states have prohibitions placed on marriages other than brother-sister or cousin-cousin. For instance, Alaska prohibits marriage to persons more closely related than the fourth degree of consanguinity, regardless of whether this relationship is whole- or half-blood. Both Rhode Island and South Carolina forbid men to marry their stepmother, grandfather's wife, wife's daughter, and various other affinal or blood kin. Both states forbid women to marry their husband's grandfather, son's daughter's husband, daughter's husband, stepfather, or grandmother's husband (Kessinger 1990). Many of these laws no doubt were influenced by persons such as Morgan, with his emphasis on social evolution and his contention that civilized people simply don't engage in incestuous pairings. It is probably also the case that the eugenics movement, of the late 1800s through roughly the early 1900s, had some effect on maintaining such laws. The word "eugenics" was coined by biologist Francis Galton and meant improvement in the human race through selective breeding (Parrinder 1997). Galton and his followers were concerned with an array of social issues that they thought could be minimized if not eradicated through eugenics. These included overpopulation caused by high fertility rates of immigrants, poor people, and other classes of people deemed to be of inferior stock. Given a belief that criminals biologically inherited their antisocial tendencies, proponents of eugenics also believed that sterilization and birth control were necessary among some people so that they did not bring forth more generations of criminals (Rafter 1997). Fears of the negative effects of human inbreeding were also high at this time, resulting in several U.S. states passing laws that made cousin-cousin marriage a criminal offense (Gibbons 1993). The eugenics movement is also associated with social Darwinism and the Nazi regime. However, both of these topics are beyond the purview of this article.

Marriage laws and moral crusades aside, endogamous unions have continued to occur in the United States. An analysis of marriage patterns in Madison County, Virginia, 1850-1939, revealed a rather high rate of marriage between first cousins and first cousins once-removed (Frankenberg 1990). Possible reasons for these patterns may include restricted physical mobility and preference for

socially similar others. Relying on U.S. Census data, another study revealed a fairly high rate of endogamy among Louisiana Cajuns (Bankston and Henry 1999). Partial explanations for this endogamy include restricted residential mobility, ethnic identity emphases of the subculture, and a preference for social-class homogamy. The rate of biologically incestuous unions among this group, however, was not made clear in the report.

The concern over genetic transmission of physical and mental defects through marriage to close relatives continues today. In an analysis of consanguineous marriage and health profiles of offspring from these unions, researchers in Dammam City, Eastern Province, Kingdom of Saudi Arabia, found that of the 1,307 unions studied, 52 percent were consanguineous. The most common consanguineous union was among first cousins. These consanguineous unions are associated with culture values and efforts to keep property and resources in the family. The study revealed that babies born of a consanguineous union tended to be smaller than those born of a nonconsanguineous union (Al-Abdulkareem and Ballal 1998). A recent American study focused on the relationship between consanguinity and childhood mortality in an Old Order Amish settlement. It was found that the Amish in the Lancaster, Pennsylvania, area are at risk for genetically transmitted defects that result in increased mortality rates. In addition, consanguineous unions resulted in higher neonatal and postnatal death rates than did nonconsanguineous unions (Dorsten et al. 1999).

Sigmund Freud focused on the incest taboo ([1913] 1950) and infantile sexuality ([1905] 1962). Through his tale of the primal horde, Freud posited that in the original family there was a jealous and violent father who engaged his daughters in incest. The jealous brothers banded together, killed, and ate the father. Horrified by their deeds, the brothers made incest taboo. For Freud, this (the moment when humans created incest rules) was when humans became social. He assumed that very young (Oedipal) children sexually desire their opposite-sex parent. Little boys then suffer from castration anxiety, fearing that the father will become aware of their desires and punish them. While little girls see their mothers as inferior to men, they also know them to be more powerful than they are.

Thus, little girls resignedly align themselves with their mothers, repress their incestuous impulses, and experience penis envy. Even though Freud's theories were based on conjecture and focused on the bourgeois nuclear family of his day, they have nonetheless influenced many social scientists, including French anthropologist Jean Claude Levi-Strauss ([1949] 1969) and American sociologist Talcott Parsons (1954).

Levi-Strauss agreed with Freud that humans became social with the creation of incest rules. Borrowing from the emphasis on the exchange of women, found in the works of Durkheim and Marcel Mauss, and Durkheim's assumption that women of a clan become the symbol (totem) for the clan, Levi-Strauss posited that exogamy represents a special form of alliance-creating reciprocity. Since incest involves sexuality, incest rules and the exchange of women represent a unique social connection between the biological and the cultural. Although frequently lauded by social scientists, the work of Levi-Strauss adds little to previous theories about incest.

Parsons incorporated Freud's theory of infantile sexuality into his structural-functionalist view of the American nuclear family. Like functionalists before him, he assumed that incest rules exist to prevent role confusion within the nuclear family and to encourage alliances with other families (see also Malinowski 1927, 1929). He contended that the mother was to exploit her son's Oedipal desires as if she had him on a rope. At earlier ages she was to pull him toward her, encouraging heterosexuality. At later ages she was to push him away, encouraging him to establish relationships with nonrelated females. Tugging and pulling on the rope were also designed to assist the son in internalizing society's incest rules and guide him in creating his own nuclear family. Parsons assumed that little girls also experience an erotic attachment to their mothers. He argued that as the mother severed this attachment, it was her responsibility to instill in the daughter an aversion to both father-daughter and brother-sister incest. Failure to do this would result in family disorganization and in the daughter's inability to become a normatively functioning adult. Parsons contended that incest aversion would be realized if the mother kept the erotic bond with her husband intact.

Parsons's analysis has thus assisted in perpetuating what child abuse researchers have sought to eradicate: placing blame on mothers when fathers incestuously abuse their daughters (see Finkelhor 1984; Russell 1986; Vander Mey and Neff 1986).

Psychologist Edward Westermarck ([1891] 1922) contended that family and clan members develop a sexual aversion to one another due to the dulling effects of daily interaction and the sharing of mundane tasks. This aversion prompts the development of laws and customs proscribing incest among persons with a shared ancestor and set of obligations based on clan membership. Westermarck noted that failure to develop this aversion and the propensity to violate incest laws were due to alcoholism, membership in the lower social classes, inability to control the sex drive, lack of alternative sexual outlets, social and geographic isolation, and failure to have developed normative, family-like feelings of duty. Variations on Westermarck's thesis appear in writings today. Some focus on a learned aversion that Oedipal children develop because their fantasies cannot be realized, due to their lack of full sexual maturation. Westermarck's thesis is weak because if there were a natural aversion to incest, then laws prohibiting it would be unnecessary. Furthermore, a growing body of literature suggests that persons who share mundane tasks and interact daily do not necessarily develop a sexual aversion to one another (see Vander Mey and Neff 1986).

Sociobiologists (e.g., Parker 1976; van den Berghe 1980) variously incorporate Freudian theses, Westermarck's thesis, and the assumption that human social behavior follows a fitness maxim to argue that groups establish prescriptive or proscriptive incest rules to enhance a group's survival by facilitating genetic advantages. Sociobiologists typically use the functionalist assumption that incest rules regulate the internal dynamics of the procreating group and encourage affiliation with other nuclear groups. Violations of incest rules are assumed to be caused by the factors identified by Westermarck. Many sociobiologists link intercourse directly to procreation. The assumption has no merit when one tries to explain why adult-child and child-child intercourse occurs. Equally problematic is the frequent reliance on the rules prescribing incest among royals. Dismissed is the role that

ethnocentrism plays in such rules. Finally, serious questions arise when research on primates, birds, or other nonhumans is extrapolated to human social behavior and organization. In sum, it is questionable whether the sociobiological arguments for, and data supporting the existence of, incest-avoidance mechanisms in humans can be confidentially accepted (Leavitt 1990).

A new twist on sociobiological discourse on the incest taboo comes from the field of genetic psychology. It is proposed that the taboo exists to protect children's "mating-strategy template." It is assumed that humans have a mating-strategy template that effectively gears females toward attracting men who can provide for them and their children. The human biogram also orients males toward being able to garner valuable resources for themselves, which they can then share with mates and children. It is assumed that should adult-child incest occur, male and female children who experience it will have disruptions in their mating-strategy template. In effect, if incest occurs, then the developing girl child will suffer various psychopathologies and will be unable to adequately compete for suitable mates. This will make her less attractive as a potential mate. Boys who experience father-son incest will lose the ability to be instigators of mating since it already has been imposed on them. Boys then become targets. It is assumed that by lowering children's attractiveness as mates, their reproductive success is also lowered. With this lowered reproductive success, future generations are negatively affected as well. Therefore, it is argued that incest taboos have been instituted as a way to avoid these long-term consequences for greater intergenerational success in mating outside one's own group (Immerman and Mackey 1997).

Prefiguring more recent research on incest as child abuse, Durkheim wrote that incest violations are most likely to occur in families in which members do not feel morally obligated to be dutiful to one another and to practice moral restraint ([1898] 1963). In many parts of the world today, incest is seen as a serious and severe type of child abuse. As child abuse, incest is any form of sexual touching, talking, or attempted or actual intercourse between an adult and a child or between two children when the perpetrator is significantly older than the victim or forces the victim to engage in actions against his or her will. The perpetrator and victim

are related either by consanguinity or affinity. Incest is abuse because it harms the victim and violates the child's basic human rights. Its negative effects linger throughout victims' lives (Russell 1986). That the sexual victimization of children is nothing new, but rather ancient, is well documented (deMause 1982; Rush 1980). Public recognition of incest as child abuse, however, has its roots in research and public discourse dating from roughly the early 1970s.

Had Freud not abandoned his first psychoanalytical paper, "The Aetiology of Hysteria" ([1896] 1946), he could have been heralded as the savior of children. In this paper he described incest and other sexual abuse experienced by eighteen patients when they were children. He linked his adult patients' problems to their experiences of childhood sexual trauma. However, when his paper was coldly received and then ignored by senior psychologists, and Freud faced professional ostracism, he dismissed these findings (see Jurjevich 1974; Masson 1984). He reasoned that his patients were incorrectly recalling masturbatory childhood fantasies of sexual encounters with adults (Freud [1905] 1962). Freud then associated adult neurosis and hysteria with unfulfilled childhood incest fantasies.

Freud's revamped theories quickly became popular. Several writers then recounted case studies of children who had experienced sex with an adult. The children were described as seductive and provocative. They were not seen as victims; rather, they were seen as sexual initiators (see Vander Mey and Neff 1986). Although Swedish sociologist Svend Riemer (1940) and American sociologist S. Kirson Weinberg (1955) tried to bring attention to incest as a form of family deviance that sometimes resulted in harm to victims, Freudian theory held fast until the late 1960s. The "discovery" of child abuse, in conjunction with the civil rights movement, the anti-Vietnam War movement, and the resurgence of the women's rights movement, refocused attention such that the physical abuse of children and rape of women became seen as serious wrongs inflicted on other humans and as special types of social problems (Vander Mey and Neff 1986). Since that time, theories and research have explored in great detail why and how incestuous abuse occurs and its lasting negative effects. Laws protecting children against such abuse have been enacted. These laws mandate that

treatment services be available for victims, perpetrators, and families (Fraser 1987). They also stipulate penalties (fines, imprisonment) for convicted abusers. These operate on a sliding scale, with harsher penalties applied to cases in which the victim is very young or has suffered serious physical or psychological trauma because of the abuse. Some cases are handled in family courts. Others, especially those involving serious harm to the child, are heard in criminal courts. Although variable in scope, laws in European nations prohibit and punish incestuous child abuse (Doek 1987). As with statutes in the United States, applicable penalties often depend on the victim-perpetrator relationship, the age of the victim, and the seriousness of harm to the victim. Efforts to refine legal statutes and penalties continue. Laws are also needed that more specifically define and punish child-child incestuous abuse. While the bulk of research on incestuous abuse has focused on adult-child incest, there is accumulating evidence to the effect that sibling incest is more common than previously thought and also can be traumatizing (Adler and Schutz 1995; Vander Mey 1988).

Feminists usually focus on father-daughter incest. They refer to it as rape to emphasize the specific type of abuse that it is. The term "rape" also illustrates the point that if an adult male forced another person to engage in intercourse with him, he could be arrested. However, if he rapes his own child, it is called incest, which is often seen as a disgusting, private, family problem (Brownmiller 1975). Feminists see incest as commonplace, originating in and perpetuated by patriarchy (Rush 1980). A general feminist approach to father-daughter incest incorporates a discussion of the sex-role socialization of males and the male-as-superior patriarchal ideology as causal factors in incest. The contention is that females ultimately are rendered second-class citizens, the property of men, and sexual outlets for men (see Rush 1980). Problems with a general feminist perspective on incest include a focus limited to father-daughter incest, the monocausal linkage of patriarchy to incest, the portrayal of all incest perpetrators as male and all incest victims as female, and the oversimplified view of male sex-role socialization (Vander Mey 1992; see also Nelson and Oliver 1998). American feminist sociologist Diana Russell (1986) extends and refines the general feminist approach to incest. She argues that males are

socialized to sexualize the power given them by virtue of the fact that they are male. This includes sexualizing the power they have over their own children. Moreover, Russell recognizes that mothers and children can be incest perpetrators. She advocates androgynous socialization of children, more equality between men and women, and more public awareness of the harsh realities of incestuous abuse.

Welsh psychologist Neil Frude (1982) relied on existing empirical and clinical research to articulate a five-factor explanatory model of father-daughter incest. These five factors are sexual need (of the perpetrator), attractive partner, opportunity, disinhibition, and sexual behavior. A strength of Frude's model is the attention paid to the interweaving of sex and power. Weaknesses include the fact that incest perpetrators are not usually sexually deprived, victim attractiveness is often irrelevant, and families are not usually closed systems today. American sociologist David Finkelhor (1984) offers a somewhat similar model, although it differs from Frude's model in that Finkelhor pays keen attention to the myriad ways in which larger social forces and cultural ideology are related to child sexual abuse. Finkelhor's model has the added strength of being applicable to several types of incest and other sexual abuse.

American sociologists Brenda Vander Mey and Ronald Neff (1986) constructed a research-based ecological model of father-daughter incest. They begin with the assumption that there is no incest taboo. Rather, there are rules proscribing adult-child incest. They contend that characteristics of the society, the neighborhood, the family and the marital dyad, and the father-daughter dyad differentially affect the probability that a daughter will be sexually abused by her father. These levels of influence interact in complicated ways. Father-daughter incest is associated with dominance in society and in the family, residence in a violent neighborhood, social isolation of the family, family disorganization, and the father's lack of empathy for his wife and children. Mitigating factors decreasing the likelihood of incest include a father's conformity to rules against incest, sex education of the daughter, and media attention to adult-child incest as wrong and illegal. This model is strong in its reliance on research and theoretical principles. Although this model is limited to father-daughter incest, it does provide

information that can assist in identifying children at risk for incestuous abuse.

Incestuous abuse of children, male and female alike, is associated with an array of problems for victims. These problems include bulimia, self-mutilation, alcohol abuse, criminal behavior such as assault and shoplifting, psychological impairment, sexually aggressive behavior, and sexual hyperarousal (Araji 1997; Green 1993; Kinzl and Biebl 1992; Ryan et al. 1996; Wonderlich et al. 1996).

Incestuous abuse is seen as an international social problem today. At least four journals frequently carry the latest research on this topic: *Child Abuse and Neglect*, *Journal of Family Violence*, *The Journal of Interpersonal Violence*, and *The Journal of Child Sexual Abuse*.

However, not everyone sees adult-child sexual contact as abusive. It has been argued that American and European cultures have long tended to deny that children have a sexuality. That is, experts on child development have for many decades focused on every other aspect of human development but have ignored child sexuality. In addition, audiences have been less than receptive to the idea that children might enjoy sexual contact with other children—or with adults—and that this contact might be conducive to their development in a constructive way. It has been argued that taking the cultural constraints away from children's need to develop their sexuality is part of a child's right (Martinson 1973, 1994; Yates 1978). Due to the very sensitive nature of the topic, few empirical studies on child sexuality have been conducted. Of the few available studies, one was a longitudinal study of sexual experiences in early childhood. It revealed that most of the children focused on in the study had engaged in some form of sex play by age six (Okami et al. 1997). However, another extensive study of college students revealed that among those who had experienced sexual contact while children, sexual contact with an adult was more often seen as unwanted, unwelcome, and/or abusive (Nelson and Oliver 1998).

Organizations such as the North American Man Boy Love Association contend that the child abuse laws constrain children's liberty. Some parents, particularly in the United States and Europe, have resisted the child sexual abuse laws as infringing on their right to rear their children in ways they see as most appropriate (for elaboration, see Beckett

1996; Hechler 1988). Thus, just as marriage laws and cultural groups vary in their definitions of what constitutes incestuous marriage, we now are seeing public discourse on how best to define sexual contact with children.

## REFERENCES

- Adler, Naomi A., and Joseph Schutz 1995 "Sibling Incest Offenders." *Child Abuse and Neglect* 19:811-819.
- Al-Abdulkareem, Abdulkareem A., and Seifeddin G. Ballal 1998 "Consanguineous Marriage in an Urban Area of Saudi Arabia: Rates and Adverse Health Effects on the Offspring." *Journal of Community Health* 23:75-83.
- Araji, Sharon Kay 1997 *Sexually Aggressive Children: Coming to Understand Them*. Thousand Oaks, Calif.: Sage.
- Bankston, Carl, III, and Jacques Henry 1999 "Endogamy Among Louisiana Cajuns: A Social Class Explanation." *Social Forces* 77:1317-1338.
- Beckett, Katherine 1996 "Culture and the Politics of Signification: The Case of Child Sexual Abuse." *Social Problems* 43:57-76.
- Briffault, Robert 1930 *Rational Evolution: The Making of Humanity*. New York: Macmillan.
- Brownmiller, Susan 1975 *Against Our Will: Men, Women and Rape*. New York: Simon and Schuster.
- deMause, Lloyd 1982 *Foundations of Psychohistory*. New York: Creative Roots.
- Doek, Japp E. 1987 "Sexual Abuse of Children: An Examination of European Criminal Law." In P. B. Mrazek, and C. H. Kempe, eds., *Sexually Abused Children and Their Families*. New York: Pergamon.
- Dorstein, Linda Eberst, Lawrence Hotchkiss, and Terry M. Kind 1999 "The Effect of Inbreeding on Early Childhood Mortality: Twelve Generations of an Amish Settlement." *Demography* 36:263-272.
- Durkheim, Emile (1893) 1963 *The Division of Labor in Society*, trans. G. Simpson. New York: Free Press.
- (1898) 1963 *Incest: The Nature and Origin of the Taboo*, trans. A. Ellis. New York: Lyle Stuart.
- Finkelhor, David 1984 *Child Sexual Abuse: New Research and Theory*. New York: Free Press.
- Firth, Raymond 1936 *We, the Tikopia: A Sociological Study of Kinship in Primitive Polynesia*. London: Allen Unwin.
- 1994 "Contingency of the Incest Taboo." *MAN* 29:712-713.
- Frankenberg, Susan R. 1990 "Kinship and Mate Choice in a Historic Eastern Blue Ridge Community, Madison County, Virginia." *Human Biology* 62:17-35.
- Fraser, Brian G. 1987 "Sexual Child Abuse: The Legislation and the Law in the United States." In P. B. Mrazek and C. H. Kempe, eds., *Sexually Abused Children and Their Families*. New York: Pergamon.
- Frazer, Sir James George 1927 *Man, God and Immortality*. New York: Macmillan.
- Freud, Sigmund (1896) 1946 "The Aetiology of Hysteria." In E. Jones, ed., and J. Riviere, trans., *Collected Papers*. New York: International Psycho-Analytic Press.
- (1905) 1962 "My Views on the Part Played by Sexuality in the Aetiology of the Neurosis." In P. Reiff, ed., *Sexuality and the Psychology of Love*. New York: Collier Books.
- (1913) 1950 *Totem and Taboo*, trans. J. Strachey. New York: Norton.
- Frude, Neil 1982 "The Sexual Nature of Sexual Abuse: A Review of the Literature." *Child Abuse and Neglect* 6:211-223.
- Gibbons, Ann 1993 "The Risks of Inbreeding." *Science* 259(5099): 1252.
- Green, Arthur H. 1993 "Child Sexual Abuse: Immediate and Long-Term Effects and Intervention." *Journal of the American Academy of Child and Adolescent Psychiatry* 32:890-992.
- Hechler, David 1988 *The Battle and the Backlash: The Child Sexual Abuse War*. Lexington, Mass.: Lexington Books.
- Immerman, Ronald S., and Wade C. Mackey 1997 "An Additional Facet of the Incest Taboo: A Protection of the Mating-Strategy Template." *Journal of Genetic Psychology* 158:151-164.
- Jurjevich, Ratibor-Ray M. 1974 *The Hoax of Freudism: A Study of Brainwashing the American Professionals and Laymen*. Philadelphia: Dorrance.
- Kessinger, Roger A. 1990 *Marriage Licensing Laws: A State by State Guide*. Boise, Idaho: Kessinger Publishing Company.
- Kinzl, Johannes, and Wilfried Biebl 1992 "Long-Term Effects of Incest: Life Events Triggering Mental Disorders in Female Patients with Sexual Abuse in Childhood." *Child Abuse and Neglect* 16:576-573.
- Leavitt, Gregory C. 1990 "Sociobiological Explanations of Incest Avoidance: A Critical Review of Evidential Claims." *American Anthropologist* 92:971-993.
- Lester, David 1972 "Incest." *The Journal of Sex Research* 8:268-285.
- Levi-Strauss, Claude (1949) 1969 *The Elementary Structure of Kinship*, trans. J. H. Bell, J. R. von Sturmer, and R. Needham. Boston: Beacon Press.
- Malinowski, Bronislaw 1922 *Argonauts of the Western Pacific*. New York: E. P. Dutton.

- 1927 *Sex and Repression in Savage Society*. London: Routledge and Kegan Paul.
- 1929 *The Sexual Life of Savages in North-Western Melanesia*. London: Routledge and Kegan Paul.
- Martinson, Floyd M. 1973 *Infant and Child Sexuality: A Sociological Perspective*. St. Peter, Minn.: The Book Mark.
- 1994 *The Sexual Life of Children*. Westport, Conn.: Greenwood Press.
- Masson, Jeffrey M. 1984 *The Assault on Truth: Freud's Suppression of the Seduction Theory*. New York: Farrar, Straus, and Giroux.
- McLennan, John F. (1865) 1876 *Primitive Marriage: An Inquiry into the Origin of the Form of Capture in Marriage Ceremonies*. London: Bernard Quaritch.
- Morgan, Lewis Henry (1877) 1964 *Ancient Society*, ed. Leslie A. White. Cambridge, Mass.: The Belknap Press of Harvard University Press.
- Murdock, George Peter 1949 *Social Structure*. New York: Macmillan.
- Nelson, Andrea, and Pamela Oliver 1998 "Gender and the Construction of Consent in Child-Adult Sexual Contact: Beyond Gender Neutrality and Male Monopoly." *Gender and Society* 12:554-577.
- Okami, Paul, Richard Olmstead, and Paul R. Ambramson 1997 "Sexual Experiences in Early Childhood: 18-Year Longitudinal Data from the UCLA Family Lifestyles Project." *The Journal of Sex Research* 34:339-347.
- Ottenheimer, Martin 1990 "Lewis Henry Morgan and the Prohibition of Cousin Marriage." *Journal of Family History* 15:325-334.
- Parker, Seymour 1976 "The Precultural Basis of the Incest Taboo: Toward a Biosocial Perspective." *American Anthropologist* 78:285-305.
- Parrinder, Patrick 1997 "Eugenics and Utopia: Sexual Selection from Galton to Morris." *Utopian Studies* 8:1-11.
- Parsons, Talcott 1954 "The Incest Taboo in Relation to Social Structure and the Socialization of the Child." *British Journal of Sociology* 5:101-117.
- Rafter, Nichole Hahn 1997 *Creating Born Criminals*. Champaign: University of Illinois Press.
- Riemer, Svend 1940 "A Research Note on Incest." *American Journal of Sociology* 45:566-575.
- Rush, Florence 1980 *The Best Kept Secret: Sexual Abuse of Children*. New York: McGraw-Hill.
- Russell, Diana E. H. 1986 *The Secret Trauma: Incest in the Lives of Girls and Women*. New York: Basic Books.
- Ryan, Gail, Thomas J. Miyoshi, Jeffrey L. Metzner, Richard D. Krugman, and George E. Fryor 1996 "Trends in a National Sample of Sexually Abusive Youths." *Journal of the American Academy of Child and Adolescent Psychiatry* 35:17-25.
- Seligman, Brenda Z. 1929 "Incest and Descent: Their Influence on Social Organization." *Royal Anthropological Institute of Great Britain* 59:231-272.
- 1932 "The Incest Barrier: Its Role in Social Organization." *British Journal of Psychology* 22:250-276.
- Spencer, Herbert (1876) 1898 *The Principles of Sociology*, vol. 1. New York: D. Appleton and Company.
- Sumner, William Graham 1906 *Folkways: A Study of the Sociological Importance of Usages, Manners, Customs, Mores, and Morals*. Boston: Ginn.
- Tylor, Edward B. 1889 "On the Method of Investigating the Development of Institutions; Applied to Laws of Marriage and Descent." *Journal of the Royal Anthropological Institute* 18:245-269.
- van den Berghe, Pierre 1980 "Incest and Exogamy: A Sociobiological Reconsideration." *Ethology and Sociobiology* 1:51-162.
- Vander Mey, Brenda J. 1988 "The Sexual Victimization of Male Children: A Review of Previous Research." *Child Abuse and Neglect* 12:61-72.
- 1992 "Theories of Incest." In W. O'Donohue, and J. Geer, eds., *The Sexual Abuse of Children: Research, Theory, and Therapy*. Hillsdale, N.J.: Erlbaum.
- , and Ronald L. Neff 1986 *Incest as Child Abuse: Research and Applications*. New York: Praeger.
- Weinberg, S. Kirson 1955 *Incest Behavior*. New York: Citadel Press.
- Westermarck, Edward (1891) 1922 *The History of Human Marriage*. New York: Allerton.
- Wonderlich, Stephen, Mary Ann Donaldson, David K. Carson, Dennis Staton, Linda Gertz, Laurie R. Leach, and Maureen Johnson 1996 "Eating Disturbances and Incest." *Journal of Interpersonal Violence* 11:195-207.
- Yates, Alayne 1978 *Sex Without Shame: Encouraging the Child's Healthy Sexual Development*. New York: William Morrow.

BRENDA J. VANDER MEY

## INCOME DISTRIBUTION IN THE UNITED STATES

As the twentieth century drew to a close, the gap between rich and poor in the United States, and between the wealthy and impoverished nations of the world, continued to widen. This entry extends the analysis presented in the 1992 edition of the



*Encyclopedia* (largely based on 1988 data) to include information on income, wealth, and poverty gathered from studies conducted in the mid- and late 1990s by various government departments, academic economists, and private research institutions.

### INCOME DISTRIBUTION

Data on money income for households, families, and persons in the United States are collected yearly by the Bureau of the Census, based on a national probability sample of approximately 50,000 households. The official definition of “income” is money income before taxes, including some government transfers (Social Security benefits, welfare payments, worker’s compensation), returns on investments, and pensions, but excluding capital gains or health insurance supplements paid by employers. In recent years, the Census Bureau has also published computations based on fourteen alternative definitions of income; for example, after deducting a range of taxes or after adding the earned income credit, the value of noncash government benefits such as food stamps, school lunches, and Medicare or Medicaid reimbursements to providers. Some of these alternative measures reduced the level of income inequality, while others increased it, but the general finding is that government transfers have a significantly stronger effect on reducing inequality than does the current tax system. For the sake of simplicity and consistency, this entry will primarily use data based on the official definition.

**Household Income.** In 1997, the median income for all 102 million U.S. households was \$37,005, an increase of almost 2 percent over 1996, continuing an upward trend from a low of \$34,700 in 1993. Adjusted for inflation, however, the 1997 median was slightly lower than the previous high of \$37,303 in 1989. As Table 1 indicates, there are significant differences among subgroups in median income, and while all but male nonfamily householders and Asian/Pacific Islanders experienced an increase in real income over the previous year, data covering the entire period 1989–1997 present a mixed picture. Gains between 1989 and 1997 were highest for black households, for both married-couple family households and those headed by a female householder, and for persons aged

55–64. Losing ground over the eight years were male family householders without a wife, male nonfamily householders, Hispanic households, and households headed by young adults age 15–24 as well as by those 35–54 years old.

With respect to region and place of residence, households outside metropolitan areas and in the South and Midwest experienced income gains; metropolitan households and those in the Northeast and West suffered declines. In absolute numbers, however, the South continues to lag well behind the Northeast, while at the extremes of the distribution, median household income in Alaska is almost twice that in West Virginia.

Because a “household” can consist of one person, often a young adult at the onset of the work life or an elderly retiree, household medians are typically lower than those for families, officially defined as two or more persons related by blood, marriage, or adoption. The advantages of family households, especially those comprising a married couple, over nonfamily households is shown in Table 1. Clearly, some of the other differences seen in Table 1—by race, ethnicity, region, and residence—can be accounted for variations in the proportion of married-couple family households.

**Per Capita Income.** Table 1 also reports on per person income, which has risen by close to 7 percent since 1989 in constant dollars, with black gains outstripping those for whites, although white per capita income remains roughly 40 percent higher than that of blacks. In addition, that part of per capita income which is accounted for by earnings continues to show a gender gap in wages, but one that has narrowed from a 31 percent advantage for male workers in 1989 to 25.8 percent in 1997. In part, this narrowing is due to a 3 percent increase in real earnings for women over the eight years, but also an even more significant decline of more than 4 percent for men during the same period. The loss in real earnings for men reflects the general decline in employment in high-paying unionized jobs in the blue-collar sector as well as corporate downsizing at the managerial level.

Although not shown in Table 1, per capita income also varies directly with years of schooling. Since 1980, the adjusted median wage for workers with only a high school education fell by 6 percent, while the median for college graduates rose by 12

## INCOME DISTRIBUTION IN THE UNITED STATES

## Summary Measure of Income by Selected Characteristics: 1989, 1996, and 1997

CHARACTERISTICS	1997	Median income (in 1997 dollars)		Percent change in real income 1996 to 1997	Percent change in real income 1989 to 1997
	Median income (dollars)	1996	1989		
<b>HOUSEHOLDS</b>					
All households	37,005	36,306	37,303	*1.9	-0.8
<b>Type of Household</b>					
Family households	45,347	44,071	44,647	*2.9	*1.6
Married-couple families	51,681	51,002	49,925	*1.3	*3.5
Female householder, no husband present	23,040	22,059	22,315	*4.4	3.3
Male householder, no wife present	36,634	36,476	39,108	0.4	*-6.3
Nonfamily households	21,705	21,454	22,221	1.2	*-2.3
Female householder	17,613	16,774	17,865	*5.0	-1.4
Male householder	27,592	27,892	29,036	-1.1	*-5.0
<b>Race and Hispanic origin of householder</b>					
White	38,972	38,014	39,241	*2.5	-0.7
White, not Hispanic	40,577	39,677	40,166	*2.3	1.0
Black	25,050	24,021	23,583	*4.3	*6.2
Asian and Pacific Islander	45,249	44,269	46,611	2.2	-2.9
Hispanic origin <sup>1</sup>	26,628	25,477	28,192	*4.5	*-5.5
<b>Age of Householder</b>					
15 to 24 years	22,583	21,930	24,027	3.0	*-6.0
25 to 34 years	38,174	36,711	38,442	*4.0	-0.7
35 to 44 years	46,359	45,439	48,554	*2.0	*-4.5
45 to 54 years	51,875	51,630	53,738	0.5	*-3.5
55 to 64 years	41,356	40,729	39,946	1.5	*3.5
65 years and over	20,761	19,894	20,402	*4.4	1.8
<b>EARNINGS OF FULL-TIME, YEAR-ROUND WORKERS</b>					
Male	33,674	32,882	35,179	*2.4	*-4.3
Female	24,973	24,254	24,237	*3.0	*3.0
<b>PER CAPITA INCOME</b>					
All races	19,241	18,552	17,999	*3.7	*6.9
White	20,425	19,621	19,088	*4.1	*7.0
White, not Hispanic	21,905	20,991	(NA)	*4.4	(X)
Black	12,351	12,172	11,231	1.5	*10.0
Asian and Pacific Islander	18,226	18,332	(NA)	-0.6	(X)
Hispanic origin <sup>1</sup>	10,773	10,279	10,605	*4.8	1.6

Table 1

SOURCE: U.S. Bureau of the Census (1998d): vii.

percent. This differential signifies the growing wage gap between unskilled and skilled workers in an economy increasingly geared to processing information rather than to manual labor. For example, according to the Bureau of Labor Statistics, in 1982 workers in the top decile earned roughly four times the wages of workers in the lowest decile; by 1996, this difference had spread to almost five times (Pierce 1998).

**Income Dispersion by Households.** In 1997, 11 percent of households had income below \$9,000, while 9.4 percent reported income of \$100,000 or more. Adjusted for inflation, these numbers indicate a decline of 9 percent in households at the lowest level and an increase of 21 percent in the proportion at the top over the decade since 1987. Also showing a slight increase were households with incomes in both the \$75,000–\$99,999 and \$10,000–\$14,999 brackets. In contrast, the proportion of households with incomes between \$15,000 and \$50,000 declined from 45.2 percent of the total to 44.5 percent. It would appear, then, that the “middle” continues to “disappear,” but that the overall distribution of household income moved upward, with the greatest gains among those in the highest income categories (U.S. Bureau of the Census 1998d, Table B-2).

**Income Dispersion by Families.** A similar pattern characterizes the distribution of income by families rather than households: a modest decline in the proportion with incomes below \$10,000, from 7.3 percent in 1987 to 6.8 percent in 1997, a 20 percent increase, from 9.3 to 11.8 percent, at the \$100,000+ level, and declines in the proportion in the \$25,000–74,999 range (U.S. Bureau of the Census 1998d, Table B-4). For both families and households at the lower end of the distribution, much of the gain in income over the decade can be attributed to the longer hours worked by both wives and husbands, as well as the tight labor market that allowed workers to bargain for higher wages (Mishel et al. 1998). Nonetheless, the major gains in family income accrued to the top 10 percent, who realized 85 percent of the increase in the value of stocks.

**Income Inequality.** Although the Census Bureau’s press releases in the late 1990s focused almost entirely on the overall gain in median household income since 1994, especially for black

and female-headed households, the data also showed a marked increase in inequality on the two measures tabulated by the Bureau.

1. *Shares of aggregate income received by households.* When income groups are divided into quintiles, the share of aggregate income received by the lowest one-fifth of households fell from a high of 4.4 percent in 1977 to a low of 3.6 percent in 1997, while the share going to the highest fifth rose from 43.6 to 49.4 percent. Declines over the past two decades also characterized the three middle quintiles. In other words, the top 20 percent of households accounted for an ever-higher share of the nation’s aggregate income, with that of the top 5 percent increasing from 16.1 to 21.7 percent.
2. *Index of income concentration.* Another way of measuring inequality is through an index of income concentration called the Gini coefficient or ratio. A Gini score of 0 would indicate complete equality, where there are no differences among units; a score of 1 indicates complete inequality, whereby one unit has all of whatever is being measured and the other units have none. Figure 1 tracks the percentage change in Gini ratios for household income since 1967. As you can see, the period from 1967 to 1980 was one in which income differences among households rose slowly, held in check largely by increases in government transfers to the poor and a relatively progressive tax structure. Inequality rose more sharply between 1978 and 1989, stabilized briefly during the recession of 1990–1992, and then shot upward in the mid- and late 1990s. Some of this increase can be accounted for by a change in Census methodology, but most is very real, and there was no indication that the trend had abated in 1998.

As striking as are the data on household income inequality shown in Figure 1, the story for *family* income is even more dramatic. After declining during the 1960s, the percent change in Gini coefficients for family income has risen sharply higher than for households, reaching the 20–25

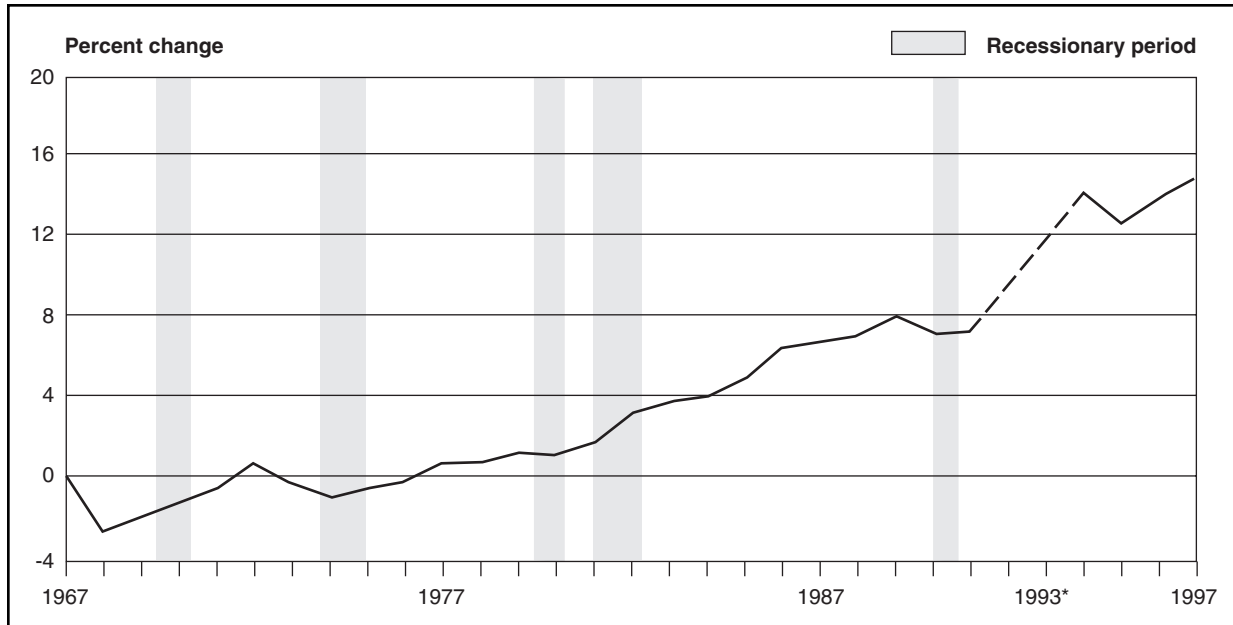


Figure 1. Percent change in Household Gini Coefficients: 1967–1997

SOURCE: U.S. Bureau of the Census (1998d): xiii.

NOTE: \*Computer-assisted personal interviewing (CAPI) was introduced in January 1994. As part of the conversion, increases were made in the limits for some income sources. This change in methodology increased measured income in 1993 for the highest income households by considerably more than their actual incomes rose. See Current Population Reports, Series P60–191, “A Brief Look at Postwar U.S. Income Inequality.”

percent range by the mid-1990s (U.S. Bureau of the Census 1996, 1998a).

**Other Benefits.** In addition to money income and government transfers, the well-being of households and families is also affected by employee benefits such as health insurance coverage, vacation pay, sick leave, and pensions. Although it was once thought that fringe benefits would tend to reduce inequality within the labor force, since such compensation represents a larger proportion of lower-wage workers’ income than of high-income earners, quite the opposite appears to be the case.

While government-sponsored programs such as Social Security and workers’ compensation insurance do indeed have an equalizing effect throughout the labor force, employer-sponsored benefits, in contrast, have “substantially increased compensation dispersion,” especially at the lower end of the wage scale, according to data from the Bureau of Labor Statistics (Pierce 1998). This is so because less skilled workers are most likely to be employed

in jobs that do not provide benefits, particularly in the small businesses providing most of the new jobs and in workplaces where employers can find a number of ways to limit coverage and/or exclude dependents. Benefits can be denied to entry-level employees and to those classified as “contingent.” This latter category—covering all types of part-time employment—also accounts for growing numbers of white-collar workers without employer-sponsored health insurance and pension plans.

As Table 2 indicates, even the proportion of workers at the top of the wage scale with health insurance coverage has declined since 1982, while that of low-wage employees has shrunk from half to one-fourth.

**Taxes and Transfers.** Thus, although government transfers, most notably Social Security and the earned income credit, have had a moderating effect on income dispersion, the impact of the tax system has been less so, as rates have become less progressive than in the 1960s. At the federal level, the Taxpayer Relief Act of 1997 might benefit

**Percentage of Employees with Employer-Provided Health Insurance**

	1982	1996
Bottom 10% of wage earners	49	26
Middle 50% of wage earners	90	84
Top 10% of wage earners	98	90

*Table 2*

SOURCE: Pierce (1998); see also U.S. Bureau of the Census (1998b).

lower-income taxpayers through a \$500 per child tax credit, but the major new programs—education and retirement savings accounts—will lighten the tax burden only of those who can afford to put money aside for the future. The big winners are the top 1 percent of wealthholders, who will receive, on average, more than \$7,300 in tax relief from a cut in the capital gains rate, compared with savings of about \$7 for the lower 60 percent of taxpayers (Congressional Budget Office 1998).

In addition, the FICA payroll tax, which takes a bigger bite out of the incomes of a majority of American families than does the federal income tax, is regressive in that most workers will pay the tax on 100 percent of income, while higher-income earners will be taxed on only the first \$72,600 of wages. Also regressive in their impact are the types of levies currently favored by state governments—sales and sin taxes. As a consequence, the current tax structure has done more to reinforce than to moderate the trend toward income inequality in the United States.

**Explanatory Variables.** A number of other secular trends have been advanced to explain the long-term increase in income inequality in family and household income.

1. *Labor-market factors.* One set of explanatory variables focuses on changes in the economy: (a) the shift in employment from manufacturing to service-related jobs, and the further division among service jobs between high- and low-skilled; (b) a concomitant decline in organized labor and the power of unions to negotiate favorable wage and benefit packages; (c) global competition and immigration patterns that depress wages of low-skill workers; and (d)

growth of the contingent labor force, such as temporary, part-time, and contract workers, who are typically ineligible for fringe benefits.

All these trends have contributed to an extremely skewed wage structure in which a few at the top crowd out the rest of the field, a “winner take all” situation that characterizes all occupations but is most notable in professional sports and in the compensation package of chief executive officers (Frank and Cook 1995). With the exception of a handful of millionaire athletes, almost all the benefits of economic growth and changes in tax system since the mid-1970s have been reaped by a small stratum of Americans already enriched by education, opportunity, and social capital.

2. *Lifestyle factors.* A second set of variables concerns long-term changes in patterns of marriage and living arrangements: (a) an increase in nonfamily and single-parent households due to later age at first marriage, high rates of divorce and separation, nonmarital births, and increased life expectancy, especially for widows in single-person households; and (b) the tendency toward endogamy among high-earning men and women, thus concentrating incomes and widening the split between the few at the top and the rest of the population.

Interestingly, the widening gap between top and bottom in income shares is not fully reflected in data on the distribution of wealth.

**WEALTH**

The methodological difficulties in measuring income are minor compared to those encountered in the study of *net worth*: the total value of all assets owned by a household, family, or person, less the debt owed by that unit. Such assets include investment portfolios, bank accounts, trusts, businesses, real estate, homes and their furnishings, insurance policies, annuities, pension equity, vehicles, works of art, jewelry, and other contents of safe deposit boxes. Because public records of such assets are minimal and/or difficult to trace, researchers are largely dependent on self-reports. In addition,

because very few extremely wealthy units would appear in a national random sample, special frames must be constructed. As a result, data reported by the Bureau of the Census, the Federal Reserve Board, and the Internal Revenue Service are not always comparable because of differences in sampling, the type of assets being counted, and the way in which they are measured and weighted. In this section, we will briefly review the history of studying wealth in America, current data, and the generalizations that can be most confidently drawn.

The first systematic study of wealthholding was conducted in 1963 by the Federal Reserve Board (FRB), which found that the wealthiest one-half of 1 percent of households (“superrich”) owned 25 percent of the aggregate net assets of the nation. The next one-half of 1 percent (“very rich”) accounted for an additional 7 percent of net worth, and the next 9 percent (plain “rich”), accounted for one-third of the total, leaving 35 percent of the total net worth in the hands of the remaining 90 percent of households.

Although comparable data were not collected until 1983, evidence from the Internal Revenue Service (IRS) estate tax records suggests that the share of assets owned by the superrich declined between 1965 and 1976 to a low of 14.4 percent. This drop was due in part to an extended stock market slump and in part to changes in tax policy, as well as the growth of social welfare programs such as Aid to Families with Dependent Children, Medicare and Medicaid, and the liberalization of Social Security benefits, all of which shifted resources from the affluent to the more needy.

Interest in research on wealth picked up again in the mid-1980s, when both the Bureau of the Census and the FRB conducted studies designed to yield data comparable to the 1963 survey, although with different sampling frames and asset measures. The common finding, however, was that between 1976 and 1983 the downward trend of asset ownership by the superrich was dramatically reversed: from owning less than 15 percent of aggregate wealth to accounting for more than 30 percent just six years later. This doubling of asset ownership reflected an upward swing in the value of stocks, reinforced by Reagan administration policies on taxes and welfare favoring the more affluent. It must be noted, however, that these

numbers are subject to considerable error due to sampling, nonresponse, and missing data.

**Recent Studies of Household Wealth.** In a major effort to standardize research findings, the FRB adopted a consistent weighting formula for adjusting data on household wealth from the Board’s Survey of Consumer Finances for 1989, 1992, and 1995 (Kennickell and Woodburn 1997). According to these calculations, shown in Table 3, the share of net worth held by the superrich remained constant between 1989 and 1992 at about 23 percent, then rose to 27.5 percent between 1992 and 1995. At the other end of the distribution, the share owned by the 90 percent less well-off households also remained stable—at about 32 percent—over the entire 1989–1995 period. The 1992–1995 increase in the share of net worth of the top one-half of one percent, then, has come largely from the share of the 90–99 percentile “plain rich,” which declined from 37 to 33 percent. In other words, there is little evidence that the rich have become richer at the expense of nonaffluent households.

There are marked contrasts between the few at the top and the rest of American households in the types of assets held. Wealth for the bottom 90 percent consists primarily of a principal residence, vehicles, and cash-value life insurance, and has been considerably diluted by rising debt (mortgage and credit card) in the 1990s. In contrast, stocks and bonds, trusts, and equity in businesses account for the bulk of the accumulated wealth of the top decile. By 1997, when stock ownership replaced the value of real estate as the leading component of aggregate wealth, the top 1 percent held more than half of all such investment instruments, with the bottom 80 percent holding 3 percent of the total value (Wyatt 1998). Yet even at this relatively low level of stock ownership, primarily through pension plans, low-wealth households will be especially vulnerable to sudden downturns in the market value of their investments.

Another set of numbers comes from the Bureau of the Census, which has, since 1990, conducted a panel study of households—the Survey of Income and Program Participation (SIPP)—that permits following the economic status of units over time. These data differ from those collected by the FRB in three crucial respects: the sample

**Percentage of Total Net Worth Held by  
Different Percentiles: 1989, 1992, and 1995**

YEAR	Percentile of Net Worth			
	0-89.9	90-99	99-99.5	99.5-100
1989	32.5	37.1	7.3	23.0
1992	31.9	36.9	7.5	22.7
1995	31.5	33.2	7.6	27.5

Table 3

SOURCE: Kennickell and Woodburn (1997): 22.

frame does not yield many very high income households; a different set of assets are counted; and the distribution of wealth is measured differently. In the SIPP study, the distribution of “asset ownership” is computed on the basis of the median net worth of households at each quintile of monthly income. For both 1991 and 1993, the one-fifth of households with the lowest monthly income owned about 7 percent of the total net worth of all households, while those in the top fifth owned 44 percent (U.S. Bureau of the Census 1995). Unfortunately, comparable data for the next wave (1995-1996) have not yet been published, nor can the SIPP numbers be used for historical comparisons.

**Personal Wealth.** In addition to studies of household wealth by the Census and Federal Reserve Board, the Internal Revenue Service periodically publishes reports on individual wealth-holding, based on estate tax returns. These data are also subject to error—from sampling, from calculating mortality rates, and from underreporting, since not all assets can be tracked and since high-income earners have ways of dispersing and hiding assets prior to death. The most recent data come from surveys of estate tax returns of the very wealthy carried out by the IRS’s Statistics of Income Division (SID) in 1992 and 1995 (Johnson 1997).

In 1992, it was estimated that 3.7 million adults, or 2 percent of Americans aged 21 or older, had gross assets of at least \$600,000, which accounted for 28 percent of the aggregate personal wealth of the nation. In 1995, the number of wealthy persons had increased (to 4.1 million), as had their total net worth, but once adjusted for inflation, these differences were minimal. Indeed, looking at the very wealthiest—persons with a net

worth of at least \$1 million—in terms of numbers, total assets, and net worth, there was a sharp decline between 1989 and 1992, due primarily to the recession of 1990-1992. Between 1992 and 1995, however, the number of millionaires increased slightly (to 1.32 million), as did their total assets and net worth, but both still remained below the levels of 1989.

The SID also computed the percent of total U.S. net worth held by the top 1 percent and one-half of 1 percent of individual wealthholders. As shown in Figure 2, the share of total wealth owned by the wealthiest individuals rose from 1989 to 1992 before declining to 1989 levels. Clearly, according to these data, there was no dramatic shift in net worth from the less affluent to the top, although the dollar amount of their assets did appreciate. Between 1989 and 1995, then, the pattern for individual wealthholders was similar to that for household net worth in showing minimal increases in the concentration of wealth.

At the time of publication, detailed reports from the IRS and FRB on trends between 1995 and the end of the decade were unavailable, but material from *Forbes* magazine’s yearly compilation of the 400 wealthiest individuals suggests that the fortunes of the very affluent have increased significantly (*Forbes*, October 12, 1998). In 1998, the minimum needed to appear on the list of the 400 wealthiest individuals in the United States was \$500 million, up 5 percent from 1997 and more than double that of a decade earlier. Most dramatic was the increase in the proportion with a net worth of more than \$1 billion, up from 170 persons in 1997 to 186 in 1998, almost half the total, compared to only 23 in 1986. Ten of the billionaires had fortunes in excess of \$10 billion, including the five children of the founder of Walmart Stores, and Bill Gates, the founder of Microsoft, whose net worth almost equals the gross domestic product of New Zealand. Thus, while the very rich may not be getting richer at the expense of the less affluent, they received the lion’s share of wealth created in the economic boom years of the mid-1990s.

## POVERTY

Interesting changes are also taking place at the lower end of the income distribution, among the persons, families, and households officially designated as living below the poverty threshold. In the



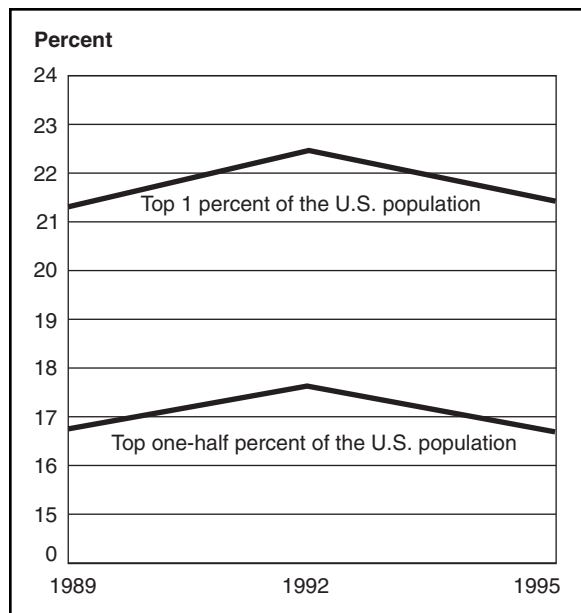


Figure 2. Percent of Total U.S. Net Worth Held by the Top 1 Percent and Top 1/2 Percent of the U.S. Population

SOURCE: Johnson (1997: 79).

early 1930s, Franklin Delano Roosevelt could speak movingly of one-third of the nation being ill-housed, ill-clad, and ill-nourished in a society without an extensive social welfare system. Yet three decades later, in the early 1960s, despite the introduction of Social Security and other programs designed to minimize the effects of unemployment, more than 20 percent of Americans could still be defined as lacking a minimally adequate standard of living.

**Defining Poverty.** By 1964, as the nation geared up for Lyndon Johnson's War on Poverty, the Social Security Administration (SSA) was pressed to construct clear parameters for measuring impoverishment. The SSA turned to research conducted by the Department of Agriculture (DoA) in 1955, which found that families typically spent one-third of their income on food and which also computed the cost of a least expensive nutritionally adequate food plan. The SSA simply multiplied the cost of the DoA's food basket by three and, with some corrections for family size, age and sex of householder, and rural/urban residence, arrived at a dollar figure for yearly income—that neatly demarcated the poor from the nonpoor.

The value of the minimal food plan is adjusted each year to the cost-of-living index, but it no longer distinguishes rural from urban residence (the rural threshold had been higher, since it was thought that country folk could grow some of their own food) or female from male heads of household (on the grounds that women ate less than men). Only number of children and age of householder have been retained in the calculations. Nor has there been an adjustment for the fact that the cost of housing today typically exceeds that for food.

By the official yardstick, the poverty threshold for a single person in 1997 was \$8,183, slightly higher for those under age 64 and slightly lower for someone aged 65 or older on the assumption that an older person eats precisely \$276 worth less food per year than does a younger person. The poverty line was \$12,802 for a family of three and \$16,400 for a family of four. These are the dollar amounts considered adequate to house, feed, and clothe household members. Income in excess of the threshold officially lifts the unit out of poverty and therefore makes the unit ineligible for additional benefits, including both income supports (the former Aid to Families with Dependent Children and its successors, and Supplemental Security Income) and in-kind programs (food stamps, Medicaid, school lunch assistance, and housing subsidies). At the urging of conservative critics, the Bureau of the Census also publishes computations that include the cash value of these in-kind benefits in the definition of income, thus automatically reducing the poverty rate by about 25 percent. Nonetheless, at this writing, the official poverty level is still calculated on the basis of money income earned or received.

As Figure 3 indicates, the number of people below the poverty level was cut in half between 1959 and 1973—from 22.4 to 11.1 percent of the population—as a result of federal programs designed to assist the elderly, low-income families, and single-parent households. As the domestic War on Poverty fell hostage to the war in Vietnam, poverty rates began to rise, reaching a high of 15.2 percent during the recession of the early 1980s, and again in 1993, before declining to the current 13.3 percent, or 35.5 million persons.

In its 1998 report, the Census Bureau drew special attention to the fact that poverty declines in 1995–1997 have been much steeper for black



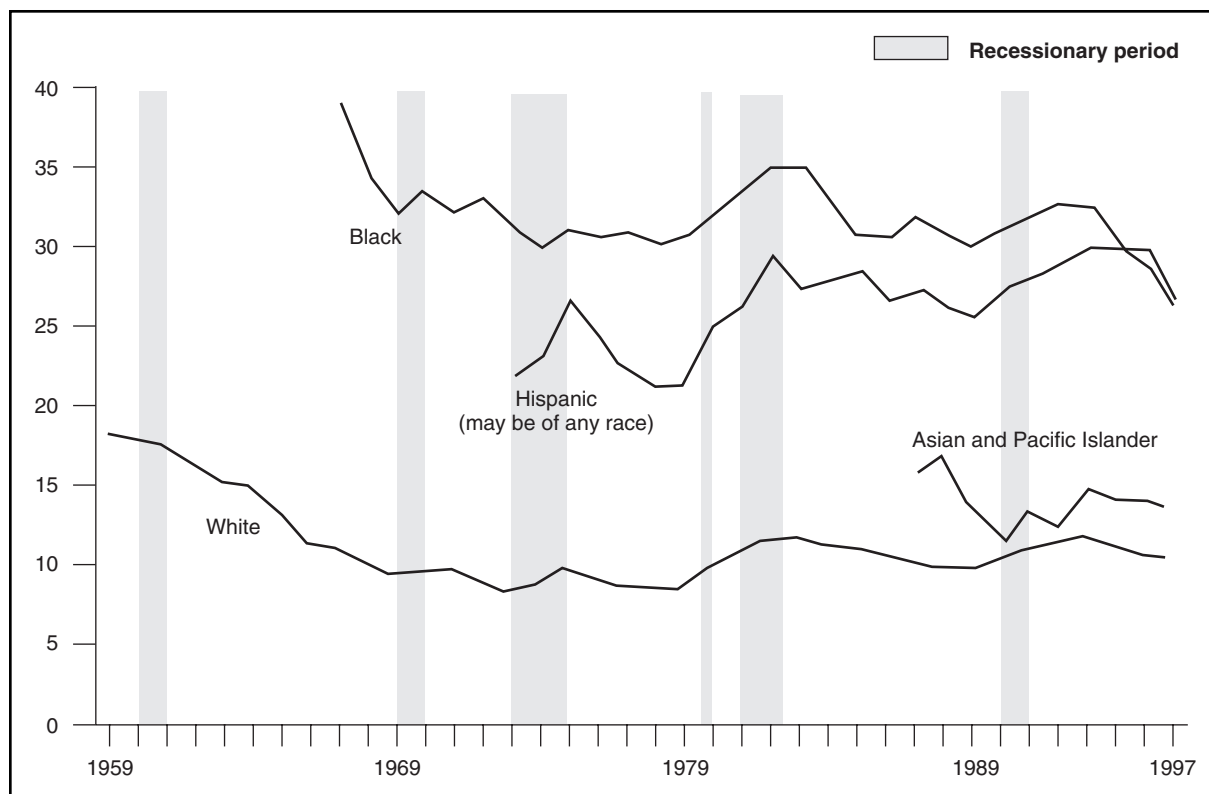


Figure 3. Percent of Persons Below the Poverty Level, by Race and Ethnicity

SOURCE: U.S. Bureau of the Census (1998):45.

and Hispanic persons than for Asian/Pacific Islanders and whites, but this effect is partly due to the fact that blacks and Hispanics were much further behind to begin with, so that any income increase will translate into a higher percentage compared to those initially less disadvantaged. Nonetheless, poverty among African Americans is at an all-time low, though the rate remains more than twice that for whites. The three factors most responsible for declining poverty rates today are: (1) The rise in the minimum wage that took effect in 1996, which accounts for the increased incomes of black and Hispanic single mothers, most of whom are hourly workers; (2) the earned income tax credit (EITC) now available to low-income workers, which has been particularly helpful to low-wage married-couple families; and (3) the economic boom that has generated a large number of jobs at the lower-skill level of the service sector, thus reducing unemployment and allowing low-wage workers to bargain for higher wages and benefits. In 1998, however, Congress voted down

further raises in the minimum wage and failed to expand the EITC, thus leaving low-skill workers increasingly dependent on market forces.

The Census also computes the “ratio of income to poverty level”; that is, the number of Americans in families whose income is under half the poverty level, the “severely poor,” as well as those with incomes 25 percent above the threshold, the “near poor.” In 1997, 14.6 million persons, or 41 percent of the poor, were “severely poor.” Another 12.3 million were “near poor,” including many full-time workers, since the earnings of a year-round, full-time minimum-wage-worker would still fall below the poverty threshold for a couple with one child.

Who are America’s poor? Forty percent are children under age 18, whose poverty rate of about 20 percent is unchanged since 1989. Especially disadvantaged are the 59 percent of children under age six living with a female householder, no husband present. And while the poverty rate for all

such households has hovered at slightly under one-third since 1987, there was a significant decline in poverty from 1996 to 1997 for black female householders. Even so, four in ten black female householders and almost half of Hispanic female householders were officially poor in 1997.

The most powerful variable affecting poverty status, however, is not race or ethnicity, but sex. It is women of all ages, especially those without husbands, from teenage mothers to elderly widows, who are most at risk of being poor. If economic well-being depends on working full time, remaining married, and being free of child-care responsibilities, then single mothers with limited education and job skills will be especially disadvantaged (U.S. Bureau of the Census 1998e). Yet it is precisely these women, many of whom were dependent for survival upon the income provided by Aid to Families with Dependent Children (AFDC), who came to be blamed for their own misfortune and to symbolize the failure of public welfare programs.

**The Poverty Debate of the Mid-1990s.** Although poverty programs absorbed less than 5 percent of the national and state budgets, they became a focal point for political debate in the mid-1990s. The public had come to believe, contrary to empirical studies, that AFDC families remained on the welfare rolls for generations, that poor women had additional children in order to increase their monthly benefit, and that it was characteristics of the poor themselves (laziness, substance abuse, sexual immorality) that accounted for poverty. In fact, fewer than 10 percent of the poor can be considered long-term welfare recipients, primarily women who entered the system as very young unwed mothers and who have been unable to develop the job skills or to find employment near their home that pays enough to lift the family above the poverty level. Nor is there any consistent relationship between benefit levels and the fertility of poor women.

As the SIPP data clearly show, poverty is a transitory state for the great majority who fall below the threshold. For example, although 21.4 percent of Americans were poor at some point in 1994, the proportion who were poor for all of 1993 and 1994 was only 5.3 percent. Almost half of all spells in poverty (47.3 percent) lasted 2–4 months, and 75 percent lasted less than one year (U.S.

Bureau of the Census 1988f). People fall into poverty when a marriage ends, when employment stops, or when children become ill; they are lifted out of poverty when they remarry, when they are employed, and when family members are restored to health.

As part of their attack on all government programs to which entire classes of citizens are entitled, conservative critics of the American social welfare system have long maintained that there is a “culture of poverty,” whereby maladaptive attitudes toward work and family are transmitted across generations (e.g. Murray 1984). In this view, welfare dependency itself is the problem, and the solution is to remove income supports, so that employment becomes the more attractive alternative. In contrast, most sociologists would focus on the structural conditions that produce unemployment, family dissolution, educational failure, and homelessness, with people’s behavior perceived more as a response to than as a cause of their situation.

The debate was won by those seeking radical change in the welfare system, and in 1996 the Personal Responsibility and Work Opportunity Act replaced AFDC and several other federal assistance programs with block grants to the states for Temporary Assistance for Needy Families (TANF). The states were given relative freedom to construct their own programs for moving former welfare recipients into the paid labor force.

It is too early to tell what effect the new systems will have on poor women and their children. Some states will be more effective than others in providing the job training and child care required for full-time labor-force participation; others will be more or less punitive. It will also be very difficult to disentangle the effects of the business cycle from those of public policy, or to determine whether people leaving welfare would have done so anyway given the transitory nature of most poverty spells. At the moment, welfare rolls have dropped dramatically, as a function of both economic growth and tighter eligibility requirements. And although the poverty rate also fell in 1997, the decline was minimal. Much will depend on whether the private sector can continue to generate jobs that former recipients can find and hold—and that will pay a family wage.

## COMPARATIVE PERSPECTIVES

Any discussion of wealth, poverty, and income inequality within a society must take into account the vast differences between modern industrial and the less developed nations. Poor people in high-income societies are rarely as deprived of the basic necessities of survival as are most people in the Third World, where income inequality is typically much greater than in the industrialized countries. But Americans do not compare themselves to Sudanese; rather, they compare themselves to other Americans to whom they feel similar. It is a sense of relative rather than absolute deprivation that tends to fuel resentment. At this writing, despite the extent of income inequality in the United States, there are few appreciable signs of anger directed toward the top wealthholders. Rather, whatever ill will has been generated by blocked opportunity appears to be directed toward racial and ethnic minorities at the same or lower social-class level.

Yet, when compared with other Western democracies, the United States has the most unequal distribution of income and highest poverty rate. For example, while the percentage of American children in poverty was among the highest in the developed world, proportionately fewer were lifted out of poverty by government aid (Atkinson et al. 1995; United Nations 1998). This is so because the United States has the least extensive social welfare system of any modern state, and since 1996 even this limited “safety net” has been reduced. Alone among its industrialized peers, the United States is without a comprehensive family policy, lacks a national health insurance system, and provides minimal assistance to the most needy. As a consequence, there are few institutionalized mechanisms other than Social Security—which has also come under attack from those who would turn it into a private rather than public responsibility—for the redistribution of income that might narrow the gap between the very rich and very poor or that might substantially reduce both the likelihood and impact of poverty. In the absence of a revitalization of a sense of collective responsibility, income inequality will continue to characterize the United States. Indeed, all signs point to a continuation of economic, social, and political trends that elevate individual over collective interests and that reinforce the power of the market and weaken that of government.

## REFERENCES

- Atkinson, Anthony B., Lee Rainwater, and Timothy M. Smeeding 1995 *Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study*. Washington, D.C.: OECD (Organization for Economic Cooperation and Development) Publications and Information Center.
- Congressional Budget Office 1988 Data provided to Citizens for Tax Justice, Washington, D.C.
- Frank, Robert H., and Philip J. Cook 1995 *The Winner Take All Society: How More and More Americans Compete for Ever Fewer and Bigger Prizes, Encouraging Economic Waste, Income Inequality, and an Impoverished Cultural Life*. New York: Free Press.
- Johnson, Barry W. 1997 “Personal Wealth, 1992–1995.” *Internal Revenue Service Statistics of Income Bulletin* 16(3):70–95.
- Kennickell, Arthur B., and R. Louise Woodburn 1997 *Consistent Weight Design for the 1989, 1992, and 1995 SCFs, and the Distribution of Wealth*, Revision II. Washington, D.C.: Federal Reserve Board.
- Mishel, Lawrence, Jared Berstein, and John Schmitt 1998 *The State of Working America*. Washington, D.C.: Economic Policy Institute.
- Murray, Charles 1984 *Losing Ground*. New York: Basic.
- Pierce, Brooks 1998 *Compensation Inequality*. Washington, D.C.: Bureau of Labor Statistics.
- United Nations 1998 *Human Development Report 1998*. New York: United Nations.
- U.S. Bureau of the Census 1995 *Asset Ownership of Households: 1993*. Current Population Reports, Series P70, No. 47. Washington, D.C.: U.S. Government Printing Office.
- 1996 *A Brief Look at Postwar U.S. Income Inequality*. Current Population Reports, Series P60, No. 191. Washington, D.C.: U.S. Government Printing Office.
- 1998a *Changes in Median Household Income: 1969 to 1996*. Current Population Reports, Special Studies, Series P23, No. 196. Washington, D.C.: U.S. Government Printing Office.
- 1998b *Health Insurance Coverage: 1997*. Current Population Reports, Series P60, No. 202. Washington, D.C.: U.S. Government Printing Office.
- 1998c *Measuring 50 Years of Economic Change Using the March Current Population Survey*. Current Population Reports, Series P60, No. 203. Washington, D.C.: U.S. Government Printing Office.
- 1998d *Money Income in the United States: 1997*. Current Population Reports, Series P60, No. 200. Washington, D.C.: U.S. Government Printing Office.

- 1998e *Moving Up and Down the Income Ladder*. Current Population Reports, Series P70, No. 65. Washington, D.C.: U.S. Government Printing Office.
- 1998f *Trap Door? Revolving Door? Or Both?* Current Population Reports, Series P70, No. 63. Washington, D.C.: U.S. Government Printing Office.
- Wyatt, Edward 1988 "Share of Wealth in Stock Holdings Hits 50-year High." *New York Times* January 11, p. 1ff.

BETH B. HESS

## INDEX

*See* Measurement.

## INDIAN SOCIOLOGY

The reviewers of Indian sociology generally trace its origin to the works of several British civil servants, missionaries, and Western scholars during the eighteenth and nineteenth centuries (Dhanagere 1985; Mukherjee 1979; Rao 1978; Singh 1986; Srinivas and Panini 1973). British administrators wanted to understand the customs, manners and institutions of the people of India to ensure the smooth running of their administration. Christian missionaries were interested in learning local languages, folklore, and culture to carry out their activities. The origin, development, and functioning of the various customs and traditions, the Hindu systems of caste and joint family, and the economy and polity of the village/tribal community were some of the prominent themes of study by the British administrators and missionaries as well as other British, European, and Indian intellectuals. The first all-India census was conducted in 1871. Several ethnographic surveys, monographs, census documents, and gazetteers produced during this period constitute a wealth of information that is of interest to sociologists even today. Mukherjee (1979) observes that the works of the civil servants, missionaries, and others during the colonial rule in India "provided the elements from which the British policy for ruling the subcontinent crystallized and also in turn helped to produce the pioneers in Indian sociology" (p. 24). Further, the available studies of Indian society and

culture became an important source for testing various theories by scholars such as Marx and Engles, Maine, and Weber.

Although the first universities in India were established in 1857 in Bombay, Calcutta and Madras, formal teaching of sociology began only in the second decade of the twentieth century—at the University of Bombay in 1914, at Calcutta University in 1917 and at Lucknow University in 1921. By this time the observations and impressions of the census commissioners and civil service officials about the caste system, family structure and functions, age at marriage and widowhood, and so forth, had already begun to be published, along with statistical tables based on the population counts. Prior to India's independence in 1947, only three other universities (Mysore, Osmania, and Poona) were teaching sociology. There was no separate department of sociology; it was joined with the department of economics (Bombay and Lucknow), economics and political science (Calcutta), anthropology (Poona), or philosophy (Mysore). Almost all the pioneers in sociology in the first half of the twentieth century were trained in disciplines other than sociology. Only a limited number of courses in sociology, as fashioned by teachers according to their interest, were taught. Sociology courses included such topics as social biology, social problems (such as crime, prostitution, and beggary), social psychology, civilization, and prehistory. "In the case of teaching of Indian social institutions the orientation showed more Indological and philosophical emphasis on the one hand and a concern for the social pathological problems and ethnological description on the other. Strong scientific empirical traditions had not emerged before Independence" (Rao 1978, pp. 2-3).

Although many of the pioneers in sociology were educated at Calcutta, substantial impact on Indian sociology during the first half of the twentieth century was made at Bombay University and Lucknow University. Patrick Geddes, the first chairperson of the department of sociology and civics at Bombay University, was a city planner and human geographer. His reports on the city planning of Calcutta, Indore, and the temple cities of south India contain much useful information and demonstrate his keen awareness of the problems of urban disorganization and renewal (Srinivas and Panini 1973, p. 187). G. S. Ghurye succeeded

Geddes in 1924. During his thirty-five year teaching career at Bombay University, he guided about eighty research students. Several of his students (e.g., M. N. Srinivas, K. M. Kapadia, I. P. Desai, Y. B. Damle, A. R. Desai, and M. S. A. Rao) later on had a great impact on the development of sociology in India. Trained as a social anthropologist at Cambridge University, he addressed a wide range of themes in his research work and writings: from castes, races, and tribes in India to cities and civilization, from Shakespeare on conscience and justice to Rajput architecture, and from Indian Sadhus to sex habits of a sample of middle-class people of Bombay. He drew attention to several unexplored dimensions of Indian society, culture, and social institutions.

R. K. Mukherjee and D. P. Mukherji, both trained in economics at Calcutta University, taught sociology at Lucknow University. R. K. Mukherjee made a series of micro-level analyses of problems concerning rural economy, land, population, and the working class in India as well as the deteriorating agrarian relations and conditions of the peasantry, intercaste tensions, and urbanization. D.P.Mukherji's interest was diverse; they ranged from music and fine arts as peculiar creations of the Indian culture to the Indian tradition in relation to modernity. A professed Marxist, he attempted a dialectical interpretation of the encounter between the Indian tradition and modernity, which unleashed many forces of cultural contradiction during the colonial era (Dhanagare 1985).

B. N. Seal and B.K. Sarkar were two of the leading sociologists of that time at Calcutta University. Seal was a philosopher, interested in comparative sociology. He wrote on the origin of race, positive sciences, and the physicochemical theories of the ancient Hindus, as well as made a comparative analysis of Vishnavism and Christianity. He stressed the need for a statistical approach, inductive logic, and methodology to appraise the contextual reality comprehensively. Sarkar, a historian and economist by training, opposed the persistent general belief that Hinduism is "other worldly." He was also one of the few who discussed Marx, Weber, and Pareto at a time when they were not fashionable with sociologists either in India or abroad (Mukherjee 1979). S. V. Ketkar and B. N. Dutt, both of whom specialized in Indological studies in United States, and K.P.

Chattopadhyay, a social anthropologist trained in the United Kingdom, are some of the other noteworthy pioneers of Indian sociology.

Mukherjee (1979) points out that the goals set by the pioneers ranged from an idealized version of Oriental culture to the materialist view of social development. They were involved in bibliographical research to establish the historical database and strongly advocated empirical research. But in the days of the pioneers, the interaction between theoretical formulations and the database remained at a preliminary stage.

Some of the outstanding trends in the development of Indian sociology since India's independence have been the organization of professional bodies of sociologists, a lack of rigid distinction between sociology and social anthropology, debates regarding the need for indigenization of sociology and the relevance of Indian sociology, diversification and specialization into various subfields, and participation of sociologists in interdisciplinary research.

There has been a tremendous increase in the number of universities, colleges, and institutes teaching sociology. In most universities, the teaching of sociology started first at the graduate level and then at the undergraduate level. Universities and institutions offering degrees in interdisciplinary areas—such as management, rural development, planning, communication, and nursing—include some sociology courses in their training program. In the 1990s, some states have also introduced sociology courses at the higher secondary level. There was no professional body of sociologists during the colonial period. Ghurye was instrumental in establishing the Indian Sociological Society in 1951, and R. N. Saksena was instrumental in organizing the first All-India Sociological Conference in 1956. These organizations merged in 1967 as a single all-India professional body of sociologists. The Indian Sociological Society has more than a thousand life members. Several regional associations of sociologists was also formed during the 1980s and 1990s.

The development of sociology in India, from the viewpoint of theory, methodology, and research interests, has been significantly influenced by sociology in Western countries. Several Western scholars, a majority of them initially from the

United Kingdom and Europe, and later on also from the United States, have carried out studies in India. Similarly, many of the leading sociologists in India have been trained in the United Kingdom and the United States. However, with a rise in the cost of higher education and a fall in the availability of financial assistance in the Western universities, there has been of late a decline in the number of Indian students going to the United Kingdom and the United States for advanced study in sociology. There has, however, been a steady increase in the participation of Indian sociologists in various international seminars, workshops, and conferences. The holding of the eleventh World Congress of Sociology in New Delhi in 1986 is indicative of the recognition of the development of Indian sociology and its contribution.

Two journals of sociology—*The Indian Journal of Sociology*, started in 1921 by Alban G. Widgery (a British professor at Baroda College), and *The Indian Sociological Review*, started in 1934 with R. K. Mukherjee as its editor—were short-lived. There are now only a few all-India journals of sociology: *Sociological Bulletin* (a biannual journal of the Indian Sociological Society since 1952), *Contributions to Indian Sociology* (edited by two French scholars, Louis Dumont and D. F. Pocock, from its inception in 1957 to 1963, when its editorship was passed on to Indian sociologists), and *Social Change* (published by the Council for Social Development since 1971). Occasionally, articles with sociological content and relevance are published in other journals, such as *Economic and Political Weekly* and the journals published by some universities and regional associations. Several sociological articles are published in the journals of some institutions and university departments with a focus on interdisciplinary training and research. For example, since the beginning of the 1980s, the National Institute of Rural Development in Hyderabad has published a quarterly (*Journal of Rural Development*), and the National Institute of Urban Affairs in New Delhi has published a biannual journal (*Urban India*). Also, the Center for Women's Development Studies in New Delhi publishes a biannual journal called *Indian Journal of Gender Studies*.

Initially, no rigid distinction was made between social anthropology and sociology, but they separated as teaching disciplines in the 1950s. In the field of research, however, the distinction

between social anthropology and sociology continues to be blurred. Ghurye, Srinivas, S. C. Dube, and Andre Beteille, among others, have argued that sociologists in the Indian context cannot afford to make any artificial distinction between the study of tribal and folk society on the one hand and advanced sections of the population on the other; nor can they confine themselves to any single set of techniques. Yogesh Atal (1985) points out that this is true of several countries in Asia and the Pacific; social anthropologists have extended the scope of their investigation to micro communities in rural as well as urban settings in their own country, and sociologists have found the anthropological method of fieldwork and participant observation useful in their research. Even the Indian Council of Social Science Research (ICSSR) treats both these disciplines together in its two surveys of research, the first covering the period upto 1969 (ICSSR 1972–1974) and the other from 1969 to 1979 (ICSSR 1985–1986); the same approach continues for the third survey, now under way, for the period 1979 to 1987.

There have been continuous debates regarding the need for indigenization of sociology, or “sociology for India” and the relevance of Indian sociology (Sharma 1985; Unnithan et al. 1967). One side of the debate started with the suggestion of Dumont and Pocock (1957) that “in principle, . . . a sociology of India lies at the point of confluence of Sociology and Indology” (p. 7). The proponents of an Indological approach in sociology emphasize that the contextual specificity of Indian social realities could be grasped better from the scriptural writings. Gupta (1974) points out the need for separating normative and actual behavior. Oommen (1983) pleads that “if sociology is to be relevant for India as a discipline it should endorse and its practitioners should internalize the value-package contained in the Indian Constitution” (p. 130), that is, socialism, secularism, and democracy rather than hierarchy, holism, pluralism, and so forth as pointed out by the Indologists. Another side of the debate is identified with a paradigm of Indian sociology free from academic colonialism—that is, borrowed packages of concepts and methods from other cultures, particularly the West, that supposedly do not have relevance for the Indian social, historical, and cultural situation (Singh 1986). However, although most sociologists are not hostile to using Western concepts,

models, and analytical categories, they want their adaptations to suit the Indian sociocultural setting. Singh (1986) analyzes the contents and salient orientations of the presidential addresses delivered by M. N. Srinivas, R. N. Saksena, Ramkrishna Mukherjee, S. C. Dube, A. R. Desai, and M. S. Gore at the conferences of the Indian Sociological Society. He observes that there is a deep concern with the issue of relevance in the contexts of social policies, normative analysis of these policies, and the role of sociologists in understanding, critical appraisal, and/or promotion of these normative objectives of development and change in India.

In the 1950s and 1960s, several micro-level studies of caste, joint families, and village communities, mostly from the viewpoint of structural-functional aspects and change, were carried out. Srinivas introduced the concepts of dominant caste, Sanskritization, Westernization, and secularization to understand the realities of intercaste relations and their dynamics (Dhanagare 1985). Change in the structural and functional aspects of family in different parts of India was the focal point of most studies in the area of marriage, family, and kinship. The village studies focused on stratification and mobility, factionalism and leadership, the *jajmani* (patron-client) relationship, contrasting characteristics of rural and urban communities, and linkages with the outside world.

Indian sociology in the last quarter of the twentieth century shows both continuity and change in research. Caste and stratification, village communities, and social change have continued to be themes of research, but the approach has shifted from the functional to the conflict viewpoint. Descriptive studies of a single village community or other unit in a single social setting have been replaced by analytical comparative studies of social structure across time and space. Interest in the area of marriage, family, and kinship has declined. Women's studies have increased greatly. Several studies have been conducted in the fields of education, urban sociology, social movements, voting behavior, communication, and industrial relations. Sociologies of medicine, law, science, and other professions have also begun to develop. Now the thrust is on studying various processes. For example, with a concern for equality and distributive justice, there is an increasing emphasis on examining the process of education as a vehicle of social

change as it affects the existing system of stratification, women, and less favored segments of the population.

A steady trend of out-migration of entrepreneurial and educated Indians, particularly to Western countries, has led to a modest beginning of sociological studies of the Indian diaspora (Motwani et al. 1993). Such studies attempt to understand the sociocultural dynamics of the Indian diaspora. Some of these studies are influenced primarily by the phenomenological and the symbolic interactionist perspectives (Jayaram, 1998). Interest in the study of changing patterns of marriage and family relations due to international migration, both among the out-migrants and among the aged and others who continue to reside in India, is slowly increasing. A genealogical study of an Asian Indian, covering nineteen generations, is a notable illustration of efforts to trace roots and study intergenerational socioeconomic changes (Desai 1997).

India started its first Five-Year Plan in 1952. Since then social scientists, particularly economists, sociologists, and demographers, have been involved in conducting diagnostic, monitoring, and evaluative studies concerning a variety of developmental programs at micro as well as macro levels. For example, the interest and financial support of the Indian government and international agencies have been instrumental, since the early 1950s, in encouraging and sponsoring research in the field of population and family planning (Visaria and Visaria 1995, 1996). Policies and programs concerning urban and rural community development, Panchayati Raj, education, abolition of untouchability, uplift of weaker sections (scheduled castes, scheduled tribes, and other backward castes), and rehabilitation of people affected by large-scale projects (constructions of large dams, industrial estates, capital cities, etc.) have been some of the other important areas of research by sociologists. At times, the various ministries of the central and the state governments, the ICSSR, and other funding agencies have sponsored all-India studies that have tended, albeit in a small way, to strengthen interdisciplinary approaches in social research. For example, in 1975-1976 the Indian Space Research Organization conducted a one-year satellite instructional television experiment in 2,330 villages spread over twenty districts of six states (Agrawal et al. 1977); the ICSSR sponsored a

nationwide study of the educational problems of students from scheduled castes and tribes (Shah 1982).

During the 1970s and 1980s, several social research institutes were established in different parts of India. Also, many universities established interdisciplinary women's studies. Most prominent sociology departments and/or social research institutes are located in Delhi, Bombay, Ahmedabad, Jaipur, Chandigarh, Poona, Bangalore, Hyderabad, and Trivandrum.

Several universities have gradually switched over to the use of the regional language as a medium of instruction at the undergraduate level, and some at the graduate level also; however, inadequate availability of textbooks in regional languages has been a major handicap in the teaching-learning process. Statistics has as yet not become an integral component of sociology curricula in a large number of colleges and universities. Although the survey approach is widely used in sociological research, most research publications hardly go beyond the use of descriptive statistics. There has been a strong plea for developing concepts and measurements that fit the Indian situation, but concerted efforts in this matter are still lacking. At the end of the twentieth-century, there is a frequently mentioned concern about an emergent paradoxical trend in sociological teaching and research in India. On the one hand, sociology has been accepted as one of the core subjects in almost all colleges and universities and several interdisciplinary institutions; on the other hand, a severe problem obstructing the growth and development of sociology in India due to a dearth of qualified teachers and falling standards, which may be due to a lack of teaching material in the regional languages and the inability of a large number of students to read textbooks and reference material in English. With a democratic polity, a developing economy, and a socioculturally diversified population, questions of applying even fairly good social science research findings get shrouded in the complex processes of the state policy and administration.

## REFERENCES

Agrawal, Binod C., J.K.Doshi, Victor Jesudason, and K.K.Verma 1977 *Satellite Instructional Television Experiment: Social Evaluation-Impact on Adults, Parts I-II*. Bangalore: Indian Space Research Organization.

- Atal, Yogesh 1985 "Growth Points in Asian and Pacific Sociology and Social Anthropology." In *Sociology and Social Anthropology in Asia and the Pacific*. Paris: UNESCO. New Delhi: Wiley Eastern.
- Desai, D. K. 1997 *A Genealogical Study of an Asian Indian*. Delhi: Himalaya Publishing House.
- Dhanagare, D. N. 1985 "India." In *Sociology and Social Anthropology in Asia and the Pacific*. Paris: UNESCO. New Delhi: Wiley Eastern.
- Dumont, Louis, and D. F. Pocock 1957 "For a Sociology of India." *Contributions to Indian Sociology* 1:7-22.
- Gupta, Krishna Prakash 1974 "Sociology of Indian Tradition and Tradition of Indian Sociology." *Sociological Bulletin* 23:14-43.
- Indian Council of Social Science Research (ICSSR) 1972-1974 *A Survey of Research in Sociology and Social Anthropology*, 3 vols. Bombay: Popular Prakashan.
- 1985-1986 *Survey of Research in Sociology and Social Anthropology 1969-1979*. 3 vols. New Delhi: Satvahan Publications.
- Jayaram, N. 1998 "Social Construction of the Other Indian: Encounters Between Indian Nationals and Diasporic Indians." *Journal of Social and Economic Development* 1:46-63.
- Motwani, Jagat K., Mahin Gosine, and Jyoti Barot-Motwani, eds. 1993 *Global Indian Diaspora: Yesterday, Today and Tomorrow*. New York: Global Organization of People of Indian Origin.
- Mukherjee, Ramkrishna 1979 *Sociology of Indian Sociology*. New Delhi: Allied Publishers.
- Oommen, T. K. 1983 "Sociology in India: A Plea for Contextualisation." *Sociological Bulletin* 32:111-36.
- Rao, M. S. A. 1978 "Introduction." In *Report on the Status of Teaching of Sociology and Social Anthropology, Part I: Recommendations*. New Delhi: University Grants Commission.
- Shah, Vimal P. 1982 *The Educational Problems of Scheduled Caste and Scheduled Tribe School and College Students in India*. New Delhi: Allied Publishers.
- Sharma, Surendra 1985 *Sociology in India: A Perspective from Sociology of Knowledge*. Jaipur: Rawat Publications.
- Singh, Yogendra 1986 *Indian Sociology: Social Conditioning and Emerging Concerns*. New Delhi: Vistar Publications.
- Srinivas, M. N., and M. N. Panini 1973 "The Development of Sociology and Social Anthropology in India." *Sociological Bulletin* 22:179-215.
- Unnithan, T. K. N., Yogendra Singh, Narendra Singhi, and Indra Deva, eds. 1967 *Sociology for India*. New Delhi: Prentice Hall.



Visaria, Leela, and Pravin Visaria 1995 *India's Population in Transition*. (Population Bulletin 50: 3). Washington, D.C.: Population Reference Bureau.

— 1996 *Prospective Population Growth and Policy Options for India, 1991–2101*. New York: Population Council.

VIMAL P. SHAH

## INDIGENOUS PEOPLES

For well over a century claims have been advanced that Native Americans, and indigenous peoples in general, are about to vanish (Bodley 1990, 1994; Dippie 1982). Apparently, however, indigenous peoples have neither read nor followed these scripts. In fact, quite the opposite has occurred. In the United States and around the world there has been a resurgence of indigenous consciousness, political mobilization, and cultural renewal (Cornell 1988; Nagel 1996; Snipp 1988a, 1989, 1992; Thornton 1987; Wilmer 1993). Groups in Canada, the United States, Australia, Brazil, Southeast Asia, and elsewhere are making land claims, petitioning for political rights, and demanding control of resources with remarkable success given their nearly universal paucity of votes, money, or military means. Interestingly, these indigenous movements are occurring while the number of people who live “traditional” or “tribal” lifestyles has diminished under the onslaught of expanding national and global industrialism and capitalism. A full description and explanation of these contradictory trends would require several volumes. Here we offer a summary of current understanding of the state of indigenous peoples in North America and around the world, and we suggest why such issues are of vital concern to sociology.

### TERMINOLOGY

Ethnic terminology is notoriously politically controversial and loaded. Terms such as “tribe,” “clan,” “ethnic entity,” and “nation” have been used over the last few centuries as weapons of both the strong and the weak in wars of words, laws, and often guns, to attack or defend the rights and survival of indigenous peoples. Even the term “indigenous peoples” is problematic—after all, everyone is indigenous to some place, and indigenousness sometimes can be a matter of

when the clock starts. We use the term “indigenous” to refer to those peoples who either live, or have lived within the past several centuries, in nonstate societies, although these indigenous societies may well have existed within the boundaries of state societies. We eschew the term “pre-state” because it implies, even sometimes unintentionally, that there is a necessary, progressive evolution from nonstate to state societies. We hasten to add that although some states have existed for several millennia (Sanderson 1999) in a variety of forms, the diversity of types among nonstate societies is far greater. This why there is a plethora of terms to refer to them: clans, bands, macrobands, tribelets, tribes, chiefdoms, segmentary lineages, and so forth. Virtually all these terms entail an attempt to organize this diversity (for discussion of the terms and concordance of the various meanings, see Chase-Dunn and Hall 1997, chs. 4, 7).

The most problematic in this litany of terms is “tribe.” As much as anthropologists and others have argued that “tribe” has become so general as to be useless (Fried 1975), it often is used in efforts to communicate with the general public or beginning students. The problem is especially salient for those peoples who inhabited North America before Europeans arrived, since the tribe–nation distinction has often been used politically to support or to deny autonomy or sovereignty for indigenous groups. To further confuse matters some, indigenous communities officially and legally call themselves “tribes,” though many have replaced “tribe” with “nation.” Even the designation “Native American” is not without problems, since legally, anyone born in the United States is a “native”[-born] American. For all these reasons we use “nonstate society” as a generic term, and for those peoples indigenous to the Americas, we alternate among Native Americans, American Indians, native peoples, and indigenous peoples.

When referring to a specific indigenous community, we use the name of the group—but even that is often problematic. There are four broad problems in regard to group names that we discuss here because of the light these difficulties shed on the issues facing indigenous peoples generally. Our discussion deals specifically with North America, although the issues we raise often are faced by native peoples elsewhere. First, membership in indigenous groups may change over time as various forms of identity and political organization

change in response to internal processes or encounters with outsiders. In early periods of contact with Europeans, North American native peoples often shared a broad sense of identity but were not ruled by any single social or political organization (Cornell 1988). Through years of contact, this situation frequently reversed itself. The need for unified resistance to European, then American, encroachments often necessitated the formation of sociopolitical structures that encompassed individuals and communities that had not necessarily shared historical cultures or identities (for examples, see Champagne 1989, 1992; Dunaway 1996; Faiman-Silva 1997; Fenelon 1998; Hall 1989; Himmel 1999; Meyer 1994).

Second, many historical indigenous cultures and communities have been destroyed, either by outright genocide, the devastations of disease, assimilation into European societies, or merger or amalgamation with other indigenous groups. As a survival strategy, many native groups found themselves greatly transformed, in particular through the consolidation of diverse individuals and communities. At times these amalgamated communities represented a form of “ethnogenesis,” that is, the creation of new native groups.

Third, a great deal of ethnographic and ethnohistorical investigation shows that the symbolic, demographic, and social boundaries of nonstate groups are extremely permeable (as, indeed, are those of many states and empires). Thus, the expectation of fixed, clear, rigid boundaries or borders is an artifact of the creation of the modern European nation-state, and of the needs of European and American negotiators to identify “leaders” of native societies for purposes of treaty-making and land acquisition. Hence naming a group often gave a false sense of unity, solidity, and organization.

Finally, there are the historical accidents of naming and the vagaries of spelling that stemmed from a lack of clear understanding of indigenous languages. One of the most notorious is the naming of the Lakota peoples as “Sioux,” which is a French corruption of an Anishinaabe [Chippewa or Ojibwa] word, “nadowasieux,” which translates into something like “slimy snake people” (Tanner 1987, p. 4)—certainly, not a name many Lakotas would wish to be known by. Many Native American groups are shifting back to their own names

for themselves, rather than continuing to use those assigned them by outsiders. For instance, Diné is increasingly used to refer to Navajo institutions and people, and the former “Winnebagos” of Wisconsin are now officially the “Ho Chunks.”

### SOCIOLOGICAL SIGNIFICANCE OF INDIGENOUS PEOPLES

There are several reasons that the study and understanding of indigenous peoples and the challenges they face are of special interests to sociology and to sociologists. First, in the United States (and every country in the Americas and in many others elsewhere) indigenous peoples are part of an ethnically diverse social, political, cultural, and economic landscape. Despite their often relatively small numbers, placing indigenous peoples into the ethnic mosaic of contemporary nations and states puts later settler and immigrant groups into a more accurate and larger historical context. This is a point that goes beyond “political correctness” or broad “multiculturalism.” The history and current conditions of native peoples often highlight questions of group rights, nation formation, justice, and social change that are relevant to all ethnic communities, not just indigenous groups. To ignore any group in academic discussion of majority–minority relations is itself a form of racism that denies the existence or legitimacy of that group.

Second, of considerable importance in developing general theoretical accounts of intergroup relations is the fact that indigenous peoples present a wide variety of social structures that are not found among state groups, such as Americans, or immigrant groups within states, such as Cuban Americans. Thus, all theories of intergroup relations that study only state peoples will lack dimensions unique to indigenous peoples, such as particular spiritual traditions or patterns of social relations. For instance, most sociologists are familiar with the U.S. racial classification norm called the “one-drop rule”: If an individual has any African ancestry, then that person is considered African American no matter what the person’s appearance or skintone. But for Native Americans—another colonized, conquered, and oppressed group—the same rule does not apply. Often standards require much more than “one drop” for an individual to be officially considered an “Indian.”

A common standard requires one-quarter Indian ancestry (or “blood quantum”), that is, that an individual must have one grandparent or two great-grandparents who are American Indian to qualify as a “real” Indian. Sometimes such ancestry rules are official U.S. government regulations, and more than occasionally they are tribal government rules (see Meyer 1994 for a detailed discussion). No such “one drop” or “blood quantum” rules apply to other U.S. immigrant or ethnic groups.

Third, since the formation of the United States, Native Americans have had a very special political and social relationship with the U.S. government. They are the only ethnic community that has legal right to direct federal action and accountability that bypass city, county, and state governmental authority. This “government-to-government” tribal-federal relationship generates many politically and sociologically interesting interactions and exceptions. One example can be found in the controversies about gaming on Indian reservations that emerged in the last quarter of the twentieth century; another involves treaty-based Indian hunting and fishing rights.

Fourth, theories of long-term social change and social evolution that do not include analysis of indigenous peoples will be inherently defective due to a biased sample of societies examined. There is another danger—that of assuming that surviving indigenous people, even those who live “traditionally,” are models or “living artifacts” of earlier societies. Contemporary indigenous peoples have survived centuries—and in parts of Asia, millennia—of contact and interaction with state societies. Their contemporary social structures have been shaped by their responses to those interactions. Indeed, Ferguson and Whitehead (1992) argue persuasively that these state/nonstate contacts so profoundly change both types of societies that scholars must view even the earliest first-hand accounts of indigenous societies with considerable skepticism. This is because typically by the time a representative of state society who produces written records observes an indigenous group, there has already been considerable interaction, and what that observer sees already has been shaped by that interaction. This is not denying that there are occasional observations that reflect very little interaction and change, but they are very rare. This raises questions, for example, about the accuracy

of depictions of western U.S. tribes by such early travelers (1804) as Meriwether Lewis and William Clark.

In summary, to ignore indigenous peoples leads to bad sociology in the form of theories and explanations that are based on biased samples, that cover only a truncated range of processes and relations found in societies, and that lead to a distorted picture of ethnic, cultural, social, political, and economic diversity in any society with indigenous communities. This is why it is important to overcome the nineteenth century legacy of the division of labor between anthropology and sociology in their objects of study. This is no easy task given the great diversity of indigenous groups in the world, or even in the continental United States.

### DEMOGRAPHIC ISSUES

One of the more formidable problems in studying indigenous peoples is describing their demography. Here, too, there are several interesting challenges, even if discussion is restricted to the United States. First, there is the politics of numbers. This derives from changes, and in some cases, improvements in historical demography in the twentieth century and the uses to which numbers are put. Stiffarm and Lane (1992) argue persuasively that there is an inherent tendency to minimize the historical population of Native Americans prior to European contact in order to support the argument that destruction of the indigenous American population was not extreme and mostly accidental. While estimates for the indigenous population of North American (United States and Canada) range from 1 million to 30 million, Thornton (1987) argues for a figure in the neighborhood of 7 million, based on careful reconstruction of population densities, early population counts, and the effects of known epidemics. From 1492 on, Native populations declined drastically, primarily, but not exclusively, because of exposure to “Old World” diseases. For the continental United States, the absolute population nadir of about a quarter million was reached around the turn of the twentieth century. Thereafter, population has grown steadily, so that at the end of the twentieth century the Native American population is approximately 2 million, or between one-third and

one-half of what it likely was in 1492. An important point here is that various studies show that more than disease was involved in the initial depopulation. That outright genocide contributed to the nearly total destruction of the Native American population is now well-documented (see Thornton 1987; Stannard 1992).

The impressive population recovery of the post-Second World War period is worth some mention. From 1960 to 1970, the number of Americans who reported their race to be "American Indian" in the U.S. census grew by 51 percent (from 523,591 to 792,730); from 1970 to 1980, the American Indian population grew faster, by 72 percent (to 1,364,033); and from 1980 to 1990, the American Indian population increased by 37 percent (to 1,878,285). Several reasons are given for this growth, including improved enumeration techniques, a decreasing death rate, and an increasing willingness of individuals to identify themselves as Native American. An important feature of the contemporary Native American population is the extensive intermarriage of indigenous peoples with non-Indians. Intermarriage has given rise to three distinct types of U.S. Indian population (Snipp 1986). First, there are "American Indians," persons who claim to be Indian racially and ethnically (having a specific tribal identification). Second, there are "American Indians of multiple ancestry," persons who claim to be Indian racially but have significant non-Indian ancestry. Third, there are "Americans of Indian descent," who do *not* claim to be Indian racially but report an Indian component in their background. The second two categories contain a number of individuals whose ethnic and racial identities readily shift with political, economic, and social contexts. By 1990 many, if not most, Indians were marrying outside their tribal group, and many were marrying non-Indians. As the number of individuals of ambiguous, and often ambivalent, Indian identity has increased, questions about membership in Indian tribes and definitions of who is and is not really an Indian have been raised by tribal governments, federal officials, and Indian communities and individuals. "Indianness" has become an empirical measurement issue, a political issue, and a theoretical issue. With the financial successes of some native community enterprises (in gaming, natural resources, and tourism), questions of tribal membership have become an economic issue as well.

## ECONOMIC DEVELOPMENT

In the twentieth century, political incorporation, assimilation, and economic interaction have tended to attenuate cultural processes and heighten political processes of tribal governments (Cornell 1988; Cornell and Kalt 1999). Access to wealth from mineral resources, gaming, and tourism has helped economic development. Snipp (1988b) shows, however, that many differences between Indian nations with energy resources and those without such resources tend to be minimal. A key problem in economic development—and one that is especially salient among indigenous communities—is how to participate in and benefit from economic development without simultaneously undermining or destroying traditional Indian values (Cornell and Kalt 1992; Ward and Snipp 1996, especially Ward's chapter). Not surprisingly, these issues are enmeshed in indigenous political action globally (Wilmer 1993).

## AMERICAN INDIAN POLITICAL ACTIVISM

The urbanization, intermarriage, education, increased participation in the paid labor force, and bicultural character of the American Indian population during the post-Second World War period gave rise to the most politically active period in American Indian history. The 1960s, 1970s, 1980s, and 1990s saw Native Americans organize themselves into political activist organizations (the American Indian Movement, Women of All Red Nations), protest movements ("Red Power," Camp Yellow Thunder in the Black Hills), legal defense organizations (Native American Rights Fund, Native Action), and lobbying groups (National Congress of American Indians, National Tribal Chairmen's Association). These organizations and movements were established and grew in the fertile political soil of the civil rights era ethnic politics in the United States. The following decades were marked by a range of Native American protest events, from the "fish-ins" in the Pacific northwest in the mid-1960s, to the nineteen-month occupation of Alcatraz Island beginning in 1969, to the seventy-one-day siege at Wounded Knee on the Pine Ridge Reservation in South Dakota in 1973, to the occupation of Camp Yellow Thunder in the Black Hills in the 1980s, to the protests against Indian athletic mascots of the 1980s and 1990s (see Johnson 1996; Nagel 1996). Against this

backdrop of marches, occupations, protests, and sometimes open conflict, many legal battles were waged in U.S. and tribal courtrooms across the country. Out of both the protest and the legal battles came a new “self-determination” era in federal Indian policy. Self-determination opened the way to increased tribal control of budgets and decision making, to the development of tribally owned natural resources, to the establishment of casinos and gaming on tribal land, and to opportunities for self-rule and economic development by Indian communities. These political and economic opportunities have raised social questions about the rules for tribal membership, which will remain a subject for debate well into the twenty-first century because of the projected continued growth of the Indian population.

#### **OTHER CONTEMPORARY NATIVE ISSUES**

Two of the major contemporary issues facing Native Americans are debates over gaming and renewed interest in indigenous religion. Because of their special relationship with the U.S. federal government, reservation governments are able to sponsor gaming over the opposition of state authorities; to sell gasoline and cigarettes without paying local and state taxes; and to sell other products, such as fireworks, that are typically regulated or made illegal by state or local governments. Of all these, the most controversial has been the establishment of gaming facilities. Indeed, these bingo parlors and casinos have spawned social movements that are nominally antigaming but are often very thinly disguised anti-Indian movements, and in some cases reflecting a conflict of interest, as when state authorities see Indian gaming as unfair competition to state-run lotteries and other gaming enterprises. Similar non-Indian opposition has resulted from renewed Indian land claims (such as by the Oneidas in upstate New York in 1998–1999). These controversies have heightened the stakes of identity politics both within native groups and between native groups and the general population.

Initially, one might expect that renewed interest in and presumably respect for native religions by non-Indians might have been received positively by Native Americans. However, this is generally *not* the case. Often non-Indian appropriation of Indian spiritual traditions is perceived by native

people as a final theft. After stealing Indian land, mineral rights, water rights, and fishing rights, the final non-Indian assault on native peoples is to usurp and subvert Indian culture. This has not been helped by the number of charlatans and hucksters (a few of whom are of native ancestry) involved in assorted “New Age” appropriations of Indian cultural elements, typically lifted entirely out of their indigenous context (see Churchill 1994, 1996; Rose 1992). The spread of New Age and “world” music, which uses elements and occasionally performers from various indigenous populations, has spawned analogous controversies at a global level (Feld 1991). Not the least of the subcontroversies is that it is non-Indian performers and producers who are making the large profits from the use of indigenous instruments, themes, music, and performances. Such controversies will not disappear quickly. They have, however, generated a new interest in relations with indigenous peoples and new attempts to reexamine the long and often tawdry history of Indian/non-Indian relations.

#### **CONCLUSION**

The sociology of indigenous peoples, including their relations to state peoples, is multifaceted, and fascinating. It is also a vital and necessary component to the study of the sociology of intergroup relations. Many complex and important sociological processes occur primarily, and sometimes only, in relations between state and nonstate peoples. In order to develop robust theories of intergroup relations, of social movements, and of globalization and resistance to its negative consequences, the study of indigenous peoples is indispensable. Theories built solely on the study of immigrant populations are necessarily flawed; those that include indigenous peoples are much richer.

#### **A NOTE ON REFERENCES**

The literature on indigenous peoples is immense, even in sociology. We have cited works which contain extensive bibliographies. We also suggest that those interested in this area use the World Wide Web to search for information on Native American studies, American Indian studies, and individual tribes—many of which have their own Web pages. The History Net list, H-AMINDIAN, is also an excellent resource. Some useful scholarly

journals are: *American Indian Culture and Research Journal*, *American Indian Quarterly*, *Native America*, *Wicazo Sa Review*, and *Cultural Survival Quarterly*.

## REFERENCES

- Bodley, John H. 1990 *Victims of Progress*, 3rd ed. Mountain View, Calif.: Mayfield Publishing.
- 1994 *Cultural Anthropology: Tribes, States, and the Global System*. Mountain View, Calif.: Mayfield Publishing.
- Champagne, Duane 1989 *American Indian Societies: Strategies and Conditions of Political and Cultural Survival*. Cambridge, Mass.: Cultural Survival.
- 1992 *Social Order and Political Change: Constitutional Governments Among the Cherokee, the Choctaw, the Chickasaw, and the Creek*. Stanford, Calif.: Stanford University Press.
- Chase-Dunn, Christopher, and Thomas D. Hall 1997 *Rise and Demise: Comparing World-Systems*. Boulder: Westview Press.
- Churchill, Ward 1994 *Indians Are Us: Culture and Genocide in Native North America*. Monroe, Maine: Common Courage Press.
- 1996 "Spiritual Hucksterism: The Rise of the Plastic Medicine Men." In Ward Churchill, ed., *From a Native Son: Selected Essays on Indigenism, 1985–1995*. Boston: South End Press.
- Cornell, Stephen 1988 *The Return of the Native: American Indian Political Resurgence*. New York: Oxford University Press.
- Cornell, Stephen, and Joseph Kalt 1992 *What Can Tribes Do?: Strategies and Institutions in American Indian Economic Development*. Los Angeles: American Indian Studies Center, University of California.
- 1999 "Sovereignty and Nation-Building: The Development Challenge in Indian Country Today." *American Indian Culture and Research Journal* 22(3):87–214.
- Dippie, Brian W. 1982 *The Vanishing American: White Attitudes and U.S. Indian Policy*. Lawrence: University Press of Kansas.
- Dunaway, Wilma A. 1996 *The First American Frontier: Transition to Capitalism in Southern Appalachia, 1700–1860*. Chapel Hill: University of North Carolina Press.
- Faiman-Silva, Sandra L. 1997 *Choctaws at the Crossroads: The Political Economy of Class and Culture in the Oklahoma Timber Region*. Lincoln: University of Nebraska Press.
- Feld, Steven 1991 "Voices of the Rainforest." *Public Culture* 4(1):131–140.
- Fenelon, James 1998 *Culturicide, Resistance, and Survival of the Lakota (Sioux Nation)*. New York: Garland Publishing.
- Ferguson, R. Brian, and Neil L. Whitehead, eds. 1992 *War in the Tribal Zone: Expanding States and Indigenous Warfare*. Santa Fe, N.Mex.: School of American Research Press.
- Fried, Morton 1975 *The Notion of Tribe*. Menlo Park, Calif.: Cummings.
- Hall, Thomas D. 1989 *Social Change in the Southwest, 1350–1880*. Lawrence: University Press of Kansas.
- Himmel, Kelly D. 1999 *The Conquest of the Karankawas and the Tonkawas: A Study in Social Change, 1821–1859*. College Station: Texas A & M University Press.
- Johnson, Troy 1996 *The Indian Occupation of Alcatraz Island: Self-Determination and the Rise of Indian Activism*. Urbana: University of Illinois Press.
- Meyer, Melissa L. 1994 *The White Earth Tragedy: Ethnicity and Dispossession at a Minnesota Anishinaabe Reservation, 1889–1920*. Lincoln: University of Nebraska.
- Nagel, Joane 1996 *American Indian Ethnic Renewal: Red Power and the Resurgence of Identity and Culture*. New York: Oxford University Press.
- Rose, Wendy 1992 "The Great Pretenders: Further Reflections on Whiteshamism." In M. Annette Jaimes, ed., *The State of Native America: Genocide, Colonization, and Resistance*. Boston: South End Press.
- Sanderson, Stephen K. 1999 *Social Transformations: A General Theory of Historical Development*, 2nd ed. Boulder: Rowman and Littlefield.
- Snipp, C. Matthew 1986 "Who Are American Indians? Some Observations About the Perils and Pitfalls of Data for Race and Ethnicity." *Population Research and Policy Review* 5:237–252.
- (ed.) 1988a *Public Policy Impacts on American Indian Economic Development*. Albuquerque: Native American Studies (Development Series No 4).
- 1988b. "Public Policy Impacts and American Indian Economic Development." In C. Matthew Snipp, ed., *Public Policy Impacts on American Indian Economic Development*. Albuquerque: Native American Studies (Development Series No 4).
- 1989 *American Indians: The First of This Land*. New York: Russell Sage Foundation.
- 1992. "Sociological Perspectives on American Indians." *Annual Review of Sociology* 18:351–370.
- Stannard, David E. 1992 *American Holocaust: Columbus and the Conquest of the New World*. New York: Oxford University Press.
- Stiffarm, Lenore A., with Lane, Phil, Jr. 1992 "The Demography of Native North America: A Question

of American Indian Survival." In M. Annette Jaimes, ed., *The State of Native America: Genocide, Colonization, and Resistance*. Boston: South End Press.

Tanner, Helen H., ed. 1987 *Atlas of Great Lakes Indian History*. Norman: University of Oklahoma Press.

Thornton, Russell 1987 *American Indian Holocaust and Survival*. Norman: University of Oklahoma Press.

Ward, Carol, and C. Matthew Snipp, eds. 1996 *Research in Human Capital and Development: Vol. 10, American Indian Economic Development*. Greenwich, Conn.: JAI Press.

Wilmer, Franke 1993 *The Indigenous Voice in World Politics: Since Time Immemorial*. Newbury Park, Calif.: Sage.

THOMAS D. HALL  
JOANE NAGEL

## INDIVIDUALISM

Individualism is a doctrine concerning both the composition of human society and the constitution of sociocultural actors. The term was invented in the 1820s, apparently, in France (Swart 1962). Its first appearance in English dates from the 1835 translation of Alexis de Tocqueville's study of the United States (Tocqueville [1850] 1969, p. 506). The basic notion conveyed by the newly coined word, that the individual is sovereign vis-à-vis society, was intensely controversial, for it stood on the grave of one established order, proclaiming the rise of another. As an early French critic saw it, individualism "destroys the very idea of obedience and of duty, thereby destroying both power and law," leaving nothing "but a terrifying confusion of interests, passions and diverse opinions" (cited in Lukes 1973, p. 6).

Individualism should be distinguished from historically specific constitutions of the individuality of human beings. The word "individual," used to discriminate a particular human being from collectivities ("family," "state"), had been in circulation for centuries prior to Tocqueville (albeit mainly as an adjective), and individualizations had been practiced under one description or another long before that, at least as evidenced in the oldest surviving texts of human history. However, premodern constitutions of individuality did not become the foci of a distinctive doctrine of individualism. That development came in response to

the profound changes of social structure and consciousness that had been slowly accumulating during the seventeenth and eighteenth centuries. In the transformation from a medieval to a modern world, new transparencies of meaning evolved—among the most important, a particular conception of "the individual." The enormous power of that conception is reflected in the fact that people of modern society have generally had no doubt as to what an "individual" is. The reference has been self-evident because the object referred to, an *individual*, has been self-evident, pregiven, natural.

But one must remember that "the individual" is a construct. Like all constructs, it is historically variable. The meaning of individualism's "individual" was formed under specific historical circumstances that, in practice as well as in ideology, increasingly prized values of rational calculation, mastery, and experimentation; deliberate efforts toward betterment of the human condition; and a universalism anchored in the conviction that "human nature" is basically the same everywhere at all times and that rationality is singular in number. These commitments were manifested in the doctrine of individualism (as, indeed, in the formation of the modern social sciences). By the time of Tocqueville and the newly coined word, individualism's individual had become integral to much of the practical consciousness of modern society. Human beings were being objectified as instances of "the individual"—that is, as instances of a *particular kind* of individuality.

The forces created during that formative period wrought great changes in the fabric of society, many of which continue to reverberate. Of course, as historical circumstances have changed, both "the individual" of individualism and the constitution of individuality have changed. Nonetheless, a certain transparency of meaning remains still today in our practical consciousness of "the individual," and it is still informed by a doctrine of individualism. Thus, when a sociologist says that "a natural unit of observation is the individual" (Coleman 1990, p. 1), he can assume without fear of failure that most of his readers will know exactly what he means.

The remainder of this article offers brief accounts of (1) the development of individualism during the seventeenth, eighteenth, and nineteenth

centuries, (2) recent shifts of emphases in individualism's conception of "the individual," and (3) some current issues and concerns. More extended treatments can be found in Macpherson (1962), Lukes (1973), Abercrombie and colleagues (1986), Heller and colleagues (1986), and Hazelrigg (1991), among others.

### THE SELF-REPRESENTING INDIVIDUAL

While elements of individualism can be seen in expressions of practical affairs as early as the twelfth-century renaissance (Macfarlane 1978; Ullmann 1966), the first more or less systematic statement of the doctrine came during the 1600s. Scholars such as René Descartes, Thomas Hobbes, and John Locke believed that in order to understand a whole (e.g., society) one had first to understand the parts of which it is composed. In the case of society, those parts, the building blocks of a society, were instances of "the individual." Although disagreeing on various specific issues—for example, whether human agency is distinct from a natural world of causal necessity (Descartes) or a product of that causal necessity (Hobbes)—these seventeenth-century scholars displayed a remarkable confidence in their understandings of "the individual" as a presocial atom. Their individual was a highly abstract being, squatting outside the world.

In the premodern order of European society, social relations had been organic, corporate, and mainly ascriptive. Sovereignty was a complex relation of duty, responsibility, and charity, focused on a specific location in the hierarchical order of organic community. Certainly members of the community were individualized, but the distinction was constituted primarily by ascriptive position in the hierarchical order. It is clear from surviving documents of the twelfth century, for example, that one individual knight was discriminable from any other in ways that we would describe as "personality." With rare exceptions, however, the discrimination was local. Otherwise, knights were discriminable mainly by pedigree, lines of fealty, and quality of chivalry. Individuals could rise (and fall) through gradations of rank, but vertical movement was first within the family or household group. A knight who aspired to still higher standing had first to be retained in another, more powerful family. Similarly, while there is no reason

to doubt that residents of a twelfth-century village or town could reliably discriminate one another by physiognomic features, the portrayal of human beings in paintings concentrated on matters of costume, placement, and posture to signal social distinctions among virtually lifeless mannequins. Even in the revolutionary works attributed to Giotto (1267–1337), whose use of subtle gestures and glances began to individualize portrayals sufficiently to be called portraits in the modern sense, such facial features as warts, wens, moles, wrinkles, lines, scars, and sagging skin were still quite irrelevant to, and therefore absent in, the visualized identity or character of a person. But by the end of the 1400s we see, most notably in Domenico Ghirlandaio's *Portrait of Old Man and Boy* (c. 1490), depictions that to the modern sensibility count as realistically individualized faces. The art of portraiture gradually developed into advertisements of a new actor, front and center—modernity's sovereign individual (Haskell 1993).

In this new order, by contrast to its predecessor, social relations were conceived as contractual rather than organic, based on achieved rather than ascribed traits of individuals liberated from constraints of community. City life was once again the center of gravity in territorial organization, having displaced the manorial system. What would become the modern nation-state was beginning to take shape during the period of relative peace inaugurated by the several Treaties of Westphalia in 1648. Within this context of invention and experimentation with new (or renewed) organizational forms, the new individual was conceived as a wholly separate entity of self-identical integrity, a "bare individual" who could freely consent to enter into concert with other, equivalently constituted individuals, each propelled by self-interest. This was, as Macpherson (1962) described it, the advent of "possessive individualism," and it correlated well with the developing motivations of capitalism.

By the end of the eighteenth century, individualism had attained mature statement in treatises by David Hume, Adam Smith, and Immanuel Kant, among others. This mature statement, worked out in the context of rapidly changing political-economic institutions, emphasized the centrality of a "self-representing individual." The chief claim—that "every individual appears as the autonomous subject of his [or her, but primarily his] decisions



and actions” (Goldmann [1968] 1973, p. 20)—served as the linchpin to formalized explanations of the political and economic rights of members of society, especially the propertied members. Expressions of the chief claim in moral and legal rights of the individual became enshrined in newly invented traditions, in legitimizing principles such as “popular sovereignty” and “inalienable rights,” and in documents of public culture such as the Declaration of the Rights of Man and the Constitution of the United States of America (Hobsbawm and Ranger 1983; Morgan 1988). The prayerful injunction “God bless the squire and his relations and keep us in our proper stations” had been replaced by the almost wholly secular “I pledge allegiance to the flag” (i.e., to an abstract sign). While the claim of autonomy emphasized the universality of rights and the particularity of the “I,” the practical emphasis on a self-representing individual was formulated in political-economic terms that “necessitated” elaborate definitions and procedures for the defense of “property rights” long before equivalent attention would be given to, say, “rights of the handicapped.”

Much like the individual of organic community, the self-representing individual is a substantial presence, manifest as the embodiment of a uniform human nature, and as such is the bearer of various traits, dispositions, and predications. However, the site of the self-representing individual’s capacity of agency and potential for autonomy is neither the community nor the accumulated traits, dispositions, and predications. Rather, it is deeply interior to what became a new “inner nature” of the human being. Beneath the faculty of reason, beneath all feeling and emotion and belief, there is “the will.” Emile Durkheim ([1914] 1973) described it as the egoistic will of the individual pole of *homo duplex*, for George Herbert Mead (1934) it was “the principle of action.” But before either of those sociologists, the master theorizer of the self-representing individual, Immanuel Kant, had formulated the basic principle as the pure functioning of the “I” through time. Only because I can unite a variety of given representations of objects in *one* consciousness, Kant ([1787] 1929, B133) argued, is it possible that I can “represent to myself the *identity of consciousness*,” throughout those representations. In other words, the very possibility of a knowledge of the external world is dependent on the temporal continuity of the “I.”

The individual is absolute proprietor of this pure functionality, this willing of the “I” as basic principle of action; the individual owes absolutely nothing to society for it.

By conceiving the essential core of human individuality to be a deeply interiorized, radically isolated pure functionality, connections between the individual and the substantive traits that he or she bears become arbitrary. The individual is formally free to exercise choice in which traits to bear, free to be mobile geographically, socially, culturally, personally. Ascriptive traits are devalued in favor of achieved traits, and one set of achieved traits can always be exchanged for yet another set. This principle of freely exchangeable traits, an aspiration directing progress toward “the good society,” depended on new means of socialization (or “internalization of norms”), so as to insure sufficient regularity in processes of exchange. Indeed, the self-representing individual was central to a distinctive regimen of behavior, a new practical meaning of “discipline” (Foucault [1975] 1977). As the doctrine of individualism saw it, the contractual, associative forms of social relation, though looser fitting in their normative constraints than the old organic community, were complemented by the figure of “the self-made man” who had internalized all the norms of rectitude and propriety so nicely as to merit life in an unprecedentedly free society. When reality failed the image, there were courts and legal actions for deciding conflicts of interest and the clash of individuals’ rights. Notably, not a single book on the law of torts had been published in English by the mid-1800s; the explosive growth of tort law and third-party rules was only beginning (Friedman 1985, pp. 53–54).

Because the doctrine of individualism provided a set of answers to questions that were foundational to sociology (as to the social sciences in general)—What is the individual? How is society possible? and so forth—virtually every topic subsequently addressed by sociology has in one way or another involved aspects of individualism. Given the composition of individualism’s self-representing individual, the most prominent issues have often centered on questions of relationship between the rise of individualism and the development of new forms of political-economic organization as manifested in capitalism, bureaucracy, and the modern state. Indeed, that relationship was

focus of one of the great controversies occupying many early sociologists (Abercrombie et al. 1986). Hardly anyone doubted the existence or importance of a relationship. Rather, the debates were about such issues as causal direction (which caused which?), periodizations (e.g., when did capitalism begin?), and whether ideas or material conditions (each category conceived as devoid of the other) were the primary motive force. In many respects the debates were a continuation of the struggles they were about.

Other, more specific topics addressed by sociologists have also involved aspects of the rise of individualism. Several have already been mentioned (e.g., a new regimen of discipline). Additional examples are the development of sectarian (as opposed to churchly) religions, followed by an even more highly privatized mystical-religious consciousness of the isolated individual; changes in domestic architecture, such as greater emphasis on individualized spaces of privacy and functionally specialized rooms; changes in table manners, rules of courtesy, and other “refinements of taste”; the emergence of a “confessional self” and practices of diarykeeping; increased emphasis on romantic love (“affective individualism”) in mate selection; new forms of literary discourse, such as the novel and autobiography; the rise of professionalism; and the rise of the modern corporation as an organizational form that, having gained standing as a legal actor comparable to a flesh-and-blood person, cast into doubt reliance on understandings of “the will” as foundation and motive force of contractual relations (see Abercrombie et al. 1986; Horwitz 1992; Perrot [1987] 1990).

### THE SELF-EXPRESSING INDIVIDUAL

The figure of the self-representing individual proved to be unstable, even as the meaning of representation gradually changed. This was mainly because the same factors that had produced this version of individualism’s individual also led to dissolution of the transparent sign. For example, whereas clothing, manners, bodily comportment, and similar traits had been, in the old order, reliable signs (representations) of a person’s rank or station in life, the sign became increasingly arbitrary in its relationship to ground. This loosening of the sign, together with a proliferation of signs in exchange,

led to a new universalism of “the empty sign.” The prototype was money and the commodity form: Devoid of intrinsic value and capable of representing everything, it represents nothing in particular. As one recent scholar has described the process, borrowing a clause from Karl Marx, “all that is solid melts into air” (Berman 1983).

At the same time, the rhetoric of transhistorical forms of value (e.g., the commodity, inalienable rights) allows for an enormous amount of individual variation in the sociocultural conditions under which it can succeed. Individualism’s emphasis on the bare individual was being increasingly generalized, further reducing the import of group-based relations and traits. In the mid-1800s, for instance, the distinction between public affairs and private matters was drawn at the door of home and family. Family life provided the chief “haven” of organic relations, nurturant domesticity, and refuge from the trials of work and politics. But soon the haven itself became a site of struggle toward still greater individuation. Of the many factors contributing to this rebellion against the traditional restraints of family, one of the most important was a new culture of sexuality, which manifested a more general and growing concern for the interior interests and needs of the individual.

Precedent for this concern can be seen in Kant’s conception of the self-representing individual (because of the transcendental “I,” every individual has in common a potential for *self-actualization*) as well as in romanticist movements of the early 1800s. However, the development of a new “inner discourse of the individual” has been mostly a twentieth-century phenomenon. The psychology of Sigmund Freud and his disciples formed part of that development, certainly; but another part was formed by the conception of a new “social citizenship” (Marshall 1964), which emphasized an individual’s rights of social welfare in addition to the earlier mandates of political and economic rights. A new figure of individualism’s individual gradually emerged, the “self-expressing individual.”

The individualism of the self-representing individual promoted the idea that all interests are ultimately interests of the bare individual. The new version of individualism both extends and modifies that idea. Whereas the self-representing individual puts a premium on self-control and

hard work, the self-expressing individual generalizes the value of “freedom of choice” from political-economic exchange relations to matters of personal lifestyle and consumption preferences (Inglehart 1990). The central claim holds that “each person has a unique core of feeling and intuition that should unfold or be expressed if individuality is to be realized” (Bellah et al. 1985, p. 336), and each person has the right to develop his or her unique capacities of self-expression. A recent change in divorce law partly illustrates the import of that claim. Prior to the 1960s one of the few organic relations still surviving in modern society was the marital bond; few conditions were deemed grave enough to have legal standing as grounds for breaking it. With the invention of “no fault” divorce (relatively noncontroversial legislation that spread rapidly from state to state; Jacob 1988), the marital relation became a civil contract much like any other, and a spouse’s freedom to choose divorce in the interest of satisfying unfulfilled needs of self-expression gained recognition.

Another manifestation of this self-expressing individual is the recent development of a specialized field, the sociology of emotions (e.g., Barbalet 1998; Thoits 1989). Certainly earlier scholars (e.g., Georg Simmel) had recognized emotive dimensions of sentiment, tradition, trust, and the like, and had assumed motivations such as fear of power, anxiety about salvation, envy of success, frustrated ambition, and eponymous glory. No one imagined that persons of premodern societies did not experience emotions—though usually this would have been in an idiom of “the passions,” and scholars disagreed whether these were in fact ahistorical, noncultural formations. But emotions had rarely been treated as important topics of inquiry by sociologists until the latter decades of the twentieth century. The emotive dimensions of life had typically been regarded as subordinate to other dimensions, just as the passions had been treated as dangerous when unbridled—energies which belonged properly in the harness of reason or “the rational faculty.” (Most of the seven deadly sins, remember, were emotional states or effects of emotional states.) When rightly yoked by reason, emotive energies of the self-representing individual could achieve highly valued public outcomes even if the particular emotion itself was classified as a vice. Thus, avowed Adam Smith among others, private greed or avarice could become, through

the device of market transaction, a public benefit. For the self-expressing individual, on the other hand, emotive aspects of life are, or should be, valued in themselves, not only for what might result from them. Whereas the eighteenth-century “pursuit of happiness” (one of the self-representing individual’s inalienable rights) was idiomatic for the unfettered formal liberty of the individual to pursue self-interest in commerce (an inherently outward-looking, social activity), for the self-expressing individual the pursuit of happiness refers to a much more introspective, privately evaluated emotional state, “being happy with who one is.”

Individualism’s self-expressing individual remains a trait-bearing substantial entity, to be sure. The variety of bearable traits is greatly expanded by the shift in emphasis from self-control to self-expression through lifestyle experimentation (“the person as work-in-progress-from-within,” as it were). Moreover, this shift in emphasis is accompanied by the stipulation that only those traits that an individual can freely choose to assume, and then jettison, should be relevant criteria by which to discriminate and evaluate individuals. Criteria falling outside the bounds of individual choice (“immutable” traits, whether biological or sociocultural) are deemed to be both irrelevant and, increasingly, a violation of an individual’s rights. In conjunction with the “entitlements” logic of social citizenship, this stipulation of a radically individualistic freedom of reversible choice has been linked to the emergence of a generalized expectation of “total justice” (Friedman 1985).

By the same token, the doctrine of individualism has always contained a large fictive component. Long after the doctrine proclaimed the sovereignty of the bare individual, for example, the actual individuality of human beings continued to be heavily marked by ascriptive traits (e.g., gender, race) and by sociocultural inheritances from one’s parents. The shift to an expressive individualism reflects efforts to situate the agency of a “free individual” outside the separately conceived domain of relations of domination. Rather than attempt to change those relations, the self-expressing individual would “transcend” them by concentrating on a logic of rights pertaining to the free expression of individual will in a domain of “personal culture” (Marcuse [1937] 1968).

Fictions can be productive in various ways, however. The fictions of individualism have often been made taskmasters, as women, African Americans, persons with disabilities, and other human beings discriminated primarily by ascriptive or group-based criteria have struggled to make reality conform to doctrinal image.

### SOME CURRENT ISSUES

Because individualism has been one of the professional ideologies of the social sciences (“methodological individualism”), a perennial issue concerns the proper structure of explanation—specifically, whether explanation of any sociocultural phenomenon must ultimately refer to facts about individuals, and if so, what precisely that means (Coleman 1990; Hazelrigg 1991; Lukes 1973). No one denies that collectivities are composed of individuals. But that truism settles neither the question of how “composition” is to be understood nor the question of what constitutes “the individual.” In short, the methodological issue involves a number of theoretical-conceptual issues, including several that are located at the intersection between individualism’s “individual” and historical variations in the actual constitution of individuality (Heller et al. 1986). The doctrine of individualism has consistently conceptualized “the individual” as a distinct and self-contained agent who acts within, yet separate from, a constraining social structure. Rather than being an ensemble of social relations, individualism’s individual is the substantial atom out of which any possible social relations are composed. This has certain implications for the empirical field.

How, for instance, does one understand the category “rational action”? What minimal criteria must be satisfied in order that a particular action can count as “rational”? The individual who stands forth in individualism tends to the heroic—self-made, self-reliant, and self-governing, rising above circumstances, taking charge of one’s own destiny, and, in the aggregate of similar atoms, building a better world. This individual was soon accorded central place as the chief fount of rational action. Whereas late eighteenth-century scholars such as Adam Smith continued to remind readers that rational action could and did stem from nonrational motivations (e.g., moral sentiment, the ethos of tradition, even the simple inertia of habit), pride of

place in the list of motivational sources increasingly shifted to that part of the “faculties of mind” called “the rational faculty,” conceived as an instrument of the will. The emphasis on rationality as faculty, an “instrumental rationality,” was no doubt encouraged by, and at the same time promoted, a growing list of successes in the inventiveness of carefully deliberated, calculated designs, plans, and projects of human engineering. (Recall the success of Dutch efforts, beginning on an ever larger scale in the 1600s and 1700s, to push back the sea and create thousands of square miles of new land—for most of us a barely remembered fact of history, but in those earlier times a truly audacious undertaking.) Are the limits of rational action therefore as closely circumscribed as the limits of an individual actor’s rational deliberations, calculations, and intentions? Most sociologists today are agreed in answering “no” to that question, although they diverge, sometimes sharply, in particulars of the answer (Coleman 1990; Kuran 1995; Sica 1988).

Social action can be highly rational in the aggregate, both in outcome and in process, even when individual actors, attending more to nonrational or irrational than to rational motivations and intentions, behave in ways that hardly fit the doctrinal image of “heroic actor.” People do learn from experience (if fitfully and slowly), and part of the accumulated fund from that learning consists in organizational forms that have rationalities built into them. This “rationality as form”—a materialized intelligence in the same way that a hand-held calculator, an airplane, or a magnetic resonance imager is a materialized solution to a set of problems—enables many more people to use, benefit from, and even operate the rationalities built into such devices than have the understanding (“rationality as faculty”) required to design their architectures or to convert design into working product. Moreover, whereas the examples just cited can be readily interpreted as immediate and punctual products of some specific individual’s rational faculty (thus fitting the heroic-actor model of the inventor or discoverer as genius), many other examples of rationality-invested organizational forms—family structures, markets, bureaucracy, and so forth—are as much or more the gradual accretions of indirection, sentiment, habituation, and happenstance than the intended consequences of rationally deliberated, calculated, instrumental

actions of particular individuals. The doctrine of individualism has often slighted the importance of these latter wellsprings of rational action, as if the actions they motivate do not quite count, or do not count in quite the same way, as action that directly manifests the willful force of an individual's reasoned intentions.

In a related vein, if the individual owes nothing to society for the "I" as principle of action, where should the distinction be drawn between determinants of action that are social and those that are psychological? Consider, for instance, a person suffering the characteristics clinically (and thus socially) categorized as "depression." Are these characteristics proper to the individual only, or are they also in some way descriptive of a social condition that is integral to the individual so characterized? If Willy Loman, of Arthur Miller's *Death of a Salesman*, is depressed, does that description say anything about the circumstances of a life Loman shared with countless others, or is it the description only of a misery interior to an isolated individual? Each of these accounts of what is under description has had its proponents. The popularity of a medicative regimen that emphasizes direct palliation or amelioration of psychic state (as in chemical applications, whether Prozac or psilocybin) suggests a growing preference for the second account.

Some sociologists contend that individualism's self-expressing individual is an accurate depiction of contemporary reconstitutions of individuality, and that in this new form of "the individual" the substance of selfhood, an individual's self-identical integrity, is being evacuated. Bellah and colleagues (1985) indict the emergence of "a language of radical individual autonomy" in which people "cannot think about themselves or others except as arbitrary centers of volition (p. 81)." Others argue that the emphasis on individual autonomy and separation is an expression of masculine, patriarchal values, as contrasted to feminine values of social attachment (Gilligan 1982). Still others see the development of an entirely new "order of simulacra" (Baudrillard [1976] 1983), in which simulation or the simulacrum substitutes for and then vanquishes the real (e.g., television images establish the parameters of reality). The alleged result is a collapse of "the social" into the indifference of "the masses," who no longer care to discriminate among "messages" (beyond their

entertainment effects) since one simulation is as good as another.

Equally contentious issues surround the evident growth in people's sense of entitlement and in the array of legal rights claimed and often won on behalf of "individual choice." Individualism's figure of the self-expressing individual is held by some to be the harbinger of a new age of democracy, by others to be the confirmation of a continuing trend toward greater atomization (Friedman 1985). Both assessments point to the emergence of a "rights industry" that promotes the invention of new categories of legal right pertaining to everything from a guaranteed freedom to experiment with unconventional lifestyles without risk of discrimination or retribution, to the rights of animals both individually and at the species level, to the possibility of endowing genes with "subject-like powers" and thus legal standing (Glendon 1991; Norton 1987; Oyama 1985). Some critics contend that the expansion of concern for increasingly particularized and "arbitrary" individual choices comes at the expense of a diminished concern for social outcomes. "Unless people regain the sense that the practices of society represent some sort of natural order instead of a set of arbitrary choices, they cannot hope to escape from the dilemma of unjustified power" (Unger 1976, p. 240). The conviction recalls that of the early French critic quoted in the opening paragraph.

## REFERENCES

- Abercrombie, Nicholas, Stephen Hill, and Bryan S. Turner 1986 *Sovereign Individuals of Capitalism*. London: Allen and Unwin.
- Barbalet, J. M. 1998 *Emotion, Theory, and Social Structure*. Cambridge, U.K.: Cambridge University Press.
- Baudrillard, Jean (1976) 1983 *In the Shadow of the Silent Majorities*. New York: Semiotext (e).
- Bellah, Robert N., Richard Madsen, William M. Sullivan, Ann Swidler, and Steven M. Tipton 1985 *Habits of the Heart*. Berkeley: University of California Press.
- Berman, Marshall 1983 *All That Is Solid Melts into Air*. London: Verso.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Durkheim, Emile (1914) 1973 "The Dualism of Human Nature and Its Social Conditions." In Robert Bellah, ed., *Emile Durkheim on Morality and Society*. Chicago: University of Chicago Press.

- Foucault, Michel (1975) 1977 *Discipline and Punish*, trans. Alan Sheridan. New York: Pantheon.
- Friedman, Lawrence 1985 *Total Justice*. New York: Russell Sage Foundation.
- Gilligan, Carol 1982 *In a Different Voice*. Cambridge, Mass.: Harvard University Press.
- Glendon, Mary Ann 1991 *Rights Talk*. New York: Free Press.
- Goldmann, Lucien (1968) 1973 *Philosophy of the Enlightenment*, trans. Henry Maas. Cambridge, Mass.: MIT Press.
- Haskell, Francis 1993 *History and Its Images*. New Haven, Conn.: Yale University Press.
- Hazelrigg, Lawrence 1991 "The Problem of Micro-Macro Linkages: Rethinking Questions of the Individual, Social Structure, and Autonomy of Action." *Current Perspectives in Social Theory* 11:229-254.
- Heller, Thomas C., Morton Sosna, and David E. Wellbery 1986 *Reconstructing Individualism*. Stanford, Calif.: Stanford University Press.
- Hobsbawm, Eric, and Terence Ranger, eds. 1983 *The Invention of Tradition*. Cambridge, U.K.: Cambridge University Press.
- Horwitz, Morton J. 1992 *The Transformation of American Law, 1870-1960*. New York: Oxford University Press.
- Inglehart, Ronald 1990 *Cultural Shift in Advanced Industrial Society*. Princeton, N.J.: Princeton University Press.
- Jacob, Herbert 1988 *Silent Revolution*. Chicago: University of Chicago Press.
- Kant, Immanuel (1787) 1929 *The Critique of Pure Reason*, 2nd ed., trans. Norman Kemp Smith. London: Macmillan.
- Kuran, Timur 1995 *Private Truths, Public Lies*. Cambridge, Mass.: Harvard University Press.
- Lukes, Steven 1973 *Individualism*. Oxford: Basil Blackwell.
- Macfarlane, Alan 1978 *The Origins of English Individualism*. Oxford: Basil Blackwell.
- Macpherson, C. B. 1962 *The Political Theory of Possessive Individualism*. Oxford: Oxford University Press.
- Marcuse, Herbert (1937) 1968 "The Affirmative Character of Culture." In *Negations*, trans. Jeremy J. Shapiro. Boston: Beacon.
- Marshall, T. H. 1964 *Class, Citizenship, and Social Development*. New York: Doubleday.
- Mead, George Herbert 1934 *Mind, Self, and Society*, ed. Charles W. Morris. Chicago: University of Chicago Press.
- Morgan, Edmund S. 1988 *Inventing the People*. New York: Norton.
- Norton, Bryan G. 1987 *Why Preserve Natural Variety?* Princeton, N.J.: Princeton University Press.
- Oyama, Susan 1985 *The Ontogeny of Information*. Cambridge, U.K.: Cambridge University Press.
- Perrot, Michelle, ed. (1987) 1990 *A History of Private Life*, vol.4, trans. Arthur Goldhammer. Cambridge, Mass.: Harvard University Press.
- Sica, Alan 1988 *Weber, Irrationality, and Social Order*. Berkeley: University of California Press.
- Swart, Koenraad W. 1962 "'Individualism' in the Mid-Nineteenth Century (1826-1860)." *Journal of the History of Ideas* 23:77-90.
- Thoits, Peggy A. 1989 "The Sociology of Emotions." *Annual Review of Sociology* 15:317-342.
- Tocqueville, Alexis de (1850) 1969 *Democracy in America*, 13th ed., J. P. Mayer, ed., George Lawrence, trans. New York: Doubleday.
- Ullmann, Walter 1966 *The Individual and Society in the Middle Ages*. Baltimore: Johns Hopkins University Press.
- Unger, Roberto Mangabeira 1976 *Law in Modern Society*. New York: Free Press.

LAWRENCE HAZELRIGG

## INDUSTRIAL SOCIOLOGY

The term *industrialization* connotes the development of social organizations, generally in nation-states in which large manufacturing enterprises loom large and in which adjunctive legal institutions supporting laissez-faire philosophies of market relationships and providential public and private services grow with them in tandem. Adam Smith's *Wealth of Nations* became early on, after its appearance in 1776, a kind of founding blueprint for the exploitation of private interests and initiatives in the service of the "general good." Industrial sociological studies accordingly focus on the "causes" or prerequisites for, the correlates of (in family, community, and other social settings), and the consequences of the industrialization process, largely in social systems governed in accord with eighteenth-century liberalism, with its emphases on the economic rights of property owners and their agents.

The key consequences under examination have been the "rights, privileges, and immunities" of parties to labor contracts; the characteristic social

and economic relationships among investors, producers, workers, and consumers in Western societies and, in turn, the circumstances of dependents, communities, and regions; and, finally, the resulting stabilities and changes in the social, economic, and political structures of the nations in which all these persons live.

Those whose research and teaching agendas include several of these broad domains draw very heavily on specialists in “world systems,” economic development, economic determinism, legal institutions, social stratification, public policy, organizations, urban sociology and community studies, labor markets, and economic sociology. A number of issues treated elliptically in these lines of research, especially those addressed by French, German, British, Italian and Japanese scholars, are discussed in relevant articles in this Encyclopedia.

Industrial sociology as a distinguishable if very broad field of study may truly be said to have been born in the late 1770s, so continuous with their forebearers’ interests have been many of the specific subjects of concern of modern investigators. Systematic studies of social organizations in industrial societies date from the long-celebrated if not always carefully read analyses by Adam Smith, who addressed philosophical issues having to do with economic activities and related “moral sentiments” early in his writings and then turned to his widely known and enduring work on the social (macroscopic) and organizational (microscopic) roles in society of capitalists and, more famously, of labor. Smith regarded labor and its effective mobilization and utilization, not the fruits of mercantilism (i.e., the quantity of gold and silver imported from colonies and the exportation of finished goods), as the real sources of what he termed the wealth of nations. Smith’s studies of the superiority of specialized divisions of labor in a pin factory became a template for studies of the personal and social costs and benefits of manufacturing enterprises by scholars of all political colors.

Furthermore, Smith’s analyses and urgings about the benefits of free trade, of constraining the role of government in economic affairs, of the critical roles of increasingly differentiated divisions of labor, and of the emergence of the “factory system” in late-seventeenth and early-eighteenth-century England staked out most of the basic subjects of modern industrial sociologists’ research

studies: economic organizations; managers’ and workers’ ways; the correlates of technology and divisions of labor; economic exchange and trade; questions about the roles of central governments in economies, and about innumerable regulatory measures affecting workplaces; and, finally, the distributive effects of industrialization on national stratification systems.

Industrial sociologists can also trace their roots to Charles Dickens’s very popular literary treatments of life in early-nineteenth-century England and to other early and more pointedly social, economic, and political reformers in Europe. Thus, investigators now undertake studies with debts to English and French critics, commentators, and philosophers whose works are captured in the term “the Enlightenment”; to Henri St. Simon’s urgings about the application of scientific rationality to social organization; to Karl Marx’s historicist treatments of capitalism’s problematical “maturer” states; to the pros and cons in arguments about nascent trade union movements; to debaters arguing over alternative welfare, trade, and other initiatives by governments; and to assessments of the philosophical legitimacy of the emerging stratification of societies, as variant forms of aristocracy gave way to variant forms of more representative governments. Dickens’s characterization of an employers’ association that resisted safety guards on moving machine parts as “The Association for the Mangling of Operatives” and his and his audiences’ concerns about the improvement of working conditions presaged a great deal of the work of industrial sociologists in the twentieth century, most of whom have been critical of unfettered markets. Continuities since the nineteenth century, meantime, in all but methods of research, have been notable. The subjects pursued include the forces that virtually compel ever more differentiated divisions of labor; the effects of different patterns of national income distribution; and the contributions of these forces to the rise, expansion, and characters of both “blue-collar” labor forces and, overlapping with them, the expansion of middle classes in urban centers; the forces that contribute to the rise of trade unions and the resulting growth of new systems of law that have changed the nature of property and property claims to include more than physical capital and realty; the forces that have contributed to the decline in small agricultural holdings, small towns, and rural

areas; and, withal, the forces that generally make social life more secular, political life more democratic or undemocratic, and working life more bureaucratic.

During the period from 1900 to 1950, much of the macro-level research that continued to add to present-day conceptions of the body of industrial sociological literature was actually performed by institutional economists—in the United States principally students of labor, management, and the “legal foundations of capitalism”, in the tradition of the “Wisconsin” or (J.R.) “Commons School” (Commons and associates 1926, 1935, 1936). These investigators also founded the field of industrial relations, whose practitioners’ work has perennially overlapped with that of industrial sociologists. In more recent times, sociologists have essentially inherited what has come to be seen as the “institutional tradition” in economics, a longtime tradition in that discipline that has given way, in the post-World War II world, to mathematical modeling and econometrics. Parallel works in western Europe have similarly been informed by Sydney and Beatrice Webb and by pre-World War I Christian democratic and secular left-liberal traditions in France and Germany.

In the period after World War I, industrial sociologists gave increasing attention to the positive and negative social effects of business cycles: “Busts” undoubtedly helped to purge national economies of many types of waste and inefficiencies and thus contributed to subsequent “booms.” At the same time, however, busts sparked unwelcome recessions and depressions, with their subversive effects on growth, employment levels, and thus on the living standards, attitudes, and behavior of most citizens. Researchers’ attentions, in the 1930s especially, focused correlatively on the growth of large corporations in the Western democracies; the pros and cons of aggressive price competition; and the possibly leavening effects of unions and of interventions by the “positive state” against the centralization of economic and thus, potentially, of political power in private hands, such that some of the less admirable correlates of capitalist systems could be controlled. Generally the researchers sought not so much to traduce “big business” as to identify constraints on private economic power that would not reduce the benefits to societies of large corporations’ productive capacities. The emergence of fascism in Germany in the

1930s embodied widespread apprehensions about both big government and the reach of the communist revolution in what became the USSR.

Systematic attention toward the middle of this period was also given to what is now termed the global economy, often in the form of studies of the effects of the international trade barriers that contributed so significantly to the collapse of the Western industrial economies—the so-called Crash of 1929 and the Great Depression. The political and economic interdependencies among nations—and those among classes of economic actors within them—became increasingly palpable to researchers, who began to see economies more clearly in terms of the structures of individual industries and industry groupings—new chapters in the all-important story of the causes and correlates of the division of labor (Brady 1943). Additionally, unions were gaining in their appeal, beyond skilled tradesmen, to industrial, “mass-production” assembly workers. Among the questions pursued by sociologists were those having to do with the prospects that a coherent working class would emerge in the United States out of the ferment of the pre-World War II era. Incipient “class warfare” could be tamed, if not quashed, by reforms that are collectively referred to nowadays as revisions in the (Jean-Jacques Rousseau’s) “social contract” and by the construction of social “safety nets” of types that are currently being critically assessed in both the United States and in western Europe; renewed and very vigorous industrial price competition has contributed significantly to concerns about expansion of central government “welfare” budgets in competition with funds for private investment.

Finally, by the 1920s and into the 1940s, there were growing concerns about whether Benito Mussolini’s “corporate state” and Adolf Hitler’s Third Reich represented a kind of conspiracy of fascists with Italy’s and Germany’s financial and manufacturing leaders (Neumann 1942) against those, especially in the three Communists Internationals, who were perceived by conservatives to be urging workers to turn to the revolutionary political left. This subject earned much increased attention after World War II; joined by historians’ work in the 1950s, the social scientific literature on the social sources of totalitarian systems has become extensive, one of the recurring questions in this literature being whether the seeds



of “totalitarianism of the right” germinate especially well in “capitalist” soil (Moore 1973). The answers from sociologists’ studies, however, have tended to focus on the fragility of support for democratic institutions in most societies. These studies have analyzed the vulnerabilities of democracies to losses of the confidence of large numbers of citizens victimized by major and long-enduring economic collapses, rather than on markets’ preoccupation with putative contradictions in capitalism per se, as a key reason for the appeal of social movements on the political right. Suffering from traumatic social, political, and economic discontinuities, sociologists reported, could all too easily contribute to the delivery of the loyalties of many—from Rome and Naples in Italy, to Marienthal in Austria, to Berlin in Germany, and to Buenos Aires in Argentina—to demagogues and “scapegoaters”. Indeed, these studies led social psychologists among sociologists to important and widely influential discoveries about scapegoating phenomena that could generate poisonous race and ethnic relations (Adorno et al. 1950).

At what we may call the “mesoscopic”—middle—levels of analyses, industrial sociologists, in loose confederation with the more institutionally oriented scholars among labor economists, turned their attention in the pre-World War II period to the implications of the levels of aggregate demand (and deficiencies thereof) for the structures and experiences of communities before and during the Great Depression [as in Muncie, (“Middletown”), Indiana, in the United States]; to the disinclinations of unemployed Americans to “blame the system” (rather than themselves) for their sad circumstances during the Great Depression; and to the pros and cons of government intervention in national economies and of the construction of the aforementioned “safety nets” for victims of economic downturns.

Not least among sociologists’ interests were those in the stratification of societies, especially respecting the distribution of wealth and income among individuals and families; the behavior of economic elites in local communities; and the causes and correlates of social mobility. In the latter matter, evidence mounted that the economic successes and failures of individuals could not be attributed simply to their ambition, “drive,” and hard work—or their lack of these oft-praised qualities. Indeed, many successes and failures are

grounded in opportunity structures, not the least important of which are individuals’ access to third-party “bonding” agents and access to formal education and training. While most economists have equated their educational achievement, or lack thereof, to earners’ incomes without further ado, sociologists have urged that imperfections in markets enabled American employers to assume that better-educated workers were more productive and to pass on the costs of such an assumption to consumers (Berg 1970). For economists, better-educated workers are paid more because they are more productive, but their measure of productivity is income, not actual output, as a supposedly valid measure of output.

The advent of World War II brought many industrial sociologists to focus, not without a degree of patriotic fervor, on organizational arrangements in relation to the need to heighten productivity gains in what President Franklin D. Roosevelt called the world’s arsenal of democracy. Studies of workers’ groups in industry in the late 1920s, many conducted by applied sociologists linked to academic business schools (at Harvard University especially), had already indicated that most workers were as responsive to work-group “peer pressures” to fix production quotas at rates below those contemplated by managers and by their time-and-motion experts as they were to economic incentives. These findings were replicated many times between 1929 and 1985. This view of workers’ motives continues to inform the urging by many sociologists that we go beyond conventional economic studies of wage administration in seeking to understand the ways of citizens in their workplaces. At issue are the respective roles of communitarian and of individualistic predispositions among workers.

With a few very noble exceptions, and for the majority’s own good reasons, economists conceive of production organizations as essentially unproblematic “givens” into which resources are pumped and from which outputs flow. Sociologists’ lessons about organizations as social systems, which balanced the valuable lessons taught by economists, were widely applied in the electrical appliances industry, the airframes industry, and in automobile, steel, and other industrial settings. The lesson: Identify work groups and win over workers’ groups and their leaders to corporate aims; then reap the benefits of the support of not-so-rugged

individualists in workplaces, who often respond more favorably to group norms than to group-disrupting industrial incentive-type pay plans. As it turned out, these lessons were perhaps applied somewhat more assiduously from 1960 to 1990 in Japan than in the United States, where worker participation in organizational decision making and work reforms have appealed to comparatively fewer employers despite the possible benefits of “humanized” personnel policies. Researchers have discovered that U. S. managers’ interest in worker participation tends to wax when there labor market conditions are “tight”—that is, high demand for workers—and wane when they are “loose.” In Western managers’ theory at least, human resources are factors of production—commodities—and the treatments of these factors are functions more of short-run market conditions than of long-term concern about the qualities of work life (O’Toole 1974). The latter concerns find more expression in collective bargaining relationships and in remedial legislation, with the endorsement and help of many sociologists who are skeptical of managers as patrons, than in initiatives by employers who are skeptical of the benefits of “codeterminative” relations with workers, such as those requiring worker representations on boards of directors under laws in Germany and the Scandinavian countries.

### THE CONTEMPORARY SCENE

In present-day industrial sociology, the main subjects of study have been (1) the increasingly important roles of education and training in shaping Americans’ opportunities (Jencks et al. 1979); (2) the relationships between the significant differences in circumstances between managers and workers in “core” industries and managers and workers in “internal labor markets” on the one hand, and the lower-paid, appreciably less capitalized, smaller, more vulnerable establishments in the “peripheries” of economies on the other (Kalleberg 1983); (3) in the correlates—income and otherwise—of discrimination against women and minorities; and (4) the problems of those who simply are not well integrated into the labor force, the so-called underclass (Wilson 1987). These newer topics have taken their places with continuing studies of work groups; industrial conflict; mobility patterns; the evolving roles of public policies and the state; the politics of income distribution;

and intramural studies of organizations, their decision-making arrangements, and, more recently, their “cultures.”

Studies by industrial sociologists are increasingly comparative in character, as these researchers seek to identify cultural and political factors—such as belief systems and constitutional arrangements, respectively—that influence the effectiveness of different nations’ populations in efforts to mobilize human and other resources, motivate leaders and their human charges, design productive organizations, and make and provide goods and services (Cole 1989; Lincoln and Kalleberg 1990).

Overall, industrial sociologists have contributed to the delineation of options facing leaders in government but less so to enterprises, unions, and urban communities. During the period until 1970, industrial sociologists’ investigations moved in increasingly specialized directions.

One thrust brought a large group of the field’s leaders to concentrate on organizations (Coleman 1982; Stinchcombe 1990; Thompson 1967). In rich elaborations and embellishments on Max Weber’s pioneering work on bureaucracies, sociologists in the United States and western Europe have ventured into “the Japanese factory”; the Tennessee Valley Authority; banks; mental hospitals; British coal mines; a gypsum mine; a state employment agency; schools, prisons and equivalent “total institutions”; a foundry; the U.S. Military Academy; steel mills (in the United States and Europe); German, Soviet, and Czech manufactories; the military establishment; French family firms; social movements; labor unions; merchant ships; American soldiers’ organized experiences in and out of World War II combat; and the YMCA—to mention just a few contexts and populations about which studies were completed (Hall 1987; Perrow 1986).

While the vast body of literature produced by the observers in this disparate array of organizations has received little acclaim in the media, it is a sign of the importance of these investigators’ findings, assessments, and consequent theories that their work is basic to the curricula of the very influential graduate schools of business and management, from Harvard to Berkeley, from Seattle to Miami, and from Maine to Los Angeles—and in Scandinavia; Germany; France; the United Kingdom; Japan; and, by the 1990s, Moscow. The

lessons learned: It is possible to design a great many optional variations on the specific structures of hierarchical organizations, their intramural arrangements, and the “production relations” therein, to meet the exigencies confronting managers and their charges in dealings with each other; with competitors, clients, customers, suppliers, subcontractors, regulators, third-party insurers, and community forces; and with labor market developments.

Industrial sociologists—whether “majors” from colleges or holders of masters and doctoral degrees, some of whom work as consultants or as technicians or managers in corporate settings—offer prescriptions for improving employee morale and marketing programs (from demographic assessments to surveys of customer attitudes); for designing optimal “mixes” of wage and salary schedules with supplementary benefits; for reducing supplementary benefits; for reducing absenteeism and turnover; for productivity “enhancement”; for designing therapeutic (rather than custodial) environments in mental health care agencies; and for constructing occupational safety programs, grievance machinery, and quality control programs.

A second group moved away from these more microscopic studies of organizations to study the social, economic, and political development of whole societies, some in historical terms during the post-World War II era, including India’s, China’s, Italy’s, Japan’s, and Germany’s (actually “redevelopment” in the latter two cases), and the USSR’s systems. Among the lessons were important guidelines to understanding the stabilities of some and the flexibilities of other social-cultural values that gave distinctive national shape to individual countries’ brands of industrialization. It is clear that while the “common denominators” in the paths to both growth and development—the latter a matter of the degree of distributive justice in a society and the former a matter of increases in gross national product—are numerous, there are instructive differences as well (Inkeles and Smith 1974).

Indeed, a consortium of scholars, many of them sociologists studying “industrialization and industrial man” in comparative-international terms, produced well over forty volumes and a great many shorter pieces on the convergences and

divergences among industrial and industrializing nations over the period 1955–1975 (Dunlop et al. 1975; Kerr and Dunlop 1973). These works of scholarship have helped thousands of leaders in governments, large corporations, labor unions, and international agencies to make judgments about investments (both public and private); social, political, and economic policies; and the aptness of designs of organizations in what is now truly a global economy in which nations’ planning efforts are turning, more and more, toward market and away from command economies (Yergin and Stanislaw, 1998). The conclusions at the end of 1990, in a continuing body of research following the preceding twenty-four months of changes in eastern Europe and the USSR, were that the “marketizing” and democratization movements in previously planned economies would assuredly reduce divergences among industrial systems but would by no means eliminate entirely the influence of discrete national cultures in shaping the practices and institutions, from child rearing to legal structures, that help, in turn, to shape social relations in a given nation’s enterprises, as some sociologists have long argued.

A third constituency moved “below” organizational levels to study the dynamics of work groups within organizations, picking up on the work of the previously mentioned human relations school before, during, and after World War II. This group of scholars drew heavily on earlier sociologists’ insights and theories—from Georg Simmel, Charles Cooley, and George Herbert Mead, and especially from the massive number of post-World War II publications of data and analyses from one of the first very-large-sample and sophisticated social scientific surveys of wartime American soldiers (Inkeles 1964). These reports gave abundant corroboration to the findings in industry, by earlier human relations investigators, concerning the critical importance of small groups and their norms in efforts to understand individual attitudes and behavior (Homans 1950). The applications of these findings—in studies of satisfaction, leadership, morale, productivity, grievances, absenteeism, turnover, and incentive systems—have become staples in training programs for supervisors and foremen in the United States and in delineating jobs and designing work flows across American industries (Perrow 1986; Porter et al. 1975). A new group of business school educators in western Europe has

generated very similar programs tailored to take account of the historical and cultural imperatives of managers and workers in different countries. The growing integration of Europe's economies and the mobility of their citizen promises to add to the convergence in the ways and means of financing and directing economic organizations.

A fourth group has focused on industries and occupations as special and highly significant aspects of organizations' "external environments" and as subsets of America's systems of social stratification. Sociological studies of whole industries—their personnel and collective bargaining policies especially—have often informed public agencies' regulations, legislators' bills, and judges' decisions. Analyses of differences among industries and occupations have helped leaders in government, business, and labor unions to understand better the dynamics of industrial conflict; the character and effectiveness of organizations; and the complexities in identifying the effects of physical technology—capital—in the spinning of "webs of rules" (Kerr et al. 1973) that, for all of their *de facto* and even informal character, function very much like governance systems in the workaday world (Kalleberg and Berg 1987). In their most formal states, these systems—arbitration procedures, for example—sometime mature into what the U. S. Supreme Court in 1960, in a "trilogy" of cases that defined the role of arbitration in industrial relations, called "systems of industrial common law" (i.e., as legal systems virtually unto themselves). Otherwise, more implicitly, as with work rules that establish "how fast is fast, how fair is fair, and how reasonable is reasonable", the webs of rules define relationships and codes applicable to both employers and employees that afford a kind of lubrication to the mechanics of human interactions in bureaucratic machines, with their close tolerances, involving millions of persons in hundreds of thousands of workplaces. Sociological studies of the costs and benefits to employees and employers (and ultimately to the public) of work rules, for example, have helped transform emotionally charged arguments about "featherbedding", "soldiering," and "goldbricking" (all efforts to escape irksome chores) into coherent and constructive debates about nonmonetary dimensions of working conditions. Studies of work rules suggest that "informal organizations" within parent organizations are really not so much informal as they are what Durkheim

long ago called "the noncontractual element of contracts"; sociologists have demonstrated that though these patterned, enduring, and bilaterally honored arrangements do not appear on an organization's wall charts, they are significant components of organizational life unto themselves, not mere shadows of more familiar and more palpable structures.

Meanwhile, the discoveries in international comparisons of data on grievances and strikes—that there are numerous short strikes and many grievances in the United States and few but long strikes and virtually no grievance procedures in western Europe for example—have led researchers, employers, and union leaders to appreciate the value of expeditious "on-site" bargaining relationships, on a day-to-day basis, such that emotional affect in disagreements may be drawn off and tensions relaxed before out-and-out conflicts disrupt production and social relations. It seems clear that these day-to-day adjustments and accommodations occur even as the efficacy of unions has declined in the United States and will likely wane in the "Euro" countries.

At the same time, the costs as well as the benefits of federal laws requiring that unions be democratic have helped us to make more realistic estimates about democratic arrangements' capacities to function as panaceas; democracy, for example, offers no guarantees against corruption, nor does it assure harmonious relations between parties to collective bargaining agreements. Sociologists have also documented a kind of (perhaps understandable) hypocrisy regarding democracy: Many lay observers and most labor columnists are delighted by unionists who vote to ratify contracts or to "decertify" their bargaining agents but are appalled by "strike votes."

Still another group of specialists have concentrated their attentions on worker satisfactions, dissatisfactions, and work experiences, by use of survey research designs that involve both periodic "snapshots" of different working Americans (and Japanese and west Europeans) and repeated observations of these same respondents, in "panel studies," over long time periods. These designs also make it possible to study "cohort effects," that is, the effects of reaching a given age in different time periods, each with their different qualities regarding a variety of social realities (Quinn et al.

1974). Thus there are significant differences, for example, in the experiences (and their attitudes about them) of workers, depending on whether they entered the labor force in 1960, 1970, 1980, or 1990. Sociologists can accordingly raise thoughtful questions about the implications, for public and private policies, of changing *social* definitions of aging, for example, in juxtaposition and contrast with essentially arbitrary public *policies* that fix eligibilities for a number of services and benefits on the basis of the chronological ages of individuals; not all those now 65 years old think, act, or want to be treated as a homogeneous class of senior citizens, nor have they had the same life histories, on average, as those who reached that age in 1940, 1960, or 1980.

Finally, the advent and continuing engagements of the civil rights movements have sparked the expenditure of a great deal of research effort on the comparative socioeconomic opportunities of men and women, and of minority group members in these groups (Jacobs 1989; Jaynes and Williams 1989). The findings by sociologists in these investigations have figured prominently in the drafting of legislation, legal suits, and employment policies, as well as in landmark civil rights decisions in courts at all levels.

In their work, as noted at the outset, a shrinking population of traditional industrial sociologists have drawn on work by sociologists in virtually every one of the profession's own major areas of interest with alacrity and have, in turn, seen much of their work inform the work of these other specialists. There has been similar intellectual commerce with social psychologists, industrial relations practitioners, and with full-time nonacademic social scientist practitioners in private enterprises, public agencies, universities, and research and other organizations in foreign lands, especially in the United Kingdom; Yugoslavia; Germany; Japan; France; Scandinavia; Canada; Italy; and what has, for so long, been called eastern Europe (Adams 1991; Barr, 1994; Freeman, 1994).

Industrial sociologists will likely grow in numbers in the years ahead, despite the growth of service sectors and information technology across the globe, as the global economy's structures will tend to be more related to multilateral regional pacts among nations with shared currencies, laws, and macroeconomic policies that will reshape the

political, economic, and social characters of member nations.

## REFERENCES

- Adams, Roy J., ed. 1991 *Comparative Industrial Relations: Contemporary Research and Theory*. London: HarperCollins Academic.
- Adorno, T. W., et al. 1950 *The Authoritarian Personality*. New York: Harper.
- Barr, Nicholas, ed. 1994 *Labor Markets and Social Policy in Central and Eastern Europe*. New York: Oxford University Press.
- Berg, Ivar 1970 *Education and Jobs: The Great Training Robbery*. New York: Praeger.
- Brady, Robert A. 1943 *Business as a System of Power*. New York: Columbia University Press.
- Cole, Robert E. 1989 *Strategies for Learning: Small Group Activities in American, Japanese, and Swedish Industry*. Berkeley: University of California Press.
- Coleman, James 1982 *The Asymmetric Society*. Syracuse, N.Y.: Syracuse University Press.
- Commons, John R., et al. 1926, 1935, 1936 *History of Labor in the United States*. New York: Macmillan.
- Dunlop, John, et al. 1975 *Industrialism and Industrial Man Reconsidered*. Princeton, N.J.: Inter-University Study of Human Resources in National Development.
- Freeman, Richard B., ed. 1994 *Working Under Different Rules*. New York: Russell Sage Foundation.
- Hall, Richard 1987 *Organizations: Structures, Process and Outcomes*. Englewood Cliffs, N.J.: Prentice-Hall.
- Homans, George 1950 *The Human Group*. New York: Harcourt, Brace.
- Inkeles, Alex 1964 *What is Sociology?* Englewood Cliffs, N.J.: Prentice-Hall.
- , and David H. Smith 1974 *Becoming Modern: Industrial Change in Six Developing Countries*. Cambridge, Mass.: Harvard University Press.
- Jacobs, Jerry 1989 *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, Calif.: Stanford University Press.
- Jaynes, Gerald David, and Robin M. Williams, Jr., eds. 1989 *A Common Destiny: Blacks and American Society*. Washington, D.C.: National Academy Press.
- Jencks, Christopher, et al. 1979 *Who Gets Ahead? Determinants of Economic Success in America*. New York: Basic Books.
- Kalleberg, Arne, ed. 1983 "Capital Labor and Work: Determinants of Work-Related Inequalities." (special issue). *International Sociology Journal*. August.

- , and Ivar Berg 1987 *Work and Industry Structures, Markets, and Processes*. New York: Plenum.
- Kerr, Clark, John T. Dunlop, Fredrick Harbison, and Charles A. Meyers 1973 *Industrialization and Industrial Man*. London: Penguin.
- Lincoln, James R., and Arne L. Kalleberg 1990 *Culture, Control and Commitment: A Study of Work Organization and Work Attitude in the U. S. and Japan*. Cambridge, U.K.: Cambridge University Press.
- Moore, Barrington 1973 *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Neumann, Franz 1942 *Behemoth: The Structure and Practice of National Socialism*. New York: Oxford University Press.
- O'Toole, James, ed. 1974 *Work and the Quality of Life: Resource Papers for Work in America*. Cambridge, Mass.: MIT Press.
- Perrow, Charles 1986 *Complex Organizations: A Critical Essay*. New York: McGraw-Hill.
- Porter, Lyman W., et al. 1975 *Behavior in Organizations*. New York: McGraw-Hill.
- Quinn, Robert P., et al. 1974 *Job Satisfaction: Is There a Trend?* Monograph 30. U.S. Department of Labor, Washington D.C.: U.S. Government Printing Office.
- Stinchcombe, Arthur L. 1990 *Information and Organization*. Berkeley: University of California Press.
- Thompson, James D. 1967 *Organizations in Action*. New York: McGraw-Hill.
- Wilson, William J. 1987 *The Truly Disadvantaged: The Inner City, the Underclass and Public Policy*. Chicago: University of Chicago Press.
- Yergin, Daniel, Joseph Stanislaw 1998 *The Commanding Heights*. New York: Simon and Schuster.

IVAR BERG

## INDUSTRIALIZATION IN LESS DEVELOPED COUNTRIES

In the two and a half centuries since the Industrial Revolution in England, the process of industrialization has perhaps had more impact on all the nations of the world than any other complex set of forces. This process has not been uniformly introduced in all countries, nor has it occurred at the same time or at the same rate. Despite the common features of industrialization, these differences in its introduction and adoption have produced

inequities among nations and among people on a scale never before experienced.

In describing various countries and regions of the world, certain terms have been adopted, first by official agencies such as the United Nations and national governments, and then more generally by scholars, journalists, and those interested in making sense of international relations. According to a now commonly used United Nations classification, *more developed countries* (MDCs, or developed countries) comprise all of Europe, North America (excluding Mexico), Japan, Australia, and New Zealand. Other countries (e.g., Singapore, Taiwan, and Israel) constitute recent additions, while many of the former Soviet-bloc countries (including the Russian Federation) are now in a developing, or "transition," phase. *Less developed countries* (LDCs, or developing countries) make up the remainder. The distinction between MDCs and LDCs mirrors the famous "North-South divide," a phrase coined by former West German chancellor Willy Brandt (1980) in his Commission's report to the World Bank. LDCs have also been referred to as the Third World, a term devised in post-World War II Europe to distinguish the politically nonaligned, underdeveloped nations of the world from the industrialized capitalist nations (First World) and the industrialized communist countries (Second World) (Worsley 1984, pp. 306-315).

In some cases, the underlying variable upon which these distinctions are based is economic, in other cases it is political, and in still others it is unspecified. However, generally speaking, MDCs are "rich" and LDCs are "poor." In 1996, the per capita gross national product (GNP) among all MDCs was US\$25,870, while in the LDCs it was only US\$1,183 (World Bank 1998, p. 38). The major explanation for this vast discrepancy is that MDCs are fully industrialized whereas LDCs are not. In 1994, the industrial market economies produced 81.4 percent of total world manufactures (World Bank 1997, p. 152). Considering that LDCs comprise 84 percent of the world's population (World Bank 1997, p. 36), their industrial output and, consequently, their standard of living are dramatically lower than in MDCs.

### INDUSTRIALIZATION DEFINED

Industrialization is a complex process comprised of a number of interrelated dimensions (Hedley

1992, pp. 128–132). Historically, it represents a transition from an economy based on agriculture to one in which manufacturing represents the principal means of subsistence. Consequently, two dimensions of industrialization are the work that people do for a living (economic activity) and the actual goods they produce (economic output). Other dimensions include the manner in which economic activity is organized (organization), the energy or power source used (mechanization), and the systematic methods and innovative practices employed to accomplish work (technology). Table 1 specifies these dimensions and also lists indicators commonly used to measure them.

According to these indicators, MDCs are fully industrialized. On average, close to one-third of the labor forces in these countries are employed in industry (three-fifths work in the service sector); manufacturing makes up approximately one-quarter of the gross domestic product; the overwhelming majority of workers are employees of organizations; commercial energy consumption is high (5,100 kilograms of oil equivalent per capita); and professional and technical workers comprise on average 15 percent of the work force. Furthermore, more than 95 percent of all receipts for royalty and license fees are collected in the MDCs (Hedley 1992, pp. 128–133; United Nations Development Programme [UNDP] 1998; World Bank 1997). Industrial activity and the services associated with it constitute the major driving force and source of income in these more developed economies.

In contrast, none of the LDCs is fully industrialized as measured by these five dimensions of industrialization. Although manufacturing accounts for a significant proportion of many of these countries' total output, most do not achieve industrial status on any of the other dimensions. Manufacturing in these countries is accomplished largely by traditional methods that have varied little over successive generations. Consequently, although manufacturing (transforming raw materials into finished goods) is an essential component of industrialization, there is considerably more to the process. Because industrialization is multidimensional, it cannot be measured by only one indicator.

In general, LDCs may be classified into three major groups according to how industrialized they are. The first and smallest group, referred to as

*newly industrializing countries* (NICs), contains the most industrialized countries in that they achieve industrial status on at least two dimensions listed in Table 1. Located mainly in East Asia (e.g., South Korea, Malaysia, Thailand, Indonesia) and Latin America (e.g., Mexico, Brazil, Argentina, Venezuela), these eight NICs accounted for more than 40 percent of all merchandise exports from developing countries in 1995 (World Bank 1997, pp. 158–160). Although China and, to a lesser degree, India (because of their huge population bases) contribute significantly to the merchandise exports of LDCs, they have not developed their industrial infrastructures to the same extent as these NICs and therefore do not belong in the most industrialized group of LDCs.

A subgroup of NICs are high-income, oil-exporting nations (e.g., the United Arab Emirates, Qatar, Bahrain, Kuwait, and Saudi Arabia). Although they do not have large manufacturing bases, they do have significant proportions of their labor forces involved in industry (oil exploration and refining), a substantial component of professional and technical workers (many of them imported), and high per capita commercial energy consumption (World Bank 1998, pp. 34–35, 42–43). Although concentrated in just one industry, they are more industrialized than most other LDCs according to the criteria specified in Table 1. As a result of their petrodollars, they have acquired an industrial infrastructure that in other countries has taken many decades to establish.

The second, very large group of LDCs in terms of industrialization are those with a traditionally strong manufacturing base that also have a substantial agricultural component. Their economies straddle the agricultural and industrial modes of production. China and India are in this group, as are most of the non-European nations that form the Mediterranean basin. The goods that these LDCs predominantly manufacture (e.g., apparel, footwear, textiles, and consumer electronics) are essential to their own domestic markets and, because they are labor-intensive, also compete very well in the international market. In addition, they export natural resources and agricultural products. Other countries included in this semi-industrial group are most of the nations in Central and South America as well as many in South and East Asia.

### Dimensions and Measures of Industrialization

#### 1. Economic Activity

- a. Percentage of labor force in manufacturing
- b. Percentage of labor force in industry

#### 2. Economic Output

- a. Manufacturing as a percentage of gross domestic product (GPD)
- b. Industry as a percentage of gross domestic product
- c. Gross output per employee in manufacturing
- d. Earnings per employee in manufacturing

#### 3. Organization

- a. Wage and salary earners as a percentage of the labor force
- b. Number of manufacturing establishments employing fifty or more workers per capita

#### 4. Mechanization

- a. Commercial energy consumption per capita
- b. Total cost of fuels and electrical energy per employee in manufacturing

#### 5. Technology

- a. Percentage of professional and technical workers in labor force
- b. Registered patents in force per capita
- c. Registered industrial designs in force per capita

Table 1

The third and final group of LDCs are not industrialized on any of the five dimensions listed in Table 1. On average, less than 10 percent of their labor forces are employed in industry; most (76 percent) work in agriculture. Manufacturing contributes only 20 percent to their national economies; the bulk of income derives from natural resources and cash crops grown exclusively for export. Per capita gross national product is very low (US\$215). Most of these nonindustrial LDCs are located in sub-Saharan Africa and Asia (UNDP 1998).

Of these groups of LDCs, the semi-industrial cluster of nations is by far the largest, constituting just over half the world's population. China and India alone make up two-thirds of this group. The second-largest group, comprising between 10 and 15 percent of the world population, is the nonindustrial countries; NICs (including high-income oil exporters) comprise less than 10 percent. Thus, approximately one-quarter of the world is fully industrialized, another 10 percent are industrializing, half are semi-industrial, and the remaining 15 percent are nonindustrial.

### CORRELATES OF INDUSTRIALIZATION

Research has demonstrated that industrialization is directly related to national and individual income, urbanization, the development of an

infrastructure (e.g., communication and transportation networks, education, and health and welfare programs), and the overall quality of life (Hedley 1992, pp. 133–146). These relationships occur because industrialization results in huge productivity gains, which in turn raise individual and national income. For example, when Britain was industrializing, total national income increased by more than 600 percent between 1801 and 1901 (Mitchell 1962, p. 366). In 1850, workers in industrial nations earned eleven times more than their counterparts in nonindustrial countries. Moreover, the advantages of industrialization have been cumulative: today per capita income is more than fifty-two times greater in developed than less developed countries (World Bank 1995, p. 53), thus magnifying what one author has termed a *Global Rift* (Stavrianos 1981). For a variety of reasons, the direct relationship between industrialization and income is continuing to increase.

In an annual survey of all countries in the world, the United Nations Development Programme (UNDP) measures “human development” based on the combination of three, admittedly crude, criteria: longevity, knowledge, and a decent standard of living. Table 2 provides a summary of its most recent results grouped by level of industrialization. Of the 174 countries examined, MDCs consistently scored at the top of all three dimensions that comprise human development, while LDCs, particularly the *least* developed (nonindustrial)



countries, scored at the bottom of the scales. The composite Human Development Index ranged from a high of 0.960 for Canada to a low of 0.185 for Sierra Leone.

According to the UNDP (1998, p. 2):

*Well over a billion people are deprived of basic consumption needs. Of the 4.4 billion people in developing countries, nearly three-fifths lack basic sanitation. Almost a third have no access to clean water. A quarter do not have adequate housing. A fifth of children do not attend school to grade 5. About a fifth do not have enough dietary energy and protein. Micronutrient deficiencies are even more widespread. Worldwide, 2 billion people are anaemic, including 55 million in industrial countries. In developing countries only a privileged minority has motorized transport, telecommunications and modern energy. Inequalities in consumption are stark. Globally, the 20% of the world's people in the highest-income countries account for 86% of total private consumption expenditures—the poorest 20% a minuscule 1.3%. More specifically, the richest fifth:*

- *Consume 45% of all meat and fish, the poorest fifth 5%.*
- *Consume 58% of total energy, the poorest fifth less than 4%.*
- *Have 74% of all telephone lines, the poorest fifth 1.5%.*
- *Consume 84% of all paper, the poorest fifth 1.1%.*
- *Own 87% of the world's vehicle fleet, the poorest fifth less than 1%.*

Although annual GNP growth rates were higher overall in LDCs (4.1 percent) than MDCs (2.2 percent) between 1980 and 1995 (UNDP 1998, p. 210), the *least* developed of these LDCs did not fare so well. Not only were their GNP growth rates (2.1 percent) lower than those of MDCs, they were also significantly below their 1970–1995 average annual population growth rate (2.6 percent) (UNDP 1998, p. 209), which means that these countries are actually falling behind in what little progress they have made (Estes 1988). For example, of the 50 countries that comprise sub-Saharan Africa

(Cunningham 1998), two-thirds have been evaluated as “least developed,” that is, targeted for priority international development assistance by the Nations Nations (UNDP 1998, p. 226). Consequently, some experts have openly questioned whether these countries will ever reach the level of industrialization and quality of life now enjoyed in the capitalist countries of the North (South Commission 1990, p. 19).

#### **DEVELOPMENT GOALS: A VIEW FROM THE SOUTH**

World War II marked the end of one era and the beginning of another. Among the more notable turning points following the war were the establishment of the United Nations (and its subsequent Universal Declaration of Human Rights), the onset of the Cold War between East and West, the emergence of many new independent states following the end of colonial rule, and the sudden realization that the income and development gaps between MDCs and LDCs were growing at a precipitous rate. In turn, these events sparked the generation of theories on international development (see, for example, *Modernization Theory*, *Convergence Theories*, *Dependency Theory*, *Global Systems Analysis*), the creation of organizations such as the International Monetary Fund (IMF) and the World Bank to design an open and stable global monetary system and to establish development and investment programs (IMF 1985), and the formation of various world commissions to address issues of global development (e.g., Brandt Report 1980; Pearson Report 1969; World Commission on Environment and Development 1987).

These activities were initiated largely in MDCs by Northern-based scholars and practitioners. Although there have been several prominent writers from the South who have viewed relations between MDCs and LDCs from a Southern perspective, (Cardoso and Faletto [1969] 1979; Said 1993; Sen 1990, 1992); it was not until 1990 that delegates from the South expressed their concerns in a representative and comprehensive manner. The South Commission (1990), initially formed in 1987, was comprised of twenty-eight members from twenty-six LDCs located on all continents of the South. Chaired by Julius K. Nyerere, former president of Tanzania, the South Commission “has its origins in a recognition within the South that developing

**Human Development Index (HDI) by Level of Industrialization (1995)**

HDI Measures	Industrial Countries <sup>1</sup> (N = 50)	Developing Countries (N = 124)	Least Developed Countries <sup>2</sup> (N = 43)	World (N = 174)
Life expectancy at birth (years)	74.2	62.2	51.2	63.6
Adult literacy rate (%)	98.6	70.4	49.2	77.6
Combined 1st; 2nd; and 3rd-level gross enrolment ratio (%)	82.8	57.5	36.4	61.6
Real GDP per capita (purchasing power parity US\$)	16,337	3,068	1,008	5,990
Human Development Index	0.91	0.54	0.34	0.77

Table 2

SOURCE: Adapted from United Nations Development Programme (1998): p. 130.

NOTE: <sup>1</sup>Including the former Soviet-bloc nations.

NOTE: <sup>2</sup>Designated by the United Nations as LDCs targeted for priority international development assistance.

countries have many problems and much experience in common, but that no one in the South was responsible for looking at these things in a comprehensive manner, or at the lessons about appropriate development strategies which could be drawn from them” (Nyerere 1990, p. v). Consequently, *The Challenge to the South* (South Commission 1990) is an attempt to articulate from a Southern perspective what needs to be done “to help the peoples and governments of the South to be more effective in overcoming their numerous problems, in achieving their ambition of developing their countries in freedom, and in improving the lives and living conditions of their peoples” (Nyerere 1990, p. vi).

As part of its initial mandate, the South Commission (1990) first defined its development goals:

*Development is a process of self-reliant growth, achieved through the participation of the people acting in their own interests as they see them, and under their own control. Its first objective must be to end poverty, provide productive employment, and satisfy the basic needs of all the people, any surplus being fairly shared. This implies that basic goods and services such as food and shelter, basic education and health facilities, and clean water must be available to all. In addition, development presupposes a democratic structure of government, together with its supporting individual freedoms of speech, organization,*

*and publication, as well as a system of justice which protects all the people from actions inconsistent with just laws that are known and publicly accepted.* (p. 13–14)

To achieve these development goals, the South Commission (1990) identified three interrelated task areas:

1. *National self-reliance.* Individual countries of the South must realize their own unique potential through united and sustained efforts on the part of *all* people working toward clearly defined interim and long-term objectives. However, because there are external factors (e.g., the current structure of the global economy) that impede progress, solidarity among nations of the South is crucial.
2. *South–South cooperation.* “By joint endeavours to use to the maximum their different resources of expertise, capital, or markets, all would be able to address their separate and differing needs more effectively, thereby widening their development options . . . By exploiting these openings for co-operation, the South as a group can also become stronger in its negotiations with the North” (p. 16).
3. *An organized South for meaningful North–South negotiations.* In their relations with the North, the South must “establish common priorities in keeping with the

development interests of all,” “share technical and negotiating expertise,” and “hold constructive South–South discussions in advance of negotiations” (p. 21). Also, the South should actively support the growing initiative to establish international regulatory frameworks for the enforcement of global economic relations (including finance and trade) that are in the best interests of *all* nations and *all* peoples.

From this brief review of the report of the South Commission, it is possible to identify several key features of sustained development, *as perceived by those affected*:

- The South must play an active role in its own self-reliant development,
- Democratic grassroots participation is essential,
- Cooperation must occur at several levels (e.g., community, nation, and region),
- International regulatory codes must be renegotiated between North and South.

Although it is too soon to assess what results the recommendations of the South Commission will achieve, there are indications that some progress is being made. For example, President James Wolfensohn of the World Bank proposed a *Challenge of Inclusion* (1997) during a recent annual meeting of the Bank. Wolfensohn’s “challenge” is similar to the objectives of the South Commission: “Our goal must be to reduce . . . disparities across and within countries, to bring more and more people into the economic mainstream, to promote equitable access to the benefits of development regardless of nationality, race, or gender” (p. 6). Moreover, the key elements of his challenge also resemble those offered by the South Commission:

- First and foremost, the governments and the people of developing countries must be in the driver’s seat—exercising choice and setting their own objectives for themselves . . .
- Second, our partnership must be inclusive—involving bilaterals and multilaterals, the United Nations, the European Union, regional organizations, the World Trade

Organization, labor organizations, NGOs [nongovernmental organizations], foundations, and the private sector . . .

- Third, we should offer our assistance to all countries in need. But we must be selective in how we use our resources. There is no escaping the hard fact: More people will be lifted out of poverty if we concentrate our assistance on countries with good policies than if we allocate it irrespective of the policies pursued . . .
- Finally, all of us in the development community must look at our strategies anew (Wolfensohn 1997, pp. 9–11).

Whether *The Challenge to the South* and *The Challenge of Inclusion* represent realistic steps toward meaningful North–South dialogue, cooperation, and beneficial action remains to be seen. However, if we do not tear down the many institutionalized structures of exclusion that are currently in place, all of us may eventually lose.

#### A NEW TECHNOLOGICAL REVOLUTION

Another set of challenges confronting LDCs is the new information and communications technology (ICT) revolution. Beginning in the late 1960s with breakthroughs in microelectronics and fiber-optic transmission (Diebold 1990; Gilder 1989), the ICT revolution is still very much in its genesis and is limited primarily to the developed countries (Hedley 1998). In 1994, just five G-7 nations (the United States, Japan, Germany, France, and United Kingdom) accounted for 80 percent of the information technology market, and American and Japanese corporations dominate the industry (Organization for Economic Co-Operation and Development [OECD] 1996, pp. 7, 37). Table 3 presents the stark facts of information communication by both traditional and modern means. Particularly with regard to telephones and personal computers, essential ingredients for electronic communication, the less developed, low- and middle-income countries are barely represented. In sub-Saharan African countries, for example, 51 percent of all telephones are located in the largest cities (7 percent in MDCs), and the average waiting time for a telephone connection is 15.2 years (Aggor 1998, p. 9). Inadequate infrastructure, low purchasing power, lack

## Information Communication by National Level of Income

Means of Information Communication	High-Income Countries	Middle-Income Countries	Low-Income Countries	World
Daily newspapers per 1,000 people (1994)	303	62	12	98
Television sets/1,000 (1996)	611	252	47	211
Telephone main lines/1,000 (1996)	540	78	11	133
Mobile telephones/1,000 (1996)	131	8	0	28
Fax machines/1,000 (1995)	33.7	1.5	0.2	6.3
Personal computers/1,000 (1996)	224	2	..	50
Internet hosts/10,000 (July 1997)	203.5	2.4	0.1	34.8

Table 3

SOURCE: World Bank (1997): p. 286; World Bank (1999): p. 227.

of technological skills, and poor competition and regulation mean that particularly in the least developed countries, it will take years to realize much benefit from the ICT revolution (World Bank 1999, pp. 62–64). By some measures, the knowledge and information gap between MDCs and LDCs is larger than the income gap.

Various development scenarios have been proposed in anticipation of the impact of a full-scale ICT revolution (Hedley 1999; Howkins and Valantin 1997). The basic question addressed is whether this revolution will “help level the international playing field in terms of opportunities for social and economic development” or whether it will “lead to increasing disparities in incomes and information access” (Baranshamaje et al. 1995). In order to address this question, a high-powered workshop comprised of twenty-seven experts in technology and development from five MDCs and eight LDCs was convened in 1996 (Howkins and Valantin 1997). Sponsored by the U.N. Commission on Science and Technology for Development and the Canadian-based International Development Research Centre, the delegates attempted to establish the parameters of worldwide ICT development. Four possibilities emerged, based on the inclusivity and openness of the global community and the proactive or reactive responses of individual countries (Howkins and Valantin):

1. *The march of follies*. “The global community is exclusive and fragmented,” and “most developing countries respond only partially and reactively to acquisition and use of ICTs.” In this scenario, the global rift between MDCs and LDCs widens.

2. *Cargo cult*. “The global community is inclusive and supportive,” but “most developing countries respond only partially and reactively to the acquisition and use of ICTs.” This scenario results in cultural imperialism in which a few transnational corporations dominate the ICT industry. Developing countries have little input.
3. *Netblobs*. “The global system is exclusive and fragmented”; however “developing countries take an active approach to the acquisition and use of ICTs and develop a complete set of policies.” This scenario produces regional groups or blocs based on shared cultures, religions, and languages. “At the end of the scenario period, the blocs have achieved much. They have created information societies and economies that reflect their own histories, traditions, cultures, and ways of doing business. But their insistence on their own regional laws, regulations, and trading principles creates centripetal forces that lead to a highly unstable situation.”
4. *Networld*. “The global community is inclusive and supportive,” and “developing countries have a complete and proactive set of policies toward the acquisition and use of ICTs.” This is a global village scenario in which corporations from developed countries operate out of “enlightened self-interest as they seek ways of working with companies and institutions in the developing world . . . Their awareness is matched by a realization in

developing countries that they should work with global corporations to create their own national information society and economy.”

The workshop delegates did not attempt to predict which of these four development scenarios was most probable. Rather, based on their two sets of parameters, they indicated the various possibilities, what would be required to reach each, and some of the development consequences that would likely flow from them. In other words, if the global community is not inclusive and supportive, and if developing countries do not develop a complete and proactive set of policies toward the acquisition and use of ICTs, then we may expect various consequences that are inimical to comprehensive and sustained development worldwide. As Herbert Simon (1987, p. 11) has observed:

*Technological revolutions are not something that “happen” to us. We make them, and we make them for better or for worse. Our task is not to peer into the future to see what computers will bring us, but to shape the future we want to have—a future that will create new possibilities for human learning, including, perhaps most important of all, new possibilities for learning to understand ourselves.*

Thus, if we wish to take up *The Challenge of Inclusion*, there are certain courses of action to follow. If we do not take up this challenge, then, according to the founding editor of *Scientific American*, the very survival of our species is at stake (Piel 1992).

## REFERENCES

- Aggor, Adolphine Y. 1998 *African Development via the Internet: A New Wave or More of Same?* Master's thesis, University of Victoria, Canada.
- Baranshamaje, Etienne, et al. 1995 *Increasing Internet Connectivity in Sub-Saharan Africa: Issues, Options, and World Bank Group Role*. Available on the Internet at <http://www.worldbank.org/aftdr/connect/incrint.htm>.
- Brandt Report 1980 *North-South: A Program for Survival*. Report of the Independent Commission on International Development Issues. Cambridge, Mass.: MIT Press.
- Cardoso, F. H., and E. Faletto (1969) 1979 *Dependency and Development in Latin America*. Berkeley: University of California Press.
- Cunningham, R. 1998 *Sub-Sahara Africa: Links to Sites by Country*. Available on the Internet at <http://www.staff.uiuc.edu/~rcunning/ssa.htm>.
- Diebold, John 1990 “The Fiber-Optics Breakthrough: The Communications Society.” In John Diebold, ed., *The Innovators: The Discoveries, Inventions, and Breakthroughs of Our Time*. New York: Truman Talley Books/Plume.
- Estes, Richard J. 1988 *Trends in World Social Development: The Social Progress of Nations, 1970–1987*. New York: Praeger.
- Gilder, George 1989 *Microcosm: The Quantum Revolution in Economics and Technology*. New York: Simon and Schuster.
- Hedley, R. Alan 1992 *Making a Living: Technology and Change*. New York: HarperCollins.
- 1998 “Technological Diffusion or Cultural Imperialism? Measuring the Information Revolution.” *International Journal of Comparative Sociology* 39(2):198–212.
- 1999 “The Information Age: Apathoid, Cultural Imperialism, or Global Village?” *Social Science Computer Review* 17(1):78–87.
- Howkins, John, and Robert Valantin (eds.) 1997 *Development and the Information Age: Four Global Scenarios for the Future of Information and Communication Technology*. Ottawa: International Development Research Centre and the United Nations Commission on Science and Technology for Development. Available on the Internet at <http://www.idrc.ca/books/835/index.html>.
- International Monetary Fund (IMF) 1985 *The Role and Function of the International Monetary Fund*. Washington, D.C.: International Monetary Fund.
- Mitchell, B. R. 1962 *Abstract of British Historical Statistics*. Cambridge, U.K.: Cambridge University Press.
- Nyerere, Julius K. 1990 “Chairman's Preface.” In South Commission, *The Challenge to the South: The Report of the South Commission*. Oxford, U.K.: Oxford University Press.
- Organization for Economic Co-operation and Development (OECD) 1996 *Information Technology Outlook 1995*. Paris: Organisation for Economic Co-operation and Development.
- Pearson Report 1969 *Partners in Development: Report of the Commission on International Development*. New York: Praeger.
- Piel, Gerard 1992 *Only One World: Our Own to Make and Keep*. New York: Freeman.
- Said, Edward W. 1993 *Culture and Imperialism*. New York: Knopf.

- Sen, Amartya K. 1990 *The Political Economy of Hunger*. Oxford, U.K.: Clarendon Press.
- 1992 *Inequality Reexamined*. New York: Russell Sage Foundation.
- Simon, Herbert A. 1987 “The Steam Engine and the Computer: What Makes Technology Revolutionary?” *Computers and People* 36(11–12):7–11.
- South Commission 1990 *The Challenge to the South: The Report of the South Commission*. Oxford, U.K.: Oxford University Press.
- Stavrianos, L. S. 1981 *Global Rift: The Third World Comes of Age*. New York: William Morrow.
- United Nations Development Programme (UNDP) 1998 *Human Development Report 1998*. New York: Oxford University Press.
- Wolfensohn, James D. 1997 “The Challenge of Inclusion.” Address to the Board of Governors, Hong Kong, China, September 23, 1997.
- World Bank 1995 *World Development Report 1995*. Oxford, U.K.: Oxford University Press.
- 1997 *1997 World Development Indicators*. Washington, D.C.: International Bank for Reconstruction and Development, World Bank.
- 1998 *1998 World Bank Atlas*. Washington, D.C.: International Bank for Reconstruction and Development, World Bank.
- 1999 *World Development Report: Knowledge for Development*. Oxford, U.K.: Oxford University Press.
- World Commission on Environment and Development 1987 *Our Common Future*. Oxford, U.K.: Oxford University Press.
- Worsley, Peter 1984 *The Three Worlds: Culture and World Development*. London: Weidenfeld and Nicolson.

R. ALAN HEDLEY

## INFANT AND CHILD MORTALITY

Mortality affects the volume of a population. Deaths are not equally distributed among all groups; rather many unique patterns have been identified. For example, the probability of dying is high among both extremes of a population's age structure—the very young and the very old. As a general pattern, the death rate is relatively high at age zero, reaches a minimum in the range from ages ten to fifteen, and then begins to increase gradually with increasing age. This increase becomes marked after age forty-five or fifty (Coale 1965).

Infant and child mortality are important because the largest mortality risk differentials between a society with high mortality and one with low mortality are always found within infancy and childhood. Because infant and child mortality are often related to general levels of health and living conditions, they are often thought of as measures by which nations can gauge their current level of socioeconomic development and societal cohesiveness. Comparisons can then be made with past mortality estimates, along with those of neighboring nations, allowing a country to make projections concerning its future development and public policy programs. Assuming this general precept, researchers investigate historical mortality trends in an attempt to isolate general patterns of infant and child mortality. The goal is to provide information for policy makers, raise standards of living for residents of a nation, thereby reducing early childhood mortality.

### MEASURES AND DATA SOURCES

The *infant mortality rate* (IMR) is defined by demographers as the probability that a newborn infant will die before it reaches its first birthday. This probability is generally computed by calculating the ratio of the number of infant deaths under one year of age in a given year per number of births in that given base year. This approach is termed a *period* IMR because it only utilizes data from one year to arrive at its estimate. This method is not entirely accurate in that it 1) utilizes infant deaths of those born in the preceding year and 2) misses those infant deaths that occur in the following year. If the number of births and deaths does not change drastically from year to year, then this method results in a close approximation of an infant's probability of survival to age one (Shryock and Siegel 1976).

A more accurate measure of infant mortality is a ratio that involves the use of infant birth and death data for two years. In this instance, a researcher is able to calculate more precisely the true probability of infant death by identifying all infant births in the base year of interest and linking them to their respective infants' deaths at less than one year of age in *both* the base year and the following year. This method is referred to as the *cohort* IMR for a given year and population. Infant mortality

rates are expressed in terms of infant deaths per 1,000 live births (Shryock and Siegel 1976).

In order to capture different causes that are known to affect the timing of infant deaths within the first year of life, researchers often decompose infant mortality into various components. *Neonatal mortality* usually refers to deaths of infants under 28 days of age, whereas *postneonatal mortality* involves the deaths of infants aged 28 to 364 days. Due to the large proportion of pregnancies that result in fetal deaths, and the existing similarity between the causes of death for fetuses and for neonates, other complementary measures have been constructed. *Fetal mortality* refers to fetal deaths at a gestational age greater than 28 weeks, and *perinatal mortality* is defined as fetal mortality plus infant deaths at less than 7 days of age. All corresponding rates are constructed in much the same manner as the IMR, including the use of either the period or cohort method of estimation. For both the fetal and perinatal mortality rates, researchers usually include the number of fetal or perinatal deaths (whichever may be the rate of interest) within the denominator of the rate along with the number of births for the population of interest (Shryock and Siegel 1976).

Conventionally, the number of deaths that occur among children aged one to four is a general measure of *child mortality*. In some cases, though, the child mortality rate may include the IMR (in which case it takes into account all deaths under five years of age). As a rule, most researchers interested in child mortality separate out infant mortality in order to arrive at a clearer picture of the specific causes of death that affect children who survive infancy. Child mortality rates are generally calculated by dividing the number of deaths of children aged one to four by the total number of children aged one to four in the population and multiplying by a 1,000 (Shryock and Siegel 1976).

The most common source of data on infant and child mortality in more developed countries (MDCs) is national vital statistics registries, which were established in some countries as early as the beginning of the eighteenth century (e.g., Sweden). Researchers are able to access detailed birth and death records for both infants and children and analyze their trends and patterns. Throughout the world, the most stable data on infant and child mortality have come from those countries

with established vital registration programs. For those MDCs that do not have adequate systems of vital registration and for the majority of less developed countries (LDCs) that do not have a vital statistics system in place, other techniques exist for estimating infant and child mortality (Palloni 1981). In most of these countries, population censuses and surveys of a sample of the total population are utilized to indirectly estimate mortality rates and trends (Hill 1991). Techniques developed in the late 1960s and early 1970s have been refined and improved upon in a variety of ways (United Nations 1990). Most techniques involve the use of birth histories collected from women in the population.

### EPIDEMIOLOGIC TRANSITION AND INFANT AND CHILD MORTALITY

Early research on infant mortality revealed an array of factors that presumably led to infant mortality declines in certain countries, specifically certain Western European MDCs. The experiences of these nations are the basis for a set of general hypotheses regarding the decline of infant mortality. These nations serve as test cases because of the nature of the available data and the familiarity of researchers with them. Ideas have been developed concerning the relationship among certain economic, social, and technological developments and the reduction of adult, childhood, and infant mortality. Epidemiologic transition theory, as elucidated by Omran (1971, 1977), gives researchers a theoretical framework for understanding these processes as they have occurred throughout different time periods and within different nations. This theory outlines several stages that specific nations have moved through on their way to their current mortality patterns. Many have used this theory to investigate current infant mortality levels among various countries and within specific ones (Frenk et al. 1989; Pampel and Pillai 1986).

In applying the precepts of epidemiologic transition theory to the case of infant mortality, researchers have used diverse historical sources. Epidemiologic transition theory outlines three stages through which various Western European nations moved through as they advanced toward a higher state of economic and technological development. The first two stages, termed the Era of Pestilence and Famine and the Era of Receding

Pandemics, characterize most of human existence. Within these eras, infant and child mortality were relatively high and primarily due to the prevalence of infectious disease. In the third stage, the Age of Degenerative and Man-Made Diseases, infectious diseases receded and came under medical control. Infant and child mortality fell accordingly. Other causes of death now became important, and interest in them rose. In this sense, infectious infant mortality is currently associated with countries that are less developed and populations within countries that live at disadvantaged socioeconomic levels (Omran 1971, 1977).

To better understand changes in infant mortality according to epidemiologic transition theory, researchers must be able to access data that detail the cause of infant mortality death. But as has been noted, very few countries possess the data repositories, either current and past, to adequately examine infant mortality declines and their component causes. The use of neonatal and postneonatal IMRs were one way of moving forward with these types of studies without having the necessary data for detailed analyses. Neonatal and postneonatal mortality became associated with cause-of-death groupings according to general preconceived notions of etiologic mechanisms.

Bourgeois-Pichat (1951) presumed that infectious infant mortality was primarily due to environmental factors that were wholly amenable without medical intervention. This category was termed “exogenous” infant mortality (e.g., infectious, parasitic, and respiratory diseases), while “endogenous” infant mortality was defined as those causes of infant death that were primarily due to natal and antenatal factors (e.g., congenital malformations, prematurity). Neonatal mortality became a proxy for endogenous causes of death while postneonatal mortality was associated with exogenous mortality. All that was now minimally required to continue analyses of infant mortality transitions was information on the timing of infant death (Bogue 1969; Bouvier and Van der Tak 1976). Although these conceptual distinctions probably held true throughout most of human history, debates exist as to the current validity of these conceptualizations and the continued value of their use. The introduction of neonatal technologies and other factors in MDCs have led to further discussion and reconceptualization (Sowards 1997).

## INFANT AND CHILD MORTALITY THROUGH HISTORY

Among historians such different academic spheres as demography, sociology of the family, and public health, there is an interesting debate concerning the timing and factors that have led to a decline in infant and childhood mortality. There are some interpretative similarities that may offer us a foundation with which to begin this discussion.

Historically, the size of the human population remained fairly constant until the first half of the eighteenth century. One hypothesis states that until then, although periods of low mortality were common, a population was still vulnerable to famines and epidemics, thus rendering population size stable. Before the Neolithic period (c. 9500 b.c.), it is estimated that infants had a 50 percent chance of surviving to adulthood (Stockwell and Groat 1984). It is thought that the first real decline in infant and child mortality came after the Neolithic revolution, when agriculture and animal domestication became more common, despite the increased spread of infectious diseases among now clustered and sedentary populations (Creighton-Zollar 1993). Within the framework of demographic transition theory, the staggering growth of the human population registered during the last 250 years is mainly explained by a combination of significant decrements in mortality and constant fertility rates (Kitagawa 1977; Stolnitz 1955).

Interestingly, some historians have found a link between the decline in infant mortality and the greater significance of children within the family organization. In the second half of the twentieth century, a discussion about the history of the family properly began with Philippe Ariès's *L'Enfant en la Vie Familiale sous L'Ancient Régime* (Van de Walle and Van de Walle 1990). Ariès suggests that the recognition of childhood as an independent stage of the life cycle occurred first during the sixteenth and seventeenth centuries. Other family historians have argued that infant mortality was so high that parents preferred not to “invest too much emotional capital in such ephemeral beings” (quoted by Van de Walle and Van de Walle 1990, p. 151).

For these authors, the British aristocracy's movement toward the increased domesticity of women redefined the role of the mother within



the domestic economy. British aristocratic mothers now spent an increasingly significant proportion of their time with their infants and children, forming strong parent-child bonds, which had previously been weak. These mothers were the first social group to conceive of a child as a treasure above and beyond the small utilitarian value children may have previously possessed (Van de Walle and Van de Walle 1990).

These authors suggest that from the seventeenth through the nineteenth centuries, the interest in children became increasingly mother-centered. Feelings of empowerment over the health and well-being of infants and children rose to such a level that women began demanding that the medical community and government pay more attention to children. Children thus passed from a "private" realm of familial responsibility to a "public" one, with the community now expressing a vested interest in improving the health and well-being of children. These historical changes in the power structure of the family may have operated in conjunction with various other social and economic factors that led to the decreasing rate of infant and child mortality among Western developed countries (Van de Walle and Van de Walle 1990).

#### **TWENTIETH-CENTURY MORTALITY TRENDS IN DEVELOPED COUNTRIES**

Developed and developing countries are comprised of different population histories worthy of individual historical analysis, even though the only available and reliable documents that allow the reconstruction of demographic transitions are found in Europe. For some local Western European populations, studies have revealed that mortality levels decreased slightly between 1840 and 1860, but several scholars agree that the radical decline of mortality among the MDCs took place during the period 1880-1910 or even up to 1930 (Kitigawa 1977; Preston 1992; Stolnitz 1956). Although it is generally believed that this decrease in mortality levels was primarily due to a reduction of deaths during childhood, the causes of that change are still nonetheless a matter of controversy.

Epidemiologic transition theory's main argument is that the socioeconomic development achieved by those countries was fundamental in the reduction of mortality (Omran 1971). From

this perspective, socioeconomic factors such as improved health and hygienic habits, diet and standards of living, were the main determinants of the transition from high to low mortality in Western nations. The sanitary revolution played an auxiliary role within this process of transition. Some other scholars suggest that although the improvement of economic conditions in Western nations was important, but this was more a permissive element than a precipitating factor. For them, the introduction of hygienic measures that helped to control communicable diseases was the leading cause of the decreased mortality rate (Preston 1992; Preston and Haines 1991; Spiegelman 1956; Stolnitz 1956).

According to Stolnitz (1956), in the mid-1800s reliable estimates for France and England show that real per capita income was rising rapidly but declines in mortality were still rather negligible. On the contrary, when mortality declined clearly after 1870, the West had already experienced a diverse set of economic circumstances. For instance, in France and Sweden the decrease in mortality appears to have occurred at a time when economic growth was slower than in earlier decades, while in Great Britain standards of living may have been slightly declining. Simultaneously, great changes occurred in the field of bacteriology, including the discoveries of Pasteur, and the large-scale public health programs were implemented. At the end of the nineteenth century in England, the main reason for the reduction in mortality levels was due to the lower prevalence of diseases resulting from these health innovations.

A study of the importance of different causes of death during the transition to a lower mortality rate shows interesting findings (Preston and Nelson 1974). Among 165 local populations (from 43 countries around the world) from the period 1861-1964, approximately 60 percent of the total decline in death rates was attributed to declining mortality from infectious diseases—25 percent due to influenza, pneumonia, and bronchitis; 10 percent due to respiratory tuberculosis; 10 percent due to diarrheal diseases; 15 percent due to other infectious and parasitic diseases; another 20 percent was attributed to a decrease in vascular diseases.

Many Western countries experienced a steady decline in mortality rates after 1930. Spiegelman (1956) suggests that during the period 1930-1950,

these nations were experiencing a reduction in mortality at all stages of life. According to the author, from 1946 to 1954 most countries of Western Europe achieved a reduction of up to 33 percent in their IMRs. In Table 1 we reconstruct infant mortality trends for some selected countries during a period of forty-five years. In 1948, Sweden already had a relatively low IMR (23.2 per 1,000); by 1993 it had decreased to 4.8

Japan is a striking case. Japan, Australia, and New Zealand are included in Table 1 among the developed countries following the suggestion of the World Bank (1983), although they do not share the historical experience and culture of the Western countries included in the same group. The reduction of infant mortality has been explained by specific health measures first taken during the U.S. occupation after World War II and continued later by the Japanese government. During the 1930s, life expectancy at birth (life expectancy at birth represents the average number of years to be lived for a newborn, given a specific pattern of mortality in a country in a given moment) in Japan was similar to that in Western countries at the end of the nineteenth century. However, between 1948–1950 and 1951–1953, the increase in life expectation was about 5.5 years at age 0, 4 to 5 years at age 15, and 2 years at age 45. This was far and away the greatest average annual increase in demographic history (Stolnitz 1956). Table 1 shows a Japanese decrease in infant mortality from 61.7 per 1,000 in 1948 to 30.4 in 1960. This accelerated decrease continued to 13.1, 7.4, and 4.3 in 1970, 1980 and 1993, respectively, giving Japan the lowest IMR in the world.

In the United States, child mortality had been declining in the two decades before 1900 for the total population, although the decline was greater among whites than among blacks. It is difficult to estimate the magnitude of this differential because of data problems (Preston and Haines 1991). Preston and Haines find that between 1899 and 1900, 88 percent of deaths among those under age 15 were children aged 0 to 4 and 59 percent of these were infant deaths. Regarding the three main causes of infant mortality, 25 percent of deaths were due to gastrointestinal diseases, 20 percent to respiratory diseases (e.g., influenza, pneumonia, and bronchitis), and 27 percent to malformations and a wide array of causes, such as premature birth, debility and atrophy, inanition, and

hydrocephalus. The main causes of child mortality (deaths below age five) were basically the same, but with a slightly different distribution: 21 percent were due to gastrointestinal diseases, 23 percent respiratory diseases, and 20 percent to malformations.

In comparing eighteen developed countries, Pampel and Pillai (1986) found the United States to have a relatively high IMR despite its leading economic position. Indeed, the United States has the highest IMR among developed countries. Table 1 shows that in 1948 the American IMR was half that of Japan. By 1993, however, that ratio had reversed; for every Japanese infant death there were 2 in the United States. Although this 2:1 does not exist between the United States and other developed countries, most of the countries shown in the table had higher IMR rates than the United States before World War II and have achieved significant declines since that time.

To briefly highlight some other developed countries, we take Spain as an example. According to Spiegelman (1956), acute infectious diseases were not totally under control there before World War II. This country began the forty-five-year period covered in Table 1 with an IMR of 70 in 1948. Infant mortality has continually decreased, reaching 6.7 in 1993.

Another case that may be worthy of highlighting is that of Germany, which is included in the group of Western countries because of its reunification and the general adoption of a market economy. Interestingly, in 1948 the Democratic Republic (East Germany) had a considerably higher IMR than the Federal Republic (West Germany), 94.0 versus 70.1. East Germany's rate was slightly higher in 1960, but in 1970 the trend reversed. In later years, both countries experienced equal rates. In 1993, unified Germany's IMR was 5.8, the third lowest of this selected group of developed countries.

#### TWENTIETH-CENTURY MORTALITY TRENDS IN EASTERN EUROPE

It is difficult to obtain comparable data for Eastern Europe; Table 1 shows only a few examples. In Romania in 1948 the IMR was particularly high. From 1948 to 1960 the IMR was almost halved, decreasing from 142.7 to 75.7. That pace of decrease could not be sustained, and in 1993 it

## INFANT AND CHILD MORTALITY

## Infant Mortality Rates for some Selected Countries and Years

	<i>Infant Mortality Rate</i>				
	1948	1960	1970	1980	1993
<b><i>Developed Countries</i></b>					
Canada	44.4	27.3	18.8	10.4	6.3
United States	32.0	26.0	20.0	12.5	8.4
Japan	61.7	30.4	13.1	7.4	4.3
Austria	76.2	37.5	25.9	14.1	6.5
France	55.9	27.4	18.2	9.9	6.5
Germany					5.8
Democratic Republic	94.0	38.8	18.5	12.1	
Federal Republic	70.1	33.8	23.6	12.6	
Italy	72.1	43.9	29.6	14.3	7.1
Spain	70.0	43.7	26.5	11.1	6.7
Sweden	23.2	16.6	11.0	6.7	4.8
United Kingdom	36.0	22.5	18.4	12.2	6.3
Australia	27.8	20.2	17.9	11.0	6.1
New Zealand	27.5	22.6	16.7	30.2	7.2
<b><i>Eastern Europe</i></b>					
Czechoslovakia	83.5	23.5	22.1	16.6	8.5
Poland	110.7	56.1	33.2	21.2	16.2
Romania	142.7	75.7	49.4	29.3	23.3
Yugoslavia	102.1	87.7	55.5	32.8	21.9
<b><i>Latin America</i></b>					
Argentina	69.5	62.4	58.9	33.2	22.9
Bolivia	131.3	74.3	—	124.4*	75.0
Brazil	—	—	—	70.6*	47.0
Costa Rica	90.4	74.3	61.5	20.2	13.7
Cuba	40.1	35.4	38.7	19.6	9.4
El Salvador	100.4	76.3	66.6	42.0	44.0
Mexico	101.7	74.2	68.5	53.0*	34.0
Peru	109.0	92.1	64.6	98.6*	74.9
Venezuela	97.8	53.9	49.3	27.7*	22.0
<b><i>Africa</i></b>					
Egypt	138.6	109.3	116.3	76.0	117.0
Kenya	—	—	81.2	80.2*	71.0
Liberia	—	—	159.2	132.5*	200.0
Mauritius	186.2	70.9	58.5	28.4*	19.9
Mozambique	48.8	34.2	41.6	153.5*	118.0
South Africa				83.3	53.0
Asian	77.1	59.6	36.4		
Black	133.2	128.6	132.6		
White	36.0	29.6	21.6		
<b><i>Asia</i></b>					
Afghanistan	—	—	—	193.8*	163.0
Bangladesh	—	—	—	128.2*	91.0
China	—	—	—	39.3*	44.0
India	130.1	86.5	122.0	113.9	74.0
Iraq	97.6	17.7	19.7	77.1*	127.0
North Korea	—	—	—	54.1*	28.8
South Korea	—	—	—	29.7*	12.3
Philippines	114.4	73.1	60.0	50.6	40.0
Saudi Arabia	—	—	—	66.1*	29.0
Singapore	80.8	34.8	19.7	11.7	4.7
Sri Lanka	92.1	56.8	45.1	34.4	16.5

Table 1

SOURCE: United Nations (1979, 1982, 1987, and 1996).

NOTE: \*Data from 1983.

reached 23.3, the highest IMR among the Eastern European countries shown.

In the cases of Poland and Yugoslavia, both countries had a high IMR in the immediate postwar period (110.7 and 102.1, respectively). Despite its lower starting level, Yugoslavia consistently experienced higher rates thereafter: 87.7 in 1960, 55.5 in 1970, 32.8 in 1980, and 21.9 in 1993, whereas in Poland such rates have been 56.1, 33.2, 21.2, and 16.2 for the same years.

In 1948, Czechoslovakia had the lowest IMR among the Eastern bloc countries shown in Table 1, but it was still higher than that of Western countries. The decrease observed from 1948 to 1960 is very significant: from 83.5 to 23.5 in twelve years. This pace decelerated but continued to decline. By 1993, it had achieved an IMR similar to that of the United States: 8.5 deaths 1,000/live births.

#### TWENTIETH-CENTURY MORTALITY TRENDS IN DEVELOPING COUNTRIES

In the developing countries of Latin America, Asia, and Africa, reduction in mortality levels began around the beginning of the twentieth century. Scholars have found only fragmented information to reconstruct the demographic history of such countries. They can affirm, however, that the pace of declining mortality was slow until 1940 and accelerated after World War II (Kitagawa 1977; Meslé and Vallin 1996; Preston and Nelson 1974). According to Stolnitz (1956), the decline in mortality in developing countries during the postwar period, which was highly accelerated, has no parallels in the Western experience. The common explanation, found in numerous studies (Arriaga 1970; Kitigawa 1977; Preston 1976; Stolnitz 1956), is that access to medical innovations coming from developed countries and the implementation of specific programs of disease control were the main cause of the mortality reduction, especially in infant mortality.

Since the changes in those regions have not been uniform, currently there is a convergence of the mortality patterns in developing countries, whereas in the developed areas there is a marked divergence (Meslé and Vallin 1996). Indeed, trends

of mortality just described for developed countries show variations, but among developing countries such heterogeneity is even greater. It would be impossible in this space to describe the mosaic of infant mortality experiences in developing countries; however, we can point out some major tendencies and outstanding cases during the forty-five years period covered by Table 1.

In order to understand the scope of differences among developing countries, it is important to consider a general characteristic of populations with high mortality: The younger the age, the larger the absolute change in mortality rates (Arriaga 1970). At the beginning of this transition toward a low mortality pattern, drastic changes are observed in infant and child mortality among different populations. For many developing countries, infant and child mortality trends throughout history cannot be known because there are no available data. Although this issue will be discussed below, usually the more socioeconomically disadvantaged countries are the ones for whom the data are lacking.

**Latin America.** During the 1930s Latin America had a high mortality pattern. For instance, a newborn baby had less than a 50 percent chance of living to the age of thirty; consequently, no country had a life expectancy more than forty years at birth (Arriaga 1970). The main decline was registered between 1940 and 1960. The largest decline could be found during the 1940s; there was a slower pace of decrease in the 1950s (Arriaga 1989; Camposortega Cruz 1989a).

Brazil and Mexico are the two most highly populated countries in Latin America, and they experienced similar processes of industrialization up until the 1970s. Brazil had a military government and Mexico did not; however, both have high economic inequality, possibly Brazil the more so. Unfortunately there are no data available for Brazil prior to the 1980s, but Table 1 shows that its IMR was higher than that of Mexico in both 1983 and 1993. Mexico's IMR decreased by 25 percent between 1948 and 1960; it continued to decline even in the midst of a drastic economic recession that began in the 1980s. By 1993, the persistently high rate had been reduced to 34.0 infant deaths per 1,000 live-births; the Brazilian IMR was 47.0 in that year.

Argentina provides an example of early industrialization as well as early demographic transition in the region. During the immediate postwar period, this country had an IMR similar to those of several Western countries (69.5 in 1948); considering this fact, the rate in 1993 (22.9) was relatively high, even higher than the IMR (22.0) of Venezuela, a country that had a rate 30 percent higher than Argentina in 1948.

El Salvador, Bolivia, and Peru, some of the poorest countries in the region, had high IMRs in 1948. For example, in El Salvador a tenth of all infants died in 1948, but this rate had decreased 25 percent by 1960 to 76.3 infant deaths per 1,000 live births. This pace of decline slowed in subsequent years, reaching 42.0 in 1980. However, by 1993 this indicator of general well-being had reversed, and the IMR had increased slightly to 44.0, perhaps because of severe economic conditions during the 1980s.

In Bolivia, if the data in Table 1 are correct, the decrease in infant mortality between 1948 and 1960 was spectacular: from 131.3 deaths to 74.3 deaths per 1,000 within twelve years. Information for the 1970s is unavailable. In 1983 the rate was again high (124.4), but ten years later it had dropped to 75.0. These significant shifts give rise to concerns about the quality of the data. In any case, the IMR in 1993 was higher than the Salvadorian rate and similar only to the Peruvian one.

Bolivia and Peru have the highest IMRs among the selected countries (75.0 and 74.9, respectively). In this region, such a rate might be surpassed only by Haiti, for which information is not available. Interestingly, Peru had a decrease in its IMR from 1948 (109.0) to 1970 (64.6). But in 1983 the IMR was even higher than it had been twenty years earlier: In 1983 it was 98.6, up from 92.1 in 1960. Although the IMR had decreased to 74.9 in 1993, Peru still has one of the highest rates of infant mortality in all of Latin America.

In contrast, Cuba and Costa Rica have the lowest and the second lowest IMRs, respectively. The public health systems in both countries have good reputations throughout the region. Costa Rica shows a steady decrease from 1948, when its IMR was 90.4. By 1970, it had declined 30 percent, reaching 61.5. Contradicting Arriaga (1989) and Camposortega Cruz (1989a), this country achieved a steeper decrease in infant mortality during the

1970s: In 1970 the rate was 61.5 and in 1980, 20.2, showing a decrease of 33 percent. In 1993, the Costa Rican IMR was 13.7.

The Cuban IMR was already the lowest in the region in 1948 (40.1) among those countries shown in Table 1. Up until 1970 there was only a small and unsteady decrease to (38.7). During the 1970s and afterward, the country achieved significant decreases: The rate reached 19.6 in 1980 and 9.4 in 1993. The decrease in infant mortality has been one of the main goals and achievements of the Cuban health system, despite the country's unfavorable economic situation. Cuba now has the lowest IMR in Latin America, being only slightly higher than that of Czechoslovakia and the United States.

**Asia.** Infant mortality rates in Asia are quite diverse. As regards of India, Jain and Visaira (1988) argue that even though several studies suggest that the infant mortality rate in 1920 was 240 per 1,000 live births, this might have been understated, and the IMR could have been between 200 and 225 during the 1940s. This finding is in contrast to the estimations of Arriaga (1989) and other researchers. Using different sources of information and techniques, Jain and Visaira construct a trend of infant mortality with higher rates than those shown in Table 1, with the rates coinciding only since the 1970s. Jain and Visaira (1988) and Arriaga (1989) found that as a result of implementing a specific program of hygiene among mothers, infant mortality did decline after 1970. Between that year and 1993, the IMR decreased from 122.0 to 74.0.

Although there is no complete direct information for China that allows us to construct a complete infant mortality trend since the postwar period, some estimations are provided in a study by Banister (1986). From the mid-1960s to 1977-1978, the author and her colleagues found a constant decrease in infant mortality (from 70 to 40 deaths), although the trend was later reversed. Their calculation for 1981-1982 was 61 deaths per 1,000 live births, which does not match the information from the United Nations: 39.3 in 1983 and 44.0 a decade later (Table 1). Banister and her team examined the possibility of measurement errors in order to discern whether the estimation was fatally biased and changed the direction of the trend. They confirmed that such was not the case

(Banister 1986). There may have been an over- or underestimation between both calculations, but either way, the direction of the tendency toward an increasing IMR seem to be correct.

According to Banister (1986), there are three reasons that may explain this inversion. First, in 1978 a general economic reform began and the availability of preventive and curative health services was somehow affected, which may have been detrimental to the survival of some infants. Second, at the same time China's severe family-planning reached rural couples. In 1981, 53 percent of all births corresponded to second and higher-order births; therefore, if mothers of these unauthorized babies lacked prenatal care, not surprisingly their infants would experience higher risk of early death. Additionally, the provision of medical care had been tilted intentionally toward the only child, which may have implied the deterioration of the health care available for children in multichild families. Third, a decrease in the proportion of breast-feeding combined with the use of contaminated water used to mix formula might have adversely affected infant survival. Although these researchers did not have precise information on the proportion of breast-feeding, based on fragmentary data that proportion seems to have decreased (in 1982 22 percent of the babies in urban areas and 60 percent in rural settlements were breast-fed). A possible fourth factor is infanticide, but the distribution of infant deaths by sex did not show a bias toward more female deaths (it is said that there exists in China a preference for male babies). Hence, if infanticide played an important role in increasing the IMR, Banister's evidence suggest either that the incidence of female infanticide is not statistically significant or that respondents reported neither the birth nor the death of female victims of infanticide.

Sri Lanka is a rare case of a developing country's retaining records from nearly the beginning of the twentieth century. Stolnitz (1956) refers to Sri Lanka (formerly Ceylon) as one example of rapidly declining infant mortality; more recently, Meegama (1986) offers an overview of mortality trends since the beginning of the twentieth century. According to this latter author, during the second half of the nineteenth century, the IMR started to decrease because of a reduction in the

number of famines and the control of cholera epidemics.

During 1911–1915, the IMR was 201 deaths per 1,000 live births, fifteen years later it had decreased to 175 deaths. In 1945, at the end of the World War II, the IMR was recorded at 140 infant deaths per 1,000 live births. Prior to the war, there was a decrease in the mortality rate for both sexes and all age groups in both endemic and nonendemic malarial zones. One of the main reasons for this decline was the expansion of maternity and child welfare services, which decreased both infant and maternal mortality (Meegama 1986). It might have been the combination of that preexisting decline and the introduction of a wider range of medical measures that caused a steady decrease in mortality after the war. As shown in Table 1, in 1948 Sri Lanka had an IMR of 92.1. This was halved by the beginning of the 1970s (45.1) and was again reduced to a third of the 1970 rate by 1993 (16.5).

South Korea and Singapore are two cases of recent industrialization and fast decline in mortality patterns. Information is not available to reconstruct the South Korean mortality history; however, what can be seen in Table 1 is a decrease of more than 50 percent from 1983 to 1993. In 1983 the country had an IMR of 29.7, which had declined to 12.3 ten years later.

Singapore is another impressive case of accelerated reduction in infant mortality. In 1948, it had a relatively low IMR among Asian countries (80.8), but its rate was greater than that of Japan. By 1960, the rate had been reduced by more than 50 percent, reaching 34.8. The trend continued, and by 1993 Singapore's IMR was 4.7, similar to the Swedish and very close to the Japanese rate.

**Africa.** Mortality in Africa is the highest of any major region in the world and historically has had the least reliable information, with certain exceptions. Through the application of special techniques, mortality trends have been reconstructed. The pattern of mortality in Africa is characteristic of populations with a high mortality schedule: It is dominated by mortality among children under five years of age, and infectious, parasitic, and respiratory diseases are the major causes of death. Concerning infant mortality in Africa, we examine some cases from this region displayed in Table 1.

Kenya lacks a usable vital registration system, but the application of indirect techniques has helped in reconstructing mortality trends there. In 1948, when the country began its transition toward lower mortality, it might have had an IMR of 184 per 1,000 live births, according to some estimates (Ewbank et al. 1986). In 1970 that rate had dropped to 81.2, but in the following decades the pace of change was considerably slower; in 1993 the rate was 71.0. For an African country, this is a low mortality pattern, but it is comparable to some Asian countries and to the poorest in Latin America.

Mauritius has a mortality pattern that is gradually approaching that of Western Europe (United Nations 1984). It has the lowest IMR of African countries shown in Table 1. In the period immediately after the World War II, Mauritius had an IMR of 186.2 deaths per 1,000 live births; this decreased 62 percent in twelve years to 70.9 in 1960. That steady rate of decrease was maintained, with the IMR falling to 19.9 by 1993.

The South African case presented some different data analysis challenges because prior to the 1980s information on mortality was not available for the country as a whole but rather by ethnic group, showing the effect of apartheid on vital statistics but also allowing us to see ethnic differentials in the levels of infant mortality. As shown in Table 1, the white population had always been better off than their fellow South Africans; the IMR of the black population held relatively constant until 1970 and the IMR of the Asian population was located between that of the white and black populaces. The ratio of deaths between ethnic groups illustrates social differences: In 1948 there were approximately 4 black infant deaths for each white one and 1.7 black infant deaths for each Asian one. The ratio for Asian to white deaths was approximately 2-to-1 for babies under the age of one year. In 1970 the gap between blacks and the other races widened: There were 6.1 black infant deaths for each white death, 3.6 black infant deaths for each Asian death, and 1.7 Asian infant deaths for each white infant death. By the 1980s the information from the *Demographic Yearbook* was no longer classified by ethnic group; it shows that the general IMR decreased between 1983 and 1993, from 83.3 to 53.0, but it does not show the ethnic differential during those years.

## DIFFERENTIALS IN INFANT AND CHILD MORTALITY

Numerous studies have documented and examined various differences in infant and child mortality among different groups along various socioeconomic and ecological spectrums. For most of the earliest periods under investigation, ecological and geographic factors seemed to account for most of the observed differences in infant and child mortality rates (Preston and Haines 1991). Current research on this topic is guided by a theoretical framework put forth by Mosley and Chen (1984). Their original intent was to outline an analytical framework for studying child survival in developing countries, but many researchers have found this framework to be extremely suitable for studying infant and mortality differentials within various contexts.

Mosley and Chen (1984) propose that all socioeconomic determinants, or exogenous factors, of child mortality operate through a set of biological mechanisms that they call "proximate determinants" or "intermediate variables." That is, social and economic factors influence infant and child mortality through a set of biological factors that an infant is born with. Many studies have identified numerous variables and factors that influence both the biological viability of an infant at birth and the health and well-being of the infant *ex utero*, thereby increasing or decreasing the probability that an infant will die before its first birthday or that a child will die before its fifth. These social and economic factors range from individual-level characteristics to household- and community-level factors.

Research has identified three main infant birth outcomes that exert major influences on infant survival chances and through which many exogenous factors operate: infant birthweight, gestational age, and the interaction of the two, sometimes referred to as intrauterine growth retardation (IUGR). Infants weighing less than 2,500 grams at birth (low birthweight, LBW) and infants of short gestation (less than thirty-seven weeks' gestation=premature) have been consistently found to be at dramatically higher risk of infant mortality (Hummer et al. 1999). Interactions between these two outcomes have also been found to affect infant mortality rates significantly (Frisbie et al. 1996). These relationships hold for both MDCs and LDCs (UN 1984).

Of the set of proximate and exogenous determinants originally outlined by Mosley and Chen (1984), most have been identified as having significant effects on infant mortality. Smoking, alcohol, and illicit drug use have all been shown to adversely affect infant and child mortality (Chomitz et al. 1995; Petitti and Coleman 1990). Maternal weight gain has been found to be negatively correlated with infant mortality risk; that is, the more weight a woman gains during her pregnancy, the lower the infant mortality risk (Hummer et al. 1999). The birth order of the infant is associated with differential infant mortality risks, while shorter interpregnancy intervals are linked to higher mortality risks (Kallan 1993; Miller 1991). Male infants have always had higher mortality risks (Rumbaut and Weeks 1996). The role of prenatal care has been debated in recent times, but most researchers agree that adequate prenatal care produces beneficial effects for the health and well-being of mothers, infants, and ultimately young children (Fiscella 1995). The effect on infant mortality risk of maternal age had previously been thought to be constantly negative in direction, but more recent analyses indicate that many interaction effects may work specifically through maternal age, making the conceptual picture much more complex (Geronimus 1987).

Exogenous factors have also been investigated and analyzed. One factor that has captured the continued interest of researchers is the role that race and ethnicity play in explaining infant and child mortality differentials. In both MDCs and LDCs consisting of multiracial and/or multiethnic populations, racial and ethnic infant and child mortality differentials often exist. For example, in 1996 the black population in the United States experienced an IMR of 14.1 while the white population experienced an IMR of 6.1 (MacDorman and Atkinson 1998). In 1990, the indigenous population of Mexico experienced a higher risk of infant mortality 1.5 times higher than the nonindigenous population (Fernández-Ham 1993). Although the role that race and ethnicity may play in IMR differentials is complex, there is agreement that such differences are based on social, economic and historical inequalities (Cooper and David 1986; David 1990). In LDCs those historical inequalities are also related to place of residence, which in most of these countries imply a more disadvantaged position for rural populations (Camposortega Cruz 1989b).

More direct socioeconomic indicators are currently being tested to identify their effects on mortality risks. Increases in maternal education have consistently been shown to decrease both infant and child mortality risk at both the individual and national levels (Hummer et al. 1992; Pampel and Pillai 1986). Marital status has also been linked to differential infant mortality risks; unmarried mothers have a higher risk for infant mortality. This may reflect the fact that single mothers often must bear the costs of both child-rearing and day-to-day living expenses with limited assistance (Frisbie et al. 1997). Studies have also identified the important role that income plays in influencing infant and child mortality risk (Geronimus and Korenman 1988). Cramer (1995) finds evidence of this association even when controlling for a wide array of other sociodemographic risk factors. Many others have explored the effects of various public health assistance and intervention programs, either by governmental or nonprofit agencies. In both MDCs and LDCs, some evidence exists to support the creation and implementation of such intervention programs, such as the Women, Infants and Children (WIC) program in the United States and governmental health programs in Guatemala (Moss and Carver 1998; United Nations 1984). According to Frenk and associates (1989), in certain developed countries there may exist two populations divided by their differential infant and child mortality risk. These researchers point to existing differences between insured and uninsured populations in Mexico as an example of the divergent character of national infant mortality profiles in both high- and middle-income countries (Frenk et al. 1989).

With the development of new theoretical frameworks and statistical methodologies, infant and child mortality research has traveled down new paths. One interesting development has been analyses based upon characteristics of both the individual and the neighborhood or community in which individuals live. Determinants of infant and child mortality are hypothesized to emerge from conflation of personal, household, and neighborhood factors. Collins and David (1997) find that violent neighborhoods may adversely effect an expectant mother's pregnancy outcome, in spite of the measured personal characteristics of the mother. Guest (1998) finds that the neighborhood unemployment level has an important effect on



infant mortality beyond effects due to other socioeconomic indicators and race.

### PROSPECTS

The search for solutions to reduce levels of infant and child mortality will require the joint participation of multiple actors: governments, international agencies, nongovernmental organizations, and individuals within the scientific and policy communities. Predicting future trends in infant and child mortality remains a rather difficult task given the complexity of this phenomenon. However, scientific knowledge gathered at both the aggregate and individual levels should make reduction of the significant gap between the developed and developing world possible. The multiple differences that exist in mortality patterns within groups in both developed and developing countries are illustrated by the current age structure of mortality. In 1990-1995, life expectancy at birth was 74.4 years in developed countries and 62.3 years in the less developed world. This indicator is consistent with the corresponding IMRs: 10 versus 70 infant deaths per 1,000 live births for MDCs and LDCs, respectively.

It is noteworthy that infant and child mortality rates fell so drastically in the twentieth century. But even so, there remains the need to determine the next course of action in order to combat the mortality differentials between less and more developed nations and among various socioeconomic groups within nations. Continued successes will depend on identifying problem areas and creating innovative solutions to these health, and ultimately, social problems.

### REFERENCES

- Arriaga, E. E. 1970 *Mortality Decline and its Demographic Effects in Latin America* (Population Monograph Series No. 6). Berkeley: University of California Press.
- 1989 “Changing Trends in Mortality Decline during the Last Decades.” In L. Ruzicka, G. Wunsch, and P. Kane, eds., *Differential Mortality: Methodological Issues and Biosocial Factors*. Oxford: Clarendon Press.
- Banister, J. 1986 *China: Recent Trends in Health and Mortality*. Washington, D.C.: Center for International Research, U.S. Bureau of the Census.
- Bogue, D. J. 1969 *Principles of Demography*. New York: John Wiley and Sons.
- Bourgeois-Pichat, J. 1951 “La Mesure de la Mortalité Infantile. I. Principes et Methodes.” *Population* 2:1-17.
- Bouvier, L. F., and J. Van der Tak 1976 “Infant Mortality: Progress and Problems.” *Population Bulletin* 31:3-33.
- Camposortega Cruz, S. 1989a “La evolución de la mortalidad en México, 1940-1980.” *Estudios Demográficos y Urbanos* 4(2):229-264.
- 1989b “Mortalidad en México. Algunas Consideraciones Sobre Los Diferenciales Urbano-Rurales.” *Estudios Demográficos y Urbanos* 4(3):573-593.
- Chomitz, V. R., L. W. Cheung, and E. Lieberman 1995 “The Role of Lifestyle in Preventing Low Birth Weight.” *Future of Children* 5(1):121-138.
- Coale, A. J. 1965 “Birth Rates, Death Rates, and Rates of Growth in Human Population.” In Mindel C. Sheps and Jeanne C. Ridley, eds., *Public Health and Population Change*. Pittsburgh, Pa.: University of Pittsburgh Press.
- Collins, J. W., and R. J. David 1997 “Urban Violence and African-American Pregnancy Outcomes.” *Ethnicity and Disease* 7:184-190.
- Cooper, R., and R. David 1986 “The Biological Concept of Race and its Application to Public Health and Epidemiology.” *Journal of Health Politics, Policy and Law* 11:97-116.
- Cramer, J. C. 1995 “Racial and Ethnic Differences in Birthweight: The Role of Income and Family Assistance.” *Demography* 27:413-430.
- Creighton-Zollar, A. 1993 “Infant Mortality by Socioeconomic Status and Race: Richmond, VA 1979-1981.” *Sociological Spectrum* 10:133-142.
- David, R. 1990 “Race, Birthweight, and Mortality Rates [editorial].” *Journal of Pediatrics* 116:101-102.
- Ewbank, D., R. Henin, and J. Kekovole 1986 “An Integration of Demographic and Epidemiologic Research on Mortality in Kenya.” In United Nations, *Determinants of Mortality Change and Differentials in Developing Countries* (Population Studies No. 94). New York: United Nations.
- Fernández-Ham, P. 1993 “La Mortalidad Infantil en la Población Indígena: Atrasos y Contrastes.” *Demos* 6:12-13.
- Fiscella, K. 1995 “Does Prenatal Care Improve Birth Outcomes? A Critical Review.” *Obstetrics & Gynecology* 85(3):468-479.
- Frenk, J., J. Bobadilla, J. Sepulveda, and M. Lopez 1989 “Health Transition in Middle-Income Countries: New Challenges for Health Care.” *Health Policy and Planning* 4(1):29-39.
- Frisbie, W. P., M. Biegler, P. deTurk, D. Forbes, and S. G. Pullum 1997 “Determinants of Intrauterine Growth

- Retardation and Other Compromised Birth Outcomes: A Comparison of Mexican Americans, African Americans, and Non-Hispanic Whites." *American Journal of Public Health* 87(12):1977-1983.
- Frisbie, W. P., D. Forbes, and S. G. Pullum 1996 "Compromised Birth Outcomes and Infant Mortality Among Racial and Ethnic Groups." *Demography* 33(4):469-481.
- Geronimus, A. T. 1987 "On Teenage Childbearing and Neonatal Mortality in the United States." *Population and Development Review* 13(2):245-279.
- , and S. D. Korenman 1988 "Comment on Pampel and Pillai's 'Patterns and Determinants of Infant Mortality in Developed Nations, 1950-1975.'" *Demography* 25(1):155-158.
- Guest, A. M. 1998 "The Ecology of Race and Socioeconomic Distress: Infant and Working-Age Mortality in Chicago." *Demography* 35(1):23-34.
- Hill, K. H. 1991 "Approaches to the Measurement of Childhood Mortality: A Comparative Approach." *Population Index* 57(3):368-382.
- Hummer, R. A., M. Biegler, P. deTurk, D. Forbes, W. P. Frisbie, Y. Hong, and S. G. Pullum 1999 "Race/Ethnicity, Nativity and Infant Mortality." *Social Forces* 77(3):1083-1118.
- Hummer, R. A., I. A. Eberstein, and C. B. Nam 1992 "Infant Mortality Differentials Among Hispanic Groups in Florida." *Social Forces* 70(4):1055-1075.
- Jain, A. K., and P. Visaira (eds.) 1988 *Infant Mortality in India. Differentials and Determinants*. New Delhi: Sage Publications.
- Kallan, J. E. 1993 "Race, Intervening Variables, and Two Forms of Low Birthweight." *Demography* 30:489-506.
- Kitagawa, E. M. 1977 "On Mortality." *Demography* 14(4):381-389.
- MacDorman, M. F., and J. O. Atkinson 1998 "Infant Mortality Statistics from the 1995 Period Linked Birth/Infant Death Data Set." *Monthly Vital Statistics Report* 46(12s):1-24.
- Meegama, S. A. 1986 "The Mortality Transition in Sri Lanka." In United Nations, *Determinants of Mortality Change and Differentials in Developing Countries* (Population Studies No. 94). New York: United Nations.
- Meslé, F., and J. Vallin 1996 *Mortality in the World: Trends and Prospects*. Paris: Centre Français sur la Population et le Développement.
- Miller, J. E. 1991 "Birth Intervals and Perinatal Health: An Investigation of Three Hypothesis." *Family Planning Perspectives* 23:62-70.
- Mosley, W. H., and L. C. Chen 1984 "An Analytic Framework for the Study of Child Survival in Developing Countries." *Population Development and Review* 10(supplement):25-48.
- Moss, N. E., and K. Carver 1998 "The Effect of WIC and Medicaid on Infant Mortality in the United States." *American Journal of Public Health* 88(9):1354-1361.
- Omran, A. R. 1971 "The Epidemiologic Transition. A Theory of Population Change. Part 1." *The Milbank Fund Quarterly* 49(4):509-538.
- Omran, A. R. 1971 "Epidemiologic Transition: A Theory of the Epidemiology of Population Change." *Milbank Memorial Fund Quarterly* 49(4):509-538.
- Omran, A. R. 1977 "Epidemiologic Transition in the U.S.: The Health Factor in Population Change." *Population Bulletin* 32(2):3-42.
- Palloni, A. 1998 "A Review of Infant Mortality Trends in Selected Underdeveloped Countries: Some New Estimates." *Population Studies* 35(1):100-119.
- Pampel, F. C., and V. K. Pillai 1986, "Patterns and Determinants of Infant Mortality in Developed Nations, 1950-1975." *Demography* 23(4):525-542.
- Petitti, D. B., and C. Coleman 1990 "Cocaine Use and the Risk of Low Birthweight." *American Journal of Public Health* 80:25-27.
- Preston, S. H. 1976 *Mortality Patterns in National Populations*. New York: Academic Press.
- 1992 "Infant and Child Mortality." In E. F. Borgatta and M. L. Borgatta, eds., *Encyclopedia of Sociology*, vol. 2. New York: Macmillan.
- , and M. R. Haines 1991 *Fatal Years: Child Mortality in Late Nineteenth-Century America*. Princeton, N.J.: Princeton University Press.
- , and V. E. Nelson 1974 "Structure and Change in Causes of Death: An International Summary." *Population Studies* 28(1):19-51.
- Rumbaut, R. G., and J. R. Weeks 1996 "Unraveling a Public Health Enigma: Why Do Immigrants Experience Superior Perinatal Health Outcomes." *Research in the Sociology of Health Care* 13B:337-391.
- Shryock, Henry S., and J. Siegel 1976 *The Methods and Materials of Demography*. Washington, D.C.: U.S. Government Printing Office.
- Sowards, K. A. 1997 "Premature Birth and the Changing Composition of Newborn Infectious Disease Mortality: Reconsidering 'Exogenous' Mortality." *Demography* 34(3):399-409.
- Spiegelman, M. 1956 "Recent Trends and Determinants of Mortality in Highly Developed Countries." In F. G. Boudreau and C. V. Kiser, eds., *Trends and*

- Differentials in Mortality*. New York: Milbank Memorial Fund, pp. 26–34.
- Stockwell, E. G., and H. T. Groat 1984 *World Population: An Introduction to Demography*. New York: Franklin Watts.
- Stolnitz, G. J. 1955 “A Century of International Mortality Trends: I.” *Population Studies* 9(1):24–55.
- 1956 “Comparison Between Some Recent Mortality Trends in Underdeveloped Areas and Historical Trends in the West.” In F. G. Boudreau and C. V. Kiser, eds., *Trends and Differentials in Mortality*. New York: Milbank Memorial Fund, pp. 51–60.
- United Nations 1979 *Demographic Yearbook. Historical Supplement*. New York: United Nations.
- 1982 *Demographic Yearbook*. New York: United Nations.
- 1984 “The Influence of Developmental-Related Change on Infant and Child Mortality in Africa.” In United Nations, *Mortality and Health Policy: Proceedings of the Expert Group on Mortality and Health Policy, Rome, 30 May to June 1983*. New York: United Nations.
- 1987 *Demographic Yearbook*. New York: United Nations.
- 1990 *Step-by-Step Guide to the Estimation of Child Mortality*. New York: United Nations.
- 1996 *Demographic Yearbook*. New York: United Nations.
- Van de Walle, E., and F. Van de Walle 1990 “The Private and the Public Child.” In J. Caldwell, S. Findley, P. Caldwell, G. Santow, W. Cosford, J. Braid, and D. Broers-Freeman, eds., *What We Know About Health Transition: The Cultural, Social and Behavioural Determinants of Health*, Vol. 1. Australia: The Australian National University, pp. 150–164.
- World Bank 1983 *World Development Report*. New York: Oxford University Press.

GEORGINA ROJAS-GARCÍA  
SAMUEL ECHEVARRÍA

## INFORMAL ECONOMY, THE

Informal economic activity can only be understood in relationship to its counterparts in the institutions of the formal economy, most notably in markets for labor, goods, and services. In most modern, industrial, capitalist societies, economic activity occurs in the form of exchanges of cash for labor, consumer goods, and services within well-regulated economic institutions located within the

*formal economy*. The labor market, in particular, provides the framework for regulating the exchange of work and wages between sellers and buyers of labor (workers and employers) under the auspices of a legal and regulatory environment enforced by the state. At the same time, it is also widely recognized that livelihood strategies may entail all manner of exchange outside these formal institutions and regulations. That is, exchanges may occur in the informal economy. Goods and services may be bartered, traded for nonmonetary compensation, or produced under conditions that flaunt health and safety regulations. Earned income may go unreported and untaxed. Individuals may self-provision, using their own resources to produce and consume household necessities even when they are readily available for purchase. Credit may be extended on the basis of kinship and trust. Numerous other types of unrecognized exchanges create value “under the table,” that is, in the informal economy.

In some societies the informal economy is the dominant mode of economic activity. It is especially prominent in developing countries where the majority of the population may engage in various types of economic exchanges outside any formal regulation, cash economy, or state supervision. Yet informal activity is not limited to developing nations. Increasingly, economists, sociologists, and others who study work and labor markets recognize that the informal economy is very much present in first world industrial and postindustrial economies. The *shadow*, *black*, *gray*, *underground*, *subterranean*, or *parallel economy* are among the common euphemisms for an informal economy that exists side by side with its more studied and theorized counterparts in developed as well as developing nations, first and third worlds, urban and rural communities, and everything in between.

### DEVELOPMENT OF THE CONCEPT

**History.** The term *informal economy* has numerous uses and meanings and few precise definitions. Its original formulation is attributed to an anthropologist, Keith Hart, studying emerging urban labor markets in Africa for the International Labour Office (ILO) (Hart 1973). Hart distinguished between waged employment in large firms or government agencies and self-employment in his studies of the urban labor force, noting the wide

variety of small-scale entrepreneurial activities that were central to the livelihood strategies and economic life of residents of urban centers in Africa. Small traders, food vendors, and sellers of goods or services from shoe shines to haircuts are the familiar mainstays of urban economies. Subsequent use of the term by the ILO came to mean self-employment and small-scale family enterprise correlated with poverty, underemployment, and low productivity, particularly in poor, developing nations. This view was widely adopted by analysts and policy makers for whom the informal economy was a phenomenon of poverty and underdevelopment, with the implication that development of a modern economy decreases the need for and significance of informal activity. Ultimately, informal work should be absorbed into the regulated economy.

More recently, assumptions about the scope, location, and importance of informal work have come under scrutiny, since it has been recognized that informal work is neither unique to impoverished Third World countries nor necessarily diminishing in size or importance. Several developments in diverse areas of economic sociology, but all of them linked to global economic restructuring, have led to renewed interest in informal work and a growing understanding that it is both more widespread and less understood than previous accounts suggested. Studies of urban ethnic enclaves, patterns of employment for new immigrants, and the explosion of the global division of labor in "world cities" have generated interest in informal work. Similarly, there has been a corresponding interest in livelihood and income-generating strategies of both the urban and rural poor, and in wide-scale economic restructuring that sought increased flexibility in production processes and labor practices, as well as recognition of the persistence of noncash exchanges in communities of all sizes and locations. Additionally, periodic headline-grabbing media accounts of sweatshops, child and immigrant labor abuses, and other gross violations of labor law and workplace protections have fueled broad awareness of the persistence and pervasiveness of informal activity.

**Definitions.** Despite growing awareness of the ubiquity of the informal sector, a precise definition remains elusive, and numerous issues remain

unresolved. Typically, informal activity is defined by what it is not; that is, it is not formal, it is not regulated, and it is not counted in official statistics and national accounting schemes. Castells and Portes (1989) provide an influential definition using this approach: "The informal economy is . . . a process of income-generation characterized by one central feature: it is unregulated by the institutions of society, in a legal and social environment in which similar activities are regulated" (p. 12).

While broadly inclusive, this definition provides little guidance for distinguishing between different types of activities conducted under different circumstances. For example, should overtly criminal activities, such as drug dealing, be classified in the same way as work that is otherwise legal, such as flea market sales, except that is unreported and untaxed? Should both production and consumption processes be classified as part of the informal sector? In other words, are the goods that are produced in sweatshops comparable to those produced at home for private consumption? When do self-employment and the use of family labor become informal work? These questions illustrate both the diverse approaches found in the literature and the confusion that exists there.

Other analysts emphasize specific defining characteristics of the informal economy rather than focusing on what it is not. For example, Roberts (1994), while accepting the basic outline of the standard definition, argues that the key feature of informality is not the absence of regulation but instead the existence of a specific type of regulation that dominates the activity. Thus, the informal sector is marked by the dominance of regulation based on personal relations and networks embedded in family, community, friendship, or ethnicity rather than on regulation organized on formal, legal, or contractual bases. Similarly, Mingione (1991) suggests that income strategies organized around reciprocal networks, such as those found in family and household labor and relations, characterize the informal sector. These analysts emphasize the interpersonal networks that create obligations and responsibilities that permit exchanges that bypass formal institutions. This approach has found expression in the many studies of immigrant labor and ethnic enclave enterprise where kinship and co-ethnic personal ties are the means for conducting business.

A number of taxonomies and typologies have been proposed in an effort to codify characteristics of informal activity. Feige (1990) distinguishes between illegal, unreported, unrecorded, and informal categories of a more encompassing *underground economy*. While there is much overlap, these categories appear to represent a continuum ranging from those most thoroughly in violation of the law to those that merely forgo its protections. Castells and Portes (1989) clearly differentiate informal from illegal by separating processes of production and distribution from the product itself. Thus, a perfectly legitimate product or service may be produced informally in violation of the law. Cappuchi (1989) cross-classifies three types of labor markets (nonmonetary, informal monetary, official) with the equivalent three markets for goods and services to derive a nine-cell typology. Informality in either labor or product markets creates five types of informality.

Still other analysts argue that, fundamentally, informality can only be understood on a case-specific basis. For example, Gaughan and Ferman (1987) contend "that the term *informal economy* will mean different things in advanced industrial nations from what it does in developing countries. . . . poor communities as opposed to middle-class communities, or in an urban as opposed to a rural setting" (p. 18). In keeping with this notion of the variability of forms of informal activity, Miller (1987) discusses informal economies in the plural rather than the singular.

Regardless of specific emphasis, virtually all analysts agree that informal economies, sectors, activities, and labor can only be understood vis-à-vis their relationship to their formal counterparts and in the context of their relationship to the state (Roberts 1994). There is no such thing as an informal economy in the absence of a formal economy. Portes (1994) even suggests that the extent of informal activity depends on the degree of state regulation, with greater informal development representing a response to an overly restrictive formal sector. Furthermore, while many analysts discuss informal activity as outcomes or in terms of structural elements that comprise a form of economic institution, there appears to be greater utility in conceptualizing it as a process that varies in time and space. Thus, the concepts of informality and informalization, in which economic activity is evaluated in terms of its degree of

adherence to state regulation and particular labor practices, permit an assessment of variation in levels and development of informal sectors and their shares of the larger economy.

**Forms of Informal Activity.** Empirical studies of informal activity identify a wide range of actual practices. Ultimately, what is included and what is excluded depends on the specific definition and the specific context for the activities, as discussed above. Nevertheless, certain broad categories are regularly found in the literature. They can be categorized in terms of what type of exchange relationship is involved and the specific kinds of goods and services transacted. Thus, goods and services can be exchanged for income that goes unreported in any official venue, as exemplified by the babysitter or maid who gets paid in cash that is not taxed or reported by either employer or employee. They may be bartered and exchanged directly for the same or other goods or services, as in the trade of child-care services or child care for firewood. They may be supplied and consumed to save money or to self-provision, as when members of a household raise a garden and preserve food for home consumption.

Specific goods, services, and activities vary from place to place. Urban dwellers are unlikely to raise livestock or cut wood for home consumption, but they are no less likely than their rural counterparts to engage in child care or household and auto repair (Tickamyer and Wood 1998). Sweatshops and street vendors are most likely to be found in large urban labor markets (Gaber 1994; Sassen-Koob 1989), but industrial subcontracting and home assembly operations are well documented in rural communities as well (Gringeri 1994).

One of the problems in specifying types of informal activity is distinguishing informal from related forms of work and exchange. At the margins it may be extremely difficult to draw this line. For example, where does engaging in an activity for recreation or personal satisfaction end and self-provisioning begin? Should the avid gardener whose vegetable consumption is a by-product of the hobby be classified as engaged in informal work? Where should volunteer work done for charitable purposes be classified? Informal labor overlaps with many other forms of waged and unwaged labor, including social and generational

reproduction, self-provisioning and subsistence activities, consumption work, and community service as well as simple commodity production, self-employment, and subcontracting.

### **SOCIAL, SPATIAL, AND TEMPORAL DIMENSIONS**

The expanding literature demonstrates that informality is not limited to a particular time, place, type of economy, or population group. Nevertheless, clear themes that focus on specific groups and places emerge in the research literature. Some of it is empirically based; some is speculative and yet to be tested or verified regarding who participates, where, and under what circumstances; and some is contradictory or unfounded. The most common assumption is that informal labor involves categories of workers who are disadvantaged in some manner and, therefore, are shut out from the opportunities available in the formal sector. For example, informal workers have been identified as particularly prevalent among racial and ethnic minorities, immigrants, women, and children (Hoyman 1987; Portes et al. 1989; Sassen-Koob 1984). Other studies would add rural residents (Fitchen 1981; Jensen et al. 1995) and the urban poor (Edin and Lein 1998).

While there are many specific accounts and case studies to illustrate the importance and workings of informal labor and sectors among these specific populations and locations, there is little evidence at this time to suggest that informal work is any more prevalent in one group compared to another. Furthermore, it is not always clear from these accounts whether informality adheres to the group or to the type of jobs certain groups are more likely to hold. For example, women are sometimes identified as more likely to operate in the informal sector, as is work that is typically female, such as child care and domestic service. If it is the case that women are more likely to work informally, is it because of gender or because they hold gendered jobs?

Informalization is often assumed to be increasing in both developing and developed countries as a response to the social and economic restructuring that has transformed and linked core and peripheral economies in the last several decades. Globalization, the growth of transnational

capitalism, and increased international competition have led capital to seek more flexible production processes, to escape from state regulation and taxation, and to reduce labor costs. This puts pressure on labor–capital agreements, benefits available from the welfare state, and the general availability of employment that pays a living wage, has social benefits, and observes rules and regulations. The result is increased self-employment, subcontracting, and, presumably, informal employment as employers seek to cut the costs of doing business and employees are forced to find any means of subsistence. Yet here, too, there is little empirical confirmation of what are largely untested hypotheses. For example, it is not yet established whether informal activity is more prevalent among different class and income groups, whether it is in fact a substitute for formal work, or whether it serves as a supplement to increase or stretch income. Instead, there is a rich tradition of case studies of different regions, communities, and industries that are informative but case-specific, lacking both comparability or comprehensiveness. Thus, it is premature to conclude that informality is increasing or that it is especially prevalent in some groups or places compared to others.

### **RELATIONSHIP TO THE ETHNIC ECONOMY**

While informal economic activity is not limited to a particular market segment or group, researchers have shown particular interest in those forms of informal activity that occur in ethnic economies. Part of the reason for this is that the extensive literature on the ethnic economy offers a unique view of informal work given by researchers whose main purpose is not to examine the informal economy per se but to study ethnic exchange relationships (rather than class relationships) in the secondary labor market. This research shows that ethnic economies provide opportunities for unique forms of informal activity not commonly found in other sectors. However, the importance that is placed on informalization in the ethnic economy literature should in no way imply that the two economies are always congruent, although some enclave economies are almost entirely informal (Stepick 1989).

Ethnic economies are labor markets characterized by sectoral specialization of industries owned

and operated by entrepreneurs from the same ethnic or immigrant group and their co-ethnic workers. While ethnic and informal economies are not synonymous, the prevalence of work outside the formal sector within ethnic economies suggests that conditions in these labor-market areas necessitate, or at least are particularly well suited for, informal economic activity.

The reasons for this vary. Since most workers in the ethnic economy are immigrants from developing countries or ethnic minorities that face discrimination in the formal labor market, they often provide a pool of available cheap labor for the informal market. That is, the labor supply is cheap as long as the work remains part of the informal economy, since the cost of labor rises when wages are regulated. Enterprising businessmen and women can make use of this inexpensive labor, keeping their costs low by hiring "off the books" or paying piece rates for contracted work that may or may not be reported by the worker as income.

Since ethnic economies are also marked by subcultural norms and preferences, informal activity may also allow for cultural variation in working conditions. In the case of Latin American immigrant women, for example, the myriad home working opportunities within the ethnic economy allow mothers to earn needed income while still tending to their children and other domestic tasks. In fact, there is evidence that some home-based industries emerged in response to the demands of experienced female laborers who wanted to earn extra income but expressed a cultural preference to stay at home (Fernandez-Kelly and Garcia 1989).

While the magnitude of informal economic activities within ethnic economies is not known, studies of the garment industry in New York and Los Angeles, a sector that is almost exclusively part of the informal economy, suggest that many of the workers in the industry are unregistered and working out of their homes (Fernandez-Kelly and Garcia 1989; Sassen-Koob 1989; Stepick 1989). This is also true of the home-based electronics production industry (Lozano 1989). Latin American immigrant women comprise the largest portion of this work force that engages in the unreported assembling of electronic components at home on a piece-rate basis. The informal economy may also be responsible for the viability of these industries.

While the garment industry faltered in the 1980s, textile businesses owned by immigrants who sub-contracted homeworkers thrived.

Ethnic enclaves are a special condition of the ethnic economy that arises when an ethnic economy is located in a geographically concentrated area. This geographic concentration can also give rise to unique opportunities within the informal sector. In enclaves such as Washington Heights in New York City, Chinatown in Los Angeles, and Little Havana in Miami, the demand for capital for entrepreneurial ventures has created possibilities for informal money-lending operations. These tend to take the form of either rotating credit associations, whereby residents contribute small sums of money that are pooled and then loaned at low interest rates to other residents, or personal loans from one wealthy entrepreneur to a potential small-business owner. These lending transactions occur outside of state and federal lending laws and are often made without any written agreement by the participants. Social pressure alone seems sufficient to guarantee repayment. This is consistent with Roberts's (1994) contention that interpersonal networks regulate the informal economy.

Ethnic enclaves also generate a heightened demand for goods typical of those widely available in an immigrant's home country. This demand for ethnically defined goods creates unique opportunities for informal work. Enterprising men and women often produce foodstuffs and other products for sale out of their homes. Some also find customers among other immigrants who agree to resell their products in their retail shops and *bodegas*. The informal production of goods insures that costs to both the retailer and the customer remain low, insuring future demand.

#### WAYS OF MEASURING AND STUDYING

This brief discussion of informal activity in ethnic economies illustrates only one variation in the form and scope of this phenomenon and underscores the importance of further research to gain a greater understanding of how informal economic processes impact individuals, households, and labor markets. As previously noted, recognition that informal economic activity is widespread in developed as well as developing, urban as well as rural,

contexts is just beginning to emerge. Consequently, what is known about the size of the informal economy, its participants, its determinants, and its relationship to the formal sector remains limited.

Most of what we currently know about informal work comes from ethnographic and field studies using in-depth interviews and other qualitative data-collection techniques. These provide a rich and varied collection of information about the nature and meaning of informal work in different regions. These regions include remote rural areas such as the Ozarks (Campbell et al. 1993), rural Pennsylvania (Jensen et al. 1995), and the Southwest (Roberts 1994). Other researchers have also looked at urban subpopulations, including Latin American immigrants in New York City (Sassen-Koob 1989) and in Miami (Stepick 1989). These studies offer an intensive look at the kinds of informal activities that occur, who does them and why, and when people engage in informal economic activities.

Case studies and other qualitative methodologies have provided a rich source of data on specific groups and locales, but their limited geographic scope means that there is little information on informal economic activity in larger populations and subpopulations, thus limiting our understanding of the prevalence of such activities across rural and urban places. Furthermore, there appear to be inconsistencies about the nature of informal work across areas, and case studies can only begin to explore these seeming differences. For example, studies of informalization in rural areas suggest that bartering and self-provisioning are both a means of supplementing a meager income and a normative way of life. On the other hand, urban studies depict informal work as an activity engaged in when the avenues to work in the formal labor market are blocked.

Attempts to quantify informal activity for use in policy making have been made by several national and international agencies, although the methods employed are often crude. Portes (1994) separates these strategies into categories based on the following approaches: (1) the labor market, (2) the very small enterprise, (3) macroeconomic discrepancies, and (4) household consumption.

The first three approaches rely on data taken from industry reports and measures of national

economic activity such as the gross domestic product. The labor-market approach entails assigning all labor as informal if it is reported as self-employment in areas not typically assumed to be high status or high paying. The very small enterprise (VSE) approach, like the labor-market approach, is also based on industry data. Instead of focusing on self-employment, however, VSE-based estimates are predicated on the notion that all businesses employing fewer than ten workers are engaging in informal activity. In the third approach, discrepancies in different measures of similar economic activity are attributed to the informal economy.

The household consumption approach was developed in 1987 in an effort to surpass the limitations involved in estimating informal activity based on reported, presumably formal, production activities. Estimates of informal economic activity, using this approach, are based on consumer reports of "the amounts spent . . . on goods and services acquired off the books or on the side" (Smith 1987). This method is unique in that it focuses on consumption; however, it is also limited in that it assumes that all products are consumed by individuals rather than firms.

In recognition of the inability to generalize to larger geographic areas based on ethnographic data, in acknowledgment of the limitation of the assumptions entailed in current estimation strategies, and in order to limit the costs entailed in labor-intensive qualitative methodologies, some researchers are turning to more direct survey methodology to examine informal economic processes. Recent studies using survey methodology include work by Jensen and colleagues (1995), Tickamyer and Wood (1998), and Tolbert and colleagues (1996). They note that in order to provide systematic, representative, and comparable data across different locales on who participates and how extensively, where and when participation occurs, and under what circumstances, different data are needed. The goal of these researchers is to provide the baseline data that will enable future researchers to determine if informal activity is increasing or decreasing over time and with changing economic conditions. They are also interested in exploring the relationship between informal work and other social processes, including the relationship to formal labor-market participation, to household, family, and community structures and processes, and to human and social capital formation.



These newer studies are significant because they demonstrate that collecting direct information on informal work using structured survey techniques is possible despite the elusive nature of such activities, their often semilegal status, and respondents' seeming lack of awareness of the extent of their participation in informal work. These studies also underscore the need for better, more extensive tests of survey methods both to determine optimal strategies for using survey methodology on this topic and to provide definitive answers to the many questions posed about informal labor.

### CONCLUSIONS

While sociologists, economists, and other researchers continue to provide information on informal work, there are many issues that remain unaddressed. These issues include questions regarding the prevalence, importance, and nature of informal activities. Why do people participate in the informal economy? How important is it to their economic survival? Is it vital or is it more socially driven, carrying little expectation for immediate economic reward? What percent of households engage in informal work, and which members of the household are most likely to be the workers? What are the economic returns to participation? Do formal and informal work complement or substitute for each other? Do people turn to the informal economy as a last resort in times when formal work is curtailed, or are certain types of people likely to engage in both formal and informal labor? And, finally, how is the probability of informal work shaped by individual, household, and labor-market characteristics?

Future research will need to address these issues. Furthermore, more studies will be needed at the national and international level (including comparative research) in order fully to understand the importance of the informal economy in shaping the larger economy as a whole. In order to understand this, it is essential to understand where the informal economy fits vis-à-vis the formal economy and how it has evolved. Is the informal economy merely a pure market response to unwanted government interference in commerce, or is it merely a reflection of expected social relationships in a given place and time? Most likely, the informal economy arises from a combination of

factors including poverty and normative pressures to help neighbors and for self-provision. A purely formal economic focus, however, limits our understanding of the range of possibilities entailed in informal work and the impact that such work has on shaping interpersonal processes. Furthermore, by limiting our measures of economic growth, household and individual well-being, inequality, and production to official statistics of formal economic activity, we see only a fraction of the entire picture. This is particularly problematic when policies that attempt to address these issues are shaped from these fragmented images.

Since modern economic policy has been shaped almost exclusively by focusing on the formal economy, conventional definitions of employment, unemployment, poverty, and economic growth may be insufficient to explain the myriad of issues that emerge when informal work is added to the picture. As it has become increasingly recognized that informal economic activity is vital to many individuals and households in both developed and developing countries, the need for more research has become more pressing. By addressing many of the unanswered questions that remain about the informal economy, policy makers can begin to address the labor-market problems that may force workers into informal economic activity, to regulate those activities that exploit certain segments of the work force, and to understand the importance of both formal and informal work as a means of improving living standards.

### REFERENCES

- Campbell, R. R., J. C. Spencer, and R. G. Amonker 1993 "The Reported and Unreported Missouri Ozarks: Adaptive Strategies of the People Left Behind." In T. Lyson and W. Falk, eds., *Forgotten Places: Uneven Development in Rural America*. Lawrence: University Press of Kansas.
- Cappechi, Vittorio 1989 "The Informal Economy and the Development of Flexible Specialization in Emilia-Romagna." In A. Portes, M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Castells, Manuel, and Alejandro Portes 1989 "World Underneath: The Origins, Dynamics, and Effects of the Informal Economy." In A. Portes, M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies*

- in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Edin, Kathryn, and Laura Lein 1998 *Making Ends Meet*. New York: Russell Sage Foundation.
- Feige, Edgar 1990 "Defining and Estimating Underground and Informal Economics: The New Institutional Economics Approach" *World Development* 18(7):989-1002.
- Fernandez-Kelly, M. P., and A. M. Garcia 1989. "Informalization at the Core: Hispanic Women, Homework, and the Advanced Capitalist State." In A. Portes, M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Fitchen, Janet M. 1981 *Poverty in Rural America: A Case Study*. Boulder, Colo.: Westview.
- Gaber, John 1994 "Manhattan's fourteenth Street Vendors' Market: Informal Street Peddlers' Complementary Relationship with New York City's Economy." *Urban Anthropology*. 23(4):373-407.
- Graughan, Joseph, and Louis Ferman 1987 "Toward an Understanding of the Informal Economy." *Annals of the American Academy of Political and Social Science* 493 (September): 15-25.
- Gringeri, Christina 1994 *Getting By: Women Homeworkers and Rural Economic Development*. Lawrence: Kansas University Press.
- Hart, Keith 1973 "Informal Income Opportunities and Urban Employment in Ghana." *Journal of Modern African Studies* 11:61-89.
- Hoyman, Michele 1987 "Female Participation in the Informal Economy: A Neglected Issue." *Annals of the American Academy of Political and Social Science* 493 (September):64-82.
- Jensen, Leif, Gretchen T. Cornwell, and Jill L. Findeis 1995 "Informal Work in Nonmetropolitan Pennsylvania." *Rural Sociology* 60(1):67-107.
- Lozano, B. 1989 *The Invisible Work Force: Transforming American Business with Outside and Home-Based Workers*. New York: Free Press.
- Miller, S. M. 1987 "The Pursuit of Informal Economies." *Annals of the American Academy of Political and Social Science* 493 (September):26-35.
- Mingione, Enzo 1991 *Fragmented Societies: A Sociology of Economic Life Beyond the Market Paradigm*. Cambridge, Mass.: Basil Blackwell.
- Portes, Alejandro 1994 "The Informal Economy and Its Paradoxes." In N. J. Smelser and R. Swedberg, eds., *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press.
- , Manuel Castells, and Lauren A. Benton, eds. 1989 *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Roberts, Brian 1994 "Informal Economy and Family Strategies." *International Journal of Urban and Regional Research* 18(1):6-23.
- Sassen-Koob, Saskia 1989 "New York City's Informal Economy." In A. Portes, M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Smith, James D. 1987 "Measuring the Informal Economy." *Annals of the American Academy of Political and Social Science* 493 (September):83-99.
- Stepick, Alex 1989 "Miami's Two Informal Sectors." In A. Portes, M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore: Johns Hopkins University Press.
- Tickamyer, Ann R., and Teresa A. Wood 1998 "Identifying Participation in the Informal Economy Using Survey Research Methods." *Rural Sociology* 63(2):323-339.
- Tolbert, Charles, L. Tobin, and D. Haynie 1996. "Informal Economic Activity in Louisiana Farm and Rural Households." Paper presented at the annual meetings of the Rural Sociological Society, Des Moines, Iowa.

ANN TICKAMYER  
STEPHANIE BOHON

## INFORMATION SOCIETY

Increased reliance on activities directly associated with the production, distribution, and utilization of information has led to characterizing many advanced countries of the world as information societies. The term *information society* and related concepts, such as information age and knowledge economy, describe a social system greatly dependent on information technologies to produce and distribute all manner of goods and services. In contrast to the industrial society, which relied on internal combustion engines to augment the physical labor of humans, the information society relies on computer technologies to augment mental labor.

Trends in labor-force composition both define and measure the extent to which a nation can be described as an information society. Machlup

(1962) was perhaps the first to describe U.S. society in these terms. He estimated that nearly one-third of the labor force in 1958 worked in information industries such as communications, computers, education, and information services, which accounted for 29 percent of the gross national product (GNP). Using a slightly different methodology, Porat (1977) estimated that information activities had risen to just under half of the U.S. GNP by 1967.

A defining attribute of the information society is the search for improvements in productivity through substituting information for time, energy, labor, and physical materials. In practical terms, this means supplying workers with computerized workstations that are networked to other workstations through intranets as well as the Internet. It allows the use of software to reprogram equipment in distant locations, and it often eliminates the physical delivery of messages and even products. These changes are aimed at making organizational production, distribution, and management decisions more efficient. An early indicator of the extent to which industries sought productivity improvements through the use of information equipment is that whereas only 10 percent of all U.S. investments in durable equipment was spent on the purchase of computers and communications equipment in 1960, that investment increased to 40 percent in 1984 (U.S. Congress 1988) and is now much higher.

The concept of postindustrial society, developed most notably by Daniel Bell (1973), anticipated development of the information society. The term *post-industrial* described the decline of employment in manufacturing and an increase in service and professional employment noted by Machlup (1962) and Porat (1977). Knowledge and information were viewed by Bell as the strategic and transforming resources of postindustrial society, just as capital and labor were the strategic and transforming resources of industrial society.

Advances in the capabilities of information technologies to process large quantities of information quickly have been a crucial factor in the development of the information society. These technologies are of two types, computer power and transmission capability. Development of inexpensive silicon integrated circuits containing as many as a million transistors on a single chip had

already been achieved by the mid-1980s, making it possible to pack enormous information processing power into very little space. As a result, desktop microcomputers gained the processing power comparable to the largest mainframe computers of the previous decade. Computational power continued to increase by a factor of ten every five to seven years in the 1990s (Martin 1995). Improvements in microprocessor technology, coupled with developments in parallel processing, storage methods, input-output devices, and speech recognition and synthesis, have continued to increase dramatically the nature, scale, and speed of tasks that can be accomplished on computers. All this has happened while prices for computers declined in an equally dramatic way.

Corresponding advancements have occurred in photonics as a result of the development of laser technology and ultra-pure glass fiber. These developments resulted in the ability to transmit enormous quantities of information long distances on tiny optic fibers without amplification. By the mid-1980s, AT&T Laboratories had transmitted 420 million bits per second of information over 125 miles without amplification. Advancements of this nature, as well as the use of satellites, made it possible for computers located thousands or tens of thousands of miles apart to share large amounts of information nearly as quickly and effectively as would happen if they were located in the same building. Like the price for computing power, the price for transmitting large amounts of information from one place to another also has declined.

These changes in computational and transmission power have made possible new ways of interacting and doing business. Automatic teller machines located on one continent can dispense cash from a bank located on another continent. Cash registers and gasoline pumps are connected to a telecommunications system so that credit card balances can be checked before making a sale. By pressing numbers on a touch tone telephone and without talking to another human being, products can be purchased, library books can be renewed, newspaper delivery can be started or stopped, survey questionnaires can be answered, and money can be transferred from one account to another. These examples illustrate not only the substitution of information technology applications for

human labor, but also the creation of services that could not previously be provided.

However, these developments pale beside the huge capability being unleashed by development of the World Wide Web. Once a system for the exchange of simple text messages among scientists, it has now expanded to a required form of communication for many, if not most, businesses and professionals. It is estimated that as many as 160 million users are now connected to the Internet, of which nearly half are in the United States and Canada. The rapid growth in Internet use in the mid-1990s has led to increases in connections among geographically dispersed work groups and to new methods for the selling of goods and services.

Development of the information society happened neither suddenly nor without warning. According to Beniger (1986), its roots go back to a crisis of control evoked by the Industrial Revolution in the late 1800s. Industrialization speeded up material-processing systems. However, innovations in information processing and communications lagged behind innovations in the use of energy to increase productivity of manufacturing and transportation systems. Development of the telegraph, telephone, radio, television, modern printing presses, and postal delivery systems all represented innovations important to the resolution of the control crisis, which required replacement of the traditional bureaucratic means of control that had been depended on for centuries before.

However, an entirely new stage in the development of the information society has been realized through advances in microprocessing technology and the convergence of mass-media telecommunications and computers into a single infrastructure of social control (Beniger 1986). An important factor in this convergence was digitalization of all information, so that distinctions between types of information such as words, numbers, and pictures become blurred, as does communication between persons and machines on the one hand, and between machines on the other. Digitalization, therefore, allowed the transformation of information into a generalized medium for processing and exchange by the social system, much as common currency and exchange rates centuries ago did for the economic systems of the world. Combining

telephone, television, and computer into a single device represents an important likely and practical consequence of this convergence.

Quite different views exist about the possible effects of the development of a full-fledged information society (Lyon 1988). One view is that it will empower workers, providing direct access to opportunities unavailable to them in an industrial society except by high organizational position and proximity to centralized positions of power. In 1985, Harlan Cleveland described information as being fundamentally different from the resources for which it is being substituted; for example, it is not used up by the one who consumes it, hence making its use possible by others. It is also easily transportable from one point to another, a characteristic made strikingly clear by the rapid rise of the World Wide Web. Cleveland argued that the information society would force dramatic changes in longstanding hierarchical forms of social organization, terminating taken-for-granted hierarchies based on control, secrecy, ownership, early access, and geography. A similar view was provided by Masuda (1981), writing in a Japanese context, who envisioned the development of participatory democracies, the eradication of educational gaps between urban and rural areas, and the elimination of a centralized class-based society.

A more pessimistic view of the consequences of knowledge as the key source of productivity was offered by Castells (1989). Fundamentally, the new information infrastructure that connects virtually all points of the globe to all others allows for great flexibility in all aspects of production, consumption, distribution, and management. To take advantage of the efficiencies offered by full utilization of information technologies, organizations plan their operations around the dynamics of their information-generating units, not around a limited geographic space. Individual nations lose the ability to control corporations. Information technologies, therefore, become instrumental in the implementation of fundamental processes of capitalist restructuring. In contrast to the view offered by Cleveland, the stateless nature of the corporation is seen as contributing to an international hierarchical functional structure in which the historic division between intellectual and manual labor is taken to an extreme. The consequences

for social organization are to dissolve localities as functioning social systems and to supersede societies.

There can be no doubt that the use of information technologies is significantly changing the structures of advanced societies. Yet it would be a mistake to think of the use of information technologies as a cause only, and not a consequence, of changes in societal structures. Laws have evolved in the United States in an attempt to regulate rates that can be charged for cable television, how much can be charged by providers for telecommunications services to schools and libraries, and what levels of telephone transmission services must be provided to individual households as a part of universal service. The influence of people's values is also being exerted on the extent and means by which confidentiality of data records must be protected; it also is being exerted through state and local laws mandating the accessibility of computers to schoolchildren.

Our rapid evolution to an information society poses many important sociological questions about how our increased dependence upon information technologies influences social interaction and other aspects of human behavior. The ability to transmit work across national boundaries, even the high likelihood that information essential to the operation of a nation will be stored outside rather than within a country, raises important questions about what is essential for preserving national sovereignty. The ability to control operations at long distance encourages an even greater division of labor among nations. As a result, labor unions may become powerless in the face of the ability of corporations to move production activities across national boundaries (Lyon 1988). And, just as elements of national society have weakened in the face of globalization, a set of counterforces have been unleashed whereby identity-based social movements compete to fill the void of power and control (Castells 1997). It is important for sociologists to seek an answer to the question of how the increased reliance on information technologies affects the sovereignty of individual nations and related social movements.

The information age provides new challenges for nearly all areas of sociology. It influences how and from whom we learn, with lifelong distance education changing the once essential learning

triangle of professor, student, and classroom. New types of crime, such as creating and spreading computer viruses, have been elevated from curiosities to major threats to the functioning of organizations, society, and the world order. The impact on people's self-concepts may also be substantial. From preschool on, computers have become part of the sociological and psychological development of children, the potential effects of which have yet to be fully understood (Turkle 1984). This interaction with computers now extends via the Internet to others with computers, so that the core sociological concept of social interaction must make room for electronic, long distance interaction and its consequences. Many adults now spend far more of their lives in interchange with computers than with another technological development, the automobile, which also dramatically changed people's lives in the industrial society.

Information technologies also have the potential for breaking down boundaries of individual communities, making it possible for people to bypass forming traditional community ties, unless extraordinary efforts are made to maintain them (Allen and Dillman 1994). Thousands of new job titles are added to the occupational structure of countries, while other job titles disappear. Robert Reich, for example, describes the evolution of jobs into three broad types—routine work, in-person service workers, and symbolic analysts (1991). The latter are theorized to be the creators of "value" in the information age, replacing land, plants, and equipment as the most valued production resource. These anticipated changes, to the extent they occur, provide the basis for evolution of a new class structure in society, based more upon educational accomplishment than upon the ownership of material resources.

Some have argued that we are evolving into a world of the information-rich and information-poor, with computer access and skills forming the great divide (Castells 1997; Lyon 1988). Even though computers seem omnipresent in society, they are present in only a minority (about 40 percent) of U.S. households and only half of those households have e-mail or Web access (National Telecommunications Information Administration [NTIA] 1998). To the extent that computers with Web connections shift from an optional way of

accessing important information and purchasing good and services to a mandatory means of obtaining competitive prices, a case can be made that class differences will expand.

It's appropriate that Daniel Bell, besides being one of the earliest prognosticators of the information age, also has more recently described quite different ways in which it could evolve (1989). He points out that the telecommunications revolution makes possible an intense degree of centralization of power if the society decides to use it in that way. However, because of the multiplicity, diversity, and cheapness of the modes of communication, decentralization is also possible. One of the important challenges for sociology is to understand which of these visions will prevail and why.

REFERENCES

Allen, John C. and Don A. Dillman 1994 *Against All Odds: Rural Community in the Information Age*. Boulder, Colo.: Westview Press.

Bell, Daniel 1973 *The Coming of Postindustrial Society: A Venture in Social Forecasting*. New York: Basic Books.

— 1989 "Communication Technology: For Better or for Worse." In Jerry L. Salvaggio, ed., *The Information Society: Economic Social and Structural Issues*. Hillsdale, N.J.: Lawrence Erlbaum Associations.

Beniger, James R. 1986 *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, Mass.: Harvard University Press.

Castells, Manuel 1989 *The Information City*. Cambridge, U.K.: Basil Blackwell.

Castells, Manuel 1997 *The Information Age Economy, Society and Culture: Volume II. The Power and Identity*. Malden, Mass.: Blackwell Publishers.

Cleveland, Harlan 1985 "The Twilight of Hierarchy: Speculation on the Global Information Society." *Public Administration Review* 45:185-196.

Lyon, David 1988 *The Information Society: Issues and Illusions*. London, U.K.: Polity Press.

Machlup, Fritz 1962 *The Production and Distribution of Knowledge in the United States*. Princeton, N.J.: Princeton University Press.

Martin, William J. 1995 *The Global Information Society*. London: Aslib Gower.

Masuda, Yoneji 1981 *The Information Society as Post-industrial Society*. Tokyo: Institute for the Information Society; Bethesda, Md.: World Future Society.

National Telecommunications Information Administration 1999 "Falling Through the Net: Defining the Digital Divide." Available at: <http://www.ntia.doc.gov/ntiahome/net2/falling.html>

Porat, M.U., et al. 1977 *The Information Economy. OT Special Publication 77-12*. Washington, D.C.: U.S. Department of Commerce.

Reich, Robert 1991 *The Work of Nations*. New York: Alfred A. Knopf.

Turkle, Sherry 1984 "*The Second Self*": *Computers and the Human Spirit*. New York: Simon and Schuster.

U.S. Congress, Office of Technology Assessment 1988 *Technology and the American Transition: Choices for the Future*. OTA-TET-ZA3. Washington, D.C.: U.S. Government Printing Office.

DON A. DILLMAN

INHERITANCE

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

Statutory and common laws governing inheritance have a profound effect on the formation or dissolution of household structures and the patterns of inheritance transfers over generations. A society, for any political or social reason, can initiate and promulgate a law or civil code controlling inheritance; inevitably, new laws or codes cause drastic change in the structure and functions of family systems.

In order to understand the subject of inheritance, it will be necessary to define many terms. To *inherit* is by law to receive property, resources, or, often, status from an ancestor at her or his decease or to take by intestate succession or by will. *Intestate succession* is a transfer of resources according to legal procedures that control distribution of the resources when there is no will. A person is *intestate* when he or she has not made a will. In a *will*, usually a written document, a person (or *testator*) makes a deposition of his or her property, and the deposition takes effect, at least in modern societies, upon the testator's death. A will is changeable and revocable during the lifetime of a testator.

*Takers* are successors or beneficiaries. *Impartible inheritance* is a situation, established by statute,

in which the property and resources are indivisible and are given to one person (*devisee*); this type of inheritance is likely to occur within family households primarily in rural areas of most historic societies, especially in Eastern and Western Europe. *Partible inheritance*, dividing assets for conveyance to one's heirs, is linked to nuclear family households found in those locations where there is accelerated industrialization and urbanization.

Any system of inheritance promotes the continuity of family and societal structures over generations. The transfer of resources from the older to the younger generation helps maintain a family's position and power in the social order. Such transfers also provide stability to existing societal caste and class arrangements and ordinarily are enconced deeply in tradition and myth. It is more likely that inheritance systems function to perpetuate existing social structures than to change society's organizations and institutions.

#### INHERITANCE, LAWS, AND CHANGE

Those most affected by changes in laws are persons with little power in society. This is evident in eighteenth- and nineteenth-century European societies, in which the promulgation of laws and decrees drastically affected the structure and practices of rural families. For example, Gaunt (1983) indicates that in nineteenth-century central Europe, in a section that is now part of Czechoslovakia, the decree to conscript unmarried men for military duty caused parents of serf families to encourage early marriage of their children. As a consequence, the incidence of complex family households increased, and there were changes in the patterns of inheritance succession. During a labor shortage in eighteenth-century Poland, fiefdom rulers encouraged endogamous marriage and attempted to restrict their chattels from leaving their villages. The economy was based on household unit sharecropping. The peasants encouraged complex family households consisting of two or more conjugal units, while the Polish lords favored neolocal household residence (Kochanowicz 1983). The inheritance outcomes of these efforts to circumvent political, economic, and social measures, often repressive in intent, affected existent family structures, their forms, and inheritance patterns.

The possession of equities, property, possessions, resources, and land (especially in the case of European peasants) determined whether these would compose an impartible or partible inheritance. For example, an impartible inheritance pattern is most likely when land is the primary family asset. In addition, the transfer of authority over these rights depended upon the timing of "stepping down," a process deeply embedded in the cultural traditions of the society (Gaunt 1983; Plakans 1989). Stepping down usually occurred through a retirement contract, which was essentially a will that indicated the conditions of the transfer from parents to children. This *inter vivos* phenomenon, a conveyance of property and other equities while the individual was alive and engaged in stepping down, was "one made of preserving intergenerationally the match between family and the property that provides its livelihood" (Sorensen 1989, p. 199). Stepping down, or disengaging, usually occurred when the oldest son married or when the parents were near or at retirement age. Stepping down was invoked by law or tradition, and it resulted in variations in inheritance patterns and in the organization of the life course of various family members.

Sorensen (1989) provides a detailed retirement contract of his great-grandfather on his mother's side, a prosperous farm family in the western part of Jutland, Denmark. The transfer of this medium-sized farm, substantially undervalued, when Laurids Poulsen was fifty-six, was apparently to make certain that the property remained in the family to buttress its position in the society.

It is not explicitly stated in the Poulsen retirement agreement that the heir, Alfred, would care for his parents in their declining years. The contract's provisions enabled Laurids and his wife, Maren, to be independent. Yet in this situation and, as Sorensen (1989) indicates, in Scandinavia since before the Middle Ages, transferring property to a son implied the promise of care in old age. Thus, there existed two motivations for conveying property to a son, but the stronger proved to be the hope of maintaining the property in the family over generations.

In former ages, conveying property and its accompanying position, status, and power was the intent of inheritance patterns. The preservation and maintenance of the family unit was the goal.

Impartible inheritance expressed this intent of maintaining the family property and social position, but it created obvious winners and losers. The transfer of a family's property to a single heir meant that other family members could leave to seek opportunities elsewhere or remain with the family enterprise in a subservient position.

The possibilities for fuller expression of family members' abilities, where talent and skills and not family membership determined the individual's life course trajectory, would come at a later period. Partible inheritance reduced the requirement that family members subordinate their desires, interests, and expectations to those of the family unit. Changes in a society's demographics, such as fertility and mortality, and changes in the means of producing goods and services resulted in lessened need for impartible inheritance. Partible inheritance became identified with the modern period of Western civilization.

### MODERN PERIOD

Rules for succession and rules for inheritance of property are related but distinct from one another. Inheritance of property usually follows lines of succession to social position. Codified systems of secular law governing inheritance and status succession emerge in complex societies and are sufficiently precise and uniform to meet the majority needs of the population (Radcliffe-Brown 1935).

For urbanites in complex societies such as those in Europe and the United States, the transfer of land, dwelling unit, tools, and equipment is less critical to the survival and maintenance of the family over generations than it is for the rural resident. Economic assets other than land and buildings—that is, personal mementos and possessions—become the content of such transfers.

Rural landholders of modest means were unable to effect a pattern of impartible inheritance. Increasing societal complexity nurtured corporate agriculture in preference to the family farm. The reduction in the number of family-owned farms, the increasing dependence of farm family members upon the larger, mostly urban society for jobs that are not located close to the farm dwelling, price supports, and payment for nonproduction have diminished the possibilities for maintaining

the family farm over generations. Partible inheritance has become the norm.

This shift to partible inheritance under new rules of succession meant that for the first time the claims of the surviving spouse outweighed those of the surviving kin of one's lineage, a reversal of the pattern found in eighteenth- and nineteenth-century Europe and in primitive societies (Benedict 1936; Hoebel 1966). Under U.S. state statutes governing intestacy the surviving spouse and children share in the estate. The spouse receives at least one-third, depending on the number of surviving children and the specifics of the state law governing inheritance. If there is no surviving spouse the children share and share alike. Where there is no surviving spouse or children the estate passes to grandchildren. In the absence of grandchildren the next to receive are grandparents, then siblings, then more distant relatives.

Where testacy exists the estate usually passes to the surviving spouse. This horizontal transfer in the generational line of succession is uniquely modern and represents an evolution from the Roman definition of inheritance as "succession to the entire legal position of a deceased man" (Maine 1963, p. 208). In Rome the heir functioned as a guardian or executor of the estate to perpetuate the honor and status of the deceased and family survivors and to keep intact and extend the estate's holdings. The stepwise shift from decedent to surviving spouse and subsequent vertical transfer of equities to children after the death of the surviving spouse resulted in a family system based more on individual relationships, feelings, perceptions, and interactions and less on tradition, primogeniture, and maintaining properties and estates.

The pattern of conveying all property to the spouse, to be discussed further in this article, is a practice seemingly not in consonance with the concept and exercise of testamentary freedom.

### TESTAMENTARY FREEDOM

Testamentary freedom, the individual's right to will away property to persons outside the family or to distribute to a number of heirs and legatees related by blood, consanguinity, or adoption, is a fundamental Anglo-American concept of the U.S. inheritance system. The primogeniture system, which passed all property to a single heir and was



most suited for the wealthy, was replaced in eighteenth-century England by a new law of the land—testamentary freedom. This occurred at the time England disposed its feudal land tenure system. The 1789 French Revolution, in keeping with its ideology of justice, freedom, equality, and fraternity, and for more concrete social and political reasons, enacted laws requiring equal distribution of a deceased person's assets among surviving children. Colonial settlers in the United States undoubtedly inspired by and enamored of the changes in English and French societies, brought with them this notion of testamentary freedom as part of their intellectual and cultural heritage.

At first blush the practice of testamentary freedom would seem at least to contradict if not destroy the major intent of an inheritance system. According to Edmund Burke, "the power of perpetuating our property in families is one of the most valuable and interesting circumstances belonging to it, and that which tends the most to the perpetuation of society itself" (1910, p. 121). The major question is whether testamentary freedom as it is practiced negates what Burke suggests is a most critical process for generational and societal continuity.

Testamentary freedom, like justice or liberty, is a relative and not an absolute condition. In practice it accommodates to family continuity over generations; a multilineal descent system; a highly differentiated society where the majority of assets owned by individuals are moveable; and values that espouse rationality, choice, freedom, and democracy. The right of an individual to dispose of property according to her or his wishes is recognized if the individual disposes property in a responsible fashion—when one takes care of one's kin, thus maximizing the possibilities of family continuity and orderly social relationships among family members. Testamentary freedom is not exercised absolutely. It accommodates to the norms of responsibility. Empirical data support the idea that compromise occurs among the interests of the individual's family and state in the exercise of testamentary freedom (Sussman, Cates, and Smith 1970).

Courts, social norms, and societal economic patterns limit the expression of testamentary freedom. Courts use the soundness of mind principle

in determining if an individual acts in a responsible fashion. Knowing what one possesses, the nature of the business in which one is engaged, and the natural objects of one's bounty are the essential components of the legal definition of a sound mind. In practice, courts almost universally consider the well-being of the family in addition to considering whether an individual has knowledge of his or her assets and successors (Cates and Sussman 1982). Neglecting or abandoning the family is viewed as unnatural, and being unnatural is equated with being of unsound mind. The media abound with cases regarding will contests involving decedents' bequests to loved pets, charities, strangers, or acquaintances. Preventing a distant relative from taking from or receiving an adequate share of an estate may result in legal action. Sussman, Cates, and Smith (1970) report a case in which charities were the major beneficiaries. The sole surviving relative, a niece, contested the will and lost, but she received a large out-of-court settlement of \$150,000. Courts, plaintiffs, and defendants normally favor out-of-court settlements because of lower economic and psychic costs. In 1965, when this case occurred, a settlement of \$150,000 was judged to be ample compensation. The settlement was also an indication that the well-being of a distant family member had been considered.

Prevalent social norms foster concern for the well-being of surviving family members. Believing that families should take care of their own and that family members should not be pauperized, the state has more than a legal interest in seeing that testamentary freedom be exercised with regard for the well-being of the family.

#### **PATTERNS OF GENERATIONAL DISTRIBUTIONS**

Complex societies have developed systems of state-wide resource transfers that have replaced in part some of the functions of the family inheritance system. These large-scale transfer systems, based on the principle of serial service, are society's way of taking care of those deemed dependent, those unable to contribute to the gross national product through gainful employment. Preponderant numbers of the very young and the elderly receive support from such transfers. The preeminent philosophical notion is that the young adult and middle

age generations, individuals ages sixteen to sixty five, pay through their earnings for government-initiated entitlement and needs programs such as Social Security, Medicare, Medicaid, educational grants, welfare payments, and so forth, programs to maintain and enhance the lives of the less fortunate. These generations may participate in these programs somewhat grudgingly, but they do so with the knowledge that they were supported during their childhood and with the expectation of being supported by these same programs in their old age. This pattern of society-wide transfers is characteristic of serial service.

Such massive transfers have reduced the economic burden of families in caring for dependent members. In some instances these programs have diminished the need for family members to provide extensive and intensive social and emotional support (Kreps 1965). Social Security pensions and other vested retirement programs, in both the private and public sectors, provide sufficient income for an increasing number of retirees. Most of them will be able to live independently or with minimal support from their relatives during their later adult years. Being economically independent or quasi-independent in old age with little need to rely upon the family's financial resources is a radical shift from earlier periods, when older relatives depended on the determinations of the inheritance system.

The society-wide transfer system based on universal taxation has not completely replaced the family inheritance system. One can view the former as an impersonal, bureaucratized, and universal system enacted primarily by statute or charter and monitored by official regulations. Participants become part of a large formal system and once they qualify are treated in a uniform manner. The use of computers and identification numbers increases the impersonality of support systems.

The family inheritance system, on the other hand, is influenced but not controlled by large-scale economic transfers. It exists and functions within a set of norms that extoll interpersonal relationships, continuity of marriage and blood lines, symbolic meanings, feelings of filial piety, nurturance, support, care, distributive justice, and reciprocity.

These institutional systems and programs that mutually support families probably condition the

bases upon which individual inheritance dispositions are made. Sussman, Cares, and Smith (1970) indicate that as a consequence of the growth of society-wide transfer programs,

*Inheritance transfers may be less a consequence of acts of sexual reciprocity, based upon what specific individuals of one generation in a family do for others of another generation, but more a function of serial service. Serial service (a concept elaborated by William Moore, 1967) involves an expected generational transfer that occurs in the normal course of events. It is expected that parents have to help their young children, and middle aged children may be called upon to give care or arrange for care of an aged and often ailing parent. This is within the cycle of life, and services of this kind are expected and are not based upon reciprocal acts. Whatever parents have in the way of worldly possessions will in due course be passed on to lineal descendants. (p. 10)*

Since the 1970s there has been a strong movement to reduce, or at best not increase, taxation to support benefit systems that assure society-wide economic transfers across generational lines. Those who support a reduction in such transfers strongly believe that the society has reached its absorptive capacity to pay increasingly higher costs for retirement and a variety of services to dependent family members. The cutback in government programs coupled with the exhortation that the government "should get off the back of family members" and the reglorification of the myth that families in ancient times cared for their own have resulted in increasing burdens for families in the care of their elderly and dependent members. This shift away from society to family responsibility for family members suggests a new look at family economic transfers and the role such inheritance plays in intergenerational relationships and in the solution of long-term care of aged and other dependent members.

#### INHERITANCE ISSUES FOR THE 1990S

Believing they do not want to be a burden to their children, older adults with economic resources and few relatives for whom to function as primary caregivers when needed are spending their inheritance on travel and other leisure activities and for

total medical and physical care as they move from independent to dependent living. A likely result is the diminishing of available funds to heirs and legatees and increased importance of family heirlooms. These gifts express the meaning and significance of the relationships between family members prior to the death of the testator.

The increasing unavailability of relatives, especially children, to care for aged family members will result in the assignment of heirs and legatees who are not related by blood or marriage. These are friends and service providers who supply social support, care, and nurturance and are "like family." Elders in the latter part of this century and well into the twenty-first, that is, members of the World War II baby boom generation, will be searching for a relative or someone to care for them in their declining years. Their drastically low reproduction rates and the consequences of the gender revolution begun in the 1960s will result in the availability of a severely limited number of immediate family members or distant relatives able or willing to provide care. Elderly people will turn to persons not related by blood or marriage—members of the individual's wider family (Marciano 1988). Wider family members are not a certain age, nor must they conform to traditional social norms; they are not related by blood or marriage. Bondings express deep friendship and voluntary informal contracted obligations and expectations. Wider family members today are on call and respond to requests for assistance immediately. They act and feel that they are at home in each other's household.

The consequence is an increasing incidence in the naming of such persons in wills and the probability that courts in the future will uphold their right to share the estate even in the case of intestacy. The courts may rule that traditional patterns of distribution to surviving family beneficiaries be modified to include those who provided care and nurturance to the deceased. The basis for such action is the notion of distributive justice, which invokes fairness. Those who have voluntarily provided services, intimacy, friendship, and care should be recognized and even rewarded. They have fulfilled the role of filial responsibility.

The dark side of this pattern is the exploitative friend, surrogate, or service provider who manipulates the care receiver and takes over the estate

through an *inter vivos* transfer or by a rewriting of the will. This sorcery, coupled with an increasing number of will contests, will keep attorneys in good financial stead and courts very busy for a long time.

Distributive justice, a just and fair distribution of resources as perceived by a testator or promulgated by the laws of intestacy, will be characteristic of transfers to family and kin as well as wider family members in the coming decades. A component of distributive justice is equality (Piaget 1932), a standard that is invoked by statutes governing intestacy. Equality is determined by the degree of relationship of the deceased to the survivor. Thus, upon the death of a parent intestate, equal shares are given to surviving children. If a spouse survives, she or he usually receives one-third of the estate, and the remaining two-thirds are distributed to surviving children.

Equity is another component of distributive justice. It implies a just and fair condition. It is not the same as equality. It is the "distribution of rewards and costs between persons" (Homans 1961, p. 74). Equity is the component most suited to arrangements for payment of rewards for incurred costs. Hence, both family and nonfamily members can expect rewards in property inheritance in relation to the costs incurred in caregiving and related activities. Such exchanges will become normative and generally accepted. Reciprocity is to be recognized and rewarded.

Portending the future, testators will increasingly enter into contracts with either family or nonfamily (wider) members, setting forth conditions and expectations of needed supports and caregiving (Hanks and Sussman 1990). Such contracts, seemingly as legal as any other contract, are likely to be challenged in the courts by those who stand to inherit under the laws of descent and succession. The rootedness of these contracts in distributive justice, with its fairness and equity principles, suggests that their validity will be sustained.

Contracts can involve *inter vivos* transfers or declarations in a will. For example, a middle-aged testator can give resources to a young relative or member of a wider family with a pay-back arrangement of care and service when needed by the benefactor. A will provision can readily be made.

The exact instrument that will enhance its legality needs to be developed in consultation with attorneys. Such contracts, similar to prenuptial agreements, drawn with both parties of sound mind, should be legal. These contracts will have provisions for modification and cancelation, like any other contract. The major point is that such contracts can reduce the concerns of an aging population regarding their life-style in very old age and diminish their total dependence upon institutional forms of care. The "inheritance contract" fosters independence and utilization of one's resources in meeting health and social needs. Caretaking as a family enterprise can reduce potentially the outlay of public monies currently expended for health and social service programs. This reduction in the economic burden of government is not tantamount to the elimination of current universal support programs. At best the inheritance contract helps provide meaningful relationships, establishes a procedure for planning and expending one's resources, and encourages individuals to rely less upon governmental institutions.

Openness in discussing wills is a new departure from the past and leads to forthrightness in discussing other family matters. Hanks (1989) reports in her sample of 111 corporate family members that 88 percent discussed their wills and 60 percent their funeral arrangements. Such openness with relatives can reduce future will contests and result in conversations and negotiations regarding foreseeable generational transfers and caregiving arrangements. The symbolic meaning attached to the passing on of jewelry, paintings, furniture, books, and other household items has been given very little attention (Sussman, Cates, and Smith 1970). Allocations of these items demonstrated to the recipients the kind of emotional and affective relationship they had with the deceased. More individuals experienced pain and depression from not receiving a promised or expected heirloom than from receiving a lesser share of property or equities. Strong feelings are aroused because of the memories connected with these heirlooms, and the deceased cannot be asked regarding her or his feelings and emotions toward the heir. Openness in discussing wills and related matters can reduce misunderstandings and misinterpretations regarding the meaning attached to the transfer of heirlooms and the trauma of not knowing.

In the 1990s and future decades inheritance will continue its patterns of transfer and distribution of properties and equities over generations. It will be a different system from that found in rural areas in historic or current time. It will continue to be insignificant in the generational transfer of status and power except for the wealthy classes. Things that will distinguish inheritance in the future from that of the past is the emergence of the inheritance contract; openness in discussing will contents with potential beneficiaries; greater inclusion of distant family relatives in wills; increased number of will contests; and increased incidence in the number of older adults who spend money on leisure activities or on their own health care rather than contribute it to their children's inheritance.

(SEE ALSO: *Intergenerational Resource Transfers*)

REFERENCES

- Benedict, Ruth 1936 "Marital Property Rights in Bilateral Society." *American Anthropologist* 38:368-373.
- Burke, Edmund 1910 *Reflections on the French Revolution and Other Essays*. New York: DuHon.
- Cates, Judith, and Marvin B. Sussman 1982 *Family Systems and Inheritance Patterns*. New York: Haworth Press.
- Gaunt, David 1983 "The Property and Kin Relationships of Retired Farmers in Northern and Central Europe." In Richard Wall, Jean Robin, and Peter Laslett, eds., *Family Forms in Historic Europe*. Cambridge: Cambridge University Press.
- Garrett, Mario D. 1995 "Filial Piety or a Filial Pie: Capital Transfer from the Elderly to Their Offspring within Southeast and Eastern Asia." In Stanley R. Ingman, Pei Xiaomei, Carl D. Ekstrom, Hiram J. Friedsam, and Kristy R. Bartlett, eds., *An Aging Population in an Aging Planet, and a Sustainable Future*. Denton, Tex.: Center Texas Studies.
- Gross, Stephen J. 1996 "Handing Down the Farm: Values, Strategies, and Outcomes in Inheritance Practices among Rural German Americans." *Journal of Family History* 21:192-217.
- Hanks, Roma 1990 "Inheritance and Caregiving: Perceptions in Veteran and Corporate Families." Unpublished paper, College of Human Resources, University of Delaware.
- , and Marvin B. Sussman 1990 "Inheritance Contracting: Policy Implications of Inheritance and Caregiving Patterns." Unpublished paper, College of Human Resources, University of Delaware.

- Hill, Gretchen J. 1995 "Inheritance Law in an Aging Society." *Journal of Aging and Social Policy* 7:57-83.
- Hoebel, Edgar A. 1966 *Anthropology*. New York: McGraw-Hill.
- Homans, George C. 1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace and World.
- Hooyman, Nancy R. 1989 "Women as Caregivers of the Elderly: Implications for Social Welfare Policy and Practice." In David E. Bregel and Arthur Bloom, eds., *Aging and Caregiving*. Newbury Park, Calif.: Sage.
- Kochanowicz, Jasalav 1983 "The Peasant Family as an Economic Unit in the Polish Feudal Economy of the Eighteenth Century." In Richard Wall, Jean Robin, and Peter Laslett, eds., *Family Forms in Historic Europe*. Cambridge: Cambridge University Press.
- Kreps, Juanita 1965 "The Economics of Intergenerational Relationships." In Gordon F. Streib and Ethel Shanas, eds., *Social Structure and the Family*. Englewood Cliffs, N.J.: Prentice-Hall.
- Maine, Henry S. 1963 *Ancient Law*. Boston: Beacon Press.
- Marciano, Teresa D. 1988 "Families Wider Than Kin." *Family Science Review* 1:115-124.
- Niraula, Bhanu B. 1995 "Old Age Security and Inheritance in Nepal: Motives versus Means." *Journal of Biosocial Science* 27:71-78.
- Peart, Nicola 1996 "Towards a Concept of Family Property in New Zealand." *International Journal of Law, Policy and the Family* 10:105-133.
- Piaget, Jean 1932 "Retributive and Distributive Justice." In *Moral Judgement of the Child*. New York, Harcourt, Brace and World. Reprinted in Edgar Borgatta and Henry J. Meyer, eds., *Sociological Theory*. New York: Alfred A. Knopf, 1956.
- Plakans, Andrejs 1989 "Stepping Down in Former Times: A Comparative Assessment of Retirement in Traditional Europe." In David I. Kertzer and K. Warner Schaie, eds., *Age Structuring in Comparative Perspective*. Hillsdale, N.J.: L. Erlbaum.
- Radcliffe-Brown, Alfred R. 1935 "Patrilineal and Matrilineal Succession." *Iowa Law Review* 20:286-303.
- Rosenfeld, Jeffrey P. 1991 "The Heir and the Spare: Evasiveness, Role-complexity, and Patterns of Inheritance." In Judith R. Blau and Norman Goodman, eds., *Social Roles and Social Institutions: Essays in Honor of Rose Laub Coser*. Boulder, Colo.: Westview Press.
- Schwartz, T. P. 1996 "Durkheim's Prediction about the Declining Importance of the Family and Inheritance: Evidence from the Wills of Providence, 1775-1985." *Sociological Quarterly* 37:503-519.
- Sorensen, Aage B. 1989 "Old Age, Retirement, and Inheritance." In David I. Kertzer and K. Warner Schaie, eds., *Age Structuring in Comparative Perspective*. Hillsdale, N.J.: L. Erlbaum.
- Sussman, Marvin B., Judith N. Cates, and David T. Smith 1970 *The Family and Inheritance*. New York: Russell Sage Foundation.
- Tsikata, Dzodzi 1996 "Gender, Kinship and the Control of Resources in Colonial Southern Ghana." In Rajni Palriwala and Carla Risseuw, eds., *Shifting Circles of Support: Contextualizing Kinship and Gender in South Asia and Sub Saharan Africa*. Walnut Creek, Calif.: AltaMira.
- Van-Houtte, Jean, Marc Keuleneer, and Kathelijne Vandenberg 1995 "Inheritance Law and Practice in the USA and Belgium." *Sociologia-del-Dritto* 22(3):75-106.

MARVIN B. SUSSMAN

## IN-LAW RELATIONSHIPS

See American Families; Intergenerational Relationships; Kinship Systems and Family Types.

## INSTITUTIONS

See American Society.

## INTEGRATION

See African-American Studies; Segregation and Desegregation.

## INTELLECTUALS

**NOTE:** Although the following article has not been revised for this edition of the *Encyclopedia*, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

Modern societies face a growing dilemma posed by the fact that key institutions and their elites are increasingly dependent upon intellectuals, particularly those in universities, research institutes, and the cultural apparatus generally. Yet, the leaders in these same social units are among the major critics of the way in which the society operates, sometimes calling into question the legitimacy of the social order and its political structure. A ruling

elite, even one that is conservative and anti-intellectual, cannot respond to such challenges by crushing the intellectuals, unless it is willing to incur the punitive costs which such suppression entails. As the Polish “revisionist” philosopher Leszek Kolakowski (1968, p. 179) wrote while still a member of the Communist party, “the spiritual domination of any ruling class over the people . . . depends on its bonds with the intelligentsia . . . ; for the less one is capable of ruling by intellectual means, the more one must resort to the instruments of force.” Decades earlier, the classically liberal (*laissez-faire*) economist and sociologist Joseph Schumpeter (1962, p. 150) argued that under capitalism the dominant economic class must protect the intellectuals, “however strongly disapproving” they are of them, because they cannot suppress intellectual criticism without initiating a process of repression which will undermine their own freedom.

The word *intellectual* is fraught with ambiguities. The meanings attached to it are diverse (Lipset and Dobson 1972, pp. 137–140). In the loosest sense in which the word is used in common parlance today, intellectuals may be said to be all those who are considered proficient in and are actively engaged in the creation, distribution, and application of culture. Edward Shils (1968, p. 179) has suggested a comprehensive definition: “Intellectuals are the aggregate of persons in any society who employ in their communication and expression, with relative higher frequency than most other members of their society, symbols of general scope and abstract reference, concerning man, society, nature and the cosmos.” For analytic purposes, however, it is desirable to distinguish between several types. It is particularly useful to emphasize the much smaller category of “creative intellectuals,” whose principal focus is on innovation, the elaboration of knowledge, art, and symbolic formulations generally. Included in this group are scholars, scientists, philosophers, artists, authors, some editors, and some journalists, as distinguished from the more marginally intellectual groups who distribute culture, such as most teachers, clerics, journalists, engineers, free professionals, and performers in the arts, as well as those who apply knowledge in the course of their work, such as practicing physicians, lawyers, and engineers. To differentiate them from the intellectuals, they may be categorized as the *intelligentsia*.

The creative intellectuals are the most dynamic group within the broad stratum: Because they are innovative, they are at the forefront in the development of culture. The intelligentsia are dependent upon them for the ideational resources they use in their work. Much of the analytic literature dealing with intellectuals has emphasized their seemingly inherent tendency to criticize existing institutions from the vantage point of general conceptions of the desirable, ideal conceptions which are thought to be universally applicable. Thus, Joseph Schumpeter (1950, p. 147) stressed that “one of the touches that distinguish [intellectuals] . . . from other people . . . is the critical attitude.” Raymond Aron (1962, p. 210) argued that “the tendency to criticize the established order is, so to speak, the occupational disease of the intellectuals.” Richard Hofstadter (1963, p. 38) noted: “The modern idea of the intellectual as constituting a class, as a separate social force, even the term *intellectual* itself, is identified with the idea of political and moral protest.” Lewis Coser (1970, p. viii) in defining the term stated: “Intellectuals are men who never seem satisfied with things as they are. . . . They question the truth of the moment in higher and wider truth.”

These concerns are iterated by the fact that “intelligentsia” and “intellectuals,” the two words most commonly used to describe those in occupations requiring trained or imaginative intelligence, were used first in the context of describing those engaged in oppositional activities. “Intelligentsia,” first began to be used widely in Russia in the 1860s, referring to the opposition to the system by the educated strata. It was generally defined as “a ‘class’ held together by the bond of ‘consciousness,’ ‘critical thought,’ or ‘moral passion.’” (Malia 1961 p. 5) “Intellectual” as a noun first secured wide usage in France during the infamous Dreyfus case in 1898. A protest against Dreyfus’s unjust imprisonment (after a biased court-martial), signed by a variety of writers and professors, was published as the “Manifesto of the Intellectuals.” The anti-Dreyfusards then tried to satirize their opponents as the self-proclaimed “intellectuals” (Bodin 1962, pp. 6–9; Hofstadter 1963, pp. 38–39). The term was picked up in the United States in the context of characterizing opponents to World War I.

The American intellectual also has been seen as a source of unrest. Many have called attention to this phenomenon, seeing it as a continuing one in

American history (May 1963; Hayek 1949, pp. 417–433). Richard Hofstadter (1965, pp. 111–112) described their stance of alienation as “historical and traditional,” and pointed out that “even the genteel, established intellectuals of the mid-nineteenth century were in effect patrician rebels against the increasing industrialization and the philistinism of the country. So that it has been the tradition of American intellectuals of all kinds and stamps to find themselves at odds with American society: this, I think, to a degree that is unusual elsewhere.” A century ago, Whitelaw Reid (1873, pp. 613–614), the editor of the *New York Tribune*, pointed to the role of the American “Scholar in Politics” as a foe of the “established,” and a leader of the “radicals.”

The reader should not get the impression that intellectual and student involvement in protest is confined to left-wing or progressive movements. This is not true, as witnessed, for example, by the intellectuals and students who constituted a core segment of activist support for the Fascist party of Mussolini, and of the National Socialist party of Hitler, before they took power, as well as among fascist and assorted anti-Semitic right-wing extreme groups in France and various countries in Eastern Europe up to World War II (Hamilton 1971; Röpke 1960, pp. 346–347). In Eastern Europe and the Soviet Union, intellectuals have been in the forefront of the struggles against Communist regimes, behavior that is perceived as left, i.e. opposition to statism, authoritarianism, and severe stratification.

In the United States, although scattered groups of right-wing intellectuals have emerged at times, the record seems to validate Richard Hofstadter’s (1963, p. 39) generalization that for almost all of this century the political weight of American intellectuals has been on the progressive, liberal, and leftist side. Quantitative data derived from attitude surveys, the earliest dating back to before World War I, plus assorted other reports of the political orientations of the American professoriate, down to the present, strongly indicate that American intellectuals have consistently leaned in this direction (Lipset and Dobson 1972, pp. 211–289; Lipset 1991). This bias, to a considerable extent, reflects the absence or weakness of a legitimate national conservative tradition in America. National identity and national ideology are linked to a value system that emphasizes egalitarianism and populism, stemming from an elaboration of

principles enunciated in the Declaration of Independence. Thus, when American intellectuals point up the gap between the real and the ideal, whether the latter is represented by what was in a bygone Jeffersonian laissez-faire era (a utopia of equal yeoman farmers) or what it should be (a classless participatory future), they challenge the system for not fulfilling the ideals implicit in the American Creed.

Still, the argument is frequently made that inherent in the structural changes since World War II, which have been described as leading to a “postindustrial society,” has been a growing interdependence between political authority and intellectualdom, which should have reduced the critical stance. Modern developed socioeconomic systems are highly dependent on superior research and development resources, which mean better support for universities, and research centers, and the much larger component of persons who have passed through the higher education system—thus creating a mass, high-culture market that pays for the institutions and products of the artistic community (Bell 1973). Governments are increasingly a major source of financing for intellectualdom, ranging from artists to scientists. Recognition and financial rewards from the polity conceivably should help to reduce the historic tensions and the intellectual’s sense of being an outsider. A further trend pressing in this direction is the fact that the complexities involved in “running” an advanced industrial or postindustrial society forces laymen, both political and economic leaders, to seek advice in depth from, to defer to, the scholarly-scientific community (Dahl 1989, pp. 334–335; Gagnon 1987). Many, therefore, have seen these trends as fostering the role of the intellectual as participant, as leading to the “interpenetration” of scholarship and policy (Shils 1968; Brint 1991).

These developments, however, have not led to the decline in the critical role of the intellectual in the United States, although patterns elsewhere appear somewhat different (Lerner, Nagai, and Rothman 1990, p. 26). A number of analyses of different American scholarly disciplines have emphasized the significant presence of political radicals among them and their greater alienation from the powers than in Western European societies (Lipset 1991). In seeking to explain this trans-Atlantic difference, a Swedish scholar, Ron Eyerman (1990) notes that unlike the situation in America,

Swedish (and I would add European) intellectuals are not an “alienated stratum” opposed to the state “because the avant-garde tradition was absorbed through a reformism [that] put intellectual labor to use in service of society,” that is, largely through the labor and socialist movements. While the overwhelming majority of American intellectuals view themselves as outsiders, and have had little experience in directly influencing power that could moderate their sense of alienation, many European intellectuals have worked in a somewhat more integrated context. And most are, it should be noted, still on the left politically, though less radical relative to their national spectrums than their American counterparts.

The dilemma remains for intellectuals everywhere of how to obtain the resources necessary to pursue creative activity without “selling out,” without tailoring creative and intellectual work to the demands of employer, patron, or consumer. In modern times in the West, the emphasis on originality, on innovation, and on following the logic of development in various creative fields—be they painting, music, literature, physics, or sociology—has been responsible for a recurrent conflict between intellectuals and those who pay for their works or exert control through the state, churches, businesses, the market, or other institutions. Intellectuals have often felt themselves to be dependent on philistines while wanting to do whatever they liked according to the norms of their field.

Much of the discussion has focused on the tensions created for unattached intellectuals. It has been asserted “that free-lance intellectuals are more receptive to political extremism than are other types of intellectuals . . . [since] the freelance intellectual . . . has been dependent on an anonymous and unpredictable market. . . . Rewards are much less certain to be forthcoming for the free-lance intellectual, the form of reward less predictable, and the permanence of the recognition more tenuous. . . . [They] tend to be more dependent on their audience, over which they have relatively little control, and to feel greater social distance from it (Kornhauser 1959, pp. 186–187).”

To understand the continued anti-establishment emphasis of intellectuals, even when well rewarded, it is important to recognize the relationship of this emphasis to their concern for creativity

or innovation. The capacity for criticism, for rejection of the status quo, is not simply a matter of preference by some intellectuals. Rather, it is built into the very nature of their occupational role. The distinction between integrative and innovative roles implies that those intelligentsia involved in the former, like teachers, engineers, and exponents of mass culture, use ideas—scholarly findings—to carry out their jobs; those in the latter activities, like scholars, poets, and scientists, are concerned with the creation of *new* knowledge, *new* ideas, *new* art. To a considerable extent, in such endeavors, one is much more rewarded for being original than for being correct—an important fact, a crucial aspect of the role insofar as we consider that such intellectuals tend to be socially critical.

(SEE ALSO: *Postindustrial Society*)

#### REFERENCES

- Allen, Norm R., Jr. 1996 “Religion and the New African American Intellectuals.” *Nature, Society, and Thought* 9:159–187.
- Aron, Raymond 1962 *The Opium of the Intellectuals*. New York: Norton.
- Beaud, Paul, and Francesco Panese 1995 “From One Galaxy to Another: The Trajectories of French Intellectuals.” *Media Culture and Society* 17:385–412.
- Bell, Daniel 1973 *The Coming of Post-Industrial Society*. New York: Basic Books.
- Bodin, Louis 1962 *Les Intellectuels*. Paris: Presses Universitaires de France.
- Brint, Stephen 1991 “The Powers of the Intellectuals.” In William Julius Wilson, ed., *Sociology and the Public Agenda*. Newbury Park, Calif.: Sage.
- Coser, Lewis 1970 *Men of Ideas*. New York: The Free Press.
- Dahl, Robert A. 1989 *Democracy and Its Critics*. New Haven: Yale University Press.
- Drake, W. Avon 1997 “Black Intellectuals and the Politics of Race: The Affirmative Action Debate.” *Research in Race and Ethnic Relations* 10:147–168.
- Eyerman, Ron 1990 “Intellectuals and the State: A Framework for Analysis; with special reference for the United States and Sweden.” Unpublished paper, Department of Sociology, University of Lund, Sweden.
- Gagnon, Alain (ed.) 1987 *Intellectuals in Liberal Democracies: Political Influence and Social Involvement*. Westport, Conn.: Greenwood Press.



- Hamilton, Alastair 1971 *The Appeal of Fascism: A Study of Intellectuals and Fascism, 1919–1945*. London: Anthony Blond.
- Hofstadter, Richard 1963 *Anti-Intellectualism in American Life*. New York: Alfred A. Knopf.
- 1965 “Discussion.” In A. Alvarez, ed., *Under Pressure*. Baltimore: Penguin Books.
- Hollander, Paul 1998 *Political Pilgrims: Western Intellectuals in Search of the Good Society*. New Brunswick, NJ: Transaction.
- “Intellectuals and Social Change in Central and Eastern Europe.” 1992 *Partisan Review* 59:525–751.
- Karabel, Jerome 1996 “Towards a Theory of Intellectuals and Politics.” *Theory and Society* 25:205–233.
- Kellner, Douglas 1997 “Intellectuals, the New Public Spheres, and Techno-Politics.” *New Political Science* 169–188.
- Kempny, Marian 1996 “Between Politics and Culture. Is a Convergence between the East-European Intelligentsia and Western Intellectuals Possible?” *Polish Sociological Review* 4:297–305.
- Kolakowski, Leszek 1968 *Marxism and Beyond*. London: Pall Mall Press.
- Kornhauser, William 1959 *The Politics of Mass Society*. New York: Free Press.
- Lepenies, Wolf 1994 “The Future of Intellectuals.” *Partisan Review* 61:111–119.
- Lerner, Robert, Althea K. Nagai, and Stanley Rothman 1990 “Elite Dissensus and Its Origins,” *Journal of Political and Military Sociology* 18 (Summer): 25–39.
- Lipset, Seymour Martin 1991 “No Third Way: A Comparative Perspective on the Left.” In Daniel Chirot, ed., *The End of Leninism and the Decline of the Left*. Seattle: University of Washington Press.
- , and Richard Dobson 1972 “The Intellectual as Critic and Rebel: With Special Reference to the United States and the Soviet Union.” *Daedalus* 101 (Summer 1971):137–198.
- Malia, Martin 1961 “What is the Intelligentsia?” In Richard Pipes, ed., *The Russian Intelligentsia*. New York: Columbia University Press.
- May, Henry 1963 *The Discontent of the Intellectuals*. Chicago: Rand McNally.
- Misra, Kalpana 1997 “From ‘Capitalist Roaders’ to ‘Socialist Democrats’: Intellectuals in Post-Mao China.” *China Report* 33:267–295.
- Petras, James 1991 “The Metamorphosis of Latin America’s Intellectuals.” *International Journal of Contemporary Sociology* 28:3–4.
- Reid, Whitelaw 1873 “The Scholar in Politics.” *Scribner’s Monthly*, vol. 6.
- Röpke, Wilhelm 1960, “National Socialism and the Intellectuals.” In George B. de Huszar, ed., *The Intellectuals*. New York: Free Press.
- Royce, Edward 1996 “The Public Intellectual Reconsidered.” *Humanity and Society* 20:3–17.
- Schumpeter, Joseph 1950 *Capitalism, Socialism and Democracy*. New York: Harper Torchbooks.
- Shils, Edward 1968 “Intellectuals.” In David L. Sills, ed. *International Encyclopedia of the Social Sciences*, vol. 7. New York: Macmillan and Free Press.
- Smith, James-Allen 1991 *The Idea Brokers: Think Tanks and the Rise of the New Policy Elite*. New York: The Free Press.
- Torpey, John C. 1995 *Intellectuals, Socialism, and Dissent: The East German Opposition and Its Legacy*. Minneapolis: University of Minnesota Press.
- Weijers, Ido 1997 “Educating the Modern Intellectuals.” *International Journal of Contemporary Sociology* 34:81–91.

SEYMOUR MARTIN LIPSET

## INTELLIGENCE

In everyday life people commonly refer to each other as being smart or slow. The perception that individuals differ widely in mental adeptness—in intelligence—long preceded development of the IQ test, and there is indeed a large vernacular for brilliance, stupidity, and the many points in between. There has been much sparring over the scientific meaning and measurement of intelligence, both in the rowdy corridors of public debate and in the sanctums of academe. But what do we actually know about intelligence? A lot more in the last decade, and some of it surprising even to experts. Moreover, the data form a very consistent pattern showing that differences in intelligence are a biologically grounded phenomenon with immense sociological import.

### MEASUREMENT OF INTELLIGENCE

The effort to measure intelligence variation among individuals is a century old. Two strategies for measuring such differences have emerged, the *psychometric* and the *experimental*. Both spring from

the universal perception that, although all people can think and learn, some are notably better at both than others. Accordingly, intelligence research focuses on how people *differ* in cognitive competence, not on what is common to all of us. (Other disciplines such as neuroscience and cognitive psychology specialize in the commonalities.) The aim of intelligence research is thus much narrower than explaining the intricacies of how brains and minds function. These intricacies are relevant to intelligence experts, but generally only to the extent that they illuminate why people in all cultures differ so much in their ability to think, know, and learn.

**Psychometric (Mental Testing) Strategy.** The IQ test represents the psychometric approach to measuring intelligence. Alfred Binet devised the first such test in France to identify children who would have difficulty profiting from regular school instruction. Binet's idea was to sample everyday mental competencies and knowledge that were *not* tied to specific school curricula, that increased systematically throughout childhood, and that could reliably forecast important differences in later academic performance. The result was a series of standardized, age-graded test items arranged in increasing order of difficulty. A child's score on the test compared the child's level of mental development to that of average children of the same age. Binet's aim was pragmatic and his effort successful.

Innumerable similar tests have been developed and refined in the intervening century (Anastasi 1996; Kaufman 1990). Some are paper-and-pencil tests, called *group tests*, that can be administered cheaply to many individuals at once and with only a small sacrifice in accuracy. Others, such as the various Wechsler tests, are *individually administered tests* that require no reading and are given one-on-one. Today, individually administered intelligence tests are typically composed of ten to fifteen subtests that vary widely in content. The two major categories are the *verbal* subtests, such as vocabulary, information, verbal analogies, and arithmetic, which require specific knowledge, and the *performance* subtests, such as block design, matrices, and figure analogies, which require much reasoning but little or no knowledge. The highly technical field that develops and evaluates mental tests, called *psychometrics*, is one of the oldest and most rigorous in psychology. Its products have

been found useful in schools, industry, the military, and clinical practice, where they are widely used.

Professionally developed mental tests are highly *reliable*, that is, they rank people very consistently when they are retested. A great concern in earlier decades was whether mental tests might be culturally biased. *Bias* refers to the systematic over- or underestimation of the true abilities of people from certain groups—a “thumb on the scale”—favoring or disfavoring them. There are many specific techniques for uncovering test bias, and all mental tests are screened for bias today before being published. IQ tests generally yield different average scores for various demographic groups, but the consensus of expert opinion is that those average differences are not due to bias in the tests. The consensus among bias experts, after decades of research often trying to prove otherwise, is that the major mental tests used in the United States today do not systematically understate the developed abilities of native-born, English-speaking minorities, including American blacks. The American Psychological Association affirmed this consensus in its 1996 task force report, “Intelligence: Knowns and Unknowns” (Neisser et al. 1996).

The biggest remaining question about IQ tests today is whether they are *valid*, that is, whether they really measure “intelligence” and whether they really predict important social outcomes. As will be shown later, IQ tests do, in fact, measure what most people mean by the term “intelligence,” and they predict a wide range of social outcomes, although some better than others and for reasons not always well understood.

**Experimental (Laboratory) Strategy.** The experimental approach to measuring differences in general intelligence is older than the psychometric but little known outside the study of intelligence. It has produced no tests of practical value outside research settings, although its likely products could someday replace IQ tests for many purposes. The approach began in the late 1800s when the great polymath Francis Galton proposed that *mental speed* might be the essence of intelligence. He therefore set out to measure it by testing how quickly people respond to simple sensory stimuli such as lights or tones. Galton's measures did not clearly correlate with “real-life” indicators of mental ability, such as educational success, so his *chronometric* approach was quickly dismissed as

wrong-headed and far too simplistic to capture anything important about the beautiful complexity of human thought.

Advances in statistics after the mid-twentieth century, however, showed that Galton's data actually had shown considerable promise. New medical and computer technology have since allowed researchers to measure elements of mental processing with the necessary precision that Galton could not. The revival of his approach in the 1970s has revolutionized the study of intelligence. It is the new frontier in intelligence research today. No longer producing "fool's gold" but the real thing, the study of elementary cognitive processes has attracted researchers from around the world. It now appears that some differences in complex mental abilities may, in fact, grow from simple differences in how people's brains process information, including their sheer speed in processing.

There is no single experimental approach, but perhaps the dominant one today is the chronometric, which includes studies of inspection time (IT) and reaction time (RT). Chronometric tasks differ dramatically from IQ test items. The aim is to measure the *speed* of various elementary perceptual and comprehension processes. So, instead of scoring how well a person performs a complex mental task (such as solving a mathematics problem or defining a word), chronometric studies measure how quickly people perform tasks that are so simple that virtually no one gets them wrong. These *elementary cognitive tasks* (ECTs) include, for example, reporting which of two briefly presented lines is the longer or which of several lights has been illuminated. In the former, an IT task, the score is the number of milliseconds of exposure required to perceive the difference. In the latter, an RT task, the score is the number of milliseconds the subject takes to *release* a "home button" (called "decision time") in order to press the lighted response button (called "movement time").

Both average speed and variability in speed of reaction are measured over many trials. It turns out that brighter people are not only faster but more consistent in their speed of stimulus apprehension, discrimination, choice, visual search, scanning of short-term memory, and retrieval of information from long-term memory. In fact, variability in speed is more highly correlated with IQ (negatively) than is average speed. ECT performance

correlates more highly with IQ as the tasks become more complex, for example, when the number of lights to distinguish among increases from two to four to eight (respectively, one, two, and three "bits" of information). Composites of various speed and consistency scores from different ECTs typically correlate  $-.5$  to  $-.7$  with IQ (on a scale of  $-1.0$  to  $1.0$ , with zero meaning no relation), indicating that both chronometric and psychometric measures tap much the same phenomena. Psychometric and chronometric measures of mental capacity also trace much the same developmental curve over the life cycle, increasing during childhood and declining in later adulthood. Debates among the experimentalists concern how many and which particular elementary cognitive processes are required to account for differences in psychometric intelligence.

#### MEANING OF INTELLIGENCE.

The meaning of intelligence can be described at two levels. Nonexperts are usually interested in the *practical meaning* of intelligence as manifested in daily life. What skills does it reflect? How useful are they in school, work, and home life? In contrast, intelligence researchers tend to be interested in the more *fundamental nature* of intelligence. Is it a property of the brain and, if so, which property exactly? Or is it mostly a learned set of skills whose value varies by culture? Personnel and school psychologists, like other researchers concerned with the practical implications of mental capability, are often interested in both levels.

**Practical Definitions of Intelligence.** The practical meaning of intelligence is captured well by the following description, which was published by fifty-two leading experts on intelligence (Gottfredson 1997a). It is based on a century of research on the mental behavior of higher- versus lower-IQ people in many different settings.

*Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—"catching on," "making sense" of things, or "figuring out" what to do. (p. 13)*

The concept of intelligence refers specifically to an *ability* that is *mental*. It does not encompass many of the other personal traits and circumstances that are important in people's lives. It does not include, for instance, strictly physical skills, creativity, or traits of personality and character such as conscientiousness and drive. IQ tests are not intended to measure these other traits. Three practical definitions that are more specific may illuminate better what intelligence means in daily affairs. Each can be translated into the others, but each highlights a different practical aspect of intelligence: the ability to deal with complexity (Gottfredson 1997b), learn (Carroll 1997), and avoid making cognitive errors (Gordon 1997).

**Intelligence as the ability to deal with complexity.** IQ test items vary widely in content and format, but they often seem esoteric or narrowly academic. Many people in the past took these superficialities as guides to the nature of what IQ tests measure and therefore mistakenly concluded that they cannot be measuring anything of real consequence, at least outside schools. IQ tests' superficial characteristics, however, are irrelevant to their ability to measure intelligence. What matters is the *complexity*, the amount of mental manipulation, their tasks require: contrasting, abstracting, inferring, finding salient similarities and differences, and otherwise turning things over in one's mind to accomplish the mental task. Complexity is the active ingredient in tests that call forth intelligence. People who score higher on IQ tests are people who deal better with complexity, that is, are more adept at understanding and effectively solving more complex mental challenges.

Any kind of test *vehicle* or content (words, numbers, figures, pictures, symbols, blocks, mazes, and so on) can be used to create different levels of complexity. IQ tests typically do, in fact, contain subtests with different kinds of content. Forward and backward digit span (two memory subtests) illustrate clearly the notion of mental manipulation and task complexity. In digits forward, individuals are asked to repeat a string of from two to nine digits (say, 3-2-5-9-6) that is presented orally at one digit per second. In digits backward, the individual simply repeats the numbers in reverse order (in this case, 6-9-5-2-3). The one extra element in the second task (mentally reversing the list) greatly increases its complexity, nearly doubling its correlation with IQ.

Number series completion subtests can also seem trivial, but they illustrate how the same simple content can be varied to build increasingly complex mental demands. Consider the following three series: 4, 6, 8, 10, 12,— (easy item); 2, 4, 5, 7, 8, 10,— (moderate); and 9, 8, 7, 8, 7, 6,— (difficult). One must discern the relations between succeeding numbers in order to complete the series, and those relations become increasingly complex across the three series (respectively, add 2 to each successive digit; add 3 to each successive set of two digits; subtract 1 from each successive set of three digits). These are similar to the items found in one of the fifteen subtests of the Stanford-Binet Intelligence Scale (SBIS-IV) for school-aged youth. They require very little knowledge. Instead, their challenge is to use that simple information effectively—to contrast and compare, find relations, and infer rules—in order to solve logical problems in the test setting. IQ tests that require this on-the-spot problem solving are referred to as tests of *fluid intelligence*—of mental horsepower, if you will.

Some IQ subtests require test takers to bring considerable knowledge into the test setting in order to perform well, but they, too, illustrate the principle that the active ingredient in IQ tests is the complexity of their mental demands. Vocabulary, for example, is one of the very best subtests for measuring intelligence. The reason is that people do not learn most words (*love, hate*) by memorization or direct instruction, but rather by *inferring* their meanings and their fine nuances in meaning (*love, affection, infatuation, devotion, and ardor; hatred, loathing, abhorrence, antipathy, and contempt*) from the way other people use them in everyday life. Learning vocabulary is largely a process of distinguishing and generalizing concepts in natural settings.

Table 1 illustrates how vocabulary level reflects differences in the ability to deal with complexity. These results are from an earlier version of the Wechsler Adult Intelligence Scale (WAIS). All the adults tested were able to provide at least a tolerable definition of concrete items such as *bed, ship, and penny*, but passing rates dropped quickly for more abstract and nuanced concepts such as *slice* (94 percent), *sentence* (83 percent), *domestic* (65 percent), and *obstruct* (58 percent). Only half could define the words *remorse, reluctant, and calamity*. Fewer than one in five knew the words *ominous* and

*tirade*, and only 5 percent could provide even a partial definition of *travesty*. Anyone who has attended high school, read newspapers and magazines, or watched television will have encountered these words. Vocabulary tests thus gauge the ease with which individuals have routinely “picked up” or “caught onto” concepts they encounter in the general culture. So, too, do the general information subtests that are included in many IQ test batteries (“Why do homeowners buy home insurance?”).

Vocabulary, information, and other tests that require considerable prior knowledge are referred to as tests of *crystallized intelligence* because they measure the knowledge that has formed or crystallized from past problem solving. The greater the mental horsepower, the greater the accumulation. Only knowledge that is highly general and widely available is assessed, however, because otherwise IQ tests would also be measuring the opportunity to learn, not success when given the opportunity to do so. Tests of fluid and crystallized intelligence correlate very highly, despite their very different content, because the key active ingredient in both is the complexity of the problems people must solve.

*Intelligence as the ability to learn.* One of life’s unremitting demands is to learn—that is, to process new information sufficiently well to understand it, remember it, and use it effectively. This is especially so in education and training, but it is also the case in meeting the challenges of everyday life, from learning to use a new appliance to learning the subtle moods of a friend or lover.

IQ level is correlated with speed, breadth, and depth of learning when learning requires thinking, specifically, when it is intentional (calls forth conscious mental effort), insightful (requires “catching on”), and age-related, that is, when older children learn the material more easily than do younger children (because they are mentally more mature) and when the material to be learned is meaningful and hierarchical (mastering earlier elements is essential for learning later ones, as in mathematics). Learning is also correlated with intelligence level when the learning task permits using past knowledge to solve new problems, the amount of time for learning is fixed, and the material to be learned is not unreasonably difficult or complex (which would cause everyone to fall back on trial-and-error learning). In short, intelligence is the

**Percentage of Adults Age 16–65 Passing<sup>a</sup> WAIS Vocabulary Items**

ITEM	% PASSING	ITEM	% PASSING
1. Bed	100	21. Terminate	55
2. Ship	100	22. Obstruct	58
3. Penny	100	23. Remorse	51
4. Winter	99	24. Sanctuary	49
5. Repair	98	25. Matchless	47
6. Breakfast	99	26. Reluctant	50
7. Fabric	92	27. Calamity	50
8. Slice	94	28. Fortitude	36
9. Assemble	90	29. Tranquil	36
10. Conceal	87	30. Edifice	22
11. Enormous	89	31. Compassion	29
12. Hasten	87	32. Tangible	30
13. Sentence	83	33. Perimeter	26
14. Regulate	80	34. Audacious	20
15. Commemce	79	35. Ominous	20
16. Ponder	64	36. Tirade	17
17. Cavern	68	37. Encumber	19
18. Designate	63	38. Plagiarize	13
19. Domestic	65	39. Impale	14
20. Consume	61	40. Travesty	5

*Table 1*

SOURCE: Matarazzo (1972), Table 5, p. 514.

NOTE: <sup>a</sup>Passing includes getting at least partial credit.

ability to learn when the material to be learned is moderately complex (abstract, multifaceted, and so on), as distinct from learning by rote or mere memorization.

People learn at very different rates. In school, the ratios of learning rates are often four or five to one, and they can go much higher depending on the material. The military has likewise found that recruits differ greatly in how well they learn, which it calls *trainability*. One 1969 (Fox, Taylor, and Caylor 1969) study done for the U.S. Army found, for example, that enlistees in the bottom fifth of ability needed two to six times as many teaching trials and prompts as did their higher-ability peers

to reach minimal proficiency in rifle assembly, monitoring signals, combat plotting, and other basic soldiering tasks. Figure 1 illustrates the major differences in trainability at different levels of IQ. People with IQs of about 115 and above can not only be trained in a college format but can even gather and infer information largely on their own. Training for people with successively lower IQs, however, must be made successively less abstract, more closely supervised, and limited to simpler tasks. Low levels of trainability limit not only how much can be learned in a set amount of time but also the complexity of the material that can be mastered with unlimited time.

*Intelligence as the ability to avoid common cognitive errors.* Intelligence can also be conceived, for practical purposes, as the probability of *not* making cognitive errors. The notion is that all people make cognitive errors but that brighter people make fewer of them in comparable situations. They make fewer errors in learning, for example, because they learn more quickly and thoroughly. And they make fewer errors of judgment in new and unexpected situations because they are better able to look ahead, assess the likely consequences of different actions and events, spot incongruities and problems, factor more information into their decision making, and perceive alternative courses of action.

Just as items on intelligence tests are scored right versus wrong or better versus worse, so, too, can many decisions in everyday life be classified in this manner. And just as intelligence tests must use many items to assess intelligence level accurately, so, too, does the meaning of intelligence in daily life manifest itself in the accumulation of good and bad decisions, large and small, throughout one's life. The lifetime advantages of higher intelligence are explored later. The point here is simply that intelligence can also be described as the ability to avoid making common errors in judgment and accumulating a harmful record of them.

These three workaday definitions give an intuitive sense of what it means at the level of personal experience to be more versus less intelligent. Intelligence researchers seek to understand intelligence differences in their more fundamental sense, below the surface of everyday observation. As described next, most have adopted a new working definition of intelligence for this purpose.

**Psychometric *g* (Not IQ) as the Research Definition of "Intelligence."** The psychometric approach to measuring intelligence cannot by itself tell us what intelligence is fundamentally, say, neurologically. However, it has greatly narrowed the possibilities. Most importantly, it has shown that intelligence is a highly general ability and that it is the backbone or supporting platform for the more specific mental abilities. This finding rests in turn on the discovery of a single, common, and replicable means for isolating for study what most people mean by intelligence. As explained, it is not the IQ but *g*, which is short for the *general mental ability factor* (Jensen 1998). The latter has replaced the former as the gold standard for measuring intelligence. Researchers do not yet know exactly what aspect of mind or brain *g* represents, but *g* has become the de facto definition of intelligence for most intelligence researchers, some of whom would drop the term "intelligence" altogether.

From the earliest days of mental testing, researchers observed that people who do well on one test tend to do well on all others. That is, all mental tests intercorrelate to some degree. This prompted Charles Spearman, Galton's student and one of the earliest theorists of intelligence, to invent the statistical technique of factor analysis to isolate that common component from any set of mental tests. Once the common factor, *g*, is statistically extracted from a large, diverse set of tests, each individual's standing on it (the person's *g* level) can then be calculated. So, too, conversely, can the ability of different tests to measure *g* (their *g-loadings*). Among mental tests, IQ tests provide the most accurate measures of *g*. Scores on the great variety of IQ tests are all highly *g*-loaded, that is, they all correlate highly with *g* (with .9 being a typical value for tests of the Wechsler variety). This high correlation means that IQ scores are quite adequate for most practical purposes; therefore, *g* scores are generally actually calculated only for research purposes.

*The replicability of g.* Research reveals that the same *g* dimension characterizes all demographic groups yet studied. Virtually identical *g* factors have been extracted from all large, diverse sets of mental tests, regardless of which method of factor analysis was used and regardless of the age, sex, or race of the test takers. The same *g* is called forth by tests that require much cultural knowledge as by ones requiring virtually none. It can be called up

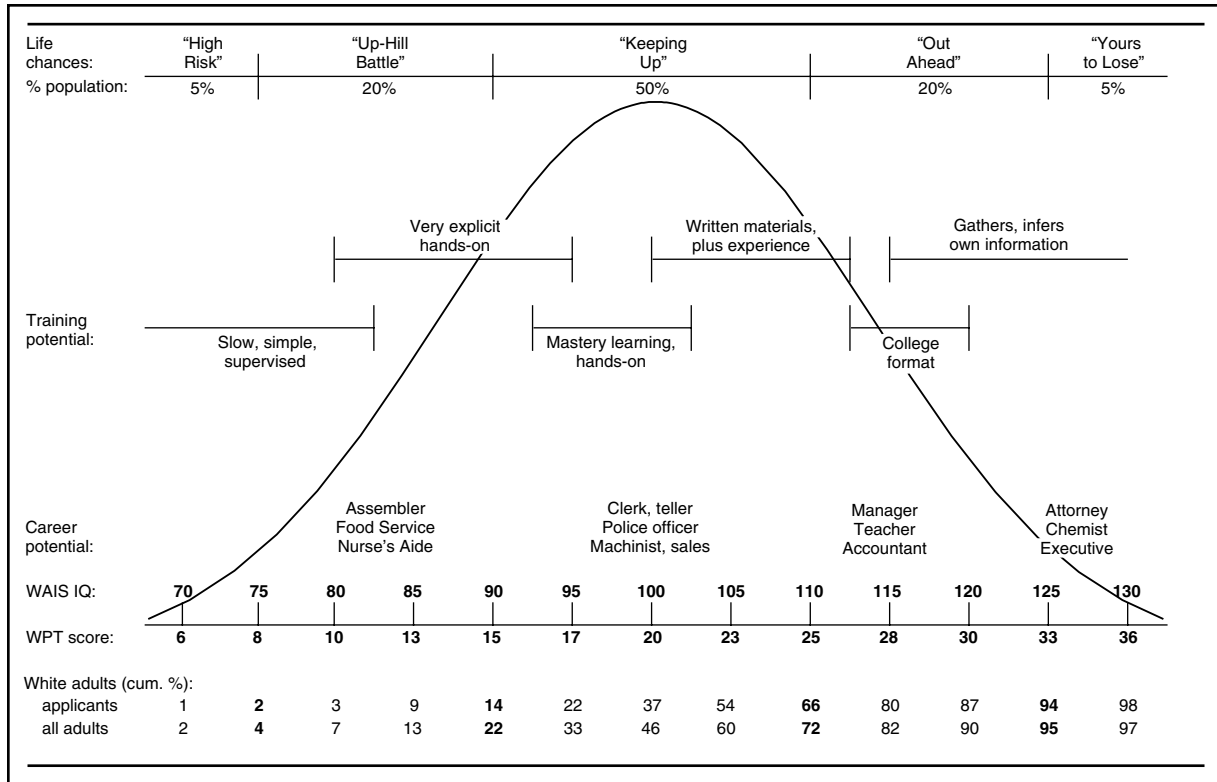


Figure 1

by any kind of item content (numbers, letters, shapes, pictures, blocks, and the like), a phenomenon that Spearman called *indifference of the indicator*.

Mental tests often measure more specific aptitudes in addition to *g* (say, verbal or spatial ability), but *g* is the crucial backbone of all mental tests. Efforts to create useful mental tests that do not measure *g* (for example, verbal aptitude tests that do not tap *g*) have all failed. Although mental tests are suffused by a common factor, no analogous common factor can be found among different personality tests (which test for extroversion, conscientiousness, sociability, and so on). The absence of a general personality factor illustrates that the general mental ability factor *g* is not an artifact of factor analysis but a real phenomenon.

To be sure, the existence of the *g* factor can be obscured by inappropriate testing (for example, when some test takers do not know the language well) and by narrow sampling (when all test takers are similar in intelligence). When allowed to manifest itself, however, the *g* factor clearly shows itself

to transcend the particulars of content and culture. This is not to say that culture cannot affect the development of *g* or its social significance, but only that culture does not determine its fundamental nature. The nature of *g* seems to be surprisingly independent of culture, as other sorts of research have confirmed.

*The generality of g.* The great generality of *g* is perhaps psychometrics' most crucial discovery about the nature of intelligence. As noted, the identical *g* factor is the major distinction in mental abilities in all groups of people and tests, regardless of cultural context or content. As also noted, all mental ability tests measure mostly *g*, no matter what specific abilities they were intended to measure (verbal aptitude, mathematical reasoning, memory, intelligence, and so on). The manifest skills most associated with intelligence in both fact and public perception—reasoning, problem solving, abstract thinking, and learning—are themselves highly general, context-independent thinking skills. The psychometric vehicles (tests and test

items) for measuring  $g$  are necessarily culture-bound to some degree, but the  $g$  abstracted from them appears not to be.

There are, of course, other mental aptitudes, but, unlike  $g$ , they seem specific to particular domains of knowledge or activity (language, music, mathematics, manipulating objects in three-dimensional space). Moreover, none of these narrower abilities seem so integral as  $g$  to the expression of all the others. Many decades of factor-analytic research on human abilities have confirmed what is called the hierarchical structure of mental abilities (Carroll 1993). As shown in the simplified version in Figure 2, abilities are arrayed from the top down, with the most general placed at the top. Research always finds  $g$  at the top of this generality hierarchy for mental abilities.

The generality of intelligence was less clear when researchers relied on IQ as their working definition of intelligence. The reason is that all IQ tests are imperfect measures of  $g$  and each often captures the flavor of some specialized ability or knowledge in addition to  $g$ . That is, all IQ tests share a large  $g$  component, but their small non- $g$  components often differ in size and content. Attempting to understand intelligence by studying IQ scores has been akin to chemists trying to understand the properties of a particular chemical element by each studying samples that were impure to different degrees and with different additives. This ensured a muddled and fractious debate about the essence of intelligence. In contrast, the  $g$  factor is a stable, replicable phenomenon. When researchers study  $g$ , they can be confident they are studying the same thing, even when the  $g$ 's they use were extracted from different sets of tests. Moreover,  $g$  has the advantage over IQ that it cannot be confused with the attributes or contents of any particular test, because  $g$  is always extracted from some large, mixed set of them. One must look below the surface characteristics of IQ tests, to  $g$ , to explain the core phenomenon they measure.

*The  $g$ -loadings of tests and tasks.* The ability to classify tests according to their correlation with  $g$  is also a major advance in the study of intelligence. It allows research on why tasks vary in their ability to call forth  $g$  and thus helps predict where in life higher levels of intelligence are most useful. Stated another way, mental tests can now be used to

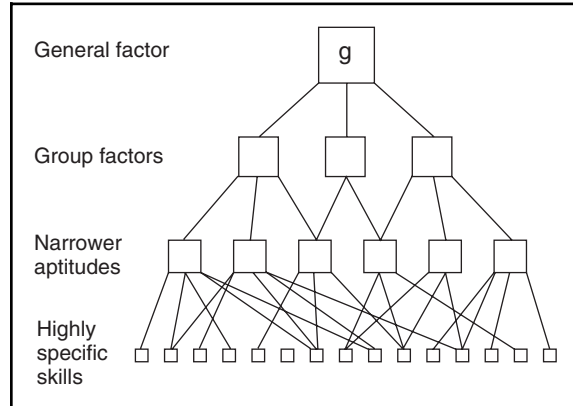


Figure 2

compare environments, not just people, and figure out why some environments are more cognitively demanding than others.

Evidence suggests that tasks are more  $g$ -loaded when they require more complex information processing, for example, when there are more pieces of information, when there are more operations to perform, and when the information is abstract, nested or incomplete. For instance, spelling and arithmetic tests pose much less complex and  $g$ -loaded tasks for adolescents and adults than do vocabulary and mathematical reasoning tests. Spelling and computing well in adolescence and beyond depends less on  $g$  level than does comprehending higher-level verbal and mathematical concepts, despite their superficially similar content.

As will be seen, many work tasks and occupations have been ranked in their demands for  $g$ . In theory, a  $g$  loading can be calculated for virtually everything we do in daily life. Life is like a series of mental tests in the sense that its demands vary considerably in complexity and consequent  $g$ -loading. This means that the advantages of being brighter will vary systematically across different life settings according to their cognitive complexity.

The finding that the subtests in an IQ test battery differ systematically in their ability to measure  $g$  has been cleverly used to explore the biological as well as the sociological meaning of  $g$ . By the method of *correlated vectors*, the  $g$ -loadings of IQ subtests are themselves correlated with other attributes of the subtests. For example, tests'  $g$ -loadings have been found to predict the genetic



heritability of their scores, degree of inbreeding depression, and the subtests' correlations with brain size, faster glucose metabolism in the brain, and greater complexity and speed of onset of various electroencephalogram (EEG) brain waves. This pattern of correlations reinforces other findings which suggest that *g* is a biologically grounded capability to process complex information regardless of its explicit content.

Mental test scores—including the IQ—are composed of both *g* and non-*g* components, however. The non-*g* component might reflect more specific abilities, specific bits of cultural knowledge, aspects of personality or the testing situation, or other unspecified impurities that are independent of *g*. The decomposition of test scores into their *g* versus non-*g* components is also an enormously important development for understanding the meaning of intelligence. For example, it has been shown that it is almost exclusively the *g* component, not the non-*g* components, of tests that accounts for their ability to predict later school achievement and job performance. This considerably narrows the range of possible explanations for why IQ tests predict differences in individuals' later achievement. The explanation cannot reside mostly in the context-specific bits of knowledge that an IQ might reflect, but in the highly general mental capability that *g* represents in all contexts and cultures.

#### **Experimental Study of the Components of *g*.**

If psychometrics has discovered that *g* is a very general information-processing capability, laboratory studies of intelligence are aimed at teasing out its components or building blocks. The debate among experimentalists has been about whether individual differences in general intelligence are more like differences in computer hardware or computer software. Both views, however, perceive differences in *g* or IQ as differences among individuals in the speed and quality of their information processing.

The “software” view argues that differences in intellectual performance originate in the better or worse use of the same hardware, for example, in the use of better strategies or algorithms for using information and solving problems. These *metacognitive* skills might include better allocation of time to the different components of a problem, monitoring of

progress or responding to feedback, and otherwise better controlling how the different components of a task are executed. Such studies might look, for example, at the kinds of planning subjects use in solving verbal analogies or the ways they use their time in comprehending a passage of text. In this view, the general factor *g* reflects not a general underlying ability but the greater conscious use of separate planning and control strategies of general value, in all of which individuals could presumably be trained.

The “hardware” view postulates that differences in the speed and quality of information processing originate in differences in basic brain physiology, such as nerve conduction velocity. The great enthusiasm over the “top-down” software view during the 1970s and 1980s waned as research began more and more to support the claims of the “bottom-up” hardware view of intelligence. People can indeed be observed to use different strategies in solving problems, but differential motivation, effort, or strategy use do not seem to account for IQ differences, and the successful strategies are fairly task-specific.

Although research has not yet proven that differences in lower-level information processing abilities actually *cause* differences in higher-level ones, measures closer to the physiological level offer more promising explanations of *g* (Vernon 1993). For example, simultaneous recordings of subjects' RTs and brain-wave activity (specifically, average evoked potentials [AEP] measured by the EEG) have shown that speeds of ECT responses are moderately to highly correlated with complexity and speed of onset of certain brain waves, both of which occur in less time than required for *conscious* awareness of a stimulus. Much other research shows that both ECT and AEP responses are, in turn, moderately to highly correlated with IQ scores and, most importantly, with *g* itself. The *g* factor is the only mental ability with which ECT scores correlate.

Accordingly, some intelligence researchers now argue that intelligence may not be an ability per se, but rather a chemical, electrical, or metabolic property of the brain. Specific aptitudes, such as verbal and spatial ones, appear to reside in particular regions of the brain, but *g* may represent a global property permeating all regions. Nerve conduction velocity is currently being investigated as

one such possible global property. Differences in velocity may in turn result from differences in nerve myelination (myelin is the fatty sheath around nerve axons). While still speculative, the velocity and myelination hypotheses are consistent with a well-established pattern of differences that any putative cause of intelligence will have to explain, namely, both the steady rise and then fall of fluid intelligence over the life cycle as well as the enduring differences in *g* among people at any single age.

**Popular Contending Theories.** Any theory of intelligence must take into account the basic facts about intelligence, whether it is measured as IQ or *g*. These include its high generality, heritability (discussed shortly), and correlations with elementary perceptual and physiological processes of the brain. Some of the theories that are most popular outside expert circles contradict or ignore these facts and thus are not viable contenders to the emerging *g* theory of intelligence. Others remain untested hypotheses. The major contenders to *g* theory can be characterized as either *specificity* or *multiplicity* theories of intelligence.

*Specificity theories.* Some scholars have argued that intelligence is not an underlying ability but merely the accumulation of specific *bits of knowledge*. For them, being “smart” is nothing more than knowing a lot, no matter how much time and effort went into that learning or what was learned. It is akin to the accumulation of marbles in a jar, signifying nothing other than that many marbles have been collected by whatever means. The apparent assumption is that people do not differ in their ability and efficiency in gathering marbles. However, intelligence has to be more than knowledge per se because, among other reasons, differences in intelligence show up on tests that require no knowledge whatsoever. Moreover, as noted, people differ greatly in their ability to acquire knowledge even when given the same opportunity to learn. There are “fast” students and “slow” students, irrespective of motivation and quality of instruction. For many experts, differences in the ability to *acquire* knowledge are at the heart of intelligence.

Another variant is the *cultural specificity* theory, which is that intelligence is merely the display of traits, whatever they may be, that are highly regarded in a particular culture. For example, one

claim is that because IQ tests are typically developed by white European males, they inevitably measure beliefs, behavior, and knowledge that white European males value but that may have no intrinsic value. Intelligence, they say, might be defined completely differently in another culture, such as skill at hunting, navigating, or cooperating for the general good. The first claim is false and the second is irrelevant, even if true. As noted, the same *g* is extracted from all diverse sets of mental tests and for all cultural groups. (Besides, Asians tend to do better than whites on tests developed by the latter.) Whether different cultural groups recognize, value, and reward the phenomenon represented by *g* is interesting and important, but it does not erase the phenomenon itself as a scientific fact any more than rejecting the concept of evolution brings evolution to a halt.

Perhaps the best-known variant is the *academic specificity* theory, which says that IQ and intelligence are simply “book smarts,” a narrow “academic” skill that is useful inside but not outside schools and bookish jobs. According to this theory, intelligence may be an enduring personal trait, but only a narrow one. As will be shown, *g* is indeed highly useful in education and training. However, the very generality of *g*—the ability to deal with complexity, to learn, and to avoid mistakes—argues against the narrow “book smarts” conception of intelligence. So, too, does much research, discussed later, on the many practical advantages conferred by higher levels of *g*. Carpenters as well as bank tellers, sales agents as well as social scientists, routinely deal with complexity on the job and are aided by higher levels of *g*.

*Multiplicity theories.* Robert Sternberg (1985) argues that there are several intelligences, including “analytical,” “practical,” and “creative.” Howard Gardner (1983) is famous for postulating eight and possibly nine intelligences: linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, intrapersonal, interpersonal, naturalist, and (possibly) existential. Daniel Goleman’s 1995 book on “emotional intelligence” has taken the country by storm. All three theories are engaging, are popular in lay circles, and describe undeniably important skills, knowledges, and achievements. All three theories suggest that *g*, if it exists, is only one of various coequal abilities. This is, indeed, why multiple intelligence theories are so popular. They are

often interpreted—wrongly—as suggesting that everyone can be smart in some useful way.

The question, however, is whether the “intelligences” these theories describe are actually comparable to *g* in any fundamental way. Specifically, are they even abilities, or might they be the *products* (literary, scientific, or artistic) of exercising *g* together with specific abilities in specific settings with specific kinds of training and experience? Are the purported intelligences even *mental* rather than, say, physical abilities or aspects of personality? And for those that are mental abilities, are they comparable to *g* in their *general applicability*? Unfortunately, the research necessary for answering these questions credibly has not been conducted. Almost none of the “multiple intelligences” has actually been measured, and none have been shown independent of *g* in representative samples of the population. Verbal descriptions of them leave many experts doubtful that they are comparable to *g* in any important way.

Some of them, like emotional intelligence, seem to be a combination of many different traits, some being abilities and others not, some being mental and others not. Verbal definitions suggest that practical intelligence (like “street smarts”) may be the accumulation of highly context-specific knowledge gathered through strictly informal experience (for example, knowing the special argot and norms of a particular neighborhood, occupation, or other subculture). Gardner’s intelligences are different forms of highly valued cultural accomplishment. As such, they require not only the ability to succeed but also the personality traits, such as drive and persistence, needed to transform that potential into a valued product. This is not to deny that personality is important for great accomplishment, but only that it is useful to distinguish between the separate ability, personality, and other factors contributing to it.

In addition, some of Gardner’s intelligences seem to mirror more specific and psychometrically well-studied traits, such as verbal, mathematical, and spatial aptitude. Much research has shown that these so-called group factors are highly correlated with *g* but appear below it in the hierarchical structure of human mental abilities (see Figure 2). Gardner himself has stated that exemplary levels of all his intelligences require IQ levels over 120,

meaning that the eight intelligences are not alternatives to *g* but narrower abilities pervaded by it. In short, they appear to be different cultural playgrounds for the cognitively rich. All the purported “multiple intelligences” are important topics for study, but they cannot be assumed to be comparable to *g* in either generality or practical importance by virtue of being labeled “intelligences.”

#### HERITABILITY AND ENVIRONMENTALITY OF INTELLIGENCE

Behavioral genetics is a method for studying the influence of *both* genes and environments on human behavior. In recent decades the field has shown that mental abilities, personality, vocational interests, psychopathology, and even social attitudes and life events are shaped by both genes and environments (Loehlin 1992; Plomin et al. 1997). More research has been conducted on the heritability of intelligence than on any other psychological trait, and much of it has been longitudinal.

Behavioral genetics focuses on explaining *variation* in a particular population. Its basic method is to look at similarities between relatives of different degrees of genetic and environmental relatedness: identical twins reared apart, adopted siblings reared together, identical versus fraternal twins, and so on. Such research can also test, among other things, whether *specific* environmental factors create IQ similarities and differences and, if the research is longitudinal, whether change (and stability) in IQ ranking is due to the operation of genes, environments, or both. It can also test whether two heritable traits or behaviors, such as IQ and academic achievement, share the same genetic and environmental roots.

Such research does not reveal *how* genes affect intelligence, only that they do. Explanations of how genes influence intelligence will come from molecular genetics, which has only recently isolated the first gene for intelligence. Molecular genetic research also holds promise for detailing exactly how environments might influence the actions of genes.

**Individual Differences.** Behavioral genetics has focused historically on explaining differences among individuals *within* a population. The following such findings should be generalized only to the sorts of populations studied so far, most of

them Western and none extreme in terms of either deprivation or privilege.

*IQ is substantially heritable.* Heritability ( $h^2$ ) refers to the percentage of observed differences in a trait (in *phenotypes*) that is due to differences in genes (*genotypes*). Estimates of the heritability for IQ typically range between .4 and .8 (on a scale from 0 to 1.0). This means that from 40 percent to 80 percent of the observed differences in individuals' IQs are due to the genetic differences among them. This means, conversely, up to 20 percent to 60 percent of IQ differences are environmental in origin. Aptitudes measured by the most *g*-loaded tests are the most heritable. Aptitudes measured by tests of more specific abilities, such as verbal and spatial visualization, are moderately heritable, but less so than *g*.

*IQ heritability rises with age.* This discovery was a surprise even to behavioral geneticists because, like virtually all social scientists, they had assumed that environmental effects cumulate over a lifetime to reduce the influence of genes. Not so, apparently. The heritability of IQ is about .4 in the preschool years, rises to .6 by adolescence, and increases to about .8 in late adulthood. The reason for this increase is unclear. The major hypothesis, however, is that "genes drive experience" and lead people to seek different social niches. That is, different genotypes tend to choose, create, and elicit different environments in childhood and beyond, which in turn shape intellectual development. For example, bright and dull youth receive different encouragement and opportunities. They also tend to choose different experiences for themselves, especially as they become more independent of parents, teachers, and other authorities. As individuals take a greater hand in shaping their environments, for better or worse, their IQ phenotypes begin to mirror their IQ genotypes more closely. The correlation between IQ phenotypes and genotypes (which is the square root of heritability) rises to .9 by later adulthood.

The surprising rise in heritabilities is consistent with the disappointing results of socioeducational interventions (similar to Head Start) that were designed to raise low childhood IQs. To date, all have exhibited *fade-out*, meaning that the initial improvements in IQ dissipated within a few years. Improvements in more malleable outcomes (such as fewer children being held back a grade) may be

observed, but permanent rises in *g* are not. The same IQ fade-out occurs with genetically at risk children adopted into more advantaged families: By adolescence, their early favorable IQs fall back to the average for their nonadopted biological relatives.

*IQ-relevant environments are partly genetic in origin.* Social scientists have tended to think of environments as conditions strictly "out there" to which people are passively "exposed." Children's environments correlate with their genes, however, partly because they passively receive both from their parents. People's environments are also heritable to some degree because people choose, make, remake, elicit, and interpret them. Because people's genetic proclivities help shape their environments, real and perceived, behavioral geneticists often refer to people's proximal environments as, in effect, their *extended phenotypes*. That is, people's near environments are to some degree an extension of themselves because they are partly *products* of the person's genotype for intelligence, personality, and the like.

When people's environments are studied with the same behavioral genetic techniques as are their psychological traits and behaviors, research consistently shows that rearing environments, peer groups, social support, and life events are, in fact, moderately heritable. For example, one measure of individual infant and toddler rearing environments found that those environments were 40 percent heritable. Moreover, half of the environmental measure's ability to predict cognitive development could be accounted for by that measure's genetic component. In other words, IQ-relevant environments are partly genetic in origin. This is an example of what behavioral geneticists refer to as the operation of nature *via* nurture.

*Shared family effects on IQ dissipate by adolescence.* Behavioral genetic research confirms that environments have substantial influence in creating IQ differences. However, providing yet another surprise, the research showed that environmental influences had been completely misunderstood. Psychologists-behavioral geneticists David Rowe (1994) and Sandra Scarr (1997) call this mistaken view, respectively, "family effects theory" and "socialization theory." This is the still widespread but false assumption that differences between families in their socioeconomic circumstances (income,

parental education, occupation, income, and so on) and child-rearing styles (cold, authoritative, and so on) create differences between their children in ability and personality. These presumed effects are called *shared* or *between-family* influences because they affect all children in the family in the same way and thus make children in the same families more alike and children in different families less alike.

As it turns out, such shared effects influence IQ (but not personality) in early childhood, but they disappear by adolescence. Nor is it known what these temporary influences are. The only environmental effects that continue to influence IQ beyond adolescence are *nonshared* or *within-family* effects on IQ. Nonshared effects are factors that influence one sibling but not others in a family. What they consist of regarding IQ is not yet known, but they could include random biological events, illness, and differential experiences in parent-child or sibling relationships. Nonshared effects help to explain why biological siblings who grow up together are so different in IQ. They differ by about 12 IQ points, on the average, compared to the average 17-point IQ difference between any two random strangers. Much of that difference is due to their genetic differences, however, because biological siblings share, on the average, only 50 percent of their segregating genes.

The dissipation of “family effects” and the rising influence of genes with age can be seen clearly in adoption research. The IQs of adopted siblings are similar in childhood but not in adolescence. By adolescence, their IQs also cease to resemble the IQs of their adoptive parents but become more like the IQs of the biological parents they have never known.

*Special abilities, ECTs, and school achievement have common genetic roots with g.* As noted, there are many mental abilities, whether at the level of ECTs, such as choice reaction time, or at the level of group factors, such as verbal ability. However, they all correlate with *g*. To the extent that they overlap each other and *g* phenotypically, that overlap is due almost entirely to a common genetic source. Conversely, only a small portion of the genetic component of specific aptitudes—such as verbal skills, memory skills, and speed of processing—is not *g*-related. The same general pattern is

found for the sizable correlation between academic achievement and IQ. To the degree that they correlate, that similarity is almost entirely genetic; to the degree that they diverge, the cause is mostly environmental.

*IQ stability is mostly genetic in origin whereas age-to-age change in IQ rank originates mostly in nonshared environments.* Rank in IQ relative to age-mates is highly stable. Genes and shared environments both contribute mostly to IQ stability rather than to age-to-age change. It is the nonshared environment that causes age-to-age change. Marked change is rare and tends to be idiosyncratic, transient, and difficult to attribute to any particular event.

*Cautions in interpreting heritabilities.* High heritabilities do *not* mean that a trait is not malleable. Heritability and malleability are separate phenomena. Certain heritable conditions (such as diabetes) are treatable and certain nongenetic effects (such as those of lead poisoning) are not. All that a high heritability means is that current differences in environmental conditions do not create much intelligence variation beyond that owing to genetic differences. If environments were equalized for everyone, phenotypic variation might be reduced somewhat, but heritability would rise to 100 percent. In contrast, if environments could be individually tailored to compensate for genetic differences (by providing insulin for diabetics, changing the diets of those with phenylketonuria, providing the best education to the least intelligent, and the like), both heritability and variability would fall.

Moreover, heritability is the degree to which genes explain phenotypic *variance* in a trait, so a high heritability does not rule out shifts in population *averages* over time. Something that affects everyone can change a group’s average without changing its variability. Recent generations have been getting taller, but height is still highly heritable within generations. The same is true for IQ levels, which have been increasing several points a decade this century in developed countries. Both increases are still scientific puzzles, but some scholars have suggested a common explanation—societywide improvements in nutrition, reduction in disease, and the like. Researchers have yet to establish, however, to what extent the rises in IQ reflect increases in the *g* versus non-*g* components of mental tests and thus an increase in *g* itself.

What is clear, however, is that *shared* family environments that vary within the *normal* range of family environments in the developed world do *not* have create lasting differences in IQ. Within the normal range of variation, different families have basically the same effects in promoting mental growth. The key to understanding how environments create IQ differences among age peers lies in understanding *nonshared* effects. These are the environments, whether biological or social, and both within and outside family settings, that affect siblings differently and make them less alike. The shattering of shared effects theory as an explanation for adult differences in IQ is a revolutionary development, albeit one yet to be accepted by many social scientists. The discovery of lasting nonshared influences opens exciting new ways of thinking about IQ-relevant environments. We may have been looking in all the wrong places. Behavioral genetics provides the best tools at present for ferreting out what those nongenetic factors are.

To reiterate a cautionary note, we do *not* know the effects of environments that are extreme or that do not allow individuals the personal freedoms that most Westerners enjoy. We do not know, either, what the effects of entirely novel environments or interventions would be, whether social or biological. We can predict, however, that any social or educational intervention would have to fall outside the normal range of variation already studied in order to change the distribution of IQs very much. For instance, supplying a typical middle-class family environment to all lower-class children cannot be expected to narrow the average IQ gap between middle- and lower-class adolescents. Middle-class children themselves range across the entire IQ spectrum (as do lower-class children) despite the advantages (or absence thereof) of middle-class life.

**Group Differences.** There is little scientific debate anymore about whether valid *phenotypic* differences exist among races, ethnicities, and social classes. Average group differences in IQ are the rule, not the exception, both worldwide and in the United States. To the extent that the matter has been investigated, group IQ differences appear to reflect differences in *g* itself and are mirrored by group differences in performance on the simple laboratory tasks described earlier.

Group IQ differences can be pictured as the displacement of the IQ bell curves of some social groups somewhat upward or downward on the IQ continuum compared to others. All groups' bell curves overlap greatly; the differences consist in where along the IQ continuum each group is centered. Ashkenazic Jews tend to score as far above average (about IQ 112) as American blacks score below average (about IQ 85), with most other groups spread in between. It should be noted, however, that black cultural subgroups differ among themselves in average IQ, as do the constituent subgroups of Jews, gentile whites, Asians, Hispanics, and Native Americans.

The most contentious debate regarding intelligence is whether average group IQ differences are partly genetic in origin. The favored assumption in the social sciences for the last half-century has been that race differences are entirely environmental. However, research designed to prove this has not done so. It has succeeded in finding environmental factors that might possibly explain at most a third of the American black-white average difference. This failure does not rule out an entirely environmental explanation based on factors yet to be assessed. It does rule out several factors, however, that were once assumed to account for the bulk of the average difference, namely, family income and social class. Large average IQ differences between black and white children are found at all levels of family income and social class.

Behavioral geneticists have recently developed statistical methods for estimating the extent to which average differences among social groups (races, sexes, and so on) might be due to genetic differences among them. Perhaps not surprisingly, few behavioral geneticists have actually applied those methods to available data, and those who have been willing to do so have experienced unusual barriers to publishing their results. As a result, there is little direct published evidence one way or the other. When surveyed in 1988, about half of IQ experts reported a belief that race and class differences result from *both* genetic and environmental influences. This should be considered a reasonable but unproven hypothesis.

The earlier caution should be repeated here. Research has so far studied only the normal range of environmental variation within any race or ethnic group. American minority children may

more often grow up in extremely deprived environments. Studies of very low-IQ Appalachian communities suggest that biologically unhealthy and cognitively retarded family environments can permanently stunt cognitive development.

Some people fear that any evidence of genetic differences between groups would have dire social consequences. This fear is unwarranted. A demonstration of genetic differences would not dictate any particular political reaction. Both liberal and conservative social policy can humanely accommodate such an eventuality, as some policy analysts and behavioral geneticists have illustrated (Kaus 1992; Rowe 1997). Depending on one's politics, for example, genetic differences by race could be used to argue either for forbidding or for requiring racial preferences in education and employment. Moreover, environmentalism and hereditarianism have both on occasion helped undergird tyrannical regimes that practiced mass murder, for example, respectively, the Stalinist Soviet Union and Nazi Germany. Political extremism (or moderation) is neither guaranteed nor precluded by scientific conclusions one way or the other. Scientific facts and political reactions to them are independent issues. Developing *effective* social policy does depend, however, on working in concert with the facts, not against them, whatever they may be.

#### SOCIAL CORRELATES AND CONSEQUENCES OF DIFFERENCES IN INTELLIGENCE

Much research has focused on how *individuals'* own behavior and life outcomes are affected by their intelligence level. There has been little research yet on what may ultimately interest sociologists more, namely, the ways in which *interpersonal contexts* and *social institutions* are shaped by the cognitive levels of the individuals populating them.

**Individual Level.** American adults clearly value intelligence highly because they rate it second only to good health in importance. Differences in intelligence do, in fact, correlate to some extent with just about everything we value, including mental and physical health, success in school and work, law-abidingness, emotional sensitivity, creativity, altruism, even sense of humor and physical

coordination. The scientific question, however, is whether differences in intelligence actually *cause* any of these differences in people's behaviors and outcomes. Or might intelligence as often be their consequence as their cause?

*Questions of causality.* Most IQ variability is genetic from adolescence on, meaning that it cannot be mostly "socially constructed." Moreover, to the extent that it has nongenetic sources, evidence leans against their being the usual suspects in social research (parents' income, education, child-rearing practices, and the like). If intelligence is not caused (much) by its major social correlates, does it cause them?

Pieces of an answer are available from experimental and quasi-experimental research conducted by educational, employment, and training psychologists in public, private, and military settings for more than a half-century. Differences in *prior* mental ability are strong—in fact, the strongest—predictors of *later* performance in school, training, and on the job when tasks are at least moderately complex. Moreover, the correlations are stronger with objective than subjectively measured performance outcomes (for example, standardized performance rather than teacher grades or supervisor ratings). The military services also have extensive experience attempting to nullify the effects of ability differences on recruits' later performance in training and on the job. Their failed attempts testify to the stubborn functional import of such differences—as does the failure of lengthy job experience to neutralize differences in worker intelligence.

IQ is moderately highly correlated with a nexus of good outcomes—higher education, high-status jobs, and income growth over a career. In view of the *g*-loadedness of the educational and occupational worlds, it would be surprising were IQ not found to be an important precursor to these outcomes. IQ is, in fact, the best predictor of later educational level attained, and it helps predict occupational status and income growth even after controlling for education and family background.

IQ is also correlated to varying degrees (negatively) with a nexus of bad outcomes—dropping out of school, unemployment, incarceration, bearing illegitimate children, dependence on welfare, and living in poverty as an adult. This nexus of

social pathology has been the focus of recent lively debates about the role of intelligence, where protagonists typically pit intelligence against an array of external factors, including various aspects of family background, to see which is the stronger predictor. Intelligence generally equals or exceeds the predictive ability of any small set of such variables, although the relations tend to be modest in both cases. One possible explanation for the relation of IQ to social pathology is that lack of socioeconomic competitiveness may precipitate a downward social spiral.

However, IQ may play a direct role, too. Committing crimes, bearing illegitimate children, and other such personal behavior may result in part from errors of judgment in conducting one's life, perhaps due in part to lack of foresight and ability to learn from experience. Conversely, higher *g* may help insulate people from harmful environments. Research has shown, for instance, that higher intelligence is a major attribute of "resilient" children, who prosper despite terrible rearing conditions, and of those who avoid delinquency despite living in delinquent environments. The hypothesis is that their greater ability to perceive options and solve problems constitutes a buffer.

Either the genetic or nongenetic components of phenotypic intelligence might be responsible for its causal impact. Because intelligence is highly genetic, it is reasonable to assume that its causal impact is mostly due to its genetic component. This has, in fact, been found to be the case with its effect on standardized academic achievement. The latter shares all its genetic roots with IQ. Similar *multivariate genetic analyses* are now accumulating for various socioeconomic outcomes that depend on mental competence. Educational and occupational level are both moderately genetic in origin, with estimates (for males) ranging from .4 to .7 for education and .3 to .6 for occupation. Part of that genetic portion overlaps the genetic roots of IQ. In the best study so far (Lichtenstein and Pedersen 1997), occupational status was more than half genetic in origin. Half that *genetic* portion was shared jointly with the genetic roots of *both* IQ and years education, and half was independent of both. The remaining variability in phenotypic occupational status was split between *nonshared* environmental effects that (1) were shared with education (but not IQ) and (2) were unique to occupation.

One of the biggest confusions in the debate over the causal role of intelligence results from the mistaken equating of intelligence with genetic factors and of social class with nongenetic factors by some of the most visible protagonists in the debate. While the former assumption has some justification owing to the high heritability of *g*, it nonetheless muddies the conceptual waters. The latter assumption is even less warranted, however, because many social "environments" turn out to be moderately genetic. All social "environments" must now be presumed partly genetic until proven otherwise. Not being genetically sensitive, virtually all current research on the effects of social and family environments is actually uninterpretable. Progress in the causal analysis of environments and their relation to *g* will come only when more social scientists begin using genetically sensitive research designs.

*Principles of importance.* Although the causal role of intelligence has yet to be clarified, research leaves no doubt that people's life chances shift markedly across the IQ continuum. Those shifts in specific life arenas will be discussed later, but it would help first to state four principles that summarize what it means for intelligence to have *practical importance* in individuals' lives.

First, importance is a matter of better *odds*. Being bright is certainly no guarantee of happiness and success, nor does being dull guarantee misery and failure. Most low-IQ people marry, work, have children, and are law-abiding citizens. Being brighter than average does, however, systematically tilt the odds toward favorable outcomes. Higher levels of intelligence always improve the odds, sometimes only slightly but often substantially, depending on the outcome in question.

Second, importance *varies systematically* across different settings and life arenas. Intelligence level tilts the odds of success and failure more in some arenas of life (such as academic achievement) than others (such as good citizenship). For instance, the correlations of IQ with years of schooling completed (.6) and composites of standardized academic achievement (.8) are over twice that for IQ correlations with delinquency (-.2 to -.3). Correlations in the same life arena can also vary depending on complexity of the tasks involved. For instance, correlations of job performance with test scores



range from .2 in unskilled work to .8 in the most cognitively demanding jobs.

Third, importance is *relative* to other known influences and one's particular aims. Many personal traits and circumstances can affect the odds of success and failure in different arenas of life. Intelligence is never "everything" in the practical affairs of life. Depending on the outcome in question, personality, experience, peers, family background, and the like can tilt the odds of success, sometimes more than intelligence does and sometimes less. As noted earlier, IQ predicts standardized achievement better than it does persistence in education, probably because personality and circumstances affect the latter much more than the former. Weak prediction at the individual level does not mean the predictor is unimportant in a pragmatic sense, as illustrated by the relation between delinquency and social class. The correlation is generally below  $-.2$  but usually thought quite important for policy purposes, as is the similarly low correlation at the individual level between smoking and various health risks.

Fourth, importance is *cumulative*. Small individual effects can be quite important when they cumulate across the many arenas and phases of one's life. Many of  $g$ 's daily effects are small, but they are consistent and ubiquitous. Like the small odds favoring the house in gambling, people with better odds win more often than they lose and can thus gradually amass large gains. Likewise, although smart people make "stupid" mistakes, they tend to accumulate fewer of them over a lifetime. Although the odds of any particular unfavorable outcome may not always be markedly higher in the lower IQ ranges, lower-IQ people face worse odds at every turn in life, meaning that their odds for experiencing *at least one* destructive outcome may be markedly higher.

*Education and training.* Schooling is the most g-loaded setting through which citizens pass en masse. Its unremitting demand is to learn and, moreover, to learn increasingly complex material as young people progress through it. It therefore highlights the intellectual distinctions among citizens better than does any other life setting and in ways plainly visible to the layperson. To be sure, schools enhance everyone's cognitive development, but they currently seem to have little impact on making

people either more alike or less alike in intelligence. Sociologist Christopher Jencks estimated in 1972 that if quantity and quality of schooling were made identical for everyone, such equalization would reduce the variance in test scores by only 20 percent. When Poland's Communist government rebuilt Warsaw after World War II, it allocated housing, schools, and health services without regard to residents' social class. This far-reaching equalization of environments did little or nothing either to equalize the IQs of the next generation of children or to reduce the correlation of their IQs with parental education and occupation (Stein, Susser, and Wald 1978).

As already noted, brighter students and military recruits learn much more from the same learning opportunities and often require less than one-fourth the exposure than do their less able peers for the same degree of learning. This difference in ability to capitalize on learning opportunities also greatly influences the maximum level of attainment youngsters are likely to reach. People with an IQ of 75 (the threshold for mental retardation) have roughly only a 50-50 chance of being able to master the elementary school curriculum; an IQ of about 105 is required for the same odds of getting grades good enough in high school to enter a four-year college; and an IQ of 115 is required for 50-50 odds of doing well enough in college to enter graduate or professional school.

Figure 3, similarly to Figure 1, summarizes accumulated employer experience about the most effective sorts of training for people at different ranges of IQ. Figure 3 is based on research with the Wonderlic Personnel Test (WPT), a short group intelligence test. Above the WPT equivalent of IQ 115-120 (which includes about 10 to 15 percent of the general white population), people can basically train themselves; the middle half of the IQ distribution (IQ 91-110) can learn routines quickly with some combination of written materials, experience, and mastery learning; but people below IQ 80 (10 percent of the general white population) require slow, concrete, highly supervised, and often individualized training. The military is prohibited by law from inducting anyone below this level because of inadequate trainability, and current minimum standards exclude anyone below the equivalent of IQ 85.

INTELLIGENCE

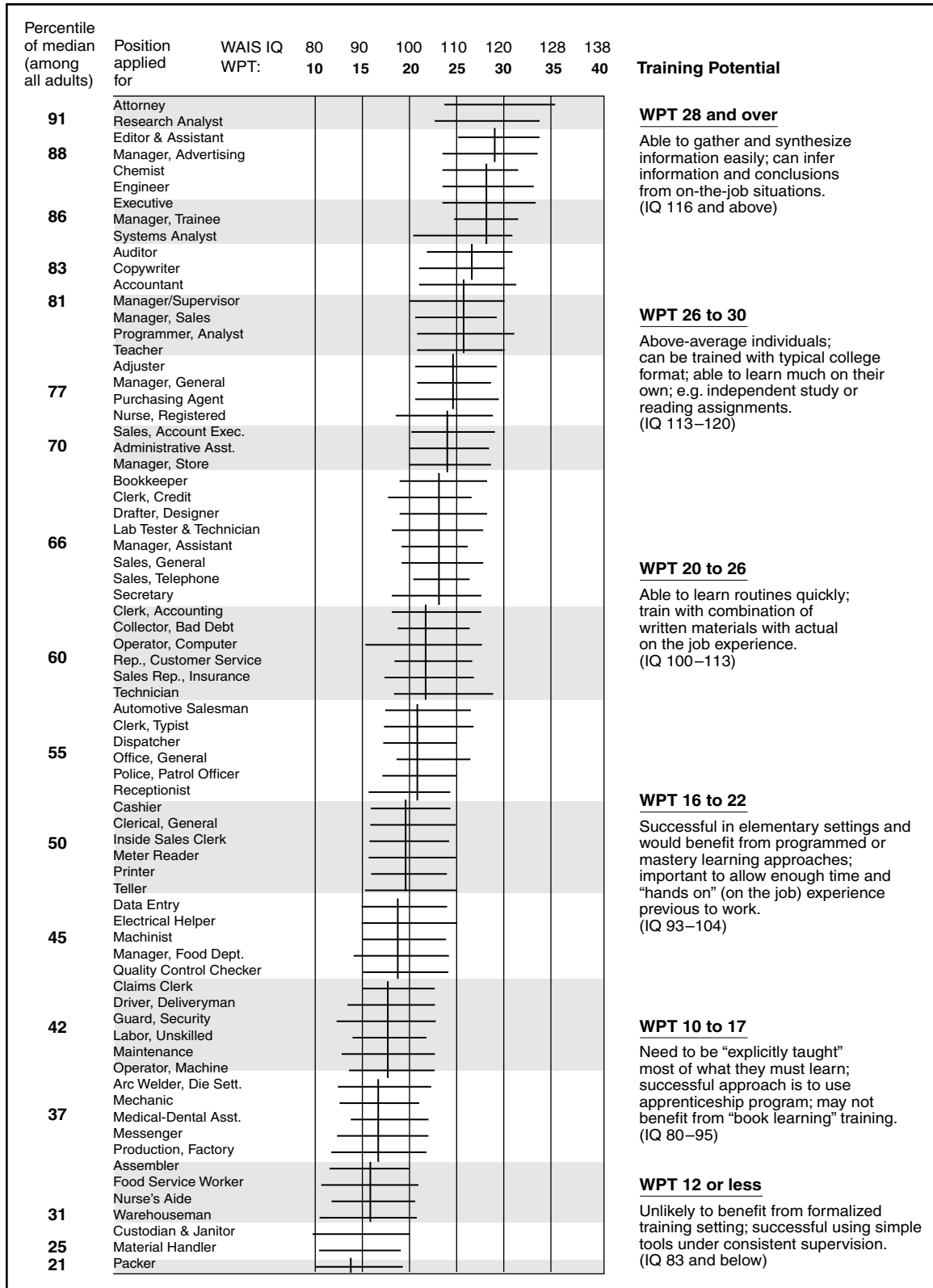


Figure 3

*Employment.* Many studies have found that the major distinction among occupations in the U.S. economy is the cognitive complexity of their constituent tasks. The most complex jobs are distinguished by their greater requirements for dealing with unexpected situations, learning new procedures and identifying problems quickly, recalling task-relevant information, reasoning and making judgments, and similar higher-order thinking skills that are prototypical of intelligence. Other job attributes that correlate highly with the *occupational complexity factor* include writing, planning, scheduling, analyzing, decision making, supervising, negotiating, persuading, instructing, and self-direction. So, too, do responsibility, criticality, and prestige. As already noted, mental ability test scores correlate most highly with performance in the most complex jobs. That is, differences in intelligence have a bigger impact on performance—"more bang for the buck"—when work is more g-loaded.

Not surprisingly, then, an occupation's overall complexity level correlates extremely highly with the average IQ of its incumbents. As Figure 3 illustrates, all occupations draw applicants from a wide range of IQ, but the minimum and average IQ levels rise with job level. Although wide, the typical IQ recruitment ranges for different occupations do not overlap at the extremes of job level (professional versus unskilled). The average IQ of applicants to middle-level jobs (about IQ 105), such as police work, is 15 IQ points (one standard deviation) lower than for applicants to professional jobs but 15 IQ points higher than for applicants to semiskilled work, such as food service worker. No occupation seems to recruit its workers routinely from below IQ 80.

The foregoing results for jobs suggest a high practical utility for higher intelligence in other aspects of life. Many jobs (child care, sales, accounting, teaching, managing) pose the same mental challenges (persuading, instructing, organizing, and ministering to people) that pervade nonpaid activities (parenting, home and financial management, civic responsibilities, friendships, and so on).

*Daily life.* Daily life has become considerably more complex during the twentieth century. Increased size, bureaucratization, and regulation of social institutions and services, together with greater

reliance on continually changing information technologies, have greatly increased the cognitive complexity of daily life. Life may be physically easier, healthier, and more pleasant today, but it has become mentally more challenging in developed societies. Some of this complexity is captured well by the U.S. Department of Education's 1992 National Adult Literacy Survey (NALS; Kirsch et al. 1993). Although the NALS was not designed as an intelligence test, it closely mimics the key attributes of an IQ test battery: Its intent was to measure complex information-processing skills by sampling a broad range of tasks from universally relevant contexts and contents; the relative difficulty of its items stems from their complexity, not their manifest content; and its three scales reflect one general factor.

Figure 4 illustrates items at different levels of the three NALS subscales; Figure 1 translates the NALS scores into IQ equivalents. These items do not involve esoteric "book smarts" but represent practical, everyday skills in dealing with banks, restaurants, transportation systems, and social agencies; understanding the news and one's options; and providing basic information about oneself. Nonetheless, about 15 percent of white adults and 40 percent of black adults routinely function no higher than Level 1 (225 or less), which corresponds to 80 percent proficiency in skills such as locating an expiration date on a driver's license and totaling a bank deposit. Another 25 percent of whites and 36 percent of blacks routinely function no higher than Level 2 (226–275), which includes proficiency in such skills as locating an intersection on a street map, entering background information on an application for a Social Security card, and determining the price difference between two show tickets.

These are examples of the myriad daily tasks that require some independent learning and reasoning as one navigates life. None may be critical by itself, but the more often one fails such tasks, the more one is hampered in grasping opportunities, satisfying one's needs and desires, and assisting family and friends. A national education panel concluded, in fact, that Level 1 and 2 skills are not sufficient for competing successfully in a global economy or exercising fully the rights and responsibilities of citizenship. Consistent with this, the

INTELLIGENCE

PROSE	DOCUMENT	QUANTITATIVE
149 Identify country in short articles	69 Sign your name	191 Total a bank deposit entry
210 Locate one piece of information in sports article	151 Locate expiration date on driver's license	
224 Underline sentence explaining action stated in short article	180 Locate time of meeting on a form	
	214 Using pie graph, locate type of vehicle having specific sales	
226 Underline meaning of term given in government brochure on supplemental security income	232 Locate intersection on a street map	238 Calculate postage and fees for certified mail
	245 Locate eligibility from table of employee benefits	246 Determine difference in price between tickets for two shows
250 Locate two features of information in sports article	259 Identify and enter background information on application for social security card	270 Calculate total costs of purchase from an order form
275 Interpret instructions from an appliance warranty		
280 Write a brief letter explaining error made on a credit card bill	277 Identify information from bar graph depicting source of energy and year	278 Using calculator, calculate difference between regular and sale price from an advertisement
304 Read a news article and identify a sentence that provides interpretation of a situation	296 Use sign out sheet to respond to call about resident	308 Using calculator, determine the discount from an oil bill if paid within 10 days
316 Read lengthy article to identify two behaviors that meet a stated condition	314 Use bus schedule to determine appropriate bus for given set of conditions	
	323 Enter information given into an automobile maintenance record form	325 Plan travel arrangements for meeting using flight schedule
328 State in writing an argument made in lengthy newspaper article		
347 Explain difference between two types of employee benefits	342 Identify the correct percentage meeting specified conditions from a table of such information	331 Determine correct change using information in a menu

(continued)

(continued)

PROSE	DOCUMENT	QUANTITATIVE
359 Contrast views expressed in two editorials on technologies available to make fuel-efficient cars	348 Use bus schedule to determine appropriate bus for given set of conditions (a harder question than the similar one above)	350 Using information stated in news article, calculate amount of money that should go to raising a child
362 Generate unfamiliar theme from short poems		368 Using eligibility pamphlet, calculate the yearly amount a couple would receive for basic supplemental security income
374 Compare two metaphors used in poem		375 Calculate miles per gallon using information given on mileage record chart
382 Compare approaches stated in narrative on growing up	379 Use table of information to determine pattern in oil exports across years	382 Determine individual and total costs on an order form for items in a catalog
410 Summarize two ways lawyers may challenge prospective jurors	387 Use table comparing credit cards. Identify the two categories used and write two differences between them	405 Using information in news article, calculate difference in times for completing a race
423 Interpret a brief phrase from a lengthy news article	396 Using a table depicting information about parental involvement in school survey to write a paragraph summarizing extent to which parents and teachers agree	421 Using calculator, determine the total cost of carpet to cover a room

Figure 4

NALS study found that, compared to adults with Level 5 skills (376–500, reached by about 4 percent of whites and less than 0.5 percent of blacks), adults with Level 1 skills were five times more likely to be out of the labor force, ten times more likely to live in poverty, only 40 percent as likely to be employed full time, and 7 percent as likely to be employed in a managerial or professional job—if employed at all.

Two daily activities where mental competence may have life-and-death implications are driving and health behavior. A large longitudinal study of Australian servicemen found that the death rate from motor vehicle accidents for men with IQs above 100 (52 per 10,000) was doubled at IQ 85–100 (92 per 10,000) and tripled at IQ 80–85 (147

per 10,000). The study authors suggested that the higher death rates might be due to poorer ability to assess risks. Medical research has likewise documented that many nonretarded patients have difficulty reading labels on prescription medicine and following simple physician instructions about self-care and future appointments.

*Nexus of social pathology.* Table 2 shows how the odds of social pathology fall (or rise) the further one’s IQ exceeds (or falls below) the average IQ. It shows the percentages of young white adults in five successive IQ ranges who experience certain bad outcomes. As shown, the odds of incarceration, illegitimate births, poverty as an adult, and the like all double at each successively lower IQ range. The ratios in the last column show how the

odds of bad outcomes thus differ greatly even for people who are only somewhat below average (IQ 76–90) versus somewhat above average (IQ 111–125) in IQ. Among these young white adults, for instance, 17 percent of the former IQ group versus only 3 percent of the latter live in poverty as adults, for a ratio of about 5:1. The odds are less discrepant for some bad outcomes (3:2 for divorce and unemployment) but more discrepant for others (7:1 for incarceration and 88:1 for dropping out of high school). The disparities in odds across IQ groups are even more extreme at the extremes of IQ. Good and bad outcomes can be found at all IQ levels, but what is *typical* differs enormously, as was also illustrated with the NALS data.

Moreover, the odds of dropping out of school, illegitimate births, poverty, and welfare dependence all increase with lower IQ among siblings within the very same family and even when the families are intact and not poor. There is something about below-average IQ itself that puts individuals at serious social risk, whatever their family circumstances.

*Overall life chances.* Figure 1 and Table 2 together paint a vivid picture of how greatly overall life chances differ by IQ level. People with IQs below 75 are clearly in the “high-risk” zone, where trainability and employability are very low and the odds of various social pathologies are much elevated. Although risks fall substantially for individuals with IQs only somewhat below average (IQ 76–90), these people still face an “uphill battle” because they are not very competitive for many training programs and jobs. The middle 50 percent of the population (IQ 91–110) is competitive for many of a modern economy’s jobs but likely only to be just “keeping up” relative to others. Their brethren of somewhat above-average IQ (IQ 111–125) are more likely “out ahead” socioeconomically because they are highly trainable and competitive for better jobs. Their rates of pathology are also very low. People with IQs above 125 are so competitive cognitively and so seldom hobbled by *g*-related social pathology that socioeconomic success is truly “theirs to lose.”

**Interpersonal Context.** One of the most fascinating questions in the study of intelligence has received virtually no attention: How does the *mix* (average and variability) of intelligence levels in a

setting—its *IQ context*—affect behavior in that setting? How might one’s fate be affected by the intelligence level of the *other* people in one’s interpersonal settings—of one’s parents, siblings, neighbors, friends, and other close compatriots?

The basic issue is this: A difference in IQ of one standard deviation (about 15 points) is socially perceptible and meaningful. Interpersonal communication becomes fraught with increasing difficulty beyond this distance because of larger gaps in vocabulary, knowledge, and ability to draw inferences or “catch on,” as well as the emotional discomfort such gaps create. Figure 1 reveals how IQ ranges of about one standard deviation also mark off substantial differences in options for education, training, and career, and thus the likelihood of entering different social niches. As shown in the figure, the normal range of intelligence (IQ 70–130, which includes roughly 95 percent of the general white population) spans four standard deviations of IQ. Socially and cognitively, that is an enormous difference. How, then, do people communicate and congregate across the IQ continuum in their daily lives? The average difference between siblings and spouses is about 12 IQ points, which means that most people in a biological family fall within the range of ready cognitive communicability. Any two random people in the population, however, differ by 17 IQ points, which represents the borderline for communicating effectively and as social equals.

*Communication, cooperation, and reciprocity.* The ability to communicate as equals constitutes a social tie, as does the ability to trade information and assistance. Such reciprocity is the basis of longer-term cooperation. Lack of reciprocity creates not only social distance but also animosity where reciprocity had been expected. There are many bases for cooperation and reciprocity, but sharing information and helping to solve problems is crucial in many settings. Ethnographic studies of middle school children, for instance, show how patterns of mutual assistance and friendship, rather than resentment and unwillingness either to provide help to classmates or to seek it from them, evolve from similarities and differences in students’ competence in answering homework and test items. Similar *g*-driven interpersonal relations can be expected in many workgroups

**Percentage of Young White Adults with Particular Life Outcomes, by IQ Level**

<b>LIFE OUTCOME</b>	<b>IQ: 75 and Below "Very dull"</b>	<b>76-90 "Dull"</b>	<b>91-110 "Normal"</b>	<b>111-125 "Bright"</b>	<b>Over 125 "Very Bright"</b>	<b>Ratio of Dull to Bright</b>
Out of labor force 1+ mo/yr (men)	22	19	15	14	10	4:3
Unemployed 1+ mo/yr (men)	12	10	7	7	2	3:2
Divorced in 5 yrs	21	22	23	15	9	3:2
% of children below IQ 75 (mothers)	39	17	6	7	–	2:1
Had illegitimate child (women)	32	17	8	4	2	4:1
Lives in poverty	30	16	6	3	2	5:1
Went on welfare after first child (women)	55	21	12	4	1	5:1
Ever incarcerated/doing time (men)	7	7	3	1	0	7:1
Chronic welfare recipient (mothers)	31	17	8	2	0	8:1
High school dropout	55	35	6	0.4	0	88:1

*Table 2*

SOURCE: Herrnstein and Murray (1994): (respectively) 158, 163, 174, 230, 180, 132, 194, 247/248, 194, 146.

and other settings in which teammates depend on one another for technical competence.

People of markedly different ability levels also tend to have different interests, which further impedes their ability to develop rapport. Assortative mating studies show that individuals explicitly seek mates of similar IQ levels and that spouses' IQs are, in fact, moderately correlated (about .4), perhaps more so than any other personal characteristic (except gender). Cognitive incompatibility is certainly responsible for the extreme social isolation often experienced by both the mentally retarded and the highly gifted. Extremely gifted children, who may be four standard deviations or more above average (IQ 160 and above), often feel, and are treated as, alien. These children are as different from the borderline gifted (IQ 130) as the latter are from the average child (IQ 100). With extraordinary vocabularies for their age, the highly gifted speak virtually a different language from their agemates. Although less extreme, the same type of alienation develops across much smaller

gaps in IQ. In short, cognitive similarity seems to affect the formation of social bonds, which themselves are the building blocks of "social structure."

*Social separation and segregation.* Because rough similarity in *g* promotes interpersonal reciprocity and rapport, it should not be surprising that people segregate themselves somewhat by cognitive ability when free to do so, marriage being the most intimate example. Segregation occurs along IQ lines for other reasons as well, many related to the functional value of intelligence in obtaining higher education and better work.

In the typical school, students enter grade 1 spanning at least mental ages four to nine, which translates quickly into markedly different grade-equivalent achievement levels. By reducing *g* variability within learning groups, ability grouping and tracking represent schools' attempt, albeit a perennially controversial one, to accommodate students of different cognitive levels. Its pedagogical merits aside, grouping reinforces friendships within IQ ranges and is but the first of many ways by

which schools and employers direct individuals toward different occupational and income groups, and thence into residential neighborhoods, partly along IQ lines.

A 1933 epidemiological survey in New York City documented that the *average* IQ levels of white school children across a large sample of the city's 273 Health Areas ranged from 74 to 118, a range of three standard deviations. The parents of these children would differ even more in average IQ. Consistent with genetic expectations, parents of any ability level produce children at virtually all ability levels, but their children's average IQ is closer to the population average than is their own.

Social clustering along IQ lines can be expected to increase familiarity, communication, and mutual assistance by enhancing within-group similarity, at least when the groups are minimally competent. Enhanced similarity can elevate the risks of low IQ, however, when IQ clustering results in a *critical mass* of individuals below some *critical threshold* in IQ. That threshold may be IQ 75, which is the level below which individuals need considerable assistance from family, friends, or social agencies to live independently in modern societies. When critical mass is reached in a family or community, networks of competent help become overwhelmed by sticky webs of bad judgment, which in turn produce a physically unhealthy and socially dysfunctional environment for all members, as sympathetic social anthropologists have documented.

In any case, greater *within-group* similarity produces greater *between-group* dissimilarity and distance. A contested but reasonable hypothesis of Richard Herrnstein and Charles Murray's 1994 book, *The Bell Curve*, is that society is becoming *increasingly* stratified along cognitive lines, jeopardizing national unity. That specter raises much anxiety in democratic societies, perhaps accounting for the quick distaste the thesis roused in many quarters. Any societal divisions that *g* creates would, however, be softened somewhat by *g*'s genetic basis. The laws of genetics guarantee that many children will differ substantially from their parents, producing intergenerational mobility across IQ and social-class lines and thereby assuring some cross-group ties. Whether or not it is increasing over time or permeable in nature, social clustering by *g* is nonetheless considerable. It is therefore a

perennial matter of public debate, whether the question be where to locate Section 8 or other public housing or how to integrate social classes and races in educational settings.

*Social networks and subcultures of attitudes, behavior, and knowledge.* *The Bell Curve*'s thesis about the dangers of cognitive stratification rests in its assumption that different cognitive strata create distinct and somewhat discordant cultures. Sociologist Robert A. Gordon (1997) has outlined at the level of small groups how different IQ contexts do actually represent different subcultures. These different subcultures in turn expose their members to different experiences, risks, knowledge, opinions, assistance, and expectations, as suggested earlier. IQ-discrepant subgroups, for example, differ not so much in the social ideals they espouse as in tolerance for their violation. They also differ in the degree to which they diffuse news and information from the broader culture rather than propagate rumor, misinformation, and even the AIDS virus.

The New York City neighborhoods mentioned earlier differed not only in IQ but also in rates of birth, death, infant mortality, and juvenile delinquency, illustrating that different IQ contexts probably constitute notably different social milieus for developing children. Children of, say, IQ 100 surely live different lives with different opportunities when raised in IQ contexts of 85 versus 115, both of which are common in the United States. Not only is such a child substantially above average in the first context while below average in the second, which creates its own opportunities and obstacles for the child, but there are also significant differences across the two contexts in the quality of ambient advice, information, and personal examples. Children's IQ levels seem not to be permanently affected by their IQ contexts, but their more malleable behaviors and outcomes may be, as studies of youthful career aspirations and delinquency suggest. Epidemiological analyses of the *g*-related contagion of certain risky health and social behaviors would further illuminate how risks rise or fall according to the level of "local intelligence" in which one is embedded.

**Societal Level.** The interpersonal contexts that influence an individual's behavior are themselves shaped partly by the *g* levels of the people inhabiting them, as just described. IQ contexts thus represent an impact of *g* on an individual level that is



above and beyond the effects of that individual's own IQ. IQ contexts have "macro" as well as "micro" effects in a society, however, because they create gradients of information flow, status and stigma, power and influence across a nation. These societal-level effects of  $g$ , via IQ contexts, may be the most important of all for a society, and they cry out for sociological analysis. Only a few such analyses have been done, but they illustrate the promise of a *sociology of intelligence*.

*Evolution of social structures.* If knowledge is power, then brighter people can be expected to advance further in any society freely allowing its accumulation. What is less obvious, except in hindsight, is that the routes to success may themselves be shaped by enduring variation in  $g$  within a population. Wide dispersion in  $g$  is a biological fact that all societies must accommodate. What norms and institutions evolve to promote such accommodation, especially where  $g$  has high functional value?

Consider the occupational hierarchy, that gradient of occupations from high to low in status, income, and educational requirements, which sociologists have shown to be replicated around the world. The consequences for individuals of their placement in it is clear, but its evolution is not. As described earlier, the major dimension underlying the hierarchy seems to be the complexity, not the content, of the tasks comprising the occupations arrayed along it. The occupational hierarchy is, then, a set of stable task configurations ranked in desirability according to their  $g$ -loadedness.

The structural question is how tasks gradually become sorted over time by their  $g$ -loadings into more  $g$ -homogeneous sets *within* occupations, thereby creating sharper distinctions in  $g$ -loading *between* occupations. This segregation of tasks by  $g$ -loading into a  $g$ -based occupational hierarchy most likely gradually arises from the natural sorting and reassignment of people *and* tasks to each other in the effort to improve aggregate performance of an organization's or society's essential functions. When workers are sorted more consistently by  $g$  level into occupations, occupational content can evolve to better fit the typical incumbent. For example, employers can gradually remove easy tasks from, and add complex tasks to, jobs whose usual incumbents are bright, and do the opposite for

jobs typically peopled by less bright workers (Gottfredson 1985).

Of course,  $g$  is hardly the only contributor to job performance, and job performance is not the only basis for how work and workers are organized in firms and societies. But to the extent that  $g$  is the most functionally important worker attribute overall *and* that people become sorted to work by  $g$  level, there will arise a  $g$ -based occupational hierarchy whose distinctions gradually expand or contract when the  $g$ -related efficiency of sorting workers rises or falls. This theory illustrates how the biological fact of differences in  $g$  can constrain the evolution of social institutions. That biological fact clearly rules out common utopian fantasies in which all citizens are assigned, rotated through, or ascend to jobs of equal difficulty and status.

*Racial politics.* When two social groups differ substantially in average  $g$  and  $g$  has functional value, they can also be expected to differ in  $g$ -related outcomes. The average difference in outcomes will depend on, among other factors, the size of the average group difference in  $g$  and the  $g$ -loading of the outcome in question. The  $g$ -generated differences in outcome have many sociopolitical reverberations, because they are pervasive, frequently large, and sometimes involve races once subjugated. The societal-level reverberations have the power to alter many aspects of a nation's culture. This can be illustrated by the national effort in recent decades to eliminate racial disparities in education and employment despite continuing racial disparities in  $g$ .

A key practical dilemma for educators and employers is that unbiased, valid measures of mental ability are generally the best predictors of school and job performance but, owing to phenotypic differences in  $g$  across racial groups, they have considerable *disparate impact*. That is, they screen out disproportionate numbers of candidates from some races. Unless group disparities in  $g$  are eliminated, there will continue to be a trade-off between selecting the most able applicants and selecting a racially balanced student body or work force, especially in highly  $g$ -loaded settings such as graduate school and the professions. In both employment law and public perceptions, unequal selection rates by race constitute *prima facie* evidence of illegal discrimination, often making it risky to use  $g$ -loaded predictors.

This combination of scientific facts and legal constraints has precipitated in personnel selection psychology a desperate but unsuccessful search for non-*g* substitutes for mental tests. There turns out to be no substitute for higher-order thinking skills. This failure created additional pressure on the field to reduce employers' legal vulnerability while retaining mental tests by instituting racial preferences. Eventually the U.S. Congress banned the most efficient such "solution" as an undisguised quota (the race-norming of employment tests, which means ranking applicants on separate racial curves). That ban in turn increased the pressure to covertly reduce or eliminate the *g* component of tests (to remove crucial mental demands), the results of which led to enormous controversy—and litigation—in personnel selection psychology. The same controversial effort to reduce the *g*-loading of employee selection criteria is now occurring for college admissions in states where racial preferences have been banned or might be. Being the most *g*-loaded predictor of student performance, the SAT has been the first target. In short, *g*-related group differences in outcomes have long been driving widespread changes in standards for admission, hiring, promotion, and more, sometimes improving selection and sometimes not, but always causing controversy.

Selection psychology is only one microcosm for observing the sorts of societal waves created by *g*-related group disparities. Virtually every school practice, from instructional and grouping practices to discipline to teacher assignment and funding, has been modified in recent decades to neutralize either the reality or the appearance of racial differences in phenotypic intelligence and their real-world effects. Keen disappointment at the failure of these modifications to accomplish that neutralization has itself sparked mutual recriminations between blacks and whites, led to more expansive definitions of discrimination and racism, and in many other ways shifted national politics. As is apparent, the societal-level ramifications of group differences in *g* hinge critically not only on how large they are and whom they affect but also on how a society explains and reacts to the differences.

*Inequality and the democratic paradox.* A population's IQ bell curve may bunch up or spread out

somewhat with environmental change, and it may shift a bit up or down the IQ continuum over time. Nonetheless, it will remain as much a biological fact as are differences in height. The bell curves for different demographic groups may also shift somewhat relative to each other along the IQ continuum, but gaps will likely persist.

As indicated in Figure 1, the IQ continuum represents a gradient of functional advantage for the individuals and groups arrayed along it. Happiness and regard may be available to all, but money, power, and prestige all tend to flow up the continuum, especially in a free society. Accordingly, envy flows up and stigma down. The IQ continuum is thus a strong current deep within the body politic that influences its surface dynamics in myriad ways and can frustrate efforts to steer a society in certain directions. Perhaps for this reason, political efforts to regulate or defy those dynamics have sometimes been violent in spirit if not in act. A 1980 analysis of genocides earlier in the century found that all but one of the targeted groups (Gypsies) were of apparently *higher* average intelligence than those seeking to exterminate them, for instance, the Jews in Germany, Armenians in Turkey, Ibos in Nigeria, and the educated classes in Cambodia.

Any humane society will moderate the effects of unequal biological and social advantage, preventing unbridled competition and the degradation of its weaker members. If resources naturally flow up the IQ continuum, societies can consciously redistribute some of them back down it—in a word, by helping. Such is the realm of charity and, increasingly, social policy, although such measures are seldom conceived in terms of helping the less "able" because that in itself would be stigmatizing. More often today, help is couched in terms of assisting the "deprived," as though all social inequality were the result of some social groups illegitimately expropriating from others what would have otherwise naturally accrued to them. Some inequality may be, but much is not.

Extreme egalitarianism is as problematic, however, as unbridled individualism, for it hobbles talent and deadens ambition. John Gardner outlined the trade-offs between promoting individual merit and equalizing social outcomes in his 1984 book, *Excellence: Can We Be Equal and Excellent Too?*

In that eloquent little book, he asked the question that writers from both the political left and right have since tried to answer in more detail: How can we create a valued place for people of all ability levels and bring out the best in all? The proffered answers differ primarily in the difficult trade-offs the authors settle for among personal liberty, equality of socioeconomic outcomes, and an emphasis on human excellence and productivity, three principles that are somewhat inconsistent owing to meaningful differences among people.

If such are the political dilemmas that the deep current of *g* inevitably creates, the debates over their resolution seldom seem cognizant of the dilemma's roots in human variation. Democracy is itself a social leveler because it grants *political* equality to people who are in numerous ways *biologically* unequal. But this strength is also its torment, because democracy excites the desire for yet more leveling, to which biological inequalities—especially intelligence differences—pose an obstacle. Mother Nature is no egalitarian. As Alexis de Tocqueville observed almost 200 years ago ([1835, 1840] 1969), “When there is no more hereditary wealth, class privilege, or prerogatives of birth, and when every man derives his strength from himself alone, it becomes clear that the chief source of disparity between the fortunes of men lies in the mind . . . [T]here would still be inequalities of intelligence which, coming directly from God, will ever escape the laws of man” (pp. 457–458, 538).

Biological diversity in *g* is a core challenge to democratic societies and to the scholars who are responsible for helping citizens understand how their society works. The challenge is exacerbated as technology advances, because such advance favors higher-*g* over lower-*g* people owing to their better ability to capitalize on it. Western democracies view democracy and technology as their twin engines of progress, however, and so haplessly seek solutions to inequality by pursuing yet more of both. That is the *democratic paradox*. The answer to the dilemma lies not in pursuing the opposite strategy—that is, curtailing both democracy and technology, as is sometimes hinted—but most likely in better understanding how differences in *g* orchestrate and constrain social life, to the extent that they do. For sociologists of intelligence, there is much to do.

## REFERENCES

- Anastasi, A. 1996 *Psychological Testing*. New York: Plenum.
- Brown, H., R. Prisuta, B. Jacobs, and A. Campbell 1996 *Literacy of Older Adults in America: Results from the National Adult Literacy Survey*. NCES 97-576. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Carroll, J. B. 1993 *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York: Cambridge University Press.
- 1997 “Psychometrics, Intelligence, and Public Perception.” *Intelligence* 24(1):25–52.
- de Tocqueville, A. (1835, 1840) 1969 *Democracy in America*. New York: Harper Perennial.
- Fox, W. L., J. E. Taylor, and J. S. Caylor 1969 *Aptitude Level and the Acquisition of Skills and Knowledges in a Variety of Military Training Tasks*. (Technical Report 69-6). Prepared for the Department of the Army by the Human Resources Research Office. Washington, D.C.: The George Washington University.
- Gardner, H. 1983 *Frames of Mind*. New York: Basic Books.
- Gardner, J. 1984 *Excellence: Can We Be Equal and Excellent Too?* New York: Harper.
- Goleman, D. 1995 *Emotional Intelligence*. New York: Bantam.
- Gordon, R. A. 1997 “Everyday Life as an Intelligence Test: Effects of Intelligence and Intelligence Context.” *Intelligence* 24(1):203–320.
- Gottfredson, L. S. 1985 “Education as a Valid but Fallible Signal of Worker Quality: Reorienting an Old Debate About the Functional Basis of the Occupational Hierarchy.” In A. C. Kerckhoff, ed., *Research in Sociology of Education and Socialization*, vol. 5. Greenwich, Conn.: JAI Press.
- Gottfredson, L. S. 1997a “Mainstream Science on Intelligence: An Editorial with 52 Signatories, History, and Bibliography.” *Intelligence* 24(1):13–23.
- Gottfredson, L. S. 1997b “Why Intelligence Matters: The Complexity of Everyday Life.” *Intelligence* 24(1):79–132.
- Herrnstein, R. J., and C. Murray 1994 *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Jencks, C., M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson 1972 *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Harper and Row.
- Jensen, A. R. 1998 *The g factor: The Science of Mental Abilities*. Wesport, Conn.: Praeger.

- Kaufman, A. S. 1990 *Assessing Adolescent and Adult Intelligence*. Boston: Allyn and Bacon.
- Kaus, M. 1992 *The End of Inequality*. New York: Basic Books.
- Kirsch, I. S., A. Jungeblut, L. Jenkins, and A. Kolstad 1993 *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*. Princeton, N.J.: Educational Testing Service.
- Lichtenstein, P., and N. L. Pedersen 1997 "Does Genetic Variance for Cognitive Abilities Account for Genetic Variance in Educational Achievement and Occupational Status? A Study of Twins Reared Apart and Twins Reared Together." *Social Biology* 44:77-90.
- Loehlin, J. C. 1992 *Genes and Environment in Personality Development*. Newbury Park, Calif.: Sage.
- Matarazzo, J. D. 1972 *Wechsler's Measurement and Appraisal of Adult Intelligence*, 5th ed. Baltimore: Williams and Wilkins.
- Neisser, U., G. Boodoo, T. J. Bouchard, A. W. Boykin, N. Brody, S. J. Ceci, D. F. Halpern, J. C. Loehlin, R. Perloff, R. J. Sternberg, and S. Urbana 1996 "Intelligence: Knowns and Unknowns." *American Psychologist* 51:77-101.
- Plomin, R., J. C. DeFries, G. E. McClearn, and M. Rutter 1997 *Behavioral Genetics*, 3rd ed. New York: W. H. Freeman.
- Rowe, D. C. 1994 *The Limits of Family Influence: Genes, Experience, and Behavior*. New York: Guilford.
- 1997 "A Place at the Policy Table? Behavioral Genetics and Estimates of Family Environmental Effects on IQ." *Intelligence* 24(1):133-158.
- Scarr, S. 1997 "Behavior-Genetic and Socialization Theories of Intelligence: Truce and Reconciliation." In R. J. Sternberg and E. Grigorenko, eds., *Intelligence, Heredity, and Environment*. New York: Cambridge University Press.
- Stein, Z., M. Susser, and I. Wald 1978 "Cognitive Development and Social Policy." *Science*, 200 23:1357-1362.
- Sternberg, R. J. 1985 *Beyond IQ: A Triarchic Theory of Human Intelligence*. New York: Cambridge University Press.
- Vernon, P. A. 1993 *Biological Approaches to the Study of Human Intelligence*. Norwood, N.J.: Ablex.
- Wonderlic Personnel Test, Inc. 1992 *Wonderlic Personnel Test and Scholastic Level Exam: User's Manual*. Libertyville, IL: Author.

LINDA S. GOTTFREDSON

## INTERGENERATIONAL RELATIONS

Throughout recorded history, concern has been expressed about relations among the generations. Historians have identified changing patterns of relationships between the old and the young, pointing out that in some epochs veneration of the aged was common, while in other eras, the aged were more likely to be held up to scorn and ridicule. In contemporary American society, these contrasts are muted, and themes of both consensus and conflict are present.

Sociologists have explored intergenerational relations extensively, using both macrosociological and microsociological approaches. Scholars who have taken a macrosociological approach have examined the discontinuity caused by the succession of different groups of individuals who were born during the same time period and therefore age together (Foner 1986). Sociologists refer to such groups as "cohorts." Many important questions have been raised regarding relations among cohorts, including: How do people differ as a result of membership in a specific cohort? How and why do cohorts come into conflict with one another? Does a "generation gap" exist?

In contrast, sociologists who have taken a microsociological approach have focused on intergenerational relations within families. These scholars have examined the content and quality of relationships among family members in different generations, posing such questions as: How much contact do adult children have with their parents? What kinds of exchanges occur between older and younger generations? What is the role of grandparents in families? Under what circumstances does conflict among the generations in families occur? To fully understand intergenerational relations, it is essential to study both levels and to draw connections between them.

### MACROSOCIOLOGICAL PERSPECTIVES

**Mannheim's View of Generations.** Karl Mannheim provided one of the most enduring analyses of relations between cohorts (he used the term "generation," however, instead of the contemporary sociological term "cohort"). Mannheim argued that the individuals born into a given cohort experience the same set of sociopolitical events while

they are growing up; this distinguishes them as a special social group. Merely by their location in a given cohort, members are likely to have certain similarities, since they are endowed with “a common location in the historical process” (1952, p. 290).

Thus, position within a cohort—like position within the socioeconomic structure—limits members to a narrow range of possible experiences and predisposes them to characteristic modes of thought. These differences can lead to conflict between the cohorts, as younger cohorts try to impose their views on society. The older cohort, on the other hand, has a major stake in preserving the existing social order. The interaction between these divergent cohort groups, according to Mannheim, is a critically important aspect of human social life.

To be sure, Mannheim did not compare belonging to a cohort with belonging to a more concrete group—such as a family—in that a cohort lacks a clear organizational framework. Further, he noted that differences may exist within cohorts. That is, within the same cohort, subgroups (in Mannheim’s terms, “generational units”) can form that are different from, and may even be antagonistic toward, one another. Nevertheless, Mannheim viewed location in a cohort as a powerful influence on people’s lives in much the same way that class position is an influence on their lives. This concern with the continual succession of cohorts, and its effects on social life, has found its clearest contemporary expression in age-stratification theory.

**Age-Stratification Theory.** Age-stratification theory begins with the fundamental assumption that to understand intercohort relations, we need to see society as *stratified* by age. Consistent with Mannheim, this view holds that society is divided into a hierarchy of socially recognized age strata. Each stratum consists of members who are similar in age and whose behavior is governed by the same set of norms for behavior appropriate for their age group. Further, members of various age strata differ in their abilities to obtain and control social resources. For example, young people in most societies have less power and fewer resources than middle-aged adults.

The duties, obligations, and privileges associated with age strata vary according to individual attributes, but they are always influenced by the structural aspect of age. Thus, various cohorts in a

society may have greatly divergent views on filial responsibility, expectations for independence of children, and values on other issues.

Sociologists see such differences among cohorts as the basis for possible conflicts of interest in society. In fact, conflict regarding continuity and discontinuity has been a major theme in macrosociological approaches to intergenerational relations. As Vern Bengtson (1989, p. 26) notes, members of the older cohort desire continuity: They want to transmit to younger cohorts “what is best in their own lives.” Correspondingly, they fear discontinuity: that young people will choose to live by very different sets of values. What has come to be known as the “problem of generations” reflects “the tension between continuity and change, affirmation and innovation, in the human social order over time” (Bengtson 1989, p. 26).

As an example, an age-stratification perspective on intergenerational relations can be applied to the political realm. When the age-stratification system is viewed as analogous to other stratification systems (e.g., class or gender), it follows that group solidarity may develop within each age cohort and that conflict—both overt conflict and conflict of interests—may occur between two different cohorts.

On a basic level, members of older and younger cohorts find different political issues more salient; for example, the elderly are more likely to focus on old-age pensions and health benefits, while the young are concerned with issues such as educational loans or the military draft (Riley et al. 1988). Such differences in salience may result in differences in voting behavior; for example, older individuals may be less likely to support tax increases for educational spending (Button 1992).

Political ideology among older and younger cohorts has been examined, with the fairly consistent finding that the current aged cohort is generally more conservative than younger cohorts. (However, these cross-sectional differences do not highlight the fact that people change their political attitudes over time, in line with changes in the society as a whole; for a discussion of political shifts during the life course, see Alwin 1998.)

Do divergent interests and attitudes result in age-related collective political action? Sociologists

have examined whether the aged constitute a self-defined political group that sets its agenda against those of other age strata. This question may be particularly important because studies have shown that older people are more likely to register and vote than younger individuals (Binstock 1997). Certainly Washington politicians and the media view the elderly as a political powerhouse; however, it is not clear that a voting block can be organized around old-age interests, as political attitudes vary greatly within the cohort. In addition, the elderly are a heterogeneous group; there are differences in socioeconomic status and racial and ethnic background within the aged cohort. Finally, the interests of the "young-old" (65-75) may differ from those of the "old-old" (over 75), with the former more concerned with retirement issues, income maintenance, and leisure opportunities, and the latter more interested in funding for medical services and long-term care. The overall evidence shows that the aged are willing to act together on some issues (like Social Security) but not on others.

A political development of the 1990s, however, may create more polarization among cohorts and thus lead to an upswing in age-based politics. This is the rising concern over "generational equity." Gordon Streib and Robert Binstock (1988) have summarized the issue in the following way. The elderly in developed nations were seen as a disadvantaged group from the 1950s to the early 1970s. They were portrayed as having low economic and social status, compared with younger persons. However, by the late 1970s, some scholarly and popular literature began to assert that old people had in fact *overbenefited*, and the elderly came to be seen as potentially burdensome economically to younger generations.

The older generation has at times been viewed as a scapegoat for a number of problems (Binstock 1983). In particular, the elderly have come to be seen as demanding resources for themselves, thus depriving children of quality schooling, health care, and other services. Organizations have arisen whose goal is to advocate the interests of the young at the expense of the older cohort. These groups have the expressed goal of establishing "generational equity," by transferring resources back to the young.

In the face of such attacks, it is conceivable that the elderly will begin to coalesce into a more

unified generational unit. However, in response to these claims, a countermovement has developed that encourages cooperation among advocates for youth *and* the elderly. On the level of practice, this interest has led to the development of intergenerational programs that bring old and young together.

While the macrosociological approach just discussed provides important insights into intergenerational relations and has obvious relevance for public policy on aging issues, it is also important to examine intergenerational relations on the microsocial level.

#### MICROSOCIOLOGICAL APPROACHES: A FOCUS ON THE FAMILY

Many sociologists have also focused on the smaller world of the family in an attempt to further our understanding of intergenerational relations. The family resembles the larger society, in that it is the locus of both intergenerational consensus and conflict. There is considerable family solidarity, indicated by feelings of affection and attachment that result from a shared history and close contact (Bengtson et al. 1990; Silverstein and Bengtson 1997), as well as inequalities of power and social resources. These twin themes of solidarity and conflict are evident throughout sociological research on the topic.

The way in which these themes are worked out in families has been affected by the dramatic changes in the age structure of American society. In particular, average life span has increased, which means that family members will spend more time than ever before occupying intergenerational family roles. Further, increased life expectancy leads to a greater likelihood that families will spend longer periods of time caring for disabled elderly relatives.

Societal changes have also increased the complexity of intergenerational relations. For example, the high divorce rate found in contemporary American society raises the likelihood that adult children will return to their parents' homes, often bringing their own young children with them. Women's unprecedented participation in the labor force and their return to college in great numbers may also affect intergenerational relations. To be sure, the acquisition of these nonfamilial roles provides new and enriching opportunities for women; however, it may alter the time that has

traditionally been devoted to “kinkeeping” between the generations.

It is only possible to comment on a few major themes in this review. The most widely studied area is that of parent–child relations in later life, including patterns of intergenerational contact and factors that affect the quality of adult child–elderly parent relationships.

**Contact Between Parents and Children.** A major concern of researchers has been to understand patterns of contact between adult children and older parents. Research on this issue has gone through two major phases. First, there was a period in which the nuclear family was held to be isolated. This view was based in part on Talcott Parsons’s analyses of family relations, which held that modernization had brought about the decline of the extended family (DeWit and Frankel 1988). During this period, it was widely believed that because of the geographic mobility of children, families abandoned their elderly relatives. This view also held that most elderly persons rarely saw their children and that family members no longer provided care for older relatives (Shanas 1979).

In the second phase, many prominent researchers devoted considerable effort to demonstrating that this view is inaccurate. Investigators such as Ethel Shanas, Marvin Sussman, and Eugene Litwak, as well as later researchers, clearly established that older persons have frequent contact with family members and that few are totally isolated from kin. Further, most aged family members are involved in a network of emotionally and instrumentally supportive relationships.

Studies have shown that in most cases parents and adult children have relatively frequent and regular contact (Umberson 1992); in fact, 40 to 50 percent of adult children see their parents at least once a week (Rossi and Rossi 1990). A major factor in determining the frequency of contact is physical proximity. Numerous studies have found that the frequency of intergenerational contact is greatly affected by geographical distance between households, with more distant children interacting less often with parents (cf. Eggebeen 1992). However, it is clear that many geographically distant children continue to interact to a significant degree with parents and that parents and children are able to maintain close ties despite being separated by great distances.

**Determinants of Quality of Parent–Child Relationships.** Researchers have moved beyond simply establishing patterns of contact to examining factors that affect the quality of parent–child relations in later life. For a review of this literature, see Suitor and colleagues (1995). A number of factors appear to have an impact on relationship quality.

*Increased parental dependency* is frequently cited as a factor that negatively affects the quality of aged parent–adult child relations. Studies have highlighted imbalanced exchanges and perceptions of inequity between the generations as major causes of family disharmony. For example, several investigations have suggested that an increase in parents’ dependence upon their adult children may reduce positive feelings between the generations. Other studies have found that adult children’s feelings of closeness and attachment are reduced when parents’ health declines. As parents’ health deteriorates, adult children are likely to need to increase their levels of support to previously independent parents, as well as to accept a lessening or termination of the parents’ provision of support—thus disrupting the previously established flow of support between the generations.

A second factor affecting parent–child relations is *gender*. Research has shown that women tend to have stronger and more supportive ties with parents than do men (Silverstein et al. 1995; Spitze et al. 1994). Further, studies of the effects of gender consistently demonstrate stronger affectional ties between mothers and daughters than any other combination (Rossi and Rossi 1990). Mothers report more positive affect with adult daughters than sons and are more likely to rely on daughters than sons as confidantes and comforters. Daughters in turn report greater feelings of closeness to mothers and are more likely to turn to them as confidantes than to fathers.

Third, the *age of the child* affects relationships with parents. Theories of adult development and intergenerational relations lead to the expectation that a child’s age will be negatively related to parent–child conflict and positively related to closeness. This literature suggests that maturational changes are likely to reduce differences between parents and adult children, thus minimizing the bases for conflict between them. For example, Bengtson (1979) suggests that as children mature,

their orientations become more similar to those of their parents. Similarly, Gunhild Hagestad (1987) posits both that differences between parents and children become muted across time and that there is greater tolerance for differences that remain. Empirical findings have consistently supported these assertions.

Fourth, changes in the degree of *status similarity* between adult children and their parents may affect their relationship. In particular, some studies have found a pattern of increased closeness in intergenerational relations when children begin to share a larger number of adult statuses with their parents. For example, the mother–daughter relationship appears to assume greater importance from the daughters’ perspective when they themselves become mothers (Spitze et al. 1994; Umberson 1992). Conversely, decreases in status similarity may negatively affect adult child–parent relations. For example, the status dissimilarity that develops when daughters surpass their mothers educationally may have particular potential for creating difficulties between the generations (Suitor 1987). For a review of research regarding parent–child relationships across the life course, see Pillemer and Suitor (1998) and Suitor and colleagues (1995).

Two other areas that have received considerable attention are relations with grandparents in the family and the importance of changing dependencies among the generations over the life course.

**Grandparenthood.** In recent years, sociologists have shown increasing interest in studying the role of grandparents in families (see Robertson 1995 for a review of contemporary grandparenting). The demographic shifts noted above have brought about changes in the nature of grandparenthood in several ways. First, more people now survive to become grandparents than ever before. Second, the entry into grandparenthood is likely to occur in midlife, rather than in old age; thus, the duration of grandparenthood may extend to four decades or more. Third, the role of grandparent is not clearly defined in American society, and the normative expectations, privileges, and obligations are ambiguous (Hagestad 1985).

Studies have revealed several consistent findings about grandparenthood. Contrary to popular stereotypes, most grandparents do not wish to

take on a parental role toward their grandchildren. Rather, they generally prefer a more distant role in the grandchildren’s lives. Nevertheless, grandparents are often a critical resource for families in times of trouble (Hogan et al. 1993).

Perhaps the most consistent finding in more than a quarter-century of research is diversity in grandparenting styles. A number of typologies have been identified, which usually array grandparents along a continuum from intense involvement and assumption of parental responsibilities, on the one hand, to relative alienation from grandchildren, on the other (Neugarten and Weinstein 1964).

The most ambitious study to date used a representative survey to examine styles of grandparenting and to uncover factors that determine the adoption of a particular style (Cherlin and Furstenberg 1986). Andrew Cherlin and Frank Furstenberg were able to identify five basic grandparenting styles. *Detached* grandparents have little contact with their grandchildren. *Passive* grandparents visit somewhat more frequently but carefully maintain a distance from their grandchildren’s lives. *Supportive* grandparents have more contact and focus on providing services to the grandchildren. *Authoritative* grandparents exert parent-like influence to a relatively great degree. Finally, *influential* grandparents combine involvement both by providing services and by adopting a parental role.

Cherlin and Furstenberg found that geographical distance was the most critical factor in determining which style of grandparenting developed. Detached grandparents were much more likely to live far away, while the most involved grandparents lived in close proximity to the grandchildren. Further, grandparents practiced “selective investment” in their grandchildren. They had more intense relationships with some grandchildren and more distant relationships with others. Other studies have affirmed this finding and have also pointed to the importance of gender: Grandmothers tend to have closer relationships with grandchildren than grandfathers.

Research interest has also highlighted the effects of children’s marital disruption on the relationship between grandparents and grandchildren. Despite suggestions in the popular media that relations between grandchildren and grandparents are damaged when adult children divorce,



studies have shown that this is not necessarily the case. Cherlin and Furstenberg (1986) found that “custodial grandparents” (that is, those whose adult children were awarded custody of the grandchildren) tended to maintain very close ties with their grandchildren, while “noncustodial grandparents” were less likely to maintain such ties. Since custody of minor children continues to be awarded more frequently to mothers than fathers, this means that ties with maternal grandparents are more likely to be maintained or strengthened following a divorce, while ties with paternal grandparents are more vulnerable—particularly in terms of frequency of interaction.

Jeanne Hilton and Daniel Macari (1997) also found that geographic proximity and whether the grandparents were related to the custodial parent were important factors in structuring the grandparent–grandchild relationship. In a study of New York families, Hilton and Macari (1997) found that grandparents were more involved with their grandchildren when they lived nearby and were related to the custodial parent; they also found that grandmothers were more involved with grandchildren than were grandfathers. Further information on the effect of divorce on parent, adult-child, and grandchild relationships can be found in Johnson (1993).

**Changing Dependencies.** Social scientists have convincingly demonstrated that children and parents continue to depend on one another for both emotional and instrumental support throughout the life course (Spitze and Logan 1992). The literature on intergenerational relations, however, has focused more heavily on children’s support to elderly parents than the reverse. Thus, before turning to a brief discussion of the issue of caregiving to the elderly, it is important to emphasize the *reciprocal* nature of intergenerational assistance.

Marvin Sussman (1985) has provided a model of parent–child relations across the life span that emphasizes the cyclical shift in relations with parents. In the beginning, parents provide a substantial amount of assistance to their offspring, even into the children’s early married life. Then, as children become more independent—and possibly move away—there is a decrease in intergenerational helping. Finally, as elderly parents begin to decline in health, they come to depend more

heavily on their children. This cyclical view of support is important, in that it stresses patterns of *mutual* aid between the generations (see also Rossi and Rossi 1990). Further, it should be noted that it is usually only very late in the life course that children’s provision of support exceeds that of parents, and even then only when parents become frail and disabled (Spitze and Logan 1992).

Throughout the 1990s, several hundred articles were published examining the experiences of families at this stage of the life course. Research has documented the strains experienced by middle-aged children when parents become dependent, including practical problems of managing competing demands on their time and energy as well as emotional stress, increased social isolation, guilt, and feelings of inadequacy.

Besides establishing both the prevalence of support for aged parents and problems in providing care, researchers have attempted to determine who is most likely both to become a caregiver and to experience the greatest stress from caregiving. Two of the most consistent findings involve the issues of gender and the relationship to the elderly person. First, women are substantially more likely to become caregivers than are men, and the role of caretaking appears to result in more intensive “hands-on” activities for women than men (cf. Allen 1994; Arber and Ginn 1995; Coward and Dwyer 1990; Finley 1989; Lee et al. 1993). Second, when the elderly person is married, it is the spouse who almost always becomes the primary caregiver (Stoller 1992). Thus, adult children generally become caregivers only when the care recipient’s spouse is not available to occupy this status. In addition, the nature of caregiving is different for spouses than for other caregivers. Spouses are more likely to have sole responsibility for caregiving and to have responsibility for providing assistance with routine activities of daily living (Stoller 1992). Further, spouses appear to experience greater physical and financial strain than do adult child caregivers and the are more likely to continue caregiving at the expense of their own well-being (Horowitz 1985). For a further discussion of caregiving, see Dwyer (1995).

Although research has focused on the detrimental consequences of caregiving, some more recent evidence suggests that most caregivers can also identify positive consequences of caregiving.

These positive aspects usually involve feelings of gratification derived from helping someone they love and fulfilling expectations of filial responsibility. Thus, the issue of changing dependencies in later life reflects the twin themes of consensus and conflict evident throughout theory and research on intergenerational relations.

## REFERENCES

- Allen, Susan M. 1994 "Gender Differences in Spousal Caregiving and Unmet Need for Care." *Journal of Gerontology: Social Sciences* 49:S187-S195.
- Alwin, Duane F. 1998 "The Political Impact of the Baby Boom: Are There Persistent Generational Differences in Political Beliefs and Behavior." *Generations* 12:46-54.
- Arber, Sara, and Jay Ginn 1995 "Gender Differences in Informal Caring." *Health and Social Care in the Community* 3:19-31.
- Bengtson, Vern L. 1979 "Research Perspectives on Intergenerational Interaction." In Pauline K. Ragan, ed., *Aging Parents*. Los Angeles: University of Southern California Press.
- 1989 "The Problem of Generations: Age Group Contrasts, Continuities, and Social Change." In V. L. Bengtson and K. W. Schaie, eds., *The Course of Later Life: Research and Reflections*. New York: Springer.
- , Carolyn Rosenthal, and Linda Burton 1988 "Families and Aging: Diversity and Heterogeneity." In R. H. Binstock and L. K. George, eds., *Handbook of Aging and the Social Sciences*, 3rd ed. San Diego, Calif.: Academic Press.
- Binstock, Robert H. 1983 "The Aged as Scapegoat." *The Gerontologist* 23:136-143.
- 1997 "The 1996 Election: Older Voters and Implications for Policies on Aging." *The Gerontologist* 37:15-19.
- Button, J. W. 1992 "A Sign of Generational Conflict: The Impact of Florida's Aging Voters on Local School and Tax Referenda." *Social Science Quarterly* 73:786-797.
- Cherlin, Andrew J., and Frank F. Furstenberg, Jr. 1986 *The New American Grandparent: A Place in the Family, A Life Apart*. New York: Basic Books.
- Coward, R. T., and J. W. Dwyer 1990 "The Association of Gender, Sibling Network Composition, and Patterns of Parent Care By Adult Children." *Research on Aging* 12:158-181.
- DeWit, David J., and B. Gail Frankel 1988 "Geographic Distance and Intergenerational Contact: A Critical Assessment and Review of the Literature." *Journal of Aging Studies* 2:25-43.
- Dwyer, Jeffery W. 1995 "The Effects of Illness on the Family." In Victoria Bedford and Rosemary Blieszner, eds., *Handbook of Aging and the Family*. Westport, Conn.: Greenwood Press.
- Eggebeen, David J. 1992 "Family Structure and Intergenerational Exchanges." *Research on Aging* 14:427-447.
- Finley, Nancy J. 1989 "Theories of Family Labor as Applied to Gender Differences in Caregiving to Elderly Parents." *Journal of Marriage and the Family* 51:79-86.
- Foner, Anne 1986 *Aging and Old Age: New Perspectives*. Englewood Cliffs, N.J.: Prentice-Hall.
- Hagestad, Gunhild O. 1985 "Continuity and Connectedness." In V. L. Bengtson and J. F. Robertson, eds., *Grandparenthood*. Newbury Park, Calif.: Sage.
- 1987 "Dimensions of Time and the Family." *American Behavioral Scientist* 29:679-694.
- Hilton, Jeanne M., and Daniel P. Macari 1997 "Grandparent Involvement Following Divorce: A Comparison in Single-Mother and Single-Father Families." *Journal of Divorce and Remarriage* 28:203-224.
- Hogan, Dennis P., David J. Eggebeen, and Clifford C. Clogg 1993 "The Structure of Intergenerational Exchanges in American Families." *American Journal of Sociology* 98:1428-1458.
- Horowitz, Amy 1985 "Family Caregiving to the Frail Elderly." In C. Eisdorfer et al., eds., *Annual Review of Gerontology and Geriatrics*. New York: Springer.
- Johnson, Colleen L. 1993 "Divorced and Reconstituted Families: Effects on the Older Generation." In L. Burton, ed., *Families and Aging*. San Francisco: Baywood Press.
- Lee, Gary R., Jeffery W. Dwyer, and Raymond T. Coward 1993 "Gender Differences in Parent Care: Demographic Factors and Same-Gender Preferences." *Journal of Gerontology: Social Sciences* 48:59-516.
- Mannheim, Karl 1952 *Essays on the Sociology of Knowledge*, trans. Paul Kecskemeti. London: Routledge and Kegan Paul.
- Neugarten, Bernice L., and K. K. Weinstein 1964 "The Changing American Grandparent." *Journal of Marriage and the Family* 26:199-204.
- Pillemer, Karl, and J. Jill Suitor 1998 "Baby Boom Families: Relations with Aging Parents." *Generations* 12:65-69.
- Riley, Matilda White, Anne Foner, and Joan Waring 1988 "Sociology of Age." In *Handbook of Sociology*. Newbury Park, Calif.: Sage.

- Robertson, Joan F. 1995 "Grandparenting in an Era of Rapid Change." In Victoria Bedford and Rosemary Blieszner, eds., *Handbook of Aging and the Family*. Westport, Conn.: Greenwood Press.
- Rossi, Alice S., and Peter H. Rossi 1990 *Of Human Bonding: Parent-Child Relations Across the Life Course*. New York: Aldine de Gruyter.
- Shanas, Ethel 1979 "Social Myth as Hypothesis: The Case of the Family Relations of Old People." *The Gerontologist* 19:3-9.
- Silverstein, Merrill, and Vern L. Bengtson 1997 "Intergenerational Solidarity and the Structure of Adult Child-Parent Relationships in American Families." *American Journal of Sociology* 103:429-460.
- Silverstein, Merrill, T. M. Parrott, and Vern L. Bengtson 1995 "Factors That Predispose Middle-Aged Sons and Daughters to Provide Social Support to Older Parents." *Journal of Marriage and the Family* 57:465-475.
- Spitze, Glenna, and J. R. Logan 1992 "Helping as a Component of Parent-Adult Child Relations." *Research on Aging* 14:291-312.
- , D. Glenn, and S. Zerger 1994 "Adult Children's Divorce and Intergenerational Relationships." *Journal of Marriage and the Family* 56:279-293.
- Streib, Gordon, and Robert H. Binstock 1988 "Aging and the Social Sciences: Changes in the Field." In R. H. Binstock and L. K. George, eds., *Handbook of Aging and the Social Sciences, Third Edition*. San Diego, Calif.: Academic Press.
- Suitor, J. Jill 1987 "Mother-Daughter Relations When Married Daughters Return to School: Effects of Status Similarity." *Journal of Marriage and the Family* 49:435-444.
- , Karl Pillemer, Shirley Keeton, and Julie Robison 1995 "Aging Parents and Aging Children: Determinants of Relationship Quality." In Victoria Bedford and Rosemary Blieszner, eds., *Handbook of Aging and the Family*. Westport, Conn.: Greenwood Press.
- Sussman, Marvin 1985 "The Family Life of Older People." In R. Binstock and E. Shanas, eds., *Handbook of Aging and the Social Sciences*. New York: Van Nostrand Reinhold.
- Umberson, Debra 1992 "Relationships Between Adult Children and Their Parents: Psychological Consequences for Both Generations." *Journal of Marriage and the Family* 54:664-674.

KARL PILLEMER  
SHIRLEY A. KEETON  
J. JILL SUITOR

## INTERGENERATIONAL RESOURCE TRANSFERS

The transfer of resources between individuals of different generations occurs on both the societal level, as the outcome of public policy or within the context of the private sector, and the family level, as in the exchange of emotional support and material goods. This discussion of the intergenerational exchange of resources will consider, first, the flow of resources within the context of the family; second, intergenerational transfers on the societal level; and last, the issue of equity in the transfer of resources intergenerationally. In addition, brief mention will be made of special circumstances, such as cross-cultural differences or the special case of divorce and stepfamilies.

The meaning of the word *generation* varies from setting to setting. In the family environment, the reference is primarily to lines of descent (grandparents, parents, children, etc.). However, each family member is also a member of a particular birth cohort, or group of individuals born during the same period of historical time, such as 1920-1924, and, at any given time, a member of a particular age group, such as sixty-five- to sixty-nine-year-olds. Each of these is also reflected in the intergenerational exchange within the context of the family and in the intergenerational exchange of a society's resources. Finally, the exchange process may also include the concept of generation that connotes a group of individuals, usually part of the same birth cohort, who share a common set of political or social beliefs (e.g., the "Woodstock Generation").

### RESOURCE TRANSFERS WITHIN THE CONTEXT OF THE FAMILY

Resources exchanged among family members vary greatly in size and type. Examples range from the ordinary, everyday exchange of care involved in child rearing or household chores to the bequest of substantial financial or material resources to descendants in a will. The latter behavior has been explored as part of a broader practice referred to as a "legacy" (Kane 1996), which includes last wills and testaments, the transfer of property, and, even more broadly, how people want to be remembered.

Over the last several decades a number of perspectives have developed regarding such familial exchanges. One such perspective views the resource-transfer process in terms of *reciprocity of exchanges among family members*. Analysts have examined reciprocity as a motivator of interdependence over time, at different points in the family structure (e.g., parent-child, sister-brother), and within the context of individual family members' perceptions of how much each member has given or received in the past. *Exchange theory* specifically views all kinds of social interactions as the exchange of rewards between individuals where the group or individual with the greater amount of *social power* regulates the exchange process (Dowd 1975). Regarding resources to older family members, Horowitz and Shindelman (1983) view reciprocity in terms of the "credits earned" by the older individual for providing resources in the past to the family member currently on the giving end.

In contrast, the *life-cycle model* of family intergenerational transfers maintains that the distribution of resources among the generations in the family takes on a curvilinear shape: Individuals in the middle generations, and most likely middle-aged, transfer the bulk of family resources to those who are either younger or older. (For an illustration, see the works of Reuben Hill [1965, 1970] regarding his study of three-generational families in the Minneapolis-St. Paul area). In contrast, the *role continuity model of family intergenerational transfers* asserts that, except in families where the older generations are financially strained, older family members redistribute their wealth to successive generations in the family (for further discussion, see Covey 1981; Kalish 1975; Riley and Foner 1971). Still another perspective used to explore the giving and receiving of resources in the family is that of *hierarchical/sequential* activities. From this perspective, individuals first seek aid from members of their nuclear family; failing that, from members of their extended family; and, only as a last resort, from institutional sources such as government programs, banks, and the educational system (viz., Morgan 1983). A final perspective is that of altruism. *Altruism* is a helping behavior that benefits the recipient but provides no benefits, and may incur costs, to the provider. In this context, then, resources are transferred from one generation to another with regard only to the

benefit they provide to the recipient, without regard to the costs incurred by the provider. Research has investigated the relative importance of exchange versus altruism as motivators for intergenerational transfers (Cox and Jakubson 1995; Cox and Rank 1992; Schoeni 1997). This research finds that intergenerational transfers are more likely to be part of an exchange process rather than behaviors conducted out of altruism.

A great variety of resources, both material and psychosocial in nature, are exchanged among members of nuclear and extended families. These include the giving and receiving of material gifts of all manner and size; the provision of help during emergencies, as when a family member is ill or in need of sudden and immediate shelter; providing physical or financial resources in time of need, such as babysitting or nursing, offering advice, taking care of household maintenance, and providing financial support; the passing on of family legacies, including pictures, artifacts, recipes, and family lore; and inheriting assets of all kinds upon the death of family members.

In Western societies prior to industrialization, inheritance was a mechanism for passing on from father to son the means of family economic support, the farm or business (milling, smithing, etc.) that traditionally remained within the family for generations. With the arrival of industrialization and the movement of economic support to the factory or the office, the content and impact of inheritance grew more varied. While monetary and property assets are still very common, items with more symbolic significance, such as mementos or heirlooms, are often the most cherished (Rosenfeld 1979; Schorr 1980; Sussman et al. 1970).

The provision of care to family members is a resource transfer that occurs in every emotionally healthy family the world over. Most caregiving involves the many ordinary, everyday activities that are necessary to the survival and well-being of family members, for example, the typical care and feeding of infants and children, the performance of daily chores and house repairs, the giving of advice and emotional support, and the nursing of relatives during episodes of acute illness. In some instances family members give or receive care for conditions that are quite taxing and of an indefinite duration, such as Down's syndrome in a child

or Alzheimer's disease or physical impairment or frailty in a spouse or older parent.

The intergenerational nature of this transfer is underscored in data from numerous studies, including nationally representative surveys, such as the one conducted by Louis Harris and Associates (1975) in conjunction with the National Council on the Aging, as well as many studies of family caregiving, including the National Long-Term Care Survey and the Longitudinal Survey of Aging. The general findings of this research indicate that the vast majority (perhaps as many as 80 percent) of older adults who need help with activities of daily living receive such help from relatives, particularly children or grandchildren (Stone et al. 1987; Stull 1995). A variety of motives for such care provision have been explored. Biegel and associates (1991) suggest that there are two types of explanations for the motives to provide care. Egoistic explanations argue that helping others is done with the anticipation that rewards will be forthcoming for such care provision. This includes payment for care, avoiding censure, gaining social approval, or complying with social norms. Some evidence for payments for care yields inconclusive results. For example, McGarry and Schoeni (1997) report that they found no evidence that parents provide financial assistance to their children in exchange for caregiving. On the other hand, Lillard and Willis (1997) report that in Malaysia, a country with neither Social Security nor Medicare, children are an important source of financial security in old age and that this is, in part, a repayment for parents' investments in children's education. They find further that parents and children engage in the exchange of time help for money.

A second set of explanations for providing care encompasses empathy and altruism. In these motivations, the interest is in benefiting the other, not oneself. As noted earlier, altruistic explanations of intergenerational transfers have not found much empirical support.

Noninheritance economic transfers within the family take a variety of forms: gifts, loans, payment of bills or down payments, and emergency financial help, for example. The directionality of this assistance, in both nuclear and extended families, is from parents of all ages to children of all ages and from adult children to older parents. Evidence (Schoeni 1997) indicates that people in

their twenties and thirties receive more time and money assistance than do people of other ages, including the very old.

The transfer of goods and services among family members includes sharing the same household; making major purchases such as furniture, a car, or a large appliance; making smaller purchases such as clothing; and doing major repairs. Data show that substantial intergenerational transfers occur. For example, 90 percent of those Americans 65 years of age and older who had children or grandchildren gave those offspring gifts, while between 92 percent and 96 percent of adults 18 to 64 years old gave gifts to parents or grandparents (Harris and Associates 1975).

**INTERGENERATIONAL RESOURCE  
TRANSFERS ON THE SOCIETAL LEVEL:  
FROM ONE BIRTH COHORT/AGE GROUP  
TO ANOTHER**

On the societal level, the concept of intergenerational transfers shifts from resource exchanges among specific individuals of the same lineage to transfers to and from large groups of people of one birth cohort or age group to large groups of people in the same or other birth cohort or age group. These exchanges take place in several arenas: (1) the public policy arena, where the transfer is usually in the form of resources defined in statutes and laws (examples in the United States include the Social Security system, Head Start, Temporary Assistance to Needy Families, school bonds); (2) the private sector, where transfers are in the form of wages or goods (examples include jobs that support workers and nonworkers, corporate profits that supply tax revenue); and (3) the creative arena (examples of these transfers include art, music, literature, and the fruits of social and physical science research).

**Products of Public Policy.** The outcomes of public policy are transfers that affect everyone in a society. Frequently, these resources are multifaceted, affecting a broad range of individuals. For instance, in the United States the term *Social Security* is commonly used to refer to Old-Age and Survivors Insurance retirement income and the death survivors' component of a larger piece of legislation that also includes a disability insurance program, an unemployment insurance component,

and the public assistance Supplemental Security Income program. As a totality, then, these components of the Social Security legislation go to a wide range of individuals in a number of different circumstances. Moreover, it is important to note that, while resources stemming from public policy may be received by specific individuals at a particular point in time or over a particular span of time, in another sense each has a larger impact that far exceeds that exchange. For example, the issuance of a school bond in a community directly and immediately provides educational resources to a particular cohort of school-age residents. However, indirectly, those resources affect the lives of other members of the family, resulting in improved quality of life for them as well. Additionally, from a longitudinal perspective, the resources invested in educational infrastructure and materials benefit future generations of the community's students.

Discussion about the relative contribution of public and private transfers has led to conflicting expectations. For example, if one expects altruistic motivations, then private transfers would neutralize the impact of public transfers. On the other hand, if exchange motivates transfers, then private transfers would reinforce the effects of public transfers on economic well-being. Cox and Jakubson (1995) find that private transfers can enhance public transfers, the reverse of previous expectations that public transfers merely supplant private ones.

**Outcomes of the Private Sector.** Intergenerational resource transfers resulting from private sector activity affect both the *active working population*, composed of individuals who are primarily in their young adult and middle years, and the so-called *dependent population*, composed of children and older people supported by the active working population. This transfer occurs through both the wages that workers earn and the tax revenues, funneling through the public sector, those workers provide. The precise age boundaries for these classifications vary from society to society according to culture and the primary means of economic support (agricultural or industrial). However, even within the age span of the active working population, one finds individuals who are part of the dependent population, such as those who are

disabled or unemployed in their age groups. Moreover, within the age span of the dependent population, one finds working school-age children and retirees working part-time or part-year.

In the United States, the *total dependency ratio* is generally considered to be the number of persons under eighteen and those age sixty-five and over (together, the two dependent populations) for every 100 working-age individuals eighteen to sixty-four years of age. The *youth dependence ratio* is generally the number of persons under eighteen for every 100 individuals eighteen to sixty-four, while the *old-age dependency ratio* is the number of persons sixty-five and older for every 100 persons eighteen to sixty-four years of age. In the United States, while the youth dependency ratio has declined from 51 in 1950 to 34 in 1994, the old-age dependency ratio climbed from 13.3 in 1950 to 20 in 1994. Additionally, the old-age dependency ratio is expected to increase to 37 by 2030, at the height of the so-called baby boom retirement years (U.S. Bureau of the Census 1996).

However, none of these ratios take into consideration the possible impact of the following on intergenerational resource transfers: (1) the current and future contributions of elderly persons and youth who are employed; (2) the potential for increased labor-force participation among those sixty-five and older; (3) the changes in the Social Security Act to gradually increase the age of eligibility for full Social Security retirement benefits to sixty-seven over a twenty-seven-year period, beginning in the year 2000; (4) the growing number of working women, who have different and, as yet, somewhat uncharted labor-force participation patterns; and (5) the growing economy (Crown 1985; Kingson et al. 1986).

**Current Cohorts of Children and the Heterogeneity of the Elderly: Two Special Considerations of Intergenerational Transfers.** In the United States, intergenerational resource transfers that are the products of public policy and private sector activity support individuals of all ages. Of growing importance, however, is a consideration of the effect of the nature of such transfers on two groups, current cohorts of children and the elderly, at any time when they are viewed as a heterogeneous group. Regarding the former, it is clear that, at present, many children live in poverty or near poverty and, thus, are not receiving from older

generations (either older family members or older generations on the societal level) the resources they need to live satisfactory lives as children. Moreover, limited access during childhood to such resources as health care and education, in particular, leads to reduced economic opportunities and productivity in adulthood. Thus, without the infusion of such resources, children are unable to prepare themselves for a productive adulthood in which they are capable of successfully joining the active working population to sustain themselves. Moreover, they will also be unable to provide support for those individuals, young and old, who will then compose a large part of the dependent population of the future. For instance, persons age seventeen and under in 1990 will be twenty to thirty-seven years of age in 2010, when the first wave of the baby boom cohort will begin to retire, and forty to fifty-seven years of age in the year 2030, when the full weight of the baby boom retirement will be felt.

Regarding the heterogeneity of the elderly, it is important to keep in mind that older people in Western societies vary greatly not only in chronological age (forty marks the onset of protection under the U.S. Age Discrimination in Employment Act; eighty-five is generally considered the beginning age of the "oldest-old") but also in income, health, and activity status. In addition, the elderly vary, on these and other factors, according to race and gender. As a result, in old age *chronological age alone is not an adequate indicator of the need for a substantial infusion of public or private sector resources*. Indeed, some older people are still greater producers than consumers of society's resources. Likewise, for any particular individual, certain times *during* old age require a greater amount of societal resources than other times.

**Outcomes of Creativity.** The fruits of research, music, literature, systems of jurisprudence, and other cultural and scientific products represent very important types of societal-level intergenerational transfers. As long-term transfers that cross the generations of a society and frequently move from one society to another, these are resources that traverse both time and geography. (For instance, disease-resistant hybrid strains of rice developed in the United States and used in agricultural settings around the world benefited not only Americans alive during and after the mid-1950s but also people in other societies over succeeding

decades.) Frequently, especially regarding the products of scientific research, those in the same birth cohort as the originator(s) or creator(s) do not feel the full impact, positive or negative, of the outcome themselves; this becomes part of the birthright, for better or worse, of the generations that follow. For example, those in the generation of the developers of the diphtheria/pertussis/tetanus vaccine, given routinely to infants, faced the risk of these diseases themselves during childhood. On the other hand, individuals in the same birth cohort as the developers of the atomic bomb have not lived the entire span of their lives, including childhood, coping with the implications of this scientific outcome.

### THE ISSUE OF EQUITY IN THE INTERGENERATIONAL TRANSFER OF RESOURCES

In the mid-1980s, the equity of intergenerational transfers was a prominent topic of discussion in both the press and various academic settings. This topic is receiving renewed interest. Consideration usually centers on the allocation process for resources resulting from public policy. However, it is important to keep in mind that much of the time these resources affect individuals within the context of the family. While there are many perspectives regarding the justification and procedure for transferring resources across generations, four particular factors appear to underlie each of these perspectives, each one subject itself to multiple meanings. When there is disagreement about the procedure for and direction of resource distribution across generations, it is frequently due to disagreement regarding the meaning of one or more of these factors (Hirshorn 1991). These are (1) the concept of the generation, discussed previously; (2) the issue of whether the resource transfer results in one party's experiencing an absolute or relative gain or loss in comparison with others; (3) the perception of the resource "pie" to be allocated (e.g., Is it constant in size? expandable?); and (4) the idea of *distributive justice*, or fairness in the rationale for deciding how to allocate resources (e.g., transfers can be based on level of need, on merit, or on equal shares for all, among other criteria).

**Intergenerational Inequity Perspective.** One of the most discussed perspectives in recent years

argues that there is *intergenerational inequity* in the transfer of resources in the United States, particularly resources resulting from public policy. It focuses on the relative welfare of current cohorts of youth and current cohorts of the elderly. This perspective maintains, moreover, that, in the United States in recent decades, the welfare of the young, especially those under the age of eighteen, has diminished as a result of the enhanced status of the elderly, who have been accumulating sizable proportions of the nation's personal wealth and, at the same time, been on the receiving end of the bulk of the resources stemming from social policy (Medicare and Social Security retirement funds are singled out especially). Moreover, those adhering to this view maintain, the absolute size of public expenditures directed at the elderly is a major cause of the current budget deficits and other economic problems facing the United States. The assumption is that sufficient funds will not exist in the future to ensure that those who are currently children and young adults will receive their fair share of these very resources—Social Security retirement and public sector health care support—in their own old age (Cornman and Kingston 1996). Among the problems with this perspective is that it assumes that the elderly, as a group, are all doing well in absolute terms; thus it does not take into consideration the variation within the older population that makes this group, which is quite heterogeneous regarding health status, economic status, living arrangements, and other factors. Moreover, this perspective relies on the idea that the correct concept for use in assessing the equity of public resources transferred across generations—one that would result in social justice—is *numerical equality*. Yet equal expenditures do *not* result in equal levels of return, no matter who receives the transfer. Finally, given the wide range of types of resources transferred within the context of the family and in the public and private sectors, it is meaningless even to try to arrive at an accurate measure of which generation/birth cohort/age group is faring better as a whole (Kingston et al. 1986).

**Common Stake Perspective.** An alternative view of intergenerational resource transfers notes the *common stake* that all generations/birth cohorts/age groups have in the wide variety of intergenerational exchanges. This perspective emphasizes the importance of using the concept of the *life*

*course* in assessing the giving and receiving of resources at all levels and in all contexts, societal and familial. Thus, it emphasizes that, at some points along the life course, one generally takes more of certain kinds of resources than one gives, while at other points along the life course the opposite is true (e.g., hands-on care is very strong in infancy and sometimes in old age; use of tax revenues for education is very prominent during childhood). Moreover, generally we do not give the same resources from the very same people, not within the context of the family or on the societal level. For instance, usually we do not provide the unstinting, continuous, and comprehensive care we received in infancy to our *parents*, but we do to our own children.

This perspective also stresses the fact that the same intergenerational resource transfer affects some individuals directly and others indirectly. For example, the public sector program Temporary Assistance to Needy Families, provides *direct* support to children in families headed by parents who are unable to provide sufficient work-related income; thus, family funds are freed to purchase such items as needed medical supplies for, say, an ailing grandparent. *Indirectly* affected by another public transfer that flows intergenerationally is the schoolchild, living with a grandparent, whose lunch money comes from the latter's monthly Social Security check. Finally, the common stake viewpoint underscores the need for current and future generations of the elderly to concern themselves with the welfare of children and young people and vice versa—for the sake of all concerned (Kingston et al. 1986).

(SEE ALSO: *Altruism; Filial Responsibility; Inheritance; Intergenerational Relations; Social Mobility*)

#### REFERENCES

- Biegel, David, E., E. Sales, and R. Schultz 1991 *Family Caregiving in Chronic Illness*. Newbury Park, Calif.: Sage.
- Cornman, James M., and E.R. Kingston 1996 "Trends, Issues, Perspectives, and Values for the Aging of the Baby Boom Cohort." *The Gerontologist* 36:15–26.
- Covey, H. 1981 "A Reconceptualization of Continuity Theory." *Gerontologist* 21:628–633.
- Cox, D., and G. Jakubson 1995 "The Connection Between Public Transfers and Private Interfamily Transfers." *Journal of Public Economics* 57:129–167.



- Cox, D., and M.R. Rank 1992 "Inter-Vivos Transfers and Intergenerational Exchange." *Review of Economics and Statistics* 74:305-314.
- Crown, William H. 1985 "Some Thoughts on Reformulating the Dependency Ratio." *Gerontologist* 25:166-171.
- Dowd, James J. 1975 "Aging as Exchange: A Preface to Theory." *Journal of Gerontology* 30:584-594.
- Harris, Louis, and Associates, Inc. 1975 "The Myth and Reality of Aging in America." Washington, D.C.: National Council on the Aging, Inc.
- Hill, Reuben 1965 "Decision Making and the Family Life Cycle." In E. Shanas and G. Streib, eds., *Social Structure and the Family*, Englewood Cliffs, N.J.: Prentice Hall.
- 1970 *Family Development in Three Generations*. Cambridge, Mass.: Schenkman.
- Hirshorn, Barbara 1991 "Multiple Views of the Intergenerational Flow of Society's Resources." *Marriage and Family Review*, special edition on intergenerational relations.
- Horowitz, Amy, and L. Shindelman 1983 "Reciprocity and Affection: Past Influences on Current Caregiving." *Journal of Gerontological Social Work* 5(3):5-20.
- Kalish, Richard 1975 *Late Adulthood*. Monterey, Calif.: Brooks/Cole.
- Kane, Rosalie A. 1996 "From Generation to Generation: Thoughts on Legacy." *Generations*, 20:5-9.
- Kingson, Eric R., B.A. Hirshorn, and J.M. Cornman 1986 *Ties That Bind: The Interdependence of Generations* (report from the Gerontological Society of America). Washington, D.C.: Seven Locks Press.
- Lillard, L.A., and R.J. Willis 1997 "Motives for Intergenerational Transfers: Evidence from Malaysia." *Demography* 34:115-134.
- McGarry, K., and R.F. Schoeni 1997 "Transfer Behavior Within the Family: Results from the Asset and Health Dynamics Study." *Journals of Gerontology Series B—Psychological Sciences and Social Sciences* 52:82-92, Special Issue.
- Morgan, James N. 1983 "The Redistribution of Income by Families and Institutions and Emergency Help Patterns." In G.J. Duncan and J.N. Morgan, eds., *Five Thousand American Families: Patterns of Economic Progress*. (Institute of Social Research 16). Ann Arbor: University of Michigan Press.
- Riley, Matilda W., and A. Foner 1971 "Social Gerontology and the Age Stratification of Society, Part 1." *Gerontologist* 11:79-87.
- Rosenfeld, J.P. 1979 *The Legacy of Aging: Inheritance and Disinheritance in Social Perspective*. Norwood, N.J.: Ablex.
- Schoeni, R.F. 1997 "Private Interhousehold Transfers of Money and Time: New Empirical Evidence." *Review of Income and Wealth* 4:423-448.
- Schorr, Alvin 1980 ". . . Thy Father and Thy Mother . . . A Second Look at Filial Responsibility and Family Policy." Washington, D.C.: Social Security Administration.
- Stone, Robyn, Gail L. Cafferata, and Judith Sangl 1987 "Caregivers of the Frail Elderly: A National Profile." *The Gerontologist* 27:616-626.
- Stull, Donald E. 1995 "Caregivers (Family)." In A.R. Davis, J. Boondas, and A. Lenihan, eds., *Encyclopedia of Home Care for the Elderly*.
- Sussman, Marvin, J. Cates, and D. Smith 1970 *The Family and Inheritance*. New York: Russell Sage Foundation.
- U.S. Bureau of the Census 1996 "Population Projections of the United States by Age, Sex, Race, and Hispanic Origin: 1995 to 2050." *Current Population Reports, Series P-25, No. 1130*. Washington, D.C.: U.S. Government Printing Office.

DONALD E. STULL  
BARBRIK HIRSHORN

## INTERGROUP AND INTERORGANIZATIONAL RELATIONS

The focus here is first on the meaning of intergroup relations and next on interorganizational relations. Although the emphasis in both cases is mainly on research and conceptualizations based in the United States, we consider to some extent research based in other nations and cultures. We also occasionally consider the uses of sociological research.

### INTERGROUP RELATIONS

A *group* is a collection of persons who have shared problems and act together in response to those problems, have shared expectations, and have a sense of common destiny. There are many kinds of groups, ranging from informal friendships to ethnic groups, to societies, and even to intersocietal groups. *Intergroup relations* refers to patterns of relationships that develop between groups. Extensive reviews of the literature in intergroup relations may be found in Seeman (1981) and Stephan (1985).

Although intergroup relations refers to all types of groups, it is not possible to avoid focusing on ethnic-racial group relations, because this has been the central concern in the United States since the social sciences and sociology, in particular, became established academic disciplines. Intergroup relations has been approached from the level of analysis of the group on the one hand, and from a social psychological perspective on the other.

This latter approach examines intergroup relations from the point of view of the individual and his or her relations with a group. In the latter part on the nineteenth century, the United States experienced huge immigration from southern, central, and eastern Europe. Since the new immigrants arrived from cultures that were markedly different from those of Americans who were already established here, conflicts developed, a pattern in some ways not materially different from conflicts that now exist between U.S. citizens and recent immigrants from Central and Latin American and Asia. The history of research on intergroup relations in sociology shows a profound emphasis on problems created by these massive immigrations.

Liebertson (1980) has thoroughly studied the question of why those who migrated to the United States from south, central and eastern Europe after 1880 in large numbers (such as Italians, Russians, Lithuanians, Poles) have been so much more successful than blacks. He found that blacks, in part because of their visibility, confronted much more serious social and competitive disadvantages than did these groups, even though the European groups did have many obstacles to overcome. Because of the slave period and the initial contacts with Africans, white society had and still has very unfavorable dispositions toward blacks, much more unfavorable than their dispositions toward white Europeans and even more unfavorable than their dispositions toward Asian immigrant groups. The competitive threat of Asian groups at the same time period was not nearly as great, since the Asian immigration was much smaller. Blacks also faced greater barriers than did Europeans from labor unions. Thus, Liebertson asserts that blacks and Europeans confronted intrinsically different situations that produced very different sets of opportunities for socioeconomic advancement.

Conflict and competition between groups is only one—albeit important—pattern that may develop, although it is the adaptation that is most likely to make headlines and be reported on the television news. William Graham Sumner (1906) applied the concept of ethnocentrism to explain intergroup conflict. Ethnocentrism is a “view of things in which one’s own group is the center of everything, and all others are scaled and rated with reference to it” (p. 13). Sumner believed that ethnocentrism served to highlight and then exaggerate differences between groups and hence contributed to in-group cohesion and strong hostility to the out-group.

One of the oldest social-psychological theories of intergroup relations stresses personality determinants. The leading exponent of this view is Gordon Allport (1954). Allport’s definition of prejudice has two components: attitude and belief. Allport defines ethnic prejudice as “an antipathy based upon a faulty and inflexible generalization” (p. 9). In his book, Allport refers to the generality of prejudice. A number of studies in this tradition report large intercorrelations of prejudices—that is, persons who are anti-Semitic are also anti-Catholic and antiblack (Epstein and Komorita 1965). The claim is that these views stem from a deep-seated personality syndrome, called by Allport “the prejudiced personality.” The prejudiced person is highly moralistic and has a need for definiteness, among other features. As Seeman (1981) has noted, the evidence on cross-group generality and its personality basis is easily challenged on both methodological and theoretical grounds. In general, the trait approach has slowly given way to a more situationalist view.

A note on discrimination and prejudice is in order, since much of the intergroup relations literature concerns these concepts. *Discrimination* is typically defined as treating people unequally due to their group membership, while *prejudice* is often seen as “a rigid emotional attitude toward a human group” (Simpson and Yinger 1972, p. 24). As Seeman (1981) points out, in prejudice persons are categorized wrongly or prejudged, and this process is a complex rather than a simple one:

*The error comes in misconceiving and misjudging such a group, and the individual members thereof, as a consequence of misreading the*

*nature of the category involved. Often enough the misreading occurs because the cultural and historical sources of supposed category qualities are not taken into account or are attributed to irrelevant features of the category, that is, to blackness, Jewishness, and so on. What makes all this extremely tricky is that (a) it is difficult to demonstrate what, in fact, the appropriate characterizations are for the social categories we find it necessary to employ; and (b) given the powerful control that majorities exercise, pressures are generated that tend to socialize the members of a given category into the very features we discern: to make Jews "intellectual," blacks "hostile," Chicanos "indolent," and women "dependent." Thus, though demonstrable relevance (correctness) of the attributed qualities to the category is critical, there is typically a seeming relevance (a misread relevance) that beclouds the issue both for the participant and the analyst (p. 380).*

Robert E. Park was one of the first leading sociologists of race relations and is identified with ecological theory (Park and Burgess 1921). Park viewed human beings as competing for territory much like plants and other animals. Ecological processes foster competition for limited resources. The distribution of the population is itself shaped by competition over scarce resources, and this competitive process structures the economic interdependence of groups and individuals. Competition also makes people aware of their status and induces them to view themselves as of superior or of inferior status depending on their social situation. Those who see themselves as superior express their consciousness of felt superiority and seek to maintain their privileged position through prejudice. Hence, the intergroup competitive process expresses itself through the development of moral and political order. This order is a product of such processes, in addition to competition, as conflict, accommodation, and assimilation. Racial consciousness is therefore seen as developing out of the competitive process, which is born in the competitive struggle for status. An "inferior" status group might not wish to compete and instead establish a niche within the division of labor, and this might lead to a stable equilibrium.

A dominant intergroup relations paradigm is derived from Marxist thought and is called conflict

theory. In this view, race relations and their consequences emerge from the system of social stratification in the society. Societies are seen as constantly changing. Societies distinguish between and among their members and award some greater rewards, such as power, prestige, and money, than others. The result is social inequality, which becomes an essential part of the stratification system. Hence, the stratification system is simply the structured inequalities of groups or categories or individuals. The different groups in the society, such as classes or ethnic-racial groups, compete for the limited resources available. Three conditions are necessary for intergroup conflict and inequalities to result (Cox 1948; Vander Zanden 1983).

First, there must be at least two identifiable groups. People must be aware of their group and another group on the basis of some characteristic or set of characteristics. These may be physical or not—that is, beliefs or values are sufficient.

Second, the two groups must compete with one another or feel they are competing for a limited pool of resources for themselves, if necessary at the expense of members of the other group.

Third, the two groups cannot have exactly the same amount of power, so that one group can claim an advantage in obtaining resources that another group also seeks. Under these conditions, one group becomes more dominant as the competition develops. The more powerful group defines the other group as inferior. As the group with less power seeks to protect and assert its interests, the dominant group may feel threatened and aggrandize to an even greater degree, and tension may mount. Most members of the dominant group soon find it easy and appropriate to view the other group in very negative terms.

*Assimilation* involves adapting another culture in place of one's native culture and usually is applied to the process that occurs when person adjust to a new society and culture by adopting it. For example, many Asian groups have recently immigrated to the United States. Typically their children learn English, dress in American style, eat American food, and are seen as and regard themselves as Americans. However, many adult Asians, although completely or partially bilingual, will retain an affinity for their native culture and understandably are more comfortable conversing in their original tongue.

*Accommodation* refers to a decision by two or more groups to put aside a significant difference that exists between them in order to stress common interests. This leads to cultural pluralism—that is, a number of different cultural patterns coexisting in the same society. The United States is a pluralist society in that it permits many distinctive religious, ethnic, and racial groups to exist side by side. The need of new Asian groups to retain their cultural identities is reflected in the existence in many large cities of Chinatowns, Koreatowns, and Japantowns. Some areas in large cities have signs only in Chinese or in both Chinese and English. This has produced some conflict with non-Chinese residents in these communities and has sparked “English-only” movements and resistance to bilingual instruction in public schools.

A great deal of research has been undertaken examining the impact of intergroup contact on intergroup hostility and prejudice (see reviews by Stephan 1985; Williams 1977). Many of the early studies looked at naturally occurring intergroup contacts. A substantial number of laboratory and field investigations have been undertaken focusing on those characteristics of intergroup contacts that foster positive intergroup outcomes. Stephan (1985) summarizes the findings with regard to this problem in a list of thirteen propositions, such as: “Cooperation within groups should be maximized and competition between groups should be minimized” and “Members of the in-group and the out-group should be of equal status both within and outside the contact situation” (p. 643).

There is a wealth of information and research on intergroup relations. However there is a definite lack of application of this information when dealing with intergroup conflicts, especially in the area of public policy. Brewer (1997) identifies several reasons for this gap between research and practice in reducing intergroup conflicts.

There have traditionally been different approaches to researching the processes involved in intergroup conflict. Research traditions focus on different levels of aggregation, with some focusing on interpersonal processes and others focusing on the group level of analysis. Additionally there are theoretical perspectives that study intergroup conflict with a primary focus on concepts in the cognitive, affective, or behavioral realm. These different

approaches tend to generate literatures that remain isolated, rarely citing research outside their own perspective.

While these separate research traditions might use different conceptual frameworks, they do have one thing in common that contributes to a lack of direct participation in the policy arena. Science has a norm of objectivity, and this leads many researchers to avoid advocating for specific action by governments or groups. Scientific research is seen as producing facts, and the role of the scientist is to produce those facts, not to decide what to do about them. The expert is hesitant to become the advocate.

History may contribute to this feeling, both for the researchers and the policy makers. In the 1950s and 1960s, much social science research (specifically the “contact hypothesis”) was used as the basis for public policy designed to reduce racial tension and conflict through the integration of schools. The research outcomes on desegregation were mixed, due in part to an oversimplified application of theoretical ideas that were very specific and conditional in nature. Additionally, the social science research on which desegregation was in part based was developed in carefully controlled laboratory experimentation. In real-life situations, which are far more chaotic and complex, the mechanisms of the theories might be overwhelmed by other factors, factors the theories were never intended to deal with. While the research community might be happy to learn from the failure of experiments and to argue about the failure of the assumptions of models to be met, the policy maker sees failure of a program and an increase in conflict among constituents. So the scientists see the politicians as understanding neither the restricted nature of most theories nor the process of the growth of knowledge as including failures. The policy makers see only that the experts were wrong and that their advice created conflict (or a perception of increased conflict) when a reduction was expected.

While research on desegregation was based on the contact hypothesis and ideas of assimilation, current research is based more on ideas of pluralism and multiculturalism. This perspective focuses on promoting positive in-group attitudes by emphasizing group identities and characteristics, and trying to make these identities respected

and recognized by other groups; to be proud of one's own ways while recognizing the pride of other's for their own ways. This creates some tensions and potential pitfalls when using this branch of theory and research as a basis for public policy.

Much of intergroup theory has focused on and developed from the situation in the United States; however, recently more attention and research have been applied to the problem in other areas of the world. With increasing migration and diversity in western Europe has come a concurrent increase in tension and conflict, and this has generated a surge of research. Pettigrew (1998) shows that "despite sharp differences in national histories, political systems, and minorities, this new work reveals considerable consistency across the nations of western Europe. It also largely replicates and extends, rather than rebuts, the North American literature" (p. 98). He shows similarities such as the higher level of "subtle" prejudice compared with "blatant" prejudice. Blatant prejudice is tied to perceived biological differences between groups and is "hot, close and direct" (p. 83). Subtle prejudice is tied to the "perceived threat of minority groups to traditional values" (p. 83) and the "lack of positive feeling towards minorities" (p. 83). In addition, the mechanisms of intergroup contact and relative deprivation seem to function in similar ways in many North American and European populations.

Given the tension between Arabs and Jews in the Middle East, it should not be surprising that theories of intergroup relations should be tested there in attempts to reduce conflicts. One such study attempted to use the contact between Jewish and Arab citizens of Israel that occurred in joint medical teams to mitigate intergroup stereotypes and prejudice (Desivilya, 1998). The study found that while the contact reduced prejudice in the local work situation, that reduction was not carried into the larger societal context and overall national image and ethnic stereotypes were not changed.

### INTERORGANIZATIONAL RELATIONS

An *organization* is a group with three main features: a goal or set of goals, a boundary, and a

technology (Aldrich 1979). Although we say organizations have goals, what is meant is that much of what organizations actually do seems as if it is directed to a shared objective or set of objectives. This may be for the sake of appearance, and/or the organization may really be goal-oriented. The boundary feature simply refers to the distinction that an organization makes between members and nonmembers. Finally, technology refers to the organization's division of labor or to the set of activities that the organization performs as part of its daily routines in processing new materials or people.

Each of these characteristics can be illustrated by the university. Its goals are often set forth, albeit in glowing and idealized terms, in its general catalogue. These typically include teaching, research, and public service. One must apply to become a member of the university—student, faculty or staff. And such statuses are frequently difficult to come by. The university's technology includes its classrooms and laboratories as well as the lecture and discussion methods of instruction.

*Interorganizational relations* refers to the relations between or among two or more organizations. There have been several overviews of the field of interorganizational relations (Aldrich 1979; Aldrich and Whitten 1981; Galaskiewicz 1985; Mulford 1984; Van de Ven and Ferry 1980).

Every organization has relationships with other organizations. In the case of the university, if it is to function it must have students, and to recruit them it must have relationships with high schools, junior colleges, and other universities. These students (and faculty and staff) must eat, work, and play, so the university has relationships with food, housing, energy, and other suppliers of various kinds in the community. And, of course, the university needs other resources, especially funds, and therefore must relate to government agencies and alumni to obtain them (Clark 1983).

Organizations are ambivalent about establishing an interorganizational relationship to obtain resources (Yuchtman and Seashore 1967). On the one hand, they want and need resources if they are to survive; but on the other hand, organizations wish to maintain their autonomy, and insofar as they establish an interorganizational tie, they will

be expected to reciprocate, and hence their freedom will be constrained. It is assumed that organizations want their autonomy from other organizations, but their survival needs induce them to relinquish some autonomy.

Galaskiewicz (1985) claims that interorganizational relations take place for three major reasons: to obtain and to allocate resources, to form coalitions to enhance power, and to achieve community acceptance or legitimacy.

Interorganizational relations research has been undertaken at three levels: the dyad (Hall et al. 1977), the action set (Hirsch 1972), and the network (Burt 1983; Galaskiewicz 1985). The simplest form of interorganizational relation is the dyad, which simply refers to the relationships of two organizations to each other. The action set concept developed from Merton's (1957) notion of role sets. Caplow (1964) and Evan (1966) took Merton's idea and applied it to the relationships between a focal organization, such as a university, and its pairwise relationship with other organizations with whom it interacts. One might examine the relationship between a university and the office of the mayor of the city within which it is located, and then study the effects of changes in this relationship as they influence other relationships in the set of organizations (Van de Ven and Ferry 1980). Aldrich (1979) has termed the group of organizations that constitute a temporary alliance for a particular or limited goal the "action set." Networks of organizations contain the complete set of ties that connect all the organizations in a population of organizations (Aldrich 1979; Hall 1987; Van de Ven and Ferry 1980). Although the approaches of Aldrich and Van de Ven and Ferry are not identical conceptually, both orientations toward networks focus on identifying all connections of specified kinds that take place within a particular organizational population. Hence, the analysis of networks is far more complex than that of action sets or dyads.

The body of knowledge in the area of interorganizational relationships is not extensive, and what there is has focused on social services. There exists quite a bit of theoretical information but very few large databases. With the exception of research on corporate board of directors' interlocks (Burt 1983; Burt et al. 1980), there is little work on the private sector.

An early area of interest to theorists was the general state of the organizational and interorganizational environment. Aldrich (1979) identified six dimensions of environments: capacity, homogeneity/heterogeneity, stability/instability, concentration/dispersion, domain consensus/dissensus, and turbulence. Capacity refers to the relative level of resources available in the organization's environment. A rich environment refers to one where resources are plentiful, while a lean environment is the opposite.

Homogeneity/heterogeneity refers to the extent to which organizations, individuals, or even social forces that influence resources are relatively similar or different. For example, does a focal organization deal with a relatively uniform and a highly heterogeneous population? If one contrasted the labor forces that a Japanese and an American firm draw from to recruit, one would find that the Japanese firm confronts a more homogeneous environment than does its American counterpart. This is, of course, because American workers are much more heterogeneous than are Japanese workers in education, ethnic-racial background, and many other features (Cole 1979).

Stability/instability concerns the degree of turnover in various elements of the environment. Again, if Japanese and American firms are compared, we would anticipate greater turnover in the latter than in the former. The advantage of low turnover or a stable environment is that it permits the organization to develop fixed routines and structures.

Aldrich (1979) refers to the extent to which resources are distributed evenly in the environment or concentrated in a particular area as concentration/dispersion. For example, the RTD is the major bus company in Los Angeles, and its potential ridership is dispersed over an area of more than 400 square miles. Such long transportation lanes present major problems, in contrast, for example, with the Santa Monica Bus Company, whose ridership is concentrated in a much smaller and geographically homogeneous area.

Organizations differ in the extent to which their claim to a specific domain is contested or acknowledged by other organizations. Domain consensus refers to a situation wherein an organization's claim to a domain is recognized, while domain

dissensus refers to a situation where disagreement exists over the legitimacy of an organization's domain.

The final organizational dimension is turbulence. This term refers to the extent to which there are increasing environmental interconnections; the more interconnections, the greater the turbulence. Areas where many new organizations are emerging are generally areas of greater turbulence.

Some factors have been shown to be involved in the dissolution of interorganizational relations. Institutional forces, power, and competition are factors involved in the stability of these relations (Baker et al. 1998). Institutional forces reduce the likelihood of ties being broken. Personal relationships and structural attachments (such as coordination of accounting methods) represent "sunk costs" for the organization. To break ties with one organization and forge them with another takes significant effort and time. Power may increase or decrease the tendency for relations to dissolve. Given a client-provider relation, if the client has high power it increases the likelihood of the relation being broken, while high power for the provider decreases the tendency for the relation to break. Competition tends to destabilize ties between organizations, although the effect of competition is weak relative to power and structural attachments. In situations where there is much competition, there are more opportunities for defection from a relation and more incentive to do so.

A great deal of the work on interorganizational relations has concerned delivery systems and stressed coordination (Mulford 1984; Rogers and Whetten 1982). This is because a central problem in service delivery involves overcoming the segmentation and fragmentation of services created by the large number of organizations with overlapping responsibilities and jurisdictions. Bachrach (1981) has identified a number of factors that discourage coordination among organizations serving the chronically mentally ill, including budget constraints, lack of a mandate to engage in interorganizational planning, and confusion due to separate funding streams for care. Other factors also discourage coordination, such as differences in organizational activities and resources; multiple network memberships and consequent conflicting obligations felt by constituent organizations; and a lack of complementary goals and role exceptions (Baker and O'Brien 1971).

Each organization in a delivery system relies on the other organizations in the system, since no single unit can generate all the resources necessary for survival. Hence, the organization in a system enters into exchanges with other organizations and consumers. It is assumed that each organization or system seeks to better its bargaining position. This perspective on delivery systems as interorganizational networks is generally labeled the resource dependence perspective (Pfeffer and Salanick 1978).

Contingency theory (Lawrence and Lorsch 1967) assumes that organizational functioning depends on the intertwining of technological and environmental constraints and the structures that emerge to deal with these constraints. The theory assumes, as does system theory more generally, that there is no single most effective way to organize (Katz and Kahn 1966), that the environment within which an organization functions influences the effectiveness of an organization, and that different organizational structures can produce different performance outcomes.

Scott (1998) uses systems theory to describe how organizations manage their task environments and their institutional environments, and adapt to their changing organization-environment interdependencies. Scott stresses the differences between technical and institutional environmental controls.

Although there is not a great deal of comparative research on interorganizational relations, some comparisons have been done, for example, on the differences in the patterns of relations in Japan and the United States. American companies tend to be connected to more organizations, have a more formalized and more extensive body of rules for the relationship, and exchange more information across the relations than their Japanese counterparts (Aldrich et al. 1998; Jang 1997). Claims such as these must be taken with a degree of caution, as there is a great deal of influence between Japanese and American firms. For example, Japanese auto assembly plants, which started in the United States, brought both intraorganizational and interorganizational patterns of organization, for example, team-based work groups and "just-in-time" delivery of parts needed in product assembly (Florida and Kenney 1991). Since that time,

these organizational characteristics have become more common in the United States, and this increasing interdependence and connection between organizations can be seen in the impact of labor disputes extending through the economy at an increasingly rapid rate. A strike at a parts production facility that makes transmissions can shut down numerous auto production plants in a few days. A strike at Federal Express or United Parcel Service leads to economic impacts nation- and world wide literally overnight.

This globalization of economic relations mirrors a globalization of international relations. Both on an ongoing basis (e.g., the G-7, or Group of Seven Economic Conferences; the European Union; the United Nations) and on a temporary basis (e.g., the Gulf War, U.N. peacekeeping missions), interorganizational relations are becoming more and more common and vital to the peace and prosperity of the world (Alter and Hage 1993).

## REFERENCES

- Aldrich, Howard 1979 *Organizations and Environments*. Englewood Cliffs, N.J.: Prentice-Hall.
- , Michele Kremen Bolton, Ted Baker, and Toshihiro Sasaki 1998 "Information Exchange and Governance Structures in U.S. and Japanese R&D Consortia: Institutional and Organizational Influences." *IEEE Transactions on Engineering Management*. 45(3):263-276.
- , and D.A. Whetten 1981 "Organizational Sets, Action Sets, and Networks: Making the Most of Simplicity." In P.C. Nystrom and W.H. Starbuck, eds., *Handbook of Organization Design*, vol. I. New York: Oxford University Press.
- Allport, Gordon W. 1954 *The Nature of Prejudice*. Reading, Mass.: Addison-Wesley.
- Alter, Catherine, and Jerald Hage 1993 *Organizations Working Together*. California: Sage Publications.
- Bachrach, Leona L. 1981 "Continuity of Care for Chronic Mental Patients: A Conceptual Analysis." *American Journal of Psychiatry* 138(11):1449-1456.
- Baker, Frank, and Gregory O'Brien 1971 "Intersystem Relations and Coordination of Human Service Organizations." *American Journal of Public Health* 61(1):130-137.
- Baker, Wayne E., Robert R. Faulkner, and Gene A. Fisher 1998 "Hazards of the Market: The Continuity and Dissolution of Interorganizational Market Relationships." *American Sociological Review* 63(2):147-178.
- Brewer, Marilyn B. 1997 "The Social Psychology of Intergroup Relations: Can Research Inform Practice?" *Journal of Social Issues* 53(1):197-212.
- Burt, Ronald S. 1983 *Corporate Profits and Cooptation: Networks of Market Constraints and Directorate Ties in the American Economy*. New York: Academic Press.
- , K.P. Christman, and H.C. Kilburn, Jr. 1980 "Testing a Structural Theory of Corporate Cooptation: Interorganizational Directorate Ties as a Strategy for Avoiding Market Constraints on Profits." *American Sociological Review* 45:821-841.
- Caplow, Theodore 1964 *Principles of Organization*. New York: Harcourt, Brace, Jovanovich.
- Clark, Burton 1983 *The Higher Education System*. Berkeley: University of California Press.
- Cole, Robert E. 1979 *Work, Mobility, and Participation*. Berkeley: University of California Press.
- Cox, Oliver C. 1948 *Caste, Class and Race: A Study in Social Dynamics*. New York: Monthly Review Press.
- Desivilya, Helena S. 1998 "Jewish-Arab Coexistence in Israel: The Role of Joint Professional Teams." *Journal of Peace Research* 35(4):429-453.
- Epstein, R., and S. S. Komorita 1965 "Parental Discipline, Stimulus Characteristics of Outgroups, and Social Distance in Children." *Journal of Personality and Social Psychology* 2:416-420.
- Evan, William 1966 "The Organization Set: Toward a Theory of Interorganizational Relations." In J. Thompson, ed., *Approaches to Organizational Design*. Pittsburgh, Pa.: University of Pittsburgh Press.
- Florida, Richard, and Martin Kenney 1991 "Transplanted Organizations: The Transfer of Japanese Industrial Organization to the U.S." *American Sociological Review* 56(3):381-399.
- Galaskiewicz, Joseph 1985 "Interorganizational Relations." *Annual Review of Sociology* 11:281-304.
- Hall, Richard H. 1987 *Organizations: Structures, Processes and Outcomes*, 4th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- , J. Clark, P. Giordano, P. Johnston, and M. Van Roekel 1977 "Patterns of Interorganizational Relationships." *Administrative Science Quarterly* 22:457-474.
- Hirsch, Paul M. 1972 "Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems." *American Journal of Sociology* 77:639-659.
- Jang, Ha-Yong 1997 "Cultural Differences in an Interorganizational Network: Shared Public Relations Firms Among Japanese and American Companies." *Public Relations Review* 23(1):327-342.
- Katz, Daniel, and Robert L. Kahn 1966 *The Social Psychology of Organizations*. New York: Wiley.



- Lawrence, Paul R., and J. W. Lorsch 1967 *Organization and Environment: Managing Differentiation and Integration*. Cambridge, Mass.: Graduate School of Business Administration, Harvard University.
- Lieberson, Stanley 1980 *A Piece of the Pie: Black and White Immigrants Since 1880*. Berkeley: University of California Press.
- Merton, Robert K. 1957 *Social Theory and Social Structure*. New York: Free Press.
- Mulford, C. L. 1984 *Interorganizational Relations: Implications for Community Development*. New York: Human Sciences Press.
- Park, Robert E., and E. W. Burgess 1921 *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.
- Pettigrew, Thomas F. 1998 "Reactions Toward the New Minorities of Western Europe." *Annual Review of Sociology* 24(1):77-103.
- Pfeffer, Jeffery, and G. R. Salaneik 1978 *The External Control of Organizations: A Resource Dependence Perspective*. New York: Harper and Row.
- Rogers, David L., and D. A. Whetten 1982 *Interorganizational Coordination: Theory, Research, Implementation*. Ames: Iowa State University Press.
- Scott, W. R. 1998 *Organizations: Rational, Natural, and Open Systems*, 4th ed. Upper Saddle River, N.J.: Prentice-Hall.
- Seeman, Melvin 1981 "Intergroup Relations." In M. Rosenberg and R.H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Simpson, George E., and J. M. Yinger 1972 *Racial and Cultural Minorities*, 4th ed. New York: Harper and Row.
- Stephan, Walter G. 1985 "Intergroup Relations." In Gardner Lindzey and Eliot Aronson, eds., *Handbook of Social Psychology*, vol. II. New York: Random House.
- Sumner, William Graham 1906 *Folkways*. Boston: Ginn.
- Van de Ven, Andrew H., and D. L. Ferry 1980 *Measuring and Assessing Organizations*. New York: Wiley.
- Vander Zanden, J.W. 1983 *American Minority Relations*, 4th ed. New York: Alfred A. Knopf.
- Williams, Robin M. 1977 *Mutual Accomodation: Ethnic Conflict and Cooperation*. Minneapolis: University of Minnesota Press.
- Yuchtman, Ephraim, and S. Seashore 1967 "A Systems Resource Approach to Organizational Effectiveness." *American Sociological Review* 32:891-903.

OSCAR GRUSKY  
JEFF ERGER

## INTERMARRIAGE

Intermarriage among people of different races, religions, nationalities, and ethnicities would be a subject of little concern in many societies (Degler 1971). It should be expected of a culturally diverse society such as the United States. Indeed, the United States is the most racially and culturally diverse nation in the Western, industrialized world. The heterogeneous composition of the United States should lend itself to a high degree of tolerance and acceptance of diversity in marriage patterns among its constituent groups (Spickard 1989). Since the 1960s, we have seen a rise in the number of intermarriages between different racial groups (U.S. Bureau of the Census 1998) and also an increasing number of interfaith marriages (Kalmijn 1993; Lehrer 1998). In order to discuss past, current, and future intermarriage trends, this article examines historical and contemporary trends in black/white intermarriages, past and future directions in Asian American intermarriages, the state of interfaith marriages, and reasons for the increasing number of intermarriages.

### HISTORY OF BLACK/WHITE INTERMARRIAGES

Slavery had its greatest impact on the interracial relations of the Africans brought to the United States. Most of the slaves who came in the beginning were males, with the number of black females not equal to the number of males until 1840. As a result, the number of sexual relations between black slaves and indentured white women was fairly high. Some of these interracial relationships were more than casual contacts and ended in marriage. The intermarriage rate between male slaves and free white women increased to the extent that laws against them were passed as a prohibitive measure. Before the alarm over the rate of intermarriages, male slaves were encouraged to marry white women, thereby increasing the property of the slavemaster, since the children from such unions were also slaves (Jordan 1968).

The end of slavery did not give the black woman any right to sexual integrity. What slavery began, racism and economic exploitation continued to impose on the sexual lives of black women. In the postbellum South, black women were still at

the mercy of the carnal desires of white men. According to historians, black women were forced to give up their bodies like animals to white men at random. Many have noted that many southern white men had their first sexual experience with black women. In some cases, the use of black women as sexual objects served to maintain the double standard of sexual conduct in the white South. Many white men did not have sexual relations with white women until they married. Some southern white men were known to joke that until they married, they did not know that white women were capable of sexual intercourse (Cash 1960; Dollard 1957).

It was the protection of the sexual purity of white women that partially justified the establishment of racially segregated institutions in the South. The southern white man assumed that black men had a strong desire for intermarriage and that white women would be open to proposals from black men if they were not guarded from even meeting with them on an equal level. As Bernard (1966) writes, "The white world's insistence on keeping Negro men walled up in the concentration camp was motivated in large part by its fear of black male sexuality" (p. 75).

The taboo on intermarriage was mostly centered on black men and white women. One reason for this is that white men and black women had engaged in coitus since the first black female slaves entered this country. Some black slave women were forced to engage in sexual relations with their white masters; others did so out of desire. Children resulting from these interracial sexual unions were always considered black, and the prevalent miscegenation of black women and white men has produced much lighter skinned American blacks than their African ancestors.

Traditionally, white fear of interracial relations has focused on the desire to avoid mongrelization of the races. Such a fear lacks any scientific basis, since many authorities on the subject of racial types seriously question that a pure race ever existed on this planet. Most authorities note it as an actual fact that the whole population of the world is hybrid and becoming increasingly so. At any rate, the rate of miscegenation in the past almost certainly casts doubt on any pure race theory for the United States (Day 1972).

Intermarriage is certainly nothing new in the United States. Its meaning and dynamics have, however, changed over the 400 years since blacks entered this country. In the era before slavery, black male and white female indentured servants often mated with each other. During the period of black bondage, most mixed sexual unions took place between white men and female slaves, often involving coercion by the white partner. A similar pattern of miscegenation occurred after slavery, with a white man and a black woman as the typical duo. When blacks moved to larger cities outside the South, the black male-white female pairing became more common. As is commonly known, legal unions between the races were prohibited by law in many states until 1967. But legal prohibitions were not the only deterrent to such biracial unions. This country's history is replete with acts of terror and intimidation of interracial couples who violated the society's taboos on miscegenation. While blacks and whites came together in love and marriage over the years, it was generally at a high cost, ranging from death to social ostracism (Stember 1976).

#### **CONTEMPORARY BLACK/WHITE MARRIAGE**

Between 1960 and 1990, black/white intermarriages increased fourfold (U.S. Census Bureau 1998). Among the reasons for this increase has been the desegregation of the public school system, the work force, and other social settings. Around 1968, American society witnessed the first significant increase in interracial dating. This was the year that blacks entered predominantly white college campuses in comparatively large numbers. Contemporaneous with this event was the sexual and psychological liberation of white women. While white society disapproved of all biracial dating, the strongest taboo was on the black male-white female bond. These bonds became the dominant figures in the increments of biracial dating. The college campus became an ideal laboratory for experiments in interracial affairs. In the university setting, the blacks and whites who dated were peers, with similar educational backgrounds, interests, and values. Young white women, who were not as racist as their parents, were liberated from parental and community control. Their student cohorts were more accepting of or indifferent to

their dating across racial lines. Those changes in interracial dating practices coincided with the civil rights movement and a greater white acceptance of blacks as racial equals.

In the 1970s through 1990s, integration of work settings, neighborhoods, schools, and other public settings has meant that blacks and whites interact much more as equals than in the past. According to a Gallup poll (1997), a majority of blacks go to school, live, and work in places where the population is at least half white or even predominantly white. Only 15 percent of blacks work with mostly or all blacks and 41 percent of blacks live in a mostly or all-black neighborhood. Therefore, it is not surprising that with increasing social interaction there would also be an increasing number of interracial unions.

In addition to these systemic changes, there has been a major change in public attitudes toward biracial couples. According to a Gallup poll (1997), a majority of blacks and whites under the age of 50 say they accept and approve of interracial unions between blacks and whites. Of all the different race-related trends, this change in attitude is the most significant. In 1958, only 4 percent of white Americans approved of marriages between blacks and whites; in 1997, 61 percent approved. However, negative attitudes toward black/white intermarriages still persist. About a quarter of blacks and about 40 percent of whites say they disapprove of interracial marriage. Much of this can be attributed to how different generations view interracial marriages, as well as regional differences. Younger people approve of such marriages, while older black and white Americans are less likely to approve. According to Wilson and Jacobson (1995), age and education are strong predictors of those who are accepting of black/white intermarriages. In their study, they found younger, educated cohorts to be more accepting of such unions, compared to older, less educated cohorts. Moreover, acceptance of such marriages is much lower in certain regions of the United States, such as the South (Gallup Organization 1997). Although these changes in attitudes toward intermarriage seem positive, this poll result could be misleading. Many people tend to give the liberal answer they think is proper or expected when asked about controversial issues such as interracial marriage. However, when confronted with the issue on a very personal level, their response may be much different.

In the past, white families in particular frequently refused to have anything to do with children who entered into interracial marriages (Golden 1954; Porterfield 1982; Spickard 1989). According to Rosenblatt and colleagues (1995), white families compared to black families were most often in opposition. Opposition by white family members was most often based in racist assumptions and stereotypes, but also based in concern about the racism that the couple would face from society at large. Moreover, white families were concerned that marrying interracially meant a poor economic future. Other concerns raised by white families included issues of safety and well-being, as well as the issue of raising a biracial child. The authors found that there was less opposition in black families. Close family members might have been militantly against the marriage, but mothers in black families played a key role in providing acceptance of the interracial marriage. On the other hand, in white families, fathers were the key person in providing acceptance. In both black and white families, families were particularly concerned about their daughters' marrying interracially. Rosenblatt and colleagues (1995) explain that "for white families, the roots may include the racism of a dominant group and fear of loss in status. For black families, the roots may include the fear and pain of being connected with the oppressor" (p. 69). St. Jean (1998) discusses how black males in families may be more ambivalent about such unions. In previous studies, black women more than black men had tended to disapprove of such intermarriages (Paset and Taylor 1991); however, St. Jean (1998) found quite the opposite. She found that black men had a more difficult time accepting such marriages.

Although many black/white couples are drawn together because of nonracial factors, such as common interests and personal attraction (Lewis et al. 1997), race remains a major factor in their interactions with family, friends, and society at large. Given the persistence of racism, many interracial marriages face rough going. Based on her qualitative research findings, St. Jean (1998) found that although for the "couples the salience of color seems to diminish after marriage, race consciousness does not diminish. In their everyday lives, they are reminded by Blacks, Whites, relatives and nonrelatives of the inappropriate nature of their association" (St. Jean, p. 12). It is a fact that the

scars of nearly 400 years of the worst human bondage known are not healed, and disapproval by many black and white people of interracial love affairs is one of the wounds.

#### PAST AND FUTURE DIRECTIONS IN ASIAN AMERICAN INTERMARRIAGES

Since 1990, fueled by immigration from Asia, the Asian and Pacific Islander population has grown at a rate of 4.5 percent per year (U.S. Bureau of the Census 1995). By the year 2000, this population had reached 12.1 million. As this group increases in size, intermarriage will probably occur more frequently (Lee and Fernandez 1998). According to Gordon (1964), the acculturation to American beliefs and values by new immigrants has meant that intermarriage would follow and is an important sign of the assimilation process. Hwang and colleagues (1997) report that Asian Americans with high levels of acculturation and structural assimilation have a high incidence of intermarriage: "Asian Americans who speak fluent English and who have lived longer in the US were found to have a higher tendency to marry persons from different racial and ethnic groups" (p. 770).

Despite the increase in population growth, the intermarriage rate for Asian Americans has declined overall. However, there are several explanations that account for these changes. According to Lee and Fernandez (1998), Asian-American outmarriage from 1980 to 1990 declined significantly from 25 percent to 15 percent based on their analysis of the 5 percent Public Use Microdata Sample (PUMS) of the 1990 U.S. census. They note that intermarriage levels dropped for Koreans, Filipinos, and Vietnamese. While outmarriage rates among American-born Asians increased, they declined for those foreign-born. They also note that Asian-American women were more likely to outmarry than men and that outmarriage was more common among Japanese Americans, Filipinos, Chinese, and Asian Indians. Lee and Fernandez (1998) further suggest that even though the Asian-American outmarriage rate overall has fallen, Asian interethnic marriages, made up of partners from two Asian ethnicities, has increased from 11 to 21 percent.

According to Shinagawa and Pang (1996), these marriage patterns can be explained through a

sociohistorical framework. They define five different historical cohorts: Pre-World War II and World War II (prior to 1946), post-World War II (1946-1962), the civil rights era (1963-1974), the post-1960s (1975-1981) and the Vincent Chin cohorts (1982-1990). Because the pre-World War II and World War II cohort experienced antimiscegenation laws, they intermarried with other nonwhites. Those in the post-World War II cohort lived in an era when antimiscegenation laws were struck down and Asian immigrants could now become citizens. At the same time, American soldiers fighting wars in Asia brought back Asian wives. The civil rights era cohort experienced radical changes in terms of race. Racial groups, no longer segregated from each other, were interacting more, and laws were enacted to bar racist discrimination. Asians could now intermarry with whites, and this era saw an increasing number of Asian women marrying white men. The post-1960s cohort experienced the beginning politicization of the Asian-American community. However, intermarriages were still mostly between whites and Asians. Major changes started to happen in the Vincent Chin cohort. In 1982, Vincent Chin, a Chinese American autoworker in Detroit, was beaten to death by two whites who mistakenly identified him as a Japanese American, angering the Asian-American community and fueling community mobilization and politicization. The murder, which became symbolic of anti-Asian hate crimes and discrimination, brought diverse Asian communities together to eradicate discrimination against Asians and bring justice to Vincent Chin's family (Espiritu 1992). This era was defined by an emerging increasingly coalesced Asian-American community as well as by a growth of the Asian Pacific American population. According to Shinagawa and Pang (1996), these dramatic developments contributed to the growth of interethnic Asian marriages.

Shinagawa and Pang (1996) hypothesize six major reasons why pan-ethnic intermarriages and pan-ethnicity among Asian Americans are on the ascent: (1) the growth of the Asian-American population; (2) the growth of personal and social networks due to these population increases; (3) the growth in the number of educated, middle-class Asian Americans; (4) the acculturation of Asian Americans; (5) shared racial identity; and (6) most importantly, the growing racial consciousness

among Asian Americans. According to Espiritu (1992), the construction of pan-Asian ethnicity arose out of a need for political strength and power. Shinagawa and Pang (1996) stress the importance of a pan-Asian ethnic consciousness that has shaped not only political but also marital patterns: "Deindustrialization, white flight, increased economic competition, anti-immigrant sentiments, hate violence against Asians, the growing sense of despair and hopelessness in the inner cities, and interracial conflicts not only between Whites and Asian Americans but also between racial minority groups all signify racial consciousness by Asian Americans" ( p. 144).

### TRENDS IN INTERFAITH INTERMARRIAGE

Although interracial marriages have increased, interfaith marriages are much more commonplace. Inter-marriage between white ethnics is quite the norm (Lieberson and Waters 1988). Marriages between members of different faiths also happens much more frequently and seems to carry less stigma than interracial coupling. In fact during 1990, 52 percent of Jews were married to non-Jews, while Protestant/Catholic intermarriages have increased significantly (Kalmijn 1991). According to Kalmijn (1991), the increasing importance of similar educational levels in spouse selection and the declining importance of religious differences explains the increase in Protestant/Catholic intermarriages. Using data from the 1987-1988 National Survey of Families and Households, Lehrer (1998) also suggests that intermarriages between Protestants and Catholics have increased steadily. She identifies key variables that play a role in the decision to intermarry. She reports that those with higher levels of education are more likely to intermarry than those with lower levels of education. She also discusses how a premarital pregnancy will increase the likelihood of an intermarriage. Those who are strongly committed to their faith are least likely to intermarry. Despite these increases, Lehrer and Chiswick (1993) report that interfaith marriages are more likely to end in divorce, attributing these high divorce rates to the different religions of the couple. However, they also suggest that conversion by one spouse to the faith of the other produces less conflict and a more harmonious relationship.

### FACTORS IN INTERMARRIAGES

Although the increasing trend toward intermarriage across ethnicity, religion, and race can be attributed to the increasing interaction between diverse individuals and to the elimination of institutional barriers, there are also other sociological, demographic, and psychological factors at work. According to psychotherapist Joel Crohn (1995), the decision to intermarry is not based on one single element—there may be many psychological influences operating. From his work as a therapist, he identifies four reasons that individuals from different religious, cultural, or ethnic backgrounds may be attracted to each other. First, stereotypes about a particular group may attract persons to each other; for example, black men are masculine, Jewish men are good providers, and Asian women are sexy.

Second, Crohn (1995) suggests that outmarriage can also be due to a partner's struggle with his or her identity. Those who outmarry may find members of the opposite sex from their ethnic, racial, or religious group unappealing and unattractive. Marriage can then be the arena in which individuals deal with their ambivalent attitudes toward their racial, ethnic, or religious identity: "Marriage to an outsider represents the ultimate strategy in trying to erase the stigma of a minority identity" (Crohn 1995, p. 52). At the same time, minority groups may outmarry into the dominant group to gain acceptance by that group. According to Pang (1997), Asian-American women who choose relationships with white men "participate in a language of assimilation that minimizes essential core parts of their self" (Pang 1997, p. 295). She found that Asian-American women who outmarry do not place importance on their race or culture and rather take on the identity of their white husband. According to Pang (1997), assimilation and incorporation into white society for these Asian-American women was of importance.

Third, Crohn (1995) suggests that whites may feel like they lack a particular cultural tradition and thus be attracted to a partner with a certain cultural and ethnic distinctiveness. Crohn further suggests that intermarriage is also a way for adult children to separate themselves from their family emotionally and/or physically. In fact, other studies have documented that those who chose to intermarry were more rebellious and had huge

conflicts with family due to objections to their marital choice (Sung 1990).

Fourth, Crohn discusses how attraction across religious, ethnic, or racial barriers may be grounded in cultural values and traditions, such as collectivism or individualism, that the individual is attracted to.

According to social exchange theory, relationships are exchanges of valued resources and involve an analysis of costs and benefits. Pierre Van den Berghe (1960) theorizes that racial or ethnic intermarriage is an instance of such an exchange, in this case that of hypergamy. *Hypergamous* means that women marry up in status; while *hypogamous* means that the racial and ethnic status of the husband is lower than that of his wife. In the past, most black men who married white women were of a higher social status than their wives. In fact, this marrying down was so common that sociologists formulated a theory about it, hypothesizing that the black groom was trading his class advantage for the racial caste advantage of the white bride (Davis 1941; Merton 1941). Kalmijn's (1993) study of marriage license data from 33 states between 1968 and 1996 indicates that black/white intermarriages have been on the rise, most prominently between black males and white women. Kalmijn notes that these marriages involved white women and high-status black men, meaning that white women moved up in socioeconomic status.

Other factors may propel people into an interracial marriage. Some students of the subject assert that uneven sex ratios are a basic cause. Whenever a group near another group has an imbalance in sex ratio, there is a greater likelihood of intermarriage. If the groups have a relatively well-balanced distribution of the sexes, members will marry more within their own group (Guttentag and Secord 1983; Parkman and Sawyer 1967).

As for the sociocultural factors that promote or deter interracial marriages, several explanations have been put forth to explain the variation in intermarriage patterns in the United States. Tucker and Mitchell-Kernan (1990) hypothesized that certain environments are more racially tolerant of intermarriage than are others. Their hypothesis is based on findings from U.S. census data showing that interracial marriage rates are highest

in the West and lowest in the South (U.S. Bureau of the Census 1985). Similar to their explanation is the argument by Blau and Schwartz (1984) that the larger the group size as a proportion of the population, the less likely it is that members will marry outside their group. Second, they suggested that the more heterogeneous an area's population, the more likely it is that people will marry outside their group. Both the aforementioned propositions imply that intermarriage is a function of environmental forces, not individual motivations.

In addition to the normal problems of working out a satisfactory marital relationship, interracial couples must cope with social ostracism and isolation. While the motivation for an interracial marriage may or may not differ from that of intraracial marriages, there are problems that are unique to interracial marriages. When researchers studied interracially married couples they discovered that courtship in most cases had been carried on clandestinely and, further, that many of them were isolated from their families following the marriage. Other outstanding social problems encountered by the couples centered on such factors as housing, occupation, and relationships with family and peers. Several of the spouses lost their jobs because of intermarriage, while others felt it necessary to conceal their marriages from their employers.

In addition to these strains, intermarried couples also face stressors within the family. Conflicts with in-laws contributed to marital instability (Chan and Smith 1995). Moreover, intermarried couples experienced not only the greatest conflict with in-laws but also differences that arose in terms of child rearing. According to Chan and Smith (1995), intermarried couples may face more problems because of their concerns with raising biracial children. They may worry about the children's psychosocial development because of their mixed heritage. The children may look more black than white and "this may create more stress for the mother who is most likely to be the primary caretaker and have to deal with the prejudice others might have about her children" (Chan and Smith 1995, p. 383).

Further research needs to be done to investigate the factors involved in intermarriages, as well to assess marital stability and instability among

intermarried couples. However, there are several data problems in conducting research on intermarriages. First, marriage data collected by the National Center for Health Statistics (NCHS) is incomplete, since not all states are required to submit data and only about forty states submit such information (Besharov and Sullivan 1996). Moreover, many states do not ask questions about race on marriage licenses. Currently, only thirty-two states report race information to the NCHS (Besharov and Sullivan 1996). Even if data on race and ethnicity are available, the analysis is quite limited when conducting research on the second-generation intermarriage experience (Roy and Hamilton 1997). Moreover, data on interfaith marriage are also sparse, since the census and other government surveys rarely ask a respondent's religion (Salins 1997). In addition, the General Social Survey, which collects information on attitudes toward intermarriages, is quite limited because only two questions are presented in this survey: "How do you feel about having a close relative or family member marry a Black person?" And "Do you think there should be laws against marriages between Blacks and Whites?" According to St. Jean (1998), these questions do not adequately address current attitudes on intermarriages. Rather, St. Jean (1998) suggests that qualitative research, such as focus groups, provide more insight and information than the existing quantitative research on the topic.

### SUMMARY

As barriers between different groups of people come down, intermarriage will continue to take place. This article has discussed the history and trends of black/white intermarriages, past and future directions of Asian-American intermarriages, trends in interfaith marriages, and factors involved in intermarriages. High intermarriage rates have been mostly among white ethnics (Salins 1997). Despite the increases in interracial marriage, black/white intermarriages are still quite low compared to other interracial and interfaith marriages. Racist attitudes still persist. Of these various types of intermarriage, black/white intermarriage is still one that is fraught with controversy. Those who oppose it often combine a hostility toward racial equality with invidious assessments of the private

thoughts and lives of interracial couples. Many men and women mate for no more complex reasons than meeting, liking each other as individuals, and choosing to transcend the societal barriers to their relationship. Only in societies similar to that of the United States does a biracial union take on any greater significance. For centuries, Latin American nations have undergone such a fusion of the races that only nationality, language, and religion remain as sources of identity. But the painful history of race relations in North America militates against the natural mixing of individuals from different races. Instead of regarding interracial dating and marriage as a matter of personal choice, many minorities have taken up the call for racial purity so common to their white supremacist adversaries of the past.

Despite the opposition to biracial unions, they will continue to increase as long as the social forces that set intermarriage in motion exist. There is, for example, the class factor. As long as middle-class blacks occupy token positions in the upper reaches of the job hierarchy, most of the people they meet in their occupational world will be white. Considering the fact that the job setting is the paramount place for meeting mates, it is only natural that many blacks will date and marry whites. Those whites will be the people they most often encounter and with whom they share common values, interests, and lifestyles. In the 1950s, E. Franklin Frazier (1957) predicted:

*The increasing mobility of both white and colored people will not only provide a first-hand knowledge of each for the other but will encourage a certain cosmopolitanism. That means there will be a growing number of marginal people who will break away from their cultural roots. These marginal people will help create not only an international community but an international society. In becoming free from their local attachments and provincial outlook, they will lose at the same time their racial prejudices, which were a product of their isolation. Many of these marginal people will form interracial marriages because they are more likely to find suitable marriage partners in the cosmopolitan circles than within their native countries.*

A careful reading of history indicates that the intermarriage rate often rises or falls for reasons

related more to political and economic factors than individual desire. Within the developed world, the United States has the lowest rate of interracial marriage because it is the most racially diverse of all those nations. Due to a combination of economic factors, many white Americans feel less secure about the racial privileges they have long enjoyed and the impact intermarriage would have on their life chances and that of their children. The increasing presence of people of color and their competition for the best jobs, houses, incomes, and lifestyles heighten these insecurities. Increasingly, they are willing to turn back those challenges by racial scapegoating in the political arena, resulting in a number of reactionary voting patterns by the Anglo majority on affirmative action and immigrant issues, which largely affect African Americans, Asian Americans, and Latino citizens. In defense of their right to participate equally in the political and economic life of the United States, people of color have forged movements of racial and ethnic solidarity to counter the white backlash.

However, racial divisions are doomed to eventual failure in large part because race itself is a social construct used more as an opiate or divisive strategy than for any other purpose. Just as religion is the divisive force in much of the world, racial differences serve to mask the underlying causes of class inequality in the United States. The political and economic elite have found it expedient to exploit minor differences in physical traits among groups to detract from public consciousness of the most pronounced case of economic inequality in the developed world. Only when economic justice becomes a social reality can we expect to see dating and marriage choices based on merit and not a group's standing in the racial and class hierarchy.

## REFERENCES

- Bernard, Jessie 1966 *Marriage and Family Among Negroes*. Englewood Cliffs, N.J.: Prentice-Hall.
- Besharov, Douglas, and Timothy Sullivan 1996 "One Flesh: America Is Experiencing an Unprecedented Increase in Black-White Intermarriage." *The New Democrat*, July/August, 8(4):20-22.
- Blau, Peter, and Joseph E. Schwartz 1984 *Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations*. Orlando, Fla.: Academic Press.
- Cash, Wilbur Joseph 1960 *The Mind of the South*. New York: Vintage.
- Chan, Anna Y., and Ken R. Smith 1995 "Perceptions of Marital Stability of Black-White Intermarriage." In Cardell K. Johnson, ed., *American Families: Issues in Race and Ethnicity*. New York: Garland Publishing.
- Crohn, Joel 1995 *Mixed Matches: How to Create Successful Interracial, Interethnic, and Interfaith Relationships*. New York: Ballantine Books.
- Davis, Kingsley 1941 "Intermarriage in Caste Societies." *American Anthropologist* 43:376-395.
- Day, Beth 1972 *Sexual Life Between Blacks and Whites: The Roots of Racism*. New York: World Publishing.
- Degler, Carl N. 1971 *Neither Black nor White*. New York: Macmillan.
- Dollard, John 1957 *Caste and Class in a Southern Town*. Garden City, N.Y.: Doubleday.
- Espiritu, Yen Le 1992 *Asian American Panethnicity: Bridging Institutions and Identities*. Philadelphia: Temple University Press.
- Frazier, E. Franklin 1957 *Race and Cultural Contacts in the Modern World*. New York: Alfred A. Knopf.
- Gallup Organization 1997 *Special Report: Black-White Relations in the US*. Princeton, N.J.: Gallup Organization.
- Golden, Joseph 1954 "Patterns of Negro-White Intermarriage." *American Sociological Review* 19:144-147.
- Gordon, Milton 1964 *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York, Oxford University Press.
- Guttentag, Marcia, and Paul F. Secord 1983 *Too Many Women? The Sex Ratio Question*. Beverly Hills, Calif.: Sage.
- Hare, Nathan, and Julia Hare 1984 *The Endangered Black Family*. San Francisco: Black Think Tank.
- Hwang, Sean-Shong, Rogelio Saenz, and Benigno E. Aguirre 1997 "Structural and Assimilationist Explanations of Asian American Intermarriage." *Journal of Marriage and the Family* 59(August):758-772.
- Jordan, Winthrop D. 1968 *White Over Black: American Attitudes Toward the Negro 1550-1812*. Chapel Hill: University of North Carolina Press.
- Kalmijn, Matthijs 1993 "Trends in Black/White Intermarriage." *Social Forces* 72(1):119-146.
- Kalmijn, Matthijs 1991 "Shifting Boundaries: Trends in Religious and Educational Homogamy." 56(6):786-800.
- Kosmin, Barry A., Sidney Goldstein, Joseph Waksberg, Nava Lere, Ariella Keysay, and Jeffrey Scheckner 1991 *Highlights of the CJF 1990 National Jewish Population Survey*. Washington D.C.: Council of Jewish Federations.



- Lee, Sharon, and Marilyn Fernandez 1998 "Trends in Asian American Racial/Ethnic Intermarriage: A Comparison of 1980 and 1990 Census." *Sociological Perspectives* 41(2):323-343.
- Lehrer, Evelyn L. 1998 "Religious Intermarriage in the United States: Determinants and Trends." *Social Science Research* 27(3):245-263.
- Lehrer, Evelyn L., and Carmel U. Chiswick 1993 "Religion as a Determinant of Marital Stability." *Demography* 30(3):385-404.
- Lewis, Richard, Jr., George Yancey, and Siri S. Bletzer 1997 "Racial and Nonracial Factors that Influence Spouse Choice in Black/White Marriages." *Journal of Black Studies* 28(1):60-78.
- Lieberson, Stanley, and Mary Waters 1988 *From Many Strands: Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage.
- Merton, Robert 1941 "Intermarriage and Social Structure: Fact and Theory." *Psychiatry* 4:361-374.
- Pang, Gin Yong 1997 *Asian American Intermarriage and the Language of Assimilation*. Unpublished doctoral dissertation, University of California, Berkeley.
- Parkman, Margaret A., and Jack Sawyer 1967 "Dimensions of Ethnic Intermarriage in Hawaii." *American Sociological Review* 32:593-608.
- Paset, Pamela, and Ronald D. Taylor 1991 "Black and White Women's Attitudes Toward Interracial Marriage." *Psychological Reports* 69(3):753-754.
- Porterfield, Ernest 1982 "Black-American Intermarriage in the United States." In Garv A. Cretser and Joseph J. Leon, eds., *Intermarriage in the United States*. New York: Haworth Press.
- Rosenblatt, Paul C., Terri A. Karis, and Richard D. Powell 1995 *Multiracial Couples: Black and White Voices*. Thousand Oaks, Calif.: Sage Publications.
- Roy, Parimal, and Ian Hamilton 1997 "Interethnic Marriage: Identifying the Second Generation in Australia." *International Migration Review* 31(1):128-143.
- Salins, Peter D. 1997 *Assimilation, American Style*. New York: Basic Books.
- Shinagawa, Larry Hajime, and Gin Yong Pang 1996 "Asian American Panethnicity and Intermarriage." *Amerasia Journal* 22(2):127-152.
- Spickard, Paul R. 1989 *Mixed Blood: Intermarriage and Ethnic Identity in Twentieth-Century America*. Madison: University of Wisconsin Press.
- St. Jean, Yanick 1998 "Let People Speak for Themselves: Interracial Unions and the General Social Survey." *Journal of Black Studies* 28(3):398-415.
- Stember, Charles Herbert 1976 *Sexual Racism*. New York: Elsevier.
- Sung, Betty L. 1990 "Chinese American Intermarriage." *Journal of Comparative Family Studies* 21(3):337-352.
- Tucker, M. Belinda, and Claudia Mitchell-Kernan 1990 "New Trends in Black American Interracial Marriage: The Social Structural Context." *Journal of Marriage and the Family* 52:209-218.
- U.S. Bureau of the Census 1985 *Census of the Population, 1980: Vol. 2. Marital Characteristics*. Washington, D.C.: U.S. Government Printing Office.
- 1998 *Intermarried Married Couples: 1960-present*. Washington, D.C.: U.S. Government Printing Office.
- Van den Berghe, Pierre 1960 "Hypogamy, Hypergenation and Miscegenation." *Human Resources* 13:83-89.
- Wilson, Deborah, and Cardell K. Johnson 1995 "White Attitudes Toward Black and White Interracial Marriage." In Cardell K. Johnson, ed., *American Families: Issues in Race and Ethnicity*. New York: Garland Publishing.

GRACE J. YOO  
ROBERT E. STAPLES

## INTERNAL MIGRATION

Migration is the relatively permanent movement of individuals or groups over varying distances to change places of residence; permanence and distance are its major defining dimensions. Internal migration occurs within the boundaries of a given country. (International migrants, not considered here, are called immigrants.) Internal migration, therefore, is a type of geographic mobility status.

### DEFINITIONS

The following definitions are standard in the field of social demography (Bogue 1985):

*Mobility status.* A classification of the population based on a comparison between the place of residence (destination) of each individual in a census enumeration or survey and the place of residence (origin) at some specified earlier date. Mobility status in terms of the distance of the move falls into four main categories: nonmovers, local movers, intrastate migrants, and interstate migrants.

They may be examined more specifically in the list below:

- I. *Nonmovers*, or nonmobile persons, live in the same house at the time of the census as at the date of origin.
- II. *Movers*, or mobile persons, live in a different house and are further classified as to where they were living at the earlier date.
  - a. *Local movers* are mobile persons who live in the same county at census time as at the date of origin.
  - b. *Internal migrants* are mobile persons who live in a different county at census time than at the date of origin. Internal migrants may be further subclassified:
    1. *Intrastate migrants* live in a different county but within the same state.
    2. *Interstate migrants* live in a different state.
    3. *Interregional migrants* live in a different geographic division or census geographic region; they are also interstate migrants.

*Mobility interval.* The lapsed time between the date specified for previous residence and the date of enumeration is usually either one year or five years. Recent census enumerations specify five years, and the Current Population Surveys have specified intervals of one, two, three, four, and five years.

*Metropolitan mobility.* A system of subdividing mobile persons into categories by place of residence at the beginning and the end of the mobility interval and, according to metropolitan statistical areas (MSAs), is as follows:

1. Within the same MSA
2. Between MSAs
3. From outside MSAs to MSAs
4. From MSAs to outside MSAs
5. Outside MSAs at both dates

*Mobility rates.* The number of persons in a specified mobility status per 100 or 1,000 in the

population of the area in which they resided at the end of the mobility interval is a mobility rate. Such rates may refer to any of the categories of nonmobile or mobile persons specified above. Mobility rates may be specific for age, race, sex, or other traits. The denominator may also be the origin date or the midpoint of the migration interval.

*Migration flows.* The key distinction of flows is that either the origin or the destination is unknown. There are two types of flows:

1. *In-migration* is comprised of migrants arriving at a particular place of destination, with no reference to the place of origin. In-flows could also arrive at specified types of places, such as central cities or metropolitan areas.
2. *Out-migration* is comprised of migrants departing from a particular area, with no reference to the place of destination. Out-flows may also depart from specified types of places, such as places outside MSAs or suburban metropolitan rings of MSAs.

*Migrations streams.* These connect an origin to a destination. There are three types of migration streams:

1. *Specific streams.* Streams that connect particular places within a category, such as streams between specific cities, counties, states, or regions. This is the major use of the term.
2. *Typological streams.* Streams that connect types of places, such as streams between all central cities and suburbs in a state or the nation.
3. *Counterstreams.* When a stream between two places endures, it usually generates a counterstream, a smaller stream in the opposite direction. The stream and counterstream are referred to as an *exchange*.

*Net migration.* This is the difference obtained when the number of out-migrants is subtracted from the number of in-migrants in a particular place or type of place. A location that experiences a loss of population through migration is said to have a negative net migration; one that gains

population through migration has a positive net migration. Because of its birth and death rates, an area may have a negative net migration and continue to have a growing population. There is no such thing as a net migrant, however.

*Return migration.* The census contains an item that identifies the state of birth. Return migrants are those persons who return to their state of birth during the mobility interval. There is no way of knowing how long they have been away from their state of birth when they return.

### WHY STUDY MIGRATION?

Migration is important to social scientists because an increase or decrease in the size of a population, due to excess in- or out-migration, causes many social conditions to change. Community infrastructure, such as highways and schools, may become overburdened due to population growth, while public services may become difficult to maintain when population declines. Furthermore, social scientists study the equilibrating effects of population movement on national and regional economic systems. Growth or decline in the local economy is an incentive for people to move, which redistributes the population to balance the system.

The ability to predict the impacts of population growth or decline on the institutional sectors of a community and the ability to understand regional population dynamics, of course, provide many practical benefits to government and business planners.

### MIGRATION RESEARCH

Net migration rates before 1940 were estimated using a survival-rate method. This method takes the population in one census as its base. It adjusts the number by adding births and subtracting deaths during the next decade. The amount of population change not accounted for is attributed to migration (Bogue and Beale 1961). The 1940 census was the first to include a mobility item. It asked where persons lived five years earlier. In 1950, after World War II, there was so much population movement that a one-year interval was substituted in the census. In 1960 the five-year mobility interval was restored and has been retained in subsequent decades. Because of these measurement

changes, the 1960 and 1970 censuses were the first from which decade changes could be derived. Thus several landmark studies appeared in the 1960s, breaking new ground and setting patterns for future migration research (Long 1988). Shryock's (1964) work showed the importance of studying gross migration flows in addition to the prevailing dependence on net migration. Lowry (1966) introduced econometric modeling to migration research. Finally, Lansing and Mueller (1967) helped introduce survey approaches to analyzing internal migration.

### U.S. MOBILITY

Americans are unusually mobile (Bogue 1985). Only Canada and Australia have populations as mobile as that of the United States. In a single year, from March 1995 to March 1996, 17 percent of U.S. inhabitants moved from one domicile to another and about 6 percent changed their county of residence. At current mobility rates, average Americans live at fourteen different addresses during their lifetimes. Of these thirteen moves, three are as a dependent moving with parents and ten are of their own volition. People who have lived their entire lives at the same address account for no more than 2 or 3 percent of the adult population. Perhaps no more than 10 to 15 percent of people spend their entire lives in their county of birth.

When using the five-year mobility interval, mobility rates are not five times as large as those for a single year because persons who move several times within the interval are counted only once. Nearly one-half of the population is mobile over a five-year period, and more than one-fifth are migrants. Since 1980 there appears to have been no diminution in the tendency to migrate, but there has been an apparent reduction in local mobility.

One can discover contradictory findings in mobility literature. These contradictions are often due to the specific databases under analysis. Some databases use mortgage data and leave out renters; others, such as the Annual Housing Survey, use households; and some, such as most census publications, use individuals as units of analysis, each database giving somewhat different results. In addition, some data sources offer little information on the characteristics of migrants. The individual

master file of the Internal Revenue Service includes state and county migration data but no personal characteristics, and several large moving companies provide data on their customers also without personal characteristics (Kahley 1990).

**Reasons for Migration.** Migration may occur in response to changing economic, social, or political conditions. Push factors are conditions in the sending population that impel or stimulate migration. Conditions that attract in-migrants are classified as pull factors (Ravenstein 1889).

Declining economic opportunities, political instability, or the weakening of place ties may stimulate out-migration. Expanding economic opportunities, potential for advancement, the presence of family members and friends, or previous vacationing or residential experience tend to attract migrants. Not surprisingly, rural communities with high birthrates and regions with limited opportunities are areas of high out-migration, whereas urban, industrial regions and communities with expanding opportunities tend to have high in-migration (Prehn 1986). Marriage, divorce, increasing or decreasing family size, and housing adequacy top the list in surveys. A sizable majority of respondents to the Annual Housing Survey reported housing or family dynamics as reasons for their move (Gober 1993).

The average age at which young adults leave home declined from the low twenties to the upper teens between 1920 and 1980, and then the median age began to rise again. These trends mirror another one; for the Vietnam cohort of young adults forward, those who return to live at home at some time holds at about 40 percent. About 25 percent had moved back in earlier cohorts. The expectation of a permanently empty nest for parents of young adults now seems a less certain one (Goldscheider and Goldscheider 1994).

Zelinski (1971) proposed a macro-level three-stage model of national internal migration. First, with the onset of modernization, the overall level of migration increases, primarily in the form of rural-to-urban moves. Second, as industrialization and modernization spread to more regions, migration may continue to increase; improved transportation and communication increase the availability of information and decrease the uncertainty of

moving. Interurban moves become the majority of all moves. Finally, at advanced stages, when level-of-living differences among areas have diminished, there may be more urban-to-rural movement and more “consumer-oriented” migration toward warm climates or locations with other amenities (Long 1988).

**Differential Migration.** What population characteristics predict migration? Characteristics that indicate less entanglement with social obligations, greater need for employment, and higher job skills are good predictors. Men are more mobile with respect to residence than women, although the difference is small. The single migrate at higher rates than the married. For several decades, blacks have been more mobile than whites. However, in 1980 whites migrated at higher rates than blacks, although blacks continued to be more mobile locally. Hispanics migrated internally at a rate between those of the black and white populations. Persons with higher levels of education are more likely to migrate than those who are less well educated.

**Age and Mobility.** The shape of the age profile of migrants in the United States has been consistent for decades, changing only gradually over time. The younger children are, the more likely they are to migrate. The migration rate of children bottoms out in the early teens and does not increase rapidly until the late teens. More than one-third of Americans in their young adult years, ages twenty to twenty-four, the peak migration years during the life course, moved at least once between 1982 and 1983, and nearly one-half of this mobility was migratory. Not surprisingly, this age corresponds with college graduation and marriage for many. The increasing age of children in the home, particularly once they begin their formal schooling, dampens the attractiveness of migration for parents. The age-specific migration rate declines slowly at first, then more steeply until age thirty-five, after which it slowly declines throughout the middle years to a life-course low point just before the retirement years. The retirement migration hump between ages sixty and seventy is small by comparison to the early adulthood migration bulge. The final increase in age-specific migration rises at the end of life and relates largely to health issues. The elderly as a broad category are

only about one-half as mobile as the general population.

### MIGRATION AND REGIONAL DISTRIBUTION OF POPULATION

Three large interregional flows of internal migration have been occurring in the United States for many decades.

**Westward Movement.** For a long time, there was a high-volume flow of persons into the Pacific region, principally California, as well as a high-volume flow into the mountainous southwestern states. The 1970–1980 decade had a higher volume of westward movement than any previous one. Mountain states that previously suffered losses all made positive gains, and Colorado, Nevada, and Arizona continued the large gains of the previous decade. In the 1990s, there was a net flow out of California, largely to other western states, reversing a long-term trend for that state.

**Northward Movement from the South.** The southern region lost population heavily between the close of the Civil War and 1950. Industrial centers in the northeast and east-north-central regions absorbed a very large share of the migrating population. Both white and black migrants flowed along these channels in great numbers. Some southern states, however, particularly Florida and Texas, were exceptions. Between 1970 and 1980 the net outflow from the South completely disappeared. Those who left the South preferred the West to the North as a destination, and in-migrants to the South balanced the out-migrants. Every state in the northeast and north-central regions suffered a net migration loss during the decade, resulting in a major regional migration turnaround (Bogue 1985). By 1990 there were no net flows out of the South to other regions, but the Northeast, Midwest, and West all contributed to the southern region (Gober 1993).

**The Southward Movement to the Gulf Coast and the Southern Atlantic Seaboard.** The entire Gulf Coast, from the mouth of the Rio Grande in Texas across the coastal portions of lower Louisiana, Mississippi, and Alabama and on to include all of Florida, experienced much more rapid and intensive economic development than the southern and southeastern parts of the United States

lying away from the coast. Although this trend is a very old one, it accelerated rapidly in the 1970s.

As of 1980 there were only two regional migration streams instead of three: movement toward the South and Southwest and movement toward the West. The northeast and the north-central regions are the sources from which these migrants came (Bogue 1985). But in the 1980s, the South gained more through net migration than did the western states (Weeks 1996), a trend that accelerated by 1990. The geographic redistribution of knowledge-based industries of the information age carries in its wake a college-educated work force to the Sunbelt, including the South (Frey 1995).

**Metropolitan Deconcentration.** One of the macro-level processes that affects geographic mobility in our time is metropolitan deconcentration. Many nonmetropolitan counties in the United States experienced a slowing of population decline in the 1960s, and in the 1970s their net migration rates climbed above the break-even point, which signaled a genuine and widespread “rural–urban turnaround.” Older people seem to have been in the vanguard of migration to nonmetropolitan counties; the turnaround for them happened in the 1960s rather than the 1970s. This reversal of a long-term trend of rural-to-urban migration is of great interest to demographers. Mounting evidence now indicates that although deconcentration continues in nonmetropolitan America as a whole, by the late 1980s metropolitan counties were outgrowing nonmetropolitan ones (Long and DeAre 1988). In the 1990s there is an uneven urban revival, with a few metropolitan areas with more flexible and diverse economies, mostly outside the Northeast and Midwest, gaining migrants. The new dominance of the suburbs over the central city is key to metropolitan deconcentration in the 1980s and 1990s. During this period, the suburbs are capturing the bulk of employment and occupational growth (Frey 1995).

### RETIREMENT MIGRATION

Demography traditionally tends to focus on youthful migration, and labor-force migration in particular. Increasing attention, however, is being given to non-labor-force-motivated migration, particularly to the migration of persons of retirement

age (Longino 1996). For the elderly, interstate flows are highly channelized—that is, half of the interstate migrants, regardless of their origin, flow into only eight of the fifty states. Florida dominates the scene, having received about one-quarter of all interstate migrants aged sixty or over in the five years preceding the 1960, 1970, 1980, and 1990 censuses. Although Florida, California, Arizona, and North Carolina have different major recruitment areas, they are the only states that attract several unusually large streams from outside their regions. Florida and North Carolina draw primarily from east of the Mississippi River, and Arizona and California draw from west of it. Among the elderly, the special characteristics of the destination tend to be more important than the distance. Warm climate, economic growth, and lower cost of living are still important pull factors.

Distance selectivity of elderly migration has been studied. Local movers are generally not as economically and socially well-off as nonmovers, and migrants are more well-off. Interstate migrants tend to have the most positive characteristics.

Permanence is an important but difficult dimension of migration to study. The census assumes that one's "usual place of residence" is not temporary. In reality, however, much of the migration among older people may be temporary. So far, studies of elderly seasonal migrants show them to be relatively advantaged, attracted by non-labor-force issues such as climate, cost of living, and the locations of family members and friends.

Metropolitan-to-metropolitan migration predominates among the elderly. Of the one-third who changed environmental types, no increase occurred between the 1960 and 1980 censuses in the proportion moving out of metropolitan areas in each decade. However, the movement in the opposite direction, up the metropolitan hierarchy, declined, both among older intrastate and interstate migrants. The net difference made it appear as though the flow from cities increased. Metropolitan-to-metropolitan migrants, especially those moving longer distances, tend to have more income, to be married, and to live in their own homes. A higher proportion of nonmetro-to-metro migrants is older, widowed, and living dependently, especially with their children. Coding revisions in

the 1990 census made updates of these comparisons impossible to calculate.

The cycle of migration for a job when one is young and returning to one's roots after retirement is an appealing notion to theorists. In contrast, Rogers (1990) demonstrated that elderly persons are not any more likely to return home than are the nonelderly; in fact, the probabilities of return migration by the elderly are lower than those of the general population, even after controlling for the different mobility levels of the two populations. There is wide state variability, however. The southeastern region is unusually attractive to older return migrants, and return migration is uncommonly high among the older black population moving into that region. Some evidence from the 1990 census shows that regional return migration patterns are shifting away from the Sunbelt states. Some migrants, apparently, return to their home states after an earlier retirement move (Longino 1995).

Some dubbed retirement migration as the growth industry of the 1990s. The amount of income transferred between states through retirement migration is quite substantial. Not surprisingly, economic development agencies are mounting efforts to attract mature migrants. This is leading to sharp competition among destinations for these migrants as new residents. The impact of elderly migration as a social phenomenon has yet to generate enough research to provide definite statements.

#### INTERNATIONAL COMPARISONS OF INTERNAL MIGRATION

Little research exists to compare countries on internal migration because measures, data sources, and units of analysis differ widely among nations. Consequently, international organizations have not published compendia of national comparative data on migration as they have on fertility and mortality. In addition, certain types of cultures conceive internal migration differently. In some small countries, such as England, lack of new housing stock limits residential movement. Migration is also limited in countries such as France, where transportation routes primarily connect the peripheral towns to a central national capital for historical reasons. Conversely, internal migration is amplified and culturally expected in nations of

immigrants with widely dispersed regional centers and major cities, such as the United States, Canada, and Australia.

Nonetheless, existing studies provide some tentative generalizations that compare internal migration in the United States with that in other countries (Long 1988). The U.S. national average for moves is higher than that of most other countries because (1) cities in the South and West are growing; (2) a relatively large minority of people who repeatedly move elevates the U.S. average for lifetime moves above that for most other countries; and (3) during the 1980s and 1990s the baby boom generation in the United States has moved through the life-cycle stages that have the highest rates of geographic mobility.

Comparative studies also give attention to older migrants, although their mobility rates are lower than for the young. Rogers (1989) argues that as the populations of industrialized nations age, the internal migration patterns of elderly persons will change. Elderly migration levels are low in countries in the first stage of this population transition. In the second stage of the transition, large, long-distance flows to particular principal destination regions appear. The third stage continues to exhibit large numbers of elderly migrants, but their moves now include a significant number of short-distance moves to more dispersed inland regions. Rogers and colleagues (1990) argue from comparative data that England is in the third stage, the United States is transitioning between the second and the third stages, Italy is well into the second stage, and Japan is in the first stage.

Since 1970, for most developed countries, population aging has brought declining national rates of internal migration (Long 1988). For the United States, the decline appears to be greater for local moves than for long-distance movement. Urbanization was the dominant redistribution trend in the 1950s in fourteen European countries studied by Fielding (1989). The relationship between net migration and settlement size began to break down, however, in the 1960s—first in the countries of northwestern Europe in the mid-1960s, then in the countries and regions of the southern and western European periphery through the 1960s, and in the case of Spain, into the 1970s. By the 1970s, most of the countries of western Europe were recording counterurbanization, where

the net flow was away from cities and toward small settlements. That counterurbanization became less dominant in the early 1980s but was not replaced by urbanization. Only in West Germany and Italy did the counterurbanization relationship persist. The United States experienced a similar pattern of long-term urbanization, reversed in the 1970s and then nearly reversed again in the 1980s (Frey 1990).

### PREDICTING FUTURE MIGRATION

Migration rooted in labor movement will change in the future as the geographical basis of the economy changes. Robust new industry will attract migrants. Such developments in the southern region may extend several more decades into the future. On the other hand, migration not related to the labor force, such as retirement migration, is more sensitive to lifestyle issues. Eventual overcrowding, which results in a decline in the quality of life of local residents, will tend to discourage retirement migration. Florida's dominance of retirement destinations in 1990 lost 2 percent of the migrant retiree market.

Migration for better jobs increases in times of economic expansion. Thus, studies of the late 1990s may find migration to have increased in response to an improved economy. Other trends could also increase migration rates. First, the age composition is always shifting. In the 1980s, there were more persons in the twenty-to-thirty age range, their prime mobility years. The baby boom generation has a lower migration rate for long moves than others. However, because of its large size, a large number of baby boomers migrated. In the 1990s, the incidence of moving will likely dampen as baby boomers age and leave their prime mobility years. Second, the rising level of education may increase migration. Each new cohort of adults has a higher level of education than its predecessor. The third factor, household change, contains counterindicators. Married couples are increasingly likely to divorce, a situation favoring migration, but at the same time there are more dual-career couples in the population, a situation favoring nonmobility (Long 1988).

As we have seen, many factors motivate migration. These factors need further study, which will certainly generate new research hypotheses to be tested by migration researchers in the twenty-first century.

(SEE ALSO: *Population; Retirement*)

## REFERENCES

- Bogue, Donald J. 1985 *The Population of the United States: Historical Trends and Future Projections*. New York: Free Press.
- , and Calvin L. Beale 1961 *Economic Areas of the United States*. New York: Free Press.
- Fielding, A.J. 1989 "Migration and Urbanization in Western Europe Since 1950." *The Geographical Journal* 155:60-69.
- Frey, William H. 1990 "Metropolitan America: Beyond the Transition." *Population Bulletin* 45:1-51.
- 1995 "The New Geography of Population Shifts." In R. Farley, ed., *State of the Union: America in the 1990s: Vol. 2. Social Trends*. New York: Russell Sage Foundation.
- Gober, Patricia 1993 "Americans on the Move." *Population Bulletin* 48:2-39.
- Goldscheider, Frances, and Calvin Goldscheider 1994 "Leaving and Returning Home in Twentieth Century America." *Population Bulletin* 48:2-35.
- Kahley, William J. 1990 "Measuring Interstate Migration." *Economic Review* 75(2):26-40.
- Lansing, John B., and Eva Mueller 1967 *The Geographic Mobility of Labor*. Ann Arbor: Institute for Social Research, University of Michigan.
- Long, Larry H. 1988 *Migration and Residential Mobility in the United States*. New York: Russell Sage Foundation.
- , and D. DeAre 1988 "U.S. Population Redistribution: A Perspective on the Nonmetropolitan Turnaround." *Population and Development Review* 14:433-450.
- Longino, Charles F., Jr. 1995 *Retirement Migration in America*. Houston: Vacation Publications.
- 1996 "Migration." In James E. Birren, ed., *Encyclopedia of Gerontology: Age, Aging, and the Aged*, vol. 2. San Diego, Calif.: Academic Press.
- Lowry, Ira S. 1966 *Migration and Metropolitan Growth: Two Analytic Models*. San Francisco: Chandler.
- Prehn, John W. 1986 "Migration." In *The Encyclopedia Dictionary of Sociology*, 3rd ed. Guilford, Conn.: Dushkin.
- Ravenstein, E.G. 1889 "The Laws of Migration." *Journal of the Royal Statistical Society* 52:245-301.
- Rogers, Andrei 1989 "The Elderly Mobility Transition: Growth Concentration and Tempo." *Research on Aging* 11:3-32.
- 1990 "Return Migration to Region of Birth Among Retirement-Age Persons in the United States." *Journal of Gerontology: Social Sciences* 45:S128-S134.
- , John F. Watkins, and Jennifer A. Woodward 1990 "Interregional Elderly Migration and Population Redistribution in Four Industrialized Countries: A Comparative Analysis." *Research on Aging* 12:251-293.
- Shryock, Henry S. 1964 *Population Mobility Within the United States*. Chicago: Community and Family Study Center, University of Chicago.
- Weeks, John R. 1996 *Population: An Introduction to Concepts and Issues*, 6th ed. Belmont, Calif.: Wadsworth.
- Zelinski, Wilbur 1971 "The Hypothesis of the Mobility Transition." *Geographical Review* 61:219-249.

CHARLES F. LONGINO, JR.

## INTERNATIONAL ASSOCIATIONS IN SOCIOLOGY

Founded by Rene Worms in Paris in 1893, the International Institute of Sociology (IIS) is the oldest continuous sociological association of any kind in the world. The IIS also happens to be the oldest continuous international association in the social sciences. Indeed, Worms's models for the IIS were the recently formed international institutes in law and in statistics. In the same year that he founded the IIS, Worms (1869-1926) also started the *Revue internationale de sociologies* (two years before Durkheimians would found the *Annee sociologique*) and a book series, the *Bibliothèque Sociologique Internationale*, which would in time publish more than fifty volumes.

In 1893 Worms was twenty-four years old. His organizational skills are even more impressive when we consider where the IIS's founding falls in relation to that of other major associations in the discipline of sociology, including the discipline's only other major international association. In 1895, Worms himself founded the *Société de Sociologie de Paris*, an association that held monthly meetings. A full ten years later, in 1905, C. W. A. Veditz of George Washington University founded the American Sociological Society (now the American Sociological Association) at a constituent meeting



held at Johns Hopkins University. Another five years later, in 1910, the German Sociological Society held its first meeting—in Frankfurt, with Simmel presenting the opening paper. Three years after that, in 1913, Tongo Takebe founded the Japan Institute of Sociology (now the Japan Sociological Society).

Then some time passed before other major associations were formed. The Italian Society of Sociology was not founded until 1937, by Corrado Gini (who in 1926 established Italy's equivalent of the Census Bureau and who developed the Gini Scale). The International Sociological Association, which is discussed below began after World War II, in 1948—under broad sponsorship from UNESCO and informed by Cold War rivalries. Finally, the British Sociological Society was founded even later, in 1951; indeed, Britain had only one chair in sociology from 1907 through World War II, housed at the London School of Economics.

Like so many of sociology's founding figures, from Auguste Comte to George Herbert Mead, Worms was trained in philosophy and began his academic career in that discipline. The son of political economist Emile Worms, he earned his philosophy degree at the Ecole Normale Supérieure and then also degrees in law and economics. His first teaching assignment was at the secondary school level while he also substituted for Henri Bergson at the Collège de France. And like other philosophically trained sociologists of the mid- and late nineteenth century, including Auguste Comte and Herbert Spencer, Worms's early way of thinking about society was by analogy to biological organisms. Within a decade, however, Worms moved, under increasing criticism from Gabriel Tarde (and also Durkheim), to a far more eclectic theoretical position, one that treated sociology essentially as the philosophy of the social sciences. After 1907 he taught an annual course in the history of sociology at the Ecole des Hautes Etudes Sociales. During this period he also taught on the law faculty at the University of Paris and served on important advisory bodies of the French government. By 1924 he had risen to *Conseiller d'Etat*.

The first World Congress of Worms's IIS was held a year after its founding, in September 1894, and the first volume of the Institute's *Annales* appeared a year after that. Not only scholars but

also politicians and other influentials throughout the West found both the *Revue* and the IIS to be tolerant and supportive of heterogeneous ideas. This hospitable environment reflected Worms's own broad work experience and interdisciplinary training—along with his own general openness to differing theoretical approaches. Worms respected the work being done, for instance, by the circle around Durkheim, but he also kept a healthy independence from its influence. Thus, we find among the IIS's earliest members major figures not only in sociology but also in economics, the other social sciences, and the natural sciences: Franz Brentano, Enrico Ferri, Ludwig Gumplowicz, Achille Loria, Alfred Marshall, Carl Menger, Edward Ross, Gustav Schmoller, Georg Simmel, Albion Small, Gabriel Tarde, Edward Taylor, Ferdinand Tönnies, Alexandre Tchouprov, Thorsten Veblen, Lester Ward, Sydney and Beatrice Webbs, and Wilhelm Wundt. Indeed, in the early years sociologists were a minority among IIS members. Yet, with the exception of Durkheim, all major sociologists, Weber included, participated in the Institute at one time or another in its first three decades.

In casting his net well beyond the academy, Worms was known to play on the vanity of the famous from many walks of life in order to bring them into his international organization (D'Antonio 1994, citing Gephart). Often he granted influentials an honorary membership or officership in order to exempt them from paying dues. The first president of the IIS was Sir John Lubbock, vice president of the Royal Society of London, president of the London Chamber of Commerce, and a member of the House of Commons. A year later, Albert Schaeffle became the second IIS president, reflecting the fact that until 1914 IIS Congresses tended to be held annually. Aside from exchanging ideas frequently with Tönnies, Durkheim, and Simmel, Schaeffle served as director of the *Zeitschrift für gesammte Staatswissenschaft* of Tübingen. The IIS's third president was a senator from Russia, Paul de Liliensfeld. One founding member, T.G. Masaryk of Prague, became both president of Czechoslovakia and president of the IIS. IIS vice presidents were no less successful in other pursuits: Bernardino Machado, Raymond Poincaré, and Woodrow Wilson were past presidents, respectively, of Portugal, France, and the United States.

The first sociologist to serve as IIS president was Lester Ward, in 1903; two years later, Ward became the first president of the American Sociological Society. Indeed, Ward, Franklin Giddings, and Albion Small were all active in the IIS and would be prime movers in the formation of the American Sociological Society. All would serve as IIS presidents and, together, they comprised three of the American organization's ten original officers.

Worms himself was always more a technical director than a research sociologist, serving as IIS secretary general and treasurer for thirty-three years. During his tenure, eleven different countries were represented in the presidency of the IIS (and in 1924 Worms was made an honorary member of the American Sociological Society). The first French sociologist to serve as IIS president was Charles Letourneau.

In the interwar years, IIS Congresses were held biannually and IIS leadership was strongly influenced by Célestin, Bougle, Secondo Bouthoul, Corrado Gini, Maurice Halbwachs, Robert MacIver, William Ogburn, Joseph Schumpeter, Pitrim Sorokin, and Luigi Sturzo. Each World War disrupted the frequency of Congresses and, as it turned out, each of these periods also marked general passages for the institute as an organization. For instance, a Congress scheduled for Vienna in 1915 was canceled because of war, and then planning did not resume until 1925. In that year Worms became ill, and he died the following year: the Congress ended up being held in 1927. A decade passed. Then the 14th Congress, scheduled for Bucharest in 1939, had to be canceled—again due to war. It would not be held until 1950, now relocated to Rome and placed under the leadership of Corrado Gini. Earlier, in 1932, Gini had founded the Italian section of the IIS.

At the turn of the century, international associations in all scholarly disciplines typically called themselves “institutes”; they were more “traveling academies, as Scheuch (1997) put it, than bureaucratically established organizations. All drew membership by invitation, all placed numerical limits on total membership, and all organized plenary sessions around particular themes. The IIS continues this tradition to this day. Thus, the format for its biannual Congresses is quite distinct from that of today's meetings of the International Sociologi-

cal Association (ISA) held every four years. In the first place, the IIS tries to keep attendance to around 500 total participants (whereas ISA meetings draw anywhere from 3,000 to 5,000 participants). In the second place, all participants are encouraged to attend morning plenary sessions organized around a particular conceptual theme selected by each Congress's program committee. The afternoons are then devoted to working groups in which members present papers spanning a far broader range of topics.

This “traditional” format is designed to serve both interpersonal and intellectual ends: Interpersonally, IIS Congresses stage events that are designed to create a sense of community among scholars from different national traditions of training and thereby to encourage collaboration and other scholarly exchanges. This is the case not only with the collective morning plenary sessions in which presentations are simultaneously translated in three languages—the language of the host country, French, and English. It is also the case with collective lunches, dinners, receptions, and local sightseeing trips. Intellectually, themes of IIS Congresses are intended to have a cumulative impact on social science disciplines.

The *Annales* served as the IIS's official publication through 1931 (and the 10th Congress held in Switzerland in 1930), at which time it was discontinued. From this period until 1989, IIS activities were reported in the *Revue*, which also informally took over publication of IIS Congress papers. In 1989 the *Annales* were revived and the *Revue* became independent of the IIS. Recent World Congresses were held in Morelia, Mexico (1982), Albufeira, Portugal (1986), Rome (1989), Kobe (1991), Paris (1993), Trieste (1995), Cologne (1997), and Tel Aviv (1999).

The International Sociological Association was founded in 1948 under an initiative of the Social Science Department of UNESCO—as part of a broad effort that spanned comparable associations in economics, law, and political science. All of these new international bodies very much reflected the geopolitical situation of the day, and all were both considerably larger in membership and looser in thematic focus than the IIS. Today, the major scholarly contribution of the ISA revolves not around its plenary sessions but rather around

nearly fifty specialized Research Committees, some of which rival in size an entire IIS Congress.

After WWII, the Allies were interested in undercutting the conditions that could foster any return to fascism, and they saw the dissemination of the social sciences as one important factor in fostering democracy. Indeed, UNESCO had already opened an Institute for Social Research in Cologne, the Rockefeller Foundation had financed another in Dortmund, and the American government still another in Darmstadt. Given that UNESCO was based in Paris, the first discussions leading to the ISA were held in that city beginning in October 1948 under the leadership of Arvid Brodersen. Also attending this first meeting were Georges Davy, Gurvitch, Gabriel Le Bras, Arie den Hollander, Rene Konig, Louis Wirth, Paul Lazarsfeld, Erik Rinde, and Otto Klineberg.

At the time, national sociological associations existed in eight countries (Belgium, Brazil, China, Germany, Italy, Japan, Netherlands, and the United States). In four other countries sociology was included in other social science associations, and in another ten the discipline was organized as institutes on the IIS model—with membership limited to selected individuals (Platt 1998). In September 1949, ISA organizers invited leading sociologists in twenty-one countries to a Constituent Congress in Oslo. Louis Wirth, the noted American sociologist of urban life, was named the first president, and Erik Rinde in Oslo was appointed Executive Secretary and Treasurer (Platt 1998). Unlike the IIS, membership in the ISA was by national association (including general social science association), not by individuals directly. As such, the rise of the ISA itself contributed to the founding of eleven national affiliates, including the British Sociological Association and its counterpart in Mexico (Platt 1998).

The ISA's first World Congress was held in Zurich in 1950, organized by Rene Konig with Wirth serving as president. One hundred fifty-four individuals participated, and eleven national associations and eighteen other bodies were admitted as members. In part, this Congress was held jointly with the newly formed International Political Science Association (and the Research Committee on Political Sociology remains a joint committee of both associations). Shortly thereafter, the ISA's

first Research Committee was formed, focusing on social stratification and mobility. This Committee's issues dominated the 2nd World Congress in 1953 in Liege, Belgium. By 1959 other Research Committees were being formed, at first referred to as subcommittees. Throughout the 1950s, another seventeen national associations were founded and joined the ISA, and then in the 1960s twelve more were added (Platt 1998). By 1952 the first issue of *Current Sociology*, the journal of the ISA, had appeared, and papers of early World Congresses were published separately as *Transactions*—a practice that ended in 1970 due to the cost and effort involved (Platt 1998). Originally, ISA World Congresses were held every three years, but beginning in 1962 they were moved to today's four-year schedule. Tracing the locations of World Congresses to this point, the 1956 Congress was in Amsterdam, the 1959 Congress in Stresa on Lake Maggiore in Italy, and the 1962 Congress in Washington, D.C.

Because the ISA relied on UNESCO for funding, its activities were heavily influenced by the agenda of UNESCO's Social Science Department (SSD). Early ISA presidents often had served earlier on SSD staff (Platt 1998). One early item on the SSD agenda was using the teaching of sociology to promote international understanding (Platt 1998). Only in the 1960s did teaching steadily lose priority to research, and this same period of transition, not coincidentally, marked the rise to preeminence of the Research Committees—at the expense of the Congress's general program topic and plenary sessions (Platt 1998). By 1970 and the World Congress in Varna, Bulgaria, there were seventeen different Research Committees (Platt 1998).

One issue facing the ISA leadership in its early years was how to deal with the earlier, well-established international association, the IIS. IIS activities had lapsed during WWII, and after the war some of its leaders were tainted with affiliations with defeated Axis regimes (Platt 1998). The same controversy split the national sociological association in Germany by 1960 but not the national associations in Italy or Japan. Regardless, as early as 1953 the leaders of sociology's two international organizations agreed to "friendly collaboration," including a willingness to schedule meetings at different times and to exchange proceedings. Relations between the two organizations have been

relatively amicable ever since. Both international bodies have had to deal with the state of sociology in communist countries during the Cold War, the rising prominence of sociology in Third World countries, and addressing cases of political and scholarly repression in selected countries (Platt 1998).

In 1970 the ISA made two major changes in its founding statutes, one that introduced general individual membership and another that opened up the governance of the Research Committees. ISA leadership saw the Research Committees as a force moving the organization beyond national associations toward a grander internationalism. One result, however, is that the Research Committees became the equivalent of the internally pluralistic “sections” of the American Sociological Association (ASA), as opposed to remaining more coherent research groups. Thus, like ASA sections, the number of Research Committees has increased and many “working groups” and “thematic groups” within each Committee operate relatively independently of each other. In 1994, more than 3,000 sociologists attended the ISA World Congress in Bielefeld, Germany (76 percent from North America and Western Europe), and forty-five national associations were admitted. By 1997 there were fifty-nine identifiable research groups of one kind or another. Concerns abound, therefore, that the organization’s centrifugal tendencies are overwhelming whatever centripetal forces remain (Platt 1998).

One consequence of the increasing internal pluralism of the ISA is that its funding base became and remains “increasingly problematic” (Platt 1998). ISA dues have been raised considerably for individual members, particularly for those living in economically advanced societies (Platt 1998). Throughout the 1970s and 1980s, communist bloc countries continued to support the voting rights of collective members, along with subsidies by affiliated governments. Still, the only ISA World Congress held in a socialist country was the 1970 meeting in Bulgaria. Looking back in 1994, the chair of a Constitutional Revision Committee held that the ISA’s founding statutes had been premised on Cold War politics. His proposal then was that the organization needed to stress even more the importance of membership by dues-paying individuals, as opposed to relying on government

subsidies. Two years earlier, in 1992, the Research Council had voted to give the Research Committees power equal to that of member national associations. For the first time, the Research Committees voted for ISA officers (Platt 1998).

Increasingly in the 1980s, the ISA was concerned about the dominance of First World sociology to the detriment of contributions from sociologists in the Third World. With this in mind, the journal *International Sociology* was founded in 1986 with the mandate to favor scholarly submissions from “disadvantaged areas” (Platt 1998). The first ISA World Congress held in a Third World country was in 1982 in Mexico City (followed by the 1986 Congress in New Delhi).

## REFERENCES

- Clark, Terry N. 1967 “Marginality, Eclecticism, and Innovation: Rene Worms and the Revue Internationale de Sociologie from 1893 to 1914.” *Revue Internationale de Sociologie* 3:3–18.
- 1968 “Rene Worms.” In *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- D’Antonio, William V. 1994 “Sociology and the IIS.” *Annals of the International Institute of Sociology*, Paris, France, pp. 1–19.
- Platt, Jennifer 1998 *History of the ISA, 1948–1997*. International Institute of Sociology.
- Rhoades, Lawrence J. 1981 *A History of the American Sociological Association 1905–1980*. Washington D.C.: American Sociological Association.
- Scheuch, Erwin 1997 “Closing Report.” Address delivered to the 33rd IIS World Congress, Cologne.

DAVID SCIULLI

## INTERNATIONAL LAW

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

International law is the system of rules and principles governing relations at the interstate

level. It originally developed in response to the needs of states but in recent times has grown to include international organizations and, to some extent, individuals.

International law as a systematic body of rules began in Europe in the seventeenth century. Before then, and from earliest history, rules existed governing the interrelations of various groups of people (Nussbaum 1958). But the rules were systematized in Europe only when the contacts among peoples became regular and frequent and the idea of a single ruler for all known society foundered. That occurred with the collapse of the Holy Roman Empire during the Thirty Years War (1618–1648). The state system developed in its place, characterized by a number of kingdoms and principalities, each equal to the others, sovereign within its own borders, and subject to no outside sovereign. Hugo Grotius, a Dutchman who lived during this time, wrote a seminal book, *De Jure Belli ac Pacis* (1620–1625), describing legal rules, derived from natural law, by which these states could achieve peaceful coexistence and, when they failed, how they could conduct their wars with some semblance of humanity. His book popularized international law, and he is generally considered the founder of international law.

The rules and principles Grotius described reflected the characteristics of states. The fundamental notion that states were sovereign and equal became a principle of international law. And because the rulers of states respected that principle as law, they were less likely to wage war or to annex their neighbors. International law helped create the success of the state system and in turn reflected the features of the system.

Grotius derived his rules from natural law, thus suggesting that nature was superior to states. By the nineteenth century, however, theorists abandoned natural law as a source of international law. Instead, they looked at the behavior of the states themselves as the source of international law. International lawyers became positivists, and the state became for international law the ultimate political entity. The rules of international law could guide and could set out regular procedures to ensure the smooth and peaceful conduct of international relations, but they permitted states wide prerogatives. Governments acquiesced in international legal rules because their states benefited

from an orderly system of international relations in which they gave up few of the attributes of absolute sovereignty. For example, the rules did not speak to what states did internally, no matter how egregious. Individuals were not considered subjects of international law. Nor did the rules restrain states from the use of force. International law described permissible uses of force, but states could in effect use force whenever they chose.

By the twentieth century pressures for change began to develop. Technological advances in war and communications accounted for movements aimed at restraining states in their use of force and abuse of human beings. People such as Elihu Root, U.S. Secretary of State in Theodore Roosevelt's administration, wanted international law to provide the vehicle for restraint. Governments started experimenting with dispute settlement through arbitration and courts. They formed the League of Nations to help them control states' uses of force. International law and international institutions were being substituted for unbridled state sovereignty. The concept was a radical departure from the past and came about only over the course of half a century and two world wars.

After World War II, governments were willing in theory to contemplate real restraints on their ultimate sovereign prerogatives. The United Nations was created as an entity under international law. Its charter committed states to uphold human rights, to cooperate in solving world problems, to abandon the use of force, and to follow the commands of the organization itself. The idea was to lessen state sovereignty for the good of the whole world community. Thus, the state began to lose its place as the ultimate political entity almost exactly 300 years after its rise.

Certain international legal norms are now theoretically superior to the wills of the states. In other words, a certain amount of natural law now characterizes the system again. In addition, the state is making way for other types of political institutions such as regional arrangements, although if the state should ever become finally obsolete so will international law. The current form of international law and many of its rules and principles presuppose a system of coequal entities without a single sovereign.

While the state system remains intact, however, international law has taken on an increasingly

important role in governing the relations of states in an interdependent, technologically linked world. It does this even though international law has never had the institutions typical of domestic law: a legislature for making law, an executive to enforce it, and a judiciary to adjudicate and interpret it. In some respects international law still functions as it did in the nineteenth century because the system benefits the state. Nevertheless, law does get made, enforced, and adjudicated, and social movements are at work putting demands on states to form and live by new norms of international law.

Because of the informal condition in which international law exists, however, some legal thinkers argue it is not law at all but rather moral precepts or mere guidelines. Most prominent of these thinkers was John Austin, who described law as a series of commands backed by sanctions. International law has no overarching authority to issue commands, and the sanctions are irregular. But Austin's criticism depends on his definition of law. If law is defined as behavior or behavioral restraint induced by a sense of obligation, international law, in its sphere, is law. Positivist international lawyers also point out that in the end the states acknowledge that international law is law, and that is the relevant indicator.

International legal rules have two basic sources: custom and treaty. Rules of customary international law are created through practices that states engage in because they believe they have a legal obligation to do so. Treaties are the explicit agreements states make with each other to be bound. As the need for international law has grown, states have relied on treaties as a law-making vehicle more and more. The general multilateral treaty has become a common form of law making for important international concerns. For almost ten years nearly every state of the world attended a conference to negotiate a comprehensive treaty on the law of the sea. Because so many states attended and the treaty was so long in the making, the treaty began to take on the characteristics of customary international law, irrespective of its status as a treaty. A similar conference will convene in 1992 to discuss climate, and suggestions have been made to convene such a conference for trade.

When no rule of custom can be found and no treaty exists, international courts have in some disputes turned to a third source of law—general principles of law. These are principles commonly found in domestic legal systems and can serve to fill any gaps in international law, which suffers from its ad hoc law-making process. Theorists consider general principles a subsidiary source of international law because general principles are not made in a positive sense by all the states of the system, and they may be applied to a state that did not wish to be bound by them. Custom and treaty, however, generally allow states to opt out of a rule, thus reflecting the traditional view that the states are superior to the system of law. In the last thirty years, however, states have accepted that certain principles cannot be derogated from because they are considered peremptory norms or *jus cogens*. Examples of such norms are the prohibitions on genocide, the slave trade, and the use of force to advance a state's political agenda. With the concept of *jus cogens*, international law has again taken on some elements of natural law.

*Jus cogens* also exemplifies the extent to which international law has overcome cultural relativity. As new states emerged in the 1960s, scholars from these states questioned whether international law should be binding on them since it was a European product that had aided in perpetuating colonialism. These criticisms have faded, however, because it became clear that international law also created the thing desired most by newly independent countries—statehood. Moreover, because international law is made by states, the new majority could begin to re-create international law. The process of international law has succeeded to the point of bringing states together in accepting that certain principles are overriding, despite the particular value systems of individual states.

Like the law-making system, the law-adjudicating system in international law depends on states volunteering to use it. The system does have courts, in particular, the International Court of Justice. But no state needs to subject itself to the court unless it wishes to do so. The court does have limited compulsory jurisdiction in the case of states that agree to submit cases to it in advance of disputes arising. Nicaragua brought a case against the United States under such an agreement in

1986. States can also agree on an ad hoc basis to submit disputes to arbitral tribunals. The vast majority of international law is adjudicated informally, however. If a state violates international law, such as when Iraq invaded Kuwait in 1990, most states in the system will pronounce their views regarding the legality of the action. Assessing these evaluations leads to conclusions regarding lawfulness. Thus, states make and adjudicate the law themselves.

States also enforce international law themselves. International law is notorious for being poorly enforced. In fact, however, most international law is in fact observed most of the time. Because the states must agree to the law, they tend to make only the laws they want and are willing to live by. Otherwise they opt out of the rules, as the United States has done for some of the new law of the sea. Law is not so well observed, however, in those areas that make headlines—war, human rights, terrorism—which perhaps accounts for international law's poor reputation for enforcement.

When a rule of international law is violated, the state that is harmed is allowed to take action against the perpetrator. For example, if a fishing treaty between state A and state B is violated because state A's fishermen overfish in state B's waters, state B might be entitled to terminate the treaty and prevent future fishing by state A. This system works to some extent, but states have tried to improve on it in recent decades by, first, giving the United Nations and, in particular, the Security Council authority to police some violations of international law and to expand the ability of domestic institutions in enforcing international law by expanding the concept of universal jurisdiction.

The Security Council has authority to maintain peace. In article 42 of the U.N. Charter, it is given the power to call on member states to contribute troops to fight at the direction of the council. The idea comes very close to having an international police force. It has only been used once, however—in Korea in the 1950s. Other attempts have been stymied by the cold war antagonism of the United States and the Soviet Union, each of which has a veto over invocation of article 42. As a sort of substitute, the Secretary General has regularly sent troops, contributed voluntarily by U.N. members, to serve a peacekeeping role. Peacekeeping troops are not supposed to

take enforcement action. But enforcement action may be a possibility again with the end of the cold war. Following Iraq's invasion of Kuwait in 1990, the council ordered worldwide economic sanctions and permitted the use of force, both to enforce the sanctions and to push Iraqi troops out of Kuwait.

Another solution to enforcement has been the widening of universal jurisdiction. To prevent states from interfering in each other's affairs, international law contains principles of jurisdiction defining when and where a state may enforce rules of domestic or international law. In some instances any state may take action. This concept of universal jurisdiction is as old as international law. It was originally developed to handle the problem of piracy. Pirates are defined as persons who commit crimes for profit on the high seas. Generally they do not fly any state's flag, and they act outside the territorial jurisdiction of any state. Typically, the state where the act took place or the state of the pirate's nationality would have jurisdiction, but those categories often do not exist for pirates. The state of the victim might have jurisdiction, but the international system developed the rule that any state may board a vessel that fails to fly a flag and that any state may arrest, try, and punish pirates.

After World War II, universal jurisdiction was expanded to include the concept of crimes against humanity. The victorious allied powers tried German and Japanese individuals, holding them personally responsible for human rights abuses, characterized as crimes against humanity and thus crimes for which any state in the world could take jurisdiction. The Nuremberg and Japanese War Crimes Trials broke new ground in international law by holding that individuals had rights and responsibilities not only under their nation's law but under international law and by expanding the scope of a state's jurisdiction.

Individual responsibility and expanded state jurisdiction are being included today in a variety of treaties, especially related to human rights, narcotics, and terrorism. Customary international law now permits universal jurisdiction over persons who have committed genocide or war crimes. The International Court of Justice has also suggested that important human rights may be enforced by any state regardless of its connection with the

violation because the obligation to respect human rights is an obligation owed to all people; it is a right *erga omnes*.

International law will need these improvements in enforcement. The scope of questions now covered by international law grows annually with the increasing interdependence of the world and the technological advances that bring peoples into conflicting contact. The need to protect the global environment is the newest challenge for international law. States may soon decide they need an international organization to regulate the world trading system. The problems of development, health, communications, education, population, and use of space on earth and in outer space are all new problems in need of attention. Add to them the old problems of war, territorial disputes, governing international organizations, treaties, dispute resolution, refugees, human rights, diplomatic immunity, law of the sea, air space, recognition of new states, and so on, and the growing importance of international law becomes apparent. International law will continue to serve as a means of conducting smooth international relations, its traditional role, but it will also continue to assume new importance as a means of solving problems. In order to achieve this, however, international law must improve its institutions and be accepted by more states, whose own sovereignty will diminish as international law advances.

(SEE ALSO: *Genocide; Law and Legal Systems; Sociology of Law; War*)

## REFERENCES

- Abeyratne, Ruwantissa I. R. 1996 "International Politics and International Justice: Unity in Diversity?" *International Journal of Politics, Culture and Society* 10:291-316.
- Bassiouni, M. Cherif 1993 "The History of the Draft Code of Crimes against the Peace and Security of Mankind." *Israel Law Review* 27:247-267.
- Berry, David S. 1998 "Conflicts between Minority Women and Traditional Structures: International Law, Rights and Culture." *Social and Legal Studies* 7:55-75.
- Binder, Leonard 1996 "The Moral Foundation of International Intervention and the Limits of National Self-Determination." *Nationalism and Ethnic Politics* 2:325-359.
- Brownlie, Ian 1990 *Principles of Public International Law*. Oxford: Clarendon Press.
- Buergethal, Thomas 1997 "The Normative and Institutional Evolution of International Human Rights." *Human Rights Quarterly* 19:703-723.
- Chinkin, Christine 1995 "Violence against Women: The International Legal Response." *Gender and Development* 3:23-28.
- Corntassel, Jeff J. and Tomas Hopkins Primeau 1995 "Indigenous 'Sovereignty' and International Law: Revised Strategies for Pursuing Self-Determination." *Human Rights Quarterly* 17:343-365.
- Dezalay, Yves and Bryant Garth 1995 "Merchants of Law as Moral Entrepreneurs: Constructing International Justice from the Competition for Transnational Business Disputes." *Law and Society Review* 29:27-64.
- Dezalay, Yves, G. Garth Bryant, Pierre Bourdieu, R. C. Austin, David Bonner, Noel Whitty, Sandra Coliver, and Heinz Klug 1996 *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order*. Chicago: University of Chicago Press.
- Gros Espiell, Hector 1998 "Universality of Human Rights and Cultural Diversity." *International Social Science Journal* 50: 525-534.
- Janis, Mark 1988 *An Introduction to International Law*. Boston: Little, Brown.
- Kash, Douglas A. 1998 "An International Legislative Approach to 21st-Century Terrorism." In Kushner, Harvey W., ed., *The Future of Terrorism: Violence in the New Millennium*., Thousand Oaks, CA: Sage
- Knoppers, Bartha M. and Sonia LeBris 1991 "Recent Advances in Medically Assisted Conception: Legal, Ethical and Social Issues." *American Journal of Law and Medicine* 17:329-361.
- Lansky, Mark 1997 "Child Labour: How the Challenge Is Being Met." *International Labour Review* 136:233-257.
- Lecuona, Rafael A. 1997 "International Law, Cuba, and the United States of America." *International Journal on World Peace* 14:37-49.
- Lowe Joseph, D 1994 *Dictionary of Diplomatic, International Law, International Relations Terms*, Berkeley: Joseph D Lowe Publisher.
- Max Planck Institute for Comparative Public Law and International Law 1981 *Encyclopedia of Public International Law*. New York: North-Holland.
- Mendlovitz, Saul and John Foušek 1996 "Enforcing the Law on Genocide." *Alternatives* 21:237-258.
- Nippold, Otfried and S. Hershey Amos 1998 *The Development of International Law after the World War*. Holmes Beach: Gaunt-Incorporated.



Nussbaum, Arthur 1958 *A Concise History of the Law of Nations*. New York: Macmillan.

Otto, Dianne 1996 "Subalternity and International Law: The Problems of Global Community and the Incommensurability of Difference." *Social and Legal Studies* 5:337-364.

Parry, Clive, and John Grant 1988 *Encyclopedic Dictionary of International Law*. New York: Ocean Publications.

Stone, Julius 1983 "A Sociological Perspective on International Law." In R. St. J. McDonald and D. M. Johnston, eds., *The Structure and Process of International Law: Essays in Legal Philosophy Doctrine and Theory*. The Hague: Martinus Nijhoff.

MARY ELLEN O'CONNELL

## INTERNATIONAL MIGRATION

*International migration* is a term used to refer to change of usual residence between nations. The number of international migrants is always much smaller than the total number of persons traveling across international frontiers, because the overwhelming majority of such travelers do not intend to change their usual residence. International migration is contrasted with *internal migration*, which refers to a change of usual residence within a nation. The term *immigration* is used to denote the flow of persons establishing a usual residence in a given nation whose last residence was in some other nation. The term *emigration* is used to denote the flow of persons relinquishing a usual residence in a given nation to establish residence in some other nation. *Net international migration* denotes the difference between the number of persons immigrating to a given nation in a given period and the number emigrating from that nation in the same period.

Immigratory and emigratory events constitute two of the four components of national population change; the other two components are births and deaths. For most nations, population change is determined predominantly by the balance of births and deaths (natural increase). However, for a few nations in certain periods, the net international immigration has also been an important component of the total population change.

In determining the number of persons who have changed residence among nations, national

statistical agencies must specify the meaning of a change in usual residence. The United Nations (1978) suggests that international movements with an intended stay of more than one year be classified as international migration. Unfortunately, there is considerable lack of uniformity among nations with respect to how international migration is defined. For example, according to data of the Mexican government, some 46,000 Mexicans emigrated to the United States in 1973; according to data of the U.S. government, the number of permanent legal immigrants from Mexico was about 72,000 (United Nations 1978). Also, many governments, including that of the United States, collect data on immigration but not on emigration. Finally, all data on immigration published by governments refer to legal immigration only. Data on illegal or undocumented immigration cannot be tabulated.

Certain terms useful for the study of either international or internal migration will now be explained. A *migration stream* is defined as the total number of migratory events from place A to place B during a given time. The *counterstream* is defined as the total number of migratory events from place B to place A. The sum of events in the stream and counterstream is termed the *gross interchange* between A and B. The *effectiveness of migration* is defined as the ratio of the net migration between A and B and the gross interchange between the two places. The effectiveness of migration can therefore vary from a low of 0 to a high of 1. For most pairs of geographic units, the effectiveness of migration tends to be much closer to 0 than to 1.

Petersen (1975) makes very useful distinctions among the concepts of free, impelled, and forced migrations. In free migration, the will of the migrant is the main factor. In impelled migration, the will of the migrant is subordinated to the will of other persons. In forced migration, the will of other persons is paramount, and the will of the migrant is of no weight at all. Another useful term is *return migration*, defined as migration back to a place in which one had formerly resided. *Chain migration* (MacDonald and MacDonald 1964) is also a frequently used concept. It refers to the common pattern whereby a given individual migrates to a particular destination in which he or she already has kin or friends who have previously migrated from his or her own area of origin.

### **MIGRATION DIFFERENTIALS**

It is universally observed that the propensity for international migration is strongest among young adults. Other differentials in migration tend to be limited to particular cultures or locales. Because the highest propensity for international migration is among young adults, the contribution of international migration to population change is often considerably greater than the net international migration by itself. This is because the birthrate for migrants is higher than for the total population, and the death rate is lower.

### **DETERMINANTS OF THE VOLUME OF INTERNATIONAL MIGRATION**

Demographers analyze the determinants of the volume of a migratory stream into two components. The first concerns the specific propensity to migrate for individuals of each given type. The second concerns the number of individuals of each given type. The volume of a migratory stream can be calculated as the sum of the products obtained by multiplying the specific propensity to migrate for individuals of each given type by the number of individuals of that type.

The determinants of the propensity to migrate may conveniently be analyzed in terms of a preference system, a price system, and the total amount of resources available for all goals (Heer 1975, 1996). The preference system describes the relative attractiveness of various places as goals for potential migrants, compared to other goals that their resources would allow them to pursue. An area's attractiveness is the balance between the positive and negative values it offers.

Among the most important of the positive values is the prospect of a better-paying job. Other positive values achieved by migration include the chance to live in a more favorable climate, freedom from persecution, opportunity for marriage, and continuation of marital ties. In the case of forced migration, the positive value achieved is simply to save one's own life.

However, international migration also creates negative values. A major disincentive to migration is that it involves a disruption of interpersonal

relationships with kin and old friends. Chain migration is so attractive precisely because it mitigates the disruption of such relationships (Massey et al. 1987). Other negative values created by international migration are the necessity of learning new customs and, often, a new language. Laws restraining legal entry or departure are also, of course, very important deterrents to international migration and will be discussed later in more detail.

The price system describes the costs in money, energy, and time, which cannot be used in the pursuit of other goals, imposed by a given migration decision. Since the cost of international migration generally varies in direct proportion to the distance traveled, the number of immigrants to a given place tends to vary inversely with the distance.

The total resources available for all goals also affects the decision to migrate. If the only drawback to migration is the expense of the move, an increase in monetary income should increase the probability of migration.

### **MAJOR STREAMS OF INTERNATIONAL MIGRATION**

Certain major streams of international migration deserve mention, either because they have had important historical consequences or because they otherwise exemplify unusual patterns. One of the earliest streams of international migration with historical significance was the westward movement of nomadic tribes in Europe and Central Asia at the time of the fall of the Roman Empire. The many tribes that moved westward during this period included those speaking Celtic, Germanic, and Ural-Altai languages. As the easternmost tribes moved westward, they pushed forward the tribes in front of them. One suggested explanation for this extensive migration is that the grasslands of Central Asia had become desiccated. A second possibility is that an expanding Chinese Empire disrupted the life of the nomadic tribes near its borders and, thus, provoked the movement of all the other tribes (Bury 1928; Huntington 1924; Teggart 1939).

The European and African migrations to North America, South America, and Oceania have probably had more important historical consequences

than any other migratory stream. This flow began slowly after Columbus's voyage to America in 1492. It has been estimated that more than 60 million Europeans have left for overseas points in the centuries since then. However, net migration was lower, since many of those leaving Europe later returned (United Nations 1953, pp. 98-102). The migration from Africa to the New World was almost wholly a forced migration of slaves. The first slaves were brought to the colony of Virginia in 1619, and the slave trade in the United States was not legally ended until 1808. During the period of slave trade, about 400,000 Africans were brought to the United States (U.S. Bureau of the Census 1909). The impact of the migration of slaves is revealed by the fact that, in 1790, 20 percent of the 4 million persons in the United States were black.

During the twentieth century, the origin of immigrants to the United States shifted drastically away from Europe and toward Asia and the Americas. This change is illustrated in Figure 1, which shows the number and percentage distribution by region of last residence of immigrants to the United States by decade from 1891-1900 to 1991-96.

The emigration from Puerto Rico to the mainland United States, of major magnitude in the years following World War II, is of interest because it exemplifies an extremely high rate. According to the 1970 census, the combined total of the population of Puerto Rico and of persons in the United States of Puerto Rican birth or parentage was about 4.1 million, of which around 1.4 million were in the United States. Thus 33.9 percent of all Puerto Ricans were on the mainland (U.S. Bureau of the Census 1971, 1973).

Immigration into Israel following World War II is likewise noteworthy because it exemplifies an extremely high rate. In 1948, when independence was established, the total population of Israel was 650,000; by 1961, after the influx of more than 1 million immigrants, it had risen to 2.2 million (Bouscaren 1963; United Nations 1966).

Perhaps the world's largest gross interchange in a short time took place in India and Pakistan following the 1947 partition of British India and the establishment of these two areas as independent states. This migration is also of interest because it was impelled rather than free. In the face

of violence, Hindus and Sikhs in Pakistan moved to India and Muslims in India moved to Pakistan. From 1947 through 1950, 10 million persons migrated from Pakistan to India and 7.5 million from India to Pakistan (Spate 1957).

The two most recent major streams of international migration exemplify what has been termed *labor migration*. Labor migration is said to occur when immigrants are legally admitted to a nation for defined time periods in order to alleviate a shortage of labor. Labor migrants are not given the right of permanent residence. The first stream was the large-scale migration of workers into the prosperous nations of northern and western Europe from poorer nations in the Mediterranean region such as Italy, Spain, Portugal, Yugoslavia, Greece, Turkey, Algeria, and Morocco (Massey et al. 1998). This stream began around 1960 and ended in 1973, following the sudden elevation of petroleum prices by the Organization of Petroleum Exporting Countries (OPEC). The proportion of the total population that was foreign increased substantially in all of the northwest European nations. For example, from 1960 to 1970, the foreign population of the German Federal Republic increased from 1 percent to 5 percent and that of Switzerland from 9 percent to 16 percent (Van de Kaa 1987). The second major stream was the large-scale migration of workers into the major oil-producing nations in the Persian Gulf region out of such nations as Jordan, Egypt, Yemen, Pakistan, and India (Massey et al. 1998). For example, from 1957 through 1975, 70 percent to 75 percent of the total labor force in Kuwait consisted of foreigners (Birks and Sinclair 1981).

#### CONSEQUENCES OF INTERNATIONAL MIGRATION

One may examine the possible consequences of international migration for the individual, the area of net emigration, the area of net immigration, and the larger social system, which includes areas of net emigration and net immigration. The discussion must be in part speculative, since knowledge about these topics is incomplete.

Before a move, an immigrant will have anticipated a net balance of favorable consequences. Sometimes, however, reality will fall short of expectations, and dissatisfaction will provoke the

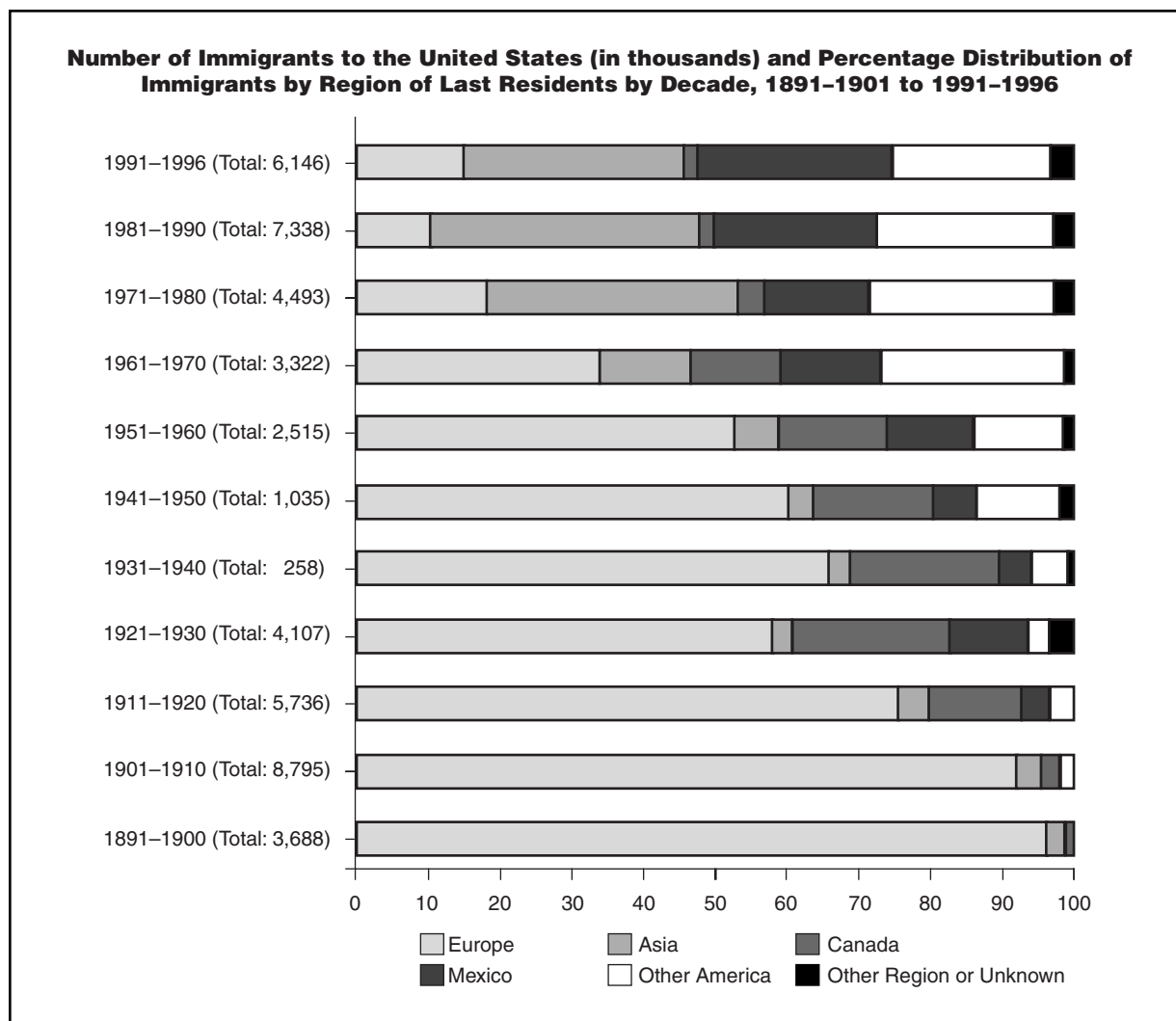


Figure 1

immigrant to either return to the nation of origin or, on occasion, to move on to some other nation.

Net emigration may have several important consequences for an area. By relieving population pressure, it may cause the average level of wage and salary income to rise. Remittances from emigrants may also be helpful. On the other hand, net emigration may cause the value of land and real estate to decline. Moreover, areas of net emigration suffer the loss of investments made to raise and educate children who spend their productive years elsewhere. This loss is particularly large when the individual receives a higher education prior to emigration. Such a loss is termed *brain drain*.

Finally, since emigration rates are selective by age, nations with substantial net emigration may have relatively few young adults relative to the number of children and the aged.

Net immigration may also have important consequences. If the area is definitely underpopulated, the resultant population increase may help the nation to achieve economies of scale (reduction in the cost of goods obtainable by increasing the scale of production and marketing) and, thus, raise the general standard of living. Under other circumstances, net immigration may result in some decline in average wage and salary income. In either case, a net flow of immigrants tends to raise

the price of land and real estate. Furthermore, in general, net immigration increases the proportion of young adults in the total population. Dependent on their composition, immigrants may receive either more or less in government benefits than the amount of their tax payments. Finally, net immigration may make the population more heterogeneous with respect to race, religion, or language.

For the system comprising the nations of both net inflow and net outflow, the direct effect of international migration is of course to promote a redistribution of population. If migrants have been responsive to differences in job opportunities, this redistribution may further the economic development of the total system. Moreover, a substantial amount of international migration responsive to job opportunities might also induce a decline in the degree to which there is economic inequality among nations. Unrestricted international migration might make the poor nations richer and the rich nations less prosperous. However, there is substantial disagreement among scholars as to what the effect of unrestricted immigration from the poor nations might be on the prosperity of the rich nations such as the United States. Simon (1984) believes that net immigration to the United States will serve to increase its average income. Davis (1981), on the other hand, takes a much more pessimistic view. Additional analysis of the consequences of immigration from poor nations to the United States is found in a study sponsored by the National Research Council (Smith and Edmonston 1997, 1998).

#### **LEGISLATION AFFECTING IMMIGRATION AND EMIGRATION**

National laws concerning immigration have varied from almost complete prohibition to positive encouragement. Laws restricting emigration are now relatively rare but have been of important consequence in modern times for at least one nation, the USSR.

In the seventeenth and eighteenth centuries, a mercantilist ideology, which saw a large population as the key to national wealth and power, encouraged many of the governments of Europe to attempt to prohibit emigration and to encourage immigration. In the late seventeenth century, the French minister Colbert enacted legislation

prescribing the death penalty for persons attempting to emigrate or helping others to emigrate, except to a French colony. In 1721, Prussia passed a similar law. Moreover, the Prussian emperor Frederick the Great invested state funds in subsidizing the settlement of immigrants. In eighteenth-century Russia, both Peter the Great and Catherine the Great subsidized colonists from abroad, mostly from Germany (Glass 1940).

The nineteenth century, influenced by the economic doctrine of *laissez faire*, was the great period of unrestricted international migration. During this century the European governments freely permitted emigration, and the newly independent United States of America welcomed millions of immigrants.

After World War I, the United States took a more active role in restricting international immigration. The major events in this connection were changes in immigration law in 1921 and 1924 that greatly restricted the number of immigrants to the United States, establishing a quota for each of the countries outside the Western Hemisphere. Furthermore, the nations of northwestern Europe was given much larger quotas relative to their populations than those of southern and eastern Europe. This was done even though, in the immediately preceding years, rates of immigration from southern and eastern Europe had been much higher than those from northwestern Europe. The justification made at the time for the quota differentials was the presumed greater ease with which immigrants from northwestern Europe could assimilate (Eckerson 1966).

By the 1960s, a changing climate of opinion with respect to the inferiority or superiority of different ethnic groups made it possible for President Kennedy to advocate the abolition of the discriminatory national-origins quota system, and a law accomplishing this was enacted in 1965 under the Johnson administration. The 1965 law called for the abolition of the national-origins quota system as of July 1, 1968, but, nevertheless, imposed an overall annual quota of 170,000 immigrants from outside the Western Hemisphere and 120,000 from within it (exclusive of immediate relatives of U.S. citizens). This legislation granted preference to persons with relatives already in the United States, to persons with needed occupational skills, and to refugees. Additional legislation

passed in 1976 abandoned the separate quotas for the two hemispheres and imposed a 20,000 limit on immigrants from any nation in the Western Hemisphere (the 20,000 limit had previously been in existence only for Eastern Hemisphere nations). The major effect of the 1976 legislation was to make it more difficult for Mexicans to enter the United States legally.

Some nations, while placing severe restrictions on immigrants in general, have made use of positive inducements to encourage immigration from selected nations or groups. Currently the best example of such legislation is that of Israel, which has committed itself to encouraging the immigration of Jews from anywhere in the world. Formerly, Canada and Australia also exemplified such policies of selective encouragement. Each of these nations subsidized immigrants from European nations while placing severe restrictions on the immigration of nonwhites (Bouscaren 1963; Petersen 1964).

The United States, Australia, Canada, and Israel have laws that allow immigrants permanent residence leading to citizenship. Since 1950, other nations, particularly in northwestern Europe and in the Persian Gulf, have had policies that encourage only labor migration, that is, migration of workers, mostly male, who were supposed to return to their native countries following a fixed term. In the case of the European nations at least, these policies had unintended results. The contract workers were allowed to bring their dependents to live with them. Hence, they became permanent immigrants even though this was not the result intended by original policy (United Nations 1982).

In many nations of the world, particularly in the United States, a major phenomenon is the existence of illegal, or undocumented, immigrants. A necessary condition for the existence of illegal immigration is a lack of congruence between the laws regulating the supply of legal immigrant opportunities and the demand for them. For example, the demand to immigrate to the United States from a particular nation should be a reflection of that nation's population size and the average propensity to immigrate to the United States if there were no legal restrictions. Accordingly, nations with large population size are likely to have more

immigrant demand than nations with small populations; yet all nations, without regard to population size, have the same annual quota. Furthermore, national differences in individual propensity to immigrate to the United States should be a function of such variables as the difference in standard of living compared with that of the United States, difference in degree of public safety compared to the United States, proximity to the United States, degree of similarity with the language and culture of the United States, and prior existence of immigrants that allows for chain migration. According to the official estimate of the U.S. Immigration and Naturalization Service, there were 5 million undocumented immigrants in the United States in October 1996. Of these, 2.7 million, more than half the total, were from Mexico. An additional 335,000 were from El Salvador and 165,000 from Guatemala (U.S. Immigration and Naturalization Service 1997, p. 198). Given Mexico's population size and presumed high average propensity for immigration to the United States, one can easily explain why such a very large proportion of all undocumented persons in the United States have been from that nation (Heer, 1990). An additional fact of interest is the high proportion of undocumented immigrants residing in California. According to the official Immigration and Naturalization Service estimate for October 1996, 40 percent of all undocumented immigrants were living in that state (U.S. Immigration and Naturalization Service 1997).

Rising concern over the extent of undocumented immigration into the United States led the U.S. Congress to enact the Immigration Reform and Control Act of 1986. The intent of the legislation was to eliminate the presence of undocumented aliens in the United States either by legalizing their status or by forcing them to leave the country. This act had two key provisions. The first was the imposition of sanctions upon employers who knowingly employed illegal aliens. The second was the provision of a process whereby undocumented persons who had lived in the United States continuously since January 1, 1982, or had worked in U.S. agriculture for ninety days in the period from May 1985 to May 1986 were allowed to become temporary legal residents. After a short time, they would be allowed to become permanent legal residents. The success of employer sanctions is problematic because sanctions can be applied only

if the employer knowingly hires illegal aliens and because it is relatively easy for an undocumented person to present to the potential employer either fake documents or the documents of some other legally resident person. On the other hand, more than 3 million persons applied for legalization of status after the act was passed, among whom about 2.3 million were from Mexico (U.S. Immigration and Naturalization Service 1990).

In 1998 more than 80 percent of the world's 5.9 billion persons lived in one of the less developed nations. The annual rate of natural increase in these nations was around 1.7 percent; in the developed nations it was only 0.1 percent. In the developed nations per capita gross national product was \$20,240; in the less developed nations, only \$1,230 (Population Reference Bureau 1998). These facts imply a strongly increasing demand for immigration to the developed nations from the less developed. Given the current barriers to legal immigration imposed by the developed nations, undocumented immigration will be of increasing prevalence unless the governments of the developed nations take extraordinary measures to curb it.

## REFERENCES

- Birks, John, and Clive A. Sinclair 1981 "Demographic Settling Amongst Migrant Workers." In *International Population Conference, Manila 1981*, vol. 2. Leige, Belgium: International Union for the Scientific Study of Population.
- Bouscaren, Anthony T. 1963 *International Migration Since 1945*. New York: Praeger.
- Bury, B. 1928 *The Invasion of Europe by the Barbarians*. London: Macmillan.
- Davis, Kingsley 1981 "Emerging Issues in International Migration." In *International Population Conference, Manila 1981*, vol. 2. Leige, Belgium: International Union for the Scientific Study of Population.
- Eckerson, Helen F. 1966 "Immigration and National Origins." *Annals of the American Academy of Political and Social Science* 367:4-14.
- Glass, David V. 1940 *Population Policies and Movements in Europe*. Oxford: Clarendon Press.
- Heer, David M. 1975 *Society and Population*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- 1990 *Undocumented Mexicans in the United States*. New York: Cambridge University Press.
- 1996 *Immigration in America's Future: Social Science Findings and the Policy Debate*. Boulder, Colo.: Westview Press.
- Huntington, Ellsworth 1924 *Civilization and Climate*. New Haven, Conn.: Yale University Press.
- MacDonald, John S., and Leatrice D. MacDonald 1964 "Chain Migration, Ethnic Neighborhood Formation, and Social Networks." *Milbank Memorial Fund Quarterly* 52: 82-97.
- Massey, Douglas, Rafael Alarcon, Jorge Durand, and Humberto Gonzalez 1987 *Return to Aztlan: The Social Process of International Migration from Western Mexico*. Berkeley: University of California Press.
- Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1998. *Worlds in Motion: Understanding International Migration at the End of the Millennium*. Oxford: Clarendon Press.
- Petersen, William 1964 *The Politics of Population*. Garden City, N.Y.: Doubleday.
- 1975 *Population*, 3rd ed. New York: Macmillan.
- Population Reference Bureau 1998 *World Population Data Sheet*. Washington, D.C.: Population Reference Bureau.
- Simon, Julian 1984 "Immigrants, Taxes, and Welfare in the United States." *Population and Development Review* 10:55-69.
- Smith, James P., and Barry Edmonston, eds. 1997 *The New Americans: Economics, Demographic, and Fiscal Effects of Immigration*. Washington, D.C.: National Academy Press.
- eds. 1998 *The Immigration Debate: Studies on the Economic, Demographic, and Fiscal Effects of Immigration*. Washington, D.C.: National Academy Press.
- Spathe, O.H.K. 1957 *India and Pakistan: A General and Regional Geography*. New York: Dutton.
- Teggart, Frederick J. 1939 *Rome and China: A Study of Correlations in Historical Events*. Berkeley: University of California Press.
- United Nations 1953 *The Determinants and Consequences of Population Trends*. New York: United Nations.
- 1966 *United Nations Demographics Yearbook, 1965*. New York: United Nations.
- 1978 *United Nations Demographic Yearbook, 1977*. New York: United Nations.
- 1982 *International Migration Policies and Programmes: A World Survey*. New York: United Nations.
- U.S. Bureau of the Census 1909 *A Century of Population Growth in the United States: 1790-1900*. Washington, D.C.: U.S. Government Printing Office.

—1971 *U.S. Census of Population: 1970*. Final Report PC(1)-A1. Washington, D.C.: U.S. Government Printing Office.

—1973 *U.S. Census of Population: 1970*. Final Report PC(2)1E. Washington, D.C.: U.S. Government Printing Office.

U.S. Immigration and Naturalization Service 1990 *Provisional Legalization Application Statistics, January 9, 1990*. Washington, D.C.: U.S. Immigration and Naturalization Service.

—1997 *1996 Statistical Yearbook of the Immigration and Naturalization Service*. Washington, D.C.: U.S. Government Printing Office.

Van de Kaa, Dirk J. 1987 "Europe's Second Demographic Transition." *Population Bulletin* 42:1-59.

DAVID M. HEER

## INTERNET

The Internet is a curious phenomenon. It's a vast international institution of critical and growing importance, yet in another sense its properties are so evanescent that it's tempting to say that Internet does not exist; there's no *there* there except the process itself. The reason is that at heart Internet is nothing more than a highly specific collection of rules or standards for computers called communication protocols. Even the name comes from the U.S. Department of Defense standard TCP/IP (terminal control protocol/internet protocol). What we now see as institutional Internet—the fiber optics and copper cabling, the billions of dollars of computer hardware and software, the business and governmental organizations, and body of human skill and knowledge that's both on the Net and in the user community—emerged and grew with volcanic force inside the purely formal boundaries of the protocol system. What the protocols allowed, the world's social and economic system instantiated.

A critical part of this development was that the protocols in question were public and open, that they were standards and not products per se. This is in contrast to the course of other communications systems, such as the telephone, radio, and television, for which key components were held proprietary or otherwise regulated by patents,

secrecy, and both governmental and private monopolies. No one would claim that what has become the Internet was created for the common welfare of all humankind, or that it is guaranteed to work for that welfare in the future. In fact, some of the key components of the Internet came from the bowels of the Pentagon at the height of the Cold War, and the overall moral valence of the Internet is as unsettled a question as one could imagine. But it is useful to note that in many ways the Internet is government work that made good and that its ongoing success appears to depend on public and open standards. Attempts to commercially supplant the Internet or take it proprietary have so far been unsuccessful, despite the huge upsurge of Net-based commerce.

## HISTORY

The history of the Internet is partly the history of technology, the hardware of the system. But more important are the social aspects of computer-mediated communication, the human networks created by the hardware system. There are two polar points of view on the social impact of the Internet on communities and society. On the one hand, enthusiasts have argued that the Internet removes barriers that have historically divided people and opens the way to unprecedented equality in social interactions and for social opportunities (Rheingold 1993). On the other hand, critics contend that the Internet removes humanity from social interaction and strands people in an impersonal virtual world without touch, dignity, or personhood (Stoll 1995). The earliest empirical research to address the actual uses and impact of the Internet in social life seems to suggest that people experience Internet interactions in ways similar to their face-to-face social experience. But the technology is young and the scope of its use has broadened only in the 1990s. A decade into the twenty-first century, the prognosis may become clear.

**Hardware.** If the Internet is anything physical it is a network of computers—actually, a network between networks, an internetwork. Computer networking is not new. Many of the house-sized 1960s-style computers, called mainframes, were often linked into larger networks, although this typically required that computers be in the same



building if not the same room. This constraint was partly overcome by using radio-frequency messaging over coaxial cables similar to those now used for cable TV.

Computer networking using coaxial cable became common and useful in the 1960s. But a problem arose when the University of Hawaii decided to adopt a multicampus computer network. Cable connections in Hawaii would not only have to span long distances, they would have to span those distances deep under the Pacific Ocean. The cost was too high even to contemplate. Computer scientists began a search for alternatives and settled on radio-based computer communications—network messages would be broadcast on special channels. The result was called AlohaNet.

A problem immediately arose. The distance between islands in Hawaii is typically only a few hundred miles. It took only a few thousandths of a second (or less) for radio messages to span that distance, but that's a fairly long time for a computer. This meant that two computers at remote locations could start talking at once, each not hearing the other immediately, which garbled their messages. Garbling of this sort was called a message collision. To solve this problem, the AlohaNet designers had to build collision-detection circuitry into their network controllers. A computer would not only have to send messages and listen for other's messages, it would have to check for message collisions, and re-send if a collision occurred.

Such a scheme sounded awkward, but in practice it was elegant. Collision detection and retransmission overhead turned out to be a tiny part of the overall cost in time and speed of the network. Broadcast-style, collision-detecting networks have a number of advantages. First, the number of computers in network (usually called nodes) is not necessarily fixed. If rules are set up to allow it, any computer in a broadcast system can chime in and negotiate to be part of the network. Once that computer is known to the network, messages can be routed to it as if it had been there all along. Second, as long as there are two or more computers up and running in a given network segment, they can talk; their peers could go on- or off-line as it suited them without bringing down the net. The flexibility of broadcast-style networks led to a revolution in computer communication.

## **ETHERNET AND LOCAL AREA NETWORKS**

Of course, networks actually broadcasting would soon fill up the available radio-frequency spectrum. Consequently, most systems that actually were put in use contained their transmissions inside dedicated runs of coaxial cable, or, later, telephone-like twisted-pair conductors. The most popular of these was the Ethernet system developed by a consortium of computer makers. The Ethernet system allowed a fairly high-speed (1 million bytes per second) network based on coaxial cable to span a physical space one kilometer end-to-end, longer with repeaters to boost the signal. This meant that all but the largest of buildings and many whole complexes and campuses could be served by a single, fast local area network (LAN). Originally, in the early 1980s, Ethernet was used mostly to tie together collections of mainframes and smaller minicomputers, but within a few years it was possible inexpensively to integrate most organizations' growing collections of personal computers (PCs—IBMs and clones, Macintoshes, and others) into an Ethernet system. The easy availability of fairly fast, fairly cheap Ethernet LAN systems, and the rising tide of PCs that could use them, was half the nascent Internet equation. You could call it the demand side. Computers increasingly were used not so much as solo workstations but as components in LANs; their interoperability became one of their principal virtues.

**ARPAnet and DOD Connectivity.** The supply side of the Internet equation actually preceded the growth of LANs by a few years. In the late 1960s, the Department of Defense (DOD) began pumping increasing amounts of money into the softer side of computer science: artificial intelligence, computer graphics, computer voice and hearing, visual processing, natural-language processing, computer-based language translation, and computer networking. The source of the funding was the Defense Advanced Research Projects Administration (DARPA, or more commonly, ARPA). The mode of funding was unusual for DOD, in that the research was mostly unclassified and public, in the form of grants to universities and think-tanks.

Public (or partly public) projects seemed to call for a method of openness and interchange different from the Pentagon's historic need-to-know procedures. The method that was suggested

was the development and use of a computer network to accomplish document and data interchange and what would now be called electronic mail (e-mail). ARPAnet was built by the Massachusetts Institute of Technology (MIT) and a think-tank—Bolt, Baranek and Newman (BBN)—to link together ARPA grant sites and allow ARPA investigators and their staffs to communicate electronically. Physically, ARPAnet used high-speed and expensive leased AT&T telephone long lines to connect to fairly large computers, like the DECsystem 10s and 20s and PDP 11/70s from BBN's neighbor Digital Equipment Corporation (DEC). Human users typically used simple display terminals to interact with to their local host computers, which sent the messages out over the network. Specialized and proprietary interface gear made the whole thing work. Only members of the ARPA grant club were allowed to participate, which led to hard feelings among those excluded—although cost would have excluded many others anyway. Even in the early 1980s, after ARPAnet had been taken over by the National Science Foundation and become NSFnet (the Internet's immediate ancestor), the cost of participation included about \$10,000 up front for network gear, substantial use of a \$150,000 minicomputer, and quite a few hundred dollars each month for leased lines.

At least two university-based networking systems sprang up in the late 1970s and early 1980s in an attempt to provide ARPAnet-like connectivity to non-ARPA institutions. Both used a mail-drop scheme to avoid the high costs of leased phone lines. In other words, the participating computers used modest-speed, dial-up modems (like those of most contemporary home PCs in the 1990s, only slower) to phone a predetermined list of their peers at fixed intervals, usually once or twice a day after long-distance rates went down. It worked like an activist organization's telephone tree: each computer called *N* neighbors, who called *N* neighbors, and so on, until all the messages were delivered. The process was slow—e-mail sent this way might take overnight or longer to arrive—but it was much less expensive than the full-time alternatives. Mail-drop networks tended to be organized by host-machine type. Universities with big IBM mainframes tended to support Bitnet operation, while sites using the AT&T Unix operating system or clones typically adopted a scheme called Usenet.

The obvious advantages of intersite networking to university researchers and others created a strong demand for this kind of service. Full-time live service such as that provided by ARPAnet was particularly advantageous, since it allowed sharing of computer facilities and easy access to scarce computing resources, such as supercomputers. When the U.S. National Science Foundation (NSF) began supporting such supercomputer facilities at about the same time DOD was losing interest in its specialized networking, NSF essentially took over the ARPAnet infrastructure and renamed it NSFnet. (ARPAnet lived on for a while inside NSFnet.) NSFnet was initially set up along the same lines as ARPAnet—only institutions that were NSF grant recipients were welcome as members. But interest from other academic and research entities (as well as the broader scope of NSF) caused the system to grow into what is now the Internet.

The Internet in the 1980s and early 1990s was still for the most part a network between networks, the lower-level networks being the university and commercial LANs. Real participation on the Internet by a PC required a dedicated LAN connection, usually a coaxial or twisted-pair link using Ethernet technology, and a high-speed leased link to the Internet backbone. These were available in most university environments but were prohibitively costly to arrange at home. Software developers soon came up with a solution: an Ethernet emulator that ran over ordinary telephone lines using relatively low-cost commercial modems. Serial Line Interface Protocol (SLIP) and its successors allowed ordinary PCs in ordinary homes to dial into Internet Service Providers (ISPs) and function as full-fledged peers on the Internet system. Now, commercial ISPs brought the Internet into homes and small businesses at prices comparable to ordinary voice telephone service.

NSF decommissioned NSFnet in 1995 amidst political fanfare, creating a generic Internet and opening the door to a growing swell of private and commercial development. The electronic marketplace surged, with on-line shopping in the 1998 holiday season topping \$1.2 billion over America On-Line (AOL) alone. Private citizens routinely communicated electronically with families and strangers, and located services and information readily from PCs in libraries, schools, businesses, and homes.

Meanwhile, NSF turned its attention to development of the next wave of innovation. In partnership with MCI WorldCom, NSF created and supports the Very High Performance Backbone Network Service (vBNS), also known as Internet 2 and Next Generation Internet. This network links the two leading-edge supercomputing sites and at the end of the 1990s, connected 150 research institutions nationwide. Internet 2 seeks to accelerate Internet development and enable a new generation of applications to improve media integration, real-time collaboration, and interactivity. Although not yet generally accessible by ordinary users, Internet 2 clearly suggests that real time multimedia virtual interaction will soon be possible around the globe. (See King, Frinter, and Pickering 1997, for more detailed history of the Internet.)

### INTERNET MODALITIES

**E-mail.** Probably the most common form of Internet communication is electronic mail, which is text-based messaging from a single computer user to one or more recipients. Internet routing services are used to send the message from a sender's address (e.g. gwjones@msn.com) to a recipient's (e.g., tbsmith@udom.edu). Individual addresses represent dedicated mailboxes—actually, special computer files—maintained by the users' ISPs. The mailbox address precedes the @ symbol, and the ISP's address follows it. E-mail has traditionally been plain, unformatted text, but newer e-mail software has the capacity to handle styled text, display graphics, even sound and video. Although these facilities are not yet universal, they suggest that e-mail will become more fluid, stylized, and expressive. In text-based e-mail, emotions are commonly expressed with “emoticons,” symbol combinations with a cartoonish character such as :- ) for humor or :-( for sadness.

E-mail can take a mass or bulk form. For example, persons with a common interest can sign up to be members of a mailing list under the control of a computer program called a list server, usually called a listserv or mail list. Mail sent to the listserv is relayed to all list members, so that the resulting interchanges have the form of a public broadcast to other listserv members. Some listservs automatically broadcast all messages received, while others are moderated by a member who screens messages for conformity to the group's purpose.

Sometimes mass mailings take a hostile or abusive turn, as in the case of unsolicited commercial e-mail (UCE), usually labeled spam. (“Spam” is not an acronym but an unflattering reference to the lunchmeat; apparently, the usage is an allusion to a Monty Python skit in which a restaurant patron can't get any meal that doesn't contain Spam. Most precisely, spam is abusive mass broadcasting to a Usenet newsgroup, as described below, but the term has been borrowed by the e-mail world.) UCE, or spam, is bulk mail sent inappropriately to computer users who have not requested it. UCE is abusive because it is analogous to junk surface mail sent with postage due—the UCE victim, after all, has to pay for the connection over which the message is delivered, and all Internet users pay indirectly for the infrastructure that carries spam.

Spamming is one example of cultural conflicts that emerge from the migration of commercial practice to the Internet. The Internet grew up in military and academic subcultures where norms and expectations for behavior are explicitly not-for-profit. With the commercialization of the Internet in the late 1990s, many business practices were translated to the new medium without adaptation to existing prescriptions and proscriptions. “Netiquette”—etiquette for the Internet—continues to evolve and issues of social control over the Internet are largely unsettled at the turn of this century.

People tend to assume that e-mail is a relatively private medium. Historically, no one knew an e-mail address unless the user disclosed it. Today, institutions and ISPs publish e-mail addresses, and search engines on the World Wide Web (WWW) can locate many user addresses. In some ways, the e-mail address is coming to replace the telephone number as a direct personal identifier. People seem increasingly to use e-mail instead of the telephone to reach friends, family, and business contacts. With the rise of telephone answering machines and whole-family participation in the work force in the 1990s, delayed responding to any kind of messaging has become normative. E-mail allows people to receive and respond to messages at their convenience, 24 hours per day. Because it is written, messages can be clearly stated and thoughtfully framed. For communication across long distances, e-mail is less expensive than the telephone and can communicate documents

much more rapidly than any postal or delivery service. These uses plainly are not revolutionary, just partial replacements of existing communication methods (c.f. Kraut, Patterson, Lundmark, Kiesler, Mukophadhyay, and Scherlis 1998).

Unlike telephone calls, e-mail is usually archived by the servers that send and receive messages, the nodes between corresponding users. In the United States, the legal precedent is that information on the server, or on machines owned by business or government, is owned by the owner of the machine and is not private information. E-mail correspondence may be subject to review by superiors, may be subpoenaed by courts, and may even be considered public records in some states. The privacy protection of telephone wiretapping laws was not extended to e-mail.

E-mail and listservs have become important tools for passing announcements and other information among members of existing face-to-face groups. Listservs are common forums for professional communications and announcements and are often used in classroom study groups to encourage open questioning and discussion. Researchers use listservs to disseminate new findings, share resources, locate grant programs, and even to recruit study participants.

For professional and other formal communications, e-mail has successfully facilitated social contact, social networking, and information distribution (Wellman and Gulia 1999). However, in private communications with family and close friends, e-mail has increased the amount of contact people have, but has not replaced the telephone or face-to-face meetings. People sometimes e-mail even to those they see daily. But they do not replace their visits with e-mail. This suggests that people are missing something when e-mailing to their close others. Perhaps it is the experience of seeing another's face, or hearing another's voice. It remains to be seen whether voice- or video-enhanced e-mail will become accepted as a face-to-face surrogate.

**Usenet.** Usenet, often called the Internet news facility, is a giant messaging system through which loosely organized thematic discussion takes place on-line. Worldwide, millions of users linked to hundreds of thousands of computer sites write messages that they post to one of more than 30,000 topical newsgroups. These posts are then

distributed by Usenet-serving computers to Usenet users. Usenet began in 1979 as a mail-drop messaging alternative to the then-exclusive ARPAnet (see earlier) but evolved into a broadcast-based medium on the Internet.

Usenet topical areas or newsgroups vary by seriousness, from the highly focused to the totally whimsical. Newsgroups with names like "sci.archaeology" or "comp.lang.c++" tend to be largely topic-oriented, while many others have no real purpose other than to offer forums for freewheeling electronic banter and disreputable, nuisance advertising. Newsgroups are loosely organized and vary considerably in their scope and traffic. There are newsgroups offering social support to the ill and grieving and newsgroups offering sexual information to the naive and inexperienced. Information on starting newsgroups is available in the newsgroup news.groups.

Users are not required to wade through thousands of newsgroups to find messages of interest. Most Web browsers now have news-reading functions, or specialized programs are available for this purpose. People typically subscribe to selected newsgroups that they want to examine regularly, so that when they go on-line they will see these 10 or 30 or 200 newsgroups, not 30,000. Messages posted to Usenet are organized in discussion threads, that is, messages related by reference to the same initial post or topic. Users may skip some threads on their subscribed newsgroups, reading only those threads of personal interest.

Usenet has a character similar to that of e-mail-based listservs, although its explicitly public nature gives it a somewhat different feel. Usenet newsgroups historically predate listservs, so that much of netiquette evolved from these public discussion forums. Admonitions against spam originated on Usenet, as did the tendency for people to flame norm violators and discussion rivals. Flaming involves posting a hostile and insulting message to the group, intended to shame the author of a prior post. Flaming may have had its origin in a puerile effort to enforce group norms, but it is just as often a deliberate attempt to violate them for shock value or amusement. A deliberate attempt to provoke a hostile response on Usenet is called trolling; trolling/flaming battles dominate some newsgroups.

Usenet norms are discussed and described in newsgroups devoted to this topic, and summaries are posted in news.announce.newsusers. Additionally, most newsgroups have a document called the Frequently Asked Questions list (the FAQ) that deals with local norms and standards. Beyond these formal statements, something not unlike a community character emerges and is maintained by regular participants in some newsgroups over time. New users, seeking to fit in there, need to go through a process of socialization first, usually by lurking and observing. Other newsgroups, however, are anarchic, and still others are abandoned and ignored.

Research has suggested that newsgroups provide a setting intermediate between the public and the private where stigmatized social identities can be established and supported. For example, McKenna and Bargh (1998) found that homosexuals who had never made their sexual orientation public found the courage to do so through social support on Usenet. In another domain, however, Mickelson (1997) found that mediated social support on Usenet was less helpful to parents of attention-deficit/hyperactive children than face-to-face support in therapy groups. Researchers may investigate other socially unacceptable phenomena by identifying populations and soliciting volunteers on Usenet. Moreover, the messages posted to newsgroups are considered public behavior. As such, they offer a rich resource for social researchers. There is at least one important possible problem: Usenet users may be deceiving others in their self-presentations and in their messages. However, this problem permeates all self-report research and is not specific to Internet phenomena.

**Chat.** Computer chat evolved from grassroots networking attempts somewhat oblique to the development of the Internet. Chat was initially a feature of old-fashioned local bulletin-board systems (BBS), in which a system operator, or sysop, with (usually) a DOS PC allowed a dozen or so outside parties to dial in with modems and read Usenet-like delayed messages, or to communicate with each other in real time, by typing messages that appeared on the screen with labels indicating from whom they came. Some of the early national time-sharing services such as CompuServe offered this feature, too, often with a name like “CB simulator,” indicating the metaphor of the period,

citizen-band radio (CB). The idea was that people who were basically strangers would adopt “handles” (later called screen names) that concealed actual identity while advertising proclivities, and that these strangers would engage in streams of banter like CB radio operators out on the highway.

With the advent of mass-marketed national services such as AOL, Microsoft Network, and the like, chat facilities—now organized into topical areas or chat rooms—became one of the principal attractions. Chat and chat metaphors soon became second only to the World Wide Web as the public image of the Internet. Most users access chat facilities through commercial providers where it is implemented on private servers. There are also pure peer-to-peer Internet chat programs, such as Internet Relay Chat (IRC), that don’t require a host chat provider. Participation in IRC and related free-floating chat is probably dwarfed by use of AOL and similar commercial services.

Like Usenet, topical options in chat are numerous and variously populated. Users interact in real-time with minimal constraints. As with Usenet, the degree of group norm enforcement in chat rooms varies from high to nonexistent. On commercial providers’ chat facilities, broad norms of conduct often are officially enforced by staff members and user volunteers. Other hand-me-downs from Usenet were adapted to the constraints of chat. For example, the chat equivalents of Usenet’s emoticons are one-letter abbreviations in angle brackets, like <g> for grin. Apparently, these are both easier to type and to parse than graphic emoticons when one is engaged in chat activity.

With the emergence of chat came a wave of real-life chatter about the development of close relationships between people who met on the Internet. People arranged marriages to others they had never met face-to-face but with whom they had chatted endlessly. The self-disclosure and narrow scope of interactions, as well as the sheer number of encounters, provided a fertile breeding ground for social relationships. Many marriages and friendships founded there have endured, while others have dissolved. While the surge of new relationships and the awe of onlookers have stabilized, chat shows no signs of losing popularity.

Another chat phenomenon, cybersex, involves graphic verbal descriptions of sex acts, exchanged

in real time by couples and groups. It is the Internet equivalent of a graphic romance novel, except it is interactive and participatory. Participants may self-stimulate while generating messages to other users. Touted as an opportunity to free the libido from social restriction and personal inhibition, users may engage in unusual acts with minimal personal risk. Cybersex has been controversial. Like arguments over pornography in print, issues of access and censorship remain unsettled as we enter the twenty-first century.

Research on chat has focused primarily on two aspects of chat: the effect of anonymity in chat rooms on the negotiation of social identity (Turkle 1995); and the development of community among interacting users (Wellman and Gulia 1999). Most of this research has involved qualitative analysis or sometimes merely impressionistic characterization of chat activity. Many case studies have been conducted that describe activity in novel chat rooms or the chat experiences of particular individuals. Like Usenet, some chat is public behavior that is ripe for empirical study.

Chat is engaging, just as conversation with similar others is engaging in face-to-face interactions. But with chat, users hear all the conversations in the room, not only their own. Moreover, self-presentations can be well controlled to avoid sharing unflattering information. For these reasons, some people become obsessed with chat activity and seem to develop a dependency on this type of social contact. Others, particularly those who are homebound, have found fulfillment of their social needs through chat. Chat seems to attract a different type of user, relative to Usenet. The chat user seeks the interaction, perhaps more than the information available from topical discussion. Usenet users may seek the information and prefer not to be distracted by the interaction.

**Multi-User Domains (MUDs).** In the late 1970s and early 1980s, as the Net was still in its infancy, the fantasy role-playing game of Dungeons and Dragons swept through high schools and colleges. MUDs (Multi-User Domains or Multi-User Dungeons) originated as electronic platforms for similar fantasy games. They have become more elaborate and more free of goal-based game character as they have grown in popularity. They are based on different kinds of software (specified by names

such as MUSE, MOO, or MUSH) that can be accessed through the Internet.

Users, known as players, join MUDs through a command that connects their computer to the computer running the game program. Connection gives players access to a shared database of rooms, exits, and other virtual objects, which are created and manipulated through simple commands. Players create a character to play (called an avatar) by giving the character a name and a description, and enter rooms. Players interact with each other using simple commands such as “say” to talk to others in the room, “whisper” to talk to specific others, or “emote” for nonverbal expressions. Avatars grow and develop through interaction and game experience.

While some MUDs have graphical interfaces that allow characters and objects to be represented by icons, most interfaces so far have been textual. MUDs require the player’s imagination to create the objects, the actions, and the outcomes of the process. There is no necessary goal, and rules can be fluid. The interactions are the thing of interest. The Usenet news group `rec.games.mud` periodically lists Internet-accessible MUDs with their complete network addresses.

There has been tremendous interest in MUDs because they are the most unusual modality on the Internet, most different from RL, that is, real life. Avid players remain connected for days, cycling between RL and several different MUDs. Players may be especially susceptible to Internet dependency as some are drawn deeper into the fantasy world of the MUDs. There, identity is self-described and under the complete control of the player. One may be whoever one wishes to be. And actions are free of RL consequences, so freedom is perfected in the MUDs. Finally, anonymity is complete, with no means available to identify players in RL. These features allow MUDs to offer a utopian existence in a virtual world that is forever changing and changeable.

Researchers interested in the effects of simulation and role play on the development of identity are flocking to MUDs to investigate. Researchers interested in the effect of special types of relationships, or exposure to violence, on the family are also observing MUDs. And researchers interested in Internet dependence are focused on MUDs as well (e.g. Turkle 1995).

While interesting and entertaining environments, it may be premature to generalize findings based on observations of MUDs to other types of Internet experience. Although MUDs are anonymous, as are some chats, most Internet activity is identifiable and public. Although MUDs, and some chats, are user-created and malleable fantasy forums, most Internet activity is quite constrained both by norms and technical requirements, as well as considerations of cost and practicality. MUDs will likely continue to appeal to certain subpopulations, but they are not expected to become primary activities of Internet users. In fact, the trend at the turn of the century suggests that traditional games are becoming available as network software that can be shared and played interactively with others (Monopoly, bridge, Risk, etc.). They are appealing to a broad audience, both young and old. These games involve no avatars and no explicit fantasies about social identity. It is too early to tell whether they will compete effectively with MUDs for player time in RL.

**World Wide Web (WWW).** E-mail, Usenet, and chat facilities are to some extent immediate and ephemeral. They depend on user activity to generate their subject matter, and they require a constant stream—in the case of chat, a stream in real time. Other Internet-based facilities arose in an attempt to provide access to more fixed corpora of text: public-domain literature, programs, documents, and so forth.

Information protocols were an attempt to link data-storage facilities and user software to make Internet-based information search and retrieval convenient. A rudimentary form was File Transfer Protocol or FTP. FTP allowed external users to go to a foreign computer—if they knew where it was, what they were looking for, and how to sign on—and retrieve some of the files to which they were given access. Gopher, another information protocol, was similar in that it put selected parts of a computer file directory on the Net in a form suitable for browsing. A network of cooperating Gopher servers built a large-scale index of their total contents that was available as an entry point. Wide-Area Information Service (WAIS) was a more industrial-strength and business-oriented approach to the problem.

Gopher, WAIS, and to a significant extent most of the rest of the Internet have been superceded

by World Wide Web, a generalized hypertext facility that has submodalities corresponding to most other computer communication tasks. The idea of hypertext had been around since the 1960s—a computer-driven book with interconnected parts. Users could browse the hyperbook by following the linkages. Hypertext had been implemented on single computers or networks in various ways in the 1970s and 1980s. In 1990, at the European high-energy physics lab CERN (Conseil Européen pour la Recherche Nucléaire, Council for European Nuclear Research), Tim Berners-Lee developed a method of doing this over the Internet, which he called World Wide Web. WWW, or the Web as it is more often called, uses coded tags and descriptors (e.g. URLs, HTTP, HTML) to specify dynamic behaviors on a user's computer screen when links are invoked. This selection opens a file or otherwise triggers an event, spilling information to the user's screen.

WWW caught on explosively when inexpensive ISP access and Web service became widely available. Many individuals established WWW sites to describe their personal lives, their work, families, friends, and even their pets. For academics, Web sites typically described professional interests. Many course materials were posted to Web sites where students could retrieve them at any time. Businesses posted their personnel directories, and cities posted their public officials. Cinemas posted their playing times. It became a kind of instant, paperless desktop publishing mixed with features of a home broadcasting studio.

WWW is not socially interactive in the sense that e-mail or Usenet are interactive. Instead, interaction is scripted and restricted to some fixed set of user actions, somewhat like a video game. But the coupling of WWW with database applications and complicated scripting languages gave it a dynamic feel. Animations, sounds, and other types of sensory stimulus contribute to this feeling, and these frills will become the norm with Internet 2.

While mainstream business, institutions, and interest groups have WWW sites describing themselves, so do deviant interests. Because interaction is controlled by the Webmasters, criticism for deviance can be avoided. Militant and extremists groups, pornographers, and political interests abound and freely argue their unpopular opinions

on WWW. Issues of censorship and access restrictions, particularly for children, are largely unsettled.

Information is located on the Web by search engines that use logical operations to specify criteria for matching against hypertext found on registered servers. Different search engines, such as Yahoo or AltaVista, work with different subsets of WWW sites, so search results vary. Webmasters link together related sites and improve their odds of appearing in the search results of interested users. And respectable information is very prevalent, although sometimes difficult to distinguish from the propaganda. WWW does not provide a reference librarian or tour guide, and the only requirement to post information is access to a Web-serving computer. For this reason, educators were encouraged to emphasize strategies for evaluating source credibility and information integrity in their curricula, and researchers are scrambling to understand how people may effectively discern valid information.

Professionals already disseminate their research results and theories on WWW. Many electronic journals are available only on-line; some print journals are also published electronically. Many researchers post their own work to a personal Web page for others to easily access and reference. However, publishing a work on the Web may interfere with journals' claims to copyright on the same material, so many professional organizations are discouraging the posting of research papers prior to mainstream publication. Still, post-print abstracts and full-text articles are available in on-line databases and libraries. This wealth of information has made literature searching simpler and more useful than ever before.

Researchers have also established Web sites for data collections for studies of almost everything imaginable. Research materials (e.g., questionnaires, stimuli) are posted to a site accessed by a code provided by researchers to participants. Some studies simply recruit all comers to participation. Some research sites are scripted, so that the user follows a specified sequence of activities. Others require relatively few and simple actions. Participants prefer the convenience of WWW administrations to face-to-face meetings with researchers. But researchers cannot verify that participants are who they say they are, and some control over

the situation is lost. Early comparisons of pencil-and-paper questionnaires to WWW questionnaires found no differences in responding due to administrative mediums for short, self-report data where sampling was controlled (Kardas and Milford 1996).

WWW as a platform for propaganda, commerce, and entertainment has not been lost on the public. In fact, many people are betting large sums of money in development and investment that WWW is revolutionary. Already, users can read about a new vocalist, find her latest recording, listen to it, buy it, and order tickets to her upcoming concert—all accomplished in a few minutes without leaving home or waiting in lines! For many users, WWW has replaced their telephone books, mail-order catalogues, newspapers, and libraries, as all of this information is easily accessed on the Web.

Since live audio, live video, live telephone transmissions, and similarly active content can be transmitted and manipulated inside WWW, on the Internet, we would seem to have arrived at a stage at which function does not necessarily follow the network form. The old Internet, with text-based messaging and 48-hour relay turnarounds, seems hopelessly outdated, even quaint. What we've reached seems to be a stage of pure mediation: The Internet, through the mechanism of WWW, has the capacity to be a conduit for nearly any form of information, limited only by the available network bandwidth, transmission speed, and user patience. The extent to which the social conduct of life on the Net carries traces of its ancestry remains to be seen.

#### REFERENCES

- Kardas, E. P., and T. M. Milford 1996 *Using the Internet for Social Science Research and Practice*. New York: Wadsworth.
- King, J., R. E. Grinter, and J. M. Pickering 1997 "The Rise and Fall of Netville: The Saga of a Cyberspace Construction Boomtown in the Great Divide." In S. Kiesler, ed., *Culture of the Internet*. Mahwah, N.J.: Erlbaum.
- Kraut, R., M. Patterson, V. Lundmark, S. Kiesler, T. Mukophadhyay, and W. Scherlis 1998 "Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-Being?" *American Psychologist* 53:1017-1031.
- McKenna, K. Y. A., and J. A. Bargh 1998 "Coming Out in the Age of the Internet: Identity 'Demarginalization'



Through Virtual Group Participation.” *Journal of Personality and Social Psychology* 75:681–694.

Mickelson, K. D. 1997 “Seeking Social Support: Parents in Electronic Support Groups.” In S. Kiesler, ed., *Culture of the Internet*. Mahwah, N.J.: Erlbaum.

Rheingold, H. 1993 *The Virtual Community*. New York: HarperCollins.

Stoll, C. 1995 *Silicon Snake Oil: Second Thoughts on the Information Highway*. New York: Doubleday.

Turkle, S. 1995 *Life on the Screen*. New York: Simon-Schuster.

Wellman, B., and M. Gulia 1999 “Net-Surfers Don’t Ride Alone: Virtual Communities as Communities.” In B. Wellman, ed., *Networks in the Global Village*. Oxford, England: Westview.

DIANA ODOM GUNN  
CHRISTOPHER W. GUNN

## INTERPERSONAL ATTRACTION

Everyone meets many people. With some, there is a natural fit; with other, there isn’t. Liking a person is quite different from liking chocolate or liking to ski. Liking someone implies feelings of warmth, intimacy, and consideration and a desire to spend time together. Interpersonal attraction plays a large part in the formation of all relationships except those into which a person is born, that is, all nonascriptive relationships. Everyone uses tactics that are expected to recruit potential partners; the specific tactics used in presenting oneself, as well as the characteristics an individual looks for in others, will vary depending on whether what is sought is friendship or love or a good working partner (McCall 1974). In earlier studies, questions of affiliation were confused with questions of attraction. You may be attracted to many people, but only those who are available in terms of physical proximity and who are defined as appropriate by social norms will actually become interaction partners (Berscheid and Reis 1998, p.204).

Even though liking someone is based on many factors that can’t always be defined, a person does know, upon meeting someone, whether he or she is in fact liked. This perceived liking in turn draws us toward the other (Sprecher and Hatfield 1992).

Men and women operate differently in the area of choosing people as being attractive. For example, men are more inclined to reject a person who disagrees with them than are women and more likely to choose the same type of person as a friend and as a marriage partner (Lindzey and Aronson 1969).

First impressions don’t necessarily last. Nisbett, reanalyzing Newcomb’s data in 1989, found that people’s liking of other people after sixteen weeks’ acquaintance was not predicted very well by their initial liking of these others after one week’s acquaintance (Nisbett and Smith 1989, p.72)

## THEORETICAL EXPLANATIONS OF INTERPERSONAL ATTRACTION

Homans, working from the perspective of exchange theory, states that people consider the rewards versus the costs of any potential relationship (Lindzey and Aronson 1969) and are attracted to those people who provide the most reward at the least cost. From this perspective, the ideal relationship is one in which both participants have equal costs and rewards, so that neither feels cheated or exploited. Newcomb asserts that frequency of interaction is an important determinant of attraction, a view known as the *propinquity perspective*. The basic assumption is that the more frequently one interacts with others, the more attractive they become. It is expected that frequency of interaction will lead to increasing similarity of beliefs and values and that this *assumed similarity* will in turn lead to increased attraction. This perspective ignores the possibility that getting to know a person better may actually reveal many differences (Lindzey and Aronson 1969). Despite the idea’s appeal to common sense, there is little evidence of increasing reciprocity of interpersonal attraction over time (Kenny and LaVoie 1982).

People do prefer those who are similar in background, interests, and values. They want to talk about things that interest them and do things familiar to them. A person who can provide social support by having similar beliefs and values is a likely potential friend. Despite the folk wisdom that opposites attract, similarity is more powerful than complementarity. The exceptions are those with strong needs on either end of the dominance–submission continuum or the nurturance–succorance continuum (Argyle 1969, p. 213); when

strong needs exist in these areas, complementarity is more powerful.

### TRYING TO ATTRACT OTHERS

Though different factors come into play when one is evaluating someone as a potential friend or a potential work partner or a potential romantic partner, there seem to be inferred qualities that make a stranger appear to be likable or not likable. One study found that when videotapes of women were shown to males and females to judge, those most often chosen were apt to be described as sociable, cheerful, and positive emotionally; the underchosen were more apt to be described as negative and moody (Hewitt and Goldman 1982).

In a culture, like the United States, that values openness, psychological awareness, and emotional vulnerability, self-disclosure increases likability. Those who disclose little are less apt to be found attractive by others (Montgomery 1986).

Revealing yourself to another person is a sign that you like and trust them. It also signals that you trust them to respond appropriately. There seem to be three specific links between self-disclosure and attraction and likeability: (1) the more you disclose about yourself, the more you are liked; (2) people disclose more to those they initially like; and (3) the very act of disclosing makes you like the person to whom you disclosed (Collins and Miller 1994). A modern form of self-presentation that tells quite a bit about what people think makes them appear attractive is the personal ad. No longer are these dismissed as being for the desperate; rather, they are seen as just another way to introduce oneself. A study of the responses to different sets of physical characteristics referred to in ads showed that tall men and thin women received the greatest number of responses (Lynn and Shurgot 1984). In these ads it can also be seen that people present themselves as happy, able, capable, and very successful. It is interesting to note that the richer a man claims to be, the younger, taller, and prettier a woman he wants. The younger and prettier a woman presents herself as being, the more successful a man she wishes to meet.

### FORMING RELATIONSHIPS

First meetings proceed cautiously. In every cultural group there are conventions about how long the

preliminaries must last. These conventions vary depending upon the age and gender of the participants, as well as where the meeting takes place. The goals of the encounter determine the interpersonal attraction tactics used. For example, when characteristics of potential dating partners were varied along two dimensions, physical attractiveness and personality desirability, undergraduate males chose physical attractiveness as the deciding variable (Glick 1985). Therefore, a female hoping for a date would find that increasing physical attractiveness would be more effective than showing what a nice personality she had. Not all initial encounters develop into relationships. In general, whether the individual develops an expectation that future encounters will be rewarding is critical to the continuation of relationship. Several studies have shown that perceiving the relationship as being better than others' relationships facilitates commitment to and satisfaction with the relationship (Rusbult et al. 1996).

Relationship satisfaction also positively correlates with evaluating your partner more positively than he or she evaluates self. Such evaluations are also positively related to more effectively resolving any conflicts that occur and thus to continuation of the relationship (Murray and Holmes 1996).

The goals of the encounter determine the interpersonal attraction tactics used. For example, when characteristics of potential dating partners were varied along two dimensions, physical attractiveness and personality desirability, undergraduate males chose physical attractiveness as the deciding variable (Glick 1985). Therefore, a female hoping for a date would find that increasing physical attractiveness would be more effective than showing what a nice personality she had.

### PLAYING AND WORKING TOGETHER

Competition has an interesting relationship with interpersonal attraction. Rees found that during intragroup competition, football players reported the most liking and respect for those who played their own position yet outperformed them (Rees and Segal 1984). Riskin also found that males, when given background data indicating both the degree of competitiveness and the degree of work mastery in target males, rated the most competitive as most attractive, as long as they were also

seen as having ability. In addition, these competitive males were assumed by the male subjects to be more attractive to women (Riskin and Wilson 1982). Numerous studies have shown that emergent leaders are given high interpersonal-attraction ratings by both sexes.

The workplace provides a setting where qualities of competitiveness, ability, and leadership are displayed. It might be assumed that this leads to the formation of romantic attachments. Though this does in fact occur, the work setting also provides for additional complexity in the handling of personal attraction. Attraction and intimacy must be seen in the context of outsiders' view of the relationship. Attempts must be made to balance the demands of the job and those of the relationship. Role relationships within the workplace are expected to contain a degree of distance that is at odds with the demands of "getting closer." Despite these problems, people do get romantically involved with co-workers. One study of 295 adults (average age: thirty-two) revealed that 84 had been involved in a romantic relationship with someone at work and 123 had been aware of a romance in their workplace. Such relationships are more likely to occur in less formal organizations, especially those that are very small or very large. The person most likely to enter into such a relationship is a female who is young, new, and of low rank (Dillard and Witteman 1985, p. 113).

### FRIENDSHIP

Being perceived as friendly, pleasant, polite, and easy to talk to increases a person's ability to attract potential friends. If in addition similar values, interests, and backgrounds are present, the likelihood of friendship is even greater (Johnson 1989). In ongoing relationships, friendship has nothing to do with the participants' rating of each other's physical appearance. Nevertheless, at the initial meeting stage a person judged as being too physically attractive will be avoided. In one study, sixty undergraduate males were shown a male target population (previously rated from 1 to 5 by a male and female sample) and were asked who among this group they would like to meet. The most attractive were chosen less frequently; they were judged to be more egocentric and less kind. It was the moderately attractive who were seen as being the type of person most of the subjects would like

to meet. Explaining these findings in terms of exchange theory, one would say that most people rate themselves as bringing moderate attractiveness to a relationship and feel that extreme attractiveness throws off the equality (Gailucci 1984).

Though it is often assumed that young people do not see older people as potential friends, a review of forty research reports reveals that perceived agreement in attitudes tends to neutralize young adults' general perception of older adults as unattractive. Elders may perceive young people as attractive or unattractive, but they still prefer to associate with individuals who are middle-aged or older (Webb et al. 1989).

### SEXUAL ATTRACTION AND ROMANTIC RELATIONSHIPS

While males and females alike differ in their ability to distinguish between friendly and sexually interested behavior, males are more likely to see sexual intent where females see only friendly behavior. When shown videotapes of five couples, each showing a male and a female behaving in either a friendly or a sexually interested way, males consistently saw more sexual intent (Shotland and Craig 1988).

Men and women also differ as to the relative importance of physical features and personal qualities in determining the choice of romantic partners. Even though both sexes rated personal qualities as being more important than physical features, males placed greater emphasis on the physical than did women (Nevid 1984). Despite this, there seems to be a point at which attempting to increase physical appeal by dressing to reveal the body has a negative effect on one's appeal as a marital partner. Hill reports that when male and female models wore very tight clothes that displayed a great deal of skin, they were rated as being very attractive as potential sex partners but their marital potential was lowered. High-status dressing had the opposite effect for both males and females: Ratings of physical, dating, sexual, and marital attractiveness all increased as the status of clothing increased (Hill et al. 1987).

A shared sense of humor is another important component of loving and liking. When a humor

test comprising cartoons, comic strips, and jokes was given to thirty college couples, along with a test that measured how much the partners loved and liked each other, a strong correlation between shared humor and a predisposition to marry was found (Murstein and Brust 1985). It can probably be assumed that the shared humor comes before the relationship, as well as serving as a factor that enhances it. How people feel about their romance at any given time tends to cause them to rewrite history. For example: When people involved in romantic relationships were asked, once a year for four years, to describe how the relationship had changed during the past year, it was found that current feelings had more to do with the ratings that actual changes (Berscheid and Reis 1998, p. 211).

(SEE ALSO: *Courtship: Exchange Theory; Love; Mate Selection Theories; Personal Relationships; Social Psychology*)

REFERENCES

- Argyle, Michael 1969 *Social Interaction*. Chicago: Aldine.
- Berscheid, Ellen, and Harry T. Reis 1998 "Attraction and Close Relationships." In D.T. Gilbert et. al., eds., *The Handbook of Social Psychology*, 4th ed. New York: McGraw-Hill.
- Collins, N. L., and L. C. Miller 1994 "Self-Disclose and Liking: A Meta-Analytic Review." *Psychological Bulletin* 116, 457-475.
- Dillard, James P., and Hal Witteman 1985 "Romantic Relationships at Work: Organizational and Personal Influences." *Human Communication Research* 12:99-116.
- Gailucci, N. 1984 "Effects of Men's Physical Attractiveness on Interpersonal Attraction." *Psychological Reports* 55:935-938.
- Glick, Peter 1985 "Orientations Toward Relationships: Choosing a Situation in Which to Begin a Relationship." *Journal of Experimental Social Psychology* 21:544-562.
- Hewitt, J., and Morton Goldman 1982 "Traits Attributed to Over- and Under-Chosen Women." *Psychological Reports* 5:431-439.
- Hill, Elizabeth, Elaine Nocks, and Lucinda Gardner 1987 "Physical Attractiveness: Manipulation by Physique and Status Displays." *Ethnology and Sociobiology* 8:143-154.
- Johnson, Martin R. 1989 "Variables Associated with Friendship in an Adult Population." *Journal of Social Psychology* 129:379-390.
- Kenny, David A., and Lawrence LaVoie 1982 "Reciprocity of Interpersonal Attraction: A Confirmed Hypothesis." *Social Psychology Quarterly* 45:54-58.
- Lindzey, Gardner, and Elliot Aronson 1969 *The Handbook of Social Psychology*, Vol. 3. Reading, Mass.: Addison-Wesley.
- Lynn, Michael, and Barbara A. Shurgot 1984 "Responses to Lonely Hearts Advertisements: Effects of Reported Physical Attractiveness, Physique, and Coloration." *Personality and Social Psychology Bulletin* 10:349-357.
- McCall, George J. 1974 "A Symbolic Interactionist Approach to Attraction." In T. L. Huston, ed., *Foundation of Interpersonal Attraction*. New York: Academic Press.
- Montgomery, Barbara 1986 "Interpersonal Attraction as a Function of Open Communication and Gender." *Communication Research Reports* 3:140-145.
- Murstein, Bernard I., and Robert G. Brust 1985 "Humor and Interpersonal Attraction." *Journal of Personality Assessment* 49:637-640.
- Murray, S. L., and J. G. Holmes 1996 "The Construction of Relationship Realities." In G. Fletcher and J. Fitness, eds., *Knowledge Structures in Close Relationships: A Social Psychological Approach*. Mahwah, N.J.: Erlbaum.
- Nevid, Jeffrey F. 1984 "Sex Differences in Factors of Romantic Attraction." *Sex-Roles* 11:401-411.
- Nisbett, Richard E., and Michael Smith 1989 "Predicting Interpersonal Attraction from Small Samples: A Re-Analysis of Newcomb's Acquaintance Study." *Social Cognition* 7:67-73.
- Rees, C. Roger, and Mady-Wechsler Segal 1984 "Intragroup Competition, Equity and Interpersonal Attraction." *Social Psychology Quarterly* 47:328-336.
- Riskin, John, and David Wilson 1982 "Interpersonal Attraction for the Competitive Person: Unscrambling the Competition Paradox." *Journal of Applied Social Psychology* 12:444-452.
- Rusbult, C.E. et. al. 1996 "An Interdependence Analysis of Accommodation Processes." In G. Fletcher and J. Fitness, eds., *Knowledge Structures in Close Relationships: A Social Psychological Approach*. Mahwah, N.J.: Erlbaum.
- Shotland, R., and Jane Craig 1988 "Can Men and Women Differentiate Between Friendly and Sexually Interested Behavior?" *Social Psychology Quarterly* 51:66-73.
- Sprecher, Susan, and Elaine Hatfield 1992 "Self-Esteem and Romantic Attraction: Four Experiments." *Recherches de Psychologie Sociale* 4:61-81.

Webb, Lynn, Judith Delaney, and Lorraine Young 1989  
 "Age, Interpersonal Attraction and Social Interaction: A Review and Assessment." *Research on Aging* 11:107-123.

ARDYTH STIMSON

## INTERPERSONAL CONFLICT RESOLUTION

Most of us understand conflict as a negative aspect of social interactions and therefore are inclined to avoid it if possible. Yet theorists contend that conflict is an inevitable part of human association and, to some extent, a necessary one (Straus 1979). In fact, conflict, either intrapsychic or interpersonal, is a pivotal concept in theories of human development (Shantz 1987). Even with its acknowledged importance in human development, however, research on conflict has been relatively recent. This may perhaps have been due to the perception of conflict as largely a negative occurrence and the equating of conflict with aggression (Shantz 1996). If interpersonal conflict is not aggression, what is it? How do individuals respond to interpersonal conflict? What is our theoretical understanding of individual behavior in interpersonal conflict, and what are the factors that are related to how individuals resolve conflicts with others? This entry addresses the above questions with a focus on childhood and adolescence. The term *conflict* will refer to interpersonal conflict unless otherwise stated.

### WHAT IS CONFLICT?

*Interpersonal conflict* is an event that occurs between two individuals in the course of interactions. Thus, it requires at least two individuals for conflict to occur; conflict is not an attribute of a single individual (Shantz 1987). According to conflict theory, which draws from symbolic interactionism, exchange theory, and systems theory, conflict is defined as "a confrontation between individuals, or groups, over scarce resources, controversial means, incompatible goals, or combinations of these" (Sprey 1979, p. 134). From the sociolinguistic perspective, conflict is a "social activity, created and conducted primarily by means of talking" (Garvey and Shantz 1992, p. 93). Thus, the main elements of conflict are an interactional

context, which is at least dyadic, and the presence of behavioral opposition (Joshi 1997).

An important condition for sustained interactions is the maintenance of interpersonal equilibrium. Conflict causes disequilibrium, and the individuals involved in the conflict can use different ways to move the interactions to an equilibrium. These different ways are conflict resolution strategies and can be defined as "sets of behaviors that seem to subserve a social goal. They may be either conscious, planned means to foreseeable ends (as these terms usually connote), or unconscious, automatic or habitual behaviors that have the effect of subserving goals" (Shantz 1987, p. 289).

### THEORETICAL MODELS RELATED TO INTERPERSONAL CONFLICT RESOLUTION

There are two theoretical models that attempt to explain how individuals respond in conflict situations—the social information processing model and the interpersonal negotiation skills model. The social information processing model focuses on the intraindividual cognitive processes that take place in social situations. The INS model, on the other hand, describes the developmental changes in conflict resolution strategies that occur with cognitive changes.

**Social Information Processing.** According to the social information processing model (Dodge et al. 1986), individuals go through a set of steps in which they process social information contained in social situations. However, the person is generally not aware of these information-processing steps. The individual must respond skillfully at every step to successfully negotiate the situation. The first step is to encode the social cues in the situation, which means that the person assesses the situation in terms of what exactly is happening. The next step is to interpret the cues in the light of previous knowledge, thus forming mental representations of the cues. Thus, the understanding of the situation—why is this conflict happening? why is the other person behaving in this way?—is constructed. Once the cues are interpreted, the individual must generate possible ways of responding to the situation, which in conflict situations would be conflict resolution strategies. This is followed

by the choice of an appropriate response, and finally the individual must enact the chosen response behaviorally.

A revision of the model (Crick and Dodge 1994) highlighted the role of arousal regulation, especially between the steps of cue interpretation and generation of strategies. This means that the emotions experienced by the individual may influence the goals and responses generated, and eventually chosen. Thus, the social information processing model focuses on the intrapersonal processes that underlie interpersonal conflict resolution.

#### **Interpersonal Negotiation Strategies Model.**

The interpersonal negotiation strategies (INS) model (Selman and Demorest 1984; Yeats et al. 1990) draws from both the cognitive-developmental and the information-processing perspectives. It predicts that cognitive development is accompanied by an increased sophistication in the ability to coordinate the social perspective of the self and the other. Three factors implicated in the choice of an interpersonal negotiation strategy determine the developmental level of the strategy—the understanding of the self's and the other's perspective, the purpose of the strategy, and the affective control, or the way in which individuals understand and balance the affective disequilibrium. A fourth factor, the orientation, determines the type of strategy that the individual uses, namely, self-transforming or other-transforming. Children move through four developmental levels of social perspective taking—undifferentiated/egocentric, differentiated/subjective, self-reflective/reciprocal, and third-person/mutual. At the undifferentiated/egocentric level, children are not able to differentiate between the thoughts and feelings of self and the other. Their strategies, therefore, tend to be unreflective and impulsive, using either force or obedience. The differentiated/subjective level is characterized by the recognition that different persons have different perspectives on the situation. However, children understand the perspectives only from one person's point of view. The next level, called the self-reflective/reciprocal level, is the level at which a child is able to mentally step outside the self and take the other person's perspective. Children also understand that people's actions do not have to reflect their thoughts and feelings. Finally, at the third-person/mutual level, the individual is truly able to take a third-person

perspective and understand the situation in terms of mutual goals.

One of the strengths of this model is that it views interpersonal negotiation strategies in an interpersonal context, juxtaposing the understanding of the self in relation to the other. However, the focus remains on the individual, not on the interpersonal process of conflict resolution.

#### **FACTORS RELATED TO CONFLICT RESOLUTION STRATEGIES**

**Age.** There appear to be developmental trends in the use of certain conflict resolution strategies. In the case of aggression as a strategy, the evidence is somewhat contradictory, especially during preschool years. Some findings indicate the use of aggression (Dawe 1934), while others indicate a predominant use of insistence and rare use of aggression (Laursen and Hartup 1989). Overall, the use of power assertion decreases with age and the use of compromise increases with age. The use of compromise, however, may be restricted to hypothetical situations (Laursen and Collins 1994). Thus, adolescents may offer compromise as a solution to hypothetical situations but may report using disengagement as often as compromise (Collins and Laursen 1992). For example, a study of conflict occurring in family talk (Vuchinich 1990) revealed that adolescents are more likely to use standoff (disengagement from the conflict) or withdrawal (physical disengagement from the situation). The use of disengagement may be evident as early as middle childhood (Joshi 1997). Also, the use of compromise may emerge during middle childhood (Joshi 1998). Longitudinal studies are needed to confirm developmental patterns evinced by cross-sectional studies.

**Gender.** The clearest gender differences are in the use of aggression, with girls being less likely to use physical aggression than boys. During middle childhood, girls may be more likely to end conflicts with friends with disengagement (Joshi 1997). Using the INS framework, an investigation of adolescents reporting hard-drug use suggested that boys are more likely to use other-transforming strategies, whereas girls use both self-transforming as well as other-transforming strategies (Leadbeater et al. 1989). Overall there appear to be no clear gender differences other than in the use of aggression.

**The Relationship.** The relationship as the context of the conflict is another factor that influences how the conflict is resolved. Relationships that are defined by kinship norms may be considered by individuals as a “safer” context in which to disagree and argue, since there is no immediate threat of dissolution of the relationships (Berscheid 1985; Laursen 1993). Thus, adolescents may be more likely to use compromise with friends than with adult family members (Laursen and Collins 1994). Another dimension along which relationships vary is the distribution of power. Peer relationships are more egalitarian, while a hierarchy typifies parent–child relationships (Maccoby and Martin 1983). We could, therefore, hypothesize that children are more likely to use negotiation and compromise with peers than with parents. How does the resolution of conflict differ between peers and between parents and children?

#### CONFLICT RESOLUTION WITH PEERS

Research on conflict resolution with peers reveals some general patterns. In one of the earliest observational studies on conflict resolution, Dawe (1934) found that in naturally occurring conflicts in 2- to 5-year-olds, the use of aggression increased with age. Aggression was expressed in behaviors such as pushing, pulling, and striking. Talking during conflict also increased with age. Children resolved most conflicts themselves. When preschoolers’ conversations were examined, Eisenberg and Garvey (1981) found that children used verbal strategies such as insistence and repetition, reasoning, and asking for explanations. Subsequent observations of preschoolers (Killen and Turiel 1991) indicated that in a natural setting as well as a lab setting, children ended more conflicts without active resolutions. Laursen and Hartup (1989) found somewhat similar results from their observations of preschoolers. Insistence was the most frequently used strategy, negotiation and aggression were rarely used, and most conflicts were resolved without adult intervention.

At an older age (7 years), children report using predominantly direct strategies (50 percent), such as leaving the scene of interaction, using physical force, or getting help from a third person (Shantz 1993). About 28 percent of the strategies reported were conventional, such as saying “please” or

apologizing, and about 20 percent of the strategies used involved reasoning and/or compromise. As mentioned earlier, the use of compromise increases during adolescence but may be restricted to hypothetical situations.

#### CONFLICT BETWEEN FRIENDS AND NONFRIENDS

Children, at least until middle childhood, may not be able to understand that friendships can be both supportive and conflictual (Berndt and Perry 1986). Conflict between friends and nonfriends does appear to proceed differently, especially the strategies used to resolve the conflict. An observational study (Hartup et al. 1988) of young children (approximately 3.5 years to 5.5 years of age) revealed that the frequency with which children used disengagement strategies, such as mutual turning away, was greater with friends than with nonfriends. This resulted in different outcomes, with equality occurring with greater frequency with friends than with nonfriends. Also, conflict with friends was more intense than conflict with nonfriends.

In another study, Hartup and colleagues (1993) found that 9- and 10-year-olds showed differences in management of conflict with friends and nonfriends in a closed-field situation, that is, “a situation in which the children cannot choose with whom, what and where their interaction will occur and how long their interaction will last” (p. 446). Children in a dyad were taught conflicting versions of a board game. In this situation, friends disagreed more frequently than nonfriends, and their disagreements lasted longer. Conflict resolution per se was not a focus of the study.

Regarding conflict resolution strategies specifically, young adolescents aged 11 to 14 years showed a greater skill at generating alternatives for resolving conflicts by reacting to hypothetical situations with friends than with nonfriends (Caplan et al. 1991). Similarly, in another study investigating differences in conflict resolution strategies used with friends and acquaintances, Vespo and Caplan (1993) found that children were more likely to use conciliatory gestures with friends than with acquaintances.

At least three reasons can contribute to differences in young adolescents’ conflict resolution skills with friends and nonfriends (Caplan et al.

1991). First, friendship is characterized by intimacy and tolerance. Second, children learn and practice conflict resolution in the context of friendship. Third, friendships shape young adolescents' understanding of conflict and expectations regarding outcome (p. 105).

Thus, no clear patterns emerge regarding developmental changes in conflict resolution strategies, except for a decline in the use of aggression. Perhaps individuals use a range of strategies depending on contextual factors, such as the issue of the conflict. Whether this is true is an important question that needs to be answered. Research has indicated that children who are able to resolve conflicts skillfully are more likely to be socially better adjusted (Asher et al. 1982). Theoretically, the use of compromise is considered both desirable and developmentally more sophisticated. How is it, then, that compromise is not one of the strategies frequently used until adolescence? Skillful conflict resolution may be more a matter of the using a strategy more appropriate for a particular conflict than of using a particular conflict resolution strategy.

#### **CONFLICT RESOLUTION WITHIN THE FAMILY**

The context of family relationships is a powerful influence in the development of social skills. Interactions with different family members foster the development of the ability to understand others' feelings (Dunn 1988). During the second year of life, children begin to get into conflict with their parents and siblings with increased intensity and frequency (Dunn and Slomkowski 1992). During the elementary school years, children are most likely to use assertion or insistence at the beginning of a conflict and end the conflict with submission (Joshi 1997). As noted earlier, adolescents may use standoff or withdrawal in their conflicts with parents and siblings (Vuchinich 1990). There is a need for replication of these findings to trace developmental trends, if any, in the use of conflict resolution strategies with family members.

While individuals might use different strategies to resolve conflicts in different relationships, it could be argued that conflict resolution skills learned in the family may transfer to other contexts. Every relationship is embedded in a network of relationships, and thus it must affect and be

affected by other relationships (Hartup and Laursen 1991; Hinde 1981; Lewis et al. 1984). How does conflict resolution in one context relate to conflict resolution in another?

#### **LINKAGES IN CONFLICT RESOLUTION BETWEEN CONTEXTS**

Studies indicate that patterns of interactions with parents serve as models for children that they can use with peers (Parke et al. 1992). The level of conflict in the family was found to be predictive of adjustment in school (Tesser et al. 1989). Most of the research related to conflict resolution between parents and children uses parental disciplinary style as the indicator of parents' conflict resolution strategies with children. Also, these studies are conducted within the social information processing framework. Overall, the findings indicate that children who experience power-assertive styles of discipline are likely to misinterpret social cues and use ineffective conflict resolution strategies (Dodge et al. 1990; Hart et al. 1990; Weiss et al. 1992). However, these studies examine conflict resolution at a more general level, with some emphasis on aggression. When the correspondence between the conflict resolution strategies parents use with children and strategies that children use with peers is examined, some interesting findings are revealed (Joshi 1997). First, parents use different strategies depending on the issue of the conflict. Second, both children and parents use more than one strategy to resolve conflict. For example, a child may use assertion to begin with and then resort to reasoning, and if the conflict does not get resolved, eventually withdraw from the conflict. The use of multiple strategies makes the one-on-one mapping of conflict resolution strategies used in two contexts a difficult task. The one correspondence that was found was between the last strategy that mothers used with children and the last strategy that children used with peers. Specifically, children who used reasoning as the last strategy with a friend had mothers who used reasoning as a last strategy with them (Joshi 1997).

#### **FUTURE DIRECTIONS IN RESEARCH AND THEORY**

One of the features of the research and theory in the area of interpersonal conflict resolution, at



least during childhood and adolescence, is that it focuses more on intraindividual processes than it does on interindividual processes. We still need more research on how individuals involved in a conflict situation influence each other's behavior. The risk of focusing on the individual lies in its inevitable outcome—that the strategy that the person uses comes to be denoted as the person's attribute, which is a stable characteristic. If we know that individuals may use different strategies in different conflict situations, and that they use a combination of strategies, then it would be erroneous to describe conflict resolution styles in terms of single strategies. Thus, the important variables that need to be examined along with conflict resolution strategies are the issue of the conflict, the relationship between the individuals, and the strategies of the other individual.

The conceptualization of interpersonal conflict resolution as a process will be better accomplished by measurement that reflects process. Therefore, interpersonal interactions must be characterized by the assessment of the behavior of one person as contingent on the behavior of the other person. This would require the development of newer models that would predict individual behavior in conflict based on the aforementioned contextual variables.

Besides achieving the goal of depicting conflict resolution with greater accuracy, the study of contextual variables will help us better understand what exactly constitutes general conflict resolution skills. This knowledge will be particularly useful in planning interventions for individuals who find interpersonal conflict situations challenging.

REFERENCES

- Asher, S. R., P. D. Renshaw, and S. Heymel 1982 "Peer Relations and the Development of Social Skills." In S. G. Moore, ed., *The Young Child: Reviews of Research*, vol. 3. Washington, D.C.: National Association for the Education of Young Children.
- Berndt, T. J., and T. B. Perry 1986 "Children's Perceptions of Friendships as Supportive Relationships." *Developmental Psychology* 22(5):640-648.
- Berscheid, E. 1985 "Interpersonal Attraction." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*. New York: Random House.
- Caplan, M., L. Bennetto, and R. P. Weissberg (1991). "The Role of Interpersonal Context in the Assessment of Social Problem-Solving Skills." *Journal of Applied Developmental Psychology* 12:103-114.
- Collins, W. A., and B. Laursen 1992 "Conflict and Relationships During Adolescence." In C. U. Shantz and W. W. Hartup, eds., *Conflict in Child and Adolescent Development*. New York: Cambridge University Press.
- Crick, N. R., and K. A. Dodge 1994 "A Review and Reformulation of Social Information-Processing Mechanisms in Children's Social Adjustment." *Psychological Bulletin* 115(1):74-101.
- Dawe, H. C. 1934 "An Analysis of Two Hundred Quarrels of Preschool Children." *Child Development* 5:139-157.
- Dodge, K. A., J. E. Bates, and G. Pettit 1990 "Mechanisms in the Cycle of Violence." *Science* 250(4988):1678-1683.
- Dodge, K. A., G. S. Pettit, C. I. McClaskey, and M. Brown 1986 "Social Competence in Children." *Monographs of the Society for Research in Child Development*, 51 (2, serial no. 213).
- Dunn, J. 1988 *The Beginnings of Social Understanding*. Cambridge, Mass.: Harvard University Press.
- , and C. Slomkowski 1992 "Conflict and the Development of Social Understanding." In C. U. Shantz and W. W. Hartup, eds., *Conflict in Child and Adolescent Development*. New York: Cambridge University Press.
- Eisenberg, A. R., and C. Garvey 1981 "Children's Use of Verbal Strategies in Resolving Conflicts." *Discourse Processes* 4:149-170.
- Garvey, C., and Shantz, C. U. 1992 "Conflict Talk: Approaches to Adversative Discourse." In C. U. Shantz and W. W. Hartup, eds., *Conflict in Child and Adolescent Development*. Cambridge, U.K.: Cambridge University Press.
- Hart, C. H., G. W. Ladd, and B. R. Bursleson 1990 "Children's Expectations of the Outcomes of Social Strategies: Relations with Sociometric Status and Maternal Disciplinary Styles." *Child Development* 61:127-137.
- Hartup, W. W., D. C. French, B. Laursen, M. K. Johnston, and J. R. Ogawa 1993 "Conflict and Friendship Relations in Middle Childhood: Behavior in a Closed-Field Situation." *Child Development* 64:445-454.
- Hartup, W. W., and B. Laursen 1991. "Relationships as Developmental Contexts." In R. Cohen and A. W. Siegel, eds., *Context and Development*. Hillsdale, N.J.: Lawrence Erlbaum.

- Hartup, W. W., B. Laursen, M. I. Stewart, and A. Eastenson 1988 "Conflict and the Friendship Relations of Young Children." *Child Development* 59:1590-1600.
- Hinde, R. 1981 "The Bases of a Science of Interpersonal Relationships." In S. Duck and R. Gilmour, eds., *Personal Relationships: Studying Interpersonal Relationships*. London: Academic Press.
- Joshi, A. 1997 *Children's Conflict Resolution Strategies with Friends: Relations to Parent-Child and Interparental Conflict Resolution Strategies*. Unpublished doctoral dissertation, Purdue University, West Lafayette, Ind.
- 1998 "On the Brink? Some Initial Findings on the Beginnings of the Use of Compromise in Interpersonal Conflict." Poster presented at the meeting of the Society for Research on Adolescence, San Diego.
- Killen, M., and E. Turiel 1991 "Conflict Resolution in Preschool Social Interactions." *Early Education and Development* 2(3):240-255.
- Laursen, B. 1993 "The Perceived Impact of Conflict on Adolescent Relationships." *Merrill-Palmer Quarterly* 39(4):535-550.
- , and W. A. Collins 1994 "Interpersonal Conflict During Adolescence." *Psychological Bulletin* 115(2):197-209.
- , and W. W. Hartup 1989 "The Dynamics of Preschool Children's Conflicts." *Merrill-Palmer Quarterly* 35(3):281-297.
- Leadbeater, B. J., I. Hellner, J. P. Allen, and J. L. Aber 1989 "Assessment of Interpersonal Negotiation Strategies in Youth Engaged in Problem Behaviors." *Developmental Psychology* 25(3):465-472.
- Lewis, M., C. Feiring, and M. Kotsonis 1984 "The Social Network of the Young Child: A Developmental Perspective." In M. Lewis, ed., *Beyond the Dyad*. New York: Plenum Press.
- Maccoby, E. E., and J. A. Martin 1983 "Socialization in the Context of the Family: Parent-Child Interaction. In P. H. Mussen (series ed.) and E. M. Hetherington (vol. ed.), *Handbook of Child Psychology: Vol. 4. Socialization, Personality, and Social Development*. New York: Wiley.
- Parke, R. D., J. Cassidy, V. M. Burks, J. L. Carson, and L. Boyum, 1992 "Familial Contribution to Peer Competence Among Young Children: The Role of Interactive and Affective Processes." In R. D. Parke and G. W. Ladd, eds., *Family-Peer Relationships: Modes of Linkages*. Hillsdale, N.J.: Lawrence Erlbaum.
- Selman, R. L., and A. P. Demorest 1984 "Observing Troubled Children's Interpersonal Negotiation Strategies: Implications of and for a Developmental Model." *Child Development* 55:288-304.
- Shantz, C. U. 1987 "Conflicts Between children." *Child Development* 58:283-305.
- 1993 "Children's Conflicts: Representations and Lessons Learned." In R. R. Cocking and K. A. Renninger, eds., *The Development and Meaning of Psychological Distance*. Hillsdale, N.J.: Lawrence Erlbaum.
- (1996). "Introduction." *Merrill-Palmer Quarterly* 42(1):iii-vii.
- Sprey, J. 1979 "Conflict Theory and the Study of Marriage and the Family." In W. R. Burr, R. Hill, F. I. Nye, and I. L. Reiss, eds., *Contemporary Theories About the Family*, vol. 2. New York: Free Press.
- Straus, M. A. 1979 "Measuring Intrafamily Conflict and Violence: The Conflict Tactics (CT) Scales." *Journal of Marriage and the Family* 41:75-88.
- Tesser, A., R. Forehand, G. Brody, and N. Long 1989 "Conflict: The Role of Calm and Angry Parent-Child Discussion in Adolescent Adjustment." *Journal of Social and Clinical Psychology* 8(3):317-330.
- Vespo, J. E., and M. Caplan 1993 "Preschoolers' differential Conflict Behavior with Friends and Acquaintances." *Early Education and Development* 4(1):45-53.
- Vuchinich, S. 1990 "The Sequential Organization of Closing in verbal Family Conflict." In A.D. Grimshaw, ed., *Conflict Talk: Sociolinguistic Investigations of Arguments in Conversations*. Cambridge, U.K.: Cambridge University Press.
- Weiss, B., K. A. Dodge, J. E. Bates, and G. S. Pettit 1992 "Some Consequences of Early Harsh Discipline: Child Aggression and a Maladaptive Social Information Processing Style." *Child Development* 63:1321-1335.
- Yeats, K. W., L. H. Schultz, and R. L. Selman 1990 "Bridging the Gaps in Child-Clinical Assessment: Toward the Application of Social-Cognitive Developmental Theory." *Clinical Psychology Review* 10:567-588.

ANUPAMA JOSHI

## INTERPERSONAL POWER

In its broadest sense, interpersonal power refers to any cause of any change in the behavior of one actor, B, which can be attributed to the effect of another actor, A. It sometimes refers to the capacity to cause such change (Weber [1918] 1968), sometimes to actual use of that capacity (Dahl 1957; Simon 1953) but always to overcoming the "resistance" of B (Weber [1918] 1968), hence causing B to do something B would not otherwise do (Dahl 1957). Interpersonal power is therefore

the power of one individual "over" another as opposed to an individual's power to do something, the capacity of an actor to attain some goal (as in Russell 1938). "Power over" always implies a relation between two actors rather than referring to an attribute of an actor. It is sometimes thought of as "micro" power and contrasts with "power to," which is attributed to collectives (Hawley 1963; Parsons 1963) and is thought of as "systemic" or "macro" power.

Adequate description of a power relation will typically refer to: (1) the *bases* of power (the bases of A's power over B are the resources that are possessed by A that are instrumental to the goals of B); (2) the *means* of power (the ways in which A uses these resources to change the behavior of B); (3) the *strength* of power (the costs to B if B does not comply with demands by A); (4) the *costs* of power (the costs of A of having to exercise power over B); (5) the *amount* of power (the extent to which A is able to get B to do something that B would not otherwise have done); (6) the *scope* of power (the acts with respect to which the amount of A's power over B is greater than zero); and (7) the *domain* of power (the persons over whom the amount of A's power is greater than zero). However, there is a great deal of disagreement over which of these constitute power and what kinds of bases, means, costs, and particularly compliance the word covers.

### CONCEPTS OF POWER

In his work on power, Simon (1953) treats the effects rather than the causes of power and covers by the one word a whole family of concepts describing the effects of all the human causes of human conduct. Force, power, influence, authority, and manipulation all have a strong family resemblance to each other, and Dahl (1957), March (1955), and Simon (1953) have all treated them as one process. But French and Raven (1959) have pointed out important ways in which the dynamics of different kinds of power differ, and March (1966) has shown compellingly that treating them all as one unitary process leads only to a dead end.

The laws of *force*, for example, differ in a fundamental way from the laws governing power. In using force, A does not require B to choose between compliance or noncompliance. A kills B,

imprisons B, drags B, wrestles B to the ground, but does not require choice by B. Threat of force requires a choice, but threat of force is a different process (Goode 1972). A threat may be futile if B has decided to die for a cause; the threat may not accomplish its purpose. Actual force will accomplish its purpose whether B chooses to comply or not. It simply removes B as a factor opposed to A's wishes.

Like force, *manipulation* does not require that B choose between complying or not complying with the wishes of A. A may control B by controlling the information at B's disposal, by preventing certain choices being open to B, or by activating motives of B known to lead B to do what A wants. In some sense, B may be said to make choices (hence the difference from force), but A does not require B to choose between compliance and noncompliance. A controls the conditions that govern how B analyzes the choice to be made.

Although there are more than one narrow senses of the term *power*, all are distinguished from force and manipulation by the fact that power involves a choice by B between compliance and noncompliance. But this is true also of influence and authority. What distinguishes power is that it involves external sanctions. In Blau (1964) B is coerced to do something, X, by threat of a penalty for noncompliance. In Festinger (1953), B is induced to do X by a promise of reward for compliance or coerced to do X by threat of a penalty for noncompliance. Both rewards and penalties are external to the actor. They make no internal change, no change in the actor's state of mind. B does not change views privately, even if B conforms publicly. *Influence*, on the other hand, persuades B that X is right according to B's own interests; hence B complies privately as well as publicly. Compliance is willing in the case of influence, "forced" (Festinger 1953) in the case of power. If compliance were not observable to A, B would still comply in the case of influence but would not in the case of power. Hence, power is highly dependent on observability, but influence is not (French and Raven 1959).

*Authority* differs from both: It refers to a claim by A, accepted by B, that A has a legitimate right to expect compliance by B, even if compliance runs counter to B's own preferences (Barnard 1938). It differs from influence because whether B likes or does not like X is irrelevant. It differs from power

in that B is expected to comply (and, if B accepts A's authority, does comply) because it is right, not because it is expedient. If B accepts A's legitimate authority, then B complies with A's commands whether compliance is observable to A or not. Furthermore, the power that legitimate authority makes possible has different effects because it is legitimate. Power exercised outside the scope of legitimate authority creates "reactance" (Brehm 1966), but power exercised inside the scope of legitimate authority does not (French and Raven 1959).

French and Raven (1959) also distinguish reward from punishment power and hence inducement from coercion. Coercion causes reactance; inducement does not. But they are so indissolubly connected that it becomes difficult to treat them as distinct processes. A reward forgone is a cost, and a penalty forgone is a reward. Hence, withholding a reward is equivalent to imposing a penalty, and withholding a penalty is equivalent to giving a reward. In order to distinguish the two, one would have to be able to separate giving something from withholding something. This might be possible in principle, but not in the case of power. Power depends on the contingency of sanctions; that is, A must be able to give *or* withhold something, depending on whether B does not comply. It is difficult to have one without the other. Thus, probably the most satisfactory concept of interpersonal power is Festinger's forced compliance (1953), compliance that is public but not private, caused either by threat of penalty or promise of reward.

This still leaves the question of whether power is potential or actual. Weber ([1918] 1968) emphasized the capacity of one actor to overcome the resistance of another whether or not the capacity is actually used. Simon (1953) and Dahl (1957) have objected to making inferences from a potential that might or might not actually be used. Research in fact consistently shows that not all potential power is used (see below), and Dahl in particular has argued that power has an effect only if used. On the other hand, Friedrich (1937) has argued that B might comply with the preferences of A because of anticipated reward or punishment, even if A says or does nothing that overtly demands compliance by B. Bachrach and Baratz (1963) go further and argue that certain kinds of acts that involve B doing nothing (nondecisions) occur without B even needing to know what A will

do, simply because B knows what A *could* do (a hypothesis confirmed by Zelditch and Ford, 1994). Both arguments imply that sheer existence of potential power has an effect whether used or not. The effect does not even have to be intended.

The issue of intentionality is another of the disputed questions in conceptualizing power, but not only can potential power have an effect without being used, there are even theories, such as those of Cartwright (1959) and Emerson (1962, 1972), in which "use" of power occurs without necessarily being intended by anyone.

The dispute over potential versus actual power refers more to causes, the dispute over broad versus narrow concepts of power more to effects. Thus, another ambiguity of power is whether one is referring to causes, effects, or the process relating the two.

#### THE CAUSAL APPROACH TO POWER

As process, power most typically refers to an exchange of "resources" (characteristics or objects instrumental to the attainment of goals). The only exception is the "causal" approach to power. Simon (1953) pointed out the similarity between the concept of "power over" and causality, leading to a longstanding tradition in which the former is defined in terms of the latter (as in Dahl 1957; Nagel 1975). But this led to so many difficulties that March (1966) concluded that the concept of power was superfluous. Although Nagel (1975) attempted to give the concept more explanatory power by narrowing it (to the effects of A's preferences on B), this led to little further advance, possibly because causal modeling gives one little insight into the process relating causes to effects.

But "process" theories themselves divide into four types: field theory; rational choice, or decision, theories of exchange; behavioral exchange theory; and a neo-Weberian "resistance" theory of exchange.

#### THE FIELD THEORY OF POWER

In field theory, power is the ability to activate forces in the life space of another in the direction of X (Cartwright 1959). The capacity to activate such forces depends on the motive base of B,

thought of as a tension system, which, when activated by A, produces a vectored force that reduces the tension. Thus, power depends on the needs of B as much as on the wishes of A. It was this approach that gave rise to the concept of a “resource”: Transferable resources give rise to power. But use of power by A faces not only opposing forces leading B away from X but also resistance due to the exercise of power itself. (This resistance was later termed *reactance* by Brehm 1966 to distinguish it from resistance due to B’s dislike of doing X itself.) A’s “control” over B is the outcome of the opposing forces acting on B. Thus, the outcome of the reduction of tension in the force field is distinguished from A’s power in a manner corresponding to the distinction between actual versus potential power. In field theory, therefore, success—that is, actual compliance by B—is not the measure of A’s power since there are many forces other than A affecting B’s behavior.

Field theory, like the causal approach to power, treats the whole family of power concepts as power, excluding only physical force. But research in this tradition was much concerned with differences between different types of power, including reward power, punishment power, expert power, legitimate power, and referent power (power arising from B’s attraction to A, which causes a desire to be similar to A). It was found that both reward and punishment power depend on A’s ability to observe B’s behavior, while the other forms of power do not. Coercion increased but reward decreased reactance. Legitimacy normatively constrained power so that using it outside its legitimate scope reduced its effect. Use of coercion when it is legitimate decreased reactance, while use of rewards when they are not legitimate increased reactance (French and Raven 1959). Because of reactance, use of coercion depends on restraints that prevent B from leaving the field.

**EXCHANGE THEORIES OF POWER**

The most important difference between field theory and exchange theories is that in the latter alternatives become the most important factor in the analysis of power. The reason is that, assuming voluntary relations, an actor may do things that she or he would prefer not to do because they are nevertheless preferable to any available alternative. Thus, an abused wife may not leave her

husband because separation or divorce are either even less desirable or even more costly. But exchange theories themselves come in three somewhat distinct forms: rational choice, or decision, theory; behavioral exchange theory; and resistance theory.

**Rational Choice Theories of Exchange and Power.** Thibaut and Kelley (1959) offer perhaps the earliest rational choice model of the exchange process applied to power. They treat all interaction by analogy to the exchange of goods and services, each actor’s participation in exchange being determined by choice among alternatives, the outcomes of which are characterized by their rewards and costs. Rewards are simply the positive values, costs the negative values, associated with the consequences of choosing an alternative. The value that results from the algebraic sum of rewards and costs is the payoff for a given act. Given interaction between two actors, A and B, payoffs are a joint function of the choices made by each. A matrix of joint payoffs for alternative actions by A and B formulates the conditions determining interaction between the two, in particular whether a relation forms and persists. Thibaut and Kelley treat payoffs in terms of their subjective value for the actors, hence in terms of their “utility” (although they do not use the term). In general, actors are assumed to maximize utility and continue in a relation only if for each actor a course of action is better than the best available alternative.

Thibaut and Kelley deal largely with stable relations rather than with particular acts. Power in a relation arises out of the “dependence” of the actors on each other. It can take either (or both) of two forms: A may be able to control the outcomes of B independently of any act by B, hence controlling B’s fate, or A may be able to make rewards and costs contingent on B’s behavior, hence controlling B’s behavior. But an important principle of the theory is that control of an individual’s fate can be converted into behavioral control.

There are many variants of the decision-theoretic approach to exchange, each differing with respect to how they treat the choice that underlies it. Value can be treated as objective (e.g., in monetary terms) or subjective (e.g., in terms of the meaning of money to the actor). The latter gives rise to a utility theory, such as in Thibaut and Kelley (1959), Blau (1964), and Bacharach and

Lawler (1981). The relation between choice of an alternative and its outcome can be treated deterministically (as is true of all three of the utility theories just cited) or probabilistically. If value is treated objectively and outcomes probabilistically, one has an *expected utility* theory as in Harsanyi (1962). Among probabilistic theories, probability itself can be treated objectively or subjectively: In March (1955) the probabilities are objective, in Nagel (1968) and Tedeschi and colleagues (1973) they are subjective. In the latter case one has a *subjective expected utility* theory. In all these variants of decision theory, the central axiom is that the actor maximizes whatever the theory's criterion of choice.

Despite the variations among these theories, in general they predict that, given a conflict of interest, A is more likely to use power when more is at stake, when the cost of exercising power is less, when the likelihood of compliance is greater, and when there are fewer alternative means of obtaining compliance. The costs of exercising power include the relative depletion of A's stock of resources, the likelihood of retaliation, and the effect of reactance on subsequent relations with B. In general, the evidence supports these predictions, although "use" sometimes refers to whether or not A prefers to impose his or her will on B; sometimes to whether or not, if A does have such a preference, it is accomplished by manipulating rewards and punishments rather than by other means; and sometimes to whether or not A's behavior is overt and explicit, communicating threats or promises, or is covert. Rational choice theories will also in general predict that B is more likely to comply with A's preferences the larger the reward attached to compliance, the larger the penalty attached to noncompliance, the greater the credibility of A's promises of reward or threat of penalty, and the fewer the alternatives available to B. Again, the evidence largely supports these predictions. But it is well to keep in mind that with respect to both use and compliance, the evidence deals largely with voluntary relations.

Unlike other decision theories, subjective expected utility theory (especially Nagel 1968) makes use of Friedrich's (1937) "law of anticipated reactions," in which compliance occurs without directives, threats, or promises by A. In the case of punishment, not only is there no visible exercise of power by A, but compliance by B may appear to be

willing because there is no visible resistance. In addition, if one takes into account that the law of anticipated reactions can be applied to A as well as to B, A may know in advance that B will comply and therefore know that it is unnecessary visibly to exercise power in order to achieve A's objective (Samuel and Zelditch 1989).

Bacharach and Lawler (1981) have pointed out that even though relative power in a relation is necessarily zero-sum, meaning that a gain by one actor implies loss by the other, relative power can be distinguished from absolute power, the sheer quantity of resources controlled by an actor. Hence, relative power can be distinguished from total power, the sum of the resources possessed by A and B. Total power, unlike relative power, is a variable that can increase for both actors at the same time. Bacharach and Lawler argue that total power has an effect independent of relative power, especially on the use of punitive power capabilities. Given a conflict of interest, if punitive power is an available tactic, the likelihood of its use by the more powerful actor increases as his or her relative power increases but decreases as total power increases.

Thibaut and Kelley note that all these ideas are more useful as the pattern of exchange becomes more stable; hence, the focus of exchange theory is often not only on the fact that power is relational, which is true even at the level of a unit act, but on the structure of relations arising from enduring patterns of repeated exchanges. The most influential theory of this kind has been Emerson's theory of power-dependence relations (1962, 1972). Though it originally grew out of Thibaut and Kelley's work, this theory has come to be associated more recently with behavioral exchange theory, a theory originating in the work of Homans.

**Behavioral Exchange Theory.** Homans (1961) behaviorized exchange theory by transforming rewards and costs into reinforcement contingencies and the process of exchange into mutual operant conditioning. Many fundamental concepts of decision theory, however, survive in behavioral exchange theory. Choices are determined by maximizing profit, which is the difference between rewards and costs. Costs in Homans take a form more like costs in economics than in Thibaut and Kelley, becoming opportunity costs, that is, profit forgone by choosing X over not-X. But, as in

Thibaut and Kelley, exchange occurs at a point at which for both actors exchange is more profitable than each actor's best available alternative. Thus, A uses power in Homans—where “use” means attempts to direct the behavior of another—when it is more profitable than any alternative available to A, and B complies when compliance is more profitable than noncompliance. Homans, however, gives more emphasis than do Thibaut and Kelley to the effect of satiation on the value of a reward: Accumulation of rewards brings diminishing returns. Hence, poor actors are more responsive than are rich ones to offers of any given level of reward for doing an otherwise undesired act.

Emerson's power-dependence theory, perhaps the most influential theory of interpersonal power, was originally a utility theory of relations like Thibaut and Kelley's (Emerson 1962). But Emerson (1972) followed Homans in behaviorizing exchange theory, except that, unlike Homans, Emerson again dealt with relations rather than acts. Power-dependence theory is formulated in terms of two actors whose social relations entail ties of mutual dependence. “Power” is the amount of B's resistance that can be overcome by A: It is potential rather than actual power that is used to describe the relation. Power is a function of dependence, which arises from the control by one actor of resources on which another depends for achieving his or her goals. Dependence varies directly with an actor's motivational investment in goals but inversely with the availability of alternative sources of resources outside the relation. The “power advantage” of one actor over another is a function of the net balance of each one's dependence on the other. If the net balance is zero, power is equal, or “balanced.” A relatively unique feature of Emerson's theory is that balanced power is assumed to be stable, while imbalance creates pressures to change the power relation in the direction of balance. There are four kinds of balancing operations. B may (1) reduce motivational investment in the goals mediated by A, withdrawing from the relation; (2) increase the number of alternative sources of the resource (extend networks); (3) increase A's motivational investment in goals mediated by B (e.g., by offering status to A); or (4) deny to A alternative sources of resources mediated by B (coalition formation). Attention to these operations leads one to distinguish use of power from change of power—increasing or decreasing

dependence in the relation. These assumptions about balance have had a major impact on the study of organizations (beginning with Thompson 1967), but other exchange theories have tended to reject the balance assumption, especially Blau (1964), who instead treats the balance mechanisms as means by which A maintains or increases power over B.

“Use” in power-dependence theory comes to mean something quite different from what it means in decision or field-theoretic formulations: It refers now to asymmetries in the outcomes of reciprocal exchange. It is assumed that use of power, in the more usual sense, increases until an equilibrium is reached, at which point the less powerful actor is receiving no more benefits than the best alternative source could provide. While power imbalance does in fact increase asymmetry of exchange, it has been consistently found that its use is suboptimal. It is constrained in part by search costs (Molm 1987s) and in part by commitment to the other and concern for equity (Cook and Emerson 1978). Furthermore, use of punishment power depends not only on punishment capability but also on reward capability: Those weaker in reward power are, other things being equal, more likely to use punishment power (Molm 1997). Molm argues they have more to gain and less to lose by the use of coercion than those with more reward power: It is muggers, who gain nothing at all without coercion, who have nothing to lose by using it.

Molm also finds that, given mutual dependence, the effectiveness of coercion is directly proportional to its use, providing that its use is: (1) instrumental rather than symbolic, that is, a specific consequence of a specific behavior rather than a personal challenge; (2) highly contingent on the undesired behavior; and (3) consistent, that is, the punishment of equivalent behavior is equivalent. Coercion is perceived as more symbolic if it is noncontingent, inconsistent, or threatened but not actually used, increasing the side effects of coercion (resistance, negative affect, retaliation, loss of the relation). If it is perceived as instrumental, on the other hand, a high probability of punishment for a specific behavior, because it is effective in the short run, actually increases the net frequency of rewards in the long run, while “weak”—that is, sporadic—coercion decreases the

long-run net frequency of rewards and increases the side effects of coercion.

The most significant development in the behavioral exchange theory of power has been its extension to networks of dyadic relations. Interpersonal power is not necessarily confined to a dyadic relation, as Coleman (1973) has shown. But except for Coleman, interpersonal power has usually been treated as dyadic. More complex structures are possible, however, by connecting dyads into networks (Emerson 1972). Study of such networks has rapidly developed as a central preoccupation of research on interpersonal power, especially preoccupation with the question of determinants of power at a position within a network. Most of this research has been concerned with negatively connected networks. A negative connection is one such that exchange in one relation decreases the value of exchange in another, while a positive connection increases its value. Thus, if A must choose which of two offers of exchange to accept from two others, B and C, AB is negatively connected to AC because choice of AB excludes exchange between A and C. Subsequent research has concentrated on the effects of varying structures of negatively connected networks. Centrality of a position turns out not to predict asymmetry of exchange very well; hence, much theory and research has been directed at discovering what alternative concept of power of a position does predict asymmetry of exchange. Cook and Emerson (1978) proposed "vulnerability" as an alternative and defined it as the effect that removing a position would have on the total quantity of resources available for exchange in a network. An intuitively appealing concept, this idea works well for many networks, though Markovsky and colleagues (1988) have objected that there are certain structures for which it does not accurately predict asymmetries. Their alternative conception derives from resistance theory.

**Resistance Theory.** Willer's (1981) resistance theory objects to the importance of satiation in behavioral exchange theory and as an alternative goes back to Weber's concept of power as overcoming resistance. Willer conceives of interaction as exchange of sanctions. The value of a sanction is a function of its objective value multiplied by its quantity, but value is not itself a function of quantity; hence, there is no satiation effect. Willer focuses instead on the "preference alteration state" of

the actor ( $P_A$ ), consequent on sanctions, and defines resistance as the ratio of two differences: the best possible outcome minus the value of  $P_A$  is divided by  $P_A$  minus the worst possible outcome B will accept, called  $P_A$  at confrontation. Exchange occurs at equiresistance.

The emphasis on resistance leads the theory to incorporate coercion and conflict more easily within the same framework as voluntary exchange. But in addition it leads to a different understanding of power at a position in a network. As in Cook and Emerson (1978), power at a position depends on alternatives, but Markovsky and colleagues (1988) propose a graph-theoretic index of power based on exclusion—B, bidding for exchange with A, is "excluded" if, having no alternatives, B does not exchange at all if outbid by another. The difference between the two is that in Markovsky and colleagues, exclusion is relative to the number of exchanges sought, while in Cook and Emerson actors seek to make a single exchange at a given point in time. It is "advantageous" in resistance theory to be connected to positions that have few alternatives relative to the number of exchanges sought and disadvantageous to be connected to positions that have many alternatives relative to the number sought. In exclusive networks, an odd number of positions is advantageous because in them a position has alternatives or a partner's alternatives have alternatives. An even number is disadvantageous because a position has one or more rivals for exchange with a partner. The ratio of advantageous to disadvantageous paths leading out from a point in a network (counting overlapping paths only once) gives the graph-theoretic index of power at that point in the network. The logic of the argument can be extended to inclusion as well as to exclusion. An inclusive connection is one such that exchange in one cannot occur until exchange in another also occurs. Inclusive relations, for example, increase the power of peripheral positions in a network. However, in mixes of the two, exclusion overwhelms inclusion and power is as centralized as in exclusive networks (Szmatka and Willer 1995).

Exclusion is more certain in some networks than others, and the magnitude of outcome differentials is sensitive to the certainty of exclusion. Outcome differentials are greater in "strong power" networks, in which the less powerful have no



alternatives if they reject the terms of exchange of the more powerful, than in "weak power" networks, in which no position is assured of excluding another without cost to itself (Markovsky et al. 1993). Markovsky and colleagues argue that, in general, the strength of power is inversely proportional to the density of a network because the greater the density of a network, the more alternative opportunities for exchange. Hence, outcome differentials should be inversely proportional to the density of a network. They also show that weak power networks are more sensitive to strategy. Strong power maximizes outcome differentials even if the powerful are passive. In weak power networks, structural forces keep the division of profits more nearly equal, but strategy makes a difference: The more experienced demand more if they are in powerful positions and yield more if they are in less powerful positions, thus magnifying outcome differentials. Weak power networks are also more sensitive to the actual experience of exclusion (Skvoretz and Willer 1993): Potential exclusion, determined by structure, determines outcomes in strong power networks, but (given full information) outcomes are sensitive to both the potential and actual experience of exclusion in weak power networks.

(SEE ALSO: *Decision-Making Theory and Research; Exchange Theory*)

REFERENCES

- Bacharach, Samuel B., and Edward J. Lawler 1981 *Bargaining: Power, Tactics, and Outcomes*. San Francisco: Jossey-Bass.
- Bachrach, Peter, and Morton S. Baratz 1963 "Decisions and Nondecisions: An Analytical Framework." *American Political Science Review* 57:632-642.
- Barnard, Chester I. 1938 *The Functions of the Executive*. Cambridge, Mass.: Harvard University Press.
- Blau, Peter M. 1964 *Exchange and Power in Social Life*. New York: Wiley.
- Brehm, Jack W. 1966 *A Theory of Psychological Reactance*. New York: Academic Press.
- Cartwright, Dorwin 1959 "A Field Theoretical Conception of Power." In Dorwin Cartwright, ed., *Studies in Social Power*. Ann Arbor: University of Michigan Press.
- Coleman, James S. 1973 *The Mathematics of Collective Action*. Chicago: Aldine.
- Cook, Karen S., and Richard M. Emerson 1978 "Power, Equity, and Commitment in Exchange Networks." *American Sociological Review* 43:721-739.
- Dahl, Robert A. 1957 "The Concept of Power." *Behavioral Science* 2:201-215.
- Emerson, Richard M. 1962 "Power-Dependence Relations." *American Sociological Review* 27:31-41.
- 1972 "Exchange Theory." In Joseph Berger, Morris Zelditch, and Bo Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.
- Festinger, Leon 1953 "An Analysis of Compliant Behavior." In Musafir Sherif and Milbourne O. Wilson, eds., *Group Relations at the Crossroads*. New York: Harper.
- French, John R. P., and Bertram Raven 1959 "The Bases of Social Power." In Dorwin Cartwright, ed., *Studies in Social Power*. Ann Arbor: University of Michigan Press.
- Friedrich, Carl J. 1937 *Constitutional Government and Politics: Nature and Development*. New York: Harper and Bros.
- Goode, William J. 1972 "The Place of Force in Human Society." *American Sociological Review* 37:507-519.
- Harsanyi, John C. 1962 "Measurement of Social Power, Opportunity Costs, and the Theory of Two-Person Bargaining Games." *Behavioral Science* 7:67-80.
- Hawley, Amos 1963 "Community Power and Urban Renewal Success." *American Journal of Sociology* 68:422-431.
- Homans, George C. 1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace Jovanovich.
- March, James G. 1955 "An Introduction to the Theory and Measurement of Influence." *American Political Science Review* 49:431-451.
- 1966 "The Power of Power." In David Easton, ed., *Varieties of Political Theory*. Englewood Cliffs, N. J.: Prentice-Hall.
- Markovsky, B., J. Skvoretz, D. Willer, M. Lovaglia, and J. Erger 1993 "The Seeds of Weak Power: An Extension of Network Exchange Theory." *American Sociological Review* 58:197-209.
- Markovsky, Barry, David Willer, and Travis Patton 1988 "Power Relations in Exchange Networks." *American Sociological Review* 53:220-236.
- Molm, Linda 1987 "Linking Power Structure and Power Use." In Karen S. Cook, ed., *Social Exchange Theory*. Newbury Park, Calif.: Sage.
- 1997. *Coercive Power in Social Exchange*. New York: Cambridge University Press.
- Nagel, Jack H. 1968 "Some Questions About the Concept of Power." *Behavioral Science* 13:129-137.

- 1975 *The Descriptive Analysis of Power*. New Haven, Conn.: Yale University Press.
- Parsons, Talcott 1963 "On the Concept of Political Power." *Proceedings of the American Philosophical Society* 107 (3):232–262.
- Russell, Bertrand R. 1938 *Power: A New Social Analysis*. London: Allen and Unwin.
- Samuel, Yitzhak, and Morris Zelditch, Jr. 1989 "Expectations, Shared Awareness, and Power." In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress: New Formulations*. Newbury Park, Calif.: Sage.
- Simon, Herbert 1953 "Notes on the Observation and Measurement of Political Power." *Journal of Politics* 15:500–516.
- Skvoretz, J., and D. Willer 1993 "Exclusion and Power: A Test of Four Theories of Power in Exchange Networks." *American Sociological Review* 58:801–818.
- Szmatka, J., and D. Willer 1995 "Exclusion, Inclusion, and Compound Connection in Exchange Networks." *Social Psychology Quarterly* 58:123–131
- Tedeschi, James T., Barry R. Schlenker, and Thomas V. Bonoma 1973 *Conflict, Power, and Games*. Chicago: Aldine.
- Thibaut, John W., and Harold H. Kelley 1959 *The Social Psychology of Groups*. New York: Wiley.
- Thompson, James D. 1967 *Organizations in Action*. New York: McGraw-Hill.
- Weber, Max (1918) 1968 *Economy and Society*. Guenther Roth and Claus Wittich, eds.; trans. Ephraim Fischhoff et al. New York: Bedminster Press.
- Willer, David 1981 "Quantity and Network Structure." In David Willer and Bo Anderson, eds., *Networks, Exchange, and Coercion*. New York: Elsevier/Greenwood.
- Zelditch, M., and Ford, J. 1994 "Uncertainty, Potential Power, and Nondecisions." *Social Psychology Quarterly* 57:64–73.

MORRIS ZELDITCH, JR.

## ITALIAN SOCIOLOGY

The birth of sociology in Italy is variously dated, depending on the causes that are adduced for it or on the perspective from which it is viewed. Indeed, even the conceptualization of sociology differs; hence, in a certain sense it is more important to determine when *each* "sociology" was born. We can distinguish two types of sociology and two

corresponding ways of providing cultural and professional training. The first type is positivist (or neopositivist) sociology, tied to quantitative empirical research, which aims to discover the laws and the causal relationships that can be drawn from the data and from experience; the second type is humanistic sociology, which interprets its role as a critical science, raises questions, is more an approach to science than a science proper, and places social phenomena in their historical context. The two places for training professional sociologists are the research center and the university, respectively; each type of institution has an ambivalent and fluctuating relationship with the two sociologies.

Both types of sociology and both kinds of training, together with their origins, take on specific meanings according to the historical period; and hence they depend on the process of complexification of Italian society and the parallel development of different Italies (at least the three indicated by Bagnasco 1977).

### THE ORIGINS OF ITALIAN SOCIOLOGY

Many historians of sociology see Italian sociology as deriving from the political thought of Niccolò Machiavelli (1469–1527), because of his interest in leadership and its connection with the structures within which the prince must exercise his will. The political sociology of Vilfredo Pareto, Gaetano Mosca, Scipio Sighele, Roberto Michels, and Camillo Pellizzi can also be seen as belonging to this current of thought. This school, largely academic in origin, is based in political science, the philosophy of politics, and the philosophy of law. Associated with and grafted onto it is a sociology of law, also rooted in the universities, reflected by the great number of articles published in the first issues of *Quaderni di sociologia* (*Notebooks of Sociology*), a review founded in 1951. Philosophy (Roberto Ardigò), political science (Gaetano Mosca), and law (Carlo Francesco Gabba, Rodolfo Laschi, Enrico Ferri, Icilio Vanni) became "godfathers" to Italian sociology. This may be explained by its dominant themes: the legitimation, explanation, and order that accrue to leadership and political structures as they develop. There were also economists, often (though not always) with a socialist background and orientation—Ginseppe Toniolo, Achille Loria,

and Pareto—always attempting to study society in macroscopic terms in a way closely linked to political science.

Besides these currents concerned with the sphere of public action, there was the study of the private sphere as it deviated from the established order, as examined from the perspectives of political science, philosophy, and economics. Cesare Lombroso, Enrico Morselli, Scipio Sighele, and Alfredo Niceforo—some of them with medical training—studied the criminal personality, constructed typologies of the “delinquent man,” elaborated pseudobiological explanations of deviant behavior, and thus attempted to interpret the relationship between society and the deviant (the duty of society being to lock up and prevent the deviant from causing further harm).

The sociological bent of these Italian protosociologists, who were generally university professors, was rooted in the work of Auguste Comte and Herbert Spencer. Hence, in their work we find a search for the laws underlying the social phenomena studied, the confidence that one can understand these phenomena by means of an inductive method borrowed from the natural sciences, and the belief that it is possible to describe and explain social reality with the positivist method. Italian sociological positivism (from 1865 to the rise of Fascism) worked out a paradigm of social analysis that was fairly homogeneous and characterized by naturalistic determinism in the form of evolutionary organicism. It was carried out by means of a positivist inductive method, although the basis of determinism varied during the history of positivism. For Lombroso, Niceforo, and others, the bases of such determinism were in biology and evolution (the theories of instinct and atavism), while for Pareto the structure of social activity took its model from economics. Another distinctive characteristic of this positivism was to consider society, social phenomena, and subjects in normative but exclusively objective terms. Roberto Ardigò worked along these lines as a theorist of human action, but after a certain point he became increasingly aware of the importance of “nonlogical” actions, of interiorization, and of socialization—the possibility of a voluntaristic theory. It is thus possible, in his case, to speak of a positivism of the subject, which foreshadows the work of Pareto.

The positivist approach was used by Socialist and Catholic scholars, as well as by liberals. This perspective was shared by the founders of *Rivista italiana di sociologia* (*Italian Review of Sociology*), which was published from 1897 to 1922.

Positivism, mainly a French and English current of thought, began to decline when its naturalistic–determinist presuppositions were left behind and other variables had been introduced, thus leading to interpretations of society in more complex terms (Mosca, Pareto, Michels, Gaetano Salvemini); also, it was questioned when a progressive and optimistic determinism was replaced by a pessimistic determinism (as in Michels, with his iron law of oligarchy). Positivism was reborn with the formation of the Chicago School of urban and ecological studies at the end of the 1930s.

In Italy, positivism was replaced by the idealism of Benedetto Croce and Giovanni Gentile, who were strongly critical of positivism’s main ideas and naturalistic research methods. The polemical debate in 1900–1901 between Pareto and Croce in the pages of *Il giornale degli economisti* (*The Journal of Economists*) is emblematic.

The technologically hard core of positivism is the research method, particularly “the experimental method that has given such brilliant results in the natural sciences,” as Pareto wrote in his “Discorso per il Giubileo” (Borgatta 1917; “Discourse for the Jubilee”).

This explains why, among the positivist sociologists, there were many demographers and statisticians (including Angelo Messedaglia, Loria, Niceforo, and Livio Livi) during the years of Fascism (1920–1940) and why the preeminence of idealism—the point of reference for sociology—lay primarily in statisticians such as Corrado Gini, Vittorio Castellano, Marcello Boldrini, and Nora Federici.

Gini, in particular, was active in founding the Istituto Centrale di Statistica (Central Institute of Statistics) in 1926; this agency was set up to organize census taking in Italy. In 1928 he founded the Comitato Italiano per lo Studio dei Problemi della Popolazione (Italian Committee for the Study of Population Problems), the school of statistics at the University of Rome (1929), the Italian section of the Institut International de Sociologie (1932), the Società Italiana di Sociologia (Italian Society of

Sociology) in 1937, and the Facoltà di Scienze Statistiche, Demografiche e Attuariali (Faculty of Statistical, Demographic, and Actuarial Sciences) at the University of Rome (1936). In this period Gini also founded, or was a founding member of, many journals, such as the *Bollettino bibliografico di scienze sociali e politiche* (1924; *Bibliographical Bulletin of Social and Political Sciences*), *La vita economica italiana* (1926; *Italian Economic Life*), and *Gems* (1934).

To sum up, Italian sociology was born through the importation of positivist ideas, paradigms, and methodological creeds, and it tried to solve the problems that society or the classes in power considered important in the second half of the nineteenth century and the first half of the twentieth—the existing forms of power and their legitimation, as well as the control of deviance. Its content therefore concerned the philosophy of law, political doctrines, socialism, the social thought of the Catholic Church, criminality and alienation, and related matters. In the first half of the twentieth century, in particular during the Fascist period, the study of demographics and statistics remained alive, as did the study of the charismatic components and fatal distortions of power (Michels) and investigation of the irrational and subjective element in society, in the form of residues and derivations (Pareto).

The themes and problems of concern were closely linked to the period in which Italian sociology developed. Socially, Italy was still traditional, and sociology was developed within the universities, in the context of more or less formally recognized courses.

#### 1945 TO THE 1960S

The postwar period witnessed a profound rupture in Italian society, in its culture, and hence in its sociology. That society became much more complex, and its emerging culture became part of sociological study. Thus problems such as those in the South of Italy—urbanization, migration, the large concentrations of workers—were brought to the forefront. In addition, Italian sociologists were very receptive to foreign ideas and schools of thought. Thus, alongside the survival in academic circles of the influence of Comte, Spencer, and Darwin and the work of Gini, Livi, and Castellano, the American influence became dominant and the

German tradition regained importance, primarily in research institutes.

In other words, there was first an accumulation of theory and research that was incorporated into university sociology courses; later came the first chair of sociology, which was awarded to Franco Ferrarotti, at Rome in 1962. Scholars did not yet have sociological training, since no academic structures for that purpose existed; therefore they came from fields such as classical studies, political science (still very much oriented toward law), economics and commerce, medicine, law, and so forth.

One effect of this complexity was the subdivision of sociology into many branches: development and modernization; urban and rural sociology; the sociology of labor, of the economy, of migrations.

This period lasted from the end of World War II to the founding of the first faculty of sociology at Trento and the reform of the faculties of political sciences; it saw the incubation of a new direction in Italian sociology. Since that time there has been the development of the faculty of sociology at the University of Rome, and many other universities have acquired significant faculty strength in sociology. Let us now examine these currents.

**The Sociology of Development and the Mezzogiorno.** One of the problems the new Italy had to face was how to reduce the developmental gap between North and South, in particular how to enable the society of the South and of Sicily and Sardinia to overcome their state of underdevelopment. Even before World War II the “Southern question” had been posed and announced by the actions and writings of Guido Dorso and Salvemini, among others. But it was only after 1946 that the problem became a subject of study and institutional intervention. Some of the institutions that deal with intervention and training, as well as with sociological research on the problems of the Mezzogiorno, are the Associazione per lo Sviluppo dell’Industria nel Mezzogiorno (SVIMEZ; Association for the Development of Industry in the Mezzogiorno), the CENTRO per la Formazione e studi per il Mezzogiorno (Center for Professional Training and Studies of the Mezzogiorno), the Associazione Nazionale per gli Interessi del Mezzogiorno (ANANI; National Association for the Interests of the

Mezzogiorno), the Movimento di Collaborazione Civica (Movement of Civic Collaboration), the Unione Nazionale per la Lotta contro l'analfabetismo (National Union for the Fight Against Illiteracy), and UNRRA-Casas (Istituto per lo Sviluppo dell'edilizia Sociale; the Institute for the Development of Public Housing).

Projects and plans were drawn up within and in cooperation with these agencies. They include the pilot project for the Abruzzi, the Sardinia project, the Center for Studies and Initiatives in western Sicily, the Center for Community Development at Palma di Montechiaro, and the Molise and Avigliano projects. It was in these institutions that sociological and anthropological studies on development, modernization, and the community were carried out. Tullio Tentori (1956), Guido Vincelli (1958), Danilo Dolci (1955, 1957, 1960), Gilberto Antonio Marselli (1963), Lidia de Rita (1964), Maria Ricciardi Ruocco (1967), Luca Pinna (1971), Gualtiero Harrison and Maria Callari Galli (1971), and Giovanni Mottura and Enrico Pugliese (1975) are among those whose works reveal a social commitment together with testing of the theories of modernization. They show an awareness of their limits as well, as elaborated by the English-language sociological tradition; moreover, there was a strong element of utopianism in this mixture of social commitment and scientific rigor.

It was not just the ideas of American and northern European scholars that were studied and applied to the modernization of the South of Italy; scholars themselves were committed to the task. At the same time, some classic research was carried out, important both because it was cited by Italian scholars and because it was rejected by them (in particular by Alessandro Pizzorno and Marselli). One of the books that became an object of controversy was Edward C. Banfield's *The Moral Basis of a Backward Society* (1958). Other frequently cited foreign research works included Joseph Lopreato's "Social Stratification and Mobility in a South Italian Town" (1961), Feliks Gross's "Value Structure and Social Change" (1970), and Johann Galtung's *Members of Two Worlds: A Development Study of Three Villages in Western Sicily* (1971).

Sociological study of modernization also entered the universities through sociology courses that took their place alongside courses in rural economics. The most important center for these

courses was the Faculty of Agriculture at Portici, and the most significant sociologists were Marselli, Pugliese, Mottura, and Emanuele Sgroi.

**The Sociology of Work and Economics.** The sociology of work and economics arose in the research agencies, in research centers of large industries, and in labor unions.

One very important research center for the training of Italian sociologists was Olive Hi's Ufficio Studi e Relazioni Sociali (Research and Social Relations Office) at Ivrea. Adriano Olivetti, building a company and a community, surrounded himself with sociologists, economists, communications experts, and planners. Ferrarotti, Pizzorno, Luciano Gallino, Paolo Ceri, and Antonio Carbonaro worked and studied at the facility. They, and many more, devoted themselves to industrial relations and the rationalization of staff selection in large industries operated by enlightened and paternalistic entrepreneurs like Olivetti. Very often, however, the left-wing slant of their training and their impression of being manipulated led these sociologists to leave the research center after a short stay and to continue their sociological training at universities in the United States. In this cultural climate, which also was indebted to nineteenth-century models of integration of the company, the community, and territorial planning, the review *Comunità* (Community) published essays with a strong cultural and social commitment. The publishing house of the same name made available to the Italian public the classics of German and, especially, American sociology.

Another forum for the training of sociologists (especially in the sociology of labor) was the research offices of the labor unions (and, to a lesser extent, of the political parties). In particular, future sociologists such as Aris Accornero, Guido Baglioni, Gian Primo Cella, and Guido Romagnoli were directed toward studies of factory work. The factory and labor union conflicts were the main subjects of their volume, with a consequent tendency to identify the "organization" with the factory (a pamphlet on the sociology of organizations by G. Bonazzi is, significantly, entitled *Dentro e fuori la fabbrica* (1982); *Inside and Outside the Factory*) and the labor union. These sociologists had few contacts with foreign scholars; when they did, such contacts were oriented toward France and particularly toward the Institut des Sciences Sociales du

Travail (Institute of the Social Science of Labor) in Paris.

**Urban Sociology and the Sociology of Planning.** Other institutions in which sociologists (this time urban sociologists) were trained in the 1950s and 1960s were the planning offices set up by the territorial governments (large municipalities and provinces in northern Italy), particularly the Istituto Lombardo di Studi Economici e Sociali (ILSES; Lombard Institute of Economic and Social Studies) in Milan and the Centro Studi Sociali e Amministrativi (CSSA; Center for Social and Administrative Studies) in Bologna. The former trained the sociologists Pizzorno, Gianni Pellicciari, Massimo Paci, Alberto Martinelli, Guido Martinotti, Paolo Guidicini; the latter, Achille Ardigò, Paolo Guidicini, Giuliano Piazzi, and Pietro Bellasi. There was also the Ufficio Studi Sociali e del Lavoro (Office for Social and Labor Studies) in Genoa, where Luciano Cavalli carried out his *Inchiesta sugli abituri* (1957; *Survey on Slums*).

ILSES dealt with research on neighborhoods, on participation in neighborhoods, and, in general, on the structure of the city. The approaches were borrowed from the concepts and research of the Chicago School and the group that had formed around Paul Henry Chombart de Lauwe. Thus, among other things, Ernest W. Burgess's model of concentric areas was verified for Italian cities (in Rome and Milan), as was Hoyt's model of sectors. The results of these researches found expression in anthologies on the Chicago School: such as Guido Martinotti's *Città e analisi sociologica* (1968; *The City and Sociological Analysis*); manuals of urban sociology, such as Franco Demarchi's *Società e spazio* (1969; *Society and Space*); and a manual of urban research based largely on that carried out by ILSES and CSSA, Paolo Guidicini's *Manuale della ricerca sociologica* (1968, 1987; *Manual of Sociological Research*).

The sociologists who worked at ILSES developed close contacts with the United States (particularly Martinotti, Paci, and Martinelli) and with the United Kingdom (Guidicini).

**International Sociology.** Sociology entered the universities where both theoretical and empirical sociologists were trained but no attention was given to international sociology. International sociology in Italy was developed by the Istituto di

Sociologia Internazionale di Gorizia (ISIG; Institute of International Sociology of Gorizia), founded by Demarchi in 1969. This branch of sociology seeks its identity in the synthesis of ideas from political sociology; the sociology of international relations; and the sociology of ethnic relations, of borders, of towns, and of territories. Relations were developed with scholars from the United States, from Eastern Europe, and from the countries of the European Economic Community. From the United States, the institute took the research methodology of multivariate analysis, under the direction of Edgar F. Borgatta; it was one of the first institutions in Italy to adopt these techniques. Its researchers later held important positions in various universities: Renzo Gubert, Alberto Gasparini, Raimondo Strassoldo, Bruno Tellia, Bernardo Cattarinussi, and Giovanni Delli Zotti.

Among the cited research organizations, ISIG is one of the few that carries out an active program that is not encompassed by university activities. There are at least two reasons for this. First, Italian sociology has tended not to be interested in international relations. Second, there has been an acceleration of international interdependence; in recent years there has been great change in the relations between Western and Eastern Europe, and between Europe and the rest of the world. Under the direction of Alberto Gasparini, ISIG has responded to the need for such study, with an emphasis on Eastern Europe and the countries of the former USSR.

**Academic Institutionalization of Italian Sociology.** The most intense activity of the research agencies occurred in the 1950s and 1960s; it was scientific in the full sense of the word, since concrete problems were tackled empirically, starting from a theory (generally developed abroad) and ultimately returning to the theory. This way of doing research, and of training for research, is quite different from the method adopted in the universities, because it is tied to concrete problems and specific deadlines.

Moreover, in this period, together with the ideas and plans designed to establish sociology in the universities (the first chair in sociology was awarded to Ferrarotti in 1962), sociology was taking on an identity as an academic discipline. The first issue of *Quaderni di sociologia* (*Notebooks of Sociology*) was published in 1951, edited by Nicola

Abbagnano and Ferrarotti, and in 1957 the Associazione Italiana di Scienze Sociali (AISS; Italian Association of Social Sciences) was founded. In 1959 the Centro Nazionale di Prevenzione e Difesa Sociale (National Center for Prevention and Social Defense) in Milan, together with AISS, organized the Fourth World Congress of Sociology at Stresa; and from then until 1974 Italian sociologists (Angelo Pagani and Guido Martinotti) held the post of secretary of the International Sociological Association. Italian sociology was now mature and ready to enter the universities with its own sociologists, their empirical experience, and their theoretical preparation; hence it was able to extend the discipline.

### SOCIOLOGY IN THE UNIVERSITIES (SINCE 1960)

Italian sociology, as a discipline and as scientific research, has developed strongly. The heart of this development lies more in the universities and university teachers than in the nonacademic institutes that played such a large role until the 1960s. All this happened with the consolidation of sociology in the university system (which had the function of training young researchers and became the channel for legitimizing the scientific character of sociological research). Subsequent developments included the publication of dictionaries and encyclopedias of sociology, the foundation of the Associazione Italiana di Sociologia (Italian Association of Sociology), and the formation and coordination of sociological interests in specific areas.

**Sociology in the Italian Universities.** The complexity of Italian society requires a strong sociological reading of reality, and the nonacademic agencies discussed above could not long meet this need. Moreover, a strong impulse to legitimize sociological analysis came from the student protest movement, which in the mid-1960s reached Europe from the United States.

Thus sociology entered the universities mainly with the establishment in Trento of the Istituto Universitario di Scienze Sociali (University Institute of Social Sciences, later the Faculty of Sociology) in 1962 and with the reform of political science faculties in 1968.

**The Faculty of Sociology at Trento.** The Istituto Universitario di Scienze Sociali was officially founded on September 12, 1962, on the initiative of the Provincia Autonoma di Trento (the Autonomous Province of Trento), the Istituto Trentino di Cultura (the Trento Institute of Culture), and of Professor Giorgio Braga, lecturer at the Università Cattolica in Milan. The governing body is composed of ten professors: three jurists, two economists, one statistician, one mathematician (the director, Mario Volpato), one ethicist, and the sociologists Braga (vice director) and Franco Ferrarotti. The first sociologists to teach there were not from the Milan area. Later Milanese came to Trento: Francesco Alberoni and Guido Baglioni. The first teachers at Trento were from several regions: Giorgio Braga (from the Università Cattolica in Milan), Franco Ferrarotti (Rome), Filippo Barbano (Turin), Sabino Acquaviva (Padua), Franco Demarchi (Trento), and Achille Ardigò (Bologna). The students who enrolled in the first year (226), and even more in the immediately following years, were strongly motivated to study social problems and came increasingly from regions far from Trento.

The Faculty of Sociology, with its four-year program, grants two types of degrees in sociology: general and special. The general course in the sociological disciplines trains teachers and researchers who will work in universities, international institutions, and centers of research on economic and social problems. The special sociology course prepares students for management careers in public administration and in private firms (in particular, for research, public relations, and personnel), social insurance offices, agencies for agricultural development, welfare agencies, labor unions and political parties, business consultancy agencies, marketing research offices, and town planning bodies.

As can be seen, this university planning aimed at the extension of the university; moreover, it set out to deal with a society that was both complex and predictable in its organization of problems and phenomena to be studied and, if possible, solved.

Things changed with the arrival of the student protest movement. At the University of Trento, different models of teaching and organization were experimented with, leading it to occupy a unique position in the Italian university system. A "critical university," managed by a joint committee of teachers and students, was formed. The director of the

institute and experimenter with this model was Francesco Alberoni, who attracted other teachers interested in this project, in particular from Milan. This situation lasted until 1970; meanwhile, it triggered experimentation at other Italian universities and to some extent contributed to the reform of the faculties of political science. In subsequent years, when the institute was transformed into the first Faculty of Sociology in the Italian university system, *corsi di laurea* (degree courses) in sociology emerged, often within existing *facoltà di magistero* (education faculties) at Rome, Naples, Salerno, and Urbino.

**The Reform of the Faculties of Political Science.** Sociology was offered in twenty-three Italian universities in 1964–1965, a total of thirty-eight courses taught by twenty-seven professors. It was episodically and marginally introduced into the faculties of law, letters, arts, economics, and political science, and often the same person was asked to teach several courses; for example, Alberoni taught four different courses in sociology at the same time at the Università Cattolica of Milan. The courses were scattered throughout Italy: four courses each in Rome and Milan (Università Cattolica); three courses each in Bari and Florence; two courses each in Bologna, Cagliari, Milan (Università Statale), Naples, and Palermo; and one course each at the universities of Catania, Ferrara, Genoa, Messina, Milan (Polytechnic), Padua, Pavia, Pisa, Salerno, Siena, Turin, Trieste, Urbino, and Venice. For the most part the teachers were “masters”: Achille Ardigò, Gianfranco Morra, Anna Anfossi, Franco Leonardi, Camillo Pellizzi, Giovanni Sartori, Luciano Cavalli, Renato Treves, Francesco Alberoni, Sabino Acquaviva, Eugenio Pennati, Agostino Palazzo, Franco Ferrarotti, Vittorio Castellano, Antonio Carbonaro, Filippo Barbano, Angelo Pagani, and Alessandro Pizzorno.

At the end of the 1990s, the situation has completely changed, both in terms of the number of universities in which sociology is taught (now forty-two), obviously expanded by the establishment of new universities—and in terms of the number of courses and teachers (now 827). Moreover, some very substantial centers of sociology have been established, and sociology is taught in all the Italian universities. The universities with the greatest numbers of sociology courses are (in descending order): Rome (eighty-four), Bologna (thirty-nine), Turin (sixty-three), Milan (forty-three),

and Trento (fifty-two). Those with fewer than twenty but at least ten courses are Padua (seventeen), Naples (sixteen), Calabria (fourteen), Florence (thirteen), Salerno (twelve), Catania (eleven), and Palermo (ten).

One very important technical reason for these profound changes in the teaching of sociology was the reform of the faculties of political science. The first faculty of political science in Italy was established in 1875 at Florence (and called Cesare Alfieri). It was designed to train public officials and, in general, it prepared men for an active life and public debate. As time passed, while it kept these functions, this faculty tended to become transformed into a faculty of social science. After World War II the need was increasingly felt to reform the faculties of political science, detaching them from other faculties (particularly from the law faculties) and giving them an updated cultural core. Subsequently such faculties were divided into two two-year courses of study: The first is based on fundamental courses (including sociology), and the second is organized into five *indirizzi* (courses of study)—international, historical, economic, administrative, and social. In this last course (political-social) some true sociological curricula are offered; they are obviously more concentrated in the faculties of political science at Rome, Bologna, Turin, Padua, and Milan. At present, the political-social course of studies is available in the following universities: Turin, Milan (Università Statale and Università Cattolica), Pavia, Trieste, Padua, Bologna, Florence, Pisa, Cambrino, Rome (LUISS), Naples, Bari, Messina, Catania, and Palermo.

#### **The Future of Sociology in the Universities.**

At present, the teaching of sociology is concentrated in the Faculties of Sociology at Trento, Rome, Naples, Milan, Urbino, Salerno; in the degree courses in sociology at others universities; and in the several faculties of political science. Sociology courses are also offered by other faculties (such as economics and commerce, architecture, medicine, letters and philosophy, and arts), but the current tendency is to reduce this spread. There are at least two reasons for this: the “zero growth” of sociology (i.e., sociology departments have the right only to replace sociologists who leave the university or switch to other disciplines) and other disciplines’ expansion of their number of chairs.



In the future, there will therefore be a tendency to strengthen sociology where it is already strong (in the faculties of political science) or through the creation of new degree courses in sociology; however, it is likely to disappear or to be excluded from the mainstream where it is peripheral (in the faculties of economics and commerce, architecture, and others). There are other ways in which the universities can train sociologists, such as Ph.D. programs (they are at the moment thirty-one: thirteen in the North; twelve in the center of the country; five in the South, one in Sicily and Sardinia) courses or schools of specialization for those who already have a degree and special-purpose schools or diploma courses for those who plan to attend the university for only two years after their secondary schooling.

#### **THE 1990S: IN THE SHADOW OF INTERNATIONAL CHANGE AND SOCIAL CHANGE IN ITALY**

The themes explored by Italian sociology, like those of any national sociology, are closely tied up with the contemporary state of society.

Italian society in the 1990s is undergoing profound changes. Of these, the most evident has resulted from the fall of communism in the Soviet Union and its satellites. The Christian Democratic party, whose *raison d'être* was as a bastion against communism, lost this purpose and no longer found itself "condemned to govern." Because of the increasing secularization of Italian Catholicism, the Christian Democrats also lost their function as the party of the Church, and it became widely acknowledged that the party was corrupt, as were the other parties. "Tangentopoli" ("Bribe City"—the name given to the massive system of the political fixing of public contracts in return for cash from industry) was uncovered, and "Mani pulite" ("Operation Clean Hands"—the attempt to eradicate it) was launched by the magistrature. The alliances between large organizations and their joint vested interests (large-scale capitalists, big trade unions, protected labor), the "Soviet" component of the state economy (state-owned industries), and the big political parties lost legitimacy with the birth of new political players (the Northern League), new economic players (Small and Medium-Sizes Enterprises), and new trade unions

(smaller, autonomous, and more trade-specific) that claimed resources and reduced the size and authority of the large organizations.

The fall of the Iron Curtain has de-ideologized international relations between Eastern and Western Europe, establishing national or ethnic identity as the criterion of statehood. But in Italy, the fall of the great central organizations has made obvious the importance of the role of highly fragmented local political, social, cultural, and economic forces. The burgeoning importance of ethnicity and local parties (more in the North than the South) and demands for federalism (especially in taxation) and even independence point to a new society, with the marks of the postmodern.

Growing of affluence has put an end to internal emigration and emigration abroad and directed interest toward women's rights, communications, and information technology, and new religious solutions within and outside the traditional Western religions. By contrast, however, the welfare state, designed and constructed by the society of large organizations and by "big" government, has entered a massive solvency crisis, burdened as it is with a level of debt that is unsustainable in comparison with other developed countries and that it cannot pay. Solutions have been sought in cuts in public spending and pensions, the growth of the service sector, the involvement of individual citizens, and the campaign for convergence to the Maastricht criteria.

Sociology has responded to these changes with the expansion of some areas of study and the contraction of others. The former have shifted to incorporate related disciplines such as political science, theology, computer science, constitutional law, and migration; the latter have been subsumed into disciplines such as economics. Research institutes have regained an important position, both inside and outside universities, and the spread of sociology courses in universities has stopped.

One relatively new theme for sociological research and reflection is national identity, both in terms of establishing whether an Italian national identity actually exists in the minds of Italians and in terms of exploring the local identities that may drive toward claims for independence or forms of federalism. A classical study in this regard is Gian

Enrico Rusconi's "*Se Cessiamo di Essere una Hazioue*" ("If We Cease to be a Nation"), which carries the emblematic subtitle *Regional Ethno-Democracies and European Citizenship*. Here the theme of national identity is combined with ethnic and European identity, and accentuated localism with accentuated cosmopolitanism. Work has also been produced in this field by Carlos Barbe, Loredana Sciolla, Carlo Tullio-Altan, and Riccardo Scartezzini. The futures studies journal *Futuribili* has devoted an edition to the future of Italy and the Italians entitled "The Italians Are Here, When Will Italy Be?"

Another theme of political sociology centres on changes in the structure of parties and consensus in the wake of the "discovery" that the existing political system is illegal. Much research has been conducted into the mechanisms of political corruption (Donatella Della Porta, Mauro Magatti, Vaclav Belohradsky), the renewal of the political class, and the formation of new ruling classes as a result of the transformation of the original illegalities into new legalities (Belohradsky 1994). These political transformations have changed the way voters feel about politics. Increases in the numbers of floating voters and nonvoters have given new life and direction to electoral sociology. Opinion polls have become a tool of fundamental importance, and sociologists are studying voting behavior in specialized institutes set up outside universities. Another dimension in the relaunching of political sociology is international relations, especially the study of post-communist Eastern Europe, including scenario building and forecasts of the workings, duration, and results of the transition there. In this field, too, research and studies have been carried out and conferences organized by institutes, both independent of universities or based in them. The funds used for these initiatives are provided by European programmes—such as Phare, Tacis, Intas, and Democracy—with the aim of creating international scientific and fact-finding networks.

Such scientific institutions have been set up in a large number of Italian towns, first and foremost ISIG in Gorizia, which has specialized for the last thirty years in the ethnic, border, and international questions in the Balkan–Danube area and the former Soviet area. ISIG has established a Forum for European Border Towns, an Ethnic Minorities Observatory, an Italian Futures Studies Academy,

and a Permanent Forum for peace initiatives specializing in pilot studies, databases, and the organization of parallel diplomatic meetings for the solution of conflicts. These initiatives frequently result in the publication of journals and books. Heading the journals are *Futuribili*, (along the lines of *Futuribles* and *Futures*), devoting each issue to special themes. *ISIG Magazine* is a journal, published in English and Italian, that also deals with particular themes on a monographic basis. Books frequently tackle international subjects: *International Solidarity and National Sovereignty* (Picco and Delli Zotti 1995) discusses the increasing intrusion upon national sovereignty by international organisations (the UN, NATO, and the EU); "The Future of the Moment Before" (Gasparini 1993) deals with Russia and its future. Other studies analyze: the factors driving toward the formation of a new Russian empire and the persistent character of Russian religiosity and mysticism; the Balkans and the solution of conflicts through international solidarity to bring in a new order based on protectorates (Mroljub Radojkovic, Kosta Barjaba, Alberto Gasparini); ethnic groups and borders in Europe (Luca Bregantini, Alessandro Pannuti, Moreno Zago, Alberto Gasparini); Arab and Israeli nationalism in the Middle East (Elie Kallas); the transition of the former socialist countries in Eastern Europe.

Other themes undergoing strong development are connected with the modern affluent society: They include women and their position with regard to work, equality, self-fulfilment outside the family, and the family as an institution (Chiara Saraceno, Laura Balbo, Pierpaolo Donati); poverty, and extreme poverty in cities rather than widespread poverty (Paolo Guidicini, Giovanni Pieretti); and the role played by poverty in immigration from the Balkans, Eastern Europe, and the Mediterranean Arab countries (Umberto Melotti, Maurizio Ambrosini, ISMU). Connected to these developments and to the relationship between politics and civil society as seen in an increasingly crisis-bound welfare state are studies (Pierpaolo Donati, Ivo Colozzi, Guido Lazzarini, Giuliano Giorio) that seek solutions to the crisis through the service sector and that explore Europe and its new forms of citizenship. Other emerging themes indicating change in Italian society and the need for certainties in an uncertain situation are to be found in the relaunching of studies on Italian religion and

religiosity (Roberto Cipriani, Franco Garelli, Stefano Martelli, Salvatore Abruzzese) and the solutions offered by traditional religions and new movements (also dealt with in an edition of *Futuribili*).

Another important theme is the relationship between communications and a changing society, with specific reference to the role of communications in mass emergency risk management (Bruna De Marchi, Luigi Pellizzoni, Daniele Ungaro, Alberto Abruzzese, Laura Bovone, Mario Morcellini). Community studies show a prevalence of territorial belonging and a decline in metropolitan expansion in favor of restricted urban areas forming a city system in which individuals retain their own identity while relating to other people (Alberto Gasparini, Reuzo Gubert, Gabriele Pollini).

Finally, a theme reemerging after the explosion and subsequent decline of the studies launched in the early 1970s by the Club of Rome is prediction. Here again, ISIG and *Futuribili* are at the fulcrum heart of an interest that is particularly strongly felt at a time of uncertainty and transition toward a society that is as yet unclear (Eleonora Masini, Giorgio Nebbia, Alberto Gasparini, Luca Bregantini, Moreno Zago).

In contrast to these themes, produced by contingent and transient situations, others have been declining in the 1990s. Less attention has been devoted to large companies, and the factory and the trade union movement have lost some of their importance as an area of study. The same applies to the study of social classes. These fields have felt the effects of a crisis in the culture of the defense of workers' rights and in the paradigm—very often a Marxist one—use to explain social classes and the relations between them.

In short, the profound external and internal changes overtaking Italian society have worked to the advantage of some themes of study and undermined the previous importance of others. As was the case with postwar sociology, the driving force in this development has been institutes outside universities (especially in international studies) or at least consortia and niches inside universities but only loosely tied to their teaching function. This has produced a strong movement to strengthen existing journals and create new ones; these include *Futuribili*, *Ikon*, *Polis*, *Quaderni di sociologia* (*Notebooks of Sociology*), *Rassegna italiana di sociologia* (*Italian Review of Sociology*), *RES* (*Ricerca e sviluppo*

*per le politiche sociali: Research and Development for Social Policies*), *Religioni e società* (*Religions and Society*), *La Società* (*The Society*), *Sociologia del lavoro* (*The Sociology of Work*), *Sociologia della comunicazione* (*The Sociology of Communication*), *Sociologia e politiche sociali* (*Sociology and Social Policies*), *Sociologia e professione* (*Sociology and Profession*), *Sociologia e ricerca sociale* (*Sociology and Social Research*), *Sociologia urbana e rurale* (*Urban and Rural Sociology*), *Studi di sociologia* (*Sociology Studies*), and *Teoria sociologica* (*Sociological Theory*).

REFERENCES

- Accornero, Aris, and Francesco Carmignani 1986 *I paradossi della disoccupazione*. Bologna: Il Mulino.
- Acquaviva, Sabino, and Enzo Pace, eds. 1984 *Dizionario di sociologia e antropologia culturale*. Assisi: Cittadella.
- Alberoni, Francesco ed., 1966 *Questioni di sociologia*. Brescia: La Scuola.
- 1968 *Statu nascenti*. Bologna: Il Mulino
- Amendola, Giandomenico, and Antonio Tosi, eds. 1987 “La sociologia dell’abitazione.” *Sociologia e ricerca sociale* 22:1–173.
- Ammassari, Paolo 1970 *Worker Satisfaction and Occupational Life*. Rome: University of Rome.
- Ardigò, Achille 1980 *Crisi di governabilità e mondi vitali*. Bologna: Cappelli.
- 1988 *Per una sociologia oltre il post-moderno*. Bari: Laterza.
- Ardigò, Roberto 1918 *Opere filosofiche*. Padova: Draghi.
- Associazione Italiana di Sociologia 1989 *L’Italia dei sociologi: Problemi, prospettive, indirizzi di ricerca*. Bologna: AIS.
- Bagnasco, Arnaldo 1977 *Tre Italie*. Bologna: Il Mulino.
- Banfield, Edward C. 1958 *The Moral Basis of a Backward Society*. Glencoe, Ill.: Free Press.
- Barbano, Filippo, and Giorgio Sola 1985 *Sociologia e scienze sociali in Italia, 1861–1890*. Milan: Angeli.
- Barberis, Corrado 1988 *La classe politica municipale*. Milan: Angeli.
- Barjaba, Kosta, ed. 1996 “Albania. Tutta d’un pezzo, in mille pezzi . . . e dopo?” *Futuribili* 2–3:453.
- Belohradsky, Vaclav, ed. 1994 “Illegalità-legalità e nuovi patti sociali. Dalla Russia all’Italia per l’Europa.” *Futuribili*, 3:239.
- Bestuzhev-Lada, Igor, and Luca Bregantini, eds. 1995 “Balcanizzazione euro-asiatica vs. nuovo impero russo.” *Futuribili* 1–2:336.

- Bettin Lattes, Gianfranco, ed. 1995 *La società degli europei*. Bologna: Monduzzi.
- Boileau, Anna Maria, and Emidio Sussi 1981 *Dominanza e minoranze*. Udine: Grillo.
- Borgotta, G. 1917 *L'opera sociologica e le feste giubilari di V. Pareto*, Torino: SE.
- Bonazzi, Giuseppe 1982 *Dentro e fuori della fabbrica*. Milan: Angeli.
- Bregantini, Luca and Alessandro Pannuti, eds. 1997 "Etnia? Sia se volete che sia." *Futuribili*, 1-2:351.
- Burgalassi, Silvano, and Gustavo Guizzardi, eds., 1983 *Il fattore religione nella società contemporanea*. Milan: Angeli.
- Caizzi, Bruno, ed. 1962 *Nuova antologia della questione meridionale*. Milan: Comunità.
- Castellano, Vittorio 1968 *Introduzione alla sociologia e primi elementi de morfologia sociale*. Rome: Ilardi.
- Cattarinussi, Bernardo, and Carlo Pelanda, eds. 1981 *Disastro e azione umana*. Milan: Angeli.
- Cavalli, Luciano 1957 *Inchiesta sugli abituri*. Genoa: Office of Social and Labor Studies.
- Cella, Gian Primo, ed. 1979 *Il movimento degli scioperi nel XX secolo*. Bologna: Il Mulino.
- Cipolla, Costantino 1997 *Epistemologia della tolleranza*. Milan: Angeli.
- Crespi, Franco, ed. 1987 *Sociologia e cultura*. Milan: Angeli.
- De Finis, Giorgio and Riccardo Scartezini, eds., 1996 *Universalità e differenza*. Milan: Angeli.
- De Marchi, Bruna, ed. 1994 "Rischio, gestione del rischio, comunicazione del rischio" (Risk, risk management, risk communication). *Isig-Magazine*, III:4-16.
- De Rita, Lidia 1964 *I contadini e la televisione*. Bologna: Il Mulino.
- Delli Zotti, Giovanni 1983 *Relazioni transnazionali e cooperazione transfrontaliera*. Milan: Angeli.
- Demarchi, Franco 1969 *Società e spazio*. Trento: Istituto Superiore di Scienze Sociali.
- , Aldo Ellena, and Bernardo Cattarinussi, eds., 1987 *Nuovo dizionario di sociologia*. Rome: Paoline.
- Dolci, Danilo 1955 *Banditi a Partinico*. Bari: Laterza.
- 1957 *Inchiesta a Palermo*. Turin: Einaudi.
- 1960 *Spreco*. Turin: Einaudi.
- Donati, Pierpaolo 1983 *Introduzione alla sociologia relazionale*. Milan: Angeli.
- 1993 *La cittadinanza societaria*. Bari: Laterza.
- Ferrarotti, Franco 1972 *Trattato di sociologia*. Turin: Utet.
- Gallino, Luciano 1978 *Dizionario di sociologia*. Turin: Utet.
- 1980 *La società. Perché cambia, come funziona*. Turin: Paravia.
- , ed. 1992 *Percorsi della sociologia italiana*. Milan: Angeli.
- , ed. 1992 *Teoria dell'attore e processi decisionali*. Milan: Angeli.
- Galtung, Johan 1971 *Members of Two Worlds: A Development Study of Three Villages in Western Sicily*. Oslo: Universitetforlaget.
- Garelli, Franco, and Marcello Offi 1996 *Fedi di fine secolo*. Milan: Angeli.
- Gasparini, Alberto 1975 *La casa ideale*. Venice: Marsilio.
- 1982 *Crisi della città e sua reimmaginazione*. Milan: Angeli.
- , and Paolo Guidicini, eds., 1990 *Innovazione tecnologica e nuovo ordine urban*. Milan: Angeli.
- 1993 *The Future of the Moment Before*. Gorizia: ISIG.
- , and Miroљub Radojkovic, eds. 1994 "Oltre le guerre balcaniche. Cosa può succedere quando i piccoli dei hanno grandi sogni". *Futuribili* 2:22.
- , and Moreno Zago, eds. 1995 "Al di là dei nuovi muri. L'Europa collaborativa che viene dalle città di confine." *Futuribili* 1-2:352.
- , and Vladimir Yadov, eds. 1995 *Social Actors and Designing the Civil Society of Eastern Europe*. Greenwich, Conn.: JAI Press.
- Gini, Corrado 1960 *Organismo e società*. Rome: Ilardi.
- Giorio, Giuliano 1985 *Società e sistemi sociali*. Milan: Angeli.
- Gross, Feliks 1970 "Value Structure and Social Change." *Revue Internationale de Sociologie* 1(3):85-120.
- Gubert, Renzo 1972 *La situazione confinaria*. Trieste: Lint.
- 1976 *L'identificazione etnica*. Udine: Del Bianco.
- , ed. 1999 *Territorial Belonging Between Ecology and Culture*. Trento: Università degli Studi di Trento.
- Guidicini, Paolo 1987 *Manuale per le ricerche sociali sul territorio*. Milan: Angeli (Revised Edition).
- , and Giovanni Pieretti, eds. 1998 *Città globale e città degli esclusi*. Milan: Angeli.
- Harrison, Gualtiero, and Maria Callari Galli 1971 *Né leggere, né scrivere*. Milan: Feltrinelli.
- Izzo, Alberto 1977 *Storia del pensiero sociologico*. Bologna: Il Mulino.
- La Rosa, Michele 1986 "La sociologia del lavoro in Italia." *Sociologia del lavoro* 26-27:1-137.
- Lentini, Orlando 1981 *La sociologia italiana nell'età del positivismo*. Bologna: Il Mulino.

- Lopreato, Joseph 1961 "Social Stratification and Mobility in a South Italian Town." *American Sociological Review* 4.
- Lotti, Luigi, and Gianfranco Pasquino 1980 *Guida alla facoltà di scienze politiche*. Bologna: Il Mulino.
- Magatti, Mauro, ed. 1997 *Per la società civile*. Milan: Angeli.
- Marselli, Gilberto Antonio 1963 "American Sociologists and Italian Peasant Society: With Reference to the Book of Banfield." *Sociologia Rurale* 3:15-32.
- Martinotti, Guido, ed. 1968 *Città e analisi sociologica*. Padua: Marsilio.
- Melotti, Umberto, ed. 1988 "Dal terzo mondo in Italia. Studi e ricerche sulle migrazioni straniere." *Quaderni del Terzo Mondo* 31-32: 1-18.
- Mosca, Gaetano 1962 *Storia delle dottrine politiche*. Bari: Laterza.
- Mottura, Giovanni, and Enrico Pugliese 1975 *Agricoltura, Mezzogiorno e mercato del lavoro*. Bologna: Il Mulino.
- Negrotti, Massimo 1979 *Uomini e calcolatori*. Milan: Angeli.
- Niceforo, Alfredo 1953 *Avventure e disavventure della personalità e della umana società*. Milan: Bocca.
- Pagani, Angelo, ed. 1960 *Antologia di scienze sociali*. Milan: Angeli.
- Pareto, Vilfredo 1964 *Trattato di sociologia generale*. Milan: Comunità.
- Pellizzi, Camillo 1964 *Rito e linguaggio*. Rome: Armando.
- Picco, Giandomenico, and Giovanni Delli Zotti, eds., 1995 *International Solidarity and National Sovereignty*. Gorizia: ISIG.
- Pinna, Luca 1971 *La famiglia esclusiva: parentela e clientelismo in Sardegna*. Bari: Laterza.
- Pinto, Diana, ed. 1981 *Contemporary Italian Sociology*. Cambridge, U.K.: Cambridge University Press.
- Pirzio Ammassari, Gloria 1997 *L'Europa degli interessi*. Roma: Euroma.
- Pizzorno, Alessandro 1960 *Comunità e razionalizzazione*. Turin: Einaudi.
- Pollini, Gabriele 1987 *Appartenenza e identità*. Milan: Angeli.
- Regini, Marino, ed. 1988 *La sfida della flessibilità*. Milan: Angeli.
- Ricciardi Ruocco, Maria 1967 *Inchiesta a Marsala*. Manduria: Lacaita.
- Rusconi, Gian Enrico 1993 *Se cessiamo di essere una nazione*. Bologna: Il Mulino.
- Savelli, Asterio 1989 *Sociologia del turismo*. Milan: Angeli.
- Scaglia, Antonio 1988 *Sociologia: Dalle scienze della natura alla scienza dell'agire umano*. Milan: Angeli.
- 1992 *La sociologia europea del primo Novecento*. Milan: Angeli.
- Scivoletto, Angelo, ed. 1983 *Sociologia del territorio*. Milan: Angeli.
- Sgroi, Emanuele 1991 *La questione ambientale da allarme a progetto: Le nuove professionalità*. Naples: CUEN.
- Strassoldo, Raimondo 1976 *Ambiente e società*. Milan: Angeli.
- 1979 *Temi di sociologia delle relazioni internazionali*. Gorizia: ISIG.
- 1993 *Le radici dell'erba*. Napoli: Liguori.
- Sturzo, Luigi 1960 *La società, sua natura e leggi*. Bologna: Zanichelli.
- Sussi, Emidio, and Danilo Sedmak 1984 *L'assimilazione silenziosa*. Trieste: Editoriale Stampa Triestina.
- Tentori, Tullio 1956 *Il sistema di vita nella comunità materana*. Rome: UNRA-Casas.
- Tullio-Altan, Carlo 1995 *Ethnos e civiltà*. Milan: Feltrinelli.
- Ungaro, Daniele 1997 *La transizione italiana*. Roma: Armando.
- Urpis, Ornella 1998 *Nuove identità e secolarizzazione*. Gorizia: ISIG.
- Vincelli, Guido 1958 *Una comunità meridionale. Montorio dei Frentani*. Turin: Taylor.
- Visentini, Luciano 1984 *Tra mestiere e vocazione: La sociologia del lavoro in Italia*. Milan: Angeli.

ALBERTO GASPARINI

# Encyclopedia of Sociology

---

Second Edition

# Encyclopedia of Sociology

---

Second Edition

VOLUME 3

**Edgar F. Borgatta**

*Editor-in-Chief*

*University of Washington, Seattle*

**Rhonda J. V. Montgomery**

*Managing Editor*

*University of Kansas, Lawrence*

**Macmillan Reference USA**

*an imprint of the Gale Group*

*New York • Detroit • San Francisco • London • Boston • Woodbridge, CT*

# J

---

## JAPANESE SOCIOLOGY

Japanese sociology divides roughly into four stages of development: pre-World War II, with emphasis on theoretical and philosophical orientations, influenced primarily by European (especially German) sources; post-World War II, with growing emphasis on empirical orientations, influenced primarily by the United States; diversification, with emphases on both theoretical and empirical orientations (on various aspects of the history of Japanese sociology, see, e.g., Halmos 1966; Koyano 1976; Odaka 1950); and globalization, with emphasis on theoretical orientations and an increasing number of empirical orientations, some encompassing cross-national and foreign area studies. In a general sense, the development of Japanese sociology reflects the country's social and cultural change, as well as shifting national policies. The significant Western influence generally exhibits a time lag in terms of its expression in Japanese sociology.

### PRE-WORLD WAR II STAGE (1893–1945)

Japanese sociology began as a European import and reflected a conservative stance. This occurred shortly after the Meiji Restoration of 1868. E. F. Fenollosa (1853–1908), an American professor, first taught sociology at the University of Tokyo in 1878. Three years later, Masakazu Toyama (1848–1900) began teaching at the same university; in 1893 (just one year after the founding of the University of Chicago's sociology department), he

became the first professor of sociology in Japan and is regarded as the founder of Japanese sociology. Toyama, and later Nagao Ariga (1860–1921), a student of Fenollosa and the first sociologist in Japan to publish, both introduced aspects of Herbert Spencer's organic analogy for society. The works of Spencer and John Stuart Mill were particularly significant during these early years and were translated frequently.

Tongo Takebe (1871–1945), successor to Toyama in 1898, introduced Auguste Comte to Japan, combining Comte's positivism with Confucian philosophy and social thought to fit Japanese society. In 1913 Takebe also founded the Japan Institute of Sociology, an organization replaced by the Japan Sociological Society in 1924. A new approach began to take hold in the 1910s—the psychological approach initiated by Ryukichi Endo (1874–1946), who drew on Franklin Giddings's theory of consciousness of kind to explain social phenomena.

During the 1910s, other Western sociological theories came to Japan, largely through the work of Shotaro Yoneda (1873–1945). Yoneda, who looked at society and culture from a sociopsychological perspective, was an important teacher who introduced the ideas of many Western sociologists to Japan, including those of Gabriel Tarde, Emile Durkheim, Georg Simmel, and Franklin Giddings. Yoneda laid the groundwork for the subsequent strong influence of the German school of sociology.

From this point forward, until the end of World War II, the German school dominated



Japanese sociology. There were two major divisions that grew out of the German school: *formale Soziologie* (formal sociology) and, later, *Kultursoziologie* (cultural sociology). The major proponent of the former was Yasuma Takata (1883–1972), a student of Yoneda. Takata (1922, 1989) successfully changed the view of sociology from that of a synthesis of the social sciences to one in which sociology stood as separate and independent, drawing in particular on the work and influence of Max Weber, Georg Simmel, Ferdinand Tönnies, and Robert MacIver.

New influences, however, emerged in the 1920s. Formal sociology was deemed abstract and out of touch with the real world. As a consequence, cultural sociology gained a stronger foothold in both Germany and Japan. Pioneering the work in cultural sociology in Japan was Eikichi Seki (1900–1939). No doubt a reaction to the Depression of 1929, cultural sociology gained popularity for its closer ties with the social realities of the day. Although a theory of cultural sociology fitting the Japanese society seemed imminent, it never really unfolded.

While there were also French and American influences on Japanese sociology during the prewar period, they were minor compared with those of Germany. Jyun'ichiro Matsumoto (1893–1947) saw a need to synthesize formal and cultural sociology into what he would call “general sociology.” At the same time, Masamichi Shimmei (1898–1988) sought to take Matsumoto's thoughts and combine them with Simmel's general sociology and the thinking of Karl Mannheim.

Because Western theory and thought dominated Japanese sociology in the prewar period, little of the work analyzed Japanese society. There were, however, a handful of notable empirical studies, especially in family and rural sociology, a tradition begun at the University of Tokyo by Teizo Toda (1887–1955). Toda had studied at the University of Chicago, where he learned about survey methodologies being used in the United States. Toda analyzed statistics on the Japanese family structure, using census and other then-current and historical data. Kizaemon Aruga (1897–1979) worked in the area of rural sociology, linking his findings with previous folklore studies and working toward clarifying the condition of social strata in prewar Japan. Lack of financial support,

however, hindered the development of empirical research during this time.

Two phenomena in particular worked against the development of Japanese sociology prior to World War II. First, Japanese sociology focused on European sociology rather than on studies of its own society. The second phenomenon, bolstered by government officials and scholars inclined toward nationalistic militarism, involved a distorted public image: that sociology and sociologists were associated with socialism because of the two words' similarity in the Japanese language (“sociology,” *shakaigaku*; “socialism,” *shakaishugi*). Many thought that sociology was the study of socialism or social revolution and that sociologists were socialists and, therefore, a sinister threat to national security. As World War II grew closer, and during the war, publications were often censored, academic freedom was severely curtailed, and meetings and conventions were forbidden.

#### POSTWAR STAGE (1946–1960s)

Defeat and U.S. occupation brought drastic social changes to Japan. The traditional family system collapsed, and land reform became the order of the day. Favorite prewar survey subjects centering on village and family were replaced by issues related to land reform and revision of traditional family values. Indeed, the traditional Japanese value system was pulled out from under the nation. “Democratization” was the new buzzword. The term “sociology” was released from taboo. Educational reforms in the 1950s now required sociology courses as part of the general university education, especially for freshmen and sophomores. More and more departments of sociology or sociology programs within other departments were formed, particularly at private colleges and universities. Suddenly, many sociologists were needed. American influences were rampant in all areas of Japanese society, and sociology was no exception. Many American sociological theories came to influence Japanese sociology, the strongest being that of Talcott Parsons.

As the importance of empirical study was growing in the United States, Japanese sociologists also began to develop a strong interest in empirical

work. Social research, positivism, and functionalism were key words. Marxism had significant impact on Japanese sociology as well. "Democratization" and "modernization" were major fundamental themes in sociological studies. Japanese sociologists studied American research methods, ultimately leading to a rapid increase in surveys and research based on the results of these surveys. To many, the empirical studies of Japanese sociology moved the entire discipline from one of art and humanity to one of social science. However, many surveys were carried out for fact-finding purposes rather than hypothesis testing for theory construction.

Tadashi Fukutake (1917–1989) significantly influenced the postwar stage and the subsequent stage of Japanese sociology. Fukutake's studies focused on rural sociology (see, e.g., 1967) in the context of Japanese society's postwar democratization. Also influential was Kunio Odaka (1908–1993), a positivist, who played an important role in industrial sociology (see, e.g., 1975) as well as general sociology, especially during the period of extraordinary economic development from 1955 through 1965. Odaka and Fukutake were two of Japan's leading empirical researchers conducting field research in real social settings during this stage in the development of Japanese sociology.

Mentioned earlier, the Japan Sociological Society is a nationwide organization for Japanese sociologists. It holds annual meetings and publishes a journal, and it joined the International Sociological Association (ISA) in 1950. Two years later, a survey on social stratification and mobility was conducted under the auspices of the Japan Sociological Society, led by Odaka, in cooperation with the ISA. This survey was repeated three years later on a nationwide scale and subsequently every ten years.

In 1954, the Institute of Statistical Mathematics in Tokyo began a nationwide time-trend survey of the Japanese national character, a survey conducted every five years since, with the objective of analyzing changes (or lack thereof) in general social attitudes among the Japanese since World War II (Hayashi 1998). This ongoing survey pioneered the use of identical questions over time and as such became a model for the General Social Survey of the National Opinion Research Center at the University of Chicago. Both these ongoing

surveys continue as the most well-known nationwide social surveys in Japan.

Although the postwar period saw great social change in Japan, within academic circles, senior sociologists, most of whom belonged to the prewar generation, prevailed. A generational change among leading Japanese sociologists occurred in the 1960s, marking the end of the postwar period (Koyano 1976).

#### DIVERSIFICATION STAGE (1960s–1990s)

Since the 1960s, American sociology has gained an ever-stronger influence on Japanese sociology. With the exception of Talcott Parsons and his structural-functionalism, however, no major American sociologists have significantly influenced the theoretical aspects of Japanese sociology. Whereas Marxist sociology had tended to influence many of the younger Japanese sociologists from a theoretical perspective, "the entire history of [Japanese] sociological development to the 60s was criticized and thrown into examination by more or less radical criticism. . . . Thus many talented younger sociologists turned from Marxism to structuralism or structuralist social theory . . . [or] alternately accepted rather subjective methodologies and theories such as phenomenological sociology, symbolic interactionism and ethnomethodology . . . [i.e.] phenomenological trends" (Shoji 1996). Interests among some Japanese sociologists also shifted from macro- to micro-sociological analyses. Multi-dimensional paradigms became prevalent, such as those of Michel Foucault, Jürgen Habermas, Niklas Luhmann, Pierre Bourdieu, Anthony Giddens, and Alfred Schutz, all of whom have consequently influenced Japanese sociology. Much as the prewar emphases of the theoretically oriented Japanese sociologists tended toward the purely theoretical, the new attractions continued the inclination toward speculative and interpretive theory—more like social philosophy—as against the empirical science tradition represented by Fukutake and Odaka. Though Japanese sociologists have been quite keen on the general trends in Western social and sociological thought, there is about a ten-year lag from introduction to translation and analysis of the works of Western sociologists.

Along with the vast changes in Japanese society in the postwar period came a nearly unlimited

number of topics for sociological study and investigation, particularly from an empirical standpoint. This, too, accounted for diversification in Japanese sociology and signaled the establishment of a number of subdisciplines in the field. Thus, as a result of American influences, economic development, and a host of other factors, Japanese sociology continuously diversified from the early 1960s on. This is partially attributable to the fact that the industrialization of Japan, into the early 1970s, led to serious social and environmental problems, which in turn led to student uprisings, increases in delinquency and other expressions of social unrest, and, in response, the emergence of environmental-protection, feminist, and other movements, even though many of these social phenomena and movements waned in the 1980s.

These circumstances brought forth a wide variety of challenging research topics for sociologists and coincidentally created a situation in which there are no especially influential figures in Japanese sociology, although each subdiscipline does have its major proponents. These scholars include, among others, Eiichi Isomura (1903–1997) in urban sociology; Michio Nagai, who later became Japanese Minister of Education, in the sociology of education (cf. 1971); Kazuo Aoi in the field of small groups; Kiyomi Morioka in the sociology of religion and the sociology of the family (cf. 1975); Saburo Yasuda (1925–1990) in sociological methodology (cf. 1964); Akira Takahashi in social movements; Joji Watanuki in political sociology (cf. 1976); Ken'ichi Tominaga in social stratification (cf. 1969); and Tamito Yoshida in communication.

From the 1960s on, the number of sociologists in Japan grew markedly, to the point where, based on memberships in national sociological organizations, there are more sociologists in Japan than in any other country except the United States. While there were about 300 sociologists teaching at colleges and universities in Japan in the 1970s, their number grew to about 1,000 by the late 1980s. The Japan Sociological Society's membership rosters totaled 870 in 1957; 1,931 in 1985; 1,945 in 1988; 2,200 in 1990; 2,450 in 1992; and 3,034 in 1999.

In 1988, Japanese sociology had about thirty subdisciplines. On the basis of first-, second-, and third-choice subdiscipline selections by the 1,945 members of the Japan Sociological Society at that time, the most prominent were: (1) rural sociology

and community studies, 17.7 percent; (2) sociology of the family, 17.2 percent; (3) general sociological theories, 16.4 percent; (4) social welfare, social security, and medical sociology, 16.2 percent; (5) social thought and the history of sociology, 14.9 percent; (6) management, industry, and labor, 13.0 percent; (7) social pathology and social problems, 12.5 percent; (8) culture, religion, and morality, 12.4 percent; and (9) the sociology of education, 11.6 percent. Among other things, we see that sociology of the family and rural sociology held their significance from the prewar era.

Sociologists who study industrial sociology—including management; urban sociology; and social welfare, social security, and medical sociology—have increased in number continuously since World War II. And, over time, foreign sociologists have shown more and more interest in industrial sociology; the sociology of education; and social welfare, social security, and medical sociology, as these are seen as particularly successful elements of Japan's economic and social development.

Derived from lists of publications in the *Japanese Sociological Review* between 1984 and 1988, publication of articles originating from the various subdisciplines broke down as follows: (1) sociology of the family, 7.4 percent; (2) social thought and the history of sociology, 7.2 percent; (3) general social theories, 6.5 percent; (4) the sociology of education, 6.5 percent; (5) urban sociology, 6.2 percent; (6) rural sociology and community studies, 5.8 percent; (7) industrial sociology and management, 5.2 percent; (8) social pathology and social problems, 5.1 percent; and (9) social welfare, social security, and medical sociology, 5.1 percent. Articles totaled 7,426 (books, 927) during the five-year period, most of which appeared in Japanese.

Unlike in the United States, in Japan there is no rigid screening or referee system for publications, with the exception of a few well-known journals such as the *Japanese Sociological Review* (*Shakaigaku Hyoron*, the official journal of the Japan Sociological Society), *The Study of Sociology* (*Shakaigaku Kenkyu*), and *Sociology* (*Shoshioroji*). With regard to presentations at meetings of the Japan Sociological Society, the five regional associations, and the associations of the various subdisciplines, there have been, in many cases, no rigid referee systems. Heated debate is rare, and

thus academic stimulation from published or presented controversies is quite limited.

By the late 1980s, 33 of Japan's 501 colleges and universities had doctoral programs in sociology. The major institutions with such programs included the public universities of Hokkaido, Tohoku, Tokyo, Hitotsubashi, Tokyo Metropolitan, Nagoya, Kyoto, Osaka, Kobe, and Kyushu, and the private universities of Waseda, Keio, and Hosei. Also during the period, there were about 700 graduate students studying sociology, 490 of whom were doctoral candidates (see Committee on Education for Sociology 1988). In general, two years are required to obtain a master's degree and an additional three years to finish coursework for doctoral programs. A much higher percentage of those who complete a master's program at public universities go on to a doctoral program than do those at private universities. Most who obtain master's degrees do not complete their doctoral theses within three years. Rather, after finishing their doctoral coursework, they obtain teaching or research positions and often complete their doctorates at a later stage in their careers. Forty-one persons obtained doctoral degrees in sociology during the period 1977–1986. During this time, there was a surplus of sociology graduates versus the number of teaching positions available. In the late 1980s, of those teaching at Japanese universities who obtained their doctoral degrees from Japanese universities, twenty-four sociology professors held doctoral degrees in literature, about thirty held doctoral degrees in sociology, and several others held doctoral degrees in related fields. Compared to other social sciences such as economics, the number of professors who obtained their doctoral degrees in sociology from foreign educational institutions is limited. For instance, only about twenty Ph.D. holders who taught at Japanese colleges and universities in the late 1980s obtained their degrees in the United States, although their numbers have been increasing.

#### GLOBALIZATION STAGE (1990s AND BEYOND)

As a whole, Japan has recently seen substantial movement toward globalization (or internationalization), and Japanese sociology is no exception. However, before exploring the implications of

globalization for Japanese sociology, it is appropriate to look at the changes that took place in the 1990s.

In 1997, out of 586 universities and colleges in Japan, 65 had master's programs and 47 had doctoral programs in sociology. In 1997, more than forty-five Ph.D. degree holders (obtained from non-Japanese universities) in sociology were teaching at Japanese universities and colleges. Also, more than sixty-five doctoral degree holders (obtained from Japanese universities) in sociology were teaching at various universities and colleges in Japan. It should be noted that these numbers are not at all significant compared to those in the West.

The Japan Sociological Society polled its members in 1998 for their preferred subdisciplines and research fields. They identified (1) general sociological theories, 15.6 percent; (2) sociology of the family, 15.2 percent; (3) communications and information, 15.1 percent; (4) social thought and the history of sociology, 14.7 percent; (5) social welfare, social security, and medical sociology, 14.3 percent; (6) culture, religion, and morality, 13.6 percent; (7) social psychology and social attitudes, 13.1 percent; (8) rural sociology and community studies, 12.2 percent; and (9) cross-national and foreign area studies, 11.3 percent. Note that the first three choices, including communications and information, are effectively tied for first place. We can see, compared with the figures from 1988, that management, industry and labor, social pathology and social problems, and the sociology of education all dropped out of the top nine fields, having been replaced by communications and information, social psychology and social attitudes, and cross-national and foreign area studies. That communications and information made such a showing is particularly notable and reflects the changes in information infrastructure (the Internet, among others) that are in large part responsible for globalization.

During the period 1989–1996, publication of articles originating from the various subdisciplines broke down as follows: (1) social thought and the history of sociology, 8.7 percent; (2) the sociology of education, 7.2 percent; (3) sociology of the family, 6.6 percent; (4) general sociological theories, 6.6 percent; (5) urban sociology, 6.2 percent; (6) communications and information, 5.6 percent;

(7) culture, religion, and morality, 5.6 percent; (8) management, industry, and labor, 5.4 percent; and (9) rural sociology and community studies, 5.2 percent. These figures were derived from lists of publications in the *Japanese Sociological Review*. Notice that, compared to the publications listed from 1984 through 1988, social pathology and social problems and social welfare, social security, and medical sociology dropped from the top nine while communications and information, and culture, religion, and morality appeared. That communications and information did not rise to the top here is not surprising, as publishing traditionally carries with it varying degrees of time lag.

As in the United States and western Europe, aging has become a serious issue; in 1998, about 16 percent (i.e., 20 million persons) of the Japanese population was over 65. Therefore, more and more sociologists are involving themselves in this field and contributing to national and local policy formation.

In Japan, there are no major university research centers such as the Institute for Social Research at the University of Michigan or the National Opinion Research Center at the University of Chicago, nor does Japan have any colleges or universities that are especially well known for their sociology departments or programs. However, there are some survey sections within organizations (e.g., research institutions, government organizations, the press and mass media) that have carried out major surveys since the 1950s, some on a continuing basis, including surveys targeting trends in social attitudes. Among others, these include the Japanese Prime Minister's Office; the Institute of Statistical Mathematics, as mentioned above; the Mainichi Press; Jiji Press; and Nippon Hoso Kyokai, a Japanese broadcasting organization (see Sasaki and Suzuki 1991). Each one of these endeavors is independent and generally does not provide its data to outside researchers, making secondary analysis of such data a difficult task in Japan. This hinders those in graduate training, as they often lack access to such data for thesis work.

On occasion, surveys, including cross-national studies, are funded by agencies such as the Japanese Ministry of Education, the Japan Society for

the Promotion of Science, and the Toyota Foundation. On a regional or local basis, funds are sometimes provided by prefectural or municipal governments. Findings from some of these studies have had significant impacts on policy formation at the local and national levels. Most research grants for sociologists, for both domestic and international (cross-national and foreign area) studies, are provided by the Japanese Ministry of Education. The Japan Society for the Promotion of Science provides grants for foreign area and cross-national research projects, for joint research conferences and seminars, as well as travel allowances for visiting scholars.

Japanese sociology does not enjoy a wide-ranging reputation in the rest of the world. This has been attributed to lack of integration and coordination, as well as a descriptive rather than analytic focus. Indeed, Japanese sociology has relied substantially on foreign influence, particularly that which is *au courant*, and has not excelled in the development of original theoretical or empirical ideas, rather having a stronger commitment to theory interpretation than theory testing and theory building. With the exception of a few research groups doing cross-national studies, in general Japan's sociologists have had limited contact with researchers in other nations. This can be attributed primarily to the language barrier and lack of experience in exchanging ideas. Despite these apparent shortcomings, sociology is comparatively popular in Japan, where there is a strong demand for books on the subject. As a consequence, Japanese scholars often feel little need to publish in foreign languages.

Broader publication—in English, in particular—will be essential to the mutual exchange of ideas and research results now and in the future, as well as to enhancing the reputation of Japanese sociology. To encourage scholars, the Japan Sociological Society has published the *Bibliography of Japanese Sociological Literature in Western Languages* in 1982, 1986, 1990 and 1994 (see Japan Sociological Society 1982, 1986, 1990, 1994). Research Committee meetings of the ISA have been held from time to time in Japan, and occasionally Japanese sociologists have served as ISA Executive Committee members. Japanese sociologists will need to host, and invite their foreign colleagues to, more international meetings and conventions, as occurred in 1991 when the 30th World Congress of

the International Institute of Sociology (IIS; founded in 1893 and the oldest sociological association in the world) met in Kobe, Japan, for the first time in the history of Japanese sociology. In 1998, a Japanese sociologist became the first IIS president ever elected from Asia.

Japan's globalization has fostered opportunities for sociologists through studies of Japanese communities with growing contingents of foreign workers and studies of communities with Japanese administrative and managerial personnel in foreign countries. In these instances, Japanese sociologists are able to examine social, ethnic, and multicultural issues stemming from these circumstances. In this respect, Japanese sociologists are able to contribute to policy implications for community formation in the globalizing environment. Recently, too, the number of Japanese sociologists conducting cross-national and/or foreign area studies has increased (see, e.g., Japan Sociological Society 1997). In 1992, the Japan Sociological Society began publishing the *International Journal of Japanese Sociology*, its only English-language journal and one of the few in Japan to employ a referee system.

Although the language barrier also hinders foreign scholars from coming to Japan to study, again, this is changing, and more and more such activity has been observed recently. In 1998, there were about 340 Japanese sociologists who reported internationally comparative sociological research or foreign area studies as one of their top three research interests, although many of these were not officially collaborating with other nations' researchers. Of these, the most popular areas of emphasis were ethnicity and nationalism; rural sociology; community studies; social history and ethnology; and sociology of the family, culture, and religion. The popular locations for study were Asia, the United States, and western Europe.

There are also less pragmatic reasons why Japanese sociology has not enjoyed a wide-ranging reputation in the rest of the world. Despite a history now spanning more than a century, Japanese sociologists have made few efforts to integrate their considerable empirical research findings with sociological theory. Japanese sociology has had a long history of importing "fashionable" theories from the West. Japanese sociologists with theoretical orientations have tended more toward

social philosophy, with its emphasis on pure theory. Japanese sociological researchers, on the other hand, have tended to limit their studies to specific features of Japanese society, often without sufficient hypothesizing aimed at investigating the underlying social structures and processes thus potentially revealed. This lack of originality, this reluctance to carry through to a complete synthesis of empirical findings and structures and processes, has contributed to Japanese sociology's limited outside appeal. The foreign influences inevitable in and attendant with globalization will no doubt encourage Japanese sociology to finally integrate empirical research results with original thought about Japanese social structures and processes, toward the construction of viable theories of Japanese society as well as society as a globalizing whole. This suggests the need for greater theoretical and methodological training—in the context of their synthesis—in Japanese graduate sociology programs.

While Japanese sociology commands a significant amount of useful empirical data, it is nonetheless disparate. The establishment of a central data archive would be imperative for secondary analysis, graduate training, and empirical study in general. As data gathering becomes increasingly expensive, the usefulness and need for such archives will become even more important.

In conclusion, in terms of the disciplines within Japanese sociology that will take on greater and greater importance, cross-national and foreign area studies will certainly become more popular. Along these same lines, time-trend studies will increase in popularity, in an effort to discover what changes and what does not change in society. Major research funding is likely to continue to center around empirical studies, both within Japan and comparatively with other nations. Finally, whereas Japanese sociologists traditionally have seldom been consulted in the industrial, business, and governmental environments, this will change as Japanese sociologists acquire greater methodological skills and theoretical knowledge, as well as empirical research experience and findings. And, from a strictly practical standpoint, growth in the number of teaching positions in sociology has not kept pace with the growth of sociology's popularity in Japan, nor with the output of sociologists from Japanese and foreign graduate schools, forcing

many to seek teaching positions in junior high and high schools or nonacademic positions in government, research institutes, and industry. This diffusion of the discipline outside the traditional academic environment will no doubt impact Japanese society as a whole, and it could well become a two-way medium for idea exchange.

While Japanese sociology will no doubt become more pervasive and useful in Japanese society, it will nevertheless take time. The same is true outside of Japan, as Japanese sociology becomes more internationalized and demonstrates greater tangible offerings to the sociological community in the rest of the world. Certainly more and better cross-national and foreign area studies will emerge, as this is an area where Japanese sociology has already made particularly beneficial contributions. In turn, in the tradition of comparativists such as Machiavelli, Marx, Weber, and Durkheim, who sought to construct universalized theory based on comparison, Japanese sociology may finally be able to make its mark. Given the methodological expertise demonstrated by Japanese sociology, these efforts may ultimately assist Japanese sociologists in positing social theories of use to the rest of the world's sociological community. Of course, such a vision requires that Japanese sociology strive to broaden its horizons internationally—for Japanese sociologists to recognize that they must cooperate and join with the global sociological community to achieve these objectives.

REFERENCES

Committee on Education for Sociology, Japan Sociological Society 1988 "Research Report on a Survey of the Problem of Post-Graduate Students." *Japanese Sociological Review* 39:314-334. (In Japanese)

Fukutake, Tadashi. 1967 *Japanese Rural Society*. Ithaca, N. Y.: Cornell University Press.

Halmos, P., ed. 1966 "Japanese Sociological Studies." *Sociological Review Monograph*, No. 10. Keele, U.K.: University of Keele.

Hayashi, Chikio. 1998 "The Quantitative Study of National Character: Interchronological and International Perspectives." In Masamichi Sasaki, ed. *Values and Attitudes Across Nations and Time*. Leiden: Brill.

Japan Sociological Society 1982, 1986, 1990, 1994 *Bibliography of Japanese Sociological Literature in Western Languages*. Tokyo: Japan Sociological Society.

— 1997 "Japanese Sociology: Its Continuous Efforts for the Development of Inter-Societal and Inter-Cultural Knowledge and Information." *International Journal of Japanese Sociology* 6:1-162.

Koyano, Shogo 1976 "Sociological Studies in Japan: Prewar, Postwar and Contemporary Stages." *Current Sociology* 24(1):1-196.

Morioka, Kiyomi 1975. *Religion in Changing Japanese Society*. Tokyo: University of Tokyo Press.

Nagai, Michio 1971 *Higher Education in Japan*. Tokyo: University of Tokyo Press.

Odaka, Kunio 1950 "Japanese Sociology: Past and Present." *Social Forces* 28:400-409.

Odaka, Kunio 1975 *Toward Industrial Democracy*. Cambridge, Mass.: Harvard University Press.

Sasaki, Masamichi, and Tatsuzo Suzuki 1991 "Trend and Cross-National Study of General Social Attitudes." *International Journal of Comparative Sociology* 31:193-205.

Shoji, Kokichi 1996 "Institutionalization of Sociology in Japan." In Su-Hoon Lee, ed., *Sociology in East Asia And Its Struggle for Creativity*, Proceedings of the ISA Regional Conference for Eastern Asia, Seoul, Korea, November 22-23, 1996. Madrid, Spain: International Sociological Association, pp. 33-43.

Takata, Yasuma 1922 *Shakaigaku Gairon*. Tokyo: Iwanami Shoten. (In Japanese) Translated into English and published in 1989 as *Principles of Sociology*. Tokyo: University of Tokyo Press/New York: Columbia University Press.

Tominaga, Ken'ichi 1969 "Trends Analysis of Social Stratification and Social Mobility in Contemporary Japan." *The Developing Economy* 7(4):471-498.

Watanuki, Joji 1976 *Politics in Postwar Japanese Society*. Tokyo: University of Tokyo Press.

Yasuda, Saburo 1964 "A Methodological Inquiry into Social Mobility." *American Sociological Review* 29:16-23.

MASAMICHI SASAKI

## JUVENILE DELINQUENCY AND JUVENILE CRIME

The twentieth century ended amid an explosion of violence in all corners of the globe. However, juvenile violence has probably been center stage more in the United States than in any other industrialized nation. As the 1998-1999 school year ended, we had just experienced a massive display

of violence by two young men at a high school in Littleton, Colorado. Both seniors, these two killed twelve of their fellow students and a teacher before turning their assault weapons on themselves. The decade of the 1990s ended in the way it started, with a public grown increasingly apprehensive of its youth. This increase was not limited to the United States; it has been detected in many other countries—England, Spain, France, Italy, and Germany, among others.

It is important to note here also that this anxiety and disquiet induced by youth is not limited to the modern area. Citing the code of Hammurabi, which dates back to 2270 B.C., Regoli and Hewitt (1991) note that “legal prohibitions of specific behavior by juveniles is centuries old” (p.6). Still, they note that in the Middle Ages “little distinction was made between juveniles and adults who were older than 12” (p. 6). In a comment made in 1959, but which is still relevant to those who are intrigued, frightened, or perplexed by the “heedlessness” of today’s youth, Teeters and Matza (1959) stated: “It has always been popular for each generation to believe its children were the worst” (p. 200). We are also reminded by them that “Sir Walter Scott in 1812 deplored the insecurity of Edinburgh where groups of boys between the ages of 12 and 20 years scoured streets and knocked down and robbed all who came in their way” (p. 200). Apropos of delinquency, such remarks underscore the relativity of opinions and the brevity of trends. They also remind us that while juvenile delinquency is a relatively new legal category that subjects children to court authority, it is also a timeless and ubiquitous part of life. Regoli and Hewitt (1991) suggest that postcolonial American delinquency was similar to that found in Spain in the seventeenth century and in Britain in the eighteenth century. By the mid-1800s, teenage gangs were frequently found in the larger cities in the United States. “The habits of hanging out on street corners, verbally abusing pedestrians, and even pelting citizens with rocks and snowballs were among the least threatening of their behaviors. More serious were the violent gangs of juvenile robbers” (p. 7). They also note that “the latter decades of the nineteenth century saw a number of changes in the public’s understanding of the causes of delinquency and [of] appropriate approaches to its control and treatment . . . the

common law distinction between child and adult had changed for purposes of criminal prosecution” (p. 7). The change contrasted greatly with practices in the early part of the century: In America, children as young as three years of age could be brought before the court, while in England a girl of seven was hanged. In Massachusetts, in 1871, 1,354 boys and 109 girls were handled by the courts. Reform schools proliferated during the nineteenth century and were criticized for failing to prevent the apparent increase in delinquency. Reformers—called “child savers”—believed that juveniles required noninstitutional treatment that would reflect the natural family (Platt 1969). This legal and humanitarian concern for the well-being of children led to the establishment of the first juvenile court in Cook County, Illinois, in 1899. By 1925, all but two states had followed the Illinois example. Thus, it seems fair to say that the idea of “juvenile delinquency” is a relatively modern construction, a notion shared by writers such as Gibbons and Krohn (1991), Empey (1982), and Short (1990). The data on delinquency, however, are not limited to the legal status of “juvenile delinquent,” because sociologists are just as interested in unofficial as in official acts of delinquency. More specifically, it is well known that much of the behavior defined by law as delinquent is not detected, not reported, or not acted on by legal agents.

Moreover, different jurisdictions have different legal definitions of delinquency. In the United States, for example, while the statistics defining delinquency are similar in the fifty states and District of Columbia with respect to age and type of offense requiring juvenile court control, there are more differences than similarities. First, laws vary in terms of the age limits of juvenile court jurisdiction: thirty-one states and the District of Columbia set seventeen years of age as the upper age limit, twelve states set sixteen years, six set fifteen years, and one sets eighteen years. Moreover, in many states the delinquency laws empower the juvenile court to remand youths under the maximum juvenile court age to criminal courts. In such cases, the offenses are often those for which adults may be arrested: index crimes (see below). In addition, some states have passed legislation that requires certain cases, such as homicides, or youths charged with other serious offenses, to be dealt with by the criminal court. In these cases, the



juvenile acquires the legal status of criminal. Finally, it should be noted that all U.S. state jurisdictions contain an omnibus clause or provision, referred to as status offenses, that awards the court jurisdiction over youths who have behaved in ways not forbidden by criminal law. While these provisions differ from state to state, it is of interest to note a few examples of these conditions. They include engaging in indecent behavior, knowingly associating with vicious or immoral persons, growing up in idleness or crime, being incorrigible, and wandering the streets at night. Critics note that these behavior categories are so vaguely defined that nearly all youngsters could be subjected to them.

Such different procedures and practices caution us against making easy generalizations both within and between countries when examining official data. Indeed, other shortcomings likewise warn against drawing firm conclusions when unofficial data are examined. Although methodological shortcomings may exist in the study of delinquency, there may be advantages in utilizing all the data of delinquency (official and unofficial) in pursuit of its understanding. Thus, the study of official delinquency data places much of the focus on the actions of official agents of control (the police, the courts), while the study of unofficial—including hidden—delinquency often allows students to examine the processes leading to the behavior. Moreover, as Vold and Bernard (1986) and others have noted, unofficial data, especially self-reports, frequently focus on trivial offenses, while the more serious offenses often do not appear in self-reports but are limited to reports of official agencies.

In sum, our concern here will be to discuss those topics of delinquency that are of the greatest concern: the frequency, severity, and duration of delinquency. Attention will also be devoted to trends. In the following section, the focus is on the extent of delinquent behavior.

#### EXTENT OF AND TRENDS IN DELINQUENCY

In addressing the matter of the extent of delinquency, it is important to note the admonitions of Empey and Erickson (1966), Hirschi (1969), Matza (1964), and others that delinquency is not only transient but also widespread. Many juveniles engage in delinquency only occasionally, but some

engage in it more frequently. Gibbons and Krohn (1991) call delinquency “a sometime thing,” while Matza (1964) describes the process of drifting into and out of delinquency. Moreover, it should be kept in mind that some acts of delinquency are serious acts of criminality and others are petty, trivial acts. As we consider both official and unofficial data on juvenile delinquency and juvenile crime, we will encounter these various clarifying factors.

**Official Delinquency.** The most serious crimes committed by youths and adults in the United States are referred to as index crimes; data on them are compiled by the FBI, based on reports of law enforcement agencies throughout the country. These index crimes, reported in Uniform Crime Reports, are divided into two major types: violent (homicide, rape, robbery, and aggravated assault) and property (burglary, larceny-theft, motor-vehicle theft, and arson). Nonindex offenses are those considered to be relatively petty, such as liquor law violations, disorderly conduct, sex offenses (except forcible rape, prostitution, and commercialized vice), and drug-abuse violations.

Readers are warned that various official reports of offenses should be treated with caution since there are inconsistencies in reporting processes. While great efforts are being made by federal agencies to improve reporting mechanisms, it is well to remember that the number of reporting agencies fluctuates regularly, and the number of cases involving persons with unknown characteristics (such as age, sex, and race) also fluctuates (sometimes by as much as 30 percent). In 1987, the U.S. Department of Justice reported that there were 10,747 contributing agencies covering approximately 84 percent of the U.S. population. This number appears to be higher than the number reporting in 1996. Moreover, the news media have occasionally reported that some jurisdictions have been accused of irregular reporting habits in apparent efforts to demonstrate improved efficiency.

Index crimes are reported annually by the FBI; sometimes estimates for the United States as a whole are derived from a sample of reporting agencies. In 1996 the FBI's Uniform Crime Reports (UCR) indicated that of all arrests in the United States in 1996, approximately 18.9 percent were of youths under eighteen years of age; 18.7 percent of the arrests for violent crimes in 1996

(murder, manslaughter, forcible rape, robbery, and aggravated assault) were of persons under eighteen years of age. For property crimes, youth under 18 years of age comprised 35 percent of offenders. For index crimes as a whole, the FBI Uniform Crime Report indicates that 30.9 percent were contributed by persons under 18 years of age in 1996. (U.S. Department of Justice 1996).

Perhaps it is surprising to learn that over the last quarter of the twentieth century, contrary to fears, delinquency has declined. According to the *Sourcebook of Criminal Justice Statistics* for 1997, in 1971 arrests of youths under age 18 comprised 45 percent of all arrests for index crimes reported to the UCR program, 23 percent of violent crime arrests, and 51 percent of property crime arrests. In 1996 persons under age 18 comprised 31 percent of index offense arrests, 18.8 percent of violent crime arrests, and just 35 percent of property crime arrests. Thus juvenile crime, as measured by arrest data, has decreased during this quarter-century.

During this same period, a number of criminologists have argued that fluctuations in the delinquency crime rates are due to changes in the age structure of the population. The UCR presented data showing a substantial decrease in delinquency arrests between 1979 and 1988, following an increase in such arrests between 1965 and 1977. Kratcoski and Kratcoski (1990) suggest that this dramatic change could be due to "movement of the 'baby boom' segment of the population through the high offense years during the mid-70s followed by a reduction in the under 18 population since 1978" (p. 15). Gibbons and Krohn (1991) also suggest that such trends may be due to shifts within the delinquency-eligible youth group. Further, the latter authors suggest that such trends may reflect increased or decreased concern about youthful misconduct.

There was apparently a substantial increase in juvenile crime between 1988 and 1991, followed by another dip in rates after 1991. Blumstein (1995) suggests that a projected increase in the proportion of the population in the high-crime-rate age group will probably lead to higher crime rates by early in the twenty-first century. Still, the FBI UCR for 1996 suggests that there has been a strong decline in juvenile delinquency in recent years.

**Juvenile Crimes of Violence.** In contrast to this apparent improvement, an examination of murder and non-negligent manslaughter shows a different picture. While the national murder rate in 1996 (7.4 per 100,000 inhabitants) was down 11 percent since 1987, the proportion of murder and non-negligent manslaughter offenders who were under age 18 increased from 9.7 percent in 1987 to a whopping 15.2 percent in 1996, an increase of about 50 percent. Editors of the *Sourcebook of Criminal Justice Statistics* for 1997, and other analysts such as Conklin (1998), Gall and Lucas (1996), and Blumstein (1995) have commented on the increasingly serious nature of crimes by juveniles.

Conklin (1998), reviewing data presented by Fox (1995) and by Blumstein (1995), says that between 1985 and 1993 "the homicide arrest rate for people twenty-five and over declined by 20 percent, while the rate for eighteen-to-twenty-four year olds increased by 65 percent and the rate for fourteen-to-seventeen year olds rose by 165 percent" (p. 124). An increasing number of researchers fear that there will be a continued surge in violent crimes by the under-eighteen population.

This increase in youth violence in the United States has apparently been mirrored in a number of industrialized nations. In a study of juvenile violence in Great Britain, Oliver (1997) states that the number of violent juveniles increased by 34 percent between 1987 and 1993, an increase he attributes to the growth of inequality and family stress in Britain. In France, Bui-Trong (1996) has written about the recent escalation of juvenile violence, noting especially the increased severity of scale. This author, noting a decline in family morals, predicts that this problem can be expected to worsen. During the 1980s even Japan experienced an increased rate of juvenile delinquency, which some observers attributed to the stress of competition for academic success (Conklin 1998).

School violence has been most dramatic, of course, in the United States and is a focus of great national attention. Presenting data gathered by the National School Safety Center, the *Boston Globe* reported in 1999 that since the 1992-1993 school year there had been 248 deaths from violent acts in schools. ("U.S. Police Chiefs" 1999). Although many school shootings and other violent acts have not resulted in deaths, they *have* resulted in injuries and heightened fears. The 248 school deaths

between 1992 and 1999 included cases of multiple deaths in seventeen states. Many of these incidents involved the use of assault weapons. This increase in school violence has elicited a spate of explanations, including weakened parental supervisions, ostracism by schoolmates, bullying and its effects (Farrington 1993), the easy availability of guns, and the poor examples set by our leaders. The search for the causes of violence among the young will, no doubt, have to be accompanied by an examination of violence in the adult population as well. Certainly the official data on violence show that the population over age eighteen is also engaged in a substantial amount of violent behavior.

**Cohort Study.** One of the first longitudinal studies of delinquency was conducted by Wolfgang and colleagues (1972) in Philadelphia. They traced the police contacts of all boys born in 1945 who lived in the city between their tenth and eighteenth birthdays. One of their aims with this cohort of 9,945 was to trace the volume and frequency of delinquent careers up to age eighteen. They found that 35 percent of these boys (3,475) were involved with the police at least once between their tenth and eighteenth birthdays. Of these 3,475 boys with police contacts, 54 percent were repeaters. The total number of delinquent events (offenses) for the 3,475 delinquent boys amounted to 10,214 through age seventeen. It is clear that the number of offenses far outnumbers the number of offenders in the cohort. One must note, then, that longitudinal (i.e., over an extended period) studies of delinquents yield data with important differences from those obtained when cross-sectional (i.e., single point in time) studies of persons arrested are conducted. Examples of other longitudinal studies include the Provo Study, the Cambridge-Somerville Study, the Vocational High Study, and, in Britain, the National Survey of Health and Development.

**Self-Reports: Offender Reports and Victim Reports.** Official reports of crime and juvenile delinquency have been criticized for years because they are widely believed to underreport the volume of offenses. Moreover, many scholars, especially those with a conflict perspective, believed that official reports underreported middle-class crime and delinquency. In an effort to detect "hidden" delinquency, sociologists developed a technique designed to produce a more accurate picture. The technique used by Short and Nye

(1958) in a number of studies of hidden delinquency consisted of having juveniles in a school or other population complete questionnaires on the extent to which they engaged in law-violating behavior. They found that delinquency was widespread throughout the juvenile population. Subsequently, Williams and Gold (1972) and Empey and Erickson (1966) embarked on studies employing self-reports. These writers found that 88 percent and 92 percent of their study groups, respectively, had engaged in violations. Hindelang and colleagues (1981) present a similar volume of law-violating behavior in their Seattle study. Criticisms of the self-report method followed many of these studies, centering on issues of respondent misrepresentation, respondent recall, and inappropriateness of study groups. A major criticism (by Nettler, 1984, for example) was that self-reports elicited admission of only minor or petty infractions for the most part. Because of its obvious utility, the self-report technique has been greatly improved in recent years, becoming an important, if not the dominant, method of measurement in studies focusing on the extent and cause of delinquency.

A number of students of delinquency agree that many of the improvements in self-report studies have been contributed by Elliott and Ageton (1980): (Bartol and Bartol 1989; Gibbons and Krohn 1991; Regoli and Hewitt 1991). Elliott and colleagues have created a panel design that employs periodic interviews instead of questionnaires. The study, called the National Youth Survey (NYS), utilized a 5-year panel design with a national probability sample of 1,726 adolescents aged eleven to seventeen and covered more than a hundred cities and towns in the United States. In contrast with earlier self-report studies, Elliott asked his respondents about a full range of activities designed to get at serious as well as minor infractions. In addition, his respondents were asked whether they were caught when engaging in delinquent and criminal activities.

Another attempt to ascertain the volume of delinquency in the United States is represented by the National Crime Survey (NCS). This survey, an effort to determine the extent of victimization in the population of the United States, was begun in 1973 after an initial study sponsored by the President's Commission on Law Enforcement and Administration of Justice in 1967. Interviews are conducted semiannually by the Bureau of the

Census with a large national sample of 60,000 households (Bartol and Bartol 1989). The survey was intended to supplement the Uniform Crime Report data and measures the extent to which persons and households are victims of rape, robbery, assault, burglary, motor-vehicle theft, and larceny. Binder and colleagues (1988) note that one of the major findings of the 1967 study of victims by the President's Commission was the revelation that "actual crime was several times that indicated in the UCR" (p. 34). In the current victim interviews, if the respondent has been victimized, he or she will be asked questions about both self-characteristics and characteristics of the offender. Binder and colleagues warn about the difficulties of age discrimination by a victim under stress and suggest that this method cannot be relied on too heavily in measuring delinquency.

Nevertheless, Laub (1983) has found NCS data useful in addressing the issue of the extent and change in delinquency volume. In an analysis of NCS data obtained between 1973 and 1980, he found no increase in juvenile crime over those years. He further noted that data from the National Center for Juvenile Justice supported this conclusion. It would seem, then, that the NCS data, UCR data, and juvenile court data have been fairly consistent regarding the volume of delinquency. While self-report data indicate that violations are consistently widespread in American society, it is important to note that these reports involve primarily minor violations. Indeed, to the extent that almost everyone engages in minor violations, it may make sense to focus mainly on serious violations. Nettler (1984), noting that self-report studies find a large number of minor infractions, suggests that such violators are best described as lying on a continuum rather than as being "delinquent" or "nondelinquent." The cohort studies of Wolfgang and colleagues (1972) show that only a small proportion of the study groups were involved in serious violations.

#### FACTORS RELATED TO DELINQUENCY

**Age and Gender.** In the United States, Britain, and other European countries where delinquency is recognized and studied, there is general agreement that it peaks in adolescence (ages fifteen to eighteen) rather than in childhood. (This is not to

say, however, that delinquency is not on the increase among younger children. A study by the FBI in 1990 found that the arrest rate for rape among males aged twelve years and under had more than tripled since 1970 (Parmley 1991). The UCR shows that 18.9 percent of all arrests in 1996 involved persons under eighteen years of age. This age group accounted for 31 percent of arrests for index crimes, however. Male arrests peaked at age eighteen in 1996 and female arrests at age sixteen. Earlier studies and analyses by Empey and Erickson (1966), Wolfgang (1983), and Braithwaite (1981) are consistent with this picture, with age sixteen being the peak year for juvenile misconduct.

It has consistently been found that males outnumber females in UCR arrest data. Thus in both 1988 and 1989, among those under eighteen, males were arrested four times more frequently than females. For those under eighteen in both 1988 and 1989, males outnumbered females 8-to-1 in arrests for violent offenses. In 1990, Short pointed out that the gender ratio had declined substantially. Between 1980 and 1989 violent crime increased by 4.5 percent for males under eighteen and by 16.5 for comparable females. Considering all arrests for males and females under eighteen, males showed a *decline* of 8.5 percent between 1980 and 1989, while females showed an *increase* of 1.1 percent for the same period.

Indeed, as reported in the Uniform Crime Reports, female delinquency, compared to male delinquency, has been rising steadily since the 1960s. In 1960, there was a 6-to-1 ratio of male to female juvenile arrests. The ratio has declined steadily since then and in 1987 was 3.4-to-1. In 1996 the FBI data revealed a male-female ratio of just 3-to-1.

Moreover, the ratio for violent offenses has dropped from 8-to-1 in 1989 to 5.5-to-1 in 1996, according to FBI data reported in the *Sourcebook* for 1997. Finally, it must be mentioned that FBI data reveal a decline of 15.3 percent in the murder and manslaughter charges for males under age 18 between 1995 and 1996 but reveal no change for females in that age group. It should also be noted, however, that for those aged thirteen to nineteen, 92 percent of murder offenders in the UCR in 1996 were males and only 8 percent were females. Also sobering is the fact that 29 percent of all

murder victims in the under-eighteen age group were females in 1996.

Perhaps Hagan's power-control theory is relevant here. Hagan and colleagues (1987) suggest that child-rearing styles in the home (the power structure) are determined in part by the nature of the parents' occupations. The two main types of child-rearing styles are patriarchal and egalitarian. The two types of occupations are command (managerial) and obey (subject to others' authority). In the egalitarian family where both parents work in authority positions, the mother's authority means she has a substantial amount of power in the home, and this leads to daughters having increased freedom relative to sons. This situation is reversed in patriarchal families, which are controlled by fathers and sons. In the egalitarian family, the adolescent daughter has an increased willingness to take risks. Hagan assumes that willingness to take risks is a fundamental requirement for delinquency. He also predicts that female delinquency will be high in mother-only homes. The absence of a father leaves a void in male power, allowing the adolescent girl more freedom, greater risk taking, and an increased tendency to deviate. The theory needs to be tested more fully.

With respect to self-reports, the reports of the NYS suggest that gender was not strongly related to involvement in delinquent acts. Although males admitted to more infractions than did females, the differences were much less pronounced than those seen in the UCR. Again, it should be emphasized that efforts are being made to enhance the ability of self-report studies to elicit information on more serious infractions.

**Race and Class.** UCRs for 1985 and 1989 present data on arrests by race in the United States. Among persons arrested and under eighteen years of age, the number of blacks increased from 23 percent in 1985 to 28 percent in 1989, but declined slightly to 27 percent in 1996; the number of whites decreased from 75 percent in 1985 to 70 percent in 1989 and remained at 70 percent in 1996. The remainder of the arrests were categorized by race as Native Americans, Asians, Pacific Islanders, and Alaskan Natives. The proportion of violent index crimes committed by blacks under age eighteen increased from 52 percent in 1985 to 53 percent in 1989 but declined to 47 percent in 1996. The comparable proportions for whites were

46 percent, 45 percent, and 50 percent. Also of interest are the figures for murder and non-negligent homicide. Here, blacks under eighteen accounted for 51 percent, 61 percent, and 57.5 percent of these crimes in 1985, 1989, and 1996, respectively; whites accounted for 48 percent, 37 percent, and 39 percent during the same three years. This slight decline in the proportion of murder and non-negligent manslaughter accounted for by blacks under eighteen is accompanied by the fact that the number of such crimes increased by 25 percent for blacks and by 18 percent for whites under eighteen during the period between 1988 and 1996. This is a decided improvement over the period between 1985 and 1989, when there was a doubling of the number of such crimes for blacks and a 30 percent increase for whites under age eighteen. Aggravated assaults also increased substantially during this period for both whites and blacks under eighteen, increasing for blacks by 25 percent and for whites by 50 percent during the eight-year period ending in 1996. Forcible rapes by blacks under age eighteen showed an improving picture, declining by 12 percent between 1988 and 1996. On the other hand, forcible rapes by whites under age eighteen increased by 15 percent during this time frame. Overall, despite short-term improvements in the murder rate for those over age 18, there is no denying the fact that serious and violent crime increased among those under age 18 between 1988 and 1996. This violence has been growing among white youth as well as black youth.

William J. Wilson (1996) and other sociologists have long suggested that social class and inequality are important concepts in the attempt to understand the increase in violence among blacks. The move of businesses to suburban locations with consequent loss of job opportunities, educational inequities, and the growth of female-headed households have combined to limit the coping resources in inner cities where most black youth reside. Wilson suggests that the resulting stresses hit lower-class blacks particularly hard and contribute to the high rates of violent crime in these areas. In addition, Bernard (1990) suggests that anger is reinforced by the social isolation that comes from being confined to communities that have high rates of crime. Policies by local departments of justice with respect to police routines, bail practices, and sentencing as these apply to

blacks have also come in for considerable comment (Teele 1970b). In 1999 news media reported that the U.S. Department of Justice was investigating allegations of racial profiling—targeting minorities for an unjustified amount of police attention—by police in New Jersey, Michigan, and Florida (“US Police Chiefs” 1999).

While there are those who say that the nation has not paid enough attention to the causes of black teenage violence, an increasing number of analysts say that we have paid a lot of, perhaps too much, attention to black violence in urban areas while ignoring the rising tides of violence by white youths in suburban and rural areas. While income inequality has seemed appropriate to an explanation of urban violence, it may appear less relevant in the analysis of violent behavior in our affluent suburban areas. This situation brings to mind Toby’s (1967) classic work on delinquency in affluent society as well as Durkheim’s (1951) discussion of rising expectations in his theory of anomie.

In Toby’s case, he was trying to account for the rise in theft crimes in a variety of countries. Although he took special note of adolescent crime in industrial countries such as Japan, Sweden, and Great Britain, he also included developing countries such as Nigeria and India in his analysis. Toby suggests that the resentment of poverty is likely to be greater among the relatively poor in an affluent society than among the poor in a poor society. He suggests, however, that envy is at work not only in the more affluent societies, such as Japan, but also in countries with rising standards of living, such as India and Nigeria. Moreover, he suggests that not only adults but also the young are subject to rising expectations. It may be that relative deprivation can account for at least some of the rise in youth violence in the United States. Toby’s work also suggests that the presence of such envy could be heightened where the inhibiting effect of schools or families or jobs is missing in either affluent or developing societies. Here Toby’s thesis greatly resembles control theory. The increase in single-parent families among blacks, persisting educational inequalities for blacks, and chronic employment problems for black youth may tend to lessen social controls and could be factors in the greater increase in violence among them.

Toby’s discussion of “resentment” seems somehow applicable to the school violence discussed

earlier, but with a different twist. There seems present in these cases something besides income inequality—especially since in many of these instances the youth are from apparently affluent families, as in the Littleton, Colorado, case. While assault weapons were certainly available and parents seemed loathe to invade their sons’ privacy, there was at school, perhaps, a certain intolerable inequality that alienated the assaulters and pulled them into circles of violent schemes, feelings, and, eventually, behavior. Most of the youths—usually white—felt rejected by their peers, their teachers, their parents, or their girlfriends. In Colorado they resented the athletes and those who were generally in the “in crowd.” There apparently were feelings of jealousy toward those who did not invite them to join their group. While such feelings, involving relative deprivation, seem to be far removed from considerations of wealth, they fit the phenomenon of inequality. This form of inequality is probably anathema to children who are deprived of neither their basic needs nor considerable luxuries.

Social class has been by far the most controversial of all the factors studied in connection with juvenile delinquency. The argument seems to revolve around both method and theory. Some argue about the impact of social class, others debate the measurement of social class, and a few argue about both. Several writers have attempted to review the research on class and delinquency or crime. Tittle and colleagues (1978), noting that nearly every sociological theory of crime or delinquency had class as a key factor, reviewed thirty-five such studies. Their findings suggested that the class and crime–delinquency connection might be a “myth” because the relationship could not be confirmed empirically.

Subsequently, Braithwaite (1981) criticized Tittle’s study not only as incomplete but also as having come to the wrong conclusions. He reviewed fifty-three studies that used official data and forty-seven studies that used self-reports in the study of delinquency. Braithwaite forcefully argues that the class–crime relationship is no myth. Of the studies using official records, Braithwaite found that the vast majority (forty-four) showed lower-class juveniles to have substantially higher offense rates than middle-class juveniles. Of the forty-seven self-report studies, he concluded that eighteen found lower-class juveniles reported higher

levels of delinquent behavior, seven reported qualified support for the relationship, and twenty-two found no relationship. Braithwaite is critical of self-report studies when (1) they do not closely examine the lowest group on the social-class continuum (the lumpen proletariat) and (2) they do not include serious offenses and chronicity in their data gathering. While the argument may continue, Braithwaite and others seem to be less critical of self-report studies when they correct these apparent shortcomings.

Apparently the work of Elliott and Ageton (1980) has done much to defuse this issue. They found, for example, that the relationship between class and self-reported delinquency is totally a consequence of the difference between the lowest class group and the rest of the sample, with no difference between the working and middle classes. Writers such as Messner and Krohn (1990), Hagan and Palloni (1990), and Colvin and Pauly (1983) have apparently profited from these debates; their work shows an inclination to refine the "objective" measure of class, using insights from conflict theory as they formulate explanations of delinquency. Indeed, it is safe to say that social class is alive and well, but it is more broadly conceptualized now; many of the new theories include patterns of child rearing, job experiences, and family structure that are incorporated into the framework of a more radical neo-Marxist perspective. The effort by sociologists in the United States and in other countries to better understand juvenile delinquency appears to have entered a new and more urgent phase.

## REFERENCES

- Bartol, Curt, and Anne Bartol 1989 *Juvenile Delinquency: A System Approach*. Englewood Cliffs, N.J.: Prentice-Hall.
- Bernard, Thomas J. 1990 "Angry Aggression Among the 'Truly Disadvantaged.'" *Criminology* 28:73-96.
- Binder, Arnold, Gilbert Geis, and Bruce Dickson 1988 *Juvenile Delinquency: Historical, Cultural, Legal Perspectives*. New York: Macmillan.
- Blumstein, Alfred 1995 "Violence by Young People: Why the Deadly Nexus?" *National Institute of Justice Journal* (August):2-9.
- Braithwaite, John 1981 "The Myth of Social Class and Criminality Reconsidered." *American Sociological Review* 46:36-57.
- Bui-Trong, Lucienne 1996 "Resurgence de la Violence en France." *Futuribles* 206:5-20.
- Colvin, Mark, and John Pauly 1983 "A Critique of Criminology: Toward an Integrated Structural-Marxist Theory of Delinquency Production." *American Journal of Sociology* 89:513-551.
- Conklin, John E. 1998 *Criminology*, 6th ed. Boston: Allyn and Bacon.
- Durkheim, Emile 1951 *Suicide*, trans. J.A. Spaulding and George Simpson. New York: Free Press.
- Elliott, Delbert, and Suzanne Ageton 1980 "Reconciling Race and Class Differences in Self-Reported and Official Estimates of Delinquency." *American Sociological Review* 45:95-110.
- Empey, Lamar T. 1982 *American Delinquency*, rev. ed. Homewood, Ill.: Dorsey.
- , and Maynard Erickson 1966 "Hidden Delinquency and Social Status." *Social Forces* 44:546-554.
- Farrington, David 1993 "Understanding and Preventing Bullying." In Michale Tonry, ed., *Crime and Justice: A Review of Research*, vol. 17. Chicago: University of Chicago Press.
- Fox, James A. 1995 "A Disturbing Trend in Youth Crime." *Boston Globe* (June 1):9.
- Gall, Timothy, and Daniel Lucas (eds.) 1996 *Statistics on Weapons and Violence*. New York: Gale Research.
- Gibbons, Don, and Marvin Krohn 1991 *Delinquent Behavior*, 5th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Hagan, John, and Alberto Palloni 1990 "The Social Reproduction of a Criminal Class in Working-Class London, Circa 1950-1980." *American Journal of Sociology* 96:265-299.
- , J. Simpson, and A. R. Gillis 1987 "Glass in the Household: A Power-Control Theory of Gender and Delinquency." *American Journal of Sociology* 92:788-816.
- Hindelang, M. J., Travis Hirschi, and Joseph Weis 1981 *Measuring Delinquency*. Beverly Hills, Calif.: Sage.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- Kratcoski, Peter, and Lucille Dunn Kratcoski 1990 *Juvenile Delinquency*, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Laub, John 1983 "Trends in Serious Juvenile Crime." *Criminal Justice and Behavior*. 10:485-506.
- Matza, David 1964 *Delinquency and Drift*. New York: Wiley.
- Messner, Steven, and Marvin Krohn 1990 "Class, Compliance Structures, and Delinquency: Assessing Integrated Structural-Marxist Theory." *American Journal of Sociology* 96:300-328.

- Nettler, Gwynn 1984 *Explaining Crime*, 3rd ed. New York: McGraw-Hill.
- Oliver, James 1997 *Juvenile Violence in a Winner-Loser Culture: Socio-Economic and Familial Origins of the Rise of Violence Against the Person*. London: Free Association Books.
- Parmley, Suzette 1991 "Children Who Molest Children." *Boston Globe* (January 14):25.
- Platt, Anthony M. 1969 *The Child Savers*. Chicago: University of Chicago Press.
- Regoli, Robert, and John Hewitt 1991 *Delinquency in Society*. New York: McGraw-Hill.
- Short, James. 1990 *Delinquency in Society*. Englewood Cliffs, N.J.: Prentice-Hall.
- , and F. Ivan Nye 1958 "Extent of Unrecorded Juvenile Delinquency: Tentative Conclusions." *Journal of Criminal Law, Criminology and Police Science* 49:296-302.
- Sourcebook of Criminal Justice Statistics 1997 Washington, D.C.: U.S. Government Printing Office.
- Teele, James E. 1970 "Social Pathology and Stress." In Sol Levine and Norman Scotch, eds., *Social Stress*. Chicago: Aldine Press.
- Teeters, Negley, and David Matza 1959 "The Extent of Delinquency in the United States." *Journal of Negro Education* 28:200-213.
- Tittle, Charles R., Wayne Vilemez, and Douglas Smith 1978 "The Myth of Social Class and Criminality: An Empirical Assessment of the Empirical Evidence." *American Sociological Review* 43:643-656.
- Toby, Jackson 1967 "Affluence and Adolescent Crime." 1967 *President's Commission on Law Enforcement and Administration of Justice: Task Force Report on Juvenile Delinquency and Youth Crime*. Washington, D.C.: U.S. Government Printing Office. Reprinted in James E. Teele, ed., *Juvenile Delinquency: A Reader*. Itasca, Ill.: Peacock, 1970.
- U.S. Department of Justice 1997 *Crime in the United States, 1996*. Washington, D.C.: U.S. Government Printing Office.
- "U.S. Police Chiefs Examining Race, Traffic Stops" 1999 *Boston Globe* (April 10):11.
- Vold, George, and Thomas Bernard 1986 *Theoretical Criminology*. New York: Oxford University Press.
- Williams, Jay, and Martin Gold 1972 "From Delinquent Behavior to Official Delinquency." *Social Problems* 20:209-229.
- Wilmot, Peter 1966 *Adolescent Boys in East London*. London: Routledge and Kegan Paul.

Wilson, William J. 1996 *When Work Disappears*. New York: Random House.

Wolfgang, Marvin 1983 "Delinquency in Two Birth Cohorts." *American Behavioral Scientist* 27:75-86

———, Robert Figlio, and Thorsten Sellin 1972 *Delinquency in a Birth Cohort*. Chicago: University of Chicago Press.

JAMES E. TEELE

## JUVENILE DELINQUENCY, THEORIES OF

The topic of juvenile delinquency is a fertile area for construction of sociological theory. Three major sociological traditions, including structural functionalism, symbolic interactionism, and conflict theory, contribute to the explanation of delinquency. Much of the work in this area seeks to explain why officially recorded delinquency is concentrated in the lower class, or in what is today more often called the underclass. This entry considers the most prominent theories of delinquency under the theoretical rubrics noted above.

### STRUCTURAL FUNCTIONALISM AND DELINQUENCY

Structural-functional theories regard delinquent behavior as the consequence of strains or breakdowns in the social processes that produce conformity. These theories focus on institutions, such as the family and school, that socialize individuals to conform their behavior to values of the surrounding society and on the ways in which these institutions can fail in this task. Wide agreement or consensus is assumed about which behaviors are valued and disvalued in society. The question structural-functional theories try to answer is: Why do many individuals during their adolescence behave in ways that challenge this consensus? That is, why do many adolescents violate behavioral norms that nearly all of us are assumed to hold in common?

**Anomie Theory.** The roots of functional theory are found in Durkheim's notion of *anomie* ([1897] 1951). To Durkheim, this term meant an absence of social regulation, or normlessness. Merton (1938, 1957) revived the concept to describe the consequences of a faulty relationship between goals and the legitimate means of attaining them. Merton



emphasized two features of social and cultural structure: culturally defined goals (such as monetary success) and the acceptable means (such as education) to their achievement. Merton argued that in our society success goals are widely shared, while the means of or opportunities for attaining them are not.

Merton's theory is used to explain not only why individual adolescents become delinquents but also why some classes are characterized by more delinquency than others. Since members of the lower- or underclass are assumed to be most affected by the disparity between the goals and the means of attaining success, this class is expected to have a higher rate of delinquent behavior. Merton outlined a number of ways individuals adapt when faced with inadequate means of attaining their goals. Among these, *innovation* revolves substituting illegitimate for legitimate means to goal attainment; it is the resort to this adaptation that is thought to account for much theft among adolescents from the underclass.

**Subcultural Theory.** Group-based adaptations to the failure to attain success goals involve the *delinquent subculture*. Cohen (1955) suggests that children of the underclass, and potential members of a delinquent subculture, first experience a failure to achieve when they enter school. When assessed against a "middle-class measuring rod," these children are often found lacking. A result is a growing sense of "status frustration." Underclass children are simply not prepared by their earliest experiences to satisfy middle-class expectations. The delinquent subculture therefore emerges as an alternative set of criteria or values that underclass adolescents can meet.

Cohen argues that these subcultural values represent a complete repudiation of middle-class standards: the delinquent subculture expresses contempt for a middle-class lifestyle by making its opposite a criterion of prestige. The result, according to Cohen, is a delinquent subculture that is "nonutilitarian, malicious, and negativistic"—an inversion, of middle-class values. Yet this is only one possible type of subcultural reaction to the frustration of failure. As we see next, many subcultural responses are elaborated in the theoretical tradition of structural functionalism.

**Differential Opportunity Theory.** Cloward and Ohlin (1960) argue that to understand the

different forms that delinquent and ultimately criminal behavior can take, we must consider the different types of illegitimate opportunities available to those who seek a way out of the underclass and where these opportunities lead. Different types of community settings produce different subcultural responses. Cloward and Ohlin suggest that three types of responses predominate, each one leading to its own respective subculture: a stable criminal subculture, a conflict subculture, and a retreatist subculture.

The *stable criminal subculture* offers, as its name suggests, the most promising (albeit still illegitimate) prospects for upward economic mobility. According to Cloward and Ohlin, this subculture can emerge only when there is some coordination between those in legitimate and in illegitimate roles—for example, between politicians or police and the underworld. One pictures the old-style political machine, with protection provided for preferred types of illegal enterprise. Only in such circumstances can stable patterns be established, allowing opportunities for advancement from adolescent to adult levels of the criminal underworld. When legitimate and illegitimate opportunity structures are linked in this way, the streets become safe for crime, and reliable upward-mobility routes can emerge for aspiring criminals.

Violence and conflict, on the other hand, disrupt both legitimate and illegitimate enterprise. When both types of enterprises coexist, violence is restrained. However, in the "disorganized slum," where these spheres of activity are not linked, violence can reign uncontrolled. Cloward and Ohlin see these types of communities as producing a *conflict subculture*. A result of this disorganization is the prevalence of adolescent street gangs and their violent activities, making the streets unsafe for more profitable crime.

The *retreatist subculture* includes adolescents who fail in their efforts in both the legitimate and illegitimate opportunity structures. These "double failures" are destined for drug abuse and other forms of escape.

Cloward and Ohlin's theory played a role in encouraging the Kennedy and Johnson administrations of the 1960s to organize the American War on Poverty, which attempted to open up legitimate opportunities for youth and minorities in the underclass (see Moynihan 1969). However,

another important variant of structural-functional theory argued that the most important cause of delinquency was not a strain between goals and means but rather a relative absence of goals, values, commitments, and other sources of social control.

**Social Disorganization Theory.** The earliest North American efforts to explain crime and delinquency in terms of social control focused on the absence of social bonds at the community level. Entire neighborhoods were seen as being socially disorganized, as lacking the cohesion and constraint that could prevent crime and delinquency. This work began in the late 1920s, when Clifford Shaw and Henry McKay (1931, 1942) sought to identify areas of Chicago that were experiencing social disorganization. They explored the process that characterized these communities. What they found were indications of what they assumed to be social disorganization— truancy, tuberculosis, infant mortality, mental disorder, economic dependency, adult crime, and juvenile delinquency. In Chicago, the rates of these conditions were highest in the slums near the city center; they diminished in areas farther away from the center. Since these problems were assumed to be contrary to the shared values of area inhabitants, they were taken as indications that these areas were unable to realize the goals of their residents. In other words, they were taken as indicators of social disorganization.

Shaw and McKay also attempted to determine the sorts of community characteristics that were correlated with delinquency so that they could infer from these characteristics what the central components of social disorganization were and how they caused delinquency. Three types of correlates were identified: the economic status of the community, the mobility of community residents, and community heterogeneity. The implication was that poverty, high residential mobility, and ethnic heterogeneity led to a weakening of social bonds or controls and, in turn, to high rates of delinquency. All of this was being said of the neighborhoods Shaw and McKay studied; it was left to later theories to spell out the meaning of weakened neighborhood bonds or controls for individuals.

**Control Theory.** At the level of individuals, to have neither goals nor means is to be uncommitted and thus uncontrolled. Hirschi (1969) has

argued that the absence of control is all that really is required to explain much delinquent behavior. There are other types of controls (besides *commitment* to conformity) that may also operate: *involvement* in school and other activities; *attachments* to friends, school, and family; and belief in various types of values and principles. Hirschi argues that delinquent behavior is inversely related to the presence of these controls. Alternatively, as these controls accumulate, so too does conformity. According to control theory, the more committed, attached, involved, and believing individuals are, the greater is their bond to society. Again, Hirschi's point is that no special strain between goals and means is necessarily required to produce delinquent behavior; all that is required is the elimination of the constraining elements of the social bond.

In each of the theories that we have considered thus far, values or beliefs play some role in causing delinquency. It is argued that the presence of success goals or values without the means to obtain them can produce deviant behavior, as can the absence of these goals or values in the first place. It is an emphasis on these values, and the role of the school and family in transmitting them, that ties the structural-functional theories together.

### SYMBOLIC INTERACTIONISM AND DELINQUENCY

Symbolic-interactionist theories of delinquency are concerned less with values than with the way in which social meanings and definitions can help produce delinquent behavior. The assumption, of course, is that these meanings and definitions, these symbolic variations, affect behavior. Early versions of symbolic-interactionist theories focused on how adolescents acquired these meanings and definitions from others, especially peers; more recently, theorists have focused on the role of official control agencies, especially the police and courts, in imposing these meanings and definitions on adolescents. The significance of this difference in focus will become apparent as we consider the development of the symbolic-interactionist tradition.

**Differential Association Theory.** Edwin Sutherland (1939, 1949) anticipated an emphasis of the symbolic-interactionist perspective with his early use of the concept of *differential association*. This concept referred not only to associations among

people but also, and perhaps even more important, to associations among ideas. Sutherland's purpose was to develop a general theory that explained delinquency as well as adult criminality. He argued that people violate laws only when they define such behavior as acceptable and that there is an explicit connection between people and their ideas (that is, definitions). So, for example, delinquent behavior is "learned in association with those who define such behavior favorably and in isolation from those who define it unfavorably," and this behavior occurs when "the weight of the favorable definitions exceeds the weight of the unfavorable definitions."

Although Sutherland intended his theory to be general and explicitly to include the explanation of delinquency, his best-known applications of the theory were in his famous studies of professional theft and white-collar crime. Nonetheless, Sutherland's emphasis on white-collar illegality was important for the study of delinquency because it stressed the ubiquity of criminality, and, as we see next, it helped to mitigate delinquency theory's preoccupation with underclass delinquency.

**Neutralization Theory.** While most of the theories we have considered to this point portray the delinquent, especially the underclass delinquent, as markedly different from "the rest of us," Sykes and Matza (1957, 1961) follow Sutherland's lead in suggesting that the similarities actually outnumber the differences. Their argument is based in part on the observation that underclass delinquents, like white-collar criminals, usually exhibit guilt or shame when detected violating the law.

Sutherland had argued that individuals become white-collar criminals because they are immersed with their colleagues in a business ideology that defines illegal business practices as acceptable. Sykes and Matza (1957) argue that the delinquent, much like the white-collar criminal, drifts into a deviant lifestyle through a subtle process of justification. "We call these justifications of deviant behavior techniques of neutralization," they write, "and we believe these techniques make up a crucial component of Sutherland's definitions favorable to the violation of law" (p. 667).

Sykes and Matza list four of these *neutralization techniques*: denial of responsibility (e.g., blaming a bad upbringing), denial of injury (e.g., claiming that the victim deserved it), condemnation of

the condemners (e.g., calling their condemnation discriminatory), and an appeal to higher loyalties (e.g., citing loyalty to friends or family as the cause of the behavior). Sykes and Matza's point is that delinquency in the underclass, as elsewhere, is facilitated by this kind of thinking. A question lingered, however: Why are these delinquencies of the underclass more frequently made the subjects of official condemnation?

**Labeling Theory.** Franklin Tannenbaum (1938) anticipated a theoretical answer to this question. He pointed out that some aspects of juvenile delinquency—the play, adventure, and excitement—are a normal part of teenage street life and that, later in their lives, many nostalgically identify these activities as an important part of their adolescence. But others see such activities as a nuisance or as threatening, so they summon the police.

Tannenbaum's concern is that police intervention begins a process of change in the way the individuals and their activities are perceived. He suggests that there is a gradual shift from defining specific acts as evil to defining the individual as evil. Tannenbaum sees the individual's first contact with the law as the most consequential, referring to this event as a "dramatization of evil" that separates the child from his or her peers for specialized treatment. Tannenbaum goes on to argue that this dramatization may play a greater role in creating the criminal than any other experience. The problem is that individuals thus singled out may begin to think of themselves as the type of people who do such things—that is, as delinquents. From this viewpoint, efforts to reform or deter delinquent behavior create more problems than they solve. "The way out," Tannenbaum argues, "is through a refusal to dramatize the evil." He implies that the less said or done about delinquency the better.

Sociologists have expanded Tannenbaum's perspective into what is often called a labeling, or societal reactions, theory of delinquency and other kinds of deviance. For example, Lemert (1967) suggests the terms *primary deviance* and *secondary deviance* to distinguish between acts that occur before and after the societal response. Acts of primary deviance are those that precede a social or legal response. They may be incidental or even random aspects of an individual's general behavior. The important point is that these initial acts

have little impact on the individual's self-concept. Acts of secondary deviance, on the other hand, follow the societal response and involve a transformation of the individual's self-concept, "altering the psychic structure, producing specialized organization of social roles and self-regarding attitudes." From this point on, the individual takes on more and more of the "deviant" aspects of his or her new role (Becker 1963, 1964). The societal response has, from this viewpoint, succeeded only in confirming the individual in a deviant role; for example, by potentially making adolescent delinquents into adult criminals through the punitive reactions of the police, courts, and others.

In the end, symbolic interactionists do not insist that all or even most delinquent behavior is caused by officially imposed labels. Being labeled delinquent is thought, rather, to create special problems for the adolescents involved, often increasing the likelihood that this and related kinds of delinquent behavior will be repeated. The point is that not only the actor but also reactors participate in creating the meanings and definitions that generate delinquency. The symbolic interactionists note that poor are more likely than the rich to get caught up in this process. This point is further emphasized in conflict theories.

### CONFLICT THEORY AND DELINQUENCY

The most distinctive features of conflict theories include attention to the role of power relations and economic contradictions in generating delinquency and reactions to it. For example, conflict theories have focused on the role of dominant societal groups in imposing legal labels on members of subordinate societal groups (Turk 1969). The fact that subcultural groups typically are also subordinate groups ties this work to earlier theoretical traditions discussed above.

**An Early Group-Conflict Theory.** George Vold (1958) was the first North American sociologist to write explicitly about a group-conflict theory of delinquency. He began with the assumption that criminality involves both human behavior (acts) and the judgments or definitions (laws, customs, or mores) of others as to whether specific behaviors are appropriate and acceptable or inappropriate and disreputable. Of the two components, Vold regarded judgments and definitions as more significant. His salient interest was in how groups

impose their value judgments by defining the behaviors of others as illegal.

Vold regarded delinquency as a "minority group" behavior. For example, he argues that "the juvenile gang . . . is nearly always a 'minority group', out of sympathy with and in more or less direct opposition to the rules and regulations of the dominant majority, that is, the established world of adult values and powers" (p. 211). In this struggle, the police are seen as representing and defending the values of the adult world, while the gang seeks the symbolic and material advantages not permitted it under the adult code. At root, Vold argues, the problem is one of intergenerational value conflict, with adults prevailing through their control of the legal process.

**A Theory of Legal Bureaucracy.** According to this viewpoint, determining which groups in society will experience more delinquency than others may be largely a matter of deciding which laws will be enforced. Chambliss and Seidman (1971) observe that in modern, complex, stratified societies such as our own, we assign the task of resolving such issues to bureaucratically structured agencies such as the police. The result is to mobilize what might be called the primary principle of legal bureaucracy. According to this principle, laws will be enforced when enforcement serves the interests of social control agencies and their officials; and laws will not be enforced when enforcement is likely to cause organizational strain. In other words, the primary principle of legal bureaucracy involves maximizing organizational gains while minimizing organizational strains.

Chambliss and Seidman conclude that a consequence of this principle is to bring into operation a "rule of law," whereby "discretion at every level . . . will be so exercised as to bring mainly those who are politically powerless (e.g., the poor) into the purview of the law" (p. 268). Theoretical work of this kind coincided with important research on the policing of juveniles (e.g., Reiss 1971). According to the conflict theorists, poor minority youth appear disproportionately in our delinquency statistics more because of class bias and police and court prejudice than because of actual behavioral differences.

**Recent Structural Theories.** Some recent theories of delinquency have combined conflict theory's structural focus on power relations with

etioloical questions about sources of delinquent behavior as well as reactions to it. Thus Spitzer (1975) begins the formulation of a Marxian theory of delinquency (and deviance more generally) with the observation, "We must not only ask why specific members of the underclass are selected for official processing, but also why they behave as they do" (p. 640).

One effort to answer behavioral questions with insights from conflict theory is an "integrated structural-Marxist theory" proposed by Colvin and Pauly (1983). This theory integrates elements of control theory and Marxian theory. The theory is comprehensive, and only some of its most striking features can be outlined here. These features include a Marxian focus on working-class parents' experiences of coerciveness in the workplace, which Colvin and Pauly suggest lead to coerciveness in parenting, including parental violence toward children. In turn, Colvin and Pauly argue that such children are more likely to be placed in coercive control structures at school and to enter into alliances with alienated peers. All of these experiences make delinquent behavior more likely, including the violent and instrumental kinds of delinquent behavior that may be precursors of adult criminality.

Power-control theory is another recent structural formulation (Hagan 1989) that attempts to explain large and persistent gender differences in delinquency by taking power relations into account. Power relations in the family are the starting point of this theory. The cornerstone of the theory is the observation that, especially in more patriarchal families, mothers more than fathers are involved in controlling daughters more than sons. A result of this intensified mother-daughter relationship is that daughters become less inclined to take what they perceive as greater risks of involvement in delinquency. Police and other processing agencies act on stereotypes that extend these gender differences in officially recorded delinquency. Power-control theory generally predicts that in more patriarchal families, sons will be subjected to less maternal control, develop stronger preferences for risk taking, be more delinquent, and more often be officially labeled for being so. More recently, this theory has been elaborated to emphasize that in less patriarchal families mothers may become more involved in the control of their

sons and this can reduce their sons' involvement in risk taking and delinquency (McCarthy and Hagan 1999).

These structural approaches illustrate an ongoing trend toward theoretical integration in this tradition and elsewhere in the study of delinquency (e.g., Hagan and McCarthy 1997; Messner et al. 1989; Sampson and Wilson 1995; Tittle 1995). These integrations involve theories that are often thought to be in apposition if not opposition to one another. Yet the trend toward integration in delinquency theory has been apparent for more than a decade, and it seems likely to continue.

### TROUBLESOME QUESTIONS

Despite the richness of sociological theories of delinquency and the emerging sense of convergence among previously competing theoretical traditions, there is a new awareness that delinquency theories remain incomplete in their capacity to explain and sometimes even address basic micro- and macro-level questions. For example, a classic issue that persists despite its recognition is the question of why most delinquents discontinue their delinquency before or during their transition to adulthood. We know that most delinquents "age out" of delinquency, but we have not adequately explained why this happens (Hirschi and Gottfredson 1983). Our theories are much more attentive to why young people become delinquent than to why they stop being so.

A seemingly related but only more recently apparent question involves the decline since the early 1990s in the violent forms of delinquency, such as robbery and homicide, that we are best able to measure and monitor statistically over time. This trend is strikingly apparent as we head toward the millennium (see, e.g., Blumstein and Rosenfeld 1998). None of the prominent sociological theories of delinquency predicted or can easily account for this decline in violent delinquency. Again, our theories have focused more on increases in delinquency than on its decline. Sociological theories of delinquency confront new as well as continuing questions in the new century.

(SEE ALSO: *Crime, Theories of; Criminology; Juvenile Delinquency and Juvenile Crime*)

REFERENCES

- Becker, Howard 1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- 1964 *The Other Side: Perspectives on Deviance*. New York: Free Press.
- Blumstein, Alfred, and Richard Rosenfeld 1998 “Explaining Recent Trends in U.S. Homicide Rates.” *Journal of Criminal Law and Criminology* 88:1175–1216.
- Chambliss, William, and Robert Seidman 1971 *Law, Order and Power*. Reading, Mass.: Addison-Wesley.
- Cloward, Richard, and Lloyd Ohlin 1960 *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press.
- Cohen, Albert 1955 *Delinquent Boys*. New York: Free Press.
- Colvin, Mark, and John Pauly 1983 “A Critique of Criminology: Toward an Integrated Structural-Marxist Theory of Delinquency Production.” *American Journal of Sociology* 89(3):512–552.
- Durkheim, Émile (1897) 1951 *Suicide*, trans. John Spaulding and George Simpson. New York: Free Press.
- Hagan, John 1989 *Structural Criminology*. New Brunswick, N.J.: Rutgers University Press.
- , and Bill McCarthy 1997 *Mean Streets: Youth Crime and Homelessness*. New York: Cambridge University Press.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- , and Michael Gottfredson 1983 “Age and the Explanation of Crime.” *American Journal of Sociology* 89:552–584.
- Lemert, Edwin 1967 *Human Deviance, Social Problems and Social Control*. Englewood Cliffs, N.J.: Prentice-Hall.
- McCarthy, Bill, and John Hagan 1999 “In the Company of Women: An Elaboration and Further Test of a Power-Control Theory of Gender and Delinquency.” *Criminology* in press.
- Merton, Robert 1938 “Social Structure and Anomie.” *American Sociological Review* 3:672–682.
- 1957 *Social Theory and Social Structure*. New York: Free Press.
- Messner, Steven, Marvin Krohn, and Allen Liska 1989 *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*. Albany: State University of New York Press.
- Moynihan, Daniel P. 1969 *Maximum Feasible Misunderstanding*. New York: Free Press.
- Reiss, Albert 1971 *The Police and the Public*. New Haven, Conn.: Yale University Press.
- Sampson, Robert, and William Julius Wilson 1995 “Toward a Theory of Race, Crime, and Urban Inequality.” In John Hagan and Ruth Peterson, eds., *Crime and Inequality*. Stanford, Calif.: Stanford University Press.
- Shaw, Clifford, and Henry McKay 1931 *Social Factors in Juvenile Delinquency*. Washington, D.C.: National Commission of Law Observance and Enforcement.
- 1942 *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Spitzer, Steven 1975 “Toward a Marxian Theory of Deviance.” *Social Problems* 22:638–651.
- Sutherland, Edwin 1939 *Principles of Criminology*. Philadelphia: Lippincott.
- 1949 *White Collar Crime*. New York: Dryden.
- Sykes, Gresham, and David Matza 1957 “Techniques of Neutralization: A Theory of Delinquency.” *American Sociological Review* 26:664–670.
- 1961 “Juvenile Delinquency and Subterranean Values.” *American Sociological Review* 26:712–719.
- Tannenbaum, Frank 1938 *Crime and the Community*. Boston: Ginn.
- Tittle, Charles 1995 *Control Balance: Toward a General Theory of Deviance*. Boulder, Colo.: Westview Press.
- Turk, Austin 1969 *Criminality and the Legal Order*. Chicago: Rand McNally.
- Vold, George 1958 *Theoretical Criminology*. New York: Oxford University Press.

JOHN HAGAN

# K

---

## KINSHIP SYSTEMS AND FAMILY TYPES

Kinship systems are mechanisms that link conjugal families (and individuals not living in families) in ways that affect the integration of the general social structure and enhance the ability of the society to reproduce itself in an orderly fashion. Kinship performs these social functions in two ways. First, through relationships defined by blood ties and marriage, kinship systems make possible ready-made *contemporaneous networks* of social ties sustained during the lifetimes of related persons and, second, they enable the *temporal continuity* of identifiable family connections over generations, despite the limited lifespan of a family's members. Variations in norms governing the structure of contemporaneous networks and the modes of temporal continuity compose the basis for the typologies of kinship systems described in this article.

In conceptualizing connections between kinship systems and family types, social scientists have applied either of two approaches. Some have developed typologies from historical analyses (and evolutionary schemes) that depict the transition of Western societies from ancient or medieval origins to modern civilizations. Other social scientists construct typologies that cut across diverse historical periods. Each historical era then constitutes a unique medium in which the structural typologies are expressed.

## MODERNITY, FAMILY PATTERNS, AND KINSHIP SYSTEMS

There are at least three ways to develop historical typologies related to kinship and family. One way is to hypothesize a linear historical progression, which includes a family type existing at the beginning point in time, a particular historical process that will act upon the family and kinship structures (e.g., urbanization or industrialization), and a logical outcome at the end of the process. A second approach builds upon the above approach by positing a transitional family type that emerges during the historical process and gives way in the final stages of the process to another family type. A third approach, which includes devising a family type based upon a configuration of attributes peculiar to a particular historical era (e.g., the Victorian family, the American colonial family), implies that any historical era represents a unique convergence of diverse factors.

**Bipolar Typologies.** By and large, sociologists have drawn a connection between kinship and family on the basis of a distinction between traditionalism and modernity. Generally, this distinction draws upon Henry Maine's ([1861] 1963) depiction of the transformation of social relations in early societies. Maine argued that social relations changed from those based on ascriptive status (deriving from birth) to relations created and sustained through voluntary contractual arrangements. Maine's theory has evoked a series of typologies that, in large measure, refine the status-contract distinction. For instance, an ideal type

developed by Ferdinand Toennies ([1887] 1957) has provided a backdrop for later typologies. The Toennies typology itself refers to a shift from *Gemeinschaft* (community) as a form of social organization based upon an existential will (*Wessenwille*), which is suited to feudalism and peasant society, to *Gesellschaft* (society) as a social form based upon rational will (*Kurwille*), which fits an urban environment under modern capitalism. Contemporary family typologies, in building upon Toennies's conceptual scheme, portray a weakening of kinship obligations and constraints.

One position, rooted in George P. Murdock's (1949) analysis of cross-cultural archives, has resulted in the main sequence theory of social change in kinship structure (Naroll 1970). Main sequence theory pertains to the way differential gender contributions to production of material resources affects the use of kindred as human resources/property. This theory holds that basic changes in kinship are initiated by a shift in the relative importance of men and women to the economic life of the society. First, there is a modification in the economic division of labor by gender. (For example, in hoe cultures, women tend to do the farming; when plows are introduced, men become the farmers.) Second, the shift in sexual division of labor generates a change in married couples' choices of residence, the major alternatives being near the husband's relatives (patrilocal), the wife's (matrilocal), or anywhere the couple desires (neolocal). (Plow cultures tend toward patrilocal residence.) Third, the change in choice of residential site affects the line of descent and inheritance favored in the kinship system: the husband's side (patrilineal), the wife's (matrilineal), or both sides (bilateral). (In line with the shift in residence, plow cultures show a greater inclination toward patrilinearity than do hoe cultures.) Fourth, the transfer to lineage affiliation generates a change in kinship terminology, particularly in ways that show tribal or clan membership, or, in modern societies, the dissolution of larger kinship structures. As applied to the emergence of modernity, main sequence theory predicts a continual emancipation from kinship constraints. An increase in the proportion of women in the labor force will produce a trend toward neolocal residence, which in turn will lead to increased emphasis upon bilaterality, weakening sibling ties and obligations to both

sides of the extended family, and in the long run to changes in kin terminology and identity [e.g., voluntarism in choice of surnames as an indicator of preference as to line(s) of descent].

In a variation of main sequence theory, urban sociologists such as Wirth (1956) and Burgess and associates (1963) wrote on the effects of transferring the economic base of societies from the land to urban centers. The theme of their work is to be found in the German proverb "Stadt Luft macht frei" ("city air makes one free"). For example, Burgess and associates described a progression from what they named the institutional family to the companionship family. In this conceptualization, the institutional family, embedded in a larger kinship group, is characterized by patriarchy, clearly defined division of household labor by sex, and high fertility. Its unity is derived mainly from external constraints—social mores, religious authority, fixity in location, position in the social structure, and the value of familism (i.e., values giving priority to the collective welfare of the family over that of individual members). Burgess and associates regarded the institutional family as an adaptation to relatively immobile, rural, agricultural societies and believed its way of life was fixed over time. By way of contrast, urban society, which is characterized by mobility, anonymity, and change, makes inoperative the social control mechanisms developed to maintain stable, rural societies. With the withering of these external controls on rural family life, Burgess, Locke, and Thomes proposed that the companionship family is bound together by internal forces—mutual affection, egalitarianism, a sense of belonging, common interests—and affords freedom from the demands of traditional family and kinship ties.

Unlike the urban sociologists, structural functionalists such as Talcott Parsons (1954) place considerable emphasis on the interaction of subsystems in the larger social system. In part, structural functionalists are concerned with economic and kinship factors in structuring nuclear family relationships. Parsons described American kinship as "a 'conjugal' system in that it is made up exclusively of interlocking conjugal families" (1954, p. 180) and is multilineal (i.e., bilateral) in descent. Parsons associates kinship solidarity with unilineal descent, that is, with a "structural bias in favor of



solidarity with the ascendant and descendant families in any one line of descent" (1954, p. 184). The absence of such bias in the American descent system, Parsons suggests, is in large measure responsible for "the structural isolation of the individual conjugal family" (i.e., its autonomy).

The importance Parsons attributes to unilinearity as a factor in facilitating strong dependence upon kin ties is exemplified by his highlighting two exceptions to the structural isolation of the conjugal family in America—the upper-class elements, whose status depends on the continuity of their patrilineages' solidarity, and the lower-class elements, in which there is "a strong tendency to instability of marriage and a 'mother-centered' type of family structure" (Parsons 1954, p. 185). However, Parsons regards the urban middle class as characterizing "the focal American type of kinship." Since in the middle class the residence of the conjugal family typically is neolocal, and the conjugal family is economically independent of "the family of orientation of either spouse," the role of the conjugal family in U.S. society can be, for theoretical purposes, understood as master of its own destiny, rid of the impediments of extended-family ties.

In reaction to those sociologists who see modernity as inimical to bonds of kinship, other social scientists (e.g., Adams 1968; Firth et al. 1969; Litwak 1985; Moge 1976; Shanas et al. 1968; Sussman 1959) turn their attention to the attenuated functions of kinship in contemporary society. Just as Goode (1963) notes a "fit" between the needs of modern capitalist society for a socially and geographically highly mobile population and the flexibility of the isolated conjugal family system, the revisionists indicate a similar fit between the existence of a highly mobile population and the presence of kin who give emergency aid and social support to relatives. The revisionists shift our attention away from constraints imposed by kinship loyalties and obligations and direct it instead to sources of services, goods, and emotional support that cannot readily be supplied by bureaucracies, markets, or other agencies. In his typology, Litwak (1960a, 1960b) distinguishes the isolated nuclear family (without kinship resources) from the traditional extended family (implying a hierarchy of authority), on the one hand, and from the

modified extended family (which consists of a network of related but autonomous nuclear families), on the other. Although the revisionists have not destroyed the foundation of the bipolar family typologies, they do focus on a previously neglected area of analysis.

**Three-Stage Typologies.** Some modernization typologies introduce a third, transitional stage between traditional and modern kinship and family structures. These typologies accept the position that initially there is an emancipation from traditional kinship constraints and obligations, but they also propose that at some point new values of modernity emerge to fill the vacuum left by the dissipation of the old kinship constraints. For example, building on the work of LePlay, Zimmerman and Frampton (1966) offer a scheme of transformation in which families change from a patriarchal form to a stem-family structure and thence to an unstable family type. Zimmerman and Frampton begin with the premise that each social organization derives its "essential character" from a triad of "imperishable institutions"—family, religion, and property. However, in their view, "familism is necessary in all complete social organization to a degree more imperative than the need for property" (1966, p. 14; 1947). Zimmerman and Frampton regard the patriarchal family as the most familistic form. The patriarchal type is rooted in idealistic religious values and is characterized by a common household of a patriarch and his married sons and their families, wherein the property is held in the name of the "house," with the father as trustee. They identify the patriarchal form as having been prevalent among agriculturists in the Orient, in rural Russia, and among Slavonic peasants.

With urbanization and industrialization, however, the unstable family becomes predominant. Zimmerman and Frampton associate the unstable family with materialism and individualism and the resulting atomization of social life. Individuals are "freed from all obligations toward their parents and relatives" (1966, p. 15), and the identity of each conjugal family as a social unit ends with the death of the parents and the dispersal of the children.

The stem family represents a transitional state between the patriarchal and unstable forms. The stem family extends branches into urban centers

while retaining its roots in the ancestral lands. As a result, the stem family provides a balance between the security of the traditional influences and resources of the “house” and the freedom and resources of the cities. (However, historical researchers yield less idyllic descriptions of the stem family than the Zimmerman and Frampton portrait. See Berkner 1972.)

A less romantic depiction of a transitional family type is drawn by Lawrence Stone (1975) in his typology of the English family’s movement from feudalism to modernity. Stone posits the existence of a dual historical process. He places the decline of the importance of kin ties in the context of the emergence of a powerful, centralized state, and he then regards the rise of the modern family as an ideological emergence accompanying the development of capitalism.

According to Stone’s typology, feudal England emphasized (1) kin-group responsibility for crimes and treasonable acts of members and (2) the institution of cousinship with its broad obligations. As political and economic power moved away from the traditional, landed elite to the state and the entrepreneurial class, the common law of the courts no longer recognized criminal and civil deviance as a kin-group responsibility, and cousinship lost its effectiveness. To fill the vacuum left by the decline of kinship as a factor in one’s destiny, the relatively denuded conjugal family had to take over the task of guiding the destiny for its members. Consequently, by the sixteenth century, as an intermediate step toward the modern family, there was a trend toward authoritarianism in husband-wife interaction, and governance in the conjugal family took the form of patriarchy.

Stone (1975, p. 15) suggests that it was not until the eighteenth century that the spread of individualism and utilitarianism gave rise to a more companionate and egalitarian family structure. This last family form has been designated by Alan Macfarlane (1986) as the Malthusian marriage system, in which (1) marriage is seen as ultimately the bride’s and groom’s concern rather than that of the kin group; (2) marital interaction is supposed to be primarily companionate; and (3) love is supposed to be a precursor of marriage. Functionally, the Malthusian system yields relatively fewer children—by choice—than earlier family forms.

### **The Problem of Structure in Modernity**

**Typologies.** Typologies depicting historical transformations in family and kinship place much emphasis on the “fit” between the needs of modern industrial society and the presence of the conjugal family type (Litwak 1960a, 1960b; Parsons 1954). Despite this conjecture, Parsons (1954, p. 184) suggests that in Western society an “essentially open system” of kinship, with its “primary stress upon the conjugal family” and its lack of larger kin structures, has existed for centuries, long before the modern period. Like Macfarlane (1986), Parsons dates its establishment in late medieval times “when the kinship terminology of the European languages took shape.” Moreover, Goode’s (1963) analysis of family trends in eleven societies indicates that acceptance of modern, conjugal family ideology may precede economic and industrial development rather than come as a subsequent adaptation. Such findings cast doubt on the validity of the dichotomy between traditional societies and modernity as providing a theoretical basis for the typologies discussed above.

Parsons argues that (1) there is an incompatibility between corporate kinship and multilineal systems, and (2) in large measure, this incompatibility accounts for the prevalence of highly adaptable, structurally independent conjugal households in modern societies. However, findings by Davenport (1959), Mitchell (1963), Pehrson (1957), Peranio (1961), and others that corporate structures of kinship (such as clans) do exist in some multilineal kinship systems undercut Parsons’s argument that such structures are to be found only in unilineal systems. Nevertheless, if multilateral kinship systems can accommodate corporate structures, then they can also include other kinship elements that sustain loyalties to descent groups and facilitate segmentation of the society.

### **The Problem of Connecting Kinship and Family in Modernity Typologies.**

Revisionists of the isolated conjugal family position have presented considerable evidence of residual elements of kinship ties in contemporary society. However, they do not adequately explain the connections between types of kinship systems and variation in performance of family functions in different parts of the social structure. Their main concern is with changes in kinship and family, changes that are consistent with the general loosening of tradition

in modern society. But their focus on emancipation from tradition diverts their attention from (1) the influence of emerging ethnic, religious, or class interests upon patterns of integration of family networks in the larger social structure and (2) the temporal dimensions of kinship, which go beyond living kin to departed ancestors and generations yet to come.

Additionally, given the fact that the family-kinship typologies described above have their roots in the distinction between tradition and modernity, they overlook those nonindustrial, primarily nonurban societies in which families approach the companionship model as well as those ethnic and religious segments of industrial, primarily urban societies where strong familistic tendencies persist. Except for Stone (1975) and Zimmerman and Frampton (1966), these typologies are based on the concept of emancipation from tradition, and they do not deal explicitly with the emergence of new family values (other than flexibility and freedom). Most of all, their emphasis on emancipation from the constraints of tradition precludes their explaining why cohesive forces of family and kinship may remain strong (or increase in strength) in the face of an economic and social environment that is hostile to stable family life. (Exceptions are Sennett 1970 and Harris and Rosser 1983.)

#### **Critical Commentary on Historical Typologies.**

Family typologies describing historical trends from one period of history to another are vulnerable to criticism of their teleological assumptions. Criticisms often involve (1) the definition of polar concepts and (2) the problem of inevitability.

(1) *Definition of polar concepts.* The definition of polar concepts depends upon the value commitments of the analyst. For example, those analysts who view trends in kinship and family as movement toward liberation from traditional constraints and from obstacles to personal independence define the original state as confining and generally unjust and the future state as enabling emancipation from these obsolete social structures.

Family-theorist Ernest Burgess and associates (1963) view the evolution of family structure as going “from institution to companionship”—from external community constraints upon family relations to voluntaristic unity that derives from affection, domestic peace, and common goals. Similarly, Marxists define the transition as being away

from family structures required to sustain an economic system based on unearned rewards of the dominant class and suppression of the laboring class. Their claim is that following the rise of future true communism, the dissolution of economic classes would liberate family life from the constraints and suffering imposed by economic position; for Frederick Engels ([1885] 1942), under true communism, family life would be liberated from economic demands, and, founded on personal bonds, families would endure only as long as these bonds lasted.

By way of contrast, analysts favoring traditional values define the trend in family life as a steady decay of family structure. Pitirim Sorokin (1937, vol. 4, p. 776), upon whose work Zimmerman and Frampton base their typology, notes that “the family as a sacred union of husband and wife, of parents and children, will continue to disintegrate . . . [T]he home will become a mere overnight parking place mainly for sex-relationship.” From another perspective, the behaviorist John Watson (1927) predicted that “in fifty years [1977], unless there is some change, the tribal custom of marriage will no longer exist. Family standards have broken down . . . The mystery and beauty of marriage and the rearing of children has pretty well broken down.” In the Aldous Huxley’s science fiction novel *Brave New World* (1955), all functions now performed by families would be community undertakings, and the word “mother” would be regarded as obscene.

The distinction between typologies focusing on personal liberation and those portraying decay highlights the fact that each approach deals with partial realities. Liberation typologies tend to slight disruptive activities of emerging family structure (e.g., spouse abuse, child abuse, splitting into factions, isolation from resources of kin and family). Instead, they tend to associate these activities with traditional family structures. In the Soviet Union, family problems were generally attributed to survivals of the traditional pre-Revolution family forms. Decay typologies do the opposite. They tend to underestimate unifying elements and personal satisfactions associated with the emerging family types and to overestimate the chaos associated with these types.

(2) *Problem of inevitability.* Some typologies posit a straight-line progression from a beginning

state to an end state, while others at least imply a degree of indeterminacy in movement. For any particular typology, the degree of indeterminacy depends on the analyst's conception of history. For example, the Burgess institution-to-companionship family typology assumes there to be a major historical evolution from fairly isolated rural communities to societies with a high degree of industrialization; the family and kinship institutions evolve through a long series of adaptations to keep up with the industrialization of societies and a movement from a rural to an urban way of life. Similar typologies—like Toennies's *Gemeinschaft* to *Gesellschaft* ([1887] 1957) and Redfield's folk society to civilization (1947)—are based on comparable assumptions deriving from the dictum of Henry Maine, namely, that, as an overarching historical trend, the basis for social relations has been evolving from status to contract.

Despite the inevitability of trends implicit in the definition of polarities of family and kin structure in typologies of liberation (or decay), with the passage of time, definitions of polarities change. For example, analysts have redefined the concept of companionship as an end-state. In the 1940s, Burgess (1948; Burgess et al. 1963) regarded the future end-state as one in which the husband and wife (1) would be married without interference from family and community constraints, (2) would remain united through affection and common interests, (3) would maintain an equality in decision making and other aspects of family status, and (4) would orient their parenthood toward producing children with healthy personalities. But variations in family life included under the "companionship family" definition have been broadly expanded over time. By the end of the twentieth century, the end-state of the companionship family (as well as the unstable-family concept) has been redefined to include a diversity of household arrangements, such as (1) couples living together without formal marriage, (2) same-sex couples and their children (by adoption or by birth from previous or supplementary liaisons), and (3) voluntary single-parent households. Given these modifications in the concept of the companionship family, the very nature of the typology has been transformed. As a result, it is difficult to determine what family and kinship theorists will consider to be the evolutionary outcome twenty-five years from the present.

According to Murdoch's (1949) depiction of main sequence theory (described earlier), the changing pattern of employment has facilitated the widespread movement of women into broad sectors of occupations. This change has affected the composition of residences and, subsequently, will affect the descent structure and eventually kinship terminology. Evidence of this development can readily be seen. Around 1960, in an offhand comment during a lecture, Murdoch predicted that the control over wealth in America (1) would flow increasingly into the hands of women, (2) would at some point create shifts in household patterns, and (3) in the long run would produce a kinship structure dominated by women. Regardless of the accuracy of Murdoch's prediction, changes in practices pertaining to kinship are appearing in various ways: (1) Newspapers obituaries have routinely begun to include "life companions" (of either gender) in the list of related survivors; (2) public policy pertaining to health insurance coverage has been modified in some communities to include unmarried domestic partners; (3) in some countries (e.g., Russia, Israel), intestacy laws have been amended to include unrelated household residents; (4) the issue of legally recognizing same-sex marriages (or domestic partners) as a valid arrangement has emerged in a wide range of communities.

Unlike the theoretical inevitability of collectively rational adaptations assumed by evolutionary theorists, the typologies formulated by cyclical theorists lead away from regarding their end-states as inevitable. In their portrayal of historical processes, the cyclical theorists have the burden of explaining conditions for triggering reversals in historical cycles. These reversals imply that critical periods arise through cultural innovations and conflicts. The effects of novelty and conflict in these critical periods introduce an indeterminacy into the historical process. This indeterminacy brings to the foreground the problem of the inhibition of change: What introduces a new cycle, and what brings the cycle to a halt? For example, as discussed earlier, Zimmerman and Frampton (1947) see the history of the family as a series of repetitive cycles: a decay from corporate family forms (based on idealistic values) to unstable, chaotic families (based on materialistic values and individualism), followed by a regeneration of familism. Their

scheme of analysis explains the oscillations between various degrees of familism and individualism in terms of a conflict between maintaining an enduring, traditional social structure and attending to persistent personal yearnings. The idealism of religious or ascetic values facilitates social stability in corporate family settings. However, the stifling of personal aims and desires, without idealism, encourages the adoption of materialistic values and sensuality associated with the unstable family. Hence, there is no guarantee that an old cycle will end or that new ideals supporting familism will again emerge.

### A TRANSHISTORICAL TYPOLOGY OF KINSHIP AND FAMILY SYSTEMS

For well over a millennium, church intellectuals have been aware of variations in marital selection and their implications for family structure and kinship ties as well as for social structure. Early in the fifth century, in his *De Civitate Dei (City of God)*, Saint Augustine of Hippo (1984, pp. 623–625) noted that in early biblical times demographic insufficiencies made it necessary for Jews to practice kinship endogamy. However, he proposed that marrying close relatives, and thereby creating multiple family ties with the same people, restricted the potential expanse of social circles that could be tied into a coherent community. Kinship endogamy tends to divide societies into segments. On the other hand, marrying persons from previously unrelated families would “serve to weld social life securely” by binding diverse peoples into an extensive web of relationships. Later, in the twelfth century, Gratian suggested that God commanded the Hebrews to select relatives as mates “because the salvation of man was realized in the pure Jewish race” but that the Christian faith, which could be readily spread through teaching, made kinship endogamy obsolete (Chodorow 1972, p. 74).

Gratian’s argument suggests that the differences between Judaic and Christian marriage systems have broad implications for contemporaneous functions of kinship as well as for temporal functions, connecting past and future generations. The discussion that follows presents a kinship and family typology derived ultimately from Augustine’s and Gratian’s depictions of marriage systems as well as from issues pertaining to descent.

This typology involves theoretical concerns drawn from sociology and anthropology.

**Contemporaneous and Temporal Functions of Kinship Systems.** Both marriage systems and descent rules affect the character of links between contemporaneous networks of families. A major controversy that at one time occupied many social anthropologists was whether marriage systems (i.e., marital alliances between groups) are more fundamental in generating forms of social organization than are descent rules or vice versa. At stake in the controversy was the issue of whether the social solidarity undergirding descent rules is more fundamental than the ideas of reciprocity and exchange involved in marriage systems. In the end, Africanists favored descent rules, while Asianists leaned toward marital alliances. In their assessment of the controversy, Buchler and Selby (1968) found evidence for the validity of both views.

However, despite the chicken-and-egg character of the controversy, the alliance–descent issue highlights the contradictory nature of kinship structure. This contradiction is depicted in the opposing views of structuralists such as Claude Lévi-Strauss (1963), who supports the alliance position, and functionalists such as Meyer Fortes (1969), who argues for the descent position.

Alliance theories of kinship systems identify the primary function of kinship as the integration of networks of related families into the contemporaneous social fabric. Alliance adherents begin with marriage as the central element in structuring the way kinship operates. To alliance theorists, the significance of marriage lies in the idea that marriage is essentially a mode of exchange whose primary reason for existence is to inhibit conflict in society. In their view, kin groups exist as organized entities to effect marital exchanges. According to Lévi-Strauss, the leading figure in alliance theory, “exchange in human society is a universal means of ensuring the interlocking of its constituent parts” (1963, p. 2). In unilineal systems, women are exchanged for equivalent valuable property, services, or both; in bilateral systems (which by their nature become multilateral in the long run), commitments to each other’s relatives are exchanged. In bilateral kinship, bride and groom are of presumably equivalent value. Thus, in general,

alliance theorists regard descent groupings primarily as a necessary ingredient for sustaining the marriage exchange system over the generations.

The descent theory of kinship systems rests on the assumption that the continued welfare of kindred over the generations is the primary function of kinship. In particular, Fortes regards "filiation"—being ascribed the status of a child of one's parents, with all the lifetime rights and obligations attached to that status (1969, p. 108)—as the "crucial relationships of intergenerational continuity and social reproduction" (pp. 255–256). He proposes that, as a concomitant of filiation, "the model relationship of kinship amity is fraternity, that is sibling unity, equality, and solidarity" (p. 241), and he provides a biblical example of the tie between David and Jonathan. But he also notes that "the Euro-American kinship institutions and values of Anglo-Saxon origin are imbued with the same notion of binding force of kinship amity" (p. 242), and he cites the mother–daughter relationship in England (in research findings by Young and Willmott 1957) as exemplifying that same moral code of diffuse but demanding reciprocal obligations.

On the one hand, alliance theory postulates that the basic drive in kinship organization is derived externally, from the kind of alliances appropriate to the structure of power in the community. Collectively, marital alliances create between families a network of links that integrate them in reference to overarching religious, economic, and political institutions. On the other hand, descent theory ascribes the bases of organization to internal demands, structural factors in the persistence of the kindred: rules governing residential location, division of labor and authority among members, and the various economic and political functions to be performed by the kinship system (Buchler and Selby 1968, p. 129).

Given the contradiction in the impulse for kinship organization, there is an apparent "impasse between the alliance and filiation point of view" (Buchler and Selby 1968, p. 141). What appears to be at issue is the depiction of the kinds of reciprocity norms that define the character of kinship. Descent theory presumes that an axiom of amity (i.e., prescriptive altruism or general reciprocity) is basic to the coherence of kin groups; alliance theory holds that balanced reciprocity

(i.e., the rightness of exchanges for overt self-interest, opportunistic individualism, or noumenal norms) is in the final analysis the glue that integrates families and kin groups into a coherent whole.

The contradiction is apparent in many ways. For example, in biblical references and religious writings, the Ten Commandments enjoin one to honor parents and, conversely, to "cleave" to one's spouse and maintain peace in the household. In terms of kinds of reciprocity, one commandment involves unconditional giving or honoring, while the other concerns maintaining domestic peace (implying fair give-and-take).

Similarly, contemporary writers on marriage generally find the concept of balanced reciprocity appropriate in describing the quality of husband–wife ties. For example, Walster and Walster (1978) report that marriages work best when both husband and wife (as well as lovers) believe that each is receiving a fair exchange for what he or she offers in the relationship. Moreover, in their review of research on the quality of marriage, Lewis and Spanier (1982) note the importance of the symmetry of exchange in establishing and maintaining strong marital ties. However, in the socialization of children and in the allocation of resources, the rule of amity (or prescriptive altruism) is supposed to prevail. For example, parents are ordinarily expected to make "sacrifices" for their children when necessary; to do otherwise is to be a "bad" parent. In the American court system, the general rule for the disposition of children in cases of divorce, child neglect or abuse, or adoption is that the court should base its decision on the welfare of the child rather than on the interests of the parents or other parties.

To some extent, the descent–marriage contradiction can be obscured by compartmentalizing marital, parental, and filial conduct and by dividing responsibilities of husband and wife. However, conflicts in norms for dealing with family members and kindred may occur for several reasons, but they occur principally because of scarcities of time and resources required to carry out duties and obligations in the face of a wide range of simultaneous and conflicting demands. Since the resulting dilemmas are widespread in the society, there is a need for a general rule. Because contradictory alliance and descent impulses are operative, each group is pushed to establish a coherent

kinship scheme that gives priority to one impulse over the other or at least establishes some form of compromise between them.

There is evidence that rules governing marital functions conflict with those pertaining to descent functions, paralleling the alliance–descent controversy in kinship systems. Where descent functions are given precedence in family organization, marital functions are subordinated (and vice versa). Examples of this inverse relationship are (1) if husband–wife unity is central, then the unity between siblings is peripheral (and the reverse), and (2) if marriage between close affines is forbidden, first-cousin marriage is permitted (and vice versa). These examples are discussed in the sections that follow.

**Marital Unity Versus Unity of the Sibling Group.** Comparisons between societies indicate that ties between siblings have an inverse relationship to husband–wife ties. Where descent is valued over alliance or marriage in kinship relations, brother–sister bonds are particularly close (Parsons 1954), while the husband–wife relationship is relatively distant. In such family systems (whether or not its therapeutic implications are true), parents are expected to remain together for the sake of the children, and this expectation expresses the priority of descent over marital ties. Conversely, in family systems where the marriage function is more valued, the husband–wife relationship is intense (e.g., the importance of the give-and-take of love and of companionship for marriage) and the brother–sister relationship is competitive, distant, or both and the incest taboo justifies their apartness (see Lopata 1973 on widows and their brothers). In societies where priority is given to marital bonds over descent ties, the presence of children is of less importance in dissolving an unhappy marriage, and there is greater ambiguity about what is best for the children. The mere fact that the strength of brother–sister ties and that of marital ties vary inversely in different societies lends support to the proposition that there is a contradiction in the family system between its marital functions and its descent functions.

**Affines and Cousins in American Marriage Law.** The opposition between marital and descent functions in the family is also illustrated by the inverse relationship in American law of marriages

considered to be incestuous: As a general tendency, states that forbid second marriages between a person and certain affines (such as that person's parents-in-law and sons- or daughters-in-law) allow first cousins to marry, while those that permit marriage between close affines forbid first-cousin marriage (Farber 1968). If the preferred function of marriage is to reinforce close consanguineous kinship ties, then this pattern of marital prohibitions signals a subordination of affinal bonds to those of consanguinity. Marrying into the family of the former spouse will not reinforce any of the other existing bonds of consanguinity. Consequently, although first-cousin marriage is to be permitted in order to reinforce intimate kinship ties, marriage with close affines should be avoided. However, if marriage is considered to be primarily a mechanism for creating new bonds between previously unrelated families, then a second marriage into the same family merely serves to maintain the affinal bonds initiated in the first marriage.

**Social Structure and Kinship Systems.** The presence of contradictory impulses in organizing kinship ties produces a predicament in establishing priorities between them. This contradiction evokes a question: Which circumstances lead some societies (and ethnic and religious subgroups) to give priority to descent and others to favor alliance assumptions in their kinship and family organization (Farber 1975)? In their analyses of the relationship between kinship organization and social structure, both Paige (1974) and Swanson (1969) distinguish between societies that feature the legitimacy of special interests—factionalism—in organizing social life and those that feature the importance of common interests—communalism—as an organizing theme.

Factions are a means for gathering forces and mobilizing members for conflict or competition with other factions. They emerge as a reaction to perceived danger to their well-being from other groups (cf. Douglas 1966). Factions emerge where either (1) special interest groups vie for superiority over other groups for access to power, wealth, or some other property, or (2) groups sense a danger to their continued autonomous existence as an ethnic or religious entity.

In kinship organization, the continual mobilization of family and kin results in the generation of

norms that are centripetal in nature, that is, they facilitate the pulling inward of human, symbolic, and material resources. This centripetal tendency permits each kin group to separate itself from competing groups in order to endure. As a result, centripetal kin groups favor norms strengthening descent relationships over norms facilitating new alliances with other groups through marriage. Insofar as descent-group norms are rooted in the axiom of amity, one would expect centripetal kinship organization to feature the norm of prescriptive altruism over balanced reciprocities in kinship and family relations (see Farber 1975).

Jewish family norms provide some insight into the relationship between centripetal kinship systems and the application of the axiom of amity. In its basic ideology and in the code of laws supporting that ideology, Judaism assigns a major significance to the concept of nurturance (Farber 1984). Since nurturance is a central feature of maternal giving, it can be regarded as a metaphor for the axiom of amity. *The Code of Jewish Law (Shulkhan Arukh)* offers numerous instances that signify the place of nurturance in Judaism (Ganzfried 1963). For example, the code sublimates feeding and eating into sacred, ritualistic acts. The act of eating is invested with holiness, to be enjoyed in abundance, particularly on feast days and the Sabbath. A connection is made in the code between providing food and giving gifts and charity. It proposes that festive occasions are also times for charity to the needy and for sending gifts. In addition to drawing a connection between food and charity, the code applies the metaphor of the parent-child relationship to charity giving and assigns a priority to family in its general concept of nurturance: First parents, then offspring, and “other kinsmen take precedence over strangers” (Ganzfried 1963, chap. 34). The injunction to nurture children involves an emphasis not only on food but on other aspects as well (for example, an exaggerated emphasis on elaborated linguistic codes for use in child rearing). Zena Smith Blau (1974) writes that “whatever Jewish mothers did for their children—and they did a great deal—was accompanied by a flow of language, consisting of rich, colorful expressive words and phrases” (p. 175). The aim of socialization is presumably to turn the child into a *Mensch*—to transform the child from a receiver of nurture to a giver of nurture (Zborowski and Herzog 1952). Hence, in traditional Judaism, the concept of

nurturance seems to tie together the kinship emphasis on descent and the axiom of amity in organizing family relationships.

If nurturing the next generation is a form of prescriptive altruism, this nurturing can also occur in symbolic form. Like the transmission of physical wealth and nurturing, the parents can also transmit a “symbolic estate” to the next generation. In a real sense, along with material resources, people inherit a collection of living and dead relatives connected to them by birth and/or marriage. These relatives constitute a trove of heroes and villains whose personal qualities, exploits, and ideas are remembered in socializing succeeding generations. This “symbolic estate” defines for individuals (1) a sense of belonging to an identifiable “family,” (2) role models to emulate (or disown), (3) a legitimation of one’s place in community and society (Farber 1971). Craig (1979) sees the symbolic estate as a vehicle for achieving personal and familial immortality. Implicitly, it is one’s duty in centripetally-oriented kinship systems to contribute to the symbolic estate by living an exemplary life (however this way of life is defined in particular historical circumstances).

To be operative as memorials (or reminders), the content of symbolic estates must have some bearing upon the personal identities (or destinies) of family members. In Judaism, historically this meant assessing the “quality” of one’s ancestry (*yachas*), however defined; this assessment was particularly important in eras of arranged marriages. However, Yerushalmi (1982) notes the general importance of collective memory for the endurance of Judaism. Its centrality is suggested by the appearance of the verb *zakhar* (to remember) “in the [Hebrew] Bible no less than one hundred sixty-nine times” (Yerushalmi 1982, p. 5). Moreover, numerous memorials have been incorporated into holy day observances (e.g., the retelling of the story of the Exodus annually at the family *seder* at Passover). Especially significant for sustaining symbolic estates among Jews is the ritualizing of the remembrance of dead relatives through (1) memorial prayer services (*yizkor*) on four major holy days, and (2) partly as a means to continue to honor one’s parents after their death, the recitation of the prayer for the dead (*kaddish*) on anniversaries of the death of each family members.



The concept of symbolic estates connects collective family memories—such as legends, myths, and moral ideas—to the continuity of “family” from one generation to the next. In her study of Genesis, Steinmetz (1991) applies the concept of “symbolic estates” to the succession from father to son of the obligation to ensure the realization of God’s command to found and then maintain a Jewish nation. She regards the entire structure of Genesis as resting upon the transfer of this ideal to worthy heirs in the family line. Her emphasis upon the transmission of “symbolic estates” is echoed in an investigation by Bendor (1996) of the social structure of ancient Israel. Bendor concludes that Israeli social stratification is derived to a large extent from the kinship ideology of familial perpetuity—rather than from the influence of economic factors upon kinship and family life. The findings on ancient Israel by Steinmetz and Bendor bear upon historical and contemporary studies of kinship and family. Examples are the research reports by Pina-Cabal (1997) on family legends in urban Portugal, Attias-Donfut (1997) on home-sharing in France, Hastrup (1982) on establishing Icelander ethnicity, and Weigert and Hastings (1977) on maintaining family archives of photographs, old records, letters, and other memorials. The focus in these studies is upon symbolic mechanisms for sustaining family continuity. (The discussion of centrifugal kinship systems in the next section will describe obstacles to the perpetuation of “symbolic estates.”)

As opposed to factionalism, communalism implies a situation in which special interests are subordinated to common concerns of diverse groups. In stateless societies, these common concerns may well emerge from economic interdependence or the presence of a common enemy. In societies with a centralized government, the state presumably symbolizes a concern for the common welfare of the populace. Other unifying concerns may exist as well, for example, the presence of a universal church (as opposed to competing sects and denominations), nationalism (as opposed to ethnic self-determination), a centralized bureaucracy or market (as opposed to regional competition for dominance), and so on. The common concerns would best be served if members of kin groups were to be dispersed by marriage to previously unrelated people living throughout the society. This dispersal would maximize the number

of diverse kin groups with which any family is connected, and it would thereby scatter kinship loyalties, obligations, and property as widely as possible. Consequently, this kind of kinship system, associated with communalism, can be identified as applying an outward pressure upon its constituents; it is centrifugal in nature.

In contrast to the centripetal system, the centrifugal system subordinates kinship ties to conjugal family ties and extends marital prohibitions widely in order to inhibit marriages that would merely reinforce existing consanguineous ties. According to the theory outlined above, in centrifugal kinship systems, in which marriage functions are given priority over descent functions, the appropriate norm for defining family interaction is balanced reciprocity—exchange rather than the axiom of amity.

In the United States, although the centrifugal kinship system appears in a wide range of socioeconomic, religious, and ethnic groups, it is found disproportionately at lower socioeconomic levels, where families seek improved integration into the larger society (Farber 1981).

The application of balanced exchange as a norm in family and kinship is exemplified in a study of poor families by Stack (1974). She describes the prevalence of “swapping” as a named, bartering norm governing both ties between kin and between family members in their struggle for survival. Stack notes that “reciprocal obligations last as long as both participants are mutually satisfied” and that they continue such exchange relationships as long as they can “draw upon the credit they accumulate with others through swapping” (p. 41). Indeed, according to Stack, “those actively involved in domestic networks swap goods and services on a daily, practically an hourly, basis” (p. 35). But this exchange does not constitute a playing out of the axiom of amity since “the obligation to repay carries kin and community sanctions” (p. 34) and it extends beyond family and kin to friends. Although swapping may involve some element of trust, it exists to ensure exchanges in the lean times that predictably recur in domestic networks that are too marginal in resources to be magnanimous. It pays to create numerous bartering arrangements rather than to accumulate obligations within a very small network of intimate kin. Thus, in its own way, swapping mimics the proliferation of

networks of previously unrelated families characteristic of centrifugal kinship systems.

In contrast to the importance of “symbolic estates” for facilitating the “immortality” of families in centripetal kinship systems, families in centrifugal systems are often characterized by a “legacy of silence.” This silence may signify the existence of shameful or immoral acts of relatives, or it may simply reflect an emphasis upon individualism in these families. In Germany after World War II, this “legacy of silence” functioned to erase the collective memory of parental activities and ideas they held during the Nazi era (Larney 1994, pp. 146–162). For victims of torture and displacement under the Nazi regime, the legacy of silence enabled them to wipe their degradation from memory (Bar-On 1989). In either case, whereas symbolic estates provide a vehicle for family continuity, the legacy of silence established a discontinuity.

The German experience may result in a single break in family continuity—to permit starting afresh. However, the institutionalization of the legacy of silence in centrifugal kinship systems perpetuates this discontinuity between generations of nuclear families. This legacy has been found to be prevalent in low socioeconomic-level families populating urban slums (Farber 1971). Families tend to exchange little information about one another; in fact what is hidden may permit closer ties between kin than the revelation of illicit or immoral acts. Then too, in families where welfare agencies and police intrude, silence serves to maintain the privacy of the household.

The tacit norm of collective forgetting in these centrifugal kinship systems places the onus for kinship unity upon mutual assistance, friendship, and availability of kin. Yet, in her study of kinship among poor racial and ethnic minorities, Roschelle (1997) found that degree of mutual assistance between families and extent of interaction among relatives depend largely upon availability of kin. Migrant families frequently are isolated in time of need and the legacy of silence may thereby be enhanced.

In a society marked by much internal migration and social mobility, there are many opportunities for a proliferation of centrifugal tendencies in kinship. Whether centrifugal systems actually emerge through mobility may depend upon a

variety of factors. Gullestad (1997) notes a shift in the meaning of kinship in urban Norway. She describes a social transformation from norms regarding “being of use” and social solidarity to self-realization and “finding oneself” (or “being oneself”), that is, from norms sustaining family continuity to norms fostering separation and discontinuity. She attributes this shift to “transformed modernity” involving “fundamental restructurings of home and neighborhood because women and children are not present in the same way or to the same extent as before” (Gullestad 1997, p. 210). Transformed modernity, as well as advances in reproductive technology, is identified also as a factor in the proliferation of diverse forms of kinship structure in contemporary society (Strathern 1992).

One can interpret the emergence of feminist movements as both stimulating and stimulated by the “transformed modernity” cited by Gullestad. For some forms of feminism, post-modern thought provides a rationale for denigrating traditional symbolic estates. Post-modern writings propose that the framing of “factual” and theoretical statements have an exclusionary element—that is, they mark a population segment for exclusion from free participation. These “factual” statements justify this exclusion. As “factual” statements, posing as objective discourses, these statements have a hidden core. This core reflects the special interests of those with the power to define “truth” for the society. (See Foucault [1971] 1996.) Certain feminists claim that the hidden core of meaning in statements justifying exclusion of women from full participation in society is to promote male dominance in social structure (Barnard 1993). The symbolic estates that facilitate the endurance of existing lines of descent are thus seen as supporting patriarchy. Consequently, they are regarded as an obstacle to the full participation of women in society. Yet, as women’s participation in economic and political spheres continues to expand, it is likely that symbolic estates will eventually be infused with a marked increase in content pertaining to exploits and interests of women.

**Kinship-Map Typology.** Although mapping of kinship ties cannot express all aspects of kinship relations, it can generate models expressing general orientations implicit in various patterns of kinship structure. Basically, genealogical maps of

consanguineal (“blood”) relationships merely locate positions in an ideal web of biological connections. Variations in mapping come into play when these maps are used to describe how one’s obligations and proscriptions vary in different kinship structures. Such obligations and proscriptions pertain to marriage, remarriage, birth and adoption, inheritance, relations between generations and genders, and so on.

One advantage of models of genealogical mapping is that these models express the logical connections between functions of kinship in a particular society and priorities assigned to different kin statuses. For instance, a kinship type with a prohibition to marry a first cousin generally has a different function in society as compared to one permitting such marriage. This pattern of marital prohibitions will likely be related to priorities in inheritance. For instance, in American state laws, permitting first-cousin marriage would be associated with giving a niece or nephew precedence over a grandparent in intestate inheritance (i.e., when there is no written will). The opposite will likely be true where first-cousin marriage is forbidden.

The relationship between genealogical mapping and functions of kinship has a long history in Western civilization. Atkins (1974) has explored a wide range of formulae for generating different patterns of priorities in mapping genealogical relationships. Implied in genealogical mapping is the principle that the smaller the number of links (by birth or marriage) between relatives, other things being equal, the greater is the degree of obligation between them. Thus, in such matters as succession to estates, when a choice is to be made among kin, genealogically close relatives are presumed to be given priority over more distantly related kin. However, since the various formulae differ in the patterns of priority among kin generated, choice of an appropriate pattern of mapping depends on the role of kinship in the particular society.

In general, three patterns of priority for mapping kin have been applied in the Western world (mainly in laws of intestacy and marriage). However, in practice, each society makes modifications in these patterns to fit its needs. Examples of these patterns occur in (1) Catholic canon law and the state of Georgia, (2) the civil code of the Twelve Tables of the Roman Republic and more recently in Napoleonic Code and Louisiana law, and (3) the

parentela orders in the Hebrew Bible and in abbreviated form in Israel, Germany, and various states (e.g., Arizona) (Farber 1981).

At one pole, the canon law of the Catholic Church stipulates that a function of the church is to create a unity that ties together diverse segments of its constituency in a web of extensive relationships (including family bonds). This aim implies that collateral ties between families are equal in importance to ties between ascendants and descendants (i.e., between generations). Under such conditions, ties between are extended outward in a centrifugal fashion. In laws governing marital prohibitions, marriage is discouraged within the second degree of distance of collateral kin (i.e., first cousins). In earlier generations, marital prohibitions in Canon Law were even more inclusive; for example, in thirteenth century, consanguineous marriages were prohibited within the fourth degree of relatedness.

In assigning distances from Ego in the canon law genealogical model (e.g., for priorities in inheritance), (1) all consanguineal members of Ego’s nuclear family (parents, siblings, and children) are one degree of distance from Ego, (2) relatives just outside the nuclear family are two degrees of distance (grandparents, aunts and uncles, first cousins, nieces and nephews, and grandchildren), and so on. The canon law model thereby expresses the general principle that neither line of descent nor collateral distance is given special emphasis—only degree of distance from one’s nuclear family is significant.

At the opposite pole, the parentela orders genealogical model places much emphasis upon line of descent (and among collateral relatives, the closeness of line of descent). Like the sociobiological ideal, the parentela orders model is oriented toward the survival of any given line of descent (or failing that, the next closest line of descent). This model expresses centripetal tendencies in kinship structure. In marriage law, collateral prohibitions are minimal, and marrying someone in the closest line of descent (first cousin) is preferred.

As the parentela orders model is applied to intestacy law, the centripetal principle is expressed in the Hebrew Bible in Numbers 27:8–11 and 36:7–9. In this model, priorities among relatives are allocated by line of descent: (1) Direct descendants of Ego are given first priority (children,

grandchildren, etc.); (2) if there are no direct descendants, those of Ego's parents are given next priority (siblings, nieces and nephews, etc.); (3) next in priority are descendants of Ego's grandparents (aunts and uncles, first cousins, etc.)—and so on. In theory, Ego's estate will be passed on to the closest survivor in the closest line of descent to Ego's.

Between the extremes of centrifugality of the canon law model and the centripetality of the parentela orders model stands the civil law model. This model gives somewhat more weight in assigning closeness in kinship distance to direct-line ascendants and descendants than to collateral relatives (i.e., those related to Ego through a common ancestor). In computing kinship distance from Ego, the civil law model counts generations between Ego and the common ancestor as well as generations between the other relative and the common ancestor; for direct-line relatives, only those generations between Ego and the other relative need be counted. Obviously, the nearer the common ancestor is to Ego, the closer is the collateral relative in genealogical distance (and vice versa). Thus, Ego's grandparent is closer genealogically than a niece or nephew; the reverse is true for parentela orders priorities, and both are equidistant from Ego in the canon law model.

Several social surveys have been undertaken to test empirically the above propositions about ways in which people's conceptions about priorities assigned to different relatives in kinship mapping are actually reflected in their lives—religious affiliation, socioeconomic status, minority status, and so on. In these surveys, the respondents were asked to choose priorities among kin (for which the kinship-map models differ) if they were to write a law to govern intestacy (i.e., where there is no written will). For ten pairs of relatives for whom the kinship models differed in assigning a priority, within each pair, the respondents were to select the relative they thought should have precedence (as a general rule). (Equal priority was one alternative.) Respondents were then classified according to the kinship model to which a majority of their choices conformed.

The first surveys were undertaken in the United States (Farber 1977, 1979). The results indicate that Jewish respondents do indeed tend to view priorities from the perspective of the parentela orders model, while Catholics tend to be

overrepresented in the canon law category. Of course, these are tendencies and not blanket findings covering all Jews or Catholics. For example, the degree to which a religious grouping adheres to scripture and/or ritual practices seems important in influencing kinship mapping. When religious branch was taken into account, responses of Jews who identified themselves as Conservative (a fairly traditional branch) tended to conform to the parentela orders model and none conformed to the canon law model, while those in the Reform category more often conformed to the canon law model than to the parentela orders model. Similarly, among Mormons whose marriage was sealed in the Temple, their responses were like those of the Conservative Jews, whereas those whose marriage was not sealed for time and eternity responded like Reform Jews. Moreover, neofundamentalist Protestants were the only other religious grouping overrepresented in the parentela orders category (Farber 1981, pp. 73–75).

Taken together, the above findings suggest that the parentela orders model tends to be prevalent in groupings where endurance of the particular religious community into the distant future may be problematic. The community would then be motivated to intensify its inward pull—its centripetal incentive—to keep succeeding generations within the fold.

In the course of one investigation (Farber 1981), a reanalysis of findings yielded a fourth kinship model. This model, whose computation is the reverse of the parentela orders model, emphasizes obligations to ancestors who have been responsible for preparing the groundwork for Ego's place in society. In the serendipitous model, Ego's direct ancestors are given priority over any descendants—first priority is given to parents, grandparents, and so on; the next set of priorities consists of Ego's children, then Ego's brothers and sisters, aunts and uncles, great-uncles and great-aunts, and so on; following these, Ego's grandchildren, nieces and nephews, first cousins, and on and on (Farber 1981, p. 50).

The serendipitous model was disproportionately prevalent in several sectors of respondents—nonminority Protestants, those in professional and managerial occupations and at higher income levels, and those persons with U.S.-born fathers. In

addition, persons who conformed to this model tended to come from smaller families (Farber 1981, p. 217) and expected to have fewer offspring than did other respondents (Farber 1981, p. 147). Since almost half the sample studied conformed to this model, it seemed appropriate to name it the Standard American model. The U.S. findings on the standard American model are consistent with Alexis de Tocqueville's observation made almost two centuries ago in *Democracy in America* ([1850] 1945), namely, that compared with Continental Europeans, Americans live in the present and show little interest in the perpetuation of family lines.

European data on the genealogical models throw further light on differences in the conception of kinship priorities between U.S. and Continental populations. An investigation in central Europe (Vienna, Bremen, and Cologne) shows parentela orders to be by far the most prevalent kinship model, especially among those families at upper socioeconomic levels (Baker 1991). By way of contrast, Baker's (1991) data from Dublin, Ireland, tend to be similar to the American findings: Jews display a strong tendency to conform to the parentela orders model, while Protestants and Catholics favor the standard American model (called by Baker the intercultural bourgeois model). Despite all the changes that have occurred over the generations, traditional perceptions of priorities in kinship claims still persist.

The typology of kinship maps (or collaterality models) is a heuristic for understanding an implicit theory of the workings of kinship structure. In itself, the typology is too simplistic to denote the complexity of norms and values and the operation of mechanisms involved. But these criticisms about the heuristic character of the types of collaterality models can be applied to all typologies used in kinship analysis. They are merely methodological tools for gaining insight into what is going on. To gain this insight, one forgoes the many nuances that give color to understanding the functioning of kinship.

**Related Transhistorical Typologies.** Variations on issues pertinent to the structural contradiction typology have been developed in other transhistorical schemes associated with the role of marriage and descent systems in organizing family and kinship

systems. For instance, Guichard (1977) distinguishes between Eastern/Islamic and Western/Christian kinship systems. According to his typology, in the Eastern system, (1) descent is patrilineal; (2) marital ties are weak, and polygyny and easy divorce are permitted; (3) close ties exist between kin related through male lineage groups; (4) strong preference is given to endogamy within patrilineages; and (5) the sexes are segregated and women are relatively secluded within the home. In contrast, in the Western system, (1) kinship is bilinear or bilateral/multilateral, with ties to the maternal family considered important and with an emphasis on affinal connections as well; (2) marital bonds are the dominant unifying feature in family and kinship, with monogamy as prescribed and with extended kin ties as weak; (3) kin ties are defined according to individual connections rather than by lineage groups, with an emphasis on the ascending line rather than the descending line and with little importance attached to lineal continuity or solidarity; (4) kinship exogamy is prescribed, with endogamy permitted primarily for economic reasons; and (5) interaction between the sexes occurs in a wide range of circumstances.

In his reaction to Guichard, Goody (1983) revives the anthropological controversy between alliance theory and descent theory. Goody criticizes Guichard for basing his typology on marital norms (i.e., the endogamy-exogamy distinction) and suggests that by not starting with descent factors (i.e., inheritance practices), Guichard has overlooked a more fundamental distinction—that between kinship systems in which property is passed from one generation to the next through both sexes (by means of inheritance and dowry) and those systems in which property is transmitted unisexually (usually through males). Goody contends that passing property down unisexually encourages the development of corporate kinship groups (e.g., African systems). However, the use of bilateral devolution discourages such corporate structures, and Goody places both Eastern and Western systems in Guichard's dichotomy in the bilateral category. He faults Guichard for overstating the existence of corporate structures in Eastern kinship and proposes that Guichard's Western type represents merely a later historical development away from its roots in the Eastern system. Goody sees the primary problem of explaining the character of family and kinship in Western society

as one of discerning how European societies shifted from preferred kinship endogamy (e.g., first-cousin marriage) to prescribed exogamy.

In his analysis of European kinship, Goody considers the changes introduced by the Christian (i.e., Roman Catholic) church from its beginnings to the late medieval period. He interprets the shift from kinship endogamy to exogamy mainly as a strategic move by the church to gain control over the lives of its members. As part of this effort, it had to wrest access to resources (especially productive land) from enduring control by family and kin. As a result, church laws evolved favoring those norms that might enhance allegiance to the church and weaken competition from the family and the state. In consequence, the church favored (1) the use of testation permitting bequests to the church; (2) the prescription of kinship exogamy as a means for inhibiting both the reinforcement of close kin ties and the passing down of resources exclusively within lineages; (3) the requirement of the consent of both bride and groom in marriage; (4) late marriage as a means for weakening family control over mate selection; (5) prohibition of divorce even for childless couples; and so on.

Goody seems to overstate his case in trying to interpret the shifts in kinship in ways that are consistent with his basic typology. For example, in giving primacy to inheritance patterns, Goody asserts that the ban on divorce in Roman Catholicism was devised primarily to encourage bequeathing estates to the church in case of childlessness. But, in fact, when there were no children, bequests usually were made “to brothers and sisters and to nieces and nephews” (Sheehan 1963, p. 75). Moreover, Goody’s explanation of the ban ignores the widespread practice of bequeathing a portion of one’s estate to the church even when one left a widow, children, or both. Sheehan (1963) reports that these bequests were made for the good of the soul: “Among the Anglo-Saxons, bequests to the parish church became so general that they were eventually required by law” (p. 292). This practice was not restricted to England. According to Sheehan, “Christians in the Mediterranean basin had developed the practice of bequeathing part of their estate in alms” (p. 303). Thus, church heirship in medieval Christian Europe was tied to repentance regardless of the existence of familial beneficiaries. Since church acquisition did not have to depend on bequests from childless couples, it is

unlikely that the ban on divorce derives primarily from the desire of the church for additional benefices.

In addition, Goody dismisses the intermittent presence of kinship endogamy in medieval Europe as opportunistic deviations from the moral injunctions of the church. Yet, as Duby (1977) indicates, in medieval Europe the ebb and flow in kinship endogamy was tied to the amount of emphasis given to strengthening lines of descent. For example, Duby notes that in northern France, from before the tenth century to about the middle of the eleventh century, there was little utilization of the concept of lineage and only vague awareness of genealogy and knowledge about ancestors. Prior to that time, even members of the aristocracy considered their family to consist of “a horizontal grouping” of neighbors and kin “whose bonds were as much the result of marriage alliances as of blood” (Duby 1977, p. 147). Then, beginning in the tenth century, there was a change in ideas and norms regarding kinship—a conscious strengthening of lineage by controlling marriage, which frequently took place between close relatives despite impediments in canon law (Canon Law Society 1983). To summarize, Goody’s argument is that medieval deviation from canon law consisted of opportunistic economic decisions and did not derive from a different set of norms. But Duby describes the coordination of kinship endogamy with the emerging notion of the legitimacy of lineage—a complex of ideas that requires a consensus among the kin in order to be effective. Hence, it appears that the change in marriage rules and the significance of lineage signaled more than ad hoc departures from church law.

There is still another reason for questioning Goody’s conclusions: Goody makes the point that through bequests the Catholic church became the largest landowner in Europe. In his focus on the growth of exogamy as a consequence of the devolution of estates to both sexes, he has overlooked the church’s own involvement as a major heir in the inheritance system. Particularly in the light of the church’s view that ties through faith are equivalent to blood ties, the church is identified with spiritual kinship (Goody 1983, pp. 194ff). However, if it is legitimate to consider the church as an heir on a par with familial heirs, the system becomes one of *trilateral* devolution—sons, daughters, *and* the church. In that case, the European

system differs markedly from the Eastern kinship system described by Guichard. Indeed, in contrast to Judaism and Islam, Christianity, at least until the end of the medieval period, saw family and kinship ties as *competitive* with church interests, and the strategies the church applied to weaken these ties altered both the marriage and the inheritance systems. The data imply that, despite their contradictory implications, the marriage, the alliance component, and the descent component should be addressed as equal factors in organizing family life. A task that remains is to integrate typologies of the emergence of modern kinship systems with transhistorical, structural typologies.

(SEE ALSO: *Alternative Life Styles; American Families; Family and Household Structure; Family Roles*)

## REFERENCES

- Adams, Bert N. 1968 *Kinship in an Urban Setting*. Chicago: Markham.
- Atkins, John R. 1974 "On the Fundamental Consanguineal Numbers and Their Structural Basis" *American Ethnologist* 1:1-31.
- Attias-Donfut, Claudine 1997 "Home-Sharing and the Transmission of Inheritance in France." In M. Gullestad and M. Segalen, eds., *Family and Kinship in Europe*. London: Pinter.
- Augustine, Saint 1966 *The City of God Against the Pagans*. New York: Penguin Books.
- Baker, David J. 1991 *Conceptions of Collaterality in Modern Europe: Kinship Ideologies from Companionship to Trusteeship*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.
- Barnard, Malcolm 1993 "Economy and Strategy: The Possibility of Feminism." In Chris Jenks, ed., *Cultural Reproduction*. New York: Routledge.
- Bar-On, Dan 1989 *The Legacy of Silence: Encounters with Children of the Third Reich*. Cambridge, Mass.: Harvard University Press.
- Bendor, S. 1996 *The Social Structure of Ancient Israel*. Jerusalem: Simor.
- Berkner, Lutz 1972 "The Stem Family and the Developmental Cycle of a Peasant Household: An Eighteenth-Century Example." *American Historical Review* 77:398-418.
- Blau, Zena Smith 1974 "The Strategy of the Jewish Mother." In Marshall Sklare, ed., *The Jew in American Society*. New York: Behrman House.
- Buchler, Ira R., and Henry A. Selby 1968 *Kinship and Social Organization*. New York: Macmillan.
- Burgess, Ernest W. 1948 "The Family in a Changing Society." *American Journal of Sociology* 53:417-422.
- , Harvey J. Locke, and Mary Margaret Thomes 1963 *The Family: From Institution to Companionship*. New York: American Book Company.
- Canon Law Society of Great Britain and Ireland 1983 *The Code of Canon Law*. London: Collins Liturgical Publications.
- Chodorow, Stanley 1972 *Christian Political Theory and Church Politics in the Mid-Twelfth Century*. Berkeley: University of California Press.
- Craig, Daniel 1979 "Immortality through Kinship: The Vertical Transmission of Substance and Symbolic Estate." *American Anthropologist* 81:94-96.
- Davenport, W. 1959 "Nonunilinear Descent and Descent Groups." *American Anthropologist* 61:557-572.
- Douglas, Mary 1966 *Purity and Danger*. London: Routledge and Kegan Paul.
- Duby, Georges 1977 *The Chivalrous Society*. London: Edward Arnold.
- Engels, Frederick (1885) 1942 *The Origin of Family, Private Property, and the State*. New York: International Publishers.
- Farber, Bernard 1968 *Comparative Kinship Systems*. New York: Wiley.
- 1971 *Kinship and Class: A Midwestern Study*. New York: Basic Books.
- 1975 "Bilateral Kinship: Centripetal and Centrifugal Types of Organization." *Journal of Marriage and the Family* 37:871-888.
- 1977 "Social Context, Kinship Mapping, and Family Norms." *Journal of Marriage and the Family* 39:227-240.
- 1979 "Kinship Mapping Among Jews in a Midwestern City." *Social Forces* 57:1107-1123.
- 1981 *Conceptions of Kinship*. New York: Elsevier.
- 1984 "Anatomy of Nurture: A Structural Analysis of the Contemporary Jewish Family." Paper presented at Workshop on Theory Construction and Research Methodology, National Council on Family Relations, San Francisco, October.
- Firth, Raymond, Jane Hubert, and Anthony Forge 1969 *Families and Their Relatives*. New York: Humanities Press.
- Fortes, Meyer 1969 *Kinship and Social Order*. Chicago: Aldine.

- Foucault, Michel (1971) 1996 "The Discourse on Language." In R. Kearney and M. Rainwater, eds., *The Continental Philosophy Reader*. New York: Routledge.
- Ganzfried, Solomon 1963 *Code of Jewish Law (Kitzur Shulkhan Aruhh)*, rev., annot. ed. New York: Hebrew Publishing Company.
- Goode, William J. 1963 *World Revolution and Family Patterns*. New York: Free Press.
- Goody, Jack 1983 *The Development of the Family and Marriage in Europe*. New York: Cambridge University Press.
- Guichard, P. 1977 *Structures sociales 'Orientales' et 'occidentales' dans l'Espagne musulmane*. Paris: Mouton.
- Gullestad, Marianne 1997 "From 'Being of Use' to 'Finding Oneself': Dilemmas of Value Transmission between Generations in Norway." In M. Gullestad and M. Segalen, eds., *Family and Kinship in Europe*. London: Pinter.
- Harris, C. C., and Colin Rosser 1983 *The Family and Social Change*. Boston: Routledge and Kegan Paul.
- Hastrup, Kirsten 1982 "Establishing an Ethnicity: The Emergence of the 'Icelanders' in the Early Middle Ages." In David Parkin, ed., *Semantic Anthropology*. New York: Academic Press.
- Huxley, Aldous 1955 *Brave New World*. New York: Bantam.
- Larney, Barbara Elden 1994 Children of World War II in Germany: A life course analysis. Unpublished doctoral diss.
- Levi-Strauss, Claude 1963 *Structural Anthropology*. New York: Basic Books.
- 1969 *The Elementary Structures of Kinship*. Boston: Beacon Press.
- Lewis, Robert A., and Graham B. Spanier 1982 "Marital Quality, Marital Stability and Social Exchange." In F. Ivan Nye, ed., *Family Relationships: Rewards and Costs*. Beverly Hills, Calif.: Sage.
- Litwak, Eugene 1960a "Occupational Mobility and Extended Family Cohesion." *American Sociological Review* 25:9–21.
- 1960b "Geographical Mobility and Extended Family Cohesion." *American Sociological Review* 25:385–394.
- 1985 *Helping the Elderly: The Complementary Roles of Informal Networks and Formal Systems*. New York: Guilford Press.
- Lopata, Helena Znaniecki 1973 *Widowhood in an American City*. Cambridge, Mass.: General Learning Press.
- Macfarlane, Alan 1986 *Marriage and Love in England: Modes of Reproduction 1300–1840*. New York: Basil Blackwell.
- Maine, Henry S. (1861) 1963 *Ancient Law*. Boston: Beacon Press.
- Mitchell, William E. 1963 "Theoretical Problems in the Concept of the Kindred." *American Anthropologist* 65:343–354.
- Mogey, John 1976 "Content of Relations with Relatives." In J. Caisenier, ed., *The Family Life Cycle in European Societies*. Paris: Mouton.
- Murdock, George Peter 1949 *Social Structure*. New York: Macmillan.
- Naroll, Rauol 1970 "What Have We Learned from Cross-Cultural Surveys?" *American Anthropologist* 75:1227–1288.
- Paige, Jeffery M. 1974 "Kinship and Polity in Stateless Societies." *American Journal of Sociology* 80:301–320.
- Parsons, Talcott 1954 "The Kinship System of the Contemporary United States." In Talcott Parsons, ed., *Essays in Sociological Theory*. New York: Free Press.
- Pehrson, R. N. 1957 *The Bilateral Network of Social Relations in Konkama Lapp District*. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.
- Peranio, R. 1961 "Descent, Descent Line, and Descent Group in Cognatic Social Systems." In V. E. Garfield, ed., *Proceedings of the Annual Meeting of the American Ethnological Association*. Seattle: University of Washington Press.
- Pina-Cabral, Joao de 1997 "Houses and Legends: Family as a Community of Practice in Urban Portugal." In Marianne Gullestad and Martine Segalen, eds., *Family and Kinship in Europe*. London: Pinter.
- Redfield, Robert 1947 "The Folk Society." *American Journal of Sociology* 52:293–308.
- Roschelle, Anne R. 1997 *No More Kin: Exploring Race, Class, and Gender in Family Networks*. Thousand Oaks, Calif.: Sage Publications.
- Sennett, Richard 1970 *Families Against the City: Middle Class Homes of Industrial Chicago, 1872–1890*. Cambridge, Mass.: Harvard University Press.
- Shanas, Ethel, Peter Townsend, Dorothy Wedderburn, Henning Friis, Paul Milhoj, and Jan Stehouwer 1968 *Old People in Three Industrial Countries*. New York: Atherton Press.
- Sheehan, Michael M. 1963 *The Will in Medieval England: From the Conversion of the Anglo-Saxons to the End of the Thirteenth Century*. Toronto: Pontifical Institute of Medieval Studies.
- Sorokin, Pitirim 1937 *Social and Cultural Dynamics*. Four Volumes. New York: Harper.
- Stack, Carol B. 1974 *All Our Kin*. New York: Harper Colophon Books.



- Steinmetz, Devora 1991 *From Father to Son: Kinship, Conflict, and Continuity in Genesis*. Louisville, KY: Westminster/John Knox Press.
- Stone, Lawrence 1975 "Rise of the Nuclear Family in Early Modern England: The Patriarchal Stage." In Charles E. Rosenberg, ed., *The Family in History*. Philadelphia: University of Pennsylvania Press.
- Strathern, Marilyn 1992 *After Nature: English Kinship in the Late Twentieth Century*. New York: Cambridge University Press.
- Sussman, Marvin 1959 "The Isolated Nuclear Family: Fact or Fiction?" *Social Problems* 6:333-340.
- Swanson, Guy E. 1969 *Rules of Descents: Studies in the Sociology of Parentage*. Anthropological Papers, no. 39. Ann Arbor: Museum of Anthropology, University of Michigan.
- Tocqueville, Alexis de (1850) 1945 *Democracy in America*. New York: Knopf.
- Toennies, Ferdinand (1887) 1957 *Community and Society*. East Lansing: Michigan State University Press.
- Walster, Elaine, and G. William Walster 1978 *A New Look at Love*. Reading, Mass.: Addison-Wesley.
- Watson, John 1927 *Chicago Tribune*. March 6, p. 1.
- Weigert, Andrew J., and Ross Hastings 1977 "Identity Loss, Family, and Social Change." *American Journal of Sociology* 82:1171-1185.
- Wirth, Louis 1956 *Community Life and Social Policy*. Chicago: University of Chicago Press.
- Yerushalmi, Yosef Hayim 1982 *Zakhor: Jewish History and Jewish Memory*. Seattle: University of Washington Press.
- Young, Michael, and Peter Willmott 1957 *Family and Kinship in East London*. London: Routledge and Kegan Paul.
- Zborowski, Mark, and Elizabeth Herzog 1952 *Life Is with People: The Culture of the Stettin*. New York: Schocken Books.
- Zimmerman, Carle C., and Merle E. Frampton 1947 *Family and Civilization*. New York: Harper.
- 1966 "Theories of Frederic LePlay." In Bernard Farber, ed., *Kinship and Family Organization*. New York: Wiley.

BERNARD FARBER

# L

---

## LABELING THEORY

See Deviance Theories.

## LABOR FORCE

Although labor-force concepts were originally designed to study economic activity and guide government policies, economic activities are a form of *social* behavior with numerous social determinants and consequences. Hence labor-force behavior has been the subject of a substantial body of sociological research.

### MEASUREMENT

The U.S. Bureau of the Census developed the labor-force concept to measure the number of working-age people who were economically active during a particular time period—the calendar week preceding the sample interview (Cain 1979; U.S. Bureau of Labor Statistics 1982). It has two components: (1) *The employed*: those who, during the reference week, did any work at all as paid employees, were self-employed, or worked as unpaid family workers at least fifteen hours in a family-operated enterprise; included also are those who were employed but on vacation, home sick, etc. (2) *The unemployed*: those who were not employed during the reference week but who were available for work and had actively sought employment sometime within the preceding four-week period. All those who are neither employed nor unemployed

are defined as being out of the labor force and primarily include students, housewives, the retired, and the disabled. Since the size of the population affects the number of people who work, labor-force measures are usually expressed in ratio form. The *labor-force participation rate* is the percentage of the total working-age *population* that is in the labor force, while the *unemployment rate* is the percentage of the *labor force* that is unemployed. In order to compare particular subgroups in the population, analysts compute group-specific measures such as the percentage of all women versus men who are in the labor force or the percentage of black versus white labor-force members who are unemployed.

The Census Bureau developed the labor-force concepts during the Great Depression of the 1930s in response to the government's difficulty in charting the severity of unemployment during that crisis. Prior to 1940, measures of economic activity were collected only at the time of the decennial censuses, making it impossible to track business-cycle fluctuations in unemployment; for example, most of the Great Depression came between the 1930 and the 1940 censuses. Hence, in order to provide ongoing unemployment data, the Census Bureau initiated the monthly Current Population Survey in the 1940s.

A second problem was the ambiguity of the previously used measure of economic activity—the “gainful worker concept”—which was designed to ascertain individuals' *usual* occupation, if they had one, rather than whether they were actually working at any given time (Hauser 1949). In fact,

census enumerators were often specifically instructed to record an occupation, even if the individual was currently unemployed, thus *overstating* the number employed. On the other hand, some kinds of employment were often *underestimated* because people who considered their market work to be secondary to their other activities, such as taking care of the home and children or going to school, were less likely to report themselves as employed in response to a question on their usual occupation. Misreporting of this type is unlikely with the labor-force measure since most people will remember whether they had worked at all the previous week or, if not, whether they had been looking for a job. Information on occupation and on other important characteristics of their employment, such as hours worked, was then obtained separately in response to additional questions.

While labor-force concepts are relatively unambiguous measures of current economic activity, they too exhibit problems. One general concern is the adequacy of the unemployment measure. The extent of unemployment may be understated if persistently unemployed persons eventually give up trying and drop out of the labor force. The Census Bureau has therefore included additional questions to try to ascertain the number of such “discouraged workers” as well as to measure additional aspects of unemployment or underemployment (Cain 1979; Bregger and Haugen 1995). Clifford Clogg and Teresa Sullivan have extended this approach to address the larger question of “underemployment” (Sullivan 1978; Clogg and Sullivan 1983). They have developed and applied a variety of indicators of underemployment in order to achieve a more extensive assessment of the problem. In addition to the usual unemployment rate and estimates of discouraged workers, they use three other indicators—a measure of involuntary part-time work (due to economic factors); low work-related income relative to the poverty level, and a measure of the proportion of workers who are “overeducated” (“mismatched”) for the jobs they hold. The “adequately employed” are all those who are not underemployed in any one of these five categories. Their results indicate that underemployment is more common among the young and the old, and appears to have increased in recent years. While several of the indexes, particularly the mismatch measure, are somewhat

controversial (Keyfitz 1981), the work of Sullivan and Clogg represents an important innovation in the multidimensional measurement of underemployment.

A major characteristic of the labor-force concept is that it is a measure of *market-oriented* economic activities. People are considered employed only if they work for pay (or in the production of goods or services for sale). Yet there is a considerable amount of economic production for *home* consumption. Hence, labor-force status per se is an imperfect indicator of whether an individual is economically productive; for example, full-time homemakers are never counted as employed although they usually put in long hours producing goods and services for their families. However, if the market-oriented nature of the measure is kept in mind, this limitation is not too serious in a modern industrial society. Increases in married women’s labor-force participation can then be interpreted as indicating their growing participation in the *market* sector of the economy, usually in addition to their *home* productive activities, although working wives do less housework than nonworking wives (Vanek 1974; Berardo et al. 1987).

More serious problems arise in comparing societies at different levels of economic development (Moore 1953). Preindustrial subsistence economies produce few goods or services for a market. As societies develop economically, an increasing proportion of labor is sold in the marketplace, and the goods and services families consume are also increasingly purchased rather than home-produced. It is often difficult to undertake a meaningful comparison of labor force or unemployment rates among such different economies. The measurement of agricultural employment, especially that of women and youth, can be particularly problematic in countries with a large subsistence sector, and measurement inconsistencies are common (Dixon 1982). Moreover, *unemployment*, especially in rural areas, is often manifested as *underemployment*, and its extensiveness is difficult to determine.

Although best suited for examining whether people are currently economically active, labor force and employment status measures have also been invaluable in the analysis of more complex sociological and economic concepts because of

the ready availability of these data in time series. However, here it is important to recognize the ambiguities and limitations of such measures at the same time we exploit their utility. One example of this problem has to do with charting the extensiveness of changes in married women's economic role in the family; another concerns the measurement of life-course transitions such as the transition to work or to retirement. In the first case, the frequently reported time series of the changes in the average proportion of married women who were in the labor force in any given week each year, or even of the proportion who had worked at some time during a year, are valuable but also imperfect indicators of how extensively women's economic role has changed over time. For example, the proportion of married women with children under age 18 who were employed at some point during the year rose from 51 to 73 percent between 1970 and 1990. However, although the proportion who worked full-time, year-round had also increased considerably, it had only reached 34 percent by 1990; of women with children under age 6, only 28 percent had worked full-time year-round in 1990 (Bianchi 1995, p. 117). Hence, the commonly used time series of married women's labor-force status during as short a time period as a week will exaggerate the magnitude of the changes in wives' economic role in the family. And neither annual labor-force participation rates nor the weeks and hours worked during a whole year provide longitudinal data on the extensiveness of individual women's labor-market involvement over their adult life courses. Yet it is this sort of information that we would really like to have in assessing the changing nature of women's economic roles.

Although commonly used for this purpose, labor-force or employment status data also have their limitations as indicators of the timing of life-course transitions. The problem is that life-course transitions are not as clear-cut as the data make them appear (Oppenheimer and Kalmijn 1995; Oppenheimer et al. 1997). Students are increasingly likely to be working, at least part time, and young people go in and out of the labor force before they are able—or willing—to make a regular commitment to year-round full-time employment. Hence, the proportion of young people employed in any given week overstates whether they have completed the transition to work. Changes

in the proportion of *older* persons who are currently employed is also an ambiguous indicator because of a fair amount of labor-market turnover among this age group as well. Hence, some studies of retirement use information on when individuals start to receive pensions. However, older persons may be receiving private pensions and/or Social Security but still be working, if only part time; moreover, the retirement these pensions signify may not be entirely voluntary, complicating our interpretation of the phenomenon. So the goal of measuring when "permanent" withdrawal from work occurs can be quite elusive (Guillemard and Rein 1993; Henretta 1992).

### DETERMINANTS AND CONSEQUENCES OF LABOR-FORCE CHANGES

The size and rate of growth of the labor force are dependent on three factors:

1. *The size and rate of population growth.* A large and/or rapidly growing population will produce a large and/or growing labor force.
2. *The propensity of the population to enter the labor force and how this varies among population subgroups.* Age and sex, and what these signify biologically and socially, are the major reasons for varying propensities. Infants and young children do not work, but, generally starting in adolescence, labor-force participation increases with age, peaking for those in their late thirties and early forties and starting an accelerating decline thereafter. Married women, particularly mothers of young children, have historically had lower labor-force participation rates than adult males, although this is much less so now than in the past (U.S. Bureau of the Census 1998, pp. 408–409).
3. *The composition of the population:* Since different population segments have different work propensities, the composition of the population will affect the overall proportions who are in the labor force. Sharp *short-run* fluctuations in the U.S. birthrate have led to corresponding variations in the relative size of the working-age population. Baby booms greatly increase

the number of new labor-force entrants after 16–18 years, while baby busts reduce this number. However, the overall *long-run* declines in U.S. fertility, combined with declines in mortality among the elderly, have increased the relative number of elderly in the population, an age group with low work propensities. On the other hand, foreign migration to the United States somewhat counteracts the effects of an aging population, since it has historically been disproportionately composed of young working-age adults, drawn to the United States by job opportunities.

One important long-term trend in labor-force participation in the United States has been the decline in the employment rate of young men, on the one hand, and of older men, on the other. In part, the decline for *younger* males has been due to more extended schooling, although this was somewhat offset by a rise in student employment (U.S. Bureau of Labor Statistics 1988). However, an additionally significant factor has been a decrease in the employment of moderately to less educated males, itself just one aspect of a trend in rising labor market inequality, encompassing declines in earnings as well as levels of employment (Burtless 1990; Levy 1998).

Another important long-term shift is the decline in the labor-force participation of older males, primarily due to the institution and spread of the social security system combined with the greater availability of disability benefits and private pensions. However, these declines have been observed not only for men aged 65 and older but also for men in their fifties and early sixties, although not to the same extent. Thus the labor-force participation rates for men aged 65–69 decreased from 64 percent in 1950 to 26 percent in 1990, but during these years the rates of men 60–64 and 55–59 also substantially declined, from 83 to 56 and from 90 to 80 percent, respectively (Gendell and Siegel 1992, p. 24). The increasing coverage of Social Security benefits for men retiring at age 65 and the institution of early retirement, at age 62, under the Social Security Act of 1962 has played the major role in the decreasing rates for those in their sixties. However, the declines for men in their fifties must be for other reasons. One important factor seems to be the growth of private pension

plans and employers' utilization of early retirement provisions to help downsize and restructure their firms (Guillemard and Rein 1993; Henretta 1992).

An important question is: How will the rapid rise in the older population affect the labor-force participation of older people? There is by no means an obvious answer to this question. For one thing, given the sensitivity of older persons' employment behavior to when Social Security is available, future changes in how old individuals must be in order to qualify for Social Security benefits will play an important role. Already the 1983 amendments to the Social Security Act have set in motion a rise in the age of entitlement to a full pension—from age 65 to 66 by 2009, and to 67 by 2027—and a higher rate of reduction will gradually be applied to pensions for those who retire earlier than 65. These changes in the law should operate to increase labor-force participation among the elderly. Other factors may also increase the employment of men in their fifties and sixties. Traditionally, the decline in employment over time has been disproportionately concentrated among the less educated who have fewer marketable skills and for whom the Social Security pension is relatively more attractive. However, with the rising educational attainment of the population, older adults in the twenty-first century might be expected to remain employed longer (Besl and Kale 1996). On the other hand, the state and structure of the economy is an important factor in the extent to which employers use private pension plans to encourage earlier retirement in periods of downturns or rapid structural change.

Probably the most substantial postwar change in employment behavior has been the enormous increase in married women's labor-force participation. While paid employment used to be generally limited to the period between school and marriage, since the 1940s married women's employment has become so prevalent that by 1997 between 66 and 76 percent of those in the 20–44 age groups were in the labor force. Moreover, 64 percent of married mothers of children under age 6 were also in the labor force (Oppenheimer 1970; U.S. Bureau of the Census 1998).

There are several reasons for this rapid rise in women's employment. One is that the bureaucratization of government and industry has raised the

demand for clerical workers; population growth, prosperity, and rising living standards have greatly expanded the consumption of services. This, in turn, raises the demand for sales workers and for those in the “helping” professions, such as teachers and social workers, as well as nurses and others in health-related occupations. All these are occupations that have been dominated by women workers for over a century. However, because the great majority of young, single, out-of-school women have worked throughout the twentieth century, the result of this increasing demand in the postwar period has been a strong and continuing demand for a previously underutilized source of female labor—married women (Oppenheimer 1970).

Married women have often had several major reasons for wanting to work—the need, early in marriage, to help set up a new household and perhaps to save money for a down payment on a house; the increasing importance of saving for children’s schooling; the couple’s aspiration to achieve a high level of living; and the desire for greater personal economic security and autonomy. Periodic rises in the cost of living and the stagnating or even declining economic position of many men since the 1970s have also increased the importance of having two earners in a family (Levy 1998).

The effects of changing labor-force behavior are not just limited to the economic realm; a number of sociologists (as well as economists) have argued that this behavior has also had an important impact on marriage and the family. Since the late 1960s the average age at marriage has risen substantially, after having first declined through most of the twentieth century; nonmarital cohabitation has become increasingly prevalent; and marital instability has accelerated its long-term upward trend after a sharp reversal in the early post-World War II period. However, the divorce rate appears to have stabilized recently. Two competing employment-related explanations for these trends are currently under debate. In one, the argument is that married women’s rapidly rising labor-force participation has increased their economic independence of males (Becker 1981; Espenshade 1985; Goldscheider and Waite 1986; Farley 1988; McLanahan and Casper 1995). The result, it is argued, is a decreasing desire on the part of women to remain in an unhappy marriage or even to marry at all. In addition, since women’s

traditional time- and energy-consuming familial roles of childbearing and childrearing compete with the pursuit of individual career goals, more women are either forgoing childbearing entirely or settling for one or two children at most.

While the women’s “independence” hypothesis appears plausible and the juxtaposition of time series data on marriage and family behavior with that of women’s labor-force participation appears to support it, this is largely because the time series utilized have typically been limited to the postwar period, the period during which married women’s employment was rising rapidly. The problem with these comparisons is that they use the family behavior of the early 1950s as the model of “traditional” family behavior against which to compare subsequent trends. However, the marriage and fertility behavior of the 1950s was by no means traditional (Cherlin 1992; Oppenheimer 1994). This was the baby boom era when, after a 150-year decline, the total fertility rate reversed itself and rose so much that, by the 1950s, it was back up to the level of 1900. Moreover, age at marriage had been decreasing throughout the twentieth century, at the same time that women’s employment was rising, so that the early postwar age at marriage was much younger than what was “traditional” before married women’s employment began its historic climb (Cherlin 1992; Oppenheimer 1994). In sum, the changes since the 1960s are more of a return to traditional patterns than a major departure from them. The independence hypothesis has also not held up to more recent micro-level empirical analyses using longitudinal data. For example, as summarized by Oppenheimer (1997), the evidence from several studies indicates that single women’s labor-market position tends to have little effect on marriage formation, but that what effect it does have is positive.

A characteristic of the women’s independence explanation of recent trends in family behavior is that it ignores the possible role of *men’s* changing economic position in these changes. Yet it is well known that marriage and family behavior are related to men’s employment characteristics (Cherlin 1979; Goldscheider and Waite 1986; Ross and Sawhill 1975; Teachman et al. 1987). Furthermore, there is a long demographic tradition, dating back to Malthus, which argues that changes in men’s economic position has an effect on marriage and fertility behavior. Supporting the view that these

changes might also be an important factor in the rise in both men's and women's age at marriage since the early 1970s is the well-documented finding that there has been a large absolute and relative decline in the labor-market position of young men with a high school education or less. In addition, there has been an increase in economic inequality within each educational group, including those with a college education (Levy 1998; Juhn et al. 1993). The reasons for these trends appear to be quite complex, and, to date, no general consensus has yet been reached regarding the relative importance of several proposed explanations. A variety of factors appear to be making a contribution. The globalization of manufacturing and the resulting competition from cheap semi-skilled labor in developing countries may be decreasing the demand for less skilled American workers, workers who previously were able to command relatively high wages in manufacturing; moreover, this same globalization has weakened the position of unions and hence their ability to protect such workers from an erosion in their job security and wages. All this has fostered the continued decline in the proportion of workers in manufacturing and the rapid rise in the proportion in service industries. However, the growth of service industries per se cannot be driving these changes because the growth in some service industries has increased the demand for more highly skilled labor while the expansion of other service industries has only resulted in a rising demand for low-wage unskilled labor, leaving the semiskilled in an increasingly poor position. Moreover, it does not appear that industrial restructuring alone can satisfactorily account for all or perhaps even most of these changes because, *within* industries, there is evidence of a sharp rise in the demand for more skilled labor pointing to an important role for technological change in both the manufacturing process and the production of services (Levy 1998; Meisenheimer 1998).

Whatever the reasons for the declining economic position of young, less educated males, recent research provides evidence that men in a poorer labor-market position do tend to delay marriage (Mare and Winship 1991; Lloyd and South 1996; Oppenheimer et al. 1997). Oppenheimer's work, in particular, has indicated that substantial inequalities in the length and difficulty of the career-entry process exist both within

and between race-schooling groups and lead to substantial differences in the marriage timing of young men.

This article has reviewed the history of labor-force measures as well as several important current issues in labor-force analysis. First, the study of the labor force reveals the changing significance of work in the lives of different segments of the population. Second, since economic behavior impacts on other social systems, such as the family and stratification systems, labor-force analysis will continue to be an essential field for sociological analysis.

## REFERENCES

- Becker, Gary S. 1981 *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.
- Berardo, Donna Hodgkins, Constance L. Shehan, and Gerald R. Leslie 1987 "A Residue of Tradition: Jobs, Careers, and Spouses' Time in Housework." *Journal of Marriage and the Family* 49:381-390.
- Besl, John R., and Balkrishna D. Kale 1996 "Older Workers in the 21st Century: Active and Educated, a Case Study." *Monthly Labor Review* 119:18-28.
- Bianchi, Suzanne 1995 "Changing Economic Roles of Women and Men." Pp. 107-154 in Reynolds Farley ed., *State of the Union: America in the 1990s*, Vol. 1. New York: Russell Sage.
- Bregger, John E., and Steven E. Haugen. 1995. "BLS Introduces New Range of Alternative Unemployment Measures." *Monthly Labor Review* 118:19-26.
- Burtless, Gary, ed. 1990 *A Future of Lousy Jobs?* Washington: Brookings Institution.
- Cain, Glenn C. 1979 "Labor Force Concepts and Definitions in View of Their Purposes." *Concepts and Data Needs-Appendix*, vol. 1. Washington: National Commission on Employment and Unemployment Statistics.
- Cherlin, Andrew J. 1979 "Work Life and Marital Dissolution." Chapter 9 in George Levinger and Oliver C. Moles, eds., *Divorce and Separation: Contexts, Causes and Consequences*. New York: Basic Books.
- 1992 *Marriage, Divorce, and Remarriage*. Cambridge, Mass.: Harvard University Press.
- Clogg, Clifford C., and Teresa A. Sullivan 1983 "Labor Force Composition and Underemployment Trends, 1969-1980." *Social Indicators Research* 12:117-152.
- Dixon, Ruth B. 1982. "Women in Agriculture: Counting the Labor Force in Developing Countries." *Population and Development Review* 8:539-561.

- Espenshade, Thomas J. 1985 "Marriage Trends in America: Estimates, Implications, and Underlying Causes." *Population and Development Review* 11:193-245.
- Farley, Reynolds 1988 "After the Starting Line: Blacks and Women in an Uphill Race." *Demography* 25:477-495.
- Gendell, Murray, and Jacob S. Siegel. 1992. "Trends in Retirement Age by Sex, 1950-2005." *Monthly Labor Review* 115:22-29.
- Goldscheider, Frances Kobrin, and Linda J. Waite 1986 "Sex Differences in the Entry into Marriage." *American Journal of Sociology* 92:91-109.
- Guillemard, Anne-Marie, and Martin Rein 1993 "Comparative Patterns of Retirement: Recent Trends in Developed Societies." *Annual Review of Sociology* 19:469-503.
- Hauser, Philip M. 1949 "The Labor Force and Gainful Workers—Concept, Measurement and Comparability." *American Journal of Sociology* 54:338-355.
- Henretta, John C. 1992 "Uniformity and Diversity: Life Course Institutionalization and Late-Life Work Exit." *The Sociological Quarterly* 33:265-279.
- Juhn, Chinhui, Kevin Murphy, and Brooks Pierce 1993 "Wage Inequality and the Rise in Returns to Skills." *Journal of Political Economy* 101:410-442.
- Keyfitz, Nathan 1981 "Review of *Measuring Underemployment*, by C. C. Clogg." *American Journal of Sociology* 86:1163-1165.
- Levy, Frank 1998 *New Dollars and Dreams*. New York: Russell Sage.
- Lloyd, Kim M., and Scott J. South 1996 "Contextual Influences on Young Men's Transition to First Marriage." *Social Forces* 74:1096-1119.
- Mare, Robert D., and Christopher Winship 1991 "Socioeconomic Change and the Decline of Marriage for Blacks and Whites." Pp. 175-202 in Christopher Jencks and Paul E. Peterson, eds., *The Urban Underclass*. Washington: Brookings Institution.
- McLanahan, Sara, and Lynne Casper 1995 "Growing Diversity and Inequality in the American Family. In Reynolds Farley, ed., *State of the Union: America in the 1900s*, vol 1. New York: Russell Sage Foundation.
- Meisenheimer, Joseph R., II 1998 "The Services Industry in the 'Good' versus 'Bad' Jobs Debate," *Monthly Labor Review* 121:22-47.
- Moore, Wilbert E. 1953 "The Exportability of the Labor Force Concept." *American Sociological Review* 18:68-72.
- Oppenheimer, Valerie Kincade 1970 *The Female Labor Force in the United States: Demographic and Economic Factors Governing Its Growth and Changing Composition*. Population Monograph Series, no. 5. Institute of International Studies, University of California, Berkeley.
- 1994 "Women's Rising Employment and the Future of the Family in Industrial Societies." *Population and Development Review* 20:293-342.
- 1997 "Women's Employment and the Gain to Marriage: The Specialization and Tradition Model." *Annual Review of Sociology* 23:431-453.
- , and Matthijs Kalmijn 1995 "Life-Cycle Jobs." *Research in Social Stratification and Mobility* 14:1-38.
- , Matthijs Kalmijn, and Nelson Lim 1997 "Men's Career Development and Marriage Timing during a Period of Rising Inequality." *Demography* 34:311-330.
- Ross, Heather L., and Isabel V. Sawhill. 1975. *Time of Transition: The Growth of Families Headed by Women*. Washington, : Urban Institute.
- Sullivan, Teresa A. 1978 *Marginal Workers, Marginal Jobs: The Underutilization of American Workers*. Austin: University of Texas Press.
- Teachman, Jay D., Karen A. Polonko, and Geoffrey K. Leigh 1987 "Marital Timing: Race and Sex Comparisons." *Social Forces* 66:239-268.
- U.S. Bureau of Labor Statistics 1982 *BLS Handbook of Methods*, vol 1. Washington: Government Printing Office.
- 1988 *Labor Force Statistics Derived from the Current Population Survey*, Bulletin 2307. Washington: Government Printing Office.
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States: 1998*. Washington: Government Printing Office.
- Vanek, Joan 1974 "Time Spent in Housework." *Scientific American* 231:116-120.

VALERIE KINCADE OPPENHEIMER

## LABOR MOVEMENTS AND UNIONS

Labor movements are collective activities by wage and salaried workers in market societies to improve their economic, social, and political status. The main manifestations of such movements are labor unions and political parties, but sometimes they include producer and consumer cooperatives; credit unions; newspapers; and educational, welfare, cultural, and recreational organizations. Labor movements and unions need sociological analysis because they are integral parts of two



major and related institutions of society, the economy and polity. Apart from bringing about changes in these institutions, they are the main vehicles for mobilizing the class interests of wage and salaried employees. No other social science discipline offers such a broad perspective of study.

Yet, in the United States since the inception of sociology, labor movements have received surprisingly little attention. From their founding up to 1999, three main journals—*The American Journal of Sociology*, *The American Sociological Review*, and *Social Forces*—together representing 219 years of publication, published only sixty-three articles whose titles mention labor movements, unions, or strikes. American sociological research on the topic has fluctuated with labor's fortunes. As union membership grew from 1940 to 1960, research expanded and then lagged, with falling membership in the 1970s. The more rapid decline of labor since 1980 has recently stimulated research on the causes.

The bulk of American labor research has been done by historians and labor economists. To be sure, sociologists have made contributions while working on other topics such as social stratification (e.g., working-class formation, income inequality), organizations (leadership turnover), race and gender (discrimination in unions), political sociology (party preferences of union members), case studies of industry (shop-floor life, the labor process), and social movements (Jenkins 1985). Combining these contributions with those of labor economists and historians, a sizable literature is now available (see the bibliography in Stern and Cornfield 1996). In Europe, the bulk of labor research has been done by sociologists and historians.

### ORIGIN OF THE LABOR MOVEMENT

Labor movements and unions emerged with the rise of capitalism, the Industrial Revolution, and free labor markets in eighteenth-century Europe. Historians agree that labor unions did not evolve from medieval guilds, which were status groups of master-owners whose monopoly of skills was protected by public authorities. Journeymen, apprentices, and laborers had fewer or no privileges (Pirenne 1932; Lederer 1932). In early capitalism, factory workers created ad hoc organizations to withhold their labor from employers, to control production and job training, and to protect

wages and working conditions. Re-created with recurring crises, the organizations eventually became permanent (Jackson 1984). As product markets grew and spread, unions were forced to organize new locals in those markets to prevent wage competition among communities. The labor movement became larger, more institutionalized, and more diversified as it organized workers in different occupations, industries, and regions (Sturmthal 1974).

This fragmented response to threats forced labor leaders to press for a more united and centralized organization to respond to threats wherever and whenever they appeared. Invariably, some unions were reluctant to commit their resources for the welfare of a vague "movement." Failure to consider this persistent resistance has led many scholars to equate labor movement growth with the formation of the working class and with working-class politics. Although the two are related, their linkage varies enormously in different times and places. Where labor movements first emerged, they were not class movements, but efforts by a minority of skilled workers to protect their traditional privileges (Calhoun 1982). Even when unions expanded to include most workers, internal factions remained, based on skill, industry, status, and influence (Form 1985, p. 96). The emergence of class movements, on the other hand, involved complex processes of linking labor movements to other special interest groups and political parties (Katznelson and Zolberg 1986).

The character and strength of labor movements must be explained in the context of the societies in which they emerge, especially the ways they relate to distinctive traditions, economies, and political and governmental systems. Autonomous labor movements survive best in capitalist democratic industrial societies. To sustain free collective bargaining, labor, management, and government must exhibit considerable independence but not exert overwhelming power in the tripartite relationship. Where union membership is compulsory and universal, where unions are completely dominated by government and/or enterprise managers, unions lack the autonomy and strength to advance the special interests of workers. Paradoxically, where labor has total control over government and the economy, it lacks opposition and the attributes of free labor movements (Sturmthal 1968).

### TYPES OF LABOR MOVEMENTS

Labor movements vary in structure and behavior according to their relationships to government and other institutions. Movements fall into roughly five types. In the independent type, as in the United States, labor is independent of all major institutions, especially parties, government, and religion. Although labor seeks political influence, it participates in a shifting multiclass coalition of a particular party. Corporate labor movements, as exhibited in Britain, Scandinavia, Germany, Italy, and France, are often formally incorporated into the political system, sometimes playing a dominant role in labor, religious, socialist, or social democratic parties. When such parties win electoral victories, labor participates in governmental bodies—cabinets, legislatures, and government agencies. In the third or party-dominant movement, sometimes found in developing societies with mixed economies, labor is part of a permanent ruling coalition—for instance, the Institutional Revolutionary Party of Mexico. Here labor loses some freedom to push for the special interests of workers because it is rarely strong enough to resist decisions of the coalition. In totalitarian fascist and communist regimes, labor movements are totally subordinate to the ruling party and exhibit the least independence. Finally, in enterprise labor movements, labor functions primarily in a consultative capacity at the enterprise level. Nationally, labor belongs to a loose federation of unions with weak links to major parties, organizations, and institutions.

In independent movements, labor seeks economic gains primarily by bargaining with management. In the political realm, it seeks governmental protection for the right to organize and bargain as well as protective legislation such as unemployment, old age, and medical insurance. In the corporate type, apart from bargaining with employers, labor makes gains through legislation that forces management to deal with labor in arriving at enterprise policies regarding job rights, the organization of work, investment and other decisions, as in codetermination in Germany and Scandinavia (Nutzinger and Backhaus 1980). Public ownership of certain industries (often mines, public utilities, and transportation) is also an option. Labor strength is highest in this type of movement. In the party-dominant movement, labor at best makes gains for a minority of organized workers,

stratifying the working class. Under totalitarian regimes, labor may be given certain functions, such as assigning housing or supervising cooperatives, but its ability to bargain with government and enterprise management is severely limited (Lane and O'Dell 1978). In enterprise unionism, organized workers may gain employment security and career rewards in exchange for loyalty to the enterprise. The unorganized are exposed to the vicissitudes of the market (Okochi et al. 1974).

### AMERICAN EXCEPTIONALISM

Scholars have long tried to explain why the American labor movement has not developed into the sort of corporatist socialist movement found in other advanced industrial democracies (Sombart 1906). Most scholars agree that European societies, compared to the United States, have had longer and stronger links to past institutions. When landed aristocrats, business, military, and religious elites resisted worker participation in the political system, class-oriented parties appeared. In such class environments, labor unions developed or joined other parties to obtain voting rights for workers, government protection of unions, benefits unavailable through collective bargaining, and eventual public ownership of industries. In addition to supporting class parties, labor movements developed structures attuned to their interests—for example, an intellectual elite, cooperatives, newspapers, banks, schools, and recreational clubs. In response to socialist and communist movements, Catholics not only launched their own labor movements and parties, but also organized schools, hospitals, newspapers, and clubs to embrace workers in a harmonious class-inclusive environment (Knapp 1976). Rejecting “captured” socialist, communist, and religious labor movements, liberals sought to organize free or neutral unions and parties that appealed to some workers and middle-class adherents (Sturmthal 1968).

The labor movement in the United States faced a different environment. The absence of an agricultural aristocracy and traditional governmental, military, and religious elites dampened class sentiments. Early extension of suffrage to all adult males removed that objective as a rallying cry for labor parties. The rapid expansion of industry into new cities, high rates of internal migration and immigration, the separation of ethnic groups in

neighborhoods, and religious diversity slowed the formation of multiple working-class bonds. Moreover, an aggressive capitalist class, not bound by traditional obligations toward subordinates, fashioned laws and courts to protect property rights and suppress unions as conspiratorial monopolies (Dougherty 1941, pp. 635–677).

Even so, trade assemblies and craft unions emerged in several cities in the decade after the Civil War. In the 1880s, the Knights of Labor tried to organize unions that included all workers, even the white-collar and small businesses. This attempt lasted roughly a decade. Beginning in 1905, the militant International Workers of the World (IWW) organized workers of all skills to join unions and engage in political action to destroy capitalism (Dougherty 1941, pp. 317–349). The IWW too lasted a decade. For a few years, splintered socialist parties tried to support class-oriented unions. These efforts failed largely because the American political system places structural limitations on the development of third parties. The presidential system, decentralized state structures, constitutional barriers to the federal government making national economic policy, and the electoral college system favor a multiclass two-party system (Lipset 1977).

Weaknesses of class-oriented unions favored the American Federation of Labor (AFL), which successfully began to organize the skilled trades in 1891. The AFL concentrated on the skilled, because skilled workers dominated their trades and were capable of the sustained solidarity needed to win strikes. As an elite minority of the working class, the Federation focused largely on wage gains, better working conditions, and remained “neutral” in party politics. Yet, without legislative protection, union gains were periodically eroded by market downturns and antiunion employer drives. Thus, as a percentage of the labor force, the AFL experienced robust growth in the prosperous years before and during World War I. Membership declined rapidly in the depressions following the war; rose dramatically along with the formation of the Committee on Industrial Organizations just before and during World War II; declined slowly after the war for two decades; and then declined more rapidly in the economic recessions of the 1970s and 1980s, from which it has not recovered. That the decline was not more precipitous after World War II was due to the legislative protection

that labor received in the late 1930s under the Democratic administration’s New Deal.

### **POSSIBLE CONVERGENCE OF LABOR MOVEMENTS**

A combination of events has tilted the American labor movement toward the European social democratic model and the latter toward the American. During the Great Depression in the United States, the Democratic Party came to power with the backing of urban-industrial and middle-class voters. The party quickly enacted legislation to increase labor’s purchasing power, reduce price competition in industry, protect union organizing, and restore economic order. The National Labor Relations Act (1935) gave unions legal protection to organize, and labor conducted a militant drive for members. The recruitment of many semiskilled workers into new unions organized by the Committee on Industrial Organization (CIO) of the AFL threatened the dominance of the skilled trades in the Federation, leading to a withdrawal of the CIO from the Federation. Yet, eager to protect and extend recent gains, the divided labor movement began to abandon its traditional nonpartisan political stand. After World War II, both labor movements created electoral organizations to support Democratic candidates and mobilize their members to vote. The two labor movements merged in 1955, as did their electoral arms, to form the Committee on Political Education (COPE), formally independent of the Democratic party, but essentially functioning as part of it.

This labor-party rapprochement, committed to a social security program and a welfare state, led some scholars to conclude that the American labor movement was no longer exceptional, because it had helped create a welfare state similar to that forged by the social democratic parties in Europe (Greenstone 1977). The claim is strained because neither American labor nor the major parties ever embraced a socialist framework (Marks 1989, Chap. 6). More important, American labor has never been formally incorporated into the party and government, nor has government passed laws giving labor consultative rights in work plant operations, both central features of European corporatist labor movements.

In Europe, experiments by labor-dominated governments to nationalize industries have met with limited success (Panitch 1976), and some parties have abandoned or severely curtailed nationalization programs. The matter has been and remains an issue within parties, with some unions strongly opposing it (Currie 1979). To obtain power or maintain it by democratic means, the parties found it necessary to obtain the support of middle-class employees who want the social security guarantees of a welfare state, but not nationalization of industry. To maintain a vigorous economy, social democratic parties have enacted pro-business policies, not unlike those of the Democratic Party in the United States. For example, pro-business legislation by the Clinton administration or free trade with Latin America opposed by labor. In short, conservative forces both in and outside social democratic parties have pushed European political economies toward the American model.

Where labor is part of a coalition of a permanent ruling party as in the Mexico or the former Yugoslavia, labor has sometimes shown independence and initiated strikes despite government opposition (Bronstein 1995). And even in Japan, some enterprise unions have abandoned traditional consensus policies and challenged management and government with strikes and political turmoil (Okochi et al. 1974; Kuruvilla et al. 1990). Korea has moved even farther in this direction (Deyo 1997).

If present trends continue, perhaps a slow convergence of labor movement structures will take place. Sufficient research is not available to uncover all the causes of this plausible trend, but several play a role. In totalitarian or enterprise labor movements, when governments or managements create organizations that resemble labor movements (unions, elections, bargaining sessions, and consultation), appearances may become realities during leadership crises, especially in face of turbulent external events. Thus, in Poland, during authority crises of the ruling party and the state in the late 1980s, unions began to assert control over working conditions, wage determination, and political choice (Martin 1997). With the help of clergy, intellectuals, farmers, and others, unions defied central authority and instigated a movement to bring about a democratic party and state.

These authority crises in the party-dominant, totalitarian, and enterprise labor movements often result from changes in their external environments. Top labor officials become acquainted with the independent and corporate types of movements while participating in international agencies like the International Labor Office of the United Nations. The ability of independent and corporate movements to gain visible economic and political rewards for their members has not escaped the notice of labor leaders of other types of movements. Thus, when they confront authority crises, they have a vision of the kind of changes that would help them.

More important, rising global trade and economic interdependence fosters convergence. For example, in the 1970s, Japanese and Korean auto manufacturers began to enlarge their share of the automobile markets of the United States and western Europe increasing their unemployment. While management and unions both called for tariff protection, Japanese and Korean and labor demanded and received pay increases (Deyo 1989). On making a partial recovery, American and European corporations, in pursuit of higher profits, began to outsource production in countries with lower wage rates—for example, Mexico and Brazil. Then, with limited success, American labor leaders pressed government to place tariffs on auto imports and urged labor leaders in the exporting countries to demand higher wages. However, international labor cooperation remains puny compared to growth of world trade.

### **TRENDS IN LABOR MOVEMENTS**

In the advanced industrial capitalist democracies since the 1970s, the proportion of union members has declined in manufacturing and risen in the services and government. Some movements have shrunk rapidly, while others have remained relatively stable. Unions that had earlier won legal rights in the conduct of enterprise (codetermination, administering unemployment insurance, production planning, national bargaining) lost fewer members, for instance, corporate types of labor movements as found in Scandinavia and Germany rather than the independent type as in the United States.

In a study of eighteen advanced capitalist countries from 1970 to 1990, Western (1995) found

increasing dispersion in the rates of union decline, with the highest declines in countries with the lowest initial union density—such as the United States, Britain, and France. By 1990, the declines, wherever they occurred, were traced to unfavorable global economic conditions, decentralization of collective bargaining institutions, and electoral weakening of labor-oriented parties.

In the United States, the percentage of the labor force that was unionized shrank from 37 percent in 1945 to 14 percent in 1999. Union density in the private sector declined to one-tenth of the labor force, while public sector union density grew to over two-fifths. As unions were forced into a defensive position, strike rates declined precipitously. The causes of labor's decline are complex. The most common explanation, the shift from a manufacturing to a service economy, is inadequate. In a study of eleven capitalist democracies, Lipset (1986) found that union decline did not vary with manufacturing decline. Goldfield's study (1987) showed that after eliminating changes in the economy, industries, and occupations as possible causes, employer antiunion drives under Republican administrations tilted decisions of the National Labor Relations Board (NLRB) against organized labor. But Freeman and Medoff (1984) emphasized that white-collar unions grew rapidly despite employer resistance. Finally, some of the union decline is traced to labor's spending less money on organizing drives.

Undoubtedly, other factors were also involved in the decline, notably the outsourcing of manufacturing to other countries and the national conservative trend attending Republican electoral victories and consequent antilabor policies. Moreover, several new constituencies in the Democratic Party (blacks, women, educators, business) competed for party influence. Despite some Democratic electoral victories since 1992 and greater organizing efforts by the new AFL-CIO leadership, labor's slow downward trend has continued.

Labor movements trends in other parts of the world defy easy generalization. The best descriptions of their recent experiences appear in the *International Labour Review*. Clearly, free labor movements did not automatically emerge with the demise of totalitarian regimes in former soviet states. For example, in Poland, where the union-sponsored Solidarity Party gained governmental

control, unions almost ceased behaving like unions at the plant level, failing to bargain with management in support of the government's anti-inflation policy. In contrast, unions in the Czech Republic, by not participating in the Civic Forum which gained control of the government, bargained with government to win codetermination rights in industry and reduce unemployment (Ost 1997). In Romania, miners have episodically threatened violence against the government to win back-pay and wage increases. In other ex-soviet countries, especially Russia, the collapse of the economy virtually stalled the formation of a free labor movement.

In east Asia, union density has been declining recently not only in the industrializing tigers of Taiwan, Korea, Hong Kong, and Singapore but also in the more recent industrializing countries of Thailand, Indonesia, Malaysia, China, and Vietnam (Deyo 1997). Where unions and the working class have improved their positions, labor-market shortages, favorable government policies, and paternalism have been primarily responsible, rather than strong unions. Patriarchal and patrimonial regimes have excluded unions from decision making at the enterprise level or have weakened them when they appeared. Only in Korea, where industrial workers are rather homogeneous, are residentially concentrated, and have developed autonomous organization, have unions intermittently exhibited strong independence (Deyo 1989).

In Latin America, recent declines in autocratic regimes have encouraged the rise of autonomous labor movements. Yet, three factors that vary enormously by country are union autonomy from the state; the amount of collective bargaining; and dialogue among unions, the state, and management. Although some progress toward free-trade unionism has appeared in Argentina, Brazil, Mexico, and Chile, government pressure to achieve economic stability in the face of inflation, unemployment, poverty, and the legacy of autocratic military regimes has slowed basic changes in the tripartite relations of labor, state, and management (Bronstein 1995).

#### DEMOCRACY OR OLIGARCHY IN THE LABOR MOVEMENT

Even labor movements that are large, strong, and autonomous face constant problems. Unlike most

institutions, all labor movements claim to be democratic in ideology, structure, and behavior. Union officers are supposed to be leaders, not bosses, and, their primary task is to improve the well-being of membership. Sociologists have long pondered whether large and complex democratic organizations can escape becoming oligarchies with self-serving officers. The most famous proponent of this proposition was Roberto Michels (1911), who studied the history of the Socialist Party in pre-World War I Germany. He found that its officers had become a self-perpetuating elite who controlled communication with the membership, appointed their staff and successors, and pursued their self-interests, often oblivious of member needs and concerns.

Several case studies of unions have challenged Michels's thesis. Lipset, Trow, and Coleman (1956) found that the unique party system of the International Typographical Union fostered electoral competition, officer turnover, and membership involvement in union affairs. Edelman and Warner's (1979) study of fifty-one international unions revealed that officer turnover varied with constitutional provisions, such as frequency of elections, percentages of officers elected, and frequency of conventions. Cornfield (1989) found that substantial ethnic turnover among officers of the United Furniture Workers resulted from changes in the economy, the regional dispersal of the industry, political disputes among officers, ethnic tolerance of the membership, and a tradition of membership involvement in union affairs.

While most case studies of turnover among union leaders have focused on relatively small unions whose members exhibit rather homogeneous skills and earnings, the studies do not reveal the extent to which these conditions apply to the universe of unions. Marcus's study (1964) of all major unions in the country revealed that the larger and more heterogeneous the union, the slower the leadership turnover, the less frequently conventions were held, and the more decision making was concentrated in the officers.

Other dimensions of stratification within unions persist over their life histories. In a rare study, Bauman (1972) demonstrated that cleavages along skill lines persisted during the entire history of the British labor movement. At the beginning of the Industrial Revolution, skilled workers formed their

own societies. From 1850 to 1890, they formed craft unions that gained recognition. Between 1890 and 1924, they became an elite sector of the labor movement and maintained their status and influence in the large industrial unions. With the development of the Liberal and Labour parties, trade union leaders became subordinated to university-trained, middle-class intellectuals who dominated the parties, seats in Parliament, and high government positions. Herman Benson (1986) argued that this pattern also applies to the American labor movement, and Alain Touraine (1986) argued it is universal. In short, the rising size and organizational complexity of the labor movement are accompanied by increasing internal stratification and a less responsive bureaucracy, confirming Michels's original argument.

### CONCLUSIONS

Although the future of the American labor movement is difficult to predict, the movement will surely survive and change. Michels argued that as movements become institutions, they lose the loyalty and commitment of the founding generations. Although economic gains always remain paramount goals of unions, money is not uppermost in the minds of members while they are working. Whatever changes take place in the economy or polity, workers live where they work. The historic union goal to improve working conditions erodes with the bureaucratization of enterprises and the labor movement. Therefore, improving the quality of work life must become a renewed priority of labor leaders, so as to invigorate member loyalty and commitment. This involves greater worker participation in the control of work organization, a challenge that management will surely resist, but union leaders must relentlessly pursue.

When labor was a larger constituency of the American Democratic Party, it had more influence in party affairs despite COPE's weak electoral organization. Form (1995) has shown that especially at the grass-roots level, COPE is severely fragmented along occupational and industrial lines; more important, union members are hardly aware of labor's political goals and electoral efforts. Understandably, Democratic Party elites have responded more to other, better-organized constituencies. Unless labor makes common cause with some of them in the workplace as well as in the political

arena, labor's party influence will continue to decline. Paradoxically, the more selflessly labor supports other constituencies, the more politically influential it will become. Labor's natural allies are African Americans, Hispanics, educators, women's movements, and environmentalists. The challenge of labor leaders is to convince their members to support this strategy.

Finally, like business, labor must become a worldwide movement. Self-interest requires labor movements in advanced capitalist economies to assist foreign labor movements in making economic, social, and political gains. This surely is the toughest assignment, but unless progress is made on this front, labor may continue to decline.

(SEE ALSO: *Labor Force, Social Movements*)

## REFERENCES

- Bauman, Zygmunt 1972 *Between Class and Elites*. Manchester, England: Manchester University Press.
- Benson, Herman 1986 "The Fight for Union Democracy," in Seymour Martin Lipset, ed., *Unions in Transition*. San Francisco: Institute for Contemporary Studies.
- Bronstein, Arturo S. 1995 "Societal Change and Industrial Relations in Latin America: Trends and Prospects." *International Labour Review* 139:163-186.
- Calhoun, Craig 1982 *The Question of Class Struggle*. Chicago: University of Chicago Press.
- Cornfield, Daniel B. 1989 *Becoming a Mighty Voice: Conflict and Change in the United Furniture Workers of America*. New York: Russell Sage Foundation.
- Currie, Robert 1979 *Industrial Politics*. Oxford: Clarendon.
- Deyo, Frederic C. 1989 *Beneath the Miracle: Labor Subordination in the New Asian Industrialism*. Berkeley: University of California Press.
- 1997 "Labor and Post-Fordist Industrial Restructuring in East and Southeast Asia." *Work and Occupations* 24:97-118.
- Dougherty, Carroll R. 1941 *Labor Problems in American Industry*. Boston, Mass.: Houghton Mifflin.
- Edelstein, J. David, and Malcolm Warner 1979 *Comparative Union Democracy*. New Brunswick, N.J.: Transaction Press.
- Form, William 1985 *Divided We Stand: Working-Class Stratification in the United States*. Urbana: University of Illinois Press.
- 1995 *Segmented Work, Fractured Labor: Labor Politics in American Life*. New York: Plenum.
- Freeman, Richard B., and James I. Medoff 1984 *What Do Unions Do?* New York: Basic Books.
- Goldfield, Michael 1987 *The Decline of Organized Labor in the United States*. Chicago: University of Chicago Press.
- Greenstone, J. David 1977 *Labor in American Politics*. Chicago: University of Chicago Press.
- Jackson, Kenneth Robert 1984 *The Formation of Craft Markets*. New York: Columbia University Press.
- Jenkins, J. Craig 1985 *The Politics of Insurgency: The Farm Workers' Movement in the 1960s*. New York: Columbia University.
- Katznelson, Ira, and Aristride R. Zolberg (eds.) 1986 *Working-Class Formation: Nineteenth Century Patterns in Europe and the United States*. Princeton, N.J.: Princeton University Press.
- Knapp, Vincent J. 1976 *Europe in the Era of Social Transformation: 1700-Present*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kuruwilla, Sarush, Daniel G. Gallagher, Jack Fiorito, and Mitusuru Wakabayushi 1990 "Union Participation in Japan: Do Western Theories Apply?" *Industrial and Labor Relations Review* 43:366-373.
- Lane, David, and Felicity O'Dell 1978 *The Soviet Worker*. New York: St. Martin's Press.
- Lederer, Emil 1932 "Labor" *Encyclopedia of the Social Sciences*. New York: Macmillan.
- Lipset, Seymour Martin 1977 "Why No Socialism in the United States?" In Seweryn Bialer and Sophis Suzlar, eds., *Sources of Contemporary Radicalism*. Boulder, Colo.: Westview.
- 1986 "North American Labor Movements." In Seymour Martin Lipset, ed., *Unions in Transition: Entering the Second Century*. San Francisco: Institute of Contemporary Studies.
- , Martin A. Trow, and James A. Coleman 1956 *Union Democracy*. New York: Free Press.
- Marcus, Philip H. 1964 "Organizational Change: The Case of American Trade Unions." In George Zollschan and Walter Hirsch, eds., *Explorations in Social Change*. Boston: Houghton Mifflin.
- Marks, Gary 1989 *Unions in Politics: Britain, Germany, and the United States in the 19th and Early 20th Centuries*. Princeton, N.J.: Princeton University Press.
- Martin, Brian D. 1997 Institutional change and strike mobilization in pre- and post-transition Poland. Ph.D. Diss., Ohio State University, Columbus.
- Michels, Robert (1911) 1959 *Political Parties*. New York: Dover.

- Nutzinger, Hans G., and Jurgen Backhaus 1980 *Co-Determination*. Berlin: Springer-Verlag.
- Okochi, Kazuo, Bernard Karsh, and Solomon B. Levine (eds.) 1974 *Workers and Employers in Japan: The Japanese Employment Relations System*. Princeton, N.J.: Princeton University Press.
- Ost, David 1997 "Can Unions Survive Communism?" *Dissent* Winter: 21-27.
- Panitch, Leo 1976 *Social Democracy and Industrial Militancy*. New York: Cambridge University Press.
- Pirenne, Henri 1932 "Guilds, European." *Encyclopedia of the Social Sciences*. New York: Macmillan.
- Seidman, Jack, J. London, B. Karsh, and Daisy L. Tagliacozzo 1958 *The Worker Views His Union*. Chicago: University of Chicago Press.
- Sombart, Werner 1906 *Warum gibt es in den Vereingten Staaten keinen Sozialismus?* Tubingen: Mohr.
- Stern, Robert N., and Daniel Cornfield 1996 *The U.S. Labor Movement: References and Resources*. New York: G. K. Hall.
- Sturmthal, Adolf F. 1968 "Labor Unions: Labor Movements and Collective Bargaining in Europe." *International Encyclopedia of the Social Science*. New York: Macmillan and Free Press.
- 1974 "Trade Unionism." *Encyclopedia Britannica*. Chicago: William Benton.
- Touraine, Alain 1986 "Unionism as a Social Movement." In Seymour Martin Lipset, ed., *Unions in Transition*. San Francisco: Institute for Comparative Studies.
- Western, Bruce 1995 "A Comparative Study of Working-Class Disorganization: Union Decline in Eighteen Advanced Capitalist Countries." *American Sociological Review* 60:179-201.

WILLIAM FORM

## LABOR THEORY OF VALUE

See Marxist Sociology.

## LATIN AMERICAN STUDIES

Sociological research on Latin American societies has focused on the understanding of the causes and consequences of different patterns of development. In the past decades, this research has been guided by the counterpoint between two different theoretical approaches and the findings

generated with their help. In the 1950s and 1960s, the field was dominated by what came to be known as the modernization approach. In the 1970s and 1980s, dependency and world-system theories became prevalent (Klaren and Bossert 1986; Valenzuela and Valenzuela 1978). Each of these paradigms spawned useful lines of research, but eventually they became unsatisfactory, either because some of their assumptions were inconsistent with the facts, or because they were incapable of encompassing important areas of social reality. The field is now ripe for a new conceptual framework, which could incorporate useful aspects of the previous ones. In the past few years, there has been a shift toward a state-centered approach, but it is still unclear whether it will develop into a synthetic paradigm.

## THE REGION

The study of Latin American societies is complicated by the heterogeneous nature of the region. The nations that compose Latin America share some common traits, but they also have important differences.

Countries vary in terms of their economy and social structure, their ethnic composition, and their political institutions. The economic differentials are very substantial. The region includes Argentina and Uruguay, whose per capita gross national products (GNPs) are \$8,570 and \$6,020, respectively, higher than those of the Czech Republic and Hungary but lower than those of Portugal and Greece, the poorer countries in the European Union. At the other extreme, there are countries like Haiti and Honduras, with per capita gross national products (GNPs) of \$330 and \$700, respectively, which are comparable to those of the Central African Republic and the Congo, respectively (World Bank 1999). Social structures vary accordingly: About 90 percent of the Argentine and Uruguayan population is urban, but the percentages in Haiti and Guatemala are 35 percent and 44 percent (PNUD 1997); manufacturing accounts for about 20 to 25 percent of the GDP of Argentina, Brazil, and Mexico (about the same as in France, Germany, and Italy), but the proportion is only 3 percent in Bolivia and 9 percent in Panama, whose levels of industrialization are comparable to those of Cambodia and Bangladesh, respectively. Enrollment ratios in post-secondary



education vary from about 40 to 50 percent of the 20 to 24-year-olds in Argentina and Uruguay to about 8 to 9 percent in Honduras or Paraguay (World Bank 1992).

Latin American societies differ widely in their ethnic composition. The three basic components, Iberian settlers and other European immigrants, Indians, and blacks, are found in different proportions and mixes in different societies. Some are relatively homogeneous: most of the Argentine and Uruguayan populations are of European origin (the greater part are the product of transatlantic immigration at the turn of the century), and the population of Haiti is basically African. Other societies, like Mexico, Peru, and Guatemala, have maintained the colonial pattern of ethnic stratification, with mostly "white" elites ruling over largely Indian citizenries. Most of the population in Chile, Colombia, Venezuela, the Dominican Republic, El Salvador, and Honduras is the product of miscegenation. Brazil has a very heterogeneous population, with large contingents of all the ethnic groups and their different mixes (Ribeiro 1971; Lambert 1967).

With respect to their political institutions, countries in the region vary as well. Some, like Costa Rica, Chile, and Uruguay, have had long histories of constitutional rule (punctuated, by authoritarian episodes), while others, like Argentina, Brazil, and Peru, have wavered between instability and authoritarianism for much of the postwar period. All over the region, military dictatorships gave way to constitutional governments in the 1980s (O'Donnell and Schmitter 1986), and the 1990s has been the decade of democratic consolidation (Huntington 1991; Linz and Stepan 1996). The legitimacy and overall potential for institutionalization of these new governments vary according to the strength of liberal democratic traditions, the subordination of the state apparatus to the government, the vitality of the party system, and the dynamism of the economy. Authoritarian rule and state corporatism are still being dismantled in Mexico, and guerrillas operate in Colombia and Peru. Finally, Cuba is one of the few remaining state socialist polities in the world.

Nevertheless, there are economic, social, and political commonalities, besides the obvious cultural and religious ones. As far as the economy is concerned, and in spite of the variability noted

above, Latin American countries share four important traits. First, all these societies belong, in terms of their per capita product, to the low or lower-middle ranks in the world system. Second, all Latin American nations have basically been, and most still are, in spite of the considerable industrialization that took place in the most advanced countries, exporters of commodities and importers of manufacturing products. Third, industrialization, from the Depression of the 1930s up to the recent past, has been based on import-substitution policies. These policies have led to low growth rates, and even outright stagnation, once the domestic markets were saturated, a stage that the most advanced countries in the region reached in the 1970s and 1980s. Fourth, Latin American societies are highly dependent, in most instances because large segments of their economies are under the control of external actors (multinational corporations in particular), and in practically all cases because of their high levels of indebtedness to the advanced industrial countries.

There are many differences among Latin American social structures and political institutions, but all these societies except Haiti originated as Spanish and Portuguese colonies. The pillars of the original institutional matrix were the organization of the economy around the large agrarian property, a highly centralized political system in which representative features were weak, and a cultural system centered in the church, and in which the toleration of pluralism was extremely low (Lambert 1967; Veliz 1980). This institutional core disintegrated as a consequence of the economic and social changes of the nineteenth and twentieth centuries: New social actors were formed, and old ones were transformed by urbanization, industrialization, and the expansion of education, and also by the irruption of external economic, political, and cultural forces. Further, in some societies there were substantial changes in the composition of the population. However, the common origins still account for important similarities in the historical trajectories of the countries of the region.

In the realm of politics, commonalities are also evident in the post-independence period, especially after the Great Depression of the 1930s. At that time, liberal-democratic regimes, most of which had not integrated the lower classes into the political system, collapsed throughout the region.

In the following decades, most Latin American countries wavered between unstable democracy and nondemocratic forms of rule. South America has been especially prone to two of these: populist-corporatist regimes and bureaucratic military dictatorships (Malloy 1977; Pike and Stritch 1974; Stepan 1978; O'Donnell 1973, 1988). Many South American polities evolved cyclically in that period: populist-corporatist regimes (such as those headed by Juan Peron in Argentina and Getulio Vargas in Brazil) were frequent from World War II to the 1960s; military regimes and their coercive counterparts in society, guerrilla warfare and terrorism, predominated in the 1960s and 1970s (in Argentina, Brazil, Chile, Uruguay, and Peru); and democratization has swept the area in the 1980s.

### MODERNIZATION THEORY

Modernization perspectives shared two core assumptions that distinguished them from dependency and world-system theories. The first was that the central variables for understanding the development of a society are internal to the society; the second, that development is an evolutionary process, whose main characteristics are common to all societies. These commonalities underlie different approaches, which can be classified in terms of the variables they have considered central for the analysis of development, because of their primary causal weight, and in terms of the nature of the evolutionary process societies were supposed to undergo.

A first approach, sometimes of Parsonian or anthropological inspiration, privileged the value system, or culture, as the basic part of society, in the sense that values were expected to determine the basic traits of the economy or the polity. Change in values or culture appeared then as a key dimension in the process of transition from "traditionality" to "modernity." Examples of this perspective are Seymour Martin Lipset's analysis of entrepreneurship in Latin America (1967) and Howard Wiarda's argument about the influence of the corporatist Iberian tradition on the political evolution of Latin American societies (1973).

The second approach was structural, in the sense that it viewed the social or economic structures as the determining part of society, and that it saw changes in these structures as the central

aspect of the process of development. The most elaborate versions of this perspective were inspired in Durkheim and in the most structural version of Parsonian theory: They regarded the degree of integration of society as the key characteristic of the social order and focused on the processes of differentiation and integration as the keys for the understanding of development, still conceptualized in terms of the "traditionality-modernity" continuum. Gino Germani's model of Latin American modernization in terms of the counterpoint between mobilization and integration is a good example of this perspective (1962, 1981; see also Kahl 1976). The typology of stages of economic development made popular by the influential Economic Commission for Latin America (ECLA) of the United Nations (Rodriguez 1980), shared these presuppositions.

These approaches spawned important types of sociological research in the 1950s and 1960s, but reality showed that the assumptions behind the evolutionary paradigm were problematic. Toward the end of the 1960s, it became clear that the predictions derived from the model of the traditionality-modernity continuum in relation to the social structure and the political system were inconsistent with the facts. In the social structure, the expansion of capitalist agriculture and the development of industry in societies with large peasantries did not lead to the dissolution of "traditional," or precapitalist, rural social relations. Its survival (Stavenhagen 1970; Cotler 1970) and the emergence of what came to be known as "dualistic," or structurally heterogeneous, societies meant that the patterns of social development of Latin America were not replicating those of the western European countries that were considered as the universally valid models of modernization.

Political processes in the 1960s and early 1970s were also paradoxical in relation to this theory: Lipset's thesis that liberal democracy was a correlate of development (as measured by variables such as income, urbanization, industrialization, and education) ([1960] 1981, pp. 27-63) seemed to be contradicted by the fact that it was precisely the most "modern" countries (Argentina, Uruguay, and Chile) that were the most prone to instability and authoritarianism. On the basis of this propensity, Guillermo O'Donnell (1973) argued that, in the peculiar situation of Latin American industrialization, countries of this type were

the most likely candidates for bureaucratic authoritarianism. This proposition was itself questioned from different perspectives (see Collier 1979), but the lack of correlation between development and democracy seemed nevertheless obvious at that time. Later, the wave of democratization that swept Latin America in the 1980s and the revolutions of 1989–1990 in eastern Europe would indicate that the relationship proposed by Lipset does exist, even though it does not seem to be linear.

### DEPENDENCY AND WORLD-SYSTEM THEORIES

Approaches based on dependency and world-system theories were the contemporary elaboration of themes found in Marxism, the theory of imperialism in particular (Lenin 1968; Luxemburg 1968). They were based on two propositions that sharply contradicted the assumptions of modernization theory. The first was that internal characteristics of a society, such as its level of development or the nature of its political system, are determined by the society's position in the international system. The second assumption was a corollary of the first: Since the main causes of development are external, there is no reason to expect that all societies will follow the same evolutionary trajectory. On the contrary: The world system is organized into a more dynamic, developed, and powerful core, and a more passive, backward, and dependent periphery. Beyond these commonalities, there were differences between the dependency and world-system approaches.

Dependency theory, which in itself was a Latin American product, emphasized the radical distinction between core and periphery, and it took dependent societies as its unit of analysis. Its focus was on the mechanisms it claimed were responsible for the preservation of underdevelopment, the transfer of economic surplus to the center via trade and investment in particular (Cardoso and Faletto 1979; Frank 1969, 1972). Wallerstein's world-system approach (1974), on the other hand, singled out the world economy as the unit of analysis, and it explored the functions of different types of society (core, semiperiphery, and periphery) in that system.

These theories have generated, since the late 1960s, important pieces of research. Fernando H. Cardoso and Enzo Faletto's comparative analysis

of Latin American societies (1979) showed how different positions in the world economy generated different class structures and different patterns of development (see also Kahl 1976). Peter Evans (1979), in a study of Brazil, showed how the peculiarities of the country's development were shaped by the interaction among the state, foreign capital, and the domestic bourgeoisie. Gary Gereffi (1983) explored the consequences of dependency in a key industry. From a world-system perspective, Alejandro Portes and his colleagues (1981, 1985, 1989) conducted important studies of the labor flows in the world economy, and of the articulation between the formal and informal economies.

These approaches improved our understanding of Latin American development in two ways. First, by focusing on the effects of external factors and processes (those in the economy in particular, but the political and cultural ones as well), dependency and world-system perspectives corrected the assumption characteristic of modernization theory that societies could be studied in relative isolation. Second, since these theories constructed models of the world economy that were either pericentric (Doyle 1986), as in dependency theory; or that emphasized the analysis of peripheral and semiperipheral social structures alongside that of the core, as in world-system theory, they balanced the emphasis on core structures and processes characteristic of most preexisting studies of imperialism, which tended to ignore the institutions and dynamism of the periphery.

However, the basic tenets of dependency and world-system theories have also led in some cases (but not in the best research spawned by these perspectives) to simplistic assumptions, in particular to a tendency to consider internal structures and processes, especially the political and cultural ones, as transmission belts for external economic and political forces. The pendulum swung to the other extreme: from the neglect of external economic and political determinants of social change, characteristic of modernization theory, to a disregard for internal factors.

### TOWARD A NEW PARADIGM

In the late 1980s the basic assumptions of dependency and world-system theories also became problematic. In the first place, research on the role of

the state in the development process led to its reconceptualization as an autonomous actor (see Evans et al. 1985; Evans 1995; and Waisman 1987 for a Latin American case). This challenged the conceptions characteristic of most approaches close to the Marxist tradition—dependency and world-system theories included—which tended to see the state, claims of “relative autonomy” notwithstanding, as basically an instrument of either domestic ruling classes or foreign forces. Second, there was a rediscovery of social movements in the region (a consequence that the redemocratization processes of the 1980s could not fail to produce). More often than in the past, agency came to be seen as relatively independent in relation to its structural and institutional context (see the essays in Eckstein [1989], for instance), a conceptualization that differs sharply from the structuralist assumption inherent in dependency and world-system perspectives. Finally, the proposition according to which the central determinants of the development of Latin American countries are external has also been challenged (Waisman 1987; Zeitlin 1984). In the 1990s there is a growing consensus, as there was in the 1970s, that existing paradigms do not encompass our current understanding of Latin American development.

A new theoretical synthesis should integrate the valid components of the modernization, dependency, and world-system approaches, and should also allow for the autonomy of the state and the role of agency. A point of departure should be the recognition of three facts. First, the division of the world into core, periphery, and semiperiphery is not sufficient to encompass the diversity of developmental situations and trajectories relevant for the study of the countries of Latin America. It is necessary to construct a typology of more specific kinds of peripheral societies (and of core societies as well: the core-periphery distinction is still too abstract), and to describe more systematically the structural “tracks” that have crystallized in the region at different stages of development of the world economy.

Second, the different developmental paths followed by Latin American societies have been determined by empirically variable constellations of external and internal processes, and of economic, political, and ideological-cognitive ones. Few propositions that privilege the causal role of specific factors are likely to be generally valid. Rather

than seeking propositions of this type, as modernization, dependency, and world-system approaches have tried to do, it would be more productive to map the specific bundles of factors that have influenced developmental outcomes in critical situations. Third, development is a discontinuous process (an instance of punctuated equilibrium, in Stephen D. Krasner’s [1984] words), but we lack a theory of the transition points, that is, of the crossroads where acceleration, stagnation, retrogression, and changes of developmental tracks have taken place. The Latin American experience indicates that these are the points in which major changes in the world system, such as depressions, wars, important technological developments, organizational changes in production or trade, and restructuring of the economic or military balance of power, have interacted with domestic processes economic, political, and cultural processes.

#### **LATIN AMERICA IN THE PAST DECADE: THE DOUBLE TRANSFORMATION**

In the past two decades, Latin America has been undergoing two simultaneous transformations: economic liberalization and democratization (Waisman 1998). The pervasive nature of these transitions is indicated by the fact that they have been independent of level of development and of political regime. Most countries in the region, from relatively underdeveloped Bolivia and Paraguay to relatively industrialized Brazil and Argentina, have moved from military rule to liberal democracy in the past decade. As for economic liberalization, it took place in authoritarian Argentina and also in democratic Argentina, in authoritarian Chile and also in democratic Chile, in state-corporatist Mexico, in social-democratic Venezuela and also in social-Christian Venezuela. Like the depressions of the 1870s or the 1930s, this is a critical juncture, a reshaping of domestic economic and political institutions, the regional manifestation of the processes of economic globalization and reconstitution of the world order following the collapse of communism.

Economic liberalization was the result of a constellation of internal and external determinants. In most cases, the former seem to be central, but in any case exogenous factors constituted a powerful set of incentives and constraints pushing countries toward the liberalization road. These

external causes can be classified into economic, political, and cultural.

The debt crisis of the 1980s was, for many Latin American countries, the central economic constraint (Felix 1990; Kahler 1986; Nelson 1990; Stallings and Kaufman 1989; Wesson 1988; Wionczek 1987). This crisis forced highly indebted countries to increase and diversify exports, which presupposed a higher rate of investment, and to reduce government spending. These objectives could have been sought with variable mixes of industrial policy ("picking the winners") and trade liberalization, but the policies of creditors and international lending agencies produced strong incentives for privatization, deregulation, and the opening up of the economy. Saying that the debt crisis has been in the 1980s both a powerful constraint and a substantial incentive for economic liberalization does not mean, of course, that the relationship between crisis and reform was necessary or sufficient. In fact, liberalization had begun in the 1960s and 1970s, as a consequence of perceived export opportunities, of export-oriented foreign investment, or of deliberate attempts by governments to overcome stagnating tendencies in the economy. The difference is that what had been a developmental option in previous decades had now become an imperious necessity.

As for political factors, the end of the Cold War facilitated not only political liberalization, which is more or less obvious, but also large-scale economic reform, especially in democratic settings. First, the collapse of communism led to the organizational and ideological disarmament of the radical left, and this rendered armed subversion less likely. Second, and more important, in the new international conditions the fear, be it realistic or paranoid, of exogenously induced revolution either weakened or disappeared. This fear, often unrealistic, was not only a central determinant of the exclusionary policies carried out by local elites and a reason for American and European support for authoritarian regimes, but it was also a constraint on elite behavior. During the Cold War, important segments of the state and political elites in many countries espoused the view that populist and protectionist economic policies were an antidote to communism, and many believed that the drastic social consequences of

the dismantling of autarkic capitalism would be a breeding ground for revolution (Waisman 1987).

Finally, the international demonstration effects received by the segments of Latin American societies most open to outside influences (economic and political elites, middle classes, and the intelligentsia) have favored, since the early 1980s, both economic and political liberalization. The economic success of capitalist economies and the crisis in the socialist ones, as well as the eastern European revolutions, have conferred an aura of empirical validity on two propositions: first, that capitalism and liberal democracy are both efficacious, while socialism as both an economic and a political formula is not; and second, that a capitalist economy and a democratic polity presuppose each other.

These external factors interacted with endogenous determinants. The most important of these was the realization, by state and political elites, but also by economic and intellectual ones, that autarkic capitalism had run its course, and caused in many countries stagnation or retrogression in the economy and instability and illegitimacy in the polity. It was simply no longer viable as an institutional model.

Autarkic capitalism was carried to its extreme in the countries of the Southern Cone: Argentina, Chile, and Uruguay. The long-term effects of radical import-substituting industrialization could only be stagnation, or at least stagnating tendencies, because industrialization was based on the transfer of resources from the internationally competitive export sector to a manufacturing sector shielded from the world by extraordinarily high levels of protection, both tariff and nontariff (outright prohibitions to import included). Advanced autarkic industrialization led to a situation in which most of the capital and labor in the society was committed to this captive domestic market. The outcome of this massive misallocation of resources could only be, in the long run, stagnation, and even retrogression.

Instability and illegitimacy were the long-term outcome of autarkic industrialization because this pattern of growth eventually produced an explosive combination: a stagnated economy and a society made up of highly mobilized and organized social forces, labor and the intelligentsia in particular. Stagnation led to the mobilization of large

segments of the lower classes and the intelligentsia. In some countries, such as those of the Southern Cone in the 1960s and 1970s, truly revolutionary situations, in the Leninist sense of the term, appeared. Eventually, they triggered, or justified, the establishment of authoritarian regimes.

The demise of autarkic capitalism was the result, as was the case with communism, of a revolution from above. In both situations, it was the elites who dismantled the existing regimes, after having reached the conclusion that the stagnation and illegitimacy experienced by their societies were the consequence of existing institutions, and that there was no solution to the economic and political crises within these institutions. In some cases, such as Argentina and Chile, this realization took place when radical import substitution had approached its limits, and produced not only stagnation but also a revolutionary regime in Chile, and guerrillas and food riots in Argentina. In other cases, such as Brazil, the switch is taking place before these consequences appear, largely as a consequence of the demonstration effects of other Latin American countries.

#### **THE CONTRADICTION LOGICS OF ECONOMIC AND POLITICAL TRANSFORMATION**

There are three possible sequences between economic and political liberalization (Bresser Pereira et al. 1993; Haggard and Kaufman 1995; Lijphart and Waisman 1997). To give them Latin American names, they would be the Chilean model, in which privatization, deregulation, and the opening up of the economy preceded democratization; the Venezuelan pattern, in which the liberalization of the economy takes place in the context of political institutions that have been in place for several decades; and finally the Argentine sequence, in which economic liberalization and the consolidation of democracy take place more or less at the same time. The three are fraught with dangers, but the third is the least favorable. The reason is that economic liberalization and the consolidation of democracy are governed by opposing social logics (Lijphart and Waisman 1997; Waisman 1998).

Privatization, deregulation, and the opening up of the economy are governed by the logic of differentiation. Their first impact on the society is

the increase of social differentiation in both the vertical and the horizontal senses: Polarization between rich and poor as a whole intensifies, and so does polarization between “winners” and “losers” within each social class, and also between sectors of the economy and regions of the country.

On the other hand, the consolidation of democracy is governed by the logic of mobilization. The political context of a new democracy renders the mobilization of those affected by economic liberalization more likely, due to the lowering of the costs of political action in relation to what was the case in the predemocratic period. Moreover, new democracies create incentives for political entrepreneurship, for in the new political conditions political and labor activists must secure and organize social and political bases. Economic liberalization presents them with an inventory of grievances that easily translates into a political agenda. Thus, both from below and from above, institutional factors are conducive to the articulation of movements of resistance to economic liberalization.

These two logics have the potential for inhibiting each other, and consequently for blocking or derailing economic liberalization, or the consolidation of democracy, or both. And yet, economic liberalization and the consolidation of democracy are occurring simultaneously in many Latin American countries, and so far no serious blockage or derailing has occurred as a consequence of the interaction between these processes. In several countries, there is an impressive consensus, involving state and economic elites, and even labor and some of the parties in the left and center that had supported state socialism or economic nationalism in the past.

The reason that, so far, the two opposing logics discussed above have not clashed lies in the operation of three “cushions” or moderating factors, which inhibit political mobilization. These can be classified into structural, institutional, and cognitive-ideological.

The structural cushion consists simply in the fact that economic liberalization itself weakens and destroys the power bases of the coalitions supporting the old regime (rent-seeking entrepreneurs and unions, and in some cases managers and workers in the public sector) and in general of those hurt by trade liberalization, privatization,

and deregulation. On the one hand, these processes of economic transformation generate insecurity and economic deprivation, and thus discontent. On the other, impoverishment and marginalization also reduce social groups' capacity for organization and mobilization in defense of their interests.

Business elites, both private and public, face the choice between recycling into competitive capitalists and dropping out from the elites. This group is usually differentiated, and its most entrepreneurial elements, and/or those endowed with large amounts of economic and social capital, usually join, or become, the elite in the open economy. Labor is also dualized: Workers in the competitive sectors of the economy, or in sectors insulated from foreign competition (e.g., those producing nontradeable goods) are likely to join the train of economic conversion, and to engage in bargaining-oriented unionism. Sectors weakened by deindustrialization, by the increase of unemployment, and in general by the growth of poverty, on the other hand, are more likely to develop confrontational forms of political action, but in the end their impact is likely to be modest, because of the factors discussed above. In many countries, the weakening of labor actually began before economic liberalization, for impoverishment and the growth of the labor reserve were manifestations of the crisis of the old regime itself.

In the second place, the clash between the two opposing logics discussed above has been inhibited by institutional change. The design or redesign of some of the basic economic and political institutions is in many cases a necessary component and in general an important contributing factor to successful transitions.

Finally, cognitive-ideological factors have both external and internal sources. As I argued above, economic nationalists as well as leftists have been affected by the collapse of communism (and also by the apparent success of the Thatcher-Reagan economic policies), and by a process of political learning, triggered by the experience of the economic and political consequences of the old regime. The cumulative effect of these cognitive processes has been not only the abandonment of autarkic capitalism, but also the acceptance of the liberal model, either under the form of active support, or because of ideological surrender produced by the exhaustion of alternatives.

Given this quasi-consensus, the negative consequences of the economic transformation are not conceptualized by large segments of economic, political, and intellectual elites as necessary or permanent attributes of the emerging institutional model, but to a large extent as the product of the failure of autarkic capitalism. And the new course is not perceived as a leap in the dark, but as a transition, whose undesirable side effects are in any case the inevitable and temporary cost of the only process that can ensure, in the long run, affluence and even a higher measure of equality of opportunity in the society. The most significant leftist and radical arguments, on the other hand, focus on the espousal of social democracy in opposition to the most radical forms of the market economy, often characterized as "Darwinian" or "savage" capitalism.

Both the strengthening of civil society and the enhancement of state capacity are necessary conditions not only for the consolidation of democracy but also for the success of the economic transformation.

The typical situation in the countries that are undergoing the double transformation is that the state is shrinking while its effectiveness remains questionable in many areas. At the same time, the corporatist apparatus that was characteristic of autarkic capitalism is being dismantled. Civil society has been weakened by the authoritarian experience, and by the increased social differentiation that follows economic liberalization. The consensus supporting the new course could erode unless the new model is institutionalized, something that cannot happen without a more effective state and a stronger civil society.

This is, thus, an inflection point in the relationship between state and society, a conjuncture that could be conceptualized both a danger and an opportunity. The danger resides in the fact that, in these conditions, the structural, institutional, and cognitive "cushions" discussed above could gradually lose their effectiveness. The social dislocation produced by economic liberalization would either increase the atomization of the society, and lead to mass anomie; or facilitate "praetorian" mobilization. The first outcome would entail a deterioration of the already modest quality of the new democracies, while the second would lead to disturbances. In both cases, the legitimacy of the

policies of economic transformation would fall precipitously. But this situation also represents and opportunity for social and political agency, whose goal would be the strengthening of civil society and the enhancement of state capacity.

## REFERENCES

- Bresser Pereira, Luis C., Jose M. Maravall, and Adam Przeworski 1993 *Economic Reforms in New Democracies: A Social-Democratic Approach*. New York: Cambridge University Press.
- Cardoso, Fernando H., and Enzo Faletto 1979 *Dependency and Development in Latin America*. Berkeley: University of California Press.
- Collier, David (ed.) 1979 *The New Authoritarianism in Latin America*. Princeton, N.J.: Princeton University Press.
- Costin, Harry, and Hector Vanolli (eds.) 1998 *Economic Reform in Latin America*. Fort Worth, Tex.: Hartcourt Brace.
- Cotler, Julio 1970 "The Mechanisms of Internal Domination and Social Change in Peru." In Irving L. Horowitz, ed., *Masses in Latin America*. New York: Oxford University Press.
- Doyle, Michael W. 1986 *Empires*. Ithaca, N.Y.: Cornell University Press.
- Eckstein, Susan 1989 *Power and Popular Protest: The Latin American Social Movements*. Berkeley: University of California Press.
- Evans, Peter B. 1979 *Dependent Development: The Alliance of Multinational, State, and Local Capital in Brazil*. Princeton, N.J.: Princeton University Press.
- 1995 *Embedded Autonomy: States and Industrial Transformation*. Princeton, N.J.: Princeton University Press.
- , Dietrich Rueschemeyer, and Theda Skocpol (eds.) 1985 *Bringing the State Back In*. Cambridge, England: Cambridge University Press.
- Felix, David (ed.) 1990. *Debt and Transfiguration: Prospects for Latin America's Economic Revival*. Armonk, N.Y.: M. E. Sharpe.
- Frank, Andre G. 1969 *Capitalism and Underdevelopment in Latin America*. New York: Monthly Review.
- 1972 *Lumpenbourgeoisie—Lumpenddevelopment*. New York: Monthly Review.
- Gereffi, Gary 1983 *The Pharmaceutical Industry and Dependency in the Third World*. Princeton, N.J.: Princeton University Press.
- Germani, Gino 1962 *Politica y sociedad en una epoca de transicion (Politics and Society in an Epoch of Transition)*. Buenos Aires: Paidos.
- 1981 *The Sociology of Modernization: Studies of Its Historical and Theoretical Aspects with Special Regard to the Latin American Case*. New Brunswick, N.J.: Transaction.
- Haggard, Stephan, and Robert R. Kaufman 1995 *The Political Economy of Democratic Transitions*. Princeton, N.J.: Princeton University Press.
- Horowitz, Irving L. (ed.) 1970 *Masses in Latin America*. New York: Oxford University Press.
- Huntington, Samuel P. 1991 *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.
- Kahl, J. A. 1976 *Modernization, Exploitation, and Dependency in Latin America*. New Brunswick, N.J.: Transaction.
- Kahler, Miles (ed.) 1986 *The Politics of International Debt*. Ithaca, N.Y.: Cornell University Press.
- Klaren, Peter F., and Thomas J. Bossert (eds.) 1986 *Promise of Development*. Boulder, Col.: Westview.
- Krasner, Stephen D. 1984 "Approaches to the State: Alternative Conceptions and Historical Dynamics." *Comparative Politics* 16(2):223–246.
- Lambert, Jacques 1967 *Latin America: Social Structures and Political Institutions*. Berkeley: University of California Press.
- Lenin, V. I. 1968 *Imperialism: The Highest Stage of Capitalism*. Moscow: Progress.
- Linz, Juan J., and Alfred Stepan 1996 *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore, Md.: Johns Hopkins University Press.
- Lijphart, Arend, and Carlos H. Waisman (eds.) 1997 *Institutional Design in New Democracies*. Boulder, Colo.: Westview.
- Lipset, Seymour M. 1967 "Values, Education, and Entrepreneurship." In Seymour M. Lipset and Aldo Solari, eds., *Elites in Latin America*. New York: Oxford University Press.
- (1960) 1981 *Political Man: The Social Bases of Politics*. Baltimore, Md.: Johns Hopkins University Press.
- Luxemburg, Rosa 1968 *The Accumulation of Capital*. New York: Monthly Review.
- Malloy, James M. (ed.) 1977 *Authoritarianism and Corporatism in Latin America*. Pittsburgh: University of Pittsburgh Press.



- Nelson, Joan (ed.) 1990 *Economic Crisis and Policy Choice: The Politics of Adjustment in the Third World*. Princeton, N.J.: Princeton University Press.
- O'Donnell, Guillermo A. 1973 *Modernization and Bureaucratic-Authoritarianism: Studies in South American Politics*. Berkeley, Calif.: Institute of International Studies.
- 1988 *Bureaucratic Authoritarianism: Argentina, 1966-73 in Comparative Perspective*. Berkeley: University of California Press.
- , and Phillippe C. Schmitter 1986 *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore, Md.: Johns Hopkins University Press.
- Pike, Frederick B., and Thomas Stritch (eds.) 1974 *The New Corporatism*. Notre Dame: University of Notre Dame Press.
- PNUD (Programa de las Naciones Unidas para el Desarrollo) 1997 *Informe sobre el desarrollo humano (Human Development Report)* Madrid: Mundi-Prensa.
- Portes, Alejandro, and John Walton 1981 *Labor, Class, and the International System*. New York: Academic Press.
- Portes, Alejandro, and Robert L. Bach 1985 *Latin Journey: Cuban and Mexican Immigrants in the United States*. Berkeley: University of California Press.
- Portes, Alejandro, Manuel Castells, and Lauren A. Benton (eds.) 1989 *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore, Md.: Johns Hopkins University Press.
- Ribeiro, Darcy 1971 *The Americas and Civilization*. New York: E. P. Dutton.
- Rodriguez, Octavio 1980 *La teoria del subdesarrollo de la CEPAL (The ECLA Development Theory)* Mexico City: Siglo XXI.
- Stallings, Barbara, and Robert R. Kaufman (eds.) 1989 *Debt and Democracy in Latin America*. Boulder, Colo.: Westview.
- Stepan, Alfred 1978 *The State and Society: Peru in a Comparative Perspective*. Princeton, N.J.: Princeton University Press.
- Stavenhagen, Rodolfo 1970 "Classes, Colonialism, and Acculturation." In Irving L. Horowitz, ed., *Masses in Latin America*. New York: Oxford University Press.
- Valenzuela, J. Samuel, and Arturo Valenzuela 1978 "Modernization and Dependency: Alternative Approaches in the Study of Latin American Underdevelopment." *Comparative Politics* 10(2):535-557.
- Veliz, Claudio 1980 *The Centralist Tradition in Latin America*. Princeton, N.J.: Princeton University Press.
- Waisman, Carlos H. 1987 *Reversal of Development in Argentina*. Princeton, N.J.: Princeton University Press.
- 1998 "The Political Dynamics of Economic Reform in Latin America." In Harry Costin and Hector Vanolli, eds., *Economic Reform in Latin America*. Fort Worth, Tex.: Harcourt Brace.
- Wallerstein, Immanuel 1974 *The Modern World System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic.
- Wesson, Robert G. (ed.) 1988 *Coping with the Latin American Debt*. New York: Praeger.
- Wiarda, Howard J. (ed.) 1973 "Towards a Framework for the Study of Political Change in the Iberic-Latin Tradition: The Corporative Model." *World Politics* 25:206-235.
- Wionczek, Miguel S. 1987 *La crisis de la deuda externa en America Latina (The Latin American External Debt Crisis)*. Mexico: Fondo de Cultura Economica.
- World Bank 1992 *World Development Report 1998/99*. New York: Oxford University Press.
- 1999 *World Development Report 1998/99*. New York: Oxford University Press.
- Zeitlin, Maurice 1984 *The Civil Wars in Chile*. Princeton, N.J.: Princeton University Press.

CARLOS H. WAISMAN

## LAW AND LEGAL SYSTEMS

The term *law* is surprisingly difficult to define. Perhaps the best-known definition within the sociology of law community is that of Max Weber: "An order will be called *law* if it is externally guaranteed by the probability that coercion (physical or psychological), to bring about conformity or avenge violation, will be applied by a *staff* of people holding themselves specially ready for that purpose" (1954, p. 5). Similar definitions include Donald Black's terse statement: "Law is governmental social control" (1976, p. 2). While these types of definitions have sometimes been attacked as employing a Westernized conception, appropriate for developed states but inappropriate for other societies, Hoebel advances a similar definition of law in all societies: "The really fundamental *sine qua non* of law in any society—primitive or civilized—is the legitimate use of physical coercion by a socially authorized agent" (1954, p. 26).

Definitions such as these are more interesting for what they exclude than for what they include. Weber and Hoebel each attempt to draw a line

where the boundary between law and something else is fuzziest. By including the term *legitimate*, Hoebel's definition is intended to distinguish law from the brute exercise of force. The leader of a criminal gang who forces people to give him money may be doing many things, but he is not enforcing the law. He is not a socially authorized agent, and his use of force is not legitimate. Legitimacy itself is a slippery concept, and disagreements about when it is present give rise to questions such as whether the Nazis governed under the rule of law.

The inclusion of coercion and specialized agents of enforcement in both Weber's and Hoebel's definitions is meant to distinguish law from customs or norms, the breach of which either is not sanctioned or is sanctioned only by members of the group against which the breach occurred. The internal rules (norms and customs) governing a family's life or an organization's life are not law unless they are reinstitutionalized, that is, unless they are "restated in such a way that they can be applied by an institution designed (or at very least, utilized) specifically for that purpose" (Bohannan 1965, p. 36). Some have rejected such definitions and argued that law consists of the regularized conduct and patterns of behavior in a community or society (Ehrlich 1936; Malinowski 1926). They fear that if the study of law is restricted to legal rules enforced by specialized legal staffs, it will exclude much of what legal anthropologists and legal sociologists may find interesting. Regularized conduct definitions in turn suffer from their inability to offer a clear boundary between legal norms and other norms in the society. However, they do capture the idea that there is law-stuff everywhere. Families and organizations do generate rules and do coerce or induce compliance. These groups constitute what Moore (1973) calls "semi-autonomous fields." Not only are the rules of these organizations interesting in their own right, the interaction of these rules and the state rules we call law helps to shape the fundamental choice between avoidance and compliance that is faced by all to whom rules are addressed.

The boundary maintenance functions of definitions such as those of Weber, Hoebel, and Bohannan undoubtedly have their place, but law's empire is so large that the border skirmishes occurring out on its frontiers have limited influence on our shared understanding of the words *law* and

*legal system*. Wherever it occurs, law is a body of rules that speak to how people should behave in society (substantive law) and how the legal system itself should proceed (adjective law). The volume and complexity of rules may be expected to parallel the size and complexity of the society of which they are a part. But broad categories of substantive law—tort law, property law, criminal law—apparently exist in all legal orders, as do the fundamentals of adjective law—procedure and evidence. The various definitions of "law" exist in an uneasy tension (Tamanaha 1997). This tension can serve us well if we follow Griffiths's advice (1984) and view "legalness" as a variable rather than thinking of "law" as a special, definable phenomenon. The complex body of substantive and adjective rules at different levels comprise a legal system.

## LEGAL SYSTEMS

The comparative study of law might trace its roots to Aristotle's comparison of Greek city-state constitutions. A more recent example is Montesquieu, who, in *The Spirit of the Laws* ([1748] 1962), attempted to explain legal diversity in terms of various factors in the social setting. Interspersed between these efforts were comparisons of canon law with Roman law in Europe and with the common law in England. Despite these precursors, the modern study of comparative legal systems has become a topic of sustained academic interest only during the last 100 to 150 years.

The history of comparative law is set forth in a number of works, including Zweigert and Kotz (1987) and David and Brierley (1985). The present essay discusses a small part of this history, focusing on what Zweigert and Kotz call scientific or theoretical comparative law rather than legislative comparative law, in which foreign laws are examined and invoked in the process of drafting new nation-state laws.

Early theoretical efforts, exemplified by Maine's *Ancient Law* ([1861] 1963), adopted evolutionary theories of legal development. In Maine's famous formulation, legal systems, following changes in social arrangements, move from *status*, wherein one's rights and duties are determined by one's social niche (the law of feudalism), to *contract*, wherein one's rights and duties are determined by

oneself and the contracts one enters into (eventually the law of capitalism).

A second well-known developmental theory of changes in legal systems is that of Durkheim ([1893] 1964). A societal movement from mechanical to organic solidarity is accompanied by a movement from repressive law (law that punishes those who violate a shared moral understanding) to restitutive law (law that attempts to facilitate cooperation and to return people to a *status quo ante* when rule violations occur).

From the sociological point of view, perhaps the most important contributor to the early development of comparative law was that preeminent lawyer-social scientist, Max Weber. Weber's contribution was in three parts. First, he developed the device of an ideal type, a stylized construct that represents the perfect example of a phenomenon. The ideal type acts as a yardstick against which we might measure actual legal systems. Second, using ideal types, he provided a typology of legal systems classified by the formality and the rationality of their decision-making processes. Ideally, legal systems could be thought of as formal or substantive, rational or irrational. A legal system is formal to the extent that the norms it applies are intrinsic to the system itself. Substantive law, as the term was used earlier, should not be confused with the substantive dimension of Weber's typology. A legal system is substantive in Weber's sense to the extent that the source of the norms it applies is extrinsic to the legal system. For example, a legal system would be substantive if a court resolved disputes by reference to a religious rather than a legal code.

A legal system is rational if it yields results that are predictable from the facts of cases; that is, if case outcomes are determined by the reasoned analysis of action in light of a given set of norms. A legal system is irrational when outcomes are not predictable in this way. Basically, a legal system is rational to the extent that similar cases are decided similarly.

A formally irrational system exists when the legal order produces results unconstrained by reason. Classic examples are judgments following consultation with an oracle or trial by ordeal. Substantive irrationality exists when lawmakers and finders do not resort to some dominant general norms but, instead, act arbitrarily or decide

upon the basis of an emotional evaluation of a particular case. Weber apparently had in mind the justice dispensed by the Khadi, a Moslem judge who, at least as Weber saw him, sat in the marketplace and rendered judgment by making a free and idiosyncratic evaluation of the particular merits of each case.

A substantively rational legal system exists when lawmakers and finders follow a consistent set of principles derived from some source other than the legal system itself. Again, Weber thought that Moslem law tended toward this type insofar as it tried to implement the thoughts and commands of the Prophet.

Western legal systems, especially those of civil law countries such as France and Germany, most nearly approximate the formally rational ideal, a legal system where the generality of legal rules is high and where the legal rules are highly differentiated from other social norms.

The relationship between formal and substantive law is obviously more complex than can be reflected in these four Weberian types. For example, legal systems may be procedurally quite formal while incorporating substantive norms rooted in nonlegal institutions. Moreover, rational systems may incorporate potentially irrational components, as when the final judgment in a case is left to a lay jury. Nevertheless, as ideal types Weber's categories help to locate idealized Western law in a wider universe of possible legal systems.

The importance of Weber's categories, like those of Maine, resides in large part in his efforts to link types of rationality with different types of societies and different ways of organizing legal systems. Weber associated an irrational legal order with domination by a charismatic leader. Formal rationality, on the other hand, accompanies the rise of the bureaucratic style of organization. Weber regarded logically formal rationality as the most "advanced" kind of legal ordering and as particularly hospitable to the growth of the capitalist state.

Weber's third contribution to comparative legal studies was his insight that the nature of a society's legal system is shaped by the kinds of individuals who dominate it. On the European continent, in the absence of a powerful central

court, domination fell into the hands of the university law faculties who strove, through the promulgation and interpretation of authoritative texts, to create and understand the legal system as a general and autonomous set of rules. The common law in England, on the other hand, grew under the tutelage of a small elite judiciary and an accompanying centralized bar, more concerned with pronouncing rules for the settlement of disputes than with developing generalized rules of law (Weber 1954). In time, the differences in the legal systems created by these different sets of legal actors helped to spur interest in comparative legal systems.

Overall, Weber's contribution was part of a general movement away from comparing the legal codes of various societies and toward a comparison of the legal solutions that "are given to the same actual problems by the legal systems of different countries seen as a complete whole" (Zweigert and Kotz 1987, p. 60, quoting Ernst Rabel). From this perspective, legal systems confront similar problems, and if we examine the whole system we will uncover fundamental differences and similarities in their various solutions. The effort to uncover these similarities and differences has taken several different paths.

**Macrocomparisons.** One path has involved attempts to develop macrocomparisons of entire legal systems. This effort has resulted in a number of taxonomies of legal systems in which the laws of nations are grouped by what are commonly called "legal families." The criteria for classification and the ultimate categories of family types have varied from scholar to scholar. Among the factors that have been used are historical tradition, the sources of law, the conceptual structure of law, and the social objectives of law. Socialist writers have traditionally focused on the relationship of law to underlying economic relations and a society's history of class conflict (Szabo and Peteri 1977; Eorsi 1979), although more recent efforts paint a more complex picture that threatens some of the presumed differences between socialist and "capitalist" law (Sypnowich 1990). David (1950) and David and Brierley (1985) base their classification on ideology (resulting from philosophical, political, and economic factors) and legal technique. Zweigert and Kotz (1987, p. 69) base their classification on a multiple set of criteria they call the "style" of law. Legal style includes: historical background and

development, predominant modes of thought in legal matters (contrasting the use of abstract legal norms in civil law versus the narrow, reasoning by analogy typical of the common law), distinctive concepts (such as the trust in the common law and the abuse of right in civil law), the source of law (statutory or case law), and ideology (e.g., the ideology of socialist and Western legal families).

Given the wide variety of criteria used by various scholars, perhaps it is surprising that the resulting "families" tend to be quite similar. To provide but one example, Zweigert and Kotz (1987) divide the world into the following eight families: (1) Romanistic family (e.g., France); (2) Germanic family (e.g., Germany); (3) Nordic family (e.g., Sweden); (4) common law family (e.g., England); (5) socialist family (e.g., Soviet Union); (6) Far Eastern family (e.g., China); (7) Islamic systems; and (8) Hindu law. While some taxonomies may have fewer civil law divisions, this set of categories shares with many others a Eurocentric emphasis and a resulting inability to fit non-European legal systems easily into the taxonomy (see Ehrmann 1976; David and Brierley 1985), although the rise of non-Western societies such as Japan should help to redress this imbalance in time (see Institute of Comparative Law, Waseda University 1988).

The Eurocentric and Western emphasis is not simply a matter of greater particularity in describing differences between the legal traditions of Europe. It is also reflected in the concepts used to make distinctions. The categories of the various typologies are based primarily on a comparison of private law rather than on public or constitutional law and on substantive law rather than on adjective law. A different focus may lead to different family configurations. For example, American and German constitutional law are in some ways more similar to each other than to French or English constitutional law. The focus on private substantive law has the additional result that it overemphasizes legal doctrine while underemphasizing the degree to which legal systems are a product of the surrounding society. The consequence is to understate similarities in Western legal arrangements that may be captured by the idea of a legal culture.

One alternative designed to avoid this tendency is found in Merryman's concept of legal traditions (1969). Legal traditions are:

*a set of deeply rooted, historically conditioned attitudes about the nature of law, about the role of law in the society and the polity, about the proper organization and operation of a legal system, and about the way law is or should be made, applied, studied, perfected, and taught. The legal tradition relates the legal system to the culture of which it is a partial expression.* (Merryman 1969, p. 2)

From this perspective the Western legal tradition may be usefully compared to and contrasted with legal systems in other cultures (Barton et al. 1983).

A second alternative to the “legal families” approach is taxonomies that are not based on differences in substantive law. One recent example, closer to the Weberian heritage, is that of Damaska (1986). Like Weber, Damaska uses two dimensions to develop ideal-typical legal orders. The first dimension divides legal orders into activist and reactive systems of justice. Activist states attempt to use law to manage society, whereas reactive states attempt only to provide a legal framework for social interaction. At the heart of the image of law of the activist state is the state decree, spelling out programs, assigning tasks, and distributing welfare to citizens. At the heart of the reactive state are devices facilitating agreement, contracts, and pacts. While it might be thought that this dimension is designed primarily to distinguish capitalist and socialist legal orders, Damaska observes that not all types of socialist models follow the state socialism that has dominated the Soviet Union and eastern Europe. Yugoslavian self-management concepts speak to this reactive tradition in socialism. Likewise, capitalist societies exhibit considerable differences in their commitment to an activist state.

Damaska’s second dimension divides legal orders into hierarchical and coordinate systems of judicial organization. In the hierarchical ideal officials are professionals who are arranged in a strict hierarchy and who employ special, technical standards of decision making. The coordinate ideal describes a more amorphous machine in which legal functionaries are amateurs who are arranged in relationships of relatively equal authority and who do justice based on prevailing ethical, political, or religious norms. Weber’s vision of the Moslem Khadi applying substantive (religious) law would appear to describe this type of legal order.

There are other strong parallels between Damaska’s and Weber’s ideal types. Their categories are less obviously Eurocentric and, more important, employ a set of concepts that facilitate an understanding of ways in which the relationship between the state and society is mediated through law. Both analyses are inclined toward a functional approach. Rather than beginning with individual legal histories and doctrines and grouping them into families, this approach begins with a set of problems—how to mediate the relationship of the state and society and how to organize the structure of legal actors—and arranges legal systems according to how they address these problems.

Damaska’s distinction between the hierarchical and coordinate ideal and Weber’s distinction between formal and substantive rationality direct our attention to a central issue concerning law—the degree to which different legal systems are autonomous. Formalist theories of law posit a self-contained enterprise separate from the rest of society (Kelsen 1967), while most Marxist theories view law solely as an instrument of domination (Spitzer 1983).

More recent theoretical discussions of autonomy include those of the Critical Legal Studies Movement, Niklas Luhmann (1985), and Pierre Bourdieu (1987). Critical Legal Studies focuses on law’s indeterminacy and on the role of social forces and power relations as the actual determinants of legal outcomes (Kelman 1987). Luhmann’s theory, to the contrary, views the legal system as autopoietic. An *autopoietic system*, like a living organism, produces and reproduces its own elements by the interaction of its elements (Teubner 1988). Bourdieu offers a complex view of the autonomy of the “juridical field.” Legal system autonomy is the result of the constant resistance of the law to other forms of social practice. One way this is accomplished is by requiring those who wish to have their disputes resolved in court to surrender their ordinary understandings and experiences. Actions and actors brought into the legal system are dealt with only after they and their dispute are translated into a set of legal categories (e.g., debtor—unsecured creditor; lessor—lessee;). Western courts tend to treat as irrelevant and inappropriate those accounts that attempt to introduce the details of litigants’ social lives (Conley and O’Barr 1990).

Interest in the question of legal autonomy has been reinvigorated by the collapse of socialist regimes in central and eastern Europe. The legal systems in the formerly communist societies of central and eastern Europe enjoyed relatively less autonomy than their Western counterparts. For example, judges were sometimes told how to decide politically sensitive cases. Because the directive usually came via telephone, the practice acquired the pithy name “telephone law.” In the last decade much attention has been devoted to the question of how to establish the rule of law—that is, greater legal system autonomy—in Russia and eastern Europe (Hendley 1996). One would be wrong, however, to conclude that legal systems in capitalist societies enjoy complete autonomy. The intrusion of politics into the legal systems of these societies is more subtle but nevertheless substantial (Jacob et al. 1996). What is different is the way judicial field is structured: for example, the relative independence of the judiciary and the organization of the legal profession. These structures, as Bourdieu notes, make it easier for the legal system to resist penetration. Indeed, intrusion is not a one-way street. The “centrality of law” thesis argues that the legal system increasingly penetrates into other areas of modern society (Hunt 1993). Modern states have “more” law, larger state bureaucratic legal institutions, and growing numbers of legal professionals. Moreover, legal forms are exported to other spheres of life, as when the workplace becomes infused with due process requirements.

It should be clear by now that legal autonomy is a multifaceted phenomenon. Systems differ in their degree of judicial independence and judicial formalism, the extent to which their laws are status neutral, and whether those forced into the legal arena enjoy equal legal competence (Lempert 1987). One of the strengths of Weber’s and Damaska’s typologies is that they suggest dimensions along which legal system autonomy may vary. For example, systems that reflect Damaska’s hierarchical ideal will be more likely to exhibit some of the features of greater autonomy.

**Microcomparisons.** Microcomparisons of legal systems are concerned with the details of specific legal rules and institutions rather than with entire legal systems (Rheinstein 1968). The functional approach is even more pronounced at this

level. Scholars often begin with a specific social problem and seek to discover the various ways in which legal systems solve it, or they begin with a specific legal institution and examine how it operates in various systems. For example, Shapiro (1981) makes a comparative analysis of the court as an institution in common law, civil law, imperial Chinese, and Islamic legal systems.

The most valuable work done at this level has been that of legal anthropologists. By examining the dispute-processing activities of African, Latin American, and Asian legal tribunals, they have provided new insights into the connection between a society’s social relationships and the way in which it processes disputes. Ethnographies by Gluckman (1967), Gulliver (1963), Nader (1969), and others exposed a general pattern wherein tribunals confronted with disputes among individuals who are in multiplex and enduring relationships are more likely to widen the range of relevant evidence and to search for outcomes that allow flexibility, compromise, and integration. Tribunals confronted with disputes among individuals who are in one-dimensional and episodic relationships are more likely to narrow the range of relevant evidence and to provide binary outcomes in which one side clearly wins and the other loses.

Legal ethnographies have also supported the earlier observation based on macrocomparisons that the organization of courts and judges plays a role in determining styles of dispute processing. Fallers (1969), for instance, found that the Soga, a society in many ways very similar to the Barotse studied by Gluckman, tended to craft decisions that were narrower and that resulted in “legalistic” rulings. His explanation was that the “judiciary” in the two societies differed in at least one key respect. The Soga courts were more purely “judicial” bodies without administrative and executive functions. A specialized legal staff was more likely to issue narrower opinions. Moreover, because binary outcomes result in a judgment to be enforced against a losing party, the availability of a coercive judicial apparatus may facilitate this type of dispute resolution (Lempert and Sanders 1986).

Perhaps because of the seminal work by Llewellyn and Hoebel (1941) on the Cheyenne, the work of legal anthropologists, more than most macro approaches to the study of legal systems,

builds on the sociological jurisprudence and the legal realist traditions (Pound 1911–1912; Oliphant 1928; Llewellyn 1930; Arnold 1935). It is concerned with the law in action, with the actual experience of the legal staff and the disputants (Merry 1990). As a consequence, legal anthropology has had a substantial influence on the sociological study of disputing and what has come to be called alternative dispute resolution in Western societies (Greenhouse 1986; Abel 1981). Postmodern legal anthropology has grown increasingly preoccupied with the problem that confronts all comparativist work—understanding the effect of the observers' own backgrounds on the ways in which we distinguish legal systems (Comaroff and Comaroff 1992). Undoubtedly, however, the ethnographic tradition has provided rich detail to our understanding of the differences among legal systems.

Recently, anthropologists have come to appreciate the degree to which African and other consensual legal systems are themselves partly the outgrowth of colonial experience and of the distribution of power in society (Starr and Collier 1989). This observation underlines a more general point that has been noted by macro and micro scholars alike. Nearly all existing legal systems are, to a greater or lesser extent, externally imposed, and therefore all legal systems are layered (Watson 1974). In many societies layering occurs because of the existence of a federal system creating an internal hierarchy of rules, some of which are imposed from above. Layered legal systems also occur when nations such as Turkey (the Swiss code) or Japan (the German code) shop abroad and adopt the laws of another nation as the basic framework for substantial parts of their own legal system. In some situations the imposition is done wholesale and involuntarily, as when colonial powers impose a legal system. The result can be considerable social dislocation (Burman and Harrell-Bond 1979). In time, multiple layers may exist, as in Japan, where indigenous law has been overlaid by both the adopted German code and American constitutional law concepts imposed after World War II (Haley 1991).

In each of these situations a society's legal system is unlikely to fit easily within any of the legal families. For instance, a society may borrow another's substantive and adjective law for commercial

law purposes but retain the existing law of domestic relations. Frequently, such societies are said to have a "dual legal system." However, to the degree that this phrase describes a situation in which two equal systems stand side by side and rarely interact, it fails to capture the rich variety of hierarchical structures in layered systems. An important task for the students of legal systems is to understand the process by which individuals and groups use law at different levels and in so doing transform both.

At the uppermost layer of legal systems are legal arrangements that are multinational or transnational in scope. Within the European Economic Community, following the Treaty of Rome in 1957 and the Single European Act in 1987, the adoption or imposition of a multinational regime is proceeding rapidly. The process requires the harmonization of a large body of law including corporate, intellectual property, environmental, tax, products liability, banking, transportation, product regulation (e.g., food and pharmaceuticals), and antitrust law. Member states must conform their national laws to comply with community directives, inevitably leading to the homogenization of European law. This process, along with the substantial alterations in property and contract law accompanying the economic changes in the Soviet Union and eastern Europe, suggests that the differences among legal systems of European origin will diminish over the next few decades, especially differences among laws governing commercial and economic transactions.

Indeed, the existence of a global economic order promotes some similarities in all laws governing economic transactions. Islamic law has been compelled to create a number of legal devices and fictions designed to avoid direct confrontation with several teachings of the Koran, such as the prohibition against charging interest that would make participation in a modern economic order difficult (David and Brierley 1985, p. 469). The emerging global economy has also created a new layer of transnational legal actors who at once attempt to export their nationalist version of law and to create a set of transnational institutional arrangements that sometimes complement and sometimes are in opposition to national legal structures. For example, Dezalay and Garth (1996) describe the emergence of a cadre of international commercial

arbitrators and their creation of an international legal field with its own networks, hierarchical relationships, expertise, and rules.

A number of additional global issues also create pressures toward the creation of transnational legal arrangements. These include transnational crime; ethnic and racial conflict; world population and migration patterns; labor flows; and, perhaps most significant, environmental regulation. Common legal structures created to address these issues and demands that nation-state legal systems enact and enforce appropriate compliance mechanisms may lead to the rebirth of the ideal of international legal unification that was popular at the beginning of the century. As can be seen in the European example, such unification inevitably involves some imposition of law.

Because pressures to build a more complex body of transnational law coincide with the diminution of differences in Western legal systems, over the next few decades one of the most interesting issues in the study of legal systems will involve movements toward and resistance to a transnational legal order premised on the hegemony of Western legal systems and Western legal concepts. The ongoing task of comparative law is to understand the processes of borrowing, imposition, and resistance, both among nations and between levels of legal systems.

(SEE ALSO: *Court Systems and Law*; *Law and Society*; *Social Control*; *Sociology of Law*)

## REFERENCES

- Abel, Richard 1981 *The Politics of Informal Justice*, 2 vols. New York: Academic.
- Arnold, Thurman 1935 *The Symbols of Government*. New Haven, Conn.: Yale University Press.
- Barton, John, James Gibbs, Victor Li, and John Merryman 1983 *Law in Radically Different Cultures*. St. Paul, Minn.: West Publishing.
- Black, Donald 1976 *The Behavior of Law*. New York: Academic.
- Bohannon, Paul 1965 "The Differing Realms of the Law." *American Anthropologist* 67 (6, point 2):133-142.
- Bourdieu, Pierre 1987 "The Force of Law: Toward a Sociology of the Juridical Field." (Richard Terdiman, trans.) *Hastings Law Journal* 38: 805-853.
- Burman, Sandra, and Barbara Harrell-Bond, eds. 1979 *The Imposition of Law*. New York: Academic.
- Comaroff, John, and Jean Comaroff 1992 *Ethnography and the Historical Imagination*. Boulder, Col.: Westview.
- Conley, John M., and William M. O'Barr 1990 *Rules versus Relationships: The Ethnography of Legal Discourse*. Chicago: University of Chicago Press.
- Damaska, Mirjan 1986 *The Faces of Justice and State Authority: A Comparative Approach to the Legal Process*. New Haven, Conn.: Yale University Press.
- David, René 1950 *Traité élémentaire de droit civil comparé*. (Introduction to the Principles of Comparative Civil Law.) Paris: Librairie Generale de Droit et de Jurisprudence.
- , and John Brierley 1985 *Major Legal Systems in the World Today*. London: Stevens.
- Dezalay, Yves, and Bryant G. Garth 1996 *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order*. Chicago: University of Chicago Press.
- Durkheim, Emile (1893) 1964 *The Division of Labor in Society*. New York: Free Press.
- Ehrlich, Eugen 1936 *Fundamental Principles of the Sociology of Law*. Cambridge, Mass.: Harvard University Press.
- Ehrmann, Henry 1976 *Comparative Legal Cultures*. Englewood Cliffs, N.J.: Prentice-Hall.
- Eorsi, Gy 1979 *Comparative Civil Law*. Budapest: Akademiai Kiado.
- Fallers, Lloyd 1969 *Law without Precedent: Legal Ideas in Action in the Courts of Colonial Busoga*. Chicago: University of Chicago Press.
- Gluckman, Max 1967 *The Judicial Process among the Barotse of Northern Rhodesia*. Manchester, England: Manchester University Press.
- Greenhouse, Carol 1986 *Praying for Justice*. Ithaca, N.Y.: Cornell University Press.
- Griffiths, John 1984 "The Division of Labor in Social Control." In Donald Black, ed., *Toward a General Theory of Social Control*. New York: Academic.
- Gulliver, P. M. 1963 *Social Control in an African Society: A Study of the Arusha*. London: Routledge and Kegan Paul.
- Haley, John 1991 *Authority without Law: Law and the Japanese Paradox*. New York: Oxford University Press.
- Hendley, Kathryn 1996 *Trying to Make Law Matter: Legal Reform and Labor Law in the Soviet Union*. Ann Arbor: University of Michigan Press.
- Hoebel, E. Adamson 1954 *The Law of Primitive Man*. Cambridge, Mass.: Harvard University Press.



- Hunt, Alan 1993 *Explorations in Law and Society*. New York: Routledge and Kegan Paul.
- Institute of Comparative Law, Waseda University, ed. 1988 *Law in East and West*. Tokyo: Waseda University Press.
- Jacob, Herbert, Erhard Blankenburg, Herbert Kritzer, Doris Marie Provine and Joseph Sanders. 1996 *Courts, Law and Politics in Comparative Perspective*. New Haven, Conn.: Yale University Press.
- Kelman, Mark 1987 *A Guide to Critical Legal Studies*. Cambridge, Mass.: Harvard University Press.
- Kelsen, Hans 1967 *The Pure Theory of Law*. Berkeley: University of California Press.
- Lempert, Richard 1987 "The Autonomy of Law: Two Visions Compared." In Gunther Teubner, ed., *Autopoietic Law: A New Approach to Law and Society*. New York: Walter de Gruyter.
- , and Joseph Sanders 1986 *An Invitation to Law and Social Science*. Philadelphia: University of Pennsylvania Press.
- Llewellyn, Karl 1930 "A Realistic Jurisprudence: The Next Step." *Columbia Law Review* 30:431–465.
- , and E. Adamson Hoebel 1941 *The Cheyenne Way: Conflict and Case Law in Primitive Society*. Norman: University of Oklahoma Press.
- Luhmann, Niklas 1985 *A Sociological Theory of Law*. London: Routledge and Kegan Paul.
- Maine, Henry (1861) 1963 *Ancient Law: Its Connection with the Early History of Society and Its Relation to Modern Ideas*. Boston: Beacon.
- Malinowski, Bronislaw 1926 *Crime and Custom in Savage Society*. New York: Harcourt, Brace.
- Merry, Sally Engle 1990 *Getting Justice and Getting Even: Legal Consciousness among Working-Class Americans*. Chicago: University of Chicago Press.
- Merryman, John Henry 1969 *The Civil Law Tradition: An Introduction to the Legal Systems of Western Europe and Latin America*. Stanford, Calif.: Stanford University Press.
- Montesquieu (1748) 1962 *The Spirit of the Laws*, 2 vols. New York: Hafner.
- Moore, Sally Falk 1973 "Law and Social Change: The Semi-Autonomous Social Field as an Appropriate Subject of Study." *Law and Society Review* 7:719–746.
- Nader, Laura 1969 "Styles of Court Procedure: To Make the Balance." In Laura Nader, ed., *Law in Culture and Society*. Chicago: Aldine.
- Oliphant, Herman 1928 "A Return to *Stare Decisis*." *American Bar Association Journal* 14:71, 159.
- Pound, Roscoe 1911–1912 "The Scope and Purpose of Sociological Jurisprudence," Parts 1, 2, and 3. *Harvard Law Review* 24:591–619; 25:140–168; 25:489–516.
- Rheinstein, Max 1968 "Legal Systems." Pp. 204–210 in David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 9. New York: Macmillan.
- Shapiro, Martin 1981 *Courts: A Comparative and Political Analysis*. Chicago: University of Chicago Press.
- Spitzer, Stephan 1983 "Marxist Perspectives in the Sociology of Law." *Annual Review of Sociology* 9:103.
- Starr, June, and Jane Collier, eds. 1989 *History and Power in the Study of Law: New Directions in Legal Anthropology*. Ithaca, N.Y.: Cornell University Press.
- Sypnowich, Christine 1990 *The Concept of Socialist Law*. Oxford: Clarendon Press.
- Szabo, I., and Z. Peteri, eds. 1977 *A Socialist Approach to Comparative Law*. Leyden, The Netherlands: A. W. Sijthoff.
- Tamanaha, Brian Z. 1997 *Realistic Socio-Legal Theory: Pragmatism and A Social Theory of Law*. Oxford, England: Clarendon.
- Teubner, Gunther, ed. 1988 *Autopoietic Law: A New Approach to Law and Society*. Berlin: Walter de Gruyter.
- Watson, Alan 1974 *Legal Transplants: An Approach to Comparative Law*. Charlottesville, Va.: University Press of Virginia.
- Weber, Max 1954 *On Law in Economy and Society*. New York: Simon and Schuster.
- Zweigert, Konrad, and Hein Kotz 1987 *Introduction to Comparative Law*, vol. 1, *The Framework*. Oxford, England: Clarendon.

JOSEPH SANDERS

## LAW AND SOCIETY

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

The concepts of *law* and *society* refer to macrostructural phenomena. Is there a macro-oriented theory of law and society or a macro-sociolegal theory to guide this field? As an interdisciplinary endeavor, the sociology of law relies upon, or is influenced by, the intellectual assumptions and propositions of general sociology and legal theory. This article will therefore consider the relationship of this field to both parent disciplines.

## RELATIONSHIP TO GENERAL SOCIOLOGY

It is no exaggeration to state that the field of sociology lacks a systematically developed and precise theory of society. Although interest in macro-sociological theory building has been in evidence for the past two decades, particularly among those concerned with comparative sociology (Eisenstadt and Curelaru 1977), no such theory has yet been developed in sufficient detail and precision to guide empirical research. This is not to deny the fact that such macrotheorists as Marx, Durkheim, Weber, and Parsons have exerted a pervasive influence on various specialties within sociology, including the relationship between law and society.

**Marx.** Marx conceived of law as a component of the “superstructure” of a capitalist society. As an epiphenomenon of the superstructure, it provides a rationale or ideology for preserving the existing class relations in a capitalist economy. Concepts of property and contract, for example, become instrumentalities for maintaining and reproducing class hegemony. In other words, legal concepts and doctrines reinforce the position of the ruling class and, at the same time, become the constituents of the “false consciousness” from which the working class suffers. Implicit in this theory of law as a weapon wielded by the state in a capitalist society against the working class is the assumption that if private property were abolished and a classless socialist society were ushered in, the state would “wither away” and, with it, law would “wither away” as well.

As usually formulated, the Marxian theory of law and society is not empirically verifiable. It does not follow, however, that this theory is devoid of any empirical implications. Questions can be raised—and have been raised—concerning class bias in the adjudication of civil and criminal cases, in the emergence of significant legal norms—for example, those regarding inheritance—and in the recurrent failure of agrarian reform laws. Likewise, it is possible to investigate a proposition counter to the Marxian thesis, namely, that the passage of laws in a capitalist state can potentially diminish the power of the ruling class vis-à-vis the working class. A case in point is the enactment of the National Labor Relations Act of 1935 in the United States, which institutionalized the rights of employees to unionize and to engage in collective bargaining with employers. Research questions

such as those cited above would test the validity of some propositions derivable from the Marxian theory of law and society.

**Durkheim.** Turning to Durkheim’s contribution to this field, one of necessity reverts to his *Division of Labor in Society* (1933), in which he argued that in societies characterized by “mechanical solidarity” there is a predominance of repressive laws, whereas in societies characterized by “organic solidarity” there is a predominance of restitutive laws. A number of social scientists have subjected Durkheim’s thesis to empirical tests and have found it wanting (Schwartz and Miller 1964). It is a testament, however, to the intriguing character of Durkheim’s thesis that it continues to evoke the interest of researchers (Baxi 1974; Schwartz 1974; Sheleff 1975). A more general formulation of Durkheim’s thesis would be that societies differing along various dimensions of societal development—of which the division of labor is but one—will exhibit systematic differences in their legal systems (Evan 1968).

In the course of developing his thesis that the division of labor is the principal source of social solidarity, Durkheim formulated his seminal idea of an “index” (Durkheim 1933, pp. 64–65). Apart from his fame as the “father of modern sociology,” Durkheim is the originator of the concept of an “index”, that is, an indirect and “external” measure of a complex dimension of social structure such as social solidarity. That he developed the concept of an index in connection with “juridical rules” and types of laws is of particular interest to sociologists of law and legal scholars. Under the circumstances, it is indeed surprising that to date, with few exceptions (Evan 1965, 1968, 1980; Merryman, Clark, and Friedman 1979; Lidz 1979), this facet of Durkheim’s work has been neglected. The concept of a “legal index” or a “legal indicator” merits systematic attention if we are to become more precise in our understanding of the role of law in social change.

**Weber.** In comparison with the work of Durkheim and Marx, Weber’s contributions to the sociology of law are appreciably more diverse and complex. Embedded in an intricate mosaic of ideal types and comparative and historical data on the emergence of legal rationality in Western civilization and on the role of law in the origins of capitalism (Weber 1950; Rheinstein 1954; Trubek

1972; Collins 1980), Weber's welter of legal conceptualizations poses a difficult challenge to the empirically oriented researcher. For example, his famous typology of lawmaking and lawfinding suggests possible research leads for comparative and historical analysis. Rheinstein, who edited and translated Weber's work on the sociology of law, lucidly summarizes his typology in the following manner:

1. *irrational, i.e., not guided by general rules*
  - a. *formal: guided by means which are beyond the control of reason (ordeal, oracle, etc.)*
  - b. *substantive: guided by reaction to the individual case*
2. *rational, i.e., guided by general rules*
  - a. *substantive: guided by the principles of an ideological system other than that of the law itself (ethics, religion, power, politics, etc.)*
  - b. *formal:*
    - (1) *extrinsically, i.e., ascribing significance to external acts observable by the senses*
    - (2) *logically, i.e., expressing its rules by the use of abstract concepts created by legal thought itself and conceived of as constituting a complete system.*

(Rheinstein 1954, p. 1)

Assuming that the meaning of each of these ideal type categories can be clarified and that legal indicators can be developed for each of the types, a comparative study could be undertaken to explore differences in lawmaking and in lawfinding of such major legal systems as common law, civil law, socialist law, and Moslem law. Equally challenging would be a study of long-term trends within each of these legal systems. The findings of such an inquiry would shed light on the occurrence of the evolutionary stages postulated by Weber.

*The general development of law and procedure may be viewed as passing through the following stages: first, charismatic legal revelation through "law prophets"; second, empirical creation and finding of law by legal honoratores; . . . third, imposition of law by secular or theocratic powers; fourth and finally, systematic elaboration of law and professionalized administration of justice by persons who have received their legal training*

*in a learned and formally logical manner.*  
(Rheinstein 1954, p. 303)

Another significant thesis in Weber's corpus of writings on law is the innovative role he attributes to "legal honoratores" or "legal notables" (Bendix 1960). Is Weber's thesis more valid for civil law systems, with its heavy immersion in Roman law, than it is for common law, let alone for socialist law or Moslem law? Once again it would be necessary to develop appropriate legal indicators to measure the degree to which legal notables—lawyers, judges, and high-level civil servants—introduce new rules and new interpretations of existing legal norms in the course of administering justice.

**Parsons.** For decades, Parsons was the leading macrosociological theorist in the United States, making singular contributions to structural functionalism and to a general theory of action. Focusing on the action of social systems, Parsons developed a "four-function paradigm." According to Parsons, every society faces four subsystem problems: adaptation, goal attainment, integration, and pattern maintenance or latency (AGIL). The societal subsystems associated with these four functional problems are, respectively, the economy, the polity, law, and religion and education.

Following Weber, Parsons treats law as a rational-legal system consisting of a set of prescriptions, proscriptions, and permissions. The legal system, especially in highly differentiated modern societies, performs the functions of a "generalized mechanism of social control" (Parsons 1962). This function is performed vis-à-vis the economy, the polity, and pattern maintenance or latency. The net effect of the pervasive normative regulation is the integration of society. As Parsons puts it: "The legal system . . . broadly constitutes what is probably the single most important institutional key to understanding . . . problems of societal integration" (Parsons 1978, p. 52).

With his four-function paradigm, Parsons addresses the nexus between law and society with the aid of "generalized media of interchange." The economy in a developed and differentiated society uses the medium of money for transactions. Functionally analogous media of exchange operate in each of the other subsystems—power in the polity, value commitment in pattern maintenance, and influence in law.

Suggestive as Parsons's framework is for understanding interinstitutional relations, the generalized media of interchange have not, as yet, been operationalized so as to explain how the legal system interacts with other societal subsystems. In other words, since Parsons has not explicated specific linkages between the legal and nonlegal subsystems, it is difficult to discern what hypotheses can be tested against any body of data. Hence, a reasonable conclusion is that Parsons's macro-sociological theory, in its present form, is actually a metatheory.

The foregoing review of some sociological theories of law and society raises two common themes: (1) each of the theorists endeavored to comprehend the macrostructural relationships between law and other institutional systems of a society, and (2) if the hypotheses implicit in these theories are to be empirically tested, systematic attention would have to be devoted to the development of a body of legal indicators. The current generation of sociologists of law has yet to face up to the problems engendered by both of these themes.

#### RELATIONSHIP TO LEGAL THEORY

Is the relationship between the sociology of law and the field of legal theory any less problematic than it is with general sociology? On its face, the question should be answered in the affirmative because the sociologist of law must take some of the legal scholars' subjects as objects of inquiry. In actuality, because of the traditions of legal scholarship, legal scholars do not generally provide an analytical basis for sociological research. Legal scholarship tends to be preoccupied with legal rules, legal principles, and their application to a multitude of specific conflict situations. As a consequence, the scholarly literature—apart from being intellectually insular—is almost entirely verbal and idiographic, with virtually no interest in a *nomothetic*, let alone *quantitative*, analysis of legal phenomena. Furthermore, there is a high degree of specialization within legal scholarship such that most scholars tend to devote their entire careers to a particular body of law, be it labor law, criminal law, contract law, family law, and so forth, in their own country. Those scholars specializing in comparative law are inclined to study a particular specialty, for example, family law, by comparing case studies from two or more countries (Glendon

1975). Relatively few legal scholars seek to study the legal system of an entire society, such as the work of Hazard (1977) and Berman (1963) on the Soviet legal system. And fewer still have had the temerity to undertake systematic comparisons of total legal systems or families of legal systems, as exemplified in the work of David and Brierly (1968) and Wigmore (1928); and those who make no effort tolerate characteristics of total legal systems to the social-structural attributes of the societies in which they are embedded.

Surveying current legal theory, three distinct theoretical perspectives can be discerned: the theory of legal autonomy, critical legal studies, and autopoietic law. Each of these perspectives will be briefly reviewed and appraised for their implications for a theory of law and society.

**Legal Autonomy.** Traditional conceptions of the legal order and "sources of law" are based on two assumptions, the first being that the law is a "seamless web," a relatively "closed system." Whatever processes of change occur in the law are generated from within the legal system, not from without. In other words, processes of change are immanent or endogenous and are not externally induced. The second assumption is that the legal system is, by definition, autonomous from other systems or institutions of a society. Therefore, it is unnecessary to inquire into how the legal system interacts with other subsystems of a society or into what degree of autonomy a given legal system actually has from other societal subsystems.

Perhaps the most quintessential articulation of the theory of legal autonomy in recent years can be found in the work of Watson, a renowned legal historian and comparative law scholar. Watson has repeated his thesis of legal autonomy in a number of monographs and articles (Watson 1974, 1978, 1981, 1983, 1985, 1987). He contends that the growth and evolution of the law is determined largely by an autonomous legal tradition, which exists and operates outside the sphere of societal needs.

*To a large extent law possesses a life and vitality of its own; that is, no extremely close, natural or inevitable relationship exists between law, legal structures, institutions and rules on the one hand and the needs and desires and political economy of the ruling elite*

*or of the members of the particular society on the other hand. If there was such a close relationship, legal rules, institutions and structures would transplant only with great difficulty, and their power of survival would be severely limited.* (Watson 1978, pp. 314–315)

*Law is largely autonomous and not shaped by societal needs; though legal institutions will not exist without corresponding social institutions, law evolves from the legal tradition.* (Watson 1985, p. 119).

Unlike the Marxist view of law, Watson's is that the law does not advance the interests of the ruling class; instead, it reflects the "culture" of the legal elite. He bolsters his provocative thesis with a study of legal borrowing, which he refers to as "legal transplants" (1974). The fact that the individual statutes, legal doctrines, and entire codes have been borrowed by countries differing in cultural, political, economic, and other respects provides evidence, according to Watson, in support of his thesis of legal autonomy.

The concept of "legal transplant" has a naturalistic ring to it as though it occurs independent of any human agency. In point of fact, however, elites—legal and nonlegal—often act as "culture carriers" or intermediaries between societies involved in a legal transplant. Legal scholars who are associated with political elites may be instrumental in effecting a legal transplant. Moreover, many instances of legal borrowing involve the "imposition" of a foreign body of law by a colonial power (Burman and Harrell-Bond 1979). Hence, it is a mistake to describe and analyze the diffusion of law as if it were devoid of human agency. If human volition is involved, it is indeed questionable whether the borrowed legal elements do not perform a societal function—at the very least on behalf of the legal elite.

**Critical Legal Studies.** Unlike Watson's internalist focus on the legal system and its autonomous development, the critical legal studies (CLS) movement appears to pursue a dual strategy: externalist as well as internalist. CLS is externalist in its critique of the social order and of the values dominating judicial decision making. It is internalist in its fundamental critique of traditional jurisprudence and legal reasoning.

The CLS movement emerged in the late 1970s in American law schools. It brought together a diverse group of scholars with a left-of-center ideology concerned about inequality and injustice in American society. Although lacking any consensus regarding societal transformation, CLS scholars sought to identify the impact of society's dominant interests on the legal process and the impact of social and political values on legal decision making.

In his introduction to a volume of essays by CLS authors, David Kairys discusses the "basic elements" of the legal theory of this movement. Three of these elements are externalist in nature:

*We place fundamental importance on democracy, by which we mean popular participation in the decisions that shape our society and affect our lives . . . We reject the common characterization of the law and the state as neutral, value-free arbiters, independent of and unaffected by social and economic relations, political forces, and cultural phenomena. The law's ultimate mechanism for control and enforcement is institutional violence, but it protects the dominant system of social and power relations against political and ideological as well as physical challenges.* (Kairys 1982, pp. 3–5)

These three externalist principles of the CLS movement have a familiar ring to them; namely, they are reminiscent of criticisms leveled by Marxists and neo-Marxists against the legal order of capitalist societies.

By far the most distinctive contribution of the CLS movement has been its elaborate internalist critique of legal reasoning and legal process. As Kairys puts it:

*We reject . . . the notion that a distinctly legal mode of reasoning or analysis characterizes the legal process or even exists . . . There is no legal reasoning in the sense of a legal methodology or process for reaching particular, correct results. There is a distinctly legal and quite elaborate system of discourse and body of knowledge, replete with its own language and conventions of argumentation, logic, and even manners. In some ways these aspects of the law are so distinct and all-embracing as to amount to a separate culture; and for many lawyers the courthouse, the law firm, the language, the*

*style, become a way of life. But in terms of a method or process for decision making—for determining correct rules, facts, or results—the law provides only a wide and conflicting variety of stylized rationalizations from which courts pick and choose. Social and political judgments about the substance, parties, and context of a case guide such choices, even when they are not the explicit or conscious basis of decision.* (Kairys 1982, p. 3)

Not only do critical legal scholars reject the notion of legal reasoning, they also reject other idealized components constituting a “legal system,” in particular, that law is a body of doctrine, that the doctrine reflects a coherent view of relations between persons and the nature of society, and that social behavior reflects norms generated by the legal system (Trubek 1984, p. 577).

The general conclusion CLS writers draw from “unmasking” the legal system, “trashing” mainstream jurisprudence, and “deconstructing” legal scholarship (Barkan 1987) is that “law is simply politics by other means” (Kairys 1982, p. 17). Such a conclusion, on its face, does not hold out any promise for developing a new, let alone heuristic, approach to a theory of law and society. On the contrary, its antipositivism combined with its search for a transformative political agenda has prompted CLS writers to view with increasing skepticism the sociology of law and research into the relationship between law and society (Trubek and Esser 1989).

**Autopoietic Law.** Similar in some respects to Watson’s theory of legal autonomy, but fundamentally different from the theory of the CLS movement, autopoietic law claims to be a challenging new theory of law and society (Teubner 1988a). For the past few years several continental social theorists, who are also legal scholars, have enthusiastically developed and propagated the theory of autopoietic law. A complex cluster of ideas, this theory is derived from the work of two biologists, Maturana and Varela (Varela 1979; Maturana and Varela 1980).

In the course of their biological research, Maturana and Varela arrived at some methodological realizations that led them to generalize about the nature of living systems. Maturana coined the term *autopoiesis* to capture this new “scientific epistemology” (Maturana and Varela 1980, p. xvii).

“This was a word without a history, a word that could directly mean what takes place in the dynamics of the autonomy proper to living systems.” Conceptualizing living systems as machines, Maturana and Varela present the following rather complex and abstract definition:

*Autopoietic machines are homeostatic machines. Their peculiarity, however, does not lie in this but in the fundamental variable which they maintain constant . . . an autopoietic machine continuously generates and specifies its own organization through its operation as a system of production of its own components, and does this in an endless turnover of components under conditions of continuous perturbations and compensation of perturbations.* (Maturana and Varela 1980, pp. 78–79)

Another definition of autopoiesis is presented by Zeleny, one of the early advocates of this new theory:

*An autopoietic system is a distinguishable complex of component-producing processes and their resulting components, bounded as an autonomous unity within its environment, and characterized by a particular kind of relation among its components, and component-producing processes: the components, through their interaction, recursively generate, maintain, and recover the same complex of processes which produced them.* (Zeleny 1980, p. 4)

Clearly, these definitions and postulates are rather obscure and high-level generalizations that, from a general systems theory perspective (Bertalanffy 1968), are questionable. Especially suspect is the assertion that autopoietic systems do not have inputs and outputs. The authors introduce further complexity by postulating second- and third-order autopoietic systems, which occur when autopoietic systems interact with one another and, in turn, generate a new autopoietic system (Maturana and Varela 1980, pp. 107–111). Toward the end of their provocative monograph, Maturana and Varela raise the question of whether the dynamics of human societies are determined by the autopoiesis of its components. Failing to agree on the answer to this question, the authors postpone further discussion (Maturana and Varela 1980, p. 118). Zeleny, however, hastens to answer this question

and introduces the notion of “social autopoiesis” to convey that human societies are autopoietic (Zeleny 1980, p. 3).

Luhmann, an outstanding German theorist and jurist, has also gravitated to the theory of autopoiesis. According to Luhmann, “social systems can be regarded as special kinds of autopoietic systems” (1988b, p. 15). Influenced in part by Parsons and general systems theory, Luhmann applied some systems concepts in analyzing social structures (1982). In the conclusion to the second edition of his book *A Sociological Theory of Law* (1985), Luhmann briefly refers to new developments in general systems theory that warrant the application of autopoiesis to the legal system. Instead of maintaining the dichotomy between closed and open systems theory, articulated by Bertalanffy, Boulding, and Rapoport (Buckley 1968), Luhmann seeks to integrate the open and closed system perspectives. In the process he conceptualizes the legal system as self-referential, self-reproducing, “normatively closed,” and “cognitively open”—a theme he has pursued in a number of essays (1985, 1986, 1988c).

This formulation is, to say the least, ambiguous. Given normative closure, how does the learning of the system’s environmental changes, expectations, or demands get transmitted to the legal system? Further complicating the problem is Luhmann’s theory of a functionally differentiated modern society in which all subsystems—including the legal system—tend to be differentiated as self-referential systems, thereby reaching high levels of autonomy (Luhmann 1982). Although Luhmann has explicitly addressed the issue of integrating the closed and open system perspectives of general systems theory, it is by no means evident from his many publications how this is achieved.

Another prominent contributor to autopoietic law is the jurist and sociologist of law Gunther Teubner. In numerous publications, Teubner discusses the theory of autopoiesis and its implications for reflexive law, legal autonomy, and evolutionary theory (Teubner 1983a, 1983b, 1988a, 1988b). One essay, “Evolution of Autopoietic Law” (1988a), raises two general issues: the pre-requisites of autopoietic closure of a legal system, and legal evolution after a legal system achieves autopoietic

closure. With respect to the first issue, Teubner applies the concept of *hypercycle*, which he has borrowed from others but which he does not explicitly define. Another of his essays (Teubner 1988b) reveals how Teubner is using this concept. For Teubner, all self-referential systems involve, by definition, “circularity” or “recursivity” (1988b, p. 57). Legal systems are preeminently self-referential in the course of producing legal acts or legal decisions. However, if they are to achieve autopoietic autonomy their cyclically constituted system components must become interlinked in a “hypercycle,” “i.e., the additional cyclical linkage of cyclically constituted units” (Teubner 1988b, p. 55). The legal system components—as conceptualized by Teubner, “element, structure, process, identity boundary, environment, performance, function” (1988b, p. 55)—are general terms not readily susceptible to the construction of legal indicators.

The second question Teubner addresses, legal evolution after a legal system has attained autopoietic closure, poses a similar problem. The universal evolutionary functions of variation, selection, and retention manifest themselves in the form of legal mechanisms.

*In the legal system, normative structures take over variation, institutional structures (especially procedures) take over selection and doctrinal structures take over retention.*  
(Teubner 1988a, p. 228)

Since Teubner subscribes to Luhmann’s theory of a functionally differentiated social system, with each subsystem undergoing autopoietic development, he confronts the problem of intersubsystem relations as regards evolution. This leads him to introduce the intriguing concept of *co-evolution*.

*The environmental reference in evolution however is produced not in the direct, causal production of legal developments, but in processes of co-evolution. The thesis is as follows: In co-evolutionary processes it is not only the autopoiesis of the legal system which has a selective effect on the development of its own structures; the autopoiesis of other subsystems and that of society also affects—in any case in a much more mediatory and indirect way—the selection of legal changes.* (Teubner 1988a, pp. 235–236)

Given the postulate of “autopoietic closure,” it is not clear by what mechanisms nonlegal subsystems of a society affect the evolution of the legal system and how they “co-evolve.” Once again, we confront the unsolved problem in the theory of autopoiesis of integrating the closed and open systems perspectives. Nevertheless, Teubner, with the help of the concept of co-evolution, has drawn our attention to a critical problem even if one remains skeptical of his proposition that “the historical relationship of ‘law and society’ must, in my view, be defined as a co-evolution of structurally coupled autopoietic systems” (Teubner 1988a, p. 218).

At least three additional questions about autopoietic law can be raised. Luhmann’s theory of a functionally differentiated society in which all subsystems are autopoietic raises anew Durkheim’s problem of social integration. The centrifugal forces in such a society would very likely threaten its viability. Such a societal theory implies a highly decentralized social system with a weak state and a passive legal system. Does Luhmann really think any modern society approximates his model of a functionally differentiated society?

A related problem is the implicit ethnocentrism of social scientists writing against the background of highly developed Western societies where law enjoys a substantial level of functional autonomy, which, however, is by no means equivalent to autopoietic closure. In developing societies and in socialist countries, many of which are developing societies as well, this is hardly the case. In these types of societies legal systems tend to be subordinated to political, economic, or military institutions. In other words, the legal systems are decidedly *allopoietic*. To characterize the subsystems of such societies as *autopoietic* is to distort social reality.

A third problem with the theory of autopoietic law is its reliance on the “positivity” of law. This fails to consider a secular legal trend of great import for the future of humankind, namely, the faltering efforts—initiated by Grotius in the seventeenth century—to develop a body of international law. By what mechanisms can autopoietic legal systems incorporate international legal norms? Because of the focus on “positivized” law untainted by political, religious, and other institutional values, autopoietic legal systems would have a

difficult time accommodating themselves to the growing corpus of international law.

Stimulating as is the development of the theory of legal autopoiesis, it does not appear to fulfill the requirements for a fruitful theory of law and society (Blankenburg 1983). In its present formulation, autopoietic law is a provocative metatheory. If any of its adherents succeed in deriving empirical propositions from this metatheory (Blankenburg 1983), subject them to an empirical test, and confirm them, they will be instrumental in bringing about a paradigm shift in the sociology of law.

The classical and contemporary theories of law and society, reviewed above, all fall short in providing precise and operational guidelines for uncovering the linkages over time between legal and nonlegal institutions in different societies. Thus, the search for a scientific macro sociolegal theory will continue. To further the search for such a theory, a social-structural model will now be outlined.

#### A SOCIAL-STRUCTURAL MODEL

A social-structural model begins with a theoretical amalgam of concepts from systems theory with Parsons’ four structural components of social systems: values, norms, roles, and collectivities (Parsons 1961, pp. 41–44; Evan 1975, pp. 387–388). Any subsystem or institution of a societal system, whether it be a legal system, a family system, an economic system, a religious system, or any other system, can be decomposed into four structural elements: values, norms, roles, and organizations. The first two elements relate to a cultural or normative level of analysis and the last two to a social-structural level of analysis. Interactions between two or more subsystems of a society are mediated by cultural as well as by social-structural elements. As Parsons has observed, law is a generalized mechanism for regulating behavior in the several subsystems of a society (Parsons 1962, p. 57). At the normative level of analysis, law entails a “double institutionalization” of the values and norms embedded in other subsystems of a society (Bohannan 1968). In performing this reinforcement function, law develops “cultural linkages” with other subsystems, thus contributing to the degree of normative integration that exists in a society. As disputes are adjudicated and new legal norms are enacted, a



value from one or more of the nonlegal subsystems is tapped. These values provide an implicit or explicit justification for legal decision making.

Parsons's constituents of social structure (values, norms, roles, and organizations) are nested elements, as in a Chinese box, with values incorporated in norms, both of these elements contained in roles, and all three elements constituting organizations. When values, norms, roles, and organizations are aggregated we have a new formulation, different from Parsons's AGIL paradigm, of the sociological concept of an institution. An institution of a society is composed of a configuration of values, norms, roles, and organizations. This definition is applicable to all social institutions, whether economic, political, religious, familial, educational, scientific, technological, or legal. In turn, the social structure of a society is a composite of these and other institutions.

Of fundamental importance to the field of the sociology of law is the question of how the legal institution is related to each of the nonlegal institutions. A preliminary answer to this question will be set forth in a model diagramming eight types of interactions or linkages between legal and nonlegal institutions (see Figure 1).

On the left-hand side of the diagram are a set of six nonlegal institutions, each of which is composed of values, norms, roles, and organizations. If the norms comprising the nonlegal institutions are sufficiently institutionalized, they can have a direct regulatory impact on legal personnel as well on the citizenry (interaction 4, "Single institutionalization"). On the other hand, according to Bohannan (1968), if the norms of the nonlegal institutions are not sufficiently strong to regulate the behavior of the citizenry, a process of "double institutionalization" (interaction 1) occurs whereby the legal system converts nonlegal institutional norms into legal norms. This effect can be seen in the rise in the Colonial period of "blue laws," which were needed to give legal reinforcement to the religious norms that held the Sabbath to be sacred (Evan 1980, pp. 517-518, 530-532). In addition, the legal system can introduce a norm that is not a component of any of the nonlegal institutions. In other words, the legal system can introduce an innovative norm (interaction 2) that does not have a counterpart in any of the nonlegal institutions (Bohannan 1968). An example of such

an innovation is "no-fault" divorce (Weitzman 1985; Jacob 1988).

The legal system's regulatory impact (interaction 3) may succeed or fail with legal personnel, with the citizenry, or with both. Depending on whether legal personnel faithfully implement the law, and the citizenry faithfully complies with the law, the effect on the legal system can be reinforcing (interaction 7) or subversive (interaction 5), and the effect on nonlegal institutions can be stabilizing (interaction 9) or destabilizing (interaction 6).

In systems-theoretic terms, the values of a society may be viewed as goal parameters in comparison with which the performance of a legal system may be objectively assessed. The inability of a legal system to develop "feedback loops" and "closed loop systems" to monitor and assess the efficacy of its outputs makes the legal system vulnerable to various types of failures. Instead of generating "negative feedback," that is, self-corrective measures, when legal personnel or rank-and-file citizens fail to comply with the law, the system generates detrimental "positive feedback" (Laszlo, Levine, and Milsum 1974).

## CONCLUSION

What are some implications of this social-structural model? In the first place, the legal system is not viewed as only an immanently developing set of legal rules, principles, or doctrines insulated from other subsystems of society, as expressed by Watson and to some extent by Luhmann and Teubner. Second, the personnel of the legal system, whether judges, lawyers, prosecutors, or administrative agency officials, activate legal rules, principles, or doctrines in the course of performing their roles within the legal system. Third, formally organized collectivities, be they courts, legislatures, law-enforcement organizations, or administrative agencies, perform the various functions of a legal system. Fourth, in performing these functions, the formally organized collectivities comprising a legal system interact with individuals and organizations representing interests embedded in the nonlegal subsystems of a society. In other words, each of the society's institutions or subsystems—legal and nonlegal—has the same structural elements: values, norms, roles, and organizations.

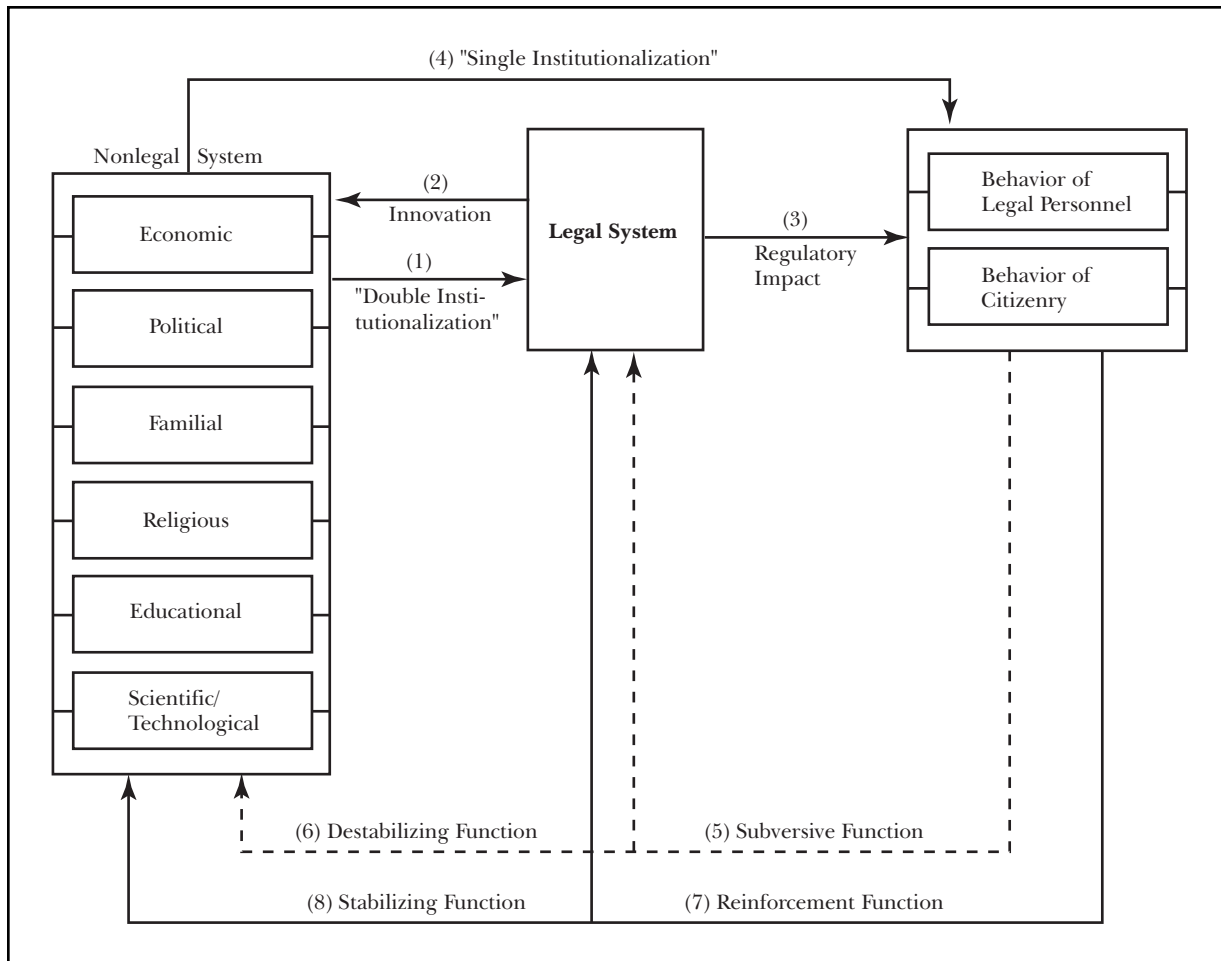


Figure 1

NOTE: A Social-Structural Model of the Interactions of Legal and Nonlegal Institutions

Interinstitutional interactions involve an effort at coupling these structural elements across institutional boundaries. A major challenge to the sociologists of law is to discover the diverse coupling or linkages—cultural and social-structural—between the legal system and the nonlegal systems in terms of the four constituent structural elements. Another challenge is to ascertain the impact of these linkages on the behavior of legal personnel and on the behavior of the citizenry, on the one hand, and to measure the impact of “double institutionalization” on societal goals, on the other.

A serendipitous outcome of this model is that it suggests a definition of law and society or the sociology of law, that is, that the sociology of law deals primarily with at least eight interactions or

linkages identified in Figure 1. Whether researchers accept this definition will be determined by its heuristic value, namely, whether it generates empirical research concerning the eight linkages.

(SEE ALSO: *Law and Legal Systems*; *Social Control*; *Sociology of Law*)

#### REFERENCES

- Abraham, David 1994 “Persistent Facts and Compelling Norms: Liberal Capitalism, Democratic Socialism, and the Law.” *Law and Society Review* 28:939–946
- Barkan, Steven M. 1987 “Deconstructing Legal Research: A Law Librarian’s Commentary on Critical Legal Studies.” *Law, Library Journal* 79:617–637.

- Barnett, Larry D. 1993. *Legal Construct, Social Concept: a Macrosociological Perspective On Law*. Hawthorne, N.Y.: Aldine De Gruyter.
- Baxi, U. 1974 "Comment—Durkheim and Legal Evolution: Some Problems of Disproof." *Law and Society Review* 8:645–651.
- Bendix, Reinhard 1960 *Max Weber: An Intellectual Portrait*. New York: Doubleday.
- Berman, Harold J. 1963 *Justice in the U.S.S.R: An Interpretation of the Soviet Law*. Cambridge, Mass.: Harvard University Press.
- Bertalanffy, Ludwig von 1968 *General System Theory*. New York: G. Braziller.
- Blankenburg, Erhard 1983 "The Poverty of Evolutionism: A Critique of Teubner's Case for 'Reflexive Law.'" *Law and Society Review* 18:273–289.
- Bohannon, Paul 1968 "Law and Legal Institutions." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- Buckley, Walter (ed.) 1968 *Modern Systems Research for the Behavioral Scientist*. Chicago: Aldine.
- Burman, Sandra, and Barbara E. Harrell-Bond (eds.) 1979 *The Imposition of Law*. New York: Academic Press.
- Chambliss, William J., and S. Zatz Marjorie 1993 *Making Law: The State, the Law, and Structural Contradictions*. Bloomington: Indiana University Press.
- Collins, Randall 1980 "Weber's Last Theory of Capitalism: A Systemization" *American Sociological Review* 45:925–942.
- Cooney, Mark 1995 "Legal Sociology and the New Institutionalism." *Studies in Law Politics, and Society* 15:85–101.
- Cotterell, Roger 1997 *Law's Community: Legal Theory In Sociological Perspective*. New York: Oxford University Press.
- 1988 "Why Must Legal Ideas Be Interpreted Sociologically?" *Journal of Law and Society* 25:171–192.
- David, Rene, and John E. C. Brierley 1968 *Major Legal Systems in the World Today*. London: Free Press and Collier-Macmillan.
- Durkheim, Emile 1933 *The Division of Labor in Society*. New York: Free Press.
- Eisenstadt, S. N., and M. Curelaru 1977 "Macro-Sociology: Theory, Analysis and Comparative Studies." *Current Sociology* 25:1–112.
- Evan, William M. 1965 "Toward A Sociological Almanac of Legal Systems." *International Social Science Journal* 17:335–338.
- 1968 "A Data Archive of Legal Systems: A Cross-National Analysis of Sample Data." *European Journal of Sociology* 9:113–125.
- 1975 "The International Sociological Association and the Internationalization of Sociology." *International Social Science Journal* 27:385–393.
- 1980 *The Sociology of Law*. New York: Free Press.
- 1990 *Social Structure and Law: Theoretical and Empirical Perspectives*. Thousand Oaks, Calif.: Sage Publications.
- Glendon, Mary Ann 1975 "Power and Authority in the Family: New Legal Patterns as Reflections of Changing Ideologies." *American Journal of Comparative Law* 23:1–33.
- Habermans, Jurgen, and William Regh 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge: MIT Press.
- Hazard, John N. 1977 *Soviet Legal System: Fundamental Principles and Historical Commentary*, 3rd ed. Dobbs Ferry, N.Y.: Oceana Press.
- Jacob, Herbert 1988 *Silent Revolution: The Transformation of Divorce Law in the United States*. Chicago: University of Chicago Press.
- Kairys, David (ed.) 1982 *The Politics of Law*. New York: Pantheon.
- Laszlo, C. A., M. D. Levine, and J. H. Milsum 1974 "A General Systems Framework for Social Systems." *Behavioral Science* 19:79–92.
- Lidz, Victor 1979 "The Law as Index, Phenomenon, and Element: Conceptual Steps Towards a General Sociology of Law." *Sociological Inquiry* 49:5–25.
- Luhmann, Niklas 1982 *The Differentiation of Society*. New York: Columbia University Press.
- 1985 *A Sociological Theory of Law*. trans. Elizabeth King and Martin Albrow. London: Routledge and Kegan Paul.
- 1986 "The Self-Reproduction of Law and Its Limits." In Gunther Teubner, ed., *Dilemmas of Law in the Welfare State*. Berlin: Walter de Gruyter.
- 1988a "The Sociological, Observation of the Theory and Practice of Law." In Alberto Febrajo, ed., *European Yearbook in the Sociology of Law*. Milan: Giuffre Publisher.
- 1988b "The Unity of Legal Systems." In Gunther Teubner, ed., *Autopoietic Law: A New Approach to Law and Society*. Berlin: Walter de Gruyter.
- 1988c "Closure and Openness: On Reality in the World of Law." In Gunther Teubner, ed., *Autopoietic Law: A New Approach to Law and Society*. Berlin: Walter de Gruyter.

- McIntyre, Lisa J. 1994 *Law in the Sociological Enterprise: a Reconstruction*. Boulder, Colo.: Westview Press.
- Maturana, Humberto R., and Francisco J. Varela 1980 *Autopoiesis and Cognition*. Dordrecht, The Netherlands: D. Reidel Publishing.
- Merryman, John Henry, David S. Clark, and Lawrence M. Friedman 1979 *Law and Social Change in Mediterranean Europe and Latin America*. Stanford, Calif.: Stanford Law School.
- Parsons, Talcott 1961 "An Outline of the Social System." In Talcott Parsons, Edward Shils, Kasper D. Naegle, and Jesse R. Pitts, eds., *Theories of Society*. New York: Free Press.
- 1962 "The Law and Social Control." In William M. Evan, ed., *Law and Sociology*. New York: Free Press.
- 1978 "Law as an Intellectual Stepchild." In Harry M. Johnson, ed., *Social System and Legal Process*. San Francisco: Jossey-Bass.
- Rheinstein, Max (ed.) 1954 *Max Weber on Law in Economy and Society*. Cambridge, Mass.: Harvard University Press.
- Schwartz, R. D. 1974 "Legal Evolution and the Durkheim Hypothesis: A Reply to Professor Baxi." *Law and Society Review* 8:653–668.
- , and J. C. Miller 1964 "Legal Evolution and Social Complexity." *American Journal of Sociology* 70:159–169.
- Sheleff, L. S. 1975 "From Restitutive Law to Repressive Law: Durkheim's *The Division of Labor in Society* Revisited." *European Journal of Sociology* 16:16–45.
- Tamanaha, Brian Z. 1996 "The Internal/External Distinction and the Notion of a "Practice" in Legal Theory and Sociological Studies." *Law and Society Review* 30:163–204.
- 1999 *Realistic Socio-Legal Theory: Pragmatism and A Social Theory of Law*. New York: Oxford University Press.
- Teubner, Gunther 1983a "Substantive and Reflexive Elements in Modern Law." *Law and Society Review* 17:239–285.
- 1983b "Autopoiesis in Law and Society: A Rejoinder to Blankenburg." *Law and Society Review* 18:291–301.
- 1988a *Autopoietic Law: A New Approach to Law and Society*. Berlin: Walter de Gruyter.
- 1988b "Hypercycle in Law and Organization: The Relationship Between Self-Observation, Self-Constitution, and Autopoiesis." In Alberto Febrajo, ed., *European Yearbook in the Sociology of Law*. Milan: Giuffrè Publisher.
- 1997 "The King's Many Bodies: The Self-Deconstruction of Law's Hierarchy." *Law and Society Review* 31:763–787.
- Travers, Max 1993 "Putting Sociology Back into the Sociology of Law." *Journal of Law and Society* 20:438–451.
- Trubek, David M. 1972 "Max Weber on Law and the Rise of Capitalism." *Wisconsin Law Review* 730:720–753.
- 1984 "Where the Action Is: Critical Legal Studies of Empiricism." *Stanford Law Review* 36:575.
- , and John Esser 1989 "'Critical Empiricism' in American Legal Studies: Paradox, Program, or Pandora's Box?" *Law and Social Inquiry* 14:3–52.
- Varela, Francisco J. 1979 *The Principle of Autonomy*. New York: North-Holland.
- Watson, Alan 1974 *Legal Transplants: An Approach to Comparative Law*. Charlottesville: University of Virginia Press.
- 1978 "Comparative Law and Legal Change." *Cambridge Law Journal* 37:313–336.
- 1981 *The Making of the Civil Law*. Cambridge, Mass.: Harvard University Press.
- 1983 "Legal Change, Sources of Law, and Legal Culture." *University of Pennsylvania Law Review* 131:1,121–1,157.
- 1985 *The Evolution of Law*. Baltimore, Md.: Johns Hopkins University Press.
- 1987 "Legal Evolution and Legislation." *Brigham Young University Law Review* 1987:353–379.
- Weber, Max 1950 *General Economic History*. New York: Free Press.
- Weitzman, Lenore J. 1985 *The Divorce Revolution: The Unexpected Social and Economic Consequences for Women and Children in America*. New York: Free Press.
- Wigmore, John Henry 1928 *A Panorama of the World's Legal Systems*. 3 vols. St. Paul, Minn.: West Publishing.
- Zeleny, Milan 1980 *Autopoiesis, Dissipative Structures, and Spontaneous Social Orders*. Boulder, Colo.: Westview Press.

WILLIAM M. EVAN

## LAW ENFORCEMENT

See Criminology; Penology; Police.

## LEADERSHIP

The concept of leadership has been the focus of research and discussion of scholars in a variety of

disciplines. Literary authors and philosophers provided the initial descriptions and guidance for leaders of their time. With the evolution of social sciences, scholars of political science, anthropology, sociology, psychology, and business have all explored the nature of leaders and the process of leadership. Although each of these various approaches did add a different perspective through the decades, there seem to be some who perceive a merging of disciplines in understanding leadership as we approach the close of the twentieth century. However, there are others who still maintain that leaders in different settings are fundamentally different.

As leadership is a vast field of research, only the most prevalent and unique approaches will be acknowledged in the following sections. Readers can find a more thorough review of leadership research and theories in texts such as Bass (1990), Chemers (1997), Chemers and Ayman (1993), and Dansereau and Yammarino (1998).

### HISTORICAL REVIEW

The definition of leadership has varied across time and cultures. In ancient times, the focus was on kings and rulers, who received guidance from philosophers. In many civilizations around the world, authors have written essays on good and bad leadership. For example, Confucius and Mencius wrote essays on leaders' proper behavior during the end of the fourth century and the beginning of the fifth century B.C. in China. Aristotle's book, *The Politics*, describes the characteristics of the kings and kingship in ancient Greece (fourth century B.C.). In eleventh-century Iran, Unsuru'l-Ma'ali wrote *Qabus-Nameh* and Nezam Mulk Tussi wrote *Siyassat Nameh*, advising the kings of the time in effective governance. Machiavelli wrote *The Prince*, in Florence, Italy, during the sixteenth century, guiding European rulers in politics. Ibn-e-khaldun from Tunisia provided his observations and guidance to the ruling groups of North Africa in his famous book *Muqiddimah* in the fourteenth century.

From the end of the eighteenth century to the beginning of the twentieth century, during the evolution of industrialization, the concept of "leader" also came to include leaders of industry. Only a hundred and fifty years ago the first chief executive officer (CEO) appeared (Smoler 1999). Smoler

stated that the increased volume of economic activity required the increased efficiency and administrative coordination that gave rise to this kind of leadership. The twentieth century saw steep growth in the numbers of CEOs and of top executive management positions.

At the close of the twentieth century there is a debate on differentiating between leaders and managers, and between leadership and management (Kotter 1988). The saying that some leaders and managers both lead and manage, but most leaders do not manage and most managers do not lead, may make this distinction. A popular phrase that clarifies the relation between leader and managers is, "Good managers do things right and good leaders do the right thing."

At the beginning of the twentieth century, two schools of thought on leadership evolved (Ayman 1993; Bass 1990). In Britain, Thomas Carlyle (1907) presented the "great-man theory," and in Germany, Karl Marx (1906) explained leadership from a zeitgeist approach (i.e., as being in the spirit of the time). The great-man approach focuses on the unique characteristics of the individual as the basis for being a leader. In this approach there is a strong belief that some people have leadership qualities and some do not; thus the born-leader concept. In the zeitgeist approach, the assumption is that the situation provides the opportunity for the individual to become a leader. Thus, it is not the person but the circumstances or the waves of time that puts people in leadership positions. These two approaches have been the foundation of leadership research throughout the twentieth century.

As the end of the century is approaching, the fascination with leaders seems to be at an all-time high, based on the number of publications about and for leaders (Ayman 1997). Many of these publications are inspirational and guiding. The scholars of leadership maintain a steady effort to refine the conceptualization of leadership and enhance the methodologies used in the study of leadership.

The twentieth century was the beginning of scientific and systematic study of leadership. Many disciplines have contributed to this endeavor, such as political science, anthropology, sociology, social psychology, and, most recently, industrial and organizational psychology. Each discipline has explored this phenomenon uniquely, guided by its

own perspective and orientation. The psychologists' focus was initially influenced by the "great-man theory," whereas the sociologists and social psychologists were more interested in the situational perspective. As the investigations proceeded and the disciplines evolved, the interplay of these approaches and the complexity of leadership became apparent. The following sections will briefly review various approaches: the trait and behavior approaches, which evolved from the great-man theory; the situational approach, which evolved from the zeitgeist approach; and, finally, the contingency approach, which combined the two approaches. At the end of this article, some challenges facing leadership practice and research are delineated.

### TRAIT APPROACH

The initial work on leadership was influenced by the *great-man theory*, which focuses on the characteristics of individuals in leadership situations. The primary focus of this approach is to compare the characteristics of those who are leaders and those who are not. The majority of trait research began in the early part of the twentieth century and declined somewhat in the late 1940s. In a majority of these studies, there seemed to be a search for one trait that best differentiated leaders from nonleaders. Stogdill (1948) reviewed studies in various settings and identified a list of traits that were most commonly studied to identify leaders. The traits that differentiated leaders from nonleaders included "sociability, initiative, persistence, knowing how to get things done, self-confidence, alertness to and insight into situations, cooperativeness, popularity, adaptability, verbal facility" (Bass 1990, p. 75). However, Stogdill (1948) concluded that research on leader's traits is inconclusive and that, depending on the situation, one trait may be more important than another.

Early industrial psychologists interested in selection maintained their focus on the trait approach. Therefore, primary pursuit of trait approach was by practitioners who were interested in selection and succession planning of managers in work settings. Later, Lord and colleagues (1986), with the assistance of meta-analytical techniques, reanalyzed past research by Stogdill (1948) and Mann (1959) and, across studies, identified certain

characteristics more associated with leaders than with nonleaders. The result of these studies was to identify intelligence, masculinity, and dominance as strongly related to leadership, whereas adaptability, introversion/extroversion, and conservatism had a weaker relationship. The adjustment and flexibility competency has received attention in recent years. In various studies, adaptability, as operationalized by the self-monitoring scale (Snyder 1979), has also been used to predict leader emergence and effectiveness with some degree of success (Zaccaro et al. 1991; Ayman and Chemers 1991). Hogan and colleagues (1994) revived interest in the trait approach by introducing a multitrait approach to the study of leaders. The traits included in the profile are known as the "big five" and consist of extroversion, emotional stability, openness, intellect, and surgency. In addition to the big five, there are also other measures (e.g., Myers Briggs) that are used to identify leader's profiles. However, results for predicting effective leaders from these traits are inconclusive. Overall, most of the trait research has focused on differentiating the characteristics of leaders from nonleaders, and on measuring leadership potential.

### BEHAVIORAL APPROACH

The first known study focusing on leader behavior examined the differing effects of democratic, autocratic, and laissez-faire leadership behaviors (Lewin et al. 1939; Lippitt and White 1943). These studies were conducted with groups of Boy Scouts led by trained graduate students. The results of these studies showed that groups with democratic leaders had the best-satisfied members. Those with autocratic leaders demonstrated the highest level of task activity, but only when the leader was present. Starting in the 1950s, the interest of the U.S. scholar gravitated toward understanding the leader's behaviors and the relationship between these behaviors and effectiveness. Three main research centers concurrently studied leader behavior in small teams at Ohio State University, led by Stogdill and his associates; at the University of Michigan, led by Likert and his colleagues; and at Harvard, led by Bales and his collaborators. Although these studies were conducted in different work settings such as the military, education, insurance companies, car manufacturing, and laboratory settings, they all found the same results.

These results led to two categories of leader behavior. One category dealt with behaviors that establish and maintain relationships, commonly referred to as considerate, people-oriented, or socio-emotional behaviors. The second category focused on behaviors that get the task accomplished, commonly referred to as initiating-structure, production-oriented, or task-focused behaviors. Although many measures were designed to assess the leader's behaviors, the most prominent are three Ohio State measures: the Leader Behavior Description Questionnaire (LBDQ), the Subordinate Behavior Description, and the Leader Opinion Questionnaire (Cook et al. 1985).

The leadership behavior researchers in other countries identified two similar categories of structure and consideration, but also found some additional categories. Unlike the findings in the United States, the two behaviors in Iran were found to be intertwined, resulting in the concept of "benevolent paternalism," or the "father figure" leader (Ayman and Chemers 1983). In India, Sinha (1984) identified "nurturant-task" behavior as an addition to the two main behavioral categories. In Japan, Misumi (1985) introduced a measure that had behavioral categories similar to those in the U.S. findings, but the behaviors were assessed in context. He referred to this model as maintenance-production (MP) behaviors.

In the 1960s and 1970s, researchers investigated the relationship between these two categories of behaviors and various indexes of effectiveness (e.g., team satisfaction, performance, turnover, and grievance) in different work settings. Fisher and Edward's meta-analysis (1988) is among the studies that supported the overall effects of these two behaviors (cited in Bass 1990).

Since the middle of the 1980s, a paradigm shift in the study of leadership behavior emerged, which moved scholars from focusing on consideration and initiating structure behaviors, often referred to as "transactional leadership," to what is referred to as "transformational leadership" behavior. This movement started with the work of McGregor Burns (1978) and House (1977), and was further developed by two groups of scholars, Bass and Avolio (see Bass 1985); Avolio and Bass (1988) in the United States, and Conger and Kanungo (1987) in Canada. In this new approach,

the leader provided guidance to ensure that the work was done properly and that subordinates were happy, and also was responsible for developing a relationship with employees based on mutual trust. It was proposed that transformational leadership not only changes the subordinate, but is also evolutionary for the leader and the task. A recent meta-analysis has provided a promising review of the effects of transformational leadership behavior in a variety of settings (Lowe et al. 1996). Also, it was through these approaches that the emphasis on empowerment in leadership evolved (Conger and Kanungo 1988b). Empowering, coaching, and facilitating are behaviors that were not noted extensively in leadership literature until the middle of the 1980s. During this period many inspirational and guiding documents were presented by various authors (e.g., Kouzes and Posner 1987).

Three main categories of criticism have been directed at the behavioral approaches. The first, advanced by Korman (1966) and by Kerr and colleagues (1974), argued that situational factors moderate the relationship between leader's behavior and outcomes. In addition, Fisher and Edwards (1988) more recently reviewed the studies using LBDQ and outcome variables, and acknowledged that there is a need for studying moderators.

The second criticism is based on the fact that almost all leader behavior measures are based on perception of self or the perception of others (Ayman 1993). In the 1980s, a series of studies demonstrated that perceived behavior is contaminated with various factors that are not necessarily related to the leader's behavior (Lord and Maher 1991). In these studies it was demonstrated that people's memory of the behaviors that occurred can be affected by what they think of the leader (e.g., Larson 1982; Philips and Lord 1982). The implication of this criticism is critical in exploration of diversity and cross-cultural leadership, topics that will be discussed later.

The third criticism focused on the level of conceptualization and analysis. This criticism is based on the premise that leaders may treat different people differently. That is to say, is leadership a group-level phenomenon, dyadic, or is it in the eyes of the beholder? Recently, a multilevel approach has been investigated in relation to various

leader behavior paradigms (see Dansereau and Yammarino 1998).

### **SITUATION AND LEADERSHIP**

In the 1950s, parallel to the development of leader behavioral studies, many scholars examined situational factors that gave rise to the identification of a leader without examining their characteristics. Some examined seating arrangement, or distance of the individuals (e.g., Bass et al. 1953; Howells and Becker 1962). The results of these studies demonstrated that individuals at the head of the table and who commanded more space were more often identified as the leader. Other studies examined communication patterns (Leavitt 1951) in the work team. The results demonstrated that those who have access to most of the information and are the center of the information flow are most often identified as leaders. The frequency of contribution to the team's goal achievement and verbal visibility was another contributor to the emergence of the leader (Bavelas et al. 1965). The position in the organization and assigned roles were also demonstrated to predict how a new leader behaved (Shartle 1951). That is, the best predictor of leadership behavior was the behavior of the leader's boss or predecessor. By the 1960s, there was some evidence that different tasks elicited different behaviors from the leaders (e.g., Morris and Hackman 1969).

In the 1980s, substitutes for leadership theory (Kerr and Jermier 1978; Howell et al. 1986) diverted the attention of leadership researchers again to the situation and provided a taxonomy for the role of the situation in leadership. This new approach categorized situations as neutralizers, enhancers, supplements, and substitutes (Schriesheim 1997). This approach has received mixed results, but the contribution of its taxonomy and methodological clarification has clarified the role of the situation in leadership research (Howell 1997).

In conclusion, the situational studies demonstrated that regardless of the leader's values, traits, or other characteristics, situational factors can influence the emergence of the leader and the behavior of the leader. Visibility and control of information are some of the necessary situational factors that assisted an individual being recognized as a leader. The more recent work on the

situation has further extended the role of the situation in leadership effectiveness.

### **LEADERSHIP EMERGENCE**

Early situational research and trait approaches focused on leadership emergence. Most studies on leader behavior attended to the effectiveness of leaders and teams. Hollander's concept of idiosyncrasy credit (1958) examined the exchange that occurs in the work team and leader emergence process. This concept gave rise to a series of studies (e.g., Hollander 1978; Hollander and Julian 1970) that demonstrated that, in a team, the members who contribute to the team's goal achievement and are loyal to the values held by the team establish credibility or idiosyncrasy credits. These studies demonstrated that idiosyncrasy credits then could be spent as the individual took risks to lead the group. Hollander's work may be seen as the first work that brought leadership and the management of innovation and change close together.

Hollander's work also demonstrated that the exchange process in teams resulted in the emergence of the member with the highest credits as the team leader. Later, the work of Graen (1976) further expanded the nature of this exchange in the vertical-dyad linkage (VDL) model. In the VDL model, the leader and the subordinates establish their relationship by negotiating their roles. Therefore, in work groups, there are some members who form the core group and are closer to the leader. Also, there may be other members who stay on the periphery of the work group.

The results of the initial studies were criticized, and unfortunately this line of research was abandoned due to methodological limitations (Dansereau et al. 1975). However, the notion of exchange between the leader and his or her subordinates is still a fertile area of research, known as leader-member exchange theory (e.g., Graen and Uhl-bien 1995).

### **CONTINGENCY APPROACHES**

Korman (1966) and Stogdill (1948) criticized past research on leader behavior and traits, and called for the study of leadership to be situation-based. In the early 1970s, two main categories of contingency approaches emerged, one based on the leader's trait and the other based on the leader's behavior.



The first scholar who operationalized the contingency approach to leadership and presented a trait-based model in this paradigm was Fred Fiedler. Several narrative reviews of the model have been written (see Ayman et al. 1998; Fiedler 1978). This model has been an impetus for many studies and some have criticized it over the years. However, several meta-analyses have provided some validating support on the thirty years of research (e.g., Strube and Garcia 1981; Peters et al. 1985).

The model is composed of three elements: one is the leader's motivational orientation, as measured by the "Least Preferred Coworker" scale; the second is the degree of leader's control and influence in a situation; and the third is the match between the situation and the leader's trait. In this model, three aspects of the situation were examined: leader-members relationship, task structure, and position power. The model's assumption is that leaders who are more focused on getting their job done than their relationship with people, do better than their counterparts in situations of high and low control and they are known to be in-match. Similarly, in situations of moderate control, those leaders with more focus on relationships will be more effective and are known as being in-match. When task oriented leaders are in moderate control condition and relationship oriented leaders are in high or low situational control conditions, they are referred to as being out-of-match (Ayman et al. 1995 and Ayman et al. 1998).

The strength of this model has been primarily in predicting team performance. More recently, subordinates' satisfaction and leader's experience of stress have also been included as outcomes examined by the model. Another strength of the model is that the information about the dependent variables and the independent variables are gathered from different sources. Leaders usually provide their motivational orientation by completing a quasi-projective test of Least Preferred Coworker. Sometimes the leader and other times the experimenter assessed the situation. A third party primarily has provided the assessments of the performance.

One area of potential improvement for the model is to include more than one trait. Recently, additional traits have been incorporated into the contingency model research. In 1993, Fiedler proposed the cognitive resource theory. This theory

included the leader's intelligence and experience and proposed that their contribution to the team's success may have contradictory effects depending on the stressfulness of the situation. Also, Ayman and Chemers (1991) included the leader's self-monitoring capability in their study, which was conducted with Mexican managers. Results showed that subordinates' work satisfaction was higher for out-of-match leaders who are high self-monitors than for those who were lower self-monitors. This may indicate that self-monitoring trait compensate for leaders who are out of match based on the contingency model of leadership effectiveness. However, the subordinates of in-match, high, self-monitoring leaders were not as satisfied with their work as the low self-monitoring in-match leaders. Therefore, self-monitoring seemed to only work for those who find themselves in situations that are not conducive to their internal state. As this is only one study, future work on these or other traits is warranted.

The contingency models of leader behavior examined the moderators that influence the relationship between leaders' behaviors and the outcome variables. As there are multiple leader behaviors and multiple ways that the situation can be conceptualized, many models have emerged. Two models have received substantial empirical attention (i.e., path goal theory [House 1971]) and the normative model of decision making [Vroom and Yetton 1973; Vroom and Jago 1974]). However, other models have made conceptual contributions, such as the multiple-influence model of leadership (Hunt 1991) and the multiple linkage model (Yukle 1989). These models' contributions have been in the development of taxonomy for situational factors and managerial behavior (see Chemers 1998). The behavior domains studied in this category of contingency models are managerial, supervisory, and decision-making behavior. However, most managers and leaders are also involved with managing interpersonal conflict, and this behavior domain has not received much attention by leadership researchers. In this section we will primarily review path goal theory and the decision-making model, and will acknowledge the relevance of conflict management to leadership.

Path goal theory (House 1971) was based on Vroom's expectancy model of motivation. Its proposition was that the leader as a supervisor needed to help the subordinates by clearing their path to goal

attainment. Therefore, the model predicted that the leader's behavior would be dependent on the situation, which was determined by the characteristics of the subordinates and the nature of their work. Therefore it predicted that subordinates with high ability were more satisfied and less stressed when their leaders were perceived as behaving considerately. Subordinates with complex tasks were more satisfied with leaders who were perceived as more directive. The model acknowledged four behavioral choices for the leaders: participative, supportive, directive, and achievement-focused. However, the majority of the research used Ohio State's measure of LBDQ-XII to define the leader's behavior as either considerate or structuring. In most of the studies, the descriptions of the situation, the subordinates' behavior, and the leader's behavior were assessed according to the subordinates' perceptions. Research on path goal theory received mixed results. In the majority of the studies either the subordinates' characteristics or their tasks were investigated. Also, the studies were primarily supportive, predicting subordinates' satisfaction but not their performance.

The decision-making model (Vroom and Yetton 1973; Vroom and Jago 1974) was mostly assessed from the leader's perspective. It provided the leader with choices of team decision-making strategies (e.g., very participative, partially participative, and very autocratic). The success of these choices depended on the situations. The model provided strategic questions that would help a leader determine the extent to which he or she should be participative or autocratic based on eight different criteria. These questions assessed the leader's knowledge and the subordinates' supportiveness in a given situation. The key finding was that subordinates who were part of the decision-making process were more satisfied, compared to those who were not part of this process.

Studies that have examined conflict management strategies of leaders are limited. The most commonly used measures are those by Rahim (1985) and by K. W. Thomas (1992). The two approaches identify five conflict management styles (competitive, withdrawing, accommodating, compromising, and collaborating) that represent the degree to which leaders are focused on themselves or the task, or the extent to which they are concerned about the subordinates' welfare.

In the same vein, the subordinates' choice of action with the leader was also examined. The situational factors studied so far have been the status of the actor in relationship to the target and the presence of an interpersonal conflict. The evidence seems to show that leaders have been perceived as using competitive and withdrawal strategies and subordinates tend to use more accommodating and compromising strategies. This line of research is new and requires more development. There is little evidence connecting choices of conflict management strategies to outcomes such as subordinate satisfaction or leader's influence or team performance.

Overall, the strength of behavioral approaches lies in their focus on the leader's behavior, which is assumed to present an aspect of the leadership process. These approaches are also sensitive to the moderating factor of situational elements; that is, they assume that not all successful or effective leaders behave the same way across situations. The behavioral approach provided some clarity in leadership process that allowed trainers to design training and development programs guiding leaders in how to behave.

These models have also faced challenges. First, the assessment of leader behavior and the situation in these models were dependent on the perception of the leader or the subordinates, thus susceptible to social cognitive influences, which will be briefly discussed in the next section. Second, most of these studies were based on a single source of information, that is, the same person who provided the predictors (e.g., the leader's behavior descriptions and the situation) also provided the assessment of the outcome and the moderator (i.e., situational factors). Finally, these models initially examined leadership only at a team level; thus assumed leaders treated everyone similarly. However, the criticism is that leader behavior may vary with each individual.

## NEW TOPICS AND FUTURE DIRECTIONS

After a century of empirical research, there are theoretical and methodological issues that have drawn the attention of the leadership researchers, theorists, and practitioners alike. Theoretically, three issues that need to be explored are: (1) the role of social cognition in leadership, (2) the role

of diversity and multiculturalism in leadership, and (3) moving toward an integrative theory. From a methodological perspective, four issues need to be considered: single sources variance, level of conceptualization and analyses, a systematic conceptualization and measurement of the situation, and measurement of leadership effectiveness.

Many scholars have recognized the role of social cognition in leadership studies (e.g., Calder 1977; Martinko and Gardner 1987). Some have argued that leadership is a product of our imaginations (Miendl 1990), whereas others have respected the role of perception in the study of leadership (Ayman 1993; Lord and Maher 1991). From these discussions and the studies that have transpired, it has been substantiated that people have an image of a leader in "their mind's eye." This schema or the mental map does affect their perception and evaluation of leaders they meet (Larson 1982; Lord et al. 1978).

Research on diversity and multiculturalism in leadership has gained momentum in the later part of the century. Initially many studies in the United States focused on gender. This may have been partially due to sample availability. Studies have demonstrated that overall, in the United States and other cultures, across age and gender, the image of a leader is more similar to the image of a man than a woman (e.g., Ayman-Nolley et al. 1993; Schein et al. 1989, 1996). Although many meta-analyses have reflected that, across studies, men and women leaders' behavior differences are not very noticeable (Eagly and Johnson 1990), and that the differences that were found were mostly on task behaviors. However, the authors identified various moderators in determining a leader's gender and behavior relationship, such as self/subordinate rating, male-dominated setting, and role congenial position (Eagly and Johnson 1990). In measuring leader effectiveness across indexes such as overall satisfaction and performance, no gender difference was found. However, women leaders were favored more than men were when subordinates' satisfaction was measured, and men did better than women when performance was measured. In addition, the moderator of the gender role congeniality of the job was identified as a major predictor of this relationship. Therefore, it is possible that the dominant masculine image of a leader (i.e., more prevalent in traditional male-dominated roles) has been the major challenge of

women in breaking the glass ceiling (Morrison and Van Glinow 1990).

As leadership process is dependent on perception and social judgement, the saliency of the stimuli's characteristics is important in this process. Although most of the diversity studies on leadership have been on gender differences, other diversity indexes also need to be examined. In studying leaders who are diverse, it is critical that the superficial diversity elements (i.e., those that can be seen, such as sex and color) and the deep diversity elements (i.e., those that cannot be seen, such as traits and education) be considered separately (Harrison et al. 1998). For example, Korabik (1997) has presented evidence that women may perceive themselves differently based on their sex-role orientation. But observers perceive women mostly based on the feminine stereotypes.

In addition, the process of social interaction between leaders and subordinates involves the leader as the actor and others as observers who judge this leader. In this case, both the leader's characteristics and the characteristics of subordinates need to be considered. In gender studies already reviewed, the gender composition of the group has been shown to affect the selection of a man or a woman as the leader (Eagly and Karau 1991) and to impact on men and women leaders' evaluation (Eagly et al. 1995). Research has demonstrated that in mixed-gender teams, men are more often favored than women. So, in studying leaders of other cultures or ethnic groups, it is not sufficient to study their behavior or expectations with their own culture; rather, it is critical to examine their interaction with people from other cultures (Ayman 1997). For example, Thomas and Ravlin (1995) found that American autoworkers rated more negatively on the dimension of trust when Japanese managers behaved like American leaders, but perceived these Japanese managers as more similar and effective.

Recently, several researchers and scholars have conducted cross-paradigm studies and provided integrated models of leadership. In the last five years there have been studies that examined the relationship between traits and leader behavior, such as Myers-Briggs and transformational behavior (Roush and Atwater 1992) and Bem's sex-role inventory and transformational leadership (Hackman et al. 1992). There are also studies connecting

leader member exchange with transformational leadership (Basu and Green 1997). More efforts to connecting various leadership research approaches are warranted. Chemers (1997) has presented an integrative model of leadership enabling researchers to approach leadership research with a more comprehensive understanding in the future. That is, although each study may be limited in scope, using a model like that of Chemers can help place it in the universe of leadership research. Doing cross-theory studies of leadership will enhance the understanding of this complex phenomenon.

The methodological challenge facing leadership research may be categorized in three topics (i.e., single-source variance, level issues, and measurement of the situation and the effectiveness of the leader). Many models reviewed in this article were validated on a single source of information. In other words, the same person who provided the predictor (e.g., the leader's behavior description) also provided the criterion (e.g., satisfaction or performance). This can artificially inflate the relationships under investigation (Yukle and Van Fleet 1992). One solution that some have used is that one subordinate describes the leader's behavior and another provides the effectiveness measure (Avolio et al. 1991). Although this is a step in the right direction, still both the leader and the subordinates have a vested interest in the relationship and thus may provide a biased perception.

Level of conceptualization and analysis is another methodological challenge that recently became the focus of leadership researchers' attention (Dansereau et al. 1995). These authors propose that the level of variables used in leadership studies needs to consider whether it is assessing the phenomenon at an individual, a dyadic, a group, or an organizational level. Also, it is important to consider the level issues in design and analysis of a study. In some studies the measures and the analyses may stay in the same level or go across levels to examine a meso-level relationship. The gravity of this issue and the methodologies involved in this procedure are reflected in a two-volume book edited by Dansereau and Yammarino (1998). In these volumes, many leadership approaches and theories are examined with the multilevel question in mind. It is not necessary for all models to be validated at all levels. However, it is valuable to

understand the parameters of the various models in relation to the level issue.

The third challenge is the assessment of leadership situation and leadership effectiveness. Although in situational studies and in contingency models some element of the situation is operationalized, a more systematic, cohesive model for evaluating these variables across theories would assist the field's evolution. Fiedler and Chemers (Ayman and Romano 1998) have identified three elements in the situation that affect the leaders' perceived control (leader-member relation, task structure, and position power). These aspects of situational control parallel French and Raven's sources of power (Ayman et al. 1998). The relative degree of importance of these aspects of situation has been substantiated both in contingency model research (Fiedler 1978) and in sources-of-power research (Podsakoff and Schriesheim 1985). Also, Hunt (1991) identified different levels within the organization as another situational factor affecting leadership. What is important to remember is that the prominent models of leadership reviewed here were validated in various work settings and industries. These models have been able to explain leadership regardless of whether the person is a political leader, a hospital administrator, a factory manager, a retail manager, a school principal, or a military leader. Consequently, future work is needed on a framework for identifying elements and dimensions that are important to consider in leadership situations. For example, one dimension that has been identified is the leader's power base. What are others?

Leadership effectiveness (e.g., group performance, leader's performance, group cohesion, subordinate satisfaction, organizational commitment, employee turnover rate) has been measured by various indexes and by various sources such as self report, superiors, and subordinates. Some models were more effective in predicting leaders emergence than effectiveness. Others have been more successful in predicting the well-being and cohesion of the group, such as subordinate satisfaction, turnover in the group, and subordinates stress. Several have predicted group performance measures by subordinates' evaluation or a superiors' assessment, or have used performance criteria such as financial indexes, sales, or goals met (e.g., sport teams' scores). It is important to consider the

parameter of models and the measures used to assess leadership. Some models may be better in predicting one criterion than another.

In conclusion, some have said, "Leadership is paradoxical; popular thinking . . . emphasizes the importance of . . . leadership establishing excellent organizations, but many academic publications assert that leadership is consequential" (Day and Lord, in Hunt 1991, p. 1). However, as Fiedler said in a personal comment to the author: "This is a pretzel shape universe that requires a pretzel shape model." The complexity of the world and the challenges it presents may call for a different form of leadership from the one-person-focused leadership. Since the team approach is being used in many organizations, it may be that the future of leadership will also be a team leadership approach. Compared to many fields of science, study of leadership is very young. Although challenges are present and the phenomenon is nebulous, leaders would attest that there are some things going on that make things work, but they are hard to explain. Future researchers are charged to systematically overcome the challenges and to collectively gain control of understanding of leadership in a global dynamic environment.

## REFERENCES

- Avolio, B., and B. M. Bass 1988 "Transformational Leadership, Charisma and Beyond." In J. G. Hunt, B. R. Balaga, H. P. Dachler, and C. Schriesheim, eds., *Emerging Leadership Vistas*. Elmsford, N.Y.: Pergamon.
- Avolio, B. J., F. J. Yammarion, and B. M. Bass 1991 "Identifying Common Methods Variance with Data Collected from a Single Source: An Unresolved Sticky Issue." *Journal of Management* 17:571-587.
- Ayman, R. 1993 "Leadership Perception: The Role of Gender and Culture." In M. M. Chemers and R. Ayman, eds., *Leadership Theory and Research: Perspectives and Directions*. New York: Academic.
- 1997 "Organizational Leadership." In S. W. Sadava and D. R. McCreary, eds., *Applied Social Psychology*. Upper Saddle River, N.J.: Prentice Hall.
- , and M. M. Chemers 1983 "Relationship of Supervisory Behavior Ratings to Work Group Effectiveness and Subordinate Satisfaction." *Journal of Applied Psychology* 68:338-341.
- 1991 "The Effect of Leadership Match on Subordinate Satisfaction in Mexican Organizations: Some Moderating Influence of Self-Monitoring." *Applied Psychology: An International Review* 40:299-314.
- , and F. Fiedler 1998 "The Contingency Model of Leadership Effectiveness: Its Levels of Analysis." In F. Dansereau and F. J. Yammarino, eds., *Leadership: The Multiple-Level Approaches—Classical and New Wave*. Stamford, Conn.: JAI.
- , and R. Romano 1998 "Appendix: Measures and Assessments for the Contingency Model of Leadership Effectiveness." In F. Dansereau and F. J. Yammarino, eds., *Leadership: The Multiple-Level Approaches—Classical and New Wave*. Stamford, Conn.: JAI.
- Ayman-Nolley, S., R. Ayman, and J. Becker 1993 "Gender Affects Children's Drawings of a Leader." Paper Presented At The Annual Convention Of American Psychological Society, Chicago (June).
- Bass, B. M. 1985 *Leadership and Performance beyond Expectations*. New York: Free Press.
- 1990 *Bass and Stogdill's Handbook of Leadership: Theory, Research, and Managerial Applications*, 3rd ed. New York: Free Press.
- , S. Kulbeck, and C. R. Wurster 1953 "Factors Influencing the Reliability and Validity of Leaderless Group Discussion Assessment." *Journal of Applied Psychology* 37:26-30.
- Basu, R., and S. G. Green 1997 "Leader-Member Exchange and Transformational Leadership: An Empirical Examination of Innovative Behaviors in Leader-Member Dyads." *Journal of Applied Social Psychology* 27:477-499.
- Bavelas, A., A. H. Hastorf, A. E. Gross, and W. R. Kite 1965 "Experiments on the Alteration of Group Structure." *Journal of Experimental Social Psychology* 1:310-320.
- Burns, J. M. 1978 *Leadership*. New York: Harper and Row.
- Calder, B. J. 1977 "Attribution Theory of Leadership." In B. M. Staw and G. R. Salanick, eds., *New Directions in Organizational Behavior*. Chicago: St. Clair.
- Carlyle, T. (1841) 1907 *Heroes and Hero Worship*. Boston: Adams.
- Chemers, M. M. 1997 *An Integrative Theory of Leadership*. Mahwah, N.J.: Lawrence Erlbaum.
- , and R. Ayman 1993 *Leadership Theory and Research: Perspectives and Directions*. New York: Academic.
- Conger, J. A., and R. A. Kanungo 1987 "Towards a Behavioral Theory of Charismatic Leadership in Organizational Settings." *Academy of Management Review* 12:637-647.
- 1988 "The Empowerment Process: Integrating Theory and Practice." *Academy of Management Review* 13:471-482.

- Cook, J. D., S. J. Hepworth, T. D. Wall, and P. B. Warr 1985 *The Experience of Work: A Compendium and Review of 249 Measures and Their Use*. New York: Academic.
- Dansereau, F., G. Graen, and W. J. Haga 1975 "A Vertical Dyad Linkage Approach to Leadership in Formal Organizations." *Organizational Behavior and Human Performance* 13:46-78.
- , and F. J. Yammarino (eds.) 1998 "Leadership: The Multiple-Level Approaches," vols. I and II. Stamford, Conn.: JAI.
- Dansereau, F., F. J. Yammarino, and S. E. Markham 1995 "Leadership: The Multiple-Level Approaches." *Leadership Quarterly* 6:97-110.
- Eagly, A. H., and B. T. Johnson 1990 "Gender and Leadership Style: A Meta-Analysis." *Psychological Bulletin* 108:233-256.
- , and S. J. Karau 1991 "Gender and the Emergence of Leaders: A Meta-Analysis." *Journal of Personality and Social Psychology* 60:685-710.
- , and M. G. Makhijani 1995 "Gender and the Effectiveness of Leaders: A Meta-Analysis." *Psychological Bulletin* 117(1):125-145.
- Fiedler, F. E. 1978 "The Contingency Model and the Dynamics of the Leadership Process." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 11. New York: Academic.
- 1993 "The Leadership Situation and the Black Box in Contingency Theories." In M. M. Chemers and R. Ayman, eds., *Leadership Theory and Research: Perspectives and Directions*. New York: Academic.
- 1995 "Cognitive Resource and Leadership." *Applied Psychology: An International Review* 44:5-28.
- , and M. M. Chemers 1984 *Improving Leadership Effectiveness: The Leader Match Concept*, 2nd ed. New York: John Wiley.
- Fisher, B. M., and J. E. Edwards 1988 "Consideration and Initiating Structure and Their Relationships with Leader Effectiveness: A Meta Analysis." In *Best Papers Proceedings, Academy of Management*. Anaheim, Calif.: Academy of Management.
- Graen, G. 1976 "Role Making Processes within Complex Organizations." In M. D. Dunnette, ed., *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- Graen, G. B., and Uhl-Bien 1995 "Relationship-Based Approach to Leadership: Development of Leader-Member Exchange (LMX) Theory of Leadership over 25 Years: Applying a Multi-Level Multi-Domain Perspective." *The Leadership Quarterly* 6:219-247.
- Hackman, M. Z., A. H. Furniss, M. J. Hills, and T. J. Paterson 1992 "Perceptions of Gender-Role Characteristics and Transformational and Transactional Leadership Behaviors." *Perceptual and Motor Skills* 75:311-319.
- Harrison, D. A., K. H. Price, and M. P. Bell 1998 "Beyond Relational Demography: Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion." *Academy of Management Journal* 41:96-107.
- Hogan, R., G. J. Curphy, and J. Hogan 1994 "What We Know about Leadership: Effectiveness and Personality." *American Psychologist* 49(6):493-504.
- Hollander, E. P. 1958 "Conformity, Status, and Idiosyncrasy Credit." *Psychological Review* 65:117-127.
- 1978 *Leadership Dynamics: A Practical Guide to Effective Relationships*. New York: Free Press.
- , and J. W. Julian 1970 "Studies in Leader Legitimacy, Influence, and Innovation." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 5. New York: Academic.
- House, R. J. 1971 "A Path Goal Theory of Leader Effectiveness." *Administrative Science Quarterly* 16:321-338.
- 1977 "A 1976 Theory of Charismatic Leadership." In J. G. Hunt and L. L. Larson, eds., *Leadership: The Cutting Edge*. Carbondale, Ill.: Southern Illinois University Press.
- Howell, J. P. 1997 "Substitutes for Leadership: 'Their Meaning And Measurement'—An Historical Assessment." *Leadership Quarterly* 8:113-116.
- , P. W. Dorfman, and S. Kerr 1986 "Moderator Variables in Leadership Research." *Academy of Management Review* 11:88-102.
- Howells, L. T., and S. W. Becker 1962 "Seating Arrangement and Leadership Emergence." *Journal of Abnormal and Social Psychology* 64:148-150.
- Hunt, J. G. 1991 *Leadership: A New Synthesis*. Newbury Park, Calif.: Sage.
- Kerr, S., and J. M. Jermier 1978 "Substitutes for Leadership: Their Meaning and Measurement." *Organizational Behavior and Human Performance* 22:374-403.
- Kerr, S. C., C. J. Schriesheim, C. J. Murphy, and R. M. Stogdill 1974 "Towards a Contingency Theory of Leadership Based upon the Consideration and Initiating Structure Literature." *Organizational Behavior and Human Performance* 12:62-82.
- Korabik, K. 1997 "Applied Gender Issues." In S. W. Sadava and D. R. McCreary, eds., *Applied Social Psychology*. Upper Saddle River, N.J.: Prentice Hall.

- Korman, A. K. 1966 "Consideration, Initiating Structure, and Organizational Criteria—A Review." *Personnel Psychology* 19:349–361.
- Kotter, J. P. 1988 *The Leadership Factor*. New York: Free Press.
- Kouzes, J. M., and B. Z. Posner 1987 *The Leadership Challenge: How To Get Extraordinary Things Done in Organizations*. San Francisco: Jossey-Bass.
- Larson, J. R., Jr. 1982 "Cognitive Mechanisms Mediating the Impact of Implicit Theories of Leader Behavior on Leader Behavior Rating." *Organizational Behavior and Human Performance* 29:129–140.
- Leavitt, H. J. 1951 "Some Effects of Certain Communication Patterns on Group Performance." *Journal of Abnormal and Social Psychology* 46:38–50.
- Lewin, K., R. Lippitt, and R. K. White 1939 "Patterns of Aggressive Behavior in Experimentally Created Social Climates." *Journal of Social Psychology* 10:271–301.
- Lippitt, R., and R. K. White 1943 "The Social Climate of Children's Groups." In R. G. Baker, J. S. Kounin, and H. F. Wright, eds., *Child Behavior and Development*. San Francisco: Freeman.
- Lord, R. G., J. F. Binning, M. C. Rush, and J. C. Thomas 1978 "The Effect of Performance Cues and Leader Behavior on Questionnaire Rating of Leadership Behavior." *Organizational Behavior and Human Performance* 21:27–39.
- Lord, R. G., C. L. Devader, and G. M. Alliger 1986 "A Meta-Analysis of the Relation between Personality Traits and Leadership Perceptions: An Application of Validity Generalization Procedures." *Journal of Applied Psychology* 71:402–410.
- Lord, R. G., and K. J. Maher 1991 *Leadership and Information Processing: Linking Perceptions and Performance*. Boston: Harper Collins.
- Lowe, K. B., K. G. Kroeck, and C. A. Campo 1996 "Effectiveness Correlates of Transformational and Transactional Leadership: A Meta-Analytic Review of the MLQ Literature." *The Leadership Quarterly* 7:385–426.
- Mann, R. D. 1959 "A Review of the Relationship between Personality and Performance in Small Groups." *Psychological Bulletin* 56:241–270.
- Martinko, M. J., and W. L. Gardner 1987 "The Leader/Member Attribution Process." *Academy of Management Review* 12:235–249.
- Marx, K. 1906 *Capital*. Chicago: Charles H. Kerr.
- Miendl, J. R. 1990 "On Leadership: An Alternative to the Conventional Wisdom." In B. A. Staw, ed., *Research in Organizational Behavior*, vol. 12. New York: JAI.
- Misumi, J. 1985 *The Behavior Science of Leadership: An Interdisciplinary Japanese Research Program*. Ann Arbor: University Of Michigan Press.
- Morris, C. G., and J. R. Hackman 1969 "Behavioral Correlates of Perceived Leadership." *Journal of Personality and Social Psychology* 13:350–361.
- Morrison, A. M., and M. A. Von Glinow 1990 "Women and Minority in Management." *American Psychologist* 45:200–208.
- Peters, L. H., D. D. Hartke, and J. F. Pohlmann 1985 "Fiedler's Contingency Theory of Leadership: An Application of the Meta-Analysis Procedures of Schmitt and Hunter." *Psychological Bulletin* 97:274–285.
- Philips, J. S., and R. G. Lord 1982 "Schematic Information Processing and Perception of Leadership in Problem-Solving Groups." *Journal of Applied Psychology* 67:486–492.
- Podsakoff, P. M., and C. A. Schriesheim 1985 "Field Studies of French and Raven's Bases of Power. Critique, Reanalysis, and Suggestions for Future Research." *Psychological Bulletin* 97:387–411.
- Rahim, M. A. 1985 "Referent Role and Styles of Handling Interpersonal Conflict." *The Journal of Social Psychology* 126:79–86.
- Roush, P. E., and L. Atwater 1992 "Using the MBTI to Understand Transformational Leadership and Self-Perception Accuracy." *Military Psychology* 4:17–34.
- Schein, V. E., R. Mueller, and C. Jacobson 1989 "The Relationship Between Sex-Role Stereotypes and Requisite Management Characteristics among College Students." *Sex Role* 20:103–110.
- Schein, V. E., R. Mueller, T. Lituchy, and J. Liu 1996 "Think Manager—Think Male: A Global Phenomenon?" *Journal of Organizational Behavior* 17:33–41.
- Schriesheim, C. A. 1997 "Substitute for Leadership Theory: Development and Basic Concepts." *Leadership Quarterly* 8:103–108.
- Shartle, C. L. 1951 "Studies in Naval Leadership." In H. Guetzkow, ed., *Groups, Leadership, and Men*. Pittsburgh: Carnegie.
- Sinha, J. B. P. 1984 "A Model of Effective Leadership Styles in India." *International Studies of Management and Organization* 14:86–98.
- Smoler, F. 1999 "The History of the CEO." *Worth* 5:150–154.
- Snyder, M. 1979 "Self-Monitoring Process." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 12. New York: Academic.
- Stogdill, R. M. 1948 "Personal Factors Associated with Leadership: A Survey of the Literature." *Journal of Psychology* 25:35–71.

- Strube, M. J., and J. E. Garcia 1981 "A Meta-Analysis Investigation of Fiedler's Contingency Model of Leadership Effectiveness." *Psychological Bulletin* 90:307-321.
- Thomas, D. C., and E. C. Ravlin 1995 "Responses of Employees to Cultural Adaptation by a Foreign Manager." *Journal of Applied Psychology* 80:133-147.
- Thomas, K. W. 1992 "Conflict and Negotiation Processes in Organizations." In M. D. Dunnette and L. M. Hough, eds., *Handbook of Industrial and Organizational Psychology*, 2nd ed., vol. 3. Palo Alto, Calif.: Consulting Psychologists.
- Vroom V. H., and A. G. Jago 1974 "Decision-Making as a Social Process: Normative and Descriptive Models of Leader Behavior." *Decision Sciences* 5:743-769.
- Vroom V.H., and P. W. Yetton 1973 *Leadership and Decision-Making*. Pittsburgh: University Of Pittsburgh Press.
- Yukle, G. 1989 *Leadership in Organizations*, 2nd ed. Englewood Cliffs, N.J.: Prentice Hall.
- , and D. D. Van Fleet 1992 "Theory and Research on Leadership in Organizations." In M. D. Dunnette and L. M. Hough, eds., *Handbook of Industrial and Organizational Psychology*, 2nd ed., vol. 3. Palo Alto, Calif.: Consulting Psychologists.
- Zaccaro, S. J., R. J. Foti, and D. A. Kenny 1991 "Self-Monitoring and Trait-Based Variance in Leadership: An Investigation of Leader Flexibility across Multiple Group Situations." *Journal of Applied Psychology* 76:179-185.

ROYA AYMAN, PH.D.

## LEARNING THEORIES

See Behaviorism; Socialization; Social Psychology.

## LEGISLATION OF MORALITY

In *The Division of Labor in Society* ([1893] 1984), Emile Durkheim advanced the idea that the distinctive sociological feature of crime is society's reaction to it. Durkheim was writing at a time when Lombroso's view on the heritability of criminality dominated scientific and popular opinion. Science sought the etiology of crime in the biology of the criminal, in *atavism*—crime viewed as a reversion to primitive, ancestral characteristics.

From this perspective, the harmfulness of criminality was taken for granted. Given his more comparative and anthropological outlook, Durkheim rejected the idea that crime was condemned by society because of its harmful consequences or because the deviant act was in itself evil. For Durkheim, many things that attracted the severest reprimands of society were objectively quite harmless, such as the neglect of the food taboos or the neglect of religious observance. In his view, the major issue was the integrative function of law, not the individual sources of deviance. Indeed, in *Suicide* ([1897] 1951) he characterized individual acts of despair as expressions of social structural pathologies. He also noted that the integrative function of law extracted far more from the condemnation of the lower class thief than from the middle-class embezzler—even though the latter's financial gain was greater and the social consequences of his acts much more adverse. From these observations Durkheim concluded that the societal reactions to crime and deviance were not based on rational models of the incapacitation or deterrence of the offender. In contrast to such utilitarian thinking, the societal reaction to crime was marked by an impassioned moral condemnation. This common feeling of moral outrage was a mark of the collective consciousness of society, particularly in highly cohesive, simple, or "tribal" societies. The condemnation of crime exercised the collective outlook, and it reinforced the group's collective beliefs and values. In short, the criminal was a scapegoat, and if he did not exist in fact, given his functional importance, he could be invented.

With the advance of more competitive and technologically sophisticated societies, the division of labor resulted in a partitioning of the collective consciousness across various elements of society and resulted in the rise of legal codes with less appetite for vengeance and with a greater investment in the reconciliation of competing interests. Hence, for Durkheim, punitive criminal law was the hallmark of primitive societies, which enjoyed a "mechanical" division of labor (i.e., a homogeneity of work functions), while reconciliatory civil law characterized advanced societies with an "organic" division of labor (i.e., a diverse but mutually dependent specialization of tasks). The former societies were characterized by brutal executions, the latter by written contracts.



Other nineteenth-century authorities tackled the role of law in society. Marx and Engels shared Durkheim's insight that the thing that needed explaining was not why people break laws—the problem of criminology—but why people make laws—the problem of the sociology of law (Cain and Hunt 1979). In their view, law was primarily an instrument of control and a source of both mystification and legitimation. As the urban merchant classes expanded their private estates into the countryside, these classes employed laws to transform the existing common rights of the rural peasants to the harvest of the forest—to deer, fish, pasture, firewood, and so forth—into crimes of theft from private property (Thompson 1975). In addition, the rise of capitalist agriculture, which was associated with the international trade in wool, cleared the British and European common lands of subsistent peasant farmers. Under contract law, the disenfranchised peasants became independent juridical subjects, able to engage in agreements to sell their labor for wages in the cities, although the urban working classes, having been displaced from subsistence on rural estates, had little choice, aside from starvation, but to work in factories.

Durkheim's position on law as an expression of the collective consciousness and Marx and Engel's views on law as an instrument of control continue to exert influence on contemporary thinking about law and morality. Where Durkheim treated the law as an expression of general social consciousness, Marx and Engels viewed it as an expression of a class consciousness, although, to be fair, Durkheim's model of collective consciousness held for a sort of mythic primitive society—akin to Hobbes on the “natural” state of man confronting the war of all against all or akin to Veblen or Freud's primal horde. In each case, the ideal type was an imaginative, indeed a fictional, reconstruction of human origins. The division of consciousness by class would have been possible for Durkheim in advanced societies—even if it constituted a state that he classified in *The Division of Labor in Society* as “abnormal” and even if it was the sort of awareness he wanted to overcome through the “corporations” that cut across class divisions and that Durkheim speculated might form the political nucleus of a future society.

In North America, sociological theories about the legislation of morality took a distinctive turn with investigations of “moral panics,” “victimless

crimes,” and “moral entrepreneurs”—all of which figured importantly in the labeling theories of the 1960s. The studies investigated distinctively morally charged issues and stressed the influence of both social movements and the mass media in the construction of the criminal law. By way of example, Sutherland (1950) examined the role of “moral crusaders” in the context of the sexual psychopathy laws that were enacted in 1937 and thereafter in eleven northern states, California, and the District of Columbia. These laws provided for indefinite incarceration in hospitals for the criminally insane for anyone pronounced by a psychiatrist to be a sexual psychopath. Retrospectively, this diagnosis is not based on a discernible medical disease. In fact, many psychiatric categories are only labels for things we do not like. Nonetheless, a wide series of jurisdictions, with one eye to reforming sexual deviants, enacted laws based on this “disease.”

There was a threefold process underlying the creation of these laws. First, the laws were enacted after a state of fear and hysteria followed newspaper accounts of several sex crimes committed coincidentally in quick succession. In these states there was a rush to buy guns, guard dogs, and locks and chains, reflecting widespread evidence of public fears. Second, these fears were fueled by news coverage of related sex crimes in other areas and other times in history and of sex-related behaviors and morality (including questions regarding striptease) and by letters to newspaper editors and statements made by public figures. The third phase was the creation of committees to study “the facts” of the sex crimes and “the facts” of sexual psychopathology. These committees, though initially struck on the basis of collective terror, persisted long after the fear and news stories subsided. They resulted in a presentation of briefs to legislative bodies, particularly by a new class of medical experts—psychiatrists. The community hysteria was legitimated by the identification of the crimes as a form of disease by the psychiatric professionals. The legislatures responded to public fears by passing laws to control the offending parties, in this instance by making sexual psychopaths subject to incarceration under psychiatric supervision.

A second important illustration is the history of the American prohibition movement, reported in the classic sociological study by Joseph Gusfield—*Symbolic Crusade* (1963). From 1919 to 1933 an

amendment to the American Constitution outlawed the manufacture and sale of alcoholic beverages. The law was passed after decades of lobbying by members of the numerous “temperance movements,” who originated predominantly from fundamentalist Protestant stock, largely from rural areas, and initially from middle-class backgrounds. They were the upholders of the Protestant ethic and valued hard work, self-reliance, and sobriety.

Gusfield argues that as America industrialized during the nineteenth century, the traditional rural settlers experienced a loss of social status compared to the new urban classes and experienced a challenge to their traditional values as urban development introduced increasingly nontraditional immigrants. The temperance movements were a form of status politics designed to reaffirm the prestige of a lifestyle. In the early period (1825–1875) the temperance movements pursued their objective through education, persuasion, and the reform of social conditions that caused excessive drinking. With increasing urbanization and greater European immigration, however, the tactics changed from a policy of socialization to one of coercive reform. The temperance movement became a prohibition movement. And when conservative forces came to dominate the political scene in America following World War I, the value of sobriety was to be achieved coercively—by outlawing booze. In Canada prohibition was approved in a national plebiscite in 1898 and enabled by federal law in 1916. The federal law allowed provinces to implement their own legislation, and all but Quebec outlawed booze, although booze continued to be manufactured for export—usually illicitly to the United States—and was available in Canada by prescription in pharmacies. The Canadian experiment in forced temperance ended in the mid-1920s, and the United States repealed Prohibition in 1933.

Gusfield’s work is important for several reasons. It stressed the idea that legislation was “symbolic”—that it reflected the values of distinctive social groups and that such groups struggled to have their views insinuated into state legislation. Unlike Marx, who viewed political struggle as class politics, Gusfield stressed “status” politics—the advancement of group respectability or status by seeking to have the group values enshrined in law. Gusfield further implied that laws, particularly in the field of “victimless crimes,” were symbolic in another sense. Instrumentally, prohibition was a

terrible failure. It created huge markets for organized crime, and liquor consumption appears to have proceeded in spite of the law. Hence, the condemnation of insobriety was merely ritualistic (i.e., symbolic) because it did not extirpate drunkenness nor seriously inhibit liquor consumption. Similar arguments have been raised regarding the criminalization of narcotics, prostitution, and pornography—classical “victimless crimes,” the markets for which thrive on a demand that seems little impeded by legal proscription.

The labeling theories of the 1960s developed the Durkheimian emphasis on the importance of societal reaction to crime (Kitsuse 1962; Lemert 1967) and identified some of the processes by which conduct was successfully labeled as criminal or how it successfully evaded such a label (Becker 1963). Theorists explored the self-fulfilling prophecies of deviant labels and the role of stigma in stabilizing rejected identities (Goffman 1961, 1963). The role of moral entrepreneurs was identified; these Weberian charismatics personified the struggle against evil and expedited legal change by campaigning in the media on the dire consequences of everything from dance halls to crime comics to television violence and marijuana use. Research on societal reaction proliferated. Studies focused on the manufacture of mental illness (Scheff 1999; Szasz 1961, 1970), the creation of witchcraft as a form of labeling (Erikson 1966; Currie 1968), the selective policing and the resulting social construction of delinquents (Cicourel 1968), the political hysteria that underlay political show trials (O’Connor 1972), the role of hysterical stereotypes in the criminalization of narcotics (Cook 1969; Lindesmith 1965), and the mythification of the Hell’s Angels in the popular press (Thompson 1966).

The studies of this period had an insurgent flavor and a barely concealed contempt for legal structures and institutions that drew the line between conformity and deviance in an arbitrary fashion. This was particularly true in the context of legislation outlawing “victimless crimes.” John Stuart Mill had established the ideological pedigree of this critique when he suggested that the democratic state’s sole justification for limiting the freedom of an individual was that the person’s behavior was harmful to others. Legal critics in the 1960s and 1970s objected to the extension of the law to cover vices and immoral activities since the participants in these acts, it was argued, were autonomous

beings who had chosen them voluntarily and so were mischaracterized as either criminals or victims. The argument was that participants were not criminals since their behavior was not harmful to others, and they were not victims since their fates were chosen voluntarily (Schur 1965). Since these activities flowed from free choice in democratic states, how could they have been made unlawful? The answer lay in the fact that conservative forces led by moral entrepreneurs had stampeded democratic governments through wild allegations that the activities in question were so harmful that they struck at the very fabric of society and that however voluntary or self-inflicted the corruption, its control and eradication was justified for the sake of society itself. More recent studies have taken a different turn. Studies of the law-and-order campaigns associated with the resurgence of political conservatism have led sociologists to ask different sorts of questions.

Where Gusfield stressed that legislation might take on a symbolic aspect that reflects the interests of discrete social groups, recent work suggests that laws may *mystify* the social conditions that give rise to them and consequently may result in social control that is sought for altogether different reasons than the ones identified in the legislation. Taylor (1982) argues that many crime waves, moral panics, and anticrime campaigns are orchestrated in response to basic social conflicts and shifting economic realities and that they function to misdirect our attention away from social contradictions. Three noted British investigations lend support to this view. Hall and colleagues (1978) argue that the moral panic in British papers over street mugging in the early 1970s that resulted in calls for tougher jail sentences and more law and order in England mystified a basic structural shift in employment patterns that occurred during attempts to dismantle the British welfare state. In the late 1960s and early 1970s there was a conservative political movement in both America and Great Britain to curb state investments in public welfare—to “down-size” governments and to privatize public institutions. There was a certain amount of conscious political resistance in the labor movement and in socialist political quarters. However, the control culture succeeded in redefining the political resistance into an issue of individual lawlessness that required tougher policing, particularly of those persons who experienced the greatest amount of

social dislocation as a result of fiscal restraint—poor youth, minority groups, and immigrants. The reports of street “muggings” imparted to England the imagery of lawlessness from the American ghetto, creating the impression that there were dramatic increases in street crime, that the crimes were disturbingly “un-British,” and that they were symptomatic of a wider threat to Great Britain’s collective security as witnessed by violence in industrial disputes (fights at picket lines) and misconduct among soccer fans.

Having defined the problem as individual lawlessness, the conservative solution of more law and order, greater investment in policing, longer jail sentences, and so forth, occluded the problem of youthful unemployment that *resulted* from conservative fiscal policies and that contributed to theft and other petty rackets in the first place. In this interpretation, moral panics have a material foundation in everyday experience, but where the earlier labeling theorists viewed the panics as *causes* of legal change, the British theorists suggest that panic was the *result* of social change. Sensationalism over crime in the popular media legitimates the introduction of coercive legislation that frequently suspends normal democratic liberties and replaces parsimonious forms of control with more punitive measures. In the case reported by Hall and colleagues (1978), penalties for petty street crimes increased dramatically as the public, the judiciary, and the politicians were treated to hysterical excesses in the popular press of violent youth running amok.

A similar structural argument is advanced by Stanley Cohen in *Folk Devils and Moral Panics* (1972). This study examines the role of the media in the re-creation of the “mods” and “rockers.” There have been a series of distinct trends in youth culture in both America and Great Britain over the last four or five decades—the zoot suiters, the Teddy Boys, the beats, the skin heads and punks, the soccer hooligans, and the motorcycle rebels—all of whom expressed antiauthoritarianism and youthful rebellion. What seems to have set off the postwar youth cultures was society’s relative affluence, which created conditions for distinctive consumption patterns supporting unique fashion styles, musical tastes, and recreational opportunities. The mods and rockers were motorcycle-riding youth groups who appeared in the British seaside resorts

in the middle 1960s. As Hunter Thompson discovered in his study (1966) of the California Hell's Angels, the British newspapers made a feast of the mods' antics, typically exaggerating and reporting spuriously on their activities. Cohen (1972) stresses that the function of the news hysteria was to create a kind of rogue's gallery of folk devils—vivid, even fearsome images that registered collective fears about the youth, their mobility, and their independence, which served as collective reminders of what youth should not be.

In this approach, moral panics are an ongoing, recurrent, and predictable aspect of the popular culture. From this perspective, they are not considered the handiwork of individual moralists who may or may not decide for personal reasons to pursue a moral campaign. Moral panics are orchestrated when individuals feel compelled by their sense of the collective consciousness to repair a breach in the ideological fabric of society. Typically, either such persons are affiliated with institutions that have a longstanding investment in moral control—church organizations, community groups, political parties, and so forth—or they occupy roles as self-appointed arbiters in democratic societies—academic experts, professional journalists, and publishers.

Also, moral panics and the legislative changes they engender track important shifts in social structure fairly closely. This is the conclusion of Pearson's *Hooligan: A History of Respectable Fears* (1983). Pearson reports that the current "public" appetite for law and order, for stiffer justice, and for a return to the past are recurrent themes in Western history. The historical perspective suggests that the feelings that, from the perspective of crime and delinquency, things are at an unprecedented nadir, that the present conduct of society compares poorly with the way things were in a previous golden age, that the family is falling apart, and that the popular entertainments of the lower classes are criminogenic, are recurrent in every major period of British history from the 1750s to the present. Over time, the rhetoric of decline has an uncanny similarity. Pearson's point is that crime fears or moral panics occur when the legitimacy of state control over the working class is somehow challenged or brought into question. In the nineteenth century the British working class was only partially integrated into the political process. Consequently, their consent to government—which

was dominated by the propertied classes—was usually testy. In the 1840s in particular, there is an apprehension in philanthropical writings about the working-class "dodgers," bold, independent delinquents viewed as potential revolutionaries who might fuel the Chartist movement and overthrow the private ownership of property. In this context, philanthropists prescribed education and policing to create internal restraints in the interest of protecting property.

From this perspective, class tensions were misinterpreted in the public press and by middle-class politicians as "rising crime and delinquency," and the middle-class solutions were recurrently more "law and order" on the one hand and education on the other—that is, more police repression and control, a return to birching, a removal of the un-British elements of the population, a repudiation of meretricious American culture, and a return to the tranquility of the golden age when people knew their place.

It is difficult to give an overall assessment of how the processes of legislating morality discovered in the labeling period in America can be integrated with the class conflict approaches stressed in the recent British studies. Certainly, there is no reason to believe that the class antagonisms that characterize British society are as developed in North America, nor that these would be a monolithic source of legal change. On the other hand, there is also no reason that the search for systematic sources of moral panics as effects of social change could not be undertaken in North American studies. Since the early 1980s it has become clear that the preoccupation with abortion, pornography, the funding of acquired immunodeficiency syndrome (AIDS) research, gay bashing, and the Equal Rights Amendment in U.S. public discourses is related to important trends in family composition and female labor force participation. For the tradition minded, the suppression of access to abortion, pornography, and the Equal Rights Amendment (ERA) as well as lethargy in dealing with AIDS, all seem to be an attempt to turn the clock back toward patriarchal families. For more progressive feminists, access to abortion on demand, suppression of degrading pornography, and the confrontation of the epidemics of sexual and physical violence against women and children are essential ideological matters for the insurance

of greater female social and economic advancement. In both cases, the public imagery of epidemics of fetal massacres on the one hand, and epidemics of incest and female abuse on the other, are not the products of idiosyncratic moral campaigns but arise in the context of profound social structural transformations—suggesting the need to develop a convergence of theories of symbolic interests and structural shifts that have developed independently to date.

As we enter the new millennium in American politics, it is clear that the contemporary issues in moral discourse are distinct from the earlier focus on the criminalization of things such as drugs and alcohol. Political observers have coined the term “culture war” to characterize the attempt of the “Moral Majority” of Christian right-wing politicians to topple “the moral decline of the nation” and to substitute “moral sanity” in the form of widespread laws consistent with their religious beliefs (Scatamburlo 1998). Despite the dominance in the U.S. House of Representatives and the Senate by leading proponents of “born-again” politics over the past decade (Shupe and Stacey 1982), the New Right has effectively capitulated and virtually despaired of incorporating its views into American public and personal life (CBS 1999). Its proponents have failed repeatedly to achieve significant change. Specifically, they have failed to reintroduce prayer into the public school system—at the very time when American schools were racked by senseless shootings by disgruntled students. They have failed to curb the use of abortion as a method of birth control, failed to oust a philandering president despite a relentless investigation by a special prosecutor, and failed to curb the spread of sexual explicitness and capricious violence in everything from video games to Hollywood movies in spite of the fear that a great deal of youthful violence appears to be copied from mass media images. Further, they have failed to protect the hegemonic status of the traditional, heterosexual family and its access to spousal benefits from incursions by lesbian and gay lobbyists. They have made headway in downsizing the welfare state, but this has been achieved to varying degrees throughout the Western democracies without the same moral agenda.

The other side of the culture war has been associated with the rise of “identity politics”—a

growth in social movements among *racial minorities*, women, and sexual minorities during the past two decades. In an attempt to recenter knowledge and power on the basis of experiences overlooked by male, Eurocentric, and heterosexual perspectives, claims have sometimes been made that gynocentric, Afrocentric, and gay voices have epistemological and moral standpoints that are superior to those which have traditionally overshadowed them (Seidman 1994, p. 234). Though such claims are extravagant, they signal cultural conflicts over basic moral and social questions and foster power politics resulting from the emancipation of racial minorities, women, and gays and lesbians. This explains in part the Moral Majority’s sense of crisis, its backlash against political correctness, and its desire to return to the previous equilibrium. Schlesinger argues that the crisis is further fueled by the feeling that the common presuppositions that are required for democracy are imperiled by the spread of “strident multiculturalism,” radical ethnocentrism, and the undermining of the very possibility of a common American identity (Schlesinger 1998). The matter is only intensified by the shifting racial composition of America. Maharidge’s *The Coming White Minority* (1996) documents the apprehension among white Californians about becoming a minority group in their “own” state, a trend true of America as a whole in the larger global perspective. The social processes are not unlike those identified by Gusfield in the analysis of the temperance movement, suggesting that the social forces explaining the cultural conflicts are not unfamiliar to students of moral panic. What is less well acknowledged, and what might be important to keep in mind for resolution of what is regrettably labeled the “culture war” is the following. From the start, the Moral Majority campaign tended to be exclusionary and fundamentalist (Snowball 1991). Democracy requires coalitions and a broad basis of mutual support. When politicians in a democracy take their policies from religion, moral condemnation overshadows political dialogue. Political decisions are not arrived at through compromise and mutual dependence but through faith. Given these tendencies, “average citizens” have failed to assent to the moral absolutes required by politics based on religious righteousness, and the movement appears to have run its course (Weyrich 1999). Whether the culture war is amenable to an alternative, more inclusive, multicultural resolution remains to be seen.

(SEE ALSO: *Deviance; Social Control*)

## REFERENCES

- Becker, Howard S. 1963 *Outsiders*. New York: Free Press.
- Cain, M., and A. Hunt 1979 *Marx and Engels on Law*. New York: Academic.
- Cicourel, Aaron 1968 *The Social Organization of Juvenile Justice*. London: Heinemann.
- Cohen, Stanley 1972 *Folk Devils and Moral Panics*. Oxford: Basil Blackwell.
- Cook, S. 1969 "Canadian Narcotics Legislation 1908–1923: A Conflict Approach." *Canadian Review of Sociology and Anthropology* 6:36–46.
- Columbia Broadcasting System (CBS) 1999 "The Moral Minority." *60 Minutes*, report by Leslie Stahl, May 16.
- Currie, Elliott P. 1968 "Crimes without Criminals: Witchcraft and Its Control in Renaissance Europe." *Law and Society Review* 3(1):7–132.
- Durkheim, Emile (1897) 1951 *Suicide*. New York: Free Press.
- (1893) 1984 *The Division of Labor in Society*. New York: Free Press.
- Erikson, Kai T. 1966 *Wayward Puritans*. New York: John Wiley.
- Goffman, Erving 1961 *Asylums*. New York: Doubleday/Anchor.
- 1963 *Stigma*. Englewood Cliffs, N.J.: Prentice-Hall.
- Gusfield, Joseph R. 1963 *Symbolic Crusade: Status Politics and the American Temperance Movement*. Urbana: University of Illinois Press.
- Hall, S., C. Critcher, T. Jefferson, J. Clarke, and B. Roberts 1978 *Policing the Crisis: Mugging, the State, and Law and Order*. London: Macmillan.
- Kitsuse, John I. 1962 "Societal Response to Deviance: Social Problems of Theory and Method." *Social Problems* 9:247–256.
- Lemert, Edwin 1967 "The Concept of Secondary Deviation." In E. Lemert, ed., *Human Deviance, Social Problems, and Social Control*. Englewood Cliffs, N.J.: Prentice-Hall.
- Lindesmith, A. 1965 *The Addict and the Law*. Bloomington: Indiana University Press.
- Maharidge, Dale 1996 *The Coming White Minority: California's Eruptions and the Nation's Future*. New York: Times Books/Random House.
- O'Connor, Walter D. 1972 "The Manufacture of Deviance: The Case of the Soviet Purge, 1936–1938." *American Sociological Review* 37:403–413.
- Pearson, Geoffrey 1983 *Hooligan: A History of Respectable Fears*. London: Macmillan.
- Scatamburlo, Valerie 1998 *Soldiers of Misfortune: The New Right's Culture War and the Politics of Political Correctness*. New York: Peter Lang.
- Scheff, Thomas 1999 *Being Mentally Ill*. New York: Aldine.
- Schlesinger, Arthur M. 1998 *The Disuniting of America: Reflections on a Multicultural Society*, rev. ed. New York: Norton.
- Schur, Edwin 1965 *Crimes without Victims*. Englewood Cliffs, N.J.: Prentice-Hall.
- Seidman, Steven 1994 *Contested Knowledge: Social Theory in the Postmodern Era*, Cambridge, Mass.: Blackwell.
- Shupe, Anson, and William S. Stacey 1982 *Born Again Politics and the Moral Majority: What Social Surveys Really Show*. New York: Edwin Mellen.
- Snowball, David 1991 *Continuity and Change in the Rhetoric of the Moral Majority*. New York: Praeger.
- Sutherland, Edwin 1950 "The Diffusion of Sexual Psychopath Laws." *American Journal of Sociology* 56:142–148.
- Szasz, Thomas 1961 *The Myth of Mental Illness*. New York: Heuber-Harper.
- 1970 *The Manufacture of Madness*. New York: Harper and Row.
- Taylor, Ian 1982 "Moral Enterprise, Moral Panic, and Law-and-Order Campaigns." In M. M. Rosenberg, R. A. Stebbins, and A. Turowitz, eds., *The Sociology of Deviance*. New York: St. Martin's.
- Thompson, E. P. 1975 *Whigs and Hunters*. Harmondsworth, England: Penguin.
- Thompson, Hunter S. 1966 *Hell's Angels*. New York: Random House.
- Weyrich, Paul 1999 "A Conservative Throws in the Towel." *FrontPage Magazine*, February 16. <http://www.frontpagemag.com/archives/politics/weyrich2-17-99.htm>.

AUGUSTINE BRANNIGAN

## LEGITIMACY

See Interpersonal Power; Organizational Structure.

## LEISURE

Since 1930 the sociology of leisure in North America and Europe has not developed in a linear or cumulative fashion. Rather, research agendas, the

accepted premises for research and theory, and the “common wisdom” of the field have been revised and challenged. Change did not come in one great overturning, but in a sequence of revisions. A dialectical model seems to be most appropriate to follow the sequence. Through the 1950s, there was an accepted consensus as to both issues and premises. This common wisdom was eroded as well as challenged by new research. The “revised consensus” expanded agendas for both research and theory without completely overturning earlier developments.

Since 1985, a more critical antithesis with multiple sources has emerged to subject the second consensus to a more thoroughgoing revision. The sources of this antithesis have included conflict or neo-Marxist theory, gender-focused critiques, non-Western perspectives, and various poststructural analytical approaches. Critiques are associated with concepts such as hegemony and power, commodification, cultural and social fragmentation, gender and patriarchal structures, imperialism, world views, symbol systems, ideologies, and existential action. Now a central question concerns the kind of synthesis that will be developed in the ongoing process.

The dialectical sequence provides a dynamic framework for a review of central areas in the study of leisure. Although many issues and lines of research can be identified, four have consistently been most salient. In a highly abbreviated form, we will summarize the dialectics of theory and research in relation to (1) work and time, (2) family and community, (3) aging and the life course, and (4) the nature of leisure.

### WORK AND TIME

**Leisure and Work Domains.** When sociologists turned their attention to leisure in the 1960s, three perspectives were adopted. The first, based on earlier community studies, approached leisure as a dimension of the social organization of the community (Lundberg, Komarovsky, and McInerney 1934; Dumazedier 1967). The second, exemplified by David Riesman and initiated at the University of Chicago, viewed leisure as social action that created its own worlds of meaning. The third, the one that came to shape domain assumption and research agendas, emerged from the sociology of work. Its fundamental premise was that economic

institutions are central to the society and economic roles the primary determinants of other roles. Especially leisure was assumed to be secondary and derivative. As a consequence, various models of determination by work were proposed that modeled leisure as similar to work (“spillover” or identity), contrasting (compensation), or separate (Wilensky 1960; Parker 1971). The bias, however, was clearly toward some kind of determination rather than segmentation.

As research proceeded, the “long arm of the job” was found to be both shorter and less powerful than expected as only limited, modest, and sometimes inconsistent relationships were found between leisure styles and occupational level and type (Wilson 1980). In a fuller perspective, on the other hand, it was evident that economic roles are determinative of the social context of adult lives—schedules, control of resources, autonomy, and other basic conditions (Blauner 1964). Leisure is part of the reward structure of a social system with differential access to resources based largely on socioeconomic position.

A second revision of the common wisdom concerned time available for leisure. The long-term reduction in the average workweek from as high as eighty hours in the early days of the industrial revolution to about forty hours in the post-World War II period along with the five-day workweek and paid vacations for many workers had led to an unquestioned assurance that more and more leisure time would be the product of increased economic productivity. In the 1970s, however, the declining rate of the decrease moving toward stability produced a revised consensus suggesting segmented time scarcity and a variety of social timetables. Most recently, analysis of labor statistics indicates that workers in high-pressure occupations and some services may have average workweeks much longer than forty hours (Schor 1991) even though time-diary research identifies a small overall increase in time for leisure (Robinson and Godbey 1997). Considerable attention has also been given to those impacted by the time scarcities of those, mostly women, with multiple work, household, and caregiving roles.

A next challenge to the early common wisdom was a recognition of leisure as a dimension of life with its own meaning and integrity. Leisure is more than leftover and derivative. It has its own

place in the rhythm and flow of life. First, leisure came to be defined more as activity than as empty time. Among the themes emerging were relative freedom of choice, distinction from the obligations of other roles, and the variety of meanings and aims that might be sought in such activity. Just as important as the revised definition, however, was the identification of leisure as something more than a derivation of work. Social life could not be divided into a work versus leisure dichotomy, but consisted of multiple sets of intersecting roles. Leisure, although it has a particular relationship with the bonding of family and other immediate communities (Cheek and Burch 1976), had multiple contexts, connections, and meanings (Kelly 1981).

**The Challenge of Critical Theory.** The domain assumptions of functional sociology have been challenged by critical analyses with roots in neo-Marxist cultural studies (Clarke and Critcher 1986; Rojek 1985), historical study that focuses on power and the struggles of the working class, and social construction approaches that take into account the interpretive symbolic activity of social actors (Rojek 1995).

The central theme of the critical challenge is social control by ruling elites. Leisure is seen as a critical element in the hegemony of ruling elites in a capitalist society. In order to assure compliance in the routinized "Fordist" workplace, the political arena, and the marketplace, leisure has emerged as central to the capitalist reward and control system. Leisure is, from this critical perspective, a market-mediated instrument that binds workers to the production process and to roles that support the reproduction of the capital-dominated social system. Leisure is defined as a commodity that must be earned and is indissolubly connected to what can be purchased and possessed.

A number of themes are gathered in this critique. The power to enforce compliance is masked behind an ideology in which "freedom" comes to be defined as purchasing power in the marketplace of leisure. Such "commodity fetishism" (Marx 1970) of attachment to things defines life and leisure in terms of possessions. Leisure, then, becomes "commodified" as the consumption of marketed goods and services, entertainment in contrast with commitment to involving,

challenging, and developmental activity (Kelly and Freysinger 1999). Social status is symbolized by leisure display (Veblen 1899). Absorption in mass media, especially low-cost and easy-access television (Robinson and Godbey 1997), legitimates consumption-oriented values and worldviews (Habermas 1975). What appear to be varying styles of leisure reflect the profoundly different conditions of work, family, and leisure assigned by class, gender, and race (Clarke and Critcher 1986).

**A Prospective Synthesis.** The fundamental presupposition of any sociology of leisure is that leisure is a thoroughly social phenomenon. It is a part of the culture and a product of the social system. Leisure is not separate and secondary, but embedded in the institutional structures, social times, and power allocations of the society. In a complex social system, both individual self-determination and institutional control differ by economic and social position. Leisure is not segmented, but woven into the system. Out of the current dialectic between the consensus and the critique, a number of issues call for attention.

The first agenda is to move beyond ideologies to examine the lived conditions of poor, excluded, and disinherited children, women, and men. Their struggles for life in the present, and their struggles for a future, are reflected in what they do to express themselves, create community, fill ordinary hours and days, and seek new possibilities.

The second issue is to identify the ways in which economic roles provide contexts, resources, limitations, and orientations for the rest of life—family and community as well as leisure. In a "post-Fordist" global economy with a loss of linear work careers and fragmented cultural schemes, the question is not the simple determination of life and leisure by work, but how determinative definitions of both the self and society are learned in a power-differentiated social context.

The third issue is meanings. Purchasing is not necessarily commodification and owning is not fetishism. What are the commitments, symbols, meanings, self-definitions, and worldviews that are the cognitive context of decisions and actions? What are the meanings and outcomes of leisure-related spending, media use, packaged entertainment and travel, and images of pleasure? Possession may be a way of life or an instrument of



activity. Does leisure reflect a culture of possession? Or is there a deep paradox between alienation and creation that permeates the entire society?

The fourth issue revolves around time. It is necessary to discard misleading models about average workweeks and the "more or less leisure" argument. Rather, what are the actual patterns and varieties of time structure and allocation? How do these patterns and possibilities vary by economic role, gender, life course, family conditions, ethnicity and race, location, and other placement factors? Time remains a basic resource for leisure action, one that not only varies widely but is one index of the possibility of self-determination.

### FAMILY AND COMMUNITY

**Leisure as a Context for Family Bonding.** If leisure is not just activity determined by and complementary to work, then is there some other critical relationship to the social system? The evident connection is to family and other immediate communities (Roberts 1970). Most leisure is in or around the home. The most common leisure companions are family and other close friends and intimates.

The basis of the first common wisdom was the series of community studies beginning in the 1930s (Lundberg, Komarovsky, and McNerny 1934; Lynd and Lynd 1956). Leisure was found to be a web of ordinary activity, mostly social interaction and tied to the institutions of the community from the family outward to status-based organizations. From this perspective, the later work of Cheek and Burch (1976) argued that the primary function of leisure is to provide a context for social bonding, especially that of family and ethnic community.

An anomaly in family leisure began a revision to the first consensus. The family context was reaffirmed and the centrality of the family to leisure supported (Kelly 1983). Despite the traditional focus on freedom as the primary defining theme of leisure, activity with major components of obligation was found to be most important to most adults. The major theme was that leisure was closely tied to central roles, not separate from them.

Leisure, then, is bound to both the roles and the developmental requirements of life (Rapoport and Rapoport 1976). In fact, from this perspective it may be quite central to life, not residual or

secondary at all. It is a primary setting for social bonding and expression as well as for human development. The implied issue, on the other hand, is the consequences for the nature of freedom and choice in leisure. No activity embedded in primary role relationships can ever be free of accompanying obligations and responsibilities (Kelly 1987a, chap. 6). Leisure might be more central but is also less pure and simple.

**Power and Self-determination in Leisure Roles.** First, there is the challenge posed by changing family patterns now that an unbroken marriage and family through the life course has become a minority probability. The most radical response and antithesis, however, begins with the suggestion that leisure, like other areas of life, has roles. That is, the expectations and power differentials that characterize family, work, and community roles are found in leisure as well. Currently the most salient source of this challenging antithesis is the focus on gender, especially from a feminist perspective (Henderson et al. 1996).

The critique calls for sociologists to go beneath the leisure rhetoric of freedom and self-expression to the realities of lives with limited power of self-determination. From this perspective, the history of the culture is characterized by male domination of women in profound and multifaceted ways that permeate every aspect of life (Deem 1986). Women have been repressed in where they are permitted to go and what they are allowed to do in leisure. Men have had the power to sexualize social contexts to objectify women and their bodies. Physical power and even violence have rendered many leisure times and places unsafe for women.

Even in the home, women's leisure is fundamentally different from that of men. It is usually women who are expected to do the work that makes "family leisure" possible. It is the "hidden work" of women that offers relative freedom to much of the leisure of men and children. Women both enable men's leisure and are leisure for men.

What, then, is the meaning of freedom and self-determination for any subordinate population segment? What about the poor, the racially and ethnically excluded, those cut off from opportunity in abandoned urban areas, and even many of the old? Again, many potential leisure venues in the inner city are unsafe, especially for children and

the old. The resources of time, money, access, and autonomy are evidently unevenly distributed in any society.

In this antithesis, the connections of leisure to nonwork roles and resources, especially family and community, and the positive evaluation of how leisure contributes to development through the life course are brought up against a critical model of society. Leisure may indeed be indissolubly tied to family and community, but in ways that reflect social divisions and dominations as well as expressive action.

**Leisure's Immediate Context: a New Agenda.** Leisure takes place in its small worlds, but also in the larger scale of the society. Further, its actualization is in the midst of real life. Research may be based on premises of systemic integration and the benefits of leisure as well as challenged by critiques reflecting ideologies of subjugation and alienation. A new agenda for research, however informed, should be directed toward the actual lived conditions of decisions and actions, relationships and roles. In such an agenda related to community and family, several themes are highlighted by critiques of the common wisdom:

First, the realities of leisure as a struggle for action and self-determination in the midst of acute differences in power and access to resources will receive more attention. Especially gender, race, and poverty will reconstitute research strategies and frames past the easy assumptions that leisure is equally free and beneficial for all. The realities of family instability and crisis as well as of community divisions and conflicts will be taken into account as the immediate communities of leisure are reformulated.

Second, underlying the new agenda is the theme of differential power, not only power to command resources but to determine the course of one's life and what is required of others. In the action of leisure, there is both a relative openness for action and modes of repression that stimulate submission and resistance.

Third, the pervasiveness of sexuality, gender roles, and sexual orientation throughout life in society will gain greater prominence in leisure sociology. One set of issues revolves around sexuality and related emotions to leisure. Simple models of explanation based solely on rational action

will be recognized as inadequate. Also, gender will be seen as negotiated identity and power of self-determination rather than a simple dichotomized category.

Fourth, the danger of leisure's becoming increasingly privatized, bound only to immediate communities and the small worlds of personal life construction, is a perspective that runs counter to the functional view of leisure as a context for social bonding. There may be a negative side to a focus on the family basis of leisure activity and meanings. As technologies increasingly make the home a center of varied entertainment, leisure could become more and more cut off from larger communities.

In general, leisure is surely not peripheral to the central concerns and relationships of life. That, however, does not lead simply to bonding without domination, to development without alienation, or to intimacy without conflict.

## AGING AND THE LIFE COURSE

**Continuity and Change in the Life Course.** The earliest common wisdom was simply that age indexed many kinds of leisure engagement. In a simple model, age was even referred to as a cause of decreased rates of participation. It was assumed that something decremental happened to people as they aged. The rates of decline varied according to activity: rapid for sports, especially team sports; more gradual for travel and community involvement. Attention given to those in their later years, generally their sixties and seventies but sometimes their fifties as well, suggested that such "disengagement" might even be functional. Perhaps older people needed to consolidate their activity and recognize their limitations.

The revised common wisdom began by recasting age as an index of multiple related changes rather than an independent variable. Further, the revised framework became the life course rather than linear age (Neugarten 1968). A number of themes emerged:

First, in the Kansas City study of adult life, normative disengagement was replaced with activity (Havighurst 1961). Instead of making a necessary or desirable withdrawal from activity, older people were found to revise their patterns and commitments in ways that fit their later life roles

and opportunities. Leisure was conceptualized as multidimensional in meaning as well as in forms. More recently, this approach has led to a discovery of the "active old," those before and in retirement who adopt lifestyles of engagement in a variety of leisure activities and relationships. Further, such engagement has been consistently found to be a major factor in life satisfaction (Cutler and Hendricks 1990).

Second, the model of inevitable decrement was challenged by research that failed to measure high correlations between age and functional ability. Rather, a model of aging that stressed *continuity* rather than loss and change was applied to leisure as well as other aspects of life (Atchley 1989). A return to earlier socialization studies provided a base for a revised model that identified lines of commitment rather than age-graded discontinuity. Especially the "core" of daily accessible activity and interaction remains central to time allocation through the life course (Kelly 1983).

Third, the life course also provided a perspective in which intersecting work, leisure, and family roles and opportunities were related to developmental changes (Rapoport and Rapoport 1976). Leisure is not a list of activities dwindling with age, but a social environment in which many critical issues of life may be worked out. Developing sexual identity for teens, expressing intimacy for those exploring and consolidating family commitments, reconstituting social contexts after midlife disruptions, and ensuring social integration in later years are all central requirements of the life course that are developed in leisure. Not only interests, but also significant identities are often found in leisure as well as in family and work (Gordon, Gaitz, and Scott 1976).

In the revised consensus, then, the life course with its interwoven work, family, and community roles was accepted as a valuable framework for analyzing both the continuities and the changes of leisure. Leisure was seen as tied not only to role sequences but also to developmental preoccupations. The life course was found to incorporate revisions and reorientations rather than being simply an inevitable downhill slide measured by participation rates in selected recreation pursuits.

**An Integrated View of Life . . . and Leisure.** The regular and predictable transitions of the life-course model, however, seem to gloss over many

of the realities of contemporary life. A majority of adults in their middle or later years have experienced at least one disrupting trauma in health, work, or the family that has required a fundamental reconstituting of roles and orientations (Kelly 1987b). Further, conditions are not the same for all persons in a social system. Race, gender, class, and ethnicity designate different life chances.

In this perspective of continuity and change in a metaphor of life as journey, a number of issues call for attention. First, salient differences in life conditions are more than variations in starting points for the journey. Rather, deprivation and denial are cumulative in ways that affect every dimension of life. Second, individuals come to define themselves in the actual circumstances of life, not in an abstracted concept. Identities, the concepts of the self that are central to what we believe is possible and probable in our lives, are developed in the realities of the life course. Third, the structures of the society, including access to institutional power, provide forceful contexts of opportunity and denial that shape both direction and resources for the journey.

In this revised life-course approach, leisure remains as a significant dimension, tied to family, work, education, community, and other elements of life. Changes in one may affect all the others. Leisure, then, is distinct from the product orientation of work and the intimate bonding of the family, and yet is connected to both.

#### **Leisure and the Life Course: New Agendas.**

From the perspective of the life course, research focusing on leisure now requires several revised issues. Among the most significant are:

First, leisure is woven through the life course. It is existential in a developmental sense. That is, leisure is action that involves *becoming*, action in which the actor becomes something more than before. For example, leisure is central to changing early socialization in the increased activity scheduling, electronic entertainment and interaction, and professional supervision of upper- and middle-class children. It is the main context for exploring sexuality and romance for teens and young adults.

Second, the developmental orientation of some leisure is highlighted by this perspective that recognizes lines of action as well as singular events and episodes. What has been termed "serious

leisure” by Robert Stebbins (1979) is activity in which there is considerable personal investment in skills and often in equipment and organization. Such investment places serious leisure in a central position in identity formation and expression. Leisure identities may provide continuity through the transitions and traumas of the life course. Yet, how women and men define themselves and take action toward redefinition has been a subject of speculation more than research.

Third, what is the place of leisure in the schema of life investments and commitments? Further, how do those investments differ according to the life conditions of men and women as they make their way through the shifting expectations and possibilities of the life course? Xavier Gallier (1988) presents a model of the life course that emphasizes disruptions rather than linear progress. In an irregular life journey, work, family, and leisure may rise and fall both in salience and in the “chunks” of time they are allocated. He proposes that education, production, and leisure become themes woven through life rather than discrete sequential periods.

### THE NATURE OF LEISURE

As already suggested, perspectives on the nature of leisure have changed in the modern period of scholarly attention from the 1930s to now. The change is not self-contained, but reflects shifts in theoretical paradigms as well as drawing from other disciplines, especially social psychology.

**Leisure as Free Time and Meaning.** Despite repeated references to Greek roots and especially Aristotle, the first accepted operational definition of leisure was that of time. Leisure did not require that all other role obligations be completed, but that the use of the time be more by choice than by requirement. How choice was to be measured was seldom addressed. Concurrently, international “time-budget” research quantified leisure as one type of activity that could be identified by its form (Szalai 1974). Leisure was assumed to be clearly distinguished from work, required maintenance, and family responsibilities.

The first consensus, although persisting in many research designs, did not endure long without amendment. To begin with, it was obvious that any activity might be required, an extension of

work or other roles. Further, even such simple terms as “choice” and “discretionary” implied that the actor’s definition of the situation might be crucial.

In the 1970s, the field claimed more attention from psychologists, who focused on attitudes rather than activities. Leisure was said to be defined by attitudes or a “state of mind” that included elements such as perceived freedom, intrinsic motivation, and a concentration on the experience rather than external ends (Neulinger 1974). Attention was directed toward meanings, but wholly in the actor rather than in definitions of the social context. Such psychological approaches were one salient influence on sociologists, who added at least three dimensions to the earlier time- and activity-known to common definitions.

First, in the 1950s, the Kansas City research (Havighurst 1961) along with the community studies tied leisure to social roles. The satisfactions anticipated in an activity involved meanings and relationships brought to the action context as well as what occurred in the time frame.

Second, the immediate experience might be the critical focus for leisure, but it occurs in particular environments that involve social learning, acquired skills and orientations (Csikszentmihalyi 1981), and interaction with components imported from other role relationships (Cheek and Burch 1976). Freedom is perceived, or not, in actual circumstances.

Third, although the dimension of freedom recurs in the literature, studies of experiences and activity engagements found that leisure seldom is monodimensional. The meanings, outcomes, motivations, and experiences themselves are multifaceted (Havighurst 1961; Kelly 1981).

Leisure, then, in the revised approaches is a more complex phenomenon than either the earlier sociologists or the psychologists proposed. In fact, the consensus broke down under the weight of multiple approaches that ranged from individualistic psychology to functional sociology, from presumably self-evident quantities of time to interpretive self-definitions and lines of action, and from discrete self-presentations (Goffman 1967) to actions embedded in life-course role sequences (Rapoport and Rapoport 1976).

**Revolt against the Abstract.** Antithetical themes came from several directions.

First, which is fundamental to accounting for life in society, the interpretive acts of the individual or the social context in which the action takes place (Giddens 1979)? Further, since the forms and symbols by which action is directed are learned and reinforced in the society, can action be prior to the context? The nature of leisure, then, is neither an acontextual nor a determined social role. Rather, it is actualized in processual action. And this process has continuities that extend beyond the immediate to personal development and the creation of significant communities (Kelly 1981).

Second, a number of critical analysts have raised questions about the positive cast usually given to leisure. Such positive approaches seem to presuppose resources, options, perspectives, and self-determination that are in fact unequally distributed in societies (Clarke and Critcher 1986). Do the unemployed and the poor have enough resources for discretion and choice to be meaningful concepts? Do histories of subjugation and life-defining limits for women in male-dominated societies make assumptions of self-determining action a sham? Such opportunity differences are most substantive in a market system of buying, renting, or otherwise acquiring resources. The real contexts of leisure are not voids of time and space, but are extensions of the structures of the society and ideologies of the culture. There is clearly an "other side" to leisure that includes many kinds of activity with destructive potential such as gambling, substance use, and sexual exploitation. There are also negative elements in other activities such as physical violence and racial stereotyping in sport, sexual violence in socializing, and even turning driving into a contest endangering others. All social forms of exploitation and exclusion are found in leisure (Rojek 1995; Kelly and Freysinger 1999).

Third, a consequence of this distorted and constricted context of leisure is alienation. Leisure is not entirely free, creative, authentic, and community-building activity. It may also be, perhaps at the same time, stultifying and alienating. It may separate rather than unite, narrow rather than expand, and entrap rather than free. It may, in short, be negative as well as positive. It is not a rarified ideal or a perfect experience. It is real life,

often struggle and conflict as well as development and expression.

The dialectic between expression and oppression that characterizes the rest of life in society is the reality of leisure as well. Being role-based in a stratified society means being limited, directed, and excluded. The contexts of any experience, however free and exhilarating, are the real culture and social system. The multiple meanings of leisure include separation as well as community, determination as well as creation, and routine as well as expression. The former simplicity of leisure as essentially a "good thing" becomes alloyed by situating it in the real society with all its forces, pressures, and conflicts.

**Leisure as a Dimension of Life.** The question, then, is what does such extension and critique do to any conceptualization of the nature of leisure? Leisure encompasses both the existential and the social. It has myriad forms, locales, social settings, and outcomes. Leisure is neither separated from social roles nor wholly determined by them. Leisure has developed amid conflict as well as social development, in division as well as integration, with control as well as freedom. It may involve acquiescence as well as resistance, alienation as well as authenticity, and preoccupation with self as well as commitment to community. Leisure, then, is multidimensional and cannot be characterized by any single or simple element.

A further issue is whether leisure is really a domain of life at all. Is leisure clearly distinguished from work, family, community, church, and school: or is it a dimension of action and interaction within them all? In the Preparation period, leisure is a social space for the exploration and development of sexual identities as well as working out the issues of peer identification and independence from parents and the past. It also stresses the theme of expression that is central to developing a sense of selfhood, of personal identity among emerging social roles. In the Establishment period, leisure adds the dimension of bonding to intimate others, especially in the formation and consolidation of the family. In the Third Age, leisure has meanings tied to both integration with significant other persons and maintenance of a sense of ability when some work and community roles are lost (Kelly 1987b). Leisure, then, might

be conceptualized as being woven into the intersecting role sequences of the life course rather than being a segregated realm of activity. Productivity is not limited to work, nor bonding to the family, nor learning and development to education, nor expression to leisure. Production, bonding and community, learning and development, and relative freedom and self-authenticating experience may all be found in any domain of life.

Yet there must also be distinguishing elements of leisure or it disappears into the ongoing round of life. Further, those elements should be significant in relation to central issues of life such as production and work, love and community, sexuality and gender, learning and development, emotion and involvement. Leisure should connect with the lived conditions of ordinary life rather than being an esoteric and precious idea to be actualized only in rare and elite conditions.

Leisure, then, may be more a dimension than a domain, more a theme than an identifiable realm (Kelly 1987a). That dimension is characterized by three elements: First, it is action in the inclusive sense of doing something, of being an intentioned and deliberate act. Such action is existential in producing an outcome with meaning to the actor. Second, this action is focused on the experience more than on the result. It is done primarily because of what occurs in the defined time and space. Third, leisure as a dimension of life is characterized by freedom more than by necessity. It is not required by any role, coercive power, or repressive ideology. Leisure is not detached from its social and cultural contexts, but is a dimension of relatively self-determined action within such contexts. Its meaning is not in its products as much as in the experience, not in its forms as much as in its expression.

**The Sociology of Leisure in the Future.** Leisure sociology, then, is not a closed book or a finalized product. Rather, central issues are currently being raised that promise to reform the field in its premises as well as conclusions. No common wisdom will go unchallenged, no consensus remain unchanged, and no theoretical formulation be above conflict. Yet, every challenge, every conflict, and every developing synthesis provides a new basis for at least one conclusion: Leisure is a significant dimension of life that calls for both disciplined and innovative attention. From this

perspective, a number of issues are likely to receive greater attention in the new century (Kelly and Freysinger 1999):

The first issue is the ascendancy of the market sector as the primary leisure resource provider, with an estimated 97 percent of total spending. In a global economy, leisure including tourism is attracting more investment capital with a significant bias toward upscale markets, big-ticket toys, sport as business and spectacle, and entertainment with multiple entrance fees. This bias combined with media images of a commodified "good life" may underlie trends away from skill-based physical activity and "serious" leisure with high time costs. Is there a fundamental conflict between developmental and consumptive leisure?

The second issue is the emergence of a global culture. The dominant direction of the dispersal is currently from the West through the mass media. However, as communication links and business and cultural contacts become more common, both the concepts and the practices of leisure in the West will become more affected by other cultures.

Third, a focus on gender is leading away from male-oriented "reasoned action" modes of leisure decisions and toward the significance of emotions and especially sexuality. Since all social interaction is gendered and most has deep dimensions of sexuality, leisure will be understood more as a multidimensional process rather than a singular choice. Leisure, then, is both contextual and contested.

Fourth, leisure becomes more a part of "ordinary life" rather than segregated activities with special designations. In a more fragmented social milieu, elements of leisure may be located in almost any social context. Further, if work itself loses familiar continuities, then leisure may become more central to identities and persistent lines of meaning as individuals seek to make sense of their lives.

Fifth, there will likely be concerns over many negative aspects of leisure. Will easy entertainment lessen personal investments in challenging leisure? Will available and affordable electronics damage the social fabric of associations and intimate relationships? Will leisure increasingly become privatized at the cost of community exploitative of the poor and powerless? Will leisure become

spectacular rather than engaging, violent rather than sharing, destructive of natural environments, and divisive rather than integrating?

The basic questions, of course, are those of the kind of society that is emerging and the kind of people who will live in it. It is clear, whatever is ahead, that leisure will be a significant dimension in a variety of forms and contexts.

(SEE ALSO: *Life Course, Gender, Social Class, Social Identity*)

## REFERENCES

- Atchley, Robert 1989 "A Continuity Theory of Normal Aging." *The Gerontologist* 29:183-190.
- Blauner, Robert 1964 *Alienation and Freedom: The Factory Worker and His Industry*. Chicago: University of Chicago Press.
- Cheek, Neil, and William Burch 1976 *The Social Organization of Leisure in Human Society*. New York: Harper and Row.
- Clarke, John and Chas Critcher 1986 *The Devil Makes Work: Leisure in Capitalist Britain*. Champaign: University of Illinois Press.
- Csikszentmihalyi, Mihaly 1981 "Leisure and Socialization." *Social Forces* 60:332-340.
- Cutler, Stephen, and Jon Hendricks 1990 "Leisure and Time Use across the Life Course." In R. Binstock and L. George, eds., *Handbook of Aging and the Social Sciences*, 3rd ed. New York: Academic Press.
- Deem, Rosemary 1986 *All Work and No Play: The Sociology of Women and Leisure*. Milton Keynes, United Kingdom: Open University Press.
- Dumazedier, Joffre 1967 *Toward a Society of Leisure*. New York: Free Press.
- Gallier, Xavier 1988 *La Deuxieme Carriere: Ages, Emplois, Retraite*. Paris: Editions du Seuil.
- Giddens, Anthony 1979 *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis*. Berkeley: University of California Press.
- Goffman, Erving 1967 *Interaction Ritual*. New York: Anchor Books.
- Gordon, Chad, C. Gaitz, and J. Scott 1976 "Leisure and Lives: Personal Expressivity across the Life Span." In R. Binstock and E. Shanas, eds., *Handbook of Aging and the Social Sciences*. New York: Van Nostrand Reinhold.
- Habermas, Jurgen 1975 *Legitimation Crisis*. Boston: Beacon Press.
- Havighurst, Robert 1961 "The Nature and Values of Meaningful Free-time Activity." In R. Kleemeier, ed., *Aging and Leisure*. New York: Oxford University Press.
- Henderson, Karla, M. Deborah Bialeschki, Susan Shaw, and Valeria Freysinger 1996 *Both Gains and Gaps: Feminist Perspectives on Women's Leisure*. State College, Pa.: Venture Publishing.
- Kelly, John R. 1981 "Leisure Interaction and the Social Dialectic." *Social Forces* 60:304-22.
- 1983 *Leisure Identities and Interactions*. London: Allen and Unwin.
- 1987a *Freedom to Be: a New Sociology of Leisure*. New York: Macmillan.
- 1987b *Peoria Winter: Styles and Resources in Later Life*. Lexington, Mass.: Lexington Books.
- Kelly, John R., and Valeria Freysinger 1999. *21st Century Leisure: Current Issues*. Boston: Allyn and Bacon.
- Lundberg, George, Mirra Komarovsky, and M. McInerney 1934. *Leisure: A Suburban Study*. New York: Columbia University Press.
- Lynd, Helen, and Robert Lynd 1956 *Middletown*. New York: Harcourt Brace.
- Marx, Karl 1970 *The Economic and Philosophical Manuscripts of 1844*. London: Lawrence and Wishart.
- Neugarten, Bernice 1968 *Middle Age and Aging*. Chicago: University of Chicago Press.
- Neulinger, John 1974 *The Psychology of Leisure*. Springfield, Ill.: C. C. Thomas.
- Parker, Stanley 1971 *The Future of Work and Leisure*. New York: Praeger.
- Rapoport, Rhona, and Robert Rapoport 1976 *Leisure and the Family Life Cycle*. London: Routledge.
- Roberts, Kenneth 1970 *Leisure*. London: Longmans.
- Robinson, John, and Geoffrey Godbey 1997 *Time for Life: The Surprising Ways Americans Use Their Time*. University Park: Pennsylvania University Press.
- Rojek, Chris 1985 *Capitalism and Leisure Theory*. London: Tavistock.
- 1995 *Decentering Leisure*. London and Thousand Oaks, Calif.: Sage.
- Schor, Juliet 1991 *The Overworked American*. New York: Basic Books.
- Stebbins, Robert 1979 *Amateurs: On the Margin between Work and Leisure*. Beverly Hills, Calif.: Sage.
- Szalai, Alexander 1974 *The Use of Time: Daily Activities of Urban and Suburban Populations in Twelve Countries*. The Hague: Mouton.
- Veblen, Thorstein (1899) 1953 *The Theory of the Leisure Class*. New York: New American Library.

Wilensky, Harold 1960 "Work, Careers, and Social Integration." *International Social Science Journal* 12:543-560.

Wilson, John 1980 "Sociology of Leisure." *Annual Review of Sociology* 6:21-40.

JOHN R. KELLY

## LESBIANISM

See Alternative Life Styles; Sexual Orientation.

## LEVELS OF ANALYSIS

Determining the level of analysis is usually straightforward, but whether to, or how to, draw inferences from one level of analysis to another is a difficult problem for which there is no general solution. The cases used as the units in an analysis determine the level of analysis. These cases may be quite varied, for example, countries, political parties, advertisements, families, or individuals. Thus, analysis may occur at the individual level, family level, advertisement level, and so forth.

The types of variables used at any one level of analysis, however, may be quite different. As an example, in studying the determinants of individuals' attitudes toward public education, the individuals (the units of analysis) may be described in terms of their sex and race (measures of individual properties), whether they attended a public or private college, and the region of the country in which they reside (measures of the collectives to which they belong). The analysis in this example is at the individual level because the cases used are individuals who are described in terms of individual properties and the properties of the collectives to which they belong.

This article focuses on (1) the types of variables used to describe the properties of collectives and members and the use of these variables at different levels of analysis; (2) problems that arise when using relationships at one level of analysis to make inferences about relationships at another level of analysis; (3) a brief discussion of a statistical model that explicates these problems; (4) some useful data analytic techniques to use when data at two or more levels of analysis are available, and (5)

proposed solutions or partial solutions to the problem of cross level inference when data at only a single level of analysis are available.

### TYPES OF VARIABLES USED TO DESCRIBE COLLECTIVES AND MEMBERS

Lazarsfeld and Menzel (1969) propose a typology of the kinds of properties (variables) that describe "collectives" and "members." For example, in discussing the properties of collectives, Lazarsfeld and Menzel distinguish between analytical, structural, and global properties.

*Analytical properties* are obtained by performing some mathematical operation upon some property of each *single* member. These properties are typically referred to as aggregate variables. Examples are the percentage of blacks in cities, the sex ratio for different counties, and the Gini Index as a measure of inequality of incomes in organizations.

*Structural properties* of collectives are obtained by performing some operation on data about the relations of each member to some or all of the others. Such measures are common in network analysis. Friendship density, for example, could be defined as the relative number of pairs of members of a collective who are directly connected by friendship ties. Since the total number of potential ties in a group with  $N$  members is  $N(N-1)/2$ , one measure of density is the total number of ties divided by this number.

*Global properties* of collectives do not use information about the properties of individual members either singly or in relationship to one another. Having a democratic or nondemocratic form of government is a global property of collectives. Being a private rather than a public school is a global property of a school. The proportion of gross national product (GNP) spent on education is a global property of countries.

Thus, variables that describe collectives can be based on summary data concerning single members of those collectives, the relationships of members to other members, or some global characteristic of the collective itself. Turning to variables that describe the properties of members of collectives, there are four major types: absolute, relational, comparative, and contextual.



*Absolute properties* are obtained without making use either of information about the characteristics of the collective or of information about the relationships of the member being described to other members. Thus, sex, level of education, and income are absolute properties of individuals.

*Relational properties* of members are computed from information about the substantive relationships between them and other members. For example, the number of friends an individual has at school or the number of family members is a property of the individual based on other members in the collective.

*Comparative properties* characterize a member by a comparison between his or her value on some (absolute or relational) property and the distribution of this property over the entire collective to which the person belongs. A person's class rank and birth order are comparative properties.

*Contextual properties* describe members by a property of the collective to which they belong. For example, being from a densely populated census tract or a school with a certain percentage of nonwhite students is a contextual property describing the context in which the member acts. Contextual properties are characteristics of collectives that are applied to members.

Contextual variables remind us that the level of analysis is determined by the cases used as the units of analysis, not by the level of the phenomena described by a particular variable. Thus, all the variables that describe a collective may be used at the individual level of analysis as well as those that describe individual properties; for example, a person's attitude may be predicted on the basis of the percentage of blacks in the person's school (contextual/analytic), whether or not the school is private or public (contextual/global), the density of friendships at the school (contextual/structural), the person's sex (absolute), his or her class standing (comparative), and the number of friendship choices he or she receives (relational).

### INFERENCES FROM ONE LEVEL TO ANOTHER

The section above describes abstractly two different levels of analysis, the collective (or aggregate) and the member (or individual). Sociologists typically distinguish levels concretely depending on

the units of analysis; for example, the units may be schools, advertisements, children's stories, or riots. To make inferences from relationships discovered at one level of analysis to relationships at another level is not logically valid, and sociologists have labeled such inferences "fallacies." Still, at times one may be able to argue for the reasonableness of such inferences. These arguments may be based on statistical considerations (Achen and Shively 1995; Duncan and Davis 1953; Goodman 1953, 1959; King 1997) or on rationales that closely tie relationships at one level with those at another (Durkheim [1897] 1966; Dornbusch and Hickman 1959).

*Disaggregative fallacies* (often called ecological fallacies) are the classic case of cross level fallacies. Robinson (1950) brought them to the attention of sociologists. He cites two cases of cross level inferences, both of which involve making inferences about relationships at the individual level based on relationships discovered at the aggregate level. Robinson noted that the Pearson product moment correlation between the percent black and the percent illiterate in 1930 for the Census Bureau's nine geographical divisions was 0.95 and for states it was 0.77, while the correlation (measured by phi) on the individual level between being illiterate or not and being black or not was only 0.20. The relationship between percent illiterate and percent foreign-born was negative for regions and states (-0.62 and -0.53, respectively), while the relationship between being illiterate and being foreign-born at the individual level was positive (0.12).

Robinson demonstrated that relationships at one level of analysis do not have to be the same as those at another level. To assume that they must be the same or even that they must be quite similar is a logical fallacy.

*Aggregative fallacies* occur in the opposite direction, that is, when one assumes that relationships existing at the individual level must exist at the aggregate level. Robinson's results show that the positive relationship between being foreign born and being illiterate at the individual level may not be mirrored at the state level.

*Universal fallacies* (Alker 1969) occur when researchers assume that relationships based on the total population must be true for subsamples of the whole. It may be true, for example, that the

relationship between population density and the crime rates of cities for all cities in the United States is not the same for southern cities or for cities with a population of over one million. Here, the fallacy is to assume that a relationship based on the total population must hold for selected subpopulations.

Similarly, one might commit a *selective fallacy* (Alker 1969) by assuming that relationships based on a particular sample of cities must hold for all cities. If the selected cities are a random sample of cities, this is a problem of statistical inference, but if they are selected on some other basis (e.g., size), then making inferences to all cities is a selective fallacy.

*Cross-modality fallacies* occur when the inference is from one distinct type of unit to another distinct type of unit. A cross-modality fallacy occurs when trends in advertisement content are used to make inferences about trends in the attitudes of individuals, or designs on pottery are used to infer the level of need for achievement in different cultures. (Aggregative and disaggregative fallacies are cross-modality fallacies because groups and individuals are distinct units. But these fallacies have traditionally been classified separately.)

*Cross-sectional fallacies* occur when one makes inferences from cross-sectional relationships (relationships based on units of analysis from a single point in time) to longitudinal relationships. For example, if unemployment rates and crime rates are positively related at the city level, this fallacy is committed by inferring that increases over time in the unemployment rate are related to increases in the crime rate over time.

*Longitudinal fallacies* occur when one makes inferences from longitudinal relationships (relationships based on units of analysis across time units) to cross-sectional ones, that is inferring from a relationship between unemployment rates and crime rates over time to the relationship between these rates over units such as cities, counties, or states at a given point in time.

In all of their varied manifestations, cross-level inferences are not logically valid inferences (Skrms 1975). That is, relationships on one level of analysis are not necessarily the same as those on another level. They may not even be similar. In the final section of this article, however, we note that data at

one level of analysis may serve as evidence for relationships at another level of analysis even if they do not strictly imply such a relationship.

**STATISTICAL ANALYSIS OF  
DISAGGREGATIVE AND AGGREGATIVE  
INFERENCES**

This section presents the results of a mathematical demonstration of why disaggregative and aggregative inferences are fallacies, that is, why results at the aggregate level are not necessarily mirrored at the individual level. The derivation of this model is not shown here but may be found in several sources (Duncan et al. 1961; Alker 1969; Hannan 1971; and Robinson 1950). Readers who prefer can skip to the next section without loss of continuity.

The individual level or total correlation ( $r^t_{xy}$ ) between two variables ( $X$  and  $Y$ ) can be written as a function of the correlation between group means (the aggregate level correlation:  $r^b_{xy}$ ), the correlation of individual scores within groups (a weighted average of the correlations within each of the groups:  $r^w_{xy}$ ), and the correlation ratios for the two variables,  $X$  and  $Y$ . The correlation ratio is the ratio of the variance between groups (the variance of the group means:  $V^b_x$ ) to the total variance (variance of the individual scores:  $V^t_x$ ). Thus, the individual level or total correlation can be written

$$r^t_{xy} = r^w_{xy} \sqrt{1 - \left(\frac{V^b_x}{V^t_x}\right)^2} \sqrt{1 - \left(\frac{V^b_y}{V^t_y}\right)^2} + r^b_{xy} \left(\frac{V^b_x}{V^t_x}\right) \left(\frac{V^b_y}{V^t_y}\right) \tag{1}$$

Similarly, the individual level regression coefficient can be written as a function of the within-group regression coefficient, the group level (between group) regression coefficient, and the correlation ratio for variable  $X$  (equation 2):

$$b^t_{yx} = b^w_{yx} + \left(\frac{V^b_x}{V^t_x}\right)^2 (b^b_{yx} - b^w_{yx}) \tag{2}$$

It is a simple algebraic exercise to derive formulas for  $r^b_{xy}$  and  $b^b_{yx}$  (the aggregate level correlation and regression coefficients) in terms of correlation ratios, and correlation and regression coefficients at other levels.

These formulas clearly demonstrate why one cannot use the ecological or group level correlation or regression coefficients to estimate individual level relationships: The individual level relationships are a function of group level relationships, within-group level relationships, and correlation ratios. This approach can be extended to include other levels of analysis, for example, individuals on one level, counties on another, states on another, and time as yet another level (see, e.g., Alker 1969; Duncan et al. 1961).

### SEPARATING AGGREGATE LEVEL AND INDIVIDUAL LEVEL EFFECTS

Obtaining data at the different levels of analysis solves the problem of inferring relationships from one level to another. Researchers in this situation know the relationship at both levels for their data. Such data also provide additional information about the relationships at different levels of analysis.

O'Brien (1998) shows that when individual level data are aggregated to create summary measures at the aggregate level (e.g., means or rates for aggregates), then it is possible to estimate the reliability of the aggregate level measures. Further, when two or more of the aggregate level measures are based on samples of the same respondents within each aggregate, correlated errors between aggregate level measures are likely to occur. This correlated error can be measured and the aggregate level relationships can be corrected for this spurious correlation as well as for unreliability in the aggregate level measures.

In some situations one may want to argue that the best measures of individual level "effects" (in a causal sense) are provided by analyses at the individual level that include as predictors relevant individual properties and the properties of the collective to which the individuals belong (contextual variables). Estimates of these individual level relationships are then "controlled" for group level effects (Alwin and Otto 1977).

Since the relationship between group level means may reflect nothing more than the relationship between variables at the individual level, it has been suggested that the best estimate of group level "effects" compares the regression coefficient for the group means and for the individual scores. Lincoln and Zeitz (1980) show how this may be

done in a single regression equation while at the same time controlling for other relevant variables. For both of these techniques to work, some stringent assumptions must be met, including assumptions of no measurement error in the independent variables and of a common within-group regression coefficient. Most importantly, these techniques depend upon having data from different levels of analysis.

The introduction of Hierarchical Linear Models (HLM) provides a flexible method for examining the relationship of individual level variables to group level variables (Bryk and Raudenbush 1992). These models allow for the relationships between individual level variables to vary within different groups and for differences between these relationships to be predicted by group level characteristics. It might be the case, for example, that the relationship between socioeconomic standing and student achievement differs depending upon class size and whether the school is private or public. These models allow for the prediction of different relationships for different individuals based on group level characteristics. These models may be extended to several levels of analysis in which members are nested within collectives.

### SOLUTIONS TO THE PROBLEM OF INFERENCES FROM ONE LEVEL TO ANOTHER

Even when data at only one level of analysis are available, cross level inferences can be and often are made by sociologists. There is no absolute stricture against making such inferences, but when researchers make them, they need to do so with some awareness of their limitations.

While both Duncan and Davis (1953) and Goodman (1953, 1959) maintain that it is generally inappropriate to use aggregate level (ecological) relationships to make inferences about individual level relationships, they each propose strategies that set bounds on the possible relationships that could exist at one level of analysis given relationships that exist at the other. The bounds are designed for use with aggregated data (analytic measures), and in some circumstances these techniques are useful.

Goodman (1953, 1959) suggested a technique called "ecological regression," which became the

most widely used method for making inferences from aggregate level data to individual level relationships when group level variables are “analytical properties of collectives.” Goodman’s method has been extended by a number of authors and summarized in the work of Achen and Shiveley (1995). King (1997) has proposed a statistical “solution to the problem of ecological inference,” but the success of that solution is controversial (Freedman et al. 1998).

“Theory” may also allow one to make cross level inferences. For example, Dornbusch and Hickman (1959) tested Riesman et al.’s contention (1950) that other-directedness in individuals declined in the United States during the first half of the twentieth century. They obviously could not interview individuals throughout the first half of the century, so they turned to advertisements in a women’s magazine (*Ladies Home Journal*, 1890–1956) to examine whether these ads increasingly used themes of other-directedness. Their units of analysis were advertisements, but they explicitly stated that they wanted to make inferences about changes in the other-directedness of individuals. Is this justified? The answer is no, on strictly logical grounds, and this constitutes a cross-modality fallacy. Certainly changes in the contents of advertisements do not demonstrate changes in individuals’ personalities. But Dornbusch and Hickman (1959) convincingly argue that advertisements (in this case) are likely to reflect aspects of other-directedness in the targets of the advertisements (individuals). They recognize the need for other tests of this hypothesis, using other types of data.

Perhaps the classic case in sociology of an analysis built on the ecological fallacy is Emile Durkheim’s analysis of suicide ([1897] 1966). One factor that Durkheim sees as “protecting” individuals from suicide is social integration. When individuals are married, have children, are members of a church that provides a high degree of social integration (e.g., Catholic rather than Protestant), or live at a time when their countries are in crisis (e.g., a war or electoral crisis), they are seen as more integrated into social, religious, and political society and less likely to commit suicide. Much of the data available to Durkheim did not allow an analysis on the individual level. There was no “suicide registry” with detailed data on the sex, religion, family status, and so forth of those committing suicide. There were, however, census data

on the proportion of Catholics, the proportion married, and the average family size in different regions. Other sources could be used to ascertain the rate of suicide for different regions. Using these data, Durkheim showed that Catholic countries had lower suicide rates than Protestant countries, and that, France and Germany, Catholic cantons exhibited lower suicide rates than Protestant cantons. Further, departments in France with larger average family sizes had lower suicide rates, and the suicide rate was lower during the months of electoral crises in France than during comparable months of the previous or following year. He combined this evidence with other evidence dealing with individuals (e.g., suicide rates for married versus unmarried men), and it was all consistent with his theory of suicide and social integration.

One could dismiss these aggregate level relationships by arguing that perhaps in Protestant countries those of other religions kill themselves at such a high rate that the suicide rates are higher in Protestant countries than in Catholic countries (and similarly in Protestant cantons in France and Germany). Isn’t it possible that in departments with relatively small average family sizes there is a tendency for those in large-size families to kill themselves relatively more often? This would create a relationship at the aggregate level (department level) in which smaller average family size is associated with higher rates of suicide. It is possible, because relationships at one level of analysis are not necessarily mirrored at another level of analysis. But Durkheim’s results are not easily dismissed.

Strict logic does not justify cross level inferences. But strict logic is not the only rational way to justify inferences. If a series of diverse relationships that are predicted to hold at the individual level are found at the aggregate level, they do not prove that the same relationships would be found at the individual level, but they are not irrelevant. It is incumbent on the critic of a study such as Durkheim’s to give a series of alternative explanations explaining why the relationships at the aggregate level should differ from those at the individual level. If the alternative explanations are not very convincing or parsimonious, researchers are likely to find Durkheim’s evidence persuasive. To the extent that social scientists are convinced that a set of advertisements is designed to appeal to motivations in their target population, that the target

population of the magazine in which the advertisements appears represents the population of interest, and so on, they will find Dornbusch and Hickman's cross level inference persuasive. That a relationship at one level of analysis does not imply a relationship at another level of analysis does not mean that it cannot be used, along with other evidence, to help infer a relationship at another level of analysis.

Persuasion is a matter of degree and is subject to change. Sociologists would want to examine additional studies based on, for instance, other populations, modalities, and periods. These data might strengthen cross level inferences.

## REFERENCES

- Achen, Christopher H., and W. Phillips Shively 1995 *Cross-Level Inference*. Chicago: University of Chicago Press.
- Alker, Hayward R. 1969 "A Typology of Ecological Fallacies." In Mattei Dogan and Stein Rokkan, eds., *Quantitative Ecological Analysis in the Social Sciences*. Cambridge, Mass.: MIT Press.
- Alwin, Duane F., and Luther B. Otto 1977 "High School Context Effects on Aspirations." *Sociology of Education* 50:259-272.
- Bryk, Anthony S., and Stephen W. Raudenbush 1992 *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage Publications.
- Dornbusch, Sanford M., and Lauren C. Hickman 1959 "Other-Directedness in Consumer-Goods Advertising: A Test of Riesman's Historical Theory." *Social Forces* 38:99-102.
- Duncan, Otis D., Ray P. Cuzzort, and Beverly Duncan 1961 *Statistical Geography: Problems in Analyzing Areal Data*. Glencoe, Ill.: Free Press.
- Duncan, Otis D., and Beverly Davis 1953 "An Alternative to Ecological Correlation." *American Sociological Review* 18:665-666.
- Durkheim, Emile (1897) 1966 *Suicide*. New York: Free Press.
- Freedman, D. A., S. P. Klein, M. Ostland, and M. R. Roberts 1998 "A Solution to the Ecological Inference Problem (Book Review)." *Journal of the American Statistical Association* 93:1518-1522.
- Goodman, Leo A. 1953 "Ecological Regression and Behavior of Individuals." *American Sociological Review* 18:663-664.
- 1959 "Some Alternatives to Ecological Correlation." *American Journal of Sociology* 64:610-625.
- Hannan, Michael T. 1971 *Aggregation and Disaggregation in Sociology*. Lexington, Mass.: Lexington Books.
- , and Leigh Burstein 1974 "Estimation from Grouped Observations." *American Sociological Review* 39:374-392.
- King, Gary 1997 *A Solution to the Ecological Inference Problem*. Princeton, N.J.: Princeton University Press.
- Lazarsfeld, Paul F., and Herbert Menzel 1969 "On the Relation Between Individual and Collective Properties." In Amitai Etzioni, ed., *A Sociological Reader on Complex Organizations*. New York: Holt, Rinehart, and Winston.
- Lincoln, James R., and Gerald Zeitz 1980 "Organizational Properties from Aggregate Data: Separating Individual and Structural Effects." *American Sociological Review* 45:391-408.
- O'Brien, Robert M. 1998. "Correcting Measures of Relationship Between Aggregate-Level Variables for Both Unreliability and Correlated Errors: An Empirical Example." *Social Science Research* 27:218-234.
- Reisman, David, Nathan Glazer, and Reuel Denney 1950 *The Lonely Crowd*. New Haven, Conn.: Yale University Press.
- Robinson, William S. 1950 "Ecological Correlations and the Behavior of Individuals." *American Sociological Review* 15:351-357.
- Skyrms, Brian 1975 *Choice and Chance*. Encino, Calif.: Dickenson.

ROBERT M. O'BRIEN

## LIBERALISM/CONSERVATISM

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

"Is (or was) Blank a liberal?" The precise reply to this question inevitably begins with a throat-clearing preface such as, "It all depends on the period you have in mind—and the place. Are you speaking of someone in nineteenth-century England, the United States during the Franklin Roosevelt New Deal days, contemporary Great Britain, continental Europe, or contemporary U.S.A.?"

For Americans nurtured on the "liberal" tradition of Franklin Delano Roosevelt and the Democratic party, the significance of the *L* word was quite clear. The private business establishment,

left to its own devices, had brought about the economic collapse of 1929. "Rugged individualism" had demonstrated its inadequacies even for many rugged individualists themselves. Almost two-thirds of a century after that collapse, the Great American Depression still retains the power to evoke terrifying memories of mass unemployment, bank failures, small business bankruptcies, soup kitchens, and popular ditties like "Brother can you spare a dime?" These were all nostalgic but chilling reminiscences of the desolation to which "nonliberal," or "conservative," social, economic, and political policies had led. There seems to persist an enormous reluctance to release this era and its memories to the historians as just another noteworthy episode in American history like the War of 1812 or the Panic of 1873.

For other Americans, some of whom matured perhaps too rapidly during the late 1960s and whose memories were filled with visions of the war in Vietnam, the battles for Civil Rights, and the turmoil in urban ghettos, things were much less clear. Stalwarts of liberalism like Hubert Humphrey (senator, vice president, and unsuccessful Democratic candidate for President) seemed to become part of the very establishment toward which hostility was directed. Government, instead of representing a liberating force, increasingly was seen as the source of existing difficulties. Liberalism, for many, was no longer the solution; it had become part of the problem.

Liberalism has often been identified with the struggle to free individuals from the confining embraces of other persons, institutions, or even habitual ways of behaving. It is seen by its advocates as a liberating orientation. Thoughts that, in various times and places, have been called liberal, all seem to have as their common denominator this fundamental notion of freedom for individual human beings.

It was during the period between the Reformation and the French Revolution that a new social class emerged. Political control by a landowning aristocracy was challenged by this new class whose power lay only in the control of movable capital. It consisted of such people as bankers, traders, and manufacturers. Science began to replace religion as the source of ultimate authority; contract replaced status as the legal foundation of society. The ideas of individual initiative and individual

control became the basis for a new philosophy: liberalism (Laski 1962, p. 11).

Liberalism began as the champion of freedom and the foe of privilege derived from inherited class position. But freedom was not fought for on behalf of everyone in society. The constituency of liberalism consisted of those who had property to defend. Its supporters sought to limit the range of political authority and tried to develop a system of fundamental rights not subject to invasion by the state (Laski 1962, p. 13). Most of the population (e.g., factory and agricultural workers) were not initially included in the concerns of liberalism. But redefinitions of governmental concerns and authority inevitably had consequences for all members of society. Thoughtful economists and others soon realized this.

Thus, from the perspective of the closing years of twentieth-century America, it is easy to forget that classical economists like Adam Smith and David Ricardo, far from being simply reflex ideologists for the business establishment, were deeply concerned about liberty for individuals and saw their classical economic doctrines as the best way to insure it. In this sense, many contemporary Americans could legitimately view them as representatives of the liberal, rather than the conservative, genre. The *free enterprise* game was originally a liberal game.

For many participants in this great game of free enterprise, it became increasingly more obvious that the game was "fixed," that somehow participants did not emerge with either their liberty or their pocketbooks intact. To insure even a modicum of freedom and liberty for individuals, it became necessary to provide them with a helping hand from a source outside the marketplace. This source, a powerful and beneficent outsider, the national government, could help monitor the rules and to some extent the play of the economic game. It *could* provide, for heavy losers or potential losers, social benefits in the form of such consolations as unemployment insurance, old age pensions, insured bank accounts, and coordinated measures to combat environmental pollution. It *could* help protect consumers from the harmful effects of adulterated food and drug products; it *could* protect workers from the hazards of unregulated workplaces. In short, the liberal ethos in twentieth-century America incorporated as one of its tenets

reliance on government action to protect the liberty of individuals.

The conservative ethos, on the other hand, during the same period has been characterized by a hands-off posture with respect to many aspects of government policy. This constitutes a fascinating reversal of traditional conservatism. The shades of Edmund Burke and other historical conservative spirits might well cringe at the characterization of their philosophy as “less government.”

For some observers, this liberal-conservative contrast seems to be based on different psychological sets, or frames, of mind. Thus, we have been told that liberals are more hospitable to change, more willing to reexamine institutions and established practices in the light of new problems and needs. From this perspective, conservatives are those who appeal to the experience of the past but do not learn from it. Liberals are presented as being less reverent of the status quo, more venturesome in the realm of ideas, and much more optimistic about the possibilities available through exploration and discovery (Girvetz 1966).

Willingness to change becomes meaningful only when viewed against the nature of the status quo. When “liberal” governments adopt measures to provide for persons who are aged, ill, unemployed, or handicapped (as they have in some Scandinavian countries), to be “liberal” may mean to *retain* the status quo. To eliminate the measures—to engage in change—may be precisely what Western conservatives might demand.

For many observers of the dramatic events occurring at the end of the 1980s and in the early 1990s in Eastern Europe, members of communist parties who struggled to maintain their power in government were labeled “conservatives.” Those fighting for *free enterprise* and other capitalist-like changes were called “liberal reformers.” Communists for many years had been viewed by Western conservatives as the ultimate enemy; liberals were characteristically referred to by many of these conservatives as thinly disguised “Reds.” In a dramatic and even comic reversal, dedicated Stalinists were now described as “conservatives.”

At one time or another a wide variety of ideas have been called conservative. There have been efforts, however, to reduce the essence of these ideas to a limited number of more or less well

defined notions. Clinton Rossiter and Russell Kirk, two ardent defenders of the conservative faith, once prepared a summary that received widespread approval. Although not all conservatives would necessarily agree with everything appearing in their summary, “it would be exceedingly difficult to find a conservative who did not agree with a good deal of what Professors Kirk and Rossiter impute to their tradition” (Witonski 1971, pp. 34–35).

The imputation begins with an assumption about the more or less immutable character of human nature. Behind the curtain of civilized behavior there exists, in human beings, wickedness, unreason, and an urge to engage in violence. In addition, a great deal of emphasis is placed upon the conviction that people are not naturally equal in most qualities of mind, body, and spirit. Liberty is more important than social equality. This leads to the conclusion that society must always have its classes. It is futile to level or eliminate them. Accordingly, societies will always require ruling aristocracies; efforts to have majority rule will lead to errors and potential tyranny.

Human beings, this conservative ethos insists, are not born with rights. These are earned as a result of duties performed. Service, effort, obedience, cultivation of virtue and self-restraint are the price of rights. Of primary importance to liberty, order, and progress, is the institution of private property. Inherited institutions, values, symbols, and rituals are indispensable and even sacred. Human reason is subject to error and severely limited; the surest guide to wisdom and virtue is historical experience.

Beyond this, fundamental to conservatism is the belief that all political problems are fundamentally religious in nature. Narrow rationality cannot, in itself, satisfy human needs. Both society and conscience are ruled by divine intent. Conservatives, it is asserted, have an affection for the traditional life; others (presumably liberals) favor narrow uniformity, egalitarianism, and pursuit of utilitarian goals. The only true equality is moral equality; civilized society requires orders and classes. Property and freedom are inseparable. If property is separated from private possession, liberty disappears.

This basic conservative creed goes on to say that human beings have anarchic impulses and are governed more by emotion than by reason. It

therefore becomes necessary to place controls on human appetite. The correct method of accomplishing this is through tradition and sound “prejudice.”

Finally, the proper instrument for social change is Providence. Some change is necessary from time to time to conserve society. It is like the perpetual renewal of the human body. But this change must always be accomplished slowly and with an awareness of the direction in which Providence is moving social forces (Witonski 1971, pp. 32–34).

Undergirding the entire structure of this conservative doctrine is a more or less explicit version of the *chain of being*. This metaphor is designed to express the enormous extent, variety, order, and unity of “God’s Creation.” The chain stretches from the foot of God’s throne to the meanest of inanimate objects. Every speck of creation is a link in this chain, and every link except those at the two extremities is simultaneously bigger and smaller than another; there can be no gap (Tillyard 1959). The classic examination of the history of this idea was written by Arthur S. Lovejoy (1960).

Every category of things excels at *something*. Plants are higher than stones, but stones are stronger and more durable. Animals (“Beasts”) are above plants, but plants can assimilate nourishment better. Human beings are above beasts but inferior to them in physical energy and desires.

All this suggests a sort of interdependency among all objects and living creatures. Central to the idea of the chain is the concept that every object, animal or person, is part of an all-encompassing whole. Basic to the metaphor is an implicit, if not always explicit, view of the universe as an organism (strong traces of this continue to be found in some formulations of contemporary systems theory).

Pushed to its logical conclusion, this suggests that all creatures and objects are equally important. Every existing part of the cosmic organism might well be seen as necessary (perhaps in some unknown way) for the survival of the whole. This, in turn implies that, although they do different things, all human beings, for example, are equally necessary. This logic could lead to dangerous subversive doctrines about social equality. Another feature was required to make the chain-of-being notion acceptable to those searching for reasons

to justify existing inequalities in human societies. This was found in the *primacy doctrine*, the idea that, within each category of objects and creatures, there is one above the others, a *primate*. The eagle is first among birds; the whale or dolphin, among “fishes”; the rose, among flowers; the fire, among elements; the lion or elephant, among beasts; and, naturally, the emperor, among men (Tillyard 1959, pp. 29–30).

More recently, the distinguished conservative sociologist Robert Nisbet has insisted that conservatism is simply one of three major political ideologies of the eighteenth and nineteenth centuries. The other two are *liberalism* and *socialism*. Interestingly, an ideology for him, in addition to having a “reasonably” coherent body of ideas, has a *power base* that makes possible a victory for the body of ideas. It extends over a period of time and has “major” advocates or spokespersons as well as a “respectable” degree of institutionalization and charismatic figures. For conservatism, these figures would include people like Edmund Burke, Benjamin Disraeli, and Winston Churchill. Liberals have their own counterparts to these. The philosophical substance of conservatism dates from 1790 with the publication of Edmund Burke’s *Reflections on the Revolution in France* (Nisbet 1986; Burke 1855).

There is a fascinating contradiction to be found in conservative doctrine. Methodologically, Nisbet offers it as a champion of historical method. This he approves of as an alternative to the liberal utilitarianism of Jeremy Bentham that is “soulless,” “mechanical,” and even “inhuman.” Bentham’s doctrine idolizes pure reason, but human beings require a different mode of thought, one based on feelings, emotions, and long experience, as well as on pure logic.

In attacking liberal utilitarianism, Nisbet is attacking a doctrine essentially abandoned by twentieth-century American liberals. Sociologist L. T. Hobhouse must be credited with having made the most serious effort to reformulate liberal doctrine.

Utilitarianism, as fashioned by Jeremy Bentham and his followers, was the visible core of nineteenth-century liberalism. It has been defined as “nothing but an attempt to apply the principles of Newton to the affairs of politics and morals” (Halevy 1972, p. 6). The principle of utility, the notion that



every possible action of human beings is either taken or not depending upon whether it is seen as resulting in either pleasure or pain, was the basis for an “objective science” of behavior modeled on the physical sciences. As such, in common with other alleged sciences, it was perhaps congenitally soulless, somewhat mechanical, and potentially inhuman. Using a “rational” (but not empirically derived) model, it announced that, by pursuing his or her own pleasure and avoiding pain, each person would maximize happiness for everyone. Society was simply a collection of separate individuals operating in their individual self-interests. Government action, or “interference,” constituted a disservice to these individuals and should be rejected on “scientific” grounds.

Hobhouse took issue with this view but modestly presented John Stuart Mill as the transition figure between the old and new liberalisms. Mill, reared on Benthamite doctrine, continually brought it into contact with fresh experience and new trains of thought. As a result, Mill is, “the easiest person in the world to convict of inconsistency, incompleteness, and lack of rounded system. Hence also his work will survive the death of many consistent, complete, and perfectly rounded systems” (Hobhouse 1964, p. 58).

Hobhouse noted that, although the life of society is, ultimately, the life of individuals as they act upon each other, the lives of individuals would be quite different if they were separated from society. He stressed the fact that collective social action does not necessarily involve coercion or restraint. The state is simply one of many forms of human association, a form to which individuals owe much more of their personal security and freedom than most people recognize. “The value of a site in London,” he pointed out, “is something due essentially to London, not to the landlord. More accurately, a part of it is due to London, a part to the British empire, a part, perhaps we should say, to Western civilization” (Hobhouse 1964, p. 100). “Democracy,” he tells us, “is not founded merely on the right or the private interest of the individual. . . It is founded equally on the function of the individual as a member of the community” (Hobhouse 1964, p. 116).

In sum, Hobhouse helped provide a theoretic basis for a twentieth-century liberalism severed

from the constricted framework of its origins and aimed at the liberation of *all* members of society. It continues to be very much concerned with feelings, emotions, and historical experience. Twentieth-century *conservatism* assumed the mantle of rigid utilitarianism shorn of its humanistic aspirations.

The contradictions in conservative doctrine become even more apparent when the matter of prejudice is considered. As Nisbet explains it, prejudice has its own intrinsic wisdom anterior to intellect. It can be readily applied in emergencies and does not leave one indecisive at the moment of decision. It sums up in an individual mind tradition’s authority and wisdom (Nisbet 1986, p. 29).

This seems to epitomize the “mechanical” thought attributed to utilitarian or enlightenment liberals. Prejudice, in these terms, is, in effect, a mode of preprogrammed decision making. It insists upon shackling the human mind when confronted with new or unforeseen situations. It demands that such situations be dealt with through the use of what may well be outdated modes of thought, with strategies that were perhaps once useful but have lost their relevance. It denies a role for human creativity.

It is but a step from this doctrine to the prejudice castigated by civil rights activists and others. We meet a man whose skin is black or yellow; we meet a woman whose features tell us she is Jewish. We have preprogrammed responses to each of these, based on the “wisdom” of tradition, informing us that they, in various ways, are inferior creatures who must be dealt with accordingly.

The feudal origins of conservative doctrine are seen most clearly in its adamant stand on the issue of inheritance and property. Not only does it fight all efforts to loosen property from family groups by means of taxation, it fights all other efforts to redistribute wealth, ranging from special entitlements to affirmative action programs. It insists upon the indestructibility of existing hierarchical structures in society, seeing all efforts to modify them as attacks on cultural and psychological diversity.

Further complicating the distinction between liberalism and conservatism has been what some might refer to as a fringe movement within conservatism, called *neoconservatism*.

One explanation for the emergence of this phenomenon begins by noting that in the mid-1960s American universities, as well as literature and art in general, had become increasingly radicalized. Some cold war liberals, others who were uncomfortable with black power politics, and some critics of the counterculture disengaged themselves from liberalism. Prominent names in this group include Irving Kristol, Norman Podhoretz, and sociologists Seymour Martin Lipset, Nathan Glazer, Daniel Bell, and James Coleman. Some of these continue to reject the label but “it is as though by some invisible hand their writings and lectures gave help to the conservative cause when it was needed” (Nisbet 1986, p. 101).

Periodicals most strongly identified with neoconservatism are *The Public Interest* and *Commentary*. Neoconservatives have had close ties with the Scaife, Smith Richardson, John M. Olin, and other foundations. They have appeared as resident scholars and trustees of think tanks like the Hoover Institution, The Heritage Foundation, and the American Enterprise Institute (Gottfried and Fleming 1988, p. 73).

Unlike more traditional conservatives, neoconservatives are not irrevocably opposed to some possible versions of a welfare state. They do not depend upon historical methodology; on the contrary, they seem to show a marked preference for quantification.

Thus, the difference between contemporary liberalism and conservatism is apparently not to be found in issues of methodology, personality, or individual items of public policy. Yet there seems to remain an ineradicable core of difference. Ultimately this must be sought in the structure of material interests on the one hand and values on the other.

Historically, conservative doctrine was formulated as an intellectual defense of feudal property rights against the onslaughts of an emerging, business-oriented, industrializing bourgeoisie. Liberalism, with its early defense of individualism, championed the enemies of conservative doctrine and properly (from its perspective) fought against dominance by a central governing authority. Subsequently (as in the case of New Deal liberalism), it discovered that the logic of unhampered individualism led to serious economic difficulties for large

numbers of the population. Experiments with varieties of the welfare state since the days of Bismarck probably can all be traced in considerable measure to the fear of more drastic consequences that might follow any effort to persist in an uncontrolled laissez-faire economy.

In the contemporary world complexity has become compounded. It has probably always been the case that some people refuse to act in accordance with their own more or less self-evident material interests. These days, however, the arts of propaganda and advertising have been raised to a level of effective applied science. It is difficult for many to recognize exactly where their own self-interest lies. Beyond this, numerous techniques have been developed with the practical effect of inducing many to adopt positions in marked contrast to their own existing material interests. Government-sponsored lotteries hold out the hope of dramatic changes in class position. Tales of fabulous profits to be made in real estate or the stock market induce many to identify with the interests toward which they aspire rather than with those they hold and are likely to retain.

This is not the only source of confusion. Although many intellectuals, as well as others, having “made it,” subsequently become concerned with the maintenance of an order that has been good to them (Coser 1977), there are many others who, despite coming from very well-to-do backgrounds, maintain values consistent with economic deprivation. There are indeed generous souls as well as villains to be found among adherents of both liberalism and conservatism.

The shape of values is by no means always coterminous with the shape of existing material interests. Labels like liberalism and conservatism are uncertain predictors of specific actions, as are existing material interests. A variety of social forms, or structures, can be used to serve either selfish or communally oriented values. It is these values that are ultimately more reliable auguries.

In recent years there have been efforts to develop more useful theoretical frameworks for conceptualizing value configurations and more empirically based social policy alternatives. For example, Amitai Etzioni has elaborated an “I-We” paradigm as a substitute for both “unfettered” individualism and organismic views of society (Etzioni 1988). S. M. Miller and his colleagues

eschew the *liberal* label in favor of *progressive*, a term less saddled with conflicting conceptual baggage. They have provided an agenda for social policy issues requiring considerable rethinking by those unwilling to accept either contemporary conservative and neoconservative doctrine or traditional socialist formulations (Ansara and Miller 1986).

It is clear that serious linguistic difficulties serve to exacerbate the problem of distinguishing clearly between liberalism and conservatism. Each term tends to represent a generic symbol for a range of widely diversified issues, values, and interests. Surrogate expressions are used extensively in popular speech to convey either finer shades of meaning or degrees of opprobrium.

For conservatives these expressions may include “right-wing crazies,” “extreme right,” “right-wing,” “supply-siders,” “libertarians,” “filthy rich,” “Republicans,” and others. In contemporary America, portions of the conservative spectrum voice strong opinions not only on economic issues but also on such “social” issues as pornography, abortion, and affirmative action, although opinion on these issues is by no means unanimous. Traditional conservatism would, of course, back state-supported cultural standards in speech, art, literature, and entertainment. Nineteenth-century liberalism would oppose these in the name of individual “rights.” Is abortion a woman’s “right” or is it an “offense” against society? Many conservatives might define it as an offense; others might well define it as a right—a traditional liberal position.

Among liberals, surrogate expressions in use include “Red,” “parlor-pink,” “left-liberal,” “bleeding heart,” “tax and spender,” “Democrat,” and “Progressive.” The liberal spectrum tends to support rights of women to have abortions or, more generally, to maintain control over their own bodies. It opposes abrogation of rights of self-expression through pornography legislation, censorship, or other efforts to monitor art, literature, theatre, and other communication vehicles. On the other hand, American liberals favor limitations on the asserted “rights” of corporations or individual business persons to discriminate in employment, housing, and other areas on the basis of skin color, age, religion, or physical disability—property rights that many conservatives insist are inviolable.

As one might expect, survey researchers have made and continue to make strong efforts to detect empirical differences between persons who are called or who call themselves liberal or conservative. Data from a variety of survey research studies indicate that, despite the existence of important philosophical differences between liberals and conservatives, changing social and economic conditions have at times compelled them to alter their positions on certain economic, social, and political issues without altering their underlying philosophies (McCloskey and Zaller 1984, p. 191). In general, these studies seem to confirm stereotypical images of both liberals and conservatives.

Liberals show a marked preference for social progress and human betterment, especially for the poor and powerless. Some believe personal happiness and success depend heavily upon institutional arrangements. Most liberals are “inveterate reformers”; they continually look for ways to improve the human condition by remodeling social, economic, and political institutions. Underlying this pursuit of change, social reform, and benevolence is faith in the potential perfectability of human beings and their capacity to manage their own affairs in a responsible and reasoned fashion.

Conservatives have a different notion of what constitutes the good society and how it can be achieved. Survey research data show that they have a more pessimistic view of human nature and its perfectability. They feel that people need strong leaders, firm laws and institutions, and strict moral codes to keep their appetites under control. Firm adherence to conventional norms and practices is essential for human well-being. They believe that those who fail in life must bear primary responsibility for its consequences. They are far less likely than liberals to support movements that have as their objectives the eradication of poverty, the better treatment of oppressed minorities, or the alleviation of social distress generally. They maintain that these movements, by disrupting existing institutions, do more harm than good (McCloskey and Zaller 1984, pp. 190–191).

Conventional public opinion polling techniques encounter increasing difficulties in this area. Thus, one study examined a hypothesized inverse relationship between socioeconomic status and conservatism on a wide range of so-called social issues.

(Many observers felt that, during the 1970s and 1980s, as well as during the late 1960s, the main support for liberal and left political parties in the United States and other Western industrialized countries came from youthful members of the upper or middle strata.) It found that, in practice, many issues defied a neat separation of interests and values or of economic versus social arenas. Many social issues such as environmental protection, nuclear power, defense spending, and race or gender problems have an economic dimension, to the extent that they influence opportunities for jobs or profits. Conversely, government domestic spending, a classic economic issue, may have a social dimension if it is seen as involving welfare or aid to minorities. No consistent relationship between socioeconomic status and conservatism was found, with one exception. Liberalism on social issues tends to increase with education, but even here the relationship varies considerably from issue to issue. The authors suggest that lack of a consistent relationship reflects both the diversity of social issues and the fuzziness of the social/economic distinction (Himmelstein and McCrae 1988).

Another study examined what appeared to be an anomaly in this area. It postulated that high socioeconomic status remains one of the best predictors of Republican party support and conservative attitudes in the United States, that Republicans are wealthier, more educated, and hold higher status jobs than Democrats and independents. Jewish liberalism, however, confounds this general relationship. American Jews are generally wealthier, better educated, and hold higher status jobs than average Americans but continue to be the most liberal white ethnic group in the United States (Lerner, Nagai, and Rothman 1989, p. 330).

The emergence of a small cadre of Jewish neoconservative intellectuals has raised questions about Jewish liberalism. A sample of Jewish elites was compared with their Gentile counterparts. The study found that Jewish elites continue to be more liberal. Despite a plethora of competing explanations, the study concludes that Jewish liberalism is a product of a family tradition of liberalism that developed in response to European conditions. Specifically, the authors suggest that the Jewish elites inherited a tradition of responding in particular ways to felt marginality. This raises the

question as to whether a realignment might occur when this cohort of American Jews loses its prominence and is replaced by a cohort with different patterns of socialization. In a concluding footnote, the authors raise the open question of whether events in the Middle East and the emergence in the Democratic party of an increasingly powerful African-American presence less supportive of Israel can transform the liberalism of American Jews (Lerner, Nagai, and Rothman 1989).

A study of public opinion on nuclear power concludes that assessing public opinion through responses to survey questions with fixed categories presents serious difficulties. It compares these difficulties with those arising from the effort to "impose elite dichotomies such as 'liberal' and 'conservative' on a mass public whose beliefs are not organized by such dimensions" (Gamson and Modigliani 1989, p. 36).

Perhaps the more general difficulty is not to be found solely in the insufficiency of measuring instruments but in the increasingly more truncated vistas of the "mass publics." Fundamental philosophical positions and value orientations seem to have become increasingly more obscured by the exigencies of short-range decision making.

In societies where immediate job opportunities, social pressures, and short-range profits have serious implications not only for the quality of life but also for existence itself, it is scarcely surprising to find that many public issues and even individual values are filtered through the prisms of short-run individual economic concerns and ethnic identification. Manipulating perceptions of vital interests through sophisticated media technology does much to resolve the recurrent riddle, "Is Blank a liberal—or a conservative?"

(SEE ALSO: *Attitudes; Individualism; Public Opinion; Value Theory and Research; Voting Behavior*)

#### REFERENCES

- Ansara, Michael, and S. M. Miller 1986 "Opening Up of Progressive Thought." *Social Policy* 17:3-10.
- Bell, Daniel A. 1997 "Liberal Neutrality and Its Role in American Political Life." *Responsive Community* 7:61-68.
- Burke, Edmund 1855 "Reflections on the Revolution in France." In *Works*, Vol. 1. New York: Harper.

- Cohen, Mitchell 1997 "Why I'm Still 'Left'." *Dissent* 2:43-50.
- Cohen, Steven M., and Charles S. Liebman 1997 "American Jewish Liberalism: Unraveling the Strands." *Public Opinion Quarterly* 61:405-430.
- Coser, Lewis A. "Introduction." In Coser, Lewis A., and Irving Howe, eds., 1977 *The New Conservatives: A Critique from the Left*. New York: Meridian.
- Etzioni, Amatai 1988 *The Moral Dimension: Toward a New Economics*. New York: Free Press.
- Favela, Alejandro, and Miriam Calvillo 1996 "Liberalism and Postmodernity; Liberalismo y posmodernidad." *Estudios Politicos* 13:55-69.
- Gamson, William A., and Andre Modigliani 1989 "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach." *American Journal of Sociology* 95:1-37.
- Girvetz, Harry K. 1966 *The Evolution of Liberalism*. New York: Collier.
- Gottfried, Paul, and Thomas Fleming 1988 *The Conservative Movement*. Boston: Twayne.
- Halevy, Elie (1928) 1972 *The Growth of Philosophic Radicalism*, trans. Mary Morris. London: Faber and Faber.
- Himmelstein, Jerome L., and James A. McRae, Jr. 1988 "Social Issues and Socioeconomic status." *Public Opinion Quarterly* 52:492-512.
- Hobhouse, L. T. (1911) 1964 *Liberalism*. New York: Oxford University Press.
- Laski, Harold J. (1936) 1962 *The Rise of European Liberalism*. New York: Barnes & Noble.
- 1997 *The Rise of European Liberalism (New Edition)*. New Brunswick, N.J.: Transaction.
- Lerner, Robert, Althea K. Nagai, and Stanley Rothman 1989 "Marginality and Liberalism Among Jewish Elites." *Public Opinion Quarterly* 53:330-352.
- Lovejoy, Arthur 1960 *The Great Chain of Being: A Study of the History of an Idea*. New York: Meridian.
- McCloskey, Herbert, and John Zaller 1984 *The American Ethos: Public Attitudes Toward Capitalism and Democracy*. (A Twentieth Century Fund Report) Cambridge, Mass.: Harvard University Press.
- Miller, Alan S., and Takashi Nakamura 1997 "Trends in American Public Opinion: A Cohort Analysis of Shifting Attitudes from 1972-1990." *Behaviormetrika* 24:179-191.
- Nisbet, Robert 1986 *Conservatism*. Minneapolis: University of Minnesota Press.
- Sandel, Michael J. 1996 *Democracy's Discontent: America in Search of a Public Philosophy*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Stone, Brad Lowell 1997 "Classical Liberalism and Sociology." *Sociological Forum* 12:497-512.
- Tillyard, E. M. W. 1959 *The Elizabethan World Picture*. New York: Vintage.
- Witonski, Peter (ed.) 1971 *The Wisdom of Conservatives*, vol. 1. New York: Arlington.

ROBERT BOGUSLAW

## LIBRARY RESOURCES AND SERVICES FOR SOCIOLOGY

### INTRODUCTION

Libraries have a long history of providing access to the resources sociologists and other social scientists have needed and used, and to the literature they produce. In addition, libraries have established a strong tradition of providing reference and instructional services. Although the recent and pervasive growth in information technology has led to major changes in the way these resources and services are provided, libraries will continue to serve an important function to the discipline. Knowing how libraries are organized and how they work, what research tools and services are available, and how to use these tools and services effectively can help researchers at various levels to be more efficient and productive.

### AN OVERVIEW OF ACADEMIC LIBRARIES

Many of the thousands of academic libraries in the United States and other parts of the world support sociological course work and research, and researchers near major urban centers can also obtain important materials and substantial research help from some of the larger nonacademic public and research libraries. However, each library's ability to provide these kinds of support is affected greatly by the size of its budget, collections, and staff.

At the larger end of the size continuum, libraries can be quite complex organizationally, although most users will be aware of only the parts and functions of direct relevance to them. In the United States, libraries are organized so that users can use them effectively on their own, identifying what is needed through a public catalog and an open

“stacks” area. Almost all users have contact with staff members at a circulation desk, who check out books and sometimes periodicals to them. These staff members also help users locate items they may not be able to find on their own and help make resources available by enforcing loan periods, recalling books from other users, setting up reserve reading rooms for heavily used materials assigned for course readings, and so on. Academic libraries also typically organize their journals into periodical reading rooms or stack areas, and may provide related support services there. Interlibrary loan departments provide access to resources owned by other libraries.

Almost all libraries provide reference service in a variety of ways, and often in different settings. The role of the reference librarian typically extends far beyond the answering of informational questions, and may include providing individualized help in using electronic or complex print resources, or organizing a literature search, as well as speaking to classes about major disciplinary research tools and strategies. Reference librarians often have special areas of subject expertise and may perform liaison or collection development work with academic departments for book and journal purchasing. Where this is the case, graduate students may find it helpful to get to know the librarian responsible for sociology. College and university libraries will generally also have special sections or departments for government publications. Because these resources can be quite specialized, and access to them complex, staff help can be especially important.

Regardless of size, most academic libraries must deal with a variety of interrelated budgetary and technological pressures. For the last several years and for a variety of reasons, the costs of providing periodicals and other serials have been rising more rapidly than other indicators of inflation. (Ketcham–Van Orsdel and Born 1998). This trend has resulted both in a larger share of budgets being devoted to serials and in an ongoing need to evaluate and sometimes cancel subscriptions to journals. At the same time, there has been very rapid growth in book title production in the United States (Bosch 1998), with the result that libraries typically buy a decreasing share of the domestic books published in most disciplines. Economic pressures and increased book production abroad

have also made it more difficult for research libraries to buy as large a share of foreign language books and journals as in the past. Accordingly, libraries rely increasingly on cooperative buying and resource sharing. Paradoxically, library users may now actually have better access to some resources than in the past when locally owned resources were relied on more exclusively.

Key developments in information technology such as the CD-ROM and the World Wide Web, have also introduced new options for delivering information. On-line catalogs with sophisticated search features and other functionality have quickly replaced card catalogs. In addition, many periodical indexes and abstracts and the text of key journals are also now available online. These services make it possible for researchers to locate information quickly and easily—even from their homes or offices—but they almost always cost substantially more to purchase and support than do comparable print resources. Together, these factors make for a complex, rapidly changing environment for academic libraries and their users.

### **LIBRARIES AND THE LITERATURE OF SOCIOLOGY**

Because of sociology’s great topical and methodological diversity, it is difficult to characterize its literature. This diversity is well represented in the hundreds of articles in this encyclopedia, and in recent discussions of sociology’s most influential books (Clawson and Zussman 1998; Gans 1998; Marwell 1998; Sullivan 1994). Sociologists rely on and use a range of publications and information sources for their research, including books, journals, statistical publications and data sets, and governmental and other specialized reports (Zabel 1996; Shapiro 1985). Of these, the most important outlets for sociological writings have been books and scholarly journals, which appear to assume somewhat greater or lesser importance for different research communities. Some have suggested, for example, that there are two fairly distinct research cultures in sociology—one a “book sociology” that relies on, values, and publishes primarily in books, and the other an “article sociology” that relies on and publishes in scholarly journals. To some extent article sociology may be more characteristic of the scientific end of the discipline,

and book sociology more characteristic of its humanistic, historical, ethnographic, and theoretical emphases and traditions (Sullivan 1994).

The sociological journal literature is quite extensive. Of the 5965 items judged significant enough to be included in the 1996 volume of the *International Bibliography of Sociology*, roughly three-quarters were periodical articles, and even more articles are indexed annually in *Sociological Abstracts*. According to the 1998–1999 *Ulrich's International Periodicals Directory*, several hundred periodicals relevant to the field are published worldwide, although this count includes newsletters and other less substantial publications. Hargens (1991) estimated that roughly 250 of these might be characterized as research journals publishing original reports of relevant research findings. Lists of the more prominent or important titles have ranged in size from a couple of dozen (Sociology Writing Group 1998) to sixty or so (Katz and Katz 1997; Aby 1997), to well over a hundred (Bart and Frankel 1986). As in many other disciplines, a handful of journals (primarily the *American Journal of Sociology*, *American Sociological Review*, and *Social Forces*) are widely regarded as the most prestigious and/or influential. However, it appears that publication of important research is less likely to be concentrated in these few journals than in some other disciplines, and use of and reliance on review serials like the *Annual Review of Sociology* is relatively low (Hargens 1991). As in other disciplines, articles submitted to journals in sociology are typically refereed, or subjected to critical evaluation by peers who thus serve an important gatekeeping function (Mullins 1977; Osburn 1984; Simon 1994).

Many journals are now available electronically as part of larger full-text services to which libraries subscribe, and archives of the journals published by the American Sociological Association are now available electronically via the World Wide Web through the innovative JSTOR program (Guthrie and Lougee 1997; see also <http://www.jstor.org/>). In addition, a few sociological journals are available exclusively via the Web, and there has been much recent speculation that journal publishing may soon undergo fundamental change in response to the Web's interactive possibilities. As in the field of physics, for example, articles in other disciplines may commonly circulate on the Web

prior to formal publication. The statistical data on which an article is based may also be linked to the Web version of the article for further analysis—which may have special significance for sociology.

As noted earlier, books are also important to the discipline and annual production numbers at least in the thousands. The 1996 volume of the *International Bibliography of Sociology*, mentioned earlier, listed about 1,300 titles from around the world, but this number did not include textbooks or popular titles. A broader count of relevant titles published or distributed in the United States that same year put the figure at 4,186 (Bosch 1998), up more than 50 percent from 1989. As with journals, hardly any library can build a comprehensive book collection, and libraries select on the basis of relevance to local emphases, user requests, judgments of quality, cost, and other factors. Often book vendor “approval programs,” which match published books with formalized interest profiles, are relied on, since they can supply new books quickly and efficiently. Standard book review publications like *Choice* (aimed at undergraduate libraries) and *Contemporary Sociology* (a publication mainly for sociologists) may also be systematically utilized for selection, but they lack the space to be comprehensive. (During 1997, for example, *Choice* reviewed only 216 of the available titles, and *Contemporary Sociology* roughly 500.) At this writing relatively few newly published books or monographs are being made available electronically, but this seems likely to change in the near future.

There are a variety of other outlets for sociological research in addition to books and journal articles, and sociologists use many other sources of information in their research. For example, those sociologists who work for government agencies often write formal reports, which may be published by the agency. Statistical reports of various kinds that are of interest to sociologists (the U.S. Census being a prime example) are often also published in printed form by government agencies and distributed to academic libraries. These data are also increasingly being made available via the Web. A few academic libraries also pay for and coordinate their institutions' memberships in the Interuniversity Consortium for Political and Social Research (ICPSR), and make the associated survey data and codebooks available, although it is more typical for other campus agencies to do so.

And, of course, the dissertations written by doctoral students are typically acquired and cataloged by their institution's libraries, and sold through Bell and Howell Information and Learning in Ann Arbor, Michigan.

The published literature in sociology can also be seen as part of a broader professional communication system which includes less formal interchanges like presentation of papers at regional and national conferences (Osburn 1984). Developments in information technology have fostered the growth of such "invisible colleges"—especially in sociology, where computers have long been important tools. Sociologists have, for many years, made use of data analysis packages the Statistical Package for the Social Sciences (SPSS) and now quite commonly have their own personal computers with word processing software and connections to the Internet. As a result, formerly disparate activities like data analysis, writing, and publishing have begun to merge (Anderson 1998). E-mail listserv discussion groups, and the ease with which writings can be posted to the World Wide Web have fostered efficient communication among sociologists sharing research interests, and seem to hold considerable promise for the discipline (Bainbridge 1995). Libraries participate in and help foster these developments in a variety of ways. For example, many libraries participate in the JSTOR program, subscribe to full text services, and provide their users with on-line access to journal indexes like *Sociological Abstracts* and the *Social Sciences Citation Index*. Many also support the development of exclusively electronic journals by directing users to them through their on-line catalogs and Web pages.

### GENERAL STRATEGIES FOR LITERATURE SEARCHING IN SOCIOLOGY

It is difficult to provide good general strategies for location information in sociology for a variety of reasons. As noted, any number of topics and approaches may be pursued, and a large number of journals and other sources may contain important articles or other information. In addition, a doctoral student "terrified by the literature" in anticipation of preliminary exam questions (Becker 1986) will need to use different research strategies than will an undergraduate student writing the

typical library term paper. Rapid changes in information technology make it even harder to suggest tactics that will be valid five years from now. It is also useful to remember that, although librarians tend to view literature searching as something of a structured and rationalized activity, it can and should often take place in a more open, informal, serendipitous, or even mysterious way—especially in the early stages of a research project. Nevertheless, being alert to the following issues and suggestions can help make literature searching more effective and more efficient. Users are urged to consult their local library staff for additional advice and guidance, and for current information on local resources.

1. *Overview or summarizing tools.* Although sociologists tend not to use or rely on review serials like the *Annual Review of Sociology*, these can often prove helpful by summarizing and evaluating the main themes of recent research and setting them in a broader context. A similar function is played by subject encyclopedias like this one, by disciplinary handbooks (such as Smelser 1988; Smelser and Swedberg 1994; Gilbert et al. 1998), and to some extent by sociology textbooks. Some summarizing sources like these can be found in general and specialized reference books and guides (Aby 1997; Balay 1996; Wertheimer 1986; Zabel 1996). In addition, searches of electronic databases can sometimes be limited to "review articles."
2. *Differences in indexing terminology.* Periodical indexes and library catalogs quite often use a defined list of subject terms, which may vary from those commonly used in an area of literature, by a community of scholars, or by an individual student trying to describe a topic. Successful use of these tools often requires matching an idea to indexing terminology. This can save time by helping to eliminate irrelevant citations.
3. *Techniques for searching electronic resources.* The scholarly communication system in all disciplines seems destined to be tied increasingly to developments in information technology, and researchers will need to understand how to interact with and



use electronic tools effectively. Although most users now know that rough “keyword” searching can give them relevant citations, it is also important to know how to use “Boolean operators,” such as AND, OR, and NOT to combine and manipulate terms (Sociology Writing Group 1998) and to incorporate subject indexing terms into a search strategy.

4. *Evaluation of sources.* The gatekeeping role played by peer reviewers of journal articles helps guarantee that an article has passed a test of quality or adherence to accepted research norms. This is less apt to be the case with other information sources—especially those found on Web pages or via Web search engines. It is consequently useful to develop a generally skeptical outlook on information sources, and to evaluate such sources on the basis of such things as credibility of origin, scope and coverage, currency, and reputation (Sociology Writing Group 1998).
5. *Library collections as a linked system.* Since it seems unlikely that the financial constraints facing academic libraries will ease in the foreseeable future, libraries will continue their efforts to share resources efficiently. As a result, library users should assume that the collections available to them extend far beyond what their local libraries own. Graduate students and faculty, especially, will need to be aware of how their libraries are making these broader resources available and how long it will take to obtain publications from elsewhere.

#### LOCATING PERIODICAL LITERATURE

The most important tools for locating relevant periodical articles are known by librarians as abstracting and indexing services, because they abstract (or summarize) articles and index them (apply subject terms to them) according to a vocabulary developed for the purpose. These tools are frequently available in both printed and electronic form, and the electronic versions offer powerful searching capabilities and other features such as links to the full text of articles. The most important for researchers in sociology are the *Social*

*Sciences Index* (and some competing products), *Sociological Abstracts*, and *Social Sciences Citation Index*.

The *Social Sciences Index* and several competing products are important because they provide students with relatively easy access to a more manageable subset of the available literature than do the other two main tools. Because even smaller academic libraries will tend to subscribe to the majority of periodicals indexed in them, they are especially helpful for beginning students. For many years the printed *Social Sciences Index* was the primary general tool for social science researchers. It has gradually grown in coverage and now indexes roughly four hundred important social science journals, including about fifty titles in sociology, and is available in a few different electronic versions with abstracts and the full text of some of the indexed articles. Articles are carefully indexed according to subject, with “see” and “see also” references pointing users to other relevant subjects.

Within the last several years a few companies like EBSCO, the Gale Group, and Bell and Howell have offered some other, more general periodical indexes aimed at the college or undergraduate library market which cover largely the same range of social science literature. Although the sociology journals indexed are fairly similar, these products do differ in their indexing and abstracting practices, how far back their indexing and full text coverage extends, which titles are provided in full text, and in their search capabilities and limitations.

In contrast to these general sources, *Sociological Abstracts* covers the sociological literature much more comprehensively. Since this source is so fundamental to literature searching in sociology, it is worth quoting the publisher’s description at length:

*Sociological Abstracts* provides access to the world’s literature in sociology and related disciplines, both theoretical and applied. The database includes abstracts of journal articles selected from over 2,500 journals, abstracts of conference papers presented at various sociological association meetings, relevant dissertation listings from *Dissertation Abstracts International*, enhanced bibliographic citations of book reviews, and abstracts of selected sociology books published in *Sociological Abstracts* (SA) and *Social*

*Planning/Policy and Development Abstracts*  
(SOPODA) since 1974.

Approximately 2,500 journals in thirty different languages from about fifty-five countries are scanned for inclusion, covering sociological topics in fields such as anthropology, economics, education, medicine, community development, philosophy, demography, political science, and social psychology. Journals published by sociological associations, groups, faculties, and institutes, and periodicals containing the term "sociology" in their titles, are abstracted fully, irrespective of language or country of publication. Noncore journals are screened for articles by sociologists and/or articles of immediate interest or relevance to sociologists.

The abstracts provided are typically lengthy and detailed, and articles are indexed by author and a sufficient number of indexing terms from the *Thesaurus of Sociological Index Terms* to describe their content. The *Thesaurus* has been developed over time with sociological concepts and terminology in mind, and this tool—which may be integrated into an electronic version of the publication—is a key to making searching more efficient and effective. As shown by the following sample entry for the term "Satisfaction," the *Thesaurus* indicates what subject terms are available for searching, and what relationships they have with one another.

**Satisfaction****DC** D740400**SN** A context-dependent term for an individual's positive assessment of self or circumstances. Select a more specific entry or coordinate with other terms.**HN** Formerly (1963–1985) DC 403350**UF** Fulfillment (1969–1985)**BT** Attitudes**NT** Community Satisfaction  
Job Satisfaction  
Life Satisfaction  
Marital Satisfaction**RT** Discontent  
Emotions  
Happiness  
ImprovementNeeds  
Quality  
Self Esteem

The most important codes shown in this example are as follows. **SN** stands for a "scope note," or definition of the term. **BT** indicates that the "broader term" of "Attitudes" can be used. **NT** stands for "narrower" or more specific terms, and **RT** for "related terms." **UF** means that "Satisfaction" is "used for" Fulfillment (in other words, fulfillment is not used as a subject term). (For a complete discussion of the codes and their meaning and use, consult the *Thesaurus*.)

Users of both the printed and electronic versions of *Sociological Abstracts* would be able to find abstract entries using these terms. The printed index refers users to an abstract number, which is then looked up in another part of the volume to actually locate the abstract, whereas the electronic version would provide a set of abstracts that could be reviewed and printed out. The following example was found on line by searching for "Life Satisfaction" as a subject term using a version of Sociological Abstracts produced by Silver Platter Information in Norwood, Massachusetts.

**TI:** Marital Status, Gender, and Perception of Well-Being**AU:** Mookherjee,-Harsha-N.**IN:** Dept Sociology Tennessee Technological U, Cookeville 38505**SO:** Journal-of-Social-Psychology; 1997, 137, 1, Feb, 95–105.**IS:** 0022-4545**CO:** JSPSAG**DT:** aja Abstract-of-Journal-Article**LA:** English**CP:** United-States**PY:** 1997**AB:** Draws on data from the combined 1982–1991 National Opinion Research Center's General Social Surveys (total N = 12,168 adults) to reexamine relationships among marital status, gender, & perception of well-being. ANOVA revealed that marriage significantly enhances perception of well-being for both men & women, though in general, women express more satisfaction than men. Well-being perceptions were significantly affected by

race & financial status, regardless of marital status. 3 Tables, 44 References.

Adapted from the source document

DEM: \*Marital-Status (D491150); \*Well-Being (D916500); \*Sex-Differences (D758100); \*Single-Persons (D771900); \*Life-Satisfaction (D463800)

DES: United-States-of-America (D890700)

SH: social psychology; personality & social roles (individual traits, social identity, adjustment, conformism, & deviance) (0312)

The “DEM” label in this entry stands for “major descriptors”, and indicates that the main topics of this article are Marital Status, Well-Being, Sex-Differences, Single-Persons, and Life-Satisfaction. The “DES” label stands for “minor descriptors,” and shows that although the article has been indexed under “United-States-of-America,” this is not an important focus.

As noted earlier, a key technique in using this and other electronic indexes is combining terms using Boolean operators to make searching more precise and to limit the amount of material that must be reviewed. In this case, this article and other entries could have been found by searching for the subject terms “Life Satisfaction” AND “Sex Differences.” A slightly more complex search strategy or statement might have been to search for “Life Satisfaction OR Well-Being” AND “Marital Status OR Single Persons.” This kind of search statement can be made as elaborate as necessary to the situation. It is also possible to limit search results by date, language, journal, and type of publication (such as book review, journal article, or conference paper), and to search on virtually any combination of words found in the article title and abstract entry. It is, of course, impossible or impractical to do so with the printed publication. Because of its comprehensiveness, many users of *Sociological Abstracts* will need to adjust their search strategies to exclude conference papers and dissertations (which are often difficult to obtain) and foreign language publications.

The *Social Sciences Citation Index (SSCI)* does not cover journals in sociology quite as extensively as *Sociological Abstracts*, but it covers journals in other social science disciplines more completely than the latter, and much more comprehensively

than does the *Social Sciences Index*. In 1997, for example, the *SSCI* indexed 1,725 journals “completely,” including 92 titles in sociology, and indexed another 1,371 on a selective basis. During that year, 72,665 articles and another 39,412 book reviews in all social science disciplines were indexed. Unlike *Sociological Abstracts*, abstracts have only recently begun to be provided in *SSCI* (in the electronic versions only), and articles are not indexed according to a fixed indexing vocabulary. Instead, heavy reliance is made in the print version on words from the titles of articles. Articles are also indexed by author, of course, although the publisher uses only authors’ first and middle initials, rather than their full names, which can occasionally cause confusion.

What is uniquely valuable about this source is that it enables users to search “by citation,” or to locate articles that have cited an earlier author or article. Eugene Garfield, the founder of the Institute for Scientific Information, which publishes *SSCI*, described the idea behind citation indexes as follows:

*The concept of citation indexing is simple. Almost all the papers, notes, reviews, corrections, and correspondence published in scientific journals contain citations. These cite—generally by title, author, and where and when published—documents that support, provide precedent for, illustrate or elaborate on what the author has to say. Citations are the formal, explicit linkages between papers that have particular points in common. A citation index is built around these linkages. It lists publications that have been cited and identifies the sources of the citations. Anyone conducting a literature search can find from one to dozens of additional papers on a subject just by knowing one that has been cited. And every paper that is found provides a list of new citations with which to continue the search. (Garfield 1979, p. 1)*

Books as well as journal articles in sociology can be, and often are, cited by the journal articles indexed in *SSCI* (Sullivan 1994) and the ability to search on cited references may provide a large number of additional “access points” for locating an article. The article on marital status that was used as the sample record from *Sociological Abstracts*, for example, cited forty-four earlier articles

and studies. Users of *SSCI* could search on any of those citations and locate the article that way.

One difficulty with performing citation searches lies in knowing when it will be useful to do so. In sociology, although it is possible to perform citation searches on the handful of authors in the recognized pantheon (i.e., Marx, Durkheim, Weber, and some others), such an approach will often result in far too many references with only tenuous subject relationships to one another. It may consequently be necessary to combine a result set from a citation search with some other group of search terms. Where citation searching seems most useful in sociology is when tracing a methodological article, or in other somewhat narrow circumstances. To do a citation search for a journal article—using either the printed *SSCI* or its electronic counterpart—requires that a researcher know the author's name and initials, and preferably the volume, volume number, year, and page of the publication. For example, Clifford Clogg and Gerhard Arminger's 1993 article titled "On Strategy for Methodological Analysis" from the publication *Sociological Methodology* (vol. 23, pp. 57-74) is listed in *SSCI* as:

CLOGG CC SOCIOL  
METHODOL 23 57 1993

By searching the printed or electronic versions of *SSCI* under this entry, it would be possible to find other articles which cited this article. It would also be possible to look at the list of sources cited by Clogg and Arminger and find other articles which have cited them. As with *Sociological Abstracts*, the electronic version of *SSCI* makes for far more efficient searching—especially of citations. Another interesting use to which citation searching is often put is comparing faculty members' research productivity—especially for promotion and tenure consideration. This procedure is controversial for a variety of reasons—especially when comparing faculty members from different disciplines, which may not share the same publication and citation patterns.

Depending on the particular topic being researched, and the depth with which the search must be carried out, a number of other tools in neighboring or related fields may prove useful, including *Anthropological Literature*, *EconLit*, *Popline*, *Psychological Abstracts/PsycInfo*, and *Social Work Abstracts*. In addition, researchers may be able to avail

themselves of journal article "alerting" services, such as *Uncover Reveal*, which automatically sends the tables of contents of specified journals and article citations containing specified keywords to users via e-mail. Although some of the electronic versions of the indexes discussed also provide linkages to a local library's journal holdings information or to on-line full text, accurately determining which articles are available locally still typically requires a separate step.

### LOCATING BOOKS, JOURNALS, AND OTHER RESOURCES IN LIBRARIES

As noted earlier, most libraries now provide on-line versions of their catalogs. Unlike card catalogs, on-line catalogs can be searched via keyword, and with Boolean operators or combinations. They may also provide direct links to electronic resources on the World Wide Web, as well as the circulation status of a book. In addition, many now permit users to view lists of the books they have checked out and to reserve items checked out to other users. Despite these and other less visible changes, books are catalogued in much the same way they have been for years: by author, title, and subject. Few college students will be unaware that books can be searched by author and title, but they may also mistakenly assume that a journal article can be found by its author or title in a library catalog. Instead, the title of the journal must be searched, and then the specific volume and page located on the shelf within that journal.

Although many students will also realize that books can be located by subject, few will understand how the subject heading system works. The system used for this in most academic libraries was developed at the Library of Congress in Washington, D.C., and the headings themselves are called Library of Congress Subject Headings (LCSH). These headings are published in multivolume sets by the same name, which now closely resemble database thesauri like the *Thesaurus of Sociological Index Terms*. They can be used in the same way to find relevant subject headings for searching. For example, the entry for "Satisfaction" in the 1997 edition of LCSH looks like this:

**Satisfaction**  
BT Self  
RT Self-realization  
NT Consumer satisfaction

Contentment	HQ-HT	Social Groups
Housing—Resident satisfaction	HQ	Family, Marriage, Woman
Job satisfaction	HS	Societies: Secret, Benevolent, etc., Clubs
Libraries—User satisfaction	HT	Communities, Classes, Races
Office buildings—Tenant satisfaction	HV	Social Pathology. Social and Public Welfare; Criminology
Patient satisfaction		
Public housing—Resident satisfaction		
Rental housing—Resident satisfaction		
— <b>Religious aspects</b>		
— — <b>Buddhism, [Christianity, etc.]</b>		

Like the entry from the *Thesaurus*, this shows the availability of Broader (BT), Related (RT), and Narrower (NT) terms. Unlike terms from the *Thesaurus*, though, *LCSH* lists a large number of terms that are subdivisions of more general terms. In other words, if a researcher were to look under “Libraries” as a subject term, she or he might find books listed under many subdivisions, including “User satisfaction.” This is also true of subject headings that are more directly related to sociology. For example, the term “Sociology” can be subdivided by a country or other geographic term, or by such terms as “—Methodology”, “—Philosophy,” or “—Statistical methods.” Subject headings can also be modified by adding a comma and an adjective to a term. For instance, “Sociology, Islamic” would be used instead of “Islamic Sociology.” Although it is not necessary to thoroughly understand this system, it is useful to realize that it exists, and that some help from library staff in identifying useful subject headings may be needed. As with other databases, noticing what subject headings have been applied to relevant, known books may be a good start, and may lead to the finding of other important books.

Most academic libraries now also physically organize their book and journal collections using another system developed at the Library of Congress: its call number scheme. Unlike the Dewey Decimal system that is commonly used in school and public libraries, Library of Congress call numbers start with one or two letters. For example, sociology and economics have both been assigned the letter H, to which a second letter is added for a further breakdown. For sociology, the primary classes are:

#### **SOCIOLOGY**

7HM	General Works, Theory
HN	Social History and Conditions; Social Problems, Social Reform

Numbers make these breakdowns more specific. For example, before extensive changes were introduced recently, the numbers between HM 1 and HM 299 were used for topics in sociology. Under the revised scheme, HM numbers 401 and higher will be used instead, which will allow a finer topical breakdown and arrangement. For example, under the old scheme the numbers from HM 251 to HM 299 were assigned to various topics in Social Psychology, such as:

HM	251	General
	255	Instinct in social psychology
	261	Public Opinion
	263	Publicity. Propaganda.

A particular book in general social psychology would have been assigned the subject call number HM 251, to which additional combinations of letters, numbers and possibly dates designating the author or work would be applied. For example, the 1998 edition of the *Handbook of Social Psychology* was assigned the number HM 251 H224 1998, where H224 represents the title, and 1998 the edition. Under the revised system, books in social psychology will receive numbers HM 1000 and above, and the same *Handbook* cataloged under it might have the call number HM 1033 H34 1998. This will obviously and unfortunately cause books on similar topics but classed under the older and newer systems to be separated on library shelves. Again, it is not necessary to fully understand the call number system, but having a general sense of how it works can make locating books a little easier.

Many researchers will also need to develop an understanding of the kinds of materials that may be available in a given library but not listed in its catalog. The single largest category of such material, in most cases, consists of government publications or documents—especially those things published by the U.S. federal government, but also by state and local governments, the governments of other countries, and the United Nations or other international agencies. Because there are so many

of these publications and documents, very few libraries can provide thorough and complete coverage of them. As a result, other indexes and similar sources are often necessary. For the federal government publications of the United States, the primary tool is the *Monthly Catalog of United States Government Publications*, although there are a number of other supplementary sources available. For example, the *American Statistics Index* provides very detailed indexing coverage of federal statistical publications. Both of these tools were originally published in paper, but are now available both in electronic versions. Access to publications of U.S. state governments and those of other countries are also unlikely to be catalogued, and other tools may be needed.

## REFERENCES

- Aby, Steven 1997 *Sociology: A Guide to Reference and Information Resources*, 2nd ed. Englewood, Col.: Libraries Unlimited.
- American Statistics Index*. 1974– Washington, D.C.: Congressional Information Service.
- Anderson, Ronald E. 1998 “Computing in Sociology.” Pp. 52–65 in *Encyclopedia of Library and Information Science*, vol. 63, sup. 26. New York: Marcel Dekker.
- Annual Review of Sociology* 1975– Palo Alto, Calif., Annual Reviews.
- Bainbridge, William Sims 1995 “Sociology on the World-Wide Web.” *Social Science Computer Review* 13:508–523.
- Balay, Robert 1996 *Guide to Reference Books*, 11th ed. Chicago: American Library Association.
- Bart, Pauline, and Linda Frankel 1986 *The Student Sociologist's Handbook*, 4th ed. New York: Random House.
- Becker, Howard S. 1986 *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*. Chicago: University of Chicago Press.
- Bosch, Steven 1998 “Prices of U.S. and Foreign Published Materials.” Pp. 495–520 in *The Bowker Annual of Library and Book Trade Information, 1998*. New York: R. R. Bowker.
- Choice* 1964– Chicago: American Library Association.
- Clawson, Dan, and Robert Zussman 1998 “Canon and Anti-Canon for a Fragmented Discipline.” Pp. 3–17 in Dan Clawson, ed., *Required Reading: Sociology's Most Influential Books*. Amherst: University of Massachusetts Press.
- Contemporary Sociology* 1972– Washington, D.C.: American Sociological Association.
- Dissertation Abstracts*. 1952– Ann Arbor, Mich.: University Microfilms.
- Gans, Herbert J. 1998 “Best-Sellers by American Sociologists: An Exploratory Study.” Pp. 19–27 in Dan Clawson, ed., *Required Reading: Sociology's Most Influential Books*. Amherst: University of Massachusetts Press.
- Garfield, Eugene 1979 *Citation Indexing—Its Theory and Application in Science, Technology, and the Humanities*. New York: Wiley.
- Gilbert, Daniel T., Susan T. Fiske, and Gardner Lindzey 1998 *The Handbook of Social Psychology*, 4th ed. Boston: McGraw-Hill.
- Guthrie, Kevin M., and Wendy P. Lougee 1997 “The JSTOR Solution: Accessing and Preserving the Past.” *Library Journal* 122 (February 1): 42–44.
- Halstead, Kent 1998 “Price Indexes for Public and Academic Libraries.” Pp. 440–456 in *The Bowker Annual of Library and Book Trade Information, 1998*. New York: R. R. Bowker.
- Hargens, Lowell L. 1991 “Impressions and Misimpressions about Sociology Journals.” *Contemporary Sociology* 20 (May): 343–349.
- International Bibliography of Sociology* 1955– London: Tavistock Publications; Chicago: Beresford.
- Katz, William, and Linda Sternberg Katz 1997 *Magazines for Libraries*, 9th ed. New Providence, N.J.: R. R. Bowker.
- Ketchum–Van Orsdel, Lee, and Kathleen Born 1998 “E-Journals Come of Age: LJ's 38th Annual Periodical Price Survey,” *Library Journal* (April):40–45.
- Library of Congress 1997 *Library of Congress Subject Headings*, 20th ed., vol. 1–4. Washington, D.C.: Library of Congress.
- Marwell, Gerald 1998 “Sociological Politics and *Contemporary Sociology's* Ten Most Influential Books.” Pp. 189–195 in Dan Clawson, ed., *Required Reading: Sociology's Most Influential Books*. Amherst: University of Massachusetts Press.
- Monthly Catalog of United States Government Publications*. 1951– Washington, D.C.: Cataloging Branch, Library Division, Library Programs Service, Superintendent of Documents, Government Printing Office.
- Mullins, Carolyn J. 1977 *A Guide to Writing and Publishing in the Social and Behavioral Sciences*. New York: Wiley
- Osburn, Charles B. 1984 “The Place of the Scholarly Journal in the Scholarly Communications System.” *Library Resources and Technical Services* 28 (October): 315–324.
- Shapiro, Beth J. 1985 “Sociology.” Pp. 188–198 in Patricia A. McClung, ed., *Selection of Library Materials in the*

- Humanities, Social Sciences, and Sciences*. Chicago: American Library Association.
- Simon, Rita J. 1994 "An Effective Journal Editor: Insights Gained from Editing the *American Sociological Review*." Pp. 33–44 in Rita J. Simon and James J. Fyfe, eds., *Editors as Gatekeepers: Getting Published in the Social Sciences*. Lanham, Md.: Rowman and Littlefield.
- Smelser, Neil J. 1988 *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- , and Richard Swedberg 1994 *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press.
- Social Science Index* 1974 ff. New York: H.W. Wilson Co.
- Social Sciences Citation Index* 1956– Philadelphia, Pa.: Institute for Scientific Information.
- Social Sciences Citation Index Guide and Lists of Source Publications, 1997 1998* Philadelphia, Pa.: Institute for Scientific Information.
- Sociological Abstracts* 1953– Bethesda, Md.: Cambridge Scientific Abstracts.
- Sociology Writing Group 1998 *A Guide to Writing Sociology Papers*, 4th ed. New York: St. Martin's.
- Sullivan, Teresa A. 1994 "Genre in Sociology: The Case for the Monograph." Pp. 159–175 in Rita J. Simon and James J. Fyfe, eds., *Editors as Gatekeepers: Getting Published in the Social Sciences*. Lanham, Md.: Rowman and Littlefield.
- Thesaurus of Sociological Indexing Terms*, 4th ed. 1996 San Diego, Calif.: Sociological Abstracts.
- Ulrich's International Periodicals Directory*, 37th ed. 1999 New Providence, N.J.: R. R. Bowker.
- Wertheimer, Marilyn L. 1986 "Sociology." Pp. 275–331 in William H. Webb, ed., *Sources of Information in the Social Sciences: A Guide to the Literature*. Chicago: American Library Association.
- Zabel, Diane 1996 "Sociology." In Nancy L. Herron, ed. *The Social Sciences: A Cross-Disciplinary Guide to Selected Sources*, 2nd ed., Englewood, Colo.: Libraries Unlimited.

TIMOTHY D. JEWELL

## LIFE COURSE, THE

### INTRODUCTION

The study of lives represents an enduring interest of sociology and the social sciences, reflecting important societal changes and their human consequences. Most notably, developments after World War II called for new ways of thinking about

people, society, and their connection. In the United States, pioneering longitudinal studies of children born in the 1920s became studies of adults as the children grew up, thereby raising questions about the course they followed to the adult years and beyond. The changing age composition of society assigned greater significance to problems of aging and their relation to people's lives. Insights regarding old age directed inquiry to earlier phases of life and to the process by which life patterns are shaped by a changing society.

This essay presents the life course as a theoretical orientation for the study of individual lives, human development, and aging. In concept, the life course refers to a pattern of age-graded events and social roles that is embedded in social structures and subject to historical change. These structures vary from family relations and friendships at the micro level to age-graded work organizations and government policies at the macro level. Life-course theory defines a common domain of inquiry with a framework that guides research in terms of problem identification and formulation, variable selection and rationales, and strategies of design and analysis. Beginning in the 1960s, this theoretical orientation has diffused across substantive domains and disciplinary boundaries in the social and behavioral science.

It has uniquely forged a conceptual bridge between developmental processes, the life course, and ongoing changes in society, one based on the premise that age places people in the social structure and in particular birth cohorts. To understand this conceptual bridge, it is useful to distinguish among three levels of the life course and their interplay over a person's life: (1) *institutionalized pathways* in society, as established by state policies, education, the workplace, and so on; (2) the *individual life course* that is formed by the individual's choices and constraints, frequently in terms of a career or trajectory; and (3) the *developmental or aging trajectory of the individual*, defined, for example, by intellectual functioning or self-confidence.

Each of these levels are illustrated by Spilerman (1977) in terms of work. He used the concept of "career line" to refer to pathways that are defined by the aggregated work histories of individuals. Career lines are patterned by industry structures and the labor market. A person's work life is one

part of the individual life course, and it varies by the career requirements of firm and marketplace. At the psychological level, changes in work life have consequences for personal feelings of efficacy (Bandura 1997).

In this essay, I first take up concepts that have been used interchangeably—the life course, life cycle, life history, and life span. Then I turn to the emergence of life-course theory since the 1960s and its paradigmatic principles.

### CLARIFICATION OF CONCEPTS

A number of concepts have been applied interchangeably to lives (life course, life cycle, life history, and life span), but each makes a distinctive contribution that deserves notice in mapping this domain (Elder 1998). The *concept of life course* is defined by trajectories that extend across the life span, such as family or work; and by short-term changes or transitions, such as entering and leaving school, acquiring a full-time job, and the first marriage. Each life-course transition is embedded in a trajectory that gives it specific form and meaning. Thus, work transitions are core elements of a work-life trajectory; and births are key markers along a parental trajectory. Multiple marriages and divorces are elements of a marital trajectory.

Multiple roles of this kind become interlocking trajectories over time. These linked trajectories may define the life course of a parent and her child. Goode (1960) argues that an individual's set of relationships at any point in time is both "unique and overdemanding," requiring strategies that minimize demands by rescheduling transitions (such as entry into work, the birth of a second child), where possible. The synchronization of role demands may entail a spreading out of commitments or obligations, as in the transition to adulthood or in the family formation years. Among dual-earner couples, the timing of retirement has become a synchronization issue in working out an appropriate action for each partner and their relationship (O'Rand et al. 1991). The synchronization of lives is central to life-course planning in families.

Major transitions in the life course typically involve multiple life changes, from entry into the diverse roles of adulthood (Modell 1989) to later-life changes in work, residence, and family (Hareven 1978; Kohli 1986). These transitions may also

entail a sequence of phases or choice points. The transition to unwed motherhood thus involves premarital sexual experience followed by decisions not to have an abortion, not to give the child up for adoption, and not to marry the father. Causal influences vary across choice points. Early transitions can have developmental consequences by affecting subsequent transitions, even after many years and decades have passed. They do so through behavioral consequences that set in motion cumulative disadvantages or advantages, with radiating implications for other aspects of life (Furstenberg et al. 1987). For example, early teenage childbearing may curtail education and work-life prospects.

The social meanings of age give structure to the life course through age norms, sanctions, and age-graded relationships. In theory, a normative concept of social time specifies an appropriate time or age for marriage, childbearing, and retirement (Neugarten and Datan 1973). This concept also provides a guideline on the meaning of career advancement, whether accelerated or lagging relative to one's age. Empirical findings are beginning to cumulate on event timing, sequences, and durations, although the knowledge base is thin on causal mechanisms (Shanahan in press). Beyond these social distinctions, age has historical significance for the life course as it locates people in historical context according to birth cohorts.

Family connections invariably place the life course in a broader matrix of kinship relationships, one that extends beyond the boundaries of the immediate family to in-laws, grandparents, uncles, aunts, and cousins (Rossi and Rossi 1990). Within the life course of each generation, unexpected and involuntary events occur through life changes in related generations. Thus, a thirty-year old woman becomes a grandmother when her adolescent daughter has a first baby. People lose their status as grandchildren when their grandparents pass away, and their roles as sons or daughters when their parents die. They become the oldest generation in the family. Ties to family members are part of the normative regulation of life-course decisions.

The *life-cycle concept* is frequently used to describe a sequence of life events from birth to death, though its more precise meaning refers to an intergenerational sequence of parenthood stages over the life course, from the birth of the



children to their own departure from home and childbearing (O'Rand and Krecker 1990). This sequence, it should be noted, refers to a reproductive process in human populations. Within a life cycle of generational succession, newborns are socialized to maturity, give birth to the next generation, grow old, and die. The cycle is repeated from one generation to the next, though only within the framework of a population. Some people do not have children and consequently are not part of an intergenerational life cycle.

The life cycle is commonly known in terms of a family cycle, a set of ordered stages of parenthood defined primarily by variations in family composition and size (Hill 1970). Major transition points include marriage, birth of the first and the last child, the children's transitions in school, departure of the eldest and the youngest child from the home, and marital dissolution through death of one spouse. The stages are not defined in terms of age, as a rule, and typically follow a preferred script of a marriage that bears children and survives to old age, an increasingly rare specimen in view of the divorce rate. The life-cycle concept tells us about the sequence of family events, but it does not indicate how closely spaced the events are or when the sequence began in a woman's life, whether in adolescence or in the late thirties. A rapid sequence of births produces a different family process from that of widely dispersed births. The life stage of the mother also has relevance to the meaning of a birth sequence. Moreover, some women do not bear children.

*Life history* commonly refers to a lifetime chronology of events and activities that typically and variably combines data records on education, work life, family, and residence. A life history may also include information on physical health, social identity change, and emotional well-being. These records may be generated by obtaining information from archival materials or from interviews with a respondent. Some interviews are prospective and focus on the present; others are retrospective and enable the investigator to obtain information that was not collected in the past (Giele and Elder 1998). The accuracy of these reports of the past depends on the type of information requested. Subjective states in the past cannot be recovered accurately in retrospective reports. They are interpreted in terms of the present.

A retrospective life history or calendar is based on an age-event matrix (Freedman et al. 1988; Caspi et al. 1996). It records the age (year and month) at which transitions occur in each activity domain, and thus depicts an unfolding life course in ways uniquely suited to event-history analyses (Mayer and Tuma 1990) and to the assessment of time-varying causal influences. The advantages and disadvantages of retrospective life histories and prospective reports are discussed by Scott and Alwin (1998). In developing societies especially, retrospective life calendars are typically the only sources of information on prior life experience.

The term "life history" also refers to a self-reported narration of life, as in Thomas and Znaniecki's famous life history of Wladek, a Polish peasant, in *The Polish Peasant in Europe and America* (1918–1920). Narrative accounts are frequently recorded on tape and then transcribed. Another common approach assigns the interviewer a more active editorial role in actually putting together a life history. In *American Lives* (1993), Clausen interviewed six adults in their later years and prepared life histories, which the respondents later reviewed for accuracy. He has written about this method in two essays (1995, 1998) that discuss the difference between life stories and life histories.

This qualitative approach to life histories is increasingly considered one part of a multimethod approach to the study of lives. In their pathbreaking longitudinal study of juvenile delinquents, Laub and Sampson (1998) show how the qualitative life histories provide critical insight, when combined with quantitative data, on events that turned men's lives around toward productive work and good citizenship. Events that are turning points change the direction of lives, they are often "course corrections." Examples include marriage, military service, and higher education.

*Life span* specifies the temporal scope of inquiry and specialization, as in life-span psychology or sociology. A life-span study extends across a substantial period of life and generally links behavior in two or more life stages. Instead of limiting research to social and developmental processes within a specific life stage, such as adolescence or the middle years, a life-span design favors studies of antecedents and consequences. Sociologists tend to focus on the social life course, in which "life stage" refers to either socially defined positions,

such as the age of young adulthood, or to analytically defined positions, such as the career stage of men or women at age forty (i.e., where they are in their career at this age). Thus, men who differ in age when they encounter work-life misfortune occupy different life stages at the time.

Developmental stages and trajectories are the focus of life-span developmental psychology. Examples of a stage theory include Erik Erikson's (1963) psychosocial stages, such as the stage of generativity in the later years. Life-span developmental psychology gained coherence and visibility through a series of conferences at the University of West Virginia beginning in the late 1960s (Baltes and Reese 1984). The approach is defined by a concern with the description and explanation of age-related biological and behavioral changes from birth to death.

All the above concepts have a place in studies of the life course. Contemporary inquiry extends across the life span and frequently draws upon the life records and life cycles of successive generations. The life course takes the form of a multidimensional and intergenerational concept: a dynamic life pattern of interlocking trajectories and transitions, such as work, marriage, and being parents. Within this context, misfortune and opportunities are intergenerational as well as personal life events. Failed marriages and work lives can lead adult offspring back to the parental household and alter the parents's life plans for their later years (Goldscheider and Goldscheider 1993). Conversely, parents' economic setbacks and marital dissolution may impede their children's transition to adulthood by postponing higher education and marriage. Each generation is bound to fateful decisions and events in the other's life course.

### THE EMERGENCE OF LIFE-COURSE THEORY

When pioneering investigators followed children born before 1930 into their adult years, they encountered major limitations in conventional approaches to human development, including those associated with a child-based model. Three such limitations and their challenges, in particular, played a major role in the genesis of life course theory and appropriate methods (Elder 1998):

1. To formulate concepts that apply to development and aging across the life span as a replacement for child-based, growth-oriented accounts.
2. To conceptualize how human lives are socially organized and evolve over time.
3. To relate lives to an ever-changing society, with emphasis on the linking processes and mechanisms and the developmental effects of changing circumstances.

Responses to the first challenge in the 1970s led to the formulation of more life-span concepts of development and aging (Baltes and Baltes, 1990; Baltes et al. 1998), especially with the field of life-span developmental psychology. Life-span development is conceived as a life-long adaptational process. Some processes are discontinuous and innovative, while others are continuous and cumulative. Biological resources tend to decline over the life span, whereas cultural resources may increase, as in the growth of wisdom. Theorists stress the lifelong interaction of person and social context, the relative plasticity and agency of the aging organism, and the multidirectionality of life-span development (Lerner 1991). Psychologist Paul Baltes at the Berlin Max Planck Institute is a leading figure in the programmatic effort to study development and aging across the life span.

This emergence of life-span thinking on human development and aging occurred with little attention to a well-established "relationship" view of human lives. Dating back to the nineteenth century, social scientists have viewed the individual's life pattern in terms of multiple role sequences and their transitions (Cain 1964; Kertzer and Keith 1984). Changes in social roles, such as from dependency to marriage and parenthood, represent changes in social stage across the life cycle. Commitments to a line of action arise from obligations to significant others. Stable relationships ensure a measure of personal stability, just as entry into such relationships can stabilize a person's life and minimize involvement in unconventional and illegal activities (Robins 1966). A change in relationships may produce a turning point, a redirection of the life course.

Life-cycle theory helped to contextualize people's lives by emphasizing the social dynamic of "linked lives." These connections extend across the generations and across people's lives through

convoys of friends and relatives—others who remain part of their social network as they age (Kahn and Antonucci 1980). One of the earliest proponents of a life-cycle view of lives was sociologist William I. Thomas. With Florian Znaniecki, he used life-record data to study the emigration of Polish peasants to European and American cities around the turn of the century (1918–1920). The societies they left and entered presented contrasting “lines of genesis” or role sequences for individual lives.

For many decades, the life cycle perspective offered a valuable way of thinking about the social patterning of lives, though insensitive to age, time, and historical context. During the 1960s, the life-cycle approach began to converge with new understandings of age to draw upon the virtues of each tradition; of *linked lives* across the life span and generations, and of *temporality* through an age-graded sequence of events and social roles, embedded in a changing world. The emerging theory of the life course was informed as well by life-span concepts of human development.

Three discoveries of variation in lives were based on this new understanding of age, and together they gave rise to life-course thinking. First, studies in the 1960s began to identify substantial *variation in age patterns* across lives: Contrary to established views (Eisenstadt 1956), people of the same age varied in the pace and sequencing of transitions in their lives. During the late 1950s and early 1960s, Bernice Neugarten (1968) developed a research program that featured normative timetables and individual deviations from such expectations. Ever since this pioneering work and the growth of social demography, the study of differential timing and order among events has been one of the most active domains in life-course study. However, we still know little about age expectations in large populations.

Second, *social relations and kinship* especially emerged as a primary source of variation and regulation of life trajectories. Lives are lived interdependently among members of family and kin. The most significant integration of age and relationship distinctions is found in *Of Human Bonding* (Rossi and Rossi 1990). Using a three-generation sample in the Boston region, the study investigates the interlocking nature of the life

course within family systems, with particular focus on the relationship between individual aging and kin-defined relationships across the life span.

Third, the new work on aging made visible the role of *historical variation* as a source of life variations. Studies of social change and life patterns had been conducted up to the 1960s as if they had little in common, an assumption effectively challenged by Norman Ryder’s concept of the interaction of individual lives and history (1965). Ryder proposed the term “cohort” as a concept for studying the life course—the age at which the person enters the system. With its life-stage principle, Ryder’s essay provided a useful point of departure for understanding the interaction between social change and the life patterns of birth cohorts. The impact of a historical event on the life course reflects the life or career stage at which the change was experienced. The publication of *Aging and Society* (Riley et al. 1972) strengthened this sensitivity to the historical setting of lives through membership in a particular birth cohort.

The three streams of life-course theory (social relations–life cycle, age, and life-span concepts of development) came together in a study of children who were born in the early 1920s, grew up in the Great Depression, and then entered service roles in World War II. In *Children of the Great Depression* (Elder, [1974] 1999), the study began with ideas from the relationship tradition, such as generation and social roles, and soon turned to the analytic meanings of age for ways of linking family and individual experience to historical change, and for identifying trajectories across the life course. The study tested Ryder’s life-stage principle by comparing the effects of drastic income loss during the Great Depression on the family and individual experience of the Oakland study members with that of a younger cohort born toward the end of the 1920s (Elder et al. 1984). Consistent with the life-stage hypothesis, the younger boys in particular were more adversely influenced by family hardship, when compared to the older boys. But similar effects were not observed among the girls.

By the 1990s, the life course had become a general theoretical framework for the study of lives, human development, and aging in a changing society. This advance is coupled with the continued growth of longitudinal studies and the

emergence of new methodologies for the collection and analysis of life history data (Giele and Elder 1998; Mayer and Tuma 1990).

### PARADIGMATIC PRINCIPLES

Life course theory is organized around paradigmatic principles that guide inquiry on issues of problem identification, model formulation, and research design. Four principles are primary: 1) the interplay of human lives and development with changing times and places; 2) the social timing of lives; 3) the interdependence of lives; and 4) human agency in choice-making and actions (Elder 1998).

1. *The life course of individuals is embedded in and shaped by the historical times and places they experience over their lifetime.*

When societies are undergoing rapid change, different birth years expose individuals to historical worlds with different opportunity systems, cultures, and structural constraints. Historical influences on life trajectories take the form of a cohort effect when social change differentiates the life course of successive cohorts, such as older and younger men before the World War II. History also takes the form of a period effect when the impact of social change is relatively uniform across successive birth cohorts. Birth year and cohort membership locate people in relation to historical change, such as the economic recession of 1982–83, but they do not indicate actual exposure to the change or the process by which historical influences are expressed. Direct study of such change and its influences is essential for identifying explanatory mechanisms.

Life stage informs such mechanisms; people of unlike age and those who occupy different roles are differentially exposed to and influenced by particular types of social change. Four other components of an explanatory mechanism (for linking social change to lives) are worth noting (Elder 1998, pp. 959–61). One linking factor is defined by *social imperatives*, the behavioral demands or requirements of new situations. The more

demanding the situation, the more individual behavior is constrained to meet role expectations. Studies of worklives by Kohn and Schooler (1983) suggest that the establishment of a new set of occupational and workplace imperatives can alter how workers think and function. A second factor involves the dynamics of a *control cycle*. A loss of personal control occurs when people enter a new world and this sets in motion efforts to regain this control. These efforts may entail new choices that construct a different direction in life, a different life course. Historical influences are often expressed through a network of relationships. *Interdependent lives* are thus a linking factor in the explanatory analysis. The last factor involves *accentuation*. When a social transition heightens a prominent attribute that people bring to the new role or situation, the process accentuates the effect of the change. Our understanding of such change is enhanced by the *principle of timing in lives*.

2. *The developmental impact of a life transition or event is contingent on when it occurs in a person's life.*

This principle subsumes the concept of life stage; social change affects the individual according to when the exposure occurred over the life course. Recruitment into the military illustrates this point (Sampson and Laub 1996). Mobilization immediately after high school occurs before marriage and worklife obligations, and during World War II, entry into the service at this time was less disruptive than it was for males who entered in their late twenties or thirties. Studies of veterans from World War II indicate that late entry into the armed forces significantly increased the risk of divorce, worklife disruption, and an accelerated pattern of physical health decline (Elder and Chan 1999). By comparison, early entry provided access to educational and job opportunities without the costs of life disruption.

Across the life course, Neugarten's (1996) emphasis on the consequences of timing variations has been followed by

extensive research on differential timing patterns, from marriage and births to retirement (Shanahan in press). The impact of a life event, such as a personal loss, depends in part on when the event occurs. But to fully assess this impact, we must consider also the *principle of linked lives*.

3. *Lives are lived interdependently and social and historical influences are expressed through this network of shared relationships.*

Social relationships represent a vehicle for transmitting and amplifying the effects of stressful change, as in families under economic stress. For example, Depression hardship tended to increase the explosiveness of fathers who were inclined toward irritability before the crisis (Elder et al. 1986). And the more explosive they became under stress, the more adversely it affected the quality of marriage and effectiveness as a parent. Unstable family relations (marital, parent-child) became mutually reinforcing dynamics across the life course and generations. These dynamics tended to persist from one generation to the next through a process of individual continuity and intergenerational transmission.

Linked lives tend to transmit the life course implications of an ill-timed event in a person's life. One of the clearest examples of this phenomenon comes from a study of female lineages in Los Angeles (Burton and Bengtson 1985). The birth of a child to the teenage daughter of a young mother created a large disparity between age and kinship status, between being young and facing the prospects of grandparental obligations. Four out of five mothers of young mothers actually refused to accept these new child-care obligations, shifting the burden up the generational ladder to the great-grandmother, who in many cases was carrying a heavy load. By comparison, the women who became grandmothers in their late forties or so were eager for the new role; in this lineage, a daughter's timely transition to motherhood set in motion her mother's timely transition to grandmotherhood. In

both cases, the *principle of human agency* addresses the process by which lives and life courses are socially constructed.

4. *Individuals construct their own life course through the choices and actions they have taken within the constraints and opportunities of history and social circumstances.*

Human agency has been a central theme in the study of biographies. Lives are influenced by social structures and cultures, but human agency can shape lives through the choices that are made under such conditions. In Clausen's *American Lives* (1993), the central question is not how social systems made a difference in the life course, but why people made certain choices and thereby constructed their own life course. Clausen focused on the primary role of planful competence in late adolescence, that competent young people who think about the future with a sense of efficacy are more effective in making sound choices and in implementing them. This study found substantial support for this hypothesis in the lives of men in particular from adolescence to the retirement age of 65. However, longitudinal research (Shanahan et al. 1997) shows that planful competence makes a difference only when opportunities are available. Planful competence was not expressed in education or occupational career among men who entered the labor market in the depressed 1930s.

The emergence of life course theory and methods can be viewed as a response to pressing questions in the 1960s. The rapidly developing field of human development and aging needed a conceptual framework for thinking about the patterning of human lives in changing societies and its consequences. As life course theory addressed issue of this kind, it gained prominence as a theoretical orientation among fields of gerontology, criminology, social history, medical studies, developmental science, education, and social stratification. Noting such developments, Colby (Giele and Elder 1998, pp. x), refers to this approach as "one of the

most important achievements of social science in the second half of the 20th century." Only time will tell about the accuracy of this appraisal. In the meantime, the field of life course studies continues to flourish.

## REFERENCES

- Baltes, Paul B., and Margaret M. Baltes, eds. 1990 *Successful Aging: Perspectives from the Behavioral Sciences*. New York: Cambridge University Press.
- , and Hayne W. Reese 1984 "The Life-Span Perspective in Developmental Psychology." In Marc H. Bornstein and Michael E. Lamb, eds., *Developmental Psychology: An Advanced Textbook*. Hillsdale, N.J.: Erlbaum.
- , Ulman Lindenberger, and Ursula M. Staudinger 1998 "Life Span Theory in Developmental Psychology." In William Damon, general ed., and Richard M. Lerner, volume ed., *Handbook of Child Psychology, Volume 1: Theoretical Models of Human Development*, 5th ed. New York: Wiley.
- Bandura, Albert 1997 *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Burton, Linda M., and Vern L. Bengtson 1985 "Black Grandmothers: Issues of Timing and Continuity of Roles." In Vern L. Bengtson and Joan F. Robertson, eds., *Grandparenthood*. Beverly Hills, Calif.: Sage.
- Cain, Leonard 1964 "Life Course and Social Structure." In Robert E. L. Faris, ed., *Handbook of Modern Sociology*. Chicago: Rand McNally.
- Caspi, Avshalom, Terrie E. Moffitt, Arland Thornton, Deborah Freedman, James W. Amell, Honalee Harrington, Judith Smeijers, and Phil A. Silva 1996 "The Life History Calendar: A Research and Clinical Assessment Method for Collecting Retrospective Event-History Data." *International Journal of Methods in Psychiatric Research* 6:101–114
- Clausen, John A. 1993 *American Lives: Looking Back at the Children of the Great Depression*. New York: Free Press.
- . 1995. "Gender, Contexts, and Turning Points in Adults' Lives." In Phyllis Moen, Glen H. Elder, Jr., and Kurt Lüscher, eds., *Examining Lives in Context: Perspectives on the Ecology of Human Development*. Washington: APA Press.
- . 1998. Life Reviews and Life Stories." In Janet Z. Giele and Glen H. Elder, Jr., eds., *Methods of Life Course Research: Qualitative and Quantitative Approaches*, Thousand Oaks, Calif.: Sage.
- Colby, Anne 1998 "Forward: Crafting Life Course Studies." In Janet Z. Giele and Glen H. Elder, Jr., eds., *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Thousand Oaks, Calif.: Sage.
- Eisenstadt, Shmuel N. 1956 *From Generation to Generation: Age Groups and Social Structure*. Glencoe, Ill.: Free Press.
- Elder, Glen H., Jr. 1974 *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press. (25th Anniversary Edition, enlarged, 1999, Boulder, Col.: Westview.)
- . 1998 "The Life Course and Human Development." In William Damon, general ed., and Richard M. Lerner, volume ed., *Handbook of Child Psychology, Volume 1: Theoretical Models of Human Development*, 5th ed. New York: Wiley.
- , Avshalom Caspi, and Geraldine Downey 1986 "Problem Behavior and Family Relationships: Life Course and Intergenerational Themes." In Aage B. Sørensen, Franz E. Weinert, and Lonnie R. Sherrod, eds., *Human Development and the Life Course: Multidisciplinary Perspectives*. Hillsdale, N.J.: Erlbaum.
- , and Christopher Chan (1999) "Wars Legacy in Men's Lives." In Phyllis Moen and Donna Dempster-McClain, eds., *A Nation Divided: Diversity, Inequality and Community in American Society*. Ithaca, N.Y.: Cornell University Press.
- , Jeffrey K. Liker, and Catherine E. Cross 1984 "Parent-Child Behavior in the Great Depression: Life Course and Intergenerational Influences." In Paul B. Baltes and Orville G. Brim, Jr., eds., *Life-Span Development and Behavior*, vol. 6 New York: Academic Press.
- Erikson, Erik H. 1963 *Childhood and Society*, 2nd ed. New York: Norton.
- Freedman, Deborah, Arland Thornton, Donald Camburn, Duane Alwin, and Linda Young-DeMarco 1988 "The Life History Calendar: A Technique for Collecting Retrospective Data." *Sociological Methodology* 18:37–68.
- Furstenberg, Frank F., Jr., Jeanne Brooks-Gunn, and S. P. Morgan 1987 *Adolescent Mothers in Later Life*. New York: Cambridge University Press.
- Giele, Janet A., and Glen H. Elder, Jr., eds. 1998 *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Thousand Oaks, Calif.: Sage.
- Goldscheider, Frances K., and Calvin Goldscheider 1993 *Leaving Home before Marriage: Ethnicity, Familism, and Generational Relationships*. Madison: University of Wisconsin Press.
- Goode, William J. 1960 "A Theory of Role Strain." *American Sociological Review* 25(4):483–496.

- Hagestad, Gunhild O. 1990 "Social Perspectives on the Life Course." In Robert H. Binstock and Linda K. George, eds., *Handbook of Aging and the Social Sciences*, 3rd ed. New York: Academic Press.
- Hareven, Tamara K. 1978 *Transitions: The Family and the Life Course in Historical Perspective*. New York: Academic.
- Hill, Reuben 1970 *Family Development in Three Generations*. Cambridge, Mass.: Schenkman.
- Kahn, Robert L., and Toni C. Antonucci 1980 "Convoys Over the Life Course: Attachment, Roles, and Social Support." In Paul B. Baltes and Orville G. Brim, Jr., eds., *Life-Span Development and Behavior: Volume 3*. New York: Academic Press.
- Kertzner, David I., and Jennie Keith, eds. 1984 *Age and Anthropological Theory*. Ithaca, N.Y.: Cornell University Press.
- Kohli, Martin 1986 *The World We Forgot: A Historical Review of the Life Course*. In Victor W. Marshall, ed., *Later Life: The Social Psychology of Aging*. Beverly Hills, Calif.: Sage.
- Kohn, Melvin L., and Carmi Schooler 1983 *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, N.J.: Ablex.
- Laub, John H., and Robert J. Sampson 1998 "Integrating Quantitative and Qualitative Data." In Janet Z. Giele and Glen H. Elder, Jr., eds., *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.
- Lerner, Richard M. 1991 "Changing Organism-Context Relations as the Basic Process of Development: A Developmental Contextual Perspective." *Developmental Psychology* 27(1):27-32.
- Mayer, Karl U., and Nancy B. Tuma (eds.) 1990 *Event History Analysis in Life Course Research*. Madison, Wis.: University of Wisconsin Press.
- Merton, Robert K. 1968 *Social Theory and Social Structure*. New York: Free Press.
- Modell, John 1989 *Into One's Own: From Youth to Adulthood in the United States 1920-1975*. Berkeley: University of California Press.
- Neugarten, Bernice L. 1968 *Middle Age and Aging: A Reader in Social Psychology*. Chicago: University of Chicago Press.
- , and Nancy Datan 1973 "Sociological Perspectives on the Life Cycle." In Paul B. Baltes and K. W. Schaie, eds., *Life-Span Developmental Psychology: Personality and Socialization*, New York: Academic Press.
- 1996 *The Meanings of Age: Selected Papers of Bernice L. Neugarten*, edited, and with a foreword by Dail A. Neugarten. Chicago: University of Chicago Press.
- O'Rand, Angela M., and Margaret L. Krecker 1990 "Concepts of the Life Cycle: Their History, Meanings and Uses in the Social Sciences." *Annual Review of Sociology* 16:241-262.
- O'Rand, Angela M., John C. Henretta, and Margaret L. Krecker 1991 "Family Pathways to Retirement: Early and Late Life Family Effects on Couples' Work Exit Patterns." In Maximiliane Szinovacz, D. Ekerdt, and Barbara H. Vinick, eds., *Families and Retirement: Conceptual and Methodological Issues*. Newbury Park, CA: Sage.
- Riley, Matilda W., Marilyn E. Johnson, and Anne Foner, eds. 1972 *Aging and Society: A Sociology of Age Stratification*, vol. 3. New York: Russell Sage Foundation.
- Robins, Lee. 1966. *Deviant Children Grown Up*. Baltimore, Md.: Williams and Wilkins.
- Rossi, Alice S., and Peter H. Rossi 1990 *Of Human Bonding: Parent-Child Relations across the Life Course*. New York: Aldine.
- Ryder, Norman B. 1965 "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30(6):843-861.
- Sampson, Robert J., and John H. Laub 1996 "Socioeconomic Achievement in the Life Course of Disadvantaged Men: Military Service as a Turning Point, Circa 1940-1965." *American Sociological Review* 61(3):347-367.
- Scott, Jacqueline, and Duane Alwin 1998 "Retrospective versus Prospective Measurement of Life Histories in Longitudinal Research." In Janet Z. Giele and Glen H. Elder, Jr., eds., *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Thousand Oaks, Calif.: Sage.
- Shanahan, Michael J. (In press) "Pathways to Adulthood in Changing Societies; Variability and Mechanisms in Life Course Perspective." *Annual Review of Sociology*.
- Shanahan, Michael J., Glen H. Elder, Jr., and Richard A. Miech 1997 "History and Agency in Men's Lives: Pathways to Achievement in Cohort Perspective." *Sociology of Education* 70(1):54-67.
- Spilerman, Seymour 1977 "Careers, Labor Market Structure, and Socioeconomic Achievement." *American Journal of Sociology* 83(3):551-593.
- Thomas, William I., and Florian Znaniecki 1918-1920 *The Polish Peasant in Europe and America*, vol. 1-2. Urbana: University of Illinois Press.

GLEN H. ELDER, JR.

## LIFE CYCLE

The *life cycle* is the socially defined, age-related sequence of stages individuals pass through beginning with birth and ending with death. Underlying the life cycle is the recognition that humans are biological organisms that are born, mature, and die. As with other biological organisms, reproduction is a key feature of human maturation, ensuring the persistence of the species.

### FORMS OF THE LIFE CYCLE

In very simple societies, the life cycle may consist simply of two stages—infant and adult. Once infant survivorship is reasonably certain (typically by about age 6) young persons participate in adult work life, doing jobs that are suitable to their physical strength or as apprentices learning more complex skills. Work continues until death. But such a simple definition of the life cycle rarely endures the complexities attendant on reproduction. Among women, physical maturation separates childhood from the age when childbearing is possible. For men, marriage entails responsibility for supporting a family and guaranteeing their safety, something that typically must await completion of puberty and the achievement of economic viability. Even in societies with a low life expectancy, some adults survive to the point at which they are no longer able to work.

These examples illustrate how individual *social roles* (such as work or having a child) define a human life cycle that is more complex than the biological minimum. These roles are almost always defined as age related, and typically are also different for men and women. The concept of age-appropriate roles enables societies to regulate or prohibit behavior that is occurring “too early.” These societies also use the concept of age appropriateness to move individuals along in their maturation process, urging the adoption of a social role before it is “too late.”

**Age.** In most societies, chronological age is a handy proxy for maturity, with particular age groups assigned certain responsibilities and rights. In some societies (such as in postwar Japan) the age appropriateness of the sequence of social roles is rigorously defined by cultural values and enforced by

social institutions (such as schools or labor markets) which impose strict age rules on entry, promotion, and exit. In African age-set societies the system is even more rigid: Groups of persons born during contiguous years are defined as members of a particular age group (age set). These age sets experience together the transition from one life cycle stage to the next, under community traditions that specify the formal requirements and ceremonies necessary to move from one set of social roles to another. Typically this process is one of considerable dispute—the moving up of one age set causes all members of the society to move to the next life-cycle stage so that one group will have to give up preferred adult roles for old age.

#### **Age Stratification and Cohort Succession.**

Individual childbearing and the aging of individuals ready to assume new age-appropriate roles drive the societal process of age stratification and cohort succession (Riley 1985). The more complex the society, the more social roles that need to be filled. Most such roles in the society are gender linked and age stratified (defined as age appropriate and differing markedly from age to age). The use of chronological age rather than maturational capacity to construct the age-stratification system mandates that individuals as they age will move from one age stratum to the next, with an implicit societal mandate of assuming new roles. The birth of persons in contiguous years (what demographers call a “birth cohort”) reinforces the dynamic of the age-stratification system by producing new role entrants who can only be accommodated by the movement of all age strata to the next life-cycle stage. In the United States we can see this system at work in age-graded schooling: When one group achieves high school graduation, the remaining students are promoted from one grade to the next. This opens entry-level spaces for a new cohort to begin school. Universities develop a variety of incentives to get elderly faculty to retire so that newly trained and presumably more innovative faculty can be hired.

**The Life Cycle in Social Science.** To summarize, the lives of humans from birth to death are organized as socially defined, age-related sequence of stages individuals pass through over their lifetime. These stages are inherently age related, with individuals maturing from one life-cycle stage to the next. Reproduction is a key feature of human maturation, distinguishing the roles of men and



women and linked to the age-related biological capacity to bear children. An ongoing flow of new births ensures the persistence of human populations. Accommodating these new members of society also drives the dynamic of life-cycle change by necessitating the movement of earlier cohorts to more mature positions in the age-stratification system. This process of cohort succession is, in turn, a major source of societal innovation and change, as new cohorts take a fresh look at the content and form of the age strata they have just reached. In this way, the life-cycle concept links individual aging, the organization of roles in society, reproduction, and societal innovation and change.

### CONTRIBUTIONS OF THE LIFE CYCLE

The life cycle has proved to be a powerful and flexible tool for the analysis and explanation of human lives, used by researchers from a number of different disciplines. Anthropologists have focused on the process of socialization by which one age stratum is taught to succeed the next over the life cycle, linking social roles to the cultural system of beliefs and values. Age-set societies have been intensively described because of the very visible structure of age stratification and the explicit group-level patterns of life-cycle stages. Rites of passage are the symbolic counterparts of age-set transitions from one age stratum to the next, marking the personal change and announcing it to the entire community.

Developmental psychologists have used the life cycle as an organizing principle for specifying the steps in human development. The process of aging drives this principle, which is defined and structured by social organizations and individual roles. A prominent example is Erikson's eight stages of life (1968). In this model, psychosocial stages are identified, consisting of times in which opportunities for success and the risk of failure are present. For example, the young adult stage is marked by the capacity for intimacy versus isolation, and integration, while a choice between wisdom and despair marks old age. The passage from one of the eight life stages to the next is regarded as a *turning point* that is fraught with vulnerability and heightened potential.

The life cycle forced gerontologists to recognize that the study of old age in isolation from the

prior life cycle is not viable. The economic resources, health, knowledge, and family situations of the elderly result from the cumulating of life-cycle experiences. These same factors influence the chronological age at which people take on characteristics of the aged. Gerontology as a field has expanded to encompass the dynamics of life-cycle transitions. Immediately noteworthy when one adopts this aging approach is the fact that persons currently reaching old age have prior life-cycle experiences that may better prepare them for becoming old than did prior cohorts. The sociological interest in aging motivates gerontologists to attend to both the lifelong process of aging and potentially dramatic intercohort changes in successive cohorts of the aged.

**Life Cycle Squeeze.** Economists have relied on the life cycle and the gendered division of labor to study household and family economics. Wages follow a curvilinear pattern over the life cycle: Young workers receive the lowest wages; wages increase over the life cycle, peaking at midlife; while workers older than 55 tend to experience stability or even a decline in earnings. Women's earnings show much less of an age profile, both because young women frequently interrupt or reduce labor-force involvement when children are born, and because women's jobs are less likely to take the form of careers in which progression upward from one job to the next occurs.

This life-cycle pattern of earnings does not always match family income needs. Early in the life cycle, children are net consumers of income. As societies require a more educated population, youth and adolescents also become net income consumers. While this has long-term payoff for the society and for new cohorts of workers, it increases the costs and reduces the economic value of children to families. These costs most often occur when workers are at the low point in their earnings. Later in the life cycle, earnings are higher but the cost of a college education and assisting children in getting started also may be high. At the same time, elderly parents may become a social and financial obligation for children, leading to the powerful concept of the "life-cycle squeeze."

Cohorts facing this new mix of obligations with the traditional earnings profile have acted to reduce desynchronies in the stages of the life cycle.

Credit mechanisms (for example, long-term mortgages, home equity loans to pay for college education) smooth income and costs over the lifetime. The intercohort upgrading of the situations of the elderly in the United States means that parents become dependent on their adult children at an older age. The government has intervened to bear much of the cost of the elderly. Life-cycle pressures have also resulted in intercohort changes in the content of the life cycle of men and women—delays in marriage and first birth, a reduction in family size, the shift to two-earner families, and increases in divorce.

Sociologists have devoted considerable attention to the cohort-level study of life-cycle transitions. Turning points imply inevitability and potential crisis. Sociologists study the form of the transition (for example, cohabitation or marriage), whether or not a transition occurs (for example, parenthood), and the average age and variability in age of a transition across individuals of the same cohort. This approach recognizes (building on the idea of an age-stratification system) that a variety of transitions are crosscutting (for example, work and marriage, birth of a first child and marriage).

**The Life Cycle in Demographic Models.** It is the educational, labor-force, and family outcomes that interest demographers. The study of family life has proceeded with the measurement of marriage and then the progression to first birth, second birth, and so forth (taking into account both the number and the timing of births). A special tool called *parity progression analysis* has enabled demographers to identify turning points in fertility decisions, and how these have changed across cohorts. This approach to the study of demographic life-cycle stages, along with the recognition that fertility is inherently a biological process, has led demographers to develop population models of fertility that take into account marriage patterns, the level of marital fertility, and birth limitation. The life-cycle model also has informed research on age patterns of migration, and its regularities across time and place.

The life-cycle perspective has produced a variety of unexpected results. The American baby boom of the late 1940s and the 1950s was largely due to the temporal coincidence of childbearing by successive cohorts, rather than to a dramatic increase in family size. During the baby boom,

women 35 and older made up childbearing that had been delayed by the Great Depression and World War II. Women reaching adulthood during the baby boom years responded to favorable economic conditions for young families by having their first child at a younger age and having subsequent children more quickly.

**Family Life Cycle.** Family sociologists made a great leap in developing the family life cycle as a variant form of the life cycle (Glick 1965). The family life cycle is unusual in that it focuses on family formation and childbearing, ignoring such linked transitions as completion of schooling and work. The family life cycle stretched the life-cycle model to incorporate role changes associated with the transitions of other individuals. (For example, a husband makes the transition to marriage at the same time as the wife. Only when all children grow up and leave the home does the family experience an “empty nest.”) The family life cycle became a predominant research paradigm in family studies.

The family life cycle can be a useful analytic tool for understanding the succession of family roles in populations in which families predominate over individual interests. The family life-cycle model works only for those populations in which marriage precedes childbearing, the ages of each are specified within a narrow time band, and marriages do not end (by widowhood or divorce) before the last child leaves home.

None of these assumptions are even approximately satisfied for the United States. First births often precede marriage (among blacks this is the typical pattern). Many couples postpone childbearing within marriage, and as many as one-fifth remain voluntarily childless. Over half of all first marriages end within twenty years. Remarriage often follows. This degree of inconstancy in household membership begs the question of how to define the family whose life cycle is being described, and followed over time. The family life cycle is now widely regarded as a useless conceptual tool because it utterly fails to capture the realities of contemporary family life.

## LIFE CYCLE AND LIFE COURSE

The life cycle defines pathways for individuals as they age from birth to death, specifying usual expectations about the sequence and timing of

roles (for example, a first birth when married and at age 18 or older). Empirical research that uses the life cycle to analyze the lives of population *cohorts* typically find that these life-cycle stages, as socially defined and demarcated, follow expected patterns.

Research on the transitions of *individuals* over their lifetime has demonstrated the essential incorrectness of this supposition. For many individuals, childbearing precedes marriage, parenthood occurs when the parents are not yet economically self-sufficient, and adult children return to their parental household after they have assumed (and sometimes failed at) adult family and economic roles. This has caused researchers to consider whether new life-cycle patterns are emerging, or whether a group of individuals is somehow “deviant” from the established life cycle.

Clearly there have been marked intercohort changes in the life cycle, reflecting the varying opportunity structures of time periods and cultural change. For example, the availability of the GI Bill for college education and interruptions in education associated with wartime military service allowed many men to marry and have children before they finished school and became economically established. In recent cohorts of young women there is a decreasing emphasis on the necessity of marrying before having a child, and, on the part of all adults, a greater readiness to assist rather than condemn single mothers.

An even greater source of departure from the population life-cycle model is the large number of persons in each cohort who never make a transition to a given life-cycle stage (e.g., those not marrying or not becoming parents), who retreat from a given life-cycle stage to an earlier stage (e.g., fathers who divorce and abandon their families, and retired persons who return to work), and who are not part of the typical life cycle (e.g., the severely disabled, persons who die before reaching old age). These features of the life cycle vary by such primary sociological variables as social class origin, education, race and ethnic group, and place of residence.

Among Americans the seeming conformity of cohorts to the life-cycle patterns masks the overwhelming number and frequency of individual departures. In this situation the life cycle seems to be a far less useful analytic device. Social scientists

have adopted in its place the more sophisticated and flexible “life-course” perspective. The life course sees individual lives as a series of trajectories (such as family or career) that are socially recognized and defined. Age is significant in the life-course approach because it is an indicator of biological aging and locates individuals in historical context through birth cohorts. The social meaning of age helps define life-course pathways (recognized routes of trajectories) through age norms and sanctions, and social timetables for the occurrence and order of events. Transitions (leaving home, getting a job, marrying) define trajectories. Interlocking transitions and their trajectories lead to multiple roles that define the individual life course from birth to death.

Because of the emphasis on variations in trajectories across individuals, every individual life course has the potential to be unique. Much of the population-level research with the life-course model has focused on transitions—the proportion of cohorts making a transition, average age at transitions, and the range of ages at which cohorts typically make these transitions. The life cycle gives analytic meaning to these life-course transitions by providing a standard against which to measure how transitions vary across cohorts and differ among key population groups within cohorts.

Causal analyses of the life course are usually done at the individual level, typically with a class of statistical methods called “event history” or “hazards” models. These statistical methods enable investigators to examine empirical data on individual transitions, modeling the age-graded pattern of transitions from one social role to another and identifying “heterogeneity” (sources of variation in transitions at the individual level).

*The Necessity for Life Cycle in Studies of the Life Course.* This points to a dilemma for social scientists—the very life-cycle model that the life-course approach undermines provides the essential theoretical framework that gives meaning to individual behaviors. Because social scientists are part of the societies they study, they also carry in their own heads models of the life cycle—what should be done when it should be done, and what denotes success or failure. The life-course perspective on pathways that define typical trajectories and the social meanings of age capture the essence of a life-cycle model. The apparent tension between the

life-cycle and life-course conceptualizations is perhaps overdrawn.

The life-cycle model retains many valuable features that are typically missing from life-course studies. Life-course transitions focus on individuals at particular times during their lives (adolescents becoming adults, older workers becoming disabled or retired). While past experiences and current opportunities are often included in life-course models, the life-course perspective has not lent itself well to viewing transitions at particular ages in the context of the lifelong process of aging, an idea that is innate in the life-cycle approach. Economic research, which is theoretically driven and uses a life-cycle model, has been most successful at integrating findings from transitions at a given age into the lifetimes of individuals.

With such an approach it is also possible to simulate the effects of changes in transition rates (resulting from heterogeneity in rates and changes in population composition) on cohort-life cycle behaviors. The life-course perspective typically views cohorts as proxies for age-specific experiences with the social structure; the life-cycle model brings a necessary emphasis on intercohort change as a method of social innovation.

The life-course approach emphasizes variability in transitions and trajectories to such an extent that many social scientists have neglected regularities and consistencies in behaviors, and the advantages they may entail. There is a great deal of research on adolescent mothers but relatively little research on why married women have children. There is more interest in unemployment and poverty than in the advantages of paid employment and career lines. While children of welfare mothers disproportionately go on to become welfare mothers themselves (the subject of much research), the overwhelming majority grow up to be free of welfare dependence (a subject about which we know very little). Nor do we have a strong explanation of the reasons persons marry. The life-cycle approach draws attention to these questions of social organization, and the matching of individual behaviors to necessary social roles.

#### FUTURE OF LIFE CYCLE

The life cycle thus remains a viable and valuable conceptual tool for studying human lives. In much

of the developing world, transitions and trajectories are sufficiently universal and age regulated that the life-cycle model remains a highly useful tool for social science. In societies where such regularities are no longer the norm, the life-course approach is the more appropriate. To be meaningful, the life course must be interpreted in light of the life cycle—the underlying beliefs about the shape and timing of the life stages to understand the social meanings of age, identify alternative pathways for life trajectories, draw attention to the strong regularities in transition behaviors and linkages, and direct attention to intercohort stability and change. The concept of the life cycle thus will continue to be a valuable and necessary tool for the social sciences.

(SEE ALSO: *Life Course*)

#### Further Readings

- Erikson, Erik H. 1968 "Life Cycle." Pp. 286–292 in David L. Sills, ed., *International Encyclopaedia of the Social Sciences*, vol. 9. New York: Mcmillan Free Press.
- Glick, Paul C. and Robert Parke, Jr. 1965 "New Approaches in Studying the Life Cycle of the Family." *Demography* 2:187–202.
- Greenwood, M. J. 1997 "Internal Migration in Developed Countries." Pp. 647–720 in Mark R. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, vol. 1B. Amsterdam: Elsevier.
- Hotz, V. J., J. A. Klerman, and R. J. Willis 1997 "The Economics of Fertility in Developed Countries: A Survey." Pp. 275–347 in Mark R. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, vol. 1A. Amsterdam: Elsevier.
- O'Rand, Angela M. and Margaret L. Krecker 1990 "Concepts of the Life Cycle: Their History, Meanings, and Uses in the Social Sciences." *Annual Review of Sociology* 16:241–262.
- Riley, Matilda White 1985 "Age Strata in Social Systems." Pp. 369–411 in Robert H. Binstock and Ethel Shanas, eds., *Handbook of Aging and the Social Sciences*, 2nd ed. New York: Van Nostrand Reinhold.

DENNIS P. HOGAN

## LIFE EXPECTANCY

*Life expectancy* (or the expectation of life) is the average length of life remaining to be lived by a

population at a given age. It is computed in the process of building a life table and can be computed for any age in the life table. Life expectancy at birth is the most commonly presented value because this measure provides a succinct indicator of mortality that reflects mortality conditions across the age range and is unaffected by the age structure of the actual population and thus can be compared across populations. The symbol used to represent life expectancy is  $e_x$  where  $x$  represents an exact age.

### LIFE EXPECTANCY IN THE UNITED STATES

In 1996, life expectancy at birth,  $e_0$ , in the United States was 76.1 years; at age 65,  $e_{65}$  was 17.5 years; and at age 85,  $e_{85}$  was 6.1 years (Anderson 1998). These figures can be interpreted to mean that if a baby born in 1996 were exposed to the mortality conditions existing at each age of the life span in 1996, the baby with an average length life would live 76.1 years.

### PERIOD AND COHORT VALUES OF LIFE EXPECTANCY

The 1996 U.S. life table is a period life table, based on cross-sectional data collected over a year; thus, this life table indicates the mortality experience of a hypothetical cohort. No actual cohort ever experiences the mortality in a period or cross-sectional life table; rather, the table indicates mortality conditions if the mortality levels of each age group at the period of time used as a reference were experienced by the hypothetical cohort. Because mortality has been falling over time, period life tables for a cohort's year of birth have indicated an average expected length of life that is lower than that actually achieved by the cohort. For instance, in 1900 the cross-sectional life table for the United States showed life expectations of 46 for males and 49 for females. On the basis of their actual experience up through the age of 80, the 1900 birth cohort had an average length of life of 52 years for males and 58 years for females (Faber and Wade 1983).

Generation or cohort life tables, like the one mentioned above, based on the experience of an actual cohort are sometimes constructed. These

indicate the average length of life actually lived after specific ages for a real cohort. The major difficulty faced in building cohort life tables is obtaining population and death data for a cohort from birth until the last survivors have died—over a 100-year period.

A mistaken notion held by many people is that life expectancy at birth is a good indicator of the age at which an older individual will die. This notion has undoubtedly led to some poor planning for old age because a person who has already reached older adulthood on average will die at an age that exceeds life expectancy at birth by a significant amount. As mentioned above, expectation of life in 1996 was 17.5 years for 65-year-olds, 11.1 for 75-year-olds, and 6.1 for 85-year-olds. With this number of years remaining to be lived on average, 65-year-olds should expect to live to 83 on average. Those who live to 75 should expect to live to 86, and those who live to 85 can expect to live to 91 on average. While expectation of life decreases as age increases, the expected age at death increases for those who survive.

### CHANGES IN LIFE EXPECTANCY OVER TIME

As noted above, life expectancy has been increasing over time. This has probably been going on since some time in the last half of the nineteenth century, although reliable data for large sections of the country are not available to track the increase before 1900. In 1900, life expectancy at birth for both sexes was 47.3 years (U.S. Bureau of the Census 1975). This indicates an increase in life expectancy between 1900 and 1996 of 28.8 years. Most of this increase in life expectancy since 1900 is due to declines in mortality among infants and children. These mortality declines were primarily due to the diminishing force of infectious and parasitic diseases which were the most important causes of death among children.

Because life expectancy was low in the past, people often hold the mistaken notion that very few people ever reached old age under high mortality conditions. Yin and Shine (1985) have demonstrated that this mistaken notion was so prevalent that it was commonly incorporated into gerontology textbooks. The fact is that even under conditions of low life expectancy, once childhood

is survived, the chances of living to old age are quite high. This is indicated by the fact that life expectancy at the older years has not increased over time nearly as much as life expectancy at birth. For instance, while life expectancy at birth for white males has increased almost 26 years since 1900, from 48.2 to 73.9 years, life expectancy for white males at age 40 has increased almost 9 years between 1900 and 1996, from 27.7 years to 36.4 years; at age 70, the increase for males has been just over 3 1/2 years, from 9.0 to 12.6 (Anderson 1998).

It should be noted, however, that in the past three decades the pace of improvement in life expectancy at the oldest ages has increased. In 1970 expectation of life for white males at age 70 was 10.5 years, indicating an improvement of 1.5 years in the 70 years between 1900 and 1970. Between 1970 and 1998, the increase was 2 years—significantly greater than the improvement during the first seven decades of the century. This reflects the new era of mortality decline in which decreases in mortality are due to decreased mortality from chronic conditions and are concentrated among the old.

A number of authors have studied the relationships between changes in age-specific mortality and life expectancy. Vaupel (1986) concludes that a reduction in the force of mortality of 1 percent at all ages would not produce as much gain in life expectancy today as it did in 1900. This is because we have already made so much progress in lowering infant and child mortality, the ages that have the greatest effect on life expectancy. Vaupel also shows that as mortality moves to lower levels, more progress is made in increasing life expectancy from mortality declines at older ages rather than at younger ages. At the level of mortality now experienced in the United States, much of the future increase in life expectancy will come from mortality declines occurring at ages over 65. This is true because of the prior success in reducing mortality at earlier ages to such low levels.

#### CALCULATION OF LIFE EXPECTANCY WITHIN THE LIFE TABLE

These observations about changes in life expectancy should make clear that life expectancy at birth is heavily weighted by mortality conditions at the

youngest ages. A brief explanation of the life table and how life expectancy is calculated demonstrates why this is the case.

The life table is a statistical model that provides a comprehensive description of the mortality level of a population. Life table measures are particularly valuable because they are succinct indicators of mortality that reflect mortality conditions across the age range, are unaffected by the age structure of the actual population, and thus can be compared across populations. Life table measures can also be used to describe the characteristics of the stationary population that would result from an unchanging schedule of age-specific mortality rates in a closed population with a constant number of births.

There are a number of functions that appear in most life tables and for which conventional notation is widely recognized:  $q_x$ ,  $l_x$ ,  $d_x$ ,  $L_x$ ,  $T_x$ , and  $e_x$ . Each of these measures provides information useful in describing some aspect of the mortality conditions and/or characteristics of the stationary population. The definitions and interpretations of the life table functions follow below. In order to clarify the interpretation of the abridged life table functions, the life table for the U.S. population for 1996 is used as an example (Table 1).

${}_nq_x$  is the probability of dying between exact age  $x$  and  $x + n$ . As shown in Table 1, the probability of dying in the first year of life is 0.00732. This is higher than at subsequent ages until age 60 to 65, when the probability of death is 0.06649.

$l_x$  is the number of survivors reaching exact age  $x$  out of the original life table population. The size of the original life table population, the radix or  $l_0$ , is usually assumed to be 100,000; however, this is a convention and other values can be used. Mortality conditions in 1996 were such that out of 100,000 births, 99,268 would reach age 1. This column of the life table can be used to compute how many people who reach a given age will survive to a later age. For instance, among the 80,870 people who reach age 65, 33,629 people or 42 percent will reach age 85 with mortality conditions as shown in Table 1.

${}_nd_x$  is the number of deaths in the life table population between exact age  $x$  and  $x + n$ .

LIFE EXPECTANCY

**Abridged Life Table: United States, 1996**

Age Interval	Proportion Dying	Of 100,000 Born Alive		Stationary Population		Average Remaining Lifetime
Period of Life between Two Exact Ages Stated in Years	Proportion of Persons Alive at Beginning of Age Interval Dying during Interval	Number Living at Beginning of Age Interval	Number Dying during Age Interval	In the Age Interval	In This and All Subsequent Age Intervals	Average Number of Years of Life Remaining at Beginning of Age Interval
(1) $X$ to $X+n$	(2) ${}_nq_x$	(3) $l_x$	(4) ${}_nd_x$	(5) ${}_nL_x$	(6) $T_x$	(7) $e_x$
<b>Total</b>						
0-1	0.00732	100,000	732	99,370	7,611,825	76.1
1-5	0.00151	99,268	150	396,721	7,512,455	75.7
5-10	0.00097	99,118	96	495,329	7,115,734	71.8
10-15	0.00118	99,022	117	494,883	6,620,405	66.9
15-20	0.00390	98,905	386	493,650	6,125,522	61.9
20-25	0.00506	98,519	499	491,372	5,631,872	57.2
25-30	0.00544	98,020	533	488,766	5,140,500	52.4
30-35	0.00710	97,487	692	485,746	4,651,734	47.7
35-40	0.00944	96,795	914	481,820	4,165,988	43.0
40-45	0.01283	95,881	1,230	478,549	3,684,168	38.4
45-50	0.01801	94,651	1,705	469,305	3,207,619	33.9
50-55	0.02733	92,946	2,540	458,779	2,738,314	29.5
55-60	0.04177	90,406	3,776	443,132	2,279,535	25.2
60-65	0.06649	86,630	5,760	419,530	1,836,403	21.2
65-70	0.09663	80,870	7,814	385,659	1,416,873	17.5
70-75	0.14556	73,056	10,634	339,620	1,031,214	14.1
75-80	0.21060	62,422	13,146	280,047	691,594	11.1
80-85	0.31754	49,276	15,647	207,474	411,547	8.4
85 and over	1.00000	33,629	33,629	204,073	204,073	6.1

Table 1

SOURCE: R. N. Anderson 1988 *United States Abridged Life Tables, 1996. National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 47(13):5. Hyattsville, Md.: National Center for Health Statistics.

In the sample life table, 732 of the 100,000 births would die between ages 0 and 1 and 7,814 would die between ages 65 and 70.

${}_nL_x$  is the total number of years lived by the life table population between exact age  $x$  and  $x + n$ . Between birth and age 1, the life table population represented in Table 1 would live 99,370 years. This column also can be interpreted as the number of people in the stationary population at each year of age.

$T_x$  is the total number of years lived after exact age  $x$  by the life table population surviving to age  $x$ , or the number of people in the stationary population age  $x$  and older. The 100,000 entrants to the life table in Table 1 would live a total of 7,611,825

years, and the 80,870 who reach age 65 would live a total of 1,416,873 more years.

$e_x$  is the expectation of life at exact age  $x$  or the average length of life remaining to be lived for the life table population which survives to exact age  $x$ .  $e_x$  is computed from the  $T_x$  and  $l_x$  columns of the life table:  $e_x = T_x/l_x$ . As indicated earlier, at birth, the life table population in Table 1 has a life expectancy of 76.1 years.

**DIFFERENTIALS IN LIFE EXPECTANCY**

There are large differentials in life expectancy among demographic and socioeconomic groups in the United States. Males have lower life expectancies than females throughout the age

range. Males' lower chances for a longer life are thought to result from a combination of biological differences and lifestyle factors. In 1996,  $e_0$  was 73.1 for males and 79.1 for females (Anderson 1998). By age 50, the difference is narrowed to 4.3 years, with a life expectancy of 27.2 for men and 31.5 for women. At age 85, men can expect to live another 5.4 years, while women can expect to live 6.4 years.

There is also a significant difference in life expectancy between whites and African Americans in the United States. This is assumed to result primarily from the difference in socioeconomic status and accompanying life circumstances that exist between African Americans and whites in the United States. In 1996, life expectancy at birth was 76.8 for whites and only 70.2 for blacks. At age 65, white life expectancy was 17.6 years; while for blacks of that age, it was 15.8 years. At the oldest ages, a crossover in mortality rates by race has been observed in the past. After the age of crossover, African-American mortality rates are lower than white mortality rates. In 1987 this was true at ages above 83. In the past, this crossover has shown up repeatedly in comparisons of African-American and white mortality in the United States and has been attributed to the "survival of the fittest" among the black population (Manton and Stallard 1981). Recently, however, doubt has been raised as to whether the crossover is real or is a statistical artifact resulting from age misstatement by older African Americans in both the census and vital records of deaths (Coale and Kisker 1986; Elo and Preston 1994; Preston et al. 1996). Interestingly, Hispanics appear to have life expectancy values that are higher than non-Hispanic whites (Anderson et al. 1997).

#### INTERNATIONAL DIFFERENCES

In general, the life expectancy of a country is related to its level of socioeconomic development. Most countries that are classified as "more developed" have higher levels of life expectancy at birth than most of the countries classified as "developing"; however, within each of these groups of countries there is quite a bit of variability in life expectancy. While the United States has a high level of life expectancy compared to that of the developing countries of the world, the United States ranks quite low in life expectancy among

developed countries and relative to its income level. A recent United Nations listing of the developed countries by level of life expectancy at birth ranks U.S. males as nineteenth and U.S. females as fourteenth (United Nations 1997). The countries with higher life expectancy for women include Japan and the Scandinavian countries. For men, most European countries including some in southern Europe have higher life expectancies at birth than the United States. The low ranking of the United States is attributed, in part, to the inequities in mortality among subgroups of the population, especially the high level among African Americans, and also to the high level of violent deaths. In recent years Japan has become the world leader in life expectancy at birth with values of  $e_0$  of 76.4 for men and 82.9 for women in 1995 (Ministry of Health and Welfare, Japan 1999). These values exceed 1996 U.S. values by 3.3 years for men and 3.8 years for women. The success of the Japanese in raising their levels of life expectancy has been due to large declines in mortality from cerebrovascular disease and maintenance of low levels of heart disease relative to other developed countries (Yanagishita and Guralnik 1988).

#### RELATED CONCEPTS

There are some other concepts that are related to life expectancy and are sometimes confused with life expectancy. One is "life span." The *life span* of a species is the age to which the longest-lived members survive. The life span of humans is thought to be approximately 115 years; however, Madame Jeanne Calment, whose age was well documented, died in 1997 at the age of 122. Current thinking is that while life expectancy has increased dramatically over the last century, the life span of humans has not changed over time; however, this does not mean it will never change. If discoveries are made in the future that enable us to retard the aging process, it may be possible to lengthen the human life span in the future.

"Life endurancy" is a related concept that, like life expectancy, is computed from the life table. This is the age at which a specified proportion of the life table entry cohort is still alive. For instance, in 1990 the age at which 0.1, or 10 percent, of the life table population remained alive was 90 years for men and 96 years for women. Life endurancy has been increasing over time and is



expected to continue to change with changes in survival rates. In 1900 the 10 percent survival age was 81 and 82 for men and women, respectively (Faber and Wade 1983).

Finally, "healthy or active life expectancy" is a subset of total life expectancy. Total life expectancy at any age is the sum of two parts: healthy life expectancy and unhealthy life expectancy. While the concept of health life expectancy was introduced in the 1960s (Sanders 1964) and developed in the 1970s (Sullivan 1971a, 1971b), it has only become widely adapted by governments and international organizations in the 1990s.

Interest in healthy life expectancy has grown recently as people have recognized that gains in total life expectancy today may not mean the same thing as in the past. Past gains in life expectancy came about largely because fewer people died of infectious diseases, either because they did not get the diseases or they received treatment that prevented death. People thus saved from death were generally free of the disease. Under these circumstances gains in life expectancy were accompanied by better health in the population surviving. Now, with gains in life expectancy being made because of declining death rates from chronic diseases especially among the old, it is not clear that the surviving population is a healthier population. This is because generally there is no cure for the chronic diseases, and for many their onset has not yet been prevented. People may be saved from death but they live with disease. This is the basis for questioning whether the additions to life expectancy are healthy or unhealthy years.

Crimmins and colleagues (1997) estimated that healthy life expectancy or disability-free life expectancy at birth in the United States in 1990 was 58.8 years for men and 63.9 years for women. The difference between blacks and whites in disability-free life expectancy at birth was even greater than the difference in total life expectancy. In 1990 black disability-free life expectancy for males at age 20 was 37.9 years while that for whites was 45.8 years (Hayward and Heron 1999). Studies that addressed the issue of changes in healthy life expectancy for the 1970 and 1980 period generally found that when *healthy life* was defined as nondisabled life, active life expectancy had not increased (Wilkins and Adams 1983; Crimmins et al. 1989). More recent studies have found increases in active

life expectancy (Crimmins et al. 1997; Robine and Mormiche 1994).

*Healthy life expectancy* can be defined in many ways. Examples include average length of life free from a disability that causes a person to alter his or her normal activity; average length of life free of dependency on others for the performance of basic activities necessary to living, such as eating, bathing, and getting in and out of bed; and average length of life without disease (Bebbington 1988; Colvez et al. 1986; Crimmins et al. 1997; Crimmins et al. 1994; Rogers et al. 1989). Some measures of healthy life combine multiple indicators of health using weights; for instance, the U.S. National Center for Health Statistics measure combines self-assessed health and disability in its indicator of healthy life (Erickson et al. 1995).

There are multiple methodological approaches to estimating health expectancy. Most can be described under one of two headings: the Sullivan method or the multistate method (Sullivan 1971a; Schoen 1988). Microsimulation techniques have also been employed recently (Laditka and Wolf 1998).

#### REFERENCES

- Anderson, R.N. 1998 *United States Abridged Life Tables, 1996. National Vital Statistics Reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 47(13). Hyattsville, Md.: National Center for Health Statistics.
- , K. D. Kochanek, and S. L. Murphy 1997 *Report of Final Mortality Statistics, 1995. Monthly Vital Statistics Report*, 45(11), suppl. 2. Hyattsville, Md.: National Center for Health Statistics.
- Bebbington, A. C. 1988 "The Expectation of Life without Disability in England and Wales." *Social Science and Medicine* 27:321–326.
- Coale, A. J., and E. E. Kisker 1986 "Mortality Crossovers: Reality or Bad Data?" *Population Studies* 40:389–401.
- Colvez, A., J. M. Robine, D. Bucquet, F. Hatton, B. Morel, and S. Lelaidier 1986 "L'espérance de Vie Sans Incapacité en France en 1982." (Expectation of Life without Disability in France in 1982.) *Population* 41:1025–1042.
- Crimmins, E. M., M. D. Hayward, and Y. Saito 1994 "Changing Mortality and Morbidity Rates and the Health Status and Life Expectancy of the Older Population." *Demography* 31:159–175.
- Crimmins, E. M., Y. Saito, and D. Ingegneri 1989 "Changes in Life Expectancy and Disability-Free Life

- Expectancy in the United States." *Population and Development Review* 15:235–267.
- 1997 "Trends in Disability-Free Life Expectancy in the United States, 1970–1990." *Population and Development Review* 23:555–572.
- Elo, I. T., and S. H. Preston 1994 "Estimating African-American Mortality from Inaccurate Data." *Demography* 31:427–458.
- Erickson, P., R. Wilson, and I. Shannon 1995 "Years of Healthy Life." *Healthy People 2000: Statistical Notes* 7:1–15.
- Faber, J., and A. Wade 1983 *Life Tables for the United States: 1900–2050*. Actuarial Study No. 89. Washington: U.S. Department of Health and Human Services, Social Security Administration, Office of the Actuary.
- Hayward, M. D., and M. Heron 1999 "Racial Inequality in Active Life among Adult Americans." *Demography* 36:77–91.
- Laditka, S. B., and D. A. Wolf 1998 "New Methods for Modeling and Calculation Active Life Expectancy." *Journal of Aging and Health* 10:214–241.
- Manton, K. G., and E. Stallard 1981 "Methods for Evaluating the Heterogeneity of Aging Processes in Human Populations Using Vital Statistics Data: Explaining the Black/White Mortality Crossover by a Model of Mortality Selection." *Human Biology* 53:47–67.
- Ministry of Health and Welfare, Japan 1999 "Abridged Life Tables for Japan, 1995." <http://www.mhw.go.jp/english/database/lifetbl/part6.html>.
- Preston, S. H., I. T. Elo, I. Rosenwaike, and M. Hill 1996 "African-American Mortality at Older Ages: Results of a Matching Study." *Demography* 33:193–209.
- Robine, J. M., and P. Mormiche 1994 "Estimation de la Valeur de l'Espérance de Vie Sans Incapacité en France en 1991." (Estimation of Expectation of Life without Disability in France in 1991.) *Les Français et Leur Santé* 1:17–36.
- Rogers, R., A. Rogers, and A. Belanger 1989 "Active Life among the Elderly in the United States: Multistate Life Table Estimates and Population Projections." *Milbank Quarterly* 67:370–411.
- Sanders, B. 1964 "Measuring Community Health Level." *American Journal of Public Health* 54:1063–1070.
- Schoen, R. 1988 *Modeling Multigroup Population*. New York: Plenum.
- Sullivan, D. F. 1971a "A Single Index of Mortality and Morbidity." *HSMHA Health Reports* 86:347–354.
- 1971b *Disability Components for an Index of Health. Vital and Health Statistics* 2 (42). Rockville, Md.: National Center for Health Statistics.
- United Nations 1997 "Life Expectancy and Infant Mortality Rate, 1995–2000: Developed Regions." <http://www.un.org/Depts/unsd/gender/3-1dev.htm>.
- U.S. Bureau of the Census 1975 *Historical Statistics of the U.S., Colonial Times to 1970. Bicentennial Edition. Part 2*. Washington, D.C.: U.S. Government Printing Office.
- Vaupel, J. 1986 "How Change in Age-Specific Mortality Affects Life Expectancy." *Population Studies* 40:147–157.
- Wilkins, R., and O. B. Adams 1983 "Health Expectancy in Canada, Late 1970s: Demographic, Regional, and Social Dimensions." *American Journal of Public Health* 73:1073–1080.
- Yanagishita, M., and J. Guralnik 1988 "Changing Mortality Patterns that Led Life Expectancy in Japan to Surpass Sweden's: 1972–1982." *Demography* 25:611–624.
- Yin, P., and M. Shine 1985 "Misinterpretations of Increases in Life Expectancy in Gerontology Textbooks." *Gerontologist* 25:78–82.

EILEEN M. CRIMMINS

## LIFE HISTORIES AND NARRATIVES

The life history approach to social research and theory subsumes several methodological techniques and types of data. These include case studies, interviews, use of documents (letters, diaries, archival records), oral histories, and various kinds of narratives. The popularity of this approach has waxed and waned since the early 1900s. It was used extensively in the 1920s and 1930s and was identified with the Chicago sociology of W. I. Thomas, Robert Park, Clifford Shaw, and others. The succeeding generation of sociology witnessed the solidification of quantitative measurement techniques coupled with survey data collection, and the increased use of those approaches paralleled a relative decline in life history research. In the 1970s, however, there began a resurgence of interest in life history research not only in the United States but in Europe. The work of some sociologists, such as Howard S. Becker and Anselm Strauss, has maintained the early Chicago tradition, while newer generations of scholars—such as Norman Denzin and Michal McCall (United States), Ken Plummer (England), Daniel Bertaux (France), and Fritz Schütze (Germany)—have augmented life history research. This resurgence has been accompanied by the creation of Research Committee 38 (Biography and Society) of the International Sociological

Association in the late 1970s and has included a broadened interdisciplinary base through the incorporation of narrative theory and methods from other disciplines. In this broadened use there has been a transition from using the approach as purely a methodological device to using it as method, theory, and substance. It is this transition that frames this article.

The main assumptions of this approach are that the actions of individuals and groups are simultaneously emergent and structured and that individual and group perspectives must be included in the data used for analysis. Accordingly, any materials that reveal those perspectives can and should be regarded as essential to the empirical study of human social life. Life history materials, as described above (see Denzin 1989b, chap. 8; Plummer 1983, chap. 2; Gottschalk et al. 1945), contain first-, second-, and third-order accounts of past actions, as well as plans and expectations regarding future actions. Those materials will reveal significant information concerning the author's (writer and speaker) meanings. Invariably, these materials pertain also to the processual character of social life, and thus there is a major emphasis on temporal properties such as sequence, duration, and tempo. These assumptions and emphases have been characteristic of the vast majority of life history studies.

The first major empirical study in American sociology that systematically combined explicit theory and method was Thomas and Znaniecki's *The Polish Peasant in Europe and America* (1918–1920). The purpose of these researchers was to investigate Polish immigrants in America, especially their problems in adjusting to American urban life. The researchers used their famous attitude-value scheme as an explanatory framework. In this scheme, *attitudes* referred to individual subjective meanings and *values* to objective societal conditions. Thomas and Znaniecki proposed a set of causal explanations based on how the relations of attitudes and values were interpreted by individuals and groups. In their five-volume, 2,200-page work, they presented almost 800 pages of life history data in support of their conclusions and generalizations. The data included newspaper articles, letters to family members, records from courts and social work agencies, and a 300-page biography of one person that was presented as a representative case (Blumer 1939).

This research, which the Social Science Research Council in 1938 voted as the most outstanding in sociology to that date, depended solely on life history data. Because of its systematic incorporation of theory and method, it stimulated and became an exemplar for a long series of similar studies. These included research on race relations, delinquency, housing, mass media, migration, occupations, and other issues centered primarily in the areas of ethnic and urban studies (Bulmer 1984). The emphasis during this period was on the contributions of life history methods to sociology as an empirical and scientific discipline. Accordingly, researchers using this approach focused on methodological problems such as reliability, validity, hypothesis formation, and the making of generalizations, although comparatively less concern was given to sampling (Gottschalk et al. 1945). Reflecting the major issue of pre-World War II sociological work, the focus was on the adequacy of this approach for discovering lawlike behavior or empirically valid generalizations. The emphasis, in short, was on the approach as a research tool.

The developments in this approach since the early 1970s, perhaps stimulated in part by increased interest in historical sociology and in part by the articulation of insoluble problems in statistical approaches, have been more interdisciplinary, international, and sophisticated than the early works (McCall and Wittner 1990; Jones 1983; Roth 1987). It is increasingly recognized that all social science data, whether represented in discursive or numeric form, are interpretations (Denzin 1989a; Gephart 1988). This recognition is one of the central tenets of the narrative approach to social research (Fisher 1987; Reed 1989; Richardson 1990; Maines 1993), which makes the ontological claim that human beings are inherently storytellers. This shift in emphasis concerning the subject matter of sociology, in which human behavior is conceptualized as significantly communicative and narrative in nature, is precisely what has reframed the utility and potential of the life history approach.

Current uses of life history research display considerable variation as well as more precise conceptual distinctions. Terms such as *life story*, *biography*, *discourse*, *history*, *oral history*, *personal experience narratives*, *collective narratives*, and *sagas* are now distinguished from one another (Denzin 1989b, pp. 184–187), and frameworks for linking types of verbal accounts to types of generalizations

have been developed (Sperber 1985, pp. 9–34). Moreover, these developments have occurred within and across different theoretical approaches and disciplines.

It is now common to regard life histories as a legitimate form of data in which currency is established through the propositions contained in narrative theory. Some of the uses found in contemporary work include the following. Schütze (1983) has developed what he calls the narrative interview. This approach focuses on establishing event sequences across the life course on the basis of interview data. These sequences are derived from detailed analyses of biographical materials, with special attention to the structural factors that have shaped the person's life. Analytical summaries are developed and, through analytical inductive procedures, are compared to subsequently developed summaries. The goal is to produce theoretical interpretations centering on various analytical interests such as life course transitions, career models, or natural histories. Riemann and Schütze (1991) provide a substantive application of this method in the area of chronic illness.

Bertaux (1981; Bertaux and Kohli 1984) has long been an advocate of life history research and was the primary organizer of the Biography and Society Section of the International Sociological Association. He has conducted a number of projects that have goals similar to those of Schütze. His collaborative research on social movements (1990), for example, used life history data from members of students movements in the United States, England, Ireland, Italy, West Germany, and France. There he shows the application of the method in large-scale comparative research projects. The epistemological approach was not to gather data on lived, biographical experiences of the activists but to analyze those data in collective, generic terms. That is, his strategy was to focus on similarities rather than differences across nations and to ground empirically theoretical statements about, for example, processes of commitment to social-movement ideologies. The contention of this research is that biographical and life history data from ordinary people will reveal those similarities and thus make contributions to cross-national research.

Dolby-Stahl (1989), a folklorist, has developed a variation of the life history approach. She calls it

*literary folkloristics*, and it focuses on personal narrative data. She uses reader response theory to develop an interpretive method for studying the interdependence of personal narratives (stories) and collective narratives (e.g., ethnic group folklore). Her procedures entail locating the respondent (storyteller) in large collectivities (e.g., single parents), identifying salient themes (e.g., day care), and connecting personal to collective narratives (e.g., the respondent's accounts of day care and media or community accounts). The assumption of this approach is that personal and collective narratives are inherently connected, and thus a personal story is always in some way a collective story. Further, the assumption that the researcher in varying ways is part of the collective story requires facing the interpretive nature of data collection. In this respect, the researcher draws on her own shared cultural experiences to analyze the life history or narrative data provided by the respondent. These procedures locate the life history approach squarely in interpretive social theory, in which credible interpretation is the goal as opposed to, say, producing explanations justified by measures of reliability and validity.

Similarly, Denzin (1989a) has developed an interpretive approach that draws conceptually from postmodernism and phenomenology, and methodologically from Clifford Geertz's advocacy of thick description. He calls his method *interpretive biography*, and it is designed to study the turning points of problematic situations in which people find themselves during transition periods. Data include documents, obituaries, life histories, and personal experience stories, with the emphasis on how such information is read and used. The basic question he asks concerns how people live and give meaning to their lives, and how meaning is represented in written, narrative, and oral forms. His approach thus addresses an enduring problem in sociology, which C. Wright Mills located at the intersection of biography and history, as well as the newer problem articulated by interpretive theories regarding the interpretations of texts, cultural forms, and personal acts.

Most of the developments in the 1990s continue to be flamed largely under the rubric of narrative inquiry, signified by the 1991 inception of the *Journal of Narrative and Life History*. These recent lines of work can be organized loosely into four categories. First, there have been a considerable

increase in substantive studies. Ezzy (1998) joins a rather large number of scholars (e.g., Angrosino 1995) studying narrative identity, which Orbuch (1997) moves more generically into the arena of accounts. Gubrium (1993) uses narrative data in his examination of quality of life among nursing home residents. Plummer (1995) uses similar data to study sexuality. Maines and Bridger (1992) studied the narrative character of land-use decisions. Randall (1999) has studied the narrative aspects of intelligence. Eheart and Power (1995) assess processes of success and failure in families with adopted children. Maines (1999) has provided data about how racial attitudes about justice are embedded in larger narrative structures. TenHouten (1999) has used life history interviews in a comparative analysis of temporality among Australian aboriginals and Europeans.

These substantive topics overlap with the second line of work pertaining to historical sociology. Barry Schwartz (1996, 1997) has rekindled sociological interest in collective memory to show how historical processes contribute to changing cultural representations (see also Wertsch and O'Conner 1994). Another strand of work has pertained to causal analysis. Abbott (1992) articulated his version as *narrative positivism* which focuses on the properties of events and sequence to move quantitative sociology from the study of variables to the study of actual events. Griffin (1993) uses data on lynchings to propose a similar approach, utilizing computer-assisted event structure analysis that focuses directly on temporality. Gotham and Staples (1996) follow these studies with a theoretical analysis of narrative in the relations of agency and social structure, as does Berger (1995) in his analysis of Jewish Holocaust survivors. Mahoney (1999) addresses nominal, ordinal, and narrative dimensions of causal analysis, noting that "narrative analysis has the obvious strength of allowing the analyst to show sensitivity to detail, process, conjuncture, and causal complexity" (p. 1168).

The third area concerns methods of data collection and analysis. Holstein and Gubrium (1995) treat the interview as a social encounter that itself constructs the data gathered. In these encounters, respondents are regarded as narrators who tell the story of their "own past attitudes, feelings, and behaviors" (p. 32). McMahan and Rogers (1994) likewise present oral history interviewing as an interactive, negotiated process. They synthesize a

large amount of materials about potential biases, and make concrete suggestions for developing skills to deal with them. On the other hand, Lieblich and associates (1998) present a model for analyzing life story data. Their model is composed of two continua: holistic versus categorical (whether a life story is taken as a whole or dissected into parts or categories) and content versus form (analysis of the substance or structure of life story accounts). This model is suggested as useful for guiding analyses for varying purposes. Mishler (1995) also addresses the diversity of narrative inquiry, and proposes a typology of narrative analyses. These include modes of analysis that focus on (1) the correspondence between temporal sequences of actual events and their textual representation, (2) how types of stories acquire structure and coherence, and (3) the content and function of stories. This typology organizes a large literature on narrative, and serves as a heuristic for discerning similarities and differences in analytical approaches and purposes. Finally, Atkinson (1998) describes procedures for the life story interview, whose "product is entirely a first-person narrative, with the researcher removed as much as possible from the text" (p. 2). These procedures would encourage respondents to see and tell their lives as a whole, and Atkinson has developed a growing archival data base of over 300 documents suitable for various analytical purposes.

The life story interview merges with the fourth line of work, which focuses on biographies and autobiographies. Smith (1994) provides an excellent overview of biographical methods. He begins with Charles Darwin's biography and brings the reader to contemporary issues that cut across various disciplines in the human sciences. In doing so, he emphasizes that biographical analysis remains an unfinished project that is filled with both conflict and creativity. Inside that space of conflict and creativity rests the new genre of *auto-ethnography*, or literally the study of oneself. Bochner and Ellis (1992) provide one of the more interesting instances of this approach. Each author was involved in the same first-time event. Wanting to fully explore the meaning of that event, they first wrote extensive personal accounts of it. Second, they shared those accounts with one another and with friends; and third, they wrote a joint account from the standpoint of their relationship. Through these procedures, they were able to preserve their own

biographical perspectives on the event as well as depicting its intersubjective, consensual meanings.

The auto-ethnography stimulates attention to writing as a way of knowing, which is an issue widely discussed in the last decade among narrative scholars. Richardson (1994) discusses an array of issues pertaining to writing as inquiry, and then provides a detailed autobiographical account of how context affects writing (Richardson 1997). Her account includes her family, academic departments, networks of colleagues, and students, and shows how her relationships and experiences with these people were part and parcel of how and what she wrote as a professional sociologist. Her writing, she shows, was intimately tied to constructions of knowledge.

While Richardson and others illustrate what Polkinghorne (1988) has termed "narrative knowing" Diane Bjorklund (1998) treats autobiographies as sociological subject matter. She analyzed a sample of the 11,000 American autobiographies written from 1800 to 1980 to provide a historical analysis of representations of the self. She shows how social conventions and vocabularies for describing oneself have changed through four discernible eras of American history. By taking autobiographies seriously, this analysis powerfully locates the "texts" of personhood in cultural and social structural contexts.

Scholars working in the area of life history research accept that all social science data are made up of human interpretations and that nearly all such data are reconstructions or representations of past events and experiences. Because of its development of techniques for gathering, coding, and analyzing explicitly reconstructive data, the life history approach is suitable for studying not only the subjective phases of social life but the historical and structural aspects as well. It can be used for a wide variety of topics and purposes, ranging from research on the trajectories of personal biographies to organizational functioning to migration patterns. It invariably leads to development of theories emphasizing social processes. Recent scholarship on narrative has shown three dominant and related trends. The first, which represents an extension of the life history research of the 1920s, pertains to the linking of collective and personal narratives. The second, a departure from the earlier work, recognizes scholars as active

narrators themselves and thus creators of narrative knowing. The third trend is the broad appeal of this area of work. It is found in all disciplines of the human sciences, it uses a wide variety of theoretical approaches, and it incorporates the array of analytical methods typical of sociological work found among American and European scholars.

(SEE ALSO: *Case Studies*; *Qualitative Methods*)

## REFERENCES

- Abbott, Andrew 1992 "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods and Research* 20:428-455.
- Angrosino, Michael 1995 "Metaphors of Ethnic Identity: Projective Life History Narratives of Trinidadians of Indian Descent." *Journal of Narrative and Life History* 5:125-146.
- Atkinson, Robert 1998 *The Life Story Interview*. Thousand Oaks, Calif.: Sage.
- Berger, Ronald 1995 "Agency, Structure, and Jewish Survival of the Holocaust: A Life History Study." *Sociological Quarterly* 36:15-36.
- Bertaux, Daniel 1990 "Oral History Approaches to an International Social Movement." In Else Oyen, ed., *Comparative Methodology: Theory and Practice in International Social Research*. London: Sage.
- , ed. 1981 *Biography and Society: The Life-History Approach in the Social Sciences*. London: Sage.
- , and Martin Kohli 1984 "The Life-Story Approach: A Continental View." *Annual Review of Sociology* 10:149-167.
- Bjorklund, Diane 1998 *Interpreting the Self: Two Hundred Years of American Autobiography*. Chicago: University of Chicago Press.
- Blumer, Herbert 1939 *An Appraisal of Thomas and Znaniecki's "The Polish Peasant in Europe and America"*. New York: Social Science Research Council.
- Bochner, Arthur, and Carolyn Ellis 1992 "Personal Narrative as an Approach to Interpersonal Communication." *Communication Theory* 2:165-172.
- Bulmer, Martin 1984 *The Chicago School of Sociology*. Chicago: University of Chicago Press.
- Denzin, Norman 1989a *Interpretive Biography*. Newbury Park, Calif.: Sage.
- 1989b *The Research Act*. Englewood Cliffs, N.J.: Prentice-Hall.
- Dolby-Stahl, Sandra 1989 *Literary Folkloristics and the Personal Narrative*. Bloomington: Indiana University Press.

- Eheart, Brenda Krause, and Martha Bauman Power 1995 "Adoption: Understanding the Past Present, and Future through Stories." *Sociological Quarterly* 36:197-216.
- Ezzy, Douglas 1998 "Theorizing Narrative Identity: Symbolic Interactionism and Hermeneutics." *Sociological Quarterly* 39:239-252.
- Fisher, Walter 1987 *Human Communication as Narrative*. Columbia: University of South Carolina Press.
- Gephart, Robert 1988 *Ethnostatistics: Qualitative Foundations for Quantitative Research*. Newbury Park, Calif.: Sage.
- Gotham, Kevin Fox, and William Staples 1996 "Narrative Analysis and the New Historical Sociology." *Sociological Quarterly* 37:481-501.
- Gottschalk, Louis, Clyde Kluckhohn, and Robert Angell 1945 *The Use of Personal Documents in History, Anthropology, and Sociology*. New York: Social Science Research Council.
- Griffin, Larry 1993 "Narrative, Event Structure, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98:403-427.
- Gubrium, Jaber 1993 *Speaking of Life: Horizons of Meaning for Nursing Home Residents*. Hawthorne, N.Y.: Aldine de Gruyter.
- Holstein, James, and Jaber Gubrium 1995 *The Active Interview*. Thousand Oaks, Calif.: Sage.
- Jones, Gareth 1983 "Life History Methodology." In Gareth Morgan, ed., *Beyond Method: Strategies for Social Research*. Beverly Hills, Calif.: Sage.
- Lieblich, Amia, Rivka Tuval-Mashiach, and Tamar Zilber 1998 *Narrative Research*. Thousand Oaks, Calif.: Sage.
- McCall, Michal, and Judith Wittner 1990 "The Good News about Life History." In Howard S. Becker and Michal McCall, eds., *Symbolic Interaction and Cultural Studies*. Chicago: University of Chicago Press.
- McMahan, Eva, and Kim Lacy Rogers 1994 *Interactive Oral History Interviewing*. Hillsdale, N.J.: Lawrence Erlbaum.
- Mahoney, James 1999 "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis." *American Journal of Sociology* 104:1154-1196.
- Maines, David 1993 "Narrative's Moment and Sociology's Phenomena: Toward a Narrative Sociology." *Sociological Quarterly* 34:17-38.
- 1999 "Information Pools and Racialized Narrative Structures." *Sociological Quarterly* 40:317-326.
- , and Jeffery Bridger 1992 "Narrative, Community, and Land Use Decisions." *Social Science Journal* 29:363-380.
- Mishler, Elliot 1995 "Models of Narrative Analysis: A Typology." *Journal of Narrative and Life History* 5:87-124.
- Orbuch, Terri 1997 "People's Accounts Count: The Sociology of Accounts." *Annual Review of Sociology* 23:455-478.
- Plummer, Ken 1983 *Documents of Life*. London: Allen and Unwin.
- 1995 *Telling Sexual Stories: Power, Change, and Social Worlds*. London: Routledge.
- Polkinghorne, Donald 1988 *Narrative Knowing and the Human Sciences*. Albany: State University of New York Press.
- Randall, William Lowell 1999 "Narrative Intelligence and the Novelty of Our Lives." *Journal of Aging Studies* 13:11-28.
- Reed, John Sheldon 1989 "On Narrative and Sociology." *Social Forces* 68:1-14.
- Richardson, Laurel 1990 "Narrative and Sociology." *Journal of Contemporary Ethnography* 19:116-135.
- 1994 "Writing: A Method of Inquiry." In Norman Denzin and Yvonna Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- 1997 *Fields of Play: Constructing an Academic Life*. New Brunswick, N.J.: Rutgers University Press.
- Riemann, Gerhard, and Fritz Schütze 1991 "Trajectory as a Basic Theoretical Concept for Analyzing Suffering and Disorderly Social Processes." In David R. Maines, ed., *Social Organization and Social Processes: Essays in Honor of Anselm Strauss*. Hawthorne, N.Y.: Aldine de Gruyter.
- Roth, Paul 1987 *Meaning and Method in the Social Sciences*. Ithaca, N.Y.: Cornell University Press.
- Schütze, Fritz 1983 "Biographieforschung und narrative Interview." *Neue Praxis* 3:283-293.
- Schwartz, Barry 1996 "Memory as a Cultural System: Abraham Lincoln in World War II." *American Sociological Review* 61:908-927.
- 1997 "Collective Memory and History: How Abraham Lincoln Became a Symbol of Racial Equality." *Sociological Quarterly* 38:469-496.
- Smith, Louis 1994 "Biographical Method." In Norman Denzin and Yvonna Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Sperber, Dan 1985 *On Anthropological Knowledge*. Cambridge, England: Cambridge University Press.
- TenHouten, Warren 1999 "Text and Temporality: Patterned-Cyclical and Ordinary-Linear Forms of Time-Consciousness, Inferred from a Corpus of Australian

Aboriginal and Euro-Australian Life-Historical Interviews." *Symbolic Interaction* 22:121-137.

Thomas, W. I., and Florian Znaniecki 1918-1920 *The Polish Peasant in Europe and America*, 5 vols. Chicago: University of Chicago Press.

Wertsch, James, and Kevin O'Conner 1994 "Multivoicedness in Historical Representation: American College Students' Accounts of the Origin of the United States." *Journal of Narrative and Life History* 4:295-309.

DAVID R. MAINES

## LIFE TABLES

See Demography; Life Expectancy.

## LIFESTYLES AND HEALTH

Lifestyles are a major determinant of who shall live and who shall die (Fuchs 1974; McKinlay and Marceau 1999). Mechanic (1978, p. 164) argues that the concept of lifestyles refers to a diverse set of variables, including nutrition, housing, health attitudes and beliefs, risk-taking behavior, health behavior and habits, and preventive health behavior.

### CONFIRMING THE LINK BETWEEN LIFESTYLES AND HEALTH

Establishing the causal linkage between lifestyles and health is not a simple task. Variables included in lifestyles interact with each other (Mechanic 1978), making it difficult to adjust for confounding variables such as race and ethnicity, gender, social class, and psychological distress.

In addition, there are problems in specifying the nature of the etiological relationship between lifestyles and disease. Not every person who engages in an unhealthy lifestyle will die prematurely. For example, some heavy smokers do not develop lung cancer. Genetic predisposition, comorbidities, other health habits, and access to adequate medical care are factors that may intervene in the relationship between host and disease. For certain conditions, it may be complicated to determine the precise role of risky lifestyle behaviors in the development of disease.

The most convincing models of the relationship between lifestyles and health are those built on triangulated evidence from animal, clinical, and epidemiological studies. As an example, consider the link between tobacco use and cancer. In controlled randomized trials using animal subjects that are genetically the same, the experimental group of animals is exposed to tobacco smoke while the control group is not. If the experimental group has a higher incidence of cancer than the control group, the study provides evidence to link tobacco smoke to cancer. Another strategy to link lifestyle behavior to health involves clinical studies with human subjects. Lung tissue of smokers is compared with lung tissue of nonsmokers. If more smokers than nonsmokers have cancerous cells in the lung tissue, this provides additional data to confirm that smoking causes lung cancer. A final strategy uses epidemiological methods. In prospective studies, separate groups of smokers and nonsmokers are followed over time to ascertain the risk of developing cancer within each group. All else being equal, if smokers develop more cases of cancer, the causal relationship between smoking and cancer is confirmed.

With these methodological issues in mind, four selected behaviors are used below to illustrate how lifestyles impact on health: tobacco use, alcohol consumption, diet, and sexual behavior and injection-related practices that increase the risk of acquiring the human immunodeficiency virus (HIV). For each behavior, extensive animal, clinical, and epidemiological data exist to support the causal relationship between each agent and disease. Gender differences in these behaviors are used to illustrate how lifestyles vary across social groups. Men are more likely to engage in these risky behaviors compared to women; this helps explain why women live longer than men do (Croese 1997). Finally, examples of successful efforts to change risky life-style behaviors are provided.

### TOBACCO USE

Tobacco use has been defined as the most important single preventable cause of death and disease in society. Smoking causes an average of 430,700 deaths per year in the United States. One in every five deaths is smoking related (Centers for Disease Control and Prevention 1997a).



Careful epidemiological studies have determined that tobacco use increases the risks for heart disease, lung cancer, emphysema, and other lung diseases. Smoking during pregnancy increases the risk for premature births, complications of pregnancy, low-birthweight infants, stillbirths, and infant mortality. Fetuses exposed to smoke in utero and young infants exposed to secondhand smoke are at increased risk for sudden infant death syndrome, poor lung development, asthma, and respiratory infections (Centers for Disease Control and Prevention 1997b; Environmental Protection Agency 1992; Floyd et al. 1993). Secondhand smoke is also associated with adult illnesses (Centers for Disease Control and Prevention 1997a). Increased risks for lung cancer and heart disease are reported for nonsmokers who live with smokers.

Age-adjusted rates indicate that 27 percent of men smoked in 1995, compared to 23 percent of women (National Center for Health Statistics 1998). The gender gap used to be wider. In 1965, men were 1.5 times more likely to smoke than women were. Today, the gender difference is only 17 percent. The decline in cigarette smoking has been greater among men. While the prevalence of smoking among men dropped by almost half from 1965 to 1995, the rate for women dropped by only one-third. Traditionally, smoking by women was not condoned, but over time these attitudes have changed. As a result, the smoking rates for men and women are nearly the same. This translates into a 400 percent increase in deaths from lung cancer among women between 1960 and 1990 (Centers for Disease Control and Prevention 1997a).

Various prevention efforts have reduced the prevalence of smoking in the United States. Smokefree workplaces are the norm rather than the exception. Warning labels appear on cigarette packages, and billboards advertising cigarettes are banned. Community-based public education campaigns and worksite programs have been successful in smoking reduction (COMMIT Research Group 1995).

## ALCOHOL

The National Institute on Alcohol Abuse and Alcoholism (NIAAA 1997) estimates that alcohol use is

responsible for 100,000 deaths in year in this country. About 44 percent of the motor vehicle fatalities in 1994 were alcohol related (National Highway Traffic Safety Administration 1994). Alcohol use is frequently implicated in accidental injuries and deaths from falls, drowning, interpersonal and family violence, occupational hazards, and fires (NIAAA 1997).

From 10 to 20 percent of heavy drinkers develop cirrhosis of the liver, which was the tenth leading cause of death in the United States in 1996 (DeBakey et al. 1995; National Center for Health Statistics 1998). The liver is the primary site of alcohol metabolism, and drinkers are at risk for other forms of liver disease, including alcoholic hepatitis and cancer.

Heavy alcohol use causes loss in heart muscle contractile function, arrhythmias, degenerative disease of the heart muscle, and heart enlargement, and also increases the risk for hypertension and stroke. Alcohol is implicated in esophageal, breast, and colorectal cancer, and it may increase the risk for other types of cancer as well. NIAAA concludes, "The range of medical consequences of alcohol abuse is both immense and complex—virtually no part of the body is spared the effects of excessive alcohol consumption" (1990, p. 127).

Alcohol also functions as a teratogen, producing defects in the human fetus in utero. The possible effects of alcohol on the fetus include gross morphological defects as well as cognitive and behavioral dysfunctions. Alcohol ingestion during pregnancy causes a variety of birth defects, including fetal alcohol syndrome, alcohol-related birth defects, and alcohol-related neurodevelopmental disorder. Fetal alcohol syndrome is the most severe consequence of the mother's heavy drinking during pregnancy and is characterized by craniofacial anomalies, mental retardation, central nervous system dysfunction, and growth retardation. Fetal alcohol syndrome is one of the leading causes of preventable birth defects (Stratton et al. 1996).

Men are more likely to drink alcohol than women are. Twenty-two percent of men are lifetime abstainers from alcohol, versus 45 percent of women. About 56 percent of men are current drinkers, compared to 34 percent of women. Heavy drinking is gender related. Almost 12 percent of

men average more than fourteen alcoholic drinks per week, compared to less than 4 percent of women (NIAAA 1998).

Various prevention efforts are directed toward decreasing alcohol consumption. The alcohol beverage warning label, implemented in 1989, warns drinkers about birth defects, drunk driving, operating machinery, and health problems; however, its impact has been modest. Other alcohol prevention programs have been implemented, including dram shop liability (servers are legally responsible for damage or injury caused by drunk patrons), training servers of alcohol to avoid selling to intoxicated persons and to minors, lowering the allowable blood alcohol concentration levels for drivers, changing the availability of alcohol, enforcing impaired driving laws, and designating one person in a car as the nondrinking driver. These community intervention programs have reduced alcohol-related traffic deaths significantly (NIAAA 1997).

### DIET

Dietary factors have been linked to mortality from cardiovascular disease (heart disease and stroke) and cancer; these diseases are the leading causes of death in the United States. Diet affects four of the major risk factors for cardiovascular disease: hypertension, obesity, diabetes, and high cholesterol. Obesity, as well as diets high in saturated fats, trans fatty acids, and cholesterol, raise blood cholesterol concentration and blood pressure, thus increasing the risk for coronary heart disease and stroke. Diets low in saturated fats but high in fiber and some omega-3 fatty acids (found in walnuts, certain oils [fish, canola, soybean], and green leafy vegetables) lower the risk for heart disease (Hu et al. 1997; Ascherio et al. 1996).

Diet also affects cancer risk. While the link between dietary intake and breast cancer is controversial, there is evidence that a diet high in saturated fat plays an etiological role (Kolonel 1997; Hankin 1993). High fat intake increases the risk for prostate cancer, lung cancer, and colorectal cancer (Kolonel 1997). On the other hand, diets high in fruits, vegetables, and fiber lower the probability of developing various types of cancer, including breast, colorectal, stomach, and lung (Ziegler 1991; Hankin 1993).

Gender differences in diet and obesity exist. For example, women are more likely to be overweight compared to men (39 percent versus 36 percent). Among the poor, the gender gap is larger: 46 percent of poor women versus 31 percent of poor men suffer from obesity (National Center for Health Statistics 1998). Despite this increased risk for obesity, women consume healthier food than men do. While men consume more meat, saturated fat, and high-calorie foods, women eat more fruits, vegetables, whole grains, and lower-calorie foods (Croese 1997). Successful programs have been developed to encourage healthier eating habits for men and women, including interventions in worksites and families (Sorensen et al. 1999). The public has responded to these efforts, as evidenced by a 4 percent drop in blood cholesterol levels from 1978 to 1990 (National Heart, Lung, and Blood Institute 1996).

### SEXUAL AND INJECTION-RELATED PRACTICES

The human immunodeficiency virus (HIV) causes acquired immune deficiency syndrome (AIDS), the eighth leading cause of death in the United States in 1996 (Peters et al. 1998). HIV is transmitted through blood products, bodily fluids, and breast milk. Transmission of HIV can be prevented by the use of safe sexual practices and sterile needles.

There are clear gender differences in the way AIDS is contracted, and the infection rate among women has been rising. Forty-eight percent of AIDS patients acquire the disease through male homosexual contact. Another 10 percent of AIDS cases appear in men who have sexual relations with a male or female partner and also use intravenous drugs. Almost 26 percent become ill solely through infected needles, and 6 percent through heterosexual contact that was not related to intravenous drug use. Among women, the most common route of infection is intravenous drug use (by herself or her partner), 60 percent or unprotected sex with an infected partner, 22 percent (Centers for Disease Control and Prevention 1998a).

While some HIV prevention programs promote abstention from sex and cessation of drug injecting, most programs define harm reduction

as the goal (Kelly 1999). This policy promotes using condoms correctly on a consistent basis, having sex with uninfected partners, and cleaning needles with bleach or exchanging contaminated syringes for sterile ones.

Programs designed to decrease (1) the rate of unprotected anal sex among men who have sex with men and (2) the proportion of men having unprotected sex with multiple male partners contributed to the reduction in AIDS cases among homosexuals during the 1980s and early 1990s. However, 1997 data suggest that unprotected anal sex among gay men and unprotected sex with multiple partners is increasing, especially among younger gay men (Centers for Disease Control and Prevention 1999).

The second major mode of HIV transmission, especially among women, is by infected needles. Syringe exchange programs or the use of bleach to clean syringes prevent the spread of HIV among injecting-drug users. The number of syringe exchange programs has grown rapidly; 17.5 million needles were exchanged in 100 programs in 1997 (Centers for Disease Control and Prevention 1998b). Other communities have distributed bleach kits that reduce the spread of HIV (CDC AIDS Community Demonstration Projects Research Group 1999).

## DISCUSSION

Four examples were selected to illustrate the role of lifestyles and health. There are other lifestyles that are related to health, which are beyond the scope of this article. For example, stressful lifestyles have been linked to mental illness, gastrointestinal illness, and heart disease. Regular use of seatbelts reduces the likelihood of death or injury in automobile accidents. Proper dental hygiene decreases the rate of dental caries. Childhood immunizations prevent measles, mumps, and polio.

When discussing lifestyles and health, several unresolved issues remain. First, the role of lifestyles in health is still evolving. As Becker (1993) argues, what is said to be bad for us one day may be determined to be good for us the next (and vice versa). For example, while researchers have documented the deleterious effects of alcohol on health, there is some recent evidence that red wine consumption may lower cholesterol levels (NIAAA 1997).

Second, there is continuing debate about the pros and cons of changing an individual's lifestyle versus changing the social milieu (Kelly 1999). Should we invest in programs designed to encourage an individual to stop smoking? Is it better to develop strategies that change societal norms about the acceptability of smoking? Should we do both?

Third, policy makers note that healthy lifestyle programs must be tailored to the individual, the subgroup (gender, age, race or ethnicity, and social class) and the particular community at risk (Kelly 1999). Thus, designing successful interventions to alter lifestyles is challenging.

Finally, it must be emphasized that while a healthy lifestyle may be a necessary condition for longevity; it is not a sufficient condition. Many variables interact with lifestyles to protect against disease and death. For example, evidence is mounting that genetic predispositions are very important in the etiology of certain diseases. Thus, the models predicting who shall live and who shall die involve complicated interactions of lifestyles, preventive health behavior, genetic risk, sociodemographic characteristics, and so on. Nonetheless, individuals who abstain from smoking and injecting drugs, drink alcohol in moderation, reduce the intake of saturated fat and cholesterol, and use safe sex practices have a better chance of survival than those who eat and drink excessively and do not follow recommended safe sex practices.

## REFERENCES

- Ascherio, Alberto, Eric B. Rimm, Edward L. Giovannucci, Donna Spiegelman, Meir Stampfer, and Walter C. Willett 1996 "Dietary Fat and Risk of Coronary Heart Disease in Men: Cohort Follow Up Study in the United States." *British Medical Journal* 313:84-90.
- Becker, Marshall 1993 "A Medical Sociologist Looks at Health Promotion." *Journal of Health and Social Behavior* 34:1-6.
- Centers for Disease Control AIDS Community Demonstration Projects Research Group 1999. "Community-Level HIV Intervention in 5 Cities: Final Outcome Data from Demonstration Projects." *American Journal of Public Health* 89:336-345.
- Centers for Disease Control and Prevention 1997a *Facts about Cigarette Mortality*. Atlanta, Ga.: CDC, May 23.
- 1997b *Fact Sheet, Smoking and Pregnancy*, Atlanta, Ga.: CDC, November 7.

- 1998a *HIV/AIDS Surveillance Report*, No. 10. Atlanta, Ga.: CDC.
- 1998b “Update: Syringe Exchange Programs—United States, 1997.” *Morbidity and Mortality Weekly Report* 47:652–655.
- 1999 “Increases in Unsafe Sex and Rectal Gonorrhea among Men Who Have Sex with Men—San Francisco, California, 1994–1997.” *Morbidity and Mortality Weekly Report* 48:45–48.
- COMMIT Research Group 1995 “Community Intervention Trial for Smoking Cessation (COMMIT): I. Cohort Results from a Four Year Community Intervention.” *American Journal of Public Health* 85:183–192.
- Cruse, Royda 1997 *Why Women Live Longer than Men . . . and What Men Can Learn from Them*. San Francisco: Jossey-Bass.
- DeBakey, S. F., F. S. Stinson, B. F. Grant, and M. C. Dufour 1995 *Liver Cirrhosis Mortality in the US 1970–1992*, Surveillance Report No. 37. Rockville, Md.: NIAAA Division of Biometry and Epidemiology.
- Environmental Protection Agency, Office of Research and Development 1992 *Respiratory Health Effects of Passive Smoking*, EPA/600/6-90/006F. Washington, D.C.: EPA.
- Floyd, R. L., B. K. Rimer, G. A. Giovano, P. D. Mullen, and S. E. Sullivan 1993 “A Review of Smoking in Pregnancy: Effects on Pregnancy Outcomes and Cessation Efforts.” *Annual Review of Public Health* 14:379–411.
- Fuchs, Victor 1974 *Who Shall Live? Health Economics and Social Change*. New York: Basic Books.
- Hankin, Jean H. 1993 “Role of Nutrition in Women’s Health: Diet and Breast Cancer.” *Journal of the American Dietetic Association* 93:994–999.
- Hu, Frank B., Meir J. Stampfer, JoAnn E. Manson, Eric Rimm, Graham A. Colditz, Bernard A. Rosner, Charles H. Hennekens, and Walter C. Willett 1997 “Dietary Fat Intake and the Risk of Coronary Heart Disease in Women.” *New England Journal of Medicine* 337:1491–1499.
- Kelly, Jeffrey A. 1999 “Community-Level Interventions Are Needed to Prevent New HIV Infections.” *American Journal of Public Health* 89:299–301.
- Kolonel, Laurence N. 1997 “Fat and Cancer: The Epidemiologic Evidence in Perspective.” Pp. 1–19 in American Institute for Cancer Research, ed., *Dietary Fat and Cancer*. New York: Plenum.
- McKinlay, John B., and Lisa D. Marceau 1999 “A Tale of 3 Tails.” *American Journal of Public Health* 89:295–298.
- Mechanic, David 1978 *Medical Sociology*, 2nd ed. New York: Free Press.
- National Center for Health Statistics 1998 *Health, United States, 1998 with Socioeconomic Status and Health*, DHHS Pub. No. 98-1232. Hyattsville, Md.: Centers for Disease Control and Prevention.
- National Heart, Lung, and Blood Institute 1996 *Facts about Blood Cholesterol*, NIH Pub. No. 94-2696. Bethesda, Md.: U.S. Department of Health and Human Services.
- National Highway Safety Administration 1994 *Traffic Safety Facts 1994: A Compilation of Motor Vehicle Crash Data from the Fatal Accident Reporting System and the General Estimates System*. DOT HS 808 169. Washington, D.C.: Department of Transportation National Center for Statistics and Analysis.
- National Institute on Alcohol Abuse and Alcoholism 1990 *Seventh Special Report to the U.S. Congress on Alcohol and Health*, DHHS Pub. No. (ADM) 90-1656. Rockville, Md.: U.S. Department of Health and Human Services.
- 1997 *Ninth Special Report to the U.S. Congress on Alcohol and Health*. DHHS Pub. No. (ADM) 97-4017. Rockville, Md.: U.S. Department of Health and Human Services.
- 1998 *Drinking in the United States: Main Findings from the 1992 National Longitudinal Alcohol Epidemiologic Survey*. DHHS Pub. No. 99-3519. Bethesda, Md.: U.S. Department of Health and Human Services.
- Peters, Kimberley D., Kenneth D. Kochanek, and Sherry L. Murphy 1998 *Deaths: Final Data for 1996; National Vital Statistics Report*, vol. 47, no. 9. Atlanta, Ga.: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.
- Sorensen, Glorian, Anne Stoddard, Karen Peterson, Nancy Cohen, Mary Kay Hunt, Evelyn Stein, Ruth Palombo, and Ruth Lederson 1999 “Increasing Fruit and Vegetable Consumption through Worksites and Families in the Treatwell 5-A-Day Study.” *American Journal of Public Health* 89:54–60.
- Stratton, Kathleen, Cynthia Howe, and Frederick Battaglia, eds. 1996 *Fetal Alcohol Syndrome: Diagnosis, Epidemiology, Prevention, and Treatment*. Washington, D.C.: National Academy Press.
- Ziegler, Regina G. 1991 “Vegetables, Fruit, and Carotenoids and the Risk of Cancer.” *American Journal of Clinical Nutrition* 53:251S–295S.

JANET HANKIN

**LINGUISTICS**

See Sociolinguistics.

## LITERATURE AND SOCIETY

Interest in the relationship between literature and society is hardly a new phenomenon. We still read and refer to the ancient Greeks in this regard. In *The Republic*, for example, Plato presages both Mme. de Staël's treatise of 1800, which was the first to discuss cross-national differences in literature, and later notions of literary reflection with his idea of imitation. What is new, however, is the relative legitimacy of the study of literature within the discipline of sociology. This is due both to the increasing interest in culture in sociology after years of marginalization (Calhoun 1989) and to the increasing influence of cultural studies on sociology and throughout the academy.

A broader interest in and acceptance of cultural sociology has meant that the types of research questions and methods common to sociological studies of literature are now more widely accepted within the field. Sociology has extended its methodological boundaries in response to both attacks on the dominance of positivism and the rising power of alternative stances suggested by postmodernism. At the same time, changes in the goals, and sometimes the methods, of studying literature sociologically have moved the area closer to what is still the mainstream of the discipline. Thus the sociology of literature has benefited from a twofold movement in which (1) sociology as a discipline has become more interested in and accepting of research questions pertaining to meaning (cf. Wuthnow 1987, however, for a particularly strong attack on meaning from within the culture camp) and employing qualitative methods; and (2) the sociology of literature has evolved in the direction of more mainstream sociological areas through the merging of quantitative with qualitative methods and of empirical with hermeneutic research questions.

### TRADITIONAL APPROACHES

As recently as 1993, Wendy Griswold maintained that the sociology of literature was a "nonfield" and "like an amoeba . . . lack[ing] firm structure" (1993, p. 455). Certainly the sociology of literature has been a marginal area in the discipline of sociology. As such, it has generally failed to attract the kind of career-long commitments common to

more central areas of the discipline. Many scholars writing on the sociology of literature see the area as a sideline and produce only a single book or article on the subject. This has exacerbated the lack of structure in the development of the field. Even so, it is surprising just how much sociological research has been done on literature and on literature's relationship to social patterns and processes.

**Reflection Theory.** Traditionally, the central perspective for sociologists studying literature has been the use of literature as information about society. To a much lesser degree, traditional work has focused on the effect of literature in shaping and creating social action. The former approach, the idea that literature can be "read" as information about social behavior and values, is generally referred to as *reflection theory*. Literary texts have been variously described as reflecting the "economics, family relationships, climate and landscapes, attitudes, morals, races, social classes, political events, wars, [and] religion" of the society that produced the texts (Albrecht 1954, p. 426). Most people are familiar with an at least implicit reflection perspective from journalistic social commentary. For instance, when *Time* magazine put the star of the television show *Ally McBeal* on its cover, asking "Is Feminism Dead?" (1998), it assumed that a television show could be read as information on Americans' values and understanding of feminism.

Unfortunately, "reflection" is a metaphor, not a theory. The basic idea behind reflection, that the social context of a cultural work affects the cultural work, is obvious and fundamental to a sociological study of literature. But the metaphor of reflection is misleading. Reflection assumes a simple mimetic theory of literature in which literary works transparently and unproblematically document the social world for the reader. In fact, however, literature is a construct of language; its experience is symbolic and mediating rather than direct. Literary realism in particular "effaces its own status as a sign" (Eagleton 1983, p. 136; see also Candido [1995, p. 149] on the "liberty" of even naturalist authors). Literature draws on the social world, but it does so selectively, magnifying some aspects of reality, misspecifying others, and ignoring most (Desan et al. 1989). The reflection metaphor assumes a single and stable meaning for literary texts. Anyone who has ever argued about what a book "really" meant knows what researchers have worked hard to demonstrate—textual meaning is contingent,

created by active readers with their own expectations and life experiences that act in concert with inherent textual features to produce variable meanings (Jauss 1982; Radway 1984; Griswold 1987).

Despite repeated demonstrations of reflection's myriad failings (e.g., Noble 1976; Griswold 1994; Corse 1997), the idea of literature as a mirror of society still seems a fundamental way of thinking about why sociologists—and indeed many other people as well—are interested in literature. A relatively crude reflection approach remains common for teaching sociology department courses on literature, and also in certain types of journal articles whose main interest is not the sociology of literature *per se*, but the illumination of some sociological theory or observation through literary “evidence” (e.g., Corbett's article [1994] advocating the use of novels featuring probation officers to teach courses on the sociology of occupations, or the continuing stream of articles examining gender portrayals in children's literature [e.g., Grauerholz and Pescosolido 1989]). Convincing research arguing for literary evidence of social patterns now requires the careful specification of how and why certain social patterns are incorporated in literature while others are not (e.g., Lamont 1995), thorough attention to comparative data across either place or time (e.g., Long 1985), and a detailed consideration of the processes that transform the social into the literary (e.g., Corse 1997).

**Structural Reflection.** A more sophisticated but still problematic type of reflection argues that it is the form or structure of literary works rather than their content that incorporates the social: “successful works . . . are those in which the form exemplifies the nature of the social phenomenon that furnishes the matter of the fiction” (Candido 1995, p. xiii). The “humanist” Marxist Georg Lukács is perhaps the seminal figure in the development of a Marxist literary sociology. Marxism is the only one of the three major strands of classical theory to have generated a significant body of work on literature. Lukács (1971) argued that it is not the content of literary works but the categories of thought within them that reflect the author's social world.

Goldmann (1964, 1970), Lukács's most prominent student and the one most influential for American sociology, proposed the concept of a

homologous relationship between the inherent structure of literary works and the key structures of the social context of the author. Goldmann justified his focus on the canonical works he studied by arguing that lesser works fail to achieve the necessary clarity of structure that allows the sociologist to see the homologies present in works by, for example, Racine and Pascal (1964). In the 1960s Louis Althusser challenged the preeminence of Lukács's tradition through, in part, his emphasis on the autonomy of literature. Thus Goldmann's work, though it was influential at the time of its publication, has been eclipsed as newer theories have made more problematic the notion that literature embodies a single meaning that is reducible to an expression of class consciousness.

#### **The High Culture/Popular Culture Divide.**

Traditionally in the United States sociologists have left the study of high culture to specialists in literature, art, and music. This attitude was partially a product of sociologists' discomfort with aesthetic evaluation. Popular culture, on the other hand, was seen as simply unworthy of attention or study. To the extent that sociologists did consider literature, they tended to focus on high-culture literature, in part because of the largely Marxist orientation of many early sociologists of literature. Marxist thought defines literature as part of the ideological superstructure within which the literatures of elites are the ruling ideas since culture serves to legitimate the interests of the ruling class.

The tendency to concentrate on high-culture literature was intensified by the Frankfurt School, which understood “mass” culture as a destructive force, imposed on a passive audience by the machinery of a capitalist culture “industry” (e.g., Horkheimer and Adorno 1972). Lowenthal's ([1961] 1968) analysis of popular magazine biography, for example, stressed the increasing focus on leisure-time consumption over production and on personality over business and political achievement, as the private lives of movie stars and sports figures came to dominate magazine biographies. This approach highlighted the passivity and docility of audiences, tying mass culture to the increasing apathy of the public. Thus this work saw literature both as a reflection of changing social patterns and as a force shaping those patterns. Although researchers now rarely use the term “mass” culture, the Frankfurt School's critique continues to inform much of current cultural sociology, although

often it does so on an implicit level as researchers react either positively or negatively to this understanding of popular culture.

One response to the critique of mass culture was articulated by the scholars of the Birmingham School. This line of research shared earlier understandings of culture as a resource for the powerful, but focused in large part on the potential for active participation on the part of cultural receivers. Work in the Birmingham School tradition drew heavily on feminist approaches and demonstrated how “mass” audiences of popular cultural forms might engage in resistance, undermining earlier arguments of cultural hegemony and of passive cultural “dopes” (e.g., Hall et al. 1980; Hebdige 1979). This interest in resistance and the meaning-making activity of readers remains an important line of research, particularly for studies of popular culture (e.g., Radway 1984). The continued relevance of the distinction between high and popular culture, however, is now under debate, as some charge that the hierarchical dichotomy is no longer the most powerful conceptualization of cultural differences (e.g., Crane 1992; DiMaggio 1987).

**Sociology through Literature.** A final type of traditional sociological interest in literature also stems from an implicit reflectionist approach. This type of work sees literature as exemplary of sociological concepts and theories or uses literature simply as a type of data like any other. While Coser’s (1972) anthology exemplifies the former tradition, the recent ASA publication *Teaching Sociology with Fiction* demonstrates the persistence of the genre. Examples of the latter are altogether too numerous, including, for example, an article testing recent Afrocentric and feminist claims of differing epistemological stances across genders and races by coding differences in the grounding of knowledge in novels for adolescent readers (Clark and Morris 1995). Such work ignores ignoring the mediated nature of literary “reality.” These discussions, although common, are not properly part of the sociology of literature.

### SOCIOLOGICAL ADVANCES

The 1980s saw the institutionalization of sociological studies of cultural objects and processes as most prominently indicated by the establishment

of the Culture Section of the American Sociological Association (ASA)—now one of the largest sections of the ASA with over one thousand members. This groundswell of interest in culture did not produce an equally large increase in interest in the sociology of literature, but it certainly created a more favorable climate for such work, as well as reenergizing research within the field.

Wendy Griswold is the key figure in the contemporary sociological study of literature in the United States. Her early research (1981, 1983, 1987) set the stage for a new synthesis that both takes seriously the issue of literary meaning and recognizes the importance of extratextual variables, while deploying the empirical data demanded by much of the discipline. By balancing these often-competing claims, Griswold allows for a study of literature that is sociological in the deepest sense of the word. Her concern for what she has called a “provisional, provincial positivism” (1990, p. 1580) has legitimated the sociology of literature to other sociologists and has articulated to nonsociologists the unique power of literary sociology. By publishing repeatedly in *American Journal of Sociology* and in *American Sociological Review*, Griswold made the sociology of literature visible to an extent previously unknown.

Griswold’s work (1981) began with a critique of reflection theory’s exclusive focus on “deep” meaning, demonstrating the importance of production variables such as copyright legislation for explaining the diversity of books available in a market. A second project (1983, 1986) investigated the determinants of cultural revival, arguing that Elizabethan plays are revived most frequently when the social conditions of the day resonate with those the plays originally addressed. In 1987, Griswold published the results of a third project centrally located in the new reception of culture approach. This innovative work used published reviews as data on reception, thus allowing Griswold (1987) to address reception across time and across three very diverse audiences—an impossible strategy in the first instance and a prohibitively expensive strategy in the second when using interviews to gather data on audience interpretation. The 1990s saw Griswold (1992) beginning a large-scale project on the literary world of Nigeria, a project that returned Griswold to her initial interest in nationalism and literature among other concerns.

Griswold's impact on the sociology of literature has been powerful because she has systematically developed a methodological approach to studying literature and other cultural products and because her substantive research integrates a concern for meaning and the unique properties inherent in literary texts with an equal interest in social context, in the actors, institutions, and social behaviors surrounding texts.

Griswold's concern for the integration of literary content with social context is shared by many. Janet Wolff, although she works primarily in visual arts rather than literature, has repeatedly challenged sociological students of culture to take content and aesthetics seriously, allying these concerns with their traditional specialty in social context and history (e.g., 1992; see also Becker in Candido 1995, p. xi). Priscilla Parkhurst Clark/Ferguson (e.g., 1987) has written extensively on the literary culture of France, combining a study of specific works and authors with detailed analyses of literary institutions and social processes, in addition to her normative writings on improving the sociology of literature (1982). Corse (1995, 1997) combines a detailed reading of three types of American and Canadian novels with a historical consideration of the two nations' canon development and a survey of the respective publishing industries to create a full picture of cross-national literary patterns and the explanation thereof. These works draw upon several important new approaches developed in the last twenty years.

**The Production of Culture.** The production of culture approach was the earliest of the new paradigms reinvigorating the study of culture in sociology. It stemmed from the growing interest of several prominent organizational sociologists in the sociology of culture (e.g., Hirsch 1972; Peterson 1976). These scholars made the now obvious insight that cultural objects are produced and distributed within a particular set of organizational and institutional arrangements, and that these arrangements mediate between author and audience and influence both the range of cultural products available and their content. Such arguments stand in stark contrast to earlier nonsociological conceptions of artistic production that featured artists as romantic loners and inspired geniuses with few ties to the social world. Art, in this view, is the product of a single artist and the content of

artistic works and the range of works available are explained by individual artistic vision. Becker's influential *Art Worlds* (1982) effectively refuted such individualistic conceptions of cultural producers, at least in sociological research. Researchers in the production of culture tradition have showed conclusively that even the most antisocial artistic hermits work within an art world that provides the artistic conventions that allow readers to decode the work. Artists are free to modify or even reject these conventions, but the conventions are a crucial component of the work's context. Art worlds also provide the materials, support personnel, and payment systems artists rely upon to create their works.

The social organization of the literary world and the publishing industry became obvious focuses for sociological investigations, from the production-of-culture approach. Walter W. Powell initiated a major research project with his dissertation, which was followed by his work on *Books: The Culture and Commerce of Publishing* (Cosser et al. 1982) and *Getting into Print* (Powell 1985). This stream of research demonstrates how production variables, such as the degree of competition in the publishing industry, the web of social interactions underlying decisions about publication, and the fundamental embeddedness of publishing in particular historical and social circumstances, affect the diversity of books available to the public.

Peterson (1985) outlines six production factors constraining the publishing industry. Berezin (1991) demonstrates how the Italian fascist regime under Mussolini shaped the theatre through bureaucratic production. Long (1986) situates the concern with economic concentration in the publishing industry in a historical perspective, and argues that a simple relationship between concentration and "massification" is insufficient for understanding contemporary publishing. Similarly, although as part of larger projects, Radway (1984), Long (1985), and Corse (1997) analyze the publishing industry and its changes as a backdrop for an understanding of particular literary characteristics. Radway traces the rise of mass-market paperbacks and the marketing of formulaic fiction to help explain the success of the romance genre (1984; chapter 1). Long (1985; chapter 2) acknowledges the importance of post World War II changes in the publishing houses and authorial demographics



in her analysis of the changing visions of success enshrined in best-selling novels, although she grants primary explanatory power to changes in the broader social context. Corse (1997, chapter 6) provides a cross-national study of Canada and the United States, arguing that the publishing industry in the latter dominates the former because of market size and population density. Canada's publishing industry has become largely a distributive arm of the American publishing industry, despite governmental subsidies and other attempts to bolster Canadian publishing. The result is that American novels dominate the Canadian market (Corse 1997, pp. 145–154).

One important focus of production approaches is gender. Tuchman (1989) analyzes the movement of male authors into the previously female-dominated field of British novel publishing during the late 1800s as the field became increasingly remunerative. Rogers (1991), in her ambitious attempt at establishing a phenomenology of literary sociology, notes the gendered construction of both writers and readers. Rosengren's (1983) network analysis of authorial references in book reviewing demonstrates, among other suggestive findings, the persistence of the literary system's underrepresentation of female authors.

**Reception Theory and the Focus on Audience.** A second fundamental shift in the sociology of literature occurred as sociologists became familiar with the work of German reception theorists. Reception theory, and several other strains of similar work, shifted scholarly attention to the interaction of text and reader. The central figures in Germany in the late 1960s and 1970s were Hans Robert Jauss and Wolfgang Iser. In *Toward an Aesthetic of Reception* (1982) Jauss presents his main argument: that literature can be understood only as a dialectical process of production and reception in which equal weight is given to the text and the reader. Iser's (1978) central focus is the act of reading itself.

Janice Radway's (1984) seminal *Reading the Romance* introduced reception theory with its central interest in audience interpretation to many American sociologists, as well as to many scholars in related fields. To those already familiar with the work of reception theorists, Radway's work powerfully demonstrated the potential of reception approaches for the sociology of literature. Radway's

interviews with "ordinary" readers of genre romance novels (1984) uncovered multiple interpretations, instances of resistance, and fundamental insights into literary use and gender in a genre previously scorned as unworthy of serious scholarly attention.

Reception theory has generated a fruitful line of research in the sociology of literature. Long (1987) has examined women's reading groups and their acceptance or rejection of traditional cultural authority in the selection and interpretation of book choices. Howard and Allen (1990) compare the interpretations made by male and female readers of two short stories in an attempt to understand how gender affects reception. Although they find few interpretive differences based solely on gender, they find numerous differences based on "life experience" and argue that gender affects interpretation indirectly through the "pervasive gender-markings of social context" (1990, p. 549). DeVault (1990) compares professional readings to her own reading of a Nadine Gordimer novel to demonstrate both the collective and the gendered nature of reception. Lichterman (1992) interviewed readers of self-help books to understand how such books are used as what he describes as a "thin culture" that helps readers with their personal lives without requiring any deep personal commitment to the book's advice.

Griswold (1987) innovatively applied the reception perspective to a study of the cross-national range of published reviews of a single author, generating another fruitful line of research. Bayma and Fine (1996) analyze 1950s reviews of Vladimir Nabokov's *Lolita* to demonstrate how cultural stereotypes of the time constructed reviewers' understandings of the novel's protagonist. Corse and Griffin (1997) analyze the history of reception of Zora Neale Hurston's *Their Eyes Were Watching God*, analyzing the different positionings of the novel over time and detailing how various "interpretive strategies" available to critics construct the novel as more or less powerful.

**Stratification.** One final area of growth centers on the relationship between cultural products and stratification systems. Perhaps the central figure is Pierre Bourdieu (1984, 1993), whose analyses of class-based differences in taste, concepts of cultural capital and habitus, and examination of

the distinction between the fields of “restricted” and “large-scale” production have profoundly affected sociological thinking. Bourdieu (1984) has demonstrated how constructed differences in capacities for aesthetic judgment help reproduce the class structure. This fundamentally affects the conditions under which types of culture are produced, interpreted, and evaluated (1993). Bourdieu’s theoretical insights have inspired many researchers, although few work in literary sociology directly. For example, Corse (1997) examined the use of high-culture literature in elite programs of nation building, Halle (1992) investigated class variations in the display of artistic genres in the home, and DiMaggio and Mohr (1985) correlated cultural capital and marital selection. Cultural consumption and use are also stratified across categories other than class, for example, gender, race, and ethnicity. These categories have received even less attention than class in the sociology of literature, although some work has been done in gender (e.g., Simonds and Rothman 1992; Wolff 1990; Radway 1984).

Bourdieu, among others, has also highlighted the need for sociological understanding of aesthetic evaluation as a social process and for a recognition of the contested nature of the cultural authority manifested in aesthetic judgments (e.g., DiMaggio 1991). Although this is not a new point (e.g., Noble 1976), sociology is finally coming to terms with literary evaluations and the codified hierarchy of value as objects of sociological attention (Lamont 1987; Corse and Griffin 1997; Corse 1997).

**International Approaches.** Obviously much of the material discussed so far is international, primarily European, in origin. European social theory has always been part of American sociology—the “fathers” of sociology are, after all, European—but there are cycles of more and less cross-fertilization. Historically, European sociologists certainly evinced greater interest in the sociology of literature than did their American counterparts; an example is the ongoing series of articles in *The British Journal of Sociology* debating the state of literary sociology (e.g., Noble 1976). The reasons for European sociology’s greater interest in the sociology of literature are several: the relatively greater influence of Marxist and neo-Marxist traditions; methodological differences that legitimate

qualitative and hermeneutic traditions; and the tighter link between sociology and the humanities compared to the “science-envy” and concomitant embrace of positivism characterizing much of American sociology.

These historical differences have at least residual remains. Marxist and hermeneutic approaches and methods more reminiscent of the humanities are still more prevalent in Europe. For example, there is greater acceptance of work looking at a single novel, an approach rarely seen in American sociology (e.g., Wahlforss’s 1989 discussion of the success of a best-selling Finnish love story). Differences have decreased, however, primarily from the American embrace of European theories and methods rather than from the opposite movement.

One important group in the sociology of literature also proves a major exception to the historic differences in method between American and European sociologies of literature. The Marketing and Sociology of Books Group at Tilburg University in the Netherlands specializes in an institutional approach to understanding “the functioning of literary and cultural institutions . . . [and] the various aspects of consumer behavior towards books and literary magazines” (Verdaasdonk and van Rees 1991, p. 421; see also, for example, Janssen 1997). The group includes Cees van Rees, editor of the journal *Poetics*, which lives up to its subtitle—*Journal of Empirical Research on Literature, the Media and the Arts*. The International Association for the Empirical Study of Literature (IGEL) sponsors an annual conference concentrating on such work (see Ibsch et al. 1995).

### BROADER IMPLICATIONS

The sociology of literature has implications for wider social issues. In the debate over the opening of the canon—the question of what should be considered “great literature” and therefore required in school—people on both sides assume that reading X is different in some important way than reading Y. If not, it wouldn’t matter what was taught. Sociology of literature illuminates the process of canon formation helping to explain why certain books are canonized rather than others (Corse 1997; Corse and Griffin 1997); it sheds analytic light on processes of cultural authority

detailing who gets cultural power and how (DiMaggio 1991); and it elucidates the meaning-making activities of readers, showing what different audiences draw from particular texts (Griswold 1987). Sociological studies can help explain why people read, what they make of their reading, and how reading affects their lives. The relevance of literary sociology to the canon debates and its foundational arguments regarding the importance of extraliterary processes and structures can be seen in the increasing interest scholars outside sociology are showing in sociological variables and studies of literature (e.g., Tompkins 1985; Lauter 1991).

Similarly, many of the same questions of interest to sociologists of literature inform debates on media effects, debates such as whether watching cartoon violence causes children to act violently. This debate—and similar ones about the danger of rap music lyrics or the value of reading William Bennett's *Book of Virtues* rather than cyberpunk or social fears about Internet chat rooms—centers on the core question of what effect art and culture have on their audiences. Radway (1984), for example, asks whether reading romance novels teaches women to expect fulfillment only through patriarchal marriage—and demonstrates that the answer is a qualified yes. Corse (1997) argues that reading canonical novels is used to help construct national identities and feelings of solidarity among disparate readers. Griswold (1992) shows how the “village novel” establishes a powerful yet historically suspect sense of Nigerian identity. The question of the effect of reading—and the related question of literary use—is central to a complete sociology of literature. Although recent developments have moved us closer to answers, these are the key questions the sociology of literature needs to answer in the future.

## REFERENCES

- Albrecht, Milton C. 1954 “The Relationship of Literature and Society.” *American Journal of Sociology* 59:425–36.
- Bayma, Todd, and Gary Alan Fine 1996 “Fictional Figures and Imaginary Relations.” *Studies in Symbolic Interaction* 20:165–78.
- Becker, Howard S. 1982 *Art Worlds*. Berkeley: University of California Press.
- Berezin, Mabel 1991 “The Organization of Political Ideology.” *American Sociological Review* 56:639–51.
- Bourdieu, Pierre 1984 *Distinction*, transl. Richard Nice. Cambridge, Mass.: Harvard University Press.
- 1993 *The Field of Cultural Production*. New York: Columbia University Press.
- Calhoun, Craig 1989 “Social Issues in the Study of Culture.” *Comparative Social Research* 11:1–29.
- Candido, Antonio 1995 *Antonio Candido: On Literature and Society*, translated, edited and with an introduction by Howard S. Becker. Princeton, N.J.: Princeton University Press.
- Clark Priscilla Parkhurst 1982 “Literature and Sociology.” In Jean-Pierre Barricelli and Joseph Gibaldi, eds., *Interrelations of Literature*. New York: Modern Language Association.
- 1987 *Literary France: The Making of a Culture*. Berkeley: University of California Press.
- Clark, Roger, and Leanna Morris 1995 “Themes of Knowing and Learning in Recent Novels for Young Adults.” *International Review of Modern Sociology* 25(1):105–123.
- Corbett, Ronald P., Jr. 1994 “‘Novel’ Perspectives on Probation: Fiction as Sociology.” *Sociological Forum* 9:307–114.
- Corse, Sarah M. 1995 “Nations and Novels: Cultural Politics and Literary Use.” *Social Forces* 73:1279–1308.
- 1997 *Nationalism and Literature: The Politics of Culture in Canada and the United States*. Cambridge, Eng.: Cambridge University Press.
- , and Monica D. Griffin 1997 “Reception, Evaluation, and African-American Literary History: Re-Constructing the Canon.” *Sociological Forum* 12:173–203.
- Coser, Lewis 1972 *Sociology through Literature*, 2nd ed. Englewood Cliffs, N.J.: Prentice Hall.
- , Charles Kadushin, and Walter W. Powell 1982 *Books: The Culture and Commerce of Publishing*. New York: Basic Books.
- Crane, Diana 1992 “High Culture versus Popular Culture Revisited.” In Michèle Lamont and Marcel Fournier, eds., *Cultivating Differences*. Chicago: University of Chicago Press.
- de Staël, Madame (Anne-Louise-Germaine) [1800], 1998, *De la littérature considérée dans ses rapports avec les institutions sociales*. Paris: Infomédia Communication.
- Desan, Philippe, Priscilla Parkhurst Ferguson, and Wendy Griswold, eds. 1989 *Literature and Social Practice*. Chicago: University of Chicago Press.
- DeVault, Marjorie L. 1990 “Novel Readings.” *American Journal of Sociology* 95:887–921.
- DiMaggio, Paul 1987 “Classification in Art.” *American Sociological Review* 52:440–455.

- 1991 "Social Structure, Institutions, and Cultural Goods." In Pierre Bourdieu and James S. Coleman, eds., *Social Theory for a Changing Society*. Boulder, Col.: Westview.
- , and John Mohr 1985 "Cultural Capital, Educational Attainment, and Marital Selection." *American Journal of Sociology* 90:1231–1261.
- Eagleton, Terry 1983 *Literary Theory*. New York: Basil Blackwell.
- Goldmann, Lucien 1964 *The Hidden God*. London: Routledge and Kegan Paul.
- 1970 "The Sociology of Literature." In Milton Albrecht, James H. Barnett, and Mason Griff, eds., *The Sociology of Literature*. New York: Praeger.
- Grauerholz, Elizabeth, and Bernice A. Pescosolido 1989 "Gender Representation in Children's Literature." *Gender and Society* 3:113–125.
- Griswold, Wendy 1981 "American Character and the American Novel." *American Journal of Sociology* 86:740–765.
- 1983 "The Devil's Techniques." *American Sociological Review* 48:668–680.
- 1986 *Renaissance Revivals*. Chicago: University of Chicago Press.
- 1987 "The Fabrication of Meaning." *American Journal of Sociology* 92:1077–1117.
- 1990 "A Provisional, Provincial Positivism: Reply to Denzin." *American Journal of Sociology* 95:1580–1583.
- 1992 "The Writing on the Mud Wall: Nigerian Novels and the Imaginary Village." *American Sociological Review* 57:709–724.
- 1993 "Recent Moves in the Sociology of Literature." *Annual Review of Sociology* 19:455–467.
- 1994 *Cultures and Societies in a Changing World*. Thousand Oaks, Calif.: Pine Forge.
- Hall, Stuart, D. Hobson, A. Lowe, and Paul Willis 1980 *Culture, Media, Language: Working Papers in Cultural Studies*. London: Hutchinson with the Centre for Contemporary Culture Studies, University of Birmingham.
- Halle, David 1992 "The Audience for Abstract Art." In Michèle Lamont and Marcel Fournier, eds., *Cultivating Differences*. Chicago: University of Chicago Press.
- Hebdige, Richard 1979 *Subculture: The Meaning of Style*. London: Methuen.
- Hirsch, Paul 1972 "Processing Fads and Fashions." *American Journal of Sociology* 77:639–659.
- Horkheimer, Max, and Theodor W. Adorno 1972 "The Culture Industry." In *Dialectic of Enlightenment*. New York: Herder and Herder.
- Howard, Judith A., and Carolyn Allen 1990 "The Gendered Context of Reading." *Gender and Society* 4:534–552.
- Ibsch, Elrud, Dick Schram, and Gerard Steen, eds. 1995 *Empirical Studies of Literature: Proceedings of the Second IGL Conference, Amsterdam 1989*. Amsterdam: Rodopi.
- Iser, Wolfgang 1978 *The Act of Reading*. Baltimore, Md.: John Hopkins University Press.
- Janssen, Susanne 1997 "Reviewing as Social Practice." *Poetics* 24(5):275–297.
- Jauss, Hans Robert 1982 *Toward an Aesthetic of Reception*, transl. Timothy Bahti. Minneapolis: University of Minnesota Press.
- Lamont, Michèle 1987 "How to Become a Dominant French Philosopher." *American Journal of Sociology* 93:584–622.
- 1995 "National Identity and National Boundary Patterns in France and the United States." *French Historical Studies* 19:349–365.
- Lauter, Paul 1991 *Canons and Contexts*. Chicago: University of Chicago Press.
- Lichterman, Paul 1992 "Self-Help Reading as Thin Culture." *Media, Culture, and Society* 14:421–447.
- Long, Elizabeth 1985 *The American Dream and the Popular Novel*. Boston: Routledge and Kegan Paul.
- 1986 "The Cultural Meaning of Concentration in Publishing." *Book Research Quarterly* 1:3–27.
- 1987 "Reading Groups and the Postmodern Crisis of Cultural Authority." *Cultural Studies* 1:306–327.
- Lowenthal, Leo (1961) 1968 "The Triumph of Mass Idols." In *Literature, Popular Culture, and Society*. Palo Alto, Calif.: Pacific Books.
- Lukács, Georg 1971 *History and Class Consciousness*. London: Merlin Press.
- Noble, Trevor 1976 "Sociology and Literature." *British Journal of Sociology* 27:211–224.
- Peterson, Richard A., ed. 1976 *The Production of Culture*. Beverly Hills, Calif.: Sage.
- 1985 "Six Constraints on the Production of Literary Works." *Poetics* 14:45–67
- Powell, Walter W. 1985 *Getting into Print: The Decision-Making Process in Scholarly Publishing*. Chicago: University of Chicago Press.
- Radway, Janice 1984 *Reading the Romance*. Chapel Hill: University of North Carolina Press.

- Rogers, Mary F. 1991 *Novels, Novelists, and Readers*. Albany, N.Y.: State University of New York Press.
- Rosengren, K. E. 1983 *The Climate of Literature*. Report No. 5 of *Cultural Indicators: The Swedish Symbol System*. Lund: Studentlitteratur.
- Simonds, Wendy, and Barbara Katz Rothman 1992 *Centuries of Solace: Expressions of Maternal Grief in Popular Literature*. Philadelphia: Temple University Press.
- Time* 1998. Cover (June 29).
- Tompkins, Jane 1985 *Sensational Designs*. New York: Oxford University Press.
- Tuchman, Gaye (with Nina E. Fortin) 1989 *Edging Women Out: Victorian Novelists, Publishers, and Social Change*. New Haven, Conn.: Yale University Press.
- Verdaasdonk, H., and C. J. van Rees 1991 "The Dynamics of Choice Behavior towards Books." *Poetics* 20:421-437.
- Wahlforss, Jaana 1989 "Soita minuelle, Helena!—An Interpretation of the Content and Reception of a Popular Novel." *Sosiologia* 26(4):277-285.
- Wolff, Janet 1990 *Feminine Sentences*. Berkeley: University of California Press.
- 1992 "Excess and Inhibition: Interdisciplinarity in the Study of Art." In Lawrence Grossberg, Cary Nelson, and Paula A. Treichler, eds., *Cultural Studies*. New York: Routledge.
- Wuthnow, Robert 1987 *Meaning and Moral Order: Explorations in Cultural Analysis*. Berkeley: University of California Press.

SARAH M. CORSE

## LONG-TERM CARE

Long-term care (LTC) includes the full range of health, personal care, and social services provided at home and in the community for a continuing period to adults who lack or have lost the capacity to care fully for themselves and remain independent (Dearborn Financial Publishing 1997). A narrower but fairly common view is that LTC encompasses principally *residential* and institutional accommodation for older persons with special care needs. However, "long-term care" is a broad term, variously defined and interpreted in different countries and even within individual health and social care systems. As distinct from acute

care, LTC is not primarily curative but aims to maintain people either in the community or in residential (institutional) settings of various sorts.

Much discussion of LTC has focused, for demographic reasons, on care for elderly people. They are the main user groups in LTC, although long-term care users also include adults with special needs because of physical disability and mental illness or handicap. There are policy, finance, and practice issues inevitably raised by the mixed clientele of LTC (Binstock et al. 1996). A popular image is that LTC comprises mainly formal facilities and qualified personnel or trained volunteers delivering services in institutions and the community. However, probably the most common form of LTC worldwide and especially in developing countries is that provided by informal caregivers, particularly the families of those needing care. This can be a source of both pride and frustration for many families. There is increasing recognition of the burden and strain on informal family caregivers. In many cultures, the tradition of family care is enconced in the notion of filial piety. While this is a particularly Chinese and East Asian concept of (reciprocal) duty of care between parents and their children, the realities of family responsibility and interactions in care are readily recognizable in all cultures.

The numbers and proportions of people, elderly persons in particular, living in formal LTC *institutions* vary considerably from one country to another. The size of the population group in question is not generally a very good indicator of the need for LTC, or its likely provision, in any given country. More important are the evolution and philosophy of health and welfare of provision. Epidemiologically and demographically, the gender balance in any given population is also important, as patterns of disability and service use tend to differ considerably between males and females. However, data on LTC are notoriously difficult to compare internationally. There are considerable differences in both definition and provision of the various types of residential facilities, nursing homes, and hostels, as well as of duration of stay. In Europe, for example, the percentages of people aged 65 and over living in an institution in the early 1990s ranged from 1.8 percent in Greece, 2.4 percent in Spain, 5.1 percent in France and the United Kingdom, and 5.4 percent in Denmark and

Germany to 9.1 percent in the Netherlands. The percentages of those living with their children showed an even greater range, from 7 percent in Denmark, and 18 percent in Britain and France to 35 percent in Greece and 48 percent in Spain. In the United States, on any one day, it has been estimated that 5 percent of the older population lives in a nursing home. However, between 25 percent and 43 percent of 65-year-olds (depending on estimates) can expect to use a nursing home at some time in their lives. Indeed, the older the person, the greater the chance of living in a nursing home. A person aged 85 or over in the United States in 1985 had almost a twenty times greater chance than someone aged 65–74 of living in a nursing home (Novak 1997).

### ASSESSING NEED FOR LTC

An individual adult's *need* for LTC typically arises when physical or cognitive abilities impair the ability at any age to perform basic activities of daily living (ADL) such as bathing, dressing and toileting and, increasingly, the ability to conduct instrumental ADL such as shopping and house cleaning. To help assess need for and provision of LTC, various forms of ADL scales have been developed. Two of the best known are by Katz et al. (1963) and Mahoney and Barthel (1965). Other important instruments include the Functional Independence Measure (FIM); the principal component of the Uniform Data Set (UDS); and the Minimum Data Set (MDS), a component of the Residential Assessment Instrument (RAI). The nursing home RAI includes a set of core assessment items (the MDS) for assessment and care screening for nursing home residents on, say, admission and periodically thereafter. This enables the identification of any significant change in status and the development of individualized restorative plans of care and it has been successfully applied in the United States and many other countries, for a range of specific purposes in LTC. Another relevant scale is the SF-36 for health-related quality of life, although it is not primarily designed for use in LTC settings.

Aggregate need for LTC is influenced by a wide range of factors. Aging of populations is an important factor, but the health status of populations is perhaps even more important. Expectation of life at birth is increasing almost everywhere, and, more important, expectation of life at

age 60 and over is also increasing. The significance for LTC can be in at least two directions, depending on whether the expansion of morbidity or compression of morbidity hypotheses hold true. The “expansion” hypothesis suggests that further reductions in old-age mortality will expose the remaining population to a longer duration in which the nonfatal diseases of senescence are likely to emerge. The disabled portion of elderly life span will increase faster than the healthy and active portion. If this hypothesis holds true, with growing proportions of the “older-old” in the populations of most countries, there will be a steadily growing demand for community and institutional LTC provision. (See Table 1.)

Alternatively, the compression of morbidity hypothesis holds that mortality and morbidity will be simultaneously compressed and that lifestyle changes will reduce the risk of death from fatal and nonfatal diseases of senescence. A compression of morbidity is likely to have crucial implications for LTC need and provision, but precise details are difficult to assess. It implies that fewer people will need to live for long periods in institutions; conversely, there will probably be an increased need for social and community LTC for increasingly elderly population. Many of these people will be single and may not be very well provided for financially. This is especially true for single (usually widowed) elderly women, among whom many aspects of poverty are concentrated in many societies. They may disproportionately require low-cost or publicly funded community-based LTC. However, while there is some evidence, it is still early to state categorically that compression of morbidity will occur uniformly or continuously. Rather, there is increasing evidence of health *variations* both within and between populations and subgroups.

### SERVICES AND PROVIDERS IN LTC

**Nursing and Residential Homes.** These are the institutions many people associate with long-term care. Definitions of nursing homes vary internationally, but, we will consider a home to be in this category (rather than as a retirement home, as old people's home, a board-and-care home, or the like) if it provides a specified amount of actual nursing and personal care and attention. Internationally, a nursing home has been defined as “an

### Factors That May Affect the Demand for Long-Term Care

- Demographic:** Increase in the number of very old people; increased morbidity with increasing age; in particular, increase in numbers of older people suffering from dementing illnesses.
- Social:** Changes in the pattern of family structures and responsibilities at work and at home; increased tendency for some families to live at a distance from each other.
- Economic and consumer:** Improved financial position of many older people; older people making a positive choice about long-term care.
- Service:** Increased pressure on long-stay hospital beds; more effective use of acute hospital beds; closure of psychiatric hospitals.
- Political:** Initial stimulation through public funding of private and voluntary provision via supplementary benefit; community care legislation; attack on residential provision. Transfer of state funding to cash-limited local authority budgets since April 1993.
- Ideological:** Increasing popular support for a pluralist approach to welfare during the 1980s. Increasing reliance during the late 1980s and the 1990s on the market within health and social welfare services.

Table 1

SOURCE: Peace et al. (1997, p. 23)

institution providing nursing care 24 hours a day, assistance with ADL and mobility, psychosocial and personal care, paramedical care, such as physiotherapy and occupational therapy, as well as room and board” (Ribbe et al. 1997, p. 4). In general, nursing homes provide the highest level of nursing cum medical care outside acute hospitals. In many countries, especially in the developed world, they are licensed and inspected and subject to a variety of rules and standards. Some nursing homes have a medical or nursing teaching affiliation. Although the majority of long-term care is home-based and provided by informal caregivers (see below), nursing homes still epitomize the popular view of LTC, especially among those responsible for financing and legislating in this area. Indeed, this view of institutional care as the dominant form of LTC is rather difficult to change and represents a challenge to social scientists.

*International Variations in Nursing Home Provision* The proportions of people living in nursing homes and the per capita provision manifestly vary internationally, compounded by differences in definitions of nursing and residential homes. (See Table 2.) However, relative preponderance can be illustrated from a multicountry study of provision and structure of LTC systems. The countries studied are developed nations with high life expectancies; they include Sweden (with the oldest population in the study), Iceland (with the youngest), and Japan (forecast to have the highest aging rates in the coming three decades). Between 2 and

5 percent of elderly people were found to live in nursing homes, with variable percentages in residential homes (although the definitions also varied). Iceland, with the youngest age structure, actually showed the highest rates of institutionalization and nursing home residence, while Sweden, with the oldest, had a relatively lower percentage. This study found little correlation between the aging status of a country and the number of nursing home beds. Institutionalization rates differ as much according to population age structure and need as to differences in organization and financing of LTC services, the responsibility assumed for the care of elderly people by various sectors and the availability of LTC beds. Cultural factors and traditions of family care also influence strongly these levels of provision and uptake in any given country.

**Residential Homes.** Residential homes offer lower levels of care compared with nursing homes and they, rather than nursing homes per se, often provide the bulk of serviced accommodation for elderly people and adults with moderate disabilities. The terminology, scale, standards, registration, and licensing of “residential homes” vary considerably internationally. This category can include old people’s homes, old-age homes, homes for the elderly, residential homes, board-and-care homes, boarding homes, and elderly hostels, among others. Many are small-scale, board-and-lodging establishments with few residents, no specialist staff, and very few facilities. Others are specialized

**Percentage of People Over 65 Years of Age Living at Home and in Institutions  
(prevalence data; different years in the early 1990s)**

Place of Residence	Country								
	U.S.	Japan	Iceland <sup>a</sup>	Sweden	Denmark	Netherlands	U.K.	France	Italy
Own home, independently or with informal and/or formal care (including domestic help and home nursing)	–	94.0	87.0	94.0	85.0	90.0	93.0	94.0	96.0
Residential homes, home for the aged, old people's homes (low levels of care)	1.5 <sup>b</sup>	0.5	5.0	3.0	10.5 <sup>c</sup>	6.5	3.5 <sup>d</sup>	4.0	1.0
Nursing homes (high levels of care)	5.0	1.5	8.0	2.0	4.0	2.5	2.0	– <sup>e</sup>	<2.0
Hospitals (intensive medical care)	–	4.0	–	<1.0	<1.0	<1.0	1.5	–	1.0

Table 2

NOTE: <sup>a</sup>Including only elderly of  $\geq 67$  years.

<sup>b</sup>Including only residential care homes and not group facilities such as board and care homes.

<sup>c</sup>Including some sheltered housing and other special dwellings for elderly.

<sup>d</sup>Including some young disabled.

<sup>e</sup>No facilities described as nursing homes; 2 percent of elderly reside in nursing-home-like facilities.

SOURCE: Ribbe et al. (1997, p.6)

and provide high-standard, often high-cost, residential care with full meal service, facilities, grounds, and services. In some systems, most are licensed and inspected (above a certain size); in other systems, regulations are far more relaxed.

Admission to a residential setting often stems from an inability to manage at home because of difficulties with activities of daily living (ADL). In some residential homes, assistance is given with basic activities of daily living. In others, admission may be for social reasons of company or a lack of wish to maintain a home. However, while most residents will be basically ambulatory, in many aging nations there is a growing overlap among residents in residential homes and those in nursing homes. Indeed, increasing age and deteriorating health status can alter the case mix of some residential homes so that they sometimes resemble nursing homes but lack specialized facilities and staff. This can present a serious problem to the delivery of appropriate and quality care.

**Sheltered Housing.** Sheltered housing is generally purpose-built accommodation in which residents live in their own unit, with their own front

door, but in a group development, with a system of linkages or alarms and served (usually full time) by a supervisor or warden. A range of services such as cleaning, shopping and entertainment, and common facilities may be provided on site, but these vary a great deal among sheltered housing schemes. Some schemes are small scale and involve only a handful of houses and residents. At the extreme, retirement villages established for older adults, increasingly common in North America and Australia, may be regarded as a type of sheltered housing. In many countries, various types of congregate housing are central to LTC.

Sheltered housing may be publicly provided, but it can be individually owned and provide a means for elderly people to remain in owner-occupancy. Increasingly, in countries such as the United Kingdom, a distinction is emerging between “ordinary” sheltered housing and “very sheltered” accommodation. In sheltered accommodation, residents generally require little more than suitable housing with the moral support of an alarm system to call emergency assistance. Very sheltered accommodation provides a greater intensity of services more akin to those in residential



care settings. Many people feel that sheltered or very sheltered housing is an ideal form of accommodation for LTC delivery, and it certainly does have many positive aspects when the designs are appropriate and residents able largely to live independently. However, the question of what happens when tenants become older and more frail is difficult, as are the potential for ghettoization of elderly people (Tinker 1992, 1997). In a high-density accommodation society such as Hong Kong, sheltered housing has been adapted to accommodate unrelated single elderly people for whom the provision of totally independent homes has proved difficult. Since the late 1980s, under the Housing for Senior Citizens scheme, public housing apartments have been provided. Usually three elderly people share an apartment. Each has an individual bedroom, and they share a communal living room and kitchen. Alarm systems and a warden provide continuity of contact, and certain communal activities are arranged for residents.

**Care for Younger Adults: Psychiatric and Group Homes.** Younger adults are often in need of LTC because of physical or mental incapacity. Sometimes, specialist care and accommodation is required because of, for example, spinal cord injuries. Younger LTC residents often also have different needs and perspectives from older recipients: a goal might be to find ways to assist younger people get out and about, commute to work, enjoy a full range of activities, and be contributing members of society. Home- and community-based services are as important for the younger as for the older persons in need of LTC (Binstock et al. 1996). However, some younger adults, for example, with severe mental or physical problems, might remain in LTC psychiatric accommodation or community-based group homes or hostels for a very long period of time, while receiving various forms of adult day care or respite care (see below).

**Home and Community Care.** Community and home-based services have the main aim of enabling people to continue living in their own homes or in the community for as long as possible. They involve a wide range of types of services and facilities provided by the formal and informal sectors. Home health care programs deliver health care and related services to people's homes; they have been called "hospitals without walls." Again, as in LTC generally, definitions of "home care" vary among countries. A recent study of fifteen

countries in Europe found considerable variation and defined *home care* as care provided at home by professional home nursing organizations and home help services. Other professionals—such as general practitioners (family doctors), occupational therapists, and physiotherapists—were excluded, although it is recognized that they do have a clear function in delivering care and often enable people to continue living at home (Hutten and Kerkstra 1996). *Home nursing services* include rehabilitative, promotive, preventive, and technical nursing care, with an emphasis mainly on the nursing of sick people at home. *Home help services* can provide a wide range of care, including shopping, cooking, cleaning, and laundering. They sometimes help with dressing and washing, and they often help care recipients to do administrative paperwork, pay bills, collect pensions, and the like. Many studies show that a home helper is also valued for providing company and someone to talk to. Clients are generally elderly (for example, in Britain, almost 90 percent). In some countries, the growing frailty of clients has led to the development of more intensive home care schemes that provide personal care in addition to doing simple cleaning and other tasks.

Several factors have raised the importance of home-based care: *increasing demand* from aging populations; *policies of substitution* of home care for institutional care (hospital care for the sick and residential care for elderly people) because of health care costs; and the *changing nature of home care* itself, with its increasing ability to deliver innovative services at home. However, while home is becoming the venue for delivery of care for many people of all ages, it is not necessarily a cheaper option than institutional care, especially for those requiring twenty-four-hour care. In addition, its future is not necessarily troublefree. The increasing numbers of older-old mean that complex home and family settings may evolve. For example, there will be increasing numbers of adult children in their sixties caring for parents in their eighties or nineties; likewise, there is an increasing number of elderly parents caring for adult disabled children at home. This is likely to require a very complex mix of support services to be delivered in the community and to people's homes.

**Community-Based Care Facilities.** These are also very varied and can provide care or more

informal meeting, contact, and social support. *Adult day care*, variously titled, is a community-based provision that provides health, social, and related support services in a protective setting but for less than 24 hours on a daily or less frequent basis. Such facilities are becoming increasingly important in the spectrum of LTC services. Adult day care may follow the medical model or the social model or some combination of both. Day care is sometimes provided in a clinic or hospital, where it may have a specific clinical or rehabilitation aim; it may also be provided in a residential home, a day center, or a club. Purpose-built and converted day centers or clubs often provide social contact, recreation, and education. Some offer meal service, and they may act as a base for home delivery of services such as meals on wheels. Many adult day care centers provide transportation, and some have medically trained staff members. Day care is often helpful in relieving relatives from the care of an elderly person or a disabled adult for a few hours a day; recent innovative services provide care in the evenings or at night. An extension of this is respite services (see below) which may be based in the community or in an institution.

**Family and Informal Caregivers.** The majority of informal caregivers are family members, of closer or more distant relationship, but some are friends and neighbors. While increasing attention is paid to the professionalization of community-based LTC in most developed-world countries, informal caregivers undoubtedly provide the major portion of care at all levels of need at home. Ironically, ideological shifts from hospital to community in the process of deinstitutionalization have, as a by-product, increased the importance of informal and family caregivers, but often in the absence of full-fledged community care systems. In some societies, especially in the developing world, the family is popularly and sometimes officially regarded as the main and only source of long-term care. Considerable shame and disapprobation can descend on the family that neglects the care of its elderly members, in more traditional societies. Indeed, even in highly developed societies, there is often popular pressure or at least an assumed expectation that the family will take responsibility for LTC of its immediate members, even if this is provided with paid assistants. This can hold true even when the individual is resident

in formal LTC settings and family members have visiting or care-providing roles.

Informal caregivers in LTC are predominantly female. American studies show that family caregivers are wives (23 percent), daughters (29 percent), other females (20 percent), husbands (13 percent), sons (8 percent), and other males (7 percent). This pattern holds true in many cultures. In China, for example, it provides an underpinning for the strong belief in the importance of having a son: Once a daughter marries, she traditionally lives with her husband and his parents, thus depriving her own parents of her potential future LTC help and domestic or financial assistance. Research in the United Kingdom and elsewhere indicates that family care is central to the lives of a substantial proportion of older people. In many countries, however, social change—including smaller families; women increasingly in the workforce; greater longevity; and social factors such as divorce, remarriage, and migration—all render the caring potential of the family more difficult. In developing countries, there is often very strong rural-to-urban migration of young working-age people. This can often render rural elders bereft of children who would provide day-to-day help, especially in the absence of formal LTC provision.

The reciprocal nature of family care is increasingly recognized—care flowing from parents to children, grandparents to grandchildren, and in reverse. This is often intricately bound up with exchange relationships, duties, and inheritance patterns. Elderly people themselves also have key roles as caregivers and, as noted above, among retired couples, the principal LTC provider is generally a spouse, often as old as or older than the recipient. Elderly parents often shoulder great responsibility for the LTC of their adult children with developmental disabilities. Indeed, the aging of children with mental retardation and severe physical disabilities has become an acknowledged research and policy issue only over the last two decades, with the shift from an emphasis on early childhood concerns to those of life-span issues. The progressive aging of almost all societies has increased the average age of caregivers of both elderly people and disabled adults. This necessitates important shifts in policy toward, for example, rendering housing more amenable and providing aging generations with greater frequency

and variety of support, including respite care (see below).

In spite also of the strong emotional preference for being cared for by family members, many older people and their families, as well as the families of disabled younger adults, increasingly favor the support of professional caregivers when there are extensive care needs. This can be a recognition of *caregiver burden*—the stresses placed on family members by their caring responsibilities—but it is also a practical recognition of the limitations of families and informal caregivers. Researchers are increasingly focusing on the problems of family caregivers, especially those coping with physical and mental illness and especially those associated with disorders such as dementia. The impacts of providing LTC on family life, privacy, and on other members of the family besides the primary caregiver are also increasingly being recognized.

**Respite Care.** Respite care is temporary, short-duration, usually residential care with supervision and/or nursing provided for dependent adults, who typically cannot be left alone to live. The purpose is to provide caregivers (usually family members but often formal live-in caregivers) with temporary relief or respite from their caring roles. This clearly recognizes the family context within which much LTC is provided and acknowledges the burden that can be placed on caregivers. It is also a pragmatic recognition that, given appropriate breaks and holidays, family members may be able to continue to provide LTC support that would otherwise not be possible.

Respite care has the explicit aim of relieving the primary caregivers of impaired elderly persons, and its use has been extended to other caregivers of handicapped nonelderly adults. Many caregivers report that the main service they would like is to be given “a break” from caring duties and the chance for a holiday or participation in social activities. While this is undoubtedly the case, evidence is not conclusive that these services significantly reduce caregiver stress. In addition, a range of other services, such as meals on wheels, transport services, befrienders, and home help services might provide some of the functions of short-term respite care, yet they are not classed as such and are not really in this category. The major problem with respite care is that it is only as a rule available

on a local basis; it has not generally received high funding priority, and caregivers therefore often feel they cannot rely on it to be available if and when they require it. In addition, there is evidence that, for various reasons, respite care places are often not fully occupied, a feature which renders the service vulnerable when budget reviews are undertaken.

### FUNDING OF LTC

Sources of financial support for and cost of LTC vary considerably internationally and are likely to become even more variable as many countries undertake health sector reform and reassess public expenditure on welfare. Classically, LTC falls between health and social care; by its very nature, it lies at the interface between the two. Many public or insurance schemes have been able or willing to pay only for eligible clients who are in the residential or nursing home medical end of the LTC spectrum. By contrast, community-based LTC services for individuals and client groups have emerged piecemeal and have often been excluded from private insurance or public welfare payments. They can be very much a patchwork, dependent on local historical provision and the evolution of national programs. Care by the family, neighbors, and volunteers have also been of extreme importance for LTC recipients in the community, so individual resources and funding remain critical in many systems. At least five public policy options (other than total private reliance) exist for funding LTC. These include means-tested public funding; private insurance; public-private partnerships; social insurance; and funding from general taxation. Increasingly, there is an international focus not purely on value for money or efficiency and quality of LTC (as these are notoriously difficult to define) but on the concept of “best value.”

Funding of LTC is fraught with emotion among the public and politicians in many countries, and the moral panic (excess concern about future numbers of elderly people, in particular, and escalation of costs) has frequently made rational debate difficult. The issue is of major concern in the United States, where policies and, by implication, funding, are under considerable debate (O'Brien and Flannery 1997). Welfare states such as the United

Kingdom have in the past been based on a combination of free National Health Service provision for health care aspects and social welfare payments for the social care aspects. Increasingly until the early 1990s, there was assistance from local authorities and from generous publicly underpinned funding of private sector homes for elderly people. Private LTC homes flourished during the 1980s. Much of this has changed in the mid-1990s since an extensive review of the nature and funding mechanisms of community care and the retrenchment of more or less automatic public sector support for residents of private old people's homes. A new approach to privatization has resulted in more stringent assessment of residents in private homes who are to receive public support to target the most needy with care packages. Many private old people's homes have subsequently gained vacancies following this policy shift (Bartlett and Phillips 1996).

Comparative analysis of international systems of LTC can be valuable. A study of the evolution of the United Kingdom's LTC policies and those of Israel, which are based on a system of national insurance, attempts to draw some lessons for the United States (Cox 1997). Israel's 1988 Community Long Term Care Insurance Law sought to ensure that all of the state organizations involved in LTC of elderly people would be integrated into one system. Israel attempted to calculate of costs of alternatives to institutionalization. It was estimated that some 60 percent of disabled elderly people on waiting lists for institutional care could, through a basket of services, home care, day care, and personal care to relieve the burden on families, be cared for in their homes at substantially lower costs than those of the institutions. This is a thorny issue, rarely adequately addressed. Modeling exercises have suggested that, in the case of the United States, for example, efficient allocation of home care services can produce some net LTC cost savings. Greene et al. (1998) estimated that reducing nursing home use and using an optimal allocation of home care services could achieve savings of around 10 percent in the overall LTC costs for an identified frail elderly population. However, this involved targeting and a more medically oriented mix of services than has been implemented to date. It appears that the expansion of personal care services may not be significantly

justified in terms of cost containment or potential to reduce nursing home use.

Some countries have addressed the philosophy and financing of LTC systematically. The United Kingdom appointed a Royal Commission to review the long and short-term options for a sustainable system of funding LTC for older persons, both in their homes and other settings (Royal Commission on LTC 1999). They recommended coverage by "risk pooling," rather than private insurance, and services underwritten by general taxation. The report of the Commission provides useful sources of costs and means of delivery. *Japan*, for example, likely in the next two decades to become proportionately the most elderly country in the world, demonstrates the social, economic, and political importance of LTC. Japan has experienced increased lengthening and seriousness of LTC, with one out of every two bedridden persons being bedridden for two years or more. More than half the care attendants are themselves aged 60 or older. The percentage of elderly people living with their children has fallen to just over one-half, and the proportion of women working is increasing. The Japanese Ministry of Health and Welfare estimates that the percentage of people aged 65 and over who need care will rise from 11.8 percent in 1993 to 14 percent in 2010 and even 16 percent in 2025.

As a result of these pressures, in December 1997, legislation was passed for a new public LTC insurance scheme—*Kaigo Hoken*—to be fully operational from 2000. This makes Japan only the third country, after Holland and Germany (see below), to provide such insurance. These proposals are very significant because they depart radically from the Japanese tradition that families are primarily responsible for long-term care. Eligibility criteria will no longer take into account the extent of informal care available to patients; and ultimate responsibility for care will lie with the state rather than with families. The scheme separates LTC services from medical care insurance and will pay for institutional and home-based care not only for those aged 65 or more, but also for people over 40 years old with "age-related" diseases such as dementia. Each municipal government is deemed a provider, and levels of services are decided by the patients' impairment. Half the

funding is from monthly premiums levied on people over 40, with a 10 percent co-payment at the point of service; rates are altered for those on low income. The rest of the funding comes from general taxation.

The scheme has two potential drawbacks. First, health and social services professionals in each municipal government will have to assess eligibility and decide on care plans, skills that have long been neglected in Japan. Second, the mechanism of quality assurance has not been clearly defined. The scheme nevertheless is a major departure in Japanese social policy (Arai 1997). It aims to underpin the development of a more diverse system in which users have more choice and may use the services they wish. It also hopes to promote the change from the excessive reliance on geriatric hospitals to sanatorium-type wards with suitable environments for long-term recuperation.

*Germany* introduced a new LTC insurance program in 1995. This required a considerable shift in focus from traditional rights to curative medical procedures toward long-term residential and home care, which previously had to be paid privately or from means-tested social assistance. There was political agreement that LTC costs threatened ordinary citizens with impoverishment in later life, a contrast with high levels of social expenditure available for health and pensions. After extensive public debate, a public insurance system was chosen, which requires equal contributions from all employers and employees; pensioners also contribute. Benefits from the LTC insurance may be taken in kind or in cash for home care and in kind for institutional care. Benefit payment levels are graded according to fairly narrow dependency definitions: dependent, seriously dependent (in need of extensive care), and very seriously dependent (needing extensive twenty-four-hour care) (Evers 1998). The cash alternative for home care is designed to allow private family-based arrangements to be made. Various additional benefits include short-term respite care to allow family caregivers four weeks' holiday, subsidized housing improvements, and contributions to the social security system. The new scheme, ostensibly a simple extension of the German insurance-based welfare approach, does create reliable rights for everyone in serious need of care as opposed to

means testing. However, there have been discussions about the appropriateness of linking care costs to employer and employee payments. In addition, the LTC fixed-sum payments are likely to meet only part of the total costs of care. Residents of nursing homes in particular still have to spend individual resources, albeit at a slower rate. Due to the high proportion of recipients of home care initially opting for cheaper cash reimbursements, the system went into early profit and has amassed a surplus. This preference for cash has been taken to indicate the importance of the family in *Germany's* LTC, but some higher-cost medical care services can still be reimbursed from health insurance.

By contrast, even in some countries with advanced approaches to welfare, full-fledged support for long-term care remains highly dependent on local variations. In *Canada*, for example, what is regarded as continuing care has been called a "patchwork quilt" of long-term care, mainly institutional, and home or community care, mainly noninstitutional. The institutional components are widely provided across the country, with public and private funding. The noninstitutional components are developing, unevenly provided and also funded publicly and privately. While access to acute care is guaranteed, this is not the case for continuing care. The amalgam of services that compose continuing care involve the community-based and residential LTC services outlined above. Funding varies by province in *Canada*, with decreasing federal support, and is typically still complicated by the situation of LTC at the interface of health and social welfare. Various types of continuing care are cost-shared between province and municipality of residence. Ownership of facilities nationally falls under three categories, and the relative balance varies considerably between provinces: public (49 percent), private for-profit (26 percent), and private nonprofit (25 percent) ownership. For illustration, *Quebec* has 78 percent and *Nova Scotia* 36 percent publicly owned facilities on average, but 100 percent of care facilities for advanced chronic mental and physical conditions are publicly owned in *Nova Scotia*. While funding comes from a variety of sources, there is a move toward universal rather than means-tested benefits, and cost sharing with users is the norm. In most provinces, people are not required to spend down their assets to receive public support,

though they will usually have to co-pay daily rates for lower levels of care. Community-based nursing services are virtually universally supplied in Canada, although home care and support vary considerably. The picture is evolving, with increasing emphasis on noninstitutional care and budget reallocation to this sector.

### QUALITY ISSUES

Quality of provision in LTC is of increasing international interest and revolves around standard setting, regulation, inspection, and ethics. It can require legislation and a mature and sensitive appreciation of care needs on the part of LTC providers and the many professions involved. Service quality in LTC residential settings has at least five dimensions (Duffy et al. 1997): tangibles (physical facilities, equipment); reliability (ability to perform promised services); responsiveness (willingness and promptness); assurance (conveyance of confidence and trust by personnel); empathy (caring, individualized attention provided by the facility). There is also an important economic aspect to quality related to costs of services delivered: value and satisfaction.

In addition to broader quality issues, specific features such as the use of restraints also give indications of the quality and respect for individuals in a LTC system. In this and other respects, international standards vary. For example, an eight-country study of use of restraints in nursing homes (Denmark, France, Iceland, Italy, Japan, Spain, Sweden, and the United States) found considerable variation. Total use of restraints was less than 9 percent in Denmark, Iceland, and Japan; between 15 and 17 percent in France, Italy, Sweden, and the United States; but almost 40 percent in Spain. The intensity of use of restraints and the types applied varied. In all countries, there was a constant increase in the use of restraints with increasing ADL difficulties and cognitive dysfunction (Ljunggren et al. 1997). It is clear that financial factors concerned with staffing levels, staff culture and training, the balance between better and less well trained staff, and the physical nature of facilities influence levels of use of restraints. Cultural variations and the thresholds for applying restraints also vary greatly internationally. The same is undoubtedly true of many other facets of LTC.

### LTC-ISSUES OF POLICY AND PRACTICE

LTC systems have generally emerged as a series of incremental responses to a growing problem or set of problems, over a number of political eras and in many jurisdictions. The result is often a medley of noncomplementary and even conflicting programs for people with chronic disability, which are typically ineffective and inefficient. In many countries, especially in the developing world, there is still enormous emotional and even official attachment to care by the family. At the extreme, placing one's parents or relative into an LTC residential facility can be anathema. Even in countries such as the United States, with expensive health care systems and an established range of community-based programs, LTC does not work well.

It is doubtful whether a universal blueprint for LTC can be devised. The development and financing of long-term care services in many countries has been at best haphazard. Until the Japanese and German experiences, and perhaps those in Australia and the United Kingdom at present, elsewhere it has rarely been asked what the ultimate LTC system should look like. Most systems have evolved piecemeal and have suffered from being at the interface of technically expanding and expensive health care and lower-technology, but fragmented, social care.

Basic policy issues and questions arise. What should LTC encompass? How can the role of housing policy in LTC be recognized? Is LTC an individual, a family, an employer, or a state responsibility, or is it (increasingly) some complex combination of all of these? What is the role of the family, and how can it be supported? Should LTC be provided to various disability groups and ages? If so, how? What is the ideal balance and relationship between LTC and the acute health care sector? What should be the balance between the various residential, social care, nursing, ancillary services, and informal care? What resources should be devoted to home care, assisted living, and other options? Underpinning many of these questions is the basic philosophy within specific countries regarding collective and individual responsibility for vulnerable citizens.

Housing is clearly crucial to the success of LTC. The internal home environment and its interaction with the local external environment adds to the challenge of life for many people with

chronic disabilities, and their ADL are compounded by their environment. Initiatives have appeared such as sheltered housing (assisted daily living), with wardening and support services, residential homes, and various schemes to assist adaptation of existing dwellings, but, in most countries, there has been little coherent policy development in this area. Ironically, many initiatives to build specialist housing for elderly or handicapped people have come from private sector developers, who have identified a new and underserved market segment for specialist housing. Aging in place and care in the community can often be enabled by relatively minor housing improvements. These may be administrative; for example, in public housing, there may be provision for flexibility enabling residents to exchange current units for more appropriate, often smaller ones. Many academic disciplines, including sociology, gerontology, planning, geography, environmental psychology, and architecture, can contribute to this important aspect of policy development.

What will future LTC look like? It can be provided in a number of different settings, by a range of personnel, and to people with varying types and levels of disability. Nevertheless, in many systems, institutional care such as nursing homes still represent the majority of expenditure on LTC. Residential and nursing homes do indeed often represent an important care option, but they would be unlikely to occupy such an important position in a newly designed LTC system. As many systems have gradually emerged piecemeal from a range of initiatives, they do not meet the needs of the recipients or providers of LTC today, nor the needs of those who pay for care. Internationally, there is little agreement about potential policy solutions. In many countries, thoroughgoing debate and reappraisal need to be initiated.

## REFERENCES

- Arai, Yumiko 1997 "Insurance for Long-term Care Planned in Japan." *The Lancet* 350(9094):1831.
- Bartlett, Helen, and David R. Phillips 1996 "Policy Issues in the Private Health Sector: Examples from Long-Term Care in the U.K." *Social Science and Medicine* 43(5):731-737.
- Binstock, Robert H., Leighton E. Cluff, and Otto von Mering 1996 *The Future of Long-Term Care: Social and Policy Issues*. Baltimore, Md.: Johns Hopkins University Press.
- Cox, Carole 1997 "Long-Term Care: A Comparison of Policies and Services in Israel and the United Kingdom and Implications for the United States." *Journal of Aging and Social Policy* 9(2):81-99.
- Dearborn Financial Publishing 1997 *Long-Term Care*, 2nd ed. Chicago: Dearborn Financial Publishing.
- Duffy, Jo Ann, Michael Duffy, and William Kilbourne 1997 "Cross National Study of Perceived Service Quality in Long-Term Care Facilities." *Journal of Aging Studies* 11(4):327-336.
- Evers, Adalbert 1998 "The New Long-Term Care Insurance Program in Germany." *Journal of Aging and Social Policy* 10(1):77-98.
- Greene, Vernon L., Jan Ondrich, and Sarah Laditka 1998 "Can Home Care Services Achieve Cost Savings in Long-Term Care for Older People?" *Journal of Gerontology: Social Sciences* 55B(4):S228-238.
- Hutten, Jack, and Ada Kerkstra (eds.) 1996 *Home Care in Europe: A Country-Specific Guide to its Organization and Financing*. Aldershot, England: Arena Publications.
- Katz, S., A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe 1963 "Studies of Illness in the Aged: The Index of ADL: A Standardized Measure of Biological and Psychosocial Function." *Journal of the American Medical Association* 185:914-919.
- Ljunggren, Gunnar, Charles D. Phillips, and Antonio Sgadari 1997 "Comparison of Restraint Use in Nursing Homes in Eight Countries." *Age and Ageing* 26(S2):43-47.
- Mahoney, F. I., and D. W. Barthel 1965 "Functional Evaluation: The Barthel Index." *Maryland State Medical Journal* 14:61-65.
- Novak, Mark 1997 *Issues in Aging: An Introduction to Gerontology*. New York: Addison Wesley Longman.
- O'Brien, Raymond C., and Michael T. Flannery 1997 *Long-Term Care: Federal, State and Private Options for the Future*. New York: Haworth Press.
- Peace, Sheila, Leonie Kellaher, and Dianne Willcocks 1997 *Re-evaluating Residential Care*. Buckingham, England: Open University Press.
- Ribbe, Miel W., Gunnar Ljunggren, Knight Steel, Eva Topinkova, Catherine Hawes, Naoki Ikegami, Jean-Claude Henrard, and Palmi V. Jonnson 1997 "Nursing Homes in 10 Nations: A Comparison Between Countries and Settings." *Age and Ageing* 26 (S2):3-12.
- Royal Commission on Long Term Care 1999 *With Respect to Old Age: Long Term Care-Rights and*

*Responsibilities*. Report by the Royal Commission on LTC. London: The Stationery Office.

Tinker, Anthea 1992, 1997 *Elderly People in Modern Society*. London: Longman.

DAVID R. PHILLIPS

## LONG-TERM CARE FACILITIES

### THE SPECTRUM OF LONG-TERM CARE

While the nursing home remains the most prevalent option for elderly persons requiring long-term care, in recent years there has been a great upsurge in alternative options. Changes in the spectrum of long-term care services provided have come about primarily based on changes in the demographic profile of long-term care users, in health care financing, and in policies impacting the provision of long-term care.

Nursing homes reflect the major long-term care option that has been available in the United States under Medicare and Medicaid financing. However, recent trends reflect new developments targeting self-financed long-term care, particularly in the form of assisted living, continuing care retirement communities, and special need facilities focusing on dementia or hospice care. In addition, discharge policies of acute care hospitals have resulted in the development of subacute care and rehabilitation facilities sponsored by hospitals or nursing homes. There has also been an increase in the provision of home care services to frail older adults who can forestall entry to sheltered living arrangements, including nursing homes, through the use of these services. Growing recognition of the diversity of elders requiring long-term care is leading to the diversification of settings to meet these needs, and to a blurring of boundaries between home care and residential care (R. A. Kane, 1995–1996).

In spite of the proliferation of new residential options for frail and old-old adults, those requiring extensive assistance or supervision continue to rely on nursing homes as the only viable publicly financed alternative. Consequently, this essay will only briefly touch on some of the more recent options in the spectrum of long-term care, such as

assisted living facilities and continuing care retirement communities. Rehabilitative or subacute care units in nursing homes will not be addressed because they reflect mechanisms for providing short-term care. Nor will we discuss senior housing sites or retirement communities, which serve as residential options to well elders. After considering assisted living facilities and continuing care retirement communities, the major focus of our discussion will be on the nursing home.

### ASSISTED LIVING FACILITIES

Assisted living facilities (ALFs) have taken over some of the original functions of nursing homes, providing services to moderately frail elders in the context of a homelike residential setting that includes less surveillance and regimentation, and more privacy and autonomy for residents than the traditional nursing home (R. A. Kane 1995). ALFs combine housing with personal support services such as meals, laundry, housekeeping, and maintenance services. Board and care homes and personal care homes may also be included under the ALF designation. Generally, residents in ALFs pay on a month-to-month basis, although financing of home health services may occur under Medicare (Pearce 1998). Assisted living facilities are an outgrowth of the group home tradition, generally reflecting small business operations. In recent years, however, assisted living facilities have increasingly been sponsored and built by hospitals, nursing homes, and large publicly traded companies (Meyer 1998).

Consistently with the model of social care (R. A. Kane and Wilson 1993), ALF services are geared toward meeting resident preferences and wishes, and consider residents to be consumers rather than patients or clients to be taken care of according to staff-based directives. Assisted living arises from a commercial model of fees for services, and its consumer-driven orientation differs from nursing home care, which generally reflects third-party payment dynamics. While these advantages make assisted living a highly attractive long-term option for frail elders with some means, potential problems include lack of regulation and inaccessibility of this type of care to the poor. Furthermore, assisted living facilities can serve only the needs of older adults with limited disabilities (R. A. Kane 1995–1996).



### CONTINUING CARE RETIREMENT COMMUNITIES

Continuing care retirement communities (CCRCs) may also be termed "life care communities," and they typically offer a continuum of services from independent living to assisted living and nursing care. These facilities are appealing to well old-old persons who are still able to live independently but want to have the security of planning for increasing service needs as they arise. These older adults want to know that their changing personal and health care needs can be met within an organizational framework, which they select. Most elderly persons initially enter independent living units of CCRCs. Between 50 percent and 70 percent of entering residents eventually utilize assisted living and/or nursing home components of the CCRC (Newcomer et al. 1995).

CCRCs vary greatly in both the types and the range of services offered and in sponsorship and management. The fees may be structured as a refundable entry fee along with a monthly service fee, or it can be based on an endowment, a rental, or a condominium purchase (Pearce 1998). While fee structures differ, arrangements for these services have generally moved from exclusive reliance on entrance fees to requiring monthly maintenance fees along with fees paid at the time of initiation. With the rapid growth of CCRCs around the country, states have enacted legislation to protect older adults, particularly against bankruptcy of the organizations sponsoring the CCRC (Netting and Wilson 1994). These facilities are generally considered and regulated as insurance programs, and there has been great inconsistency in ways of implementing their regulation across different states.

CCRCs represent not only an emerging comprehensive long-term care alternative, but may constitute a new paradigm for long-term care (Vladeck 1995). Managed care CCRCs can offer flexibility and a more community-responsive paradigm for delivering long-term care. This form of care draws on the know-how of diverse industries, including housing, hospitality, insurance, and health care. This allows for a cross-fertilization of different value orientations regarding care, and different approaches to financing and management.

Large-scale national research on CCRCs finds (Sherwood et al. 1997) that such facilities tend to

be utilized by well-educated, middle-class, white, elderly persons aged 75 and over, with women being overrepresented among the users. More than 85 percent had been white-collar workers (Sherwood et al. 1997). When compared with community residents, CCRC-dwelling elderly were found to be in comparable or better health and to have comparable social supports, although community-dwelling elders interact more frequently with their family members. Social research on assisted living and CCRC communities has not been extensive, and generally considers administrative and organizational dimensions of care, rather than aspects of social life and social interactions. Studies comparing CCRC residents and elderly people living in the community have noted that the former are more likely to utilize nursing units after acute hospital stays and outpatient surgery. At the same time, CCRC residents have lower rates of hospital admissions than do community-dwelling elders (Newcomer et al. 1995).

### NURSING HOMES

Nursing homes are of interest to sociologists on both the macro and the micro levels. On the macro level they reflect society's orientation to financing, regulating, and delivering long-term care services to its frail citizens who are no longer capable of fully autonomous community living. Sociologists have been particularly interested in those factors affecting long-term care delivery that reflect social construction of the reality of the lives of marginal individuals such as aged persons.

On a micro level, nursing homes represent formal organizations that regulate and control the daily lives of frail elders, while providing them with medical and social care. The tension between expectations for rehabilitative or prosthetic services and actual delivery of palliative care or of dependency-inducing treatment, with iatrogenic consequences, has been a particular source of fascination and concern to gerontologists (Baltes 1996; E. Kahana, Kahana, and Riley 1989; Kane 1995-1996). Complementing this interest has been the search for understanding person-environment interactions and interpersonal and intergroup relationships involving residents, staff, and families within the institutional context (Diamond 1992; Gubrium 1975, 1993; E. Kahana, Liang, and Felton, 1980).

The following discussion about nursing homes includes an introductory section outlining models of nursing home care, followed by a discussion of demographic profiles of residents and an analysis of nursing home characteristics and financing. Efforts to ensure high quality of care are discussed in terms of both regulatory efforts and formal interventions. We next consider the effects of institutionalization on the social lives of residents, and their adaptation to the nursing home. Discussion then turns to new initiatives in formal program development and effects of interventions on resident life in the nursing home. The article concludes with a discussion of emerging trends in nursing home care and the future of long-term care facilities in the United States.

Even as we aim to review sociological contributions to the understanding of nursing homes, it is important to note that information about this subject has been obtained through the research efforts of diverse disciplines, with epidemiologists, political scientists, social workers, psychologists, nurses, and physicians all contributing to the literature. In fact, medical sociologists have focused their investigation primarily on acute health care delivery in general hospitals, and many of the theoretical constructs they have developed are more applicable to acute illness than to chronic disability (Cockerham 1998; Charmaz and Paterniti 1999). Consequently, this discussion of the nursing home draws upon multidisciplinary sources.

### MODELS OF CARE IN NURSING HOMES

Those requiring institutional care represent a highly select group of the most vulnerable aged persons (Hing 1987), particularly since there are now many options which can keep frail elders in the community. Any therapeutic efforts to enhance the functioning of the very frail older adults living in nursing homes thus represent a regimen of limited objectives (E. Kahana, B. Kahana, and Chirayath 1999). Nevertheless, nursing home settings can serve important therapeutic and rehabilitative functions and can enhance the quality of life of frail elders by ensuring comfort and even some measure of autonomy (Agich 1993). Moving away from an exclusive focus on poor quality of care in nursing homes, recent research has begun to consider criteria for high quality of care and even

excellence in nursing home settings (Andersen 1987; Brittis 1996; Groger 1994; Looman et al. 1997).

Conceptualizations of models of care typically reflect the philosophies and traditions of the disciplines from which they originate. Sociology, medicine, nursing, the allied health professions, psychology, and social work—each has a distinct set of traditions that influence its respective conceptualizations of long-term care. Interventions arising from these diverse traditions may include therapeutic interventions (psychology), programs that enhance function (nursing or medicine), and interventions that empower patients (sociology and social work).

The *medical model* of care considers nursing homes as health care institutions designed to deliver high-quality chronic care to patients. As such, nursing homes are expected to provide competent and well-trained personnel to meet health care needs of patients (R. L. Kane et al. 1994). In principle, the criterion for successful care is improved health or at least minimal health decline of the patient. In practice, quality of care is generally approached by the organization through adherence to certain standards of care delivery, such as adequate staff recruitment and training, and high staff-patient ratios (Nyman and Geyer 1989; L. Z. Rubenstein and Wieland 1993).

The organizational climate of managed care has brought with it a new emphasis on *management-centered* interventions and care. Management-oriented approaches often treat patients as objects of intervention who must be manipulated to achieve desired objectives. One such approach, continuous quality improvement (CQI), is an organizational management framework based on older business models such as total quality management (TQM) (Schnelle et al. 1997). The goal of CQI is to make systematic improvements by employing multiple short cycles of designing, evaluating, and implementing interventions. These cycles of continuous intervention and assessment drive long-range systematic change. Attempts at CQI in nursing homes have included interventions in incontinence and physical restraints, but have been hindered by a lack of information technology and by care standards legislated by the Omnibus Budget Reconciliation Act (OBRA) of 1987 (Schnelle et al. 1997).

The *social model* (R. L. Kane and R. A. Kane 1978; R. A. Kane et al. 1998) considers the ideal model of long-term care to be a sheltered housing arrangement for older adults who can no longer function independently in the community. The goal of such sheltered living arrangements is ensuring a good quality of life for residents through encouragement of maximum autonomy in a homelike setting. The criterion for successful care is defined in terms of sustained psychosocial well-being, self-esteem, and life satisfaction of residents.

While medical definitions of care dictate provision of diagnostic and treatment activities, social definitions suggest emphasis on comfort, choice, and adaptation. Advocates of quality-of-life approaches have often argued for major redirection in the way long-term care is handled. They seek greater support for the home-based rather than the institution-based model of care (Leutz et al. 1992). The growing popularity of assisted living options in long-term care reflects this orientation. Empowering older clients to act as consumers provides a major mechanism for the social care model (Burger et al. 1996; Cox and Parsons 1994). Since residents living in nursing homes are typically too frail and powerless to demand active consumer involvement in their care, there is little evidence of actual empowerment-based approaches in nursing home settings (E. Kahana 1999).

The *Patient-Responsive Care Model* (Kahana, Kahana, Kercher, and Chirayath 1999) has been developed to recognize social, psychological, and physical needs of frail patients who are not able to function in residential settings and who require nursing home care. The Patient-Responsive Care Model is an ecological model guided by communitarian principles of sociologists (Bellah 1996; Etzioni 1993). It calls for a systematic and empathetic discernment by staff of resident perspectives and preferences. This approach is based on earlier work on person-environment fit as a determinant of resident well-being in nursing homes (E. Kahana 1973; E. Kahana 1982). Environments may be matched to resident needs by altering the physical or social environment and by enhancing staff responsiveness in an effort to improve residents' quality of life (E. Kahana, B. Kahana, and Riley 1989). This model is also consistent with recently proposed "cultural" models of

nursing home care proposed by Scandinavian scholars who advocate individualized care based on "knowing the patient" (Evans 1996).

The major requisite of patient-responsive care is an understanding and an empathetic appreciation of nursing home life from the patient's perspective. Four major areas of need, loosely based on Maslow's hierarchy of needs (1970), are to be met by staff: reducing physical distress, meeting basic physiological needs, meeting emotional needs, and meeting social needs. Since cognitively impaired and frail elders are limited in their ability to articulate their needs, a major challenge of patient-responsive care is to elicit expressions of resident needs and preferences. This model is predicated on empathetic listening to the patient's lived experience (Gubrium 1993; Savishinsky 1991) as a basis for developing Patient-Responsive Care, and does not rely on need assessments based on objective test data.

Staff behaviors along dimensions of responsiveness can range from total lack of involvement (ignoring or dismissing a problem) to hearing and validating patient concerns. Staff who are responsive to patients provide reassurance and, whenever possible, take constructive actions toward resolving problems. Eliciting information about patient needs and validating those needs are proposed as concrete requisites of providing patient-responsive care. Action components of responsive care also involve communicating with family members and with other staff in advocating on the patient's behalf. Responsiveness involves three central actions by staff: eliciting the patient's definition of need, validating the need, and acting or advocating on the patient's behalf. The pattern of responsiveness thus involves observing, questioning, listening, and responding. Exemplifying this approach, aides who care for the elderly may be trained to redefine their jobs from providing custodial care to providing therapeutic measures. For example, personnel delivering meals could easily be trained to engage in conversation with residents during meals, explaining what the resident is eating and contributing to the enjoyment of the meal. Staff providing responsive care may be empowered through this very initiative. Providing responsive care can serve to enhance staff's self-efficacy and sense of competence (E. Kahana et al. 1999).

While none of the models described above fully fit, the complex realities of frail older persons living in institutional settings within the constraints of managed care, each helps us appreciate the multifaceted nature of efforts to meet resident needs in the framework of long-term care.

### PROFILE OF NURSING HOME RESIDENTS

In 1995, about 1.4 million American elderly persons over the age of 65 were residing in 16,700 nursing homes. While 4.5 percent of the elderly are in nursing homes at any given time, 30 percent of older adults can expect to spend some time during their lives in a nursing home. The majority of residents enter a nursing home immediately after discharge from an acute care hospital (Lewis et al. 1985). Most of these patients need ongoing medical care and some may require rehospitalization (Densen 1987). Others enter a nursing home directly from the community, due to multiple risk factors including physical and cognitive frailty and nonavailability of caregivers. Additionally, during the era of deinstitutionalization of the mentally ill, some elderly were transferred to nursing homes from mental hospitals. Among social risk factors for institutionalization, studies have identified that those elderly persons who were never married, those who are widowed, and those who do not have children nearby are overrepresented (R. A. Kane 1995–1996).

A demographic profile of nursing home residents has been provided by the U.S. Bureau of the Census (1998) based on a 1995 survey of residents in U.S. nursing homes (National Center for Health Statistics 1997). According to the report, 24.7 percent of nursing home residents are men and 75.3 percent are women.

Older women are three times as likely as older men to reside in nursing homes, both because of their longevity and because they are more likely to be widowed, and moreover without caregivers.

Nursing home placements are generally due to a combination of functional impairments, mental infirmity and unavailability of caregivers (R. L. Kane et al. 1998; Rovner and Katz 1993).

Of those elderly living in a nursing home, 8 percent were between ages 65 and 74; another 42 percent were between ages 75 and 84, and 40

percent were over age 85. Although the vast majority of nursing home residents are elderly, about 10 percent are younger than age 65 (U.S. Bureau of the Census 1998).

Although representation of minorities in long-term care facilities has been increasing over time, such groups continue to be underrepresented (Bonifazi 1998). Racial composition of nursing home residents is: 89.5 percent white, as compared to 8.5 percent black and 2.3 percent Hispanic (R. L. Kane et al. 1998; U.S. Bureau of the Census 1998). Both the greater willingness of black and Hispanic families to care for elders and barriers to the utilization of nursing home by minorities have been cited as possible reasons for these racial differences (Mui and Burnette 1994; Wallace et al. 1998).

Based on the 1995 National Nursing Home Surveys, the majority of elderly persons residing in nursing homes need significant assistance with both activities of daily living (ADLs) and instrumental activities of daily living (IADLs) (National Center for Health Statistics 1997). The study reports that 96 percent of residents in a nursing home need help with showering, 87 percent need help with dressing, 45 percent need help with eating, 24 percent need help with transferring, and 58 percent need help with toileting. Limitations were also found in IADL function, with 78 percent of respondents needing assistance with care of personal possessions and 69 percent needing help with managing money. High levels of physical impairment in this group are also reflected in extensive use of assistive devices. Thus, 64 percent of all elderly in nursing homes utilize a wheelchair, 25 percent use a walker, and 78 percent occupy a hospital bed. Sensory and communicative impairments are also prevalent among nursing home residents, with 51 percent of residents having moderate to severe hearing loss (Garahan et al. 1992) and 24 percent exhibiting visual problems (National Center for Health Statistics 1997).

There is increasing recognition that the prevalence of diagnosable neuropsychiatric disorders in nursing homes is high, with estimates ranging up to 80 percent or greater (Kim and Rovner 1996; Rovner and Katz 1993). Elders with schizophrenia or other psychoses compose only a very small proportion of the nursing home population. The most prevalent neuropsychiatric disorders include

dementia (primarily Alzheimer's disease) and depression.

Over 50 percent of elderly persons admitted to nursing homes will live out their lives in a nursing facility, with about 10 percent dying during the first year of their nursing home stay. The average length of stay of long-term care residents is two to three years. In any given year, 25 percent of residents return to the community from a nursing home (Cohen-Mansfield et al. 1999). The majority of those returning to the community have spent time in short-term, rehabilitative units of nursing homes.

### CHARACTERISTICS AND DISTRIBUTION OF NURSING HOMES

The number of nursing homes in the United States has decreased from 19,100 in 1985 to 16,700 in 1995. This represents a decline of 2,400 homes over the past decade. The number of nursing home residents has increased only slightly over the same decade from 1.49 million residents in 1985 to 1.55 million residents in 1995 (U.S. Bureau of the Census 1998). In 1995, there were about 1.77 million nursing home beds. The average number of beds per nursing home increased from 85 in 1985, to an average of 106 beds per nursing home by 1995. These changes parallel organizational shifts toward the entrance of large nursing home chains into the market. At the most recent 1995 survey, 66 percent of nursing homes were proprietary, 26 percent were nonprofit, and 8 percent were sponsored by the government (U.S. Bureau of the Census 1998). Fifty-five percent of nursing homes were run by commercial organizations operating chains of facilities. Of the total nursing homes, 66 percent were certified by both Medicare and Medicaid, 20 percent by Medicaid only, 6 percent by Medicare only, and 4 percent were not certified by either Medicare or Medicaid (U.S. Bureau of the Census 1998). Nursing homes generally have high occupancy rates (87 percent) in part based on newly proliferating special units such as respite care, rehabilitative services, hospice, dementia units, and acquired immune deficiency syndrome (AIDS) special care units (U.S. Bureau of the Census 1998). Nevertheless, there are some indications that frail older adults are not entering nursing homes at the same rate they

previously did. The increase in assisted living settings and other residential options is likely to contribute to this trend (Strahan 1997).

There are regional differences in both distribution and utilization of nursing homes. Older adults in the Midwest and the South have the largest number of nursing homes available to them, each with 33 percent of the total number of nursing homes in the United States. The Northeast, containing 17 percent of the country's total number of nursing homes, has the highest occupancy rate compared to the Midwest, South, and West. The West, in contrast, has a small proportion of nursing homes and has the lowest occupancy rates (U.S. Bureau of the Census 1998). Migration patterns may influence nursing home occupancy rates. Many older adults, upon retiring, move to warmer Sunbelt states seeking to improve their quality of life. This group is likely to enter nursing homes as they become frail in old-old age, thus contributing a high demand for nursing homes in the South (Longino 1998). At the same time, when they require more extensive care, significant numbers of older adults may return from the Sunbelt to the states from which they originally migrated. These countermoves may also affect the utilization rates of nursing homes in certain regions. Rural-urban comparisons of nursing home distribution reveal that metropolitan areas have a larger supply of nursing home beds per capita than do rural areas (62 versus 45 certified beds per 100 elders, respectively) (Shaughnessy 1994).

### FINANCING OF LONG-TERM CARE

It has been argued that, in the United States, funding mechanisms have been largely responsible for shaping the delivery of health care in general and the delivery of long-term care in particular (R. A. Kane et al. 1998). In 1995, nursing home expenses totaled \$77.9 billion, sharply up from \$36 billion in 1985 (Levit et al. 1996). Medicare, a universal age-based health care program, has traditionally covered the acute and rehabilitative health care needs of older adults and reimburses only a very limited number of nursing home expenses. Medicaid, a need-based, state-administered program, on the other hand, covers nursing home expenses for older adults after their financial resources are depleted. Thus, while Medicare paid 9 percent of nursing home expenses in

1995, Medicaid paid 47 percent of the expenditures. Thirty-seven percent of nursing home expenses were paid out of pocket, whereas private insurance paid only 3 percent of nursing home expenses (Levit et al. 1996).

While private nursing home insurance is increasingly advocated as an important protection for older adults, only very few adults are currently benefiting from such coverage. This limited role played by private insurance may be attributable to high cost, limited coverage afforded by such policies, and lack of information about the availability of insurance (Cohen 1998). Most elderly individuals start their nursing home stays as privately paying patients. Nursing home care, in fact, represents the major portion (82.5 percent) of out-of-pocket health care expenses of the elderly (Rice 1989). After personal resources are depleted (usually in less than a year), patients become eligible for Medicaid financing (Lusky and Ingman 1994).

There is continuing pressure for cost containment to reduce burdens of nursing home costs. The major response to these needs has been the proliferation of assisted living facilities, CCRC facilities, and home health care. In 1995, the average annual cost of a nursing home stay was over \$38,000, while out-of-pocket costs for home health care averaged \$370 a month, or \$4,440 annually (Cohen 1998). There have also been reductions in reimbursable nursing home services by Medicare and Medicaid. Efforts directed at cost containment through reduced payments for care pose a serious threat to the provision of high-quality services. In addition, the profit motive inherent in proprietary health care may be seen as posing a conflict of interest with the provision of high-quality care.

### **ENSURING STANDARDS OF NURSING HOME CARE**

Ensuring high quality of care and standards for services has posed a major challenge to the nursing home industry (Vladeck 1980; R. A. Kane 1995–1996). Problems have been encountered in terms of poor staff training (R. A. Kane 1995–1996), high staff turnover rates (Banazac-Holl and Hines 1996), and limited physician involvement (Fortinsky and Raff 1995–1996). Nurses' aides, the

staff members with the most limited training and education and at the lowest end of the pay scale, provide the majority of patient care in nursing homes (IOM 1996). Higher ratios of registered nurses (RNs) to patients are associated with higher patient survival rates, increased functional status, and increased numbers of patients discharged from the nursing home (IOM 1996). Nevertheless, pay differentials between hospitals and nursing homes have contributed to a paucity of RNs working in nursing homes and an overrepresentation of nurses with limited educational backgrounds working in long-term care facilities (IOM 1996). It is noteworthy that nursing assistants employed in nursing homes are also less well paid than their counterparts working in hospitals (BLS 1995). Nursing assistants have limited education, with 46 percent having high school diplomas, while nearly 18 percent have not graduated from high school (IOM 1996). In addition to morale problems resulting from low-status jobs with poor pay, many nursing assistants also face personal problems exacerbated by their low socioeconomic status (Coons and Mace 1996). These factors are related to very high turnover rates among nursing assistants, resulting in poorer quality of care (Schnelle et al. 1993).

Generally speaking, assessing quality of care involves an inquiry into three key ingredients of health care quality: structure, process, and outcomes. Structural measures gauge the presence of certain provider characteristics that are thought to produce good-quality outcomes. Process measures compare the actual care delivered to standards or norms of practice. Outcome measures serve to indicate the results of the care received (e.g., death, functional change). Reliable measures of appropriate outcomes of care are sought by quality-assurance programs in long-term care (R. A. Kane, R. L. Kane, and Ladd 1998). In the context of improvement in care quality, outcome assessments may be used to allocate resources to those areas that require remediation.

Although diverse quality assurance and enhancement programs have been advocated, regulation continues to serve as the major approach to insuring high quality of care. Until Medicare and Medicaid were enacted in 1965, nursing home regulation was each state's responsibility. Once federal programs began paying for nursing home

services, the federal government became more involved in regulating nursing homes. It has been argued that Medicare and Medicaid had a major influence on delivery of care in nursing homes due to the regulations placed on facilities and the standards established for certification and for eligibility for payments (R. A. Kane, R. L. Kane, and Ladd 1998; Lusky and Ingman 1994). Lack of uniformity from state to state in care delivery arises as states decide who is eligible for Medicaid. They also impose their own diverse standards for quality care and regulation of such care through inspections.

Alternative approaches to ensuring standards for services have been advanced through regulation or free market economy (Nyman and Geyer 1989). Regulation seeks to enforce high quality of care through staffing standards, care plans, and result audits. In general, regulation aims to ensure the most basic aspects of quality of care, such as appropriate medical care, sanitary living conditions, sufficient exercise, adequate diet, and at least limited privacy for residents. Regulatory efforts are widely employed and have had at least limited success in defining and monitoring quality of health care. However, regulation is seen by many as a costly and often ineffective approach that is dependent on enforcement of a limited set of universally agreed-upon standards (Nyman and Geyer 1989).

In free market competition, consumers do the work of raising quality by making informed choices and purchasing high-quality services. Ensuring standards through the use of the free market economy rests on the assumption that the consumer can identify good-quality care and has furthermore researched the market and will choose to reside in a facility that provides the best care. However, third-party payment systems limit the effectiveness of consumers in exercising market choices. Severely impaired nursing home patients (who may also lack family or advocates) are limited in their ability to exercise sufficient rational market choice to ensure nursing home quality.

Alternative approaches to improving quality of care have been noted in addition to those of regulation or competition. For example, increased involvement by volunteers, family members, and other representatives of the community may enhance care by increasing public awareness and

accountability. Community advocacy programs that encourage local citizens to press for patient rights and for improved care delivery have been found to be useful (Williams 1986). In addition, there has been increasing education of the public to have patients and families serve as their own advocates in choosing high-quality nursing homes and in seeking high-quality care through awareness of nursing home residents' rights (Burger et al. 1996). There has also been evidence of the useful roles played by nursing home ombudsmen in helping to resolve disputes involving residents (J. Kahana 1994).

In a 1986 report, the Institute of Medicine called for improvements in the quality of care in nursing homes. A landmark development following this report was the Omnibus Budget Reconciliation Act (OBRA) of 1987 setting forth guidelines by the Health Care Financing Administration for nursing home care. This act aimed to protect resident rights and improve residents' quality of life through a broad set of regulations. These included training and certification of nursing assistants, establishment of quality assurance committees, mandated resident assessments to allow for individualized care planning, reduction of physical and chemical restraints and preadmission screening for mental illness (Hamme 1991). OBRA also enforces a regulation that sets standards of care by increasing financial sanctions for noncompliance. Accordingly, noncompliant homes could be subject to fines and have Medicaid and Medicare payments withheld (Hamme 1991).

The enactment of OBRA raised high hopes about ensuring a high quality of care in nursing homes, and its implementation has resulted in improvements in quality of care. There is evidence of reduced use and overuse of drugs, in particular psychotropic medications (Borson and Doane 1997; Lantz et al. 1996), as well as a reduction in the use of physical restraints in nursing homes since the implementation of OBRA (Dunbar et al. 1996; Siegler et al. 1997). Research has also documented improvements in the accuracy of information in residents' medical records and in comprehensiveness of care plans (Hawes et al. 1997). There has also been an increase in positive programs such as presence of advanced directives and participation in activities by residents (Hawes et al. 1997). As an outgrowth of OBRA legislation, psychiatric assessments and screening programs have resulted in

some improvement in matching residents to appropriate mental health services (Borson and Doane 1997). Nevertheless, many problems and challenges have remained in the wake of OBRA. Thus, for example, the OBRA-mandated freedom of residents to select their own physicians can seldom be implemented, because most physicians will not follow their patients to the nursing home to deliver care. Lack of funding for inspections allows violations to go undetected (Day 1996). There have also been only limited advances made in the treatment of the depressed elderly (Snowden and Roy-Byrne 1998).

### **RECENT TRENDS IN PROGRAMS AND CARE DELIVERY IN NURSING HOMES**

In addition to mandated programs to enhance nursing home quality, there have also been notable developments in recent years to introduce innovative or at least new approaches to caring for older adults in nursing homes and other long-term care facilities. We will review two major approaches to such developments: special units and formal intervention programs to improve quality of care and of residential life.

**Special Units.** Special units within nursing homes are based on homogeneous groupings of people with special needs. The assumption underlying this movement is that specially trained staff may best meet the unique needs of specific patient groups. Segregating populations of residents with stigmatized or disruptive characteristics may also be seen as benefiting other residents by limiting their exposure to these populations.

Special programs can be targeted to unique groups such as dementia patients and patients in the final stages of life. Each of these will be discussed below. Although there is increasing recognition of the role of nursing homes in meeting the special needs of populations such as patients with AIDS (Zablosky and Ory 1995) or patients with brain injuries (Parsons 1997), programs for such individuals have been sporadic and limited. Often these special needs groups include higher proportions of younger patients and models of nursing home care geared to the elderly are poorly matched to their needs. Furthermore, research evaluating such programs is generally lacking.

*Dementia Care Units* Special care units in nursing homes have been proliferating for dementia patients. About 20 percent of nursing homes had dementia-specific special care units in the mid-1990s (Aronson 1994). These units are based on presumed benefits of functionally homogeneous resident groupings and generally target programs for confused but ambulatory residents.

It has been argued that dementia patients need “high-touch” rather than “high-tech” interventions (Aronson 1994). Many programs targeting dementia patients, are not reimbursable under skilled care guidelines. Special units for dementia care have been in the forefront of the development of certain innovative programs and particularly in integrating family members in patient care (Gaston 1994). Family members’ prior knowledge of the older adult when he or she was functioning well can enable expressions of emotional support, which in turn facilitates maintenance of the dignity and self-worth of these elders. Effective formal programs in dementia units that involve families in the care of elders range from family stories workshops (Hepburn et al. 1997) to the inclusion of families in staff-initiated treatment programs (Grower et al. 1994).

*Hospice Units* As greater numbers of older adults are dying in nursing home settings, the provision of hospice care is becoming a more central function of nursing homes. Between 1992 and 1995, the number of nursing homes with hospice units increased 100 percent, although the development of such units is constrained by market forces such as Medicaid reimbursement (Castle 1998). The value of the hospice as a useful model for end-of-life care has been increasingly recognized. This approach to care emphasizes holistic patient care aimed at diminishing pain and enhancing comfort and meaning. This orientation brings both dignity and autonomy to the process of dying for the elderly resident (Hayslip and Leon 1992).

The hospice model also points the way to recognition that the nursing home often serves as the context where older adults prepare for dying. Accordingly, a much-needed and often-absent function of nursing homes is helping residents come to terms with awareness and acceptance of their finitude (Johnson and Barer 1997).



**Formal Interventions to Improve Quality of Life in Nursing Homes.** While sociologists have generally focused on the specification of broad social features of nursing homes that impact on the life and welfare of the resident, psychologists, nurses, and social workers have been more involved in implementing specific and circumscribed intervention efforts in nursing homes. Direct interventions to improve resident functioning and/or the quality of resident life in nursing homes have been limited by the absence of systematic theories, on the one hand, and lack of resources, on the other. Diverse efforts to improve care have generally yielded some success, suggesting that almost any type of intervention can improve the quality of life or functioning of residents (E. Kahana, B. Kahana, and Chirayath 1999). Interventions may be broadly classified into two types: those aiming to improve the physical or social environment of the setting, and those aiming to improve the coping strategies, psychosocial well-being, or cognitive functioning of residents. Interventions may alternatively be directed at the resident, the environment, family, or staff.

Interventions to improve health and physical functioning of nursing home residents typically include programs that attempt to reduce impairment and disability (e.g. nutritional problems, urinary incontinence) or the risk of falling. Some interventions aimed at improving health and physical functioning expand their scope by attempting to improve overall quality of life for the resident. Such interventions include programs which grant residents greater autonomy and locus of control (Wagner et al. 1994), and those that incorporate self-image enhancements (Plautz and Timen 1992). Intergenerational programs that promote interaction between nursing home residents and children or young adults have been found to be successful in enhancing quality of life for nursing home patients as well as young children (Gaston 1994; Newman 1985). Such programs have generally been found to improve the activities and social interactions of the elderly.

Interventions targeting cognitively impaired residents include divergent approaches. Montessori methods have also shown success in improving the cognitive functioning of demented elders (Camp and Mattern 1999). Cueing and reality orientation have been therapeutic strategies aimed at reinforcing orientation to time, place, and person among

cognitively impaired elderly (Aronson 1994; Whanger 1980). Fantasy validation therapy takes a divergent view, providing staff acceptance of nursing home patients through expressing empathy in response to unrealistic beliefs or behaviors (Feil 1982).

Behavioristic approaches to the treatment of psychiatric problems in nursing homes include reinforcement of appropriate behaviors through token economies or habit training (Whanger 1980). Educational opportunities to enhance competent coping strategies have also been advocated (E. Kahana and B. Kahana 1983). Regarding the physical environment of the nursing home, prosthetically designed environments have been found to retard decline among mentally impaired elderly (Lawton et al. 1984). Milieu therapy is a systematic approach to enhancing all aspects of the social and physical environment in order to encourage social interactions among residents (Soth 1997).

As illustrated in the above discussion of systematic interventions in nursing homes, much of the empirical work relevant to person-environment transactions focuses on only one of the two related influences: the patient or the milieu. Thus, while conceptual frameworks recognize the dynamic nature of person-institution transactions, these complex interactions have not yet been incorporated in therapeutic interventions.

**Focus on the Life of the Nursing Home Resident: Effects of Institutionalization.** We will now consider those aspects of institutionalization that impact on the experiences, lifestyles, and well-being of the individuals who enter even the best of nursing homes. It is primarily this area of inquiry, relating social processes to the life experiences of the individual patient, to which sociologists have addressed their research on nursing homes. Considering the individual resident in the context of the physical and social milieu of a given institution, we can appreciate both the factors that induce negative reactions and those resources which facilitate positive responses to institutionalization. On this level of analysis, the sociologist moves away from considering the patient as a mere object of care and notes the interactive nature of the encounter between the institutionalized person and diverse elements within the nursing home environment.

The problems brought about by institutionalization go beyond problems of quality health care

and in fact, may be inherent in the very nature of congregate care. Accordingly, sociologists have recognized that the nursing home, by its very nature, represents a unique social context with homogenizing qualities, and that there are alterations in the normal patterns of interaction and social exchange, even in facilities providing high-quality care (Linz et al. 1993). Holistic analyses of life in the nursing home have been conducted primarily in a qualitative tradition. They range from firsthand accounts to in-depth interviews and participant observations of the nursing home (Diamond 1992; Gubrium 1993; Henderson and Vesperi 1995; Laird 1979; Savishinsky 1991; Shield 1988). There have been few, if any, quantitative research projects to address the complex fabric of residents' experiences of nursing home life in the past two decades. Conceptual developments have also been limited, highlighting the seminal nature of Goffman's original conceptualizations of the total institution as the most comprehensive and best model in our field.

Goffman's classical depiction of the total institution (1961) still serves as a standard for understanding the problems of institutional living. The *total institution* is described as a place where inmates are brought together under a common authority, are stripped of their normal identities, and are expected to engage in activities of daily living according to formal rules and a rational plan that regiments them. Activities of work, play, and sleep overlap and are typically conducted in the presence of others. The institution (which is often located at a distance from friends and from the previous community of the resident) also effectively cuts residents off from social ties in the outside world. It has been suggested that, among nursing home residents, isolation from society and loss of control over one's life lead to learned helplessness (Coons and Mace 1996; Baltes 1996; Baltes and Baltes 1986).

Frail older persons who typically enter nursing homes are particularly vulnerable because of physical and mental infirmities, sensory impairments, and loss of social supports that have created the need for such placement (Resnick et al. 1997; E. Kahana, B. Kahana, and Kinney 1990). Such vulnerable individuals are particularly sensitive to environmental change, and adverse living conditions (Lawton 1980). Elderly persons living

in the community generally fear institutional placement and seldom plan for a move to an institution (E. Kahana, B. Kahana, and Young 1985; Schoenberg and Coward 1997). The transition to living in a nursing home is typically involuntary, with patients seldom playing major roles in the relocation decision or choice of facility (Reinardy 1992). Furthermore, the new institutional setting is unfamiliar to the resident in terms of both physical features and social expectations. Fewer than half of family members visit facilities prior to placement of an elder (Lieberman and Kramer 1991). Unpredictability and uncontrollability are major risk factors accounting for the negative outcomes of institutional relocation (Schultz and Brenner 1977). Lack of involvement in decision making has also been associated with adjustment problems (Rubenstein et al. 1992). Alternatively, positive adjustment can be facilitated where older adults participate in decision-making about relocation (Armer 1993).

It is difficult to establish conclusively which elements of institutionalization are responsible for negative outcomes among residents because the effects of morbidity, relocation, and institutionalization occur concurrently and are difficult to separate (Lieberman and Kramer 1991). Yet the negative personal consequences of life in nursing homes have been documented. Institutionalization has been described in terms of dismantling one's home with its comforts, memories, and freedoms (Savishinsky 1991). Institutionalized elders have shown loss of self-esteem and identity (Tobin 1991), and often manifest withdrawal, apathy, and depression (Gubrium 1975, 1993; R. A. Kane 1995-1996; Vladeck 1980).

Gubrium (1993) conducted qualitative interviews with nursing home residents and found great disjunctures in meaning between the workings of the nursing home as an organization and the textured realities of the "lived experience" of residents. Residents live in a rich world of subjective meaning, which is seldom recognized and validated by the nursing home. In fact, nursing homes may actively resist responding to subjective realities of residents, in an effort to introduce an organizational rationality (Diamond 1992). It has been documented that institutional environments providing limited control over the daily lives of residents result in negative resident outcomes,

including diminished life satisfaction (R. L. Kane et al. 1990; Timko et al. 1993).

In spite of the evidence of adverse effects of institutionalization, there is a growing body of research and clinical observation that documents positive features and potential benefits of residential life in nursing homes (Patchner and Balgopal 1993; Pynoos and Regnier 1991). For isolated older persons who can no longer care for themselves, the nursing home can offer protection, improved living conditions, and even homelike qualities (Groger 1994). Advantages of living in a nursing home can also include behavioral expectations that are well matched to the competencies of the frail resident (Baltes and Werner 1992; Werner et al. 1994; Lawton 1980).

Nutritious meals, regular medical care, and supervised administration of medications may maximize health for frail elders living in nursing homes. Proximity to other residents may allow for social needs to be met, and organized activities can lead to meaningful social participation (Bitzan and Kruzich 1990). Indeed, there is some evidence that resident satisfaction subsequent to institutionalization exceeds expectations (E. Kahana, B. Kahana, and Young 1985). Some early studies have noted improved morale (Spasoff et al. 1978) and enhanced family relationships subsequent to institutionalization of elders (Smith and Bengtson 1979).

There are indications that elderly nursing home residents find their lives to be meaningful and that their sources of meaning do not differ significantly from those of community-living older adults (DePaola and Ebersole 1995). Well-being among residents subsequent to institutionalization may reflect not so much the positive influences of institutional life as the resilience and survival skills of residents. Accordingly, Lieberman and Tobin (1983) demonstrated that even in the face of major involuntary environmental changes such as institutionalization, many elderly persons continue to preserve a coherent and consistent self-image.

**Adaptation in Institutions.** There is growing evidence supporting the view that residents are active agents who attach meaning to and impact actively on their environment, continuing to take personal initiatives to remain socially engaged in the face of personal and environmental obstacles for doing so (E. Kahana, B. Kahana, and Chirayath 1999; Mor et al. 1995; Gubrium 1993).

Goffman (1961) described a range of adaptive responses among inmates of total institutions. *Withdrawal* refers to the resident's efforts to curtail interaction with others and to withhold emotional investment in his or her surroundings. *Intransigence* is a response that challenges institutional authority through noncooperation. These two modes of responding are likely to result in further alienation and to invite negative responses from staff and other residents. *Colonization* represents a strategy of maximizing satisfactions within the confines of the institution by accepting the rules and norms of institutional life. *Conversion* represents an identification with both the outward characteristics and the values of staff. Patients who opt for conversion submerge their identities into their patient roles. Although Goffman's conceptualization and description provide the earliest and possibly the richest sociological efforts to understand resident adaptation in nursing homes, there has been little follow-up research to confirm the typologies that he proposed.

Research has documented that there are active efforts even among frail institutionalized elderly to adapt to demands and stresses of institutional living and to remain socially engaged even while living in an institution (Mor et al. 1995). Instrumental coping strategies have been associated with maintenance of psychological well-being subsequent to institutionalization, whereas affective modes of coping have been related to decline in morale (E. Kahana, B. Kahana, and Young 1987).

Appraisals of life in a nursing home may contribute greatly to perceptions of stress, to coping responses, and ultimately to adaptive outcomes. In fact, an understanding of the interpretive meaning of institutional life may help integrate conflicting findings about effects of selection, relocation, and institutionalization on psychosocial well-being of elders residing in nursing homes (Gubrium 1993).

Aspects of physical frailty that create a need for institutional placement, along with perceived or real abandonment by family, require major reappraisals of both one's worldview and one's self-concept. Given a vulnerable self and loss of intimacy with significant others, the safety of one's physical and social milieu becomes a critical concern. To the extent that new residents of nursing homes can appraise their physical and social environment as safe, they will feel protected. To the

extent that they feel that other residents who compose their new reference group are helpful, social integration may be possible, and depersonalization may be avoided. There is evidence that residents in assisted living facilities engage in reciprocal helping social interaction and derive satisfaction from being providers of assistance to others (Litwin 1998). This research supports earlier work (E. Kahana, B. Kahana, Sterin, Fedirko, and Brittis 1990) which demonstrated that perceptions of even minor acts of helpfulness by other residents help the institutionalized elderly reinterpret their surroundings as benign. Formation of social ties with other patients represents an important mode of positive adaptation for nursing home residents (Mor et al. 1995).

Personal backgrounds of residents as well as environmental influences affect the nature of adaptations that residents make in institutions. Lack of mental impairment and few mobility limitations and sensory deficits are associated with maintenance of close social ties within the nursing home (Bitzan and Kruzich 1990). Personality and cognitive traits such as impulse control have been found to be associated with psychosocial well-being subsequent to institutionalization (B. Kahana and E. Kahana 1976).

In order to better understand and operationalize person-environment transactions in nursing homes, several conceptual models have been articulated that take into account both personal and environmental features. Lawton's ecological model (1980) focuses on the importance of matching environmental elements to personal competencies of frail elders. E. Kahana's person-environment congruence model (1982) emphasizes the role of individual differences in needs and environmental preferences, and specifies alternative formulations for expected outcomes based on oversupply, undersupply, or congruence of environmental characteristics, such as stimulation or homogeneity, in relation to personal preferences. It is notable that there has been very limited attention to advancing conceptual models in this arena during the past two decades.

Interactionist perspectives lead to a better understanding of social influences in nursing homes by calling attention to the importance of both personal reactions and environmental presses. Environmental design and intervention approaches

in nursing homes have focused on improving adjustment of residents to nursing homes by providing environments that benefit residents in general, or interventions that improve fit between the environment and personal needs of residents. Research has demonstrated that resident characteristics interact with environmental features of the nursing home to predict outcomes (Baltes et al. 1991; Timko et al. 1993). Accordingly, supportive physical features and assistive services were found to benefit impaired residents while policies permitting resident control were most likely to benefit independent residents. Furthermore, specification of salient dimensions of the institutional milieu has been one important area of progress toward designing better nursing home environments. Moos and Lemke (1996) have conducted pioneering work in providing reliable and valid indicators of social dimensions of the institutional environment.

Specific components of the institutional environment determine the demands, constraints, and benefits of institutional life for residents. They include the administrative structure, the physical environment, and the social environment. The social environment may be further subdivided into staff environment; patient environment; and community representatives such as volunteers, or friends and family, who visit the resident in the institution. There is growing evidence that the perspectives of residents, staff, and families diverge with regard to quality of life in nursing homes (Brennan et al. 1988). Research on administrative structure has focused primarily on size, financing, and type of ownership. Although it has been argued that proprietary ownership may result in poorer quality of care, the link between type of ownership and level of care has not been conclusively established. Similarly, suggested links between size of home or proportion of Medicaid patients and quality of care have not been fully documented (Shapiro and Tate 1995).

Similarities in ethnic, cultural, and social backgrounds of staff and residents appear to facilitate positive interactions, whereas discrepancies in cultural values have been found to hamper communication and mutual understanding (E. Kahana, B. Kahana, Sterin, Fedirko, and Taylor 1993; Harel 1987). Institutional norms as well as formal policies have been found to shape the impact of institutions on residents (Kiyak et al. 1978). Research

also suggested that perceived social support from family not only deters institutional placement but also relates to higher self-esteem and diminished depression among elderly nursing home residents (McFall and Miller 1992).

*Resident-to-Resident Interactions* There has been very little research focusing on the social milieu of nursing homes in terms of resident-to-resident interactions. It is indeed noteworthy that, in one study of nursing home residents, nearly 50 percent of all residents were found never to talk to their roommates, typically because of barriers to communication such as hearing or speech problems (Kovach and Robinson 1996). Nevertheless, among those residents able to talk to their roommates, rapport with the roommate was found to be a significant predictor of life satisfaction. It is notable that communication rules about talking often inhibit communication and contribute to living in silence among nursing home residents (Kaakinen 1992).

The salience of friends appears to be limited for nursing home residents, as there are many barriers to elderly friends maintaining contact with their noninstitutionalized friends. Accordingly, perceived social support from friends did not significantly relate to positive outcomes of nursing home residents (Commerford and Reznikoff 1996).

*Family-Resident Interactions* Interactions with family also play an important role in social integration of residents. Research has underscored that the majority of institutionalized elderly maintain meaningful ties and interactions with family members (Schwartz and Vogel 1990; High and Rowles 1995; Smith and Bengtson 1979). Ties to children who visit most frequently appear to be closest, followed by ties to other family members and friends (Bitzan and Kruzich 1990). Proximity of family members and previous history of extensive social interactions facilitate continued contact between residents and kin (York and Calsyn 1977). In turn, visitation by families and friends has been associated with enhanced residential functioning and well-being (Greene and Monahan 1982).

Broader roles for family members for being involved in the support and direct care of patients have been discussed, particularly in the social work literature (LaBrake 1996). However, systematic investigations of the efficacy of such efforts are

sparse. Some family support group activities as well as family counseling programs have reported success in bringing family and staff closer together (Campbell and Linc 1996). There have also been some educational efforts directed at assisting family members in coming to terms with psychological issues that they face regarding institutionalization of the elder relative (Drysdale et al. 1993).

Even while there is limited indication of systematic and welcoming programs initiated by the nursing home for the involvement of families in the care of residents, there are indications about naturally occurring involvement by family members in the lives and care of institutionalized older adults. Thus, for example, studies indicate that families are highly involved in assisting their relatives with decision making, and in actually making decisions when needed (High and Rowles 1995). Although researchers had expected to find less participation in decision making by families over time (since it was assumed that they would gradually defer to staff in making decisions), it was found that families do remain highly involved in their relatives' lives and well-being, even after four years. Families reported involvement in a broad range of decisions, ranging from those made during crises to those concerning the physical environment and treatment decisions.

There is also evidence, from surveys of family perspectives on nursing home care, that family members are typically sensitive and understanding about the constraints under which nursing home staff work and the difficulties posed by caring for frail elders, and particularly those with dementia (Looman et al. 1997). This research found that families also appreciate positive interpersonal ties between nursing assistants and residents.

The potential of constructive family involvement for improving resident life in nursing homes is yet to be recognized. It could have a major positive influence on helping break down barriers between the outside world and those of the institutions, addressing Goffman's classic challenge to the total institution (1961).

**Future Prospects for Policy and Long-Term Care Delivery.** It has been argued that, in the future, home care may evolve into a model of personal assistance services, and thus could blur the boundaries between institutional (nursing

home) and community-based care (R. A. Kane 1995). To the extent that personal care and housing-related services in nursing homes could be separated, residents and families could gain greater control of their lives. With the support of flexible home care options, many frail older persons could remain in innovative community residential options, such as assisted living arrangements, for the remainder of their lives. Such potential developments are attractive, but assume major changes in the financing of long-term care and the building of private accommodations, which include baths and kitchenettes, in housing for nursing home residents. Such new trends in financing and delivery of care would have to be based on dramatic expansion of personal long-term care insurance and/or availability of savings to finance long-term care. Alternatively, they could be implemented based on major new universal services furnished through social insurance programs (Kingson 1996). However, there are no clear indications that any of these financing options is likely to materialize in the near future. Furthermore it is likely that in the future most new and even existing long-term care services will be implemented by individual states. Increasing involvement in delivery of long-term care by states is likely to lead to increasing variability in the quality of care delivered (R. L. Kane et al. 1998).

In spite of expected increases in alternative long-term care facilities, the population of old-old adults residing in nursing homes is expected to increase dramatically, due largely to the ever-increasing age structure of the U.S. population. It is notable in this regard that research projecting future nursing home use demonstrates that better health in future cohorts of the old-old will only slightly decrease the proportion of time older adults will spend in nursing homes, or the proportion of this cohort who will enter nursing homes (Laditka 1998). Thus the nursing home as an organization is here to stay for the foreseeable future.

Furthermore, it is important to recognize that all the numerous alternatives proposed to address the shortcomings in the current system involve potential problems and tradeoffs (R. A. Kane et al. 1998). Ultimately, the planning, financing, delivery, and oversight of long-term care challenges our values and ingenuity as a society. Accordingly, sociologists, who have generally opted out of the

study of long-term care, are very much needed to get involved in this area, if we are to gain deeper understandings and develop systematic research based guidelines for improving services in this field.

## CONCLUSION

The foregoing discussion has highlighted a series of counterpoints in consideration of nursing home care on the macro and micro levels. On the macro level, distinctions between medical and social models of care, quality-of-care, and quality-of-life issues have been discussed. Societal needs for cost containment have been juxtaposed with the need to invest greater resources in long-term care to ensure provision of high-quality care. Regulation and free market competition have been presented as alternative strategies to improve standards of care. Uncertainty about service models, along with a great concern about costs of care, have resulted in a stalemate in the field of long-term care (Vladeck 1995).

On the micro level, we have noted evidence for depersonalizing aspects of institutional living, along with data about protective features and benefits of long-term care environments. Furthermore, residents of nursing homes have been described as frail and vulnerable on the one hand, and as adaptable and resilient on the other. The ultimate well-being of nursing home residents is seen as a function of the environment, of the person, or of transactions between the two. These dualities are useful to propel dialogue and to permit a thorough examination of nursing home care. At the same time they hold the danger of oversimplification of issues that may be approached from a unidimensional framework as proponents advocate one pole of the duality or the other. In fact, a sociological understanding of nursing homes underscores the complexity and multidimensionality of the social context and the social world of the nursing home. Thus, there is great benefit in attempting to integrate insights gained from both poles of the dualities discussed.

The nursing home resident of the future is likely to be ever more frail, especially if we succeed in developing more home-based alternatives to care. Hence, we cannot reject the medical model in favor of a social model of care, or focus exclusively on quality of life rather than on quality of

care. Instead, we need to complement concerns of high-quality health care with those of high standards for social care. Similarly, just as proprietary care is likely to remain a part of health care in the United States, so regulation is here to stay. Although prospective nursing home residents are likely to be ever more frail, they are also likely to be more highly educated and more conscious of their rights as consumers. Financing mechanisms that enhance the ability of the consumers of long-term care to exercise control over their lives can complement regulatory efforts to upgrade quality of care in nursing home settings. On the broadest societal level, decisions about both commitment of resources and development of creative alternatives to institution-based, long-term care are likely to shape the parameters and qualities of nursing homes of the future.

The experience of a given individual in being cared for in a nursing home must ultimately be understood in the context of the complex matrix of influences posed by institutional living. Accordingly, it is not fruitful to focus exclusively on either the ill effects or the benefits of institutionalization. Reviews of nursing homes continue to focus on normative understandings, generally highlighting poor quality of care in such facilities. As we have noted in this essay, empirical support for such negative conclusions is generally derived from qualitative research. Quantitative studies generally provide less support for expectations of decline and adverse patient reactions. In an effort to understand conflicting conclusions of different genres of research in this area, it is useful to focus on personal as well as environmental and situational influences that moderate the effects of institutional living. More carefully designed nursing home-based research, utilizing quantitative as well as qualitative approaches, is needed to specify conditions of both person and environment that maximize the well-being of the individual requiring institutional care.

Newer health care options, such as CCRCs, assisted living, and special care units, have been touted as more patient-responsive solutions to care than traditional nursing homes. Nevertheless, true consumer control and resident outcomes are still largely contingent on the resident's power to demand and advocate for appropriate care options. If an older adult becomes physically frail or

mentally impaired, such options are likely to diminish, or even evaporate. Consumers are once again at the mercy of bureaucratic decision making about their best interests.

Even advocates of alternative forms of long-term care acknowledge that nursing home care will continue to be needed and utilized by the increasing segment of old-old citizens. Sanctions, incentives, and intervention programs have all been shown to be beneficial, at least to a limited extent, in improving the quality of care and the quality of life in nursing homes.

Enhancing the quality of nursing home care creates a challenge for society to commit greater resources generated by currently productive citizens to the care of those who have made previous contributions. The resources society devotes to long-term care ultimately mirror the value placed by society on its frail or dependent citizens. Thus, a devaluing of older people is likely to result in a devaluing of institutions that care for them, along with a devaluing of the providers of their care. Conversely, more positive societal attitudes toward frail elders are likely to be translated into increasing involvement by high-caliber, trained professionals in the care of the institutionalized elderly. In addition, positive societal attitudes should bring family members and community representatives into closer contact with institutions, and should help to break down barriers between the nursing home and community living.

#### REFERENCES

- Agich, G. J. 1993 *Autonomy and Long-Term Care*. New York: Oxford University Press.
- Andersen, B. R. 1987 "What Makes Excellent Nursing Homes Different from Ordinary Nursing Homes?" *Danish Medical Bulletin, Special Supplement Series 5*:7-11.
- Armer, J. M. 1993 "Relocation and Adjustment in the PGC Morale Scale and the IRA." *Clinical Gerontologist 13*(3):77-81.
- Aronson, M. K. (ed.) 1994 *Reshaping Dementia Care: Practice and Policy in Long-Term Care*. Thousand Oaks, Calif.: Sage.
- Baltes, M. M. 1996 *Many Faces of Dependency in Old Age*. Cambridge, Eng. Cambridge University Press.
- , and P. B. Baltes 1986 *Psychology of Control and Aging*. Hillsdale, N.J.: Lawrence Erlbaum.

- Baltes, M. M., and W. H. Werner 1992 "Behavior System of Dependency in the Elderly: Interaction with the Social Environment." In M. G. Ory, R. P. Ables, and P. D. Lipman, eds., *Aging, Health, and Behavior*. Newbury Park, Calif.: Sage.
- , and R. Monika 1991 "Successful Aging in Long-Term Care Institutions." In S. K. Warner and M. P. Lawton, eds., *Annual Review of Gerontology and Geriatrics*. New York: Springer.
- Banaszak-Holl, J., and M. A. Hines 1996 "Factors Associated with Nursing Home Staff Turnover." *The Gerontologist* 36(4):512-517.
- Bellah, R. N. 1996 *Habits of the Heart: Individualism and Commitment in American Life: Updated Edition with a New Introduction*. Berkeley: University of California Press.
- Bitzan, J. E., and J. M. Kruzich 1990 "Interpersonal Relationships of Nursing Home Residents." *The Gerontologist* 30(3):385-390.
- Bonifazi, W. L. 1998 "Changing Population." *Contemporary Long-Term Care* 21(12):54-58.
- Borson, S., and K. Doane 1997 "The Impact of OBRA-87 on Psychotropic Drug Prescribing in Skilled Nursing Facilities." *Psychiatric Services* 48(10):1289-1296.
- Brennan, P. L., R. H. Moos, and S. Lemke 1988 "Preferences of Older Adults and Experts for Physical and Architectural Features of Group Living Facilities." *Gerontologist* 28:84-90.
- Brittis, S. 1996 "Sharing Destinies: Staff and Residents' Perspectives on Excellence in High Quality Nursing Homes in London, England and New York City, U.S.A." Unpublished doctoral dissertation, Case Western Reserve University.
- Bureau of Labor Statistics (BLS) 1995 *Special Tabulations Prepared for the Institute of Medicine Committee on the Adequacy of Nurse Staffing in Hospitals and Nursing Homes*. Washington, D.C.: U.S. Department of Labor.
- Burger, S. G., V. Fraser, S. Hunt, and B. Frank 1996 *Nursing Homes: Getting Good Care There*. San Luis Obispo, Calif.: Impact.
- Camp, C., and J. Mattern 1999 "Innovations in Managing Alzheimer's Disease." In *Innovations in Practice and Service Delivery Across the Lifespan*. New York: Oxford University Press.
- Campbell, J., and L. Linc 1996 "Support Groups for Visitors of Residents in Nursing Homes." *Journal of Gerontological Nursing* 22(2):30-35.
- Castle, N. G. 1998 "Innovations in Dying in the Nursing Home: The Impact of Market Characteristics." *Omega* 36(3):227-240.
- Charmaz, K., and D. A. Paterniti 1999 *Health, Illness, and Healing: Society, Social Context, and Self*. Los Angeles: Roxbury.
- Cockerham, W. C. 1998 *Medical Sociology*, 7th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Cohen, M. A. 1998 "Emerging Trends in the Finance and Delivery of Long-Term Care: Public and Private Opportunities and Challenges." *Gerontologist* 38(1):80-89.
- Cohen-Mansfield, J., M. S. Marx, S. Lipson, and P. Werner 1999 "Predictors of Mortality in Nursing Home Residents." *Journal of Clinical Epidemiology* 52(4):273-280.
- Commerford, M. C., and M. Reznikoff 1996 "Relationship of Religion and Perceived Social Support to Self-Esteem and Depression in Nursing Home Residents." *Journal of Psychology* 130(1):35-50.
- Coons, D., and N. Mace 1996 *Quality of Life in Long-Term Care*. New York: Hayworth.
- Cox, E. O., and R. J. Parsons 1994 *Empowerment-Oriented Social Work Practice with the Elderly*. Pacific Grove, Calif.: Brooks/Cole.
- Day, W. V. 1996 "There 'Oughta Be a Law' to Eliminate so Many Laws." *Nursing Homes Long-Term Care Management* 45(10):16-18.
- Densen, P. M. 1987 "The Elderly and the Health Care System: Another Perspective." *Milbank Memorial Fund Quarterly* 65:614-638.
- DePaola, S. J., and P. Ebersole 1995 "Meaning in Life Categories of Elderly Nursing Home Residents." *International Journal of Aging and Human Development* 40(3):227-236.
- Diamond, T. 1992 *Making Grey Gold: Narratives of Nursing Home Care*. Chicago: University of Chicago Press.
- Drysdale, A. E., C. F. Nelson, and N. M. Wineman 1993 "Families Need Help Too: Group Treatment for Families of Nursing Home Residents." *Clinical Nurse Specialist* 7(3):130-134.
- Dunbar, J. M., R. R. Neufeld, H. C. White, and L. S. Libow 1996 "Retrain, Don't Restrain: The Educational Intervention of the National Nursing Home Restraint Removal Project." *Gerontologist* 36(4):539-542.
- Etzioni, A. 1993 *The Spirit of Community*. New York: Touchstone.
- Evans, L. K. 1996 "Knowing the Patient: The Route to Individualized Care." *Journal of Gerontological Nursing* 3:13-19.
- Feil, N. 1982 *V/F Validation, the Feil Method: How to Help Disoriented Old-Old*. Cleveland, Ohio: Edward Feil.



- Fortinsky, R. H., and L. Raff 1995–1996 “Changing Role of Physicians in Nursing Homes.” *Generations* 19(4):30–35.
- Garahan, M. B., J. A. Waller, M. Houghton, W. A. Tisdale, and C. F. Runge 1992 “Hearing Loss Prevalence and Management in Nursing Home Residents.” *Journal of the American Geriatrics Society* 40:130–134.
- Gaston, P. 1994 “Families as an Integral Part of Dementia Care.” In M. K. Aronson, ed., *Reshaping Dementia Care: Practice and Policy in Long-Term Care*. Thousand Oaks, Calif.: Sage.
- Goffman, E. 1961 *Asylums: Essays on the Social Situation of Mental Patients*. New York: Anchor.
- Greene, V. L., and D. Monahan 1982 “The Impact of Visitation on Patient Well-Being in Nursing Homes.” *Gerontologist* 22(4):418–423.
- Groger, L. 1994 “Decision as Process: A Conceptual Model of Black Elders’ Nursing Home Placement.” *Journal of Aging Studies* 8(1):77–94.
- Grower, R., C. Wallace, G. Weinstein, K. Lazar, S. Leventer, and T. Martico-Greenfield 1994 “Approaches to Social Programming.” In M. K. Aronson, ed., *Reshaping Dementia Care: Practice and Policy in Long-Term Care*. Thousand Oaks, Calif.: Sage.
- Gubrium, J. 1975 *Living and Dying at Murray Manor*. New York: St. Martin’s Press.
- 1993 *Speaking of Life: Horizons of Meaning for Nursing Home Residents*. Hawthorne, N.Y.: Aldine de Gruyter.
- Hamme, J. M. 1991 “Federal Nursing Home Reform: An Overview.” In *The Long-term Care Handbook: Legal, Operational, and Financial Guideposts*. Washington: National Health Lawyers Association.
- Harel, Z. 1987 “Ethnicity and Aging: Implications for Service Organizations.” In C. H. Hayes, R. A. Kalish, and D. Guttman, eds., *European-American Elderly*. New York: Springer.
- Hawes, C., V. Mor, C. D. Phillips, E. Fries-Brant, J. N. Morris, E. Steel-Friedlob, A. M. Greene, and M. Nennstiel 1997 “OBRA-87 Nursing Home Regulations and Implementation of the Resident Assessment Instrument: Effects on Process Quality.” *Journal of the American Geriatrics Society* 45(8):977–985.
- Hayslip, B., and J. Leon 1992 *Hospice Care*. Newbury Park, Calif.: Sage.
- Henderson, J. N., and M. D. Vesperi 1995 *The Culture of Long-Term Care: Nursing Home Ethnography*. Westport, Conn.: Bergin and Garvey.
- Hepburn, K. W., W. Caron, M. Luptak, S. Ostwald, L. Grant, and J. M. Keenan 1997 “The Family Stories Workshop: Stories for Those Who Cannot Remember.” *Gerontologist* 37(6):827–832.
- High, D. M., and G. D. Rowles 1995 “Nursing Home Residents, Families, and Decision Making: Toward an Understanding of Progressive Surrogacy.” *Journal of Aging Studies* 9(2):101–117.
- Hing, E. 1987 “Use of Nursing Homes by the Elderly: Preliminary Data from the 1985 National Nursing Home Survey.” *NCHS Advancedata* 135:1–11.
- Institute of Medicine (IOM) 1986 *Improving the Quality of Care in Nursing Homes*. Washington, D.C.: National Academy Press.
- 1996 *Nursing Staff in Hospitals and Nursing Homes: Is it Adequate?* In G. S. Wunderlich, F. A. Sloan, and C. K. Davis, eds., Washington, D.C.: National Academy Press.
- Johnson, C. L., and B. M. Barer 1997 *Life Beyond 85 Years: The Aura of Survivorship*. New York: Springer.
- Kaakinen, J. W. 1992 “Living with Silence.” *Gerontologist* 32(2):258–264.
- Kahana, B., and E. Kahana 1976 “The Relationship of Impulse Control to Cognition and Adjustment Among Institutionalized Aged Women.” *Journal of Gerontology* 30(6):679–687.
- Kahana, E. 1973 “The Humane Treatment of Old People in Institutions.” *Gerontologist* 13:282–289.
- , J. Liang, and B. Felton 1980 “Alternative Models of Person-Environment Fit: Prediction of Morale in Three Homes for the Aged.” *Journal of Gerontology* 35:584–595.
- 1982 “A Congruence Model of Person-Environment Interactions.” In M. P. Lawton, P. G. Windley, and T. O. Byerts, eds., *Aging and the Environment: Theoretical Approaches*. New York: Springer.
- Kahana, E., and B. Kahana 1983 “Environmental Continuity, Discontinuity, and Futurity, and Adaptation of the Aged.” In G. Rowles and R. Ohta, eds., *Aging and Milieu: Environmental Perspectives on Growing Old*. New York: Academic.
- , and H. Chirayath 1999 “Innovations in Institutional Care from a Patient-Responsive Perspective.” In D. E. Biegel and A. Blum, eds., *Innovations in Practice and Service Delivery across the Lifespan*. New York: Oxford University Press.
- Kahana, E., B. Kahana, K. Kercher, and H. Chirayath 1999 “A Patient-Responsive Model of Hospital Care.” In J. J. Kronenfeld, ed., *Research in the Sociology of Health Care* 16:31–54.
- Kahana, E., B. Kahana, and J. M. Kinney 1990 “Coping among Vulnerable Elders.” In Z. Harel, P. Ehrlich,

- and R. Hubbard, eds., *Understanding and Servicing Vulnerable Aged*. New York: Springer.
- Kahana, E., B. Kahana, and K. Riley 1989 "Person-Environment Transactions Relevant to Control and Helplessness in Institutional Settings." In P. S. Fry, ed., *Psychological Perspectives of Helplessness*. Alberta, Canada: Elsevier Science.
- Kahana, E., B. Kahana, G. Sterin, T. Fedirko, and S. Brittit 1990 "Patterns of Mutual Assistance and Well-Being among Ethnic Nursing Home Residents." Paper presented at the meeting of the Society for Traumatic Stress, New Orleans.
- Kahana, E., B. Kahana, G. Sterin, T. Fedirko, and R. Taylor 1993 "Adaptation to Institutional Life among Polish, Jewish, and Western European Elderly." In C. Barresi and D. Stull, eds., *Ethnicity and Long-Term Care*. New York: Springer.
- Kahana, E., B. Kahana, and R. Young 1985 "Social Factors in Institutional Living." In W. Peterson and J. Quadagno, eds., *Social Bonds in Later Life: Aging and Interdependence*. Beverly Hills, Calif.: Sage.
- 1987 "Strategies of Coping and Post-Institutional Outcomes." *Research on Aging* 9:182-199.
- Kahana, J. 1994 "Reevaluating the Nursing Home Ombudsman's Role with a View toward Expanding the Concept of Dispute Resolution." *Journal of Dispute Resolution* 94:217-233.
- Kane, R. A. 1995 "Expanding the Home Care Concept: Blurring Distinctions among Home Care, Institutional Care, and Other Long-Term Care Services." *Milbank Quarterly* 73(2):161-186.
- 1995-1996 "Transforming Care Institutions for the Frail Elderly: Out of One Shall Be Many." *Generations* 4:62-68.
- , R. L. Kane, and R. C. Ladd 1998 *The Heart of Long-Term Care*. New York: Oxford University Press.
- , and K. B. Wilson 1993 *Assisted Living in the United States: A New Paradigm for Residential Care for Frail Older Persons?* Washington, D.C.: American Association for Retired Persons.
- Kane, R. L., J. G. Evans, and MacFadyen 1990 *Improving the Health of Older People: A World View*. Oxford, England: Oxford University Press.
- , and R. A. Kane 1978 "Care of the Aged: Old Problems in Need of New Solutions." *Science* 200(26):913-919.
- , —, and R. C. Ladd, W. N. Veazie 1998 "Variation in State Spending for Long-Term Care: Factors Associated with More Balanced Systems." *Journal of Health Politics, Policy, and Law* 23(2):363-390.
- , J. G. Ouslander, and I. B. Abrass 1994 *Essentials of Clinical Geriatrics*, 3rd ed. New York: McGraw-Hill.
- Kim, E., and B. Rovner 1996 "Psychiatric Care in the Nursing Home." In W. E. Reichman, and P. R. Katz, eds., *The Nursing Home as a Psychiatric Hospital*. New York: Oxford University Press.
- Kingson, E. R. 1996 "Ways of Thinking about the Long-Term Care of the Baby-Boom Cohorts." *Journal of Aging and Social Policy* 7(3-4):3-23.
- Kiyak, A., E. Kahana, and N. Lev 1978 "The Role of Informal Norms in Determining Institutional Totality in Homes for the Aged." *Long-Term Care and Health Administration Quarterly* 2(4):102-110.
- Kovach, S. S., and J. D. Robinson 1996 "The Roommate Relationship for the Elderly Nursing Home Resident." *Journal of Social and Personal Relationships* 13(4):627-634.
- LaBrake, T. 1996 *How to Get Families More Involved in the Nursing Home: Four Programs That Work and Why*. New York: Haworth.
- Laditka, S. B. 1998 "Modeling Lifetime Nursing Home Use under Assumptions of Better Health." *Journal of Gerontology: Social Sciences* 53B(4):S177-S187.
- Laird, C. 1979 *Limbo: A Memoir about Life in a Nursing Home by a Survivor*. Novato, Calif.: Chandler and Sharp.
- Lantz, M. S., V. Giambanco, and E. N. Buchalter 1996 "A Ten-Year Review of the Effect of OBRA-87 on Psychotropic Prescribing Practices in an Academic Nursing Home." *Psychiatric Services* 47(9):951-955.
- Lawton, M. P. 1980 *Environment and Aging*. Monterey, Calif.: Brooks/Cole.
- , M. Fulcomer, and M. H. Kleban 1984 "Architecture for the Mentally Impaired Elderly." *Environment and Behavior* 16(6):730-757.
- Leutz, W. N., J. A. Capitman, M. Macadam, and R. Abrahams 1992 *Care for Frail Elders: Developing Community Solutions*. Westport, Conn.: Greenwood.
- Levit, K. R., H. C. Lazenby, B. R. Braden, C. A. Cowan, P. A. McDonnell, L. Sivarajan, J. M. Stiller, D. K. Won, C. S. Donham, A. M. Long, and M. W. Stewart 1996 "National Health Expenditures, 1995." *Health Care Financing Review* 18(1):175-214.
- Lewis, M. A., S. Cretin, and R. L. Cane 1985 "The Natural History of Nursing Home Patients." *Gerontologist* 25(4):382-388.
- Linz, C. W., L. Fischer, and R. M. Arnold 1993 *The Erosion of Autonomy in Long-Term Care*. New York: Oxford University Press.
- Lieberman, M. A., and J. H. Kramer 1991 "Factors Affecting Decisions to Institutionalize Demented Elderly." *Gerontologist* 31(3):371-374.

- Lieberman, M. A., and S. S. Tobin 1983 *The Experience of Old Age: Stress, Coping and Survival*. New York: Basic Books.
- Litwin, H. 1998 "The Provision of Informal Support by Elderly People Residing in Assisted Living Facilities." *Gerontologist* 38(2):239-246.
- Longino, C. F. 1998 "Geographic Mobility and the Baby Boom." *Generations* 22(1):60-64.
- Looman, W. J., L. S. Noelker, D. Schur, C. J. Whitlatch, and K. Ejaz-Farida 1997 "Nursing Assistants Caring for Dementia Residents in Nursing Homes: The Family's Perspective on the High Quality of Care." *American Journal of Alzheimer's Disease* 12(5):221-226.
- Lusky, R. A., and S. R. Ingman 1994 "Medical Care in Residential Settings: The Nursing Home in Transition." In W. E. Folts and D. E. Yeatts, eds., *Housing and the Aging Population: Options for the New Century*. New York: Garland.
- Maslow, A. H. 1970 *Motivation and Personality*. New York: Harper and Row.
- McFall, S., and B. Miller 1992 "Caregiver Burden and the Continuum of Care: A Longitudinal Perspective." *Research on Aging* 14(3):376-398.
- Meyer, H. 1998 "The Bottom Line on Assisted Living." *Hospitals and Health Networks* (July 20).
- Moos, R. H., and S. Lemke 1996 *Evaluating Residential Facilities: The Multifacet Environmental Assessment Procedure*. Thousand Oaks, Calif.: Sage.
- Mor, V., Kenneth Branco, J. Fleishman, and C. Hawes 1995 "The Structure of Social Engagement among Nursing Home Residents." *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences* 50B(3):P1-P8.
- Mui, A. C., and D. Burnette 1994 "Long-Term Care Service Use by Frail Elders: Is Ethnicity a Factor?" *Gerontologist* 34(2):190-198.
- National Center for Health Statistics (NCHS) 1997 "An Overview of Nursing Homes and their Current Residents: Data from the 1995 National Nursing Home Survey." *Advance Data from Vital and Health Statistics*, no. 280. Bethesda, Md.: NCHS.
- Netting, F. E., and C. C. Wilson 1994 "CCRC Oversight: Implications for Public Regulations and Private Accreditation." *Journal of Applied Gerontology* 13:250-266.
- Newcomer, R., S. Preston, and S. S. Roderick 1995 "Assisted Living and Nursing Unit Use among Continuing Care Retirement Community Residents." *Research on Aging* 17(2):149-167.
- Newman, S. J. 1985 "Housing and Long-Term Care: The Suitability of the Elderly's Housing to the Provision of In-Home Services." *Gerontologist* 25(1):35-40.
- Nyman, J. A., and C. R. Geyer 1989 "Promoting the Quality of Life in Nursing Homes: Can Regulations Succeed?" *Journal of Health Politics, Policy and Law* 14(4):797-816.
- Parsons, Y. 1997 "No Shades of Gray." *Contemporary Long-Term Care* 20(9):42-47.
- Patchner, M. A., and P. R. Balgopal 1993 *Excellence in Nursing Homes*. New York: Springer.
- Pearce, B. W. 1998 *Senior Living Communities: Operations Management and Marketing for Assisted Living, Congregate, and Continuing Care Retirement Communities*. Baltimore, Md.: Johns Hopkins University Press.
- Plautz, R., and B. Timen 1992 "Positioning Can Make the Difference." *Nursing Homes* 41(1):30-33.
- Pynoos, J., and V. Regnier 1991 "Improving Residential Living Environments for the Frail Elderly: Bridging the Gaps between Theory and Application." In J. E. Birren, J. E. Lubber, J. C. Rowe, and D. E. Deutchman, eds., *The Concept and Measurement of Quality of Life in the Frail Elderly*. New York: Academic.
- Reinardy, J. R. 1992 "Decisional Control in Moving to a Nursing Home: Post-Admission Adjustment and Well-Being." *Gerontologist* 32(1):96-103.
- Resnick, H. E., B. E. Fries, and L. M. Verbrugge 1997 "Windows to Their World: The Effect of Sensory Impairments on Social Engagement and Activity Time in Nursing Home Residents." *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences* 52B(3):S135-S144.
- Rice, T. 1989 "The Use, Cost, and Economic Burden of Nursing-Home Care in 1985." *Medical Care* 27(12):1133-1147.
- Rovner, B. W., and I. R. Katz 1993 "Psychiatric Disorders in the Nursing Home: A Selective Review of Studies Related to Clinical Care." *International Journal of Geriatric Psychiatry* 8:75-87.
- Rubenstein, R. L., J. C. Kilbride, and S. Nagy 1992 *Elders Living Alone: Frailty and the Perception of Choice*. New York: Aldine.
- Rubenstein, L. Z., and D. Wieland 1993 *Improving Care in the Nursing Home: Comprehensive Reviews of Clinical Research*. Newbury Park, Calif.: Sage.
- Savishinsky, J. S. 1991 *The Ends of Time: Life and Work in a Nursing Home*. New York: Bergin and Garvey.
- Schnelle, J. F., M. P. McNeese, S. F. Simmons, M. E. Agnew, and V. C. Crooks 1993 "Managing Nurse Aides to Promote Quality of Care in the Nursing Home." In L. Z. Rubenstein and D. Wieland, eds., *Improving Care in the Nursing Home: Comprehensive Reviews of Clinical Research*. Newbury Park, Calif.: Sage.

- Schnelle, J. F., J. G. Ouslander, and P. A. Cruise 1997 "Policy Without Technology: A Barrier to Improving Nursing Home Care." *Gerontologist* 37(4):527-532.
- Schoenberg, N. E., and R. T. Coward 1997 "Attitudes about Entering a Nursing Home: Comparisons of Older Rural and Urban African-American Women." *Journal of Aging Studies* 11(1):27-47.
- Schultz, R., and G. Brenner 1977 "Relocation of the Aged: A Review and Theoretical Analysis." *Journal of Gerontology* 32:323-333.
- Schwartz, A. N., and M. E. Vogel 1990 "Nursing Home Staff and Resident's Families Role Expectations." *Gerontologist* 30(1):49-53.
- Shapiro, E., and R. B. Tate 1995 "Monitoring the Outcomes of Quality of Care in Nursing Homes Using Administrative Data." *Canadian Journal on Aging* 14(4):755-768.
- Shaughnessy, P. W. 1994 "Changing Institutional Long-Term Care to Improve Rural Health Care." In R. T. Coward, C. N. Bull, G. Kukulka, and J. M. Gallihier, eds., *Services for Rural Elders*. New York: Springer.
- Sherwood, S., H. S. Ruchlin, C. C. Sherwood, and S. A. Morris 1997 *Continuing Care Retirement Communities*. Baltimore, Md.: Johns Hopkins University Press.
- Shield, R. R. 1988 *Uneasy Endings: Daily Life in an American Nursing Home*. Ithaca, N.Y.: Cornell University Press.
- Siegler, E. L., E. Capezuti, G. Maislin, M. Baumgarten, L. Evans, and N. Strumpf 1997 "Effects of a Restraint Reduction Intervention and OBRA '87 Regulations on Psychoactive Drug Use in Nursing Homes." *Journal of the American Geriatrics Society* 45(7):791-796.
- Smith, K., and V. Bengston 1979 "Positive Consequences of Institutionalizations: Solidarity between Elderly Parents and Their Middle-Aged Children." *Gerontologist* 19(5):438-447.
- Snowden, M., and P. Roy-Byrne 1998 "Mental Illness and Nursing Home Reform: OBRA-87 Ten Years Later." *Psychiatric Services* 49(2):229-233.
- Soth, N. B. 1997 *Informed Treatment: Milieu Management in Psychiatric Hospitals and Residential Treatment Centers*. Lanham, Md.: Medical Library Association/Scarecrow.
- Spasoff, R. A., A. S. Kraus, E. J. Beattie, D. E. Holden, and J. S. Lawson 1978 "A Longitudinal Study of Elderly Residents of Long-Stay Institutions." *Gerontologist* 18(3):281-291.
- Strahan, G. W. 1997 "An Overview of Nursing Homes and Their Current Residents: Data from the 1995 National Nursing Home Survey." *Advance Data from National Center for Health Statistics, Number 280*. Atlanta: Centers for Disease Control and Prevention.
- Timko, C., A. T. Q. Nguyen, W. Williford, and R. H. Moos 1993 "Quality of Care and Outcomes of Chronic Mentally Ill Patients in Hospitals and Nursing Homes." *Hospital and Community Psychiatry* 44(3):241-246.
- Tobin, S. S. 1991 *Personhood in Advanced Old Age: Implications for Practice*. New York: Springer.
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States*, 118th ed. Washington, D.C.: National Technical Information Services.
- Vladeck, B. C. 1980 *Unloving Care—The Nursing Home Tragedy*. New York: Basic.
- 1995 "Long-Term Care: The View from the Health Care Financing Administration." In J. M. Wiener, S. B. Clauser, and D. L. Kennell, eds., *Persons with Disabilities: Issues in Health Care Financing and Service Delivery*. Washington, D.C.: Brookings Institution.
- Wagner, L., V. Wahlberg, and A. M. Worning 1994 "Drug Consumption among Elderly—A Four-Year Study." *Scandinavian Journal of Caring Sciences* 8(2):113-117.
- Wallace, S. P., L. Levy-Storms, R. Kington, and R. M. Andersen 1998 "The Persistence of Race and Ethnicity in the Use of Long-Term Care." *Journal of Gerontology: Social Sciences* 53B(2):S104-S112.
- Werner, P., V. Koroknay, J. Braun, and J. Cohen-Mansfield 1994 "Individualized Care Alternatives Used in the Process of Removing Physical Restraints in the Nursing Home." *Journal of American Geriatrics Society* 42:3.
- Whanger, A. D. 1980 "Treatment within the Institutions." In *Handbook of Geriatric Psychiatry*. New York: Van Nostrand Reinhold.
- Williams, C. 1986 "Improving Care in Nursing Homes Using Community Advocacy." *Social Science and Medicine* 23(12):1297-1303.
- York, J., and R. Calsyn 1977 "Family Involvement in Nursing Homes." *Gerontologist* 17(6):500-505.
- Zablotsky, D. L., and M. G. Ory 1995 "Fulfilling the Potential: Modifying the Current Long-Term Care System to Meet the Needs of Persons with AIDS." *Research in the Sociology of Health Care* 12:313-328.

EVA KAHANA

## LONGITUDINAL RESEARCH

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

According to Heckman and Singer, "Longitudinal data are widely and uncritically regarded as a panacea . . . The conventional wisdom in social science equates 'longitudinal' with 'good' and discussion of the issue rarely rises above that level" (1985, p. ix).

There is probably no methodological maxim in sociology more often repeated than the call for longitudinal data. From the work of David Hume more than 250 years ago, to the exhortations for a "radical reformation" in the work of Stanley Lieberman (1985, p. xiii), the importance of longitudinal data has been emphasized and reemphasized. Yet it is doubtful that there is an area of sociological method in which more disagreement exists both as to rationale and as to method. Until relatively recently, it has been possible to ignore the problem because longitudinal data have been relatively rare, and methods for their analysis quite sparse. Since 1965, however, there has been a virtual flood of new longitudinal data and a concomitant increase in sophisticated methods of analysis. In large part the computer has made both developments possible, permitting the management of data sets of enormous complexity along with analyses undreamed of only a short while ago.

At the micro level, numerous longitudinal studies have tracked individuals over a good share of their lifetimes. For example, some participants in the Oakland and Berkeley studies of growth and development have been sporadically studied from birth until well into their seventies (Elder, this volume). On a more systematic basis, the Panel Study of Income Dynamics (PSID) has interviewed a panel based on an original sample of five thousand families (households) on an annual basis since the mid 1960s, supplementing the sample with new households formed as split-offs from the original families (Duncan and Morgan 1985). Many other large-scale panel studies, some extending over periods of thirty years and longer, are in progress (Migdal et al. 1981).

At the macro level, extended time series on various social and economic indicators such as GNP, fertility, mortality, and education are gradually becoming available in machine-readable form from virtually all industrialized societies and from many that are less developed. In some cases, data series, particularly for vital statistics, go back for decades. In other cases, such as China and the

Soviet Union, modern-era data are gradually being accumulated and linked to earlier series. Descriptions of many such data sets can be found in the annual guide published by the Inter-University Consortium for Political and Social Research (ICPSR 1991).

Perhaps the most exciting developments are at the nexus of macro- and micro-level analysis. In the United States, for example, the General Social Survey (GSS) has obtained data on repeated cross-sectional samples of the population (that is, the population is followed longitudinally but specific individuals are not) on an annual basis (with two exceptions) since 1972. More recently, annual surveys modeled on the GSS have been started in a number of other countries (Smith 1986). Because of the careful replication, these surveys permit one to track aggregate responses of the changing population on a wide variety of issues (such as on attitudes toward abortion or capital punishment) over time. As the time series becomes ever longer, it is possible to carry out a multilevel analysis, linking micro and macro variables. For example, using the GSS, DiPrete and Grusky (1990) attempt to link short-term changes in rates of social mobility to macro-level changes in the U.S. economy.

Although the size and complexity of the longitudinal data base has expanded rapidly, so have the statistical tools with which to analyze it. Perhaps the most exciting development is in the area of "event history models," which permit one to relate an underlying rate of transition in continuous time to a set of "covariates" or independent variables. However, event models are by no means the only development. New and powerful approaches to the analysis of means over time, to structural equation models, and to various forms of categorical data analysis have given the researchers unprecedented power. Standard computer packages now make routine what was impossible in the 1980s.

#### **THE RATIONALE FOR LONGITUDINAL RESEARCH**

Longitudinal studies are carried out in virtually every area of the social sciences. Although studies of infant development and national development share certain similarities, it is not likely that a single rationale, design, or approach to analysis will simultaneously apply to every area in which longitudinal data might be collected. At the most

abstract level, there are three basic reasons for conducting a longitudinal study.

First, in any area in which development and change over time are at issue, there is, almost by definition, a necessity to obtain time-ordered data. Questions pertaining to rate and sequence of change, and to variability in rates and sequences are at the heart of longitudinal research. At one level, these questions are essentially descriptive, and getting an adequate descriptive handle on time-ordered data is an essential first step in coming to any understanding of change.

A second reason involves the role of temporal priority in causal analysis. There are few things on which philosophers of science agree when it comes to causation (see Marini and Singer 1988 for a superb review), but one is that A must precede B in time if A is to be taken as a cause of B. It is natural to assume that observing A before B means that A precedes B. Unfortunately, designs that actually allow one to establish temporal, let alone causal, priority are not as easily arrived at as one might think.

Related to the issue of temporal priority is the cross-sectional fallacy. Lieberman (1985, pp. 179–183) argues that assertions of causality based on cross-sectional data must necessarily imply a longitudinal relationship. To show that city size “leads to crime” based on cross-sectional data implies a dynamic relationship that may or may not be true. The problem is particularly acute in cross-sectional age comparisons that are necessarily confounded with cohort differences. All cross-sectional attempts to ascertain the “effect” of age on income confound cohort differences in average income with age differences.

A third reason, particularly relevant to sociologists, is the necessity to distinguish gross change from net change. A census taken at two points in time shows changes in the distribution of a variable, say occupation, at the macro level. We might find that the proportion of the population in service-oriented occupations increases over the period of a decade. That indicator of net change conceals myriad patterns of gross change at the individual level. The count of persons in service occupations at two points in time consists of persons who were occupationally stable over the interval and persons who changed in various ways. Of course the population itself changed over the interval due to age-related changes in the labor

force, migration, and differing levels of labor-force participation. All of this is masked by repeated cross-sectional samples.

Finally, although not really a “rationale” for longitudinal research, observation plans with repeated measures on the same individuals can offer certain statistical advantages. Cook and Ware (1983) discuss these in detail.

### TYPES OF LONGITUDINAL DATA

For many sociologists, the term *longitudinal* connotes a particular design, usually referred to as a panel study, in which individual subjects are measured repeatedly over time. A more general definition is desirable. Longitudinal data consist of information that can be ordered in time. Such data can be obtained in a variety of ways: by measuring the subject prospectively at repeated intervals, by obtaining a retrospective history in one or more interviews, from institutional records, or various combinations of these approaches. “Strong” longitudinal data preserve exact time intervals, while “weak” data provide sequence and order but lose interval. The distinction is parallel to that between interval and ordinal measurement.

As Featherman (1977) notes, under some circumstances, retrospective data collection may have substantial advantages of time and cost. Most surveys collect retrospective data of one kind or another, such as educational attainment, family background, and marital histories. Using structured interviewing methods in which the respondent is provided with a time-oriented matrix, it is possible to collect quite accurate information on many aspects of the life course. For example, as part of an ongoing panel, Freedman et al. (1988) obtained retrospective reports of family structure and other demographic variables that had previously been measured contemporaneously. The results are highly encouraging; when retrospective questions of this kind are asked carefully and interviewers are well trained, respondents can provide accurate and detailed information.

Of course not all variables can be measured retrospectively. Most researchers would argue that any attempt to measure past psychological states is invalid on its face. Reporting of the timing and frequency of events, even those which are quite salient, such as hospitalization, appears to suffer

from serious recall problems. Subjects tend to forget them or to telescope them in time. Reports of exact earnings, hours worked, and other economic variables may suffer from similar problems. The truth is that we don't have much information on what can and cannot be measured retrospectively. A systematic research program on these issues is becoming more and more necessary as new methods of data analysis that depend on the exact timing of events continue to evolve.

Another serious weakness of the retrospective design is that it represents the population that survives to the point of data collection and not the original population. In some situations this kind of selection bias can be quite serious—for example, in intervention studies that are subject to high levels of attrition.

Prospective studies in which a subject is measured at some baseline and then at repeated intervals are usually referred to as panel studies. Panel designs have a number of strengths along with several significant weaknesses. The primary strength, at least in principle, is accuracy of measurement and correct temporal referents. Depending on the exact design of data collection, subjects are measured at a point close enough in time to the event or status in question to permit reliable and valid assessment. Under certain circumstances temporal priority may be more firmly established.

Second, the prospective design provides greater leverage on attrition. Besides measuring a population defined at baseline, preattrition information can be used to determine the extent to which attrition is "random" and perhaps can be used to correct estimates for selection bias. There is a trade-off, however. Frequent measurement has two potentially undesirable side effects. First, subjects may tire of repeated requests for cooperation and drop out of the study. Second, "panel conditioning" may result in stereotypic responses to repeated questions. Thus, relative to the retrospective design, there may actually be a net *decrease* in data quality.

On the surface, prospective designs that extend in time are far more costly than retrospective designs. There is a clear cost/quality trade-off, however, that cannot be easily evaluated without consideration of the purposes of the survey. In obtaining population estimates over time, the panel may actually be less expensive to maintain than

resampling the population repeatedly. On the other hand, using a panel for this purpose brings problems of its own in the form of attrition and panel conditioning.

### QUASI-EXPERIMENTAL AND DESCRIPTIVE APPROACHES

The large-scale, multiwave surveys so common now have rather diffuse origins. Paul Lazarsfeld introduced the panel design (Lazarsfeld and Fiske 1938). In his hands, the panel study was basically a *quasi-experimental design*. A panel of subjects was recruited and measured repeatedly, with the foreknowledge that a particular event would occur at a particular time. The most famous application is to election campaigns. As such, the design is a simple example of an interrupted time series.

A second source of current designs is the child development tradition. Baltes and Nesselroade (1979) cite examples going back to the late eighteenth century, but systematic studies date to the 1920s (Wall and Williams 1970). The best of these studies emphasized cohort-based sampling of newborns and systematic assessment of development at carefully determined intervals. In the tradition of experimental psychology, investigators paid attention to careful control of the measurement process, including the physical environment, the raters and observers, and the measurement instruments. The development of age-specific norms in the form of averages and variation about them was seen as a primary goal of the study. Unanticipated events, such as an illness of either mother or child or factors affecting the family, were seen as undesirable threats to the validity of measurement rather than as opportunities to assess quasi-experimental effects.

Large-scale multiwave panel studies of the kind described above combine aspects of both traditions, often in somewhat inchoate and potentially conflicting ways. On the one hand, investigators are interested in describing a population as it evolves. Often basic descriptive information on rates, variability, and sequence is unknown, calling for frequent measurement at fixed intervals. On the other hand, there is also interest in evaluating the impact of specific transitions and events, such as childbearing, retirement, and loss of spouse. Meeting these two objectives within the constraints of a single design is often difficult.

## DESIGN ISSUES

Although it might be argued that the ideal longitudinal study should take a huge sample of the population without age restriction, measure it constantly, and follow it in perpetuity with systematic supplements to the original sample, cost and logistics intervene. Although there is an infinite range of potential designs, longitudinal studies can be classified on various dimensions including (a) the consistency of the sample over time, (b) population coverage, particularly with regard to age, and (c) measurement protocols, including not only choice of variables but also timing, interval, and frequency. These factors will influence the extent to which the study can be used for descriptive purposes, relating information to a well-defined population, and/or drawing causal inferences on temporal and quasi-experimental aspects of the design.

**Consistency of the Sample.** The following classification is based on Duncan and Kalton (1987) and on Menard (1991).

1. *Repeated Cross-Sectional Surveys.* A new sample is drawn at each measurement point. The GSS, described above, is an example. This is a longitudinal study at the macro level. It describes a dynamic population.
  2. *Fixed-Sample Panel.* A sample is drawn at baseline and measured repeatedly over time. No new subjects enter the sample after baseline. Several examples are described above. The sample refers only to the cohorts from which it was drawn. It may continue to represent them adequately if panel attrition is completely at random.
  3. *Dynamic Sample Panel.* After baseline, subjects are added to the panel in an attempt to compensate for panel attrition and represent changes in the underlying population. The PSID is a major example.
  4. *Rotating Panels.* A sample is drawn and interviewed for a fixed number of waves and then dropped. At each measurement point a new sample is also drawn so that samples enter and leave on a staggered basis. The best-known example is the Current Population Survey carried out by the U.S. Bureau of the Census. At any given time, the sample consists of subjects who have been in the panel from one to four months.
  5. *Split Panels.* In addition to a basic panel survey, a coordinated cross-sectional survey is drawn at each measurement point. In effect, this is a quasi-experimental design in which comparisons between samples permit tests of panel conditioning, among other things. This design is rare.
- Population Definition.** The broader the population sampled, the wider the potential generalization. On the other hand, homogeneity provides greater leverage for some kinds of causal inference. The following rough classification is useful. See Campbell (1988) for elaboration.
1. *Unrestricted Age Range.* A sample of the entire (adult) population is selected and followed.
  2. *Restricted Age Range.* A sample of persons in a narrow age band, such as adolescents in developing nations, is selected, with resulting homogeneity of developmental process.
  3. *Event-Based.* A sample is selected on the basis of a particular event. Birth is the prime example; others are motherhood, school completion, business start-up, and administrative reorganization. Subjects can be members of a cohort experiencing the event at a fixed time or can be drawn on a floating baseline, as is the case when each new patient in a given clinic is followed prospectively.
  4. *Population at Risk.* A sample is selected on the likelihood (but not the certainty) that it will experience a particular event. Although similar to an event-based sample, it is less controlled. Age-restricted samples are usually at risk for certain events, which is one reason for restricting age in the first place. An interesting example at the macro level is a sample of cities likely to experience a disaster such as an earthquake or a hurricane.
- Measurement Protocols.** What variables should one measure with what frequency at what time



intervals? Answering such a question requires a clear appreciation of the linkage between the substantive purpose of an investigation and the mode of data analysis. For example, if one's intent is to study labor-force participation using event history models, then frequency of measurement would be dictated primarily by (a) the frequency of change of labor-force status among subjects, (b) the amount of variability in individual rates of change, (c) the necessity to obtain exact dates of transitions for the analysis, and (d) the maximum time interval at which subjects can provide reliable and valid recall data. If the investigators have explanatory models in mind, then the measurement protocol will have to deal with similar issues for the regressors.

If one is, however, interested in the effects of widowhood, which might be studied either descriptively or as a quasi-experiment using an interrupted time series approach, very different measurement strategies would be optimal. At the very least, one would try to schedule measurements at fixed intervals *preceding and following* the event. If economic effects were the primary focus, annual measurement might be sufficient; but if the grief process was the focus, more frequent measurement would be required.

The more undifferentiated the sample, and the more multipurpose the study, the more difficult it is to achieve an effective measurement strategy. Many of the large-scale longitudinal studies carried out since the 1960s have one or more of the following characteristics that tend to interfere with analysis:

1. Multiple substantive foci that result in attempts to measure scores, if not hundreds, of variables.
2. Nonoptimal measurement intervals because of conflicting demands of different kinds of variables and topics. A secondary problem is that intervals are often determined by administrative and funding criteria rather than by substantive intent.
3. Measurement strategies that are chosen without regard to statistical analysis. The problem is acute for event history models

that require dated transitions rather than reports of status at fixed intervals. Other examples are failure to acquire multiple indicators of constructs for LISREL models and intersubject variation in measurement intervals that interferes with growth curve estimation.

4. Weak identification of temporal sequence. This problem is most often a result of a "snapshot" orientation to measurement in which the subject's status is ascertained at a sequence of fixed measurement points. Knowing that at time 1 someone is married and prosperous and that at time 2 that person is single and poverty-stricken doesn't tell us much about causal order or effect.

#### **CAUSAL INFERENCE AND LONGITUDINAL DESIGN**

As noted, the extent to which longitudinal studies allow one to establish causal effects is the subject of some controversy. Indeed, the whole issue of causal inference from nonexperimental data is extremely controversial (Berk 1988; Freedman 1991), with some taking the position that it is impossible in any circumstance other than a randomized design involving an experimentally manipulated stimulus. Although that may be an extreme position, the assumption that one can use time-ordered observational data to establish causal order is difficult to defend.

Cross-sectional analyses and retrospective designs suffer from the fact that variables which are supposed to have time-ordered causal effects are measured simultaneously. As a result, various competing explanations of observed associations, particularly contaminated measurement, cannot be ruled out. Asking about educational aspirations after completion of schooling is an example. Panel designs at least have the advantage of measuring presumed causes prior to their effects. Even in that case, establishing temporal order is not as easy as it might appear. This is particularly true when one attempts to relate intentions, attitudes, and similar

cognitive variables to behaviors. Marini and Singer (1988) give an example based on education and marriage. Consider two women who leave school and get married. One may have decided to terminate her education in light of a planned marriage, and the other may have decided to marry following a decision to terminate education. The observed sequence of events tells us nothing. They note:

*Because human beings can anticipate and plan for the future, much human behavior follows from goals, intentions and motives, i.e., it is teleologically determined. As a result, causal priority is established in the mind in a way that is not reflected in the temporal sequences of behavior or even in the temporal sequence of the formation of behavioral intentions. (Marini and Singer 1988, p. 377)*

Because of the many varieties of longitudinal research design and the many controversies in the field, it is difficult to give hard-and-fast rules about when causal inference may be justified. Dwyer (1991), writing for epidemiologists, provides a useful way to approach the problem based on whether there is variance in presumed causes and effects at baseline.

Variation in Independent Variable

		NO	YES
Variation in Dependent Variable	NO	I	II
	YES	III	IV

Examples include the following:

- I. Follow crime-free baseline sample through possible arrest, incarceration, and later recidivism.
- II. Relate aspirations of eighth graders to eventual level of educational attainment.
- III. Carry out an experimental intervention, such as a job training program, that attempts to increase existing skill levels.
- IV. Relate organizational characteristics at baseline to later levels of productivity.

Cases I and III can be observational or (quasi-) experimental. Cases II and IV are strictly observational. In each case, although variation may not

exist in the variable of direct interest, there may be variation in closely related variables. In the experimental case, this corresponds to lack of randomization and, thus, uncontrolled variation in potential causal variables. The same is true in purely observational studies. In case II, although there is no variation in the direct outcome of interest, there is often variation in related variables at baseline. In the example given, although there is no variation in educational attainment in terms of years of schooling completed, aspirations are certainly based in part on the child's prior academic success in the grade school environment. Case IV presents the most difficult situation because it picks up data in the middle of an ongoing causal sequence. This is precisely the situation where many researchers believe that panel data can untangle causal sequence, but neither temporal-sequenced observations nor complex analysis is likely to do so.

To reiterate an important point about design, in large-scale longitudinal research each of these four cases is typically embedded in a complex design, and the degree of causal inference depends on many factors ranging from the timing of measurement to attrition processes.

**COMMON PROBLEMS IN LONGITUDINAL DATA ANALYSIS**

Those who collect and analyze longitudinal data face certain generic problems. Virtually any researcher will face some or all of the following difficulties.

**Conceptualizing Change.** James Coleman notes that the concept of change is "a second order abstraction . . . based on a comparison, or difference, between two sense impressions, and, simultaneously, a comparison of the times at which the impressions occurred" (1968, pp. 428-429). It is particularly difficult to think about the *causes* of change, and this difficulty has been reflected in arguments about how to model it. In particular, there has been a running debate about the use of change scores, computed as a simple difference ( $\Delta Y = Y_2 - Y_1$ ) versus a regression approach in which

the time 2 variable is regressed on time 1 plus other variables. In an influential paper, Bohrnstedt (1970) showed that simple gain scores tended to have very low reliability relative to either variable composing them. In light of that and other problems, Cronbach and Furby (1969) argued that the best way to model change was to treat "residualized gain scores" as dependent variables in regression analysis. The basic equation is

$$Y_2 = a + \beta_1 Y_1 + \beta_2 X_1 + \dots + \beta_k X_k + \varepsilon$$

where  $Y_1$  is the baseline measure and the  $X$ 's are any set of independent variables. Hence the effect of  $X$  is net of the baseline score. This method has become standard in many fields of inquiry.

More recently, a number of papers have appeared that question the use of residualized gain scores. Liker, Augustyniak, and Duncan (1985) argue that equations in which one takes differences on both sides of the model ( $\Delta Y = \alpha + \beta(\Delta X) + \varepsilon$ ) have strong advantages, particularly when one wants to "difference out" unchanging characteristics of subjects. Allison (1990) argues that in some cases the difference score as a dependent variable is not only acceptable but necessary. The issue is not purely statistical; it depends in large part on exactly what process one is attempting to model. Suffice it to say here that an issue which was once thought to be resolved has been reopened for serious examination.

Related to the issue of change scores and their reliability is the problem of regression toward the mean. Whenever a variable is positively correlated over time, subjects with high scores at time 1 will tend to have somewhat lower scores at time 2, while those with lower time 1 scores will tend to have higher scores at time 2. As a result, gain scores will be negatively correlated with initial values. This problem is exacerbated by random measurement error and was a primary reason for the predominance of residualized gain models in the past. Again, however, the issue is one of the underlying model. There are cases where feedback processes do indeed result in regression to the mean (Dwyer 1991) and the regression is by no means an "artifact." There is no question that one

always needs to correct for measurement error, but regression to the mean should not necessarily be treated as an annoyance to be gotten rid of by statistical manipulation.

**Lack of Independence.** A standard assumption in elementary statistical models such as least squares regression is that errors are independent across observations. In the equation we typically assume  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ . If the same subject is observed at two or more points in time, the independence assumption is almost certain to be violated due to omitted variables, correlated errors of measurement, or other problems. This difficulty permeates virtually all statistical approaches to longitudinal data and requires complex methods of estimation.

**Articulating Analysis with Observation Plan.** Ideally, one should have in mind a model for how the underlying process of change in a variable articulates with the observation plan. In part, this requires some knowledge of causal lag or how long it takes for changes in  $X$  to result in changes in  $Y$ . The data analysis method chosen should be congruent with both the underlying process of change and the observation plan. Observation plans may obtain exact dates of transitions (when did you quit your job?), locate the subject's state at time  $t$  (are you currently working?), or obtain retrospective information relative to some other time point (were you working as of June 30?). Studies that have multiple goals are difficult to optimize in this respect.

**Construct Validity.** In long-term longitudinal studies, the validity of constructs may degrade over time, sometimes through changes in the subjects themselves, sometimes through measurement problems such as conditioning, and sometimes through changes in the environment. For example, the meaning of terms used in political research such as *liberal* may change with time, or references to specific issues such as abortion may become more or less salient. In growth studies, subjects may "age out" of certain variables. This issue is quite understudied, at least by sociologists.

**Measurement Error.** Measurement error is a serious problem for statistical analysis, particularly regression, because it results in biased estimates.

The problem is particularly acute when independent variables are subject to error. In the longitudinal case, errors of measurement tend to be correlated across subjects over time, causing even greater difficulty. A major reason for the popularity of structural equation packages like LISREL (Jöreskog and Sörbom 1988; Bollen 1989) is that they permit one to model correlated error structures. This is not to say that LISREL is an all-purpose solution; indeed, its error-handling capabilities probably lead to its use in situations where other approaches would be better.

**Time-varying Independent Variables.** In a typical situation one is faced with a set of fixed independent variables such as age, sex, and race, and a set of variables whose values may change over the course of a study. In studies of income, for example, educational levels and family structure may change between observations. Although work on this problem has been going on in economics for some years (Hsiao 1986), sociologists have been slow to respond. The problem always leads to great statistical and computational complexity.

**Missing Data, Attrition, Censoring, and Selection Bias.** Attrition in panel studies is inevitable. In rare cases attrition is completely random, but more commonly it is associated with other variables. Standard listwise and pairwise missing data “solutions” are rarely appropriate in this situation. Three types of solutions are available. First, one can develop weights to correct for attrition. Second, one can model the attrition process directly, using “selection bias” models (Heckman 1979; Berk 1983). Finally, one can impute estimates of missing data (Little and Rubin 1987). This entire area is controversial and under rapid development. A special case of attrition occurs when observations are *censored* in such a way that measurement stops before a transition of interest, such as ending a panel study of fertility before all subjects have completed childbearing.

#### APPROACHES TO DATA ANALYSIS

The period since 1975 or 1980 has seen enormous increases in the power and sophistication of longitudinal data analysis. Singer (1985) is a useful

overview; Dwyer and Feinleib (1991), especially chapter 1, provides a comprehensive review, as does Von Eye (1990). This material is much too complex to cover in this brief review. A more reasonable goal is to provide a few examples of the kinds of substantive questions that one can ask with longitudinal data and appropriate methods for answering them.

**Outcomes at Fixed Points in Time.** Example: “predicting” a person’s savings at age sixty-five. This is the simplest longitudinal question; indeed, in one sense it is not longitudinal at all. The dependent variable involves change over time very indirectly, and the independent variables usually involve life cycle variables like the nature of labor-force participation. These data can often be collected retrospectively, and the only advantage of a *repeated measures* approach is control over measurement error.

It is not uncommon to see models for fixed outcomes at arbitrary time points—for example, level of education attainment as of the last available wave of an ongoing panel study. This is a particular problem when the sample consists of subjects who vary in age or is not defined on the basis of some reasonable baseline. When the outcome is a discrete transition, such as marriage, censoring is often a problem.

**Means over Time.** Example: Comparison of aspiration levels of boys and girls over the middle school years. With a sequence of independent samples, this is a straightforward analysis-of-variance problem. If the same individuals are measured repeatedly, a number of problems arise. First, the observations are not independent over time; we have to assume that the usual statistical assumption of independence in the error terms will be violated. Within the ANOVA tradition, this problem can be approached via multivariate analysis of variance or as a classic univariate design. Hertzog and Rovine (1985) compare the two approaches. Sociologists have tended to ignore mean comparisons, preferring to work with structural equations; however, analyses of means are often far more direct and informative. See Fox (1984) for an interesting application to an interrupted time series.

Classic ANOVA ignores individual-level heterogeneity. Over the course of five observation points, students vary about the group means. The description of change at the individual level rather than at the group level may be of some interest. A simple approach is to use “blocking variables” such as race to account for additional heterogeneity. Adding covariates to the model is another. A more sophisticated approach is to model the individual-level growth curve directly. Conceptually, analysis begins by estimating an equation to describe each respondent’s score with respect to time, often using a polynomial model. The coefficients of these equations are then treated as dependent variables in a second-stage analysis. There are a number of different statistical approaches to this kind of analysis. Rogosa (1988) has argued forcefully in favor of approaching longitudinal analysis in this way. Dwyer (1991) provides an interesting example. McCardle and Aber (1990) deal with such models in a LISREL context.

**Structural Equation Systems.** Figure 1, based on Blau and Duncan’s classic path model of occupational mobility, is a longitudinal analysis, although it is not often thought of in those terms. For our purposes, the important point is that the exact timing of variables was not assumed to be of great import, although the sequence was. Education was taken to intervene between measures of family background and occupational attainment, and its timing was not specified. The latter variable was assumed to reach some plateau relatively early in the occupational career; again, timing was not important. Models of this kind, with rather vague time referents and in which it is assumed that the order of events does not vary across subjects, have played an important role in sociology for some time, adding great clarity to the notion of intervening variables. It is natural to assume that structural equation models are a natural way to analyze multiwave panel data.

Kessler and Greenberg (1981) deal at length with the application of structural equation models to panel data, particularly to cross-lagged structures where both variables vary at baseline. They show that attempting to estimate relative causal effects must be handled in the context of a formal

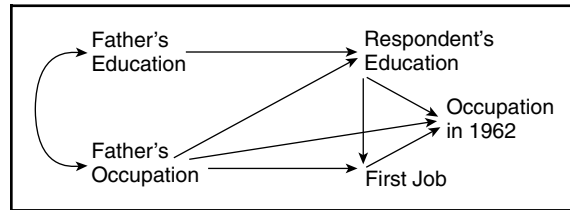


Figure 1

NOTE: A Basic Model of Attainment (Adapted from Blau and Duncan 1967)

model rather than by comparing cross-lagged partial correlations. Figure 2, based on Campbell and Mutran (1982), is a multiwave panel model with intervening variables. Here, a number of statistical and conceptual difficulties become obvious. First, what is the lag time for the effect of health on income satisfaction and vice versa? Second, how does one deal with the potentially complex error structure of observed variables? Third, if illnesses intervene (as measured by times in hospital), how does the timing of events relative to the timing of observations affect the results? Fourth, how does one deal with discrete intervening variables that fail to meet standard distributional assumptions? Finally, how does one handle correlated errors in equations over time? These are by no means the only questions that could be raised, and it is not clear that models of this kind are the most effective way to approach multiwave panel data.

**Event History Models.** This class of models treats the underlying rate of change in a discrete variable as the dependent variable in a regression-like framework. Typically, the dependent variable is conceived of as the instantaneous risk of transition at time  $t$  conditional on prior history. The formulation is dynamic rather than static, and time enters the analysis explicitly. Allison (this volume) discusses a number of such models in detail. At first glance, the model may seem mathematically esoteric, but the basic ideas are straightforward. The underlying concepts are closely related to the life table and to Markov models. The Markov model, as applied in the classic labor-force studies of Blumen, Kogen, and McCarthy (1955), assumes a transition process that is constant over time and invariant across subjects. Event history

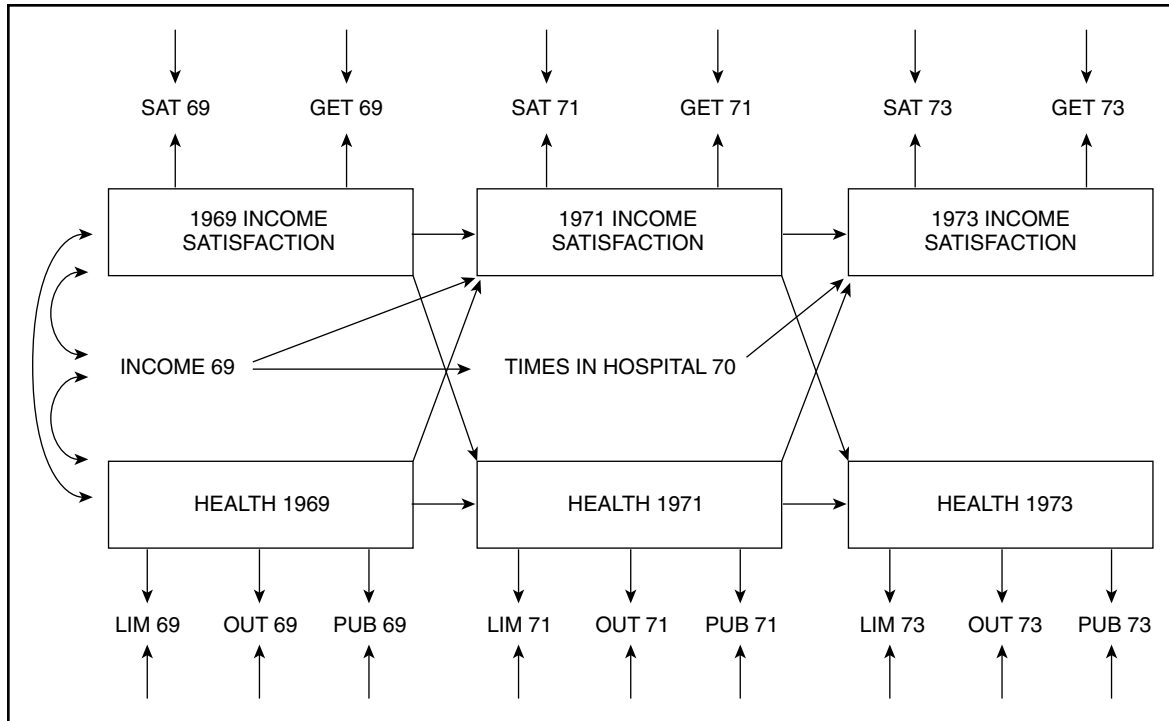


Figure 2

NOTE: A Model for Health and Income Satisfaction (Adapted from Campbell and Mutron 1982)

models allow one to relax both assumptions, specifying a time-dependent rate of change and a set of individual-level “covariates” that allow transition processes to vary with characteristics of the subjects. Covariates can be allowed to change with time as well, at the cost of substantial computational complexity.

In one of the first applications in sociology, Tuma, Hannon, and Groeneveld (1979) showed that the rate of marital dissolution among subjects in an income-maintenance experiment depended on the treatment condition, thus demonstrating that event models could be used for causal analysis. Event models are inherently dynamic and take the timing of transitions explicitly into account. They are having enormous impact not only on how sociologists analyze data but also on how they conceptualize problems and design data collection.

**Differential Equation Models.** Coleman (1964, 1968, 1981), among others, has argued that the appropriate way to study change in continuous

variables is via differential equation models. A simple differential equation relating  $Y$  to its own level and to an exogenous variable takes the form

$$\frac{dY}{dt} = a + b_1Y + b_2X_t$$

where  $dY/dt$  is the rate of change in  $Y$  with respect to time, that is, the derivative. The rate of change is not directly observed, of course; and to estimate the coefficients of the model, it is necessary to integrate the equation. Coleman (1968) shows that for this particular model, it is possible to obtain estimates of  $a$ ,  $b_1$ , and  $b_2$  by first estimating a regression equation of the form

$$Y_t = a^* + b_1^*Y_{(t-1)} + b_2^*X_t$$

using ordinary regression and then transforming the regression coefficients to obtain the coefficients of the differential equation model. In this particular case the coefficients of the differential equation model are

$$a = \frac{a^* C^*}{\Delta t}$$

$$b_1 = \frac{\ln b_1^*}{\Delta t}$$

$$b_2 = \frac{b_2^* C^*}{\Delta t}$$

where

$$C^* = \frac{\ln b_i^*}{(b_i^* - 1)}$$

The resulting coefficients describe the time path of  $Y$  as a function of its initial value and the value of  $X$ . Note that this model assumes that  $X$  is not changing over time. One implication of this model is that if one wishes to assume that  $Y$  at time  $t$  does not depend on its initial value, then residualized gain score models of the kind described above are inappropriate.

Applications of differential equation models in sociology have been relatively rare, although they would seem to be a natural way to approach the analysis of change. Many of the seemingly endless arguments about the representation of change in regression models stems from the application of static methods to dynamic processes. The difficulty is that it is not easy or always possible to transform the differential equations in such a way that they can be estimated by simple regression. Doing so requires considerable mathematical sophistication on the part of the researcher. Tuma and Hannon (1984) discuss these and related models at length. Arminger (1986) provides an example of how to recast a standard three-wave LISREL-based panel analysis into a differential equation model.

### CONCLUSION

It is no accident that words like *controversial* and *under discussion* recur frequently in this review. Many important issues regarding the collection, analysis, and interpretation of longitudinal data remain to be resolved. Frequently, what are in fact conceptual issues are argued in statistical terms. The literature is technical and mathematically demanding. But important progress is being made

and will continue as new methods of research continue to emerge.

(SEE ALSO: *Causal Inference Models*; *Cohort Analysis*; *Event History Analysis*; *Time Series Analysis*; *Quasi-Experimental Research Designs*)

### REFERENCES

- Allison, Paul D. 1984 *Event History Analysis: Regression for Longitudinal Event Data*. Newbury Park, Calif.: Sage.
- 1990 “Change Scores as Dependent Variables in Regression Analysis.” In Clifford C. Clogg, ed., *Sociological Methodology*, Vol. 20. London: Basil Blackwell.
- 1996 “Fixed-Effects Partial Likelihood for Repeated Events.” *Sociological Methods and Research* 25:207–222.
- Anderson, Edward R. 1993 “Analyzing Change in Short-Term Longitudinal Research Using Cohort-Sequential Designs.” *Journal of Consulting and Clinical Psychology* 61:929–940.
- Armingier, Gerhard 1986 “Linear Stochastic Differential Equation Models for Panel Data with Unobserved Variables.” In Nancy B. Tuma, ed., *Sociological Methodology*, Vol. 16. Washington, D.C.: American Sociological Association.
- Baltes, Paul B., and John R. Nesselroade 1977 “History and Rationale of Longitudinal Research.” In John R. Nesselroade and Paul B. Baltes, eds., *Longitudinal Research in the Study of Behavior and Development*. New York: Academic Press.
- Barton, Thomas R., Vijayan K. Pillai, and Tracy J. Dietz 1996 “Program Evaluation Using Event History Analysis.” *Evaluation Practice* 17:7–17.
- Berk, Richard A. 1983 “An Introduction to Sample Selection Bias in Sociological Research.” *American Sociological Review* 48:386–398.
- 1988 “Causal Inference in Social Science.” In Neil J. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- Bijleveld, Catrien C. J. H. et al. 1998 *Longitudinal Data Analysis: Designs, Models and Methods*. London: Sage Publications.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Blossfeld, Hans-Peter, Goetz Rohwer, and Gotz Rohwer 1995 *Techniques of Event History Modeling: New Approaches to Causal Analysis*, vol. 31. Mahwah, N.J.: Lawrence Erlbaum Associates.



- Blumen, Isadore, M. Kogan, and P. J. McCarthy 1955 *The Industrial Mobility of Labor as a Probability Process*. Ithaca, N.Y.: Cornell University Press.
- Bohrnstedt, George W. 1969 "Observations on the Measurement of Change." In Edgar F. Borgatta, ed., *Sociological Methodology*, Vol. 2. San Francisco: Jossey-Bass.
- Bollen, Kenneth A. 1989 *Structural Equations with Latent Variables*. New York: Wiley.
- Campbell, Richard T. 1988 "Integrating Conceptualization, Design, and Analysis in Panel Studies of the Life Course." In K. W. Schaie, R. T. Campbell, W. Meredith, and S. C. Rawlings, eds., *Methodological Issues in Aging Research*. New York: Springer.
- , and Elizabeth Mutran 1982 "Analyzing Panel Data in Studies of Aging." *Research on Aging* 4:3–41.
- Coleman, James S. 1964 *Introduction to Mathematical Sociology*. New York: Free Press.
- 1968 "The Mathematical Study of Change." In Hubert M. Blalock, Jr., and Ann B. Blalock, eds., *Methodology in Social Research*. New York: McGraw Hill.
- 1981 *Longitudinal Data Analysis*. New York: Basic Books.
- Collins, Linda M. 1996 "Measurement of Change in Research on Aging: Old and New Issues From an Individual Growth Perspective." In J. E. Birren and K. W. Schaie, eds., *Handbook of the Psychology of Aging*. San Diego: Academic Press.
- Cook, Nancy R., and James H. Ware 1983 "Design and Analysis Methods for Longitudinal Research." *Annual Review of Public Health* 4:1–23.
- Cronbach, Lee J., and Leta Furby 1970 "How We Should Measure Change—Or Should We?" *Psychological Bulletin* 74:32–49.
- DeShon, Richard P., Robert E. Ployhart, and Joshua M. Sacco 1998 "The Estimation of Reliability in Longitudinal Models." *International Journal of Behavioral Development* 22:493–515.
- DiPrete, Thomas A., and David B. Grusky 1990 "Structure and Trend in the Process of Stratification for American Men and Women." *American Journal of Sociology* 96:107–143.
- Duncan, Greg J., and Graham Kalton 1987 "Issues of Design and Analysis of Surveys Across Time." *International Statistical Review* 55:97–117.
- Duncan, Greg J., and James Morgan 1985 "The Panel Study of Income Dynamics." In Glen H. Elder, ed., *Life Course Dynamics: Trajectories and Transitions, 1968–1980*. Cambridge: Cambridge University Press.
- Duncan, Terry E. et al. 1999 *An Introduction to Latent Variable Growth Curve Modeling: Concepts Issues, and Applications*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Dwyer, James 1991 "Overview of Models for Longitudinal Data." In James Dwyer and Manning Feinleib, eds., *Statistical Models for Longitudinal Studies of Health*. Oxford: Oxford University Press.
- , and Manning Feinleib (eds.) 1991 *Statistical Models for Longitudinal Studies of Health*. Oxford: Oxford University Press.
- Featherman, David L. 1977 "Retrospective Longitudinal Research: Methodological Considerations." *Journal of Economics and Business* 32:152–169.
- Fox, John 1984 "Detecting Change in Level and Slope in Repeated Measures Analysis." *Sociological Methods and Research* 12:263–278.
- Freedman, David A. 1991 "Statistical Models and Shoe Leather." In Peter V. Marsden, ed., *Sociological Methodology*. Oxford: Basil Blackwell.
- Freedman, Deborah, Arland Thornton, Donald Camburn, Duane Alwin, and Linda Young-DeMarco 1988 "The Life History Calendar: A Technique for Collection of Retrospective Data." In Clifford C. Clogg, ed., *Sociological Methodology*, Vol. 18. Washington D.C.: American Sociological Association.
- Goodman, Jodi S., and Terry C. Blum 1996 "Assessing the Non-Random Sampling Effects of Subject Attrition in Longitudinal Research." *Journal of Management* 22:627–652.
- Gottman, John-Mordechai 1995 *The Analysis of Change*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Heckman, James J. 1979 "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–161.
- , and Burton Singer 1985 *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press.
- Hertzog, Christopher, and Michael Rovine 1985 "Repeated-Measures Analysis of Variance in Developmental Research: Selected Issues." *Child Development* 56:787–809.
- Hsiao, Cheng 1986 *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Inter-University Consortium for Political and Social Research 1991 *Guide to Resources and Services*. Ann Arbor, Mich.: Institute for Social Research.
- Jöreskog, Karl G., and Dag Sörbom 1988 *LISREL 7: A Guide to the Program and Applications*. Chicago: SPSS.
- Kessler, Ronald C., and David F. Greenberg 1981 *Linear Panel Analysis: Models of Quantitative Change*. New York: Academic Press.



- Lazarsfeld, Paul R., and M. Fiske 1938 "The 'Panel' as a New Tool for Measuring Opinion." *Public Opinion Quarterly* 2:596-612.
- Lieberson, Stanley 1985 *Making It Count*. Berkeley: University of California Press.
- Liker, Jeffrey, Susan Augustyniak, and Greg J. Duncan 1985 "Panel Data and Models of Change: A Comparison of First Order Difference and Conventional Two-Wave Models." *Social Science Research* 14:80-101.
- Little, Roderick J. A., and Donald B. Rubin 1987 *Statistical Analysis with Missing Data*. New York: Wiley.
- Magnusson, David et al. 1994 *Problems and Methods in Longitudinal Research: Stability and Change*. Cambridge: Cambridge University Press.
- Marini, Margaret M., and Burton F. Singer 1988 "Causality in the Social Sciences." In Clifford C. Clogg, ed., *Sociological Methodology*, Vol. 18. Washington, D.C.: American Sociological Association.
- Maxwell, Scott E. 1998 "Longitudinal Designs in Randomized Group Comparisons: When will Intermediate Observations Increase Statistical Power?" *Psychological Methods* 3:275-290.
- McArdle, John J. 1998 "Modeling Longitudinal Data by Latent Growth Curve methods." In G. A. Marcoulides, ed., *Modern Methods for Business Research. Methodology for Business and Management*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- , and Mark S. Aber 1990 "Patterns of Change within Latent Variable Structural Equation Models." In Alexander von Eye, ed., *Statistical Methods in Longitudinal Research*, Vol 1. San Diego, Calif.: Academic Press.
- McPhee, Robert D. 1998 "Alternate Approaches to Integrating Longitudinal Case Studies." In George P. Huber and Andrew H. Van de Ven, eds., *Longitudinal Field Research Methods: Studying Processes of Organizational Change*. Thousand Oaks, Calif.: Sage.
- Menard, Scott 1991 *Longitudinal Research*. Newbury Park, Calif.: Sage.
- Migdal, Susan, Ronald P. Abeles, and Lonnie R. Sherrod 1981 *An Inventory of Longitudinal Studies of Middle and Old Age*. New York: Social Science Research Council.
- Nielson, Francois, and Rachel A. Rosenfeld 1981 "Substantive Interpretations of Differential Equation Models." *American Sociological Review* 46:159-174.
- Petersen, Trond 1993 "Recent Advances in Longitudinal Methodology." *Annual Review of Sociology* 19:425-454.
- Pitts, Steven C., Stephen G. West, and Tein Jenn-Yun 1996 "Longitudinal Measurement Models in Evaluation Research: Examining Stability and Change." *Evaluation and Program Planning* 19:333-350.
- Raykov, Tenko 1993 "On Estimating True Change Interrelationships with Other Variables." *Quality and Quantity* 27:353-370.
- Rogosa, David 1988 "Myths About Longitudinal Research." In K. W. Schaie, R. T. Campbell, W. Meredith, and S. C. Rawlings, eds., *Methodological Issues in Aging Research*. New York: Springer.
- , and Hilary Saner 1995 "Longitudinal Data Analysis Examples with Random Coefficient Models." *Journal of Educational and Behavioral Statistics* 20:149-170.
- Rovine, Michael J., and Pete C. M. Molenaar 1998 "A Nonstandard Method for Estimating a Linear Growth Model in LISREL." *International Journal of Behavioral Development* 22:453-473.
- Singer, Burton 1988 "Longitudinal Data Analysis." In Samuel Kotz and Norman Johnson, eds., *Encyclopedia of Statistical Sciences*, Vol. 5. New York: Wiley.
- Smith, Tom W. 1986 "The International Social Survey Program." *Journal of Official Statistics* 2:337-338.
- Speer, David C., and Paul E. Greenbaum 1995 "Five Methods for Computing Significant Individual Client Change and Improvement Rates: Support for an Individual Growth Curve Approach." *Journal of Consulting and Clinical Psychology* 63:1044-1048.
- Thornberry, Terence P., Beth Bjerregaard, and William Miles 1993 "The Consequences of Respondent Attrition in Panel Studies: A Simulation Based on the Rochester Youth Development Study." *Journal of Quantitative Criminology* 9:127-158.
- Tuma, Nancy Brandon, and Michael T. Hannan 1984 *Social Dynamics: Methods and Models*. Orlando, Fla.: Academic Press.
- Vermunt, Jeroen K. 1997 *Log-linear Models for Event Histories*. Thousand Oaks, Calif.: Sage.
- von Eye, Alexander 1990 *Statistical Methods in Longitudinal Research*. San Diego: Academic Press.
- Wall, W. D., and H. L. Williams 1970 *Longitudinal Studies and the Social Sciences*. London: Heinemann.

RICHARD T. CAMPBELL

## LOVE

Sociologists agree that love is one of the most complex and elusive concepts to deal with from a scientific point of view. Indeed, they often point out that poets, novelists, and musical composers are much more adept at producing eloquent expressions about this pervasive sentiment. Dictionary definitions are of limited use in categorizing

the essential ingredients of love, except to connote its many variations as an attitude, an emotion, or a behavior. No one definition can capture all the dimensions of love, which can involve a wide range of elements such as romantic obsession, sexuality, caring, even irrationality. Indeed, some have argued: "There is no single, subjective meaning of love that everyone experiences in the same way" (Hendrick and Hendrick, 1992). Part of the difficulty is that individuals and their cultures define love very differently, depending on particular relationships and circumstances.

### **HISTORICAL CONCEPTIONS OF LOVE**

Conceptions of love have varied not only from one culture to another, but also from one historical era to another (Murstein 1974; Hunt 1959). Prominent among these are courtly love, Romanticism, and Victorian-era love (Hendrick and Hendrick 1992), as well as the modern era of love (Seidman 1991). Various forms of courtly love appeared in the twelfth century. This marked the beginning of the transformation of love from a philosophical or theological ideal to a practical way of relating between men and women. Sexual desire and expression were seen as one of the goals of love, and the love relationship was seen as intense and passionate. This bond was not limited to one's spouse, however. Marriage was seen as a more mundane and practical relationship. In contrast, courtly love was impractical, as the time commitment necessary to follow its elaborate rules and customs limited involvement to the wealthier or aristocratic classes. It often found its expression—sometimes sexual, but often not—in the idealized devotion of a knight or nobleman to a lady of nobility, who symbolized the perfect partner. Men and women did not relate in the day-to-day interaction associated with more recent intimate relationships.

In the late eighteenth to early nineteenth centuries, Romanticism replaced courtly love. In Romanticism it is possible to recognize the origins of some of our modern ideas about love: concern with similarities between partners, equality in the relationship, and the experience of the emotional side of love. This form furthered the courtly love ideal of sexual expression between partners as a worthy goal, with the added recognition that this fusion of emotion and sexuality could (and ideally should) occur with one's spouse.

The Victorian era (approximately 1830–1900) brought great changes to the idea of romantic love. This is hardly surprising, as people of this time were also adjusting to the changes in work and community life brought about by the Industrial Revolution. In the shift from the home economy to paid labor, the status of women declined, and new myths and ideas about love and its expression emerged. The partnership-oriented focus of Romanticism all but disappeared. Women in the Victorian era were seen as weaker and less intelligent than men. Within marriage, this assumption defined women as mothers and helpers, not individuals who might have complementary interests and a rightful concern for equality within a relationship. Women were also seen as childlike and asexual during this time, and were thought to need their husband's protection. This is an abrupt departure from earlier modes of loving, which acknowledged sexuality as an important aspect of the experience of love. In the Victorian era, sexual matters were not discussed between spouses.

In the modern era (1890–1960), the literature on love and marriage once again recognized the importance of an intimate bond between partners. Marital advice from this time shows a preoccupation with couples' happiness and the security of the marital bond. During this period, romantic love and sexual compatibility were inseparable, and this sexualization of love (Seidman 1991) brought the perception of sexual expression in romantic relationships full circle. Sexual expression was no longer seen as a component of love (whether an idealized or neglected one) but the very basis of love itself.

Sexual impulses and romantic love are often directed at the same person, but they are not the same. People are known to pursue sexual encounters in the absence of any romantic feelings. Sexual impulses are derived from our biological heritage, and romantic love is a learned cultural pattern. The two bear a relationship to one another, however, in that romance often encompasses sexually motivated behavior plus a cluster of cultural expectations (Hatfield and Rapson 1993)

### **CULTURAL DIFFERENCES**

The range of psychological and social meanings attached to love would, therefore, appear to be limitless. However, we experience and express

love mostly according to the culture and subcultures in which we have formed our sentiments. The formation of these sentiments begins very early and evolves through the physical and emotional attachments that characterize the parent-child relationship. While idiosyncratic patterns exist, love scripts by and large reflect the influence of social conditioning. In short, love is largely a learned response.

Despite cultural differences, several aspects of love seem to be universal (Sternberg 1998). Everywhere love involves four ideals: the suitable partner, the emotional experience of loving, the mental experience or thought process of loving, and the actions deemed acceptable and expected between the lover and his or her partner. Although these components seem to be found in all societies, how they are defined is culturally specific. The suitable partner may be of the opposite or the same sex, younger or older than oneself (or the same age), and tall and thin or short and heavy. Similarly, thoughts about love may be pragmatic or passionate, and emotions may range from deeply affectionate to reserved and respectful. The potential combinations that add up to "love" vary widely, and might be unrecognizable to a person from another culture. For example, in some tribes in Africa there is a traditional system known as "sweethearting" in which young people choose partners based on emotional and physical attraction (Goode, 1963). Even within this system there is variation. Some tribes allow sweethearting to lead to marriage, while others permit their young people to take lovers by choice but maintain the practice of arranging marriages for girls when they are still children. In the Arab world, love may also take secondary importance to familial concerns in taking a spouse. Many unmarried adults in Arab countries report that love is their ideal basis for marriage. However, married adults report varying levels of emotional closeness with their spouse prior to marriage, including love, acquaintance, and kinship ties without either love or acquaintance.

Cultural norms are internalized and program us to fall in love with specific types of people, within certain social contexts, to the exclusion of others. However, love is not necessarily related to marriage. There is a saying, for example, that in the West one falls in love and then gets married, whereas in the East one marries and then falls in love. In some societies, arranged marriages were

and still are contracted. The emotional intensity a couple feel toward each other is given little or no consideration. Instead, emphasis is given to the sociopolitical implications of the marital alliance for the families and kinship groups involved.

In the United States, on the other hand, love is viewed as an important condition to marriage. We are generally suspicious of anyone who would marry for any other reason. We are not comfortable with the idea that in some cultures a man marries his mother's brother's daughter because that is the prescribed pattern. Even among our upper classes, where concern about protecting family resources leads to a greater emphasis on practical considerations in mate selection, couples are expected to espouse mutual love as the basis for their marriage, or else their motives become suspect. Revelation of marital alliances designed to preserve or enhance family wealth often receive a cynical response from the public at large.

Families recognize that courtship and mate selection merge not only two individuals but also two different kinship lines, which in turn may affect their socioeconomic and political stature. Consequently, families invest considerable energy and resources to control love (Goode 1959). Several mechanisms to accomplish this have been identified, including the direct control provided by (1) child marriages, in which betrothal may occur before puberty or even before the child is born; (2) defining the pool of eligibles, that is, delineating whom one can and cannot marry; (3) physical or social isolation to limit the probabilities of contact; and (4) various indirect controls such as moving to preferred residential neighborhoods, enrolling children in appropriate schools, joining select organizations, or attending certain churches. The latter mechanism is most characteristic of Western societies.

### THE ROMANTIC LOVE COMPLEX

In American society a *romantic love complex* exists, and this complex posits love as a central prerequisite to marriage. The basic components of this complex are assimilated through the mass media—through romantic stories in novels, magazines, television, and movies. In this way we are psychologically prepared to fall in love. The major

characteristics of romantic love include *romantic democracy*; that is, cultural differences between couples are minimized or ignored because “love and love alone” is sufficient. Indeed, it involves the notion that romantic love thrives on such differences. Romantic love also includes *romantic intensity*; that is, people are expected to fall in love instantly (to experience love at first sight) and deeply, with great emotional attachment. Finally, romantic love includes *romantic monopoly* in that once the “bolt from the blue” strikes, the couple presume exclusive emotional and social rights to each other, in perpetuity (Merrill 1959). A person experiencing the full thrust of this complex is, supposedly, consumed by constant thoughts about the beloved, a longing to spend all one’s time with that person, a sad pining in the beloved’s absence, and a feeling that life would not be worth living without him or her (Tennov 1980).

There is disagreement about the extent to which people adhere to the tenets of the romantic love complex and whether it actually influences mate selection. Sociologists have generally viewed it as a poor basis for the establishment of permanent unions, inasmuch as it involves an element of capricious choice based upon an unpredictable emotion. Moreover, romantic love is not completely rational. Part of its credo is that there is one and only one true love or ideal mate. Yet we know that people fall in and out of love several times in a lifetime.

People caught up in the romantic complex often idealize their partners. Some argue that this process of *idealization*, in which a distorted positive picture of the love partner is constructed, results from the blockage of sexual impulses by cultural prohibitions. If this were true, then one would expect a liberalization of our sexual mores to be accompanied by a decline in romanticism. However, this does not appear to have happened, at least in Western societies. What does happen is that the elements of the romantic love complex, including the idealized picture of the partner, are modified to fit the reality of that person that emerges through close and intimate interaction over time. Couples unable to make this accommodation are apt to suffer disillusionment, and their marriages may encounter persistent difficulties. For marriage to succeed, the overly romanticized notions and idealizations of romantic love must eventually be replaced by *conjugal love*, which is

based upon habits, common interests, mutual acceptance, and mature companionship derived from a shared history.

### THE MASS MEDIA AND LOVE

The mass media undoubtedly encourages many romantic myths about love. This can be problematic for several reasons (Ellis 1985). Fictional lovers, whether in the movies or in novels, can magically go about their romantic business without the mundane everyday constraints faced by real people. Fictional portrayals of romantic love also tend to idealize the partners as beautiful, perfect people who are in the earliest passionate stages of love. It is assumed that this stage of love will last forever for the couple. In reality, few of us are unencumbered by family expectations or blessed with perfect beauty. Perhaps more damaging, the passionate stage of love shown on the silver screen tends to mature and deepen in real life into a more companionate form of love—which may then be taken as not being “real” love by those who subscribe to the romanticized media ideal.

Media portrayals of love and family life may also set up a template for behavior that is difficult or impossible for a couple to enact and sustain, while at the same time making them feel wrong somehow for failing to meet this expectation (Coontz 1992). Many intimate partners do not choose to follow the homemaker-breadwinner pattern in their relating. For some others who would like to be able to do so, this may be a financial impossibility. To make matters even more complicated, mass media presents other images of lifestyles that are at odds with the nuclear family image. Messages of consumerism and mass consumption in television shows and commercials encourage individuals to keep their choices open and to accept only the best. This message of individualism makes intimacy and commitment between romantic partners more difficult, as some may shy away from choosing a partner out of fear that something better may come along. Our kinship system does not provide a protective mechanism against this individualism. Because we trace kinship bilaterally through both the mother and the father, lineage concerns do not keep couples together (Farber 1964). As a result, marital status and presumed romantic exclusivity do not prevent

partners from being available as a potential mates to others.

### GENDER AND LOVE

There appear to be gender differences with respect to love and loving. It is generally assumed that such differences reflect culturally defined sex roles. Women have been stereotypically portrayed as starry-eyed romantics, while men are viewed as exploitative realists. However, research shows that men fall in love more quickly than women do and with less deliberation, score higher on scales of romanticism, express stronger romantic attitudes, and suffer greater emotional stress when relationships are terminated. Compared to men, women are more apt to exhibit a companionate rather a passionate approach to love. The difference between the two centers around the element of emotional intensity. Companionate love "is a calm, steady, relaxed state; passionate love is an emotional roller coaster, with intense highs and lows" (Fehr 1995). Because women may have potentially more to lose from a social and an economic standpoint, they tend to be more prudent or discriminating in establishing and maintaining love relationships (Hochschild 1983). They are more apt to take into account practical considerations regarding mate selection. Hence, in comparison to men, women are more likely to terminate relationships and are able to disengage emotionally more easily when they break up (Rubin 1973).

Men and women also have different expectations of love (Cancian 1987). The common denominator among many of these differences seems to be communication styles and preferences. Women have an advantage in communicating emotionally, because traditional gender norms have allotted the jobs of nurturing and relationship maintenance to them. They also tend to have more emotionally close (but not sexually intimate) relationships than do men. This suggests that women have less difficulty with emotional vulnerability and communication. Women tend to want more intimacy through verbal communication from their partners. This can cause a host of problems, as men tend to express feelings of love actively rather than verbally.

Love is also experienced differently by men and women partly because of varying levels of

dependency. Although a large proportion of women now hold paid employment, economic dependence on men is still a factor in many wives' and mothers' experience. They are more likely to put their careers on hold to raise children than are men, to earn less than their spouses, and to relocate due to a spouse's career needs. The critical link between gendered experience of love and financial concerns can be seen in socioeconomic class differences. Working- and middle-class men are less likely to identify strongly with their work than are men in the upper middle class. As a result, it is in the upper middle class that conflict between wives' needs for intimacy and husbands' desires for self-fulfillment and career achievement may be felt most acutely.

The gendered differences inimical to romantic love are found in communication as well (Gray 1992). Men and women have vastly different communication styles, and problems arising from miscommunication may lead to difficulties within a relationship. For a man to feel truly loved he must feel that he is needed. For a woman to feel loved she must feel understood and appreciated as well. These requirements stem from men's desire to feel competent and women's desire to feel socially and emotionally connected to others. When partners relate, however, they tend to speak to one another in ways that they would like to be spoken to themselves because they do not recognize these differing bases for feeling loved. Women tend to air their problems and difficulties with their partners in order to feel validated and understood. In response, men tend to offer solutions to women's problems or to minimize these problems—offering the vote of confidence in their partners' competency that they would want in similar circumstances. These differences can potentially undermine feelings of love experienced by the partners (Gray 1993). A woman is left feeling that her partner does not truly listen or understand her when he offers solutions to her grievances, and she feels abandoned when he withdraws to consider his own problems in privacy. A man can feel blamed by his partner's complaints and stunned by her negative reaction to his proposed solutions, and tends to feel hounded if she attempts to penetrate the silence he needs to work issues out for himself.

Social scientists have explored the sequences through which the dimensions of love relationships develop and have constructed numerous

theories, models, and typologies of love (Fehr 1995). They have also detected and described a number of styles of loving (Kemper 1988; Lee, 1988; Hendrick and Hendrick, 1986; Lasswell and Lasswell 1980). Love is said to begin typically with physical symptoms—palpitations of the heart, rapid breathing, sweating, and so on. At this stage the symptoms are essentially similar to those associated with other emotions such as fear. Next the person proceeds to label his arousal as a love response. This labeling process gains impetus from social pressures and cultural dictates, which prod one to define the experience as love and to follow its ritualistic patterns.

The optimum conditions for love to flourish require that the couple be equally involved in and committed to each other and the relationship. Where there is unequal involvement, the person with the strongest commitment may be vulnerable to exploitation. This is known as the *principle of least interest*, in which the partner with the least interest has the most control. Few relationships that are based on this principle can endure.

(SEE ALSO: *Courtship; Interpersonal Attraction; Mate Selection, Theories*)

## REFERENCES

- Cancian, Francesca M. 1987 *Love in America: Gender and Self-Development*. Cambridge and New York: Cambridge University Press.
- Coontz, Stephanie 1992 *The Way We Never Were: American Families and the Nostalgia Trap*. New York: Basic Books—HarperCollins.
- Ellis, Albert 1985 "Romantic Love." In John F. Crosby, ed., *Reply to Myth: Perspectives on Intimacy*. New York: Wiley. pp. 218–227.
- Farber, Bernard 1964 *Family: Organization and Interaction*. San Francisco: Chandler.
- Fehr, Beverley 1995 "Love." In David Levinson, ed., *Encyclopedia of Marriage and the Family*, vol. 2. New York: Simon and Schuster—Macmillan.
- Goode, William J. 1959 "The Theoretical Importance of Love." *American Sociological Review* 24:38–47.
- 1963 *World Revolution and Family Patterns*. New York: Free Press.
- Gray, John 1992 *Men Are from Mars, Women Are from Venus*. New York: HarperCollins.
- 1993 *Men, Women and Relationships*, rev. 2nd ed. Hillsboro, Oreg.: Beyond Words Publishing.
- Hatfield, E., and Rapson, R. L. 1993 *Love, Sex, and Intimacy: Their Psychology, Biology, and History*. New York: HarperCollins.
- Hendrick, Clyde, and Susan S. Hendrick 1986 "A Theory and Method of Love." *Journal of Personality and Social Psychology*. 50:392–402.
- Hendrick, Susan S. and Clyde Hendrick 1992 *Romantic Love*. Newbury Park, Calif.: Sage.
- Hochschild, Arlie R. 1983 *The Managed Heart*. Berkeley: University of California Press.
- Hunt, Morton 1959 *The Natural History of Love*. New York: Knopf.
- Kemper, Theodore 1988 "Love and Like and Love and Love." In David Franks, ed., *The Sociology of Emotions*. Greenwich, Conn.: JAI Press.
- Lasswell, Marcia, and Norman M. Lasswell 1980 *Styles of Loving*. New York: Doubleday.
- Lee, J. A. 1988 "Love Styles." In R. J. Sternberg and M. L. Barnes, eds., *The Psychology of Love*. New Haven, Conn.: Yale University Press.
- Merrill, Francis E. 1959 *Courtship and Marriage*. New York: Holt-Dryden.
- Murstein, Bernard I. 1974 *Love, Sex, and Marriage through the Ages*. New York: Springer.
- Rubin, Zick 1973 *Loving and Liking*. New York: Holt, Rinehart and Winston.
- Schwartz, Gary, Don Mertem, Fran Beham, and Allyne Rosenthal 1980 *Love and Commitment*. Beverly Hills, Calif.: Sage.
- Seidman, Steven 1991 *Romantic Longings: Love in America, 1830–1980*. New York and London: Routledge.
- Sternberg, Robert J. 1998 *Cupid's Arrow: The Course of Love through Time*. Cambridge and New York: Cambridge University Press.
- Tennov, Dorothy 1980 *Love and Limerence: The Experience of Being in Love*. New York: Stein and Day.

FELIX M. BERARDO  
ERICA OWENS

# M

---

## MACROSOCIOLOGY

The term “macro” denotes “large”; thus *macrosociology* refers to the study of large-scale social phenomena. This covers a very broad range of topics that includes groups and collectivities of varying sizes, the major organizations and institutions of one or more societies, cross-sectional or historical studies of a single society, and both comparative and historical analyses of multiple societies. At the grandest level it may cover all human society and history.

Sociologists distinguish macrosociology from microsociology, which focuses on the social activities of individuals and small groups. The micro-macro distinction forms one of the central dualisms characterizing divergent sociological perspectives. Seemingly polar opposites such as conflict-consensus, stability-change, structure-agency, subjective-objective, and materialist-idealist, as well as micro-macro, provide a shorthand method for denoting differences in central assumptions, subjects, and models. As with many other oppositional concepts, however, the boundary between microsociology and macrosociology is not clearly distinguished, and at the margins there is much room for overlap.

Typically, micro-level studies examine individual thought, action, and interaction, often coinciding with social-psychological theories and models, whereas macro-level investigations target social structures and those forces that organize as well as divide individuals into political, social or religious organizations, ethnic populations, communities, and nation-states. Nevertheless, in defining these

terms there is major conceptual ambiguity that can be formulated as a question: Should the distinction be based on substantive criteria (specialty and subdisciplinary areas within sociology such as social change and development), theoretical criteria (e.g., functionalist, Marxist), metatheoretical criteria (type of paradigm, epistemology), or methodological criteria (type of research design and analysis techniques)? Since sociologists often use the terms “micro” and “macro” quite casually as convenient devices for categorizing broad areas of theory and research, each of these criteria can be found in the literature, and quite often they are seriously confounded.

A useful means of distinguishing between the two approaches is based on the concept of “units of analysis.” Macrosociology uses as its subjects structural-level units of analysis or cases that are larger than observations of individual action and interaction. Even here, however, there is ambiguity, since it is quite possible to make observations on smaller units (e.g., individuals) with the intention of analyzing (making inferences about) larger entities (e.g., groups, classes). Also, the issue of where to draw the line remains. Rather than attempting to draw any hard-and-fast line delineating macro-level from micro-level phenomena, it is helpful to conceptualize a continuum of the subject matter of sociology with “micro” and “macro” defining two end points and with societal-level phenomena clearly placed at the macro pole. George Ritzer, for example, describes one “level of social reality” as a micro-macro continuum moving from individual thought and action through

interaction, groups, organizations, and societies to culminate in world systems (Ritzer 1988, pp. 512–518).

Since the macro end of the continuum focuses on social structure, it is important to clarify the use of this term. In a review essay, Neil Smelser (1988, pp. 103–129) describes structure as patterned relationships that emerge from the interaction of individuals or groups over time and space. Institutions and identifiable collectivities are the outcomes of systematically related structures of activities. Structure is dually defined as located in collective actors and in their interaction. Thus social class is an example of social structure, as are the relationships between classes whose locus is the economy. The study of social class and the study of the economy are examples of macrosociology.

Other examples emerge from the macrosociological focus on large-scale structural arrangements and activities of a great number of individuals in large-scale geographical space over long periods of time. Thus macroscopic questions in sociology conventionally revolve around the largest social, spatial, and temporal processes, such as the rise and decline of civilizations; the origins and development of modern nation-states, social movements, and revolutions; and the origins and consequences of social, political, economic, and cultural transformations. Examples include the rise and spread of secular ideologies and religious belief systems, democratic transitions, and the nature and effects of large-scale institutions and organizations. Macro-level analysis is usually embedded in structural and conflict theories, and in studies of societal dynamics and epochal transformations of cultures and social structures. Topics are located within numerous subfields of sociology, including but not limited to stratification and inequality, resource mobilization, political and economic sociology, world systems, human evolution, and ecology. They are equally likely to cross or link disciplinary boundaries to incorporate history, geography, political economy, and anthropology.

### HISTORICAL BACKGROUND

The concern with macro-level phenomena is as old as the discipline of sociology and arguably is the primary motivation for the creation of classical sociological theory and research. The men generally accorded honored places in the pantheon of

sociology's founders, such as Auguste Comte, Herbert Spencer, Karl Marx, Emile Durkheim, and Max Weber (and additional historical figures such as Alexis de Tocqueville), all included macro-level phenomena among their dominant concerns. The traditions they established retain their definitive role for the central issues of sociology in general and macrosociology in particular.

The themes pursued by these and other classical theorists are found in subsequent theory and research. For example, the evolutionary perspectives on the development of human society advanced by early theorists have been modified, revised, and developed by contemporary evolutionary theorists such as Lenski (Lenski et al. 1995) and in the modern functionalist and neo-functionalist theories of Talcott Parsons (1966), Niklas Luhmann (1982), and Jeffrey Alexander (1998). Marx's historical materialist explanation of the unfolding of capitalism has spawned numerous offspring, including dependency and world system theories (Amin 1976; Frank 1967; Wallerstein 1974), and studies of the rise of the modern state and class conflict by Moore (1966). Similarly, Weber's comparative and historical studies of social stratification and the development of modern states are reflected in the works of Reinhard Bendix (1977; 1978), Theda Skocpol (1979) and Michael Mann (1986; 1993). Emile Durkheim's analysis of the division of labor in modern societies as well as the sources of societal integration underlie all modernization theory and functional perspectives on race and ethnic relations, as well as most contemporary studies of occupational structures. Alexis de Tocqueville's comparative study of democracy has remained an inspirational source for contemporary theories of democracy and social change. In short, the macrosociological problems defined early in the history of sociology remain major focuses of current sociological research.

Also located in these early works but often overlooked in subsequent interpretations is an issue that is the current central project of many social theorists: the links between macro- and micro-level phenomena. At least in the writings of Marx, Weber, and Durkheim, to a greater or lesser degree, efforts are made to connect individual and structural level activities in some coherent fashion. For example, Marx is often considered the quintessential macrosociologist, providing the foundation for much current macrosociology. Yet as



Bertell Ollman (1976) and others point out, there is a distinct social psychology anchored in Marx's concept of alienation that in turn motivates and is motivated by his macro-level modes of productive relations and class conflict. This concern with linkage has often been ignored or forgotten in the distinctive development of different schools of sociological thought. After years of separate development and sometimes acrimonious debate, efforts to conduct research and develop social theories that include both ends of the micro-macro continuum now constitute a major agenda for many sociologists.

### THEMES IN MACROSOCIOLOGICAL THEORY AND RESEARCH

Macrosociological studies vary in both subject and theoretical orientation, but the two are closely related. For example, large-scale studies of single total societies or particular societal institutions often operate from a functionalist or systems perspective in which the effort is to understand how component parts fit together and serve larger social goals. On the other hand, studies of social change, either within a single society or across cultures, more often use one of the many variants of conflict—Marxist, neo-Marxist, and Weberian perspectives. They do so because such theories are better equipped to explain conflict and change than the relatively static models promoted by functionalism, and because functionalism no longer dominates sociology. These are broad generalizations, however, which invite counterexample.

Given the sweeping scope of macrosociology, it is not possible to provide comprehensive coverage of all the topics and theories subsumed under this approach. The next section will illustrate key concerns of macrosociologists by describing exemplary theory and research in some major areas of macrosociology.

**Societal Evolutionary Change.** The numerous approaches to the study of societal change illustrate the diversity of sociological perspectives. At the most sweeping level, evolutionary theories take all human history and society as their subject, but there are numerous variants to this approach. For example, evolutionary theory has gone hand in hand with functionalism, as in the later work of

Talcott Parsons on human societies (1966), which features the basic assumptions of evolutionary theory in terms of holism (the whole unit rather than its parts), universalism (natural and perpetual change), and unidirectionality (progressive and cumulative change). An idealist version of an evolutionary perspective can be found in Jurgen Habermas (1979), who uses an evolutionary model to explain the development of normative structures and forms of rationality. Alternatively, it has also taken a materialist form, as developed by anthropologists (Harris 1977) and a few sociologists (Lenski, 1966, Lenski et al. 1995), to explain inequality and uneven distribution of social resources. Another version of societal evolutionary change that deviates somewhat from the mainstream of progressive evolution are the cyclical dynamics of societal and cultural change proposed in works of Pitirim Sorokin (1962). Evolutionary analysis also was once popular in the fields of human ecology (Hawley 1971), modernization (Smelser 1964), and structural and cultural assimilation of different racial groups in modern society (Gordon 1964).

Currently, there are relatively few sociologists who operate on this scale or who find it useful for analyzing more confined periods of historical change. Nevertheless, contemporary theories of human evolution have been influential in providing comparative evidence for the material and normative bases of different forms of social organization and for describing the broadest patterns of societal change. These include the distribution of societal goods and services, enduring forms of inequality (e.g., patriarchy), and normative systems.

**Modernization and Development.** Sociologists often limit their study of change to the emergence of modern industrial society, either to trace the paths taken by mature industrialized societies to reach their current state of development or to investigate the problems of developing nations. Here, too, different approaches emerge from different theoretical perspectives.

Modernization theory, which until the 1960s dominated accounts of development and change, grew out of functionalism and evolutionary perspectives. In the version articulated by economist W. W. Rostow (1960), nonindustrial societies,

through diffusion and a natural developmental sequence, were expected to follow a series of stages previously traversed by fully industrialized nations to attain the significant characteristics of modern societies considered prerequisites for development. This process required breaking from traditional social norms and values to build institutions based on "modern" values such as universalism, rationalization, and achievement orientation.

Although today largely abandoned in favor of more historically and materially grounded theories, modernization theory was highly influential among both scholars and policy makers of the post-World War II era. In fact, it can be argued that the influence of modernization theory in part explains its repudiation, since students of and from emerging developing nations viewed it as an instrument of continued colonial domination and capitalist exploitation. Their search for tools to provide a better explanation for their disadvantaged and subordinate position in the international arena led to the adoption of Marxist-based models of dependency, underdevelopment, and world systems to replace modernization as the dominant approach to change and development within the modern era. As summarized by Peter Evans and John Stephens (1988, p. 740), these "approaches turned the modernization theorists' emphasis on diffusion . . . on its head, arguing instead that ties to 'core' countries were a principal impediment to development."

In an influential early formulation, dependency theorist Andre Gunder Frank (1967) maintained that the experience of most nonindustrial nations is explained by the "development of underdevelopment." In other words, the exploitation of peripheral Third World nations by capital in the core, developed world increased the economic, social, and political misery experienced by the majority populations of those Third World countries. Alliances between local and international elites actively worked to defend the status quo distribution of power and privilege at the expense of peasant- and working-class majorities. Later versions refined the models of class conflict and competition or, as in the writings of Samir Amin (1976), elaborated the model of the relationships between center and periphery economies to show how underdevelopment grows out of the exploitive links between the two types of systems.

All versions contribute to a refutation of the trajectory of development described by modernization theory.

A more global approach to development issues was formulated by Immanuel Wallerstein (1974, 1980, 1988) and his followers. World system theory elaborates the Marxist model of economic domination into a system in which exploitation occurs worldwide. Wallerstein broadens the focus on class relations among and across nations to examine the development of an international division of labor in the capitalist world economy where core industrial nations exploit peripheral regions as sources of raw materials and labor. This approach has been both enormously influential and controversial, generating massive amounts of research on the model itself, particular spatial and historical portions of the world system, and particular subsectors and groups. A helpful overview that charts the intricacies of this perspective can be found in a text by Thomas Shannon (1989).

The emphasis on First World as well as Third World development found in world-system theory provides a bridge to a slightly different tradition that focuses on the emergence of the core industrial nations and their political systems. Much of this literature is concerned with development of modern political as well as economic systems. For example, while Barrington Moore's study (1966) of the transformation of agrarian societies into modern industrial states remains firmly anchored in a Marxist emphasis on class relations and productive systems, it is also concerned with the political roles played by antagonistic classes and the political outcomes of their confrontation. Numerous other studies pursue a similar comparative perspective on the upheavals that accompanied the emergence of modern Western industrial nations (cf. Tilly et al. 1975).

One other type of study in this tradition deserves mention. These are studies of social and political change that occur within a particular society at various stages in the industrialization process. John Walton (1987) provides a convenient typology based on cross-classifying epochs and processes of industrialization. The resulting types range from protoindustrialization through deindustrialization. Studies from early periods focus on the emergence of particular classes, on class

conflict, and on the influence of classes on the historical development of modern nations, as in E. P. Thompson's and John Foster's influential accounts of English class formation (Thompson 1963; Foster 1974), Ron Aminzade's analysis of nineteenth-century France (1984), and Herbert Gutman's studies of American class culture and conflict (1966). Influential studies of transitions in later periods of industrial development examine the consolidation of control of the labor process (Burawoy 1979; Edwards 1979), deindustrialization (Bluestone and Harrison 1982), informalization of labor markets (Portes et al. 1989), and post-Fordist production systems (Mingione 1991).

Finally, while beyond the scope of this review, there are also other important traditions that have strong links to one or the other approaches described above. One of these is found in a vast literature on social movements that has many points of intersection with the work on comparative and historical social and political change discussed here. Another is work that applies the theories of dependency and uneven development to regional development problems internal to particular societies. Finally, there are structural and poststructural approaches to the development of major social institutions and forms of repression, as found in the complex but influential work of Michel Foucault (1979, 1980, 1985). While this last example could as easily be classified under studies of social institutions and processes, it is included here because of its focus on changes in historical times that have produced modern social forms.

**State Formation and State Breakdown.** The study of state formation and state breakdown has always been a central focus of macrosocial inquiry, especially in the area of comparative historical sociology. Studies of state formation examine the nature of state power and the processes by which it develops. While some sociologists have seen the state as emerging from internal dynamics of society, largely in terms of the interests and struggles of social classes, others have turned their attention to the external dynamics of society along with the market forces of the capitalist system. Tilly (1975, 1990) demonstrates that the modern states were created in the process of capital concentration and consolidated under the pressure of increased international military competition (war and preparation for war).

Michael Mann (1986, 1993) examines the nature of power in human societies by focusing on the interrelations of four principal sources of social power—economic, ideological, military, and political resources—and relates them to the rise of city-states, militaristic empires, modern nation-states, and nationalism. Another important theory on the relations between the state and society is Robert Wuthnow's (1989) analysis of how conflicts between the state, elites, and cultural entrepreneurs caused the great ideological movements to challenge the status quo in the development of modern society.

Contemporary theory of revolution and state breakdown starts with Barrington Moore (1966) who proposed a model of agrarian class politics. Drawing upon both Marxian and Weberian theoretical perspectives, Skocpol's (1979) breakthrough analysis introduced the state-centered theoretical paradigm of revolution in her case study of French, Russian, and Chinese revolutions. Treating the state as an "autonomous entity," Skocpol argues that the state has its own military and fiscal interests, and that under certain circumstances, state interests necessarily are in deep conflict with the interests of social classes. State breakdown occurs when the state experiences high levels of fiscal crisis induced by strain on resources from both internal elite conflicts and external military pressure. In Skocpol's theory, the state thus becomes the central actor and the location of crisis in revolutionary situations.

This state-centered theme developed by Skocpol has been further expanded by Goldstone (1991), who uses a structural-demographic approach to indicate that the early modern boom in population led to strain on state resources associated with the taxation system and economic development. Goldstone argues that in a system tied to agricultural output, the agrarian state depends mainly on land taxes for revenue. As growing population places pressure on the agricultural economy, rising grain prices result in inflation that erodes state revenues, leads to higher taxes, and exacerbates elite conflict. Rising prices generate profits for some elites who are quick to take advantage of commercial opportunities, but hurts other elites who are slow to adjust and lag behind in social mobility. Revolution is the ultimate outcome of the state's failure to meet its obligations.

While Skocpol's and Goldstone's models emphasize either structural or demographic sources of strain on the state, interest in geopolitical principles and strains became increasingly prominent during the 1980s, inspired by Paul Kennedy's analysis (1987) of the rise and fall of great powers. Randall Collins's geopolitical theory (1986, 1995) offers another route to state breakdown. Bringing in the Weberian principle that legitimacy of the governing apparatus at home depends on the state's power and prestige abroad, Collins's analysis, given validation by his prediction of the collapse of the Soviet Union, demonstrates that a state's geopolitical position has a crucial effect on its ability to mobilize critical resources and manage internal politics. In Collins's model, geopolitical strains result in inability to maintain fiscal health. A state that suffers the geopolitical disadvantage of being surrounded by multiple enemy states experiences logistical overextension and fiscal crisis, and thus tends to decline and disintegrate to the point of revolution and state fragmentation.

#### **Social Structures, Processes, and Institutions.**

The research described above incorporates investigation of many of the major social structures, processes, and institutions that form the core subject matter of sociology. Studying change in economic and political systems requires scrutiny of economies, politics, and other social institutions and their major organizational manifestations and constituencies. However, other theoretical and substantive approaches subsumed under macrosociology either have fallen outside the scope of these large-scale studies of social change and development or are at their periphery. Theoretical perspectives include relatively recent developments such as structural, poststructural, postmodern, and feminist theories. Important substantive areas are defined by cumulating empirical bases of knowledge about power structures; work structures; social stratification and mobility; labor markets; household and family arrangements; and the intersections of race, class, gender, and nationality.

While it is impossible to survey each of these areas, the explosive growth of feminist theories to investigate both gender stratification and economic change and development provides a prime example of new influences on macrosociology. Feminist theorists argue that gender analysis must be

integrated with class, race, ethnicity, nationality, and other sources of social cleavage, and that analyses that ignore the system of gender relations embedded in society are incomplete. Feminist theories have contributed to macrosociology by demonstrating how theories of social reproduction must be joined to theories of economic production to understand social life fully, thus delineating the ways patriarchy coexists with particular economic and political systems to explain the position of women in society.

For example, the subordination of women is predicated on the allocation of tasks that exist outside formal labor markets such as household and reproductive labor and consumption activities as well as labor market work. Heidi Hartmann's early, influential, socialist feminist analysis of the intersection of capitalism and patriarchy (1981) explains women's disadvantaged status in both the labor market and the household in late capitalism as the outcome of an uneasy alliance between the two systems. With increasing demand for women's labor in the second half of the twentieth century, the intersection of the two systems has taken the form of the double and even the triple day—that is, women burdened by responsibility for formal labor market activity; household work; and, frequently, informal work as well.

Similar insights from feminist perspectives have informed studies of developing nations and processes of industrialization and globalization. For example, Ester Boserup's critique of conventional development theories (1970) demonstrates the pitfalls for development projects resulting from ignoring women, as well as the ways women have been marginalized by development scholars and practitioners. Numerous feminist scholars have built on this and related work, combining it with other theoretical perspectives such as world systems and globalization theories, to expand knowledge of the gendered social consequences of core nation exploitation of the periphery (Ward 1990) and the general pattern of ignoring women in large-scale societal accounting schemes (Beneria 1981). Postcolonial theories and "Third World feminism" further explore the intersections of race, class, and gender as they influence different populations in the global economy (Alexander and Mohanty 1997). Finally, the historical research of Louise Tilly and Joan Scott (1978), among

others, has been important in an understanding of how the shift from household economies to wage labor affected working-class women and their families. Unfortunately, much of this work remains underutilized and unincorporated in the kind of macro-level analyses reviewed in previous sections, representing parallel developments rather than integrated studies of macrosocial processes.

### RESEARCH METHODOLOGY OF MACROSOCIAL INQUIRY

In the past decades, research methodology in macrosociology has been widely discussed among sociologists. Both quantitative and qualitative methods are used extensively, often in the same larger study. Virtually any methodological tool available to social science is found in macrosocial analysis, ranging from survey research to hermeneutic inquiry.

Quantitative approaches include quantification of documentary and archival data, such as analysis of the lists of grievances, or *cahiers de doléances* (Markoff 1996); analysis of official socioeconomic, demographic, and political data aggregated for larger geopolitical units such as counties, states, or nations; time series of such data for a single state or nation; and standard survey research techniques interpreted to represent structural and contextual process (Coleman et al. 1970). Trend analysis using survey data is one method frequently used by sociologists to establish long-term patterns of change by examining historical change in statistical data. Quantitative methods that use longitudinal designs of panel and cohort analyses to conduct observations at two or more points in time have been extensively employed in the assessment of social change and development at the local, national, and global levels.

Historical and comparative methods are featured prominently in macrosociological analysis and have been consistently used by the most prominent classical and contemporary sociologists. This approach develops ideal-typical case studies of large-scale organizations, nations, and civilizations across time and space. Thus, social and cultural differences manifest in temporal processes and contexts are the focal point of macrohistorical studies that, as Skocpol (1984, p. 1) summarizes:

(1) “address processes over time, and take temporal sequences seriously in accounting for outcomes,” (2) “attend to the interplay of meaningful actions and structural contexts in order to make sense of the unfolding of unintended as well as intended outcomes in individual lives and social transformations,” and (3) “highlight the particular and varying features of specific kinds of social structures and patterns of change.”

In the existing literature on macrosociological research, historical and comparative methods, with their focus on case studies devoted to understanding the nature and effects of large-scale structures and fundamental processes of change, have proven to be an effective approach to macrosociological explanations of macrosocial phenomena. While most historical and comparative research still involves qualitative analyses using available documents and records, more and more research attempts to employ both qualitative and quantitative approaches.

In advocating “moving beyond qualitative and quantitative strategies” Charles Ragin (1987) points out that, in macrosociological analysis, there are two basic strategies: the case-oriented strategy and the variable-oriented strategy. The former is very much evidence oriented, while the latter is theory centered. The goals of case-oriented investigation, with its extensive use of ideal types, often are both historically interpretive and causally analytic. “Investigators who used case-oriented strategies often want to understand or interpret specific cases because of their intrinsic values” (Ragin 1987, p. 35). Work by Bendix (1977, 1978) exemplifies this approach.

Unlike the case-oriented strategy, the variable-oriented strategy tests hypotheses derived from theory, often using quantitative techniques such as multivariate statistical analysis. In macrosocial analysis, a typical variable-oriented study “examines relationships between general features of social structures conceived as variables. Social units, such as nation-states, have structural features which interact in the sense that changes in some features produce changes in other features, which in turn may produce changes in others” (Ragin 1987, p. 55). For example, a cross-national study of modernization by Delacroix and Ragin (1978) is a typical example of variable-oriented research, and

this approach has remained quite popular in the study of development issues as well as in macro-level studies of organizations.

Each of these two strategies has its strengths and weaknesses. The case-oriented research enables investigators to comprehend diversity and address complexity by examining causal processes more directly in historical and comparative context. In variable-oriented research, by contrast, generality is given precedence over complexity when investigators are able to digest large numbers of cases. In some macrosociological studies, scholars combine the two approaches, as in Jeffrey Paige's *Agrarian Revolution* (1975).

### THE FUTURE OF MACROSOCIOLOGY

Macroscopic analysis of human society stands as a foundational area of research in sociology, and it is safe to predict that it will continue to grow and expand its scope of inquiry. In an increasingly global economy, marked by shifting boundaries and allegiances, and linked by rapidly advancing communications and information technology, there will be pressing need for explanation and analysis of the major historical and contemporary social movements and upheavals. Events of state formation, transition, and breakdown; revolution and devolution; conflicts based on gender, race, ethnicity, religion, region, and class; global movements of populations; and numerous other large-scale processes that increasingly mark the post-cold war era will provide the raw materials for scholarly and policy relevant analysis.

At the same time, in the interests of advancing social theory, sociologists will continue to seek ways to link macroprocesses to microprocesses. One of the perennial debates that surfaces among sociologists is whether macroprocesses or microprocesses have primacy in explaining social life. A variant revolves around the issue of whether microprocesses can be derived from macroprocesses or vice versa. Those who believe that the macro has causal priority risk being labeled structural determinists. Those who think that macrophenomena can be derived from microprocesses are dismissed as reductionists. Quite often an uneasy truce prevails in which practitioners of the two types of sociology go their own ways, with little interaction or mutual influence.

Despite pendulum swings that alternately emphasize one approach over the other, there are ongoing efforts to construct theory and conduct research built on genuine principles of micro-macro linkage. These have come from a variety of theoretical traditions and perspectives, including those with both macro and micro foundations. While many of these efforts ultimately result in de facto claims for theoretical primacy of one or the other approach, they nonetheless represent an interesting effort to create uniform and widely applicable sociological theory (Huber 1991).

Ultimately, most of the efforts to integrate micro and macro levels reflect the initial concerns of the theorist. For example, Randall Collins's efforts (1998) begin with a microfocus on interaction to derive macrophenomena, while neofunctionalist Jeffrey Alexander (1985) gives primacy to subjective forms of macrophenomena. Perhaps the most highly developed integrative effort is found in Anthony Giddens's theory of structuration (1984) in which social structure is defined as both constraining and enabling human activity, as well as being both internal and external to the actor.

The efforts to link microphenomena and macrophenomena are mirrored in a growing body of empirical research. Such work appears to follow Giddens's view of the constraining and enabling nature of social structure for human activity and the need to link structure and action. It appears safe to say that, while macrosociology will always remain a central component of sociological theory and research, increasing effort will be devoted to creating workable models that link it with its micro counterpart.

### REFERENCES

- Alexander, M. Jacqui, and Chandra Talpade Mohanty, eds. 1997 *Feminist Genealogies, Colonial Legacies, Democratic Futures*. New York and London: Routledge.
- Alexander, Jeffrey (ed.) 1985. *Neofunctionalism*. Beverly Hills, Calif.: Sage.
- 1998 *Neofunctionalism and After*. Malden, Mass.: Blackwell.
- Amin, Samir 1976 *Unequal Development: An Essay on the Social Formations of Peripheral Capitalism*. New York: Monthly Review Press.

- Aminzade, Ron 1984 "Capitalism Industrialization and Patterns of Industrial Protest: A Comparative Urban Study of Nineteenth-Century France." *American Sociological Review* 49:437–453.
- Beneria, Lourdes 1981 "Conceptualizing the Labour Force: The Underestimation of Women's Economic Activities." *Journal of Development* 17:10–28.
- Bendix, Reinhard 1977 *Nation-Building and Citizenship: Studies of Our Changing Social Order*. Berkeley: University of California Press.
- 1978 *Kings or People: Power and the Mandate to Rule*. Berkeley: University of California Press.
- Bluestone, Barry, and Bennet Harrison 1982 *The Deindustrialization of America: Plant Closings, Community Abandonment, and the Dismantling of Basic Industry*. New York: Basic Books.
- Boserup, Ester 1970 *Women's Role in Economic Development*. New York: St. Martin's Press.
- Burawoy, Michael 1979 *Manufacturing Consent: Changes in the Labor Process under Monopoly Capital*. Chicago: University of Chicago Press.
- Coleman, James S., Amitai Etzioni, and John Porter 1970 *Macrosociology: Research and Theory*. Boston: Allyn and Bacon.
- Collins, Randall 1986 "Future Decline of the Russian Empire." In Randall Collins, ed., *Weberian Sociological Theory*. New York: Cambridge University Press.
- 1995 "Prediction in Macrosociological Theory: The Case of the Soviet Collapse." *American Journal of Sociology* 100: 1552–1593.
- 1988 "The Micro Contribution to Macro Sociology." *Sociological Theory* 6:242–253.
- Delacroix, Jacques, and Charles Ragin 1978 "Modernizing Institutions, Mobilization, and Third World Development: A Cross-National Study." *American Journal of Sociology* 84:123–150.
- Edwards, Richard 1979 *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*. New York: Basic Books.
- Evans, Peter, and John Stephens 1988 "Development and the World Economy." In Neil Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- Foster, John 1974 *Class Struggle and the Industrial Revolution: Early Industrial Capitalism in Three English Towns*. London: Weidenfeld and Nicholson.
- Foucault, Michel 1979 *Discipline and Punish: The Birth of the Prison*. New York: Vintage.
- 1980 *The History of Sexuality, Volume 1: An Introduction*. New York: Vintage.
- 1985 *The History of Sexuality, Volume 2: The Uses of Pleasure*. New York: Pantheon.
- Frank, Andre Gunder 1967 *Capitalism and Underdevelopment in Latin America*. New York: Monthly Review Press.
- Giddens, Anthony 1984 *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley: University of California Press.
- Goldstone, Jack A. 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.
- Gordon, Milton 1964 *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
- Gutman, Herbert G. (1966) 1977 *Work, Culture, and Society in Industrializing America*. New York: Vintage Books.
- Habermas, Jurgen 1979 *Communication and the Evolution of Society*. Boston: Beacon Press.
- Harris, Marvin 1977 *Cannibals and Kings: The Origins of Cultures*. New York: Random House.
- Hartmann, Heidi 1981 "The Family as the Locus of Gender, Class, and Political Struggle: The Example of Housework." *Signs* 6:366–394.
- Hawley, Amos H. 1971 *Urban Society: An Ecological Approach*. New York: Wiley.
- Huber, Joan (ed.) 1991 *Macro-Micro Linkages in Sociology*. Newbury Park, Calif.: Sage.
- Kennedy, Paul 1987 *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House.
- Lenksi, Gerhard 1966 *Power and Privilege*. New York: McGraw-Hill.
- Lenksi, Gerhard, Patrick Nolan, and Jean Lenksi 1995 *Human Societies*, 7th ed. New York: McGraw-Hill.
- Mann, Michael 1986 *The Sources of Social Power, Vol. 1: A History of Power from the Beginning to A.D. 1760*. New York: Cambridge University Press.
- 1993 *The Sources of Social Power, Vol. 2: The Rise of Classes and Nation-States, 1760–1914*. Cambridge: Cambridge University Press.
- Luhmann, Niklas 1982 *The Differentiation of Society* (trans. S. Holmes and C. Larmore). New York: Columbia University Press.
- Markoff, John 1996 *The Abolition of Feudalism: Peasants, Lords and Legislators in the French Revolution*. University Park: Pennsylvania State University.
- Mingione, Enzo 1991 *Fragmented Societies: A Sociology of Economic Life beyond the Market Paradigm* (trans. Paul Goodrick). Oxford: Basil Blackwell.

- Moore, Barrington, Jr. 1966 *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Ollman, Bertell 1976 *Alienation*, 2nd ed. Cambridge: Cambridge University Press.
- Paige, Jeffrey 1975 *Agrarian Revolution: Social Movements and Export Agriculture in the Underdeveloped World*. New York: Free Press.
- Parsons, Talcott 1966 *Societies: Evolutionary and Comparative Perspectives*. Englewood Cliffs, N.J.: Prentice-Hall.
- Portes, Alejandro, Manuel Castells, and Lauren Benton, eds. 1989 *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore, Md.: Johns Hopkins University Press.
- Ragin, Charles C. 1987 *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ritzer, George 1988 *Sociological Theory*, 2nd ed. New York: Knopf.
- Rostow, W. W. 1960 *The Stages of Economic Growth*. Cambridge: Cambridge University Press.
- Shannon, Thomas Richard 1989 *An Introduction to the World System Perspective*. Boulder, Col.: Westview Press.
- Skocpol, Theda 1979 *States and Social Revolutions*. Cambridge: Cambridge University Press.
- (ed.) 1984 *Vision and Method in Historical Sociology*. New York: Cambridge University Press.
- Smelser, Neil 1964 "Toward a Theory of Modernization." Pp. 268–284 in Amitai Etzioni and Eva Etzioni, eds., *Social Change*. New York: Basic Books.
- 1988 "Social Structure." In Neil Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- Sorokin, Pitirim 1962 *Social and Cultural Dynamics*. New York: Bedminster Press.
- Thompson, E. P. 1963 *The Making of the English Working Class*. New York: Vintage.
- Tilly, Charles 1975 *The Formation of National States in Western Europe*. Princeton, N.J.: Princeton University Press.
- 1990 *Coercion, Capital, and the Rise of European States A.D. 990–1990*. Cambridge, Mass.: Blackwell.
- , Louise Tilly, and Richard Tilly 1975 *The Rebelious Century, 1830–1930*. Cambridge, Mass.: Harvard University Press.
- Tilly, Louise, and Joan Scott 1978 *Women, Work, and Family*. New York: Holt, Rinehart, and Winston.
- Wallerstein, Immanuel 1974 *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.
- 1980 *The Modern World-System II: Mercantilism and the Consolidation of the European World-Economy, 1600–1750*. New York: Academic Press.
- 1988 *The Modern World-System III: The Second Era of Great Expansion of the Capitalist World-Economy, 1730–1840*. San Diego, Calif.: Academic Press.
- Walton, John 1987 "Theory and Research on Industrialization." *Annual Review of Sociology* 13:89–108.
- Ward, Kathryn, ed. 1990 *Women Workers and Global Restructuring*. Ithaca, N.Y.: Industrial and Labor Relations Press.
- Wuthnow, Robert 1989 *Communities of Discourse: Ideology and Social Structure in the Reformation, the Enlightenment, and European Socialism*. Cambridge, Mass.: Harvard University Press.

ANN R. TICKAMYER AND JIELI LI

## MAJOR PERSONALITY THEORIES

Problems of definition arise with the terms *personality* and *personality theories*. Personality is understood by some people to mean self-concept; by others, the consensus of other people's opinions about one's character, and by others, one's true character. Some personality theories have elaborate coordinated concepts discussing how personality originates and develops from conception to senescence, taking up cognitive, conative, and affective aspects of the mind as well as free will, holism, philosophy, and other issues. On the other hand, there are relatively simple, one-dimensional theories of personality that pay little attention to what seems important to other theorists.

This topic is complicated not only by its complexity and variations but also by intellectual belligerence among those who favor one theory over another and those who differ about the same theory. The analogy to religions is inescapable.

In view of this situation, personality theories will be handled in an unusual way. Sentences in italics are reprinted from *Personality Theories, Research, and Assessment* (Corsini and Marsella 1983).



**Some Personality Theories and Their Originators**

Abelson, R.P.	Least effort	Lowen, Alexander	Bio-energetics
Allport, Gordon W.	Personalism	Maltz, Albert	Psychocybernetics
Angyll, Andreas	Organismic theory	Maslow, Abraham	Self-actualizations
Assiogoli, Roberto	Psychosynthesis	May, Rollo	Existentialism
Berne, Eric	Transactional analysis	Mead, G.H.	Social interaction
Binwangers, Ludwig	Daseinanalysis	Miller, Neal	Learning theory
Branden, Nathaniel	Biocentrism	Meyer, Adolf	Psychobiological theory
Burrow, Trigant	Phyloanalysis	Moreno, J.L.	Sociometry
Bühler, Charlotte	Humanistic psychology	Mowrer, O.H.	Two-factor theory
Bühler, Karl	Funktionlust	Murphy, Gardner	Biosocial theory
Boss, Medard	Daseinanalysis	Murray, H.A.	Need-press theory
Cattell, Raymond	Multivariate theory	Osgood, Charles	Congruity theory
Combs, Arthur	Phenomenology	Perls, Frederick	Gestalt theory
Ellis, Albert	Rational-emotive theory	Piaget, Jean	Developmental theory
Erikson, Erik	Developmental theory	Rank, Otto	Will theory
Eysenck, Hans	Developmental theory	Reich, Wilhelm	Character analysis
Frankl, Victor	Logotherapy	Rolf, Ida	Structural integration
Fromm, Erich	Humanistic psychoanalysis	Rotter, Julian	Social learning
Heider, Fritz	Balance theory	Sarbin, Theodore	Role theory
Horney, Karen	Sociopsychological theory	Sheldon, William	Morphological theory
Jackson, Don	Systems theory	Sullivan, H.S.	Interpersonal theory
Kelly, Charles	Neo-Reichian theory	Van Kaam, Adrian	Transpersonal psychology
Korsybski, Alfred	General semantics	Werner, Heinz	Developmental theory
Lecky, Philip	Self-consistency	Wolpe, Joseph	Behavior theory
Lewin, Kurt	Topological psychology		

Table 1

NOTE: Some Personality Theories and Their Originators

They contain quotes of selected assertions about the various theories written by authorities of nine major systems. Additional sources presenting comparative information on personality theories include: Burger 1993; Cloninger 1993; Corsini and Wedding 1995; Drapela 1995; Engler 1999; Ewen 1997; Schultz and Schultz 1994.

Table 1 is a list of a number of other important personality theories.

**PSYCHOANALYSIS (SIGMUND FREUD, 1856–1939)**

Psychoanalysis is both a theory of personality and a form of psychotherapy (see Freud 1952–1974). Highly controversial throughout Freud's lifetime, it continues to be so.

Freud saw personality as a dynamic conflict within the mind between opposing instinctual and social forces. *The topographical hypothesis views the mind in terms of three systems.* They are: the unconscious, the preconscious, and the conscious. *The mind is composed of the id, ego, and superego.* The id

consists of primitive instinctual demands, the superego represents society's influence restricting the id's demands, and the ego is dynamically in between the two. *Fundamental motives are instinctual.* Instincts are the basic forces (drives) of the psyche. The aim of drives is their satisfaction. *All instincts are basically sexual.* Freud's concept of sexuality was equivalent to physical pleasures. *There is a series of built-in stages of sexual development.* Freud postulated that people went through three sexual stages: An oral stage following the primary infantile narcissistic stage, then an anal phase, and finally a phallic phase. *Children develop libidinal attitudes towards parents.* This notion of the Oedipus and Electra complex of children having sexual attractions to parents of the opposite sex has especially generated controversy.

*The psyche develops a number of defenses.* To survive, the human being's ego develops a number of processes intended to repress awareness of conflicts. Repression is the main mental mechanism, but others defenses are related to it, including rationalization, displacement, identification

and conversion. *Dreams have meaning and purpose.* According to Freud, dreams are disguised desires permitting people to sleep by permitting expressions of illicit desires disguised by various symbolisms.

#### INDIVIDUAL PSYCHOLOGY (ALFRED ADLER, 1870–1937)

Alfred Adler's personality theory is distinguished by its common sense and simple language (see Adler 1956). In contrast to Freud and Jung, Adler's views demonstrate social concern.

*Man, like all forms of life, is a unified organism.* This basic holistic notion contradicts Freud's classifications and opposing theses and antitheses. Adler viewed the individual as an indivisible totality that could not be analyzed or considered in sections. *Life is movement, directed towards growth and expansion.* Adler took a dynamic and teleological attitude toward life, that people were always striving toward goals of personal self-improvement and enhancement. *Man is endowed with creativity and within limits is self-determined.* Instead of taking the usual position that only biology and society were to be considered in the formation of personality, Adler posited a third element: personal creativity or individual responsibility, akin to the concept of free will. Adler accepted that we all have certain biological and social givens and what is made of them is the responsibility of individuals.

*Man lives inextricably in a social world.* Adler had a social personality theory. *Individuale* in German does not have the same denotation as *individual* in English but rather denotes indivisibility or unity. Adler did not see humans apart from society. *The important life problems—human relations, sex, occupation—are social problems.* Adler believed that to be successful in life all humans had to complete the life tasks of socialization, family, and work.

*Social interest is an aptitude that must be consciously developed. Social interest is the criterion of mental health. Social interest is operationally defined as social usefulness.* This trio of related statements is an explicit philosophy unique for personality theories. Adler believed that psychological normality

depended on *Gemeinschaftsgefühl*—social interest. He saw all human failures, such as criminals, the insane, and neurotics, as lacking this element.

#### ANALYTICAL PSYCHOLOGY (CARL G. JUNG, 1875–1961)

Jung's analytical psychology stresses unconscious mental processes and features elements in personality that derive from mankind's past (Jung 1953–1972).

*Personality is influenced by potential activation of a collective transpersonal unconscious.* Jung believed that individuals upon conception came with something from the past that directed their personalities, a concept somewhat like Lamarckism relative to physical heredity. *Complexes are structured and energized around an archetypical image.* This is an extension of the first assertion. Complexes refer to important bipolar aspects of personality, such as introversion–extraversion. Complexes, directed by archetypes, are seen as innate and universal capacities of the mind to organize human experiences. Archetypes are considered innate potentials of the mind derived from the experiences of ancestors, a kind of directing blueprint of one's character.

*The ego mediates between the unconscious and the outside world.* According to Jung, a strong, well-integrated ego is the ideal state for a person. *Unconscious psychic reality is as important as the outside world.* Jung stressed the importance of phenomenology in contrast to overt behavior. He explored people's inner realms with great diligence. He even exceeded Freud in concentrating on the importance of the unconscious. *Personality growth occurs throughout the life cycle.* Jung saw individuals in constant growth and development with imperceptible stages that sometimes, as in the case of adolescence and midlife crises, became evident. *The psyche spontaneously strives towards wholeness, integration, and self-realization.* This last statement is echoed in many different ways by a number of other theorists, including the two just considered, and is made a central point by some theorists such as Carl Rogers and Kurt Goldstein.

### PERSON-CENTERED THEORY (CARL ROGERS 1902–1987)

Carl Rogers developed his theory as part of his system of client-centered or nondirective therapy (see Rogers 1951). He had a lifelong abiding faith in the potentials of people to correct the errors of their past if a therapeutic environment could be created in which the client felt understood and accepted by a neutral nonevaluative therapist. His system emerges from one central theme, the first assertion below.

*Each person has an inherent tendency to actualize unique potential.* Rogers viewed each person as having a built-in tendency to develop all his or her capacities in ways that serve to maintain or enhance the organism. *Each person has an inherent bodily wisdom which enables differentiation between experiences that actualize and those that do not actualize potential.* Rogers's trust in people is indicated here: There is a wisdom of the body in that everyone knows what is best for one's self in terms of the ultimate goal of self-realization.

*It is crucially important to be fully open to all experiences. Experiencing becomes more than bodily sensing as one grows older. Through complex interactions with our body and with other persons we develop a concept of self.* These three assertions belong together, and in them Rogers is taking up the nature-nurture, heredity-environment controversy. Essentially, his position is that personality is a function of bodily wisdom and the effect of others (primarily parents).

*One can sacrifice the wisdom of one's own experiences to gain another's love.* Rogers as a therapist came to the conclusion that a great deal of human suffering is due to the tendency of people to sacrifice their own body wisdom to gain positive regard from others. Children, in order to gain acceptance by their parents, will too often agree with them, accept their premises, and maintain them throughout life, generating problems thereby if the premises are incorrect. His therapeutic system was intended to get people to understand their historical processes and to be able to revise the history of their life. *A rift can develop between*

*what is actually experienced and the concept of self.* The same theme is here elaborated. A person may deny reality to gain approval from others, and this bifurcation can generate a host of problems. *When the rift between experiencing and self is too great, anxiety or disorganized behavior can result.* Once again, the same theme is emphasized. We all want to be loved and accepted, but the continued pursuit of acceptance may separate us from reality. *Validating experiencing in terms of others can never be completed. All maladjustments come about through denial of experiences discrepant with the self-concept.* And so, one must depend on one's self for reality and not on others. Adler believed that maladjusted people lacked social interest, while Rogers stated that maladjustment essentially came from people listening to others rather than to their own bodily wisdom.

### PERSONAL CONSTRUCTS THEORY (GEORGE A. KELLY, 1905–1967)

Kelly was a highly original thinker. He developed a unique cognitive system that called for the use of idiosyncratic language (see Kelly 1955). While his personal constructs theory covers all of psychology from the ideographic point of view, he bypassed usual terms and concepts such as *learning* and *emotions* and paid no attention to the environment or heredity.

*All our interpretations of the universe are subject to revision.* Kelly starts with a skepticism about beliefs and takes the position that there is no absolute reality. He took the position of constructive alternativism to indicate that people with differences of opinions could not necessarily be divided in terms of right and wrong. Two people can view the same situation in quite different ways and both can be right, both can be wrong, or one or the other may be right. *No person needs to be a victim of his own biography.* Here we have a statement of the freewill concept in a different form.

*A person's processes are psychologically channelized by the ways in which he anticipates events.* This is Kelly's fundamental postulate. Essentially, this viewpoint states that what is important is how events

are interpreted rather than the events themselves. This assertion leads naturally to Kelly's major contribution to personality theory, a series of other personal constructs, relative to how people view reality. We need not attempt to cover all of his constructs, but a few of them will give the reader a sense of Kelly's thinking: *A person anticipates events by construing their replication.* (The construction corollary.) *Persons differ from one another in their construction of events.* (The individuality corollary.) *A person may successively employ a variety of construction subsystems which are inferentially incompatible with each other.* (The fragmentation corollary.) This last corollary relates directly to Carl Rogers's theme that maladjustment comes from divergent forces: from within and from without.

*Many of the important processes of personality and behavior arise as a person attempts to change or is threatened with forced change in his construct system.* Kelly's point here is echoed by many other theorists, that one establishes some sort of life pattern or life-style, but changes in thinking about one's self and others will disrupt the individual.

Kelly's system is the purest cognitive system of any discussed here, solely dependent on perceptions and interpretations.

#### **OPERANT REINFORCEMENT THEORY (B.F. SKINNER 1904–1990)**

Skinner has denied that his operant reinforcement is a personality theory, but rather that it covers all aspects of overt human behavior (Skinner 1938). In contrast to those theorists who view personality as essentially phenomenological, Skinner decries the term mind and concerns himself solely with overt behavior. As a radical behaviorist, Skinner does not deny internal processes but considers them not relevant to psychology as an objective science of behavior.

*Personality is acquired and maintained through the use of positive and negative reinforcers.* Skinner applies operant reinforcement to all aspects of human behavior. We tend to repeat what works and to give up what does not work, to continue behavior that leads to pleasant consequences and

to discontinue behavior that leads to unpleasant consequences. *Behavior may be altered or weakened by the withholding of reinforcers.* If other people change their ways of operating towards an individual, this in turn will affect that person's behavior and consequently his personality.

*Personality develops through a process of discrimination.* In life, we experience all kinds of consequences, and we have to make decisions about our future behavior to these consequences. *Personality becomes shaped or differentiated.* Over time, our personalities are shaped by generalizations about ways that lead to the achievement of goals.

#### **SOCIAL LEARNING THEORY (ALBERT BANDURA, 1925–)**

Bandura, like Skinner, came to his opinions about personality mostly through research (Bandura and Walters 1963). His system is of the cognitive-learning type stressing the capacity of individuals to generalize in terms of symbols.

*The causes of human behavior are the reciprocal interaction of behavioral, cognitive, and environmental influences.* Bandura believes personality is a function of how we think and act and our responses of the environment's reactions to our behavior. In terms of the three elements of biology, society, and creativity, Bandura stresses the latter two. Heredity is discounted as a major determiner in personality development: How a person thinks and acts and how the environment responds to a person's behavior determines one's personality. *Behavior can be self-governed by means of self-produced consequences (self-reinforcement).* This assertion also emphasizes the importance of reciprocity: life is interaction: the individual versus the world, with the individual changing the world and the world changing the individual.

*Individuals may be influenced by symbols which act as models.* Reality to people need not only be direct stimuli, such as a smile or a slap, but reality can also be via symbols, such as pictures or words. Bandura's major research studies called for children to watch the behavior of others. He found that if a person considered to be a model acted in an aggressive

manner and got what he wanted, that observers were likely to imitate the model. Consequently, not only direct stimuli and responses (as per Skinner) but symbolic experiences also determine personality. *Reinforcements (and punishment) can operate in a vicarious manner.* This is more of the above. Various kinds of behavior can be changed by seeing what happens to others. We learn not only by doing and getting responses but also by observing.

### EXISTENTIAL PERSONALITY THEORY

Existential psychology is a loosely organized and ill-defined set of concepts mostly based on the work of philosophers and theologians (see Blackham 1959; Grimsley 1955). Essentially, existentialists see individuals as being in search of meaning. People are also seen as striving to achieve authenticity.

*Personality is primarily constructed through attribution of meaning.* Essentially, this point of view is similar to Kelly's concept of constructs. *Persons are characterized by symbolization, imagination, and judgment.* These are seen as attempts to find meaning. The human being is always trying to make sense out of existence, others, and self and uses mental processes in interaction with self and the world.

*Life is best understood as a series of decisions.* The human individual not only has to make evident decisions such as what to eat, but more subtle and important ones, such as who he or she really is. One has to decide what the world is like, what is real, what is important, and how to participate in the world. *Personality is a synthesis of facticity and possibility.* Facticity means the givens of heredity and environment and possibility becomes the creative aspect of personality. The facts of reality limit behavior variations.

*A person is always faced with the choice of the future, which provokes anxiety, and the choice of the past, which provokes guilt.* The human condition is such that people looking backwards in time can find reasons to be guilty and looking forward can find reasons to be afraid. Existentialists see anxiety and guilt as essential elements of the human being.

*Ideal development is facilitated by encouraging individuality.* Here we find traces of Carl Rogers's

concept of the importance of listening to one's own body or Adler's and Kelly's requirement for personal courage. A human problem is to escape the effects of one's early environment, especially the effects of one's family.

### CONSTITUTIONAL THEORIES

The oldest theories of personality formation are the constitutional that state that personality is a function of the nature of one's corporeal body. Aristotle (1910) in his *Physiognomica*, for example, stated that the "ancients" had a variety of theories to explain differences in human character. The Greek physician Galen took Hippocrates's physiological explanation of bodily health as a function of the balance between certain bodily fluids and stated that various personality types were a function of excesses of these fluids. Gall and Spurzheim (1809) extolled phrenology (the shape of the human head) in establishing personality. Kretschmer (1922) declared that people with certain kinds of body types tended to have particular types of mental conditions. Lombroso (1911) declared that criminal types were distinguished by a number of physiological anomalies. The list goes on and on. At present there are a variety of constitutional personality theories, some of which will be discussed below.

**Structural Approach.** William Sheldon (see Sheldon and Stevens 1942) classified individuals in terms of body shapes claiming that there was a positive correlation between various structural variations and personality types. He spent many years in doing basic research to find evidence for his theory. He found strong evidence to support the validity of his views. Other investigators also found supporting evidence but not to any useful degree.

*The somatotype provides a universal frame of reference for growth and development that is independent of culture.* This statement by implication discounts society and creativity. Born with a particular body type and you will have a specific personality type. *Three polar extremes called endomorphy, mesomorphy, and ectomorphy identify the essential components of the somatotype.* Sheldon had a somewhat complex classificatory system with three main body types:

mesomorphs had an excess of muscle, endomorphs an excess of fat, and ectomorphs were relatively thin. For example, mesomorphs were considered to be bold, endomorphs to be extraverted and ectomorphs to be introverted.

**Experiential Approach.** This particular constitutional position is championed by Schilder (1950) and Fisher (1970) among others. It is a combined learning/physiological approach, referring to the nature of the experiences that a person has via contact during life, between the inner viscera, the skin, and the environment's effect on the body.

*Body sensations provide the primary basis for initial differentiation of self from environment.* The basic notion is that an unborn infant is only aware of internal sensations, but following birth, now becomes aware of stimuli from the outside world. Thus, the body surface becomes the locus of separation of self from the environment and the child now becomes able to identify the self and the outer world. *The development of the body image proceeds through stages, each of which has a lasting effect upon the body image as a whole.* This assertion has elements of the Freudian sexual stages and of Skinner's behaviorism in that contact with the outside world not only establishes the world but also the individual's personality.

**Holistic Approach.** Kurt Goldstein, who worked primarily with brain-injured patients, is primarily identified with this viewpoint (see Goldstein 1939). In working with various cases of physical pathology, such as stroke victims, he came to the realization of the importance of a human's attempt to maximize and organize potentials to survive and to enhance one's situation.

*The normal human organism is equipped to maximize self-actualization, provided environmental forces do not interfere.* This statement is accepted in a variety of ways by a number of other personality theorists, but Goldstein made this his central point. Of those theorists already discussed, Adler, Jung, and Rogers would have agreed completely. *Self-actualization is manifested by maximum differentiation and by the highest possible level of complexity of an integrated system.* This statement follows from the

prior one and gives emphasis to the concept of the wisdom of the body. *The key to effective behavior is adequate functioning of part-whole relations.* Goldstein used Gestalt concepts of figure and ground to give evidence of the importance of understanding behavior as a totality, and consequently he can be considered an holistic theorist.

## SUMMARY

At present there are a considerable number of personality theories, each working as it were completely independently of one another. There is lack of a common vocabulary that in turn leads to different people saying the same thing in different words. A complete eclectic theory would consider all elements mentioned, taking up the issue of personality in terms of the issues of heredity, environment and creativity, self and the environment.

(SEE ALSO: *Personality and Social Structure; Social Psychology*).

## REFERENCES

- Adler, A. 1956 *The Individual Psychology of Alfred Adler*. H. L. Ansbacher and R. Ansbacher, eds., New York: Basic Books.
- Aristotle 1910 *Physiognomica*. Oxford: Oxford University Press.
- Bandura, A., and R. Walters 1963 *Social Learning Theory and Personality Development*. Englewood Cliffs, N.J.: Prentice-Hall.
- Blackham, G. W. 1959 *Six Existential Thinkers*. New York: Harper and Row.
- Burger, Jerry M. 1993 *Personality*. Pacific Grove, Calif. Brooks/Cole Publishing
- Cloninger, Susan C. 1993 *Theories of Personality: Understanding Persons*. Englewood Cliffs, N.J.: Prentice-Hall
- Corsini, R. J., and A. J. Marsella 1983 *Personality Theories, Research, and Assessment*. Itasca, Ill.: Peacock.
- Corsini, Raymond J. and Danny Wedding 1995 *Current Psychotherapies*, 5<sup>th</sup> ed. Itasca, Ill.: Peacock.
- Drapela, Victor J. 1995 *A Review of Personality Theories*, 2<sup>nd</sup> ed. Springfield Il.: Charles C. Thomas
- Ewen, Robert B. 1997 *An Introduction to Theories of Personality*, 5<sup>th</sup> ed. Hillsdale, N.J.: Erlbaum Associates.

- Fisher, S. 1970 *Body Experience in Fantasy and Behavior*. New York: Appleton-Century-Crofts.
- Freud, S. 1952–1974 *The Complete Psychological Works of Sigmund Freud*. London: Hogarth Press (24 volumes).
- Gall, F. J., and J. C. Spurzheim 1809 *Recherches sur la système nerveux*. Paris: Schoell.
- Goldstein, K. 1939 *The Organism*. New York: American Book Co.
- Grimsley, R. 1955 *Existentialist Thought*. Cardiff: University of Wales Press.
- Jung, C. G. 1953–1972 *The Collected Works of C. G. Jung*. Princeton, N.J.: Princeton University Press.
- Kelly, G. A. 1955 *The Psychology of Personal Constructs*. New York: W. W. Norton.
- Kretschmer, E. 1922 *Physique and Character*. London: Paul, Trench Trubner.
- Lombroso, C. 1911 *The Criminal Man*. Boston: Little, Brown.
- Monte, Christopher F. 1995 *Beneath the Mask: An Introduction to Theories of Personality*, 5<sup>th</sup> ed. Ft. Worth, TX: Harcourt Brace.
- Rogers, C. R. 1951 *Client-Centered Therapy*. Boston: Houghton Mifflin.
- Schilder, P. 1950 *The Image and Appearance of the Human Body*. New York: International Universities Press.
- Schultz, Duane P. and Sydney Ellen Schultz 1994 *Theories of Personality* 5<sup>th</sup> ed. Pacific Grove, Calif.: Brooks/Cole Publishing.
- Sheldon, W. H., and S. S. Stevens 1942 *Varieties of Human Temperament*. New York: Harper and Row.
- Skinner, B. F. 1938 *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

RAYMOND J. CORSINI

## MALTHUSIAN THEORY

See Demographic Transition; Human Ecology and Environmental Analysis.

## MARGINAL EMPLOYMENT

### INTRODUCTION

In industrialized countries the principal way in which adults and their families make ends meet

economically is through employment in the formal labor force. In the United States, earnings in the form of wages and salaries account for 71 percent of household income, on average. Even among individuals in poor families, half (49 percent) of all household income comes from earnings—their single most important source of income (Current Population Survey 1998). The centrality of work for economic well-being also is reflected in the close connection between finding gainful employment and exiting poverty on the one hand, and losing employment and sliding into poverty on the other. While formal employment is a necessary condition for economic well-being among almost all nonwealthy, able-bodied adults and their families, it is not a sufficient condition. Not all jobs pay well, many come without health and other benefits that workers must either pay for through their wages or do without, some are merely part-time, and some jobs are highly unstable and offer little long-term security. *Marginal employment can be conceptually defined as the circumstance in which the formal employment of adults (or groups of adults within families) fails to generate the earnings needed to achieve a minimally acceptable standard of living, either because they work too few hours (insufficient labor supply), and/or because their wages are too low.*

Two trends motivate increasing concern over marginal employment. The first is a shift in social welfare systems which, as in the United States, have come to place greater emphasis on formal employment as the route out of poverty and reliance on the state. In recent years industrialized nations have reoriented their social welfare programs to encourage employment and work preparedness (Cornea and Danziger 1997). In the United States, the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 has placed strict time limits on Public Assistance receipt. The image of thrusting people off the welfare rolls to fend for themselves in the open labor market has generated deep concern over their employment prospects. Second, in recent decades the structure of employment opportunities has changed in ways detrimental to low-skilled workers. Whether due to changing technology (e.g.,

computerization), which is displacing labor; to the globalization of production and trade, which is leading to outsourcing of low-skilled work; to competition through increased immigration; or to changes in wage-setting institutions (e.g., declines in unionization or real minimum wages), low-skilled workers have slipped further behind their better-skilled counterparts in terms of wages and total compensation (Freeman and Gottschalk 1998). The point is, just when many governments are foisting more responsibility for the economic well-being of citizens on the labor market and workers themselves, questions are being raised about the implications of macroeconomic changes for the quality of jobs available, especially for low-skilled groups. As such, marginal employment—its severity, causes, and consequences—has become an important concern for sociologists and policy makers alike.

#### MARGINAL EMPLOYMENT AS UNDEREMPLOYMENT

Marginal employment per se is not as yet a conventional concept in sociology and thus there are no standard empirical definitions of this term. As noted, we conceptualize it as consisting of both the degree of attachment to the labor force and the quality of jobs. A composite measure that incorporates both these dimensions is *underemployment*. Drawing on the Labor Utilization Framework of Phillip Hauser (1974), Clifford Clogg (1979) developed a measure of underemployment that consists of the following categories. The *working poor* are those people who are working full-time, full-year, but whose labor market earnings are less than 125 percent of the poverty threshold for single individuals. *Involuntary part-time workers* include those who are working fewer than full-time hours (thirty-five-hours per week) only because they are unable to find full-time work. The *unemployed* are those who are not currently working and are either actively looking for work or on layoff. (The unemployment rate—a standard aggregate measure of the health of the economy—is simply calculated as the number unemployed divided by the number in the labor force, where the latter is the sum of those employed and those unemployed.) Finally, *discouraged workers* (sometimes referred to as the

*subunemployed*) are neither working nor looking for work, but would nonetheless like to be working if they could find a job. All other adults, those who are not working and do not want to be, are considered *not in the labor force* and are excluded from analyses of underemployment. It is important to recognize that this operationalization of underemployment is an imperfect empirical measure of marginal employment as conceptually defined above. Discouraged or involuntary part-time workers, for example, could be in families that are otherwise well off economically. While this caveat needs to be borne in mind, the statistical patterns described below suggest underemployment is a reasonable measure of marginal employment.

Table 1 shows the distribution of underemployment and its types for the U.S. population as a whole, and for important demographic groups. The data are from the March 1998 Current Population Survey, a nationally representative survey of 64,659 U.S. households and a key source of official employment data (Current Population Survey 1998). In 1998, of all U.S. adults aged 18–64 who either were working or would like to have been working, 18.0 percent were underemployed. We stress that by this definition, over one in six U.S. workers or would-be workers are marginally employed. Of these, 4.7 percent were unemployed, 6.3 percent were working part-time but wanted full-time work, 6.1 percent were working full-time but for near- or below-poverty wages, and 0.8 percent had given up in the search for work, despite wanting a job.

Underemployment is not distributed equally across demographic groups, but parallels well-known correlates of economic well-being generally. At over one-third, the young (those aged 18–24) are far more likely to be underemployed than older adults. A more refined age breakdown for all adult age groups (not shown) suggests a curvilinear pattern, with young adults and elderly Americans at greater risk of underemployment than those in their prime working ages. Women have long been disadvantaged in the U.S. labor market with respect to wages and occupational prestige, so it is little surprise that there also is gender inequality in underemployment. In relative terms,



MARGINAL EMPLOYMENT

**Percentage distribution of underemployment by selected characteristics**

	Underemployed					Adequately Employed
	Total	Low income	Low hours	Unemployed	Discouraged	
<b>Total</b>	18.0%	6.1%	6.3%	4.7%	0.8%	82.0%
<b>Age</b>						
18-24	33.8	1.34	9.0	9.5	1.9	66.2
25-34	17.7	5.5	6.6	5.0	0.6	82.3
45-54	15.5	5.0	5.9	3.9	0.6	84.5
55-64	14.4	5.4	5.4	5.3	3.0	85.6
<b>Gender</b>						
Men	16.9	4.9	6.4	5.0	0.6	83.1
Women	19.2	7.5	6.2	4.5	1.0	80.8
<b>Race/ethnicity</b>						
Non-Hispanic white	15.2	5.4	5.6	3.7	0.6	84.8
Non-Hispanic black	25.3	7.1	7.5	8.8	1.9	74.7
Hispanic	28.3	10.2	10.0	7.1	1.1	71.7
Asian	17.5	6.5	6.2	3.9	1.0	82.5
Native American	19.4	9.1	8.8	9.8	1.7	70.6
<b>Marital status</b>						
Married	13.2	4.6	5.1	3.1	0.5	86.8
Divorced/separated	19.7	5.7	7.8	5.3	0.9	80.3
Widowed	23.1	8.4	9.3	4.1	1.3	76.9
Never married	17.1	9.6	8.1	8.0	1.5	72.9
<b>Immigrant status</b>						
Third (or higher) generation	17.0	5.7	6.0	4.6	0.8	83.0
Second generation	18.1	5.9	5.9	5.2	1.0	81.9
First generation	24.6	9.2	8.8	5.6	1.0	75.4
<b>Education</b>						
Less than high school	36.3%	12.2%	11.5%	10.5%	2.2%	63.7%
High school or GED	20.8	7.0	7.4	5.5	0.9	79.2
Some college	16.5	6.1	5.8	4.0	0.7	83.5
College degree or more	7.5	2.4	3.1	1.8	0.2	92.5
<b>Residence</b>						
Nonmetropolitan	21.6	8.2	7.4	5.3	0.8	78.4
Metropolitan	17.2	5.7	6.1	4.6	0.8	82.8
Central city	20.5	6.7	6.8	5.7	1.2	79.5
Non-central city	15.0	5.0	5.5	3.9	0.6	85.0
Not identified	18.3	5.9	6.8	5.0	0.6	81.7

Table 1

NOTE: Percentage Distribution of Underemployment by Selected Characteristics.

SOURCE: Original calculations from the March 1998 Current Population Survey.

women are 14 percent (2.3 percentage points) more likely than men to be underemployed, a difference that is due primarily to their greater risk of being working poor. Blacks, Hispanics and, to a lesser degree, Native Americans, are distinctly disadvantaged with respect to non-Hispanic whites. Indeed, about one-fifth of Native Americans, and over one-quarter of blacks and Hispanics, are underemployed, which compares to only 15 percent of non-Hispanic whites. Being married carries an advantage, while being divorced, separated, or widowed increases the risk of underemployment.

With respect to immigrant status, first-generation adults (immigrants themselves), are far more likely than the native-born to be underemployed, and among natives, those with a foreign-born parent (the second generation), are slightly disadvantaged compared to those with native-born parents (the third and higher generations). (This pattern should not be taken as direct evidence of economic assimilation, however, given immigrant-cohort differences in factors that determine the risk of underemployment.) The rising wage inequality of the 1980s and 1990s is due largely to deteriorating

conditions among low-skilled workers (Freeman and Gottschalk 1998). Indeed, as shown in Table 1, education is strongly associated with underemployment. Whereas 36.3 percent of those with less than a high school education are underemployed, only 7.5 percent of college graduates are similarly situated. Finally, the risk of underemployment is not distributed evenly across space. Despite legitimate and well-chronicled concern over declining employment opportunities in the inner city (Wilson 1996), at 21.6 percent the prevalence of underemployment is higher in nonmetropolitan (nonmetro) counties of the United States, than in metropolitan areas. Indeed, the nonmetro risk is even higher than that for those living in central cities of metro areas (20.5 percent). Thus, marginal employment is both an inner city and a rural problem.

### THE DYNAMICS OF MARGINAL EMPLOYMENT

Underemployment and its component parts are countercyclical. They rise in prevalence when the state of the wider economy declines, and they fall as the economy heats up. For example, above we noted the underemployment rate in 1998—a relatively good year economically—was 18.0 percent. In 1968, another good year, it stood at 15.2 percent. Amid recession in 1983, by comparison, 24.3 percent were underemployed. The rise and fall in the prevalence of marginal employment over time reflects the dynamic nature of labor markets. In large numbers individuals are constantly moving into and out of the labor force, between part-time and full-time work, and between good jobs and bad jobs.

While cross-sectional studies are useful for describing the broad contours of marginal employment and identifying sociodemographic groups that are more vulnerable, sociologists have begun to turn their attention to the nature and determinants of the movement between labor-force states. In an analysis of data from the Current Population Surveys of 1968 through 1993, Jensen and colleagues (1999) explored the determinants of transitions between adequate employment and underemployment. They found that adequately employed adults are more likely to slide into underemployment from one year to the next if they are black or Hispanic, female, less well educated, or young, or

if they live in a nonmetropolitan area. The risk of becoming underemployed also was higher amid a souring economy, and lower when the economy was strengthening. Once underemployed, these same groups were found to be disadvantaged with respect to their likelihood of finding adequate employment. Similarly, using data from the Survey of Income and Program Participation, Hsueh and Tienda (1996) explored transitions between being employed, unemployed, and out of the labor force altogether. They document greater labor-force instability among women, blacks and Hispanics, young adults, those with only a high school degree or less, and those who are not married.

That minorities and women, particularly those who are less well educated, have greater labor-force instability is important for understanding persisting gender and racial or ethnic inequality in society. The checkered work history that comes from frequent transitions into and out of the labor force means less opportunity to gain valuable work experience, and may signal to prospective employers a questionable commitment to work (Tienda and Stier 1996). The negative implications of work instability can accumulate over time, darkening future labor market outcomes and increasing the likelihood of marginal employment (Clogg et al. 1990).

### JOB DISPLACEMENT

Even a stable work history is no guarantee against future employment hardship. An important determinant of labor-force instability is job displacement. Displaced workers are often technically defined as “persons 20 years and older who lost or left a job [which they had held for three or more years] because their plant or company closed or moved, there was insufficient work for them to do, or their positions or shifts were abolished” (Hipple 1997, p. 38). Sometimes referred to as “structurally unemployed” (because their unemployment is due to structural rather than cyclical shifts in the economy), displaced workers have little prospect of getting their old jobs back: Their job loss is permanent. To get a sense of the magnitude of the

problem, consider that between 1993 and 1995, 15 percent of U.S. workers experienced job displacement (Kletzer 1998). The underlying causes of displacement are not well documented, but are likely rooted in technological change, rising foreign competition, and declining productivity within industries (Kletzer 1998). More is known about differences across sociodemographic groups in the risk of displacement. Recent evidence indicates that those who are black, less well educated, and employed in production occupations (as craftsmen, operatives, or laborers) and in manufacturing industries have higher rates of displacement (Kletzer 1998). Once displaced, there also is inequality in the prospects of finding new work after displacement. Here it helps to be white, male, and better educated.

#### CONTINGENT AND TEMPORARY WORKERS

Also contributing to volatility of labor-force attachment is the emergence of contingent labor and temporary workers. In the 1970s, international competition had begun to challenge U.S. industries, marking an end to two decades of economic expansion and growing corporate profits. In response U.S. firms began to restructure in order to maintain their competitive edge. By the mid-1980s, terms such as “mergers,” “acquisitions,” “deindustrialization,” and “downsizing” had become commonplace (Harrison and Bluestone 1988). In the process, the traditional relationship between employer and employee began to change. Today, for a growing percentage of U.S. workers the postwar model of relatively stable and secure full-time employment with fringe benefits, steady wages, and the opportunity, if not expectation, of internal advancement within a firm, is no longer a reality (Henson 1996). Instead an increasing proportion of the American workforce is now engaged in *contingent work*.

To date there is no conventional definition of contingent work. The term has been used broadly in the literature to refer to a range of nonstandard employment practices including “part-time work, work performed by independent contractors and

on-call workers, and work done by temporary workers, hired either directly for limited-duration projects or through temporary help firms” (Blank 1998, p. 258). In 1989, the Bureau of Labor Statistics (BLS) developed the following definition: “Contingent work is any job in which an individual does not have an explicit or implicit contract for long-term employment” (Polivka 1996, p. 3). In sum, despite disagreements over specifics, contingent work usually refers to nonstandard employment relationships that are associated with temporary or insecure employment duration, inadequate wages and benefits, and a conditional relationship that limits the attachment between employer and employee. Compared to permanent employees, contingent workers earn less and are less likely to be covered by employer health insurance or pension plans (Cohany et al. 1998; Hipple 1996). Depending upon the definition used, contingent workers composed between 2.2 and 16 percent of the labor force in 1995 (Barker and Christensen 1998).

The use of contingent workers has grown rapidly in recent decades. This growth has been driven primarily by a combination of employee desires (supply-side explanations) and employer strategies (demand-side explanations). On the supply side, contingent work often has features that some workers prefer, such as limited and flexible hours, independence, and variety. And some workers use contingent work as a way to shop around for good jobs and employers, in the hopes of obtaining more permanent work (Blank 1998). However, the idea that the rise in contingent work is being driven by worker preferences should not go too far since, by one estimate, two-thirds of contingent workers would prefer permanent work (Cohany et al. 1998). Demand-side factors include fluctuations in product demand which give rise to the need for flexibility in hiring and firing, reduced labor costs, and the ability to screen employees for permanent, full-time work (Blank 1998). Thus, in response to the growth of global competition and increasingly volatile consumer demand, employers have used contingent workers to shift the risk of market uncertainties from the firm to the worker. By reducing the obligation of the employer to the employee, firms are able to be

more flexible in adjusting labor inputs and reducing labor costs. One of the primary reasons that firms have been able to cut costs at the expense of labor is the transition in the balance of power between labor and management in recent decades (Golden and Applebaum 1992). As labor unions have gradually lost bargaining power, management has increasingly been able to exercise cost cutting strategies at the expense of workers.

A subgroup of contingent workers that appear to be particularly marginalized are temporary help workers or "temps." Temps are typically employees of temporary help supply (THS) firms, which provide workers to other businesses on a contractual basis (Ofstead 1998). In most cases, THS firms remain the worker's legal employer while the worker provides services to another establishment. THS firms emerged as labor market actors in the 1940s, but have seen tremendous growth in recent decades. Research by the Bureau of Labor Statistics shows that the number of temporary jobs filled daily by THS firms has increased more than two and a half times since 1982 (Golden and Applebaum 1992). Further, research has found that over 90 percent of employers use temporary workers and that more than five million workers seek employment through THS firms in a given year (Golden and Applebaum 1992). THS firms typically specialize in worker placement in one of four occupational areas: office, industrial, medical, or engineering and technical, with office workers comprising over half of all temporary workers (Carey and Hazelbaker 1986).

Temporary workers are subject to an earnings penalty in comparison to permanent workers. Research has shown that temporary workers typically earn between 80 and 52 percent of the wage earned by their permanent counterparts (Barker and Christensen 1998). Further, research shows that the temporary workforce is disproportionately composed of individuals that face disadvantage in the labor force in general: women, minorities, and the young (Barker and Christensen 1998; Howe 1986). The predominance of women and the young among temporary workers is likely due in part to the flexibility allowed by the employment relationship. However, limited alternative

employment opportunities and employer perceptions of what types of workers prove to be risky investments as permanent employees, also may be at play.

As it continues, the growth of contingent labor and proliferation of temporary workers will likely contribute significantly to the segmentation and bifurcation of the labor force. Permanent workers are able to realize higher wages, receive fringe benefits, receive on-the-job training, and have opportunities to climb the occupational ladder. On the other hand, contingent workers earn lower wages; receive few, if any, benefits; and enjoy little opportunity for advancement. In the case of those who engage in contingent employment because of personal preference, this may not present a problem. However, that many of those engaged in contingent work do so involuntarily is of concern. There is potential for the expansion of a fringe class of workers disenfranchised and marginalized from the mainstream labor force. At the very least, the growing use of contingent workers represents a fundamental shift in the relationship between employer and employee.

To summarize, marginal employment is a concept of increasing importance for sociologists. Structural economic shifts that are giving rise to job displacement and deterioration in the employment circumstances of low-skilled workers, combined with the increasing tendency for industrialized countries to link government help with work and/or work preparedness, have thrown a spotlight of concern on those most vulnerable to employment hardship. In addition to simply documenting the nature and trends in marginal employment, sociologists have begun to adopt more dynamic empirical approaches to model the movement of workers into and out of various states of marginal employment. All this work points to significant and persisting difficulties faced by women, racial and ethnic minorities, immigrants, young adults and elders, those with low achieved human capital, those who are unmarried, and rural or inner-city residents. Sociologists have also begun to explore the institutional changes that have contributed to the persistence of marginal employment and its unequal distribution across

sociodemographic groups. Key among these is the rise in contingent work which is characterized by part-time and temporary employment arrangements, often at lower wages and lacking in the fringe benefits enjoyed by permanent workers.

## REFERENCES

- Barker, Kathleen, and Kathleen Christensen 1998 *Contingent Work: American Employment Relations in Transition*. Ithaca, N.Y.: Cornell University Press.
- Blank, Rebecca M. 1998 "Contingent Work in a Changing Labor Market." In Richard B. Freeman and Peter Gottschalk, eds., *Generating Jobs: How to Increase Demand for Less-Skilled Workers*. New York: Russell Sage Foundation.
- Carey, Max L., and Kim L. Hazelbaker 1986 "Employment Growth in the Temporary Help Industry." *Monthly Labor Review* 109(4):37-43.
- Clogg, C. C. 1979 *Measuring Underemployment: Demographic Indicators for the United States*. New York: Academic.
- , Scott R. Eliason, and Robert Wahl 1990 "Labor Market Experiences and Labor Force Outcomes." *American Journal of Sociology* 95:1536-1576.
- Cohany, Sharon R., Steven F. Hipple, Thomas J. Nardone, Anne E. Polivka, and Jay C. Stewart 1998 "Counting the Workers: Results of a First Survey." In Kathleen Barker and Kathleen Christensen, eds., *Contingent Work: American Employment Relations in Transition*. Ithaca, N.Y.: Cornell University Press.
- Cornea, Andrea Giovanni, and Sheldon Danziger 1997 *Child Poverty and Deprivation in the Industrialized Countries, 1945-1995*. Oxford: Clarendon.
- Current Population Survey 1998 Conducted by the Bureau of the Census for the Bureau of Labor Statistics. Washington, D.C.: U.S. Bureau of the Census, March, 1998. (Machine-readable data file.)
- Freeman, Richard B., and Peter Gottschalk 1998 *Generating Jobs: How to Increase Demand for Less-Skilled Workers*. New York: Russell Sage Foundation.
- Golden, Lonnie, and Eileen Applebaum 1992 "What Was Driving the 1982-99 Boom in Temporary Employment?" *American Journal of Economics and Sociology* 51(4):473-493.
- Harrison, Bennett, and Barry Bluestone 1988 *The Great U-Turn: Corporate Restructuring and the Polarizing of America*. New York: Basic.
- Hauser, Philip M. 1974 "The Measurement of Labor Utilization." *Malayan Economic Review* 19:1-17.
- Henson, Kevin D. 1996 *Just a Temp*. Philadelphia: Temple University Press.
- Hipple, Steven F. 1996 "Earnings and Benefits of Contingent and Noncontingent Workers." *Monthly Labor Review* 119(10):22-30.
- 1997 "Worker Displacement in an Expanding Economy." *Monthly Labor Review* (December):26-39.
- Howe, Wayne J. 1986 "Temporary Help Workers: Who They Are, What Jobs They Hold." *Monthly Labor Review* 109(11):45-47.
- Hsueh, Sheri, and Marta Tienda 1996 "Gender, Ethnicity, and Labor Force Instability." *Social Science Research* 25:73-94.
- Jensen, Leif, Jill L. Findeis, Wan-Ling Hsu, and Jason P. Schachter 1999 "Slipping Into and Out of Underemployment: Another Disadvantage for Nonmetro Workers?" *Rural Sociology* 64(3):417-438.
- Kletzer, Lori G. 1998 "Job Displacement." *Journal of Economic Perspectives* 12(1):115-136.
- Ofstead, Cynthia 1998 *Temporary Help Firms and the Job Matching Process*. Unpublished doctoral dissertation. Department of Sociology, University of Wisconsin, Madison.
- Polivka, Anne E. 1996 "Contingent and Alternative Work Arrangements Defined." *Monthly Labor Review* 119(10):3-9.
- Tienda, Marta, and Haya Stier 1996 "Generating Labor Market Inequality: Employment Opportunities and the Accumulation of Disadvantage." *Social Problems* 43(2):147-165.
- Wilson, William J. 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.

LEIF JENSEN  
TIM SLACK

## MARITAL ADJUSTMENT

Marital adjustment has long been a popular topic in studies of the family, probably because the concept is believed to be closely related to the stability of a given marriage. Well-adjusted marriages are expected to last for a long time, while

poorly adjusted ones end in divorce. Simple as it seems, the notion of marital adjustment is difficult to conceptualize and difficult to measure through empirical research. After more than half a century of conceptualization about and research on marital adjustment, the best that can be said may be that there is disagreement among scholars about the concept, the term, and its value. In fact, several scientists have proposed abandoning entirely the concept of marital adjustment and its etymological relatives (Lively 1969; Donohue and Ryder 1982; Trost 1985).

### CONCEPTUAL ISSUES

Scientists have long been interested in understanding which factors contribute to success in marriage and which to failure. As early as the 1920s, Gilbert Hamilton (1929) conducted research on marital satisfaction by using thirteen clusters of questions. In 1939, Ernest Burgess and Leonard Cottrell published *Predicting Success or Failure in Marriage*, in which they systematically discussed marital adjustment. They defined *adjustment* as “the integration of the couple in a union in which the two personalities are not merely merged, or submerged, but interact to complement each other for mutual satisfaction and the achievement of common objectives” (p. 10).

Researchers have not agreed upon the use of any one term. To describe the seemingly same phenomenon, some have used the terms “marital quality,” “marital satisfaction,” and “marital happiness.” Robert Lewis and Graham Spanier have defined *marital quality* as “a subjective evaluation of a married couple’s relationship” (1979, p. 269)—a concept similar to that of “marital adjustment.” There have been numerous definitions of “marital adjustment” and “marital quality” (Spanier and Cole 1976), and it may not be fruitful to attempt to define the concept in a sentence or two. Rather, the following description of the factors that constitute marital adjustment or quality may prove more meaningful.

Since Burgess and Cottrell’s formulation, scientists have examined extensively the factors constituting marital adjustment. Although there has

been no consensus among researchers, factors constituting marital adjustment include agreement, cohesion, satisfaction, affection, and tension. Agreement between spouses on important matters is critical to a well-adjusted marriage. Though minor differences may broaden their perspectives, major differences between the spouses in matters such as philosophy of life, political orientations, and attitudes toward gender roles are detrimental to marital adjustment. In addition, agreement on specific decisions about family matters must be reached in good accord. *Marital cohesion* refers to both spouses’ commitment to the marriage and the companionship experienced in it. In a well adjusted marriage, both spouses try to make sure that their marriage will be successful. They also share common interests and joint activities. In a well-adjusted marriage, both spouses must be satisfied and happy with the marriage. Unhappy but long-lasting marriages are not well-adjusted ones. Spouses in well-adjusted marriages share affection, and it is demonstrated as affectionate behavior. Finally, the degree of tension in a well-adjusted marriage is minimal, and when tension arises it is resolved amicably, probably in discussion, and the level of tension and anxiety is usually low.

The core component of marital adjustment is marital satisfaction, and it has been extensively studied as a stand-alone concept. As such, it deserves separate consideration. *Marital satisfaction* has been defined as:

*The subjective feelings of happiness, satisfaction, and pleasure experienced by a spouse when considering all current aspects of his marriage. This variable is conceived as a continuum running from much satisfaction to much dissatisfaction. Marital satisfaction is clearly an attitudinal variable and, thus, is a property of individual spouses. (Hawkins 1968, p. 648).*

Again, scientists disagree about the definition. Some scholars conceptualize satisfaction rather as “the amount of congruence between the expectations a person has and the rewards the person actually receives” (Burr et al. 1979, p. 67). Because marital satisfaction is influenced not only by the

congruence between expectations and rewards but also by other factors, the former definition is broader than the latter and thus is adopted here.

Although Hawkins's definition of marital satisfaction subsumes happiness, marital happiness is usually considered a distinct variable. According to Campbell et al. (1976), happiness is similar to satisfaction, but these two qualities do differ in one important aspect:

*[A] term like "happiness" seems to evoke chiefly an absolute emotional state, whereas "satisfaction" implies a more cognitive judgment of a current situation laid against external standards of comparison such as "other people I know" or more private levels of aspiration. (p. 31)*

Since happiness (and marital happiness) denotes an emotional state, it has been known to be affected by the mood swing of the respondent. For that reason, this article does not specifically use the concept "marital happiness." Since marital happiness, marital satisfaction, marital quality, and marital adjustment are highly related to each other, interchangeable use of these terms is relatively common.

Although many scientists treat marital satisfaction as a factor of marital adjustment, there exist possibly major differences between these two concepts about the unit of analysis. Because satisfaction is a subjective property of an actor, there are two kinds of marital satisfaction in a marriage, the husband's and the wife's, and they are conceptually distinct. As Jessie Bernard (1972) stated, there are always two marriages in a family; the husband's marriage and the wife's marriage. Then, do these two marital satisfactions go hand in hand, or are they independent of each other? Research has produced mixed findings. In general, the more satisfied one spouse is with the marriage, the more satisfied is the other, but the correlation between the husband's and the wife's marital satisfactions is far from perfect (Spanier and Cole 1976). Marital adjustment or quality, on the other hand, can be either an individual or a dyadic property. When we say "a well-adjusted marriage," we refer to the dyad, while when we say, "She is well adjusted to

her marriage," we refer to the individual. No one has proposed valid measurement techniques for examining marital adjustment as a dyadic property, although some observational methods might be considered.

Another difference between marital satisfaction and marital adjustment is that while the former is a static product, the latter can be a dynamic process. In fact, marital adjustment is sometimes defined as a dynamic process, and marital satisfaction is listed as one of the outcomes of the adjustment process (Spanier and Cole 1976, pp. 127-128). It has also been proposed that marital adjustment be defined as a dynamic process and yet be measured as a state at a given point in time, a "snapshot" conception (Spanier and Cole 1976). Nevertheless, this connotation of dynamic process in the term "adjustment" has been criticized (Trost 1985) as a confusion of its meaning, because no measure of "adjustment" involves dynamic change, such as negotiation between the spouses.

Without agreeing on either which term to use or on the definition of such a term, researchers have tried for decades to measure marital adjustment, quality, or satisfaction. Burgess and Cottrell (1939) created one of the first measures of marital adjustment from twenty-seven questions pertaining to five subareas (agreement; common interests and joint activities; affection and mutual confidences; complaints; and feelings of being lonely, miserable, and irritable). Along with numerous attempts at measuring marital adjustment, Locke and Wallace (1959) modified Burgess and Cottrell's measure and called it the Marital Adjustment Test. Based on factor analysis, the test consists of fifteen questions ranging from the respondent's overall happiness in the marriage, the degree of agreement between the spouses in various matters, how they resolve conflicts, and the number of shared activities, to the fulfillment of their expectations about the marriage.

This measure was widely used until a new measure, the Dyadic Adjustment Scale, was proposed (Spanier 1976). It is composed of thirty-two questions and four subscales. The dyadic satisfaction subscale is composed of ten questions such as

“How often do you and your partner quarrel?” and “How often do you discuss or have you considered divorce, separation, or terminating your relationship?” The dyadic cohesion subscale is made up of five questions, “Do you and your mate engage in outside interests together?” and “How often would you say the following events occur between you and your mate?—have a stimulating exchange of ideas, laugh together, calmly discuss something, or work together on a project.” The dyadic consensus subscale is based on thirteen questions on “the extent of agreement or disagreement between you and your partner.” Items range from handling family finances, religious matters, and philosophy of life, to household tasks. Finally, the affectional expression subscale is composed of four questions related to affection and sex, two of which are agreement questions on demonstration of affection and sex relations.

All the above measures have been criticized as lacking a criterion against which the individual items are validated (Norton 1983; Fincham and Bradburn 1987). Some scholars have argued that only global and evaluative items, rather than content-specific and descriptive ones, should be included in marital adjustment or quality measures, because the conceptual domain of the latter is not clear. What constitutes a well-adjusted marriage may differ from one couple to another as well as cross-culturally and historically. Whether or not spouses kiss each other every day, for example, may be an indicator of a well-adjusted marriage in the contemporary United States but not in some other countries. Thus, marital adjustment or quality should be measured by the spouses’ evaluation of the marriage as a whole rather than by its specific components. Instead of “How often do you and your husband (wife) agree on religious matters?” (a content-specific description), it is argued that such questions as “All things considered, how satisfied are you with your marriage?” and “How satisfied are you with your husband (wife) as a spouse?” (a global evaluation) should be used. By the same reasoning, the Kansas Marital Satisfaction (KMS) scale has been proposed. This test includes only three questions: “How satisfied are you with your (a) marriage: (b) husband (wife)

as a spouse: and (c) relationship with your husband (wife)?” (Schumm et al. 1986).

Traditional indexes also have been criticized for their lack of theoretical basis and the imposition of what constitutes a “successful marriage.” On the basis of exchange theory, Ronald Sabatelli (1984) developed the Marital Comparison Level Index (MCLI), which measures marital satisfaction by the degree to which respondents feel that the outcomes derived from their marriages compare with their expectations. Thirty-six items pertaining to such aspects of marriage as affection, commitment, fairness, and agreement were originally included, and thirty-two items were retained in the final measure. Because this measure is embedded in the tradition of exchange theory, it has strength in its validity.

### PREDICTING MARITAL ADJUSTMENT

How is the marital adjustment of a given couple predicted? According to Lewis and Spanier’s (1979) comprehensive work, three major factors predict marital quality; social and personal resources, satisfaction with lifestyle, and rewards from spousal interaction.

In general, the more social and personal resources a husband and wife have, the better adjusted their marriage is. Material and nonmaterial properties of the spouses enhance their marital adjustment. Examples include emotional and physical health, socioeconomic resources such as education and social class, personal resources such as interpersonal skills and positive self-concepts, and knowledge they had of each other before getting married. It was also found that good relationships with and support from parents, friends, and significant others contribute to a well-adjusted marriage. Findings that spouses with similar racial, religious, or socioeconomic backgrounds are better adjusted to their marriages are synthesized by this general proposition.

The second major factor in predicting marital adjustment is satisfaction with lifestyle. It has been found that material resources such as family income positively affect both spouses’ marital adjustment. Both the husband’s and the wife’s satisfaction with their jobs enhances better-adjusted



marriages. Furthermore, the husband's satisfaction with his wife's work status also affects marital adjustment. The wife's employment itself has been found both instrumental and detrimental to the husbands' marital satisfaction (Fendrich 1984). This is because the effect of the wife's employment is mediated by both spouses' attitudes toward her employment. When the wife is in the labor force, and her husband supports it, marital adjustment can be enhanced. On the other hand, if the wife is unwilling to be employed, or is employed against her husband's wishes, this can negatively affect their marital adjustment. Marital adjustment is also affected by the spouses' satisfaction with their household composition, by how well the couple is embedded in the community, and the respondent's health (Booth and Johnson 1994).

Parents' marital satisfaction was found to be a function of the presence, density, and ages of children (Rollins and Galligan 1978). Spouses (particularly wives) who had children were less satisfied with their marriages, particularly when many children were born soon after marriage at short intervals. The generally negative effects of children on marital satisfaction and marital adjustment could be synthesized under this more general proposition about satisfaction with lifestyle.

It has been consistently found that marital satisfaction plotted against the couple's family life-cycle stages forms a U-shaped curve (Rollins and Cannon 1974; Vaillant and Vaillant 1993). Both spouses' marital satisfaction is quite high right after they marry, hits the lowest point when they have school-age children, and gradually bounces back after all children leave home. To illustrate this pattern, publicly available data from the National Survey of Families and Households (collected in 1987 and 1988) are analyzed here. Figure 1 shows the result. While husbands' average marital satisfaction hit the lowest point when their oldest child was between 3 and 5 years old, wives satisfaction was lowest when their oldest child was between 6 and 12 years old.

This pattern has been interpreted as a result of role strain or role conflict between the spousal, parental, and work roles of the spouses. Unlike

right after the marriage and the empty-nest stages, having children at home imposes the demand of being a parent in addition to being a husband or wife and a worker. When limited time and energy cause these roles to conflict with each other, the spouses feel strain, which results in poor marital adjustment (Lavee et al. 1996). Along this line of reasoning, Wesley Burr et al. (1979) proposed that marital satisfaction is influenced by the qualities of the individual's role enactment as a spouse and of the spouse's role enactment. They argue further, from the symbolic-interactionist perspective, that the relationship between marital role enactment and marital satisfaction is mediated by the importance placed on spousal role expectations.

As seen above, the concept of family life cycle seems to have some explanatory power for marital adjustment. Researchers and theorists have found, however, that family life cycle is multidimensional and conceptually unclear. Once a relationship between a particular stage in the family life cycle and marital adjustment is identified, further variables must be added to explain that relationship—variables such as the wife's employment status, disposable income, and role strain between spousal and parental roles (Crohan 1996; Schumm and Bugaighis 1986). Furthermore, the proportion of variance in marital adjustment "explained" by the family's position in its life cycle is small, typically less than 10 percent (Rollins and Cannon 1974). In the case of our analysis above, it is only 3 percent for both husbands and wives. Thus, some scholars conclude that family life cycle has no more explanatory value than does marriage or age cohort (Spanier and Lewis 1980).

The last major factor in predicting marital adjustment is the reward obtained from spousal interaction. On the basis of exchange theory, Lewis and Spanier summarize past findings that "the greater the rewards from spousal interaction, the greater the marital quality" (Lewis and Spanier 1979, p. 282). Rewards from spousal interaction include value consensus; a positive evaluation of oneself by the spouse; and one's positive regard for things such as the physical, mental, and sexual attractiveness of the spouse. Other rewards from spousal interaction include such aspects of emotional gratification as the expression of affection;

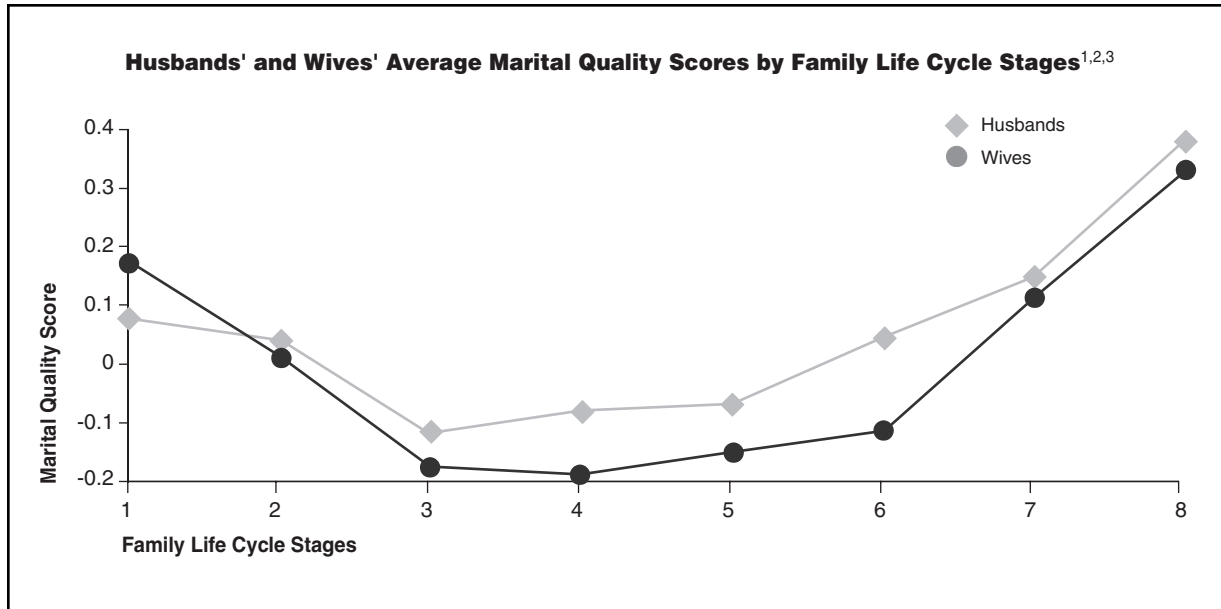


Figure 1. Husbands' and Wives' Average Marital Quality Scores by Family Life Cycle Stages<sup>1,2,3</sup>

NOTE: 1. See Sweet, Bumpass, and Call (1988) for the structure and data of the survey. Included in this figure are first-time married couples with at least one child or no child but married for less than 10 years.

2. Family life cycles stages are defined as follows (see Duvall 1977):

Stage 1 (Married Couples): Married for less than 10 years with no children

Stage 2 (Childbearing Families): The oldest child younger than 30 months

Stage 3 (Families with Preschool Children): The oldest child younger than 6

Stage 4 (Families with School Children): The oldest child younger than 13

Stage 5 (Families with Teenagers): The oldest child younger than 20

Stage 6 (Families Launching Young Adults): The oldest child 20 or older

Stage 7 (Middle-Aged Parents): The youngest child 20 or older

Stage 8 (Aging Families): The youngest child 20 or older and one of the spouses is 60 or older

3. Marital quality is measured by three questions: "Taking things all together, how would you describe your marriage? Very happy (7) to Very unhappy (1)." "Do you feel that your marriage might be in trouble right now? Yes (2) or No (1)." "What do you think the chances are that you and your husband/wife will eventually separate or divorce? Very low (5) to Very high (1)." The last two items are rescaled in such a way that the three items are given the same weight and the mean is set to 0. The final score is the average of these three scores.

respect and encouragement between the spouses; love and sexual gratification; and egalitarian relationships. Married couples with effective communication, expressed in self-disclosure, frequent successful communication, and understanding and empathy, are better adjusted to their marriages (Erickson 1993). Complementarity in the spouses' roles and needs, similarity in personality traits, and sexual compatibility all enhance marital adjustment. Finally, frequent interaction between the spouses leads to a well-adjusted marriage. The lack of spousal conflict or tensions should be added to the list of rewards from spousal interactions.

In this context, it is interesting to compare the average marital quality between men and women (Figure 1). The husbands' average marital quality is higher than the wives' *except* before they have their first child. The wife's marital quality is initially higher than the husband's, but it decreases after the arrival of her first baby. After that, it tends to be lower than that of her husband during the entire span of their marriage. Given that women perform most of the child care and household work, the steep decline in the average marital quality for women from Stage 1 to Stage 3 can be interpreted as a result of this burden of household

work and child care. Women's rewards from marriage are lower than those of men, on the average, and this differential appears as a gender difference in marital adjustment over the family life-cycle stages.

Symbolic interactionists also argue that relative deprivation of the spouses affects their marital satisfaction: If, after considering all aspects of the marriage, spouses believe themselves to be as well off as their reference group, they will be satisfied with their marriages. If they think they are better off or worse off than others who are married, they will be more or less satisfied with their marriages, respectively (Burr et al. 1979).

### **CONSEQUENCES OF MARITAL ADJUSTMENT: PERSONAL ADJUSTMENT AND MARITAL STABILITY**

It has been widely shown that married persons tend to be better adjusted in their lives than either never-married, separated, divorced, or widowed persons. This seems true not only in the area of psychological adjustments such as depression and general life satisfaction, but also in the area of physical health. Married people are more likely to be healthy and to live longer. Two factors should be considered to account for this relationship. First, psychologically and physically well-adjusted persons are more likely to get married and stay married. Second, the favorable socioeconomic status of married persons may explain some of this relationship. Nevertheless, scholars generally agree that marriage has a positive effect on personal adjustment, in both psychological and physical aspects.

If marriages in general affect personal adjustment in a positive fashion, it is likely that well-adjusted marriages lead to well-adjusted lives. Past research shows just this, though the findings should be cautiously interpreted. Some people tend to favorably answer "adjustment" questions, whether the questions are about their marriages, their personal lives in general, or their subjective health. The apparent positive relationship may be spurious. Nevertheless, if the psychological adjustment

is a composite of the adjustments in various aspects of life (i.e., marriage, family, work, health, friendship, etc.), high marital adjustment should lead to high psychological adjustment. In addition, positive effects of well-adjusted marriages on physical health may be accounted for, in part, by psychosomatic aspects of physical health.

This relationship provides an important policy implication of marital adjustment. Well-adjusted marriages may reduce health service costs, involving both mental and physical health. This is in addition to the more obvious reduction in social service costs derived from unstable and/or unhappy marriages. Children of divorce who need special care and domestic violence are just two examples through which poorly adjusted marriages become problematic and incur social services expenses.

Does marital adjustment affect the stability of a marriage? Does a better-adjusted marriage last longer than a poorly adjusted one? The answer is generally yes, but this is not always the case. Some well-adjusted marriages end in divorce, and many poorly adjusted marriages endure. As for the latter, John Cuber and Peggy Harroff conducted research on people whose marriages "lasted ten years or more and who said that they have never seriously considered divorce or separation" (1968, p. 43). They claim that not all the spouses in these marriages are happy and that there are five types of long-lasting marriages. In a "conflict-habituated marriage," the husband and the wife always quarrel. In a "passive-congenial marriage," the husband and the wife take each other for granted without zest, while "devitalized marriages" started as loving but have degenerated to passive-congenial marriages. In a "vital marriage," spouses enjoy together such things as hobbies, careers, or community services, while in a "total marriage," spouses do almost everything together. It should be noted that even conflict-habituated or devitalized marriages can last as long as vital or total marriages. For people in passive-congenial marriages, the conception and the reality of marriage are devoid of romance and are different from other people's.

What then determines the stability of marriage and how the marital adjustment affects it? It

is proposed that although marital adjustment leads to marital stability, two factors intervene; alternative attractions and external pressures to remain married (Lewis and Spanier 1979). People who have both real and perceived alternatives to poorly adjusted marriages—other romantic relationships or successful careers—may choose divorce. A person in a poorly adjusted marriage may remain in it if there is no viable alternative, if a divorce is unaffordable or would bring an intolerable stigma, or if the person is exceptionally tolerant of conflict and disharmony in the marriage. Nevertheless, it should be emphasized that even though marital stability is affected by alternative attractions and external pressures, marital adjustment is the single most important factor in predicting marital stability. Lack of large-scale longitudinal data and adequate statistical technique have hampered scholars' efforts to establish this link between marital adjustment and stability. Given recent availability of longitudinal and technological development, this area of research holds a high promise.

(SEE ALSO: *Divorce; Family Roles; Interpersonal Attraction; Marriage*)

## REFERENCES

- Bernard, Jessie 1972 *The Future of Marriage*. New York: World Publishing.
- Booth, Alan, and David R. Johnson 1994 "Declining Health and Marital Quality." *Journal of Marriage and the Family* 56:218-223.
- Burgess, Ernest W., and Leonard Cottrell, Jr. 1939 *Predicting Success or Failure in Marriage*. New York: Prentice-Hall.
- Burr, Wesley R., Geoffrey K. Leigh, Randall D. Day, and John Constantine 1979 "Symbolic Interaction and the Family." In W. Burr, R. Hill, F. I. Nye, and I. Reiss, eds., *Contemporary Theories about the Family*. New York: Free Press.
- Campbell, Angus, Philip E. Converse, and Willard L. Rodgers 1976 *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*. New York: Russell Sage Foundation.
- Crohan, Susan E. 1996 "Marital Quality and Conflict Across the Transition to Parenthood in African American and White Couples." *Journal of Marriage and the Family* 58:933-944.
- Cuber, John F., and Peggy B. Harroff 1968 *The Significant Americans: A Study of Sexual Behavior among the Affluent*. Baltimore: Penguin Books.
- Donohue, Kevin C., and Robert G. Ryder 1982 "A Methodological Note on Marital Satisfaction and Social Variables." *Journal of Marriage and the Family* 44:743-747.
- Duvall, Evelyn M. 1977 *Marriage and Family Development*, 5th ed. Philadelphia: Lippincott.
- Erickson, Rebecca J. 1993 "Reconceptualizing Family Work: The Effect of Emotion Work on Perception of Marital Quality." *Journal of Marriage and the Family* 55:888-900.
- Fendrich, Michael 1984 "Wives' Employment and Husbands' Distress: A Meta-Analysis and a Replication." *Journal of Marriage and the Family* 46:871-879.
- Fincham, Frank D., and Thomas N. Bradbury 1987 "The Assessment of Marital Quality: A Reevaluation." *Journal of Marriage and the Family* 49:797-809.
- Hamilton, Gilbert V. 1929 *A Research in Marriage*. New York: A. and C. Boni.
- Hawkins, James L. 1968 "Associations Between Companionship, Hostility, and Marital Satisfaction." *Journal of Marriage and the Family* 30:647-650.
- Lavee, Yoav, Shlomo Sharlin, and Ruth Katz 1996 "The Effect of Parenting Stress on Marital Quality: An Integrated Mother-Father Model." *Journal of Family Issues* 17:114-135.
- Lewis, Robert A., and Graham B. Spanier 1979 "Theorizing about the Quality and Stability of Marriage." In W. Burr, R. Hill, F. I. Nye, and I. Reiss, eds., *Contemporary Theories About the Family*. New York: Free Press.
- Lively, Edwin L. 1969 "Toward Concept Clarification: The Case of Marital Interaction." *Journal of Marriage and the Family* 31:108-114.
- Locke, Harvey J., and Karl M. Wallace 1959 "Short Marital-Adjustment and Prediction Tests: Their Reliability and Validity." *Marriage and Family Living* 21:251-255.
- Norton, Robert 1983 "Measuring Marital Quality: A Critical Look at the Dependent Variable." *Journal of Marriage and the Family* 45:141-151.
- Rollins, Boyd C., and Kenneth L. Cannon 1974 "Marital Satisfaction over the Family Life Cycle: A Reevaluation." *Journal of Marriage and the Family* 36:271-282.
- , and Richard Galligan 1978 "The Developing Child and Marital Satisfaction of Parents." In R. M.

- Lerner and G. B. Spanier, eds., *Child Influences on Marital and Family Interaction*. New York: Academic Press.
- Sabatelli, Ronald M. 1984 "The Marital Comparison Level Index: A Measure for Assessing Outcomes Relative to Expectations." *Journal of Marriage and the Family* 46:651-662.
- Schumm, Walter R., and Margaret A. Bugaighis 1986 "Marital Quality over the Marital Career: Alternative Explanations." *Journal of Marriage and the Family* 48:165-168.
- Schumm, Walter R., Lois A. Paff-Bergen, Ruth C. Hatch, Felix C. Obiorah, Janette M. Copeland, Lori D. Meens, and Margaret A. Bugaighis 1986 "Concurrent and Discriminant Validity of the Kansas Marital Satisfaction Scale." *Journal of Marriage and the Family* 48:381-387.
- Spanier, Graham B. 1976 "Measuring Dyadic Adjustment: New Scales for Assessing the Quality of Marriage and Similar Dyads." *Journal of Marriage and the Family* 38:15-28.
- , and Charles L. Cole 1976 "Toward Clarification and Investigation of Marital Adjustment." *International Journal of Sociology and the Family* 6:121-146.
- Spanier, Graham B., and Robert A. Lewis 1980 "Marital Quality: A Review of the Seventies." *Journal of Marriage and the Family* 42:825-839.
- Sweet, James A., Larry L. Bumpass, and Vaughn R. A. Call 1988 *A National Survey of Families and Households*. Madison: Center for Demography and Ecology, University of Wisconsin.
- Trost, Jan E. 1985 "Abandon Adjustment!" *Journal of Marriage and the Family* 47:1072-1073.
- Vaillant, Caroline O., and George E. Vaillant 1993 "Is the U-Curve of Marital Satisfaction an Illusion? A 40-Year Study of Marriage." *Journal of Marriage and the Family* 55:230-239.

YOSHINORI KAMO

## MARRIAGE

The current low rates of marriage and remarriage and the high incidence of divorce in the United States are the bases of deep concern about the future of marriage and the family. Some have used these data to argue the demise of the family in

American Society (Popenoe 1993). Others see such changes as normal shifts and adjustments to societal changes (Barich and Bielby 1996). Whatever the forecast, there is no question that the institution of marriage is currently less stable than it has been in previous generations. This article explores the nature of modern marriage and considers some of the reasons for its vulnerability.

Marriage can be conceptualized in three ways: as an institution (a set of patterned, repeated, expected behaviors and relationships that are organized and endure over time); as rite or ritual (whereby the married status is achieved); and as a process (a phenomenon marked by gradual changes that lead to ultimate dissolution through separation, divorce, or death). In the discussion that follows we examine each of these conceptualizations of marriage, giving the greatest attention to marriage as a process.

### MARRIAGE AS INSTITUTION

From a societal level of analysis the institution of marriage represents all the behaviors, norms, roles, expectations, and values that are associated with the legal union of a man and woman. It is the institution in society in which a man and woman are joined in a special kind of social and legal dependence to found and maintain a family. For most people, getting married and having children are the principal life events that mark the passage into mature adulthood. Marriage is considered to represent a lifelong commitment by two people to each other and is signified by a contract sanctioned by the state (and for many people, by God). It thus involves legal rights, responsibilities, and duties that are enforced by both secular and sacred laws. As a legal contract ratified by the state, marriage can be dissolved only with state permission.

Marriage is at the center of the kinship system. New spouses are tied inextricably to members of the kin network. The nature of these ties or obligations differs in different cultures. In many societies almost all social relationships are based on or mediated by kin, who may also serve as allies in times of danger, may be responsible for the transference of property, or may be turned to in times

of economic hardship (Lee 1982). In the United States, kin responsibilities rarely extend beyond the nuclear family (parents and children). There is the possible exception of caring for elderly parents, where norms seem to be developing (Eggebeen and Davey 1998). There are no normative obligations an individual is expected to fulfill for sisters or brothers, not to mention uncles, aunts, and cousins. Associated with few obligations and responsibilities is greater autonomy and independence from one's kin.

In most societies the distribution of power in marriage is given through tradition and law to the male—that is, patriarchy is the rule as well as the practice. For many contemporary Americans the ideal is to develop an egalitarian power structure, but a number of underlying conditions discourage attaining this goal. These deterrents include the tendency for males to have greater income; higher-status jobs; and, until recently, higher educational levels than women. In addition, the tradition that women have primary responsibility for child rearing tends to increase their dependency on males.

Historically, the institution of marriage has fulfilled several unique functions for the larger society. It has served as an economic alliance between two families, as the means for legitimizing sexual relations, and as the basis for legitimizing parenthood and offspring. In present-day America the primary functions of marriage appear to be limited to the legitimization of parenthood (Davis 1949; Reiss and Lee 1988) and the nurturance of family members (Lasch 1977). Recently, standards have changed and sexual relationships outside marriage have become increasingly accepted for unmarried people. Most services that were once performed by members of a family for other members can now be purchased in the marketplace, and other social institutions have taken over roles that once were assigned primarily to the family. Even illegitimacy is not as negatively sanctioned as in the past. The fact that marriage no longer serves all the unique functions it once did is one reason some scholars have questioned the vitality of the institution.

### MARRIAGE AS RITE OR RITUAL

Not a great deal of sociological attention has been given to the study of marriage as a rite or ritual that

transfers status. Philip Slater, in a seminal piece published in 1963, discussed the significance of the marriage ceremony as a social mechanism that underscores the dependency of the married couple and links the new spouses to the larger social group. Slater claims that various elements associated with the wedding (e.g., bridal shower, bachelor party) help create the impression that the couple is indebted to their peers and family members who organize these events. He writes,

*[F]amily and friends [are] vying with one another in claiming responsibility for having "brought them together" in the first place. This impression of societal initiative is augmented by the fact that the bride's father "gives the bride away." The retention of this ancient custom in modern times serves explicitly to deny the possibility that the couple might unite quite on their own. In other words, the marriage ritual is designed to make it appear as if somehow the idea of the dyadic union sprang from the community, and not from the dyad itself. (p. 355)*

Slater describes the ways in which rite and ceremony focus attention on loyalties and obligations owed others: "The ceremony has the effect of concentrating the attention of both individuals on every OTHER affectional tie either one has ever contracted" (Slater 1963, p. 354). The intrusion of the community into the couple's relationship at the moment of unity serves to inhibit husband and wife from withdrawing completely into an intimate unit isolated from (and hence not contributing to) the larger social group.

Martin Whyte (1990) noted the lack of information on marriage rituals and conducted a study to help fill this gap. He found that, since 1925, wedding rituals (bridal shower, bachelor party, honeymoon, wedding reception, church wedding) have not only persisted but also increased in terms of the number of people who incorporate them into their wedding plans. Weddings also are larger in scale in terms of cost, number of guests, whether a reception is held, and so on. Like Slater, Whyte links marriage rituals to the larger social fabric and argues that an elaborate wedding serves several functions. It

*serves notice that the couple is entering into a new set of roles and obligations associated with marriage, it mobilizes community support behind their new status, it enables the families involved to display their status to the surrounding community, and it makes it easier for newly marrying couples to establish an independent household. (p. 63)*

### MARRIAGE AS PROCESS

Of the three ways in which marriage is conceptualized—institution, rite or ritual, and process—most scholarly attention has focused on process. Here the emphasis is on the interpersonal relationship. Changes in this relationship over the course of a marriage have attracted the interest of most investigators. Key issues studied by researchers include the establishment of communication, affection, power, and decision-making patterns; development of a marital division of labor; and learning spousal roles. The conditions under which these develop and change (e.g., social class level, age at marriage, presence of children) and the outcomes of being married that derive from them (e.g., degree of satisfaction with the relationship) are also studied. For illustrative purposes, the remainder of this article will highlight one of these components, marital communication, and one outcome variable, marital quality. We also address different experiences of marriage based on sex of spouse: “his” and “her” marriage.

**The Process of Communication.** The perception of a “failure to communicate” is a problem that prompts many spouses to seek marital counseling. The ability to share feelings, thoughts, and information is a measure of the degree of intimacy between two people, and frustration follows from an inability or an unwillingness to talk and listen (Okun 1991). However, when the quality of communication is high, marital satisfaction and happiness also are high (Holman and Brock 1986; Burlinson and Denton 1997; Gottman 1994).

The role of communication in fostering a satisfactory marital relationship is more important now than in earlier times, because the expectation

and demands of marriage have changed. As noted above, marriage in America is less dependent on and affected by an extended kin network than on the spousal relationship. One of the principal functions of contemporary marriage is the nurturance of family members. Perhaps because this function and the therapeutic and leisure roles that help fulfill it in marriage are preeminent, “greater demands are placed on each spouse’s ability to communicate” (Fitzpatrick 1988, p. 2). The communication of positive affect and its converse, emotional withdrawal, may well be the essence and the antithesis, respectively, of nurturance. Bloom and colleagues (1985) suggest that one important characteristic of marital dissatisfaction is the expectation that marriage is a “source of interpersonal nurturance and individual gratification and growth” (p. 371), an expectation that is very hard to fulfill.

In the 1990s many studies focused on the relationship between communication and marital satisfaction (Burlinson and Denton 1997). The findings from this body of research suggest that there are clear communication differences between spouses in happy and in unhappy marriages. Patricia Noller and Mary Anne Fitzpatrick (1990) reviewed this literature, and their findings can be summarized as follows: Couples in distressed marriages report less satisfaction with the social-emotional aspects of marriage, develop more destructive communication patterns (i.e., a greater expression of negative feelings, including anger, contempt, sadness, fear, and disgust), and seek to avoid conflict more often than nondistressed couples. Nevertheless, couples in distressed marriages report more frequent conflict and spend more time in conflict. In addition, gender differences in communication are intensified in distressed marriages. For example, husbands have a more difficult time interpreting wives’ messages. Wives in general express both negative and positive feelings more directly, and are more critical. Spouses in unhappy marriages appear to be unaware that they misunderstand one another. Generally, happily married couples are more likely to engage in positive communication behaviors (agreement, approval, assent, and the use of humor and laughter), while unhappy couples command, disagree, criticize,

put down, and excuse more. Recently, Burleson and Denton (1997) explored the complexity of the communication–marital satisfaction relationship and found a variety of moderating factors: skills in communicating (realizing the communication goal, producing and receiving messages, social perception), the context or setting in which communication takes place, and the cognitive complexity of each spouse. They suggest that communication problems are best viewed as a symptom of marital difficulties and should not be seen merely as a diagnostic tool for distressed relationships.

Communication patterns may be class linked. It has long been found that working-class wives in particular complain that their husbands are emotionally withdrawn and inexpressive (Komarovskiy 1962; Rubin 1976). Olsen and his colleagues (1979) assign communication a strategic role in marital and family adaptability. In their conceptualization of marital and family functioning, communication is the process that moves couples along the dimensions of cohesion and adaptability. In another study the absence of good communication skills is associated with conjugal violence (Infante 1989).

Differences between the sexes have been reported in most studies that examine marital communication. The general emphasis of these findings is that males appear less able to communicate verbally and to discuss emotional issues. However, communication is not the only aspect of marriage for which sex differences have been reported. Other components of marriage also are experienced differently, depending on the sex of spouse. The following paragraphs report some of these.

**Sex Differences.** In her now classic book, *The Future of Marriage* (1972), Jessie Bernard pointed out that marriage does not hold the same meanings for wives as for husbands, nor do wives hold the same expectations for marriage as do husbands. These sex differences (originally noted but not fully developed by Emile Durkheim in 1897 in *Le Suicide*) have been observed and examined by many others since Bernard's publication (Larson 1988; Thompson and Walker 1989; Kitson 1992). For example, researchers have reported differences between husbands and wives in perceptions

of marital problems, reasons for divorce, and differences in perceived marital quality; wives consistently experience and perceive lower marital quality than do husbands.

Sex differences in marriage are socially defined and prescribed (Lee 1982; Blaisure and Allen 1995). One consequence of these social definitions is that sex differences get built into marital roles and the division of labor within marriage. For example, it has been observed that wives do more housework and child care than husbands (Thompson and Walker 1989; Presser 1994). Even wives who work in the paid labor force spend twice as many hours per week in family work as husbands (Benin and Agostinelli 1988; Coltran and Ishii-Kuntz 1992; Demo and Acock 1993). Wives are assigned or tend to assume the role of family kin keeper and caregiver (Montgomery 1992). To the extent that husbands and wives experience different marriages, wives are thought to be disadvantaged by their greater dependence, their secondary status, and the uneven distribution of family responsibilities between spouses (Baca Zinn and Eitzen 1990).

All these factors are assumed to affect the quality of marriage—one of the most studied aspects of marriage (Adams 1988; Berardo 1990). It will be the subject of our final discussion.

**Marital Quality.** Marital quality may be the “weather vane” by which spouses gauge the success of their relationship. The reader should be sure to differentiate the concept of marital quality from two other closely related concepts: family quality and the quality of life in general, called “global life satisfaction” in the literature. Studies show that people clearly differentiate among these three dimensions of well-being (Ishii-Kuntz and Ihinger-Tallman 1991).

Marriage begins with a commitment, a promise to maintain an intimate relationship over a lifetime. Few couples clearly understand the difficulties involved in adhering to this commitment or the problems they may encounter over the course of their lives together. More people seek psychological help for marital difficulties than for any other type of problem (Veroff et al. 1981). For a



large number of spouses, the problems become so severe that they renege on their commitment and dissolve the marriage.

A review of the determinants of divorce lists the following problems as major factors that lead to the dissolution of marriage: "alcoholism and drug abuse, infidelity, incompatibility, physical and emotional abuse, disagreements about gender roles, sexual incompatibility, and financial problems" (White 1990, p. 908). Underlying these behaviors appears to be the general problem of communication. In their study of divorce, Gay Kitson and Marvin Sussman (1982) report lack of communication or understanding to be the most common reason given by both husbands and wives concerning why their marriage broke up. The types of problems responsible for divorce have not changed much over time. Earlier studies also list nonsupport, financial problems, adultery, cruelty, drinking, physical and verbal abuse, neglect, lack of love, in-laws, and sexual incompatibility as reasons for divorce (Goode 1956; Levinger 1966).

Not all unhappy marriages end in divorce. Many factors bar couples from dissolving their marriages, even under conditions of extreme dissatisfaction. Some factors that act as barriers to marital dissolution are strong religious beliefs, pressure from family or friends to remain together, irretrievable investments, and the lack of perceived attractive alternatives to the marriage (Johnson et al. 1999).

One empirical finding that continues to be reaffirmed in studies of marital quality is that the quality of marriage declines over time, beginning with the birth of the first child (Glenn and McLanahan 1982; Glenn 1991; White et al. 1986). Consequently the transition to parenthood and its effect on the marital relationship has generated a great deal of research attention (Cowen and Cowen 1989; McLanahan and Adams 1989). The general finding is that marital quality decreases after the birth of a child, and this change is more pronounced for mothers than for fathers. Two reasons generally proposed to account for this decline are that the amount of time couples have to spend together decreases after the birth of a child, and that sex

role patterns become more traditional (McHale and Huston 1985).

In an attempt to disentangle the duration of marriage and parenthood dimensions, White and Booth (1985) compared couples who became new parents with nonparent couples over a period of several years and found a decline in marital quality regardless of whether the couple had a child. A longitudinal study conducted by Belsky and Rovine (1990) confirmed the significant declines in marital quality over time reported in so many other studies. They also found the reported gender differences. However, their analysis also focused on change scores for individual couples. They reported that while marital quality declined for some couples, this was not true for all couples: It improved or remained unchanged for others. Thus, rather than assume that quality decline is an inevitable consequence of marriage, there is a need to examine why and how some couples successfully avoid this deterioration process. The authors called for the investigation of individual differences among couples rather than continuing to examine the generally well-established finding that marital quality declines after children enter the family and remains low during the child-rearing stages of the family life cycle.

Finally, the overall level of marital satisfaction is related to the frequency with which couples have sex (Call et al. 1995). The argument states that happy couples have sex more frequently, leading to more satisfying marriage, and that satisfaction in marriage leads to a greater desire for sex and the creation of more opportunities for sex.

Many students of the family have found it useful to consider marital development over the years as analogous to a career that progresses through stages of the family life cycle (Duvall and Hill 1948; Aldous 1996). This allows for consideration of changes in the marital relationship that occur because of spouses' aging, the duration of marriage, and the aging of children. In addition to changes in marital quality, other factors have been examined, such as differences in the course of a marriage, when age at first marriage varies (e.g., marriage entered into at age 19 as opposed to in

the mid-30s), varied duration of childbearing (few versus many years), varied number of children (small family versus large), and the ways in which consumer decisions change over the course of marriage (Aldous 1990).

### CONCLUSION

If the vitality of marriage is measured by the extent to which men and women enter marriage, then some pessimism about its future is warranted. Marriage rates are currently lower than during the early depression year 1931 (*Statistical Abstracts* 1998, 1938)—which was the lowest in our nation's recent history (Sweet and Bumpass 1987). One conclusion that might be drawn from reading the accumulated literature on marriage, especially the writings that discuss the inequities of men and women within marriage, the increasing incidence of marital dissolution, cohabitation as a substitute for marriage, and the postponement of marriage, is that the institution is in serious trouble. These changes have been interpreted as occurring as part of a larger societal shift in values and orientations (Glick 1989) that leans toward valuing adults over children and individualism over familism (Glenn 1987; White 1987; see also the entire December 1987 issue of *Journal of Family Issues*, which is devoted to the state of the American family). Supporting this perspective are the data on increased marital happiness among childless couples and lower birth rates among married couples.

Yet every era has had those who wrote of the vulnerability of marriage and the family. For example, earlier in this century Edward Alsworth Ross wrote, "we find the family now less stable than it has been at any time since the beginning of the Christian era" (1920, p. 586). Is every era judged to be worse than previous ones when social institutions are scrutinized? More optimistic scholars look at the declining first-marriage rate and interpret it as a "deferral syndrome" rather than an outright rejection of the institution (Glick 1989). This is because, in spite of declines in the overall rate, the historical 8 percent to 10 percent of never-married people in the population has remained constant: Almost 90 percent of all women

in the United States eventually marry at least once in their lifetimes. Also, projections that about two-thirds of all first marriages in the United States will end in divorce (Martin and Bumpass 1989) do not deter people from marrying. In spite of the high divorce rate, an increased tolerance for singleness as a way of life, and a growing acceptance of cohabitation, the majority of Americans continue to marry. Marriage is still seen as a source of personal happiness (Kilbourne et al. 1990).

More fundamentally, marriage rates and the dynamics of marital relationships tend to reflect conditions in the larger society. What appears clear, at least for Americans, is that they turn to marriage as a source of sustenance and support in a society where, collectively, citizens seem to have abrogated responsibility for the care and nurturance of each other. Perhaps it is not surprising that divorce rates are high, given the demands and expectations placed on modern marriages.

(SEE ALSO: *Alternative Life-Styles; Courtship; Family Roles; Heterosexual Behavior Patterns; Intermarriage; Marital Adjustment; Remarriage; Sexual Behavior and Marriage*)

### REFERENCES

- Adams, Bert N. 1988 "Fifty Years of Family Research: What Does It Mean?" *Journal of Marriage and the Family* 50:5-17.
- Aldous, Joan 1990 "Family Development and the Life Course: Two Perspectives on Family Change." *Journal of Marriage and the Family* 52:571-583.
- 1996 *Family Careers: A Rethinking*. Thousand Oaks, Calif.: Sage.
- Baca Zinn, Maxine, and D. Stanley Eitzen 1990 *Diversity in Families*, 2nd ed. New York: Harper & Row.
- Barich, Rachel Roseman, and Denise D. Bielby 1996 "Rethinking Marriage: Change and Stability in Expectations, 1967-1994." *Journal of Family Issues* 17:139-169.
- Barry, W. A. 1970 "Marriage Research and Conflict: An Integrative Review." *Psychological Bulletin* 73:41-54.
- Belsky, Jay, and Michael Rovine 1990 "Patterns of Marital Change across the Transition to Parenthood: Pregnancy to Three Years Postpartum." *Journal of Marriage and the Family* 52:5-19.

- Benin, Mary H., and Joan Agostinelli 1988 "Husbands' and Wives' Satisfaction with the Division of Labor." *Journal of Marriage and the Family* 50:349-361.
- Berardo, Felix M. 1990 "Trends and Directions in Family Research in the 1980s." *Journal of Marriage and the Family* 52:809-817.
- Bernard, Jessie 1972 *The Future of Marriage*. New York: Bantam.
- 1982 *The Future of Marriage*. New Haven, Conn.: Yale University Press.
- Blaisure, Karen, and Katherine R. Allen 1995 "Feminists and the Ideology and Practice of Marital Equality." *Journal of Marriage and the Family* 57:5-19.
- Bloom, Bernard L., Robert L. Niles, and Anna M. Tatacher 1985 "Sources of Marital Dissatisfaction among Newly Separated Persons." *Journal of Family Issues* 6:359-373.
- Burleson, Brant R., and Wayne H. Denton 1997 "The Relationship between Communication Skills and Marital Satisfaction: Some Moderating Effects." *Journal of Marriage and the Family* 59:884-902.
- Call, Vaughn, Susan Sprecher, and Pepper Schwartz 1995 "The Incidence and Frequency of Marital Sex in a National Sample." *Journal of Marriage and the Family* 57:639-652.
- Coltrane, Scott, and Masako Ishii-Kuntz 1992 "Men's Housework: A Life Course Perspective." *Journal of Marriage and Family* 54:43-57.
- Cowen, Phillip, and C. Cowen 1989 "Changes in Marriage during the Transition to Parenthood: Must We Blame the Baby?" In G. Michaels and W. Goldberg, eds., *The Transition to Parenthood: Current Theory and Research*. New York: Cambridge University Press.
- Davis, Kingsley 1949 *Human Society*. New York: Macmillan.
- Demo, David H., and Alan C. Acock 1993 "Family Diversity and the Division of Domestic Labor: How Much Have Things Really Changed?" *Family Relations* 42:323-331.
- Durkheim, Emile 1897 *Le Suicide*. Paris: F. Alcan.
- Duvall, Evelyn M., and Reuben Hill 1948 "Report of the Committee on the Dynamics of Family Interaction." Paper delivered at the National Conference on Family Life, Washington.
- EGGEBEEN, David J., and Adam Davey 1998 "Do Safety Nets Work? The Role of Anticipated Help in Times of Need." *Journal of Marriage and the Family* 60:939-950.
- Fincham, Frank D., and Thomas N. Bradbury 1987 "The Assessment of Marital Quality: A Reevaluation." *Journal of Marriage and the Family* 49:797-809.
- Fitzpatrick, Mary Anne 1988 "Approaches to Marital Interaction." In P. Noller and M. A. Fitzpatrick, eds., *Perspectives on Marital Interaction*. Clevedon, England: Multilingual Matters.
- Glenn, Norval D. 1987 "Tentatively Concerned View of American Marriage." *Journal of Family Issues* 8:350-354.
- 1991 "Qualitative Research on Marital Quality in the 1980s: A Critical Review." In A. Booth, ed., *Contemporary Families: Looking Forward, Looking Back*. Minneapolis, Minn.: National Council on Family Relations.
- , and Sara McLanahan 1982 "Children and Marital Unhappiness: A Further Specification of the Relationship." *Journal of Marriage and the Family* 44:63-72.
- Glick, Paul C. 1989 "The Family Life Cycle and Social Change." *Family Relations* 38:123-129.
- Goode, William J. 1956 *After Divorce*. New York: Free Press.
- Gottman, John Mordechai 1994 *What Predicts Divorce? The Relationship between Marital Processes and Marital Outcomes*. Hillsdale, N.J.: Lawrence Erlbaum.
- Holman, Thomas B., and Gregory W. Brock 1986 "Implications for Therapy in the Study of Communication and Marital Quality." *Family Perspectives* 20:85-94.
- Infante, D. A. 1989 "Test of an Argumentative Skill Deficiency Model of Interpersonal Violence." *Communication Monographs* 56:163-177.
- Ishii-Kuntz, Masako, and Marilyn Ihinger-Tallman 1991 "The Subjective Well-Being of Parents." *Journal of Family Issues* 12:58-68.
- Ishii-Kuntz, Masako, and Scott Coltrane 1992 "Remarriage, Stepparenting, and Household Labor." *Journal of Family Issues* 13:215-233.
- Johnson, Michael P., John P. Caughlin, and Ted L. Huston 1999 "The Tripartite Nature of Marital Commitment: Personal, Moral, and Structural Reasons to Stay Married." *Journal of Marriage and the Family* 61:160-177.
- Kilbourne, Barbara S., Frank Howell, and Paula England 1990 "Measurement Model for Subjective Marital Solidarity: Invariance across Time, Gender and Life Cycle Stage." *Social Science Research* 19:62-81.
- Kitson, Gay C. 1992 *Portrait of Divorce: Adjustment to Marital Breakdown*. New York: Guilford Press.
- , and Marvin B. Sussman 1982 "Marital Complaints, Demographic Characteristics and Symptoms of Mental Distress in Divorce." *Journal of Marriage and the Family* 44:87-102.

- Komarovsky, Mirra 1962 *Blue Collar Marriage*. New York: Vintage.
- Larson, Jeffrey H. 1988 "The Marriage Quiz: College Students' Beliefs in Selected Myths about Marriage." *Family Relations* 37:3-11.
- Lasch, Christopher 1977 *Haven in a Heartless World*. New York: Basic.
- Lee, Gary 1982 *Family Structure and Interaction: A Comparative Analysis*, 2nd ed., rev. Minneapolis: University of Minnesota Press.
- Levinger, George 1996 "Sources of Marital Dissatisfaction among Applicants for Divorce." *American Journal of Orthopsychiatry* 36:803-807.
- Martin, Teresa C., and Larry L. Bumpass 1989 "Recent Trends in Marital Disruption." *Demography* 26:37-51.
- McHale, Susan M., and Ted L. Huston 1985 "Men and Women as Parents: Sex Role Orientations, Employment, and Parental Roles." *Child Development* 55:1349-1361.
- McLanahan, Sara, and J. Adams 1989 "The Effects of Children on Adults' Psychological Well-Being." *Social Forces* 68:124-146.
- Montgomery, Rhonda J. 1992 "Gender Differences in Patterns of Child-Parent Caregiving Relationships." In J. W. Dwyer and R. T. Coward, eds., *Gender, Families, and Elder Care*. Newbury Park, Calif.: Sage.
- Noller, Patricia, and Mary Anne Fitzpatrick 1990 "Marital Communication in the Eighties." *Journal of Marriage and the Family* 52:832-843.
- O'Donohue, William, and Julie L. Crouch 1996 "Marital Therapy and Gender-Linked Factors in Communication." *Journal of Marital and Family Therapy* 22:87-101.
- Okun, B. F. 1991 *Effective Helping, Interviewing, and Counseling Techniques*, 13th ed. Monterey, Calif.: Brooks/Cole.
- Olsen, David H., D. Sprenkle, and C. Russell 1979 "Circumplex Model of Marital and Family Systems I: Cohesion and Adaptability Dimensions, Family Types and Clinical Applications." *Family Process* 18:3-28.
- Popenoe, David 1993 "American Family Decline, 1960-1990: A Review and Appraisal." *Journal of Marriage and the Family* 55:527-555.
- Presser, H. B. 1994 "Employment Schedules among Dual-Earner Spouses and the Division of Household Labor by Gender." *American Sociological Review* 59:348-364.
- Reiss, Ira L., and Gary Lee 1988 *Family Systems in America*. New York: Holt, Rinehart, and Winston.
- Rosenthal, Carolyn 1985 "Kinkeeping in the Familial Division of Labor." *Journal of Marriage and the Family* 47:965-974.
- Ross, Edward Alsworth 1920 *The Principles of Sociology*. New York: Century.
- Rubin, Lillian 1976 *Worlds of Pain: Life in the Working Class Family*. New York: Basic.
- Ruble, Diane, A. Fleming, L. Hackel, and C. Stangor 1988 "Changes in the Marital Relationship during the Transition to First-Time Motherhood: Effects of Violated Expectations Concerning the Division of Labor." *Journal of Personality and Social Psychology* 55:78-87.
- Slater, Philip 1963 "On Social Regression." *American Sociological Review* 28:339-364.
- Statistical Abstracts of the United States* 1939 Washington: Bureau of the Census, Department of Commerce.
- 1998 Washington: Bureau of the Census, Department of Commerce.
- Sweet, James A., and Larry L. Bumpass 1987 *American Families and Households*. New York: Russell Sage Foundation.
- Thompson, Linda, and Alexis J. Walker 1989 "Gender in Families: Women and Men in Marriage, Work, and Parenthood." *Journal of Marriage and the Family* 51:845-871.
- Veroff, Joseph, Richard Kulka, and Elizabeth Douvan 1981 *Mental Health in America: Patterns of Help-Seeking from 1957 to 1976*. New York: Basic.
- White, Lynn K. 1987 "Freedom versus Constraint: The New Synthesis." *Journal of Family Issues* 8:468-470.
- 1990 "Determinants of Divorce: Review of Research in the Eighties." *Journal of Marriage and the Family* 52:904-912.
- , and Alan Booth 1985 "The Transition to Parenthood and Marital Quality." *Journal of Family Issues* 6:435-449.
- , and John N. Edwards 1986 "Children and Marital Happiness: Why the Negative Correlation?" *Journal of Family Issues* 7:131-147.
- Whyte, Martin K. 1990 *Dating, Mating, and Marriage*. New York: Aldine de Gruyter.

MARILYN IHINGER-TALLMAN

## MARRIAGE AND DIVORCE RATES

Marriage and divorce rates are measures of the propensity for the population of a given area to become married or divorced during a given year. Some of the rates are quite simple, and others are progressively more refined. The simple ones are

called *crude rates* and are expressed in terms of the number of marriages or divorces per 1,000 persons of all ages in the area at the middle of the year. These are the only marriage and divorce rates available for every state in the United States. They have the weakness of including in the base not only young children but also elderly persons, who are unlikely to marry or become divorced. But the wide fluctuations in crude rates over time are obviously associated with changes in the economic, political, and social climate.

More refined rates will be discussed below, but the following illustrative crude rates of marriage for the United States will demonstrate the readily identifiable consequences of recent historical turning points or periods (NCHS 1990a, 1990b, 1998).

Between 1940 and 1946 the crude marriage rate for the United States went up sharply, from 12.1 per 1,000 to 16.4 per 1,000, or by 36 percent, showing the effects of depressed economic conditions before, and disarmament after, World War II.

Between 1946 and 1956 the rate went down rapidly to 9.5 per 1,000, or by 42 percent, as the baby boom peaked. The unprecedented increase in the number of young children was included in the base of the rate, and this helped to lower the rate.

Between 1956 and 1964 the rate declined farther to 9.0 per 1,000, or by 5 percent, as the baby boom ended and the Vietnam War had begun. Between 1964 and 1972 the rate went up moderately to a peak of 10.9 per 1,000, or by 21 percent, as the Vietnam War ended and many returning war veterans married.

Between 1972 and 1990 the crude marriage rate declined irregularly to 9.8 per 1,000 persons of all ages and on down to 8.9 in 1997. The factors involved are discussed below.

At the time of this writing, vital statistics annual reports present only *crude* marriage and device rates for the United States and individual states (NCHS 1998). More *refined* rates were last published for 1987 on some subjects and for 1990 on other subjects.

*Refined marriage rates* show the propensity to marry for adults who are eligible to marry. They exclude from the base all persons who are too young to marry and may also limit the base to an age range within which most marriages occur. The conventional practice of basing these rates on the number of women rather than all adults has the advantage of making the level of the rates correspond approximately to the number of couples who are marrying. Moreover, the patterns of changes in rates over time are generally the same for men and women.

Changes over time in the tendency for adults to marry are more meaningful and may fluctuate more widely if they are reported in refined rather than crude rates. To illustrate, the crude marriage rate declined between 1972 and 1980 from 10.9 per 1,000 population to 10.6, or by only 3 percent, while the refined rate (marriages per 1,000 women 15 to 44 years old) declined from 141.3 to 102.6, or by 27 percent. A change that appeared to be small when measured crudely turned out to be large when based on a more relevant segment of the population. Persons who want to have others believe that the change was small may cite the crude rates, and persons who want to demonstrate that the marital situation was deteriorating rapidly may cite the refined rates. But persons interested in making a balanced presentation may choose to cite both types of results and explain the differences between them.

Still greater refinement can be achieved by computing marriage rates according to such key variables as age groups and previous marital status. Examples appear in Table 1 and Figures 1 and 3 for the United States from 1971 to 1990. Table 1 and Figure 1 show first marriage rates by age for the only marital status category of eligible persons, namely, never-married adults (women) 15 years old and over. Figure 2 shows divorce rates by age for married women, and Figure 3 shows remarriage rates by age for divorced women. The low remarriage rates for widows are not shown here but are treated briefly elsewhere in the article. Rates of separation because of marital discord are also not presented here: Separated adults are still legally married and are therefore included in the base of divorce rates.

MARRIAGE AND DIVORCE RATES

**First Marriage Rates per 1,000 Never-Married Women, Divorce Rates per 1,000 Married Women, and Remarriage Rates per 1,000 Divorced Women, by Age: United States, 1970-1990**

AGE (YEARS)	1970a	1975	1980b	1983	1985	1987	1990
<b>First Marriage Rate</b>							
15 or over	93	76	68	64	62	59	58
15 to 17	36	29	22	16	13	12	11
18 to 19	150	115	92	73	67	58	53
20 to 24	198	144	122	107	102	98	93
25 to 29	131	115	104	105	104	105	109
30 to 34	75	62	60	61	66	69	71
35 to 39	48	36	33	38	37	42	47
40 to 44	27	26	22	22	24	22	20
45 to 64	10	9	8	8	11	10	NA
65 or over	1	1	1	1	1	1	1
<b>Divorce Rate</b>							
15 or over	14	NA	20	19	19	19	19
15 to 19	27	NA	42	48	48	50	49
20 to 24	33	NA	47	43	47	46	46
25 to 29	26	NA	38	36	36	34	37
30 to 34	19	NA	29	28	29	27	28
35 to 44	11	NA	24	25	22	22	NA
45 to 54	5	NA	10	11	11	11	NA
55 to 64	3	NA	4	4	4	4	NA
<b>Remarriage Rate</b>							
15 or over	133	117	104	92	82	81	76
20 to 24	420	301	301	240	264	248	252
25 to 29	277	235	209	204	184	183	200
30 to 34	196	173	146	145	128	137	138
35 to 39	147	117	108	99	97	92	93
40 to 44	98	91	69	67	63	69	69
45 to 64	47	40	35	31	36	37	NA
65 or over	9	9	7	5	5	5	5

*Table 1*

NOTE: <sup>a</sup>First marriage and remarriage rates for 1971.

<sup>b</sup>First marriage and remarriage rates for 1979.

SOURCE: National Center for Health Statistics 1990b, 1990c, 1995a, 1995b.

MARRIAGE AND DIVORCE RATES

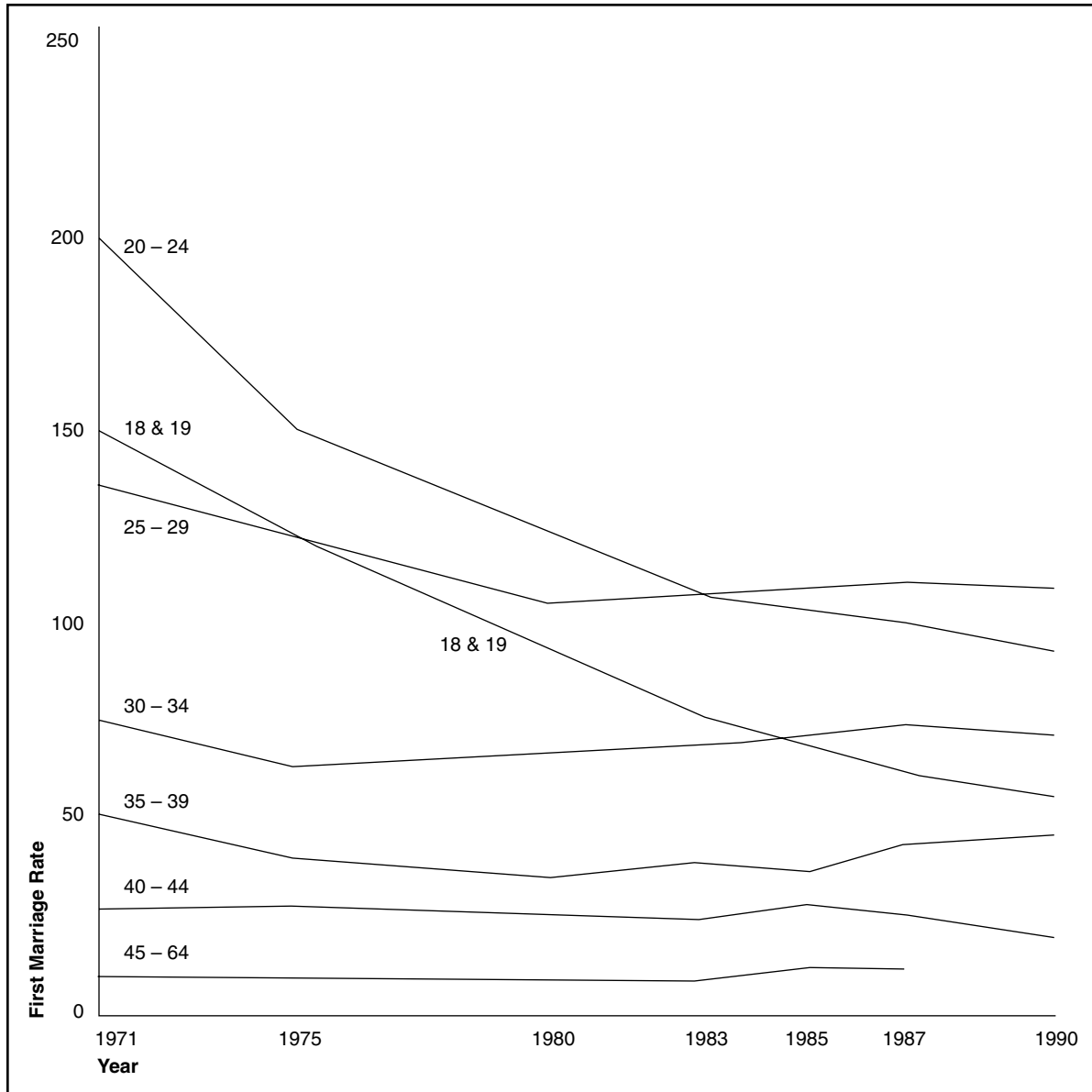


Figure 1

NOTE: First marriage rate per 1,000 never-married women, by age: United States, 1971-1990.

The marriage and divorce rates in Table 1 were based on data from reports published by the National Center for Health Statistics (NCHS). These reports contain information, obtained from central offices, of vital statistics in the states that are in the Marriage Registration Area (MRA) and the Divorce Registration Area (DRA). In 1990 the District of Columbia and all but eight states were in the MRA, while the District of Columbia and

only thirty-one states were in the DRA. Funding for the central offices is determined by each state's legislature. But for states not in the MRA or DRA, the NCHS requests the numbers of marriages and divorces from local offices where marriage and divorce certificates are issued. The reports on divorce include the small number of annulments and dissolutions of marriage. Bases for the marriage and divorce rates in Table 1 were obtained

from special tabulations made by the U.S. Bureau of the Census from Current Population Survey data. These are tabulations of adults in MRA and DRA states and classified by marital status, age, and sex. Because not all the population of the United States is included in the MRA and DRA, the detailed marriage and divorce statistics published by the NCHS constitute approximations of the marital situation in the country as a whole. This article contains much numerical information that was published in one or more of the NCHS reports listed in References.

### FIRST MARRIAGE RATES BY AGE

Illustrations of first marriage rates appear in table 1. For the United States, the first marriage rates per 1,000 never-married women 15 years old and over were 93 in 1971 and 58 in 1990. In effect, 9.3 percent of the never-married women in 1971 and 5.8 percent in 1990, became married for the first time. For men the corresponding rates were 68 in 1971 and 47 in 1990.

First marriage rates tend to decline with age, and the rates for most of the age groups shown in Table 1 were declining over time. The rates for the age groups under 20 years of age were among the highest, but they dropped so sharply that they were only about one-third as high in 1990 as they had been in 1971. At the oldest ages, the change appears to have been slight. Obviously, the propensity to marry was falling far more abruptly among the young than among the older singles, probably in reaction to the suddenly changing cultural climate.

The generally downward trend in the marriage rate for each young age group was especially rapid during the early 1970s. By 1975, the veterans of the Vietnam War had already entered delayed marriages, and the upsurge in cohabitation outside marriage was only beginning to depress the first marriage rate. During the 1980s the slight upturn in the first marriage rate for women over 30 years of age probably reflected an increase in marriages among women who had delayed marrying for the purposes of obtaining a higher education and becoming established in the workplace.

Research has produced evidence that women who marry for the first time after they reach their thirties are more likely to have stable marriages than those who marry in their twenties (Norton and Moorman 1987). Although first marriage rates among adults in their forties are relatively low, they are by no means negligible.

As first marriage rates declined between the mid-1960s and the late 1980s, the median age at first marriage rose at an unprecedented pace over this short period of time. According to vital statistics, the median age at first marriage for women went up from 20.3 years in 1963 to 24.0 in 1990, for men it went up from 22.5 years to 25.9 years. As age at first marriage increased, the distribution of ages at first marriage also increased (Wilson and London 1987).

One of the consequences of the great delay of first marriage has been a very sharp rise in premarital pregnancy. Only 5 percent of births in 1960 occurred to unmarried mothers, but this increased to 24 percent in 1987 and to 32 percent in 1996. Research by Bumpass and McLanahan (1989) showed that one-half of nonmarital births during the late 1980s were first births, and about one-third occurred to teenagers. Moreover, about one-tenth of brides were pregnant at first marriage. Thus, about one-third of the first births during the late 1980s were conceived before marriage.

As the first marriage rate declined, the proportion of all marriages that were primary marriages (first for bride and groom) also declined. In 1970, two-thirds (68 percent) were primary marriages, but by 1987 the proportion was barely over one-half (54 percent). Meantime, marriages of divorced brides to divorced grooms nearly doubled, from 11 percent to 19 percent, while marriages of widows to widowers went down from 2 percent to 1 percent.

Men and women, regardless of previous marital status, tend to marry someone whose age is similar to their own. But men who enter first marriage when they are older than the average age of men at marriage have a reasonable likelihood of marrying a woman who has been divorced.

Procedures have been developed for projecting the proportion of adults of a certain age who



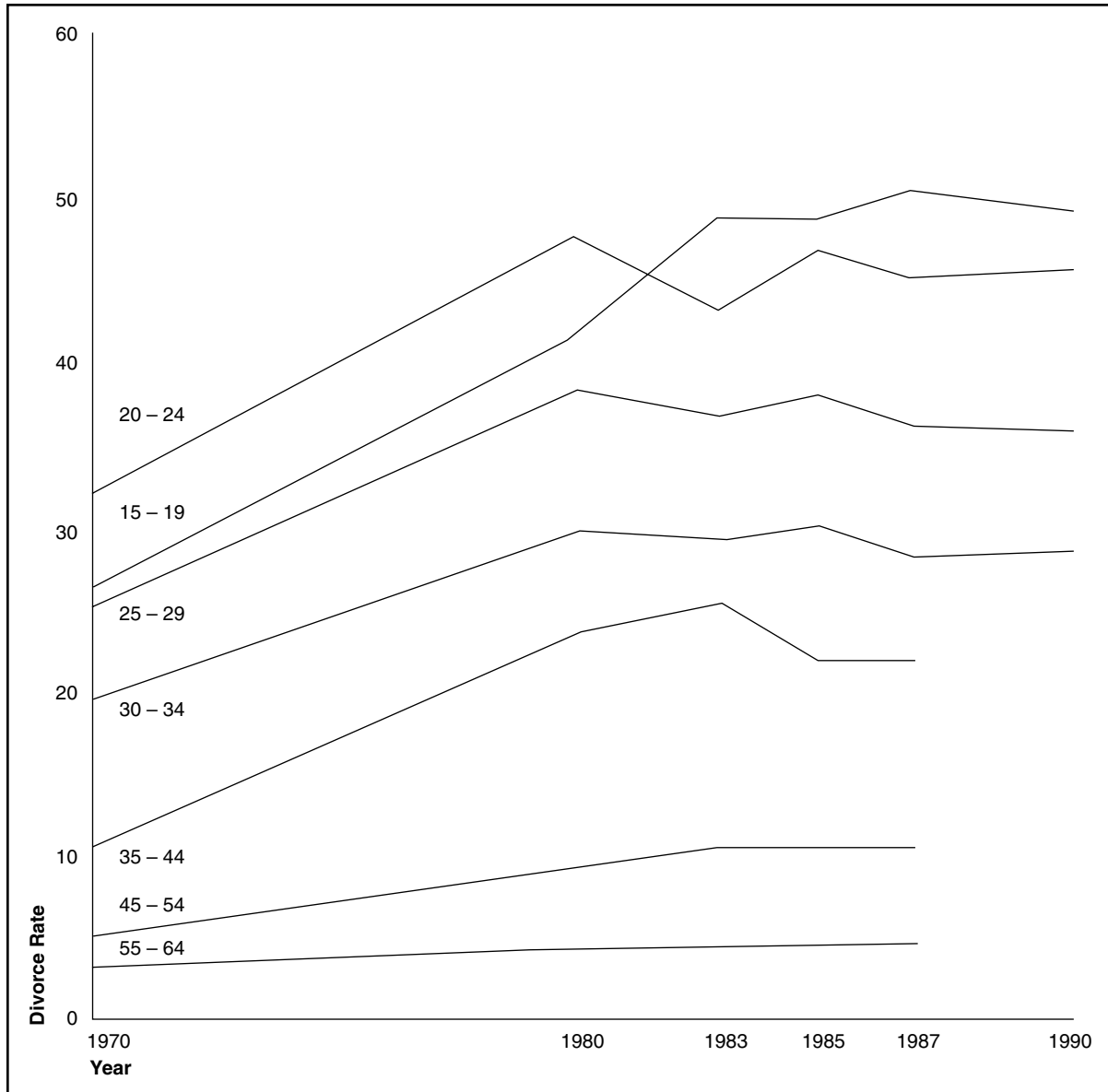


Figure 2

NOTE: Divorce rate per 1,000 married women, by age: United States, 1970-1990.

are likely to enter first marriage sometime during their lives. This measure is, in effect, “a lifetime first marriage rate.” One of these procedures was used by Schoen and colleagues (1985) to find that about 94 percent of men who were born in the years 1948 to 1950 and who survived to age 15 were expected to marry eventually; for women, it was 95 percent. Their projections for those born in 1980 were significantly lower, 89 percent for men

and 91 percent for women. Despite the implied decline, a level of nine-tenths of the young adults deciding to marry at least once is still high by world standards.

#### DIVORCE RATES BY AGE

As indicated above, *crude* divorce rates are the only rates available annually for the United States and

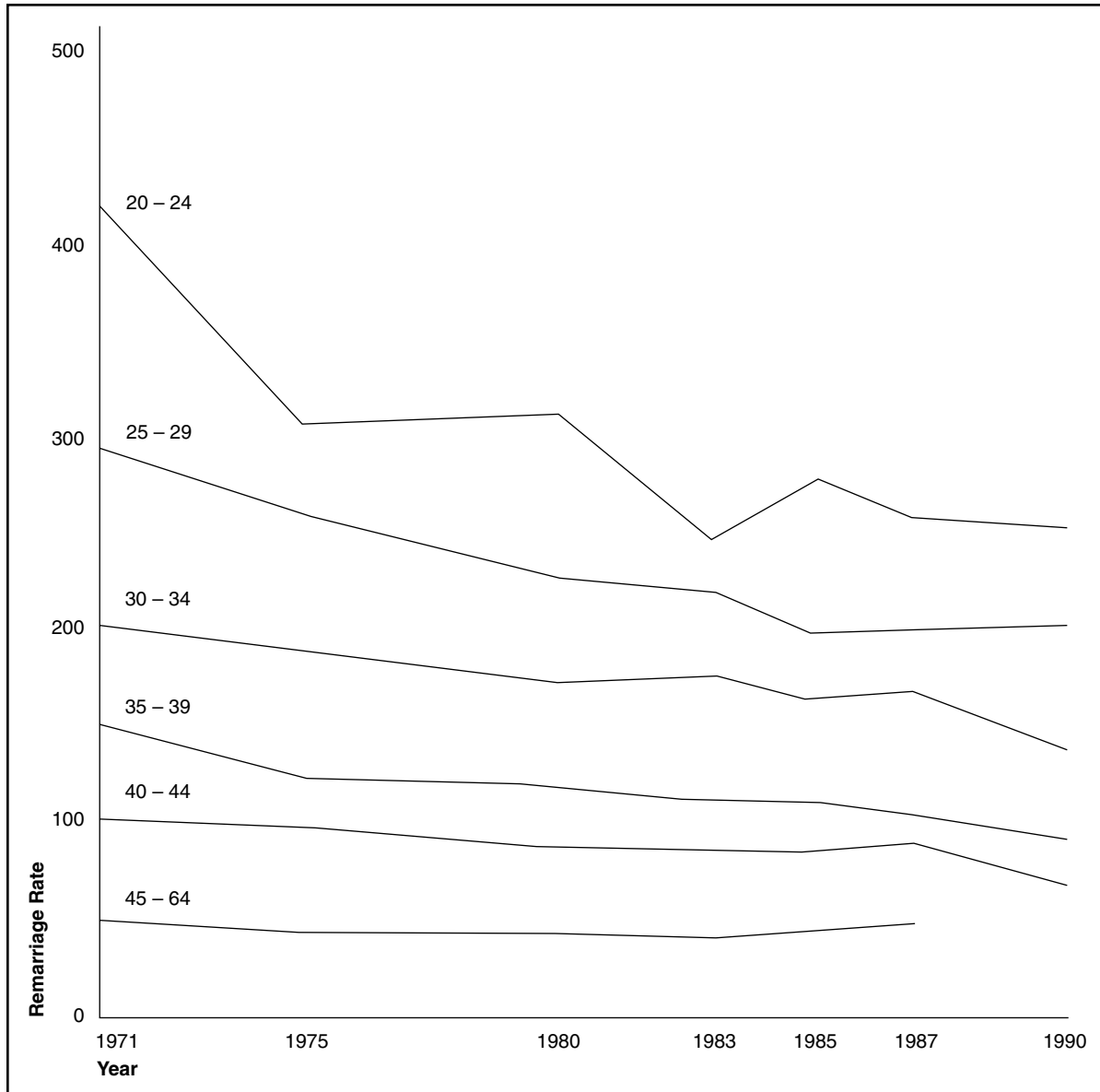


Figure 3

NOTE: Remarriage rate per 1,000 divorced women, by age: United States, 1971-1990.

individual states. As recently as 1965, the rate was only 2.5 divorces per 1,000 persons in the United States, but by 1979 the rate reached a peak more than twice that high, 5.3. Then the rate declined gradually until 1997, when it was down to 4.3, the lowest rate since 1973.

The most recent *refined* divorce rates available include those by age and sex in Table 1 for 1990. The report for that year also presented other

*refined* divorce statistics on the number of divorces occurring among persons in their first marriage, their second marriage, and their third or subsequent marriage (NCHS 1995b). But the required bases for computing first divorce rates and redi-orce rates are not available. During the 1980s, about three-fourths of the divorces were obtained by adults in their first marriage, about one-fifth by

those in their second marriage, and one-twentieth by those who had been married at least three times.

The divorce rates in Table 1 provide illustrations of the magnitude of the rates for the period 1970 to 1990. The divorce rate per 1,000 married women was 14 in 1970 and 19 in 1990. Corresponding rates for men were the same. The rate reached a peak of 20 in 1980, nearly half again as high as in 1970, and declined slightly to 19 in 1987, a level still well above that in 1970.

The divorce rate for married women rose dramatically in every age group during the 1970s and changed relatively little from then through the 1980s. In January 1973 the Vietnam War ended, and at that time the norms regarding the sanctity of marriage were being revised. The advantages of a permanent marriage were being weighed against the alternatives, including freedom from a seriously unsatisfactory marital bond and the prospect of experimenting with cohabitation outside marriage or living alone without any marital entanglements.

By 1994, one-fourth of divorced persons were cohabiting outside marriage, one-third were living alone, and most of the rest were in single-parent families (U.S. Bureau of the Census 1996).

Married women under 25 years of age have consistently high divorce rates resulting largely from adjustment difficulties associated with early first marriage. Noteworthy in this context is the finding that 30 percent of the women entering a first marriage in 1980 were in their teens, while 40 percent of those obtaining a divorce had entered their marriage while teenagers.

More divorces during the 1980s were occurring to married adults 25 to 34 years of age than to those in any other ten-year age group. A related study by Norton and Moorman (1987) concluded that women in their late thirties in 1985 were likely to have higher lifetime divorce rates (55 percent) than those either ten years older or ten years younger. This cohort was born during the vanguard of the baby boom and became the trend setter for higher divorce rates. The lower rate for those ten years younger may reflect their concern caused by the adjustment problems of their older divorced siblings or friends.

Married women over 45 years of age have quite low divorce rates. Most of their marriages

must still be reasonably satisfactory, or not sufficiently unsatisfactory to persuade them to face the disadvantages that are often associated with becoming divorced. Yet, the forces that raised the divorce rate for younger women greatly after 1970 also made the small rate for the older women increase by one-fourth during the 1970s and remain at about the same level through the 1980s.

The median duration of marriage before divorce has been seven years for several decades. This finding is not proof of a seven-year itch. In fact, the median varies widely according to previous marital status from eight years for first marriages to six years for second marriages and four years for third and subsequent marriages. The number of divorces reaches a peak during the third year of marriage and declines during each succeeding year of marriage.

Among separating couples, the wife usually files the petition for divorce. However, between 1975 and 1987, the proportion of husband petitioners increased from 29.4 percent to 32.7 percent, and the small proportion of divorces in which both the husband and the wife were petitioners more than doubled, from 2.8 percent to 6.5 percent. These changes occurred while the feminist movement was becoming increasingly diffused and the birthrate was declining, with the consequence that only about one-half (52 percent) of the divorces in 1987 involved children under 18 years of age and one-fourth (29 percent) involved only one child. It is not surprising that nine-tenths of children under 18 living with a divorced parent live with their mother, far more often in families with smaller average incomes than those living with divorced fathers.

Although only about one-tenth of the adults in the United States in 1988 were divorced (7.4 percent) or separated (2.4 percent), the lifetime experience of married persons with these types of marital disruption is far greater. Based on adjustments for underreporting of divorce data from the Current Population Survey and for underrepresentation of divorce from vital statistics in the MRA, Martin and Bumpass (1989) have concluded that two-thirds of current marriages are likely to end in separation or divorce.

### REMARRIAGE RATES BY AGE

Remarriage rates published by NCHS include separate rates for remarriages after divorce and after widowhood as well as for all remarriages. In this article attention is concentrated on remarriages after divorce, which constitute about nine-tenths of all remarriages. As mentioned above, the remarriage rate after divorce is a measure of the number of divorced women who marry in a given year per 1,000 divorced women at the middle of the year. For example, Table 1 shows that the remarriage rate after divorce for the United States was 133 in 1970 and 76 in 1990. Therefore, about 13 percent of the divorced women in 1970 became remarried in that year, as compared with 8 percent in 1990, only three-fifths as much as in 1970.

Divorced men have far higher remarriage rates than divorced women (in 1990, 106 versus 76). This situation and women's greater longevity largely account for the number of divorced women 45 years old and over in 1994 being one and one-half times the number of divorced men of that age (U.S. Bureau of the Census 1996).

Like first marriage rates, remarriage rates tend to decline with age, and the rate for each age group declined after 1970. The especially sharp drop during the first half of the 1970s for women under 40 years of age resulted from the compounding effect of a rapid increase in divorce and a rapid decline in remarriage during that period.

A large majority of divorced persons eventually remarry. In 1980, among persons 65 to 74 years old, 84 percent of the men and 77 percent of the women who had been divorced had remarried (U.S. Bureau of the Census 1989). But because of the declining remarriage rates, a projection based on information from the National Survey of Families and Households conducted in 1987 and 1988 shows that 72 percent of the recently separated persons are likely eventually to remarry (Bumpass, et al. 1989). About 6 percent of those who become separated never become divorced and therefore are not eligible for remarriage. Two-thirds of the remarriages in recent years occurred to women who entered first marriages as teenagers, according to the same study. Moreover, the rate of remarriage declines as the number of young children increases. Among married parents of young children with one or both parents remarried, about

one-third of the children were born after the remarriage, and the others are stepchildren with the usually accompanying adjustment problems (Glick and Lin 1987).

In a given year about two of every three adults who marry are marrying for the first time, but this includes those marrying after widowhood as well as divorce. Among those who remarry after divorce, about three-fourths have been married only once, one-fifth have been married twice, and one-twentieth three or more times. Some couples who remarry had been married to each other previously. According to unpublished data from the National Survey of Families and Households, this occurred in about 3.3 percent of all marriages and closer to 5 percent for those in the age range when most remarriages after divorce occur.

Men are older than women, on the average, when they marry for the first time or when they remarry. Moreover, the gap is considerably wider at remarriage than at first marriage, but it narrowed somewhat between 1970 and 1987. Thus, successive marriages have been happening at older ages but with shorter intervals between them. In this context, the wider gap at remarriage than at first marriage may be less socially significant than an identical gap would have been at first marriage.

A woman in her second marriage is likely to be married to a man who is about ten years older than her first husband when she married him. Therefore, her second husband was probably more advanced in his occupation than her first husband was when she married him. But research has established that her first husband was probably about as far advanced ten years after her first marriage as her second husband was when she married him (Jacobs and Furstenberger 1986).

### OTHER MARRIAGE AND DIVORCE RATES

Another rate that differs from the rates shown in Table 1 is the *total marriage rate*. This rate is intended to show the number of marriages that group of 1,000 men and women would have if they experienced in their lives the age-sex marriage rates observed in a given year (NCHS 1990b). It is therefore a hypothetical rate analogous to a total fertility rate. Both first marriage rates and remarriage rates have as the base the total population of the

United States without regard to previous marital status.

The total first marriage rate for the United States in 1987 implied that only 69 percent of men and 70 percent of women would eventually marry. The corresponding remarriage rates were 45 percent for men and 41 percent for women. Because both rates are based on the population regardless of marital status, they are additive. Therefore, the (combined) total marriage rate for 1987 implied that men are likely to have 1.14 marriages during their lifetime and women 1.11. These results may seem low because of the assumptions involved. For instance, if currently about 90 percent of every 100 adults marry, if one-half of the first marriages end in divorce, and if 70 percent of the divorced persons remarry, this would mean that 100 young adults in the 1990s are likely to have 90 first marriages and 32 second marriages after divorce ( $.90 + .90 \times .50 \times .70$ ). In addition, many will redi-  
vorce and remarry again, and others will become widowed and remarry. Thus, realistically, the average young adult in the 1990s who marries is likely to have more than 1.1 marriages.

Remarriage among widowed adults is not featured in this article. However, the remarriage rate per 1,000 widows declined from 7 to 5 in 1990, while the rate for widowers declined from 32 to 24. The remarriage rate for widowers is much higher than the rate for widows because about one-half of them marry widows, who outnumber them five to one (U.S. Bureau of the Census 1989). Of all men and women who married in 1987, 68 percent had never previously married, 29 percent were divorced, and only 3 percent were widowed.

Divorce rates per 1,000 involving children under 18 published by NCHS shows that in 1987, the rate was 17, implying that 1.7 percent of the children under 18 years of age in the United States were involved in parental divorces in that year. This finding implies further that, if the same rate continued for eighteen years, 30.6 percent of the children would likely experience parental divorce before they reached 18 years of age. Data from the U.S. Bureau of Census used by Norton and Glick (1986) put the estimate at 40 percent. They also estimated that about an additional 20 percent of children become members of one-parent families because of premarital birth, parental separation

that does not end in divorce, or death of a parent. Therefore, about 60 percent of the children born in the 1990's and later may expect to spend a significant amount of time in a one-parent family before they become 18 years of age.

Marriage rates vary among countries as a reflection of dissimilar social, demographic, and economic conditions. An analysis of marriage rates in fifteen developed countries revealed that between 1965 and 1980 the rate per 1,000 unmarried women declined in all but two of the countries (Glick 1989). The marriage rates in English-speaking countries and Israel were above the average for the entire group, but the rates were below the average in the Germanic, French, and Scandinavian countries. A special reason for the differences was the extent to which cohabitation had been accepted as at least a temporary alternative to marriage. The generally downward trend among the marriage rates shows that the changing social conditions related to the propensity toward marriage have become widely diffused.

Divorce rates per 1,000 married women in almost all of the fifteen countries went up between 1965 and 1980 and doubled in the majority of them. Most of the countries with divorce rates above the overall average were English speaking. High divorce rates tended to be associated with high marriage rates because remarriages after divorce make an important contribution to the level of the marriage rates per 1,000 unmarried women.

Remarriage rates have been falling in most of the fifteen countries. In Canada a part of the decline between 1965 and 1980 was attributed to a change in the divorce laws and an increase in the delay of remarriage. The remarriage rate in New Zealand actually rose slightly in the context of a baby boom and an increase in immigration of young adults with subsequent high divorce rates. Countries with the highest remarriage rates were English speaking, and those with the lowest were Scandinavian countries, France and certain other European countries where cohabitation outside marriage had risen sharply.

Variation in the level of marriage and divorce rates among the American regions and states can be documented only by the use of crude rates. The Northeast and Midwest had consistently lower marriage rates than the South and West. The

general pattern is similar for divorce rates. Some of the variations in the rates by states result from differences in the strictness of residence requirements for obtaining a marriage or divorce license. About seven of every eight marriages occur in the state where both the bride and the groom have their usual residence.

#### SELECTED VARIABLES RELATED TO MARRIAGE AND DIVORCE

Marriage and divorce rates are not presented in the NCHS reports by education and race of those involved. However, some reports do show distributions of first marriages and remarriages by several categories of education and race. The 1987 report documents that those marrying for the first time had more education, on the average, than those who were remarrying. Information from the 1980 census showed that women 25 to 34 years of age who had exactly four years of college training had distinctly the largest proportion of intact first marriages; those with graduate school training had a somewhat smaller proportion; and those with an incomplete college education had a smaller proportion than those in any other education category (U.S. Bureau of Census 1985).

The 1990 report on marriage provided evidence that white adults tend to marry at a younger age than black adults. The difference was two years for both brides and grooms at first marriage and two years at remarriage after divorce. Information for 1988 from the U.S. Bureau of the Census (1989) indicated that 18 percent of black adults had not married by the time they were 40 years old, as compared to 8 percent for white adults of the same age. The pressure to marry and to remain married evidently tends to be less for black adults than for white adults.

Interracial marriage occurs between a small but socially significant proportion of those who marry. In 1987, 2 percent of black brides married white grooms, and 6 percent of black grooms married white brides. Also, nearly 2 percent of black women obtaining divorces were married to white men, and more than 4 percent of black men obtaining divorces were married to white women. Thus, among those who intermarried, the marriages of black men to white women tend to be

more stable than marriages of white men to black women.

Cohabitation outside marriage increased dramatically from only one-half million heterosexual couples in 1970 to 3.5 million in 1993, with about one million of their households maintained by the woman. This numerical growth occurred primarily among adults below middle age and has contributed importantly to the decline in marriage rates as well as to the increase in the number of separated persons. According to the 1987-1988 National Survey of Families and Households, "almost half of the persons in their early thirties and half of the recently married have cohabited" (Bumpass and Sweet 1989, p. 615). In order to provide some balance on the issue, Thornton has concluded that "even though cohabitation will be experienced by many, most people will continue to spend substantially more time in marital unions than in cohabiting unions" (1985, p. 497).

The health of adults is related to marital selection and marital stability, but NCHS does not provide marriage and divorce rates by the health status of those involved. The center does, however, publish current information on several indicators of the health condition of adults by marital status. Nearly all the indicators confirm that people with more signs of good health are likely to marry and remain married or to remarry after marital dissolution (Wilson and Schoenborn 1989). Parental divorce tends to be related to health problems of children but largely through the custodial parents' loss of income and time to spend with the children after divorce (Mauldon 1988). And stress prior to an event such as divorce or premarital breakup may actually reduce the impact of the life transition (Wheaton 1990).

Current trends in marriage and divorce rates do not necessarily indicate whether the rates will tend to stabilize at or near their 1990 levels, to resume their movement in historical directions, or to continue fluctuating in response to future social developments. A few more decades of observing the impact of past changes in marriage and divorce rates on the persons involved may be necessary before a definitive evaluation can be made concerning the longtime effect of these changes on family and child welfare.

(SEE ALSO: *Divorce; Marriage; Remarriage*)

## REFERENCES

- Bumpass, Larry, and Sara McLanahan 1989 "Unmarried Motherhood: Recent Trends, Composition, and Black-White Differences." *Demography* 26:279-286.
- Bumpass, Larry, and James A. Sweet 1989 "National Estimates of Cohabitation." *Demography* 26:615-625.
- Bumpass, Larry, James A. Sweet, and Teresa Castro Martin 1989 "Changing Patterns of Remarriage." Working Paper 89-02. Madison: Center for Demography and Ecology, University of Wisconsin.
- Glick, Paul C. 1989 "The Family of Today and Tomorrow." In K. Ishwaran, ed., *Family and Marriage: Cross-Cultural Perspectives*. Toronto: Wall and Thompson.
- Glick, Paul C., and Sung-Ling Lin 1987 "Remarriage after Divorce: Recent Changes and Demographic Variations." *Sociological Perspectives* 30:162-179.
- Jacobs, Jerry A., and Frank F. Furstenberg, Jr. 1986 "Changing Place: Conjugal Careers and Women's Marital Stability." *Social Forces* 63:714-732.
- Lin, Sung-Ling 1987 "Marital Selection and Child-Bearing and Companionship Functions of Marriage and Remarriage." Ph.D. dissertation, Arizona State University, Tempe.
- Martin, Teresa Castro, and Larry L. Bumpass 1989 "Recent Trends in Marital Disruption." *Demography* 26:37-51.
- Mauldon, Jane 1988 "The Effect of Marital Disruption on Children's Health." Paper presented at the Population Association of America meeting, New Orleans. National Center for Health Statistics (NCHS) 1990a "Births, Marriages, Divorces, and Deaths for 1989." *Monthly Vital Statistics Report* 38, no. 12.
- 1990b "Advance Report of Final Marriage Statistics, 1987." *Monthly Vital Statistics Report* 38, no. 12, Supplement.
- 1990c "Advance Report of Final Divorce Statistics, 1987." *Monthly Vital Statistics Report* 38, no. 12, Supplement 2.
- 1995a "Advance Report on Final Divorce Statistics, 1989 and 1990." *Monthly Vital Statistics Report* 43, No. 9, Supplement.
- 1995b "Advance Report on Final Marriage Statistics, 1989 and 1990." *Monthly Vital Statistics Report* 43, No. 12, Supplement.
- 1998 "Births, Marriages, Divorces, and Deaths for 1997." *Monthly Vital Statistics Report* 46, No. 12.
- Norton, Arthur J., and Paul C. Glick 1988 "One-Parent Families: A Social and Economic Profile." *Family Relations* 35:9-17.
- Norton, Arthur J., and Jeanne E. Moorman 1987 "Current Trends in Marriage and Divorce among American Women." *Journal of Marriage and the Family* 49:3-14.
- Schoen, Robert, William Urton, Karen Woodrow, and John Baj 1985 "Marriage and Divorce in Twentieth-Century American Cohorts." *Demography* 22:101-114.
- Thornton, Arland 1988 "Cohabitation and Marriage in the 1980s." *Demography* 25:492-508.
- U.S. Bureau of the Census 1985 *1980 Census of Population: Marital Characteristics*. Washington: U.S. Government Printing Office.
- 1996 "Marital Status and Living Arrangements: March 1994" *Current Population Reports* P-20, no. 484 Washington: U.S. Government Printing Office.
- Wheaton, Blair 1990 "Life Transitions, Role Histories, and Mental Health." *American Sociological Review* 55:209-223.
- Wilson, Barbara F., and Kathryn A. London 1987 "Going to the Chapel." *American Demographics* 9:26-31.
- , and Charlotte Schoenborn 1989 "A Healthy Marriage." *American Demographics* 11:40-43.

PAUL C. GLICK

## MARXIST SOCIOLOGY

The concept of a Marxist sociology does not refer to a clearly defined approach to social research; indeed, it is "now employed so widely that it has begun to lose all meaning" (Abercrombie et al. 1988, p. 148). The ambiguity of the term stems from the multiplicity of interpretations of the work of Karl Marx and Friedrich Engels, whose approach to social theory is usually termed *historical materialism*. The most well-known version is that of Soviet communism (Marxism-Leninism) and is identified with the worldview of *dialectical materialism* (a term never used by Marx and Engels), which has served largely the interests of Soviet ideology. The influence of historical materialism in modern sociology, however, stems primarily from the lesser-known independent tradition of European "Western Marxism" and the resulting forms of Marxist sociology (Agger 1979).

Contemporary usage of the term *Marxist sociology* varies considerably. In the United States, for example, the term *Marxist* is often used rather

loosely, to designate virtually any type of radical or critical approach influenced by Marxian concepts (e.g., Ollman and Vernoff 1982; Flacks 1982). In societies with more strongly developed social democratic labor movements, as in Europe, the term is more closely identified with communist parties. Given the inherent ambiguity of the term, it is therefore useful to define Marxist sociology rather narrowly and concretely as a specific form of conflict theory associated with Western Marxism's objective of developing a positive (empirical) science of capitalist society as part of the mobilization of a revolutionary working class.

It is also useful to distinguish three basic types of relations between sociologists and Marxist sociology: those who work directly within the Marxist tradition (Marxist sociology proper) but "incorporate sociological insights, findings, and methodologies"; those who are Marxist-influenced in the sense of being stimulated by its historical approach and the "big questions" Marxists have posed but remain indifferent to "whether the best explanatory answers turn out to be Marxist" (Burawoy and Skocpol 1982, p. vii); and those identifying with highly revisionist critical theories (sometimes still in the name of Marxism) that seek to preserve the emancipatory vision of the Marxist tradition despite abandonment of the conventional notion of working-class revolution (Held 1980; Kellner 1989). It should also be stressed that Marxist sociology in this first sense refers to a historically identifiable—but widely contested—interpretation of the sociological implications of Marx's approach. As the most well-known British Marxist sociologist has concluded with particular reference to the German Frankfurt tradition of critical theory: "The tasks of a Marxist sociology, as I conceive it, are therefore very different from those of a neo-critical theory of society" (Bottomore 1984, p. 81). Marxist sociology in this strict sense thus tends in its most rigid form to resist any eclectic appropriation of sociological concepts: "Marxism has been courted by virtually every conceivable non-Marxist ideology: by existentialism, phenomenology, critical academic sociology, and by several variants of theology. To raise the question of a Marxism of Marxism is to take a resolute stand against all attempts to capture and exploit Marx for non-Marxist purposes; and to adopt as a guiding principle . . . the claim that Marx himself made for his work: that Marxism is a specific science, related to the

working class as a guide to socialist revolution" (Therborn 1976, p. 40).

The distinction between Marxism as "science" and as "critique" provides another way of describing the aspirations of Marxist sociology, which most commonly seeks to develop an objective, political economic science of society rather than a critical philosophy of praxis (Bottomore 1975; Gouldner 1980). Such a project is inherently interdisciplinary and often referred to under the heading of *political economy*, a term designating Marxist-oriented research that may be carried out in various disciplines: economics, political science, and history, as well as sociology (Attewell 1984). Though those identifying with neo-Weberian conflict or critical theories accept many of the empirical findings of neo-Marxist sociology and political economy, they tend to disagree about their broader interpretation and relation to political practice and social change.

## ORIGINS

All forms of Marxist sociology trace themselves to the general theoretical approach of historical materialism as developed by Marx and Engels (Marx and Engels 1978; Bottomore and Goode 1983; Bottomore et al. 1983). Standard accounts of Marx's theoretical program stress that it does not constitute a unified system so much as diverse, though interrelated, modes of theorizing.

- Early writings that outline a theory of philosophical critique, an analysis of alienated labor, and a normative vision of human emancipation;
- A general sociology in the form of historical materialism (i.e., a theory of modes of production) as an approach to historical evolution;
- A specific account of capitalism and its economic contradictions deriving from this general theory; and
- A political philosophy and theory of praxis concerned with translating objective crisis tendencies in capitalism into a revolutionary transformation that would bring about a new form of "socialist" and eventually



“communist” society (Giddens 1971, pp. 1–64; Bottomore 1975).

The precise relationship among these areas remains controversial. Though each level of theorizing has sociological implications, Marxist sociology has drawn primarily from the general sociology suggested by historical materialism and the more historically specific analysis of capitalist development that is the empirical focus of Marx’s social theory and historical sociology. The key concept of a mode of production serves as the comparative framework for analyzing different social formations in terms of the contradictory relationship between their forces of production (primarily technology) and relations of production (forms of work organization and exploitation). The resulting mode of production directly shapes the specific structures of the class system and the manner in which the economic base determines the cultural superstructure composed of the state and various ideological institutions such as the mass media, education, law, religion, political ideologies, and so forth. In capitalist societies the contradiction deriving from the unresolvable polarization between labor and capital becomes the basis for revolutionary change under conditions of economic crisis.

As a positive science of society, Marxist sociology emerged in the period following Marx’s death in 1883 through World War I. Strongly influenced by Engels’s conception of “scientific socialism” as elaborated by Karl Kautsky in Germany, it was institutionalized in German Social Democracy and the Second International. Particular stress was placed upon how this approach provided an account of the historical “laws” that explained the causes of changes in modes of production and class formation and struggle. According to these laws a transition to socialism could be deduced from the “necessary” breakdown of capitalism. Extensive further development of such a scientific socialism was carried out by the Austro-Marxists, who deepened the logical analysis of Marxism as a form of causal explanation by drawing upon contemporary debates in the philosophy of science; as well, they extended Marx’s theory to new phenomena such as the analysis of nationalism and the ethical foundations of Marxist sociology (Bottomore 1975, 1978; Bottomore and Goode 1978).

These two traditions of scientific socialism, however, did not develop much beyond the level codified in Bukharin’s textbook of 1921 (*Historical Materialism: A System of Sociology*, translated in 1925). The primary reason was that there was considerable resistance to Marxism in the academy (related in part to identification of Marxism with the Soviet Union), ignorance of the richness of the suppressed tradition of Western Marxism, and a broad post-World War II institutionalization of sociology that virtually excluded Marxist sociology, despite some marginal influences of research in social stratification and change (e.g., Ralf Dahrendorf in West Germany and C. Wright Mills in the United States). For all practical purposes the resurgence of Marxist sociology (often identified as “political economy” as opposed to “critical theory”) coincides with the parallel emergence of radical and critical theories in the late 1960s, along with the recovery of the deeper foundations of Marxian theory with the proliferation of translations of Western Marxist texts in the 1970s.

It is customary to distinguish two basic starting points of a Marxist sociology based on different interpretations of the base-superstructure metaphor that underlies the concept of a mode of production: economistic or instrumental approaches as opposed to structuralist reproduction models. Economistic interpretations—sometimes associated with the idea of orthodox or “vulgar” Marxism—are based on a more or less reductionistic, causal account of the effects of the economic base or infrastructure upon the cultural superstructure, especially as causally derived from the assumed objective consequences of class interests.

Structuralist theories of social and cultural reproduction, in contrast, argue that the base-superstructure relation is more complex, involving functional relations that ensure the relative (if variable) autonomy of cultural factors, even if the economic is determinant in the last instance. Such structuralist interpretations derive primarily from the concept of the social reproduction of labor power developed in Marx’s later works, the theory of cultural hegemony developed by the Italian Marxist leader and theorist Antonio Gramsci (1971) in the 1930s, and the reinterpretation of both Marx and Gramsci by the French philosopher Louis Althusser in the 1960s (Althusser 1971; Althusser and Balibar [1968] 1979).

### CONTEMPORARY THEMES IN MARXIST SOCIOLOGY

The most convenient way to speak of contemporary themes in Marxist sociology—thus differentiating it from conflict or critical theory generally—is to restrict the concept to research that continues to adhere to the basic principles of neo-Marxist theory. Such a Marxist sociology would sometimes include adherence to the labor theory of value (which holds that labor is the only source of profit) but more essentially the primacy of economic and class factors, the priority of objective structures over subjectivity and consciousness, and the privileged role of the working class in a transition to socialism as defined by direct state ownership of the means of production (Anderson 1984; Wood 1986; Archibald 1978).

With respect to more recent developments in Marxist sociology, the immense literature and wide national variations make it appropriate to focus primarily—if not exclusively—on the remarkable resurgence of Marxism in American sociology. This phenomenon stems from the 1970s and has been traced to four key influences: the broadened audience of the journal *Monthly Review*, which was founded in 1949 and pioneered the application of Marxist economic theory for an analysis of the United States and its “imperial” role in world politics; Marxist historians who developed a critique of American liberalism and proposed a class-based reinterpretation of American history; the Hegelian Marxism and Critical Theory of the Frankfurt School tradition; and, finally, structuralist Marxism of largely French inspiration that sought to reestablish the credentials of Marx’s theory as a science of society (Burawoy 1982, pp. 4–6). To illustrate the concerns of recent Marxist sociology, it is instructive to review some of the most representative examples of research in four key areas: work and the division of labor, class structure, the state and crisis theory, and culture and ideology. Research on the political economy of the world system and dependency theory have also been important but will not be discussed here (see, however, So 1990).

**The Labor Process and the New International Division of Labor.** Until the late 1960s, Marxist theory was associated primarily with either communist ideologies or economic theory (Sweezy [1942] 1968). Not surprisingly, the pioneering,

technically sophisticated Marxist research in the United States was concerned with an economic analysis of the new form of “monopoly capitalism” (Baran and Sweezy 1966). Drawing out the sociological implications of such, political economy is most closely associated with subsequent pathbreaking work involving the rediscovery of the labor process (i.e., Braverman 1974; Burawoy 1979) that gave work the central place that had been lost with the sense of consensus in postwar labor relations (Thompson 1983; Attewell 1984, pp. 93–141). Particular attention was given to the logic of capital’s need to cheapen labor costs in ways that degrade and divide workers. More recent work has attempted to emphasize the incorporation of a subjective dimension to the labor process; further, it has developed a comparative perspective through the analysis of factory regimes in different types of economic systems and in relation to the new international division of labor (e.g., Burawoy 1985).

**Class.** Class analysis of course remains the key aspect of any Marxist sociology, but the focus contrasts sharply with conventional sociological approaches, even those (e.g., Max Weber) that acknowledge the importance of class conflict. Marxist sociology strongly insists on the primacy of the relations of production over the market processes stressed by neo-Weberians (Grabb 1990). The most influential empirical research in this area has stressed the importance of contradictory class locations and of reconnecting the objective and subjective dimensions of class with a theory of exploitation (Grabb 1990, pp. 152–163; Wright 1978, 1985).

Recent Marxist class analysis has been confronted by a number of challenges that pose serious problems for orthodox approaches; for example, the problem of the urban question, the role of ethnicity or race and especially gender as independent sources of domination (Shaw 1985), the emergence of the middle strata and the decline of the traditional “working class” (Walker 1978), the failure of class consciousness and actions to develop in the ways required for revolutionary transition, and other issues. The relation between Marxist and feminist theory has proved most controversial. More orthodox Marxist sociologists have attempted to incorporate gender into the theory of class and modes of production through the concept of unpaid “domestic labor” that contributes to the overall process of social reproduction (Fox

1980). But many feminists have abandoned Marxist sociology precisely because of its insistence on the primacy of class at the expense of gender (and other sources of domination).

**The State and Crisis Theory.** The contributions of Marxist sociology to the theory of the state have been wide-ranging and influential (Carnoy 1984; Jessop 1982; Holloway and Picciotto 1978). As well, they illustrate most clearly the issues involved in the debate between instrumental and structuralist interpretations of the base-superstructure model. In the context of theories of the state, this issue takes the form of whether the dominant class controls the state directly through its elite connections or indirectly through the functional economic and political imperatives that constrain public policy, regardless of who happens to hold power in a democratic regime. For example, the so-called Miliband-Poulantzas debate sharply defined the different empirical consequences of these two approaches. Miliband ([1969] 1973) stressed the actual empirical link between economic elites and political power of the dominant class, hence the role of the state as an instrument of class rule. Poulantzas ([1968] 1978), on the other hand, analyzed the state as a factor of cohesion that requires relative autonomy in order indirectly to serve the process of social reproduction in the long run.

Another central theme of Marxist sociology has been the relationship between economic crisis tendencies and the state in advanced capitalism (Attewell 1984, pp. 142–206). Research has focused especially on the concept of the “fiscal crisis of the state” (O’Connor 1973). It has been argued that the state is caught between the contradictory pressures of ensuring capital accumulation and legitimating the negative effects of the economy with the safety nets of the welfare state and that these conflicts become the potential basis for the emergence of new class-based oppositional movements.

**Culture and Ideology.** The most recent flourishing area of research has been in Marxist-influenced analyses of cultural phenomena as manifestations of ideology. The central focus has been on how cultural hegemony (or domination) is formed and the types of resistance that oppressed groups may develop against it. The stress of economic Marxist has been upon the way in which the cultural superstructure of society “reflects” or

“mirrors” economic processes and class relations. This approach often resulted in crude, reductionistic political and class analyses that were often unsatisfactory to those intimately acquainted with both high and popular culture. With the emergence of more sophisticated structuralist models of cultural reproduction, however, the relative autonomy of cultural forms could be acknowledged without obscuring their origins in “material” social relations. In the more extreme form represented by structuralist Marxism, it has been more generally held that all of the cultural institutions of society (e.g., the media, family, law, arts, etc.) functioned in the last instance as “ideological state apparatuses” that served the long-term interests of capital (Althusser 1971). Research based on both instrumentalist and structuralist approaches has been applied to the range of cultural activities (e.g., art, literature, law, sport, etc.), but education and the mass media figure most prominently, and they can serve as illustrations.

In the case of the Marxist sociology of education the result was a shift from an instrumentalist perspective (i.e., the role of capitalist ideology in directly using the educational system to shape consciousness in its interests) to one based on the idea of social reproduction as an indirect form of social control. Initially, structuralist approaches put particular stress upon the formal “correspondence” between the economic base and the hegemonic superstructure, despite the autonomy of the latter. Hence, structuralist research on education attempted to demonstrate the way in which the hidden curriculum of the school “corresponded” to the type of labor required by capital (Bowles and Gintis 1977; Cole 1988). Research on the political economy of the media was more strongly represented by instrumentalist perspectives that stressed the role of the media as instruments of “mind control” on the part of the dominant class and the broader dominance of the American media globally (Schiller 1971; 1973); others argued from a more structuralist perspective that the primary function of the media is the “selling of audiences,” irrespective of the specific ideological content of programming (Smythe 1981).

#### THE CRISIS OF MARXIST SOCIOLOGY

As a specific theoretical approach that seeks to discover the role of economic and class factors in

social change, Western Marxist sociology will certainly endure, though its significance will vary with the type of social formation and topic examined. As a philosophy of history or general theory of modes of production, and more especially as part of a particular theory of working-class revolution, orthodox Marxist sociology has been seriously called into question, especially in advanced capitalism. It has already collapsed in Eastern bloc countries, where a completely new tradition of social science is in the process of formation. This is not to say that economic and class factors or Marx become irrelevant, though the broader crisis of historical materialism suggests their significance and relation to social, political, and cultural processes will have to be interpreted in more self-critical, flexible, and historically specific ways (Aronowitz 1981). The complex outcome of the crisis of Marxist sociology in advanced capitalism is suggestively anticipated in the response of three contemporary countertendencies.

So-called analytical Marxism is defined more by its methodological stance than its substantive content. It thus differentiates itself from traditional Marxism in its commitment to abstract theorizing (as opposed to more concrete historical analysis), a search for rethinking the foundations by asking heretical questions, and “using state-of-the-art methods of analytical philosophy and ‘positivist’ social science” (Roemer 1986, pp. 1–2). Though these developments will undoubtedly have some impact on social theory, they are clearly too heterogeneous and revisionist to fall under the heading of Marxist sociology in the sense used here.

A second, opposing “poststructuralist” strategy is evident in the work of some former Marxists who have retreated from orthodox class concepts, arguing that a “post-Marxism” is required that involves eliminating the notion of the working class as a “universal class” and resurrecting a new conception of socialist democracy (Laclau and Mouffe 1985). Not surprisingly, this (partial) “retreat from class” has been treated with hostility by many neo-Marxists (Wood 1986), but it has provoked important debates about the role of new social movements and democratic processes in any defensible conception of socialist transformation.

A third tendency has been loosely referred to as “cultural Marxism.” Such critics of the functionalist

tendencies of structuralist Marxism put particular stress upon the contested and uneven character of cultural reproduction in capitalist societies. In particular, various researchers have pointed to how dominated groups resist cultural domination in ways that often become the basis of counterhegemonic social movements. Further, it is argued that a crucial feature of contemporary “postmodern” societies is the distinctive role of the “cultural.” The result has been a flowering of cultural research often identified, especially in the British context, with the notion of “cultural studies” (e.g., the work of Raymond Williams, Richard Johnson, Stuart Hall, et al.; see Brantlinger 1990). A related tendency has been the cultural Marxist historiography of E. P. Thompson (Kaye and McClelland 1990) that has influenced a major reinterpretation of Marx as a historical sociologist (Sayer 1983; Corrigan and Sayer 1985). A distinctive aspect of the cultural Marxist tradition, an aspect that has led it away from Marxist sociology in its more restrictive sense, has been its ability actively to engage in debate with and appropriate concepts from a wide variety of non-Marxist approaches. Much of the recent work carried out in the name of cultural Marxism thus increasingly blends with poststructuralist and critical theories of culture, reflecting the circumstance that “Marxism is no longer a single coherent discursive and political practice” (Nelson and Grossberg 1988, p. 11). One consequence is that it is no longer “possible to talk unproblematically of a ‘Marxist’ sociology, since Marxism has become a major contributor to sociology in general, while Marxist-influenced sociologists increasingly identify with their discipline (Shaw 1985, p. 16).

#### **THE 1990S: THE DECLINE OF MARXIST SOCIOLOGY AND RECONSTRUCTIONS OF MARXIAN CRITICAL THEORY**

By the end of the 1990s, several key shifts had redefined the fate of Marxist theory in the social sciences: (1) the collapse of the Soviet bloc; (2) a proliferation of metatheoretical debates—loosely grouped as postmodernist—that called into question the capacity of theory generally, and Marxist theory in particular, to provide general explanations or ethical critiques of social life; (3) a blurring of disciplinary boundaries that contributed to the

flow of Marxist-related concepts into the humanities, especially under the heading of cultural studies; and (4) a substantive focus on globalization as a framework for rethinking the Marxian theory of society and change. The outcome of these tendencies was a decline of Marxist sociology as a clearly identifiable research orientation. Yet these developments also reinforced the three tendencies previously visible in the 1980s: (1) the continuing development of analytical Marxism (Jacobs 1996); (2) the proliferation of poststructuralist critiques of class essentialism that attempted to include standpoints such as race and gender in Marxist theory; and above all, (3) the continuing development of cultural Marxism, especially in the humanities. Many of those who continued to develop such questions had strayed so far from classical Marxist positions that they more often identified with an expanded notion of “critical theory” than with “Marxism” in the strict sense; rather than referring primarily to the Frankfurt School tradition, such an ecumenical reference to critical theory reflected a shared aspiration for a fundamental rethinking of the Marxian tradition.

The disintegration of the Soviet bloc culminating in 1989 marked the end of the failed experiment of “actually existing socialism.” Though Western Marxists had generally rejected the Soviet model as an expression of Marx’s utopian vision, its sudden demise came as a shock and unleashed a complex theoretical debate (Magnus and Cullenberg 1995). Western Marxists greeted this momentous historical transformation with mixed emotions. On the one hand, they were relieved to see the end of this tragic example of state socialism. On the other hand, they were perturbed by the uncritical embracing of the market model of development as the only alternative and the triumphalist heralding of these changes as the “end of Marxism.” Within the ex-Soviet bloc the disillusionment with Marx was so intense that the very idea of Marxist sociology was generally repudiated; partly as a consequence the “Marxist” interpretation of these events (e.g. the important work of Michael Burawoy) has been undertaken, paradoxically, by Western Marxists (Kennedy and Galtz 1996). In Latin America, though Marx remained in influential intellectual force, political debate shifted from revolutionary rhetoric toward pragmatic questions of “democratic transition” (Castañeda 1993). Though many Marxist social scientists moved toward variants of

post-Marxist critical theory in response to these events, others insisted that the collapse of these regimes did not constitute “proofs of the bankruptcy of Marxism as a tradition of social scientific practice,” though conceding that “it must be reconstructed in various ways” (Wright 1996, pp. 121–122). Though Marxist sociology declined precipitously as a specific orientation in the 1990s—hence its absence from leading social theory anthologies, discussion of Marxist theory continued to flourish in a few journals, most notably the *New Left Review* in Britain and *Rethinking Marxism* in the United States. Despite sporadic attempts to defend Marxism as a science (Burawoy 1990) or “test” classic Marxist claims (e.g., Boswell and Dixon 1993; Cockshott et al. 1995; Smith 1994), the revitalization of empirical Marxist research appeared increasingly dependent upon working out the implications of major theoretical reconstructions (e.g., Postone 1993; Wright et al. 1992). Such efforts have also been influenced by “Weberian Marxism,” despite the resistance of more orthodox Marxists (Gubbay 1997; Lowy 1996).

A second major shift that has contributed to the decline of Marxist sociology in the 1990s can be identified with the elusive concept of “postmodernism” and related poststructuralist tendencies. The Marxist tradition was a specific target of epistemological postmodernism which called into question all “meta-narratives” or grand theories of society and history, whether as explanations or as the basis of universalistic ethical claims. Related poststructuralist critiques of essentialism challenged Marxist theories of the working class and called for a decentered conception of the subject that acknowledged the diversity of subject positions, a theme especially central to feminist, gender, and racial research. Some Marxists simply defended classical Marxist theory against these developments, but many felt compelled to revise their positions in the direction of an “emancipatory postmodernism,” a “postmodern Marxism,” or a “poststructuralist materialism.” In this context, various efforts were made to reconcile selective poststructuralist and postmodernist arguments with a reconstructed neo-Marxist, critical theory (e.g., Best and Kellner 1997). On the one hand, such postmodernist-influenced approaches acknowledged that the fundamental cultural shifts linked to the mass media and new information technologies have profound implications for a “postmodern”

society that was not foreseen by Marx. On the other, they also conceded that classic Marxist theory (though not necessarily Marx himself) relied on an inadequate, teleological philosophy of history, economic reductionism, and an essentialist conception of the working-class subject.

A third tendency that has become more clearly defined over the past decade has been the partial erosion of disciplinary boundaries and the diverse impact of Marxist theorizing on the humanities, especially under the heading of cultural studies and poststructuralist textual theories. Paradoxically, though the influence of Marxist theory declined within sociology in the 1990s, it had a continuing renaissance and diffuse effect in other disciplines, especially literary studies. Much of this work has in turn influenced cultural Marxism, as well as sociologists of culture and the media. In its most ambitious form, such neo-Marxist cultural theory has culminated in an account of postmodernity as an historical epoch (Jameson 1991). Another outcome was an expansion and revision of the canons of classical Marxian theory. Though his status as a “Marxist” has been disputed, the Soviet scholar Mikhail Bakhtin (1895–1975) was rehabilitated as a major theorist of culture whose dialogical theory of language anticipated many of the later developments in poststructuralist critiques of Marxism (Bell and Gardiner 1998). Similarly, Antonio Gramsci’s cultural and linguistic writings have been reinterpreted in terms of his poststructuralist anticipations (Holub 1992).

Finally, the most sustained empirical development in the 1990s was the emergence of the problematic of globalization (and aspects of postcolonial theory) as a reference point for an historical materialist account of social change (Hoogvelt 1997; Larrain 1994). From this perspective, “postmodernism” could be read as a symptom of processes better understood from the perspective of a political economy of globalization. Hence it was argued that Marx was the pioneer of globalization theory, given his account of the ceaseless worldwide expansion of markets, a theme elaborated in Wallerstein’s world-system theory. But many other Marxists have joined critical theorists in emphasizing the distinctiveness of more recent developments since World War II, especially the transformation of production processes (e.g., the shift from “Fordism” to “post-Fordism”) and the impact of information technologies, resulting

in changed relations between space and time (Harvey 1989). Unlike those who uncritically celebrated the advent of an “information society” and globalization as the inevitable outcomes of capitalism, such Marxists and critical theorists have pointed to the selective characteristics of this “modernization” and its inadequate vision of the alternatives (Webster 1995). The most elaborate empirical effort to describe this new “informational mode of production” in Marxist-influenced but also innovative sociological terms can be found in Manuel Castells’ three-volume *The Information Age* (Castells 1996–1998). As against various postmodernist currents, neo-Marxist conceptions of globalization (and related variants of critical theory) are united by their political opposition to neoliberalism (or what is often referred to as neoconservatism in the United States) and an empirical concern with analyzing capitalist globalization in terms of its grave implications for marginalized groups and societies, as well as nature itself. For the most part, such researchers no longer appeal to the classic conception of working-class revolution, though they hold out hope for new social movements as the basis of constructive forms of resistance and democratic renewal. As a consequence, Marxist sociologists have increasingly moved toward positions previously staked out by the Frankfurt tradition of critical theory and related poststructuralist French developments (Morrow 1994), though often without consistently acknowledging the full implications for the “death of Marxism” as an oppositional political discourse (Fraser 1998).

(SEE ALSO: *Critical Theory*; *Macrosociology*; *Materialism*; *Socialism*)

#### REFERENCES

- Abercrombie, Nicholas, Stephen Hill, and Bryan S. Turner 1988 *The Penguin Dictionary of Sociology*, 2nd ed. Harmondsworth, England: Penguin.
- Agger, Ben 1979 *Western Marxism: An Introduction—Classical and Contemporary Sources*. Santa Monica, Calif.: Goodyear.
- Althusser, Louis 1974 “Ideology and Ideological State Apparatuses.” In *Lenin and Philosophy and Other Essays*, Ben Brewster, trans. New York: Monthly Review.
- , and Etienne Balibar (1968) 1979 *Reading Capital*. Ben Brewster, trans. London: Verso.
- Anderson, Perry 1984 *In the Tracks of Historical Materialism*. London: Verso.

- Archibald, W. Peter 1978 *Social Psychology as Political Economy*. Toronto: McGraw-Hill.
- Aronowitz, Stanley 1981 *The Crisis of Historical Materialism: Class, Politics, and Culture in Marxist Theory*. New York: Praeger/J. F. Bergin.
- Attewell, Paul A. 1984 *Radical Political Economy since the Sixties: A Sociology of Knowledge Analysis*. New Brunswick, N.J.: Rutgers University Press.
- Baran, Paul A., and Paul M. Sweezy 1966 *Monopoly Capital: An Essay on the American Economic Order*. New York: Monthly Review.
- Bell, Michael Mayerfeld, and Michael Gardiner, eds. 1998 *Bakhtin and the Human Sciences: No Last Words*. London: Sage.
- Best, Steve, and Douglas Kellner 1997 *The Postmodern Turn*. New York: Guilford.
- Boswell, Terry, and William J. Dixon 1993 "Marx's Theory of Rebellion: A Cross-National Analysis of Class Exploitation, Economic Development, and Violent Revolt." *American Sociological Review* 58 (October:681-702).
- Bottomore, Tom 1975 *Marxist Sociology*. London: Macmillan.
- , 1978 "Marxism and Sociology." In Tom Bottomore and Robert Nisbet, eds., *A History of Sociological Analysis*. New York: Basic.
- 1984 *The Frankfurt School*. New York: Ellis Horwood/Tavistock.
- , and Patrick Goode, eds. 1978 *Austro-Marxism*, T. Bottomore and P. Goode, trans. Oxford: Clarendon.
- 1983 *Readings in Marxist Sociology*. Oxford, England: Oxford University Press.
- Bottomore, Tom, et al., eds. 1983 *A Dictionary of Marxist Thought*. Cambridge, Mass.: Harvard University Press/Basil Blackwell.
- Bowles, Samuel, and Herbert Gintis 1977 *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic.
- Brantlinger, Patrick 1990 *Crusoe's Footprints: Cultural Studies in Britain and America*. New York: Routledge.
- Braverman, Harry 1974 *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review.
- Burawoy, Michael 1979 *Manufacturing Consent: Changes in the Labor Process under Monopoly Capitalism*. Chicago: University of Chicago Press.
- 1982 "Introduction: The Resurgence of Marxism in American Sociology." In Michael Burawoy and Theda Skocpol, eds., *Marxist Inquiries: Studies of Labor, Class, and States*. Chicago: University of Chicago Press.
- 1985 *The Politics of Production*. London: Verso.
- 1990 "Marxism as Science: Historical Challenges and Theoretical Growth." *American Sociological Review* 55:775-793.
- , and Theda Skocpol, eds. 1982 *Marxist Inquiries: Studies of Labor, Class, and States*. Chicago: University of Chicago Press.
- Castañeda, Jorge G. 1993 *Utopia Unarmed: The Latin American Left after the Cold War*. New York: Knopf.
- Castells, Manuel 1996-1998. *The Information Age: Economy, Society and Culture*, 3 vols. Oxford, England: Basil Blackwell.
- Cockshott, Paul, Allin Cottrell, and Greg Michaelson 1995 "Testing Marx: Some New Results from UK Data." *Capital and Class* 55:103-129.
- Cole, Mike, ed. 1988 *Bowles and Gintis Revisited: Correspondence and Contradiction in Educational Theory*. New York: Falmer.
- Corrigan, Philip, and Derek Sayer 1985 *The Great Arch: English State Formation as Cultural Revolution*. Oxford, England: Basil Blackwell.
- Flacks, Richard 1982 "Marxism and Sociology." In Bertell Ollman and Edward Vernoff, eds., *The Left Academy: Marxist Scholarship on American Campuses*. New York: McGraw-Hill.
- Fox, Bonnie, ed. 1980 *Hidden in the Household: Women's Domestic Labour under Capitalism*. Toronto: Women's Press.
- Fraser, Nancy 1998 "A Future for Marxism." *New Politics (New Series)* 6 (Winter):95-98.
- Giddens, Anthony 1971 *Capitalism and Modern Social Theory*. Cambridge, England: Cambridge University Press.
- Gouldner, Alvin W. 1980 *The Two Marxisms: Contradictions and Anomalies in the Development of Theory*. New York: Seabury.
- Grabb, Edward G. 1990 *Theories of Social Inequality: Classical and Contemporary Perspectives*, 2nd ed. Toronto: Holt, Rinehart and Winston.
- Gramsci, Antonio 1971 *Selections from the Prison Notebooks*, Quintin Hoare and Geoffrey Nowell Smith, eds. and trans. New York: International.
- Gubbay, Jon 1997 "A Marxist Critique of Weberian Class Analyses." *Sociology* 31:73-89.
- Harvey, David 1989 *The Condition of Postmodernity*. Oxford, England: Basil Blackwell.
- Held, David 1980 *Introduction to Critical Theory: Horkheimer to Habermas*. Berkeley: University of California Press.

- Holloway, John, and Sol Picciotto, eds. 1978 *State and Capital: A Marxist Debate*. London: Edward Arnold.
- Holub, Renate 1992 *Antonio Gramsci: Beyond Marxism and Postmodernism*. New York: Routledge.
- Hoogvelt, Ankie 1997 *Globalization and the Postcolonial World: The New Political Economy of Development*. Baltimore, Md.: Johns Hopkins University Press.
- Jacobs, Lesley A. 1996 "The Second Wave of Analytical Marxism." *Philosophy of the Social Sciences* 26:279–292.
- Jameson, Fredric 1991 *Postmodernism, or, The Cultural Logic of Late Capitalism*. Durham, N.C.: Duke University Press.
- Jessop, Bob 1982 *The Capitalist State: Marxist Theories and Methods*. Oxford: Martin Robertson.
- Kaye, Harvey J., and Keith McClelland, eds. 1990 *E. P. Thompson: Critical Perspectives*. Philadelphia: Temple University Press.
- Kellner, Douglas 1989 *Critical Theory, Marxism, and Modernity*. Baltimore, Md.: Johns Hopkins University Press.
- Kennedy, Michael D, and Naomi Galtz 1996 "From Marxism to Postcommunism: Socialist Desires and East European Rejections." *Annual Review of Sociology* 22:437–458.
- Laclau, Ernest, and Chantal Mouffe 1985 *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*, Winston Moore and Paul Cammack. trans. London: Verso.
- Larrain, Jorge 1994 *Ideology and Cultural Identity: Modernity and the Third World Presence*. Cambridge: Polity.
- Lowy, Michael 1996 "Figures of Weberian Marxism." *Theory and Society* 25 (June):431–446.
- Magnus, Bernd, and Stephen Cullenberg 1995 *Whither Marxism? Global Crises in International Perspective*. New York: Routledge.
- Marx, Karl, and Friedrich Engels 1978 *The Marx-Engels Reader*, 2nd ed. Robert C. Tucker, ed. New York: Norton.
- Miliband, Ralph (1969) 1973 *The State in Capitalist Society: The Analysis of the Western System of Power*. London: Quartet.
- Morrow, Raymond A. 1994 *Critical Theory and Methodology*. Newbury Park, Calif.: Sage.
- Nelson, Cary, and Lawrence Grossberg, eds. 1988 *Marxism and the Interpretation of Culture*. Urbana, Ill.: University of Illinois Press.
- O'Connor, James 1973 *The Fiscal Crisis of the State*. New York: St. Martin's.
- Ollman, Bertell, and Edward Vernoff, eds. 1982 *The Left Academy: Marxist Scholarship on American Campuses*. New York: McGraw-Hill.
- Postone, Moishe 1993 *Time, Labor, and Social Domination: A Reinterpretation of Marx's Critical Theory*. New York: Cambridge University Press.
- Poulantzas, Nicos (1968) 1978 *Political Power and Social Classes*. Timothy O'Hagan, trans. London: Verso.
- Roemer, John, ed. 1986 *Analytical Marxism*. Cambridge, England: Cambridge University Press.
- Sayer, Derek 1983 *Marx's Method: Ideology, Science, and Critique in "Capital"*. Atlantic Highlands, N.J.: Harvester/Humanities.
- Schiller, Herbert I. 1971 *Mass Communications and the American Empire*. Boston: Beacon.
- 1973 *The Mind Managers*. Boston: Beacon.
- Shaw, Martin, ed. 1985 *Marxist Sociology Revisited: Critical Assessments*. London: Macmillan.
- Smith, Murray E.G. 1994 *Invisible Leviathan: The Marxist Critique of Market Despotism beyond Postmodernism*. Buffalo, N.Y.: University of Toronto Press.
- Smythe, Dallas W. 1981 *Dependency Road: Communications, Capitalism, Consciousness, and Canada*. Norwood, N.J.: Ablex.
- So, Alvin Y. 1990 *Social Change and Development: Modernization, Dependency, and World-System Theories*. Newbury Park, Calif.: Sage.
- Sweezy, Paul M. (1942) 1968 *The Theory of Capitalist Development: Principles of Marxian Political Economy*. New York: Monthly Review.
- Therborn, Göran 1976 *Science, Class, and Society: On the Formation of Sociology and Historical Materialism*. London: New Left.
- Thompson, Paul 1983 *The Nature of Work: An Introduction to Debates on the Labour Process*. London: Macmillan.
- Walker, Pat, ed. 1978 *Between Labor and Capital*. Montreal: Black Rose.
- Webster, Frank 1995 *Theories of the Information Society*. New York: Routledge.
- Wood, Ellen Meiksins 1986 *The Retreat from Class: A New "True" Socialism*. London: Verso.
- Wright, Erik Ohlin 1978 *Class, Crisis, and the State*. London: Verso.
- 1985 *Classes*. London: Verso.
- 1996 "Marxism after Communism." Pp. 121–145 in S. P. Turner, ed., *Social Theory and Sociology*. Cambridge, Mass.: Blackwell.
- , Andrew Levine, and Elliott Sober 1992 *Reconstructing Marxism*. London: Verso.

RAYMOND A. MORROW



## MASCULINITY

*See* Femininity/Masculinity.

## MASS MEDIA RESEARCH

The interest of sociologists in mass communication was stimulated by developments in technology allowing the reproduction and speedy transmission of messages. It began with the rise of the popular press, followed by the invention of film, sound broadcasting (or radio), and the audiovisual, including television and cable television. In the past decade, this interest has grown to embrace computer-influenced adaptations of these traditional mass media, the latest being the World Wide Web (www), which is part of the Internet, or the Information Superhighway.

All of us live in a world of media-constructed images that, presumably, significantly influence what we think and how we partition our attention, time, and other scarce resources. So pervasive has been the media presence that issues relating to these influences have also drawn the attention of researchers from disciplines other than sociology.

It is to Harold Lasswell (1947), an empirically oriented political scientist, that the social science community owes a succinct formula that lays out the major elements within the field of communication research: *Who* says *what* in which *channel* to *whom* and with what *effects*? Only some channels lend themselves to *mass* communication, which can be defined, in the terms of the above formula, as the transmission by professional communicators (*who*) of a continuous flow of a uniform content (*what*) by means of a complex apparatus (*channel*, or *how*) to a large, heterogeneous, and geographically dispersed audience (*to whom*), the members of which are usually anonymous to the communicator and to each other.

Not included in this definition of mass communication are its effects or, more broadly speaking, its consequences, toward which most of the sociological research effort has been directed. The physical or electronic transmission of message content does not in itself suffice for communication. Communication is indisputably social, in that it consists of a meeting of minds between communicator and audience, in the sense of mutual accommodation. Yet the nature and extent of effects

have been, over the years, the central problem of sociological interest in media research.

Effects, however, do not stand alone as a separate and independent dimension for research. More accurately, scholars have investigated the effects of each element in Lasswell's formula. For example, there has been analysis of what effects the different kinds of communicator controls may have, whether the *who* be defined by demographic characteristics and professional values of the communicator (Weaver and Wilhoit 1996; Weaver with Wu 1998) or as the growing big-business controls over mass media through concentration of ownership by not more than ten corporations (Bagdikian 1997). In other research, the effects of different kinds and amounts of media content have been analyzed to determine what was likely to have influenced particular audience behavior, as has been what difference it makes whether news is obtained from radio, television, or newspapers.

Media effects have been studied on three levels: the atomistic, the aggregate, and the societal. Effects on the atomistic level involve the cognitive processes and behavioral responses of *individuals* who make up the various mass audiences. By contrast, aggregate measures take into account only *distributions* that produce changes in averages usually expressed as net effects. Consequences for society have more to do with the political, cultural, and other *institutional* changes that represent cumulative adaptations over time to the dominance of a particular mass medium. Inferences based on the observation of effects on one level when ascribed to effects on a different level have often turned out to be invalid.

### THE ATOMISTIC LEVEL

Much of the media research effort has been a response to the operational needs of communicators and propagandists, or of those who wished to defend the public against what was perceived as the pernicious influence of the media. The basic problem has been that of precisely pinpointing effects: What were the characteristics of the potential audience? Who among them was susceptible? What were the determinants of their reactions?

To answer these and similar questions, audience research has typically focused on the situations in which mass communications are received

and on the habits and cognitive processes that underlie the responses of individuals either to specific media messages or to some significant part of the media fare. The responses under scrutiny have ranged from the arousal of interest, gains in knowledge, the recognition of dangers, changes of opinion, and other attitudinal measures to such behavioral indicators as consumer purchases, electoral decisions, and the "elevation" of cultural taste.

Precisely because of its focus on the individual, this line of research tends to stress the diversity of ways in which individuals relate to media content. First of all, audiences are found to be stratified by education, interest, taste, habits, gender, and age. Taken together, education and age tend to account for considerable variance in media use. The observation that some content had only minimal audience penetration helped explain why some information campaigns failed. Consistent patterns of exposure to different kinds of content further suggested that members of the mass audience, by and large, found what suited their needs and interests.

Second, even common exposure turns out to be a less strong predictor of response than expected. Not everyone understands or understands fully, and reactions are affected by the preconceptions with which people approach the content, by preconceptions rooted in past socialization experience but also reflecting the perspective of groups with which they are associated or identify themselves. Audiences are obstinate and people have options in how they orient themselves to any particular set of messages. They can ignore, misunderstand, accept, find fault with, or be entertained by the same content. In other words, there is no assurance that anyone other than those, for whatever reason, already so disposed will accept the facts, adopt the opinion, or carry out the actions suggested by the mass communicator (Schramm 1973).

This downplaying of the importance of content elements by a methodical partitioning of the mass audience received systematic formulation in the "minimal effects" theorem, derived from Joseph Klapper's review (1960) of certain empirical research findings of studies conducted mostly during the 1940s and 1950s. He generalized that certain factors, such as audience characteristics and a pluralistic media structure, which mediated

between content and response, worked primarily in the service of reinforcement of prior attitudes. Changes triggered by exposure were pretty much limited to people whose situations already impelled them to move in that direction. Klapper did, however, acknowledge the power of mass communication to move people on matters with which they were unfamiliar and concerning which they had no distinct views of ingrained habits.

Strong evidence in favor of not-so-minimal effects has come from observations made in the laboratory, especially through the series of experimental studies on children, reported in *Television and Social Behavior*, conducted under the auspices of the U.S. Surgeon General (Murray and Rubinstein 1972). After exposure to programs that included "violent" behavior, subjects often engaged in similar behavior during their play, and were more likely than subjects not so exposed to commit other violent acts.

Such experiments generally have been set up to maximize the possibility of demonstrating direct effects. Thus, in this instance, children, especially young children, would be inclined to model their own behavior on what they see. Moreover, such findings of short-term effects observed in a play situation have to be considered within the context of the whole socialization experience over many years. Longitudinal studies and experimental studies of older children in a more natural setting have yielded results that are more ambiguous (Milavsky et al. 1973). Laboratories do not fully replicate communication situations of real life (Milgram 1973).

Over years of study, conflicting evidence emerged about the relation between mass media (largely televised) violence and its influence on aggressive behavior. Operational definitions of key variables, research designs, and support for research having come from the television industry or not, all have been viewed as possible influences on the research results themselves.

The media-violence effects issue is one in which research results have played an important role in public policy debates involving the broadcast industry often pitted against public interests. Decades of debate eventually led to passage of the Telecommunications Act of 1996, authorizing an electronic device called the V-chip, that would

allow blocking programs with violent programming from television viewing (Rowland 1997).

In recent years, concern over the media-violence effects issue has been “exported” to the rest of the world through the multiple means of transmission and aggressive marketing by multinational corporate media interests. Public opinion surveys in all corners of the globe indicate disturbance by the ubiquitousness of television violence (Maherzi 1997).

The challenges posed by experimental studies have to be faced. Casual but repeated exposure to televised messages results in incidental learning. For example, content of advertising appeals gradually intrude into our consciousness until we associate a product with a particular brand name, or issues dominating the news become the criteria by which we measure the effectiveness of a political leadership. Insofar as the various mass media sources transmit similar content and play on similar themes, such limited effects, if they are cumulative, can produce shifts of significant proportions.

### AGGREGATE EFFECTS

Because the responses of persons are so diverse, the effect of communication en masse has to be conveyed in some kind of summary measure—as an average, a trend, a general movement. From this perspective, the magnitude of the shift in the responses of individuals, or whether this represents reinforcement or a reversal, matters less than the general picture, taken over time.

How differently effects can appear when viewed from different perspectives, and at different periods of time, may be illustrated by reference to studies of the diffusion of innovations (Rogers 1995). Detailed documentation in the more than fifty-year tradition of diffusion studies has spanned many disciplines and many nations, bringing revision and fine-tuning to the original concepts. The early diffusion model was applied to post-World War II development programs in agriculture, family planning, public health, and nutrition. Today, the model is being applied to areas as diverse as acquired immunodeficiency syndrome (AIDS) and the Internet.

The process by which innovations are adopted and spread suggests that early adopters, also called

“influentials” or “opinion leaders,” depending on context, are more cosmopolitan in their orientation and hence more attuned to certain media messages. They select from the total stream the messages that best meet their needs and interests. Others will adopt an innovation only after its success has been demonstrated or, if that is precluded—as it would be in most political decisions—out of trust in the expertise of the pacesetters. One can account for the different behavior of leaders and followers in such situations—that is, why one person moves ahead and another is content to wait—in terms of personal characteristics and social relationships. Aggregate effects, on the other hand, have to do with whether or not there has been a general movement toward acceptance or rejection of the innovation.

The most direct measures of aggregate effects are to be found in two-variable relationships designed to show cause and effect, with one variable functioning as an indicator of media presence and the other representing the response. Many such combinations are possible. One can use media penetration (e.g., newspaper circulation, or the proportion of homes with a television set) or content characteristics (e.g., the number of violent acts in children’s programs, editorial endorsements, or issues emphasized in the news). This rules out media behavior, which is voluntary for individuals in the audience, and may bring into question the influence of still other, often unmeasured, variables that also account for the presumed effect.

As the age of television dawned, opportunities for “controlled” observation—comparing two matched areas, one receiving television and the other not yet within reach of the broadcast signal—were never fully exploited. Rarely did findings about the advent of television go beyond documenting the rather obvious fact that television viewing cut into the use of some other media, especially radio and to a lesser degree movie attendance and children’s comic book reading. Nor were the consequences of this reallocation of time at all clear.

A study of children in “Teletown” and “Radiotown,” the latter community still without television but comparable in other respects, concluded that before television, many children (had gone) through the same type of change as today

from fantasy-seeking media behavior toward reality-seeking media behavior (Schramm et al. 1961).

For another natural experiment, a researcher was able to identify a Canadian town that was to get television within a year, after being unable to obtain reception because of geographic location. Comparisons were made with two other communities, one that would receive a second television channel and one that remained the same that year, with four channels. Results showed that the arrival of television did make a difference in children's reading skills, aggression levels, sex role stereotyping, and attendance at social events outside the home (Williams 1986).

Any such cause-effect evidence from natural experiments is lacking on matters relating to citizen participation among adults in national elections. Systematic comparisons between the turnout and overall responsiveness to "party" issues during the 1952 presidential election in counties with high TV penetration and low TV penetration revealed no consistent differences, probably because other media were already saturated enough with campaign material to have produced a high level of interest. Situations subject to such "ceiling" effects prevent further research in response to the presence of a new medium.

Variations in content, when they occur, have offered far more opportunities for controlled observations, many of which have challenged the conventional wisdom. That voters on the west coast of the United States would be dissuaded from voting in the presidential election once television, based on early returns, had declared a winner seemed only logical. Yet studies showed that westerners continued to cast ballots in roughly the same proportions as their compatriots in states where polls had already closed. In voting, they were evidently moved by considerations other than practical utility and by other competing media messages. Whatever the effects of such broadcast returns on the decision to vote or not to vote, they have been too small for detecting with present techniques of measurements (Lang and Lang 1984, ch. 5). As regards editorial endorsements, where the range of variations is greater, research has shown that such support gives candidates for minor offices, many of whom are only names on a ballot, an incremental but nevertheless

distinct advantage over other minor candidates on the same slate.

Correlations that pair media use variables with some measure of response always imply change over time. The alternative is to conduct before-and-after studies in response to events as they are being communicated via the mass media: a televised speech by a political leader, the announcement of an unexpected reversal of government policy, news of foreign crisis, or simply the flow of information about economic conditions and problems facing the country.

Polls before and after an appropriately timed speech have documented the power of a head of state to move opinion through appeals directed to the public. Speeches can create greater awareness. They are designed to focus attention on those issues and actions from which the politician stands to benefit. The effectiveness of such media events is apt to be greatest when an issue is just surfacing. Leaders also have the ability to make news. Even without an undisputed success, their public appearances, diligently reported by the news media, are used to dramatize their own role in promoting solutions to matters believed to be of general concern, thereby conferring status on themselves, as well as on the matters they seek to promote.

Careful analysis of the impact of many such events over the years again challenges the conventional wisdom. Neither speeches nor foreign travel by American presidents over the many administrations have, by themselves or in combination, *uniformly* shored up public support. Public response to these events has been highly dependent on the political context—that is, on whether they coincided with other events that tended to enhance the president's standing or whether his administration was plagued by intractable problems, such as public concern about a declining economy, an indignation over American hostages whose release it could not effect, or revelations of governmental wrongdoing such as those that surfaced during Watergate and led to the resignation of Richard Nixon as chronicled by Lang and Lang (1983).

What stands out in a long line of studies is the general correspondence over time between the overall amount of attention a topic, an issue, or a personality enjoys in the media and the audience's awareness, interest in, and concern about these. Mass communication influences not so much *what*

people think (opinion) but what they think *about* (recognition). Insofar as there is enough common emphasis, the media perform an “agenda-setting” function (McCombs and Shaw 1972, 1976; Iyengar and Kinder 1987). Collectively, the media and the audience define the terms of public debate. Media attention also confers status on some of the many voices clamoring to be recognized.

Agenda-setting research requires at least two steps; that is, content analysis is conducted to define the media agenda, and surveys are used to identify the public or audience agenda. This simple formulation, which has been a central focus of effects research in political communication for some years, attributes to the media at one and the same time too much and too little influence. On the one hand, the media do not, all on their own, dictate or control the political agenda. Neither public awareness nor recognition of a problem is sufficient to stir a controversy on which people take sides. On the other hand, access to the media is a major resource for the advocates of particular policies. Concerns become issues through discussion in which political leaders, government officials, news and commentary in the press, and the voices of citizens reciprocally influence one another in a process more aptly characterized as “agenda building” (Lang and Lang 1984). In fact, three agendas—those of the media, the public, and the policy makers—all influence each other in a rather interactive political process (Rogers and Dearing 1988).

### THE SYSTEMIC PERSPECTIVE

The dissemination of content via the mass communication system occurs in a highly selective fashion. Some information is privileged; other information is available only to those with the interest and resources to pay for it, thereby stratifying societies into the “haves” and have-nots,” or into the information rich and the information poor. The upshot of all the efforts to direct communication flow is a repertoire of images of events and ongoing social activities. Media organizations are themselves producers of content. The mass communications through which the world is brought into focus are, in the broadest sense, cultural creations that incorporate the perspectives of the

producers and of others whose views have to be taken into account. This influence of the communications system on content, intended or not, is a source of bias often unrecognized by those responsible for it.

One has to differentiate between two sources of bias: technological and social. Technological bias stems from the physical characteristics of the medium. Harold Innis (1951), the Canadian institutional economist, distinguished between bias toward space and bias toward time. Paper, he averred, because of its light weight, was easily transported but also perishable, and so supported the development of centralized administration. The uniformity thus imposed over a given area (space) was usually at the expense of continuity (time). A more flexible medium could adapt to an oral tradition that favored spontaneous cooperation among autonomous units. Indeed, signs continue to point in the direction that cheaper and smaller electronic devices may be radically increasing the control individuals have over the information available to them and what to do with it (Beniger 1986). Information is central to the control individuals have over the social events of the world in which they live.

Applied mechanistically, without regard to who is in control, these categories lead to a simple-minded media determinism. Social bias has to do with how the capabilities intrinsic to the dominant medium are exploited. Television, in and of itself, may not have had a demonstrable effect on voter turnout but nevertheless contributed indirectly. Thus, a nonpartisan political coverage, designed not to offend but with an insatiable appetite for scandal, was implicit in the economic logic of aiming at the largest possible audience while the American regulatory system opened the way for a well-financed candidate could buy nearly unlimited time to air well-targeted, and often negative, political messages. Both trends fed into an already existing distrust of government. Meanwhile, campaign strategies adapted to the medium of television helped undermine the power of party machines to deliver votes. The nominating conventions in which political bosses once traded votes have been transformed into showpieces, played for a national audience as the curtain raiser for the U.S. presidential election campaign.

More generally, its penetration into the spheres of other institutions—political, cultural, educational, and so on—is what makes mass communication a potentially powerful influence on the societal level. It hardly matters whether the media are viewed as a resource or as a threat. The publicity generated through mass communication brings the norms of the larger society to bear on actions that once might have been considered privileged or at least shielded from public scrutiny. Conversely, the competition for visibility is an inducement for elite institutions to adopt at least some of the conventions of the media culture.

There remains the question of who sets these norms and standards. Some scholars have argued that repeated exposure to a sanitized media culture results in “mainstreaming” (Gerbner et al. 1986), or what has been commonly referred to as homogenizing consequences. Accordingly, all but a few of the diverse currents that feed into the kaleidoscope of minority United States cultures, including women and racial and ethnic minorities, receive comparatively little or no recognition. Lacking an effective institutional representation, they are more readily marginalized. Content analyses of character portrayal and the values espoused by heroes and villains, victims and perpetrators of violence in the popular entertainment fare, as well as by surveys of adults, adolescents, and children in the United States, have lent some support to the charge of “mainstreaming.” Such cultivation analysis in cross-cultural settings also is finding support for these kinds of media effects (Signorielli and Morgan 1990). Despite the premium on novelty, most media organizations are inclined not to stray too far from what is popularly accepted but will be likely to eagerly imitate any demonstrated success.

The representation of political views is similarly constrained. Despite the independence of the press and a few celebrated instances where a small number of persistent journalists initiated an inquiry, as when the Woodward and Bernstein team pursued the Watergate story, the more typical pattern is to wait until political actors have highlighted a problem. Usually, it is they, rather than the press, who define the terms of controversies over policy, as was the case in the much-studied limited role of the media, largely American and mostly Cable News Network (CNN), in the 1991 Persian Gulf conflict (Dennis et al. 1991; *Media*

*Development*, 1991). To paraphrase W. Lance Bennett (1988), when the institutional voices speaking in protest are stilled, all but the more radical media are inclined to drop the issue as well. Discussion and diversity exist but usually within self-imposed limits.

It should be clear that major media organizations, though important players, are less than fully separate from other establishments. Their influence on events is greatest when they act in conjunction with other agencies.

### INTERNATIONAL, OR GLOBAL, RESEARCH

Into the 1950s, mass communication research was very much a product of the United States: U.S. media programming and news events, U.S. media practices and policies, with necessary consideration of U.S. politics and society. There was a tendency for mass media research to concentrate on one aspect of the communication process originally outlined by Lasswell, namely effects. Later, research refocused some attention on other elements of the communication process—on what was produced and how it was produced.

Research also began exploring dimensions of mass media influence outside the United States. Some studies were international extensions of research already being conducted in the United States, as with the media-violence issue, diffusion of innovations, and cultivation analysis. For other research, the international perspective was the question under investigation: What should be considered when news, television entertainment programming, and other media products cross geopolitical boundaries?

In the 1950s, the identification of four theoretical perspectives of the press—authoritarian, communist, libertarian, and social responsibility—focused attention on the social, political, cultural, and philosophical underpinnings that determine how and why the mass media of various parts of the world are different (Siebert et al. 1956). Here, the press was defined as all available mass media responsible for dissemination of news. Some country-specific case studies have helped to elucidate these differences (Mickiewicz 1988; Alot 1982; Howkins 1982).

The post-World War II era was concerned with reconstruction and development of the so-called Third World and spawned a developmental perspective of the mass media (Schramm 1964; Lerner and Schramm 1967). A fifth concept of the press focused on revolutionary goals of the mass media, with China's transitions over the past century and the recent emergence of terrorism as examples of present-day extensions of this perspective (Chu 1977; Lull 1991; Nacos 1994).

Debate over international communication issues has been strongly influenced by the political and economic imbalances among the nations of the world. So-called North-South debates juxtapose the concerns of the politically independent, industrialized, and developed part of the world against those of the newly independent countries, which have been trying to shed their colonial pasts and, because they have missed out on the Industrial Revolution, are still developing. Other debates between the East and the West have been reflections the cold-war conflicts between the former Soviet Union and the United States, in which an emergent group of nonaligned countries, largely in the developing world, have refused to take sides.

In the 1970s, the United Nations Educational, Scientific and Cultural Organization (UNESCO) became an outspoken supporter of efforts by this developing Third World to create a New World Information and Communication Order (NWICO) as the counterpart to calls for a less well known New World Economic Order. UNESCO formed an advisory commission of international scholars, diplomats, and specialists, headed by Irishman Sean MacBride, to study the world's communications problems, and to address Third World protests against the domination of news flow from industrialized countries of the West to the rest of the world. It was also concerned with issues as diverse as governmental controls, censorship, cultural dominance or imperialism, concentration of media ownership and powerful control of transnational corporations, freedom and responsibility of the press, commercialization of mass media, protection of journalists when they are at work in countries other than their home, access to technology and infrastructure development, and rights to communication (International Commission for the Study of Communication Problems 1980).

Studies proliferated in relation to the various issues posited by the UNESCO report. "Cultural imperialism" that reflected political imperialism of older models from historical colonialism was the theoretical perspective from which empirical studies documented imbalances in the world's media infrastructures and contents. The "media imperialist" was decidedly Western, and largely American. UNESCO together with professional organizations of scholars and practitioners supported large multicountry studies of international news flow (Sreberny-Mohammadi 1984; Kirat and Weaver 1985). These studies analyzed images, distortion in content, and the role of news agencies, and also explored factors that determine what becomes news (Galtung and Ruge 1965).

This research showed that certain factors do influence news production, such as economic considerations in distribution of news-gathering resources and ease of reporting; importance to the country producing the news in terms of trade, social and cultural ties, geographic proximity, and relative standing or status of nations in the eyes of others. Such findings have influenced shifts in mass media structures and output, such as the increase in national and regional news through agencies like the New China News Agency, also known as Xinhua, and the Caribbean News Agency (CANA).

In these nearly twenty years, the UNESCO debate over NWICO has faded and resurfaced more than once, and UNESCO-sponsored round table discussions continue on a regional basis to provide a forum for discussion of research on global communication issues. A recent UNESCO report on the status of the world's mass media and how they are handling the challenges of new technologies claims that the problems identified by the MacBride Commission "still remain a burning issue" and are likely to continue to be the focus of research efforts at national, regional, and international levels (Maherzi 1997).

One issue highlighted in the original UNESCO report that has received considerable national and multinational research attention is the inequality of women and girls in societies around the globe, their disproportionate representation in media-related employment, and their stereotyped images in news and entertainment media portrayals. More than a decade behind the U.S. civil rights and

women's movements that triggered considerable mass media research with implications for local media practices, programming decisions and U.S. policy, this UNESCO report was to become pivotal in pitching these issues to the rest of the world for research and analysis.

The United Nations Fourth World Conference on Women, held in Beijing in 1995, included women and media among the thirteen major substantive areas for consideration. There were reports on country-specific surveys of media practices, characteristics and attitudes of those holding jobs in the media, and program content (All-China Journalists Association and Institute of Journalism, Chinese Academy of Social Sciences 1996) and studies that involved the coordination of data collection in many countries. For example, the Global Media Monitoring Project gathered data from news media—radio, television and newspapers—in seventy-one countries for representation and portrayal of women at one point in time: January 18, 1995. Based on the data, an overall report and separate regional reports were produced. The study provides a large database with policy implications for UNESCO and for media systems in many countries. A benchmark for measuring future change, the Media Monitoring Project's study remains the most extensive survey of portrayals of women in the world's news media that has been undertaken to date (Media Watch 1995).

On balance, however, it remains unclear from available research whether women and men working in the mass media take different approaches on issues or events, even if women and men have a different range of interests they select for media attention. Though there has been an increase in the number of women working in the media in the past few years, surveys also show that the world of media remains strongly male (Maherzi 1997).

Most recently, research questions emerging in international communications are centering on the concept of globalization of mass media producers and their products. The concentration and control of media industries into fewer and fewer transnational corporations has long been a concern in research that points to the negative effects of increasing global "homogenization." It is not without some irony that, on the one hand, the results of mass media investigations have brought

about international pressure for diversifying media products with voices of women and minorities of all kinds, while on the other hand, ownership of the world's media is concentrated in the hands of a few corporate giants.

### NEW MEDIA TECHNOLOGIES

One of these corporate giants in the media world is Time Warner. Years ago, Warner Brothers produced films and Time produced a news magazine. Together, they are now also involved with other aspects of the media environment, including newspapers, magazines, television, cable television, radio, video games, telephones, and computers in the United States and around the world. Understanding why such a corporate giant was created is key to understanding the impacts of new media technologies: "Convergence" of technical infrastructures is leading to consolidation of media products and services.

The Internet, or the Information Superhighway, with the World Wide Web (www), is the most recent development, which has the potential to offer any or all of the aforementioned media products and services through on-line information networks. Current Internet services are transmitted via computer-telephone connections, but explorations of other methods of transmissions are under way, such as cable television and direct satellite links. Obviously, the corporate giants wish to be positioned to take their share of any new media offerings and to provide new delivery systems. What technology will deliver news and entertainment, as well as other information and services, in the future remains an open question.

Regardless of how the infrastructure question is decided, will the definition of "news" remain the same? Because of the vast storage capacity of computers, the role of editors and gatekeepers may change. With the ability of computers to allow more personalized reception of news offerings, will the "newspaper" still be considered a *mass* medium?

Internet-related services and e-mail already are providing new possibilities for on-line public opinion polls. But so far, such surveys are similar to volunteer call-in or write-in polls and are limited to Internet users. While the number of new Internet users grows daily, by the year 2000 the numbers of



U.S. adults on-line is expected to reach only 60 million (Pew Research Center 1996). Increasingly, on-line users are “decidedly mainstream,” according to a nationwide telephone survey, and for now, television is losing in their allocations of time (Pew Research Center 1998).

At the atomistic level, researchers have conducted ethnographic and clinical observation on how people relate to computers and how they may be reconstructing their basic sense of identity (Turkle 1995).

How will the traditional mass communication conceptualizations of diffusion of innovations, agenda-setting and agenda-building, cultivation, and mainstreaming help us to understand the new media environment?

Is our new world more global or more local (Sreberny-Mohammadi et al. 1997)? Are we moving into Marshall McLuhan’s long-promised “global village”? Or is it a global megalopolis? Are we moving from UNESCO’s “many voices, one world” to many worlds, one voice?

These are some of the many questions that will occupy media researchers in the near future.

REFERENCES

All-China Journalists Association, and Institute of Journalism, Chinese Academy of Social Sciences 1996 “Survey on the Current Status and Development of Chinese Women Journalists.” Beijing.

Alot, Magaga 1982 *People and Communication in Kenya*. Nairobi: Kenya Literature Bureau.

Bagdikian, Ben H. (1983) 1997 *The Media Monopoly*, 5th ed. Boston: Beacon Press.

Beniger, James R. 1986 *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, Mass.: Harvard University Press.

Bennett, W. Lance 1988 *News: The Politics of Illusion*. New York: Longman.

Chu, Godwin C. 1977 *Radical Change through Communication in Mao’s China*. Honolulu: East-West Center/University Press of Hawaii.

Dennis, Everette E., ed. 1991 *The Media at War: The Press and the Persian Gulf Crisis*. New York: Gannett Foundation Media Center.

Galtung, Johan, and Mari Holmboe Ruge 1965 “The Structure of Foreign News,” *Journal of Peace Research* 2:64–91.

Gerbner, George, Larry Gross, Michael Morgan, and Nancy Signorielli 1986 “Living with Television: The Dynamics of the Cultivation Process.” In J. Bryant and D. Zillmann, eds., *Perspectives on Media Effects*. Hillsdale, N.J.: Lawrence Erlbaum.

Howkins, John 1982 *Mass Communication in China*. New York: Longman.

Innis, Harold A. 1951 *The Bias of Communication*. Toronto: University of Toronto Press.

International Commission for the Study of Communication Problems 1980 *Many Voices, One World: Communication and Society, Today and Tomorrow*. New York: Unipub.

Iyengar, Shanto, and Donald R. Kinder 1987 *News That Matters: Television and Public Opinion*. Chicago: University of Chicago Press.

Kirat, Mohamed, and David Weaver 1985 “Foreign News Coverage in Three Wire Services: A Study of AP, UPI and the Nonaligned News Agencies Pool,” *Gazette* 35:31–47.

Klapper, Joseph T. 1960 *The Effects of Mass Communication*. New York: Free Press.

Lang, Gladys Engel, and Kurt Lang 1984 *Politics and Television Re-Viewed*. Beverly Hills, Calif.: Sage.

—, and — 1983 *Battle for Public Opinion: The President, the Media, and the Polls During Watergate*. New York: Columbia University Press.

Lasswell, Harold D. 1947 “The Structure and Function of Communication in Society.” In L. Bryson, ed., *The Communication of Ideas*. New York: Harper.

Lerner, Daniel, and Wilbur Schramm 1967 *Communication and Change in the Developing Countries*. Honolulu: East-West Center Press.

Lull, James 1991 *China Turned On: Television, Reform and Resistance*. New York: Routledge.

McCombs, Malcom E., and Donald L. Shaw 1972 “The Agenda-Setting Function of Mass Media.” *Public Opinion Quarterly* 36:176–187.

— 1976 “Structuring the ‘Unseen Environment.’” *Journal of Communication* (Spring):18–22.

Maherzi, Lotfi 1997 *World Communication Report: The Media and the Challenge of New Technologies*. Paris: UNESCO.

*Media Development* 1991 “Reporting the Gulf War,” Special Issue, October. London: World Association for Christian Communication.

Media Watch 1995 *Global Media Monitoring Project: Women’s Participation in the News*. Toronto and Ontario: National Watch on Images of Women in the Media (Media Watch) Inc.

- Mickiewicz, Ellen 1988 *Split Signals: Television and Politics in the Soviet Union*. New York: Oxford University Press.
- Milavsky, J. Ronald, Horst Stipp, Ronald C. Kessler, and William S. Rubens 1982 *Television and Aggression: A Panel Study*. New York: Academic.
- Milgram, Stanley 1973 *Television and Anti-Social Behavior: A Field Experiment*. New York: Academic.
- Murray, John P., and Eli A. Rubinstein 1972 *Television and Social Behavior: Reports and Papers*, 5 vols. Washington, D.C.: National Institute for Mental Health.
- Nacos, Brigitte L. 1994 *Terrorism and the Media: From the Iran Hostage Crisis to the World Trade Center Bombing*. New York: Columbia University Press.
- Pew Research Center for the People and the Press 1996 "One in Ten Voters Online for Campaign." September. <http://www.people-press.org>.
- 1998 "Online Newcomers More Middle-Brow, Less Work-Oriented: The Internet News Audience Goes Ordinary." October-December. <http://www.people-press.org>.
- Rogers, Everett M. (1962) 1995 *Diffusion of Innovations*, 4th ed. New York: Free Press.
- , and James W. Dearing 1988 "Agenda-Setting Research: Where Has It Been, Where Is It Going?" In J. Anderson, ed., *Communication Yearbook 11*. Newbury Park, Calif.: Sage.
- Rowland, Willard D., Jr. 1997 "Television Violence Redux: The Continuing Mythology of Effects." In M. Baker and J. Petley, eds., *Ill Effects: The Media/Violence Debate*. New York: Routledge.
- Schramm, Wilbur 1964 *Mass Media and National Development*. Stanford, Calif.: Stanford University Press.
- 1973 *Men, Women, Messages and Media*. New York: Harper and Row.
- , Jack Lyle, and Edwin B. Parker 1961 *Television in the Lives of Our Children*. Stanford, Calif.: Stanford University Press.
- Siebert, Fred, Theodore Peterson, and Wilbur Schramm 1956 *Four Theories of the Press*. Urbana: University of Illinois Press.
- Signorielli, Nancy, and Michael Morgan, eds. 1990 *Cultivation Analysis: New Directions in Media Effects Research*. Newbury Park, Calif.: Sage.
- Sreberny-Mohammadi, Annabelle 1984 "The 'World of the News.'" *Journal of Communication* (Winter):121-134.
- , Dwayne Winseck, Jim McKenna, and Oliver Boyd-Barrett, eds. 1997 *Media in Global Context: A Reader*. New York: Arnold/Hodder Headline Group.
- Turkle, Sherry 1995 *Life on the Screen: Identity in the Age of the Internet*. New York: Simon and Schuster.
- Weaver, David H., and G. Cleveland Wilhoit 1996 *The American Journalist in the 1990s: U.S. News People at the End of an Era*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- , with Wei Wu 1998 *The Global Journalist: News People Around the World*. Cresskill, N.J.: Hampton Press.
- Williams, Tannis, ed. 1986 *The Impact of Television*. New York: Academic Press.

JANICE M. ENGSBERG  
KURT LANG  
GLADYS ENGEL LANG

## MASS SOCIETY

The mass society theory, in all its diverse formulations, is based on a sweeping general claim about "the modern world," one announcing a "breakdown of community." The leading nineteenth-century proponents of this position were Louis de Bonald, Joseph de Maistre, and, from a different perspective, Gustave Le Bon. These formulations argue the collapse of the stable, cohesive, and supportive communities found in the days of yore. In modern times, as a consequence, one finds rootlessness, fragmentation, breakdown, individuation, isolation, powerlessness, and widespread anxiety (Giner 1976; Halebsky 1976).

The original formulations of this position, those of the nineteenth century, were put forth by conservatives, by persons identified with or defending the old regime. These were critiques of the liberal theory or, more precisely, of liberal practice. The basic aim of the liberals was to free individuals from the restraints of traditional institutions. That aim was to be accomplished by the dismantling of the "irrational" arrangements of the old regime. Liberals, understandably, were enthusiastic about the achievement: Free men could do things, achieve things, create things that were impossible under the old arrangement. The collective benefits, they argued, were (or would be) enormous. The conservative critics agreed about some aspects of the history. They agreed about the general process of individuation. They, however, called it fragmentation or a decline of community. More important, they provided very different assessments of the consequences. At its simplest, the

liberals argued an immense range of benefits coming with the transformation, a conclusion signaled, for example, in Adam Smith's title, *The Wealth of Nations*. The mass society theorists agreed with the basic diagnosis but drew strikingly opposite conclusions pointing to a wide, and alarming, range of personal and social costs.

The modern world begins, supposedly, with an enormous uprooting of populations. Ever greater numbers are forced from the small and stable communities into the large cities. In place of the strong, intimate, personal supports found in the small community, the large cities were characterized by fleeting, impersonal contacts. The family was now smaller. The isolated nuclear family—father, mother, and dependent children—was now the rule, replacing the extended family of farm and village. The urban neighborhoods were and are less personal. The frequent moves required in urban locales make deep, long-lasting friendships difficult if not impossible. As opposed to the support and solidarity of the village, instrumental and competitive relationships are typical in the large cities and this too makes sustained social ties problematic. In the mass society people are “atomized.” The human condition is one of isolation and loneliness. The claims put forth in this tradition are typically unidirectional—the prediction is “more and more.” There is ever more uprooting, more mobility, more societal breakdown, more isolation, and more anxiety.

The nineteenth-century versions of this theory focused on the insidious role of demagogues. In those accounts, traditional rulers, monarchs, aristocracy, and the upper classes did their best to govern fundamentally unstable societies. But from time to time, demagogues arose out of “the masses,” men who played on the fears and anxieties of an uneducated, poorly informed, and gullible populace. The plans or programs offered by the demagogues were said to involve “easy solutions.” But those, basically, were unrealistic or manipulative usages, ones providing no solutions at all. The demagogues brought revolution, which was followed by disorder, destruction, and death. The traditional patterns of rule were disrupted; the experienced and well-meaning leaders were displaced, either killed or driven into exile. The efforts of the demagogues made an already-desperate situation worse.

Conservative commentators pointed to the French Revolution as the archetypical case with Robespierre and his associates as the irresponsible demagogues. Mass society theorists also pointed to the experience of ancient Greece and Rome. There too the demagogues had done their worst, overthrowing the Athenian democracy and bringing an end to the Roman republic. The republic was succeeded by a series of emperors and praetorians, men who, with rare exceptions, showed various combinations of incompetence, irresponsibility, and viciousness.

The lesson of the mass society theory, in brief, was that if the masses overthrew the traditional leaders, things would be much worse. The “successes” of liberalism, the destruction of traditional social structures, the elimination of stable communities, and the resulting individualism (also called “egoism”) could only worsen an already precarious situation. The theory, accordingly, counseled acceptance or acquiescence.

It is easy to see such claims as ideological, as pretense, as justifications for old-regime privilege. Such claims were (and are) given short shrift in the opposite liberal dramaturgy and, still later, in the dramaturgy of the left. In those opposite accounts, the old regime is portrayed as powerful. The rulers, after all, had vast wealth and influence; they controlled the police and the ultimate force, the army.

In private accounts, however, the leaders of the old regime reported a sense of powerlessness. Their “hold” on power, they felt, was tenuous; they stood on the edge of the abyss. Chateaubriand, the French ambassador, congratulated Lord Liverpool on the stability of British institutions. Liverpool pointed to the metropolis outside his windows and replied: “What can be stable with these enormous cities? One insurrection in London and all is lost.” The French Revolution itself proved the flimsiness of “established” rule. In 1830, the restored monarchy in France collapsed after only a week of fighting in the capital. In 1848, Louis Philippe's regime fell after only two days of struggle. A month later, the Prussian king and queen, effectively prisoners of the revolution, were forced to do obeisance to the fallen insurgents. The queen's comment—“Only the guillotine is missing.” More than a century later, the historian J. R. Jones

declared that “during long periods of this time, many conservatives felt that they were irretrievably on the defensive, faced not with just electoral defeat but also doomed to become a permanent and shrinking minority, exercising a dwindling influence on the mind and life of the nation.”

Early in the twentieth century, sociologists in Europe and North America developed an extensive literature that also argued a loss-of-community thesis. Among the Europeans, we have Ferdinand Tönnies, Georg Simmel, and, with a difference, Emile Durkheim. Simmel’s essay, “The Metropolis and Mental Life,” had considerable influence in North America, especially in the development of sociology at the University of Chicago. The Chicago “school” was founded by Robert Ezra Park who had studied under Simmel. Park’s essay, “The City: Suggestions for Investigation of Human Behavior in the Urban Environment,” provided the agenda for generations of sociologists. Another central work in the Chicago tradition was Louis Wirth’s 1938 article, “Urbanism as a Way of Life.” The city, Wirth wrote, is “characterized by secondary rather than primary contacts. The contacts of the city may indeed be face to face, but they are nevertheless impersonal, superficial, transitory, and segmental . . . . Whereas the individual gains, on the one hand, a certain degree of emancipation or freedom from the personal and emotional controls of intimate groups, he loses, on the other hand, the spontaneous self-expression, the morale, and the sense of participation that comes with living in an integrated society. This constitutes essentially the state of *anomie*, or the social void, to which Durkheim alludes” (p. 153).

Writing almost a half-century after Wirth, sociologist Barrett A. Lee and his coworkers—in an important challenge to those claims—commented on this tradition as follows: “Few themes in the literature of the social sciences have commanded more sustained attention than that of the decline of community . . . . In its basic version, the thesis exhibits a decidedly antiurban bias, stressing the invidious contrast between the integrated small-town resident and the disaffiliated city dweller” (pp. 1161–1162). Those sociologists do not appear to have had any clear political direction. Their work was value-neutral. It was pointing to what they took as a basic fact about modern societies without proposing any specific remedies.

Later in the twentieth century, a new version of the mass society theory made its appearance. This may be termed the *left variant*. All three versions of the theory, right, neutral, and left, agree on “the basics,” on the underlying root causes of the modern condition, all agreeing on the “decline of community.” But the right and left differ sharply in their portraits of the rulers, of the elites, the upper classes, or the bourgeoisie. In the rightist version, the rulers face a serious threat from below, from the demagogues and their mass followings. Their control is said to be very tenuous. In the left version, the rulers are portrayed as skillful controllers of the society. The key to their successful domination is to be found in their adept use of the mass media.

The bourgeoisie, the ruling class, or its executive agency, the “power elite,” is said to control the mass media of communication, the press, magazines, motion pictures, radio, and television, using them for their purposes. News and commentary, much of it, is said to be self-serving. It is essentially ideological, material designed to justify and defend “the status quo.” The entertainment provided is diversionary in character, intended to distract people from their real problems. Advertising in the media serves the same purposes—distraction, creation of artificial needs, and provision of false solutions. The bourgeoisie, it is said, owns and controls “the media.” With their vast resources, they are able to hire specialists of all kinds, market researchers, psychologists, and so forth, to aid in this manipulative effort. The near-helpless audience (as ever, atomized, powerless, and anxious) is psychologically disposed to accept the “nostrums” provided.

Elements of this position appeared in the writings of the Italian Marxist, Antonio Gramsci, with his concept of ideological hegemony. Some writers in the “Frankfurt school,” most notably Herbert Marcuse, also argued this position. It appeared also in the work of C. Wright Mills, in his influential book, *The Power Elite*. Many others have offered variants of this position.

The left mass society theory provided a third “revision” of the Marxist framework, that is, after those of Bernstein and Lenin. It is the third major attempt to explain the absence of the proletarian revolution. Marx and Engels assigned no great importance to the mass media. They occasionally

referred to items in the “bourgeois” press, adding sardonic comments about its “paid lackeys.” But newspaper reports were treated as of little importance. They could not stop or reverse the “wheel of history.” But in this third revision, “the bourgeoisie” had found the means to halt the “inevitable” course. The controllers of the media were able to penetrate the minds of “the masses” and could determine the content of their outlooks. The masses were said to be drugged or, to use a favored term, they were “narcotized.”

In the 1950s, in the Eisenhower era, the mass media were unambiguously affirmative about “society” and its major institutions. Families were portrayed as wholesome and happy; the nation’s leaders, at all levels, were honorable and upstanding. It was this “affirmative” content that gave rise to the argument of the media as manipulative, as distracting. In the late 1960s, media content changed dramatically. Programs now adopted elements of the mass society portrait, dwelling on themes of social dissolution. Families, neighborhoods, and cities were now “falling apart.” Many exposés, in books, magazines, motion pictures, and television, in the news and in “investigative reports,” told tales of cunning manipulation. Unlike the right and left versions of the mass society theory, these critics do not appear to have any clear political program. They appear, rather, to be driven by an interest in “exposure.” No evident plan, directive, or call for action seems to be involved. Studies indicate that most of the participants are modern-day liberals, not socialists or Marxists.

The mass society theory has had a peculiar episodic history, a coming-and-going in popularity. It had a wave of popularity in the 1940s when Karl Mannheim, Emil Lederer, Hannah Arendt, and Sigmund Neumann, all German exile-scholars, attempted to explain the major events of the age. A sociologist, William Kornhauser presented an empirically based synthesis in 1959, but this effort, on balance, had little impact. In the 1960s, the wave of “left” mass society theorizing appeared, beginning with the influential work of Herbert Marcuse. In 1970, Charles Reich’s *The Greening of America* appeared, a book destined to have, for several years, an enormous influence. It provided a depiction of the nation that was entirely within the mass society framework: “America is

one vast, terrifying anti-community. The great organizations to which most people give their working day, and the apartments and suburbs to which they return at night are equally places of loneliness and isolation. Modern living has obliterated place, locality, and neighborhood” (p. 7).

Few research-oriented social scientists have given the mass society theory much credence in the last couple of decades, this for a very good reason: virtually all the major claims of the theory have been controverted by an overwhelming body of evidence (Campbell et al. 1976; Campbell 1981; Fischer 1981; Fischer 1984; Hamilton and Wright 1986).

The mass society portrait is mistaken on all key points. Most migration is collective; it is serial, chain migration, in which people move with or follow other people, family and friends, from their home communities. Most migration involves short-distance moves; most migrants are never very far from their “roots.” Cities do grow through the addition of migrants; but they also grow through annexation, a process that does not disturb established social ties. The typical mass society account, moreover, is truncated, providing an incomplete narrative. The “lonely and isolated” migrants to the city supposedly remain that way for the rest of their lives. Those lonely people presumably have no capacity for friendship; they are unable to get together with others to overcome their powerlessness, and so forth.

Many academics in other fields, however, continue to give the theory considerable credence. It is a favorite of specialists in the literary sciences, of those in the humanities. The theory, as noted, is also a favorite of journalists, of social affairs commentators, of writers, dramatists, and poets.

This paradoxical result requires some explanation. The literature dealing with “the human condition” has a distinctive bifurcated character. The work produced by research-oriented scholars ordinarily has a very limited audience, most of it appearing in limited-circulation journals for small groups of specialists. Those specialists rarely attempt to bring their findings to the attention of larger audiences. Attempts to correct misinformation conveyed by the mass media are also infrequent. The producers of mass media content show an opposite neglect: they rarely contact academic

specialists to inquire about the lessons found in the latest research.

Those who argue and defend mass society claims, on the whole, have an enormous audience. Writing in 1956, Daniel Bell, the noted sociologist, stated that apart from Marxism, the mass society theory was probably the most influential social theory in the Western world. Four decades years later, the conclusion is still valid. The intellectual productions based on this theory reach millions of susceptible members of the upper and upper-middle classes, most especially those referred to as the “intelligentsia.”

The mass society theory proves well-nigh indestructible. It continues to have wide and enthusiastic support in some circles regardless of any and all countering evidence. Some people know the relevant evidence but engage in various “theory-saving” efforts, essentially ad hoc dismissals of fact. Some people, of course, simply do not know the available evidence, because of the compartmentalization of academia. Some academics do not make the effort required to find out what is happening elsewhere. Some others appear to be indifferent to evidence.

## REFERENCES

- Bell, Daniel 1961 “America as a Mass Society: A Critique.” *End of Ideology: On the Exhaustion of Political Ideas in the Fifties*. New York: Collier.
- Campbell, Angus, Philip E. Converse, and Willard L. Rodgers 1976 *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*. New York: Russell Sage Foundation.
- 1981 *The Sense of Well-Being in America: Recent Patterns and Trends*. New York: McGraw-Hill.
- Fischer, Claude S. 1981 *To Dwell Among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.
- 1984 *The Urban Experience*, 2nd ed. San Diego, Calif.: Harcourt Brace Jovanovich.
- Giner, Salvador 1976 *Mass Society*. London: Martin Robertson.
- Halebsky, Sandor 1976 *Mass Society and Political Conflict: Toward a Reconstruction of Theory*. Cambridge, England: Cambridge University Press.
- Hamilton, Richard F. forthcoming “Mass Society, Pluralism, Bureaucracy: Explication, Critique, and Assessment.” Westport, Conn.: Praeger.

—, and James D. Wright 1986 *The State of the Masses*. New York: Aldine.

Jones, J. R. 1966 “England.” In Hans Rogger and Eugen Weber, eds., *The European Right: A Historical Profile*. Berkeley, Calif.: University of California Press.

Kornhauser, William 1959 *The Politics of Mass Society*. Glencoe: Free Press.

Lee, Barrett A., R. S. Oropesa, Barbara J. Metch, and Avery M. Guest 1984 “Testing the Decline-of-Community Thesis: Neighborhood Organizations in Seattle, 1929 and 1979.” *American Journal of Sociology* 89:1161–1188.

Nisbet, Robert A. (1953) 1962 *The Quest for Community: A Study in the Ethics of Order and Freedom*, New York: Oxford University Press. Reissued as *Community and Power*. New York: Oxford University Press.

Reich, Charles 1970 *The Greening of America*. New York: Random House.

Sennett, Richard, ed. 1969 *Classic Essays on the Culture of Cities*. New York: Appleton-Century-Crofts.

RICHARD F. HAMILTON

## MATE SELECTION THEORIES

Social scientists who study the family have long been interested in the question “Who marries whom?” On one level, the study of mate selection is conducted from the perspective of family as a social institution. Emphasis is placed on the customs that regulate choice of mates. A counterperspective views the family as an association. This perspective centers instead on the couple and attempts to understand the process of marital dyad formation. Both of these perspectives generate an abundance of knowledge concerning mate selection. Beginning primarily in the 1920s, theoretical and empirical work in the area of mate selection has made great advances in answering the fundamental question “Who marries whom?”

### INSTITUTIONAL PERSPECTIVES ON MATE SELECTION

The purview of anthropologists has centered on kinship structures as they relate to mate selection in arranged marriage systems. Sociological inquiry that sees the family as a social institution in the context of the larger society focuses instead on the

evolution of courtship systems as societies modernize. In this respect, it is important to note the contributions of scholars such as Bernard Murstein (1974, 1976) who have pointed out the importance of cultural and historical effects on courtship systems that lead to marriage.

Historical evidence suggests that, as a society modernizes, changes in the courtship system reflect a movement toward autonomous courtship systems. Thus, parentally arranged marriages diminish in industrialized cultures, since arranged marriages are found in societies in which strong extended kinship ties exist or in which the marriage has great significance for the family and community in terms of resources or status allocation. As societies modernize, arranged marriages are supplanted by an autonomous courtship system in which free choice of mate is the preferred form. These autonomous courtship systems are also referred to as “love” marriages, since the prerequisite for selection of a mate has shifted from the need to consolidate economic resources to that of individual choice based on love. Of course, family sociologists are quick to point out that the term “love marriage” is somewhat of a misnomer, since many other factors operate in the mate selection process.

Family social scientists have tried to understand the human mate selection process by using a variety of data sources and theoretical perspectives. The most global or macro approaches have made use of vital statistics such as census data or marriage license applications to study the factors that predict mate selection. Attention has been placed on social and cultural background characteristics such as age, social class, race, religion, and educational level.

### **THEORY BEHIND THE MARRIAGE MARKET**

Before considering individual background characteristics and interpersonal dynamics of the mate selection process, it is important to note the increasing attention given to the marriage market and the marriage squeeze. The term “marriage market” refers to the underlying assumption that we make choices about dating and marriage partners in a kind of free-market situation. Bargaining and exchange take place in contemporary selection processes, and these exchanges are based on

common cultural understandings about the value of the units of exchange. The basis for partner selection plays out in a market situation that is influenced by common cultural values regarding individual resources, such as socioeconomic status, physical attractiveness, and earning potential. Numerous studies have concluded that gender roles play a significant part in the marriage market exchange process, with men trading their status and economic power for women’s attractiveness and domestic skills. But changes in contemporary gender roles suggest that as women gain an economic viability of their own, they are less likely to seek marriage partners (Waite and Spitze 1981). Thus, the marriage market and the units of exchange are not constant but subject to substantial variation in terms of structure and selection criteria.

The premise that marital partners are selected in a rational choice process is further extended in the study of the effects of the marriage squeeze. The “marriage squeeze” refers to the gender imbalance that is reflected in the ratio of unmarried, available women to men. In theory, when a shortage of women occurs in society, marriage and monogamy are valued. But when there are greater numbers of women, marriage as an institution and monogamy itself take on lesser importance. Similarly, when women outnumber men, their gender roles are thought to be less traditional in form (Guttentag and Secord 1983).

The marriage squeeze has important effects for theoretical consideration, especially in studying the lower rates of marriage among African-American women in today’s society. Due to a shortage of African-American men, coupled with greater expectations on the part of African-American women of finding mates with economic resources (Bulcroft and Bulcroft 1993), the interplay between the marriage squeeze and motivational factors to marry suggest that future research needs to disentangle the individual and structural antecedents in mate selection. These studies also point to the complexity of mate selection processes as they take place within both the social structure and cultural gender role ideologies.

The marriage squeeze is further exacerbated by the marriage gradient, which is the tendency for women to marry men of higher status. In general, the trend has been for people to marry within the

same socioeconomic status and cultural background. But men have tended to marry women slightly below them in age and education (Bernard 1982). The marriage gradient puts high-status women at a disadvantage in the marriage market by limiting the number of potential partners. Recent changes in the educational status of women, however, suggest that these norms of mate selection are shifting. As this shift occurs, one can speculate that the importance of individual characteristics such as physical attractiveness, romantic love, and interpersonal communication will increasingly come to play important roles in the mate selection process in postmodern society (Beck and Beck-Gersheim 1995; Schoen and Wooldredge 1989).

Norms of endogamy require that people marry those belonging to the same group. Concomitantly, exogamous marriages are unions that take place outside certain groups. Again, changes in social structures, ethnic affiliations, and mobility patterns have dramatically affected the modern marriage market. More specifically, exogamy takes place when marriage occurs outside the family unit or across the genders. Taboos and laws regulating within-family marriage (i.e., marriages considered to be incestuous that occur between brother and sister, mother and son, etc.) and marriage to same-sex partners are examples of the principle of endogamy. Recent attempts have been made to legally recognize same-sex marriages, thus suggesting that norms of endogamy are tractable and subject to changes in the overall values structure of a society or social group.

In addition to endogamy and exogamy, the marriage market is further defined by norms of homogamy and heterogamy. Mate selection is considered to be homogenous when a partner is selected with similar individual or group characteristics. When these characteristics differ, heterogamy is evidenced. The norm of homogamy continues to be strong in American society today, but considerable evidence suggests we are in a period of change regarding social attitudes and behaviors with regard to interracial and interfaith unions.

Recent data suggest that the number of interracial marriages for African-Americans has increased from 2.6 percent in 1970 to 12.1 percent in 1993 (Besharov and Sullivan 1996). But African-American mate selection operates along lines of

endogamy to a larger degree than do the mate selection processes of Asian-American, Native American, or other nonwhite groups. Nearly one-half of all Asian-Americans marry non-Asians (Takagi 1994) and over half of all Native-Americans marry non-Native Americans (Yellowbird and Snipp 1994).

Similarly, rates of interfaith marriage have increased. For example, only 6 percent of Jews chose to marry non-Jewish partners in the 1960s. Today nearly 40 percent of Jews marry non-Jewish partners (Mindel et al. 1988).

The background characteristics of age and socioeconomic status also demonstrate norms of endogamy. The Cinderella story is more of a fantasy than a reality, and self-help books with titles such as *How to Marry a Rich Man (Woman)* have little basis for success.

The conditions of postmodern society are shaping mate selection patterns as they relate to endogamy and homogamy. The likelihood of marrying across social class, ethnic, and religious boundaries is strongly affected by how homogeneous (similar) the population is (Blau et al. 1982). In large cities, where the opportunity structures are more heterogeneous (diverse), rates of intermarriage are higher, while in small rural communities that demonstrate homogeneous populations, the norm of endogamy is even more pronounced.

Again, the complex interplay between the marriage market and individual motives and preferences is highlighted. The extent to which marriage outside one's social group is the result of changing preferences and attitudes or largely the result of shifting opportunity structures, known as marriage market conditions, is not clear at this time (Surra 1990).

The factors that operate in the selection process of a mate also function in conjunction with opportunity structures that affect the potential for social interaction. The evidence suggests that proximity is an important factor in determining who marries whom. Thus, those who live geographically proximate to each other are more likely to meet and marry. Early work by James Bossard (1932) shows that at the time of the marriage license application, about 25 percent of all couples live within two city blocks of each other.



Bossard's Law, derived from his empirical findings, states "the proportion of marriages decreases steadily and markedly as the distance between the consenting parties increases." Or, put more simply, "Cupid's wings are best suited for short flights." Of course, current American society has changed since the time Bossard studied mate selection patterns in Philadelphia, and there is a tendency to think that as society becomes more mobile propinquity plays less of a role in the choice of a mate. Propinquitous mate selection does not mean nonmobility, however. It is simply the case that the influence of propinquity shifts as the individual geographically shifts. Thus, one is likely to marry someone who is currently near than someone previously propinquitous. The overriding effect of propinquity is that people of similar backgrounds will meet and marry, since residential homogamy remains a dominant feature of American society. However, changing marriage patterns, such as delaying age of first marriage, will impact the strength of propinquity in the mate selection process by expanding the opportunity structures and breaking down homogenous marriage markets.

One interesting area of research that often goes overlooked in discussions of the correlates of mate selection concerns homogamy of physical attractiveness. Based on the equity theory of physical attractiveness, one would expect that persons who are similar in physical attractiveness levels would marry. Many experimental designs have been conducted to test the effects of physical attractiveness on attraction to a potential dating partner. In general, the experimental conditions have yielded the findings that the more highly attractive individuals are the most desired as dating partners. But studies of couples actually involved in selecting a mate or who are already married support the notion that individuals who are similar in attractiveness marry on their own level. Thus, while attractiveness is a socially valued characteristic in choice of a mate, the norms of social exchange dictate that we select a partner who is similar in attractiveness and is thus attainable. It is only when other highly valued factors such as wealth, wit, or intelligence compensate for deficits in attractiveness that inequity of physical attractiveness in mate selection might occur.

In review, theories of mate selection are more often applied to the study of personality characteristics or process orientations than to marriage

market conditions. It is important to note, however, that the basic assumption is that the marriage market operates in a social exchange framework. Men and women make selections under relative conditions of supply and demand with units of exchange. The market is further shaped by cultural norms such as endogamy and homogamy that can further restrict or expand the pool of eligibles.

### **NEED COMPLEMENTARITY**

While earlier work on the correlates of mate selection focused on homogamy of background characteristics, the work of Robert Winch (1958) set the stage for further investigation into the hypothesis that "opposites attract." That is, persons of dissimilar values or personality traits would marry. While value theorists speculated that similarity of values and personality would lead to great affiliation and propensity to marry, Winch posited that persons select mates whose personality traits are complementary (opposite) to their own. Inherent in Winch's theoretical work is the notion that certain specific trait combinations will be gratifying to the individuals involved. For example, a submissive person would find it gratifying or reciprocal to interact with a mate who had a dominant personality. Winch developed twelve such paired complementary personality traits, such as dominant-submissive and nurturant-receptive, for empirical testing using a very small sample of recently married couples. In Winch's work, as well as the work of others, the notion that complementarity of traits was the basis for marriage was not supported by the data.

Although empirical support for need complementarity is lacking, the concept remains viable in the study of mate selection. The appeal of the concept rests in its psychological origins, as work prior to Winch's focused primarily on structural and normative influences in mate selection. The work of Winch set the stage for research commencing in the 1960s that began to examine the processes of mate selection on the dyadic level.

### **PROCESS THEORIES OF MATE SELECTION**

The process of selecting a mate received considerable attention beginning in the 1970s. The basic form these theories take follows the "filter theory"

of Alan Kerckhoff and Keith Davis (1962). Kerckhoff and Davis found empirical support that individuals, having met through the channels of propinquity and endogamy, proceed through a series of stages or steps in the development of the relationship. According to their theory, social status variables such as social class and race operate early on in the relationship to bring people together. The next stage involved the consensus of values, during which time the couple determines the degree of similarity in their value orientations. Couples who share similar values are likely to continue to the third stage, need complementarity. However, the data collected by Kerckhoff and Davis offered only weak support for need complementarity as part of the process of mate selection.

Development of process theories of mate selection continued into the 1970s and is exemplified in the work of Ira Reiss (1960), Bernard Murstein (1970), Robert Lewis (1973), and R. Centers (1975). While these theoretical perspectives differ in terms of the order and nature of the stages, they have much in common. Melding these theories of mate selection, the following assumptions can be made concerning the stages of dyad formation that lead to marriage:

1. There are predictable trajectories or stages of dyadic interaction that lead to marriage.
2. The social and cultural background of a couple provides the context for the interpersonal processes.
3. Value similarity leads to rapport in communication, self-disclosure, and the development of trust.
4. Attraction and interaction depend on the exchange value of the assets and liabilities that the individuals bring to the relationship.
5. Conditional factors such as age, gender, or marital history may influence the order or duration of the stages, or the probability that the relationship will end in marriage.

All the studies of the mate selection process have struggled with methodological difficulties. Most studies have relied on small, volunteer samples of couples. Most have used college-age, never-married couples. Finally, most studies have made extensive use of retrospection in assessing the

process of dyad formation rather than collecting longitudinal data. These methodological difficulties may, in part, account for the recent decline in the number of studies examining the process of mate selection.

Furthermore, these stages may or may not result in marriage, but the primary focus of the research is on relationships that endure or terminate in marriage. Therefore, relatively little is known about the mate selection process as it pertains to rejection of a potential mate or how such terminations of relationships affect subsequent mate selection processes.

More current research has begun to shift away from antecedents that lead to legal marriage and turn instead to disentangling the trajectory of relationship development over the life course. More attention will turn to the formation and development of interpersonal relationships that may move through stages of romance, cohabitation, friendship, marriage, divorce, and so forth. Emphasis on relationship quality and durability, gender role negotiations, commitment processes, and romantic love have recently taken on increased importance in social science studies of mate selection (Surra and Hughes 1997; Houts et al. 1996; Surra 1990).

Many of the theories have also overlooked the influence of peer groups and family members in the mate selection process. The theoretical and empirical inquiry that has paid attention to peer and kin influences is restricted to studies of dating. Unfortunately, studies of dating and studies of mate selection have not been sufficiently integrated to provide the field with adequate data concerning the interrelationships between dating and mate selection processes.

Yet another area of research that has the potential for contribution to further understanding of the mate selection process is studies of romantic love. Process theories of mate selection seldom examine love as the basis, or even as a stage, in the development of a heterosexual relationship. While there is a large body of empirical and theoretical work on romantic love, conceptually the studies of love have been treated as quite distinct from the research on mate selection. Contrary to popular opinion, the relationship between love and marriage is not well understood.

### FUTURE DIRECTIONS

As the family system changes in American society, so too the direction of research on mate selection shifts. As more couples delay first marriage, examination of courtship cohabitation becomes more salient. Future studies of courtship cohabitation will most likely examine the association between increasing rates of cohabitation and decreasing rates of marriage. On the individual level, the effects of the cohabitation experience on the decision to marry also warrant attention.

Research is just beginning on the mate selection process of remarriage (Bulcroft et al. 1989; Rodgers and Conrad 1986; Spanier and Glick 1980). While some factors that predict first marriage may remain constant in remarriage, such as endogamy and propinquity, other factors may come into play in remarriage. For example, age homogeneity may be less of a factor in remarriage since the pool of eligible mates is impacted by sex ratio imbalance. The exchange relationship in the mate selection process also differs in remarriage, since presence of children, prior marital history, and the economic liabilities of child support and alimony bring new dimensions to considerations of remarriage. Of particular interest are barriers to remarriage in the middle and later years of the life cycle, such that cohabitation or serious dating may offer more long-term rewards to the couple than legal marriage might provide. Thus, the strong profamilial norms that encourage the younger members of society to marry dissipate at mid and later life. Low rates of remarriage for individuals over the age of 50, in part, indicate that societal pressure to marry is greatly reduced.

Last, it has generally been assumed that homogamy of background characteristics leads to similarity of values, shared marital role expectations, rapport, and intimacy in the process of mate selection. But due to changing gender role expectations, this assumption may no longer be valid. As a result, more attention needs to be given to the process of role negotiation as part of the mate selection process.

In summary, studies of mate selection began with understanding the correlates of mate selection. Social scientists began by studying demographic data on homogamy in religion, social class, age, and other factors as these variables related to who married whom. For a brief period in the

1960s through the early 1980s, attention was turned to theories and data that examined the process of mate selection. Current research in the 1990s has not abandoned the study of the correlates and theories of mate selection, but as the nature of the family system changes, researchers have begun to consider that the generalizability of theories and findings may be limited when a researcher is trying to explain mate selection at a point later than young adulthood. Recent studies on the courtship processes of divorced (O'Flaherty and Workman 1988) and later life mate selection (Veevers 1988) point to the future focus of theories and research on mate selection processes.

### REFERENCES

- Beck, Ulrich, and Elisabeth Beck-Gernsheim 1995 *The Normal Chaos of Love*. Cambridge, Mass.: Blackwell.
- Bernard, Jessie 1982 *The Future of Marriage*. New York: Columbia University Press.
- Besharov, D., and T. Sullivan 1996 "Welfare-Reform and Marriage." *Public Interest* 125:81-94.
- Blau, Peter, Terry Blum, and Joseph Schwartz 1982 "Heterogeneity and Intermarriage." *American Sociological Review* 47:45-62.
- Bossard, James 1932 "Residential Propinquity as a Factor in Marriage Selection." *American Journal of Sociology* 38:219-224.
- Bulcroft, Kris, Richard Bulcroft, Laurie Hatch, and Edgar F. Borgatta 1989 "Antecedents and Consequences of Remarriage in Later Life." *Research on Aging* 11:82-106.
- Bulcroft, Richard, and Kris Bulcroft 1993 "Race Differences in Attitudinal and Motivational Factors in the Decision to Marry." *Journal of Marriage and the Family* 55:338-355.
- Centers, Richard 1975 *Sexual Attraction and Love: An Instrumental Theory*. Springfield, Ill.: C. C. Thomas.
- Guttentag, M., and P. Secord 1983 *Too Many Women?* Thousand Oaks, Calif.: Sage.
- Houts, Renate, Elliot Robins, and Ted Houston 1996 "Compatibility and Development of Premarital Relationships." *Journal of Marriage and the Family* 58:7-20.
- Kerckhoff, Alan, and Keith Davis 1962 "Value Consensus and Need Complementarity in Mate Selection." *American Sociological Review* 27:295-303.
- Lewis, Robert 1973 "A Longitudinal Test of a Developmental Framework for Premarital Dyadic Formation." *Journal of Marriage and the Family* 35:16-27.

- Mindel, Charles, Robert Haberstein, and Roosevelt Wright 1988 *Ethnic Families in America: Patterns and Variations*. 3rd ed., New York: Elsevier North Holland.
- Murstein, Bernard 1974 *Love, Sex, and Marriage through the Ages*. New York: Springer.
- 1976 *Who Will Marry Whom? Theories and Research in Marital Choice*. New York: Springer.
- O'Flaherty, Kathleen, and Laura E. Workman 1988 "Courtship Behavior of the Remarried." *Journal of Marriage and the Family* 50:499-506.
- Reiss, Ira 1960 "Toward a Sociology of the Heterosexual Love Relationship." *Marriage and Family Living* 22:139-145.
- Rodgers, Roy, and Linda Conrad 1986 "Courtship for Remarriage: Influences on Family Reorganization after Divorce." *Journal of Marriage and the Family* 48:767-775.
- Schoen, Robert, and John Wooldredge 1989 "Marriage Choices in North Carolina and Virginia 1969-71 and 1979-81." *Journal of Marriage and the Family* 51:465-481.
- Spanier, Graham, and Paul Glick 1980 "Paths to Remarriage." *Journal of Divorce* 3:283-298.
- Surra, Catherine 1990 "Research and Theory on Mate Selection and Premarital Relationships in the 1980s." *Journal of Marriage and the Family* 52:844-856.
- , and Debra Hughes 1997 "Commitment Processes in Accounts of the Development of Premarital Relationships." *Journal of Marriage and the Family* 59:5-21.
- Takagi, Diana 1994 "Japanese American Families." In R. L. Taylor, ed., *Minority Families in the United States: A Multicultural Perspective*. Engelwood Cliffs, N.J.: Prentice Hall.
- U.S. Bureau of the Census 1986 *Statistical Abstract of the United States*. Washington, D.C.: Government Printing Office.
- Veevers, Jean 1988 "The 'Real' Marriage Squeeze: Mate Selection Mortality and the Marriage Gradient." *Sociological Perspectives* 31:169-189.
- Waite, L., and G. Spitze 1981 "Young Women's Transition to Marriage." *Demography* 18:681-694.
- Winch, Robert 1958 *Mate Selection: A Study of Complementary Needs*. New York: Harper and Row.
- Yellowbird, Michael, and C. Matthew Snipp 1994 "American Indian Families." In R. L. Taylor, ed., *Minority Families in the United States: A Multicultural Perspective*. Engelwood Cliffs, N.J.: Prentice Hall.

KRIS BULCROFT

## MATERIALISM

*Materialism* posits the epistemological primacy of matter over ideas, mind, values, spirit, and other incorporeal phenomena. Philosophical perspectives stressing the fundamental importance of physical conditions and needs have grown more elaborate with the increasing differentiation and autonomy of secular knowledge from religion. Materialists oppose magical, religious, and metaphysical explanations of worldly affairs, criticizing their role as mystifications and socioeconomic or political legitimations. The enduring debate over materialism and idealism (which gives primacy to ideas) centers on the two approaches' relative effectiveness as guides to scientific, technical, and sociopolitical practices. In social science, materialism refers to often tacit metatheories, or heuristic devices, which frame distinctive types of research problems, hypotheses, concepts, and theories stressing the causal force of physical realities on sociocultural matters.

Two contrasting threads of the materialist tradition have divergent consequences for the behavioral and social sciences. *Reductionists* posit that phenomena are determined strictly by physical causes. Their view that existential knowledge is a reflection of corporeal conditions denies the autonomy of psychological and sociocultural factors and, thus, suggests that the behavioral and social sciences have no distinct content. *Nonreductionists* hold that psychological and sociocultural phenomena arise from and are dependent upon physical substrata. They also accord "primacy" to physical realities and corporeal impulses, motives, values, representations, and interests, seeing them as basic constraints that "determine," or channel, the direction of human practices. By contrast to reductionists, however, they treat the sociocultural realm as an "emergent," "sui generis," "relatively autonomous" domain having distinct properties, processes, and laws and exerting reciprocal causality with the material realm. Implying interpenetrating sociocultural and material spheres, they include socially constituted entities and processes (e.g., technology and labor) as prime "material" determinants.

Modern materialism is rooted in ancient Greek conceptions of elementary bodies. Atomistic philosophers held that all existing things are composed of indivisible, ultimate objects of the same

material in perpetual motion in empty space. They argued that perceptible objects derive from atoms of various sizes and shapes colliding, getting entangled, and forming different combinations, and that sensations arise from atoms passing through the sense organs and impacting on the soul (also composed of atoms). From the start, materialists considered knowledge to be a “subjective” manifestation of “objective” reality, holding that physical realities ultimately determine individual experiences and sociocultural constructions. However, claims about the scope of this determination varied widely with the degree to which the thinker adhered to reductionist or nonreductionist presuppositions.

Atomistic materialism reemerged as a major cultural force during the Renaissance science revolution. Galileo and Newton again portrayed physical reality as ultimate particles moving in empty space, but their distinction between precisely measurable, primary sensory qualities (i.e., length, width, weight, figure) and nonmathematizable secondary qualities (e.g., color, smell, taste, texture) established a sharp boundary between objective and subjective experience and decisive methodological standard for distinguishing science from metaphysical, aesthetic, or sociocultural thought. The capacity of diverse observers, employing the same experimental techniques, to arrive at similar findings supported materialist claims about the primacy of the physical world and certainty of objective knowledge.

The extraordinary success of Newtonian science contributed greatly to an extensive secularization of knowledge that reduced barriers to materialist approaches in human affairs. For example, Hobbes argued that social actions are also effects of matter in motion; material primacy is manifested in the dominant drive for self-preservation, all-pervasive power struggles, and subsequent need for absolute monarchy. Locke’s Newtonian theory of mind held that primary sense qualities reflect external objects and that complex ideas merely combine the simple ones received directly from sense experience. However, Descartes’ dualistic vision of a materialist physical world and an autonomous mind blessed with innate ideas of divine origin exemplifies the seventeenth century tendency to provide separate grounds for science, which avert direct subversion of religion. The power of the Roman Catholic Church and its censors, who saw

materialism as a subversive force, was already inscribed powerfully in the earlier trial and conviction of Galileo and in his concession to dualism. Even Hobbes left space for spiritual realities. The strong subjectivist currents and subject-object dualism in Western philosophy derived from its earlier religious roots and its lack of autonomy in relation to the Church.

Enlightenment thinkers fashioned a new cultural space for free inquiry and autonomous science, subverting the power of religion and paving the way for modern materialism. La Mettrie, D’Holbach, and Diderot held that all experience has material causes. Idealizing Newtonian mechanics, the philosophes believed that naturalistic explanation of all phenomena would demystify religious and metaphysical superstition, limit the rule of the Church and nobility, and animate scientifically guided social reform. Revolutionary advances in eighteenth-century medicine, chemistry, and biology upheld their faith in materially based rationality, disenchantment, and progress. The later social revolutions against the ancien régime, secularization of political power, and gradual rise of liberal democracy favored the spread of materialist thinking in new social domains.

Manifesting Enlightenment culture, Karl Marx set the agenda for modern materialism. Even today’s debates about the topic center on different interpretations of his work. Following Feuerbach, Marx charged that religion and its secular, idealist substitutes (e.g., Hegelian philosophy) are “inverted,” or “alienated,” projections of human capacities and potentialities. As “natural beings,” he argued, people must satisfy their needs by appropriating and shaping physical objects. Although he attacked Hegel’s speculative history of “spirit,” Marx retained his view that people create themselves and their societies through their labor. Young Marx called for a “true” materialism, or a new science of humanity focusing on “social relationships.” He wanted to illuminate ideologically obscured forms of socially structured production and exploitation, which, in his view, constitute the “real history” of “corporeal” human beings.

Divergent strains of Marxian materialism—“dialectical materialism” and “historical materialism”—are rooted ultimately in tensions between the political and scientific sides of Marx’s and

Engels' works. The terms were coined after Marx's death and codified into a variety of schematic orthodoxies. *Dialectical materialism* is a philosophy of nature developed first by Engels, elaborated by Plekhanov and Lenin, and fashioned by Marxist-Leninists into an eschatology of the communist movement. Serving primarily as a political ideology and metalanguage that justified communist states and insurgencies, it did not have much impact on social science outside the communist regimes and movements. By contrast, *historical materialism*, Marx's main metatheory of social development, has much relevance for sociology. Yet even this approach was stated variously by him, and gave rise to conflicting interpretations. The elder Engels ([1890–1894] 1959, pp. 395–400), who earlier helped Marx frame his theories, berated first-generation “Marxists” for using materialism “as an excuse for *not* studying history,” or as a dogmatic Hegelian “lever for construction,” instead of as “a guide to study.” Accepting partial responsibility for the later vulgar interpretations, he admitted that Marx and he, in the heat of political battle, often spoke with too much certainty. However, he insisted that Marx and he opposed the idea of all-encompassing, mechanistic causality by narrowly conceived economic factors and that their materialism meant nothing more than that “the production and reproduction of real life” is “the *ultimately* determining element in history.”

Engels implied a problematic split in Marx's thought. On the one hand, Marx sometimes implied, in Hegelian fashion, progressively unfolding stages of history, animated by relentless technological advance and ever more unified classes rationally taking account of material factors, controlling them, and speeding history to an “inevitable” emancipatory conclusion. Such points were warranties for his political program. On the other hand, the scientific side of Marx's work breaks from teleological thinking. His *historical materialism* stresses empirical inquiry about specific material conditions and practices and their sociocultural consequences among particular groups in finite space and time, and calls for reconstruction of theory and practice in light of the findings. This approach has more affinity for the empirical and secular thrust of the Darwinian revolution than for earlier Enlightenment ideas of progress and science, which implied a new faith or civil religion. After the failed revolutions of 1848 and the rise of

the new Napoleonic dictatorship in France, Marx qualified his formerly highly optimistic views about “making history,” holding that we do not make it in accord with circumstances or outcomes that we freely choose. This insight was precursory to his mature work on capitalism, which, although retaining certain Hegelian taints, manifested clearly his historical materialism.

Following Adam Smith, Marx held that a specialized division of labor enhances productive powers geometrically. However, by contrast to Smith, he specified that capitalism's exceptional productivity derives expressly from its historically specific forms of complex cooperation, which end the relative isolation of peasants and other independent producers. Marx charged that economists mystify this process by obscuring the fundamental role of labor and portraying capitalist growth as if it were animated by a “fantastic . . . relation between things,” rather than by a specific type of “social relation” (Marx [1887] 1967, pp. 71–83). He saw the unequal relationship between capitalists and workers as the “secret” of capitalist accumulation. Treating associated labor and related forms of social organization, extraction, and domination as the decisive “material” forces in human affairs, Marx averted the reductionist physicalism of the early atomists, awkward dualism of the Renaissance thinkers, and mechanistic determinism of the dialectical materialists.

Marx and Engels ([1845–1846] 1964, pp. 31–32) said that the “first premise of all human history” is the production of “the means of subsistence” and a characteristic “*mode of life*.” As productive forces are refined, they held, growing surpluses offer provision beyond subsistence and allow certain individuals and strata to be freed from direct production for other activities, including the creation of fresh technical knowledge and new productive forces. Marx and Engels contended that incremental advances in production generate increasingly elaborate socio-cultural differentiation, class structure, and domination and more universal “class struggles” over the productive forces. In their view, material determination in the “last instance” means that socially structured patterns of production, extraction, and disposition of surplus shape social development. A materialist sociology focuses on these processes and their consequences.

The central tenet of historical materialism is that the “base,” or “mode of production,” determines “superstructure,” or “the forms of intercourse.” The base is composed of “productive forces,” which contribute directly to production for material needs (i.e., natural resources, tools, technical knowledge, labor power, and modes of cooperation), and “property relations,” which provide certain “classes” effective control over production and relegate others to direct labor. “Superstructure” is composed of nonproductive types of social intercourse (i.e., noneconomic forms of public organization, private association, and thought), which are determined by and, in turn, maintain, or reproduce, the base.

There is a core technological root to Marx’s materialism, but he also implied reciprocal causality between the physical and sociocultural realms. Thus, he was not a reductionist. Although granting productive forces ultimate theoretical primacy, his historical analyses dwelled more on property relations and class dynamics. Moreover, he did not hold that material determination operated with the same intensity throughout society. Rather, he emphasized that its strongest force is exerted in institutions that play a direct role in reproducing the mode of production and that the rest of society and culture has much more autonomy. For example, under capitalism, mainstream public discourses about property rights or free trade are central to ideology, but most music and fiction, as well as Marx’s critical writings, bear only the broad imprint of the level and form of productive development. Marx envisioned society as composed of interdependent social structures and cultural forms, implying that each element has a distinct impact on the whole. Yet he still saw the base as the most decisive factor shaping overall society and its internal relations. He treated the productive activities of subordinate producers and surplus extraction by superordinate classes as the most central and obscured processes in social life and the root of the most major institutions, conflicts, and transitions.

In a much-debated argument about the capitalist mode of production, Marx held that the rate of profit would fall under highly mechanized production. He believed that the trend toward automation in heavy industry reduces the proportion of direct laborers to fixed capital, diminishing the ultimate source of capitalist profit and requiring hyperexploitation of remaining workers to pay for

vastly increased technical investment. Marx argued that the classes would be compressed into a tiny monopoly stratum of owners and a huge impoverished mass of de-skilled workers and permanently unemployed people. However, he also held that science and technology, having replaced direct labor as the leading productive force, would provide the material basis for proletarian revolution and an emancipated, postcapitalist order (i.e., scientific production and administration, greatly reduced work and material necessity, and uncoerced cooperation). This hopeful scenario about collective agency and planning as replacing blind history, and material determinism as becoming truly “in the last instance,” went beyond the scientific thrust of historical materialism and reflected the Hegelian residue in his thought.

By the early twentieth century, pathbreaking ideas about electromagnetic fields, relativity theory, and quantum mechanics relativized space and time, eroded the borders between energy and mass, and fashioned a new relational cosmology that undermined Newtonianism. Similarly, Nietzsche, James, Dewey, Heidegger, and Wittgenstein subverted the foundations of modern Western philosophy. Major social changes accompanied the intellectual shifts. A more complex system of classes and subclasses, new types of workplaces, indirect forms of ownership and control, and state intervention altered the structure of capitalism and blurred the line between base and superstructure. Stalinism; fascism; resurgent irrationalism; and persistent racial, ethnic, gender, and religious splits also cried out for new types of theory. In this climate, Marxists tended to drift toward new types of cultural theory or to embrace dialectical materialism, which remained the official ideology of global communism.

In the 1930s, when sociology was being transformed into a professionally specialized discipline in the United States, Talcott Parsons relegated Marx to the status of a footnote in the history of social theory and gave a strong primacy to values and ideas that turned Marx on his head. Parsons’s leading role in post-World War II sociology helped harden mainstream opposition to materialism. Claiming that class was no longer a prime basis of association, a determinant of political and cultural beliefs, or a source of major social conflicts, numerous North American sociologists and many

thinkers from other liberal democracies argued that materialism did not come to terms with the nascent “postindustrial society.” Although certain critics dissented, most mainstream sociologists dismissed materialism as a crude ideology, or simply ignored the approach.

Materialist themes appeared much more prominently in later 1960s social theory, animated, in part, by New Left attacks on functionalism and positivism. Very influential among European theorists, Lewis Althusser’s “structuralist” interpretation of Marx radically revised materialism. However, G. A. Cohen’s *Karl Marx’s Theory of History* (1978) helped stimulate wider efforts to reconstruct historical materialism (e.g., Anderson 1974a, 1974b; Shaw 1978; McMurtry 1978; Roemer 1982; Bhaskar 1989). Dispensing with overt political facets and overcoming conceptual gaps, the new works were more systematic versions of the original approach. However, they did not have much impact on mainstream sociology.

Materialist theory fell on hard times in a climate where “post-materialist” identity politics were ascendent over labor-centered or class politics, and failed communism and resurgent neoliberalism seemed to doom the socialist “alternative” (Anderson 1983; Antonio 1990). In this climate, a new “cultural sociology” blossomed as part of a broader interdisciplinary turn to “representation” and “discourse.” Postmodern arguments about the primacy of culture were often pitched directly against Marxian materialism, which they declared moribund. These new cultural theories tended to shift the focus among “critical sociologists” from “structural determinants” (i.e., production and labor) to “cultural surfaces,” especially mass entertainment and other signifiers of mass consumption. Many of these theorists argued that a defining feature of postmodern culture is its increased, or, even, total autonomy from material underpinnings.

Although materialist theory has declined in recent years, historical-comparative and empirical work, manifesting tacit materialism and clear connections to nondoctrinaire forms of Marxism, was well established in North American sociology by the 1980s. It still flourishes among significant subgroups in the discipline. These thinkers do not directly employ the base-superstructure model

(which is now associated with dialectical materialism and dogmatic Marxism), but they share historical materialist presuppositions, albeit often fused with other traditions. Recent historical changes are intensifying interest in materialist themes. First, neoliberal globalization and restructuring stimulate heated debates over multinational firms, international finance, socioeconomic dislocations, and related public policy initiatives. Second, the collapse of eastern European communism and global decline of communism raise major questions about socioeconomic reconstruction in postcommunist society and detach materialism from communist politics. Third, new materialist critiques counter claims about the primacy of culture and representation in interdisciplinary theory circles. Historical materialism is a heuristic device that poses problems and hypotheses, and is perhaps best judged on this account. Consider the current “materialist” issues below.

#### MATERIALIST ANALYSES OF THE NEW PHASE OF CAPITALISM

As in Marx’s time, today’s materialism arises in a climate of perceived rupture, or the end of the post-World War II socioeconomic system and its distinctive patterns of growth, regulation, and geopolitics. Materialists focus on the shifts in the system of production and their impacts on other aspects of social life, especially on inequality, conflict, and politics.

**Globalization and Restructuring.** Recent materialist work focuses on the new “flexible” forms of production, network firms, international division of labor, and multinational economic blocs—for example, the North American Free Trade Association (NAFTA) and the European Union (EU). The core debates among these analysts concern the shape, permanence, and consequences of the new global economy and the degree to which it departs from postwar capitalism (i.e., they ask whether a new multinational, or “post-Fordist,” phase of capitalism has replaced “Fordism,” or simply modified it.) (e.g., Harvey 1989; Harrison 1994; Gordon 1996).

**Income and Wealth Inequality.** Neoliberal deregulation, free trade, and recommodification of public goods have increased global disparities



of wealth, income, jobs, and life chances (Braun 1997; Davis 1992). Inequality is most extreme between the richest and the poorest nations, but it has also grown considerably within societies, including wealthy ones, such as the United States. Materialists debate the scope of economic inequality, its relation to global restructuring, and its consequences.

**Class, Labor, and Politics.** Materialists also focus on the impact of restructuring and polarizing labor markets on organized labor. For example, they ask: Are new types of unions arising among formerly unorganized low-wage service workers and professional groups? Will transnational labor organizations arise in the emergent economic blocs? Are new fusions of class politics and cultural politics arising?

**The State.** Materialists raise questions about the impact of capital mobility and new forms of international financial and labor markets on the state's capacity to protect and regulate its various environments and maintain public goods (e.g., health, education, welfare, retirement). They also study the role of the state's police and military arms in the global political economy.

**Environmental Issues and Sustainability.** Increasing threats of resource depletion, global warming, hazardous waste, and overall environmental destruction, the need for alternative forms of energy and technologies, and the communist regimes' dismal environmental records have stimulated sharp criticism of earlier materialism's "productivism" and inattention to the costs, risks, and limits of growth. They pose questions about sustainable growth and material limits to political aspirations.

**Communication and Information Technologies.** These industries play a major role in globalization and restructuring and, thus, are a central focus in fresh materialist inquiries.

#### FOUNDATIONAL ISSUES IN MATERIALIST THEORY

Although certain thinkers (e.g., Wallerstein 1991; Postone 1993; Wood 1995; Wright 1997) have begun reframing materialism in light of the recent historical changes, more fundamental rethinking

of the tradition is likely, especially in the two broad areas discussed below.

**The Role of Technology.** This area has long been a central focus of materialism. In the late nineteenth century, Marx held that science and technical knowledge were already becoming the primary productive force, altering fundamentally the capitalist mode of production and materialist dynamics of all preceding epochs. Certain contemporary theorists argue that information and communication technologies are giving rise to an "informational society" (e.g., Poster 1990; Castells 1996). These thinkers raise basic questions about previous ideas of production, property, labor, organization, and other aspects of earlier materialism, which could lead to a new generation of theories framed around a knowledge-driven logic of material development.

**Materialism and Politics.** The effort to unify theory and practice, and the consequent fusion of science and politics, generate distinct resources (i.e., they provide a "critical" thrust) as well as major sources of tension (i.e., between science and ideology) in Marxian materialism. As "critical theorists," materialists focus on the contradictions between democratic ideology's claims about freedom, equality, abundance, and participation, on the one hand, and actual conditions of capitalist and state socialist societies, on the other. They claim to illuminate the determinate, historical possibilities for progressive social change (i.e., "emancipatory" structural factors, cultural conditions, and social movements), which favor reconstructing modern societies according to revised versions of their own democratic ideals. However, the failure of proletarian revolution and communism have blurred materialist political aims. Some materialists abandon critical theory completely, fashioning a strictly empirical sociology, while others substitute highly generalized ideals of "discursive" or "radical" democracy and cultural politics for socialism. It is hard to predict the fate of critical theory, but increasing inequalities in wealth, polarizing labor markets, and declining social benefits raise major questions about unmet needs and the material bases of inclusive democratic citizenship. The bankruptcy of soviet-style regimes and the erosion of social democracies make materialist political goals an open question, but the consequent loss of certainty undercuts the tradition's

dogmatic side and increases chances for theoretical reconstruction and restored vitality.

REFERENCES

- Anderson, Perry 1974b *Lineages of the Absolutist State*. London: New Left Books.
- 1974a *Passages from Antiquity to Feudalism*. London: New Left Books.
- 1983 *In the Tracks of Historical Materialism*. London: New Left Books.
- Antonio, Robert 1990 "The Decline of the Grand Narrative of Emancipatory Modernity: Crisis or Renewal in Neo-Marxian Theory?" In George Ritzer, ed., *Frontiers of Social Theory: The New Syntheses*. New York: Columbia University Press.
- Bhaskar, Roy 1989 *Reclaiming Reality*. London: New Left Books.
- Braun, Denny 1997 *The Rich Get Richer: The Rise of Income Inequality in the United States and the World*, 2nd ed. Chicago: Nelson-Hall Publishers.
- Castells, Manuel 1996 *The Information Age: Economy, Society and Culture: Vol. 1: The Rise of Network Society*. Malden, Mass.: Blackwell.
- Cohen, G. A. 1978 *Karl Marx's Theory of History: A Defence*. Princeton, N.J.: Princeton University Press.
- Davis, Mike 1992 *City of Quartz: Excavating the Future in Los Angeles*. New York: Vintage Books.
- Engels, Friedrich (1890–1894) 1959 "Letters on Historical Materialism." Pp. 395–412 in Lewis S. Feuer, ed., *Karl Marx and Friedrich Engels: Basic Writings on Politics and Philosophy*. Garden City, N.Y.: Anchor Books.
- Gordon, David M. 1996 *Fat and Mean: The Corporate Squeeze of Working Americans and the Myth of Managerial "Downsizing"*. New York: Free Press.
- Harrison, Bennett 1994 *Lean and Mean: The Changing Landscape of Corporate Power in the Age of Flexibility*. New York: Basic Books.
- Harvey, David 1989 *The Condition of Postmodernity: An Inquiry into the Origins of Cultural Change*. Cambridge, Mass.: Blackwell.
- McMurtry, John 1978 *The Structure of Marx's World-View*. Princeton, N.J.: Princeton University Press.
- Marx, Karl (1887) 1967 *Capital: A Critique of Political Economy*, vol. 1. New York: International Publishers.
- , and Frederick Engels (1845–1846) 1964 *The German Ideology*. Moscow: Progress Publishers.
- Poster, Mark 1990 *The Mode of Information: Poststructuralism and Social Context*. Chicago: University of Chicago Press.

Postone, Moishe 1993 *Time, Labor, and Social Domination: A Reinterpretation of Marx's Critical Theory*. New York: Cambridge University Press.

Roemer, John E. 1982 *A General Theory of Exploitation and Class*. Cambridge, Mass.: Harvard University Press.

Shaw, William H. 1978 *Marx's Theory of History*. Stanford, Calif.: Stanford University Press.

Wallerstein, Immanuel 1991 *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Cambridge, U.K.: Polity.

Wood, Ellen Meiksins 1995 *Democracy Against Capitalism: Renewing Historical Materialism*. New York: Cambridge University Press.

Wright, Erik Olin 1997 *Class Counts: Comparative Studies in Class Analysis*. New York: Cambridge University Press.

ROBERT J. ANTONIO

## MATHEMATICAL SOCIOLOGY

*Mathematical sociology* means the use of mathematics for formulating sociological theory more precisely than can be done by less formal methods. The term thus refers to an approach to theory construction rather than to a substantive field of research or a methodology of data collection or analysis; it is also not the same as statistical methods, although it is closely related. Mathematical sociology uses a variety of mathematical techniques and applies to a variety of different substantive research fields, both micro and macro.

Theory involves abstraction from and codification of reality; formulation of general principles describing what has been abstracted; and deduction of consequences of those formulations for the sake of understanding, predicting, and possibly controlling that reality. When social phenomena can be described in mathematical terms, the deductive power of mathematics enables more precise and more detailed derivations and predictions based on original premises.

Mathematical expression also enables sociologists to discover that the same abstract forms and processes sometimes describe what seem to be diverse social phenomena. If the same type of formulation describes both the spread of a disease and the adoption of an innovation, then a common type of process is involved and the theorist

can search further for what generates that commonality. Ideally, therefore, mathematics provides the basis for very general and powerful integrative theory.

The vigor of mathematical sociology varies widely over the different subfields of sociology. Precise formulation requires precise observations and careful induction of general patterns from those observations. Some sociologists have rejected any attempts to quantify human behavior, either on the grounds that what is important is in principle not subject to precise measurement or from a philosophical unwillingness to consider human behavior to be in any way deterministic. That issue will not be addressed here, but clearly it is much easier to obtain precise information in some areas of inquiry than in others. For example, census data provide reasonably precise counts of many sociological interesting facts, with the consequence that mathematical demography has long flourished.

Mathematics can be thought of as an elegant logic machine. Application of mathematics to any substantive discipline involves careful translation of the substantive ideas into mathematical form, deriving the mathematical consequences, and translating the results back into a substantive interpretation. There are three key aspects to the process: (1) finding a satisfactory way of expressing the substantive ideas in mathematical terms, (2) being able to solve the mathematical puzzle, and (3) eventually being able to compare derivations from the model with data from the substantive application.

The mathematical expression of a substantive theory is called a *model*. Often, models are created with primary emphasis on their being tractable, or readily solved. Some types of models appear frequently in substantive literature because they are widely known, are relatively simple mathematically, and have easy solutions. Basic models and applications can be found in Coleman (1964), Fararo (1973), and Leik and Meeker (1975). Mathematics and sociological theory are discussed in Fararo (1984). A wealth of more complex models can be found in specialized volumes and in periodicals such as the *Journal of Mathematical Sociology*.

Simple models have the advantage of presenting an uncluttered view of the world, although derivations from such models often do not fit

observed data very well. When the goal is heuristic, a simple model might be preferable to one that matches more closely the reality being modeled. However, when the goal is accurate prediction and possible control, then more complex models are typically needed.

There are two general questions to be raised in deciding whether developing a mathematical model has been useful. One is whether the mathematics of the model lead to new ideas about how the system being modeled operates. This is purely a theoretical question, concerned with understanding reality better by creating an abstraction of it that enables us to think more clearly about it. The second question concerns how well the model fits that reality, and is a statistical question.

Statistical models are concerned solely with fitting an underlying mathematical model to data from a sample of real-world cases. The underlying model may be complex or simple, but the statistical concerns are whether the sample can be assumed to represent adequately the population from which it was selected and whether the parameter (equation constant) estimates based on that sample can be considered accurate reflections of how the variables of the underlying model are related in that population.

#### A HISTORY OF RECENT APPROACHES TO FORMAL THEORY

During much of this century, there has been concern in sociology over the relationship between theory and research. Whereas theory was abstract and typically discursive, research increasingly employed statistical methods. The gulf between verbal statements about theoretical relationships and statistical tests of empirical patterns was great. Beginning in midcentury, some sociologists suggested that one way to bridge this gulf is to translate theoretical ideas into mathematics. Proposals for ways that theoretical ideas could be represented mathematically included those by Simon (1957) and several summarized by Berger et al. (1962).

Another approach to more formal presentation of theories that seemed to offer a bridge over the chasm between verbal statements and statistical tests was Zetterberg's concept of axiomatic theory (1965). The popularized result was *axioms*

in the form of monotonic propositions (“The greater the  $X$ , the greater the  $Y$ ”) and hypotheses derived solely from concatenating (multiplying) the signs of the relationships. For example, the axioms “As  $A$  increases,  $B$  increases” and “As  $B$  increases,  $C$  decreases” lead to the hypothesis that “As  $A$  increases,  $C$  decreases.” Standard statistical tests of the deduced hypotheses were presumed to be proper tests of the theory that generated the axioms. The approach was quick, convenient, and readily understood, and did not require expertise in mathematics or statistics. Consequently, it quickly became popular.

Numerous inadequacies with such theories soon became apparent, however (see for example critiques in Hage 1994), and interest turned to *path models* based on earlier work by Wright (1934). Path models assume that empirical measurements coincide exactly with theoretical concepts; that all variables are continuous (or reasonably close to continuous); that all relationships (paths) are bivariate, linear, and causal; that there is no feedback in the system (the recursive assumption); and that all relevant variables have been included in the model (the “closed-system” assumption). The underlying theoretical model is therefore very simplistic but does allow for the introduction of assumptions about multiple causal paths, and for different types of causal relationships including intervening, spurious, modifying, and counteracting.

If the assumptions are reasonable, then the causal effects, or path coefficients, are equal to ordinary multiple regression coefficients. That is, the underlying mathematical model feeds directly into a well-known statistical model, and the tie between theory and research seems well established. Are the assumptions reasonable?

The closed-system assumption, for example, implies that there is no correlation between the prediction errors across the various equations. If error correlations appear, then path coefficient estimates based on the statistical model will be in error, or biased, so the theory will not be tested properly by the statistics. Only two solutions are possible: (1) add more variables or paths to the model or (2) develop a statistical model that can accommodate correlated errors.

The assumption that measurement equals concept poses a different problem whenever various

scales or multiple indicators are used to represent a theoretical concept not readily assessed in a simple measure. Traditionally, the statistical model called *factor analysis* has been used to handle this measurement-concept problem, but factor analysis was not traditionally linked to the analysis of theoretical systems.

Recent years have seen very extensive development and elaboration of statistical models for linear systems, and these models address both the correlated error and the measurement-concept problems while allowing departure from recursiveness. They are called *linear structural models* (Joreskog 1970; Hayduk 1987). The underlying mathematical model still said that all variables are continuous and all relationships bivariate, linear, and causal. This development focused entirely on technical statistical questions about bias in parameter estimates. Furthermore, although the theory generally supposed that one variable affects another over time, the data are normally from only one or a very few points of time. From a general theory point of view, the underlying linear model’s assumptions, which were imposed for tractability, are highly restrictive.

### MOVING BEYOND LINEAR MODELS

Consider the assumption of linearity. If a dynamic process is being modeled, linear relationships will almost always prove faulty. How change in some causal factor induces change in some consequent system property is typically constrained by system limits and is likely to be altered over time through feedback from other system factors.

As a disease like acquired immune deficiency syndrome (AIDS) spreads, for example, the rate at which it spreads depends on how many people are already infected, how many have yet to be infected, and what conditions allow interaction between the not-yet-infecteds and the infecteds. With very few infected the rate of spread is very small because so few cannot quickly infect a very large number of others. As the number of infecteds grows, so does the rate of spread of the disease, because there are more to spread it. On the other hand, if nearly everyone had already been infected, the rate of spread would be small because there would be few left to spread it to. To complicate

theoretical matters further, as the disease has generated widespread concern, norms governing sexual contact have begun to change, influencing the probabilities of transmission of the virus. In mathematical terms, the implication is that the rate of change of the proportion infected (i.e., the rate of spread of the disease) is not constant over time, nor is it a constant proportion of change in any variable in the system. In short, the process is inherently nonlinear.

Nonlinear models in mathematics take many forms. If the variables are conceived as continuous over time, and the primary theoretical focus is on how variables change as a consequence of changes in other variables, then the most likely mathematical form is differential equations. The substantive theory is translated into statements about rates of change (Doreian and Hummon 1976). Most diffusion and epidemiology models, like the AIDS problem just noted, use differential equations. So do a number of demography models. Recently some of the numerous differential equations models dealing with topics such as conflict and arms control have been applied to theoretical questions of cooperation and competition in social interaction (e.g., Meeker and Leik 1997).

For relatively simple differential equations models, once the model is developed it is possible to determine the trajectory over time of any of the properties of the system to ascertain under what conditions the covariation of system properties will shift or remain stable, and to ask whether that system will tend toward equilibrium or some other theoretical limit, oscillate in regular patterns, or even “explode.”

None of these questions could be asked of a linear model because the mathematics of the linear model leaves nothing about the model itself to be deduced. Only statistical questions can be asked: estimates of the regression coefficients that fit the model to the data and the closeness of that fit. To the extent that sociology addresses questions of process, appropriate theory requires nonlinear models.

What about the assumption of continuous variables? For cross-sectional data, or data from only two or three time points, new techniques of statistical analysis suitable for categorical data have been developed. However, these focus once again

on technical statistical problems of bias in estimation of parameters and generally rely on assumptions of linearity. Within mathematical sociology, other approaches exist. If time or time-related variables were to be treated in discrete units, there are at least three different approaches available. For handling dynamic systems without the calculus of differential equations, difference equations are the appropriate form. Huckfeldt and colleagues (1982) provide a convenient overview of this approach.

For extensive time series with relatively few variables, there are Box-Jenkins and related types of models, although these have seen relatively little use in sociology. Because they are closer to statistical models than general theoretical models, they are only mentioned in passing.

Many theories treat systems as represented by discrete states. For example, over a lifetime, an individual is likely to move into and out of several different occupational statuses. Are certain status transitions more likely than others? In simplest form, the implied theoretical model involves a matrix algebra formulation that specifies the probability of moving from each of the states (occupational statuses for this example) to each of the other states. Then the mathematics of matrix algebra allows deduction of a number of system consequences from this “transition matrix.” Such a treatment is called a *Markov chain*. An early application by Blumen and colleagues (1955) demonstrated that certain modifications of a Markov chain were needed for the theory to fit the data they had available on occupational transitions. Their work was the initial inspiration for a distinguished string of mathematical models of social mobility. Another classic is by White (1970), who conceptualized mobility of vacant positions as well as of individuals into and out of positions.

Another type of substantive problem that deals with discrete data is the analysis of social networks such as friendship structures. These can be modeled using the mathematics of *graph theory*, the basic concepts of which include *nodes* (or *points*) and *relationships between pairs of nodes* (or *lines*). One of the most vigorous modeling areas in sociology, network analysis has produced a rich and elaborate literature addressing a wealth of substantive issues. Typically, network data consists of

whether or not any two cases (nodes in the network) are linked in one or more ways. The resulting data set, then, usually consists of presence or absence of a link of a given type over all pairs of nodes.

Early network analyses concerned friendships, cliques, and rudimentary concepts of structurally based social power. With the introduction of directed graph theory (Harary et al. 1965), random or probabilistic net theory, and block modeling, powerful tools have been developed to approach social structure and its consequences from a network point of view. As those tools emerged, the range of questions addressed via network analysis has greatly expanded. Over the past twenty years, numerous articles in the *Journal of Mathematical Sociology* have dealt with networks. The journal *Social Networks* also publishes work in this area. Overviews and examples can be found in Burt and Minor (1983), and in Wellman and Berkowitz (1988).

Graph theory may be applied to other theoretical issues; for example, a graph-theoretic model developed by Berger and colleagues (1977) helps explain the processes by which people combine information about the various characteristics of themselves and others to form expectations for task performance.

Small group processes have also generated a variety of mathematical formulations. Because observations of groups often generate counts of various types of acts (for example amount of talking by members of a discussion group) that display remarkable empirical regularity, these processes have intrigued model builders since the early 1950's. Recent developments, combining network analyses with Markov chains, include Robinson and Balkwell (1995) and Skvoretz and Fararo (1996).

At the most micro level of sociology, the analysis of individual behavior in social contexts has a long tradition of mathematical models of individual decision making. Recent developments include the satisfaction-balance decision-making models of Gray and Tallman (Gray et al. 1998).

An exciting aspect of these different levels of development is that, increasingly, inquiries into microdynamics based on social exchange theory are working toward formulations compatible with the more general network structural analyses. These

joint developments, therefore, promise a much more powerful linking of micro-system dynamics with macro-system structural modeling (Cook 1987; Willer 1987). Recent work on power as a function of the linkages that define the exchange system is an example. The use of mathematics to express formal definitions has enabled researchers to pinpoint where there are theoretical differences, leading to productive debate (Markovsky et al. 1988; Cook and Yamagishi 1992; Friedkin 1993; Bonacich 1998).

There are other examples of work in mathematical sociology involving attempts to develop appropriate mathematical functions to describe a theoretically important concept. One is Jasso's innovative work on models of distributive justice (e.g., 1999). Similarly, affect control theory (e.g., Smith-Lovin and Heise 1988) represents mathematical formulation in an area (symbolic interaction and sociology of emotions) typically considered not subject to such treatment.

One other area of vigorous development deserves attention; the treatment of strings of events that constitute the history of a particular case, process, or situation (Allison 1984; Tuma and Hannan 1984; Heise 1989). Event history analysis has some of its origins in traditional demographers' life tables, but methods and models have experienced a great deal of attention and growth in recent years. If one had lifetime data on job placements, advancements, demotions, and firings (i.e., employment event histories) for a sample of individuals, then event history methods could be used for examining what contributes to differential risks of one of those events occurring, how long someone is likely to be in a given situation ("waiting time" between events), and so forth.

An important recent development is the use of complex computer simulations for developing and exploring mathematically expressed theories. We have noted above the tension between simple models that are mathematically tractable and more complex models that may be more realistic as for example including feedback loops and random processes. Computer simulation is a way of showing what can be derived from the assumptions of a model without an analytic mathematical solution; examples include Macy and Skvoretz (1998), Carley (1997), and Hanneman (1995).

Like many other areas of social and behavioral science, mathematical sociology has been influenced by developments in game theory, from early work by Rapoport (1960) to the more recent idea of evolutionary games introduced by Axelrod (1984) and extensive interest in problems of collective action (Marwell and Oliver 1993). These consider how actions of individuals may produce unintended outcomes because of the logic of their interdependence with actions of others. The related theoretical area of rational choice theory also has a strong mathematical component (Coleman 1990); see also recent issues of the journal *Rationality and Society*.

A notable feature of current work in mathematical sociology is that the development, testing, and refinement of mathematical models is located within substantive research programs. Mathematical formulations appear in mainstream sociology journals and are becoming accepted as one component of continuing programs of research along with development and refinement of theory and collection and analysis of empirical data (several examples can be found in Berger and Zelditch 1993).

One indication of continuing interest in mathematical sociology is the recent formation of the Mathematical Sociology Section of the American Sociological Association. There is also a large amount of work internationally, including in Japan (Kosaka 1995) and in England and Europe (Hegselmann et al. 1996). Mathematical work in sociology is alive and vigorous. It truly does promise a higher level of theoretical precision and integration across the discipline.

(SEE ALSO: *Paradigms and Models*; *Scientific Explanation*)

REFERENCES

Allison, Paul D. 1984 *Event History Analysis: Regression for Longitudinal Event Data*. Beverly Hills, Calif: Sage.

Axelrod, Robert M. 1984 *The Evolution of Cooperation*. New York: Basic Books.

Berger, Joseph, Bernard P. Cohen, J. Laurie Snell, and Morris Zelditch, Jr. 1962 *Types of Formalization in Small Group Research*. Boston: Houghton Mifflin.

———, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch, Jr. 1977 *Status Characteristics and Social Interaction: An Expectation-States Approach*. New York: Elsevier.

———, and Morris Zelditch, Jr., eds. 1993 *Theoretical Research Programs: Studies in the Growth of Theory*. Stanford, Calif: Stanford University Press.

Blumen, Isadore, Marvin Kogan, and Philip H. McCarthy 1955 *The Industrial Mobility of Labor as a Probability Process*. Ithaca, N.Y.: Cornell University Press.

Bonacich, Phillip 1998 "A Behavioral Foundation for a Structural Theory of Power in Exchange Networks." *Social Psychology Quarterly* 61:185–198.

Burt, Ronald S., and Michael J. Minor 1983 *Applied Network Analysis*. Beverly Hills, Calif.: Sage.

Carley, Kathleen, ed. 1997 *Computational Organizational Theory*. New York: Gordon and Breach (special issue of the *Journal of Mathematical Sociology*).

Coleman, James S. 1964 *Introduction to Mathematical Sociology*. New York: Free Press.

——— 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.

Cook, Karen S., ed. 1987 *Social Exchange Theory*. Newbury Park, Calif.: Sage.

———, and Toshio Yamagishi 1992 "Power in Exchange Networks: a Power-Dependence Formulation." *Social Networks* 14:245–266.

Doreian, Patrick, and Norman P. Hummon 1976 *Modeling Social Processes*. New York: Elsevier.

Fararo, Thomas J. 1973 *Mathematical Sociology*. New York: Wiley.

———, ed. 1984 *Mathematical Ideas and Sociological Theory*. New York: Gordon and Breach (special issue of the *Journal of Mathematical Sociology*).

Friedkin, Noah E. 1993 "An Expected Value Model of Social Exchange Outcomes." Pp. 163–193 in E. J. Lawler, B. Markovsky, K. Heimer, and J. O'Brien, eds., *Advances in Group Processes*, vol. 10. Greenwich, Conn.: JAI Press.

Gray, Louis N., Irving Tallman, Dean H. Judson, and Candan Duran-Aydintug 1998 "Cost-Equalization Applications to Asymmetric Influence Processes." *Social Psychology Quarterly* 61:259–269.

Hage, Jerald, ed. 1994 *Formal Theory in Sociology: Opportunity or Pitfall?* Albany: State University of New York Press.

Hanneman, Robert A., ed. 1995 *Computer Simulations and Sociological Theory*. Greenwich Conn.: JAI (special issue of *Sociological Perspectives*).

Harary, Frank, Robert Z. Norman, and Dorwin Cartwright 1965 *Structural Models*. New York: Wiley.

Hayduk, Leslie A. 1987 *Structural Equation Modeling with LISREL*. Baltimore, Md.: Johns Hopkins University Press.

- Hegselmann, Rainer, Ulrich Meuller, and Klaus G. Troitzsch 1996 *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. Dordrecht, The Netherlands: Kluwer Academic.
- Heise, David R. 1989 "Modeling Event Structures." *Journal of Mathematical Sociology* 14:139-169.
- Huckfeldt, Robert R., C. W. Kohfeld, and Thomas W. Likens 1982 *Dynamic Modeling: An Introduction*. Beverly Hills, Calif.: Sage.
- Jasso, Guillermina 1999 "How Much Injustice Is There in the World? Two New Justice Indexes." *American Sociological Review* 64:133-167.
- Joreskog, Karl G. 1970 "A General Method for Analysis of Covariance Structures." *Biometrika* 57:239-232.
- Kosaka, Kenji, ed. 1995 *Mathematical Sociology in Japan*. New York: Gordon and Breach (special issue of the *Journal of Mathematical Sociology*).
- Leik, Robert K., and Barbara F. Meeker 1975 *Mathematical Sociology*. Englewood Cliffs, N.J.: Prentice-Hall.
- Macy, Michael W., and John Skvoretz 1998 "The Evolution of Trust and Cooperation between Strangers: A Computational Model." *American Sociological Review* 63:638-660.
- Markovsky, Barry, David Willer, and Travis Patton 1988 "Power Relations in Exchange Networks." *American Sociological Review* 53:220-236.
- Marwell, Gerald, and Pamela Oliver 1993 *The Critical Mass in Collective Action: A Micro-Social Theory*. Cambridge, England: Cambridge University Press.
- Meeker, Barbara F., and Robert K. Leik 1997 "Computer Simulation: an Evolving Component of Theoretical Research Programs." Pp. 47-70 in J. Szmatka, J. Skvoretz, and J. Berger, eds., *Status, Network and Structure: Theory Construction and Theory Development*. Palo Alto, Calif.: Stanford University Press.
- Rapoport, Anatol 1960 *Fights, Games, and Debates*. Ann Arbor: University of Michigan Press.
- Robinson, Dawn T., and James W. Balkwell 1995 "Density, Transitivity, and Diffuse Status in Task-Oriented Groups." *Social Psychology Quarterly* 58:241-254.
- Simon, Herbert A. 1957 *Models of Man: Social and Rational*. New York: Wiley.
- Skvoretz, John, and Thomas J. Fararo 1996 "Status and Participation in Task Groups: a Dynamic Network Model." *American Journal of Sociology* 101:1355-1414.
- Smith-Lovin, Lynn, and David R. Heise 1988 *Affect Control Theory: Research Advances*. New York: Academic.
- Tuma, Nancy B., and Michael T. Hannan 1984 *Social Dynamics: Models and Methods*. New York: Academic.
- Wellman, Barry, and S. D. Berkowitz, eds. 1988 *Social Structures: A Network Approach*. New York: Gordon and Breach.
- White, Harrison C. 1970 *Chains of Opportunity; System Models of Mobility in Organizations*. Cambridge, Mass.: Harvard University Press.
- Willer, David 1987 *Theory and the Experimental Investigation of Social Structures*. New York: Gordon and Breach.
- Wright, Sewell 1934 "The Method of Path Coefficients." *Annals of Mathematical Statistics* 5:161-215.
- Zetterberg, Hans 1965 *On Theory and Verification in Sociology*. Totowa, N.J.: Bedminster.

BARBARA F. MEEKER  
ROBERT K. LEIK

## MEASUREMENT

There are many standards that can be used to evaluate the status of a science, and one of the most important is how well variables are measured. The idea of measurement is relatively simple. It is associating numbers with aspects of objects, events, or other entities according to rules, and so measurement has existed for as long as there have been numbers, counting, and concepts of magnitude. In daily living, measurement is encountered in myriad ways. For example, measurement is used in considering time, temperature, distance, and weight. It happens that these concepts and quite a few others are basic to many sciences. The notion of measurement as expressing magnitudes is fundamental, and the observation that *if something exists, it must exist in some quantity* is probably too old to attribute to the proper authority. This notion of quantification is associated with a common dictionary definition of *measurement*: "The extent, capacity, or amount ascertained by measuring."

A concept such as distance may be considered to explore the meaning of measurement. To measure distance, one may turn to a simple example of a straight line drawn between two points on a sheet of paper. There is an origin or beginning point and an end point, and an infinite number of points between the beginning and the end. To measure in a standard way, a unit of distance has to be arbitrarily defined, such as an inch. Then the



distance of any straight line can be observed in inches or fractions of inches. For convenience, arbitrary rules can be established for designating number of inches, such as feet, yards, and miles. If another standard is used—say, meters—the relationship between inches and meters is one in which no information is lost in going from one to the other. So, in summary, in the concept of measurement as considered thus far, there are several properties, two of which should be noted particularly: the use of arbitrary standardized units and the assumption of continuous possible points between any two given points. The case would be similar if time, temperature, or weight were used as an example.

There is another property mentioned above, a beginning point, and the notion of the beginning point has to be examined more carefully. In distance, if one measures 1 inch from a beginning point on a straight line, and then measures to a second point 2 inches, one may say that the distance from the beginning point of the second point is twice that of the first point. With temperature, however, there is a problem. If one measures temperature from the point of freezing using the Celsius scale, which sets 0 degrees at the freezing point of water under specified conditions, then one can observe temperatures of 10 degrees and of 20 degrees. It is now proper to say that the second measurement is twice as many degrees from the origin as the first measure, but one cannot say that it is twice the temperature. The reason for this is that the origin that has been chosen is not the origin that is required to make that kind of mathematical statement. For temperature, the origin is a value known as absolute zero, the absence of any heat, a value that is known only theoretically but has been approximated.

This problem is usually understood easily, but it can be made more simple to understand by illustrating how it operates in measuring distance. Suppose a surveyor is measuring distance along a road from A to B to C to D. A is a long distance from B. Arriving at B, the surveyor measures the distance from B to C and finds it is 10 miles, and then the distance from B to D is found to be 20 miles. The surveyor can say that the distance is twice as many miles from B to D as from B to C, but he cannot say that the distance from A to D is twice the distance from A to C, which is the error one

would make if one used the Celsius temperature scale improperly. Measuring from the absolute origin for the purpose of carrying out mathematical operations has become known as *ratio level measurement*.

The idea of “levels of measurement” has been popularized following the formulation by the psychologist S. S. Stevens (1966). Stevens first identifies scales of measurement much as measurement is defined above, and then notes that the type of scale achieved depends upon the basic empirical operations performed. The operations performed are limited by the concrete procedures and by the “peculiarities of the thing being scaled.” This leads to the types of scales—nominal, ordinal, interval, and ratio—which are characterized “by the kinds of transformations that leave the ‘structure’ of the scale undistorted.” This “sets limits to the kinds of statistical manipulation that can legitimately be applied to the scaled data.”

*Nominal* scales can be of a type like numbering individuals for identification, which creates a class for each individual. Or there can be classes for placement on the basis of equality within each class with regard to some characteristic of the object. *Ordinal* scales arise from the operation of rank ordering. Stevens expressed the opinion that most of the scales used by psychologists are ordinal, which means that there is a determination of whether objects are greater than or less than each other on characteristics of the object, and thus there is an ordering from smallest to largest. This is a crucial point that is examined below. *Interval* scales (equal-interval scales) are of the type discussed above, like temperature and these are subject to linear transformation with invariance. There are some limitations on the mathematical operations that can be carried out, but in general these limitations do not impede use of most statistical and other operations carried out in science. As noted, when the equal-interval scales have an absolute zero, they are called *ratio* scales. A lucid presentation of the issue of invariance of transformations and the limitations of use of mathematical operations (such as addition, subtraction, multiplication, and division) on interval scales is readily available in Nunnally (1978).

What is important to emphasize is that how the scales are constructed, *as well as how the scales*

are used, determines the level of measurement. With regard to ordinal scales, Nunnally makes a concise and precise statement that should be read carefully: "With ordinal scales, none of the fundamental operations of algebra may be applied. In the use of descriptive statistics, it makes no sense to add, subtract, divide, or multiply ranks. Since an ordinal scale is defined entirely in terms of inequalities, only the algebra of inequalities can be used to analyze measures made on such scales" (1978, p. 22). What this means is that if one carries out a set of operations that are described as making an ordinal scale, the moment one adds, subtracts, divides, or multiplies the ranks, one has treated the scale as a particular type of interval scale. Most commonly, the type of scale that is de facto created when ordinal data are subject to ordinary procedures like addition, subtraction, division, and/or multiplication is through the assumption that the difference between ranks are equal, leading to sets like 1, 2, 3, 4, 5, 6, and thus to the treatment for ties such as 1, 2, 4, 4, 4, 6. This is sometimes called a flat distribution with an interval of one unit between each pair of ordered cases. Effectively, this is the same kind of distribution in principle as the use of ordered categories, such as quartiles, deciles, or percentiles, but it is a more restrictively defined distribution of one case per category. To repeat, for emphasis: The use of addition, subtraction, division, and/or multiplication with ordinal data automatically requires assumptions of intervals, and one is no longer at the level of ordinal analysis. *Thus, virtually all statistical procedures based on collected ordered or rank data actually assume a special form of interval data.*

### ISSUES ON LEVEL OF MEASUREMENT

A number of issues are associated with the notion of levels of measurement. For example, are all types of measurement included in the concepts of nominal, ordinal, interval, and ratio? What is the impact of using particular statistical procedures when data are not in the form of well-measured interval scales? What kind of measurement appears (epistemologically) appropriate for the social and behavioral sciences?

The last question should probably be examined first. For example, are measures made about attributes of persons nominal, ordinal, or interval?

In general, we cannot think of meaningful variables unless they *at least* imply order, but is order all that one thinks of when one thinks about characteristics of persons? For example, if one thinks of heights, say of all males of a given age, such as 25, does measurement imply ordering them on an interval scale? We know that height is a measure of distance, so we assume the way one should measure this is by using a standard. For purposes of the example here, an interval scale is proposed, and the construction is as follows. The shortest, 25 year-old male (the category defined as 25 years and 0 days to 25 years and 365 days of age) and the tallest are identified. The two persons are placed back to back, front to front, and every other possible way, and the distance between the height of the shortest and the tallest is estimated on a stick. Many estimates are made on the stick, until the spots where agreement begins to show discretely are evident; and so, with whatever error occurs in the process, the locations of beginning and end are indicated on the stick. Now the distance between the beginning and the end is divided into equal intervals, and thus an interval scale has been created. On this scale it is possible to measure every other male who is 25 years old, and the measure can be stated in terms of the number of intervals taller than the shortest person. Note that all possible values can be anticipated, and this is a continuous distribution.

Now if a million persons were so measured, how would they be distributed? Here the answer is on the basis of naive experience, as follows. First, there would be very few people who would be nearly as short as the shortest or as tall as the tallest. Where would one expect to find most persons? In the middle of the distance, or at some place not too far from it. Where would the next greatest number of persons be found? Close to the biggest. With questions of this sort one ends up describing a well-distributed curve, possibly a normal curve or something near it.

It is proper now to make a small diversion before going on with answering the questions about the issues associated with level of measurement. In particular, it should be noted that there are many sources of error in the measurement that has just been described. First, of course, the age variable is specified with limited accuracy. At the limits, it may be difficult to determine exact age

because of the way data are recorded. There are differences implied by the fact that where one is born makes a difference in time, and so on. This may seem facetious, but it illustrates how easily sources of error are bypassed without examination. Then it was noted that there were different estimates of the right location for the point of the shortest and the tallest person as marked on the stick. This is an error of observation and recording, and clearly the points selected are taken as mean values. Who are the persons doing the measuring? Does it make a difference if the person measuring is short or tall? These kinds of errors will exist for all persons measured. Further, it was not specified under what conditions the measurements taken or were to be taken. Are the persons barefoot? How are they asked to stand? Are they asked to relax to a normal position or to try to stretch upward? What time of day is used, since the amount of time after getting up from sleep may have an influence? Is the measurement before or after a meal? And so forth.

The point is that there are many sources of error in taking measures, even direct measures of this sort, and one must be alert to the consequences of these errors on what one does with the data collected. Errors of observation are common, and one aspect of this is the limit of the discriminations an observer can make. One type of error that is usually built into the measurement is rounding error, which is based on the estimated need for accuracy. So, for example, heights are rarely measured more accurately than to the half-inch or centimeter, depending on the standard used. There is still the error of classification up or down, by whatever rule is used for rounding, at the decision point between the intervals used for rounding. Rounding usually follows a consistent arbitrary rule, such as "half adjusting," which means keeping the digit value if the next value in the number is 0 to 4 (e.g., 24.456 = 24) or increasing the value of a digit by 1 if the next value is 5 to 9 (e.g., 24.789 = 25). Another common rounding rule is simply to drop numbers (e.g., 24.456 = 24 and 24.789 = 24). It is important to be aware of which rounding rule is being used and what impact it may have on conclusions drawn when the data collected are used.

The use of distribution-free statistics (often called nonparametric statistics) was popularized beginning in the mid-1950s, and quickly came to

be erroneously associated with the notion that most of the measurement in the social and behavioral sciences is of an ordinal nature. Actually, the use of the distribution-free statistics was given impetus because some tests, such as the sign test, did not require use of all the information available to do a statistical test of significance of differences. Thus, instead of using a test of differences of means, one could quickly convert the data to plus and minus scores, using some arbitrary rule, and do a "quick-and-dirty" sign test. Then, if one found significant differences, the more refined test could be carried out at one's leisure. Some of the orientation was related to computing time available, which meant time at a mechanical calculator. Similarly, it was well known that if one used a Spearman rank correlation with larger samples, and if one were interested in measuring statistical significance, one would have to make the same assumptions as for the Pearson product moment correlation, but with less efficiency.

However, this early observation about distribution-free statistics suggests that measures can be thought of in another way. Namely, one can think of measures in terms of how much they are degraded (or imperfect) interval measures. This leads to two questions that are proper to consider. First, what kind of measure is implied as appropriate by the concept? And second, how much error is there in how the measure is constructed if one wants to use procedures that imply interval measurement, including addition, subtraction, multiplication, and division?

What kind of measure is implied by the concept? One way of answering this is to go through the following procedure. As an example, consider a personal attribute, such as aggressiveness. Is it possible to conceive of the existence of a least aggressive person and a most aggressive person? Obviously, whether or not such persons can be located, they can be conceived of. Then, is there any reason to think that persons cannot have any and all possible quantities of aggressiveness between the least and the most aggressive persons? Of course not. Thus, what has been described is a continuous distribution, and with the application of a standard unit, it is appropriately an interval scale. *It is improper to think of this variable as intrinsically one that is ordinal because it is continuous. In fact, it is difficult to think of even plausible examples of*

*variables that are intrinsically ordinal.* As Kendall puts it, “the essence of ranking is that the objects shall be orderable, and the totality of values of a continuous variate cannot be ordered in this sense. They can be regarded as constituting a range of values, but between any two different values there is always another value, so that we cannot number them as would be required for ranking purposes” (1948, p. 105). While a few comments were published that attempted to clarify these issues of measurement in the 1960s (Borgatta 1968), most methodologists accepted the mystique of ordinal measurement uncritically.

Often measures of a concept tend to be simple questions with ordered response categories. These do not correspond to ordinal measures in the sense of ordering persons or objects into ranks, but the responses to such questions have been asserted to be ordinal level measurement because of the lack of information about the intervals. So, for example, suppose one is attempting to measure aggressiveness using a question such as “When you are in a group, how much of the time do you try to get your way about what kinds of activities the group should do next?” Answer categories are “never,” “rarely,” “sometimes,” “often,” “very often,” and “always.” Why don’t these categories form an interval scale? The incorrect answer usually given is “because if one assumes an interval scale, one doesn’t know where the answer categories intersect the interval scale.” However, this *does not* create an ordinal scale. It creates an interval scale with unknown error with regard to the spacing of the intervals created by the categories.

Thus, attention is now focused on the second question: How much error is involved in creating interval scales? This question can be answered in several ways. A positive way of answering is by asking how much difference it makes to distort an interval scale. For example, if normally distributed variables (which are assumed as the basis for statistical inference) are transformed to flat distributions, such as percentiles, how much impact does this have on statistical operations that are carried out? The answer is “very little.” This property of not affecting results of statistical operations has been called robustness. Suppose a more gross set of transformations is carried out, such as deciles. How much impact does this have on statistical operations? The answer is “not much.” However,

when the transformations are to even grosser categories, such as quintiles, quartiles, thirds, or halves, the answer is that because one is throwing away even more information by grouping into fewer categories, the impact is progressively greater. What has been suggested in this example has involved two aspects: the transformation of the shape of the distribution, and the loss of discrimination (or information) by use of progressively fewer categories.

### CONSEQUENCES OF USING LESS THAN NORMALLY DISTRIBUTED VARIABLES

If one has normally distributed variables, the distribution can be divided into categories. The interval units usually of interest with normally distributed variables are technically identified as standard deviation units, but other units can be used. When normally distributed variables are reduced to a small number of (gross) categories, substantial loss of discrimination or information occurs. This can be illustrated by doing a systematic exercise, the results of which are reported in Table 1. The data that are used for the exercise are generated from theoretical distributions of random normal variables with a mean of 0 and a standard deviation of 1. In the exercise, one aspect is examining the relationship among normally distributed variables, but the major part of the exercise involves the data in ordered categorical form, much as it is encountered in “real” data. The exercise permits reviewing several aspects associated with knowledge about one’s measurement.

As noted, the exercise is based on unit normal variables (mean = 0, standard deviation = 1) that are sampled and thus are subject to the errors of random sampling that one encounters with “real” data. The theoretical underlying model is one that is recommended in practice, specified as follows: (1) A criterion variable is to be predicted, that is, the relationships of some independent (predictor) variables are to be assessed with regard to the criterion variable.

(2.) For each sample in which the relationship of the independent variables is assessed, the theoretical underlying relationship of the independent variables is specified, and thus for the purposes of the exercise is known. For the exercise, four levels of underlying theoretical relationship are product

**Median Correlation of Scores with Criterion Variable, Range of 9 Samples (N=150) of Unit Normal Deviates, Additive Scores Based on Four Items with Theoretical Correlations of .8, .6, .4, and .2 with the Criterion Variable in the Population**

	XO	XOR	XOA	XOAR	XOB	XOBR	XO <sup>2</sup>	XOA <sup>2</sup>	XOB <sup>2</sup>
SumX.8	.94	.93-.96	.75	.69-.78	.66	.52-.73	.88	.56	.44
SumA.8	.92	.90-.94	.76	.74-.80	.62	.56-.67	.85	.58	.38
SumB.8	.89	.87-.92	.68	.63-.74	.62	.58-.68	.78	.46	.38
SumC.8	.85	.81-.86	.78	.69-.84	.51	.46-.59	.72	.61	.36
SumDL.8	.72	.64-.74	.54	.48-.60	.72	.64-.77	.52	.29	.52
SumDR.8	.71	.64-.76	.55	.49-.58	.26	.23-.30	.50	.30	.07
SumX.6	.84	.78-.88	.65	.56-.75	.57	.49-.71	.71	.42	.32
SumA.6	.82	.77-.85	.66	.60-.71	.58	.48-.62	.67	.44	.34
SumB.6	.79	.70-.81	.62	.55-.65	.51	.43-.60	.62	.38	.26
SumC.6	.74	.70-.79	.65	.56-.70	.48	.40-.54	.55	.42	.23
SumDL.6	.62	.56-.69	.53	.45-.58	.56	.45-.67	.38	.28	.31
SumDR.6	.62	.57-.67	.52	.45-.56	.32	.23-.36	.38	.27	.10
SumX.4	.65	.59-.71	.53	.44-.57	.48	.31-.54	.42	.28	.23
SumA.4	.66	.51-.72	.52	.47-.60	.42	.27-.55	.44	.27	.18
SumB.4	.61	.48-.67	.47	.37-.55	.41	.33-.46	.37	.22	.17
SumC.4	.59	.49-.66	.48	.39-.58	.38	.27-.43	.35	.23	.14
SumDL.4	.45	.32-.52	.39	.30-.47	.33	.15-.51	.20	.15	.11
SumDR.4	.50	.45-.59	.43	.38-.49	.26	.22-.27	.25	.18	.07
SumX.2	.39	.25-.51	.31	.20-.41	.25	.12-.45	.15	.10	.06
SumA.2	.31	.26-.42	.23	.16-.34	.27	.22-.31	.10	.05	.07
SumB.2	.32	.20-.49	.29	.12-.29	.23	.10-.29	.10	.08	.05
SumC.2	.28	.23-.42	.22	.15-.40	.17	.10-.24	.08	.05	.03
SumDL.2	.29	.17-.35	.16	.12-.24	.22	.16-.35	.08	.03	.05
SumDR.2	.24	.19-.37	.20	.04-.31	.17	.07-.30	.06	.04	.03

Table 1

moment correlation coefficients of .8, .6, .4, and .2, between the predictor variables and the criterion variable. This represents different levels of relationship corresponding to magnitudes commonly encountered with “real” social and behavioral data.

(3.) For each sample that is used in the exercise, totally independent distributions are drawn from the theoretical distributions. That is, each sample in the exercise is created independently of all other samples, and in each case the criterion and independent predictor variables are drawn independently within the definition of the theoretical distributions.

(4.) Corresponding to a common model of prediction, assume that the independent variables are of a type that can be used to create scores. Here we will follow a common rule of thumb and use four independent variables to create a *simple additive score*; that is, the values of the four independent variables are simply added together. There are many ways to create scores, but this procedure has

many virtues, including simplicity, and it permits examining some of the consequences of using scores to measure a concept.

(5.) The independent variables are modified to correspond to grouped or categorical data. This is done before adding the variables together to create scores, as this would be the way variables encountered in research with “real” data are usually defined and used. The grouped or categorical independent variables are modified in the following ways:

A. Four groups or categories are created by using three dividing points in the theoretical distribution, -1 standard deviation, the mean of 0, and +1 standard deviation. The values for four variables, now grouped or categorical data, are added to give the score used in the correlation with the criterion variable. Thus, looking at Table 1, the row SumA.8 involves samples in which four independent variables based on data having four categories as defined above, and

having an underlying theoretical correlation coefficient for each independent variable to the criterion variable of .8, have been used to create a score. The variables defined by these cutting points correspond to a notion of variables with response categories, say, such as “agree,” “probably agree,” “probably disagree,” and “disagree,” and as it happens the responses theoretically are well distributed, roughly 16 percent, 34 percent, 34 percent, and 16 percent.

B. Three groups or categories are created by using two dividing points in the theoretical distribution, -1 and +1, corresponding to a notion of variables with response categories, say, such as “agree,” “don’t know or neutral,” and “disagree,” and these are again theoretically well distributed, but in this case the center category is large (about 68 percent of responses). The scores for these samples are identified as SumB.8 for the case of the underlying theoretical distribution of the independent variables having a correlation coefficient with the criterion variable of .8.

C. Two groups or categories are created by using one dividing point at the mean of the theoretical distribution, and this corresponds to variables with response categories such as “agree” and “disagree,” or “yes” and “no.” Division of a variable at the midpoint of the distribution is usually considered statistically to be the most efficient for dichotomous response categories.

DL. Two groups or categories may be created that are not symmetrical, unlike the case of C above. Here we create two sets of samples, one identified by Sum DL with the cutting point at -1 standard deviation, or the left side of the distribution as it is usually represented. Presumably, when variables are not well distributed, more information is lost in the data collection if the skew is due to the way the question is formulated.

DR. Two groups or categories are created with the skew on the right side of the distribution at +1 standard deviation. SumDR scores are created in parallel to the SumDL scores.

Finally, the rows of SumX are those where scores are computed and no conversion of the variables has been carried out, so the independent variables are drawn from the theoretically formulated unit normal distributions at the level of correlation coefficient between the independent variables and the criterion variable.

In summary, for each sample the predictor score is an additive score based on four variables, each related to the criterion variable in a theoretical distribution at a given level, with the examples in the exercise the levels chosen as product moment correlation coefficients of .8, .6, .4, and .2.

The rationale for the choice of four variables in each score is drawn arbitrarily from the theoretical distribution of reliability coefficients, which, for variables of equal reliability, is a curve of diminishing return with each increment of the number of variables. At four variables a score is usually considered to be at an efficient point, balancing such things as time and effort needed to collect data, independence of variable definitions, and the amount of improvement of reliability expected. Using more variables in the “real” world usually involves using less “good” items progressively from the point of view of content, as well as the diminishing efficiency even if they were equally as good as the first four.

With regard to the criterion variable, the dependent variable to be predicted, we have generated three as follows: First, the variable is drawn from a theoretical unit normal distribution without modification, and this involves correlation coefficients reported in column XO. Second, the dependent variable is drawn from the theoretical distribution dichotomized at the mean of 0, creating a symmetric dichotomous variable (two categories), and this involves correlation coefficients reported in column XOA. Third, the variable is drawn from the theoretical distribution dichotomized at -1 standard deviation, creating an asymmetrical dichotomous variable and this involves correlation coefficients reported in column XOB.

As noted above, we are dealing with generated data from theoretical distributions, so all that is done is subject to sampling error, which permits in a limited way for variation of results to be shown. This is useful in illustrating the variation that occurs simply by random processes in the study of variables. For the exercise there were 648 samples

of 150 cases each. The choice of 150 cases was selected for several reasons, convenience in the data set generated being one, but also because 150 cases was, as a rule of thumb, a reasonable number of cases proposed for doing research at an earlier time in the discipline when a Guttman scale was anticipated, and too much “shrinkage” in the subsequent use of the Guttman scale was to be avoided. Shrinkage referred to the experience of finding out that the scale in a subsequent use did not have the same characteristics as in the first (generating) use. The samples are grouped in the six procedures, as noted above, of the full normal data and of five created sets of categories, each with nine repetitions (i.e., samples) carried out to permit examination of distribution of results for each theoretical level of correlation coefficient and for each of the three criterion variable definitions.

A sense of the stability of the relationships involved is provided by using many samples, and *the values that are in columns XO, XOA, and XOB are the median product moment correlation coefficients between a Sum score and a criterion variable.* In the table there are two other sets of columns. XOR, XOAR, and XOBR are the actual range of values for the product moment correlation coefficients between a Sum score and the criterion variable for nine independently drawn samples of 150 cases each. The additional (last three columns) at the right of the table are the squared values of XO, XOA, and XOB, and these values represent the amount of variance that is accounted for by the correlation coefficient.

The table is somewhat complex because it includes a great deal of information. Here we will only review a number of relatively obvious points, but with emphasis. First, note the relationship between SumX.8 and XO, which is .94. Recall that the theoretical correlation coefficient between each independent predictor variable and the criterion variable was defined as .8 in the exercise. Thus, we illustrate rather dramatically the importance of building scores rather than depending on a single predictor variable. The actual improvement in prediction is from a theoretical value of 64 percent of the variance being accounted for by a single predictor variable to the median value of 88 percent of the variance in the median sample of the exercise, an extremely impressive improvement! Examining the correlation coefficient between SumX.6, SumX.4, and SumX.2, it is seen that the

advantage of using scores rather than single items is actually even more dramatic when the relationship between the criterion and the predictor variables is weaker.

Now, examine the relationships between all the Sum variables and XO. It is noted that having the SumA.8 as a well distributed categorical variable lowers the correlation to XO somewhat (.92) compared to the full normal SumX.8 (.94), and, indeed, with SumB.8 and SumC.8 the correlation coefficients are still lower. This illustrates that the more information that is available in the categorical data, the more variance can be predicted for a criterion variable. Note further, that for SumDL.8 and SumDR.8, which are based on dichotomous but skewed variables, the correlation coefficients are still lower. Note that the results for the weaker variables SumX.6 to SumDR.6, for SumX.4 to SumDR.4, and for SumX.2 to SumDR.2 are in general roughly parallel to the results noted, but some irregularity is noted when the relationship between the criterion and the predictor variables is weaker.

The results we have been examining are the median values of nine samples in each case. It is appropriate to look at column XOR and to note that there is quite a bit of variation in the correlation coefficients based on the nine samples of size 150 cases each.

One additional finding needs to be pointed out and emphasized. Compare the correlation coefficients of SumDT.8 and XOB, and of SumDR.8 and XOB. In the former case, the criterion variable and predictor variables are both asymmetric but matched (at the same cutting point), and the median correlation coefficient value is .72. In the latter case they are asymmetric but unmatched, and the median correlation coefficient is .26. This indicates the need to know how cutting points (difficulty levels) of variables can affect the values of correlation coefficients that are encountered.

The reader is urged to examine this table in detail, and it may be useful to keep the table as a guideline for interpreting results. How? If the researcher finds as a result a correlation coefficient of a given size, where does it fit on this table, taking into consideration the characteristics of the criterion variable and the scores based on the predictor variables? Doing this type of examination should help the researcher understand the

meaning of a finding. However, this is but one approach to the issues of measurement. The articles in this encyclopedia that are vital for researchers to consider are those on *Reliability*, *Validity*, and *Quasi-Experimental Research Design*. Additionally, measurement issues are discussed in a variety of types of texts, ranging from specialized texts to books on research methods to general statistics books, such as the following: Agresti and Finaly 1997; Babbie 1995; De Vellis 1991; Knoke and Bohrnstedt 1994; Lewis-Beck 1994; Neuman, 1997; and Traub 1994.

### OTHER MEASURES

The consideration of measurement thus far has concentrated on interval measurement, with some emphasis on how it can be degraded. There are many other issues that are appropriately considered, including the fact that the concepts advanced by Stevens do not include all types of measures. As the discussion proceeds to some of these, attention should also be given to the notion of nominal scales. Nominal scales can be constructed in many ways, and only a few will be noted here. By way of example, it is possible to create a classification of something being present for objects that is then given a label A. Sometimes a second category is not defined, and then the second category is the default, the thing not being present. Or two categories can be defined, such as male and female. Note that the latter example can also be defined as male and not male, or as female and not female. More complex classifications are illustrated by geographical regions, such as north, west, east, and south, which are arbitrary and follow the pattern of compass directions. Such classifications, to be more meaningful, are quickly refined to reflect more homogeneity in the categories, and sets of categories develop such as Northeast, Middle Atlantic, South, North Central, Southwest, Mountain, Northwest, and Pacific, presumably with the intention of being inclusive (exhaustive) of the total area. These are complex categories that differ with regard to many variables, and so they are not easily ordered.

However, each such set of categories for a nominal scale can be reduced to dichotomies, such as South versus “not South”; these variables are commonly called “dummy variables.” This

permits analysis of the dummy variable as though it represented a well-distributed variable. In this case, for example, one could think of the arbitrary underlying variable as being “southernness,” or whatever underlies the conceptualization of the “South” as being different from the rest of the regions. Similarly, returning to the male versus female variable, the researcher has to consider interpretatively what the variable is supposed to represent. Is it distinctly supposed to be a measure of the two biological categories, or is it supposed to represent the social and cultural distinction that underlies them?

Many, if not most, of the variables that are of interest to social and behavioral science are drawn from the common language, and when these are used analytically, many problems or ambiguities become evident. For example, the use of counts is common in demography, and many of the measures that are familiar are accepted with ease. However, as common a concept as city size is not without problems. A city is a legal definition, and so what is a city in one case may be quite different from a city in another case. For example, some cities are only central locations surrounded by many satellite urban centers and suburbs that are also defined as cities, while other cities may be made up of a major central location and many other satellite urban centers and suburbs. To clarify such circumstances, the demographers may develop other concepts, like Standard Metropolitan Areas (SMAs), but this does not solve the problem completely; some SMAs may be isolated, and others may be contiguous. And when is a city a city? Is the definition one that begins with the smallest city with a population of 2,500 (not even a geographical characteristic), or 10,000 population, or 25,000 population? Is city size really a concept that is to be measured by population numbers or area, or by some concept of degree of urban centralization? Is New York City really one city or several? Or is New York City only part of one city that includes the urban complexes around it? The point that is critical is that definitions have to be fixed in arbitrary ways when concepts are drawn practically from the common language, and social concepts are not necessarily parsimoniously defined by some ideal rules of formulating scientific theory. Pragmatic considerations frequently intervene in how data are collected and what data become available for use.



An additional point is appropriate here: that in demography and in other substantive areas, important measures include *counts* of discrete entities, and these types of measures do not easily fit the Stevens classification of levels of measurement. A discussion of several technical proposals for more exhaustive classifications of types of measures is considered by Duncan (1984).

**CONSTRUCTING MEASURES**

There are obviously many ways that measures can be constructed. Some have been formalized and diffused, such as Louis Guttman’s cumulative scale analysis, so popular that it has come to be known universally as Guttman scaling, a methodological contribution that was associated with a sociologist and had an appeal for many. An early comprehensive coverage of Guttman scaling can be found in Riley and colleagues (1954). The essence of Guttman scaling is that if a series of dichotomous items is assumed to be drawn from the same universe of content, and if they differ in difficulty, then they can be ordered so that they define scale types. For example, if one examines height, questions could be the following: (1) Are you at least five feet tall? (2) Are you at least five and a half feet tall? (3) Are you at least six feet tall? (4) Are you at least six and a half feet tall? Responses to these would logically fall into the following types: a + to indicate yes and a - to indicate a no:

1	2	3	4
+	+	+	+
+	+	+	-
+	+	-	-
+	-	-	-
-	-	-	-

The types represent “perfect” types, that is, responses made without a logical error. The assumption is that within types, people are equivalent. In the actual application of the procedure, some problems are evident, possibly the most obvious being that there are errors because in applying the procedure in studies, content is not as well specified as being in a “universe” as is the example of height; thus there are errors, and

therefore error types. The error types were considered of two kinds: unambiguous, such as - + + +, which in the example above would simply be illogical, a mistake, and could logically be classed as + + + + with “minimum error.” The second kind is ambiguous, such as + + - +, which with one (minimum) error could be placed in either type + + - - or type + + + +.

Experience with Guttman scaling revealed a number of problems. First, few scales that appeared “good” could be constructed with more than four or five items because the amount of error with more items would be large. Second, the error would tend to be concentrated in the ambiguous error type. Third, scales constructed on a particular study, especially with common sample sizes of about one hundred cases, would not be as “good” in other studies. There was “shrinkage,” or more error, particularly for the more extreme items. The issue of what to do with the placement of ambiguous items was suggested by an alternative analysis (Borgatta and Hays 1952): that the type + + - + was not best placed by minimum error, but should be included with the type + + + - between the two minimum error locations. The reason for this may be grasped intuitively by noting that when two items are close to each other in proportion of positive responses, they are effectively interchangeable, and they are involved in the creation of the ambiguous error type. The common error is for respondents who are at the threshold of decision as to whether to answer positively or negatively, and they are most likely to make errors that create the ambiguous error types.

These observations about Guttman scaling lead to some obvious conclusions. First, the scaling model is actually contrary to common experience, as people are not classed in ordered types in general but presumably are infinitely differentiated even within a type. Second, the model is not productive of highly discriminating classes. Third, and this is possibly the pragmatic reason for doubting the utility of Guttman scaling, if the most appropriate place for locating nonscale or error types of the common ambiguous type is not by minimum error but between the two minimum error types, this is effectively the same as adding the number of positive responses, essentially reducing the procedure to a simple additive score. The remaining virtue of Guttman scaling in the logical placement of unambiguous errors must be

balanced against other limitations, such as the requirement that items must be dichotomous, when much more information can be gotten with more detailed categories of response, usually with trivial additional cost in data collection time.

In contrast with Guttman scaling, simple addition of items into sum scores, carried out with an understanding of what is required for good measurement, is probably the most defensible and useful tool. For example, if something is to be measured, and there appear to be a number of relatively independent questions that can be used to ascertain the content, then those questions should be used to develop reliable measures. *Reliability* is measured in many ways, but consistency is the meaning usually intended, particularly internal consistency of the component items, that is, high intercorrelation among the items in a measure. Items can ask for dichotomous answers, but people can make more refined discriminations than simple yeses and nos, so use of multiple (ordered) categories of response increases the efficiency of items.

The question of whether the language as used has sufficient consistency to make refined quantitative discriminations does not appear to have been studied extensively, so a small data collection was carried out to provide the following example. People were asked to evaluate a set of categories with the question "How often does this happen?" The instructions stated: "Put a vertical intersection where you think each category fits on the continuum, and then place the number under it. Categories 1 and 11 are fixed at the extremes for this example. If two categories have the same place, put the numbers one on top of the other. If the categories are out of order, put them in the order you think correct." A continuum was then provided with sixty-six spaces and the external positions of the first and the last indicated as the positions of (1) always and (11) never. The respondents were asked to locate the following remaining nine categories: (2) almost always (3) very often; (4) often; (5) somewhat often; (6) sometimes; (7) seldom; (8) very seldom; (9) hardly ever; and (10) almost never. It is not surprising that average responses on the continuum are well distributed, with percent locations respectively as 9, 15, 27, 36, 48, 65, 75, 86, and 93; the largest standard deviation for placement location is about 11 percent. Exercises with alternative quantitatively oriented

questions and use of a series of six categories from "definitely agree" to "definitely disagree" provide similar evidence of consistency of meaning. In research, fewer than the eleven categories illustrated here usually used, making the task of discrimination easier and faster for respondents. The point of emphasis is that questions can be designed to efficiently provide more information than simple dichotomous answers and thus facilitate construction of reliable scores.

### MEASUREMENT IN THE REAL WORLD

Many variations exist on how to collect information in order to build effective measurement instruments. Similarly, there are alternatives on how to build the measuring instruments. Often practical considerations must be taken into account, such as the amount of time available for interviews, restraints placed on what kinds of content can be requested, lack of privacy when collecting the information, and other circumstances.

With the progressive technology of computers and word processors, the reduced dependence of researchers on assistants, clerks, and secretaries has greatly facilitated research data handling and analysis. Some changes, like Computer Assisted Telephone Interviewing (CATI) may be seen as assisting data collection, but in general the data collection aspects of research are still those that require most careful attention and supervision. The design of research, however, still is often an ad hoc procedure with regard to the definition of variables. Variables are often created under the primitive assumption that all one needs to do is say, "The way I am going to measure XXX is by responses to the following question." This is a procedure of dubious worth, since building knowledge about the measurement characteristics of the variables to be used should be in advance of the research, and is essential to the interpretation of findings.

A comment that is commonly encountered is that attempting to be "scientific" and developing a strict design for research with well-developed measures forecloses the possibility of getting a broad picture of what is being observed. The argument is then advanced that attempting to observe and accumulate data in systematic research is not as revealing as observing more informally (qualitatively) and "getting a feel" for what is going on.

Further, when systematic research is carried out, so goes the argument, only limited variables can be assessed instead of “getting the complete picture.” This, of course, is the ultimate self-delusion and can be answered directly. If positive findings for the theory do not result from more rigorous, well-designed research, then the speculative generalizations of more casual observation are never going to be any more than that, and giving them worth may be equivalent to creating fictions to substitute for reality. *The fact that attempted systematic empirical research has not produced useful findings does not mean that more intuitive or qualitative approaches are more appropriate.* What it means is that the theory may not be appropriate, or the design of the research may be less than adequate.

Further, this does not mean that there is anything wrong with informal or qualitative research. What it does mean is that there is a priority order in the accumulation of knowledge that says that the informal and qualitative stages may be appropriate to produce theory, which is defined as speculation, about what is being observed, and this may then be tested in more rigorous research. This is the common order of things in the accumulation of scientific knowledge.

If a sociological theory has developed, then it must be stated with a clear specification of the variables involved. One cannot produce a volume on the concept of anomie, for example, and then use the word “anomie” to mean twenty different things. The concept on which one focuses *must* be stated with a clear specification of one meaning, and there are two elements that go into such a definition. The first is to indicate how the concept is to be measured. The second is more commonly neglected, and that is to specify how the concept is differentiated from other concepts, particularly those that are closely related to it in meaning.

The development of well-measured variables in sociology and the social sciences is essential to the advancement of knowledge. Knowledge about how good measurement can be carried out has advanced, particularly in the post-World War II period, but it has not diffused and become sufficiently commonplace in the social science disciplines.

It is difficult to comprehend how substituting no measurement or poor measurement for the best measurement that sociologists can devise can

produce better or more accurate knowledge. Examples of the untenable position have possibly decreased over time, but they still occur. Note for example: “Focus on quantitative methods rewards *reliable* (i.e., repeatable) methods. Reliability is a valuable asset, but it is only one facet of the value of the study. In most studies, reliability is purchased at the price of lessened attention to theory, validity, relevance. etc.” (Scheff 1991). Quite the contrary, concern with measurement and quantification is concern with theory, validity, and relevance!

Finally, it is worth emphasizing two rules of thumb for sociologists concerned with research, whether they are at the point of designing research or interpreting the findings of a research that has been reported. First, check on how the variables are specified and ask whether they are measured well. This requires that specific questions be answered: Are the variables reliable? How does one know they are reliable? Second, are the variables valid? That is, do they measure what they are supposed to measure? How does one know they do? If these questions are not answered satisfactorily, then one is dealing with research and knowledge of dubious value.

(SEE ALSO: *Levels of Analysis: Nonparametric Statistics; Reliability, Validity; Quasi-Experimental Research Design*)

## REFERENCES

- Agresti, Alan, and Barbara Finaly 1997 *Statistical Methods for Social Sciences*. Englewood Cliffs, N. J.: Prentice Hall.
- Babbie, Earl 1995 *The Practice of Social Research*. Belmont, Calif.: Wadsworth.
- Borgatta, Edgar F. 1968 “My Student, the Purist: A Lament.” *Sociological Quarterly* 9:29–34.
- , and David G. Hays 1952 “Some Limitations on the Arbitrary Classifications of Non-Scale Response Patterns in a Guttman Scale.” *Public Opinion Quarterly* 16:273–291.
- De Vellis, Robert F. 1991 *Scale Development*. Newbury Park, Calif.: Sage.
- Duncan, Otis Dudley 1984 *Notes of Social Measurement*. New York: Russell Sage Foundation.
- Herzog, Thomas 1997 *Research Methods and Data Analysis in the Social Sciences*. Englewood Cliffs, N.J.: Prentice Hall.

- Kendall, Maurice G. 1948 *Rank Correlation Methods*. London: Griffin.
- Knoke, David, and George W. Bohrnstedt 1994. *Statistics for Social Data Analysis*. Itasca, Ill.: Peacock.
- Lewis-Beck, Michael, ed. 1994 *Basic Measurement*. Beverly Hills, Calif.: Sage.
- Neuman, Lawrence W. 1997 *Social Research Methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon.
- Nunnally, Jum C. 1978 *Psychometric Theory*. New York: McGraw-Hill.
- Riley, Matilda White, John W. Riley, Jr., and Jackson Toby 1954 *Sociological Studies in Scale Analysis*. New Brunswick, N.J.: Rutgers University Press.
- Scheff, Thomas J. 1991 "Is There a Bias in ASR Article Selection." *Footnotes* 19(2, February):5.
- Stevens, S. S., ed. 1966 *Handbook of Experimental Psychology*. New York: John Wiley.
- Traub, Ross E. 1994 *Reliability for the Social Sciences*. Beverly Hills, Calif.: Sage.

EDGAR F. BORGATTA  
YOSHINORI KAMO

## MEASUREMENT INSTRUMENTS

See Factor Analysis; Measurement; Quasi-Experimental Research Designs; Reliability; Survey Research; Validity.

## MEASURES OF ASSOCIATION

Long before there were statisticians, folk knowledge was commonly based on statistical associations. When an association was recognized between stomach distress and eating a certain type of berry, that berry was labeled as poisonous and avoided. For millennia, farmers the world over have observed an association between drought and a diminished crop yield. The association between pregnancy and sexual intercourse apparently was not immediately obvious, not simply because of the lag between the two events, but also because the association is far from perfect—that is, pregnancy does not always follow intercourse. Folk knowledge has also been laced with superstitions, commonly based on erroneously believed statistical associations. For example, people have

believed that there is an association between breaking a mirror and a long stretch of bad luck, and in many cultures people have believed that there is an association between certain ritual incantations and benevolent intervention by the gods.

Scholarly discussions sometimes focus on whether a given association is actually true or erroneously believed to be true. Is there an association between gender and mathematical ability? Between harsh punishment and a low incidence of crime? Between the size of an organization and the tendency of its employees to experience alienation? In contemporary discussions, questions and conclusions may be expressed in terms of "risk factors." For example, one might seek to find the risk factors associated with dropping out of school, with teen suicide, or with lung cancer. "Risk factors" are features that are associated with these outcomes but that can be discerned prior to the outcome itself. Although a reference to "risk" suggests that the outcome is undesirable, researchers may, of course, explore factors that are associated with positive as well as negative outcomes. For example, one could examine factors associated with appearance in *Who's Who*, with the success of a treatment regimen, or with a positive balance of international trade.

Referring to a "risk factor" entails no claim that the associated factor has an effect on the outcome, whereas a statistical association is sometimes erroneously interpreted as indicating that one variable has an effect on the other. To be sure, an interest in the association between two variables may derive from some hypothesis about an effect. Thus, if it is assumed that retirement has the effect of reducing the likelihood of voting, the implication is that retirement and non-voting should be statistically associated. But the reverse does not hold; that is, the fact that two variables are statistically associated does not, by itself, imply that one of those variables has an effect on the other. For example, if it is true that low attachment to parents encourages involvement in delinquency, it should be true that low attachment and delinquency are statistically associated. But a statistical association between low attachment and delinquency involvement might arise for other reasons as well. If both variables are influenced by a common cause, or if both are manifestations of the same underlying tendency, those variables will be statistically associated with each other, even if there is no effect of

one on the other. Finding a statistical association between two variables, even a strong association, does not, in itself, tell the reason for that association. It may result from an effect of one variable on another, or from the influence of a common cause on both variables, or because both variables reflect the same underlying tendency. Furthermore, if the association is transitory, it may appear simply because of an accident or coincidence. Discovering the reason for a statistical association always entails inquiry beyond simply demonstrating that the association is there.

### WHY MEASURE ASSOCIATION?

The focus here is on measures of association for categorical variables, with brief attention to measures appropriate for ordered categories. Quantitative variables, exemplified by age and income, describe each case by an amount; that is, they locate each case along a scale that varies from low to high. In contrast, categorical variables, exemplified by gender and religious denomination, entail describing each case by a category; that is, they indicate which of a set of categories best describes the case in question. Such categories need not be ordered. For example, there is no inherent ordering for the categories that represent region. But if the categories (e.g., low, medium, and high income) are ordered, it may be desirable to incorporate order into the analysis. Some measures of association have been designed specifically for ordered categories.

The degree of association between categorical variables may be of interest for a variety of reasons. First, if a weak association is found, we may suspect that it is just a sampling fluke—a peculiarity of the sample in hand that may not be found when other samples are examined. The strength of association provides only a crude indication of whether an association is likely to be found in other samples, and techniques of statistical inference developed specifically for that purpose are preferred, provided the relevant assumptions are met.

Second, a measure of the degree of association may be a useful descriptive device. For example, if the statistical association between region and college attendance is strong, that suggests differential access to higher education by region. Furthermore, historical changes in the degree of

association may suggest trends of sociological significance, and a difference in the degree of association across populations may suggest socially important differences between communities or societies. For example, if the occupations of fathers and sons are more closely associated in Italy than in the United States, that suggests higher generational social mobility in the latter than in the former. Considerable caution should be exercised in comparing measures of association for different times or different populations, because such measures may be influenced by a change in the marginal frequencies as well as by a change in the linkage between the variables in question. (See Reynolds 1977)

Third, if a statistical association between two variables arises because of the effect of one variable on the other, the degree of association indicates the relative strength of this one influence as compared to the many other variables that also have such an effect. Unsophisticated observers may assume that a strong association indicates a causal linkage, while a weak association suggests some kind of noncausal linkage. But that would be naïve. The strength of association does not indicate the reason for the association. But *if* an association appears because of a causal link between two variables, the strength of that association provides a rough but useful clue to the relative importance of that particular cause relative to the totality of other causes. For example, income probably influences the tendency to vote Democratic or Republican in the United States, but income is not the only variable affecting the political party favored with a vote. Among other things, voting for one party rather than another is undoubtedly influenced by general political philosophy, by recent legislative actions attributed to the parties, by specific local representatives of the parties, and by the party preferences of friends and neighbors. Such multiple influences on an outcome are typical, and the degree of association between the outcome and just one of the factors that influence it will reflect the relative “weight” of that one factor in comparison to the total effect of many.

Fourth, if a statistical association between two variables arises because both are influenced by a common cause, or because both are manifestations of the same underlying tendency, the degree of association will indicate the relative strength of the common cause or the common tendency, in

comparison to the many other factors that influence each of the two variables. Assume, for example, that participation in a rebellious youth subculture influences adolescents to use both alcohol and marijuana. If the resulting association between the two types of substance use is high, this suggests that the common influence of the rebellious youth subculture (and perhaps other common causes) is a relatively strong factor in both. On the other hand, if this association is weak, it suggests that while the rebellious youth subculture may be a common cause, each type of substance use is also heavily influenced by other factors that are not common to both types.

Fifth, the degree of association indicates the utility of associated factors (“risk factors”) as predictors, and hence the utility of such factors in focusing social action. Assume, for example, that living in a one-parent home is statistically associated with dropping out of high school. If the association between these two variables is weak, knowing which students live in a one-parent home would not be very useful in locating cases on which prevention efforts should be concentrated for maximum effectiveness. On the other hand, if this association is strong, that predictor would be especially helpful in locating cases for special attention and assistance.

In summary, we may be interested in the degree of statistical association between variables because a weak association suggests the possibility that the association is a fluke that will not be replicated, because changes in the degree of association may help discern and describe important social trends or differences between populations, because the degree of association may help determine the relative importance of one variable that influences an outcome in comparison to all other influences, because the degree of association may reflect the degree to which two variables have a common cause, and because the degree of association will indicate the utility of associated factors in predicting an outcome of interest.

#### MEASURING THE DEGREE OF ASSOCIATION

The degree of statistical association between two variables is most readily assessed if, for a suitable set of cases, the relevant information is tallied in a cross-classification table. Table 1 displays such a

**College Attendance by Race in a Sample of 20-Year-Olds in Centerville, 1998 (Hypothetical Data)**

ATTENDING COLLEGE?	WHITE	BLACK	ASIAN-AMERICAN	TOTAL
Yes	400	60	80	540
No	300	140	20	460
Total	700	200	100	1,000

*Table 1*

table. For a contrived sample of young adults, this table shows the number who are and who are not attending college in each of three racial groupings. Hence the two variables represented in this table are (1) race (white, black, Asian-American) and (2) attending college (or not) at a given time. The frequencies in the cells indicate how the cases are jointly distributed over the categories of these two variables. The totals for each row and column (the “marginal frequencies” or simply the “marginals”) indicate how the cases are distributed over these two variables separately.

If young adults in all racial groupings were equally likely to be attending college, then there would be no association between these two variables. Indeed, the simplest of all measures of association is just a percentage difference. For example, blacks in this set of cases were unlikely to be attending college (i.e., 30 percent were enrolled), while Asian-Americans were very likely to be attending (80 percent). Hence we may say that the percentage attending college in the three racial groupings represented in this table ranges from 30 percent to 80 percent, a difference of 50 percentage points. In this table, with three columns, more than one percentage difference could be cited, and the one alluded to above is simply the most extreme of the three comparisons that could be made between racial groupings. Generally speaking, a percentage difference provides a good description of the degree of association only in a table with exactly two rows and two columns. Even so, citing a percentage difference is a common way of describing the degree of statistical association.

Leaving aside the difference between percentages, most measures of association follow one of two master formulas, and a third way of assessing

association provides the basis for analyzing several variables simultaneously. The oldest of these master formulas is based on the amount of departure from statistical independence, normed so that the measure will range from 0 (when the two variables are statistically independent and hence not associated at all) to 1.0 or something approaching 1.0 (when the cross-classification table exhibits the maximum possible departure from statistical independence). The several measures based on this master formula differ from each other primarily in the way the departure from statistical independence is normed to yield a range from 0 to 1.

The second master formula is based on the *improvement* in predictive accuracy that can be achieved by a "prediction rule" that uses one variable to predict the other, as compared to the predictive accuracy achieved from knowledge of the marginal distribution alone. The several measures based on this master formula differ from each other in the nature of the "prediction rule" and also in what is predicted (e.g., the category of each case, or which of a pair of cases will be higher). When such a measure is 0 there is no improvement in predictive accuracy when one variable is predicted from another. As the improvement in predictive accuracy increases, these measures of association will increase in absolute value up to a maximum of 1, which indicates prediction with no errors at all.

A third important way of assessing association, used primarily when multiple variables are analyzed, is based on the difference in *odds*. In Table 1, the odds that an Asian-American is attending college are "4 to 1"; that is, 80 are in college and 20 are not. If such odds were identical for each column, there would be no association, and the ratio of the odds in one column to the odds in another would be 1.0. If such ratios differ from 1.0 (in either direction), there is some association. An analysis of association based on *odds ratios* (and more specifically on the logarithm of odds ratios) is now commonly referred to as a *loglinear analysis*. This mode of analysis is not discussed in detail here.

**Departure from Statistical Independence.** The traditional definition of statistical independence is expressed in terms of the probability of events; that is, events  $A$  and  $B$  are statistically independent if, and only if:

$$P(A/B) = P(A) \quad (1)$$

" $P(A)$ " may be read as the probability that event  $A$  occurs. This probability is usually estimated empirically by looking at the proportion of all relevant events ( $A$  plus not- $A$ ) that are  $A$ . Referring to Table 1, if event  $A$  means attending college, then the probability of event  $A$  is estimated by the proportion of all relevant cases that are attending college. In Table 1, this is .54 (i.e., 540 were attending out of the table total of 1,000).

" $P(A|B)$ " may be read as the conditional probability that event  $A$  occurs, given that event  $B$  occurs, or, more briefly "the probability of  $A$  given  $B$ ." This conditional probability is usually estimated in a manner parallel to the estimation of  $P(A)$  described above, except that the relevant cases are limited to those in which event  $B$  occurs. Referring to Table 1 again, if event  $A$  means attending college and event  $B$  refers to being classified as Asian-American, then  $P(A|B) = .80$  (i.e., 80 are attending college among the 100 who are Asian-American).

As indicated above, the traditional language of probability refers to "events," whereas the traditional language of association refers to "variables." But it should be evident that if "events" vary in being  $A$  or not- $A$ , then we have a "variable." The difference between the language used in referring to the probability of "events" and the language used in referring to a statistical association between "variables" need not be a source of confusion, since one language can be translated into the other.

If we take as given the "marginal frequencies" in a cross-classification table (i.e., the totals for each row and each column), then we can readily determine the probability of any event represented by a category in the table; that is, we can determine  $P(A)$ . Since  $P(A|B) = P(A)$  if statistical independence holds, we can say what  $P(A|B)$  *would be* for any  $A$  and  $B$  in the table if statistical independence held. Otherwise stated, if the marginal frequencies remain fixed, we can say what frequency would be *expected* in each cell if statistical independence held. Referring again to Table 1 and assuming the marginal frequencies remain fixed as shown, if statistical independence held in the table, then 54 percent of those in each racial grouping would be attending college. This is because, if statistical



independence holds, then  $P(A)$  (i.e., .54) must equal  $P(A|B)$  for all  $B$  (i.e., the proportion enrolled in each of the columns). Hence, if statistical independence held, we would have 378 whites in college (i.e., 54 percent of the 700 whites in the table), 108 blacks in college (i.e., 54 percent of the 200 blacks in the table), and 54 Asian-Americans in college (i.e., 54 percent of the 100 Asian-Americans in the table.) Evidently, the number *not* enrolled in each racial grouping could be obtained in a similar way (i.e., 46 percent of each column total), or by subtraction (e.g., 700 whites minus 378 whites enrolled leaves 322 whites not attending college). Hence, the frequencies that would be “expected” for each cell if statistical independence held can be calculated, not just for Table 1 but for any cross-classification table.

If the “expected” frequencies for each cell are very similar to the “observed” frequencies, then the departure from statistical independence is slight. But if the “expected” frequencies differ greatly from the corresponding “observed” frequencies, then the table displays a large departure from statistical independence. When the departure from statistical independence reaches its maximum, an ideally normed measure of association should then indicate an association of 1.0.

A quantity called *chi square* is conventionally used to reflect the degree of departure from statistical independence in a cross-classification table. Chi square was originally devised as a statistic to be used in tests against the null hypothesis; it was not designed to serve as a measure of association for a cross-classification table and hence it does not range between 0 and 1.0. Furthermore, it is not well suited to serve as a measure of association because it is heavily influenced by the total number of cases and by the number of rows and columns in the cross-classification table. Even so, calculating chi square constitutes the first step in calculating measures of association based on departure from statistical independence. For a cross-classification table, this statistic will be zero when the observed frequencies are identical to the frequencies that would be expected if statistical independence held, and chi square will be progressively larger as the discrepancy between observed and expected frequencies increases. As indicated below, in the calculation of chi square, the differences between the frequencies observed and the

frequencies expected if statistical independence held are squared and weighted by the reciprocal of the expected frequency. This means, for example, that a discrepancy of 3 will be more heavily weighted when the expected frequency is 5 than when the expected frequency is 50. These operations are succinctly represented in the following formula for chi square:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{2}$$

where  $\chi^2$  = chi square

$\Sigma$  is the instruction to sum the quantity that follows over all cells

$O_i$  = the frequency observed in the  $i$ th cell

$E_i$  = the frequency expected in the  $i$ th cell if statistical independence holds

To illustrate equation (2), consider Table 2, which shows (1) the observed frequency ( $O$ ) for each cell as previously shown in Table 1; (2) the frequency expected ( $E$ ) for each cell if statistical independence held (*in italics*), and, (3) for each cell, the squared difference between observed and expected frequencies, divided by the expected frequency (in bold type). When the quantities in bold type are summed over the six cells—in accord with the instruction in equation 2—we obtain a chi square of 76.3. There are various ways to norm chi square to create a measure of association, although no way of norming chi square is ideal since the maximum possible value will not be 1.0 under some commonly occurring conditions.

The first measure of association based on chi square was Pearson’s *Coefficient of Contingency* ( $C$ ), which is defined as follows:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \tag{3}$$

This measure can never reach 1.0, although its maximum possible value approaches 1.0 as the number of rows and columns in the table increases. For Table 1,  $C = .27$ .



An alternative measure of association based on chi square is Cramer's  $V$ , which was developed in an attempt to achieve a more appropriately normed measure of association.  $V$  is defined as follows:

$$V = \sqrt{\frac{\chi^2}{(N) [\text{Min} (r-1, c-1)]}} \tag{4}$$

The instruction in the denominator is to multiply the table total ( $N$ ) by whichever is smaller: the number of rows minus 1 or the number of columns minus 1 (i.e., the "Min," or minimum, of the two quantities in parentheses). In Table 1, the minimum of the two quantities is  $r - 1 = 1$ , and the denominator thus becomes  $N$ . Hence, for Table 1,  $V = .28$ .

While this measure *can* reach an upper limit of 1.0 under certain conditions, it cannot be 1.0 in some tables. In Table 1, for example, if *all* of the 540 persons attending college were white, and all blacks and Asian-Americans were not in college, chi square would reach a value of approximately 503 (and no other distribution that preserves the marginals would yield a larger chi square). This maximum possible departure from statistical independence (given the marginal frequencies) yields a  $V$  of .71.

In the special case of a cross-classification table with two rows and two columns (a "2 x 2 table")  $V$  becomes phi ( $\Phi$ ), where

$$\Phi = \sqrt{\frac{\chi^2}{N}} \tag{5}$$

The maximum possible value of phi in a given table is 1.0 if and only if the distribution over the two row categories is identical to the distribution over the two column categories. For example, if the cases in the two row categories are divided, with 70 percent in one and 30 percent in the other, the maximum value of phi will be 1.0 if and only if the cases in the two column categories are also divided, with 70 percent in one and 30 percent in

**College Attendance by Race: Observed Frequencies (from Table 1), Expected Frequencies Assuming Statistical Independence (in *Italic*), and  $(O-E)^2/E$  for Each Cell (in **Bold Type**)**

ATTENDING COLLEGE?	WHITE	BLACK	ASIAN-AMERICAN	TOTAL
Yes	400 <i>378</i> <b>1.3</b>	60 <i>108</i> <b>21.3</b>	80 <i>54</i> <b>12.5</b>	540
No	300 <i>322</i> <b>1.5</b>	140 <i>92</i> <b>25.0</b>	20 <i>46</i> <b>14.7</b>	460
Total	700	200	100	1,000

$$\chi^2 = 1.3 + 21.3 + 12.5 + 1.5 + 25.0 + 14.7 + 76.3$$

$$C = 0.27$$

$$V = 0.28$$

Table 2

the other. Some consider measures of association based on chi square to be flawed because commonly encountered marginals may imply that the association cannot possibly reach 1.0, even if the observed frequencies display the maximum possible departure from statistical independence, given the marginal frequencies. But one may also consider this feature appropriate because, if the degree of statistical association in a cross-classification table were perfect, the marginal distributions would not be disparate in a way that would limit the maximum value of the measure.

A more nagging concern about measures of association based on the departure from statistical independence is the ambiguity of their meaning. One can, of course, use such measures to say that one association is very weak (i.e., close to zero) and that another is relatively strong (i.e., far from zero and perhaps close to the maximum possible value, given the marginals), but "weak" and "strong" are relatively crude descriptors. The measures of association based on chi square may also be used in making comparisons. Thus, if a researcher wished to compare the degree of association in two populations,  $C$  or  $V$  could be compared for the two populations to determine whether the association was approximately the same in both and, if not, in which population the association was stronger. But there is no clear interpretation that can be

attached to a coefficient of contingency of precisely .32, or a Cramer's  $V$  of exactly .47.

**Relative Reduction in Prediction Error.** We shift now to measures of association that reflect the relative reduction in prediction error. Since such measures indicate the proportion by which prediction errors are reduced by shifting from one prediction rule to another, we follow common practice and refer to them as *proportional reduction in error* (PRE) measures. Every PRE measure has a precise interpretation, and sometimes it is not only precise but also clear and straightforward. On the other hand, the PRE interpretation of some measures may seem strained and rather far removed from the common sense way of thinking about the prediction of one variable from another.

The basic elements of a PRE measure of association are:

1. a specification of what is to be predicted, and a corresponding definition of prediction error. For example, we might say that what is to be predicted is the row category into which each case falls, and a corresponding definition of prediction error would be that a case falls in a row category other than that predicted. Referring again to Table 1, if we are predicting whether a given case is attending college or not, a corresponding definition of prediction error would be that our predicted category (attending or not) is not the same as the observed category for that case.
2. a rule for predicting either the row variable or the column variable in a cross-classification table from knowledge of the marginal distribution of that variable alone. We will refer to the prediction error when applying this rule as  $E_1$ . For example, if what is to be predicted is as specified above (i.e., the row category into which each case falls), the rule for predicting the row variable from knowledge of its marginal distribution might be to predict the modal category for every case. This is not the only possible prediction rule but it is a reasonable one (and there is no rule based only on the marginals that would have higher predictive accuracy). Applying this rule to Table

1, we would predict "attending college" (the modal category) for every case. We would then be wrong in 460 cases out of the 1,000 cases in the table, that is, each of the 460 cases not attending college would be a prediction error, since we predicted attending college for every case. Hence, in this illustration  $E_1 = 460$ .

3. A rule for predicting the same variable as in step (2) from knowledge of the joint distribution of both variables. We will refer to the prediction error when applying this rule as  $E_2$ . For example, continuing with the specifications in steps (1) and (2) above, we specify that we will predict the row category for each case by taking the modal category for each column. Thus, in Table 1, we would predict "attending college" for all whites (the modal category for whites), "not attending college" for all blacks (the modal category for blacks), and "attending college" for Asian-Americans. The prediction errors are then 300 for whites (i.e., the 300 not attending college, since we predicted attending for all whites), 60 for blacks (i.e., the 60 attending college, since we predicted nonattending for all blacks), and 20 for Asian-Americans, for a total of 380 prediction errors. Thus, in this illustration  $E_2 = 380$ .
4. The calculation of the proportion by which prediction errors are reduced by shifting from the rule in step (2) to the rule in step (3), that is, the calculation of the proportional reduction in error. This is calculated by:

$$\text{PRE} = \frac{E_1 - E_2}{E_1} \tag{6}$$

The numerator in this calculation is the amount by which error is reduced. Dividing this amount by the starting error indicates what proportion of the possible reduction in prediction error has actually been achieved. Utilizing the error calculations above, we can compute the proportional reduction in error.

$$\text{PRE} = \frac{460 - 380}{460} = \frac{80}{460} = .174 \quad (7)$$

This calculation indicates that we achieve a 17.4 percent reduction in prediction error by shifting from predicting the row category that is the marginal mode for all cases to predicting the column-specific modal category for all cases in a given column.

The PRE measure of association with prediction rules based on modal categories (illustrated above) is undoubtedly the simplest of the many PRE measures that have been devised. This measure is called lambda ( $\lambda$ ), and for a given cross-classification table there are two lambdas. One of these focuses on predicting the row variable ( $\lambda_r$ ), and the other focuses on predicting the variable that is represented in columns ( $\lambda_c$ ). In the illustration above, we computed  $\lambda_r$ ; that is, we were predicting college attendance, which is represented in rows. Shifting to  $\lambda_c$  (i.e., making the column variable the predicted variable), we find that the proportional reduction in error is 0. This outcome is evident from the fact that the modal column for the table ("White") is also the modal column for each row. Thus, the prediction errors based on the marginals sum to 300 (i.e., the total who are not "white") and the prediction errors based on row-specific modal categories also sum to 300, indicating no reduction in prediction error. Thus, the proportional reduction in prediction error (as measured by lambda) is not necessarily the same for predicting the row variable as for predicting the column variable.

An alternative PRE measure for a cross-classification table is provided by Goodman and Kruskal's tau measures ( $\tau_r$  and  $\tau_c$ ) (1954). These measures are based on prediction rules that entail distributing predictions so as to recreate the observed distributions instead of concentrating all predictions in the modal category. In doing so, there is an expected number of misclassified cases (prediction errors), and these expected numbers are used in calculating the proportional reduction in error. In Table 1,  $\tau_r = .25$  and  $\tau_c = .05$ . Although  $\lambda_c$  was found to be zero for Table 1 because the modal column is the same in all rows,  $\lambda_r$  is not zero because the percentage distributions within rows are not identical.

Other PRE measures of association have been developed, and some have been designed specifically for a cross classification of ordered categories (ordinal variables). For example, if people were classified by the highest level of education completed (e.g., into the categories pre-high school, high school graduation, bachelor's degree, higher degree) we would have cases classified into a set of ordered categories and hence an ordinal variable. If the same cases were also classified into three levels of income (high, medium, and low) the result would be a cross classification of two ordinal variables. Although several measures of association for ordinal variables have been devised, the one now most commonly used is probably Goodman and Kruskal's gamma ( $\gamma$ ). Gamma is a PRE measure, with a focus on the prediction of order within pairs of cases, with order on one variable being predicted with and without knowledge of the order on the other variable, disregarding pairs in which there is a tie on either variable.

These and other PRE measures of association for cross-classification tables are described and discussed in several statistics texts (See, for example, Blalock 1979; Knoke and Bohrnstedt 1991; Loether and McTavish 1980; Mueller et al. 1977). In some instances, the prediction rules specified for a given PRE measure may closely match the specific application for which a measure of association is sought. More commonly, however, the application will not dictate a specific kind of prediction rule. The preferred measure should then be the one that seems likely to be most sensitive to the issues at stake in the research problem. For example, in seeking to identify "risk factors" associated with a relatively rare outcome, or with a very common outcome, one of the tau measures would be more appropriate than one of the lambda measures, because the modal category may be so dominant that lambda is zero in spite of distributional differences that may be of interest.

### MEASURES OF ASSOCIATION IN APPLICATION

When the initial measures of association were devised at the beginning of the twentieth century, some regarded them as part of a new mode of inquiry that would replace speculative reasoning

and improve research into the linkages between events. At the end of the twentieth century, we now recognize that finding a statistical association between two variables raises more questions than it answers. We now want to know more than the degree to which two variables are statistically associated; we want also to know why they are associated: that is, what processes, what conditions, and what additional variables are entailed in generating the association?

To a limited degree, the measures of association discussed above can be adapted to incorporate more than two variables. For example, the association between two variables can be explored separately for cases that fall within each category of a third variable, a procedure commonly referred to as “elaboration.” Alternatively, a new variable consisting of all possible combinations of two predictors can be cross-classified with an outcome variable. But the traditional measures of association are not ideally suited for the task of exploring the reasons for association. Additional variables (e.g., potential sources of spuriousness, variables that mediate the effect of one variable on another, variables that represent the conditions under which an association is weak or strong) need to be incorporated into the analysis to yield an improved understanding of the meaning of an observed association—not just one at a time but several simultaneously. Additional modes of analysis (e.g., loglinear analysis; see Goodman 1970; Knoke and Burke 1980) have been developed to allow an investigator to explore the “interactions” between multiple categorical variables in a way that is roughly analogous to multiple regression analysis for quantitative variables. Computer technology has made such modes of analysis feasible.

The same technology has generated a new use for relatively simple measures of association in exploratory data analysis. It is now possible to describe the association between hundreds or thousands of pairs of variables at very little cost, whereas at an earlier time such exhaustive coverage of possible associations would have been prohibitively expensive. Measures of association provide a quick clue to which of the many associations explored may identify useful “risk factors” or which associations suggest unsuspected linkages worthy of further exploration.

## REFERENCES

- Blalock, Hubert M., Jr. 1979 *Social Statistics*, 2nd ed. New York: McGraw-Hill.
- Bohrnstedt, George W., and David Knoke 1988 *Statistics for Social Data Analysis*, 2nd ed. Itasca, Ill.: F. E. Peacock Publishers.
- Costner, Herbert L. 1965 “Criteria for Measures of Association.” *American Sociological Review* 30:341–353.
- Fienberg, S. E. 1980 *The Analysis of Cross-Classified Categorical Data*. Cambridge, Mass.: MIT Press.
- Goodman, Leo A. 1970 “The Multivariate Analysis of Qualitative Data: Interactions among Multiple Classifications.” *Journal of the American Statistical Association* 65:226–257.
- 1984 *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.
- Goodman, Leo A., and William H. Kruskal 1954 “Measures of Association for Cross Classifications.” *Journal of the American Statistical Association* 49:732–764.
- 1959 “Measures of Association for Cross Classifications: II. Further Discussion and References.” *Journal of the American Statistical Association* 54:123–163.
- 1963 “Measures of Association for Cross Classifications: III. Approximate Sampling Theory.” *Journal of the American Statistical Association* 58:310–364.
- 1972 “Measures of Association for Cross Classifications: IV. Simplification of Asymptotic Variances.” *Journal of the American Statistical Association* 67:415–421.
- Kim, Jae-on 1984 “PRU Measures of Association for Contingency Table Analysis.” *Sociological Methods and Research* 13:3–44.
- Knoke, David, and George W. Bohrnstedt 1991 *Basic Social Statistics*. Itasca, Ill.: F. E. Peacock Publishers.
- Knoke, David, and Peter Burke 1980 *Loglinear Models*. Beverly Hills, Calif.: Sage.
- Loether, Herman J., and Donald G. McTavish 1980 *Descriptive Statistics for Sociologists: An Introduction*. Boston: Allyn and Bacon.
- Mueller, John H., Karl F. Schuessler, and Herbert L. Costner 1977 *Statistical Reasoning in Sociology*. Boston: Houghton Mifflin.
- Reynolds, H. T. 1977 *Analysis of Nominal Data*. Beverly Hills, Calif.: Sage.

HERBERT L. COSTNER

## MEDIA

See American Society; Mass Media Research; Popular Culture.

## MEDICAL SOCIOLOGY

Over the past several decades medical sociology has become a major subdiscipline of sociology, at the same time assuming an increasingly conspicuous role in health care disciplines such as public health, health care management, nursing, and clinical medicine. The name *medical sociology* garners immediate recognition and legitimacy and, thus, continues to be widely used—for instance, to designate the Medical Sociology Section of the American Sociological Association—even though most scholars in the area concede that the term is narrow and misleading. Many courses and texts, rather than using the term “sociology of medicine,” refer instead to the sociology of health, health and health care, health and illness, health and medicine, or health and healing. The study of medicine is only part of the sociological study of health and health care, a broad field ranging from (1) *social epidemiology*, the study of socioeconomic, demographic, and behavioral factors in the etiology of disease and mortality; to (2) studies of the *development and organizational dynamics* of health occupations and professions, hospitals, health maintenance and long-term care organizations, including interorganizational relationships as well as interpersonal behavior, for example, between physician and patient; to (3) the *reactions of societies* to illness, including cultural meanings and normative expectations and, reciprocally, the reactions of individuals in interpreting, negotiating, managing, and socially constructing illness experience; to (4) the *social policies, social movements, politics, and economic conditions* that shape and are shaped by health and disease within single countries, as well as in a comparative, international context.

The rise of contemporary medical sociology can be traced back to the immediate post-World War II period, when science and medicine were dominant cultural forces, fueling a modern optimism that many of society's ills could be eliminated. Several key contributions during the 1950s gave credibility and spurred scholarly interest in the newly developing subfield. Koos's *The Health of Regionville* (1954) and Hollingshead and Redlich's

*Social Class and Mental Illness* (1958) addressed the connections between social circumstances and health status, and were instrumental in establishing a strong tradition of sociological research focusing on the social determinants of health. The finding that individuals in the lower socioeconomic levels of society experience greater morbidity and mortality has turned out to be one of the most consistent of these patterns. Also during this time, a number of sociology's most prominent theorists turned their attention to health and health care. They approached the topic not because their primary interest was in health care or medicine, but out of a generic interest in authority and the maintenance of social order. Robert Merton, Everett Hughes, and Anselm Strauss all studied professional organizations and socialization during the 1950s, focusing primarily on physicians and the process of medical education (Merton et al. 1957; Becker et al. 1961).

The theoretical work of the 1950s most influential for medical sociology was undoubtedly Talcott Parsons's *The Social System* (1951). In it, Parsons recognized illness as a major threat to the stability and productivity of societies and introduced the “sick role” concept to describe the social regulation of sickness and explain the mechanism through which individuals are induced to return to productive activity. Parsons argued that because sick persons were unable to perform their expected social roles, they were subject to being negatively sanctioned. On the other hand, if they had not intended to become ill and were motivated to get well, then, according to Parsons's analysis, they could claim and be granted temporary exemption without blame from normally expected role responsibilities. Rather than being held accountable for failure to perform, they would be excused as sick. Parsons's work generated enormous sociological interest because of its analysis of illness and medical care in terms of their broad social consequences and because of its focus on the structure and functions of social roles. His work also expanded the theoretical foundations of medical sociology by provoking equally compelling work from contrasting perspectives. Elliot Freidson in *Profession of Medicine* (1970) analyzed the dominance of the medical profession, suggesting that power relations in health care were fundamentally contentious. He saw physicians as rising to dominate health care through a process of

struggle with competitors in which they prevailed largely because they gained the support of political institutions, limiting the role of competing occupations. In contrast to the fixed roles in structural-functional theory, Freidson argued that illness definitions and illness behavior were socially constructed through a process of negotiation. The debate over structure and agency represented in these early contributions laid theoretical pathways for subsequent scholarship and solidified medical sociology's ties to some of the central issues of the discipline.

Medical sociology became established in only a few sociology departments during its early years, typically in elite universities. It was not until the 1970s that most graduate departments of sociology began to offer medical sociology. Today, sociology courses on health and medicine can be found in nearly every graduate program in the United States as well as in many other nations, notably the United Kingdom and Germany (Bloom 1986). Research funding to support the growth of medical sociology in many countries has come from government sources. In the 1960s and 1970s, U.S. medical sociology expanded in part because social science research was held in favor by the federal government as well as by influential private foundations. Major funding sources at that time included the National Institute of Mental Health (NIMH) and, later, the National Center for Health Services Research (NCHSR).

It has been argued that the fortunes of medical sociology have shifted in relation to the social-medical environment (Pescosolido and Kronenfeld 1995). Until the 1980s, medical sociology experienced relatively fertile conditions due in part to the fact that the health care system was dominated by professional medicine. Access to health care was the primary health policy concern, while research funding priorities focused on the biomedical and psychosocial aspects of disease, disease prevention, and patient care. This environment encouraged medical sociologists to pursue quantitative research, including surveys, national-level studies, and multivariate statistical models that predicted utilization of health services and the effects of risk factors and other variables. Two particular lines of medical sociology research gained prominence as a result of this focus. The first involved researchers studying utilization patterns for health services. There were two groups, each

using a somewhat different explanatory model. Marshall Becker and his colleagues employed the Health Belief Model, a cognitive framework originated by Rosenstock (1966) and eventually applied in research, to explain a wide variety of preventive and health-related behaviors (Becker and Maiman 1975). Ronald Andersen developed the somewhat broader sociobehavioral model (1995), which included health beliefs but also emphasized economic factors and health needs. The second line of research concerned quantitative studies of social stress. David Mechanic, one of the founders of medical sociology, pioneered sociological research on stress and mental health as early as the 1960s (Mechanic and Volkart 1961). The "stress process" group that emerged in the late 1970s, however, was closer to an interface of psychology and sociology. Using multivariate analyses, they examined the relationships among stress (Aneshensel 1992), social support (Turner and Marino 1994), and coping (Pearlin and Schooler 1978). Much of this research was published in the American Sociological Association's *Journal of Health and Social Behavior*, beginning in the late 1970s and continuing into the present (Thoits 1995).

The social-medical environment in the United States changed dramatically in the 1980s, threatening the autonomy and authority of physicians (Starr 1982). The federal government's increasing role in financing health care (through the Medicare and Medicaid programs) combined with rapidly escalating health care costs and the concern expressed by business, leading to a major federal policy shift. Rather than inequality in access and social factors in illness, public policy attention was now placed on cost control and the cost effectiveness of care. NIMH support for medical sociology was weakened, and soon afterward, the NCHSR became the Agency for Health Care Policy and Research with an agenda of research focused on managed care and evidence-based medicine. Research funding priorities gravitated from the behavioral and social sciences to economics and clinical medicine and epidemiology. No doubt these changes contributed to critical claims in the late 1980s and the 1990s, that medical sociology research had become fragmented.

The significance of health system changes for the profession of medicine became a hotly debated topic among medical sociologists during the

1980s. The controversy was sparked in 1985 with the publication of McKinley and Arch's "Toward the Proletarianization of Physicians," in which the authors argued that historical processes of bureaucratic rationalization were finally reaching medicine, irreversibly eroding the functional autonomy of physicians. This directly challenged Freidson's medical dominance perspective (1970). Also part of the debate was the hypothesis, introduced by Marie Haug (1976), that physicians had lost authority due to the increasing knowledge and medical sophistication of patients. A plethora of articles appeared, identifying and discussing at length various hypothesized changes in medical dominance and authority and culminating, though by no means ending, with a special issue of the *Milbank Quarterly* in 1988.

Much of the early growth of medical sociology can be attributed to scholars located outside sociology departments in medical schools, nursing schools, schools of public health, and health administration programs. These individuals addressed research concerns and questions that were of paramount importance in their respective settings, such as the reasons people engage in health-promoting behavior, define themselves as sick, use health services, and comply with medical treatment. They contributed to medical and health care disciplines by bringing attention to the significance of culture and human interaction in producing the *meaning* of illness and shaping illness-related behavior (Zola 1966; Mechanic 1995). They dispelled the image of the physician as a purely rational scientist. Sociologists also contributed to the development of social epidemiology, mapping the social patterns of disease, and adding social factors to the causal understanding of mortality and chronic diseases (Berkman and Syme 1979). A third group studied hospitals and health care organizations, bringing an organizational sociology perspective into the field of health services research (Flood and Fennell 1995).

Robert Straus, a medical school sociologist, introduced in 1957 what became for many years a popular way of dividing the subfield. Sociologists such as those described above were designated "sociologists *in* medicine" in contrast to sociologists *of* medicine who were typically based in sociology departments. According to Straus, sociologists *of* medicine used medical settings to address

questions of sociology while sociologists *in* medicine used sociological knowledge to address questions of medicine. Today, the boundaries between those working in health care settings and those in academic departments of sociology are blurred; sociologists in both venues conduct applied research as well as research that contributes to basic sociological theorizing. In fact, it is quite common for medical sociologists to have multiple academic appointments. On the other hand, the distinction remains valid in the pressure to conduct research that reflects the priorities of the dominant group. Medical sociologists in medicine often engage in research shaped by medical issues and a biomedical approach, whereas those in sociology have an easier time posing sociological questions grounded in sociological theory. In its early years, medical sociology was sometimes dismissed by other academic sociologists as "applied" sociology, based on the rather elitist assumption that its research did not contribute to the basic body of knowledge of the discipline and that it lacked a theoretical body of its own. Today, there is greater understanding of the links between basic sociological theory and medical sociology (Gerhardt 1989). Medical sociology concepts such as "medicalization" have added to the broader understanding of social order and social control (Conrad 1992). Medicine and the other health care disciplines recognize sociology as a valuable discipline that can contribute much to the understanding and application of health care. Academic sociology has come to regard the sociology of medicine as a fruitful area of specialization.

It is in their role of social critic that medical sociologists encounter the greatest resistance from mainstream medicine and health care. Critical medical sociology emerged from both Marxist and social constructionist traditions within the discipline (Waitzkin 1989; Brown 1995). Symbolic interactionists and labeling theorists in the 1960s saw that, despite the Parsonian notion of the sick role, many types of illness and disability were responded to socially as forms of deviance. Goffman's concept of stigma (1963) explored the relationship between labeling and identity as a process of managing spoiled identity. One of the most powerful explanatory concepts in medical sociology, stigma has been used for decades to capture the experience of mental illness, alcoholism, physical disability, and many types of chronic illness. Goffman



(1961) and Zola (1972), among others, turned the standard notion of medical care as a service on its head by arguing that medicine functions as an institution of social control. Despite strong microsociological interest in the social construction and social consequences of medical labels (i.e., diagnoses), the professional power of physicians made it exceedingly difficult for sociologists to study these processes until the 1980s. What could be studied, however, using the broader, cultural meaning of social construction, were processes of medicalization. Building on the social control perspective of Zola and others, a number of studies examined the processes through which nonmedical phenomena—such as childbirth, excessive drinking, children’s active behavior, and menstrual distress—became medical phenomena, with diagnostic criteria and specific medical treatments.

Bias in medicine and social inequality in health care have been concerns of critical medical sociology as well as of corresponding social movements initiated to improve health care. Gender analyses, especially those from a feminist perspective, offered a critical, alternative perspective on the medical profession (Lorber 1984) and the health care system (Zimmerman and Hill 1999), as did research on the women’s health movement (Ruzek 1978; Weisman 1998). This work examined the relationship between cultural ideas about gender, medical knowledge, and gender stratification systems; pointed out that the division of labor in medicine is also a gendered division of labor; and observed that the factors that often make women sick are linked to their social roles and disadvantaged social circumstances. Other critical perspectives were offered by disability researchers (Zola 1982) and by researchers focusing on the health and health care of racial and ethnic minorities (Hill 1992; Williams and Collins 1995).

The critical perspective in medical sociology was fortified by Mishler’s (1981) critique of the biomedical model, in which he argued that medicine was itself a culture, based as much on customs, social norms, and values as on scientific fact. Mishler’s view of medicine as socially constructed led to a concern with medical discourse analysis (1984) and, for some researchers, to the study of illness narratives. Departing from the political and critical concerns of the 1980s and 1990s, these

scholars have conducted in-depth, qualitative studies of illness experience, incorporating aspects ignored by their predecessors, such as emotions and the body (Charmaz 1991; Weitz 1991). The “postmodern turn” that swept over academic humanities and social science departments in the latter decades of the twentieth century influenced a number of symbolic interactionist and social constructionist medical sociologists. Working at the interface of constructionism and postmodernism, these scholars created new ways to explore the relationship between illness and identity (Frank 1995; Hall 1998).

Reviewing the literature of medical sociology reveals an unusually broad range of topics, theoretical perspectives, and research methodologies. Beyond the contributions reviewed above, medical sociologists are also active in international comparative research studying health systems or specific health care sectors within them. They are involved in health policy research both at the federal and at the local community level; they are studying alternative health care providers and their clients as well as various forms of folk medicine and lay care; and they are doing research on informal caregivers and the process of care work. Even these additions do not exhaust the parameters of the field. Medical sociology has enriched and continues to enrich the discipline of sociology, as well as making unique and valuable contributions to important policy issues and to the needs of health care professionals, managers, and patients.

#### REFERENCES

- Andersen, Ronald 1995 “Revisiting the Behavioral Model and Access to Care: Does It Matter?” *Journal of Health and Social Behavior* 36:1–10.
- Aneshensel, Carol S. 1992 “Social Stress: Theory and Research.” *Annual Review of Sociology* 18:15–38.
- Avison, William R., and R. Jay Turner 1988 “Stressful Life Events and Depressive Symptoms: Disaggregating the Effects of Acute Stressors and Chronic Strains.” *Journal of Health and Social Behavior* 29:253–264.
- Becker, Howard S., Blanche Geer, Everett C. Hughes, and Anselm Strauss 1961 *Boys in White: Student Culture in Medical School*. Chicago: University of Chicago Press.
- Becker, Marshall, and Lois Maiman 1975 “Sociobehavioral Determinants of Compliance with Health and Medical Care Recommendations.” *Medical Care* 13:10–24.



- Berkman, Lisa, and Leonard Syme 1979 "Social Network, Host Resistance and Mortality: A Nine Year Follow-up Study of Alameda County Residents." *American Journal of Epidemiology* 190:186-204.
- Bloom, Samuel W. 1986 "Institutional Trends in Medical Sociology." *Journal of Health and Social Behavior* 27:265-276.
- Brown, Phil 1995 "Naming and Framing: The Social Construction of Diagnosis and Illness." *Journal of Health and Social Behavior* (Extra Issue):34-52.
- Charmaz, Kathy 1991 *Good Days, Bad Days: The Self in Chronic Illness and Time*. New Brunswick, N.J.: Rutgers University Press.
- Conrad, Peter 1992 "Medicalization and Social Control." *Annual Review of Sociology* 18:137-162.
- Flood, Ann Barry, and Mary L. Fennell 1995 "Through the Lenses of Organizational Sociology: The Role of Organizational Theory and Research in Conceptualizing and Examining Our Health Care System." *Journal of Health and Social Behavior* (Extra Issue):154-169.
- Frank, Arthur 1995 *The Wounded Storyteller: Body, Illness and Ethics*. Chicago: University of Chicago Press.
- Freidson, Elliot 1970 *Profession of Medicine*. New York: Dodd, Mead.
- Gerhardt, Uta 1989 *Ideas about Illness: An Intellectual and Political History of Medical Sociology*. New York: New York University Press.
- Goffman, Erving 1961 *Asylums*. New York: Doubleday.
- 1963 *Stigma: Notes on the Management of Spoiled Identity*. Englewood Cliffs, N.J.: Prentice Hall.
- Hall, Lisa Cox 1998 "Illness, Identity and Survivorship: Modern and Postmodern Breast Cancer Narratives." *Illness, Crisis and Loss* 6:255-274.
- Haug, Marie 1976 "The Erosion of Professional Authority: A Cross-Cultural Inquiry in the Case of the Physician." *Milbank Memorial Fund Quarterly* 54:83-106.
- Hill, Shirley A. 1992 *Managing Sickle Cell Disease in Low Income Families*. Philadelphia: Temple University Press.
- Hollingshead, August B., and Frederick Redlich 1958 *Social Class and Mental Illness*. New York: John Wiley.
- Koos, Earl L. 1954 *The Health of Regionville*. New York: Columbia University Press.
- Lorber, Judith 1984 *Women Physicians: Careers, Status and Power*. New York: Tavistock.
- McKinley, John, and Joan Arches 1985 "Toward the Proletarianization of Physicians." *International Journal of Health Services* 15:161-195.
- Mechanic, David 1995 "Sociological Dimensions of Illness Behavior." *Social Science and Medicine* 41:1207-1216.
- , and E. H. Volkart 1961 "Stress, Illness Behavior and The Sick Role." *American Sociological Review* 26:51.
- Merton, Robert K., G. G. Reeder, and Patricia Kendall 1957 *The Student Physician*. Cambridge Mass.: Harvard University Press.
- Mishler, Elliot G. 1981 "Critical Perspectives on the Biomedical Model." In E. G. Mishler, Lorna R. Amara Singham, Stuart T. Hauser, Ramsay Liem, Samuel D. Osherson, and Nancy E. Waxler, eds., *Social Contexts of Health, Illness, and Patient Care*. New York: Cambridge University Press.
- 1984 *The Discourse of Medicine: Dialectics of Medical Interviews*. Norwood, N.J.: Ablex.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- Pearlin, Leonard I., and Carmi Schooler 1978 "The Structure of Coping." *Journal of Health and Social Behavior* 19:2-21.
- Pescosolido, Bernice A., and Jennie J. Kronenfeld 1995 "Health, Illness, and Healing in an Uncertain Era: Challenges from and for Medical Sociology." *Journal of Health and Social Behavior* (Extra Issue):5-33.
- Rosenstock, Irwin M. 1966 "Why People Use Health Services." *Milbank Memorial Fund Quarterly* 44:94-106.
- Ruzek, Sheryl Burt 1978 *The Women's Health Movement: Feminist Alternatives to Medical Control*. New York: Praeger.
- Starr, Paul 1982 *The Social Transformation of American Medicine*. New York: Basic Books.
- Strauss, Robert 1957 "The Nature and Status of Medical Sociology." *American Sociological Review* 2:200-204.
- Thoits, Peggy A. 1995 "Stress, Coping and Social Support Processes: Where Are We? What Next?" *Journal of Health and Social Behavior* (Extra Issue):53-79.
- Turner, J. Jay, and Franco Marino 1994 "Social Support and Social Structure: A Descriptive Epidemiology." *Journal of Health and Social Behavior* 35:193-212.
- Waitzkin, Howard 1989 "A Critical Theory of Medical Discourse: Ideology, Social Control, and the Processing of Social Context in Medical Encounters." *Journal of Health and Social Behavior*. 30:220-239.
- Weisman, Carol S. 1998 *Women's Health Care: Activist Traditions and Institutional Change*. Baltimore, Md.: Johns Hopkins University Press.
- Williams, David R., and Chiquita Collins 1995 "U.S. Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology* 21:349-386.
- Weitz, Rose 1991 *Life with AIDS*. New Brunswick, N.J.: Rutgers University Press.

- Zimmerman, Mary K., and Shirley A. Hill 1999 "Health Care as a Gendered System." In J. S. Chafetz, ed., *Handbook of the Sociology of Gender*. New York: Plenum.
- Zola, Irving K. 1966 "Culture and Symptoms: An Analysis of Patients' Presenting Complaints." *American Sociological Review* 31:615-630.
- 1972 "Medicine as an Institution of Social Control." *Sociological Review* 20:487-504.
- 1982 *Missing Pieces: A Chronicle of Living with a Disability*. Philadelphia, Pa.: Temple University Press.

MARY K. ZIMMERMAN

## MEDICAL-INDUSTRIAL COMPLEX

The concept of the medical-industrial complex was first introduced in the 1971 book, *The American Health Empire* (Ehrenreich and Ehrenreich 1971) by Health-PAC. The medical-industrial complex (MIC) refers to the health industry, which is composed of the multibillion-dollar congeries of enterprises including doctors, hospitals, nursing homes, insurance companies, drug manufacturers, hospital supply and equipment companies, real estate and construction businesses, health systems consulting and accounting firms, and banks. As employed by the Ehrenreichs, the concept conveys the idea that an important (if not the primary) function of the health care system in the United States is business (that is, to make profits) with two other secondary functions, research and education.

Since that time, a number of authors have examined the medical-industrial complex: Navarro (1976, pp. 76, 80), Relman (1980), Estes and colleagues (1984), Wohl (1984), and McKinlay and Stoeckle (1994). Himmelstein and Woolhandler (1990) argue that health care facilitates profit making by (1) improving the productivity (health) of workers, (2) ideologically ensuring the social stability needed to support production and profit, and (3) providing major opportunities for investment and profit (p. 16). The last function, profit, is now "the driving force," as health care has fully "come into the age of capitalist production" (p. 17).

Arnold Relman (1980), Harvard medical professor and editor of the *New England Journal of*

*Medicine*, was the first mainstream physician to write about the medical-industrial complex, observing that the corporatization of medicine is a challenge to physician authority, autonomy, and even legitimacy for the doctors who become health care industry owners. Ginzberg (1988) and others (Andrews 1995; Estes et al. 1984; Himmelstein and Woolhandler 1990) have written about the monetarization, corporatization, and proprietarization of "health" care. By the mid-1980s, the author of a book appearing with the title *The Medical Industrial Complex* (Wohl 1984) did not see the need to define it but, rather, began with "the story of the explosive growth of . . . corporate medicine" and focused on "medical moguls," monopoly, and a prescription for profit.

While the health care industry has certainly contributed to improvements in the health status of the population, it has also strengthened and preserved the private sector and protected a plurality of vested interests. In U.S. society, the medical-industrial complex functions economically as a source of growth, profit accumulation, investment opportunity, and employment (Estes et al. 1984, pp. 56-70). It also contributes to the human capital needed for productivity and profit by preserving an able-bodied workforce whose work is not sapped by illness (Rodberg and Stevenson 1977), although another interpretation suggests that private capital's stability is built upon the appropriation of the working-class population's health (see Navarro 1976, 1982, 1995).

## STRUCTURE OF THE HEALTH CARE INDUSTRY

*Industry Components* Today's medical-industrial complex consists of more than a dozen major components: hospitals; nursing homes; physicians (salaried and fee-for-service); home health agencies; supply and equipment manufacturers; drug companies; insurance companies; managed care organizations (HMOs, PPOs, IPAs); specialized centers (urgi, surgi, dialysis); hospices; nurses and all other health care workers; administrators, marketers, lawyers, and planners; and research organizations. In addition to these entities, thousands of other organizations are springing up in long-term care (e.g., case management, respite care, homemaking/chore, independent living center) and other

services for the disabled and aging, including social services that have incorporated health care components such as senior centers.

**Changes in the Structure of the Industry.**

There were a number of significant changes in the structure of the health care industry between the 1970s and 1990s, including (1) rapid growth and consolidation of the industry into larger organizations; (2) horizontal integration; (3) vertical integration; (4) change in ownership from government to private, nonprofit, and for-profit organizations; and (5) diversification and corporate restructuring (Starr 1982; McKinlay and Stoeckle 1994). These changes occurred across the different sectors, which are dominated by large hospital, insurance, and managed care organizations.

*Rapid Growth and Consolidation* Health care has long been moved from its cottage industry stage with small individual hospitals and solo physician practitioners to large corporate enterprises. Health care corporations are diverse and growing in terms of size and complexity. Hospitals are the largest sector of the health care industry, and while the growth rate in hospital expenditures was increasing rapidly, the number of community hospitals actually declined from 5,830 in 1980 to 5,194 in 1995 (a decrease of 11 percent) (American Hospital Association 1996) (see Table 1). The number of community hospital beds also began to decline going down to 988,000 in 1980 and continuing so that by 1995 there were only 873,000 (a decrease from 12 percent) (AHA 1996).

Nursing homes grew rapidly in numbers of facilities and beds after the passage of Medicaid and Medicare legislation. In 1996, there were 17,806 licensed nursing facilities with 1.82 million beds (Harrington et al. 1998). The number of facilities increased by 25 percent and the number of beds increased by 37 percent between 1978 and 1996. More recently, their overall growth has leveled off, so that growth is not keeping pace with the aging of the population (Harrington et al. 1998).

Relatively new and influential corporate forces in the health industry are the managed-care organizations such as health maintenance organizations (HMOs), preferred provider organizations (PPOs), and independent practice associations (IPAs). There has been a large growth in HMOs, which provide health care services on the basis of fixed monthly charges per enrollee. In 1984, there

were only 337 HMOs with 17 million enrollees. By 1988, there were 31 million members enrolled in 643 HMOs (InterStudy 1989). Managed-care enrollees grew to more than 50 million in 1996 and are expected to reach 100 million by the year 2002. Nearly 75 percent of U.S. workers with health insurance now receive that coverage through an HMO, a PPO, or a point of service plan (PSP) (McNamee 1997). There have been numerous rounds of mergers and acquisitions among HMOs, and some nonprofit HMO corporations have established profit-making operations (Gallagher 1999).

PPOs are modified HMOs that provide health care for lower costs when the enrollee uses participating providers who are paid on the basis of negotiated or discount rates (U.S. DOC 1990). In 1988 there were about 620 PPOs with about 36 million members.

Private health insurance companies constitute another large sector of the health industry. In 1988, the United States had over 1,000 for-profit, commercial health insurers and 85 Blue Cross/Blue Shield plans (Feldstein 1988). These private insurance organizations, along with HMOs, PPOs, and other third-party payers, paid for 32 percent (\$348 billion out of \$1,092 billion) of the total expenditures in 1997 (Srinivasan et al. 1998).

Physician practice patterns changed rapidly between the 1970s and 1990s, moving toward larger partnerships and group practices. Eighteen percent of physicians were in group practices (with three or more physicians) in 1969, compared to 28 percent in 1984 (Andersen and Mullner 1989). It is estimated that about 75 percent of all practicing physicians are part of at least one qualified health management organization (U.S. DOC 1990). Thus, physicians are moving toward larger and more complex forms of group practice. In addition, physicians are actively involved in the ownership and operation of many of the newer forms of HMOs, PPOs, IPAs, and other types of corporate health care activities (Relman 1980; Iglehart 1989).

*Horizontal Integration* The major changes in corporate arrangements have been the development of multiorganizational systems through horizontal integration. The formation of multihospital systems has grown tremendously within the industry. Ermann and Gable (1984) estimated there

**Community Hospitals and Beds by Ownership, 1980, 1990, and 1995**

Type of Ownership	1980		1990		1995	
	Hospitals	Beds	Hospitals	Beds	Hospitals	Beds
Nonprofit	3,322 (57%)	692 (70%)	3,191 (60%)	657 (71%)	3,092 (60%)	610 (70%)
Investor	730 (13%)	87 (9%)	749 (14%)	101 (11%)	752 (14%)	106 (12%)
State and local government	1,778 (30%)	209 (21%)	1,444 (27%)	169 (18%)	1,350 (26%)	157 (18%)
Total	5,830	988	5,384	927	5,194	873

Table 1

NOTE: Excludes federal psychiatric, tuberculosis, and other hospitals.

SOURCE: Adapted from American Hospital Association. *Hospital Statistics, 1989-1990 and 1996-1997 Editions*. Chicago: AHA, 1989, 1996.

were 202 multihospital systems controlling 1,405 hospitals and 293,000 beds in 1975 (or 24 percent of the hospitals and 31 percent of all beds). In 1997, there were 280 multihospital systems controlling 1,514 hospitals, and 543,588 beds (Table 2). This represents a 39 percent increase in the number of multihospital systems, a small (7 percent) increase in the number of hospitals, and an 86 percent increase in the number of beds between 1975 and 1997.

Multihospital corporations are becoming consolidated, with large companies controlling the largest share of the overall hospital market. Most of the recent increase in these systems has been the result of purchases or leases of existing facilities and mergers of organizations, rather than of construction of new facilities.

*Vertical Integration* Vertical integration involves the development of organizations with different levels and types of organizations and services. One such type of integration has involved the linkage of hospitals and health maintenance organizations and/or insurance companies. For example, National Medical Enterprises owned hospitals, nursing homes, psychiatric hospitals, recovery centers, and rehabilitation hospitals (Federation of American Health Systems 1990). There has also been an increase in the number of academic medical center hospitals that have relationships with proprietary hospital firms (Howard S. Berliner and Burlage 1990, p. 97). Many of the major investor-owned

health care corporations are diversified, with many different types of health care operations.

*Changes in Ownership* Between the 1970s and the 1990s the organizational side of health care witnessed a surge in the growth of both for-profit and not-for-profit health care delivery corporations, initially in hospitals and later extending to other types of health organizations. The ownership of hospitals shifted from public to nonprofit and for-profit organizations (see Table 1). The percentage of government-owned community hospitals dropped from 30 percent of the total community hospitals in 1980 to 26 percent in 1995, and the percentage of total beds declined from 21 percent to 18 percent during the same period (AHA 1996). In contrast, the percentage of proprietary facilities increased from 13 percent to 14 percent, and the percentage of proprietary beds increased from 9 percent to 12 percent, of the total during the 1980-1995 period. The percentage of total U.S. hospitals owned by nonprofit corporations increased from 57 percent to 60 percent during the period, while the percentage of beds remained at 70 percent (AHA 1996). Of the total 280 multihospital systems in 1997 (down from 303 systems in 1988), investor-owned systems controlled 40 percent of the hospitals and 27 percent of the beds, compared to nonprofit facilities (AHA 1997). The federal government controlled 9.8 percent of hospitals and 13 percent of beds, while nonprofit organizations controlled 50 percent of hospitals and 59 percent of beds (AHA 1997). We note that this represents a 3 percent decrease in the number

**Hospitals and Beds in Multihospital Health Care Systems, by Type of Ownership and Control, 1997**

Type of Ownership	TOTAL NOT-FOR-PROFIT		INVESTOR-OWNED		ALL SYSTEMS	
	Hospitals	Beds	Hospitals	Beds	Hospitals	Beds
Owned, leased, or sponsored	1,343 (53%)	309,216 (62%)	887 (35%)	119,466 (24%)	2,525 (100%)	501,724 (100%)
Contract-managed	171 (36%)	12,901 (31%)	301 (64%)	28,963 (69%)	472 (100%)	41,864 (100%)
Total	1,514 (50%)	322,117 (59%)	1,188 (40%)	148,429 (27%)	2,997 (100%)	543,588 (100%)

Table 2

SOURCE: Adapted from American Hospital Association. *Guide to the Health Care Field. 1997-1998 Edition*. Chicago: AHA, 1997, Table B3.

of hospitals and a 6 percent decrease in the number of beds in multihospital systems controlled by investor-owned systems since 1988.

Nursing homes have the largest share of proprietary ownership in the health field (except for the drug and medical supply industries). In 1997, some 65 percent of all nursing homes were profit-making, 28 percent were nonprofit, and 7 percent were government-run (Harrington et al. 1999). By 1997, chains owned 54 percent of the total nursing home facilities.

For-profit companies dominate the health maintenance organization (HMO) market. Between 1981 and 1997, for-profit HMOs grew from representing 12 percent to 62 percent of total HMO enrollees and from 18 percent to 75 percent of health plans (Srinivasan et al. 1998). Investor-owned corporations have also established themselves in many other areas of health care, ranging from primary-care clinics to specialized referral centers and home health care. The number of proprietary home health corporations is increasing rapidly, while the number of traditional visiting nurse associations is declining (Estes et al. 1992). In 1982 it was estimated that 14 percent of the Medicare home health charges were by proprietary agencies, 26 percent by nonprofit organizations, 32 percent by visiting nurse associations, 15 percent by facility-based agencies, and 14 percent by other agencies (U.S. Department of Health and Human Services 1989). By 1996, proprietary agencies accounted for 44 percent of total Medicare agencies, nonprofit care for 37 percent, and government and others for 19 percent (U.S. DHHS 1997).

Forty-four percent of home health agencies were part of a multifacility chain. This represents a dramatic shift in ownership structure within a six-year period. The changes brought about by the for-profit chains are more extensive than their proportionate representation among health care providers might suggest (Bergthold et al. 1990; Estes and Swan 1994). By force of example and direct competition, for-profit chains have encouraged many nonprofit hospitals and other health entities to combine into chains and convert to for-profit status (Dube 1999).

**Diversification, Restructuring, and Growth.**

Diversification of health care corporations is continuing to occur. Some large hospital corporations have developed ambulatory care centers (such as Humana, which later sold its centers), while others have developed their own HMOs or insurance. By the mid-1980s, many experts expected America's health care system to be dominated by the four largest for-profit hospital chains: Hospital Corporations of America (HCA), Humana, National Medical Enterprises, and American Medical International. By the late 1990s, only Humana and HCA were left standing, and HCA had already merged with Columbia. Eventually HCA Columbia almost collapsed as the result of a scandal over fraud during the late 1990s (*Multinational Monitor* 1998). Economic problems in the late 1980s resulted in some industry restructuring, by scaling down operations and spinning off substantial segments (Ginzberg 1988).

In the 1990s, this cycle repeated itself, as the frenzy of mergers and acquisitions has produced

ever-greater desires for cost cutting and restructuring. One report states that much of this drive for cost containment stemmed from drug companies' raising prices and from an increase in patient visits to doctors (utilization costs) (Hayes 1997). High stock values and the desire to improve market share have catalyzed many health care firms to seek growth through mergers and acquisitions. The number of mergers among health services (483) and HMO companies (33) peaked in 1996. These mergers were valued at \$27 billion and \$13.3 billion, respectively (Hayes 1997).

As HMOs grow, it has become clear that their primary goals are market control, profit making, and cost containment (through capitation and other mechanisms). For example, the merger of Aetna and U.S. Health care in 1996 put a "corporate giant in control of the care of 1 in 12 people in the United States" (Slaughter 1997, p. 22). In the years 1994–1999 a dozen companies were merged or acquired by six of the biggest firms: Aetna, Cigna, United HealthCare, Foundation Health Systems, Pacificare, and Wellpoint Health Networks.

As managed-care organizations became the dominant player in the health care industry in the 1990s, both doctors and patients began to voice complaints about the system. Many patients felt that they were no longer able to receive the quality time and personal care of a primary physician, because the physicians had to provide hurried treatment to patients in order to maintain efficiencies demanded by HMOs (Managed Care Improvement Task Force 1998). For their part, many doctors argued that capitation and other structures introduced by HMOs limited their freedom to make treatment decisions.

Many doctors and patients argued that these trends were producing lower-quality care (Dao 1999, p. A1). Additionally, registered nurses were increasingly being used in place of doctors to lower labor costs. Registered nurses, in turn, were also being replaced with less skilled and lower-paid personnel. In response, both doctors and nurses have begun a fervent effort to unionize (Slaughter 1997). While some critics argue that the impacts of this growth at all costs by both for-profit and nonprofit organizations are often devastating to communities (Bond and Weissman 1997; Kassirer 1997), others find few negative effects (Fubini and Limb 1997).

**Financial Status and Profits.** The private health care sector was marked by great volatility and growth in the 1990s. *Forbes's* annual report on investor-owned health corporations shows that the median five-year average return-on-equity for health corporations was 14.6 percent, well above the 10.5 percent for all U.S. industries (Condon 1998) (see Table 3). Median health industry sales for investor-owned companies grew 8.8 percent for 1997 and at a 11.1 percent rate for the five-year average. Earnings per share were 15.5 percent in the most recent twelve months, compared with 8.6 percent for the five-year average. The earnings per share were higher than the 14.9 percent earnings for all U.S. industries in the most recent twelve months in 1997 (Condon 1998).

The *Forbes* financial reports for the largest health corporations are shown in Table 3 for three different sectors of the industry: health care services, drugs, and medical supply companies (Condon 1998; Hayes 1998). The most profitable health care service corporation in 1989 was Humana, which owns both hospitals and insurance companies. In 1989 its group health insurance division had almost 1 million members and a \$4 billion operating profit (Fritz 1990). The most profitable health care service corporation in 1997 was HBO and Company, while Oxford Health Plans had the strongest five-year average.

While large investor-owned HMOs are growing each year, the 1990s were tumultuous financially. As *Forbes* Annual Report on American Industry put it, "Health care providers are supposed to make people well, but many of these companies are very sick themselves" (Hayes 1998, p. 180). Although Oxford Health Plans had the higher five-year average, between July 1997 and January 1998 its stock lost over 80 percent of its value. Similarly, the number two-ranked company, Mid Atlantic Medical Services, saw its stock drop by more than 50 percent of its value in a year's time. Rapid growth, through mergers, acquisitions, and internal sales, "eventually outstripped management's ability to run these companies" (Hayes 1998, p. 180).

Earnings per share of drug companies were at 15.8 percent in 1997, which was up from the five-year average of 11.9 percent (Condon 1998). Return on equity reported for drug companies was at 11.7 percent in 1997, significantly less than the 14.3 percent on average over the previous five

MEDICAL-INDUSTRIAL COMPLEX

**Selected U.S. Health-Care Investor Corporations, 1998**

Company	RETURN ON CAPITAL		PROFITABILITY GROWTH				NET INCOME
	5-Year Average (%)	Latest 12 Months (%)	SALES		EARNINGS PER SHARE		Latest 12 Months \$ million
			5-Year Average (%)	Latest 12 Months (%)	5-Year Average (%)	Latest 12 Months (%)	
<b>Health Care Services:</b>							
Oxford Health Plans	36.5	4.2	116.0	46.2	74.7	-72.0	25
Mid Atlantic	34.4	2.7	21.7	2.9	NM	-67.6	5
Health Management	20.8	22.3	25.0	25.4	33.2	27.6	108
WellPoint Health	19.8	15.0	15.1	43.0	6.0	17.2	203
Sun Healthcare	18.9*	7.4	89.9	28.2	NM	D-P	49
PacifiCare Health	18.3	18.2	33.7	69.5	8.9	37.0	124
United HealthCare	17.4	10.0	62.2	19.0	24.2	64.1	436
HBO & Co	17.0	23.6	35.8	43.0	NM	63.6	145
Quorum Health	14.5	11.3	50.3	28.4	32.1	20.0	88
Humana	12.3	11.0	22.2	15.3	32.7	270.8	147
<b>Medians</b>	<b>10.4</b>	<b>7.4</b>	<b>24.9</b>	<b>22.3</b>	<b>6.0</b>	<b>23.5</b>	<b>52</b>
<b>Drugs:</b>							
Schering-Plough	49.6	58.1	8.9	15.9	15.1	17.5	1,378
Abbott Labs	38.1	35.2	9.5	10.8	13.3	12.8	2,038
Amgen	35.4	29.6	25.1	10.6	37.4	0.9	643
Bristol-Myers Squibb	34.9	41.8	6.4	10.2	3.6	57.4	3,119
Warner-Lambert	32.1	19.7	7.6	6.2	32.5	8.6	805
<b>Medians</b>	<b>14.3</b>	<b>11.7</b>	<b>14.7</b>	<b>10.7</b>	<b>11.9</b>	<b>15.8</b>	<b>198</b>
<b>Medical Supplies:</b>							
Medtronic	28.1	31.5	16.5	8.7	27.2	15.5	564
Johnson & Johnson	26.5	26.2	11.5	7.3	14.6	14.7	3,229
Patterson Dental	24.4	20.4	18.9	19.9	30.6	24.3	37
Stryker	20.1	16.2	20.7	8.8	24.2	21.0	119
Perkin-Elmer	19.7	26.9	5.4	10.5	13.6	271.9	107
<b>Medians</b>	<b>14.6</b>	<b>13.4</b>	<b>8.7</b>	<b>5.9</b>	<b>4.7</b>	<b>15.5</b>	<b>79</b>
Industry Medians	14.6	12.7	11.1	8.8	8.6	15.5	107
All Industry Medians	10.5	10.3	8.9	7.9	6.7	14.9	96

Table 3

NOTE: : D-P: Deficit to profit. NM: Not meaningful. \*Four-year average.

SOURCE: Adapted from *Forbes*, January 12, 1998, pp. 176-182.

years (Condon 1997). On the other hand, earnings per share of medical supply companies were doing well at 15.5 percent in 1997, far better than their five-year median earnings of 4.7 percent. In 1989 a number of large drug company mergers occurred, particularly between U.S. firms and foreign corporations such as Genentech, Inc., and Roche Holding, Ltd., of Switzerland (Southwick 1990). These international mergers continued into the 1990s.

Although the biotechnology industry did not show overall profits in 1989, the sales growth rates were strong, and some companies had high profit rates, such as Diagnostic Products, with a 22.3 percent earnings per share and 23.6 percent return on equity in 1989 over the previous year (Clements 1990, p. 182). Biotechnology saw an upsurge of economic growth and media coverage when several new developments emerged in the 1990s. The first was the introduction of “gene therapies” whereby scientists could modify a person’s genetic makeup to fight against otherwise deadly or incurable diseases such as cancer and Alzheimer’s. The second was the launching of the Human Genome Project, a massive effort by scientists in government and industry to “map” the structure of the human genetic code. The third, was the announcement, by a Scottish scientist in 1997, that he had successfully cloned a sheep from another sheep’s DNA.

While all these developments signaled the importance of biotechnology in the future of health care research and policy, some observers were critical of these technological “advances.” First, the cloning of animals like sheep leaves open the distinct possibility that human beings might soon be cloned. The prospect of this event raised dire concerns among bio-ethicists, politicians, scientists, and religious leaders during 1997, and President Bill Clinton issued a worldwide call asking scientists to voluntarily refrain from any such activities. Second, some critics have noted that the Human Genome Project has been associated with efforts to “locate” genes believed not only to be the cause of certain diseases like breast cancer, but also those genes believed to be associated with certain types of deviant, or criminal, behavior. Focusing on the genetic “causes” of certain diseases and social behaviors raises many problematic scenarios for public policy (for example, defining breast cancer as genetically based, rather than

being rooted in social structures and the production of toxins by industry). Third, some biotechnology companies have been associated with efforts to patent life forms around the globe, including parts of the human body—prompting some critics to label this practice “biopiracy” or “biocolonization” (Kimbrell 1996). Taken together, these charges suggest that, through biotech, the medical-industrial complex is charting revolutionary territory that has allowed private interests to define, claim ownership over, and even create life on this planet. Despite these criticisms, biotechnology stocks continue to rise. As noted, the pharmaceutical industry alone traded upward of \$110 billion globally in 1997, while the overall health expenditures topped \$1 trillion by the end of the 1990s. Some analysts project health costs to more than double by 2015 to \$2.3 trillion, of which the government share will be between 25 and 50 percent (Pardes et al. 1999).

In summary, the 1970s, 1980s, and 1990s were decades of enormous growth in health care spending and the rationalization of health care service delivery, with the formation of large, complex, bureaucratically interconnected units and arrangements that reached well beyond the hospital and permeated virtually all sectors of the health care industry. At the same time, new sectors emerged (e.g., genetic research and subacute care), bringing additional industry developments. This vertical and horizontal integration of medical organizations and industries, combined with the revival of market ideologies and government policies promoting competition and deregulation, have profoundly altered the shape of U.S. health care delivery. As we enter the new millennium, these changes continue to signal a fundamental transformation of American medicine and a rationalization of the system under private control that was described by Paul Starr (1982).

**The Need for Regulation.** The federal government has been playing and continues to play a crucial role in the development of the medical-industrial complex. After World War II, the federal role expanded as Congress enacted legislation and authorized money for research, education, training, and the financing of health services. The passage of Medicare and Medicaid in 1965 was pivotal in expanding the medical-industrial complex, as government became the third-party payer for health care services (Estes et al. 1984). As a consequence, public demand for health care among



the aged, blind, disabled, and poor (all previously limited in access) was secure. Medicare and Medicaid provided the major sources of long-term capital financing for hospitals and contributed to the marked increase in service volume and technology, as well as to the continued oversupply of physicians (McKinlay and Stoeckle 1994). Thus, federal financing of health care has performed the very important functions of sustaining aggregate demand through health insurance programs, protecting against financial risks, subsidizing research and guaranteeing substantial financial returns, supporting the system's infrastructure through training subsidies and capital expansion, and regulating competition through licensure and accreditation (LeRoy 1979).

In addition to government spending, third-party insurance offered by Blue Cross/Blue Shield and private commercial companies covered most of the remaining inpatient hospital expenditures and a significant proportion of physician costs. The cost-based service reimbursement by private insurers, Blue Cross, and Medicare created and sustained strong cash flows in the hospital industry (Ginzberg 1988). With public and private sector third-party payments covering 90 percent of all inpatient hospital expenditures, the hospital business had become virtually riskfree.

In the 1980s and 1990s, two other forces were responsible for the dramatic changes in the medical-industrial complex: a change in the ideological climate with the election of President Ronald Reagan, President George Bush, and later President Bill Clinton, and changes in state policies to promote privatization, rationalization, and competition in health care (Estes 1990). These changes contributed to increases in the proportion of services provided by proprietary institutions (Schlesinger et al. 1987).

While policies of the 1960s and 1970s encouraged a form of privatization built on the voluntary sector (Estes and Bergthold 1988), President Reagan, President Bush, and President Clinton shifted the direction and accelerated privatization. In the 1980s and 1990s, the form of privatization was government subsidy of a growing proportion of for-profit (rather than nonprofit) enterprises (Bergthold et al. 1990). There was also privatization in the form of a transfer of work from the formal sector of the hospital to the informal sector of

home and family with ambulatory surgery and shortened lengths of hospital stays (Binney et al. 1993). Regulatory and legislative devices were important in stimulating and accelerating privatization in the health and social services. The Omnibus Reconciliation Act of 1980 and the Omnibus Budget Reconciliation Act of 1981 contributed to competition and deregulation, private contracting, and growth of for-profits in service areas that were traditionally dominated by nonprofit or public providers (e.g. home health care).

President Clinton introduced a health care reform plan in 1993 that ultimately failed, giving way to a private sector-driven market reform managed care, promoting a system that many contend benefits investors over patients, doctors, and community hospitals (Andrews 1995). Given the long-term historical role of the private, nonprofit sector in U.S. health and social services since the earliest days of the republic and the rapid organizational changes of the 1980s and 1990s, vertical and horizontal integration have blurred boundaries between the nonprofit and for-profit health care sectors. For-profit entities have nonprofit subsidiaries, and vice versa, and conceptual and structural complexities have multiplied, rendering impossible the simple differentiation of public from private. It is noteworthy that government-initiated privatization strategies did not reduce public sector costs (see the section entitled "Financial Status and Profits").

The distinction between for-profit and nonprofit is less meaningful when both organizational forms appear to be pursuing greater revenues through cost cutting and mergers. Eight of the ten largest health care systems (by net patient revenues) in 1997 were nonprofits; that same year, four out of the ten largest health care systems (by number of hospitals owned) were also nonprofits (Bellandi and Jaspén 1998, p. 36). Whether not-for-profits are still oriented toward the needs of the community is unclear, as many of these organizations (both insurance plans and hospitals) are undergoing "conversions" to for-profit status (Marsteller et al. 1998).

The concern that many communities have is that these conversions may mean less attention to the health needs of local residents. The increase in conversions has "heightened the need for accountability regarding the accurate determination and

disposition of assets developed with the assistance of tax subsidies for nonprofit medical entities such as Blue Cross” (Estes and Linkins 1997, p. 436). Federal and state laws require that their assets remain in the charitable sector and continue to be used for the community’s benefit. Twenty-three states now have conversion laws clarifying the authority of attorneys general to regulate these conversions. Conversion laws mandate varying degrees of public participation and public disclosure, but often are best implemented when there is an active community-based activist presence to monitor the organization’s practices.

From social movements to the federal government, institutions across the nation are recognizing the increasing need for monitoring and regulation of the myriad branches of the medical-industrial complex. As noted earlier, physicians are advocating new legislation and even beginning to unionize against HMOs. Unfortunately, often when regulation exists, it is easily circumvented. For example, the Health Insurance Portability and Accountability Act of 1996 was intended to protect Americans who change or lose their jobs by assuring portability of plans across groups and into the individual market. It was also intended to protect people against denial of coverage for preexisting conditions. However, many insurers have skirted this law by denying commissions to their agents who sell insurance to people with medical problems (Pear 1997).

In other cases, there is little to no regulation of purchasers, such as large employers who self-insure under the Employment Retirement Income Security Act (ERISA) of 1974. These employers are exempt from state insurance laws and are bound by no federal regulation in this area. More than 125 million Americans who have HMO coverage cannot sue their providers for punitive damages. As *Time* magazine recently reported, this represents a “clear subordination to corporate interests” (Howe 1999, p. 46). Furthermore, the continuing rapid pace of mergers and acquisitions in the health care industry has created a consolidation of markets that raise questions about the need for antitrust policies directed at this sector.

At the federal level, the President’s Advisory Commission on Consumer Protection and Quality in the Health Care Industry has called for a “consumer bill of rights,” while others clamor for

“patient’s bill of rights.” These proposals would provide, for example, the right to sue HMOs for damages, prohibitions against negative financial incentives, and external reviews of patient complaints. Whether this proposal will ever become national legislation is unclear. But as, under the “New Federalism,” previously federal responsibilities are “devolved” to the states, it is certain that fifty separate governments will have great difficulty coherently regulating a health care system for the entire nation. This is especially troublesome because of the growing numbers of the uninsured and those in need of long-term care (Estes and Linkins 1997).

Long-term care (LTC) is an area of health policy in need of greater attention as the age distribution of the population changes, so that the number of persons over 65 and over 85 is increasing rapidly. Medicare, Social Security and other entitlement programs affecting the elderly have been the focus of the “devolution revolution” (Estes and Linkins 1997) and raise serious questions about the quality and accessibility of LTC under managed care. There are profits to be made in this sector of the industry as well. One publication referred to 1997 as the “year of assisted living” because seven of the top ten health care provider organizations were assisted-living companies and half of them posted returns of 50 percent or greater (*SeniorCare Investor* 1998). LTC will have to meet the “bottom line” criterion of “cost-saving” or “profit generation,” and this is likely to produce problems for those in need of these services. Indeed the needs of all less powerful groups—the elderly, the poor, the working and middle classes, women, and people of color—are increasingly being confronted by these business directives as well. Regulation of health care institutions in the interests of these marginalized groups is especially difficult when the American Medical Association is second to none in money spent on lobbying the Congress and state legislatures (Jaspen 1999).

**The U.S. Medical-Industrial Complex in a Global Context.** The U.S. health and health care systems rank near the bottom of all industrialized nations on a number of key dimensions. By comparison, the health of U.S. citizens is poorer and the number of underinsured and uninsured individuals is greater than in any other industrialized nation. Among the top 24 industrialized nations,

the United States ranks sixteenth in life expectancy for women, seventeenth for men, and twenty-first in infant mortality (Andrews 1995, p. 38). A Harris poll indicated that the citizens of Canada, western Europe, the United Kingdom, and Japan report much higher satisfaction with their health care systems than do Americans (Isaacson 1993). Additionally, compared with other nations, doctors in the United States receive much higher incomes relative to the average worker. For example, in 1987 the income ratio of doctors to the average worker in the United States was 5.4, while in Canada it was 3.7 and in Japan and the United Kingdom it was 2.4 (Isaacson 1993). This poor performance of the United States on health indicators is ironic, given that the United States spends more money on its health care system than any other nation in the world.

Multinational health enterprises are an increasingly important part of the medical-industrial complex, with investor-owned and investor-operated companies active not only in the United States but also in many foreign countries. In 1990 a report showed 97 companies reporting ownership or operation of 1,492 hospitals with 182,644 beds in the United States and 100 hospitals with 11,974 beds in foreign countries (FAHS 1990, pp 16–17). The four largest for-profit chains owned two-thirds of the foreign hospitals (Berliner and Regan 1990). Pharmaceutical firms have also become major global corporate players. In 1990, foreign control over pharmaceutical production was 72 percent in Australia, 61 percent in the United Kingdom, 57 percent in Italy, and 30 percent in the United States (Tarabusi and Vickery 1998). The total value of global pharmaceutical exports and imports is estimated to be in excess of \$110 billion (Tarabusi and Vickery 1998). The effects of these developments in foreign countries and the profit potential of these operations are not clearly understood (Berliner and Regan 1990).

Because of the pluralistically financed health care system in the United States, administrative costs are much higher than those of the national and publicly financed health care systems of virtually all other Western industrialized nations, with the exception of South Africa. U.S. health care expenditures were increasing at an alarming rate until around 1993, when the rate of growth in expenditures began to slow. However, this trend is expected to reverse itself, and one study projects

that “health spending is expected to rise as a share of gross domestic product (GDP) beginning in 1988, climbing from 13.6 percent in 1996 to an estimated 16.6 percent by 2007” (Smith et al. 1998, p. 128).

Examining the U.S. medical-industrial complex in a comparative context provides an understanding of the role of the welfare state and government vis-à-vis civil society and private capital. What has become clear is that the unique problems the U.S. medical-industrial complex has created are rooted in the subordination of the state and civil society to corporate interests. Other nations whose health care systems are much more effective are marked by the state’s taking an active role to restrict the profit motive in health insurance, “or they simply never let a commercial market develop” (Andrews 1995, p. 36). This is because voluntary insurance in many other countries historically preceded public legislation, and these insurance funds were linked to labor unions, political groups, and religious groups—not to private companies and health care providers, as in the United States. In these cases, government policy was, and remains, heavily influenced by nongovernmental organizations (NGOs), namely religious and labor groups. The medical-industrial complex in particular and corporate-civil society relations in general in the United States are much less democratic largely because of the lower levels of mobilization by trade unions and other NGOs. Government agencies in the United States can learn from other nations and begin to implement policies that leverage the power of the state and NGOs in ways that bring a greater balance among the stakeholders in the medical-industrial complex.

*Issues Raised by the Medical-Industrial Complex*  
*Commodification. Commercialization, proprietarization, and monetarization* are terms used to describe an increasingly salient dynamic in the medical-industrial complex: the potentially distorting effects of money, profit, and market rationality as a (if not *the*) central determining force in health care. After three decades devoted to market rhetoric, cost containment, and stunning organizational rationalization, the net result is the complete failure of any of these efforts to stem the swelling tide of problems of access and cost. For example, while national health care expenditures make up around 15 percent of the GDP, the number of uninsured

Americans was fully 43.4 million, or 16.1 percent of the population in 1997—the highest level in a decade (Kuttner 1999). Moreover, there are alarming increases in the uninsured population among African-Americans, Latinos, and the middle class (Carrasquillo et al. 1999). Of those Americans who do have insurance, a recent study found that the number of persons insured by the private sector is much less than previously believed. While many studies had estimated that 61 percent of the insured received coverage through the private sector, Carrasquillo and colleagues (1999) found that the public sector subsidizes much of this coverage so that, in fact, only 43 percent of the population receives insurance through the private sector. Thus, not only is the burden on the state greater than previously thought because of this subsidy to the private sector, but the general decline in private insurance coverage will also produce further strains on the government's budget.

The rapidly growing health care industry is creating strains on the economic system while it also is creating a financial burden on government, business, and individuals through their payments for health services. These strains are occurring simultaneously when, in 1999, a huge federal deficit has been turned into a surplus, an event of historic significance. The budget surplus has produced a combination of euphoria and vigorous debate over what the government should do with it. This surplus emerged against the backdrop of an unusually high economic growth rate and strong general U.S. economy as we enter the new millennium. Responses to these deficit strains and fluctuations have included cutbacks in services and reimbursements; cost shifts onto consumers; and alterations in the structure of the health care system itself to accord better with a competitive, for-profit model. The competition model as a prescription for the nation's health-care woes has restricted access to health care and raised questions of quality of care (Bond and Weissman 1997; Harrington 1996; Kassirer 1997). Cost shifting to consumers is increasingly limiting access to needed services for those with less ability to pay. Managed care has not delivered the cost savings it promised, and the Health Care Financing Administration acknowledges that Medicare does not benefit from cost reductions from HMO enrollment of elders due to continuing adverse risk selection (DePearl 1999).

The juxtaposition of the commercial ethos familiar in fast-food chains with health care collides with traditional images of medicine as the embodiment of humane service. Investor-owned health care enterprises have elicited a number of specific criticisms. It has been argued that commercial considerations can undermine the responsibility of doctors toward their patients; can lead to unnecessary tests and procedures; and, given other financial incentives, can lead to inadequate treatment. The interrelationships among physicians and the private health care sector, particularly for-profit corporations, raises many issues about the effects on quality of care and health care utilization and expenditures. Many have argued that the potential for abuse, exploitation, unethical practices, and disregard of fiduciary responsibilities to patients is pervasive (Iglehart 1989). Legislation has even been introduced in Congress that would prohibit physicians from referring patients to entities in which they hold a financial interest and from receiving compensation from entities to which they refer patients (Iglehart 1989). In the late 1990s several versions of a "patient's bill of rights" were considered at the state and federal levels of government.

Critics of for-profits argue that such ownership drives up the cost of health care, reduces quality, neglects teaching and research, and excludes those who cannot pay for treatment. Opponents of the market model for health care reflect diverse interests, including members of the medical profession seeking to preserve their professional autonomy, advocates for access to health care for the poor and uninsured, those concerned about the impact of profit seeking on quality of care, and many others. As government and business attempt to restrain health care spending, cutting into profits and forcing cost reductions, these concerns intensify.

The medical-industrial complex is an inherently fascinating topic for sociological analysis because it underscores many of the less obvious dimensions of health care. Profits, power, and market control are not terms that have traditionally come to mind when the average person in the United States talked about their health care provider (although this is changing as we observe the managed-care "backlash"). Yet, these are some of the primary goals of those organizations administering our medical care. Sociology itself arose as

an effort to wrestle with the myriad social impacts of nineteenth century industrialization (Durkheim [1893] 1984; Marx [1867] 1976). Early sociologists were especially concerned about how changes in communities were created by industrial production. Over the last century, sociologists have maintained a particular interest in the connection between advanced capitalism and the emergence of specific types of work, formal organizations, political systems, families, and cultural beliefs. An examination of the medical-industrial complex is therefore one of the more recent attempts to refocus core sociological questions about community, power, stratification, and social change.

Underlying this early sociological research on the impact of industrialization was an interest in power. How does power get accumulated and applied in a market-based society? How does it get distributed unevenly to social groups? To what ends do empowered social groups apply their power? One school of thought has been that industrial production skews power to an elite class (Domhoff 1998; Gramsci [1933] 1971). Max Weber's theory of bureaucracy, for example, suggested that this was an organizational form that pervades all social institutions. Its emphases on rationality, efficiency, predictability, calculability, and control (Weber [1921] 1961) were in part seen as a social advance over arbitrary religious, charismatic, and personalized forms of authority. Later analysts of bureaucracy, in contrast, saw this formalization of organizations as a dehumanizing and antisocial mechanism.

Issues for sociological investigation include the systematic identification of the ways in which the new commercial practices and organization of health care affect health care delivery. Organizational studies are needed to disentangle the effects of organizational characteristics (e.g., tax status and system affiliation) on the outcomes of equity, access, utilization, cost, and quality of care. The effects on provider-patient interactions of these structural and normative changes in health care require investigation as well. A general sociological theory of the professions will emerge from understanding the ways in which the dominant medical profession responds to the ongoing restructuring of health care and accompanying challenges to its ability to control the substance of its own work, erosions in its monopoly over medical knowledge, diminishing authority over patients

resulting from health policy changes, major technological and economic developments, and changes in the medical-industrial complex. Finally, sociologists must confront the coming biotechnological revolution and its impacts on society, human health, and the environment. The corporatization of health care and health-related research and the medical industrial complex are topics of great interest to scholars studying social movements, organization behavior, stratification, health and illness, and science and technology.

(SEE ALSO: *Health Care Financing; Health Policy Analysis; Health Services Utilization; Medical Sociology*)

REFERENCES

- American Hospital Association 1989 *Hospital Statistics, 1989-90*. Chicago: AHA.
- 1996 *Hospital Statistics, 1996-97*. Chicago: AHA.
- 1997 *AHA Guide to the Health Care Field 1997-1998*. Chicago: AHA.
- Andersen, Ronald M., and Ross M. Mullner 1989 "Trends in the Organization of Health Services." Pp. 144-165 in H. E. Freeman and S. Levine, eds., *Handbook of Medical Sociology*, 4th ed. Englewood Cliffs, N.J.: Prentice Hall.
- Andrews, Charles 1995 *Profit Fever: The Drive to Corporatize Health Care and How to Stop it*. Monroe, Me.: Common Courage.
- Bellandi, Deanna, and Bruce Jaspen 1998 "While You Weren't Sleeping." *Modern Healthcare* (May 25): 35-42.
- Bergthold, Linda A. 1990 "Business and the Pushcart Vendors in an Age of Supermarkets." In J. W. Salmon, ed., *The Corporate Transformation of Health Care: Issues and Directions*. Amityville, N.Y.: Baywood.
- , Carroll L. Estes, and A. Villanueva 1990 "Public Light and Private Dark: The Privatization of Home Health Services for the Elderly in the United States." *Home Health Services Quarterly* 11:7-33.
- Berliner, Howard S., and R. K. Burlage 1990 "Proprietary Hospital Chains and Academic Medical Centers." In J. W. Salmon, ed., *The Corporate Transformation of Health Care: Issues and Directions*. Amityville, N.Y.: Baywood.
- Berliner, Howard W., and C. Regan 1990 "Multi-National Operations of U.S. For Profit Hospital Chains: Trends and Implications." In J. W. Salmon, ed., *The Corporate Transformation of Health Care Issues and Directions*. Amityville, N.Y.: Baywood.

- Binney, Elizabeth A., Carroll L. Estes, and Susan E. Humphers 1993 "Informalization and Community Care for the Elderly." Pp. 155-170 in Carroll L. Estes, James Swan, and Associates, eds., *The Long Term Care Crisis*. Newbury Park, Calif.: Sage.
- Bond, Patrick, and Robert Weissman 1997 "The Costs of Mergers and Acquisitions in the U.S. Health Care Sector." *International Journal of Health Services* 27:77-87.
- Carrasquillo, Olveen, David Himmelstein, Steffie Woolhandler, and David Bor 1999 "Going Bare: Trends in Health Insurance Coverage, 1989-1996." *American Journal of Public Health* 89:36-42.
- Clements, J. 1990 "Insurance." *Forbes* (January 8):184-186.
- Condon, Bernard 1998 "Annual Report on American Industry: Health Care Products." *Forbes* (January 12):176-178.
- Dao, James 1999 "Concern Rising about Mergers in Health Plans." *New York Times* (January): A1.
- DePearl, Nancy Aun Min 1999 "The Future of Medicare," presentation at the Commonwealth Fund Conference on the Future of Medicare, Washington, D.C., February.
- Domhoff, G. William 1998 *Who Rules America? Power and Politics in the Year 2000*. Mountain View, Calif.: Mayfield.
- Dube, Monte 1999 "Lighten Your Load: In the Race to Compete, Public Hospitals Shed Excess Baggage." *Trustee* (January):17-19.
- Durkheim, Emile (1893) 1984 *The Division of Labor in Society*. New York: Free Press.
- Ehrenreich, Barbara, and John Ehrenreich 1971 *The American Health Empire: Power, Profits and Politics*. New York: Vintage.
- Ermann, Dan, and Jon Gabel 1984 "Multihospital Systems: Issues and Empirical Findings." *Health Affairs* 3:50-64.
- Estes, Carroll L. 1990 "The Reagan Legacy: Privatization, the Welfare State and Aging." In J. Quadagno and J. Myles, eds., *Aging and the Welfare State*. Philadelphia: Temple University Press.
- , and Linda A. Bergthold 1988 "The Unravelling of the Nonprofit Service Sector in the U.S." In J. I. Nelson, ed., *The Service Economy* (special issue of *International Journal of Sociology and Social Policy*) 9:18-33.
- , Lenore E. Gerard, Jane Sprague Zones, and James H. Swan 1984 *Political Economy, Health, and Aging*. Boston: Little, Brown.
- , and Karen Linkins 1997 "Devolution and Aging Policy: Racing to the Bottom in Long-Term Care." *International Journal of Health Services Research* 27:427-442.
- , and J. H. Swan 1994 "Privatization and Access to Home Health Care." *Milbank Quarterly* 72:277-298.
- , ———, and associates 1993 *The Long Term Care Crisis*. Newbury Park, Calif.: Sage.
- , ———, L.A. Bergthold, and P. Hanes-Spohn 1992 "Running as Fast as They Can: Organizational Changes in Home Health Care." *Home Health Care Services Quarterly* 13:35-69.
- Federation of American Health Systems 1990 *1990 Directory*. Little Rock, Ark.: FAHS.
- Feldstein, Paul J. 1988 *Health Care Economics*. New York: John Wiley.
- Fritz, M. 1990 "Health." *Forbes* (January 8):180-182.
- Fubini, Sylvia, and Stephanie Limb 1997 "The Ties that Bind." *Health Systems Review* (September/October):44-47.
- Gallagher, Leigh 1999 "The Big Money is in Small Towns." *Forbes* (January 11):182-183.
- Ginzberg, Eli 1988 "For Profit Medicine: A Reassessment." *New England Journal of Medicine* 319:757-761.
- Gramsci, Antonio (1933) 1971 *Selections from the Prison Notebooks*. New York: International.
- Harrington, Charlene 1996 "The Nursing Home Industry: Public Policy in the 1990s." In Phil Brown, ed., *Perspectives in Medical Sociology*, 2nd ed. Prospect Heights, Ill.: Waveland.
- , Helen Carrillo, Susan Thollaug, and Peter Summers 1999 *Nursing Facilities, Staffing, Residents, and Facility Deficiencies, 1991-1997*. San Francisco: University of California Press.
- , J. Swan, Carrie Griffin, et al. 1998 *1996 State Data Book on Long Term Care Programs and Market Characteristics*. San Francisco: University of California Press.
- Hayes, John 1997 "Annual Report on American Industry." *Forbes* (January 13):166-168.
- , 1998 "Annual Report on American Industry." *Forbes* (January 12):176-182.
- Himmelstein, David U., and Steffie Woolhandler 1986 "Cost without Benefit: Administrative Waste in the U.S." *New England Journal of Medicine* 314:440-441.
- , 1990 "The Corporate Compromise: A Marxist View of Health Policy." *Monthly Review* (May):14-29.
- Howe, Robert 1999 "The People vs. HMOs." *Time* (February 6):46-47.

- Iglehart, John K. 1989 "The Debate over Physician Ownership of Health Care Facilities." *New England Journal of Medicine* 321:198-204.
- 1999 "The American Health Care System: Expenditures." *New England Journal of Medicine* 340:70-76.
- InterStudy 1989 "Findings on Open-Ended HMOs Reports by InterStudy." *InterStudy Press Release*, March 7. Excelsior, Md.: InterStudy.
- Isaacson, Elisa 1993 "Prescription for Change." *San Francisco Bay Guardian* (April 14).
- Jaspen, Bruce 1999 "AMA Firing Impolitic—or Just Politics?" *Chicago Tribune* 5 (January 31):1.
- Kassirer, Jerome 1997 "Mergers and Acquisitions—Who Benefits? Who Loses?" *New England Journal of Medicine* 334:722-723.
- Kimbrell, Andrew 1996 "Biocolonization: The Patenting of Life and the Global Market in Body Parts." In Jerry Mander and Edward Goldsmith, eds., *The Case against the Global Economy and for a Turn Toward the Local*. San Francisco: Sierra Club.
- Kuttner, Robert 1999 "The American Health Care System: Health Insurance Coverage." *New England Journal of Medicine* 340:163-168.
- LeRoy, Lauren 1979 "The Political Economy of U.S. Federal Health Policy: A Closer Look at Medicare," unpublished manuscript. University of California, San Francisco.
- McKinlay, John B., and John D. Stoeckle 1994 "Corporatization and the Social Transformation of Doctoring." In Peter Conrad and Rochelle Kern, eds., *The Sociology of Health and Illness: Critical Perspectives*. New York: St. Martin's.
- McNamee, Mike 1997 "Health-Care Inflation: It's Baaack!" *Business Week* (March 17):28-30.
- Managed Health Care Improvement Task Force 1998 "Public Perceptions and Experiences with Managed Care: Background Paper." (January).
- Marsteller, Jill, Randall Bovbjerg, and Len Nichols 1998 "Nonprofit Conversion: Theory, Evidence, and State Policy Options." *Health Services Research* 33:5.
- Marx, Karl (1867) 1976 *Capital: A Critique of Political Economy*, vol 1. Middlesex, England: Penguin.
- Multinational Monitor* 1998 "Blowing the Whistle on Columbia/HCA: An Interview with Marc Gardner." *Multinational Monitor* (April):17-20.
- Navarro, Vicente 1976 *Medicine under Capitalism*. New York: Prodist.
- 1982 "The Labor Process and Health: A Historical Materialist Interpretation." *International Journal of Health Services* 12:5-29.
- 1995 "Why Congress Did Not Enact Health Care Reform." *Journal of Health Politics, Policy and Law* 20:455-461.
- Pardes, Herbert, Kenneth G. Manton, Eric S. Lander, H. Dennis Tolley, Arthus D. Ullian, and Hans Palmer 1999 "Effects of Medical Research on Health Care and the Economy." *Science* (January):36-37.
- Pear, Robert 1997 "Health Insurers Skirting New Law, Officials Report." *New York Times* (October 5, 1997):A1.
- Relman, Arnold S. 1980 "The New Medical-Industrial Complex." *New England Journal of Medicine* 303:963-970.
- Rodberg, L., and G. Stevenson 1977 "The Health Care Industry in Advanced Capitalism." *Review of Radical Political Economics* 9:104-115.
- Schlesinger, Mark, Theodore R. Marmor, and Richard Smithey 1987 "Nonprofit and For-Profit Medical Care: Shifting Roles and Implications for Health Policy." *Journal of Health Politics, Policy and Law* 12(3):427-457.
- SeniorCare Investor* 1998 "The Public Market." 10:1.
- Shadle, M., and M. M. Hunter 1988 *National HMO Firms 1988*. Excelsior, Minn.: InterStudy.
- Slaughter, Jane 1997 "Doctors Unite: Corporate Medicine and the Surprising Trend of Doctor Unionization." *Multinational Monitor* (November):22-24.
- Smith, Sheila, Mark Freeland, Stephen Heffler, David McKusick, and the Health Expenditures Projection Team 1998 "The Next Ten Years of Health Spending: What Does the Future Hold?" *Health Affairs* 17:5.
- Southwick, Karen 1990 "More Merger Mania among Drugmakers." *Healthweek* 4:1-51.
- Srinivasan, Srija, Larry Levitt, and Janet Lundy 1998 "Wall Street's Love Affair with Health Care." *Health Affairs*. 17:4.
- Starr, Paul 1982 *The Social Transformation of American Medicine*. New York: Basic Books.
- Sussman, David 1990 "HMOs Are Still Riding a Wave of Profitability." *Healthweek* 4:12.
- Tarabusi, Claudio Casadio, and Graham Vickery 1998 "Globalization of the Pharmaceutical Industry, Parts I and II." *International Journal of Health Services* 28:67-105, 281-303.
- U.S. Department of Commerce, International Trade Administration 1990 "Health and Medical Services." *U.S. Industrial Outlook 1990*. Washington, D.C.: U.S. DOC.
- U.S. Department of Health and Human Services 1989 *Health United States, 1989*, DHHS 90-1232. Hyattsville, Md.: U.S. DHHS.

- 1997 *Home Health Agencies*. Health Care Financing Administration, unpublished data. Baltimore, Md.: U.S. DHHS.
- U.S. National Center for Health Statistics [U.S. NCHS], E. Hing, E. Sekscenski, and G. Strahan 1989 "National Nursing Home Survey: 1985 Summary for the United States." *Vital and Health Statistics Series 18*. Hyattsville, Md.: Public Health Service.
- U.S. Office of National Cost Estimates (U.S. ONCE) 1990 "National Health Expenditures, 1988." *Health Care Financing Review* 11:1-41.
- Weber, Max (1921) 1961 *General Economic History*. New York: Collier.
- Wohl, Stanley 1984 *The Medical Industrial Complex*. New York: Harmony.

CARROLL L. ESTES  
CHARLENE HARRINGTON  
DAVID N. PELLOW

## MENTAL HEALTH

See Positive Mental Health.

## MENTAL ILLNESS AND MENTAL DISORDERS

After years of empirical research and theoretical activity, social scientists still do not agree about what mental illness actually is, let alone about what its primary causes are or about the efficacy of various treatments. Sociologists disagree about whether or not mental disorder is truly a disease that some people have and other people do not have, thus fitting a medical model of health and illness. They disagree about the relative importance of genetics, biochemical abnormalities, personality characteristics, and stress in the onset and course of psychiatric impairment. Most sociologists do, however, agree that definitions of mental illness are shaped by the historical, cultural, and interpersonal contexts within which they occur. They argue that the significance of any particular set of psychological or behavioral symptoms to a diagnosis of mental disorder lies in part with the actor and in part with the audience. Given this understanding of mental illness, sociologists are often as interested in understanding the consequences of being labeled mentally ill as they are in understanding the causes. Sociologists do, indeed,

study the social distribution and determinants of mental disorder. However, they also study social reactions to mental illness and the mentally ill and investigate ways in which mental health professionals and institutions can come to serve as agents of social control.

### CLASSIFICATION AND DIAGNOSIS

Although psychiatrists themselves have difficulty defining mental illness, the official system for classifying and diagnosing mental disorder in the United States is produced by the American Psychiatric Association (APA). It is known as the *Diagnostic and Statistical Manual of Mental Disorder* (DSM) and was first published in 1952. In its earliest form, DSM-I included a list of 60 separate mental illnesses. By the second edition in 1968, psychiatric definitions of mental illness had changed so markedly that 145 different types of mental disorder were included. Despite the attempt in DSM-II to define the parameters of mental illness more precisely, critics from inside and outside psychiatry pointed out that diagnoses of mental disorder were extremely unreliable. When different psychiatrists independently used DSM-II to diagnose the same patients, they did so with substantially different results. Studies conducted during the 1960s and 1970s indicated that there was poor agreement about what disease classification was appropriate for any given patient; studies also found that clinicians had difficulty in differentiating normal persons from mental patients and that they frequently disagreed about prognosis and the clinical significance of particular symptom patterns (Loring and Powell 1988).

After years of debate, some of which was quite heated, the APA published a third edition of DSM in 1980. *Mental disorder* was defined in DSM-III as "a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is associated with either a painful symptom (distress) or impairment in one or more areas of functioning (disability)" (APA 1980, p. 6); deviant behaviors and conflicts between individuals and society were specifically excluded from this definition *unless* they were symptoms of another diagnosable disorder. DSM-III took a purely descriptive approach to diagnosis, outlining the essential and associated features of each disorder but making no attempt to explain the etiology of



either symptoms or illnesses. At the time of the most recent revision (DSM-IV) in 1994, there was considerable debate about whether or not to actually retain the phrase “mental disorder” in the title of the DSM. The argument was that doing so implied a false distinction between mind and body, and between mental disorders and physical or general medical conditions. For lack of a better term, however, DSM-IV retained the same terminology and definition of mental disorder that had been used in DSM-III. DSM-IV contains hundreds of mental diagnoses, including such disorders as caffeine intoxication, circadian rhythm sleep disorder, and hypoactive sexual desire disorder. It also contains a section on “other conditions that may be a focus of clinical attention” but that are not mental disorders themselves. Included here are such conditions as relational problems connected with a family member’s mental disorder, noncompliance with treatment, religious or spiritual, academic, occupational, acculturative, and phase-of-life problems.

The use of DSM-III and DSM-IV criteria has vastly improved the overall reliability of psychiatric diagnoses, thereby enabling psychiatry to meet one of the major criticisms of the medical model of mental disorder. The inclusion of more and more categories of illness in each succeeding version of DSM has led to more precise and consequently more reliable diagnoses. However, some scholars have argued that this expansion of mental diagnoses has less to do with problems of disease classification than with “problems” of third-party reimbursement (Kirk and Kutichins 1992; Mirowsky and Ross 1989). Each increase in the number of disorders listed in DSM has increased the scope of psychiatric practice. As the number of patients with recognized illnesses increases, so too does the amount of compensation that psychiatrists receive from insurance companies.

Even the firmest supporters of DSM-IV recognize that the classification of mental disorder is influenced by nonmedical considerations. In fact, DSM-IV itself includes a discussion of specific culture, age, and gender features which should be taken into account for each diagnosis. Pressures from outside psychiatry also influence diagnostic classifications. In order to reduce their payment liabilities, insurance companies have lobbied the APA to reduce the number of diagnoses. Changes in public attitudes toward sexual preference issues

led, in 1974, to dropping homosexuality from the list of mental disorders, and veterans’ groups successfully pressed for the inclusion of posttraumatic stress syndrome (Scott 1990). The storm of controversy that surrounds the issue of whether premenstrual syndrome is a medical or a psychological condition, or whether it is socially unacceptable behavior, is another example of the intersection of political, social, economic, and diagnostic concerns (Figert 1994); at the insistence of feminists inside and outside psychiatry, premenstrual dysphoric disorder has been relegated to an appendix of DSM-IV on the basis that there is currently little scholarly evidence to support such a diagnosis and that its social implications dangerously feed “the prejudice that women’s hormones are a cause of mental illness” (Tavris 1993, p. 172). In sum, there is a less than perfect correspondence between some disease-producing entity or syndrome and the diagnosis of mental disorder; psychiatric diagnosis is based partly in the reality of disordered behavior and emotional pain, and partly in the evaluations that society makes of that behavior and pain. Thus, questions about the validity of psychiatric diagnosis are as troubling for DSM-IV as they were for DSM-I. As one observer has noted, “We have learned how to make reliable diagnoses, but we still have no adequate criterion of their validity” (Kendell 1988, p. 374). Given the problems scholars have in defining “mental disorder” and given the validity problems that ensue, it is not surprising that epidemiologists have used a number of different strategies to estimate rates of psychiatric impairment. These different research methodologies often have led to quite different interpretations of the role of social factors in the etiology of mental illness.

## MEASUREMENT

The earliest sociological research on mental disorder relied on data from individuals receiving psychiatric care. In a classic epidemiological study, Faris and Dunham (1939) reviewed the records of all patients admitted to Chicago’s public and private mental hospitals between 1922 and 1934. They found that admission rates for psychosis were highest among individuals living in the inner city. Several years later, researchers used a similar design to study the social class distribution of mental disorder in New Haven. In contrast to the

Chicago study, which focused only on individuals who had been hospitalized, Hollingshead and Redlich (1958) included individuals receiving outpatient care from private psychiatrists in their study. Results from the New Haven study confirmed the earlier findings; the lower the social class, the higher the rate of mental disorder. More recent studies have also used information on treated populations. Studies of patient populations provide useful information to be sure; findings shed light on the social factors that influence the course of mental health treatment. Individuals receiving treatment for mental disorder, however, are not a random subset of the population of individuals experiencing psychological distress. Everybody who has potentially diagnosable mental disorder does not receive treatment. Furthermore, pathways to mental health care may be systematically different for individuals with different social characteristics. Consequently, research based on treated rates of mental disorder seriously underestimates the true rate of mental illness in a population. (In 1994 the Institute of Medicine Committee on the Prevention of Mental Disorders estimated, for instance, that only 10–30 percent of those with a mental disorder receive any treatment.) Furthermore, such research may confuse the effects of variables such as social class, gender, place of residence, and age on psychiatric treatment with the impact of those same variables on the development of psychiatric impairment.

An alternative strategy for studying the epidemiology of mental disorder is the community survey. Early studies such as the Midtown Manhattan study (Srole et al. 1962) used symptom checklists with large random samples to estimate the amount of psychiatric impairment in the general population. Although such studies provided less biased estimates of the prevalence of psychological distress than did research on patient populations, they were subject to a different set of criticisms. The most serious limitation of the early community studies was that they used impairment scales that measured global mental health. Not only did the scales fail to distinguish different types of disorders, they confounded symptoms of physical and psychological disorder (Crandall and Dohrenwend 1967), measured relatively minor forms of psychiatric impairment, and frequently failed to identify the most serious forms of mental illness (Dohrenwend and Crandall 1970). Since it

was not clear what relationship psychological symptom scales bore to cases of actual psychiatric disorder, it was also not clear how results from those studies contributed to an understanding of the social causes of mental illness.

Since the early 1980s, symptom scales that measure specific forms of impairment have largely replaced the early global scales with the consequence that the reliability and validity of community survey research has been vastly improved. The CES-D, for instance, is a twenty-item depression scale that can accurately distinguish clinical from normal populations and depression from other psychiatric diagnoses (Weissman et al. 1977). Consistently with the overall improvement in diagnostic reliability that has accompanied the development of DSM-III and DSM-IV, methods have been developed that provide reliable psychiatric diagnoses of many disorders among community residents. The most widely used diagnostic instrument of this sort was developed by a team of researchers at Washington University as part of the National Institute of Mental Health (NIMH) Division of Biometry and Epidemiology's Catchment Area Program. Called the Diagnostic Interview Schedule (DIS), the instrument can be administered by nonpsychiatrists doing interviews with the general population. Using DSM-IV criteria, it provides both current and lifetime diagnoses for many adult psychiatric disorders (Eaton et al. 1985). Enormous amounts of time and money have been devoted to the development of the DIS, and research that makes use of it promises to provide a vital link between studies of clinical and community populations. Nevertheless, even instruments like the DIS have shortcomings.

Mirowsky and Ross (1989) have challenged the DIS and the DSM upon which it is based on the grounds that psychiatric diagnosis is a weak form of measurement and that it is of questionable validity. These authors claim that psychiatric disorders are dimensional, not categorical. By collapsing a pattern of symptoms into a single diagnostic case, valuable information is lost about the nature of the disorder. As a result, the causes of mental, emotional, and behavioral problems are obscured. Mirowsky and Ross go on to suggest that the reliance on diagnosis does not give a true, that is, a valid, picture of psychiatric distress. Instead, psychiatrists use diagnosis because it allows them to

receive payment from insurance companies who will pay only for cases and because it establishes mental distress as a problem that can be treated only by a physician. Although their criticisms are harsh, these authors reestablish the important distinction between the social construction of psychiatric diagnoses and the social causes of psychological pain. It is the latter issue, however, that most sociological research has addressed.

### THE EPIDEMIOLOGY OF MENTAL DISORDER

**Socioeconomic Status.** The inverse relationship between socioeconomic status and mental disorder is now so well established that it has almost acquired the status of a sociological law. The relationship is surprisingly robust; it holds for most forms of mental disorder, no matter how socioeconomic status is measured, and for both patient populations and community samples. The relationship is strongest and most consistent for schizophrenia, personality disorders, and medically based syndromes. Findings for the major affective disorders are somewhat less consistent. Studies tend to report weak to moderate inverse relationships between social class and the incidence of major affective disorders such as anxiety and depression (Kessler et al. 1994). However, studies sometimes report no relationship (Weissman et al. 1991) or a positive class gradient (Weissman and Myers 1978). Evidence on the class distribution of minor depression is more clear-cut, with studies almost universally showing higher levels of depressive symptomatology among the lower strata. Similarly, research consistently shows that the highest levels of general distress are also found among those with the lowest income, education, or occupational status.

There are two general qualifications to the pattern outlined above. First, even though socioeconomic status is negatively associated with most types of mental disorder, the relationship is probably not linear. Extremely high rates of disorder are typically found in the lowest stratum. Higher strata do have progressively lower rates, but variation is considerably less between them than between the lowest and next-to-lowest tier. Some scholars have claimed, therefore, that serious mental illness is primarily an underclass phenomenon. Second, the inverse relationship between social

class and mental disorder may be stronger in urban than in rural areas and is probably stronger in the United States than in other societies. (For a comprehensive review of this literature, see Ortega and Corzine 1990.)

Most, if not all, of the major sociological theories of mental illness begin with the empirical observation that psychological disorder is most prevalent among those individuals with the fewest resources and the least social power. Until recently, researchers focused almost exclusively on one dimension of inequality—social class. Indeed, the dominant paradigms in the sociology of mental health have derived primarily from the attempt to explain this relationship; hypotheses regarding the effects of gender, age, or marital status on psychological distress are often simple elaborations of models derived from the study of social class and mental disorder. Three general models of the relationship between social resources and mental illness have been suggested. These are (1) the “social causation” hypothesis; (2) the “social selection” or “drift” hypothesis; and (3) the “labeling” or “societal reaction” approach.

**Social Causation.** *Social causation* is a general term used to encompass a number of specific theories about the class-linked causes of mental disorder. Perhaps the most common version of social causation explains the higher rates of mental disorder among the lowest socioeconomic strata in terms of greater exposure to stress. According to this perspective, members of the lower class experience more stressful life events and more chronic strains (Turner et al. 1995). In addition, they are more likely to experience physical hazards in the environment, blocked aspirations, and status frustration (Cockerham 1996). Taken together, these stresses produce elevated rates of psychiatric impairment. In another version of social causation, scholars have argued that class differences in coping resources and coping styles are at least as important in the etiology of mental disorder as are class differences in exposure to stress (Pearlin and Schooler 1978). In this view, poverty increases the likelihood of mental illness because it (1) disrupts precisely those social networks that might effectively buffer the effects of stressful events and (2) inhibits the development of an active, flexible approach to dealing with problems. For both social and psychological reasons, then,

the lower classes make use of less effective coping strategies. Finally, part of the class difference in mental disorder, especially rates of treated disorder, may stem from class differences in attitudes toward mental illness and psychiatric care. Because of more negative attitudes toward mental illness and because of inadequate access to appropriate psychiatric care, the lower classes may be more seriously ill when they first come in contact with the mental health care system, and thus they may be more likely to be hospitalized (Rushing and Ortega 1979).

**Social Selection and Drift.** This perspective implies that, rather than causing mental disorder, low socioeconomic status is a result of psychological impairment. Two mobility processes can be involved. According to the drift hypothesis, the onset of mental disorder adversely affects an individual's ability to hold a job and generate income. As a result of psychological disorder, then, individuals experience downward intragenerational mobility and physical relocation to less socially desirable neighborhoods (Eaton 1980). Social selection, on the other hand, occurs when premorbid characteristics of the mentally disordered individual prevent him or her attaining as high a social status as would be expected of similar individuals in the general population. Here, the focus is on intergenerational mobility (Kendler et al. 1995).

**Labeling or Societal Reaction.** Based on the work of Thomas Scheff (1966), this approach holds that much of the class difference in mental disorder stems not from any real difference in mental illness but rather from a tendency to diagnose or label a disproportionate number of lower-class individuals as psychologically impaired. According to Scheff, the process works as follows. Psychiatric symptoms have many different causes and many people experience them. Only a few individuals, however, are ever labeled as mentally ill. People who are so labeled are drawn from the ranks of those least able to resist the imputation of deviance. Once an individual is identified as mentally ill, a number of forces work to reinforce and solidify a mentally ill self-identity. Once labeled, individuals are encouraged by family and mental health professionals to acknowledge their illness. They are rewarded for behaving as "good" patients should, a task made easier by virtue of the fact that individuals learn the stereotypes of mental illness in early childhood. When individuals are

discharged from the mental hospital, or when they otherwise terminate treatment, they may be rejected by others. This rejection has psychological consequences that simply reinforce a mentally ill identity. The process is self-fulfilling, leading Scheff to conclude that attachment of the mentally ill label is the single most important factor in the development of chronic mental disorder.

**An Assessment.** After two decades or more of acrimonious debate, the search for unitary explanations of the relationship between mental disorder and social class has largely been abandoned. Sociologists seldom claim that mental illness is derived only from medical factors, is caused only by features of the social environment, or stems purely from societal reaction. Most scholars now believe that different types of disorders require different types of explanations. Genetic and other biomedical factors are clearly involved in schizophrenia, and certain forms of depression. However, genetics, brain chemistry, and other medical factors do not provide the entire answer since, even among identical twins, concordance rates for mental illness fall only in the range of 30 to 50 percent. Thus, the causes of mental disorder must also be sought in the social environment. Research does suggest a modest relationship between the social stressors attendant to lower-class status and the onset of some forms of mental disorder. The evidence is clearest, however, for anxiety, substance abuse, and relatively minor forms of depression or psychological distress. For the more severe forms of mental illness and for conduct disorders, the drift and selection hypothesis appears to have the most empirical support. (See Miech et al. 1999 for a comprehensive review and data bearing on these points.) Although labeling is not the only cause of chronic mental illness, it is clear that the mental illness label does have negative consequences. In what has come to be called modified labeling theory, researchers have demonstrated that the status of ex-mental patient and the discrimination that follows from it, coupled with the ex-patient's expectation of rejection by others, adversely affects earnings, work status, and subsequent mental health (Link 1987; Markowitz 1998). Thus, labeling is one of the processes through which drift occurs. As researchers continue to refine the definitions and measurement of various mental disorders and as they more clearly delineate the processes of social causation, drift, and

labeling, it is likely that further theoretical convergences will be identified.

**Gender.** It is not yet clear whether there are significant gender differences in overall rates of mental illness. There is little doubt, however, that certain types of disorders occur more frequently among women than among men. Research clearly shows that women are more likely to suffer from major and minor depression and anxiety than are men (Kessler et al. 1994); men, however, are usually found to have higher rates of antisocial personality disorders and the various forms of substance abuse and dependence (Aneshensel et al. 1991). The sex ratio for some forms of mental disorder may be age dependent; males have higher rates of schizophrenia prior to adolescence and females have higher rates in later adulthood (Loring and Powell 1988). Studies also find that male-female differences in levels of depression are most pronounced among young adults (Dean and Ensel 1983). Furthermore, gender effects appear to interact with those of marital, occupational, and parental roles.

Scholars continue to disagree about the precise form of the interaction effects of gender and marital status on mental illness. Virtually all studies report that gender differences are most pronounced among married persons; married women consistently show higher levels of depression and anxiety than married men. Evidence on the unmarried, however, is mixed. Research based on treated populations often finds higher rates of disorder among single men. Studies based on community samples more frequently report higher distress levels among unmarried women. The interaction between gender and marital status is further complicated by the presence of children, work outside the home, or both.

Some research on married persons finds that gender differences are reduced when both husbands and wives are employed. Studies comparing groups of women often find that employment has modest, positive effects on mental health. However, other studies report no difference between employed women and housewives (Carr 1997) and a few report that married, full-time homemakers with children have fewer worries and more life satisfaction (Veroff et al. 1981). These apparently contradictory findings stem, in part, from the

different measures of mental health and illness used. It is possible, for instance, that small children can simultaneously increase their mothers' life satisfaction and their overall levels of anxiety and distress. However, two substantive factors also appear to be involved. First, it is the demands created by children and employment, rather than by parental or employment status per se, that cause elevated levels of distress among married women (Rosenfeld 1989). The level of demands varies, of course, depending upon the level of male responsibility for child care and housework. Second, employment decreases gender differences in distress only when it is consistent with both the husband's and the wife's desires. Married men's distress levels may, in fact, surpass married women's when wives work but their husbands prefer them not to (Ross et al. 1983).

As is true for social class, explanations for gender differences in mental disorder fall into three broad classes: social causation, social selection, and labeling. Because of the consistency of gender effects (at least for depression and anxiety) and improvements in the reliability of psychiatric diagnoses, most recent work has focused on the ways in which the social and psychological correlates of male and female roles cause variation in rates of mental disorder. Some have argued that differences in sex-role socialization make females more likely to direct frustration inward, toward themselves, rather than towards others, as males might. Thus, women are more likely to develop introjective disorders, whereas men are more likely to behave in antisocial ways (Loring and Powell 1988). Others have argued that women are more attached to others and are more sensitive to others' needs than are men. As a result, not only is women's mental health influenced by their own experiences but, also in contrast to men, they are more psychologically vulnerable to the stresses or losses of loved ones (Kessler and McLeod 1984). Although the empirical literature is far from clear on this point (cf. Umberson et al. 1996), according to this perspective, women experience more stressful events and are more psychologically reactive to them than are men. Other explanations—for both direct and interactive effects—of gender on mental illness have focused on male-female differences in power, resources, demands, and personal control. Insofar as employment increases women's

power and resources, it is likely to have positive effects on mental health. Well-educated employed women have fewer mental symptoms than nonworking women; among working-class and lower-class women, however, employment may actually increase anxiety and depression because it elevates demands at the same time that it produces only marginal increases in resources (Sales and Hanson Frieze 1984). Since employed women generally retain full responsibility for children, the demands of caring for children, particularly those under the age of 6, exacerbate work-related stress. Thus, male-female differences in power and resources produce differences in ability to control demands. Gender differences in control, in turn, shape perceptions of personal mastery; personal mastery is the psychological mechanism that connects gender differences in resources and demands to gender differences in mental illness (Rosenfield 1989).

The social selection perspective is valid only for explaining male-female differences in the relationship between marital status and mental disorder. The argument is that mental illness is more likely to select men out of marriage than women. (See Rushing 1979 for a related discussion.) According to this perspective, male forms of mental disorder—psychosis and antisocial personality, for example—prevent impaired men from satisfactorily discharging the traditional male obligation to be good economic providers, making them ineligible as marriage partners. In contrast, female forms of psychiatric impairment may go undetected for long periods of time and may not seriously interfere with a woman's ability to fulfill the traditional housekeeping role. Thus, the higher rates of female disorder among the married may be a partial artifact of the differing probabilities of marriage for mentally disordered men and women.

The labeling explanation for male-female differences in psychiatric impairment begins by challenging the notion that women actually experience more symptoms and disorders than men do. Labeling theorists argue that women are overdiagnosed and overmedicated because of biases on the part of predominantly male psychiatrists and because of the male biases inherent in psychiatric nomenclature. Coupled with the greater willingness of females to admit their problems and to seek help for them, these biases simply

produce the illusion that women are more likely to be disordered than men. Scholars using the labeling–societal reaction–critical perspective argue that the effects of gender biases are not benign and that they have consequences at two levels. First, individual women are unlikely to receive appropriate services for their real mental health problems. Second, and at a societal level, critics argue that psychiatry simply legitimates traditional gender roles, thereby buttressing the status quo (Chesler 1973).

**A Theoretical Assessment.** With the development of DSM-IV and with increases in the number of female mental health professionals, concern over the issues raised by labeling theorists has diminished somewhat. Trusting that the most blatant instances of sexism have been eliminated, researchers have turned their attention toward specifying the social psychological dynamics of the gender–mental health equation; considerable progress has been made in elucidating the circumstances under which women are most likely to experience symptoms of mental disorder. Nevertheless, it may be premature to close the question of gender bias in psychiatric disorders. In one study, male clinicians appeared to overestimate the prevalence of depressive disorders among women, a tendency that is certainly consistent with gender stereotypes. In the same study, black males were most likely to be diagnosed as paranoid schizophrenics, a view consistent with both gender and racial stereotypes (Loring and Powell 1988). In yet another study, male and female psychiatrists made similar diagnoses of male and female patients presenting severe Axis I conditions but made significantly different diagnoses for male and female patients with Axis II conditions, such as personality disorders (Dixon et al. 1995). Thus, advances of DSM-III (and IV) notwithstanding, the authors of these studies conclude that sex and race of client and psychiatrist continue to influence diagnosis even when psychiatric criteria appear to be clear-cut.

**Age.** Among adults, and with the exception of some types of dementia and other syndromes due to general medical conditions, rates of mental illness decrease with age. Rates of schizophrenia, manic disorder, drug addiction, and antisocial personality all peak between the ages of 25 and 44 (Robins et al. 1984). Furthermore, an older person

with a serious mental disorder is likely to have had a first psychiatric episode in young or middle adulthood. At least 90 percent of older schizophrenics experienced the onset of the disorder in earlier life. Similarly, about two-thirds of older alcoholics have a long history of alcohol abuse or dependence (Hinrichsen 1990). Depression is the disorder most likely to occur among the elderly, and a substantial proportion of older community residents do report some of its symptoms. In general, the relationship between age and depression appears to be curvilinear, with depression lowest among the middle aged, higher among younger and older adults, and highest among the oldest (Mirowsky and Ross 1992). Nevertheless, relatively few of these older individuals meet criteria for clinical depression (Blazer et al. 1987), and rates of major depression are lower among older adults than in younger age groups. Some older persons, however, are more vulnerable to depression than others. As is true throughout the life cycle, women, individuals with health problems, the unmarried, and those with lower socioeconomic status are at greater risk of depression in late life than their peers. Estimating the true prevalence of depression among the elderly is especially problematic because its symptoms are frequently confused with Alzheimer's disease or other forms of dementia.

According to some estimates, two to four million older Americans suffer some form of mental disorder due to a general medical condition. Of these, roughly half are diagnosed with Alzheimer's, a disease that involves an irreversible, progressive deterioration of the brain. Approximately half of all nursing home residents are estimated to suffer from some form of dementia. Because there is no known treatment for most of these disorders, older mental patients receive little psychiatric care. Critics suggest, however, that many older persons are improperly diagnosed as having disorders of general medical origin. A sizable minority may actually be depressed; others may have treatable forms of dementia caused by medications, infection, metabolic disturbances, alcohol, or brain tumors. In some instances, then, the stereotype that senility is a concomitant of the aging process prevents appropriate diagnosis, intervention, and treatment.

Most explanations of the age-mental health relationship have focused on specific age groups.

Clinicians suggest, for instance, that anxiety and depression in middle age are a consequence of hormonal change or of changes in family and occupational roles. The personality disorders of young adulthood are often explained in terms of the stresses produced by the transition from adolescence to full adult roles. Among the elderly, explanations have focused on either organic or environmental factors. The dementias have recognized organic causes. Although neither is a normal part of the aging process, the two major causes of these disorders are (1) the deterioration of the brain tissue that is associated with Alzheimer's disease and (2) cerebral arteriosclerosis. However, environmental factors also contribute to the onset of the dementias. They do so, in part, by increasing the likelihood of stroke or heart attack. In contrast, primary mental disorders, such as depression, personality disorders, and anxiety, depend more directly upon environmental factors. Some types of depression appear to have a genetic component, but the genetic link appears to be stronger in early- than in late-onset cases. Individuals who have their first episode of clinical depression prior to the age of 50, for instance, are more likely to have relatives with depression than those who become depressed in later years (Hinrichsen 1990). Consequently, losses typical of late life—losses of health, occupation, income, and loved ones—appear to be the primary causes of mental health problems among older adults.

Clearly, no single theory can adequately explain the etiology of mental disorder; at each stage of the life cycle, variables that are relevant to the onset of one type of disorder may be insignificant in the onset of other illnesses. Similarly, no single variable or set of variables is likely to explain age differences in overall rates of mental disorder. Nevertheless, efforts are under way to systematically explain the inverse relationship between age and primary psychiatric impairment. Gove and his associates have suggested that psychological distress decreases with age because individuals are able, over time, to find and settle into an appropriate social niche; as individuals move through life, they become less emotional and less self-absorbed, function more effectively in their selected roles, and generally become more content with themselves and with others. As a result, rates of mental disorder decrease from late adolescence through late life (Gove 1985; Gove et al. 1989).

**Place of Residence.** Sociologists have commonly assumed that rates of mental disorder are higher in urban than in rural areas. However, this assumption is based more on the antiurban bias of much sociological theory than it is on empirical research. In a thoughtful and systematic review, Wagenfeld (1990) has argued that there is little evidence in the mental health literature to suggest the superiority of rural life. In several of the rural community studies Wagenfeld cites, researchers report a “probable” case rate of depression and anxiety of 12 to 20 percent. Studies that explicitly compare rural and urban communities generally find that rates of psychosis are higher in rural communities and that rates of depression are somewhat higher in urban areas. Residents of metropolitan communities also appear more likely to have multiple diagnoses than do rural residents, leading Kessler et al. (1994) to conclude that urban-rural differences in the prevalence of mental disorder probably reflect differences in comorbidity rather than differences in rates of individuals having a psychiatric condition. Differences in case definition and diagnosis, differences in how “rural place of residence” is defined and measured, and differences in the time period during which studies were conducted make it difficult, overall, to assess whether rural communities have significantly higher overall rates of pathology than urban areas. Results are sufficient, however, to suggest that rural life is not as blissful as it is often claimed to be. Recent declines in the rural economy, the out-migration of the young and upwardly mobile, and the relative paucity of mental health services are likely to be major contributing factors in the etiology of rural mental health problems.

**Other factors.** Epidemiologists have also explored the relationships between the incidence or prevalence of mental disorder and such variables as race and ethnicity, migration, social mobility, and marital status. In each case, results generally support the view that individuals with the fewest resources—both economic and social—are most likely to experience psychiatric impairment. However, most research has adopted a rather static view; few studies have assessed the extent to which relationships between each of these variables and mental disorder have changed over time. Given the significant changes in diagnostic practices and in the mental health professions over the last decades, this is a striking omission.

#### AN AGENDA FOR FUTURE RESEARCH

Since the early 1960s, psychiatric sociology has undergone enormous changes. During the 1960s and 1970s, much of the literature was sharply critical of psychiatry and of medical models of madness. Although sociologists were divided about the relative importance of labeling processes in the etiology of mental illness, most agreed that psychiatric diagnoses were unreliable and were influenced by social status and social resources, that long-term institutionalization had detrimental effects, and that at least some patients were hospitalized inappropriately. Such criticisms provided one impetus for the substantial change that took place in psychiatric care during the same period; laws were changed to make involuntary commitment more difficult; steps were taken to deinstitutionalize many mental patients; and a major effort was made to improve the reliability of mental diagnoses. By the time DSM-III was published in 1980, the most flagrant abuses and the sharpest criticism of psychiatry seemed to have disappeared. Consequently, many sociologists shifted their attention from concerns about the lives of people with serious mental disorder to the social correlates of psychological distress among the general population (Cook and Wright 1995). Using what is basically a medical model of impairment, researchers have focused on delineating the relationship between social variables (such as gender, age, race, social class, place of residence, life events, and stress) and specific diagnoses (most often depressive symptoms, anxiety, and substance abuse). Indeed, the psychiatric view of mental disorder is so well established in sociology that the growing literature on homelessness has generally accepted the assertion of mental health professionals that most of the homeless are simply individuals who have fallen through the cracks of the mental health care system. (For notable exceptions, see Bogard et al. 1999; Snow, Baker, and Anderson 1986.) It is surprising that sociologists have been so uncritical in their acceptance of this position; it is also surprising that in the decade of the 1990s, declared by the National Institute of Mental Health to be the “Decade of the Brain,” they have been so ready to accept the view that mental illness is primarily a problem of genetics or brain chemistry and that it can be treated just like any other disease. It is certainly true that enormous strides have been made in the diagnosis and psychopharmacological



treatment of mental disorder. It is also certainly true that biomedical factors are causally involved in some types of mental illness. Sociologists must, therefore, continue their efforts to develop a model of mental disorder that integrates medical, psychological, and social factors.

As some critics point out, however, the current emphasis on diagnoses, cases, and the medical model of mental illness has limitations. Acceptance of the psychiatric view of mental disorder leads to the acceptance of policy recommendations that are not yet firmly grounded in empirical research. It is far from clear, for instance, that deinstitutionalization of the mentally ill is the primary cause of homelessness in America. As Bogard et al. (1999) point out, conventional wisdom notwithstanding, very few homeless mothers are mentally ill; it can be reasonably argued, then, that the enormous resources that have been directed toward providing them with mental health care might more appropriately and effectively be used to provide safe, affordable housing. In a similar vein, Link and Phelan (1995) note that current attention to individual risk factors in disease gives rise to "personal policy" recommendations that leave totally unaddressed the fundamental social conditions that cause differential exposure to risk.

Furthermore, few studies have assessed the extent to which changes in psychiatric diagnosis or changes in the civil rights guarantees of mental patients have affected the delivery and quality of mental health services. Consumers and families have voiced concern that the powerful new psychopharmacological drugs are being inappropriately used as forms of social control and chemical restraint at the same time that research continues to show that it is racial and ethnic minority consumers who are most likely to be so restrained (Cook and Wright 1995). Aside from the field trials used in their formulation, few studies have assessed the reliability and validity of DSM-IV diagnoses. However, results from several studies show that nonclinical factors such as gender, race, the availability of viable community housing and the presence of reliable caretakers significantly affect not only diagnosis but treatment protocols and outcomes. (See Cook and Wright [1995] for a review of these studies and these concerns.) It is far from clear, then, that lower-class women are any more likely to receive

appropriate care in 1999 than they were in 1950 or 1970. It is unclear whether urban-rural differences in rates of mental disorder have changed over time and, if so, to what extent changes in diagnostic systems or service availability are implicated. Evidence that rural residents may actually experience mental illness at approximately the same rates as urban residents coupled with an acute shortage of rural mental health providers suggest the importance of understanding the diagnostic practices of primary-care physicians and of providing appropriate training to them.

Research in the next century must adopt a more dynamic or process view of mental health issues. The consequences of changes in psychiatric diagnosis, of the increased reliance on drug therapies, of changes in mental health law and policy, and in the availability of mental health services must be assessed. Changes in the mental health system must be linked to changes in the composition of the pool of "potential clients" and to issues regarding the development of gender, age, class, and culturally appropriate systems of care.

REFERENCES

- American Psychiatric Association 1980 *Diagnostic and Statistical Manual of Mental Disorders III*. Washington: APA.
- 1994 *Diagnostic and Statistical Manual of Mental Disorders IV*. Washington: APA.
- Aneshensel, Carol S., Carolyn M. Rutter, and Peter A. Lachenbruch 1991 "Social Structure, Stress, and Mental Health." *American Sociological Review* 56:166-178.
- Blazer, Dan, Dana C. Hughes, and Linda K. George 1987 "The Epidemiology of Depression in an Elderly Community Population." *Gerontologist* 27:281-287.
- Bogard, Cynthia J., J. Jeff McConnell, Naomi Gerstel, and Michael Schwartz 1999 "Homeless Mothers and Depression: Misdirected Policy." *Journal of Health and Social Behavior* 40:46-62.
- Carr, Deborah 1997 "The Fulfillment of Career Dreams at Midlife: Does It Matter for Women's Mental Health?" *Journal of Health and Social Behavior* 38:331-344.
- Chesler, Phyllis 1973 *Women and Madness*. New York: Avon.
- Cockerham, William 1996 *Sociology of Mental Disorder*, 4th ed. Upper Saddle River, N.J.: Prentice Hall.
- Cook, Judith A., and Eric R. Wright 1995 "Medical Sociology and the Study of Severe Mental Illness:

- Reflections on Past Accomplishments and Directions for Future Research." *Journal of Health and Social Behavior* (Extra Issue):95-114.
- Crandall, D. L., and B. P. Dohrenwend 1967 "Some Relations among Psychiatric Symptoms, Organic Illness, and Social Class." *American Journal of Psychiatry* 1527-1538.
- Dean, Alfred, and Walter M. Ensel 1983 "Socially Structured Depression in Men and Women." In James Greenley, ed., *Research in Community Mental Health*, vol. 3. Greenwich, Conn.: JAI.
- Dixon, Jo, C. Gordon, and T. Khomusi 1995 "Sexual Symmetry in Psychiatric Diagnosis." *Social Problems* 42:429-448.
- Dohrenwend, Bruce P., and D. L. Crandall 1970 "Psychiatric Symptoms in Community, Clinic, and Mental Hospital Groups." *American Journal of Psychiatry* 126:1611-1621.
- Eaton, William W. 1980 *The Sociology of Mental Disorder*. New York: Praeger.
- , M. M. Weissman, J. Anthony, L. Robins, D. Blazer, and M. Karno 1985 "Problems in the Definition and Measurement of Prevalence and Incidence of Psychiatric Disorder." In W. Eaton and R. Kessler, eds., *Epidemiologic Field Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program*. Orlando, Fla.: Academic.
- Faris, Robert E., and H. Warren Dunham 1939 *Mental Disorders in Urban Areas*. Chicago: University of Chicago Press.
- Figert, Anne 1994 *Women and the Ownership of PMS: The Structuring of a Psychiatric Diagnosis*. Hawthorne, N.Y.: Aldine de Gruyter.
- Gove, Walter R. 1985 "The Effect of Age and Gender on Daviant Behavior: A Biopsychosocial Perspective." In Alice Rossi, ed., *Gender and the Life Course*. New York: Aldine de Gruyter.
- , Suzanne T. Ortega, and Carolyn Briggs Style 1989 "The Maturational and Role Perspectives on Aging and Self Through the Adult Years: An Empirical Evaluation." *American Journal of Sociology* 94:1117-1145.
- Hinrichsen, Gregory A. 1990 *Mental Health Problems and Older Adults*. Santa Barbara, Calif.: ABC-CLIO.
- Hollingshead, A. B., and F. C. Redlich 1958 *Social Class and Mental Illness*. New York: John Wiley.
- Institute of Medicine, Committee on Prevention of Mental Disorders, Division of Biobehavioral Sciences and Mental Disorders 1994 *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington: National Academy Press.
- Kendell, R. E. 1988 "What is a Case? Food for Thought for Epidemiologists." *Archives of General Psychiatry* 45:374-376.
- Kendler, Kenneth S., Ellen E. Walters, Michael C. Neale, Ronald C. Kessler, Andrew C. Heath, and Lindon J. Eaves 1995 "The Structure of the Genetic and Environmental Risk Factors for Six Major Psychiatric Disorders in Women: Phobia, Generalized Anxiety Disorder, Panic Disorder, Bulimia, Major Depression, and Alcoholism." *Archives of General Psychiatry* 52:374-383.
- Kessler, Ronald C., Katherine A. McGonagle, Shanyang Zhao, Christopher B. Nelson, Michael Hughes, Suzanne Eshleman, Hans-Ulrich Wittchen, and Kenneth S. Kendler 1994 "Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States: Results from the National Comorbidity Study." *Archives of General Psychiatry* 51:8-19.
- , and Jane McLeod 1984 "Sex Differences in Vulnerability to Undesirable Life Events." *American Sociological Review* 49:620-631.
- Kirk, Stuart A., and Herb Kutchins 1992 *The Selling of DSM*. Hawthorne, N.Y.: Aldine de Gruyter.
- Link, Bruce 1987 "Understanding Labeling Effects in the Area of Mental Disorders: An Assessment of Expectations of Rejection." *American Sociological Review* 52:96-112.
- , and Jo Phelan 1995 "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* (Extra Issue):80-94.
- Loring, Marti, and Brian Powell 1988 "Gender, Race, and DSM-III: A Study of the Objectivity of Psychiatric Behavior." *Journal of Health and Social Behavior* 29:1-22.
- Markowitz, Fred E. 1998 "The Effects of Stigma on the Psychological Well-Being and Life Satisfaction of Persons with Mental Illness." *Journal of Health and Social Behavior* 39:335-347.
- Miech, Richard A., Avshalom Caspi, Terrie E. Moffitt, Bradley R. Entner Wright, and Phil A. Silva 1999 "Low Socioeconomic Status and Mental Disorders: A Longitudinal Study of Selection and Causation during Young Adulthood." *American Journal of Sociology* 104:1096-1131.
- Mirowsky, John, and Catherine E. Ross 1992 "Age and Depression." *Journal of Health and Social Behavior* 33:187-205.
- , and Catherine E. Ross 1989 "Psychiatric Diagnosis as Reified Measurement." *Journal of Health and Social Behavior* 30:11-25.
- Ortega, Suzanne T., and Jay Corzine 1990 "Socioeconomic Status and Mental Disorders." In James

- Greenley, ed., *Research in Community and Mental Health*, vol. 6. Greenwich, Conn.: JAI.
- Pearlin, Leonard, and Carmi Schooler 1978 "The Structure of Coping." *Journal of Health and Social Behavior* 19:2-21.
- Robins, Lee N., J. E. Helzer, M. M. Weissman, H. Orvaschel, E. Gruenberg, J. D. Burke, Jr., and D. A. Regier 1984 "Lifetime Prevalence of Specific Disorders in Three Sites." *Archives of General Psychiatry* 41:949-958.
- Rosenfield, Sarah 1989 "The Effects of Women's Employment: Personal Control and Sex Differences in Mental Health." *Journal of Health and Social Behavior* 30:77-91.
- Ross, Catherine E., John Mirowsky, and Joan Huber 1983 "Dividing Work, Sharing Work, and In-Between: Marriage Patterns and Depression." *American Sociological Review* 48:809-823.
- Rushing, William A. 1979 "The Functional Importance of Sex Roles and Sex-Related Behavior in Societal Reactions to Residual Deviants." *Journal of Health and Social Behavior* 20:208-217.
- , and Suzanne T. Ortega 1979 "Socioeconomic Status and Mental Disorder: New Evidence and a Sociomedical Formulation." *American Journal of Sociology* 84:1175-1200.
- Sales, Esther, and Irene Hanson Frieze 1984 "Women and Work: Implications for Mental Health." In L. E. Walker, ed., *Women and Mental Health Policy*. Beverly Hills, Calif.: Sage.
- Scheff, Thomas J. 1966 *Being Mentally Ill: A Sociological Theory*. Chicago: Aldine de Gruyter.
- Scott, Wilbur J. 1990 "PTSD in DSM-III: A Case in the Politics of Diagnosis and Disease." *Social Problems* 37:294-310.
- Snow, David, Susan Baker, and Leon Anderson 1986 "The Myth of Pervasive Mental Illness among the Homeless." *Social Problems* 33:407-423.
- Srole, L., T. S. Langner, S. T. Michel, M. D. Opler, and T. C. Rennie 1962 *Mental Health in the Metropolis: The Midtown Manhattan Study*. New York: McGraw-Hill.
- Tavris, Carol 1993 "Do You Menstruate? If So, Psychiatrists Think You May Be Nuts." *Glamour* (November):172.
- Townsend, John M. 1980 "Psychiatry versus Societal Reaction: A Critical Analysis." *Journal of Health and Social Behavior* 21:268-278.
- Turner, R. Jay, Blair Wheaton, and Donald A. Lloyd 1995 "The Epidemiology of Social Stress." *American Sociological Review* 60:104-125.
- Umberson, Debra, Meichu D. Chen, James S. House, Kristine Hopkins, and Ellen Slaten 1996 "The Effect of Social Relationships on Psychological Well-Being: Are Men and Women Really So Different?" *American Sociological Review* 61:837-857.
- Veroff, Joseph, Elizabeth Douvan, and Richard A. Kulka 1981 *The Inner American*. New York: Basic.
- Wagenfeld, Morton O. 1990 "Mental Health and Rural America: A Decade Review." *Journal of Rural Health* 6:507-522.
- Weissman, Myrna M., Martha L. Bruce, Phillip J. Leaf, Louis P. Florio, and Charles E. Holzer 1991 "Affective Disorders." In Lee N. Robins and Darrel A. Regier, eds., *Psychiatric Disorders in America: The Epidemiological Catchment Area Study*. New York: Free Press.
- , and Jerome K. Myers 1978 "Affective Disorders in a U.S. Urban Community." *Archives of General Psychiatry* 35:1304-1311.
- , D. Scholomskas, M. Pottenger, B. Prusoff, and B. Locke 1977 "Assessing Depressive Symptoms in Five Psychiatric Populations: A Validation Study." *American Journal of Epidemiology* 106:203-214.

SUZANNE T. ORTEGA  
SHARON L. LARSON

## MERITOCRACY

See Affirmative Action; Equality of Opportunity.

## META-ANALYSIS

*Meta-analysis* is the practice of statistically summarizing empirical findings from different studies, reaching generalizations about the obtained results. Thus, "meta-analysis" literally refers to analysis of analyses. Meta-analysis, a term coined by Glass (1976), is also known as *research synthesis* and *quantitative reviewing*. Because progress within any scientific field has always hinged on cumulating empirical evidence about phenomena in an orderly and accurate fashion, reviews of studies have historically proved extremely influential (e.g., Mazela and Malin 1977). With the exponential growth in the numbers of studies available on a given social scientific topic, the need for these reviews has increased proportionally, meaning that reviews are potentially even more important each day. The empirical evidence, consisting of multiple studies examining a phenomenon, exists as a literature on

the topic. Although new studies rarely replicate earlier studies without changing or adding new features, many studies can be described as conceptual replications that use different stimulus materials and dependent measures to test the same hypothesis, and still others might contain exact replications embedded within a larger design that adds new experimental conditions. In other instances, repeated tests of a relation accrue in a less systematic manner because researchers sometimes include in their studies tests of particular hypotheses in auxiliary or subsidiary analyses.

In order to reach conclusions about empirical support for a phenomenon, it is necessary to compare and contrast the findings of relevant studies. Therefore, accurate comparisons of study outcomes—reviews of research—are at the very heart of the scientific enterprise. Until recently these comparisons were nearly always made using informal methods that are now known as *narrative reviewing*, a practice by which scholars drew overall conclusions from their impressions of the overall trend of the studies' findings, sometimes guided by a count of the number of studies that had either produced or failed to produce statistically significant findings in the hypothesized direction. Narrative reviews have appeared in many different contexts and still serve a useful purpose in writing that does not have a comprehensive literature review as its goal (e.g., textbook summaries, introductions to journal articles reporting primary research). Although narrative reviewing has often proved useful, the method has often proved to be inadequate for reaching definitive conclusions about the degree of empirical support for a phenomenon or for a theory about the phenomenon. One indication of this inadequacy is that independent narrative reviews of the same literature often have reached differing conclusions.

#### COMMON PROBLEMS WITH NARRATIVE REVIEWS

Critics of the narrative reviewing strategy (e.g., Glass et al. 1981; Rosenthal 1991) have pointed to four general faults that frequently occur in narrative reviewing: (1) Narrative reviewing generally involves the use of a convenience sample of studies, perhaps consisting of only those studies that the reviewer happens to know. Because the parameters of the reviewed literature are typically

not explicit, it is difficult to evaluate the adequacy of the definition of the literature or the thoroughness of the search for studies. If the sample of studies was biased, the conclusions reached may also be biased. (2) Narrative reviewers generally do not publicly state the procedures they used for either cataloging studies' characteristics or evaluating the quality of the studies' methods. Therefore, the review's claims about the characteristics of the studies and the quality of their methods are difficult to judge for their accuracy. (3) In cases in which study findings differed, narrative reviewing has difficulty in reaching clear conclusions about whether differences in study methods explain differences in results. Because narrative reviewers usually do not systematically code studies' methods, these reviewing procedures are not well suited to accounting for inconsistencies in findings. (4) Narrative reviewing typically relies much more heavily on statistical significance to judge studies' findings than on the *magnitude* of the findings. Statistical significance is a poor basis for comparing studies that have different sample sizes, because effects of identical magnitude can differ widely in statistical significance. Because of this problem, narrative reviewers often reach erroneous conclusions about a pattern in a series of studies, even in literatures as small as ten studies (Cooper and Rosenthal 1980).

As the number of available studies cumulates, the conclusions reached in narrative reviews become increasingly unreliable because of the informality of the methods they use to draw these conclusions. Indeed, some historical scholars have attributed crises of confidence in central social scientific principles to apparent failures to replicate findings across studies (e.g., Johnson and Nichols 1998). Clearly, there will be practical limitations on the abilities of scholars to understand the vagaries of a literature containing dozens if not hundreds of studies (e.g., by 1978, there were at least 345 studies examining interpersonal expectancy effects, according to Rosenthal and Rubin 1978; and by 1983, there were over 1,000 studies evaluating whether birth order is related to personality, as reported by Ernst and Angst 1983). From this perspective, the social sciences might be considered victims of their own success: Although social scientists have been able to collect a myriad of data about a myriad of phenomena, they were forced to rely on their intuition when it came to assessing

the state of the knowledge about popular topics. Since the 1980s, however, as scholars have gained increasing expertise in reviewing research literatures, literatures that once appeared haphazard at best and fragile at worst now frequently are shown to have substantial regularities (Johnson and Nichols 1998). For example, although scholars working in the 1950s and on through the 1970s frequently reached conflicting conclusions about whether men or women (or neither) are more easily influenced by others, reviewers using meta-analytic techniques have found highly reliable tendencies in this same literature. For example, Eagly and Carli's meta-analysis (1981) showed that men are more influenced than women when the communication topic is feminine (e.g., sewing), and that women are more influenced than men when the topic is masculine (e.g., automobiles). Moreover, contemporary meta-analysts now almost routinely move beyond relatively simple questions of whether one variable relates to another to the more sophisticated question of *when* the relation is larger, smaller, or reverses in sign. Thus, there is, indeed, a great deal of replicability across a wide array of topics, and inconsistencies among study findings can often be explained on the basis of methodological differences among the studies.

#### META-ANALYTIC REVIEWS OF EVIDENCE

Because of the importance of comparing study findings accurately, scholars have dedicated considerable effort to making the review process as reliable and valid as possible and thereby circumventing the criticisms listed above. These efforts highlight the proposition that research synthesis is a scientific endeavor—there are identifiable and replicable methods involved in producing reliable and valid reviews (Cooper and Hedges 1994). Although scientists have cumulated empirical data from independent studies since the early 1800s (see Stigler 1986), relatively sophisticated techniques for synthesizing study findings emerged only after the development of such standardized indexes as *r*-, *d*-, and *p*-values, around the turn of the twentieth century (see Olkin 1990). Reflecting the field's maturation, Hedges and Olkin (1985) presented a sophisticated version of the statistical bases of meta-analysis, and standards for meta-analysis have grown increasingly rigorous. Meta-analysis is now quite common and well accepted because scholars realize that careful application of

these techniques often will yield the clearest conclusions about a research literature (Cooper and Hedges 1994; Hunt 1997).

Conducting a meta-analysis generally involves seven steps: (1) determining the theoretical domain of the literature under consideration, (2) setting boundaries for the sample of studies, (3) locating relevant studies, (4) coding studies for their distinctive characteristics, (5) estimating standardized effect sizes for each study, (6) analyzing the database, and (7) interpreting and presenting the results. The first conceptual step is to specify with great clarity the phenomenon under review by defining the variables whose relation is the focus of the review. Ordinarily a synthesis evaluates evidence relevant to a single hypothesis; the analyst studies the history of the research problem and of typical studies in the literature. Typically, the research problem will be defined as a relation between two variables, such as the influence of an independent variable on a dependent variable (e.g., the influence of silicon breast implants on connective tissue disease, as reported by Perkins et al. 1995). Moreover, a synthesis must take study quality into account at an early point to determine the kinds of operations that constitute acceptable operationalizations of these conceptual variables. Because studies testing a particular hypothesis typically differ in the operations used to establish the variables, it is no surprise that these different operations were often associated with variability in studies' findings. If the differences in studies' operations can be appropriately judged or categorized, it is likely that an analyst can explain some of this variability in effect size magnitude.

The most common way to test competing explanations is to examine how findings pattern across studies. Specifically, a theory might imply that a third variable should influence the relation between the independent and dependent variables: The relation should be larger or smaller with a higher level of this third variable. Treating this third variable as a potential moderator of the effect, the analyst would code the studies for their status on the moderator. This meta-analytic strategy, known as the moderator variable approach, tests whether the moderator affects the examined relation across the studies included in the sample. This moderator variable approach, advancing beyond the simple question of *whether* the independent variable is related to the dependent variable,

addresses the question of *when* the magnitude or sign of the relationship varies. In addition to this moderator variable approach to synthesizing studies' findings, other strategies have proved to be useful. In particular, a theory might suggest that a third variable serves as a mediator of the critical relation because it conveys the causal impact of the independent variable on the dependent variable. If at least some of the primary studies within a literature have evaluated this mediating process, mediator relations can be tested within a meta-analytic framework by performing correlational analyses that are an extension of path analysis with primary-level data (Shadish 1996).

Clearly, only some studies will be relevant to the conceptual relation that is the focus of the meta-analysis, so analysts must define boundaries for the sample of studies, the second step in conducting a meta-analysis. Decisions about the inclusion of studies are important because the inferential power of any meta-analysis is limited by the methods of the studies that are integrated. To the extent that all (or most) of the reviewed studies share a particular methodological limitation, any synthesis of these studies would be limited in this respect. As a general rule, research syntheses profit by focusing on the studies that used stronger methods to test the meta-analytic hypotheses. Nonetheless, it is important to note that studies that have some strengths (e.g., manipulated independent variables) may have other weaknesses (e.g., deficiencies in ecological validity). In deciding whether some studies may lack sufficient rigor to include in the meta-analysis, it is important to adhere to methodological standards within the area reviewed. Although a large number of potential threats to methodological rigor have been identified (Campbell and Stanley 1963; Cook and Campbell 1979), there are few absolute standards of study quality that can be applied uniformly in every meta-analysis. As a case in point, although published studies are often thought to be of higher quality than unpublished studies, there is little basis for this generalization: Many unpublished studies (e.g., dissertations) have high quality, and many studies published in reputable sources do not. It is incumbent on the analyst to define the features of a high-quality study and to apply this definition to all studies in the literature, regardless of such considerations as the reputation of the journal.

Analysts often set the boundaries of the synthesis so that the methods of included studies differ dramatically only on critical moderator dimensions. If other, extraneous dimensions are thereby held relatively constant across the reviewed studies, moderator variable analyses can be more clearly interpreted. Nonetheless, an analyst should include in the sample all studies or portions of studies that satisfy the selection criteria, or, if an exhaustive sampling is not possible, a representative sample of those studies. Following this principle yields results that can be generalized to the universe of studies on the topic.

Because including a large number of studies generally increases the value of a quantitative synthesis, it is important to locate as many studies as possible that might be suitable for inclusion, the third step of a meta-analysis. To ensure that a sufficient sample of studies is located, reviewers are well advised to err in the direction of being extremely inclusive in their searching procedures. As described elsewhere (e.g., Cooper 1998; White 1994), there are many ways to find relevant studies; ordinarily, analysts should use all these techniques. Because computer searches of publication databases seldom locate all the available studies, it is important to supplement them by (1) examining the reference lists of existing reviews and of studies in the targeted literature, (2) obtaining published sources that have cited seminal articles within the literature, (3) contacting the extant network of researchers who work on a given topic to ask for new studies or unpublished studies, and (4) manually searching important journals to find some reports that might have been overlooked by other techniques.

Once the sample of studies is retrieved, analysts code them for their methodological characteristics, the fourth step in the process. The most important of these characteristics are potential moderator variables, which the analyst expects on an a priori basis to account for variation among the studies' effect sizes, or which can provide useful descriptive information about the usual context of studies in the literature. In some cases, reviewers recruit outside judges to provide ratings of methods used in studies. Because accurate coding is crucial to the results of a meta-analysis, the coding of study characteristics should be carried out by two or more coders, and an appropriate index of interrater reliability should be calculated.

To be included in a meta-analysis, a study must contain some report of a quantitative test of the hypothesis that is under scrutiny in order to convert summary statistics into effect sizes, the fifth step of the process. Most studies report the examined relation by one or more inferential statistics (e.g., *t*-tests, *F*-tests, *r*-values), which can be converted into an effect size (see Cooper and Hedges 1994b; Glass et al. 1981; Johnson 1993; Rosenthal 1991). The most commonly used effect size indexes in meta-analysis are the standardized difference and the correlation coefficient (see Rosenthal 1991, 1994). The standardized difference, which expresses the finding in standard deviation units, was first proposed by Cohen (1969) in the following form:

$$g = \frac{M_A - M_B}{SD} \quad (1)$$

where  $M_A$ ; and  $M_B$ ; are the sample means of two compared groups, and  $SD$  is the standard deviation, pooled from the two observations. Because this formula overestimates population effect sizes to the extent that sample sizes are small, Hedges (1981) provided a correction for this bias; with the bias corrected, this effect estimate is conventionally known as  $d$ . Another common effect size is the correlation coefficient,  $r$ , which gauges the association between two variables. Because the sampling distribution of a sample correlation coefficient tends to be skewed to the extent that the population correlation is large, it is conventional in meta-analysis to use a logarithmic transform of each correlation in statistical operations (Fisher 1921). The positive or negative sign of the effect sizes computed in a meta-analysis is defined so that studies with opposite outcomes have opposing signs. When a study examines the relation of interest within levels of another variable, effect sizes may be calculated within the levels of this variable as well as for the study as a whole. In addition to correcting the raw  $g$  and  $r$  because they are biased estimators of the population effect size, analysts sometimes correct for many other biases that accrue from the methods used in each study (e.g., unreliability of a measure; see Hunter and Schmidt 1990). Although it is unrealistic for analysts to take into account all potential sources of bias in a meta-analysis, they should remain aware of biases that may be important within the context of their research literature.

Once the effect sizes are calculated, they are analyzed, the sixth step of the process, using either fixed- or random-effects models. Fixed-effects models, which are the most common analysis used, assume that there is one underlying, but unknown, effect size and that study estimates of this effect size vary only in sampling error. Random-effects models assume that each effect size is unique and that the study is drawn at random from a universe of related but separate effects (see Hedges and Vivea 1998 for a discussion). The general steps involved in the analysis of effect sizes usually are: (1) to aggregate effect sizes across the studies to determine the overall strength of the relation between the examined variables; (2) to analyze the consistency of the effect sizes across the studies; (3) to diagnose outliers among the effect sizes; and (4) to perform tests of whether study attributes moderate the magnitude of the effect sizes. Although several frameworks for modeling effect sizes have been developed (for reviews, see Johnson et al. 1995; Sánchez-Meca and Marín-Martínez 1997), the Hedges and Olkin fixed-effect approach (1985) appears to be the most popular and therefore will be assumed in the remainder of this discourse. These statistics were designed to take advantage of the fact that studies have differing variances by calculating the nonsystematic variance of the effect sizes analytically (Hedges and Olkin 1985). Because this nonsystematic variance of an effect size is inversely proportional to the sample size of the study and because sample sizes typically vary widely across the studies, the error variances of the effect sizes are ordinarily quite heterogeneous. These meta-analytic statistics also permit an analysis of the consistency (or homogeneity) of the effect sizes across the studies, a highly informative analysis not produced by conventional, primary-level statistics. As the homogeneity calculation illustrates, analyzing effect sizes with specialized meta-analytic statistics rather than the ordinary inferential statistics used in primary research allows a reviewer to use a greater amount of the information available from the studies (Rosenthal 1991, 1995).

As a first step in a quantitative synthesis, the study outcomes are combined by averaging the effect sizes with each weighted by its sample size. This procedure gives greater weight to the more reliably estimated study outcomes, which are in general those with the larger sample sizes (see

Hedges et al. 1992; Johnson et al. 1995). As a test for significance of this weighted mean effect size, a confidence interval is typically computed around this mean, based on its standard deviation,  $d_{\pm} \pm 1.96 \sqrt{v_{\pm}}$ , where 1.96 is the unit-normal value for a 95 percent confidence interval (CI) (assuming a nondirectional hypothesis). If the CI includes zero (0.00), the value indicating exactly no difference, it may be concluded that aggregated across all studies there is no significant association between the independent and dependent variables (X and Y). For example, Perkins and colleagues (1995) found no evidence across thirteen studies that silicone breast implants increased risk of connective tissue disease. In a different literature, He and colleagues (1999) found that, across eighteen studies, nonsmokers exposed to passive smoke had a higher relative risk of coronary heart disease than nonsmokers not exposed to smoke.

Once a meta-analysis has derived a weighted mean effect size, it, and other meta-analytic statistics, must be interpreted and presented, which is the seventh step of conducting a meta-analysis. If the mean effect is nonsignificant and the homogeneity statistic is small and nonsignificant, an analyst might conclude that there is no relation between the variables under consideration. However, in such cases, it is wise to consider the amount of statistical power that was available: If the total number of research participants in the studies integrated was small, it is possible that additional data would support the existence of the effect. Even if the mean effect is significant and the homogeneity statistic is small and nonsignificant, concerns about its magnitude arise. To address this issue, Cohen (1969, 1988) proposed some guidelines for judging effect magnitude, based on his informal analysis of the magnitude of effects commonly yielded by psychological research. Cohen intended “that medium represent an effect of a size likely to be visible to the naked eye of a careful observer” (1992, p. 156). He intended that small effect sizes be “noticeably smaller yet not trivial” (p. 156) and that large effect sizes “be the same distance above medium as small is below it” (p. 156). As Table 1 shows, a “medium” effect turned out to be about  $d = 0.50$  and  $r = .30$ , equivalent to the difference in intelligence scores between clerical and semiskilled workers. A “small” effect size was about  $d = 0.20$  and  $r = .10$ , equivalent

**Cohen's (1969) Guidelines for Magnitude of  $d$  and  $r$**

Size	Effect size metric		
	$d$	$r$	$r^2$
Small	0.20	.10	.01
Medium	0.50	.30	.09
Large	0.80	.50	.25

Table 1

to the difference in height between 15- and 16-year-old girls. Finally, a large effect was about  $d = 0.80$  and  $r = .50$ , equivalent to the difference in intelligence scores between college professors and college freshmen.

Another popular way to interpret mean effect sizes is to derive the equivalent  $r$  and square it. This procedure shows how much variability would be explained by an effect of the magnitude of the mean effect size. Thus, a mean  $d$  of 0.50 produces an  $R^2$  of .09. However, this value must be interpreted carefully because  $R^2$ , or variance explained, is a directionless effect size. Therefore, if the individual effect sizes that produced the mean effect size varied in their signs (i.e., if the effect sizes were not all negative or all positive), the variance in  $Y$  explained by the predictor  $X$ , calculated for each study and averaged, would be larger than this simple transformation of the mean effect size. Thus, another possible procedure consists of computing  $R^2$  for each individual study and averaging these values.

When the weighted mean effect size and the CI are computed, the homogeneity of the  $d$ 's is statistically examined, in order to determine whether the studies can be adequately described by a single effect size (Hedges and Olkin 1985). If the effect sizes can be so described, then they would differ only by unsystematic sampling error. If there is a significant fit statistic, the weighted mean effect size may not adequately describe the outcomes of the set of studies because it is likely that quite different mean effects exist in different groups of studies. Further explanatory work would be merited, even when the composite effect size is significant. The magnitude of individual study outcomes would differ systematically, and these



differences may include differences in the direction (or sign) of the relation. In some studies, the independent variable might have had a large positive effect on the dependant variable, and in other studies, it might have had a smaller positive effect or even a negative effect. Even if the homogeneity test is nonsignificant, significant moderators could be present, especially when the fit statistic is relatively large (for further discussions, see Johnson and Turco 1992; Rosenthal 1995). Nonetheless, in a meta-analysis that attempts to determine the relation of one variable to another, rejecting the hypothesis of homogeneity could be troublesome, because it implies that the association between these two variables likely is complicated by the presence of interacting conditions. However, because analysts usually anticipate the presence of one or more moderators of effect size magnitude, establishing that effect sizes are not homogeneous is ordinarily neither surprising nor troublesome.

To determine the relation between study characteristics and the magnitude of the effect sizes, both categorical models and continuous models can be tested. In *categorical models*, analyses may show that weighted mean effect sizes differ in magnitude between the subgroups established by dividing studies into classes based on study characteristics. In such cases, it is as though the meta-analysis is broken into sub-meta-analyses based on their methodological features. For example, He and colleagues (1999) found that risk of coronary heart disease was greater for women than for men nonsmokers who were exposed to passive smoke. If effect sizes that were found to be heterogeneous become homogeneous within the classes of a categorical model, the relevant study characteristic has accounted for the systematic variability between the effect sizes. Similarly, *continuous models*, which are analogous to regression models, examine whether study characteristics that are assessed on a continuous scale are related to the effect sizes. As with categorical models, some continuous models may be completely specified in the sense that the systematic variability in the effect sizes is explained by the study characteristic that is used as a predictor. Continuous models are least squares regressions, calculated with each effect size weighted by the reciprocal of its variance (sample size). For example, He and colleagues (1999) found that risk of coronary heart disease was greater to the extent that nonsmokers had greater exposure to passive

smoke. Goodness-of-fit statistics enable analysts to determine the extent to which categorical or continuous models, or mixtures of these models provide correct depictions of study outcomes.

As an alternative analysis to predicting effect sizes using categorical and continuous models, an analyst can attain homogeneity by identifying outlying values among the effect sizes and sequentially removing those effect sizes that reduce the homogeneity statistic by the largest amount (e.g., Hedges 1987). Studies yielding effect sizes identified as outliers can then be examined to determine whether they appear to differ methodologically from the other studies. Also, inspection of the percentage of effect sizes removed to attain homogeneity allows one to determine whether the effect sizes are homogeneous aside from the presence of relatively few aberrant values. Under such circumstances, the mean attained after removal of such outliers may better represent the distribution of effect sizes than the mean based on all the effect sizes. In general, the diagnosis of outliers should occur prior to calculating moderator analyses; this diagnosis may locate a value or two that are so discrepant from the other effect sizes that they would dramatically alter any models fitted to effect sizes. Under such circumstances, these outliers should be removed from subsequent phases of the data analysis. Alternatively, outliers can be examined following categorical or continuous models (e.g., finding those that deviate the most from the values predicted by the models).

#### TRENDS IN THE PRACTICE OF META-ANALYSIS

Although the quality of meta-analyses has been quite variable, it is possible to state the features that compose a high-quality meta-analysis, including success in locating studies, explicitness of criteria for selecting studies, thoroughness and accuracy in coding moderators variables and other study characteristics, accuracy in effect size computations, and adherence to the assumptions of meta-analytic statistics. When meta-analyses meet such standards, it is difficult to disagree with Rosenthal's conclusion (1994) that it is "hardly justified to review a quantitative literature in the pre-meta-analytic, prequantitative manner" (p. 131). Yet merely meeting these high standards does not

necessarily make a meta-analysis an important scientific contribution. One factor affecting scientific contribution is that the conclusions that a research synthesis is able to reach are limited by the quality of the data that are synthesized. Serious methodological faults that are endemic in a research literature may well handicap a synthesis, unless it is designed to shed light on the influence of these faults. Also, to be regarded as important, the review must address an interesting question. Moreover, unless the paper reporting a meta-analysis “tells a good story,” its full value may go unappreciated by readers. Although there are many paths to a good story, Sternberg’s recommendations to authors of reviews (1991) are instructive: Pick interesting questions, challenge conventional understandings if at all possible, take a unified perspective on the phenomenon, offer a clear take-home message, and write well. Thus, the practice of meta-analysis should not preclude incorporating aspects of narrative reviewing, but instead should strive to incorporate and document the richness of the literature.

One reason that the quality of published syntheses has been quite variable is that it is a relatively new tool among scholars who practice it. Yet, as the methods of quantitative synthesis have become more sophisticated and widely disseminated, typical published meta-analyses have improved. At their best, meta-analyses advance knowledge about a phenomenon by explicating its typical patterns and showing when it is larger or smaller, negative or positive, and by testing theories about the phenomenon (see Miller and Pollock 1994). Meta-analysis should foster a healthy interaction between primary research and research synthesis, at once summarizing old research and suggesting promising directions for new research. One misperception that scholars sometimes express is that a meta-analysis represents a dead end for a literature, a point beyond which nothing more needs to be known. In contrast, carefully conducted meta-analyses can often be the best medicine for a literature, by documenting the robustness with which certain associations are attained, resulting in a sturdier foundation on which future theories may rest. In addition, meta-analyses can show where knowledge is at its thinnest, to help plan additional, primary-level research (see Eagly and Wood 1994). As a consequence of a carefully conducted meta-analysis, primary-level studies can

be designed with the complete literature in mind and will therefore have a better chance of contributing new knowledge. In this fashion, scientific resources can be directed most efficiently toward gains in knowledge. As time passes and new studies continue to accrue rapidly, it is likely that social scientists will rely more on quantitative syntheses to inform them about the knowledge that has accumulated in their research. Although it is possible that meta-analysis will become the purview of an elite class of researchers who specialize in research integration, as Schmidt (1992) argued, it seems more likely that meta-analysis will become a routine part of graduate training in many fields, enabling students to develop the skills necessary to ply the art and science of meta-analysis and to integrate findings across studies as a normal and routine part of their research activities.

## REFERENCES

- Bond, R., and P. B. Smith 1996 “Culture and Conformity: A Meta-Analysis of Studies Using Asch’s (1952b, 1956) Line Judgment Task.” *Psychological Bulletin* 119:111–137.
- Campbell, D. T., and J. T. Stanley 1963 *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Cohen, J. 1969 *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic.
- 1988 *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, N.J.: Erlbaum.
- 1992 “A Power Prime.” *Psychological Bulletin* 112:155–159.
- Cook, T. D., and D. T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.
- Cooper, H. 1998 *Integrative Research: A Guide for Literature Reviews*, 3rd ed. Newbury Park, Calif.: Sage.
- , and L. V. Hedges 1994a. “Research Synthesis as a Scientific Enterprise.” Pp. 3–14 in H. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis*. New York: Russell Sage.
- , and L.V. Hedges, (eds.) 1994b *The Handbook of Research Synthesis*. New York: Russell Sage.
- , and Rosenthal, R. (1980). “Statistical versus Traditional Procedures for Summarizing Research Findings.” *Psychological Bulletin* 87:442–449.
- Eagly, A. H., and L. Carli 1981 “Sex of Researchers and Sex-Typed Communications as Determinants of Sex

- Differences in Influenceability: A Meta-Analysis of Social Influence Studies." *Psychological Bulletin* 90:1-20.
- , and W. Wood 1994 "Using Research Syntheses to Plan Future Research." Pp. 485-500 in H. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis*. New York: Russell Sage.
- Ernst, C., and J. Angst 1983 *Birth Order: Its Influence on Personality*. New York: Springer-Verlag.
- Fisher, R. A. 1921 "On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample." *Metron* 1:1-32.
- Glass, G. V. 1976 "Primary, Secondary, and Meta-Analysis of Research." *Educational Researcher* 5:3-8.
- , B. McGraw, and M. L. Smith 1981 *Meta-Analysis in Social Research*. Beverly Hills, Calif.: Sage.
- He, J., S. Vupputuri, K. Allen, M. R. Prerost, J. Hughes, and P. K. Whelton 1999 "Passive Smoking and the Risk of Coronary Heart Disease—A Meta-Analysis of Epidemiologic Studies." *New England Journal of Medicine* 340:920-926.
- L. V. Hedges 1981 "Distribution Theory for Glass's Estimator of Effect Size and Related Estimators." *Journal of Educational Statistics* 6:107-128.
- 1987 "How Hard Is Hard Science, How Soft Is Soft Science? The Empirical Cumulativeness of Research." *American Psychologist* 42:443-455.
- 1994 "Statistical Considerations." Pp. 29-38 in H. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis*. New York: Russell Sage.
- , H. Cooper, and B. J. Bushman 1992 "Testing the Null Hypothesis in Meta-Analysis: A Comparison of Combined Probability and Confidence Interval Procedures." *Psychological Bulletin* 111:188-194.
- , and I. Olkin 1985 *Statistical Methods for Meta-Analysis*. Orlando, Fla.: Academic.
- , and J. L. Vevea 1998 "Fixed- and Random-Effects Models in Meta-Analysis." *Psychological Methods* 3:486-504.
- Hunt, M. 1997 *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage.
- Hunter, J. E., and F. L. Schmidt 1990 *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park, Calif.: Sage.
- Johnson, B. T. 1993 *DSTAT 1.10: Software for the Meta-Analytic Review of Research Literatures*. Hillsdale, N.J.: Erlbaum.
- , B. Mullen, and E. Salas 1995 "Comparison of Three Major Meta-Analytic Approaches." *Journal of Applied Psychology* 80:94-106.
- , and D. R. Nichols 1998 "Social Psychologists' Expertise in the Public Interest: Civilian Morale Research during World War II." *Journal of Social Issues* 54:53-77.
- , and R. Turco 1992 "The Value of Goodness-of-Fit Indices in Meta-Analysis: A Comment on Hall and Rosenthal." *Communication Monographs* 59:388-396.
- Mazela, A., and M. Malin 1977 *A bibliometric Study of the Review Literature*. Philadelphia: Institute for Scientific Information.
- Miller, N., and V. E. Pollock 1994 "Meta-Analysis and Some Science-Compromising Problems of Social Psychology" Pp. 230-261 in W. R. Shadish and S. Fuller, eds., *The Social Psychology of Science*. New York: Guilford.
- Olkin, I. 1990 "History and Goals." In K. W. Wachter and M. L. Straf, eds., *The Future of Meta-Analysis*. New York: Russell Sage.
- Perkins, L. L., B. D. Clark, P. J. Klein, and R. R. Cook 1995 "A Meta-Analysis of Breast Implants and Connective Tissue Disease." *Annals of Plastic Surgery* 35:561-570.
- Rosenthal, R. 1990 "How Are We Doing in Soft Psychology?" *American Psychologist* 45:775-777.
- 1991 *Meta-Analytic Procedures for Social Research*, rev. ed. Beverly Hills, Calif.: Sage.
- 1994 "Parametric Measures of Effect Size." Pp. 231-244 in H. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis*. New York: Russell Sage.
- 1995 "Writing Meta-Analytic Reviews." *Psychological Bulletin* 118:183-192.
- , and D. Rubin 1978 "Interpersonal Expectancy Effects: The First 345 Studies." *Behavioral and Brain Sciences* 3:377-415.
- Sánchez-Meca, J., and Marín-Martínez, F. 1997 "Homogeneity Tests in Meta-Analysis: A Monte-Carlo Comparison of Statistical Power and Type I Error." *Quality & Quantity* 31:385-399.
- Schmidt, F. L. 1992 "What Do Data Really Mean? Research Findings, Meta-Analysis, and Cumulative Knowledge in Psychology." *American Psychologist* 47:1173-1181.
- Shadish, W. R. 1996 "Meta-Analysis and the Exploration of Causal Mediating Processes: A Primer of Examples, Methods, and Issues." *Psychological Methods* 1:47-65.
- Sternberg, R. J. 1991 Editorial. *Psychological Bulletin* 109:3-4.
- Stigler, S. M. 1986 *History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, Mass.: Harvard University Press.

White, H. D. 1994 "Scientific Communication and Literature Retrieval." In H. Cooper and L. V. Hedges eds., *The Handbook of Research Synthesis*, pp. 41-55. New York: Russell Sage.

B. T. JOHNSON

## METATHEORY

Metatheory in sociology is a relatively new specialty that aims to describe existing sociological theory systematically, and also, to some degree, to prescribe what future sociological theories ought to be like. It leaves to other specialties—most notably the sociology and history of sociology and the logic of theory construction—the problems of explaining and predicting how such theories have been, and can be, formulated.

There are two broad varieties of metatheory. One variety, *synthetic*, classifies whole theories according to some overarching typology; the other variety, *analytic*, first dissects theories into their underlying constituents and then classifies these constituents into types.

Some typologies encountered in synthetic metatheory refer to the time periods when the theories were originated, for example, forerunner, classical, and contemporary (Timasheff and Theodorson 1976 and Eisenstadt 1976 provide examples). Some refer to the places where the theories were originated, for example, France, Germany, Italy, and the United States (Bottomore and Nisbet 1978 and Gurvitch and Moore 1945 provide examples). Some refer to the substantive themes of the theories, for example, structural-functional, evolutionist, conflict, and symbolic interactionist (Turner 1986 and Collins 1988 provide examples). Some refer to the ideologies supported by the theories, for example, pro-establishment and anti-establishment (Martindale 1979 provides an example). Some refer to various combinations of all the above differences (Wiley 1979 and Ritzer 1983 provide examples).

Analytic metatheory is divisible into two broad classes: one in which the constituents of theories are required to have empirical referents, either directly or indirectly, and another in which these constituents are required or permitted to have nonempirical referents. Thus, one sociologist claims our theory should be brought "closer to nonempirical

standards of objectivity" (Alexander 1982), while another claims "sense-based inter-subjective verification is indispensable [to sociology]" (Wallace 1983). (This difference in kind of analytic metatheory reflects an applicability of the synthetic-analytic distinction to *metamethod* in sociology: In the synthetic variety of metamethod, whole methods are characterized as *empirical* or *nonempirical*, *positivistic* or *hermeneutical*, *experimental* or *participant-observational*, and so on. In the analytic variety, such methods are dissected into their underlying constituents, which are then classified as *measurement*, *interpretation*, *speculation*, *comparison*, *test*, *generalization*, *specification*, *deduction*, *induction*, and so on.)

Some types of underlying constituents encountered in empirical analytic metatheory are "control" (Gibbs 1989); "individual actors," "corporate actors," "interests," and "rights" (Coleman 1990); "rational action," "nonrational action," "individualist order," and "collectivist order" (Alexander 1982); "social and cultural structures," "spatial and temporal regularities," "instinct," "enculture," "physiology," "nurture," "demography," "psychical contagion," "ecology," and "artifacts" (Wallace 1983, 1988); and general causal images like "convergence," "amplification," "fusion," "fission," "tension," "cross-pressure," "dialectic," and "cybernetic" (Wallace 1983, 1988). Some types of underlying constituents encountered in nonempirical analytic metatheory have been called (so far, without further explication or specification), "moral implication," "moral commitments," and "moral preferences" (Alexander 1982).

Metatheory in general has been sweepingly condemned as a dead end leading only to the study of "the grounds of other people's arguments rather than substantive problems" (Skocpol 1987), and as holding "little prospect for further developments and new insights" (Collins 1986). Against such characterizations, however, certain unique and indispensable contributions of both synthetic and analytic metatheory to sociology should not be overlooked.

Synthetic metatheory plays obviously central roles in descriptive classifications of sociological theory (e.g., textbooks and course outlines), but they are no less central to the sociology and history of sociology, where efforts to account for the rise and fall of schools, or perspectives, in sociological analysis require systematic conceptualization of

such groupings. The contributions of nonempirical analytic metatheory remain unclear (as mentioned, the kinds of ideological commitment and moral foundation to which it refers, and their consequences for sociological theory, have yet to be specified) and, therefore, will not be examined here. The contributions of empirical analytic metatheory will occupy the rest of this article.

### THREE CONTRIBUTIONS OF EMPIRICAL ANALYTIC METATHEORY

Empirical analytic metatheory can aid (1) systematic cumulation of the *end product* of sociological investigation (namely, collectively validated empirical knowledge about social phenomena); (2) systematic construction of new versions of the principal *means* employed in generating that end product (namely, collectively shared theory and method); and (3) a sense of discipline-wide *solidarity* among sociologists of all theoretical traditions, all specializations and, eventually one hopes, among all social scientists.

**Cumulation of Sociological Knowledge** Knowledge can only cumulate when new knowledge of a given phenomenon is added to old knowledge of that same phenomenon (or, rather, insofar as no phenomenon is ever repeated exactly, that same *type* of phenomenon). The key to holding such objects of investigation constant is, of course, communication. That is to say, only the communication to investigator B of the identity of the exact phenomenon investigator A has examined, together with the exact results of that examination, can enable investigator B systematically to add new knowledge to A's knowledge.

*Disciplinary Communication* Now it may be imagined that we already possess such communication in sociology, but we do not. Consider the terms *social structure* and *culture*. One can hardly doubt that, by denoting the substantive heart of our discipline, they indicate what the entire sociological enterprise is about. By virtually all accounts, however, each term signifies very different kinds of phenomena to different sociologists.

Thus, *social structure* has been authoritatively said, at various times for over two decades, to be "so fundamental to social science as to render its uncontested definition virtually impossible" (Udy 1968); to attract "little agreement on its empirical

referents" (Warriner 1981); and to possess a meaning that "remains unclear" (Turner 1986). "Few words," it has been said, "do sociologists use more often than 'structure,' especially in the phrase 'social structure.' Yet we seldom ask what we mean by the word" (Homans 1975). In a more detailed statement, one analyst asserts that

*The concept of social structure is used widely in sociology, often broadly, and with a variety of meanings. It may refer to social differentiation, relations of production, forms of association, value integration, functional interdependence, statuses and roles, institutions, or combinations of these and other factors.*  
(Blau 1975)

Indeed, we can still read that "sociologists use the term ['social structure'] in diverse ways, each of which is either so vague as to preclude empirical application or so broad as to include virtually all collective features of human behavior" (Gibbs 1989). As recent evidence of this diversity, it is noteworthy that, where one sociologist claims that "for sociologists, the units of social structure are conceived of . . . as *relational* characteristics" (Smelser 1988), another refers, without explanation, to a type of "social structure" in which the participants "have *no relations*" (Coleman 1990).

The situation is no different with the term *culture*. Some years ago it was said that "by now just about everything has been thrown into 'culture' but the kitchen sink," and the author of this remark then reflected that "The kitchen sink has been thrown in too" as part of "material culture" (Schneider 1973). Years later, it has again been pointed out that "Theorists of culture remain sorely divided on how best to define culture" (Wuthnow et al. 1984) and "values, orientations, customs, language, norms, [and religion]" have been referred to as though they were all somehow different from "culture" (Coleman 1990). No wonder at least one sociologist has simply given up: "[A]ny definition" of culture, he claims, "will be (1) inclusive to the point of being meaningless, (2) arbitrary in the extreme, or (3) so vague as to promise only negligible empirical applicability" (Gibbs 1990).

More recently still, Gilmore affirms that "there is no current, widely accepted, composite resolution of the definition of culture," and claims that as a result "the contemporary concept of culture

in sociology does not exclude any particular forms of [collective] activity.” This difficulty notwithstanding, there has arisen, Gilmore says, “a new appreciation of the salience of culture as an explanatory perspective in contemporary sociological research” (1992). To the extent that these judgments are true, one can only wonder what contribution an explanatory perspective that lacks even a rudimentary and tentative definition of its own central variable (and which expresses a kitchen-sink inclusiveness that “does not exclude any particular empirical forms of activity”) can possibly make to social science.

Regarding social structure, Rytina asserts his conviction that “Social structure is a general term for any collective social circumstance that is unalterable and given for the individual” and that such social structure “is the same for all and is beyond the capacity for alteration by any individual will” (1992). Apart from noticing, again, the kitchen-sink inclusiveness of this claim (“*any* collective social circumstance”), one wonders what good can come of conceptualizing social structure in a way that rules out the possibility of variable individual *power* (i.e., “the probability that one actor within a social relationship will be in a position to carry out his own will” [Weber 1978]). See, however, Sewell (1992) for explicit inclusion of “power” in structure—although Sewell lumps “resources” and “schemas” (i.e., what others would distinguish as components of social structure and culture, respectively) together into a single undifferentiated “structure” concept.

Some other discussions of the ongoing social structure versus culture problem may be found in Emirbayer and Goodwin (1994), Hays (1994), Holmwood and Stewart (1994), Wallace (1986), and Whitmeyer (1994). It will also be noticed that in addition to “social structure” and “culture” the concept “agency” appears in some of these discussions, and a brief comment on it may be useful. Human “agency” is said, by Sewell, to refer to “the efficacy of human action,” and to arise from the actor’s “ability to apply [known schemas] to new contexts” and to act “creatively” (1992, pp. 2, 20). It is not easy to understand, however, why a special anthropocentric term is needed for such a phenomenon inasmuch as *all* action, by *whatever* agent, is (by any physical definition of “action”) efficacious, and *all* “applications,” by definition, occur in “new” contexts and are thus “creative.”

In response to such expressions of disciplinary decline (and acknowledging their strong evidential basis), empirical analytic metatheory falls back on Durkheim’s argument that insofar as “Every scientific investigation concerns a specific group of phenomena which are subsumed under the same definition,” it follows that “[the] sociologist’s first step must . . . be to define the things he treats so that we may know—he as well—exactly what his subject matter is. This is the prime and absolutely indispensable condition of any proof or verification” (1982).

Empirical analytic metatheory, then, seeks a common disciplinary language for sociologists everywhere, regardless of their specializations. Its proponents believe that only with the adoption of some such language can our discipline begin solving its central problems, namely, systematic knowledge cumulation, theory innovation, and solidarity enhancement.

(SEE ALSO: *Epistemology*; *Scientific Explanation*)

#### REFERENCES

- Alexander, Jeffrey C. 1982 *Positivism, Presuppositions, and Current Controversies*. Berkeley: University of California Press.
- Blau, Peter M. 1975 “Parameters of Social Structure.” In Peter M. Blau, ed., *Approaches to the Study of Social Structure*. New York: Free Press.
- Bottomore, Tom, and Robert Nisbet 1978 *A History of Sociological Analysis*. New York: Basic.
- Brinkerhoff, David B., and Lynn K. White 1985 *Sociology*. St. Paul Minn.: West.
- Broom, Leonard, Philip Selznick, and Dorothy Broom Darroch 1981 *Sociology*. New York: Harper and Row.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Collins, Randall 1986 “Is 1980s Sociology in the Doll-drums?” *American Journal of Sociology* 91:1336–1355.
- 1988 *Theoretical Sociology*. San Diego: Harcourt Brace Jovanovich.
- Durkheim, Emile 1982 *The Rules of Sociological Method*. New York: Free Press.
- Eisenstadt, S. N., with M. Curelaru 1976 *The Form of Sociology: Paradigms and Crises*. New York: Wiley.
- Emirbayer, Mustafa, and Jeff Goodwin 1994 “Network Analysis, Culture, and the Problem of Agency.” *American Journal of Sociology* 99(6):1411–1454.

- Gibbs, Jack P. 1989 *Control, Sociology's Central Notion*. Urbana: University of Illinois Press.
- Gilmore, Samuel 1992 "Culture." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*, 1st ed. New York: Macmillan.
- Gurvitch, Georges, and Wilbert E. Moore 1945 *Twentieth Century Sociology*. New York: Philosophical Library.
- Hays, Sharon 1994 "Structure and Agency and the Sticky Problem of Culture." *Sociological Theory* 12(1):57-72.
- Holmwood, John, and Alexander Stewart 1994 "Synthesis and Fragmentation in Social Theory: A Progressive Solution," *Sociological Theory* 12(1):83-100.
- Homans, George C. 1975 "What Do We Mean by 'Social Structure'?" In Peter M. Blau, ed., *Approaches to the Study of Social Structure*. New York: Free Press.
- Martindale, Don 1979 "Ideologies, Paradigms, and Theories." In William E. Snizek, Ellsworth R. Fuhrman, and Michael K. Miller, eds., *Contemporary Issues in Theory and Research*. Westport, Conn.: Greenwood.
- Orenstein, David Michael 1985 *The Sociological Quest: Principles of Sociology*. St. Paul, Minn.: West.
- Ritzer, George 1983 *Contemporary Sociological Theory*. New York: Knopf.
- Rytina, Steven L. 1992 "Social Structure." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*, 1st ed. New York: Macmillan.
- Schneider, Louis 1973 "The Idea of Culture in the Social Sciences: Critical and Supplementary Observations." In Louis Schneider and Charles M. Bonjean, eds., *The Idea of Culture in the Social Sciences*. Cambridge, England: Cambridge University Press.
- Sewell, William H., Jr. 1992 "A Theory of Structure: Duality, Agency and Transformation." *American Journal of Sociology* 98(1):1-29.
- Skocpol, Theda 1987 "The Dead End of Metatheory." *Contemporary Sociology* 16:10-12.
- Smelser, Neil J. 1988 "Social Structure." In Neil H. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- 1989 "Reviewing the Field of Sociology: A Response." *Contemporary Sociology* 18:851-855.
- Timasheff, Nicholas S., and George A. Theodorson 1976 *Sociological Theory*, 4th ed. New York: Random House.
- Tumin, Melvin M. 1973 *Patterns of Society*. Boston: Little, Brown.
- Turner, Jonathan 1986 *The Structure of Sociological Theory*. Chicago: Dorsey.
- 1989 "The Disintegration of American Sociology." *Sociological Perspectives* 32:419-433.
- Udy, Stanley H., Jr. 1968 "Social Structure: Social Structural Analysis." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 14. New York: Free Press.
- Wallace, Walter L. 1983 *Principles of Scientific Sociology*. Hawthorne, N.Y.: Aldine.
- Wallace, Walter L. 1986 "Social Structural and Cultural Structural Variables in Sociology." *Sociological Focus* 19(April):125-138
- 1988 "Toward a Disciplinary Matrix in Sociology." In Neil J. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- Warriner, Charles K. 1981 "Levels in the Study of Social Structure." In Peter M. Blau and Robert K. Merton, eds., *Continuities in Structural Inquiry*. Beverly Hills, Calif.: Sage.
- Weber, Max 1978 *Economy and Society*. Berkeley: University of California Press.
- Whitmeyer, Joseph M. 1994 "Why Actor Models Are Integral to Structural Analysis." *Sociological Theory* 12(2):153-165.
- Wiley, Norbert 1979 "The Rise and Fall of Dominating Theories in American Sociology." In William E. Snizek, Ellsworth R. Fuhrman, and Michael K. Miller, eds., *Contemporary Issues in Theory and Research*. Westport, Conn.: Greenwood.
- Wuthnow, Robert, James Davison Hunter, Albert Bergesen, and Edith Kurzweil 1984 Introduction." In Robert Wuthnow et al., eds., *Cultural Analysis*. Boston: Routledge and Kegan Paul.

WALTER L. WALLACE

## MEXICAN STUDIES

Mexican history since the country achieved its independence from Spain in 1821 is marked by a set of dramatic events that created the context for fundamental social and economic changes. These events shaped contemporary Mexican society and established much of the current agenda of Mexican social and political studies. Three issues in particular have long historical antecedents. These are the ethnic question, the role of the state in development, and relations with the United States. Despite these continuities, social trends and pre-occupations in Mexico have also differed according to historical period. In order to handle both

continuity and change, we will organize the discussion in terms of four main periods, each initiated by major events, and will consider the dominant trends and intellectual orientations of each period. The time between Mexican independence and the Mexican-American War, the first period is marked by economic stagnation and political fragmentation. The second period lasts from the aftermath of the Mexican-American War until the Mexican Revolution of 1910. Rapid but uneven economic development and political unification distinguish the second period. The third period begins with the consolidation of the Mexican Revolution and goes through the 1920s and on into the debt crises of the early 1980s. Nationalism, political centralization, and high rates of urbanization and industrialization characterize this period. Finally, we consider the two last decades of the twentieth century, shaped by the external opening of the economy with its accompanying economic and political liberalization.

This historical context and the unevenness of Mexico's development make U.S.-Mexican relations different, in significant respects, from relations between other neighboring countries. The U.S.-Mexican border is the only international frontier that divides a developing country and a highly developed country. The contradictory trends of Mexico's development shape U.S.-Mexican relations (Weintraub 1990). The two populations have been brought closer to each other in terms of trade, tourism, labor migration, and consumer aspirations. This increasing proximity can also create social distance as each population develops opposing images of the other. These are often based on the misunderstandings that result from wide disparities in income, standards of living, and political and family culture (Pastor and Castañeda 1988). For many Mexicans, the history of conflictive relations with the United States continues to negatively color perceptions of U.S. political and economic intentions toward Mexico.

Independence from Spain left Mexico an impoverished and fragmented nation. The main sources of external revenue—mining and specialized plantation agriculture—declined substantially. Internal markets were not strongly developed and were mainly concentrated in a few major cities. Thus, neither internal nor external trade provided enough stimuli to expand local agricultural or

industrial production. The economy was organized on a local and regional basis rather than on a national one. Within this context, politics was dominated by a decentralized system of *caudillos* (regional leaders, often military) and *caciques* (local bosses and power brokers). The formal government alternated between a centralist and a federalist republic, after the short-lived Empire of Iturbide. The proponents of a centralist republic tended to be conservatives, supporting the maintenance of corporatist institutions such as the church, the army, the guilds, and the indigenous communities. In a sociological account of the processes shaping the Independence movement, the intellectual leader of the conservatives, Lucas Alamán ([1847] 1942), advocated a strong central power, economic protectionism, the remedying of class inequality, and populating the empty spaces to the north through European Catholic colonization. In contrast, the federalists tended to be economically and politically liberal, arguing for individual property rights, free markets, and individual civil and political equality. The radical liberal parliamentarian, José María Luis Mora, went so far as to argue for the complete incorporation of the indigenous population into the market as a means to remedy backwardness and poverty (Lira Gonzalez 1984). A leading liberal politician and intellectual, Lorenzo de Zavala, admired the United States and considered it a model to emulate. He also viewed favorably the colonization of the northern territories by Protestant colonists from the United States (Lira Gonzalez 1984). Later, Zavala became the first vice president of the independent Republic of Texas. In order to govern, both the federalists and centralists relied on the *caudillo-cacique* power structure. Thus, the paradigmatic *caudillo* of this period, Santa Anna, was president several times, under both centralist and federalist banners.

The loss of half of the national territory, first in the Texas War of Independence of 1835–1836 and then in the Mexican-American War of 1846–1848, demonstrated dramatically the costs of Mexico's political and economic fragmentation. The conservatives entrenched themselves further in their corporate and hierarchical beliefs, leading finally to the ill-fated, conservative-backed empire of Maximilian of Hapsburg (1864–1867). The aftermath of the Mexican-American war was perhaps more traumatic for the liberals. They saw themselves as betrayed by the United States, which



they had regarded as a friend and an ally in modernizing Mexico. Thereafter, particularly with the experiences of the French intervention, Mexican liberals, while supporting market and individual freedoms and the abolition of corporations, were to become the main proponents of a strong state and of nationalism in Mexico.

Beginning with Benito Juárez (president, 1861–1863 and 1867–1872), liberals were to dominate the politics of Mexico until the Mexican Revolution, even in the paradoxical guise of the dictator, Porfirio Díaz (president, 1877–1880, 1884–1911). In face of the difficulties of achieving spontaneous economic and social progress, liberalism was modified, however, in an authoritarian and positivist direction. A disciple of Comte, Gabino Barreda, reorganized the educational system of the country to ensure a uniform and centralized system (Zea 1974). Justo Sierra (1969 [1901–1903]), one of Díaz's ministers, wrote a new version of Mexican history as a political evolution that needed a period in which political rights would be limited. He saw economic progress as depending on civic education, as well as individualizing property rights in agriculture, industrial modernization, foreign investment, and trade. Though corporate rights to land had been abolished under Juárez, Díaz made this effective, particularly for the indigenous communities. Railways were built and roads improved. Political control was centralized through a bureaucratic system of patronage that gave the central government a more effective control of state and local governments. For the first time in fifty years, Mexico experienced a long period of peace and stability.

The Díaz period was one of considerable economic expansion in Mexico, but the style of development resulted in a series of social and economic dislocations that were to contribute to the Mexican Revolution of 1910. Commercial haciendas expanded, depriving many peasants of access to land. The new economic opportunities and improvements in communication affected communities throughout Mexico, encouraging villagers to give up traditional livelihoods for wage labor in mines, plantations, or as sharecroppers. As dependence on the market grew, so too national and regional economic downturns caused increasing discontent and hardship. Advocates of progress, such as Andrés Molina Enríquez ([1909] 1978), one of the intellectual leaders of the Mexican

Revolution, denounced the economic inequities produced by land concentration. Instead, he placed his hopes on the emerging *mestizo* middle class of medium-scale farmers, merchants, manufacturers, and professionals. With Molina Enríquez, there is a clear statement of the ethnic basis of social change. Both the European-origin upper class and the indigenous peasantry are seen as incapable of progress. He sees the *mestizo* as the crucial Mexican identity capable of achieving economic development and social equity. By the time of the Mexican Revolution, the indigenous population, defined in census terms by "race" and non-Spanish-language use, constituted nearly half the Mexican population.

In spite of political centralization, the economy remained regionalized during the Díaz regime. The Mexican Revolution was essentially a set of regional movements and leaders, that responded to the particular economic problems and possibilities that the Porfirian expansion brought to local economies. These regional contexts included the sugar economy of Morelos; the *henequen* economy of Yucatan; the sharecropping, ranching, and mining economy of Chihuahua and Coahuila; and the emerging commercial farming economy of Sonora (Knight 1986).

The years from after the Mexican Revolution until approximately 1980 transformed Mexico from a society in which three-quarters of the population was agricultural and lived in small villages, to one in which less than a quarter of the population worked in agriculture and most people lived in large cities. This transformation was shaped by two outcomes of the revolution. One was an agrarian reform that destroyed the *hacienda* system, replacing it with small and medium-scale commercial farming and a system of social property (the *ejido*) that reinstated and, at times, extended the rights to land of peasant communities. In the first thirty years after the Revolution, Mexican peasant agriculture was sufficiently strong to feed the growing urban population and help the country's industrialization.

The second major legacy of the revolution was the consolidation of a centralized, modern political system around the official party of the revolution, known as the PRI (*Partido Revolucionario Institucional*) since 1948. The PRI based its control on corporate affiliations by industrial and peasant

unions. Also, the massive extension of state employment in education, health, social and administrative services modernized the country and served to create an extensive web of government patronage (Eckstein 1977). Part of the program of social incorporation was the movement known as *Indigenismo*, which was founded by the anthropologist Manuel Gamio ([1935] 1987). Though the movement sought to protect the indigenous population, its final objective was to incorporate them into the dominant *mestizo* culture through education, health campaigns, and local development programs sponsored by specialized government agencies. By 1940, the indigenous population, defined in terms of those who speak an Indian language, had declined according to the census of that year to just over 14 percent of the population. In general, the indigenous population remained poorer than the *mestizo* peasant population.

Economically, the PRI was nationalist, protecting local industries and restricting foreign investment. The state intervened in the economy, managing industries in sectors which were considered vital to the national interest, such as the oil industry, railroads, certain branches of transport, and telephones. However, it tolerated and, at times, promoted private ownership in many industrial sectors, by including entrepreneurs in the corporate alliances through which the PRI governed. For Mexico, the 1940–1970 decades were the “golden years” of development, when annual rates of gross domestic product (GDP) growth of 6 percent and more were common (Leopoldo Solís, 1970). It was also a period in which a strong political nationalism accompanied economic nationalism. Under various presidents, Mexico made clear its independence from U.S. political and economic influence, as in the nationalization of oil and as in its continuing relations with Castro’s Cuba.

The understanding of these processes was aided by the development of social sciences in Mexico, stimulated by an interest in issues of national identity and culture in face of rapid modernization. Anthropology developed under the auspices of official nationalism in the National Institute of Anthropology and History and the National Indigenist Institute, promoting the restoration and preservation of the pre-Hispanic archaeological patrimony as symbols of Mexican nationality (Vazquez 1995). Gamio’s ([1931] 1969) interest in national identity led him to document

the living and working conditions of Mexican labor migrants in the United States. He and the American anthropologist Paul Taylor (1928–1934) demonstrated that the experiences of these migrants would often reinforce Mexican nationalism through ill treatment, prejudice, and summary deportations.

Within Mexico, anthropologists undertook empirical studies of acculturation and cultural integration, viewed as irreversible trends (Aguirre Beltrán 1957). By the 1970s, sociologists and political scientists working at the National University and at El Colegio de México were questioning the Mexican style of development, focusing on issues of class and ethnic inequality (Stavenhagen 1969; Instituto de Investigaciones Sociales 1970–1973; Benítez Zenteno 1977). Some openly questioned the institutional authoritarianism of the political system and the vertical control that the government exercised over the peasantry and industrial workers through the PRI’s corporatist structure (González Casanova, 1972). The critical tone of Mexican social science increased after the student movements of 1968 and their repression by the government.

The anthropologists and economists who concentrated on the rural sector showed that, despite the propeasant rhetoric of the Mexican Revolution and of its successor governments, the lot of the small, subsistence-oriented farmer—the peasant farmer—did not significantly improve (Bartra 1974; Esteva 1983; Warman 1980; Hewitt de Alcantara 1976). Mexican and foreign social scientists demonstrated that agrarian reform had only limited success in creating employment opportunities in agriculture, in diversifying the rural economy, and in stimulating development of rural market and service centers. The expense of producing efficiently for high-risk markets deterred many members of the *ejido* from continuing in agriculture (de la Peña 1981). The deterrence increased with the advent of new technological packages that offered substantially improved yields, but which indebted the small producer and created dependence on state bureaucracy or on private sector agro-industries (Barkin 1990). Agricultural modernization in Mexico meant the displacement of the peasant farmer and the increasing importance of commercial cultivation, often linked to agro-industry (Arroyo 1989; Sanderson 1986).

One major factor that changed the pattern of political incorporation was the extension of government bureaucracy from the 1960s onward (Grindle 1977). In agriculture, this resulted in a considerable increase in government personnel administering various programs of rural development, credit, and technical advice. Direct contacts with government officials enabled peasants to bypass local power brokers, thus weakening the broker's position; but it created a new issue, the interface between bureaucracy and the peasant. Bureaucratic officials implemented central government policy but pursued their own career goals, and both objectives often came into conflict with the strategies of peasant farmers.

Demographic factors also conditioned the processes of social change, geographical mobility, and concentration of wealth. The Mexican population has had high rates of growth as a result of significant declines in mortality without offsetting declines in fertility (Alba and Potter 1982). High rates of population growth combined with the lack of economic opportunities in the countryside result in substantial rural-urban and international migration (Massey et al. 1987; Arroyo 1986). The urban growth rate of Mexico between 1940 and 1980 averaged 4.7 percent a year, with approximately a third of urban growth due to migration from rural areas. This rapid increase and the high urban primacy that resulted in Mexico City having a population six times larger than the next-largest city, Guadalajara, attracted the attention of many researchers, often in collaboration with U.S. or European social scientists. Led by Luis Unikel (1977), researchers at the Colegio de Mexico focused on the "urban bias" of Mexican development policies in which subsidies encouraged urban concentration. Garza (1985) showed that without these subsidies the profitability of Mexico City enterprises would have been less than that of enterprises in smaller cities. Other researchers focused on rural-urban migration showing the high social mobility of rural migrants in the urban labor markets, despite the importance of the informal economy (Balán et al. 1973; Muñoz et al. 1977; Escobar 1986). However, other studies looked at the social and political implications of this rapid urban growth in the spread of squatter settlements, the aggravation of poverty, and the beginnings of urban social movements in demand for

basic services (Montaño 1976; Lomnitz 1977; Alonso 1980; Cornelius 1985; Ward 1998).

The last period is that following the debt crises of the early 1980s. These crises reflected the exhaustion of the import-substituting model of development in Mexico. The national industries were not competitive internationally and Mexico desperately needed foreign investment to continue its modernization. From the mid-1980s onward, under strong pressure from international agencies and the United States, the Mexican government adopted a free market approach to development, marked by its accession to the General Agreement on Tariffs and Trade (GATT) in 1985 and its joining with the United States and Canada in the North American Free Trade Agreement (NAFTA) in 1993. Commercial interests in Mexico saw *ejidal* property rights as discouraging foreign and national investment in agriculture and in the cities where *ejidal* land hemmed in city boundaries. In 1992, laws were passed that permitted the private sale of *ejidal* land, thus effectively dismantling the agrarian reform system (de Janvry et al. 1997).

Accompanying these economic policies was a political liberalization, which included electoral reform and a greater role for opposition parties. The regime's new economic policies undermined much of its corporate political support. The abolition of the *ejido* effectively weakened the power of PRI's National Peasant Confederation in the countryside. Deregulation, particularly of labor markets, weakened the main industrial unions. The political challenge to the PRI became greater in this period because economic liberalization also created economic instability (Otero 1996). Income inequality grew and employment became unstable, as indicated by the growth of the urban informal economy.

The growth of the urban population continued, but less rapidly than earlier as a result of declines in fertility and in a diminishing pool of rural migrants (Demos 1997). According to the population count of 1995, 63 percent of Mexico's 91 million inhabitants live in places of more than 10,000 people. The fastest growing cities in the period 1990–1995 (3.1 percent growth annually) were the ones that began the period with between 100,000 and 1 million inhabitants (Solís 1997). An important component of the new population dynamic is a gradual change in the pattern of internal

migration in Mexico. From a predominantly rural-urban migration focusing mainly on Mexico City, Mexican migration patterns have become increasingly interurban, with a movement of population from the center and south of the country toward the cities of the north (Lozano et al. 1997).

The changes in migration patterns reflect the changing economic geography of Mexico resulting from the development of export-oriented manufacturing mainly in the north of the country along or close to the U.S. border. By 1999, the in bond industry (*maquiladora*) sector generated over a million jobs, mainly in the cities of the north. In contrast, in the 1980s manufacturing lost jobs in the old industrial regions of the center of the country, although industrial restructuring in the Federal District of Mexico in the 1990s has reasserted Mexico City's importance in industrial production. The industrial transformation is a complex one. There are assembly-line operations that use cheap, mainly female labor (Fernández-Kelly 1983; Young and Fort 1994). There are also plants with high levels of technology that require workers with high levels of qualification and that are prepared to offer attractive conditions of work. The Mexican car industry has shifted to northern Mexico as part of an integrated North American production system, with Mexico shipping cars as well as parts to the U.S. market (Carrillo 1989). These processes are creating a distinct border region straddling the frontier, neither fully Mexican nor fully American (Bustamante 1981; Tamayo and Fernandez 1983).

The Mexican economy is not only highly regionalized and predominantly urban, but is increasingly based on service employment (García 1988). By 1995, the majority of employment (53 percent) was in services and commerce, with manufacturing and construction contributing 21 percent and agriculture 25 percent (Pacheco 1997). The same source shows that these shifts in sectoral employment have also been accompanied by increasing rates of female economic participation, growing from 21.5 percent in 1979 to 34.5 percent in 1995. Also, after 1979, the employment categories associated with the urban informal economy—self-employed and unpaid family workers—increased from 33.7 percent to 38.3 percent of employment, whereas employees declined from 62.9 percent to 57.2 percent (Pacheco 1997).

New intellectual tendencies are emerging in face of these realities. There is now less emphasis on rural studies than in the past, although there is a growing focus on the consequences for rural areas of economic integration with the United States, including international migration (Cornelius and Myhre 1998). Mexican and U.S. scholars are studying the impact of the U.S. market on production and commercialization in Mexico with respect to crops such as avocados, mangos, and tomatoes (González Chávez 1991). Particular focuses are the difficulties faced by small-scale farmers, given the capital and technology needed to compete in the U.S. market. Also, there is a greater concern with environmental issues than in the past as market development puts greater competitive pressures on local agriculture or creates incentives to develop ecologically fragile areas for tourism (Tudela 1989; Moreno 1992; Toledo 1994).

Studies of international migration have become increasingly numerous, showing how migration to the United States has become the major means by which poor rural households complement inadequate incomes, attracting even the skilled and better educated (Massey et al. 1994). The existence of large Mexican-origin communities in cities throughout the United States, but particularly in Los Angeles and other cities of the Southwest, provides bridgeheads facilitating immigration. The importance of the U.S. migration experience has led scholars to document extensively the life histories of migrants, sometimes explicitly comparing their experiences with those recorded by Gamio and Taylor in the 1930s (Durand 1996). Mexican and U.S. researchers have also focused on the significance for U.S.-Mexico relations of transnational migrant communities where migrants settled in the U.S. maintain active links with their Mexican communities of origin, even to the extent of participating in local politics and development projects (Roberts et al. 1999).

In the realm of urban research, there is now considerable attention given to studies of family, gender and female economic participation (Benería and Roldán 1987; García and Oliveira 1997). These studies began with an emphasis on household coping strategies in face of poverty, particularly the pooling of resources and placing more members in the labor market (González de la Rocha

1994; Chant 1991; de la Peña et al. 1990; Selby et al. 1990; Roberts 1989). They have increasingly taken up issues of the changing role of women within the household as they gain a certain economic independence. The researchers question assumptions about the strength of the Mexican family and its ability to cope on its own with the fragmenting pressures coming from economic adjustment and migration. Domestic violence has become an important topic of research. In contrast with these demythifying visions of the Mexican family, other studies that focus on elites and entrepreneurial groups show that family cooperation and hierarchy are crucial for upper-class social mobility and business success, particularly among foreign immigrants and innovators willing to take risks (Lomnitz and Pérez Lizaur 1987; Ramírez 1994; Padilla Dieste 1997).

There has also been considerable research on the impact of international economic integration and “neoliberal” policies on two phenomena: poverty and political participation. The new faces of poverty and inequality have been examined through the analysis of representative surveys and case studies, but also from the perspective of the state’s retreat from welfare and social services (González de la Rocha and Escobar 1991; Cortés and Rubalcava 1991; Boltvinik 1994; Brachet-Márquez 1996; Escobar 1986). The changing political scene has been studied particularly through electoral studies and the analysis of emerging political cultures. Widespread social discontent from the magnitude of fraud in the 1988 elections, as well as the government’s search for international legitimacy in the context of economic globalization, led to the strengthening of opposition parties and the reform of electoral legislation (Molinar 1989, 1992; Azíz and Peschard 1993; Crespo 1993). The process of Mexican democratization is also related to the decline of clientelistic practices, which nowadays proves to be ineffective for handling demands in both urban and rural situations (Middlebrook 1995). The political awakening of an increasingly active civil society is reflected in new social movements that in addition to their demands for services and social justice nowadays challenge the official party and focus on issues of citizenship (Krotz 1996). Middle-class actors, whose standard of living is also declining rapidly, have a considerable presence in such movements (Sergio Tamayo 1994).

The quest for democracy and citizenship is crucial for the understanding of ethnic movements in Mexico. Their importance was dramatically highlighted by the eruption of the Chiapas rebellion on January 1, 1994—the very date of the formal enactment of NAFTA—but urgent ethnic demands have been put forward by dozens of nonviolent indigenous organizations since the 1970s (Campbell 1994; Zárate 1993). Such organizations criticized official Indigenista policies because of their ineffectiveness in eradicating poverty but also because of their assimilationist practices and their blatant disregard for the cultural and political rights of the indigenous population (Arizpe 1978). This critique was reinforced by a new anthropological literature emphasizing persisting ethnicity and cultural resistance instead of acculturation and national homogeneization (Bonfil 1981; Boege 1988). A radical statement of this position argues for the existence of an Indian “deep Mexico”—the real Mexico that has been artificially covered and repressed by authoritarianism and a spurious imitation of European modernity (Bonfil 1996). Even the National Indigenist Institute adopted a discourse of cultural plurality, and the Ministry of Education reinforced its programs of bilingual education (Instituto Nacional Indigenista 1988). In 1991, the National Congress surprisingly passed a constitutional reform that recognized multiculturalism as an essential component of the Mexican nation. If this reform may be explained as related to the government’s search for international legitimacy when NAFTA was being negotiated, it nevertheless opened the way to new indigenous demands and to a widespread discussion of the issue of autonomy for autochthonous peoples, where the participation of Indian representatives and authors is as important as that of non-Indian scholars (Warman and Argueta 1993; Díaz-Polanco 1997; Villoro 1998).

## REFERENCES

- Aguirre Beltrán, Gonzalo 1957 *El proceso de aculturación*. México, D.F.: UNAM.
- Alaman, Lucas (1847) 1942 *Disertaciones*. México, D.F. Editorial Jus.
- Alba, Francisco, and Joseph Potter 1982 “Population and Development in Mexico since the 1940s: An Interpretation.” *Population and Development Review* 12:47–73.

- Alonso, Jorge 1980 *Lucha urbana y acumulación de capital*. México, D.F.: Ediciones de la Casa Chata, CISINAH.
- Arizpe, Lourdes 1978 *El reto del pluralismo cultural*. Mexico City: Instituto Nacional Indigenista.
- Arroyo, Jesús 1986 *Migración a centros urbanos*. Guadalajara: Universidad de Guadalajara.
- 1989 *El abandono rural*. Guadalajara: Universidad de Guadalajara.
- Azíz, Alberto, and Jacqueline Peschard, eds. 1993 *Las elecciones federales de 1991*. Mexico City: Miguel Angel Porrúa.
- Balán, Jorge, Harley Browning, and Elizabeth Jelin. 1973. *Men in a Developing Society: Geographical and Social Mobility in Monterrey, Mexico*. Austin: University of Texas Press.
- Barkin, David 1990 *Distorted Development: Mexico in the World Economy*. Boulder, Colo.: Westview.
- Bartra, Roger 1974 *Estructura agraria y clases sociales en México*. México, D.F.: Ediciones Era.
- Benería, Lourdes, and Marta Roldán 1987 *The Cross-Roads of Class and Gender*. Chicago: University of Chicago Press.
- Benítez Zenteno, Raúl, ed. 1977 *Clases sociales y crisis política en América Latina*. Mexico City: Universidad Nacional Autónoma de México/Siglo Veintiuno Editores.
- Boege, Eckart 1988 *Los mazatecos ante la nación*. Mexico City: Siglo Veintiuno Editores.
- Boltvinik, Julio 1994 "La satisfacción de las necesidades esenciales en México en los setenta y ochenta." In Pablo Pascual Moncayo and José Woldenberg, eds., *Desarrollo, desigualdad y medio ambiente*. Mexico City: Cal y Arena.
- Bonfil, Guillermo 1996 *México profundo*. Austin: University of Texas Press.
- , ed. 1981 *Utopía y revolución*. Mexico City: Nueva Imagen.
- Brachet-Márquez, Viviane 1996 *El pacto de dominación: estado, clase y reforma social en México*. Mexico City: El Colegio de México.
- Bustamante, Jorge A. 1981 "La interacción social en la frontera Mexico-Estados Unidos: Un marco conceptual por la investigación." In Roque González, ed., *La frontera del norte*. México, D.F.: El Colegio de México.
- Campbell, Howard 1994 *Zapotec Renaissance: Ethnic Politics and Cultural Revivalism in Southern Mexico*. Albuquerque: University of New Mexico Press.
- Carrillo, Jorge V., ed. 1989 *La nueva era de la industria automotriz en México*. Tijuana, B.C.: El Colegio de la Frontera Norte.
- Chant, Sylvia 1991 *Women and Survival in Mexican Cities*. Manchester, England: Manchester University Press.
- Cornelius, Wayne 1975 *Politics and the Migrant Poor in Mexico City*. Stanford, Calif.: Stanford University Press.
- , and David Myhre, eds. 1998 *The Transformation of Rural Mexico. Reforming the Ejido Sector*. La Jolla: Center for U.S.-Mexican Studies. San Diego: University of California.
- Cortés, Fernando, and Rosa María Rubalcava 1991 *Autoexplotación forzada y equidad por empobrecimiento*. Mexico City: El Colegio de México (Jornadas, 120).
- Crespo, José Antonio 1994 "Legitimidad política y comportamiento electoral en el Distrito Federal." In Jorge Alonso, ed., *Cultura política y educación cívica*. Mexico City: Miguel Angel Porrúa/Universidad Nacional Autónoma de México.
- De Janvry, Alan, Gustavo Gordillo, and Elisabeth Sadoulet 1997 *Mexico's Second Agrarian Reform. Household and Community Responses*. La Jolla: Center for U.S.-Mexican Studies, University of California, San Diego.
- De la Peña, Guillermo 1981 *A Legacy of Promises. Agriculture, Politics and Ritual in the Morelos Highlands of Mexico*. Austin: University of Texas Press.
- , Juan Manuel Durán, Agustín Escobar, and Javier García de Alba, eds. 1990 *Crisis, conflicto y sobrevivencia: Estudios sobre la sociedad urbana en México*. Guadalajara: Universidad de Guadalajara/CIESAS.
- Díaz-Polanco, Héctor 1997 *La Rebelión Zapatista y la Autonomía*. Mexico City: Siglo Veintiuno.
- Durand, Jorge, ed. 1996 *El norte es como el mar: Entrevistas a trabajadores migrantes en Estados Unidos*. Guadalajara: Universidad de Guadalajara.
- Eckstein, Susan 1977 *The Poverty of Revolution: The State and the Urban Poor in Mexico*. Princeton, N.J.: Princeton University Press.
- Escobar, Agustín 1986 *Con el sudor de tu frente: Mercado de trabajo y clase obrera en Guadalajara*. Guadalajara: El Colegio de Jalisco.
- Esteva, Gustavo 1983 *The Struggle for Rural Mexico*. South Hadley, Mass.: Bergin and Garvey.
- Fernández-Kelly, Maria Patricia 1983 *For We Are Sold: I and My People*. Albany: State University of New York Press.
- Gamio, Manuel (1931) 1969 *The Mexican Immigrant*. New York: Amo.
- (1935) 1987 *Hacia un México nuevo*. México, D.F.: Instituto Nacional Indigenista.
- García, Brígida 1988 *Desarrollo económico y absorción de fuerza de trabajo en México, 1950-1980*. México, D.F.: El Colegio de México.

- , and Orlandina de Oliveira 1994 *Trabajo femenino y vida familiar en México*. Mexico City: El Colegio de México.
- Garza, Gustavo 1985 *El proceso de industrialización en la ciudad de México, 1821–1970*. México, D.F.: El Colegio de México.
- González Casanova, Pablo 1972 *Democracy in Mexico*. London: Oxford University Press.
- González Chávez, Humberto 1991 “Los empresarios en la agricultura de exportación en México: Un estudio de caso.” *European Review of Latin American and Caribbean Studies* 50 (June):87–114.
- González de la Rocha, Mercedes 1994 *The Resources of Poverty: Women and Survival in a Mexican City*. Oxford, England: Blackwell.
- , and Agustín Escobar, eds. 1991 *Social Responses to Mexico's Economic Crisis of the 1980's*. La Jolla: Center for U.S.-Mexican Studies, University of California, San Diego.
- Grindle, Merilee S. 1977 *Bureaucrats, Politicians and Peasants in Mexico: A Case Study of Public Policy*. Berkeley: University of California Press.
- Hewitt de Alcantara, Cynthia 1976 *Modernizing Mexican Agriculture*. Geneva: UNRISD.
- Instituto de Investigaciones Sociales, ed. 1970–1973 *El perfil de México en 1980*, 3 vols. Mexico City: Universidad Nacional Autónoma de México/Siglo Veintiuno Editores.
- Instituto Nacional Indigenista, ed. 1988 *INI. 40 Años*. Mexico City: INI.
- Knight, Alan 1986 *The Mexican Revolution*. Cambridge, England: Cambridge University Press.
- Krotz, Esteban, ed. 1996 *El Estudio de la Cultura Política en México: Perspectivas Disciplinarias y Actores Políticos*. Mexico City: Consejo Nacional para la Cultura y las Artes/CIESAS.
- Lira Gonzalez, Andres 1984 *Espejo de discordias: La sociedad Mexicana vista por Lorenzo de Zavala, J. M. Luis Mora and Lucas Alaman*. México, D.F.: Secretaría de Educación Pública.
- Lomnitz, Larissa Adler 1977 *Networks and Marginality. Life in a Mexican Shantytown*. New York: Academic.
- , and Marisol Pérez-Lizaur 1987 *A Mexican Elite Family, 1820–1980: Kinship, Class, and Culture*. Princeton, N.J.: Princeton University Press.
- Lozano, Fernando, Frank Bean, and Bryan Roberts 1997 “The Interconnectedness of Internal and International Migration.” *Soziale Welt* 12:163–179.
- Massey, Douglas S., Rafael Alarcón, Jorge Durand, and Humberto González 1987 *Return to Aztlan*. Berkeley: University of California Press.
- , Luin Goldring, and Jorge Durand 1994 “Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities.” *American Journal of Sociology* 99(6):1492–1533.
- Middlebrook, Kevin 1995 *The Paradox of Revolution: Labor, the State, and Authoritarianism in Mexico*. Baltimore, Md.: Johns Hopkins University Press.
- Molina Enriquez, Andres (1909) 1978 *Los grandes problemas nacionales*. México, D.F.: Ediciones Era.
- Molinar, Juan 1989 “The Future of the Electoral System.” In Wayne A. Cornelius, Judith Gentleman, and Peter H. Smith, eds., *Mexico's Alternative Political Futures*. La Jolla: Center for U.S.-Mexican Studies, University of California, San Diego.
- 1992 *Elecciones, autoritarismo y democracia en México*. Mexico City: Cal y Arena.
- Montaño, Jorge 1976 *Los pobres de la ciudad en las asentamientos espontaneos*. México, D.F.: Siglo Veintiuno Editores.
- Moreno, José Luis, ed. 1992 *Ecología, recursos naturales y medio ambiente en sonora*. Hermosillo: El Colegio de Sonora.
- Muñoz, Humberto, Orlandina de Oliveira, and Claudio Stern, eds. 1977 *Migración y desigualdad social en la ciudad de México*. México, D.F.: Instituto de Investigaciones Sociales/El Colegio de México.
- Otero, Gerardo, ed. 1996 *Neo-Liberalism Revisited: Economic Restructuring and Mexico's Political Future*. Boulder, Colo.: Westview.
- Pacheco, Edith 1997 “La población económicamente activa 1900–1995.” *Demos* 10:30–31.
- Padilla Dieste, Cristina 1997 *Todo queda en familia: El mercado de abastos de Guadalajara*. Guadalajara: Universidad de Guadalajara.
- Pastor, Robert A., and Jorge G. Castañeda 1988 *Limits to Friendship: The United States and Mexico*. New York: Alfred A. Knopf.
- Ramírez, Luis Alfonso 1994 *Secretos de familia: Libaneses y élites empresariales en Yucatán*. Mexico City: Consejo Nacional para la Cultura y las Artes.
- Roberts, Bryan 1989 “Employment Structure, Life Cycle, and Life Chances: Formal and Informal Sectors in Guadalajara.” In A. Portes, M. Castelis, and L. Benton, eds., *The Informal Economy—Comparative Studies in Advanced and Third World Countries*. Baltimore, Md.: Johns Hopkins University Press.
- , Reane Frank, and Fernando Lozano 1999 “Transnational Migrant Communities and Mexican Migration to the US.” *Ethnic and Racial Studies* 22.
- Sanderson, Steven E. 1986 *The Transformation of Mexican Agriculture: International Structure and the Politics*

- of Rural Change*. Princeton, N.J.: Princeton University Press.
- Selby, Henry A., Arthur D. Murphy, and S. A. Lorenzen 1990 *Urban Life in Mexico: Coping Strategies of the Poor Majority*. Austin: University of Texas Press.
- Sierra, Justo 1969 (1900–1902) *The Political Evolution of the Mexican People*. Austin: University of Texas Press.
- Solís, Leopoldo 1970 *La realidad economica Mexicana: Retrosivision y perspectivas*. Mexico City: Siglo Veintiuno Editores.
- Solís, Patricio 1997 “Población urbana y la población rural.” *Demos* 10:6–7.
- Stavenhagen, Rodolfo 1969 *Las clases sociales en las sociedades agrarias*. México, D.F.: Siglo Veintiuno.
- Tamayo, Jesús, and José Luis Fernandez 1983 *Zonas fronterizas*. México, D.F.: CIDE.
- Tamayo, Sergio 1994 “The 20 Mexican Octobers: A Study of Citizenship and Social Movements.” Ph.D. Dissertation, University of Texas at Austin.
- Taylor, Paul S. 1928–1934 *Mexican Labor in the United States*, vols. 1–111. Berkeley: University of California Press.
- Toledo, Alejandro 1994 *Riqueza y Pobreza en la Costa de Chiapas y Oaxaca*. México: Centro de Ecología y Desarrollo.
- Tudela, Fernando, ed. 1989 *La modernización forzada del trópico: El caso de Tabasco*. México: El Colegio de México.
- Unikel, Luis (with Crescencio Ruiz Chiapetto and Gustavo Garza) 1978 *El desarrollo urbano de México. Diagnóstico e implicaciones futuras*. Mexico City: El Colegio de México.
- Vazquez, Luis 1995 *El leviathan arqueológico*. Leiden: University of Leiden.
- Villoro, Luis 1998 *Estado plural, pluralidad de culturas*. Mexico City: Paidós/Universidad Nacional Autónoma de México.
- Ward, Peter M. 1998 *Mexico City*. New York: John Wiley.
- Warman, Arturo 1980 “We Come to Object”: *The Peasants of Morelos and the National State*. Baltimore, Md.: Johns Hopkins University Press.
- , and Arturo Argueta 1993 *Movimientos indígenas contemporáneos en México*. Mexico City: Miguel Angel Porrúa/Universidad Nacional Autónoma de México.
- Weintraub, Sidney 1990 *A Marriage of Convenience: Relations between Mexico and the United States*. New York: Oxford University Press.
- Young, Gay 1994 “Household Responses to Economic Change: Migration and ‘Maquiladora’ Work in Ciudad Juarez, Mexico.” *Social Science Quarterly* 75(3):656–671.
- Zárate, Eduardo 1993 *Los señores de la utopía. Etnicidad política en una comunidad purhépecha*. Zamora: El Colegio de Michoacán.
- Zea, Leopoldo 1974 *Positivism in Mexico*. Austin: University of Texas Press.

GUILLERMO DE LA PEÑA  
BRYAN R. ROBERTS

## MIDDLE EASTERN STUDIES

In the closing decades of the twentieth century, successive crises and cataclysmic events in the Middle East, though obviously not unique to this region, have been widely publicized by the American mass media. The 1973 Arab-Israeli War and the subsequent oil crisis, the civil war in Lebanon, the Iranian revolution of 1978–1979 and the ensuing hostage crisis, the Iraq-Iran War of 1980–1988, the Intifadah in the West Bank and Gaza, the Gulf War, the continued bombing of Iraq, the Oslo Agreement between Palestinians and Israelis, and the World Trade Center bombing have increased Americans’ awareness of the Middle East but not necessarily their understanding of its culture and society. Unfortunately, the ways in which the media describe and interpret these events have created negative stereotypes about this region which are reinforced by movies such as *The Siege*. On the other hand, extensive research by anthropologists, historians, and political scientists has provided many insights into Middle Eastern societies and cultures that help counteract these stereotypes. What has been the contribution of sociologists?

More than twenty years ago, a review of a few sociological journals led me to conclude that “with the possible exception of demographers, the scholarly output of American sociologists on the Middle East is still very modest” (Sabagh 1976, p. 523). For the period 1963–1973, of the twelve articles on the Middle East published in the *American Sociological Review* and the *American Journal of Sociology*, eight were on Israel, three on Egypt, and one on Turkey. For the same period *Sociology and Social Research* had even fewer articles on the Middle East, but they were more evenly distributed (two on Israel, two on Turkey, two on Iraq, and one on Israel and Jordan).

One interpretation of this situation, which is still valid today, is that “as long as the research



findings tend to be area-oriented rather than hypothesis-oriented they are not likely to be published in American sociological journals" (Sabagh 1976, p. 523). Is the state of Middle Eastern studies in sociology today any better than in the early 1970s? The objective of this review is to assess the changes that have occurred since the mid-1970s, more particularly in the 1980s and the 1990s. Four topics will be considered: (1) an overview of the salient socioeconomic and demographic characteristics of Middle Eastern countries and their implications for sociological analysis, (2) trends and distributions in the number of sociologists in the United States specializing in the Middle East, (3) a brief review of the main substantive or methodological concerns of these sociologists, and (4) some of their theoretical contributions studies to American sociology. Contributions to the sociological study of Islam will not be reviewed here since they are extensively discussed in the article "Sociology of Islam."

#### **SOCIOECONOMIC AND DEMOGRAPHIC CHARACTERISTICS OF THE MIDDLE EAST**

The Middle East includes Iran, Israel, Turkey, and the Arab countries of North Africa and southwest Asia. As shown in Table 1, countries in this vast region vary greatly, not only in population but also in levels of income, education, urbanization, and stages of demographic transition. The most telling differences are those between rich and poor countries, unmatched by the experience of any other Third World region. The range in gross national product (GNP) per capita in 1997 was from \$169 for Somalia to \$17,360 for the United Arab Emirates (UAE) and to \$22,110 for Kuwait, which approaches the figure for the United States. This gap was even greater before the decline in oil prices in the mid-1980s. The least-populated countries (Kuwait, UAE, and Qatar) have the highest incomes, exceeding \$10,000, and two of the most populous countries (Egypt and Iran) have lower incomes, in the range \$1,000–\$2,000. This income differential is in part responsible for a massive labor migration from the poor Arab countries, particularly Egypt and Yemen, to the rich Arab countries (Amin and Awny 1985; Owen 1989). One measure of the importance of this migration is provided by the figures on the share of workers' remittances in the

GNP of labor-exporting countries. In 1992, net workers' remittances constituted 19.5 and 16.5 percent of the GNP of Jordan and Egypt, respectively. Estimates of the size of the Arab labor migration streams vary widely (Amin and Awny 1985; Ibrahim 1982), but it increased rapidly in the 1970s and was substantial in 1980. For Egypt alone, one estimate places the number of workers abroad at over one million in 1980, compared with around 400,000 in 1975 (Amin and Awny 1985; Fergany 1987). There was also a considerable migration from the poor countries of southern Asia and eastern Asia to Gulf countries, with an estimated 700,000 contract workers migrating in 1987 (Stahl and Asam 1990). Prior to the Iraq-Kuwait War in 1990, there were an estimated 2.6 millions migrants in these two countries, but at least 1 million of them left within two months after the outbreak of this war (United Nations, 1996, p. 209). In addition, there was a massive exodus of Yemenis from Saudi Arabia, leading to a sharp drop of net worker remittances for Yemen within a fairly short period. There is no doubt that this massive exodus of labor migrants from Iraq, Kuwait, and Saudi Arabia created new economic problems for Arab labor-exporting countries.

The characteristics and implications of the massive labor migration of Middle Easterners and Asians to the rich Arab oil-producing countries have been analyzed by a few American demographers and sociologists (Arnold and Shah 1984; Sabagh 1993; Sell 1988). These and other studies (Amin and Awny 1985; Fergany 1987) have contributed to the analysis of (1) the process of migrant settlement in countries with stringent legislation against settlement, (2) the impact of this migration on social mobility, and (3) the consequences of labor migration on countries of origin.

While a few North Africans and Turks emigrated to the Gulf, most were attracted by expanding economic opportunities in the European community. In 1996, there was an estimated 2.6 million Turks, 1.1 million Moroccans, and about 1 million Algerians and Tunisians in Europe (OECD 1998, p. 34). Most of the Turks were in Germany, and most of the Algerians and Tunisians and half of the Moroccans resided in France.

Migration streams from the Middle East to the United States go back to the 1980s, but political

**Selected Indicators: Middle Eastern Countries and the United States, 1995–1998**

Country	GNP per Capita (\$U.S.) mid-1997	Population (Millions), mid-1997	Total Fertility Rate, 1997	Life Expectancy, at Birth, Males, 1998	Life Expectancy, at Birth, Females, 1998	Annual Rate of Growth Total, 1995-2000	Population Urban, 1995-2000	Enrollment Ratios (%) Secondary School		Female Economic Activity Rate as % of Male, 1995
								Female, 1995	Male, 1995	
Somalia	\$169	7.1	7	45	48		5.3	6	4	NA
Yemen	\$270	8.5	7.1	54	54	3.7	6.3	63	14	38
Sudan	\$280	28	5.7	55	57	2.2	5	21	19	40
Mauritania	\$450	2	5.4	52	55	2.5	4.9	22	11	48
Syrian AR	\$1,150	11.4	4	66	71	2.5	3.4	45	40	35
Egypt	\$1,180	60	3.3	64	67	1.8	2.4	80	70	40
Morocco	\$1,250	27	3.3	64	68	1.8	3.1	44	34	53
Algeria	\$1,490	29	3.4	68	72	2.3	3.6	65	62	32
Jordan	\$1,520	4	4.8	71	75	3.3	4.1	NA	NA	27
Iran IR	\$1,780	51	4.3	67	70	2.2	3.1	79	65	32
Tunisia	\$2,090	9	2.8	69	71	1.8	2.9	67	64	44
Turkey	\$3,130	64	2.5	66	71	1.6	3.2	74	50	56
Lebanon	\$3,350	4	2.7	68	71	1.8	2.3	78	86	39
Oman	\$4,950	2	5.1	69	73	4.2	6.3	68	65	16
Libyan AJ	\$5,621	5	6.2	63	68	3.3	3.9	NA	NA	26
Saudi Arabia	\$6,790	20	6.2	69	71	3.4	4.1	65	57	15
Bahrain	\$7,820	0.5	3	72	78	2.1	2.5	91	98	24
Qatar	\$11,570	0.4	3.5	71	77	1.8	2	79	77	15
Israel	\$15,810	6	2.7	75	79	1.9	2	89	87	60
UAE	\$17,360	3	4.7	74	76	2	2.5	77	74	15
Kuwait	\$22,110	2	3.4	75	79	3	3.1	65	64	45
Iraq	NA	22	6.1	67	69	2.8	3.4	50	32	22
West Bank/ Gaza	NA	3	4.9	71	74	4.2	4.2	NA	NA	NA
United States	\$28,740	268	2.1	74	80	0.8	1.1	98	99	82

Table 1

NOTE: NA=not available.

SOURCE: United Nations Development Programme 1998 *Human Development 1998*. New York: Oxford University Press; World Bank *World Development Report 1998–99*. Washington: World Bank; United Nations Home Page 1999 Statistics Division Home Page. Social Indicators Home Page. New York: United Nations.

upheavals such as the Iranian revolution, the Arab-Israeli wars of 1967 and 1973, and the Lebanese civil war as well as changes in American immigration legislation in the 1960s have led to substantial and more recent immigration to the United States. Thus, of the 1.3 million persons of Middle Eastern origin (including those of Arab, Armenian, Assyrian, Iranians, Israelis, and Turkish origins) in the United States in 1990, half were immigrants and 325,000 were very “new” immigrants who had arrived between 1980 and 1990. (U.S. Bureau of the Census 1993). As indicated by Bozorgmehr (1998, p. 5) “The Iranian revolution and its aftermath have contributed to the growth of the Iranian diaspora worldwide,” so that by 1990 the United States had

285,000 Iranians “defined as persons born in Iran or Iranian ancestry.” This represents a substantial growth from “small beginning of perhaps no more than 15,000 individuals in 1965” (Bozorgmehr and Sabagh 1988, p. 32).

In almost all Middle Eastern countries, rates of population growth of urban areas are higher, and in some cases much higher, than total population growth. This means that rural-urban migration constitutes an important feature of both rich and poor Middle Eastern societies. This rural exodus is most marked in some of the poorest countries but tends to be lower in poor countries with substantial international labor migration.

Differences in secondary school enrollment levels are as wide as income differences. Enrollment ratios in secondary schools in 1996 varied between a minimum of 6 for males and 4 for females in Somalia to a maximum 91 for males and 98 for females in Bahrain, almost equal to the United States (Table 1). While Bahraini's high enrollment ratios could be partly attributed to its relatively high income, wealth is no guarantee of higher education. For Kuwait, which has the highest income level, male secondary enrollment is noticeably lower (65) than it is for Egypt (80) and the Islamic Republic of Iran (79), which have much lower income. The ratios for females are about the same for these three countries, but the female-male gap is greater in Egypt and Iran. It should be noted that in 15 of the 25 Middle Eastern countries in Table 1, the female-male gap in secondary enrollment ratios is 10 percent or less. In Bahrain and Lebanon, females have a higher secondary school enrollment. These findings suggest gender equalization in education. Nevertheless, in some countries, notably Somalia, Yemen, and, surprisingly, Turkey, the gender gap is substantial. Since educational opportunities are fewer in rural areas, particularly for women, this gap could partly reflect the greater importance of these areas. Even in Middle Eastern countries where the educational gender gap is almost nil, women are still much less likely to have gainful employment outside the home than men. Nationwide data on women's employment, however, have to be interpreted carefully, since they include rural areas and also reflect different definitions of employment. Women in Israel, Turkey, and Morocco are the most likely to be employed, but even in these countries the ratio of women's to men's employment is in the range 50–60 percent (Table 1). The ratios between 40 percent and 50 percent in Kuwait, Tunisia, and Mauritania and between 30 percent and 40 percent in seven other Middle Eastern countries. By contrast, with the noticeable exception of Kuwait, in the rich Arab oil-exporting countries with little or no gender inequality in access to education, the ratio is much lower and is mainly around 15 percent. Increasing gender equality in education combined with continued gender inequality in employment in the Middle East have many consequences for women's roles and have been the subject of many analyses (Hatem 1994; Moghadam 1995; Obermeyer 1996). These analyses reveal the complex ways in which family structure, religious

tradition, economic development, and state policies affect the education and employment of women.

It should be noted that even in countries with high labor-force participation by women, there is a glaring gender gap in earned income. Thus, in the United States, while there are 82 gainfully employed women for every 100 gainfully employed men, these women receive only 40 percent of men's earnings. The comparable figures are 50 women and 22 percent for Israel, 57 women and 35 percent for Turkey, and 53 women and 28 percent for Morocco. Surprisingly, this lag in working women's income is less marked in some Middle Eastern countries than in the United States (for sources, see Table 1).

During the last twenty years, fertility and mortality trend in the Middle East indicate that parts of in this region have experienced a rapid demographic transition. In 1980, all countries in this region, except Israel, Lebanon, and Turkey, had total fertility rates (TFR) of 5 or more children per women, often reaching a level above 6 children. Mortality had already declined but was still relatively high. By the end of the twentieth century, life expectancy for males had increased to about 63 years at birth and, even higher, to 74–75 years in Kuwait, Israel, and the UAE, equal to that of the United States (Table 1). There was a similar trend among females, most of whom outlive men by as much as 6 years. In nine Middle Eastern countries there was a marked decline in TFR to fewer than 3.5 children per woman, reaching as low as 2.5–2.7 in Israel, Lebanon, and Turkey in 1997 (Table 1). This is only slightly higher than the TFR for the United States. Clearly some Middle Eastern countries are rapidly undergoing a demographic transition to noticeably lower fertility and mortality, but in some other countries, even though mortality is low, fertility remains high. This is a challenge to the demographic transition theory and has been extensively analyzed by demographers and sociologists (Obermeyer 1992 and 1995).

Clearly, the Middle East is a region worthy of consideration in any analysis of the social and demographic impact of economic modernization and rising levels of income. The sudden increase in the wealth of some Arab countries is part of what Ibrahim (1982) has called the "new Arab social order," which involves the appearance of new social forces, new values, and new behavior

patterns. This has resulted in a great deal of social chaos and the emergence of new social problems. Partly as a result of this social chaos, there have been significant political and social movements and revolutions that need to be analyzed from a comparative sociological perspective.

By making systematic and comparative analyses of various features of Middle Eastern societies, American sociologists could provide a foundation for a better understanding of these societies. This assumes, however, that these sociologists have incorporated the history and experience of Middle Eastern countries into their substantive concerns and theoretical models. The sections that follow will review the trends in the number of sociologists of the Middle East and in the United States, and the substantive focus of their work.

#### **NUMBER OF SOCIOLOGISTS OF THE MIDDLE EAST IN THE UNITED STATES**

It is significant that an American sociologist, Monroe Berger of Princeton University, was active in the formation of the Middle East Studies Association of North America (MESA) in 1964 and was elected its first president. Nevertheless, in the mid-1970s there were only nine sociologists teaching courses on the sociology of the Middle East in American departments of sociology (Sabagh 1976, p. 524). While growth was slow, the 1980s and 1990s were marked by a noticeable increase in sociologists teaching or doing research on the Middle East. Sociologist members of MESA, including students, increased from 12 in 1968 to 25 in 1972, 37 in 1984, 52 in 1986, 90 in 1990, and 121 in 1998 (Sabagh 1976; Bonine 1986; MESA 1998; and special tabulation by MESA). Not all these sociologists, however, resided in the United States. In 1986 only 3 percent of MESA members were sociologists, compared to 32 percent who were historians and 21 percent who were political scientists (Bonine 1986, p. 159). In the mid-1980s Bonine could still state that "compared to Latin America, for instance, there are few sociologists and demographers specializing on the Middle East" (1986, p. 160).

While an increasing number of Middle Easterners have come to the United States to obtain Ph.D.'s in sociology, most of them have gone back home; but, as a result of the Iranian revolution and its negative effect on the teaching of Western-type

sociology in Iran, many Iranians who obtained Ph.D.'s in sociology have remained in the United States as intellectual exiles. The Iranian revolution also provided a real impetus for Iranian students to major in sociology so as to gain a better understanding of the revolution.

Dr. Anne Betteridge, executive secretary of MESA, and Nancy B. Dishaw, also of MESA, kindly provided a special tabulation of the countries of birth of 1998 MESA members who indicated sociology as their discipline. There were 62 members who were residents of the United States or Canada with Ph.D.'s in sociology, of whom 44 had received their degrees since 1980, indicating a noticeable increase in the number of sociologists actively interested in Middle East studies. Most of this increase, however, may be attributed to sociologists born in the Middle East, whose numbers more than doubled, from 10 to 25. A comparison of the later and earlier Ph.D. cohorts shows that the number of sociologists born in Iran quadrupled from 3 to 12, the number of those born in Turkey increased from 0 to 7, and the number born in the United States increased more slowly from 8 to 15. By contrast, the number of sociologists born in an Arab country decreased from 7 to 6. The creation of the Association of American University Graduates in 1967 may be partly responsible for this decline.

About half of the 62 sociologists who are members of MESA and who reside in the United States or Canada are in graduate departments of sociology. By contrast, nearly all the 24 sociology student MESA members are in graduate departments of sociology. Not all sociologists with a research or teaching interest in the Middle East are members of MESA. A search of the *Guide to Graduate Departments of Sociology* of the American Sociological Association (ASA) (1999) shows that there were at least 12 who were not members of MESA.

#### **SUBSTANTIVE FOCUS OF SOCIOLOGISTS OF THE MIDDLE EAST**

It is safe to assume that sociologists who belong to MESA will focus some of their research and theoretical interest on the Middle East. The directory of ASA members (ASA 1990) provides information on fields of specialization. Unfortunately, there

are no tabulations available that would allow comparisons with all sociologists or sociologists who have an interest in other world areas. There were 26 nonstudent and 9 student members of ASA in 1989 who were also members of MESA. Of the 26 nonstudent members, 15 were in graduate departments of sociology. Members were asked to check four substantive or methodological areas of sociology in order of priority, of which the first and second are below.

The three most important areas of specialization are (1) development and social change, (2) comparative sociology/macro, and (3) political sociology. Next in importance are (1) demography, (2) race/ethnic/minority, (3) sociology of sex roles, and (4) social movements. Unfortunately, no questions were asked about the extent to which sociologists focused on a given region, country, society, or community. For this, it would be necessary to analyze the publications and papers of the 35 scholars who were members of both the ASA and MESA. A survey was made of publications and papers by these scholars that were cited in *Sociological Abstracts* for the period 1985–1990. The three most frequent substantive or theoretical topics were (1) the Iranian revolution, (2) historical-sociological analyses of Turkey and Iran, with an emphasis on dependency or world system approaches, and (3) Iranian immigrants in the United States. Other topics included (1) Egyptian international migration, (2) social distance in Egypt, (3) child nutrition, (4) a review of Arab sociology, and (5) the Intifadah. It is likely, however, that there are many more contributions to Middle Eastern studies by sociologists who are not members of MESA. The task of assessing this literature is clearly beyond the scope of this article. One way to obtain this information is through a survey of these sociologists. In 1983–1984 such a survey was carried out for Arab sociologists or sociologists studying the Arab world (Sabagh and Ghazalla 1986). Ten of the sociologists residing in the United States considered social change and development and the role of Islam to have the highest priority for research on the Arab world.

One issue that needs to be considered is the extent to which sociological research on the Middle East is cited in major American sociological journals. An analysis was made of articles on the Middle East cited in *Sociological Abstracts* for the years 1985–1990 and published in the general and

more specialized sociological journals. Since the number of papers on Israel was substantial, they were left out of the analysis. The results of this analysis showed that the situation was even worse now than it was in the period 1963–1973 (Sabagh 1976). On the other hand, in the 1980s a substantial number of articles on the Middle East were published in the newer or more specialized sociological journals.

An analysis of *Sociological Abstracts* for the period 1985–1990 showed that 31 specialized journals had published 104 articles pertaining to the Middle East or to Middle Easterners. The distribution of articles by country or region (exclusive of Israel) was as follows: Iran, 31; Turkey/Ottoman Empire, 14; Arab world/Middle East, 15; Egypt, 14; other Middle Eastern countries, 19; Palestinians, 5; Arab and Iranian immigrants in the United States, 6. The large number of articles on Iran may be explained by the analyses of the Iranian revolution and by the increasing number of Iranian sociologists in the United States. One issue of *State, Culture, and Society* (1985) included 8 articles devoted to “The Sociological and Ideological Dimensions of the Iranian Revolution.” There were four articles on the Iranian revolution (or comparing the Iranian revolution with other revolutions) in *Theory and Society*. Five specialized journals published 10 articles on marriage and the family and the status of women, 5 of which were on Iran. The period 1985–1990 was also characterized by a substantial number of articles on the Ottoman Empire and Turkey published in the Fernand Braudel Center’s *Review*. The winter 1985 issue of the *Review* included 3 articles on “From Ottoman Empire to Modern State,” and the spring 1988 issue was devoted to the “Ottoman Empire: Nineteenth-Century Transformations.” These articles present critical analyses of the process of incorporation of the Ottoman Empire/Turkey into the world economy. Clearly, sociologists whose research focuses on the Middle East or Middle Easterners have a better chance of getting their articles accepted in specialized sociological journals than in the major and mainstream sociological journals.

In 1996, the editor and board of *Contemporary Sociology* decided to “make a greater effort to cover sociology around the world” (Clawson 1997, p. vii). As a consequence, this journal published assessments of the development of sociology in Egypt and the Arab World (Ibrahim 1997; Naim-Ahmed

1998), Turkey (Oncu 1997), Israel (Ben-Yehuda 1997), and Palestine (Hammami 1997). Earlier, in a section of *Contemporary Sociology* entitled "Journals in Review," Mirsepassi (1995) presented an analysis of Middle Eastern journals. This is perhaps the most promising development of the 1990s, and it is hoped that it will lead to a greater impact of Middle Eastern studies on American sociology in the twenty-first century.

Although the numbers are still small, there has been a surprising increase in articles devoted to the Middle East (excluding Israel) published in the 1990s in the three major sociological journals. From 1990 to 1997, there were 7 articles, all of them devoted to Iran or the Ottoman Empire, in the *American Sociological Review*, the *American Journal of Sociology*, and *Social Forces*. This compares to only 1 article in the same journals in the 1980s.

#### **THEORETICAL AND METHODOLOGICAL CONTRIBUTIONS OF MIDDLE EASTERN STUDIES TO SOCIOLOGY**

There are two major ways in which Middle Eastern studies can contribute to sociology: (1) dramatic events in the Middle East, such as wars and revolutions, present a puzzle to sociologists studying such events, and (2) sociologists specializing in Middle Eastern studies focus on substantive issues that have consequences for sociological theory and methodology. Examples of both types of contributions will be given and discussed.

The Iranian Islamic revolution is clearly one of the most dramatic events of the last decades of the twentieth century. For Arjomand, it is a "cataclysm as significant and as unprecedented in world history as the French revolution of 1789 and the Russian revolution of 1917" (1988, p. 3). As could be expected, American sociologists of Iranian origin have contributed many important insights into the causes and consequences of this revolution, and these insights have been and will be incorporated into sociological theories about revolutions (see, e.g., Arjomand 1988; Ashraf and Banuazizi 1985; Ashraf 1988; Dabashi 1984; Moghadam 1989; Parsa 1989). For other sociologists who have studied the nature and sources of revolution, the Iranian revolution presents a real challenge. This is expressed as follows by Theda Skocpol:

*The recent overthrow of the Shah, the launching of the Iranian revolution between 1977 and 1979, came as a sudden surprise to outside observers. . . . All of us have watched the unfolding of current events with fascination and, perhaps, consternation. A few of us have also been inspired to probe the Iranian sociopolitical realities behind those events . . . . Its unfolding . . . challenged expectations about revolutionary causation that I developed through comparative-historical research on the French, Russian and Chinese revolutions. (1982, p. 265)*

Skocpol applied her earlier structural analysis of the causes of revolutions to the Iranian case, pointing to the involvement of urban masses in this revolution, the role of the rentier state, and the structural consequences of Shia Islam. When challenged by Nichols to the effect that she is "ready to concede the potentially revolutionary content of traditional religious teachings" (1986, p. 182). Skocpol replied that "far from offering a 'subjectivist' or 'ideational' analysis, I point out that Shi'a religion had the cultural potential to facilitate either a rebellion against or passive acquiescence to secular authority" (1986, p. 193). Arjomand criticizes Skocpol for not recognizing the normative factor of legitimacy" (1988, p. 191).

Arjomand, Ashraf, and other sociologists indicate the importance of analyzing the 1977-1979 revolution by considering its historical roots, particularly the structural, ideological, and leadership factors in previous uprisings and rebellions. Thus, for Ashraf, "The urban uprising of 1963 combined the leadership of a militant charismatic leader, the political resources of the bazaar-mosque alliance, and the sympathy of activist university students . . . a prelude to the Islamic revolution of 1977-79" (1988, p. 550). Arjomand (1988) points to the involvement of "high-ranking members" of the Shiite "clergy" or hierocracy in the constitutional revolution of 1905-1906. Arjomand (1984a, 1984b, 1988) not only traces the history of the relationship between the Shiite hierocracy, the state, and other social groups but also places his analysis of the Iranian revolution in a comparative perspective. With passage of time, one might expect a decline in the number of sociological analyses of the Iranian revolution. In fact, the opposite is true, and this revolution continues to fascinate some

American sociologists challenged by its implications for theories of revolution, social protests, and social movements. There is also increasing interest in the consequences of this revolution. Kurzman (1994) uses the material about the Iranian revolution to raise some important theoretical questions about resource mobilization theory. In a later paper (1996) he also uses this revolution to challenge the theory about structural and perceived opportunities. His argument is based on a number of interviews he carried out with Iranian expatriates in Turkey. Moaddel (1992) uses the Iranian revolution to test a the role of ideology, suggesting that ideology is a language that is used to express ideas about social problems and their solution. What is probably the first quantitative analysis of events preceding the Islamic revolution was carried out by Karen Rasler, in which she tests a number of hypotheses about the effects of governmental repression and concessions on the number of political protests from December 1, 1977, to February 14, 1979 (1996, p. 138).

Studies of other political protest movements and revolutions in Iran have been brought together in a book edited by John Foran (1994). This book includes analyses of the Tobacco Movement of 1890–1892, the constitutional revolution of 1905–1911, the nationalist movement in Azerbaijan and Kurdistan of 1941–1946, the oil nationalization movement of 1939–1953, and social conflict in the 1960s. American historians of Iran such as Nikki Keddie and Ervand Abrahamian are, of course, very familiar with these events, but the book should inform and hopefully challenge sociologists interested in social movements and revolutions. In one chapter Foran provides an analysis of the 1977–1979 revolution, showing the importance of comparing it to those that occurred in other countries. He concludes that “its deep causes are to be sought in the same processes of dependent development, state repression, political cultures of resistance, economic downturn, and world-systemic opening that underlay other successful revolution in Mexico, Cuba, and Nicaragua” (Foran 1994, p. 181). Foran and Goodwin (1993) and Parsa (1995a) have published studies comparing the Iranian and Nicaraguan revolutions. Parsa (1995a) shows that ideological conversions and tactical considerations explain the support that entrepreneurs and workers gave to the Iranian and Nicaraguan revolution. In another paper, Parsa (1995b) presents a critical

analysis of the role of the bourgeoisie in supporting the process of democratization in Iran under the shah and in the Philippines under Marcos. Moaddel (1996) analyzes the rise of Islamic fundamentalism in Iran and Syria. Sohrabi (1995) has carried out an extensive comparative analysis of the constitutional revolutions in Iran, the Ottoman Empire, and Russia in the early 1990s. Systematic comparisons between the Iranian and other revolutions and their outcomes are the most promising development of the 1990s.

One interesting development of the last twenty years has been the application of the world system theoretical model to the Ottoman Empire and Turkey. An early issue of the Fernand Braudel Center's *Review* includes three articles on the Ottoman Empire and the world economy, one of which provides a program of research delineated by Immanuel Wallerstein (1979). The objective of this program was to analyze the process of incorporation of the Ottoman Empire into the world economy, the consequences of this incorporation, and the degree and nature of the peripheralization of the Ottoman Empire and Turkey. This program was carried out by the Research Working Group on the Ottoman Empire and was coordinated by Caglar Keyder and Immanuel Wallerstein. The work of Kasaba (1987, 1988); Keyder and Tabak (1986); Keyder (1988); Pamuk (1988); and other sociologists, economists, and historians with a world-system perspective has documented the timing and the process of incorporation of the Ottoman Empire into the world economy, the role of the Ottoman bureaucracy, and the emergence of a minority non-Muslim bourgeoisie. Kasaba (1987, p. 842) argues that “between ca. 1750 and ca. 1820 the Ottoman Empire as a whole was incorporated into the capitalist world-economy . . . [and] the history of the Ottoman Empire after the 1820's was that of its peripheralization” (Pamuk's p. ). However, analysis of various indexes of foreign trade and investment (1988) shows that until 1914 the degree of integration of the Ottoman Empire into the world economy was noticeably lower than that of Latin America or that of Algeria and Egypt. Also, within the Ottoman Empire different sectors of the economy were incorporated in different periods (Cizacka 1985).

One consequence of the process of incorporation into the world economy was the emergence of

a new bourgeois class of non-Muslim merchants (Kasaba 1987; Keyder 1988). While there was a social class conflict between these merchants and the Ottoman bureaucracy, this conflict "appeared on the Ottoman agenda to be acted out as Moslem-Christian ethnic-religious conflict" (Keyder 1988, p. 162). The periphery-core status of Safavid Iran and the Ottoman Empire in the seventeenth century is analyzed by Foran, who concludes that these states "were far too strong to be colonized and dominated by the core, and yet too weak to compete with Europe in the new peripheries of Asia, Southeast Asia, and Africa, not to mention Latin America" (1989, p. 113; italic in original). Thus, these states were neither core nor periphery.

The Fernand Braudel Center *Review* continues to publish studies of the Ottoman Empire and Turkey with a world-system perspective. Of particular interest to sociologists interested in the application of this model to urbanization is the special issue of the *Review* (fall 1993) devoted to the role of port cities in the process of incorporation of the Ottoman Empire in the nineteenth century. Included are studies of Izmir by Kasaba, Trabzon by Turgay, and Beirut by Ozveren. Keyder's study of Istanbul in the 1940s (1993) was also published in *Review*, later in the same year.

One important theoretical and methodological addition to the socio-historical studies of the Ottoman Empire is the research by Karen Barkey on conflict and contention in Turkish villages in the seventeenth century (1997). Barkey statistically analyzes court records about contentions pertaining to taxes and land. She argues that "local and small, are the most basic and common form of contention, closely linked to discontent, often preceding and underlying collective form of discontent (1997, p. 1346).

Studies of Arab immigrants in the United States provide important insights into the process of ethnicity maintenance among immigrants, showing, for example, the importance of minority status in the country of origin, religion, and entrepreneurship in the United States (Abu-Laban 1989; Abraham et al. 1983; Hagopian and Paden 1969; Sawaie 1985; Sengstock 1982; Suleiman and Abu-Laban 1989; Swan and Saba 1974). Pulcini (1993) provides a detailed review of this extensive literature.

The recent immigration from Arab countries, from Israel, and particularly from Iran, documented above, has stimulated sociological studies of the adaptation patterns of these immigrants. Bozorgmehr and Feldman (1996) edited *Middle Eastern Diaspora Communities in America*, which includes chapters on all Middle Easterners, Iranians, Israelis, and Arab Muslims. It is significant that a volume edited by Waldinger, Der-Martirosian, and Bozorgmehr, which is devoted to analyses of the social and economic trajectories of major immigrant groups in the Los Angeles metropolitan region, includes a chapter entitled "Middle Easterners: A New Kind of Immigrants" (Bozorgmehr et al. 1996). It is also notable that *Iranian Studies*, which publishes only articles on Iran, devoted an entire recent issue, edited by Bozorgmehr, to the "Iranian diaspora" (1998). The articles by Bozorgmehr, Ali Modarres, Hamid Naficy, Shideh Hanassab, and Ali Akbar Mahdi cover a wide range of topics, including demographic and socioeconomic trends, patterns of settlement, music videos, sexuality and dating, and ethnic identity.

The high socioeconomic status of Iranian immigrants in the United States and their ethnoreligious diversity provide a theoretical challenge for the study of the process of immigrant adaptation. Iranian immigrants include Christian Armenians and Assyrians, Baha'is, Jews, Muslims, and Zoroastrians. A challenge for the sociological analysis of immigrant adaptation is the fact that some of these groups have their non-Iranian religious counterparts in the United States. This diversity was the focus of the first extensive study of the adaptation of Iranian immigrants in Los Angeles, which has the largest concentration of Iranians in the United States (Bozorgmehr and Sabagh 1989). Bozorgmehr (1997) has developed the concept of "internal ethnicity" for the analysis of the process of adaptation of ethnically diverse immigrant groups. This concept is applicable not only to Iranians but also to many other early and recent immigrant groups. He shows that the neglect of internal ethnicity in the literature has led to the simplification of a complex process of immigrant adaptation.

Bozorgmehr's recent comprehensive review of the studies of first- and second-generation Iranians (1998) shows the wide range of topics covered by this literature, including comparisons between immigrants and exiles, ethnicity and ethnic identity, professionals and entrepreneurs, gender, and



assimilation. As the sons and daughters of Iranian immigrants are reaching adulthood, it is time for "research to direct its attention" to this second generation (Bozorgmehr 1998, p. 26). It is also time to compare the experiences of the Iranian second generation to the sons and daughters of other immigrants with equally high human and social capital, whether they are from the Middle East or from other regions of the world.

### CONCLUSION

One unexpected consequence of the Iranian revolution has been an increase in the number of Iranian sociologists in the United States who have developed extensive and theoretically relevant analyses of this revolution—which, analyses in turn, have stimulated other sociologists to study this revolution. Studies of the Iranian revolution are now appearing in the mainstream sociological journals.

The world-system theoretical model has greatly benefited from extensive studies of the Ottoman Empire, especially by sociologists of Turkish origin. Ottoman archival documents are also being used to test other theoretical perspectives.

The rapid growth in the numbers of Middle Eastern immigrants in the United States in the last twenty-five years has led to the development of new concepts and insights that help us better understand the adaptation of high-status immigrants.

Still, there is much progress to be made. A continuing need exists for incorporating the experiences of this region into sociological models of wars and revolutions, of social and demographic change, of social movements, of gender roles, and the impact of religion on society. As suggested by Miresepassi (1995, p. 324), accomplishment of this work should also lead to the sociology of Middle Eastern studies becoming more relevant.

### REFERENCES

- Abraham, Sameer Y., Nabeel Abraham, and Barbara C. Aswad 1983 "The Southend: An Arab Muslim Working-Class Community." In Sameer Y. Abraham and Nabeel Abraham, eds., *Arabs in the New World: Studies on Arab-American Communities*. Detroit: Wayne State University Press.

- Abu-Laban, Sharon M. 1989 "The Coexistence of Cohorts: Identity and Adaptation among ArabAmerican Muslims." *Arab Studies Quarterly* 11:45–83.
- American Sociological Association 1990 *Biographical Directory of Members*. Washington: ASA.
- 1999 *Guide to Graduate Departments of Sociology*. Washington: ASA.
- Amin, Galal A., and Elizabeth Awny 1985 *The International Migration of Egyptian Labour. A Review of the Suite of the Arts*. Ottawa: International Development Research Centre.
- Arjomand, Said A., ed. 1984a *From Nationalism to Revolutionary Islam*. Albany: State University of New York Press.
- 1984b *The Shadow of God and the Hidden Imam*. Chicago: University of Chicago Press.
- 1988 *The Turban and the Crown: The Islamic Revolution in Iran*. New York: Oxford University Press.
- 1993 *The Political Dimension of Religion*. Albany: State University of New York Press.
- Arnold, Fred, and Nasra M. Shah 1984 *Asian Labor Migration: Pipeline to the Middle East*. Honolulu: East-West Center.
- Ashraf, Ahmad 1988 "Bazaar-Mosque Alliance: The Social Basis of Revolts and Revolutions." *Politics, Culture, and Society* 1:538–567.
- , and Ali Banuazizi 1985 "The State, Classes and Modes of Mobilization in the Iranian Revolution." *State and Society* 1:3–40.
- Barakat, Halim 1993 *The Arab World: Society, Culture, and State*. Berkeley: University of California Press.
- Barkey, Karen, and Ronan Van Rossem 1997 "Networks of Contention: Villages and Regional Structure in the Seventeenth-Century Ottoman Empire." *American Journal of Sociology* 102:1345–1382.
- Bayat, Asef 1999 "Revolution without Movement, Movement without Revolution: Comparing Islamic Activism in Iran and Egypt." *Comparative Studies in Society and History* 40:136–169.
- Ben-Yehuda, Nachman 1997 "The Dominance of the External: Israeli Sociology." *Contemporary Sociology* 26:271–274.
- Bonine, Michael E. 1986 "MESA and Middle East Studies." *Middle East Studies Association Bulletin* 20:155–170.
- Bozorgmehr, Mehdi 1997 "Internal Ethnicity: Iranians in Los Angeles." *Sociological Perspectives* 40:387–408.
- 1998 "Iranians in America" *Iranian Studies* 31:1–96.
- , and Alison Feldman, eds. 1996a *Middle Eastern Diaspora Communities in America*. New York: Hagop

- Kevoorkian Center for Near Eastern Studies of New York University.
- , Claudia Der-Martirosian, and Georges Sabagh. 1996. "Middle Easterners: A New Kind of Immigrant." In Roger Waldinger and Mehdi. Bozorgmehr, eds., *Ethnic Los Angeles*. New York: Russell Sage.
- , and Georges Sabagh 1988 "High Status Immigrants: A Statistical Profile of Iranians in the United States." *Iranian Studies* 21:5–36.
- Cizakca, Murat 1985 "Incorporation of the Middle East European World-Economy." *Review* 8:353–377.
- Clawson, Dan 1997 "Editorial." *Contemporary Sociology* 26(3):vii.
- Dabashi, H. 1984 "The Revolutions of Our Time: Religious Politics in Modernity." *Contemporary Sociology* 13:673–676.
- Dickens, David 1989 "The Relevance of Domestic Traditions in the Development Process: Iran 1963–1979." *International Journal of Contemporary Sociology* 26:55–70.
- Fergany, Nader 1987 *Differential Labour Migration: Egypt (1974–1984)*. Cairo: Cairo Demographic Centre.
- Foran, John 1989 "The Making of an External Arena: Iran's Place in the World-System, 1500–1722." *Review* 12:71–120.
- , ed. 1994 *A Century of Revolution: Social Movements in Iran*. Minneapolis: University of Minnesota Press.
- 1997 "The Future of Revolutions at the Fin-de-siecle." *Third World Quarterly* 18:791–820.
- , and Jeff Goodwin 1993 "Revolutionary Outcomes in Iran and Nicaragua: Coalition Fragmentation, War, and the Limits of Social Transformation." *Theory and Society* 22:209–247.
- Hagopian, Elaine C., and Ann Paden, eds. 1969 *The Arab-Americans: Studies in Assimilation*. Wilmette, Ill.: Median University Press International.
- Hammami, Rema, and Salim Tamari 1997 "Popular Paradigm: Palestinian Sociology." *Contemporary Sociology* 27:275–276.
- Hasso, Frances S. 1998 "The 'Women's Front': Nationalism, Feminism, and Modernity in Palestine." *Gender and Society* 12:441–465.
- Hatem, Mervat F. 1994 "Women in Arab Society: Work Patterns and Gender Relations in Egypt, Jordan and the Sudan." *Signs* 19:535–539.
- Ibrahim, Saad E. 1982 *The New Arab Social Order: A Study of the Social Impact of Oil Wealth*. Boulder, Colo.: Westview.
- 1997 "Cross-Eyed Sociology in Egypt and the Arab World." *Contemporary Sociology* 26:547–551.
- Kasaba, Resat 1987 "Incorporation of the Ottoman Empire: 1750–1820." *Review* 10:805–847.
- 1988 "Was There a Compradore Bourgeoisie in Mid-Nineteenth Century Western Anatolia?" *Review* 11:215–228.
- Keyder, Caglar 1988 "Bureaucracy and Bourgeoisie: Reform and Revolution in the Age of Imperialism." *Fernand Braudel Center Review* 11:151–165.
- 1994 "Globalization of a Third-World Metropolis: Istanbul in the 1980's." *Review* 17:383–421.
- , and Faruk Tabak 1986 "Eastern Mediterranean Port Cities and Their Bourgeoisie: Merchants, Political Projects, and Nation-States." *Fernand Braudel Center Review* 10:131–135.
- Kurzman, Charles 1994 "A Dynamic View of Resources: Evidence from the Iranian Revolution." *Research in Social Movements, Conflicts and Change* 53–84.
- 1996 "Structural Opportunity and Perceived Opportunity in Social-Movement Theory: The Iranian Revolution of 1979." *American Sociological Review* 61:153–170.
- Middle East Studies Association 1998 *Roster of Members 1998*. Tucson, Ariz.: MESA.
- Mirsepassi, Ali 1995 "Middle Eastern Studies and American Sociology." *Contemporary Sociology* 24:324–328.
- Moaddel, Mansoor 1992 "Ideology as Episodic Discourse: The Case of the Iranian Revolution." *American Sociological Review* 57:353–379.
- 1996 "The Social Bases and Discursive Context of the Rise of Islamic Fundamentalism: The Cases of Iran and Syria." *Sociological Inquiry* 330–355.
- Moghadarn, Val 1989 "Populist Revolt and the Islamic State in Iran." In Terry Boswell, ed., *Revolution in the World System*. New York: Greenwood.
- 1995 "Women's Employment Issues in Contemporary Iran: Problems and Prospects in the 1990s." *Iranian Studies* 28:175–202.
- Naim-Ahmed, Samir 1998 "Sociology in the Arab World: A Self-Criticism." *Contemporary Sociology* 27:327–328.
- Nichols, Elizabeth 1986 "Skocpol on Revolution: Comparative Analysis vs Historical Conjecture." *Comparative Social Research* 9:163–186.
- Obermeyer, Carla Makhlof 1992 "Islam, Women, and Politics: The Demography of Arab Countries." *Population-and-Development-Review* 18:33–60.
- , ed. 1995 *Family, Gender, and Population in the Middle East*. Cairo: American University in Cairo.
- Oncu, Ayse 1997 "Crossing Borders into Turkish Sociology with Gunder Frank and Michel Foucault." *Contemporary Sociology* 26:267–270.

- Organization for Economic Cooperation and Development 1998 *Continuous Reporting System on Migration. SOPEMI 1998*. Paris: OECD.
- Owen, R. 1989 "The Movement of Labor In and Out of the Middle East over the Last Two Centuries: Peasants, Patterns, and Policies." In Georges Sabagh, ed., *The Modern Economic and Social History of the Middle East in World Context*. Cambridge, England: Cambridge University Press.
- Pamuk, Sevket 1988 "The Ottoman Empire in Contemporary Perspective." *Fernand Braudel Center Review* 11:127-151.
- Parsa, Misagh 1989 *Social Origins of the Iranian Revolution*. New Brunswick, N.J.: Rutgers University Press.
- 1995a "Conversion or Coalition? Ideology in the Iranian and Nicaraguan Revolutions." *Political Power and Social Theory* 9:1000-1022.
- 1995b "Entrepreneurs and Democratization: Iran and the Philippines." *Comparative Studies in Society and History* 37:803-830.
- Pulcini, Theodore 1993 "Trends in Research on Arab Americans." *Journal of American Ethnic History* 12:27-60.
- Rasler, Karen 1996 "Concessions, Repression, and Political Protest." *American Sociological Review* 61:132-151.
- Sabagh, Georges 1976 "Sociology." In Leonard Binder, ed., *The Study of the Middle East*. New York: John Wiley.
- 1993 "Labor Migrants in Kuwait: Will They Always Be Sojourners?" In Gema Martin Munoz, ed., *Democracia y Derechos Humanos en el Mundo Arabe*. Madrid: ICMA.
- , and Iman Ghazalla 1986 "Arab Sociology Today: A View from Within." *Annual Review of Sociology* 12:373-399.
- , and Mehdi Bozorgmehr 1987 "Are the Characteristics of Exiles Different from Immigrants? The Case of Iranians in Los Angeles." *Sociology and Social Research* 71:77-84.
- Sawaie, Mohammed, ed. 1985 *Arabic-Speaking Immigrants in the United States and Canada*. Lexington, Ky.: Mazda.
- Sell, Ralph R. 1988 "Egyptian International Labor Migration and Social Processes: Toward Regional Integration." *International Migration Review* 22:87-108.
- Sengstock, Mary C. 1982 *Chaldean-Americans: Changing Conceptions of Ethnic Identity*. New York: Center for Migration Studies.
- Skocpol, Theda 1982 "Rentier State and Shi'a Islam in the Iranian Revolution." *Theory and Society* 11:265-283.
- 1986 "Analyzing Causal Configurations in History: A Rejoinder to Nichols." *Comparative Social Research* 9:187-194.
- Sohrabi, Nader 1995 "Historicizing Revolutions: Constitutional Revolutions in the Ottoman Empire, Iran, and Russia, 1905-1908." *American Journal of Sociology* 100:1383-1447.
- Stahl, Charles W., and Farooq-i-Asarn 1990 "Counting Pakistanis in the Middle East: Problems and Policy Implications." *Asian and Pacific Population Forum* 4:1-10, 24-28.
- Suleiman, Michael W., and Baha Abu-Laban 1989 "The Arab Tradition in North America." *Arab Studies Quarterly* 11:1-13.
- Swan, Charles L., and Leila B. Saba 1974 "The Migration of a Minority." In Barbara C. Aswad, ed., *Arabic-Speaking Communities in American Cities*. New York: Center for Migration Studies.
- United Nations 1996 *World Population Monitoring 1993*. New York: UN.
- U.S. Bureau of the Census 1993 *1990 Census of Population Ancestry of the Population of the United States*. Washington: U.S. Government Printing Office.
- Waldinger, Roger, and Mehdi Bozorgmehr, eds. 1996 *Ethnic Los Angeles*. New York: Russell Sage Foundation, 1996.
- Wallerstein, Immanuel 1979 "The Ottoman Empire and the Capitalist World-Economy: Some Questions for Research." *Review* 2:389-408.

GEORGES SABAGH

## MIGRATION

See Internal Migration; International Migration.

## MILITARY SOCIOLOGY

Military sociology has been a relatively minor field in American sociology. Few sociologists conduct research and write on military topics, and few university departments offer courses of study in this field. To the layperson this circumstance may seem rather curious, given the sheer size and complexity of the military enterprise and the effects of war on individual societies and on the world order as a whole. The Vietnam War of the 1960s and the Persian Gulf War of 1991 are reminders of the impact that military institutions and operations have on American society.

There is no single reason for this lack of emphasis. One reason may be ideological aversion to military matters. Most sociologists fall on the

liberal side of the political spectrum, which has tended to emphasize the peace rather than the war aspect of international relations. Another reason may be that support for most military research requires clear policy applications, and applied research is of less interest to the majority of sociologists. Unlike other social sciences, sociology has not developed a sizable cadre of practitioners who conduct research and policy studies in those nonacademic settings where most military research takes place.

There are notable exceptions. One is the classic work of Samuel Stouffer and colleagues during World War II, *The American Soldier* (1949). Its studies of cohesion, morale, and race relations among combat units made critical contributions to both sociological theory and military personnel policies. The work of Morris Janowitz on organizational and occupational changes in the military is also a significant exception to the rule (Janowitz 1960). Charles Moskos at Northwestern University and David and Mady Segal at the University of Maryland are upholding this tradition today, and their universities are among the few that offer courses on military sociology.

In recent years most of the studies that could be classified as military sociology have been carried out in private research institutes or in military agencies, such as the Rand Corporation, the Brookings Institute, the Human Resources Research Organization (HumRRO), the Army Research Institute, and the Office of the Secretary of Defense. Study teams are generally interdisciplinary, with the fields of psychology, economics, political science, and management frequently represented. While these studies have a sociological complexion, they are not limited to sociological concepts or theory, and they usually have a military policy focus.

Much of the published work in this field consists of reports from these various organizations and agencies. No journal is devoted specifically to military sociology, although there is an American Sociological Association Section on Peace and War. *Armed Forces and Society* is an interdisciplinary journal that focuses on social science topics in the military, and *Military Psychology*, published by the Division on Military Psychology of the American Psychological Association, also covers topics of sociological interest.

This review emphasizes contemporary issues and studies in military sociology, most of which focus on the American military. For reviews that include historical perspectives the reader is referred elsewhere (D. R. Segal 1989; D. R. Segal and M. W. Segal 1991).

### THE BASIS OF SERVICE

Without question the most significant issue in American military sociology following World War II is the shift from conscription to voluntary service in 1973. Indeed, many contemporary issues in military sociology—recruiting policies, social representation, and race and gender concerns—flow directly or indirectly from this change in the basis of service.

The notion of voluntary service is not new; indeed, compulsory military service has been the exception rather than the rule in the history of most Western societies, although the meaning of “involuntary” and the nature of service has changed over time. Until the Civil War, U.S. military manpower was raised through state militias, which technically encompassed all “able-bodied men” and hence might be thought of as involuntary. The raising and maintenance of militia was left to the discretion and policies of individual states, and they varied greatly in their representation and effectiveness. Although the desirability and feasibility of national conscription has been debated since the beginning of American history, until 1948 it was employed only during the Civil War and during the World Wars I and II. At all other times enlistment was voluntary (Lee and Parker 1977).

The peacetime conscription adopted in 1948 was a significant and historic departure for U.S. military policy. It was prompted by a combination of factors, including Cold War tensions with the Soviet Union, the critical role being played by the United States in defense alliances, and a perception that America had been inadequately prepared for World War II. These conditions led to proposals for a large standing, or “active-duty,” military force that most believed could not be maintained by voluntary methods. The policy was supported by the American public at that time, and it was consistent with the longstanding value that all “able-bodied men” should serve their country.

The end of the peacetime draft occurred in 1973 at the end of the Vietnam War. The debate over draft policies had been intensifying during the 1960s, particularly over the issue of equity. Given the growth in the American youth population (from the post-World War II baby boom), not all men were needed for the military, and a succession of draft-exemption policies—college, occupational, marriage—fueled the debate over the fairness and equity of the draft. Not only were all able-bodied men not serving, but those exempted from service tended to be from more affluent classes. In an attempt to solve the equity problem, nearly all exemptions were eliminated during the late 1960s, and a national lottery system was adopted in 1969. These changes did not quell the debate; indeed, they probably doomed the draft, since they coincided with one of the most unpopular and unsuccessful wars in America's history. Ending exemptions for college students at a time when college students were leading the movement against the Vietnam War simply fanned the flames of opposition.

On the other side of the issue, most defense and military leaders, including many members of Congress, initially opposed ending the draft. The conditions leading to a large peacetime military force (over two million active-duty personnel) were unchanged and were not under debate. If such a large force was maintained by voluntary means, it was argued, it would be either (1) a mercenary force lacking patriotic motivation, which might be either adventuresome or ineffective in large-scale combat; or (2) socially unrepresentative, thereby placing the burden of combat on minorities, the poor, and the uneducated; or (3) too costly and therefore unaffordable (D. R. Segal 1989).

The shift to voluntary service followed definitive recommendations from a presidential commission, known as the Gates Commission, which concluded that a large peacetime force could be adequately maintained with an all-voluntary force (AVF), provided that basic pay was increased to be competitive with comparable civilian jobs. The commission argued that paying enlisted personnel below-market wages not only was unfair but entailed "opportunity" costs in the form of lost productivity of persons who are compelled to work at jobs they would not voluntarily perform. If military pay and benefits were competitive, the commission's studies concluded, a volunteer force would be socially representative and would be as

effective as a drafted force. Although some attention was given to the social costs of the draft, there is little question that economic analysis and arguments formed the central basis of the commission's report (President's Commission 1970).

### THE ALL-VOLUNTEER FORCE

The AVF seemed to work well for the first several years, and the dire predications of those who opposed ending the draft did not materialize. Enlistment requirements were met, and neither quality nor representation appeared much different from draft-era enlistments. The predictions of the Gates Commission seemed to be valid (Cooper 1977).

During the late 1970s, however, this positive picture faded. Military pay and benefits did not keep up with the civilian sector, and post-Vietnam War antiwar attitudes contributed to a generally negative image of the military among youth. The number and quality of new recruits began declining, and the services were forced to enlist more applicants with low aptitude scores and without high school diplomas. To make matters worse, a misnorming error was discovered in the military aptitude test (the Armed Forces Qualifying Test [AFQT]), and the quality of personnel was even lower than had been thought. By 1980 the Army was severely affected, with enlistments falling significantly below requirements and with the proportion of low-aptitude recruits—those reading at fourth- and fifth-grade levels—reaching 50 percent (Eitelberg 1988).

Some claimed that the Gates Commission had been wrong after all, and called for a return to the draft. Others, including the Secretary of the Army, proposed eliminating the aptitude and education standards that had been used since the draft era, arguing that they were not important for military job performance and were potentially discriminatory. Yet others argued that both the AVF and quality standards could be maintained but only by increasing pay and benefits to enable competition with the civilian sector. Critical to this third argument were the policy studies led by the Rand Corporation showing a relationship between aptitudes and military job performance and the feasibility of economic incentives for enlistment (Armor et al. 1982; Polich et al. 1986).

Substantial increases in pay and other benefits in the early 1980s, as well as changes in recruiting techniques including advertising, had a dramatic impact on recruiting success. By the late 1980s, all military services, including the Army, were not only meeting requirement quotas but setting new records in the quality of personnel, surpassing even the draft years in the proportion of recruits with high school diplomas and with higher aptitudes. Most military leaders, many of whom had been skeptical about ending the draft, became strong supporters of the AVF and claimed that morale and skills were at all-time highs and discipline problems were at all-time lows. Although this success did require significant increases in the military personnel budget, both the Congress and the public seemed prepared to pay this cost (Bowman et al. 1986).

The one argument against voluntary service that had never been tested empirically was the combat effectiveness of a market-motivated military. The dramatic success of the American forces in the 1991 Persian Gulf War put an end to any doubt about the ability of a volunteer military to operate successfully. Although high-technology weapons systems received much of the publicity during the war, most military leaders gave as much credit to the skill and morale of their troops, especially in comparison to the poorly trained and demoralized Iraqi soldiers.

Two major events, the collapse of the Soviet Union and the U.S. budget deficit crisis, combined to force major reductions in the size of the active military, which also meant fewer new recruits needed to sustain the smaller force. In spite of lower recruiting targets, recruit quality began edging downward again after the Persian Gulf War. The main causes were thought to be a combination of less positive attitudes toward military on the part of youth, and an increase in the proportion of youth deciding to attend college (Office of the Assistant Secretary of Defense 1998). Although some concerns have surfaced in the military community about these downward trends, by 1998 overall recruit quality was still similar to the successful levels seen during the late 1980s. A mitigating factor is the trend in the number of 18-year-olds, which declined between 1980 and 1992 but has been increasing since then and is expected to continue to grow until 2010.

## **SOCIAL REPRESENTATION AND ACCESS**

Social representation is the degree to which an armed force represents the population from which it has been drawn. Although representation has been a longstanding issue in the United States, concerns were heightened during the AVF era, especially since that era happened to coincide with increasing emphasis on ethnic and gender diversity throughout American society. Racial representation was an issue in both the Vietnam War and the Persian Gulf War, when some civil rights leaders alleged that African Americans were overrepresented in armed forces and hence carried an unfair burden of casualties. Gender representation has developed as a major issue in concert with the women's rights movement, generating major policy changes concerning the participation of women in the military. Finally, with the election of President Bill Clinton and his campaign commitment to overturn the military ban on gays and lesbians, for the first time the issue of sexual orientation became the focus of a major military policy debate.

Debates about social representation usually divide over positions on three basic principles: equity, effectiveness, and responsiveness. The equity principle has two aspects. One is the equity of obligation, meaning that the burden of service and the risks of combat should be borne by all segments of the citizenry. Since a national military force defends the interests of an entire country and entails the risk of casualties, this obligation should be shared uniformly by all citizens, or at least all able-bodied persons. Placing this burden unequally upon certain socioeconomic or racial groups not only violates the canons of fairness in democratic societies but also potentially undermines troop morale and motivation if they perceive the disproportion as unfair (Congressional Budget Office 1989). The other aspect is the equity of access, which often means that racial minorities, women, and homosexuals should be allowed to serve in the military according to their abilities and interests, without artificial restrictions or ceilings.

The principle of effectiveness is raised because most military experts believe that the quality of military personnel in terms of education and aptitude directly influence training and combat effectiveness, and therefore a force that

underrepresents quality personnel will be less effective than a representative force. Although the armed services have attained their quality goals since the late 1980s, there is no guarantee that they can always do so, particularly if military pay and benefits do not keep up with civilian compensation. The effectiveness principle is also invoked by some military experts as a counterargument to the equity position, arguing that allowing women in combat units or open homosexuals in any type of unit will undermine unit morale and cohesion, thereby reducing military effectiveness.

Finally, a representative armed force is sometimes seen as a requirement for military acceptance of civilian control as well as respect for democratic values and institutions. A military force that overrepresents particular social strata or groups (not necessarily lower strata) might place parochial interests or values over the interests and values of the society as a whole. A force drawn proportionately from all major sectors of a society—all regions, races, and social classes—is viewed as one most likely to respect and advance the shared values and goals of the total society.

**Racial Representation.** One of the most frequently expressed concerns about representation during the AVF era has been the overrepresentation of African Americans and lower socioeconomic groups in the armed services, especially in the Army. Given the early experience with the AVF, this concern is not without some justification, although the situation improved considerably during the 1980s.

Table 1 shows the percentage of black recruits for selected years since the end of the draft in 1973. Black overrepresentation reached its peak in 1979, before the aptitude misnorming error was discovered. Black representation declined during the early 1980s after the military corrected its screening test. It edged back up during the late 1980s, but it declined again during the early 1990s, when recruiting was scaled back in response to the smaller permanent force. Even during the draft years, black representation exceeded the black proportion of the youth population, due in part to higher voluntary enlistment and—prior to the Vietnam War—in part to college exemption policies that applied disproportionately to white youth. During most of the 1990s, black representation in the total active forces was only a few points higher

than a comparable civilian population, although the trend was moving upward again toward the end of the decade, particularly in the Army.

It should not be inferred from these data that black soldiers have a risk of death or injury in combat that is disproportionately higher than that of white soldiers. In fact, black enlisted members in the active-duty forces tend to be underrepresented in combat occupations such as infantry, gun crews, and combat ships and overrepresented in administrative, clerical, and supply occupations. As a result, black casualty rates in wartime are in rough proportion to their share of the general civilian population. For example, in the Vietnam War black fatalities were approximately 12 percent of all American fatalities (Moskos and Butler 1996).

Some argue that these differences in representation present a policy problem, although there is no consensus on what, if anything, should be done about it. While some see black overrepresentation as a problem of burden, others see it as an opportunity for advancement in one of the best-integrated occupational sectors in American society. Higher black representation reflects a greater interest in military careers among African Americans, and any attempt to limit black enlistments could be viewed as a denial of equal opportunity. Indeed, the military in general, and the Army in particular, has been held up as a model for racial integration that should be emulated by other institutions (Moskos and Butler 1996).

**Women in the Military.** Another topic that has received considerable attention from military sociologists is the role of women in the military. Until 1967 women's participation in the military was limited to 2 percent by law, and most military jobs were closed to women. Even after the restriction was lifted by Congress, the representation of women did not change until the AVF policy and the opening of more noncombat jobs to women. After 1973 the percentage of women enlistees increased steadily, reaching a plateau during the 1980s of about 12 percent. Following major policy changes regarding combat restrictions in the early 1990s, the percentage of women recruits rose again, reaching a second plateau of about 18 percent during the late 1990s. This representation of women in the U.S. military is the highest among all NATO countries (Office of the Assistant Secretary of Defense 1998; Stanley and Segal 1988).

**Percentage of Black Recruits, Active Forces**

Fiscal Year	Army	Navy	Marine Corps	Air Force	Total Active Forces	U.S. Population, Ages 18–24
1973	19	11	22	14	17	12
1979	37	16	28	17	26	13
1983	22	14	17	14	18	14
1989	26	22	18	12	22	14
1991	20	17	14	11	17	14
1993	20	17	12	13	17	14
1995	22	20	13	14	18	14
1997	23	20	14	17	20	14

*Table 1*

The growing participation of women in the military reflects several forces, including the need for manpower in the AVF era as well as the increased demands for equity in the treatment of men and women in the workplace. Although all military services recruited more women to help alleviate shortfalls during the 1970s, the Department of Defense has also been pressured by Congress and various interest and advisory groups to enlarge the opportunities for women. The primary barriers to expanding the role of women were statutory restrictions on the assignment of women to combat units and missions. These statutes were repealed by Congress in 1993, and by 1994 the Department of Defense had substantially modified combat restriction regulations. All combat aircraft and combat ships were opened to women (except submarines, for privacy reasons), and the only units that could be closed to women were those that involved “direct ground combat” such as infantry, armor, field artillery, and special forces (Armor 1996).

In spite of the increased participation of women, gender representation continues to be a topic of debate. Arguments against allowing women in combat jobs include physical and emotional differences, privacy and sexual behavior, impacts on unit cohesion and morale, and a basic moral position that women should be protected from the high risk of death, injury, capture, or torture faced by those in ground combat units. Proponents of opening combat jobs to women counter that (1) women should be evaluated individually and not as a group for physical and emotional suitability for combat; (2) experience with women in noncombat jobs has shown no serious adverse impacts on unit cohesion or morale; and (3) women currently

serve in jobs and units that are exposed to increased risks of death, injury, or capture, thereby rendering the moral argument moot and in conflict with current policies (M. W. Segal 1982). Consistent with this latter view, recent reviews on the integration of women in the military report substantial progress (Harell and Miller 1997).

A second major issue involving women in the military is sexual harassment. This problem is obviously not unique to the military; it continues to be an issue throughout American society. It may be more acute, however, in a traditional male institution like the military where women constitute small minorities in many types of jobs. Little hard evidence was available on the extent of sexual harassment in the military until a comprehensive survey was conducted at the request of a 1988 Task Force on Women in the Military. The results indicated rather serious levels of sexual harassment; altogether, 64 percent of women experienced some form of sexual harassment during a one-year period, and 5 percent reported actual or attempted rape or sexual assault (Martindale 1990). These rates of harassment were considerably higher than those found in surveys of federal civilian employees, and they presaged several serious incidents during the 1990s involving sexual harassment or other sexual improprieties in three of the four services: the 1991 Navy Tailhook incident; the 1996 Army Aberdeen Proving Grounds scandal; and the 1997 Air Force dismissal of Kelly Flinn, its first female bomber pilot, for adultery with the husband of an enlisted woman.

A follow-up survey on sexual harassment was conducted in 1995 after the military services had tightened regulations, clarified policies and practices, and increased training with regards to sexual



harassment. The new survey showed some improvement, but rates of sexual harassment were still high, with 55 percent of women reporting some form of unwanted sexual behavior or advances and 4 percent reporting actual or attempted rape or some other form of sexual assault (Bastion et al. 1996).

These high rates of sexual harassment may be due to a number of factors, but are most likely associated with the increased presence of women in a traditional male workplace, and a workplace where a traditional "male" culture may be insensitive if not hostile to women. Another factor may be the attitudes and actions of military leaders and supervisors. Only about half of the women responding to the 1988 survey believed that the senior leadership of their service and their installation made "honest and reasonable" attempts to prevent sexual harassment, and only 60 percent said their immediate supervisor did so. In 1995 these rates had increased by only a few percentage points. Whether or not these figures truly reflect the behavior of military leaders, there is a perception among many military women that the military leadership is still not giving enough emphasis to the problem of sexual harassment.

**Policies on Sexual Orientation.** Since World War I, persons who acknowledge homosexual status or behavior have been excluded from U.S. military service, either by antisodomy provisions of the Uniform Code of Military Justice or by explicit Department of Defense regulations. This policy was challenged by President Bill Clinton in 1993, who directed the secretary of defense to prepare a "draft of an Executive Order ending discrimination on the basis of sexual orientation in determining who may serve in the Armed Forces" (Rostker and Harris 1993, p. 1). Most senior military officers opposed the change, expressing beliefs that open homosexuals would hamper maintenance of good order and discipline and would undermine unit cohesion and combat effectiveness. Various studies were commissioned, including a major study by the Rand Corporation which concluded that, based on experience in other countries and on certain U.S. police and fire departments, the homosexual ban could be lifted without serious adverse effects on unit cohesion and military performance (Rostker and Harris 1993). The Rand study also noted that nondeclared homosexuals had always served in some number in the

United States and foreign militaries without undue consequences.

The center of the debate shifted to Congress, and particularly to Democratic Senator Sam Nunn, Chairman of the Armed Services Committee, who had made it clear before the 1992 election that he was skeptical about this change and would hold hearings on the matter. Based on these hearings, including testimony from both civilian experts and military leaders, a compromise proposal was crafted, which became known as the "Don't ask, don't tell" policy; this proposal became a federal statute in 1994. The statute states that sexual orientation is a private and personal matter, and is not a bar to military service. It prohibits routine questions about sexual orientation on recruiting forms, but also says that engaging in homosexual conduct is grounds for discharge, and that a voluntary statement by a person acknowledging a gay or lesbian orientation is considered homosexual conduct (Office of the Under Secretary of Defense 1998).

Needless to say, gay and lesbian activists were not satisfied with this compromise, and especially with the provision that service members continue to be discharged for merely telling someone that they are homosexual (Rimmerman 1996). The "Don't ask, don't tell" policy has been challenged in federal courts, but with the exception of several cases where a discharged gay or lesbian has been reinstated because information was obtained improperly, the statute itself has been upheld by several appellate courts.

## OTHER ISSUES

There are a number of other topics that are studied by military sociologists or that raise important sociological issues. While they have not received the same degree of attention as the topics reviewed above, they deserve mention.

**Sociology of Combat.** A major focus of *The American Soldier* studies during World War II, the sociology of combat deals with the social processes involved in combat units, such as unit cohesion and morale, leader-troop relations, and the motivation for combat. Recent works include studies of the "fragging" incidents in the Vietnam War and the role of ideology in combat motivation (Moskos 1988).

**Family Issues.** The proportion of the military personnel who are married increased from around 40 percent during the draft years to 57 percent by 1995. The current active force also has a higher proportion of career personnel than the draft-era forces (50 percent compared to 25 percent), which means more families and more family concerns. Much of the increase in military benefits is related to families—housing improvements, medical insurance, overseas schools, child care. Family policy issues include the role and rights of spouses (especially officer spouses) and the issue of child care when single-parent members are deployed in a conflict (M. W. Segal 1986).

**The Military as Welfare.** Somewhat at odds with the social representation issue, some argue that the military should provide opportunity for educational and occupational advancement to the less advantaged in society. The most dramatic example was Project 100,000, begun in 1966 as part of President Lyndon Johnson's War on Poverty, whereby 100,000 men (mostly black) who did not meet education and aptitude requirements were offered enlistments. According to recent follow-up studies of this group as well as a group of low-aptitude recruits who entered the AVF during the misnorming era, military training and experience do not appear to offer advantages when compared to civilian experiences (D. R. Segal 1989).

**Military Social Organization.** Given changes in military organization at several levels—from draft to AVF, from combat-intensive jobs to technical and support jobs, and from leadership to rational management—some have argued that the military is changing from an institution or “calling” legitimized by normative values to an occupation legitimized by a market orientation (Moskos and Wood 1988). While some of these changes apply to the military as well as to many other American institutions, others suggest that the role of institutional values and traditions is still a dominant characteristic of the American military (D. R. Segal 1989).

**War and Peace.** The most profound impacts of national security policies and their associated military forces are on the relations between whole nations. There is very little sociological work at this level, although some of the issues would seem to be fairly critical for sociological theory (e.g., the effectiveness of deterrence policies during the

Cold War; the role of effects of military alliances; the consequences of war for societal changes). One of the few sociological discussions of these broader issues is found in D. R. Segal (1989).

**Comparative Perspectives.** As in many other fields of sociology, comparative studies of sociomilitary issues across nations are relatively scarce. Exceptions are some comparative studies of whether the military is an institution or an occupation (Moskos and Wood 1988) and of women in military forces (Stanley and Segal 1988).

## CONCLUSION

Although military sociologists are few in number, it is not a small field. The military is the largest single government agency, and it truly represents a microcosm of the larger society. The types of military issues that can be addressed by social scientists have important ramifications for military policy as well as for the development of sociology as a discipline. On the one hand, sociological concepts and perspectives have contributed to rationalizing the basis of service, studying the problems of social representation, and improving the role of women in the military. On the other hand, studies of the structure and processes of military institutions and policies can enhance sociological understanding and insights about important forms of social behavior, thereby contributing ultimately to advances in social theory.

(SEE ALSO: *Gender; Race; Peace; War*)

## REFERENCES

- Armour, D. J. 1996 “Race and Gender in the U.S. Military.” *Armed Forces and Society* 23:7-27
- , R. L. Fernandez, Kathy Bers, and Donna Schwarzbach 1982 *Recruit Aptitudes and Army Job Performance*, R-2874-MRAL. Santa Monica, Calif.: Rand.
- Bastian, L. D., A. R. Lancaster, and H.E. Reyst 1996 *1995 Sexual Harassment Survey*. Arlington, Va.: Defense Manpower Data Center.
- Bowman, W., R. Little, and G. T. Sicilia 1986 *The All-Volunteer Force after a Decade*. Washington: Pergamon-Brassey's.
- Congressional Budget Office 1989 *Social Representation in the U.S. Military*. Washington: Congress of the United States.

- Cooper, Richard V. L. 1977 *Military Manpower and the All-Volunteer Force* R-1450-ARPA. Santa Monica, Calif.: Rand.
- Eitelberg, Mark J. 1988 *Manpower for Military Occupations*. Alexandria, Va.: Human Resources Research Organization.
- Harrell, M. C., and L. L. Miller 1997 *New Opportunities for Military Women: Effects upon Readiness, Cohesion, and Morale*, MR-896-OSD. Santa Monica, Calif.: Rand.
- Janowitz, Morris 1960 *The Professional Soldier*. Glencoe, Ill.: Free Press.
- Laurence, J. H., and P. F. Ramsberger 1991 *Low-Aptitude Men in the Military: Who Profits, Who Pays?* Alexandria, Va.: Human Resources Research Organization.
- Lee, G. C., and G. Y. Parker 1977 *Ending the Draft: The Story of the All-Volunteer Force*, FR-PO-77-1. Alexandria, Va.: Human Resources Research Organization.
- Martindale, Melanie 1990 *Sexual Harassment in the Military: 1988*. Arlington, Va.: Defense Manpower Data Center.
- Moskos, C. C. 1988 *Soldiers and Sociology*. Alexandria, Va.: U.S. Army Research Institute for the Behavioral and Social Sciences.
- , and J. S. Butler 1996 *All That We Can Be*. New York: Basic.
- , and F. R. Wood 1988 *The Military: More Than Just a Job?* Washington: Pergamon-Brassey's.
- Office of the Assistant Secretary of Defense (Force Management) 1998 *Population Representation in the Military Services, FY 97*. Washington: U.S. Department of Defense.
- Office of the Under Secretary of Defense (Personnel and Readiness) 1998 *Review of the Effectiveness of the Application and Enforcement of the Department's Policy on Homosexual Conduct*. Washington: U.S. Department of Defense.
- Polich, J. M., J. N. Dertouzos, and S. J. Press 1986 *The Enlistment Bonus Experiment*, R3353-FMP. Santa Monica, Calif.: Rand.
- President's Commission 1970 *Report of the President's Commission on an All Volunteer Force*. Washington: U.S. Government Printing Office.
- Rimmerman, C. A., ed. 1996 *Gay Rights, Military Wrongs*. New York: Garland.
- Rostker, B. D., and S. A. Harris, study directors 1993 *Sexual Orientation and U.S. Military Personnel Policy*, MR-323-OSD. Santa Monica, Calif.: Rand.
- Segal, D. R. 1989 *Recruiting for Uncle Sam*. Lawrence: University of Kansas Press.
- , and M. W. Segal 1991 "Sociology, Military." *International Military and Defense Encyclopedia*. New York: Macmillan.
- Segal, M. W. 1982 "The Argument for Female Combatants." In Nancy Loring Goldman, ed., *Female Soldiers: Combatants or Noncombatants?* Westport, Conn.: Greenwood.
- 1986 "The Military and the Family as Greedy Institutions." *Armed Forces and Society* 13:9-38.
- Stanley, S. C., and M. W. Segal 1988 "Military Women in NATO: An Update." *Armed Forces and Society* 14:559-585.
- Stouffer, S. A., et al. 1949 *The American Soldier*. Princeton, N.J.: Princeton University Press.

DAVID J. ARMOR

## MODERNIZATION THEORY

Modernization theory is a description and explanation of the processes of transformation from traditional or underdeveloped societies to modern societies. In the words of one of the major proponents, "Historically, modernization is the process of change towards those types of social, economic, and political systems that have developed in Western Europe and North America from the seventeenth century to the nineteenth and have then spread to other European countries and in the nineteenth and twentieth centuries to the South American, Asian, and African continents" (Eisenstadt 1966, p. 1). Modernization theory has been one of the major perspectives in the sociology of national development and underdevelopment since the 1950s. Primary attention has focused on ways in which past and present premodern societies become modern (i.e., Westernized) through processes of economic growth and change in social, political, and cultural structures.

In general, modernization theorists are concerned with economic growth within societies as indicated, for example, by measures of gross national product. Mechanization or industrialization are ingredients in the process of economic growth. Modernization theorists study the social, political, and cultural consequences of economic growth and the conditions that are important for industrialization and economic growth to occur. Indeed, a degree of circularity often characterizes discussions of social and economic change involved in modernization processes because of the notion,

embedded in most modernization theories, of the functional compatibility of component parts. The theoretical assumptions of modernization theories will be elaborated later.

It should be noted at the outset that the sociological concept of modernization does not refer simply to becoming current or “up to date” but rather specifies particular contents and processes of societal changes in the course of national development. Also, modernization theories of development do not necessarily bear any relationship to more recent philosophical concepts of “modernity” and “postmodernity.” Modernity in philosophical and epistemological discussions refers to the perspective that there is one true descriptive and explanatory model that reflects the actual world. Postmodernity is the stance that no single true description and explanation of reality exists but rather that knowledge, ideology, and science itself are based on subjective understandings of an entirely relational nature. While their philosophical underpinnings place most modernization theories of development into the “modern” rather than the “postmodern” context, these separate uses of the term *modernity* should not be confused.

Also, modernization, industrialization, and development are often used interchangeably but in fact refer to distinguishable phenomena. Industrialization is a narrower term than modernization, while development is more general. Industrialization involves the use of inanimate sources of power to mechanize production, and it involves increases in manufacturing, wage labor, income levels, and occupational diversification. It may or may not be present where there is political, social, or cultural modernization, and, conversely, it may exist in the absence of other aspects of modernization. Development (like industrialization) implies economic growth, but not necessarily through transformation from the predominance of primary production to manufacturing, and not necessarily as characterized by modernization theory. For example, while modernization theorists may define development mainly in terms of economic output per capita, other theorists may be more concerned about development of autonomous productive capacity, equitable distribution of wealth, or meeting basic human needs. Also, while modernization theories generally envision democratic and capitalist institutions or secularization of belief systems as components of modern society,

other development perspectives may not. Indeed, dependency theorists even talk about the “development of underdevelopment” (Frank 1966).

Each of the social science disciplines pays particular attention to the determinants of modern structures within its realm (social, political, economic) and gives greater importance to structures or institutions within its realm for explaining other developments in society. Emphasis here is given to sociological modernization theory.

Although there are many versions of modernization theory, major implicit or explicit tenets are that (1) societies develop through a series of evolutionary stages; (2) these stages are based on different degrees and patterns of social differentiation and reintegration of structural and cultural components that are functionally compatible for the maintenance of society; (3) contemporary developing societies are at a premodern stage of evolution and they eventually will achieve economic growth and will take on the social, political, and economic features of western European and North American societies which have progressed to the highest stage of social evolutionary development; (4) this modernization will result as complex Western technology is imported and traditional structural and cultural features incompatible with such development are overcome.

At its core modernization theory suggests that advanced industrial technology produces not only economic growth in developing societies but also other structural and cultural changes. The common characteristics that societies tend to develop as they become modern may differ from one version of modernization theory to another, but, in general, all assume that institutional structures and individual activities become more highly specialized, differentiated, and integrated into social, political, and economic forms characteristic of advanced Western societies.

For example, in the social realm, modern societies are characterized by high levels of urbanization, literacy, research, health care, secularization, bureaucracy, mass media, and transportation facilities. Kinship ties are weaker, and nuclear conjugal family systems prevail. Birthrates and death rates are lower, and life expectancy is relatively longer. In the political realm, the society becomes more participatory in decision-making processes,

and typical institutions include universal suffrage, political parties, a civil service bureaucracy, and parliaments. Traditional sources of authority are weaker as bureaucratic institutions assume responsibility and power. In the economic realm, there is more industrialization, technical upgrading of production, replacement of exchange economies with extensive money markets, increased division of labor, growth of infrastructure and commercial facilities, and the development of large-scale markets. Associated with these structural changes are cultural changes in role relations and personality variables. Social relations are more bureaucratic, social mobility increases, and status relations are based less on such ascriptive criteria as age, gender, or ethnicity and more on meritocratic criteria. There is a shift from relations based on tradition and loyalty to those based on rational exchange, competence, and other universally applied criteria. People are more receptive to change, more interested in the future, more achievement-oriented, more concerned with the rights of individuals, and less fatalistic.

Underlying the description of social features and changes that are thought to characterize modern urban industrial societies are theoretical assumptions and mechanisms to explain the shift from traditional to modern societal types. These explanatory systems draw upon the dominant theoretical perspectives in the 1950s and 1960s, growing out of classical evolutionary, diffusion, and structural-functionalist theories.

The evolutionary perspective, stemming from Spencer, Durkheim, and other nineteenth-century theorists, contributed the notion that societies evolve from lower to higher forms and progress from simple and undifferentiated to more complex types. Western industrial society is seen as superior to preindustrial society to the extent that it has progressed through specialization to more effective ways of performing societal functions. Diffusionists added the ideas that cultural patterns associated with modern society could be transferred via social interaction (trade, war, travelers, media, etc.) and that there may be several paths to development rather than linear evolution. Structural functionalists (Parsons 1951; Hoselitz 1960; Levy 1966) emphasized the idea that societies are integrated wholes composed of functionally compatible institutions and roles, and that societies

progress from one increasingly complex and efficient social system to another. This contributed to the notion that internal social and cultural factors are important determinants or obstacles of economic change.

Research by Smelser (1969) draws on all three traditions in describing modernization of society through processes of social differentiation, disturbances, and reintegration. In a manner similar to other conceptions of modernization, Smelser emphasizes four major changes: from simple to complex technology, from subsistence farming to commercial agriculture, from rural to urban populations, and, most important, from animal and human power to inanimate power and industrialization.

Parsons's later theoretical work (1964) also combines these perspectives in a neo-evolutionist modernization theory that treats societies as self-regulated structural functional wholes in which the main processes of change are social differentiation and the discovery (or acquisition through diffusion) of certain "evolutionary universals" such as bureaucratic organizations and money markets. These, in turn, increase the adaptive capacity of the society by providing more efficient social arrangements and often lead to a system of universalistic norms, "which, more than the industrial revolution itself, ushered in the *modern* era of social evolution" (Parsons 1964, p. 361). A similar neo-evolutionist social differentiation theory of modernization is provided by Eisenstadt (1970).

Another early influence on modernization theory was Weber's work on the Protestant ethic. This work stressed the influence of cultural values on the entrepreneurial behavior of individuals and the rise of capitalism. Contemporary theorists in the Weberian tradition include Lerner, McClelland, Inkeles, and Rostow. Lerner's (1958) empirical studies in several Middle Eastern societies identified empathy, the capacity to take the perspective of others, as a product of media, literacy, and urbanization and as a vital ingredient in producing rational individual behavior conducive to societal development. McClelland (1961) felt that prevalence of individuals with the psychological trait of high "need for achievement" was the key to entrepreneurial activity and modernization of society. In a similar vein, Inkeles and Smith (1974) used interview data from six societies to generate a set

of personality traits by which they defined “modern man.” They felt that the prevalence of individual modernity in society was determined by such factors as education and factory experience and that individual modernity contributed to the modernization of society. Finally, Rostow’s (1960) well-known theory of the stages of economic growth, which he derived from studying Western economic development, emphasized the importance of new values and ideas favoring economic progress along with education, entrepreneurship, and certain other institutions as conditions for societies to “take off” into self-sustained economic growth.

All of these versions of modernization theory depict a gradual and more or less natural transition from “traditional” social structures to “modern” social structures characteristic of Western European and North American societies. More specifically, these theories tend to share to one degree or another the views that (1) modern people, values, institutions, and societies are similar to those found in the industrialized West, that is, the direction of change tends to replicate that which had already occurred in Western industrial societies; (2) tradition is opposite to and incompatible with modernity; (3) the causes of delayed economic and social development (i.e., underdevelopment) are to be found within the traditional society; (4) the mechanisms of economic development also come primarily from within societies rather than from factors outside of the society; and (5) these internal factors (in addition to industrial development) tend to involve social structures, cultural institutions, or personality types.

In keeping with this orientation, empirical studies of sociological modernization tend to deal with the internal effects of industrialization or other economic developments on traditional social institutions or with the social, political, and cultural conditions that facilitate or impede economic growth within traditional or less-developed societies. Examples might include research on the impact of factory production and employment on traditional family relations or the effects of an indigenous land tenure system on the introduction of cash crop farming in society.

Even though modernization theory since the 1960s has been dominated by and sometimes equated with Parsons’s neo-evolutionary theory, it is

clear that there is no single modernization theory but rather an assortment of related theories and perspectives. In addition to those mentioned, other important contributors of theoretical variants include Hagan (1962), Berger, Berger, and Kellner (1973), Bendix (1964), Moore (1967), Tiryakian (1985), and Nolan and Lenski (1999). Useful reviews include Harrison (1988), Harper (1993), and Jaffee (1998).

Since the 1960s, many critiques of modernization theory and the emergence of competing theories of development have eroded support for modernization theory. Foremost among these are dependency, world systems, and neo-Marxist theories, all of which criticize the ethnocentricity of the modernization concept and the bias in favor of dominant capitalist interests. The focus of these theories is on explaining the contemporary underdevelopment of Third World countries or regions of the world in terms of colonization, imperialist interference, and neocolonial exploitation of developing countries since their gaining of independence. In these counterperspectives, both development and underdevelopment are viewed as part of the same process by which certain “center” countries or regions become economically advanced and powerful at the expense of other “periphery” areas. Rather than explaining development and underdevelopment by the presence or absence of certain internal institutions or personalities, these alternative theories argue that both result from unequal exchange relations and coalitions of interests associated with the structural position of societies in the global economy. Rather than interpreting underdeveloped societies as traditional or archaic, both underdeveloped and developed societies are contemporary but asymmetrically linked parts of capitalist expansion. Both are relatively “modern” phenomena.

Attention to modernization theory in sociology has declined as a result of the theoretical and empirical weaknesses raised especially during the 1970s. Nevertheless, it is still the dominant perspective among government officials and international agencies concerned with third world development. Hoogvelt has noted its influence on development policies as follows:

*Because modernisation theories have viewed the total transformation, that is westernisation, of*

*developing countries to be an inescapable outcome of successful diffusion of the Western economic/technological complex, by methodological reversal it is argued that a reorganization of existing social and cultural as well as political patterns in anticipation of their compatibility with the diffused Western economic/technological complex may in fact facilitate the very process of this diffusion itself. This monumental theoretical error—which to be fair was not always committed by the theorists themselves—has in fact been made and continues to be made by modernisation policy-makers such as those employed by Western government, U.N. organizations, the World Bank, and so forth. (1978, pp. 60–61)*

Thus, various indicators of social, political, and cultural development (such as degree of urbanization, high literacy rates, political democracy, free enterprise, secularization, birth control, etc.) have frequently been promoted as “conditions” for development.

Interestingly, as modern structures and institutions have spread around the world and created economic, political, social, and cultural linkages, an awareness of global interdependence and of the ecological consequences of industrial development and modern lifestyles has grown. It is now clear that finite natural resources and the nature of the global ecosystem could not sustain worldwide modern conditions and practices of European and North American societies even if modernization theory assumptions of evolutionary national development were correct. Thus, new visions and interpretations of national and global development have already begun to replace classical modernization theory.

Some selected publications readers may wish to consult on this topic include Billet (1993), Inglehart (1997), McMichael (1996), Roberts and Hite (1999), Roxborough (1988), and Scott (1995).

(SEE ALSO: *Global Systems Analysis; Industrialization in Less Developed Countries*)

REFERENCES

Bendix, Reinhold 1964 *Nation-Building and Citizenship: Studies of Our Changing Social Order*. New York: John Wiley.

Berger, Peter L., Brigitte Berger, and Hansfried Kellner 1973 *The Homeless Mind: Modernization and Consciousness*. New York: Vintage.

Billet, Bret L 1993 *Modernization Theory and Economic Development: Discontent in the Developing World*. Westport, Conn: Praeger.

Eisenstadt, S. N. 1966 *Modernization: Protest and Change*. Englewood Cliffs, N.J.: Prentice-Hall.

— 1970 “Social Change and Development.” In S. N. Eisenstadt, ed., *Readings in Social Evolution and Development*. Oxford: Pergamon.

Frank, Andre Gunder 1966 “The Development of Underdevelopment.” *Monthly Review* 18(4):17–31.

Hagen, Everett E. 1962 *On the Theory of Social Change*. Homewood, Ill.: Dorsey.

Harper, Charles L. 1993 *Exploring Social Change*, 2nd ed. Englewood Cliffs, N.J.: Prentice Hall.

Harrison, David 1988 *The Sociology of Modernization and Development*. London: Unwin Hyman.

Hoogvelt, Ankie M. M. 1978 *The Sociology of Developing Societies*, 2nd ed. London: Macmillan.

Hoselitz, Berthold F. 1960 *Sociological Aspects of Economic Growth*. New York: Free Press.

Inglehart, Ronald 1997 *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton, NJ: Princeton University Press.

Inkeles, Alex, and David H. Smith 1974 *Becoming Modern*. Cambridge, Mass.: Harvard University Press.

Jaffee, David 1998 *Levels of Socio-Economic Development Theory*. Westport, Conn: Praeger.

Lerner, Daniel 1958 *The Passing of Traditional Society: Modernizing the Middle East*. New York: Free Press.

Levy, Marion, Jr. 1966 *Modernization and the Structures of Societies*, vol. 1. Princeton: Princeton University Press.

McClelland, David C. 1961 *The Achieving Society*. New York: Free Press.

McMichael, Philip 1996 *Development and Social Change: A Global Perspective*. Thousand Oaks, Calif.: Pine Forge.

Moore, Barrington 1967 *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Nolan, Patrick, and Gerhard E. Lenski 1999 *Human Societies: An Introduction to Macrosociology*, 8th ed. New York: McGraw-Hill.

Parsons, Talcott 1951 *The Social System*. New York: Free Press.

- 1964 “Evolutionary Universals in Society.” *American Sociological Review* 29:339–357.
- Roberts, J. Timmons, and Amy Hite 1999 *From Modernization to Globalization: Social Perspectives on International Development*. Malden, Mass.: Blackwell.
- Rostow, Walt W. 1960 *The Stages of Economic Growth: A Non-Communist Manifesto*. London: Cambridge University Press.
- Roxborough, Ian 1988 “Modernization Theory Revisited.” *Contemporary Studies of Society and History* 30(4):753–762.
- Scott, Catherine V. 1995 *Gender and Development: Rethinking Modernization and Dependency Theory*. Boulder, Colo.: L. Rienner.
- Smelser, Neil 1966 “The Modernization of Social Relations.” In Myron Weiner, ed., *Modernization: The Dynamics of Growth*. New York: Basic.
- So, Alvin 1990 *Social Change and Development: Modernization, Dependency and World-Systems Theories*. Newbury Park, Calif.: Sage.
- Tiryakian, Edward A. 1985 “The Changing Centers of Modernity.” In Erik Cohen, Moshe Lissak, and Uri Almagor, eds., *Comparative Social Dynamics: Essay in Honor of S.N. Eisenstadt*. Boulder, Colo.: Westview.

J. MICHAEL ARMER  
JOHN KATSILLIS

## MONEY

Sociologists treat money paradoxically: On the one hand, money is considered a central element of modern society, and yet it remains an unanalyzed sociological category. In classic interpretations of the development of the modern world, money occupies a pivotal place. As “the most abstract and ‘impersonal’ element that exists in human life” (Weber [1946] 1971, p. 331), it was assumed that money spearheaded the process of rationalization. For Georg Simmel and Karl Marx, money revolutionized more than economic exchange: It fundamentally transformed the basis of all social relations by turning personal bonds into calculative instrumental ties.

But by defining money as a purely objective and uniform medium of exchange, classical social theory eclipsed money’s sociological significance.

If indeed money was unconstrained by subjective meanings and independent social relations, there was little left of sociological interest. As a result, economists took over the study of money: There is no systematic sociology of money. Significantly, the *International Encyclopedia of the Social Sciences* devotes over thirty pages to money but not one to its social characteristics. There are essays on the economic effect of money, on quantity theory, on velocity of circulation, and on monetary reform, but nothing on money as a “réalité sociale,” using Simiand’s apt term (1934).

The sociological invisibility of money is hard to pierce. For instance, the current resurgence of interest in economic sociology has led to a serious revamping of the neoclassical economic model of the market, firms, and consumption (see, e.g., Smelser and Swedberg 1994). But despite the stimulus, no full-fledged sociology of money as social process has emerged. Consider the recent literature on the culture of consumption, which boldly reverses our understanding of modern commodities. The new revisionist approach uncovers the symbolic meanings of what money buys, but, curiously, the cultural “freedom” of money itself is seldom directly challenged (see, e.g., Appadurai 1986; Bronner 1989; Brewer and Porter 1993).

A sociology of money must thus dismantle a powerful and stubborn utilitarian paradigm of a single, neutral, and rationalizing market money. It must show that money is a meaningful, socially constructed currency, continually shaped and redefined by different networks of social relations and varying systems of meanings. There is some evidence that the sociological conversion of money has begun. (See, e.g., Doyle 1992; Carruthers and Espeland 1998; Dodd 1994; Lane 1990; Mizuchi and Stearns 1994; Reddy 1987; Singh 1997; Wuthnow 1996; Mongardini 1998; Neary and Taylor 1998; Zelizer 1994, 1996.) And in anthropology, psychology, political science, geography and history there are also scattered indications that the economic model of money is starting to lose its hold. (See, e.g., Berti 1991; Bloch 1994; Cohen 1998; Guyer 1995; Heath and Soll 1996; Helleiner 1998; Kahneman and Tversky 1982; Lane 1990; Parry and Bloch 1989; Leyshon and Thrift 1997; Thaler 1990; Shafir, et al. 1997; Shell 1995.) The following two sections will first discuss the classic approach



to money and then propose the basis for a sociology of money.

### MARKET MONEY: A UTILITARIAN APPROACH TO MONEY

Many eighteenth-century thinkers saw the monetization of the economy as compatible with or even complementary to the maintenance of a morally coherent social life (see Hirschman 1977; Silver 1990). But the transformative powers of money captured the imagination of nineteenth- and early twentieth-century social theorists. Money turned the world, observed Simmel ([1908] 1950, p. 412), into an “arithmetic problem.” On purely technical grounds, the possibility of money accounting was essential for the development of impersonal rational economic markets. But traditional social thinkers argued that the effects of money transcended the market: More significantly, money became the catalyst for the generalized instrumentalism of modern social life. As Simmel ([1900] 1978, p. 346) observed: “The complete heartlessness of money is reflected in our social culture, which is itself determined by money.”

The task of social theory was thus to explain the uncontested revolutionary power of money. Presumably, it came from money’s complete indifference to values. Money was perceived as the prototype of an instrumental, calculating approach; in Simmel’s ([1900] 1978, p. 211) words, money was “the purest reification of means.” Unlike any other known product, money was the absolute negation of quality. With money, only quantity mattered. That “uncompromising objectivity” allowed it to function as a “technically perfect” medium of modern economic exchange. Free from subjective restrictions, indifferent to “particular interests, origins, or relations,” money’s liquidity and divisibility were infinite, making it “absolutely interchangeable” (pp. 373, 128, 441). Noneconomic restrictions in the use of money were unequivocally dismissed as residual atavisms. As money became nothing but “mere money,” its freedom was apparently unassailable and its uses unlimited. With money, all qualitative distinctions between goods were equally convertible into an arithmetically calculable “system of numbers” (p. 444).

This objectification of modern life had a dual effect. On the one hand, Simmel argued that a

money economy broke the personal bondage of traditional arrangements by allowing every individual the freedom of selecting the terms and partners of economic exchange. But the quantifying alchemy of money had a more ominous chemistry. In an early essay, Marx ([1844] 1964, p. 169) had warned that the transformational powers of money subverted reality: “Confounding and compounding . . . all natural and human qualities . . . [money] serves to exchange every property for every other, even contradictory, property and object: it is the fraternization of impossibilities.” As the ultimate objectifier, money not only obliterated all subjective connections between objects and individuals but also reduced personal relations to the “cash nexus.” Half a century later, Simmel ([1908] 1950, p. 414) confirmed Marx’s diagnosis, dubbing money a “frightful leveler” that perverted the uniqueness of personal and social values. And Max Weber ([1946] 1971, p. 331) pointed to the fundamental antagonism between a rational money economy and a “religious ethic of brotherliness.”

The prevailing classic interpretation of money thus absolutized a model of market money, shaped by the following five assumptions:

1. The functions and characteristics of money are defined strictly in economic terms. As a qualityless, absolutely homogeneous, infinitely divisible, liquid object, money is a matchless tool for market exchange.
2. All monies are the same in modern society. Differences can exist in the quantity of money but not in its meaning. Thus, there is only one kind of money—market money.
3. A sharp dichotomy is established between money and nonpecuniary values. Money in modern society is defined as essentially profane and utilitarian in contrast to noninstrumental values. Money is qualitatively neutral; personal, social, and sacred values are qualitatively distinct, unexchangeable, and indivisible.
4. Monetary concerns are seen as constantly enlarging, quantifying, and often corrupting all areas of life. As an abstract medium of exchange, money has not only the freedom but also the power to draw an

increasing number of goods and services into the web of the market. Money is thus the vehicle for an inevitable commodification of society.

5. The power of money to transform nonpecuniary values is unquestioned, while the reciprocal transformation of money by values or social relations is seldom conceptualized or else is explicitly rejected.

As the classic view reasons, the monetization of the economy made a significant difference to the organization of social life. For example, it facilitated the multiplication of economic partners and promoted a rational division of labor. But a link is missing from the traditional approach to money. Impressed by the fungible, impersonal characteristics of money, classic theorists emphasized its instrumental rationality and apparently unlimited capacity to transform products, relationships, and sometimes even emotions into an abstract and objective numerical equivalent. But money is neither culturally neutral nor socially anonymous. It may well “corrupt” values and social ties into numbers, but values and social relations reciprocally corrupt money by investing it with meaning and social patterns.

### TOWARD A SOCIOLOGY OF MONEY

The utilitarian model has had a remarkable grip on theorizing about money. Coleman (1990, pp. 119–131), for example, builds an extremely sophisticated analysis of social exchange yet continues to treat money as the ultimate impersonal common denominator. Even when analysts recognize the symbolic dimension of modern money, they stop short of fully transcending the utilitarian framework. Parsons (1971a, p. 241; 1971b, pp. 26–27), for instance, explicitly and forcefully called for a “sociology of money” that would treat money as one of the various generalized symbolic media of social interchange, along with political power, influence, and value commitments. In contrast to Marx’s definition of money as the “material representative of wealth” ([1858–1859] 1973, p. 222), in Parsons’s media theory, money was a shared symbolic language; not a commodity, but a signifier, devoid of use-value. Yet Parsons restricts the symbolism of money to the economic sphere. Money,

Parsons (1967, p. 358) contends, is the “symbolic ‘embodiment’ of economic value, of what economists in a technical sense call ‘utility.’” Consequently, Parsons’s media theory left uncharted the symbolic meaning of money outside the market: money’s cultural and social significance beyond utility. Giddens (1990) complains that Parsons incorrectly equates power, language, and money, whereas for Giddens money has a distinctly different relationship to social life. As a “symbolic token,” money, in Giddens’s analysis, serves as a key example of the “disembedding mechanisms associated with modernity,” by which he means the “‘lifting out’ of social relations from local contexts of interaction and their restructuring across indefinite spans of time-space” (1990, pp. 22, 25, 21). Giddens’s interpretation still ignores the fact that despite the transferability of money, people make every effort to embed it in particular times, places, meanings, and social relations.

Anthropologists provide some intriguing insights into the extraeconomic, symbolic meaning of money, but mostly with regards to primitive money. For instance, ethnographic studies show that in certain primitive communities, money attains special qualities and distinct values independent of quantity. How much money is less important than *which* money. Multiple currencies, or “special-purpose” money, using Polanyi’s term (1957, pp. 264–266), have sometimes coexisted in one and the same village, each currency having a specified, restricted use (for purchasing only certain goods or services), special modes of allocation and forms of exchange (see, e.g., Bohannan 1959), and, sometimes, designated users.

These special moneys, which Douglas (1967) has perceptively identified as a sort of primitive coupon system, control exchange by rationing and restricting the use and allocation of currency. In the process, money sometimes performs economic functions serving as media of exchange, but it also functions as a social and sacred “marker,” used to acquire or amend status, or to celebrate ritual events. The point is that primitive money is transformable, from fungible to nonfungible, from profane to sacred.

But what about modern money? Has modernization indeed stripped money of its cultural meaning? Influenced by economic models, most interpretations establish a sharp dichotomy between

primitive, restricted “special-purpose” money and modern “all-purpose” money, which, as a single currency, unburdened by ritual or social controls, can function effectively as a universal medium of exchange. Curiously, when it comes to modern money, even anthropologists seem to surrender their formidable analytical tools. For instance, twenty years ago, Douglas (1967), in an important essay, suggested that modern money may not be unrestricted and “free” after all. Her evidence, however, is puzzlingly limited. Modern money, argues Douglas (p. 139), is controlled and rationed in two situations: in international exchange and at the purely individual personal level, where “many of us try to primitivize our money . . . by placing restrictions at source, by earmarking monetary instruments of certain kinds for certain purposes.”

Modern money, however, is marked by more than individual whim or by the different material form of currencies. As François Simiand, one of Durkheim’s students, argued, the extraeconomic, social basis of money remains as powerful in modern economic systems as it was in primitive and ancient societies (1934). Indeed, Simiand warned against an orthodox rationalist approach that mistakenly ignores the persistent symbolic, sacred, and even magical significance of modern money. In recent work, sociologists, as well as anthropologists, psychologists, historians, and political scientists, have finally heeded the warning, proposing long-overdue alternatives to the standard utilitarian model of money.

Impatient with their former theoretical blinders, some anthropologists are now claiming modern money for their disciplinary terrain, casting off the fallacy of a single, culturally neutral currency. Parry and Bloch’s important collection of essays (1989) demonstrates the heterogeneity of money, showing how the multiple symbolic meanings of modern money are shaped by the cultural matrix. In psychology, new studies reject the notion that money is psychologically general, maintaining that instead money involves “multiple symbolizations” (Lea et al. 1987, p. 335). An exciting literature on “mental accounting” challenges the economists’ assumption of fungibility by showing the ways individuals distinguish between kinds of money. For instance, they treat a windfall income much differently from a bonus or an inheritance, even when the sums involved are identical.

A sociological accounting of money goes even further. Anthropologists reveal the multiple symbolic representations of modern money in societies outside the centers of capitalism, and psychologists explore individual or household-based differentiations between monies. A sociological model, on the other hand, must show how, even in the most advanced capitalist societies, different networks of social relations and meaning systems mark modern money, introducing controls, restrictions, and distinctions that are as influential as the rationing of primitive money. Special money in the modern world may not be as visibly identifiable as the shells, coins, brass rods, or stones of primitive communities, but its invisible boundaries emerge from sets of historically varying formal and informal rules that regulate its uses, allocation, sources, and quantity. How else, for instance, do we distinguish a bribe from a tribute or a donation, a wage from an honorarium, or an allowance from a salary? How do we identify ransom, bonuses, tips, damages, or premiums? True, there are quantitative differences between these various payments. But surely, the special vocabulary conveys much more than diverse amounts. Detached from its qualitative differences, the world of money becomes undecipherable.

The sociological model of money thus challenges the traditional utilitarian model of market money by introducing different fundamental assumptions in the understanding of money:

1. While money does serve as a key rational tool of the modern economic market, it also exists outside the sphere of the market and is profoundly shaped by different networks of social relations and varying systems of meaning.
2. Money is not a single phenomenon. There is a plurality of different kinds of monies; each special money is shaped by a particular set of cultural and social factors and is thus qualitatively distinct. Market money does not escape extraeconomic influences but is in fact one type of special money, subject to particular social and cultural influences.
3. The classic economic inventory of money’s functions and attributes, based on the assumption of a single general-purpose

type of money, is thus unsuitably narrow. By focusing exclusively on money as a market phenomenon, it fails to capture the complex range of characteristics of money as a social medium. A different, more inclusive coding is necessary, for certain monies can be indivisible (or divisible but not in mathematically predictable portions), nonfungible, nonportable, deeply subjective, and therefore qualitatively heterogeneous.

4. The assumed dichotomy between a utilitarian money and nonpecuniary values is false, for money under certain circumstances may be as singular and unexchangeable as the most personal or unique object.
5. Given the assumptions above, the alleged freedom and unchecked power of money become untenable assumptions. Culture and social structure set inevitable limits to the monetization process by introducing profound controls and restrictions on the flow and liquidity of money. Extraeconomic factors systematically constrain and shape (a) the uses of money, earmarking, for instance, certain monies for specified uses; (b) the users of money, designating different people to handle specified monies; (c) the allocation system of each particular money; (d) the control of different monies; and (e) the sources of money, linking different sources to specified uses.

Exploring the quality of multiple monies does not deny money's quantifiable and instrumental characteristics but moves beyond them; it suggests very different theoretical and empirical questions from those derived from a purely economic model of market money. In fact, a utilitarian theory of money had a straightforward task: explaining how money homogenized and commoditized modern social life. Its critics have a much more complex empirical agenda. The illusion of a fully commoditized world must be rectified by showing how different social relations and systems of meanings actively create and shape a plurality of qualitatively distinct kinds of money. Specifically, a sociological theory of money must come to grips with the remarkably

different ways in which people identify, classify, interpret, organize, and use money.

Consider for instance the family economy. Domestic money—which includes wife's money, husband's money, and children's money—is a special category of money. Its meanings, uses, allocation, and even quantity are partly determined by considerations of economic efficiency, but domestic money is equally shaped by ideas about family life, by power relationships, age, gender, and social class (Zelizer 1994; Pahl 1989; Singh 1997). For instance, a wife's pin money—regardless of the amount involved—was traditionally reserved for special purchases such as clothing or vacations and kept apart from the “real” money earned by her husband. Or consider the case of gift money. When money circulates among friends or kin as a personal gift for ritual events such as weddings, christenings, bar mitzvahs, or Christmas, it is reshaped into a sentimental currency expressing care and affection. It matters who gives it, when it is given, how it is presented, and how spent. Within formal institutions, money is again redefined this time partly by bureaucratic legislation (Goffman 1961).

These cases are not anomalies or exceptions to valuefree market money but typical examples of money's heterogeneity in modern society. In fact, money used for rational instrumental exchanges is simply another socially created currency, not free from social constraints, but subject to particular networks of social relations and its own set of values and norms. A sociological theory of money must explain the sources and patterns of variation between multiple monies. How, for instance, do personal monies, such as domestic and gift monies, which emerge from the social interaction of intimates, differ from the imposed institutional money of inmates? How does the social status of transactors affect the circulation of monies? What determines the relative rigidity or permeability of boundaries between monies? And what are the patterns of conversions between them?

Developing a sociological model of multiple monies forms part of a broader challenge to neoclassical economic theory. It offers an alternative approach not only to the study of money but to all other aspects of economic life, including the market. In the long run, a proper sociological

understanding of multiple monies should challenge and renew explanation of large-scale economic change and variation. It should illuminate such phenomena as aggregate expenditures on consumer durables, rates of saving, response to inflation, income redistribution, and a wide range of other phenomena in which individual consumer actions make a large macroeconomic difference. In the sociological model, economic processes of exchange and consumption are defined as one special category of social relations, much like kinship or religion. Thus, economic phenomena such as money, although partly autonomous, intertwine with historically variable systems of meanings and structures of social relations.

SEE ALSO: *Economic Sociology*

## REFERENCES

- Appadurai, Arjun, ed. 1986 *The Social Life of Things*. Cambridge, England: Cambridge University Press.
- Berti, L., et al. 1991 "Il Denaro." *Problemi Del Socialismo* 7-8.
- Bloch, Maurice 1974 "Les usages de l'argent (Money's Uses)." In Maurice Bloch, ed., "Les Usages De L' Argent." *Terrain* 23: 5-10.
- Bohannon, Paul 1959 "The Impact of Money on an African Subsistence Economy." *Journal of Economic History* 19:491-503.
- Brewer, John, and Roy Porter, eds. 1993 *Consumption and the World of Goods*. New York: Routledge.
- Bronner, Simon J. ed. 1989 *Consuming Visions*. New York: Norton.
- Carruthers, Bruce, and Wendy Nelson Espeland 1998 "Money, Meaning, and Morality." *American Behavioral Scientist* 41 (August): 1384-1408.
- Cohen, Benjamin J. 1998 *The Geography of Money*. Ithaca, N.Y.: Cornell University Press.
- Coleman, James 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Dodd, Nigel 1994 *The Sociology of Money*. New York: Continuum.
- Douglas, Mary 1967 "Primitive Rationing." In Raymond Firth, ed., *Themes in Economic Anthropology*. London: Tavistock.
- Doyle, Kenneth O., ed. 1992 "The Meanings of Money." *American Behavioral Scientist* 35:637-840.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Stanford, Calif.: Stanford University Press.
- Goffman, Erving 1961 *Asylums*. New York: Anchor.
- Guyer, Jane I., ed. 1995 *Money Matters*. Portsmouth, N.H.: Heinemann.
- Heath, Chip, and Jack B. Soll 1996 "Mental Budgeting and Consumer Decisions." *Journal of Consumer Research* 23:40-52.
- Helleiner Eric 1998 "National Currencies and National Identities." *American Behavioral Scientist* 41 (August):1409-1436.
- Hirschman, Albert O. 1977 *The Passions and the Interests*. Princeton, N.J.: Princeton University Press.
- Kahneman, Daniel, and Amos Tversky 1982 "The Psychology of Preferences." *Scientific American* 246 (January):160-173.
- Lane, Robert E. 1990 "Money Symbolism and Economic Rationality." Paper presented at the Second Annual Meeting of the Society for the Advancement of Socio-Economics, Washington, D.C., March.
- Lea, Stephen E. G., Roger Tarpy, and Paul Webley 1987 *The Individual in the Economy*. New York: Cambridge University Press.
- Leyshon, Andrew, and Nigel Thrift 1997 *Money Space: Geographies of Monetary Transformation*. London: Routledge.
- Marx, Karl (1844) 1964 "The Power of Money in Bourgeois Society." In *The Economic and Philosophic Manuscripts of 1844*. New York: International Publishers.
- (1858-1859) 1973 *Grundrisse*. New York: Vintage.
- Mizruchi, M. S., and L. Brewster Stearns 1994 "Money, Banking, and Financial Markets." Pp. 313-341 in Neil Smelser and Richard Swedberg, eds., *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press; and New York: Russell Sage Foundation.
- Mongardini, Carlo, ed. 1998 *Il denaro nella cultura moderna (Money in Modern Culture)* Roma: Bulzoni.
- Neary, Michael, and Graham Taylor 1998 *Money and the Human Condition*. New York: St. Martin's.
- Pahl, Jan 1989 *Money and Marriage*. New York: St. Martin's.
- Parry, J., and M. Bloch, eds. 1989 *Money and the Morality of Exchange*. Cambridge, England: Cambridge University Press.
- Parsons, Talcott 1967 "On the Concept of Influence." In *Sociological Theory and Modern Society*. New York: Free Press.
- 1971a "Higher Education as a Theoretical Focus." In Herman Turk and Richard L. Simpson, eds., *Institutions and Social Exchange*. New York: Bobbs-Merrill.

- 1971b “Levels of Organization and the Mediation of Social Interaction.” In Herman Turk and Richard L. Simpson, eds., *Institutions and Social Exchange*. New York: Bobbs-Merrill.
- Polanyi, Karl 1957 “The Economy as an Instituted Process.” In Karl Polanyi, Conrad M. Arensberg, and Harry W. Pearson, eds., *Trade and Market in the Early Empires*. New York: Free Press.
- Reddy, William 1987 *Money and Liberty in Modern Europe*. New York: Cambridge University Press.
- Shafir, Eldar, Peter Diamond, and Amos Tverski 1997 “Money Illusion.” *Quarterly Journal of Economics* CXII(2):341–374.
- Shell, M. 1995 *Art and Money*, Chicago: University of Chicago Press.
- Silver, Allan 1990 “Friendship in Commercial Society: Eighteenth-Century Social Theory and Modern Sociology.” *American Journal of Sociology* 95:1,474–1,504.
- Singh, Supriya 1997 *Marriage Money: The Social Shaping of Money in Marriage and Banking*. St. Leonards, Australia: Allen and Unwin.
- Simiand, François 1934 “La monnaie, réalité sociales [Money: A Social Reality] *Annales sociologiques*, ser. D, pp. 1–86.
- Simmel, Georg (1900) 1978 *The Philosophy of Money*, trans. Tom Bottomore and David Frisby. London: Routledge and Kegan Paul.
- (1908) 1950 *The Sociology of Georg Simmel*, ed. Kurt H. Wolf. New York: Free Press.
- Smelser, Neil J., and Richard Swedberg, eds. 1994 *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press; and New York: Russell Sage Foundation.
- Wuthnow, Robert 1996 *Poor Richard's Principle*. Princeton, N.J.: Princeton University Press.
- Zelizer, Viviana A. 1994 *The Social Meaning of Money*. New York: Basic Books.
- 1996 “Payments and Social Ties.” *Sociological Forum* 11 (September):481–495.

VIVIANA ZELIZER

## MORAL DEVELOPMENT

*Morality* refers to the set of values that people use to determine appropriate behavior, that is, what is right versus what is wrong. Determining which behavior is morally appropriate, or “right,” is

essentially a cognitive decision-making process called *moral judgment*.

Moral judgment is but one component of the process leading to the actual performance of morally appropriate behavior (Rest 1986). However, research on moral development over the past forty-five years has focused primarily on the development of moral judgment. This is due in large part to the influence of psychologists Lawrence Kohlberg (1969, 1971, 1976) and Jean Piaget ([1932] 1948).

Both Piaget and Kohlberg maintained that moral behavior largely depends upon how one perceives the social world and oneself in relation to it. Furthermore, they viewed moral decision making as a rational process and thus linked the development of moral judgment to the development of rational cognition. In this way, moral development is seen largely as changes in one’s *way of thinking* about questions of morality as he or she gets older.

The present discussion will provide a brief historical description of the theoretical foundation laid down by Piaget and Kohlberg; the method used by Kohlberg to assess moral development as well as alternative methods that have emerged more recently; some of the major criticisms and reconceptualizations of Kohlberg’s theory of moral development; and recent research that has pursued those criticisms in the areas of cultural differences, gender differences, and continued adult development.

### THEORETICAL FOUNDATION

Kohlberg built on Piaget’s theory of cognitive development to hypothesize a sequence of six specific stages of moral judgment in individual development. This theory of moral development is based on a fundamental idea from Piaget that the way people think about the physical and social world is the result of an “interactional” process between the human organism’s innate tendencies and influences from the environment.

This “cognitive-developmental” approach is thus distinguished from both maturational and environmental theories of development. Maturational theories (Gesell 1956) maintain that patterns of behavior express the organism’s inherent

tendencies. Development is seen as the natural unfolding of a process determined by hereditary factors. In contrast, environmental theories argue that behavior is determined primarily by external influences. From this point of view, behavior is not innately patterned but is essentially learned, whether as a result of conditioning processes that associate the behavior with particular stimuli, rewards, and punishment, or as a result of observing (and subsequently modeling) the behavior of others.

Social learning theory (Bandura 1977) has produced considerable research on how observational learning explains a variety of behaviors relevant to morality, including prosocial behavior (e.g., sharing, cooperation), aggression, resistance to temptation, and delayed gratification. More recent developments have pursued the question of how individuals exert control over their behavior, thus providing some balance to the theory's focus on environmental influences. Bandura's self-efficacy theory (1982), for example, emphasizes the individual's expectations as important to the successful performance of a behavior. However, social learning theory has not addressed moral action and moral character in terms of a broad developmental course (Musser and Leone 1986).

Cognitive-developmental theory, on the other hand, focuses on the developmental process by which people come to understand and organize, or "cognitively structure," their experience. It attempts to resolve the "nature-nurture" controversy by emphasizing the development of these cognitive structures as the result of the interaction between organismic tendencies and influences from the outside world. While particular ways of understanding experience may reflect innate tendencies, they develop in response to the individual's specific experiences with the environment.

Thus, development is not seen as primarily maturational, because experience is necessary for cognitive structure to take shape. However, neither is development thought to be primarily determined by the environment. Rather, cognitive developmentalists argue that, because the underlying thought organization at each stage is qualitatively different, cognitive development is more than the progressively greater acquisition of information. Furthermore, at any given stage, the current cognitive structure can influence how the world is perceived. Thus, cognitive structure is

seen to be "the result of an interaction between certain organismic structuring tendencies and the structure of the outside world, rather than reflecting either one directly" (Kohlberg 1969, p. 352).

### THE DOCTRINE OF COGNITIVE STAGES

Piaget's theory of cognitive development maintains that cognitive structures are periodically transformed or restructured as they become unable to account for (or assimilate) new information from the external world adequately. These periods of restructuring result in new ways of understanding that are different from the earlier mental structures as well as from those to be developed later. This allows for the differentiation of distinct cognitive stages each identifiable by a characteristic approach to processing and organizing one's experience of external reality.

Piaget (1960) identified four main characteristics of cognitive stages. Kohlberg (1969) maintains that these characteristics accurately describe his stages of moral development. The characteristics identified by Piaget are as follows:

1. Stages refer to distinct qualitative differences in the way a person thinks about an experience or solves a problem. Although the focus of attention may be the same, the mode of thinking about it is very different.
2. The progression of stages follows an invariant sequence in the development of individuals. That is, the order in which the stages occur is universal for all human beings. It is possible that the speed or timing at which one progresses through the stages may vary with individual or cultural environments—or even that development may stop at one point or another. However, a given stage cannot be followed by any other stage than the one that is next in the sequence. Conversely, the earlier stage must first be achieved before its inadequacies become apparent and the subsequent transformation to the next stage can occur.
3. The characteristic mode of thinking represents a structured whole. Specific cognitive responses to specific tasks depend upon

the overall organizational framework within which one processes information. It is this underlying cognitive structure that produces a logical consistency to one's responses. Thus, the stage is not identified by specific responses to specific stimuli, but it is the pattern in one's responses that indicates a particular underlying cognitive structure.

4. The sequence of stages is hierarchical. At each stage, the underlying structure represents a more integrated and more complex organizational system, one that adequately accounts for information that had created discrepancies within the previous structure. For example, children in the preoperational stage of cognitive development (Piaget's second stage) cannot understand that equal-sized balls of clay formed into two different shapes still have equal amounts of clay. However, children who have achieved concrete operational thinking (Piaget's third stage) understand the principle of conservation and thus recognize that the amount of clay remains the same (is conserved) for both pieces, even though the pieces have changed in shape (Piaget and Inhelder 1969). The underlying cognitive structure of concrete operational thinking differentiates between amount and shape and integrates the information to achieve a more complex understanding of the phenomenon. It is thus logically superior to preoperational thinking. That the later stages in cognitive development are also more comprehensive and more advanced introduces a hierarchical element to the sequence. The stages of cognitive development are not just different but also hierarchical, in the sense that they provide a progressively more differentiated and more integrated—and hence more adaptive—understanding of one's interaction with the environment.

#### KOHLBERG'S STAGES OF MORAL DEVELOPMENT

Kohlberg's six stages of moral reasoning are divided into three levels, each consisting of two stages. The three levels are differentiated according to

what serves as the basis for the person's moral judgment, specifically the significance given the prevailing, or "conventional," social expectations and authority. Briefly, the *preconventional* level, which is the level of most children under 9 years old, occurs prior to the individual's achievement of a full understanding of what is expected or required socially. The *conventional* level, which characterizes most adolescents and adults, refers to an understanding of the social conventions and a belief in conforming to and maintaining the established social order. The *postconventional* level is reached only by a minority of adults, who understand and generally accept the social rules and expectations but recognize that these have been established for the larger purpose of serving universal moral principles. Should the social conventions conflict with these principles, then moral judgment at this level will support the principles at the expense of the conventions.

Within each level, the second stage is a more advanced form than the first. More specifically, the preconventional level refers to judgment based not so much on a sense of what is right and wrong as on the physical consequences that any given act will have for the self. Accordingly, at the first stage within this level, characterized by the *punishment and obedience* orientation, the child will make judgments on the basis of avoiding trouble. This includes obeying authorities to avoid punishment.

At Stage 2, still in the preconventional level, the individual has a sense of the needs of others but still makes judgments to serve her or his own practical interests. This is called the *instrumental* orientation. Although the person is beginning to understand that social interaction involves reciprocity and exchange among participants, moral judgment is still determined by the significance that the action has for oneself. Thus a child may share candy to get some ice cream.

Next, in the conventional level, moral judgment is determined by what is considered "good" according to conventional standards. At this level, the individual has an understanding of what kind of behavior is expected. The first stage at this level (Stage 3) is characterized by the *good boy-good girl* orientation. Judgment as to what is right is based on living up to the expectations of others. It involves a trust in established authority and conformity for the sake of approval.



At Stage 4, the orientation is toward doing one's duty. This is called the *law-and-order* orientation. The individual personally subscribes to the existing social order and thus believes that obeying authority and maintaining the social order are good values in their own right. Whereas behaving according to the social conventions is desirable at Stage 3 because it produces approval from others, at Stage 4 the individual has successfully "internalized" these conventions, so that proper behavior is rewarding because it reinforces one's sense of doing one's duty and therefore produces self-approval.

At the postconventional level, one's understanding of what is right and wrong is based on one's personal values and a sense of shared rights and responsibilities. Morality is no longer determined simply by social definition, but rather by rational considerations. Stage 5 is characterized by the *social contract* orientation, which recognizes that conventions are determined by social consensus and serve a social function. There is an emphasis on utilitarian agreements about what will serve the most good for the most people. Here the person recognizes that rules or expectations are essentially arbitrary. The focus on agreement or contract produces an emphasis on what is legal and on operating "within the system" to achieve one's goals.

Stage 6, however, places the responsibility of a given moral decision firmly on the shoulders of the individual. The basis for moral judgment is found in *universal ethical principles* rather than socially established rules or expectations. One is guided by one's own conscience and recognizes the logical superiority of principles such as respect for human dignity. At Stage 6, it is thus possible to adopt a position that is in conflict with the prevailing social order, and to maintain this position as morally correct.

In Kohlberg's last theoretical paper, he and his colleagues attempt to articulate Stage 6 more completely (Kohlberg et al. 1990). They describe it as fundamentally characterized by a "respect for persons," specifically one that successfully integrates a sense of justice that is universal and impartial with an attitude of "benevolence" that is empathic and understanding of the individual (see also Lapsley 1996).

## MEASURING MORAL JUDGMENT

Kohlberg's procedure for assessing moral judgment involves presenting a hypothetical "dilemma" that requires the subject to make a moral choice. The most famous example refers to "Heinz," a man whose wife is dying of cancer. The woman could possibly be saved by a new drug, but the druggist who discovered it is charging an exorbitant amount of money for it, ten times what it costs him to make it. Heinz tried but could not raise enough money, so he steals the drug. Should he have done this?

Because Kohlberg's scheme emphasizes cognitive structure, an individual's stage of moral development is indicated not by the actual behavior that is advocated but rather by the pattern of reasoning behind the decision. Thus, two people may arrive at the same decision (e.g., that Heinz should steal the drug to save the life of his dying wife) but for two entirely different reasons. An individual at the preconventional Stage 2, operating within the instrumental orientation, might recommend stealing the drug because any jail term would be short and worth saving his wife. An individual at the postconventional Stage 6 might also recommend stealing the drug but with a different understanding of the dilemma: Although stealing would violate the law, it would uphold the higher principle of valuing human life and allow Heinz to maintain his self-respect.

The difference between the actual behavioral content of a decision and the cognitive structure of the decision is also illustrated when two people arrive at different decisions but for similar reasons. Thus, the decision not to steal the drug because Heinz would go to jail and probably not be released until after his wife died is also Stage 2 thinking. Even though the ultimate decision advocates the opposite behavior of what was indicated above, it is similarly based on the consideration of what would be most instrumental to Heinz's own self-interest. On the other hand, an individual at Stage 6 might recommend not stealing the drug because, although other people would not blame Heinz if he stole it, he would nonetheless violate his own standard of honesty and lose his self-respect.

Because the stage of moral development is demonstrated not by the behavioral content but

by the form of the moral judgment, the subject is allowed to respond freely to these moral dilemmas, and is asked to explain and justify his or her answer. The interviewer can probe with specific questions to elicit more information about the basis of the subject's decision. Interviewers are trained to collect relevant information without directing the subject's responses.

The subject's answers are then transcribed and coded for stage of moral development. Kohlberg identified twenty-five aspects of moral judgment, basic moral concepts that refer to such matters as rules, conscience, one's own welfare, the welfare of others, duty, punishment, reciprocity, and motives. Each of the twenty-five aspects was defined differently for each of the six stages of moral development. Originally, Kohlberg used an aspect-scoring system, whereby every statement made by the subject was coded for aspect and rated as to stage ("sentence scoring"). The subject's usage of any given stage of moral reasoning was indicated by the percentage of his or her statements that was attributed to that stage. Aspect scoring also included an overall "story rating," whereby a single stage was assigned to the subject's total response.

Coding difficulties led to the abandonment of the aspect-scoring system. Because the unit of analysis for sentence scoring was so small, coding often became dependent upon the specific content and choice of words and did not lend itself to identifying the general cognitive structure underlying the statement. Conversely, whereas story rating referred to the total response as the unit of analysis, it created some uncertainty when the subject's answer included conflicting themes.

Kohlberg and his colleagues recognized these scoring difficulties and devoted considerable attention to developing a more reliable and valid scoring system. This led to "standardized issue scoring," which relies on the use of a standardized interview format. The subject is presented with three standard dilemmas, and the interviewer probes for only two issues that are specified for each dilemma (e.g., life and punishment in the Heinz dilemma). Scoring of the subject's responses refers to a manual that describes the patterns of reasoning for Stages 1–5 on each issue (Colby et al. 1987). Stage 6 was dropped from the coding procedure, due to its empirically low incidence, but

was retained as a theoretical construct (Kohlberg et al. 1990).

Because the focus of the new scoring system is directed more toward the abstract mode of reasoning, the unit of analysis is considered larger and less concrete than the single sentence. However, because this approach focuses on specifically identified issues, norms, and elements, it is considered more precise than the global story rating. Despite the qualitative nature of this approach and its potential vulnerability to rater bias, its developers report that long-term study of its inter-rater reliability, test-retest reliability, internal consistency, and validity has produced favorable results (Colby and Kohlberg 1987).

Validity has been a major concern regarding Kohlberg's moral judgment interview. Kurtines and Grief (1974) criticized the low utility of moral judgment scores for predicting moral action. Other questions have been raised about the validity of the data collected, even for the purposes of assessing moral judgment. For one, use of the "classical" dilemmas in this research has been criticized on grounds that they are not representative: Not only do they address hypothetical—as opposed to real-life—circumstances, but they refer to a limited domain of moral issues (e.g., property and punishment). Assessment may fail to indicate the extent to which the person's moral judgment is influenced by the particular context provided by the dilemma. A related matter is whether responses are affected by the characteristics (e.g., the gender) of the story's protagonist. Also the effect of differences in interviewing style, as interviewers interact with subjects and probe for further information, needs to be considered. Of particular importance is this method's dependence on the subject's verbal expression and articulation skills for the information that is collected. To the extent that the rating might be affected by either the amount of information that is provided or the manner in which it is expressed, the validity of the scoring system is called into question. (See Modgil and Modgil 1986 for discussion of these issues.)

An alternative to Kohlberg's Moral Judgment Interview is the Defining Issues Test (DIT) (Rest 1986). This is a standardized questionnaire that presents a set of six moral dilemmas and, for each dilemma, specifically identifies twelve issues that

could be considered in deciding upon a course of action. The subject's task is to indicate, on a five-point scale, how important each issue is in deciding what ought to be done in the given situation. The subject also ranks the four most important issues.

Here, the term "issue" is used differently than it is in Kohlberg's new scoring procedure. The items are prototypical statements designed to represent considerations (e.g., "whether a community's laws are going to be upheld") that are characteristic of specific stages of moral reasoning as they are described in Kohlberg's theory. The importance assigned by the subject to items that represent a particular stage is taken to indicate the extent to which the subject's moral judgment is characterized by that stage's mode of thinking.

There are advantages and disadvantages to the DIT compared with the open-ended interview. Whereas the interview is helpful for originally identifying the considerations that may be relevant to resolving moral dilemmas, the DIT provides a more systematic assessment of the relative importance of such considerations. In the open-ended interview, it is never clear whether a specific concern is not identified because it is not important or because the subject failed to articulate it. Similarly, interviews are less comparable to the extent that subjects do not all address the same issues. These problems are avoided by the more structured DIT, because the task requires the subject only to recognize what is important rather than to identify and articulate it spontaneously. However, because recognition is an easier task than spontaneous production, it tends to allow higher-level responses. Another important difference is that the DIT measures the maturity of moral judgment as a continuous variable rather than in terms of the holistic step-by-step sequence of cognitive-developmental stages. Researchers must be aware of such differences when interpreting results.

A third instrument, the Moral Judgment Test (MJT) (Lind et al. 1981; Lind and Wakenhut 1985) similarly attempts to measure moral reasoning by the subject's endorsement of specific items. Hypothetical moral dilemmas are presented, and subjects respond to a series of twelve statements for each dilemma. Each of Kohlberg's six stages is

represented by two statements, one in favor of and one against the particular action in question. Subjects indicate how acceptable they find each of the statements.

Citing Lind's paper (1995) extensively, Rest and colleagues (1997) focus on an important distinction between the MJT and the DIT. Instead of adding ratings to indicate how much the subject prefers a particular stage's statements (*stage preference*)—as most DIT studies do—Lind emphasizes how consistently the subject responds to different statements from the same stage (*stage consistency*). Lind argues that stage consistency is a more accurate measure of true cognitive structure, whereas stage preference is more indicative of an affective (like versus dislike) response.

Rest and colleagues (1997) use the DIT statements to construct a consistency measure that is similar to the one developed by Lind. They conclude that the stage-preference measure shows greater construct validity than the stage-consistency measures in differentiating groups with different expertise and different education. The stage-preference measure correlates more highly with moral comprehension—indicating longitudinal development—predicts both prosocial and antisocial behavior, and correlates with political attitudes.

The question of scoring for preference or consistency and what construct is measured by each approach is a legitimate methodological concern with important implications for our understanding of moral development. However, both the DIT and the MJT can be scored for preference and for consistency. Thus, they each remain a viable alternative for attempting to empirically measure moral judgment.

Another measurement tool is the Sociomoral Reflection Measure (SRM) (Gibbs and Widaman 1982; Gibbs et al. 1982), and a more recent variation is the Short Form (SRM-SF) (Gibbs et al. 1992; Basinger et al. 1995; Communion and Gielen 1995; Garmon et al. 1996). This is an open-ended, group-administrable instrument that asks subjects to rate the importance of such topics as keeping promises, affiliation, life, property, and law. It does not present specific dilemmas, but instead uses "lead-in statements" that instruct subjects to generate their own example, such as "Think about when

you've made a promise to a friend," prior to providing their rating of importance. The short form consists of eleven items that can produce a score ranging from 100 (exclusively Stage 1) to 400 (exclusively Stage 4).

Proponents argue that not only is the SRM-SF suitable for assessing stages of moral judgment, but because examples are self-generated, items also can be used to assess differences in content emphasis (Basinger et al. 1995; Garmon et al. 1996). As such, it is suggested as especially useful for research on cultural differences, gender differences, everyday life (versus hypothetical) experience, and the relationship between moral judgment and moral behavior (Communian and Gielen 1995).

As discussed below, perhaps the single most influential criticism of Kohlberg's theory is Carol Gilligan's contention that it fails to describe the moral development of females (1982). Her articulation of a more female-oriented "morality of care," complete with its own sequence of stages, has led to the development of the Ethics of Care Interview (ECI) (Skoe and Marcia 1991). Similar to Kohlberg's methodology, the ECI assesses stage differences, but specifically as they are relevant to the development of care-based morality. Research with the ECI has recently been reviewed by Skoe (1998), demonstrating that the morality of care has important application to human development in general and to the development of personality in particular.

### CRITICISMS AND FURTHER RESEARCH

Besides the methodological problems discussed above, Kohlberg's theory of moral development has been criticized on a number of points. The major criticisms include the following:

1. The sequence of stages is more representative of Western culture and thus not universal or invariant across all cultures. Moreover, it is culturally biased in that it maintains the ideals of Western liberalism as the highest form of moral reasoning.
2. Like many theories of personality development, Kohlberg's theory fails to describe the development of women accurately but provides a much better understanding of

male development. This is a specific variation of the first criticism, suggesting that the theory itself reflects the sexism of Western culture.

3. Kohlberg's theory fails to describe adult development adequately. In particular, its emphasis on abstract principles fails to recognize how adult moral judgment is more responsive to the specific practical matters of everyday, real-life contexts. Also, its emphasis on cognitive structure fails to recognize that changes in the content of moral reflection may be the most important aspect of adult moral development.

**Cultural Bias.** A cornerstone of cognitive-developmental theory is invariant sequence, the notion that the given developmental progression is universal for all human beings within all cultures. Because the conceptual organization of any given stage is considered logically necessary before the cognitive structure of the next stage can develop, each stage is said to have logical priority to subsequent stages. Shweder and LeVine (1975) take issue with both the notion of logical priority and the doctrine of invariant sequence, although they do not address the development of moral judgment per se. Specifically, they analyze dream concepts among children from the Hausa culture in Nigeria and conclude that there are multiple sequences by which such concepts develop.

Shweder (1982) follows up this initial skepticism with a fuller critique of what he sees as Kohlberg's failure to recognize cultural conceptions of morality as relative to one another. He disagrees with the assertion that there is a rational basis upon which morality can be constructed objectively. Rather, he argues that the postconventional morality that Kohlberg maintains as rationally superior is simply an example of American ideology.

Similarly, others (Broughton 1986; Simpson 1974; Sullivan 1977) argue that Kohlberg's theory is necessarily culture-bound, reflective of the Western society from which it originates. Simpson suggests that the specific moral dilemmas used in the testing situation may not have the same meaning for people of different cultures and thus the scoring system may not adequately detect legitimate

cultural variations in moral structures. Thus, she maintains that the claims to universality are not valid. Sullivan goes even further, suggesting that Stage 6 reasoning is so rooted in the philosophical rationale for current Western society that it serves to defend the status quo. In doing so, it distracts attention from the injustices of such societies.

In an early response to the charge of cultural bias, Kohlberg and colleagues (1983) acknowledge the influence of Western liberal ideology on the theory. They agree there is a need to be more sensitive to cultural differences in the meaning attributed not only to the various elements of the research protocol but, consequently, also to the responses of the subjects themselves. However, they defend the claim to universality for the six-stage sequence of moral development and maintain that empirical research using the scientific method will help to determine to what extent this position is tenable.

They also maintain that, while it is appropriate to remain impartial in the study of moral judgment, this does not make it necessary to deny the relative value of certain moral positions. They assert that some positions are rationally superior to others. They thus continue to subscribe to the ideal that any given moral conflict can be brought to resolution through rational discourse.

Kohlberg's position on invariant sequence has been supported by a number of cross-cultural studies, although postconventional reasoning (Stages 5 and 6) may occur less frequently in nonurbanized cultures (Snarey 1985). However, in a sample of subjects from India, Vasudev and Hummel (1987) not only found stage of moral development to be significantly related to age, but also found postconventional thinking to occur among a substantial proportion of adults. Concluding that commonalities exist across cultures, Vasudev and Hummel also suggest there is cultural diversity in the way moral principles are expressed, interpreted, and adapted to real life.

More recent research has increasingly acknowledged the significance of cultural influences on moral development. In another study with a sample from India, Moore concludes that "the notion of justice, as defined by Western social scientists, was rarely used as a moral rationale" (1995, p.

286). Rather, results indicated moral judgment to be dependent on one's religious ideology or social status. Similarly, Okonkwo's study of Nigerians (1997) concludes that moral thinking and moral language are culture dependent. While some parallels to Kohlberg's scheme were found, some "well-articulated moral expressions could not be scored" (Okonkwo 1997, p. 117) due to the inability of Kohlberg's instrument to adequately assess certain concepts that served as the basis for subjects' moral judgments.

Ma and Cheung (1996), studying samples of Chinese, English, and American adolescents and young adults, likewise found cultural differences in the way subjects interpreted specific items on the Defining Issues Test. However, after they deleted some of the items used to indicate Kohlberg's Stage 4, their samples demonstrated a consistent hierarchical structure across the three cultures. They thus conclude that, while different cultures may encourage different perceptions of specific moral statements, there is some support for the idea of a fundamentally universal development.

Markoulis and Valanides (1997) similarly addressed the cultural bias controversy in a conciliatory fashion. Comparing students from Greece and Nigeria, they found stage differences between the two cultures, but nonetheless found that the sequence of development was similar. Again, while cultural environment is recognized as a factor, invariant sequence in development is supported.

A related concern that has been receiving more attention from researchers is the question of whether differences in political ideology within a single, larger culture may be inaccurately represented as developmental variation. Specifically, Gross (1996) compared Americans who are pro-life on the abortion issue with others who are pro-choice, and also compared Israelis who disagree on how to handle the West Bank settlement issue. As long as socioeconomic status was similar, he found no difference between the relevant ideological groups and thus concludes that there is no evidence that ideological bias is built into the stages of moral development.

Conversely, Emler and colleagues (1998) argue that differences in moral development as assessed by the DIT more accurately reflect differences in political ideology. Thoma and colleagues

(1999) acknowledge overlap between political thinking and moral judgment, but argue that Emler and colleagues and St. James (1998) provide no evidence to discount the DIT as a valid measure of moral development. Narvaez and colleagues (1999) attempt to resolve the issue with a model of moral judgment and cultural ideology as engaged in parallel development, each influencing the other to produce specific moral thinking.

Whereas recent times have been characterized by an increased sensitivity to cultural diversity and political "correctness," more attention has been drawn to the consideration of possible cultural and political bias in the theory of moral development. While some researchers have identified cultural differences in moral reasoning, this has led to an increased recognition of sociocultural factors in moral development (Eckensberger and Zimba 1997).

Shweder and colleagues (1987), for example, propose the social communication theory, which maintains that the learning of morality depends largely on the transmission of cultural ideology to children, by virtue of the evaluation and judgments that parents and others make. The point, of course, is that morality is socially constructed, not self-constructed (Emler 1998).

However, other researchers continue to maintain that, while we must make specific adjustments to our understanding of moral development, it is not necessary to abandon the general consideration of a single, universal pattern to human moral development. This area of inquiry thus promises to remain a controversial yet productive focus for several years.

**Gender Bias.** Carol Gilligan (1982) argues that the major theories of personality development describe males more accurately than females. She includes Kohlberg's theory in this assessment and points to the prevalence of all-male samples in his early research as a partial explanation. Gilligan contrasts two moral orientations. The first is the morality of justice, which focuses on fairness, rights, and rules for the resolution of disputes. The second is the morality of care, which focuses on relationships, a sensitivity to the needs of others, and a responsibility for others. Gilligan asserts that the orientation toward morality as justice is especially characteristic of males and, conversely, that

morality as care and responsibility is especially relevant to females. To the extent that Piaget, Freud, and Kohlberg each address morality as justice, they accurately represent male moral development but inadequately represent female moral development.

Gilligan argues that women are more likely to rely on the orientation of care to frame personal moral dilemmas. Furthermore, whereas the morality of care focuses on interpersonal relationship, it resembles the Stage 3 emphasis on satisfying the expectations of others. Gilligan believes this resemblance results in a high number of female responses being misrepresented with Stage 3 ratings.

Gilligan thus argues that Kohlberg's theory and scoring system are biased to favor men. However, Walker (1984), after systematically reviewing empirical studies that used Kohlberg's method, concludes that men do not score higher than women, when samples are controlled for education, socioeconomic status, and occupation. Similarly, Thoma (1986) reports that sex differences on the Defining Issues Test actually favor women but that the differences are trivial.

Kohlberg and colleagues (1983) address Gilligan's criticisms and agree that the care orientation is not fully assessed by their measurement but disagree that this leads to a biased downscoring of females. They suggest that care and justice may develop together and that Stage 6 nonetheless represents a mature integration of the care and justice moralities (see also Vasudev 1988).

Walker and colleagues (1987) found that both the care and the justice orientations were used by both males and females. Furthermore, the orientation used was related to the type of dilemma being discussed. If the dilemma was focused on personal relationships, both men and women tended to use the care orientation. If the dilemma was impersonal, both men and women tended to express a justice orientation. This suggests that observed gender differences in moral judgment may be more a reflection of the particular kind of dilemma they choose to discuss. Perhaps females tend to report more relationship-oriented dilemmas, the kind that pull for care-based judgments (Yussen 1977).

Wark and Krebs (1996) show just this pattern: Females did not score lower than males on

Kohlbergian dilemmas; however, females were more likely to report care-based dilemmas when asked to recall and describe moral conflicts from real life; this difference in the type of moral dilemmas accounted for differences in moral orientation from males and females.

Using the Sociomoral Reflection Measure-Short Form (SRM-SF), which does not rely on specific dilemmas provided by the researcher, Garmon and colleagues (1996) found support for gender differences in moral orientation, with females more likely to refer to a morality of care. However, they reject Gilligan's claim of a bias against females. In fact, results (Basinger et al. 1995) have indicated a possible female advantage in early adolescence. As measured by the SRM-SF, moral judgment was found to be higher among young female adolescents than among their male counterparts. No gender difference was found in late adolescence or young adulthood. Communian and Gielen (1995) found similar results in an Italian sample, with early adolescent girls scoring higher than early adolescent boys, but no gender differences in adults. In another study of seventh and eighth graders, Perry and McIntire (1995) found that subjects used a care mode, a justice mode, and a third narrowly concerned "selfish" mode to make moral decisions. The girls were more likely to use both the care and the justice modes, while the boys were more likely to choose the less developed selfish mode. Contrary to a bias against females, this research suggests that, at least in early adolescence, girls are more advanced in their moral development. Silberman and Snarey (1993) relate such a cognitive advantage to the earlier physical maturation of girls.

Consistent with the lack of evidence for a bias against females, Skoe (1995) found that Kohlberg's justice-based moral reasoning was unrelated to sex-role orientation, as measured by the Bem Sex Role Inventory (Bem 1974). However, this research indicates an interesting pattern for care-based moral reasoning. Using the Ethic of Care Interview (ECI), Skoe found that care-based reasoning was higher in women who were more androgynous and who indicated higher levels of ego identity. Skoe concludes that women who relinquish the traditional female gender role are more likely to develop a mature care-based morality than are women who retain this role.

This would seem to be inconsistent with Gilligan's argument (1982) that the morality of care depends on traditional female socialization. While this morality may be rooted in the traditional female role, Skoe's findings suggest that its advanced development may require a more integrated, androgynous identity.

Current researchers seem to be recognizing that the different moralities go beyond simple gender role differences. Woods (1996) argues that both Kohlberg and Gilligan represent polarized, sexist views, limited to a focus on gender differences. She suggests that researchers need to take a more comprehensive view of the multiple biological and cultural variables that impact moral development, without reducing it to a discussion of sexism. Gilligan's morality of care identifies a vital approach to morality that may have its origin and strength in feminine ideals, indeed that may be more salient in females than in males. However, women are not confined to it, nor is it confined to women, especially as gender roles become more relaxed.

**Adult Development.** A third major issue concerning Kohlberg's theory is whether or not it accurately addresses continued adult development. This issue reflects a more general concern in lifespan developmental psychology regarding the inapplicability of Piaget's model for cognitive development beyond adolescence, leading to a consideration of what has come to be called "postformal" development (Commons et al. 1984). Murphy and Gilligan (1980) found that college and postcollege subjects not only indicated a greater tendency to appreciate the importance of specific contexts in real-life dilemmas but also indicated a slight tendency to regress from Stage 5 moral reasoning on the classical dilemmas. They suggest that a more mature recognition of the significance of contextual particulars leads one to question the validity of abstract moral principles (hence the regressed score). This argument is consistent with other work suggesting that adult cognitive development in general is marked by a greater appreciation of the practical realities of day-to-day living (Denney and Palmer 1981; Labouvie-Vief 1984; W. G. Perry, Jr. 1970). Related to this emphasis on the practical is the finding of Przygotzki and Mullet (1997) that elderly adults, when attributing blame,

give more importance to the consequences of an action than to the intention of the perpetrator.

Finally, Gibbs (1979) argues that adult development is characterized more by increased reflection on such existential matters as meaning, identity, and commitment than by any structural change in the way the person thinks. Similarly, Nisan and Applebaum (1995) suggest that older adults give more weight to identity-related personal considerations when considering moral choices, unless they conflict with "a moral demand" from an "unambiguous law." Gibbs (1979) suggests that Kohlberg's postconventional stages are not structural advances over the earlier stages but would be more appropriately described in terms of existential development. In response, Kohlberg et al. (1983) maintain that Stage 5 represents a legitimate cognitive structure. However, they acknowledge the possibility of further nonstructural development in the adult years with regard to both specific contextual relativity and existential reflection. They suggest that such development could be described in terms of "soft" stages that do not strictly satisfy Piaget's formal criteria for cognitive stages.

### SUMMARY

In spite of the formidable criticisms that have been levied against it, Kohlberg's theory of moral development remains the centerpiece to which all other work in this area is addressed, whether as an elaboration or as a refutation. At the very least, Kohlberg has formulated a particular sequence of moral reasoning that adequately represents the prevalent sequence of development in traditional Western society. To that extent, it serves as a model, not only for building educational devices (see Modgil and Modgil 1986; Nucci 1989; Power et al. 1989), but also for comparing possible alternatives. Whether or not this sequence is in fact universal or relative to the particular culture—or a particular socialization process within the culture—is a debate that continues. Nonetheless, the scheme remains the prototype upon which further work in this area is likely to be based.

### REFERENCES

Bandura, Albert 1977 *Social Learning Theory*. Englewood Cliffs, N.J.: Prentice-Hall.

- 1982 "Self-Efficacy Mechanism in Human Agency." *American Psychologist* 37:122-147.
- Basinger, Karen S., John C. Gibbs, and Dick Fuller 1995 "Context and the Measurement of Moral Judgment." *International Journal of Behavioral Development* 18:537-556.
- Bem, Sandra L. 1974 "The Measurement of Psychological Androgyny." *Journal of Consulting and Clinical Psychology* 42:155-162.
- Broughton, John M. 1986 "The Genesis of Moral Domination." In S. Modgil and C. Modgil, eds., *Lawrence Kohlberg: Consensus and Controversy*. Philadelphia: Falmer.
- Colby, Anne, and Lawrence Kohlberg 1987 *The Measurement of Moral Judgment*, vol. 1, *Theoretical Foundations and Research Validation*. New York: Cambridge University Press.
- , Alexandra Hewer, Daniel Candee, John C. Gibbs, and Clark Power 1987 *The Measurement of Moral Judgment*, vol. 2, *Standard Issue Scoring Manual*. New York: Cambridge University Press.
- Commons, Michael L., Francis A. Richards, and Cheryl Armon (eds.) 1984 *Beyond Formal Operations: Late Adolescent and Adult Cognitive Development*. New York: Praeger.
- Communian, Anna Laura, and Uwe P. Gielen 1995 "Moral Reasoning and Prosocial Action in Italian Culture." *Journal of Social Psychology* 135:699-706.
- Denney, Nancy W., and Ann M. Palmer 1981 "Adult Age Differences on Traditional and Practical Problem-Solving Measures." *Journal of Gerontology* 36:323-328.
- Eckensberger, Lutz H., and Roderick F. Zimba 1997 "The Development of Moral Judgment." In J. W. Berry and P. R. Dasen, eds., *Handbook of Cross-Cultural Psychology*, vol. 2, *Basic Processes and Human Development*, 2nd ed. Boston: Allyn and Bacon.
- Emler, Nicholas 1998 "Sociomoral Understanding." In A. Campbell and S. Muncer, eds., *The Social Child*. Hove, East Sussex: Psychology Press.
- , E. Palmer-Canton, and A. St. James 1998 "Politics, Moral Reasoning, and the Defining Issues Test: A Reply to Barnett et al. (1995)." *British Journal of Social Psychology* 37:457-476.
- Garmon, Lance C., Karen S. Basinger, Virginia R. Gregg, and John C. Gibbs 1996 "Gender Differences in Stage and Expression of Moral Judgment." *Merrill-Palmer Quarterly* 42:418-437.
- Gesell, Arnold 1956 *Youth: The Years from Ten to Sixteen*. New York: Harper and Row.
- Gibbs, John C. 1979 "Kohlberg's Moral Stage Theory: A Piagetian Revision." *Human Development* 22:89-112.



- , Karen S. Basinger, and Dick Fuller 1992 *Moral Maturity: Measuring the Development of Sociomoral Reflection*. Hillsdale, N.J.: Erlbaum.
- Gibbs, John C., and K. F. Widaman 1982 *Social Intelligence: Measuring the Development of Sociomoral Reflection*. Englewood Cliffs, N.J.: Prentice-Hall.
- , and Anne Colby 1982 "Construction and Validation of a Simplified Group-Administrable Equivalent to the Moral Judgment Interview." *Child Development* 53:875–910.
- Gilligan, Carol 1982 *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, Mass.: Harvard University Press.
- Gross, Michael L. 1996 "Moral Reasoning and Ideological Affiliation: A Cross-National Study." *Political Psychology* 17:317–338.
- Kohlberg, Lawrence 1969 "Stage and Sequence: The Cognitive-Developmental Approach to Socialization." In D. Goslin, ed., *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- 1971 "From *Is* to *Ought*: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development." In T. Mischel, ed., *Cognitive Development and Epistemology*. New York: Academic.
- 1976 "Moral Stages and Moralization: The Cognitive Developmental Approach." In T. Lickona, ed., *Moral Development and Behavior: Theory, Research, and Social Issues*. New York: Holt, Rinehart, and Winston.
- , D. Boyd, and Charles Levine 1990 "The Return of Stage 6: Its Principle and Moral Point of View." In T. Wren, ed., *The Moral Domain: Essays in the Ongoing Discussion between Philosophy and the Social Sciences*. Cambridge: Massachusetts Institute of Technology Press.
- Kohlberg, Lawrence, Charles Levine, and Alexandra Hewer 1983 *Moral Stages: A Current Formulation and a Response to Critics*. Basel, Switzerland: Karger.
- Kurtines, William, and Esther B. Grief 1974 "The Development of Moral Thought: Review and Evaluation of Kohlberg's Approach." *Psychological Bulletin* 81:453–470.
- Labouvie-Vief, Gisela 1984 "Culture, Language, and Mature Rationality." In K. McCluskey and H. W. Reese, eds., *Life-Span Developmental Psychology: Historical and Generational Effects*. New York: Academic.
- Lapsley, Daniel K. 1996 *Moral Psychology*. Boulder, Colo.: Westview.
- Lind, Georg 1995 "The Meaning and Measurement of Moral Competence Revisited." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- , J. Sandberger, and T. Bargel 1981 "Moral Judgment, Ego Strength, and Democratic Orientations: Some Theoretical Contiguities and Empirical Findings." *Political Psychology* 3:70–110.
- Lind, Georg, and R. Wakenhut 1985 "Testing for Moral Competence." In G. Lind, H. A. Hartmann, and R. Wakenhut, eds., *Moral Development and the Social Environment*. Chicago: Precedent.
- Ma, Hing-Keung, and Chau-Kiu Cheung 1996 "A Cross-Cultural Study of Moral Stage Structure in Hong Kong Chinese, English, and Americans." *Journal of Cross-Cultural Psychology* 27:700–713.
- Markoulis, Diomedes, and Nicolaos Valanides 1997 "Antecedent Variables for Sociomoral Reasoning Development: Evidence from Two Cultural Settings." *International Journal of Psychology* 32:301–313.
- Modgil, Sohan, and Celia Modgil 1986 *Lawrence Kohlberg: Consensus and Controversy*. Philadelphia: Falmer.
- Moore, Erin 1995 "Moral Reasoning: An Indian Case Study." *Ethos* 23:286–327.
- Murphy, John M., and Carol Gilligan 1980 "Moral Development in Late Adolescence and Adulthood: A Critique and Reconstruction of Kohlberg's Theory." *Human Development* 23:77–104.
- Musser, Lynn M., and Christopher Leone 1986 "Moral Character: A Social Learning Perspective." In R. T. Knowles and G. F. McLean, eds., *Psychological Foundations of Moral Education and Character Development: An Integrated Theory of Moral Development*. Lanham, Md.: University Press of America.
- Narvaez, Darcia, Irene Getz, James R. Rest, and Stephen J. Thoma 1999 "Individual Moral Judgment and Cultural Ideologies." *Developmental Psychology* 35:478–488.
- Nisan, Mordecai, and Barbara Applebaum 1995 "Maintaining a Balanced and Respective Identity: Moral Choice in Late Adulthood." In G. Ben-Shakhar and A. Lieblich, eds., *Studies in Psychology in Honor of Solomon Kugelmass*. Jerusalem: Magnes.
- Nucci, Larry P. (ed.) 1989 *Moral Development and Character Education: A Dialogue*. Berkeley, Calif: McCutchan.
- Okonkwo, Rachel U. N. 1997 "Moral Development and Culture in Kohlberg's Theory: A Nigerian (Igbo) Evidence." *IFE Psychologia: An International Journal* 5:117–128.
- Perry, Constance M., and Walter G. McIntire 1995 "Modes of Moral Judgment among Early Adolescents." *Adolescence* 30:707–715.
- Perry, William G., Jr. 1970 *Forms of Intellectual and Ethical Development in the College Years*. New York: Holt, Rinehart, and Winston.

- Piaget, Jean (1932) 1948 *The Moral Judgment of the Child*. Glencoe, Ill.: Free Press.
- 1960 "The General Problems of the Psychobiological Development of the Child." In J. M. Tanner and B. Inhelder, eds., *Discussions on Child Development: Proceedings of the World Health Organization Study Group on the Psychobiological Development of the Child*. New York: International Universities Press.
- , and Barbara Inhelder 1969 *The Psychology of the Child*. New York: Basic.
- Power, Clark F., Ann Higgins, and Lawrence Kohlberg 1989 *Lawrence Kohlberg's Approach to Moral Education*. New York: Columbia University Press.
- Przygotzki, Nathalie, and Etienne Mullet 1997 "Moral Judgment and Aging." *European Review of Applied Psychology* 47:15–21.
- Rest, James R. 1986 *Moral Development: Advances in Research and Theory*. New York: Praeger.
- , Stephen Thoma, and Lynne Edwards 1997 "Designing and Validating a Measure of Moral Judgment: Stage Preference and Stage Consistency Approaches." *Journal of Educational Psychology* 89:5–28.
- Shweder, Richard 1982 "Review of Lawrence Kohlberg's *Essays in Moral Development*, vol. 1, *The Philosophy of Moral Development: Liberalism as Destiny*." *Contemporary Psychology* 27:421–424.
- , and Robert A. LeVine 1975 "Dream Concepts of Hausa Children: A Critique of the 'Doctrine of Invariant Sequence' in Cognitive Development." *Ethos* 3:209–230.
- , M. Mahapatra, and Joan G. Miller 1987 "Culture and Moral Development." In J. Kagan and S. Lamb, eds., *The Emergence of Morality in Young Children*. Chicago: University of Chicago Press.
- Silberman, M. A., and John Snarey 1993 "Gender Differences in Moral Development during Early Adolescence: The Contribution of Sex-Related Variations in Maturation." *Current Psychology: Developmental, Learning, Personality, and Social* 12:163–171.
- Simpson, Elizabeth L. 1974 "Moral Development Research: A Case Study of Scientific Cultural Bias." *Human Development* 17:81–106.
- Skoe, Eva E. 1995 "Sex Role Orientation and Its Relationship to the Development of Identity and Moral Thought." *Scandinavian Journal of Psychology* 36:235–245.
- , 1998 "The Ethic of Care: Issues in Moral Development." In E. E. A. Skoe and A. L. von der Lippe, eds., *Personality Development in Adolescence: A Cross National and Life Span Perspective*. New York: Routledge.
- , and J. E. Marcia 1991 "A Measure of Care-Based Morality and Its Relation to Ego Identity." *Merrill-Palmer Quarterly* 32:289–304.
- Snarey, John R. 1985 "Cross-Cultural Universality of Social-Moral Development: A Critical Review of Kohlbergian Research." *Psychological Bulletin* 97:202–232.
- Sullivan, Edmund V. 1977 "A Study of Kohlberg's Structural Theory of Moral Development: A Critique of Liberal Social Science Ideology." *Human Development* 20:352–376.
- Thoma, Stephen J. 1986 "Estimating Gender Differences in the Comprehension and Preference of Moral Issues." *Developmental Review* 6:165–180.
- , Robert Barnett, James Rest, and Darcia Narvaez 1999 "What Does the DIT Measure?" *British Journal of Social Psychology* 38:103–111.
- Vasudev, Jyotsna 1988 "Sex Differences in Morality and Moral Orientation: A Discussion of the Gilligan and Attanucci Study." *Merrill-Palmer Quarterly* 34:239–244.
- , and Raymond C. Hummel 1987 "Moral Stage Sequence and Principled Reasoning in an Indian Sample." *Human Development* 30:105–118.
- Walker, Lawrence J. 1984 "Sex Differences in the Development of Moral Reasoning: A Critical Review of the Literature." *Child Development* 55:677–691.
- , B. deVries, and S. Trevethan 1987 "Moral Stages and Moral Orientations." *Child Development* 58:842–858.
- Wark, Gillian R., and Dennis L. Krebs 1996 "Gender and Dilemma Differences in Real-Life Moral Judgment." *Developmental Psychology* 32:220–230.
- Woods, Cindy J. P. 1996 "Gender Differences in Moral Development and Acquisition: A Review of Kohlberg's and Gilligan's Models of Justice and Care." *Social Behavior and Personality* 24:375–384.
- Yussen, S. 1977 "Characteristics of Moral Dilemmas Written by Adolescents." *Developmental Psychology* 13:162–163.

THOMAS J. FIGURSKI

## MORTALITY

See Birth and Death Rates; Infant and Child Mortality; Life Expectancy.

## MULTICULTURALISM

See Ethnicity; Indigenous Peoples; Race.

## MULTINATIONAL CORPORATIONS

See Corporate Organizations; Transnational Corporations.

## MULTIPLE INDICATOR MODELS

A primary goal of sociology (and science in general) is to provide accurate estimates of the causal relationship between concepts of central interest to the discipline. Thus, for example, sociologists might examine the causal link between the amount of money people make and how satisfied they are with their lives in general. But in assessing the causal relationships between concepts—such as income and life satisfaction—researchers are subject to making errors stemming from a multitude of sources. In this article, we will focus on one common and especially large source of errors in making causal inferences in sociology and other social and behavioral sciences—specifically, “measurement errors.” Such errors will produce biased (under- or over-) estimates of the true causal relationship between concepts.

*Multiple indicator models* are a method of testing and correcting for errors made in measuring a concept or “latent construct.” Multiple indicators consist of two or more “alternative” measures of the same concept (e.g., two different ways of asking how satisfied you are with your life). Before examining models that use multiple indicators to assess and correct for measurement error, however, the reader should become familiar with the terms and logic underlying such models. “Latent constructs” (also described as a “latent variables”) are unobservable phenomena (e.g., internal states such as the amount of “satisfaction” a person experiences) or more concrete concepts (e.g., income) that represent the hypothetical “true” or “actual” score persons would provide if we could measure a given concept without any error (e.g., a person’s actual income or actual satisfaction, as opposed to their reported income or reported satisfaction).

Additionally, an “indicator” is simply another name for the empirical measure (“observed” score)

of a given construct. For instance, researchers might measure income through a single “indicator” on a questionnaire that asks people “How much money do you earn per year?”—with the response options being, say, five possible income levels ranging from “income greater than \$200,000 per year” (coded as “5”) to “income less than \$10,000 per year” (coded as “1”). Similarly, researchers might provide a single indicator for life satisfaction by asking persons to respond to the statement “I am satisfied with my life.” The response options here might be: “strongly agree” (coded as “5”), “agree” (coded as “4”), “neither agree nor disagree” (coded as “3”), “disagree” (coded as “2”), and “strongly disagree” (coded as “1”).

Social scientists might expect to find a positive association between the above measures of income and life satisfaction. In fact, a review of empirical studies suggests a correlation coefficient ranging from .1 to .3 (Larson 1978; Haring et al. 1984). Correlation coefficients can have values between 1.0 and -1.0, with 0 indicating no association, and 1.0 or -1.0 indicating a perfect relationship. Thus, for example, a correlation of 1.0 for income and life satisfaction would imply that we can perfectly predict a person’s life satisfaction score by knowing that person’s income. In other words, individuals with the highest income (i.e., scored as a “5”) consistently have the highest life satisfaction (i.e., scored as a “5”); those with the next-highest income (i.e., “4”) consistently have the next-highest life satisfaction (i.e., “4”), and so on. Conversely, a -1.0 suggests the opposite relationship. That is, people with the highest income consistently have the lowest life satisfaction; those with the second-highest income consistently have the second-lowest life satisfaction, and so on.

Furthermore, a correlation coefficient of .2, as possibly found, say, between income and life satisfaction, suggests a relatively weak association. A coefficient of this size would indicate that individuals with higher incomes only *tend* to have higher life satisfaction. Hence, we should expect to find many exceptions to this “average” pattern. (More technically, one can square a correlation coefficient to obtain the amount of variance that one variable explains in another. Accordingly, squaring the  $r = .2$  correlation between income and life

satisfaction produces an  $r^2$  of .04—i.e., income explains 4 percent of the variance in life satisfaction.) In sum, given a correlation of .2, we can predict life satisfaction better by knowing someone's income than if we did not have this information (we reduce our errors by 4 percent) but we will still make a lot of errors in our prediction (96 percent of the variance in life satisfaction remains unexplained).

These errors in prediction stem in part from less-than-perfect measures of income and life satisfaction analyzed (a topic covered in the next section). However, they also occur because there are many other causes of life satisfaction (e.g., physical health) in addition to income. The more of these additional causes there are, and the stronger their effects (i.e., the stronger their correlation with life satisfaction), the weaker the ability of a single construct such as income to predict life satisfaction. The same principles apply, of course, to any construct used to predict other constructs (e.g., using people's level of stress to predict the amount of aggression they will display).

Correlation coefficients and "path coefficients" are part of the "language" of causal modeling, including multiple indicator models. Like a correlation coefficient, a path coefficient describes the strength of the relationship between two variables. One can interpret a (standardized) path coefficient in a manner roughly similar to a correlation coefficient. Readers will increase their understanding of the material to follow if they familiarize themselves with these measures of strength of association. (For more information on interpreting correlation and path coefficients, see Blalock [1979]; Kline [1998].)

### RELIABILITY AND VALIDITY OF MEASURES

As noted earlier, measurement errors can bias estimates of the true causal associations between constructs of interest to sociologists and other researchers. Accordingly, it is important to use measures of high quality. Specialists in the field of measurement (often labeled "psychometricians") describe high-quality measures as having strong *reliability* and *validity* (Nunnally and Bernstein 1994). Reliability concerns the consistency with

which an indicator measures a given construct; validity assesses whether one is measuring what one intends to measure or something else.

**Reliability** A common method of assessing reliability is to determine the strength of the correlation (consistency) between alternative (multiple) indicators of the same construct—for example, the correlation between two different measures of life satisfaction. A correlation of 1.0 would suggest that both indicators are perfectly reliable measures of life satisfaction. Conversely, a correlation of, say, .3 would suggest that the two indicators are not very reliable measures of life satisfaction.

Given the subjective nature of indicators of life satisfaction (and many other constructs found in the social and behavioral sciences), we should not be surprised to find fairly low correlations (consistency) among their many possible multiple indicators. The ambiguity inherent in agreeing or disagreeing with statements like "I am satisfied with my life," "The conditions of my life are excellent," and "In most ways my life is close to my ideal" should introduce considerable measurement error. Furthermore, we might anticipate that much of this measurement error would be *random*. That is, relative to respondents' actual ("true") scores for life satisfaction, they would provide answers (observed scores) to the subjective questions regarding life satisfaction that are likely to be nonsystematic. For example, depending on the degree of ambiguity in the multiple indicators for subjective construct like life satisfaction, a person is likely to display a random pattern of giving too high and too low scores relative to the person's true score across the set of measures.

This "noise"—unreliability due to random measurement error—will reduce the correlation of a given indicator with another indicator of the same construct. Indeed, "pure" noise (e.g., completely random responses to questions concerning respondents' life satisfaction) should not correlate with anything (i.e.,  $r = 0$ ). To the extent that researchers can reduce random noise in the indicators (e.g., by attempting to create as clearly worded self-report measures of a construct as possible) the reliability and corresponding correlations among multiple indicators should increase. Even where researchers are careful however, to select the best available indicators of constructs

that represent subjective states (like life satisfaction), correlations between indicators frequently do not exceed  $r$ 's of .3 to .5.

Not only does less-than-perfect reliability reduce correlations among multiple indicators of a given construct, but, more importantly, this random measurement error also reduces the degree to which indicators for one latent construct correlate with the indicators for another latent construct. That is, unreliable measures (such as each of the multiple indicators of life satisfaction) will *underestimate* the true causal linkages between constructs of interest (e.g., the effect of income on life satisfaction). These biased estimates can, of course, have adverse consequences for advancing our scientific knowledge (e.g., perceiving income as a less important source of life satisfaction than it might actually be). (Although unreliable measures will always underestimate the true relationship between two constructs, the bias of unreliable measures is more complex in "multivariate" situations where important control variables may exhibit as much unreliability as [or more unreliability than] do the predictor and outcome variables of interest.)

Psychometricians have long been aware of this problem of "attenuated correlations" from unreliable measures. In response, traditional practice is to combine each of the multiple indicators for a given construct into a single *composite* scale (e.g., sum a person's score across each life satisfaction indicator). The random errors contained in each individual indicator tend to "cancel each other out" in the composite scale (cf. Nunnally and Bernstein 1994), and the overall reliability (typically measured with Cronbach's alpha) on scale ranging from 0 to 1.0 can improve substantially relative to the reliability of individual items within the scale. Although composite scales are a definite step in the right direction, they are still less than perfectly reliable, and often much less. Consequently, researchers are still faced with the problem of biased estimates of the causal linkages among constructs of interest.

**Validity.** Unreliable indicators are not the only source of measurement error that can bias estimates of causal linkages among constructs. Invalid indicators can also create biased estimates. As we shall see in subsequent sections, bias from invalid measures stems from different sources and

is more complex and difficult to detect and control than bias from unreliable measures.

Valid measures require at least modest reliability (i.e., correlations among indicators of a given construct cannot be  $r = 0$ ); but reliable measures are not necessarily valid measures. One can have multiple indicators that are moderately to highly reliable (e.g.,  $r$ 's = .5 to .8), but they may not measure the construct they are supposed to measure (i.e., may not be valid). For example, life satisfaction indicators may display at least moderate reliability, but no one would claim that they are valid measures of, say, a person's physical health.

This example helps clarify some differences between reliability and validity, but at the risk of obscuring the difficulty that researchers typically encounter in establishing valid measures of many latent constructs. Continuing with our example, researchers may select multiple indicators of life satisfaction that critics could never plausibly argue actually measure physical health: Critics might make a very plausible argument, however, that some or all the indicators of life satisfaction also measure a more closely related concept—such as "optimism."

Note, too, that if the life satisfaction indicators do, in fact, also measure optimism, then the correlation that income has with the life satisfaction indicators could stem entirely from income's causal links with an optimistic personality, rather than from income's effect on life satisfaction itself. In other words, in this hypothetical situation, invalid ("contaminated") measures of life satisfaction could lead to *overestimating* income's effect on life satisfaction (though, as we will see in later sections, one can construct examples where invalid measures of life satisfaction could also lead to *underestimating* income's effect).

Given the many subjective, loosely defined constructs that form the core concepts of sociology and other social and behavioral sciences, the issue of what the indicators are actually measuring (i.e., their validity) is a common and often serious problem. Clearly, our scientific knowledge is not advanced where researchers claim a relationship between constructs using invalid measures of one or more of the constructs.

The sections, below, on single-indicator and multiple-indicator models will elaborate on the

bias introduced by measurement error stemming from unreliability and invalidity, and how to use the multiple indicators and “path analysis” of structural equation modeling (SEM) to test and correct for the bias. This discussion continues to use the example of estimating the effect of income on life satisfaction in the face of measurement error.

### SINGLE-INDICATOR MODELS

Figure 1 depicts the hypothesized causal link (solid arrow labeled “ $x$ ”) between the latent (unobservable) constructs of income and life satisfaction (represented with circles), and the hypothesized causal link (solid arrows labeled “ $a$ ” and “ $b$ ”) between each latent construct (circle) and its respective empirical (observed) indicator (box). The  $D$  (disturbances) in Figure 1 represents all potential causes of life satisfaction that the researcher has not included in the causal model (stressful life events, personality characteristics, family relationships, etc.). The  $E$ ’s in Figure 1 represent random measurement error (and any unspecified latent constructs that have a “unique” effect on a given indicator).

Because each latent construct (circled “ $I$ ” and “ $LS$ ”) in Figure 1 has only one indicator (boxed “ $I_1$ ” or “ $LS_1$ ”) to measure the respective construct, researchers describe the diagram as a causal model with *single* (as opposed to multiple) indicators. Additionally, Figure 1 displays a dashed, double-headed arrow (labeled “ $r_1$ ”) between the box for income and the box for life satisfaction. This dashed, double-headed arrow represents the empirical or *observed* correlation between the empirical indicators of income and life satisfaction. Following the logic diagramed in Figure 1, this observed correlation is the product of actual income affecting both measured income through path  $a$ , and actual life satisfaction through path  $x$ . In turn, actual life satisfaction affects measured life satisfaction via path  $b$ . Stated in more formal “path equations,”  $a*x*b = r_1$ .

Note, that researchers can never directly observe the *true* causal effect (i.e., path  $x$ ) of actual income ( $I$ ) on actual life satisfaction ( $LS$ ). Researchers can only *infer* such a relationship based on the *observed* correlation ( $r_1$ ) between the empirical indicators for income and life satisfaction— $I_1$  and  $LS_1$ . In other words, social scientists use

an observed correlation— $r_1$ —to estimate an unobservable true causal effect—path  $x$ .

Notice also that in the presence of random measurement error, the observed correlation  $r_1$  will always be an *underestimate* of the unobservable path  $x$  representing the hypothesized true effect of income on life satisfaction. Of course, researchers hope that  $r_1$  will equal path  $x$ . But  $r_1$  will only equal  $x$  if our empirical measures of income and life satisfaction ( $I_1$  and  $LS_1$ ) are *perfectly reliable*—that is, have no random measurement error.

The phrase “completely reliable measures” implies that each person’s observed score for income and life satisfaction indicators reflect exactly that person’s actual or “true score” for income and life satisfaction. If the indicators for income and life satisfaction are indeed perfect measures, then researchers can attach a (standardized) path coefficient of 1.0 to each path (i.e.,  $a$  and  $b$ ) between the latent constructs and their indicator. Likewise, researchers can attach a path coefficient of 0 to each path (i.e.,  $d$  and  $e$ ) representing the effects of random measurement errors ( $E_1$  and  $E_2$ ) on the respective indicators for income and life satisfaction.

The path coefficient of 1.0 for  $a$  and  $b$  signifies a perfect relationship between the latent construct (i.e., true score) and the measure or indicator for the latent construct (i.e., recorded score). In other words, there is no “slippage” between the actual amount of income or life satisfaction people have and the amount of income or life satisfaction that a researcher records for each person (i.e., there is no random measurement error). Therefore, people who truly have the highest income will report the most income, those who truly have the lowest income will report the lowest income, and so on. Likewise, people who, in fact, have the most life satisfaction will always record a life satisfaction score (e.g., “5”) higher than those a little less satisfied (e.g., “4”), individuals a little less satisfied will always record a life satisfaction score higher than those a little bit less satisfied yet (e.g., “3”); and so on.

Under the assumption that the measures of income and life satisfaction are perfectly reliable, social scientists can use the *observed* correlation ( $r_1$ ) between the indicators  $I_1$  and  $LS_1$  to estimate the *true* causal effect (path  $x$ ) of actual income ( $I$ ) on actual life satisfaction ( $LS$ ). Specifically,  $r_1 = a*x*b$ ;

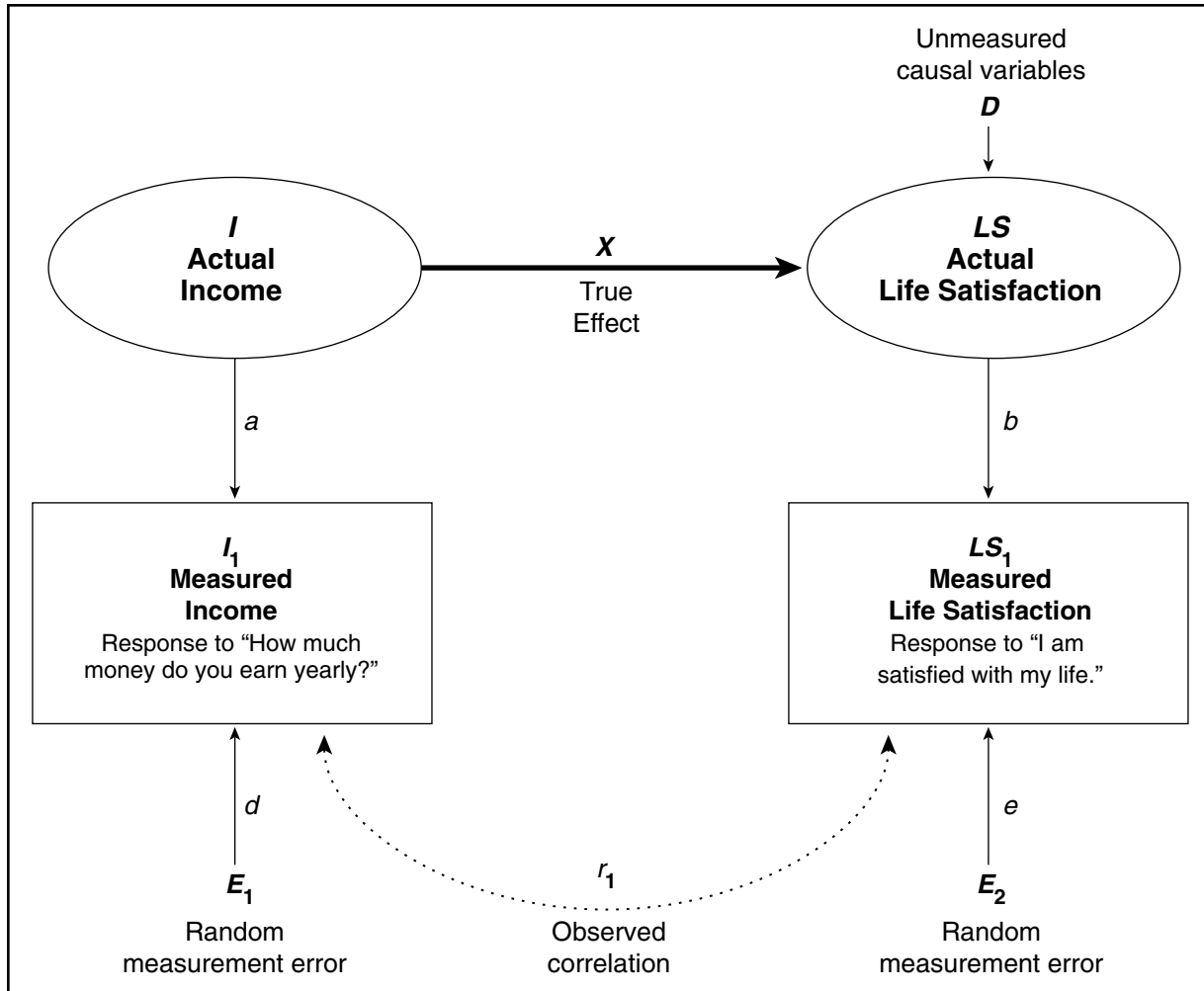


Figure 1. Single-Indicator Model for Estimating the Effect of Income on Life Satisfaction.

hence,  $r_1 = 1.0 * x * 1.0$ , or  $r_1 = x$ . Thus, if the observed correlation between income and life satisfaction (i.e.,  $r_1$ ) is, say, .2, then the true (unobservable) causal effect of income on life satisfaction (i.e., path  $x$ ) would also be .2. (For more detailed explanations of how to interpret and calculate path coefficients, see Sullivan and Feldman [1979]; Loehlin [1998].)

Of course, even if researchers were to measure income and life satisfaction with perfect reliability (i.e., paths  $a$  and  $b$  each equal 1.0), there are other possible errors (“misspecifications”) in the model shown in Figure 1 that could bias researchers’ estimates of how strong an effect income truly has on life satisfaction. That is, Figure 1 does not

depict other possible “misspecifications” in the model such as “reverse causal order” (e.g., amount of life satisfaction determines a person’s income) or “spuriousness” (e.g., education determines both income and life satisfaction and hence only makes it appear that income causes life satisfaction). (For more details on these additional sources of potential misspecification, see also Blalock [1979].)

How realistic is the assumption of perfect measurement in single indicator causal models? The answer is, “It depends.” For example, to assume a path coefficient of 1.0 for path  $a$  in Figure 1 would not be too unrealistic (the actual coefficient is likely a little less than 1.0—say, .90 or .95). That is, we would not expect many errors in

measuring a person's true income. Likewise, we would expect few measurement errors in recording, say, a person's age, sex, or race. As noted earlier, the measurement of subjective states (including satisfaction with life) is likely to occur with considerable error. Therefore, researchers would likely *underestimate* the true causal link between income and life satisfaction, if they assumed no random measurement error for the life satisfaction indicator—that is, if they assumed a coefficient of 1.0 for path *b* in Figure 1.

How badly researchers underestimate the true causal effect (i.e., path *x*) would depend, of course, on how much less than 1.0 was the value for path *b*. For the sake of illustration, assume that path *b* equals .5. Assume also that income is perfectly measured (i.e., path *a* = 1.0) and the observed correlation ( $r_1$ ) between the indicators for income and life satisfaction is .2. Under these conditions,  $1.0 * x * .5 = .2$ , and  $x = .4$ . In other words, researchers who report the observed correlation between income and life satisfaction ( $r_1 = .2$ ) would substantially *underestimate* the strength of the true effect ( $x = .4$ ) that income has on life satisfaction. Recall that an  $r$  of .2 represents an  $r^2$  of .04, or 4 percent of the variance in life satisfaction explained by income, whereas an  $r$  of .4 represents an  $r^2$  of .16, or 16 percent of the variance explained by income. Accounting for 16 percent of the variance in life satisfaction represents a much stronger effect for income than does 4 percent explained variance. Based on this hypothetical example, income goes from a weak to a relatively strong predictor of life satisfaction, if we have the appropriate information to allow us to correct for the bias of the unreliable measure of life satisfaction.

But how do researchers know what values to assign to the unobservable paths (such as *a* and *b* in Figure 1) linking a given latent construct to its single indicator? Unless logic, theory, or prior empirical evidence suggests that the constructs in a given model are measured with little error (i.e., the path between the circle and the box = 1.0) or indicate what might be an appropriate value less than 1.0 to assign for the path between the circle and box, researchers must turn from *single* indicator models to *multiple* indicator models. As noted earlier, multiple indicator models allow one to make corrections for measurement errors that would otherwise bias estimates of the causal relationships between constructs.

## MULTIPLE INDICATOR MODELS

For researchers to claim that they are using multiple indicator models, at least one of the concepts in a causal model must have more than one indicator. "Multiple indicators" means simply that a causal model contains alternative measures of the same thing (same latent construct). Figure 2 depicts a multiple indicator model in which income still has a single indicator but life satisfaction now has three indicators (i.e., three alternative measures of the same life satisfaction latent construct). In addition to the original indicator for life satisfaction—"I am satisfied with my life"—there are two new indicators; namely, "In most ways my life is close to my ideal" and "The conditions of my life are excellent." (Recall that the possible response categories range from "strongly agree" to "strongly disagree." See also Diener et. al. [1985] for a fuller description of this measurement scale.)

As in Figure 1, the dashed, double-headed arrows represent the *observed* correlations between each pair of indicators. Recall that these observed correlations stem from the assumed operation of the latent constructs inferred in the causal model. Specifically, an increase in actual income (*I*) should produce an increase in measured income ( $I_1$ ) through (causal) path *a*. Moreover, an increase in actual income should also produce an increase in actual life satisfaction (*LS*) through (causal) path *x*, which in turn should produce an increase in each of the measures of life satisfaction ( $LS_1$ ,  $LS_2$ , and  $LS_3$ ) through (causal) paths *b*, *c*, and *d*. In other words, these hypothesized causal pathways should produce observed correlations between all possible pairs of the measured variables ( $I_1$ ,  $LS_1$ ,  $LS_2$ , and  $LS_3$ ). (We are assuming here that the underlying measurement model is one in which the unobserved constructs have a causal effect on their respective indicators. There are some types of multiple indicators, however, where a more plausible measurement model would suggest the reverse causal order—i.e., that the multiple indicators each drive the latent construct common to the set of indicators. See, Kline [1998] for a discussion of these "cause" indicator measurement models, as opposed to the more traditional "effect" indicator measurement model described here.)

The use of a single indicator for income means that researchers must use logic, theory, or prior



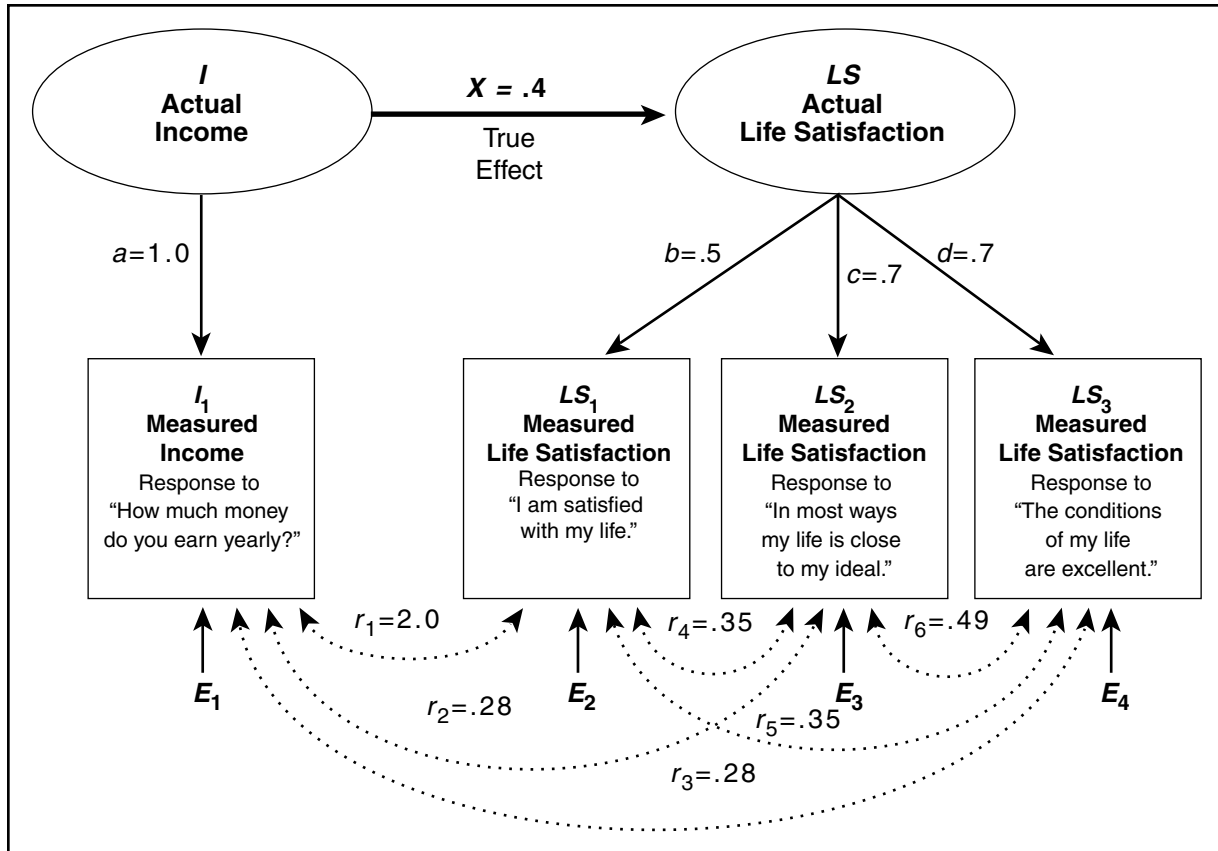


Figure 2. Multiple-Indicator Model for Estimating the Effect of Income on Life Satisfaction.

empirical evidence to assign a path coefficient to represent “slippage” between the true score ( $I$ ) and the measured score ( $I_1$ ). For income, path  $a = 1.0$  (i.e., no random measurement error) seems like a reasonable estimate, and makes it easier to illustrate the path equations. (Although a coefficient of say, .95, might be more realistic, whether we use 1.0 or .95 will make little difference in the calculations that follow.) As noted earlier, however, indicators of life satisfaction are not as easily assigned path coefficients of 1.0. That is, there is likely to be considerable random error (unreliability) in measuring a subjective state such as life satisfaction. Fortunately, however, the use of *multiple indicators* for the life satisfaction construct permits researchers to provide reasonable estimates of random measurement error based on the *empirical data in the current study*—namely, the observed correlations (i.e., consistency) among the multiple indicators. Because measurement error can vary so much from one research setting to

another, it is always preferable to provide estimates of reliability based on the current rather than previous empirical studies. Likewise, reliability estimates based on the current study are much better than those estimates obtained from logic or theory, unless the latter sources can provide a compelling case for a highly reliable single indicator (such as the measure of income used in the present example).

If the multiple indicator model in Figure 2 is correctly specified, then the observed correlations among the several pairs of indicators should provide researchers with information to calculate estimates for the hypothesized (unobservable) causal paths  $b$ ,  $c$ , and  $d$  (i.e., estimates of how much “slippage” there is between actual life satisfaction and each measure of life satisfaction). Researchers can use hand calculations involving simple algebra to estimate the causal paths for such simple multiple indicator models as depicted in Figure 2 (for

examples, see Sullivan and Feldman 1979 and Loehlin 1998). But more complicated models are best left to “structural equation modeling” (SEM) computer software programs such as LISREL (Linear Structural Relationships; Joreskog and Sorbom 1993), EQS (Equations; Bentler 1995), or AMOS (Analysis of Moment Structures; Arbuckle 1997). Kline (1998) provides a particularly excellent and comprehensive introduction to the topic. There are two annotated bibliographies that represent almost all work related to SEM up to about 1996 (Austin and Wolfe 1991; Austin and Calderon 1996). Marcoulides and Schumacker (1996) and Schumacker and Marcoulides (1998) cover even more recent advances. Smallwaters software company has a Web site that gives a wealth of information, including other relevant Web sites: <http://www.smallwaters.com/weblinks>.

In essence, these SEM computer programs go through a series of trial-and-error “iterations” in which different values are substituted for the hypothesized causal paths—in Figure 2, paths *b*, *c*, *d*, and *x*. (Recall that we assigned a value of 1.0 for path *a*, so the SEM program does not have to estimate a value for this hypothesized path.) Ultimately, the program reaches (“converges on”) a “final solution.” This solution will *reproduce* as closely as possible the *observed correlations*—in Figure 2,  $r_1$ ,  $r_2$ ,  $r_3$ ,  $r_4$ ,  $r_5$ , and  $r_6$ —among each of the indicators in the proposed causal model. In the *final solution* depicted in Figure 2, the path estimates for *b*, *c*, *d*, and *x* (when combined with the “assigned” or “fixed” value for path *a*) exactly reproduce the observed correlations. (Note that the *final* solution will reproduce the observed correlations better than will the *initial* solutions, unless, of course, the SEM program finds the best solution on its first attempt—which is not likely in most “real-world” data analyses.)

More technically, the SEM program builds a series of (simultaneous) equations that represent the various hypothesized causal paths that determine each observed correlation. In Figure 2, the correlation ( $r_1 = .20$ ) for  $I_1$  and  $LS_1$  involves the “path” equation:  $a*x*b = .20$ ; for the correlation ( $r_2 = .28$ ) of  $I_1$  and  $LS_2$ :  $a*x*c = .28$ ; for the correlation ( $r_3 = .28$ ) of  $I_1$  and  $LS_3$ :  $a*x*d = .28$ ; for the correlation ( $r_4 = .35$ ) of  $LS_1$  and  $LS_2$ :  $b*c = .35$ ; for the correlation ( $r_5 = .35$ ) of  $LS_1$  and  $LS_3$ :  $b*d = .35$ ; and for the correlation ( $r_6 = .49$ ) of  $LS_2$  and  $LS_3$ :  $c*d =$

.49. The SEM program then uses the *known* values—observed correlations and, in the causal model for Figure 2, the fixed (predetermined) value of 1.0 for path *a*—to simultaneously solve the set of equations to obtain a value for each of the causal paths that initially have *unknown* values.

Except in artificial examples (like Figure 2), however, the SEM program is unlikely to obtain final values for the causal paths such that the path equations exactly reproduce the observed correlations. More specifically, the program will attempt through its iterative trial-and-error procedures to find a final value for each of the causal paths *b*, *c*, *d*, and *x* that will *minimize* the “average discrepancy” across each of the six model-*implied* correlations (i.e., predicted by the path equations) versus the six empirically *observed* correlations.

For example, to reproduce the observed correlation ( $r_4 = .35$ ) between  $LS_1$  and  $LS_2$  (recall that empirical correlations among indicators of subjective states like life satisfaction often range between .3 and .5), the SEM program would have to start with values for causal paths *b* and *c* (i.e., estimates of “slippage” between actual life satisfaction and measured life satisfaction) considerably lower than 1.0. In other words, the software program would need to allow for some random measurement error. If, instead, the initial solution of the SEM program assumed perfect reliability, there would be a substantial discrepancy between at least some of the implied versus observed correlations among the indicators. That is, multiplying the path  $b = 1.0$  by the path  $c = 1.0$  (i.e., assuming perfect reliability of each indicator) would imply an observed correlation of 1.0 (i.e., perfect consistency) between  $LS_1$  and  $LS_2$ —an implied (i.e., predicted) correlation that far exceeds the  $r_4 = .35$  correlation we actually observe. (Keep in mind that, as depicted in Figure 2, the *best values* for *b* and *c* are .5 and .7, respectively, which the SEM program will eventually converge on as it “iterates” to a final solution).

If, at the next iteration, the SEM program were to substitute equal values for *b* and *c* of about .59 each (to allow for less than perfect reliability), this solution would exactly reproduce the observed correlation of .35 (i.e.,  $.59*.59 = .35$ ) between  $LS_1$  and  $LS_2$ . But using a value of .59 for both *b* and *c* would *not* allow the program to reproduce the observed correlations for  $LS_1$  and  $LS_2$  with  $I_1$  (the indicator for income). To obtain the observed

correlation of .20 between  $I_1$  and  $LS_1$ , the program needs to multiply the paths  $a*x*b$ . Accordingly,  $r_1$  (.20) should equal  $a*x*b$ —that is,  $1.0*x*.59 = .20$ . Solving for  $x$ , the program would obtain a path value of about .35. Likewise, to obtain the observed correlation of .28 between  $I_1$  and  $LS_2$ , the program needs to multiply the paths  $a*x*b$ . Accordingly,  $r_2$  (.28) must equal  $a*x*c$ —that is,  $1.0*x*.59 = .28$ . Solving for  $x$ , the program would obtain a path value of about .47. In other words, the program cannot find a solution for the preceding two equations that uses the *same value* for  $x$ . That is, for the first equation  $x \approx .35$ . But for the second equation  $x \approx .47$ .

Given the SEM program's need to come up with a *unique* (i.e., single) value for  $x$  (and for all the other causal paths the program must estimate), a possible compromise might be to use a value of .41. Substituting this value into the preceding two equations— $a*x*b$  and  $a*x*c$ —would provide an *implied* correlations of about .24— $1.0*.41*.59 = .24$ —for both equations. Comparing this implied correlation with the *observed* correlations of .20 and .28 for  $I_1/LS_1$  and  $I_1/LS_2$ , respectively, the discrepancy in these two situations is +/- .04.

Although this is not a large discrepancy between the implied and the observed correlations, the SEM program can do better (at least in this hypothetical example). If the SEM program subsequently estimates values of .50 and .70 for causal paths  $b$  and  $c$ , respectively, then it is possible to use the same value of  $x$  (specifically, .40) for each path equation involving  $I_1/LS_1$  and  $I_1/LS_2$ , and reproduce exactly the observed correlations for  $r_1$  (i.e.,  $1.0*.40*.50 = .20$ ) and  $r_2$  (i.e.,  $1.0*.40*.70 = .28$ ). By using these estimated values of .5 and .7 for paths  $b$  and  $c$  in place of the .6 and .6 values initially estimated, the program can also reproduce exactly the observed correlation ( $r_4 = .35$ ) between  $LS_1$  and  $LS_2$ —that is,  $.5*.7 = .35$ . Furthermore, by using an estimated value of .7 for causal path  $d$ , the program can exactly reproduce *all* the remaining observed correlations in Figure 2— $r_3 = .28$ ,  $r_5 = .35$ , and  $r_6 = .49$ —that involve path  $d$ , that is,  $a*x*d$ ,  $b*d$ , and  $c*d$ , respectively. (We leave to the reader the task of solving the equations.)

In sum, by using the fixed (a priori) value of  $a = 1.0$  and the estimated values of  $x = .4$ ,  $b = .5$ ,  $c = .7$ , and  $d = .7$ , the six implied correlations exactly

match the six observed correlations depicted in Figure 2. In other words, the hypothesized causal paths in our model provide a “perfect fit” to the “data” (empirical correlations).

Reproducing the observed correlations among indicators does not, however, establish that the proposed model is correct. In the logic of hypothesis testing, one can only disconfirm models, not prove them. Indeed, there is generally a large number of alternative models, often with entirely different causal structures, that would reproduce the observable correlations just as well as the original model specified (see Kim and Mueller 1978 for examples). It should be apparent, therefore, that social scientists must provide rigorous logic and theory in building multiple indicator models, that is, in providing support for one model among a wide variety of possible models. In other words, multiple indicator procedures require that researchers think very carefully about how measures are linked to latent constructs, and how latent constructs are linked to other latent constructs.

Additionally, it is highly desirable that a model contain more observable correlations among indicators than unobservable causal paths to be estimated—that is, the model should be “overidentified.” For example, Figure 2 has *six* observed correlations ( $r_1, r_2, r_3, r_4, r_5$ , and  $r_6$ ) but only *four* hypothesized causal paths ( $x, b, c$ , and  $d$ ) to estimate. Thus, Figure 2 is overidentified—with *two* “degrees of freedom” (df) By having an excess of observed correlations versus hypothesized causal paths (i.e., by having at least one and preferably many degrees of freedom), a researcher can provide tests of “model fit” to assess the probability that there exist alternative causal paths not specified in the original model. (Where the fit between the implied vs. observed correlations is poor, researchers typically seek to revise their causal model to better fit the data.)

Conversely, “just-identified” models will contain exactly as many observable correlations as hypothesized causal paths to be estimated (i.e., will have 0 degrees of freedom). Such models will *always* produce estimates (solutions) for the causal paths that *exactly* reproduce the observable correlations—no matter how badly misspecified the proposed causal pathways may be. In other words,

“perfect” fit is inevitable and provides no useful information regarding whether the model is correctly specified or not. This result occurs because, unlike an overidentified model, a just-identified model does not have any degrees of freedom with which to detect alternative causal pathways to those specified in the original model. Accordingly, just-identified models are not very interesting to SEM practitioners.

Finally, the worst possible model is one that is “underidentified,” that is, has fewer observable correlations than unobservable causal paths to be estimated. Such models can provide no single (unique) solutions for the unobservable paths. In other words, an infinite variety of alternative estimates for the causal paths is possible. For example, if we restricted Figure 2 to include only  $LS_1$  and  $LS_2$  (i.e., dropping  $LS_3$  and  $I_1$ ), the resulting two-indicator model of life satisfaction would be underidentified. That is, we would have *two* causal paths to estimate—from the latent construct to each of the two indicators (i.e., paths  $b$  and  $c$ )—but only *one* observed correlation ( $r_4 = .35$ ). Under this situation, there is no *unique* solution. We can literally substitute an infinite set of values for paths  $b$  and  $c$  to exactly reproduce the observed correlation (e.g., given  $b*c = .35$ , we can use  $.7*.5$  or  $.5*.7$ , or two values slightly less than  $.6$  each, and so on).

The number of indicators per latent construct helps determine whether a model will be overidentified or not. In general, one should have at least three and preferably four or more indicators per latent construct—unless one can assume a single indicator, such as income in Figure 2, has little measurement error. Adding more indicators for a latent construct rapidly increases the “overidentifying” pieces of empirical information (i.e., degrees of freedom). That is to say, observable correlations (between indicators) grow faster than the unobservable causal paths (between a given latent construct and indicator) to be estimated.

For example, adding a fourth indicator for life satisfaction in Figure 2 would require estimating *one* additional causal path (linking the life satisfaction latent construct to the fourth indicator), but would also produce *four* more observed correlations ( $LS_4$  with  $LS_3$ ,  $LS_2$ ,  $LS_1$ , and  $I_1$ ). The modified model would thus have *three* more degrees of freedom, and correspondingly greater *power* to

determine how well the model fits the empirical data (observed correlations). Including a fifth indicator for life satisfaction would produce *four* more degrees of freedom, and even more power to detect a misspecified model. (The issue of model identification is more complicated than outlined here. The requirement that an identified model have at least as many observed correlations as causal paths to be estimated is a necessary but not sufficient condition; cf. Kline [1998].)

Some additional points regarding multiple indicator models require clarification. For instance, in “real life” a researcher would never encounter such a perfect reproduction of the (noncausal) observable correlations from the unobservable (causal) paths as Figure 2 depicts. (We are assuming here that the causal model tested in “real life,” like that model tested in Figure 2, is “overidentified.” Recall that a “just-identified” model always exactly reproduces the observed correlations.) Indeed, even if the researcher’s model is correctly specified, the researcher should expect at least *some* minor discrepancies in comparing the observed correlations among indicators with the correlations among indicators predicted (implied) by the hypothesized causal paths.

Researchers can dismiss as “sampling error” (i.e., “chance”) any discrepancies that are not too large (given a specific sample size). At some point, however, the discrepancies do become too large to dismiss as “chance.” At that point, researchers may determine that they have not specified a proper model. Poor model fit is strong grounds for reevaluating and respecifying the original causal model—typically by adding and (less often) subtracting causal paths to obtain a better fitting model. Just because an overidentified model can detect that a model is misspecified does not mean, however, that it is easy to then tell where the misspecification is occurring. Finding and correcting misspecification is a complex “art form” that we cannot describe here (but see Kline 1998 for an overview).

The next section will describe how nonrandom measurement error can create a misfitting multiple indicator model and corresponding bias in estimates of causal pathways. Additionally, we will demonstrate how a just-identified model, in contrast to an overidentified model, will fail to detect

and thus correct for this misspecification, resulting in considerable bias in estimating the true effect of income on life satisfaction.

**MULTIPLE-INDICATOR MODELS WITH NONRANDOM MEASUREMENT ERROR**

We have discussed how poor quality measures—low reliability and low validity—can bias the estimates of the true effects of one latent construct on another. Our specific modeling examples (Figures 1 and 2), however, have focused on the bias introduced by unreliable measures only. That is, our causal diagrams have assumed that all measurement error is *random*. For example, Figure 2 depicts the error terms (*E*'s) for each of three multiple indicators of life satisfaction to be *unconnected*. Such random measurement error can occur for any number of reasons: ambiguous questions, coding errors, respondent fatigue, and so forth. But none of these sources of measurement error should *increase* correlations among the multiple indicators. Indeed, as noted in previous sections, random measurement error should *reduce* the correlations (consistency) among multiple indicators of a given latent construct—less-than-perfect correlations that researchers can then use to estimate reliability and thereby correct for the bias that would otherwise occur in underestimating the true effect of one latent construct on another (e.g., income on life satisfaction).

Conversely, where measurement error *increases* correlations among indicators, social scientists describe it as systematic or *nonrandom*. Under these conditions, the measurement errors of two or more indicators have a common source (a latent construct) other than or in addition to the concept that the indicators were suppose to measure. The focus now becomes the *validity* of measures. Are you measuring what you claim to measure or something else? (See the section above entitled “Reliability and Validity of Measures” for a more general discussion.)

Failure to include nonrandom—linked or “correlated”—errors in a multiple-indicator model will bias the estimates of other causal paths in the model. Figure 3 depicts such a linkage of error terms through the personality variable “optimism” (*O*). Based on the hypothetical model in Figure 3,

the observed correlation ( $r_6$ ) between the indicators  $LS_2$  and  $LS_3$  would not be entirely the consequence of the effects of life satisfaction ( $LS$ ) operating through the causal paths  $c$  and  $d$ . In fact, part of this observed correlation would occur as a consequence of the causal paths  $e$  and  $f$  (which in this hypothetical example, we have constrained to be equal). In other words, the indicators  $LS_2$  and  $LS_3$  measure some of life satisfaction but also optimism. That is, they measure something in addition to what they were intended to measure. Stated in still other words, the two indicators are not “pure” (completely valid) measures of life satisfaction because they are “contaminated” by also tapping optimism.

Note that  $r_6$  is the only observed correlation that differs in strength in comparing Figures 2 and 3. For Figure 2, this correlation equals .49; for Figure 3, this correlation equals .85. The higher observed correlation for  $r_6$  in Figure 3 stems from the “inflated” correlation produced by the effects of optimism through paths  $e$  (.6) and  $f$  (.6). Note, also, that  $.6 \times .6$  equals .36. If we add .36 to the original observed correlation for  $r_6$  (i.e., .49) in Figure 2, we obtain the observed correlation of .85 in Figure 3. All other paths estimates and observed correlation remain the same across the two figures. Furthermore, like Figure 2, Figure 3 depicts path estimates for a final solution that exactly reproduce all observed correlations.

Note, too, if we had not added the causal paths ( $e$  and  $f$ ) to represent the hypothesized effect of optimism on two measures of life satisfaction, we could not obtain a “good fit” for the observed correlations in Figure 3. Indeed, *without* these additional paths, the SEM computer program would have to increase the path coefficients for  $c$  and  $d$ —say, to about .92 each—in order to reproduce the observed correlation of .85 for  $r_6$ . But then the program would fail to reproduce the observed correlations involving  $LS_2$  and  $LS_3$  with the other indicators in the model ( $LS_1$  and  $I_1$ ). For example, the observed correlation between  $I_1$  and  $LS_2$  ( $r_2 = .28$ ) would now be *overestimated*, based on the product of the causal paths— $a \times x \times c$ —that the model in Figure 3 suggests determines  $r_2$ . That is,  $1.0 \times .4 \times .92$  results in an implied (predicted) correlation of about .37, which leaves a discrepancy of .09 relative to the correlation (.28) actually observed.

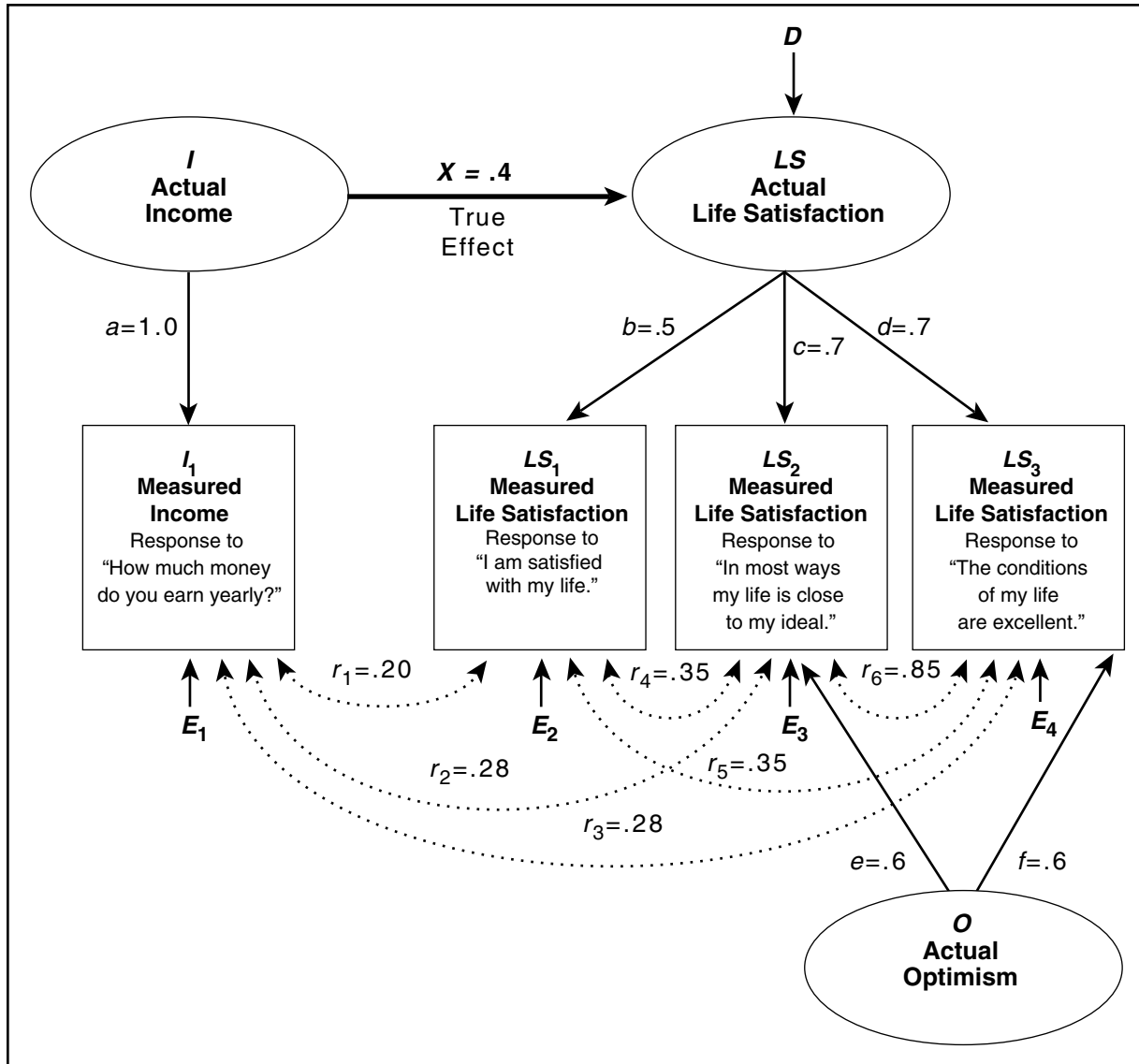


Figure 3. Multiple-Indicator Model for Estimating the Effect of Income on Life Satisfaction with Nonrandom Measurement Error for Optimism\*.

NOTE: \*To obtain an overidentified model, paths *k* and *l* are “constrained” to have equal path coefficients.

In the absence of good fit—as a consequence of not allowing paths *e* and *f* to be included in the model to be tested (i.e., not specifying the “correct” model)—the SEM program might continue to “iterate” by substituting new estimates for *c* and the other causal paths in Figure 3—*b*, *d*, and *x*—which initially have unknown values. No matter what values are estimated for each of these causal paths, however, the program will not be able to eliminate a discrepancy between the implied and

the observed correlations for at least some of the pairs of indicators—a discrepancy that, as noted previously, the program will attempt to minimize based on the criterion of finding path estimates that result in the lowest average discrepancy summed across all possible comparisons of implied versus observed correlations in the model.

Indeed, a misspecification in one part of a model (in this instance, failure to model nonrandom

measurement error in  $LS_1$  and  $LS_2$  indicators) generally “reverberates” throughout the causal model, as the SEM program attempts to make adjustments in all path coefficients to minimize the average discrepancy. In other words, estimates for each path coefficient are most often at least slightly biased by misspecified models (the greater the misspecification, the greater the bias). Most importantly, the failure to model (and thus correct for) nonrandom measurement error will typically result in the SEM program’s estimating a value for the true effect of one construct on another—in the present example, the true effect (path  $x$ ) of income on life satisfaction—that will be biased (either over- or underestimated). For example, when we use a SEM program (EQS) to calculate estimates for the hypothesized causal paths in Figure 3 (excluding the paths for the contamination of optimism), we obtain a final solution with the following estimates:  $b = .38$ ,  $c = .92$ ,  $d = .92$ , and  $x = .31$ . Thus, in the present example, the failure to model nonrandom measurement error has led to an underestimate ( $x = .31$ ) of income’s true effect ( $x = .40$ ) on life satisfaction. Interestingly, by far the largest discrepancy in fit occurs between  $I_1$  and  $LS_1$  (the implied correlation is .08 less than the observed correlation), not between  $LS_2$  and  $LS_3$  (where there is 0 discrepancy)—demonstrating that misfit in one part of the model can produce misfit (reverberate) elsewhere in the model.

The hypothetical model depicted in Figure 3 is necessarily artificial (to provide simple illustrations of SEM principles). In real-life situations, it is unlikely (though still possible) that sources of contamination, when present, would impact only some of the multiple indicators for a given construct (e.g., only two of the three measures of life satisfaction). Furthermore, researchers would be on shaky ground if they were to actually claim that the correlated measurement error modeled in the present example stemmed from an identifiable source—such as optimism. (Indeed, to avoid specifically labeling the source of the inflated correlation between  $LS_1$  and  $LS_2$ , researchers would most likely model a curved, double-headed arrow between  $E_3$  and  $E_4$  in Figure 3, and have the SEM program simply estimate the appropriate value—.36—for this new path representing correlated measurement error.) To make a legitimate claim that a construct like optimism contaminates a construct such as life satisfaction, researchers

would also need to provide traditional measures of the suspected source of contamination (e.g., include established multiple indicators for optimism).

As a final example of the need to use *overidentified* multiple indicator models in order to detect misspecified models, consider a modification of Figure 3 in which we have access to the single indicator for income ( $I_1$ ) but have only two indicators for life satisfaction available to us— $LS_2$  and  $LS_3$ . Assume, further, that we are unaware that optimism contaminates the two life satisfaction indicators. Note that this new model is *just identified*. That is, there are three observed correlations ( $r_2$ ,  $r_3$ , and  $r_6$ ) and three hypothesized causal paths to estimate ( $x$ ,  $c$ , and  $d$ ). Accordingly, this model will perfectly fit the data, despite the fact that we have not modeled (i.e., misspecified) the correlated measurement error (.36) that we’ve built into the .85 observed correlation ( $r_6$ ) between  $LS_2$  and  $LS_3$ .

More specifically, in order to reproduce the .85 correlation, the SEM program will increase the estimate of reliability for the  $LS_2$  and  $LS_3$  life satisfaction indicators by increasing the  $c$  and  $d$  paths from their original (true) value of .7 each to new (biased) values of about .92 each. To reproduce the observed correlation of .28 for both  $I_1/LS_2$  and  $I_1/LS_3$  ( $r_2$  and  $r_3$ ), the program can simply decrease the  $x$  path (i.e., the estimate of the true effect of income on life satisfaction) from its original (true) value of .40 to a new (biased) value of about .30—that is, for both path equations  $a^*x*c$  and  $a^*x*d$ :  $1.0*.3*.92 = .28$ .

In sum, our just-identified model perfectly fits the data despite the fact our model has failed to include (i.e., has misspecified) the correlated measurement error created by optimism contaminating the two life satisfaction indicators. More importantly, the failure to model the nonrandom measurement error in these two indicators has led us to a biased underestimate of the true effect of income and life satisfaction: Income explains  $.3^2 = 9\%$  of the variance in life satisfaction for the current (biased) model, but explains  $.4^2 = 16\%$  variance in life satisfaction for the earlier (correctly specified) model.

As noted earlier, the debilitating effects of nonrandom measurement error are extremely diverse in their potential forms and correspondingly

complex in their potential bias in under- or overestimating causal links among latent constructs. The present discussion only touches on the issues and possible multiple indicator models for detecting and correcting this type of measurement error. The reader is encouraged to consult other sources that provide additional details (starting with Kline 1998).

#### STRENGTHS AND WEAKNESSES OF MULTIPLE-INDICATOR MODELS USING SEM

As we have seen, by incorporating multiple indicators of constructs, structural equation modeling procedures allow more rigorous tests of causal models than possible when using single indicators. With sufficient multiple indicators to provide an overidentified model, SEM procedures are particularly powerful in detecting and correcting for the bias created by less than perfectly reliable and valid measures of latent constructs of interest. In other words, by building a more accurate *measurement model*—that is, providing a more correct specification of the causal linkages between latent constructs and their indicators—SEM researchers can obtain more accurate estimates of the causal linkages in the *structural model*—that is, the effect of one latent variable on another.

A correct specification of the measurement model is, then, a means to the ultimate goal of obtaining accurate path estimates (also described as “parameter estimates”) for the structural model (e.g., obtaining the true effect of income on life satisfaction by correcting for measurement error). Some research, however, makes the measurement model the central focus (e.g., whether the three indicators of life satisfaction correlate with each other in a manner consistent with their measuring a single latent construct). In this situation, the SEM procedure is known as “confirmatory factor analysis” (CFA). The CFA uses the typical SEM principle of positing a measurement model in which latent constructs (now described as “factors”) impact multiple indicators (the path coefficient now described as a “factor loading”). Unlike a “full-blown” SEM analysis, however, the CFA does not specify causal links among the latent constructs (factors), that is, does not specify a structural model (e.g., there is no interest in whether income affects life satisfaction). Instead, the CFA treats

potential relationships among any two factors as simply a “correlation” (typically represented in diagrams with a curved, double-headed arrow between the factors); that is, the causal ordering among latent constructs is *not specified*.

Researchers have developed especially powerful CFA procedures for establishing the validity of measures through significant enhancements to traditional construct validity techniques based on multitrait/multimethods (MTMM), second-order factor structures, and testing the invariance of measurement models across diverse subsamples (again, see Kline 1998 for an introduction). CFA has also been at the vanguard in developing “nonlinear” factor analysis to help reduce the presence of “spurious” or “superfluous” latent constructs (referred to as “methods” or “difficulty” factors) that can occur as a consequence of using multiple indicators that are highly skewed (i.e., non-normally distributed) (Muthen 1993, 1998). These nonlinear factor analysis procedures also provide a bridge to recent advances in psychometric techniques based on item-response theory (cf. Reise et al. 1993; Waller et al. 1996).

CFA provides access to such powerful psychometric procedures as a consequence of its enormous flexibility in specifying the structure of a measurement model. But this strength is also its weakness. CFA requires that the researcher specify all aspects of the measurement model a priori. Accordingly, one must have considerable knowledge (based on logic, theory, or prior empirical evidence) of what the underlying structure of the measurement model is likely to be. In the absence of such knowledge, measurement models are likely to be badly misspecified. Although SEM has available particularly sophisticated techniques for detecting misspecification, simulation studies indicate that these techniques do not do well when the hypothesized model is at some distance from the true model (cf. Kline 1998).

Under these circumstances of less certainty about the underlying measurement model, researchers often use “exploratory factor analysis” (EFA). Unlike CFA, EFA requires no a priori specification of the measurement model (though it still works best if researchers have some idea of at least the likely number of factors that underlie the set of indicators submitted to the program). Individual



indicators are free to “load on” any factor. Which indicators load on which factor is, essentially, a consequence of the EFA using the empirical correlations among the multiple indicators to seek out “clusters” of items that have the highest correlations with each other, and designating a factor to represent the underlying latent construct that is creating (is the “common cause” of) the correlations among items. Of course, designating these factors is easier said than done, given that the EFA can continue to extract additional factors, contingent on how stringent the standards are for what constitutes indicators that are “more” versus “less” correlated with each other. (For example, does a set of items with correlations of, say, .5 with each other constitute the same or a different “cluster” relative to other indicators that correlate .4 with the first cluster and .5 with each other? Extracting more factors will tend to separate these clusters into different factors; extracting fewer factors is more likely to combine the clusters into one factor.) The issue of the number of factors to extract is, then, a major dilemma in using EFA, with a number of criteria available, including the preferred method, where possible, of specifying a priori the number of factors to extract (cf. Floyd and Widaman 1995).

A potential weakness with EFA is that it works best with relatively “simple” measurement models. EFA falters in situations where the reasons for the “clustering” of indicators (i.e., interitem correlations) stem from complex sources (see Bollen 1989 for an example). As an aspect of this inability to handle complex measurement models, EFA cannot match CFA in the sophistication of its tests for validity of measures. On the other hand, because CFA is more sensitive to even slight misspecifications, it is often more difficult to obtain satisfactory model “fit” with CFA.

Exploratory and confirmatory factor analysis can complement each other (but see Floyd and Widaman 1995 for caveats). Researchers may initially develop their measurement models based on EFA, then follow that analysis with CFA (preferably on a second random subsample, in order to avoid modeling “sampling error”). In this context, the EFA serves as a crude “first cut,” which it is hoped, results in a measurement model that is close enough to the “true model” that a subsequent CFA can further refine the initial model,

using the more sophisticated procedures available with the confirmatory procedure.

Even though researchers use EFA and CFA to focus exclusively on *measurement models*, these psychometric studies are really a preamble to subsequent work in which the measures previously developed are now used to test causal linkages among constructs in *structural models*. Most of these subsequent tests of causal linkages among constructs (e.g., the effect of income on life satisfaction) use data analytic procedures, such as ordinary least squares (OLS) regression or analysis of variance (ANOVA) that assume no underlying latent structure. Accordingly, these traditional analyses of structural models work with measured (observable) variables only, can accommodate only a *single indicator* per construct of interest and, therefore, must assume (in essence) that each single indicator is perfectly reliable and valid. In other words, in the absence of a multiple indicators measurement model for each construct, there is no way to make adjustments for less-than-perfect reliability or validity.

In an attempt to enhance the reliability of the single-indicator measures that traditional data analyses procedures require, researchers often combine into a single *composite* scale the set of multiple indicators (e.g., the three indicators of life satisfaction) that prior EFA and CFA research has established as measuring a given latent construct (factor). Although, as noted earlier, a single-indicator composite scale (e.g., summing a respondent’s score on each of the three life satisfaction indicators) is generally preferable to a noncomposite single indicator (at least when measuring more subjective and abstract phenomena), a composite scale is still less than perfectly reliable (often, much less). Accordingly, using such an enhanced (more reliable) measure will still result in biased parameter estimates in the structural model.

Validity also remains a concern. Even if prior psychometric work on measures has rigorously addressed validity (though this is often not the case), there is always a question of how valid the measure of a given construct is in the context of other constructs in the current structural model (e.g., what sort of “cross-contamination” of measures may be occurring among the constructs that might bias estimates of their causal linkages).

SEM remains a viable alternative to the more traditional and widely used data analysis procedures for assessing causal effects in structural models. Indeed, SEM is so flexible in specifying the structural, as well as measurement part of a causal model, that it can literally “mimic” most of the traditional, single indicator, data analysis procedures based on the general linear model (e.g., Analysis of Variance [ANOVA], Multivariate Analysis of Variance [MANOVA], Ordinary Least Squares [OLS] regression), as well as the newer developments in these statistical techniques (e.g., Hierarchical Linear Modeling [HLM]). SEM can even incorporate the single-indicator measurement models (i.e., based on observable variables only) that these other data analysis procedures use. The power to adjust for random and nonrandom measurement error (unreliability and invalidity) is realized, however, only if the SEM includes a multiple-indicator (latent variable) measurement model.

A multiple-indicator measurement model also provides SEM procedures with a particularly powerful tool for modeling (and thus adjusting for) correlated measurement errors that inevitably occur in analyses of *longitudinal* data (cf. Bollen 1989). Likewise, multiple-indicator models allow SEM programs to combine latent growth curves and multiple subgroups of cohorts in a “cohort-sequential” design through which researchers can examine potential cohort effects and developmental sequences over longer time frames than the period of data collection in the study. By also forming subgroups based on their pattern of attrition, (i.e., at what wave of data collection do respondents drop out of the study), SEM researchers can analyze the “extended” growth curves while adjusting for the potential bias from nonrandom missing data (Duncan and Duncan 1995)—a combination of important features that non-SEM methods cannot match. Furthermore, although SEM is rarely used with *experimental* research designs, multiple-indicator models provide exceptionally powerful methods of incorporating “manipulation checks” (i.e., whether the experimental treatment manipulates the independent variable that was intended) and various “experimental artifacts” (e.g., “demand characteristics”) as alternative causal pathways to the outcome variable of interest (cf. Blalock 1985).

Although this article has focused on strengths of the SEM multiple indicator measurement model, the advantages of SEM extend beyond its ability to model random and nonrandom measurement error. Particularly noteworthy in this regard is SEM’s power to assess both contemporaneous and lagged *reciprocal* effects using nonexperimental longitudinal data. Likewise, SEM has an especially elegant way of incorporating maximum likelihood full-information imputation procedures for handling missing data (e.g., Arbuckle 1996; Schafer and Olsen 1998), as an alternative to conventional listwise, pairwise, and mean substitution options. Simulation studies show that the full-information procedures for missing data, relative to standard methods, reduce both the bias and inefficiency that can otherwise occur in estimating factor loadings in measurement models and causal paths in structural models.

So given all the apparent advantages of SEM, why might researchers hesitate to use this technique? As noted earlier in discussing limitations of confirmatory factor analysis, SEM requires that a field of study be mature enough in its development so that researchers can build the a priori causal models that SEM demands as input. SEM also requires relatively large sample sizes to work properly. Although recommendations vary, a minimum sample size of 100 to 200 would appear necessary in order to have a reasonable chance of getting solutions that “converge” or that provide parameter (path) estimates that are not “out of range” (e.g., a negative variance estimate) or otherwise implausible. Experts often also suggest that the “power” to obtain stable parameter estimates (i.e. the ability to detect path coefficients that differ from zero) requires the ratio of subjects-to-parameters-estimated be between 5:1 and 20:1 (the minimum ratio increasing as variables become less normally distributed). In other words, more complex models demand even larger sample sizes.

Apart from the issue of needed a larger  $N$  to accommodate more complex causal models, there also is some agreement among SEM experts that models should be kept simpler in order to have a reasonable chance to get the models to fit well (cf., Floyd and Widaman 1995). In other words, SEM works better where the total number of empirical indicators in a model is not extremely large (say 30 or less). Accordingly, if researchers wish to test

causal models with many predictors (latent constructs) or wish to model fewer constructs with many indicators per construct, SEM may not be the most viable option. Alternatively, one might consider “trimming” the proposed model of constructs and indicators through more “exploratory” SEM procedures of testing and respecifying models, possibly buttressed by preliminary model trimming runs with statistical techniques that can accommodate a larger number of predictors (e.g., OLS regression), and with the final model hopefully tested (confirmed) on a second random subsample. Combining pairs or larger “parcels” from a common set of multiple indicators is another option for reducing the number of total indicators, and has the added benefit of providing measures that have more normal distributions and fewer correlated errors from idiosyncratic “content overlap.” (One needs to be careful, however, of eliminating too many indicators, given that MacCallum et al. [1996] have shown that the power to detect good-fitting models is substantially reduced as degrees of freedom decline.)

An additional limitation on using SEM has been the complex set of procedures (based on matrix algebra) one has had to go through to input the model to be tested—accompanied by output that was also less than user-friendly. Fortunately, newer versions of several SEM programs can now use causal diagrams for both input and output of the structural and measurement models. Indeed, these simple diagramming options threaten to surpass the user-friendliness of the more popular mainstream software packages (e.g., Statistical Package for the Social Sciences [SPSS]) that implement the more traditional statistical procedures. This statement is not meant to be sanguine, however, about the knowledge and experience required, and the care one must exercise, in using SEM packages. In this regard, Kline (1998) has a particularly excellent summary of “how to fool yourself with SEM.” The reader is also encouraged to read Joreskog (1993) on how to use SEM to build more complex models from simpler models.

As the preceding discussion implies, using multiple-indicator models requires more thought and more complicated procedures than does using more common data analytic procedures. However, given the serious distortions that measurement errors can produce in estimating the true causal links among concepts, the extra effort in

using multiple-indicator models can pay large dividends.

REFERENCES

- Arbuckle, J. L. 1996 “Full Information Estimation in the Presence of Incomplete Data. In G. Marcoulides and R. Schumacker eds., *Advanced Structural Equation Modeling: Issues and Techniques*. Mahwah, N. J.: Lawrence Erlbaum.
- 1997 *Amos User’s Guide Version 3.6*. Chicago: SPSS.
- Austin, J., and R. Calderon 1996 “Theoretical and Technical Contributions to Structural Equation Modeling: An Updated Annotated Bibliography.” *Structural Equation Modeling* 3:105–175.
- , and L. Wolfe 1991 “Annotated Bibliography of Structural Equation Modeling: Technical Work.” *British Journal of Mathematical and Statistical Psychology* 44:93–152.
- Bentler, P.M. 1995. *EQS: Structural Equations Program Manual*. Encino, Calif.: Multivariate Software.
- Blalock, H. M. 1979 *Social Statistics*, 2nd ed. New York: McGraw-Hill.
- , ed. 1985 *Causal Model in Panel and Experimental Designs*. New York: Aldine.
- Bollen, K. A. 1989 *Structural Equations with Latent Variables*. New York: John Wiley.
- Diener, E., R. A. Emmons, R. J. Larsen, and S. Griffin 1985 “The Satisfaction with Life Scale.” *Journal of Personality Assessment* 49:71–75.
- Duncan, T. E., and S. C. Duncan 1995 “Modeling the Processes of Development via Latent Variable Growth Curve Methodology.” *Structural Equation Modeling* 2:187–213.
- Floyd, F. J., and K. F. Widaman 1995 “Factor Analysis in the Development and Refinement of Clinical Assessment Instruments.” *Psychological Assessment* 7:286–299.
- Haring, M. J., W. A. Stock, and M. A. Okun 1984 “A Research Synthesis of Gender and Social Class as Correlates of Subjective Well-Being.” *Human Relations* 37: 645–657.
- Joreskog, K. G. 1993 “Testing Structural Equation Models.” In K. A. Bollen and J. S. Long, eds., *Testing Structural Equation Models*. Thousand Oaks, Calif.: Sage.
- , and D. Sorbom 1993 *Lisrel 8: User’s Reference Guide*. Chicago: Scientific Software.
- Kim, J., and C. W. Mueller 1978 *Introduction to Factor Analysis: What It Is and How to Do It*. Beverly Hills, Calif.: Sage.
- Kline, R. B. 1998 *Principles and Practice of Structural Equation Modeling*. New York: Guilford.

- Larson, R. 1978 "Thirty Years of Research on the Subjective Well-Being of Older Americans." *Journal of Gerontology* 33:109–125.
- Loehlin, J. C. 1998 *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*, 3rd ed. Hillsdale, N.J.: Lawrence Erlbaum.
- Marcoulides, G. A., and R. E. Schumacker, eds. 1996 *Advanced Structural Equation Modeling: Issues and Techniques*. Mahwah, N.J.: Lawrence Erlbaum.
- Muthen, B. O. 1993 "Goodness of Fit with Categorical and Other Nonnormal Variables." In K. A. Bollen and J. S. Long, eds., *Testing Structural Equation Models*. Thousand Oaks, Calif.: Sage.
- 1998 *Mplus User's Guide: The Comprehensive Modeling Program for Applied Researchers*. Los Angeles, Calif.: Muthen and Muthen.
- Nunnally, J. C., and I. H. Bernstein 1994 *Psychometric Theory*, 3rd ed. New York: McGraw-Hill.
- Randall, R. E., and G. A. Marcoulides (eds.) 1998 *Interaction and Nonlinear Effects in Structural Equation Modeling*. Mahwah, N. J.: Lawrence Erlbaum.
- Reise, S. P., K. F. Widaman, and R. H. Pugh 1993 "Confirmatory Factor Analysis and Item Response Theory: Two Approaches for Exploring Measurement Invariance." *Psychological Bulletin* 114:552–566.
- Schafer, J. L., and M. K. Olsen 1998 "Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective." Unpublished manuscript, Pennsylvania State University. <http://www.stat.psu.edu/~jls/index.html/#res>.
- Sullivan, J. L., and S. Feldman 1979 *Multiple Indicators: An Introduction*. Beverly Hills, Calif.: Sage.
- Waller, N. G., A. Tellegen, R. P. McDonald, and D. T. Lykken 1996 "Exploring Nonlinear Models in Personality Assessment: Development and Preliminary Validation of a Negative Emotionality Scale." *Journal of Personality* 64:545–576.

KYLE KERCHER

## MUSIC

From the earliest days of the discipline, music has been a focus of sociological inquiry. Max Weber, for example, used the development of the system of musical notation as a prime illustration of what he saw as the increasing rationalization of European society since the Middle Ages (Max Weber 1958). Like Weber, many since have used music

and music making as a strategic research site for answering sociologically important questions. Nevertheless, music has not become the focus of a distinctive fundamental approach in sociology comparable to topics like "socialization," "organization," "deviance," and "culture." That is why the subject of this entry is the "sociology of music" and not "musical sociology" and why it takes such a long bibliography to suggest the range of work in the field.

While no musical sociology has developed, over the decades numerous aspects of music making or appreciation have been the substantive research site for addressing central questions in sociology. Broadly, these can be grouped together as six ongoing research concerns, and together they can be said to constitute the scattered but rich sociology of music. These six focuses will be considered in turn.

### MACROSOCIOLOGY

Many sociologists have been concerned with the relationship between society and culture. Talcott Parsons, among others, believed that there was a close fit between the two, so that it would be possible to compare societies and changes taking place in society by studying elements of their cultures including music. This is what led Max Weber (1958) to focus on music as an example of the rationalization of Western society and led another early sociologist, Georg Simmel ([1882] 1968), to see music as a mirror of socio-psychological processes. The most ambitious cross-national attempt to directly link the structure of society with musical expression has been that of Lomax (1968).

Contemporary scholars see the links between social structure and music not as natural but as deliberately constructed, and they focus on the circumstances of that process. For example, Cerulo (1999) has found an association between the circumstances of a nation's founding and its national anthem. Whisnant (1983) examined the politics of culture involved in constructing the idea of Appalachian folk culture, and Cloonan (1997) examines how "Englishness" is constructed in contemporary British pop music.

Another line of contemporary studies exploring the society-music link focuses on the globalization

of music culture. The specific focuses and methods of these researchers vary widely, but together they show the resilience of local forms of musical expression in the context of the globalization of the media. See, for example, Wallis and Malm (1984), Nettl (1985, 1998), Hebdige (1987), Guilbault (1993), and Manuel (1993).

### THE SOCIAL CONSTRUCTION OF AESTHETICS

In line with the constructivist orientation just mentioned, a number of scholars have examined the processes by which aesthetic standards are set and changed. The seminal works in this line are by Meyer (1956) and Becker (1982). Hennion (1993) studies the evolution of standards in the Paris music world. Monson (1996) carefully examines the conventions of jazz improvisation. Ellison (1995) shows the close interplay between country music artists and their fans in defining the country music experience, and Frith (1996) examines many of the same issues in the world of rock music.

Alternative kinds of aesthetic standards have been contrasted. The most widely compared are "fine art," "folk," and "popular" aesthetics (Gans 1974). Using the example of jazz, Peterson (1972) shows that music can evolve from folk to pop and then to fine art. Frith (1996) perceptively shows that all three standards are used simultaneously in the rock world and elsewhere.

### THE STUDY OF MUSIC WORLDS

In his early studies of interaction among dance-hall musicians, Becker pioneered the study of the conventions that musicians use in making their kind of music together (summarized in Becker 1982). This line of activity has been continued in musical scenes varying from a group of coldly calculating recording session musicians (Peterson and White 1979; Faulkner 1983) to an impassioned gay community (DeChaine 1997). See also Zolberg (1980), Etkorn (1982), Gilmore (1987), MacLeod (1993), Rose (1994), Monson (1996), Devereaux (1997), and Nettl (1998).

A related and often overlapping line of studies focuses on the socialization of performers. See, for example, Bennett (1980), Kingsbury (1988), Freeman (1996), DeChaine (1997), and Levine and

Levine (1996). The socialization of audiences has also been the focus of considerable attention. The focus of inquiry has included popular music fans (Hebdige 1979, Frith 1981, 1996), the Viennese supporters of Ludwig van Beethoven (DeNora 1995), country music fans (Ellison 1995), and hip-hop fans (Rose 1995).

### THE INSTITUTIONALIZATION OF MUSIC FIELDS

While the art-world perspective focuses on the interaction of individuals and groups in making and appreciating music, research on institutionalization has to do with the creation of the organizations and the infrastructure that support distinctive music fields. DiMaggio (1982), for example, shows the entrepreneurial processes by which the classical music field was established in the United States. Menger (1983) and Hennion (1993) show the structure of the contemporary art music fields of France. Peterson (1990) and Ennis (1992) trace the reconfiguration of the pop music field into rock in the 1940s and 1950s, and Peterson (1997) traces the institutionalization of the country music field and its articulation with the larger field of commercial music in the mid-1950s.

Based on the production of culture perspective, some researchers have shown how elements of the production process, including law, technology, industry structure, industrial organization, careers, and market structures, profoundly shape the sort of music that is produced. The interplay between these processes is traced by Peterson (1990) to show why rock music emerged in the mid-1950s. For other examples of work in the production of culture perspective, see Keil (1966), Ryan (1985), Lopes (1992), Jones (1992), Frith (1993), Manuel (1993), Freeman (1996), and Levine and Levine (1996).

A number of studies have traced the development of specific music organizations and organization fields. Ahlquist (1997) shows the structure of the early-nineteenth-century entrepreneurial opera world, and Martorella (1982) shows the remarkably different field of opera that emerged following the American Civil War. Arian (1971) and Hart (1975) focus on the classical music orchestra. Peterson and White (1979) and Faulkner (1983) discuss the organization of the apparently

purely competitive world of recording studio session musicians. Gilmore (1987) focuses on three distinct contemporary classical music fields in New York City. MacLeod traces the field of club-date musicians in the same city. Guilbault (1993) shows the shaping of World Music in the West Indies. Negus (1999) compares the very different methods the major music industry corporations use in exploiting country music, rap, and reggae.

One specific line of studies focuses on the "concentration-diversity" hypothesis. Peterson and Berger (1975) show that periods of competition in the popular music field lead to a great diversity in the music that becomes popular, while concentration of the industry in a few firms leads to homogeneity in popular music. Using later data, Burnett (1990) challenges this finding, and Lopes (1992) shows how large multinational firms are able to control the industry while at the same time producing a wide array of musical styles.

### MUSIC FOR MAKING DISTINCTIONS

All known societies employ music as a means of expressing identity and marking boundaries between groups (Lomax 1968), and a number of authors have shown the place of music in social class and status displays. DeNora (1995) shows the role that Beethoven's music played in the Vienna of 1800 by marking off the rising business aristocracy from the older court-based aristocracy that championed the music of Haydn and Mozart. DiMaggio (1982) shows how the rising commercial elite of nineteenth-century Boston used support for classical music as a status marker, and recent survey research studies show how music preferences have been used in making social class distinctions. See, for example, Bryson (1996) and Peterson and Kern (1996).

Numerous studies have explored issues of racism in the meanings ascribed to music and its creators. Reidiger (1991), for example, shows how immigrant Irish, Germans, Jews, Swedes, and the like developed the "black-faced" minstrel show to satirize African-Americans and show their own affinity with the dominant English-speaking class as part of what they came to call the "white race." Leonard (1962) and Peterson (1972) show the evolving attitudes of whites toward jazz and African-American culture. The most extensive line of work on race focuses on how jazz was created and

interpreted by African-American intellectuals as a form of resistance to the dominant white culture. See especially Jones (1963), Kofsky (1970), Vincent (1995), Devereaux (1997), and Panish (1997). The racial shaping of blues, gospel, and rap have been explored, respectively, by Keil (1966), Heilbut (1997), and Rose (1994).

Music is gendered in many ways. Men are more likely to be producers, women consumers; women are often demeaned in the lyrics of rock, rap, and heavy metal, but they are characterized as strong in country music lyrics. Yet, compared to other topics, there has not been much research on gender in music and music making. Four different books suggest the range of topics that could be explored: McClary (1991) on the gendering of opera in which strong female leads do not succeed; Buffwack and Oermann (1993) on the decades of struggle of female country music entertainers for equality with men; Lewis (1990) on the gendered nature of rock videos; and Nehring (1997) on gender issues in pop music.

At least since the "jazz age" of the 1920s, music has been one of the primary ways in which generations are defined and define themselves. In the 1990s, for example, one slogan of the young was "If it's 'too loud,' you're too old." Hebdige (1979), Frith (1981), Laing (1985), Liew (1993), Lahusen 1993, Epstein (1994), Weinstein (1994), Shevory (1995), and DeChaine (1997), among many others, show how specific contemporary youth groups use music to state distinctions not only between themselves and their parents but between themselves and other youth groups as well.

### THE EFFECTS OF MUSIC

As just noted, new musical forms are identified with each new generation, and adults have tended to equate the music of young people with youthful deviance. Jazz, swing, rock, punk, disco, heavy metal, and rap, in turn, have been seen as the causes of juvenile delinquency, drug taking, and overt sexuality. See, for example, Leonard (1962), Hebdige (1977), and Laing (1985). Jazz was under attack for most of the first half of the twentieth century, and more recently heavy metal music has come in for the most sustained attention (Raschke 1990; Walser 1993; Weinstein 1994; Binder 1993; and Arnett 1996).

One of the most persistent research topics involving music has been the content analysis of song lyrics, on the often unwarranted assumption that one can tell the meaning a song has for its fans by simply interrogating lyrics (Frith 1996). Content analyses can be divided roughly into two types, holistic and systematic.

Systematic content analyses identify a universe of songs, draw a systematic sample from the universe, and analyze the lyrics in terms of a standard set of categories, typically using statistics to represent their results. Examples of systematic content analysis include Horton (1957), Carey (1969), Rogers (1989), and Ryan et al. (1996). Most such content analyses have focused on popular music, but one of the most remarkable is John H. Muller's study of the changing repertoire of classical music over the first half of the twentieth century (1951)—a study that clearly deserves updating. Kate Muller (1973) makes a start. Holistic content analyses typically involve selecting songs that fit the research interest of their author and making judgments of the meaning of a song, or groups of songs, as a whole (Lewis 1990; McLurin 1992; Lahusen 1993; Ritzel 1998).

Finally the "effects" focus in social movements research returns the sociological study of music full circle to more macro-sociological concerns. Music has been an integral part of most social movements, as has been shown by a number of studies. For example, Denisoff and Peterson (1972) provide a collection of essays that describe the place of music in movements ranging from the International Workers of the World, Nazi youth groups, and folk protest music to avant-garde jazz, country music, rock, and religious fundamentalism. Denisoff (1971) has made an analysis of the use of music in the political protest movement of the 1960s and 1970s, and Neil Rosenberg (1993) has assembled a set of essays on the politicization of folk music in that same era. Treece (1997) describes the role of bossa nova in Brazil's popular protest movement, and Eyerman and Jamison (1998) examine the place of music in more recent social movements.

Something approaching one-half the works cited here are by scholars who research music topics relevant to sociology, but who are not themselves sociologists. This openness and diversity is

important to the continuing vitality of the sociology of music. It is to be hoped that, in forthcoming decades, self-conscious efforts to bring more order to the scatter will facilitate the systematic articulation of research questions, stimulate research on more diverse music worlds, and encourage more cumulation of research findings. These efforts may also show the ways in which music can become a fundamental perspective in sociology similar to socialization, organization, deviance, culture, and the like.

## REFERENCES

- Ahlquist, Karen 1997 *Democracy in the Opera: Music, Theater, and Culture in New York City, 1915-69*. Urbana: University of Illinois Press.
- Arian, Edward 1971 *Bach, Beethoven and Bureaucracy*. University: University of Alabama Press.
- Arnett, Jeffrey 1996 *Metal Heads: Heavy Metal Music and Adolescent Alienation*. Boulder, Colo: Westview.
- Atali, Jacques 1985 *Noise: The Political Economy of Sound*. Minneapolis: University of Minnesota Press.
- Becker, Howard S. 1982 *Art Worlds*. Berkeley: University of California Press.
- Bennett, Stit 1980 *Becoming a Rock Musician*. Amherst: University of Massachusetts Press.
- Binder, Amy 1993 "Media Depictions of Harm in Heavy Metal and Rap Music." *American Sociological Review* 58:753-767.
- Bryson, Bethany 1996 "'Anything but Heavy Metal': Symbolic Exclusion and Musical Dislikes." *American Sociological Review* 61:884-899.
- Bufwack, Mary A., and Robert K. Oermann 1993 *Finding Her Voice: The Saga of Women in Country Music*. New York: Crown.
- Burnett, Robert 1990 *Concentration and Diversity in the International Phonogram Industry*. Gothenburg, Sweden: University of Gothenburg Press.
- Carey, James T. 1969 "Changing Courtship Patterns in Popular Song." *American Journal of Sociology* 74:720-731.
- Cerulo, Karen 1999 *Identity Designs: The Sights and Sounds of a Nation*. New Brunswick, N.H.: University Press of New England.
- Cloonan, Martin 1997 "State of the Nation: 'Englishness,' Pop and Politics in the Mid-1990s." *Popular Music and Society* 21(2):47-70.
- DeChaine, D. Robert 1997 "Mapping Subversion: Queercore Music's Playful Discourse of Resistance." *Popular Music and Society* 21(1):7-38.

- Denisoff, R. Serge 1971 *Great Day Coming: Folk Music and the American Left*. Urbana: University of Illinois Press.
- , and Richard A. Peterson 1972 *The Sounds of Social Change*. Chicago: Rand McNally.
- DeNora, Tia 1995 *beethoven and the construction of Genius: Musical Politics in Vienna, 1792–1803*. Berkeley: University of California Press.
- Devereaux, Scott 1997 *The Birth of Bebop: A Social and Musical History*. Berkeley: University of California Press.
- DiMaggio, Paul J. 1982 “Cultural Entrepreneurship in Nineteenth Century Boston, Part I: The Creation of an Organizational Base for High Culture in America.” *Media, Culture and Society* 4:33–50.
- Ellison, Curtis W. 1995 *Country Music Culture*. Jackson: University Press of Mississippi.
- Ennis, Philip 1992 *The Seventh Stream*. Hanover, N.H.: University Press of New England.
- Epstein, Jonathon S. (ed.) 1994 *Adolescents and Their Music: If It's Too Loud, You're Too Old*. New York: Garland.
- Etzkorn, K. Peter 1982 “On the Sociology of Musical Practice and Social Groups.” *International Social Science Journal* 34:555–569.
- Eyerman, Ron, and Andrew Jamison 1998 *Music and Social Movements*. Cambridge, England: University of Cambridge Press.
- Faulkner, Robert 1983 *Music on Demand*. New Brunswick, N.J.: Transaction Books.
- Freeman, Robert 1996 “On the Future of America's Orchestras.” *Harmony* 3:11–22.
- Frith, Simon 1981 *Sound Effects: Youth Leisure and the Politics of Rock 'N' Roll*. New York: Pantheon.
- 1996 *Performing Rites*. Cambridge, Mass.: Harvard University Press.
- Frith, Simon (ed.) 1993 *Music and Copyright*. Edinburgh: University of Edinburgh Press.
- Gans, Herbert J. 1974 *Popular Culture and High Culture*. New York: Basic Books.
- Gilmore, Samuel 1987 “Coordination and Convention: The Organization of the Concert World.” *Symbolic Interaction*. 10:209–227.
- 1988 “Schools of Activity and Innovation.” *Sociological Quarterly* 29:203–219.
- 1990 “Tradition and Novelty in Concert Programming: Bringing the Artist Back In.” *Sociological Forum* 8:221–242.
- Guilbault, Jocelyne 1993 *Zouk: World Music in the West Indies*. Chicago: University of Chicago Press.
- Hart, Philip 1975 *Orpheus in the New World: The Symphony Orchestra as an American Cultural Institution*. New York: W.W. Norton.
- Hebdige, Dick 1979 *Subcultures: The Meaning of Style*. London: Methuen.
- 1987 *Cut and Mix: Cultural Identity and Caribbean Music*. London: Methuen.
- Heilbut, Anthony 1997 *The Gospel Sound: Good News and Bad Times*. New York: Limehouse Editions.
- hennion, Antoine 1993 *La Passion Musicale: Une Sociologie de la Mediation*. Paris: Edition Metailie.
- Horton, Donald 1957 “The Dialogue of courtship in American Popular Songs.” *American Journal of Sociology* 62:569–578.
- Jones, Leroy 1963 *Blues People: The Negro Experience in America and the Music That Developed from It*. New York: William Murrow.
- Jones, Steve 1992 *rock Formation: Music, Technology, and Mass Communication*. Newbury Park, Calif.: Sage.
- Keil, Charles 1966 *Urban Blues*. Chicago: University of Chicago Press.
- Kingsbury, Henry A. 1988 *Talent and Performance: A Conservatory Culture System*. Philadelphia: Temple University Press.
- Kofsky, Frank 1970 *Black Nationalism and the Revolution in Music*. New York: Pathfinder.
- Lahusen, Christian 1993 “The Aesthetics of Radicalism: The Relationship between Punk and the Patriotic Nationalist Movement of the Basque Country.” *Popular Music* 12:263–280.
- Laing, Dave 1985 *One Chord Wonders: Power and Meaning in Punk Rock*. Milton Keynes, England: Open University Press.
- Leonard, Neil 1962 *Jazz and the White Americans*. Chicago: University of Chicago Press.
- Levine, Seymour, and Robert Levine 1996 “Stress and Discontent in the the Orchestra Workplace.” *Harmony* 2:15–26.
- Lewis, Lisa A. 1990 *Gender Politics and MTV*. Philadelphia: Temple University Press.
- Liew, Maria van 1993 “The Scent of Catalan Rock: Els Pets' Ideology and the Rock and Roll Industry.” *Popular Music* 12:245–261.
- Lomax, John 1968 *Folk Song Style and Culture*. Washington, D.C.: American Association for the Advancement of Science.



- Lopes, Paul 1992 "Innovation and Diversity in the Popular Music Industry." *American Sociological Review* 57:56-71.
- McClary, Susan 1991 *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: University of Minnesota Press.
- McLaurin, Melton A. 1992 "The changing Image of the South in Country Music." In Melton A. McLaurin and Richard A. Peterson, eds., *You Wrote My Life: Lyrical Themes in Country Music*. Philadelphia: Gordon and Beach.
- MacLeod, Bruce A. 1993 *Club Date Musicians: Playing the New York Party Circuit*. Urbana: University of Illinois Press.
- McPhee, William 1963 "When Music becomes a Business." In Joseph Berger, Maurice Zeitlin, and B. Anderson, eds., *When Music becomes a business*. Boston: Houghton-Mifflin.
- Menger, Pierre-Michel 1983 *Le Paradoxe du Musicien: Le Compositeur, le Mélomane et l'État dans la Société Contemporaine*. Paris: Harmoniques Flammarion.
- Meyer, Leonard B. 1956 *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
- Middleton, Richard 1990 *Studying Popular Music*. Milton Keynes, England: Open University Press.
- Monson, Ingrid 1996 *Saying Something: Jazz Improvisation and Interaction*. Chicago: University of Chicago Press.
- Muller, John H. 1951 *The American Symphony Orchestra: A Social history of Taste*. Bloomington: Indiana University Press.
- Muller, Kate 1973 *Twenty Seven Major American Symphony Orchestras*. Bloomington: University of Indiana.
- Negus, Keith 1993 *Producing Pop: Culture and Conflict in the Popular Music Industry*. London and New York: E. Arnold.
- 1997 *Popular Music in Theory: An Introduction*. Hanover, N.H.: University Press of New England.
- 1999 *Music Genres and Corporate Cultures*. London: Routledge.
- Nehring, Neil 1997 *Popular Music, Gender, and Postmodernism: Anger Is an Energy*. Thousand Oaks, Calif.: Sage.
- Nettl, Bruno 1985 *The Western Impact on World Music*. New York: Schirmer.
- (ed.) 1998 *In the Course of Performance: Studies in the World of Musical Improvisation*. Chicago: University of Chicago Press.
- Panish, Jon 1997 *The Color of Jazz: Race and Representation in Postwar American Culture*. Jackson: University Press of Mississippi.
- Peterson, Richard A. 1972 "A Process Model of the Folk, Pop, and Fine Art Stages of Jazz." In Charles Nanry, ed., *American Music: From Storyville to Woodstock*. New Brunswick, N.J.: Rutgers University Press.
- 1990 "Why 1955? Explaining the Advent of Popular Music." *Popular Music* 9:97-116.
- 1997 *Creating Country Music: Fabricating Authenticity*. Chicago: University of Chicago Press.
- , and David G. Berger 1975 "Cycles in Symbol Production: The Case of Popular Music." *American Sociological Review* 40(2):158-173.
- , and Roger Kern 1996 "Changing Highbrow Taste: From Snob to Omnivore." *American Sociological Review* 61:900-907.
- , and Howard White 1979 "The Simplex Located in Art Worlds." *Urban Life* 7:411-439.
- Raschke, Carl A. 1990 *Painted Black: From Drug Killings to heavy Metal: The Alarming True Story of how Satanism Is Terrorizing Our Communities*. San Francisco: Harper and Row.
- Reidiger, David R. 1991 *The Wages of Whiteness: Race and the Making of the American Working Class*. London: Verso.
- Ritzel, Fred 1998 "Was ist aus uns geworden?—Ein Haufchen Sand am Meer: Emotions of Post War Germany as Extracted from Examples of popular Music." *Popular Music* 17:293-309.
- Rogers, Jimmie N. 1989 *The Country Music Message*. Fayetteville: University of Arkansas Press.
- Rose, Tricia 1994 *Black Noise: Rap Music and Black Culture in Contemporary America*. Hanover, N.H.: University Press of New England.
- Rosenberg, Neil V. 1985 *Bluegrass: A History*. Urbana: University of Illinois Press.
- , ed. 1993 *Transforming Tradition: Folk Music Revival Examined*. Urbana: University of Illinois Press.
- Ryan, John 1985 *The Production of Culture in the Music Industry*. Lanham, Md.: University Press of America.
- , Legarde H. Calhoun, and William M. Wentworth 1996 "Gender or Genre: Emotion Models in Rap and Country Music." *Popular Music and Society* 20(2):121-154.
- Shevory, Thomas C. 1995 "bleached Resistance: the Politics of Grunge." *Popular Music and Society* 19(2):23-48.
- Simmel, Georg (1882) 1968 *the Conflict in Modern Culture*. New York: Teachers College Press.
- Treece, Davide 1997 "Bosa Nova and Brazil's Music of Popular Protest." *Popular Music* 16:1-30.
- Vincent, Ted 1995 *Keeping Cool: The Black Activists Who Built the Jazz Age*. London: Pluto.

- 
- Wallis, Roger, and Krister Malm 1984 *Big Sounds from Small Peoples: The Music Industry in Small Countries*. Londn: Constable.
- Walser, Robert 1993 *Running with the Devil: Power, Gender, and Madness in Heavy Metal Music*. Hanover, N.H.: University Press of New England.
- Weber, Max 1958 *The Rational and Social Foundations of Music*. Carbondale: Southern Illinois University Press.
- Weber, William 1986 "The Rise of Classical Repertoire in Nineteenth-Century Orchestral concerts." In Joan Peyser, ed., *The Orchestra*. New York: Scribners.
- Weinstein, Deena 1994 *Heavy Metal: A Cultural Sociology*. New York: Lexington.
- Whisnant, David E. 1983 *All That Is Native and Fine: The Politics of Culture and an American Region*. Chapel Hill: University of North Carolina Press.
- Zolberg, Vera 1980 "Displaced Art and Performed Music: selective Innovation and the Strucure of Artistic Media." *Sociological Quarterly* 21:219-231.
- 1990 *Constructing a Sociology of the Arts*. Cambridge, England: Cambridge University Press.

RICHARD A. PETERSON

# N

---

## NATIONAL BORDER RELATIONS

Are the processes commonly described as globalization presenting us with a borderless future? Are borders, the limits of state authority, traditionally the instrument by which states and nations define themselves, changing beyond recognition at the threshold of the twenty-first century? Are borders, in the classical sense of the age of the nationstate, being abolished, as part of the erosion of the nation-state itself? Are they withering, or are they undergoing functional change? Are there regional differences between the developments in Europe (EU), America (NAFTA), Asia, and Africa?

A full answer to such very practical questions would perhaps benefit from a grounding in a “general theory of boundaries” conceived as an integral part of the “general theory of systems” (Strassoldo 1976–77, 1979, 1982). States and nations are just a genus of social (societal) systems, and it can be argued that social systems are just a specific kind of (general) systems. For instance, the study of boundary (or interface, as they are sometimes also called) processes, osmotic processes taking place through cell membranes, is one of the frontiers of biological research. In fact, the boundary became a basic concept in system theories in the 1960s and 1970s (Miller 1977). The boundary has an important place in several other disciplines, beginning perhaps with Gestalt psychology (boundary as the line that activates *Gestalten*) and philosophical anthropology, according to which

the drawing of neat boundary lines—around concepts, in the definition of categories, in the classification of phenomena, in staking out domains and setting rules—is one of the hallmarks of formal, rational discourse (Bateson 1972). That most elusive concept, form, can best be defined as a “marked space” (Brown 1969). We can only conceive objects, that is, chunks of reality that have a form, a line around them: “Epistemology is about where you draw the line” (Wilden 1972).

Restricting our view to sociological system theories, we find the concept of boundary (“boundary maintenance,” “boundary articulation,” etc.) recurring throughout Parsons’s voluminous works. He seems to have borrowed the idea from earlier anthropologists; and it was another anthropologist, Barth, who has given it renewed popularity in more recent times (Barth 1969).

The logical and ontological primacy of boundaries over any other system element has been emphasized by Luhmann, arguably the sociologists who has most systematically grappled with this problem. For him, any system first emerges as a difference—that is, a boundary—between the inside and the outside, the system and the environment. Later, a difference between a center and a periphery emerges within the system (Luhmann 1982).

It must be admitted, however, that run-of-the-mill sociology has not paid much attention to the concept and problem of boundaries, although

every time one meets such dichotomies and expressions as “in-out,” “internal-external,” “backstage-frontstage,” “cross-cutting,” “center and periphery,” “marginality,” “stranger,” “distinction,” “identity,” “closure,” and so on, a boundary is implied. Several hypotheses have been suggested for this widespread blindness of sociologists to such a pervasive reality. One is the necessity, typical of all sciences, to isolate phenomena in order to make them manageable for study; this leads researchers to focus on the “core” of social objects, dismissing their margins as though they just “die out” in a surrounding vacuum. Alternatively, it may well be that it is just the inevitable interconnectedness of social phenomena, the generally overlapping, uncertain, and fuzzy character of empirical social boundaries that has kept sociologists clear of them.

On the other hand, it is always possible to find forerunners and prophets for every sociological concept and theory. So one can be referred to Simmel’s fascinating musings on the “criss-crossing of social circles,” on the “door,” the “bridge,” and the “handle” as different means of overcoming the boundary-line between different domains. In fact, his whole “formal” approach to sociology is built on the notion that social life is a system of “frames”—that is, boundaries—that are an essential part of its dynamics. This notion, of course, has fed many other, later theorists, from Schutz to Goffman. In the case of territorial borders, one can refer to what Sorokin (1928) called the “geographic school” of nineteenth-century sociology (Le Play, Des Moulins) and to Durkheim’s claim that “frontiers” are a central feature of “social morphology.” The fullest early sociological treatise on borders is by a follower of Durkheim, the Belgian sociologist and social reformer De Greef, who developed a complex and suggestive theory of an evolutionary interaction between national, political, and territorial (“horizontal”) borders on one side, and social (“vertical”) boundaries (between groups, classes, organisations etc.) on the other (De Greef, 1908).

Borders—or frontiers, or boundaries—can be studied in a variety of ways: as limits of state sovereignty or as limits of administrative units within states, but also as cultural markers—markers of ethnicity, group or individual identity, and as form. Here, we are predominantly interested in the political borders between nations and states

rather than substate boundaries or cultural and anthropological concepts of boundaries (Cohen 1986). Borders, thus, are products of human attempts at organizing territories. They may change location as well as function. “Frontiers are inseparable from the entities which they enclose” (M. Anderson 1996, p. 178); they are “a geographical instrument . . . for the organization of space” (Guichonnet and Raffestin 1974, p. 9). Political borders are human constructs, not natural givens. “Frontiers between states are institutions and processes.” (M. Anderson 1996, p. 1).

Borders are institutionalized in legal texts and international agreements, and as such are the expression of political will and social organization. They mark the limits of political decision making, the limits of a legal space, the limits within which state identities and rights and duties of citizenship operate. Looking at borders as processes, Malcolm Anderson defines four dimensions (M. Anderson 1996, p. 2): (1) They are instruments of state policy, protecting and promoting interests; (2) the de facto control states exercise over their borders is indicative of the nature of the state; (3) frontiers are markers of identities of “imagined communities” (B. Anderson 1983); and (4) the “frontier” is a term of discourse, affecting “not only the physical flow of goods and persons, which can be measured; much more important, they affect the culture and consciousness of people, which is much more difficult to assess” (Strassoldo 1998, p. 87).

Historically, the development of the idea of the border has been closely linked to the idea of the development of the state (Breuille 1993). As the idea of the state changed, so did the functions of borders. Ancient empires, like the Greek or the Roman, had clearly defined boundaries within them—to define citizenship rights and duties—but the outer limits were ‘fuzzy’. The claim was that the limits of the empire were the limits of the (civilized) world; the empire could not be bounded. In practice, the Romans developed the idea of the *limes*, a frontier line well within the boundaries within which Roman authority ruled, with a zone of influence beyond, which would act as a buffer against the “barbarians.” Even in the Middle Ages, empires preferred to be separated by spaces—“marches” (lat. *margo*)—rather than by fixed lines. In premodern times, feudal, vertical bonds of fealty were of greater importance than territorial

frontiers, and local borders—city limits, customs and toll collection points—exceeded state borders in their practical impact on everyday life. Only the decline of the feudal and the rise of the absolutist order, and then of the nation-state, necessitated greater reliance on clearly defined and defensible boundaries.

Borders became the defining feature of the emerging territorial states, as fixed in the Westphalian system (1648), following the Thirty-Years' War, with its "permanent and unalterable" international frontiers. They were neither permanent nor unalterable. The next attempt to achieve and fix a balanced international state system was undertaken in the Treaty of Utrecht (1713); and then again at the Congress of Vienna (1815) after the Napoleonic Wars.

Social contract theories highlighted the right to exit a territory (if an individual or a group found that the territorial government had broken the contract)—a stipulation closely linked with the discovery of "uninhabited" lands, particularly in the New World. Right of entry, on the other hand, rested solely with the sovereign (with the emerging exception of diplomatic immunity).

Of particular importance was the concept of borders for the emerging European nation-states in the wake of the French Revolution: Homogenization within (i.e., erosion of cultural and linguistic boundaries) and "rational" or "natural" external frontiers exactly delineated frontiers to mark the limits of exclusive authority and sovereignty, guaranteeing and safeguarding the modern state's claim to be the "sole, exclusive fount of all powers and prerogatives of rule" (Poggi 1978, p. 92).

Nation-states established the "classical" functions of frontiers—limits of state jurisdiction, fiscal limits, lines of military protection and defense, customs borders, and sociocultural boundaries. Borders serve to safeguard stability within and protect against external threats. All these functions aim at differentiating between inside and outside, between "us" and "them"; they are means of inclusion and exclusion. In this sense, they are, first and foremost, barriers, enclosing a "security community" (Deutsch et al. 1957). Wilson and Donnan define borders as "political membranes through which people, goods, wealth and information must pass in order to be deemed acceptable by the state. Thus," they argue, "borders are agents

of a state's security and sovereignty, and a physical record of a state's past and present relations with its neighbours." In their view, borders consist of three elements: (1) "the legal borderline which simultaneously separates and joins states;" (2) "the physical structures of the state which exist to demarcate and protect the borderline, composed of people and institutions which often penetrate deeply into the territory of the state;" and (3) "territorial zones of varying width which stretch across and away from borders, within which people negotiate a variety of behaviours and meanings associated with their membership in nations and states." (Wilson and Donnan 1998, p. 9).

In the late nineteenth and early twentieth centuries, concepts of the border assumed, under the auspices of "geopolitics," a Darwinian—biological or organic—interpretation. Friedrich Ratzel (1897) argued that states were living organisms, with frontiers as their skins. And as the organism grew (or shrunk), the skin would adapt, or have to be adapted. States, in his view, were striving to secure the necessary "living space" (*Lebensraum*) for their people and the most effective borders to safeguard them. While Lucien Febvre (1922) rejected such ideas of a "natural" border, seeing it closely linked to the militarization of the modern state, Ratzel's disciple Karl Haushofer paved the way for the Nazi interpretation of the state ruthlessly pursuing the frontiers deemed necessary for expansionist policies based on racial superiority (Haushofer 1986; Murphy 1997).

In the United States, Frederick Jackson Turner, from the perspective of the settlement of the North American continent, saw the "frontier" as a moving section where wilderness and civilization meet, having profound implications for an American "frontier mentality" (Turner 1894). This dynamic concept of frontier as a moving space led to purchase of territory (Louisiana, 1803, from France; Florida, 1819, from Spain) and territorial conquest through war (New Mexico, Arizona, Nevada, Utah, California, part of Colorado, 1848, from Mexico), creating the Rio Grande border. The northwestern frontier with Canada was settled, after decades of controversy with Britain, in 1846, following the 49th parallel.

Frontiers in South America have been products of European colonialism. Yet, in contrast to

Africa, they developed over a longer period of time. South American independence, on the other hand, preceded African decolonization by about 150 years. According to the principle of *uti possidetis*, the administrative boundaries of the Spanish colonial government were generally accepted, but disputes about the exact location of frontiers have led to military conflicts up to the present day (Peru and Ecuador signed a peace treaty in 1998).

African boundaries were mostly drawn by the colonizing European powers in the late nineteenth century (Berlin Conference 1884–1885) in the “scramble for Africa,” largely ignoring ethnic, tribal, and linguistic structures on the continent (Asiwaju 1985). These artificial boundaries, sometimes modified in the implementation process *in loco*, were accepted by the Organization of African Unity (OAU) at their first postcolonial conference in Cairo, in 1964, in order to guarantee stability. Recent events, such as those in Rwanda and Somalia, have given rise to fundamental questioning of these colonial frontiers and the state system exported from Europe along with them.

The literature relating to frontiers and boundaries in recent times is extensive (M. Anderson 1983; Anzaldúa 1987; Barth 1969; Brownlie 1979; Day 1987; Foucher 1988; Herzog 1990; Heyman 1991; Koptoff 1987; Kratochwil 1986; Lamb 1968; Luard 1970; Martínez 1994a; Prescott 1987; Sahlins 1989; Strassoldo 1973; Strassoldo and Delli Zotti 1982; Tägil 1977). Particularly since the fall of the Berlin Wall, frontiers have increasingly returned to the political and academic discourse in Europe (M. Anderson 1996; M. Anderson and Bort, eds. 1998; Blake 1994; Brunn and Schmitt-Egner 1998; Donnan and Wilson 1994; Eger and Langer 1996; Eskelinen et al. 1998; Foucher 1990; Ganster et al. 1997; Kramer 1997; Martínez 1994b; Murray and Holmes 1998; Neuss et al. 1998; O’Dowd and Wilson 1996; Raich 1995; Rupich 1994; Wilson and Donnan 1994–1998a).

New international frontiers have been created (e.g., the Baltic states, the former Yugoslavia, the Czecho-Slovakian “velvet divorce,” Moldova); other boundaries have changed their function fundamentally, particularly in the case of the former Iron Curtain. Resurgent nationalisms in the former Soviet Union and in the Balkans, emulating the classic European claim to independent

nationstates in multinational contexts, have highlighted the inherent contradiction between the international community’s accepted principles of “sanctity of borders” and the “right to national self-determination (Hayden 1992; Sluga 1998).

At the same time, the rhetoric about a borderless Europe (i.e. the retreat of the classic nation state within the European Union) has been, at least partially, translated into reality. Free movement of goods and people was already envisaged in the Treaty of Rome (1957), the founding document of the European (Economic) Community. With the introduction of the Single Market in 1993, based on the 1986 Single European Act, the economic functions (customs, tariffs, etc) of borders inside the EU have been eroded. Since 1995, the 1985 Schengen Agreement and the 1990 Schengen Convention have been progressively implemented, in the process blurring the distinction between international and substate boundaries within the EU (den Boer 1998).

Political frontiers as limits of sovereign states were a European invention and were subsequently exported through colonialism and imperialism. Europe is now facing the biggest challenges to the traditional role of borders. The Schengen Convention, as of 1999 signed by all but two EU member states (the United Kingdom and Ireland) and implemented in ten member states, abolished border controls (passport controls, border police checkpoints) at the internal frontiers and transferred those border controls, standardized and supervised by the Schengen Control Committee, to the external frontiers of “Schengenland.” “The general purpose of frontiers in the sovereign state was to establish absolute physical control over a finite area and to exercise exclusive legal, administrative and social controls over its inhabitants. But the traditional attributes of ‘sovereignty’ are clearly being eroded in Europe and frontiers are losing their hard-edged clarity” (M. Anderson 1996, p. 89). Pooling of sovereignty, the legal prerogative of EU law over national law, and economic and monetary union (with the introduction of the euro in 1999) are often cited as indicators of the demise of the nation-state. This may be exaggerated. Milward (1992) has argued that the EU actually came to the rescue of the nation-state by providing the material benefits that secure its legitimacy. Yet

Milward did not take into account the developments in the 1990s, as expressed in the 1992 Maastricht and 1997 Amsterdam Treaties on European Union, the latter incorporating, for example, the Schengen *acquis* into the institutional framework of the EU.

Will the processes that can be observed in Europe have repercussions elsewhere? Commentators from areas where the European border experience was exported have taken a keen interest (Asiwaju 1996). If borders lose their symbolic and real functions in security and control, will other functions—such as markers of identity and culture—become more salient? Fernand Braudel (1985) observed how the Roman “frontier between the Rhine and the Danube was . . . a cultural frontier *par excellence*,” exercising tangible influence long after its historical demise (p. 66).

The Schengen Convention already entails a transformation of the border-line into a spatial concept of borders, which, in the light of advancing surveillance technology and the need to combat cross-border crime at the locus of its origin (or destination), rather than at the border has been seen as a return of the marches, or *limes* (Foucher 1998).

In addition, and orchestrating this process of erosion, institutionalized cross-border cooperation has become a common feature at nearly all European frontiers, having started along the German–Dutch and German–French frontiers and subsequently expanded with every phase of EU enlargement.

This, against the backdrop of the demise of the Iron Curtain, has also helped to focus attention on the fact that although borders may be barriers, they can also be (or become) points of contact, channels of communication and interaction. Martinez (1994b) describes four types of interaction at borders, arranged on a continuum from closed to open:

1. “Alienated borderlands,” characterized by political and military tensions that allow for very little, if any, exchange across the border. The border is closed; borderlanders on each side perceive of each other as aliens.

2. “Coexistent borderlands,” where contacts are possible and limited exchange takes place but long-term cooperation seems undesirable for political or military reasons.
3. “Interdependent borderlands,” where contacts are frequent, mutual trade and exchange across the frontier has assumed complementary character, and a common borderland mentality is being developed both sides of the border. The border, however, is still closely monitored and only open in so far as the states’ interests are not damaged.
4. “Integrated borderlands,” where all barriers and obstacles to cross-border communication, exchange, and movement of people, goods, services, and capital have been removed and a common cultural cross-border identity is developing.

Although these are “ideal-typical” definitions, it is not difficult to find practical examples to these four stages, which can, as envisaged by Martinez, be seen as stages in an evolution. Number 1 would be the historical example of the bipolar Cold War frontier, the Iron Curtain, as symbolized by the Berlin Wall, or, perhaps, by the “Green Line” separating Turkish and Greek Cypriots. Number 2 could be the borders between former Soviet republics, such as Belarus and Poland. Number 3 is clearly Martinez’s model for the U.S.–Mexican border, where, under the umbrella of NAFTA, goods and capital may flow relatively unhindered across the internal border of the free-trade area, but movement of people is restricted and border control is a high priority. Number 4 would be the internal frontiers of the post-Schengen European Union, classically expressed in the close cross-border relations along the German–Dutch (Euregio) or German–French borders, including the German provision of transferring sovereignty rights to institutions straddling the frontier (Beyerlin 1998).

Yet opening frontiers is not seen solely as a positive process. What is apparent is that people in Europe and the United States seem to harbor an unfocused, general anxiety about frontiers no longer providing the protection they once did. Organized cross-border crime, trafficking of drugs and other smuggled goods, and organized human trafficking seem to indicate that frontier controls are

no longer as effective as they once were. This may be changing as populations become more accustomed to the absence of frontier controls at the internal frontiers. This absence is widely welcomed in frontier regions. In general, the French—normally very sensitive to these matters—seem to have adopted a reasonably relaxed attitude about open frontiers, and those living in the frontier regions seem very pleased with the new situation. Law enforcement agencies seem to have adapted to the new situation without undue difficulty. The nature of frontiers is perceived as changing. New information technology for surveillance and identity control is widely seen as a key factor in securing efficient frontier controls.

The dangers of cross-border crime—drug trafficking, illegal weapons trade, car and cigarette smuggling, money laundering, fraud and corruption, human smuggling, and so on—must not be underestimated. At “their most extreme, substantial rises in the proportion of illegality in international economic activity can destabilize national economies” (Holmes 1999). The rise in internal and cross-border crime in Eastern Europe, particularly in the countries of the former Soviet Union, can be attributed to the difficult transitional situation in these countries: post-communist states attempting, in Claus Offe’s (1996) term, a “triple transition”: the rapid and simultaneous transformation of their political systems, their economic systems, and their boundaries and identities.

The discourse of migration control has become intricately linked with the discourses on crime and security in a process of “securization” (Bigo 1999; Huysmans 1995). Security has become a much broader concept, compared with the focus on military concerns that dominated the discourse until the changes of 1989–1990, encompassing new risks and threats to society, the economy, and the polity itself (Zielonka 1991). This constitution of a security continuum, including the control of frontiers and immigration among police activities in the fight against crime, is, Bigo argues, “not a natural response to the changes in criminality,” but rather a proactive mixing of crime and immigration issues (Bigo 1999, p. 67–68). Buzan has coined the term “societal security” to describe the shift of security concerns from protection of the state to protection against threats, or perceived

threats, against society and identity or against the identity and security of groups within a society (Buzan 1991).

The southern and eastern frontiers of the European Union, as well as the U.S.–Mexican border, demonstrate that the “promotion of borderless economies based on free market principles in many ways contradicts and undermines . . . efforts to keep borders closed to the clandestine movements of drugs and migrant labor.” (Andreas 1996, p. 51). Yet despite these efforts at tightening border controls, even erecting what has euphemistically been dubbed the “tortilla curtain”—a metal wall along the border south of San Diego—and combining military and law enforcement agencies, “many clandestine border crossers are adapting rather than being deterred.” (Andreas 1996, p. 64). Economic factors, “underlying push-pull factors” (Andreas 1996, p. 68), have frustrated repeated attempts at closing the U.S.–Mexican border to illegal migrants. Operation “Wetback” (under Richard Nixon) and, more recently, operation “Gatekeeper,” caused “immediate economic damage, tensions between social groups and [had] almost zero effect on illegal immigration.” (Bigo 1998, p. 159).

Clearly, these functional changes of states will be reflected in the functional changes of their borders, and vice versa. Borders, Foucher (1998) reminds us, “are time inscribed into space or, more appropriately, time written in territories” (p. 249). Thus, “different conceptions of the frontier as an institution existed before the modern sovereign state and other kinds will emerge after its demise” (M. Anderson 1996, p. 5). If we are not witnessing the demise of the nation-state, under the dual pressures of globalization and regional responses, the least we can state is that it is “diversifying, developing,” if “not dying” (Mann 1996).

Or will the new security architecture being created in Europe establish or cement dividing lines that will echo the maintenance of global inequalities? Is the hardening of the external frontier of—perhaps an enlarged—EU part of the Huntingtonian scenario of a “clash of civilizations” (Huntington 1996)? In view of the complexities within what Huntington loosely defines as “civilizations” as well as between them, this is unlikely (Holmes 1998), but frontiers will remain



instruments of politics and instruments for the protection of interests.

REFERENCES

- Anderson, B. 1983 *Imagined Communities: Reflexions on the Origin and Spread of Nationalism*. London: Verso.
- Anderson, M. 1996 *Frontiers: Territory and State Formation in the Modern World*. Cambridge: Polity Press.
- , ed. 1983 *Frontier Regions in Western Europe*. London: Frank Cass.
- , and E. Bort, eds. 1998 *The Frontiers of Europe*. London: Pinter.
- Anzaldúa, G. 1987 *Borderlands/La Frontera: The New Mestiza*. San Francisco: Spinsters/Aunt Lute.
- Asiwaju, A. I., ed. 1985 *Partitioned Africans: Ethnic Relations Across Africa's International Boundaries, 1884–1984*. London: Hurst and University of Lagos Press.
- 1996 "Public Policy for Overcoming Marginalization: Borderlands in Africa, North America and Western Europe." In S. Nolutshungu, ed., *Margins of Insecurity: Minorities and International Security*. Rochester, N.Y.: Rochester University Press.
- Barth, F., ed. 1969 *Ethnic Groups and Boundaries: The Social Organisation of Cultural Difference*. London: Allen and Unwin.
- Bateson, G. 1972 *Steps to an Ecology of Mind*. San Francisco: Chandler.
- Beyerlin, U. 1998 "Neue rechtliche Entwicklungen der regionalen und grenzüberschreitenden Zusammenarbeit." In G. Brunn and P. Schmitt-Egner, eds., *Grenzüberschreitende Zusammenarbeit in Europa: Theorie–Empirie–Praxis*. Baden-Baden: Nomos.
- Bigo, D. 1998 "Frontiers and Security in the European Union: The Illusion of Migration Control." In M. Anderson and E. Bort, eds., *The Frontiers of Europe*. London: Pinter.
- 1999 "The Landscape of Police Co-operation." In E. Bort and R. Keat, eds., *The Boundaries of Understanding*. Edinburgh: ISSI.
- Blake, G., ed. 1994 *World Boundaries Series*, 5 vols. London: Routledge.
- Braudel, F. 1985 *Civilisation and Capitalism: Vol 3. The Perspective of the World*. London: Collins.
- Breuilly, J. 1993 *Nationalism and the State*. Manchester, U.K.: Manchester University Press.
- Brown, G. S. 1969 *Laus of Form*. London: Allen and Unwin.
- Brownlie, I. 1979 *African Boundaries: A Legal and Diplomatic Encyclopedia*. London: Hurst.
- Brunn, G., and P. Schmitt-Egner, eds. 1998 *Grenzüberschreitende Zusammenarbeit in Europa–Theorie–Empirie–Praxis*. Baden-Baden: Nomos.
- Buzan, B. 1991 *People, States and Fear: An Agenda for International Security Studies in the Post Cold War Era*. London: Harvester Wheatsheaf.
- Cohen, A. P., ed. 1986 *Symbolising Boundaries*. Manchester, U.K.: Manchester University Press.
- Day, A. J., ed. 1987 *Border and Territorial Disputes*. London: Keesings Reference Publications.
- De Greef, G. 1908 *La structure generale de la societe, III: La theorie des frontieres et des classes*. Brussels; Larcier; Paris: Alcan.
- den Boer, M., ed. 1998 *Schengen's Final Days*. Maastricht, Netherlands: European Institute of Public Administration.
- Deutsch, K. W., et al. 1957 *Political Community and the North Atlantic Area: International Organization in the Light of Historical Experience*. Princeton, N.J.: Princeton University Press.
- Éger, G. and J. Langer, eds. 1996 *Border, Region and Ethnicity in Central Europe*. Klagenfurt, Austria: Norea.
- Eskelinen, H., I. Liikanen, and J. Oksa, eds. 1998 *Curtains of Iron and Gold: Reconstructing Borders and Scales of Interaction*. Aldershot, U.K.: Ashgate.
- Febvre, L. 1922 *La Terre et l'évolution humaine*. Paris: La Renaissance du Livre.
- Foucher, M. 1988 *Frons et frontières*. Paris: Fayard.
- Ganster, P., A. Sweedler, J. Scott, and W. D. Eberwein, eds. 1997 *Borders and Border Regions in Europe and North America*. San Diego: San Diego State University Press.
- Guichonnet, P., and C. Raffestin 1974 *Géographie des frontières*. Paris: Presses Universitaires de France.
- Haushofer, K. 1986 *De la Géopolitique*. Paris: Fayard.
- Hayden, R. M. 1992 "Constitutional Nationalism in the Formerly Yugoslav Republics." *Slavic Review* Winter.
- Herzog, L. A. 1990 *Where North Meets South: Cities, Space and Politics on the US–Mexico Border*. Austin: University of Texas Press.
- Heyman, J. 1991 *Land, Labor and Capital at the Mexican Border*. Flagstaff: University of Arizona Press.
- Holmes, L. 1998 "Europe's Changing Boundaries and the Clash of Civilizations Thesis." In P. Murray and L. Holmes, eds., *Europe: Rethinking the Boundaries*. Aldershots, U.K.: Ashgate.

- 1999 "Crime, Corruption and Politics: International and Transnational Factors." In J. Zielonka and A. Pravda, eds., *Democratic Consolidation in Eastern Europe: International and Transnational Factors*. Forthcoming.
- Huntington, S. 1996 *The Clash of Civilizations and the Remaking of the New World Order*. New York: Simon and Schuster.
- Huysmans, J. 1995 "Migrants as a Security Problem: Dangers of 'Securitizing Societal Issues.'" In R. Miles and D. Thranhardt, eds., *Migration and European Integration: The Dynamics of Inclusion and Exclusion*. London: Pinter.
- Koptyoff, I., ed. 1987 *The African Frontier*. Bloomington: Indiana University Press.
- Krämer, R. 1997 *Grenzen der Europäischen Union*. Potsdam, Germany: Brandenburgische Landeszentrale für Politische Bildung.
- Kratochwil, F. 1986 "Of Systems, Boundaries and Territoriality: An Inquiry into the Formation of the State System." *World Politics* 1: .
- Lamb, A. 1968 *Asian Frontiers: Studies in a Continuing Problem*. London: Pall Mall Press.
- Luard, E., ed. 1970 *The International Regulation of Frontier Disputes*. London: Thames and Hudson.
- Luhmann, N. 1982 "Territorial Boundaries as System Boundaries." In R. Strassoldo and G. Delli Zotti, eds., *Cooperation and Conflict in Border Areas*. Milan: Angeli.
- Mann, M. 1996 "Nation States in Europe and Other Continents: Diversifying, Developing, Not Dying." In G. Balakrishnan, ed., *Mapping the Nation*. London: Verso.
- Martinez, O. J. 1994a *Border People: Life and Society in the US-Mexico Borderlands*. Tucson: University of Arizona Press.
- 1994b "The Dynamics of Border Interaction: New Approaches to Border Analysis." In C. H. Schofield, ed., *Global Boundaries*. London: Routledge.
- Miller, J. G. 1977 *Living Systems*. New York: McGraw-Hill.
- Milward, A. 1992 *The European Rescue of the Nation State*. London: Routledge.
- Murphy, D. T. 1997 *The Heroic Earth: Geopolitical Thought in Weimar Germany, 1918-1933*. Kent, Ohio: Kent State University Press.
- Murray, P., and L. Holmes, eds. 1998 *Europe: Rethinking the Boundaries*. Aldershot, U.K.: Ashgate.
- Neuss, B., P. Jurczek, and W. Hilz, eds. 1998 *Grenzübergreifende Kooperation im östlichen Mitteleuropa*. Tübingen: Europäisches Zentrum für Föderalismus Forschung.
- O'Dowd, L., and T. M. Wilson, eds. 1996 *Borders, Nations and States*. Aldershot, U.K.: Avebury.
- Offe, C. 1996 "Capitalism by Design? Democratic Theory Facing the Triple Transition in Eastern Europe." *Social Research* 58(2):3-13.
- Poggi, G. 1978 *The Development of the Modern State*. London: Hutchinson.
- Prescott, J. R. V. 1987 *Political Frontiers and Boundaries*. London: Allen and Unwin.
- Raich, S. 1995 *Grenzüberschreitende und interregionale Zusammenarbeit in einem "Europa der Regionen"*. Baden-Baden: Nomos.
- Ratzel, F. 1897 *Politische Geographie*. München: Oldenberg.
- Rupnik, J. 1994 "Europe's New Frontiers." *Daedalus* Summer.
- Sahlins, P. 1989 *Boundaries: The Making of France and Spain in the Pyrenees*. Berkeley: University of California Press.
- Sluga, G. 1998 "Balkan Boundaries: Writing History and Identity into Territory." In P. Murray and L. Holmes, eds., *Europe: Rethinking the Boundaries*. Aldershot, U.K.: Ashgate.
- Strassoldo, R., ed. 1973 *Confini e Regioni*. Trieste, Italy: Lint.
- 1976-1977 "The Study of Boundaries: A Systems-Oriented, Multidisciplinary, Bibliographical Essay." *The Jerusalem Journal of International Relations* 2(2):81-107.
- 1979 "La teoria del confine." In *Temi di sociologia delle relazioni internazionali*. Gorizia, Italy: ISIG.
- 1982 "Boundaries in Sociological Theory, a Reassessment." In R. Strassoldo and G. Delli Zotti, eds., *Cooperation and Conflict in Border Areas*. Milan: Angeli.
- 1998 "Perspectives on Frontiers: The Case of Alpe Adria." In M. Anderson and E. Bort, eds., *The Frontiers of Europe*. London: Pinter.
- Strassoldo, R., and G. Delli Zotti, eds. 1982 *Cooperation and Conflict in Border Areas*. Milan: Angeli.
- Tägil, S. 1977 *Studying Boundary Conflicts*. Stockholm: Scandinavian University Books.
- Turner, F. J. 1894 *The Frontier in American History*. Madison: State Historical Society of Wisconsin.
- Wilden, A. 1972 *System and Structure*. London: Tavistock.
- Wilson, T. M., and H. Donnan, eds. 1994 *Border Approaches: Anthropological Perspectives on Frontiers*. Lanham, Md.: University Press of America.

—, eds. 1998 *Border Identities: Nation and State at International Frontiers*. Cambridge, U.K.: Cambridge University Press.

— 1998 "Nation, State, and Identity at International Borders." In T. M. Wilson and H. Donnan, eds., *Border Identities: Nation and State at International Frontiers*. Cambridge, U.K.: Cambridge University Press.

Zielonka, J. 1991 "Europe's Security: A Great Confusion." *International Affairs* 67(1):127-137.

RAIMONDO STRASSOLDO  
EBERHARD BORT

## NATIONALISM

Nationalism has been defined in a variety of ways at different levels of analysis (e.g., Kohn 1955, 1968; Symmons-Symonolewicz 1970; Kamenka 1973; Plamenatz 1973; Smith 1976, 1981; Snyder 1984). The concept combines a sense of identification with a people, an ideology of common history and destiny, and a social movement addressed to shared objectives. This definition raises questions about what differentiates a people or a nation from others, the nature of identification, the conditions under which nationalist ideologies develop, and the course and aims of nationalist movements.

Historically the term *nationalism* was applied to attempts to follow early European models "to make the boundaries of the state and those of the nation coincide" (Minogue 1967, p. 12), that is, to create loyalty to a nation-state (Kohn 1968). It was also applied to struggles, proliferating after World War II, to gain independence from colonial domination and join the community of sovereign states. More recently, however, analysts have found the confusion between the concepts of state and nation to be a hindrance to understanding contemporary nationalism. Only rarely, if at all, do the boundaries of a state coincide with those of a nation. By *nation* we mean an ethnic group that (1) shares one or more identifying characteristics, such as language, religion, racial background, culture, and/or territory; and (2) is politically mobilized or is amenable to such mobilization. Thus, *nationalism* should be distinguished from *patriotism*, in that the identification and loyalty in the former is to an ethnic group or nation, and in the latter to the state.

## IDENTIFICATION

Debates centering on the intrinsic nature of collective ethnic identification feature variations on two general themes—primordialism and structuralism. Implied in the primordial perspective is a deemphasis on an instrumental view of ethnic ties. These ties are seen as ends in themselves shaped by forces other than material self-interest; they are persistent and they resist the homogenization predicted by convergence and modernization theorists. The essence of these ties "is a psychological bond that joins a people and differentiates it in the subconscious conviction of its members from all other people in the most vital way" (Connor 1978, quoted in Stack 1986, pp. 3, 4). These bonds stem from "immediate contiguity and kin connection mainly, but beyond them . . . from being born into a particular religious community, speaking a particular language, or even a dialect of a language, and following particular social practice" (Geertz 1963, pp. 14, 15). Thus, there can be multiple identities—for instance, a Moroccan may identify with an ethnic group within Morocco, the Maghrebs of North Africa, the Arabic speaking people, and the Muslim people at large. These identities assume varying significance depending on the context, lending credence to the old saying "My brother and I against my cousin, and my cousin and I against the stranger." To structuralists, "ethnic identity results instead from objective intergroup differences in the distribution of economic resources and authority" (Hechter 1986, p. 109). Implied here is rational choice and self-interest, that ethnic ties are means to certain ends, and that the boundaries of ethnic groups are changeable.

Neither primordialists nor structuralists would deny the obvious variations in the intensity of identification and in the potential for nationalist movements cross-culturally and over time. Several factors are expected to contribute to these variations. Among the more important of these factors is coterminality of characteristics. In most instances multiple characteristics are involved in distinguishing among ethnic groups—in an overlapping manner at times and coterminously at others. The United States offers an example of overlapping identities where people from different racial backgrounds share the same religious orientation, people with different religious orientation share a common language, and there is no territorial exclusiveness. This overlap in ethnic identification is

credited in part with lowering intergroup tensions (Williams 1947). At the other extreme are peoples in southern Sudan, Eritrea, Tibet, the Tamils and Sinhalese in Sri Lanka, the republics of the former Soviet Union, the republics of former Yugoslavia, and other places where all or a combination of racial, religious, linguistic, cultural, and territorial maps largely coincide. The greater the number of factors that coincide, the greater the gulf or "social fault" among ethnic groups along which nationalist sentiments and tensions are likely to intensify (Nagi 1992).

In addition to differentiating attributes and geographic distributions, a number of features of the social structure contribute heavily to variance in intensity and patterns of nationalism. Hechter (1986) suggests two types of such factors. One is the institutionalization of ethnic differences in legal and normative rules, especially those governing property and civil rights, as was the case in South Africa. The other is differentiation in positions in the division of labor that shape "specialization experiences as well as material interests." Deutsch offers another factor in nationalism, attributing membership in a people essentially to a "wide complementarity of social communication" that "consists in the ability to communicate more effectively, and over a wide range of subjects, with members of one large group than with outsiders" (1953, p. 71). Communication is subject not only to commonality of language and cultural background, but also to available means. These include networks of social relations as well as the ever-advancing technological means of mass communication.

### NATIONALISM

In addition to identification and "consciousness of kind," nationalism involves ideology and mobilization for social movement and political action. The ideology stems from identification, the sense of uniqueness of group origin, history, culture, collective authority, and destiny (Smith 1981). Political mobilization and the course of nationalist movements are greatly influenced by a host of internal and external factors. Important among the internal factors are uneven economic conditions and disparities along cultural lines in control of resources, access to goods, and distribution of

positions in the occupational structure. The collective perception of an ethnic group of its deprivation or exploitation may challenge the legitimacy not only of the regime but also of the state itself. The theme of "relative deprivation" has been central to explanations of political violence (e.g., Gurr 1970).

The political system allocates power and authority. An uneven distribution among ethnic groups can be the direct result of exclusionary rules and practices, as in the case of Apartheid in South Africa, or the indirect result of socioeconomic disparities and associated discriminatory practices, as characterized in the relationships between Catholics and Protestants in Northern Ireland. Whatever the reason, a cultural distribution of power is as evocative of a nationalist sense of deprivation as a cultural distribution of resources. The close relationship between power and resources led Lasswell to observe that politics is "Who Gets What, When, How" (1936, p. ii).

Applying a developmental perspective, Huntington (1968) connects political stability or instability to the balance between "institutionalization" and "participation." This line of reasoning suggests to some analysts (e.g., Sanders 1981) a linear relationship between political development and political instability—in this case arising from nationalism, which is credited with having been the most prevalent reason for state-level violence (Said and Simmons 1976). Support for such a pattern of relationships derives also from the expectation that the national integration and political assimilation characteristic of developed societies contribute to a shift from culturally based to functionally based cleavages. The net result is a reduction in the prevalence and intensity of ethnic mobilization in the more politically developed countries. However, clearly implied in Huntington's discussion of relations between demand for participation and institutional capacity is a balance at low levels in the least developed countries and at high levels in the developed ones. Imbalances can be expected at early and middle states of development, which suggests a curvilinear pattern.

There is no unanimity on the relationships between development and nationalism. This is understandable in view of the complexity of both phenomena and the current state of research. An important voice on these relationships is that of

Walker Connor (e.g., 1972), who contends that modernization, the spread of education, and improved means of mass communication are responsible for a resurgence in ethnic nationalism. Juan Linz observes that "in the modern world the aim seems to be to build nations rather than states, a task that is probably beyond the capacity of any state that has not achieved the characteristics of a nation-state before the era of nationalism" (1978, p. 62). The distinction between "level" and "process" of modernization and development is useful in understanding the rise in nationalist sentiments. As Smith has noted, "Perhaps, then, it is not the fact of economic progress or decline that is relevant to ethnic revival, but simply economic change per se. Most change . . . is painful and uprooting" (1981, p. 34). Tensions, strains, and dislocations associated with change in the structure and distribution of power—political change—are no less painful or uprooting. Attempts by central governments to secularize and to shift loyalties from ethnic groups to the state underlie much of contemporary nationalism.

#### NATIONALIST MOVEMENTS

The literature offers two perspectives in explaining the formation of social movements—collective behavior and resource mobilization. In the former, social movements are viewed as responses to a rise in grievances and the actors are seen as "arational" if not "irrational"; in the latter they are considered as goal-oriented, rational responses dependent on organization and mobilization of resources (Jenkins 1983). Debates concerning the strengths and limitations of these two perspectives are yet to be settled. Important to a discussion of nationalism, however, is that literature on social movements, especially on resource mobilization, is primarily Western in conceptualization and empirical foundations. Significant in this respect are differences in aims that guide social movements. More common to Western societies is "changing some elements of the social structure and/or reward distribution of a society" (McCarthy and Zald 1977, pp. 1217, 1218), as compared to nationalist movements that press for autonomy, if not for secession. The first type of movement seeks change through influencing political institutions, the legitimacy of which is not in question (e.g., civil rights in the United States and labor movements in many countries). Nationalist movements, on the

other hand, often challenge, if not outright reject, the legitimacy of the state (e.g., Croatia, Slovenia, Bangladesh, Eritrea, and others that eventually seceded from larger states, as well as Biafra, Southern Sudan, Quebec, and others where no separation has occurred). This is not to cast doubt on the applicability of theories of social movements to ethnic nationalism but, rather, to point out that the influence of differences among societies in levels of development, political institutions, types of regimes, and movements' aims has not been adequately explored (see McCarthy and Zald 1977; Tilly 1978; Jenkins 1983).

In the following paragraphs we shall outline some of the important features of nationalist movements and the processes of mobilization. More specifically, we shall consider the role of grievance, resources, repertoires of expression, and aims.

The role of grievances remains unresolved. While traditional analysis places grievances stemming from structural strains associated with social change at the root of movements (e.g., Smelser 1962; Gurr 1970; Gusfield 1970), resource mobilization proponents favor structural "causal" explanations (e.g., Tilly 1978). Some feel discontent is ubiquitous and therefore, by itself, cannot explain the emergence of social movements. McCarthy and Zald go even further: "For some purposes, grievances and discontent may be defined, created and manipulated by issue entrepreneurs and organizations" (1977, p. 1215). In line with this perspective is Smith's account of "ethnic revival" (1981). Smith maintains that the unique and distinguishing rationale of an ethnic group is the emphasis on group belonging and group uniqueness that links successive generations of the group with specific origins and history. Driving the engines of "historicism" and "nationalism" are discontented intellectuals, educators, and professional intelligentsia. Blocked mobility, opposition and repression by traditional authorities, and frustrated expectations concerning recognition, especially on the part of Third World intellectuals, are factors in the radicalization of these groups, which then turn to historicism and inward to their ethnic communities (Smith 1981).

In explaining the spread of "value-oriented movements," Smelser (1962) refers to "structural conduciveness." Two elements of conduciveness

are highly applicable to ethnonationalist movements. One relates to the importance of communication in “disseminating a generalized belief”—a position consistent with that of Deutsch (1953) and Connor (1972), cited earlier. The other is “the availability of means to express grievances” during troubled or uncertain times in order to redress problems.

Attempts to explain nationalist movements must account for the mobilization of resources. McCarthy and Zald outline five central considerations: (1) “aggregation of resources (money and labor)”; (2) the form of organization these resources entail; (3) involvement of individuals and organizations outside the movement; (4) the flow of resources to and from the movement; and (5) “the importance of costs and rewards in explaining individual and organizational involvement” (1977, p. 1216).

The prevailing patterns of social relations are expected to influence the potential for, and forms of, organization and mobilization. Tilly (1978) maintains that combined strength in identification and interpersonal bonds lead to high levels of organization and to greater possibilities for mobilization. In a similar vein Oberschall (1973) offers a classification for patterns of organization. Along one dimension—relations within collectivities—he identifies three types: “communal,” “associational,” and “weakly [organized] or unorganized.” Mobilization, which is facilitated by communal and associational forms, is rendered difficult by the weakly organized and unorganized structures. Along another dimension, Oberschall distinguishes between “vertical” and “horizontal” relations to other collectives and segments of society. Social and political bonds across classes and collectivities can influence mobilization; however, the direction of influence can be expected to vary depending on the type of movement. For example, Oberschall observes: “If in a stratified society there exist [sic] strong vertical social and political bonds between upper and lower classes, mobilization into protest movements among lower classes is not likely to take place” (1973, p. 120). While this may be the case in regard to class conflicts, vertical bonds within an ethnic group can significantly facilitate mobilization.

Strength in family and kinship relations underlies Tilly’s networks of interpersonal bonds

and Oberschall’s communal organization. Houseknecht sees strength here as referring “to the extent to which family/kinship obligations and rights take precedence over their nonkinship counterparts” (1990, p. 1). She outlines important ways in which kinship relates to ethnic identification, nationalism, and the organization of movements. Early socialization builds identification with an ethnic culture and commitment to its values and norms, which continue to be strengthened and enforced through kinship ties. Commonality in cultural background facilitates social communication, and the networks of kinship ties are readily available channels for the mobilization of human and material resources. Furthermore, the traditional authority structure afforded by strong family and kinship systems provides protection in resisting pressures applied by the state’s central authorities.

While informal networks of kinship and interpersonal relations are important to nationalist movements, the role of formal organizations cannot be overstated. As pointed out by McCarthy and Zald (1977), a social movement may include more than one organization; and all organizations in a movement compose a “social movement industry.” Competition over resources can arise between a movement industry and other commitments, as well as among organizations within the same movement industry. The latter form of competition is common to nationalist movements, as in the case of the Kurds. This frequently encourages the involvement of neighboring states or major powers, and often leads to internal conflicts within movements.

These lines of reasoning concerning the role of kinship and interpersonal relations bring up the unresolved debate over the relative effectiveness of bureaucratic centralized movement organizations compared with those of an informal decentralized nature. While some argue that “a formalized structure with a clear division of labor maximizes mobilization by transforming diffuse commitments into clearly defined roles and . . . centralized decision making,” others maintain that “decentralized movements with a minimum division of labor and integrated by informal networks and an overarching ideology are more effective” (Jenkins 1983, p. 539).

Why would individuals and organizations contribute their labor and resources to ethnonationalist

movements? To primordialists, the answer is that they do so in order to preserve cultural integrity and a way of life, to maintain group solidarity and social bonds, and to advance the cause of a community from which a sense of security and pride is derived. These are valued ends in themselves. To collective behavior theorists, the answer is in the collective mood, the social contagion, and the state of mind that are engendered in response to the perception of grievances. To structuralists and resource mobilization proponents, the answer lies in rational collective action and the pursuit of interests.

Opportunities for mobilization are also enhanced by the quality of leadership and its effectiveness in articulating the interests of the group. Smelser explains the significance of charismatic leadership to value-oriented movements. Such forms as “the dreamer prophet of the cult, the nationalist crusader, and the totalitarian demagogue” (1962, p. 355) are compatible with the character of these movements in certain phases of their development. He goes on to point out, however, that “Insofar as a value-oriented movement receives material from outside sources, and insofar as it inherits an organizational structure, the need for charismatic leadership lessens” (p. 356). Useful here are categories identified by Hermann (1986) to analyze political leaders. Particularly important are such contextual factors as to whom the leaders are accountable, forms of interaction with followers, the constraints defined by constituents’ beliefs and norms, the strength and nature of opposition, and available resources.

Repertoires of expression of ethnic nationalist movements vary along a wide spectrum that includes interest articulation, passive resistance, demonstrations and riots, sabotage and terrorist acts, and internal wars. Patterns of expression are shaped by the style of leadership, the level of organization, and resources. They are also influenced by the reactions of the state as well as by external forces. These expressions represent events in the life history of movements. The points at which these movements begin and end are generally difficult, if not impossible, to ascertain. Thus, what might be referred to as beginnings or outcomes may well represent only arbitrarily defined points in a process.

Ethnic movements differ in aims and strategies. Smith identifies six types:

*1. Isolation . . . was the most common strategy for smaller ethnic communities in the past. The ethnic community chooses to stay aloof from society as a whole. 2. Accommodation. Here the ethnic community aims to adjust to its host society by encouraging its members to participate in the social and political life of the society and its state. Often, individual members try to assimilate to the host society, or at least become acculturated, for individual advancement. 3. Communalism . . . is simply a more dynamic and active form of accommodation . . . . The aim is communal control over communal affairs in those geographical areas where the ethnic community forms a demographic majority. 4. Autonomism. There are . . . various forms and degrees of autonomy . . . . Cultural autonomy implies full control by representatives of the ethnic community over every aspect of its cultural life, notably education, the press and mass media, and the courts. Political autonomy or “home rule” extends this to cover every aspect of social, political, and economic life, except for foreign affairs and defense. Ideally, autonomists demand a federal state structure, and this strategy is really only open to communities with a secure regional base. 5. Separatism. This is the classic political goal of ethnonational self-determination . . . . In each case, the aim is to secede and form one’s own sovereign state, with little or no connection with former rulers. 6. Irredentism. Here an ethnic community, whose members are divided and fragmented in separate states, seeks reunification and recovery of the “lost” or “unredeemed” territories occupied by its members. In general, this is only possible where the ethnic community has its membership living in adjoining states or areas. (1981, pp. 15–17)*

Seeking independent rule, as in secessionist or separatist movements, poses the most serious threat to the state. Such movements challenge not only the legitimacy of the government or regime, but also the integrity of the state itself.

It is reasonable to expect that factors shaping the intensity of nationalist movements will, in turn, influence the formation of secessionist goals.

Territorial coterminality with lines of ethnic identification is a strong contributor, with distance from the ruling center adding to the potential for secessionist claims (Young 1975; Islam 1985; Pankhurst 1988). Timing seems to be important to the rise of such claims as well as to their success. Based on the experiences of several African countries, Young (1975) points out that time is conducive when polity in the parent state falls into disrepute, calling its legitimacy into question because of mismanagement, corruption, and discrimination. He also concludes that times of cataclysmic events in the lives of states offer opportunities for secessionist claims because of the fluidity these events engender, the options they open, and the push for choices to be made. Young cites other antecedents that provide a basis for solidarity and mobilization, such as regional differentials in concentration of wealth and representation in power positions, and at least the minimum political resources to make independent status possible even when economic sacrifice is required. To this list of contributing factors, Islam (1985) adds the magnitude of suffering, the impact of the secession of a region on the rest of the country, whether the seceding region represents a majority or a minority of the population, and the involvement of outside powers.

#### REGIONAL AND GLOBAL INFLUENCES

The rise of nationalism and the forms and directions it takes are significantly influenced by forces external to the respective states—regional and global socioeconomic and political conditions, and by the spread of ideologies. The power vacuum created by the liquidation of colonialism after World War II tended to be filled by newly created states where borders had been frequently drawn arbitrarily vis-à-vis ethnic distributions. Regional power struggles, territorial disputes, economic competition, and ideological differences have left many regions of the world fraught with turmoil. Since cultural pluralism is characteristic of most, if not all, of the new states, ethnic nationalism often figures prominently in regional conflicts—as a cause at times, and as a consequence at others. The regional dynamics of ethnonationalism can take many forms. It has been a force behind the formation of alliances among independent states, for instance, the Arab League (1945), the Arab Magreb

Union (1989), and the Gulf Co-operation Council (1981). It was also influential in rare cases of voluntary fusion of independent states. Egypt and Syria formed the United Arab Republic (1958) which “was open to other Arab states to join, but only Yemen entered a loose association” (*Encyclopedia of World History* 1998, p. 689). A military coup in Syria ended its union with Egypt in 1961, and Yemen pulled out in 1966. Ethnonationalism was also used as justification for annexation by force as in Germany’s invasion of Austria and the Sudetenland (late 1930s); China’s occupation of Tibet (1950s); and Iraq’s invasion of Kuwait (early 1990s), which was repelled by an international coalition led by the United States and including several Arab countries.

The phenomenal advancements in means of communication and transportation have enormously increased the intensity and scope of global relations. Four features of these relations are particularly relevant to nationalism. First is the spread of ideologies related to human rights and the right to self-determination. Reports of abuses of these rights are communicated in a graphic and rapid, if not an instant, manner. They galvanize global public opinion and prod governments to intercede. The intervention of the North Atlantic Treaty Organization (NATO) countries and others in opposition to “ethnic cleansing” in Bosnia and Kosovo is a case in point. Rapid communication also brings to ethnic communities the successes of others who are engaged in struggles or have succeeded in attaining varying measures of autonomy. Second is the presence of world forums to address these issues and bring to bear the weight of the global community, such as the United Nations General Assembly, the Security Council, the International Court of Justice, and other governmental and nongovernmental organizations. Third is the increasing globalization of the economy, which makes it possible for small sociopolitical units to find multilateral niches, thereby reducing dependency on bilateral economic relations with either former colonial powers or states from which they have seceded. Fourth is the geopolitical relations among major powers. When the world was polarized between the two superpowers—the United States and the former Soviet Union—they frequently supported different sides of conflicts within pluralistic societies, either directly or through proxies. In the past, both countries have taken



positions in support of the doctrine of “self-determination.” President Wilson was a staunch spokesman for the principle during and after World War I. Around the same time, “at the Seventh All Russian Democratic Labor Party Conference of May 12, 1917 . . . in a resolution drafted by Lenin, the conference unequivocally endorsed the right of all of the nations forming part of Russia freely to secede and form independent states” (Connor 1984, p. 45). In contemporary global relations both countries have shown less commitment to the principle. This became clearly evident in the resistance of the former Soviet Union to the independence of the Baltic States, and in the use of military force to keep Chechnya within the fold of the Russian Federation. The approaches of these powers to nationalism have become subject to strategic and economic interests, tempered by the balance of force that can be brought to the situation. The ebb and flow of relations among these powers exerted considerable influence on ethnonationalism for decades.

The vested interests of other states are usually rooted in resources, trade, security, geopolitical advantage, ethnic affinity, or other ideology. More recently, humanitarian considerations have been assuming greater significance in the foreign policy of the West, most notably the United States, progressing “from the call for moral pressure in the 1970s, to economic sanctions in the 1980s, to military intervention in the 1990s” (Kissinger 1999, p. 43), as in the case of Kosovo. While the large states “are likely to intervene for instrumental reasons,” the small ones “are more likely to intervene for affective reasons” (Heraclides 1990, p. 377). Regional and global influences may strengthen or weaken the state’s means of control and repression, contribute to the movement’s resources and sanctuary, and/or raise awareness and help mobilize regional and global public opinion. These interests and their expression, directly or through international organizations, expand or inhibit opportunities for nationalist movements including those with secessionist aims.

The spread of the ideology of minority rights and self-determination plays a special role in secessionist movements. In a comparative analysis, Sigler concludes: “There is historical weight to the charge that minority rights is a guise for separatist sentiment” (1983, p. 188). He maintains that the “concept of minority rights briefly blossomed under

the minorities treaties system” (p.190) that followed from the Treaty of Versailles. The system was to protect the rights of minorities in many countries, mostly in Europe, after World War I. It gave jurisdiction to the Permanent Court of International Justice, but litigations became protracted and difficult to adjudicate, thus severely limiting the court’s role. It gave power to the Council of the League of Nations to intervene but did not guarantee that disciplinary actions would be carried out against states for infractions. Mutual reinforcement of a system of states as a basis for international relations restrained the various states from interfering in each other’s internal affairs. “The collapse of the minorities treaties system . . . has encouraged minority separatist movements . . . . In the absence of a strong international system for the protection of minority rights, resort to separatist politics may have become more prevalent” (Sigler 1983, p. 190). This reinforces Smelser’s ideas about structural conduciveness mentioned earlier.

If the odds facing secessionists are great, those facing irredentists are much greater. By definition, irredentism involves multiple secessions, which means contending with forces of more than one state. Consider the difficulties facing the Kurds for whom fulfilling the aim of a unified independent homeland (Kurdistan) is resisted by all the states among which they are divided (Azerbaijan, Iran, Iraq, Syria, and Turkey). An exception was when an ethnic group in one state seeks to attach itself to a neighboring state that includes part of the same group. This tactic is used by some separatists to obtain external support (Young 1975), as in the case of Kashmir that has attracted the backing of Pakistan in its struggle to secede from India. Potential annexation is not necessary for a state to intervene in support of related groups in another state. In the course of extending protection to the Greek and Turkish subpopulations of the island state of Cyprus, war almost erupted between Greece and Turkey.

### REACTIONS OF THE STATE

Actions by the authorities and the means of control they employ are largely shaped by the objectives of nationalist movements and greatly affect their course. When perceived as threats to the stability of state and government, they are usually

met with repression. However, nationalist movements that remain clandestine and limited under conditions of severe repression tend to gather momentum and erupt into open expression when a new regime, or change in policy, reduces coercion. Recent events in the former Soviet Union and other countries illustrate the point.

Repression is not the only means used by states in responding to threats of ethnonationalism. Some states, such as the former Soviet Union and the People's Republic of China, adopted population redistribution policies aimed at diluting the coterminality of ethnicity with territory. Language policies in education are also used to increase homogenization. Data and other information about Russia and the former Soviet Union are more readily available than about China. The "russification" of Central Asia involved "several migration waves . . . the most recent of which occurred from the mid-nineteenth century to the 1960's . . . conquest and incorporation of the region was accompanied by Russian in-migration, first of peasants in search of new farmland and later of semiskilled industrial workers who entered with regional economic development under socialism" (Kaiser 1994, p. 238). For example, between 1926 and 1959, the proportion of Russians in the population of Kazakhstan increased from 19.7 to 42.7 percent, and from 11.7 to 30.2 percent in Kurgyzstan (Koslov 1975, quoted in Kaiser 1994). These policies seem to have been counterproductive because of heightening nationalist sentiments among the native populations who did, and continue to, view the Russians as colonialists who enjoy greater power and resources (Kaiser 1994). Migration is not the only means used in population redistribution; native populations may be driven out in a process of "ethnic cleansing" as occurred in Bosnia and Kosovo. Central governments in some pluralist states may adopt redistributive policies in regard to resources to assist economic development in lagging regions. Tito's regime in Yugoslavia introduced such policies which ignited a feeling of being exploited on the part of ethnic groups in the more advanced republics (e.g., Slovenia and Croatia).

Other means in the state's arsenal are arrangements for power sharing in governance and mechanisms for conflict regulation. Nordlinger (1972) outlines six such mechanisms: (1) a stable governing coalition of political parties involving all major conflict groups, (2) proportional distribution of

elective and appointive positions, (3) mutual veto by which government decisions must be acceptable to major conflict organizations, (4) purposive depoliticization in which leaders of conflict groups agree to keep government out of policy areas that impinge upon the various segments' values and interests, (5) regulation by compromise over conflictual issues, and (6) one group granting concessions to another as a way of managing conflicts. He also sees four motives for leaders to engage in conflict regulation: (1) external threats or pressure, (2) negative effects on the economic well-being of the groups involved, (3) aversion to risking violence and human suffering that might result from unregulated conflicts, and (4) the protection of leaders' own power position.

Lijphart places equal emphasis on leaders "whose cooperative attitudes and behavior are needed to counteract the centrifugal tendencies inherent in plural societies" (1977, p. 1). The forms he outlines for "consociational" democracies overlap in many ways with Nordlinger's mechanisms for conflict regulation. One important difference, however, is Lijphart's inclusion of "segmental autonomy," that is, federalism. Nordlinger specifically excluded federalism as a mechanism for regulating conflicts because he saw in it a recipe for a breakup of the state.

Emphasis in the foregoing discussion has been on strategies to prevent the breakup of the state. An alternative, however, is to carry out an orderly and peaceful separation. The division of Czechoslovakia (1993) into two countries (the Czech and the Slovak Republics), although a rare example, renders this alternative real rather than just theoretical.

## OUTLOOK ON THE FUTURE

While forecasting is a hazardous endeavor, especially in regard to such complex and volatile phenomena as nationalism, disciplined consideration of important trends and their implications for the future is warranted. Interest here is in long-term change that may span generations and many decades of evolution. As implied earlier, the forces of modernization and development are intricately interrelated to the prevalence, severity, and modes of resolving nationalist tensions. By modernization, we mean the forces of change in institutions, organization, and behavior in adaptation to the

vast and rapid advancement in knowledge and its technological application. In a classic analysis, Black describes the awesome significance of this process:

*The change in human affairs now taking place is of the scope and intensity that mankind has experienced on only two previous occasions, and its significance cannot be appreciated except in the context of the entire course of world history. The first revolutionary transformation was the emergence of human beings, about a million years ago, after many thousands of years of evolution from primate life . . .*

*The second great revolutionary transformation in human affairs was that from primitive to civilized societies, culminating seven thousand years ago. . . . Three [early civilizations]—the Mesopotamian, the Egyptian, and the Cretan—have transmitted their knowledge and institutions to later societies.*

*The process of change in the modern era is of the same order of magnitude as that from prehuman life and from primitive to civilized societies; it is the most dynamic of the great revolutionary transformations in the conduct of human affairs. What is distinctive about the modern era is the phenomenal growth of knowledge since the scientific revolution and the unprecedented effort at adaptation to this knowledge that has come to be demanded of the whole mankind. (1966, pp. 2–4).*

Emerging through this process is change in institutions toward greater capacity for political participation, expanding and more productive economies, and increasing cultural neutrality in the formulation of laws and universality in their application. As has been noted by many analysts of social change, this transformation has been accompanied by change in value orientation toward increasing emphasis on achievement, secularization, individualism, rational choice, the rule of law, and tolerance for differences. Powering this process are mutually reinforcing systems of education, research, and development. Several points are important in explaining the connections between modernization and nationalism.

*First*, the course of modernization is highly influenced by the ability of leaders “to keep the delicate balance required for survival between the

maintenance of the traditional pattern of values that serves as the basis for cohesion and adaptation to new knowledge that requires a revision of the traditional value system” (Black 1966, p. 4). This “clash of civilizations” is common to developing and transitional societies; it is often seized upon in mobilizing nationalist sentiments. The resulting mixes tend to negate the notion that modernization means turning other societies into clones of those of the West. In fact, there are pronounced differences among the advanced countries in the configurations of institutions, organizations, and values they have evolved.

*Second*, “the challenge . . . in the societies that modernized earlier was internal, the processes of transformation took place generally over centuries. In the later modernizing societies this challenge has been increasingly external, hence more rapid and even abrupt” (Black 1966, p. 8). As pointed out earlier, rapid and abrupt change in itself can create considerable dislocation in institutions and unpredictability in behavior—conditions of anomie—that are fertile grounds for mobilizing fervor for such emotional causes as nationalism.

*Third*, with halts at times and set backs at others, modernization is a continual process with different countries, and different peoples within countries, being at varying stages of the process. Certain levels of transformation bring about institutional arrangements, modes of organization, and value orientations favorable to prospects for peaceful and orderly resolution of nationalist tensions. It is no accident that the ethnic sub-populations in Switzerland (French, German, Italian, and others), in former Czechoslovakia (Czechs and Slovaks), in the United Kingdom (English, Scottish, and Welsh), in Belgium (Flemings and Walloons), and in Canada (citizens of Quebec and others), found ways to resolve, or to go about resolving, these issues within peaceful legal frameworks. To be sure, there remain instances of violent conflicts as in the case of Northern Ireland.

The point to be made here is that, in the long run—over an extended span of history—the process of modernization can be expected to lessen ethnic tensions and promote peaceful resolution. At present and in the short run, however, the ubiquity of ethnic conflicts and the escalation in associated violence are to be expected in many

regions and countries of the world. In part, this is because of the abruptness of change in general, the pressure built over decades of ethnic differentiation in the distribution of power and resources, and fluidity in the structure of central authorities that suppressed their expression as happened in Yugoslavia and the former Soviet Union. The influence of fluidity in power structures is illustrated by Naimark in his foreword to a volume aptly entitled *The Revenge of the Past*:

*The collapse of the Soviet Union was caused by nationalism, that is, by the demands of the subject nationalities of the USSR for genuine independence and autonomy. Unified in their hostility to the Kremlin's authority, the fifteen constituent Union Republics, including the Russian republic, declared sovereignty and began to build state institutions of their own . . . . With the failure of the August 1992 putsch attempt, sovereign republics obtained their independence. Nationalism reigned supreme" (1993, p. ix).*

Another major trend expected to have significant dampening effects on violent nationalism is the development of global and regional governmental and nongovernmental organizations addressed, in part or in whole, to investigating, reporting, and intervening with cases of discrimination and abuses of human rights. Precedents, of historic importance, were introduced in 1999 with the military intervention of the NATO countries in attempting to reverse "ethnic cleansing" in Kosovo, and the indictment of a head of state and high officials in a serving government by a United Nations Tribunal for "crimes of war and against humanity." To focus on the United Nations for a moment, among its achievements over the first fifty years, Alger cites the following:

*It has invented new tools such as peace-keeping. It has extended its mission to include violent conflicts within states. It has broadened peace-keeping to include an array of supportive humanitarian operations. It has drafted a broad array of human rights conventions covering economic, social, cultural, civil, and political dimensions. (1998, p. 422).*

Finally, other trends with potential impacts on nationalism include the *voluntary* formation of

economic and political federations among states such as the European Union, and the globalization of the economy. The growing interdependence among countries brings to bear the weight of regional and global communities in containing the risks of disruptive violence. Furthermore, these trends facilitate greater autonomy for small sociopolitical units by offering multilateral niches that reduce dependence on bilateral economic relations with the countries in which they are part or from which they have seceded. In other words, globalization and regionalization tend to transfer power and authority from the states upward to the larger structures and downward to local levels, thus, increasing local autonomy. As has already been explained, a host of other factors particular to each situation are also at work.

## CONCLUSIONS

In conclusion, it can be said that the current pervasiveness of nationalism should be self-evident. The socioeconomic, political, and other human costs are frequently staggering. In the short and intermediate term, the ubiquity of the phenomenon can be expected to continue. In addition to Africa and the Balkans, there remain significant ethnic tensions, if not risks of violence, in many countries of Asia—Russia, China, and India, among others. However, in the long run, major trends of modernization, economic globalization and interdependence, and the development of international organizations and forums to address the problem, all point to conditions less conducive to the mobilization of nationalist sentiments and more tending to peaceful resolutions of such tensions.

As a major feature of social structure, ethnicity is of central theoretical importance. However, many (and often large) gaps in the state of knowledge exist, both in theory and in accumulated data. Basically, three shortcomings in the literature account for this. First, there are conceptual and theoretical limitations, especially at the intermediate levels of abstraction that connect abstract explanatory schemes—of which a notable few exist—with concrete events. Second, the preponderance of empirical work in this area is in the form of case studies that are mostly historical. These case studies offer uniformity neither in concepts nor in

evidence. The third problem with literature on ethnic nationalism is its fragmentation. A coherent picture must draw upon concepts and propositions from a number of traditions, and there are many overriding concepts with which to assemble frameworks to advance theory and to guide the collection of evidence. While they are difficult to plan and execute, there is compelling need for comparative studies to advance understanding of the underlying principles. Clear understanding of nationalism, and of the processes involved, is essential to the evolving of appropriate educational, policy, constitutional, and other legal means for accommodating cultural diversity.

(SEE ALSO: *Ethnicity*; *Social Movements*)

REFERENCES

- Alger, Chadwick F. 1998 "Conclusion: The Potential of the United Nations System." In Chadwick F. Alger, ed., *The Future of the United Nations System*. Tokyo: United Nations University Press.
- Black, Cyril E. 1966 *The Dynamics of Modernization: A Comparative Study of History*. New York: Harper and Row.
- Connor, Walker 1972 "Nation-Building or Nation-Destroying?" *World Politics* 24:319-355.
- 1978 "A Nation Is a Nation, Is a State, Is an Ethnic Group, Is a . . ." *Ethnic and Racial Studies* 1(4):377-400.
- 1984 *The National Question in Marxist-Leninist Theory and Strategy*. Princeton, N.J.: Princeton University Press.
- Deutsch, Karl W. 1953 *Nationalism and Social Communication*. Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Encyclopedia of World History* 1998 New York: Oxford University Press.
- Geertz, Clifford 1963 "The Integrative Revolution: Primordial Sentiments and Civil Politics in the New States." In Clifford Geertz, ed., *Old Societies and New States: The Quest for Modernity in Asia and Africa*. New York: Free Press.
- Gurr, Ted Robert 1970 *Why Men Rebel*. Princeton, N.J.: Princeton University Press.
- Gusfield, Joseph R. 1970 "Introduction: A Definition of the Subject." In Joseph R. Gusfield, ed., *Protest, Reform, and Revolt: A Reader in Social Movements*. New York: John Wiley.
- Hechter, Michael 1986 "Theories of Ethnic Relations." In John F. Stack, Jr., ed., *The Primordial Challenge: Ethnicity in the Contemporary World*. New York: Greenwood.
- Heraclides, Alexis 1990 "Secessionist Minorities and External Involvement." *Internal Organization* 44(3):341-378.
- Hermann, Margaret G. 1986 "Ingredients of Leadership." In Margaret G. Hermann, ed., *Political Psychology: Contemporary Problems and Issues*. San Francisco: Jossey-Bass.
- Houseknecht, Sharon K. 1990 "The Role of Family and Kinship in Ethnic Nationalist Movements." Columbus: Mershon Center, Ohio State University. Mimeograph.
- Huntington, Samuel P. 1968 *Political Order in Changing Societies*. New Haven: Yale University Press.
- Islam, Rafiqul 1985 "Secessionist Self-Determination: Some Lessons from Katanga, Biafra and Bangladesh." *Journal of Peace Research* 22(3):211-221.
- Jenkins, J. Craig 1983 "Resource Mobilization Theory and the Study of Social Movements." *Annual Review of Sociology* 9:527-553.
- Kaiser, Robert A. 1994 "Ethnic Demography and Interstate Relations in Central Asia." In Roman Szporluk, ed., *National Identity and Ethnicity in Russia and the New States of Eurasia*. Armonk, N.Y.: M. E. Sharpe.
- Kamenka, Eugene, ed. 1973 *Nationalism, the Nature and Evolution of an Idea*. London: Edward Arnold.
- Kissinger, Henry 1999 "New World Disorder." *Newsweek* (May 31):41-43.
- Kohn, Hans 1955 *Nationalism*. Princeton, N.J.: D. Van Nostrand.
- 1968 "Nationalism." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 11. New York: Macmillan/Free Press.
- Koslov, Victor 1975 *Natsional'nosti SSSR*. Moscow: Statistika: 86-88.
- Lasswell, Harold D. 1936 *Politics: Who Gets What, When, How*. New York: McGraw-Hill.
- Lijphart, Arend 1977 *Democracy in Plural Societies: A Comparative Exploration*. New Haven, Conn.: Yale University Press.
- Linz, Juan J. 1978 "The Breakdown of Democratic Regimes: Crisis Breakdown and Reequilibrium." In Juan J. Linz and Alfred Stepan, eds., *The Breakdown of Democratic Regimes*. Baltimore, Md.: Johns Hopkins University Press.

- McCarthy, John D., and Mayer N. Zald 1977 "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82:1212-1241.
- Minogue, K. R. 1967 *Nationalism*. New York: Basic.
- Nagi, Saad Z. 1992 "Ethnic Identification and Nationalist Movements." *Human Organization* 51(4):304-317.
- Naimark, Norman M. 1993 "Forward." In Ronald G. Suny, *The Revenge of the Past: Nationalism Revolution, and the Collapse of the Soviet Union*. Stanford, Calif.: Stanford University Press.
- Nordlinger, Eric 1972 *Conflict Regulation in Divided Societies*. Cambridge, Mass.: Center for International Affairs, Harvard University.
- Oberschall, A. 1973 *Social Conflict and Social Movements*. Englewood Cliffs, N.J.: Prentice-Hall.
- Pankhurst, Jerry G. 1988 "Muslims in Communist Nations: The Cases of Albania, Bulgaria, Yugoslavia, and the Soviet Union." In Anson Shupe and Jeffery Hadden, eds., *The Politics of Religion and Social Change: Religion and Political Order*, vol. 2. New York: Paragon.
- Plamenatz, John P. 1973 "Two Types of Nationalism." In Eugene Kamenka, ed., *Nationalism, the Nature and Evolution of an Idea*. London: Edward Arnold.
- Said, Abdul, and Luis Simmons 1976 *Ethnicity in the International Context*. New Brunswick, N.J.: Transaction.
- Sanders, David 1981 *Patterns of Political Instability*. London: Macmillan.
- Sigler, Jay A. 1983 *Minority Rights: A Comparative Analysis*. Westport, Conn.: Greenwood.
- Smelser, Neal 1962 *Theory of Collective Behavior*. New York: Free Press.
- Smith, Anthony D. 1976 "Introduction: The Formation of Nationalist Movements." In Anthony D. Smith, ed., *Nationalist Movements*. London: Macmillan.
- 1981 *The Ethnic Revival*. Cambridge [England]; New York: Cambridge University Press.
- Snyder, Louis L. 1984 *Macro-Nationalism: A History of the Pan-Movements*. Westport, Conn.: Greenwood.
- Stack, John F., Jr. 1986 "Ethnic Mobilization in World Politics: The Primordial Perspective." In John F. Stack, Jr., ed., *The Primordial Challenge*. New York: Greenwood.
- Symons-Symonolewicz, Konstantin 1970 *Nationalist Movements: A Comparative View*. Meadville, Pa.: Maplewood.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.

Williams, Robin M., Jr. 1947 *The Reduction of Intergroup Tensions*. New York: Social Science Research Council.

Young, M. Crawford 1975 "Nationalism and Separatism in Africa." In Martin Kilson, ed., *New States in the Modern World*. Cambridge, Mass.: Harvard University Press.

SAAD Z. NAGI

## NATURE VS. NURTURE

See Gender; Intelligence; Sex Differences; Socialization.

## NEGOTIATION OF POWER

In ordinary usage, *negotiation* signifies contracting, bargaining, and attempting to reach an agreement. Many studies of negotiation have been carried out particularly with respect to commercial, diplomatic, political, and trade union relations. Recently, acquisitions from the study of negotiation (modalities, styles, strategies, and so on) have been applied to the analysis of interpersonal and social dynamics, ranging from one-to-one relationships to more complex social situations and the management of organizations. Negotiation has also found a specific field of application in the solution of conflicts (Pruitt and Carnevale 1993).

Negotiation is, briefly, a process in which two or more parties try to reach a satisfactory solution to a shared problem. To be more specific, it is a process in which the actors define their own obligations, costs, and benefits to achieve a common result. It is important to emphasize that negotiation is a process, that is to say, it is a dynamic—and not an instantaneous situation—that may be adopted as a way of regulating relationships. Negotiation is a process of exchange (of information, threats, favors, and so on) that goes on until compromises beneficial for all parties involved begin to become apparent. It is a process that advances cautiously and methodically, so that the interests and the expectations of the parties involved are able to emerge gradually. Only when the parties have succeeded in deciphering the real interests and intentions that lie behind their respective declared positions can possible solutions be identified.

Negotiation processes can be analyzed from various points of view, using various techniques. It is possible, for example, to use macrostructural variables. In this case, the outcome is explained on the basis of variables in the context where the negotiation is carried out. Alternatively, it is possible to use psychological variables, interpreting the result of negotiation in the light of the personal characteristics of the actors and of the psychodynamics that develop between them. A third possibility is the "strategic" approach. A strategic analysis does not ignore context-defined effects and limits on the negotiation process nor the importance of the psychodynamic relationship established by the participants but it assumes that the outcome of negotiation will be primarily the result of strategic actions by the negotiators. According to this approach, a negotiation process can only be reconstructed by starting from the actual choices made from the alternative strategies available by actors in specific situations.

It is in this perspective that the special case of "interpower" negotiation (where the term "power" is used as an abbreviation for "institutional power" or "branch of government") has emerged. The topic has become important because in modern political systems the traditional concept of such powers has been modified; because the opportunities for conflict between them have multiplied; and finally because an increasing number of social and economic issues involving values of fundamental importance for individuals and communities have become the target of decisions, interventions, or actions taken or promoted by the various institutional powers.

During the twentieth century, the state has everywhere become increasingly interventionist, while respecting, in democratic societies, traditional civil and political freedoms. State intervention involves, on one hand, the economy and, on the other, various aspects of everyday life. In the economy, the state regulates private initiative relatively strictly, striving to guarantee the "social rights" of less favored groups without yielding to socialist collectivism. As well as impacting, sometimes significantly, on the lives of individuals, groups, communities, and enterprises, the intervention of the contemporary state upsets the parameters of the classic division of powers (legislative, executive, judicial), threatening its survival.

Indeed, it has led to the gradual rise of a new pattern of division of powers.

State intervention involves a massive increase in the production of regulations and therefore in the opportunities for violating such regulations and promoting legal action. It also exacerbates competition and conflicts among powers, among the various branches of state administration, and between decision-making centers and citizens (Field 1996). Normative activity has expanded so much that it can no longer remain the prerogative of parliament. For practical reasons, it is widely delegated to the executive and to administrative authorities. Moreover, a first decision on many conflicts must be referred to the administrative authorities, so that the classic principle of the exclusive attribution of fundamental functions to separate, corresponding powers is violated. The principle of the clear-cut division of powers has ceased to exist.

On the other hand, the very nature of public administration has changed. It not only carries out normative functions but also possesses wide-ranging discretionary power in defining and safeguarding the public interest. To do this, either official decrees or contractual instruments may be used.

Public administration has become a huge—indeed excessively large, according to many commentators—bureaucratic body. Formally, it continues to be subject to the law (which often leaves the administration a high degree of freedom for discretionary action) and to the directives of the executive. But there are many fields in which the law itself concedes considerable autonomy from the executive to administrative organs.

In addition, the nature of the work of the judiciary has changed. The system of sources for precedent is getting more complex and the number of regulations is mushrooming. Interpretation by judges therefore must involve strongly creative elements. There has also arisen a new area of intervention connected with the increasing number of functions assigned to public administration and the growing role of government. Judges are prompted to set aside their old self-image of being the "voice of the law" and to undertake an independent "political" role, inspired partly by new developments in legal theory, such as the law of interests and values, the sociology of law, and judicial pragmatism.

At the same time, Western states are experiencing two new and extremely important functions that are not comparable to the traditional ones. The first is that of *political orientation*. The intervention of the state must necessarily be organized around a program if it is to have significance and coherence. The program must be one on which the electorate expresses itself at elections and which an organ of the state has the task of translating into concrete provisions for the period of office of the legislature (at the next election, the electorate will decide whether to continue or to change program and governing team).

The body of the state delegated to administer the *political orientation* could only be the former executive. The executive, however, having been charged with its new task, has completely transformed its structures and authority. It has changed from a mainly executive body into a *governing power* that lasts for at least one legislature. Generally, it is the main initiator of the laws necessary to implement the *political orientation* and is followed in such action by both the legislative power and the public administration.

The second new function is that of guaranteeing the principles of the constitution. The United States provided a valid example in this respect, although for many years it was not followed by other countries. The much faster rate of change of the normative system makes it more likely that the laws themselves may violate fundamental values, as well as personal and social rights that the community wants to maintain intact. Parliaments have ceased to be reliable champions of the freedoms and the rights of the citizen. Frequently, they have been transformed into assemblies that merely represent sectarian or corporate interests. For that reason, there has emerged a demand to entrust the defense of those values and freedoms to a body that does not have to face a periodic electoral test. Constitutional courts were therefore created in nearly all political systems of the twentieth century. The interventionist state does not disclaim the division of power as such. The division must remain effective to avoid an excessive or even total concentration of power that would threaten citizens' civil and political freedoms and social rights, or would jeopardize the proper, efficient government of society and the economy. But as we have

noted, the modern state can no longer be considered a harmonic whole made up of separate powers organized around the three functions of legislative, executive, and judicial power.

The division of the powers remains, but it has assumed a different form (O'Toole 1993). It has become a Weberian ideal-type with new features that can be summarily described as follows. There are at least five powers instead of the traditional three:

1. *The governing power* carries out the function of political orientation, defining the main policies. Whether it is embodied in a president, a prime minister, a chancellor, or some other figure, the governing power is an organ that enjoys guaranteed stability, which is indispensable for translating political orientation into action. In general, the governing power will attempt to perform a broader normative function than the one attributed to it, thus entering into conflict or competition with the parliament.
2. *The legislative branch* has taken on the function of confirming the political orientation decided by the governing power when its majority belongs to the same party. Where the majority belongs to a different party, the function is more political and relationships are competitive, because the overriding aim is to win the next election. Nevertheless, obstruction of the policies of the governing power, or its formal removal, must remain absolutely exceptional eventualities. The legislative branch also has other checks and controls over the governing power (formal questioning, inquiries, and so on).
3. *Public administration* has become one of the powers because, although it is subject to law and the directives of the governing power, its sheer size and the growing range of functions assigned to it enable the public administration to pursue objectives in the public interest which it identifies independently with discretionary freedom. The aims of the public administration are achieved by regulation and sanctions.



4. *The judicial power* is the power that guarantees in complete autonomy the proper application of the law to specific cases. The judiciary contributes creatively to the evolution of the normative system. For this reason, it may perform, intentionally or not, an important "political" role.
5. *The supreme court (constitutional court)* has the function of guaranteeing supreme political and social values for all the other powers, especially the governing and legislative powers. These values have to be removed from the cut and thrust of day-to-day politics and constantly defended. In fact, the supreme court identifies these values with wide discretion.

The new model of the division of powers places the governing—and not the legislative—power at the center of the state's organization. It is the governing power that provides the driving force. The other powers act as potential brakes, the legislative branch on the political front and the supreme (constitutional) court on the legal side.

The governing power cannot therefore be compared to a medieval absolute (albeit elective) monarch. Still less can it be compared to the dictators of contemporary authoritarian and totalitarian regimes. It acts in a context of free democracy and in a system of the division of powers that places limits on it and prevents it from taking the place of other powers in the exercise of functions that do not properly belong to it. Realistically, its position within the organization of the state is probably less strong than might appear from the above model.

If we examine what has happened to the former legislative, executive, and judicial functions that today are no longer attributed to or exercised by a specific power, it will become apparent that they are distributed among the new powers without the rigid distinctions that once prevailed. Although every political system has mechanisms for blocking democracy-threatening initiatives (for example, penal norms can be only introduced by acts of parliament), the vagueness and fluidity of the new political situation demands the introduction of new procedures to regulate the relationships of the new powers. New procedures are necessary to

avoid the risk of permanent conflict, the blocking of all political plans, and the struggle to occupy ever-wider areas of influence.

Almost all the major Western nations have, to some extent, adapted to the new model. Of course, each system has chosen its own way of allowing a governing power to emerge and of defining a corresponding series of complementary authorities and checks. Indeed, a surprising variety of solutions has been devised to ensure the optimal functioning of the interventionist state. Since it is impossible to illustrate here the performance and the details of the models established by the major Western nations, we shall merely list some of the most important changes.

Great Britain in the twentieth century has transformed its classic parliamentary government into government by the prime minister, who represents the governing power. This has been possible thanks to the rigid bipartite nature of the British political system, which guarantees the alternation of parties in power; to the adoption of a first-past-the-post voting system that reinforces the two-party system; and to the strict internal discipline in the parties.

France under the Fifth Republic has overcome the disadvantages of its extreme parliamentarism, which were due to the excessive fragmentation and lack of internal discipline of the parties. This has been achieved through the direct election of the president of the republic, the reinforcement of the position of the government with respect to parliament, and the return to a majority electoral law with two ballots. The governing power in France is the president, except when the parliamentary majority is hostile (policies are decided with the prime minister). The Constitutional Council safeguards the ideals of French constitutionalism. In Germany, the governing power is represented by the chancellor. Elected by majority vote in the Bundestag, the chancellor enjoys remarkable stability, partly because of the principle of the constructive vote of no confidence. It is up to the chancellor to dictate the political orientation, which parliament normally backs. The strength of the German political parties, the federal organization of the state, and an effective constitutional court provide an effective counterweight to the power of the chancellor.

In the nineteenth century, Congress was the real center of American political life. During the twentieth century, however, the presidential system of the United States has become actual rather than nominal, and the true driving force of the nation today is the president. Congress has nevertheless kept a significant portion of its former power and is currently the most powerful of all the Western parliaments. The Supreme Court is also rather more than a mere check on the governing power.

Italy is something of an anomaly because its constitution enshrines the model of the social interventionist state, but the country has not succeeded in developing the structures for a new division of powers. The Italian parliament has jealously preserved all its normative functions, but the public administration interprets and applies the regulations with such a high degree of autonomy that it distorts the intentions of the legislator. The strict two-chamber structure slows down parliamentary activity; the government has only weak prerogatives over parliament; and the judiciary, originally conceived as a mere instrument to apply the laws, has actually been performing a largely political role in recent years. These modifications to the institutional order have been fostered by major changes in the political process, especially concerning the demands and behavior of citizens. Two aspects, in particular, deserve attention. First, increasingly frequently, public policy disputes among social or special-interest groups are taken to the legislator, the administrative branch, or the judicial arena for resolution. Second, more and more people request government intervention to solve private problems or to further group interests. The new powers therefore have to tackle new tasks and satisfy additional social demand.

The breakdown of the traditional division of powers, whose roles, functions and relationships are very clearly defined, at least formally, leads to a fluid situation in which the new powers are forced to negotiate with each other.

There are two aspects that coexist in interpower negotiation. The first is negotiation on matters of substance and—albeit generally in less explicit terms—for the attribution of roles. On fundamental issues, it is necessary to make a distinction between topics that involve nonnegotiable values

and interests on which it is possible to reach a compromise or to find some way of compensating the losing party.

Even though the present era is characterized by the spread of secularism and cultural relativism, there are topics on which people—and therefore institutions—are deeply divided. Abortion is one example, for it involves worldviews and values that are not negotiable. In such cases, the majority will impose its position without attempting to mediate an agreement with the minority. The decision adopted, whatever it is, will obviously produce resentment, the undermining of authority, and conflict. The process of developing a common-ground culture and fostering empathy on sensitive topics, such as abortion, is complex, although it may be effective with small groups of participants (LeBaron and Carstarphen 1997). Another such example is the question of worship in schools. Moreover, people are becoming very sensitive about issues such as environment, health, and safety: Divisions and conflicts go through not only interests but values, ideologies, deep-rooted ideas. If the area of decisions imposed by a majority on the minority or the minorities were to extend, the already serious fractures in society would deepen, with consequences that could threaten the very continuance of coexistence.

In addition, negotiation over the attribution of roles can develop at two different levels, the institutional and the political. The subject of negotiation is always the same, concerning the organizational model to be applied to the powers, the levels of government (federal, state, and local), and the sphere in which each power will exercise its influence. All participants claim to be pursuing the public interest but the context in which negotiation is carried out changes. In the first case, constitutional principles, values, and rights are the benchmarks. In political negotiation, however, concrete policy priority decisions are negotiated as well as the allocation of financial and other resources and standards for crucial areas such as health, public safety, and education.

It is also important to bear in mind that other factors apart from those already mentioned play a part in interpower negotiation, including each actor's unique characteristics and awareness of its source of legitimization. The president, or governing power, will be more sensitive to general public

opinion. The legislative power cannot ignore strong pressures exercised both by constituencies and by lobbies. The judiciary is the guardian of the laws, and the public administration bases its position on technical competence and a professed impartiality. The ensuing system looks rather unstable, so the problem then becomes how to combine diverse, conflicting interests, styles, and sensitivities, which all claim to be pursuing a common goal: the public interest and the satisfaction of individual and social needs.

In order to achieve this agreement, it is necessary, first of all, to establish common definitions of what is meant by the "public interest" and by "social and individual needs". Second, negotiating procedures must be agreed upon since the checks-and-balances principle, a feature of political systems with a traditional division of powers and well-defined functions, is no longer adequate. The checks-and-balances principle loses much of its value in a continuously changing situation with different actors. Third, it is necessary to seek the consent of the governed to clear the way for negotiations, enabling citizens to be informed and to exercise some kind of control (Susskind and Cruikshank 1989).

Interpower negotiation should not aim for minimum levels of agreements but should seek to create bargains that improve the working of the political system and produce a genuine gain for citizens. In contrast with normal conditions of negotiation, when powers are involved it is not appropriate to seek the victory of one or other of the parties. Instead, the actors must perceive that they should exploit their differences to achieve benefits jointly.

Taking advantage of differences of interest does not mean helping the other party in order to receive help in turn (*do ut des*) or making a simple exchange. It signifies, above all, learning how to recognize the priorities, responsibilities, and interests of each actor. The first problem to be resolved in interpower negotiation is to identify a limit beyond which it is in no one's interest to go. In the literature on the subject, some analysts propose "best alternatives to nonagreement" (BATNA) as a criterion for verifying the motivation and interests underlying negotiation (Fisher and Ury 1981). Negotiation is thus the key to improvement. If

powers wage battle, each trying to impose its own interests and strengthen its own position, the result may be to block the political system, thus losing the confidence of citizens. Conflict between the courts, the political authorities, and the public administration is not the most effective way to deal with the breakdown of the traditional division of powers.

## REFERENCES

- Aharoni, Y., ed. 1997 *Changing Roles of State Intervention in Services in an Era of Open International Markets*. Albany: State University of New York Press.
- Baruch Bush, R. A., and J. P. Folger 1994 *The Promise of Mediation*. San Francisco: Jossey-Bass.
- Bercovitch, J., and J. Rubin, eds. 1992 *Mediation in International Relations: Multiple Approaches to Conflict Management*. New York: St. Martin's Press.
- Breslin, J., and S. Rubin, eds. 1995 *Negotiation Theory and Practice*. Cambridge, Mass.: Program on Negotiation at Harvard Law School.
- Dukes, F. E. 1996 *Resolving Public Conflict: Transforming Community and Governance*. Manchester, England: Manchester University Press.
- Fisher, R., and W. Ury 1981 *Getting to Yes: Negotiation Agreement without Giving In*. Boston: Houghton Mifflin.
- Goodwin, C. S. 1996 *The Arc of the Pendulum: A Philosophy for Government in the 21st Century*. Lanham, Md.: University Press of America.
- Hanson, R. L. 1998 *Governing Partners: State-Local Relations in the United States*. Boulder, Colo.: Westview.
- Lavinia, H. *Negotiation, Strategies for Mutual Gain*. Beverly Hills, Calif.: Sage.
- Lax, D. A., and J. K. Sebenius 1986 *The Manager as Negotiator: Bargaining for Cooperation and Competitive Gain*. New York: Free Press.
- LeBaron, M., and N. Carstarphen 1997 "Negotiating Intricable Conflict: The Common Ground Dialogue Process and Abortion." *Negotiating Journal* (October):341-363.
- Leowenstein, K. 1965 *Political Power and the Governmental Process*. Chicago: University of Chicago Press.
- Lewicki, R. J., D. M. Saunders, and J. W. Minton 1997 *Essentials of Negotiation*. Homewood, Ill.: Richard D. Irwin.
- O'Toole, L. J., ed. 1993 *American Intergovernmental Relations: Foundations, Perspectives, and Issues*. Washington, D.C.: Congressional Quarterly Press.

- Pruitt, D. G., and P. J. Carnevale 1993 *Negotiation in Social Conflict*. Pacific Grove, Calif.: Brooks/Cole.
- Raiffa, H. 1985 *The Art and Science of Negotiation*. Cambridge, Mass.: Harvard University Press.
- Simeon, R. 1985 *Division of Powers and Public Policy*. Toronto: University of Toronto Press.
- Susskind, L., et al. 1999 *Negotiating Environmental Health and Safety Agreements*. Washington D.C.: Island Press.
- , and P. T. Field 1996 *Dealing with an Angry Public: The Mutual Gain*. New York: Free Press.
- , and J. Cruikshank 1987 *Breaking the Impasse: Consensual Approaches to Resolve Public Disputes*. New York: Basic Books.
- Ursel, J. 1992 *Private Lives, Public Policy: 100 Years of State Intervention in the Family*. London: Women's Press.
- Weiss, D. S. 1996 *Beyond the Walls of Conflict: Mutual Gains Negotiating for Unions and Management*. Homewood, Ill.: Irwin Professional.

BRUNO TELLIA

## NETWORK ANALYSIS

See Social Networks.

## NONPARAMETRIC STATISTICS

Parameters are characteristics of the population being studied such as a mean ( $\mu$ ), standard deviation ( $\sigma$ ) or median ( $M_d$ ). Statistics describe characteristics of a sample. Sample mean ( $\bar{x}$ ) and sample median ( $m$ ) are two examples. In traditional or classical parametric statistics, making inferences from a sample or samples to a population or populations necessitates assumptions about the nature of the population distribution such as a normal population distribution.

*Nonparametric statistics* may be defined as statistical methods that contribute valid testing and estimation procedures under less stringent assumptions than the classical parametric statistics. However, there is no general agreement in the literature regarding the exact specification of the term “nonparametric statistics.” In the past, nonparametric

statistics and the term “distribution-free methods” were commonly used interchangeably in the literature, although they have different implications. Nonparametric statistics were often referred to as distribution-free statistics because they are not based on a particular type of a distribution in the population such as a normal distribution. It is, however, generally agreed that nonparametric statistics require fewer, or less stringent, assumptions about the nature of the population distribution being studied, compared to classical parametric statistics. It is important to note that nonparametric statistics are not totally assumption-free either, since they must be based on some assumptions applicable to specific tests, such as lack of ties or the independence of two random samples included in the study. They are flexible in their use and are generally quite robust. Nonparametric statistics allow researchers to arrive at conclusions based on exact probability values and the confidence interval.

The major advantages of nonparametric statistics compared to parametric statistics are that: (1) they can be applied to a large number of situations; (2) they can be more easily understood intuitively; (3) they can be used with smaller sample sizes; (4) they can be used with more types of data; (5) they need fewer or less stringent assumptions about the nature of the population distribution; (6) they are generally more robust and not often seriously affected by extreme values in data such as outliers; (7) they have, in many cases, a high level of asymptotic relative efficiency compared to the classical parametric tests; (8) the introduction of jackknife, bootstrap, and other resampling techniques has increased their range of applicability; and (9) they provide a number of supplemental or alternative tests and techniques to currently existing parametric tests.

Many critics of nonparametric tests have pointed out some major drawbacks of the tests: (1) they are usually neither as powerful nor as efficient as the parametric tests; (2) they are not as precise or as accurate as parametric tests in many cases (e.g., ranking tests with a large number of ties); (3) they might lead to erroneous decisions about rejecting or not rejecting the null hypothesis because of lack of precision in the test; (4) many of these tests utilize data inadequately in the analysis because

they transform observed values into ranks and groups; and (5) the sampling distribution and distribution tables for nonparametric statistics are too numerous, are often cumbersome, and are limited to small sample sizes. The critics also claim that new parametric techniques and the availability of computers have reduced the need to use nonparametric statistics. Given these advantages and disadvantages, how does one choose between parametric and nonparametric tests?

Many criteria are available for this purpose. Statistical results are expected to be unbiased, consistent, or robust, or to have a combination of these characteristics. If both tests under consideration are applicable and appropriate, what other criteria may be used in the choice of a preferred test? The power and efficiency of a test are very widely used in comparing parametric and nonparametric tests.

The power of a test is the probability of rejecting the null hypothesis  $H_0$  when it is false. It is defined as  $1 - \beta$ , where  $\beta$  is the probability of a Type II error, or acceptance of a false null hypothesis. If test A and test B would have the same alpha level of significance and the same sample sizes, and test B has a higher power, then test B is considered a preferred choice. However, calculations of the power of a test are often cumbersome. Another way of addressing the issue of the power of a test is through an asymptotic relative efficiency measure. The relative efficiency of a test in comparison to another test is the ratio of the sample size needed to achieve the same power with similar assumptions about significance levels and population distribution. Asymptotic relative efficiency of a test (sometimes called Pitman efficiency) is measured without a limit on the increase in sample size of that test. Thus, for example, when an asymptotic efficiency of Kendall's  $\tau$  relative to Personian  $r$  is given as .912, it means that the  $\tau$  with a sample size of 1000 is as efficient as an  $r$  with a sample size of 912. The test for asymptotic relative efficiency is now routinely used in the literature to compare two tests.

Based on these criteria, it can be generalized that nonparametric tests and techniques tend to be preferred when: (1) nominal or ordinal data are involved; (2) robust tests are needed because of outliers in the data; (3) probability distribution or

density function of the data is unknown or cannot be assumed, as in the case of small sample sizes; (4) the effects of violations of specific test assumptions are unknown, or only weak assumptions can be made; (5) analogous parametric tests are not available, as in the case of runs tests and goodness-of-fit tests; and (6) preliminary trials, supplemental or alternative techniques for parametric techniques are needed.

### NONPARAMETRIC STATISTICAL LITERATURE

The term *nonparametric statistics* was introduced in the statistical literature in the 1940s (Noether 1984). The historical roots of nonparametric statistics are supposed to go as far back as 1710 when John Arbuthnot did a study that included the proportion of male births in London (Daniel 1990; Hettmansperger 1984) using a sign test which is a nonparametric procedure. Spearman's rank correlation coefficient, which is widely used even today, was the forerunner of the chi-square test, which was introduced in 1900.

Initial major developments in the field that introduced and popularized the study of nonparametric methods appeared in the late 1940s and early 1950s, though a few nonparametric tests were published in the late 1920s such as the Fisher-Pitman test, Kendall's coefficient of concordance, and Hotelling and Pabst's paper on rank correlation in the 1930s. Contributions by Wilcoxon and Mann-Whitney were introduced during the 1940s. The term *nonparametric* reportedly appeared for the first time in 1942 in a paper by Wolfowitz (Noether 1984). The 1956 publication of Siegel's text on nonparametric statistics popularized the subject to such an extent that the text was one of the most widely cited textbooks in the field of statistics for a few years. It was during this time that computer software for nonparametric tests was introduced.

The major emphasis in the 1960s and 1970s was on the extension of the field into regression analysis, and analysis of variance to develop tests analogous to the classical parametric tests. It was also discovered during this period that nonparametric tests were more robust and more efficient than previously expected. The basis for the asymptotic

distribution theory was developed during this time period.

A number of texts have been published recently (see, e.g., Conover 1999; Daniel 1990; Hollander and Wolfe 1999; Krauth 1988; Neave and Worthington 1988; Siegel and Castellan 1988; Sprent 1989). Some of these texts can be used without an extensive statistical background; they have excellent bibliographies and provide adequate examples of assumptions, applications, scope, and limitations of the field of nonparametric statistics. In addition, the *Encyclopedia of Statistical Sciences* (Kotz and Johnson 1982–1989) and the *International Encyclopedia of Statistics* (Kruskal and Tanur 1978) should serve as excellent sources of reference material pertaining to nonparametric statistics.

The literature on nonparametric statistics is extensive. The bibliography published in 1962 by Savage had approximately 3,000 entries. More recent bibliographies have made substantial additions to that list.

### TESTS AND TECHNIQUES

Nonparametric statistics may be divided into three major categories: (1) noninferential statistical measures; (2) inferential estimation techniques for point and interval estimation of parametric values of the population; and (3) hypothesis testing, which is considered the primary purpose of nonparametric statistics. (Estimation techniques included in the category above are often used as a first step in hypothesis testing.) These three categories include different types of problems dealing with location, dispersion, goodness-of fit, association, runs and randomness, regression, trends, and proportions. They are presented in Table 1 and illustrated briefly in the text.

Table 1, which includes a short list of some commonly used nonparametric statistical methods and techniques, is illustrative in nature. It is not intended to be an exhaustive list. The literature literally consists of scores of nonparametric tests. More exhaustive tables are available in the literature (e.g., Hollander and Wolfe 1999). The six columns in the table describe the nature of the sample, and the eight categories of rows identify the major types of problems addressed in

nonparametric statistics. Types of data used in nonparametric tests are not included in the table, though references to levels of data are made in the text. Tables that relate tests to different types of data levels are presented in some texts (e.g., Conover 1999). A different type of table provided by Bradley (1968) identifies the family to which the nonparametric derivations belong.

The first column in Table 1 consists of tests involving a single sample. The statistics in this category include both inferential and descriptive measurements. They would be used to decide whether a particular sample could have been drawn from a presumed population, or to calculate estimates, or to test the null hypothesis. The next column is for two independent samples. The independent samples may be randomly drawn from two populations, or randomly assigned to two treatments. In the case of two related samples, the statistical tests are intended to examine whether both samples are drawn from the same (or identical) populations. The case of  $k$  (three or more) independent samples and  $k$  related samples are extensions of the two sample cases.

The eight categories in the table identify the main focus of problems in nonparametric statistics and are briefly described later. Only selected tests and techniques are listed in table 1. Log linear analyses are not included in this table, although they deal with proportions and meet some criteria for nonparametric tests. The argument against their inclusion is that they are rather highly developed specialized techniques with some very specific properties.

It may be noted that: (1) many tests cross over into different types of problems (e.g. the chi-square test is included in three types of problems); (2) the same probability distribution may be used for a variety of tests (e.g., in addition to association, proportion, and goodness-of-fit, the chi-square approximation may also be used in Friedman's two-way analysis of variance and Kruskal-Wallis test); (3) many of the tests listed in the table are extensions or modifications of other tests (e.g., the original median test was later extended to three or more independent samples; e.g., the Jonckheere test); (4) the general assumptions and procedures that underlie some of these tests have been extended beyond their original scope (e.g. Hájek's extension of the Kolmogorov-Smirnov test to regression

NONPARAMETRIC STATISTICS

**Selected Nonparametric Tests and Techniques**

Type of Problem	TYPE OF DATA				
	One Sample	Two Independent Samples	Two Related, Paired, or Matched Samples	<i>k</i> Independent Samples	<i>k</i> Related Samples
Location	Sign test  Wilcoxon signed ranks test	Mann-Whitney-Wilcoxon rank-sum test  Permutation test  Fisher tests  Fisher-Pitman test  Terry Hoeffding and van der Waerden/normal scores tests  Tukey's confidence interval	Sign test  Wilcoxon matched-pairs signed rank test  Confidence interval based on sign test  Confidence interval based on the Wilcoxon matched-pairs signed-ranks test	Extension of Brown-Mood median test  Kruskal-Wallis one-way analysis of variance test  Jonckheer test for ordered alternatives  Multiple comparisons	Extension of Brown-Mood median test  Kruskal-Wallis one-way analysis of variance test  Jonckheer test for ordered alternatives  Multiple comparisons   Friedman two-way analysis of variance
Dispersion (Scale Problems)		Siegel-Tukey test  Moses's ranklike tests  Normal scores tests  Test of the Freund, Ansari-Bradley, David, or Barton type			
Goodness-of-fit	Chi-square goodness-of-fit  Kolmogorov-Smirnov test Lilliefors test	Chi-square test  Kolmogorov-Smirnov test		Chi-square test  Kolmogorov-Smirnov test	
Association	Spearman's rank correlation  Kendall's tau <sub>a</sub> tau <sub>b</sub> tau <sub>c</sub>  Olmstead-Tukey test  Phi coefficient  Yule coefficient  Goodman-Kruskal coefficients  Cramer's statistic  Point biserial coefficient	Chi-square test of independence	Spearman rank correlation coefficient  Kendall's tau <sub>a</sub> tau <sub>b</sub> tau <sub>c</sub>  Olmstead-Tukey corner test	Chi-square test of independence  Kendall's Partial rank correlations  Kendall's coefficient of agreement  Kendall's coefficient of concordance	Kendall's coefficient of concordance

*(continued)*

**Selected Nonparametric Tests and Techniques (continued)**

Type of Problem	TYPE OF DATA				
	One Sample	Two Independent Samples	Two Related, Paired, or Matched Samples	k Independent Samples	k Related Samples
Runs and Randomness	Runs test Runs above and below the median Runs up-and-down test	Wald-Wolfowitz runs test			
Regression		Hollander and Wolfe test for parallelism Confidence interval for difference between two slopes		Brown-Mood test	
Trends and Changes	Cox-Stuart test Kendall's tau Spearman's rank correlation coefficient McNemar change test Runs up-and-down test		McNemar Change test		
Proportion and Ratios	Binomial test	Fisher's exact test Chi-square test of homogeneity		Chi-square test of homogeneity	Cochran's Q test

Table 1

analysis and extension of the two-sample Wilcoxon test for testing the parallelism between two linear regression slopes); (5) many of these tests have corresponding techniques of confidence interval estimates, only a few of which are listed in Table 1; (6) many tests have other equivalent or alternative tests (e.g., when only two samples are used, the Kruskal-Wallis test is equivalent to the Mann-Whitney test); (7) sometimes similar tests are lumped together in spite of differences as in the case of the Mann-Whitney-Wilcoxon test or the Ansari-Bradley type tests or multiple comparison tests; (8) some tests can be used with one or more samples in which case the tests are listed in one or

more categories, depending on common usage; (9) most of these tests have analogous parametric tests; and (10) a very large majority of nonparametric tests and techniques are not included in the table.

Only a few of the commonly used tests and techniques are selected from Table 1 for illustrative purposes in the sections below. The assumptions listed for the tests are not meant to be exhaustive, and hypothetical data are used in order to simplify the computational examples. Discussions about the strengths and weaknesses of these tests is also omitted. Most of the illustrations are either two-tailed or two-sided hypotheses at the



0.05 level. Tables of critical values for the tests illustrated here are included in most statistical texts. Modified formulas for ties are not emphasized, nor are measures of estimates illustrated. Generally, only simplified formulas are presented. A very brief description of the eight major categories of problems follows.

**Location.** Making inferences about location of parameters has been a major concern in the field of statistics. In addition to the mean, which is a parameter of great importance in the field of inferential statistics, the median is a parameter of great importance in nonparametric statistics because of its robustness. The robust quality of the median can be easily ascertained. If the values in a sample of five observations are 5, 7, 9, 11, 13, both the mean and the median are 9. If two observations are added to the sample, 1 and 94 (an outlier), the median is still 9, but the mean is changed to 20. Typical location problems include estimating the median, determining confidence intervals for the median, and testing whether two samples have equal medians.

**Sign Test** This is the earliest known nonparametric test used. It is also one of the easiest to understand intuitively because the test statistic is based on the number of positive or negative differences or signs from the hypothesized median. A binomial probability test can be applied to a sign test because of the dichotomous nature of outcomes that are specified by a plus (+) which indicates a difference in one direction or a minus (-) sign which indicates a difference in another direction. Observations with no change or no difference are eliminated from the analysis. The sign test may be a one-tailed or a two-tailed test. A sign test may be used whenever a *t*-test is inappropriate because the actual values may be missing or not known, but the direction of change can be determined, as in the case of a therapist who believes that her client is improving. The sign test only uses the direction of change and not the magnitude of differences in the data.

**Wilcoxon Matched-Pairs Signed-Rank Test** The sign test analysis includes only the positive or negative direction of difference between two measures; the Wilcoxon matched-pairs signed-rank test will also take into account the magnitude of differences in ordering the data.

*Example:* A matched sample of students in a school were enrolled in diving classes with different training techniques. Is there a difference? The scores are listed in Table 2.

*Illustrative Assumptions:* (1) The random sample data consist of pairs; (2) the differences in pair values have an ordered metric or interval scale, are continuous, and independent of one another; and (3) the distribution of differences is symmetric.

*Hypotheses:* A two-sided test is used in this example.

$$\begin{aligned} H_0: & \text{Sum of positive ranks} = \text{sum of negative} \\ & \text{ranks in population} \\ H_1: & \text{Sum of positive ranks} \neq \text{sum of negative} \end{aligned} \quad (1)$$

*Test statistic or procedures:* The differences between the pairs of observations are obtained and ranked by magnitude. *T* is the smaller of the sum of ranks with positive or negative signs. Ties may be either eliminated or the average value of the ranks assigned to them. The decision is based on the value of *T* for a specified *N*. *Z* can be used as an approximation even with a small *N* except in cases with a relatively large number of ties. The formula for *Z* may be substituted when *N* > 25.

$$Z = \frac{T - N(N+1)/4}{\sqrt{N(N+1)(2N+1)/24}} \quad (2)$$

This formula is not applicable to the data in Table 2 because the *N* is < 25 and the calculations in Table 2 will be used in deciding whether to reject or fail to reject (“accept”) the null hypothesis. In this example in Table 2, the *N* is 7 and the value of the smaller *T* is 9.5.

*Decision:* The researchers fail to reject the null hypothesis (or “accept” the null hypothesis) of no difference between the two groups, with an *N* of 7 at the 0.05 level, for a two-sided test, concluding that there is no statistically significant difference in the two types of training at the 0.05 level.

*Efficiency:* The asymptomatic related efficiency of the test varies around 95 percent, based on the sample sizes.

Total Scores for Five Diving Trials						
Pairs	X Team A	Y Team B	Y - X Differences	Signed Rank of Differences $T_+$	Negative Ranks $T_-$	
1	37	35	-2	-1	1	
2	39	46	7	+4		
3	32	24	-8	-5.5	5.5	
4	21	34	13	+7		
5	20	28	8	+5.5		
6	9	12	3	+2		
7	14	9	-5	-3	3	
				$T_+ = 18.5$	$T_- = 9.5$	

Table 2

*Related parametric test:* The *t*-test for matched pairs.

*Analogous nonparametric tests:* Sign test; randomization test for matched pairs; Walsh test for pairs.

*Kruskal-Wallis One-Way Analysis of Variance Test* This is a location measure with three or more independent samples. It is a one-way analysis of variance that utilizes ranking procedures.

*Example:* The weight loss in kilograms for 13 randomly assigned patients to one of the three diet programs is listed in Table 3 along with the rankings. Is there a significant difference in the sample medians?

*Illustrative Assumptions:* (1) Ordinal data; (2) three or more random samples; and (3) independent observations.

*Hypotheses:* A two-sided test without ties is used in this example.

$$\begin{aligned}
 H_0: & \quad Md_1 = Md_2 = Md_3, \text{ The populations} \\
 & \quad \text{have the same median values.} \\
 H_1: & \quad Md_1 \neq Md_2 \neq Md_3, \text{ All the populations} \\
 & \quad \text{do not have the same median value.}
 \end{aligned}
 \tag{3}$$

*Test statistics or procedures:* The procedure is to rank the values and compute the sums of those ranks for each group and calculate the *H* statistic. The formula for *H* is as follows:

$$H = \left( \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{N_i} \right) - 3(N+1) \tag{4}$$

where  $N_i$  = the case in the *i*th category of rank sums  
 $R_i$  = the sum of ranks in the *i*th sample.

$$\begin{aligned}
 H &= \frac{12}{13(13+1)} \left[ \frac{(46)^2}{5} + \frac{(16)^2}{4} + \frac{(29)^2}{4} \right] \\
 &\quad - 3(13+1) = 45.9857 - 42 \\
 H &= 3.99
 \end{aligned}
 \tag{5}$$

*Decision:* Do not reject the null hypothesis, as the chi-square value for 2 df at the 0.05 level is 5.99 and the *H* value of 3.99 is less than the critical value.

*Efficiency:* Asymptotic relative efficiency of Kruskal-Wallis test to *F* test is 0.955 if the population is normally distributed.

*Related parametric test:* *F* test. Analogous nonparametric test(s): Jonckheere test for ordered alternatives.

*Friedman Two-Way Analysis of Variance* This is a nonparametric two-way analysis of variance based on ranks and is a good substitute for the parametric *F* test when the assumptions for the *F* test cannot be met.

*Example:* Three groups of telephone employees from each of the work shifts were tested for their ability to recall fifteen-digit random numbers, under four conditions or treatments of sleep

Diet Programs and Weight-Loss Rankings					
Group 1	Rank	Group 2	Rank	Group 3	Rank
2.8	3	2.2	1	2.9	4
3.5	7	2.7	2	3.1	6
4.0	11	3.0	5	3.7	9
4.1	12	3.6	8	3.8	10
4.9	13				
$R_1 = 46$		$R_2 = 16$		$R_3 = 29$	

Table 3

deprivation. The observations and rankings are listed in Tables 4 and 5. Is there a difference in the population medians?

$$F_{\tau} = \left[ \frac{12}{Nk(k+1)} \sum_{j=1}^k R_j^2 \right] - 3N(k+1)$$

$N =$  number of rows (subjects)

$k =$  number of columns (variables or conditions or treatments)

$R_j =$  sum of ranks in the  $j$ th column

$$F_{\tau} = \left[ \frac{12}{3(4)(4+1)} \right] [(11)^2 + (6)^2 + (3)^2 + (10)^2] - 3(3)(4+1)$$

$$F_{\tau} = (0.20)(266) - 45 = 8.2$$

*Illustrative Assumptions:* (1) There is no interaction between blocks and treatment; and (2) ordinal data with observable magnitude or interval data are needed.

*Hypotheses:*

$$H_0: Md_1 = Md_2 = Md_3 = Md_4. \text{ The different levels of sleep deprivation do not have differential effects.}$$

$$H_1: \text{One or more equality is violated. The different levels of sleep deprivation have differential effects.}$$

*Test statistic or procedures:* The formula and computations are listed above.

*Decision:* The critical value at the 0.05 level of significance in this case for  $N=3$  and  $k=4$  is 7.4. Reject the null hypothesis because the  $F$  value is higher than the critical value. Conclude that the ability to recall is affected.

*Efficiency:* The asymptotic relative efficiency of this test depends on the nature of the underlying population distribution. With  $k=2$  (number of samples), the asymptotic relative efficiency is reported to be 0.637 relative to the  $t$  test and is higher in cases of larger number of samples. In the case of three samples, for example, the asymptotic relative efficiency increases to 0.648 relative to the  $F$  test, and in the case of nine samples it is at least 0.777.

*Related parametric test:*  $F$  test.

*Analogous nonparametric tests:* Page test for ordered alternatives.

*Mann-Whitney-Wilcoxon Test* A combination of different procedures is used to calculate the probability of two independent samples being drawn from the same population or two populations with equal means. This group of tests is analogous to the  $t$ -test, it uses rank sums, and it can be used with fewer assumptions.

*Example:* Table 6 lists the verbal ability scores for a group of boys and a group of girls who are less than 1 year old. (The scores are arranged in ascending order for each of the groups.) Do the data provide evidence for significant differences in verbal ability of boys and girls?

**Scores of Three Groups by Four Levels of Sleep Deprivation**

Conditions	I	II	III	IV
Group 1	7	4	2	6
Group 2	6	4	2	9
Group 3	10	3	2	7

Table 4

*Illustrative Assumptions:* (1) Samples are independent and (2) ordinal data.

*Hypotheses:* A two-sided test is used in this example.

- $H_0$ :  $Md_1 = Md_2$ . There are no significant differences in the verbal ability of boys and girls.
- $H_1$ :  $Md_1 \neq Md_2$ . There is a significant difference in the verbal ability of boys and girls. (10)

*Test statistic or procedures:* Rearrange all the scores in an ascending or descending order (see Table 7). The test statistics are  $U_1$  and  $U_2$  and the calculations are illustrated below.

*Mann-Whitney Wilcoxon U Test* The following formulas may be used to calculate  $U$ .

$$U_1 = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_2 \quad (11)$$

$$U_2 = N_1N_2 + \frac{N_2(N_2 + 1)}{2} - R_1 \quad (12)$$

$$R_1 = 1 + 4 + 5 + 9 = 19$$

$$R_2 = 2 + 3 + 6 + 7 + 8 + 10.5 + 10.5 + 12 + 13 = 72$$

$R_1$  and  $R_2$  refer to the sum of ranks for group 1 and group 2, respectively. (13)

$$U_2 = (4)(9) = [9(9 + 1)/2] - 72 = 9 \text{ and}$$

$$U_1 = 27, \text{ for } U_1 + U_2 = (N_1N_2) = 36$$

*Decision:* Retain null hypothesis. At the 0.05 level, we fail to reject the null hypotheses of no differences in verbal ability. The rejection region

**Rank of Three Groups by Four Levels of Sleep Deprivation**

Ranks	I	II	III	IV
Group 1	4	2	1	3
Group 2	3	2	1	4
Group 3	<u>4</u>	<u>2</u>	<u>1</u>	<u>3</u>
$R_j$	11	6	3	10

Table 5

for  $U$  in this case is 4 or smaller, for sample sizes of 4 and 9 respectively.

*Efficiency:* For large samples, the asymptomatic relative efficiency approaches 95 percent.

*Related Parametric Test:*  $F$  test.

*Analogous Nonparametric Tests:* Behrens-Fisher problem test, robust rank-order test.

$Z$  can be used as a normal approximation if  $N > 12$ , or  $N_1$ , or  $N_2 > 10$ , and the formula is given below.

$$Z = \frac{R_1 - R_2(N_1 - N_2)(N + 1)/2}{\sqrt{N_1N_2(N + 1)/3}} \quad (14)$$

**Dispersion.** Dispersion refers to spread or variability. Dispersion measures are intended to test for equality of dispersion in two populations.

The two-tailed null hypothesis in the Ansari-Bradley-type tests and Moses-type tests assumes that there are no differences in the dispersion of the populations. The Ansari-Bradley test assumes equal medians in the population. The Moses test has wider applicability because it does not make that assumption.

Dispersion tests are not widely used because of the limitations on the tests imposed by the assumptions and the low asymptotic related efficiency of the tests, or both.

**Goodness-of-Fit.** A goodness-of-fit test is used to test different types of problems—for example, the likelihood of observed sample data's being drawn from a prespecified population distribution, or comparisons of two independent samples

**Verbal Scores for Boys and Girls Less than 1 Year Old**

Boys $N_1$ (sample A):	10	15	18	28					
Girls $N_2$ (sample B):	12	14	20	22	25	30	30	31	32

Table 6

being drawn from populations with a similar distribution. The first problem mentioned above is illustrated here using the chi-square goodness-of-fit procedures.

$X_2$ , or the chi-square test, is among the most widely used nonparametric tests in the social sciences. The four major types of analyses conducted through the use of chi-square are: (1) goodness-of-fit tests, (2) tests of homogeneity, (3) tests for differences in probability, and (4) test of independence. Of the four types of tests, the last one is the most widely used. The goodness-of-fit test and the test of independence will be illustrated in this article because the assumptions, formulas, and testing procedures are very similar to one another. The  $X_2$  test for independence is presented in the section on measures of association.

Goodness-of-fit tests would be used in making decisions based on the prior knowledge of the population; for example, sentence length in a new manuscript could be compared with other works of an author to decide whether the manuscript is by the same author; or a manager's observation of a greater number of accidents in the factory on some days of the week as compared to the average figures could be tested for significant differences. The expected frequency of accidents given in table 8 below is based on the assumption of no differences in the number of accidents by days of the week.

*Illustrative Assumptions:* (1) The data are nominal or of a higher order such as ordinal, categorical, interval or ratio data. (2) The data are collected from a random sample.

*Hypothesis:*

*Test Statistic or Procedures:* The formula for calculating this is the same as for the chi-square test of independence. A short-cut formula is also provided and is used in this illustration:

- $H_0$  : The distribution of accidents during the week is uniform.
- $H_1$  : The distribution of accidents during the week is not uniform.

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \tag{15}$$

$$\chi^2 = \sum \frac{(f_o)^2}{f_e} - N \tag{16}$$

The notation  $f_o$  refers to the frequency of actual observations and  $f_e$  is the frequency of expected observations.

$$\chi^2 = \frac{225}{30} + \frac{900}{30} + \dots + \frac{1600}{30} + \frac{2025}{30} - N = 230 - 210 = 20 \tag{17}$$

*Decision:* With seven observations, there are six degrees of freedom. The value for  $\chi^2_{v_2}$  is 12.59 at the .05 level of significance. Therefore, the null hypothesis of equal distribution of accidents over the 7 days is rejected at the .05 level of significance.

*Asymptotic Relative Efficiency:* There is no discussion in the literature about this because nominal data can be used in this analysis and the test is often used when there are no alternatives available. Asymptotic relative efficiency is meaningless with nominal data.

*Related Parametric Test.* *t* test.

*Analogous Nonparametric Tests.* The Kolmogrov-Smirnov one-sample test, and the binomial test for dichotomous variables.

The Kolmogrov-Smirnov test is another major goodness-of-fit test. It has two versions, the one-sample and the two-sample tests. It is different from the chi-square goodness-of-fit in that the

Ranked Verbal Scores for Boys and Girls Less than 1 Year Old													
Scores:	10	12	14	15	18	20	22	25	28	30	30	31	32
Rank:	1	2	3	4	5	6	7	8	9	10.5	10.5	12	13
Comp:	A	B	B	A	A	B	B	B	A	B	B	B	B

Table 7

Kolmogrov-Smirnov test, which is based on observed and expected differences in cumulative distribution functions and can be used with individual values instead of having to group them.

**Association.** There are two major types of measures of association. They consist of: (1) measures to test the existence (relationship) or nonexistence (independence) of association among the variables, and (2) measures of the degree or strength of association among the variables. Different tests of association are utilized in the analysis of nominal and nominal data, nominal and ordinal data, nominal and interval data, ordinal and ordinal data, and ordinal and interval data.

*Chi-Square Test of Independence* In addition to goodness-of-fit,  $\chi^2$  can also be used as a test of independence between two variables. The test can be used with nominal data and may consist of one or more samples.

*Example:* A large firm employs both married and single women. The manager suspects that there is a difference in the absenteeism rates between the two groups. How would you test for it? Data are included in Table 9.

*Illustrative Assumptions:* (1) The data are nominal or of a higher order such as ordinal, categorical, interval, or ratio data. (2) The data are collected from a random sample.

*Hypothesis:*

*Test statistic or procedures:* The formula for  $\chi^2$  is given below. Differences between observed and expected frequencies are calculated, and the resultant value is indicated below.

The expected frequencies are obtained by multiplying the corresponding column marginal totals by row marginal totals for each cell divided by the total number of observations. For example, the expected frequency for the cell with an observed frequency of 40 is  $(100 \times 100)/400=25$ .

$H_0$ : The two variables are independent or there is no difference between married and single women with respect to absenteeism.

$H_1$ : The two variables are not independent (i.e., they are related), or there is no difference between married women and single women with respect to absenteeism. (18)

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \tag{19}$$

$f_o$  – observed frequency,  
 $f_e$  – expected frequency

Similarly, the expected frequency for the cell with an observed frequency of 170 is  $(200 \times 300)/400=150$ .

$$\begin{aligned} &(30 - 25)^2/25 + (70 - 75)^2/75 + (40 - 25)^2/ \\ &25 + (60 - 75)^2/75 + (30 - 50)^2/ \\ &50 + (170 - 150)^2/150 \end{aligned} \tag{20}$$

$$\chi^2 = 1 + .33 + 9 + 3 + 8 + 2.67 = 24$$

$$df = (\text{number of rows} - 1) \times (\text{number of columns} - 1) = (3 - 1)(2 - 1) = 2 \tag{21}$$

*Decision:* As the critical  $\chi^2$  value with two  $df$  is 5.99, we reject the null hypothesis, at the 0.05 level. We accept the alternate hypothesis of the existence of a statistically significant difference in the ratio of absenteeism per year between the two groups of married and single women.

*Efficiency:* The asymptotic relative efficiency of a  $\chi^2$  test is hard to assess because it is affected by the number of cells in the contingency table and the sample size as well. The asymptotic related efficiency of a  $2 \times 2$  contingency table is very low,



**Frequency of Traffic Accidents for One Week during May**

Day	S	M	T	W	T	F	S	Total
Traffic Accidents	15	30	30	25	25	40	45	210
Expected Frequencies	30	30	30	30	30	30	30	210

Table 8

but the power distribution of  $\chi^2$  starts approximating closer to 1 as the sample size starts getting larger. However, a large number of cells in a  $\chi^2$  table, especially with a combination of large sample sizes, tend to yield large  $\chi^2$  values which are statistically significant because of the size of the sample. In the past, Yate's correction for continuity was often used in a  $2 \times 2$  contingency table if the cell frequencies were small. Because of the criticism of this procedure, this correction procedure is no longer widely used. Other tests such as Fisher's Exact Test can be used in cases of small cell frequencies.

*Related Parametric Test:* There are no clear-cut related parametric tests because the  $\chi^2$  test can be used with nominal data.

*Analogous Nonparametric Tests:* The Fisher Exact Test (limited to  $2 \times 2$  tables and small tables) and the median test (limited to central tendencies) can be used as alternatives. In addition, a large number of tests such as phi, gamma, and Cramer's V statistic, can be used as alternatives, provided the data characteristics meet the assumptions of these tests. The  $\chi^2$  distribution is used in many other nonparametric tests.

The chi-square tests of contingency tables allow partitioning of tables, combining tables, and using more than two-way tables with control variables.

The second type of association tests measure the actual strength of association. Some of these tests also indicate the direction of the relationship and the test values in most cases extend from  $-1.00$  to  $+1.00$  indicating a negative or a positive relationship. The values of some other nondirectional tests fall between  $0.00$  and  $1.00$ . Contingency table formats are commonly used to measure this type of association. Among the more widely used tests are the following, arranged by the types of data used: *Nominal by Nominal Data:*

Phi coefficient—limited to a  $2 \times 2$  contingency table. A square of these test values is used to interpret a proportional reduction error.

Contingency coefficient based on the chi-square values. The lowest limit for this test is  $0.00$ , but the upper limit does not attain unity (value of  $1.00$ ).

Cramer's V statistic—not affected by an increase in the size of cells as long as it is related to similar changes in the other cells.

Lambda—the range of lambda is from  $0.00$  to  $1.00$ , and thus it has only positive values.

*Ordinal by Ordinal Data:*

Gamma—uses ordinal data for two or more variables. Test values are between  $-1.00$  and  $+1.00$ .

Somer's D—used for predicting a dependent variable from the independent variable.

Kendall's tau—described in more detail below.

Spearman's rho—described in more detail below.

*Categorical by Interval Data*

Kappa—The table for this test needs to have the same categories in the columns and the rows. Kappa is a measure of agreement, for example, between two judges.

The tests described above are intended for two-dimensional contingency tables. Tests for three-dimensional tables have been developed recently in both parametric and nonparametric statistics.

Two other major measures of association referenced above are presented below. They are Kendall's  $\tau$  (the forerunner of this test is also one

**Annual Absences for Single and Married Women**

Days Absent		Married Women		Single Women		Row Totals
		$f_o$	$(f_e)$	$f_o$	$(f_e)$	
Low	(0–5)	30	(25)	70	(75)	100
Medium	(6–10)	40	(25)	60	(75)	100
High	(11+)	30	(50)	170	(150)	200
Column Totals		100		300		400

Table 9

of the oldest test statistics) and Spearman’s  $\rho$  (one of the oldest nonparametric techniques). They are measures of association based on ordinal data.

$\tau$  is a measure of rank-order correlation. It also tests for independence of the two variables observed. It assumes at least ordinal data. The test scores of  $\tau$  range between  $-1.00$  and  $+1.0$  by convention. A score of  $+1.00$  indicates a perfect agreement between rankings of the two sets of observations. A score of  $-1.00$  indicates a perfect negative correlation which means the ranking in one observation is in the reverse order of the other set of observations. All other scores fall between these two values. The test statistics for data with no ties is given below.

The test statistic:

$$\tau = \frac{S}{n(n-1)/2} \tag{22}$$

$S$  measures the difference between the pairs of observations in natural order. A number of students are ranked according to their mathematical and verbal test scores. The six students are identified by letters A to F. There is no need to know the actual scores, and if actual scores are known, they will be converted to ranks as in Table 10.

What is being measured here with  $\tau$  is the degree of correspondence between the two rankings based on mathematical and verbal scores. The calculations can be simplified if one of the values is arranged in the natural ascending order from the lowest value to the highest value. The lack of correspondence between the two rankings can be measured by just using the second set of rankings which, in this case, is verbal ability.

*Illustrative Assumptions:* (1) The data are at least ordinal. (2) Variables in the study are continuous.

*Hypothesis:*

- $H_0$ : Mathematical and verbal test scores are independent.
- $H_1$ : Mathematical and verbal test scores are related positively or negatively. (23)

The number of concordant pairs is measured by the agreement between the two pairs based on their corresponding standing in their own ranking, and discordant pairs are calculated on the basis of the disagreement between them. The example given below in Table 11 consists of no tied ranks.

$$S = 8 - 7 = 1$$

$$\tau = \frac{S}{n(n-1)/2} = \frac{8-7}{6(5)/2} = \frac{1}{15} = 0.067 \tag{24}$$

*Decision:* With  $n = 6$ , the researcher fails to reject the null hypothesis  $H_0$  at 0.05 level since  $\tau = 0.067$  is smaller than  $\tau = 0.600$ , which is the critical value for a two-tailed test.

*Efficiency:* The asymptotic relative efficiency: Both Kendall’s  $\tau$  and Spearman’s  $\rho$  have an asymptotic relative efficiency of .912 compared to the Personian  $r$  (correlation coefficient) which is a parametric test and can go up to 1.12 in exponential type of distribution data.

*Related Parametric Test:* Personian correlation coefficient (in case of interval or ratio data).

*Analogous Nonparametric Tests:* Spearman’s  $\rho$  (rho).



**Mathematical (X) and Verbal (Y) Test Scores for Six Students**

	A	B	C	D	E	F
X Math	9	15	8	10	12	7
Y Verbal	10	13	9	11	4	12

Table 10

In many studies, however, the same values may occur more than once within  $X$  observations, within  $Y$  observations, or within both observations—in which case there will be a tie. In the case of tied ranks, different formulas for  $\tau_a$ ,  $\tau_b$ ,  $\tau_c$  are used. There is, however, no general agreement in the literature about the usage of subscripts. More detailed treatment of this topic is provided in Kendall and Gibbons (1990). Kendall's  $\tau$  measures the association between two variables, and when there are more than two variables, Kendall's coefficient of concordance can be used to measure the association among multiple variables.

Spearman's  $\rho$  (rho) is another popular measure of a rank-correlation technique. The technique is based on using the ranks of the observations, and a Personian correlation is computed for the ranks. Spearman's rank correlation coefficient is used to test independence between two rank-ordered variables. It resembles Kendall's tau in that regard, and the two tests tend to be similar in their results of hypothesis testing.

**Runs and Randomness.** A run is a gambling term that in statistics refers to a sequence or an order of occurrence of event. For example, if a coin were tossed ten times, it could result in the following sequence of heads and tails: H T H H H T T H H T. The purpose of runs test is to determine whether the run or sequence of events occurs in a random order.

Another type of runs test would be to test whether the run scores are higher or lower than the median scores. A runs up-and-down test compares each observation with the one immediately preceding it in the sequence. Runs tests are not very robust, as they are very sensitive to variations in the data.

As there are no direct parallel parametric tests for testing the random order or sequence of a

series of events, the concept of power or efficiency is not really relevant in the case of runs tests.

**Regression.** The purpose of regression tests and techniques is to predict one variable based on its relationship with one or more other variables. The procedure most often used is to try to fit a regression line based on observed data. Though only a few simple regression techniques are included in Table 1, this is one of the more active areas in the field of nonparametric statistics.

The problems relating to regression analysis are similar to those in parametric statistics such as finding a slope, finding confidence intervals for slope coefficients, finding confidence intervals for median differences, curve fitting, interpolation, and smoothing. For example, the Brown-Mood method for finding a slope consists of graphing the observations on a scatter diagram; then a vertical line is drawn representing the median value, and points on both sides of the vertical line are joined. This gives the initial slope, which is later modified through iterative procedures if necessary. Some other tests are based on an extension of the rank order of correlation techniques discussed earlier.

**Trends and Changes.** Comparisons between, before, and after experiment scores are among the most commonly used procedures of research designs. The obvious question in such cases is whether there is a pattern of change and, if so, whether the trends can be detected. Demographers, economists, executives in business and industry, and federal and state departments of commerce are always interested in trends—whether downward, upward, or existing at all. The Cox-Stuart change test, or test for trend, is a procedure intended to answer these kinds of questions. It is a modification of the sign test. The procedures are very similar to the sign test.

<b>Mathematical (X) and Verbal (Y) Test Scores in Rank Order</b>			
Values (X, Y)	Rankings (X, Y)	(Y is similar in natural order to X)	Y is reverse of natural order of X)
(7, 12)	(1, 5)	1	4
(8, 9)	(2, 2)	3	1
(9, 10)	(3, 3)	2	1
(10, 11)	(4, 4)	1	1
(12, 4)	(5, 1)	1	0
(15, 13)	(6, 6)	<u>0</u>	<u>0</u>
		8	7

Table 11

*Example:* We want to test a statement in the newspaper that there is no changing trend in snowfall in Seattle between the last twenty years and the twenty years before that. The assumptions, procedures, and test procedures are similar to the sign test. These data constitute twenty sets of paired observations. If the data are not chronological, then values below the median and above the median would be matched in order. The value in the second observation in the set is compared to the first and would be coded as either positive or plus, negative or minus, and no change. A large number of positive or negative signs would indicate a trend, and the same number of positive and negative signs would indicate no trend. The distribution here is dichotomous, and, hence, a binomial test would be used. If there are seven positive signs and thirteen negative signs, it can be concluded that there is no statistically significant upward or downward trend in snowfall at the 0.05 level as  $p(k = 7/20/0.50)$  is greater,  $k$  being the number of positive signs. The asymptotic relative efficiency of the test compared to rank correlation is 0.79. Spearman's coefficient of rank correlation and Kendall's Tau may also be used to test for trends.

**Proportion.** The term "portion" or the term "share" may be used to convey the same idea as "proportion" in day-to-day usage. A sociologist might be interested in the proportion of working and nonworking mothers in the labor force, or in comparing the differences in the proportion of people who voted in a local election and the voter

turnout in a national election. In case of a dichotomous situation, a binomial test can be used as a test of proportion; but there are many other situations where more than two categories and two or more independent samples are involved. In such cases, the chi-square distribution may also be viewed as a test of differences in proportion between two or more samples.

Cochran's  $Q$  test is another test of proportions that is applicable to situations with dichotomous choices or outcomes for three or more related samples. For example, in the context of an experimental situation, the null hypothesis would be that there is no difference in the effectiveness of treatments. Restated, it is a statement of equal proportions of success or failures in the treatments.

### NEW DEVELOPMENTS

The field of nonparametric statistics has matured, and the growth of the field has slowed down compared to the fast pace in the 1940s and 1950s. Among the reasons for some of these changes are the development of new techniques in parametric statistics and log linear probability models, such as logit analysis for categorical data analysis, and the widespread availability of computers that perform lengthy calculations. Log linear analysis has occasionally been treated in nonparametric statistics literature because it allows the analysis of dichotomous or ordered categorical data and allows the application of nonparametric tests. As indicated earlier, it was not included in the tests discussed in this article because it is a categorical response analog to regression or variance models.

There is further research being conducted in the field of nonparametric robust statistics. The field of robust statistics is also in the process of developing into a new subspecialty of its own. At the same time, development of new tools—jackknife, bootstrapping, and other resampling techniques that are used both in parametric and nonparametric statistics—has positively impacted the use of nonparametric techniques. Based on the contents of popular computer software, the field of nonparametric statistics cannot yet be considered part of the mainstream of statistics, although some individual nonparametric tests are widely used.

There are many new developments taking place in the field of nonparametric statistics. Only three major trends are referenced here. The scope of application of nonparametric tests is gradually expanding. Often, more than one nonparametric test may be available to test the same or similar hypotheses. Again, many new, appropriate, and powerful nonparametric tests are being developed to replace or supplement the original parametric tests. Last, new techniques such as bootstrapping are making nonparametric tests more versatile.

As in the case of bootstrapping, many other topics and subtopics in nonparametric statistics, have developed into full-fledged specialties or subspecialties—for example, extreme value statistics, which is used in studying floods and air pollution. Similarly, a large number of new techniques and concepts—such as (1) the influence curve for comparing different robust estimators, (2) *M*-estimators for generalizations that extend the scope of traditional location tests, and (3) adaptive estimation procedures for dealing with unknown distribution—are being developed in the field of inferential statistics. Attention is also being directed to the developments of measures and tests with nonlinear data. In addition, attempts are being made by nonparametric statisticians to incorporate techniques from other branches of statistics, such as Bayesian statistics.

**Parallel Trends in Development.** There are two parallel developments in nonparametric and parametric statistics. First, the field of nonparametric statistics is being extended to develop more tests analogous to parametric tests. Recent developments in computer software include many of the new nonparametric tests and alternatives to parametric tests. Second, the field of parametric statistics is being extended to develop and incorporate more robust tests to increase the stability of parametric tests. The study of outliers, for example, is now gaining significant attention in the field of parametric statistics as well.

Recent advancements have accelerated developments in nonparametric and parametric statistics, narrowing many distinctions and bridging some of the differences between the two types, bringing the two fields conceptually closer.

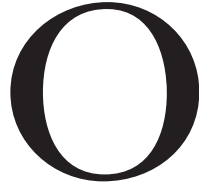
## REFERENCES

- Bradley, James B. 1968 *Distribution-Free Statistical Tests*. Englewood Cliffs, N.J.: Prentice-Hall.
- Conover, W. J. 1999 *Practical Nonparametric Statistics*, 3rd ed. New York: John Wiley.
- Daniel, Wayne W. 1990 *Applied Nonparametric Statistics*, 2nd ed. Boston: PWS-KENT.
- Kotz, Samuel, and Norman L. Johnson, editors-in-chief 1982–1989 *Encyclopedia of Statistical Sciences*, vols. I–IX. New York: John Wiley.
- Gibbons, Jean Dixon 1986 “Distribution-Free Methods.” In S. Kotz and N. L. Johnson, eds., *Encyclopedia of Statistical Sciences*, vol. 2. New York: John Wiley.
- Hettmansperger, Thomas P. 1984 *Statistical Inference Based on Ranks*. New York: John Wiley.
- Hollander, Myles, and Douglas A. Wolfe 1999 *Nonparametric Statistical Methods*, 2nd ed. New York: John Wiley.
- Kendall, Maurice G., and Jean Dickinson Gibbons 1990 *Rank Correlation Methods*, 5th ed. London: Oxford University Press.
- Kotz, Samuel, Norman L. Johnson, Campbell, and B. Read 1982–1989 *Encyclopedia of Statistical Sciences*, vols. 1–9 and supplement. New York: John Wiley.
- Krauth, Joachim 1988 *Distribution-Free Statistics: An Application-Oriented Approach*. Amsterdam: Elsevier.
- Kruskal, William H., and Judith M. Tanur, eds. 1978 *International Encyclopedia of Statistics*, vols. I–II. New York: Free Press.
- Neave, H. R., and P. L. Worthington 1988 *Distribution-Free Tests*. London: Unwin Hyman.
- Noether, Gottfried E. 1984 “Nonparametrics: The Early Years—Impressions and Recollections.” *American Statistician* 38(3):173–178.
- Savage I. R. 1962 *Bibliography of Nonparametric Statistics*. Cambridge, Mass.: Harvard University Press.
- Siegel, Sidney 1956 *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.
- Siegel, Sidney, and N. John Castellan, Jr. 1988 *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. New York: McGraw-Hill.
- Sprenst, P. 1989 *Applied Nonparametric Statistical Methods*. London: Chapman and Hall.

SUBHASH R. SONNAD

## NURSING HOMES

See Long Term Care Facilities.



## **OBJECTIVITY**

*See* Epistemology; Measurement; Scientific Explanation.

## **OBSERVATION SYSTEMS**

The most sensitive, sophisticated, and flexible instrument of observation available today is the human being. All the recent advances in technology have not changed this central fact; we can monitor communications, transcribe conversations using language processing software, and conduct computer-assisted content analysis of the results. Still, at some point a person who understands the social context must infer the meaning of the communications. Systematic methods of observation may vary the unit of analysis, shift the boundaries of categories, and adjust the level of judgment allowed, but sociologists are ultimately left with the basic reality of human beings watching other human beings. The role of the methodologist is to make this process more systematic.

Most sociological data are filtered through the perceptions of informants in an idiosyncratic manner. Retrospective accounts of events, opinion polls, and surveys measure the output of the social perception process. Only systematic observation, with valid and reliable instruments, provides a record of the events themselves rather than the retrospective reconstruction of the events. The more rigorously defined the categories, the more confident the researcher can be that the data

reflect the events and not just the biases and preconceptions of the informants.

While some of the systems that are described below were developed for specific purposes such as the observation of business case-study groups or the diagnosis of psychiatric patients, most attempt to capture the full range of social behavior and may thus be applied to a wide range of settings. Not included here are specialized systems that have been developed for single contexts, such as the classroom behavior of small children, the responses of subjects in tightly controlled laboratory experiments, and the evaluation of employees. One fast-food restaurant, for example, has developed a thirty-one-category observation system that managers can use to observe and evaluate counter staff. Items include "There is a smile," "The bag is double folded," and "Change is counted efficiently."

Observation systems have been used for a wide variety of purposes over the years. Early uses included psychiatric diagnosis, job placement, and basic research into group process and development. As corporate assessment centers came into widespread use for the selection of executives, early observation systems reappeared for the analysis of leaderless-group exercises. More recent applications have included research and consulting on team building and training, the evaluation of social workers, the prediction of success and failure of military cadets, the study of leadership networks in large corporations, the evaluation and treatment of problem children in the classroom, the evaluation of psychiatric interventions, the analysis of delinquent behavior, and resocialization,

and consultation on mergers and consolidations (Polley et al. 1988). In the past ten years, direct observation has been in decline. As a result of the high costs in terms of time and money, systematic observation has often given way to retrospective rating systems (Bales 1998). Such methods are, however, a poor substitute for direct observation.

### SINGLE-CATEGORY SYSTEMS

Elliot D. Chapple introduced the *interaction chronograph* in 1940. It was a simple device that consisted of two telegraph keys. Observers were instructed to press key A when person A spoke and key B when person B spoke. A record of the conversation was kept on a moving paper tape. Not surprisingly, inter-rater reliabilities were nearly perfect. Twenty-five years later, the human observers were replaced by voice-activated microphones attached to analog-digital converters (Wiens et al. 1965). That such a simple device could replace the human observer suggests that the systems were not taking full advantage of the observers' capabilities. In reality, the decision to record such objective and basic information simply shifted the burden of interpretation from the observer to the researcher. Elliot Chapple (1940) and his successors developed elaborate schemes for interpreting the patterns of lines and blanks that appeared on their paper tapes. At the peak of its popularity, the interaction chronograph was used for everything from psychiatric diagnosis to employee placement.

### MULTIPLE CATEGORY SYSTEMS

Chapple's work serves as an important benchmark. The near-perfect reliability is achieved at the cost of validity. The observation systems that followed it generally traded off a measure of reliability for greater validity. More meaningful sets of categories will almost certainly be harder to employ with any degree of inter-rater reliability.

Interaction Process Analysis (IPA) was one of the earliest attempts to devise more meaningful categories. Bales (1950) began with an encyclopedic list of behaviors and gradually consolidated them into an elegant set of twelve basic categories. Fifty years later, his original system is still one of the most widely applied observation methods.

The IPA category list was evolving at the same time that Parsons and Bales (1955) were developing a model of family socialization. They saw family leadership in the 1940s and 1950s as divided between the father and the mother. IPA perpetuates this somewhat dated division through its primary dimension; six categories are provided for coding task-oriented behaviors and six are for coding socioemotional behaviors (Table 1). When Slater (1955) suggested that this role differentiation could be extended to a general model of effective leadership in groups, he sparked a continuing controversy. But it must be remembered that this conceptualization preceded any serious consideration of either androgyny or flexibility in sex roles.

Additional symmetries are built into IPA. Three of the six socioemotional categories carry positive affect, and each has a direct counterpart on the negative side. Task-oriented behavior is seen as the process of asking questions and offering answers, though the answers—in the form of suggestions, opinions, and orientation (or information)—may be in response to questions asked or implied. Finally, the functionalist orientation of Parsons and Bales (1955) appears in the identification of six problems faced by groups: communication, evaluation, control, decision, tension reduction, and reintegration.

Following Chapple's lead, Bales developed and marketed a moving paper-tape recording device. The interaction recorder allowed for observation of groups rather than just dyads; the tape was divided into twelve rows and moved past a window at a constant speed so that the observer could write a code, indicating who was speaking to whom, within a category-by-time sector on the tape. This added complexity and enlarged the unit of analysis. The interaction recorder provided a continuous on-off record while IPA recorded discrete acts. The coding unity was, however, kept small. IPA coders often record two or three acts for a single sentence and are expected to record *all* acts. The first serious challenge to Bales's IPA system was Borgatta's Interaction Process Scores (IPS) system (1963). Borgatta argued that the twelve categories failed to make some crucial distinctions. His redefinition of the boundaries resulted in an eighteen-category system that had the advantage of greater precision and the disadvantages of greater

<b>Interaction Process Analysis</b>			
Category	Emotion	Task	Problem
1. Shows solidarity	Positive		Reintegration
2. Shows tension release	Positive		Tension reduction
3. Agrees	Positive		Decision
4. Gives suggestion		Answer	Control
5. Gives opinion		Answer	Evaluation
6. Gives orientation		Answer	Communication
7. Asks for orientation		Question	Communication
8. Asks for opinion		Question	Evaluation
9. Asks for suggestion		Question	Control
10. Disagrees	Negative		Decision
11. Shows tension	Negative		Tension reduction
12. Shows antagonism	Negative		Reintegration

*Table 1*

SOURCE: Adapted from Bales (1950) p. 14.

complexity and a lack of symmetry. The former problem was largely solved by the availability of a self-training workbook (Borgatta and Crowther 1965). Such a manual has never been widely available for IPA.

The next logical step after categorization is the organization of categories. IPA had an implicit organization that was lacking in IPS. This, and the fact that IPA had twelve rather than eighteen categories, made IPA the easier system to learn and use. As Weick (1985) points out, however, there is a problem of "requisite variety." A system for understanding a phenomenon must be at least as complex as the phenomenon itself. Weick uses the metaphor of a camera with variable focal length. In order to photograph objects at twenty different distances, the camera must have at least twenty focal settings or the pictures will not all be of equal clarity. This creates real problems for the interaction chronograph. Clearly, social behavior is more complex than Chapple's on-off category system. Unfortunately for researchers, it is also more complex than IPA's twelve-category system. This creates a dilemma. The mind can hold only so many categories at once; even with twelve categories, most observers tend to forget the rarer ones in an attempt to simplify their task.

## MULTIDIMENSIONAL SYSTEMS

Timothy Leary (1957) proposed one of the first observation systems based primarily on dimensions rather than categories. His interpersonal diagnosis system placed sixteen categories at the compass points of a two-dimensional circumplex. The dimensions, dominance-submission and love-hate, would prove to be the two most common dimensions among the systems that followed. While intended primarily for the diagnosis of psychiatric disorders, the system also identified the "normal," or less intense, variant of each behavior as well as the likely response that each behavior would generate in other people. For example, the general behavior in the direction of dominance is "manage, direct, lead"; this behavior is likely to "obedience" in others. The extreme version of the category is "dominate, boss, order." This falls into the larger category of behavior that is described as "managerial" when in the normal range and "autocratic" when in the abnormal range. More than a method of observation, the interpersonal diagnosis system was a remarkably well-articulated theory of interpersonal relations. Were it not for Leary's well-publicized advocacy of lysergic acid diethylamide (LSD), the system might very well be in common

use today. While the research was briefly resurrected by McLemore and Benjamin (1979), it never had a major impact on either social psychology or psychiatry. Leary claimed that the system worked so well for psychiatric diagnosis that a receptionist trained in its use could provide diagnoses based on observations in the waiting room that rivaled those obtained by psychiatrists after conducting a diagnostic interview.

Leary's two dimensions were theoretically derived. In a 1960 doctoral dissertation, Arthur Couch pioneered the application of factor analysis to interpersonal behavior, thus offering an empirical alternative for the derivation of dimensions. Couch's six dimensions of interpersonal behavior were derived from the factor analysis of a vast amount of data on twelve groups of five undergraduates each. The individuals were given a large battery of personality tests and the twelve groups were observed participating in a wide variety of tasks across five meetings each. While the factor analysis is not independent of the original categories of measurement and observation, Couch's data set was so exhaustive as to deserve credence.

### MULTILEVEL SYSTEMS

The three dimensions of Bales and colleagues' SYMLOG (System for the Multiple Level Observation of Groups) (1979) owe much both to IPA and to Couch's empirical work. They also reflect the legacy of Leary's circumplex. Dominant-submissive (up-down, or U-D) was the first factor from Couch's analysis; in IPA it corresponds to total amount of talking rather than to interaction in any specific set of categories; Leary's interpersonal diagnosis system had already identified the primary dimension as dominant-submissive. In essence, this was also the dimension that Chapple's interaction chronograph measured. Friendly-unfriendly (Positive-negative, or P-N), Couch's second factor, was represented in IPA by the three positive and three negative socioemotional categories, and it corresponds to love-hate in Leary's system. Task-oriented-emotional expressive (forward-backward, or F-B), the controversial distinction from IPA, was a compromise that created one bipolar dimension out of two of Couch's remaining factors. These three dimensions, generally referred to by the code letters shown above, define a three-dimensional conceptual space. Following Leary, Bales

and his colleagues then defined the compass points of the space. In this case, definitions were produced for all twenty-six vectors in the three-dimensional space. Thus, a dominant, unfriendly, task-oriented act (UNF) is defined as "authoritarian and controlling," and a submissive, unfriendly, emotional act (DNB) is described as "withdrawn and alienated." This was a creative solution to the dilemma of providing requisite variety while keeping the number of categories reasonable. While there are twenty-six categories, they are organized into three dimensions, so coders can hold the three-dimensional space rather than the twenty-six categories in mind. The twenty-six category descriptions then become a reference to be used when first learning the system and trying to understand where specific behaviors fit.

In addition to the basic level of behavior described above, SYMLOG also provides definitions for the twenty-six vectors on the level of nonverbal behavior. This level is coded when unintentional messages are sent through nonverbal behaviors or when the nonverbal—or paralinguistic—cues are at variance with the overt verbal cues. In developing his descriptions of facial expression and nonverbal cues, Bales drew on the work of the eighteenth-century French Encyclopedists. He found that Diderot and his colleagues had developed a more sophisticated understanding of facial expression and nonverbal nuance than have modern social scientists.

SYMLOG also allows for the coding of verbal content. Theoretically, the two levels of behavior—overt and nonverbal—could be coded without reference to the content of the message. Conversely, the content of messages could be coded from written transcripts without reference to the behaviors of the speakers. The content coding level attends only to evaluative statements. Each statement is first coded PRO (for statements in favor of something) or CON (statements against something). The content of the value statement is then coded in the same three-dimensional space that was used for behavior. Finally, the level of the value statement is recorded. Levels begin with the self and move outward: self, other, group, situation, society, and fantasy.

When the full SYMLOG system is used for coding a conversation, the result is a set of simple



sentences written in code. For example, Figure 1 records a brief exchange that took place from 1:23 to 1:24. Joe ordered Bill to stop contradicting him. The behavior was authoritarian (UNF) and the value statement was against negativity in Bill (CON N BIL). Bill responded in a rebellious manner (UNB) and made a value statement against Joe's authoritarianism (CON UNF JOE). Ron intervened in a "purposeful and considerate" manner (UPF) and made a value statement in favor of reducing the level of conflict (PRO DP GRP). If we see contradictions between the overt and nonverbal behavior, we could add lines such as 123 JOE GRP N DN. This would indicate that underneath Joe's authoritarianism is a note of insecurity and nervousness. (The first "N" in the message stands for "nonverbal," the "A" in each of the messages in Figure 1 stands for "act.") Clearly, the system requires fairly extensive training of coders. Most SYMLOG coders went through a fifteen-week course that was run as a self-analytic group. During this time, they alternated between serving as group participants and retreating behind a mirror to serve as coders. While there remain some universities in which such courses are still taught, the use of this method peaked in the late 1970s and early 1980s. Most work being done with SYMLOG at this time uses a parallel system of retrospective rating which requires no special training. This has resulted in time and cost savings, but at the expense of sacrificing the dynamic moment-to-moment process measurement that the original SYMLOG observation system made possible.

SYMLOG adds depth and sophistication to the coding of social interaction but again enlarges the unit of measurement. While IPA is an act-by-act coding system, SYMLOG is a "salient-act" coding scheme. Since it takes longer to record a SYMLOG observation, the coder must select the most important acts for recording. Clearly, this results in a loss of reliability as disagreements may result not only from two observers interpreting the same act differently but also from two observers selecting different acts for recording. Because of this, two observers are generally sufficient for IPA, but five or more are recommended for SYMLOG. Alternatively, interaction can be coded from video recordings. By going over the record repeatedly and picking up different acts with each pass, a higher percentage of acts can be picked up

by each coder, thus increasing the inter-rater reliability and decreasing the number of coders required. The only problem with this alternative is that SYMLOG does not record intensity. When multiple coders are used, more intense acts are likely to be selected as salient by more coders and will thus be weighted more heavily than the less intense acts that may be picked up by only one or two coders. Using only two coders and allowing them multiple passes in order to pick up a higher percentage of acts eliminates this implicit weighting effect.

The issue of reduced reliability is directly related to the problem of subjectivity. Moreno (1953) was an early critic of IPA; he argued that the observations of nonparticipants were meaningless because they could not possibly comprehend the life of a group to which they did not belong. While his position was extreme, it raised a difficult problem for any system that relies on the observations of outsiders. By standing outside the group, the outsider gains distance and "objectivity." Unfortunately, it is not clear that objectivity has any real meaning when speaking of interpersonal interaction. IPA sidestepped the problem by providing very clear specifications of categories. SYMLOG confronts the problem directly since observers are required to make fairly strong inferences as to the meaning of behavior. The coder is instructed to take the role of the "generalized other," or the "average" group member. When a group is polarized, this may be impossible. Half the group is likely to interpret an act in one way and the other half in a different way. This is another reason for having multiple observers. It is hoped that the various biases of the observers will cancel one another out if five or more people observe. Polley (1979) goes a step further by providing "descriptive reliabilities." Instead of simply reporting a reliability figure, descriptive reliabilities allow for a systematic analysis of observer bias. It is argued that observations tell us as much about the observer as about the observed. This is particularly true when group members are trained to serve as observers, as is the case in the self-analytic groups that are often the subject of SYMLOG observations.

Both IPA and SYMLOG have been repeatedly criticized for their use of a bipolar model of the relationship between "task-oriented" and "emotional" behavior. Two of these critiques have proposed adding a fourth dimension. Hare (1976)



Time	Who	To Whom	Act/Non	Direction	Description	Pro/Con	Direction	Level
123	JOE	BIL	A	UNF	Stop contradicting me!	C	N	BIL
123	BIL	JOE	A	UNB	Don't give me orders.	C	UNF	JOE
124	RON	GRP	A	UPF	I think we should all calm down.	P	DP	GRP

Figure 1. Sample SYMLOG Coding

went back to Parsons' original Adaptation, Goal Attainment, Integration, Latent Pattern Maintenance (AGIL) system and concluded that the task dimension should be divided into two dimensions: serious versus expressive and conforming versus nonconforming. Wish and colleagues (1976) proposed leaving the task-emotion dimension intact, even though it did not seem to be quite bipolar, and adding a fourth dimension: intensity. Polley (1987) has argued that emotionality is already captured in the friendly-unfriendly dimension and that the third dimension should be reserved for recording conventional versus unconventional behavior. This solution contends that the polarization of task and emotion is an artifact from the 1950s and has largely lost its meaning since then. If work is defined as devoid of emotional satisfaction, then task and emotion are bipolar. As soon as we recognize the possibility of having an emotional reaction—positive or negative—to work, the two dimensions become orthogonal. As Stone points out, the task-emotion polarity “implies that most work involves sublimation of the libido, and demands impulse control. Moreover, it becomes difficult to imagine management’s orientation . . . fostering individual creativity as some so-called excellent companies have been able to do” (1988, p. 18). It is becoming increasingly apparent that no “universal” scheme for coding interpersonal interaction exists totally independent of cultural and temporal context.

### SPECIALIZED SYSTEMS

While Leary's system was intended primarily for psychiatric diagnosis, and IPA was originally designed for the coding of case discussion groups at Harvard Business School, the methods described above represent attempts to develop comprehensive and general coding schemes. In addition to these all-purpose methods, observation systems

have been devised for somewhat more specific relationships or channels of communication.

Richard Mann's (1967) sixteen-category system codes only one-way relationships from members to leaders. By narrowly defining the relationships to be observed, Mann was able to provide a much more detailed picture. Coding categories are provided for four types of impulse, four different expressions of affect, three variations of dependency, and five ego states. The additional sophistication is achieved by drastically reducing the number of observed relationships. If we observe a ten-person group using Mann's leader-member relationship scheme, we are looking at nine one-way relationships. If we use one of the all-purpose methods, we are coding forty-five two-way relationships.

The most thoroughly studies channel of communication is probably the nonverbal. While this is a small part of the SYMLOG system, proxemic behavior has been exhaustively categorized by Hall (1963) in terms of posture, orientation of bodies, kinesthetic factors, touch code, visual code, thermal code, olfaction code, and voice loudness. The potential complexity of this mode of communication is further illustrated by the fact that Birdwhistell (1970) developed an equally elaborate system for the coding of movement. While Hall's system codes states, Birdwhistell's codes state-to-state transitions. Each of these coding schemes concentrates on the physical nature of nonverbal behavior. In contrast, Mehrabian's system (1970) attempts to directly record the meaning of the behavior via a three-dimensional model that closely parallels SYMLOG. Again, the difference is in whether meaning is inferred by the observer or deferred to the researcher.

The other channel of communication that has been studied in depth is content. Again, this is a small part of SYMLOG; all content that does not carry a positive or negative evaluation is ignored.

In 1977, Philip Stone and his colleagues published *The General Inquirer*, a computerized content analysis system. While many of the applications involve the coding of written text, it has also been used for the coding of transcribed conversations. In one of the first applications of the method, Dexter Dunphy used *The General Inquirer* to code descriptions that group members had written of recent sessions. The great advantage of the methods is that it allows the user to define a dictionary. Three of the earliest dictionaries were the *Harvard III Psychosocial Dictionary*, the *Stanford Political Dictionary*, and the *Need-Achievement Dictionary*. It would also be possible to develop a SYMLOG value-level dictionary. In the case of a computerized content analysis system, the dictionary designer infers meaning *before* the behavior is coded.

### RECENT DEVELOPMENTS

Observation systems only flourish when supported by a substantial research group. Because of the expense and the need for multiple coders, it is difficult to use observation systems as an isolated researcher. For three decades (the 1950s through the 1970s) such a group existed at Harvard. Bales and his colleagues had a well-equipped observation room and plenty of graduate student labor available. When Bales retired in 1985, the observation laboratory ceased to function as a research center. For most of the 1980s there were really no major research centers conducting observation work.

In the 1990s Joseph McGrath and his colleagues at the University of Illinois established a major research agenda around McGrath's theory of time, interaction, and performance (TIP) (1991). While much less method-driven than Bales's IPA and SYMLOG research groups, McGrath's group uses an observation method known as "time-by-event-by-member pattern observation" (TEMPO) (Futoran et al. 1989).

TEMPO was designed as a general observation method for the analysis of a wide range of group performance settings. It is divided into two sets of categories. The first codes on four performance functions:

1. Propose content—solutions, ideas.
2. Propose process—goals, strategies, acts.

3. Evaluate content—agreement, disagreement, clarification.
4. Evaluate process—agree, modify, disagree.

The two "proposal categories" are further classified as: (1) new—a new proposal; (2) prior—previously proposed; or (3) dictate—repeat content for clarification or emphasis. The two evaluation categories are further classified as: (1) agree, (2) clarify or modify, (3) disagree, or (4) reject/veto.

The second set of categories codes seven nonproduction function categories:

- T. Task digression
- P. Personal comments
  - I. Interpersonal comments
- R. React to experiment
- D. Digressions
- U. Uninterpretable
- S. Silence

McGrath's group devised their own methodology rather than relying on IPA or SYMLOG because they were concerned with both process and activity. They felt that the existing systems did not allow for the level of temporal pattern recognition that they were looking for. Their primary criteria for the system were that it: (1) be time-based, (2) identify individual member behavior, (3) recognize multiple acts within a single speech, (4) allow for multiple simultaneous acts and periods of inactivity, and (5) relate act to task products. Virtually all their observation studies are based on the analysis of videotaped records of interacting groups. The TIP theory and the TEMPO method are central to the line of research that McGrath and his colleagues conducted in the 1990s. *Small Group Research* devoted a special issue to six articles that resulted from a single large research project known as "The JEMCO Workshop Study" (McGrath 1993), but additional work from this research team has resulted in numerous journal articles and books.

Another major research effort that has produced its own observation system is led by Susan Wheelan. The work of Wheelan and her colleagues focused on group development. The Group Development Observation System (GDOS) is based in Wilfred Bion's concept of basic assumption and

work groups (1959). It consists of seven categories (Wheelan et al. 1993):

1. Dependency statement
2. Counterdependency statement
3. Fight statement
4. Flight statement
5. Pairing statement
6. Counter-pairing statement
7. Work statement

Like McGrath's group, Wheelan's group is primarily concerned with temporal patterns in group development. The conceptual underpinnings of the two lines of research are, however, very different. Again, GDOS has resulted in numerous journal articles and books. A special issue of *Small Group Research* was also devoted to this line of research (Wheelan 1999).

While there have been a number of other observation systems to emerge in the past decade, these two represent extended lines of research. Most of the publications that have used observation systems since 1980 have used either modifications of existing systems or idiosyncratic systems that were used once or twice and so did not generate a comparative body of work.

#### A DECADE OF SMALL GROUP RESEARCH

In 1990, *International Journal of Small Group Research* and *Small Group Behavior* merged to form a new international and interdisciplinary journal known as *Small Group Research*. This journal publishes research on small groups that ranges across sociology, psychology, organizational behavior, social work, group psychotherapy, communications, and management information systems. As the only interdisciplinary journal in the field of small groups, it is a good indicator of the popularity of the observation systems in general and also of specific systems. Table 2 presents the results of an inventory of the past decade of research published in *Small Group Research*. In addition to tabulating the total number of articles using observation systems each year, it also indicates whether the articles used direct observation, videotapes, or transcriptions and whether they used one of the

four most popular general-purpose methods (IPA, SYMLOG, TEMPO, and GDOS) or some other method.

Of the 270 articles published in *Small Group Research* over the past decade, 41 used observation methods. IPA and SYMLOG are still in use; the other most common methods were TEMPO and GDOS, described above. However, the majority of articles (28) used idiosyncratic observation methods. While many represent real contributions to the literature, the lack of a common method makes comparisons across studies difficult. The other trend that is evident from Table 1 is that relatively few studies (9) used live observation. The remaining studies were almost equally split between the coding of videotapes and the coding of transcriptions. Perhaps the real loss in this shift of methodology is that so many rely on transcripts. One of the strengths of observation methods is that they capture paralinguistic cues; these data are lost to systems that rely on transcriptions.

#### THE FUTURE OF OBSERVATION SYSTEMS

Observation systems have been used less and less in recent years for a mundane reason: cost. Training observers is time consuming, as is the actual process of observing and coding behavior. At this point, serious research using direct SYMLOG observation is being done at only two or three institutions. In contrast, the much less time-consuming method of SYMLOG retrospective rating (using either the Bales items or the Polley revisions) is currently in use in at least fifty institutions around the world. While there is some indication of a resurgence of interest in direct observation, as evidenced by the studies cited in the preceding sections, it is clear that the method requires a substantial commitment of time and money on the part of the researcher.

While the costs of direct observation are high, it is clear that there are a great many aspects of social behavior that simply cannot be understood without it. Basic research still needs to be done on group development, particularly as it relates to team building in organizations. The effects of various leadership styles and decision-making processes on group functioning are still not thoroughly understood. A wide range of styles is currently

Small Group Research										
Year	No. of Articles	No. Using Observation	No. Using Direct	No. Using Videotape	No. Using Transcripts	IPA	SYMLOG	TEMPO	GDOS	Other
1999*	22	4		1	3			1	1	2
1998	31	5	2	1	2					5
1997	27	2		1	1				1	1
1996	26	2	1		1					2
1995	25	3	1	1	1					3
1994	27	2		1	1					2
1993	28	10	3	5	2		1	4	1	4
1992	26	3	1		2		1			2
1991	28	5		4	1	1				4
1990	30	5	1	3	1	1		1		3
Total	270	41	9	17	15	2	2	6	3	28

Table 2

NOTE: \*February through August.

used by group therapists, but these styles are more often backed by rival schools of thought than empirical evidence.

As with content analysis, the future of observation systems may well lie with computerization. Johansen (1989) coined the term *groupware* to refer to computer systems for the support of groups or teams. To date, most of the systems have provided little more than an "electronic flipchart" for nominal group technique sessions. With advances in artificial intelligence, more sophisticated examples of groupware have been gradually emerging. Speech-recognition programs for the automatic transcription of meetings could greatly reduce the cost of using content analysis. Programs have been developed for recognizing emotional content in speech. If these were to be combined with voice-activated, some automatic scoring of behavior—at least at the paralinguistic level—should be possible. Until technology substantially reduces the costs, a return to the widespread use of observation systems seems unlikely.

## REFERENCES

- Bales, Robert F. 1950 *Interaction Process Analysis*. Chicago: University of Chicago Press.
- 1998 *Social Interaction Systems: Theory and Measurement*. New Brunswick, N.J.: Transaction Publishers.
- , Stephen P. Cohen, and Stephen A. Williamson 1979 *SYMLOG: A System for the Multiple Level Observation of Groups*. New York: Free Press.
- Bion, Wilfred 1959 *Experiences in Groups*. New York: Basic.
- Birdwhistell, Ray 1970 *Kinesics and Context*. Philadelphia: University of Pennsylvania Press.
- Borgatta, Edgar 1963 "A New Systematic Interaction Observation System." *Journal of Psychological Studies* 14:24-44.
- , and Betty Crowther 1965 *A Workbook for the Study of Social Interaction Processes*. Chicago: Rand McNally.
- Chapple, Elliot D. 1940 "Measuring Human Relations: An Introduction to the Study of the Interaction of Individuals." *Genetic Psychology Monographs* 27:3-147.
- Couch, Arthur S. 1960 Personality determinants of interpersonal behavior. Ph.D. diss, Harvard University.
- Futoran, Gail K., Janice R. Kelly, and Joseph E. McGrath 1989 "TEMPO: A Time-Based System for Analysis of Group Process." *Basic and Applied Social Psychology* 10(3):211-232.
- Hall, Edward T. 1963 "A System for the Notation of Proxemic Behavior." *American Anthropologist* 65:1003-1026.

- Hare, A. Paul 1976 *Handbook of Small Group Research*, 2nd ed. New York: Free Press.
- Johansen, Robert 1989 *Groupware*. New York: Free Press.
- Leary, Timothy 1957 *Interpersonal Diagnosis of Personality*. New York: Ronald.
- McGrath, Joseph E. 1991 "Time, Interaction, and Performance (TIP): A Theory of Groups." *Small Group Research* 22:147-174.
- (ed.) 1993 "Time, Task, and Technology in Work Groups: The JEMCO Workshop Study." *Small Group Research* (Special issue). 24:285-420.
- McLemore, Clinton, and Lorna Benjamin 1979 "Whatever Happened to Interpersonal Diagnosis?" *American Psychologist* 34:17-34.
- Mann, Richard D. 1967 *Interpersonal Styles and Group Development*. New York: John Wiley.
- Mehrabian, Albert 1970 "A Semantic Space for Nonverbal Behavior." *Journal of Consulting and Clinical Psychology* 35:248-257.
- Moreno, Jacob 1953 *Who Shall Survive?* Beacon, N.Y.: Beacon House.
- Parsons, Talcott, and Robert F. Bales 1955 *Family, Socialization, and Interaction Process*. New York: Free Press.
- Polley, Richard B. 1979 "Investigating Individual Perceptual Biases of Group Members in Rating and of Observers in SYMLOG Interaction Scoring." In R. F. Bales, S. P. Cohen, and S. A. Williamson, eds., *SYMLOG: A System for the Multiple Level Observation of Groups*. New York: Free Press.
- 1987 "The Dimensions of Interpersonal Behavior: A Method for Improving Rating Scales." *Social Psychology Quarterly* 50:72-82.
- , A. Paul Hare, and Philip J. Stone (eds.) 1988 *The SYMLOG Practitioner: Applications of Small Group Research*. New York: Praeger.
- Slater, Philip J. 1955 "Role Differentiation in Small Groups." *American Sociological Review* 20:300-310.
- Stone, Philip J. 1988 "SYMLOG for Skeptics." In R. B. Polley, A. P. Hare, and P. J. Stone, eds., *The SYMLOG Practitioner: Applications of Small Group Research*. New York: Praeger.
- , Dexter C. Dunphy, Marshall S. Smith, and Daniel M. Ogilvie 1966 *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Weick, Karl E. 1985 "Systematic Observation Methods." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Wheelan, Susan A. (ed.) 1999 "Group Development." *Small Group Research* (Special Issue). 30:3-129.
- , A. F. Verdi, and R. McKeage 1993 *The Group Observation Development System: Origins and Applications*. Philadelphia: PEP.
- Wiens, A. N., J. D. Matarazzo, and G. Saslow 1965 "The Interaction Recorder: An Electronic Punched Paper Tape Unity for Recording Speech Behavior during Interviews." *Journal of Clinical Psychology* 21:142-145.
- Wish, Myron, Morton Deutsch, and S. Kaplan 1976 "Perceived Dimensions of Interpersonal Relations." *Journal of Personality and Social Psychology* 33:409-420.

RICHARD B. KETTNER-POLLEY

## OCCUPATIONAL AND CAREER MOBILITY

Occupational and career mobility in adulthood is often referred to as intragenerational social mobility. It involves change in an individual's position in the labor market over the adult life course. Change is studied with respect to both type of work and the rewards derived from work. The term *career* refers to an individual's job history. Empirical *regularity* in the careers of individuals in the labor force defines what we call a "career line" or "job trajectory," since a work history common to a portion of the labor force reflects the existence of structurally determined linkages among jobs in the economy. Jobs are located in particular firms, whereas occupations and industries encompass jobs in many firms. An individual may remain in the same occupation or industry but change firms (and jobs within the same firm) any number of times. Since the process of job change does not necessarily involve a change of occupation or industry, but a change of occupation or industry always involves a job change, the process of job change provides a more detailed account of career movement. Changes in the rewards derived from work usually accompany job changes but can also occur during the course of tenure in a job.

Research on intragenerational mobility has focused on the labor force as a whole and on employees in particular occupations and firms. Research on the labor force as a whole has usually considered change in occupation and industry, as measured for detailed categories or more aggregated groupings that define broad occupational

and industrial groups. It has also considered change in the rewards derived from work, focusing primarily on occupational prestige and earnings. Research on particular occupations and firms has usually focused on job status and authority changes within organizational hierarchies and on changes in work rewards.

Early work on intragenerational mobility involved the mathematical modeling of transition probabilities, usually among a few, highly aggregated categories (see Mayer 1972 for a review). This work used Markov models and semi-Markov models to analyze transition probabilities in a sequence under the assumption that the job category an individual will occupy in the future depends only on the job category occupied in the present and not on job categories occupied previously. Although there was empirical support for this assumption in some studies, it was not found to be broadly applicable. In semi-Markov models, transition probabilities are permitted to vary with time and for subgroups of the population. These models capture declines in mobility with age or duration of stay and allow for the fact that some individuals are more likely to move than others. There is evidence that mobility is an exponentially declining function of time (Mayer 1972) and that some individuals become "movers" while others become "stayers" (Blumen et al. 1955).

Later analyses have considered more refined models of job change, focusing on job shifts as elementary acts in the mobility process. Job shifts may be either voluntary or involuntary, and may occur at a decreasing rate with time in the labor market. They are also more likely to occur within a firm than between firms as labor market experience increases (Rosenfeld 1992; DiPrete and Nonnemaker 1997). As time spent in a firm increases, both rates of promotion and rates of leaving the firm also decline (Petersen and Spilerman 1990; DiPrete and Nonnemaker 1997). When the effect of labor market experience is considered simultaneously, firm tenure may have a positive effect on within-firm job mobility, but this effect indicates that the rate of within-firm job mobility declines with labor market experience more slowly for each year of tenure with an employer than for each year of pre-employer labor market experience. Job shifts (both voluntary and involuntary) are strongly and negatively related to job tenure. As the duration of a job increases, the probability

of leaving it declines. Most of this decline occurs in the first year of a job, after which there is a leveling off (Topel and Ward 1992). Within firms, however, job tenure has a positive effect on the probability of promotion, in some instances increasing up to a point and then decreasing (Felmlee 1982; Althausser and Kalleberg 1990; Petersen and Spilerman 1990). The rate of mobility varies with characteristics of the individual, the job, the occupation, the employing organization, and the economic environment.

Change in the rewards derived from work usually accompanies a job shift (Sørensen 1974; Rosenbaum 1984; Topel and Ward 1992). Voluntary job shifts are associated with increases in prestige and wages; involuntary job shifts, with losses. Job shifts *within* firms and occupations are more often associated with gains in prestige and wages than those *between* firms and occupations (DiPrete and Kreckler 1991; DiPrete and McManus 1996; Cheng and Kalleberg 1996). In general, the rewards of changing jobs decline with labor market experience and job tenure. Wages, as well as the opportunity for wage growth and promotion, bear a negative relationship to the probability of a job shift (Petersen and Spilerman 1990; Topel and Ward 1992). Change in wages also occurs during the course of a job (Topel 1991), declining with labor market experience and, at a decreasing rate, with job tenure (Topel and Ward 1992). For the labor force as a whole, occupational prestige and earnings increase over the adult working life. The shape of these trajectories tends to be concave downward—with a rise early in the career, a plateau during the middle years, and a slight decline as the end of the work career approaches. Status and earnings trajectories have the same form, but the former is flatter (Mincer 1974; Rosenfeld 1980). There is also variation in the shape of these trajectories for those in different career lines (Spilerman 1986).

Given the heterogeneity of career lines and the important role that age, or duration since career entry, plays in shaping career lines, there have been attempts to study the path of careers in recent years. Not only does mobility decline sharply with age, and therefore with the proximity of a job to the end of a career line, but job changes that occur later in an individual's work life tend to involve jobs requiring skills that are more similar than those that occur earlier in the work life

(Spenner et al. 1982; Althauser and Kalleberg 1990). There is also evidence that early career experiences have an important effect on later career outcomes. This evidence indicates not only that those entering *different* career lines, who receive different rewards at career entry, can expect different career outcomes (Marini 1980; Spenner et al. 1982), but also that early experiences *within* a career line can condition subsequent progression and the level of reward attained relative to others who enter the same career line (Stewman and Konda 1983; Rosenbaum 1984). The career lines most often studied have been trajectories within institutional structures, but there have also been attempts to describe career lines that cross institutional boundaries (Spilerman 1977; Spenner et al. 1982; Althauser and Kalleberg 1990).

### LABOR MARKET STRUCTURE

The concept of "career line" or "job trajectory" derives from the view that the labor market is structured in a way that makes some types of job changes more likely than others. Early work on career mobility ignored this structural differentiation, estimating the overall (linear) relationship between the status and earnings of an individual's first job and the status and earnings of a job held later in the career (Blau and Duncan 1967; Coleman et al. 1972; Marini 1980). Jobs resembling each other in status, pay, and working conditions, however, are sometimes part of a career line and sometimes not, even if part of a career line can be attached to different career lines. Jobs providing similar current rewards may therefore not offer the same prospects for future mobility.

Since career lines are rooted in labor-market structure, their existence demonstrates that intragenerational mobility is influenced by the structure of the labor market and changes in that structure, as well as by the demography of the labor force and individual characteristics that affect movement within segments of the labor market. If jobs have entry requirements and confer rewards, the structure of jobs plays a critical role in establishing the link between the attributes of individuals and work rewards. Recognition that the labor market is structured in a way that produces segmentation among career lines has led to attempts not only to describe career lines but to identify the forces shaping them.

Pioneering work by White (1970) directed attention to the importance of social structure in shaping careers by modeling the way in which vacancies in a structured labor market trigger career movement. More recent models of vacancy-based movement have further delineated the ways in which vacancy chains, job distributions, and managerial staffing and hiring practice structure the relative career chances of individuals (Sørensen 1977; Skvoretz 1984; Stewman and Konda 1983; Stewman 1986). In these models the availability of job openings determines career advancement possibilities, and the shape of organizational hierarchies affects the probability of advancement. Since most organizational hierarchies are pyramidal, with many more low-level than high-level positions, the average employee's advancement must slow down over time. Aggregate age-promotion curves appear to be described empirically by an exponential-decline function where the highest promotion chances occur at the outset, declines are a fixed proportion of an individual's current chances, and promotion chances become increasingly rare but not impossible. In short, promotion favors youth, declines gradually, but does not disappear for older workers. There may also be chances for an increase in promotion during the initial career years when most on-the-job training occurs.

Although career advancement often occurs as a result of job change to fill a vacant position, not all career movement depends on or is affected by a job opening. New jobs in organizations are sometimes created for particular individuals, and jobs that are vacated may be eliminated rather than filled. New jobs can also be created by individual moves to become self-employed, and the termination of self-employment does not necessarily create a vacancy. Within organizations, career advancement occurs through upgrading as well as vacancy filling. An employee receives an upgrade promotion when a certain level of seniority or a certain performance marker is reached. Such promotions involve reclassification to a higher rank and do not depend on a vacancy's being present (Stewman and Yeh 1991; Barnett and Miner 1992). Both upgrade promotions and vacancy promotions may occur during the course of an individual's career. These alternative advancement mechanisms, however, are not necessarily found equally throughout an organizational hierarchy. In several occupational categories studied

in one organization, most lower-level promotions were vacancy based, and a higher proportion of upper-level promotions occurred through upgrading (Barnett and Miner 1992). The mechanisms available for promotion affect an individual's promotion chances. It has been found that if promotion occurs through upgrading based on performance rather than vacancy filling, salary grade level is not necessarily associated with a lower probability of promotion (Petersen and Spilerman 1990). Similarly, the effect of hiring temporary workers on the promotion chances of permanent workers depends on the extent to which promotion occurs through vacancy filling rather than upgrading (Barnett and Miner 1992).

Because the structure of the labor market affects career movement, there have been attempts to understand the structural bases of segmentation among career lines. During the 1940s and 1950s, institutional economists called for an understanding of well-defined systems of jobs and firms, drawing a distinction between internal and external labor markets. For example, Dunlop (1957) argued that within a firm there are groups of jobs, or "job clusters," each of which is linked together by technology, the administrative organization of the production process, and the social customs of the work community. A job cluster usually contains one or more key jobs and a group of associated jobs, and the wage rates for the key jobs mediate the effects of labor market influences, including union and government wage policies, and forces in the market for products on the wage structure of the firm. The distinction between internal and external labor markets was later reintroduced by Doeringer and Piore (1971), who discussed what they called a "mobility cluster." The central idea was that administrative rules and procedures tend to set up separate markets for those already hired (an internal labor market) and those seeking employment (an external labor market). A firm hires workers from the outside labor market into "entry jobs," and other jobs are filled internally as workers progress on well-defined career ladders by acquiring job-related skills, many of which are firm-specific. Thus, firms make investments in individuals, and these investments segment the workforce with respect to advancement opportunity.

The concept of the internal labor market was developed further in what was first called *dual* and

then *segmented* labor market theory (Gordon 1972; Kalleberg and Sørensen 1979). This theory draws a distinction between primary and secondary jobs, arguing that the internal labor market is only one kind of work setting. *Primary* jobs emphasize long-term attachment between workers and firms and offer built-in career ladders and promotion opportunities, whereas *secondary* jobs do not offer these advantages. The distinction between primary and secondary jobs may occur within the same firm or between firms, since primary jobs are considered more likely to be found in oligopolistic, unionized industries, and secondary jobs in competitive industries.

Dual and other segmented labor-market theories have come under attack as being too crude to meaningfully characterize the multiple dimensions on which labor markets vary. Attempts to measure labor-market segmentation by crude topologies, including industrial and occupational categories, have also been criticized. Because there is extensive heterogeneity within industrial and occupational categories, and because there is no way to link them to the kinds of job clusters hypothesized to exist, it has been argued that more disaggregated analyses of jobs and firms are needed. Many subsequent analyses have focused on particular bureaucracies or firms, and those examining larger segments of the labor market have sought to measure the characteristics on which employing organizations and occupations vary.

Attempting to develop a systematic conceptual scheme for studying labor market segmentation, Althausser and Kalleberg argued that "the concept of an internal labor market should include any cluster of jobs, regardless of occupational titles or employing organizations, that have three basic structural features: (a) a job ladder, with (b) entry only at the bottom, and (c) movement up this ladder, which is associated with a progressive development of knowledge and skill" (1981, p. 130). Based on four possible pairings of type of control and prospects for advancement, they differentiated four types of labor market structures: (1) firm internal labor markets, which are internal labor markets controlled by firms; (2) occupational internal labor markets, which are internal labor markets controlled by occupational incumbents; (3) firm labor markets, which provide



firm-specific security without advancement prospects; and (4) occupational labor markets, which provide occupational security without advancement prospects.

Within a firm or an occupation, there may be multiple job ladders that vary in length and shape, and there are structured relationships among job ladders. Job mobility occurs not only along formal organizational and occupational ladders but across them (Baron et al. 1986; Stewman 1986; DiPrete 1989). For example, job ladders often cross detailed and even aggregated occupational boundaries. DiPrete (1987) described the permeability of these boundaries as depending on several types of contingencies: (1) skill-based contingencies, (2) information-based contingencies, (3) contingencies due to the configuration of positions in an organization, and (4) contingencies that emerge from the institutionalization of formal structure. Internal labor markets limit competition from outsiders by offering a combination of closed and restricted mobility contests that involve movement on and across job ladders (DiPrete and Krecker 1991; Stewman and Yeh 1991).

Because employing organizations vary in the kinds of jobs they offer, the occupational groupings in which those jobs fall, and the wages they pay, there has been some attempt to examine the effect of organizational characteristics on job mobility. Large organizations offer better pay, have more job ladders offering opportunities for promotion and wage growth, and offer better employee benefits. Organization size is therefore positively related to mobility within an organization and negatively related to voluntary movement out of an organization (Hachen 1992; DiPrete 1993; Cheng and Kalleberg 1996). Organizations in concentrated industries and industries dominated by conglomerates experience both lower rates of within-firm mobility and lower rates of voluntary exit because they have higher average wage levels, and there is a negative relationship between the average wage level of an industry and the probability of both within-firm mobility and voluntary firm exit. High-wage industries have lower rates of within-firm mobility than low-wage industries because, net of organizational size, internal labor markets are less prevalent in firms in high-wage industries. Capital-intensive industries have higher rates of

both within-firm mobility and voluntary firm exit but lower rates of involuntary exit. These differences are linked to the greater prevalence of internal labor markets in firms in capital-intensive industries (Hachen 1992).

Although there is general agreement that the technical character of work influences organizational development, existing variation in institutional and personnel structures, especially cross-nationally, indicates that the technical character of work alone does not determine job ladders and career lines. These are affected by the historical circumstances surrounding an organization's founding, including the gender, racial, and ethnic composition of the workforce; by market and other social conditions that affect organizational change and employment growth; and by the negotiating strength of various bargaining units.

A number of explanations have been advanced for the emergence of internal labor markets. One focuses on the development of specialized knowledge and skill relevant to employment in particular occupations, jobs, and firms. Becker (1964) noted that when a worker receives training specific to the needs of an employer, the worker's value to the employer increases since a new employee would have to receive similar training before being able to perform at the same level. It is therefore in the interest of the employer to retain workers with such specific training by providing them with opportunities for promotion, salary growth, and employment security. When firm-specific skills are rewarded, workers also have an interest in remaining with the employer because the wages and other benefits they could obtain from another employer are lower. Based on this reasoning, Williamson (1975) argued that internal labor markets emerged because they were preferable to long-term contracts for maintaining the employment relationship. Thus, internal labor markets are seen as resulting from employers' needs for renewable supplies of otherwise scarce, highly skilled workers. Substantial empirical evidence is consistent with this view. Another explanation, offered by Bulow and Summers (1986), is that job hierarchies offering promotion prospects and wage increases are an important means of motivating workers when individual performance is not easily monitored.

The emergence and spread of internal labor markets is also seen as related to the development

of personnel departments, formalized rules, and the rise of bureaucratic control. Although there is some empirical support for this view, at least some internal labor markets preceded the emergence of bureaucratic control systems and rules (Althauser 1989; Stovel et al. 1996). Spilerman (1986) has noted that workers have an interest in barring lateral entry and in having high-level positions filled through promotion. Workers may also wish to limit employer discretion by having decisions about promotion and layoff tied to seniority. These worker interests cause labor unions to work to create more widespread job hierarchies and promotion. Although internal labor markets have been argued to be a result of unionization, evidence suggests that they are to some degree an alternative to unionization (Pfeffer and Cohen 1984). In manufacturing establishments, the technology of production helps to manage the workforce, binding workers to a specific organization with specific machines and production technology, and providing necessary skills and training. Unionization reduces turnover by providing workers with a voice, skewing compensation toward deferred benefits over current wages and bonuses, and paying a larger proportion of total compensation as fringe benefits valued particularly by senior workers. In the presence of a governance system under unionization, the internal labor market and other forms of bureaucratic control are redundant if not in competition with union mechanisms.

Because the prevalence of internal labor markets varies across organizations, there has been some attempt to explain why they are found more often in some organizations than others. It was initially argued that firms characterized by high profit levels, oligopolistic pricing, and large organization size were more likely to create internal labor markets because they could better afford to. However, there is evidence that internal labor markets are not merely a derivative feature of core-economy organization (Althauser 1989). Hachen (1992) has suggested that they are affected by three characteristics of employer personnel policy: the need to retain workers, the availability of alternative mechanisms for retention, and the ability to use specific mechanisms. In labor-intensive industries where labor costs are substantial and worker replacement is used as a strategy to maintain or even lower wage levels, retaining workers is less important. As a result, within-firm mobility rates

are low, and involuntary exit rates are high. When an organization desires to retain workers, offering higher wages and increasing internal opportunities are two means of retention, but organizations vary in their ability to use these mechanisms. A large organization will be better able to offer internal opportunities. An organization that can offer high wages may have less need to offer advancement opportunities.

Career lines not only involve substantial movement between job ladders within firms but often cross firm and industry boundaries rather than remaining within them. In the U.S. economy, most job change involves a change of employer (DiPrete and Krecker 1991). It has been estimated that only somewhat over a quarter of U.S. workers are continuously employed by the same employer for twenty years or more (Hall 1982). In some careers, such as the salaried professions, crafts, and "secondary" labor-market positions, firm and industry are not a locus of career line structure. Nevertheless, moves between firms and industries often occur between related positions, so that labor-market segmentation emerges even without institutional barriers.

Aspects of the structure of labor markets, other than internal labor markets, that have implications for occupational and career mobility have been documented in cross-national research. One type of cross-national variation is in the relationship between the educational system and the labor market, which varies according to the structure of the educational system. In countries with more stratified and standardized educational systems, such as Germany, Austria, France, and Norway, there is a tighter linkage between educational preparation and labor-market position than in the United States. In these countries, reliance on educational degrees and other credentials as screening devices is greater, and rates of occupational and job mobility are lower (Haller et al. 1985; Allmendinger 1989; DiPrete and McManus 1996). In the United States, where schooling is more open and comprehensive, the number of years of education attained matters more, and career mobility is higher because a larger proportion of employees are allocated to low-level entry positions in organizations. In Great Britain, which has a stratified educational system, similar to those of the other European countries, occupational mobility is higher than in the United States because

the linkage between education and training prior to labor-market entry and job placement is weaker. Workers are also more likely to attain part-time education after labor-market entry that leads to occupational change (Winfield et al. 1989).

Another aspect of labor market structure that varies cross-nationally is the degree of institutional separation between manual and white-collar work, which affects mobility across this boundary. Existing throughout Europe, but strongest in the German-speaking middle-European countries, is the institution of apprenticeship for manual work, which restricts access to jobs for skilled workers to those who complete an apprenticeship after elementary schooling. In countries where the institution of apprenticeship is less strong, as in France and the United States, access to a job as a skilled worker is more dependent on affiliation with a specific enterprise and is open to those moving up from unskilled jobs. The greater stratification of the educational system in countries where the institution of apprenticeship is strong also restricts upward mobility from manual to nonmanual jobs.

Another variation observed cross-nationally is in the degree of separation between hierarchically ranked white-collar positions. In Austria and France, there is greater separation among hierarchical layers of white-collar work than in the United States, where the higher levels of white-collar work are more open to entry from below (Haller et al. 1985). In the United States, both the highest and lowest categories in the occupational structure are less separated from the middle layers than they are in Austria and France. As a result, mobility over relatively long distances is more common in the United States. This difference may be due in part to the less stratified educational system in the United States.

Another type of variation in labor market structure found across countries is in the links of the labor market to other national institutions such as the state and centralized employer and labor associations. In some countries, such as China, the state remains the major employer (Zhou et al. 1997); in others, such as Austria and France, large enterprises and whole industrial sectors are nationalized (Haller et al. 1985). Labor markets in Europe have also been described as less flexible than those in the United States because "corporatist"

institutional arrangements of various types coordinate and regulate the labor markets. Countries described as corporatist have relatively centralized wage-setting institutions that result in less volatility in wages and smaller wage differences (DiPrete and McManus 1996; Gottschalk and Smeeding 1997). They may also have greater employment security that reduces involuntary job mobility, but restrictions on wages may result in relatively high unemployment as well.

Because the availability of vacancies in organizational hierarchies affects career advancement, and the supply of labor affects both the probability of advancement and wages, career progress in all countries is affected by demographic factors such as the size and distribution of education and skills in various age cohorts and the rate of exit from positions (Stewman and Konda 1983). It is also affected by organizational growth and contraction that result in the creation and termination of jobs. Change in the actual structure of jobs occurs in response to social and economic influences, including technological development. At the level of the firm, expansion increases the rate of promotion, and contraction decreases it (Stewman and Konda 1983; Barnett and Miner 1992). Expansion of an industry, as measured by growth in the average size of firms or the establishment of new firms, increases the rate of job shifts within and between firms in the industry and decreases the rate of involuntary exit (Hachen 1992; DiPrete 1993; Haveman and Cohen 1994; DiPrete and Nonnemaker 1997). Contraction of an industry, as measured by decline in the average size of firms or the closing of firms, decreases the rate of job shifts between firms in the industry and increases the rate of movement out of the industry, including movement to unemployment. Contraction of an industry may either decrease or increase the rate of within-firm mobility, since contraction may lead to the internal redeployment of labor as a substitute for new hiring. Internal mobility that reflects reorganization occurs primarily among lower white-collar, service, and blue-collar workers. Movement out of the industry and to unemployment occurs primarily among service and blue-collar workers (DiPrete 1993). Change that involves merger and acquisition within an industry has fewer effects, but they tend to be similar to the effects of industry contraction (DiPrete 1993; Haveman and Cohen 1994). The effects of *occupational* expansion and

contraction differ somewhat from those of *industrial* expansion and contraction. Expansion of an occupation increases the rate of mobility between firms in an industry but not within firms. Contraction of an occupation increases the rate of mobility within firms (due to the redeployment of labor) and increases the rate of movement to unemployment but does not increase the rate of movement out of the industry (DiPrete and Nonnemaker 1997). The adverse effects of economic turbulence in either an industry or occupation are dampened by employer tenure.

Since the structure of the labor market affects both job mobility and wage growth, aspects of labor market structure account for the aggregate declines in job mobility and wage gain observed over the course of workers' careers. One aspect of labor-market structure that plays a role in explaining the declining change in career rewards is the pyramidal shape of organizational and occupational hierarchies, which contain many more lower-level than higher-level positions. Given this structure and the role of vacancy filling in promotion, the chances of promotion and wage growth linked to promotion decline with labor-market experience. To the extent that upward mobility is possible, workers entering a career line are in competition with others at the same level for advancement to the next highest level. Because jobs at the next level are filled from those occupying positions in the level below, entry to a job ladder and performance at each step on the ladder affect ultimate career attainments. Rosenbaum (1984) has described the process by which employees in a cohort are progressively differentiated throughout their careers in a series of implicit competitions as being like a tournament. Selections among the members of a cohort occur continually as careers unfold, and each selection affects the opportunity to advance further. As a result, individuals are distributed more and more to "winner" and "loser" paths, and interpersonal variance in wages rises with labor-market experience and firm tenure. Those who are not promoted are more likely to leave the firm, but their gain from doing so is less than the gain of those who are promoted internally.

According to human capital theory, the shape of wage trajectories over the career course is a function of the changing productivity of workers. Wage growth through labor-market experience

results from learning on the job that increases a worker's productivity and value to the employer. On-the-job training occurs in formal training programs provided by employers or as a result of informal instruction by supervisors and coworkers and simply by doing the job. Time away from the job, in contrast, can lead to skill depreciation. Because investments in training are costly, these are concentrated in the early part of careers, leading to greater wage growth at early rather than later career stages. Formal and informal training in a particular firm and on particular jobs can provide knowledge and skills that are general or specific. General knowledge and skills are applicable in other firms and other jobs, whereas specific knowledge and skills are applicable only in the firm or job where they are acquired. As workers develop firm- and job-specific human capital, they are less likely to leave that firm or job because their productivity and resulting wages are higher than they would be in another firm or job. Although productivity may be greater in the firm or job where specific human capital is acquired, increments in productivity decline over time. As a result of this decline and biological aging, productivity may be lower at the end of a career than it was earlier.

One problem with human capital theory is that the relationship between labor market experience and worker productivity has not been adequately tested. Labor-market experience may affect wages because it is a proxy for seniority if wages rise with seniority and job tenure regardless of productivity. For women, labor-market experience may be an outcome rather than a determinant of career mobility if entry into career lines offering little opportunity for advancement affects labor-force participation (Marini 1980; Marini and Fan 1995). Evidence on the relationship between labor market experience and productivity within occupations suggests that, although labor market experience has some bearing on productivity, its effect on career advancement is largely independent of productivity (Horowitz and Sherman 1980; Medoff and Abraham 1980, 1981; Maranto and Rodgers 1984). Much variability is also seen across occupations and work contexts in the extent to which work experience affects either productivity or earnings (Horowitz and Sherman 1980; Spilerman 1986). Evidence also contradicts the "tradeoff hypothesis"—that individuals sacrifice earnings at

the beginning of their careers for better long-term career prospects. Individuals with lower earnings early in their careers actually have lower rather than higher job status and earnings later.

More recent economic explanations of the relationship between labor-market experience and wage growth have focused on new arguments about the role of productivity in shaping the desire of employers, especially those who make large investments in screening and training employees, to retain workers and motivate high performance over time (Lazear 1981; Lazear and Rosen 1981; Malcomson 1984; Bulow and Summers 1986). Promotions and wage increases are seen as a means of eliciting effort from workers when the monitoring of their efforts and outputs is prohibitively expensive. These explanations deviate from the view that labor is paid its marginal product in each short period, assuming that workers are paid their marginal product over the life cycle or in some cases in excess of their marginal product. So far, none of these explanations has an empirical basis, and, as single-factor explanations, they are unlikely to account for the diverse compensation schemes observed across occupations and work settings (Talbert and Bose 1977; Spilerman 1986). A further problem is that they ignore the role of nonmonetary incentives in retaining workers and motivating high performance.

Other economic explanations of career mobility have combined human capital theory and information theory to consider the role of imperfect information and the acquiring of additional information as influences on career decision making (e.g., Jovanovic 1979). As in human capital theory, voluntary job mobility is assumed to result from a comparison of the benefits derived from one's current job with those that could be obtained from another job. The decision to search for or accept another job is argued to be affected by new information about either the current job or a possible alternative. Because certain aspects of a job cannot be assessed prior to employment, new information is argued to be acquired through actually working on the job, and this information affects assessment of the quality of the person-job match. Jobs that survive are those evaluated as a good match, or at least as a better match than could be obtained elsewhere. With increases in labor-market experience, workers acquire more information about the labor market and are better

able to demonstrate their performance, which in turn provides more information to employers. As a result, matches improve over the course of a career, and, because it becomes increasingly less likely that a better match can be found, job mobility declines. Although the role of information is implicit, a related argument is linked to the vacancy-based model of career advancement (Sørensen 1977). According to this argument, as time in the labor market increases, people who enter the labor market in a position below their optimum level because a better vacancy is unavailable have more time to take advantage of vacancies and to close the gap between their current and potential rewards. What these models do not consider is the path dependency of careers, whereby there are long-term effects of early job placements on later career outcomes. Job placement early in a career not only shapes worker experience and training but is used as a basis for screening by employers that affects later career opportunities.

Information about the labor market and specific alternative jobs may also change with labor-market experience in ways that affect assessments of the likelihood that a better alternative is available. At labor-market entry and during the early career, both occupational aspirations and the importance attached to a variety of job attributes decline. These changes suggest a diminished view of the labor market, which makes it less likely that a better alternative to one's current job will be perceived to be available. As firm and job tenure increase, it is also likely that knowledge of specific alternative job opportunities declines, since the time since the worker's last involvement in job search is longer and contacts with others who would be sources of information are fewer.

If opportunities for career advancement and wage growth are used by employers to retain workers who have firm- and job-specific knowledge and skill, career advancement and wage growth may also decline over the course of individual careers because they become less necessary means of retaining workers. They become less necessary because increases in firm-and job-specific knowledge and skills increase the extent to which a worker will be more valuable to the current employer than to an alternative employer and because the value of accrued employee benefits, including retirement benefits, makes it increasingly costly for a worker to change employers. As firm

and job tenure increase, personal relationships and familiarity with the work environment also increase, reducing the likelihood that a worker will change employers or jobs. There is also less time left as a career advances to recover the costs of a change of employer or job and achieve the same level of functioning as was attained with the current employer or a higher level (Groot and Verberne 1997).

### WORKER CHARACTERISTICS

Given the structure of the labor market, particularly the local labor market where most job search occurs, entry into a career line is affected by worker characteristics, such as job-related credentials (e.g., education, intelligence, physical attributes), job preferences, and access to resources (e.g., information, material support, sponsorship by influential others). In some career lines, worker characteristics at entry may also influence career progression, whereas in others such characteristics may have little effect after access to a career line is obtained. If worker characteristics at entry continue to have an effect, they may do so in part because they influence subsequent on-the-job training and performance.

Most theoretical attempts to explain individual differences in career mobility have focused on individual differences in worker performance. In human capital theory, workers are seen as rational actors who make investments in their productive capacities to maximize lifetime income (Becker 1964; Mincer 1974). The investments usually studied are education and on-the-job training, although the theory applies to other investments, such as effort, job search, geographic mobility, and health. It is assumed that labor markets offer open opportunity and that earnings growth is a function of change in productivity due to how hard individuals work and to the ability, education, and training they possess. Individuals can increase their productivity not only through formal education but by learning on the job. Because investments in education and training are costly, workers concentrate their investments in the early part of their careers, sacrificing immediate earnings for better long-term career prospects. Although worker qualifications that may affect productivity have been found to be associated with advancement, these

associations are consistent with explanations other than that afforded by human capital theory. In the theory of vacancy-based competition, workers' performance-related resources are also assumed to affect career advancement, but the mechanism by which they affect advancement is left unspecified (e.g., Sørensen 1977).

Sociologists generally view the relationships between education and experience, on the one hand, and career advancement on the other, as influenced primarily by the organizational, and even the broader societal, context in which the administrative arrangements that govern promotions and salary advancement arise. Spilerman (1986) has suggested that education and experience bear weaker relationships to promotion and earnings when organizational rules rigidly prescribe the temporal paths of occupational and earnings advancement. These rigid schedules are usually found in workplaces where the majority of workers are engaged in a very few career lines, or where multiple career lines exist, but there is little opportunity for transferring among them. Such schedules often result from unionization, since labor unions seek to standardize work arrangements. Even the effects of education and experience that exist under this type of personnel system vary across cities in ways suggesting an influence of general societal beliefs that educated and experienced workers should be paid more. Because such beliefs do not specify how much more, compensation schedules vary widely. Education and experience appear to bear stronger relationships to earnings in large nonunionized organizations that encompass many occupational specialties. However, even in these organizations, societal notions of equity and custom may affect wage structures. In Japan, for example, both seniority and family size are major determinants of salary in large companies (Dore 1973).

Sociological accounts differ from economic explanations in recognizing that the rewards of work derive from job occupancy and that a relatively enduring structure of jobs determines the relationships of individual effort, ability, and performance to work rewards. In the sociological view, jobs are assigned wage rates via processes operating at the societal and organizational levels; and the mechanisms that match individuals to jobs produce associations between individual effort, ability, and training, on the one hand, and work

rewards, on the other. Jobs differ in the routes by which they are entered and in the extent to which performance can affect work rewards.

What influences advancement is not the attributes of individuals per se, but their attributes in relation to organizational positions. Organizations define the criteria by which ability is identified. To the extent that ability is not tied to those criteria, it will not be recognized and rewarded. If individuals of limited ability are able by chance or other more calculated means to meet the criteria by which ability is identified, they will be assumed to have ability by an inference process in which the direction of causality is reversed. An important consequence of this system is that factors affecting access to positions, of which ability is only one, and factors affecting performance in accordance with organizationally recognized criteria, including the willingness to conform to organizational goals and practices, have an important influence on long-term career outcomes.

Given the importance of the structure of jobs in mediating the relationship between individual attributes such as education and experience and job rewards, it is not surprising that unchanging attributes have little direct effect on rewards within job status categories. However, they have an important effect on careers via their influence on access to positions in the structure of jobs. Education is a major determinant of access to job ladders and career lines, and movement within these produces relationships between education and experience and work rewards. The organizational hierarchies to which college graduates, especially those from preferred colleges, have access increase the effects of college on career attainments over time. These hierarchies are moved through via experience. Within at least some career lines, education increases rates of promotion and wage gain (Rosenbaum 1984; Petersen and Spilerman 1990; Sicherman and Galor 1990). Education bears a positive relationship to within-firm and within-occupation mobility and a negative relationship to mobility between occupations, firms, and industries, as well as a negative relationship to moves to unemployment (Sicherman and Galor 1990; Cheng and Kalleberg 1996; DiPrete and Nonnemaker 1997). When mobility does occur between occupations and firms, education increases the likelihood that these moves are voluntary and in an upward direction. Education also helps to buffer workers

from economic turbulence. Young, educated workers are both helped less by organizational expansion and hurt less by organizational contraction (Rosenbaum 1984; DiPrete 1993; DiPrete and Nonnemaker 1997).

Over time, career histories become differentiated in the timing and occurrence of advancement. Early recognition and achievement relative to members of the same entry cohort have long-term effects on career mobility. Those receiving early promotions have more rapidly advancing subsequent careers and ultimately attain higher-level positions and higher earnings than those not receiving early recognition (Stewman and Konda 1983; Rosenbaum 1984). The age when a person enters a position bears a negative relationship at career advancement (Petersen and Spilerman 1990; Rosenfeld 1992). The early recognition of "stars" and its cumulative effect on subsequent career movement has been argued to occur as a result of labeling processes that identify such individuals as especially able—as high achievers who are likely to achieve more.

One problem in research on career mobility has been the limited availability of actual measures of worker ability and performance. Most research has studied the effects of education; formal on-the-job training; and informal training acquired through experience in the labor market, in a firm, or on a job. This set of influences is not only limited, but measures of these influences have tended to be crude and indirect. Informal training through experience has usually been inferred from the durations of time spent in the labor market, in a firm, or on a job. Since direct measures of ability and performance have rarely been obtained, knowledge of the effects of ability and performance on career mobility remains limited.

In both theory and research, there has been relatively little attention paid to the influence of worker characteristics other than ability and performance. There is evidence that individual differences in work and job values affect entry into career lines and that these change over the career course, in part as a result of experiences in the labor market (Lindsay and Knox 1984; Judge and Bretz 1992). However, the effect of workers' values, normative beliefs, preferences, and personality characteristics on career mobility has not been

studied, and individual variation in these characteristics has often been assumed to be either unimportant or nonexistent. Similarly, little attention has been paid to the influence of noncareer events and activities on career advancement. There has been some limited study of the effects of family roles, which indicates that marriage and children have different effects on the career advancement of women and men, but relatively little is known about the effects of family and other noncareer events and activities on career mobility.

Apart from ability and performance, the worker characteristics receiving the most attention in research on career mobility have been ascriptive characteristics such as gender, race, and ethnicity. The structure of the labor market mediates the relationship of these characteristics to job rewards, since these characteristics become a basis for the differentiation of job ladders and career lines, as well as a basis for access to the jobs within them (Spilerman 1977; DiPrete 1989; Marini and Fan 1995; Yamagata et al. 1997). Because other measures of worker characteristics and the process of employer evaluation and selection have received little research attention, however, understanding of the multiple factors that produce differences in career mobility among groups defined by ascriptive characteristics remains limited.

Another type of worker characteristic that has received some research attention is the social ties of workers, particularly as used in the process of job change. It has been argued that social networks provide a kind of capital—social capital—that is useful to workers in career advancement. Social ties of high status are seen as furthering career advancement more than those of low status, and it has also been argued that weak social ties are more helpful than strong social ties because weak ties provide a less redundant source of information and influence (Granovetter 1974; Lin 1990). There is evidence that workers do rely on others whom they know in the process of job change, but the extent to which social ties are a source of information and influence varies with the structure of the labor market and a worker's place in it. In China's largely state-controlled economy, a worker's social ties are not used as a source of information but are a source of influence on state authorities with decision-making power over jobs. In market economies, social ties are used as a source of both

information and influence, although the extent to which they are used varies across countries and for different types of career lines (Wegener 1991; Bian and Ang 1997). In at least some countries, workers in lower-status positions are more likely to use social ties when changing jobs than workers in higher-status positions. Evidence has been inconsistent on whether the use of social ties in obtaining jobs produces better outcomes, but when social ties are used and have an effect, obtaining help from those of high status rather than low status increases job rewards. Empirical evidence on the influence of the strength of social ties has been inconsistent, in part because the strength of social ties has been defined and operationalized in a variety of ways. In some countries strong social ties have been shown to be more helpful than weak ones. In other countries weak social ties are more helpful for workers in jobs of high status, whereas strong social ties are more helpful for workers in jobs of lower status (Wegener 1991; Bian and Ang 1997).

#### **SELECTION BY EMPLOYERS**

Movement within a career line is affected not only by the characteristics of workers but by the characteristics and actions of those empowered by employers to make hiring and promotion decisions. Evaluation of performance and ability is usually based on incomplete information. In many jobs, performance is difficult to assess. Ability is even harder to assess, because it is inferred from performance. In addition, human perceptive capabilities are limited and variable. As a result, employers tend to rely on readily available information, or "signals," such as the amount of education attained, where it was attained, the amount and types of prior job experience, observable personal attributes, and evidence of past performance such as the rate of career advancement and prior job status and earnings (Spence 1974). The use of such signals is particularly likely when job shifts are made between organizations and between job ladders within an organization rather than within an organization or a job ladder (Rosenfeld 1992). The criteria on which individuals are evaluated are therefore often superficial, and having the resources (i.e., money, knowledge, and skill) to identify the way the evaluation process works and acquire an appropriate set of signals plays an



important role in career advancement. Because of the difficulty of obtaining information, employers are susceptible to employees' attempts to supply and manipulate information about themselves as well as others.

Another influence on the evaluation of ability and performance is attitudes and beliefs previously acquired by those with decision-making authority. Information is unconsciously filtered and interpreted through that cognitive lens. The effect of prior attitudes and beliefs is evident in the prejudice and stereotyping triggered by ascriptive characteristics such as gender, race, and ethnicity (Hamilton 1981; Marini 1989), which have been a focus of theories of discrimination in the labor market (see, e.g., Blau 1984).

In addition to the difficulties that arise in assessing performance and ability, personal relationships and political coalitions influence mobility. As noted, there is growing evidence that personal contacts and relationships can constitute important sources of information and influence in gaining access to jobs. For workers seeking advancement, they can provide information about available positions and how to apply or present oneself favorably as a candidate. They can also result in preferential treatment based on personal liking and trust, shared interests, or expected gain on the part of those making hiring and promotion decisions. For employees responsible for hiring and promotion, personal contacts and relationships can provide information about candidates that reduces the risk of error in the selection process and increases the likelihood that individuals with similar interests and a sense of loyalty or indebtedness to them will be advanced. Personal contacts and relationships can also affect decisions if they lead to a desire to gain favor with others supportive of a candidate. Because those making hiring and promotion decisions can act to advance their own interests and the interests of others, a candidate's position in workplace political coalitions can affect advancement prospects.

## REFERENCES

- Allmendinger, Jutta 1989 *Career Mobility Dynamics: A Comparative Analysis of the United States, Norway, and West Germany*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Althausser, Robert P. 1989 "Internal Labor Markets." *Annual Review of Sociology* 15:143-161.
- , and Arne L. Kalleberg 1981 "Firms, Occupations and the Structure of Labor Markets: A Conceptual Analysis." In I. Berg, ed., *Sociological Perspectives on Labor Markets*. New York: Academic.
- 1990 "Identifying Career Lines and Internal Labor Markets within Firms: A Study in the Interrelationships of Theory and Methods." In R. L. Breiger, ed., *Social Mobility and Social Structure*. New York: Cambridge University Press.
- Barnett, William P., and Ann S. Miner 1992 "Standing on the Shoulders of Others: Career Interdependence in Job Mobility." *Administrative Science Quarterly* 37:262-281.
- Baron, James N., Alison Davis-Blake, and William T. Bieble 1986 "The Structure of Opportunity: How Promotion Ladders Vary within and among Organizations." *Administrative Science Quarterly* 31:248-273.
- Becker, Gary S. 1964 *Human Capital*. New York: National Bureau of Economic Research.
- Bian, Yanjie, and Soon Ang 1997 "Guanxi Networks and Job Mobility in China and Singapore." *Social Forces* 75:981-1005.
- Blau, Francine D. 1984 "Discrimination against Women: Theory and Evidence." In W. Darity, Jr., ed., *Labor Economics: Modern Views*. Boston: Kluwer-Nijhoff.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: John Wiley.
- Blumen, Isadore, M. Kogan, and P. J. McCarthy 1955 *The Industrial Mobility of Labor as a Probability Process*. *Cornell Studies in Industrial and Labor Relations*, vol. 6. Ithaca: New York State School of Industrial and Labor Relations, Cornell University.
- Bulow, J. I., and Larry H. Summers 1986 "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment." *Journal of Labor Economics* 4:376-414.
- Cheng, Mariah Mantsun, and Arne L. Kalleberg 1996 "Labor Market Structures in Japan: An Analysis of Organizational and Occupational Mobility." *Social Forces* 74:1235-1260.
- Coleman, James S., Zahava D. Blum, Aage B. Sørensen, and Peter H. Rossi 1972 "White and Black Careers during the First Decade of Labor Force Experience. I. Occupational Status." *Social Science Research* 1:243-270.
- DiPrete, Thomas A. 1987 "Horizontal and Vertical Mobility in Organizations." *Administrative Science Quarterly* 32:422-444.
- 1989 *The Bureaucratic Labor Market*. New York: Plenum.

- 1993 "Industrial Restructuring and the Mobility Response of American Workers in the 1980s." *American Sociological Review* 58:74–96.
- , and Margaret L. Krecker 1991 "Occupational Linkages and Job Mobility within and across Organizations." *Research in Social Stratification and Mobility* 10:91–131.
- DiPrete, Thomas A., and Patricia A. McManus 1996 "Institutions, Technical Change, and Diverging Life Chances: Earnings Mobility in the United States and Germany." *American Journal of Sociology* 102:34–79.
- DiPrete, Thomas A., and K. Lynn Nonnemaker 1997 "Structural Change, Labor Market Turbulence, and Labor Market Outcomes." *American Sociological Review* 62:386–404.
- Doeringer, Peter B., and M. J. Piore 1971 *Internal Labor Markets and Manpower Analysis*. Lexington, Mass.: Heath.
- Dore, Ronald P. 1973 *British Factory–Japanese Factory*. Berkeley: University of California Press.
- Dunlop, John 1957 "The Task of Contemporary Wage Theory." In G. W. Taylor and F. C. Pierson, eds., *New Concepts in Wage Determination*. New York: McGraw-Hill.
- Felmlee, Diane H. 1982 "Women's Job Mobility Processes within and between Employers." *American Journal of Sociology* 80:44–57.
- Gordon, David M. 1972 *Theories of Poverty and Underemployment*. Lexington, Mass.: Heath.
- Gottschalk, Peter, and Timothy M. Smeeding 1997 "Cross-National Comparisons of Earnings and Earnings Inequality." *Journal of Economic Literature* 35:633–687.
- Granovetter, Mark 1974 *Getting a Job: A Study of Contacts and Careers*. Cambridge, Mass.: Harvard University Press.
- Groot, Wim, and Maartje Verberne 1997 "Aging, Job Mobility, and Compensation." *Oxford Economic Papers* 49:380–403.
- Hachen, David S., Jr. 1992 "Industrial Characteristics and Job Mobility." *American Sociological Review* 57:39–55.
- Hall, Robert L. 1982 "The Importance of Lifetime Jobs in the U.S. Economy." *American Economic Review* 72:716–724.
- Haller, Max, Wolfgang König, Peter Krause, and Karin Kurz 1985 "Patterns of Career Mobility and Structural Positions in Advanced Capitalist Societies: A Comparison of Men in Austria, France, and the United States." *American Sociological Review* 50:579–603.
- Hamilton, David L. 1981 *Cognitive Processes in Stereotyping and Intergroup Behavior*. Hillsdale, N.J.: Erlbaum.
- Haveman, Heather A., and Lisa E. Cohen 1994 "The Ecological Dynamics of Careers: The Impact of Organizational Founding, Dissolution, and Merger on Job Mobility." *American Journal of Sociology* 100:104–152.
- Horowitz, Stanley A., and Allen Sherman 1980 "A Direct Measure of the Relationship between Human Capital and Productivity." *Journal of Human Resources* 15:67–76.
- Jovanovic, Boyan 1979 "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87:972–990.
- Judge, Timothy A., and Robert D. Bretz, Jr. 1992 "Effects of Work Values on Job Choice Decisions." *Journal of Applied Psychology* 77:261–271.
- Kalleberg, Arne, and Aage B. Sørensen 1979 "The Sociology of Labor Markets." *Annual Review of Sociology* 5:351–379.
- Lazear, Edward P. 1981 "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review* 71:606–620.
- , and Sherwin Rosen 1981 "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89:841–864.
- Lin, Nan 1990 "Social Resources and Social Mobility: A Structural Theory of Status Attainment." In R. L. Breiger, ed., *Social Mobility and Social Structure*. Cambridge: Cambridge University Press.
- Lindsay, Paul, and William E. Knox 1984 "Continuity and Change in Work Values Among Young Adults." *American Journal of Sociology* 89:918–931.
- Malcomson, James M. 1984 "Work Incentives, Hierarchy, and Internal Labor Markets." *Journal of Political Economy* 92:486–507.
- Maranto, Cheryl L., and Robert C. Rodgers 1984 "Does Work Experience Increase Productivity? A Test of the On-the-Job Training Hypothesis." *Journal of Human Resources* 19:341–357.
- Marini, Margaret Mooney 1980 "Sex Differences in the Process of Occupational Attainment: A Closer Look." *Social Science Research* 9:307–361.
- 1989 "Sex Differences in Earnings in the United States." *Annual Review of Sociology* 15:343–380.
- , and Pi-Ling Fan 1995 "The Gender Gap in Earnings at Career Entry." *American Sociological Review* 62:588–604.
- Mayer, Thomas 1972 "Models in Intragenerational Mobility." In J. Berger, M. Zelditch, Jr., and B. Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.
- Medoff, James L., and Katherine G. Abraham 1980 "Experience, Performance, and Earnings." *Quarterly Journal of Economics* 95:703–736.

- 1981 “Are Those Paid More Really More Productive? The Case of Experience.” *Journal of Human Resources* 16:186–216.
- Mincer, Jacob 1974 *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research.
- Petersen, Trond, and Seymour Spilerman 1990 “Job Quits from an Internal Labor Market.” In K. U. Mayer and N. B. Tuma, eds., *Event History Analysis in Life Course Research*. Madison: University of Wisconsin.
- Pfeffer, Jeffrey, and Yinon Cohen 1984 “Determinants of Internal Labor Markets in Organizations.” *Administrative Science Quarterly* 29:550–572.
- Rosenbaum, James L. 1984 *Career Mobility in a Corporate Hierarchy*. New York: Academic.
- Rosenfeld, Rachel A. 1980 “Race and Sex Differences in Career Dynamics.” *American Sociological Review* 45:583–609.
- 1992 “Job Mobility and Career Processes.” *Annual Review of Sociology* 18:39–61.
- Sicherman, Nachum, and Oded Galor 1990 “A Theory of Career Mobility.” *Journal of Political Economy* 98:169–192.
- Skvoretz, John 1984 “The Logic of Opportunity and Mobility.” *Social Forces* 63:72–97.
- Sørensen, Aage B. 1974 “A Model for Occupational Careers.” *American Journal of Sociology* 80:44–57.
- 1977 “The Structure of Inequality and the Process of Attainment.” *American Sociological Review* 42:965–978.
- Spence, Michael A. 1974 *Market Signaling*. Cambridge, Mass.: Harvard University Press.
- Spenner, Kenneth I., Luther B. Otto, and Vaughn R. A. Call 1982 *Career Lines and Careers*. Lexington, Mass.: Lexington Books.
- Spilerman, Seymour 1977 “Careers, Labor Market Structure, and Socioeconomic Achievement.” *American Journal of Sociology* 83:551–593.
- 1986 “Organizational Rules and the Features of Work Careers.” *Research in Social Stratification and Mobility* 5:41–102.
- Stewman, Shelby 1986 “Demographic Models of Internal Labor Markets.” *Administrative Science Quarterly* 31:212–247.
- , and S. L. Konda 1983 “Careers and Organizational Labor Markets: Demographic Models of Organizational Behavior.” *American Journal of Sociology* 88:637–685.
- Stewman, Shelby, and Kuang S. Yeh 1991 “Structural Pathways and Switching Mechanisms for Individual Careers.” *Research in Social Stratification and Mobility* 10:133–168.
- Stiglitz, Joseph E. 1973 “Approaches to the Economics of Discrimination.” *American Economic Review* 63:287–295.
- Stovel, Katherine, Michael Savage, and Peter Bearmen 1996 “Ascription and Achievement: Models of Career Systems at Lloyds Bank, 1890–1970.” *American Journal of Sociology* 102:358–399.
- Talbert, Joan, and Christine E. Bose 1977 “Wage Attainment Processes: The Retail Clerk Case.” *American Journal of Sociology* 83:403–424.
- Topel, Robert H. 1991 “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority.” *Journal of Political Economy* 99:145–176.
- , and Michael P. Ward 1992 “Job Mobility and the Careers of Young Men.” *Quarterly Journal of Economics* 107:439–479.
- Wegener, Bernd 1991 “Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment.” *American Sociological Review* 56:60–71.
- White, Harrison C. 1970 *Chains of Opportunity*. Cambridge, Mass.: Harvard University Press.
- Williamson, Oliver E. 1975 *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Winfield, Idee, Richard T. Campbell, Alan C. Kerckhoff, Diane D. Everett, and Jerry M. Trott 1989 “Career Processes in Great Britain and the United States.” *Social Forces* 68:284–308.
- Yamagata, Hisashi, Kuang S. Yeh, Shelby Stewman, and Hiroko Dodge 1997 “Sex Segregation and Glass Ceilings: A Comparative Statics Model of Women’s Career Opportunities in the Federal Government over a Quarter Century.” *American Journal of Sociology* 103:566–632.
- Zhou, Xueguang, Nancy Brandon Tuma, and Phyllis Moen 1997 “Institutional Change and Job Shift Patterns in Urban China, 1949 to 1994.” *American Sociological Review* 62:339–365.

MARGARET MOONEY MARINI

## OCCUPATIONAL PRESTIGE

Individuals have repeatedly demonstrated an ability to rank occupations according to their prestige, a relative social standing in a society. Occupational prestige is one of the most empirically studied aspects of stratification structure in modern societies. Social stratification theories, however, differ

in their views of the concept of prestige. Wegener points out that theories vary primarily in their suppositions of the foundation on which prestige is based, that is, achievement, esteem, honor, or charisma. Wegener also distinguishes two types of stratification theories, one that views prestige as a hierarchy of positions and the other as an attribute of socially closed groups.

*By and large, stratification theories that emphasize order in society (e.g., functionalist theories) conceive prestige as an attribute of individuals or of individual social positions that form a hierarchy. Stratification theories that emphasize conflict (e.g., Weber) think of prestige as designating social aggregates, or individuals within social aggregates, influenced by social closure processes. (Wegener 1992, p. 255)*

Despite such variation in theoretical views, most empirical studies share the notion that occupational positions are hierarchically ordered along a single dimension as judged by the individuals in the society. Efforts to measure such a concept involve a reputational approach in which respondents are asked to evaluate occupations.

The modern study of occupational prestige dates to a landmark survey fielded in 1947 by the National Opinion Research Center under the direction of Cecil C. North and Paul K. Hatt (Reiss 1961). Although others had conducted earlier investigations in the United States, NORC's national sample and broad coverage of the occupational hierarchy became the model for later inquiries. Perhaps the best known product of the North-Hatt study was Duncan's Socioeconomic Index (SEI) (1961), which assigned to each detailed occupational category a predicted prestige score based on the age-standardized education and income characteristics of occupational incumbents reported by the 1950 Census of Population. Although this index proved to have somewhat different properties than prestige, it exploited the limited occupational titles evaluated by the North-Hatt study to construct the first metric scale of socioeconomic status for all occupations.

In the 1960s, a second generation of studies was carried out by NORC (Hodge et al. 1964). Piecing together surveys from 1963, 1964, and 1965, Siegel (1971) generated the first prestige

scale for all Census occupations. This scale served for twenty years as the foundation of socioeconomic status scores, and it became the backbone of Treiman's International Prestige Scale (1977). In 1980, however, a major change in the occupational classification system employed by the U.S. Bureau of the Census called into question scores based on earlier classifications of occupational titles, since there was no sensible way to match old and new categories. In 1989, the NORC General Social Survey undertook another periodic sounding of Americans' evaluations of the general social standing of occupations (Nakao and Treas 1994). In the 1989 survey, a total of 740 occupational titles were selected to ensure reasonable coverage of the 503 detailed occupational categories of the 1980 census classification system. These titles were rated by a sample of 1,166 respondents, each of whom rated 110 occupations. The evaluation involved sorting cards, each bearing one occupational title, onto a sheet displaying a nine-rung ladder of social standing (from "1" for the lowest possible social standing to "9" for the highest possible). Based on their ratings, a score was calculated for each occupation and was converted to a scale so as to have a logical range from 0 (lowest) to 100 (highest) of prestige scores. For use in data analyses of most social surveys, in which occupations are coded according to occupational categories, scores assigned to categories, not titles, become necessary. Thus, using the scores for occupational titles in respective categories, a score was computed for each of the 503 detailed occupational categories of the 1980 census. The 1989 survey, which was the first to collect evaluations for all occupational categories at one time, yielded new prestige scores. These scores became the basis for updating the Socioeconomic Index (Hauser and Warren 1997).

Five generalizations may be drawn from the empirical research on occupational prestige.

First, very different methods for soliciting occupational evaluations yield very similar prestige hierarchies. Presenting respondents with an occupational title (e.g., electrician), the North-Hatt study asked them to "pick out the statement that best gives your personal opinion of the general standing that such a job has." Five response categories, ranging from "excellent standing" to "poor standing," were presented, along with a "don't know" option. One might readily fault the ambiguous

instructions calling for both a “personal” opinion and a reading of “general” standing. Nonetheless, the results proved virtually identical to those of the 1964 and 1989 surveys, which asked respondents to sort cards bearing occupational titles onto a “ladder of social standing,” the method that became the primary standard against which all other inquiries are evaluated (Nakao et al. 1990). The 1964 study went on to ask respondents to sort occupations onto a horizontal ruler according to another specific dimension (e.g., freedom and independence, perceived income, how interesting the work). However different the tasks, the correlation with social standing evaluations was over .90 for eight of nine dimensions.

Even when respondents are instructed to cluster occupations according to their similarity (rather than to rank them by social standing), multidimensional scaling methods reveal that one of the organizing principles behind judged similarity is a prestige hierarchy. Burton (1972) first demonstrated this with a nonrandom subsample of volunteers solicited from an advertisement in the Harvard student newspaper. Kraus et al. (1978) achieved similar results with a representative sample of 463 urban Israelis. To confirm that people view occupations in terms of an up/down classification scheme, Schwartz (1981) showed that ranking occupations according to “vertical” paired adjectives (e.g., top/bottom) yields results highly correlated with prestige scores, while rankings based on evaluative (e.g., kind/cruel), potency (e.g., big/little), and activity (e.g., slow/fast) dimensions fail to replicate prestige orderings. Studies that directly asked the respondents on what basis they rated occupations, or that asked them to rate occupations on selected dimensions such as “value to the society” or “power,” did not culminate in conclusive results. Individuals’ evaluations seem to be based on their judgments of the “overall desirability” of occupations (Goldthorpe and Hope 1974; Hauser and Featherman 1977). In short, the prestige hierarchy is so central to how we evaluate occupations that it emerges from virtually any reasonable effort to elicit it.

Second, overall prestige rankings are very stable over time. Hodge et al. (1964) reported a correlation of .99 between the 1947 North-Hatt study and their own in the mid-1960s. Nakao and Treas (1994) found a correlation of .96 between

the mid-1960s and 1989. This stability is not surprising. First, the relative income and education levels associated with various occupations are quite stable over time (Treiman and Terrell 1975). Second, to the extent that prestige is fixed by the division of labor and workplace authority, we do not expect the prestige of flight attendants to soar above that of pilots.

This is not to say that prestige never changes. Hodge et al. (1964) noted modest gains for blue-collar occupations, an upswing in scientific occupations and the “free” professions (e.g., “physician”), and a downturn in artistic, cultural, and communication occupations. Nakao and Treas (1994) compared the scores for 160 occupational titles evaluated in both 1964 and 1989 and found that the mean score moved up from 45.2 to 47.5 while the standard deviation declined from 17.3 to 15.8. They noted especially that the bottom of the American occupational prestige distribution shifted upward between 1964 and 1989. Low-status service and farming occupations especially came to be more favorably evaluated (see Table 1). This may be, in part, due to changes in the occupation itself. Farmers, for example, have undergone changes and have come to be seen as “agribusiness” owners. Thus, change in prestige may occur not only because of the succession of new generations who hold different views, but also because all age groups change their thinking about the relative standing of occupations. Changes in the general public’s familiarity with an occupation can also affect its rating, as demonstrated for “nuclear physicist” between 1947 and 1963 (Hodge et al. 1964). Thus, individual occupations change even though the overall ranking of occupations remains quite stable over time. The growing prestige of low-status occupational titles warrants further study, especially because it is inconsistent with socioeconomic trends observed in the workforce and in the workplace—the growing inequality in earnings, the decline of unionized blue-collar employment, the influx of traditionally devalued workers like women and immigrants, and the absence of systematic skill upgrading for blue-collar workers (Nakao and Treas 1994).

Third, prestige evaluations are surprisingly comparable from one society to another. Arguing that industrialization everywhere demands a similar organization and reward of work, Treiman

OCCUPATIONAL PRESTIGE

**Prestige Scores for Selected Occupations in the United States, 1964 and 1989  
(1980 Census Major Occupational Categories)**

Occupational Title	1964	1989
<b>Managerial and Professional Specialty Occupations</b>		
Department head in a state government	80	76
Banker	72	63
General manager of a manufacturing plant	64	62
Lunchroom operator	31	27
Accountant	57	65
Chemist	69	73
Public grade school teacher	60	64
Clergyman	69	67
Lawyer	76	75
Musician in a symphony orchestra	59	59
<b>Technical, Sales, and Administrative Support Occupations</b>		
Medical technician	61	68
Manager of a supermarket	47	48
Insurance agent	47	46
Travel agent	43	41
Cashier in a supermarket	31	33
Telephone solicitor	26	22
Secretary	46	46
Post office clerk	43	42
Shipping clerk	29	33
Bill collector	26	24
Bank teller	50	43
<b>Service Occupations</b>		
Housekeeper in a private home	25	34
Policeman	48	59
Bartender	20	25
Cook in a restaurant	26	34
Janitor	16	22
Barber	38	36
<b>Farming, Forestry, and Fishing Occupations</b>		
Farm owner and operator	44	53
Gardener	23	29
Logger	26	31
<b>Precision Production, Craft, and Repair Occupations</b>		
Airplane mechanic	48	53
Superintendent of a construction job	51	57
House painter	30	34
Baker	34	35
<b>Operators, Fabricators, and Laborers</b>		
Saw sharpener	19	23
Welder	40	42
Assembly line worker	27	35
Bus driver	32	32
Locomotive engineer	48	48
Filling station attendant	22	21

Table 1

SOURCE: Nakoa and Treas (1994); Siegel (1971).

(1977) assembled eighty-five prestige studies of sixty nations, tribal societies, and territories. Comparing the United States with fifty-nine other societies yielded an average intercorrelation of .837; in other words, about 70 percent of the variation in U.S. prestige evaluations is shared in common

with the “average” society available to Treiman. To be sure, the correlations ranged from .98 for Canada to .54 for Zaire. Prestige hierarchies are similar, but not identical. Notable differences relate to level of economic development (Treiman 1977) and the greater appreciation of manual

labor in socialist societies (Penn 1975; for a Chinese exception, see Lin and Xie 1988).

Fourth, subgroups within a society also tend to agree about the relative ranking of occupations. Efforts to discern differences between blacks and whites, between those employed in more versus less prestigious jobs, have typically found little effect of the respondent's social location on his or her view of the occupational hierarchy (e.g., Goldthorpe and Hope, 1972; Kraus et al. 1978). To be sure, higher-status groups assign somewhat higher absolute rankings to high-status jobs than do lower-status groups, who tend to boost lower-status jobs somewhat (Hodge and Rossi 1978). Apparently, this phenomenon does not arise because groups hold self-serving views of the social order. Instead, high-status individuals agree more among themselves and, thus, avoid random ranking errors that move both high- and low-status occupations toward the middle of the distribution. In comparison to the differences among individuals in their ratings (i.e., reported interrater correlations range from .42 to .745), the variation between groups is smaller than the variation within groups (Hodge et al. 1982). Since location in the social structure has been shown to influence so many other attitudes, it is surprising that groups agree so closely on the order of occupations. The mechanisms leading to this consensus are not well understood. By early adolescence, however, children can agree on how jobs rank (Gunn 1964).

Fifth, the main factors associated with an occupation's prestige are its education and income levels, the basic logic behind the original construction of the Socioeconomic Index (Duncan 1961). Socioeconomic scores based on these two factors account for about 80 percent of the variation in prestige attributed to different occupations (Hodge 1981). Based on the 1989 prestige scores and the occupational data from the 1990 U.S. Census, Hauser and Warren (1997) also reported that in several models they used, between 70 and 80 percent of the variance in occupational prestige was accounted for by occupational education and earnings.

In the studies of status attainment models, socioeconomic status scores were demonstrated to be superior to prestige scores in accounting for a son's occupational achievement using his father's status (Featherman et al. 1975; Featherman

and Hauser 1976; Treas and Tyree 1979). It would be a mistake, however, to dismiss occupational prestige. Subgroups of raters agree not only on the prestige of an occupation, but also on how prestige differs from the occupation's socioeconomic location (Hodge 1981). Furthermore, a more recent study by Kerckhoff et al. (1989) showed in a comparative study of status attainment in the United States and Great Britain, that differences between the two countries were illuminated in the models using prestige scores, but not in the models using the Socioeconomic Index. The British data yielded stronger intergenerational effect on the respondents' occupations than did the United States data when prestige scores were used. This implies that the prestige scale captures different aspects of the status attainment process than the Socioeconomic Index does.

However important the educational requirements and economic rewards of occupations, they are not alone in determining the prestige accorded occupations. The racial or age composition of jobs may also figure in their public evaluation. Even controlling for education and occupation, a higher proportion of nonwhites in an occupation is associated with a lower prestige rating (Siegel 1971). Occupations dominated by the very young or the very old are similarly disadvantaged (Siegel 1971). However, the proportion of women in an occupation was shown to have no or little effect on prestige (England 1979; Bose and Rossi 1983; Fox and Suschnigg 1989).

There is no conclusive evidence that American respondents consistently downgrade the status of female-gendered occupational titles (e.g., policewoman) as compared with male titles (e.g., policeman). However, male (but not female) respondents do downgrade the standing of occupations in which women find employment—a relation that holds after the income and education levels of workers are taken into account (Meyer 1978). Further studies are warranted investigating systematic associations between gender and occupational evaluation.

Empirical studies of occupational prestige are based on the notion that prestige represents a hierarchy of social standing of individuals or aggregates based on their occupations. This notion is widely used in analyses of social survey data as a

variable indicating an individual's social status. While it is a useful index, the assumption behind its application as a measure of social status is that individuals' social positions are manifested by the occupations they hold. This assumption is rooted in the fact that occupation is a means through which social and economic resources are distributed. Furthermore, occupation is salient in one's life, providing not only economic needs, but also social relations that establish one's role in a society. This assumption may hold true in most industrialized countries, yet it perhaps deserves a further consideration in future research. As to what measure is appropriate as an index of social status in what context (for example, socioeconomic scale versus prestige scale), further research is called for that would lead us to a more thorough understanding of the stratification system and the process of individuals' attainment of social positions within it.

## REFERENCES

- Bose, C. E., and P. H. Rossi 1983 "Gender and Jobs: Prestige Standings of Occupations as Affected by Gender." *American Sociological Review* 48:316-330.
- Burton, M. 1972 "Semantic Dimensions of Occupational Names." In A. Kimball Romney, Roger N. Shepard, and Sara Nerlove, eds., *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Vol. 2, Applications*. New York: Seminar.
- Duncan, O. D. 1961 "A Socioeconomic Index for All Occupations." In Albert J. Reiss, Jr., et al., eds., *Occupations and Social Status*. New York: Free Press.
- England, P. 1979 "Women and Occupational Prestige: A Vacuous Sex Inequality." *Signs* 5:252-265.
- Featherman, D. L., F. L. Jones, and R. M. Hauser 1975 "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4:329-360.
- Featherman, D. L., and R. M. Hauser 1976 "Prestige or Socioeconomic Scales in the Study of Occupational Achievement?" *Sociological Methods and Research* 4:402-422.
- Fox, J., and C. Suschinnig 1989 "A Note on Gender and the Prestige of Occupations." *Canadian Journal of Sociology* 14(3):353-360.
- Goldthorpe, J. H., and K. Hope 1972 "Occupational Grading and Occupational Prestige." In Keith Hope, ed., *The Analysis of Social Mobility: Methods and Approaches*. Oxford, U.K.: Clarendon Press.
- Goldthorpe, J. H., and K. Hope 1974 *The Social Grading of Occupations. A New Scale and Approach*. Oxford, U.K.: Clarendon Press.
- Gunn, B. 1964 "Children's Conceptions of Occupational Prestige." *Personnel and Guidance Journal* 558-563.
- Hauser, R. M., and J. R. Warren 1997 "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27:177-298.
- Hodge, R. W. 1981 "The Measurement of Occupational Prestige." *Social Science Research* 10:396-415.
- , and P. M. Rossi 1978 "Intergroup Consensus in Occupational Prestige Ratings: A Case of Serendipity Lost and Regained." *Sozialwissenschaftliche Annalen* 2:B59-B73.
- Hodge, R. W., P. M. Siegel, and P. Rossi 1964 "Occupational Prestige in the United States, 1925-1963." *American Journal of Sociology* 70:286-302.
- Kerckhoff, A. C., R. T. Campbell, J. M. Trott, and V. Kraus 1989 "The Transmission of Socioeconomic Status and Prestige in Great Britain and the United States." *Sociological Forum* 4:155-177.
- Kraus, V., E. O. Schild, and R. W. Hodge 1978 "Occupational Prestige in the Collective Conscience." *Social Forces* 56:900-918.
- Lin, N., and W. Xie 1988 "Occupational Prestige in Urban China." *American Journal of Sociology* 93:793-832.
- Meyer, G. S. 1978 "Sex and Marriage of Raters in the Evaluation of Occupations." *Social Science Research* 7:366-388.
- Nakao, K., R. W. Hodge, and J. Treas 1990 *On Revising Prestige Scores for All Occupations. General Social Survey Methodological Report No. 69*.
- Nakao, K., and J. Treas 1994 "Updating Occupational Prestige and Socioeconomic Scores: How the New Measures Measure Up." *Sociological Methodology* 24:1-72.
- Penn, R. 1975 "Occupational Prestige Hierarchies: A Great Empirical Invariant?" *Social Forces* 54:352-364.
- Reiss, A. J., Jr., O. D. Duncan, P. K. Hatt, and C. C. North 1961 *Occupations and Social Status*. New York: Free Press of Glencoe.
- Schwartz, B. 1981 *Vertical Classification: A Study in Structuralism and the Sociology of Knowledge*. Chicago: University of Chicago Press.
- Siegel, P. M. 1971 *Prestige in the American Occupational Structure*. Ph.D. diss, University of Chicago.
- Treas, J., and A. Tyree 1979 "Prestige Versus Socioeconomic Status in the Attainment Processes of American Men and Women." *Social Science Research* 8:201-221.
- Treiman, D. J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.



Treiman, D. J., and K. Terrell 1975 "Women, Work and Wages-Trends in the Female Occupation Structure." In Kenneth C. Land and Seymour Spilerman, eds., *Social Indicator Models*. New York: Russell Sage Foundation.

Wegener, B. 1992 "Concepts and Measurement of Prestige." *Annual Review of Sociology* 18:253-280.

KEIKO NAKAO

## OCCUPATIONS

See Professions; Work and Occupations.

## ORGANIZATIONAL EFFECTIVENESS

See Bureaucracy; Complex Organizations; Industrial Sociology; Organizational Structure; Social Organization.

## ORGANIZATIONAL STRUCTURE

Organizations are composed of a variety of elements. Perhaps the fundamental component is *organizational structure*, the set of interrelationships (social bonds) between positions. Even organizations of globe-encircling proportions, such as multinational corporations, demonstrate "the consciously coordinated activities of two or more people" (Barnard 1938, p. 73). Similarly, it may be argued that relationships between and among sets of such organizations form the social structure of whole societies.

Within an organizational structure, groups or sets of social relationships can be differentiated by task specialization, known as the *division of labor*. People are assigned to specific positions within an organizational structure to increase the specificity of tasks and the reliability with which they are performed. Organizational structure is both (1) an outcome resulting from interactive processes between elements within the organization, as well as between the environment and the organization, and (2) a determinant of those interactive processes. Organizational structure calls forth or inhibits particular behaviors by organizational participants.

Interaction among parties to a relationship results in shared understandings that become part of an organization's culture. Focusing as it does on relationships constituting organizational structure, the social systems perspective for organizational analysis has been criticized as having a static cast. By contrast, organizational theorists contend that studying social processes among constituent parties is dynamic because it examines how social change occurs as participants grant or withhold their consent for collective actions. When the volitional and cognitive exigencies among constituents change, their behavior toward each is altered. Negotiation among organizational participants leads to the creation of new relationships, fluidity in existing relationships, and the potential for breaking off longstanding relationships. Processes influencing expectation and negotiation are complicated by the structural advantages enjoyed by dominant constituent parties. Whether change in agreed-upon relations is viewed as desirable often depends on whether a given constituent party perceives change to be disproportionately beneficial to itself when compared with the benefits to other specific constituent groups or to the overall organization. Moreover, more powerful constituents can, to their own benefit, cloud less powerful constituents' perceptions of what is actually in the latter's self-interest.

Much traditional theory and research on organizations, almost always implicitly (sometimes explicitly), assumes that decisions by upper participants benefit the entire collectivity. Most analyses do not differentiate benefit to the collectivity from benefit to upper-, middle-, or lower-level participants. Most researchers simply proceed on the implicit assumption that owners and managers are prime, if not sole, legitimate participants in and *entitled* beneficiaries of organizational structure. It is assumed, moreover, that lower-level employees in particular, and lower level participants generally, are incidental to the social construction of the organization, and are thus passive contributors to and beneficiaries of organizational structure. By contrast, modern sociological theory (as expressed throughout the present discussion of organizational structure and its correlates) increasingly differentiates among upper-, middle-, and lower-level participants as well as, separately, the organization as a whole. Thus, in this view, a much broader array of constituents contributes directly

or indirectly to organizational survival and therefore merit an appropriate share of rewards. Researchers need explicitly to identify and differentiate all participants and constituents that contribute directly or indirectly to an organization. Allocation of resources to an organization's tasks, defined through the division of labor, critically affects the power balance between participants, thus raising important and politically disturbing questions. Organizational structure, composed of collectively endorsed and enacted resource allocation agreements, can be usefully understood from an "organizational justice" perspective.

Organizational justice research seeks to understand how resources are allocated in ways believed to be "fair" by various interactants. The perceived fairness of resource allocations by organizational participants is distributive justice; the perceived fairness of the process through which resource allocation decisions are made is procedural justice. Both distributive justice and procedural justice interact to influence organizational participants' perceptions of fairness. Distributive justice may follow underlying norms of allocation, including an assessment of both the quantity and the quality of participants' contribution to a given effort (thus, their earned equity in the organization), as well as the extent to which resources are evenly allocated among participants (thus, their right to equal treatment) and, moreover, how much participants *need* the resources they receive (thus, their right to basic means for survival) (Linkey and Alexander 1998). These norms underlying the allocation of resources may affect the satisfaction of organizational participants through the outcomes of allocation decisions. Procedural justice has two components: the structure and form of procedures through which resources are allocated, and the social and emotional relations between parties in an exchange relationship (interactive justice). As Brockner and Siegel (1995) point out, the difference between these two components may signal the difference between how much control participants have over the process of allocative decisions, and their intent toward others in making such decisions.

Researchers have proposed two different viewpoints to explain variations in organizational participants' satisfaction with procedural and distributive justice. The self-interest explanation is that individuals wish to maximize their resources and

are more satisfied when their rewards are greater. The group-values explanation is that individuals evaluate themselves in the context of group membership; the perception of an individual as fair is colored by perceptions of the fairness of his or her group, and how resources are allocated to an individual indicates his or her worth relative to other members of the group. Both the self-interest and the group-values explanations have helped to explain organizational participants' satisfaction with the allocation of resources. These two approaches are reflected in the focus of allocation rules based on the *attributes* and *contributions* of recipients (Cook and Yamagishi 1983). *Attributes* are personal characteristics such as gender, age, and ethnicity that are used to determine social status. *Contributions* refer to valued inputs made by participants in an exchange relationship; among these valued inputs are performance and effort. Classification is not always clear-cut; for example, *ability* might be classified as an attribute in some circumstances but as a contribution otherwise.

While most of the organizational justice literature focuses on the contributions of organizational participants, a growing number of studies focus on their attributes. Evidence suggests that the degree to which organizational participants endorse norms for resource allocation to individuals is associated with their own placement within an organizational structure. Moreover, endorsement of these norms may also vary according to demographic characteristics such as age, gender, and income, as well as cultural factors such as political affiliation and religious affiliation (Jason S. Lee and Stolte 1994; Linkey and Alexander 1998). An important question for further research is whether individuals with given attributes alter their endorsement of allocative norms as their own placement within the organization's structure varies over the course of their own tenure in the organization. Further, *satisfaction* with resource allocation outcomes is associated with structural, demographic, and cultural factors (Huo et al. 1996; Irwin 1996); although the evidence for a consistent association between gender and satisfaction is mixed (e.g., Sweeney and McFarlin 1997; Lee and Fahr 1999). Given the level of ethnic and gender segregation within the workforce, particular tasks appear to be associated with certain attributes as workers are distributed into organizational positions. Additionally, institutional factors, such as

state legislation, and cultural values may affect allocators' distribution norms. In some cases, distribution norms for external constituents related to a given organization through its impact on their culture, economy, or community may differ from distribution rules for those participating directly in an organization. These factors contribute to differential allocation of resources to participants within organizational structures, and their different perceptions of, and satisfaction with, allocation outcomes.

Blau and Scott's *cui bono* criterion (1962) explicitly raises the question of who benefits from particular policies and characteristics of organizational structure. Blau and Scott suggest a fourfold topology of organizations: (1) mutual benefit organizations, such as clubs, where presumably egalitarian *members* are prime beneficiaries; (2) business or industrial organizations, where *owners* are prime beneficiaries; (3) service organizations, such as hospitals, where *clients* are prime beneficiaries; (4) commonweal organizations, such as the State Department, in which the *public-at-large* is prime beneficiary. For each type of organization, researchers should systematically examine patterns of benefit by virtue of a constituency's location either (1) externally in an input-output exchange relationship to the organization or (2) internally as upper-, middle-, or lower-level participants in its organizational structure (Etzioni 1961). A fully developed distributive justice perspective would evaluate the benefit a constituency derives from the organization relative to its contribution to the organization's *sustained* existence (Alvarez 1979).

Not only does the allocation of resources impact organizational structure, it also affects assessments of organizational structure in terms of effectiveness and efficiency. *Organizational effectiveness* may be defined as the capacity of an organization to produce intended and unintended outcomes. An organization may unknowingly, unintendedly, or inadvertently serve the interest of a given constituency. Such "latent" functional consequences are seldom explored and documented in advance by social scientists. More usually, latent functions are discovered when they become manifest because organizational activities are dramatically altered or the organization ceases to exist, with the consequence that former latent beneficiaries are severely affected. Indeed, previously unrecognized support from such latent beneficiaries may have

been critical to organizational well-being (perhaps survival), and the disaffection of such beneficiaries might threaten future organizational survival. But the concept of "latent" functions has not yet produced a research literature.

Notwithstanding vexing conceptual problems raised by the issue of "latency," research on organizational effectiveness has traditionally overemphasized manifest purposive action, by reference to the organization's *intended* outcomes (or formal goals). It is obvious, but critical to note, that if the organization goes out of existence, it can no longer accomplish anything. Hence, perhaps the most important task (latent or manifest) for any organization is its continued survival. While some organizations are designed, and are prepared from the outset, to go out of existence upon completion of the task for which they were initially created, many, perhaps most, organizations are quick to acquire new purposes so as to maintain themselves in existence. Indeed, some organizations such as political parties and governments have been known to kill their own people in attempts to remain in existence. Thus, the "cost" per se of getting things done is, at least theoretically, of no consequence in assessing organizational effectiveness. Traditionally, assessing effectiveness requires the comparison of organizational activities to an optimal standard outlined in organizational goals. Since different organizational constituencies (or stakeholders) may have different and sometimes conflicting expectations, this traditional approach has serious limitations.

By contrast, the "cost" of getting things done is paramount in any assessment of *organizational efficiency*. "Costs" of accomplishing any given task can be of various kinds; and each type of cost may be assessed by various techniques, with varying degrees of efficacy. Unfortunately, social science has not reached any consensus on what types of costs and benefits are worth measuring. Thus, in its broadest sense, *efficiency* may be viewed as the assessment of an organization's inputs (either in kind or in the distribution of those inputs) relative to its output, in a more narrow view it is simply cost per unit of production. Traditionally, efficiency is reflected in economic analyses and cost-accounting techniques (e.g., Rossi and Freeman 1993). However, it may reflect other kinds of resources, such as opportunity costs; risk; and political, social, and social benefits. Like effectiveness, the

attributes and positions of organizational stakeholders influence how efficiency is perceived and assessed.

This review is categorized into four functional requirements for social system survival, as posited by Parsons (1960): (1) adaptation to the environment, (2) goal attainment, (3) integration of its members into a “whole,” and (4) creation of cultural understandings (often latent) among members by which the meaning of collective action can be judged. We do not suggest Parsons’s AGIL scheme is the only or the definitive way to classify organizational structures or attendant activities. Rather, AGIL focuses on the totality of a particular organization (as a social system) and raises questions of how its organizational structures come into being, change, and persist; within each category we question “who benefits” from consensual, consciously coordinated activities.

#### ADAPTATION TO THE ENVIRONMENT

“Environment” refers to a broad array of elements that are “outside” organizational boundaries but are relevant to organizational functioning. Organizational boundaries are the set of agreed-upon relationships that constitute organizational structure. Dill defines the task, or technical, environment as all features of the environment “potentially relevant to goal setting and goal attainment” (1958, p. 410). Established beliefs and practices embedded within the organization may systematically affect the shape and operation of the focal organization’s structure. Scott (1998) terms these influences the *institutional environment* of organizations; others refer to it as *organizational culture*. An organization may seek to control or reshape all or some elements in its environment as a means to lessen uncertainty about its capacity to endure. An organization might reshape its own organizational structures if it is unable to reshape the environment (e.g., Aldrich 1979). Dess and Beard (1984) suggest that Aldrich’s topology of environmental characteristics (1979) may be classified into three categories: (1) munificence, or the environmental availability of resources needed by organizations; (2) complexity, or the similarity or dissimilarity of environmental entities and their distribution across the environment; and (3) dynamism, or the degree of change in the environment.

Organizations are embedded within larger societies. The societal culture (relatively integrated sets of values and value-based orientations) of a given society directly affects the kinds of organizational structures that can be sustained by organizations. Nevertheless, organizational structures and their internal organizational cultures might be a stronger determinative force than the outer societal culture. Hence, some organizations are often viewed as determinative importers of social change into some societies (e.g., technology transfers by multinational corporations). Organizational culture and the external societal culture inevitably affect one another; exploring specific nuances at the interface between these two cultural arenas is still a matter for future research. Blau (1994) notes that differential mobility among social groups may alter organizational structures, either increasing or constraining opportunities. Gains in social equity made by groups characterized by gender, ethnicity, age, or socioeconomic status may result in the alteration of organizational structures. The ability of workers to form coalitions both within and outside organizations may alter organizational hiring, promotion, and retention practices. Such restructuring may increase the mobility of groups in gaining more equitable organizational positions, or it may retain inequalities through the eventual resegregation of occupations (Cohen et al. 1998).

Neither effectiveness nor efficiency is an unalloyed universal good. As given constituencies gain stature and power, they are able to reform, reconstitute, or create relationships accordingly; but what makes one constituency more effective and/or more efficient may impact negatively on either the effectiveness or the efficiency of another. Hence, while not its sole determinant, the *net balance of power* between constituencies is an important variable in the formation and maintenance of organizational structure. When environmental conditions remain relatively constant, the stratification of organizational structure may also remain relatively constant as it is being reified, recreated, and reenacted, given that those already well placed are advantaged within each new round of structured interactions. However, when environmental conditions are visited by strong or rapid changes for example, in political culture, technology, or rapid changes (for example, in political culture, technology, or demographic characteristics), these may precipitate renegotiation and

recalibration of social structures within the organization. Because modern organizations seldom resemble a zero-sum model, not all gains by one constituency constitute a loss for another. However, unless the organization continues to grow, prosper, and reward formerly well-rewarded constituencies, a sense of relative deprivation due to perceived loss of effectiveness and efficiency in obtaining organizational rewards may lead to intergroup conflict.

Thus, organizations are affected by the exchange of resources between organizations and various constituencies, including those providing resource inputs to an organization, the exchange of labor for wages, and those between organizations and consumers of outputs. Environmental pressures, including competition for resources, may influence organizations to change structure, including strategic alliances, mergers, joint ventures, downsizing, and divestitures. Organizations may choose to maintain or change relationships on the basis of competition, power to create stable market relations, and institutional attachments through interpersonal and interorganizational relationships (Baker et al. 1998). Such changes presumably allow organizations more flexibility in responding to turbulent and uncertain environments, and each changes the stance of an organization in relation to others that compete for related resources in a market. This is accomplished by increased organizational efficiency, in which decreased labor force size, or hierarchy, and increased organizational interdependence results in fewer costs in producing outputs. The impacts of such strategies have been consequential for workers' expectations of long-term connections to workplaces, and expectations about their careers. Workers increasingly expect duties to be assigned on the basis of experience and training rather than organizational structure (Powell 1996). Pugh and Hickson (1996) argue that as organizational environments become more turbulent and uncertain, organizations require a "redundancy of function" in which individuals are called on to have a wide variety of skills and fill multiple functions in a highly interdependent environment. Thus, workers develop flexible organizational "roles" rather than fulfilling a function based on a hierarchical position (Powell 1996). Older and younger workers may experience a sense of relative deprivation as a result of changes in organizational structure

and as their own expectations of future employment opportunities decrease (Lerner 1996). Lerner concludes that procedural justice in organizational restructuring has not led to distributive justice in terms of employment opportunities and expectations.

The *state* (the system of governance in a society) is another global element with pervasive repercussions for organizational functioning. The state may regulate the organizations directly by instituting programs within them or indirectly by state regulation of what an organization may produce or how it may transact with other organizations such as suppliers or consumers. Organizations may attempt to influence governmental actions so that public policy does not constrain them or so that it will actively benefit them. Organizations frequently influence legislation directly, as by activities of lobbyists on retainer. Organizations often mount campaigns either to achieve or to prevent the enactment of specific legislation by directly influencing general public opinion and particular voters. Even if legislation is passed over their opposition, organizations can achieve their purposes by subsequently influencing the allocation of resources for its enforcement. Fligstein and Mara-Drita (1996) note that market relations between buyers, sellers, and the state are characterized by a power struggle in which each participant mobilizes resources to enact its own interests in maintaining (or altering) current relationships.

Communities are attentive to organizations located in their midst since changes in organizational structure can have considerable repercussions for the community at large. As Scott (1998) points out, not all organizations are strongly tied to the communities in which they are located. Locally based firms have a greater vested interest in community prosperity than do geographically dispersed firms, and they may act to assure continued community prosperity. Organizations may strongly affect the allocation of public goods and services as well as specifications in local policies, such as zoning and tax laws. Organizations may affect communities in which they are located through either implicit or explicit threats of "exit" as well as by directly impacting regulatory and economic conditions. The number, size, and type of organizations located in a community also have widespread consequences for individual local residents. South and Xu (1990) compared industries

that dominate their local metropolitan economy with those that do not, finding that employees in dominant industries earn higher wages. Thus, organizations have important political, economic, and normative effects on individuals with or without organizational membership and on their community's organized power structures.

We now explore how organizations respond to and create their own environments, "Gatekeepers" are organizational participants at various levels who "selectively" permit information and people to traverse boundaries into and out of an organization. They "legitimate" particular environmental constituencies, with whom the organization then establishes institutional relations. Relations with constituencies not so "selected" become invisible, neutralized, or illegitimate. The breadth of the environmental domain that an organization claims in this manner has consequences for its stability. Narrow domains are associated with greater stability, while broad, inconsistently defined domains are associated with loss of function (Meyer 1975). At the same time, normative and regulatory forces outside organizations may exert pressures on gate-keepers to allow or disallow certain pieces of information or particular people.

"Loose coupling," or a seemingly weak relationship between parts of an organization (Pfeffer and Salancik 1978), is one of many ways in which organizations learn to deal with a broad environmental domain. This weak relationship allows change in one part of an organization to precipitate minimal or no change in other parts. Relationships between subunits or individuals in organizations, and relationships between the organization and other environmental entities, may be loosely coupled. Organizations may respond to potentially disruptive threats by very limited conformity in a specific sector, and yet this limited conformity projects an aura of complete organizational compliance. In reality, components of organizational structure affected by a given threat may be effectively uncoupled from many other components and processes, resulting merely in the appearance of compliance (DiMaggio and Powell 1983). Loose coupling can also lead to structural "inertia" in an organization's response to environmental changes, causing "lags" between environmental changes and adaptations to them on the part of various organizational structures. However, the degree to

which loose coupling is useful as a strategy depends on the particular situation and linkages involved. Loose coupling in terms of outsourcing labor, autonomous work teams, organizational networking, and increasingly separated divisional organizational forms have all been adopted in the last decade as a response to increased market competition and technical complexity. Such modifications of organizational structure are most often justified by upper-level participants as attempts to increase either or both the organization's efficiency and *responsiveness* (effectiveness) to environmental challenges perceived to threaten organizational survival. What is left unsaid, however, is that a narrow economic conception of what constitutes efficiency or effectiveness by upper-level participants, such as executives and stockholders, may have devastating consequences for middle- and lower-level participants as well as for the surrounding community. Indeed, the very fact that they are essentially disenfranchised from organizational decision making may engender a confrontational relationship with upper-level participants that may render all parties inefficient and ineffective for the process of inventing alternative problem-solving solutions that contribute to the organization's holistic and long-term well-being.

Assessing the effectiveness of an organization's capacity to adapt may be less focused on outcomes than on the processes of organizational change. Organizational theorists have suggested that organizational effectiveness is reflected in how well an organization acquires and processes information and with what flexibility and adaptability (Weick 1977). Galbraith (1977) contends that organizations reduce environmental complexity by changing communication structures within organizations. Other researchers suggest that organizations are effective when their subunits are congruent with the specific environment with which they interact, and when organizations overall are congruent with their environments (Lawrence and Lorsch 1967). Such perspectives of organizational effectiveness may not be linked to the achievement of goals but, from a managerial perspective, may rest in the survival, or profitability of organizations. As Lerner (and maybe others) demonstrate, such goals may not be satisfactory to middle- and lower-level participants whose employment opportunities and expectations may be adversely affected by a particular organizational adaptation.

Further, since organizational participation is affected by the demographic characteristics of participants, the specific demographic consequences (whether intended or unintended) of environmental adaptation must be examined.

Who benefits in the adaptation of organizations to the environment? Elites (fiduciaries, executives, and high-level managers) are only one kind of constituency vying for potential benefits derived from organizational structure. Middle- and lower-level organizational participants are often neglected in the research literature on organizational structure. Often, when research findings indicate that either an "organization" or a "community" benefits from a particular activity, what is really meant is that *upper*-level participants benefit. Certain populations (women and ethnic minorities, for example) participate differentially at upper, middle, and lower levels of organizational structure; thus, a focus on upper-level participants is insufficient to fully describe patterns of benefit.

### GOAL ATTAINMENT

"Goal" refers to a desirable future state of affairs. *Official* goals are the formal statements put forth by organizations to state their general purposes (Perrow 1961). *Operative* goals, on the other hand, refer to "what the organization is actually trying to do" (Perrow 1961, p. 855). The degree of congruency between official and operative goals is variable. It is important to distinguish organizational goals from the motives of individual organizational participants (Simon 1964). However, researchers need to clarify how particular goal activity differentially benefits specific internal or external constituencies. Goals limit and direct organizational decision making and suggest criteria by which organizational performance can be measured.

Over time, organizations establish multiple, often disparate, and sometimes conflicting goals. Kochan and colleagues (1976) argue that goal multiplicity and conflict are associated with both horizontal (number of tasks at the same level of structure) and vertical (number of levels between the "highest" and "lowest" units) differentiation of organizational structure. In the pursuit of multiple goals, coordination of effort is necessary, leading to vertical differentiation of organizational

structure. Organizations change their goals over time, for both external and internal reasons. Thompson and McEwen (1958) contend that goals vary because interaction with elements *external* to the organization can be of two kinds: competitive or cooperative. The only competitive option in their discussion we call *bounded competition*, referring to the fact that the interaction takes place within the bounds of the normative structure (institutional environment) of the larger social system. By contrast, we conceive of *raw* or *unbounded competition* as taking place outside any normative order common to the contending parties and is not accounted for by Thompson and McEwen's discussion. In the extreme absence of common normative understandings, hostility between contending parties can rise to a level wherein one party believes itself justified in attempts to annihilate another. Accordingly, Thompson and McEwen define competition as rivalry between two or more organizations mediated by a third party. Organizations compete for resources viewed as desirable for organizational functioning. Thompson and McEwen discuss three cooperative styles of interaction between organizations and external elements, each underscored by a decreasing level of hostility: co-optation, bargaining, and coalition. Co-optation is the absorption of an external element into the organization, neutralizing its potential hostility by incorporating it within the organization's structure. Bargaining is direct interaction with environmental entities in which some kind of exchange takes place so that the organization can get what it desires. Coalition is an agreement, usually of specific duration for specific collective purposes, combining the efforts of two or more organizations, and restricting the right of each to set goals unilaterally. Notice that coalition requires very low levels of hostility between an organization and its partners; indeed, a potential outcome may be loss of separate identity and structural unification. These strategies can increase or decrease the size and complexity of organizational structure and the allocation of resources within it.

Organizational goals also change for *internal* reasons. Constituencies within the organization frequently form *coalitions*, initially for self-protection against real or imagined threats to the pursuit of their own interests. Each unable to impose its will on others, but fearing imposition, makes alliances with other constituencies perceived to be

friendly. Some such alliances capture key positions of the organizational structure, thus giving greater access to the allocation of organizational resources. This dominance can be maintained over time by securing the cooperation of other elements and coalitions within the organization through the selective distribution of resources. How central a given goal is to an organization may depend on the composition of the dominant coalition and the relative balance of power within it.

This discussion has emphasized the complexity of goals and goal setting, given a variety of internal and external factors. The processes by which groups or members of organizations gain power, and the loose coupling between goals and motivations, have a large impact on goal setting. It is useful to discuss the processes of power in organizations in more depth. Many theorists have proposed definitions of power, but it was Emerson who proposed one of the most useful: *Power* resides in the dependency of one on the resources of another (1962). Resource control theorists believe that individuals or organizational subunits exercise power because they allocate resources needed by others to reduce uncertainty or because their resources are specialized or are central to the workflow of the organization (Lachman 1989). Researchers have paid much attention to structural conditions associated with power in organizations. Spaeth (1985) found that resource allocation is central to task performance since lower-level employees are assigned to produce given outputs and are provided with the necessary resources to do so. The higher the level of the worker, the more discretionary resources she or he will have to allocate. Recently, researchers have explored the relationship between technological innovation and shifts of power among organizational members. Burkhardt and Brass (1990) found that early adopters of a computerized information system in a federal agency increased their centrality and power in organizational networks. These shifts did not completely alter the power structure since those in power were not completely displaced by early technology adopters. Barley (1986) makes clear that technology provides organizational members an occasion for structuring." The same technological system may have different implications and may cause different social changes in different organizational structures. Thus, Barley disputes the claims by some researchers that

technology has objective material consequences regardless of the social contexts within which it exists.

Scott (1998) writes that organizations attempt to build structures not only to accomplish a division of labor, but also to create a structure of authority. As organizations become more formalized—that is, as procedures and rules are explicitly formulated—power differentials are built into the system and institutionalized. The distributive advantage of upper participants is not obvious since hierarchy is presumably built on specific task competence and power is vested for specific task achievement. Those with institutionalized power need not mobilize to have their interests served, since they control the flow of resources and information. The institutionalization of power contradicts resource control theory, which asserts that those who control the contingencies for change in organizations gain power. Lachman (1989) emphasizes that when the relative power of subunits changes, the previous power structure significantly affects the new one. He found that the greatest predictor of subunit power after organizational change was its degree of power before the change, regardless of its control over organizational contingencies or changes.

The complexity of goals within organizations creates a number of difficulties in assessing effectiveness. The use of *goals* as virtually the sole criterion for effectiveness places excessive focus on the outcomes, rather than the processes, of organizational activities. While some goals may be universally held within an organization, these are usually held with varying degrees of intensity by various constituencies. Moreover, multiple, and sometimes conflicting, goals are held by various organizational constituencies; notably those at different hierarchical levels, but also those defined by a variety of other factors. Goals may be difficult to operationalize for organizations such as human service agencies, and "fuzzy goals" may be adaptive in avoiding conflict or in concealing operative goals (Weiss 1972). Often, the evaluation of organizational effectiveness is skewed toward one or another constituency's view of what constitute legitimate organizational goals or the priority assigned to given goals. Also, goal displacement may occur when compliance with indicators of good job performance (such as the number of calls taken per hour at a computer help desk) supplants



the intended outcome of job performance (solving computer users' technical problems).

Accordingly, evaluation researchers have suggested a number of strategies for resolving some of the conflicts in determining and measuring goals. Stake (1975) suggests "responsive" evaluation that takes into account multiple and conflicting constituencies' goals, treating each as a separate and valid view of organizational functioning. One of the goals of responsive evaluation is to inform each group of the others' perspectives. Patton (1997) suggests a shift from goals to outcomes, focusing on expected changes, maintenance, or prevention that are the intended focus of organizational activity. Scriven (1975) suggests "goal-free" evaluation that focuses on only the actual, measurable outcomes of organizational activities compared to a profile of the demonstrated needs of participants. "Goal-free" evaluation does not include a formal elicitation of goals nor a review of formal organizational documents, but focuses instead on the needs of organizational participants, whether these are formally stated or not. The success that these alternative perspectives have in circumventing goals as the focus of effectiveness evaluation is variable, but they do constitute an acknowledgement of the difficulties in assessing organizational effectiveness solely through goal attainment.

Organizational *efficiency*, like *effectiveness*, is also tied to the attainment of organizational goals, but measures attainment against its costs. Cost-effectiveness analysis measures the monetary benefits of a program against the financial costs of implementing it. Cost-benefit analysis can estimate the broadly conceived benefits of a program versus its various costs, but usually also monetizes the benefits and compares them. Not only do these types of analyses face the same challenges described above in determining how the effectiveness of organizational activity is to be determined, they must choose a cost-accounting perspective that also reflects a particular "interest group" perspective. Freeman and Rossi (1993) describe three cost-accounting perspectives: (1) The individual level weighs costs, benefits, and effectiveness from the perspective of the unit targeted by an organization (people, groups, or organizations). (2) The program sponsor level takes the perspective of the sponsor in valuing benefits and accounting for costs, and examines the profitability of a

given organizational activity. 3) The communal (or societal) perspective assesses both the direct and the indirect effects of organizational activity in a broad context and for a wide variety of participants. Both upper-level organizational participants and researchers have overemphasized the program-sponsor perspective, emphasizing short-term impacts and economic outcomes and short-changing factors that are integral to self-perceived well-being of often less powerful constituencies, such as the impact on their psychic, social, cultural, or political condition. Recent developments in "socioeconomics" seek to develop theoretical approaches that may bring these considerations under systematic analysis (Etzioni 1990, 1996, 1998; Granovetter 1985, 1992; Granovetter and Swedberg 1992). Monetizing such factors is difficult and controversial, but the costs and benefits of broader social and cultural factors both in monetary and other forms is included in efficiency studies using the communal perspective.

Economic analyses, however, frame their results in a monetary context for use by managers or policy makers, framing outcomes in ways that are less relevant to mid- and low-level workers, who do not control the allocation of resources. Additionally, it is difficult (and controversial) to monetize benefits, especially those involving complex social contexts, such as cultural and political changes. It may be difficult to combine multiple and/or conflicting goals in economic analyses, including the unintended consequences of organizational action. Finally, by definition, economic analysis takes an incremental approach to change, assuming that benefit is achieved if the benefits outweigh the costs. For some kinds of organizational activities, an incremental approach toward cost-effectiveness or cost-benefit analysis may not be appropriate if a partial change in outcomes is not desirable. For example, many programs designed to prevent drug use do not measure outcomes in terms of decreased drug use among participants, but in the extent to which total nonuse of drugs may be effectuated.

Setting and attainment of goals are complex processes affected by factors internal and external to organizations and by processes for power distribution. The *cui bono* criterion alerts the researcher not to take formal goals at face value but to identify how key actors and groups in coalitions

differentially benefit from goal activity. In spite of its tendency to persist, organizational structure can be and is altered to reflect the power of new alliances among internal and external constituencies who benefit from new institutional arrangements. Organizations survive because powerful constituent alliances continue to derive benefit from them.

### THE INTERNAL INTEGRATION OF ORGANIZATIONAL STRUCTURE

Organizations cohere in part because some elements of organizational structure are designated to coordinate other organizational elements into a collective “whole” in the pursuit of goals. This discussion of integration will examine (1) how some organizational variables affect composition of organizational structural; (2) structure’s differential outcomes for stratified groups; and (3) who benefits from integration.

One factor important to integrating organizational members is technology, knowledge about how to get things done. Organizations often must increase coordination to achieve technically complicated tasks. Galbraith (1977) notes four mechanisms for increasing task coordination: (1) Rules standardize both acceptable actions and agreed-upon ends. (2) Schedules coordinate interdependent activities or multiple activities occurring at the same time. (3) Departmentalization routinizes the division of labor by grouping homogeneous tasks together. (4) Organizational hierarchy helps to coordinate tasks that are interrelated among departments. Organizational size can be an important factor affecting internal structure. Several measures have been used for size, including the physical plant, number of clients, or number of employees (Kimberly 1976). Most researchers have treated size as a determinant of organizational structure. It can be an indicator of demand for organizational services or products, providing constraints or opportunities for structural change.

These two variables are related to a number of structural outcomes, the most important of which are complexity or differentiation, formalization, and centralization. Complexity is the diversity of factors that must be simultaneously coordinated to get a task done. Horizontal differentiation is the

specification of component elements of tasks performed at the same level of organizational structure. Vertical differentiation is the number of levels of importance, power, and control among units. Multiple and complex technologies are associated with increased structural differentiation (Dewar and Hage 1978). Increased size (number of participants) is necessary to achieve high degrees of structural differentiation, although a large organization may be minimally differentiated and simply coordinated. Nevertheless, larger organizations are generally more structurally differentiated and more complexly coordinated (cf. Blau and Schoenherr 1971).

*Formalization* is the extent to which rules for behaviors and relations in organizations are explicitly specified for participants directly or indirectly associated with particular tasks. This means that formalized positions have standardized powers and duties, regardless of the particular individual incumbent. Organizations with routine technology have a greater degree of formalization (Dornbusch and Scott 1975) at a more minute level of detail. Glisson (1978) contends that routinized technology may produce greater formalization. Formalization, however, is only moderately associated with organizational size (e.g., Blau and Schoenherr 1971).

Centralization refers to the extent to which decision making in an organization is concentrated or dispersed. Dornbusch and Scott (1975) point out that organizations may engage multiple technologies that vary in terms of clarity, predictability, and efficacy. Tasks high on these dimensions are more likely to be centralized, but tasks low on the dimensions are likely to be allocated to specialists, decentralizing organizational authority. Centralization through rules is associated with use of routine technology (Hage and Aiken 1967). There is an inverse relationship between size and centralization: Larger organizations tend to decentralize decision making.

Recent organizational restructuring in the United States has resulted in some organizations becoming “smaller” in size than they were three decades ago. The adoption of new technologies including computer-assisted production and just-in-time inventory systems have resulted in a decrease in vertical hierarchy, and a flattened lateral

hierarchy. At the same time, the complexity of the technology and the need for increased coordination has increased interdependence of organizational subunits or “teams,” while centralization has decreased (Galbraith and Lawler 1993). Reengineering technologies has often resulted in changing the processes of work, rather than directly changing organizational structures, but has contributed to the increasingly interdependent relations among organizational subunits (Keidel 1994). Such reengineering has led to flexibility in team function, resulting in “semistuctures” in which some aspects of performance are prescribed, but the processes through which teams operate are not (Brown and Eisenhardt 1995). Scott (1998) notes, however, that the resources needed to successfully alter structure to function in a more lateral, independent fashion are considerable, and many organizations do not have the resources to make this kind of structural transformation.

Consistently with generalized values in the host society, organizational structure has differential outcomes for people of diverse identities. Labor theorists offer various explanations for *occupational segregation*, the organizational practice of reserving specific kinds of jobs for people of given ethnicity or gender. Occupational segregation results in lower earnings in jobs traditionally held by relatively powerless groups—minority men and white and minority women. Kaufman (1986) found that black men were more likely than white men to be employed in highly routinized, less skilled jobs and were less likely to be in job ladders involving increasing levels of status dominance over other workers. Tienda and colleagues (1987) found that, although women’s earnings increase as they move into jobs traditionally held by males, male incumbents experience a greater increase in earnings than their female counterparts. Some researchers contend that market forces have created an occupationally segregated labor force, while others believe that organizational practices are responsible. Bridges and Nelson (1989) argue that although market forces produce gender-related inequalities (as expected), these are exacerbated by intraorganizational decisions resulting in preservation of occupational segregation.

Although hierarchy appears to be determined rationally by high-level managers and by owners, labor theorists reveal that organizational structure

has differential outcomes for people based on ethnicity and gender. Further, the structural inequalities between wages and perquisites for workers at various levels of the organization may be quite large and not necessary to maintain a task-oriented division of labor on which an organization’s survival might depend. Marxists, work redesign theorists, and researchers exploring organizational “structuring” explicitly address how organizations benefit, or fail to benefit, worker categories. These perspectives challenge the presumption of rationality for organizational structures imbalanced in favor of a given constituency. Problems in assessing benefit are due to the assumption that managers or the environment determine what middle and lower participants do without exploring resistance to more generally beneficial institutional policies by upper participants. Researchers examining the impact of variables such as size on formalization frequently exclude human agents, and thereby reify organizational structure. Such research fails to give due consideration to how organizational structuring might account for these phenomena. To the extent that workers successfully and persistently subvert, modify, or resist organizational structuring of their activities, researchers cannot justify the assumption that structure is the sole result of owner, managerial, or environmental influences.

#### LATENT FACTORS IN STRATEGIC PLANNING

Values guide and give meaning to activities. A set of activities, however, may be variously interpreted from alternative value perspectives. Activities that are rational and meaningfully appropriate for an organizational task (in a bank, for example) may also be in substantial congruity with prescriptions and proscriptions of, say, an ostensibly unrelated religious organization serving the same client population. This hidden or latent positive affinity might make a bank clerk appear particularly productive as new customers are attracted to the bank and give favorable reactions for service received from the bank in spite of, rather than because of, role performance as bank clerk. The latent import could be negative, in which case the bank clerk might look particularly unproductive.

Although there is no assurance an organizational structure staffed with a demographically

diverse population has a higher probability of survival, rational theorists believe its chances are increased. As an organization takes in participants with demographic characteristics different from those of previous participants, modalities of personal values at different participant levels will change. This creates a latent potential for future pressure to change goals and procedures. For effective long-term strategic planning, decision makers are well advised to consider implications of alternative demographic concentrations among actual or potential external constituencies and among each internal participant level. Cultivation of new markets to absorb organizational output may require recruitment from populations with new attributes as much as recruitment from new populations to reduce labor costs may require acceptance of new types of participant contributions. Such changes inherently precipitate structural change. An organization that sets narrow limits to organizational culture may as a latent consequence inadvertently limit its capacity to adapt to altered environmental conditions and thus decrease its capacity to survive.

Research examining the link between organizational culture and effectiveness has reflected managerial interest in improving organizational performance through the control and directed change of organizational culture. Research examining the link has been mixed, finding a direct relationship between the two (e.g., Petty et al. 1995), examining culture as a set of mediating variables in organizational performance (e.g., Saffold 1988), and finding no relationship between the two (e.g., Reynolds 1986). While most studies have examined factors internal to organizations in the construction of culture, some have examined the links between culture and organizational markets, networks, and environmental factors (e.g., Burt et al. 1994). One reason for the mixed findings in the association between organizational culture and effectiveness is a difficulty in defining and measuring culture. Culture has been variously defined as underlying shared values and beliefs among participants (Schein 1991); meanings of actions among individual participants (Golden 1992), and means of governing transactions among individual participants (Ouchi 1980). Culture has therefore been measured in terms of strength, meaning the congruence of managerial beliefs to managerial practices (Denison 1990);

individual practices, affect, and cognitions (Silvester et al. 1999); and organizational types governing transactions among participants, which exhibit traits such as the degree of hierarchy or flexibility (Ouchi 1980). These differences in the definitions of culture and in its measurement account for some of the variability of findings in the organizational culture literature. Additionally, differences in organization types and criteria of effectiveness may also account for differential findings of the link between culture and organizational performance. A struggle between proponents of qualitative and quantitative methods has led to a bifurcated literature and a lack of multimethod studies examining the complexities of organizational culture (Martin and Frost 1996). Most studies on organizational culture and effectiveness focus on overall performance, profit, and productivity from the perspectives of the organizational elite, upper management, and administrators while purporting to describe a “whole” organization’s culture and outcomes.

While organizational culture is both an important *result* and a *determinant* of organizational structure, it appears to have been given both narrow and exaggerated importance by writers whose focus is on profit-making organizations. In the mid-1980s researchers contended that centralized control of corporate culture by managers was correlated with organizational success (e.g., Deal and Kennedy 1982). Meyerson and Martin (1987) call this the integrationist perspective, but argue that organizational theory has encompassed at least two other perspectives, the differentiation perspective and the fragmentation perspective. The differentiation perspective acknowledges that there may be “nested” cultures and subcultures within organizations that may embrace different forms, ideologies, and rituals. The fragmentation perspective observes that clearly conflicting cultural understandings and practices in organizations may exist, and may shift as coalitions form around specific tasks and as organization members deal with the ambiguity of their tasks.

Martin (1992) notes that one or all of these cultural characteristics may be present within organizations, but that participants’ positions throughout the organizational structure may affect how they perceive an organization’s culture. Management may take an integrationist view of culture,

seeing it as a way of controlling and coordinating organizational activity. Rank-and-file workers may be more likely to adopt a differentiation perspective instead, and those whose jobs are characterized by a high degree of ambiguity and environmental uncertainty may be more likely to adopt a fragmentation perspective. The issue here is how much of values essential to its own well-being each constituency will surrender in return for real or imagined benefits derived from continued organizational participation. Upper participants are only one constituency potentially benefiting from organizational culture. Other constituencies may use it as a vehicle to express resistance to, or interest in, a wide array of organizational and environmental conditions. If upper participants seek effectiveness in goal achievement and efficient productivity for increased profits, lower participants might seek security of employment and quality of working conditions. Participants may see the organization as a forum within which each subgroup expresses its own array of self-interests for potential incorporation into organizational culture in the context of collective, perhaps universal, concern for organizational survival. Depending on their relative power (location) within the organization's structure, constituents may be limited in the degree to which they can enact and establish understandings about the distribution of rewards (of various kinds including but not limited to economic rewards) relative to their own structural positions and to their self-perceived contributions to organizational survival.

If participants decide the distribution of rewards is unsatisfactory or inequitable relative to some other comparison group or individual, relative deprivation theorists argue that participants will then attempt systemic or individual remedies. Indeed, lower participants may choose to treat "official" cultural accounts with disbelief or to subvert them (Smircich 1983). The form that perceived unfairness takes influences the actions that people are likely to take; Skarlicki and Folger (1997) found that if participants perceive organizations as having low procedural justice and low distributive justice, they are more likely to take retributive action against an organization for perceived injustices than those who believe that only procedural or distributive justice is low. However, researchers have failed to demonstrate how unfavorable comparisons cause behavioral outcomes.

Historically, groups of workers have understood existence of ethnicity and gender inequality in the work place, but not all groups have taken steps to remedy the problems. The organizational justice perspective may allow exploration of organizational structuring as a process in which agency exists on the parts of participants and constituents of organizations at both aggregate and individual levels in accord with their cultural orientation.

### CONCLUDING REMARKS

Current theory and research on organizational structure have been characterized by both managerial determinism and by a vague environmental determinism used to exonerate owners and managers from detrimental consequences of their structural designs upon middle and lower participants within the organization. Owners and higher-level managers often overemphasize short-term gains achieved by production and managerial techniques undermining relative empowerment of lower participants. On their behalf, organizational theorists and researchers contribute unwittingly to intensification of unnecessary control over workers, consumers, community residents, and other types of participants (Clegg 1981). Research guided by an organizational justice perspective may reduce an overemphasis on the competitive advantage of upper participants (disguised as advantage by the *organization* against external adversaries) and open new vistas on cooperative actions beneficial to all participants and related to the perceived value of their contributions to organizational survival. In the final analysis, perhaps the only justification for rewards received by a category of organizational participants is the contribution it makes to the essential task of organizational survival (Alvarez 1979).

(SEE ALSO: *Complex Organizations*; *Work and Occupations*; *Professions*)

### REFERENCES

- Aldrich, Howard E. 1979 *Organizations and Environments*. Englewood Cliffs, N.J.: Prentice-Hall.
- Alvarez, Rodolfo 1979 "Institutional Discrimination in Organizations and Their Environments." In R. Alvarez and K. Lutterman, eds., *Discrimination in Organizations: Using Social Indicators to Manage Social Change*. San Francisco: Jossey-Bass.

- Baker, Wayne E., Robert R. Faulkner, and Gene A. Fisher "Hazards of the Market: The Continuity and Dissolution of Interorganizational Market Relationships." *American Sociological Review* 63(2):147-177.
- Barnard, Chester I. 1938 *The Functions of the Executive*. Cambridge, Mass.: Harvard University Press.
- Barley, Stephen R. 1986 "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments." *Administrative Science Quarterly* 31:78-108.
- Blau, Peter M. 1994 *The Structural Contexts of Opportunities*. Chicago, Ill.: University of Chicago Press.
- , and Richard Schoenherr 1971 *The Structure of Organizations*. New York: Basic Books.
- Blau, Peter, and W. Richard Scott 1962 *Formal Organizations*. New York: John Wiley.
- Bridges, William P., and Robert L. Nelson 1989 "Markets in Hierarchies: Organizational and Market Influences on Gender Inequality in a State Pay System." *American Journal of Sociology* 95:616-658.
- Brockner, Joel, and Phyllis Siegel. 1995 "Understanding the Interaction Between Procedural and Distributive Justice: The Role of Trust." In Roderick M. Kramer and Tom R. Tyler, eds., *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, Calif.: Sage.
- Brown, Shona L., and Kathleen M. Eisenhardt 1995 "Product Innovation as Core Capacity: The Art of Continuous Change." *Academy of Management Review* 20(2):343-375.
- Burkhardt, Malene E., and Daniel J. Brass 1990 "Changing Patterns or Patterns of Change: The Effects of a Change in Technology of Social Network Structure and Power." *Administrative Science Quarterly* 35:104-127.
- Burt, Ronald S., Shaul M. Gabbay, Gerhard Holt, and Peter Moran 1994 "Contingent Organization as a Network Theory: The Culture-Performance Contingency Function." *Acta Sociologica* 37:345-370.
- Clegg, Stewart 1981 "Organization and Control." *Administrative Science Quarterly* 33:24-60.
- Cohen, Lisa E., Joseph E. Brosehak, and Heather A. Haveman 1998 "And Then There Were More? The Effect of Organizational Sex Composition on the Hiring and Promotion of Managers." *American Sociological Review* 63:711-727.
- Cook, Karen S., and Toshio Yamagishi 1983 "Social Determinants of Equity judgments." In David M. Messick and Karen Cook, eds., *Equity Theory*. New York: Praeger.
- Deal, Terrence, and Allan Kennedy 1982 *Corporate Cultures*. Reading, Mass.: Addison-Wesley.
- Dess, Gregory G., and Donald W. Beard 1984 "Dimensions of Organizational Task Environment." *Administrative Science Quarterly* 29:52-73.
- Denison, D. 1990 *Corporate Culture and Organizational Effectiveness*. New York: John Wiley.
- Dewar, Robert D., and Jerald Hage 1978 "Size, Technology, Complexity and Structural Differentiation: Toward a Theoretical Synthesis." *Administrative Science Quarterly* 23:111-136.
- Dill, William R. 1958 "Environment as an Influence on Managerial Autonomy." *Administrative Science Quarterly* 2:409-443.
- DiMaggio, Paul J., and Walter W. Powell 1983 "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147-160.
- Dornbusch, Sanford M., and W. Richard Scott 1975 *Evaluation and the Exercise of Authority*. San Francisco: Jossey-Bass.
- Emerson, Richard M. 1962 "Power-Dependence Relations." *American Sociological Review* 27:31-40.
- Etzioni, Amitai 1961 *A Comparative Analysis of Complex Organization*. New York: Free Press.
- 1990 "A New Kind of Socio-Economics (The Ascent of Socio-Economics)." *Challenge* 33(1):31.
- 1996 "How to Make a Humane Market: Quality of Life or Economic Success?" *New Statesman* 127(4412):25-28.
- 1998 "A Communitarian Note on Stakeholder Theory." *Business Ethics Quarterly* 8(4):679.
- Fligstein, Neil, and Iona Mara-Drita 1996 "How to Make a Market: Reflections on the Attempt to Create a Single Market in the European Union." *American Journal of Sociology* 102:1-33.
- Freeman, David M., and Peter Rossi 1993 "Using Theory to Improve Program and Policy Education." *Social Forces* 71(4):1102-1103.
- Galbraith, Jay 1977 *Organization Design*. Reading, Mass.: Addison-Wesley.
- , and Edward E. Lawler, III (eds.) 1993 *Organizing for the Future: The New Logic for Managing Complex Organizations*. San Francisco, Calif.: Jossey-Bass.
- Glisson, Charles 1978 "Dependence of Technological Routinizations of Structural Variables in Human Service Organizations." *Administrative Science Quarterly* 23:383-395.
- Golden, K. A. 1992 "The Individual and Organizational Culture: Strategies for Action in Highly-Ordered Contexts." *Journal of Management Studies* 29:1-21.

- Granovetter, Mark 1985 "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.
- 1992 "Economic Institutions as Social Constructions: A Framework for Analysis." *Acta Sociologica* 35(1):3-12.
- , and Richard Swedberg 1992 *The Sociology of Economic Life*. San Francisco: Westview.
- Hage, Jerald, and Michael Aiken 1967 "Relationship of Centralization to Other Structural Properties." *Administrative Science Quarterly* 12:72-91.
- Huo, Juen J., Heather J. Smith, Tom R. Tyler, and E. Allan Lind 1996 "Superordinate Identification, Subgroup Identification, and Justice Concerns: Is Separatism the Problem; Is Assimilation the Answer?" *Psychological Science* 7(1):40-45.
- Irwin, Sarah 1996 "Age Related Distributive Justice and Claims on Resources." *British Journal of Sociology* 47(1):68-92.
- Kaufman, Robert L. 1986 "The Impact of Industrial and Occupational Structure on Black-White Employment Allocation." *American Sociological Review* 310-323.
- Keidel, Robert W. 1994 "Rethinking Organizational Design." *Academy of Management Executives* 8(4):12-30.
- Kimberly, John 1976 "Organizational Size and the Structuralist Perspective." *Administrative Science Quarterly* 29:571-597.
- Kochan, Thomas, George Cummings, and Larry Huber 1976 "Operationalizing the Concepts of Goals and Goal Incompatibility in Organizational Behavior Research." *Human Relations* 29:544-577.
- Lachman, Alan 1989 "Power from What?" *Administrative Science Quarterly* 34:231-251.
- Lawrence, Paul R., and Jay W. Lorsch 1967 *Organization and Environment: Managing Differentiation and Integration*. Boston: Graduate School of Business Administration, Harvard University.
- Lee, Cynthia, and Jiing-Lih Farh 1999 "The Effects of Gender in Organizational Justice Perception." *Journal of Organizational Behavior* 20:133-143.
- Lee, Jason S., and John F. Stolte 1994 "Cultural and Structural Determinants of Justice Reactions in the Economic Domain." *Social Behavior and Personality* 22(4):319-328.
- Lerner, Melvin J. 1996 "Victims without Harmdoers: Human Casualties in the Pursuit of Corporate Efficiency." In Leo Montada and Melvin J. Lerner, eds., *Current Societal Concerns about Justice*. New York: Plenum.
- Linkey, Helen, and Sheldon Alexander 1998 "Need Norm, Demographic Influence, Social Role, and Justice Judgement." *Current Psychology* 17(2-3):152-162.
- Martin, Joanne 1992 *Cultures in Organizations: Three Perspectives*. New York: Oxford University Press.
- , and Peter Frost 1996 "The Organizational Culture War Games: a Struggle for Intellectual Dominance." In Stewart R. Clegg, Cynthia Hardy, and Walter R. Nord, eds., *Handbook of Organizational Studies*. Thousand Oaks, Calif.: Sage.
- Meyer, Marshall W. 1975 "Leadership and Organizational Structure." *American Journal of Sociology* 81:514-542.
- Meyerson, D., and J. Martin 1987 "Cultural Change: An Integration of Three Different Views." *Journal of Management Studies* 24:623-647.
- Oliver, Amalya 1997 "On the Nexus of Organizations and Professions: Networking through Trust." *Sociological Inquiry* 67(2):227-245.
- Ouchi, William G. 1980 "Markets, Bureaucracies and Clans." *Administrative Science Quarterly* 25:129-141.
- Parsons, Talbot 1960 *Structure and Process in Modern Society*. New York: Free Press.
- Patton, Michael Quinn 1997 *Utilization-Focused Evaluation*. Thousand Oaks, Calif.: Sage.
- Perrow, Charles 1961 "The Analysis of Goals in Complex Organizations." *American Sociological Review* 26:854-866.
- Petty, M. M., A. Beadles, II, Christopher M. Lowery, Deborah F. Chapman, and David W. Connell 1995 "Relationships between Organizational Culture and Organizational Performance." *Psychological Reports* 76:483-492.
- Pfeffer, Jeffrey, and Gerald R. Salancik 1978 *The External Control of Organizations*. New York: Harper and Row.
- Powell, Walter W. 1996 "The Capitalist Firm in the 21st Century: Emerging Patterns." Paper presented at the Annual Meetings of the American Sociological Association, New York, August.
- Pugh D. S., and D. J. Hickson 1996 *Writers on Organizations*, 5th ed. Thousand Oaks, Calif.: Sage.
- Reynolds, P. D., 1986 "Organizational Culture as Related to Industry, Position, and Performance: A Preliminary Report." *Journal of Management Studies* 23:333-345.
- Rossi, Peter H., and Howard E. Freeman 1993 *Evaluation: A Systematic Approach*. Thousand Oaks, Calif.: Sage.
- Saffold, G. S., III 1988 "Culture Traits, Strength, and Organizational Performance." *Academy of Management Review* 12(4):546-558.
- Schein, Edgar H. 1991 "What is Culture?" In Peter J. Frost, Larry F. Moore, Meryl Reis Louis, Craig C.

- Lundber, and Joanne Martin, eds., *Reframing Organizational Culture*. Thousand Oaks, Calif.: Sage.
- Scott, W. Richard 1998 *Organizations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Scriven, Michael 1972 "Pros and Cons about Goal-Free Evaluation." *Evaluation: The Journal of Educational Evaluation* (Center for the Study of Evaluation, UCLA) 3(4):1-7.
- Silvester, Joanne, Neil R. Anderson, and Fiona Patterson 1999 "Organizational Culture Change: An Inter-Group Attributional Analysis." *Journal of Occupational and Organizational Psychology* 72:1-23.
- Simon, Herbert A. 1964 "On the Concept of Organizational Goals." *Administrative Science Quarterly* 9:122.
- Skarlicki, Daniel P., and Robert Folger 1997 "Retaliation in the Workplace: The Roles of Distributive, Procedural, and Interactional Justice." *Journal of Applied Psychology* 82(3):434-443.
- Smircich, Linda 1983 "Organizations as Shared Meanings." In Louis Pondy, ed., *Organizational Symbolism*. Greenwich, Conn.: JAI.
- South, Scott J., and Woman Xu 1990 "Local Industrial Dominance and Earnings Attainment." *American Sociological Review* 55:591-599.
- Spaeth, Joe L. 1985 "Job Power and Earnings." *American Sociological Review* 50:603-617.
- Stake, Robert E. 1975 *Evaluating the Arts in Education: A Responsive Approach*. Columbus, Ohio: Charles E. Merrill.
- Sweeney, Paul D., and Dean B. McFarlin 1997 "Process and Outcome: Gender Differences in the Assessment of Judgement." *Journal of Organizational Behavior* 18:83-98.
- Thompson, James D., and William McEwen 1958 "Organizational Goals and Environments: Goal Setting as an Interaction Process." *American Sociological Review* 23:23-31.
- Tienda, Marta, Shelley A. Smith, and Villma Ortiz 1987 "Industrial Restructuring, Gender Segregation, and Sex Differences in Earnings." *American Sociological Review* 52:195-210.
- Weick, Karl E. 1977 "Re-Punctuating the Problem." In Paul S. Goodman and Johnnes M. Pennings, eds., *New Perspectives on Organizational Effectiveness*. San Francisco: Jossey-Bass.
- Weiss, Carol H. 1972 *Evaluation Research: Methods of Assessing Program Effectiveness*. Englewood Cliffs, N.J.: Prentice-Hall.

RODOLFO ALVAREZ  
LEAH ROBIN

## ORGANIZED CRIME

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

Organized crime is considered one of the most serious forms of crime for two reasons: (1) It is so often lucrative and successful; and (2) it is so difficult to counteract. In the broadest terms, organized crime can be viewed as any form of group conduct designed to take advantage of criminal opportunities, whether on a one-time or a recurring basis. More commonly, the label *organized crime* has more restricted usage.

It should not be a surprise to find criminals associating for the purpose of committing crime. The achievement of goals through cooperative efforts is a common element of contemporary life. Association with other criminals creates an interesting dilemma for the individual criminal. Having coconspirators increases the visibility of criminal conduct, the risks of apprehension, and, upon apprehension, the risk of betrayal. On the other hand, some types of criminal opportunities can be exploited only through group behavior. Offenders who fail to join with others may thereby limit their rewards from criminal conduct. For organized crime to persist, it must function both to overcome the risks of associating with others and result in positive benefits and increased rewards for those who participate.

No one knows how much organized crime is committed every year. Rather than being measured in numbers of events, the significance of organized crime is generally recorded in the dollar volume of activities. Some have argued that revenue estimates for organized crime are advanced more in the interest of drama than of accuracy, but the estimates are nonetheless staggering. For example, annual revenue from the sale of drugs is placed in the \$40 to \$60 billion range, and annual revenue from illicit gambling operations, at \$20 to \$40 billion.

Organized crime is unlikely to be reflected in official crime statistics for three reasons: first, crime statistics record information about individual criminal events rather than about the individuals or groups committing them; second, many of



the activities of organized crime groups involve so-called victimless crimes where no report of a crime is made; and third, persons who are victimized by organized crime groups may be unlikely or unwilling to come forward and report their victimization.

Organized crime is best understood by examining the nature of ongoing criminal organizations, their activities, and societal response to the behavior of these organizations. A secondary use of the term, referring to criminal support systems that aid all offenders, is also briefly noted.

### ONGOING CRIMINAL ORGANIZATIONS

Organized crime usually refers to the activities of stable groups or gangs that commit crimes on an ongoing basis. While not all group crime is properly labeled “organized crime,” there exists no standard definition of the term. Instead, the phenomenon has been variously described as “a cancerous growth on American society” (Andreoli 1976, p. 21); “one of the queer ladders of social mobility” (Bell 1970, p. 166); “a society that seeks to operate outside the control of the American people and its governments” (President’s Commission 1967, p. 1); and “the product of a self-perpetuating criminal conspiracy to wring exorbitant profits from our society by any means” (Salerno and Tompkins 1969, p. 303).

Scholars have identified the following as elements that characterize all groups labeled as organized crime: a hierarchical organizational structure, dominated by a strong leader; a territorial imperative, exhibited in attempts to monopolize all lucrative criminal opportunities within a geographic area; a predilection to violence both to enforce internal norms and to advance economic objectives; and a desire to influence the social response to criminal conduct, as demonstrated in significant investments in public corruption. Cressey (1969) identified these latter two characteristics—the “element of enforcement” and the “element of corruption”—as the essential features of organized crime.

In the United States, organized crime has been personified in ethnically based criminal organizations and, in particular, the twenty-four fictive “families” believed to make up the American Mafia. Joseph Valachi, in testimony before the United States Senate in 1963, revealed a sordid

and secret world in which he claimed these criminal organizations operated and prospered.

While some scholars pointed to inconsistencies in Valachi’s statements, and other criminals would later question the depth of his experience and knowledge, his testimony spawned a series of books and movies that placed Italian criminal organizations in the forefront of the public’s consciousness with respect to organized crime.

The term *mafia* (with either upper- or lowercase letters) is used variously to apply to a secret criminal organization or to a life-style and philosophy that developed in sixteenth- and seventeenth-century southern Italy and Sicily in opposition to a series of foreign rulers who dominated the area. The life-style combines the idea of manliness in the face of adversity with an antagonism to authority and a closeness and reliance on family and clan.

Smith (1975) argues that applying the “mafia” label to Italian criminal organizations in the United States was both fateful and calculated. It was fateful because it imbued these organizations with international connections that there was little evidence they had. This may in turn have helped these organizations consolidate and solidify their power, in competition with other ethnic (primarily Irish) criminal organizations that were actually more prevalent in the early decades of the twentieth century. Use of the label was calculated in that it gave Italian criminal organizations a subversive and sinister character designed to ensure public support for extraordinary law enforcement efforts to eradicate these groups.

Overlooked in much of the attention paid to Mafia “families” was the fact that bands of brigands and smugglers, displaying many of the same organizational attributes as “mafia,” were well established in Elizabethan England; that criminal organizations linked to a life-style based on the primacy of male-dominated families and the concept of dignity and manliness are common in Central and Latin America; and that criminal societies and associations have been successful vehicles for social mobility in many cultures, for example, the Chinese Triads or the Japanese Yakuza.

Also overlooked were some significant changes occurring during the late 1960s and early 1970s in the major cities in which the Mafia “families” were

believed to operate. These changes signaled what Ianni (1974) called “ethnic succession” in organized crime: That is, as the populations of inner cities came to be dominated by racial minorities, so too were the ranks of those running criminal organizations. This gave rise to law enforcement characterizations of such groups as the “Mexican Mafia” or the “Black Mafia,” not because of their associations with Mafia “families” but as a short-hand way of communicating their organizational style and methods of operating.

By the 1980s, the term *organized crime* had lost its ethnic distinction. Instead, the label started being applied to many more criminal associations, from motorcycle and prison gangs to terrorist groups to some juvenile gangs. Observed in all these groups were the organizational characteristics that had first been identified as distinctive of Mafia “families.”

#### ACTIVITIES OF ORGANIZED CRIME GROUPS

The activities in which organized crime groups are involved constitute another distinctive hallmark. Broadly defined, the term *organized crime* can be used to describe the activities of a band of pickpockets, a gang of train robbers, or a cartel of drug smugglers. Practically speaking, however, use of the term is somewhat more restrictive.

Commentators, scholars, and lawmakers generally use the term when referring to criminal conspiracies of an entrepreneurial nature—in particular, enterprises focused in black-market goods and services. Black markets are those in which contraband or illegal goods and services are exchanged. In this more restrictive use of the term, the activities of drug smugglers would be included while the conduct of pickpockets or train robbers would not.

These latter groups, and others involved in various forms of theft and extortion, commit what are best termed *predatory crimes*. As such, they are generally regarded as social pariahs. Society and the forces of social control will actively seek to root out such groups and bring them to justice. This is despite the fact that such groups may display a highly evolved organizational structure, a strong sense of territory, and a tendency toward violence.

Contrast this with groups engaged in entrepreneurial conduct. These groups supply the society with goods and services that are illegal but in demand. While some in society may still view these criminals as social pariahs, many in society will not. Instead, the criminal group establishes patron–business relationships with criminal and noncriminal clients alike. The forces of social control are not so bent on eradicating these groups because of the widespread social support they garner. This support, when added to the profits reaped as criminal entrepreneurs, creates both the means and the conditions for corruption.

If any one activity can be considered the incubator for organized crime in the United States, it would be the distribution and sale of illegal alcohol during Prohibition. The period of Prohibition (1920–1933) took a widely used and highly desired commodity and made it illegal; it also created the opportunity for a number of predatory criminal gangs to evolve as entrepreneurs. These entrepreneurs then developed important client and business relationships and emerged from Prohibition as wealthy and somewhat more respectable members of their communities.

Gambling and other vices have similar social support profiles. Purveyors of these services are widely perceived as engaged in victimless crimes, that is, black-market transactions involving willing buyers and sellers. Their activities, while morally unacceptable to many, arouse little social concern.

Societal ambivalence becomes even more pronounced where gray-market goods and services are involved. These are situations where legal goods or services are being provided in an illicit manner or to persons ineligible to receive them. Abuses of wartime rations is a good example of a gray-market situation, as is the negotiation of so-called sweetheart labor contracts or the illegal disposal of hazardous waste.

Criminal groups involved in gray-market activities are very tightly meshed in the economic and social fabric of the legitimate community. The persona of such groups is more likely to be legal than illegal, and the capacity of their members to become closely affiliated with persons of power and authority is likely to be great. Public corruption becomes not only likely but inevitable.

Combined with the characteristics that groups of criminals exhibit, the activities in which they engage are also likely to define such groups as organized crime. Some commentators feel it is impossible to separate a group's character from the nature of the crimes it commits. To this way of thinking, selling drugs requires a certain level of organization, but it is impossible to tell whether a group has evolved an organizational style in order to sell drugs or began to sell drugs as a consequence of its organizational capacity.

What is clear is that the capacity to exploit one type of criminal opportunity can be parlayed into other legal and illegal endeavors. Similarly, profits from organized crime activities can permit individual criminals to climb that "queer ladder of social mobility" that wealth creates.

### **SOCIETY AND ORGANIZED CRIME GROUPS**

The final aspect of organized crime that distinguishes it from other forms of crime is its relationship with the society in which it operates. Organized crime is the one form of crime that assesses criminal opportunities in light of the probable social response as well as the possible economic and social rewards.

Cressey (1969) identified this capacity of organized crime as a "strategic planning" capability. Using this capability, organized crime groups choose "safe crime": where there is high social tolerance or at least ambivalence toward the conduct; where the chances of apprehension are therefore not great; where, even if apprehended, the chances of conviction are small; and where, even if convicted, the likelihood of serious consequences is also small.

In this assessment, society's attitudes toward various criminal activities become a key determinant in the nature of organized crime activities that will be displayed. Society's attitudes, as embodied in the criminal law, become even more significant.

Packer (1969) argued that in black- and gray-market situations, the criminal law may actually serve as a protective tariff. As such, the law limits the entry of entrepreneurs into the proscribed marketplace while guaranteeing, for those who do

enter the market, exorbitant profits. The theory of deterrence does not work in such markets because as the sanction increases so do the likely rewards.

A similar analysis by Smith (1978) suggests that the law operates in many marketplaces to segment genetic demand, labeling some legal and some illegal. By so doing, the law does not reduce demand; what it does is change the dynamics of the market, creating the "domain" or market share of organized crime. In this sense, it is society—through its legislative enactments—that generates and structures the dynamics of organized crime opportunities.

Social institutions also play a role in structuring the nature and success of organized crime. The nature of government and the underpinnings of justice systems loom large in determining how organized crime groups will operate and succeed. Anglo-American legal systems, founded on the principle of individual responsibility for criminal acts, deal at best ineptly with group crime. When faced with more sophisticated criminal conspiracies, they appear to falter.

Where criminal organizations are armed with investments in public corruption, justice systems will not operate properly. Where government operates ineffectively or unfairly, the black market will flourish. Where there is the tendency to proscribe what cannot be controlled, criminal organizations will reap social support and financial rewards.

### **CRIMINAL SUPPORT SYSTEMS**

A secondary use of the term *organized crime* refers to support systems that aid the criminal activities of all offenders. The typical list of support systems includes the tipster, the fix, the fence, and the corrupt public official. Each of these mechanisms serves to reduce the risks of criminal conduct or to lessen its consequences.

For example, tipsters function to provide criminals with information critical to committing a crime, such as the internal security schedule for a building or the timing of valuable cargo shipments. By doing so, they reduce the uncertainties the offenders face and enhance the chances for success. As a reward, tipsters receive a percentage of the "take," or proceeds of the crime.

The fix arranges for special disposition of a criminal's case, once it is in the justice system. This might mean seeing that paperwork is lost, or that a light sentence is imposed. Usually, the fix operates with the aid of corrupt public officials who are in a position to accomplish the required improper acts.

The fence serves as the market for stolen property, transforming it into cash or drugs for thieves. In the role of middleman, the fence determines what thieves steal, how much they are paid to do so, and, consequently, how often they steal. Fences provide structure and stability to a wide range of offenders. Like other criminal support systems, they impose organization on the activities of criminal groups and individual criminals.

Not all criminals can access the services of criminal support systems. These mechanisms will not act to serve the notorious or the psychotic. In this way, criminal elites are established and preserved. For those who use these support systems, crime is organized, the justice system is predictable, and success is likely.

There was a time when the "criminal underworld" was a physical place, a true sanctuary to hide and protect criminals. Now the underworld exists as a communication system, an important dimension of which involves support systems that can aid and protect offenders. The manner in which these systems function provides stability and organization to the underworld.

### SUMMARY

Organized crime refers primarily to the broad range of activities undertaken by permanent criminal organizations having the following characteristics: a hierarchical organizational structure; a territorial imperative; a predilection for violence; and the capacity and funds to corrupt public officials.

These groups tend to locate in gray or black markets where they establish patron-client relationships with much of society. The funds they earn as entrepreneurs, combined with social support for their activities, permit them to influence the social response to their acts. They may become upwardly mobile as a result of investing their profits in both legal and illegal endeavors.

Organized crime also refers to criminal support systems such as the tipster, the fix, the fence,

and the corrupt official, who impose organization and stability in the underworld.

(SEE ALSO: *Criminology*)

### REFERENCES

- Abadinsky, Howard 1981 *Organized Crime*. Boston: Allyn and Bacon.
- Anderson, Annelise Graebner 1979 *The Business of Organized Crime*. Stanford, Calif.: Hoover Institution Press.
- Andreoli, P. D. 1976 "Organized Crime Enterprises—Legal." In S. A. Yefsky, ed., *Law Enforcement Science and Technology: Proceedings of the First National Conference on Law Enforcement Science and Technology*. Chicago: IITRI.
- Arlacchi, Pino 1996 "Mafia: The Sicilian Cosa Nostra." *South European Society and Politics* 1:74–94.
- Bell, Daniel 1970 "Crime as an American Way of Life." In M. E. Wolfgang, L. Savitz, and N. Johnson, eds., *The Sociology of Crime and Delinquency*. New York: Wiley.
- Chambliss, William J. 1978 *On the Take: From Petty Crooks to Presidents*. Bloomington: Indiana University Press.
- Chubb, Judith 1996 "The Mafia, the Market and the State in Italy and Russia." *Journal of Modern Italian Studies* 1:273–291.
- Cressey, Donald R 1969 *Theft of the Nation: The Structure and Operations of Organized Crime in America*. New York: Harper and Row.
- Frisby, Tanya 1998 "The Rise of Organised Crime in Russia: Its Roots and Social Significance." *Europe Asia Studies* 50:27–49.
- Godson, Roy, and William J. Olson 1995 "International Organized Crime." *Society* 2:18–29.
- Homer, Frederic D. 1974 *Guns and Garlic: Myths and Realities of Organized Crime*. West Lafayette, Ind.: Purdue University Press.
- Huang, Frank F. Y., and Michael S. Vaughn 1992 "A Descriptive Analysis of Japanese Organized Crime: The Boryokudan from 1945 to 1988." *International Criminal Justice Review* 2:19–57.
- Huey-Long-Song, John, and John Dombrink 1994 "Asian Emerging Crime Groups: Examining the Definition of Organized Crime." *Criminal Justice Review* 19:228–243.
- Ianni, Francis A. J. 1974 *Black Mafia: Ethnic Succession in Organized Crime*. New York: Simon and Schuster.
- Joe, Karen A. 1994 "The New Criminal Conspiracy? Asian Gangs and Organized Crime in San Francisco." *Journal of Research in Crime and Delinquency* 31:390–415.

- Kelly, Robert J., Ko-Lin Chin, and Jeffrey A. Fagan 1993 "The Dragon Breathes Fire: Chinese Organized Crime in New York City." *Crime, Law and Social Change* 19:245-269.
- Kwitny, Jonathan 1979 *Vicious Circles: The Mafia in the Marketplace*. New York: Morton.
- Maas, Peter 1968 *The Valachi Papers*. New York: Putnam.
- Packer, Herbert L. 1969 *The Limits of the Criminal Sanction*. Stanford, Calif.: Stanford University Press.
- Pons, Philippe 1996 "Social Order in Modern Japan [17th-20th Centuries]. Mobster Control of Vagrancy; Ordre marginal dans le japon moderne (17e-20e siecle). Les Voyous canaliseurs de l'errance." *Annales* 51:1155-1178.
- President's Commission on Law Enforcement and the Administration of Justice 1967 *Task Force Report: Organized Crime*. Washington, D.C.: U.S. Government Printing Office.
- Robertson, Frank 1977 *Triangle of Death: The Inside Story of the Triads—The Chinese Mafia*. London: Routledge and Kegan Paul.
- Salerno, Ralph, and John S. Tompkins 1969 *The Crime Confederation: Cosa Nostra and Allied Operations in Organized Crime*. Garden City, N.Y.: Doubleday.
- Smith, Dwight C., Jr. 1975 *The Mafia Mystique*. New York: Basic Books.
- 1978 "Organized Crime and Entrepreneurship." *International Journal of Criminology and Penology* 6:161-177.
- Walsh, Marilyn E. 1977 *The Fence: A New Look at the World of Property Theft*. Westport, Conn.: Greenwood Press.

MARILYN E. WALSH

## OVERPOPULATION

See Human Ecology and Environmental Analysis; Environmental Sociology; Population.

# P

---

## PARADIGMS AND MODELS

The terms “paradigm” and “model” have enjoyed considerable popularity in sociology, in part because the terms have a range of meanings. In everyday language, *model* has two senses: (1) a replica of an object being modeled, for example, a “model of a building,” and (2) an exemplar to be emulated, as in “role model.” “Paradigm” is somewhat more esoteric in everyday usage, but has become quite important in academic disciplines including sociology, largely due to Thomas S. Kuhn’s 1962 book, *The Structure of Scientific Revolutions*.

The original meaning of *paradigm* overlaps that of “model” in the sense of exemplar. The term comes from the study of grammar where a paradigm provides a model of, for example, the way to conjugate all the regular verbs of a particular type (I love, you love, he or she loves, etc.). Its appearance in sociology predates Kuhn, but since Kuhn, it has become a much grander idea. Merton used the device of analytical paradigms for presenting codified materials, by which he means a technique for exposing the “complete array of assumptions, concepts and basic propositions employed in a sociological analysis” (1949, p. 13). Parsons treated “paradigm” in a similar way when he presented paradigms for social interaction and social change and emphasized that these are distinct from theories (1951, p. 485).

It is quite clear, however, that Kuhn’s work has brought “paradigm” from relative obscurity to a central place in the discourses of the humanities, the social sciences, and the history and philosophy

of science. Kuhn’s paradigm and his model of change in science are highly controversial; yet, they have had considerable influence in these academic fields. Some of the effects Kuhn intended; many effects, however, were unintended, causing Kuhn to disavow explicitly some of the interpretations of his work (Kuhn 1970, 1974). Crews (1986) comments that a loose reading of Kuhn’s book—which he calls the most frequently cited academic book of modern times—even justifies the rejection of science as empirical.

## WORKING DEFINITIONS

A useful definition of *model* draws on Kaplan:

*Any system A is a model of a system B if the elements and relationships of A aid in the understanding of B without regard to any direct or indirect causal connection between A and B.* (1964, p. 263)

For example, one could use ideas about political change (system A)—stable government, crisis, revolution, new government attains stability—to model change in science (system B). Models range from informal analogies or metaphors (e.g., society as an organism) to highly formal equation systems.

A model contains elements that have (or are given) properties—that is, are characterized by descriptive terms, and connections among some or all of these elements. One could construct a physical model of a social network using balls with holes in them and rods that fit into the holes. The rods would connect some but not all of the balls;

the balls (elements of the model) would represent individuals in the network and the rods (connections) would link those individuals who communicate with one another.

The critical property of a model is an isomorphism with its object, the thing being modeled. Strictly speaking, *isomorphism* is a one-to-one correspondence between the elements and relationships in the model and the elements and relationships in its object. Needless to say, strict isomorphism is not satisfied by many metaphors and analogies even though these can aid understanding; the lack of strict isomorphism, however, should serve as a caution in taking analogies too literally. Similar caution applies to models in general since it is all too easy to reify the model, that is, confuse the model with its object. Models do not capture all the properties of their objects, so that even with strict isomorphism, a model inevitably omits significant aspects of the object. As long as those concerned with a model recognize this limitation, they can employ the model fruitfully and unproblematically.

A loose usage of “model” treats it as synonymous with “theory” or as theory expressed in mathematical symbols. While it is possible to have a model of a theory, the working definition given above rules out synonymous usage because a model of  $x$  must resemble  $x$  in terms of pattern or structure, whereas a theory about  $x$ , that predicts or explains  $x$ , need not, and generally does not, resemble  $x$ . In short, theories of  $x$  do not have to be isomorphic with  $x$ ; models do.

In formulating a working definition of *paradigm*, it would seem reasonable to follow Kuhn since he is largely responsible for the contemporary significance of the idea. This presents serious difficulties, however, for as Laudan observes, “Kuhn’s notion of paradigm has been shown to be systematically ambiguous . . . and thus difficult to characterize accurately” (1977, p. 73). While constructing a definition of “paradigm” is a formidable task, it is an essential one because the paradigm is critical to an understanding of Kuhn, and Kuhn’s work plays an important role in many of the current controversies in sociology. Some of the ambiguity disappears if one distinguishes between what a paradigm *is* and what a paradigm *does*; still, it would be presumptuous to believe that one could fully capture Kuhn’s concept with all its

nuances. The following working definition has the more modest goal of representing many of his key ideas:

*A paradigm is a significant scientific achievement recognized by a particular community of scientists that provides a model from which springs a coherent tradition of scientific research and also a general way of looking at the world.*

The principal functions of a paradigm include: (1) determining what kinds of problems are appropriate objects of study, (2) specifying appropriate ways to study these problems, and (3) delimiting the types of theories and explanations that are acceptable. For example, one could consider Durkheim’s *Suicide* ([1897] 1951) as a paradigm for many sociologists that provides a model for research and a way to look at phenomena. For Durkheimians, *Suicide* is an exemplar that specifies ways to look for and at social facts, and that limits explanations of given social facts to other social facts, rather than, for example, explaining social facts with psychological ideas like motives.

It is important to distinguish “paradigm” from “theory.” While a well-developed paradigm may contain a number of specific theories, a paradigm is “metatheoretical.” A theory makes statements about the world, whereas a paradigm involves statements about the nature of acceptable theory, the appropriate entities to investigate, and the correct approaches to these investigations. As an exemplar, a paradigm serves as a normative standard; hence, it is a special type of model, a model of what the given scientific community considers to be exemplary work.

#### KUHN’S MODEL OF SCIENTIFIC CHANGE

Kuhn developed a model of change in science built around his concept of paradigm. The model employs several concepts related to “paradigm” that extend its meaning: normal science, anomaly, crisis, revolution, paradigm shift, and incommensurability. Brief explications of these ideas follow. (Page references below are to Kuhn 1970.)

**Normal Science.** When a paradigm gains acceptance in a scientific community, it is largely in terms of a promise of success if researchers in the community follow the exemplar. “Normal science

consists in the actualization of that promise, an actualization achieved by extending the knowledge of those facts that the paradigm displays as particularly revealing, by increasing the extent of the match between those facts and the paradigm's predictions, and by further articulation of the paradigm itself" (p. 24). To Kuhn, normal science involves "mopping-up operations [that] engage most scientists throughout their careers" (p. 24), and is primarily devoted to puzzle solving: "Perhaps, the most striking feature of the normal research problems . . . is how little they aim to produce major novelties, conceptual or phenomenal" (p. 35).

**Anomaly.** The term *anomaly* refers to "the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science" (p. 52).

**Crisis.** A *crisis*, a period of "pronounced professional insecurity" (p. 68–69), occurs when enough anomalies accumulate so that scientists question the appropriateness of the paradigm.

**Revolution.** Kuhn takes scientific *revolutions* to be "those non-cumulative developmental episodes in which an older paradigm is replaced in whole or in part by an incompatible new one" (p. 92).

**Paradigm shift.** Kuhn argues that scientists with a new paradigm do not merely have new interpretations for what they observe, but rather "see" things differently. This difference in perception, or paradigm shift, is in a sense similar to the Gestalt switch studied by psychologists, in which the same figure can yield a right-hand face or a left-hand face depending on how it is viewed.

**Incommensurability.** During crisis periods, rival paradigms coexist and advocates of rival paradigms cannot understand one another because they have different ways of seeing the world, different standards for appraisal, and different objectives for their scientific community; that is, there is *incommensurability*. Consequently, paradigm conflicts are not resolvable by appeal to a set of shared criteria.

These concepts allow a sketch of Kuhn's model:

Paradigm I → Normal Science → Anomalies → Crisis → Revolution → Paradigm II

When the members of a scientific community accept a paradigm, that acceptance generates a set

of shared commitments to objectives, methods for achieving those objectives and criteria for appraising theories and research. These commitments are the foundation for a period of normal science, since what to study, how to study it, and what constitutes adequate explanations are not problematic. During normal science, the community regards the paradigm itself as unalterable and immune to challenge. Scientists work to solve the puzzles that the paradigm presents and for which the paradigm guarantees solutions; in this process of puzzle solving, refutations for theories contained in the paradigm arise and empirical studies come up with unexpected findings. The more serious refutations and the more surprising findings come to be regarded as anomalies, and as these anomalies accumulate, scientists are at a loss as to how to deal with them. In the crisis period, rival paradigms arise, but because paradigms are incommensurable, scientists adhering to rivals have no common basis for choosing among the competitors. Paradigm II wins out because some scientists undergo the conversion experience of a paradigm shift where they see the world in a new way, because the adherents of Paradigm II are especially persuasive, or because the adherents of Paradigm I die off.

Kuhn's model questions ideas of cumulation and progress in science. While normal science is cumulative, revolutions do not preserve all, or even the most important, achievements of previous paradigms. "There are losses as well as gains in scientific revolutions" (p. 167). Furthermore, the incommensurability thesis argues that Paradigm II does not win out because it is better or more progressive; the lack of shared standards precludes definitive judgments.

## CRITIQUE

Kuhn's work has raised important questions for scholars in the history, philosophy, and sociology of science and has generated a large literature in these areas. Interestingly enough, it has also had broad appeal to intellectuals outside of these specialized fields. Although Kuhn presented his basic model in 1962, his ideas continue to command attention and engender controversy, both within relevant technical fields and in broader academic circles. His work seems to have had less impact in



the natural sciences, although it attacks the conventional views of many natural scientists.

Kuhn's emphasis on scientific communities and on noncognitive factors in the development of science clearly appeals to sociologists, especially sociologists of science—for instance, those who study social networks among scientists or the effects of reward systems. His model provides an antidote to the mythology of cumulative, linear growth of knowledge, the view that characterizes science as building one discovery on another and, with each increment, coming closer to total truth about nature. Nearly forty years after his initial publication, it is difficult to appreciate how dominant such mythology was when Kuhn challenged it.

If prior views overemphasized cognitive and rational factors and held to a naive and simplistic view of progress in science, many post-Kuhnians underemphasize the cognitive and the rational or overemphasize the social and the political; moreover, some totally deny the possibility of scientific progress. For many, Kuhn provides the license to remove science from its pedestal, to deny that science has any distinctive character or special claim on society's attention or support, and to demand that social sciences cease their attempts to be scientific. Close examination of a few of Kuhn's key ideas indicates how Kuhn came to be used in ways he, himself, disavowed.

The first problem is the ambiguity of the term "paradigm." Masterman (1970), a sympathetic critic, has pointed to twenty-one different senses of "paradigm" in the 1962 book, and Kuhn has added to the problem with later modifications of his ideas (1970, 1974). Laudan notes, "Since 1962 most of Kuhn's philosophical writings have been devoted to clearing up some of the ambiguities and confusions generated [by the 1962 book] . . . to such an extent that . . . [m]ore than one commentator has accused the later Kuhn of taking back much of what made his message interesting and provocative in the first place" (1984, pp. 67–68).

The metaphorical language in Kuhn's model poses a second serious problem. As Laudan puts it:

*Notoriously, he speaks of the acceptance of a new paradigm as a "conversion experience," conjuring up a picture of the scientific revolutionary as a born-again Christian, long on zeal and short on argument. At other times*

*he likens paradigm change to an "irreversible Gestalt-shift." . . . Such language does not encourage one to imagine that paradigm change is exactly the result of a careful and deliberate weighing-up of the respective strengths of rival contenders. (1984, p. 70)*

Although Laudan believes that problems of misinterpretation can be "rectified by cleaning up some of the vocabulary," and although Kuhn has assumed some responsibility for the misunderstandings due to his own rhetoric (1970, pp. 259–260), it should be noted that Kuhn's ambiguity and his vivid metaphors are important reasons for much of the enthusiastic embracing of his model. The possibility of reading into the meaning of "paradigm" and "paradigm shift" enables the model to serve a variety of agendas.

Kuhn's incommensurability thesis is both the most frequently criticized and the most frequently misused element of his model. It has provided ammunition for a radical subjectivism that denies all standards (Scheffler 1967). In attacking those who use Kuhn to justify any sweeping paradigm of their own, Crews writes, "By incommensurability Kuhn never meant that competing theories are incomparable but only that the choice between them cannot be *entirely* consigned to the verdict of theory-neutral rules and data" (1986, p. 39; emphasis added). Since Kuhn never clearly specified how rules and evidence entered into paradigm change and since he understated the degree to which different scientific communities can share rules and objectives, it is not surprising that this thesis became a rallying point for attacks on scientific rationality.

Another major criticism—and one that goes to the heart of Kuhn's model—involves the indivisible character of paradigms. Paradigm shift is an all-or-nothing process; a paradigm is accepted or rejected as a whole. Treating the parts of a paradigm as inseparable almost requires the transition from one paradigm to another to be a conversion experience; moreover, such a holistic view does not provide an accurate picture of how large-scale changes of scientific allegiance occur (Laudan 1984, pp. 71–72).

Many of these criticisms are widely held. Nevertheless, Kuhn's model remains important to many disciplines including sociology. Some scholars remain adherents to the basic features of the

model; others believe they must respond to the model's challenges to contemporary philosophy and methodology of science.

### USE OF "PARADIGM" IN SOCIOLOGY

While several leading British sociologists of science who focus on the sociology of knowledge reflect the intellectual mood of Kuhnian analysis (Collins 1983), the main applications of "paradigm" have occurred in either descriptive or normative efforts in the sociology of sociology.

Analysts have employed versions of Kuhn's concept to classify sociological activities as belonging to one or another paradigm; for example, Ritzer (1975) distinguishes three paradigms for the field at large: social facts, social definition, and social behavior. Bottomore (1975) identifies four paradigms of macrosociology: structural-functionalist, evolutionist, phenomenologist, and structuralist. Other sociologists discriminate varying numbers of paradigms from as few as two to as many as eight for a single subfield. Some writers classify theories into different paradigms; others categorize research strategies; still others codify more general philosophical orientations. In addition to disputing the number of paradigms, sociologists also disagree about whether paradigm applies to the field in general or only to subfields and whether there are any paradigms at all in sociology, that is, whether the field is preparadigmatic.

A few studies have applied Kuhn to examine historical development of a paradigm. Colclough and Horan (1983), for example, use content analysis of status attainment studies to illustrate Kuhn's model. Their analysis finds evidence for a stage of normal science and a stage in which anomalies arise, and they suggest the onset of a crisis for status attainment research.

Sociologists who use Kuhn focus on different properties of paradigms and debate the utility of their approaches or their faithfulness to Kuhn (cf. Ritzer 1975, 1981a, 1981b; Eckberg and Hill 1979; Hill and Eckberg 1981; Harvey 1982). Compounding the confusion are the many interpreters who read their own meanings into Kuhn. One critic argues:

*[F]or the most part the use of the term paradigm in sociology fails to reflect the*

*analytic elements of Kuhn's concept . . . Arbitrary pigeon-holing schemes of varying degrees of sophistication, are constructed and theories of, ideas about, and approaches to, sociology are dropped in. The result is a personalized schematic device that . . . plays little part in providing . . . any explanation or understanding of the growth of knowledge. The labeling of such pigeon-holes as paradigms legitimizes the scheme and implies an authority it does not possess. (Harvey 1982, p. 86)*

The disarray led Bell to ask "Does the term 'paradigm' carry too many possible meanings for rigorous thinking? Has it become quasi-mystical? Is it time to review its many usages and to consider discarding it in favor of more precise terms that convey to others more accurately what we mean to say?" (1990, p. 17) Bell went on to answer yes to all three questions.

### MODELS IN SOCIOLOGY

The first basic distinction necessary to characterize the use of models in sociology is among substantive, measurement, and statistical models. The substantive modeler's objective is to model a specific theory, phenomenon, or process in order to learn (or teach) something about the specific object of the model. For example, Pfeffer and Salancik (1978) use a theory of interpersonal power and dependence (Emerson 1962) to model interorganizational relations, or Gamson (1969) creates an elaborate game that simulates the operation of key features of a society. In both examples, the model focuses on a particular problem—to learn about organizational relations in the first, or to teach students about major aspects of society in the second.

The objective of a measurement model is to create and justify a measuring instrument that has a wide range of applicability. One measurement area to which sociologists were early contributors is attitude measurement; sociologists were among the first to develop models for scaling attitudes. These scaling models provide the rationale for assigning numbers to represent different expressions of the attitude. For example, Bogardus (1925), in a classic work, created a measure of social distance on the analogy of physical distance, in which allowing a member of some ethnic group

into your family represented less social distance than only allowing that person to live in your neighborhood. Guttman (1950) developed one of the most frequently used attitude scaling models, which defines and justifies procedures for ordering a set of attitude questions.

Statistical models are the basis for the systematic quantitative analysis of sociological data, and most empirical research in the field employs one or more statistical models. These models are general tools, broadly applicable to diverse problems; the nature of the subject matter that the data reflect is largely irrelevant to the model as long as the data have certain properties (e.g., as long as they are drawn randomly from some population). A large number of different statistical models appear in the sociological literature. Some are simple, for example, using the analogy of coin tossing to evaluate the likelihood that a given event occurred by chance; others are highly complex, using systems of equations to represent the structure of relationships in a set of variables. Models for multivariate analysis have become extremely important to sociology, including, for example, structural equation models, also known as path or causal models (Bielby and Hauser 1977), and log-linear models (Goodman 1984).

In recent years, models have become a central feature of the sociological literature. Researchers have constructed verbal, mathematical, and computer models to deal with a range of substantive phenomena and measurement issues. The application of statistical models to sociological data has increased dramatically in both the diversity of models employed and their mathematical sophistication. The most significant developments, however, have occurred in the creation, elaboration, and testing of formalized models of substantive phenomena, that is, models involving mathematics or logic.

In the early years of model building, substantive sociologists did not regard formal models as central to their concerns. Many mainline researchers were dubious about the utility of formal models in advancing knowledge about a substantive problem and often considered model construction as primarily an intellectual exercise, stimulating to the creator but of little relevance to anyone else. Articles presenting models appeared mainly

in esoteric journals read by small audiences of the initiated.

A striking feature of recent sociological work is the emergence of formal models as a central feature of several research traditions. Papers presenting formal models, extensions of these models, critiques of them, and empirical studies testing the models have appeared regularly in journals such as the *American Sociological Review*. While specialized journals such as the *Journal of Mathematical Sociology* still publish many articles dealing with substantive models, the fact that general sociological journals also publish reports dealing with formal models indicates that formal modeling has become much more integrated with substantive research in several problem areas of the discipline.

Recent research involving formal models has dealt with a diversity of substantive problems, employed a variety of mathematical tools, and used many different types of data to evaluate the empirical predictions of the models. Sociologists have constructed models for collective behavior (Heckathorn 1996; Kim and Bearman 1997), demographic processes (Yamaguchi and Ferguson 1995), exchange processes (Lawler and Yoon 1996; Yamaguchi 1996; Molm 1997; Burke 1997), geographic mobility (Herting et al. 1997), organization ecology (Gábor et al. 1994; Hannan et al. 1995) performance expectations and power-prestige orders (Troyer and Younts 1997; Berger et al. 1998), and social differentiation-inequality (Rickson and Palange 1994; Nielsen 1995; Orbell et al. 1996; Mark 1998). What has changed from the past is that highly developed models in each of these lines of research have become part of the core of theory building rather than merely interesting adjuncts.

### FUNCTIONS OF MODELS

A model, first of all, is an abstract representation of what the modeler regards as important aspects of the object being modeled. It involves assumptions, explicit or tacit, about that object, its elements, and the relationships among these elements. The definition given above requires that modeling lead to increased understanding; examining the relationships assumed in the model and ascertaining those that fit the object well, those

that fit the object somewhat, and those that do not fit at all serves this most important function. In other words, scrutinizing ways the model resembles the object and ways it does not can uncover previously unrecognized features of the object. Formal models, as noted earlier, employ logical or mathematical tools to derive conclusions from the model's assumptions, and researchers then test whether these conclusions represent hitherto unknown properties of the object. A model, in general, provides the means to ask new questions about the phenomenon it represents and thus functions to generate new ideas.

A model provides a vehicle for communication. Since a model is more abstract than its object and since it omits many properties of the object, it is easier for writer and audience to share understanding of the model than of the object. Caution, however, is necessary, particularly when the model is a vague analogy (e.g., society as a living organism). The possibility of reading into analogies and metaphors can defeat shared understanding when the audience focuses on aspects that are unintended. A model may display the complexities of the object. Attempting to represent the ties among members of a social group, for example, may require the elaboration of a simple model that represents only the presence or absence of a relationship to one that distinguishes types of relationships—attraction, task interdependence, relative status, and so on. Even if a model inadequately represents its object, it can enhance the understanding of that object by exposing issues that need to be addressed. One important class of models—those known as “baseline models”—are developed in order to study how the object deviates from the model's minimal representation. Investigating discrepancies between a baseline representation and the object can direct the construction of a more complex model (Cohen 1963).

The last general function to be noted is that a model can serve to relate what appear to be different phenomena. An abstract model can apply to a variety of different objects, thus calling attention to their common aspects; sometimes, particularly with metaphors and analogies, the features of one phenomenon serve as the model for another. Using drama as a model for everyday interaction calls attention to some common elements (e.g., what it means to play a role).

Berger and colleagues (1980) present a typology of models based upon more specific functions. Although their concern is mathematical models, their typology has more general application; they distinguish: (1) explicational models, (2) representational models, and (3) theoretical construct models. Explicational models are those for which the primary goal is to explicate or render precise one or more concepts. Many social network models are explicational models in that they provide an explication of the concept of social structure (Marsden and Lin 1982). While all models are representational, in this typology the term refers to those models that attempt to represent a particular observed social phenomenon. The majority of models developed in sociology are representational in this sense as long as one has some latitude in interpreting “observed,” in many cases, the object of the model is a generalization from observed phenomena. Coleman (1964, 1973) has formulated a number of representational models dealing with social change and with collective action.

Theoretical construct models are those which formalize an explanatory theory. In sociology, these models have focused on a variety of substantive topics as the following examples illustrate: Fararo and Skvoretz (1989) have devised a model drawing on social structural theories of Blau (1977) and Granovetter (1973); Hannan and Freeman (1989) have constructed models of organizational birth and death processes based on bioecological theories; Berger and colleagues (1977) have formulated a model of status characteristic theory (Berger et al. 1966, 1972).

Sociological models are models in a sense that is close to the everyday meaning of model as a replica or representation. It is even reasonable to think of many of them in the other everyday sense, as exemplars, since we can point to sociological research traditions devoted to their development and test. Only with the benefit of more hindsight than is presently available, however, will it be possible to judge whether any of these models are paradigmatic.

## REFERENCES

- Bell, Wendell 1990 “What Do We Mean by ‘Paradigm.’”  
*Footnote* 18:17.

- Berger, Joseph, Bernard P. Cohen, and Morris Zelditch, Jr., 1966 "Status Characteristics and Expectation States." In Joseph Berger, Morris Zelditch Jr., and Bo Anderson, eds., *Sociological Theories in Progress*, vol. 1. Newbury Park, Calif.: Sage.
- 1966 "Status Characteristics and Social Interaction." *American Sociological Review* 37:241–255.
- and J. Laurie Snell 1980 *Types of Formalization in Small-Group Research*. Westport, Conn.: Greenwood.
- Berger, Joseph, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch, Jr. 1977 *Status Characteristics and Social Interaction: An Expectation-States Approach*. New York: Elsevier.
- Berger, Joseph, Cecilia Ridgeway, M. Hamit Fisek, and Robert Z. Norman 1998 "The Legitimation of Power and Prestige Orders." *American Sociological Review* 63:379–405.
- Bielby, William T., and Robert M. Hauser 1977 "Structural Equation Models." *Annual Review of Sociology* 3:137–161.
- Blau, Peter 1977 *Inequality and Heterogeneity*. New York: Free Press.
- Bogardus, Emery S. 1925 "Measuring Social Distance." *Journal of Applied Sociology* 9:299–308.
- Bottomore, Tom 1975 "Competing Paradigms in Macrosociology." *Annual Review of Sociology* 1:191–202.
- Burke, Peter 1997 "An Identity Model for Network Exchange." *American Sociological Review* 62:134–150.
- Cohen, Bernard P. 1963 *Conflict and Conformity: A Probability Model and Its Application*. Cambridge, Mass.: MIT Press.
- Colclough, Glenna, and Patrick M. Horan 1983 "The Status Attainment Paradigm: An Application of the Kuhnian Perspective." *Sociological Quarterly* 24:25–42.
- Coleman, James 1964 *Models of Change and Response Uncertainty*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1973 *The Mathematics of Collective Actions*. Chicago: Aldine.
- Collins, H. M. 1983 "The Sociology of Scientific Knowledge: Studies of Contemporary Science." *Annual Review of Sociology* 9:265–285.
- Crews, Frederick 1986 "In the Big House of Theory." *New York Review of Books*. (May 29).
- Durkheim, Emile (1897) 1951 *Suicide*. New York: Free Press.
- Eckberg, Douglas L., and Lester Hill, Jr. 1979 "The Paradigm Concept and Sociology: A Critical Review." *American Sociological Review* 44:925–937.
- Emerson, Richard 1962 "Power-Dependence Relations." *American Sociological Review* 27:31–41.
- Fararo, Thomas J., and John Skvoretz 1989 "The Biased Net Theory of Social Structures and the Problem of Integration." In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress: New Formulations*. Newbury Park, Calif.: Sage.
- Péli, Gábor, Jeroen Bruggeman, Michael Masuch, and Breannán Ó. Nualláin 1994 "A Logical Approach to Formalizing Organization Ecology." *American Sociological Review* 59:571–593.
- Gamson, William A. 1969 *SIMSOC: Simulated Society*. New York: Free Press.
- Goodman, Leo A. 1984 *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.
- Granovetter, Mark S. 1973 "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–1380.
- Guttman, Louis 1950 "The Basis for Scalogram Analysis." In Samuel A. Stouffer, et al., eds., *Measurement and Prediction*. Princeton, N.J.: Princeton University Press.
- Hannan, Michael T., Glenn R. Carroll, Elizabeth A. Dundon, and John Charles Torres 1995 "Organizational Evolution in a Multinational Context: Entries of Automobile Manufacturers in Belgium, Britain, France and Italy." *American Sociological Review* 60:509–528.
- Hannan, Michael T., and John Freeman 1989 *Organizational Ecology*. Cambridge, Mass.: Harvard University Press.
- Harvey, Lee 1982 "The Use and Abuse of Kuhnian Paradigms in the Sociology of Knowledge." *Sociology* 16:85–101.
- Heckathorn, Douglas D. 1996 "The Dynamics and Dilemmas of Collective Action." *American Sociological Review* 61:250–277.
- Herting, Jerald R., David Grusky, and Stephen E. Van Rompaey 1997 "The Social Geography of Interstate Mobility and Persistence." *American Sociological Review* 62:267–287.
- Hill, Lester, Jr., and Douglas L. Eckberg 1981 "Clarifying Confusions about Paradigms: A Reply to Ritzer." *American Sociological Review* 46:248–252.
- Kaplan, Abraham 1964 *The Conduct of Inquiry*. San Francisco: Chandler.
- Kim, Hyojoung, and Peter S. Bearman 1997 "The Structure and Dynamics of Movement Participation." *American Sociological Review* 62:70–93.

- Kuhn, Thomas S. 1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- 1970 “Postscript.” *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- 1974 “Second Thoughts on Paradigms.” In Frederick Suppe, ed., *The Structure of Scientific Theories*. Urbana, Ill.: University of Illinois Press.
- Laudan, Larry 1977 *Progress and Its Problems: Towards a Theory of Scientific Growth*. Berkeley: University of California Press.
- 1984 *Science and Values: The Aims of Science and Their Role in Scientific Debate*. Berkeley: University of California Press.
- Lawler, Edward J., and Jeongkoo Yoon 1996 “Commitment in Exchange Relations: Test of a Theory of Relational Cohesion.” *American Sociological Review* 61:89–108.
- Mark, Noah 1998 “Beyond Individual Differences: Social Differentiation from First Principles.” *American Sociological Review* 63:309–330.
- Marsden, Peter V., and Nan Lin (eds.) 1982 *Social Structure and Network Analysis*. Beverly Hills, Calif.: Sage.
- Masterman, Margaret 1970 “The Nature of a Paradigm.” In Imre Lakatos and Alan Musgrave, eds., *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Merton, Robert K. 1949 *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Molm, Linda 1997 “Risk and Power Use: Constraints on the Use of Coercion in Exchange.” *American Sociological Review* 62:113–133.
- Nielsen, Francois 1995 “Meritocratic and Monopoly Inequality: A Computer Simulation of Income Distribution.” *Journal of Mathematical Sociology* 20:319–350.
- Orbell, John, Zeng Langche, and Matthew Mulford 1996 “Individual Experience and the Fragmentation of Societies.” *American Sociological Review* 61:1018–1032.
- Parsons, Talcott 1951 *The Social System*. Glencoe, Ill.: Free Press.
- Pfeffer, Jeffrey, and Gerald R. Salancik 1978 *The External Control of Organizations*. New York: Harper and Row.
- Rickson, Roy E., and Jean-Yves Parlange 1994 “Structural Differentiation and Size in Organizations: A Thermodynamic Formulation and Generalization.” *Journal of Mathematical Sociology* 19:69–90.
- Ritzer, George 1975 *Sociology: A Multiple Paradigm Science*. Boston: Allyn and Bacon.
- 1981a “Paradigm Analysis in Sociology: Clarifying the Issues.” *American Sociological Review* 46:245–248.
- 1981b *Toward an Integrated Sociological Paradigm: The Search for an Exemplar and an Image of the Subject Matter*. Boston: Allyn and Bacon.
- Scheffler, Israel 1967 *Science and Subjectivity*. Indianapolis, Ind.: Bobbs Merrill.
- Troyer, Lisa, and C. Wesley Younts 1997 “Whose Expectations Matter? The Relative Power of First- and Second-Order Expectations in Determining Social Influence.” *American Journal of Sociology* 103:692–732.
- Yamaguchi, Kazuo 1996 “Power in Networks of Substitutable and Complementary Exchange Relations: A Rational Choice Model and an Analysis of Power Centralization.” *American Sociological Review* 61:308–332.
- , and Linda R. Ferguson 1995 “The Stopping and Spacing of Childbirths and Their Birth-History Predictors: Rational Choice Theory and Event-History Analysis.” *American Sociological Review* 60:272–298.

BERNARD P. COHEN

## PARENTAL ROLES

In the opening years of the twenty-first century, most adults are or will become parents. The ages at which they start having children, as well as the number of children they have, differ significantly from earlier generations and from culture to culture, as do the social and economic conditions of parenthood. In this article, several major aspects of parenthood in the contemporary United States—and other industrialized nations—are discussed. First, several *demographic patterns* associated with parenting are reviewed. Second, the *rewards and costs* associated with parenting are examined. Third, *changes in the responsibilities of parents*, as defined by social perceptions of the nature of childhood, are discussed. In this section, special attention is given to gender differences in parenting styles. The fourth section examines *the impact of the first child's birth* on the parents. The paper closes with a discussion of *parent-child relations in middle and later life*.

### DEMOGRAPHIC TRENDS IN PARENTING

**Average Number of Births per Woman.** One of the most dramatic changes in the nature of parenthood in industrialized nations has been the *decline*

*in the average number of births to each woman.* In the United States, this number has decreased from seven births per woman to two, over the past 200 years. This downward trend has not been steady or consistent, however. During some periods, such as the Great Depression of the 1930s, the rate of decline was much more pronounced, while at other times, most notably during the post-World War II baby boom (1946–1962), the number of births per woman actually increased. In 1936, the middle of the Depression, American women were giving birth to two children, on the average. At the peak of the baby boom, in 1957, the number reached 3.6 births per woman. For nearly twenty years following this peak, the birth rate dropped dramatically. Since the mid-1970s, the birth rate has been fairly stable. There are, however, racial and socioeconomic differences in fertility. In the United States, white women have the fewest births, followed by black and Hispanic women. White women aged 40 to 44 have given birth to an average of 1.9 children, while African-American and Hispanic women have had 2.1 and 2.6 children, respectively (U.S. Bureau of the Census 1999).

*Reasons for the Declining Birth Rate.* Why has the average number of births per woman decreased over the past 200 years? One major reason is that in the past women had little control over their childbearing. Through the eighteenth and nineteenth centuries, married women gave birth to as many children as they were biologically capable of having. This pattern persisted into the twentieth century for many women, although by the end of the nineteenth century many middle-class women were beginning to use birth control. Childbearing in the United States is now largely controlled through the use of various contraceptive methods and, to a lesser extent, through medically induced abortion. Most people have the number of children they want to have, and generally, they want no more than two children. Only a small percentage of couples want, and actually have, more than two children (Shehan and Kammeyer 1997, pp. 183–184).

A second reason for the historical decline in childbearing is that children are no longer economic assets for their families. Through much of the nineteenth century, the labor of children helped families survive. Most Americans lived on farms, where children could help out in many ways. Even

the children who lived in small towns and cities often worked at young ages to help support their families. When laws pertaining to child labor and mandatory education were passed in the late nineteenth and early twentieth centuries, children became less valuable to their parents in economic terms. Today, children require large economic expenditures from their parents, as discussed later in this article. Most people have the number of children they think can afford, and thus, the number they have is much lower than in the past (Shehan and Kammeyer 1997, p. 184).

A third reason for today's low fertility rate is the high number of women who are in the paid labor force. Economic need has led to a dramatic increase in the employment of mothers, married as well as divorced. In 1960, fewer than one in every five mothers of preschool children was employed. Today, over 60 percent are employed. The increased demand for substitute child care has become one of the most pressing social problems in industrial nations. Child-rearing duties decrease the amount of time and energy women can devote to their jobs; therefore, those who are career oriented may choose to forego childbearing altogether or to limit the number of children they have (Shehan and Kammeyer 1997).

#### **Changes in the Characteristics of Parents.**

Social change in the twentieth century has not only reduced the average number of children born per woman in industrialized nations but has also seen a change in the marital status and age at which women begin to have children. The childbearing period seems to be expanding, as some women continue to become mothers in their teen years while others wait until they are in their late twenties, thirties, or early forties. Developments in reproductive biology have even made it possible for women who have gone through menopause (which typically occurs in the early fifties) to have babies. Teenage childbearing was fairly common in the post-World War II baby boom, but most of the teenagers who had babies during that era were married. Teen birth rates fell from the mid-1950s to the mid-1980s then increased in the early 1990s. Today, the rate of teen childbearing is lower than it was at the beginning of the 1990s, but most teen births (71 percent) occur outside of marriage (Shehan and Kammeyer 1997, p. 186).

Another significant change in fertility is the increase in births to unmarried women. Today, American women spend fewer years in marriage because they are marrying at older ages and are more likely to get divorced. As a result, there is less time available to them to have children within the context of marriage. One-third of all births in the United States today are to unwed mothers. The rate is higher among blacks than it is among whites. Childbearing outside marriage actually varies widely among industrialized nations, from Japan, where only 1 percent of all babies are born to unmarried women, to Sweden, where the comparable figure is 53 percent.

As a consequence of the high out-of-wedlock birth rate and the high divorce rate, an increasing number of adults are engaging in solo parenting. In 1995, about twelve million American families—or about one-fourth of all families with children—had only one parent (U.S. Bureau of the Census 1995). While the majority of single parents are women, men are also raising children alone. Today, about one-sixth of all single-parent homes are headed by men.

#### **THE CHANGING SOCIAL CIRCUMSTANCES OF PARENTING**

The demographic shifts that have occurred in industrial nations since World War II—particularly the increases in divorce and out-of-wedlock births—have increased the likelihood that children will have little or no contact with one of their biological parents during some period of their childhood. Only one in every four children of divorced parents who live with their mothers sees his or her father once a week; and close to half have little or no personal contact of any type with their fathers after divorce. Men's contact with their biological children from previous marriages is reduced when they remarry and become involved in the lives of stepchildren or have children with their new partner. Other adults, however, may step in to fill the gap in children's access to their biological parents. For instance, three-fourths of the children who live with their fathers share their homes with other adults. About one-third live in multigenerational households (with grandparents and other relatives), and another third

live with their father and his intimate partner. Children in these homes may or may not benefit from the added financial and emotional resources that nonrelated adults can offer (Shehan and Seccombe 1996).

Single-parent families headed by women often have financial problems. Many live below the poverty level. A major contributing factor to this "feminization of poverty" is the fact that not all noncustodial fathers are ordered by the courts to pay child support after divorce, and, of those who are, fewer than half actually pay the full amount.

#### **GOVERNMENT SUPPORT FOR PARENTS**

The major changes in family patterns that have occurred over the past twenty-five years have made government leaders from around the world very aware of family problems. Concerns about declining birthrates and the aging of the population have been widespread throughout industrial nations. Many governments have adopted "pro-family" attitudes and attempted to create more supportive environments for families. But industrial nations have differed significantly in the extent to which they have provided support for families. The U.S. government has been reluctant to assume public responsibility for care of children, instead passing laws that reinforce the private nature of child support obligations. According to one indicator of "family friendliness," the United States, along with Portugal, Italy, Ireland, Spain, and Greece, is among the least generous of the industrialized nations in terms of its benefits packages for families with children (Gauthier 1999, p. 7).

As women's labor-force participation increased in industrialized nations in the 1970s, renewed attention was given to the provision of maternity leave plans. In most countries, these plans were upgraded, at least in terms of their duration, increasing by six weeks, on average across eighteen industrialized nations. Major increases in maternity pay also occurred in most countries. But in the United States, where only thirty states had provisions for maternity leave, no pay was mandated. In 1992, the Family and Medical Leave Act was passed by the U.S. Congress and signed into law by President Bill Clinton. The legislation provides public and private employees in companies with fifty or



more workers the right to take up to twelve weeks leave each year to care for a newborn or newly adopted child; or for a seriously ill child, spouse, or parent; or for their own serious health conditions (Gauthier 1999).

### THE CHILDBEARING DECISION: THE REWARDS AND COSTS OF PARENTING

Since individuals are now better able to control their reproductive functions due to advances in contraceptive technology and access to legal abortion, it is important to examine why the majority of adults still choose to become parents. For most people, becoming a parent involves a decision-making process in which the anticipated rewards and costs of parenting are weighed. One major influence on this choice is societal pressure. While most societies encourage fertility among their members, the United States is considered to be particularly *pronatalist*, which means its dominant values and attitudes promote and encourage childbearing (Schoen 1997).

**Rewards of Parenting.** In earlier historical periods, children were valued in large part because their labor could greatly enhance a family's economic survival. Today, children may continue to have economic value for parents who view them as a type of old age insurance. A small minority of parents, typically those from rural areas, also see children's ability to participate in household labor as a valuable asset. Having children may also enable a family to maintain control over property or a family-owned or family-operated farm or business. For most families, though, children may be more of an economic liability than a benefit. Their primary value to their parents is emotional and symbolic. Understanding the noneconomic benefits of child rearing helps in understanding why most people choose to become parents when they are no longer forced to do so by biological necessity.

Studies of the value of children to their parents have identified many major rewards that child rearing provides. One reward is *family continuity or personal immortality*. The birth of children ensures that the father's family name will continue into the future, at least for one more generation, and may give the parents a sense of immortality, a feeling that part of them will survive after death.

Becoming a parent can also lead to a *change in social identity*. It bestows adult status, which carries with it implications of maturity and stability. In fact, for many people, parenthood is *the* event that gives them a sense of feeling like an adult. Having children may also produce a sense of achievement, not only for the physical fact of conception but also for the challenges of raising a child. Moreover, many parents also feel a sense of accomplishment through their children's achievements. Finally, during the child-rearing years, the legitimate authority that is attached to the parent role, as well as the resulting power and influence parents have over their children's lives, may increase parents' self-esteem, especially for those who have little control over other aspects of their lives.

The most frequently mentioned reward of child rearing involves primary group ties and affection. Children help their parents establish new relationships with their own parents, grandparents, aunts and uncles, siblings, and friends; such relationships can provide emotional support and practical assistance. Children are also expected to help prevent loneliness and to provide love and companionship for their parents. Many people regard childbearing as a sacred duty and may feel that by becoming parents they are fulfilling a divine commandment. Moreover, the physical and symbolic sacrifices involved in child rearing are perceived by some as a sign that parents are more virtuous and altruistic than childless adults (Shehan and Kammeyer 1997, pp. 187-189).

**The Costs of Parenting.** There are also numerous costs associated with parenting. These costs are economic as well as social and emotional. Over the years, economists have estimated the *direct economic costs* (e.g., food, clothing, shelter, medical and dental care, toys, recreation, leisure, and education) and the *indirect economic costs* (e.g., foregone savings and investments, lower standards of living, and loss of potential income for parents who leave the labor market to care for the child) of child rearing. The direct costs of raising a child to age 18 have risen by more than 20 percent in recent decades, after adjustments for inflation and for changes in family size have been made. In 1998, raising a child to adulthood, which includes sending her or him to a state university for four years, costs middle-income parents an average of

\$459,014; that's \$301,183 in direct costs and \$157,831 in costs associated with college education (Longman 1998).

Another important economic cost of parenthood that must be considered is the current or future family income that is given up when parents leave the labor force—or reduce their hours in employment—to care for children. Virtually all parents face some opportunity costs in having children. Using estimates and procedures developed by the U.S. Department of Agriculture—which provides reports on the costs of child rearing annually—the following example of the opportunity costs encountered by parents is offered. A married woman with an annual salary of \$23,600 who stays home to care for a child from birth until kindergarten age (at which point she reenters the labor force on a half-time basis) can expect to sacrifice nearly \$1 million in lost income by the time that child reaches age 21 (Longman 1998). When opportunity costs are added to the direct costs of child rearing that were listed above, a middle-income couple can expect to spend slightly more than \$1.45 million on one child by the time she or he completes college.

In the United States and other industrialized nations, however, parents receive some governmental subsidies for raising children. In the United States in 1998, each dependent child was worth a \$2,650 deduction on his or her parents' federal income taxes. Parents are also able to take a tax credit for child care expenses—a maximum of \$720 for one child in a low-income family to a maximum of \$1,440 for two or more children (Longman 1998).

The presence of children may also be costly to adults through the restriction of their activities and the resulting loss of freedom. Parents are responsible for the mental, emotional, physical, spiritual, and social development of their offspring. Obviously, such responsibility may consume much of the parents' time and attention and may require a readjustment of lifestyle to take the children's needs and activities into account. Consequently, parenting may have negative effects on marriage. Studies have shown that the birth of children may hurt a couple's affectional and sexual relationship, due largely to frequent interruptions, a loss of privacy, and increased demands on time and energy. The birth of children often results in role

segregation, which means that spouses engage in fewer joint activities as they attempt to fulfill their parental responsibilities (Walzer 1998).

**When the Costs Outweigh the Rewards.** For a small minority of adults in industrialized nations, the anticipated costs of having children outweigh the anticipated rewards, and they decide to have no children. Throughout most of the twentieth century, the proportion of childless marriages in the United States was only 5 to 10 percent. The rates of childlessness increased in the last decades of the century, however. In the late 1990s, nearly one in every five American women aged 40 to 44 (considered the final years of the childbearing period in women) was childless (U.S. Bureau of the Census 1995, 1999).

The individuals with the highest education in our society are the ones least likely to have children, because they experience the highest opportunity costs associated with childbearing and child rearing. In the United States, middle-aged women with graduate degrees are three times more likely than those who drop out of high school to be childless. Two-income couples who earn more than \$75,000 are much more likely than those who make less than \$20,000 to be childless (Longman 1998).

Adults who choose to remain childless may believe that having children would interfere with their ability to achieve in their careers or in other types of public service. They may feel that caring for children would drain away the time and energy that could be devoted to other highly valued pursuits. Another important consideration for such adults is the expectation that the commitment to parenthood would reduce the time available to devote to their marriage. To voluntarily childless couples, marriage rather than children may be the primary source of affection and sense of belonging.

#### THE CHANGING RESPONSIBILITIES OF PARENTS

Views of children and childhood, and of parental responsibilities, have changed over the centuries. For most of human history, simple physical survival of children was the dominant issue in child rearing. Prior to the seventeenth century, parental

love was rarely identified by child-rearing experts as a critical factor in the development of a child. In medieval societies, there was no recognition of any characteristics that distinguished children from adults; children merged naturally into adult society from about the age of 7. There were no boundaries separating the adult world from the world of children (Cunningham 1995). It wasn't until the sixteenth century that childhood was viewed as a period of innocence. Children were idolized and valued as a source of amusement or escape for adults. Later in the sixteenth century, and carrying over into the seventeenth and eighteenth centuries, childhood began to be seen as a time of immaturity and children were believed to need discipline and guidance as they prepared for adulthood. Parents were expected to provide physical care, consistent discipline, and a model for proper behavior.

At the close of the nineteenth century, during the Victorian age, awareness of the vulnerability of children and a belief in the sacredness of childhood emerged. Compulsory schooling, which removed children from the labor force, was introduced and drew a sharp line between childhood and adulthood. Once children were regarded as different from—and inferior to—adults, they became dependent on older persons for survival (Shehan 1999). In some segments of American society today, there is a belief that parents alone can—and should—meet all their children's needs. Parents are responsible not only for their children's physical well-being but also for their psychological adjustment. Often, flaws in a person's adjustment are traced to a lack of parental love or some other parental shortcoming.

**The Mother Role.** To many people, parenting is synonymous with mothering, and mothering is believed by many to be an instinct found in all women. While scientists have yet to find an instinctual motive for motherhood among humans, they have demonstrated a strong learned need among women to have children. Most women, given the choice to become mothers, would choose to do so, and most women who are already mothers would choose the role again (Genevie and Margolies 1987).

The belief that motherhood is necessary for women's fulfillment and for the normal healthy development of children has waxed and waned

throughout our history, largely in response to economic conditions. When women's labor is not needed outside the home, the mother role is glorified and exalted; when women's labor is essential to the economy, the importance of the mother-child bond is downplayed (Margolis 1984). Earlier in our history, when our economy was agrarian, parenting was more of a joint venture. Child rearing was shared among a larger number of adults and the mother-child bond was not regarded as primary. Only after industrialization and urbanization changed the nature of work and family life did the role of mother in child development become preeminent. As will be discussed below, similar changes are under way in regard to the father role.

**The Father Role.** While the verb *to mother* is used to refer to the nurturance and care given to children, usually by women, the verb *to father* has a much more restricted meaning. To many, this simply refers to the male role in procreation. The responsibilities attached to the father role have traditionally been economic. To be a good father, a man had to be a good provider. Participation in the daily custodial care of the child was not expected, nor was companionship or nurturance. In recent decades, with the entrance of large numbers of mothers into the labor force, the expectations attached to the father role have begun to change. Men can no longer be good fathers simply by being good providers. They must also participate more fully in the daily care of their children and in the socialization process (Griswold 1993).

It appears that these expectations are being reflected in changed behavior. More fathers are attending childbirth education classes with their wives and are present at the births of their children, and the average amount of time men spend with their children has increased since the early 1970s. These changes support the argument that men are psychologically capable of participating in all parenting behaviors. Perhaps the most telling evidence of the extent to which Americans' ideas about the father role have changed since the 1940s can be found in the expert advice on parenting. In the first edition of his classic book about child care, which was published in 1945, Dr. Benjamin Spock reminded fathers that they need not be as involved in child care as mothers, at most preparing a bottle for the baby on Sundays. By the

1980s, Dr. Spock was admonishing fathers to take on half of all child care and housework tasks.

### **TRANSITION TO PARENTHOOD**

Some sociological studies have found that many couples experience the birth of their first child as a crisis. The changes brought about by the addition of a baby to a household are indeed extensive. Occupational commitments, particularly of the mother, may be reduced, and family economics must be reorganized as spending increases and earnings are reduced. Household space must be reallocated to accommodate the infant's lifestyle. The parents' time and attention must be redirected toward the infant. Relationships with kin, neighbors, and friends must be redefined to include the baby's schedule. The marital relationship itself may be disrupted due to the enormous demands for time, energy, and attention made by the baby. As a result, new mothers frequently report unexpected fatigue, confinement to the home and a sharp reduction in social contacts, and the loss of the satisfaction that accompanied outside employment. New fathers may feel added economic pressure (Walzer 1998).

The severity of the crisis experienced by new parents does not seem to be related to the quality of the marital relationship before the birth of the child or to the degree to which the child was planned and wanted. Instead, it may be the degree to which the parents had romanticized parenthood in conjunction with their lack of preparation for the role that leads to a feeling of crisis. As a result of the tremendous changes brought about by the presence of children and the burdens associated with child rearing, marital satisfaction appears to decline sharply around the time of the first child's birth and to remain low until children leave the home (Shehan and Kammeyer 1997, pp. 213-214).

### **PARENTING IN THE MIDDLE YEARS**

During the twentieth century, there was a steady increase in the number of young adults who returned to their parents' homes to live. When children return home to live with their parents, there are both advantages and disadvantages. Many of

the problems are caused by the ambiguity of the situation. Parents may wonder whether they should establish and enforce rules governing their children's behavior (e.g., curfews, financial obligations, use of drugs and alcohol, and guests). Both parents and their young adult children are likely to revert to the roles they played when the children were adolescents. Young adults may revert to being emotionally, physically, and financially dependent on their parents but may find it stressful being "adult teenagers." Parents, too, are likely to feel stress when their young adult children return home to live. They are often forced to make major changes in their daily routines, and to take on unwanted burdens such as extra cooking, cleaning, and laundry when children return home. There can also be financial strains for the parents, associated with larger food bills and greater utilities costs. Conflicts between parents and their adult children may arise simply because of crowded household conditions. When adult children bring their furniture, sports equipment, pets, and clothing back home, the "empty nest" can become a "cluttered nest." But in spite of the problems that can arise when young adults live with their parents, both groups typically feel that the situation has many benefits (Shehan and Kammeyer 1997, pp. 260-265).

### **PARENTING IN THE LATER YEARS OF LIFE**

The rewards of parenting can persist throughout life. Continued attachment to their children helps to minimize elderly parents' sense of isolation and loneliness. Children's readiness to take care of parents' needs helps to build a sense of security in old age and is important in day-to-day survival. However, elderly parents also want to continue to help their children. Their contributions to their adult children occur in many forms, from providing financial assistance to help with housework and care of grandchildren. Elderly persons derive a great sense of pride and satisfaction from their parental role, even when their sons and daughters are middle-aged adults with children of their own. Being recognized by their children for their value and competence as parents helps them to maintain high levels of self-esteem. When adult children provide aid and assistance to their elderly

parents—especially if they take a condescending attitude toward their parents and make all decisions for them—the latter may find it demoralizing. The ideal situation is one where elderly parents and their adult children help each other and make decisions together. All things considered, the relationships between elderly parents and their adult children are very positive (Shehan and Kammeyer 1997, pp. 273–275).

(SEE ALSO: *American Families*; *Family Roles*; *Fertility*; *Socialization*)

## REFERENCES

- Cunningham, Hugh 1995 *Children and Childhood in Western Society since 1500*. London: Longman.
- Gauthier, Anne Helene 1999 *The State and the Family: A Comparative Analysis of Family Policies in Industrialized Countries*. Oxford: Clarendon Press.
- Genevie, Lou, and Eva Margolies 1987 *The Motherhood Report: How Women Feel about Being Mothers*. New York: Macmillan.
- Griswold, Robert L. 1993 *Fatherhood in America: A History*. New York: Basic Books.
- Longman, Phillip J. 1998 "The Cost of Children." *U.S. News and World Report* 124 ( March 30):50–57.
- Margolis, Maxine L. 1984 *Mothers and Such: Views of American Women and How They've Changed*. Berkeley: University of California Press.
- Schoen, Robert 1997 "Why Do Americans Want Children?" *Population and Development Review* 23(2):333–359.
- Shehan, Constance L. 1999 "No Longer Place for Innocence: The Re-Submergence of Childhood in Post-industrial Societies." In Constance Shehan, ed., *Through the Eyes of the Child: Re-Visioning Children as Active Agents of Family Life*, vol. 1 in *Contemporary Perspectives on Family Research*. Stanford, Conn.: JAI.
- , and Kenneth C. W. Kammeyer 1997 *Marriages and Families: Reflections of a Gendered Society*. Needham Heights, Mass.: Allyn and Bacon.
- Shehan, Constance L., and Karen Secombe 1996 "The Changing Social Circumstances of Children's Lives." *Journal of Family Issues* 17(4):435–440.
- U.S. Bureau of the Census 1995 *Fertility of American Women: June 1995*, P-20-499.
- 1999 *Census Bureau Facts for Features* (CB99-FF.07), Mother's Day 1999, released on April 29.
- Walzer, Susan 1998 *Thinking about the Baby: Gender and Transitions into Parenthood*. Philadelphia: Temple University Press.

CONSTANCE L. SHEHAN

## PARTICIPANT OBSERVATION

See *Ethnography*; *Qualitative Methods*; *Sociocultural Anthropology*.

## PARTICIPATORY RESEARCH

Participatory research integrates scientific investigation with education and political action. Researchers work with members of a community to understand and resolve community problems, to empower community members, and to democratize research. The methods of participatory research include group discussions of personal experience, interviews, surveys, and analysis of public documents. Topics that have been investigated with this approach include community issues such as polluted water supplies and the school curriculum, employment issues such as working conditions and unionization, and theoretical issues about consent and resistance to domination. For social scientists who question the traditions of being detached and value-free, and who seek an approach that is less hierarchical and that serves the interests of those with little power, participatory research is a valuable alternative.

Participatory research can be identified by five characteristics: (1) participation by the people being studied; (2) inclusion of popular knowledge; (3) a focus on power and empowerment; (4) consciousness raising and education of the participants; and (5) political action. A precise definition should be avoided so that each group that does participatory research can be free to develop some of its own methods.

Participation in the research process by the people being studied is best viewed as a continuum that includes low levels of participation, such as asking people who are interviewed to read and comment on the transcripts of their interviews, as well as high levels of participation. Ideally, community members have a significant degree of participation *and control*, and help to determine the major questions and overall design of the study.

Second, participatory research validates popular knowledge, personal experience and feelings, and artistic and spiritual expressions as useful ways of knowing. If researchers are to work with community members as co-investigators, they must respect people's knowledge. Moreover, one of the rationales for community participation in research is the assumption that people understand many aspects of their situation better than outsiders do. Practitioners have used group discussions, photography, theater, and traditional tales to draw on popular knowledge (Barndt 1980; Luttrell 1988).

A focus on power and empowerment also distinguishes most participatory research. "The core issue in participatory research is power. . . the transformation of power structures and relationships as well as the empowerment of oppressed people," states Patricia Maguire in her excellent analysis of the field (1987, p. 37). Participatory researchers differ widely in their positions on empowerment, and include radicals who try to transform the power structure by mobilizing peasants to wrest land from the ruling class, as well as conservatives who ignore power relations and focus on limited improvements such as building a clinic or a collective irrigation system.

The fourth characteristic of participatory research—consciousness raising and education—is closely related to power. Group discussions and projects typically attempt to reduce participants' feelings of self-blame and incompetence, and try to relate personal problems to unequal distributions of power in the community and the society. Participants often become visibly more confident and effective as they speak out in discussions, learn that others share some of their experiences, and learn research skills and relevant technical information.

Finally, participatory research includes political action, especially action that cultivates "critical consciousness" and is oriented toward structural change, not toward adjusting people to oppressive environments (Brown and Tandon 1983). Some scholars argue that "real" participatory research must include actions that radically reduce inequality and produce "social transformation." Research and action, from this perspective, should be guided by a general theory like Marxism to help identify the underlying causes of inequality and the best strategies for changing society. Others

caution against expecting to achieve radical changes because "social transformation requires . . . organizing, mobilizing, struggle" as well as knowledge (Tandon 1988, p. 12). These researchers point to the value of small collective actions in educating people about the local power structure, creating greater solidarity and feelings of power, and providing new knowledge about how power is maintained and challenged. Many projects include little or no collective action and are limited to changing the behavior of individual participants, strengthening or "creating a community network" and "fostering critical knowledge" (Park 1978, p. 20).

In some cases, participatory research produces major changes, as exemplified by a project with residents of a small town in the state of Washington. The town was going to be destroyed by the expansion of a dam, and the U.S. Army Corps of Engineers was planning to disperse the community. But with the assistance of Professor Russell Fox and numerous undergraduates from Evergreen State College, residents clarified their own goals for a new community, learned about the planning process, and produced a town-sponsored plan for a new town. Their plan was accepted by the Corps of Engineers after prolonged struggle involving the courts and the U.S. Congress. The new town thrived and continued to involve the entire community in planning decisions (Comstock and Fox 1982).

A study of the working conditions of bus drivers in Leeds, England, illustrates the mixed results that are more typical of participatory research. As a result of greater pressure at work accompanying Prime Minister Margaret Thatcher's program of deregulation, bus drivers were experiencing increasing stress, accidents, and conflicts at home (Forrester and Ward 1989). With the help of professors from the University of Leeds who were running an adult education program for workers, a group of eight bus drivers, selected by their local union, decided to do research that would investigate stress at work and motivate the drivers' union to take action. They designed and carried out a survey of a sample of drivers and their families, studied accident records, and measured physical signs of stress. Although the report presenting their findings failed to produce the desired action by the union, the project was successful in many other ways—workers' stress became part of the

agenda for the union and the national government, and the report was used by workers in other countries to document the need for improved working conditions. The participants in the research gained research skills and knowledge about work stress, and the professors produced academic papers on work stress and participatory research. The professors had a dual accountability (as they put it) both to the bus workers and to the university; their projects produced results that were valuable to both groups.

### THE HISTORY OF THE FIELD AND RELATIONS TO OTHER FIELDS

Participatory research was developed primarily by Third World researchers, and most projects have been in Third World communities. In the 1970s it became clear that mainstream economic development projects were failing to reduce poverty and inequality. In response, researchers began to develop alternative approaches that increased the participation of the poor in development programs and aimed at empowering poor rural and urban communities as well as improving their standard of living (Huizer 1979; Tandon 1981, 1988). For example, in the Jipemoyo Project in Tanzania, researchers and villagers investigated traditional music and dance practices and developed cooperative, small-scale industries based on these traditions, such as the production of "selo drums" for sale in urban areas (Kassam and Mustafa 1982). In other projects, peasants and farmers participated with agricultural and social scientists to determine the most appropriate and productive farming methods. Several projects in Latin America, led by Orlando Fals Borda and labeled "participatory action research," integrated the knowledge of peasant activists and academics to build rural organizations and social movements.

Participatory researchers in the Third World are closely associated with Paulo Freire, an exiled Brazilian educator with roots in Marxism and critical theory. His book *Pedagogy of the Oppressed* is the most influential work in participatory research. Freire argues that teaching and research should not be dominated by experts but should be based on dialogue with a community of oppressed people. Through dialogue and collective action, people can develop critical consciousness, learn the

skills they need to improve their situation, and liberate themselves. A similar approach has been developed by the influential Highlander Research and Education Center in the southern United States. Organized by Myles Horton and others in the 1930s, Highlander has inspired many participatory researchers with its success in educating and empowering poor rural people (Gaventa and Horton 1981; Gaventa et al. 1990). Another important center of participatory research has been the Participatory Research Network in Toronto, focusing on adult education (Hall 1975, 1981).

The development of participatory research in the 1970s was also fostered by challenges to positivist social science by feminists, Marxists, critical theorists, and others (Bernstein 1983; Harding 1986). The critics emphasized the links between knowledge and power. They argued that the positivists' emphasis on objectivity, detachment, and valuefree inquiry often masked a hidden conservative political agenda, and encouraged research that justified domination by experts and elites and devalued oppressed people. The critics proposed alternative paradigms that integrated research and theory with political action, and gave the people being studied more power over the research (Carr and Kemmis 1986; Lather 1986; Rose 1983).

The development of alternative paradigms, together with the emergence of participatory research in the Third World and the political activism accompanying social movements of the 1960s and 1970s, sparked a variety of participatory research projects by North American social scientists. John Gaventa investigated political and economic oppression in Appalachian communities, and grass-roots efforts to challenge the status quo, and Peter Park criticized mainstream sociology from the perspective of participatory research and critical theory. Health-related issues such as wife-battering, health collectives, and toxic wastes were studied by Patricia Maguire and others, while researchers in education examined community efforts to improve public schools and participatory methods of teaching (Luttrell 1988). Issues at the workplace such as struggles for unionization have been investigated by many participatory researchers; they have documented the impact of factors such as ethnic divisions and women's work culture on the success of unionization (Bookman and Morgen 1988).

Participatory research is closely related to several other fields. Feminist approaches to research and teaching often closely resemble participatory research and emphasize nonhierarchical relations between researcher and researched, raising consciousness, taking action against sexism and other forms of domination, and valuing expressive forms of knowledge (see Smith 1987; Stanley and Wise 1983). Feminists have done the majority of the participatory research projects in North America, but they do not use the label, and feminists and participatory researchers rarely consult each other's work.

A similar approach has been developed by William F. Whyte, who works with representatives of management and workers to study organizational problems such as reducing production costs, or redesigning training programs. His approach differs from participatory research in that it gives little attention to power and empowerment, or to consciousness raising and education, and the action component of the projects is coordinated with management and does not directly challenge the existing power structure. Whyte labels his approach "participatory action research," which will cause confusion since the same term is used to describe Orlando Fals Borda's very different, more radical approach.

Participatory research also overlaps with several traditional social science methods, especially participant observation, ethnography, and intensive interviews, all of which rely on empathic interpretation of popular knowledge and everyday experience and that lead researchers to be engaged with the people being studied, not detached from them. Applied research also focuses on social action, but usually for the privileged—those with the money and sophistication to employ researchers or consultants.

### ISSUES IN PARTICIPATORY RESEARCH

Two issues underlie many of the problems confronted by academic participatory researchers: the relations between researchers and the researched; and the tensions between being politically active and producing objective, academic studies.

Relations between researchers and researched are problematic at each stage of a participatory

research project (on stages, see Maguire 1987; Vio Grossi 1981; Hall 1981). At the beginning of the project the researchers must consider what segments of the community will participate. Although some participatory researchers talk about "the oppressed" people in a community or "the poor" as if they were a homogeneous group, most communities are complex and internally stratified, and the more powerless people usually are more difficult to include.

The power of the researcher versus the researched also is problematic in the next stage of a project, when participants identify and discuss community problems. Researchers have specific skills in facilitating the group and obtaining information, and typically have more time, money, and other resources. Therefore, they can take more responsibility (and power) in the project, and community members often want a researcher to take charge in some areas. There are also conflicts during group discussions between validating participants' knowledge and power versus educating for critical consciousness and validating the researchers' power (Vio Grossi 1981).

When the project moves to the stage of designing research on community problems, researchers are especially likely to have a power advantage, since community members typically lack the skills and the interest to carry out this task. If community members are to be equal participants in designing a complex research project, they first need an extended educational program like the adult education classes for the Leeds bus drivers. Otherwise, the research probably will have to be fairly limited, or the researchers will control the research design, while community members participate as consultants and trained research assistants. In this case it becomes especially important that community members have substantial power in setting the research agenda (e.g., Merrifield 1989).

Conflicts between activism and involvement versus academic objectivity and detachment are another source of problems. However, many of the problems can be resolved by questioning the assumed incompatibility between being involved with the people one is studying and producing objective or valid evidence. On the one hand, involved researchers often produce valid knowledge. Sociology and anthropology include many



examples of systematic, highly regarded ethnographies and interview studies by researchers who were very involved with the community. Moreover, participating as an activist probably yields just as valid an account as being a traditional participant observer. On the other hand, research methods associated with being detached, such as surveys and quantitative analysis, often contribute to effective political action. For example, a research group in Bombay organized a participatory census of pavement dwellers in a large slum. Their results documented that slum dwellers had been underenumerated by the official census and had been unjustly denied census-dependent services. Participants also created strong community organizations and learned how to use existing services (Patel 1988). In this project, community involvement and academic standards were compatible. In other projects, participatory researchers have experienced many conflicts between serving the interests of the community being researched, and producing knowledge that is valuable to the academic community.

#### RECENT DEVELOPMENTS

Participatory research methods have gradually spread through the social sciences and related fields in recent years. Researchers in health (Cornwall and Jewkes 1995; Mishra et al. 1998), family sociology (Small 1995), community psychology (Stoecker and Beckwith 1992; Yeich and Levine 1992), and other fields are using participatory research methods. Several journals have devoted special issues to participatory research, including *American Sociologist* (Stoecker and Bonacich 1992, 1993), *The Journal of Social Issues* (Bryden-Miller 1997) and *Human Relations* (Chisholm and Elden 1993). New books on participatory research also have appeared (De Koning and Martin 1996; Fals-Borda and Roshman 1991; Park et al. 1993).

The growing numbers of studies using participatory research methods vary widely in the degree to which they depart from traditional social science methods. In most studies, control over the research design remains with the researcher, although the researched may have power over a limited component of the project. In addition, many studies do not include social action as part of the project. As participatory research continues to

develop, researchers will continue to struggle with balancing the power of the researcher and the researched, and with the conflicting demands of activism and academic standards.

(SEE ALSO: *Field Research Methods*; *Qualitative Methods*; *Social Movements*)

#### REFERENCES

- Acker, Joan, Kate Berry, and Joke Esseveld 1983 "Objectivity and Truth: Problems in Doing Feminist Research." *Women's Studies International Forum* 6:423-435.
- Barndt, Deborah 1980 *Education and Social Change: A Photographic Study of Peru*. Dubuque, Iowa: Kendal/ Hunt.
- Bernstein, Richard 1983 *Beyond Objectivism and Relativism*. Philadelphia: University of Pennsylvania Press.
- Bookman, Ann, and Sandra Morgen (eds.) 1988 *Women and the Politics of Empowerment*. Philadelphia: Temple University Press.
- Bredo, Eric, and Walter Feinberg (eds.) 1982 *Knowledge and Values in Social and Educational Research*. Philadelphia: Temple University Press.
- Brown, L. David, and Rajesh Tandon 1983 "The Ideology and Political Economy of Inquiry: Action Research and Participatory Research." *Journal of Applied Behavioral Science and Technology: An International Perspective* 19:277-294.
- Brydon-Miller, Mary, and Deborah Tolman (eds.) 1997 *Journal of Social Issues* 53(4):597-827. Special Issue on Transforming Psychology: Interpretive and Participatory Research Methods.
- Carr, Wilfred, and Stephen Kemmis 1986 *Becoming Critical: Education, Knowledge and Action Research*. London: Falmer.
- Chisholm, R. G., and M. Elden (eds.) 1993 *Human Relations* 46(2):121-298. Special Issue: Action Research.
- Comstock, Donald, and Russell Fox 1982 "Participatory Research as Critical Theory: North Bonneville U.S.A. Experience." Paper presented at 10th World Congress of Sociology, Mexico City. (Available from Russell Fox, Evergreen State College, Olympia, Wash. 98505.)
- Cornwall, Andrea, and Rachel Jewkes 1995 "What Is Participatory Research?" *Social Science and Medicine* 41:1667-1677.
- De Koning, Korrie, and Marion Martin (eds.) 1996 *Participatory Research in Health*. London: Zed.

- Fals Borda, Orlando 1988 *Knowledge and People's Power: Lessons with Peasants: Nicaragua, Mexico and Colombia*. New York: New Horizons.
- , and Mohammed Roshman (eds.) 1991 *Action and Knowledge: Breaking the Monopoly with Participatory Action Research*. New York: Apex.
- Forrester, Keith, and Kevin Ward 1989 "The Potential and Limitations: Participatory Research in a University Context." Participatory Research Conference Case Study, Division of International Development, International Centre, Calgary, Alberta.
- Freire, Paulo 1970 *Pedagogy of the Oppressed*. New York: Continuum.
- Gaventa, John 1980 *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley*. Urbana: University of Illinois Press.
- 1988 "Participatory Research in North America." *Convergence* 21:19–29.
- , and Billy Horton 1981 "A Citizen's Research Project in Appalachia, U.S.A." *Convergence* 14:30–40.
- , Barbara Smith, and Alex Willingham (eds.) 1990 *Communities in Economic Crisis*. Philadelphia: Temple University Press.
- Hall, Bud 1975 "Participatory Research: An Approach for Change." *Prospects* 8:24–31.
- 1981 "The Democratization of Research in Adult and Non-Formal Education." In P. Reason and J. Rowan, eds., *Human Inquiry*. New York: John Wiley.
- Harding, Sandra 1986 *The Science Question in Feminism*. Ithaca, N.Y.: Cornell University Press.
- Huizer, Gerrit, and Bruce Mannheim (eds.) 1979 *The Politics of Anthropology: From Colonialism and Sexism Toward a View from Below*. Paris: Mouton.
- Kassam, Yusuf, and Mustafa Kemal (eds.) 1982 *Participatory Research: An Emerging Alternative Methodology in Social Science Research*. New Delhi: Society for Participatory Research in Asia.
- Lather, Patti 1986 "Research as Praxis." *Harvard Educational Review* 56:257–277.
- Luttrell, Wendy 1988 *Claiming What Is Ours: An Economics Experience Workbook*. New Market, Tenn.: Highlander Research and Education Center.
- Maguire, Patricia 1987 *Doing Participatory Research: A Feminist Approach*. Amherst: Center for International Education, School of Education, University of Massachusetts.
- Merrifield, Juliet 1989 "Putting the Scientists in Their Place: Participatory Research in Environmental and Occupational Health." *Highlander Center Working Paper* New Market, Tenn.: Highlander Research and Education Center.
- Mies, Maria 1983 "Towards a Methodology for Feminist Research." In G. Bowles and R. Klein, eds., *Theories of Women's Studies*. London: Routledge and Kegan Paul.
- Mishra, Shirz, Leo Chavez, J. Raul Magana, Patricia Nava, R. Burciaga Valdez, and F. Allan Hubbell 1998 "Improving Breast Cancer Control among Latinas: Evaluation of a Theory-Based Educational Program." *Health Education and Behavior* 25:653–670.
- Park, Peter 1978 "Social Research and Radical Change." Presentation at 9th World Congress of Sociology, Uppsala, Sweden.
- , Mary Brydon-Miller, Budd Hall, and Edward T. Jackson (eds.) 1993 *Voices for Change: Participatory Research in the U.S. and Canada*. Westport, Conn.: Greenwood.
- , Budd Hall, and Ted Jackson (eds.) Forthcoming *Knowledge, Action and Power: An Introduction to Participatory Research*. South Hadley, Mass.: Bergin and Garvey.
- Participatory Research Network 1982 *Participatory Research: An Introduction*. New Delhi: Society for Participatory Research in Asia.
- Patel, Sheela 1988 "Enumeration as a Tool for Mass Mobilization: Dharavi Census." *Convergence* 21:120–135.
- Rose, Hilary 1983 "Hand, Brain and Heart: A Feminist Epistemology for the Natural Sciences." *Signs: Journal of Women in Culture and Society* 9:73–94.
- Shor, Ira (ed.) 1987 *Freire for the Classroom: A Sourcebook*. Portsmouth, N.H.: Heinemann.
- Small, Stephen A. 1995 "Action-Oriented Research: Models and Methods." *Journal of Marriage and the Family* 57:941–956.
- Smith, Dorothy 1987 *The Everyday World as Problematic: A Feminist Sociology*. Boston: Northeastern University Press.
- Stanley, Liz, and Sue Wise 1983 *Breaking Out: Feminist Consciousness and Feminist Research*. London: Routledge and Kegan Paul.
- Stoecker, Randy, and David Beckwith 1992 "Advancing Toledo's Neighborhood Movement through Participatory Action Research: Integrating Activist and Academic Approaches." *Clinical Sociology Review* 10:198–213.
- Stoecker, Randy, and Edna Bonacich (eds.) 1992, 1993 *American Sociologist*. Special Issues on Participatory Research, Part I, 23 no. 4 (Winter 1992); Part II, 24, no. 1 (Spring 1993).
- Tandon, Rajesh 1981 "Participatory Research in the Empowerment of People." *Convergence* 14:20–29.

- 1988 “Social Transformation and Participatory Research.” *Convergence* 21:5–18.
- Vio Grossi, Francisco 1981 “Socio-Political Implications of Participatory Research.” *Convergence* 14:43–51.
- Whyte, William F. 1989 “Advancing Scientific Knowledge through Participatory Action Research.” *Sociological Forum* 4:367–386.
- Yeich, S., and R. Levine 1992 “Participatory Research’s Contribution to a Conceptualization of Empowerment.” *Journal of Applied Social Psychology* 22:1894–1908.

FRANCESCA M. CANCIAN  
CATHLEEN ARMSTEAD

## PEACE

Humans have always prized and sought peace. The conditions believed to foster peace and the very conception of peace, however, have varied in different periods and cultures. In this article, we examine contemporary scholarly understandings of peace and how to achieve and maintain peace (Barash 1991; Galtung 1996; Stephenson 1994). In particular, we discuss the views of American sociologists and other social scientists who regard themselves as engaged in peace studies, peace research, conflict resolution, and related fields.

The concept of peace is contested. Some analysts use the term “peace” in opposition to war; this is *negative peace*, defined as the absence of direct physical violence. Other analysts stress *positive peace*, defined as social relations marked by considerable equality in life chances, by justice, or even by harmony. Some writers use the term “peace” to refer only to relations among global actors in a world system, while others include relations among persons and groups as well as among countries. Finally, some observers regard peace as a stable condition and others think of it as many never-ending processes.

In this article, we discuss certain aspects of positive peace, while focusing on negative peace. Furthermore, we emphasize international peace, but also consider large-scale relations within societies. With these focuses, we examine three categories of peace processes: (1) *building peace*, developing processes that prevent the emergence of destructive conflicts; (2) *making peace*, developing processes that contribute to deescalation and settlement of conflicts; and (3) *keeping and restoring*

*peace*, fostering processes that help maintain peace and construct equitable relations.

### BUILDING PEACE

The analysts providing the research and theorizing examined here and in the next sections vary in the relative importance they give to variables and conditions from different sources: from within one or more of the contending parties, from the relationship among them, and from their social context. Each is discussed in turn.

**Internal Factors.** Considerable work has been done about the processes and the conditions within countries that contribute to international peace or war and within large-scale groups that contribute to societal peace or destructive conflict. One such body of work stresses the role of self-serving elites in arousing, sustaining, and exacerbating antagonisms against other countries or groups. For example, during the years of the Cold War, many observers analyzed the existence and effects of a political–military–industrial complex in promoting the arms buildup in the United States and in the Soviet Union (Mills 1956; Sanders 1983).

More recently, peace workers have been directing their attention to political and intellectual elites who develop and promote ethno-nationalist ideologies. The way such ideologies are based on a socially constructed history and shared community is the subject of considerable analyses (Anderson 1991). In addition, many analysts stress the contribution that such ideologies make to the emergence and exacerbation of bitter fights and of genocide (Anthony Smith 1991). Finally, observers often examine how military, political, and intellectual leaders promote such ideologies for their personal benefit.

Peace workers believe that analyzing such processes demystifies them and their products. Furthermore, they believe that such unmasking undermines the effectiveness of those seeking to mobilize followers to wage struggles that deny legitimacy to their opponents. This kind of critical analysis is a major form of peace research.

Another large body of writing about building peace examines the education and socialization of members of a society or group in ways that promote peace. This includes research and theorizing about the ways this has been done and about the

ways that it might be done. The feminist scholarly perspective is an influential source for important contributions to this body of work. For example, using this perspective, the invisibility of women in studies of conflicts and peace processes becomes apparent, and feminist scholars provide new insights into international and domestic conflicts by paying attention to the roles women play in such conflicts (Enloe 1989). Furthermore, considerable research demonstrates consistent differences between women and men regarding support for the use of military means in international conflicts. The popular expectations, however, tend to exaggerate the degree to which men and women differ in their conduct in conflicts and negotiations (Stephenson 1996; Taylor and Miller 1994).

Feminist work tends to emphasize that the gender differences that do exist result significantly from past socialization of males and females into gender roles and from patriarchal social structures. Men learn to be relatively competitive and hierarchical, while women emphasize integrative relations. The feminist perspective fosters a vision that social relationships could be less patriarchal and therefore less unjust and less prone to destructive conflict than they generally have been.

How language and imagery are used to give meaning to conflicts helps frame conflicts and thus affects how they are waged. For example, analysts examine how the mass media and films contribute to an overreliance on violence and the threat of violence to wage conflicts (Gibson 1995). Such work also illuminates the processes of dehumanization of opponents in social conflicts, as well as revealing how such dehumanization contributes to the destructive escalation of struggles.

Since conflicts are inherent in social life, the role of social structure and culture in shaping how conflicts are waged is highly significant for building peace. Analysts are giving increasing attention to variations in the repertoire of methods used to conduct conflicts, including constructive ones, that are available for different people in different historical periods (Tilly 1978). Efforts to study and to train people in the methods of nonviolent action and problem-solving conflict resolution methods therefore contribute to building peace internationally and domestically (Kriesberg 1998).

**Relational Factors.** Several aspects of the relations among global and among societal actors

affect the likelihood that those actors will interact peacefully. One long-standing area of peace studies has been the effect of integration between societies and of sectors within societies. Integration is indicated by the high rate of exchange of goods, peoples, and ideas across societal and group lines, relative to exchanges within. Research findings support the generalization that such integration enhances mutual security and reduces the probability of countries' waging wars or threatening each other's security. Increased integration not only creates greater bonds of mutual interest and identity, but also improves communication and exchanges that parties regard as equitable. Furthermore, research on ethnic and other cleavages within societies also indicates the importance of integration, cross-cutting ties, and shared identities in preventing such cleavages from manifesting themselves in destructive conflicts (Dahrendorf 1959; Kriesberg 1998).

Considerable evidence has been reported indicating that democratic countries do not make war against each other (Gleditsch and Heegre 1997). Although the finding and particularly its interpretation are contested, the finding seems robust, given particular definitions of democracy and war. The finding may be explained by the tendency of governments in democratic societies to accord legitimacy to each other and credibility to each other's claims. Furthermore, negotiating differences may tend to be regarded as more acceptable and more skillfully practiced in democratic than in nondemocratic societies.

**Contextual Factors.** The social context within which possible adversaries interact certainly affects their relations. The context includes the social system within which adversaries interact, including the overall level of integration, the nature of institutional structures, the likelihood of external intervention in conflicts, and the kind of norms that are shared. The concepts of positive peace and structural violence help in understanding the relationship between social context and peace. Unlike personal violence, structural violence is indirect. It refers to the "avoidable denial of what is needed to satisfy fundamental needs" (Galtung 1980, p. 67). Thus, structural conditions may damage and cut short people's lives by restrictions of human rights or by malnutrition and illness, while other people using available knowledge and resources do not suffer the same deprivations. Such

inequities are built into the global order and constitute negative peace. This influential idea has stimulated various studies, particularly regarding conditions in peripheral or underdeveloped regions. The literature about the development of the world system and of colonialism obviously bears on this matter (Chase-Dunn and Hall 1997).

The expansion in the number, scope, and size of international nongovernmental organizations (INGOs) is a subject of growing sociological attention, reflecting INGOs' increasing global importance. Many kinds of transnational organizations perform activities and are arenas for interactions that supplement or even compete with states and with international governmental organizations. INGOs include multinational corporations, religious and ideological organizations, professional and trade associations, trade union federations, and ethnic associations. These groupings provide important bases of transnational identity and action (Smith et al. 1997).

Critical analysts view these developments as part of a new global order in which a transnational elite exercises hegemonic domination. Their work stresses the increasing global inequality and the development of a transnational elite that fosters globalization and profits from it. From this perspective, the U.S. government's promotion of democracy throughout the world is a method of maintaining order while promoting free markets and capitalism. Democracy, in this context, means *polyarchy*, a system in which a small group rules and mass participation is limited to choosing leaders in managed elections (Robinson 1996, p. 49).

International and supranational governmental organizations are also taking on increasing importance; witness the peacekeeping activities of the United Nations after the Cold War. Social scientists, including sociologists, have examined the conditions in which such institutions emerge and survive, how they serve to improve the quality of life, and how they may help to prevent conflicts from erupting and escalating destructively (Etzioni 1965).

The people of the world are already highly interdependent and are becoming increasingly so. This is true at the societal and at the global level. The flow of goods, capital, labor, ideas, and information is ever faster, ever less expensive, and ever more extensive. Consequently, the people of the

world share problems relating to environmental threats, governmental abuse and brutal conflicts spurring large-scale refugee flows, dislocations resulting from rapid social change, and challenging social relations among groups that are culturally different.

These phenomena contribute to the growing homogenization of the world. More and more people share images, ideas, and norms relating to consumer preferences, forms of entertainment, the protection of human rights, and economic development. But these phenomena also generate particularistic reactions and threaten destructive conflicts within and between societies. Experiences with these phenomena around the world and within each society are not the same for everyone. Some people reject the spreading secularism and the dominance of Western, particularly American, ideas and power. The empirical contradictions and the moral dilemmas arising from these developments are increasingly matters of inquiry among peace workers (Boulding 1990).

### MAKING PEACE

Recent peace work has focused on limiting the destructive escalation of conflicts, fostering transitions toward deescalation, and conducting negotiations that help end conflicts constructively. Internal, relational, and contextual factors contribute to these ways of making peace.

**Internal Factors.** Among internal factors, sociological work attends particularly to popular forces that pressure governments to move toward accommodations with external adversaries. This interest combines with the growth in theory and research about social movements to generate many studies of peace movements (Lofland 1993; Marullo and Lofland 1990). Analyses of campaigns against nuclear weapons and other evidence indicate that, at least within the United States and western Europe, public opinion and organized public pressure have influenced governments, often in the direction of peacemaking (Joseph 1993; Klandermans 1991).

In addition, certain internal structural factors can help leaders to recognize the needs of the other side and to communicate responsiveness. Such factors may include leaders who are accorded legitimacy, openness to considering alternative

courses of action, sources of good information about outside groups, and norms limiting intolerance. Such factors and specific policy-making procedures can help limit escalation, manage crises, and negotiate settlements (Wilensky 1967).

**Relational Factors.** Most work on peacemaking focuses on the relations between adversaries, including analyses of tacit bargaining, formal and informal negotiations, and providing mutual reassurance about security. The relationship between the United States and the Soviet Union, and their alliances during the Cold War, has been the subject of considerable study. Contributions have been made about the extent to which that conflict and the proxy wars associated with it were based on misunderstandings and on processes of dehumanization that fostered conflict escalation (Gamson and Modigliani 1971).

Other writing has drawn from and contributes to studies of problem-solving conflict resolution and conflict transformation. This work includes the analysis of ways of waging struggles constructively so that escalation is limited, and so that possibilities of reaching mutually acceptable accommodations are not foreclosed. This is an argument examined in studies of the use of nonviolent action, as in the American civil rights struggle and in many other conflicts (Powers and Vogeley 1997; Wehr et al. 1994; Sharp 1973).

Work on relational aspects of peacemaking also includes analyses of conciliatory gestures and other initiatives to deescalate conflicts, of the management of crises, of the transformation of intractable conflicts into tractable ones, and of strategies and techniques for negotiating mutually acceptable agreements (Patchen 1988). It also includes the efforts by persons in one camp to exchange information and possible options for peacemaking with their counterparts in the opposing camp, through conferences, dialogue groups, and ongoing workshops.

The ending of the Cold War illustrates the success of some of these methods (Kriesberg 1992). Specifically, such methods include negotiating mutual assurances that vital interests would not be threatened, as was done in the Conference for Security and Cooperation in Europe, resulting in the Helsinki Accords, signed in 1975. Included in the Accords was a shared commitment to norms, for example, about protecting human rights, which

fostered increased mutual exchanges. In addition, the American and Soviet military alliances established confidence-building measures and later restructured their military forces to be less provocative. Nonofficial channels such as the Pugwash meetings and the Dartmouth conference assisted in reaching agreements that helped to transform the Cold War. These developments have served as inspirations for efforts to limit or transform other regional conflicts.

The transformation of the conflict in South Africa about apartheid is another important illustration of the effectiveness of some of these methods (Kriesberg 1998). For example, the African National Congress (ANC), with the leadership of Nelson Mandela, consistently pursued nonracist goals, thus offering assurance that whites were and would remain recognized as South Africans. The means used in the struggle to end apartheid were considered in that light; they were initially nonviolent, and even when the decision to wage armed struggle was undertaken, terrorism was excluded. Informal and unofficial communications prepared the adversaries for working out an agreement that was acceptable to all the major adversaries in the seemingly intractable conflict in South Africa (van der Merwe 1989).

Of course, in some cases when challengers initiated nonviolent struggles, they were repressed or the conflicts escalated destructively. Such cases, as in China and in Northern Ireland, deserve and have received attention. However, there are many case studies and quantitative analyses indicating that reliance on violence and threat of violence is frequently counterproductive and often mutually destructive (Vasquez 1993).

**Contextual Relations.** One important aspect of a conflict's context, affecting its transformation and its peaceful settlement, is the involvement of intermediaries. Analysts using a sociological approach give attention to the role of nonofficial persons and groups as well as official intermediaries. Such intermediaries often provide a variety of mediating services, including helping bring adversaries to the negotiating table, facilitating meetings, aiding in developing new options, building support for an agreement, and helping to implement and to sustain an agreement that is reached (Burton 1990; Laue 1973). Persons and groups providing such services vary greatly in the way

adversaries understand their role and in the resources they bring with them. For example, unofficial mediators with relatively few material resources may be able to provide exploratory services relatively well, since their engagement generally involves low risks for the antagonists. Mediators with a major stake in the fight and with great resources can provide compensations and assurances that are relatively important in closing negotiations and implementing them.

The U.S. government and U.S. private citizens have played important mediating roles in many international and even internal conflicts in other countries (Kriesberg 1992). American mediation in Israeli-Arab conflicts has been particularly extensive and often crucially effective. U.S. secretaries of state and U.S. presidents have conducted major mediating efforts, often using considerable resources to induce the negotiating parties to conclude an agreement and to implement it. Among the notable agreements reached are those mediated by Secretary of State Henry Kissinger in 1974 between Israel and Egypt and between Israel and Syria, and the 1978 Camp David agreements between Israel and Egypt, mediated by President Jimmy Carter and leading to the two countries' signing a peace treaty.

Nonofficial persons and groups from the United States have also provided mediation services, for example in the conflict between Israeli Jews and Palestinians. Nonofficial channels have been particularly important due to the long-standing refusal of the Israeli government to recognize the Palestine Liberation Organization (PLO) as the representative of the Palestinians. Jewish and Palestinian Americans have often had the knowledge, interest, and contacts to provide useful channels for exploring possible options and ways of taking steps toward official mediation and negotiation.

Mediation efforts, obviously, do not always succeed. Many efforts never result in agreements. In some cases, agreements are reached but not ratified or not implemented. On occasion, agreements are followed by disastrous breakdowns, as happened in Rwanda in 1994. Of course, failing to mediate probably would not have yielded better results. Nevertheless, this indicates that we need to know much more about the type of mediation that tends to be effective at each stage of various kinds of conflicts.

Among the many other relevant contextual factors, we note only a few. First, changes in prevailing norms and understandings sometimes embolden one party in a conflict and undermine the faith of its opponent. The result is that a conflict that has long persisted can move toward resolution. For example, changing views about human rights and democracy contributed to ending the civil wars in Central America and apartheid in South Africa.

The context also includes a wide range of international governmental and nongovernmental organizations, with varying capabilities of contributing to peacemaking. The mass media are also increasing the global attention to especially terrible events at particular times. Finally, the increasing availability of weapons of all kinds enables more and more people to challenge existing conditions, as well as enabling those in authority to resort to violent means of control.

#### KEEPING AND RESTORING PEACE

The recent transformation and settlement of protracted international and societal conflicts and the radical transformation of previously authoritarian and repressive societies have heightened attention to the challenges of building postconflict relations that are enduring and just (Lederach 1997). Changes within one or more antagonist camps and between former antagonists are crucial in meeting these challenges. In recent years, analysts have given particular attention to the role of intermediaries, standards of human rights, and other elements of the antagonists' social context.

**Internal Factors.** A fundamental change in ways of thinking among members of one or more antagonistic sides can be a powerful factor in producing an enduring peace between them. This does sometimes happen. For example, most Germans after the defeat of Nazism repudiated what they themselves had believed and done; instead, they welcomed beliefs, values, and institutions shared with the victors. To some extent, a similar transformation occurred among Russians as the Cold War ended. As a result of the American civil rights struggle of the 1950s and 1960s, most southern American whites became convinced that they were wrong to resist ending the Jim Crow system of discrimination. Similarly, most South African

whites would now concur that the ending of apartheid was right and just.

Changes in internal social structure also are frequently crucial. Countries that have had internal conditions engendering overreliance on military means and goals that threatened vital interests of other countries may reduce their external threat only after undergoing a fundamental internal restructuring. The restructuring may entail civilian control of the military and the development of a civil society and democratic institutions. Peaceful accommodations in postconflict relations within a country may also depend upon fundamental changes in one or more sides of the past conflict. This occurs as governments change or as the leadership of an ethnic, a religious, or a class movement undergoes change.

**Relational Factors.** Traditionally, efforts to restore peace after a conflict ends include policies to redress the grievances that were viewed as the conflict's source. For communal differences within a country, this may entail more autonomy for citizens with different languages or religions and provisions for popular participation in determining the form and degree of autonomy. For example, during the early 1950s, the status of Puerto Rico in relationship to the rest of the United States was being reconstituted. A Puerto Rican nationalist group resorted to violence in seeking independence. The suppression of violent attacks while avoiding general repression, the availability of a legitimate electoral political process, social and economic improvements, and programs of integration and autonomy, including cultural nationalism, combined to produce a generally peaceful relationship in which alternative arrangements are contested within the established political system.

In the United States, a wide variety of methods and strategies are employed to redress grievances and increase equity; they include programs of affirmative action for women and minorities. Such programs, however, have become subject to challenge and have been reduced. This demonstrates the ongoing nature of conflicts related to socially constructed differences between citizens.

In recent years, peace workers have been giving considerable attention to fostering mutual understanding and tolerance among peoples with different cultural backgrounds living in the same society (Weiner 1998). This attention extends to

reconciliation between peoples who perpetrated gross human rights violations and peoples who suffered profound losses during periods of repression or of violent struggle. Reconciliation is complex, variously combining several processes: (1) acknowledging the truth of what happened; (2) administering justice for past misdeeds, and ensuring future justice and security; (3) extending forgiveness to members of the group that committed wrongs (sometimes in response to expressions of remorse); and (4) accepting responsibility by those who committed wrongs or failed to oppose them. In postapartheid South Africa, for example, the work of the Truth and Reconciliation Commission represents one way to deal with these postconflict issues.

A variety of recent developments contribute to reconciliation among the different peoples making up the United States. The truth about discrimination, violent repression, and other injustices regarding Native Americans, African Americans, and other groups has been more frequently acknowledged; this is evident in the mass media, in scholarly work, and in governmental statements. In addition, religious and other community organizations, corporations, and local governments have promoted or provided education programs, workshops, training, and dialogue groups to help persons of different communities learn about each other's experiences and perspectives.

Furthermore, long-standing policies have been instituted to strengthen a shared identity as Americans. The conception of Americans as belonging to a single ethnic group or an assortment of people melting into a single ethnicity, however, is changing. Instead, the multicultural character of America is increasingly accepted and even celebrated.

**Contextual Factors.** International organizations are increasingly expected to play critical roles in keeping and restoring the peace. United Nations and other peacekeeping forces have undertaken many more such tasks since the Cold War ended. Regional organizations and individual countries, particularly the United States, have intervened to restore and sustain peace (Moskos 1976; Segal and Segal 1993). Even after an agreement ending civil strife has been reached, the continuing engagement of external governments is crucial for the survival of the agreement and its implementation (Hampson 1996).



International nongovernmental humanitarian and advocacy organizations have grown greatly and are often helpful in restoring and maintaining peace (Lederach 1997). They may support the development of civil organizations that sustain peace. Even in the postconflict reconstruction of what was Yugoslavia, some success may be found. For example, many governmental and nongovernmental activities have helped the people in Macedonia manage external threats and the dangers of internal strife.

### CONCLUSIONS

Peace work and the ways of thinking about peace have greatly expanded in recent decades. Peace is increasingly understood to be multidimensional and dynamic. Consequently, the ways of promoting peace are also manifold, and they vary in different settings for different actors. Theory and research about aspects of peace and their promotion draws from and contributes to social theory and social practice.

Recent applied and scholarly peace work is based on past experience, but the realities of the current world necessitate fresh thinking and innovative practices. New approaches and ideas are developing, combining knowledge and experience from many new interdisciplinary fields, including conflict resolution, feminist studies, security studies, and international relations.

Much more work needs to be done to understand the nature of peace and how its various aspects can be promoted. Peace is not easily advanced, is never total, and is never wholly secure. Whatever peaceful gains may be made must be energetically defended against the inevitable threats arising from new challenges.

### REFERENCES

- Anderson, Benedict 1991 *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Barash, David P. 1991 *Introduction to Peace Studies*. Belmont, Calif.: Wadsworth.
- Boulding, Elise 1990 *Building a Global Civic Culture: Education for an Independent World*. Syracuse, N.Y.: Syracuse University Press.
- Burton, John 1990 *Conflict: Resolution and Prevention*. New York: St. Martin's.
- Chase-Dunn, Christopher, and Thomas D. Hall 1997 *Rise and Demise: Comparing World-Systems*. Boulder, Colo.: Westview.
- Dahrendorf, Ralf 1959 *Class and Class Conflict in Industrial Society*. Stanford, Calif.: Stanford University Press.
- Enloe, Cynthia 1989 *Bananas, Beaches and Bases*. Berkeley and Los Angeles: University of California Press.
- Etzioni, Amitai 1965 *Political Unification*. New York: Holt, Rinehart and Winston.
- Galtung, Johan 1980 *The True Worlds: A Transnational Perspective*. New York: Free Press.
- 1996 *Peace by Peaceful Means: Peace and Conflict, Development and Civilization*. Thousand Oaks, Calif.: Sage.
- Gamson, William A., and Andre Modigliani 1971 *Untangling the Cold War*. Boston: Little, Brown.
- Gibson, James William 1995 *Warrior Dreams: Violence and Manhood in Post-Vietnam America*. New York: Hill and Wang.
- Gleditsch, Nils Petter, and Havard Heegre 1997 "Peace and Democracy." *Journal of Conflict Resolution* 41:283-310.
- Hampson, Fen Osler 1996 *Nurturing Peace: Why Peace Settlements Succeed or Fail*. Washington, D.C.: U.S. Institute of Peace Press.
- Joseph, Paul 1993 *Peace Politics: The United States between the Old and New World Orders*. Philadelphia: Temple University Press.
- Klandermans, Bert (ed.) 1991 *Peace Movements in Western Europe and the United States*. Greenwich, Conn.: JAI Press.
- Kriesberg, Louis 1992 *International Conflict Resolution: The US-USSR and Middle East Cases*. New Haven, Conn.: Yale University Press.
- 1998 *Constructive Conflicts: From Escalation to Resolution*. Lanham, Md.: Rowman and Littlefield.
- Laue, James 1973 "Intervenor Roles: A Review." *Crisis and Change* III:4-5.
- Lederach, John Paul 1997 *Building Peace: Sustainable Reconciliation in Divided Societies*. Washington, D.C.: U.S. Institute of Peace Press.
- Lofland, John 1993 *Polite Protestors: The American Peace Movement of the 1980's*. Syracuse, N.Y.: Syracuse University Press.
- Marullo, Sam, and John Lofland (eds.) 1990 *Peace Action in the Eighties*. New Brunswick, N.J.: Rutgers University Press.

- Mills, C. Wright 1956 *The Power Elite*. New York: Oxford University Press.
- Moskos, Charles C. 1976 *Peace Soldiers*. Chicago: University of Chicago Press.
- Patchen, Martin 1988 *Resolving Disputes between Nations: Coercion or Conciliation*. Durham, N.C.: Duke University Press.
- Powers, Roger S., and William B. Voegelé (eds.) with Christopher Kruegler and Ronald M. McCarthy, associate eds. 1997 *Protest, Power, and Change: An Encyclopedia of Nonviolent Action from ACT-Up to Women's Suffrage*. New York: Garland.
- Robinson, William I. 1996 *Promoting Polyarchy: Globalization, U.S. Intervention, and Hegemony*. New York: Cambridge University Press.
- Sanders, Jerry W. 1983 *Peddlers of Crisis*. Boston: South End Press.
- Segal, David R., and Mady Wechsler Segal 1993 *Peacekeepers and their Wives: American Participation in the Multinational Force and Observers*. Westport, Conn.: Greenwood.
- Sharp, Gene 1973 *The Politics of Nonviolent Action*. Boston: Porter Sargent.
- Smith, Anthony 1991 *National Identity*. Reno: University of Nevada Press.
- Smith, Jackie, Charles Chatfield, and Ron Pagnucco (eds.) 1997 *Transnational Social Movements and Global Politics: Solidarity beyond the State*. Syracuse, N.Y.: Syracuse University Press.
- Stephenson, Carolyn M. 1994 "New Approaches to International Peacemaking in the Post-Cold War World." In Michael T. Klare, ed., *Peace and World Security Studies: A Curriculum Guide*. Boulder, Colo.: Lynne Rienner.
- 1996 "Gender Differences in Conflict Resolution." In *Report of the Expert Group Meeting on Political Decision-Making and Conflict Resolution: The Impact of Gender Difference*. Santo Domingo, Dominican Republic: United Nations Division for the Advancement of Women (EGM/PDCR/1996/rep.1).
- Taylor, Anita, and Judi Beinstein Miller (eds.) 1994 *Conflict and Gender*. Cresskill, N.J.: Hampton Press.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.
- van der Merwe, Hendrik 1989 *Pursuing Justice and Peace in South Africa*. New York: Routledge.
- Vasquez, John A. 1993 *The War Puzzle*. Cambridge: Cambridge University Press.
- Wehr, Paul, Heidi Burgess, and Guy Burgess (eds.) 1994 *Justice without Violence*. Boulder, Colo.: Lynne Rienner.
- Weiner, Eugene (ed.) 1998 *The Handbook of Interethnic Coexistence*. New York: Continuum.

- Wilensky, Harold L. 1967 *Organizational Intelligence: Knowledge and Policy in Government and Industry*. New York: Basic Books.

LOUIS KRIESBERG

## PENOLOGY

*Penology*, an applied field of sociology, is the theoretical study of prison policy, prison management, and the resulting prison culture. The sociological contributions to prison issues are applicable worldwide and offer practical solutions to problems relating to overcrowding, prison violence, and prison culture. While there is a rich European and American history of prison (e.g., Howard 1777; Beccaria 1819; Bentham 1843; Foucault 1977; Hirsh 1992), the focus here is on American prisons from the 1940s to the commencement of the twenty-first century. The following discussion begins with major sociological analyses of prison and concludes with changes in American prison policy over the past fifty years.

With a few exceptions (e.g., Goffman 1966), prison sociology has traditionally followed some variety of structural analysis. This theoretical perspective is concerned with a societal member's values, attitudes, roles, activities, and relationships that are assumed to be significantly influenced by the organization and structure of the member's environment. There are two broad types of this perspective that have been utilized in prison analysis; both focus on the structural components of prison and the members of that system. The first type, structural functionalism, focuses on how these components affect order in prison. The second type, conflict theory, focuses on how the structural components of prison and the larger society create conflict in prison.

### THEORETICAL ANALYSIS

**Structural Functionalism.** Clemmer's *The Prison Community* (1940) is an early form of structural functional analysis. Based on interviews with staff and inmates, Clemmer's exploratory study is an ethnography of prison. This pioneering work set forth the argument that administrative policy (i.e., authoritative control over inmates, or the lack of control over inmates) influenced inmate subculture. This subculture included behavior, rules, and

attitudes. Inmates could be unruly, and could prey on each other and staff if control was lacking.

This behavior would disrupt prison order and inmate adjustment to prison life. The prison administration would attempt to reestablish order by increasing control efforts, and inmates would counter these efforts through a process of "prisonization." "Prisonization" was a socialization process through which seasoned inmates inducted new inmates into the inmate subculture. The content of this socialization for new inmates included the learning of negative attitudes toward work, government, family, and inmate groups other than their own. Inmate behavior would become resistant, obstructive, and subterfugal. Clemmer saw this "prisonization" and the resultant inmate subculture as unfortunate and unintended consequences of the administrative controls. He believed the negativism and hostility perpetuated by the inmate subculture was disruptive to the inmates' reform and was "a stronger force for evil than the programs are for good" (Clemmer 1940, p. xiii).

Sykes's *The Society of Captives* (1958) was a qualitative, exploratory, and ex post facto study that also applied a structural functional analysis. Attempting to determine the cause of a series of prison riots, Sykes studied institutional records and interviewed correctional officers, civilian work supervisors, and inmates. Sykes found the cause of the prison disequilibrium to be rooted in the prison structure and values that followed from a policy of control. He argued that management's major task was to control the inmates, but the prison's system of power was flawed with structural weaknesses that left administrators serious difficulties in imposing their control regime on the inmates. These structural weaknesses involved the inmates' lack of a sense of duty to obey correctional officers, and the correctional officers' lack of legitimate rewards and punishments, with which they could encourage inmate submission.

Power based on authority, Sykes claimed, has two essential elements: the legitimacy of control efforts and a sense of duty to obey by those who are controlled (1958, pp. 46-47). He found the latter to be present down to the correctional officer level at the New Jersey State Prison, which operated under a traditional organizational hierarchy. The sense of duty to obey disappeared,

however, when the control efforts were applied to inmates. The correctional officers had to make deals (i.e., giving correctional officer duties to trusted inmates in exchange for their help in preventing trouble) and compromises (i.e., overlooking rule violations) with inmates to achieve compliance and order. Sykes argues that correctional officer corruption (i.e., reciprocity) could not be eliminated by replacing the correctional officers. New correctional officers were aggressively pressured by the inmates (i.e., threats of riots, blackmailing staff) until they, too, compromised the rules and regulations.

The problem was a weakness of the prison system. "The effort of the custodians to 'tighten up' the prison undermines the cohesive forces at work in the inmate population and it is these forces which play a critical part in keeping the society of the prison on an even keel" (Sykes 1958, p. 124). The cohesive forces are the less violent and more stable inmates who are given illicit privileges in exchange for their help in encouraging inmate cohesiveness and prison equilibrium. If the prison officials strip these inmates of their power (tighten up), the more violent and less stable inmates rise to power.

In addition to providing the structural functional analysis described above, Sykes argued that the result of imprisonment on inmate values, attitudes, and behaviors is a product of the patterns of interaction the inmate experiences on a daily basis (1958, p. 134). For example, inmates feel helpless and frustrated when staff refuse to explain bureaucratic decisions. Sykes described the hardships of imprisonment (i.e., rejection, degradation, deprivation, alienation, and lack of safety) felt by the inmates, and he acknowledged that "Somehow the imprisoned criminal must find a device for rejecting his rejectors, if he is to endure psychologically" (1958, p. 67). Sykes also depicted the correctional officers' work environment as dangerous and tense. He reported that the correctional officers were heavily outnumbered in a potentially violent setting, and were often frustrated by the administrative pressure to maintain order while they lacked unconditional compliance from the inmates.

*The Prison: Studies in Institutional Organization and Change* by Cressey (1961) and *Theoretical Studies in Social Organization of the Prison*, by Cloward

and colleagues (1960) were also structural functional analyses. Both studies depicted prisons as authoritarian systems governed by bureaucratic hierarchy and empowered to control inmates. They described how the control policy in prison affects the power structure which, in turn, influences communication and values (i.e., staff pursuit of power over inmates, inmate restraint from talking to staff). Cressey (1961) found that two different prison policies, one emphasizing custody and the other treatment, had different hierarchies with contradictory purposes. The two hierarchies, with different models of decision making, one authoritative and the other participatory, had to increase their efforts to communicate with each other to facilitate the security desired by the custody branch and the programming desired by the treatment branch. It was also reported in the Cressey study (1961) that inmates and staff can work together effectively when the role expectations of their respective groups are not involved. Effective communication deteriorates when inmates and staff allow their respective group pressures to interfere with their relationships (1961, pp. 229–259).

Webb and Morris's *Prison Guards* (1978) focused on the organization of prison structure, which Clemmer (1940) had felt contributed to the maintenance of prison order (or disorder) and determined the inmates' prison experience. They focused on the correctional officers' subculture, and concluded that the officers saw maintaining security as their main function. While discipline was expressed as the means for carrying out the policy, poor prison facilities and inexperienced administrators were seen as barriers to prison order and safety.

Webb and Morris (1978) reported that the officers' major complaints about administrators were lack of communication and failure to consult the officers when decisions had to be made. This suggests that those making the decisions were distrustful of the officers' ability to participate in decision making, and preferred to monopolize power at the top. It also means that information that the officers would know best (i.e., security issues relating to certain inmates) was not being considered by those making the decisions. The correctional officers saw their safety compromised. The frustration caused by this poor communication process led to correctional officers being

openly critical of the administration, and to their finding ways of undermining the administration's authority (i.e., doing no more than they were told).

Webb and Morris (1978) described how the officers developed their own protective society. The veteran correctional officers pressured the new to become "hardened" and "con-wise," and if, after six months to a year, the officers were not seen as "con-wise," then they were given the derogatory label "pro inmate." This pressure to be estranged and nonsympathetic in their attitudes toward inmates was the correctional officers' way of protecting themselves from "getting conned" or "being burned" by the inmates.

Kauffman's *Prison Officers and Their World* (1988) was another structural functional analysis of the correctional officer subculture. Correctional officers had been blamed by the central administration for the riots and inmate violence that had plagued the Massachusetts prison system in the 1970s. Based on interviews with sixty correctional officers, Kauffman describes the rite of passage that many new correctional officers went through: how the recruits were assigned the high-inmate-contact positions in cellblocks, abandoned by the administration and veteran officers, and aggressively tested by the inmates. The general result of this treatment was an emotional hardening toward the inmates, and the acceptance of violence by staff as a means of controlling inmates. Kauffman concludes that the power veteran officers had over staff recruits was an unintended consequence of the lack of a clear administrative policy for bringing order to the prison system. She found that the administration was determined to find someone to blame for the rite of passage many recruits went through, rather than adopting a policy "to counter the continuing legacy of chaos and violence" (1988, p. 199).

The influence of prison policy on the structure, process, behavior, and attitudes in the prison (a focus typical of both structural perspectives) is also represented in DiIulio's *Governing Prisons* (1987). He shows how prison management problems have been addressed in Texas, California, and Michigan through the development and use of different management policies. He refers to the Texas policy as one of "control," where authority is centralized and where the use of administrative-controlled inmates to control other inmates

("building tendering") prevailed until the decision in *Ruiz v. Estelle* (1980) ruled the system illegal. DiIulio claims that California has a "consensus" policy, in which each warden works to get inmate "consensus" on classification systems, there are elaborate in-service training programs for correctional officers, and there are community-related educational/vocational programs for inmates. DiIulio (1987) found that Michigan has a third alternative, which emphasizes a bureaucratic system of measuring and dispensing different levels of inmate "responsibility" and accountability. DiIulio found the quality of life for inmates inside the Michigan prisons better than what he had found in other state prisons. Staff had a difficult time accepting the responsibility model, however.

**Conflict Theory.** The second theoretical strand of prison sociology, conflict theory, saw tight controls on inmates as unjust and called for the sharing of power with them. Wright's *The Politics of Punishment* (1973) is the first of three works representing this second branch of structural analysis that will be reviewed here. Wright is more critical of prison operations than the structural functionalists. He thought rehabilitation was manipulative in the way that it attempted "to coerce the prisoner to conform to established authority" (1973, p. 325), and then use this conformity as the basic criterion for parole. Wright believed inmates were almost totally helpless to protect themselves against the dehumanization of "liberal" totalitarian rule he found in most prisons.

The prison administration's focus on rehabilitation was an individualistic solution that avoided and disguised the real causes of criminality. Wright claimed that it was the structural flaws in the capitalist system (i.e., unemployment and the lack of opportunities outside the prison), not individual failings, that were the problem. Wright believed that crime could be reduced, and prisons reformed, by moving the American capitalist society toward socialism. Within this context, prisons could be decentralized and not controlled by a "self-perpetuating, unrestrained bureaucracy" (1973, p. 342). The administration of punishment could be placed under public surveillance. The prison under this system would "serve the interests of the people rather than of the elite" (1973, p. 337). Wright suggests that prison management has typically held to the position that prison order can be

obtained only with traditional control strategies (i.e., correctional officer power over inmates), and that the effect of these strategies has been oppressive and inhumane prison conditions.

Hawkins's *The Prison: Policy and Practice* (1976) is another critical analysis of prisons and prison management. Although Hawkins's prison experience was gained in the United Kingdom, his message was relevant to American prison practitioners and researchers. Among his claims was that correctional officers had only been superficially studied by the functional branch of sociologists, who presented them as stereotypes acting out assigned roles (1976, p. 81). Hawkins saw the correctional officers in a situation of role conflict (security versus treatment) that was due to the absence of clear job descriptions and training. The results of this conflict for the correctional officers were negative attitudes toward the administration and obstructive behavior. These results, along with the conflicting goals of imprisonment (punishment versus rehabilitation), were seen by Hawkins as impediments to prison reform. Hawkins also urged that prison evaluations should be more than just "eulogistic and imprecise descriptions of success stories" (1976, p. 178).

In his assessment of the Attica prison in New York, Hawkins also found that the relationship between correctional officers and inmates was a crucial factor in the 1971 riot. The problem was mostly structural. The officers were assigned to large blocks of inmates, rather than to small units where they could have established rapport with the inmates. In the large units the officers worked with different inmates each day, thus reducing the opportunities and motivations for staff and inmates to develop mutual respect and understanding. The structural problem went further, however. Hawkins, citing a New York State Special Commission's report on Attica, reported that there were unnecessary priorities in the correctional officer recruitment policies (i.e., age, physical size, and strength were more important than skills in persuasion, leadership, and interpersonal relations).

Irwin's *Prisons in Turmoil* (1980) also represents this "conflict" branch of sociological theorizing. Irwin, an ex-felon and past leader in the prison reform movement of the 1970s, shows how American penal policy has followed changes in the economic and political conditions in the United States.

His thinking ran against the usual functionalist sociology, which explained prison policy and activity largely in terms of internal events. Irwin believed that sociologists (e.g., Clemmer 1940; Sykes 1958) who studied prison cultures were blinded by their focus on prison structure, and therefore failed to recognize the fact that when inmates enter prison, they bring their outside values with them. He also believed that if prisons are to be safe and free from malicious authoritarianism, staff and inmates should have more formal input into policy and grievance decisions.

These sociological analyses of prison were written in different phases of the development of penal policy. It is very difficult to determine whether the literature influenced the policy, or vice versa. Most likely, they shaped each other. The point of the following brief history of the development of American penal policy over the past fifty years is not to make the causal argument, but to help the reader get a sense of the changing ideologies leading to the current management model.

#### CHANGES IN AMERICAN PENAL POLICY

**Retribution to Rehabilitation.** Prisons in the United States have traditionally been centralized hierarchical organizations with authoritarian and coercive management. Policies and procedures have been formulated at the central administrative offices and passed down in military fashion to prison administrations for implementation. This is assuredly true of the management style that administered the penal philosophies of the 1950s. The decade began with a "Big House" warehousing of inmates that implemented a classical penal philosophy embracing retribution, "care," and the notion that punishment should fit the crime. The decade ended with a changed perspective on crime and punishment, emphasizing that punishment should fit the individual and the inmate should be rehabilitated. By 1958, a noncustodial treatment branch of prison management shared an uneasy coexistence with the custody branch. Three types of treatment programs (psychological, educational, and economic) were centrally administered under the rubric of the rehabilitative ideal (Irwin 1980). The idea was to treat the individual problem that "caused" the individual to commit the crime. By the end of this decade, most prison

administrators had implemented this rehabilitative model without any interference from federal and state governments.

**Rehabilitation to Reintegration.** The early 1960s saw variations of the rehabilitative or treatment model of the late 1950s sharing an uneasy coexistence with an authoritative management style that had traditionally maintained order and control in large prison bureaucracies. Public and prison administrative acceptance of the "rehabilitative ideal" as a penal policy was partly rooted in its advocates' view that the new scientifically based treatment model could provide a remedy for crime.

By the middle 1960s, the call for community involvement in rehabilitation was signaled in a report by the President's Commission on Law Enforcement and Administration of Justice. Advocating a shift to a policy of reintegration, the report advised that rehabilitation might be best done outside prison, and that community programs should be designed that could change both the inmate and society (U.S. President's Commission 1967, pp. 27-37). It took a few years to catch on, but in the late 1960s a number of community alternatives to prison were developed and heavily underwritten by federal revenues.

**Reintegration to Retribution.** This reintegration policy continued into the 1970s. Prison administrators were pressed to design reentry programs such as work and educational release, and home furloughs. But public support waned in the mid-1970s, when the "get-tough-with-criminals" and antirehabilitative mood became dominant. The second theoretical strand of prison sociology, conflict structuralism, began around this time. This orientation was influenced by the 1960s societal reaction and critical theories that denounced the 1950s conservative structural functionalism. Riots and litigation forced federal courts to scrutinize the repressive management of correctional facilities and community-based treatment programs to ensure compliance with the Eighth Amendment right to protection from cruel and unusual punishment.

If prisons were being unfair in their partiality toward manipulative rehabilitation (e.g., Wright 1973; Hawkins 1976: the liberal critique), then retribution ought to replace it as a penal goal. The mid- to late 1970s saw a readoption of the early

classical policy—"let the punishment fit the crime"—rather than the rehabilitative or positivist policy of "letting the punishment fit the criminal." A different argument was that rehabilitation rarely worked (e.g., Martinson 1974: the conservative critique). Prison technologies had not successfully rehabilitated inmates despite prison administrators' repeated claims of success. The Martinson survey (1974) of 231 treatment reports in New York concluded that treatment had not worked. The publication of this survey ended any lingering hope for serious rehabilitative programming.

The "new" word in penal policy was "retribution." Though it had been criticized by many social scientists and judges since the early 1900s, in 1972 the U.S. Supreme Court offered retribution as an appropriate reason for capital punishment (*Furman v. Georgia*, 408 U.S. 238, 1972). The neoclassical doctrine generally implies deterrence as a function of punishment. Wilson and Herrnstein (1985) are among those who take this position. There are others, however, who argue that retribution is the main function, and that there does not have to be any further utilitarian purpose such as deterrence. The argument these penologists make for retribution is that inmates are incarcerated because they deserve to be, not because they should be treated and not because their punishment should serve as a deterrent. Many penologists credit the retributive model with encouraging prison management to create an environment where inmates can serve their time safely and productively, rather than a harsh environment designed to deter crime. The deprivation of freedom was a harsh enough punishment: Prisons did not have to be punishing beyond that.

As the 1980s commenced, national policy emphasized broad reductions in domestic welfare programs. The criminal justice system warmed to the reacceptance of the classical notion that punishment should fit the crime rather than the criminal. The widespread adoption of determinate sentencing continued throughout the 1980s and 1990s, as it became increasingly difficult for social theorists to agree on definitive causes of crime. By 1999, most states had determinate sentencing, action supportive of the classical model. This, along with the conservative "get even tougher" response to violent crimes and drug offenses, was why the U.S. prison population more than tripled from 1980 to 1998. With 1,277,866 inmates in state and federal

prisons (U.S. Department of Justice 1999), the United States ranks with Russia as having the highest-known incarceration rates of any industrialized nation in the world.

**The Penal Debate.** One contemporary justification for prison is its crime deterrence function, for which there is considerable public support. Many sociologists take the neoclassical position that an appropriate and updated crime control policy should include an increase in the swiftness and certainty of punishment for street offenders, and that control efforts should focus on both the before-sentencing work of police and prosecution, and the after-sentencing work of corrections. The deterrence advocates' penal policy is based on the theory of criminality that proposes that the frequency of an individual's criminal behavior is a matter of the individual's choice, and a consequence of that behavior's reinforcement or punishment in the past. The individual's choice is based on the expectation of subsequent reward or punishment.

Despite the general public's belief that prison is a deterrent to crime, there is considerable research evidence to the contrary, leaning some sociologists against deterrence. They argue that crimes of passion are less apt to be deterred by prison sentences than crimes committed for monetary gain. The threat of prison is less likely to deter someone of lower socioeconomic status, who has less "good life" and status to lose, than someone of wealth and higher status. Additionally, the courts are so overcrowded that the possibility of swift and certain punishment is diminished.

Entering the twenty-first century, the prevailing model of sentencing is still grounded in the neoclassical deterrence ideology. When it comes to the serving of the sentence, however, the neoclassical retributive, nondeterrence doctrine prevails. This retribution policy requires a correctional system that does not abuse individual treatment, but instead makes rehabilitation programs voluntary, restricts autonomy of movement, specifies the length of sentence in advance (determinate), and implements it in a fair and safe environment. Since about 1973, prison administrators have addressed these fairness and safety issues, specified in the retribution policy, by attempting to minimize the potential violence associated with overcrowding, gang membership, and racism. They

have tried to do this by separating the inmate populations into small functional housing units (Levinson and Gerard 1973) that are supervised by a decentralized “unit management” team. Current research (i.e., Farmer 1994) has shown that, if implemented properly, this method of managing inmates can facilitate increased communication, safer conditions, a more satisfying work environment for staff, and more personal programming for inmates.

Sociologists and penologists generally expect the twenty-first-century prison situation to continue its current trends. Incarceration rates will continue to climb as sentences become longer, and as good-time awards and parole options decrease. Acquired immune deficiency syndrome (AIDS) and racism in prison will be more closely watched and more rigorously dealt with. Inmates will become more violent as confinement, with less programming, becomes more Spartan and harsher. Gang membership in prison will increase, along with management efforts to combat the membership. Prison “privatization” in areas of prison construction and management, inmate work, and services (i.e., food, health, counseling, education) will become a more widely accepted solution to increasing prison costs and decreasing prison budgets.

## REFERENCES

- Beccaria, Cesare 1819 “*Crimes and Punishments*,” transl. from Italian by M. D. Voltaire. In F. P. Williams, III and Marilyn D. McShane, ed., 1998 *Criminology Theory: Selected Classic Readings*, 2nd ed. Cincinnati: Anderson Publishing.
- Bentham, Jeremy 1843 “Introduction to the Principles of Morals and Legislation.” In F. P. Williams, III and Marilyn D. McShane, ed., 1998. *Crime Theory: Selected Classic Readings*, 2nd ed. Cincinnati: Anderson Publishing.
- Clemmer, Donald 1940 *The Prison Community*. Boston: Christopher.
- Cloward, Richard A., Donald R. Cressey, George H. Grosser, Richard McCleery, Lloyd E. Ohlin, Gresham M. Sykes, and Sheldon L. Messinger 1975 *Theoretical Studies in Social Organization of the Prison*. Millwood, N.Y.: Kraus.
- Cressey, Donald (ed.) 1961 *The Prison: Studies in Institutional Organization and Change*. New York: Holt, Rinehart, and Winston.

- DiIulio, John J. 1987 *Governing Prisons: A Comparative Study of Correctional Management*. New York: Free Press.
- Farmer, J Forbes 1994 “Decentralized Management in Prison: A Comparative Case Study.” *Journal of Offender Rehabilitation* 20 (Fall–Winter):117–130.
- Foucault, Michel 1977 *Discipline and Punish*. New York: Pantheon.
- Goffman, Irving 1966 “On the Characteristics of Total Institutions: Staff-Inmate Relations.” In D. R. Cressey, ed., *The Prison*. New York: Holt, Rinehart and Winston.
- Hawkins, Gordon 1976 *The Prison: Policy and Practice*. Chicago: University of Chicago Press.
- Hirsh, Adam J. 1992 *The Rise of the Penitentiary*. New Haven, Conn.: Yale University Press.
- Howard, John 1777 *The State of Prisons in England and Wales*. London: E. Eyres
- Irwin, John 1980 *Prisons in Turmoil*. Boston: Little Brown.
- Kauffman, Kelsey 1988 *Prison Officers and Their World*. Cambridge, Mass.: Harvard University Press.
- Levinson, Robert B., and Roy E. Gerard 1973 “Functional Units: A Different Correctional Approach.” *Federal Probation* 37(4):8–16.
- Martinson, Robert 1974 “What Works? Questions and Answers about Prison Reform” *Public Interest* 35:22–54.
- Sykes, Gresham M. 1958 *The Society of Captives*. Princeton: Princeton University Press.
- U.S. Department of Justice, Bureau of Justice Statistics 1999 <http://www.ojp.usdoj.gov/bjs>.
- U.S. President’s Commission on Law Enforcement and Administration of Justice 1967 *The Challenge of Crime in a Free Society*. Washington, D.C.: Government Printing Office.
- Webb, G. L., and David G. Morris 1978 *Prison Guards: The Culture and Perspective of an Occupational Group*. Houston, Tex.: Coker.
- Wilson, James Q., and Richard J. Herrnstein 1985 *Crime and Human Nature*. New York: Simon and Schuster.
- Wright, Erik O. 1973 *The Politics of Punishment*. New York: Harper Row.

J FORBES FARMER

**PENSION SYSTEMS**

See Retirement; Social Security Systems.



## PERSONAL AUTONOMY

*Personal autonomy* refers to a person's sense of self-determination, of being able to make choices regarding the direction of her or his own actions, including the freedom to pursue those choices. With personal autonomy, an individual is able to engage in effective self-regulation—successfully monitoring needs and values; responding adaptively to the environment, and initiating, organizing, and directing actions toward the achievement of needs. For some theorists, the psychological experience of autonomy has its origin in the organism's natural tendency to organize both itself and its environment in the pursuit of goals. In this view, a sense of autonomy requires the absence of restraining forces that can limit this natural tendency. Importantly, feelings of autonomy are not only crucial for adequate intrapersonal functioning—competent action and adequate psychological health—but are also essential for the adequate functioning of a healthy society.

### CONCEPTIONS OF PERSONAL AUTONOMY

Early personality theorists viewed autonomy as one element of a dialectical process in the developing self. Angyal (1941), for example, proposed that personality develops in the context of two conflicting pressures, autonomy and surrender (or homonomy). A pull toward autonomy leads toward differentiation from other people and the physical environment, connoting individuation, separation, independence, freedom, and the like. The tendency toward autonomy, however, is met by a countervailing pull toward surrender, felt by the individual as a desire to become part of something greater than oneself, uniting with others and with the physical environment. Surrender is reflected in concepts such as community, union, interdependence, and obligation. A similar dialectic can be seen in the theorizing of Otto Rank (1929) and David Bakan (1966). For instance, in Bakan's approach, a concept comparable to autonomy is agency, a tendency toward manipulation that results in aloofness and differentiation of the personality. Agency is viewed as occurring in conflict with communion, a tendency that pushes a person toward connectedness and personality coherence.

The dialectical view of autonomy is interesting, suggesting as it does that autonomy has little meaning outside some notion of wholeness or integration against which the individuating, segregating pressure of autonomy can push. More recent theorists also seem to understand that the concept of autonomy implies the question "Autonomy from what?" But rather than viewing autonomy as one element in an intrapsychic union of opposites, current conceptualizations focus on conflicts between an individual's need for self-determination and potential external constraints encountered in the social environment (Deci and Ryan 1985; Ryan et al. 1997). In this view, autonomy is conceptualized as reflecting the organism's natural developmental trajectory toward increasing complexity and the concomitant press toward greater organization of its environment in the process of self-development (Ryan et al. 1997). But, of course this natural development is not guaranteed. Insufficient resources—such as insecure attachment in infancy, emotional detachment in adolescence, inadequate social support in adulthood, or even neurobiological deficits in the individual—represent potential oppositional forces to the expression of autonomy. Desires for interpersonal relatedness do not stand in conflict with autonomy needs from such a perspective, but rather play a supportive role (Ryan 1991). Thus, the development of autonomy requires responsive parental nurturing, including recognition of and support for the child's expression of autonomy. In adulthood, the sense of autonomy is facilitated by an interpersonal environment that allows the individual to view his or her intentions to act as being caused by internal, personal motivations rather than being caused by external sources.

### THEORIES OF AUTONOMY

Theories relevant to an understanding of autonomy all share the assumption that individuals are motivated in some way to have the freedom to determine their own fate. For instance, deCharms (1968) proposed a general motivational tendency to strive to be an agent of causality. Individuals who initiate an intentional behavior experience themselves as the causal origin of the action and as intrinsically motivated. Individuals who do not experience personal causation, but rather view themselves as pawns being impelled by external

causes will experience action as being extrinsically motivated.

**Causality Orientations Theory.** Intrinsic motivation is viewed as a basic ongoing motivational propensity of all individuals that directs activity unless it is blocked in some way (Deci and Ryan 1985). Major constraints on experiencing intrinsic motivation can include stable or transient individual differences in the ways that individuals make sense of events. The most important causal orientations for understanding autonomy are the autonomy orientation and the control orientation. Individuals experiencing an autonomy orientation do not feel their behavior to be controlled by external contingencies. Instead, the individuals' experiences are ones of choice, flexibility, awareness of needs, effective accommodation to the environment, and responsiveness to available information. Such individuals are able to effectively seek out situations that allow them to experience autonomy, to use information in initiating action, and to be resilient in the face of difficulties. Only in an individual experiencing a "control orientation" does the tension between desired autonomy and controlling forces become salient and intrinsic motivation become reduced. In situations requiring action, such individuals experience feelings of pressure and anxiety, and action comes to be viewed as controlled by either internal factors not of their making (e.g., perceived obligations) or external forces (Deci and Ryan 1985). An individual's causal orientation is important because failure to experience autonomy is associated with reduced functioning, poorer health, and increased psychopathology (Deci and Ryan 1987; Ryan et al. 1995).

**Undermining Effects of Reward on Intrinsic Motivation.** Intrinsic and extrinsic motivation are affected not only by individuals' orientations toward explaining the causes of their actions, but also by the presence of external rewards or punishments for engaging in certain actions. Engaging in an intrinsically motivated activity means that it is the feelings of enjoyment and excitement that result from the activity that are rewarding. Presence of other rewards is unnecessary to such experiences (Deci and Ryan 1985). In fact, an important finding is that receiving rewards can be detrimental for behavior that is normally intrinsically motivated. Children or adults who begin receiving rewards for engaging in such behavior

may find motivation for that activity to become extrinsic in nature (Lepper and Greene 1978). Paradigmatic demonstrations of this phenomenon involve paying college students for performing enjoyable tasks such as solving puzzles and giving children a ribbon and gold star as a good-player award for playing with magic markers (Deci 1971; Lepper et al. 1973). Subsequent to receiving rewards, interest in the activity can be decreased once rewards are again absent. This decrease in task interest is thought to indicate a shift toward extrinsic motivation for performance of the task. Moreover, rewards can have similarly deleterious effects on tasks that require an individual to be creative.

Both the pervasiveness and the interpretation of these negative effects on task interest and creativity remain matters of heated debate (Eisenberger and Cameron 1996, 1998; Hennessey and Amabile 1998; Lepper 1998; Sansone and Harackiewicz 1998). There do, however, appear to be some reliable conditions under which tangible rewards such as money, candy, or gold stars decrease intrinsic motivation. Receiving expected rewards regardless of the quality of one's performance leads participants to subsequently spend less time engaging in a task once the reward is removed.

*Cognitive Evaluation Theory.* The primary autonomy-based interpretation for reduced task interest comes from cognitive evaluation theory (Deci and Ryan 1985). Cognitive evaluation theory suggests that, to the extent that rewards are controlling, intrinsic motivation will be decreased. Controlling events are ones that make individuals feel pressured to behave in a certain way. The presence of controlling events decreases feelings of self-determination by leading perceivers to believe that they are acting in order to receive a desired outcome. For example, tangible rewards like money or physical awards that are given regardless of the quality of the work can lead an individual to view actions as being instrumental for getting a desired outcome, and thus as controlled by those rewards. But receiving rewards can also be informative, conveying that one is competent. Such feedback about performance increases intrinsic motivation. Thus, in some contexts the informational function of a reward may override the decrease in intrinsic motivation that results from its controlling function. This is particularly the case,

according to cognitive evaluation theory, for unexpected or intangible rewards such as praise. In addition, when receipt of a reward is contingent on performance, a tangible cue highly symbolic of one's achievement can also be intrinsically motivating (Harackiewicz et al. 1984; Sansone and Harackiewicz 1998). Related research on creativity has been explained in terms of how rewards can orient the individual toward goal-relevant stimuli. For creativity tasks where one's thinking needs to be divergent and less stimulus-bound, such a goal-related focus can be counterproductive, reducing the cognitive flexibility and intense involvement in a task needed for producing novel solutions (Amabile 1983).

The findings regarding intrinsic and extrinsic motivation have important implications for the ideal socialization of members of a society (Ryan et al. 1997). One task of a society is appropriate socialization of its members, directing their behavior in productive avenues. In a sense, a major task of culture is to provide individuals with appropriate means of determining how their lives can contribute to a fuller development of humankind, offering as it were, an appropriate avenue for "heroism" (Becker 1971). Ryan and colleagues (1991, 1997) argue that creating individuals who are cooperative and not alienated from society requires that socializing agents provide opportunities for autonomy within a supportive context of belongingness. Behavior that is experienced as occurring either under social pressures or from internal forces of incompletely integrated, or introjected, societal values will not be experienced as autonomous. Successful internalization of societal values such that they result in intrinsically motivated behavior is thus ideal for effective functioning of a society.

*Reactance Theory.* A second theory that has implications for understanding decreased interest in a task after receiving controlling rewards is reactance theory. From this perspective, loss of autonomy, of ability to choose to engage in some action, means loss of freedom. Thus, when individuals begin to anticipate a reward for behavior that was previously driven solely by intrinsic motives, they may in some contexts come to see that reward as an attempt to impose on them some type of action-outcome contingency. They may feel a sense that the attempt is intended to restrict their

ability to freely engage or not engage in the activity at will. This loss of expected freedom induces a state of psychological tension known as reactance (Brehm 1993; Brehm and Brehm 1981). Reactance involves the experience of active, negative emotional states such as frustration and anger, and results in an individual's engaging in active attempts to regain the lost freedoms. From this perspective, reduced interest in a task is not a result of declining intrinsic motivation per se, but is merely an expression of a general motivation to regain lost freedom of choice. It is important to note that defiant rebellion against controlling influences does not necessarily mean regaining freedom. Automatically withdrawing effort from activities that were previously pleasing means that one's choices are being controlled in an oppositional way by those external influences (Deci and Ryan 1985).

#### **AUTONOMY IN THE CONTEXT OF CONTROL NEEDS**

The recent conception of autonomy proposed by Deci and Ryan (1985) is similar in many ways to an earlier notion of the motivation to have effective interaction with one's environment, effectance motivation (White 1959). White proposed that a variety of behaviors of the active organism, such as play, exploration, and active curiosity, can best be explained by proposing a need to engage in activities that lead to feelings of efficacy, and that allow the developing organism to become competent. Thus the notion of an effectance or mastery motive is closely related to a motive for autonomy in that both reflect a force that directs the organism toward increasing competence in managing the self and personal goals, or in managing one's environment. Indeed, the need for autonomy, mastery, and a third motive, power (Winter 1973) can all be viewed as reflections of an organism's need for personal control over various domains of their life (Marsh et al. 1998). Personal control involves a contingency between actions and outcomes, a sense of having in one's repertoire actions that can increase the likelihood of getting desired outcomes. Thus needs for power reflect a need for control over the social environment, and needs for mastery reflect a need for control over the physical and nonsocial environment; autonomy reflects intrapersonal control needs.

From such a control motivation perspective, autonomy is reflected in controlling the self, regulating emotional responses, and making decisions regarding one's actions. A key assumption of such an approach is that autonomy involves beliefs about contingency— notions about oneself as a causal force, consistent with deCharms's (1968) notion of origins versus pawns. A particularly distinguishing feature of need for autonomy as a control motive is that it becomes defined in large part as reflecting a need to resist others' controlling influence. From Deci and Ryan's (1985) perspective, however, autonomy is about the choice over what action-outcome contingencies to explore, *not* about control per se. From their perspective, need for self-determination is quite independent of control needs; it is in a sense prior to control. Thus, if control needs are about whether the individual can have a shot at winning a game through his or her actions, autonomy is about having the choice of whether to enter the game—to decide whether to try to explore the contingencies, regardless of whether or not there are contingencies once the choice is made. Such need for autonomy is a proactive, ever-present force, not a reactive force that emerges in response to *loss* of intrapersonal control. In this way autonomy has an element much like early notions of mastery needs; effectance motivation was explicitly hypothesized to not be a deficit-based need and thus was viewed as dissimilar to tissue-deficit drives like hunger (White 1959).

However, the research on the effects of reward on intrinsic motivation makes it clear how sensitive the experience of autonomy is to deficits of autonomy. Autonomy is uniquely about freedom from control, in that having the freedom to explore contingencies in the world, to attend to intrinsic experiences of action rather than having to attend to the controlling features of a situation, is, in a real sense, about the ability to have control over oneself. In fact, individuals who have greater needs for autonomy show somewhat greater ability to detect another's nonverbal displays in which affect and expression of dominance versus passivity are expressed; needs for power, in contrast, are uncorrelated with such skill (Marsh et al. 1999). Perhaps the nonverbal behavior of others is more informative to an individual with greater need for autonomy because the intentions of others might offer some threat to one's own freedom. Thus,

considering autonomy within a context of control needs, and as potentially involving homeostatic processes (Pittman and Heller 1987), seems to have utility.

In general, though, understanding autonomy and its relationship to different domains of control needs, and understanding the effects of loss of autonomy, seem to be important factors not yet fully explored by social scientists. In particular, many of the cognitive processes that have been explored with loss of control in mastery domains have been less frequently explored in the domain of autonomy. Control motivation research demonstrates that individuals have strong tendencies toward biased perceptions that events are under their personal control, and that cognitive processes function to help maintain such biases. The extent to which autonomy-based action is similarly mediated by biased cognitive processes suggests important questions for future study.

Perhaps the most fundamental question not addressed in current research on autonomy harkens back to personality theorists' beliefs that autonomy motives should be understood as standing in perpetual intrapsychic conflict with the need of the individual for surrender and communion with others. If such forces operate independently of one another, separate theories for autonomy and control and for belongingness processes are reasonable. On the other hand, early theorists may be correct in the belief that the desire for autonomy, control, and independence act in a continual creative conflict with the individual's desire for social acceptance, integration, and union. From such a perspective, the tension may result in emergent phenomena not otherwise predicted by considering the needs independently of one another. Reexamining this dynamic tension may provide a useful next direction for developing a more complete picture of the impact of autonomy needs on the individual and on his or her connection with other individuals, important social groups, and society.

## REFERENCES

- Amabile, T. M. 1983 *The Social Psychology of Creativity*. New York: Springer-Verlag.
- Angyal, A. 1941 *Foundations for a Science of Personality*. New York: Commonwealth Fund.
- Bakan, D. 1966 *The Duality of Human Existence*. Chicago: Rand McNally.

- Becker, E. 1971 *The Birth and Death of Meaning: An Interdisciplinary Perspective on the Problem of Man*, 2nd ed. New York: Free Press.
- Brehm, J. W. 1993 "Control, Its Loss, and Psychological Reactance." In G. Weary, F. Gleicher, and K. L. Marsh, eds., *Control Motivation and Social Cognition*. New York: Springer-Verlag.
- Brehm, S., and J. W. Brehm 1981 *Psychological Reactance: A Theory of Freedom and Control*. New York: Academic Press.
- deCharms, R. 1968 *Personal Causation*. New York: Academic Press.
- Deci, E. L. 1971 "Effects of Externally Mediated Rewards on Intrinsic Motivation." *Journal of Personality and Social Psychology* 18:105-115.
- , and R. M. Ryan 1985 *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.
- 1987 "The Support of Autonomy and the Control of Behavior." *Journal of Personality and Social Psychology* 53:1024-1037.
- Eisenberger, R., and J. Cameron 1996 "Detrimental Effects of Reward: Reality or Myth?" *American Psychologist* 51:1153-1166.
- 1998 "Reward, Intrinsic Interest, and Creativity: New Findings." *American Psychologist* 53:676-679.
- Harackiewicz, J. M., G. Manderlink, and C. Sansone 1984 "Rewarding Pinball Wizardry: Effects of Evaluation and Cue Value on Intrinsic Interest." *Journal of Personality and Social Psychology* 47:287-300.
- Hennessey, B. A., and T. M. Amabile 1988 "Reward, Intrinsic Motivation, and Creativity." *American Psychologist* 53:674.
- Lepper, M. R. 1998 "A Whole Much Less than the Sum of its Parts." *American Psychologist* 53:675-676.
- , and D. Greene 1978 *The Hidden Costs of Rewards*. Hillsdale, N.J.: Erlbaum.
- , and R. E. Nisbett 1973 "Undermining Children's Intrinsic Interest with Extrinsic Reward: A Test of the 'Overjustification' Hypothesis." *Journal of Personality and Social Psychology* 28:129-137.
- Marsh, K. L., S. A. Nasco, D. Hilton, G. S. Bains, and W. M. Webb 1999 "Domain-Specific Control: Individual Differences in Motivations for Mastery, Power, and Autonomy." Manuscript under review.
- Pittman, T. S., and J. F. Heller 1987 "Social Motivation." *Annual Review of Psychology* 38:461-489.
- Rank, O. 1929 *The Trauma of Birth*. New York: Harcourt, Brace.
- Ryan, R. M. 1991 "The Nature of the Self in Autonomy and Relatedness." In J. Strauss and G. R. Goethals, eds., *The Self: Interdisciplinary Approaches*. New York: Springer-Verlag.
- , J. Kuhl, and E. L. Deci 1997 "Nature and Autonomy: An Organizational View of Social and Neurobiological Aspects of Self-Regulation in Behavior and Development." *Development and Psychopathology* 9:701-728.
- Ryan, R. M., R. W. Plant, and S. O'Malley 1995 "Initial Motivations for Alcohol Treatment: Relations with Patient Characteristics, Treatment Involvement and Dropout." *Addictive Behaviors* 20:279-297.
- Sansone, C., and J. M. Harackiewicz 1998 "'Reality' is Complicated." *American Psychologist* 53:673.
- White, R. W. 1959 "Motivation Reconsidered: The Concept of Competence." *Psychological Review* 66:297-333.
- Winter, D. G. 1973 *The Power Motive*. New York: Free Press.

KERRY L. MARSH

## PERSONAL DEPENDENCY

Personal dependency is the tendency to seek support, security, reassurance, and guidance from outside the self. The object of dependency may be another person, a social unit (e.g., family, a religious group), or a symbolic belief system from which people receive positive outcomes, such as assistance, love, and/or the attainment of personal goals. The support requested can be physical (dependency on a caregiver by infants and persons who are very old, sick, or disabled), cognitive (a student's reliance on his or her teacher), and/or emotional (reliance on another person for reassurance and love). A dependent relationship implies the existence of interpersonal bonding, commitment, involvement, obligation, and trust.

An analysis of social science literature clearly reveals that personal dependency has been conceptualized in either negative or positive terms. On the one hand, personal dependency has been equated with weakness, immaturity, and passivity, and it has been viewed as an obstacle to the development of an autonomous and mature person. On the other hand, personal dependency has been viewed as a basic human motivation, which accomplishes important adaptive functions. It seems to contribute to the process of coping with life adversities and to set the basis for the

formation of close relationships and social ties as well as for social cooperation and hierarchical social structures. In this article, I first present negative and positive conceptualizations of personal dependency. I then attempt to provide a more integrative conceptualization of dependency, in which its positive and negative aspects can coexist.

### **NEGATIVE VIEWS OF PERSONAL DEPENDENCY**

During the last century, several social science professionals have viewed personal dependency in adulthood as a sign of weakness and immaturity. For example, Bornstein (1993) defines personal dependency in terms of negatively valued beliefs and emotions. Dependency-related beliefs refer to mental representations of the self and the social world that justify the tendency to seek support from other persons in times of need. Specifically, they include the perception of the self as weak, helpless, and ineffectual, as well as the beliefs that other persons are powerful and have the ability and skills to solve life problems and to control the course and outcomes of environmental events. Dependency-related emotions include the arousal of anxiety and worry upon external demands to deal with life tasks in an independent manner as well as fears of criticism, rejection, and separation. According to Bornstein (1993), these beliefs and emotions detract from independence, reinforce dependency over the life span, and facilitate its generalization across different interpersonal situations and social settings.

The emphasis on the negative aspects of personal dependency seems to reflect Western societies' values of independence, autonomy, and mastery. It also seems to reflect the view that mature, well-adjusted persons should attempt to cope with life tasks in an autonomous and self-directed way. As a result, the tendency to turn to others for support, assistance, and reassurance in times of need can be equated with immaturity, helplessness, and powerlessness, as well as with the failure to meet cultural expectations and standards. Moreover, it can be viewed as a risk factor for psychological problems, such as depression, anxiety, alcoholism, and eating disorders, as well as for negative social phenomena, such as loss of personal identity

and blind obedience to totalitarian leaders. In extreme cases, when people suffer from persistent and severe interpersonal or occupational problems, personal dependency is considered to be a specific type of diagnosable psychological disorder.

The negative view of dependency can be traced to early writings of Freud, who argued that dependency in adulthood consists of immature and infantile forms of behaviors. In his view, a dependent orientation toward life is related to gratifying and frustrating breast-feeding experiences during the first year of life, when infants' survival completely depends on their mother's goodwill. Abrupt weaning, frustrations related to rigid and insensitive feeding schedules, and/or the failure to end the nursing period are hypothesized to result in a failure to adequately deal with conflicts regarding dependency and autonomy. This failure is reflected in the arousal of anxiety every time one is required to act in an independent manner; the experience of serious doubts about one's own ability to be an autonomous person; and the longing for the infantile, dependent relationship with the feeding mother. Problematic breast-feeding experiences are also hypothesized to lead people to believe that the solution for their problems is outside and that others can take care of them in the same way that their mothers fed them.

The problem with Freud's ideas is that they are not supported by empirical findings (see Bornstein 1993, for a review). First, studies have failed to find a coherent and meaningful pattern of association between feeding experiences in infancy and self-reports of personal dependency in adulthood. Second, there is no strong evidence that personal dependency in adulthood is associated with mouth- and food-related activities as well as with oral behaviors (e.g., thumbsucking). Third, mixed results have been found in studies assessing the association between personal dependency and psychological disorders that have an oral component, such as eating disorders, alcoholism, and tobacco addiction. However, although findings do not support Freud's premise that dependency equals immaturity and infantilism, it still continues to exert a major influence in psychological writings.

A related approach to personal dependency can be found in object relation theories, which

emphasize the crucial role that the early child-parent relationship plays in social and emotional development (Greenberg and Mitchell 1983). Like Freudian theory, these theories view events occurring during the first one or two years of life as critical for the development of personal dependency. However, unlike Freud, the proponents of object relations theories do not emphasize feeding experiences. Rather, they focus on the child's relationship with his or her parents and the failure to resolve conflicts around nurture and closeness, on the one hand, and separation and autonomy, on the other. In these terms, personal dependency reflects an infantile desire to merge with other persons and to be cared for by them. Furthermore, it is related to the search for absolute love and enmeshed relationships, the use of psychotropic drugs, and the identification with strong leaders and highly cohesive groups.

One basic hypothesis derived from object relation theories is that parenting style during infancy and childhood may be critical for an understanding of the development of offspring's personal dependency. Bornstein (1993) follows this idea and contends that overprotective and/or authoritarian parenting may create a vicious circle that increases the likelihood of offspring's dependency. Specifically, overprotective and/or authoritarian parents may prevent their children from engaging in exploratory and trial-and-error activities that promote a sense of mastery and autonomy. As a result, these children may perceive themselves as weak and may tend to seek others' help when confronted with life tasks. This support-seeking tendency may elicit others' helping behaviors, which, in turn, may further reinforce personal dependency. Along this reasoning, personal dependency would be overtly expressed mainly when the other (e.g., a parent, a teacher) is perceived as a powerful authority, and may underlie the blind pursuit of strong and authoritarian leaders who can offer protection and help.

The contribution of parenting style to personal dependency has also been acknowledged in social learning theories, which focus on the type of behaviors that parents reinforce throughout childhood and adolescence (Rotter 1982). Specifically, children whose parents positively reinforce passive and dependent behaviors are hypothesized to become dependent adult persons. These children

may learn that dependent behaviors are adequate instrumental means for obtaining positive outcomes (e.g., love, esteem) from parents and that active and autonomous behaviors should be inhibited if they want to maintain a good relationship with parents. Social learning theories also hypothesize that this learning would be generalized across situations, leading people to behave in a dependent manner in a wide variety of social contexts. Like object relation theories, social learning theories suggest that overprotective and/or authoritarian parenting would lead to offspring's dependency, because such parents may reinforce passive and dependent behaviors. However, social learning theories differ from object relation theories in that they view personal dependency as an active instrumental means for obtaining positive outcomes from authority figures.

In emphasizing parental reinforcement and learning processes, social learning theories also highlight the role that the learning of gender roles—the learning of cultural norms and expectations regarding feminine and masculine traits and behaviors—may play in the development of personal dependency (Mischell 1970). This is particularly noted in Western cultures, where the equation between dependency and femininity may lead parents to reinforce dependent behaviors among girls and to punish these behaviors in boys. In this way, parents may provide differential reinforcement for boys and girls, leading children to meet cultural expectations concerning the expression of dependency. However, one should recall that the “dependency = femininity” equation reflects Western societies' norms and that other cultural contexts can produce different gender-role expectations and different patterns of parental reinforcement.

The negative aspects of dependency have been also emphasized by interpersonal theories of personality (e.g., Leary 1957). In these theories, dependency has been equated with personal characteristics of weaknesses, passivity, and helplessness. Moreover, it has been associated with suggestibility, compliance, and the adherence to others' beliefs and interests as a means for obtaining love, approval, and support. Interpersonal theories also emphasize that dependency is associated with fear of negative evaluation and test anxiety, which, in turn, may further exacerbate dependency. These

worries may divert cognitive resources away from active, self-directed behaviors and may lead people to escape or avoid any situation that demands autonomy and independence.

Like psychoanalytic theories, interpersonal theories also suggest that personal dependency in adulthood increases the risk for psychological problems. In this context, studies have examined the association between self-reports of dependency, on the one hand, and depression, anxiety, substance abuse, and eating disorders, on the other. However, results are mixed. Whereas some studies have indeed found a positive association between self-reports of dependency and psychological disorders, other studies have failed to find such an association. Moreover, there is evidence that dependency may be an outcome rather than a cause of psychological disorders (Bornstein 1993). For examples, there are studies showing that the onset of depression and the resulting increase in helplessness and powerlessness feelings promote passive and dependent behavior. Accordingly, longitudinal studies have shown that the onset of alcoholism is followed by an increase in dependent behavior.

The above findings put into question the view that dependency is a negative personal characteristic. Rather, they may imply that dependent behavior reflects a cry for help, reassurance, and support in dealing with emotional and social problems. As such, dependent behavior may be an adaptive means for overcoming life difficulties with the help of others. In these cases, avoiding support seeking and maintaining a facade of autonomy and self-reliance may have detrimental consequences for people who really need help and support to overcome their predicaments.

#### **POSITIVE VIEWS OF PERSONAL DEPENDENCY**

The positive sides of personal dependency have been emphasized by psychological and sociological theories that focus on the development and stability of social relations, contracts, and organizations. In terms of Kelley's interdependence theory (1979)—the most influential theory on interpersonal relationships—the reliance of person on a partner for the satisfaction of his or her own needs is a basic requirement for the development and

stability of close and positive relationships. In fact, Kelley suggests that there can be no stable relationship when a person is unable or unwilling to rely on his or her partner. Dependency is the psychological glue that maintains a close relationship over time. Moreover, it seems to prevent people from moving toward alternative relationships.

Along the above reasoning, personal dependency is considered to be one core component of the experience of love and to be associated with other components of this phenomenon, such as commitment, intimacy, and trust. For example, theory and research have emphasized that a person's dependency on a romantic partner precedes commitment to the relationship—precedes the intention to maintain the relationship in the future. Accordingly, dependency seems to play a critical role in the development of closeness and intimacy within love relationships. People who can rely on their partners for need satisfaction have been found to develop a sense of trust and confidence in their partner's goodwill, which, in turn, may facilitate the taking of risks in the relationship and the sharing of intimate feelings and thoughts with the partner. As a result, people may strengthen their willingness to initiate and maintain intimate patterns of communication while developing a sense of togetherness with the partner.

Personal dependency has been also viewed to play a positive role in group dynamics as well as in the stability of hierarchical relationships within social organizations. In this context, personal dependency seems to precede a person's willingness to participate in group activities and to collaborate with others in teamwork, mainly when others are at a higher rank in the organizational hierarchy. People who are unable or unwilling to rely on others for goal achievement and task completion may be reluctant to participate in teamwork, may prefer to work alone, and may react negatively to authority figures who threaten their independence. Only when people feel that they need support and guidance from others may they be positively oriented toward teamwork and authority figures. In fact, there is evidence that the level of dependency members of a group feel toward each other is a sign of group cohesion and a positive predictor of group effectiveness. In a broader perspective, personal dependency seems to be a prerequisite for the phylogenetic development of social ties and structures.



Theory and research have also highlighted the association between dependency and positive personal characteristics. First, research has shown that the equation between dependency and passivity is not always true. In fact, there are many situations in which dependency leads to active forms of behavior, such as attempts to outperform others in order to attract the attention of an authority figure. Second, positive associations have been found between dependency and sensitivity to interpersonal cues—the ability to decode and understand others’ messages. Third, there is evidence that personal dependency is related to health-promotion behaviors, such as seeking of medical treatments and compliance with them, as well as to relationship-enhancing traits, such as sociability, self-disclosure, and cooperative orientation in social interactions.

Studies in the field of stress and coping have shown that people who seek support and guidance from others in times of need possess positive mental representations of the self and the social world. In these studies, the tendency to rely on support seeking as a coping strategy has been found to be related to the perception of the self as capable of coping with stressful events and environmental demands. This tendency has been also found to be associated with optimistic beliefs about distress management as well as about others’ ability and willingness to provide support and guidance.

Personal dependency has also been found to result in adaptive behavior. The tendency to seek support from others has been conceptualized as a basic behavioral strategy that people use in coping with life adversities. In Lazarus and Folkman’s model (1984), the most influential theory in the field of stress and coping, support seeking is considered to be one of the most frequently used coping strategies. More important, research has consistently found that reliance on this coping strategy leads to positive psychological and social outcomes. Specifically, people who cope with life problems by seeking support from others tend to feel better, to experience less distress, and to show fewer problems in social functioning than people who rely only on themselves. Moreover, the belief that one can depend on others in times of need has been found to facilitate social adjustment.

Studies in the field of stress and coping have also shown that a reluctance to seek support in

times of need has negative health and adjustment outcomes. People who rely exclusively on themselves in dealing with intense and persistent stressful events have been found to experience, in the long run, high levels of distress and serious problems in physical health and social functioning. These findings may reflect the fragile nature of the “pseudo-safe” world of a person who believes that he or she can deal alone with all life problems and does not need the help of others. It seems that the lack of others to depend upon leaves this person vulnerable and helpless in face of stressful events.

The adaptive advantage of personal dependency has been particularly emphasized in Bowlby’s attachment theory (1969). In his terms, human infants are born with a prewired repertoire of behaviors aimed at maintaining proximity to other persons and seeking their support in times of need. These behaviors seem to reflect a basic human motivation and to accomplish a crucial adaptive function—to guarantee the survival of the helpless infant by eliciting helping behaviors in parents. According to Bowlby, this motivation persists over the entire life span, even among mature and autonomous adults. In fact, Bowlby does not view dependency as a sign of immaturity and infantilism, but as a healthy motivation that facilitates the process of coping with life problems and the development of positive social ties.

According to Bowlby, the tendency to rely on others in times of need is an inborn affect regulation device, which is automatically activated upon the experience of distress. In these cases, other persons function as a “haven of safety” to which people can retreat for comfort and reassurance and as a “secure base” from which they can develop their unique personalities in a loving and approving atmosphere. As a result, the overt expression of dependency needs and behaviors may have beneficial effects on the process of distress management as well as on the individual’s psychological wellbeing.

Bowlby also proposes that actual experiences related to the expression of dependency needs exert tremendous influence on social and emotional development. On the one hand, interactions with significant others who are responsive to one’s dependency needs may lead to the experience of more and longer episodes of positive affect and the development of positive feelings

toward the world and the self. On the other hand, interactions with rejecting others may elicit chronic distress, serious doubts about others' intentions, and problems around dependency-autonomy themes.

Along the above reasoning, actual experiences related to the expression of dependency needs may shape the way people cope with life adversities. Specifically, interactions with significant others who are responsive to one's dependency needs set the basis for the construction of effective distress management strategies. People may find out that acknowledgment and display of distress elicits positive responses from significant others. They may also learn that they are capable of eliciting helping responses from others and that support seeking is an effective way of coping. In this way, the satisfaction of dependency needs would lead people to regulate affect by overtly expressing distress and engaging in active seeking of support.

Interestingly, Bowlby does not conceptualize dependency and autonomy as antagonistic motivations. Rather, he suggests that the satisfaction of dependency needs might facilitate the development of an autonomous person. In his terms, autonomous activities in infancy are activated when a caregiver satisfies infants' dependency needs and when he or she is perceived as a "secure base" to which infants can retreat in case of danger. Thus, infants who can use the caregiver as a "secure base" can show a balance between dependency and autonomy. They can move away from the caregiver without being anxious about his or her availability, can return to him or her when danger arises, and can recommence autonomous activities (e.g., play, exploration) as the proximity to the caregiver is reestablished. Overall, the expression of dependency needs does not necessarily mean that the person cannot engage in autonomous behavior. Rather, the expression and satisfaction of these needs seems to set the basis for a confident and pleasurable development of autonomy.

#### **AN INTEGRATIVE VIEW OF PERSONAL DEPENDENCY**

The above-reviewed literature clearly indicates that personal dependency may have both negative and positive implications for the individual and the society. The main question here is whether and how the negative and positive aspects of personal dependency can coexist. In other words, one should

attempt to present a more integrative view of personal dependency, which can explain how this phenomenon may be an adaptive device and at the same time may lead to maladjustment and social problems.

As noted earlier, Bowlby argues that personal dependency is a healthy human behavior and that the tendency to seek support from others has positive psychological and social effects. However, this healthy pattern of behavior can become dysfunctional upon the recurrent frustration of one's cry for help. According to Bowlby, when people perceive significant others as nonresponsive to their dependency needs, they may learn that support seeking fails to bring the expected relief and that other defensive strategies should be developed. One of these strategies is a "fight" response, by which people attempt to compulsively elicit others' love and support through controlling and clinging responses. That is, these persons seem to develop an overly dependent pattern of behavior. The problem with this strategy is that it may create an anxious focus around social relationships; doubts about one's autonomy; anxious demands for proximity; fears of separation, rejection, and criticism; and inability to leave frustrating social interactions. As can be seen, this defensive strategy seems to result in all the negative consequences that psychoanalytic and social learning theories have linked to personal dependency.

Along the above reasoning, personal dependency per se is not a pathological sign of weakness and immaturity. Only when healthy dependency needs are frustrated and a person adopts an overly dependent defensive strategy is he or she caught in a vicious circle of anxiety, helplessness, maladjustment, and increasing dependency. In this view, psychoanalytic and social learning theories have in fact dealt with overdependency rather than with the normal expression of dependency needs. Moreover, most of the findings relating dependency to emotional and social problems have been obtained from self-report questionnaires that tap overdependency rather than the tendency to seek support from others in times of need. These questionnaires include items about fears of rejection, separation, and criticism; need for approval; and doubts about autonomy.

The above reasoning receives strong support in studies that focus on the interpersonal and

intrapersonal correlates of adult attachment styles. Shaver and Hazan (1993) define attachment styles as stable patterns of beliefs, emotions, and behaviors in social relationships, and divide them into three types: secure, avoidant, and preoccupied. The "secure" style is defined by feelings of comfort with dependent relationships as well as by the tendency to seek support from others. The "avoidant" style is defined by reluctance to depend on others, avoidance of close relationships, and an overemphasis on autonomy and self-reliance. The "preoccupied" style is defined by compulsive attempts to minimize distance from others via clinging behaviors and fears of rejection, separation, and criticism. Overall, these three styles can be organized along a dependency continuum, with the "avoidant" style reflecting underdependency, the "secure" style the normal expression of these needs, and the "preoccupied" style overdependency.

Studies in adult attachment styles have consistently documented the adaptive advantage of the normal expression of dependency needs, as manifested in the secure style. First, securely attached persons have been found to report on more positive interactions with parents than avoidant and preoccupied persons. Second, persons addressing the secure style have been found to show less distress in times of stress than persons who address either an avoidant or a preoccupied style. Third, secure persons have been found to have more positive and stable close relationships and to be more positively involved in social activities than avoidant and preoccupied persons. That is, the overt expression of dependency needs is related to positive personal and social outcomes. In contrast, either the inhibition of these needs or the adoption of an overly dependent pattern of behavior seems to have detrimental effects on psychological and social functioning.

In conclusion, the above line of thinking and findings emphasize the balance between dependency and autonomy needs. The overt expression of dependency needs in adulthood does not necessarily compete with or inhibit autonomy needs. Rather, as Bowlby suggests, the satisfaction of dependency needs may facilitate the expression of autonomy needs. In these cases, people can freely move back and forth between dependency and autonomy, can flexibly accommodate to social demands, and then can maintain an adequate level

of social adjustment and functioning. Only when this balance is disrupted, either among persons who overemphasize autonomy or among those who behave in an overly dependent manner, is adjustment at risk. In these cases, people may be unable and/or unwilling to flexibly adjust to the social world, in which they should act autonomously at same times and rely on others at other times.

Of course, a person's balance between dependency and autonomy needs depends not only on psychological factors. Rather, it also results from cultural and societal norms, values, and expectations. In fact, this balance would be different in societies that emphasize collectivist values (e.g., acceptance of social roles, family maintenance, and security) and in societies that emphasize individualistic values (e.g., personal achievement, mastery, freedom, and autonomy). Theory and research should attempt to provide a better understanding of the interface between cultural norms and values, a person's history of social interactions, and his or her expression of personal dependency. Moreover, they should throw away anachronistic and simplistic views of personal dependency and adopt a more integrative view of this basic human motivation.

## REFERENCES

- Bornstein, R. F. 1993 *The Dependent Personality*. New York: Guilford.
- Bowlby, J. 1969 *Attachment and Loss: Attachment*. New York: Basic Books.
- Greenberg, J. R., and Mitchell, S. A. 1983 *Object Relations in Psychoanalytic Theory*. Cambridge, Mass.: Harvard University Press.
- Kelley, H. H. 1979 *Personal Relationships: Their Structure and Processes*. Hillsdale, N.J.: Erlbaum.
- Lazarus, R. S., and S. Folkman 1984 *Stress, Appraisal, and Coping*. New York: Springer.
- Leary, T. 1957 *Interpersonal Diagnosis of Personality*. New York: Ronald Press.
- Mischell, W. 1970 "Sex Typing and Socialization." In P. H. Mussen, ed., *Carmichael's Manual of Child Psychology*, 3rd ed. New York: Wiley.
- Rotter, J. 1982 *The Development and Application of Social Learning Theory*. New York: Praeger.
- Shaver, P. R., and C. Hazan 1993 "Adult Romantic Attachment: Theory and Evidence." In D. Perlman

and W. Jones, eds., *Advances in Personal Relationships*, vol. 4. London: Jessica Kingsley.

MARIO MIKULINER

## PERSONAL RELATIONSHIPS

See Interpersonal Power; Symbolic Interaction Theory.

## PERSONALITY AND SOCIAL STRUCTURE

Three questions underlie the study of social structure and personality: What is social structure? What is personality? And, what is the relationship between the two? The history of this area and the current state of knowledge contain tremendous variability in the answers to these questions.

For example, social structure includes whole cultural configurations, social institutions such as family and the state, social stratification and class, the nature of roles, organizational structures, group dynamics, and micro-features of day-to-day interactions. Social structure also includes process and change at the group level, such as economic depression-recession, modernization, revolution, war, organizational growth and decline, human development and aging, and life-course transitions (school-to-work, retirement).

Concepts and approaches to personality also have a rich history in this area, with many variations. These include attitudes, abilities, affective and attributional styles, values, beliefs, cognitive schema, identities, aspirations, views of the self and others, and individual behaviors. The concept of personality, too, connotes both structure and process or change. Most contemporary observers would agree with a definition of personality as “regularities and consistencies in the behavior of individuals in their lives” (Snyder and Ickes 1985, p. 883). In some approaches, attitude and the self precede and determine behavior; in other approaches, people observe behavior and infer their own mental states and features of self.

Neither a simple answer nor close consensus exists among scholars on the nature of the relationship between social structure and personality, although most would agree that the relationship is

reciprocal rather than asymmetric, and modest rather than extremely strong or extremely weak (House and Mortimer 1990; Miller-Loessi 1995; Mortimer and Lorence 1995). That is, multiple areas of research provide clear evidence that variations in social structures shape components of personality, *and* that variations in personality in turn affect social structure. Humans are not completely pawns in the face of social forces, nor are they entirely independent, autonomous agents, unfettered by social influences. The study of human lives shows clear evidence of both forms.

The study of social structure and personality has its roots in the disciplines of sociology, psychology, and anthropology. Scholars whose ideas and research offer inspiration include Marx (1963), Freud (1928), Mead (1934), Lewin (1951), Gerth and Mills (1953), Inkeles and Levinson (1954), Smelser and Smelser (1963), and Turner (1956).

The focus of scholarship in the 1930s, 1940s, and early 1950s was to define the basic concepts and processes for personality, for social structure, and for the relationship between the two. This era produced and elaborated developments such as field theory, role theory, and interactionist perspectives on the self, along with concepts such as self, significant other, role taking, socialization, the authoritarian personality, modal personality, and national character.

In the 1950s, the sociological research on social structure and personality focused on macroscopic empirical studies of national character: What was national character? How did it vary? Could it be defined in terms of modal personality types? A long tradition of comparative anthropological studies of culture and personality informed these studies. For example, some of these studies considered the relationship between social class and personality, with *social class* defined as white-collar–blue-collar. *Personality* referred to some underlying continuum of “adjustment,” and the link between social class and personality occurred in socialization, in particular in child-rearing practices.

The 1960s produced major changes in the study of personality and social structure. First, the quantity of research increased significantly, concurrently with the massive growth of sociology and psychology as disciplines and with the growth of higher education. Second, research in the area

became more diffuse and more differentiated. What had been a fairly identifiable area of research scattered to subareas of scientific disciplines, such as the sociology of medicine, social stratification, small group dynamics, or attitude-behavior research. The research problems multiplied; research methods and strategies multiplied; theories and explanations multiplied; and journal outlets and books multiplied. At the same time, communication, integration, and cross-fertilization across the research fragments declined, although in recent years this may be changing. In short, during this era “social structure and personality” became an umbrella description for many different lines of investigation that were only loosely connected.

The third major change in the 1960s was a refocusing of research on social structure and personality, one that continues into the 1990s. The empirical macroscopic studies of national character, and the emphasis on holistic conceptions of culture and national character, declined. On the sociological side, the emphasis shifted to studying “*aspects* of societies in relation to *aspects* of individuals” (House 1981, p. 526). On the psychological side, a looser, multidimensional approach to personality replaced the earlier Freudian approach, which was based on a coherent dynamic system and on personality types and structures (DiRenzo 1977).

House (1981) describes this major refocusing of research in terms of three principles, which also define ideals for the investigation of personality and social structure. First, the *components* principle suggests that social structures such as roles, positions, and systems are multidimensional, and theory should specify which dimensions are important for which personality phenomena (such as stress, self-esteem, and locus of control). Second, the *proximity* principle suggests focusing first on understanding the more proximate stimuli that affect people and then mapping the causal patterns across broader levels of social structure in time and space. Third, the *psychological* principle identifies the importance of specifying the psychological processes involved when social structures and processes affect the self, personality, and attitudes. House’s three principles nicely summarize many of the recent advances in the study of social structure and personality. They also define the nature

of limitations in current knowledge, and identify research frontiers.

The contemporary landscape of research on social structure and personality in sociology is a patchwork of problems and areas. These include social stratification, work, and personality (Kohn et al. 1983); social structure and health, both physical and psychological (Mirowsky and Ross 1986); disjunctive social changes (war, economic depression) and individual adjustment (Elder 1974); role transitions and psychological changes (O’Brien 1986); variations in self-concept by structural position (Gecas and Burke 1995); human development, aging, and social change (Featherman and Lerner 1985; Alwin et al. 1991); and political and discriminatory attitudes, social institutions, and change (Kiecolt 1988), to mention just a few.

One of the most substantial and important areas of research involves the study of social stratification, work and personality, and the program of research of Kohn, Schooler, and colleagues (1983; Kohn and Slomczynski 1990; Kohn et al. 1997). The Kohn-Schooler model reflects the dominant approach in this particular area, and illustrates the major sociological approach to the study of social structure and personality. Spenner (1988a, 1988b, 1998) provides detailed review of this research. In comparison, approaches in psychology are more microscopic—in focusing on shorter intervals of time and smaller arenas of social space—and more likely to rely on experiments and lab studies, or field research versus large-scale survey research of people’s work lives and personality histories.

The Kohn-Schooler model begins with dimensions of jobs that are defined and measured as objectively as possible (versus subjective dimensions and measures of individual’s jobs). These structural imperatives of jobs include: occupational self-direction (substantive complexity of work, closeness of supervision, and routinization); job pressures (time pressure, heaviness, dirtiness, and hours worked per week); extrinsic risks and rewards (the probability of being held responsible for things outside one’s control, the risk of losing one’s job or business, job protections, and job income); and organizational location (ownership, bureaucratization, and hierarchical position). The three basic dimensions of personality in this research include intellectual flexibility, self-directedness of orientation, and sense of well-being or distress.

Among the subdimensions of these organizing dimensions are authoritarian conservatism, personally responsible standards of morality, trustfulness, self-confidence, self-deprecation, fatalism, anxiety, and idea conformity.

The type of analysis used in the Kohn-Schooler research estimates the lagged and contemporaneous reciprocal relationships between conditions of work and dimensions of personality in nonexperimental, panel, survey data. The major data come from a national sample of over 3,000 persons, representative of the male, full-time labor force, age 16 and over in 1974. About one-third of these men were reinterviewed about ten years later, with measures being taken of work conditions and personality at both points in time. Most of the studies of women in this tradition refer to wives of men in the sample. In a series of structural equation model analyses that adjust for measurement error in dimensions of jobs and personality, the authors document an intricate pattern of lagged (over time) and contemporaneous selection and socialization effects. *Selection effects* refers to the effects of personality on work and social structure; *socialization effects* refers to the effects of work (social structure) on self and personality. Most of the effects of personality on work are lagged, as workers appear to select jobs of a given type depending on measured aspects of their personality, or to slowly mold jobs to match their personalities. Conversely, the effects of jobs on personalities appear to be somewhat larger and to involve both contemporaneous and lagged effects. The largest relationships center on components of occupational self-direction, in particular, on substantive complexity of work. For example, substantive complexity of work environments increases intellectual flexibility for men by an amount that is one-fourth as great as the effect of intellectual flexibility a decade earlier, net of controls for other variables and confounding influences.

Kohn, Schooler, and colleagues interpret their findings with a "learning-generalization" explanation. In it, people learn from their jobs and generalize the lessons to spheres of their lives away from the job. Rather than using alternate psychological mechanisms such as displacement or compensation, the structural imperatives of jobs affect a worker's values; orientations to self, children, and society; and cognitive functioning. They do this

primarily through a direct process of learning from the job and generalization of what has been learned to off-job realities. The collected research shows that these generalizations appear to hold under a broad range of controls for spuriousness, alternate explanations, and extensions. The extensions include men's and women's work lives, self-direction in leisure activities, housework, and educational domains, as well as a number of replications of the basic model including careful comparisons with samples from Poland and Japan, and more recently from the Ukraine (Kohn et al. 1997).

Similar summaries exist for many other areas of research in social structure and personality, but this line of research has been one of the most important. The limitations of the Kohn-Schooler program of research illustrate some of the frontiers facing research on work and personality. First, are these conditions of work the most important dimensions of social structure? Do they combine and exert their effects in a more complicated recipe? Are there other features of context that should be considered? Second, are these the appropriate dimensions and combinations or personality? Are there left-out dimensions or other larger meta- or organizing dimensions of personality, such as flexibility-rigidity or general affectivity (Spencer 1988a) or processual dimensions of personality, that might be more important?

Third, there are many alternate explanations that replace or extend the learning-generalization explanation for how jobs and personality reciprocally relate (for review, see O'Brien 1986). They include: (1) fit hypotheses, in which the quality of the match between dimensions of personality and dimensions of social structure determines the effects of the person on the job and vice versa; (2) needs and expectancy explanations, in which additional layers of cognitive weighting, interpretation, and processing mediate the relationships among job attributes, personality dimensions, and work attitude outcomes; (3) buffering and mediational hypotheses, in which the effects of social structure on personality (or vice versa) are accentuated or damped for certain extreme combinations of work conditions, or outside influences such as social support buffer the effects of social structure on personality; and (4) social information processing and attributional explanations, which posit additional perceptual or judgmental, evaluational or

choice, or attributional processes that affect job-attitude and attitude-behavior linkages.

At more microscopic levels (shorter time intervals such as seconds or minutes, and smaller domains of social space, such as intrapsychic, or face-to-face interactions) the challenge for research on social structure and personality is to discover the meanings and processes that underlie longer-term, larger-scale correlations between the two. This challenge applies not only to how job or social structure affects personality, but also to how a domain of personality selects a worker into an occupational or another role, or serves as a catalyst for human agency and leads to attempts by people to modify their roles and circumstances. For example, if learning generalization operates as hypothesized, what does that mean? Is the learning part of the process as straightforward as textbook images of reinforcement psychology and social learning theory imply it is? Survey research designs, the dominant methodology, typically assume and rarely observe, specify, or test the social-psychological and psychological concomitants of learning generalization. What are the associated perceptual, affective, cognitive, and behavioral concomitants of learning generalization? What are the supporting and disconfirming attribution patterns and mediations? Or is the learning process below the level of cognitive operations and attributional web of inferences that people use to make sense of their world? Experimental and observational design may be more informative than the typical survey research approach. The field understands many of the ingredients but not the specific recipe.

At a mezzoscopic level—careers, the human life span, organizations, and other institutional settings and mechanisms—the challenge to research on social structure and personality is to put the snapshots of relationships in motion and understand the dynamics over longer periods of time. For example, in the Kohn-Schooler model, how are its findings nested in adult development, and how is adult development affected by the dynamics implied in this model? Age and developmental variations have received only limited attention in the Kohn-Schooler approach. Further, much of our knowledge about relationships between various social structures and personality assumes a system in equilibrium (for example, a single coefficient capturing an effect over five or ten years). We know much less about the dynamics of social

structure–personality relationships, including estimation of trajectories, threshold effects, and rates of change. Here too, different types of research and data designs will be required to advance the state of knowledge.

Finally, at a larger, macroscopic level—encompassing decades and centuries, and whole institutional spheres and societies—the challenges confronting research in social structure and personality are multiple. They include discovering the larger sociohistorical, psychological, and biological contexts and processes in which social structure–personality relationships are embedded, and then mapping and tracing the lines of influence across levels. For example, many research streams are exclusively national or subnational, in terms of generalizations. The Kohn-Schooler approach, with systematic studies in the United States, Poland, Japan, and the Ukraine, is a notable exception, but even here we are still early on in understanding how the models vary in comparative studies of national and subnational contexts. Further, the state of knowledge is young in our understanding of how historical variations in the content and composition of work, the labor process, and the family, or in organizational form and practice altered relationships between social structure and personality. This larger challenge also includes discovering how long-term variations in human personality feed back on long-term variations in social structures, shaping history and defining what is possible for the evolution of social forms and processes.

#### REFERENCES

- Alwin, D. F., R. L. Cohen, and T. M. Newcomb 1991 *Aging, Personality and Social Change: Attitude Persistence and Change over the Life-Span*. Madison: University of Wisconsin Press.
- DiRenzo, Gordon J. 1977 "Socialization, Personality and Social Systems." In A. Inkeles, J. Coleman, and N. Smelser, eds., *Annual Review of Sociology*, vol. 3. Palo Alto, Calif.: Annual Reviews.
- Elder, Glen H., Jr. 1974 *Children of the Great Depression*. Chicago: University of Chicago Press.
- Featherman, David L., and Richard M. Lerner 1985 "Ontogeneses and Sociogenesis: Problematics for Theory and Research about Development and Socialization across the Life-Span." *American Sociological Review* 50:659–676.

- Freud, Sigmund 1928 *The Basic Writings of Sigmund Freud*. Edited and translated by A. A. Brill. New York: Random House.
- Gecas, Victor, and Peter Burke 1995 "Self and Identity." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Gerth, Hans, and C. Wright Mills 1953 *Character and Social Structure*. New York: Harcourt.
- House, James S. 1981 "Social Structure and Personality." In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- , and Jeylan T. Mortimer 1990 "Social Structure and the Individual: Emerging Themes and New Directions." *Social Psychology Quarterly* 53:71–80.
- Inkeles, Alex, and Daniel Levinson 1954 "National Character: The Study of Modal Personality and Social Systems." In G. Lindzey, ed., *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley.
- Kiecolt, K. Jill 1988. "Recent Developments in Attitudes and Social Structure." In W. R. Scott, and J. Blake, eds., *Annual Review of Sociology*, vol. 14. Palo Alto, Calif.: Annual Reviews.
- Kohn, Melvin L., and Carmie Schooler, with the collaboration of J. Miller, K. Miller, C. Schoenbach, and R. Schoenberg 1983 *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, N.J.: Ablex.
- Kohn, Melvin L., and Kazimierz M. Slomczynski 1990 *Social Structure and Self-Direction: A Comparative Analysis of the United States and Poland*. Cambridge, Mass.: Basil Blackwell.
- , Kazimierz M. Slomczynski, Krystyna Janicka, Valeri Khmelko, Bogdan W. Mach, Vladimir Paniotto, Wojciech Zaborowski, Roberto Gutierrez, and Cory Heyman 1997 "Social Structure and Personality under Conditions of Radical Social Change: A Comparative Analysis of Poland and Ukraine." *American Sociological Review* 62:614–638.
- Lewin, Kurt 1951 *Field Theory in Social Science*. New York: Harper.
- Marx, Karl 1963 *Karl Marx: Early Writings*. Edited and translated by T. B. Bottomore. New York: McGraw-Hill.
- Mead, George H. 1934 *Mind, Self and Society*. Chicago: University of Chicago Press.
- Miller-Loessi, Karen 1995 "Comparative Social Psychology: Cross-Cultural and Cross-National." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Mirowsky, John, and Catherine E. Ross 1986 "Social Patterns of Distress." In R. Turner and J. Short, eds., *Annual Review of Sociology*, vol. 12. Palo Alto, Calif.: Annual Reviews.
- Mortimer, Jeylan, and John Lorence 1995 "Social Psychology of Work." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- O'Brien, Gordon E. 1986 *Psychology of Work and Unemployment*. New York: Wiley.
- Smelser, Neil J., and William T. Smelser 1963 *Personality and Social Systems*. New York: Wiley.
- Snyder, Mark, and William Ickes 1985 "Personality and Social Behavior." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology, Special Fields and Applications*, vol. II. New York: Random House.
- Spenner, Kenneth I. 1988a "Social Stratification, Work and Personality." In W. R. Scott and J. Blake, eds., *Annual Review of Sociology*, vol. 14. Palo Alto, Calif.: Annual Reviews.
- 1988b "Occupations, Work Settings and the Course of Adult Development: Tracing the Implications of Select Historical Changes." In P. B. Baltes, D. L. Featherman, and R. M. Lerner, eds., *Life-Span Development and Behavior*, vol. 9. Hillsdale, N.J.: Lawrence Erlbaum.
- 1998 "Reflections on a 30-Year Career of Research on Work and Personality by Melvin Kohn and Colleagues." *Sociological Forum* 13:169–181.
- Turner, Ralph. H. 1956 "Role-taking, Role Standpoint, and Reference Group Behavior." *American Journal of Sociology* 61:316–328.

KENNETH I. SPENNER

## PERSONALITY MEASUREMENT

### WHAT IS PERSONALITY?

"Personality" is an ambiguous term derived from the natural language, and not necessarily a scientific concept. Consensus among interested scientists as to its precise meaning has been fairly modest. One major cleavage is between views of personality as core dynamic processes inherent in all people



and views that emphasize characteristics on which individuals differ. Because the second sort of view defines personality in a way as to make it far more conducive to measurement, this article focuses on personality as certain potentially measurable characteristics of individuals.

Which measurable characteristics? There are many kinds of characteristics, and useful ways of categorizing characteristics have been developed (Norman 1967; Angleitner et al. 1990). Characteristics that fall into many of the categories do not fit common definitions of personality. Descriptors of physical characteristics (e.g., short, muscular) lack sufficient reference to psychological (behavioral, affective, cognitive) features. Descriptors invoking social roles (e.g., motherly, professional) and social effects (e.g., famous, neglected) involve social-contextualization and relativization too heavily to give inferences about an individual's personality attributes. Descriptors of emotions (e.g., elated, afraid) and many motivational and intentional states (e.g., hungry, reluctant, inspired) are too prone to reference relatively transient characteristics. And some descriptors (e.g., awful, impressive) are so purely evaluative that they provide insufficient specificity with respect to psychological features.

Among descriptors that refer to presumably more internal and enduring psychological attributes, three categories stand out. Abilities or talents (e.g., skillful, creative, athletic) refer to maximum rather than typical levels of performance on tasks. Beliefs and attitudes (e.g., religious, racist, environmentalist) concern affectively tinged habits of mind pertaining to specific objects and concepts. Although personality models have frequently contained some ability- and attitude-related content, at their core are traits (e.g., daring, patient) that are more directly related to typical behavioral patterns. Because they are expressed behaviorally, enduring motivational patterns (e.g., need for achievement) might also be easily fit within definitions of personality that emphasize typical behavior patterns. "Temperament" usually denotes the more clearly inborn and genetically derived aspects of personality, whereas "character" is often used to denote acquired moral qualities. But these terms are otherwise synonymous with personality, which can be defined as consistencies in patterns of behavior—where behavior is

defined broadly to include affect, cognition, and motivation—on which individuals differ. Overall, personality is a lay concept of sufficient importance and usefulness to have been taken up and refined by scientists.

In natural languages, personality descriptors are alternately represented as adjectives (e.g., adventurous), attribute nouns (e.g., adventurousness), or type nouns (e.g., adventurer). Because adjectives differentiate properties, personality adjectives are inherently central to personality description, although some languages lack an adjective class and carry on this adjective function in other ways (Dixon 1977). Psychologists move easily between adjectives and attribute-noun characterizations of the same trait (as with extraverted and extraversion): Either form suggests properties that exist in varying degrees. Type-noun characterizations, in contrast, imply a categorization—one either is an extravert or one is not—which in turn suggests the assumption of a bimodal frequency distribution of individuals on the trait. Such bimodal distributions appear to be rare. Although type-noun characterizations of personality traits have great popular appeal, their use by academic psychologists has become quite limited. There has been a search for categorical taxons underlying the traits related to certain mental disorders (e.g., Meehl 1995), but this task is not easy. Recent correlational studies indicate that symptoms of mental disorders, and of personality disorders in particular, are continuously distributed in the population, and have substantial overlap with measures of various personality traits (Costa and Widiger 1994).

In the last three decades, many psychologists have addressed the fundamental issue of whether personality traits are real, or whether they exist only in the eye of the beholder. Mischel's early review (1968) suggested that personality measures were at best modest predictors of relevant actual behaviors. However, a consensus has emerged that when criterion behaviors are aggregated, across time or across situations, personality measures can become quite highly predictive (Kenrick and Funder 1988). This follows, of course, from definitions of personality offered above: Personality does not denote single behaviors but rather consistent patterns in multiple behaviors. Moreover, studies involving twins and adoptees have repeatedly indicated that personality-trait scores are partially (as much as 50 percent) heritable, implying biological

underpinnings. Such behavior-genetic findings have stimulated the development of models that set forth biological explanations for personality variation, attempting to delineate phenotypic constructs so that they correspond directly to known biological mechanisms (e.g., Gray 1987; Rothbart et al. 1994; Zuckerman 1995). Nonetheless, studies with twins and adoptees also indicate that differences in experience, mainly of the type *not* shared by family members, have a profound effect on personality characteristics (Plomin 1990). There are also some indications of important effects of culture.

Although personality traits reflect behavioral consistencies across situations, a variety of research findings indicate that some situations facilitate expression of personality traits more than do other situations (Caspi and Moffitt 1993). Highly structured or ritualized social settings (e.g., a funeral home, a lecture hall) tend to attenuate the expression of personality differences, whereas relatively unstructured settings (e.g., a nightclub, a playground) bring out personality differences. This pattern has three important consequences for personality measurement. First, to the extent that a society's social milieu are age-stratified, one might expect across-time continuity in apparent traits to be somewhat "heterotypic"—leading to different surface characteristics at different ages. Second, personality characteristics are best assessed by placing the individual in a relatively unstructured situation in which responses are relatively unconstrained by social norms; therefore, the stimuli on personality measures should not have correct responses. Third, to the extent that a cultural milieu is highly structured and ritualized, one might expect to see less emphasis on personality differences than would be found in relatively individualistic cultural milieus (Miller 1984).

#### **PERSONALITY MEASUREMENT CAPTURES CONSISTENCIES**

Measurement can be defined as a set of rules for assigning numbers to entities (such as individuals) in such a way that attributes of the entities are faithfully represented. Measurement procedures typically set up, or specify, regularized administration conditions. Thus, measurement generally involves consistency of procedure; but personality measurement has an even deeper relation to consistency.

As noted above, a consensual definition of personality would emphasize characteristics that are internally rather than externally caused, psychological rather than overtly physical, and stable and enduring rather than transient. Because of the emphasis on stable and enduring qualities, reliability—relative absence of measurement error—is of first importance in personality measurement. A reliable measure by definition registers characteristics that are consistent across time, as indicated by test-retest reliability coefficients or across situations. Internal consistency, or inter-item reliability, is an analogue of cross-situational stability. Each item represents a unique situation in either of two ways: (1) its referent content may refer to a distinct situation (e.g., "I talk a lot at parties"), or (2) simply being presented with this item as distinguished from another item (e.g., "I talk very little at parties") is a unique immediate situation for the respondent. Reliability is most often measured either by a stability coefficient, indicating the correlation between scores at one time and those at another, or by an internal consistency coefficient (e.g., coefficient alpha) that represents the average correlation between pairs of all possible split halves of the test items, although a variety of alternative reliability models are available. Until cross-time stability is established, a trait's stability is only presumed and it might be better termed an attribute, since the latter term has fewer implications as to stability.

Another criterion for the reliability of personality judgments is the extent to which different observers agree in rating a target. This is a more demanding criterion; ideally, to the degree an individual has a characteristic it will be obvious to both self and observers. However, a number of influences tend to attenuate agreement between observers. Some characteristics (e.g., sociability) are highly observable, whereas others (e.g., anxiety level) are less so. Interobserver agreement is typically reduced by having observed the target at different times or in different situations. Generally, we might expect the self to be the most privileged observer, but in certain ways the self-viewpoint can be misleading. Personality characterizations have social functions and, understandably, self-observer agreement is prone to be affected by conscious impression management and unconscious self-enhancement tendencies (Paulhus and

Reid 1991). Finally, quite independently of content, observers differ in their use of measurement scales (e.g., differential tendencies to agree or disagree, respond extremely, or use the middle option if available). Given this minefield of potential difficulties, the moderate level of interobserver agreement documented in the research literature may be quite remarkable. It makes sense to capitalize on the conjoint perspective of multiple judges: the best arbiter of the degree to which an individual can be characterized with a certain trait may be the pooled judgments of several observers well acquainted with the subject (Hofstee 1994; Kolar et al. 1996), perhaps conjoined with self-ratings. Though greater acquaintance clearly increases judges' accuracy, there is, particularly for the more observable traits, surprisingly good consensus among near strangers for the traits of a target (Borkenau and Liebler 1993).

Another important, but demanding, index of consistency is across-time stability: within a sample of individuals, the extent to which one's relative standing at time 1 correlates with that at time 2. Across-time stabilities tend to be very high for short intervals (e.g., a day or a week) but diminish with greater intervals to a more moderate level. Even across long intervals, cross-time stabilities in adulthood—particularly after age 30—for most personality traits are impressively high (Costa and McCrae 1997). It appears, however, that the further the measurement intervals reach into childhood and especially infancy, the lower the stabilities become; judgments of infant temperament may not do much better than chance in predicting judgments of later adult personality. Stabilities may be held down by the incommensurability of the contexts within which infant and adult temperaments function: It is difficult to apply many adult traits (e.g., industrious, artistic, unselfish) to infants, so that any forms of continuity would have to be heterotypic. It seems likely, however, that levels of traits often do change from childhood to adulthood. Part of this change could be genetically programmed, as a different set of genes comes on-line with greater maturity, and an initial set goes off-line; on the other hand, much may change under the influence of experience. To Wordsworth's assertion that "the child is father of the man," psychometricians offer an assent beset with caveats: "usually," "in many ways," "with definite exceptions."

## PERSONALITY DATA HAS MULTIPLE SOURCES

Data on behavior patterns are most commonly elicited from self and observers or acquaintances using standardized measures of personality traits. Scores on these structured measures are compared within a sample in a "nomothetic" manner, that is, seeking generalizations that can be applied to all individuals. Historically, the dominant position of this structured, nomothetic approach stems from the success of well-known inventories like (1) the Minnesota Multiphasic Personality Inventory (MMPI), which is actually more of a psychiatric symptom inventory than a personality inventory; (2) the California Psychological Inventory (CPI), which resembles the MMPI in numerous respects but taps rather different content, with scales labeled so as to stress the presence (or absence) of adaptive traits; and (3) the Myers-Briggs Type Indicator (MBTI), a measure based on parts of C. G. Jung's typology. The MBTI has been criticized by psychometricians, ignored by academic researchers, yet bought up by the millions in other circles. Today, these older inventories have competition from numerous new inventories that in some cases are shorter and more efficient.

Nonetheless, there are potentially useful alternatives to the questionnaire. Some embody an idiographic approach—seeking individually unique constructs that are not generalized to all people—rather than a nomothetic one. For example, in George Kelly's Role-Construct Repertory Test, each testee nominates a set of personally significant acquaintances, then derives idiographic constructs by comparing subsets of them. Such idiographic measures undoubtedly have a unique contribution: They may generate results that are more meaningful to the individual measured. But knowledge of general laws illuminates understanding of the individual case; it is possible to adapt many nomothetic measures to serve idiographic ends. Thus, improved nomothetic understanding lays the groundwork for improved idiographic understandings.

Questionnaires, whether used nomothetically or idiographically, are essentially overt and direct in their measurement approach. A trait is assessed with reference to a person's behaviors, emotions, and cognitions. Descriptions, which may be at a

rather broad level, are collected. This overt method can be highly efficient, but has a significant disadvantage: Because the descriptive content provides clues to what is being measured, respondents completing the measure could, if motivated, intentionally present an inaccurate picture. Moreover, responses can be provided thoughtlessly. Some of those who are dissimulating or not paying attention can be identified using so-called validity indexes. These indexes are computed by scouring the response pattern for various signs of less than honest and accurate responding: unusual levels of agreement with unfavorable items, disagreement with favorable items, denying common vices, claiming rare virtues, responding dissimilarly to items with similar content, or responding similarly to items with contradictory content.

Projective measures, in contrast, are covert measures of personality that are more resistant to dissimulating or inattentive responding. These measures assume a “projective hypothesis” first defined by Rorschach, Jung, and others early in this century: If an individual is presented with a vague or ambiguous stimulus, that individual’s response will be determined by habitual internal tendencies, preoccupations, and cognitive styles, rather than being affected by features of the stimulus. In a word, respondents “project” their proclivities onto the stimulus. Projective measures are potentially very sensitive receptors for personality variation. As noted above, personality differences are clearest when individuals are confronted with unstructured situations; vague and ambiguous stimuli are unstructured situations. One might simply place the individual in an unstructured situation and observe which behaviors, emotions, and thoughts ensue.

The most popular unstructured stimuli for these purposes have been inkblots (e.g., the Rorschach and Holtzman stimulus sets), sets of pictures—selected for their ambiguity—about which stories are elicited (e.g., the Thematic Apperception Test [TAT] and its derivatives), and figure drawings; in the last case the individual is presented with blank paper and asked to draw a certain object (e.g., person, house, tree). Other commonly used unstructured stimuli include incomplete sentences (e.g., “Most people \_\_.”) and single words for which an association is elicited. The raw material provided by the respondent must then be

coded and interpreted with reference to response patterns of aggregate respondents. These measures capture aspects of personality covertly and indirectly; due to the ambiguity of the test materials, respondents are unlikely to guess what is being measured.

Though attractive in theory, projective measures have been problematic in practice. The rock upon which they are prone to founder is the crucial one of reliability. A first problem is that individuals’ responses to projective stimuli are affected by social context and environment, and to a considerable degree they change from one day to another. This problem may be partially solved by gathering responses to many stimuli, preferably on multiple occasions, and looking for consistent patterns across time and across stimuli. A second problem is that observers often have low levels of agreement with regard to coding and interpreting the stimuli; that is, there is often a great deal of interobserver noise obscuring any underlying signal. There have been recent attempts to create interpretive coding schemes with better reliability. The best example is Exner’s comprehensive system for the Rorschach (1986), which integrates features of several previous Rorschach scoring systems.

The Rorschach and the TAT remain fairly popular measures in clinical settings, and continue to generate a stream of research. Presently, however, many psychologists are skeptical about the usefulness of these measures, given the laborious, complex procedures for collecting and scoring data. One problem may be the sheer volume and range of the data that such free-response methods bring in; perhaps only a fraction of these data are of any importance, and we are not yet sure which fraction deserves the most attention. Projective measures might in the future become increasingly important, to the extent that they can be made more reliable, parsimonious, and efficient.

Other forms of data can be coded using the interpretive schemes developed for projective measures. For example, politician’s speeches, reports of early memories, and virtually any autobiographical material can be analyzed in much the same way as stories elicited from the TAT. Most often, such material has been analyzed in terms of implicit motivational features (e.g., achievement, power,

intimacy), and evidence suggests that such covertly measured motivation is not substantially correlated with indexes of similar content derived from structured measures of self-attributed motivation (McClelland et al. 1989). In general, autobiographical data seems to provide information outside that provided by personality questionnaires, given that individuals seem to store schematic beliefs about traits separately from autobiographical memories (Klein and Loftus 1993). Therefore, autobiographical data could become an important part of the comprehensive assessment of individuals.

### WHICH TRAITS ARE WORTH MEASURING?

Whether the measurement method is overt or covert, another crucial issue concerns the particular traits that one ought to measure. Most commonly, this issue has been handled within a scale-construction strategy that might be called “rational”: A researcher decides which trait (i.e., construct) he or she wants to measure, creates a pool of potential items, tries them out on a sample of respondents, and perhaps iterates between data and preconceived theory to create a relatively efficient measure of the construct. A second, “empirical” strategy is in some ways a variant of the first. The researcher includes in the sample of respondents one or more criterion groups (e.g., introverts, psychopaths, artists) and determines the set of items that best differentiates each criterion group from a control sample, thus leading to a “criterion-keyed” scale for the construct (e.g., introversion, psychopathy, creative temperament). In either strategy, the researcher begins with an a priori conception of what ought to be measured, but in the empirical strategy this conception is identified with a criterion group. It is not difficult to combine rational and empirical strategies, as was done in the only major revision of the original MMPI.

Unfortunately, a field that accumulates a great host of a priori conceptions can become quite chaotic, and this was the predominant state of affairs in personality measurement until at least the 1970s, when expert compendiums on personality traits could still be organized alphabetically by trait (e.g., London and Exner 1978), as if there were no other way to order them. There were

many constructs, and it was clear that some of them were related to others, but the structure underlying the whole set of constructs was unclear. From the early decades of the twentieth century, investigators seeking an ordering framework turned to a statistical technique called “factor analysis.” *Factor analysis* is a method for reducing a large number of observed variables to a smaller number of hypothetical variables (factors), by analyzing the covariances among the observed variables and identifying redundancies in the set of variables. Factor analysis can be used to identify parsimonious sets of variables within sets of items built by any scale-construction strategy. Historically, reviews of factor analyses of various collections of personality scales (e.g., French, 1953) have not led to a consensus on a common framework (Goldberg 1972).

Significant progress on the structural problem came largely by temporarily averting attention from the a priori constructs of experts in order to study those personality conceptions of laypersons that are embedded in the natural language. As noted at the outset, personality traits are socially meaningful phenomena about which laypersons comment and generalize, and the lexicon of any language is a repository of descriptors referencing a wide variety of human characteristics. The lexical hypothesis formalizes this state of affairs into a strategy for identifying necessary features for an organizing framework, or taxonomy, of personality attributes (Goldberg 1981). This hypothesis essentially states that the more important the attribute, the more likely people are to develop a word for it. The most important attributes will then be those represented by numerous terms (often representing specific aspects of broader concepts) within one language, and by recurrence across many languages. Once descriptors are gathered, for example from a dictionary, they can be used by individuals to describe themselves or others. Factor analysis of this data, in any language, can be used to search for a few dimensions underlying numerous descriptors.

Expert personality constructs are typically based on certain aspects of the lay descriptive vocabulary, but scientists may refine and extend lay distinctions in useful ways. Therefore, one cannot obtain a *sufficient* model of personality traits by

studying lexical descriptors, but one can find *necessary* features for such a model, aspects that—based on their salience to lay observers—are too important to leave out. In this respect, lexical models of personality dimensions offer minimum-content criteria for other personality models, pointing clearly to some (but not all) of the trait concepts important enough to measure. In practice, lexical models have helped focus attention on important variables previously omitted from expert-derived models.

Lexical studies involve (1) culling descriptors from a dictionary; (2) omitting descriptors that are infrequently used or, by the consensus of multiple judges, refer to categories less relevant to personality (e.g., physical traits, temporary states); (3) aggregating the remaining descriptors (typically 300 to 400) into a questionnaire format with a multipoint (e.g., 1 to 5) rating scale; (4) administering the forms so constructed to a large (usually >400) sample of respondents for description of self, a well-acquainted peer, or sometimes both; and (5) factor-analyzing the descriptors to derive an indigenous or “emic” personality structure for that language. Such studies have been conducted in over a dozen languages, including English, German, Dutch, Italian, Spanish, Hungarian, Czech, Polish, Filipino, Korean, Turkish, and Hebrew.

Findings of these lexical studies (reviewed by Saucier 1997) show some variations, probably due to differences in sampling of subjects and variables as much as to actual differences between languages. But the most common result has been a robust structure of five independent (uncorrelated) factors, with apparent cross-language universality for the three largest of these factors: extraversion (which includes sociability, activity, and assertiveness), agreeableness (which includes warmth, generosity, humility, patience, and nonaggressiveness), and conscientiousness (which includes dependability, orderliness, and consistency). The remaining two factors (one referencing aspects of emotional stability, the other aspects of intellect, imagination, and unconventionality) are generally smaller and more variant from one study to another. Despite these partial inconsistencies between one emic structure and another, the five-factor structure, often labeled the “Big Five,” has been shown to be easily translatable into a large number of languages (McCrae et al. 1998). Moreover, the

Big Five appears to capture the structure of trait judgments about children as well as adults (Digman and Shmelyov 1996).

The Big Five has had considerable influence on personality questionnaires. For example, one prominent three-factor inventory (the NEO Personality Inventory) was revised to add the two missing factors from the Big Five, in this case agreeableness and conscientiousness (McCrae and Costa 1985). Moreover, the five factors have strong relations to the constructs measured by other prominent inventories, including the Myers-Briggs Type Indicator, the 16PF inventory, and the Personality Research Form. Four of the five factors (excluding intellect/imagination) are substantially correlated with measures of personality disorder symptoms and some mood, anxiety, and impulse control disorders cataloged in current psychiatric nosologies; a safe generalization seems to be that disorders tend to co-occur with extreme scores on personality dimensions (like the Big Five) on which there is wide variation in the general population (Costa and Widiger 1994). However, psychotic syndromes map rather poorly onto the Big Five, as do a few other clearly important individual-differences constructs, like religiousness and attractiveness (Saucier and Goldberg 1998)

Although the Big Five have obtained some degree of consensus as an organizing framework for personality characteristics, there are at least five remaining issues whose resolution might lead to a different consensual structure: (1) The generalizability of the Big Five factors to languages spoken in non-Western, nonindustrialized nations, and indeed in less complex societies, is as yet uncertain. (2) The Big Five represent very broad, global trait constructs, and groups of more specific constructs are typically more useful in prediction contexts; however, there is as yet little consensus about the particular specific subcomponents that make up each of the broad factors. (3) There may be factors in related domains, such as abilities or attitudes, that could arguably be added to the model. (4) The organization of personality variables into factors based on lexical representation might be reasonably superseded by a set of factors based on a superior rationale, for example, correspondence to main lines of biological or environmental influence. And (5) there may be constructs

well represented among natural-language descriptors that will prove to be of great importance. Personality-relevant constructs with apparently meager representation in natural language descriptors, but attracting much current research interest, include those having to do with (1) defense mechanisms (Paulhus et al. 1997), (2) coping styles (Suls et al. 1996), and (3) personal goals and strivings (Pervin 1989).

One important criterion by which personality constructs might be added to, or eliminated from, a basic descriptive model is validity. Validity should be clearly distinguished from reliability: *Reliability* concerns whether a scale is measuring anything at all, and is a prerequisite to any form of validity. *Validity* concerns the meaning of a scale score, that is, the accuracy of the inferences one can make from the scale. For a construct that demonstrates validity, there is a good argument for meaningfulness and usefulness with respect to other phenomena. Validity evidence is gathered in an ongoing process, rather than in any single study.

A prior question, of course, is whether personality measures have a substantial enough degree of validity to make them worthwhile. Mischel's early critique (1968) suggested they did not. An increasingly large volume of studies documents the many ways in which they do. For example, conscientiousness is a valuable predictor of effective job-related performance (even after the predictive value of intelligence is accounted for and removed), and has also been associated with increased longevity. Low scores on emotional stability (i.e., high scores on neuroticism) are predictive of higher rates of divorce, of male midlife crises, and of health-related complaints (though not actually greater illness). Agreeableness has associations with aspects of conflict (or absence of conflict) in close relationships, and extraversion predicts variation in a wide range of social-interaction variables.

### PROSPECTS AND FURTHER APPLICATIONS

As the foregoing review indicates, the science of personality measurement is no longer at the primitive stage represented by early taxonomies of virtues (such as those of Plato or Confucius), of physiological humors (such as that of Galen), or by

the pseudoscience of astrological signs which beckons from magazine racks. Unlike these approaches, current personality measurement is far more explicit about (1) defining personality, (2) measuring attributes in a standardized, reliable manner, (3) attending to multiple sources of data, (4) checking the validity of hypothesized models, and (5) placing the plethora of possible constructs within a parsimonious and empirically justifiable organizing framework. Nonetheless, current scientific practices are inevitably based on assumptions that are subject to being overturned. Rorer (1990) provides a conceptual review of personality assessment with attention to differences between the assumptions of mainstream and alternative paradigms.

Personality as a discipline has come to be located mainly within the larger umbrella of psychology. This is reasonable, given that personality deals with the behavior, affect, and cognitions of individuals, and with individual differences. But certain aspects of personality seem to be equally relevant to other sciences. Many of the more in-born, dispositional aspects of personality are clearly rooted in biology and genetics, and personality change may well be associated with physiological changes—as cause or effect. Moreover, humans are social as well as biological beings, and personality functions within a social context. The sociological aspects of personality need more attention: Although Goffman (1972) proposed aspects of one useful sociological theory of personality, and Bellah et al. (1985) described certain social-structural contexts that may foster narcissism, generally the relation of personality variation to its broader social context is but dimly understood.

In anthropology, one finds a rich tradition of studies of "culture and personality." Past work in this area has been hindered by use of less than adequate personality measurement models. For example, many studies used projective personality measures with insufficiently developed reliability. If improvements are made in analysis of projective data, it may be possible to take a new look at old data. Moreover, linguists and anthropologists may be uniquely well equipped to gather and judge evidence pertinent to the cross-cultural universality of basic personality dimensions, and help answer several important questions: Which dimensions of interindividual variation are not only

measurable in any culture, but derivable from the indigenous language of any culture? What is the meaning of between-culture variation in the classification of personality characteristics? To what extent is "modal personality" a viable way of differentiating societies, or of mapping cultural change, and how do temperament and social structure interact? Do cultures differ in how they organize the heterogeneity that personality variation introduces?

The fascinating generalizations offered by Mead, Benedict, and others must be considered provisional, given the probably limited range and value of data upon which they are based. If waves of advancement in personality measurement were joined to waves of advancement in other social sciences, a powerful current might ensue, which would yield a far better understanding of societies in terms of the diverse range of humans within each of these societies.

REFERENCES

- Angleitner, Alois, Fritz Ostendorf, and Oliver P. John 1990 "Towards a Taxonomy of Personality Descriptors in German: A Psycho-Lexical Study." *European Journal of Personality* (4):89-118.
- Bellah, Robert, Richard Madsen, William M. Sullivan, Ann Swidler, and Steven M. Tipton 1985 *Habits of the Heart: Individualism and Commitment in American Life*. Berkeley: University of California Press.
- Borkenau, Peter, and Anette Liebler 1993 "Consensus and Self-Other Agreement for Trait Inferences from Minimal Information." *Journal of Personality* (61):477-496.
- Caspi, Avshalom, and Terrie E. Moffitt 1993 "When Do Individual Differences Matter: A Paradoxical Theory of Personality Coherence." *Psychological Inquiry* (4):247-271.
- Costa, Paul T., Jr., and Robert R. McCrae 1997 "Longitudinal Stability of Adult Personality." In Robert Hogan, John A. Johnson, and Stephen R. Briggs, eds., *Handbook of Personality Psychology*. San Diego, Calif.: Academic Press.
- Costa, Paul T., Jr. and Thomas A. Widiger (eds.) 1994 *Personality Disorders and the Five-Factor Model of Personality*. Washington, D.C.: American Psychological Association.
- Digman, John M., and Alexander G. Shmelyov 1996 "The Structure of Temperament and Personality in Russian Children." *Journal of Personality and Social Psychology* (71):341-351.
- Dixon, R. M. W. 1977 "Where Have All the Adjectives Gone?" *Studies in Language* (1):19-80.
- Exner, John E., Jr. 1986 *The Rorschach: A Comprehensive System, Vol. 1, Basic Foundations*, 2nd ed. New York: Wiley.
- French, John W. 1953 *The Description of Personality Measurements in Terms of Rotated Factors*. Princeton, N.J.: Educational Testing Service.
- Goffman, Erving 1972 *The Presentation of Self in Everyday Life*. Woodstock, N.Y.: Overlook Press.
- Goldberg, Lewis R. 1972 "A Historical Survey of Personality Scales and Inventories." In Paul McReynolds, ed., *Advances in Psychological Assessment*, vol. 2. Palo Alto, Calif.: Science and Behavior Books.
- 1981 "Language and Individual Differences: The Search for Universals in Personality Lexicons." In Ladd Wheeler, ed., *Review of Personality and Social Psychology*, vol. 2. Beverly Hills, Calif.: Sage.
- Gray, Jeffrey A. 1987 "The Neuropsychology of Emotions and Personality." In Stephen M. Stahl, Susan D. Iverson, and Elisabeth C. Goodman, eds., *Cognitive Neurochemistry*. Oxford: Oxford University Press.
- Hofstee, Willem K. B. 1994 "Who Should Own the Definition of Personality?" *European Journal of Personality* 8:149-162.
- Kenrick, Douglas T., and David C. Funder 1988 "Profiting from Controversy: Lessons from the Personality Situation Debate." *American Psychologist* 43:23-34.
- Klein, Stanley B., and Judith Loftus 1993 "The Mental Representation of Trait and Autobiographical Knowledge about the Self." In Thomas K. Srull and Robert S. Wyer, Jr., eds., *Advances in Social Cognition, vol. 5, The Representation of Trait and Autobiographical Knowledge about the Self*. Hillsdale, N.J.: Lawrence Erlbaum.
- Kolar, David W., David C. Funder, and Randall C. Colvin 1996 "Comparing the Accuracy of Personality Judgements by the Self and Knowledgeable Others." *Journal of Personality* 64:311-337.
- London, Harvey, and John E. Exner, Jr. 1978 *Dimensions of Personality*. New York: John Wiley.
- McClelland, David C., Richard Koestner, and Joel Weinberger 1989 "How Do Self-Attributed and Implicit Motives Differ?" *Psychological Review* 96:690-702.
- McCrae, Robert R., and Paul T. Costa, Jr. 1985 "Updating Norman's Adequate Taxonomy: Intelligence and Personality Dimensions in Natural Language and in Questionnaires." *Journal of Personality and Social Psychology* 49:710-721.



- , Gregorio H. Del Pilar, Jean-Pierre Rolland, and Wayne D. Parker 1998 “Cross-Cultural Assessment of the Five-Factor Model: The Revised NEO Personality Inventory.” *Journal of Cross-Cultural Psychology* 29:171–188.
- Meehl, Paul E. 1995 “Bootstraps Taxometrics: Solving the Classification Problem in Psychopathology.” *American Psychologist* 50:266–275.
- Miller, Joan G. 1984 “Culture and the Development of Everyday Social Explanation.” *Journal of Personality and Social Psychology* 46:961–978.
- Mischel, Walter 1968 *Personality and Assessment*. New York: John Wiley.
- Norman, Warren T. 1967 *2800 Personality Trait Descriptors: Normative Operating Characteristics for a University Population*. Ann Arbor: University of Michigan Press.
- Paulhus, Delroy L., and Douglas B. Reid 1991 “Enhancement and Denial in Socially Desirable Responding.” *Journal of Personality and Social Psychology* 60:307–317.
- Paulhus, Delroy L., Bram Fridhandler, and Sean Hayes 1997 “Psychological Defense: Contemporary Theory and Research.” In Robert Hogan, John A. Johnson, and Stephen R. Briggs, eds., *Handbook of Personality Psychology*. San Diego, Calif.: Academic Press.
- Pervin, Lawrence A. (ed.) 1989 *Goal Concepts in Personality and Social Psychology*. Hillsdale, N.J.: Lawrence Erlbaum.
- Plomin, Robert 1990 “The Role of Inheritance in Behavior.” *Science* 248:183–188.
- Rorer, Leonard G. 1990 “Personality Assessment: A Conceptual Survey.” In Lawrence A. Pervin et al., eds., *Handbook of Personality: Theory and Research*. New York: Guilford.
- Rothbart, Mary K., Douglas Derryberry, and Michael I. Posner 1994 “A Psychobiological Approach to the Development of Temperament.” In John E. Bates, and Theodore D. Wachs, eds., *Temperament: Individual Differences at the Interface of Biology and Behavior*. Washington, D.C.: American Psychological Association.
- Saucier, Gerard 1997 “Effects of Variable Selection on the Factor Structure of Person Descriptors.” *Journal of Personality and Social Psychology* 73:1296–1312.
- , and Lewis R. Goldberg 1998 “What Is Beyond the Big Five?” *Journal of Personality* 66:495–524.
- Suls, Jerry, James P. David, and John H. Harvey 1996 “Personality and Coping: Three Generations of Research.” *Journal of Personality* 64:711–735.

- Zuckerman, Marvin 1995 “Good and Bad Humors: Biochemical Bases of Personality and Its Disorders.” *Psychological Science* 6:325–332.

GERARD SAUCIER

## PERSONALITY THEORY

### WHAT IS PERSONALITY?

A definition of *personality*, and there are many such definitions, must precede a treatment of personality theory. One must note at the start that personality is not an entity or a thing. It is a mentalistic construct that serves as an abstract cognitive device for understanding (1) the characteristic ways human beings behave or are inclined to behave; (2) their perceptions of their defining characteristics; and in the view of many, (3) the common (in some cases measured) perception that others have of them. This view of “personality” overlaps to a great extent with William James’s view of self ([1890]1952), both psychological and social. Though personologists may at times refer to personality as if it had been reified, this stems less from their intention than from a limitation of the language used. Given this condition, one can only strive to assure as close a correspondence as is possible of the definition of this word with the complex reality it is intended to reflect and capture.

Definitions of personality vary depending on the standpoint—scientific, philosophical, humanistic, or strictly psychological—that one adopts. Personality for our purposes is one or another heuristic enabling social scientists to understand and predict human behavior, both overt and tacit, and individuals’ complex responses to the events that sweep over them daily. To speak of individuals’ personalities is to allude to all their longstanding, characterological properties that dispose them, given any set of circumstances, to respond in a predictable way. In popular parlance “to really know someone” is to understand that person’s self-concept and the idiosyncratic manner in which he or she deals with problems of daily life, whether social, political, or purely intrapsychic. In short, to know someone is to have a knowledge of their personality; but the personality of an individual is not a simple matter to assess. To realize this, one has only to consider that there is a shifting presentation of self as a function of the multifarious

situations in which one can find oneself. The same person with a single personality has a number of personae at his or her disposal for (generally) adaptive use.

It is useful to quote in this context a classic statement by William James:

*Properly speaking, a man has as many social selves as there are individuals who recognize him and carry an image of him in their mind. To wound any one of these images is to wound him. But as the individuals who carry the images fall naturally into classes, we may practically say that he has as many different social selves as there are distinct groups of persons about whose opinion he cares. He generally shows a different side of himself to each of these different groups. Many a youth who is demure enough before his teachers and parents, swears and swaggers like a pirate among his "tough" young friends. We do not show ourselves to our children as to our club-companions, to our customers as to the laborers we employ, to our own masters and employers as to our intimate friends. ([1890] 1952, chap. X, "The Consciousness of Self," pp. 189–190)*

This statement, written in the 1880s, has a modern ring, and clearly presages the later work of, say, Hartshorne and May (1928) and Walter Mischel (1968). Clearly it represents a theory of personality that is relativistic and relational. Relative to the former aspect, it invokes a double perspective. The first is the perspective of individuals as they present themselves very differently and even inconsistently in various contexts to different classes of people. The second is the perception of observers who make judgments about individuals' personalities (or characters) based on the limited sampling of behavior they have witnessed either directly or indirectly. As such, it also presents the notion of self and of personality, though these two terms do not have identical meanings, as much more complex when viewed from within by the individual than when viewed from without by society. Relative to the relational aspect, personality appears to be shaped by the demand characteristics of the social group within which the individual performs. In this Jamesian perspective, personality is the flip side of self, that

is, observers' characterizations of others as distinguished from the views that those latter individuals have of themselves. Indeed the everyday language that individuals use to describe themselves (as well as others) has provided the vocabulary (e.g., altruistic, aggressive, nurturant, inquisitive, venturesome) for defining the components of the construct, personality (e.g., Cattell 1964, 1965). Further, the equivalents of these terms, used to personologically distinguish individuals, can be found in most of the world's languages. (This notion has been labeled the "fundamental lexical hypothesis," and is addressed below.)

An individual's identity—understanding of his or her character and typical patterns of behavioral responses to social and other environmental stimuli—is the result not simply of their personal history, but of his or her construal of that history and of the either vivid or tacit memories that form the warp and weft of their self-understanding. It is generally accepted that memories are never absolutely veridical. The fact that false memories are more or less richly interlarded with relatively true memories suggests that personality as viewed from its owner's standpoint is partially self-constructed. William James stated that "False memories are by no means rare occurrences in most of us, and, whenever they occur, they distort the consciousness of the me . . . The most frequent source of false memory is the accounts we give to others of our experiences" ([1890]1952, p. 241). He asserts that this is a source of the errors in testimony that the individual has every intention of making honest. This is confirmed in more recent analyses (e.g., Laurence et al. 1998) that demonstrate that people's characterizations of themselves or others can be compellingly shaped by the situational as well as by intrapsychic demands that subjects experience when they must make judgments about personality. That these dynamics may be at work when individuals are completing even the best-validated and most reliable of personality inventories is reason to draw conclusions from such inventories with caution.

### THEORIES IN GENERAL

Theories that have been developed to explain the dimensions of personality and the way personality develops are too numerous to describe (see Table

## PERSONALITY THEORY

---

### **Selected Personality and Related Theories, and Their Originators**

---

Adler, Alfred	Developed a holistic theory, individual psychology, characterized by teleological, sociobiological, and self-actualizing dynamics. Utilized notion of unconscious, but used a rational, commonsense approach to remediating problems.
Allport, Gordon	Theorized that structural elements of personality comprised traits disposing individuals to respond to stimulus fields in predictable patterns. Traits are viewed as both individualizing and nomothetic.
Angyll, Andreas	Developed a partial theory that stressed the human's need to serve superordinate group goals, while striving for self-individualization.
Bandura, Albert	Developed a social cognitive theory of human development in which the modeling of behavior by important others shapes each person's behavior and character. This theory has a strong teleological emphasis meshed with sense of self-efficacy.
Berne, Eric	Posited a phenomenological theory of human personality, <i>transactional analysis</i> , comprising three distinct ego states: parent, adult, and child.
Binswanger, Ludwig	An existential psychologist, a disciple of Heidegger, who elaborated a phenomenological approach to the understanding of human behavior, especially in the face of the most difficult aspects of life.
Burrow, Trigant	Pioneered view of humans as profoundly social, interactive, personologically shaped by group activity. An early group therapist.
Boss, Medard	A Heideggerian whose antitheoretical approach nevertheless limns a view of human nature that is constructivist, but with traces of the psychoanalytic and the phenomenological.
Cattell, Raymond	Developed a complex trait theory and psychometric measure (16 PF) through factor-analysis of vernacular trait expressions.
Corsini, Raymond	Formulated a developmental theory of personality of Adlerian and Rogerian inspiration, involving a strong genetic component; emphasized formative influence of parental values, educational experiences, and critical stochastological events.
Digman, John	(See McDougall, W.)
Dollard, John	Developed (with Neal Miller) a theory of personality of behavioral inspiration, constituted largely of evolving habits; drives and stimulus-response connections account for dynamics of changing personality structures.
Erikson, Erik	Elaborated a stage-based, life-span model of human development that was psychosocial and psychoanalytic in inspiration. A major focus is on personal identity development.
Eysenck, Hans	Fashioned a psychometrically supported model of human personality that is biologically trait-based and hierarchically organized.
Frankl, Viktor	Developed a life-span, noninstinctual, teleological model of human development in which the dominant motivational force is the will to meaning, underpinned by values and ideals.
Freud, Sigmund	Developed a state-based, conflict theory of personality that is pansexual and deterministic. The central construct is the Oedipal conflict, the resolution of which enables, especially for the male, the full expansion of the psyche.
Horney, Karen	Developed a revisionist psychoanalysis that shed its androcentric features; her sociopsychological theory later evolved along Adlerian lines and generated a cogent feminine psychology.
Jackson, Don	With J. Haley, P. Watzlawick, J. Beavin, and others at the Mental Research Institute in Palo Alto, California, formulated a humanistic, systems-oriented approach to understanding human character and behavior.
Janet, Pierre	Founder of <i>psychological analysis</i> . Formulated a holistic psychology of the person, including a theory of rapport, the unconscious, and the complexes that drive human behavior. A nineteenth-century empiricist, he developed, inter alia, notions of transference and suggestibility.
Jung, Carl	Founder of <i>analytic psychotherapy</i> . Posited a holistic psychology of the person, including a theory of the unconscious and the complexes that drive human behavior.
Kelly, George	Developed <i>personal construct theory</i> , a constructivist system of human personality development, predicated on the assumption that individuals construe reality in light of their life history. These construals govern anticipation and realization of events.

(continued)

**Selected Personality and Related Theories, and Their Originators (continued)**

Lewin, Kurt	Developed a topological psychology predicated on Gestalt psychological principles; this is a psychophysical model for human behavior utilizing geometric constructs to conceptualize personality determinants.
Lowen, Alexander	Reichian in inspiration, Lowen developed <i>bioenergetic analysis</i> , predicated on a multifactorial model of human personality, heavily organicist in character. Therapy involves body work.
Maslow, Abraham	Fashioned the <i>psychology of being</i> , a model of human nature as self-actualizing, holistic, creative, and joyful. This nomothetic model postulates a hierarchy of psychophysical needs that will sequentially assert themselves, if not impeded.
May, Rollo	Developed an existential personology that accents the dynamic. This <i>continental existentialism</i> provided the elements and "givens" for conflicts at the core of personality development.
McDougall, W.	The <i>five-factor model</i> of personality has a distinguished lineage: Formulated by W. McDougall and developed by L. Klages, F. Baumgartner, G. Allport, E. Borgatta, D. Fiske, E. Tupes and R. Christal, and J. Digman, among other more recent investigators, it involves a strong heritability component; this model is cross-cultural and statistically and psychometrically generated.
Mead, George	Developed <i>social interaction theory</i> , a system of social behaviorism in which individuals symbolically interiorize their roles and status as a function of the interactive perceptions of others.
Miller, Neal	(See Dollard, John)
Meyer, Adolf	Developed a psychobiological theory of personality, stressing the unity of mind and body, the former a function of the latter. The theory has a strong orientation to the organismic in the genesis of psychopathology.
Murray, Henry	Elaborated a comprehensive, holistic theory that stressed motivation, environment, the psychodynamic, the idiographic-nomothetic spectrum, the personal, the life-span history, and specifically the cerebral determinants of one's ever-in-flux personality.
Piaget, Jean	Developed a partial theory of personality comprising a seminal model of cognitive development that deliberately made abstraction of emotions and social matrices. The model is characterized by an ineluctably sequenced, linear series of cognitive stages.
Rank, Otto	A polymath, he developed a humanistic personology based on minimizing primordial fears (of life and death) and developing one's will, which integrates a person's sense of who he or she is. Influenced Carl Rogers.
Rogers, Carl	Developed a theory of human nature characterized by self-actualization and holism; an innate organismic valuing process, if left unimpeded, leads individuals to live fully effloresced, authentic, and healthful lives.
Skinner, Burrhus	Developed an operant conditioning theory, which serves as a scientific heuristic for understanding the development of behavioral habits and capabilities of any organism; mistakenly thought by some to deny existence of intrapsychic realities.
Sullivan, Harry	Developed an interpersonal theory of personality that focuses on the social situation rather than the person; this theory postulates that individuals' personalities are a reflection of the assessment others make, or are imagined to make, of them.
Wolpe, Joseph	Rejecting psychodynamic methodologies, he evolved a theory of personality development that was of Pavlovian inspiration. His theory is of interest principally to psychotherapists.

Table 1

l for a partial list). They each have been zealously defended and propagated by their adherents. In spite of the paucity of empirical validation for most of them, this has not assured their demise. Though most of the large-scale, molar theories that are still studied have been developed in the twentieth century, the study of human personality has veered in the second half of the century toward the manageably molecular. The trend to fine, multivariate analyses of human behavior in the

established experimentalism of academia has created its own tension with the countervailing movement to holistic conceptualizations of personality. The synthesizing of psychosociological findings with those that are cognitive, affective, and psychoneuro-endocrinological is a work in progress. At the end of the day, this scientific project has eventuated in scholarly specialization in numerous subdisciplines, which can only, with great difficulty and some contortions, be articulated

into a unified and coherent theoretical system. The task still needs to be addressed.

### HISTORICAL PRECURSORS

It needs to be noted that theorizing about the determinants of human behavior in individuals as well as their social groups has its origins in antiquity and has progressed throughout history. There is a rich vein of philosophical speculation on this subject in the Greco-Roman civilization that was the seedbed of Western culture. Hellenistic thinkers and playwrights bequeathed a rich assortment of treatises and ideas to what would become Euro-American concepts of the nature of personality and its development. The ideas that are most salient for us and that have had the greatest impact on modern Western thought have come from the great Hellenistic thinkers that preceded the era of Roman cultural and military hegemony. Among the pre-Socratics, physiological theories of the development of personality can be traced at least to Empedocles. Later Hippocrates (fourth century B.C.E.) and still later Galen (second century C.E.) (Kagan 1994) postulated that the physical humors of the body are related to temperament. Blood, black bile, yellow bile, and phlegm were linked, as these nouns suggest, to sanguine, melancholic, choleric, and phlegmatic temperaments, respectively. The specifics of this theory are given no credence today, but the principle that personality has physiological determinants (or at least correlates) is very much alive.

Plato and Aristotle proposed powerful models of human psychological development. In *Laws*, for example, Plato proposed an environmental perspective relative to the problems that are occasioned by parents' overreactions to children's spontaneous and immature behaviors.

*The privacy of home life screens from general observation many little incidents, too readily occasioned by a child's pains, pleasures, and passions which are not in keeping with a legislator's recommendations, and tend to bring a medley of incongruities into the characters of our citizens.* (Book 7, sec. 788)

He does not exclude a genetic perspective as the following passage from the same work attests.

*Now of all wild young things a boy is the most difficult to handle. Just because he more than*

*any other has a fount of intelligence in him that has not yet "run clear," he is the craftiest, most mischievous, and most unruly of brutes. So the creature must be held in check, as we may say, by more than one bridle—in the first place, when once he is out of the mother's and the nurse's hand, by attendants to care for his childish helplessness, and then further, by all the masters who teach him anything.* (Book 7, sec. 808)

Plato's student, Aristotle, was a more entrenched environmentalist than Plato. He affirmed the principle of *tabula rasa* in his treatise "On the Human Mind."

*Mind is in a sense potentially whatever is thinkable, although actually it is nothing until it has thought. What it thinks must be in it just as characters may be said to be on a writing-tablet on which as yet nothing stands written. This is exactly what happens with the mind.* (Book 3, chap. 4, sec. 430)

That the human personality evolves as a function of the myriad contingencies that befall a person during the course of life has been entrenched since classical times. Plato asserted that

*While spoiling of children makes their tempers fretful, peevish and easily upset by mere trifles, severe and unconditional tyranny makes its victims spiritless, servile, and sullen, rendering them unfit for the intercourse of domestic and civic life.* (Book 7, sec. 791)

### RELIGION AND THE PATHOGNOMONIC

That the human character has dysfunctional and even evil propensities is a nomothetic principle that was propagated in Western thought as a salient dogma of the Christian church. Although widely rejected, it pervaded much of political and social theorizing up to and even beyond the seventeenth century. Its most radical expression was in the homilies of such divines as Jonathan Edwards. The biblical dictum "I was shapen in iniquity and in sin did my mother conceive me" (Psalm 51) founded in part the traditional Christian view that human nature is fundamentally flawed by "original sin." It found expression in various vehicles of social thought and intellectual discussion of that time. The philosophers of the Enlightenment in

Europe repudiated this doctrine, although it continued to be propagated in religious circles and adumbrated in secular writings. The notion that human nature was essentially corrupt persisted in many of the systems of philosophy that flourished in the nineteenth century.

The dominant personality theorists of that period were clinicians with medical training. This professional background reinforced the pathological slant they gave to their descriptions of the human personality. Their theories of human psychological development gave prominence to the causes of deviant behavior rather than to the conditions for normal growth. Models of human personality were larded with dispositions to aggressive and narcissistic behavior. Freud in his more mature writings designated the aggressive instinct as one of the two pillars of his drive theory; the other, of course, was the erotic, a notion conceived in variously narrow and broad terms. His expansive pleasure principle was anticipated in the eighteenth century in the philosophy of Utilitarianism. Jeremy Bentham, a member of that school, stated that humans routinely engage in a “felicific calculus” ordained to maximize pleasure and minimize pain.

Twentieth-century theoreticians, principally university-based experimentalists, have striven to develop personological models focusing on wellness rather than on the pathognomonic. Gordon Allport is one of the outstanding representatives of this movement. Nevertheless the vocabulary, grammar, and notions of dysfunction and pathology continue to be woven into modern trait psychology. For example, Eysenck’s three-factor model of personality contains a psychoticism factor and a neuroticism factor, and neuroticism is one of the factors in the influential five-factor model of human personality postulated by William McDougall and most recently associated with the work of Paul T. Costa and Robert R. McCrae.

### SELF-ACTUALIZATION

A contrary movement, often termed “humanistic,” is based on the self-actualizing and growth-oriented models of Kurt Goldstein, Otto Rank, Carl Rogers, Abraham Maslow, the cognitive psychologist Jean Piaget, and many others in which

pathology is conceptualized as a departure from normality rather than a component of it. Rogers, for example, asserted that just as plants are shaped by the directions they must take in seeking sources of light, human beings evolve throughout their life spans by reaching for the emotional and social sustenance that will allow the “organismic valuing process” to shape them into fully effloresced and functioning persons. Intrinsic wickedness and pathology have no place in these models of human nature. And although human nature can be thought of in terms of fundamental human needs, as in Maslow’s *Psychology of Being*, the needs are not directed to destruction of self and others but to the fullest expression of a creative, generous, and joyful life expression. This is clearly in conflict with the psychodynamic models of nineteenth-century psychiatry, among them the Freudian system, which explained the engine of human development in terms of libidinal drives and tension reduction.

Self-disclosure and social transparency is one of the more healthful aspects of this model of the normal personality. O. Hobart Mowrer stated that psychological health depends on conditions that ensure transparency, and he traced its historical roots as far back as the ancient practice of *exomologesis*, in which a community of believers periodically engaged in collective confession of their violations of community mores, as in the monastic communities of the (Semitic) Essenes. This principle finds its most celebrated expressions in the psychotherapeutic literature. Moritz Benedikt, in the last third of the nineteenth century, and Carl Gustav Jung, in the twentieth, made it central to their therapeutic systems. They emphasized the critical importance for patients of divulging the “pathogenic secrets” that are woven into the fabric of their lives. A modern expression of this burgeoned in the 1960s in the therapeutic community (TC) movement, in locations such as Daytop Village and Phoenix House, where each day began with a morning meeting in which each member of the “family” publicly recounted his or her failures of the preceding day.

That we know more than we can tell and that we know more than we are aware that we know has been postulated since the pre-Socratics of Hellenistic Greece. Johann Christian Reil’s seminal work, “Rhapsodien,” published in 1803, describes the phenomenon of multiple personalities, and a complex topographical model of the human psyche. It

is rarely questioned as we approach the twenty-first century that our behavior is influenced and our personality shaped by information that is sedimented in the organism, to paraphrase Merleau-Ponty, but which exists at a subsymbolic level and cannot in every instance be consciously accessed. The somatic-marker hypothesis of Antonio Damasio (1994; see chap. 8, pp. 165–201) is a recent expression of this position, which has profound implications for our conceptualization of what “personality” is and how it is shaped.

#### **DEVELOPMENTAL PSYCHOLOGY AND PERSONALITY THEORY**

It is self-evident to the psychologist that one cannot treat the construct of personality without grounding it in a scientific developmental psychology. This is true even of those theories that are simply descriptive rather than causal-explanatory. Persons are not born with personalities, although Thomas and Chess (1980), for example, have demonstrated in their widely cited New York Longitudinal Study that temperamental traits, constituents of some models of personality, are strongly influenced by hereditary factors. Personalities develop from infancy to adulthood and, indeed, across the life span. If the developmental psychology on which a personality theory is based is seriously flawed, then it is highly probable that that personality theory, itself, is also flawed. Ausubel et al. (1980) have reviewed a large number of developmental theories and their derivative personality theories. These theories range from the preformationist and quaintly theological to the most social and educationist systems that have flourished in the humanist climate of late-twentieth-century Western thought. They are, it must be emphasized, of unequal value. It flows from this that as the prevailing models of developmental psychology have evolved in the light of empirical research, personality theories have also had to evolve. For example, in the measure that the psychosexual stage theory of human development proposed by Freud a hundred years ago has been superseded and, arguably, discredited, to that extent has Freudian personality theory become superannuated. This can be stated of a number of personality theories whose value to scholars is now largely of a historical character. It is useful to add at this point that a theory of developmental psychology *is* a theory of personality as long as it

provides an integrative and broad-scale view of the way human beings develop socially, psychically, and characterologically.

#### **THE MOVEMENT TO HOLISM**

Holism, a much-used term of Greek etymology, characterizes an approach to understanding human personality that integrates all aspects of the person and, more recently, the social matrix in which it has been shaped and finds expression. Human psychology that attempts to explain behavior by appeal to only one “faculty” or one dimension of the human organism, that restricts itself to, say, simply rationalist or neuroendocrinological processes without understanding the systemic character of personality, has largely fallen into disfavor. Jan Christian Smuts, Alfred Adler, and Kurt Goldstein were early proponents of an integrative and holistic approach to the study of human beings, integrating organicist as well as social dimensions in a comprehensive view of personality and its determinants. Abandonment of fragmentary models has increased the appeal of Maturana and Varela’s work (1980). Situated in a constructivist tradition, this work is not only biological but also profoundly social-psychological in its conceptualization of the human person. For example, these researchers’ concept of “autopoiesis” is that the integrity and unity of the organism is generated from within, even as it interacts with and accommodates its surroundings. Every person is a self-organizing entity. The feed-forward mechanisms of which modern constructivists speak indicate that knowledge is to a great extent the construction of schemas involving the total organism. In this respect even the human’s immunological system is part of a larger knowing system. The congruence between this vision and the self-actualizing models of Goldstein, Maslow, and Rogers is obvious. Unidimensional systems have evolved toward multidimensionality. Personologists have increasingly pursued the goal of formulating integrative models of human experience.

“Holism,” as an approach to understanding human personality, can now be characterized more broadly as biopsychosocial. The knowledge that individuals have of themselves, whether tacit or conscious, is not simply intrapsychically determined. That knowledge is a function of what they

have been told they are by those who are most meaningful to them. The principle of Harry Stack Sullivan that our self-concept is the image that we see reflected in the eyes of those who are making appraisals of us had its antecedents in the concept of the "looking-glass" self, developed by Charles Horton Cooley, George Herbert Mead, and James Mark Baldwin at the beginning of the twentieth century. A broadly systemic view of the development of the human person has come into focus, and it is profoundly sociologized.

### THE FEMINIST DYNAMIC

It is a commonplace notion of modern psychology that nomothetic models of the person have in the past been formulated by Caucasian males who unreflectively based them on male psychological norms. If, as Nietzsche said, all theory is autobiographical, it would follow that theories of personality developed by males reflect in part their personal and culturally circumscribed experiences. Sigmund Freud's, Lawrence Kohlberg's, and Erik Erikson's theories of moral development and evolution of human identity have been identified, among others, as based on male norms. Deviations from this male template have been considered abnormal. Karen Horney was among the first to repudiate psychoanalytic assertions that women were endowed with a moral sense inferior to that of males because they could not as children resolve the "Oedipal conflict" in the univocal sense that boys could. Feminists, including Inge K. Broverman, Pat Chesler, Carol Gilligan, and Carolyn Z. Enns have detailed the distortions implicit in characterizations of the typical female personality as submissive, conformist, masochistic, depression-prone, and so forth. The internalization of social roles by both men and women inevitably results in personality profiles that reflect those roles. The argument is made that establishing a template for normality that derives from a culture-circumscribed social role for men necessarily casts the typical feminine personality in the realm of the deficient, at least to the extent that it is laden with the deficits of the roles in which women have been socialized.

Whether there are universal, that is, nomothetic psychological differences between men and women arising from genetic determinants of gender-related behavior is an empirical question that is

still being studied by behavioral geneticists. If indeed there is no significant difference in the psychological development of the two sexes independent of fluid, relativistic cultural variables, it would not seem to matter which sex one used as the nomothetic template for healthy human development. The focus of analysis would necessarily shift to the socially constructed roles that are imposed on females. Assessing women as inferior by virtue of their possession of personality traits they have internalized in a society that has prescribed them would seem patently unjust.

### NEUROPSYCHOLOGICAL INFLUENCES IN PERSONOLOGY

Just as there are no two Coke bottles and no two newly minted pennies that are perfectly identical, a fortiori there are no two human beings who are truly identical. Reminded that the preponderance of the human genome endows each of us with characteristics that are universal, we must also acknowledge that there are at least 20,000 genes, called "polymorphic genes," that come in many varieties, that get randomly assorted at conception, and that endow us in part with our individuality. The genes that define us as human are called "monomorphic genes." They account for the fact that we all look recognizably human and normally act so. Whatever nomothetic principles have been truly (or purportedly) established relative to the psychobiological and behavioral properties of the human, they derive, by definition, from these genes.

Behavioral geneticists such as Robert Plomin and Thomas Bouchard (among other researchers in this field) have established the contribution of heredity to personality. For example, Bouchard has done well-regarded studies on twins reared apart that have shed light on the heritability of personality traits. Although his findings and those of other behavioral geneticists have excited controversy and negative comment, there is a convergence of evidence to show that indeed some of the variance one finds in certain personality traits is of genetic origin. But one needs to recognize that the controversy is fraught as much with ideological and political concerns as with strictly scientific ones, and the values embedded in these conflicting opinions can tilt arguments in one direction as well as another. The flashpoint, par excellence, for



this controversy is the position taken on the personality trait “intelligence,” a construct referring to certain adaptive and creative human capabilities. That these capabilities are still imperfectly understood ensures that the various constructs purported to define them are even more imperfectly formulated. Psychometricians are placed at an even further remove from these nebulous realities when they attempt to measure them, often by paper-and-pencil instruments.

The power of environment to shape or misshape personality is incontestable. Attachment Theory, associated with the work of John Bowlby, among others, offers a cogent explanation of the conditions for healthful psychological development. This theory has an affinity with the principles evolved by such ethologists as Tinbergen, Lorenz, Hinde, and other students of the social life of animals. Failure to develop a strong infant–primary caregiver bond during the period of primary socialization (that is, between the appearance of the first social smile at several weeks of age and the appearance of stranger anxiety at, say, six to eight months of age) is thought to lay the groundwork for developmental psychopathy—and the rupture of this bond in the first years of life, is thought to lay the groundwork for later affective disorders. A related perspective is that of Object-Relations theorists, whose views on “mothering” and its influence on development are more (human) relational and less instinct- or drive-based than the classical psychoanalytic model.

Although the influence of childhood experiences, both positive and negative, on adult personality is not in dispute, the partial reversibility of these effects is asserted by Michael Rutter, Jerome Kagan, and others. The variables at issue, for example, bonding in infancy to a primary caregiver and stimulus deprivation in toddlerhood are, after all, continuous and complex. The strength and duration of the variables, and their interaction with genetic variables of an idiographic and often unknown character, produce personality effects that are not predictable with any accuracy. The principle of *neoteny*, that is, the slowing of developmental rates and the retention of plasticity and developmental capabilities well into adulthood, is at issue here. Humans enjoy a relatively long period of growth from infancy to adulthood, a period in which earlier psychological deformity can be mitigated, if not entirely undone. In the wake of a

traumatic infancy and childhood, prolonged immaturity provides humans with the affordances of redemption.

### HOMUNCULARISM

Theories of personality development range on a continuum from the most rigidly preformationist at one end to the most malleable and environmentally sensitive at the other. Historically, Western psychology has been shifting toward the latter. In his remarkable work, *Centuries of Childhood* (1962), Philippe Ariès has demonstrated the long tradition of viewing children as miniature adults, *homunculi*, so to speak. The homuncular theory in its most primitive form postulated that the human organism, a little man (i.e., homunculus), contained in the semen of the male, is deposited in the uterus of the mother. It is presumed organically and morphologically complete. It merely needs to be nurtured to a mature status. The term has been extended as we know to embrace the postulate that not only the organic features of the mature human reside in the neonate but also the psychological features.

Children are, in this view, born with preformed characters and the cognitive structures of adults. To quote Ausubel et al.,

*The basic human properties and behavioral capacities—personality, values, and motives; perceptual, cognitive, emotional, and social reaction tendencies—are not conceived as undergoing qualitative differentiation and transformation over the life-span but are presumed to exist preformed at birth.* (1980 p. 15).

Vestiges of homuncularism remain in the work of twentieth-century personality theorists who attribute to infants and older children cognitions and emotions that, in a univocal sense, can only be the product of adult mentation. Confusing infantile sexuality with adult sexuality, as in classical psychoanalysis, is an example of this (cf., e.g., Thomas and Chess 1980, chap. IV, for a broader look at this issue). Among the more important of these innate schemas are the “racial unconscious” of Carl Gustav Jung and the “phylogenetic unconscious” of Sigmund Freud. They both have important implications for personological features and specific behaviors that are predicated on them.

It is widely recognized that the younger the organism, the more plastic it is in terms of acquiring and altering the learnings and adaptations it needs in order to thrive. Early-twentieth-century theories of human development postulated rigid and clearly delineated stages. These stages are no longer considered to have the rigid boundaries they were once supposed to have. In the earlier stage models it was presumed, but never demonstrated, that personality structure was established and firmly set by the age of, say, six or seven. In this view, the principal features of the personality did not change; they were only elaborated. This prescientific homuncularism has yielded to life-span developmental models.

There exists a countervailing point of view. Research has revealed the profound and subtle possibilities for change in the human personality that exist throughout the maturational process. Critical to this process of differentiation are education and other environmental factors, not excluding the benign and nurturant conditions that need to prevail during infancy in the home. Relevant to this orientation to human development is, again, the principle of neoteny. The slowing of humans' development from infancy to adulthood, has afforded humans the opportunity for evolving the complex and diverse personalities that, in the view of these theorists, is the hallmark of this species.

#### THE NOMOTHEIC AND THE IDIOGRAPHIC

That the psychologist as scientist has, in the past, inclined to a nomothetic and deterministic understanding of human beings is understandable. After all, it is difficult to build a science on the idiographic. Nevertheless, the concern of rank-and-file persons not to see themselves simply as one of billions of "knock-offs" from a universal template is also understandable. This may underlie some of the concern about cloning. "If there are clones about, how will you know who people really are?" ask some. Though there is a misconception about the phenomenon of cloning that underlies that question (after all, identical twins meet the definition of clones, and we come to know *who they are*), it indicates the concern about personal identity found in Western society. More specifically, the need to be an individual and to be

different (but not too different) from everyone else is evident in many Western and Westernized cultures. This plays out in the theorizing of personologists who have developed personality theories that give emphasis to the environmental factors that impinge on individuals in all phases of their lives, and that individuate them. It is this diversity in personal histories that accounts to a large extent for the uniqueness of each person. That members of the same family who have similar histories still mature to adulthood with very different personality profiles attests to the interaction of genetic factors with education, not excluding stochastic events that may have powerful formative impact on development.

#### THEORIES: DICHOTOMIES AND STAGES

The study of personality has addressed the question of what dynamics operate in all human beings to shape their behavior throughout the life span. This question assumes that there are some species-wide principles governing the development of human traits from conception to demise. It also addresses the question of what factors effect the individuation of human beings such that no one person is like any other person though they all share the same monomorphic genic substrate. Social scientists will recognize the ancient dichotomy for conceptualizing the relative contributions of nature as distinguished from life experiences. The code for this derives from Shakespeare's characterization of Caliban in *The Tempest*. Prospero refers to him as "a born devil, on whose nature Nurture can never stick." The expressions "nomothetic" and "idiographic" also refer to this dual factor—to wit, the universal principles governing the set "human beings" vis-à-vis relative principles governing the development of individuals within the set. In reality these are not true dichotomies except insofar as we wish to logically make them so. As in other conceptual polarities, explaining the emergence of a human being into adulthood admits of varying theoretical frameworks. A comparable perspective is reflected in the anthropological and linguistic distinctions between the *etic* (that is, the most general, often universal, frame for analyzing cultural behavior) and the *emic* frame (which examines social subsets of the species). An example of the former is Francis Galton's hypothesis, recently labeled "fundamental lexical hypothesis" (cf. Goldberg 1990, p.

1216), which affirms that the world's languages all contain terms describing similar personality traits, albeit traits that, although they have the same meaning, are valued differently, and present themselves with varying intensity and frequencies. An example of the latter is the principle (inherent in most multicultural perspectives on personality development) that the geohistorical background of any distinct society profoundly shapes the ideals of human behavior, morals, and social conduct, which get incarnated as "personality" in the majority of its members. The Sapir-Whorf hypothesis goes even further and states that the very structure of the language that a cultural subset of the human race speaks affects the character of their ideation and the consequent conduct of the group's affairs. But earlier, Nietzsche affirmed that "philosophers within the domain of the Ural-Altai languages" look into the world differently and take different paths than "Indo-Germans and Moslems: the spell of grammatical functions is in the last resort the spell of physiological value judgments and racial conditions" ([1885]1952, p. 472)

There have been zealous partisans of theories explaining human behavior that have, for reasons ideological as much as scientific, conceptualized personality into sets of dichotomies, in which they have emphasized one polarity rather than another. This antipodal schema has been with us since antiquity. The motifs of heroism and cowardice, of altruism and selfishness, of loyalty and betrayal, of candor and duplicity, of anguish and ecstatic joy, among other descriptors of the human experience that can play out in all human lives, received their finest expression in the works of such tragedians as Sophocles and Shakespeare. On the other hand, the uniqueness of each person and the passion for establishing one's personal identity is no less prominent in the great literature of the West. One has only to review, for example, the plays of Tennessee Williams, the great oeuvre of Cervantes, the vast range of characters in the novels of Balzac or Dickens, or in, say, Thomas Mann's *Magic Mountain*, to grasp the diversity that artists have displayed in limning the contours of these distinct personalities, to which readers have been able to resonate.

That human personality evolves and becomes different as an individual proceeds through a series of life stages would seem to be self-evident.

This view presumes, of course, that the construct "stage" is grounded in something real in human development. There are a number of complications to this view. First, the scope for change gets narrowed in the measure that variance in personality factors is explained by genetic factors. Conversely, the scope for change gets broadened in the measure that it is explained by education and other experience. Obviously, stage theory is less plausible, or less coercive in its implications, in the measure that personality development is fluidly malleable and educationally structured. The opposite seems true if biologically prestructured. But there is a second-order level for explaining personality change in a genetic perspective. One can accept a measure of predetermination in the psychoneurological substrate for personality, but it is not illogical to allow room in this model for a series of hardwired developmental periods. The model of Arnold Gesell, for example, though predeterministic, allows for qualitative stage-based shifts in the cognitive, psycho-motoric, interpersonal, and affective preferences and capabilities of the person.

As Piaget noted, whether or not one "sees" human development progressing from infancy to adulthood in stages is a function of how fine-grained and molecular one makes one's analysis. It is a question of scale. Standing inches from a *pointillist* painting, say, by Seurat, one sees only a multitude of tiny dots of pigment. It is only in distancing oneself so that the entire painting or large sections of it come into focus that the transitions from beach to sea to boats to trees to bathers become apparent. The phases of the panorama stand out as one macroscopically scans the canvas. It is difficult (but not impossible) to discern the shape of a galaxy when one is part of it.

As the analysis of personality development has assumed a more scientific, microscopic focus, there has been a tendency for the stage-based theories to fall into disrepute. This may account for the decline in popularity of the schemas developed by Freud, Erikson, Piaget, Sullivan, and Kohlberg, for example. Though these conceptualizations of human personality development continue to enjoy widespread support in part, if not in whole, they are partially in eclipse by virtue of advances in the newer human sciences. Cognitive science, psychoneurology, endocrinology, and social and

developmental psychology, among other disciplines, have inevitably superannuated all these grand systems to a greater or lesser degree.

### SUMMARY

On entering the twenty-first century, the Zeitgeist will favor, it appears, theories of personality that are life-span developmental, less instinct-driven but more persons-relational, constructivist, process-oriented and dynamic, that is, Heraclitean, holistic, teleonomic, evolutionary, genetics-based, gender-equal, emic, sociological, and idiographic. Clearly, personality theories that predate World War II are not, by and large, consistent with these descriptors. It appears that the era of the grand systems is past. Personology will reconcile itself to more modest paradigms for describing, explaining, and predicting human behavior.

Students who wish to follow the development of this discipline are urged to regularly consult the *Annual Review of Psychology*, the *Journal of Personality and Social Psychology*, the *European Journal of Personality*, *Developmental Psychology*, and other respected periodicals that publish articles in this domain. An excellent and more detailed analysis of many of the issues raised above can be found in *Theories of Personality* by Hall, et al. (1998, chap.1, 15).

### REFERENCES

- Ariès, P. 1962 *Centuries of Childhood: A Social History of Family Life*. New York: Random House.
- Ausubel, D. P., E. V. Sullivan, and S. W. Ives 1980 *The Theory and Problems of Child Development*, 3rd ed. Grune and Stratton.
- Cattell, R. B. 1964 *Personality and Social Psychology*. San Diego: Knapp.
- Cattell, R. B. 1965 *The Scientific Analysis of Personality*. Baltimore: Penguin Books.
- Damasio, A. R. 1994 *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
- Goldberg, L. R. 1990 "An Alternative 'Description of Personality': The Big-Five Factor Structure." *Journal of Personality and Social Psychology* 59:1216-1229.
- Hall, C. S., G. Lindzey, and J. B. Campbell 1997 *Theories of Personality*, 4th ed. New York: John Wiley.
- Hartshorne, H., and M. A. May 1928 *Studies in the Nature of Character*, vol. I, *Studies in Deceit*. New York: Macmillan.

- James, W. (1890) 1952 *The Principles of Psychology*. vol. 53 in *Great Books of the Western World*. Chicago: Encyclopedia Britannica.
- Kagan, J. 1994 *Galen's Prophecy: Temperament in Human Nature*. New York: Basic Books.
- Laurence, J. R., D. Day, and L. Gaston 1998 "From Memories of Abuse to the Abuse of Memories." In S. J. Lynn, ed., *Truth in Memory*. New York: American Psychological Association Press.
- Mahoney, M. 1991 *Human Changes Processes: The Scientific Foundations of Psychotherapy*. New York: Basic Books.
- Maturana, H. R., and F. J. Varela 1980 *Autopoiesis and Cognition: The Realization of the Living*. Boston: Reidel.
- Mischel, W. 1968 *Personality and Assessment*. New York: John Wiley.
- Nietzsche, F. (1885) 1952 *Beyond Good and Evil*. vol. 43 in *Great Books of the Western World*. Chicago: Encyclopedia Britannica.
- Pervin, L. A. 1996 *The Science of Personality*. New York: John Wiley.
- Thomas, A., and S. Chess 1980 *The Dynamics of Psychological Development*. New York: Brunner/Mazel.

FRANK DUMONT

## PERSUASION

The average number of hours the television set is on in American households is 6.8 hours a day (Peterson 1981). When not watching their 6.8 hours of television, most people spend the bulk of their time in talk with others. Much of this talk is geared not just to making oneself understood but to convincing someone else of the value and correctness of one's viewpoint. The average adult spends the majority of his or her waking hours at work, where, depending on the job, much activity involves efforts to get others to do one's bidding or being the object of such efforts. All this television watching, conversation, and work takes place in a social and political climate that, in theory if not in practice, encourages the exchange and dissemination of ideas among large numbers of people.

These facts have led some to conclude that this is an era of persuasion in which understanding who says what to whom in what way and with what effect is of critical importance (Lasswell 1948). In fact, some argue that the current era of persuasion is one of the few periods in the four millennia of

Western history characterized by such a degree of openness to argument (McGuire 1985).

Whether or not the present era is unique in this manner, more and more people are becoming conscious of the persuasive contexts in which they spend most of their time. Indeed, if the increasingly ingenious efforts of advertisers to pique interest and shape tastes and habits are any indication, people are becoming increasingly savvy about others' efforts at persuasion. This means that we have a very practical interest in understanding just how persuasion works. It also means that social scientists, and social psychologists in particular, have an interest in understanding and explaining a pervasive social phenomenon.

As one aspect of understanding persuasion, social psychologists have long studied attitude formation and change. During the 1920s and 1930s psychologists focused on describing the attitudes people hold. This led to the development of techniques for measuring attitudes, primarily scales such as the Likert scale, which continue to be used today by persuasion researchers. The second period of interest in attitude research occurred during the 1950s and 1960s, with the main focus moving from description to the study of attitude change and the effects of attitudes on behavior. (For a review of the research on attitude-behavior consistency, see Ajzen and Fishbein 1977.) This interest waned considerably during the next two decades as social psychologists became increasingly interested in social perception, or how people selectively interpret and respond to their social environments. The resurgence of interest in attitudes, and particularly in persuasion, that followed is thus largely informed by social psychology's more general emphasis on how people process the information they take in from their environment.

One might speculate that the interest in explaining persuasion and attitude change over mere attitude description reflects the increasing influence of mass media. However, this coincidence of research interest and social change is belied by the lack of communication between those studying the effects of mass media on attitudes and those studying persuasion in more immediately interpersonal contexts (Roberts and Maccoby 1985). The discussion here reflects this split in research focuses, concentrating solely on persuasion research in face-to-face, interpersonal contexts and

dealing only peripherally with research on the effects of mass media on attitudes. (See Roberts and Maccoby 1985 for a review of this literature.) The issue of brainwashing, an extreme form and method of persuasion, is considered only when it has direct relevance to less extreme persuasion contests and processes.

Of relevance to research on persuasion is the study of normative compliance occurring in settings where no active attempt is made to influence, but people change their opinions or judgments nonetheless. Asch's (1951) research on conformity demonstrated that subjects involved in a simple task of judging the length of a line were highly influenced by the judgments of others present, even when no overt influence attempts were made. In these studies people working with the experimenter gave incorrect assessments of the relative lengths of lines viewed. Even in cases of obviously incorrect judgments, most subjects conformed to the majority's assessment. Normative compliance is found to be greater the closer to unanimity the majority view, the larger the number holding it, and the more the conforming subject is attracted to and invested in group membership.

The persistence of findings of normative compliance, even in the absence of overt influence attempts, raises an obvious question: What happens when such overt attempts at persuasion are made? This and related questions are the focus of persuasion research. The focus in persuasion research is on attitude change "occurring in people exposed to relatively complex messages consisting of a position advocated by a communicator and usually one or more arguments designed to support that position" (Eagly and Chaiken 1984, p. 256). More simply defined, *persuasion* is an effort to change people's attitudes, these being the emotional and cognitive responses they have to objects, people, experiences, and so on.

#### FACTORS AFFECTING THE LIKELIHOOD OF PERSUASION

The guiding question of who says what to whom in what way with what effect has largely determined what factors researchers look to in explaining and predicting persuasion. These factors fall into four general classes: source or communicator variables, message variables, channel variables, and receiver variables. Reflecting two decades of research on

social cognition, studies of the effects of these variables on persuasion investigate how these factors affect persuasion by shaping the way in which people process information in the persuasion context. Indeed, the term *process* reflects the computer analogy often used to capture the manner in which people perceive, interpret, and respond to their environment.

For example, in studying the effects of source characteristics (the characteristics of the person communicating the persuasive message) researchers might examine how the likability of the source leads the receiver (the object of persuasion) either to attend to or to ignore the quality of the arguments accompanying the message. Similarly, researchers examine other aspects of cognition (e.g., attention, comprehension, receptivity, retention) for the manner in which they mediate the effects on persuasion of source, message, channel, and receiver variables.

**Source Variables.** The source variable of greatest interest is the credibility of the person communicating the persuasive message, including the communicator's apparent knowledge, social class, attractiveness, and likability. Consistently with common sense, the more credible the source, the more persuasive the source and the more likely that the receiver will change his or her attitude in the direction of the persuasive message. More interesting, however, is the combination of source credibility with other factors, and their combined effect on persuasion. For example, if the target person is not personally involved with the issue at hand, source credibility is more likely to enhance persuasion than if the person is highly involved in the issue. This is because personal involvement is likely to be associated with greater argument scrutiny by the target person, reducing the likelihood of immediate acceptance of even a credible source's position (Chaiken 1980).

The effects of source credibility on persuasion are thus mediated by the extent to which the target is motivated to thoughtfully scrutinize the supporting arguments presented by the source. Personal involvement in the issue is one such motivation, but specific knowledge of the topic (without any particularly emotionally charged investment in it) and being educated in general are also factors that mediate the effect of source credibility because of their impact on the manner in which the

target processes information at his or her disposal. Level of involvement, knowledge, and education are all characteristics of the receiver. The above example thus reveals the complex relationships between the factors that affect persuasion and their joint effects on information processing.

**Message Variables.** Many aspects of the persuasive message itself have been examined by research on persuasion. These include message style, ordering of arguments presented, speed of delivery, and message repetition. The effects of message repetition on persuasion are particularly interesting because they reveal the often unexpected combined effects of variables. For example, a study by Cacioppo and Petty (1979) revealed that only if supporting arguments are strong does repetition enhance persuasion, since repetition leads to greater argument scrutiny by the receiver, which enhances persuasion only to the extent that arguments are convincing.

Common sense might tell us that the use of humor in a message will enhance its persuasiveness by increasing the attractiveness of the source or, in the case of weak supporting arguments, distracting the target's attention from the content. Researchers have hypothesized that humor should enhance persuasion with a highly credible source, but not with a less credible source, since humor is likely to further reduce credibility in the latter case. However, there is little evidence to support the expected effects of humor on persuasion. The observed effects are rarely significant, and are as often negative as positive. Furthermore, the combined effects of humor with source credibility, or humor's impact on interest, retention, or source evaluation, are not found (McGuire 1985).

**Channel Variables.** Channel variables refer to the medium in which the message is communicated. For example, message persuasiveness varies depending on whether the message is given in person or in verbal, written, audio, or video form. In addition, channel variables include factors such as distraction—either direct distraction created by the behavior of the source or indirect distraction such as repeated external noise during the communication.

Petty and Cacioppo (1981) studied the effects of distraction on persuasion and found that its effects are mediated by cognitive factors such as the target's ability or motivation to scrutinize the

arguments. If the message is accompanied by a distracting noise, for example, this will enhance persuasion if accompanying arguments are weak, since it decreases the receiver's ability or motivation to pay close attention to the supporting arguments. Conversely, if the supporting arguments are strong, distraction decreases the likelihood of persuasion, especially with knowledgeable targets, since it makes it unlikely they will pay attention to the strong arguments designed to persuade them.

Similarly, other channel variables, such as the use of catchy music in television ads, affect persuasion to the extent that they motivate the receiver to generate positive rather than negative thoughts in response to the message. Generating thoughts is distinguished from argument scrutiny because it refers to the additional arguments or ideas the receiver brings to bear in evaluating a message, not to the supporting arguments provided by the communicator. Channel or other variables enhance persuasion to the extent that they generate positive thoughts or supporting arguments in the target (Greenwald 1968). Like other factors affecting persuasion then, the impact of channel variables on persuasion is mediated by the resulting cognitive processes engaged in by the target of the persuasive message.

**Receiver Variables.** Since persuasion is oriented toward convincing someone to adopt a particular viewpoint or opinion, it makes sense that the person himself or herself has some impact on the persuasion process. Interest in the personality correlates of persuasion was high during the 1950s but waned in the following decades as individual-level explanations of social phenomena became less popular among social psychologists. However, more recent work has revived interest in the effect of receiver variables on persuasion.

A variety of receiver variables relating to personality characteristics have been studied to determine receiver susceptibility and resistance to persuasion. With respect to self-esteem, researchers predict that greater self-esteem will decrease the likelihood of persuasion to the extent that it increases the likelihood that the receiver will carefully scrutinize the arguments, and decreases the likelihood of the receiver's being swayed by a credible source in the absence of strong supporting arguments. Therefore, like the other factors

affecting persuasion, self-esteem exerts its effects on persuasion through the variables related to the manner in which the receiver processes the information available in the persuasion context.

Results of studies of self-esteem and persuasion offer mixed results, some indicating negligible effects (Barber 1964) and others suggesting that the lower the self-esteem, the greater the tendency to conform, especially when the aspect of self-esteem involved is closely related to the issue at hand (Endler et al. 1972). The effects of self-esteem are complex, especially when combined with other variables. For example, greater influenceability is associated with higher, not lower, self-esteem as the complexity of the persuasive message increases.

More consistent findings have been produced on the effects of receiver's mood on responses to persuasive attempts. Interestingly, research has shown that neutral or bad moods decrease susceptibility to persuasive attempts, largely because targets are more likely to scrutinize messages when in a neutral or bad mood than when in a good mood (Wegener et al. 1995; Rosseli et al. 1995).

The effects of authoritarianism or dogmatism have also received considerable research attention. *Dogmatism* is defined as a general inclination to be closed-minded, intolerant, and deferential to authority. Research has shown that receivers low in dogmatism are persuaded by strong, but not by weak, arguments. Dogmatic receivers, in contrast, are persuaded by strong arguments only when the source is nonexpert. With expert sources, dogmatic individuals are equally persuaded by strong and weak arguments (DeBono and Klein 1993).

Aside from personality characteristics, persuasion researchers, like other social researchers, have shown a longstanding interest in the issue of gender differences. A summary of 148 studies of the effects of gender on persuasion (Eagly and Carli 1981) indicates that women are more easily influenced than men. However, as consistent as these findings are, the magnitude of gender differences in influenceability is small enough to raise doubts about its practical significance for people's everyday lives (McGuire 1985).

However trivial these differences may be, they have inspired persuasion researchers to search for

an explanation. One plausible account is that “greater female susceptibility and greater male predictability might derive from socialization differences such that conforming pressures are exerted more strongly and uniformly on women, compressing them into a narrow band of high influenceability” (McGuire 1985, p. 288). Thus, the effects of gender on influenceability result from the different social experiences of men and women (for example, the greater likelihood that women will hold jobs in which they will receive more persuasion attempts than they themselves perform).

#### **METHODOLOGICAL AND THEORETICAL ISSUES IN THE STUDY OF PERSUASION**

While all studies of persuasion tend to focus on some combination of the variables discussed above, more general research orientations have not been uniform. One way of dividing research approaches to persuasion is to distinguish between the descriptive and the mathematical models. The former may focus on any combination of variables affecting persuasion, but their predictions tend to be stated in the form of verbal argument or a set of hypotheses. Mathematical or probabilistic models, on the other hand, cast their predictions about persuasion in the form of equations, with variables and their relationships represented in algebraic terms. While the choice of research approach does not dictate which variables the researcher focuses on, some argue that the greater precision of probabilistic models makes for a more exact understanding of the conditions likely to produce persuasion (McGuire 1985). Others (e.g., Eagly and Chaiken 1984) refer to these models as “normative,” meaning that they describe how persuasion ought to work, not necessarily how it works in reality.

Whatever the relative merits of descriptive and probabilistic approaches, there is some consensus among persuasion researchers that a more general theory of persuasion is necessary if the vast research findings in the area are to be integrated in a meaningful manner (Eagly and Chaiken 1984). As this discussion shows, the research on persuasion is blessed with a large number of well-conceptualized variables, most of which are easy to operationalize, that is, to create in a laboratory setting. The discussion also shows, however, that even when a small portion of all the possible combinations of these variables are matched with

one or two of the mediating cognitive processes, the resulting insights into persuasion are far from some of the commonsense notions alluded to throughout this discussion. This fact lends support for the call for a more unifying theory of persuasion to integrate the somewhat fragmented picture that emerges from a combining-of-variables approach.

#### **SUSCEPTIBILITY AND RESISTANCE TO PERSUASION**

This discussion raises important questions about issues of susceptibility and resistance to persuasion. When is being easily persuaded good? When is resistance good, and how can it be taught? If Americans are spending between six and seven hours a day in front of their television sets, and the rest of their waking hours in persuasive communications with others on the job or at leisure, teaching resistance to all this persuasion might become a top priority. For example, when research reveals that motivating people to generate reasons for their attitudes toward a product ultimately causes them to change their initial attitude, and even to purchase a product on the basis of their changed attitude (only to regret it later) attention is called to the conditions under which enhancing resistance might be warranted (Wilson et al. 1989). On the other hand, when research shows that two years of viewing *Sesame Street* caused both black and white children to manifest more positive attitudes towards black and Hispanics, something is learned about the conditions under which enhancing attitude change and persuasion is desirable (Bogatz and Ball 1971).

Whether one values resistance or openness to persuasion, our understanding of resistance to persuasion is enhanced by our understanding of what increases the likelihood of persuasion. As this discussion has shown, source variables (e.g., credibility), channel variables (e.g., distractions), message variables (e.g., argument strength), and receiver variables (e.g. personality characteristics) can all affect the receiver’s susceptibility or resistance to persuasion. In addition, research has shown that when people perceive persuasive messages as threats to their freedom, they resist persuasion and maintain their original position or, in some cases, adopt a position opposite to the one advocated in the message (Brehm 1966).



More recent research on values reveals another factor involved in susceptibility and resistance to persuasion. In general, people are more susceptible to messages which are consistent with their own values, and more resistant to those that are not. For example, Han and Shavitt (1994) have shown that receivers with individualistic values (e.g., Americans) are more persuaded by advertising messages emphasizing individualistic themes (e.g., "Take care of number one!"). In contrast, receivers with collectivistic values (e.g., Koreans) are more persuaded by advertising messages emphasizing collectivistic themes (e.g., "For you and your family!"). Thus, whether we are interested in enhancing persuasion or increasing resistance to it, understanding the role of values in the persuasion process is crucial.

Ultimately, resistance to the potential loss of autonomy involved in persuasion is not surprising in a culture that places a premium on such individuality and freedom. In addition, such concerns are not unwarranted, given the fact that many of the attempts at persuasion people face in natural settings are specifically designed to minimize the threatening aspects of such attempts and thus to reduce resistance (hence the use of the word "seductive" to refer to arguments, advertisements, etc.). Are people continually at risk, then, of potentially harmful persuasion?

One could argue that in people's ability to be persuaded lies the possibility of autonomy. If people are easily persuaded, then they are in similar measure unlikely to be overwhelmingly persuaded by a particular viewpoint over all others. They may be likely to resist total indoctrination, if only because they are susceptible to some other credible source sending a convincing message in a captivating medium.

## REFERENCES

- Ajzen, I., and M. Fishbein 1977 "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research." *Psychological Bulletin* 84:888-918.
- Asch, S. E. 1951 "The Effects of Group Pressure upon the Modification and Distortion of Judgments." In H. Guetzkow, ed., *Groups, Leadership and Men*. Pittsburgh, Pa.: Carnegie Press.
- Barber, T. X. 1964 "Hypnotizability, Suggestibility and Personality: V. A Critical Review of Research Findings." *Psychological Reports* 14:299-320.
- Bogatz, G. A., and S. J. Ball 1971 *The Second Year of Sesame Street: A Continuing Evaluation*, 2 vols. Princeton, N.J.: Educational Testing Service.
- Brehm, J. W. 1966 *A Theory of Psychological Reactance*. New York: Academic Press.
- Cacioppo, J., and R. Petty 1979 "Effects of Message Repetition and Position on Cognitive Response, Recall, and Persuasion." *Journal of Personality and Social Psychology* 37:97-109.
- Chaiken, S. 1980 "Heuristic versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39:752-766.
- DeBono, K. G., and C. Klein 1993 "Source Expertise and Persuasion: The Moderating Role of Recipient Dogmatism." *Personality and Social Psychology Bulletin* 19:167-173.
- Eagly, A. H. and L. L. Carli 1981 "Sex of Researchers and Sex-typed Communications as Determinants of Sex Differences in Influenceability: A Meta-analysis of Influence Studies." *Psychological Bulletin* 90:1-20.
- Eagly, A. H., and S. Chaiken 1984 "Cognitive Theories of Persuasion." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 17. New York: Academic Press.
- Endler, N. S., D. L. Weisenthal, and S. H. Geller 1972 "The Generalization Effects of Agreement and Correctness on Relative Competence Mediating Conformity." *Canadian Journal of the Behavioral Sciences* 4:322-329.
- Han, S., and S. Shavitt 1994 "Persuasion and Culture: Advertising Appeals in Individualistic and Collectivistic Societies." *Journal of Experimental Social Psychology* 30:326-350.
- Greenwald, A. G. 1968 "Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change." In A. G. Greenwald, T. S. Brock, and T. M. Ostrom, eds., *Psychological Foundations of Attitudes*. New York: Academic Press.
- Lasswell, H. D. 1948 "The Structure and Function of Communication in Society." In L. Bryson, ed., *Communication Ideas*. New York: Harper.
- McGuire, W. J. 1985 "Attitudes and Attitude Change." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*, vol. 2. New York: Random House.
- Peterson, R. A. 1981 "Measuring Culture, Leisure, and Time Use." *Annals of the American Academy of Political and Social Science* 453:1969-1979.
- Petty, R. E., and J. T. Cacioppo 1981 *Attitudes and Persuasion: Classic and Contemporary Approaches*. Debuque Ia.: W. C. Brown.

- Roberts, D. F., and N. Maccoby 1985 "Effects of Mass Communication." In G. Lindzey and E. Aronson, eds., *The Handbook of Social Psychology*, vol. 2. New York: Random House.
- Rosselli F., J. Skelly, and D. M. Mackie 1995 "Processing Rational and Emotional Messages: The Cognitive and Affective Mediation of Persuasion." *Journal of Experimental Social Psychology* 31:163–190.
- Wegener, D. T., R. E. Petty, and S. M. Smith 1995 "Positive Mood Can Increase or Decrease Message Scrutiny: The Hedonic Contingency View of Mood and Message Processing." *Journal of Personality and Social Psychology* 69:5–15.
- Wilson, T. D., D. S. Dunn, D. Kraft, and D. J. Lisle 1989 "Introspection, Attitude Change, and Attitude Behavior Consistency: The Disruptive Effects of Explaining Why We Feel the Way We Do." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 22. New York: Academic Press.

SUSAN McWILLIAMS

## PHENOMENOLOGY

Phenomenology is a movement in philosophy that has been adapted by certain sociologists to promote an understanding of the relationship between states of individual consciousness and social life. As an approach within sociology, *phenomenology* seeks to reveal how human awareness is implicated in the production of social action, social situations and social worlds (Natanson 1970).

Phenomenology was initially developed by Edmund Husserl (1859–1938), a German mathematician who felt that the objectivism of science precluded an adequate apprehension of the world (Husserl 1931, 1954). He presented various philosophical conceptualizations and techniques designed to locate the sources or essences of reality in the human consciousness. It was not until Alfred Schutz (1899–1959) came upon some problems in Max Weber's theory of action that phenomenology entered the domain of sociology (Schutz 1967). Schutz distilled from Husserl's rather dense writings a sociologically relevant approach. Schutz set about describing how subjective meanings give rise to an apparently objective social world (Schutz 1954, 1962, 1964, 1966, 1996; Schutz and Luckmann 1973; Wagner 1983).

Schutz's migration to the United States prior to World War II, along with that of other phenomenologically inclined scholars, resulted in the transmission of this approach to American academic circles and to its ultimate transformation into interpretive sociology. Two expressions of this approach have been called *reality constructionism* and *ethnomethodology*. Reality constructionism synthesizes Schutz's distillation of phenomenology and the corpus of classical sociological thought to account for the possibility of social reality (Berger 1963, 1967; Berger and Berger 1972; Berger and Kellner 1981; Berger and Luckmann 1966; Potter 1996). Ethnomethodology integrates the Parsonian concern for social order into phenomenology and examines the means by which actors make ordinary life possible (Garfinkel 1967; Garfinkel and Sacks 1970). Reality constructionism and ethnomethodology are recognized to be among the most fertile orientations in the field of sociology (Ritzer 1996).

Phenomenology is used in two basic ways in sociology: (1) to theorize about substantive sociological problems and (2) to enhance the adequacy of sociological research methods. Since phenomenology insists that society is a human construction, sociology itself and its theories and methods are also constructions (Cicourel 1964, 1973). Thus, phenomenology seeks to offer a corrective to the field's emphasis on positivist conceptualizations and research methods that may take for granted the very issues that phenomenologists find of interest. Phenomenology presents theoretical techniques and qualitative methods that illuminate the human meanings of social life.

Phenomenology has until recently been viewed as at most a challenger of the more conventional styles of sociological work and at the least an irritant. Increasingly, phenomenology is coming to be viewed as an adjunctive or even integral part of the discipline, contributing useful analytic tools to balance objectivist approaches (Aho 1998; Levesque-Lopman 1988; Luckmann 1978; Psathas 1973; Rogers 1983).

## TECHNIQUES

Phenomenology operates rather differently from conventional social science (Darroch and Silvers 1982). Phenomenology is a theoretical orientation, but it does not generate deductions from

propositions that can be empirically tested. It operates more on a metasociological level, demonstrating its premises through descriptive analyses of the procedures of self-, situational, and social constitution. Through its demonstrations, audiences apprehend the means by which phenomena, originating in human consciousness, come to be experienced as features of the world.

Current phenomenological techniques in sociology include the method of “bracketing” (Bentz 1995; Ihde 1977). This approach lifts an item under investigation from its meaning context in the commonsense world, with all judgments suspended. For example, the item “alcoholism as a disease” (Peele 1985; Truan 1993) is not evaluated within phenomenological brackets as being either true or false. Rather, a *reduction* is performed in which the item is assessed in terms of how it operates in consciousness: What does the disease notion do for those who define themselves within its domain? A phenomenological reduction both plummets to the essentials of the notion and ascertains its meanings independent of all particular occasions of its use. The reduction of a bracketed phenomenon is thus a technique to gain theoretical insight into the meaning of elements of consciousness.

Phenomenological tools include the use of introspective and *Verstehen* methods to offer a detailed description of how consciousness itself operates (Hitzler and Keller 1989). Introspection requires the phenomenologist to use his or her own subjective process as a resource for study, while *Verstehen* requires an empathic effort to move into the mind of the other (Helle 1991; Truzzi 1974). Not only are introspection and *Verstehen* tools of phenomenological analysis, but they are procedures used by ordinary individuals to carry out their projects. Thus, the phenomenologist as analyst might study himself or herself as an ordinary subject dissecting his or her own self-consciousness and action schemes (Bleicher 1982). In this technique, an analytic attitude toward the role of consciousness in designing everyday life is developed.

Since cognition is a crucial element of phenomenology, some theorists focus on social knowledge as the cornerstone of their technique (Berger and Luckmann 1966). They are concerned with how commonsense knowledge is produced,

disseminated, and internalized. The technique relies on theoretical discourse and historical excavation of the usually taken for granted foundations of knowledge. Frequently, religious thought is given primacy in the study of the sources and legitimations of mundane knowledge (Berger 1967).

Phenomenological concerns are frequently researched using qualitative methods (Bogdan and Taylor 1975; Denzin and Lincoln 1994, 1998). Phenomenological researchers frequently undertake analyses of small groups, social situations, and organizations using face-to-face techniques of participant observation (Bruyn 1966; Psathas and Ten Have 1994; Turner 1974). Ethnographic research frequently utilizes phenomenological tools (Fielding 1988). Intensive interviewing to uncover the subject’s orientations or his or her “life world” is also widely practiced (Costelloe 1996; Grekova, 1996; Porter 1995). Qualitative tools are used in phenomenological research either to yield insight into the microdynamics of particular spheres of human life for its own sake or to exhibit the constitutive activity of human consciousness (Langsdorf 1995).

Techniques particular to the ethnomethodological branch of phenomenology have been developed to unveil the practices used by people to produce a sense of social order and thereby accomplish everyday life (Cuff 1993; Leiter 1980; Mehan and Wood 1975). At one time, “breaching demonstrations” were conducted to reveal the essentiality of taken-for-granted routines and the means by which threats to these routines were handled. Since breaching these routines sometimes resulted in serious disruptions of relationships, this technique has been virtually abandoned. Social situations are video- and audiotaped to permit the painstaking demonstration of the means by which participants construct their identities, their interpretations of the meanings of acts, and their sense of the structure of the situation (Blum-Kulka 1994; Jordan and Henderson 1995). Conversational analysis is a technique that is frequently used to describe how people make sense of each other through talk and how they make sense of their talk through their common background knowledge (Psathas 1994; Schegloff and Sacks 1974; Silverman, 1998). The interrelations between mundane reasoning and abstract reasoning are also examined in great depth as researchers expose, for example, the socially constituted bases of scientific

and mathematical practice in commonsense thinking (Knorr-Cetina and Mulkay 1983; Livingston, 1995; Lynch, 1993).

### THEORY

The central task in social phenomenology is to demonstrate the reciprocal interactions among the processes of human action, situational structuring, and reality construction. Rather than contending that any aspect is a causal factor, phenomenology views all dimensions as constitutive of all others. Phenomenologists use the term *reflexivity* to characterize the way in which constituent dimensions serve as both foundation and consequence of all human projects. The task of phenomenology, then, is to make manifest the incessant tangle or reflexivity of action, situation, and reality in the various modes of *being in the world*.

Phenomenology commences with an analysis of the *natural attitude*. This is understood as the way ordinary individuals participate in the world, taking its existence for granted, assuming its objectivity, and undertaking action projects as if they were predetermined. Language, culture, and common sense are experienced in the natural attitude as objective features of an external world that are learned by actors in the course of their lives.

Human beings are open to patterned social experience and strive toward meaningful involvement in a knowable world. They are characterized by a typifying mode of consciousness tending to classify sense data. In phenomenological terms humans experience the world in terms of *typifications*: Children are exposed to the common sounds and sights of their environments, including their own bodies, people, animals, and vehicles. They come to apprehend the categorical identity and *typified* meanings of each in terms of conventional linguistic forms. In a similar manner, children learn the formulas for doing common activities. These practical means of doing are called *recipes for action*. Typifications and recipes, once internalized, tend to settle beneath the level of full awareness, that is, to become *sedimented*, as do layers of rock. Thus, in the natural attitude, the foundations of actors' knowledge of meaning and action are obscured to the actors themselves.

Actors assume that knowledge is objective and that all people reason in a like manner. Each actor assumes that every other actor knows what he or she knows of this world: All believe that they share common sense. However, each person's biography is unique, and each person develops a relatively distinct stock of typifications and recipes. Therefore, interpretations may diverge. Everyday social interaction is replete with ways in which actors create feelings that common sense is shared, that mutual understanding is occurring, and that everything is all right. Phenomenology emphasizes that humans live within an intersubjective world, yet they at best approximate shared realities. While a *paramount reality* is commonly experienced in this manner, particular realities or *finite provinces of meaning* are also constructed and experienced by diverse cultural, social, or occupational groupings.

For phenomenology, all human consciousness is practical—it is always consciousness of something. Actors intend to introduce projects into the world; they act in order to implement goals based on their typifications and recipes, their *stock of knowledge* at hand. Consciousness as an *intentional process* is composed of thinking, perceiving, feeling, remembering, imagining, and anticipating directed toward the world. The objects of consciousness, these intentional acts, are the sources of all social realities, which are, in turn, the materials of commonsense.

Thus, typifications derived from common sense are internalized, becoming the tools that individual consciousness uses to constitute a *lifeworld*, the unified arena of human awareness and action. Common sense serves as an ever-present resource to assure actors that the reality that is projected from human subjectivity is an objective reality. Since all actors are involved in this intentional work, they sustain the collaborative effort to reify their projections and thereby reinforce the very frameworks that provide the construction tools.

Social interaction is viewed phenomenologically as a process of reciprocal interpretive constructions of actors applying their stock of knowledge at hand to the occasion. Interactors orient themselves to others by taking into account typified meanings of actors in typified situations known to them through common sense. Action schemes are geared by each to the presumed projects of

others. The conduct resulting from the intersection of intentional acts indicates to members of the collectivity that communication or coordination or something of the like is occurring among them. For these members, conduct and utterances serve as *indexical* expressions of the properties of the situation, enabling each to proceed with the interaction while interpreting others, context, and self. Through the use of certain interpretive practices, members order the situation for themselves in sensical and coherent terms: In their talk they gloss over apparent irrelevancies, fill in innumerable gaps, ignore inconsistencies, and assume a continuity of meaning, thereby formulating the occasion itself.

Ongoing social situations manifest patterned routine conduct that appears to positivist investigators to be normative or rule guided. Phenomenologically, rules are indexical expressions of the interpretive processes applied by members in the course of their interactions. Rules are enacted in and through their applications. In order to play by the book, an interpreter endeavors to use a rule as an apparent guide. However, he or she must use all sorts of background expectancies to manage the fit somehow between the particular and the general under the contexted conditions of the interaction, and in so doing is acting creatively. Rules, policies, hierarchy, and organization are accomplished through the interpretive acts or negotiations of members in their concerted efforts to formulate a sense of operating in accord with a rational, accountable system. This work of doing structure to the situation further sustains its commonsensical foundations as well as its facticity.

Phenomenologists analyze the ordering of social reality and how the usage of certain forms of knowledge contributes to that ordering. It is posited that typified action and interaction become *habitualized*. Through sedimentation in layered consciousness, human authorship of habitualized conduct is obscured and the product is externalized. As meaning-striving beings, humans create theoretical explanations and moral justifications in order to legitimate the habitualized conduct. Located in higher contexts of meaning, the conduct becomes objectivated. When internalized by succeeding generations, the conduct is fully institutionalized and exerts compelling constraints over individual volition. Periodically, the institutions

might be repaired in response to threats, or individuals might be realigned if they cognitively or affectively migrate.

The reality that ordinary people inhabit is constituted by these legitimations of habitualized conduct. Ranging from commonsense typifications of ordinary language to theological constructions to sophisticated philosophical, cosmological, and scientific conceptualizations, these legitimations compose the paramount reality of everyday life. Moreover, segmented modern life, with its proliferation of meaning-generating sectors, produces multiple realities, some in competition with each other for adherents. In the current marketplace of realities, consumers, to varying degrees, may select their legitimations, as they select their occupations and, increasingly, their religions (Berger 1967).

### PRACTICE

Doing phenomenological sociology involves using procedures that are distinct from positivist research. Phenomenological practice is increasingly evident in the discipline as more subjectivist work is published. The phenomenological analysis of mass media culture content, for example, applies the elements of the approach to yield an understanding of the reflexive interplay of audience lifeworlds and program material (Wilson 1996). Thus, TV talk show discourses may be described as social texts that are refracted by programmers from commonsense identity constructs. The visual realization yields narrative images that audiences are seduced into processing using their own experiences. The viewers' lifeworlds and the TV representations are blended into reality proxies that provide viewers with schemata to use in configuring their personal orientations. Subsequently, programmers draw upon these orientations as additional identity material for new content development.

Phenomenological work with young children examines how both family interactions and the practices of everyday life are related to the construction of childhood (Davila and Pearson 1994). It reveals how the children's elemental typifications of family life and common sense are actualized through ordinary interaction. Penetrating the inner world of children requires that the phenomenological practitioner view the subjects in children's own terms, from their levels and viewpoints

(Waksler 1991; Shehan 1999). Such investigation shuns adult authoritative and particularly scientific perspectives, and seeks to give voice to the children's experience of their own worlds. Infants' and children's communicative and interactive competencies are respected and are not diminished by the drive toward higher-level functioning (Sheets-Johnstone 1996).

At the other end of the lifecycle, phenomenologists investigate how aging and its associated traumas are constituted in the consciousness of members and helpers. The struggle for meaning during aging accompanied by chronic pain may be facilitated or impaired by the availability of constructs that permit the smoother processing of the experiences. Members of cultures that stock typifications and recipes for managing aging and pain skillfully may well be more likely than others to construct beneficial interpretations in the face of these challenges (Encandela 1997). Phenomenological work encourages the helpers of the elderly to gain empathic appreciation of their clients' lifeworlds and enhanced affiliation with them through the use of biographical narratives that highlight their individuality and humanity (Heliker 1997).

The healing professions, particularly nursing, seem to be deeply imbued with a phenomenological focus on the provision of care based on a rigorous emphasis on the patient's subjective experience (Benner 1995). Substantial attention has been devoted to the ethical implications of various disease definitions, to how language shapes the response to illness, and to how disease definitions and paradigmatic models impact communication between health professionals and patients (Rosenberg and Golden 1992). Significant work on the phenomenology of disability has demonstrated how the *lived body* is experienced in altered form and how taken-for-granted routines are disrupted by invoking new action recipes (Toombs 1995). Nonconventional healing practices have also been examined, revealing how embodiment and the actor's subjective orientation reflexively interrelate with cultural imagery and discourse to transfigure the self (Csordas 1997). Further, phenomenological work has suggested that emotions are best analyzed as interpreted processes embedded within experiential contexts (Blum 1996; Solomon 1997).

## IMPLICATIONS

For phenomenology, society, social reality, social order, institutions, organizations, situations, interactions, and individual actions are constructions that appear as suprahuman entities. What does this suggest regarding humanity and sociology? Phenomenology advances the notion that humans are creative agents in the construction of social worlds (Ainlay 1986). It is from their consciousness that all being emerges. The alternative to their creative work is meaninglessness, solipsism, and chaos: a world of dumb puppets, in which each is disconnected from the other and life is formless (Abercrombie 1980). This is the nightmare of phenomenology. Its practitioners fear that positivist sociologists actually theorize about such a world (Phillipson 1972).

Phenomenologists ask sociologists to note the misleading substantiality of social products and to avoid the pitfalls of reification. For the sociologist to view social phenomena within the natural attitude as objects is to legitimate rather than to analyze. Phenomenological sociologists investigate social products as humanly meaningful acts, whether these products are termed attitudes, behaviors, families, aging, ethnic groups, classes, societies, or otherwise (Armstrong 1979; Gubrium and Holstein 1987; Herek 1986; Petersen 1987; Starr 1982). The sociological production of these fictive entities is understood within the context of their accomplishment, that is, the interview setting, the observational location, the data collection situation, the field, the research instrument, and so forth (Schwartz and Jacobs 1979). The meaning contexts applied by the analyst correlates with those of the subjects under investigation and explicates the points of view of the actors as well as expressing their lifeworlds. Phenomenological sociology strives to reveal how actors construe themselves, all the while recognizing that they themselves are actors construing their subjects and themselves.

Phenomenologically understood, society is a fragile human construction, thinly veneered by abstract ideas. Phenomenology itself is evaluatively and politically neutral. Inherently, it promotes neither transformative projects nor stabilization. In the work of a conservatively inclined practitioner, the legitimation process might be supported,

while the liberative practitioner might seek to puncture or debunk the legitimations (Morris 1975). Phenomenology can be used to reveal and endorse the great constructions of humankind or to uncover the theoretical grounds of oppression and repression (Smart 1976). Phenomenologists insist upon the human requirements for meaning, subjective connectedness, and a sense of order. These requirements may be fulfilled within existent or emancipative realities (Murphy 1986).

The phenomenological influence upon contemporary sociology can be seen in the increased humanization of theoretical works, research methods, educational assessment procedures, and instructional modes (Aho 1998; Darroch and Silvers 1982; O'Neill 1985; Potter 1996). Phenomenological thought has influenced the work of postmodernist, poststructuralist, critical, and neofunctional theory (Ritzer 1996). Notions such as constructionism, situationalism, and reflexivity that are at the core of phenomenology also provide the grounds for these recent formulations. For example, the premise of poststructuralism that language is socially constituted denying the possibility of objective meaning is clearly rooted in phenomenology. The procedure known as *deconstruction* essentially reverses the reification process highlighted in phenomenology (Dickens and Fontana 1994). The postmodernist argument that knowledge and reality do not exist apart from discourse is also clearly rooted in phenomenology. Postmodernism's emphasis on the representational world as reality constructor further exemplifies the phenomenological bent toward reflexivity (Bourdieu 1992). On the other hand, phenomenology has been used to reverse nihilistic excesses of postmodernism and poststructuralism (O'Neil 1994). The emphases of the critical school on the constitution of the liberative lifeworld by the autonomous, creative agent via the transcendence of linguistic constraint echoes a theme of phenomenological thought (Bowring 1996). Neofunctionalism, a looser and more inclusive version of its predecessor, finds room for a microsocial foundation focusing on the actor as a constructive agent (Layder 1997).

Phenomenology, while remaining an identifiable movement within the discipline of sociology, has influenced mainstream research. Inclusion of qualitative research approaches in conventional research generally expresses this accommodation

(Bentz and Shapiro 1998). The greater acceptance of intensive interviewing, participant observation, and focus groups reflect the willingness of nonphenomenological sociologists to integrate subjectivist approaches into their work. The study of constructive consciousness as a method of research has broadened and strengthened the standing of sociology in the community of scholars (Aho 1998).

Phenomenology has made a particular mark in the area of educational policy on a number of levels. The flaws of objective testing have been addressed using phenomenological tools. The issue of construct validity, the link between observation and measurement, has been studied ethnographically as a discursive activity to clarify the practices employed by education researchers to establish validity (Cherryholmes 1988). Testing of children has increasingly respected the subjectivity of the test taker (Gilliatt and Hayward 1996; Hwang 1996). Educators are more alert to the need for understanding the learner's social and cognitive processes, for taking into account the constraining parameters of consciousness, and for encouraging self-conscious reflection. Instructional practices that emphasize constructivist approaches have gained great support among professionals and have been broadly implemented to the benefit of learners (Marlowe and Page 1997).

The future impact of phenomenology will depend on its resonance with the needs and aspirations of the rising generations of sociologists. The drive of some among this emerging generation is to examine the obvious with the infinite patience and endurance that is required to come up with penetrating insight. The arena of discourse analysis perhaps holds the greatest promise of this achievement and will likely elicit substantial effort. The phenomenology of emotions also appears to entice young scholars. The reflexive analyses of popular and mediated culture in relation to identity formation will likely draw further interest, as will the study of virtuality, cyberspace, and computer simulcra. The study of children, the family, and education will increasingly be informed by an emphasis on constructive consciousness. Due to its lack of presumption and openness, the phenomenological movement in sociology has proved hardy during the closing decades of the twentieth century and is well situated to encounter the new millennium.

(SEE ALSO: *Ethnomethodology*; *Qualitative Methods*; *Post-modernism*; *Poststructuralism*)

## REFERENCES

- Abercrombie, Nicholas 1980 *Class, Structure and Knowledge*. New York: New York University Press.
- Aho, James A. 1998 *The Things of the World: A Social Phenomenology*. Westport, Conn.: Praeger.
- Ainlay, Stephen C. 1986 "The Encounter with Phenomenology." In James Davison Hunter and Stephen C. Ainlay, eds., *Making Sense of Modern Times: Peter L. Berger and the Vision of Interpretive Sociology*. London: Routledge and Kegan Paul.
- Armstrong, Edward G. 1979 "Black Sociology and Phenomenological Sociology." *Sociological Quarterly* 20:387-397.
- Benner, Patricia 1995 *Interpretive Phenomenology: Embodiment, Caring and Ethics in Health and Illness*. Thousand Oaks, Calif.: Sage.
- Bentz, Valerie Malhotra 1995 "Experiments in Husserlian Phenomenological Sociology." *Studies in Symbolic Interaction* 17:133-161.
- , and Jeremy J. Shapiro 1998 *Mindful Inquiry in Social Research*. Thousand Oaks, Calif.: Sage.
- Berger, Peter L. 1963 *Invitation to Sociology: A Humanistic Perspective*. New York: Anchor.
- 1967 *The Sacred Canopy: Elements of a Sociological Theory of Religion*. New York: Doubleday.
- , and Brigitte Berger 1972 *Sociology: A Biographical Approach*. New York: Basic Books.
- Berger, Peter L., and Hansfred Kellner 1981 *Sociology Reinterpreted*. New York: Doubleday.
- Berger, Peter L., and Thomas Luckmann 1966 *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Doubleday.
- Bleicher, Josef 1982 *The Hermeneutic Imagination: Outline of a Positive Critique of Scientism and Sociology*. London: Routledge and Kegan Paul.
- Blum, Alan 1996 "Panic and Fear: On the Phenomenology of Desperation." *Sociological Quarterly* 37:673-698.
- Blum-Kulka, Shoshana 1994 "The Dynamics of Family Dinner Talk: Cultural Contexts for Children's Passages to Adult Discourse." *Research on Language and Social Interaction* 27:1-50.
- Bogdan, Robert, and Steven J. Taylor 1975 *Introduction to Qualitative Research Methods: A Phenomenological Approach to the Social Sciences*. New York: John Wiley.
- Bowring, Finn 1996 "A Lifeworld without a Subject: Habermas and the Pathologies of Modernity." *Telos* 106:77-104.
- Bourdieu, Pierre 1992 *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.
- Bruyn, Severn T. 1966 *The Human Perspective in Sociology: The Methodology of Participant Observation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Cherryholmes, Cleo H. 1988 "Construct Validity and the Discourses of Research." *American Journal of Education* 96:421-457.
- Cicourel, Aaron V. 1964 *Method and Measurement in Sociology*. New York: Free Press.
- 1973 *Theory and Method in a Study of Argentine Fertility*. New York: John Wiley.
- Costelloe, Timothy M. 1996 "Between the Subject and Sociology: Alfred Schutz's Phenomenology of the Life-World." *Human Studies* 19:247-266.
- Csordas, Thomas J. 1997 *The Sacred Self: A Cultural Phenomenology of Charismatic Healing*. Berkeley: University of California Press.
- Cuff, E. C. 1993 *Problems of Versions in Everyday Situations*. Lanham, Md.: University Press of America.
- Darroch, Vivian, and Ronald J. Silvers (eds.) 1982 *Interpretive Human Studies: An Introduction to Phenomenological Research*. Washington, D.C.: University Press of America.
- Davilla, Roberta A., and Judy C. Pearson 1994 "Children's Perspectives of the Family: A Phenomenological Inquiry." *Human Studies* 17:325-341.
- Denzin, Norman K., and Yvonna S. Lincoln 1994 *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- (eds.) 1998 *Collecting and Interpreting Qualitative Materials*. Thousand Oaks, Calif.: Sage.
- Dickens, David R., and Andrea Fontana (eds.) 1994 *Postmodernism and Social Inquiry*. New York: Guilford Press.
- Encandela, John A. 1997 "Social Construction of Pain and Aging: Individual Artfulness within Interpretive Structures." *Symbolic Interaction* 20:251-273.
- Fielding, Nigel G. (ed.) 1988 *Actions and Structure: Research Methods and Social Theory*. London: Sage.
- Garfinkel, Harold 1967 *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.
- , and Harvey Sacks 1970 "The Formal Properties of Practical Actions." In John C. McKinney and Edward A. Tiryakian, eds., *Theoretical Sociology*. New York: Appleton-Century-Crofts.



- Gilliatt, Stephen E., and Nicholas F. Hayward 1996 "A Testing Time: The Role of Subjective Practices in Making Sense of Student Performance." *Assessment and Evaluation in Higher Education* 21:161-171.
- Grekova, Maya 1996 "Restructuring of 'The Life-World of Socialism.'" *International Sociology* 11:63-78.
- Gubrium, Jaber F., and James A. Holstein 1987 "The Private Image: Experiential Location and Method in Family Studies." *Journal of Marriage and the Family* 49:773-786.
- Heliker, Diane M. 1997 "A Narrative Approach to Quality Care in Long-Term Care Facilities." *Journal of Holistic Nursing* 15:68-81.
- Helle, Horst J. (ed.) 1991 *Verstehen and Pragmatism: Essays in Interpretative Sociology*. Frankfurt, Federal Republic of Germany: Peter Lang.
- Herek, Gregory M. 1986 "The Instrumentality of Attitudes: Toward a Neofunctional Theory." *Journal of Social Issues* 42:99-114.
- Hitzler, Ronald, and Reiner Keller 1989 "On Sociological and Common-Sense *Verstehen*." *Current Sociology* 37:91-101.
- Husserl, Edmund 1931 *Ideas: General Introduction to Pure Phenomenology*. W. R. Boyce Gibson, trans. New York: Humanities Press.
- 1970 *The Crisis of European Sciences and Transcendental Phenomenology*. David Carr, trans. Evanston, Ill.: Northwestern University Press.
- Hwang, Ahn-Sook 1996 "Positivist and Constructivist Persuasions in Instructional Development." *Instructional Science* 24:343-356.
- Ihde, Don 1977 *Experimental Phenomenology: An Introduction*. New York: Putnam.
- Jordan, Brigitte, and Austin Henderson 1995 "Interaction Analysis: Foundations and Practice." *Journal of the Learning Sciences* 4:39-103.
- Knorr-Cetina, Karin, and Michael J. Mulkey (eds.) 1983 *Science Observed*. London: Sage.
- Langsdorf, Lenore 1995 "Treating Method and Form as Phenomena: An Appreciation of Garfinkel's Phenomenology of Social Action." *Human Studies* 18:177-188.
- Layder, Derek 1997 *Modern Social Theory: Key Debates and New Directions*. London: UCL Press.
- Leiter, Kenneth 1980 *A Primer on Ethnomethodology*. New York: Oxford University Press.
- Levesque-Lopman, Louise 1988 *Claiming Reality: Phenomenology and Women's Experience*. Totowa, N.J.: Rowman and Littlefield.
- Livingston, Eric 1995 "The Idiosyncratic Specificity of the Methods of Physical Experimentation." *Australian and New Zealand Journal of Sociology* 31:1-21.
- Luckmann, Thomas (ed.) 1978 *Phenomenology and Sociology: Selected Readings*. New York: Penguin.
- Lynch, Michael 1993 *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press.
- Marlowe, Bruce A., and Marilyn L. Page 1997 *Creating and Sustaining the Constructivist Classroom*. Thousand Oaks, Calif.: Corwin.
- Mehan, Hugh, and Houston Wood 1975 *The Reality of Ethnomethodology*. New York: John Wiley.
- Morris, Monica B. 1975 "Creative Sociology: Conservative or Revolutionary?" *American Sociologist* 10:168-178.
- Murphy, John W. 1986 "Phenomenological Social Science: Research in the Public Interest." *Social Science Journal* 23:327-343.
- Natanson, Maurice 1970 "Alfred Schutz on Social Reality and Social Science." In Maurice Natanson, ed., *Phenomenology and Social Reality*. The Hague: Nijhoff.
- O'Neill, John 1985 "Phenomenological Sociology." *Canadian Review of Sociology and Anthropology* 22:748-770.
- 1994 *The Poverty of Postmodernism*. London: Routledge.
- Peele, Stanton 1985 *The Meaning of Addiction: Compulsive Experience and Its Interpretation*. Lexington, Mass.: Lexington Books.
- Petersen, Eric E. 1987 "The Stories of Pregnancy: On Interpretation of Small-Group Cultures." *Communication Quarterly* 35:39-47.
- Phillipson, Michael 1972 "Phenomenological Philosophy and Sociology." In Paul Filmer, Michael Phillipson, David Silverman, and David Walsh, eds., *New Directions in Sociological Theory*. Cambridge: Massachusetts Institute of Technology Press.
- Porter, Eileen Jones 1995 "The Life-World of Older Widows: The Context of Lived Experience." *Journal of Women and Aging* 7:31-46.
- Potter, Jonathan 1996 *Representing Reality: Discourse, Rhetoric and Social Construction*. Thousand Oaks, Calif.: Sage.
- Psathas, George 1994 *Conversation Analysis: The Study of Talk-in-Interaction*. Thousand Oaks, Calif.: Sage.
- (ed.) 1973 *Phenomenological Sociology: Issues and Applications*. New York: John Wiley.
- Ritzer, George 1996 *Modern Sociological Theory*. New York: McGraw-Hill.

- Rogers, Mary F. 1983 *Sociology, Ethnomethodology, and Experience: A Phenomenological Critique*. New York: Cambridge University Press.
- Rosenberg, Charles E., and Janet Golden (eds.) 1992 *Framing Disease: Studies in Cultural History*. New Brunswick, N.J.: Rutgers University Press.
- Schegloff, Emmanuel, and Harvey Sacks 1974 "Opening Up Closings." In Roy Turner, ed., *Ethnomethodology*. Baltimore: Penguin.
- Schutz, Alfred 1996 *Collected Papers IV*. In Helmut Wagner and George Psathas, eds., in collaboration with Fred Kersten, *Phaenomenologica Series*, vol. 136. Boston: Kluwer Academic.
- 1962 *Collected Papers I: The Problem of Social Reality*. Maurice Natanson, ed. The Hague: Nijhoff.
- 1964 *Collected Papers II: Studies in Social Theory*. Arvid Brodersen, ed. The Hague: Nijhoff.
- 1966 *Collected Papers III: Studies in Phenomenological Philosophy*. I. Schutz, ed. The Hague: Nijhoff.
- 1967 *The Phenomenology of the Social World*. George Walsh and Frederick Lehnert, trans. Evanston, Ill.: Northwestern University Press.
- 1970 *Reflections on the Problem of Relevance*. Richard Zaner, ed. New Haven: Yale University Press.
- , and Thomas Luckmann 1973 *The Structure of the Life World*. Evanston, Ill.: Northwestern University Press.
- Schwartz, Howard, and Jerry Jacobs 1979 *Qualitative Sociology: A Method to the Madness*. New York: Free Press.
- Sheets-Johnstone, Maxine 1996 "An Empirical-Phenomenological Critique of the Social Construction of Infancy." *Human Studies* 19:1-16.
- Shehan, Constance L. (ed.) 1999 *Through the Eyes of the Child: Revisioning Children as Active Agents of Family Life*. Greenwich, Conn.: JAI Press.
- Silverman, David 1998 *Harvey Sacks: Social Science and Conversation Analysis*. New York: Oxford University Press.
- Smart, Barry 1976 *Sociology, Phenomenology, and Marxian Analysis: A Critical Discussion of the Theory and Practice of a Science of Society*. London: Routledge and Kegan Paul.
- Solomon, Robert C. 1997 "Beyond Ontology: Ideation, Phenomenology and the Cross Cultural Study of Emotion." *Journal for the Theory of Social Behaviour* 27:289-303.
- Starr, Jerold M. 1982 "Toward a Social Phenomenology of Aging: Studying the Self Process in Biographical Work." *International Journal of Aging and Human Development* 16:255-270.
- Ten Have, Paul, and George Psathas 1994 *Situated Order: Studies in the Social Organization of Talk and Embodied?* Lanham, Md.: University Press of America.
- Toombs, S. Kay 1995 "The Lived Experience of Disability." *Human Studies* 18:9-23.
- Truan, Franklin 1993 "Addiction as a Social Construction: A Postempirical View." *Journal of Psychology* 127:489-499.
- Truzzi, Marcello 1974 *Verstehen: Subjective Understanding in the Social Sciences*. Reading, Mass.: Addison-Wesley.
- Turner, Roy (ed.) 1974 *Ethnomethodology: Selected Readings*. Baltimore: Penguin.
- Waksler, Frances Chaput (ed.) 1991 *Studying the Social Worlds of Children: Sociological Readings*. London: Falmer Press.
- Wagner, Helmut R. 1983 *Alfred Schutz: An Intellectual Biography*. Chicago: University of Chicago Press.
- Wilson, Tony 1996 "Television's Everyday Life: Towards a Phenomenology of the 'Televisual Subject.'" *Journal of Communication Inquiry* 20:49-66.

## Web Sites of Interest

Center for Advanced Research in Phenomenology: <http://www.fliinet.com/~carp/>

## Internet Resources

[http://mills.edu/ACAD\\_INFO/SOC/SOC153/internet.html](http://mills.edu/ACAD_INFO/SOC/SOC153/internet.html)

MYRON ORLEANS

## POLICE

For modern sociology the core problem of police has been, and continues to be, the extrication of the concept *police* from the forms and institutions in which it has been realized and the symbols and concealments in which it has been wrapped. Doing so is essential to the interpretive understanding of the idea of police and is prerequisite to mature answers to the question of what policing means, has meant, and can mean. In one form or another it is the project that has occupied sociologists of police since the early 1960s, and although there is occasional overlap and interchange, attention to it is primarily what distinguishes contributions to the sociology of police from scholarly efforts in the study of police administration, jurisprudence, criminalistics, and police science.

## THE POLICE: A SOCIOLOGICAL DEFINITION

By the end of the 1960s a small number of now-classic empirical studies of police had made it apparent that conventional understandings of the idea of police were fundamentally and irreparably flawed. In the face of large-scale studies by Reiss (1971) and Black (1971) which showed that the model tour of duty of a patrol officer in the high-crime areas of the nation's largest cities did not involve the arrest of a single person, it became impossible for sociologists to continue to speak of police as "law enforcers" or of their work as "law enforcement." Likewise, both Skolnick's *Justice without Trial* (1966) and Wilson's *Varieties of Police Behavior* (1968) illustrated dramatic differences in the way police were organized and the relationships they elected to enjoy with courts and law. Similarly, early studies of both the exercise of patrol officer discretion (Bittner 1967a, 1967b) and requests for police service (Cumming et al. 1965; Bercal 1970) cast substantial doubt on the notion that a substantial, much less a defining, activity of police was "fighting crime."

**Police Role and Functions.** The task of extricating the concept of police from these common misconceptions was assumed by Egon Bittner in *The Functions of Police in Modern Society* (1970). A fundamental theme of Bittner's work was that to define police as "law enforcers," "peacekeepers," "agents of social control," "officers of the court," or, indeed, in any terms that suppose what police should do, confuses police role and function. Throughout history, in this country and in others, police have performed all sorts of functions. In fact, the functions, both manifest and latent, that police have performed are so numerous and so contradictory that any attempt to define police in terms of the functions they are supposed to perform or the ends they are supposed to achieve is doomed to failure.

**Force as the Core of the Police Role.** Sociologically, policing cannot be defined in terms of its ends; it must be defined in terms of its means. In *Functions* Bittner advanced an approach to understanding the role of the police that was based on the single means that was common to all police, irrespective of the ends to which they aspired or were employed. The means Bittner found to define police was a right to use coercive force. Police,

said Bittner, are "a mechanism for the distribution of non-negotiably coercive force" (1971). No police had ever existed, nor is it possible to conceive of an entity that could be called police ever existing, that did not claim the right to use coercive force.

Sociologically, Bittner's formulation had three major virtues. First, it was universal. It was applicable to police everywhere and at all times, police as diverse as the sheriff's posse of the old West, the London bobby, the FBI, or the police of Hitler's Third Reich or Castro's Cuba. Second, it was politically and morally neutral. It could be used to refer to police whose behavior was exemplary as readily as it could be applied to police whose behavior was appalling. Third, it made it possible to make explicit and to probe in systematic ways a host of questions about the role of police that could not previously be explored because they had been concealed in the confusion between role and function: Why do all modern societies, from the most totalitarian and most tyrannical to the most open and democratic, have police? What does having police make available to society that no other institution can supply? What functions are appropriate to assign to police and what are best left to other institutions?

These questions are of such enormous consequence and so fundamental to an understanding of the role of the police that it is difficult to conceive of a sociology of police existing prior to their recognition.

**Why Police?** If police are a "mechanism for the distribution of non-negotiably coercive force," why should all modern societies find it necessary to create and sustain such a mechanism? What does having such a mechanism make available to modern societies that no other institution can provide?

Bittner's answer is that no other institution has the special competence required to attend to "situations which ought not to be happening and about which something ought to be done NOW!" (1974, p. 30). The critical word in Bittner's careful formulation of the role of the police is "now." What the right to distribute coercive force gives to police is the ability to resolve situations that cannot await a later resolution. The crucial element is time. Turning off a fire hydrant against the wishes of inner-city street bathers, preventing the escape

of a serial murderer, halting the escalation of a domestic dispute, or moving back the curious at the scene of a fire so that emergency equipment can pass—these and hundreds of other tasks fall to police because their capacity to use force may be required to achieve them “now.”

This view of police radically inverts some conventional conceptions. While popular opinion holds that police acquire their right to use coercive force from their duty to enforce the law, the sociology of police holds that police acquire the duty to enforce the law because doing so may require them to invoke their right to use coercive force. Similarly, focus by police on the crimes and misdemeanors of the poor and humble, and their relative lack of attention to white-collar and corporate offenders, is often promoted as reflecting a class or race bias in institutions of social control. While not denying that such biases can exist and do sometimes influence the direction of police attention, if such biases were eliminated entirely, the distribution of police effort and attention would undoubtedly remain unchanged. It would remain unchanged because the special competence of police, their right to use coercive force, is essential in enforcement efforts in which offenders are likely to physically resist or to flee. In white-collar and corporate crime investigations, the special competence of lawyers and accountants is essential, while the special competence of police is largely unnecessary.

### INSTITUTIONAL FORMS

Although all modern societies have found it necessary to create and maintain some form of police, it is obvious that any institution that bears the right to use coercive force is extraordinarily dangerous and highly subject to abuse and corruption. The danger of the institution of police would appear to be magnified when it gains a monopoly, or a near-monopoly, on the right to use coercive force, and when those who exercise that monopoly are almost exclusively direct and full-time employees of the state. Appearances and dangers notwithstanding, these are nevertheless the major terms of the institutional arrangement of police in every modern democracy. Some comment on the sociology of this institutional uniformity may be helpful.

**Avocational Policing.** For most of human history most policing has been done by individuals,

groups, associations, and organizations in the private sector. This type of private-sector policing, done by citizens not as a job but as an avocation, may be classified into at least three types, each of which offered a somewhat different kind of motivation to private citizens for doing it (Klockars 1985). Historically, the most common type is *obligatory avocational policing*. Under its terms private citizens are compelled to police by the threat of some kind of punishment if they fail to do so. In American police history the sheriff's posse is perhaps the most familiar variety of this type of policing. The English systems of frankpledge (Morris 1910) and parish constable (Webb and Webb 1906) were also of this type.

A second type of private-sector policing, *voluntary avocational policing*, is done by private citizens not because they are obliged by a threat of punishment but because they, for their own reasons, want to do it. The most familiar American example of this type of policing is vigilante groups, over three hundred of which are known to have operated throughout the United States up to the end of the nineteenth century (Brown 1975).

A third type, *entrepreneurial avocational policing*, includes private citizens who as English thief takers, American bounty hunters, French agents provocateurs, and miscellaneous paid informants police on a per-head, per-crime basis for money.

The institutional history of these avocational forms of policing is thoroughly disappointing, and modern societies have largely abandoned these ways of getting police work done. The central flaw in all systems of obligatory avocational policing is that as the work of policing becomes more difficult or demanding, obligatory avocational policing takes on the character of forced labor. Motivated only by the threat of punishment, those who do it become unwilling and resistant, a situation offering no one any reason to learn or cultivate the skill to do it well. All forms of voluntary avocational policing suffer from the exact opposite problem. Voluntary avocational police, vigilantes and the like, typically approach their work with passion. The problem is that because the passionate motives of voluntary avocational police are their own, it is almost impossible to control who and where and what form of police work they do and on whom they do it. Finally, the experience with

entrepreneurial forms of avocational policing—thief takers, bounty hunters, and paid informants—has been the most disappointing of all. The abuse and corruption of entrepreneurial avocational police has demonstrated unequivocally that greed is too narrow a basis on which to build a police system.

Sociologically, the shortcoming of all forms of avocational policing is that none of them offers adequate means of controlling the police. This observation leads directly to the question of why one might have reason to suspect that a full-time, paid police should be easier to control than its avocational precedents. What new means of control is created by establishing a full-time, paid, police vocation?

The answer to this problem is that only when policing becomes a full-time, paid occupation is it possible to dismiss, to *fire*, any particular person who makes his or her living doing it. The state can only hire entrepreneurial avocational police, bounty hunters, paid informants, and thief takers; it cannot fire them. Vigilantes are driven by their own motives and cannot be discharged from them. Obligatory avocational police are threatened with punishment if they don't work; most would love to be sacked. Because the option to fire, to take police officers' jobs away from them, is the only essential means of controlling police work that separates the police vocation from all avocational arrangements for policing, how that option is used will, more than anything else, determine the shape and substance of the police vocation.

**The Police Vocation.** The English, who in 1829 created the first modern police, were intimately familiar with the shortcomings of all forms of avocational policing. They had, in fact, resisted the creation of a paid, full-time police for more than a century, out of fear that such an institution would be used as a weapon of political oppression by the administrative branch of government. To allay the fears that the "New Police" would become such a weapon, the architects of the first modern police, Home Secretary Robert Peel and the first commissioners of the New Police, Richard Mayne and Charles Rowan, imposed three major political controls on them. Peel, Mayne, and Rowan insisted that the New Police of London would be unarmed, uniformed, and confined to preventive patrol. Each of these features shaped in profound

ways the institution of the New Police and, in turn, the police of the United States and other Western democracies that explicitly copied the English model.

**Unarmed.** Politically, the virtue of an unarmed police is that its strength can be gauged as a rough equivalent of its numbers. Weapons serve as multipliers of the strength of individuals and can increase the coercive capacity of individuals to levels that are incalculable. One person with a rifle can dominate a dozen citizens; with a machine gun, hundreds; with a nuclear missile, thousands. One person with a police truncheon is only slightly stronger than another, and that advantage can be quickly eliminated by the other's picking up a stick or a stone. In 1829 the individual strength of the 3,000 constable, unarmed New Police offered little to fear to London's 1.3 million citizens.

While this political virtue of an unarmed police helped overcome resistance to the establishment of the institution, the long-run sociological virtue of an unarmed police proved far more important. Policing is, by definition, a coercive enterprise. Police must, on occasion, compel compliance from persons who would do otherwise. Force is, however, not the only means to compel compliance. Sociologically, at least three other bases for control are possible: authority, power, and persuasion.

Unarmed and outnumbered, the New Police "bobby" could not hope to police effectively on the basis of force. Peel, Mayne, and Rowan knew that if the New Police were to coerce successfully, they would have to do so on the basis of popular respect for the authority and power of the institution of which they were a part. The respect owed each constable was not owed to an individual but to a single, uniform temperament, code of conduct, style of work, and standard of behavior that every constable was expected to embody.

In order to achieve this uniformity of temperament, style, conduct, and behavior, the architects of the New Police employed the option to dismiss with a passion. "Between 1830 and 1838, to hold the ranks of the New Police of London at a level of 3300 men required nearly 5,000 dismissals and 6,000 resignations, most of the latter not being altogether voluntary" (Lee 1971; p. 240). During the first eight years of its organization, every position on the entire force was fired or forced to resign more than three times over!

Unlike their earlier London counterparts, the new American police were undisciplined by the firing option. What prevented the effective use of the firing option by early American police administrators was that police positions were, by and large, patronage appointments of municipal politicians. In New York, for example, the first chief of police did not have the right to fire any officer under his command. So while London bobbies were being dismissed for showing up late to work or behaving discourteously toward citizens, American police were assaulting superior officers, taking bribes, refusing to go on patrol, extorting money from prisoners, and releasing prisoners from the custody of other officers.

In New York, Boston, Chicago, and other American cities the modern police began, in imitation of London's bobbies, as unarmed forces; but, being corrupt, undisciplined, and disobedient, they could not inspire respect for either their power or their authority. In controlling citizens they had no option but to rely on their capacity to use force. The difficulty with doing so unarmed is that someone armed with a multiplier of strength can always prove to be stronger. Gradually, against orders, American police armed themselves, at first with the quiet complicity of superior officers and later, as the practice became widespread, in open defiance of departmental regulations. Eventually, in an effort to control the types of weapons their officers carried, the first municipal police agencies began issuing standard service revolvers.

The long-run sociological consequence of arming the American police can be understood only by appreciating how it shaped American police officers' sense of the source of their capacity to control the citizens with whom they dealt. While the London bobbies drew their capacity for control from the profoundly social power and authority of the institution of which they were a part, American police officers understood their capacities for control to spring largely from their own personal, individual strength, multiplied if necessary by the weapon they carried on their hips. This understanding of the source of their capacity for control led American police officers to see the work they did and the choices they made in everyday policing to be largely matters of their individual discretion. Thus, the truly long-run sociological

effect of the arming of the American police has been to drive discretionary decision making to the lowest and least public levels of American police agencies. Today how an American police officer handles a drunk, a domestic disturbance, an unruly juvenile, a marijuana smoker, or a belligerent motorist is largely a reflection not of law or agency policy but of that particular officer's personal style. This is not to say that law or agency policy cannot have influence over how officers handle these types of incidents. However, one of the major lessons of recent attempts by sociologists to measure the impact of changes in law or police policy in both domestic violence and drunken driving enforcement is that officers can resist those changes vigorously when the new law or policy goes against their views of proper police response (see Dunford et al. 1990; Mastroski et al. 1988).

**Uniformed.** Politically, the requirement that police be uniformed is a guarantee that they will not be used as spies; that they will be given information only when their identity as police is known; that those who give them information, at least when they do so in public, are likely to be noticed doing so; and that they can be held accountable, as agents of the state, for their behavior. The English, who had long experience with uniformed employees of many types, understood these political virtues of the uniform completely. In fact, an incident in 1833 in which a police sergeant assumed an ununiformed undercover role resulted in such a scandal that it nearly forced the abolition of the New Police.

By contrast, the early American understanding of the uniform was totally different. Initially it was seen to be a sign of undemocratic superiority. Later it was criticized by officers themselves as a demeaning costume and resisted on those grounds. For twelve years, despite regulations that required them to do so, early New York policemen successfully refused to wear uniforms. In 1856 a compromise was reached by allowing officers in each political ward to decide on the color and style they liked best.

Despite the early resistance to the uniform and the lack of appreciation for its political virtues, American police eventually became a uniformed force. While the London bobby's uniform was explicitly designed to have a certain "homey"

quality and to reflect restraint, the modern American police officer's uniform is festooned with the forceful tools of the police trade. The gun, ammunition, nightstick, blackjack, handcuffs, and Mace, all tightly holstered in shiny black leather and set off with chromium buckles, snaps, badges, stars, flags, ribbons, patches, and insignia, suggest a decidedly military bearing. The impression intended is clearly one not of restraint but of the capacity to overcome the most fearsome of enemies by force.

#### **The Military Analogy and the War on Crime.**

To understand the sociology of the American police uniform, it is necessary to see in it a reflection of a major reform movement in the history of the American police. Around 1890 American police administrators began to speak about the agencies they administered as if they were domestic armies engaged in a war on crime (Fogelson 1977).

The analogy was powerful and simple. It drew upon three compelling sources. First, it sought to connect police with the victories and heroes of the military and to dissociate them from the corruption and incompetence of municipal politics. Second, it evoked a sense of urgency and emergency in calls for additional resources. From the turn of the century to the present day, the war on crime has proved a useful device for getting municipal governments and taxpayers to part with money for police salaries and equipment. Third and most important, the war on crime and the military analogy sought to create a relationship between police administrators and politicians at the municipal level that was similar to the relationship enjoyed by military generals and politicians at the national level. At the national level Americans have always conceded that the decision on whether to fight a war was a politicians' decision, but how that war was to be fought and the day-to-day discipline of the troops was best left to the generals. By getting the public and the politicians to accept these terms of the police-politics relationship, the early police administrators found a way to wrest from the hands of politicians the tool they needed to discipline their troops: the option to fire disobedient officers.

The uniform of the war-ready American police officer is testimony to the fact that since the 1940s, American police administrators have won the battle to conceive of police as engaged in a war

on crime. In doing so they have gained control of the option to fire for administrative purposes. However, the cost of that victory has been enormous.

A major problem is the idea of a war on crime and the expectation police have promoted that they can, in some sense, fight or win it. In point of fact, a war on crime is something police can neither fight nor win for some fundamental sociological reasons. It is simply not within the capacity of police to change those things—unemployment, the age distribution of the population, moral education, civil liberties, ambition and the social and economic opportunities to realize it—that influence the amount and type of crime in any society. These are the major social correlates of crime, and despite presentments to the contrary, police are but a small tail on a gigantic social kite. Moreover, any kind of real “war on crime” is something that no democratic society would be prepared to let its police fight. No democratic society would be able to tolerate the kinds of abuses to the civil liberties of innocent citizens that fighting any real “war” on crime would necessarily involve. It is a major contribution of the sociology of police since the 1960s to demonstrate that almost nothing police do can be shown to have any substantial effect on reducing crime.

The problems of policing in the name of crime when one cannot do much of anything about it are enormous. It is not uncommon for patrol officers to see their employers as hypocritical promoters of a crime-fighting image that is far removed from what they know to be the reality of everyday police work. They may seek to explain what they know to be their failure to do much about crime in terms of the lack of courage of their chief, the incompetence of police administration, or sinister political forces seeking to “handcuff” the police. They often close off what they regard as the disappointing reality of what they do in cynicism, secrecy, and silence—the “blue curtain,” the occupational culture of policing.

Equally problematic as a spoil of the early chiefs' victory in their war on crime is the quasi-military police administrative structure. Once heralded as a model of efficiency, it is now regarded as an organizationally primitive mode of management. It works, to the extent that it works, by creating hundreds and sometimes even thousands

of rules and by punishing departures from those rules severely. The central failing of such an administrative model is that it rests on the unwarranted assumption that employees will not discover that the best way to avoid punishment for doing something wrong is to do as little as possible. The administration can, in turn, respond by setting quotas for the minimum amount of work it will tolerate from employees before it moves to punish them, but if it does so, that minimal amount of work is, by and large, all it will get.

**Preventive Patrol.** The third major mechanism with which architects of the New Police sought to neutralize their political uses was to confine police to preventive patrol. This restriction was understood to have the effect of limiting the uniformed, patrolling constable to two relatively apolitical types of interventions: situations in which constables would be called upon for help by persons who approached them on the street and situations that, from the street, constables could see required their attention. These political virtues of patrol impressed the architects of the New Police, particularly Sir Richard Mayne. Mayne postponed the formation of any detective unit in the New Police until 1842, and during his forty years as commissioner held its ranks to fewer than 15 detectives in a force of more than 3,500.

In the early American experience, uniformed patrol served the principal purpose of imposing some semblance of order on unruly officers. Patrol offered some semblance of assurance that officers could be found at least somewhere near the area to which they were assigned. And while American police created detective forces almost immediately after they were organized, patrol has become in the United States, as in Britain and other modern democracies, the major means of getting police work done.

Sociologically, patrol has had tremendous consequences for the form and substance of policing. It has, for example, been extraordinarily amenable to the three most profound technological developments of the past century: the automobile, the telephone, and the wireless radio. And while there is no evidence that increasing or decreasing the amount of patrol has any influence whatsoever on the crime rate, each of these technological developments has made police patrol more convenient

and more attractive to citizens who wish to call for police service. It is not an exaggeration to say that the vast majority of the activity of most modern police agencies is driven by a need to manage citizen demand for patrol service.

In recent years, attempts to manage this demand have taken many forms. Among the most common are the creation of computer-aided dispatch systems that prioritize the order in which patrol officers are assigned to complaints and increasingly stringent policies governing the types of problems for which police will provide assistance. Also increasingly common are attempts to handle complaints that merely require a written report, by taking that report over the telephone or having the complainant complete a mail-in form. In no small part, such efforts at eliminating unnecessary police response and making necessary police response efficient have produced some of the increasing cost for police labor.

### REORIENTING POLICING

Despite efforts at prioritization, limitation of direct police response, and development of alternative ways of registering citizen complaints, demand for police service continues to grow. Despite the fact that individual citizens appear to want this form of police service more than any other, some contemporary approaches suggest that the entire idea of “dial-a-cop,” “incident-driven” policing requires reconsideration. Two such approaches, “community-oriented policing” (Skolnick and Bayley 1986) and “problem-oriented policing” (Goldstein 1979, 1990), have been advanced as the next generation of “reform” movements in American policing (Greene and Mastrofski 1988).

As theories of police reform, both “problem-oriented” and “community-oriented” policing are grounded in the suspicion that the traditional police response of dispatching patrol officers in quick response to citizen complaints does little to correct the underlying problem that produced the complaint. To some degree at least, this suspicion is confirmed by studies that tend to show that a fairly small number of addresses tend to generate disproportionate numbers of calls for police service, and that patrol officers commonly return to such “hot spots” again and again to attend to similar problems (Sherman et al. 1989).



Both problem-oriented and community-oriented policing offer strategies to deal with such problems that go beyond merely dispatching an officer to the scene. Problem-oriented policing offers a generic, four-step, problem-solving strategy—scanning, analysis, response, and assessment—that police can use to identify problems and experiment with solutions. Community-oriented policing, by contrast, does not offer a mechanism for problem analysis and solution. It is, however, committed to a general strategy that calls for cooperative, police-community efforts in problem solving. In such efforts it encourages the employment of a variety of police tactics—foot patrol, storefront police stations, neighborhood watch programs—that tend to involve citizens directly in the police mission.

While both approaches to reorienting policing have been heralded as revolutionary in their implications for the future of policing, both confront some major obstacles to their realization. The first is that neither problem-oriented nor community-oriented police efforts have been able to reduce the demand for traditional patrol response. Unless that demand is reduced or police resources are increased to allow it to be satisfied along with nontraditional approaches, the community- and problem-oriented policing approaches will most likely be relegated, at best, to a secondary, peripheral role.

The second problem confronting both community- and problem-oriented policing stems from the definition of police and the role appropriate to it in a modern democratic society. The special competence of police is their capacity to use force, and for that reason all modern societies find it necessary and appropriate to have them attend to situations that cannot await a later resolution. Reactive, incident-driven, dial-a-cop patrol is a highly popular, extremely efficient, and (as nearly as possible) politically neutral means of delivering that special competence. To expand the police role to include responsibility for solving the root problems of neighborhoods and communities is an admirable aspiration; but it is a responsibility that seems to go beyond the special competence of police and to require, more appropriately, the special competence of other institutions.

In the past quarter of a century, the scholarly literature on police has grown so dramatically that

a guide is quite helpful if not essential for the reader who seeks to explore the sociology of police. There are seven classic works in the sociology of police that continue to frame most modern dialogues: Michael Bantons's *The Policeman in the Community* (1968), Egon Bittner's *Functions of Police in Modern Society* (1970), Herbert Packer's *The Limits of the Criminal Sanction* (1968), Jerome Skolnick's *Justice without Trial* (1966), William Westley's *Violence and the Police* (1970), and James Q. Wilson's *Varieties of Police Behavior* (1968). These classic works are informed, dramatized, broadened, and otherwise enhanced by a second generation consisting of five extraordinary empirical and theoretical works: Donald Black's *Manners and Customs of the Police* (1980), David Bayley's *Forces of Order: Police Behavior in Japan and the United States* (1976), William Ker Muir, Jr.'s *Police: Streetcorner Politicians* (1977), Peter Manning's *Police Work: The Social Organization of Policing* (1977), and Jonathan Rubinstein's *City Police* (1973).

Beginning in the 1980s, the literature on the sociology of police may be thought of as dividing into three tracks. One track continues the general reflections on the nature of the police institution and its enterprise and includes Richard Ericson's *Making Crime* (1982), Carl Klockars's *The Idea of Police* (1983), Robert Reiner's *The Politics of the Police* (1985), and most recently Richard Ericson and Kevin Haggarty's *Policing the Risk Society* (1997). A second track includes a collection of studies that focus on particular types of or particular problems in policing. Among them are Geoffrey Alpert and Laurie Fridell's *Police Vehicles and Firearms: Instruments of Deadly Force* (1992), Gary Marx's *Undercover* (1988), Maurice Punch's *Conduct Unbecoming: The Social Construction of Police Deviance and Control* (1985), and Lawrence W. Sherman's *Policing Domestic Violence* (1992), as well as two collections, *Modern Policing*, edited by Michael Tonrey and Norval Morris (1992), *Police Violence: Understanding and Controlling Police Abuse of Force*, edited by William A. Geller and Hans Toch (1996). The third track is, of course, the enormous literature on community policing. The collection edited by Jack Greene and Stephen Mastrofski and entitled *Community Policing: Rhetoric or Reality* (1988) is a good guide to the early literature. Herman Goldstein's *Problem-Oriented Policing* (1990) is, of course, essential. Malcolm Sparrow, Mark Moore, and David

Kennedy's *Beyond 911* (1995), and George Kelling and C. Coles's *Fixing Broken Windows* (1996) are two works of some influence in contemporary discussions of the promise of community policing.

(SEE ALSO: *Criminal Sanctions*; *Criminology*; *Social Control*)

## REFERENCES

- Alpert, Geoffrey, and Laurie Fridell 1992 *Police Vehicles and Firearms: Instruments of Deadly Force*. Prospect Heights, Ill.: Waveland.
- Bayley, David 1976 *Forces of Order: Police Behavior in Japan and the United States*. Berkeley: University of California Press.
- Bercal, T. E. 1970 "Calls for Police Assistance: Consumer Demand for Governmental Service." *American Behavioral Scientist* 13, no. 2 (May–August):221–238.
- Bittner, E. 1967a "Police Discretion in Apprehension of Mentally Ill Persons." *Social Problems* 14 (Winter):278–292.
- 1967b "The Police on Skid Row: A Study of Peace Keeping." *American Sociological Review* 32 (October):699–715.
- 1970 *The Functions of Police in Modern Society*. Washington: U.S. Government Printing Office.
- 1974 "Florence Nightingale in Pursuit of Willie Sutton: A Theory of Police." In H. Jacob, ed., *The Potential for Reform of Criminal Justice*. Beverly Hills, Calif.: Sage.
- Black, D. 1971 "The Social Organization of Arrest." *Stanford Law Review* 23 (June):1087–1111.
- 1980 *Manners and Customs of the Police*. New York: Academic.
- Brown, R. M. 1975 *Strain of Violence: Historical Studies of American Violence and Vigilantism*. Oxford: Oxford University Press.
- Cumming, E., I. Cumming, and L. Edell 1965 "Police-man as Philosopher, Guide, and Friend." *Social Forces* 12(3):276–286.
- Dunford, F. W., D. Huizinga, and D. S. Elliott 1990 "The Role of Arrest in Domestic Assault: The Omaha Police Experiment." *Criminology* 28(2):183–206.
- Ericson, Richard 1982 *Making Crime*. Toronto: University of Toronto Press.
- , and Kevin Haggarty 1997 *Policing the Risk Society*. Toronto: University of Toronto Press.
- Fogelson, R. 1977 *Big City Police*. Cambridge, Mass.: Harvard University Press.
- Geller, William, and Hans Toch 1996 *Police Violence: Understanding and Controlling Police Abuse of Force*. New Haven, Conn.: Yale University Press.
- Goldstein, H. 1979 "Improving Policing: A Problem-Oriented Approach." *Crime and Delinquency* 25 (April):236–258.
- 1990 *Problem-Oriented Policing*. New York: McGraw-Hill.
- Greene, J., and S. Mastrofski 1988 *Community Policing: Rhetoric or Reality*. New York: Praeger.
- Kelling, George, and C. Coles 1996 *Fixing Broken Windows*. New York: Free Press.
- Klockars, C. B. 1985 *The Idea of Police*. Beverly Hills, Calif.: Sage.
- Lee, M. 1971 *A History of Police in England*. Montclair, N.J.: Patterson Smith.
- Manning, Peter 1977 *Police Work: The Social Organization of Policing*. Prospect Heights, Ill.: Waveland.
- Mastrofski, S., R. R. Ritti, and D. Hoffmaster 1988 "Organizational Determinants of Police Discretion: The Case of Drunk Driving." *Journal of Criminal Justice* 15:387–402.
- Morris, W. A. 1910 *The Frankpledge System*. New York: Longmans, Green.
- Muir, William, Jr. 1977 *Police: Streetcorner Politicians*. Chicago: University of Chicago Press.
- Packer, Herbert 1968 *The Limits of the Criminal Sanction*. Stanford, Calif.: Stanford University Press.
- Punch, Maurice 1985 *Conduct Unbecoming: The Social Construction of Police Deviance and Control*. London: Tavistock.
- Reiner, Robert 1985 *The Politics of the Police*. Brighton: Weatsheaf.
- Reiss, A. J., Jr. 1971 *Police and the Public*. New Haven, Conn.: Yale University Press.
- Rubinstein Jonathan 1973 *City Police*. New York: Farrar, Strauss and Giroux.
- Sherman, L. W., P. Gartin, and M. E. Buerger 1989 "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." *Criminology* 27:27–55.
- Sherman, Lawrence 1992 *Policing Modern Violence*. New York: Free Press.
- Skolnick, J. K. 1966 *Justice Without Trial*. New York: John Wiley.
- , and D. Bayley 1986 *The New Blue Line*. New York: Free Press.

- Sparrow, Malcolm, Mark Moore, and David Kennedy 1995 *Beyond 911*. New York: Basic.
- Tonrey, Michael, and Norval Morris (eds.) 1992 *Modern Policing*. Chicago: University of Chicago Press.
- Webb, S., and B. Webb 1906 *English Local Government from the Revolution to the Municipal Corporations Act: The Parish and the County*. New York: Longmans, Green.
- Westley, William 1970 *Violence and the Police*. Cambridge: Massachusetts Institute of Technology Press.
- Wilson, J. Q. 1968 *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, Mass.: Harvard University Press.

CARL B. KLOCKARS

## POLISH AND EASTERN EUROPEAN SOCIOLOGY

### CENTRAL AND EASTERN EUROPEAN SOCIOLOGY

There is no doubt that central and eastern European sociologies have similar intellectual, historical, and political roots and can be treated as one block, in contrast to western European sociologies, which are not characterized by uniformity (Nedelmann and Sztompka 1993). This holds true especially for the postwar period in the development of eastern European sociologies. The only exception to this pattern is Polish sociology, which is why we analyze the history and current state of Polish sociology separately. In this brief analysis of central and eastern European sociology we focus on Russian, Hungarian, Czech, Bulgarian, and Rumanian sociologies (see Kolaja and Das 1988; Genov 1989; Keen and Mucha 1994).

Overall, we do not evaluate these sociologies as very impressive, especially in comparison to western European sociologies, on the one hand, and American sociology, on the other. Central and eastern European sociologies have not produced important contributions either to classical tradition or to contemporary sociology. The only contribution to classical world sociology that should be mentioned here comes from Russian tradition and belongs to Pitrim Sorokin (1959, 1962).

The postwar period in the development of eastern and central European sociology is marked

by the imposition of the communist system on the societies in that region. This historical development had overwhelming impact on the development of sociology in these societies. In Russia we witness further expansion of orthodox Marxism—the development that began right after the Bolshevik revolution in 1917, when sociology was removed from universities along with “bourgeois” professors. Historical materialism was proclaimed the only true scientific sociology, whereas the critique of “bourgeois sociology” was the only way of dealing with Western social thought and adopting Western sociological ideas. In the 1950s and 1960s in Russia, a kind of “empirical” Marxist sociology was established. Because of this development, survey research on the conditions of the working class on a large scale was launched and has continued up to this day. This continuing research is atheoretical and purely descriptive.

The so-called Stalinist period, which began right after World War II and lasted until the late fifties, or in some countries even the early sixties, was marked by the almost complete defeat of academic sociology. Sociology was labeled a “bourgeois pseudo-science” (Kolosi and Szelenyi 1993, p. 146), and was abolished as an academic and autonomous discipline. It is hard to overestimate the negative outcomes of this period and the entire period of the communist system in the eastern and central European countries. The development of sociology has been substantially slowed down if not, in some cases, completely stopped. This is why it was only in the 1960s that debates about the scientific character of sociology reemerged in Hungarian and other sociologies in this region. During most of the time after the Stalinist period and until the 1990s, sociologies of this region were trying to free themselves from Marxist ideology, which was not easy since communist regimes always treated sociology as dangerous discipline. These factors are basically responsible for the retardation of these sociologies, as compared to the rest of European sociology. Another important factor that should be mentioned here is the intellectual tradition. As opposed to such countries as Germany, France, Great Britain, and even Poland, the central and eastern European countries have had no tradition of sociological thought.

In this context it is easy to see why, during the communist period and even after the collapse of

the communist system, there was little significant achievement in sociological theory and research in eastern and central European countries.

For example, in Bulgaria it was only in 1985 that sociology began as an academic discipline at few universities. The factor that ignited this development was the fact that Bulgarian Sociological Association organized the VII World Congress of International Sociological Association (ISA) in Varna in 1970. However, the organization of the World Congress of ISA was possible due to purely political decision made by Bulgarian communist government and ISA authorities, but not as the result of advancement of Bulgarian sociology itself.

In Rumania and Czechoslovakia the condition of sociology was very bad, and practically until 1989 the discipline of sociology in these countries was subjected to special controls by the communist regimes. For example, in Czechoslovakia, to have any sort of academic career, communist party membership was required. In East Germany it was only after 1980 that earning a Ph.D. in sociology was allowed for East German academics. Even in the former Yugoslavia, sociology was strictly controlled by the government; this control, in conjunction with the lack of a sociological research tradition, created a situation in which substantial development of the discipline of sociology was very difficult (Keen and Mucha 1994).

Only in Hungary were there important contributions to sociology. These were made by Gyorgy Konrad and Ivan Szelenyi in urban sociology and especially in the sociology of inequalities, classes, and intelligentsia. Their book, *The Intellectuals on the Road to Class Power*, is probably the most famous contribution of Hungarian sociology to world sociological literature (Konrad and Szelenyi 1979). Research on economical sociology by Istvan Gabor, Janos Kornai, and Elmer Hankiss, and on social structure and stratification by Tamas Kolosi, was also significant, not only for our understanding of Hungarian society but also for a general understanding of the phenomena studied.

As mentioned previously, Polish sociology is a special and different case, which is why we treat it separately. There is a rich tradition of sociological thought in Poland. Important contributions to sociology that have significance for this discipline were made by Polish sociologists. Even during the communist period, sociology in Poland remained

relatively free in terms of research, the process of institutionalization of academic life, and contact with Western sociology.

### THE ORIGINS OF SOCIOLOGY IN POLAND

To understand the past and present status of Polish sociology, one should take into account its peculiarity, namely, its particularly tight, intrinsic link with the course of Polish national history, overabundant with uprisings, wars, revivals and transformations. The nineteenth century and the period up to World War II were characterized by the reception of the dominant European trends of social thought. The organicism of Herbert Spencer is reflected in the works of Jozef Supinski (1804–1893), called the founder of Polish sociology. He formulated, for the first time, the problem of the interplay between the nation and the state, which became persistent later on in Polish sociology. Ludwik Gumplowicz (1838–1909) was one of the classic exponents of the conflict tradition and probably the only Polish sociologist of that period who entered the standard textbooks of the history of sociology. He published several works, mainly in German: *Der Rassenkampf* (1883), *Grundriss der Soziologie* (1885), *Die soziologische Staatsidee* (1892), and *Soziologie und Politik* (1892). Gumplowicz's peculiarity consisted in his being an advocate of sociologism before Émile Durkheim. His approach to social life was that the emergence and functioning of social organizations are marked by enduring conflict between social groups, for example, ethnic groups. This is why in some textbooks he is also called a social Darwinist.

The psychologism of Gabriel Tarde and Gustave Le Bon influenced the ideas of Leon Petrazycki (1867–1931), whose three fundamental works were originally published in Russian: *The Introduction to the Study of Politics and Law* (1892), *An Introduction to the Study of Law and Morality* (1905), and *The Theory of Law and State* (1907). His *Die Lehre vom Einkommen* (two volumes, 1893–1895) was published in Berlin. For Petrazycki, observation is a basic method of investigating and studying objects and phenomena. As regards psychic phenomena, the observation consists in self-observation, or introspection. The task of sociology is to detect objective tendencies of social phenomena. Unconscious adaptation processes might be replaced by deliberate steering of man's destiny with the help

of law. The ideal pursued by Petrazycki consisted in the human psyche's being so fitted to the requirements of social life that normative systems (e.g., morality) would prove unnecessary.

Another advocate of psychologism was Edward Abramowski (1868–1918). His main writings include *Individual Elements in Sociology* (1899), and *Theory of Psychological Units* (1899), in which he sketched his theory of sociological phenomenalism. Its main thesis was that the development of societies is based on the constant interaction between objective phenomena and human consciousness, which are causes and effects alternately. In three works—*Problems of Socialism, Ethics and Revolution*, and *Socialism and the State*, all written before 1899 and published in *Social Philosophy: Selected Writings* (1965)—he applied sociological phenomenalism to the analysis of the strategy of class struggle and to the realization of the socialist system. Social revolution should be preceded by “moral revolution”—a deep transformation of conscience. A cooperative is a germ of a socialist society, while the state is its enemy. A cooperative can be transformed into a real republic—a cooperative *res publica*.

Ludwik Krzywicki (1859–1941) was the foremost representative of the first Polish Marxists' generation. Among his works are: *Modern Social Issue* (1888), *Political Economy* (1899), *Sociological Studies* (1923), and *Idea and Life* (1957). Krzywicki was under substantial influence from Darwin and Comte, which led him to the idea of society as a section of natural phenomena and social evolution as a part of universal evolution. The merging of historical materialism with positivistic scientific criteria produced a natural-evolutionistic branch of Marxism comprising the canon of “iron historical laws,” of which Krzywicki himself was the representative. His conception of “historical background” allowed him to invent the original typology of social systems. Also, he was the author of original conception of “industrial feudalism,” being the precursor of the “welfare state” theory.

Another follower of Marx's ideas, Kazimierz Kelles-Krauz (1872–1906), published, among other works, *The Law of Revolutionary Retrospection* (1895), *Sociological Law of Retrospection* (1898), *Economic Basis of Primitive Forms of the Family* (1900), *A Glimpse of XIX Century Sociology* (1901), and *Economic Materialism* (1908). Kelles-Krauz defined the

sociological theory of Marxism as monoeconomism, according to which the whole of social life is determined by the mode of production. However, the central point of his sociological conception was the law of revolutionary retrospection. It referred exclusively to the sphere of social consciousness and was supposed to explain the origins of the revolutionary ideal in a way parallel to monoeconomism: The ideals by which the whole reformatory movement wishes to substitute the existing social norms are always similar to norms from the more or less distant past.

Stefan Czarnowski (1879–1937), in his *Leading Ideas of Humanity* (1928), *Culture* (1938), and *Works*, (2 vols., 1956), continued Durkheim's ideas. *Culture* is Czarnowski's top achievement, in which he claims that culture is the whole of objective elements of social heritage, common for several groups and because of its generality able to expand in space. Czarnowski overcame the dualism of Durkheim's conception, granting both society and culture the character of reality *sui generis*. Characteristic of all these conceptions was the overt impact of the actual sociopolitical conditions on the content of social theory.

The soliology of Florian Znaniecki (1882–1958) and the social (or cultural) anthropology of Bronislaw Malinowski (1884–1942) were different from the above-mentioned bodies of work in at least two respects. First, the works of both Znaniecki and Malinowski gained worldwide recognition; second, both consisted of general conceptions not restricted in scope by particular conditions of time, place, and culture. Znaniecki was coauthor (with W. I. Thomas) of *The Polish Peasant in Europe and America* (1918–1920) and author of numerous books, such as *Cultural Reality* (1919), *Introduction to Sociology* (1922), *The Laws of Social Psychology* (1925), *Sociology of Education* (2 vols., 1928–1930), *The Method of Sociology* (1934), *Social Actions* (1936), *The Social Role of the Man of Knowledge* (1940), *Cultural Sciences* (1952), and the posthumous volume *Social Relations and Social Roles*. He is well known as the author of the concept of “humanistic coefficient,” and of a theoretical system unfolding the postulate of universal cultural order and axionormatively ordered social actions. Bronislaw Malinowski, author of *Argonauts of Western Pacific* (1922), *The Sexual Life of Savages in North-Western Melanesia* (1929), *Coral Gardens and Their Magic* (1935), *A Scientific Theory of Culture and Other Essays* (1944),

and *Freedom and Civilization* (1947), among other works, found world esteem as one of the most influential scholars in establishing the functional approach in cultural anthropology.

#### BETWEEN POST-WAR YEARS AND THE COLLAPSE OF COMMUNISM, 1949–1989

The history of Polish sociology in this period has yet to be written. Among the best attempts to characterize Polish sociology under communism are Wladyslaw Kwasniewicz's articles: "Dialectics of Systemic Constraint and Academic Freedom: Polish Sociology under Socialist Regime," and "Between Universal and Native: The Case of Polish Sociology" (Kwasniewicz 1993, 1994). It was a time when sociology underwent severe criticism (including condemnation in the period 1949–1956), a time of a great shift toward Marxist orientation, but also a time of continuation of traditional lines of theorizing and of implementing in Polish sociology several novelties emerging in Western sociology (especially after 1956, when Polish sociology was brought back to life). The revival and development of Polish sociology was possible at that time thanks to the following outstanding intellectuals of the older generation: Jozef Chalasiniski (1904–1979), who wrote *Young Generation of Peasants* (1938), *Society and Education* (1948), *Young Generation of the Villagers in People's Poland* (a series of volumes, 1964–69), and *Culture and Nation* (1968). He was a prominent student of Polish intelligentsia, peasantry, and youth. Stanislaw Ossowski (1897–1963), who wrote *On the Peculiarities of Social Sciences* (1962) and *Class Structure in the Social Consciousness* (1963, English ed.). The latter contained fresh, stimulating, and critical overviews of theories of both class and social stratification. Maria Ossowska (1896–1974), who wrote *Foundations of the Study of Morality* (1947) and *Social Determinants of Moral Ideas* (1970). Andrzej Malewski (1929–1963), whose work will be mentioned in the next section. Stefan Nowak (1925–1990), whose work will also be mentioned in the next section.

While Ossowski studied class structure and stratification theoretically, Jan Szczepanski initiated empirical research around the problems of the emergence of a socialist-grown working class and an intelligentsia. His book *Polish Society* (1970) summarizes about thirty monographs that emerged from this research project between 1956 and 1965.

The period from 1956 up to the 1970s was certainly the time of a strong group of Marxist sociologists, including among others Zygmunt Bauman, Julian Hochfeld, Wladyslaw Markiewicz, and Jerzy Wiatr.

#### MAJOR CONTEMPORARY CONTRIBUTIONS

The major group of Polish sociologists consequently avoided pure theorizing. However, a large number of works present novel interpretations of contemporary sociological theories. Functionalistic orientation was extensively studied by several sociologists, among them Piotr Sztompka (*System and Function: Toward a Theory of Society*, 1974), and by social anthropologists like Andrzej Paluch (*Conflict, Modernization and Social Change: An Analysis and Critique of the Functional Theory*, 1976). Another extensively studied orientation is interactionist theory by such theoreticians as Marek Czyzewski (*The Sociologist and Everyday Life: A Study in Ethnomethodology and Modern Sociology of Interaction*, 1984), Elzbieta Halas (*The Social Context of Meanings in the Theory of Symbolic Interactionism*, 1987), Zdzislaw Krasnodebski (*Understanding Human Behavior: On Philosophical Foundations of Humanistic and Social Sciences*, 1986), Ireneusz Krzeminski (*Symbolic Interactionism and Sociology*, 1986), and Marek Ziolkowski (*Meaning, Interaction, Understanding: A Study of Symbolic Interactionism and Phenomenological Sociology as a Current of Humanistic Sociology*, 1981).

There are also good examples of innovative works within the domain of conflict theory by Janusz Mucha (*Conflict and Society*, 1978), Marxist theory by Andrzej Flis (*Antinomies of the Great Vision*, 1990), social exchange theory by Marian Kempny (*Exchange and Society: An Image of Social Reality in Sociological and Anthropological Theories of Exchange*, 1988), and "sociological theory of an individual's identity" by Zbigniew Boksanski (*Identity–Interaction–Group: Individual's Identity in Perspective of Sociological Theory*, 1989). Since social anthropology used to be treated in Poland as closely related to sociology, I should mention Piotr Chmielewski's work, *Culture and Evolution* (1988), in which he gives penetrating theoretic insight into evolutionistic theory from Darwin through his own contemporaries, and Zdzislaw Mach's book, *The Culture and Personality Approach in American*

*Anthropology* (1989), which presents a critical evaluation of this influential theoretical paradigm.

Original, creative efforts at the level of history of social thought, metatheory, and sociological theory have been quite substantial in the postwar period. In the domain of history of sociology an important achievement is the monumental, two-volume *History of Sociological Thought* by Jerzy Szacki (1979). The work is not just a simple presentation of theories of significant social thinkers from social philosophy of antiquity to contemporary sociological controversies of the 1970s. Critical analysis of each conception is accompanied by a penetrating account of its place in intellectual history, its relation to other orientations, and its role in the development of the social sciences. It can be said that this work presents the "true history of social thought." The work does not have its equivalent in world literature.

We should also mention another original, creative work in the domain of history of sociology. Edited by Piotr Sztompka, *Masters of Polish Sociology* (1984) presents a comprehensive account of Polish sociology from its beginnings up to martial law in 1981 and after—a period that has been described as initiating a search for a new perspective for Polish sociology.

The greatest achievements in the fields of metatheory and/or philosophy of social sciences include two books by Stefan Nowak and Edmund Mokrzycki. Nowak's book, *Understanding and Prediction: Essays in Methodology of Social and Behavioral Sciences* (1976), can be considered the vehicle by which Polish sociology entered metatheoretical debates of contemporary social sciences as a fully mature partner. Nowak discusses several issues crucial for sociological metatheory, such as the usefulness of the "humanistic coefficient," laws of science versus historical generalizations, inductionism versus deductionism, the time dimension, causal explanations, reduction of one theory to another, and axiomatized theories. The solutions he proposes are novel and enlightening. The same can be said about Edmund Mokrzycki's book, *Philosophy and Sociology: From the Methodological Doctrine to Research Practice* (1983). Mokrzycki argues that since early positivism began circulating in the 1950s as the methodological foundation of sociology, the result has been that empirical sociology

has lost the character of a humanistic discipline without acquiring the status of a true scientific discipline. As a way out, Mokrzycki proposes to put sociology within the framework of a broadly understood theory of culture.

In the field of sociological theory, the following achievements should be pointed out. First, we should mention the theoretical group dealing with class, social structure, and stratification. This group is headed by Włodzimierz Wesolowski, whose *Classes, Strata, and Power* (Wesolowski 1979) serves as their leading theoretical achievement. The crux of the argument is that while, theoretically, relationship to the means of production determines attributes of social position (such as income, work, and prestige), the uniformity of that relationship created under socialism makes the means of production lose their determining properties. In this circumstance, status becomes disengaged from class and tends to "decompose" so that we encounter the phenomenon of "leapfrogging" by groups along certain dimensions. This statement was the point of departure for further studies on meritocratic justice, educational meritocracy, and stratification and structure in comparative perspective (Słomczynski et al. 1981; Słomczynski 1989; Kohn and Słomczynski 1990), social mobility (Wesolowski and Mach 1986), as well as for other studies.

A second group of works deals with problems in the field of sociology but bordering on microsociology and social philosophy. Paweł Rybicki's *The Structure of the Social World* (1979) introduces to sociological debates in Poland, for the first time in a very comprehensive way, problems of micro-macro link, the problematics of a small group, and ontological dilemmas especially related to individualism versus holism controversy. On the other hand, Andrzej Malewski's work (1975), aimed primarily at modification and experimental testing of the social-psychological theories of L. Festinger, M. Rokeach, and N. E. Miller, also undertakes fundamental methodological and theoretical problems of the integration of social sciences, which Malewski tries to solve through the procedure of theoretical reduction. Jacek Szmatka's work *Small Social Structures: An Introduction to Structural Microsociology* (1989), tries to reach virtually all the same goals that his predecessors, mentioned above, tried to reach. The final result of these endeavors

is his structural microsociology, based on assumptions of emergent sociological structuralism, the structural conception of the small group, and specific conception of short- and long-range social structures.

Still another type of theorizing is present in the next two important theoretical works, Sztompka's *Theory of Social Becoming* (1990) and Jadwiga Staniszkis's *The Ontology of Socialism* (1989). The two works are very different in terms of style of theorizing and level of abstraction, but they have one goal in common: to produce theoretical conceptions that would account for tensions, problems, and processes of Polish society. Sztompka, who develops his conception around such categories as human agency and social movements, is highly abstract and stays within the Marxian tradition. Staniszkis engages in her analysis categories such as power, politics, legitimization, and ideology. She is less abstract and refers frequently to concrete societies. However, she too stays within the Marxian tradition.

*Sociology of Culture*, by Antonina Kloskowska (1983), continues vital traditions of this field in Polish sociology and also provides several theoretical innovations. The term *sociology of culture* is understood here to refer to a branch of sociological theory that is culture oriented and that operates with various types of cultural data. Basic subject matter for this theory is symbolic culture, while basic factors are conditions and functions of symbolic culture in the domain of societal culture. Kloskowska develops, among other theories, communication theory of symbolic culture and theory of symbolic culture development; one of her statements is that symbolic culture can perform its functions only when it preserves its original character, and its values remain intrinsic and autotelic, and are sought for their own sake.

The most vital and extensively cultivated, however, is empirical sociology of Polish society. Its standards of research procedures do not differ from western European ones. Polish sociology owes many important methodological and technical improvements to two prominent and in some sense classical methodologists: Stefan Nowak and Jan Lutyński. The empirical branch of Polish sociology is very diversified and multifaceted. Especially extensive are studies on several aspects of

social consciousness of Polish society (continuing Ossowski's 1963 work); its changes in time perspective (Koralewicz 1987), value system, attitudes, aspirations (Nowak 1980, 1982, 1989), class consciousness, and political participation (Ziolkowski 1988); and collective subconsciousness and the concept of collective sense (Marody 1987, 1988). Another important and vital field of research are studies on political and legal system of Polish society (Staniszkis 1987), legitimization of the social order (Rychard 1987), repressive tolerance of the political system (Gorlach 1989), local power elite (Wasilewski 1989), and deviance and social control (Kwasniewski 1987). The third domain of research consists of different aspects of social and economic organization of the Polish society, namely self-management and current economical crisis (Morawski 1987), determinants of economical interests (Kolarska-Bobinska 1988), and conditions of social dimorphism (Wnuk-Lipinski 1987). There are also interesting studies on the life values of youth in Poland (Sulek 1985), the role of the army in the Polish political and social scene (Wiatr 1988), and the birth and the role of the Solidarity movement (Staniszkis 1984). We should also mention the Polish attempt to develop the framework of sociotechnics by Adam Podgorecki (1966), which convinced many academics to switch to the study of practical applications of sociology.

#### **CURRENT TRENDS AND PERSPECTIVES IN POLISH SOCIOLOGY**

The events of 1989, which marked the beginning of an economical, political, and social transformation in Poland, gave Polish sociologists rich empirical material to study. The great interest in research on Polish transition provoked even the suppositions of overpolitization of Polish sociology. Nevertheless, new perspectives appeared in the field of social structure, stratification, and mobility (Wnuk-Lipinski 1989, 1996; Domanski 1994). Attempts to construct a model of the middle class in postcommunist societies have been made by H. Domanski and J. Kurczewski. There is a special focus on the consolidation of young central and eastern European democracies and political and party systems in studies by Ewa Nalewajko (1997). Another well-developed field is the study of power and business elites, their roots and integration by Jacek Wasilewski (Wasilewski 1997, 1998) and



Włodzimierz Wesolowski (1995, 1996). The mainstream of current Polish sociology is focused on society in transition, but this does not mean that there are no pure theoretical studies being in process. The basic science type endeavor is being pursued in the tradition of structural social psychology (i.e., group processes). The theoretical research program in network exchange theory developed by Szmátka, Mazur, and Sozanski is one of the most methodologically advanced in sociology, and their new studies on network conflict theory gained some recognition in world sociology (Szmátka et al. 1997, 1998; Szmátka and Mazur 1998).

## REFERENCES

- Domanski, Henryk 1994 "The Recomposition of Social Stratification in Poland." *Polish Sociological Review* 4:335-358.
- Genov, Nikolai (ed.) 1989 *National Traditions in Sociology*. London: Sage.
- Gorlach, Krzysztof 1989 "On Repressive Tolerance: State and Peasant Farm in Poland." *Sociologia Ruralis* 28:23-33.
- Keen, Mike Forrest, and Janusz Mucha (eds.) 1994 *Eastern Europe in Transformation. The Impact on Sociology*. Westport, Conn.: Greenwood.
- Kloskowska, Antonina 1983 *Sociology of Culture*. Warszawa: Polish Scientific Publishers. (In Polish).
- Kohn, Melvin L., and Kazimierz M. Slomczynski 1990 *Social Structure and Self-Direction: A Comparative Analysis of the United States and Poland*. Cambridge, Mass.: Basil Blackwell.
- Kolaja, Juri, and Man Singh Das (eds.) 1988 *Sociology in Eastern Europe*. India: Books and Periodicals.
- Kolarska-Bobinska, Lena 1988 "Social Interests, Egalitarian Attitudes, and the Change of Economic Order." *Social Research* 55:111-138.
- Kolosi, Tamas, and Ivan Szelenyi 1993 "Social Change and Research on Social Structure in Hungary." In Birgitta Nedelmann and Piotr Sztompka, eds., *Sociology in Europe: In Search of Identity*. New York: Walter de Gruyter.
- Konrad, Gyorgy, and Ivan Szelenyi 1979 *The Intellectuals on the Road to Class Power*. New York: Harcourt Brace Jovanovich.
- Koralewicz, Jadwiga 1987 "Changes in Polish Social Consciousness during the 1970s and 1980s: Opportunism and Identity." In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford: Berg.
- Kwasniewicz, Wladyslaw 1993 "Between Universal and Native: The Case of Polish Sociology." In Brigitta Nedelmann and Piotr Sztompka, eds., *Sociology in Europe: In Search of Identity*. New York: Walter de Gruyter.
- 1994 "Dialectics of Systemic Constraint and Academic Freedom: Polish Sociology under Socialist Regime." In Mike Forrest Keen and Jonusz Mucha, eds., *Eastern Europe In Transformation: The Impact on Sociology*. Westport, Conn.: Greenwood.
- Kwasniewski, Jerzy 1987 *Society and Deviance in Communist Poland: Attitudes to Social Control*. Oxford, England: Berg.
- Malewski, Andrzej 1975 *For a New Shape of the Social Sciences: Collected Papers*. Warszawa: Polish Scientific Publishers. (In Polish).
- Marody, Mirosława 1987 "Social Stability and the Concept of Collective Sense." In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford, England: Berg.
- 1988 "Antinomies of Collective Subconsciousness." *Social Research* 55:97-110.
- Mokrzycki, Edmund 1983 *Philosophy and Sociology: From the Methodological Doctrine to Research Practice*. London: Routledge and Kegan Paul.
- Morawski, Witold 1987 "Self-Management and Economic Reform." In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford, England: Berg.
- Nalewajko, Ewa 1997 *Proto-Parties and Proto-System? An Outline of Emerging Polish Multipartism*. Warszawa: Instytut Studiów Politycznych PAN (In Polish).
- Nedelmann, Birgitta, and Piotr Sztompka (eds.) 1993 *Sociology in Europe: In Search of Identity*. New York: Walter de Gruyter.
- Nowak, Stefan 1976 *Understanding and Prediction: Essays in Methodology of Social and Behavioral Sciences*. Dordrecht: Reidel.
- 1980 "Value System of the Polish Society." *Polish Sociological Bulletin* 2:5-20.
- 1982 "Value System and Social Change in Contemporary Poland." *Polish Sociological Bulletin* 1-4:119-132.
- 1989 "The Attitudes, Values and Aspirations of Polish Society." *Sisyphus: Sociological Studies* 5:133-163.
- Podgorecki, Adam 1966 *Principles of Sociotechnics*. Warszawa: Wiedza Powszechna. (In Polish).

- Rybicki, Pawel 1979 *The Structure of the Social World*. Warszawa: Polish Scientific Publishers.
- Rychard, Andrzej 1987 "The Legitimation and Stability of the Social Order in Poland." In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford, England: Berg.
- Slomczynski, Kazimierz M. 1989 *Social Structure and Mobility: Poland, Japan, and the United States*. Warszawa: Polish Academy of Sciences, Institute of Philosophy and Sociology.
- , Joanne Miller, and Melvin L. Kohn 1981 "Stratification, Work, and Values: A Polish-United States Comparison." *American Sociological Review* 46:720–744.
- Sorokin, Pitirim A. 1959 *Social and Cultural Mobility*. Glencoe, Ill.: Free Press.
- 1962 *Social and Cultural Dynamics*, 4 vols. Englewood Cliffs, N.J.: Bedminster Press.
- Staniszki, Jadwiga 1984 *Poland: Self-Limiting Revolution*. Princeton, N.J.: Princeton University Press.
- 1987 "The Political Articulation of Property Rights: Some Reflections on the University 'Inert Structure.'" In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford, England: Berg.
- 1989 *The Ontology of Socialism*. Warszawa: In Plus. (In Polish).
- Sulek, Antoni 1985 "Life Values of Two Generations: From a Study of the Generational Gap in Poland." *Polish Sociological Bulletin* 1–4:31–42.
- Szacki, Jerzy 1979 *History of Sociological Thought*. Westport, Conn.: Greenwood.
- Szczepanski, Jan 1970 *Polish Society*. New York: Random House.
- Szmatka, Jacek 1989 *Small Social Structures: An Introduction to Structural Microsociology*. Warszawa: Polish Scientific Publishers. (In Polish).
- , and Joanna Mazur 1998 "Power Distribution in Conflict Networks: An Extension of Elementary Theory to Conflict Networks." In John Skvoretz and Jacek Szmatka, eds., *Advances in Group Processes*, vol. 15. Greenwich, Conn.: JAI.
- Szmatka, Jacek, John Skvoretz, and Joseph Berger (eds.) 1997 *Status, Network, and Structure: Theory Development in Group Processes*. Stanford, Calif.: Stanford University Press.
- Szmatka, Jacek, John Skvoretz, Tadeusz Sozanski, and Joanna Mazur 1998 "Conflict in Networks." *Sociological Perspectives* 41:49–66.
- Sztompka, Piots 1990 *Theory of Social Becoming*. Cambridge, England: Polity.
- (ed.) 1984 *Masters of Polish Sociology*. Krakow: Ossolimeum. (In Polish).
- Wasilewski, Jacek 1989 "Social Processes of Power Elite Recruitment." *Sisyphus: Sociological Studies*. 5:205–224.
- 1997 "Elite Research in Poland: 1989–1995." In Heinrich Best and Ulrike Becker, eds., *Elites in Transition: Elite Research in Central and Eastern Europe*. Opladen, Germany: Leske and Budrich.
- 1998 "Elite Circulation and Consolidation of Democracy in Poland." In John Higley, Jan Pakulski, and Wlodzimierz Wesolowski, eds., *Postcommunist Elites and Democracy in Eastern Europe*. London: Macmillan.
- Wesolowski, Wlodzimierz 1979 *Classes, Strata, and Power*. London: Routledge and Kegan Paul.
- 1995 "Formation of Political Parties in Post-Communist Poland." *Sisyphus. Social Studies* 11(1):9–32.
- 1996 "New Beginnings of the Entrepreneurial Classes. The Case of Poland." *The Polish Sociological Review* 1(113):79–96.
- , and Bogdan W. Mach 1986 "Unfulfilled Systemic Functions of Social Mobility I: A Theoretical Scheme." *International Sociology* 1:19–35.
- Wiatr, Jerzy 1988 *The Soldier and the Nation*. Washington: Westview.
- Wnuk-Lipinski, Edmund 1987 "Social Dimorphism and Its Implications." In J. Koralewicz, I. Bialecki, and M. Watson, eds., *Crisis and Transition: Polish Society in the 1980s*. Oxford, England: Berg.
- 1989 "Inequality and Social Crisis." In Wladyslaw Adamski and Edmund Wnuk-Lipinski, eds., *Poland in the 1980s: Reassessment of Crises and Conflicts*. Warszawa: PWN.
- 1996 *Democratic Reconstruction: From Sociology of Radical Social Change*. Warszawa: PWN. (In Polish).
- Ziolkowski, Marek 1988 "Individuals and the Social System: Values, Perceptions, and Behavioral Strategies." *Social Research* 55:139–177.

KINGA WYSIENSKA  
JACEK SZMATKA

## POLITICAL AND GOVERNMENTAL CORRUPTION

Everyone knows what political corruption is, but it is notoriously hard to define. Different cultures

have different conceptions of corruption: what would be considered corrupt in Denmark might be seen as simply polite in Indonesia. This understanding also varies across time: buying office was standard procedure in eighteenth-century Britain but would be inexcusable today. Arnold Heidenheimer has divided corruption into three categories (Heidenheimer et al. 1978). “*White*” corruption includes acts that a majority of people would not consider worthy of punishment. “*Gray*” corruption includes acts that “some elements” would want to see punished, but others would not. “*Black*” corruption includes acts that a “majority consensus . . . would condemn and would want to see punished on the grounds of principle.” Heidenheimer suggests that as societies modernize, behavior that was once seen as “white” becomes “gray,” and may eventually turn “black.” Others have suggested that the concept of “corruption” develops as societies move from seeing governmental offices as private property to believing them to be public trusts.

The American political scientist V. O. Key Jr. (1936) defines *graft* as “an abuse of power for personal or party profit.” Joseph Nye, another American political scientist, calls corruption:

*[B]ehavior which deviates from the formal duties of a public role because of private-regarding (family, close family, private clique) pecuniary or status gains; or violates rules against the exercise of certain types of private-regarding influence. This includes such behavior as bribery (use of a reward to pervert the judgment of a person in a position of trust); nepotism (bestowal of patronage by reason of ascriptive relationship rather than merit); and misappropriation (illegal appropriation of public resources for private-regarding uses) (Nye 1967).*

One long-running controversy in political science is whether corruption can serve a useful purpose for society. The most traditional school of thought, the “moralist” school, insists that corruption is always a breach of the public trust and must never be tolerated. Foremost in the ranks of this school were the “muckrakers,” American journalists who wrote books, during the early twentieth century, denouncing big business and public corruption. But in a review of a muckraking classic,

Lincoln Steffens’s *The Shame of the Cities*, the American political scientist Henry Jones Ford (1904, p. 678) argued that corruption was better than “slackness and decay.” If businessmen had to corrupt politicians in order to get things done, that was an acceptable cost of progress. Ford’s heirs, the so-called revisionists, became important in social science during the 1960s. Such revisionists as Nye and Samuel Huntington argued that corruption could serve useful social purposes. Corruption could bridge gaps between otherwise antagonistic groups. It could ease the path to modernization by overcoming bureaucratic inertia. It could foster social integration by binding people closer to the state.

More recently, scholars have returned to emphasizing the ill effects of corruption. It can increase the costs of administration, bloating national budgets. Rather than cutting bureaucracy, corruption tends to expand it, as officials seek more opportunities to shake down citizens. If businessmen must bribe in order to function in a particular country, they may take their investments elsewhere. Corruption can weaken the private sector into dependence and degrade the public sector into banditry. The economic collapse of Indonesia, a country previously seen as an example of the coexistence of corruption and prosperity, has fueled this school of thought. Recent research by Paolo Mauro (1997), an economist at the International Monetary Fund, has shown that highly corrupt countries tend to invest less of their gross domestic product, spend less on education, and have lower economic growth rates than their cleaner counterparts.

#### CORRUPTION IN THE UNITED STATES

The United States is less corrupt than almost all Third World countries. Its capitalist economy and democratic government offer fewer opportunities for massive political graft than the socialist kleptocracies of Africa. American government’s dispersal of power prevents the sort of presidential corruption that has characterized Latin America. The American press, its freedom constitutionally guaranteed, is quick to jump on accusations of malfeasance, often blowing them out of proportion. The American legal system has inherited the English tradition of protecting private property and respecting individual rights.

But for most of its history, the United States has had a reputation of being more corrupt than most western European countries. British observers such as Lord Bryce expressed this sentiment most often, contrasting the probity of the Victorian era to the iniquity of the Gilded Age. The exuberant democracy of the nineteenth century spawned extensive patronage, widespread vote fraud, and a venal political class. Indeed, Bryce and Alexis de Tocqueville noted that the most able Americans went into business rather than government; when they needed political favors, they bought them from those lesser talents who had won office. Civil servants continue to have a lower status in the United States than in Europe. In addition, the American system is much more permeable than its European counterparts. Its numerous state and local governments encourage popular participation. They also give citizens more chances to corrupt bureaucrats.

As a general rule, corruption in the United States has been greater at the state and local levels than in the federal government. While no president has ever been imprisoned, many mayors and governors have. Fixated on Washington, the media often ignores what happens in city halls or, especially, state capitols. This gives boodlers more confidence that their misdeeds will go unpunished. Since city governments conduct most law enforcement, more scandals take place in local police departments than in the Federal Bureau of Investigation (FBI). Since state and local governments conduct most construction, contractors develop more shady relationships with governors or mayors than with presidents. Many defense firms, however, have cozy ties with Congress members and Pentagon officials.

While there was corruption in colonial America (even George Washington used liquor to buy votes during his candidacy for the Virginia legislature), municipal graft did not truly take off until the nineteenth century. Powered by the Industrial Revolution, and fed by waves of immigration, American cities boomed during this period. While governmental institutions were weak, a new form of organization filled this gap: the political machine. Michael Johnston defines a political machine as a “party organization within which power is highly centralized, and whose members are motivated and rewarded by divisible material incentives rather than by considerations of ideology or

long-term goals of public policy” (1982, p. 38). While they would sometimes advocate social reform if they believed it would win votes, machine politicians were primarily interested in the spoils of office: jobs and favors for their supporters and graft for themselves. Machines developed close ties to business interests seeking contracts or other concessions: construction contractors, real estate developers, insurance agencies, law firms, organized crime. Ward leaders mobilized machine supporters, often the immigrant poor, through small favors such as Christmas baskets; the fixing of minor crimes; or, for the most loyal, undemanding government jobs.

While the stereotypical machine was Democratic, there were many Republican machines until the New Deal made urban workers loyal Democrats. City machines were the most common, but there were rural machines (most notably in southern Texas), suburban machines, and even statewide machines. While one normally associates machine politics with the inner-city poor, one of the strongest remaining machines in the country is the Republican organization of Nassau County, New York—a New York City suburb that is one of the wealthiest counties in the United States. (It is probably best known as the political base of former U.S. Senator Alfonse D’Amato.)

Probably the most famous machine was Tammany Hall, which dominated Manhattan’s Democratic organization from the early nineteenth century until the early 1960s. Descended from the Sons of Liberty of the Revolutionary era, it rose to supremacy after the Civil War under the leadership of William “Boss” Tweed and Richard Croker, and reached its zenith of power around World War I under the guidance of the unusually shrewd Charles F. Murphy. In the 1930s, Murphy’s successors made a series of miscalculations that crippled Tammany. They alienated both President Franklin D. Roosevelt and Governor Herbert Lehman, and they supported John O’Brien, an inept hack, as the successor to the discredited Mayor James J. Walker. O’Brien lost the 1933 mayoral race to Fiorello LaGuardia, a nominal Republican, staunch New Dealer, and sworn Tammany foe. LaGuardia’s three terms in office helped pound the nails into Tammany’s coffin. Other Democratic organizations in New York City, particularly those in the Bronx and Brooklyn, became more important.

While Tammany enjoyed a brief comeback in the 1950s under Carmine DeSapio, its fate was sealed when DeSapio lost his race for Democratic district leader in 1961. Manhattan had become gentrified, and its new, wealthier residents had no interest in Tammany-style politics.

Many cities had machines. Until the New Deal, Chicago had rival Democratic and Republican organizations whose battles for power occasionally turned violent. After the New Deal, the Democratic organization dominated Chicago and some of its leaders, especially Mayor Richard J. Daley (1955–1976), became national figures. Philadelphia had a Republican machine of legendary rapaciousness that managed to hold onto City Hall until 1951. Once it became clear that the GOP would not return to power, many longtime Republican ward bosses switched parties and built a powerful Democratic machine. Kansas City had its Pendergast organization, which produced President Harry Truman. Jersey City's two machine mayors, Frank "I Am the Law" Hague and John V. Kenny, dominated its politics from 1917 to 1971. Machines were not confined to the North: Edward "Boss" Crump ruled Memphis for many years, while San Antonio became notorious for its vote buying and graft.

Political machines have declined throughout the twentieth century. Prompted by Lincoln Steffens's *The Shame of the Cities* (1904), Progressive reformers strengthened the civil service and reduced machine influence in city government, sometimes by making elections nonpartisan or by replacing elected mayors with the council-manager system. While Franklin D. Roosevelt rewarded those machine Democrats who backed him (most notably Chicago Mayor Edward Kelly and Edward Flynn, "the Boss of the Bronx"), some observers have argued that his New Deal eventually undermined machines by replacing their handouts with bureaucratic programs (although Richard J. Daley could show them how those could be turned to a machine's advantage).

The postwar years brought many changes to American cities that hurt machines. Television gave candidates a means to reach voters without the help of local bosses. Court decisions greatly restricted the amount of patronage leaders could bestow upon their followers. The rise of public

sector unions further restricted bosses' ability to hand out jobs, and in some cities, unions replaced party bosses as the main support for politicians. Inquiries such as the 1951 Kefauver hearings exposed the links between many urban politicians and organized crime.

Political scientists and journalists have offered varying assessments of political machines. Some praise them for integrating immigrants into American politics. Many of the newcomers had little experience with democracy and little understanding of the Anglo-American tradition. Machine politicians gave them something concrete for their vote, whether that was a small gift, help with the law, or a job. Machines have also been praised for making government function; Richard J. Daley boasted that Chicago was "the city that works." Others have criticized machines for corrupting government, enriching political hacks, and providing poor public services.

While few machines still operate, American cities still see plenty of graft. Periodic scandals rock urban police departments, most recently in New Orleans and Washington. Frequently, there are revelations about the scandalous relations between construction firms and politicians. In the mid-1980s, a scandal concerning the Parking Violations Bureau ended the careers of many New York City politicians. Marion Barry, the four-term mayor of Washington, made himself a synonym for graft and misgovernment.

There have also been many scandals concerning state government. During the late nineteenth century, businessmen, especially railroad magnates, bribed state legislators en masse. From Arizona to South Carolina, recent sting operations have found state legislators pathetically eager to accept pay-offs. The cozy relationships between state officials and the construction industry have had national implications. In 1973, Vice President Spiro Agnew resigned as part of a plea bargain stemming from bribes he accepted as governor of Maryland.

There are large regional differences in attitudes toward corruption. What is considered scandalous in Minnesota may be seen as merely inappropriate in New York and harmless in Louisiana. Daniel Elazar argues for a theory of political culture that explains these differences. Puritan New England developed a "moralistic" conception of

governance that emphasized the common good. When New Englanders settled the states of the northern tier, from Michigan to the Pacific Northwest, they brought this culture with them. These states continue to be intolerant of corruption. The Middle Colonies attracted a diverse ethnic mix that made it difficult to agree on a common good. Instead, they developed an “individualistic” political culture that saw politics as a business and tolerated politicians’ pursuit of self-interest. States in a belt running west from New York and New Jersey to Illinois and Missouri continue to accept petty graft as part of the normal way of conducting politics. The South developed politics dominated by elite factions and low public participation. Elazar calls this culture “traditionalistic.” Particular states have developed their own cultures tolerant of or antagonistic to corruption. The Scandinavians who settled the upper Midwest brought their squeaky-clean politics with them. Louisiana’s French heritage bestowed upon it the Napoleonic Code, with its emphasis on state power and a certain lassitude about graft.

#### **CORRUPTION AT THE FEDERAL LEVEL**

While there were scandals during the Republic’s early years, corruption did not become a major problem until the Jacksonian era of the 1820s and 1830s. Mass political parties brought with them the “spoils system” through which the president bestowed government jobs upon his supporters. The “spoils system” came to dominate government to the extent that some presidents spent most of their time fulfilling the demands of state party leaders for rewards for their followers. After a disappointed office seeker assassinated President James Garfield in 1881, demands for reform led Congress to pass the Pendleton Act two years later, which created the modern civil service; but it was not until the early twentieth century that merit came to predominate in government employment.

The Industrial Revolution increased business influence in government. The Civil War saw a series of military procurement scandals, beginning a long tradition in American life. During the years after the war, there were several scandals concerning the relationships between politicians and railroads (many of which received government subsidies). In addition, during this period,

many prominent officials were found to have taken bribes, most notably William Belknap, Ulysses Grant’s secretary of war.

While the Progressive era saw a general reform of American politics, the United States experienced several scandals during the 1920s. The sale and production of alcoholic beverages was forbidden during the Prohibition era (1920–1933), but the public’s continuing thirst spawned a vast underground liquor trade that fueled the rise of organized crime and corrupted many public officials. The administration of President Warren Harding (1921–1923) is usually considered the most corrupt in American history. Charles Forbes, head of the Veterans’ Bureau, was convicted of fraud, conspiracy, and bribery. In the “Teapot Dome” scandal, Interior Secretary Albert Fall secretly leased federal oil reserves to two businessmen in return for cash gifts and no-interest loans; Fall was eventually imprisoned for bribery.

The succeeding decades saw few high-level scandals. There were many congressional peccadilloes, but the worst that could be discovered about any high officials was influence peddling by some of Harry Truman’s cronies and acceptance of small gifts by Dwight Eisenhower’s chief of staff. The Watergate scandal (1972–1974) made up for lost time, bringing down President Richard Nixon and sending many of his closest allies to jail. While it is beyond the scope of this article to explain the details of this affair—there are entire books that do that—we should note that little about Watergate concerned financial corruption, primarily Nixon’s tax evasion and massive illegal contributions to his 1972 reelection campaign. Watergate was mostly a matter of abuse of power, especially Nixon’s use of the government to harass his political enemies and his cover-up thereof.

Watergate led to a flood of legislation. The 1971 Federal Election Campaign Act was extensively amended in 1974. The Ethics in Government Act of 1978 created the institution of the independent counsel to investigate official malfeasance. The Foreign Corrupt Practices Act of 1977 forbade American companies to bribe overseas officials. Prosecutors and journalists became more aggressive in their investigations of corruption. Two major scandals rocked Congress: the revelation of corrupt links between a South Korean lobbyist and several congressmen (“Koreagate”)

and the willingness of some senators and representatives to take bribes in a sting investigation (“Abscam”).

The 1980s saw their share of scandals. The Iran-Contra affair, which involved the covert sale of arms to Iran and the illegal diversion of the profits to Nicaraguan anti-communist rebels (the “contras”), tarnished the last years of the Reagan Administration. A long investigation found favoritism and influence peddling to be rife within the Department of Housing and Urban Development (HUD). The collapse of many savings and loans (S&Ls), which cost the government \$500 billion to pay back depositors, revealed suspicious relationships between some thrift operators and prominent politicians, most notably House Speaker Jim Wright. The so-called Keating Five senators lobbied for easier treatment for S&L owner Charles Keating; revelations about their activities ended the careers of three of the five.

While Bill Clinton promised to run the “most ethical administration in history,” he has instead presided over an era as sleazy as those of his predecessors. Scandals forced out HUD Secretary Henry Cisneros and Agriculture Secretary Mike Espy. There were charges that foreigners made large illegal donations to the Democratic National Committee during the 1996 election. The most explosive charges surrounded Clinton himself. Allegations of corrupt real estate deals by Clinton when he was governor of Arkansas led to an independent counsel investigation that began in 1993 and continues as of this writing. This investigation eventually produced the charge that Clinton had perjured himself and had obstructed an inquiry into his affair with a White House intern. Clinton was questioned about this affair as part of a sexual harassment lawsuit. Clinton was eventually impeached, but the Senate failed to convict him.

#### **CORRUPTION IN WESTERN EUROPE**

Italy is probably considered the most corrupt country in western Europe. Southern Italy has a long history of patron-client relationships that take precedence over written law. In this highly stratified society, landowners and professionals serve as natural patrons to their clientele of laborers and peasants. The Mafia, of course, is part of this system of

patronage. Italy’s complex administrative law undermines efficiency, and so requires mediation in order to function. Politicians and local notables help citizens, especially businessmen, navigate the maze of red tape—for a price.

After World War II, Italy developed a strong party system that permeated all of society. While Italy had many small parties, three dominated politics until 1993: the Christian Democrats, the Socialists, and the Communists. In order to keep the Communists out of power, the Christian Democrats dominated every government from 1947 to 1993; after 1962, they usually governed with the support of the Socialists. Both the Christian Democrats and the Socialists were deeply factionalized, with politicians commanding their own followings. As such, the parties could only be held together through bargaining—which often included graft. The Christian Democrats, in particular, were dependent upon the support of local leaders in the South and in Sicily (which were both far behind the rest of country economically), many of them with ties to the Mafia. Southern politicians used the national government’s resources to support their local economies. Often this meant public works boondoggles, which enriched politically connected (and Mafia-owned) construction firms and which only increased the contempt held by northern Italians for their southern countrymen.

Because many Italians feared that a communist government would bring their country into the Soviet bloc, the Christian Democrats were essentially a perpetual governing party. Proportional representation meant that no party would ever gain a majority, so coalition governed Italy. While squabbling among factions meant that particular governments usually lasted a year or less, the same parties always came out on top. The Christian Democrats and the Socialists used their power to create a vast web of patronage, which enmeshed Italy’s many state-owned firms, the national broadcasting system, executive agencies, and public banks. The lack of turnover loosened what few inhibitions Italian politicians had about exploiting their positions for financial gain. While the Communists remained out of power, they cut backroom deals with the ruling parties to enrich themselves, especially after the 1970s, when they distanced themselves from the Soviet Union and when regional autonomy gave Communists more chances to govern.

These arrangements collapsed in 1992–1993. There were a number of reasons for this change. Under the pressure of global competition, Italian business owners grew weary of the high taxes and inefficient government produced by widespread corruption. The drive for European monetary union required Italy to reduce its huge budget deficit. The end of the Cold War made communist control less unthinkable (especially when the party renounced communism and renamed itself the Democratic Party of the Left).

It was a team of prosecutors investigating corruption in Milan's construction industry who really rocked Italy's political world. Led by Antonio Di Pietro (who now sits in Italy's Senate and may be the next president), the so-called clean hands brigade demolished the Italian establishment during 1992–1994. Prosecutors exposed massive patronage, political awarding of contracts, and bribery. It became clear that organized crime had friends at the highest levels. Former Prime Minister Bettino Craxi, a Socialist, was convicted of corruption after he fled to Tunisia. Giulio Andreotti, a Christian Democrat who had served as prime minister seven times, is, as of this writing, on trial for being a member of the Mafia.

The corruption investigations touched even those who benefited from them. After the Italian party system collapsed in 1993, the media tycoon Silvio Berlusconi led a conservative coalition to victory in the March 1994 elections. Berlusconi's government collapsed when prosecutors announced that they were investigating him. In 1998, Berlusconi was convicted of bribing tax inspectors (although he was later acquitted of other charges), but he remains leader of his Forza Italia party.

Spain also has a long history of corruption and patronage. Governments have long relied on local notables to administer the law. While Francisco Franco strengthened the state during his long rule (1939–1975), he made appointments more on the basis of loyalty than competence. The return of democracy brought greater professionalism, but local patrons remain important. It was corruption on a national scale that grabbed the headlines in the 1990s. A series of scandals embarrassed the socialist government of Prime Minister Felipe Gonzalez, who led Spain from 1982 to 1996. Many observers attributed the corruption to the socialists' arrogance after so many years in power. In

1991, Deputy Prime Minister Alfonso Guerra resigned after his brother Juan was accused of using party facilities for private gain. A series of scandals surrounded Socialist party finances. The heads of the Bank of Spain and the Civil Guard were besmirched by corruption. This wave of "sleaze" helped lead to the socialists' defeat in 1996.

France has had many scandals throughout its history. Some have even threatened the stability of the regime. Conservatives opposed to the Republic used the Panama Canal scandal (1892–1993) and the Stavisky affair (1934) to attack the legitimacy of democratic government. Others have simply been embarrassing, as when it was revealed that President Valéry Giscard d'Estaing had accepted diamonds from the infamous tyrant Jean-Bedel Bokassa, president of the Central African Republic and a longtime French client. Overall, the French tend to have a cynical view of government, accepting corruption and the abuse of power as the way the world works.

The "long presidency" of François Mitterand (1981–1995) included a series of scandals, many of them concerning the finances of Mitterand's Socialist Party. Rising campaign costs forced socialists, who lacked conservatives' constituency among business, to raise money in irregular ways. Contractors who sought business with socialist-governed municipalities were told to contribute to party-controlled consultancies, which funneled the money to the national party. The exposure of this practice embarrassed the socialists, but they could plausibly claim that other parties engaged in similar schemes. The businessman Roger-Patrice Pelat, who was indicted for insider trading before his death in 1989, had questionable relationships with many politicians, including his close friend Mitterand and Prime Minister Pierre Bérégovoy, who committed suicide in 1993 after it was revealed that Pelat had given him an interestfree loan.

Two other leading French politicians have come under fire in recent years. Alan Juppé, a conservative former prime minister, has been charged with obtaining fraudulent jobs for Gaullist party activists when he worked for then-Paris Mayor Jacques Chirac, who is now president. Roland Dumas, a former foreign minister and chief of the Constitutional Court, is under investigation for his ties to a defense contractor, which hired his mistress and paid her a \$9 million commission.



Greece is another southern European country notorious for corruption and clientelistic politics. The governments of Andreas Papandreu (1981–1989, 1993–1996), longtime leader of the Pan-Hellenic Socialist Movement (PASOK), were widely perceived as crooked and patronage ridden, but Papandreu himself was acquitted of embezzlement charges in 1992. Tax evasion is common, and the economy remains dependent upon state patronage and contracts.

Many of the countries of northern Europe have become bywords for clean government. The governments of Scandinavia, the Netherlands, and Switzerland, in particular, are known for their honesty. A typical “scandal” in these countries would be a politician using party funds to pay for diapers for her baby. One major exception is Belgium, which is known as the “Italy of the North” for its corruption and cronyism. The government was shaken by a case of the pedophile murderer Marc Dutroux, who had conducted his activities in such a way that it appeared he had official protection. The 1991 murder of a Socialist Party boss remains unsolved, and many suspect the involvement of leading politicians and the Mafia. Willy Claes, former North Atlantic Treaty Organization (NATO) secretary general, was convicted in December 1998 of accepting a bribe from a defense contractor when he was economic affairs minister. Two other prominent Belgian politicians were also convicted.

Corruption was widespread in British politics during the eighteenth and early nineteenth centuries, but the Northcote-Trevelyan reforms of 1854 established the modern British civil service, which is world renowned for its honesty. Indeed, many former British colonies, such as Singapore, have benefited from their reputation for sound administration. The conservative governments of Margaret Thatcher (1979–1990) and John Major (1990–1997) saw their share of scandals, notably some involving large donations to the Conservative Party by questionable characters. (The United Kingdom does not limit donations to parties, nor does it require them to be disclosed.) Years of one-party rule led to a certain degree of arrogance and self-satisfaction, but little real corruption. While many observers attributed Major’s defeat in 1997 to “sleaze,” it appears that is just another way of saying, “The Tories had been in power for too long and it was time to give someone else a chance.”

Germany is also known as a relatively corruption-free country, although, contrary to popular belief, the Nazi regime was filled with petty graft and party patronage. In the 1980s, it was revealed that the Flick industrial group had made illegal donations to the major political parties, while this embarrassed the government of Chancellor Helmut Kohl (1982–1998), it never threatened its stability. More recently, there have been scandals about bribery in the construction industry and about the involvement of big-city police with the drug trade and prostitution. Probably the most serious threat is the large network of ex-communists in the former East Germany, which has become known for favoritism and intimidation.

The most conspicuous European scandal of the late 1990s has concerned the European Commission, the governing body of the European Union. An internal audit, leaked to members of the European Parliament, revealed waste, corruption, and cronyism in many Euro-programs. It revealed that Edith Cresson, commissioner for research policy and a former French prime minister, had appointed her dentist to coordinate acquired immune deficiency syndrome (AIDS) research and had presided over a fraud-ridden job training program. In order to stem a parliamentary drive to censure or remove the commissioners, commission president Jacques Santer appointed the Committee of Independent Experts to investigate the audit’s charges. The panel produced a report that essentially found the allegations to be true. Faced with a parliamentary revolt, and with Cresson unwilling to quit alone, all twenty European Commissioners resigned in March 1999.

### **CORRUPTION IN EAST ASIA**

During the last quarter of the twentieth century, East Asia has had the fastest-growing economy in the world, but the economic crisis of 1997–1998 exposed the vast corruption of some countries. Indonesia was the most conspicuous case. During his 1966–1998 reign, Suharto amassed a fortune that may reach into the tens of billions of dollars. His children enriched themselves by acting as middlemen between state-owned firms and contractors; serving as the Indonesian partners of many foreign investors; and acquiring interests in aviation, broadcasting, automobiles and many other industries. Businessmen close to Suharto, such

as industrialist Liem Sioe Liong and timber magnate Mohammed "Bob" Hasan, came to dominate the economy. As long as Indonesia kept growing, most citizens and businessmen accepted Suharto's banditry. Indonesia's economic and political collapse in 1998 led to demonstrations against his rule and to bloody riots in Jakarta, the capital. Suharto quit in May 1998; B. J. Habibie, his successor, found his associates implicated in a multi-million-dollar banking scandal. Abdurrahman Wahid, elected president in October 1999, announced that he would pardon Suharto should he ever be convicted of corruption.

The Philippines has a long history of corruption. A small group of families has long dominated Philippine politics, buying off their supporters through graft and patronage. Filipino culture emphasizes personal relationships, rather than formal institutions. As such, written law has little connection with the workings of the Philippine government. The Philippine civil service, police, and congress have particularly corrupt reputations. After independence in 1946, several Philippine presidents were known to be corrupt, especially Elpidio Quirino (1948–1953) and Carlos García (1957–1961). Ferdinand Marcos (1965–1986) put them all to shame, especially after he declared martial law in 1972. Philippine Airlines became a private commuter line for the Marcos family. Marcos bullied wealthy families into giving up their holdings to his family and friends. The Central Intelligence Agency (CIA) estimated that Marcos himself was worth \$10 billion, but after Marcos fell from power, investigators were able to track down only a fraction of that sum. Despite her promises of a new era, the administration of Corazon Aquino (1986–1992) was also wracked with corruption. Fidel Ramos (1992–1996), however, built a reputation for honesty and presided over a period of prosperity.

Thailand is world renowned for its tolerance for drug trafficking and prostitution, and it is widely alleged that these businesses have allies in the government. The country also has a political culture dominated by patronage, bribery, vote buying, and crony capitalism. Prime Minister Banharn Silpa-Archa (1995–1996) was particularly identified with these practices, but the crisis of 1997–1998 disrupted many of Thailand's arrangements. The country has now adopted a more

democratic constitution and has banned vote buying and insider deals.

South Korea has a political system that concentrates power in the presidency. It is not surprising that corruption has been concentrated there, too. Park Chung-hee seized power in 1961, vowing to end the sleaze of South Korea's early years. Park amassed enormous power over South Korea's economy and government. While most observers agree that he mostly used that power to his country's betterment, his dominance convinced many businessmen that they needed to donate to Park's party if they wanted to win his favor. After Park was assassinated in 1979, Chun Doo Hwan assumed the presidency. There were rumors of corruption surrounding Chun, which were confirmed in 1996 when he and his successor, Roh Tae Woo, were convicted of corruption and sedition. Chun was sentenced to life in prison; Roh received seventeen years. Kim Young Sam, who succeeded Roh in 1993, had a clean image that was soiled when his son Kim Hyun Chul was convicted in 1997 of receiving bribes. The collapse of one of the nation's leading conglomerates sparked charges that it had corrupt ties with Kim's government.

The Liberal Democratic Party (LDP) dominated Japanese politics from 1955 to 1993; it still leads most governing coalitions. On the international stage, it projected an image of serene competence. Behind the scenes, Japanese back rooms saw wheeling and dealing that would put any Chicago pol to shame. Kakuei Tanaka, prime minister from 1972 to 1974, ran a pork-barrel-fueled machine that dominated the LDP. Tanaka forced the government to extend the bullet train line into his mountainous home base, nearly bankrupting the national railroad in the process. In 1983, Tanaka was finally convicted of accepting bribes from the American aircraft manufacturer Lockheed, which had spread around \$22 million among Japanese politicians in the 1970s. While the LDP's rule was broken in 1993, it regained control in 1994, and by 1996 an old Tanaka ally, Ryutaro Hashimoto, was serving as prime minister. In recent years, scandals have rocked the Ministry of Finance, which wields great power over the Japanese economy. Finance officials have been accused of covering up banks' financial troubles and of leaking sensitive information to bank executives.

### CORRUPTION IN SOUTH ASIA

The nations of South Asia (India, Pakistan, Bangladesh, and Sri Lanka) are all former British colonies that benefited from the United Kingdom's common law tradition and bureaucratic honesty; but all have sunk into a morass of corruption. Pakistan is now considered one of the most corrupt countries in the world. Tax evasion is ubiquitous: only about half of assessed taxes are actually collected. Businessmen believe that they must pay government officials if they seek a public contract. Civil servants are poorly paid, and believe they must take bribes in order to survive. The two most recent prime ministers, Benazir Bhutto (1988–1990, 1993–1996) and Nawaz Sharif (1990–1995, 1997–present), have presided over administrations rife with corruption. In April 1999, Bhutto and her husband were convicted of taking kickbacks on government contracts. They were each sentenced to five years in prison.

A recent book on India and Pakistan calls corruption in India “pervasive, systematic, structured, and graded . . . running from the bottom to the top of the political order.” The British gave India an honest civil service, but Indira Gandhi's state of emergency (1975–1977) destroyed its integrity and morale by forcing bureaucrats to break the law. Now India's estimated fifteen million civil servants shake down citizens for every service imaginable, from installing phone lines to permitting people to enter the country. Corrupt politicians hold power on a national scale. Jayalitha Jayaram, the notorious boss of Tamil Nadu state, brought down the government of Prime Minister Atal Bihari Vajpayee through her defection in April 1999. A 1996 corruption investigation discovered that Jayaram owned 10,000 saris and 350 pairs of shoes.

### CORRUPTION IN THE MIDDLE EAST

The Middle East is a nation bereft of functioning democracies, save Israel and Turkey. The predominance of authoritarian regimes has created a situation ripe for corruption. The Saudi royal family has a reputation for corruption on a massive scale. It treats the national income as royal revenue and distributes it at whim. There are 7,000 royal princes, whom the *Economist* described as “lawless and hedonistic.” Not content with their

generous allowances, they collect huge commissions on foreign contracts and muscle in on private business. Public anger at royal corruption, particularly that of King Saud, fueled an Islamic fundamentalist movement for a time. Crown Prince Abdullah assumed management of the government in January 1996, and his honesty and ruthlessness have defused the threat from the opposition.

Egypt is world renowned for its bureaucratic corruption. Its one million civil servants are notoriously indolent and graft hungry. Bribes are necessary to engage in any government transaction. The government is unable to provide most public services, opening the door for Islamist groups to fill the void. Morocco, Tunisia, and Algeria are also known for their corruption. In Algeria, years of corrupt rule by the military and the National Liberation Front fueled an Islamist revolt that has led to a savage civil war.

Turkey's fragile democracy has been besieged by corruption. Its parties are far more concerned with seeking patronage and state contracts than with promoting ideological agendas. Former Prime Minister Tansu Ciller (1993–1996) has become vastly wealthy while in politics, and allegations of graft and cronyism have swirled about her. A staunch secularist, she nevertheless formed a coalition government with Islamic leader Necmettin Erbakan in 1996, because he promised to prevent an investigation of Ciller's finances. She made a similar deal with Erbakan's successor, Mesut Yilmaz, who had his own skeletons to hide. Yilmaz's government finally fell in November 1998 because of continuing allegations about corruption. There have also been charges that the police are in league with gangsters, right-wing hit squads, and drug traffickers.

### CORRUPTION IN SUB-SAHARAN AFRICA

With the possible exception of the former Soviet Union, sub-Saharan Africa is the most corrupt region in the world. Colonial rule left weak governments and few educated administrators. Socialist-minded leaders created state-dominated economies that failed disastrously and offered numerous opportunities for graft. With arbitrary boundaries left over from colonial times, few states have developed a sense of “nationhood,” and political actors' first loyalty is usually to their ethnic group.

“Personal rulership” is perhaps the most common political system in Africa. The ruler has almost unlimited legal competence. He is not concerned with building a following among the general public; instead he attends to the needs of his more powerful supporters, usually in the military or in state-owned firms. But state power is too limited to implement any policy except pillage. Personal rulership tends to be corrupt and arbitrary, but also weak and unstable.

The most extreme form of personal rulership is what Max Weber called a “sultanistic” regime. Such a state has no claim to legitimacy, whether based on tradition, ideology, charisma, or social order. The ruler runs the state to benefit him and his allies. He treats the public treasury as an extension of his personal fortune. He sells import and export licenses to the highest bidders, and appoints his cronies to run state-owned firms where they can exploit their nation’s resources. Because the ruler’s whims (and those of his supporters) supersede the written law, there is no predictability to the state’s actions. This discourages private enterprise, since businessmen cannot know whether their property will be confiscated. The only path to success is through winning the ruler’s fickle favors. While sultanistic states rarely function well enough to collect taxes, that did not bother many African leaders, who had their own streams of revenue: exporting natural resources through state-owned firms (especially during the high-commodity-price 1970s) and collecting foreign aid from the Cold War rivals. The French have remained active in their former colonies and have propped up many an unsavory dictator.

Mobutu Sese Seko, who ruled Zaire (now the Democratic Republic of Congo) from 1965 to 1997, was a perfect example of a sultanistic ruler who treated the state as his personal property. The *Economist* (1997, p.21) described his rule as the “systematic theft of the state from top to bottom.” Mobutu accumulated a fortune that amounted between \$4 and \$9 billion. Bribes, kickbacks, the public treasury, the returns from Zaire’s mineral wealth, French and American aid—all went into Mobutu’s Swiss bank accounts. By the end of his reign, he owned nine villas in Belgium; a Brazilian coffee plantation; a mansion in Spain; property in several African cities; and estates in Portugal, Switzerland, and the French Riviera. He also owned

tens of thousands of bottles of wine bottled in the year of his birth.

Mobutu found Zaire a poor country—and left it destitute. In the name of Zairean nationalism, he seized expatriate-owned businesses in 1973 and gave them to his cronies, who promptly ran them into the ground. (Uganda’s Idi Amin did much the same thing at the same time with Asian-owned businesses, with the same results.) Zaire is world-renowned for its mineral resources, including cobalt, diamonds and copper. Mobutu nationalized the mining industry—and wrecked it. Rather than paying them, Mobutu let his soldiers work for their money by robbing ordinary citizens. Pity the poor businessman trying to import goods into Zaire, for he had to pass a gauntlet of nine state agencies, all of them demanding bribes. The end of the Cold War meant that Mobutu’s Western allies would no longer subsidize his regime. By the mid-1990s, Mobutu’s government had stopped functioning, and a guerrilla uprising took root. With his army more interested in looting than in fighting, Mobutu fled the country in May 1997 and died soon after. By that point, Zaireans’ purchasing power was only one-fourth what it had been in the 1950s, when they still lived under Belgian rule. The Belgians left behind an excellent system of roads. By the 1990s, it had deteriorated to the point where it was no longer possible to drive across Zaire.

While Mobutu set an impressive standard for banditry, other African rulers did their best to keep up. Mali’s Moussa Traoré, Niger’s Gnassingbe Eyadema, Cameroon’s Paul Biya, and Gabon’s Omar Bongo are some of the more noteworthy kleptocrats who have misgoverned Africa over the past generation. Few have had the panache of Jean-Bedel Bokassa of the Central African Republic (1965–1979). Like Mobutu, Bokassa treated his land as his private domain. Unlike Mobutu, Bokassa went so far as to crown himself emperor in a ceremony modeled after Napoleon’s coronation. Through his control of the diamond trade, Bokassa accumulated a truly imperial fortune. His antics became an embarrassment to his French masters, who overthrew him after he personally led a massacre of children who were protesting the cost of school uniforms. The French spread a rumor that Bokassa ate his political opponents; distraught by this charge, Bokassa revealed that he had given

diamonds to leading French politicians, including then-President Valéry Giscard d'Estaing.

Of all African countries, Nigeria is probably the most notorious for corruption. Bribery is universal and scamming seems to be the national pastime. With its vast reserves of petroleum, Nigeria experienced an economic boom during the 1970s oil crisis. The returns from that wealth mostly went to the military and to corrupt politicians. The nation's schools, hospitals and roads are in dreadful condition. President Shehu Shagari's administration (1979–1983) was so corrupt that the army overthrew him so they could get in on the action. More recently, Sani Abacha stole billions of dollars from the national treasury during his 1993–1998 dictatorship. Abacha realized that distributing import licenses for fuel was a good way to reward his cronies. He allowed the nation's oil refineries to deteriorate, forcing this oil-producing nation to import gasoline, and thereby enriched his pals. Alas, poor Abacha did not get to enjoy his wealth for long: He died suddenly in June 1998, allegedly in the company of prostitutes. (Some have suggested that Abacha's rivals poisoned his Viagra.)

#### CORRUPTION IN LATIN AMERICA

While Latin America has not been pillaged as thoroughly as Africa, corruption is a serious problem in most countries in the region. Only Chile, Costa Rica, and Uruguay have reputations for being relatively clean. Latin America's corruption is rooted in its Spanish and Portuguese colonial past. Far from home, colonial bureaucrats were able to exploit their positions for profit, often treating their offices as their private possessions. Mercantilistic regulation meant that merchants and landowners often had to bribe officials in order to do business. Independence changed little. Hispanic traditions of paternalism meant extensive involvement by the state in the economy, which created great opportunities for bribery and smuggling. A foreign investment boom in the late nineteenth century strengthened local economies, but often primarily enriched a few leaders at the top. Nationalists attacked foreign investment, especially in the oil industry, and replaced it with state-owned firms, which frequently became deeply corrupt. The domestic private sector remained underdeveloped and dependent on the state.

Much of Latin America's corruption has swirled about the various strongmen who have held power over the years. Alfredo Stroessner, who dominated Paraguay from 1954 to 1989, turned his country into a gangster's paradise. Nazi war criminals and international drug traffickers flocked to live under Stroessner's protection. Stroessner treated Paraguay's government as his personal fiefdom, distributing patronage to his supporters and providing few services to ordinary citizens. Stroessner ensured his longevity by spreading the graft around. High tariffs in neighboring Argentina and Brazil spawned a thriving smuggling trade. Many military officers became millionaires through this trade, buying Rolls-Royces and building mansions. Eventually, the armed forces were wholly funded by the smuggling of cigarettes, whiskey, cocaine, and heroin. The costs of the vast Itaipú hydroelectric project mushroomed from a 1973 estimate of \$2 billion to a final figure of \$21 billion in 1991. Much of the overrun appears to have gone into officials' pockets through kickbacks and bribery.

Venezuela saw a series of *caudillos* who made huge fortunes off the nation's oil wealth. General Juan Vicente Gomez (1910–1935) accumulated \$200 million and became the nation's largest landowner. Major Marcos Pérez Jiménez did even better during his short reign (1951–1957), building a fortune of more than \$250 million. The Somoza clan ruled Nicaragua from 1936 to 1979, dominating the economy and stashing away tens of millions of dollars. Haiti has had scarcely one uncorrupt leader in its entire history, but the Duvaliers (François 1957–1971; Jean-Claude 1971–1986) stand out even in that wretched lot. When Jean-Claude fled Haiti in 1986, he was said to be worth \$800 million. Fulgencio Batista, president of Cuba from 1940 to 1944 and from 1952 to 1958, accumulated a fortune of between \$100 and \$300 million. Rafael Trujillo treated the Dominican Republic as his personal fiefdom during his 1930–1961 reign. At one point, Trujillo acquired the country's only shoe factory and then forbade Dominicans to go barefoot. Visitors to Santo Domingo noted that the country's poverty must be exaggerated: After all, everyone wore shoes! Trujillo's practice of requiring a 10 percent bribe on every governmental transaction helped him fill his bank accounts with up to \$500 million.

While presidential corruption is a long tradition in Latin America, a newer plague is infecting

states from the Rio Grande to the Andes: drug trafficking. *Norteamericanos*' insatiable appetite for cocaine, marijuana, and heroin (and their government's insistence on maintaining the ban on such substances) is undermining political stability in several Latin American countries. Cocaine cartels have been a fixture in Colombia since the 1970s. In the 1980s, the Medellín cartel threatened the stability of the Colombian government. More recently, it was revealed that the Cali cartel had partially funded the presidential campaign of Ernesto Samper (1994–1998). The Revolutionary Armed Forces of Colombia (FARC), guerrillas funded by the drug trade, now control much of the country. Many observers fear that Colombia will become the hemisphere's first "narco-state."

However, Mexico may get there first. Because of the extensive border it shares with the United States, Mexico has long been an ideal location for smugglers. When the Drug Enforcement Administration (DEA) increased surveillance in South Florida and the Caribbean during the late 1980s, the cocaine trade moved into Mexico. Since then the drug trade has boomed in Mexico, corrupting nearly every area of the Mexican government. Even General Jesús Gutiérrez Rebollo, who was leading Mexico's war on drugs, was arrested for accepting bribes from a major drug trafficker. The Mexican police are universally believed to be corrupt. State governors, army officers, party leaders—the drug traffickers have ties to them all.

Several leaders have made their countries havens for the drug trade. The most famous, of course, is Manuel Antonio Noriega, who dominated Panama from 1983 to 1989. An American invasion brought him to the United States, where he now serves a sentence for drug trafficking and racketeering. Vinicio Cerezo, president of Guatemala from 1985 to 1990, was widely suspected of being friendly to the drug trade. For a time during the 1980s, Bolivia was run by a group of military officers known as the "cocaine colonels."

Some of the larger nations in Latin America have developed large, clientelistic political party machines. Probably the most famous is Mexico's Institutional Revolutionary Party (PRI), which has (under different names) dominated Mexican politics since 1929. Since then, the PRI has won every presidential elections and most state ones, often

fraudulently. Since the 1980s, however, the National Action Party (PAN) and the Democratic Revolutionary Party (PRD) have become serious competitors for power. The PRI's influence permeates Mexican society through corporatist labor and business organizations. PRI patronage dominates many institutions. Mexico's state-owned oil company, Petroleos Mexicanos, is notoriously corrupt. Mexico has a long history of presidents who retired wealthy. José López Portillo, president from 1976 to 1982, parlayed the oil boom of those years into a fortune of at least \$5 billion. Carlos Salinas de Gortari (president, 1988–1994) moved to Ireland under widespread suspicion of massive corruption. His brother, Raul, was later convicted of arranging the murder of a prominent PRI politician. While Salinas privatized many of Mexico's state-owned firms, thereby opening up its economy, some of these companies went to PRI allies under suspicious circumstances.

From 1958 to 1998, two political parties dominated Venezuela in close cooperation. Vast oil wealth provided the necessary resources for patronage (a country of 22 million has 1.3 million government employees) and subsidies. The state nationalized many industries, providing more opportunities for patronage and graft, but also leaving the private sector weak and dependent. During the boom of the 1970s, politicians could indulge their penchant for cronyism and corruption. The decline in oil prices in the 1980s diminished the parties' ability to mollify dissent. Jaime Lusinchi, president from 1983 to 1988, had to remain in exile for several years to avoid prosecution for misuse of state property. Carlos Andrés Pérez, a fixture in Venezuelan politics since the 1950s and twice president (1974–1979, 1988–1993), was impeached in 1993 for abuse of public funds. Hugo Chávez, who had led an attempted coup d'état against Pérez, was elected president in 1998 without the support of the major parties.

Brazil lacks strong political parties, but it does not lack for patronage. Especially in rural areas, there is a tradition of machine politics dominated by local elites who use public resources to win mass support. President Fernando Collor de Mello (1990–1992) emerged from this background. While other Brazilian presidents had corrupt reputations, such as Juscelino Kubitschek de Oliveira (1956–1962) and José Sarney (1985–1990), Collor

became the first president to be impeached by Congress. He resigned before he could be removed from office. Corruption under the Collor regime included overpricing on government contracts, rigging public bidding, insider deals, and illegal fundraising. Collor's campaign treasurer, Paulo César Farias, collected \$2 billion in kickbacks from leading contractors. Collor himself spent \$2.5 million on a garden at his home. One corrupt Collor ally in the Brazilian Congress claimed his wealth came from winning the lottery 333 times; he used the lottery as a means of laundering his corrupt wealth. Some observers argued the Collor era's corruption stemmed from changes in Brazil's electoral law that made it more difficult for the president to form stable coalitions in Congress. Brazil's economic crisis reduced the public resources available for bargaining, forcing Collor to turn to private sources.

#### **CORRUPTION IN COMMUNIST AND POST-COMMUNIST STATES**

Nowhere in the world does corruption pervade society more thoroughly than in communist and post-communist states. The absolute command of resources that characterizes communist states, and many post-communist states, makes them perfect loci for massive corruption. Under Joseph Stalin and Mao Zedong, terror kept corruption from getting out of line. As communist states aged, corruption became more flagrant. In the USSR, the regime of Leonid Brezhnev (1964–1982) became a synonym for graft and inefficiency. Almost all Communist Party officials took bribes; most exploited their offices for all the privileges they were worth. Organized crime, which has a centuries-old history in Russia, established close ties with party leaders. Indeed, the black market often was all that kept the USSR from collapsing. Managers of state factories and farms had to be corrupt to survive. They lacked the material necessary to meet their goals, so they bribed officials to approve their production. Ordinary citizens had to bribe, too, in order to get consumer goods, medical care, and housing.

Other communist countries featured extensive corruption that came to light after the revolution of 1989–1991. Exposure of the lavish lifestyles

of the East German elite—who had always preached Marxist simplicity—led to massive demonstrations and helped trigger the end of the communist regime. The leaders of Bulgaria and Romania have also been revealed to be corrupt.

The communist era bequeathed a legacy that made massive corruption inevitable in most post-communist countries: a contempt for written law, a complete lack of protections for private property, a cynical public, an arrogant officialdom, a state-dominated economy, an absence of any institutions independent of the state. Some post-communist states made the transition relatively easily and with relatively little corruption: Poland, Hungary, the Czech Republic, Slovenia, and the Baltic states. But the states of the former USSR (with the exception of the Baltics) have traveled a rockier path that has mostly led in to a swamp of corruption.

When Russia privatized many of its state firms in 1992, the bulk of them went to former communist apparatchiks, ex-KGB agents, *mafia* leaders, and a few well-connected businessmen. These “oligarchs” generally paid bargain-basement prices. They have dominated Russian politics in the post-communist era and have supplied crucial support for President Boris Yeltsin. The crash of the ruble in the summer of 1998 damaged both their political credibility and their economic clout. Former Prime Minister Viktor Chernomyrdin is widely believed to have made a multi-billion-dollar fortune from the privatization of the state natural gas monopoly, Gazprom—one of the few Soviet-era companies that makes something that anybody wants. Allegations of massive corruption have swirled about former Finance Minister Anatoly Chubais, once the great hope of reform.

Paid poorly or not at all, military officers have turned to selling their own equipment. During the conflict in Chechnya, some Russian soldiers actually sold weapons to the Chechen separatists. Officers often would rather smuggle than perform their military duties. Some even sell nuclear material or chemical weaponry to rogue states. Customs officials are so corrupt and obstreperous that the U.S. government actually trains Russian businessmen in techniques for avoiding them.

Despite all the talk of Russian capitalism, the state still dominates the economy. Most Russians

still work for state firms. Businessmen must obtain licenses for most activities, which creates great opportunity for extortion. More important, Russia has not created the kind of legal system necessary for private enterprise to flourish. With few laws to govern the emerging capitalist economy, bureaucrats can take away a company's assets at a whim. With their property at the mercy of corrupt officials, even honest businessmen must bribe to stay afloat. There is little tradition of respecting private property in Russia. Even under the czars, all land belonged to the sovereign; even the wealthiest landowner was legally a tenant.

Taking advantage of this time of troubles, organized crime has taken over much of the Russian economy, with the help of their old communist and Komitet Gosudarstvennoy Bezopasnosti (Committee for State Security) (KGB) allies. The police and other law enforcement agencies are underpaid, undermanned, and demoralized. With domination of their home territory assured, Russian gangsters are now spreading their activities worldwide.

Other post-communist states have their own patterns of massive corruption. Foreign investors have shied away from Ukraine, despite its rich farmland and modern factories, because its all-encompassing graft and labyrinthine bureaucracy makes doing business almost impossible. As in Russia, former communist officials and *mafia* gangsters dominate the economy. Ex-communists made Bulgaria and Romania bywords for corruption until the voters booted them out.

China has moved toward a market economy while retaining the Communist Party's monopoly on political power. This has created many opportunities for corruption. Party leaders have plundered state firms, received state loans, and embezzled state funds. The People's Liberation Army has entered a variety of businesses, some legal, some not. The children of communist leaders—the so-called “princelings”—have used their positions to build fortunes. Some observers have argued that corruption has smoothed the transition to capitalism by circumventing bureaucratic rules. It is true that corruption has been most visible in the fastest-growing parts of China, especially in the southern province of Guangdong, where it may well have opened doors that strict legality would have kept shut. Corruption may now be harming commerce.

Like Russia, China lacks a system of property law, so businessmen remain at the mercy of officialdom and cannot settle their disputes in court. They must instead rely on bribery and *guangxi* (connections), which often translates into giving local bosses a cut of their profits. Foreign investors are losing patience. There are also signs that corruption is impairing governance. During the Yangtze River floods of the summer of 1998, water washed away many “bean-curd” dams and bridges. Officials had embezzled the money that would have strengthened these structures. Premier Zhu Rongji has used these disasters to justify his current anticorruption campaign.

North Korea has not moved away from communism one bit, so there is no private sector to shake down. Instead, the state itself has become a criminal enterprise. With its economy in ruins, North Korea has become a “gangster state” that sells narcotics and counterfeit money around the world. It can no longer pay its diplomats, so they support themselves by smuggling. North Korea's fake \$100 bills are so convincing that they led the United States to redesign its currency. To fuel its drug trade, the state runs its own methamphetamine factories and opium poppy farms.

### THE OUTLOOK FOR THE FUTURE

The global trend toward democracy and capitalism may decrease corruption in the long term, but in the short term it may increase the perception of corruption. The openness brought by democracy allows for greater disclosures of official malfeasance. Graft that would have remained secret under an authoritarian regime instead comes out, perhaps damaging the public's opinion of politicians. While capitalism reduces the state's role in the economy, thereby decreasing the opportunity for officials to enrich themselves, the privatization schemes necessary for the transition to free markets often become a bonanza for those in power.

There are several attempts to fight corruption on an international level. Transparency International, founded in 1993, organizes anticorruption efforts across the world. Its annual survey of perceptions of corruption has embarrassed many graft-ridden countries. Transparency International now has chapters in seventy countries. In 1997, twenty-nine members of the Organization for Economic Cooperation and Development (OECD) signed



the Convention on Combating Bribery of Foreign Public Officials. This accord, which went into effect in February 1999, forbids the bribery of foreign executive, legislative, and judicial officials. Countries that signed the accord had to agree to punish offenders severely, to help each other prosecute offenders, and to encourage the wider use of sound accounting and auditing practices. The International Monetary Fund (IMF) and the World Bank now often condition their loans on steps to reduce corruption and “crony capitalism.” In 1997, the World Bank established a comprehensive anticorruption program.

## REFERENCES

- Elazar, Daniel J. 1984 *American Federalism: A View from the Stars*, 3rd ed. New York: Harper and Row.
- Elliott, Kimberly Ann (ed.) 1997 *Corruption and the Global Economy*. Washington, D.C.: Institute for International Economics.
- Ford, Henry Jones 1904 “Municipal Corruption.” *Political Science Quarterly* 19:673–686.
- Heidenheimer, Arnold J., Michael Johnston, and Victor T. LeVine (eds.) 1989 *Political Corruption: A Handbook*. New Brunswick, N.J.: Transaction.
- Huntington, Samuel P. 1968 *Political Order in Changing Societies*. New Haven, Conn.: Yale University Press.
- Johnston, Michael 1982 *Political Corruption and Public Policy in America*. Monterey, Calif.: Brooks/Cole.
- Little, Walter, and Eduardo Posada-Carbó (eds.) 1996 *Political Corruption in Europe and Latin America*. New York: St. Martin's.
- Key, V. O., Jr. 1936 *The Techniques of Political Graft in the United States*. Chicago: University of Chicago Press.
- Mauro, Paolo 1997 *Why Worry about Corruption?* Washington, D.C.: International Monetary Fund.
- Noonan, John T., Jr. 1984 *Bribes*. New York: Macmillan.
- Nye, Joseph S. 1967 “Corruption and Political Development: A Cost-Benefit Analysis.” *American Political Science Review* 66 (June):417.
- Sabato, Larry J., and Glenn R. Simpson 1996 *Dirty Little Secrets: The Persistence of Corruption in American Politics*. New York: Times Books.
- Steffens, Lincoln. (1904) 1968 *The Shame of the Cities*. New York: Hill & Wany.
- “The Last Days of Mobutu.” 1997 *The Economist* March 22: 21.

RICHARD McGRATH SKINNER

## POLITICAL CORRECTNESS

## THE HISTORY OF POLITICAL CORRECTNESS

While the term *political correctness* (hereafter, PC) entered common parlance only in the mid-1990s and may soon lose its popularity, the phenomenon to which it refers, namely ideological conformity with the views of those in power or in fashion, is millenia old. To have PC, there must be ideology, and ideology developed with the formation of states. To the extent that all states use violence or the threat thereof to serve the interests of the few who control the state by extracting resources from the many who do not, states always face a problem of legitimacy. If they are to minimize their use of coercive violence to extract surplus production, states must try to convince their subjects that the state serves the interests not only of the ruling class but of the society as a whole. Such is the role of ideology, whether religious or secular.

Until the French Revolution, most states tried to legitimate themselves by identifying themselves with a religion, for instance, Confucianism in China, various branches of Christianity in Europe, and Islam from Morocco to Indonesia. In these “premodern” states, PC was virtually synonymous with religious orthodoxy, as defined by the political and religious elites in control. It was enforced by sanctions ranging from execution, torture, and expulsion to expropriation, ostracism, disfranchisement, discrimination, and segregation. In the late eighteenth century, the American and French revolutions abruptly shifted the onus of state legitimation from religious orthodoxy to a secular ideology of “liberty, equality, fraternity.” This ideology was, however, widely at variance with the structure of highly stratified societies (such as the United States, a slave society) and the practice of coercive states that were often democratic in name only. The bourgeois liberal ideology of the French Revolution gradually spread to much of the world in the nineteenth and twentieth centuries, briefly challenged by two large counter-movements, one reactionary (fascism), one radical (Marxism). Both challenges collapsed, fascism after a quarter-century, in mid-twentieth century, and Marxism in three-quarters of a century, some forty years later.

All these political movements thrived on the elaboration of ideology that served not only to legitimate state power, but to obscure the enormous discrepancy between the democratic, egalitarian ideals and the oppressive, exploitative reality. What the priesthood had been to religiously based PC, the intelligentsia and the bureaucracy became to secular PC. The actual term “political correctness” began to be used principally in the Marxist tradition, which claimed the ability to perform scientific analysis of social and political events, and thus allowed for the possibility of being correct or incorrect in one’s analysis. The “party line,” as defined by the ruling elite in communist regimes, invariably claimed correctness, of course, and invented elaborate *ex post facto* rationalizations for even the most radical policy changes (e.g., the 1939–1941 Soviet-Nazi alliance to gobble up Poland). Soon, however, the concept of PC was adopted by opponents of Marxism to ridicule the dogmatism and obvious opportunism of communist regimes, as, for instance, in George Orwell’s brilliant satire *Animal Farm*.

More recently, in the 1990s, the term PC was revived in the English-speaking world, and rapidly gained currency as a description of the self-imposed ideological conformity and censorship practiced by intellectual, business, and governmental elites in the United States, Canada, Britain, and elsewhere. While the term is principally a weapon used by conservatives to ridicule liberals, it is used across the ideological spectrum, sometimes in healthy, self-deprecatory criticism of fellow liberals or radicals.

### CONTEMPORARY POLITICAL CORRECTNESS

In the current context, PC has several interesting features, as follows:

1. While PC is often supported by state regulations, legislation, and a wide range of legal and extralegal measures (covering such things as “hate speech,” sexual harassment, homosexuality, and affirmative action), its principal focus of conformity enforcement is situated in the intellectual, bureaucratic, and business elites that “voluntarily” adhere to the PC

ideology, exercise extensive self-censorship, and practice more or less subtle forms of ostracism against mavericks.

2. PC ideology is widely at variance with what the majority of the population regards as commonsensical, and it conflicts with the interests and beliefs of many large groups (e.g. males, whites, religious fundamentalists, the poorly educated). It thus pits the liberal ruling elites against the populist masses. The colleges and universities, especially the elite institutions among them, are the most vocal and articulate proponents of PC, as are, of course, the graduates of such institutions when they accede to the command posts of government and business.
3. PC is not a coherent political movement aimed at changing the fundamental structure of society, as were Marxism and fascism, for example. Rather, it is an inchoate cultural movement aimed at imposing on a supposedly benighted populous the values, lifestyles, and speech patterns of the elites. It attacks almost exclusively the symbolic aspects of popular culture, rather than the fundamental features of social inequality and injustice. One of its principal cultural products is a rich crop of rapidly changing euphemistic neologisms, the use or nonuse of which serves as evidence of whether one is PC or not.

The central tenets of this elite PC ideology in the United States include the elements discussed below.

1. *PC is a theoretical celebration of “diversity” on many fronts: religious, ethnic, racial, linguistic, and sexual.* The corollary of this celebration of diversity is the rejection of the older ideology of the melting pot, of cultural assimilation, of English-language dominance, and an alternative vision of American society as a mosaic of racial, ethnic, linguistic, religious, and “lifestyle” groups clamoring for recognition of special status, and competing for scarce resources.

This diversity tenet has led in practice not so much to *tolerance* of “otherness,” a

widely accepted value, but to institutionalized, officialized recognition of group affiliation, rights and quota in employment, education, contracts, and so on. These “affirmative action” programs have often been “voluntary” and flexible, but they have faced massive resentment because they violated many values held dear by masses of Americans. They flew in the face of the principle that *individuals* have legal rights, not groups, and that these rights are *equal*. They also squarely violated universalistic norms, such as that merit, competence, or seniority should govern allocation of resources, rather than ethnic or racial membership. A massive backlash against affirmative action, bilingual education, and other “special group rights” programs resulted, because PC is seen, not only as creating division and dissension in American society, but as imposing a double standard of tolerance. For instance, black students are allowed to autosegregate on campuses in a way that white students are not. Hate speech codes are also asymmetrically enforced for blacks and whites, men and women, heterosexuals and homosexuals. Indeed, the very same “hate” words, such as “queer” and “nigger,” are differently evaluated depending on who uses them!

2. *PC embraces a libertarian ideology holding that individual freedom is, in principle, only limited by respect for the rights of others.* Again, if PC libertarianism were evenhandedly applied, it would meet little opposition. This, however, is not the case. Take two examples: drug use and hate speech. The PC position on drug use is one of unequivocal condemnation of tobacco use (increasingly a lower-class habit), even those forms of it that do not pollute the air (such as chewing tobacco). Alcohol use, on the other hand, is widely tolerated, even though its lethality is second only to that of tobacco. As for the illegal drugs, there is no clear PC line, although many self-styled libertarians are amazingly tolerant of heavy state repression and suspension of civil rights involved in the “war against drugs.”

Attempts to control hate speech on campuses and elsewhere is another glaring example of PC double standards. First, many PC proponents have little trouble reconciling their self-styled libertarianism and defense of free speech with attempts to restrict the latter for those with whom they disagree. Second, there is blatant lack of evenhandedness in ostracizing certain kinds of hate speech but not others. Thus, college professors may with impunity refer to their opponents as “male chauvinist pigs,” “racist honkies,” or “fascist cops,” but not as “dumb broads,” “ignorant niggers,” or “flaming faggots.” “Afrocentrist” teachers are allowed to spread anti-Semitic venom by alleging that African slavery was dominated by Jews, but psychometricians and psychologists like Arthur Jensen, Philippe Rushton, and Richard Herrnstein are virulently attacked for suggesting that part of the persistent difference in performance of whites and blacks on intelligence quotient (IQ) and other standardized tests may be genetic in origin. Examples of this double standard of PC enforcement in academia are legion, as documented in Dinesh D’Souza’s *Illiberal Education*.

The offence, it seems, lies as much in the gender, skin pigmentation, or sexual preference of the offender as in the hate speech itself. Only certain combinations are taboo. Others are tolerated, or even found amusing. (Stand-up comedy is an excellent barometer of PC, for instance, and most stand-up comics are acutely attuned to PC. Women may, with impunity, insult men; gays may insult straights; and blacks may insult whites—but not vice versa.)

3. *PC embraces conservationist, environmentalist causes; these, again, would receive wider acceptance if PC were not so elitist, selective, and sentimentalist.* PC ecologism is usually cast as a morality play between good guys (Native Americans, Aborigines, Greenpeace, the Sierra Club) and bad guys (loggers, oil companies, maquiladoras). The reality, of course, is much more nuanced. Native peoples, for

example, began to show as great a proclivity to overfish, overhunt, overgraze, overlog, and otherwise degrade their environment as that of more “advanced” peoples, as soon as they gained access to the destructive technology (steel axes, nylon nets, firearms, chain saws) that make these activities possible. There is also now a powerful backlash against ecologism among Third World intellectuals who argue, rather persuasively, that the ecological movement in the West is highly elitist. The “Green” movement is spearheaded, they say, by leisured romantics, intellectuals, biologists, and other educated members of affluent societies who, after despoiling their own countries, want to convince the poor countries to stop development in order to preserve unspoiled playgrounds for the scientists and eco-tourists of the rich countries.

4. *PC shows a concern for equality, which, once more, would be widely shared in the general population, if it had not been extended by PC proponents in two radically new directions.* The first extension concerns the transformation of the notion of *individual* equality of *opportunity* into *group* equality of *results*. The latter is as controversial as the former is almost universally accepted. The PC model of the ideal society is no longer one in which social rewards are fairly distributed to individuals according to their abilities, efforts, and ethics, but one in which socially defined racial and ethnic groups achieve proportional representation in every sphere of activity. Such a model of a quota society is not only absurd and unrealizable; it is also a prescription for perpetual conflict. The definition of groups is arbitrary and manipulable for gain. Millions of people are of mixed descent, a disallowed category in contemporary America. The population base for establishing proportionality is elastic. (For instance, should blacks be 2 percent of the students at the University of Washington, their percentage in the population of the State of Washington, which it serves, or 12 percent, the percentage of blacks in the United States?) Why should some

groups (e.g., Hispanics) be represented, but not others (e.g., Arabs)? Why is an overrepresentation of whites among physics professors objectionable, but not an overrepresentation of blacks on basketball teams, or of Hassidic Jews in the diamond trade? Why do PC liberals, who proclaim that race does not matter, defend, in the same breath, race-based affirmative action? (Even a formerly sensible Nathan Glazer reversed himself recently on this score and now believes in the necessity for a proportional black presence at elite universities. If blacks, why not poor rural whites, surely another oppressed minority?) In short, the quota society is a bad idea that deserves a quick burial. South Africa just abolished racial apartheid. Why should the United States perpetuate it?

The second controversial extension of the concept of equality by the PC proponents is on the gender front. Few Americans would contest the ideas that women’s worth is equal to that of men, that women should have equal rights with men, that they should be able to compete with men on equal terms, and that they should receive equal pay and benefits for equal work. But the PC agenda goes well beyond that consensual definition of gender equality. PC feminists are generally quite ambivalent about accepting the obvious differences between men and women; if they do accept these differences, they generally ascribe them more to nurture than to nature, and they seek either to ignore or to minimize them; and, quite inconsistently, they want to preserve gender segregation where it favors women (e.g., in sports). PC feminists, in short, want a unisex society, except where it suits them. Finally, PC feminists want the right to be protected against “sexual harassment” whenever they *feel* they have been harassed. Many, in addition, would want to subject sexual relations to a code of conduct involving repeated, explicit verbal assent (unilaterally imposed on men) for every sexual act. Some colleges have gone as far as institutionalizing this nonsense.

Needless to say, the common sense of most Americans, women as well as men,

refuses to accept an ideology that flies in the face of experience, and there is a substantial backlash against radical feminism. Yet PC feminist dogma still rules supreme in many policy domains—for instance, in the military, which tries not only to achieve full gender integration but to rule sex out of existence in its ranks. Any officer bold enough to suggest that this is impossible immediately jeopardizes his or her career.

5. *PC takes a secular outlook that has its roots in eighteenth-century Enlightenment and has dominated Western intellectual life ever since, but that arouses deep antagonism in the half or more of the American population that considers itself religious, and, even more so, in the quarter or so who are fundamentalist Christians.* PC secularism is more than Jeffersonian separation of church and state in that it also frontally attacks the “religious right” on moral issues such as contraception, divorce, and marriage, as well as most controversially—abortion and homosexuality. PC secularism is also at variance with its classical Enlightenment expression in that it does not automatically align itself with the scientific mainstream. For example, many PC secularists reject the mounting evidence supporting the partially genetic underpinning of human behavior, of gender differences, and of basic abilities and character traits. While they may ridicule biblical creationism, they espouse an extreme “social constructionist” view of human behavior and human relations, which is, in fact, a form of secular creationism.

#### THE EFFECTS OF POLITICAL CORRECTNESS

Because the current form of American PC is such an elite phenomenon, its effect has been limited to the relatively small class of literate bureaucrats and academics who take it seriously. It did stifle intellectual discourse on American campuses and did promote the teaching of a good deal of nonsense masquerading as scholarship. But there is little evidence that it makes many converts, and considerable evidence that it brings out a conservative

backlash. Indeed, PC ideology already seems on the defensive, vulnerable as it is to ridicule. American cultural products always have some resonance in other countries, especially English-speaking and Western European ones: the United Kingdom, Canada, Australia, Scandinavia, the Netherlands, and Germany in particular. But in those countries as well, its influence has been limited to the intellectual elite. In much of the rest of the world, American-style PC has had even less resonance, either because the issues it addresses have little relevance there (e.g., race relations), or because other political and economic problems (such as human rights violations or poverty) give such issues as feminism or conservation much lower priority. Indeed, many PC tenets may even clash more openly with the values of non-Western cultures than they do with the common sense of Westerners.

The probability is, thus, high that the current wave of American-style PC has already crested and that it will be ephemeral. But then, a new brand will crop up, probably no more sane than the current one. Intellectuals, alas, often prove themselves to be much more the sycophants of power than the guardians of reason.

#### REFERENCES

- Devine, Philip E. 1996 *Human Diversity and the Culture Wars*. Westport, Conn.: Praeger.
- D’Souza, Dinesh 1992 *Illiberal Education*. New York: Vintage.
- Fish, Stanley Eugene 1995 *Professional Correctness*. New York: Claredon.
- Friedman, Marilyn, and Jan Narveson 1995 *Political Correctness: For and Against*. Lanham, Md.: Rowman and Littlefield.
- Glazer, Nathan 1975 *Affirmative Discrimination*. New York: Basic.
- Orwell, George 1946 *Animal Farm*. New York: Harcourt.
- Wilson, John K. 1995 *Myth of Political Correctness*. Durham, N.C.: Duke University Press.

PIERRE L. VAN DEN BERGHE

## POLITICAL CRIME

Political crime has been more often an object of partisan assertion than of independent research.

Passions are easily aroused, facts are difficult to establish. Nevertheless, a growing number of studies have contributed to (1) articulating the issues in defining political criminality, (2) describing instances and patterns of resistance to political-legal authority, (3) cataloging and analyzing governmental efforts to prevent and counter such challenges, and (4) proposing research agendas.

### THE PROBLEM OF DEFINITION

Political criminality may be narrowly or broadly defined, with greater or lesser regard for definitions offered by laws and interpretations by authorities. Moreover, the values and politics of observers frequently influence their conceptions of what and who is politically criminal. The resulting *mélange* of definitions has led Kittrie and Wedlock to conclude pessimistically, "It may be that an objective and neutral definition of political crime is impossible, because the term seems to involve relativistic relationships between the motives and acts of individuals and the perspectives of government toward their conduct and allegiances" (1998, p. xxxvi). However, an alternative view is implicit in their understanding of political criminality as perceptual and relational, defined in interaction between opposing parties.

Whose perceptions decide what is to be called political crime? Apart from partisan and subjective answers, the empirical reality of political criminality is that it is defined by those with enough power to impose their perceptions. Insofar as a political authority structure has been established, one may argue that the dominant parties within it by definition have the power to define criminality (Turk 1982a, pp. 11–68, 1984; pp. 119–121; for more restrictive legalistic definitions, see Ingraham 1979, pp. 13, 19; Frank Hagan 1997, p. 2). However, this view leaves unsettled the question of how authorities themselves may be defined as political criminals.

The most common resolution is to expand the definition to include anyone who commits extralegal acts defending or attacking an authority structure. For example, Roebuck and Weeber consider political crime to be any illegal or disapproved act committed by "government or capitalistic agents" or by "the people against the government" (1978, pp. 16–17). This very subjective definition (disapproved by whom?) leaves one unable to distinguish either between acts against and on behalf of

authority or between legal and nonlegal behavior. A more promising approach is to recognize that criminality may be defined at different levels of political organization—international as well as national.

International conventions, treaties, and judicial decisions have gone far toward formally defining war crimes, genocide, terrorism, violations of human rights, and environmental depredations. Such international standards may be invoked as criteria for labeling not only individuals but also agencies and regimes as political criminals guilty of "state crimes" (Barak 1991; Ross 1995). Even when formal legal criteria have not been articulated, evidence of harmful consequences of governmental policies and corporate practices may be used to define as political crimes "acts of commission and omission that result in grave social harm" (Tunnell 1993, p. xiv).

Formal pronouncements and unofficial assertions, however, may not have much significance. Accusations against governmental and corporate regimes are unlikely to impact on them unless there is a significant likelihood that they will be penalized for their actions. Within nations, acts by subnational authorities have sometimes been effectively treated as political crimes by national authorities who saw them as endangering societal or elite interests. An example is the American federal government's historic crackdown on state and local violations of constitutionally guaranteed civil rights in the South.

At the international level, it is still rare for even gross violations of human rights to result in sanctioning. Though often accompanied by rhetorical accusations of criminality, expeditions such as the Kosovo intervention by the North Atlantic Treaty Organization (NATO) have typically been launched more out of geopolitical concerns than legal or moral ones, and tend to end in new political and economic accommodations rather than criminal trials. For the present, the lack or weakness of international policing and judicial institutions leaves powerful nations such as the United States free to reject or ignore such charges as practicing genocidal policies toward Native Americans, holding political prisoners, torturing detainees and convicts, and illegally authorizing operations in other countries by the Federal Bureau of Investigation (FBI), the Drug Enforcement

Agency (DEA), and the Central Intelligence Agency (CIA). With little or no chance of their enforcement, invoking international or national laws to define political crimes will continue to amount to little more than largely subjective and partisan ideological tactics.

Another question is whether political criminality is to be defined only as behavior. While one may agree that only specified harmful acts *should* be punishable, the historical fact is that, in addition to offending behavior, nonbehavioral attributes such as ethnicity and class background have frequently been used by antagonists as criteria of intolerable political deviance. Moreover, imputed as well as observed deviations or threats have been used. Anticipation as well as reaction are involved in the identification of political criminality. At the beginning of World War II, Japanese immigrants and their descendants in the United States were officially defined as security risks and forcibly relocated to prison camps, and during both world wars many German nationals and immigrants were subjected to surveillance and orders restricting their freedom to travel outside the areas where they lived.

In sum, political criminality is most realistically defined as whatever is treated as such by specified actors (usually governments, dominant groups, or their agents) who have the power to impose their conceptions in particular historical situations.

### RESISTING AUTHORITY

Resistance to political authority may be more or less deliberate, organized, or planned. As noted above, even the appearance or potential of resistance may be sufficient for authorities to act against perceived challengers. Actual resisters may be engaged in activities ranging from merely disrespectful comments to the most violent assaults, from spontaneous eruptions to carefully orchestrated attacks, from individually motivated acts to organized strategies of rebellion. Acts of resistance may be categorized as evasion, dissent, disobedience, or violence.

Resisters may simply evade the orders and demands of the powerful. Avoiding masters, bosses, tax collectors, and military conscription has generally been safer than explicit defiance—because open defiance tends to force authorities to

respond, while tacit evasion is more easily ignored or minimized. Russian serfs learned in such instances as the St. Petersburg massacre how ruthless the Cossacks could be in suppressing dissenters and petitioners—the kind of lesson that has been taught to peasants and workers over the centuries in many lands. As Karl Marx complained, peasants and industrial workers tend to settle for mere survival and small gains instead of becoming revolutionaries. Fear of brutal repression is not groundless, and those with little social power are realistically skeptical even in democratic societies (as regularly shown in legal and political attitude polls) of notions that they can expect protection because they have legal rights.

The right to speak out against authority is enshrined in many legal traditions but in practice is limited by the varying tolerance of authorities and people—decidedly lower in wartime and economic hard times. After centuries of attempting to deter such offenses as seditious libel and treasonous utterances, more democratic governments have de facto concluded that the effort is incompatible with the concept of free speech (Law Commission 1977; Stone 1983; Hurst 1983; cf. Franks 1989). However, other labels may be invoked to authorize punishing those whose dissent is especially galling, especially when authorities feel particularly threatened. U.S. Senator Joseph McCarthy, for instance, achieved the lasting notoriety of “McCarthyism” by making freewheeling accusations of “subversion” against a wide spectrum of targets—from known communists to President Dwight Eisenhower. Though McCarthy’s failure to refrain from smearing the president and the army, both icons in 1950s America, resulted in his downfall, the political climate remained one in which public objections to capitalism, American foreign policies, and even racial segregation carried the risk of being labeled a “communist” or at least a “fellow traveler.”

Dissenting is one thing, but actually disobeying rules and commands is another. Lower-class people have historically suffered many demonstrations of their vulnerability, and so have characteristically been less likely than higher-class people either to dissent or to disobey overtly (Turk 1982a, pp. 69–114). On the other hand, dissenting or disobedient higher-class, especially youthful, resisters have typically been subjected to less punitive treatment. For instance, sentences of Vietnam

draft resisters decreased in response to the political repercussions of imprisoning growing numbers of higher-class young men (Hagan and Bernstein 1979). Not surprisingly, civil disobedience has been more likely to be a higher-class than a lower-class mode of resistance—that is, a tactic of those whose backgrounds encourage them to believe, perhaps erroneously, in their own significance and efficacy and in legal rights and protections. Certainly in the 1960s, in the early stages of the civil rights movement, marchers against southern segregation were far more likely to be students, intellectuals, and “bourgeois” than laborers, farmers, and “proletarians.”

Violent resistance by individuals and small organizations seldom threatens authority structures directly, but does indirectly weaken them in that authorities facing or fearing violence typically adopt extralegal and repressive control measures—which contradict beliefs in legal restraints on governmental power. The more actually and ideologically democratic the political order, the greater the contradiction—which is associated with the greater vulnerability of democratic than despotic regimes to terrorism (Turk 1982b, p. 127).

Political violence tends to escalate from coercive to injurious to destructive forms, in an interaction spiral generated by decreasing hope on the side of one or both parties of an acceptable accommodation and/or by perceived advantages in escalation (Turk 1996). Coercive violence, which may complement or succeed verbal or legal appeals and arguments, is intended to pressure opponents to change their policies or practices—typical acts being telephoned threats, disruption of communications, and rioting. Injurious violence is aimed at causing limited damage to property or persons, with little or no risk of lethal damage, when opponents do not respond satisfactorily to coercion—as when Quebec separatists blew up mailboxes and American student radicals firebombed Reserve Officers Training Corps (ROTC) buildings on university campuses. Torture and execution mark the transition from injurious to destructive violence. Destructive violence seeks to terrify active and potential opponents and to eliminate the enemy’s resources of personnel, supplies, and facilities; examples include the assassinations of governmental (or revolutionary) leaders, as well as acts of terrorism (e.g., the Oklahoma City bombing).

Political offenders have commonly been viewed as morally or mentally defective. However, even the most violent assassins and terrorists appear unlikely to exhibit psychopathology (Turk 1983, 1984, p. 123; cf. Schafer 1974; also Robins and Post 1997). The effort to understand political resisters begins in recognizing that they may vary enormously in political consciousness and motivation, organizational involvement, and readiness to commit violent acts, in addition to other characteristics, such as class origins. Distinctions must be made among deliberate political actors, emotional reactors to climates of political instability and violence, and apolitical opportunists such as ordinary criminals who merely seek to profit from their contacts with resistance figures and movements.

### ASSERTING AUTHORITY

Political dominance is defended legally and often extralegally at all institutional levels. The highest and broadest level is typified by national (and increasingly international) policies of insulation (regulating access and mobility, e.g., through immigration and licensing laws), sanctioning (defining and enforcing behavioral rules, e.g., through criminal laws), and persuasion (disseminating favored communications, e.g., through education and censorship)—that is, through *statecraft*, governing strategies designed to ensure that potential resisters lack the opportunities, resources, and will to challenge authority (Gamson 1968). More specifically focused on controlling resistance is political policing—the organized effort to gather relevant intelligence, manipulate channels of communication, neutralize opposition, and deter challenges.

Because the value of information can never be fully anticipated, intelligence gathering is inherently limitless. Advancing technologies of surveillance and analysis enable ever more extensive and intensive monitoring of human behavior and relationships. The distinction between public and private is increasingly dubious in both law and practice. Recurring legislative and judicial efforts to impose legal restraints have occasionally slowed but never stopped the trend (Marx 1988). One of the most celebrated efforts—to end surveillance of civilians by U.S. Army intelligence officers—failed when the highest court ruled in 1972 “that it was not the business of the Supreme Court but of



Congress to monitor executive practices like surveillance” (Jensen 1991, p. 255).

Authorities have never been entirely comfortable with the ideal of free and open communications. Even where that ideal has been most firmly asserted in law, in practice the right to disseminate critical information and ideas has been limited (Kittrie and Wedlock 1998, *passim*). Openly or subtly, communications favoring the status quo have been facilitated, while dissent has been inhibited to a greater or lesser degree. For example, Eugene Debs (leader of the Socialist Party) was convicted and imprisoned not explicitly for his outspoken opposition to American involvement in World War I, but instead for “conspiracy” to obstruct military conscription and war production—a charge under the Espionage Act of 1917, which required no direct evidence of either conspiracy or obstructive actions.

When resistance is encountered, some blend of enclosure and terror tactics is used—that is, a combination of methods designed not only to suppress resistance but also to convince offenders that apprehension is inevitable and punishment unbearably severe. Psychological as well as physical coercion is accomplished by subjecting targets to sanctions varying from character assassination to torture and extermination. An example of the first is the FBI’s effort to discredit Martin Luther King, Jr., as a subversive agitator, a communist sympathizer or dupe, and an immoral womanizer. As for the second, South African apartheid forces provided many instances, including the beating death under interrogation of Steve Biko, leading voice of the “black consciousness” movement.

International covenants notwithstanding, torture and other violations of human rights continue throughout the world (regularly documented by Amnesty International and other organizations). Extreme human rights abuses are reported not only in dictatorships such as Iraq and authoritarian regimes such as Sri Lanka but also in democracies such as Britain, India, Israel, and the United States. Clearly, authorities facing serious challenges are unlikely to be restrained by legal or other norms.

The ultimate goal of political policing is general deterrence. Insofar as the subject population does not knowingly and willingly accept the political order, fear and ignorance may still ensure

acquiescence. Surveillance, censorship, and neutralization are designed not only to repress political deviance but also to discourage any inclination to question the social order. But intimidation must be supplemented by persuasion if superior power is to be legitimated—transformed into authority. A classic technique is the political trial, in which legal formalities are used to portray the accused as a threat to society, to convey the impression that political policing is legally restrained, and to reinforce the sense that political dominance is both right and irresistible. Such trials have not been limited to totalitarian states such as Nazi Germany and the Soviet Union (Kirchheimer 1961), but have also occurred with some regularity in democratic nations such as the United States. Historically, examples have occurred from the Puritans’ legalistic suppression of Anne Hutchinson to the Chicago trial for “conspiracy” of eight assorted radicals with varied causes, whose only link was that they “represented the spectrum of leftist dissidents” (Belknap 1994, p. 240; see also Christenson 1986).

#### RESEARCH AGENDAS

Studies of political criminality have raised far more questions than they have answered. In the future, it will be not isolated, small-scale investigations but ongoing research programs, that will be essential if the quest for systematic explanations of political criminality is to be successful. Such programs will necessarily be multilevel, integrating research on the political socialization of individuals, the radicalization of defenders as well as challengers of authority, the interaction of resistance and policing strategies, and the conditions under which the political organization of social life is relatively visible (low rates of both resistance and repression).

The viability of a political authority structure can in principle be objectively defined: The probability of its survival increases or decreases. Accordingly, research can identify progressive actions (which increase viability) and destructive actions (which decrease viability) by anyone involved in political conflict. Turk (1982a, pp. 181–191) hypothesizes that random violence, economic exploitation, and weakening social bonds are destructive, while nonviolent actions increasing the life chances of everyone instead of only some people are likely to be progressive.

## REFERENCES

- Barak, Gregg 1991 *Crimes by the Capitalist State*. Albany: State University of New York Press.
- Belknap, Michal R. (ed.) 1994 *American Political Trials*. Westport, Conn.: Praeger.
- Christenson, Ron 1986 *Political Trials: Gordian Knots in the Law*. New Brunswick, N.J.: Transaction.
- Franks, C. E. S. (ed.) 1989 *Dissent and the State*. Toronto: Oxford University Press.
- Gamson, William A. 1968 *Power and Discontent*. Homewood, Ill.: Dorsey.
- Hagan, Frank E. 1997 *Political Crime: Ideology and Criminality*. Needham Heights, Mass.: Allyn and Bacon.
- Hagan, John, and Ilene Bernstein 1979 "Conflict in Context: The Sanctioning of Draft Resisters, 1963-76." *Social Problems* 27:109-122.
- Hurst, James Willard 1983 "Treason" In Sanford H. Kadish, ed., *Encyclopedia of Crime and Justice*, vol. 4. New York: Free Press.
- Ingraham, Barton L. 1979 *Political Crime in Europe: A Comparative Study of France, Germany and England*. Berkeley: University of California Press.
- Jensen, Joan M. 1991 *Army Surveillance in America, 1775-1980*. New Haven, Conn.: Yale University Press.
- Kirchheimer, Otto 1961 *Political Justice: The Use of Legal Procedure for Political Ends*. Princeton, N.J.: Princeton University Press.
- Kittrie, Nicholas N., and Eldon D. Wedlock, Jr. (eds.) 1998 *The Tree of Liberty: A Documentary History of Rebellion and Political Crime in America*. Baltimore, Md.: Johns Hopkins University Press.
- Law Commission 1977 *Codification of the Criminal Law: Treason, Sedition, and Allied Offences*. Working Paper 72. London: Her Majesty's Stationery Office.
- Marx, Gary T. 1988 *Undercover: Police Surveillance in America*. Berkeley: University of California Press.
- Robins, Robert S., and Jerrold M. Post 1997 *Political Paranoia: The Psychopolitics of Hatred*. New Haven, Conn.: Yale University Press.
- Roebuck, Julian, and Stanley C. Weeber 1978 *Political Crime in the United States: Analyzing Crimes by and against Government*. New York: Praeger.
- Ross, Jeffery Ian (ed.) 1995 *Controlling State Crime: An Introduction*. New York: Garland.
- Schafer, Stephen 1974 *The Political Criminal: The Problem of Morality and Crime*. New York: Macmillan.
- Stone, Geoffrey R. 1983 "Sedition" In Sanford H. Kadish, ed., *Encyclopedia of Crime and Justice*, vol. 4. New York: Free Press.
- Tunnell, Kenneth D. (ed.) 1993 *Political Crime in Contemporary America: A Critical Approach*. New York: Garland.
- Turk, Austin T. 1982a *Political Criminality: The Defiance and Defense of Authority*. Newbury Park, Calif.: Sage.
- 1982b "Social Dynamics of Terrorism." *Annals of the American Academy of Political and Social Science* 463:119-128.
- 1983 "Assassination" In Sanford H. Kadish, ed., *Encyclopedia of Crime and Justice*, vol. 1. New York: Free Press.
- 1984 "Political Crime." In Robert F. Meier, ed., *Major Forms of Crime*. Newbury Park, Calif.: Sage.
- 1996 "La violencia politica desde una perspectiva criminologica." (Political Violence in Criminological Perspective) *Sistema* 132-133 (Junio):41-55.

AUSTIN T. TURK

**POLITICAL ORGANIZATIONS**

Investigations of the behaviors of political organizations occur at an intersection of sociology, political science, and organizations studies. This interdisciplinary perspective offers great potential for richly informed understanding and comprehensive theoretical explanation of numerous facets of these crucial social actors and their relationships with the larger society and polity in which they are embedded. Four fundamental questions have dominated research and theorizing in this field over the past several decades: (1) What sociopolitical conditions encourage the creation of political organizations? (2) How are participants recruited and induced to provide crucial resources for political action? (3) What mobilization processes enable political organizations to work together most effectively toward collective ends? (4) What strategies and tactics exert the greatest impact on public policy makers' decisions? This article attempts to provide brief answers to these questions. Although most of the empirical research on political organizations concentrates on the United States, some recent evidence from European nations is examined.

The least restrictive definition of a political organization is any formally organized, named group that tries to influence the policy decisions of public officials. Most political organizations take the form of a voluntary association of persons or

organizations that pools its members' and constituents' financial and other resources, and engages in conventional political actions to affect policy-making outcomes. Common synonyms for this type of organization are "interest groups," "pressure groups," and "collective action organizations." Ironically, political parties are not political organizations, because their primary purpose is to elect candidates to public office and only incidentally to press for specific policy agendas. Most public bureaucracies should be excluded, unless they act regularly to promote their own policy preferences within a government. Another questionable type is the social movement organization whose primary political tactics involve rallies, demonstrations, and violent forms of protest (including revolutionary actions intended to overthrow the government) rather than working within routine channels of the political system. Some social movement organizations eventually transform into conventional political interest groups, if they survive their turbulent youths as outside challengers. However, certain profit-making corporations might be considered quasi-political organizations, in instances where their government affairs officers lobby for preferential treatment from legislators and regulators (Salisbury 1994).

Political organizations—including such types as labor unions, professional societies, business and trade associations, churches, neighborhood and community organizations, fraternities and sororities, nationality and racial-ethnic federations, civic service, philanthropic, and cooperative groups, medical and legal societies, conservation leagues, and even recreational and hobby clubs—encompass a broad range of formal goals. Political purposes need not be their primary goal nor compose the majority of their activities, but the critical requirement is that they go beyond merely providing direct services to their members by seeking to change or preserve the social, economic, cultural, or legal conditions faced by their members or those on whose behalf they operate. One interesting type is the so-called citizens' group or public interest group (PIG), which purports not to benefit narrow sectarian or economic self-interests but to promote the broader collective values of the society (Berry 1977). For example, civil rights, civil liberties, environmental protection, feminist, and consumer advocacy associations frequently proclaim a disinterested agenda. A close examination

of their supporters and activities suggests that they do not differ fundamentally from other political organizations in methods of operation (Schlozman and Tierney 1986, pp. 30–35). Based on listings compiled by various American directories, perhaps as many as 23,000 voluntary associations operate at the U.S. national level (many with dozens or hundreds of chapters and branches in state and local communities). Of these, perhaps half qualify as political organizations based on their efforts to communicate their positions on national policy issues to the federal government (Knoke 1990, p. 208). They range in size from the American Association of Retired Persons (AARP), with more than twenty-five million members, to small staff organizations with fewer than a dozen operatives bankrolled by foundations or public donations.

The creation, growth, and expansion of U.S. political organizations seem to occur in cycles corresponding to national political and economic events, including shifts in the legislative, regulatory, and judicial climate (Berry 1977, p. 13; Schlozman and Tierney 1986, pp. 74–82; Gray and Lowery 1996). American labor unions established a national policy presence during the New Deal, and public interest groups blossomed during the civil rights, antiwar, and feminist social movements of the 1960s. Business advocacy associations flocked to Washington in the 1970s and 1980s in reaction to restrictions imposed by newly established federal regulatory agencies for environment, occupational safety and health, consumer protection, and equal employment opportunity (Vogel 1996). Increasingly, mass membership associations have yielded ground to institutionally based organizations, including corporations; universities; foreign firms and governments; and confederations of U.S. state and local governments, such as the National League of Cities.

Interest groups rarely form spontaneously but require leadership and resources. Interest group foundings and expansions may be best understood as involving exchanges between entrepreneurial organizers, who invest capital in a set of benefits offered to potential members as the price of membership, and members who pay dues in order to receive these benefits. Intergroup subsidies may occur; for example, the American Federation of Labor–Congress of Industrial Organization (AFL-CIO) founded the National Council of

Senior Citizens, helped to recruit its early members from unions, and continued to underwrite its activities. Although committed citizen activists, such as Ralph Nader or John Gardner, occasionally provide an energizing impetus for launching new organizations, many public interest groups rely on patronage from wealthy individuals or foundation sponsors to get launched (Walker 1983), as well as on favorable mass media treatment to bolster their legitimacy. Once an organization is formed, its survival, growth, and effectiveness depend on its ability to attract and hold new members and other organizational sponsors. Labor unions and business associations can acquire substantial war chests through dues and assessments on their members, but PIGs have more limited capacities to tap potential diffuse constituents' money. The virtual collapse of Greenpeace in the 1990s underscores the vulnerability of many activist organizations to quickly dwindling support.

For more than three decades, a central paradigm to explain member contributions has been the economic or rational choice model developed in Mancur Olson Jr.'s *Logic of Collective Action* (1965). Olson considered the conditions under which people would voluntarily contribute their resources to an organized group seeking a public good (such as a governmental farm crop subsidy) from which no eligible recipients could be excluded. Olson argued that utility-maximizing actors would refuse to pay for public goods that will be produced regardless of their contributions, and thus would take a "free ride" on the efforts of other members. As a result, the model predicts that most political organizations should fail to mobilize their potential supporters if they were to rely solely on public goods to obtain sustenance. Olson concluded that such entities are viable only if they offer "selective incentives" to prospective members in exchange for their contributions toward the organization's public-good lobbying efforts. These inducements might include magazine subscriptions, group insurance, social gatherings, certification and training programs, and similar benefits from which the organization could effectively exclude noncontributors unless they pay dues and assessments. In Olson's formulation, a political organization's policy objectives are reduced to a secondary "by-product" of its members' and supporters' interests in obtaining personal material

benefits. Despite his appealing analytical arguments, Olson's propositions were repeatedly challenged by empirical investigations of the incentive systems of real voluntary organizations. Members often respond to diverse inducements apart from personal material gains, including normative and purposive appeals and organizational lobbying for public goods (Moe 1980, pp. 201-231; Knoke 1988, 1990, pp. 123-140). For example, right-to-life organizations appeal to their supporters' religious, ideological, and emotional convictions about the illegitimacy of abortion and the necessity to take direct action to shut down clinics as well as to campaign on behalf of prolife politicians. The availability of picnics or T-shirts could hardly provide a compelling motivation for most participants in these groups. The internal economies of political organizations turn out to be more complex than originally believed. Organizational leaders have an important role in defining the conditions and prospects for their members and in persuading them to contribute to collective efforts that may run counter to the members' short-term personal interests.

The governance of political organizations is often posed as a choice between oligarchic or democratic alternatives. Persistent leadership and staff cliques in labor unions, trade associations, fraternal organizations, professional societies, and other types of associations are frequently interpreted as evidence of an inevitable "iron law of oligarchy." However, apart from labor unions (with their legal monopolies on occupational representation within certain industries), most voluntary groups are too dependent on their members for critical resources to enable officials to flout the memberships' interests in the long run. Consequently, most political organizations' constitutions provide for an array of democratic institutions, including competitive elections, membership meetings, referenda, and committee systems (Berry 1984, pp. 92-113; Knoke 1990, pp. 143-161). But actual practices of consulting members to formulate collective actions vary widely, and researchers have only begun to examine how the democratic control of political organizations shapes their capacities to mobilize their members for collective actions. The analytic task is further complicated by the complex interactions of formal governance processes with executive and leadership actions,

bureaucratic administration, environmental conditions, and the internal economy of member incentives.

Political organizations serve a dual function for a political system. First, they aggregate the interests of citizens holding similar preferences, enabling them to press their demands on government officials more effectively. By articulating member demands and pooling the scarce resources of weak individuals, interest groups fashion a louder voice that is not readily dismissed by those in positions charged with public policy making. However, not all interests are created equal. Because higher-socioeconomic-status groups are more likely to join and participate in political organizations, the pressure-group system is biased against representing the views of less organized class, race, gender, and ethnic interests (Verba et al. 1995; Van Deth 1997). Second, political organizations provide public authorities with channels to communicate policy information and provide benefits to their electoral constituencies. Adroit politicians can manipulate public opinion to some degree by selectively targeting which interest groups will receive coveted access to present a case for modifications to pending policy decisions. Public officials and political organizations have a mutual interest in delivering policy successes that permit them both to survive to play the influence game again and again (Browne 1998, pp. 226–228). The fragmentation of political power among numerous policy arenas in the American federal system offers many aggrieved groups several institutional pressure points—legislatures; executive agencies; regulatory bodies; and courts at the local, state, and national levels—through which to raise their demands and promote their preferred solutions onto the public policy agenda for debate and resolution. This duality of political organizations at the interface between the state and its citizenry assures that the interest-group system exerts a crucial, if constitutionally ambiguous, impact on shaping many outcomes of collective political action.

Researchers have made substantial progress in uncovering the evolving techniques deployed by political organizations in lobbying public policy makers on specific issues (Schlozman and Tierney 1986, pp. 261–385; Knoke 1990, pp. 187–213; Baumgartner and Leech 1998, pp. 147–167). Campaign contributions and litigation are relatively

rare methods, while contacting governmental officials (legislative, executive, regulatory), testifying at hearings, presenting research findings, and mobilizing their mass memberships are the most prevalent tactics. But mustering the appearance of grassroots support by hiring consultant firms and lobbying specialists to generate calls and letters may be quickly discredited as phony “astroturf” (Kollman 1998, pp. 157–160). The Internet and the World Wide Web are only the most recent technological innovations to be pressed into the interest-group battle. The impact of political money, primarily unlimited political action committee (PAC) “soft money” election campaign contributions, is a highly emotional topic. Some researchers conclude that a corrupt campaign financing system allows large corporations to enjoy disproportionate political access and influence (Clawson et al. 1998), while others see the corporate capacity to act in unison on political affairs as more problematic and conditional (Mizruhi 1992; Grier et al. 1994). All lobbying methods aim at gaining organizational access to policy makers by winning their attention, communicating with contacts about mutual information needs, and reinforcing for those targets the importance of continuing to pay attention to the organization’s issues (Browne 1998, pp. 68–82). However, the precise conditions under which diverse lobbying tactics exert demonstrable impacts on policy decisions remain elusive.

One increasingly important strategy is collective action by a coalition of political organizations, often competing against an opposing coalition that advocates the contrary policy position. The organizational-state conceptualization of national policy domains emphasizes the shifting nature of short-term networks among organized interest groups mobilizing and coordinating their collective resources in campaigns to pass or defeat particular legislative proposals. The processes by which interorganizational communication networks generate collective action were examined in empirical studies of the U.S. national energy, health, agriculture, and labor policy domains (Laumann and Knoke 1987; Heinz et al. 1993), and in a comparison of U.S., German, and Japanese labor policy making (Knoke et al. 1996). European political scientists have been especially energetic in applying a policy network perspective to understanding how informal bargaining between interest groups

and officials shapes policy outcomes in complex institutional settings (Peterson 1992; Verdier 1995).

Debates among European scholars about the organized representation of societal interests initially concentrated on corporatism as a distinctive form of interest intermediation. Although many definitions of corporatism and neocorporatism abound (Cox and O'Sullivan 1988), the dominant theme concerns how interest groups become incorporated into public policy-making processes through institutionalized access to the levers of state power rather than as seekers of intermittent influence and access that characterize fragmented, pluralist systems such as the United States (Baumgartner and Walker 1989). A corporatist arrangement involves explicit policy negotiations between state agencies and interest groups, followed by implementation of policy agreements through these political organizations, which enforce compliance by their members. The corporatist state takes a highly interventionist role by forming private sector "peak" (encompassing, nonvoluntary, monopolistic) interests groups; delegating to them quasi-public authority to determine binding public policy decisions; and brokering solutions to conflicts (Hirst 1995). In return for a stable share of power, privileged corporatist organizations are expected to discipline their members to accept the imposed policy decisions. Within national labor and other policy domains a closed tripartite network of state agencies, business, and labor organizations collaborates on solutions to such problems as workplace regulation and income distribution, and imposes these compromises on the society. Although much corporatist bargaining occurs primarily within the executive and regulatory sectors, the social partnership aspect of negotiated class conflicts should carry over into the parliamentary arena. The corporatist organizations representing capital, labor, and state interests jointly sponsor legislative proposals originated by agreement with the executive branch. Other interest groups are effectively excluded from participating in these corporatist agreements, resulting in a pattern of cumulative cleavages between them and the corporatist core. These disgruntled, excluded status groups are sources of new social movements against the corporatist monopolies; ecological, antinuclear, feminist, homeless, and immigrant groups are examples of these deprived segments.

This well-ordered corporatist framework seems to be breaking down as a result of the 1986 Single European Act that leads inexorably to an integrated internal market (Mazey and Richardson 1993; Fligstein and Mara-Drita 1996). Simultaneously, a "Europe of regions" is developing, with such areas as Scotland, Brittany, and the Basque country of Spain attaining formal representation and integration into European Union (EU) affairs. The Union is not yet a state, because it still lacks full sovereign power to make and enforce many types of decisions, particularly taxation. Rather, EU policy making is nonhierarchical, open, complex, conflictful, and unpredictable. With Brussels emerging as a supranational forum for resolving social, environmental, producer, and consumer conflicts, new forms of interest representation and lobbying are arising to tackle the expanding EU policy agenda. To varying degrees across different policy domains, the member states are steadily losing control over intergovernmental bargains, while "networks of actors . . . have become guardians of the policy agenda at the subsystemic level of EU governance, over which political controls are often weak or attenuated" (Peterson 1997, p. 7). In sum, the European Union is embarked on a huge, unforeseeable natural experiment that seems likely to transform traditional corporatist state-society relations into a system more closely resembling the "disjointed pluralism" of United States (Mazey and Richardson 1993, p. 24). The situation offers unbounded theoretical and research opportunities.

Despite occasional pessimistic appraisals that "interest-group studies have defined themselves into a position of elegant irrelevance" (Baumgartner and Leech 1998, p. xvii), research on political organizations is thriving at several levels of analysis, from the individual members to organizational political economies to the integration of societal interests into national and supranational polities. Analysts must exert greater effort to link these diverse focuses into a comprehensive explanation of interest organization behaviors situated within their sociopolitical environments. Especially promising avenues include developing formal models of rational social choice at the micro and macro levels; developing models of intra- and interorganizational exchange network; accounting for historical and institutional differences in interest representation processes across diverse national settings; and the functions of nongovernmental agencies

(such as the World Health Organization) and pressure groups (such as Amnesty International) in the world system. Given the vastly expanded sociopolitical functions of modern states in all their permutations, a better understanding of the roles that political organizations play as developers, mediators, expresser, and manipulators of societal interests is indispensable.

(SEE ALSO: *Political Party Systems*; *Voluntary Associations*)

## REFERENCES

- Baumgartner, Frank R., and Beth L. Leech 1998 *Basic Interests: The Importance of Groups in Politics and in Political Science*. Princeton, N.J.: Princeton University Press.
- Baumgartner, Frank R., and Jack L. Walker 1989 "Educational Policymaking and the Interest Group Structure in France and the United States." *Comparative Politics* 21:273-288.
- Berry, Jeffrey M. 1984 *The Interest Group Society*. Boston: Little, Brown.
- 1977 *Lobbying for the People: The Political Behavior of Public Interest Groups*. Princeton, N.J.: Princeton University Press.
- Browne, William P. 1998 *Groups, Interests, and U.S. Public Policy*. Washington, D.C.: Georgetown University Press.
- Clawson, Dan, Alan Neustadt, and Mark Weller 1998 *Dollars and Votes: How Business Campaign Contributions Subvert Democracy*. Philadelphia: Temple University Press.
- Cox, Andrew, and Noel O'Sullivan (eds.) 1988 *The Corporate State: Corporatism and the State Tradition in Western Europe*. Hants, England: Edward Elgar.
- Fligstein, Neil, and Iona Mara-Drita 1996 "How to Make a Market: Reflections on the European Union's Single Market Program." *American Journal of Sociology* 102:1-34.
- Gray, Virginia, and David Lowery 1996 *The Population Ecology of Interest Representation: Lobbying Communities in the American States*. Ann Arbor: University of Michigan Press.
- Grier, Kevin B., Michael C. Munger, and Brian E. Roberts 1994 "The Determinants of Industry Political Activity, 1978-1986." *American Political Science Review* 88:911-926.
- Heinz, John P., Edward O. Laumann, Robert L. Nelson, and Robert H. Salisbury 1993 *The Hollow Core: Private Interests in National Policymaking*. Cambridge, Mass.: Harvard University Press.
- Hirst, Paul 1995 "Quangos and Democratic Government." *Parliamentary Affairs* 48:341-359.
- Knoke, David 1990 *Organizing for Collective Action: The Political Economies of Associations*. New York: Aldine de Gruyter.
- 1988 "Incentives in Collective Action Organizations." *American Sociological Review* 53 (June): 311-329.
- , Franz Urban Pappi, Jeffrey Broadbent, and Yutaka Tsujinaka 1996 *Comparing Policy Networks: Labor Politics in the U.S., Germany, and Japan*. New York: Cambridge University Press.
- Kollman, Ken 1998 *Outside Lobbying: Public Opinion and Interest Group Strategies*. Princeton, N.J.: Princeton University Press.
- Laumann, Edward O., and David Knoke 1987 *The Organizational State: A Perspective on the Social Organization of National Energy and Health Policy Domains*. Madison: University of Wisconsin Press.
- Mazey, Sonia, and Jeremy Richardson 1993 "Introduction: Transference of Power, Decision Rules, and Rules of the Game." In Sonia Mazey and Jeremy Richardson, eds., *Lobbying in the European Community*. New York: Oxford University Press.
- Mizruchi, Mark S. 1992 *The Structure of Corporate Political Action: Interfirm Relations and Their Consequences*. Cambridge: Harvard University Press.
- Moe, Terry 1980 *The Organization of Interests: Incentives and the Internal Dynamics of Political Interest Groups*. Chicago: University of Chicago Press.
- Olson, Mancur, Jr. 1965 *The Logic of Collective Action*. Cambridge: Harvard University Press.
- Peterson, John 1997 "States, Societies and the European Union." *West European Politics* 20(4):1-23.
- 1992 "The European Technology Community: Policy Networks in a Supranational Setting." In David Marsh and R. A. W. Rhodes, eds., *Policy Networks in British Government*. Oxford: Oxford University Press.
- Salisbury, Robert H. 1994 "Interest Structures and Policy Domains: A Focus for Research." In William Crotty, Mildred A. Schwartz, and John C. Green, eds., *Representing Interests and Interest Group Representation*. Washington: University Press of America.
- Schlozman, Kay L., and John T. Tierney 1986 *Organized Interests and American Democracy*. New York: Harper and Row.
- Van Deth, Jan W. (ed.) 1997 *Private Groups and Public Life: Social Participation, Voluntary Associations and Political Involvement in Representative Democracies*. London: Routledge.

- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady 1995 *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, Mass.: Harvard University Press.
- Verdier, Daniel 1995 "The Politics of Public Aid to Private Industry: The Role of Policy Networks." *Comparative Political Studies* 28:3-42.
- Vogel, David 1996 *Kindred Strangers: The Uneasy Relationship between Politics and Business in America*. Princeton, N.J.: Princeton University Press.
- Walker, Jack L. 1983 "The Origins and Maintenance of Interest Groups in America." *American Political Science Review* 77:390-406.

DAVID KNOKE

## POLITICAL PARTY SYSTEMS

### DEFINITIONS

Political parties have been defined both normatively, with respect to the preferences of the analyst, and descriptively, with respect to the activities in which parties actually engage. Normative definitions tend to focus on the representative or educational functions of parties. Parties translate citizens' preferences into policy and also shape citizens' preferences. Parties are characterized as "policy seeking." Thus, Lawson (1980) defines parties in terms of their role in linking levels of government to levels of society. She states, "Parties are seen, both by their members and by others, as agencies for forging links between citizens and policy-makers." Von Beyme (1985, p. 13) lists four "functions" that political parties generally fulfill: (1) the identification of goals (ideology and program); (2) the articulation and aggregation of social interests; (3) the mobilization and socialization of the general public within the system, particularly at elections; and (4) elite recruitment and government formation.

Descriptive definitions usually stay closer to Max Weber's observation that parties are organizations that attempt to gain power for their members, regardless of constituent wishes or policy considerations. Parties are characterized as "office seeking." "Parties reside in the sphere of power. Their action is oriented toward the acquisition of social power . . . no matter what its content may be" (Weber 1968, p. 938). Schumpeter ([1950] 1975) applies this type of definition to a democratic setting. He argues that parties are organizations

of elites who compete in elections for the right to rule for a period. Or as Sartori (1976, p. 63) puts it, "a party is any political group identified by an official label that presents at elections, and is capable of placing through elections (free or nonfree), candidates for public office."

The present article employs a descriptive definition but also investigates how well parties perform functions described in the normative definitions. Thus, a party system may be characterized as the array or configuration of parties competing for power in a given polity. The focus here will be almost exclusively on Western-style democracies.

### ORIGINS

Von Beyme (1985) suggests three main theoretical approaches to explain the emergence of political parties: institutional theories, historical crisis situation theories, and modernization theories. (Also see LaPalombara and Weiner 1966.)

**Institutional Theories.** Institutional theories explain the emergence of parties as largely due to the way representative institutions function. Parties first emerge from opposing factions in parliaments. Continuity, according to such theories, gives rise to stable party constellations based on structured cleavages. These theories seem most relevant to countries with continuously functioning representative bodies, such as the United States, Britain, Scandinavia, Belgium, and the Netherlands. However, institutional theories do not explain developments well in some countries, such as France, because continuity of parliament has been absent, and the parliament's strength and independence has come repeatedly into question. The timing of the franchise is also relevant, but its effect is indeterminate because a party system has often been partly established before the franchise was fully extended. Moreover, liberal bourgeois parties that have helped establish parliamentary government have often been opposed to extending the franchise to the lower classes, while leaders such as Bismarck or Napoleon III have sometimes extended the franchise in nonparliamentary systems for tactical political reasons (von Beyme 1985, p. 16). Likewise, Lipset (1985, chap. 6) argues that a late and sudden extension of the franchise has sometimes contributed to working-class radicalism because the lower classes were not slowly



integrated into an existing party system. Voting laws can also affect the structure of the party system. Single-member districts, with a first-past-the-post plurality winner, as in the United States and in Britain, are said to encourage a small number of parties and ideological moderation (competition for the center). National lists, with proportional representation (PR), are said to encourage multipartism (fractionalization) and ideological polarization. However, PR may have this effect only if it is implemented concurrently with the extension of the franchise, because already-established parties may otherwise be well entrenched and leave little room for the generation of new parties. Lijphart (1985) notes that voting laws may also affect other features of political life, such as voter turnout and efficacy or system legitimation, but that these effects have not been extensively investigated.

**Crisis Theories.** Critical junctures in a polity's history may generate new political tendencies or parties. Crisis theories are especially associated with the Social Science Research Council's (SSRC's) project on Political Development (e.g., LaPalombara and Weiner 1966; Grew 1978). According to SSRC scholars, five such crises can be identified in political development: the crises of national identity, state legitimacy, political participation, distribution of resources, and state penetration of society. The sequence in which these crises are resolved (if only temporarily) and the extent to which they may coincide can affect the emerging party system. Thus, Britain's well-spaced sequence contributed to the moderation of its party system. The recurrent piling up of crises in Germany from the mid-nineteenth century to the mid-twentieth century, and the attempt to solve problems with penetration (strong-state measures) contributed to the fragmentation, polarization, and instability of its party system. The piling up of all five crises in mid-nineteenth century America contributed to the emergence of the Republican Party—and the second party system. From a slightly different perspective, von Beyme (1985) notes three historical crisis points that have generated parties. First, the forces of nationalism and of integration during the nation-building process have often taken on roles as political parties. Second, party systems have been effected by breaks in legitimacy as a result of dynastic rivalries, as between Legitimists, Orleanists, and Bonapartists in mid-nineteenth

century France. Third, the collapse of parliamentary democracy to fascism has produced characteristic features in the party systems of post-authoritarian democracies: "a deep distrust of the traditional right; an attempt to unify the centre right; [and] a split on the left between the socialists and the Communists" (p. 19).

**Modernization Theories.** Some theories, following the tenets of structural functionalism, argue that "parties will not in fact materialize unless a measure of modernization has occurred" (LaPalombara and Weiner 1966). Modernization includes such factors as a market economy and an entrepreneurial class, acceleration of communications and transportation, increases in social and geographic mobility, increased education and urbanization, an increase in societal trust, and secularization. LaPalombara and Weiner argue that the emergence of parties requires one, or both, of two circumstances: citizens' attitudes may change, so that they come to perceive a "right to influence the exercise of power," or some group of elites or potential elites may aspire to gain or maintain power through public support. Clearly, not all elements of modernization are necessary, since the first party systems (in the United States and Britain) emerged in premodern, agrarian, and religious societies. Also, not all modernization theories are functionalist. Thus, Moore (1966) and others have suggested the emergence of a bourgeoisie increases the probability of the emergence of democracy.

Probably the most influential theory of the origins of party systems is by Lipset and Rokkan (1966) and Lipset (1983). While ostensibly anchored in Parsonsian functionalism, theirs is a comparative-historical approach that borrows from each of the categories listed here. According to Lipset and Rokkan, the contours of the party systems for western European states can be understood in the context of the specific outcomes of three historical episodes. The three crucial junctures are (1) the Reformation, "the struggle for the control of the ecclesiastical organizations within the national territory"; (2) the "Democratic Revolution," related to a conflict over clerical/secular control of education beginning with the French Revolution; and (3) the opposition between landed and the rising commercial interests in the towns early in the "Industrial Revolution." A significant fourth struggle between owners and workers

emerges in the later stages of the Industrial Revolution. Lipset and Rokkan suggest that the shape of current party systems was largely determined during the stages of mass mobilization in the pre-World War I West.

Following Lipset and Rokkan, von Beyme (1985, pp. 23–24) lists ten types of parties that have emerged from this historical development: (1) liberals in conflict with the old regime, that is, in conflict with: (2) conservatives; (3) workers' parties against the bourgeois system (after c. 1848) and against left-wing socialist parties (after 1916); (4) agrarian parties against the industrial system; (5) regional parties against the centralist system; (6) Christian parties against the secular system; (7) communist parties against the social democrats (after 1916–1917) and anti-revisionist parties against "real Socialism"; (8) fascist parties against democratic systems; (9) protest parties in the petty bourgeoisie against the bureaucratic welfare state system (e.g., Poujadisme in France); (10) ecological parties against a growth-oriented society. No one country contains all ten sorts of parties, unless one includes splinter groups and small movements.

### PARTY SYSTEMS AND SOCIETY

Even under a purely office-seeking definition, parties in a democracy must have some connection to society since they have to appeal to voters' material or ideal interests. Yet the connection between the party system and social structure or social values is rather weak in most countries—and much weaker than would be expected under a theory that sees parties as mediating between society and the state. In many cases, organizational or institutional factors may be much more important than social factors in determining party strength.

**Social Cleavages.** The party types listed above clearly have some connection to divisions or cleavages in society. Parties may seek to represent social classes, religious denominations, linguistic communities, or other particular interests. Three types of politically relevant social cleavages may be identified:

1. Positional cleavages correspond to a party supporter's place in the social structure. This may be an ascriptive position into which one is born, such as race, ethnicity, or gender, or it may be a social structural

position, such as social class or religious denomination, which one might be able to change in the course of a lifetime. Of course, the distinction between ascriptive and social structural position is not absolute, but may itself be partly determined by social norms. Also, against Marxist expectations, class determinants of party support are generally overshadowed by racial, ethnic, religious, regional, or linguistic determinants, when these are also present. One explanation for this finding is that, while one can split differences on class (especially monetary) policies, similar compromises are much more difficult where social "identity" is concerned.

2. "Behavioral" cleavages, especially membership, generally have a greater impact on party support than positional cleavages. Studies have shown that while working-class status is mildly correlated to support for leftist parties, union membership is quite strongly correlated. And while religious denomination is correlated to support for religious parties (e.g., Catholics and Christian Democrats in Germany), strength of belief or church attendance is much more strongly correlated.
3. Ideological cleavages are preferences, values, worldviews, and the like, which may not correspond entirely to one's position in society. Indeed, ideological orientations may overshadow positional cleavages as a determinant of partisan preferences. For instance, several of the ostensibly working-class communist parties of western Europe have traditionally drawn large percentages of their support from middle-class leftists.

Not all cleavages or issues that exist in a society are politically relevant at any given time, or if they are, they may not correspond to party support. One can distinguish between latent and actual cleavages around which politics are mobilized. Some cleavages may remain latent for a very long time before becoming politicized. For instance, women's issues had been relevant for decades before the "gender gap" emerged in the elections of the 1980s. One can also consider the process of politicization as a continuum that begins when a new social division or issue emerges, develops into

a (protest) movement, then a politicized movement, and ends—at an extreme—with the creation of a new political party or the capture of an existing party. Of course, this process may be halted or redirected at any stage.

**Party Loyalty and Party System Change: Alignment, Realignment, Dealignment.** Parties may persist over time, and the party system alignment may be stable. There are several possible reasons for this:

1. The social cleavages around which a party was built may persist.
2. Voters may grow up in a stable party system and be socialized to support one or another party. Studies show that when a new cleavage line emerges in party alignment, it begins with the youngest generations. These generations then carry their new party loyalties with them throughout their lives, though perhaps to a decreasing extent if the events that originally motivated them fade over time. Likewise, older generations tend to resist alignments along newly emerging cleavage lines because they remain loyal to the parties they began to support in their own youth.
3. Parties may become organizationally entrenched and difficult to dislodge. Even if cleavages or issues emerge that cause voter dissatisfaction with existing parties, these parties may have the organizational resources to outmaneuver new movements or parties. They may be able to “steal” the new parties’ issues and absorb or coopt their constituencies, or they may be able to stress other issues that distract voters from the new issues.

However, newly emerging cleavage structures may overwhelm these inertial tendencies. The party system may respond in three ways to new social cleavages. The first two are processes of party “realignment”:

1. New parties may be formed to appeal to the new constituencies. A classical example is the emergence of the British Labour Party in the late nineteenth and early twentieth centuries when the Liberals and Conservatives did not pay sufficient attention to the concerns of the growing

working classes. The more recent emergence of Green parties in some European countries is another example. The creation of the American Republican Party in the 1850s shows the explosive impact a new party can have: Lincoln’s election precipitated the South’s secession.

2. Existing parties may change their policies to appeal to new constituencies. For instance, existing parties seem now to be in the process of killing the European Greens by adopting their issues. Perhaps the best example of this process is found in American history. Bryan’s Democrats moved to absorb the Populist Party, and Al Smith’s and Franklin Roosevelt’s Democrats moved to absorb the growing urban ethnic constituencies (Burnham 1970; Chambers and Burnham 1975).
3. If neither of these changes occurs, there may be a period of “dealignment” in which much of the population—especially new constituencies—is alienated from all parties, and turnout or political participation declines. New constituencies may organize themselves into pressure groups or social movements that fail either to form new parties or to capture existing parties. Existing parties may become internally more heterogeneous and polarized, single-issue actions may proliferate, referenda may increase, and citizen action groups may simply bypass parties. Scholars since the mid-1960s have debated whether Western politics are going through a period of realignment or dealignment (Dalton et al. 1984). Of course, both processes may be occurring: dealignment may be a way-station on the road to party realignment.

### STRUCTURAL FEATURES

Certain structural features of the party system may be important independently of parties’ connections to society.

**Representativeness.** The electoral system determines how votes are translated into seats in the legislature. The results can vary widely. At one extreme, a system of proportional representation

(PR) with a single national list enables even tiny parties to get representatives into the legislature. Thus, if 100 parties each received 1 percent of the vote, each would receive 1 seat in a 100-seat legislature. Such systems put no obstacles in the way of party system fragmentation. At the other extreme, first-past-the-post plurality voting with single-member constituencies tends to overrepresent larger parties and underrepresent smaller parties. Thus, if party A won 40 percent of the vote in every district, and parties B and C each won 30 percent of the vote in every district, party A would get all the seats in the legislature, and parties B and C would get none at all. Such systems discourage party system fragmentation. Still, regionally concentrated minority parties tend to be less underrepresented than minority parties whose support is spread across all districts. If 100 parties were completely concentrated in each of 100 districts, the electoral system could not prevent fragmentation. Some election systems combine features. German voters have two votes, one for a district candidate and one for a party list. If any candidate receives a majority in his or her district, that candidate gets a seat. The remaining seats are allocated proportionately according to the list votes. Furthermore, a party must receive at least 5 percent of the national vote to get any seats from the list portion. This system attempts to reduce party system fragmentation and at the same time to reduce overrepresentation and underrepresentation. It was once thought that PR reduces government stability and endangers democracy. However, recent research gives little support for this proposition: "electoral systems are not of overriding importance in times of crisis and even less in ordinary times" (Taagepera and Shugart 1989, p. 236).

**Volatility.** Party system volatility, or fluctuations in electoral strength, encompasses several different processes (Dalton et al. 1984; Crewe and Denver 1985). It includes the gross and net flow of voters between parties, as well as into and out of the electorate because of maturity, migration, death, and abstention. It also includes realignment and dealignment: changes in the electoral alignment of various constituencies, and the overall weakening of party attachments. Scholars have long debated whether electoral volatility contributed to the collapse of democracies in the 1930s, especially the mobilization of first-time or previously alienated voters. Recently, Zimmermann and Saalfeld

(1988) concluded that volatility encouraged democratic collapse in some, but not all, countries. Studies also show that most postwar antidemocratic "surge" parties draw support disproportionately from voters who are weakly attached to parties or weakly integrated in politically mobilized subcultures such as labor, religious, or ethnic organizations. Yet volatility and protest do not always flow in an antidemocratic direction. On the contrary, they are also normal components of democratic politics. Few would argue that the New Deal realignment harmed American democracy or that most new-left or ecology movements are antidemocratic. In order for volatility to cause trouble for democracy, it must be accompanied by antidemocratic sentiments. Indeed, massive vote switching among *democratic* parties may be the best hope for *saving* democracy during a crisis. Everything depends on the propensity of voters to support antidemocratic parties.

**Fragmentation.** In the wake of World War II, some scholars argued that the fragmentation of party systems, partly caused by proportional representation, contributed to the collapse of European democracies. In a fragmented party system, they argued, there are too many small parties for democratic representation and effective government. Citizens are confused and alienated by the large array of choices. Because parties have to form coalitions to govern, voters' influence over policy is limited, and they become further disenchanted with democracy. With so many small parties, governing coalitions can be held hostage to the wishes of very minor parties. Empirical studies show some support for these theses. Fragmentation is associated with reduced confidence in government and satisfaction with democracy. Governments in fragmented party systems tend to be unstable, weak, and ineffective in addressing major problems. However, other scholars argue that party-system fragmentation is not the main culprit. Fragmentation contributes to problems, but other factors are more important. Since fragmented party systems are often composed of blocs of parties (as in, e.g., the Netherlands and Italy), voters have less difficulty reading the terrain than alleged. Besides, party system polarization may contribute to governmental instability and ineffectiveness more than to fragmentation. Scholars have looked at this possibility in both the interwar period and the

postwar period. While the evidence is not overwhelming, it tends to support the thesis.

**Polarization.** Sartori's model of "polarized pluralism" (1966, 1976) is the most influential account of party system polarization. In a polarized party system, according to Sartori, a large (but not majority) party governs more or less permanently in unstable coalitions with various other parties. At least one extremist (antisystem) party is in quasi-permanent opposition. Extremist parties are sufficiently unacceptable to others that they cannot form alternative coalitions, but they are strong enough to block alternative coalitions that do not include themselves. Sartori argues that this leads to stagnation and corruption at the center, frustration and radicalization at the periphery, and instability among governing coalitions. He cites Weimar Germany, Fourth Republic France, and contemporary Italy as examples. Much empirical evidence supports Sartori's model. Polarization is associated with illiberal values in postauthoritarian democracies such as West Germany, Austria, Italy, and Spain.

The dynamic may also work in reverse. When intolerant and distrustful relations among political actors were institutionalized by constitutional guarantees in some postauthoritarian countries, they became crystallized in a polarized party system. Cross-national research shows that polarization harms other aspects of democracy, as well. Polarization is negatively related to democratic legitimation and trust in government, and is positively associated with cabinet instability. However, other elements of Sartori's model have been disputed. In particular, studies in the early 1980s of Italy—the model's current exemplar—called into question Sartori's claim that polarized pluralism generates extremism and thus harms democracy. These studies claimed that the Italian Communists had moderated and that the centrist Christian Democrats had become less intolerant of them. However, the studies' own evidence were not entirely persuasive, and subsequent developments—while not reversing course—do not present a decisive break with earlier patterns.

### COALITIONS

Single-party government in Western democracies is relatively rare (Laver and Schofield 1990). The multiparty systems of most countries necessitates

coalition government. Even in two-party America, a president and Congress of different parties produce a kind of coalition government. (Indeed, internal party discipline is so weak in America, as well as in some parties in Italy, Japan, and other countries, that one can characterize parties themselves as coalitions of political actors.) Most work on coalition government attempts to predict which parties get into office. One of the most influential theories predicts that "minimum connected winning" (MCW) will form most often. This theory combines office-seeking and policy-seeking approaches, predicting that parties will form bare-majority coalitions (so that the spoils can be divided among the smallest number of winners) among contiguous parties on the ideological dimension (so that there is not too much disagreement about policy). MCW theory succeeds fairly well in predicting coalitions in unidimensional party systems, but less well in multidimensional systems, which are often fragmented, polarized, and/or based on rather heterogeneous societies. Likewise, research suggests that in unidimensional systems, offices are most often allocated among the winning parties proportionately to their electoral strength. In multidimensional systems, however, offices are allocated less according to parties' electoral strength than according to their "bargaining" strength, that is, how much they are needed to complete the majority. Thus, if three parties won 45 percent, 10 percent, and 45 percent of the vote, the small party would have just as much bargaining strength as either of the larger parties.

Research also shows that party-system fragmentation and polarization and the presence of antisystem parties all contribute to cabinet instability. Theorists have sometimes posited that cabinet instability leads to instability of democracy—that it may reduce governments' capacity to solve problems effectively, and that this may reduce the regime's legitimacy. Yet research gives only mixed support for this conjecture. Investigators have found that cabinet instability tends to depress the electorate's evaluation of "the way democracy works," but its effects on other measures of democratic legitimation and confidence in government are inconsistent. Research on contemporary democracies shows that cabinet instability is related to civil disorder and governmental ineffectiveness. But research on the period between the world wars indicates that cabinet instability cannot be

definitely tied to the collapse of democracy. Cabinets in France and Belgium were as unstable as those in Germany and Austria, but only the latter democracies collapsed (British and Dutch cabinets were more stable). Why is cabinet instability not more clearly tied to problems for democracy? One possibility is that cabinet instability simply reflects the severity of problems. Just as electoral volatility may reflect citizens' desire for change, cabinet instability may reflect elites' flexible response to the problems. Neither of these need reflect a desire for a regime change, simply for a policy change. Indeed, cabinet *immobility* might be more damaging to effectiveness and democratic legitimation if problems are severe enough. In this respect, cabinet instability, like electoral volatility, probably has an indeterminate effect on democratic survival.

Oversized grand-coalition governments also have ambiguous effects on liberal democracy. The most important theory is Lijphart's (1977, 1984) model of "consociational democracies," plural societies with high levels of intercommunal conflict. In such polities, parties are unwilling to go into opposition because they risk losing too much and because party strength—closely tied to the size of the ascriptive communities—changes too slowly to make their return to office likely. Thus, formal opposition could lead to more extreme conflict. The alternative is a grand coalition government of all major parties, combined with a degree of federalism and proportional allocation of state services according to party or community size. Since potential conflict is too dangerous, open opposition is delegitimated and suppressed. In this respect, consociational procedures are intended to be a method for *reducing* extreme underlying intercommunal conflict through contact among opponents (at the elite level), which promotes trust. If these measures succeed, the "game among players" can move to one in which moderate conflict and tolerance of opponents becomes legitimated. This appears to have succeeded in the Netherlands and Austria, and failed most miserably in Lebanon. On the other hand, if grand coalitions are formed in societies *without* extreme underlying conflict, they may *initiate* a vicious circle of intolerance and delegitimation. To form a grand coalition, prosystem parties generally move closer to the center of the policy spectrum than

they would otherwise do. This move may leave their more militant (but still prosystem) constituents politically homeless, and they may seek harder positions in a more extremist party or movement. These constituents do not so much abandon their party as the party abandons them. Thus, if a grand coalition submerges a moderate competitive structure, it can generate polarization. The grand coalition government of 1966–1969 in West Germany, a country with little intercommunal conflict, was probably largely responsible for the rise of antisystem voting at the time. If the grand coalition government had not ended fairly quickly, it might have caused serious problems for West German democracy.

### RESEARCH DEVELOPMENTS IN THE 1990s

Research on political parties and party systems has continued to stream unabated in the 1990s, yet many of the basic principles outlined above continue to hold true. Three important research areas may be mentioned. First, scholars have sought to understand the role of party systems in democratization, especially in central and eastern Europe, but in other regions as well. Second, the study of political extremism has been knit more closely with the study of party systems. Third, recent stock taking in the field of political legitimation has highlighted the importance of party systems.

The "third wave" of democratization, beginning with transitions in southern Europe in the mid-1970s, and continuing with transitions in Latin America, East Asia, and central and eastern Europe, is one of the most important social and political developments of the last quarter of the twentieth century. Scholars seeking explanations for the relative success or failure of democratic transition and, especially, consolidation have generally highlighted the importance of well-functioning party systems. Thus, Huntington (1991, chap. 6) argues that party-system polarization is one of the greatest hazards to democratization (also see Di Palma 1990; Lipset 1994). Theorists of democratic transitions have pointed to the importance of "pacting" between authoritarian-regime soft-liners and democratic-opposition moderates, and to the exclusion of regime hard-liners and antiregime extremists (O'Donnell and Schmitter 1986; Karl

and Schmitter 1991). The importance of moderation during the transition period, prior to the legalization of a party system, parallels the importance of moderation of a party system within an existing democracy (Weil 1989). Empirical studies of democratization in Latin America (Remmer 1991), Central and Eastern Europe (Fuchs and Roller 1994; Toka 1996; Wessels and Klingemann 1994), and East Asia (Shin 1995) tend to support this thesis—as do general, comparative treatments of democratization (Linz and Stepan 1996).

The study of political extremism has taken party systems into account more fully in the 1990s than had perhaps previously been the case. Earlier studies often characterized extremism in terms of psychological predispositions, socialization, or economic dislocations. These accounts tended to focus on personal distress—sometimes in absolute terms, but sometimes in terms of reference groups and relative deprivation—and were often couched in functionalist theories of social dislocation in the course of social modernization. A later wave of extremism research focused more on resource mobilization within social movements. It was not deprivation (absolute or relative) that created extremism, according to this view, but the ability to organize. A third wave of extremism research has emphasized political “opportunity space,” gaps or niches in the opposition structure, which political entrepreneurs can fill if they are skillful. Extremism often arises not so much because conditions have worsened, nor because groups have newly organized, as because existing parties within the party system have vacated certain ideological positions and opened competitive opportunities or niches for extremists. Mainstream parties may vacate these niches because they enter or leave office, or because they feel they need to compete more effectively with another party. The reader will notice that it is not so much that these three accounts contradict each other as that they are nested, with the first most specific and the last most general. Perhaps the most important recent study of right-wing extremism in Western polities is Kitschelt and McGann (1995). Other useful recent collections of essays include Weil (1996) and McAdam and colleagues (1996).

Studies of legitimation, confidence, and trust continue to attend to the effects of parties and

party systems. Recent surveys of the literature show that party systems do not always or uniformly have an influence, but when they do, a moderate opposition structure is most conducive to these forms of political support. Polarization, grand coalitions, and “cohabitation” (“divided government” in America) do not tend to promote legitimation, confidence, and trust (see Fuchs et al. 1995; Listhaug 1995; Listhaug and Wiberg 1995).

Finally, a few recent general contributions to the literature may be listed. Important recent books that bring the field up to date include Ware (1996) and Mair (1997). Also, a new journal devoted to political parties and party systems, *Party Politics*, from Sage Publications, began publication in 1995 and has become a major outlet for scholarship in this field.

## REFERENCES

- Burnham, Walter Dean 1970 *Critical Elections and the Mainsprings of American Politics*. New York: Norton.
- Chambers, William Nisbet, and Walter Dean Burnham (eds.) 1975 *The American Party Systems*, 2nd ed. New York: Norton.
- Crewe, Ivor, and David Denver (eds.) 1985 *Electoral Change in Western Democracies: Patterns and Sources of Electoral Volatility*. New York: St. Martin's.
- Dalton, Russell J., Stephen C. Flanagan, and Paul A. Beck 1984 *Electoral Change in Advanced Industrial Democracies*. Princeton, N.J.: Princeton University Press.
- Di Palma, Giuseppe 1990 *To Craft Democracies: An Essay on Democratic Transitions*. Berkeley: University of California Press.
- Fuchs, Dieter, Giovanna Guidorossi, and Palle Svensson 1995 “Support for the Democratic System.” In H. D. Klingemann and D. Fuchs, eds., *Citizens and the State*. New York: Oxford University Press.
- Fuchs, Dieter, and Edeltraud Roller 1994 “Cultural Conditions of the Transformation to Liberal Democracies in Central and Eastern Europe,” WZB Discussion Paper FS III 94-202. Wissenschaftszentrum Berlin, Berlin.
- Grew, Raymond (ed.) 1978 *Crises of Political Development in Europe and the United States*. Princeton, N.J.: Princeton University Press.
- Huntington, Samuel P. 1991 *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.

- Karl, Terry Lynn, and Philippe C. Schmitter 1991 "Modes of Transition in Latin America, Southern and Eastern Europe." *International Social Science Journal* 128:269-284.
- Kitschelt, Herbert, and Anthony J. McGann 1995 *The Radical Right in Western Europe: A Comparative Analysis*. Ann Arbor: University of Michigan Press.
- La Palombara, Joseph, and Myron Weiner (eds.) 1966 *Political Parties and Political Development*. Princeton, N.J.: Princeton University Press.
- Laver, Michael, and Norman Schofield 1990 *Multiparty Government. The Politics of Coalition in Europe*. New York: Oxford University Press.
- Lawson, Kay (ed.) 1980 *Political Parties and Linkage*. New Haven, Conn.: Yale University Press.
- Lijphart, Arend 1977 *Democracy in Plural Societies*. New Haven, Conn.: Yale University Press.
- 1984 *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries*. New Haven, Conn.: Yale University Press.
- 1985 "The Field of Electoral Systems Research: A Critical Survey." *Electoral Studies* 4:3-14.
- Linz, Juan J., and Alfred Stepan 1996 *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore: Johns Hopkins University Press.
- Lipset, Seymour Martin 1983 "Radicalism or Reformism: The Sources of Working-Class Politics." *American Political Science Review* 77:1-18.
- 1994 "The Social Requisites of Democracy Revisited." *American Sociological Review* 59:1-22.
- , and Stein Rokkan (eds.) 1967 "Cleavage Structures, Party Systems and Voter Alignments." In Seymour Lipset and Stein Rokkan *Party Systems and Voter Alignments*. New York: Free Press.
- Listhaug, Ola 1995 "The Dynamics of Trust in Politicians." In H. D. Klingemann and D. Fuchs, eds., *Citizens and the State*. New York: Oxford University Press.
- , and Matti Wiberg 1995 "Confidence in Political and Private Institutions." In H. D. Klingemann and D. Fuchs, eds., *Citizens and the State*. New York: Oxford University Press.
- McAdam, Doug, John D. McCarthy, and Mayer N. Zald 1996 *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. New York: Cambridge University Press.
- Mair, Peter 1997 *Party System Change: Approaches and Interpretations*. New York: Oxford University Press.
- Moore, Barrington, Jr. 1966 *Social Origins of Dictatorship and Democracy*. Boston: Beacon.
- O'Donnell, Guillermo, and Philippe C. Schmitter 1986 *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore: Johns Hopkins University Press.
- Remmer, Karen L. 1991 "The Political Impact of Economic Crisis in Latin America in the 1980s." *American Political Science Review* 85:777-800.
- Sartori, Giovanni 1966 "European Political Parties: The Case of Polarized Pluralism." In Joseph LaPalombara and Myron Weiner, eds., *Political Parties and Political Development*. Princeton: Princeton University Press.
- 1976 *Parties and Party Systems: A Framework for Analysis*. Cambridge: Cambridge University Press.
- Schumpeter, Joseph (1950) 1975 *Capitalism, Socialism and Democracy*. New York: Harper Colophon.
- Shin, Doh Chull 1995 "The Democratization of Korean Politics and Culture in Progress and Repose: Public Opinion Survey Findings, 1988-1994." Paper presented at the International Conference, 50 Years of Korean Independence, 50 Years of Korean Politics. Sponsored by the Koreana Political Science Association, Seoul.
- Taagepera, Rein, and Matthew Soberg Shugart 1989 *Seats and Votes. The Effects and Determinants of Electoral Systems*. New Haven, Conn.: Yale University Press.
- Toka, Gabor 1996 "Parties and Electoral Choices in East-Central Europe." In G. Pridham and P. G. Lewis, eds., *Stabilising Fragile Democracies: Comparing New Party Systems in Southern and East Europe*. London: Routledge.
- von Beyme, Klaus 1985 *Political Parties in Western Democracies*. Gower, Wales: Aldershot.
- Ware, Alan 1996 *Political Parties and Party Systems*. New York: Oxford University Press.
- Weber, Max 1968 *Economy and Society*. Berkeley: University of California Press.
- Weil, Frederick D. 1989 "The Sources and Structure of Legitimation in Western Democracies: A Consolidated Model Tested with Time-Series Data in Six Countries since World War II." *American Sociological Review* 54:682-706.
- 1996 *Research on Democracy and Society: Volume 3, Extremism, Protest, Social Movements, and Democracy*. Greenwich, Conn.: JAI.
- Wessels, Bernhard, and Hans Dieter Klingemann 1994 "Democratic Transformation and the Prerequisites of Democratic Opposition in East and Central Europe," Working Paper FS III 94-201. Wissenschaftszentrum Berlin für Sozialforschung, Berlin.



Zimmermann, Ekkart, and Thomas Saalfeld 1988 "Economic and Political Reactions to the World Economic Crisis of the 1930s in Six European Countries." *International Studies Quarterly* 32:305–334.

FREDERICK D. WEIL

## POLITICAL SOCIOLOGY

Two distinct but converging intellectual traditions have defined the field of political sociology: the social stratification tradition pioneered by Karl Marx and Frederick Engels; and the organizational tradition originated by Max Weber and Robert Michels (Lipset 1981). In the first, political sociology is defined broadly as the study of social power in all institutional sectors of society with a primary emphasis on the state and its structural roots in the class system. This tradition takes a holistic view of social structure and change, arguing that the class system determines the organization of the state and political action. The state is conceived as the institutional structure whose central function is maintaining the social order and is thus examined in terms of its functions. The second tradition defines political sociology more narrowly in terms of the organization of political groups and political leadership with primary emphasis on the structure of the state and the groups that compete for control over the state. The state is conceived as the legitimate monopoly on the means of violence. This approach emphasizes the informal and formal organization of political parties, interest groups and social movements, their links to the governmental bureaucracy and formal centers of policy making, the legitimating myths that are used to justify the system of rule, the organization of the legal system, and the sources and impact of public opinion, including the organization of the mass media and electoral politics.

As societies modernize and become more complex, it becomes increasingly difficult to distinguish these two traditions. Institutions become more complex and differentiated, thus developing their own distinctive autonomy while at the same time being shaped by the larger system of power. In response, analysts have blended these approaches together, synthesizing arguments drawn from the competing perspectives. We begin with a brief

summary of the classic ideas and their bearing on contemporary work and then examine attempts to integrate these approaches.

## KARL MARX AND THE THEORY OF THE STATE

Karl Marx and Frederick Engels developed two distinct theories of the state: an instrumental theory; and a structuralist argument (Carnoy 1979). In the first, the state is a tool or instrument of the dominant class, used to protect the property system and impose order through force and ideological manipulation. The dominant class is a ruling class, meaning that it simultaneously dominates the economy and the political system. This class is socially cohesive and politically unified, which gives it direct control over the state. The state is the institutionalization of power or, to draw on Max Weber's conception (1947), the legitimate monopoly on the means for force in society and, as a tool of the upper class, consolidates upper-class power by force and fraud.

In the contemporary period, this argument inspired Mills's theory of the "power elite" (1956), defined in terms of the cohesive leadership group that unifies the corporate rich, the political directorate, and the military elite, as well as Domhoff's theory of business dominance (1979, 1990, 1998) and Useem's inner-circle thesis (1984). In these arguments, the capitalist class is seen as socially cohesive—integrated through exclusive private clubs, prep schools and universities, debutante balls, and interlocking corporate directorships. These networks create an inner-circle leadership group that controls the largest multinational corporations and is central to these various networks. Domhoff (1998) argues that this upper class rules through four processes: (1) policy planning in which leading inner-circle-controlled policy organizations, such as the Business Roundtable, the U.S. Chamber of Commerce, and the Brookings Institution, develop the major policy proposals that are eventually adopted; (2) campaign contributions that select the candidates for electoral office (Clawson et al. 1992); (3) special-interest lobbying for specific firms and industries; and (4) ideological control through publicity campaigns that control the political agenda and mold public opinion so as to minimize opposition. Thus the ruling class

constitutes an organized leadership group with considerable cohesion and political unity.

Marx and Engels also advanced a structural theory of state, treating it as the autonomous product of class struggles. Thus, in his essay *The Eighteenth Brumaire of Louis Bonaparte* ([1852] 1964), Marx argued that the political development of the working-class movement combined with disorganization and internal conflicts within the ruling class led to a military dictatorship. The capitalist class was too divided to rule directly and thus had to be protected by a military dictator. Extending this argument, Neumann (1942) explained the rise of fascist dictatorship in Germany during the 1930s in terms of a combination of working-class mobilization, a disorganized bourgeoisie and strong autocratic political traditions. Drawing on Gramsci's ideas about the "modern prince" of modern bourgeois civil society (1957), Poulantzas (1973) advanced a structuralist theory of the state, arguing that market competition disorganizes the capitalist class, requiring that the state operate as a "relatively autonomous" institution that organizes the capitalist class into a hegemonic bloc while disorganizing the working class. Critics have pointed out that this functionalist argument lacks a specific mechanism for class rule (Skocpol 1981). In response, Offe (1984) and Block (1987) argued that in capitalism the state is barred from entering profit-making enterprise, which makes state managers structurally dependent on capitalists to make investments and thereby create employment, taxes, and economic growth. State managers are thus structurally pressured to create capital accumulation and act autonomously to promote reforms that rationalize and stabilize the capitalist system.

Thus the stratification tradition has gradually incorporated arguments from the organizational tradition, focusing on leadership groups and the autonomy of political institutions. By conceiving political institutions as independent but structurally dependent on the capitalist economy, these analysts have synthesized Marxian arguments with organizational arguments. Still this position has been subjected to several criticisms. First is that social stratification is not primarily a question of class but one of prestige and exclusionary social practices (Weber 1947; Parkin 1979). As Weber (1947) argued, classes are defined by their market position, which rarely provides sufficient cohesion for successful political mobilization. Status groups,

in contrast, are based on shared values and a code of honor, thus readily mobilizing for political action. Thus the political struggle over the state is typically not about class but about the lifestyles and prestige of status communities. Second, political institutions have their own autonomous logic and interests, operating independent of the interests and action of classes and other organized groups. Thus Michels ([1911] 1962) argued that formal organization creates oligarchic leadership which is able to divert political organizations from their official goals to promoting the interests of the permanent staff. Similarly, Weber (1947) argued that rational-legal authority and bureaucratic organization tend to replace traditional and charismatic rule, thus creating the "iron cage" of the modern bureaucratic state. Both arguments underscore the independent importance of political organizations, to which we now turn.

#### MAX WEBER, ROBERT MICHELS, AND THE ORGANIZATIONAL APPROACH

The second major approach is rooted in the work of Weber and Michels as well as the classic elite theorists Gaetano Mosca ([1896] 1939) and Vilfredo Pareto ([1935] 1963). Their core argument is that rule by the few (or the elite) is inevitable. This is due to the scarcity of leadership talent, the psychological need of masses for leadership, and the imperatives of complex organization. The central political question, then, is not the elimination of a ruling class or the creation of a classless society but the institutional mechanisms that regulate elite recruitment and the effectiveness of various formulas for legitimating elite rule. Thus Michels's thesis about the "iron law of oligarchy" ([1911] 1961) took as its primary target the German Social Democratic Party (SPD), contending that, despite its formal goal of promoting working-class democracy and internationalism, the SPD was in fact controlled by the permanent staff who used it to promote their own careers and positions. This proved prescient when, a few years later, the SPD rallied to the autocratic German Kaiser to support World War I against the French working class. Similarly, Weber's arguments about the "iron cage" of bureaucracy (1947) were borne out by the creation of the Soviet Union which created a highly rationalized political economy more oppressive than any in western Europe. It thus came to dominate human conduct independent of its formal

goals. In a similar vein, Mosca ([1896] 1939) and Pareto ([1935] 1963) argued that participatory or “direct” democracy is not feasible and that, at its core, modern democracy is a system for institutionalizing elite turnover by requiring that elites peacefully compete for popular support in competitive elections. Thus democracy is really a system for selecting among competing elites and thus reducing the risks of incompetent hereditary leaders and violent struggles for power. It also legitimizes the system by encouraging mass to assume that they have a say in their rulers and thus contributes to political stability.

Schumpeter (1942) developed this further into the pluralist theory of democracy, arguing that by forcing candidates to compete for popular votes in elections with multiple political parties, providing adult suffrage, and protecting rights to speech and assembly, modern representative democracy also disperses power and creates a degree of popular accountability. Electoral competition means that elites have to mobilize popular support and thus respond to public concerns. Dahl (1957, 1982) and Rose (1968) argued that this creates a countervailing power system that ensures that all significant groups have a voice and a vehicle for countering more powerful groups. In this vein, Duverger (1962) shows that electoral rules determine the number and type of political parties. Majoritarian systems with single-member districts and “winner-take-all” elections (e.g., the United States and Britain) create two-party systems with nonideological parties because voters are unwilling to throw away their ballots to support third parties. Proportional systems that allocate legislative seats based on their proportion to the popular vote create multiple ideological parties. This, in turn, creates different opportunities for political expression with majoritarian systems experiencing more protest and conflict because they are less responsive while proportional systems have greater governmental instability due to the number and contentiousness of political parties (Powell 1982).

This pluralist thesis has come under several criticisms. First, competitive elections and rules protecting freedom of association and speech do not counteract the institutional bias of politics. Thus Schattschneider (1960) argued that the interest groups and parties in the United States overwhelmingly represent the upper class. Business and the professions are well represented, while

workers, consumers, and disadvantaged groups are weakly represented (Schlozman and Tierney 1986; Berry 1997; Form 1995). Second, pluralist institutions provide no guarantee that disadvantaged groups will be able to mobilize and secure access to the system (Gamson 1975; Piven and Cloward 1979). Thus McAdam (1982) and Jenkins (1985) show that movements on behalf of African-Americans and farm workers were politically blocked by their upper-class opponents until the 1960s, when electoral realignments and increased resources for these groups allowed them to successfully mobilize. Thus political opportunities for the excluded and disorganized are historically variable and by no means institutionally guaranteed. Third is the organizational argument that oligarchy undermines the responsiveness of political parties and interest groups to their supporters (Schattschneider 1960; McConnell 1966). Thus, despite a plurality of organizations, these are internally autocratic and largely unresponsive to member interests.

Arguing that pluralists and stratification theorists share a similar “society-centered” approach, “state-centered” analysts have analyzed the role of state managers as autonomous agents of political change (Nordlinger 1981; Evans et al. 1985; Skocpol and Amenta 1986). Thus, instead of being the tools of special-interest groups (pluralism) or of the upper class and class struggles (neo-Marxism), these state managers independently develop policies that promote their ideological visions and politico-administrative interests. This ability is rooted in the growing autonomy and resources of the state, which is becoming more rationalized and more central to contemporary political economies. Bureaucracy and professionalization insulate government agencies from outside control, thus making them more independent. Because many policies are responses to problems that have previously been addressed, policy precedents have a strong positive feedback effect on the development of new policies. Thus the greater the policy development and administrative capacities of governmental bodies, the greater the ability of state managers to initiate rationalizing reforms (Finegold and Skocpol 1995; Amenta 1998).

This perspective (also called “historical institutionalism” [Skocpol and Campbell 1995]) is a useful corrective to society-centered theories. It shares with classic elite theory the idea that policy

making can be independent of societal constraint, and it develops the idea of state capacities. It has been criticized, however, for overemphasizing the autonomy of state managers and failing to specify the conditions under which political institutions are autonomous (Domhoff 1996). In response, Amenta (1998) has synthesized state-centered arguments with resource mobilization theory (Tilly 1978; McAdam 1982; Jenkins 1983) to explain how political institutions mediate the impact of social movements on policy change. Likewise, Finegold and Skocpol (1995) show how political leaders during the New Deal mobilized business support for their policies but also became dependent on business leaders. Some contend that this perspective applies better to the more rationalized stronger states in western Europe and Japan, where state managers and coordinated corporatist bargaining have been central to post-World War II economic policy (Schmitter 1981; Katzenstein 1985). The United States is a strong case for capitalist dominance, thus supporting the stratification approach, while other states display stronger state institutions and thus support the organizational approach.

#### POSSIBILITIES FOR SYNTHESIS

Since both traditions have proved useful, several have attempted to synthesize them. One approach has been treating these theories contextually, arguing that each theory bears in particular settings or with specific aspects of political change. Thus Laumann and Knoke (1987) argue that some policy arenas are organized pluralistically with multiple competing groups and considerable fluidity, while others are more centralized with a smaller number of actors and a more hierarchical relationship among them. Thus they compare the more pluralistic health care arena with the more centralized field of energy policy. Similarly, Dye (1995) argues that pluralism better explains patterns of elite recruitment while a hierarchical "power elite" model better explains decision making. The challenge for such a contextual synthesis is to develop a rationale for the conditions under which neo-Marxist, pluralist, and state-centered arguments are relevant.

A second approach is to integrate these arguments into a more comprehensive theoretical framework. Lukes (1974) and Alford and Friedland (1985) have advanced a multidimensional theory

of political power that conceptually integrates these perspectives. The starting point is Weber's classic conception of power as the ability to secure one's will against that of others even if against their resistance (1947). Thus power is always relational, involving interactions among at least two or more parties. It is also hierarchical, in the sense that one party or group is stronger or controls the other. It is also a potential or capacity. History shows that power is often based on force—whether physical or psychological coercion—but that such force is inherently unstable. No society can be organized solely on the basis of force because, at the first opportunity, people will break the rules. Terror is not only unstable but also inefficient. Thus power holders attempt to institutionalize their power; that is, they try to make it more stable by being seen by subordinates as legitimate. Authority is power that is widely perceived by subordinates as fair and just, and thus is more likely to be stable. Manipulated authority is power that power holders have attempted to legitimize by controlling information and creating false images among subordinates. Authentic authority is freely accepted by subordinates as fair and just (Habermas 1976).

Lukes (1974) distinguishes three dimensions of power: decision-making power, agenda control, and systemic power. *Decision-making power* exists when subordinates are aware of an underlying conflict of interest with power holders but are induced to act in the power holder's interest. Resources, such as wealth, charismatic claims, or the threat of violence, allow power holders to control subjects against their will. Even if compliance is not total, claimants with superior resources will typically control the behavior of subordinates. This is the aspect of power on which pluralists and state-centered analysts focus.

In *agenda control*, power holders prevail because they control what issues will be decided and what potential issues are removed from the political agenda. Such processes constitute "nondecisions" (Bachrach and Baratz 1970) in that power holders decide not to decide. Agenda control is more subtle than decision-making power, and is difficult to study but is nonetheless central to the political system. Thus, in a community power study of Gary, Indiana, Crenson (1971) found that powerful industrialists pressured local government to ignore local pollution problems created by the

steel plants. Similarly, Perrucci (1994) found that state and local officials mounted joint publicity campaigns with Japanese auto corporations that convinced local voters that the large tax breaks granted to these foreign auto companies would create jobs and tax revenues despite unequal tax burdens. Because the mass media are central to creating public issues, they are typically key. Domhoff (1998) argues that upper-class control over the mass media, the leading universities, and the major policy organizations prevents issues such as unemployment and redistributive tax reform from becoming topics of decision making. News stories are largely created by press releases and interviews with government officials, which gives government officials considerable say over what issues will be aired in the mass media. Subordinates may be aware that power is being wielded, but they cannot force their own perspectives onto the political agenda. Thus this second dimension strongly controls the decision-making process by determining what issues and views get included or ignored. Elite theorists are the key architects of this type of work.

In *systemic power*, power holders benefit by structural arrangements, such as the distribution of wealth or superior political organization. Subordinates may not be aware of their conflicting interests with power holders and may, in fact, be ideologically indoctrinated to accept the authority of power holders. This may stem from several processes. One is the inability of subordinates to mobilize and press their interests. Disorganized and resource-poor groups are often unable to mobilize (Gaventa 1985; Jenkins 1985). Second are ideologies that legitimize power. Under slavery, slave owners successfully promoted the idea that they were paternalistic "father figures" who cared for and protected their childlike slaves. Insofar as slaves accepted this ideology, they were reluctant to rebel and resisted in ways that were controllable (Genovese 1974). Neo-Marxists advance this type of analysis.

Alford and Friedland (1985) contend that different theories of power have different objects of explanation and different time scales, with systemic arguments focusing on long-term structural change, agenda setting about moderate-term institutional changes, and decision making about short-term behavior. These are hierarchically organized

with systemic power-setting limits on the agenda-setting process, which in turn sets limits on decision making. Thus studies that focus solely on first- or second-dimensional power without including the systemic context are limited and potentially flawed. Seemingly conflicting theories of political power can be synthesized once we specify their objects of explanation and integrate them into a larger theoretical framework.

Hicks and Misra (1993) advance an alternative synthetic "political resource" theory, arguing that institutional contexts define the infrastructures or facilitative conditions that enable specific groups to mobilize and secure their interests (called "instrumental resources"). Thus neo-Marxian and pluralist arguments about the political actions of specific groups need to be combined with state-centered and elite arguments about institutional capacities and policy precedents. They show that factors identified by all these theories affect the amount of social welfare spending in Western democracies. Similarly, in a study of the adoption of public venture capital firms by state governments in the United States, Leicht and Jenkins (1998) show that corporatist bargaining between big capital and labor combine with strong administrative capacities of state governments to create new entrepreneurial economic development policies. Thus, instead of specifying a causal hierarchy among theories, this approach identifies interactive combinations of factors identified by these various power theories.

At this point, it is unclear which approach will eventually prevail. The contextual approach has considerable flexibility but needs to explain the conditions under which different power theories hold. The multidimensional power argument has conceptual breadth but lacks clear empirical relevance. Unless competing theories can be tested simultaneously in empirical analyses, it is difficult to evaluate how useful a purely conceptual synthesis is. In effect, Alford and Friedland (1985) have simply demonstrated that structural, institutional, and behavioral explanations refer to different objects. The political resource theory has strong empirical promise but needs greater specification as to the interactive combinations that integrate these power theories. Single-factor theories clearly need to be surpassed so that a broader theory of power and the state can be developed that will account for a variety of contexts as well as different

power processes within an inclusive theoretical framework.

## REFERENCES

- Alford, Robert, and Roger Friedland 1985 *The Powers of Theory*. New York: Cambridge University Press.
- Amenta, Edwin 1998 *Bold Relief*. Princeton, N.J.: Princeton University Press.
- Bachrach, Peter, and Morton Baratz 1970 *Power and Poverty*. New York: Oxford University Press.
- Berry, Jeffrey 1997 *The Interest Group Society*, 3rd ed. New York: Longman.
- Block, Fred 1987 *Revising State Theory*. Philadelphia, Pa.: Temple University Press.
- Carnoy, Martin 1984 *The State and Political Theory*. Princeton, N.J.: Princeton University Press.
- Clawson, Dan, Alan Neustadt, and Denise Scott 1992 *Money Talks*. New York: Basic.
- Crenson, Matthew 1971 *The Un-Politics of Air Pollution*. Baltimore, Md.: Johns Hopkins University Press.
- Dahl, Robert 1957 *A Preface to Democratic Theory*. New Haven, Conn.: Yale University Press.
- 1982 *Dilemmas of Pluralist Democracy*. New Haven, Conn.: Yale University Press.
- Domhoff, G. William 1979 *The Powers That Be*. New York: Random House.
- 1990 *The Power Elite and the State*. New York: Aldine DeGruyter.
- 1996 *State Autonomy or Class Dominance?* New York: Aldine DeGruyter.
- 1998 *Who Rules America?* 3rd ed. Mountain View, Calif.: Mayfield.
- Dye, Thomas 1995 *Who's Running America?* 6th ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Duverger, Maurice 1962 *Political Parties*, 2nd ed. New York: John Wiley.
- Evans, Peter, Dietrich Rueschemeyer, and Theda Skocpol (eds.) 1985 *Bringing the State Back In*. New York: Cambridge University Press.
- Finegold, Kenneth, and Theda Skocpol 1995 *State and Party in America's New Deal*. Madison: University of Wisconsin Press.
- Form, William 1996 *Segmented Labor, Fractured Politics*. New York: Plenum.
- Gamson, William 1975 *The Strategy of Social Protest*. Homewood, Ill.: Dorsey.
- Gaventa, John 1985 *Power and Powerlessness*. Urbana: University of Illinois Press.
- Genovese, Eugene G. 1974 *Roll, Jordan, Roll! The World the Slaves Made*. New York: Vintage.
- Gramsci, Antonio 1957 *The Modern Prince and Other Writings*. New York: International.
- Habermas, Jurgen 1976 *Legitimation Crisis*. Boston: Beacon.
- Hicks, Alexander, and Joya Mishra 1993 "Political Resources and the Growth of Welfare in Affluent Democracies, 1960-1982." *American Journal of Sociology* 99:668-710.
- Jenkins, J. Craig 1983 "Resource Mobilization Theory and the Study of Social Movements." *Annual Review of Sociology* 9:527-553.
- 1985 *The Politics of Insurgency*. New York: Columbia University Press.
- Katzenstein, Peter 1985 *Small States in World Markets*. Ithaca, N.Y.: Cornell University Press.
- Laumann, Edward, and David Knoke 1987 *The Organizational State*. Madison: University of Wisconsin Press.
- Leicht, Kevin, and J. Craig Jenkins 1998 "Political Resources and Direct State Intervention." *Social Forces* 76:1323-1345.
- Lipset, Seymour Martin 1981 *Political Man*, 2nd ed. Baltimore, Md.: Johns Hopkins University Press.
- Lukes, Steven 1974 *Power*. London: Macmillan.
- McAdam, Doug 1982 *Political Process and the Development of Black Insurgency*. Chicago: University of Chicago Press.
- McConnell, Grant 1966 *Political Power and American Democracy*. New York: Vintage.
- Marx, Karl (1852) 1964 *The Eighteenth Brumaire of Louis Bonaparte*. New York: International.
- Michels, Robert (1911) 1962 *Political Parties*. Chicago: Dover.
- Mills, C. Wright 1956 *The Power Elite*. New York: Oxford University Press.
- Mosca, Gaetano (1896) 1939 *The Ruling Class*. New York: McGraw-Hill.
- Neumann, Franz 1942 *Behemoth*. New York: Octagon.
- Nordlinger, Eric A. 1981 *On the Autonomy of the Democratic State*. Cambridge, Mass.: Harvard University Press.
- Offe, Claus 1984 *Contradictions of the Welfare State*. London: Hutchinson.
- Pareto, Vilfredo (1935) 1963 *Mind and Society*. New York: Dover.
- Parkin, Frank 1979 *Marxism and Class Theory*. London: Tavistock.
- Perrucci, Robert 1994 *Japanese Auto Transplants in the Heartland*. New York: Aldine DeGruyter.

- Piven, Frances, and Richard Cloward 1979 *Poor Peoples Movements*. New York: Pantheon.
- Powell, G. Bingham 1982 *Contemporary Democracies*. Cambridge, Mass.: Harvard University Press.
- Poulantzas, Nicos 1973 *Political Power and Social Classes*. London: New Left.
- Rose, Arnold 1968 *The Power Structure*. New York: Oxford University Press.
- Schattschneider, E. E. 1960 *The Semi-Sovereign People*. New York: Holt, Reinhart and Winston.
- Schlozman, Kay Lehman, and John T. Tierney 1986 *Organized Interests in American Democracy*. New York: Harper and Row.
- Schmitter, Phillipe 1981 "Interest Intermediation and Regime Governability in Contemporary Western Europe and North America." In Suzanne Berger, ed., *Organizing Interests in Western Europe*. New York: Cambridge University Press.
- Schumpeter, Joseph 1942 *Capitalism, Socialism and Democracy*. New York: Harper and Row.
- Skocpol, Theda 1981 "Political Response to Capitalist Crisis." *Politics and Society* 10:155–201.
- , and Edwin Amenta 1986 "States and Social Policies." *Annual Review of Sociology*. vol. 131–157.
- Skocpol, Theda, and John L. Campbell (eds.) 1995 *American Society and Politics*. New York: McGraw-Hill.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Pa.: Addison-Wesley.
- Useem, Michael 1984 *The Inner Circle*. New York: Oxford University Press.
- Weber, Max 1947 *From Max Weber*. Hans Gerth and C. Wright Mills, eds. New York: Oxford University Press.

J. CRAIG JENKINS

## POPULAR CULTURE

Since the 1960s, studies of popular culture in the United States have proliferated and a range of novel arguments have been proposed, linking patterns of popular culture production and consumption to systems of stratification and power. Before the 1960s in Europe, Roland Barthes ([1957] 1972) and Fernand Braudel ([1949] 1966) championed (for quite different reasons) increased attention to everyday culture and its social significance, and members of the Frankfurt school emigrating to the United States brought new theories of mass

culture to American academics (Rosenberg and White 1957; Lowenthal 1961), but American scholars still did not generally see any value in studying popular culture.

Beginning in the mid-1960s, as the American middle class began to be targeted by the mass media as the desired audience, more American educators started to show more interest in media-based popular culture, even though in much of academia, studying popular culture was either declassé or taboo (Ross 1989). A few hardy souls from sociology and literary criticism looked at popular culture as a realm of interesting fads and fashions, ephemeral cultural forms that plummeted though modern urban life with regularity, gave rise to much cultural entrepreneurship, and left ordinary citizens running to keep up with what was "happening." Sociologists found it a bit easier to justify ongoing attention to these social ephemera because of the established tradition in sociology of examining urban and suburban communities and their cultures (Park 1955; Lynd and Lynd 1929). By the mid-1960s a quite active community of scholars around Bowling Green University proliferated empirical and descriptive accounts of everything from fast-food restaurants to rock and roll (Keil 1966; Nye 1972; Cawelti 1972; Browne 1982). At roughly the same time, a small group of literary scholars drew on longstanding literary interest in the voices of the people in literature (Shiach 1989, chap. 2, 4), and argued that to understand contemporary uses of language, one had to study commercial language in popular culture (McQuade and Atwan 1974). This work did not have much success in changing either sociology or literature. In sociology, it was eclipsed conceptually by sociological work that linked patterns of popular culture to systems of institutional control (Cantor 1971; Denisoff 1974; Hirsh 1972). This work had greater legitimacy because it addressed the organizations' literature, but it also reinforced the sense that the study of popular culture was not really important enough to stand on its own.

By the end of the 1960s, as the political climate shifted, radical scholars began to champion studies of popular culture either to understand the world of "the people" (disregarded by elites) or to account for the political passivity of the working class and poor. They tried to resuscitate questions about elite distaste for popular culture itself and its relation to systems of social control. These

questions gave popular culture new importance, not as an aesthetic or commercial system but as a political actor in systems of stratification and power (Schiller 1969; Guback 1969; Aronowitz 1973; Gans 1974).

This legacy has been carried into present-day popular culture research as it has spread through sociology, literature, anthropology, history, and cultural studies. Ongoing fascination with “politics from below” has made this subfield a conceptually complex and politically “left” branch of cultural studies, not concerned so much with the moral fabric of society or the ideational sources of its integration (subjects derived from the Weberian tradition of cultural studies), but rather with the use of culture to exert or avoid systematic domination from above.

Many contemporary attempts to explain patterns of cultural domination through popular culture are indebted to (and in different ways critical of) the work on mass culture and consciousness begun by the Frankfurt school. Members of the Frankfurt Institute of Social Research originally organized themselves to examine the philosophical underpinnings of Marxism, but when Hitler came to power, since most of the leading members of the group were Jewish, this project was disrupted and many figures came to the United States. The work on mass culture that developed from this group was (not surprisingly) devoted to understanding the success of Nazism by dissecting and analyzing the psychological and political effects of mass society (Jay 1973). Members of the Frankfurt school perceived mass culture as aesthetically and politically debilitating, reducing the capacities of audiences to think critically and functioning as an ideological tool to manipulate the political sentiments of the mass public. They argued that in modern industrial societies, the pursuit of economic and scientific rationality provided an impoverished environment for human life. The realm of culture, which might have provided respite from the drudgery of everyday life, was itself being industrialized, yielding a commercial mass culture that atomized audiences and lulled them with emotionally unsatisfactory conventionality. This world of commodities only added to the dissatisfactions that deflected people from their desires. The dulling of their senses made them politically passive, and their emotional discontent made them easy targets for propaganda that addressed their

powerful inner feelings. This combination, according to theory, was what made the propaganda in Nazi Germany so effective (Horkheimer and Adorno 1972).

During the 1960s, critical theory, as the work of the Frankfurt school came to be known, continued in U.S. intellectual circles to be used to explain the political conservatism of the working class, but it was also taken up in the student movement as a critique of commercial mass culture that justified the efforts by “flower children” to create radical social change through cultural experimentation. The problem was that, for the latter purpose, critical theory was too deterministic to have much room for human agency, including cultural strategies for change. Constructivist models from the sociology of culture could be and were used to explain how ordinary people could break the hold of political institutions over their imaginations (Blumer 1969; Goffman 1959; Berger and Luckmann 1966; Schutz 1967; Becker 1963), but they did not explain how ideological control of populations by elites could work. The insights of the Italian communist political writer Antonio Gramsci (1971) about hegemony seemed a better scheme for explaining both the role of ideology in systems of power and the constructed nature of social reality. According to Gramsci, elites maintained their power and legitimacy by creating hegemonic definitions of reality that were accepted as common sense by the population. By subscribing to these views, nonelites collaborated in their own oppression. Gramsci’s work, available in English translations and popularized in the academic community in the 1970s, gave the study of culture and power in the English-speaking world new direction.

By the 1970s, much innovative work on popular culture was coming out of Great Britain. The British school of cultural studies drew attention to the role of nonelites in systems of power, but it focused more on working-class culture—particularly its role as a crucible for cultural resistance, innovation, and change. This school had its roots in the work of E. P. Thompson (1963) and Raymond Williams (1961; 1977; 1980). These authors began from the premise that the working class is not defined just by relations of production, but also by a self-consciousness bred in a class-based way of life. The working class has its own history and traditions that give its members distinct values and a complex relation to societal level systems of



power. In their own cultural enclaves, class members are active producers of their own political institutions and popular entertainments (and through them defined social realities). So while the public culture of Western industrialized societies may be dominated by elites who control the mass media and who try to use cultural systems for exerting power, their hegemonic control is circumscribed by the cultures of subordinated groups. The realm of popular culture, in this tradition, is an arena of conflict in which cultural identity, authority, and autonomy are contested. Social rifts are made manifest in the multiplicity of points of view that enter into the public sphere along with the hegemonic messages of much mass culture (Curran, Gurevitch, and Woolcott 1979; Hall and Whannel 1965).

While early British cultural studies paid greatest attention to working-class culture, the ideas about cultural resistance were easily transferred to the analysis of other subordinated groups such as women, youth, and minorities. This broader approach to cultures of resistance gave birth to the kind of subcultural analysis conducted, for example, by Dick Hebdige (1979). He argues that innovations in youth culture come from marginalized working-class youths rebelling against both their parents and hegemonic culture. New developments in music and dress are culled from the cultural possibilities made available in mass society, both in commercial commodities and local cultures. These cultural resources are mixed and reassembled to create new subcultural styles. Much innovation of this sort comes from minority communities and is picked up by middle-class kids in part because it is so offensive to their parents. The irony, of course, is that if they make these styles popular, they end up making them part of the world of mass culture, economic mainstays of the entertainment industry.

One of the most interesting literatures spawned in America by this British school comes from historians looking to the realm of popular culture to try to understand class, gender, and ethnic relations in the United States. Roy Rosenzweig (1983), Kathy Peiss (1985), and George Lipsitz (1990) look at how class, gender, and ethnic culture are sustained and dissolved over time in patterns of resistance, co-optation, mutual influence, and change. They identify ways that residues of older cultural traditions both resist and are incorporated into the cultural mainstream, and

ways that different groups have absorbed and used elements of both traditional and mass culture to fashion distinct identities and ways of life.

Rosenzweig (1983), studying the white working class in nineteenth-century America, treats popular culture as a site of resistance to work discipline in the factory. The division of life into periods of work and leisure for workers in this period was not, to Rosenzweig, the articulation of two spheres of activity, but a political division that was part of a struggle over control of time by workers.

Peiss (1985) looks at women workers in nineteenth-century cities. She demonstrates that young working women used their new economic independence to resist the constraints of the family as well as of the factory. They were able to develop new styles of dress, dancing, and play, but could not free themselves from their highly structured gender relations.

Lipsitz (1990) looks at how ethnic and class cultures have been sustained and dissolved in the late twentieth century in the United States. He sees popular culture forming a kind of popular memory, obscuring and yet reviving the U.S. working class's immigrant past and ethnic complexity. Centralized mass media such as television have helped to create and record the decline of immigrant identity under the force of consumerism. In contrast, more participatory cultural forms like street dancing and parading during Mardi Gras and some popular music forms have allowed ethnic groups to play their identities and create an urban mixed culture that simultaneously embraces and rejects traditional ethnic identity.

Another direction in the analysis of class and culture has been developed by Pierre Bourdieu (1984) and his colleagues in France. They have been looking for the mechanisms by which domination has been sustained across generations. If social constructivists are right and social life must necessarily be "created" by each new generation, then social *stability* over time needs theoretical explanation. To Bourdieu, culture is a main source of class stability. He argues that each rank has its own kind of cultural tastes, some systems of taste constituting cultural capital that can be exchanged for economic capital. People at the top of the hierarchy have a class culture with a high amount of cultural capital. They teach this culture to their

children and thereby give them an economic edge. This kind of elite culture is also taught in school, but kids from less affluent backgrounds often resist learning it. This cultural resistance by the working class is not a victory, according to Bourdieu; rather, it is a trap for reproducing the class system.

Bourdieu's theory of cultural and social stratification is interestingly unlike most models found in the United States and Britain because it has no special place for a homogenizing mass culture. Bourdieu argues that members of different social ranks may see the same films (or other forms of mass culture), but they see them in different ways and they like or dislike them for different reasons. Elite culture is more abstract and formal than working-class culture, so elite filmgoers pay more attention to film language while nonelites care more about plots and stars. These differences in cultural consumption are more significant to Bourdieu than differences of cultural production (mass versus handmade culture) because elites identify with formal approaches to culture and prefer to associate with (and hire) those who share their views.

Scholars in both Britain and the United States have been profoundly influenced by Bourdieu. Paul DiMaggio (1982), in a study of the Boston Symphony and its development in the nineteenth century, paid attention to the differentiation of tastes and social ranks at issue when concerts for elite audiences were purged of popular songs and were used to define a special repertoire of classical music. This happened when the symphony was established as an elite musical institution and drove out competing musical organizations that had more "democratic" tastes. DiMaggio argues that this cultural differentiation took place when immigrant groups grew dramatically in Boston and took over local politics there. The superiority of traditional elites was no longer visible in the public sphere, but it remained central to the economy. The creation of cultural institutions identifying this elite with elevated tastes helped to make class power visible and to sustain it over time by giving upper-class Bostonians a distinctive culture.

In Britain, Paul Willis (1977) has confirmed Bourdieu's perceptions about class reproduction through his study of the education of working-class youths. He argues that distaste for the "elevated" values of the school among working-class

youths is expressed in school by resistance to lessons. This resistance does not have the optimistic possibilities found in the theories of Williams (1977) or Hebdige (1979), but results in class reproduction. Working-class youths, in eschewing elite cultural values, end up reproducing their own domination within the class system. MacLeod (1987) in the United States finds much the same thing, although he focuses on differences between blacks and whites. Members of gangs from both ethnic communities who lived in the same housing project found difficulty escaping their social rank because of difficulties at school. The blacks believed that by going to school they could achieve mobility, while the white kids did not. Still, both groups were kept in their "places" by a lack of cultural capital.

Since the end of the 1970s, there has been a growing literature, stimulated by the women's movement, on gender stratification and popular culture. The bulk of it addresses two media—novels and film—because of the centrality of women to the economic development of these two areas of popular entertainment. As Ann Douglas (1977) pointed out in her seminal and controversial book, *Feminization of American Culture*, women readers and women writers helped to establish this form of popular novel writing in the United States during the nineteenth century. Sentimental novels were tailored to the domesticated women in the period, who had to stay home and devote their attention to familial relations and child rearing. The novels focused on interpersonal relations and problems of individuals' morals (as opposed to large issues of morality)—just the kind of thing that both fit and justified middle-class women's highly circumscribed lives. Douglas decries the role of women writers in shaping this disempowering literature for women and praises in contrast more "masculine" and male writings from the period (hence generating much controversy about her easy acceptance of the literary canon). Most important for students of popular culture, she argues that the sentimental novels were models of mass culture that have been used ever since in romance novels and soap operas.

Janice Radway (1984) questioned this easy dismissal of romance novels, and went out to study in a quasi-ethnographic fashion the readers of contemporary romance novels to see how they were affected by their reading. She found that the

novels had more mixed effects than Douglas supposed. While they taught traditional gender relations (including male violence toward women), they also celebrated the gentler side of men and (more important) were used by women readers as a reason to deflect demands on their time by husbands and children. Women claimed their reading time as their own, and used it to withdraw temporarily from the uninterrupted flow of demands on their attention.

Gaye Tuchman's book (1989) provides some interesting history that serves as a vantage point from which to view the controversy between Douglas and Radway. She shows that, around the turn of the century, publishing houses began to reject the women novelists and their sentimental novels and to favor male novelists. Publishers were central to the switch to the canons of modernism, and the "expulsion" of women from the profession of novel writing. Women readers still constituted the major market for novels, but their market had become so lucrative that high-status male poets, who had eschewed the novel before, began to be interested in it. Once this occurred, their tastes were taken as authoritative and the novel was quickly placed in their hands. Tuchman makes clear that changes in taste like this were neither arbitrary nor grounded purely in aesthetics; they were the result of economic changes in the literary market, institutional decisions about how to address them, and institutional trivialization of women and their culture.

The attention to gender and film has been inspired not by the importance of the female audience or the centrality of women to the film industry (the opposite is the case), but rather by the importance of actresses, of the faces and bodies of film stars, to the commercial and cultural success of the industry. When feminist studies of film began in the 1970s, most of the work was on the exploitation of the female body in films by male filmmakers and for a male audience. This kind of analysis stressed how commercial films used male-centered notions of sexuality and power, presenting women in films as objects of desire and/or violence (Weibel 1977; Tuchman et al. 1978). In the 1980s, researchers turned away from the study of film production and toward analyses of film language and film consumption to construct a psychology of film watching (Modleski

1982; Mulvey 1989). Much of this literature focuses on the voyeuristic pleasure film watching provides men by allowing them to gaze at women's bodies while sitting in a dark theater where the female objects of the gaze cannot look back. Scholars in this tradition examine in shot-by-shot detail how men and women are differentially presented on film: men are generally in medium shots, carrying the action of films, while women stand in the background (or are dissected in close-ups to appear as faces or other body parts, available to the male gaze).

Because this type of analysis seemed to prove so decisively that films are constructed for a male audience, feminists wondered why women seem to find so much pleasure in going to the movies. One answer is that some films contain strong and interesting female characters who address issues of concern to female audiences. Another is that interpretations of films are not so controlled by authorial intentions, and are much more a matter of audiences' active readings of messages. Drawing on Nancy Chodorow's ideas about female psychology (1978), Carol Gilligan's ideas about female reasoning (1982), Lacanian psychology, and poststructuralist views of the politics of interpretation (Eagleton 1983), a psychology of film emerged around how audiences (particularly women) construct meaning from film texts (Mulvey 1985, 1989; Erens 1979). The most recent works of media analysis have rejected altogether such a dualistic approach to gender. Learning from studies of sexuality itself, they have considered how gender categories are at stake in popular culture. They examine how gender dualism is enforced and contested in the mass media (Gamson 1998).

In the 1980s, two opposite developments in culture theory have emerged from renewed attention (in poststructuralism in general and in the film theory described above) to the multivocality of texts and the proliferation of meanings through multiple readings. The upbeat one emphasizes the liberatory nature of culture, and is related to: (1) the poststructuralist argument that asserting alternative interpretations undermines the authority of canonical readings; (2) feminist versions of reader response theory that contend that how you use culture is central to what it is; and (3) the idea from the British school of cultural studies that competing social voices enter into the public sphere and are available for readers or audiences to find.

Advocates of this position claim that efforts at social control through culture do not work very well because, in their own life worlds, people use the cultural resources around them in their own ways. These new constructivists—for example, Robert Bellah (Bellah et al. 1985) Ann Swidler (1986), Joseph Gusfield (1989), and Michael Schudson (1989)—are much like Goffman (1959) and earlier symbolic interactionists who presented everyday life as a cultural achievement, but they see the construction of meaning in everyday life (in an optimistic reversal of Foucault and other poststructuralists) as a healthy exercise of power as well as symbolic manipulation (Foucault 1970, 1975, 1979; Jameson 1984; Zukin 1988).

This optimistic view of the proliferation of meanings in everyday life is countered by students of postmodernism who derive from structuralism and poststructuralism an interest in the languages of culture and see in modern urban society a loss of meaning resulting from the multiplication of signs and their decontextualization or reappropriation. They argue that commercial culture has such a need to assign meaning to objects (in order to make sense of their consumption and use) that signs are proliferated, reappropriated, mixed, and reused until they lose their meaning. For example, as famous paintings are used to sell cosmetics, images of the Old West are used to signify the solidity of banks, and bits and pieces of past architecture are mixed to construct a new built environment, history is made meaningless. The play with signs goes on without serious thought to what this does to human life. The result is (to postmodernists) a politically debilitating alienation. Cultural production and counterproduction, the argument goes, may reduce hegemony by undermining attempts to define “common sense,” and may give people pleasure through the free play of signs, but they provide only an illusion of freedom, and breed a loss of meaning in life. This view of modern urban life contains some of the unremitting pessimism of the Frankfurt school, but it is tied to a view of cultural decentralization that is at odds with traditional critical theory.

The diverse approaches to popular culture that have developed since the 1960s seem to have produced a proliferation of meanings for popular culture itself, but the result has not been alienation. Popular culture research has gained an analytic richness that it lacked when few scholars

dared or cared to approach it. Conflicting theoretical views about what makes popular culture significant may make it more difficult to define and characterize (much less understand) the field. But all the debates consider how groups come to understand the world they live in, and how those understandings subordinate or alienate them (on the one hand) or liberate them to make meaningful lives, in spite of efforts by others to control them (Long 1997). This heritage is clear, and gives both meaning and direction to popular culture studies.

(SEE ALSO: *Mass Media Research*; *Postmodernism*; *Social Movements*)

## REFERENCES

- Aronowitz, Stanley 1973 *False Promises*. New York: McGraw-Hill.
- Barthes, Roland (1957) 1972 *Mythologies*. New York: Hill and Wang.
- Becker, Howard 1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Bellah, Robert N., et al. 1985 *Habits of the Heart*. Berkeley: University of California Press.
- Berger, Peter, and T. Luckmann 1966 *Social Construction of Reality*. Garden City, N.Y.: Doubleday.
- Blumer, Herbert 1969 *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.
- Bourdieu, Pierre 1984 *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, Mass.: Harvard University Press.
- Braudel, Fernand (1949) 1966 *The Mediterranean and the Mediterranean World in the Age of Phillip II*. New York: Harper.
- Browne, Ray 1982 *Objects of Special Devotion*. Bowling Green, Ohio: Popular Press.
- Cantor, Muriel 1971 *The Hollywood Producer*. New York: Basic.
- Cawelti, J. G. 1972 *The Six-Gun Mystique*. Bowling Green, Ohio: Popular Press.
- Chodorow, Nancy 1978 *Reproduction of Mothering*. Berkeley: University of California Press.
- Curran, James, Michael Gurevitch, and Janet Woolcott 1979 *Mass Communication and Society*. New York: Russell Sage Foundation.

- Denisoff, Serge 1974 *Solid Gold: The Popular Record Industry*. New Brunswick, N.J.: Transaction Books.
- DiMaggio, Paul 1982 "Cultural Entrepreneurship in Nineteenth-Century Boston: The Creation of an Organizational Base for High Culture in America." *Media, Culture and Society* 4:33-50.
- Douglas, Ann 1977 *Feminization of American Culture*. New York: Knopf.
- Eagleton, Terry 1983 *Literary Theory*. Minneapolis: University of Minnesota Press.
- Erens, P. 1979 *Sexual Stratagems*. New York: Horizon.
- Foucault, Michel 1970 *The Order of Things*. New York: Random House.
- 1975 "What Is an Author?" *Partisan Review* 4:603-614.
- 1979 *Discipline and Punish*. New York: Vintage.
- Gamson, Joshua 1998 *Freaks Talk Back*. Chicago: University of Chicago Press.
- Gans, Herbert 1974 *Popular Culture and High Culture*. New York: Basic.
- Gilligan, Carol 1982 *In a Different Voice*. Cambridge, Mass.: Harvard University Press.
- Goffman, Erving 1959 *Presentation of Self in Everyday Life*. Garden City, N.Y.: Anchor.
- Gramsci, Antonio 1971 *Selections from the Prison Notebooks*. New York: International.
- Guback, Thomas 1969 *The International Film Industry: Western Europe and America since 1945*. Bloomington: Indiana University Press.
- Gusfield, Joseph 1989 *On Symbols and Society*. Chicago: University of Chicago Press.
- Hall, Stuart, and Paddy Whannel 1965 *The Popular Arts*. New York: Pantheon.
- Hartounic, Valerie 1997 *Cultural Conceptions*. Minneapolis: University of Minnesota Press.
- Hebdige, Dick 1979 *Subculture: The Meaning of Style*. New York: Methuen.
- Hirsh, Paul 1972 "Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems." *American Journal of Sociology* 77:639-659.
- Horkheimer, Max, and Theodor Adorno (1944) 1972 *Dialectic of Enlightenment*. New York: Seabury.
- Jameson, Frederic 1984 "Postmodernism, of the Cultural Logic of Late Capitalism." *New Left Review* 146:53-93.
- Jay, Martin 1973 *The Dialectical Imagination: A History of the Frankfurt School and the Institute of Social Research 1923-50*. Boston: Little, Brown.
- Keil, C. 1966 *Urban Blues*. Chicago: University of Chicago Press.
- Lipsitz, George 1990 *Time Passages*. Minneapolis: University of Minnesota Press.
- Long, Elizabeth 1997 *From Sociology to Cultural Studies*. Malden, Mass: Blackwell.
- Lowenthal, Leo 1961 *Literature, Popular Culture and Society*. Englewood Cliffs, N.J.: Prentice-Hall.
- Lynd, Robert, and H. Lynd 1929 *Middletown*. New York: Harcourt, Brace, and World.
- MacLeod, Jay 1987 *Ain't No Making It*. Boulder, Colo.: Westview.
- McQuade, Donald, and Robert Atwan 1974 *Popular Writing in America*. New York: Oxford University Press.
- Modleski, Tania 1982 *Loving with a Vengeance*. Hamden, Conn.: Archon.
- Mulvey, Laura 1985 "Film and Visual Pleasure." In G. Mast and M. Cohen, eds., *Film Theory and Criticism*. New York: Oxford University Press.
- 1989 *Visual and Other Pleasures*. Bloomington: Indiana University Press.
- Nye, R. B. 1972 *New Dimensions in Popular Culture*. Bowling Green, Ohio: Popular Press.
- Park, Robert 1955 *Society*. New York: Free Press.
- Peiss, Kathy 1985 *Cheap Amusements: Working Women and Leisure in New York City, 1880 to 1920*. Philadelphia: Temple University Press.
- Radway, Janice 1984 *Reading the Romance*. Chapel Hill: University of North Carolina Press.
- Rosenberg, Bernard, and David Manning White 1957 *Mass Culture*. New York: Free Press.
- Rosenzweig, Roy 1983 *Eight Hours for What We Will: Workers and Leisure in an Industrial City, 1870-1920*. New York: Cambridge University Press.
- Ross, Andrew 1989 *No Respect: Intellectuals and Popular Culture*. New York: Routledge and Kegan Paul.
- Schiller, Herbert 1969 *Mass Communication and American Empire*. New York: Kelly.
- Schudson, Michael 1989 "How Culture Works." *Theory and Society* 18:153-180.
- Schutz, Alfred 1967 *Collected Papers*, vol. 3, *The Problem of Social Reality*. The Hague: Martinus Nijhoff.
- Shiach, Morag 1989 *Discourse on Popular Culture*. Stanford, Calif.: Stanford University Press.
- Swidler, Ann 1986 "Culture in Action: Symbols and Strategies." *American Sociological Review* 51:273-286.

- Thompson, E. P. 1963 *The Making of the English Working Class*. New York: Vintage.
- Tuchman, Gaye 1989 *Edging Women Out*. New Haven, Conn.: Yale University Press.
- , Arlene Daniels, and James Benet 1978 *Hearth and Home*. New York: Oxford University Press.
- Weibel, Kathryn 1977 *Mirror, Mirror*. Garden City, N.Y.: Anchor.
- Williams, Raymond 1961 *The Long Revolution*. New York: Columbia University Press.
- 1977 *Marxism and Literature*. Oxford: Oxford University Press.
- 1980 *Problems in Materialism and Culture*. London: Verso.
- Willis, Paul 1977 *Learning to Labor*. New York: Columbia University Press.
- Zukin, Sharon 1988 "The Post-Modern Debate over Urban Form." *Theory, Culture and Society* 5:431–446.

CHANDRA MUKERJI

## POPULATION

As a topic, *population* refers to the size and composition of social groupings and the dynamics of change in these characteristics. It extends to the determinants and consequences of levels, differentials, and changes in fertility, mortality, and spatial distribution. Within the social sciences, population study has a broad concern with the cultural, social, economic, and psychological causes and consequences of these characteristics (see Hauser and Duncan 1959; Coleman and Schofield 1986; Stycos 1987; Namboodiri 1988). Within sociology, the main concern is with linkages between social institutions and the dynamics of population change and equilibrium (see Taeuber, Bumpass, and Sweet 1978; Nam 1994; Greenhalgh 1996). *Demography* is an important component of population study that focuses on data collection, measurement, and description. (Though it would be impossible to cite all articles relevant to the general topic of population, a number of useful articles are cross-referenced at the end of this article.)

### OVERVIEW OF POPULATION HISTORY

A brief overview of major changes in the characteristics of human populations will suggest some of

the linkages between social life and these characteristics. (For a more detailed discussion see, for example, Wrigley 1969; Petersen 1975.)

During Neolithic times, hunting and gathering groups probably consisted of only a few dozen individuals. Based on the technology and the area believed to have been populated, it is estimated that in about 8000–6000 B.C. the population of the world was only five million to ten million (Yaukey 1985, p. 38). Probably about half of all children died before age 5. Mainly because of the high mortality of infants and children, life expectancy was probably only about 20 years, although some adults would have survived to advanced ages. Maternal mortality was also high, probably resulting in considerably shorter life expectancy for females than for males (United Nations 1973, p. 115). In such a setting, the social consequences of these fundamental demographic facts would have been enormous. Many of our current social values and institutions have their roots in this harsh environment.

Death was sudden, random, and frequent—at least five times as common, in a population of a given size, as in a developed country today. A major function of religion was to enable people to interpret death as a part of the life course and to surround it with rituals. New births to offset these deaths were essential for the survival of a group. There can be no doubt that many groups, of various sizes, failed in this effort, but the survivors were our ancestors. Children were important to the economy of the household and community from an early age, and young adults were crucial to the welfare of older adults. Practices and institutions to encourage fertility—of humans, as well as of the environment of plants and animals—were essential. Marital unions of some type would have been virtually universal and would have begun at an early age. It is likely that unions would have been arranged by the parents, partly because they occurred at such early ages, as well as because of the social advantages of arranged marriages for strengthening an intergenerational network of social obligations. Thus the social institutions of religion, marriage, and the extended family, among others, had some of their original impetus in the extremely high mortality of human prehistory and the consequent imperative for high fertility and child survivorship.

One of the most remarkable functions of social institutions is that they provide valves or mechanisms by which population size can be regulated to be compatible with the prevailing environmental circumstances and the level of technology (Davis 1963). As just mentioned, high mortality must be accompanied by comparably high fertility if a population is to sustain itself. For the most part, it is the level of mortality that drives the level of fertility, through various intervening mechanisms, although mortality (infanticide) has sometimes served as a means of population control. The social regulation of population size usually takes the form of increasing or decreasing the average number of births per woman, to compensate for uncontrollable influences on mortality.

Humans have a much greater capacity to reproduce than is often recognized, and even in situations of very high mortality and fertility there is usually an untapped potential for even higher fertility. There are well-documented "natural fertility" populations in which the *average* woman who survives the childbearing ages has 10 to 11 children after age 20, and it is estimated that the average after age 15 could be as high as 16 children under some circumstances. (Of course, if maternal mortality is high, many women will not survive the childbearing years.) The potential supply of births is adequate to balance even the most severe conditions; if it were not, the species would not have survived.

Over the long run, there has been a pattern of increased life expectancy with two major transitions. The transition from hunting and gathering to settled agriculture and larger human settlements produced a net increase in life expectancy, although with some shifts in causes of death. Larger settlements have a higher incidence of infectious diseases because of inadequate sanitation and sources of clean drinking water. The second major transition began in the seventeenth century with industrialization and the progressive reduction of deaths from infectious diseases. Fluctuations in mortality due to transitory influences have been superimposed on these two main transitions (Tilly 1978; Wrigley and Schofield 1981). Some of the short-term increases in mortality, due to wars, famines, and epidemics, have been devastating. For example, the Black Death in fourteenth-century Europe eliminated more than one-third of the population in several areas.

During the transition to agriculture and larger settlements, mortality fell and population size increased. It is believed that, starting in roughly 5000 B.C., world population approximately doubled every 1,000 years, reaching 100 million around 500 B.C. It doubled again in the next 1,000 years, reaching 200 million around 500 A.D., and again in the next 1,000 years, reaching 425 million in 1500 A.D. (McEvedy and Jones 1978). Most of the growth was in Europe and the Middle East. This is obviously an extremely low rate of growth by modern standards, around 0.07 percent per year, but it was a marked increase from the hunting and gathering era.

The increase in carrying capacity with a new technology and social organization probably had its principal effect on reproduction at the household level (Laslett and Wall 1972; Goody 1976, 1983). If territory for new settlements is absent or inaccessible, having too many children will lead to excessive division of land and property. Some kind of limitation on reproduction will result.

The most important social mechanism or lever for regulating fertility has probably been limitation of exposure to the risk of conception—in short, regulation of sexual activity. Thus, in preindustrial Europe, the age at marriage was high—on average, in the mid-twenties. The motivation for delaying marriage and the formation of new households probably arose because the parental generation had limited land to pass to their children. Marriage and childbearing were deferred until a viable household could be established. These household-level motives led to a general consensus that marriage at later ages was preferable. Associated with late age at marriage were voluntary rather than arranged marriages, and apprenticeships or domestic service for many young people. Alternatives to married life developed—for example, celibate religious communities. These behaviors can be viewed as mechanisms for fertility limitation, even though they certainly had more direct functions as well. Within marriage there was probably very little use, or even knowledge, of contraception.

When short-term increases in mortality occurred, as with a famine, war, or epidemic, the social response was an increase in the prevalence of marriage and/or a reduction in the mean age at

marriage. Again, these motivations operated primarily at the household level, in the sense that when mortality rose, there were increased opportunities for land division and settlement, and new households could be formed more quickly.

The attempt here is to characterize population growth and homeostasis in the broadest terms, up to the beginning of the industrial era in the West. Because of limited space, this description glosses over enormous differences worldwide in the patterns of reproduction and social structure and their linkages. There have several ethnographic and historical analyses of these variations (see, e.g., Hanley and Wolf 1985).

The growth of population that accompanied industrialization is indicated by the following estimates. World population was about 545 million in 1600; 610 million in 1700; 900 million in 1800; 1,625 million in 1900. It is expected to be about 6,100 million in 2000 (McEvedy and Jones 1978; United Nations 1998). The period saw an acceleration in the rate of growth, as well as enormous increases in sheer numbers of people. Prior to 1900, Europe increased most rapidly. In other parts of the world, most of the increase was concentrated in the twentieth century. This period of rapid growth is described as the *demographic transition*. It is discussed briefly below. (See the end of this article for relevant cross-references.)

### MORTALITY DECLINE

The transformation of mortality from a common event, occurring most often to children, to a relatively rare event, occurring most often to the elderly, arose from a confluence of technological and social developments. Most important among these was the control of infectious diseases spread by microorganisms in the air and water. An improved understanding of the etiology of these diseases, together with technical capacity and political support for public health measures, led to childhood vaccinations, clean drinking water, and improved sanitation (McKeown 1976). Standards of personal hygiene and cleanliness of clothing improved. There is little evidence that improvements in diet were important, and curative (as contrasted to preventive) medicine played a relatively small role in the main part of the transformation.

In developed countries, the infant mortality rate has steadily fallen from about 250 deaths (in the first year of life) per 1,000 births to a present level of fewer than 10 per 1,000 births (Mosley and Chen 1984). One consequence of a decline in the risk of infant and child deaths is to increase the sense of parental control over reproduction. It is more rational to develop notions of desired numbers of children when the survivorship of children is less random. Similarly, as survivorship improves, it is more rational to invest in children's future by providing them with formal education. The cost of children, to their parents and to the larger society, increases as child mortality falls and life expectancy improves (see, e.g., Easterlin 1976).

The increase in life expectancy, currently about 73 years in the developed countries, has also resulted in a rise in the mean age of the population, in a substantial increase in the proportion who are elderly or retired, and in a shift to causes of death associated with old age. These trends have broad social implications—for the employment of and advancement in opportunities for young people, the resilience of political structures, and the cost of retirement programs and medical care for the elderly, for example.

Fewer births per woman, together with the now negligible rates of maternal mortality in developed countries, have resulted in a substantially greater life expectancy for women than for men. Although at birth there are about 104 males for every 100 females, there are progressively more females, per male, for every age after about age 30 in the United States. The elderly population consists disproportionately of women. Also, because women tend to be younger than their husbands, they typically experience much longer periods of widowhood than men do.

In today's developing countries, mortality decline has been much more rapid than it was in the developed countries, because an accumulation of Western public health measures could be introduced nearly simultaneously. Most of the decline has occurred since World War II and the ensuing independence of most of these countries from colonial powers, although some of it can be traced back to earlier decades of the twentieth century. The rapidity of the mortality decline and its largely exogenous nature have been factors in the delay of a subsequent fertility decline in many cases (see



Preston 1978). (See the end of this article for relevant cross-references.)

### FERTILITY DECLINE

Within Europe, the onset of substantial reductions in fertility occurred first in France early in the nineteenth century, and last in Ireland early in the twentieth century. In the earliest cases, the onset of fertility control coincided with mortality decline rather than following it. In general, a lag between the decline in mortality and the decline in fertility resulted in substantial population growth. In the United States and Britain, 1880 is regarded as a watershed year for the widespread initiation of contraception.

With no exceptions, from the cases of France through Ireland, the initiation and the bulk of the modern fertility declines occurred mainly as a result of contraception rather than delayed marriage, and in contexts in which contraception was publicly regarded as immoral, supplies and information were illegal, and methods were primitive by today's standards. As a generalization, births were not intentionally spaced or postponed; rather, attempts were made to terminate childbearing at some earlier parity than would have been the case without intervention. Some married couples appear even to have chosen to have no children or only one child. The main contraceptive method was withdrawal (*coitus interruptus*). Abstinence was probably not infrequently used as a last resort. Rhythm may have been used, but probably incorrectly; douching was common but was probably ineffective. Sterilization was not available, although it is likely that a high proportion of hysterectomies served the same function. Condoms were not widely available until the twentieth century.

It is clear that the motivation to control fertility was both powerful and personal. It is unfortunate that it is so difficult at this distance to reconstruct the specific strategies that were employed, patterns of communication between couples, and sources of information. However, at least two generalizations can be made. One is that the practice of contraception required an ideational justification, to the effect that individual couples have a personal right to control their family size. From a modern perspective it is easy to overlook the fundamental importance of this concept. It is not just a coincidence that France was the first country

to experience contraception on a wide scale, that it was the home of the Enlightenment, and that it was also the first European country to experience a fundamental political revolution. Intervention to prevent a birth rests on the premise that it is legitimate for an individual—or a couple—to make critical choices affecting personal welfare. Contraception can be viewed as a manifestation of a value for personal freedom, even in the face of strong pronatalist pressures from both church and state.

A secondary condition for contraceptive use in the West appears to have been some degree of local development, as evidenced by higher literacy and a higher standard of living. Historical research continues into the importance of specific factors such as the relative status of women, the transition to a wage-earning class, local industrialization, improvements in social security and public welfare, and so on. (For more details on specific countries, see, e.g., Ryder [1969]; Livi Bacci [1977]; Teitelbaum [1984]. For a general discussion of these factors in Europe, see van de Walle and Knodel [1980]. For theoretical discussions, see Caldwell [1976, 1978]. See also Nam [1994].)

Turning to the transition in economically less developed countries, one to two generations of reduced mortality, combined with a traditional high level of fertility, resulted in annual growth rates of 3 percent or more. However, beginning in the late 1960s, some Asian countries, particularly Taiwan and South Korea, began to experience rapid declines in fertility. At present these countries have reached approximate equilibrium between fertility and mortality rates, although they continue to grow because of their youthful age distributions. About a decade later, Thailand, Indonesia, and several Latin American countries such as Colombia and Mexico showed rapid reductions in their fertility rates. By the late 1990s, dramatic fertility declines were under way in nearly all countries outside of sub-Saharan Africa and Pakistan, and even these countries have shown clear declines among better-educated urban couples. The declines are due in small part to delayed marriage, but for the most part to use of contraception—primarily sterilization, and secondarily reversible methods such as intrauterine devices and the pill (see Bulatao and Lee [1983]; Cleland and Hobcraft [1985]; see also the country reports published on the Demographic and Health Surveys Project by Macro International [1985–2000]).

The conditions for these fertility declines show both similarities to and differences from the Western declines. It appears critical for couples to accept the idea that it is appropriate to intervene in the procreative process. In Pakistan, for example, it is commonly held that the number of births, as well as their gender and survivorship, is in the hands of Allah, and it would be wrong to interfere with his will. (It must be noted that a stated religious rationale for high fertility often masks other factors, such as a subordinate role for women. In other Islamic countries, such as Indonesia, Bangladesh, and even Iran, family planning is considered to be consistent with Islam.) Contraception tends to be found where the concepts of political and economic self-determination are better established, particularly among women. Female education is the single strongest correlate of fertility change.

In contrast with the Western experience, however, it also appears critical to have institutional support for contraception. The countries that have shown the clearest declines in fertility had national family planning programs with visible support from the highest levels of the government. The most effective programs have integrated family planning services into a general program of maternal and child health, and provide couples with easy access to a range of alternative methods (see Lapham and Mauldin 1987). Many countries are actually passing beyond this phase of the contraceptive transition, beginning in the 1990s, so that government-subsidized programs are being replaced by privatized services, at least for the middle class.

The consequences of population growth for economic development have been much debated (Simon 1977; Birdsall 1980). There have been some cases, such as Japan and South Korea, in which rapid economic expansion occurred simultaneously with rapid increases in population. Virtually all such cases were transitional, and the fertility of those countries is currently at replacement level, so the debate is now of mainly historical interest. There is a general consensus that low growth facilitates development. A growing population has a young age distribution, with many new entrants to the labor force and relatively few old people in need of pensions and health care. These factors may stimulate economic growth, but they must be balanced against the costs of supporting

and educating large numbers of children. In several countries, such as the Philippines, the economy is unable to employ large cohorts of young people satisfactorily, especially those who are better educated, and they emigrate in large numbers. In addition, household welfare can be adversely affected by large numbers of children, even in situations of economic expansion at the macro level. The negative consequences of rapid growth extend beyond the economy and into the areas of health, social welfare, political stability, and the environment. (See the end of this article for relevant cross-references.)

### CHANGES IN POPULATION DISTRIBUTION

Enormous changes in geographical distribution have been superimposed on the major trends in population size described above. Many of the social problems attributed to rapid population growth are more accurately diagnosed as consequences of increasing concentration. Urban areas, in particular the megalopolises of developing countries such as Mexico City, Buenos Aires, and New Delhi, have been growing during the twentieth century at more than twice the rate of the countries in which they are located. Cities are centers of concentration of economic, intellectual, and political life (see Hawley 1981), but rapid growth has exacerbated the problems of inadequate housing, sanitation, transportation, schooling, unemployment, and the crime associated with urban life. It is estimated that, in the year 2000, 76 percent of the population of developed countries is urban (compared to 26 percent in 1900 and 55 percent in 1950). In the developing countries, 41 percent is urban in 2000 (compared to 7 percent in 1900 and 18 percent in 1950).

The growth of cities has resulted in part from the excess of births over deaths in rural areas. With out-migration serving as one of the household-level valves for population regulation, individuals have been displaced from areas that cannot absorb them and have moved to cities, which are perceived to have better economic opportunities. Often that perception is incorrect. With a few exceptions, fertility is lower in cities than in rural areas.

A second major type of population redistribution in recent centuries has, of course, been across

national borders. Movement to the Americas was greatest during the half-century between 1880 and 1930, and continues to the present. There are many streams of both short-term and long-term international migration, for example out of South and Southeast Asia and into the Middle East, Europe, and North America, and from Africa into Europe and North America, and the economies of several sending countries are strengthened by monthly remittances from their emigrants (United Nations 1979, 1997). (See the end of this article for relevant cross-references.)

### POPULATION OF THE UNITED STATES

The United States has a population of approximately 278 million in the year 2000 and a growth rate of somewhat less than 1 percent annually, roughly one-third of which is due to immigration and two-thirds to natural increase. Fertility is slightly below replacement level, but there are more births than deaths because of the large size of the baby boom cohort, born from the late 1940s through the early 1960s. Age-specific fertility rates are increasing gradually for women in their thirties and forties, mainly because of postponed first and second births rather than later births. Otherwise, rates have been remarkably stable since the early 1970s. Rates below age 20 were dropping in the late 1990s but are still higher than in most other developed countries. For several excellent articles on fertility levels, differentials, and trends in the United States, see Casterline and colleagues (1996). See also National Center for Health Statistics (1999) and more recent annual reports in the same series.

Perhaps the most serious issues related to fertility are the large numbers of unplanned births to young women and the high numbers of abortions—about 2 for every 5 births—that could have been averted by contraception. Few developed countries have a range of contraceptive methods as limited as that of the United States. For example, intrauterine devices (IUDs) and progesterone-based pills are the two main nonsurgical methods in the rest of the world, but IUDs are not available in the United States. As mentioned earlier, most of the fertility decline in the West occurred while contraception was considered immoral and was explicitly illegal. Although contraception and even abortion are now legal, deep cultural ambiguities

remain in the linkage between sexuality and procreation. A litigious environment has inhibited both the development of new contraceptives by American pharmaceutical companies and the marketing of new contraceptives developed elsewhere, and the U.S. government plays a minimal role in such development.

Life expectancy in the late 1990s was 73 years for males and 79 years for females (PRB 1998). Although life expectancy is increasing for both males and females, the female advantage is also gradually increasing. The female advantage was nearly 3 years at the beginning of the twentieth century, nearly 4 years at the middle, and nearly 7 years at the end of the century (Gelbard et al. 1999). The greatest improvements in mortality are in the highest age groups, especially after age 85. Because of increases in the number of elderly and projected changes in the age distribution of the labor force, the age at receipt of full Social Security benefits is scheduled to increase gradually, from 65 to 66 by the year 2009 and to 67 by the year 2027 (Binstock and George 1990).

Among males, whites have approximately a 5-year advantage over nonwhites, and among females, whites have approximately a 4-year advantage (National Center for Health Statistics 1989; and more recent annual reports in the same series). Life expectancy for black males is falling, due mainly to deaths by homicides to black males in their twenties and a greater prevalence of cardiovascular disease among blacks (Keith and Smith 1988). Infant mortality rates for nonwhite babies are increasing, due to low birth weights and inadequate prenatal care. These reversals of earlier long-term improvements are indicators of worsening conditions among poorer Americans. (For more description of the population of the United States, see Bogue [1985]; Lieberman and Waters [1988]; Sweet and Bumpass [1988]; Nam [1994].) (See the end of this article for cross-references to articles containing more discussions of the population of the United States.)

### FUTURE POPULATION

The overriding concern of world population policy in recent years has been the achievement of a new equilibrium between fertility and mortality,

so that growth will be slowed or stopped. Sometimes this policy is stated in terms of enabling couples to have the number of children they want to have, and no more, because surveys in most developing countries indicate that couples desire smaller families than they actually have (see the country reports published by Macro International [1985–2000]).

The concept of *replacement fertility* is important for understanding population projections. Normally stated in terms of the female population, the reproductive value of a woman of a given age is equal to the number of daughters she will have who will survive to (at least) this same age. If the average reproductive value is exactly 1, then fertility is at replacement level. This occurs when the average woman has about 2.1 births in her lifetime (a little above 2.0, to compensate for children who do not survive). If there has been a history of higher fertility, then the population has been growing and is relatively young, with many women in the peak ages of childbearing. As a result, there can continue to be more births than deaths (i.e., population growth) for a very long time after the net reproduction rate has reached or even fallen below the replacement value of 1.0. However, in the long term, replacement fertility will lead to a no-growth population, and below-replacement fertility will lead to population decline.

Strictly speaking, reproduction is intergenerational, but it is estimated with the net reproduction rate, a synthetic measure calculated from age-specific fertility and survival rates within an interval of time such as one year. If an artificial cohort of women is subjected to these rates and does not replace itself, then current fertility is interpreted to be below replacement.

Projected improvements in mortality would have a relatively low impact on population growth. For reproduction, it is survivorship up to and through the childbearing ages that matters. Improvements in survivorship after age 45 or so have little impact. Projections do require some assumptions about the future of the human immunodeficiency virus (HIV) and acquired immune deficiency syndrome (AIDS) epidemic, primarily in sub-Saharan Africa and also in South and Southeast Asia.

In the year 2000, world population is estimated to be 6.1 billion, with a growth rate of about 1.8

percent annually. If the current rate were to continue, world population would reach 9.2 billion by the year 2025. Even if reproduction immediately came into balance with mortality, as just described, world population would continue to increase. Population momentum, due to the youthful age distribution, would produce an excess of births over deaths (i.e., growth) for more than a century. World population would reach 7.5 billion by 2025, eventually stabilizing at nearly 10 billion. Substantial future growth is inevitable, but there is a wide range of possible scenarios.

Projections developed by the United Nations (United Nations 1998) for the year 2025 range from a low variant of 7.3 billion to a medium variant of 7.8 billion and a high variant of 8.4 billion, depending on assumptions about future mortality and (more important) future fertility. By comparison, ten years earlier (United Nations 1988) the *low* projection for 2025 was 8.5 billion, slightly above the current *high* projection. This substantial downward revision reflects a general optimism that the threat of a world population explosion has largely receded. (It also implies that little credibility should be attached to longer-term projections, and for that reason this article will not cite projections beyond 2025.)

All the current projection scenarios assume that fertility will decline. The lowest estimate is based on an assumption that it will decline to below-replacement levels very soon, after 2005, and will decline steadily thereafter. This assumption is highly improbable, but even the medium variant assumes that worldwide fertility will be below replacement by 2030. This is a remarkable change from earlier projections, because prior to the late 1990s almost all scenarios assumed that fertility would ultimately converge to replacement level, with uncertainty only about when that would occur. Now that better data are available on actual changes during the 1980s and 1990s, it is considered likely that fertility will fall *below* replacement level early in the twenty-first century.

The net reproduction rate is currently less than 1.0 (that is, fertility is below the long-term level needed to balance mortality) in virtually all the more developed countries. The one-fifth of the world's population that resides in those countries will increase scarcely at all, and much of that growth will be the result of immigration from

developing countries. Around the year 2000, compared to a standard of replacement fertility, Europe is at 68 percent of replacement level and North America is at 93 percent, while Asia is 10 percent above replacement, Latin America is 20 percent above, and Africa is 84 percent above. (These percentages come from estimates and medium-variant projections for 2000 by the United Nations [1998].) It is projected that low growth in Europe will cause the median age to rise to the mid-forties by 2025, with a fifth of the population above 65 years of age, a situation that the United States will approach a few years later. An excellent discussion of the causes and effects of future population change, and of many of the other topics in this article, can be found in Gelbard and colleagues, Haub, and Kent (1999).

Although world growth has been the preoccupation of recent decades, the possibility of population decline has long been acknowledged in Europe. In the United States, fertility has been below replacement since approximately 1970, and if it were not for high levels of immigration, this country would also face the prospect of population decline. Policies directed at increasing fertility in European countries have met with little success (see Calot and Blayo 1982; van de Kaa 1987). In urban settings with a high standard of living, children lose much of their earlier value as a source of economic activity, household wealth, and security in old age. They become increasingly expensive in terms of direct costs such as clothing, housing, and education, and in terms of opportunity costs such as forgone labor force activity by the mother.

In brief, there are probably two main reasons why fertility has not declined even further in the developed countries. One is the adherence to a powerful norm for two children that was consolidated around the middle of the twentieth century. Surveys show an overwhelming preference for exactly two children—preferably one boy and one girl, especially in the United States—with little flexibility either above or below that number. Actual childbearing often departs from the norm, more commonly by being below two children, so that fertility is below replacement. A high proportion of childbearing beyond two children is due to a desire for at least one child of each sex. Reliable methods to achieve the desired sex composition would result in a noticeable reduction of third and later births.

Second, children provide parents with a primary social group. There is no longer an expectation that they will provide support in old age, nor an important concern with carrying on the family name, but children do provide psychic and social rewards. To bear children is to emulate the behavior of one's parents and to replicate the family of orientation. However, this goal can be largely attained with only one child, as is being demonstrated in the lowest-fertility countries of Europe and in urban China. As increasing numbers of women opt for no children or only one child, it is possible that the widespread preference for two will weaken, even in the United States. For further discussion of fertility preferences in the United States, see Schoen and colleagues (1997).

Although the world as a whole is far from experiencing a decline in population, the low reproductivity of some countries and subpopulations raises questions about future mechanisms for restraining an indefinite decline in fertility, and eventually in population (see Teitelbaum and Winter 1985; Davis et al. 1986). It is reasonable to speculate on whether the cultural and social props for replacement fertility will continue to hold, or whether, as in the low and medium UN projections, worldwide fertility will decline to below-replacement levels early in the twenty-first century.

Major reductions in fertility in the past have been the result of delayed marriage and contraception and have been motivated at the level of the household. Maintenance of equilibrium in the future will require an increase in desired family size and less use of contraception. Many household-level factors associated with economic development would seem to support a projection of continued decline in fertility—for example, increased labor-force participation and autonomy for women, declines in marriage rates, increased costs for the education of children, and increased emphasis on consumption and leisure. It is easier to project a continued decline in fertility, rather than a significant upturn, in the absence of major changes or interventions in the microeconomy of the household. However, it also seems plausible that children will take on an increased (noneconomic) value as they become scarcer, or that subpopulations that favor high fertility will come to dominate, in which case the world will return to the previous pattern of long-term homeostasis at some level, or

will establish a new pattern of very gradual growth, rather than allowing an inexorable decline.

(SEE ALSO: *Birth and Death Rates; Demographic Transition; Demography; Ethnicity; Family Planning; Infant and Child Mortality; Internal Migration; International Migration; Life Expectancy; Race; Retirement; Urbanization*)

## REFERENCES

- Binstock, Robert H., and Linda K. George (eds.) 1990 *Handbook of Aging and the Social Sciences*. London: Academic.
- Birdsall, Nancy 1980 "Population Growth and Poverty in the Developing World." *Population Bulletin* 35 (5).
- Bogue, Donald J. 1985 *The Population of the United States: Historical Trends and Future Projections*. New York: Free Press.
- Bulatao, Rodolfo, and Ronald D. Lee 1983 *Determinants of Fertility in Developing Countries*. New York: Academic.
- Caldwell, John G. 1976 "Toward a Restatement of Demographic Transition Theory." *Population and Development Review* 2:321–366.
- 1978 "A Theory of Fertility: From High Plateau to Destabilization." *Population and Development Review* 4:553–577.
- Calot, G., and Chantal Blayo 1982 "The Recent Course of Fertility in Western Europe." *Population Studies* 36:349–372.
- Casterline, John B., Ronald D. Lee, and Karen A. Foote (eds.) 1996 *Fertility in the United States: New Patterns, New Theories*. *Population and Development Review*, a supplement to vol. 22.
- Cleland, John, and John Hobcraft (eds.) 1985 *Reproductive Change in Developing Countries*. Oxford: Oxford University Press.
- Coleman, David, and Roger Schofield (eds.) 1986 *The State of Population Theory*. Oxford: Basil Blackwell.
- Davis, Kingsley 1963 "The Theory of Change and Response in Modern Demographic History." *Population Index* 29:345–366.
- , Mikhail S. Bernstam, and Rita Ricardo-Campbell (eds.) 1986 *Below-Replacement Fertility in Industrial Societies: Causes, Consequences, Policies*. *Population and Development Review*, a supplement to vol. 12.
- Easterlin, Richard A. 1976 "The Conflict between Resources and Aspirations." *Population and Development Review* 2:417–425.
- Gelbard, Alene, Carl Haub, and Mary M. Kent 1999 "World Population beyond Six Billion." *Population Bulletin* 54(1): Population Reference Bureau.
- Goody, Jack 1976 *Production and Reproduction*. Cambridge: Cambridge University Press.
- 1983 *The Development of the Family and Marriage in Europe*. Cambridge: Cambridge University Press.
- Greenhalgh, Susan 1996 "The Social Construction of Population Science: An Intellectual, Institutional, and Political History of Twentieth Century Demography." *Comparative Studies in Society and History* 38:26–66.
- Hanley, Susan B., and Arthur P. Wolf (eds.) 1985 *Family and Population in East Asian History*. Stanford, Calif.: Stanford University Press.
- Hauser, Philip M., and Otis Dudley Duncan (eds.) 1959 *The Study of Population*. Chicago: University of Chicago Press.
- Hawley, Amos H. 1981 *Urban Society*. New York: John Wiley.
- Keith, Verna M., and David P. Smith 1988 "The Current Differential in Black and White Life Expectancy." *Demography* 25:625–632.
- Lapham, Robert J., and W. Parker Mauldin 1987 "The Effects of Family Planning on Fertility: Research Findings." In R. J. Lapham and G. B. Simmons, eds., *Organizing for Effective Family Planning Programs*. Washington: National Academy Press.
- Laslett, Peter, and Richard Wall (eds.) 1972 *Household and Family in Past Time*. Cambridge, England: Cambridge University Press.
- Liebersohn, Stanley, and Mary C. Waters 1988 *Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage Foundation.
- Livi Bacci, M. 1977 *A History of Italian Fertility*. Princeton, N.J.: Princeton University Press.
- McEvedy, Colin, and Richard Jones 1978 *Atlas of World Population History*. New York: Facts on File.
- McKeown, Thomas 1976 *The Modern Rise of Population*. New York: Academic.
- Macro International, Inc. 1985–2000 Country reports from the Demographic and Health Surveys Project, Calverton, Md.
- Mosley, W. Henry, and Lincoln C. Chen (eds.) 1984 *Child Survival: Strategies for Research*. *Population and Development Review*, a supplement to vol. 10.
- Nam, Charles 1994 *Understanding Population Change*. Itasca, Ill.: F.E. Peacock.
- Nambodiri, Krishnan 1988 "Ecological Demography: Its Place in Sociology." *American Sociological Review* 53:619–633.
- National Center for Health Statistics 1999 "Births: Final Data for 1997." *National Vital Statistics Reports* 47(18). See also more recent annual reports in the same series.

- Petersen, William 1975 *Population*. New York: Macmillan.
- Population Reference Bureau (PRB) 1998 *World Population Data Sheet*. Washington: PRB. See also more recent publications in the same series.
- Preston, Samuel H. 1978 *Mortality Patterns in National Populations*. New York: Academic.
- Ryder, Norman B. 1969 "The Emergence of a Modern Fertility Pattern: United States 1917-1966." In S. J. Behrman, ed., *Family Planning: A World View*. Ann Arbor: University of Michigan Press.
- Schoen, Robert, Young J. Kim, Constance A. Nathanson, Jason Fields, and Nan Marie Astone 1997 "Why Do Americans Want Children?" *Population and Development Review* 23:333-358.
- Simon, Julian 1977 *The Economics of Population Growth*. Princeton, N.J.: Princeton University Press.
- Stycos, J. Mayonne (ed.) 1987 *Demography as an Interdiscipline*. *Sociological Forum* 2 (4). Special issue.
- Sweet, James A., and Larry L. Bumpass 1988 *American Families and Households*. New York: Russell Sage Foundation.
- Tauber, Karl E., Larry L. Bumpass, and James A. Sweet (eds.) 1978 *Social Demography*. New York: Academic.
- Teitelbaum, Michael S. 1984 *The British Fertility Decline: Demographic Transition in the Crucible of the Industrial Revolution*. Princeton, N.J.: Princeton University Press.
- , and Jay M. Winter 1985 *The Fear of Population Decline*. San Diego, Calif.: Academic.
- Tilly, Charles (ed.) 1978 *Historical Studies of Changing Fertility*. Princeton, N.J.: Princeton University Press.
- United Nations 1973 *The Determinants and Consequences of Population Trends: New Summary of Findings on Interaction of Demographic, Economic and Social Factors*, vol. 1. *Population Studies* 50. New York: UN.
- 1979 *Trends and Characteristics of International Migration Since 1950*. New York: UN.
- 1988 *World Demographic Estimates and Projections, 1950-2025*. New York: UN.
- 1997 *World Urbanization Prospects: The 1996 Revision*. New York: UN.
- 1998 *World Population Prospects: The 1998 Revision*. New York: UN.
- van de Kaa, Dirk J. 1987 "Europe's Second Demographic Transition." *Population Bulletin* 42(1).
- van de Walle, Etienne, and John Knodel 1980 "Europe's Fertility Transition: New Evidence and Lessons for Today's Developing World." *Population Bulletin* 34 (6).
- World Bank 1990 *World Development Report 1990*. Washington: Oxford University Press.

- Wrigley, E. A. 1969 *Population and History*. New York: McGraw-Hill.
- , and R. S. Schofield 1981 *The Population History of England, 1541-1871: A Reconstruction*. London: Edward Arnold.
- Yaukey, David 1985 *Demography: The Study of Human Population*. New York: St. Martin's.

THOMAS W. PULLUM

## PORNOGRAPHY

Most societies have some kinds of sexually explicit words, images, and other materials that are disapproved, prohibited, illegal, or restricted in circulation; that are portrayed in a manner that is generally unacceptable; and that are classified as pornography. The term *pornography* often denotes subjective disapproval of materials rather than their content or effect.

Historically, attitudes toward pornography tend to be cyclical. The most recent expansion in the United States of the availability of such materials began in the 1960s, as a result of larger social trends, including the civil rights movement, the counterculture and antiwar movement, approval of the contraceptive pill, and liberal U.S. Supreme Court decisions providing access to previously banned novels and films. These trends contributed to continuing discussion on whether new forms of social control would be necessary to cope with pornography, and federal government investigations into pornography's impact were launched in 1968 and 1985.

Both the President's Commission on Obscenity and Pornography (1971) and the Attorney General's Commission on Obscenity and Pornography (1986) felt that it was not possible to define pornography clearly. The term has no legal definition, although certain sexual content is not protected by the free speech guarantees of the First Amendment to the U.S. Constitution and can be prosecuted under federal and state laws regulating obscenity.

Over the last several decades, the connotations of pornography have changed from a generic descriptor of explicit sexual content to more pejorative representations of sexual material. Pornography is often distinguished from erotica, which is

a more tender, humanistic, and artistic type of sexual representation.

The divergence in opinion on how to deal with the pornography problem could be seen in the different recommendations of the two federal commissions established to examine it. The President's Commission recommended that the laws on obscenity should be repealed, because the material represented an insignificant social problem and had no lasting negative effect on its users. The report of the Attorney General's Commission urged that the obscenity laws should be strictly enforced, because the material had significant negative impact and constituted an important national problem. One possible reason for this disagreement was the change in the market for explicit materials between 1970 and 1986, but another factor could have been differences in the time available, composition, and resources of the two commissions.

Because sexually explicit content is often initially excluded from mainstream channels of communication, its creators have usually been alert to newly emerging outlets for distribution. One of the first books to appear in print after Gutenberg developed printing around 1448 was Boccaccio's erotic masterpiece, *The Decameron*. Soon after Edison developed moving pictures in 1891, underground "blue" movies began to appear. Super 8mm movies presenting sexually explicit content could be found in 1970, the very first year that the new technology appeared.

In the 1980s, soon after deregulation of the telephone business, many sex telephone services offered both recorded messages and the opportunity to have individualized "audiotext" conversations. Around the same time, sexually explicit movies became available via pay-per-view cable systems servicing homes and hotel chains.

In 1979, when fewer than 1 percent of American homes had videocassette recorders (VCRs), 80 percent of VCR owners bought or rented sexually explicit tapes. Some stores boosted sales of VCRs by giving a sexually explicit videotape as a gift with every purchase. During the 1990s, perhaps one-third of adult males and one-fifth of adult females have seen a sexually explicit videotape in a given year. The VCR is especially popular for viewing sexual content because it permits the user to speed up, slow down, replay, or freeze images. America is

the world's largest producer of sexually explicit videotapes, churning out approximately 7,500 per year. The huge growth in the explicit video market during the 1980s and 1990s paralleled an increase in the number of large strip clubs around the country. The videos have become a way of promoting the careers of the actresses who appear at the clubs, who also can be seen in men's magazines.

CD-ROM and Internet income from sex-related content cannot be accurately measured but is believed to be enormous. A series of U.S. Supreme Court decisions has established the Internet as a public forum that is entitled to the highest level of protection by the First Amendment, including the representation of sexually explicit materials. Several laws that restricted access of children to the Internet have been declared unconstitutional by federal courts.

Rating systems represent one approach to social control of sexual content. The 1968 original movie rating system, by assigning sexually explicit materials to a special Adult X category and the development of similar kinds of ratings for television programs and popular music, further helped legitimate sexual content.

Within several years after the movie rating system began, X-rated films began to be shown at major theaters. The X-rated *Last Tango in Paris*, starring Marlon Brando, the country's most celebrated actor, was one of the most popular films of 1973, the same year in which *Deep Throat* became a national sensation because of its enormous commercial success and humor. The X-rated *Midnight Cowboy* had already received an Academy Award as Best Picture in 1969.

In addition to ratings, technological approaches have been used as a way of regulating children's access to sexual content. A special chip was developed so that parents could block children's access to objectionable materials on television. Such approaches attempt to strike a balance between the American abhorrence of censorship and parents' concern about their children's exposure to inappropriate material.

The ultimate weapon in blocking the availability of offensive sexual material has always been the application of the laws against obscenity. Three



criteria for obscenity were promulgated by the U.S. Supreme Court in the case of *Miller v. California* (413 U.S. 15, 1973): the material had to be patently offensive; it must appeal to the prurient interest of the average adult applying contemporary standards; and it must have no literary, artistic, political, or scientific value.

Because the Court no longer wanted to deal with the stream of cases that it had been handling, "community" was newly defined to be a local county, state, or federal court district, so that each area could have whatever level of sexual material it found acceptable. The number of indictments and convictions declined after 1973 because of juries' reluctance to convict, prosecutorial predictability, changes in the political climate that led to antipornography becoming a less salient and less attractive political theme, and communities' views that criminal justice resources should give less attention to obscenity and focus more on crimes against persons.

The decline in implementation of criminal justice sanctions encouraged new entrepreneurs to produce and distribute sex-oriented materials. Another spur to entrepreneurs, in the 1980s, was the Federal Communications Commission's decision not to regulate cable television, and this was followed in 1984 by the television networks' decision to eliminate their national code of broadcast standards. The continuing competition for audiences between media further increased the likelihood that sexual content would receive more prominence, in a situation in which administrative barriers to representation of sex content were clearly diminishing.

Instead of criminal prosecution of materials for violation of the obscenity laws, a number of cities adopted a totally different approach in the 1990s. These communities use zoning laws against theaters, strip clubs, bookstores, and other sex-related industries in order to break up their concentration and minimize visual pollution. The new approach accepted these establishments' right to sell sexually explicit products or services but challenged their right to maintain their location.

The zoning laws rely on a U.S. Supreme Court decision which held that if "adult" businesses could be shown to have adverse secondary effects

(decline in real estate values and quality of life and increase in crime), the community could make zoning changes, provided that reasonable alternative avenues of communication were available to the businesses (*Playtime Theatres v. City of Renton*, 106 U.S. 925, 1986).

An example of the zoning approach could be seen in New York City where a 1997 law required that "adult" bookstores, strip clubs, peepshows, and theaters be at least 500 feet from each other and at least 500 feet from residences and public buildings, thus eliminating the concentration of such establishments in Times Square. Any challenges to such zoning changes have been handled in the civil rather than the criminal courts.

One way in which producers of these materials attempt to minimize legal problems is to make publications, movies, and videotapes in both hard-core and soft-core versions, for targeted distribution to specific markets. The hard-core materials usually show some form of actual sexual interaction and the soft-core content generally presents simulated interaction. Any representation of sexual material involving children has been the target of intensive law enforcement and is very scarce.

There seems to be no relationship between sexually explicit materials and prostitution; consumers of these materials are not likely to be customers of prostitutes, because of differences in the gratifications provided by the two activities. There also seems to be no significant relationship between the production and distribution of such materials and organized crime. There are now so many producers and distributors and the prices are so competitively low that organized crime groups cannot make a large enough profit to justify participation in the business.

Although there is no unified and consistent feminist attitude toward pornography, feminist opposition to pornography has generated much attention. One sustained feminist attack on pornography, jointly led by writer Andrea Dworkin and attorney Catherine A. MacKinnon, argued that pornography is misogynistic and dehumanizing, graphically presenting the sexually explicit subordination of women. They developed a model ordinance for Indianapolis that represented a civil rights approach, which identified pornography as

a violation of the equality guarantee of the Fourteenth Amendment. However, the ordinance was found to be unconstitutional by a federal appeals court (*American Booksellers Association v. Hudnut*, 771 F. 2d, 7th Circuit, 1985). A comparable approach was upheld by the highest court in Canada (*Butler v. The Queen*, 1 SCR 452, 199, 1992), where it has been primarily used to prosecute gay publications.

A collateral approach involved dissemination of feminist writer Robin Morgan's slogan, "Pornography is the theory, rape is the practice." Around the same time that feminist theorists were arguing that the growing availability of misogynistic pornography was the direct cause of increasing rates of rape, a number of mass communication researchers conducted laboratory experiments that attempted to prove the connection between exposure to such materials and the development of callous attitudes toward women and rape. One of the problems with these studies is that they were usually conducted with college students, who were exposed to materials that they would not ordinarily see. There is also the possibility of an experimenter effect and the difficulty of proving the carrying over of an attitude from viewing a film to a real rape situation. Several researchers have expressed the view that to interpret their largely laboratory-based studies as confirming a causal nexus between exposure to sexually explicit films and subsequent sexually violent behavior would represent overgeneralization.

Another problem in arguing for such a nexus is that no large-scale content studies prove that there has been any increase in the amount of sexual violence in the media. An examination of the available data suggests that there has been a decline since 1977 in the amount of sexual violence in the media, but there has probably been an increase in violence in nonpornographic fare, such as action, slasher, and horror films. To single out pornography for stringent legal action and to ignore ideas about rape and violence that are otherwise so pervasive in our culture is not justified by the available evidence.

A number of natural experiments have provided information on the possible relationship between media sex content and rape and other sex crimes. In the United States, four states that eliminated all obscenity statutes and had a concomitant

increase in the availability of sexually explicit materials for varying periods between 1973 and 1986 were studied. During the period of obscenity nonprosecution, the mean annual arrest rates for sex offenses in these four states declined significantly in comparison to the other forty-six states. A study of the first three countries to legalize sexually explicit materials (Denmark, Sweden, Germany) concluded that legalization did not lead to increased sexual violence in any of them.

Researchers using other indirect measures have similarly attempted to correlate sex offenses with previous exposure to sexually explicit materials. A large-scale study by the Kinsey Institute for Sex Research of convicted sex offenders found no relation between sex crimes and offenders' exposure to such materials. Another study of nonincarcerated sex offenders reported no significant differences between the exposure to such materials of the offenders and the general adult male population. Whether on an individual, a regional, or a national basis, empirical studies have failed to find any consistent correlation between availability of and exposure to explicit materials and sex crimes against women.

A number of producers are actively distributing sexually explicit materials targeted for women consumers, and other producers have developed such materials for feminist audiences. Some feminists have produced their own explicit films and others have made videotapes for women and couples that are sold in mainstream superstores and direct mail catalogs. A number of performers who became known from appearing in explicit movies and tapes are obtaining roles in mainstream studio or independent films.

Research has identified a number of functions served by sexually explicit materials. They provide information that is not otherwise available, as well as fantasy outlets and sexual stimulation. They also provide reassurance about one's body and about shame or guilt relating to specific practices. The materials can desensitize people so that they are better able to consider and discuss sexual matters, and can facilitate communication between sexual partners. The materials may provide a socially acceptable functional equivalent for masturbation and for acting out otherwise socially harmful sexual behavior. For members of a group audience, as

in a strip club or a movie theater, the social context can reinforce the experience. The materials and performers also may provide a vehicle for connoisseurship.

Plausible next steps in sociological investigations of pornography could include cross-media content analyses, naturalistic studies of effects, systematic case studies of the role of the various interest groups involved in promoting or blocking pornography, investigation of the structure of the pornography industry, and studies of the role of the home environment in contributing to how explicit materials are used.

In Denmark, the first country to legalize pornography (1969), the market for such materials declined sharply, as a result of satiation of the audience. Tracking the cycle of acceptance of pornography in America, to see whether the Danish experience is relevant, could represent an important study. A collateral study could determine whether there are specific subgroups in the population for whom these materials are becoming more or less important. The extent to which pornography continues to be a feature of the political culture wars and a vehicle for obtaining support for larger economic and policy initiatives represents an important social indicator, as is the alliance between feminists and moral conservatives in opposing pornography. Another activity that will surely bear watching in the future is how the effort to regulate pornography on the Internet, for children or adults, can be accomplished by a state or a country that does not have legal jurisdiction over the global information network that is the Internet.

REFERENCES

Attorney General's Committee on Obscenity and Pornography 1986 *Final Report*. Washington, D.C.: U.S. Government Printing Office.

Gebhard, P. H., J. H. Gagnon, W. B. Pomeroy, and C. Christenson 1965 *Sex Offenders*. New York: Harper and Row.

Kimmel, M. (ed.) 1989 *Men Confront Pornography*. New York: Crown.

Lederer, L. (ed.) 1980 *Take Back the Night: Women on Pornography*. New York: William Morrow.

MacKinnon, C. A. 1993 *Only Words*. Cambridge, Mass.: Harvard University Press.

President's Commission on Obscenity and Pornography 1971 *Final Report*. Washington, D.C.: U.S. Government Printing Office.

Strossen, N. 1995 *Defending Pornography: Free Speech, Sex, and the Fight for Women's Rights*. New York: Scribner.

Winick, C. 1977 "From Deviant to Normative: Changes in the Social Acceptability of Sexually Explicit Material." In E. Sagarin, ed., *Deviance and Social Change*. Beverly Hills, Calif.: Sage.

Zillman, D., and J. Bryant (eds.) 1989 *Pornography: Research Advances and Policy Considerations*. Hillsdale, N.J.: Lawrence Erlbaum.

CHARLES WINICK

**POSITIVE MENTAL HEALTH**

The concept of positive mental health was developed by Marie Jahoda, who argues that positive mental health can be viewed as an enduring personality characteristic or as a less permanent function of personality and the social situation (Jahoda 1958). This article summarizes Jahoda's approach to positive mental health, reviews other discussions of the concept, describes a challenge to the assumption that positive mental health requires the accurate perception of reality, examines the value assumptions inherent in the concept, and compares the notions of normality and positive mental health.

In her classic book, *Current Concepts of Positive Mental Health* (1958), Jahoda identified the following six approaches to the definition of positive mental health, which are described in detail below: (1) attitude toward own self; (2) growth, development, and self-actualization; (3) integration; (4) autonomy; (5) perception of reality; and (6) environmental mastery.

1. Acceptance of self, self-confidence, and self-reliance characterize the mentally healthy person. An important attribute of positive mental health includes the understanding of one's strengths and weaknesses, coupled with the conviction that one's positive characteristics outweigh the negative traits. Independence, initiative, and self-esteem are other indicators of positive mental health.

2. The realization of one's potential is the underlying assumption of this dimension of positive mental health. Maslow (1954) explains that

self-actualization is a motive that encourages the person to maximize capabilities and talents. It is hypothesized that growth motivation is related to positive mental health. Rather than meeting basic human needs, self-actualization implies movement toward higher goals. This dimension of positive mental health also implies an investment in living—a concern with other people and one’s environment, rather than a primary focus on satisfying one’s own needs.

3. The person with positive mental health has a balance of psychic forces, a unified outlook on life, and resistance to stress. Psychoanalysts view integration as the balance of the id, the ego, and the superego. This balance is viewed as changeable, with flexibility as the desired end result. Positive mental health refers to integration at the cognitive level, which implies a unifying philosophy of life that shapes feelings and behaviors. Finally, resistance to stress characterizes the integrated person. The mentally healthy person can adapt to stress without deteriorating. Everyone experiences anxiety when encountering a stressful situation. A mentally healthy response to anxiety and stress suggests some tolerance of tension, ambiguity, and frustration.

4. Autonomy refers to self-determination and independence in decision making. The concept suggests that the person with positive mental health is self-directed and self-controlled. The individual acts independently of the outside world; behavior is not dictated by environmental circumstances.

Jahoda points out that some authors have a different interpretation of autonomy. Autonomy may be defined as having freedom of choice about conforming to societal norms. This perspective implies that the person is not independent of the environment, but does have free choice to decide how to respond to societal demands.

5. “As a rule, the perception of reality is called mentally healthy when what the individual sees corresponds to what is actually there” (Jahoda 1958, p. 49). Mentally healthy reality perception includes perception free from need distortion. A mentally healthy person views the world without distortions, fitting the perception to objective cues that are present, and does not reject evidence because it does not fit his or her wishes or needs.

Jahoda argues that this dimension of positive mental health implies the ability to perceive others in an empathetic manner. This social sensitivity enables a healthy person to put himself or herself in another person’s place and anticipate that person’s behavior in a given social situation.

6. Mastery of the environment refers to achieving success in some social roles and appropriate function in those roles. Positive mental health also includes the ability to have positive affective interpersonal relations. The social roles involved in environmental mastery may include sexual partner, parent, and worker. Environmental mastery suggests the ability to adapt, adjust, and solve problems in an efficient manner.

#### OTHER DEFINITIONS OF POSITIVE MENTAL HEALTH

Jourard and Landsman propose similar criteria for positive mental health: positive self-regard, ability to care about others, ability to care about the natural world, openness to new ideas and to people, creativity, ability to work productively, ability to love, and realistic perception of self (1980, p. 131).

Jensen and Bergin (1988) conducted a nationwide survey of 425 professional therapists (clinical psychologists, marriage and family therapists, social workers, and psychiatrists) to determine values associated with mental health. Eight themes were identified as important for a positive, mentally healthy lifestyle: (1) competent perception and expression of feelings (sensitivity, honesty, openness with others); (2) freedom/autonomy/responsibility (self-control, appropriate feelings of guilt, responsibility for one’s actions, increasing one’s alternatives at a choice point); (3) integration, coping, and work (effective coping strategies, work satisfaction, striving to achieve); (4) self-awareness/growth (awareness of potential, self-discipline); (5) human relatedness/interpersonal and family commitment (ability to give and receive affection, faithfulness in marriage, commitment to family needs, self-sacrifice); (6) self-maintenance/physical fitness (healthful habits, self-discipline in use of alcohol, drugs, tobacco); (7) mature values (purpose for living, having principles and ideals); and (8) forgiveness (making restitution, forgiving

others) (Jensen and Bergin 1988, p. 292). They found a high level of consensus among the practitioners. Many of these values are consistent with the six approaches identified by Jahoda in 1958.

### ILLUSIONS AND POSITIVE MENTAL HEALTH

The validity of one of the components of positive mental health has been questioned (Snyder 1989). Is accurate reality perception the hallmark of positive mental health? Taylor and Brown argue that "certain illusions may be adaptive for mental health and well being" (1988, p. 193). They explain that mentally healthy persons have an unrealistic positive self-evaluation. Normal individuals are more aware of their strengths and less aware of their weaknesses, perceiving themselves as better than the average person and viewing themselves more positively than others see them.

Another illusion held by mentally healthy persons is an exaggerated sense of self-control. Taylor and Brown (1988) cite evidence that depressed individuals are more likely to have realistic perceptions of personal control than are nondepressed persons. Positive illusions of personal control over the environment, self-worth, and hopefulness about the future imply mental health, and these illusions enable people to function in an adaptive manner.

According to Taylor and Brown (1988), illusions can promote several criteria of mental health, including happiness or contentment, the ability to care for others, and the capacity for intellectually creative and productive work. While mentally healthy people learn from negative experiences, their illusions help them to cope with stresses and strains (Taylor et al. 1989).

Taylor and Brown conclude, "the mentally healthy person appears to have the enviable capacity to distort reality in a direction that enhances self-esteem, maintains beliefs in personal efficacy, and promotes an optimistic view of the future" (1988, p. 204).

### THE ROLE OF UNDERLYING VALUE ASSUMPTIONS

Jahoda (1980) argues that the definition of positive mental health depends upon underlying value

assumptions. Schwartz and Link explain, "What is viewed as good and functional is often dependent on who is doing the viewing and what value hierarchy is being applied" (1991, p. 240). The definition of positive mental health varies across societies. In addition, there may be variance across social groups within one society (e.g., social class, gender, race, and ethnicity). It is also the case that the definition of positive mental health may be a function of the situation.

Different societies have their own definition of positive mental health. In some societies, a mentally healthy individual is supposed to be autonomous, while in other societies, the mentally healthy person is expected to be compliant, conforming to particular rules imposed by others.

Variance across social groups within one society is illustrated by evidence that there are different standards of positive mental health for men and women. Broverman et al. (1981) report sex role stereotypes in the clinical judgments of mental health among seventy-nine psychotherapists. The therapists were asked to identify the characteristics that portrayed healthy, mature, and socially competent adults. Broverman et al. found "that healthy women differ from healthy men by being more submissive; less independent; less adventurous; more easily influenced; less aggressive; less competitive" (1981, p. 92).

The definition of positive mental health is situational. While a particular behavior is mentally healthy in one situation, it may represent mentally ill behavior in another situation. For example, behavior in a life-threatening situation may be defined as adaptive, given the stresses of the environment. In normal everyday life the same behavior may be defined as bizarre. This observation leads Foote and Cottrell to ask, "What are the psychologically relevant attributes of an environment which permit the manifestations of psychologically healthy behavior?" (1959, p. 44).

Finally, according to Jahoda (1988), the definition of positive mental health is also influenced by the following four assumptions: (1) the criteria for judging health and illness are debatable; (2) neither mental illness nor mental health can be defined by the absence of the other; (3) there are degrees of mental health; and (4) a low level of mental health is not synonymous with mental illness.

**NORMALITY OR POSITIVE MENTAL HEALTH?**

There is continued debate on the definition of and relationship between normality and positive mental health. There is general agreement among researchers who study normatology that the definition of psychopathology is more precise than the definition of normal behavior (Offer and Sabshin 1991; Strack and Lorr 1997). While psychiatrists have developed sets of very specific criteria for defining mental illness (American Psychiatric Association 1994), there are no set criteria for defining either normal behavior or positive mental health. As is the case for the definition of positive mental health, the determination of normality varies across societies, subgroups within one society, and situations. What is normal (or positive mental health) in one time and place is abnormal (or mentally ill) in another. Both concepts are fluid in nature. Additional research is needed to understand how the definitions of positive mental health and normality are created. What are the underlying assumptions of these definitions, and how do they vary across societies and situations?

While normality implies the absence of psychopathology, positive mental health goes *beyond* normality. Researchers agree that positive mental health is more than the absence of mental illness; it represents the enhancement of human potential. Although a person may not have symptoms of mental illness, he or she may not have positive mental health, especially in the absence of self-confidence, self-actualization, integration, autonomy, reality perception, and environmental mastery. To have positive mental health implies fulfilling one's potential to the fullest. One might argue that positive mental health is a moving target, representing goals that are established, evaluated, and then revised as a person's circumstance change. The concept of positive mental health is utopian. While many strive to achieve positive mental health, only a few fulfill the goal of maximizing their potential.

REFERENCES

American Psychiatric Association 1994 *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. Washington: American Psychiatric Association.

Broverman, Inge K., S. R. Vogel, D. M. Broverman, F. E. Clarkson, P. S. Rosenkranz 1981 "Sex-Role Stereotypes and Clinical Judgments of Mental Health." In Elizabeth Howell and Marjorie Bayes, eds., *Women and Mental Health*. New York: Basic Books.

Foote, Nelson N., and Leonard S. Cottrell, Jr. 1959 *Identity and Interpersonal Competence: A New Direction in Family Research*. Chicago: University of Chicago Press.

Jahoda, Marie 1958 *Current Concepts of Positive Mental Health*. New York: Basic Books.

— 1980 *Current Concepts of Positive Mental Health*, rev. ed. New York: Arno.

— 1988 "Economic Recession and Mental Health: Some Conceptual Issues." *Journal of Social Issues* 44:13-23.

Jensen, Jay P., and Allen E. Bergin 1988 "Mental Health Values of Professional Therapists: A National Interdisciplinary Survey." *Professional Psychology: Research and Practice* 19:290-297.

Jourard, S. M., and T. Landsman 1980 *Healthy Personality: An Approach from the Viewpoint of Humanistic Psychology*, 4th ed. New York: Macmillan.

Maslow, Abraham H. 1954 *Motivation and Personality*. New York: Harper and Row.

Offer, Daniel, and Melvin Sabshin (eds.) 1991 *The Diversity of Normal Behavior: Further Contributions to Normatology*. New York: Basic Books.

Schwartz, Sharon, and Bruce G. Link 1991 "Sociological Perspectives on Mental Health: An Integrative Approach." In Daniel Offer and Melvin Sabshin, eds., *The Diversity of Normal Behavior: Further Contributions to Normatology*. New York: Basic Books.

Snyder, C. R. 1989 "Reality Negotiation: From Excuses to Hope and Beyond." *Journal of Social and Clinical Psychology* 8:130-157.

Strack, Stephen, and Maurice Lorr 1997 "Invited Essay: The Challenge of Differentiating Normal and Disordered Personality." *Journal of Personality Disorders* 11:105-122.

Taylor, Shelley, and Jonathan Brown 1988 "Illusion and Well Being: A Social Psychological Perspective on Mental Health." *Psychological Bulletin* 103:193-210.

Taylor, Shelley, Rebecca L. Collins, Laurie A. Skokan, and Lisa G. Aspinwall 1989 "Maintaining Positive Illusions in the Face of Negative Information: Getting the Facts without Letting Them Get to You." *Journal of Social and Clinical Psychology* 8:114-129.

JANET HANKIN

## POSITIVISM

The concept of “positivism” was originally used to denote the scientific study of social phenomena, but today the term *positivism* has become vague. Most often, it is used as a pejorative smear for certain kinds of intellectual activity in the social sciences, sociology in particular. Most frequently, at least within sociology, positivism is associated with such undesirable states as “raw empiricism,” “mindless quantification,” “antihumanism,” “legitimation of the status quo,” and “scientific pretentiousness.” With few exceptions (e.g., Turner 1985), sociologists are unwilling to label themselves “positivists.” Yet, the titular founder of sociology—Auguste Comte—used this label as a rallying cry for developing formal and abstract theory that could still be used to remake society; so, the current use of the term does not correspond to its original meaning. If anything, the term connotes almost the exact opposite of Comte’s vision (1830–1842). It is proper, therefore, to review Comte’s original conception of positivism and its use in early sociology, and then we can discover how and why the meaning of positivism changed.

In *Cours de philosophie positive*, Comte began by asserting that “the first characteristic of Positive Philosophy is that it regards all phenomena as subject to natural *Laws*” (1830–1842, p. 5). Moreover, he emphasized that “research into what are called *causes*, whether first or final,” is “in vain” (1830–1842, p. 6); and by the time he was well into *Cours de philosophie positive*, he stressed that a “great hindrance to the use of observation is the empiricism which is introduced into it by those who . . . would interdict the use of any theory whatever” because “no real observation of any kind of phenomena is possible, except in as far as it is first directed, and finally interpreted, by some theory” (1830–1842, p. 242). Rather, the goal of positivistic sociology is to “pursue an accurate discovery of . . . Laws, with a view to reducing them to the smallest possible number,” and “our real business is to analyze accurately the circumstance of phenomena, to connote them by natural relations of succession and resemblance” (1830–1842, p. 6). Comte’s exemplar for this advocacy was Newton’s law of gravitation, an affirmation of his early preference to label sociology as “social physics.” Moreover, such laws were to be used to

reconstruct society; and while Comte went off the deep end on this point, proclaiming himself, late in his career, to be the “high priest of humanity” (Comte 1851–1854), it is difficult to see Comte’s positivism as antihumanistic, as conservative, or as legitimating the status quo.

How, then, did Comte get turned on his head? The answer to this question cannot be found in nineteenth-century sociology, for the most positivistic sociologists of this period—Herbert Spencer (1874–1896) and Émile Durkheim ([1893] 1947; [1895] 1934)—could hardly be accused of “raw” and “mindless” empiricism, nor could they in the context of their times be considered antihumanistic, conservative, and apologists for the status quo (the label “conservative” for these thinkers is imposed retrospectively, through the refraction of contemporary eyeglasses). Moreover, early American sociologists—Albion Small, Frank Lester Ward, Robert Park, William Graham Sumner, and even the father of statistical methods and empiricism in American sociology, Franklin Giddings—all advocated Comtean and Spencerian positivism before World War I. Thus, the answer to this question is to be found in the natural sciences, particularly in a group of scientist-philosophers who are sometimes grouped under the rubric “the Vienna Circle,” despite the fact that several intellectual generations of very different thinkers were part of this circle.

Before the “circle” was evident, the nature of the issues was anticipated by Ernst Mach (1893), who argued that the best theory employs a minimum of variables and does not speculate on unobservable processes and forces. Mach emphasized reliance on immediate sense data, rejecting all speculation about causes and mechanisms to explain observed relations among variables. Indeed, he rejected all conceptions of the universe as being regulated by “natural laws” and insisted that theory represent mathematical descriptions of relations among observable variables. Although Mach was not a member of the Vienna Circle, his ideas framed the issues for those who are more closely identified with this group. Yet, his ideas did not dictate their resolution. Many in the Vienna Circle were concerned primarily with logic and systems of formal thought, almost to the exclusion of observation (or, at least, to the point of subordinating it to their primary concerns). A split thus developed in the Vienna Circle over the relative

emphasis on empirical observation and systems of logic; a radical faction emphasized that truth can be “measured solely by logical coherence of statements” (which had been reduced to mathematics), whereas a more moderate group insisted that there is a “material truth of observation” supplementing “formal truths” (Johnston 1983, p. 189). Karl Popper, who was a somewhat marginal figure in the Vienna Circle of the 1930s, is perhaps the best-known mediator of this split, for he clearly tried to keep the two points of emphasis together. But even here the reconciliation is somewhat negative (Popper 1959, 1969): A formal theory can never be proved, only disproved; and so, data are to be used to mount assaults on abstract theories from which empirical hypotheses and predictions are formally “deduced.”

Why did the philosopher-scientists in the Vienna Circle have any impact on sociology, especially American sociology? In Europe, of course, sociology had always been firmly anchored in philosophy, but in American sociology during the 1920s and 1930s, the rise of quantitative sociology was accelerating as the students of Franklin Giddings assumed key positions in academia and as Comtean and Spencerian sociology became a distant memory. (It should be noted, however, that Marx, Weber, and Durkheim had yet to have much impact on American sociology in the late 1920s or early 1930s.) But American sociology was concerned with its status as science and, hence, was receptive to philosophical arguments that could legitimate its scientific aspirations (Turner and Turner 1990). Mach was appealing because his advocacy legitimated statistical analysis of empirical regularities as variables; and Popper was to win converts with his uneasy reconciliation of observation and abstract theory. Both legitimated variable analyses; and for American sociologists in the 1930s and later from the 1940s through the early 1960s, this meant sampling, scaling, statistically aggregating, and analyzing empirical “observations.” Members of the Vienna circle had even developed an appealing terminology, *logical positivism*, to describe this relation between theory (abstract statements organized by a formal calculus) and research (quantitative data for testing hypotheses logically deduced from abstract statements). The wartime migration of key figures in the late Vienna Circle to the United States no doubt increased their impact on the social sciences in the United States (despite the

fact that the “logical” part of this new label for “positivism” was redundant in Comte’s original formulation). But logical positivism legitimated American empiricism in this sense: The quantitative data could be used to “test” theories, and so it was important to improve upon methods of gathering data and analyzing methodologies in order to realize this lofty goal. Along the way, the connection of theory and research was mysteriously lost, and positivism became increasingly associated with empiricism and quantification, *per se*.

There was a brief and highly visible effort, reaching a peak in the late 1960s and early 1970s, to revive the “logical” side of positivism by explaining to sociologists the process of “theory construction.” Indeed, numerous texts on theory construction were produced (e.g., Zetterberg 1965; Dubin 1969; Blalock 1969; Reynolds 1971; Gibbs 1972; Hage 1972), but the somewhat mechanical, cookbook quality of these texts won few converts, and so the empiricist connotations of positivism were never successfully reconnected to abstract theory. Even the rather odd academic alliance of functional theory with quantitative sociology—for example, Merton and Lazarsfeld at Columbia and Parsons and Stouffer at Harvard—was unsuccessful in merging theory and research, once again leaving positivism to denote quantitative research divorced from theory.

Other intellectual events, anticipated by various figures of the Vienna Circle, created a new skepticism and cynicism about the capacity to develop “objective” science, especially social science. This skepticism stressed the arbitrary nature of symbols and signs and hence their capacity to represent and denote the universe independently of the context in which such signs are produced and used. Such thinking was supplemented by Kuhn’s landmark work (1970) and by the sociology of science’s emphasis (e.g., Whitley 1984) on the politico-organizational dynamics distorting the idealized theory-data connection as advocated by Popper (1969). Out of all this ferment, a new label increasingly began to appear: *postpositivism*. This label appears to mean somewhat different things to varying audiences, but it connotes that Comte’s original vision and Popper’s effort to sustain the connection between empirical observations and theory are things of the past—just as “rationalism” and “modernity” are giving away to “postmodernism.” Thus, one hears about a



“postpositivist” philosophy of science, which, despite the vagueness and diversity of usages for this label, is intended to signal the death of positivism. Curiously, this postpositivism is meant as an obituary for the older Comtean positivism or its resurrection as logical positivism by the Vienna Circle, where abstract logic and observation were more happily joined together.

The result is that the term “positivism” no longer has a clear referent, but it is evident that, for many, being a positivist is not a good thing. It is unlikely, then, that “positivism” will ever be an unambiguous and neutral term for sociological activity revolving around the formulation and testing of theory and the use of plausible theories for social engineering (or in more muted form, for “sociological practice”). Other labels are likely to be employed in light of the negative connotations of positivism in an intellectual climate dominated by “post-isms.”

Despite this apparent eclipse of positivism by various post-isms, positivistic sociology remains a vibrant activity, albeit by other names. Because of the pejorative use of the label “positivism,” few are willing to embrace it, but many practice positivistic sociology. What, then, are the main tenets of positivism? This question can be answered under ten general points.

First, positivism assumes that there is a “real world” that can be studied scientifically. The social world is not an illusion, or a total fabrication of sociologists’ imaginations. It is there; it has properties amenable to investigation.

Second, positivism assumes that there are fundamental properties of the social universe that are always operative when people act, interact, and organize. While the properties can manifest themselves in a wide variety of forms in varying contexts, they nonetheless exist; and they are what drive the dynamics of the social universe. The goal of positivism is to uncover these fundamental properties, to see how they work, to develop theories on their operation, and to test these theories with systematically collected data.

Third, the theories developed by positivists should strive for some degree of formality. The making of formal statements need not invoke mathematics or some other system of formal argument; rather, all that is necessary is that concepts

denoting processes be explicitly defined and that relations among concepts be stated clearly. These goals can be met with ordinary language, although if they can be converted into mathematics, this is seen by most positivists as useful though not absolutely necessary.

Fourth, in defining concepts formally, these definitions should denote aspects of the social universe such that what is encompassed by the concept is clear and, equally important, what is not is also explicit. In stating relations among concepts denoting fundamental properties of the social world, these relations can be stated in three basic ways. One is functional (in the mathematical sense), whereby variation in one concept is seen to be related to another (e.g., the level of differentiation in a population is a positive function of its size). A second way to state relations is through analytical models that specify the direct, indirect, and reverse causal effects among those forces of the universe that are seen as connected. A third procedure is historical in which events at earlier points in time are seen to cause directly, or in combination with other events, an outcome. A fourth, though less desirable (and at best, preliminary), procedure is to find the place of particular forces in an abstract category system that juxtaposes phenomena (e.g., the periodic table in chemistry or Parsonian four-functions analysis).

Fifth, the goal of all positivistic theories statements is parsimony. Reducing theories to their simplest form is always desired, whether this be a simple equation, an analytical model, a historical sequence of cause, or even a simple set of categories.

Sixth, at the same time that statements move toward parsimony, they should become ever more abstract and should seek to explain as large a portion of reality as is possible. The goal is always to explain as much of the social universe with as few principles and models as can do justice to the dynamics of the social world.

Seventh, all theoretical statements must be testable, at least in principle. Some statements can be tested directly with existing methodologies; others must be transformed (e.g., from deductions to hypotheses); and still others may have to wait for new methodologies or for specific classes of events to occur. The critical criterion is that theories be testable, now or in the future.

They must suggest by their formulation ways of operationalization.

Eighth, theories can be tested by all relevant methods: historical, comparative, experimental, survey, observational, and even simulational. No one method identifies positivism; all are useful in assessing the plausibility of theories.

Ninth, tests must always be used to assess the plausibility of theories. When tests do not support the theory, the theory must be rejected and/or revised.

Tenth, theories that remain plausible constitute, for the time being, the best explanations of the social universe. And the more theories remain plausible, the more they are made parsimonious, and the more new theories are developed to explain what has not yet been explained, the more knowledge of the nature and operative dynamics of the social universe accumulates.

## REFERENCES

- Blalock, Hubert M., Jr. 1969 *Theory Construction: From Verbal to Mathematical Formulations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Comte, Auguste 1830–1842 *Cours de philosophie positive: Les Préliminaires généraux et la philosophie mathématique*. Paris: Bachelier.
- 1851–1854 *Système de politique: ou, traite de sociologie, instituant la religion de l'humanité*. Paris: L. Mathias.
- Dubin, Robert 1969 *Theory Building*. New York: Free Press.
- Durkheim, Emile (1895) 1934 *The Rules of the Sociological Method*. New York: Free Press.
- (1893) 1947 *The Division of Labor in Society*. New York: Free Press.
- Gibbs, Jack 1972 *Sociological Theory Construction*. Hinsdale, Ill.: Dryden Press.
- Hage, Jerald 1972 *Techniques and Problems of Theory Construction in Sociology*. New York: John Wiley.
- Johnston, William M. 1983 *The Austrian Mind: An Intellectual and Social History 1848–1938*. Berkeley: University of California Press.
- Kuhn, Thomas 1970 *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- Mach, Ernst 1893 *The Science of Mechanics*, trans. T. J. McCormack. La Salle, Ill: Open Court.
- Popper, Karl 1969 *Conjectures and Refutations*. London: Routledge and Kegan Paul.

— 1959 *The Logic of Scientific Discovery*. London: Hutchinson.

Reynolds, Paul Davidson 1971 *A Primer in Theory Construction*. Indianapolis: Bobbs-Merrill.

Spencer, Herbert 1874–1896 *The Principles of Sociology*, 3 vols. New York: Appleton.

Turner, Jonathan H. 1985 “In Defense of Positivism.” *Sociological Theory* 3:24–30.

Turner, Stephen Park, and Jonathan H. Turner 1990 *The Impossible Science: An Institutional Analysis of American Sociology*. Newbury Park, Calif.: Sage.

Whitley, Richard 1984 *The Intellectual and Social Organization of the Sciences*. Oxford: Clarendon Press.

Zetterberg, Hans L. 1965 *On Theory and Verification in Sociology*, 3rd ed. New York: Bedminster Press.

JONATHAN H. TURNER

## POSTINDUSTRIAL SOCIETY

Postindustrial society is a concept used to characterize the structure, dynamics, and possible future of advanced industrial societies. Like the more recent concepts of postmodern and radically modern society, the concept of postindustrial society attempts to make sense of the substantial changes experienced by advanced industrial societies since the end of World War II. In providing a depiction of the character and future of these societies, analyses usually attempt to shape the futures they describe. Such efforts illustrate an awareness among sociologists of the “reflexive” character of much social science—that is, an awareness that analyses of society become elements of the social world that have the potential to shape the future.

Social analysts have long been aware of the potential effects of their ideas, at times engaging in work precisely because it may have an effect on the future through such mechanisms as social engineering, social movements, and the application of technology. The nineteenth-century theorists of “industrial” society, like postindustrial theorists a century later, tried to make sense of the diverse changes surrounding them, oftentimes in an effort to help shape the future. In the early nineteenth century, Claude-Henri de Saint-Simon attempted to provide an image of what was then a barely emerging industrial society, an image he hoped

would enable scientists and industrialists to see the crucial roles they were to play in a society consciously directed by scientific knowledge. Similarly, in the mid-nineteenth century, Karl Marx and Fredrich Engels were outlining the characteristics of industrial capitalism, including the revolutionary role of the proletariat, when the factory system in England was still in its infancy, and the workforce “was still heavily concentrated in agriculture and domestic service, with the remainder mostly employed in the old craft industries.” (Kumar 1978, p. 133) Although theorists of industrial society may not have been fully correct in all areas, their recognition of some major correlates of industrialization provided a surprisingly accurate portrait of a form of society qualitatively different from prior modes of human organization. The emerging industrial society was one that increasingly utilized technology and machinery for work; a society with substantial increases in communications, transportation, markets, and income; a society within which urbanism became a way of life, and the division of labor became increasingly complex; a society marked by an increasing role for the state, and bureaucratization in government and the economy; and a society marked by increasing secularization and rationalization.

The concept of postindustrial society indicates significant changes in some of these central characteristics of industrial society. In probably the earliest use of the concept, the Guild Socialist Arthur Penty (1917) called for development of a postindustrial state that reversed key characteristics of industrial society. Penty called for development of a mode of organization reflecting the artisan workshop, in which work, leisure, and family would be once again brought together. Although Penty may have been the first to use the concept of postindustrial society, it was not until the 1960s that the concept took on its present character, focusing on quantitative changes separating postindustrial from industrial society. Interest in the future, and in postindustrial society, developed at this time as a response to the dramatic changes occurring in advanced industrial societies. These changes included the technological and organizational expansion accompanying economic growth in the post-World War II era, the expansion of the welfare state and an increased concern over the dark side of industrialism. An array of terms emerged to characterize the social

milieu of advanced industrial societies, including the technocratic era (Brzezinski 1970), service class society (Dahrendorf 1967), personal service society (Halmos 1970), postscarcity society (Bookchin 1971), posteconomic society (Kahn and Wiener 1967), knowledge society (Drucker 1969), postmodern society (Etzioni 1968), and postindustrial society (Touraine 1971; Richta et al. 1969). Although differing in focus, the analyses overlapped considerably with Daniel Bell’s work on postindustrial society (1973, 1989), which has been considered the best known and most complete analysis (Kumar 1978). Thus, the following effort to characterize postindustrial society uses Bell’s analysis as the organizing framework. This is followed by an examination of the related concepts of postmodern and radically modern society.

### CHARACTERISTICS OF POSTINDUSTRIAL SOCIETY

For analytical purposes Bell divides society into three parts: social structure, culture, and the polity. The concept of postindustrial society focuses primarily on changes in social structure, that is, changes in the economy, in technology, and in occupational structure. Although the social structure, polity, and culture may influence one another, it is not assumed there is a harmonious relation between the three. In fact, changes in any one may pose problems for the others (Bell 1976).

Bell’s depiction of postindustrial society (1973, 1989) focuses on two dimensions: the centrality of codified theoretical knowledge, and the expansion of the service sector, especially professional and human services. The centrality of theoretical knowledge is viewed as the most important dimension, or axial principle, of postindustrial society. The institutions that most embody this dimension are the university and the research institute. In postindustrial society major innovations are more a product of the application of theoretical knowledge (e.g., Albert Einstein’s discussion of the photoelectric effect for the development of lasers, holography, photonics), than the product of persons skilled in the use of equipment (e.g., Alexander Graham Bell and Thomas Edison). The use of theoretical knowledge increases the importance of advanced education, reflected in substantial enrollments in colleges and universities, as well as substantial numbers of scientists, engineers, and

persons with advanced degrees. For example, between 1939 and 1964, as the United States moved from an industrial to a postindustrial society, the number of scientists and engineers increased over fivefold, from 263,000 to 1,475,000 (Bell 1973). Relatedly, the percentage of 20- to 24-year-olds studying for college degrees went from 4 percent in 1900 to 15 percent in 1940, 26 percent in 1950, and 34 percent in 1960 (estimated from Bell 1973, table 3-4). Related estimates show tertiary enrollment rates going from 32 percent in 1960 to 56 percent in 1980 and 81 percent in 1993 (World Bank 1980, 1997). Accompanying the importance of science and theoretical knowledge is an occupational structure in which more persons are involved in services, with professional and helping services especially important. This includes increased employment in education, science, and engineering, which is a natural consequence of a society committed to science and education. It also includes expansion of the number of white-collar workers and professionals in government and the helping services, resulting from an expanded welfare state and increased attention to health care. Such trends are illustrated in the following data (Table 1) on employment trends in the percentage of the workforce in white-collar and professional occupations in the United States, and in seventeen developed countries (twelve western European countries, plus Canada, Australia, New Zealand, Japan, and Israel) (U.S. Bureau of the Census 1975, 1998; ILO 1965-1988, 1997). Also included are related data for agriculture, mining, and manufacturing for the United States.

Increased employment in white-collar and professional occupations occurs largely at the expense of agricultural employment and, to a lesser extent, of manufacturing employment. The economic dynamic behind such changes consists of shifts in relative demand toward services as disposable income increases (Clark 1960), and to the greater responsiveness of agriculture and manufacturing to technical innovation (Fuchs 1968). As incomes rise, the *percentage* of total income used for food and agricultural products declines, and the need for basic manufactured products is more easily met. Thus, relative demand for health, education, and an array of other services may increase. At the same time, technical innovations increase productivity in agriculture and manufacturing, lowering the demand for labor in these sectors. However,

technical change is less able to displace workers in services, even though technical advances aid productivity. This is most clearly seen in health care, where technical change may increase the services available and the need for personnel to provide new services. Thus, economic and technical developments shift relative consumer demand and labor toward services. Such economic dynamics point out that development of an increasingly service-oriented postindustrial economy does not mean agriculture and manufacturing are ignored (Cohen and Zysman 1987). It means that technology takes over much of what people formerly did, and shifts productive efforts and labor toward services.

Accompanying the rising importance of theoretical knowledge and the service sector are several other changes. These include changes in women's roles, especially increased participation in the formal labor force. Additionally, the class structure of postindustrial society comes to increasingly center around education and technical expertise, creating possible tensions between expertise and populist sentiments. New political issues and attitudes concerned with the environment and quality of life move onto the political agenda (Ingelhart 1977; Lipset 1976). Technocratic rule may begin to take hold in organizations, and confront the problem of rationalized means becoming ends. Corporations may come under pressure to take into account objectives other than profit maximization. Efforts at social planning increase, and confront the problem of establishing a rational calculus for maximizing benefits throughout the society. The society becomes more politicized and conflictive as citizenship expands and groups seek a place in the polity.

Such conditions provide the basis for changes in consciousness and cosmology, as individuals confront a world of information and expanding specialized knowledge. The world becomes less one in which the individual interacts with nature and machines, and more one of persons interacting with persons. The reciprocal consciousness of self and other becomes increasingly important in defining the world. The culture may come to show contradictions between values of self-restraint, discipline, and work as a calling, on the one hand, and emerging postindustrial values of consumption and the negation of traditional bourgeoisie life, on the other (Bell 1976).

POSTINDUSTRIAL SOCIETY

	1900	1940	1960	1970	1980	1990	1995
Percentage of White-Collar Occupations							
United States	17.6	31.1	40.1	44.8	48.0	59.9	61.8
Percentage of Professional Occupations							
United States	4.3	7.5	10.8	13.8	14.1	16.5	17.6
Seventeen Developed Countries			6.1	9.1	14.3	18.5	
Percentage of Agriculture, Mining, Manufacturing							
United States			31.4	26.5	21.3	19.2	

Table 1

The image of postindustrial society provided by Bell and other social analysts is one in which technological advances have made possible development of a society characterized by the expansion and use of theoretical knowledge, and the concomitant expansion of white-collar and especially professional employment. A commitment to human welfare and social planning facilitates expansion of the welfare state and human services, further increasing white-collar and professional employment. Yet, such a society is not without tensions and conflict, both within the social structure and between the social structure, polity, and culture. It is by pointing out such tensions and problems that postindustrial theorists attempt not only to describe the present and the future but also to shape the future. In one of the more recent contributions to the postindustrial literature, Fred Block is clear about the role social analysis plays in shaping the future. He points out that “[social theory] has real consequences, because individuals cannot do without some kind of conception of the type of society in which [they] live” (Block 1990, p. 2). Block’s book represents a return to the analysis of postindustrial society after a brief hiatus during the 1980s that Block attributes to the breakdown of mainstream and leftist social theory, and the reemergence of a tradition of economic liberalism quiescent since the Great Depression. Block’s analysis carries forward the focus of prior postindustrial theory on social structure, yet gives scant attention to the role of codified theoretical knowledge, the university and professional groups. Instead, Block looks at key aspects of the changing postindustrial economy of the United States in an effort to develop alternative possibilities for the future. He notes that a postindustrial economy utilizing advanced technology, and having a substantial service sector, would be most productive if it kept some distance from the dictates of classical

economic theory. Specifically, productivity is likely enhanced by: (1) relatively low levels of marketness, allowing for greater predictability and more accurate information; (2) labor relations emphasizing cooperation between labor and management, since skilled labor in technologically sophisticated industries and the professions has substantial knowledge and is expensive to replace; (3) treating capital not merely as a physical asset, but as part of a productive process that includes the organization of persons working with capital; (4) developing measures of economic well-being that incorporate positive and negative utilities currently excluded from measures like gross national product (e.g., child care and pollution as positive and negative utilities); and (5) using various hybrid forms of market, state, and other regulatory mechanisms to enhance quality growth.

Block’s work provides a useful extension of postindustrial theory into the more traditional economic domains of labor, capital, and measured economic output. The book may represent a revival of analyses of postindustrial society, perhaps under the alternative concept of postmodern society. However, before considering the notion of postmodern society, some criticisms of the concept of postindustrial society need to be mentioned. Much of the critical literature focuses on Bell’s work, since it is viewed as one of the best expressions of postindustrial theory. One criticism of postindustrial theory is that it overemphasizes the role of theoretical knowledge in decision making. Although critics acknowledge that formal knowledge is more important than ever before, they contend that technical experts and scientific knowledge have not come to play the central role in decision making in government or corporations that postindustrial theorists said they would. Within government, political dynamics of the industrial era persist, and within the corporation “the expert

and his knowledge are, for the most part, embedded in the corporate bureaucracy” (Cohen and Zysman 1987, p. 260). Also, critics point out that while there may be more scientists, more persons with advanced education, and more money spent on research and development, these may be only tenuously related to increases in the amount and effective use of theoretical knowledge (Kumar 1978).

Critics also question the attention given by postindustrial theorists to the service sector, and to white-collar and professional work. As with theoretical knowledge, critics agree the service sector has grown, yet question the focus of postindustrial theorists. Critics remind us that the service sector has always been a major segment of preindustrial and industrial society (Hartwell 1973), that most service employment is in low-skilled, low-paid work, and that increases in white-collar and professional employment are a consequence of dynamics embedded in industrial society. Relatedly, most of the increases in white-collar and professional employment are in clerical positions and jobs like teaching and nursing, jobs that carry less autonomy and income than traditional professional occupations. Perhaps most important, critics argue that a focus on the service sector fails to adequately acknowledge the key role that manufacturing would play in a postindustrial economy, including the dependence of the service sector on a dynamic manufacturing sector (Cohen and Zysman 1987).

One response to these and other criticisms (e.g., see Ritzer 1989; Frankel 1987) would be to point out that postindustrial theory often acknowledged points made by critics, such as the fact that the political order could check trends in the economy; that the bulk of services are in low-skilled, low-paid labor; and that a sophisticated and dynamic manufacturing sector is important for an economy as large as that of the United States. Postindustrial theory merely chose to focus on other developments that may hold insights into the future of advanced industrial society. A related response to criticisms is to point out that critics have frequently acknowledged the validity of postindustrial claims of an increasing role for formal and theoretical knowledge, the increasing importance of technical developments in information processing, and the rise of a service sector with a good number of white-collar and professional workers. The points of contention between postindustrial theory and

its critics appear to center on the general portrait of society provided by postindustrial theorists: Is this society best seen as a new type of society, or as a logical extension of advancing industrialization (Kumar 1978; 1988; 1995; Ritzer 1989; Giddens 1990)? Also, what are the implications of this image of society, not only for describing the present and future, but for shaping both?

### POSTMODERN SOCIETY

Recent discussions of the character and future of advanced industrial societies have been framed within the concept of postmodern society. Although the variety of work dealing with postmodern society makes a clear definition difficult (Smart 1990; Kumar 1995), it seems reasonable to say postmodern society differs from postindustrial society in having the more nebulous reference point of a type of society coming after “modern” society. Whereas studies of postindustrial society understandably give substantial attention to technical and economic factors, studies of postmodern society broaden the focus to bring political and especially cultural phenomenon more to the center of analysis (Kumar 1995). This shift in focus helps explain why analyses of the cultural dynamics of postindustrial society are frequently considered in terms of the dynamics of postmodern society (e.g., Baudrillard 1983), and why studies of the economic characteristics of postmodern society look very much like studies of postindustrialism (e.g., Clegg 1990). For many authors, postmodernism is the culture of a postindustrial society (e.g., Lyotard 1984; Lash 1990; Jameson 1992; Mandel 1978).

Analyses of postmodern society may differ from postindustrial analyses not only in substantive focus, but also in basic presuppositions regarding the nature of social reality. While studies of postmodern society can reflect a standard social science framework, what has come to be called postmodern theory often includes a critique of science that undermines traditional social scientific inquiry (Turner 1998; Bogard 1990). Although the central tenants of postmodern theory are still unclear (Smart 1990; Ritzer 1997; Rosenau 1992), it is possible to note some major themes. Reflecting the influence of literary criticism, postmodern theory often views “social and cultural reality, and

the social sciences themselves [as] linguistic constructions" (Brown 1990, p. 188). Authoritative images of what is real are seen to emerge through rhetorical processes in which "people establish repertoires of categories by which certain aspects of what is to be the case are fixed, focused, or forbidden" (Brown 1990, p. 191). Postmodern theory critiques the notion that signs and symbols adequately capture reality, and posits that signs and symbols are the only reality we truly know (Brown 1987). The reality of the social world described by social scientists is viewed as inseparable from the discourse through which social scientists come to signify some descriptions of the world as more legitimate than others (Lemert 1990). The traditional hierarchy in which reality occupies a privileged status separate from the symbols used to describe it is broken down, or deconstructed. The deconstruction project of postmodern theory not only critiques the separation of reality from symbols describing it, but also critiques other hierarchies, such as those separating expert knowledge from common knowledge, and high culture from low. The project even moves to deconstruct the very possibility of a generalizing social science, and the idea of a social world (Foucault 1980; Baudrillard 1983). In such bold forms a postmodern sociology becomes a contradiction in terms, undermining its own basis for existence (J. H. Turner 1998; B. Turner 1990; Smart 1990; Bauman 1988). However, although a postmodern sociology may be difficult to establish, a sociology of postmodernism that examines postmodern society, and draws from postmodern theory, is possible.

Attempts to characterize postmodern society reflect the contributions of postmodern theory, as well as more conventional sociological approaches. From postmodern theory comes a view of postmodern society as a technologically sophisticated high-speed society, with access to vast amounts of information, and fascinated by consumer goods and media images. Mass consumption of goods and information is seen as facilitating a breakdown of hierarchies of taste, and development of an explicit populism. Technology and speed blur lines separating reality from simulation, as television, videos, movies, advertising, and computer models provide simulations of reality more real than real. People may realize this hyperreality is simulation, yet be fascinated by it, and come to

make it part of their lives, thus transforming much of reality into simulation (Baudrillard 1983). Under such conditions history comes to have little meaning, and the fast-paced present becomes increasingly important. People may attempt to come to grips with such a world, but the world comes to undermine major assumptions, or grand narratives, of rationality and progress, thus generating a sense of exhaustion (Lyotard 1984). An acute sense of self-consciousness and unease may develop, as "the current age stumbles upon the very transvaluation of Western values and vocabularies that Nietzsche urged more than a century ago" (Baker 1990, p. 232).

Views of postmodern society from more conventional social science frameworks echo some of these themes. For example, both Daniel Bell (1976) and Christopher Lasch (1979) point to the development of cultural themes in postindustrial society emphasizing consumption and personal gratification at the expense of themes emphasizing self-restraint, work, commitment, and a sense of historical connection and continuity. In a Marxist analysis, Fredric Jameson (1984) points to the loss of a sense of historical connection in the consumer-oriented world of late capitalism.

One productive approach to postmodern society starts from the assumption that a central process defining modern society is differentiation. Thus, the type of society coming after modern society is viewed as reversing this process. The process of de-differentiation is illustrated in the cultural arena in the conflation of high and popular culture noted above, as well as the general deconstruction project of postmodernist theory (see also Lash 1988). Within the economic arena de-differentiation appears in the reversal of the processes of bureaucratization and the division of labor characteristic of the assembly line (Clegg 1990). Such trends have been viewed as part of the changing character of production in the world economy since 1960, and have been examined under such terms as the second industrial divide (Piore and Sable 1984), or the emergence of disorganized capitalism (Lash and Urry 1987). The changes frequently include a shift to smaller organizations or subunits of organizations, a less formalized and more flexible division of labor, increased variation in the character of products, more decentralized managerial structures, and the use of computers. Such organizations are engaged

in services, information processing, or production using computer controlled flexible production techniques (Burris 1989; Heyderbrand 1989; Clegg 1990). The postbureaucratic character of these organizations is made possible by computers that do routine tasks, as well as by the need to respond flexibly to diverse clientele and markets, and the sophisticated capabilities of employees that operate complex manufacturing equipment and provide professional services.

Although postmodern production techniques and organizations are becoming more prominent and important in advanced industrial societies, they may express themselves in different ways and do not eclipse other forms of production and organization. Low-skilled tasks in manufacturing and services will likely compose a substantial segment of the workforce far into the future, creating the possibility of increased variation in skill and income in postmodern society. Also, postmodern organizations can vary substantially among themselves. For example, Clegg (1990) argues that Japan and Sweden stand as alternative expressions of postmodernist futures. Sweden provides the more optimistic democratic scenario with fairly broad representational rights of workers in organizations. Japan represents a less optimistic view with an enclave of privileged workers formed on exclusive principles of social identity, such as gender, ethnicity, and age.

### **RADICALLY MODERN SOCIETY**

Much as critics of the concept of postindustrial society pointed out that postindustrial trends were best seen as the logical extension of major characteristics of industrial society, so Anthony Giddens (1990) has questioned the notion that postmodern trends actually represent a break with modern society. Giddens acknowledges current trends of complexity and widespread change accompanied by lack of a clear sense of progress, and would also likely admit to trends of de-differentiation in some organizations. However, Giddens does not see this as constituting postmodernity. Instead he refers to such trends as characterizing radically modern societies, for these societies represent the logical extension of three essential characteristics of modernity. The first characteristic is an increase in complexity, as the major processes of industrialization, class formation, and rationalization proceed

apace and shape one another. The second key element of modernity is the separation of much of what humans experience as social reality from concrete instances of time and space. Thus, money represents an element of social reality that designates a mechanism of exchange not bound to any specific instance of exchange. Even more abstractly, experts and expert systems represent specialized knowledge that may be called upon for a variety of purposes at a number of times and places. Science is one of the clearer expressions of such knowledge, with its esoteric and changing images of the character of the natural and physical universe. As societies become more modern, humans are viewed as increasingly experiencing the world through such abstract, or "disembedded," categories as money, expertise, and science. Realization of the disembedded nature of social life helps facilitate development of the third characteristic of modernity, its reflexivity. As noted earlier, *reflexivity* refers to the fact that efforts and information used to understand social reality become part of reality, and thus help shape the present and the future. According to Giddens, awareness of reflexivity is not confined to social scientists, but is embedded in the nature of modernity. Modern societies generate theories of what they are, which become important constitutive elements that shape society, including the future character of theories. All three of the characteristics of modernity find expression in each of four interrelated institutional complexes: capitalism, the nation-state, the military, and industrialization. Thus, radically modern societies are those in which major institutions are highly complex, abstract or disembedded, and reflexive. Such societies will be more difficult for individuals to concretely understand; will require substantial amounts of "trust" in abstract and expert systems; and will generate high levels of self-consciousness, as persons seek to use personal resources to more reflexively construct personal identity and meaning. Thus, radically modern societies may produce the unease and sense of powerlessness that some postmodern theorists see in postmodern societies. By avoiding the concept of postmodern society, Giddens is able to illustrate how current trends are part of the logic of modernity, and is able to provide a view of the future more optimistic than the view of some postmodern theorists by underscoring the reflexive and transformative power of ideas and action.



### POSTMODERN SOCIETY AND THE WORLD SYSTEM

Most analyses of postindustrial, postmodern, and radically modern societies carry with them a view of societal change occurring within an increasingly interconnected world system. Giddens is quite explicit in this regard, pointing out how the complexity, abstractness, and reflexive character of modernity moves beyond the nation-state to become a world-level phenomenon (1990). This is most clearly expressed in the development of a world capitalist system, a nation-state system, a world military order, and an international division of labor. Although Giddens does not systematically consider the dynamics of these aspects of the world system in discussing radical modernity, his conceptualization is a useful reminder of the possible utility of considering multiple dimensions of the world system in efforts to understand postmodern societies. As the review of postindustrial and postmodern literature suggested, most analyses consider the role of the world system in terms of economic exchanges (Bell 1989), new technologies, and new modes of economic organization (Piore and Sable 1984; Lash and Urry 1987; Clegg 1990).

Recently, analysts have drawn increasing attention to cultural dimensions of the world system (e.g., Featherstone 1990; Boli and Thomas 1997), pointing out how themes associated with modern, postmodern, and postindustrial society have become elements of an emerging world culture (Meyer et al. 1997; Meyer 1991; Smith 1990; Robertson 1989, 1990). Giddens touched on these themes when he noted how the abstract and reflexive character of modernity has become a global phenomenon, making theories of modernity part of contemporary culture. Meyer (1980, 1991) and others (Meyer et al. 1997) also note the role that ideologies of modernity and postindustrialism may play in the contemporary world, pointing out how rationalized conceptions of modern society, including social scientific discourse, have become part of a contemporary world cultural framework. For example, modern images of society that view economic development, national integration, and personal and societal progress as the product of individuals with specialized knowledge are said to have become a world-level assumption, and generated expansion of educational enrollments between 1950 and 1970 (Meyer et al. 1977; Boli and

Ramirez 1986; Fiala and Lanford 1987). Relatedly, postindustrial images of the role of specialized knowledge in society are viewed as an important factor affecting the worldwide expansion of higher education and professional employment in the post-World War II era (Meyer and Hannan 1979). Such images may account for part of the increase from 3.7 percent to 9.2 percent between 1960 and 1990 in the percentage of the labor force employed in professional occupations among twenty-two developing countries, a significant increase considering the substantial population and labor-force growth during this period. This contrasts with the increase from 6.1 percent to 18.5 percent for the seventeen developed countries reported earlier, and with the increase from 10.8 percent to 16.5 percent for the United States. While economic growth may account for some of these increases in professional employment, the substantial expansion for developing countries, especially compared to the United States, indicates other variables likely play a role. It seems plausible that postindustrial images regarding the importance of specialized knowledge and personnel may have an effect.

### CONCLUSION

Efforts to understand the character and future of advanced industrial societies in the latter half of the twentieth century have been done under the rubric of at least three major concepts. The concept of postindustrial society focused attention on social structural or economic dimensions of society, especially the changing character of technology, knowledge, occupations, and the market. The concept of postmodern society continued to bring attention to structural and economic aspects of society, yet gave increased attention to political, cultural, and psychological dimensions, at times with an explicit critique of standard social science methodology. The concept of radically modern society underscored the links between contemporary changes and the basic dynamics of modernity, while introducing the useful idea that radically modern societies are characterized by high levels of abstractness and reflexivity. Each of the three concepts acknowledges that changes are occurring within an increasingly interconnected world system.

The current review clearly illustrates the reflexive character of much social science. In

postindustrial theory this was largely a recognition that formal scientific knowledge may shape the future. In postmodern theory, and especially in radically modern theory, this reflexivity also became part of the culture itself, affecting the consciousness of much of the population. This creates a vastly more complex image of society, with a diverse array of possible futures, and may account for the unease postmodern theory sees as characterizing postmodern society.

While work on postmodern and radically modern society offers insightful ideas, most empirical research has been done on postindustrial society, and economic aspects of postmodern society. Future studies should attempt to examine some of the major elements of postmodern and radically modern approaches, as well as continue to clarify areas of ambiguity within postindustrial theory. For example, within the postindustrial framework, attention should be given to understanding the influence of various segments of a knowledge class on specific institutions in society, and to trying to explain national variation in the expansion of various types of service employment.

Within the postmodernist framework, it would be informative to assess claims of pessimism, self-consciousness, and declines of grand narratives, noting their variation across social groups and societies. Examination of variation across time and space in the way individuals, communities, and societies experience exposure to media might help clarify the provocative hypothesis of the increasing salience of simulation in social life. Efforts to rethink extant research in terms of the simulation hypothesis could also prove informative.

Drawing from the view of radically modern societies, it would be useful to examine the idea that modern cultures increasingly incorporate reflexive elements. Analysis of the changing role of evaluation studies, news presentations and commentaries, and economic projections and predictions could help clarify the developing reflexive character of human societies.

Last, to assess hypotheses regarding the effects of world cultural themes, research could examine the effects of an ideology of specialized knowledge and expertise on expansion of professional employment. The data presented above suggest the plausibility of such an effect, yet do not offer firm support.

Issues such as those above are but a few of the avenues of research that may help provide a clearer image of the character and future of advanced industrial societies. In posing the issues, and in providing answers, social science will likely help shape the future it describes.

## REFERENCES

- Baker, Scott 1990 "Reflection, Doubt, and the Place of Rhetoric in Postmodern Social Theory." *Sociological Theory* 8:232-245.
- Baudrillard, Jean 1983 *Simulations*. New York: Semiotext(e).
- Bauman, Zygmunt 1988 "Is There a Postmodern Sociology?" *Theory, Culture and Society* 5:217-237.
- Bell, Daniel 1989 "The Third Technological Revolution." *Dissent* 36:164-176.
- 1976 *The Cultural Contradictions of Capitalism*. New York: Basic Books.
- 1973 *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. New York: Basic Books.
- Block, Fred 1990 *Postindustrial Possibilities: A Critique of Economic Discourse*. Berkeley: University of California Press.
- Bogard, William 1990 "Closing Down the Social: Baudrillard's Challenge to Contemporary Sociology." *Sociological Theory* 8:1-15.
- Boli, John, and Francisco Ramirez 1986 "World Culture and the Institutional Development of Mass Education." In John Richardson, ed., *Handbook of Theory and Research for the Sociology of Education*. Westport, Conn.: Greenwood.
- Boli, John, and George M. Thomas 1997 "World Culture in the World Polity: A Century of International Non-Governmental Organization." *American Sociological Review* 62:171-190.
- Bookchin, Murray 1971 *Post-Scarcity Anarchism*. Berkeley, Calif.: Ramparts Press.
- Brzezinski, Zbigniew 1970 *Between Two Ages: America's Role in the Technocratic Era*. New York: Viking Press.
- Brown, Richard Harvey 1990 "Rhetoric, Textuality, and the Postmodern Turn in Sociological Theory." *Sociological Theory* 8:188-197.
- 1987 *Society as Text: Essays on Rhetoric, Reason, and Reality*. Chicago: University of Chicago Press.
- Burris, Beverly 1989 "Technocratic Organization and Control." *Organization Studies* 10:1-22.
- Clark, Colin 1960 *The Conditions of Economic Progress*, 3rd ed. New York: Macmillan.

- Clegg, Stewart R. 1990 *Modern Organizations: Organization Studies in the Postmodern World*. Newbury Park, Calif.: Sage.
- Cohen, Stephen, and John Zysman 1987 *Manufacturing Matters: The Myth of the Post-Industrial Economy*. New York: Basic Books.
- Dahrendorf, Ralf 1967 *Society and Democracy in Germany*. Garden City, N.Y.: Doubleday.
- Drucker, Peter 1969 *The Age of Discontinuity: Guidelines to Our Changing Society*. New York: Harper and Row.
- Etzioni, Amitai 1968 *The Active Society*. New York: Free Press.
- Featherstone, Mike 1990 *Global Culture: Nationalism, Globalization and Modernity*. Newbury Park, Calif.: Sage.
- Fiala, Robert, and Audri Gordon Lanford 1987 "Educational Ideology and the World Educational Revolution 1950-1970." *Comparative Education Review* 31:315-332.
- Foucault, Michel 1980 *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*. Colin Gordon, ed. New York: Pantheon.
- Frankel, Boris 1987 *The Post-Industrial Utopians*. Great Britain: Polity Press.
- Fuchs, Victor 1968 *The Service Economy*. New York: Columbia University Press.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Stanford Calif.: Stanford University Press.
- Halmos, P. 1970 *The Personal Service Society*. Great Britain: Constable.
- Hartwell R. M. 1973 "The Service Revolution: The Growth of Services in the Modern Economy." In *The Industrial Revolution, 1700-1914*. Brighton, England: Harvester Press.
- Heyderbrand, Wolf 1989 "New Organizational Forms." *Work and Occupations* 16:323-357.
- Ingelhart, Ronald 1977 *The Silent Revolution: Changing Values and Political Styles Among Western Publics*. Princeton, N.J.: Princeton University Press.
- International Labor Organization (1965-1988) 1997 *Yearbook of Labor Statistics*. Geneva: ILO.
- Jameson, Fredric 1992 *Postmodernism, or the Cultural Logic of Late Capitalism*. London: Verso.
- 1984 "Postmodernism, or the Cultural Logic of Late Capitalism." *New Left Review* 146:59-92.
- Kahn, Herman, and Anthony Wiener 1967 *The Year 2000*. New York: Macmillan.
- Kumar, Krishan 1995 *From Postindustrial to Postmodern Society*. Cambridge: Basil Blackwell.
- 1988 *The Rise of Modern Society*. New York: Basil Blackwell.
- 1978 *Prophecy and Progress: The Sociology of Industrial and Post-Industrial Society*. London: Allen Lane.
- Lasch, Christopher 1979 *The Culture of Narcissism*. New York: Warner.
- Lash, Scott 1990 *Sociology of Postmodernism*. New York: Routledge.
- 1988 "Postmodernism as a Regime of Signification." *Theory, Culture and Society* 5:311-336.
- , and John Urry 1987 *The End of Organized Capitalism*. Madison: University of Wisconsin Press.
- Lemert, Charles 1990 "General Social Theory, Irony, Postmodernism." In Steven Seidman and David G. Wagner, eds., *Postmodernism and Social Theory*. Cambridge, Mass.: Blackwell.
- Lipset, Seymour 1976 *Rebellion in the University*. Chicago: University of Chicago Press.
- Lyotard, Jean-Francois 1984 *The Postmodern Condition*. Minneapolis: University of Minnesota Press.
- Mandel, Ernst 1978 *Late Capitalism*. London: Verso.
- Meyer, John 1991 "Concluding Commentary: The Evolution of Modern Stratification Systems." In David Grusky, ed., *Social Stratification: Class, Race and Gender in Sociological Perspective*. Boulder, Colo.: Westview.
- 1980 "The World Polity and the Authority of the Nation State." In Albert Bergesen, ed., *Studies of the Modern World System*. New York: Academic Press.
- , John Boli, George M. Thomas, and Francisco O. Ramirez 1997 "World Society and the Nation State." *American Journal of Sociology* 103:144-181.
- Meyer, John, and Michael Hannan 1979 "National Development in a Changing World System: An Overview." In *National Development and the World System: Educational, Economic, and Political Change, 1950-1970*. Chicago: University of Chicago Press.
- Meyer, John, Francisco Ramirez, Richard Rubinson, and John Boli-Bennett 1977 "The World Educational Revolution, 1950-1970." *Sociology of Education* 50:242-258.
- Penty, Arthur 1917 *Old Worlds for New: A Study of the Post-Industrial State*. London: Allen and Unwin.
- Piore, Michael, and Charles Sabel 1984 *The Second Industrial Divide: Possibilities for Prosperity*. New York: Basic Books.
- Richta, Radovan, et al. 1969 *Civilization at the Crossroads: Social and Human Implications of the Scientific and Technological Revolution*, 3rd ed. New York: International Arts and Sciences Press.

- Ritzer, George 1997 *Postmodern Social Theory*. New York: McGraw-Hill.
- 1989 “The Permanently New Economy: The Case for Reviving Economic Sociology.” *Work and Occupations* 16:243–272.
- Robertson, Roland 1990 “Mapping the Global Condition: Globalization as the Central Concept.” In Mike Featherston, ed., *Global Culture: Nationalism, Globalization and Modernity*. Newbury Park, Calif.: Sage.
- 1989 “Globalization, Politics and Religion.” In J. A. Beckford and T. Luckmann, eds., *The Changing Face of Religion*. Newbury Park, Calif.: Sage.
- Rosenau, Pauline Marie 1992 *Post-Modernism and the Social Sciences: Insights, Inroads, and Intrusions*. Princeton, N.J.: Princeton University Press.
- Smart, Barry 1990 “Modernity, Postmodernity and the Present.” In Bryan S. Turner, ed., *Theories of Modernity and Postmodernity*. Newbury Park, Calif.: Sage.
- Smith, Anthony 1990 “Towards a Global Culture?” In Mike Featherstone, ed., *Global Culture: Nationalism, Globalization and Modernity*. Newbury Park, Calif.: Sage.
- Touraine, Alain 1971 *The Post-Industrial Society*. New York: Random House.
- Turner, Bryan 1990 “Periodization and Politics in the Postmodern.” In Bryan S. Turner, ed., *Theories of Modernity and Postmodernity*. Newbury Park, Calif.: Sage.
- Turner, Jonathan H. 1998 *The Structure of Sociological Theory*, 6th ed. Belmont, Calif.: Wadsworth.
- U.S. Bureau of the Census 1998 *Statistical Abstract of the United States 1998: The National Data Book*. Washington: U.S. Government Printing Office.
- 1975 *Historical Statistics of the United States, Part I. Colonial Times to 1970*. Washington: U.S. Government Printing Office.
- World Bank 1997 *World Development Report, 1997*. New York: Oxford University Press.
- 1980 *World Development Report, 1980*. New York: Oxford University Press.

ROBERT FIALA

## POSTMODERNISM

In 1959, C. Wright Mills speculated that “the Modern Age is being succeeded by a post-modern period” in which assumptions about the coherence of the Enlightenment values of scientific rationality and political freedom were being challenged (1959, p. 166). Critical theorists had earlier speculated about how the revolutionary potential

of the urban laboring classes of the nineteenth century was co-opted by the shift to a postindustrial twentieth-century society, a society characterized by mass consumerism and war economies. The characteristics of postindustrial societies were explored more recently in Daniel Bell’s analysis of contemporary capitalism (1976). In the modern information societies, Bell argues that the class forces that drove nineteenth-century social change have been replaced by new processes. Under welfare capitalist states, scientists, technicians, managers, and bureaucrats formulate social tensions as administrative and technical issues based on political consensus. For Bell, postindustrial societies with their information bases hold the key to social harmony and the end of misery.

Alain Touraine’s view (1984) is different. He also stresses how class antagonism has changed under welfare capitalism, but for Touraine the information managers, guided by technological thinking, tend to “steer the entire social order toward the perfectly programmed society, the ultimate technocratic prison” (Baum 1990, p. 5). Touraine also writes that the categories of basic sociological analysis have become out of touch with changes in contemporary societies. In particular, he rejects the modernist supposition of evolutionary progress over time, culminating in what he calls “the impoverishing homogeneity of modern civilization” (Touraine 1984, p. 38), the sociologist’s conception of the national states as units of analysis, and the gradual expunging of the cultural diversity of traditional societies. The collapse of modernism arises from a failure of sociology to keep abreast of the developing autonomy of cultural products, the globalization of capital, and the rise of new forms of social control and of public resistance to them. In this analysis, “the crisis of modernity . . . is not all-encompassing . . . the crisis in modernity is not considered to be total or terminal but limited in scope, if deep” (Smart 1990, p. 408). Though Bell and Touraine have been associated with critiques of postindustrial societies, the meaning of “postmodernism” has become far more radical, and draws on earlier sources.

De Saussure’s “structural” theory of linguistics (1959) distinguished the “sign” from the “signified” and developed a science based on the discovery that symbolic systems might have formal properties that were unrelated to the meanings of

the objects they signified but that might hold across different systems of symbols. The system for describing, for example, the animal kingdom and the protocols for siting teepees might reflect the same logic—without any inherent equivalence to the things they described or organized (Giddens 1987). “Poststructuralism” deepened this disinterest in the signified objects by dismissing any link binding words/symbols to things. Things were only the hypostatizations of language. All analysis was “deconstruction,” or unmasking of phenomena in terms of their underlying rhetorical conventions. For poststructuralists, texts only pointed to other texts. The world was viewed as an intertwining of systems of representation without any derivation from or basis within “terra firma.” Everything became text—including violent speech and violent symbolic acts such as incest, war, or spousal abuse.

The postmodern twist is the application of the linguistic implications of poststructuralism to the three core principles connecting contemporary civilizations with the project of the Enlightenment: scientific knowledge (Truth), aesthetics (Beauty), and morality (the Good). The Enlightenment project—“modernism”—refers to the rise of the Age of Reason. It was characterized by the gradual shift away from religious sensibilities and to scientific objectivity, the rational exploitation of nature for human needs, the perspectival representation of nature in art and humanistic truth in fiction, and the struggle for a humane society. Modernism was realist in its epistemology and progressive in its politics. Truth could be attained—particularly with scientific advances. Beauty could be distinguished from trash. Humanism could nurture moral conduct and decent conditions of life. And history was purposive and progressive.

Postmodernism is the sensibility that arises when the credibility of these “master narratives” is questioned. The postmodern period, as described by Lyotard (1984) and Baudrillard (1983), is the one we now confront. Though it is often dated as a creature of the post-World War II period, it is thought only to have become generalized with mass consumerism in the age of electronics and to have been initiated by dramatic changes in contemporary capitalism. Capitalism has ushered in global communication and exchange, and has created a self-sustaining cybernetic system that almost

completely transcends the ability of individual governments to control their directions and objectives. As a result, postmodernity is sometimes referred to as posthistorical society in the sense that the mission or sense of purpose that guided nation-states in the past through events such as the French Revolution has been overtaken by global consumerism. The view of history as a struggle for the gradual liberation of humanity and progressive evolution of more humane societies is dismissed by Lyotard as mythic. In addition, the transcendence of communities by electronic representations makes the idea of “the social” purely illusory (Bogard 1990). The objectivity of societies, nation-states, communities, and history is viewed only as “narratives” or “simulacra” (Baudrillard 1983). Referents disappear in favor of a world of simulations, models, performatives, and codes—in short, “information” (broadly conceived), which becomes the predominant phenomenon of exchange among the masses through the mass media.

One of the important theoretical aspects of the postmodernist position is its explicit rejection of Marxism in general and Habermas’ theorizing about creating a rational society through communicative competence. Where Habermas (1973, p. 105) speculates that a rational community might be able to “arrive at the conviction that in the given circumstances the proposed norms are ‘right,’” Lyotard rejects the myth of reason behind the Habermasian project. The project implies that there is a correct moral and scientific standard to which communities ought to aspire. Enforcement of such standards valorize conformity and, as witnessed in the fascist European states in the 1920s and 1930s, promote terrorism to extract it. This introduces a tension within postmodernism that has not been worked through (Frank 1990). On the positive side, the supposition of “multivocality” in scientific and moral discourse promotes the “excavation” of minority voices and minority experiences which have been occluded in “master” modernist ways of thinking. Every point of view can be heard, none can be privileged. On the other side, postmodernism seems ill equipped to distinguish between any particular moral or objective position and any other, including the fascist discourse that is associated with modernism. Choosing between Holocaust history and Holocaust denial would seem in principle to be a matter of rhetorical preference—which is clearly nihilistic.

When we move away from epistemology to postmodernity's impact on art and architecture, the situation is different. Postmodernism celebrates a rejection of hegemonic traditions and styles by mixing elements from competing schools and by bringing the rim of representation into focus as an organic element of the depiction. Novels exploit the discontinuities in perspective and the fragmentations in contemporary society. In architecture, postmodernism represents a repudiation of high-density ("efficient") functionality based on centralized city planning with an emphasis on no-frill construction. Decentralized planning emphasizes collage and eclecticism and the development of spaces to heighten aesthetic possibilities (Harvey 1989). The globalization of experience encourages such exchanges and experimentation. However, the replacement of "efficient" spaces like public housing, with their brutalizing side effects, with more particularistic designs is part of the emancipatory interest of modernity—so that at least here, postmodernity is still part and parcel of the Enlightenment. This raises three questions.

First, to what extent has the case been made that the changes in capitalism during the past two decades have been so profound as to represent a destruction of modernist society and a disappearance of history and "the social" (i.e., real face-to-face community)? Some critics dispute the claim (Baum 1990); others point out that Baudrillard's evidence is unconvincing since his own postmodern manner of exposition is self-consciously rhetorical, inflationary, and, consistent with the idiom, only one possible reading of recent history (Smart 1990). Second, if we have no careful (i.e., modernist) analysis of the factors that have contributed to the demise of our confidence in the "master narratives" of modernism, would it not seem to be impossible to discern which ones have been shaken, and how seriously our confidence in them has been eroded? Skepticism is insufficient. Finally, under these circumstances, is it not predictable that there is no consensus about the meaning of postmodernism? Arguably this dissensus is inherent in the perspective. For Baudrillard (1983), postmodernism spells the end of sociology (and other modernist subjects); for Bauman (1988), it is simply a new topic—a sociology of postmodernism, but not a postmodern sociology. For Brown (1990), it is an opportunity to rethink the continuing

relevance of the role of rhetoric in politics and science in order to employ knowledge in guiding human conduct. Under these circumstances, the proclamation of the death of modernism would seem premature.

Over the past decade two debates have emerged which raise questions about the postmodern project and which deal respectively with the issues of (1) epistemology, or the postmodern struggle with truth and the special standing of the natural sciences, and (2) of morality and the struggle to define and achieve the "good society" within democratic politics.

The first issue is raised by the famous hoax perpetrated by Alan Sokal, professor of physics at New York University, a left-leaning radical who taught math in Nicaragua under the rule of the Sandanistas. His "transformative hermeneutics of quantum gravity" was a spoof of the pretentious claims about science running through the leading figures of the French intellectual establishment—Derrida, Lacan, and Lyotard—spiked, according to Sokal, with the usual nonsensical and impenetrable jargon and buzzwords (Sokal 1996b). The parody was published in a special issue of *Social Text* designed by the editors to rebuke the criticisms of their antiscientific agenda by leading scientists. Sokal published an exposé at the same time in *Lingua Franca* (1996a). The hoax ignited a storm of criticism on both sides which some have referred to as "the science wars" (Natoli 1997, pp. 115 ff.; Kingwell 1999). The hoax continues to provoke debate today (see <http://www.physics.nyu.edu/faculty/sokal/index.html>). Although Sokal abused the good faith of the editors of a journal in a humanities discipline, he justified his behavior by reference to the naïveté of those editors, whose credentials could scarcely permit them to draw the conclusions about the purely rhetorical foundations of science or to discredit centuries of scientific progress. Yet in neither side was the "demonstration" definitive. The validity of the scientific method is neither established by the hoax, nor rubbished by what Richard Dawkins writing in *Nature* dubbed "the vacuous rhetoric of montebanks and charlatans" on the Left Bank (1988, p. 141). Nor is it undermined on this side of the Atlantic by Steven Seidman's claims in *Contested Knowledge* that the scientific community in eighteenth-century Europe prevailed over aristocracy and the

Church because it provided the forces of liberalism with a convenient but essentially vacuous cudgel with which to rout the older order: "Is the claim that only science ensures true knowledge . . . not merely another ruse on the part of a rising social elite wishing to legitimate their own aspirations for privilege? . . . [S]cience, just like religion, rests upon a series of assumptions that cannot be scientifically proven . . . Enlightenment science is as weak a guarantor of truth as religion." (1994, p. 24). Science for Seidman and the other postmodernists is just faith in a secular vein without any special epistemological leverage. The proponents of the Enlightenment were simply replacing priests with scientists, and religion with another form of superstition. Sokal and Bricmont's *Fashionable Nonsense* (1998) calls the postmodernists' bluff on such charges. At this point, it is premature to draw conclusions about the future of this controversy. Suffice it to say for the time being that poetry and physics should be cataloged separately.

On the issue of morality and everyday politics, John O'Neill's *The Poverty of Posmodernism* takes a similarly critical view of "the academic trendiness" of advocates of cynicism and fragmentation who have forsaken humanism and the critique of social injustice as part of the postmodern turn. "Today we are told to jettison the old-fashioned belief in unique values . . . we are wholly ruled by prejudice and politics . . . we are asked to believe that human beings are now so speciated by gender and race—though we are silent about class—that there can be no universal knowledge, politics, or morality." As O'Neill stresses, "these ideas have not grown up among the masses who have sickened of the injustices and exploitation that grinds their life . . . it is not these people who have abandoned idealism, universalism, truth and justice. It is those who already enjoy these things who have denounced them on behalf of the others" (1995, p. 1)—the new "illiterati" of the simulacrum, the Disney World theorists and the esthetes of the radically chic. "Too much of the world still starves, dies young and is wasted by systematic greed and evil for anyone to write the obituaries of philosophy, ideology and humanism" (p. 196). O'Neill argues that there are good philosophical foundations in the phenomenology of everyday life, the transcendental instinct of reason to fasten on the contradictions of existence, and commonsense scepticism

to keep alive the distinction between truth and falsity, freedom and slavery, and justice and injustice. The idea that political discourse has been reduced to mere perspective, that there is nothing outside language or that there is no "extratextuality" deny the vision of history and community, and exhaust culture with cultural industry, communication with fashion. "The claim that there are no longer any grand stories capable of under-writing common sense . . . gives comfort only to those who lack community at any level of society other than intellectual fashion" (p. 198). O'Neill's own inspiration for criticizing the antirationalism of postmodernism is founded on what he describes as the tension between language and vision, between elites and masses, between reality and possibility. The static linguisticity of postmodernism is a trap, an illusion. The current celebration of relativism by what used to be the progressive left is not only ill founded but politically nihilistic. It valorizes only minoritarianism while leaving unexamined the consequences of globalization and the dismantling of the welfare state by the political right. That state, with all its blemishes, has been an achievement in which civic democracy took root, in which politics was driven in part for reasons other than naked self-interest. In his vision, O'Neill says that postmodernism undermines our sense of cultural debt to the past and removes the duty which each generation has to pass on what it has learned to the future. "Those of us who own knowledge, who enjoy literacy, health, self-respect and social status have chosen to rage against our own gifts rather than to fight for their enlargement in the general public. We have chosen to invalidate our science, to psychiatrize our arts, to vulgarize our culture, to make it unusable and undesirable by those who have yet to know it. We honour no legacy. We receive no gifts. We hand on nothing" (1995, p. 2). Yet if O'Neill's critique of antirationalism is sound, all this changes and his self-flagellation is premature.

The moral and epistemological concerns of O'Neill and Sokal are quite different from the critiques of commodity esthetics in George Ritzer's *The McDonaldization of Society*, which provides a useful point of comparison. Following Weber's insight about the intrusiveness of rationalization, Ritzer shows how the McDonald fast-food franchise led to a widespread adoption of commodity

production and consumption based on standardization, efficiency, leveling of tastes, and the stifling of worker experiences that has come to set the model for our individual relationships with the market, if not the world generally. But McDonaldization calls forth that critical posture which Veblen labeled “the instinct of craftsmanship” ([1899] 1962, p. 75), without which waste and conspicuous consumption would be unbridled. Like Veblen, Ritzer applauds the resistance to such trends; but his perspective is Weberian and transcendental. By contrast, for students of postmodernism, scientific truths and emancipatory projects are merely further commodities, fashions, or fictions caught up in the McDonalization of empirical truth and the social good. Sokal and O’Neill represent the opening of serious dialogues on these engaging issues.

(SEE ALSO: *Popular Culture*; *Postindustrial Society*; *Social Philosophy*)

REFERENCES

- Baudrillard, Jean 1983 *In the Shadow of the Silent Majorities*. New York: Semiotext(e).
- Baum, Gregory 1990 “The Postmodern Age.” *Canadian Forum* (May):5-7.
- Bauman, Z. 1988 “Is There a Postmodern Sociology?” *Postmodernism*, special issue of *Theory, Culture and Society* 5(2-3):217-238.
- Bell, Daniel 1976 *The Cultural Contradictions of Capitalism*. New York: Basic.
- Bogard, William 1990 “Closing Down the Social: Baudrillard’s Challenge to Contemporary Sociology.” *Sociological Theory* 8(1):1-15.
- Brown, Richard H. 1990 “Rhetoric, Textuality, and the Postmodern Turn in Sociological Theory.” *Sociological Theory* 8(2):188-197.
- Dawkins, Richard 1998 “Postmodernism Disrobed.” *Nature* 394(9 July):141-143.
- de Saussure, Ferdinand 1959 *Course in General Linguistics*. C. Bally and A. Sechehaye, eds., Wade Baskin, trans. New York: McGraw-Hill.
- Frank, Arthur 1990, “Postmodern Sociology/Postmodern Review.” *Symbolic Interactionism* 14(1):93-100.
- Giddens, Anthony 1987 “Structuralism, Post-structuralism and the Production of Culture.” In A. Giddens and J. Turner, eds., *Social Theory Today*. Stanford, Calif.: Stanford University Press.
- Habermas, J. 1973 *Legitimation Crisis*. Boston: Beacon.
- Harvey, David 1989 *The Condition of Postmodernity*. Oxford: Blackwell.
- Kingwell, Mark 1999 “Peace in the Science Wars?” *Globe and Mail* (May 29). (Review Essay, Toronto.)
- Lyotard, Jean-Francois 1984 *The Postmodern Condition: A Report on Knowledge*. Minneapolis: University of Minnesota Press.
- Mills, C. Wright 1959 *The Sociological Imagination*. New York: Oxford University Press.
- Natoli, Joseph 1997 “Postmodernity’s War with Science.” In *A Primer on Postmodernity*. Malden, Mass.: Blackwell.
- O’Neill, John 1995 *The Poverty of Postmodernism*. New York: Blackwell.
- Ritzer, George 1996 *The McDonalization of Society*. Thousand Oaks, Calif.: Pine Forge.
- Seidman, Steven 1994 *Contested Knowledge: Social Theory in the Postmodern Era*. Malden, Mass.: Blackwell.
- Smart, Barry 1990 “On the Disorder of Things: Sociology, Postmodernity and ‘The End of the Social.’” *Sociology* 24:397-416.
- Sokal, Alan D. 1996a “A Physicist Experiments with Cultural Studies.” *Lingua Franca* 6(4):62-64.
- 1996b “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity.” *Social Text* 14(1, 2):217-252.
- , and Jean Bricmont 1998 *Fashionable Nonsense: Post Modern Intellectuals’ Abuse of Science*. New York: St. Martin’s.
- Touraine, Alain 1984 “The Waning Sociological Image of Social Life.” *International Journal of Comparative Sociology* 25:33-44.
- Veblen, Thorsten (1899) 1962 *The Theory of the Leisure Class*. New York: Mentor.

AUGUSTINE BRANNIGAN

**POVERTY**

Scholarly as well as ideological debate has long centered around the most elementary questions concerning poverty. What is poverty? How can it be measured? What causes it? Is it a natural phenomenon or a symptom of a poorly ordered society? Though answers to all these questions abound, there is no definitive answer to any one of them, nor can there ever be, for the questions are not purely demographic, but moral, ethical, and political as well. Poverty is a concept, not a fact, and



must be understood as such. Even though no definitive answers are possible, this does not mean that all answers are thereby equal; many are based on ignorant assumptions and ill-formed judgments. Sociologists involved in poverty research seek to make sure that all understand the meaning and consequences of various points of view, and that both theoretical and policy research is based soundly upon clear definitions and reliable data.

Even the definition of "poverty" is problematic. The word is derived from the French *pauvre*, meaning "poor." Poverty is simply the state of lacking material possessions, of having little or no means to support oneself. All would agree that anyone lacking the means necessary to remain alive is in poverty, but beyond that there is little agreement. Some scholars and policy makers would draw the poverty line at the bare subsistence level, like Rowntree's "the minimum necessities for the maintenance of merely physical efficiency" (1901, p. viii). Others argue for poverty definitions that include persons whose level of living is above subsistence but who have inadequate means; among those holding to the latter, further arguments concern the definition of adequacy. Social science cannot resolve the most basic arguments. For example, the level of living implied by the poverty threshold in the United States would be seen as desirable and unattainable in many other countries, yet few would suggest that poverty in the United States be defined by such outside standards. Sociologists can evaluate the demographic and economic assumptions underlying standards of poverty, but not the standards themselves.

### CONCEPTIONS OF POVERTY

Poverty can be defined in absolute or relative terms. The subsistence line is a good example of an absolute definition (i.e., below this line one does not have sufficient resources to survive). A criterion based on some arbitrary formula, such that poverty equals some fraction of the median income or below, is a good example of a relative definition (e.g., "All persons earning less than 25 percent of the median income are poor"). In all industrial societies an absolute definition will have far fewer persons officially in poverty than will a relative definition, creating natural political pressure for absolute definitions. For example, a study in 1976 revealed that if poverty was defined as

having 50 percent of the median income, data on income distributions would show that an unchanging 19 percent of the population had been poor for almost the past two decades (U.S. DHEW 1976). Absolute definitions show declines in poverty over time in industrial nations. There are valid arguments for both types of definitions. Some argue that relative definitions of poverty render the term meaningless in affluent societies, and make cross-national comparisons difficult—for example, in an advanced industrial society, 50 percent of national median income could leave one adequately provided for, while the same percentage in many less industrialized societies would not provide basic necessities to sustain life. On the other hand, within societies there is evidence that most people see poverty in relative terms rather than as an absolute standard (Rainwater 1974; Kilpatrick 1973). That is, popular conceptions of what level of living constitutes poverty have been found to change as general affluence goes up and down. Advocates of relative measures point out that any absolute measure is arbitrary and thus meaningless. A reasonable definition of the poor, they argue, should be one that demarcates the lower tail of the income distribution as the poor, whatever the absolute metric represented by that tail, for those persons will be poor by the standards of that time and place. As the average level of income rises and falls, they argue, what is seen as poverty will, and should, change. Advocates of absolute measures of poverty do not deny that perception of poverty is intimately tied to distributional inequality, but argue that relative definitions are too vague for policy purposes. An absolute standard, defined on some concrete level of living, is a goal that can possibly be attained. Once it is attained, they say, a new goal could be set. Eliminating poverty as defined by relative standards is a far more difficult goal, both practically and politically. T. H. Marshall noted, "the question of what range of inequality is acceptable above the 'poverty line' can only marginally, if at all, be affected by or affect the decision of where that line should be drawn" (1981, p. 52).

Relative versus absolute poverty is a distributional distinction, but there are other important distinctions as well. A social distinction, and one with considerable political import, is usually made between the "deserving poor" and the "undeserving poor." In their brief summary of the historical

origins of this distinction, Morris and Williamson (1986, pp. 6–12) maintain that it became significant in the fourteenth century, when, for a variety of reasons (the decline of feudalism, the rise of a market economy with concomitant periodic labor dislocations, bubonic plague–induced regional labor shortages), the poor became geographically mobile for the first time. Before that, the local Catholic parish, with its “Blessed are the poor” theology, was the primary caretaker of the indigent. Mobility caused an increase in the number of able-bodied individuals needing temporary assistance, and troubles arising from their presence contributed to a growing antipathy toward the able-bodied poor.

Katz (1989) also traces the origins of the “undeserving poor” in part to demographic factors. He points out that, prior to the twentieth century, poverty was a seemingly unalterable fact of life, and most people would spend their lives in it. Thus no moral taint was attached to poverty. The only policy question usually involved the locus of responsibility for aid, and the answer was a simple one: Responsibility was local, and those needy persons not belonging to the community could be “resettled.” Increased population mobility made the settlement provisions unworkable, and the original distinction between the genuinely needy and “rogues, vagabonds, and sturdy beggars” hardened into a moral distinction between the poor, who needed no public relief, and “paupers,” those needing assistance because of personal failings (Katz 1989, pp. 12–14).

Feagin (1975, chap. 2) locates the origins of negative attitudes toward the poor in the Protestant Reformation. Under Protestantism, he notes, the “work ethic”—the ideology of individualism—became a central tenet of the Western belief system. Poverty, in the extreme Calvinist version of this viewpoint, is largely a consequence of laziness and vice, and can even be regarded as just punishment from a righteous God. The rise of Puritan thought contributed to the increasing disfavor with which the unemployed and destitute were regarded. It became a matter of faith that poverty was individually caused and must thereby be individually cured. These ideas became secularized, and programs to aid the poor thereafter focused on curing the individual faults that led to poverty: Potential problems in the structure of society that caused unemployment and underemployment were

not to be scrutinized in search of a solution. The notion of poverty continues to be in flux. As Marshall (1981) noted, the concept has been with us since antiquity, but its meaning has not been constant through the ages.

Most sociologists today distinguish among three major types of explanations of poverty: individual, situational, and structural theories. Individual theories attribute the primary cause of poverty to individual failings or, more neutrally, to individual differences—the central argument being that the poor are different from the nonpoor in some significant way. Situational theories agree that the poor are different from the nonpoor, but argue that the differences are a result of poverty, not a cause of it. Structural theories see differences in individual attributes as irrelevant, and argue that poverty has only societal-level origins: The characteristics of economic systems create poverty, not the characteristics of individuals.

### THEORY AND POLICY

The epitome of the individual viewpoint in the social sciences was the once-dominant “culture of poverty” explanation for destitution. Oscar Lewis (1961, 1966) is usually credited with this idea, which sees poverty not only as economic deprivation, or the absence of something, but also as a way of life, the presence of specific subcultural values and attitudes passed down from generation to generation. Lewis saw the structure of life among the poor as functional, a set of coping mechanisms without which the poor could not survive their harsh circumstances. But there were negative consequences of the value system as well, he noted. Family life was disorganized, there was an absence of childhood as a prolonged lifecycle stage, a proliferation of consensual marriages, and a very high incidence of spouse and child abandonment—all of which left individuals unprepared and unable to take advantage of opportunities. Exacerbating the problem, the poor were divorced from participation in and integration into the major institutions of society, leading to constant hostility, suspicion, and apathy. Many have maintained that the culture-of-poverty viewpoint dovetailed perfectly with a politically liberal view of the world. It blamed the poor as a group for their poverty, but held no single person individually responsible, nor

did it blame the structure of the economy or the society. This view of the poor led to antipoverty policies directed at changing the attitudes and values of those in poverty, so that they could “break out” of the dysfunctional cultural traits they had inherited. It led political liberals and radicals to attempts to “organize the poor.” Political conservatives transformed the explanation into one that held the poor more culpable individually and the problem into one that was intractable—“benign neglect” being then the only sensible solution (Banfield 1958, 1970; Katz 1989). There were many problems with the culture-of-poverty explanation. Most serious was the fact that the cultural scenario simply doesn’t fit. Only a minority of the poor are poor throughout their lives; most move in and out of poverty. Also, a substantial proportion of those in poverty are either women with children who fell into poverty when abandoned by a spouse, or the elderly who became poor when their worklives ended: Neither event could be explainable by the culture of the class of destination. Many studies falsified specific aspects of the culture-of-poverty thesis (for a review, see Katz 1989, pp. 41 ff.), and Hyman Rodman’s influential notion of the “lower-class value stretch” (1971) offered an alternative explanation (the poor actually share mainstream values, but must “stretch” them to fit their circumstances—remove the poverty, and they fit neatly into dominant culture—attempts to alter their “culture” are unnecessary, and meaningless, since “culture” is not the problem). Nonetheless, the culture-of-poverty thesis was (and to some extent still is) a very popular explanation for poverty. This is probably so in part because it fits so well the individualistic biases of most Americans. Surveys of attitudes toward poverty have shown that most persons prefer “individualistic” explanations of poverty, which place the responsibility for poverty primarily on the poor themselves. A minority of Americans subscribe to “structural” explanations that blame external social and economic factors, and this minority consists largely of the young, the less educated, lower income groups, nonwhites, and Jews (Feagin 1975, p. 98).

A more sophisticated recent treatment incorporating some of the explanatory power of the culture of poverty argument without that theory’s untenable assumptions is William Wilson’s depiction of the “underclass” (1987). This recent work

by Wilson on the underclass has been criticized by some as a return to classical culture of poverty theory in a new guise. It is not, of course, and represents a very different type of explanation. Wilson defines the underclass as an economically disadvantaged group whose marginal economic position and weak attachment to the labor force is “uniquely reinforced by the neighborhood or social milieu” (Wilson 1993, p. 23). Changes in the geography of employment, industrial specialization, and other factors have resulted in a rise in joblessness among urban minorities, which has in turn led to an increase in other social dislocations. Primary among those other factors has been the steady out-migration of working-class and middle-class families from the inner cities, which groups would normally provide a social buffer. Wilson notes: “in a neighborhood with a paucity of regularly employed families and with the overwhelming majority of families having spells of long-term joblessness, people experience a social isolation that excludes them from the job network system that permeates other neighborhoods and that is so important . . . other alternatives such as welfare . . . are not only increasingly relied on, they come to be seen as a way of life” (1987, p. 57). The 1990 U.S. Census showed that about 15 percent of the poor lived in neighborhoods where the poverty rate was at least 40 percent (O’Hare 1996). In these neighborhoods, where few are likely to have resources or job networks, reside Wilson’s “underclass.” Unlike the culture-of-poverty theory, Wilson’s theory analyzes the structural and cultural resources of poor places, rather than the socialized attitudes and values of poor people. Wilson contends that “ghetto-specific cultural traits” are relevant in understanding the behavior of inner-city poor people, but these traits, *contra* culture-of-poverty theory, do not have a life of their own. That is, they are an effect of deprivation and social isolation, not a cause of it, and they command very little commitment—they are not self-perpetuating. “Social isolation,” Wilson’s key concept, implies not differential socialization of the inner-city poor, but rather their lack of cultural resources supporting the desirability and possibility of achieving culturally normative aspirations. The individual characteristics normally associated with a culture of poverty argument represent expected and even rational responses to adverse environments, not socialized belief systems.

While Wilson's work has been very influential in the social sciences, one of the most politically influential recent works on poverty policy has been that of Murray (1984). Murray argued that the viewpoint that individuals ultimately cause their own poverty changed in the 1960s to the viewpoint that the structure of society was ultimately responsible. This alteration in the intellectual consensus, which freed the poor from responsibility for their poverty, was fatally misguided, he argues, and caused great damage to the poor. Despite enormously increased expenditures on social welfare from 1965 on, he maintains, progress against poverty ceased at that point. In the face of steadily improving economic conditions, the period 1965–1980 was marked by increases in poverty, family breakdown, crime, and voluntary unemployment. Murray argues that this occurred precisely because of the increased expenditures on social welfare, not despite them. Work incentive declined during these years because of government policies that rewarded lack of employment and nonintact family structure. It is a standard economic principle that any activity that is subsidized will tend to increase. Murray's arguments have had policy impact, but have been subject to extensive criticism by students of the field.

As evidence of the disincentive to work brought about by social welfare payments, Murray cites the Negative Income Tax (NIT) experiments. These were large social experiments designed to assess the effects of a guaranteed income. The first NIT experiment was a four-year study in New Jersey from the late 1960s to the early 1970s. In this study, 1,375 intact "permanently poor" families were selected, and 725 of them were assigned to one of eight NIT plans. It was found that the reduction in labor-market activity for males caused by a guaranteed income was not significant, but that there were some reductions for females (5–10 percent of activity), most of which could be explained by the substitution of labor-market activity for increased child care (home employment). In a larger NIT study conducted throughout the 1970s, the Seattle-Denver Income Maintenance Experiment (usually referred to in the literature as the SIME-DIME study), much larger work disincentives were found, about 10 percent for men, 20 percent for their spouses, and up to 30 percent for women heading single-family households (see Haveman 1987, chap. 9, for an excellent summary of the

many NIT experiments). Murray offered these findings as evidence that existing welfare programs contributed to poverty by creating work disincentives. Cain (1985) pointed out that the experiments provided much higher benefits than existing welfare programs, and also noted that, given the low pay for women at that level, the 20 percent reduction for wives would have a trivial effect on family income. If it resulted in a proportionate substitution of work at home, the reduction could actually lead to an improvement in the lives of the poor. Commentators have presented arguments against almost every point made by Murray, insisting that either his measures or his interpretations are wrong. For example, Murray's measure of economic growth, the gross national product (GNP), did increase throughout the 1970s, but real wages declined, and inflation and unemployment increased—poverty was not increasing during good times, as he argues. His other assertions have been similarly challenged (for summaries, see McLanahan et al. 1985; Katz 1989, chap. 4), but though his arguments and empirical findings simply do not stand up to close scrutiny, the broad viewpoint his work represents remains important in policy deliberations, probably because they offer pseudoscientific support for the biases of many.

### MEASURES OF POVERTY

In the United States, official poverty estimates are based on the Orshansky Index. The index is named for Mollie Orshansky of the Social Security Administration, who first proposed it (Orshansky 1965). It is an absolute poverty measure, based on the calculated cost of food sufficient for an adequate nutritional level and on the assumption that persons must spend one-third of their after-tax income on food. Thus, the poverty level is theoretically three times the annual cost of a nutritionally adequate "market basket." This cost was refined by stratifying poor families by size, composition, and farm/nonfarm, and by creating different income cutoffs for poverty for families of differing types. Originally there were 124 income cutoff points, but by 1980 the separate thresholds for farm families and female-headed households had been eliminated, and the number of thresholds reduced to 48. Since 1969 the poverty line has been regularly updated using the Consumer Price Index (CPI) to adjust for increased costs. The

original index was based on the least costly of four nutritionally adequate food plans developed by the Department of Agriculture. Since a 1955 Department of Agriculture survey of food consumption patterns had determined that families of three or more spent approximately one-third of their income on food, the original poverty index was simply triple the average cost of the economy food plan. This index was altered for smaller families to compensate for their higher fixed costs, and for farm families to compensate for their lower costs (the farm threshold began as 70 percent of the nonfarm for an equivalent household, and was raised to 85 percent in 1969). Originally, the poverty index was adjusted yearly by taking into account the cost of the food items in the Department of Agriculture economy budget, but this changed in 1969 to a simple CPI adjustment (U.S. Bureau of the Census 1982).

Over the years there have been many criticisms of the official poverty measure and its assumptions (for summaries and extended discussion, see U.S. DHEW 1976; Haveman 1987). The first set of problems, some argue, come from the fact that the very basis of the measure is flawed. The economy food budget at the measure's core is derived from an outdated survey that may not reflect changes in tastes and options. Further, the multiplication of food costs by three is only appropriate for some types of families, other types must spend greater or lesser proportions on food. Some estimates indicate that the poor spend half or more of their income on food; the more well-to-do spend one-third or less. Even if the multiplier was correct, the original Department of Agriculture survey discovered it for posttax income; in the poverty measure it is applied to pretax income, though the poor pay little in taxes. Other problems often cited include the fact that the "economy budget" assumes sufficient knowledge for wise shopping—a dubious assumption for the poor—and the fact that the poor are often locked into paying much higher prices than average because of a lack of transportation. An additional problem is that the poverty thresholds are not updated by using changes in the actual price of food, but instead by changes in the CPI, which includes changes in the price of many other items such as clothing, shelter, transportation, fuel, medical fees, recreation, furniture, appliances, personal services, and many other items probably irrelevant to

the expenses of the poor. Findings are mixed, but it is generally agreed that the losses in purchasing power suffered by the poor in inflationary periods is understated by the CPI (see Oster et al. 1978, p. 25). A second set of problems derives from the fact that the definition is based on income only. Both in-kind transfers and assets are excluded. Excluding in-kind transfers means that government-provided food, shelter, or medical care is not counted. Excluding assets means that a wealthy family with little current income could be counted as poor.

### DEMOGRAPHY OF POVERTY

In 1987 the average poverty threshold for a family of four was \$11,611 per year (all figures in this paragraph are from U.S. Bureau of the Census 1989a, p. 163; 1989b, p. 166; 1990). This means the assumed annual cost of an adequate diet for four persons was \$3,870.33, or about 88 cents per meal per person. For a single person the poverty threshold was \$5,778, and the food allowance \$1.76 per meal. In 1986 the poverty threshold was \$11,203, allowing 85 cents per meal, and in 1988 it had risen to \$12,091, or 92 cents per meal. In the United States in 1987 there were 32,341,000 persons below the poverty threshold, or 13.4 percent of the population. In 1988 there were 31,878,000, or 13.1 percent, almost a half-million fewer persons below official poverty than the year before. These figures underestimate official poverty somewhat, since they are based on the Current Population Survey, which is primarily a household survey and thus does not count the homeless not in shelters. The decline from 1987 to 1988 in the number in poverty is part of a long-term trend. In 1960 there were 8 million more—39,851,000 persons—who by today's guidelines would have been counted as officially in poverty, representing 22.2 percent of the population. By the official, absolute standard, poverty has greatly decreased over the past three decades, both in terms of the actual number of persons below the threshold and, even more dramatically, by the percentage of the population in poverty (U.S. Bureau of the Census 1989). This decrease actually took place over two decades, since the number of people in poverty in 1970 had declined to only 25,420,000, or 12.6 percent of the population, and the number and percentage have risen since then, but never back as high as the 1960 levels. Poverty is not evenly spread over the population. Of those below the

official poverty level in 1988, 57.1 percent were female, 29.6 percent were black, and 16.9 percent were Hispanic. Female-headed families with children were disproportionately poor. In poverty in 1988 were 38.2 percent of all such white families and 56.3 percent of all such black families (this is a gender phenomenon, not a single-parent one, since in 1988 only 18 percent of male-headed single parent families were below the poverty threshold). The age composition of the poor population has changed. In 1968, 38.6 percent of those in poverty were of working age (18–64), while twenty years later 49.6 percent of those in poverty were of working age. From 1968 to 1988 the percentage of the poor population over 65 declined from 18.2 percent to 10.9 percent, and the percentage who were children under 18 declined from 43.1 percent to 39.5 percent. A higher percentage of working-age poor is seen by some as a sign of worse times. It almost certainly reflects not only economic downturns but also in part ideological biases toward helping the presumably able-bodied poor; most antipoverty programs have been specifically aimed at the old or the young. O'Hare (1996) points out that the poverty rate and the number of poor in the 1990s exceed those figures in the 1970s. He notes that all the dramatic postwar decline in poverty rates occurred before 1973. After that, poverty rates in the United States rose through the early 1980s, then declined, but never back to the 1973 level.

Despite extensive debate about the policy implications of various definitions of poverty, and the inherent difficulty of locating this population, one can have confidence that the poor are being counted with reasonable precision. More than one generation of social scientists have contributed to the refinement of the measures of poverty, and existing statistical series are based on data collected by the U.S. Bureau of the Census—an organization with very high technical competence. Nonetheless, there is one group, the extremely poor, whose numbers are in doubt. All current measurement relies on the household unit, and assumes some standard type of domicile. As Rossi puts it, “our national unemployment and poverty statistics pertain only to that portion of the domiciled population that lives in conventional housing” (1989, p. 73). An extremely poor person living, perhaps temporarily, in a household where other adults had sufficient income would not be counted

as being in poverty. Even more important, the literally homeless who live on the street, and those whose homes consist of hotels, motels, rooming houses, or shelters are not counted at all in the yearly Current Population Survey (the decennial census does attempt to count those in temporary quarters, but the 1990 census was the first to even attempt to count those housed in unconventional ways or not at all). The studies of Rossi and his colleagues indicate that the number of extremely poor people in the United States (those whose income is less than two-thirds of the poverty level) is somewhere between four and seven million. The number of literally homeless poor people, those who do not figure into the official poverty counts, must be estimated. The best available estimate is that they number between 250,000 and 350,000, about 5–8 percent of the extremely poor population (Rossi 1989). The number of extremely poor people has more than doubled since 1970, while the population was increasing only by 20 percent (Rossi 1989, p. 78). The extremely poor are at considerable risk of becoming literally homeless. When they do so, they will disappear from official statistics (just as the unemployed cease being officially unemployed soon after they give up the search for work). To see that official statistics remain reliable in the face of increasing extreme poverty is the most recent methodological challenge in the field.

#### POVERTY IN LOW-INCOME COUNTRIES

Most of the discussion thus far has concerned poverty in the United States. Comparing poverty across countries is a difficult enterprise, but is important to do if one is to put poverty in any individual country in perspective. Poverty in the United States and in other highly industrialized countries simply does not fall into the same category as poverty in less industrialized nations. Many of those classified as in poverty in the United States would be seen as reasonably well off by international standards. This means that, although many countries report a “percentage-in-poverty” figure for their populations, these figures cannot be sensibly compared, since the concept of what constitutes poverty varies so widely. Statistics from the United Nations Development Program 1998 Human Development Report can illustrate this stark contrast. In 1998, in the forty-four countries the U.N. classifies as “least developed,” about 29

percent of the population was not expected to survive to age 40. Compare this to the 5 percent not expected to survive to that age in the industrial countries. In the least developed countries, 43 percent of the population has no access to safe water, 64 percent no access to sanitation, and 51 percent no access to health services (considering all developing countries rather than just the poorest, those figures would be 29 percent, 20 percent, and 58 percent, respectively). This level of living is characteristic of very few people in industrial nations, making poverty comparisons almost meaningless. The World Bank has attempted to provide international comparisons of poverty by creating a measure of the percentage of a country's population living on less than \$1 a day, calculated in 1985 international prices and "adjusted to local currency using purchasing power parities" (World Bank 1999, p. 69). The \$1-a-day figure was chosen because this is the typical poverty line in low-income countries. World Bank figures indicate that about 1.3 billion people live on less than \$1 a day. Twenty-seven countries in which over 25 percent of the population lives with resources at less than this level were counted, as were fourteen countries (Guatemala, Guinea-Bissau, Honduras, India, Kenya, Lesotho, Madagascar, Nepal, Niger, Peru, Rwanda, Senegal, Uganda, Zambia) where approximately half or more of the population lives on less than \$1 a day.

It is clear that while relative poverty is a serious moral and political issue in industrial countries, absolute poverty—at levels unheard of in industrial countries—is a far more serious problem in much of the rest of the world.

The study of poverty is a difficult field, and is not properly a purely sociological endeavor. As even this brief overview shows, a thorough understanding requires the combined talents of sociologists, economists, demographers, political scientists, historians, and philosophers. All these fields have contributed to our understanding of the phenomenon.

REFERENCES

Banfield, Edward C. 1958 *The Moral Basis of a Backward Society*. New York: Free Press.  
 ——— 1970 *The Unheavenly City*. Boston: Little, Brown.  
 Cain, Glen 1985 "Comments on Murray's Analysis of the Impact of the War on Poverty on the Labor

Market Behavior of the Poor." In Sara McLanahan, et al., eds., *Losing Ground: A Critique*, Special Report No. 38. Institute for Research on Poverty, University of Wisconsin.  
 Feagin, Joe R. 1975 *Subordinating the Poor: Welfare and American Beliefs*. Englewood Cliffs, N.J.: Prentice-Hall.  
 Haveman, Robert H. 1987 *Poverty Policy and Poverty Research: The Great Society and the Social Sciences*. Madison: University of Wisconsin Press.  
 Katz, Michael B. 1989 *The Undeserving Poor: From the War on Poverty to the War on Welfare*. New York: Pantheon.  
 Kilpatrick, R. W. 1973 "The Income Elasticity of the Poverty Line." *Review of Economics and Statistics* 55:327-332.  
 Lewis, Oscar 1961 *The Children of Sanchez*. New York: Random House.  
 ——— 1966 *La Vida: A Puerto Rican Family in the Culture of Poverty—San Juan and New York*. New York: Random House.  
 McLanahan Sara, Glen Cain, Michael Olneck, Irving Piliavin, Sheldon Danziger, and Peter Gottschalk 1985 *Losing Ground: A Critique*, Special Report No. 38. Institute for Research on Poverty, University of Wisconsin.  
 Marshall, T. H. 1981 *The Right to Welfare and Other Essays*. New York: Free Press.  
 Morris, Michael, and John B. Williamson 1986 *Poverty and Public Policy: An Analysis of Federal Intervention Efforts*. Westport, Conn.: Greenwood.  
 Murray, Charles 1984 *Losing Ground: American Social Policy, 1950-1980*. New York: Basic.  
 O'Hare, William P. 1996 "A New Look at Poverty in America." *Population Bulletin* 51(2):1-48.  
 Orshansky, Mollie 1965 "Counting the Poor: Another Look at the Poverty Profile." *Social Security Bulletin* 28(1):3-29.  
 Oster, Sharon M., Elizabeth E. Lake, and Conchita Gene Oksman 1978 *The Definition and Measurement of Poverty*. Boulder, Colo.: Westview.  
 Rainwater, Lee 1974 *What Money Buys: Inequality and the Social Meaning of Income*. New York: Basic.  
 Rodman, Hyman 1971 *Lower-Class Families: The Culture of Poverty in Negro Trinidad*. London: Oxford University Press.  
 Rossi, Peter H. 1989 *Down and Out in America: The Origins of Homelessness*. Chicago: University of Chicago Press.  
 Rowntree, B. S. 1901 *Poverty: A Study of Town Life*. London: Macmillan.

- U.S. Bureau of the Census 1982 "Changes in the Definition of Poverty." *Current Population Reports*, Series P-60, No. 133. Washington, D.C.: U.S. Government Printing Office.
- 1989a "Poverty in the United States: 1987." *Current Population Reports*, Series P-60, No. 163. Washington, D.C.: U.S. Government Printing Office.
- 1989 "Money Income and Poverty Status in the United States: 1988." *Current Population Reports*, Series P-60, No. 166. Washington, D.C.: U.S. Government Printing Office.
- 1990 "Measuring the Effect of Benefits and Taxes on Income and Poverty: 1989." *Current Population Reports*, Series P-60, No. 169-RD. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Health, Education, and Welfare (DHEW) 1976 *The Measure of Poverty: A Report to Congress as Mandated by the Education Amendments of 1974*. Washington, D.C.: U.S. Government Printing Office.
- United Nations 1998 *Human Development Report*. Oxford: Oxford University Press.
- Wilson, William Julius 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- 1993 *The Ghetto Underclass: Social Science Perspectives*. Newbury Park, Calif.: Sage.
- World Bank 1999 *World Development Indicators, 1999*. Washington: International Bank for Reconstruction and Development/World Bank.

WAYNE J. VILLEMEZ

## POWER

See Bureaucracy; Interpersonal Power; Social Organization.

## PRAGMATISM

*Pragmatism*, the Greek root word of which means "action," grew out of a turn-of-the-century reaction in American philosophy to Enlightenment conceptions of science, human nature, and social order. Generally, it has sought to reconcile incompatibilities between philosophical idealism and realism. In the former, reality is conceived of as existing only in human experience and subjectivity, and is given in the form of perceptions and ideas. In the latter, reality is proposed as existing in

the form of essences or absolutes that are independent of human experience. These two traditions have grounded different approaches to empiricism and in their extremes can be found, respectively, in the solipsism of the British philosopher George Berkeley and the positivistic embracement of natural law by the French sociologist Auguste Comte.

As a response and an alternative to these traditions, pragmatism has not been developed as a unified philosophical system. Rather, it has existed as a related set of core ideas and precepts that are expressed as versions or applications of pragmatic thought. Variation within that thought ranges across the realism–idealism continuum and is largely a function of different analytical agendas. Despite that variation, however, pragmatism has become quite broad in its application to theoretical and research problems. Acknowledged as the most distinctive and profound contribution of American intellectual thought, its influence can be found in all contemporary disciplines in the humanities and social sciences. Its intellectual roots are described in Rucker (1969), Martindale (1960), and Konvitz and Kennedy (1960), and its core ideas are described in Shalin (1986), Rochberg-Halton (1986), and Rosenthal (1986).

## MAIN IDEAS AND VARIATIONS

In summary form, the main ideas embodying the thrust of pragmatism are as follows. First, humans are active, creative organisms, empowered with agency rather than passive responders to stimuli. Second, human life is a dialectical process of continuity and discontinuity and therefore is inherently emergent. Third, humans shape their worlds and thus actively produce the conditions of freedom and constraint. Fourth, subjectivity is not prior to social conduct but instead flows from it. Minds (intelligence) and selves (consciousness) are emergent from action and exist dialectically as social and psychical processes rather than only as psychic states. Fifth, intelligence and consciousness are potential solutions to practical problems of human survival and quality of life. Sixth, science is a form of adjustive intelligibility and action that is useful in guiding society. Seventh, truth and value reside simultaneously in group perspectives and the human consequences of action. Eighth, human nature and society exist in and are sustained



by symbolic communication and language. In these core ideas can be seen the neo-Hegelian focus on dialectical processes and the concurrent rejection of Cartesian dualisms, the Darwinian focus on the emergence of forms and variation through adjustive processes, and the behavioristic focus on actual conduct as the locus of reality and understanding. These ideas were embraced and developed by a variety of scholars and thinkers, including Ralph Waldo Emerson, Percy Bridgman, C. I. Lewis, Morris Cohen, Sidney Hook, Charles Morris, Charles Peirce, William James, Charles Horton Cooley, John Dewey, and George Herbert Mead. Of these, Peirce, James, Cooley, Dewey, and Mead will be reviewed here for purposes of assessing the relevance of pragmatism to social science and sociology.

Charles S. Peirce (1839–1914) is generally credited as the originator of the term *pragmatism* and the formulation of some of its basic tenets. One of his earliest statements (1877–1878) pertained to methods for resolving doubt about conclusions, in which he argued in favor of science because of its flexibility and self-correcting processes. This view contrasts sharply with the Cartesian basis of science in subjectivism and individualism proposed by Descartes. Rather than focusing on belief and consciousness as definitive, Peirce focused on probability. Both truth and scientific rationality rest in a community of opinion in the form of perpetual doubt and through a process of revisions are measured in terms of movement toward the clarification of ideas. The emphasis in this process is both evolutionary, since Peirce saw modes of representing knowledge moving from chance (firstness) to brute existence (secondness) to generality or order (thirdness), and pragmatic, since truth is meaningful only in terms of future consequences for human conduct (the “pragmatic rule”).

William James (1842–1910) was driven by the problem of determinism and free will. This problem was expressed in his monumental *Principles of Psychology* (1890), in which he established a functional view assimilating biology and psychology and treated intelligence as an instrument of human survival, and in his brilliant analysis of consciousness (1904), in which he characterized human experience as an ongoing flow instead of a series of psychic states. Pragmatist principles were sharply articulated in these works. James accepted

Peirce’s pragmatic rule that the meaning of anything resides in experimental consequences. Accordingly, the distinction between subject and object as fundamental was denied and was replaced with the idea that relations between the knower and the known are produced by and in ongoing experience. This more dialectical conceptualization informed his theory of emotions, which stressed bodily responses as producers of emotional responses (e.g., we feel afraid because we tremble) and his theory of the self, emphasizing the “I” (self as knower) and the “Me” (self as known) as tied to multiple networks of group affiliation. His focus throughout was on the operations of ongoing experience, and his formulations not only specified and elaborated pragmatist principles but anticipated or contributed to behaviorism, gestalt psychology, and operationalism.

Charles Horton Cooley (1864–1929), building on the work of James, rejected the legitimacy of all dualisms. In his famous statement that “self and society are twin-born,” he asserted the inseparability of individuals and society. Individuals, he proposed, are merely the distributive phase and societies the collective phase of the same social processes. The indissoluble connection of self and society, which was the dominant theme of Cooley’s writings, is manifested in their necessary interdependence. His theory of the social self held that self-concepts are behaviorally derived through reflected appraisals of the actions of others—the looking-glass self. Especially important in the process of self-acquisition are primary groups (family, friends), which link the person to society. Correspondingly, society significantly exists in the form of personal imaginations or mental constructs; society, Cooley stated, is an interweaving and interworking of mental selves. Cooley’s approach was thoroughly holistic and organic, with human consciousness and communication being the most critical processes, and his work added further to the pragmatist’s dismantling of the Cartesian split between mind and society.

John Dewey (1859–1952) was perhaps the most influential and prolific of the early pragmatists. Coming philosophically to pragmatism from Hegelianism, he emphasized intelligence, process, and the notion that organisms are constantly reconstructing their environments as they are being determined by them. He contributed forcefully to the critique of dualistic thought in his analysis of

stimulus-response theory (1896). Instead of constituting an arc, in which the stimulus leads to a response (a dualistic conception), Dewey argued that they are merely moments in an overall division of labor in a reciprocal, mutually constitutive process (a dialectical conception). Central to those dialectics was communication, which according to Dewey was the foundation and core mechanism of social order. He developed an instrumentalist theory of language (1925)—language as a tool—which was generalized into a broader instrumentalism. One of his central interests was moral and social repair through the application of intelligence and scientific methods. He merged theory and practice in the view that ideas are instruments for reconstructing and reconstituting problematic situations. Those ideas may be moral judgements or scientific findings, but both take the general form of hypotheses, which are proposals for action in response to difficulties. Dewey thus built upon Peirce's and James's pragmatic rule by arguing that validity and truth statements, whether theological or scientific, are established by examining the consequences of action derived from hypotheses.

George Herbert Mead (1863–1931) sought understanding of emergent human properties, such as the ability to think in abstractions, self-consciousness, and moral and purposive conduct. His central argument was that these properties are grounded in the development of language and social interaction as humans adjust to the conditions of their environments and group life (Mead 1934). His position is said to be one of social behaviorism, in which the social act is the unit of analysis and out of which minds and selves develop. The act has covert and overt phases. It begins in the form of an attitude (an incipient act), is constructed through role-taking processes (imaginatively placing oneself in the position of others), and is manifested in overt conduct. All social behavior involves a conflation of subjective and objective processes through which persons adjustively contend with the facts of their environments and simultaneously create social situations. Mead's explicit theory of time and sociality places these adjustments squarely in the dialectics of continuity and discontinuity (Maines et al. 1983).

In these five brief summaries can be seen how the early pragmatists wove together strands of

scientific method, evolutionary theory, language, and behaviorism into a radically new perspective. Pragmatism provided a clear alternative to perspectives based on Cartesian dualisms and reconstituted science, morality, aesthetics, political theory, and social development in terms of dialectical transactions. Philosophical idealism and realism were brought into a common framework in the proposition that human experience and facts of nature and society (the “world that is there,” as Mead called them) are only phases of ongoing social processes that mediate persons and their environments in terms of transacted meanings. The variation within pragmatism hinges largely on individual affinities for idealism and realism: James and Cooley tended toward idealism, Peirce toward realism, and Dewey and Mead toward a transactional midpoint between the two. Moreover, there is variation in pragmatism's influence in the social sciences and humanities. Dewey has been enormously influential in education and communication, Mead and Cooley in sociology and social psychology, Peirce in semiotics, and James in psychology. That variation, however, only represents modal tendencies, since pragmatism as a whole has had a significant impact across disciplines.

#### INFLUENCE IN SOCIAL SCIENCE

Since 1980 there has been a major resurgence of interest in pragmatism (Bernstein 1986). Its compatibilities with quantum mechanics and relativity theory have been articulated, as has its relevance for the development of a more social semiotics and discourse analysis (Perinbanayagam 1986). The relation of pragmatism to hermeneutics, from the tradition of German Idealism, has been reexamined (Dallmayr 1987), as has its relation to critical theory in the work of Jürgen Habermas (McCarthy 1984), literary criticism (Rorty 1982), phenomenology (Ricoeur 1985), cultural studies (Carey 1989), and modernization theory (Rochberg-Halton 1986). According to some, such as Richard Bernstein and more recently John Diggins (1994), this resurgence indicates that the early pragmatists were ahead of their time. Some scholars in the 1990s have continued to revise pragmatist thought on its own terms. Wiley (1995) has proposed a sophisticated modification of Mead's theory of the self through Peirce's triadic,

semiotic perspective. Joas (1993), while not writing on pragmatism, per se, has shown how pragmatist precepts have been intrinsic both to classical and contemporary theory. He also (Joas 1996) has retheorized creativity, long at the heart of pragmatism's emphasis on novelty and indeterminism, for dominant sociological models based on structural differentiation, rational adaptation, and self-enhancement.

Other scholars have sought to reinvigorate pragmatism's relevance for contemporary political and cultural agendas. Noting that the seeds of a feminist pragmatism existed in the work and practices of the classical thinkers (Mead, Dewey, James), Seigfried (1996) politicizes pragmatism in the common struggles of women, ethnic minorities, and the sexually marginalized. While retaining the traditional center of the perspective, she seeks to move it toward a reconsideration of contextual ethics and reciprocal moral responsibility and to promote the search for pragmatic truths that would emancipate people from distorted beliefs and values that become sedimented into accepted fact. These scholars also include Cornell West (1989), who articulated his version of "prophetic pragmatism" in *The American Evasion of Philosophy* and followed it with his influential *Race Matters* (1994), which was more politically engaged and contributed to a more vibrant democracy that better empowers local participation in institutional concerns. Still others have sought more radical revisions of pragmatism. Denzin (1992, 1996), drawing in part from West, provides a revision of pragmatism in postmodernist, cultural studies terms that seeks to transform it into a mechanism of cultural critique. Such works find audiences in humanistic circles, but have met with mixed reactions among social scientists. Farberman (1991) and Lyman (1997) reject the postmodernist project partly on the grounds of its nihilistic implications, Van Den Berg (1996) because of its limited vision of Enlightenment philosophy, and Maines (1996) because its core concepts are already contained in pragmatist thought. Others, such as Seidman (1996) have attempted more even-handed critiques and seek the common ground of postmodernism and pragmatism. Regardless of recent interpretations of pragmatism, the collapse of hegemonic theories and the corresponding import of post-positivistic debate in social scientific theorizing has brought

the basic tenets of pragmatism back into the search for new paradigms in social theory.

These recent influences and developments notwithstanding, there has been a long tradition of direct influence of pragmatism on social science research and theory. Sociology's first research classic was Thomas and Znaniecki's (1918–1920) study of Polish immigrant adaptation to American urban life. They were interested in questions of personal adjustment, family relations, neighborhood formation, delinquency, and cultural assimilation, and they used the principles of pragmatism, especially as expressed by G. H. Mead, to answer those questions. Their monumental five-volume work presented their attitude-value scheme as a general theory of the adjustive relations between individuals and society. "Attitudes" referred to the individual's tendencies to act and represented human subjectivity; "values" referred to the constraining facts of a society's social organization and represented the objective social environment. Both are present in any instance of human social conduct, they argued, but the relationships between the two are established in processes of interpretation that they called "definitions of the situation." Thomas and Znaniecki thus placed human agency at the center of their explanations, and they conceptualized society as the organization of dialectical transactions.

That research contained pragmatist ideas pertaining to the social psychological and social organizational aspects of human behavior. These aspects were developed during the 1920s and 1930s at the University of Chicago by Ellsworth Faris and Robert Park. Faris (1928) examined attitudes, especially in terms of the nature of their influence on behavior. He argued that human subjectivity is a natural datum for sociological research and proposed that wishes and desires, not attitudes, have a direct bearing on overt conduct. Park (1926) was more interested in large-scale historical issues such as urban organization and racial stratification. He directly applied Dewey's focus on society as communication to his research on urban communities. These communities, he argued, have objective spatial patterns, but those patterns are not separable from human consciousness. Urban ecology thus has a moral dimension composed of meanings that collectivities attribute to urban areas. The pragmatist roots of Park's sociology recently has been reemphasized

in his theory of human ecology (Maines et al. 1996), his influence on applied sociology (Reitzes and Reitzes 1992; see also Maines 1997), and as a framework for describing and theorizing urban public space (Lofland 1998).

The pragmatist themes of individual/society inseparability and human behavior as transactions were pursued by other sociologists. In 1937, Herbert Blumer coined the phrase *symbolic interaction* to refer simultaneously to how humans communicate and to a sociological perspective. He applied that perspective to a wide range of research areas such as social psychology, collective behavior, race relations, and social problems (Maines 1989). Blumer's posthumous volume on industrialization and social change (1990) elaborates the Thomas and Znaniecki formulations by presenting a conceptualization of causal influences that hinge on human agency and interpretation. Stone's research on clothing and fashion (1962) similarly focuses on human behavior as transactions. In particular, his analysis of identity establishment identifies dialectical processes of communication through which individuals are located and placed in the social organization of society. His treatment of interpersonal identities is sympathetic to Cooley's emphasis on the importance of primary groups, while his treatment of structural relations corresponds with Thomas and Znaniecki's concept of values and predates contemporary social psychological research on individuals and social structure.

Studies of social organization have been directly influenced by pragmatist principles, as previously mentioned, but that influence has been especially apparent since the early 1970s (see Hall 1987 and Fine 1993 for summaries). Anselm Strauss's research on occupations and formal organizations has been prominent in this recent work and has led to the development of the "negotiated order" perspective (Strauss 1978). Negotiations, he argues, are processes through which collective actions occur and tasks are accomplished. These processes are influenced by actor characteristics, the immediate situation, and larger structural contexts. However, those larger contexts are also influenced reciprocally by actual negotiations and their situations. Strauss's model of social organization thus is a recursive and dialectical one. Stryker (1980) presents a slightly different version but one that is no less pragmatist. He incorporates

traditional role theory to argue that social structural arrangements limit options and opportunities by channeling people into status and role positions. So located, people construct their identities in terms of social meanings that are hierarchically organized. Both Strauss's and Stryker's applications of pragmatism have stimulated considerable research and theoretical development.

Such development has continued with some vigor throughout the 1990s. Pestello and Saxton (2000) draw on explicit principles of pragmatism to reframe the study of deviance in terms of the dialectics of inclusion and exclusion. While closer to the classical statements, they share with the neo-pragmatists such as West, Seigfried, and Rorty the interest in reclaiming the vision of emancipatory democracy that has always been present in the perspective. Other scholars have used pragmatist principles to further develop areas of sociological theory. Fine (1992) addresses the relationship between agency and structure, and theorizes both the obdurate and interpretive dimensions of the relations among contexts of action. Musolf (1998) shows how agency and structure relations are expressed in an array of standard sociological areas (socialization, gender, deviance, power, society), while Hall (1997) offers an updated version of his earlier conceptual framework (Hall 1987) that allows analysts to better specify dimensions of agency-structure relations. Five dimensions are identified: strategic agency, rules and conventions, structuring situations, culture construction, and empowering delegates. He then shows how power is expressed along these dimensions and links together situations and contexts.

In a related approach to matters of agency and structure, Strauss (1993) builds on his earlier theory of negotiations to focus analysis on ordering processes. He begins by articulating over a dozen assumptions drawn from pragmatism that guide his analysis. He then presents a "processual order" theory that focuses on how larger scale processes and structures condition organizations and situations within which people interpret and construct their conduct and how local decisions have consequences for larger scale processes and structures. This theory has been used in research by Fischer and Dirsmith (1995) to examine organizational strategies and technology use in large accounting firms. Ulmer (1997) also has used the theory to

explain how state sentencing guidelines are filtered through state and local political processes, and how local court communities set “going rates” for actual criminal sentencing. Ulmer’s research is especially relevant because it simultaneously presents data on statewide sentencing outcomes and the contextual practices that produce those outcomes.

Yet another related line of development has focused on what W. I. Thomas ([1927] 1966) called “situational analysis,” which examines the interplay of actors’ interpretations and the obdurate qualities of situations. Katovich and Couch (1992) draw on Mead’s theory of time to show how pasts are used by people to become socially situated. Situations, they argue, are not merely settings in which human conduct occurs, but rather are forms of conduct themselves and exist as transactions of pasts and futures. Gusfield (1996) presents his long line of research on alcohol problems, and addresses how those problems are constructed. He discusses various claims-makers within the alcoholism movement, and then dissects actual alcohol problems in terms of their situations—situations of drinking, of driving, of accountability—and the ideologies and logics-in-use that connect them. Hall and McGinty (1997) analyze educational policy processes from the “transformation of intentions” perspective. In a statewide study of the Missouri career ladder program, they show how various contexts (state legislature, advisory committees, bureaucratic offices, school districts, local schools) influence the original intended policy. Actual policy effects, accordingly, are strongly influenced by situational contingencies that render policy processes themselves less-than-rational. Similarly, Deutscher and colleagues (1993) analyze the contradictory findings from decades of attitude-behavior research, and provide a situational approach for explaining the inconsistencies between attitudes (what people say about themselves) and behavior (what people do). To understand consistency and inconsistencies between attitudes and behavior, they argue, we must understand the situations that produce those kinds of relationships.

One of the distinctive characteristics of social science research and theory that has been grounded in pragmatism is the reluctance to give credence to dualisms such as micro-macro or individual-society. While issues of scale have always been acknowledged and used, as the work of Park, Blumer,

Stone, Stryker, Strauss, Hall, and Fine has demonstrated, the focus has been on the examination of social processes that produce, maintain, and change social orders. These social processes have generally been conceptualized as communicative in nature, and the central thrust has been on how large- and small-scale phenomena are simultaneously or similarly transacted by individuals and groups. The focus on those processes has maintained the action orientation of pragmatism, and the central precepts of the perspective are finding increasing currency and relevance in contemporary work in the social sciences and humanities.

(SEE ALSO: *Social Philosophy*; *Social Psychology*)

## REFERENCES

- Bernstein, Richard 1986 *Philosophical Profiles*. Philadelphia: University of Pennsylvania Press.
- Blumer, Herbert 1990 *Industrialization as an Agent of Social Change: A Critical Analysis*. Hawthorne, N.Y.: Aldine de Gruyter.
- Carey, James 1989 *Communication as Culture*. Boston: Unwin Hyman.
- Dallmayr, Fred 1987 *Critical Encounters between Philosophy and Politics*. Notre Dame, Ind.: University of Notre Dame Press.
- Denzin, Norman 1992 *Symbolic Interactionism and Cultural Studies*. Cambridge, Mass.: Blackwell
- 1996 “Sociology at the End of the Century.” *Sociological Quarterly* 37:743–752.
- Deutscher, Irwin, Fred Pestello, and H. Frances G. Pestello 1993 *Sentiments and Acts*. Hawthorne, N.Y.: Aldine de Gruyter.
- Dewey, John 1896 “The Reflex Arc Concept in Psychology.” *Psychological Review* 3:363–370.
- 1925 *Experience and Nature*. Chicago: Open Court.
- Diggins, John Patrick 1994 *The Promise of Pragmatism: Modernism and the Crisis of Knowledge and Authority*. Chicago: University of Chicago Press.
- Farberman, Harvey 1991 “Symbolic Interactionism and Postmodernism: Close Encounter of a Dubious Kind.” *Symbolic Interaction* 14:471–488.
- Faris, Ellsworth 1928 “Attitudes and Behavior.” *American Journal of Sociology* 33:271–281.
- Fine, Gary Alan 1992 “Agency, Structure and Comparative Contexts: Toward a Synthetic Interactionism” *Symbolic Interaction* 15:87–107.

- 1993 “The Sad Semise, Mysterious Disappearance, and Glorious Triumph of Symbolic Interactionism.” *Annual Review of Sociology* 19:61–87.
- Fischer, Michael, and Mark Dirsmith 1995 “Strategy, Technology, and Social Processes Within Professional Cultures: A Negotiated Order, Ethnographic Perspective.” *Symbolic Interaction* 18:381–412.
- Gusfield, Joseph 1996 *Contested Meanings: The Construction of Alcohol Problems*. Madison: University of Wisconsin Press.
- Hall, Peter 1987 “Interactionism and the Study of Social Organization.” *Sociological Quarterly* 28:1–22.
- 1997 “Meta-Power, Social Organization, and the Shaping of Social Action.” *Symbolic Interaction* 20:397–418.
- , and Patrick McGinty 1997 “Policy as Transformation of Intentions: Producing Program from Statute.” *Sociological Quarterly* 38:439–467.
- James, William 1890 *Principles of Psychology*. New York: Henry Holt.
- 1904 “Does Consciousness Exist?” *Journal of Philosophy, Psychology, and Scientific Method* 1:477–491.
- Joas, Hans 1993 *Pragmatism and Social Action*. Chicago: University of Chicago Press.
- 1996 *The Creativity of Action*. Chicago: University of Chicago Press.
- Katovich, Michael, and Carl Couch 1992 “The Nature of Social Pasts and Their Use as Foundations for Situated Action.” *Symbolic Interaction* 15:25–47.
- Konvitz, Milton, and Gail Kennedy (eds.) 1960 *The American Pragmatists*. Cleveland, Ohio: World.
- Lofland, Lyn 1998 *The Public Realm*. Hawthorne, N.Y.: Aldine de Gruyter.
- Lyman, Stanford 1997 *Postmodernism and the Sociology of the Absurd*. Fayetteville: University of Arkansas Press.
- McCarthy, Thomas 1984 *The Critical Theory of Jürgen Habermas*. Cambridge, England: Polity.
- Maines, David 1989 “Repackaging Blumer: The Myth of Herbert Blumer’s Astructural bias.” In Norman Denzin, ed., *Studies in Symbolic Interaction*. Greenwich, Conn.: JAI.
- 1996 “On Postmodernism, Pragmatism, and Plasterers: Some Interactionist Thoughts and Queries” *Symbolic Interaction* 19:323–340.
- (ed.) 1997 “Interactionism and Practice.” *Applied Behavioral Science Review* 5:1–139. (Special Issue.)
- , Jeffrey Bridger, and Jeffery Ulmer 1996 “Mythic Facts and Park’s Pragmatism: On Predecessor-Selection and Theorizing in Human Ecology” *Sociological Quarterly* 37:521–549.
- , Noreen Sugrue, and Michael Katovich 1983 “The Sociological Import of G. H. Mead’s Theory of the Past.” *American Sociological Review* 48:151–173.
- Martindale, Don 1960 *The Nature and Types of Sociological Theory*. Boston: Houghton Mifflin.
- Mead, George Herbert 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Musolf, Gil Richard 1998 *Structure and Agency in Everyday Life*. Dix Hills, N.J.: General-Hall.
- Park, Robert 1926 “The Urban Community as a Spatial and Moral Order.” In Ernest Burgess, ed., *The Urban Community*. Chicago: University of Chicago Press.
- Peirce, Charles 1877–1878 “Illustrations of the Logic of Science.” *Popular Science Monthly* 12:1–15, 286–302, 604–615, 705–718; 13:203–217, 470–482.
- Perinbanayagam, Robert 1986 “The Meaning of Uncertainty and the Uncertainty of Meaning.” *Symbolic Interaction* 9:105–126.
- Pestello, Frances, and Stanley Saxton 2000 “Renewing the Promise of Pragmatism: Towards a Sociology of Difference.” *Studies in Symbolic Interaction* 23:(In press.)
- Reitzes, Donald, and Dietrich Reitzes 1992 “Saul Alinsky: An Applied Urban Symbolic Interactionist.” *Symbolic Interaction* 15:1–24.
- Ricoeur, Paul 1985 *Time and Narrative*. Chicago: University of Chicago Press.
- Rochberg-Halton, Eugene 1986 *Meaning and Modernity: Social Theory in the Pragmatic Attitude*. Chicago: University of Chicago Press.
- Rorty, Richard 1982 *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press.
- Rosenthal, Sandra 1986 *Speculative Pragmatism*. Amherst: University of Massachusetts Press.
- Rucker, Darnell 1969 *The Chicago Pragmatists*. Minneapolis: University of Minnesota Press.
- Seidman, Steven 1996 “Pragmatism and Sociology: A Response to Clough, Denzin, and Richardson.” *Sociological Quarterly* 37:753–759.
- Seigfried, Charlene Haddock 1996 *Pragmatism and Feminism: Reweaving the Social Fabric*. Chicago: University of Chicago Press.
- Shalin, Dmitri 1986 “Pragmatism and Social Interactionism.” *American Sociological Review* 51:9–29.
- Stone, Gregory 1962 “Appearance and the Self.” In Arnold Rose, ed., *Human Behavior and Social Processes*. Boston: Houghton Mifflin.
- Strauss, Anselm 1978 *Negotiations*. San Francisco: Jossey-Bass.

- 1993 *Continual Permutations of Action*. Hawthorne, N.Y.: Aldine de Gruyter.
- Stryker, Sheldon 1980 *Symbolic Interactionism: A Social Structural Version*. Menlo Park, Calif.: Benjamin/Cummings.
- Thomas, W. I. 1966 "Situational Analysis: The Behavior Pattern and the Situation." In M. Janowitz, ed., *W. I. Thomas on Social Organization and Social Personality*. Chicago: University of Chicago Press. (Originally published in 1927 in *Publications of the American Sociological Society* 22:1-12.)
- , and Florian Znaniecki 1918-1920 *The Polish Peasant in Europe and America*, 5 vols. Chicago: University of Chicago Press.
- Ulmer, Jeffery 1997 *The Social Worlds of Sentencing: Court Communities under Sentencing Guidelines*. Albany, N.Y.: State University of New York Press.
- Van Den Berg, Axel 1996 "Liberalism without Reason?" *Contemporary Sociology* 24:153-158.
- West, Cornell 1989 *The American Evasion of Philosophy*. Madison: University of Wisconsin Press.
- 1994 *Race Matters*. N.Y.: Vintage.
- Wiley, Norbert 1995 *The Semiotic Self*. Chicago: University of Chicago Press.

DAVID R. MAINES

## PREDICTION AND FUTURES STUDIES

### PREDICTION AND SOCIETY

**Meaning.** Prediction (*previsione*) is "seeing beforehand" how the future will be; that is, the situation that will come about in the short, medium, and long term. This, however, is a "seeing beforehand" that is not content with simply knowing what the final situation will be but it is also concerned with knowing with how it will be reached. It is concerned not only with the end-point but also with the process that the present situation will undergo to transform itself into (or remain as) the end-point. This two-part nature of prediction (end-point and development/process) enables us to take prediction into a scientific dimension, because the process by which the end-point is reached is comprehensible by means of the elaboration of a standardized method of identifying variables and their positions in a model of relationships that

connect the before to the after, and carry the before to the after, and lead from the starting point to the end-point. Without this scientific attention to the passage from the former point to the latter, prediction could simply be a matter of taking the present and reading its evolution by means of an external "operator," such as a magic formula, casting stones, or the interpretation of animal innards.

This scientific attention to the process leading to the end-point—the prediction—is so crucial because by identifying elements of the process we can modify their developments to bring about the prediction we postulated. In other words, the variable now to be discovered and defined is not so much the prediction (with the extrapolative method) as the scientific process to be predicted so as to achieve a situation we desire (with the normative method).

In the light of the above, the term "prediction" as a name for this discipline is clearly less appropriate than others, such as "futurology," "future thinking," or "futures studies." A distinction may be drawn between two of these in that "futurology" is more of a disciplinary name and "futures studies" is more of an agenda of scientific activity aimed at achieving and discriminating between possible, probable, and desired futures.

In fact, if prediction is all of these things, it is of fundamental importance for taking action; that is, for deciding a priori how to achieve a given objective or to discover what a particular action will lead to.

Prediction is basically a scientific process in both of the senses just described, but it is also strongly related to other cognitive dimensions, such as ideology, ideas of utopia, action, and change. The theories underpinning prediction therefore combine or conflict with the theories and cognitive dimensions of reality. The links between these concepts should therefore be made clear.

**Prediction, Utopia, and the Future of Traditional Society.** A link of some kind between prediction, ideas of utopia, planning, and change certainly exists. We shall try to highlight it, starting from the problematic identified for prediction. Utopia is a "nonplace" in which is located a perfect society (perfect for he who first conceives it) dominated by a "cold synchrony," that is, by mechanical

relations that serve only to maintain the utopian system without creating the emotional “warmth” or interests, including conflicting interests, whose relational outcome is unpredictable and therefore transforms the time-one system into something different from the time-zero system. Utopia is therefore the end-point of prediction, static and without relations, whose outcome cannot be predicted—an absolute end-point beyond which is nothingness. And between the present and the utopian state is “non-sense,” that is, a black box whose contents do not interest us—the contents being the process that enables us to pass from the past to utopia.

What relation is there between prediction and utopia? We can interpret utopia as a residue of traditional society in which everything is tied to the past, but experienced in the pre-modern European society of the fifteenth and sixteenth centuries. In traditional society, real prediction, and therefore the future situation, is *deduced* by the past: Action is prescriptive, change is an aberration, organizations are similar because they all have the same structural contents and perform the same functions. Everything comes from the past: Rules are written in the past and actions are already perfected in the past. In these conditions, the process generating the future is a *dejà-vu*—not a perfect one, but a human condition and a destiny. The end-point that is different from such a prediction is not of this earthly condition but of the other life, in heaven or in hell. Utopia represents a sort of rebellion against the placing of heaven beyond earthly life—it is the secular dream of human omnipotence because it dares to place social perfection “not here” but nonetheless on earth. This utopia, an expression of the rebellion—no matter how fantastic—of traditional man, comes to represent a piece of “heaven” brought to earth, in which relations between people and social structures are so “sweet and delicate” as to strengthen the sweetness we have inside us rather than producing new situations and without automatically creating new equilibriums and new states. Because of all this, we can understand why constructors of utopia do not need to know the process enabling us to pass from the present to the future (because it is a copy of the past) and why it is dominated by a static equilibrium that does not change once it has been achieved.

Thus prediction, utopia, and change in a society in which change is an aberration mean at the

most bringing the perfection of the nonearthly world into the earthly world, but leaving it detached from reality, which also remains immutable. In other words, prediction is a game left to forces that are untameable and therefore ineluctable and at bottom mechanical, perpetuating positive and negative flows *ad infinitum*.

**Prediction and the Objective as a Reference for the Plan: Ideology and the Future of Modern Society.** Prediction becomes practically useful—able not only to reveal what will be but how this future may be controlled—when we lose sight of the perfections of the state we have called utopia and it takes on the role of an objective to be striven for, when an active value is ascribed to social ideals and single individuals’ capacity for action. Here, society activates ideology as a resource and at the same time recognizes the ability of individual action, and above all the synthesis of individual actions, to create new situations.

Modern society thus shifts the focus from *the perfection of utopia*, which needs no modification since it is by definition perfect, and which is (an unreachable) vaguely defined objective, to the *laboratory of process*, which is concerned with relations and objectives to be achieved. If such an objective happens to be clearly defined, it is so accepted as provisional and therefore “adjustable,” because certainty is only to be found in highly generic values such as justice, equality, and self-fulfillment in a fair, egalitarian, individual-enhancing society. This process of achieving the desired or probable prediction is guided by two resources activated by society. The first is ideology, the cognitive representation of the world used to guide practical action toward the objective that is the subject of prediction. The second resource is trust in the individual whose initiative may contribute to achieving the prediction, with the proviso that the action of this individual must be combined with that of other individuals to thus produce positive results for the predicted state of interpersonal and social relations.

In this view of modern society, it seems that the focus—aside from the objective defined in the probable or desired prediction that in its most complete form takes on the configuration of a plan—shifts for the most part to the process from which the prediction (and the predicted plan) emerges and thus to how this process is rationally



manifested, how it may be scientifically explained, and what may be done to bend it to the achievement of the prediction.

All this comes about in modern society because change is conceived as normal, a factor built into the trajectory toward the future, tendentially and plausibly different from the present and above all from the past. All that remains of the past is the genetic origin of the present and a certain limited influence on it. If we have thus conceptually severed the deterministic link (at least in ideological terms) between past and future, modern society clearly has to focus very sharply on the processes and interdependent relations of the present in order to predict and dominate the future.

#### **Determinism and Creativity in Prediction.**

To understand how prediction is to be orientated and manifested, rationality and the scientific method are essential because rationality and the scientific method provide the most effective and efficient ways of bringing about the realization of what we want to happen or what “must” happen. Hence the importance of method in obtaining a prediction and controlling it.

Methods have both deterministic components and creative components that are selected and embedded into techniques proper; these will be considered below. For the time being, it is sufficient simply to mention some features of these components. Deterministic components ground the formation of the end-point (prediction) in relations among the social, economic, environmental, and value structures defined in a model. Creative components include those that highlight the identification and pursuit of new and “invented” ways of controlling or accelerating the achievement of a prediction. Around such components, objective or subjective methodological techniques are developed that highlight the workings of a model and its simulation or formation of decisions.

### **METHODS OF PREDICTION**

**The Scientific Problem of Prediction.** Not only is the scientific nature of social sciences considered suspect by people outside the social sciences, but certain social scientists themselves consider the social disciplines nonscientific because the most they can do is provide a way of “reading” a social reality composed of individuals, groups, mutual

relations, and formal organizations. For such people, the social sciences are not sciences but opinions. The reason is that—apart from some concepts providing interpretative keys for human, social and organizational action, mutual relations, and the products of those relations—interpretative theories stand the test of falsifiability only for a short time, often only until an event outside the phenomenon under investigation undermines the equilibrium and internal stability achieved by the phenomenon and explained by the theory. In the short term, the interpretative weakness of the theory even throws doubt on the ability of social science to explain the phenomenon. And this weakness of explanation obviously affects the strength of the prediction and consequently what has to be done to change it, that is, what is to be done to carry the present into the future.

Yet particular attention has been focused on four of the activities or purposes of which science is composed: (1) description as a pre-scientific stage and (2) explanation, (3) prediction, and (4) control as scientific activities proper. In point of fact, causal explanation is the central activity of the scientific process, since prediction is deduced from explanation and control is a “political” manipulation (and as such outside the phenomenon explained) of the variables of the explanatory model, undertaken deductively to achieve a modification of the prediction.

It may therefore be said that induction is at the root of description and explanation and deduction is at the root of prediction and control. But it is for precisely this reason that the first two activities are “more scientific,” in that they are caught up in the bond between theory and theory-testing empirical research, whereas prediction and control are more rooted in utilization and change, in the final analysis in technical application. It is probably in this logic that we should see the contradiction between prediction as a science, whereby methods and techniques are elaborated as a deductive extension of methods for description and explanation or the perfection of methods, and prediction as techniques to help elites who have to make decisions to modify predictions and the explanatory picture deriving from them.

In other words, making predictions becomes scientific activity on the basis of data that are

absent but that are plausible or possible or probable or desired, and whose relations may give rise to situations and scenarios that are equally possible, probable, and desired, but not certain.

The scientific nature of prediction is therefore based on rationality and the logic implicit in the links between events that have already come about and implicit in the possible reactions of, or to, a behavior that may come about. It is thus a matter of *reasoning by analogy*: Such-and-such has happened before in certain situations, so it may happen now in similar situations.

The scientific nature of prediction is also based on the fact that from the level of spatial analogy (if we have verified that an effect comes about *here*, we may infer that it will come about *there* in culturally analogous conditions) we may pass to the level of temporal analogy (if we have verified in the causal explanation that something comes about *today*, we may infer that it will come about *tomorrow*).

In more general terms, in prediction there is a “low-profile science,” which becomes the rational study of what could happen in the future and above all how this might be more adequately dealt with so as best to guide or govern it.

Given these epistemological premises for prediction, the techniques that manifest its methodological paths are the result of the combination of certain features of the methods: qualitative and quantitative, those based on objective data or the opinions of elites (of power or knowledge), and extrapolative or normative. The predictive specificity of these three features increases from the first to the third. *Quantitative and qualitative* refer to the level of research and knowledge concerning a given phenomenon; the higher the level, the greater the chance of having indicators that are tried and tested and therefore more practically defined. *Objective data and elite opinions* are more closely tied to the usefulness ascribed to prediction. The objective datum reconstructs the model in a system, identifies the causal process and objective to which it leads, and starts from the assumption that it is “technically” possible to act on the structure of the process to modify its consequences. Leaders’ opinions are privileged in that the basic assumption is that it will be their ideas and expectations, “true” or “false” as they may be, that condition, or even produce, the change in the

objective/end-point. In the first of these approaches, there is extreme confidence in the scientific character of the epistemological canons of science; in the second, there is substantial lack of confidence that science can produce reality control—it becomes merely a formal exercise, though rationally useful. The *extrapolative and normative* features introduce action aimed at controlling the objective/end-point: Extrapolation is the projection of present processes and mechanisms into the future to postulate how it may possibly and probably be configured; the normative feature is fixing the desired future to identify the processes and mechanisms to achieve it. This is a dual approach to the future that is complementary rather than contradictory, in that its second (normative) part begins where the first (extrapolative) has served its purpose.

Table 1 shows the various techniques of prediction in relation to the three dimensions defined by the three features: qualitative–quantitative, objective–leaders’ opinions, and normative–exploratory.

This plotting enables us to make some statements and develop some assessments of methods, especially of their function in establishing the various dimensions of prediction in terms of future studies.

**Objective and Quantitative Techniques.** Objective and quantitative techniques are common to all research. They are standardized and consolidated in practice. Scenarios, time series, causal models, simulations, and so on provide potent instruments for translating quantitative results into explained variance and high probability, although their considerable rigidity leaves them unable to cope with interference from new and sudden exogenous variables. These highly statistical methods are more effective in short-term prediction and for circumscribed rather than global events. They are also used in combined form to achieve both exploratory and normative prediction. Here the methodology may be used to explain the causal process by means of which a given trend is manifested, after which a “mission” is decided upon—a defined objective, such as a plan—and modifications are introduced into the contextual variables and their relational flows in order to achieve the objective. An example of these combined methodologies may be found in the research carried out

<b>Methods of Prediction in the Combination of the Three Criteria</b>			
		<i>normative method</i>	<i>extrapolative method</i>
<i>objective data</i>	<i>quantitative</i>	scenarios	scenarios
		gaming and simulation	time series
			regression analysis and canonical analysis
			econometrics and causal models
			nonlinear models
	<i>qualitative</i>		trend impact analysis
			cross impact analysis
			gaming and simulation
		scenarios	scenarios
		relevance trees	gaming and simulation
<i>opinions of leaders</i>	<i>quantitative</i>	science fiction	
		scenarios	scenarios
		Delphi	Delphi
		cross impact analysis	cross impact analysis
			trend impact analysis
	<i>qualitative</i>		
		"expert group meetings," in-depth interviews, "genius forecasting"	"expert group meetings," in-depth interviews, "genius forecasting," intuitive logic
		intuitive logic	
		Delphi	Delphi
		cross impact analysis	cross impact analysis
	scenarios	scenarios	

Table 1

on the Italian situation by Alberto Gasparini under the auspices of the Istituto di Sociologia Internazionale di Gorizia (ISIG) (Gasparini et al. 1983, 1988).

The first research project set out to identify how to meet the housing needs of a metropolitan area in accordance with a preestablished norm, which in this case was the habitation standard expressed in terms of the acceptable ratio between living space and family sizes. Housing needs were identified by applying several habitation standards,

after which exploratory techniques were used, simulating natural demographic and social trends, simulating new housing markets and vacant housing markets, introducing factors such as filtering and vacancy chains. The result was the determination of how much of future housing needs (for example, ten years later) would remain with no intervention, that is, by leaving the area's housing situation to develop under its own impetus. Subsequently interventions were carried out on single variables: incentives for building new housing or for putting vacant housing on the market or for

restoring existing housing or for providing financial help for needy families, and so on. The observation of the effects of interventions on the variables of the housing needs model indicated whether the objective had been achieved and, if it had not, provided indications as to the most appropriate modifications to be applied to single variables in order to achieve the objective.

The second project was on the quality of the environment in daily life in Italian towns. The questions were the following: What type of environment is it? How can quality be defined? What is the current state of environmental quality? How can high environmental quality be achieved in the daily life of the town? To answer these questions, prediction was developed over the following stages: (1) Quality of the environment was defined according to people's expectations in terms of services (number) and their spread or concentration in the town area, all placed in relation to individual and community values expressed by the local inhabitants. Surveys were carried out in each town by giving a questionnaire to 137 samples in as many communities, for a total of 33,000 interviewees. The result was the *model of desired environmental quality for daily life* as derived from subjective data converted into objective data. (2) This desired model was applied to each town (how many and what services existed and their location), which gave the *model of environmental quality for daily life lived*. (3) Observations were carried out on the context in which the above model was placed and the variables producing it, in order to identify the variables that influence the quality of the environment in it. These independent variables (clusters of multiple variables reduced in number by factor analysis and causally related to environmental quality through canonical analysis) represented the various features of community life: population, town territory, values, economic structure, social structure, local government, endogenous resources, exogenous resources, communications with the outside world, and so on. (4) The achievement of the desired environmental quality for daily life was explored by intervening on the variables that were causally most important for environmental quality (as they emerged from stage 3) and by simulating the effects that these interventions would have on environmental quality. This may entail further interventions on single variables until the

achievement of the desired environmental quality, which is the subject of the prediction.

A third project was the definition of task environments and their dynamics in agricultural production organizations. This research was basically exploratory in nature, concluding with normative assessments. The exploration was not carried out by inquiring into how company task environments are modified over time (an inquiry into process), because the starting theory (to be subjected to verification) was that proposed by Emery and Trist (1965), whereby task environments are modified as companies expand and become increasingly causally important and disruptive, introducing irrationality into the decisions companies have to make. In this research, then, predictive exploration was not based on the projection of variables into the future, but on the investigation of two situations (small-company and large-company task environments) and their comparison in accordance with the Emery-Trist theory, reconstructing the dynamics by comparing the two potential stages of a single company that grows from a small one into a large one (Gasparini 1983). By means of defining the role of the agricultural entrepreneur and his relations with the task environment organizations, synthesising these into a few factors (by factor analysis) and linking them through canonical analysis, task environments were identified and articulated according to their influential and direct relations with the company, according to contacts not generating real influences in that relations were based on sporadic contacts, and so on. One substantial difference emerged between small and large companies. In small companies, there are few influential relations and a great many casual and sporadic contacts. In large companies, the relational task environment is very rich in relations, influences, and dependency on company decisions; the structure of the sporadic contact task environment is by contrast marked by relational links that are weak and few and far between. This exploratory projection produced by the hypothetical transformation of a small company into a large one therefore showed a radical change in the functional and power relationships of the task environment. The identification of concrete relations and contacts and their respective influence clearly leads to the intention to intervene according to the normative objective, which in this case is a rethinking of

entrepreneurship, or an operational intervention in the agricultural economy to make sure that small and large entrepreneurs retain the power and responsibility assigned to them by the theory.

These three research examples show the great versatility of quantitative and objective techniques, that they need to be integrated with one another, and that they can be used in the exploratory dimension and some normative functions. These techniques are inextricably intertwined, as is exemplified by the fact that the exploratory dimension itself must be defined by reference to the *criterion* implicit in the normative dimension.

**Leader's Opinions and Qualitative Techniques.** The methods and techniques based on leaders' opinions, be they decision makers or experts in a particular field, are fundamentally qualitative in nature, that is, they are based on assessments that can be conventionally ordered in numerical values from which relations can then be highlighted. This can be done, as in the case of cross-impact analysis, but it should not be forgotten that the quantitative values manipulated are derived from percentages attributed intuitively to the occurrence of one event rather than another. Nevertheless, there are slightly differing degrees of formalization between these methods, and they are expressed in terms of their internal logic, reasoning experience that discriminates the more possible from the less possible, the ability to progressively refine judgments (the Delphi method), the compatibility between the reasoning and the context in which it is used, and the compatibility that has to give rise to the prediction for the phenomenon placed in the context.

This type of technique also contains scenarios, but they derive more from leaders' judgments than from the (highly implicit) model at the basis of the issues at stake and therefore of the variables defining the features of the model itself.

These are thus methods that can be used for the study and prediction of phenomena whose details are not known and/or which are relatively new, which means that recourse is made to qualified individuals equipped, for one reason or another, to see their own knowledge and predictions through the prism of research experience or familiarity with decision-making processes. If such is the case, the next step might be to transform the

results of these subjective predictions into indicators and formal explanatory models, to be tested with exploratory methods and normative methods to obtain a (concrete) measurement of the projection or process required to achieve the predetermined objective-norm-criterion.

But it may also be the case that the simple results derived from these opinions are considered sufficient (expressed to various degrees of sophistication by means of in-depth interviews, the Delphi method, cross-impact analysis and the qualitative scenario), and this happens because, or probably because, the scientific component in the prediction is not held to be very important; it is considered as a set of rational instruments for reasoning about the plausibility of the prediction itself. Taking into account that these rationalized judgments come from policymakers (at the summit of the decision-making process) or opinion makers, this ascientific factor is even more worrisome.

In this case, the implicit conviction is that these are the players who will have a major role in the achievement of their own prediction—in which case, we are faced squarely with the principle of self-fulfilling prophecies.

**Techniques in a Band of Abivalence.** From Table 1 we still have to analyze the two intermediate bands in which predictive techniques are to be placed. Though these are conceptually different in some respects, they are also instrumentally contiguous, which in practice means that they often overlap, or are at least complementary, when being used. Objective qualitative methods indicate that phenomena are analyzed structurally with no measured data; methods that are quantitative but based on leaders' opinions provide judgments strongly based on facts, or at least measurable data, which involves a strong tendency to apply leaders' opinions to a concrete context.

Prediction techniques placed in these two bands of ambivalence are very similar to those devised for the predictive analysis of leaders' opinions, but they rely heavily on a detailed knowledge of context. Thus they also make use not only of the Delphi method, cross-impact analysis, and scenarios, but also simulation in objective quantitative methods. However, the most typical of these two bands are relevance trees, science fiction, and tendency impact.

An example of predictive research in this context of ambivalence is condensed in what Igor Bestuzhev-Lada (1997) calls “technological prediction.” It comprises seven procedures that use methods which are both quantitative and qualitative, objective and subjective. The procedures are: program elaboration, construction and analysis of the starting model, construction and analysis of the predictive background model, exploratory prediction, normative prediction, prediction verification, and formulation of recommendations for a proper management of technological prediction. Indicators are often measured quantitatively, but their treatment and assessment are mostly qualitative.

In summary, the combination of the three criteria detailed in Table 1 indicates the following:

1. There are more specific techniques in the objective–quantitative methods and the leaders’ opinions–qualitative methods.
2. In the ambivalence band, methods that are typically objective and bound to opinion leaders tend to extend toward the quantitative and qualitative.
3. The exploratory and normative methods are not alternatives but are fairly well integrated with one another. An exploratory projection is implicit in the normative method, and the exploratory method requires a criterion that is able to be transformed into the desired prediction–norm.
4. Many methods are multivalent in prediction in that they are used to construct many types of prediction, but also in that they are technically versatile because they can be used with measured data and opinions alike. The most important example of this is the scenario.
5. The effectiveness of the methods varies according to the type of prediction in question. For short-term predictions and those on a circumscribed subject, the quantitative–objective method is most effective. The longer the period involved and the broader the subject, the more effective are qualitative methods and methods based on leaders’ opinion. This means that a broad to medium to

long-term framework analyzed with qualitative methods contains specific short-term subjects studied with mathematical formalization.

6. The scientific nature of prediction methods therefore varies with the variation of frames and times of reference, or at least there is a variation of the forms of expression of the scientific activities of description, explanation, and control.

From the above it thus emerges that the complexity of *prediction of the future* entails a multiplicity of studies because the future expresses itself in very different ways. This is why it is more than legitimate to speak in methodological terms of *futures studies*.

#### **THE DISTRIBUTION OF FUTURES STUDIES IN TIME AND SPACE**

It is well known that studying the future becomes a strongly felt need in times of transition. The rules of the past no longer hold, and the rules for the future do not yet exist or are still untried. In addition, globalization accentuates the need to build niches within which some form of autonomy may be regained. But what does niche autonomy mean when globalization bombards it with the upheaval external to it, upheaval which is therefore experienced as irrational?

This question gives rise to the need to predict and to achieve prediction through studies on the future or futures. It is hardly surprising that many public and private institutions have set up study centers for prediction. An indicator of this growth is provided by the large number of Web sites for such centers, which are being put together to form a *Futures Studies Internet Society*.

An ISIG study (Apuzzo et al. 1999) has found that worldwide two hundred and eight institutions working in futures studies have Web sites. Of these, one hundred and forty publish papers, one hundred and twenty-three provide links, sixty-nine publish their own journals on-line and conduct training, sixty-four advertise books, fifty-seven provide on-line shopping, forty-seven pass on chat news, thirty-eight make available projects, thirty-four are concerned with software, and twenty-eight deal with methods and techniques.

Most of the sites are American (one hundred and seventeen futures studies institutes) and only twenty-three are international; fourteen are in Britain; eleven in France; eight in Australia; six in Sweden; five in Canada; three each in Germany, Finland, Belgium, Norway, and Italy; two in Switzerland; and one each in Argentina, Denmark, Zaire, Austria, Israel, and Russia.

There can be no doubt that the future, especially in the United States, is strongly perceived as a subject requiring analysis. We have interpreted this as a way of finding autonomous futures for individual niches in the context of sweeping globalization. But it may also happen that despite this intention, the incipient Futures Studies Internet Society will accentuate the very standardization of feeling, thinking about, and planning the future that such great efforts are being made to curtail.

Bibliography

Aa.Vv 1987 *Futuro e complessità*. Milan: Angeli.

Apuzzo, Gian Matteo, Bruno Maltoni, and Moreno Zago 1999 "Futures Studies Bookmarks." *Quaderni di Futuribili* 3:98.

Bell, Wendell 1994 "Futuro." In *Enciclopedia delle Scienze Sociali*, vol. IV. Rome: Istituto Enciclopedia Italiana.

— 1997 *Foundations of Futures Studies*. London: Transaction publishers. Vol. I and vol. II.

—, and J. A. Mau (eds.) 1971 *The Sociology of the Future*. New York: Russell Sage Foundation.

Bestuzhev-Lada, Igor 1997 "I fondamenti metodologici della previsione tecnologica." *Quaderni di Futuribili* 1:110.

de Jouvenel, Bertrand 1967 *The Art of Conjecture*. New York: Basic Books.

Emery, F. E., and E. L. Trist 1965 "The Causal Texture of Organizational Environment." *Human Relations* 18:21-32.

Flechtheim, Ossip K. 1966 *History and Futurology*. Meisenheim am Glau, Germany: Verlag Anton Hain.

Gasparini, Alberto 1983 *Ambiente operativo e azienda agricola*. Milan: Angeli.

—, Aldo de Marco, and Roberto Costa (eds.) 1988 *Il futuro della città*. Milan: Angeli.

Glenn, Jerome C. 1994 *Introduction to the Futures Research Methodology Series*. Washington, D.C.: Unu Millennium Project Feasibility Study.

Godet, Michel 1979 *The Crisis in Forecasting and the Emergence of the Prospective Approach*. New York: Unitar, Pergamon Policy Studies.

— 1994 *De la anticipation a la accion*. Barcelona: Editorial Marcombo.

— 1995 "Global Scenarios: Morphological and Probability Analysis." In *1995 profutures meeting*, Ipts, Jrc, Ec: 17-30.

Gordon, Theodore J. 1992 "The Methods of Futures Research." *Annals* 522(July):25-35.

Irvine, J., and B. Martin 1984 *Foresight in Science; Picking the Winners*. London: Pinter Publishers.

Jantsch, Eric 1967 *Technological Forecasting in Perspective*. Paris: OECD.

Kahn, Herman, and Antony Wiener *The Year 2000—A Framework for Speculation in the Next 33 Years*. London: Macmillan.

Land, Kenneth C., and Steph H. Schneider (eds.) 1987 *Forecasting in the Social and Natural Sciences*. Dordrecht, The Netherlands: D. Reidel.

Malaska, Pentti 1995 "Survey of the Use of the Multiple Scenario Approach in Big European Companies Since 1973." In *1995 profutures meeting*, Ipts, Jrc, Ec.: 42-46.

Mannermaa, Mika, Sohail Inayatullah, and Richard Slaughter (eds.) 1994 *Coherence and Chaos in Our Uncommon Futures. Visions, Means, Actions*. Turku, Finland: World Futures Studies Federation.

Masini, Eleonora 1993 *Why Futures Studies?* London: Grey Seal.

— (ed.) 1994 "La previsione. Idee, protagonisti, nodi problematici." *Futuribili* 1:223.

—, and Giorgio Nebbia (eds.) 1997 "I limiti dello sviluppo. 1972-2022. Che cosa resta dopo 25 anni, che cosa resterà fra 25 anni." *Futuribili* 3:193.

McHale, John 1969 *The Future of the Future*. New York: George Braziller.

Meadows, Donella H., Dennis L. Meadows, Jurgen Randers, and William W. Behrens, III 1972 *The Limits to Growth*. New York: Universe Books.

Norse, David 1979 "Scenario Analysis in Interfutures." *Futures* October.

Ortegon, Edgar, and Javier E. Medina Vasquez (eds.) 1997 *Prospectiva: construccion social del futuro*. Santiago de Cali, Colombia: Ilpes.

Pocecco, A. (ed.) 1994 *Il futuro come progetto degli scienziati sociali*. Gorizia, Italy: ISIG (Institute of International Sociology).

Polak, Fred 1971 *Prognostic: A Science in the Making*. Amsterdam: Elsevier.

— 1973 *The Image of the Future*. Amsterdam: Elsevier.

Schwartz, Peter 1996 *The Art of the Long View. Planning for the Future in Uncertain World*. New York: Currency Doubleday.

Slaughter, Richard A. (ed.) 1993 "The Knowledge Base of Futures Studies." *Futures* 25:3.

Wilson, Ian 1978 "Scenario." In *Handbook of Futures Research*. Westport, Conn.: Greenwood Press.

ALBERTO GASPARINI

## PREGNANCY AND PREGNANCY TERMINATION

Pregnancy is one of the most important events in a woman's life. For many women, pregnancy defines, enhances, or determines their economic, social, or personal value. Pregnancy and childbearing can be highly desired for personal reasons, or to meet social expectations. In some circumstances, childlessness may be grounds for suspicion, divorce, or worse. Pregnancy in the wrong situation may be undesirable to an equivalent degree. Traditionally, the woman with an unacceptable pregnancy may be abandoned, lose her prospects for marriage, be forced into prostitution, or suffer an even worse fate. Other pregnancies, even though they occur within a socially sanctioned circumstance, may result in loss of opportunities for social or economic advancement. Control over the occurrence and outcome of pregnancy is integral to women's control over their lives. This concept was strongly stated at the International Conference on Population and Development in Cairo in 1994.

Menarche, the first menstrual period, marks the beginning of a woman's reproductive period, and menopause, the last menstrual period, marks the end. Menarche occurs between ages 8 and 14 for most well-nourished women, but may come later if there is malnutrition or chronic disease. If the first menstrual period follows ovulation (release of an egg or ovum from the ovary), the woman may be fertile before menarche has occurred. For most women, fertility increases rapidly in the first year after menarche, as ovulation becomes more regular, so that as many as 95 percent of women in their late teenage years may be fertile. Infertility in the first years after menarche may be due to malnutrition, malformations of the uterus and reproductive organs, hormone abnormalities

leading to anovulation, or genetic diseases. After the early twenties, fertility declines, as sexually transmitted diseases; endometriosis (occurrence in abnormal sites of tissue normally found in the lining of the uterus); and other inflammatory, infectious, and vascular diseases affect reproductive organs. Menopause occurs between ages 40 and 55 in most women, but most women over 40 are infertile, and very few live births occur to women over 45.

## PREGNANCY

Pregnancy is typically divided into trimesters, which are uneven in length. Obstetricians invariably refer to "weeks" of pregnancy, meaning the number of weeks from last menstrual period, or when the last menstrual period should have occurred. The first trimester lasts from 3 to 12-14 weeks, the second from 12-14 to 24-26, and the third from 24-26 to delivery. The terminology is confusing, as menstrual age (the time since the last period) is not the same as gestational age (the time since fertilization, the union of the sperm and ovum). Therefore gestational age will be 2 weeks less than menstrual age. The system evolved in an era when many or most women could recall a last menstrual cycle, but few knew when conception or fertilization occurred. Now, more women may be aware of the date when conception occurred, and some pregnancies occur after assisted reproductive technologies (ART), when the time of fertilization may be known exactly. Either way, pregnancy itself begins immediately after implantation. This occurs about 3 weeks after the menstrual period and about 1 week after fertilization (therefore the first trimester starts at 3 weeks). A blood pregnancy test becomes positive within a day or two of implantation, and a urine test becomes positive a few days later. By the time the menses are missed, 4 weeks after the last menses, a urine pregnancy test will be positive if fertilization occurred at the expected time.

Between 10 percent and 40 percent of pregnancies result in *spontaneous abortion* in the first trimester. The wide variation in incidence reflects difficulty in recognizing very early abortion as well as differences in abortion rates among different age groups. Most spontaneous abortions are the result of chromosomal or developmental abnormalities of the pregnancy and are unavoidable; 1-2 percent of pregnancies abort spontaneously in



the second trimester, often due to infection, and about 1 percent of pregnancies have serious anomalies. About 1 percent of pregnancies are *ectopic*, occurring outside of the uterus. Most *ectopic pregnancies* require medical or surgical treatment to avoid hemorrhage. Ectopic pregnancy is one of the major causes of maternal mortality; in the United States, it is still associated with about 10 percent of maternal deaths (Cunningham et al. 1997).

### RESPONSE TO PREGNANCY

Women typically pass through several stages of reaction to pregnancy. The first reaction may be denial, particularly when the pregnancy is unwanted, or delight, when the pregnancy is wanted. Later, desired or planned pregnancy may still be accompanied by periods of ambivalence or anxiety (Affonso 1997). The concern may be about personal or work changes, financial stresses, health concerns, or concern about fetal health. At some point in the pregnancy, typically first or second trimester, the pregnancy is usually accepted. The fetus is visualized as a baby, and planning for birth and beyond ensues (Taylor 1980).

The pregnancy may not be accepted; the woman may choose to abort the pregnancy, or she may deny existence of the pregnancy, or she may never accept it even though she eventually gives birth. The lack of acceptance may follow a situation where abortion was unavailable; this may be an individual situation or a matter of political policy. During wartime, impregnation of "enemy" women (which may be accompanied by denial of abortion) is a form of terrorism (Swiss 1993).

Pregnancy itself does not impart any ability to deal with stress, nor does it decrease ability to manage stress. Women who have difficulty caring for themselves adequately when not pregnant, such as adolescents and women with severe psychiatric or developmental handicaps, may have difficulty caring for themselves while pregnant, to the detriment of the fetus. Pregnancy may be an incentive to stop self-abusive behavior such as drug use, but while pregnancy may be the motivating factor, it does not supply the emotional organization to change behavior. Domestic abuse of women does not usually stop with pregnancy; it often escalates (American College of Obstetricians and Gynecologists 1995b).

### PREGNANCY AND BIRTHRATES

The *pregnancy rate* is the number of pregnancies occurring per 1,000 women of reproductive age (considered to be ages 15–45 inclusive), per year. The *fertility rate* is the number of all births (liveborn and stillborn) per 1,000 women age 15–45 per year; this is the *birth rate for reproductive-age women*. The term *birthrate*, unmodified, usually means the number of live births per 1,000 population (male and female of all ages).

Pregnancy rates increase during the teenage years and generally peak during the late twenties in most societies, before declining in the thirties and forties. Birthrates follow a similar trend, although the peak birthrate may occur later than the peak pregnancy rate, and the difference between pregnancy rates and birthrates is more pronounced at the extremes of reproductive life. At the beginning and the end of reproductive life more pregnancies are likely to be terminated by abortion.

More pregnancies than are wanted occur. Some are mistimed, occurring earlier than wanted, and some are unwanted at any time. In the United States, between 40 percent and 50 percent of pregnancies are unintended (Horton 1995; Henshaw 1998a). Overall, mistimed pregnancies are more frequent than unwanted pregnancies, but the relationship may be reversed in some groups, such as women over 40. Rates of unplanned pregnancy are generally lowest in countries with wide availability of effective contraception combined with public education, such as western Europe (Paul 1999). Unplanned pregnancy may occur because of contraceptive failure or non-use of contraception. In the United States noncontracepting women, about 8 percent of women at risk for unintended pregnancy, account for about 50 percent of unplanned pregnancies (Gold 1990).

Unplanned pregnancy may result in birth, spontaneous abortion, or induced abortion. The *abortion ratio* (ratio of pregnancies aborted to pregnancies occurring) in the United States for unplanned pregnancy is about 0.5 (Gold 1990). In other countries the rates may be lower, where unplanned pregnancies are accepted or abortion is not readily available, or higher, where unplanned pregnancy is not tolerated personally, socially, or economically (Henshaw et al. 1999).

Teenage pregnancy rates have been of concern because of adverse medical and economic outcomes, which are intertwined (Fraser et al. 1995). For most of the century, pregnancy rates for U.S. teenagers have been declining. Rates began to rise in the 1980s, reaching a recent high of 117 per 1,000 (women ages 15–19) in 1990 (Alan Guttmacher Institute 1999). By 1997 rates had declined about 10 percent to 97 per 1,000. Birthrates have shown a parallel decline from 60 per 1,000 to 54 per 1,000, and abortion rate declined as well.

In contrast to teenage women, women in their twenties have higher pregnancy and birthrates. Birthrates are currently about 110 per 1,000 women in this age group (National Center for Health Statistics 1999). Crude abortion rates are similar to those of adolescents, 20–30 per 1,000 women. Since adolescents have fewer pregnancies and births, the abortion ratio is higher than for women in their twenties or thirties. Overall, more than 40 percent of teenage pregnancies resulted in induced abortion. There are large variations in pregnancy, birth, and abortion rates by state, ethnic group, and age (Horton 1998). Nineteen-year-olds have a much higher pregnancy rate than 15-year-olds. Nonwhite and black teenage pregnancy rates are approximately twice those of white teenage rates. Nonwhite and younger teenagers are more likely to have induced abortions (National Center for Health Statistics 1999).

Comparison of birth and abortion rates can be cumbersome. Distribution of women's age, ethnic group, geographic location, time of collection of statistics, event definition, and accuracy of reporting will affect rates.

### CONDUCT OF PREGNANCY

A comprehensive discussion of pregnancy is outside the scope of this encyclopedia; for more complete information, an obstetrics textbook (e.g., Cunningham et al. 1997; Gabbe et al. 1996) should be consulted.

The first trimester of pregnancy is often accompanied by nausea and vomiting, which is typically short-lived but may be severe enough to interfere with daily activities. Some women have hyperemesis, vomiting which is severe enough to result in dehydration or even death; it is treated

with intravenous fluids, antinausea drugs, and sometimes intravenous feeding. Although a link between hyperemesis and ambivalence or anxiety about pregnancy has been postulated, the strength and the relevance of any such link is uncertain.

Women often note changes in food likes and dislikes, but there is no medical reason to restrict types of food. There are often cultural restrictions and prescriptions, and there is seldom any reason to interfere with custom. In the second half of pregnancy many women have heartburn, because of the pressure of the pregnancy on the bowel and stomach, and because of relaxation of the esophagus. Changes in diet or eating habits, or using antacids, may help.

Women need additional calories during the second 2 trimesters; the amount is about 300 kcal per day for most women, the equivalent of a modest sandwich. The additional calories should consist mostly of complex carbohydrates; most Americans have adequate protein intake. Women who receive protein supplements may have higher rates of preterm birth than women receiving carbohydrate supplements or no supplements at all. The U.S. Department of Agriculture administers a food supplement program called WIC (Women, Infants, and Children), which provides food to low-income pregnant and lactating women, infants, and children up to 5 years of age. Supplementation is associated with small decreases in the rate of low-birthweight delivery and a decrease in cost of caring for newborns (Merkatz 1990).

Folic acid, a vitamin, has been shown to decrease the incidence of neural tube defects, a complex of birth defects of the brain or spine. Since the neural tube fuses in the fifth menstrual week, often before pregnancy is noticed, women should take adequate folic acid before pregnancy (Czeizel and Dudas 1992; Rosenberg 1992). The recommended daily allowance of folic acid prior to pregnancy, 400 mg, is easily achieved with a well-balanced diet. Since bread and cereals are now fortified with folic acid, it is becoming harder to avoid adequate folic acid. However, residents of closed communities, producing their own food may not have access to fortified food. Overconsumption of most vitamins is harmless, and the excess is excreted in urine. However, fat-soluble vitamins may accumulate and reach levels

that are toxic to the fetus. Iron is useful for women who are anemic because of iron deficiency.

Prenatal medical care is available to most women in the United States through a patchwork of public and private funding (Alan Guttmacher Institute 1994). Public funding has increased over the last decades as it became evident that the cost of caring for low-birthweight infants could be decreased by providing prenatal care (Merkatz 1990). The relative importance of individual facets of prenatal care is not certain (e.g., Crane et al. 1994; Higby et al. 1993), and some usual practices are probably useless but are persistent (AAP and ACOG 1997; Merkatz 1990). Efforts to assess prenatal care are complicated by the observation that women who actively seek prenatal care have better outcomes than those who do not, even when the latter group receives the same care.

### OUTCOMES OF PREGNANCY

A normal pregnancy lasts 38–42 menstrual weeks (*term*). Infants over 37 weeks are considered *full-term*. Infants under 37 weeks are considered *pre-term*. *Low birthweight* (or under 2,500 grams at birth), includes both preterm infants and full-term infants who weigh less than expected. The *low-birthweight rate* is the percentage of births under 2,500 grams.

The low-birthweight rate is an indicator of maternal health and of the effectiveness of prenatal and intrapartum care. However, the low birthweight numbers may be affected by the definition of live births. In some locations, live births are considered to be births over 28 weeks and/or 750 or 1,000 grams. In the United States the usual definition is any birth over 500 grams or showing movement or cardiac activity after birth; birth of a fetus under 500 grams without movement is classified as a spontaneous abortion. However, the definition varies by state and some states have adopted 350 grams as the threshold (Horton 1998). A 350-gram threshold will appear to raise the rate of low birthweight and the mortality rate of low-birthweight infants, as virtually no fetus born between 350 and 500 grams will survive (see ACOG 1995a).

Low birthweight is the leading cause of *perinatal death*. Perinatal deaths consist of *ante-partum deaths* (in which the fetus dies before delivery), *intra-partum death* (in which the fetus dies during labor or

delivery), and *neonatal deaths* (in which a liveborn infant dies in the first 28 days). In contrast to the low-birthweight rate, which is usually expressed as a percentage, the death rate is usually expressed as the rate per thousand births (both live and stillbirths). Death rates vary according to policies on inclusion of the smallest fetuses and infants.

A *viable fetus* is a fetus that can survive outside the uterus. Viability may affect a decision to use an intrauterine treatment versus delivery and treatment of the infant. Viability has also played a role in some debates about abortion, as the gestational age of viability has decreased with advances in perinatal care.

For most healthy women, the physical stress of pregnancy is easily managed. Pregnancy is much more dangerous to women with underlying anemia, heart, or kidney disease, who cannot manage the necessary increase in blood supply and circulation, and to women who do not have adequate nutrition. Family planning, by allowing adequate nutrition before pregnancy and recovery between pregnancies, is critical to women's health status during pregnancy (Alan Guttmacher Institute 1997). Women who are poorly nourished or ill have a much higher chance of delivering a low-birthweight baby, who in turn is more likely to have chronic illnesses and disabilities.

Delivery has additional risks to all women. In the nineteenth century, with the best available care, maternal mortality was about 1 percent per pregnancy, and it remains at that level, or higher, in some parts of the world. It is estimated that each year 600,000 women die from pregnancy and childbirth (Berg et al. 1996; WHO 1998a). The three most common causes of morbidity and mortality are hemorrhage, infection, and pre-eclampsia (high blood pressure and blood vessel disease unique to pregnancy). Morbidity and mortality from all 3 of these situations are less likely where there are trained birth attendants (midwives or physicians) with access to a short list of medications (Rooks 1997). Relatively simple interventions may make a major difference in outcome (e.g., Wallace et al. 1994). In contrast, the incremental increase of complex technology is relatively small, although it may be dramatic in some situations. The widespread application of some technologies has been harmful by creating additional problems, such as interference with normal

<b>Some Commonly Used Prenatal Tests</b>		
Test	Rationale	Limitations
Anemia screening	Allows supplementation.	
Gonorrhea	Treatment prevents infection of the fetus.	Screening cost-effective in high-prevalence populations.
Syphilis	Treatment prevents infection of the fetus.	
Human immunodeficiency virus (HIV)	Allows treatment with anti-retroviral medications during delivery, planning for infant feeding	Effective if antiretroviral drugs are available and safe alternatives to breastfeeding are available.
Bacterial vaginosis	Treatment prevents preterm labor.	Data on effectiveness are conflicting.
Glucose	Identify diabetic women early enough for treatment.	Effective in high-prevalence groups; may not be cost-effective in low-prevalence groups.
Rubella	Identification of nonimmune women identifies pregnancies at risk.	Vaccination of most young women has reduced incidence.
Hepatitis B	Identification of carrier women allows vaccination of infants.	In high-prevalence areas universal vaccination without screening is cost-efficient and simpler.
Rh typing	Identify fetuses at risk for Rh disease, and prevent sensitization.	
Alpha-fetoprotein	Identify fetuses with spine and brain malformations.	Can identify 90% of affected fetuses. 2-5% of women will need additional testing
"Triple screen"	Screen for Down syndrome.	Can identify 60% of affected fetuses; 5% of women will need additional testing.
Ultrasound	Identify twins, fetal defects, nonviable pregnancy.	Will identify most twins. Screening for birth defects not supported by randomized controlled trial.
Amniocentesis	Identify chromosomal abnormalities in fetus.	Carries some risk, expensive. CVS more versatile and faster.
Chorionic villus sampling (CVS)	Identify chromosomal and metabolic abnormalities in fetus.	May be riskier than amniocentesis, less available.

Table 1

labor by restriction of movement (Butler et al. 1993) or by diversion of resources from other beneficial programs.

Births are classified as *spontaneous* (vaginal), *operative vaginal*, or *operative abdominal*. Spontaneous birth includes births that are assisted by a birth attendant's hands. Operative vaginal births include forceps and vacuum extractor (suction cup applied to the baby's head); rates range from as low as 1 percent to 30 percent or higher in some settings. Operative abdominal births are "cesarean sections," and rates vary widely. In western

Europe rates are generally 8–15 percent of total births. In the United States, rates had risen steadily for several decades before decreasing very slightly in the last several years to about 22 percent of all births (National Center for Health Statistics 1999). In some cities throughout the world there is a local demand for elective abdominal delivery as an alternative to labor, generally restricted to women who can afford to pay for the request. With such widely varying rates of operative intervention there is no consensus on optimal rates, although there are many opinions. For instance, there has been no decline in the rate of mental retardation in the

United States, although both fetal monitoring and operative delivery were purported to prevent at least some retardation (ACOG 1992; Rosen and Dickinson 1992).

After birth, women and babies do best when they are kept together and allowed to establish lactation (WHO 1998b). Separation of mother and baby as practiced in many hospitals is not only unpleasant, but inappropriate in terms of infection control and infant nutrition. The effect of labor and delivery routines on “bonding” or attachment has been debated; whether any simple intervention can influence parenting is questionable at this point. The social and educational support of trained lay women has measurable success in some studies on labor outcome and breastfeeding success (Rooks 1997). More intensive programs of pregnancy, delivery, and postpartum peer care have not been fully evaluated (O’Connor 1998).

### TERMINATION OF PREGNANCY

Abortion is an event, other than a birth, that terminates a pregnancy. Abortion may be *spontaneous*, if it begins without intervention from the woman and without medical intervention, or *induced*, if some agent or procedure is used to cause the abortion. Induced abortion may also be classified as *legal* or *illegal*. Other terminology is neither uniform nor clear. *Therapeutic* abortion may refer to all legal induced abortions performed under medical supervision, or to those performed for medical indication such as severe illness in the woman. Spontaneous abortion may be classified as *inevitable* if it has not yet occurred but will occur, as *incomplete* if the pregnancy has been partially passed, or as *complete* if all tissue has been expelled or removed.

A *viable pregnancy* may refer to a pregnancy in which the baby is apparently healthy and growing; the pregnancy would be expected to continue if there were no intervention. A *nonviable pregnancy* will result in spontaneous abortion. Nonviable pregnancies are also called *blighted ovum* (which is technically incorrect since the ovum or egg has already divided), *empty sac*, or *fetal demise*. Abnormal pregnancies are sometimes detected by ultrasound or by blood testing, before there are any symptoms of spontaneous abortion.

Most abortions are requested because the pregnancy was unwanted, but there are other reasons, some of which are listed in Table 2.

In the United States, 90 percent of abortions occur in the first trimester; half of all women request abortion before 8 menstrual weeks. Ten percent of women request abortion after 12 weeks. Between 1 and 2 percent of all abortions performed are for fetal malformations; these are almost all second-trimester procedures. Fewer than 1 percent of abortions are performed in the third trimester. Generally, as the length of gestation increases, the cost of abortion increases, and the number of providers decreases (Gold 1990).

There are several types of procedures in use (for a more complete discussion, see Paul 1999). In the first trimester, there are both medical and surgical techniques. Surgical techniques consist of variations of suction curettage. In this procedure the cervix is stretched open if necessary; in very early pregnancy (fewer than 6 menstrual weeks) no opening may be necessary. The inside of the uterus is suctioned using a plastic tube and a vacuum created by an electric or manual pump. “Sharp” curettage, the traditional dilatation and curettage (D&C) is more traumatic and has been largely replaced by suction curettage.

Several medications are used for early medical abortion, depending on availability. Methotrexate is a medication that blocks folic acid, a vitamin. Several days later the woman takes a second drug, misoprostol (a prostaglandin drug), which makes the uterus contract and expel the pregnancy. The process is similar in both timing and feeling to a spontaneous abortion. The exact time sequence is difficult to predict; about 75 percent of women abort within 1 week. Vaginal bleeding occurs for a mean of about 10 days, but may last longer. The medication may fail to produce an abortion about 1 percent of the time, and about 1 percent of women need a suction curettage because of heavy bleeding. Methotrexate can be used up to 7 or 8 menstrual weeks, and is the most common agent in use in the United States, because it is FDA approved for other uses, and therefore easily available.

Mifepristone (formerly called RU 486) has been used in millions of women in China, France, Sweden, and the United Kingdom for medical

---

**Reasons Offered for Abortion**


---

Health concerns (e.g., severe hyperemesis)

Change in circumstances since conception:

- Abandonment by partner or family
- Change in finances or social situation
- Illness of other family member

Abnormal pregnancy

Exposure to teratogen

*Table 2*

abortion since 1986, but has had to overcome numerous political roadblocks in the United States. Mifepristone, a different class of medication from methotrexate, blocks progesterone hormone binding sites. It requires use of a second drug, a prostaglandin drug such as misoprostol, several days later to expel the pregnancy. In contrast to methotrexate, 75 percent of women abort within 2 days, and 90–95 percent abort within 1 week. Mifepristone can be used up to 7 to 9 menstrual weeks depending on the selection of the second drug.

After 7 to 9 menstrual weeks, suction curettage is the main technique. At the end of the first trimester and in second trimester, opening the cervix sufficiently becomes more challenging. Osmotic dilators are used over several hours or several days; osmotic dilators are placed in the cervix, where they absorb water and swell. Preparation for abortion in the mid- and late second trimester may take several days. Surgical procedures, often called dilatation and evacuation (D&E), are variations of early pregnancy techniques using suction and extraction instruments. After 16 weeks, medical (induction) techniques can be used to induce labor. All the agents—prostaglandin, oxytocin, and saline—are unpredictable, may take several hours to several days, and may involve a curettage to remove placenta if it is not completely expelled. Hospitalization is the rule, and therefore induction techniques are more costly than surgical techniques in an out-patient setting.

Legal abortion has very few serious complications. Term pregnancy has a death rate at least ten times higher than first-trimester abortion (10–12 deaths per 100,000 versus 0.5–1.0 deaths per 100,000 women in the United States) (Berg et al.

1996; Paul et al. 1999), and a morbidity rate (serious medical outcomes such as major operations) hundreds of times higher. There is virtually no situation in which it is safer for a woman to continue a pregnancy than to abort a pregnancy, although there are sometimes situations in which abortion should be delayed briefly. However, abortion is safest if it is performed early (Gold 1990; Paul et al. 1999). Early abortion by suction curettage does not have an effect on future fertility. Abortion itself is not associated with adverse psychological sequelae; unwanted pregnancy may have adverse associations regardless of whether the pregnancy is aborted or continued. In second trimester the medical risks are higher than in first trimester (Paul et al. 1999).

It is difficult to ascertain the numbers and types of abortions performed in the United States. In most states abortion is a reportable procedure; however, many procedures are not reported, probably for reasons of confidentiality of both provider and patient (Fu et al. 1998). Some abortions may be misclassified as treatment of a spontaneous abortion, since the suction procedure is identical; or the reverse may be true. In hospitals, some induced abortions may be classified as spontaneous abortion. The Alan Guttmacher Institute estimates abortion procedures using both governmental and nongovernmental sources. The CDC also publishes “Abortion Surveillance” as a supplement to *Mortality and Morbidity Weekly Review*. Small variations in abortion rates may be related to reporting and surveillance issues as well as varying rates of pregnancy and abortion.

Currently, early medical abortion accounts for fewer than 1 percent of all abortions in the United States, as methotrexate is the only agent available. In other countries, medical abortion may account for up to a third of all abortion procedures. In France, half of all women seeking abortion request abortion early enough to be eligible for medical abortion, and two-thirds of them choose medical abortion. This shift to medical abortion in France has been accompanied by a shift to earlier abortion in general, while overall abortion rates have remained the same (Paul et al. 1999). Medical abortion is highly acceptable to women in many countries. Acceptability studies consistently estimate that 60–80 percent of women would choose medical abortion over surgical abortion were it available (Winikoff 1994).

However, despite the preference for medical abortion over surgical abortion, surgical abortion remains the only method available to most women, if they have access to any method at all. In the United States the number of abortion providers has dropped over the last two decades. Abortion is available in a minority of counties in the United States, most of them in urban areas (Lichter et al. 1998). The majority of abortion procedures take place in free-standing medical facilities, many of them primarily devoted to abortion provision (Henshaw 1998b). Planned Parenthood clinics provide about 12 percent of all abortions. About 10 percent of abortions occur in hospitals or hospital-affiliated clinics. At least 5 percent of abortions occur in doctors' offices that are not identified as providing primarily abortion-related care.

Most physicians providing abortions in the United States are obstetrician-gynecologists, but other physicians provide abortions. Only about a third of gynecologists provide abortions, so most women needing abortion are referred to an unfamiliar caregiver. First-trimester abortions are also provided in several states by physicians' assistants (PAs) and midwives. Most states have "physician-only" statutes that limit the provision of abortion to physicians. However, these statutes were written before the increase in utilization of advanced practice nurses, such as midwives, and "physician extenders," such as PAs, who currently perform other procedures of comparable complexity and skill (Freedman et al. 1986). In a few countries (e.g., Bangladesh) midwives provide most abortion services.

Costs for abortions range from several hundred dollars for a first-trimester procedure in an office or clinic setting, to thousands of dollars for procedures performed in hospitals. Many insurance policies cover abortion, although reimbursement rates vary. Some forms of insurance do not cover abortion at all, for example, insurance provided to U.S. federal employees. Publicly funded insurance for indigent women covers "medically indicated" abortions in some states but not others. Finally, many women requesting abortion do not have any insurance at all, most commonly because their employment does not provide insurance. Women who need specialized procedures because of underlying medical illnesses may have additional barriers; there are several states in which no hospital will allow an abortion to be performed.

In the United States the majority of people polled supported the availability of abortion, but many were in favor of some restrictions. Parental consent laws for minors have been passed in nearly half the states. Any such law must contain a "judicial bypass" to be considered constitutional, so that minors can petition a judge if they cannot tell a parent (Paul et al. 1999). In practice, most minors do involve a parent with or without parental consent laws. The main effect of such laws is to delay abortion for those minors without good family support.

In the United States there is a vocal and well-funded minority with the goal of criminalizing abortion. Mandatory waiting periods and mandatory consent processes have also been passed. These increase the amount of time and expense necessary for a woman to obtain an abortion, particularly where there are few providers. Other quasi-legal attempts to decrease abortion include passage of laws that are either unenforceable or unconstitutional. These include restrictions on the type of facility that can perform abortion and "partial-birth abortion" laws. The resulting legal challenges are expensive, and this tactic can be considered a sustained economic assault (Reproductive Freedom News 1999). Antiabortion groups in the United States have been increasingly involved in terrorist activities, including harassment, arson, violence, and even murder. Several thousand incidents of harassment and arson are reported annually to the National Abortion Federation (National Abortion Federation 1999).

Abortion is generally legal in most European countries and in much of Asia, but is illegal in most of South America and Africa (WHO 1998b). The effect of government-sanctioned denial of abortion rights is not to decrease the number of abortions significantly. Abortion rates are correlated with multiple factors such as patterns of sexual activity, type of contraceptive use, desired family size, and tolerance of unwanted pregnancy. In the United States it has been estimated that the number of abortions performed annually before widespread legalization was 600,000 to 1.2 million per year (Gold 1990). The number reported the first year after legalization was 615,831 (Koonin et al. 1996). The effect of criminalizing abortion is to delay abortion for some women and to make it riskier for almost all women, resulting in increased

death rates (Population Reports 1997). As an example, after abortion was criminalized in Romania, the death rate from abortions quintupled; while maternal mortality doubled (WHO 1998b). In South American and East Africa, the rates of illegal abortion are similar, about 35 per 1,000 women of childbearing age, but death rates are dissimilar (0.3 percent and 1.5 percent, respectively, of women undergoing abortion), related to underlying health status and access to postabortion care (WHO 1998b).

Increases in the provision of contraceptive services are associated with decreases in abortion rates (Henshaw et al. 1999; Estrin 1999). However, the United States has consistently refused to fund contraceptive services adequately, particularly services in developing countries, because of right-wing political pressure (Alan Guttmacher Institute 1996).

Pregnancy carries an intrinsic risk to women's health, which can be minimized by appropriate medical care to women who are in good health to start. Family planning is essential to women's health during pregnancy, and essential to the health of their infants. Women who do not want to be pregnant will risk health and life to end the pregnancy. Denying or criminalizing abortion care results in additional health risks for women while diverting health resources that could be used for maternity services for other women.

## REFERENCES

- Affonso, Dyanne D., Chong-Yeu Lui-Chang, and Linda J. Mayberry 1997 "Worry: Conceptual Dimensions and Relevance to Childbearing Women." *Health Care for Women International* 20:227-236.
- Alan Guttmacher Institute 1994 *Uneven and Unequal: Insurance Coverage and Reproductive Health Services*. New York: Alan Guttmacher Institute.
- 1996 *Endangered: U.S. Aid for Family Planning Overseas*. New York: Alan Guttmacher Institute.
- 1997 *Family Planning Improves Child Survival and Health*. New York: Alan Guttmacher Institute.
- 1999 *Teenage Pregnancy: Overall Trends and State-by-State Information*. New York: Alan Guttmacher Institute (April).
- American Academy of Pediatrics and American College of Obstetricians and Gynecologists 1997 *Guidelines for Perinatal Care*, 4th ed. Elk Grove Village, Ill.: AAP.
- American College of Obstetricians and Gynecologists 1992 *Fetal and Neonatal Neurologic Injury*. Technical Bulletin. Washington, D.C.: ACOG.
- 1995a *Committee Opinion: Perinatal and Infant Mortality Statistics*. Washington, D.C.: ACOG.
- 1995b *Domestic Violence*. Washington, D.C.: ACOG.
- 1998 *Committee Opinion: Vitamin A Supplementation during Pregnancy*. Washington, D.C.: ACOG.
- Berg, Cynthia J., Hani K. Atrash, Lisa M. Koonin, and Myra Tucker 1996 "Pregnancy-Related Mortality in the United States, 1987-1990." *Obstetrics and Gynecology* 88:161-167.
- Crane, James P., Michael L. LeFevre, Renee C. Winborn, Joni K. Evans, Bernard G. Ewigman, Raymond P. Bain, Frederic P. Frigoletto, Donald McNellis, and the RADIUS Study Group 1994 "A Randomized Trial of Prenatal Ultrasonographic Screening: Impact on the Detection, Management, and Outcome of Anomalous Fetuses." *American Journal of Obstetrics and Gynecology* 171:382-399.
- Butler, Jane, Barbara Abrams, Jennifer Parker, James M. Roberts, and Russell K. Laros 1993 "Supportive Nurse-Midwife Care is Associated with a Reduced Incidence of Cesarean Section." *American Journal of Obstetrics and Gynecology* 168:1407-1413.
- Cunningham, E. Gary, Paul C. MacDonald, Norman F. Gant, Kenneth J. Leveno, Larry C. Gilstrap, Gary D. V. Hankins, and Steven L. Clark 1997 *Williams Obstetrics*, 20th ed. Stamford, Conn.: Appleton and Lange.
- Czeizel, Andrew E., and Istvan Dudas 1992 "Prevention of the First Occurrence of Neural-Tube Defects by Periconceptional Vitamin Supplementation." *New England Journal of Medicine* 327:1832-1835.
- Estrin, D. J. 1999 "In Three Former Soviet States, Rates of Abortion Are Declining as Contraceptive Prevalence Increases." *International Family Planning Perspectives* 25:49-50.
- Fraser, Alison M., John E. Brockert, and R. H. Ward 1995 "Association of Young Maternal Age with Adverse Reproductive Outcomes." *New England Journal of Medicine* 332:1113-1117.
- Freedman, Mary Anne, David A. Jillson, Roberta R. Coffin, and Lloyd F. Novick 1986 "Comparison of Complication Rates in First Trimester Abortions Performed by Physician Assistants and Physicians." *American Journal of Public Health* 76:550-554.
- Fu, Haishan, Jacqueline E. Darroch, Stanley K. Henshaw, and Elizabeth Kolb 1998 "Measuring the Extent of Abortion Underreporting in the 1995 National Survey of Family Growth." *Family Planning Perspectives* 30:128-133, 138.



- Gabbe, S., J. R. Niebyl, and J. L. Simpson 1996 *Obstetrics: Normal and Problem Pregnancies*, 3rd ed. New York: Churchill-Livingston.
- Gold, Rachel Benson 1990 *Abortion and Women's Health*. New York: Alan Guttmacher Institute.
- Henshaw, Stanley K. 1998a "Unintended Pregnancy in the United States." *Family Planning Perspectives* 30:24–29, 46.
- 1998b "Abortion Incidence and Services in the United States, 1995–1996." *Family Planning Perspectives* 30:263–270, 287.
- , Susheela Singh, and Taylor Haas 1999 "Recent Trends in Abortion Rates Worldwide." *International Family Planning Perspectives* 25:44–48.
- Higby, Kenneth, Elly M. J. Xenakis, and Carl J. Pauerstein 1993 "Do Tocolytic Agents Stop Preterm Labor? A Critical and Comprehensive Review of Efficacy and Safety." *American Journal of Obstetrics and Gynecology* 168:1247–1259.
- Horton, Jacqueline A. (ed.) 1995 *The Women's Health Data Book, A Profile of Women's Health in the United States*. Washington, D.C.: Jacobs Institute of Women's Health.
- 1998 *State Profiles of Women's Health*. Washington, D.C.: Jacobs Institute of Women's Health.
- Koonin, Lisa M., Jack C. Smith, Merrell Ramick, and Clarice A. Green 1996 "Abortion Surveillance—United States, 1992." *Morbidity and Mortality Weekly Review*, vol. 45, supp. 3. Atlanta, Ga.: Center for Disease Control.
- Lichter, Daniel T., Diane K. McLaughlin, and David C. Ribar 1998 "State Abortion Policy, Geographic Access to Abortion Providers and Changing Family Formation." *Family Planning Perspectives* 30:281–287.
- National Abortion Federation 1999 "Clinic Support Update." Washington, D.C.: NAF (April).
- National Center for Health Statistics 1999 *Births: Final Data for 1997*. Atlanta, Ga.: Center for Disease Control and Prevention.
- O'Connor, M. L. 1998 "Home Nurse Visits from Pregnancy until Child's Second Birthday Have Sustained Benefits for Mother and Child." *Family Planning Perspectives* 30:47–48.
- Paul, Maureen, E. Steve Lichtenberg, Lynn Borgatta, David A. Grimes, and Phillip G. Stubblefield 1999 *A Clinician's Guide to Medical and Surgical Abortion*. New York: Churchill-Livingston.
- Population Reports 1997 *Caring for Postabortion Complications: Saving Women's Lives*. Baltimore, Md.: Population Information Center (September).
- Reproductive Freedom News 1999 *Anti-Abortion Legislature Swamps Choice Activists*. New York: Center for Law and Reproductive Policy (May).
- Rooks, Judith S. 1997 *Midwifery and Childbirth in America*. Philadelphia, Pa.: Temple University Press.
- Rosen, Mortimer G., and Janet C. Dickinson 1992 "The Incidence of Cerebral Palsy." *American Journal of Obstetrics and Gynecology* 167:417–423.
- Rosenberg, Irwin H. 1992 "Folic Acid and Neural-Tube Defects—Time for Action?" *New England Journal of Medicine* 327:1875–1877.
- Swiss, S. 1993 "Rape as a Crime of War: A Medical Perspective." *Journal of the American Medical Association* 270:612–615.
- Taylor, Paul (ed.) 1980 "Parent-Infant Relationships." New York: Grune and Stratton.
- Wallace, H. M., R. P. Nelson, and P. J. Sweeney 1994 *Maternal and Child Health Practices*, 4th ed. Oakland, Calif.: Third Party.
- Winikoff, Beverly 1994 *Acceptability of First Trimester Medical Abortion*. New York: Population Council.
- World Health Organization, Division of Reproductive Health 1998a *Postpartum Care of the Mother and Newborn: A Practical Guide*. Geneva, Switzerland: WHO.
- 1998b *Unsafe Abortion*, 3rd ed. Geneva, Switzerland: WHO.

LYNN BORGOTTA

## PREJUDICE

Gordon Allport, in his classic *The Nature of Prejudice*, defined prejudice as "an antipathy based upon a faulty and inflexible generalization" (1954, p. 9). This phrasing neatly captures the notion that both inaccurate beliefs and negative feelings are implicated in prejudice. To these "cognitive" and "affective" dimensions of prejudice, some analysts add "conative," referring to action orientation (Klineberg 1972) and prescription (Harding et al. 1969). Allport's circumspection on the conative implications of prejudice—he said "(prejudice) may be felt or expressed" (1954, p. 9)—foreshadowed our growing understanding that the correspondence of behavior with cognitions and feelings is uncertain, a research issue in its own right (Schuman and Johnson 1976).

Racial and ethnic prejudice was Allport's primary interest. Emerging social issues have brought

expanded attention to other forms of prejudice—against women, the elderly, handicapped persons, AIDS patients, and others. This discussion will focus on racial prejudice among white Americans, in the expectation that parallels and points of contrast will continue to make race relations research relevant to other forms of prejudice.

### TRENDS AND PATTERNS

For many years, derogatory stereotypes, blatant aversion to interracial contact, and opposition in principle to racial equality were seen as the central manifestations of race prejudice, virtually defining the social science view of the problem. Indicators of these beliefs and feelings show a clear positive trend (Jaynes and Williams 1989; Schuman et al. 1997). White Americans' belief in the innate intellectual inferiority of blacks declined from 53 percent in 1942 to about 20 percent in the 1960s, when the question was discontinued in major national surveys. The percentage of whites who said it would make no difference to them if a Negro of equal social status moved into their block rose from 36 percent to 85 percent between 1942 and 1972. White opinion that blacks should have "as good a chance as white people to get any kind of job" climbed from 45 percent in 1944 to 97 percent in 1972. Thomas Smith and Paul Sheatsley sum up this picture without equivocation: "Looking over this forty-year span, we are struck by the steady, massive growth in racial tolerance" (1984, p. 14).

Recurrent outbursts of overt racial hostility and public acts of discrimination (Feagin 1991) serve as unfortunate reminders that some white Americans still cling to blatant prejudice. More importantly, even the majority of whites, those on whom Smith and Sheatsley focus, appear unambiguously tolerant only if attention is confined to such traditional survey indicators as those described above. A confluence of developments has broadened the study of race prejudice and transformed our understanding of white racial attitudes. First, evidence of widespread, subtle prejudice has been revealed in research using disguised, "nonreactive" methods. Second, "social cognition" scholarship, paramount for two decades in the psychological wing of social psychology, has been powerfully

applied to intergroup relations. Recent scholarship is broadened and balanced by its acknowledgment of the crucial role of affect along with cognition. Third, evolution of the struggle for racial equality in the United States has shifted attention to a new domain of racial policy-related beliefs and feelings. These perspectives provide ample evidence that white racial prejudice is not a thing of the past, but exists today in complex forms that have yet to be thoroughly charted.

**Evidence from "Nonreactive" Studies.** Given the clear dominance of "liberal" racial norms evinced in public opinion data, it might be expected that needs for social acceptability and self-esteem would lead many whites to withhold evidence of negative racial feelings and cognitions whenever possible. Disguised, "nonreactive" research (Webb 1981) provides substantial evidence that, indeed, traditional survey approaches underestimate negative racial feeling. Field experiments reveal that whites often provide less help to victims who are black (Crosby et al. 1980), sometimes redefining the situation so as to justify their lack of response (Gaertner 1976). Such elements of nonverbal behavior as voice tone (Weitz 1972) and seating proximity (Word et al. 1974) have been found to reveal negative racial feelings and avoidance. Recent reaction time and word completion studies similarly document the existence of "implicit" racial prejudice among many whites who score low on self-reported "explicit" prejudice (Dovidio et al. 1997). Thus, accumulating American evidence reveals that "microaggressions" (Pettigrew 1989) often accompany self-portrayals of liberalism. Parallel research in western Europe has uncovered similar forms of microaggression against that continent's new immigrant minorities (Den Uyl et al. 1986; Klink and Wagner 1999; Sissons 1981).

**Social Cognition Perspectives, Now Acknowledging the Role of Affect.** In recognizing aspects of prejudice as predictable outgrowths of "natural" cognitive processes, Allport (1954) was ahead of his time. A wave of social cognition research on intergroup relations was set in motion by Henri Tajfel (1969), who demonstrated that mere categorization—of physical objects or of people—encourages exaggerated perception of intragroup homogeneity and intergroup difference. Even in

“minimal groups” arbitrarily created in psychology laboratories, these effects of social categorization are often accompanied by ingroup favoritism and outgroup discrimination (Brewer 1979, 1991; Hamilton 1979). Accumulating evidence of the negative consequences of ingroup/outgroup categorization has spurred research aimed at identifying conditions of intergroup contact that are likely to decrease category salience and promote “individuation” or “decategorization” (Brewer and Miller 1988; Wilder 1978), or at least to reduce the negativity of outgroup stereotypes (Rothbart and John 1985; Wilder 1984). Recent attention to the role of motivation in guiding cognition (Fiske 1998) contributes to this effort.

The study of attributional processes (Heider 1958) also has been usefully applied to intergroup relations, calling attention to such issues as whether white perceivers believe that black economic hardship results from discrimination or lack of effort. Research evidence has linked stereotypic thinking to attributions of outgroup behavior (David L. Hamilton 1979). Specific predictions are developed in Pettigrew’s discussion of the “ultimate attributional error” (1979a), the tendency to hold outgroups personally responsible for their failures, but to “discount” their responsibility for successes, attributing successes to such factors as luck or unfair advantage.

The intense research scrutiny given cognitive factors meant that the critical affective component of prejudice was often ignored. That imbalance is being corrected. The 1993 publication of *Affect, Cognition, and Stereotyping*, edited by Mackie and Hamilton (1993), marked the dramatic shift in emphasis in social psychology. Varied research, from American laboratory experimentation (Stangor et al. 1991; Dovidio et al. 1989) to European surveys (Pettigrew 1997; Pettigrew and Meertens 1995), demonstrates that emotional factors not only are central to intergroup prejudice but have special characteristics of their own and are highly predictive of policy attitudes. One attempt at synthesis outlines a tripartite conception of prejudice as stereotypes, affect, and “symbolic beliefs” (Esses et al. 1993; Zanna 1994).

**Expanding the Racial Attitude Domain to Policy Views.** Over the past twenty-five years, evolution in the struggle for racial equality has brought new complexity to the public debate about racial

issues. Notions that barriers to black equality consist solely of white hostility and aversion, and formal denial of rights, now appear naive. Advocates insist that structural barriers far more complex and far more pervasive than formal denial of access prevent actual desegregation and equality of opportunity, making questions about acceptance by white individuals a moot point for millions of black Americans.

In the current era of U.S. race relations, traditional manifestations of race prejudice recede in relevance, and different forms of race-related belief and feeling take center stage—reactions to agitation for change, recognition and interpretation of continuing inequality, and support for proposed remedies. By all indications, such white “perceptions, explanations, and prescriptions” (Apostle et al. 1983, p. 18) show far less consensus and support for racial change than appeared in traditional race survey data. Asked about specific policies and programs designed to increase racial equality—fair housing guarantees, school desegregation plans, affirmative action in hiring and college admission—white Americans show substantially less support than they voice for racial equality in principle (Pettigrew 1979b; Schuman et al. 1997). Many white Americans exaggerate recent black gains and benefits of affirmative action (Steeh and Krysan 1996) and underestimate the remaining inequality (Kluegel and Smith 1982). There is substantial white resentment of black activism and perceived progress (Bobo 1988a; Schuman et al. 1997).

Attribution research in social psychology and earlier societal analyses (Ryan 1971; Feagin 1975) converge with recent studies of racial policy opinion to tell a clear story: Whites explain the economic plight of black Americans more often as the result of such “individualistic” factors as lack of motivation than in terms of such “structural” factors as discrimination (Apostle et al. 1983; Kluegel and Smith 1986). In addition, individualistic attributions along with denial of discrimination are linked to a variety of policy-relevant beliefs and opinions, including opposition to affirmative action (Bobo and Kluegel 1993; Kluegel and Smith 1983, 1986).

## CHARACTERIZATIONS OF WHITE RACIAL ATTITUDES

Efforts to characterize the complex pattern of racial attitudes held by white Americans emphasize an array of themes, as discussed below.

*“Natural” Cognitive Processing.* As noted earlier, social cognition analyses claim that a substantial part of the racial prejudice once thought to have sociocultural or psychodynamic roots actually stems from ordinary cognitive processing, particularly categorization (David L. Hamilton 1979). Social cognition portrayals increasingly acknowledge motivational and social influences (David A. Hamilton and Tina K. Trolier 1986; Fiske 1987, 1998). And there are recent powerful calls to acknowledge the joint influence on prejudice of cognition and affect (Esses et al. 1993; Pettigrew 1997; Smith 1993).

*Strain Between Individualism and Egalitarianism.* Current racial policy issues are said to pull whites between two cherished American values, individualism and egalitarianism (Lipset and Schneider 1978). *Qualified* support for social programs exists, in this view, because egalitarian sentiments prevail only until a proposal challenges individualistic values.

*Ambivalence.* Adding psychodynamic flavor to the individualism/egalitarianism value strain idea, some analysts describe current white feelings as an ambivalence that produces an unpredictable mix of amplified positive and negative responses (Katz et al. 1986).

*“Aversive” Racism.* A desire to avoid interracial contact, muted negative feeling, and egalitarian self-concept are the mix Kovel (1970) characterized as aversive racism. The outcome is avoidance of positive interracial behavior when the situation can be defined to permit it, and expression of negative feelings when there are ostensible nonracial justifications (Gaertner 1976; Gaertner and Dovidio 1986).

*“Symbolic” or “Modern” Racism and “Racial Resentment.”* Antiblack affect instilled by childhood socialization and the sense that racial change threatens fondly held individualistic values, *not* self-interest, are claimed as the twin foundations of Sears’s “symbolic” racism (1988). “Modern”

racism contains the added ingredient of denying continuing racial inequality (McConahay 1986). “Racial resentment” (Kinder and Sanders 1996) is a recent addition to this family of “new racism” constructs. The label serves as a reminder that whatever the roots of new racism, a primary manifestation is anger that blacks’ gains have exceeded their entitlement.

*“Subtle” Prejudice.* Using probability survey data from western Europe, Pettigrew and Meertens (1995) developed scales for both blatant and subtle prejudice that proved highly predictive of attitudes toward immigrants across four nations, six target outgroups, and seven samples. Their conception of subtle prejudice contains three components: perceived threat to traditional values (similar to symbolic racism), exaggeration of intergroup differences, and the absence of positive affect (admiration and sympathy) for the outgroup. Rejecting the claims of Sniderman and his colleagues (1991), Meertens and Pettigrew (1997) demonstrate that their subtle prejudice is distinctly different from political conservatism.

*“Dominant Stratification Ideology.”* A belief that opportunity is plentiful and equally distributed, and thus effort is economically rewarded and economic failure is deserved—these compose the “dominant stratification ideology” (Huber and Form 1973; Kluegel and Smith 1986), a sociological elaboration on the individualism theme. Although personal status and strands of American “social liberalism” also play a role, unyielding adherence to this American “dominant ideology” is portrayed as a major impediment to public support for redistributive claims in general, and to calls for racial change in particular (Kluegel and Smith 1986). On a backdrop of ignorance bred of social segregation, whites’ own experiences of economic success work to prevent recognition of the continuing barriers to full opportunity for black Americans (Kluegel 1985).

*Self-Interest.* Collective self-interest is sometimes identified as the primary basis of whites’ interracial beliefs and feelings (Jackman 1994; Wellman 1993). If zero-sum assumptions prevail, redistribution in favor of blacks will be seen as a losing proposition to whites. Self-interest is at the heart of what Bobo (1988b) called an “ideology of bounded racial change” and what Bobo and colleagues (1997) have dubbed “laissez-faire racism”: White

acceptance of racial change and efforts to promote it end when continued change is perceived to threaten the well-being of whites.

### PRESCRIPTIONS FOR MODERN PREJUDICE

When the lessons from cognitive social psychology are counterposed with those from other perspectives on modern race prejudice, an apparent dilemma is revealed. Though social cognition findings indicate that category salience can promote stereotype change under some circumstances (Cook 1984; Pettigrew 1998), much of the cognitive literature insists that categorization is a central contributor to race prejudice and negative race relations: Color consciousness is often portrayed as an evil, color blindness the ideal. From other scholars of modern prejudice, the analysis and prescription are nearly a mirror image of this view. Color blindness is said to impede forthright problem solving in desegregated institutions (Schofield 1986); to represent ignorance of the structural barriers faced by black Americans (Kluegel 1985); and to be used as a weapon by those opposing black claims of collective rights (Jackman and Muha 1984; Omi and Winant 1986). The solution implied or stated by these analysts is for whites to adopt a color consciousness that fully acknowledges the historical impact of racial subordination and the continuing liabilities of direct and indirect discrimination. The two streams of advice present this challenge: How to promote a racial understanding in the white public that minimizes the psychological liabilities of ingroup/outgroup categorization while acknowledging the full sociological implications of the past and continuing color line.

(SEE ALSO: *Discrimination*)

#### REFERENCES

- Allport, Gordon 1954 *The Nature of Prejudice*. Cambridge, Mass.: Addison-Wesley.
- Apostle, Richard A., Charles Y. Glock, Thomas Piazza, and Marijean Suelzle 1983 *The Anatomy of Racial Attitudes*. Berkeley: University of California Press.
- Bobo, Lawrence 1988a "Attitudes toward the Black Political Movement: Trends, Meaning, and Effects on Racial Policy Preferences." *Social Psychology Quarterly* 51:287-302.
- 1988b "Group Conflict, Prejudice, and the Paradox of Contemporary Racial Attitudes." In Phyllis A. Katz and Dalmas A. Taylor, eds., *Eliminating Racism*. New York: Plenum.
- , and James R. Kluegel 1993 "Opposition to Race-Targeting: Self-Interest, Stratification Ideology, or Racial Attitude?" *American Sociological Review* 58:443-464.
- , and Ryan A. Smith 1997 "Laissez Faire Racism: The Crystallization of a 'Kinder Gentler' Anti-Black Ideology." In Steven A. Tuch and Jack K. Martin, eds., *Racial Attitudes in the 1990s: Continuity and Change*. Greenwood, Conn.: Praeger.
- Brewer, Marilyn B. 1979 "Ingroup Bias in the Minimal Intergroup Situation: A Cognitive-Motivational Analysis." *Psychological Bulletin* 86:307-324.
- 1991 "The Social Self: On Being the Same and Different at the Same Time." *Personality and Social Psychology Bulletin* 17:475-482.
- , and Norman Miller 1988 "Contact and Cooperation: When Do They Work?" In Phyllis A. Katz and Dalmas A. Taylor, eds., *Eliminating Racism*. New York: Plenum.
- Cook, Stuart W. 1984 "Cooperative Interaction in Multiethnic Contexts." In Norman Miller and Marilyn B. Brewer, eds., *Groups in Contact: The Psychology of Desegregation*. New York: Academic.
- Crosby, Faye J., Stephanie Bromley, and Leonard Saxe 1980 "Recent Unobtrusive Studies of Black and White Discrimination and Prejudice: A Literature Review." *Psychological Bulletin* 87:546-563.
- Den Uyl, Roger, C. E. Choenni, and Frank Bovenkerk 1986 "Mag Het Ook Een Buitenlander Wezen? Discriminatie Bij Uitzendburo's." Utrecht, The Netherlands: National Bureau against Racism.
- Dovidio, John F., Kerry Kawakami, Craig Johnson, Brenda Johnson, and Adiaiah Howard 1997 "On the Nature of Prejudice: Automatic and Controlled Processes." *Journal of Experimental Social Psychology* 33:510-540.
- Dovidio, John F., Jeffrey Mann, and Samuel L. Gaertner 1989 "Resistance to Affirmative Action: The Implications of Aversive Racism." In Faye J. Crosby and Fletcher A. Blanchard, eds., *Affirmative Action in Perspective*. New York: Springer-Verlag.
- Esses, Victoria M., Geoffrey Haddock, and Mark P. Zanna 1993 In Diane M. Mackie and David L. Hamilton, eds., *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*. San Diego, Calif.: Academic.
- Feagin, Joe R. 1975 *Subordinating the Poor*. Englewood Cliffs, N.J.: Prentice-Hall.

- 1991 “The Continuing Significance of Race: Antiblack Discrimination in Public Places.” *American Sociological Review* 56:101–116.
- Fiske, Susan T. 1987 “On the Road: Comment on the Cognitive Stereotyping Literature in Pettigrew and Martin.” *Journal of Social Issues* 43:113–118.
- 1998 “Stereotyping, Prejudice, and Discrimination.” In Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *The Handbook of Social Psychology*, 4th ed. New York: McGraw-Hill.
- Gaertner, Samuel L. 1976 “Nonreactive Measures in Racial Attitude Research: A Focus on ‘Liberals.’” In Phyllis A. Katz, ed., *Toward the Elimination of Racism*. New York: Pergamon.
- , and John F. Dovidio 1986 “The Aversive Form of Racism.” In John F. Dovidio and Samuel L. Gaertner, eds., *Prejudice, Discrimination, and Racism*. Orlando, Fla.: Academic.
- Hamilton, David A., and Tina K. Trolie 1986 “Stereotypes and Stereotyping: An Overview of the Cognitive Approach.” In Phyllis A. Katz and Dalmas A. Taylor, eds., *Eliminating Racism*. New York: Plenum.
- Hamilton, David L. 1997 “A Cognitive-Attributional Analysis of Stereotyping.” In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 12. New York: Academic.
- Harding, John, Harold Proshansky, Bernard Kutner, and Isador Chein 1969 “Prejudice and Intergroup Relations.” In Gardner Lindzey and Elliot Aronson, eds., *Handbook of Social Psychology, Volume V*, 2nd ed. Reading, Mass.: Addison-Wesley.
- Heider, Fritz 1958 *The Psychology of Interpersonal Relations*. New York: John Wiley.
- Huber, Joan, and William H. Form 1973 *Income and Ideology*. New York: Free Press.
- Jackman, Mary R. 1994 *The Velvet Glove: Paternalism and Conflict in Gender, Class, and Race Relations*. Los Angeles, Calif.: University of California Press.
- , and Michael J. Muha 1984 “Education and Intergroup Attitudes: Moral Enlightenment, Superficial Democratic Commitment, or Ideological Refinement?” *American Sociological Review* 49:751–769.
- Jaynes, David Gerald, and Robin M. Williams, Jr. 1989 *A Common Destiny: Blacks and American Society*. Washington, D.C.: National Academy Press.
- Katz, Irwin, Joyce Wackenhut, and R. Glen Hass 1986 “Racial Ambivalence, Value Duality, and Behavior.” In John F. Dovidio and Samuel L. Gaertner, eds., *Prejudice, Discrimination, and Racism*. Orlando, Fla.: Academic.
- Kinder, Donald R., and Lynn M. Sanders 1996 *Divided by Color: Racial Politics and Democratic Ideals*. Chicago: University of Chicago Press.
- Klineberg, Otto 1972 “Prejudice. The Concept.” In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, reprint ed. New York: Macmillan Free Press.
- Klink, Andreas, and Ulrich Wagner 1999 “Discrimination against Ethnic Minorities in Germany: Going Back to the Field.” *Journal of Applied Social Psychology* 29:402–423.
- Kluegel, James R. 1985 “If There Isn’t a Problem, You Don’t Need a Solution.” *American Behavioral Scientist* 28:761–784.
- , and Eliot R. Smith 1982 “Whites’ Beliefs about Blacks’ Opportunity.” *American Sociological Review* 47:518–532.
- 1983 “Affirmative Action Attitudes: Effects of Self-Interest, Racial Affect, and Stratification Beliefs on Whites’ Views.” *Social Forces* 61:797–824.
- 1986 *Beliefs about Inequality: Americans’ Views of What Is and What Ought to Be*. New York: Aldine de Gruyter.
- Kovel, Joel 1970 *White Racism: A Psychohistory*. New York: Pantheon.
- Lipset, Seymour Martin, and William Schneider 1978 “The Bakke Case: How Would It Be Decided at the Bar of Public Opinion?” *Public Opinion* 1:38–44.
- McConahay, John B. 1986 “Modern Racism, Ambivalence, and the Modern Racism Scale.” In John F. Dovidio and Samuel L. Gaertner, eds., *Prejudice, Discrimination, and Racism*. Orlando, Fla.: Academic Press.
- Mackie, Diane M., and David L. Hamilton (eds.) 1993 *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*. San Diego, Calif.: Academic Press.
- Meertens, Roel W., and Thomas F. Pettigrew 1997 “Is Subtle Prejudice Really Prejudice?” *Public Opinion Quarterly* 6:54–71.
- Omi, Michael, and Howard Winant 1986 *Racial Formation in the United States from the 1960s to the 1980s*. New York: Routledge and Kegan Paul.
- Pettigrew, Thomas F. 1979a “The Ultimate Attribution Error: Extending Allport’s Cognitive Analysis of Prejudice.” *Personality and Social Psychology Bulletin* 5:461–476.
- 1979b “Racial Change and Social Policy.” *Annals of the American Academy of Political and Social Science* 441:114–131.
- 1989 “The Nature of Modern Racism.” *Revue Internationale de Psychologie Sociale* 2:291–305.

- 1997 “The Affective Component of Prejudice: Results from Western Europe.” In Steven A. Tuch and Jack K. Martin, eds., *Racial Attitudes in the 1990s: Continuity and Change*. Westport, Conn.: Praeger.
- 1998 “Intergroup Contact Theory.” *Annual Review of Psychology* 49:65–85.
- , and Roel W. Meertens 1995 “Subtle and Blatant Prejudice in Western Europe.” *European Journal of Social Psychology* 25:57–75.
- Rothbart, Myron, and Oliver P. John 1985 “Social Categorization and Behavioral Episodes: A Cognitive Analysis of the Effects of Intergroup Contact.” *Journal of Social Issues* 41:81–104.
- Ryan, William 1971 *Blaming the Victim*. New York: Vintage.
- Schofield, Janet Ward 1986 “Causes and Consequences of the Colorblind Perspective.” In John F. Dovidio and Samuel L. Gaertner, eds., *Prejudice, Discrimination, and Racism*. Orlando, Fla.: Academic.
- Schuman, Howard, and Michael P. Johnson 1976 “Attitudes and Behavior.” *Annual Review of Sociology* 2:161–207.
- Schuman, Howard, Charlotte Steeh, Lawrence Bobo, and Maria Krysan 1997 *Racial Attitudes in America: Trends and Interpretations*, 2nd ed. Cambridge, Mass.: Harvard University Press.
- Sears, David O. 1988 “Symbolic Racism.” In Phyllis A. Katz and Dalmas A. Taylor, eds., *Eliminating Racism*. New York: Plenum Press.
- Sissons, Mary 1981 “Race, Sex and Helping Behavior.” *British Journal of Social Psychology* 20:285–292.
- Smith, Eliot R. 1993 “Social Identity and Social Emotions: Toward New Conceptualizations of Prejudice.” In Diane M. Mackie and David L. Hamilton, eds., *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*. San Diego, Calif: Academic.
- Smith, Thomas W., and Paul B. Sheatsley 1984 “American Attitudes toward Race Relations.” *Public Opinion* 6:14–15, 50–53.
- Sniderman, Paul M., Thomas Piazza, Philip E. Tetlock, and Ann Kendrick 1991 “The New Racism.” *American Journal of Political Science* 35:423–447.
- Stangor, Charles, Sullivan, Linda A., and Thomas E. Ford 1991 “Affective and Cognitive Determinants of Prejudice.” *Social Cognition* 9:359–380.
- Steeh, Charlotte, and Maria Krysan 1996 “The Polls—Trends: Affirmative Action and the Public, 1970–1995.” *Public Opinion Quarterly* 60:128–158.
- Tajfel, Henri 1969 “Cognitive Aspects of Prejudice.” *Journal of Social Issues* 4:79–97.
- Webb, Eugene J. 1981 *Nonreactive Measures in the Social Sciences*, 2nd ed. Boston: Houghton Mifflin.
- Wellman, David T. 1993 *Portraits of White Racism*, 2nd ed. New York: Cambridge University Press.
- Weitz, Shirley 1972 “Attitude, Voice, and Behavior: A Repressed Affect Model of Interracial Interaction.” *Journal of Personality and Social Psychology* 32:857–864.
- Wilder, David A. 1978 “Reduction of Intergroup Discrimination Through Individuation of the Out-Group.” *Journal of Personality and Social Psychology* 36:1361–1374.
- 1984 “Intergroup Contact: The Typical Member and the Exception to the Rule.” *Journal of Experimental Social Psychology* 20:177–194.
- Word, Carl O., Mark P. Zanna, and Joel Cooper 1974 “The Nonverbal Mediation of Self-Fulfilling Prophecies in Interracial Interaction.” *Journal of Experimental Social Psychology* 10:109–120.
- Zanna, Mark P. 1994 “On the Nature of Prejudice.” *Canadian Psychology* 31:11–23.

MARYLEE C. TAYLOR  
THOMAS F. PETTIGREW

## PROBABILITY THEORY

Sociologists, as much as researchers in any field perhaps, use a variety of approaches in the investigation of their subject matter. Quite successful and important are the historical and exegetical approaches and those in the traditions of anthropology and philosophy. Also of great importance are the systematic approaches that use mathematical models. Here the social investigator proposes a model, a mathematical depiction of social phenomena. A successful mathematical model can be very powerful, providing not only confidence in the theory from which the model was derived, giving us an explanation of the phenomena, but producing as well a method for predicting, giving us a practical means for controlling or affecting the social phenomena.

The social mathematical model is first of all a description of the relationship of the properties of social objects—groups, states, institutions, organizations, even people. If the model is derived from a theory, or if it contains features implied by a theory, and if the model fits data (i.e., has been found to satisfy some criterion of performance), the model can in addition be regarded as evidence to support that theory. In this case we can think of

a true, underlying model that generated the observations we are studying and a proposed model that will be tested against data. Quantitative analysis begins, then, with some theoretical understanding of the properties of groups of social objects; this understanding leads to the specification of a model of the interaction of these properties, after which observations of these properties on a sample of the objects are collected. The performance of the model is then evaluated to determine to what degree the model truly describes the underlying process.

The measurement of a property of a social object is called a *variable*. Variables can be either fixed or random. Fixed variables are those determined by the investigator; they usually occur in experiments and will not be of concern in this chapter. All other variables are random. The random nature of these variables is the unavoidable consequence of two things; first, the fact that our observations are *samples*, that is, groups of instances of social objects drawn from a population (that is, a very large number of possible instances to be observed); second, the fact that our theories and data collection are often unable to account for all the relevant variables affecting the variables included in the analysis. Probability theory in social models, or, equivalently, random variables in social models, will derive from these two subtopics: sampling and the specification of residual or excluded variables in the models.

A certain philosophical difference of opinion arises among probability theorists about the nature of the true source of the randomness in nature. One group argues that these features are inherent in reality, and another argues they are simply the consequence of ignorance. The primary modeling tool of the former group is the *stochastic process* (Chung 1974), while that of the latter is the *Bayesian statistical model* (de Finetti 1974).

**MAIN CONCEPTS**

*Probability* is a name assigned to the relative frequency of an event in an event space, that is, a set of possible events. For example, we might define the event space as the two sides of a single coin labeled heads (*H*) and tails (*T*): {*T*, *H*}. The actual outcome of a coin flip is a random variable, *X*, say, and the probability of the outcome *H* is  $P(X = H)$ . The probability distribution function (or PDF)

assigns a quantity to this probability. By definition, for a fair coin  $P(X = H) = .5$ . Since the event space is composed of only two events, then  $P(X = T) + P(X = H) = 1$ , that is, one or the other event occurs for certain, and  $P(X = T) = 1 - P(X = H) = .5$ . Thus the probability of *T* is equal to the probability of *H* and the coin is fair.

In general we assign numbers to the events in our event space, allowing us to use mathematical language to describe the probabilities. For example, the event space of the number of people arriving at a bank's automatic teller machine (ATM) is {0, 1, 2, . . .} over a given time interval  $\Delta t$ . Given certain assumptions, such as that the arrival time of each person is independent of anyone else's, we can derive a theoretical PDF. For a given time interval  $\Delta t$ , the probability of the number of people *X* can be shown to be

$$P(X = k | \Delta t) = \frac{(\lambda \Delta t)^k e^{-(\lambda \Delta t)}}{k!}$$

where  $\lambda$  is the mean rate of people arriving at the ATM over the time interval  $\Delta t$ , and  $k = 0, 1, 2, . . .$

Suppose from bank records we are able to determine that 100 people per hour complete a transaction at a particular ATM during normal working hours. For  $\Delta t$  equal to one minute or 1/60 an hour, the PDF is

$$P(X = k) = \frac{1.6667^k e^{-1.6667}}{k!}$$

For some selected values of *k* we have

k	P(X = k)
0	.1889
1	.3148
2	.2623
3	.1457
4	.0607
...	...

If we assume that each person spends about a minute at the ATM, we should expect one or more people standing in line behind someone at the



ATM about 50 percent of the time since the probability of two or more people arriving during a one minute interval is

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = .4963$$

The event spaces for the examples above are discrete, but continuous event spaces are also widely used. A common PDF for continuous event spaces is the normal distribution:

$$p(X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where  $\mu$  and  $\sigma^2$ , commonly called the mean and variance respectively, are parameters of the distribution, and  $x$  is a real number greater than minus infinity and less than plus infinity. The normal PDF is the most widely used distribution in social models, first because it had advantageous mathematical properties and second because its specification in many cases can be justified on the basis of the central limit theorem (Hogg and Tanis 1977, p. 155).

Other important concepts in probability theory are the *cumulative distribution function* (or CDF), *joint distributions* (distributions involving more than one variable), and *conditional distributions*. The CDF is the probability of  $X$  being less than or equal to  $x$ , that is,  $Pr(X < x)$ . An accessible introduction to probability may be found in Hogg and Tanis (1977).

### SAMPLING

In physics all protons behave similarly. To determine their properties, any given instance of a proton will do. Social objects, on the other hand, tend to be complex, and their properties can vary considerably from instance to instance. It is not possible to draw conclusions about all instances of a social object from a given one in the same manner we might from single instance of a proton. Given equivalent circumstances, we cannot expect everyone to respond the same way to a question about their attitudes toward political issues or to behave the same way when presented with a set of options.

For example, suppose we wish to determine the extent to which a person's education affects his

or her attitudes towards abortion. Let  $A_i$  represent a measurement of the attitude of some person, labeled the  $i$ th person, scored 0 if they are opposed to abortion or 1 if not. Let  $B_i$  be the measurement of the person's education, scored 0 for less than high school, 1 for high school but no college, or 2 for at least some college.

Given measurements on a sample of people, we would find that they would be distributed in some fashion across all the six possible categories of the two variables. Dividing the number that fall into each category by the total number in the sample would give us estimates of the empirical distribution for the probabilities:  $PR(A = 0, B = 1)$ ,  $PR(A = 0, B = 2)$ , and so on. We might also model this distribution. For example, an important type of model is the loglinear model (Goodman 1972; Haberman 1979; Agresti 1990), which models the log of the probability:

$$\log PR(A = i, B = j) = \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

where  $\lambda_i^A$ ,  $\lambda_j^B$  and  $\lambda_{ij}^{AB}$  are parameters (actually sets of parameters). In this model the  $\lambda_{ij}^{AB}$  parameters represent the associations between  $A$  and  $B$ , and an estimate of these quantities might have important implications for a theory.

Given a sample distribution, computing an estimate of  $\lambda_{ij}^{AB}$  is straightforward (Bishop, Fienberg, and Holland 1975). It is important, however, to realize that such an estimate is itself a random variable, that is, we can expect the estimate to vary with every sample of observations we produce. If the sample is properly selected, in particular if it is a simple random sample in which each person has an equal chance of being included in the sample, it can be shown that the estimates of  $\lambda_{ij}^{AB}$  have, in large samples at least, a normal distribution (Haberman 1973). Our estimates, then, are themselves parameters of a distribution, usually the means of a normal distribution. It follows that the fundamental parameters upon which a theory will depend can never be directly observed and that we must infer its true value from sample data.

All research on social objects is unavoidably research on samples of observations. Therefore all such research will necessarily entail at the very least a probabilistic sampling model, and the conclusions drawn will require properly conceived statistical inference.

**MODELS WITH EXCLUDED VARIABLES**

**The Regression Model.** The most well-known and widely used statistical model is the regression model. It is a simple linear hypersurface model with an added feature: a disturbance term, which represents the effects on the dependent variable of variables that have not been measured. To the extent that the claims or implications of a theory may be put into linear form, or at least transformed into linear form, the parameters (or regression coefficients) may be estimated and statistical inference drawn by making some reasonably benign assumptions about the behavior of the variables that have been excluded from measurement. The key assumption is that the excluded variables are uncorrelated with included variables. The failure of this assumption gives rise to *spurious* effects; that is, parameters may be under- or over-estimated, and this results in faulty conclusions. The statistical inference also requires a homogeneity of variance of the disturbance variables, called *homoscedasticity*. The variation of the excluded variables must be the same across the range of the independent variables. This is not a critical assumption, however, because the consequence of the violation of this assumption is inefficiency rather than bias, as in the case of the spurious effects. Moreover, the underlying process generating the heteroscedasticity may be specified, which would yield efficient estimates, or a modified inference may be computed, based on revised estimates of the variances of the distribution of the parameter estimates (White 1980).

For example, a simple regression of income, say, on years of education may be described,  $y_i = b_0 + b_1x_i + \epsilon_i$ , where  $y_i$  and  $x_i$  are observations on income and years of education, respectively, of the  $i$ th person  $b_0$  and  $b_1$  are regression coefficients, and  $\epsilon_i$  is the disturbance term. Estimates of  $b_0$  and  $b_1$  may be found (without making any assumptions about the functional form of the distribution of  $\epsilon_i$  by using perhaps the most celebrated theorem in statistics, the Gauss-Markov theorem, and they are usually called *ordinary least squares estimates*.

If we gather the observations into matrices, we can rewrite the regression equation as functions of matrices:  $Y = XB + E$ , where  $Y$  is an  $N \times 1$  vector of observations on the dependent variable.  $X$  an  $N \times K$  matrix of observations on  $K$  independent variables,  $B$  a  $K \times 1$  vector of regression coefficients,

and  $E$  an  $N \times 1$  vector of disturbances. With this notation the estimates in  $B$  may be described  $B^{\wedge} = (X'X)^{-1} X'Y$ , where; the “ $\wedge$ ” over the  $B$  emphasizes that they are estimates of the parameters.

Our observations are samples, and since our estimates of  $B$  will vary from sample to sample, it follows that these estimates will themselves be random variables. Appealing again to the Gauss-Markov theorem, it is possible to show that the ordinary least squares estimates have a normal distribution with variance-covariance matrix equal to  $\text{VarCov}(\hat{B}) = \sigma_{\epsilon}^2 (X'X)^{-1} X'Y$ , where  $\sigma_{\epsilon}^2$  is estimated by  $\sigma_{\epsilon}^2 = (Y - XB)'(Y - XB)/(N - K - 1)$ .

**Other models.** The regression model in the previous section is a “single equation” model, that is, it contains one dependent variable. A generalization of the regression model incorporates multiple dependent variables. This model may be represented in matrix notation as  $BY = \Gamma X + Z$ , where  $Y$  is an  $N \times L$  matrix of  $L$  *endogenous* variables (i.e., variables that are dependent in at least one equation),  $B$  is an  $L \times L$  matrix of coefficients relating endogenous variables among themselves,  $X$  is an  $N \times K$  matrix of  $K$  *exogenous* variables (i.e., variables that are never dependent),  $\Gamma$  is an  $L \times K$  matrix of coefficients relating the exogenous variables to the endogenous variables, and  $Z$  is an  $N \times L$  matrix of disturbances. Techniques have been developed to produce estimates and statistical inference for these kinds of models (Judge et al. 1982; Fox 1984).

Measurement error is another kind of excluded variable, and models have been developed to incorporate them into the regression and simultaneous equation models. One method for handling measurement error is to use multiple measures of an underlying *latent* variable (Bollen 1989; Jöreskog and Sörbom 1988). A model that incorporates both measurement error and excluded variable disturbances may be described in the following way:

$$\begin{aligned} Y &= \Lambda_y \eta + \epsilon_y \\ X &= \Lambda_x \xi + \epsilon_x \\ B\eta &= \Gamma \xi + \zeta \end{aligned}$$

where  $Y$  and  $X$  are our observations on the endogenous and exogenous variables respectively,  $\Lambda_y$ , and  $\Lambda_x$  are coefficient matrices relating the underlying variables to the observed variables,  $\eta$  and  $\xi$  are the

latent endogenous and exogenous variables respectively,  $B$  and  $\Gamma$  are coefficient matrices relating the latent variables among themselves  $\varepsilon_\gamma$  and  $\varepsilon_\zeta$  are the measurement error disturbances, and  $\zeta$  is the excluded variable disturbance.

This model incorporates three sources of randomness, measurement error disturbance, excluded variable disturbance, and sampling error. Models of the future may contain a fourth source of randomness: a structural disturbance in the coefficients. These latter models are called random coefficient models and are a special case of the most general kind of probabilistic model called the *mixture model* (Judge et al. 1982; Everitt 1984).

The models described to this point have been linear. Linearity can be a useful approximation that renders the problem tractable. Nonlinearity may be an important aspect of a theoretical specification, however, and methods to incorporate nonlinearity in large-scale models have been developed (Amemiya 1985). It also appears to be the fact that most social measures are not continuous, real variables, which is what is assumed by the regression and simultaneous models described above. Thus, much work is now being devoted to the development of models that may be used with measures that are limited in a variety of ways—they are categorical, ordinal, truncated, or censored, for example (Muthén 1984; Maddala 1983). Limited variable methods also include methods for handling variations on the simple random method of sampling.

Probability theory has had a profound effect on the modeling of social processes. It has helped solve the sampling problem, permitted the specification of models with excluded variables, and provided a method for handling measurement.

## REFERENCES

- Agresti, A. 1990 *Categorical Data Analysis*. New York: John Wiley.
- Amemiya, T. 1985 *Advanced Econometrics*. Cambridge, Mass.: Harvard University Press.
- Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland 1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT Press.
- Bollen, K. A. 1989 *Structural Equations with Latent Variables*. New York: John Wiley.

- Chung, K. L. 1974 *Elementary Probability Theory with Stochastic Processes*. Berlin: Springer-Verlag.
- de Finetti, B. 1974 *Theory of Probability*, 2 vols. New York: John Wiley.
- Everitt, B. S. 1984 *An Introduction to Latent Variable Models*. London: Chapman and Hall.
- Fox, J. 1984 *Linear Statistical Models and Related Methods*. New York: John Wiley.
- Goodman, L. A. 1972 "A General Model for the Analysis of Surveys." *American Journal of Sociology* 37:28–46.
- Haberman, S. J. 1973 "Loglinear Models for Frequency Data: Sufficient Statistics and Likelihood Equations." *Annals of Mathematical Statistics* 1:617–632.
- 1979 *Analysis of Qualitative Data*, vol. 2, *New Developments*. Orlando, Fla.: Academic.
- Hogg, Robert V., and Elliot A. Tanis 1977 *Probability and Statistical Inference*. New York: Macmillan.
- Jöreskog, K. G., and Dag Sörbom 1988 *LISREL VII*. Chicago: SPSS.
- Judge, George G., R. Carter Hill, William Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee 1982 *Introduction to the Theory and Practice of Econometrics*. New York: John Wiley.
- Maddala, G. 1983 *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press.
- Muthén, B. 1984 "A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators." *Psychometrika* 49:115–132.
- Tuma, Nancy Brandon, and Michael T. Hannan 1984 *Social Dynamics: Models and Methods*. Orlando, Fla.: Academic.
- White, H. 1980 "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817–838.

RONALD SCHOENBERG

## PROBATION AND PAROLE

The criminal justice system is the primary institution responsible for the formal social control of criminal deviance. Those who violate the criminal law are subject to a variety of sanctions, ranging from the reprimand of a police officer to execution by hanging. Most offenders are not apprehended, and among those who are arrested many are not prosecuted nor convicted of a crime. For offenders who are found guilty, either by trial or

more often by negotiated guilty plea, the sentence handed down by the court typically mandates correctional supervision, usually either some form of probation or incarceration with early release to some form of parole.

Even though probation and parole have been integral components of corrections since the nineteenth century, the differences between them are not always clear. Both are postconviction alternatives to incarceration that include supervision in the community by a probation or parole officer, who, depending on the jurisdiction, may be the same person. They are conditional releases to the community that are contingent on compliance with stipulated conditions, which if violated, may lead to revocation. Many probation and parole programs are similar (e.g., intensive supervision) or they share clientele. Last, as alternatives to incarceration, both are less expensive, less punitive, and probably more effective strategies of crime control.

The major difference between probation and parole is that probationers are sentenced directly to community supervision without being incarcerated, while parolees serve part of their sentence incarcerated before they are released to parole. Parole is a conditional release from confinement, whereas probation is a conditional suspension of a sentence to confinement. In both cases, a new crime or technical violation of conditions may lead to enhanced restrictions or incarceration. A general definition of *probation* is the conditional supervised release of a convicted offender into the community in lieu of incarceration (Allen et al. 1985). *Parole* is the conditional supervised release of an incarcerated offender into the community after serving part of the sentence in confinement (Clear and Cole 1986).

### PROBATION

Before informal probation was created in Boston in 1841 by philanthropist John Augustus, and the first statewide probation law was enacted in Massachusetts in 1978, convicted offenders were typically fined or imprisoned, often serving their full sentence. Probation was instituted as an alternative to incarceration at a time when jail and prison overcrowding became a critical management and humanitarian issue. Probation was considered a front-end sentencing solution to overcrowding,

intended specifically for less serious, first-time, and juvenile offenders amenable to "rehabilitation."

Over the years, rehabilitation has remained the primary goal of probation, and to this end, probation facilitates behavioral reform in a variety of ways. First, the often negative practical and symbolic consequences of the stigma of being an "ex-con" are neutralized. As less notorious and visible "probationers," the label will have less deleterious effects on the rehabilitative process. Second, the contaminating effects of imprisonment are avoided. This is particularly important for the less experienced offender, who may learn more about crime from more sophisticated, and sometimes predatory, inmates. The "pains of imprisonment" also produce anger, resentment, hostility, cynicism, and many other dysfunctional attitudes and feelings that make it more difficult to reform. Third, probation supports the existing social integration of the offender in the free community of noncriminals, including family, neighbors, employers and coworkers, friends, teachers and classmates, and others who are critical to the informal social control of crime. The offender released from incarceration will have the more difficult task of "reintegration." Fourth, the rehabilitative programs and services available to probationers are generally less coercive and more varied, flexible, and effective than those provided for prisoners. Fifth, the implied trust in leaving an offender in the community to demonstrate the ability to conform reinforces a positive attribution of self and expectations of appropriate behavior. Probation is more likely than incarceration to contribute to a self-fulfilling prophecy of rehabilitation.

Secondary goals of probation are more punitive. Probation is a penal sanction by virtue of the restrictions placed on the freedom of the offender. The conditions range from very lenient (e.g., weekly phone contact with a probation officer) to very punitive (e.g., twenty-four-hour home confinement), depending on the nature of the offense and offender characteristics. The goal of crime control can also be addressed by enhanced monitoring of probationers' compliance with the terms of probation, particularly their whereabouts. This can be accomplished by increasing the number and duration of meaningful (namely, face-to-face) contacts between probationer and probation officer, in the department's office, at home, at work,

in a residential or nonresidential program facility, or anywhere else in the community. More comprehensive control is possible with electronic monitoring devices; for example, transmitter anklets can verify the location of a probationer within, or outside of, a stipulated free movement area. The goal of deterrence is served to the extent that rehabilitation and punishment succeed in preventing probationers from committing more crimes and returning to the criminal justice system. Finally, justice is achieved when probation is the appropriate sentence for the offense and offender, and its application is equitable and uniform across race, sex, and socioeconomic status categories (McAnany et al. 1984).

The decision to grant probation is the product of a complex organization of legal actors, sentencing procedures, decision criteria, and system capacity. The decision may be initiated by the prosecutor, who negotiates a guilty plea in exchange for a recommendation to the judge that the defendant be sentenced to probation. Or it may await conviction at trial. In either case, a presentence investigation report prepared by a probation officer may support the recommendation by providing background information on the offender and an assessment of the public risks and prospects for probation success.

There are intense organizational pressures to minimize the number of trials and to divert convicted offenders from incarceration: There are huge case backlogs in the courts (Meeker and Pontell 1985) and tremendous overcrowding in jails and prisons, as evidenced by the almost forty states in 1988 that were under court order to reduce inmate populations in order to meet a variety of correctional standards (Petersilia 1987). Incredibly, it has been estimated that more than 90 percent of convictions for felonies are the result of negotiated guilty pleas (McDonald 1985), and a high percentage of those receive probation since, by state, from 25 to 70 percent of convicted felons are sentenced to probation (Petersilia 1985). It is clearly in the personal and organizational interests of the defendant, prosecutor, judge, and even the jailer and prison superintendent to support "copping" a plea for probation.

Whether the conviction is negotiated or decided at trial, the judge sentences the offender, within

the constraints imposed by the sentencing model and guidelines used in the jurisdiction. Most states use indeterminate sentencing, where judges have substantial discretion in rendering sentences and parole authorities are responsible for release decisions of incarcerated offenders. The trend is toward determinate sentencing, where judges and parole boards have much less influence on sentence and release decisions. In both models, probation is a widely used sentencing option, especially for less serious offenders but even for many serious offenders: Nationally, as high as 20 percent of violent offenders receive probation, including 13 percent of defendants convicted of rape and 9 percent of those convicted of homicide (Lisefski and Manson 1988).

Despite the confluence of the trend toward determinate sentencing, more pervasive justice model practices in corrections, and increased public pressure to be more punitive with criminals, there are relatively more offenders on probation than incarcerated or on parole than there were two decades ago (Petersilia et al. 1985). More serious offenders are being incarcerated for more fixed sentences, but the concomitant institutional overcrowding has produced a greater utilization of probation, as well as a variety of types of probation designed to meet the needs of both less serious and middle-range offenders, who in the past would have been more likely to be incarcerated. In addition to "standard probation," characterized by assignment to a probation officer with a caseload of as many as two hundred probationers and nominal contact, supervision, and rehabilitative services, a whole range of "intermediate sanctions" has been created that includes programs that are typically more punitive, restrictive, intensive, and effective than standard probation (Petersilia 1987; Morris and Tonry 1990). Judges now have a diversity of probation alternatives at sentencing: intensive supervised probation, home confinement, electronic monitoring, residential centers (half-way houses), and split sentences (jail/probation). These alternatives are often combined and coupled with other probation conditions that require restitution to victims, employment or education, payment of program costs, random urinalysis, specialized treatment or classes (e.g., Alcoholics Anonymous), or community service. Many of the intermediate sanctions are also used in the supervision of parolees.

## PAROLE

Like probation, parole in the United States was created to relieve the serious overcrowding problem in prisons at the beginning of the nineteenth century. Years before informal probation began, some prison wardens and correctional administrators were releasing prisoners before their full sentence was served. They were either released outright, much as if they had received a pardon, or monitored informally by the police. Based on the European correctional innovations of “good time” and “ticket of leave,” formal parole emerged toward the end of the nineteenth century, with the first indeterminate sentencing law passed in 1876 in New York (Champion 1990).

Until Maine abolished parole in 1976, all states had indeterminate sentencing and parole authorities. In general, within these systems a prisoner earns good time by productive participation in institutional programs and good conduct. The accumulated good time is subtracted from the sentence to determine the time incarcerated. This decision is typically made by a parole board, which is often a group of political appointees from a variety of occupations and constituencies. The offender is then released (or awarded a leave) to the supervision of a parole officer. If an offender does not commit a new crime or violate the conditions upon which release to parole is contingent, he or she can complete the remainder of the sentence as a parolee, to the time of discharge and freedom.

The goals of parole are anchored in indeterminate sentencing and the tenets of the rehabilitative ideal. It is assumed that offenders are amenable to reformation, through both the rehabilitation provided by the prison’s treatment, educational, and vocational programs and the reintegration back into the free world facilitated by the transitional programs and services of parole. These twin primary goals of “rehabilitation” and “reintegration” drive the decisions and actions of the parole system. Parole is granted when the prisoner is considered ready for release, based on behavior during confinement, the extent to which rehabilitation is evident, and the apparent risk to public safety. In practice, many offenders spend a relatively small proportion of their sentence incarcerated; for instance, a convicted murderer with a life sentence

may “do hard time” for as few as, say, ten years and serve the rest of the sentence on parole. Parole is revoked, or modified, when there are indications that reintegration is in jeopardy or unlikely, owing to violation of parole conditions. A new crime, in particular, but even a technical violation may be sufficient for the parole board to reincarcerate a parolee. The parole board also has the discretionary authority to set dates for parole hearings, fix minimum terms and release dates, determine good time credits, and specify parole conditions and requirements.

Like probation, the secondary goals of parole include punishment, crime control, and deterrence. After all, parole is a part of the penal sanction defined by the sentence to imprisonment, and, depending on the type of parole and stipulated conditions, the parole experience can be very restrictive and quasi-custodial. Effective rehabilitation, supervision, and monitoring of parolees should also produce deterrent effects—the combination of reformation and punishment should prevent future criminal conduct among parolees.

Unfortunately, by the mid-1970s, evidence had accumulated that suggested that parole was not an especially effective crime control strategy (Martinson 1974), and the shift away from the rehabilitative ideal to a more punitive justice philosophy (von Hirsch 1976) began in earnest. About one-third of states have returned to some form of determinate sentencing, and more are likely to follow. The discretionary power of judges in rendering sentences and of parole boards in implementing them has been abridged, in order to make decisions more rational and fair by linking them to the severity of the offense rather than the characteristics of the offender. Offenders are now more likely to be serving sentences in confinement; since 1980, the rate of incarceration has increased by 76 percent (Bureau of Justice Statistics 1989). They are also less likely to be placed on parole by a paroling authority; between 1977 and 1987, releases from imprisonment decided by parole boards dropped from about 70 percent to 40 percent of all releases, while mandatory releases increased from roughly 5 percent to 30 percent of the total (Hester 1988). With determinate sentencing, many states simply do not have paroling authorities or parole supervision in the community.

These changes have also affected the types of parole that are still available in a majority of states. Although many parole and probation programs are similar, the usually more serious offenses and criminal histories of parolees and the shift toward more punitive correctional systems have led to a hardening of the conditions of (cf. U.S. Sentencing Commission 1987), less utilization of "standard parole," and a greater emphasis on protecting public safety by extending custody from the institution to the community. The goals of rehabilitation and reintegration have become less important, while crime control and deterrence have become more important. Consequently, there is greater reliance on transitional programs that maximize monitoring and supervision, while providing opportunities and services (e.g., employment, school, counseling, drug treatment) intended to facilitate reentry into the community and desistance from crime during parole and, ultimately, after discharge from correctional supervision. These programs are more intensive and custodial, often involving residential placement in a halfway house, intensive parole supervision, or home confinement with electronic monitoring. From these community bases, parolees may participate in work or school releases, home furloughs, counseling, religious services, and a variety of other reintegrative activities.

Although parole is not used to the extent that it was before the advent of determinate sentencing, and there are those who believe that it should be abolished in all states, some research suggests that determinate sentencing is no more a panacea than prior correctional reforms. There may be a leveling of sentencing disparities, and more offenders are being incarcerated. But they, on the average, are doing less time and, after release, may be as likely to be reconvicted and reincarcerated (Covey and Mande 1985).

### RESEARCH

While the research evidence on the efficacy of determinate sentencing may be sketchy, there are many studies of other issues in probation and parole that have produced more substantive results. Social scientific research on probation and parole has tended to revolve around a set of related issues that are common to both: program effectiveness, recidivism, and classification and

prediction. The overriding empirical and policy question is "What works?" Attempts to address the question vary in rigor and quality, and the answers are neither direct nor simple.

There are innumerable studies of program effectiveness, most of them not producing useful, much less compelling, evidence of program success or failure. For example, many studies conclude that a probation program is successful because 30 percent of participants recidivate or that a parole program is successful because 40 percent of participants recidivate. There are serious problems with those kinds of studies. First, they do not compare the program being evaluated with others, either with other probation or parole programs, or across correctional alternatives (e.g., release, probation, incarceration, and parole).

Second, without comparison groups, one can only evaluate program effectiveness in relation to some standard of success. But preordained acceptable levels of recidivism are determined normatively, not empirically. Normative criteria of success cannot be applied uniformly across the incredible variation in probation and parole programs. For instance, some probation programs, because of the very low risk participants, should probably generate recidivism rates that are closer to 5 percent than 30 percent.

Third, recidivism is often not defined or measured adequately. Generically, recidivism has come to mean "reoffending," particularly by offenders who have had contact with the criminal justice system, as measured typically by rearrest, reconviction, reincarceration, or some variation or combination thereof. But what does a probationer or parolee relapse to, and what is the most appropriate and accurate measure? The answers are complicated by the fact that probation and parole can be revoked if an offender commits a crime that becomes known to criminal justice authorities or by a noncriminal violation of release conditions (a "technical violation"). Paradoxically, practically all studies ignore the substantial amount of successful criminal behavior that remains hidden from officials, but many use both revocation criteria as measures of recidivism. Which measure is used can dramatically affect judgments about program effectiveness. Evaluations of three intensive probation supervision programs show that technical violations, as compared to new crimes, account for

the majority of revocations. The revocation rates due to technical violations for these programs were 56 percent, 70 percent, and 85 percent, respectively. However, if the technical violations are not counted in the recidivism rates, the rates drop to 7 percent, 8 percent, and 5 percent, respectively (Petersilia 1987). Depending on program objectives, recidivism, no matter how measured, may not be the only or most appropriate criterion of program effectiveness. What may also be useful is assessment of relative costs and savings, impacts on jail and prison overcrowding, effects on public perceptions of safety, performance at school or in the workplace, changes in offenders' attitudes and self-concept, and so on.

More rigorous studies utilize comparison groups in order to assess the relative effectiveness of different correctional strategies. A typical study compares the recidivism rates of various combinations of offenders on probation, incarcerated, and on parole, and concludes that probationers are least likely and ex-prisoners are most likely to recidivate. Of course, one would predict those results based on the differences between the groups in their risk to recidivate. The selection biases of the court place the least serious, low-risk offenders on probation and the most serious, high-risk offenders in institutions. The observed differences in recidivism do not reflect the relative effectiveness of the programs, but the original differences in the recidivism risks of the comparison groups.

Some studies attempt to produce more comparable groups by using more objective probation and parole prediction instruments to classify and then compare offenders by level of recidivism risk across programs. That is, they try improve comparability by "matching" offenders within the different program groups. For example, a study of the relative effectiveness of standard probation, intensive supervision probation, and incarceration with parole classified offenders within each group into low-, medium-, high-, and maximum-risk levels. Comparisons of recidivism, measured by rearrest, reconviction, and reincarceration, across the three program alternatives for each of the four categories of offenders (namely, least likely to most likely to recidivate) show that parole is least effective in preventing recidivism at all levels of recidivism risk. The differences between standard and intensive probation are not as consistent: no matter how recidivism is measured,

the rate is higher among intensive supervision probationers who are low- and high-risk offenders; among medium-risk offenders, it varies by the measure of recidivism; and for maximum-risk probationers, there seems to be little difference in the effectiveness of standard or intensive supervision, except for the somewhat higher reincarceration rate among intensive supervision probationers (Erwin 1986).

The equivocal findings of this and many similar studies reflect the difficulty in predicting recidivism risk with any degree of accuracy. The most comprehensive and statistically sophisticated techniques (e.g., cluster analyses, linear models, complex contingency tables) are not much more accurate than bivariate tabular procedures developed seventy years ago by Ernest Burgess. And no technique is able to predict recidivism with higher than 70 percent accuracy, with most slightly better than chance (Blumstein et al. 1986). Therefore, it is virtually impossible to make groups comparable on the basis of recidivism risk, or any other prediction criteria, which compromises the validity of the findings regarding differential program effectiveness.

The mixed results probably also reflect the paradox of intensive supervision programs in general: Increasing supervision and monitoring may increase, rather than decrease, the probability of recidivism. The offender is at greater risk to recidivate, simply because there is a better chance that unacceptable conduct will be observed. However, depending on the declared program goals, this may indicate success rather than failure: If intensive supervision of probation and parole are intended to enhance crime control and public safety, rather than to rehabilitate, higher rates of recidivism may demonstrate program effectiveness (Gottfredson and Gottfredson 1988).

Research on probation and parole effectiveness cannot produce compelling findings from studies that depend on comparisons of typically noncomparable groups. What is necessary are "equivalent" groups that are created through random assignment within experiments. Unfortunately, experimental designs are usually more expensive and more difficult to implement and complete in natural settings. Consequently, they are extremely rare in research on probation and parole. For instance, there are more than one hundred



studies of the effectiveness of intensive probation supervision, but none have an experimental design with random assignment to program conditions (Petersilia 1987). There are some current efforts to implement studies of probation and parole that have experimental designs, but if the objective is to produce valid and useful knowledge on "what works," there must be a greater commitment on the part of the criminal justice system, funding agencies, and the social science research community.

The field of probation and parole research and study will continue to be important for social scientists. While there is no indication that the description and analysis presented here is changing at this time, the concern with the issues is active. Some more recent writings add to the character of the presentation here, and are well worth reviewing for persons who wish to pursue the development of the field. (See Albonetti and Hepburn 1997; Heilbrun 1999; Geerken and Hayes 1993; Gendreau et al. 1994; McCorkle and Crank 1992; and Turner et al. 1992.)

(SEE ALSO: *Court Systems of the United States*; *Criminal Sanctions*; *Criminology*; *Social Control*)

## REFERENCES

- Albonetti, Celesta, and John Hepburn 1997 "Probation Revocation: A Proportional Hazards Model of the Conditioning Effects of Social Disadvantage." *Social Problems* 44(1):124-138.
- Allen, Harry E., Chris Eskridge, Edward Latessa, and Gennaro Vito 1985 *Probation and Parole in America*. New York: Free Press.
- Blumstein, Alfred, Jacqueline Cohen, Jeffrey Roth, and Christy Visher (eds.) 1986 *Criminal Careers and "Career Criminals"*. Washington, D.C.: National Academy Press.
- Bureau of Justice Statistics 1989 *Prisoners in 1988* (bulletin). Washington, D.C.: U.S. Department of Justice.
- Champion, Dean J. 1990 *Probation and Parole in the United States*. Columbus, Ohio: Merrill.
- Clear, Todd, and George F. Cole 1986 *American Corrections*. Belmont, Calif.: Brooks/Cole.
- Covey, Herbert C., and Mary Mande 1985 "Determinate Sentencing in Colorado." *Justice Quarterly* 2:259-270.
- Erwin, Billie S. 1986 "Turning Up the Heat on Probationers in Georgia." *Federal Probation* 50:17-24.
- Geerken, Michael, and Hennessey Hayes 1993 "Probation and Parole: Public Risk and the Future of Incarceration Alternatives." *Criminology* 31(4):549-564.
- Gendreau, Paul, Francis Cullen, and James Bonta 1994 "Intensive Rehabilitation Supervision: The Next Generation in Community Corrections?" *Federal Probation* 58(1):72-78.
- Gottfredson, Michael, and Don M. Gottfredson 1988 *Decision Making in Criminal Justice*. New York: Plenum.
- Heilbrun, Alfred 1999 "Recommending Probation and Parole." In Allen Hess and Irving Weiner, eds., *The Handbook of Forensic Psychology*. New York: John Wiley.
- Hester, Thomas 1988 *Probation and Parole, 1987*. Washington, D.C.: U.S. Department of Justice.
- Lisefski, Edward, and Donald Manson 1988 *Tracking Offenders*. Washington, D.C.: U.S. Department of Justice.
- McAnany, Patrick D., Doug Thomson, and David Fogel (eds.) 1984 *Probation and Justice: Reconsideration of a Mission*. Cambridge, Mass.: Oelgeschlager, Gunn and Hain.
- McCorkle, Richard, and John Crank 1997 "Meet the New Boss: Institutional Change and Loose Coupling in Parole and Probation." *American Journal of Criminal Justice* 21(1):1-26.
- McDonald, William F. 1985 *Plea Bargaining: Critical Issues and Common Practices*. Washington, D.C.: U.S. Department of Justice.
- Martinson, Robert 1974 "What Works? Questions and Answers about Prison Reform." *Public Interest* 35:22-54.
- Meeker, James, and Henry M. Pontell 1985 "Court Caseloads, Plea Bargains, and Criminal Sanctions: The Effects of Section 17 P.C. in California." *Criminology* 23:119-143.
- Morris, Norval, and Michael Tonry 1990 *Between Prison and Probation: Intermediate Punishments in a Rational Sentencing System*. New York: Oxford.
- Petersilia, Joan 1985 *Probation and Felony Offenders*. Washington, D.C.: U.S. Department of Justice.
- 1987 *Expanding Options for Criminal Sentencing*. Santa Monica, Calif.: Rand.
- , Susan Turner, James Kahan, and Joyce Peterson 1985 *Granting Felons Probation: Public Risks and Alternatives*. Santa Monica, Calif.: Rand.
- Quinn, James, and John Holman. "Electronic Monitoring and Family Control in Probation and Parole." *Journal of Offender Rehabilitation* 17(3-4):77-87.
- Turner, Susan, Joan Petersilia, and Elizabeth Deschenes 1992 "Evaluating Intensive Supervision Probation/Parole (ISP) for Drug Offenders." *Crime and Delinquency* 38(4):539-556.

U.S. Sentencing Commission 1987 *United States Sentencing Commission Guidelines Manual*. Washington, D.C.: U.S. Sentencing Commission.

von Hirsch, Andrew 1976 *Doing Justice*. New York: Hill and Wang.

JOSEPH G. WEIS

## PROFESSIONS

The idea of a “profession” did not exist in ancient times. Although there were people who did what is currently denoted as professional work, these “professionals” often labored in dependent positions. For example, physicians in the Roman Empire were slaves in wealthy households, and architects worked as salaried public employees. Lawyers in ancient Greece were merely friends of the litigants who spoke before a gathering of their peers. Neither lawyers nor physicians received formal training (Carr-Saunders and Wilson 1937). By medieval times, the three classic professions—medicine, law, and the clergy (which included university teaching)—began to approximate more closely the modern conception of professions. With the development of universities, then under religious auspices, would-be professionals completed lengthy training in their chosen fields. They also began to constitute a new class of intellectuals. As society increasingly secularized, the professions emerged from under religious control and began to organize professional associations. By the eighteenth century, they had achieved independent status.

In the nineteenth century, middle-class occupations such as dentistry, architecture, and engineering began to professionalize, aspiring to the gentlemanly status of the classic, learned professions (Dingwall and Lewis 1983). “Gentlemanly” was the appropriate description, since the majority of professionals were men. It was not until the 1970s that women began to make significant inroads into these occupations, and even today men predominate in the status professions.

In the nineteenth and early twentieth centuries, professionalism developed in concert with the increasing division of labor and rationalization characteristic of industrializing Europe and the United States. As competitors in market economies, occupational incumbents sought to

professionalize to improve their status. In addition, they wanted to better their economic positions by securing occupational niches for their services (Ritzer and Walczak 1986). In the United States, at least, universities played the key role of transferring both the technical know-how and the culture of professionalism to new generations of professional aspirants (Bledstein 1976).

### WHICH OCCUPATIONS ARE PROFESSIONS?

Today the term “profession” includes a range of occupations arrayed along a continuum of high to medium levels of prestige. At the high end of the continuum are the classic, or “status,” professions of medicine, law, clergy, and university teaching. Incumbents in these occupations usually receive high incomes, exercise job autonomy, and receive deference from the public and those lower in the status hierarchy. Although women have begun to gain entry into some of the professions, most professions remain predominantly male: in 1998, women’s representation among physicians, lawyers, and the clergy was 27, 28, and 12 percent, respectively. Women have made more progress moving in to college and university teaching, representing 42 percent of incumbents by 1998 (all 1998 data are from U.S. Bureau of Labor Statistics 1999, Table 11).

Somewhat lower on the continuum are the “newer” professions such as dentistry, engineering, accounting, and architecture, which also command respect and relatively high salaries. Men also predominate in these occupations: 20, 11, and 18 percent of dentists, engineers, and architects, respectively, were women in 1998. Only in accounting have women made significant inroads: 58 percent of all accountants and auditors were women in 1998. This feminization is attributable largely to accounting’s dramatic occupational growth in the 1970s (Reskin and Roos 1990, pp. 40–41).

Still lower on the professional continuum are the “marginal professions” (e.g., pharmacy, chiropractic) and the “semiprofessions” (e.g., nursing, public school teaching, social work, librarianship). These occupations exhibit some characteristics of the classic professions but have not acquired full professional status because of opposition from

established professions and an inability to convince the public that they command unique expertise. These occupations are less prestigious and their incumbents are paid less than those in either the classic or the new professions. Moreover, because they are concentrated in bureaucratic settings, these professionals exercise less job autonomy than higher-status professionals. An important difference between the marginal professions and the semiprofessions is that males are in the majority in the former (although pharmacy, an occupation that has feminized rapidly in recent years, is now 44 percent female), whereas semiprofessions have long been predominantly female (in 1998 women composed 92, 75, 68, and 83 percent of nurses, public school teachers, social workers, and librarians, respectively).

Paraprofessionals work with, but as subordinates to, members of the other professions. They are generally technicians associated with various professional occupations. Paralegals, for example, work closely with lawyers, and physicians delegate certain tasks and responsibilities to physicians' assistants. As in the semiprofessions, women tend to predominate in paraprofessional occupations. These data clearly show the sex-stratified nature of professional occupations, but these occupations are also stratified by race: in 1998, for instance, 5 percent of physicians were black, 6 percent of college and university teachers, 4 percent of lawyers, and 9 percent of clergy, notably less than their 11 percent representation in the labor force as a whole. Blacks are overrepresented, relative to their labor-force representation, in the following professions: dietitians (18 percent); respiratory therapists (12 percent); prekindergarten and kindergarten teachers (14 percent); educational and vocational counselors (13 percent); social workers (23 percent); recreation workers (16 percent); athletes (13 percent).

### THE PROCESS OF PROFESSIONALIZATION

Given the stratification of the U.S. occupational structure, it is clear why workers desire to professionalize. Professionalization brings higher income, higher prestige, greater job autonomy, and higher job satisfaction. It also protects incumbents from competition. The "process of professionalization" posits a common sequence of development that occupations undergo. Some

scholars accept Harold Wilensky's depiction of this process (1964). First, people begin to work full time at a specific set of tasks that will form the new occupation's core jurisdiction. Second, those in the occupation establish a university-affiliated training program, and some incumbents undertake the responsibility for training new generations of practitioners. Third, practitioners and teachers combine to form a professional association that identifies the occupation's core tasks and makes claims regarding skill jurisdiction. Fourth, occupational incumbents seek to protect their jurisdictional claims by political means. Professionals lobby for legal protection, in the form of licensing and certification requirements, to generate labor-market shelters that ensure their monopoly of skills. Finally, incumbents develop a formal code of ethics that embodies rules to protect clients, eliminate the unqualified, and spell out the occupation's service ideal. As we shall see, later theorists have come to question this linear sequence of events.

### APPROACHES TO THE STUDY OF THE PROFESSIONS

What distinguishes the professions from other occupations? Theoretical approaches in studying this question have changed over time and remain in flux. Scholars have also developed new methods to address these questions.

**The Trait Approach.** After World War II, the trait approach was dominant in scholarship on the professions (Freidson 1986). Scholars—mostly American academics—tried to define the professions by generating an exhaustive list of characteristics. These traits, scholars hoped, would distinguish professions from nonprofessions and higher-status professions from those of lower status. The main method used was the case study.

Scholars carefully scrutinized particular occupations to determine how well they approximated the four major criteria, or traits, of the ideal-typical profession (Hodson and Sullivan 1990). First, professionals are experts with abstract, esoteric knowledge and skills that set them apart from others. Second, because of their unique expertise, professionals exercise autonomy on the job. Codes of ethics help to ensure autonomy from outside control by permitting professionals to police misconduct internally. Third, their esoteric knowledge

allows professionals to claim authority over their clients and subordinate occupational groups. Finally, the professions are altruistic, that is, service-rather than profit-oriented. Underlying these four traits is a fifth: the public must recognize the occupation as a profession. Regardless of an occupation's claim of unique expertise, if the public does not view the occupation's knowledge as abstract, it is difficult for those working in it to claim professional status and the perquisites that accompany it.

Scholars used these criteria to differentiate among the professions, most particularly in comparing the female semiprofessions to the typically male status professions. While the semiprofessions have a body of knowledge, they lack a monopoly over that knowledge. They also have a difficult time convincing the public that their skills are professional. The public is less likely to recognize their expertise (e.g., teaching children, servicing library patrons) as particularly esoteric. Semiprofessionals typically work in bureaucratic settings and, as a consequence, are subject to heteronomy, or supervision by organizational superiors and professional colleagues. Their ability to claim autonomy is limited.

**The Power Approach.** In the 1960s, scholars in the United States and Great Britain began to criticize the trait approach as static and ideological. Recognizing the salience of culture and social structure, power theorists also shifted to historical methods to understand the sources of professional power. They argued that occupations we view as professions do not necessarily exhibit the requisite traits, but rather have the power to convince the public that they do. Power theorists viewed the professions as monopolistic organizations intent on gaining and retaining professional control and ensuring their status in the stratification system. Eliot Freidson (1986), for example, focused on how professions establish protected labor markets for their services. Magali Sarfatti Larson (1977) argued that the professions are market organizations in the capitalist economy, explicitly seeking to dominate the market for their expertise.

For these theorists, the so-called objective characteristics of the trait approach are ideological attempts to preserve the professions' status and privilege (Freidson 1986). Rather than being truly altruistic, professional incumbents create the myth

of service orientation to gain public goodwill, enhance their status, and minimize external control. According to this view, professionals sometimes abuse their autonomy by failing to police themselves, and incompetent doctors and lawyers fleece the public with little fear of reprisal from their peers. Finally, the professional's authority over clients has also declined in recent years as a more educated public has become active in activities seen as the province of professionals (e.g., getting second opinions on medical recommendations and becoming more educated consumers regarding medical and legal issues).

The historical battles between physicians, on the one hand, and pharmacists and chiropractors on the other, illustrate how an established profession exercises power against competing occupations (Starr 1982). When the pharmacist's traditional task of compounding drugs shifted to pharmaceutical companies, they lost their diagnostic expertise (and hence monopoly over their knowledge). With respect to chiropractors, the American Medical Association restricted access through licensing laws or blocked reimbursement from private insurance companies. As a consequence, pharmacy and chiropractic have yet to become fully professional.

What enables professions to wield power? Theorists point to the occupational characteristics indeterminacy and uncertainty (Ritzer and Walczak 1986). The professions that have achieved and maintained power are those whose tasks cannot be broken down or otherwise routinized (indeterminacy). Similarly, those that deal with areas of uncertainty are also likely to preserve their power. As Wilensky (1964) described it, professional knowledge involves a "tacit" dimension, in Polanyi's (1967) terminology. Their lengthy training and years of practical application ensure that physicians "know" what treatments to use for which symptoms. Similarly, lawyers "know" what legal strategies work best and college professors "know" what instructional strategies are most effective in enhancing student learning. Reading a textbook or consulting computerized data bases is not equivalent to tacit knowledge. This kind of knowledge—expertise refined by years of experience—is not easily routinized. Physicians and lawyers also deal with areas of high uncertainty for their clients, the former with physical health and the latter with legal affairs. Clients need these

professionals to translate medical and legal jargon into everyday language they can more readily understand.

**The System of Professions.** Andrew Abbott's (1988, 1995) theory of professions critiqued the notion that professions undergo a common process of development. Employing historical and comparative methods, Abbott provided a wealth of evidence that the history of the professions is much more complicated than that represented in a linear process of professionalization. Rather than a history of professions that established systems of control (e.g., schools, professional associations, licensing and certification), Abbott's account is of ongoing interprofessional competitions, squabbles over jurisdictions, and professional births and deaths.

Abbott contended that professions make up an interdependent "system of professions." Understanding modern professions entails articulating their histories of conflict with other professions. Comprehending the realities of modern medicine, for example, depends more on investigating its historical conflicts with closely related professions such as psychiatry and chiropractic than on the particulars of medieval or nineteenth-century medicine.

Recognizing the interdependence of professions is important because it clarifies that professions emerge, grow, change, and die within the historical context of competition with other professions. This competition takes the form of jurisdictional disputes over the control of abstract knowledge. Professionals can use their abstract knowledge to define a core set of tasks (their jurisdiction), defend that jurisdiction from others, or appropriate the tasks of others. Interprofessional boundary wars reflect professions' attempts to "enclose" jurisdictional tasks within the boundaries of their profession's sphere of influence (Abbott 1995, p. 553; see also Witz 1992). One recent example of such a jurisdictional dispute was pediatric medicine's partially successful attempt to expand their sphere of influence to include children's psychosocial disorders (Halpern 1990). Whether professions succeed or fail in such jurisdictional disputes, any changes reverberate throughout the system. The professions reequilibrate, with some occupations accepting a subordinate or advisory role, some agreeing to split jurisdictions or clients, and others exiting the professions altogether. To

adequately theorize and model jurisdictional disputes, Abbott (1993, p. 205) advocated a multilevel analysis that links micro-level information on careers, to meso-level data on the network structure of careers and jobs, to the macro-level work and occupational structures.

### SEX DIFFERENCES IN THE PROFESSIONS

In 1998, women composed 46 percent of the employed labor force and 53 percent of professionals, suggesting that women do quite well in the professions. However, a closer look reveals a different story. As noted, women are heavily concentrated in the semiprofessions (as nurses, public school teachers, social workers, and librarians) and men in the higher-status professions (as physicians, lawyers, and engineers). Incumbents in the former earn less and exercise less autonomy than in the latter. As in the occupational structure as a whole, the professions are highly segregated by sex.

Even within the higher-prestige professions, women work in different, lower-paying, and less prestigious jobs than men. Women lawyers, for example, work in government jobs, in research rather than litigation, and in certain specialties such as trust and estates; women physicians are more likely than men to specialize in pediatrics and to work in health maintenance organizations (HMOs); female clergy specialize in music or education (Reskin and Phipps 1988; see also Tang and Smith 1996).

Part of the reason for this differential job distribution by sex has to do with a characteristic unique to the professions. The high level of uncertainty inherent in prestigious professional jobs means that employers are careful to choose recruits who "fit in" with those already on the job. Thus, as Rosabeth Moss Kanter (1977) suggested, employers tend to recruit people much like themselves, a process she calls "homosocial reproduction." The predominance of males in the status professions thus helps to reproduce sex segregation.

Other factors reducing women's access to high-status professions are entrance restrictions such as certification and licensing. Physicians, for example, consolidated their control over medical jurisdictions by successfully pressing for legislation to outlaw midwives and prohibit the licensing of those trained at "irregular" schools, activities that

disproportionately affected women. In 1872, the Supreme Court restricted women's ability to practice law, arguing that "the natural and proper timidity and delicacy which belongs to the female sex unfits it for many of the occupations of civil life" (Reskin and Phipps 1988, p. 192). Male professionals were thus able to establish labor-market shelters to protect themselves from competition from women as well as other "undesirables."

Since 1970, women have gained greater entry into some of the professions, including occupations such as medicine, law, and pharmacy. Indications are, however, that internal differentiation within the professions perpetuates job segregation by sex (Reskin and Roos 1990). For example, Polly Phipps (1990) found that women's representation in pharmacy nearly tripled (from 12 to 32 percent) between 1970 and 1988, and has continued to increase since then, reaching 44 percent by 1998. However, women pharmacists concentrate in the lower-paying hospital sector, while men predominate in the higher-paying retail sector. Similar ghettoization exists in other professions that are admitting more women (see, for example, Reskin [1990] on book editors; Donato [1990] on public relations specialists; Roos and Jones [1993] on academic sociologists; and Roos and Manley [1996] on human resource managers and professionals).

### THE CHANGING PROFESSIONS

Some view the future of the professions as bleak, predicting that ongoing proletarianization or deprofessionalization will eliminate the professions' unique traits. The proletarianization thesis argues that an increasing division of labor and a bureaucratization within the professions are routinizing knowledge and transferring authority from professionals to organizational superiors. The deprofessionalization thesis documents declines in the professions' monopolistic control over their knowledge, their exercise of autonomy on the job, their ability to protect their jurisdiction from encroachments, and the public's deference to professional authority (Ritzer and Walczak 1986).

Some occupations, of course, have lost some of their professional status. As noted, as pharmaceutical companies increasingly absorbed the compounding of drugs, and as chains replaced independent pharmacies, pharmacists lost some of

their autonomy and monopoly of their knowledge to physicians. Taking a broad view of the professions, however, Freidson (1984; 1994, p. 9) argued that professionalism has been "reborn" in a hierarchical form in which bureaucratization and professionalization are often quite compatible. Working in organizations, he argued, has been the norm for many professions from their inception, with engineers the most obvious example. In addition, organizations that employ professionals tend to diverge enough from the ideal-typical bureaucracy to protect professional privilege. For example, professionals in organizations often exercise a lot of autonomy, working under senior members of their own profession rather than nonprofessional managers. Freidson thus portrayed most organizations as working to accommodate professionals, and operating under an "occupational principle" of authority, as opposed to an "administrative principle" of organizing and controlling work (1994, p. 61). Wallace (1995) also described how some bureaucratic work systems—specifically a "corporatist" model of control—can be quite compatible with professionals' self interests.

Like Freidson and Abbott, Brint (1994) viewed the professions as social forms that evolve historically in conjunction with other occupations and work organizations. Rather than look to professions interacting through jurisdictional disputes, however, Brint contextualized changes in the professions within the workplace itself, in the occupationally and organizationally based ties professionals have, and the markets in which they work (see also Leicht and Fennell 1997). He described the ascendance of a new, organizationally based profession—expert professions—that relaxes traditional assumptions about professional work. Pursuing profits, closer interconnections with business, and lesser attention to larger public interests, are all hallmarks of "expert professionalism" (Brint 1994, p. 8), which Brint juxtaposed with "social trustee professionalism," the traditional conception of professions as the trustees of socially important knowledge (p. 4). In Brint's view, the professional class has splintered into highly skilled experts in resource-rich organizations as opposed to professionals with less marketable skills in resource-poor organizations, especially in the public and nonprofit sectors (p. 11). These differing organizational locations have predictable consequences for incumbents' political and social views.

Even with such changes, Freidson (1984, 1994) found no evidence that the prestige of the professions as a whole has declined. Nor did he find that public trust in professionals has deteriorated relative to other American institutions. Moreover, the professions' continuing ability to erect labor-market barriers to competition is important evidence of their enduring power. Professionals today remain strong enough to exert their will against others lower in the occupational hierarchy. Professional privilege remains intact in the American occupational structure.

As the professions change, theorizing must evolve as well. Some have begun to call for a broader, synthetic theory of occupations, one that situates the professions in a comparative way within the larger occupational structure (e.g., Abbott 1993; Freidson 1994). Central to such a broader theory is a method that sets occupations not only in their historical, comparative contexts, but also in the organizational and market context in which occupational incumbents work (e.g., Brint 1994). Freidson (1994, p. 21) has argued for moving beyond a theory that defines professions by fiat, to one that recognizes the professional aspirations of a variety of other nonprofessional occupations. In Freidson's terminology, how do ordinary occupational incumbents invoke their day-to-day work activities so as to claim professional status, regardless of whether the larger public would recognize them as professionals? Such a conceptualization opens up the possibility that other occupations can organize in their own self-interest to generate their own "professional projects." Claiming a unique expertise or specialized knowledge, developing professional associations, and convincing the state that certification is appropriate are first steps toward claiming professional status. Ultimately, the true test of professional status will be to convince others that the expertise one has entitles one to professional privilege, and to the right to establish labor market shelters. Only then will professional status and earnings rise to accompany claims of professional privilege.

## REFERENCES

- Abbott, Andrew 1988 *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- 1993 "The Sociology of Work and Occupations." *Annual Review of Sociology* 19:187–209.

- 1995 "Boundaries of Social Work or Social Work of Boundaries." *Social Service Review* 4:545–562.
- Bledstein, Burton J. 1976 *The Culture of Professionalism: The Middle Class and the Development of Higher Education in America*. New York: W. W. Norton.
- Brint, Steven 1994 *In an Age of Experts: The Changing Role of Professionals in Politics and Public Life*. Princeton: Princeton University Press.
- Carr-Saunders, A. M., and P. A. Wilson 1937 "Professions." In Edwin R. A. Seligman and Alvin Johnson, eds., *Encyclopaedia of the Social Sciences*. New York: Macmillan.
- Dingwall, Robert, and Philip Lewis 1983 *The Sociology of the Professions: Lawyers, Doctors and Others*. London: Macmillan.
- Donato, Katharine M. 1990 "Keepers of the Corporate Image: Women in Public Relations." In Barbara F. Reskin and Patricia A. Roos, eds., *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia, Pa.: Temple University Press.
- Freidson, Eliot 1984 "The Changing Nature of Professional Control." *Annual Review of Sociology* 10:1–20.
- 1986 *Professional Powers: A Study of the Institutionalization of Formal Knowledge*. Chicago: University of Chicago Press.
- 1994 *Professionalism Reborn: Theory, Prophecy, and Policy*. Chicago: University of Chicago Press.
- Halpern, Sydney 1990 "Medicalization as Professional Process: Postwar Trends in Pediatrics." *Journal of Health and Social Behavior* 31:28–42.
- Hodson, Randy, and Teresa A. Sullivan 1990 *The Social Organization of Work*. Belmont, Calif.: Wadsworth.
- Kanter, Rosabeth Moss 1977 *Men and Women of the Corporation*. New York: Harper and Row.
- Larson, Magali Sarfatti 1977 *The Rise of Professionalism: A Sociological Analysis*. Berkeley: University of California Press.
- Leicht, Kevin T., and Mary L. Fennell 1997 "The Changing Organizational Context of Professional Work." *Annual Review of Sociology* 23:215–231.
- Phipps, Polly A. 1990 "Industrial and Occupational Change in Pharmacy: Prescription for Feminization." In Barbara F. Reskin and Patricia A. Roos, eds., *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia, Pa.: Temple University Press.
- Polanyi, Michael 1967 *The Tacit Dimension*. Garden City, N.Y.: Anchor.
- Reskin, Barbara F. 1990 "Culture, Commerce, and Gender: The Feminization of Book Editing." In Barbara F. Reskin and Patricia A. Roos, eds., *Job Queues, Gender*

- Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia, Pa.: Temple University Press.
- , and Polly A. Phipps 1988 "Women in Male-Dominated Professional and Managerial Occupations." In Ann H. Stromberg and Shirley Harkess, eds., *Women Working: Theories and Facts in Perspective*. Mountain View, Calif.: Mayfield.
- Reskin, Barbara F., and Patricia A. Roos 1990 *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia, Pa.: Temple University Press.
- Ritzer, George, and David Walczak 1986 *Working. Conflict and Change*, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Roos, Patricia A., and Katharine Jones 1993 "Shifting Gender Boundaries: Women's Inroads into Academic Sociology." *Work and Occupations* 20:395–428.
- Roos, Patricia A., and Joan E. Manley 1996 "Staffing Personnel: Feminization and Change in Human Resource Management." *Sociological Focus* 29:245–261.
- Starr, Paul 1982 *The Social Transformation of American Medicine: The Rise of a Sovereign Profession and the Making of a Vast Industry*. New York: Basic.
- Tang, Joyce, and Earl Smith (eds.) 1996 *Women and Minorities in American Professions*. Albany, N.Y.: State University of New York Press.
- U.S. Bureau of Labor Statistics 1999 *Employment and Earnings*, vol. 46. Washington, D.C.: U.S. Government Printing Office.
- Wallace, Jean E. 1995 "Corporatist Control and Organizational Commitment among Professions: the Case of Lawyers Working in Law Firms." *Social Forces* 73:811–839.
- Wilensky, Harold L. 1964 "The Professionalization of Everyone?" *American Journal of Sociology* 70:137–158.
- Witz, Anne 1992 *Professions and Patriarchy*. New York: Routledge.

PATRICIA A. ROOS

## PROSTITUTION

See Sexual Behavior Patterns; Legislation of Morality; Sexual Violence and Exploitation.

## PROTEST MOVEMENTS

Protest movements have been of high interest to sociological research since the inception of the

discipline in the mid-nineteenth century, during the periods of great industrial and urban development in Europe and North America. In the context of massive changes in the economic structure and mass rural-to-urban and cross-national migration, a variety of protest movements developed. They caught the attention of Comte, Le Bon, Weber, and other early sociological analysts. In the United States, the first widely used introductory sociology textbook, developed by Chicago School sociologists Robert Park and Ernest Burgess in the 1920s, was organized around the concepts of collective behavior. Protest movements occupied a substantive part of the text.

Sociologists' interest in protest movements reflects the high interest of many who are not sociologists and are not research oriented. Such movements have the potential of affecting lives in substantial ways. This is particularly so when a protest touches on wide public concerns. In American society, the recognition of the potential impact of protest movements is encompassed within the framework of the U.S. Constitution's First Amendment guarantee of the right of the people peaceably to assemble and to petition the government for redress of grievances.

The language of the First Amendment, including the right to "peaceably" protest, is in recognition that protest movements can turn violent. Rebellion on taxes and other violent protest in the 1780s had strongly influenced the desire of those at the Constitutional Convention in 1787 to expand democratic public expression while outlawing violent means of bringing about protest movement changes. The language of the First Amendment reflects the potential power of protest ideas. Any consideration of protest movements needs to include the effects of those, like James Madison and Thomas Jefferson in the case of the amendment language, whose articulated ideas about grievances provide a key element in active protest emergence.

While most protest movements in the United States and in other democratically based societies have been mostly peaceful, there is a stream of protests which have not been peaceful. There are many examples of protest movements with no or little violence, including the multiple women's protests for more rights and opportunities throughout the nineteenth and twentieth centuries; the



food and drug protection movement in the early twentieth century; the early- and late-twentieth century environmental protest movements for cleaner air, water, and protection of endangered species and open spaces; and the protest marches of 1998 in many communities and in Washington, D.C., which resulted in new record public expenditures for cancer research at a time of budget cutbacks in most government programs. In contrast, examples of protest movements that generated periodic violence include the labor movement protests to nonresponsive corporations and legislatures from the late nineteenth century through the 1930s, the racial- and ethnic-led civil rights protest since the mid-twentieth century, and the anti-Vietnam War protest movement of the late 1960s and the early 1970s.

As not all protest movements succeed, or may succeed at a frustratingly slow pace for participants, protest activists in these protests went beyond democratically legitimate means of protest. When that occurs generally, the basis exists for violent protest episodes. Protest movements may also generate violence in opposition to a protest movement from those whose perceived vested interests and way of life are threatened. Such violence may occur to intimidate people and prevent them from engaging in protests, one clear aim of the hundreds of lynchings of black citizens in the first half of the twentieth century. Violence may occur in the context of a counterprotest movement to reverse a successful movement, as in the case of the bombings and physician killings at abortion clinics following the successful establishment of the legal right of women to seek and have an abortion.

Whether violent or nonviolent, protest movements have the potential of being an interim form of collective challenge to some aspect of the social status quo. The protest continuum ranges from localized groups and crowds that organize around specific and short-term delimited grievances to mass protest movements about social conditions and perceived injustices. These mass protests are designed to generate comprehensive and fundamental changes in a society and sometimes across societies. More so than localized acting crowds and less so than systemic social movements, protest movements encompass mass behavior that extends beyond a localized situation, and they have the potential of generating social movements

when a variety of conducive conditions exists (Gusfield 1968; Smelser 1962; Tilly 1978).

The twentieth century has been characterized by a wide variety of protest movements. In the United States, industrial protests were common for the first third of the century, as were anti-immigration protests. The suffragette protest movement early in the century was a precursor to the women's movement for equal treatment and opportunity in the last third of the century. The civil rights movement, led by blacks in the 1950s and 1960s, precipitated countermovements—a common characteristic of protest movements—including the White Citizen's Council protests and the reemergence of the Ku Klux Klan. Poor people in Chile, El Salvador, Nicaragua, and other Latin American countries have protested the privileges of an elite economic class as vestiges of an unproductive and rigid class colonial structure. Such protest movements are evident globally, with protest occurrences in Africa, the Middle East, and Asia.

A common thread through the wide variety of protest movements is their political nature. In various ways governmental authority is challenged, changed, supported, or resisted in specific protest movements. To advance their prospects for success, protest movement leaders often engage in coalition politics with more powerful individuals and groups who, for their own interests and values, support the challenge raised by the movement. When protest movements succeed in generating sufficient public support to secure all or most of their goals, governments may offer policy legitimization of the movement as a means of adapting to, coopting, or modifying a movement's challenge to the state of premovement affairs.

Such political legitimation has taken a variety of forms. The labor protest movements culminated in the passage of the National Labor Relations Act of 1935, which legitimized labor-management collective bargaining and negotiated agreements. The suffragette movement resulted in passage of the Nineteenth Amendment to the Constitution, guaranteeing that the right to vote in the United States could not be denied or abridged on account of sex. The civil rights movement attained support with passage of the comprehensive Civil Rights Act in 1964 and then the Economic Opportunity Act and the Elementary and Secondary Education Act, both in 1965. All these system-modifying acts,

which affect the lives of millions in American society, have continued in effect during relatively high and low periods of public support. This gives evidence of the long-term societal legitimation of these acts, which grew out of protest movements.

Success of these and other movements is often tempered by countermovements, participants in which perceive their relative positions and interests to be threatened. For instance, the women's movement experienced a series of challenges from religious groups, often fundamentalist, that adhered to a male-dominated patriarchy. As a consequence, women's progress was slowed in winning various forms of equal treatment and opportunity in educational, economic, political, and social areas of life, and in the 1970s and the 1980s Congress failed to pass the Equal Rights Amendment (ERA) to the U.S. Constitution, which was supported by the women's movement.

More generally, after passage of civil rights legislation in the mid-1960s, a series of protest movements within the Democratic and, more extensively, the Republican parties resulted in growing administrative, legislative, and judicial resistance to equality in educational, occupational, and housing opportunities. The countermovement result has been a reentrenchment of a long-established economic structure of racism and low-income class rigidity that functions independently of personalized racist feelings and beliefs (Wilson 1987). A reflection of such countermovement pressure is the growth in perception among white males that affirmative action educational and occupational policies directed toward racial and ethnic minorities and women constitute a form of reverse, or affirmative, discrimination (Glazer 1989).

Countermovements have generated their own counterprotest movements. This countermovement variation of Hegelian dialectic does not result in a return to whatever constituted premovement normalcy. In conventional political terms the results are more conservative, reactionary, liberal, or radical than what existed before the protest movement. When protests and counterprotests result in social change, such change generally affects the participants in a specific protest movement as well as established authorities in ways often not fully anticipated. While a predominant orientation may exist among protest activists and another among established authorities, in complex, mass modern

societies, a range of political orientations is usually contending among protestors and their supporters and among established authorities and their supporters, against whom the protest is directed.

The U.S. civil rights movement of the 1950s and 1960s can be viewed in historical terms, if not contemporary terms, as a primarily conservative movement. The predominant, although not all-encompassing, aim of activists and organizations was to enable blacks and other minorities to break into the democratically value-based, but not fully practiced, political and economic system. Most protest leaders and participants did not aim to break the established system. In contrast, the late 1980s and early 1990s liberal to radical protest movements in Poland, Hungary, Rumania, and other eastern European countries did have as their goal the breakdown of the system of exclusive authoritarian political and economic domination.

In the United States, protest ideologies are largely reminiscent of established, liberal, democratic political ideals. This is evidenced in the way many protest groups adopt language from the Declaration of Independence to advance their aims. For example, the Black Panthers, popularly perceived as a radical group, adopted a statement of purpose that held, "We hold these truths to be self-evident, that all black and white [sic] men are created equal and endowed by their creator with certain unalienable rights." Similarly, the National Organization for Women inserted into its declaration of purposes the wording that "men and women" are created equal.

Protest movements attain mixed and sometimes changed results. These results occur because of institutional inertia, in which certain things have been done certain ways over a long period of time, and because of countermovements within institutional centers such as schools; businesses; and local, state, and national government offices. In the United States, reactions to the civil rights movement have resulted in private and public attitudes and behavior that have combined to support inclusion of some minority members while disadvantaging more severely the lowest-income racial and ethnic minorities (Wilson 1987). The result is that the countermovement resistance to educational, economic, and political advances for minority status groups has adversely affected the poorest racial and ethnic minority members. At

the same time, census bureau reports document a growing number and proportion of blacks, women, and other minority group members moving into educational institutions, occupational settings, and political positions from which they were formerly excluded *de jure* or *de facto*.

Examples from history and other cultures demonstrate the mixed potential and results of protest movements. The German Nazi protest movement in the 1920s illustrated that a movement could be radical *and* reactionary, in that case toward further destabilization of the Weimar Republic's democratic government, which was perceived as being decreasingly effective and legitimate by growing sectors of the German public (Shirer 1960). After the Nazis succeeded in countering various democratic and communist protest movements, Germany saw a more comprehensive institutionalization of Nazi ideological and authoritarian control during the 1930s. An historic, more recent example is the 1989 Chinese student democratic movement in Tianenmen Square that resulted in a government-sponsored countermovement that physically shattered the student protest and created a system of political, economic, and educational controls that were more comprehensively rigid than those that existed before the protest movement. Yet, the underlying educational, economic, and political forces that generated the Chinese student activists continued to affect the dynamics of Chinese society, with the potential for further protest activation.

In these and other protest movements, there is a wide range of participants and of protest methods employed. Along with the nature of the social context and historical influences, the characteristics of protest participants and the methods they employ are consequential and have been central concerns of research on such movements.

### PROTEST PARTICIPANTS AND METHODS OF PROTEST

If protest participants could alleviate their grievances or sense of injustice individually, there would be no likely motivation for them to become active in a protest movement. Protest participants thus have two central characteristics: (1) they have insufficient influence to gain a desired change in their circumstances, and (2) they seek active association with relatively like-minded persons to gain relief from their aggrieved state.

These two characteristics can be seen among protest participants over time and in different locales. In the 1960s civil rights movements in the United States, leading activists—including blacks, Hispanics, Native Americans, and women—expressed a strong sense of unequal treatment and opportunity while associating with and supporting activists to achieve equal opportunities in schools, jobs, elected offices, and other social settings. College students, the most active participants in the civil rights movement, could not generally be characterized in these minority status terms. Yet, they were not yet an established part of the economic and political order being challenged and were in a position to be critical of that order (Lipset 1971). Other participant supporters, such as labor unions, selected corporate leaders, and religiously motivated persons, often saw protest-related change needed in terms of their own long-term interests and worked either to help the civil rights movement succeed or to preempt or coopt it (Gamson 1990, pp. 28–31). The broad political support base for the comprehensive 1964 Civil Rights Act had all these protest movement participant elements.

The individuals who are most likely to initiate and support a protest movement tend to be those with long-developed grievances within a society. A case in point is Solidarity, the labor group that precipitated the successful 1980s protest movement against communist rule in Poland and that helped precipitate other successful eastern European protest movements. The initial work stoppage, instrumental in offering a political challenge to Polish and Soviet authority, occurred at the Lenin shipyard in Gdansk, a center of Cassubian ethnic residence. For generations Cassubians have held a minority status in Polish society (Lorentz 1935). As the protest movement proceeded to secure broad-based support among Polish citizens, it was no accident that Cassubians, who have experienced prejudice and discrimination beyond communist rule in Poland, would be at the forefront. It is not surprising that Solidarity was led by a Cassub, Lech Walesa. It is also noteworthy that the protest movement received strong support from another Cassub, Pope John Paul II, whose original name of Karol Wojtyła ends with a Cassubian “a” rather than the more typical Polish “ski.”

In the United States, the civil rights movement was manifestly initiated and led by blacks. Jews,

who have experienced more prejudice and discrimination than most other whites in American society, where they constitute less than 2 percent of the population, composed the largest group of whites in the movement. In the Congress of Racial Equality (CORE), one of the leading mass civil rights protest organizations, almost one-half of the white participants identified themselves as Jewish or as secularists whose parents were Jewish (Bell 1968).

The methods employed by protest participants and leaders tend to reflect a lack of institutionalized power. When such institutionalized power is available, it can be exercised to redress grievances without resorting to mass protests. Within democratic political processes in the United States and in other democratic societies much organized protest on such issues as trade policies, road construction and placement, and taxation can be viewed in more normative, adaptive terms.

When such normative activities do not result in a resolution of grievances, the potential for a protest movement increases. In such a context, legitimized guarantees of the right to protest, as embedded in the U.S. Constitution's First Amendment guarantee of the right to assemble and petition for redress of grievances, do not preclude protest strategies that go beyond legal or normative boundaries of protest behavior.

Methods of protest are related to prospects of success and levels of frustration. When a protest movement or a countermovement has broad public support and is likely to receive a positive response from targeted authorities, protest activities are likely to be peaceful and accepted by such authorities. Such is the case with pro-choice protest on the abortion issue; protest for clean air; and protest in support of Christian, Jewish, and other minority religious status groups in the Soviet Union. All these protest activities have relatively broad American support, even when they experience a minority activist opposition.

A variety of nonlegitimate strategies are used when protest movements address issues and involve participants with relatively little initial public support and active opposition. One such nonlegitimate strategy is Ghandi's nonviolent protest confrontation with British authorities in India. Adapted by Martin Luther King, Jr., and most other black civil rights protest leaders in the 1950s and early 1960s, the strategy of nonviolence was

designed to call general public attention in a nonthreatening manner to perceived injustices experienced by blacks. With such techniques as sit-ins at racially segregated lunch counters and boycotts of segregated public buses, this nonviolent method generated conflict by breaking down established social practices. The aim of such nonviolent methods is to advance conflict resolution by negotiating a change in practices that produced the protest. A particularly famous case is the 1955 Montgomery, Alabama, bus boycott, which was one of several major precipitants of the national black-led civil rights movement.

Other, violent forms of protest include both planned strategies and unplanned spontaneous crowd action. In either case such activity tends to be perceived by authorities and their supporters as disorderly and lawless mob behavior. Masses of protest participants are likely to be drawn to violent action when the general perception, or emerging norm (Turner and Killian 1986, pp. 21-25), develops that redress of felt grievances is believed to be unattainable either in normal conditions before protest activation or by peaceful means. The history of violent protest is extensive in many societies, as exemplified by the forcible occupation of farms and fields by landless French peasants in the eighteenth century, American attacks on British possessions and military posts prior to the Declaration of Independence, and bread riots by Russian urban dwellers in World War I. The particular centuries-long history of violent protest movements in American society was documented in context of the urban race riots and anti-Vietnam War protest in the 1960s by the presidentially appointed National Commission on the Causes and Prevention of Violence in the United States (Graham and Gurr 1969).

Violent protests usually concern specific issues such as taxes, conscription into the military, and food shortages, issues that are confined to particular situations and times. Although these types of protest do not evolve into major social movements, they may have severe and immediate consequences. In 1863, for instance, during the Civil War, Irish Catholics protested what they perceived as the unfair nature of the military draft in New York City. This protest left several hundred dead. Likewise, many college students in the late 1960s and the early 1970s revolted against the draft during the unpopular Vietnam War, and the

results included loss of student lives at Kent State University.

Unplanned violence may also be a form of protest. As reported by the National Advisory Commission on Civil Disorders and other research on over two dozen urban racial riots in the 1960s in the United States, these riots, which resulted in over a hundred deaths and over \$100 million in property damage, were disorganized extensions of the black civil rights movement. These violent events closely fit Davie's J-curve thesis (1974), which argues that rising expectations, such as those generated by legal successes in mid-1960s by the black civil rights movement, were frustrated by the declining urban ghetto environments and growing Vietnam War tensions, both of which were related to the fact that large numbers of blacks were being drafted while most white college students were exempted.

Overall, protest movements are more frequent in societies that legitimize the right of protest. In such societies, social conflict generated by protest movements is often functional in resolving conflict over issues between challenging and target groups (Coser 1956). Still, urban and campus riots of the 1960s illustrate that formal rights of protest do not deter democratic authorities and their public supporters from responding with police force or from beginning a countermovement. Authoritarian societies may experience fewer protest movements, but when they do occur, such movements are far more likely to be intense and to have the potential for massive social movements designed to transform the society. This could be seen in widely disparate societies, including most eastern European nations and the Soviet Union, El Salvador, Nicaragua, South Africa, Iran, and mainland China.

### CONSEQUENCES OF PROTEST MOVEMENTS

Given the long and continuing history of protest movements, there has been growing interest in the long-term consequences of such movements. Some assessments concentrate on historical, comparative analysis such as Snyder and Tilly's analysis (1972) of French collective violence in response to government-sponsored repression between 1830 and 1968, or Bohstedt and Williams's analysis

(1988) of the diffusion of riotous protests in Devonshire, England, between 1766 and 1801. Other studies of long-term consequences make empirical assessments of the aftermath of more contemporary protest movements. Examples include Gordon's community-based assessment (1983) of black and white leadership accommodations in the decade following the Detroit race riots of 1967 and Morris's assessment (1980) of the decade-long impact on national public values of the environmental movement of the late 1960s.

The need for more short- and long-term assessment of the consequences of protest movements is evident in reviews of past movements. William Gamson's consideration of fifty-three protest movements in the United States between the 1830s and 1930s illustrates the need. Gamson categorized each movement's own specific goals in one of four ways: coopted, preempted, full response success, and collapsed failure (Gamson 1990, pp. 145-453). Gamson assessed each protest movement's success in achieving its goals during its own period of organized activity. For instance, Gamson assessed such groups as the German American Fund (1936-1943), the American Proportional Representation League (1893-1932), and the Dairyman's League (1907-1920).

Of the fifty-three identified protest movements, twenty-two, the largest single proportion, were categorized as being collapsed failures, twenty as achieving full response success, six as being preempted, and five as being coopted. Protest movements categorized as collapsed failures and full response successes demonstrate the need for assessment of protest movements long beyond their activist periods. Listed under collapsed failures were major long-term successful movements including the abolitionist North Carolina Manumission Society (1816-1834) and the American Anti-Slavery Society (1833-1840). In contrast, among full response success movements was the American Committee for the Outlawry of War (1921-1929), a major force in the achievement of the international Kellogg-Briand Pact of 1928, which outlawed war between nations, a short-lived success that for at least most of the rest of the twentieth century proved a grand failure.

Successful or unsuccessful in the short or the long term, protest movements are periodically a part of social change at local, national, and global

levels and in situational, institutional, and cross-cultural contexts. In the United States and other modern mass urban societies, the long-term trend has been for protest movements to become more professionalized with increased mobilization of resources to more effectively challenge entrenched interests (Tilly 1978). Modern communication systems, global economic interdependence, and economical movement of masses of people over great distances assures that protest movements of the future will increasingly be characterized by a combination of ideas, people, and organization across all these areas of social life.

In contemporary terms, the historical end of the global Cold War at the end of the twentieth century with the collapse of the Soviet Union generated a basis for increased localized protest movements. This was initially evident in the former Soviet Union's sphere of influence in eastern European societies and within Russia itself. On the eve of the twenty-first century, evidence of this broadening localized protest activism was to be seen in many societies as exemplified by student-led protests in Indonesia, Chiapas Indian revitalization protests in Mexico, the Quebec separatist movement in Canada, the Queens' Borough protest to separate from New York City in the United States, and many others. In this context, the supply of localized protest issues is likely to proliferate. At the same time, some issues—such as protests associated with such transnational issues involving the environment and health—may touch upon many localized issues that coalesce into becoming national and global protest activities. Given massive economic, political, and social change forces in modern society and the potential for protest movements to directly affect the lives of many people, the concern of research specialists and publics with such movements can be expected to continue and to increase.

(SEE ALSO: *Segregation and Desegregation; Social Movements; Student Movements*)

REFERENCES

Bell, Inge Powell 1968 *CORE and the Strategy of Non-Violence*. New York: Random House.

Bohstedt, John, and Dale E. Williams 1988 "The Diffusion of Riots: The Patterns of 1966, 1795, and 1801 in

Devonshire." *Journal of Interdisciplinary History* 19(1):1-24.

Coser, Lewis 1956 *The Functions of Social Conflict*. New York: Free Press.

Davies, James C. 1974 "The J-Curve and Power Struggle Theories of Collective Violence." *American Sociological Review* 39:607-612.

Fogelson, Robert 1971 *Violence as Protest: A Study of Riots*. New York: Anchor.

Gamson, William A. 1990 *The Strategy of Social Protest*. Homewood, Ill.: Dorsey.

Glazer, Nathan 1989 *Affirmative Discrimination*. Cambridge, Mass.: Harvard University Press.

Gordon, Leonard 1983 "Aftermath of a Race Riot: The Emergent Norm Process among Black and White Community Leaders." *Sociological Perspectives* 26:115-135.

Graham, Hugh D., and Ted Gurr (eds.) 1969 *The History of Violence in America: Report of the National Commission on the Causes and Preventions of Violence*. New York: Bantam.

Lipset, Seymour M. 1971 *Rebellion in the University*. Boston: Little, Brown.

Lorentz, Friedrich 1935 *The Cassubian Civilization*. London: Faber and Faber.

Morris, Aldon D. 1984 *The Origins of the Civil Rights Movement: Black Communities Organizing for Change*. New York: Free Press.

Morris, Denton 1980 "The Soft Cutting Edge of Environmentalism: Why and How the Appropriate Technology Notion Is Changing the Movement." *Natural Resources Journal* 20:275-298.

Mueller, Carol, and Charles Judd 1981 "Belief Consequences and Belief Constraint." *Social Forces* 60:182-187.

Shirer, William 1960 *The Rise and Fall of the Third Reich*. New York: Simon and Schuster.

Smelser, Neil J. 1962 *Theory of Collective Behavior*. New York: Free Press.

Snyder, David, and Charles Tilly 1972 "Hardship and Collective Violence in France: 1830 to 1960." *American Sociological Review* 37:520-532.

Tilly, Charles 1978 *From Mobilization to Rebellion*. New York: McGraw-Hill.

Turner, Ralph, and Lewis M. Killian 1986 *Collective Behavior*. Englewood Cliffs, N.J.: Prentice-Hall.

Wilson, William J. 1987 *The Truly Disadvantaged: The Inner City, The Underclass, and Public Policy*. Chicago: University of Chicago Press.

## RECOMMENDED READINGS

Gamson, William A. 1990 *The Strategy of Social Protest*. Homewood, Ill.: Dorsey.

Tilly, Charles 1978 *From Mobilization to Rebellion*. New York: McGraw Hill.

LEONARD GORDON

## PUBLIC OPINION

Public opinion is characterized, on the one hand, by its form as elementary collective behavior (Blumer 1972) and, on the other, by its functions as a means of social control (Ross 1901). It comes into play in situations that are problematical or normatively ambiguous in one or more of several senses: The situation is novel and unprecedented, so that established ways of coping no longer prove adequate, people actively disagree over which of several conventionally acceptable practices should apply, or the conventions themselves have come under serious challenge by a dissident group. In the extreme case, controversy over what should be done can heat up to a point where order gives way to violent group conflict.

Interest in public opinion is historically linked to the rise of popular government. Although rulers have always had to display some minimum sensitivity to the needs and demands of their subjects, they felt little need, unlike most contemporary governments that must face voters in mandated elections, to anticipate their constituents' reactions to events that were yet to occur or to policies still to be implemented. But public opinion operates equally outside the relationship between citizens and the state. Its influence is felt throughout civil society, where, on many matters, including personal taste in dress, music, and house furnishings, people remain sensitive to the changing opinions of peers and neighbors. They court approval by showing themselves in step with the times.

Opinions have behind them neither the sanctity of tradition nor the sanctions of law. They derive their force from agreement, which is more tenuous than either of these. People do change their minds. Moreover, to label something as "opinion" implies a certain willingness to acknowledge the

potential validity of contrary views. Issues are settled by discussion and bargaining. Those who pay at least some attention and are ready to take sides make up a public. It expands in size as an issue heats up, only to contract again as the focus shifts to new problems. There are, in fact, as many publics as there are issues.

Social control through public opinion functions in subtle ways. All but the most intransigent partisans will recognize when an issue is settled to a point where further debate becomes not only superfluous but possibly disruptive of an underlying consensus. There are times when the need to display unity, perhaps in coping with a crisis, or moral fervor, whipped up in crusades against internal enemies, may unduly narrow the range of public discourse. Dissidents come under pressure to conform, at least outwardly, but not as completely as in the regimes that suppress public opinion by seeking control over all conversational channels through which ideologically deviant tendencies could spread.

Studies of public opinion have to contend with a broad range of beliefs. Located at one extreme are beliefs anchored in longstanding allegiances; at the other, the often fluctuating "gut" responses to whatever is current. The more general ideas that underlie the legitimacy of the political system have the greatest stability. They build on childhood experiences within the family, where children tend to adopt the views of their parents. Then, as the children become adults, these early beliefs are elaborated and modified in sustained contact with other major institutions, like school and church, and also (especially during major catastrophes affecting their country or its leaders) by the news media (Renshaw 1977). The content of the political culture from which opinions are derived differs from milieu to milieu.

Endemic cleavages related to position in social structure and in historical time are highlighted during controversies. Hence, public opinion often divides in predictable ways—by region, race, religion, ethnicity, class background, educational attainment, and so forth. If the differences in experience are sharp enough and each side raises nondebatable demands, public discourse can escalate to a point where the polity splits apart into two irreconcilable camps. Instead of reaching agreement, the more powerful group imposes its will.

Acquiescence does not qualify as rule by public opinion.

In the center of early sociological study of public opinion has been the question of competence. Analysts have sought to distinguish conceptually between the reasoned opinions developed in discussion and the clamor of a feared “mob” acting under the sway of emotion. Two works, one in Germany and one in America, coincidentally published the same year, analyzed the problem in structural terms.

Tönnies (1922) pointed to the press and to associations that usurped for themselves the role of articulating public opinion. He would have been even more critical of the public relations industry that has flourished since. For Lippmann (1922) there was a still more fundamental obstacle. He rejected as a false ideal the notion that ordinary citizens—even the most well-educated—had the time and incentive to acquire the expertise to grasp the complex problems of the day in the detail necessary to direct the course of public policy. Drawing on a wide range of literature, he showed how the public perceived the world through the stereotypes fed them by the press. This meant, he argued, that whenever the public attempted to intervene in the course of events, it inevitably did so as the dupe or unconscious ally of elite interests. Its role was properly limited to identifying the problem areas in need of remedial action and to deciding which party, institution, or agency should be entrusted with the solution. But the public was a potentially effective “reserve force” most effectively mobilized in support of the procedural norms of democracy. Contemporary contractualists, like Buchanan and Congleton (1998), have come to share Lippmann’s emphasis on generalized procedures about which near unanimity is more easily achieved than about policies that inevitably produce winners and losers.

The list of social scientists pointing from different perspectives to the limits of “rule by public opinion” under modern conditions includes Mannheim (1940), Schumpeter (1942), Schattschneider (1960), Bogart (1972), and Ginsberg (1986). An all-too-obvious gap between the expectation of an informed citizenry put forward by democratic theory and the discomfiting reality revealed by systematic survey interviewing is identified by Neuman (1986) as the “paradox of mass

politics.” Where pluralists see a public made up of many competing interests, each with its own leadership, Neuman discriminates among three levels of competence: an uncomfortably tiny percentage of citizens with some input into policy; uninterested and inactive know-nothings, who make up roughly one-fifth of the potential electorate; and a large middle mass whose members, if they vote, do so largely out of a sense of duty but with only a very limited understanding of the issues their vote is meant to decide. This discrepancy from the ideal, apparent in many polls on specific issues, can hardly be attributed to flaws in the method by which public opinion is ascertained.

### PUBLIC OPINION POLLING

There are nevertheless some real questions about the momentary numerical majorities obtained in an opinion poll as a valid measure of public opinion. For one thing, polls vary in quality. How closely any particular poll reflects the distribution of opinion within a larger population hinges on three general factors: (1) who is interviewed, (2) the situation in which the interview takes place, and (3) the questions asked. Insofar as elections also provide a public record of “opinion,” the utility of polls as a research tool can be tested against the actual vote count.

Two fiascos in polling history have been painstakingly diagnosed. Never again will we see anything like the wildly incorrect 1936 forecast by the *Literary Digest* that Franklin D. Roosevelt, who won reelection by a landslide, would be voted out of office. It was based on 2.3 million returns from over 10 million straw ballots mailed to names on automobile registration lists and in telephone books. Poll takers learned the hard way that, especially when respondents are self-selected, sheer numbers do not guarantee accuracy. For one thing, the well-to-do, who owned cars and lived in homes with telephones, were somewhat less supportive of Roosevelt and his New Deal than those too poor to have either or both such conveniences. A second, actually more important, source of bias was self-selection; Republicans were more strongly motivated than Democrats to mail in their straw ballot as a protest against the party in power.

When in 1948 the pollsters, despite more rigorous sampling and face-to-face interviews,



wrongly predicted the defeat of incumbent President Harry Truman, this became the occasion for one of the most extensive inquiries into polling practices by a committee of the Social Science Research Council (Mosteller et al. 1949). Its report stressed the importance of random selection, in which every voter stands the same chance of being contacted, an objective often difficult to implement. Polling techniques have come a long way since and errors of such magnitude have not reoccurred. But, as shown by an investigation by Crespi (1988) of the factors associated with accurate prediction in 430 preelection polls during the 1980s, the extra effort invested in callbacks still pays off in greater accuracy. Persons missed because they are hard to reach or because they refuse to answer often differ from the rest in ways difficult to estimate.

As to the interview situation, answering the questions of a poll taker is hardly the same as casting a vote, all the more so when the election is months away. The large margins by which Truman had been trailing in the early fall of 1948 caused several pollsters to cease polling weeks before voting day. Thus, they never registered the strong Democratic rally taking place toward the end of the campaign. Another problem is determining how firm respondents are in their convictions and who will actually vote on election day. Despite refined techniques to ferret out likely nonvoters, predicting the outcome of a low-turnout election with little-known candidates and in referendums on questions beyond the understanding of many voters can be hazardous.

Public opinion research covers much more than elections, where the "issue" boils down to a clear split between parties, candidates, or those for or against a measure on the ballot. Issues come and go, and people do not necessarily have preconstructed views on everything about which polls may ask but tend to construct their answers in an ad hoc manner from information that, at the moment, strikes them as salient (Zaller 1992). Just as pollster have learned to omit from their tabulations of preelection surveys the likely nonvoters, so various "filters" are used to screen out persons who are generally unconcerned about politics, have given the particular subject little thought, and may not have even been aware of it except for the questions put to them. To avoid eliciting

"nonattitudes," as responses from such persons are called, they should be asked no further questions about the matter. This still leaves those reluctant to admit their unfamiliarity or lack of concern. One survey that deliberately inserted a question about a nonexistent bill allegedly under consideration by Congress had significant minorities respond that they had heard of it, with some even willing to fabricate an opinion about it.

Insofar as responses depend on how a question is phrased, different polls can provide dramatically different readings of where the public stands. Various polls on impeachment during Watergate, all about the same time, recorded levels of support ranging from a high of 53 percent on a question with the condition "*if it were* "decided that President Nixon was involved in the coverup" attached to a low of 10 percent on a question that offered resignation as an alternative to impeachment. Similarly, after Gallup modified its original question, which had coupled "impeachment *and* removal from office," by first explaining impeachment and then asking, "Given the various charges brought against the president, do you think impeachment charges should be brought against him or not?" support for impeachment jumped from 37 to 53 percent.

Views on many subjects are too nuanced to be caught in response to a single question. Few people are absolutists on most issues. On abortion, a highly contentious issue, it is not just being "pro choice" or "pro life" in all situations. One can deny that abortion is a constitutional right and still allow it under certain circumstances—when conception has resulted from rape or when abortion is necessary to save the life of the mother. Some legal restrictions are likewise acceptable for supporters of *Roe v. Wade*. Similar distinctions are called for in polling on "affirmative action," a term that stands for a welter of different policies. Explicit preferences, even as a redress for past discrimination, are endorsed by no more than a minority, but outreach programs to locate qualified minorities and/or women enjoy wide support. To correctly assess what is on people's mind requires a series of probing questions.

Nor do apparent majorities always speak as clearly as the student of public opinion would like. In a 1971 poll, a time when concern over American

involvement in Vietnam was higher than any other issue on the public agenda, two out of three respondents answered the question, "Do you favor or oppose the withdrawal of all American troops from Vietnam by the end of the year?" by declaring themselves in favor of withdrawal. But on another question in the same survey, about withdrawing all troops "regardless of how the war was going," these same respondents split with a plurality of 44 percent against and only 41 percent still for withdrawal. Does this 25-point difference between the 66 percent in favor of withdrawing on the first question and the 41 percent on the second identify a group ready to take back an off-the-cuff answer when reminded of the possible consequences of such a move? Or were those giving a "consistent" response confident that South Vietnam would not fall? There is still a third possibility: A lot of people no longer cared whether or not America was forced to withdraw in defeat. An undetermined number among them may, in fact, have welcomed such an outcome.

Sometimes the public does indeed hold views that seem contradictory, but the logic people follow may differ from that of the analyst, as it did during Vietnam, when polls were recording, at one and the same time, majorities in support of both immediate withdrawal and a stepped-up air war on the North. Others who considered American intervention in Vietnam a mistake still spoke out in favor of President Lyndon Johnson's Vietnam policy. Nor, for that matter, were self-styled "doves" necessarily in sympathy with student protests against the war.

For all their potential pitfalls, polls are one of the best ways for political leaders to maintain contact with their increasingly large and diverse constituencies—a more reliable reading for sure than such alternative indicators of public opinion as letters, telegrams, phone calls, and petitions to political leaders, which may reflect nothing more than the effort of a well-organized interest group. Yet even minorities, if implacable enough, should not always be ignored. A rise in the number who refuse draft calls or desert from the armed forces; a rise in certain crimes; and an increase in demonstrations, strikes, and other forms of protest are important clues, not necessarily to general opinion but at least to the unwillingness of the groups most affected by some problem to settle for the

status quo (Tilly 1978). They signal problems or injustice that call for recognition by the rest of society.

### DYNAMICS OF PUBLIC OPINION

Changes in the basic attitudes that underlie public opinion usually occur slowly, partly through replacement and partly through the diffusion of new experience. Differential birth rates, migration, social mobility, and the succession of generations are processes that disturb the existing balance without any change on the individual level and despite evidence of significant political continuity between parents and offspring. Distinct intergenerational differences are believed to develop in response to certain critical experiences, like the encounter, in early adulthood before one's outlook has fully crystallized, with general poverty and war, or participation in social struggles (Sigel 1989).

Attitudes on race are a good example of how diffusion and replacement operate in conjunction with each other. Surveys taken over time show a distinct movement toward greater racial tolerance (Schuman and Bobo 1985). All groups moved in the same direction, even if at different speeds. The younger generation showed the way, often with tacit support from sympathetic parents not yet ready themselves to join a radical challenge to segregation. Rising levels of education and the increase in certain kinds of intergroup contact also contributed to the shift, as did mortality among the older, more conservative cohorts. Especially noteworthy is that, in regions where segregationist practices were most firmly entrenched, opinion changed more rapidly than in the rest of the country, once the full force of national public opinion had been brought to bear on them through television. The pervasive coverage the national media gave the campaign for civil rights conveyed to even the most intransigent southerners how the rest of the country viewed their continuing resistance. While issues relating to race continue to divide the polity, the terms in which they are debated were never to be the same again.

The day-to-day shifts in public opinion on matters great and small are even more subject to influence by the media of mass communication. Most of the public is not issue-oriented but reacts

to the general image of presidential performance. "Good" news of any kind tends to bolster that image, with one major exception: a national crisis. Such events typically generate a rally to the flag; critical voices are stilled, at least temporarily, in a show of patriotic unity. Effective political leadership under any condition requires access to the news media.

In their comprehensive review of American policy preferences over a half-century, Page and Shapiro (1992) show that, on thirty-two foreign and forty-eight domestic issues for which there were adequate data, the media coverage explained a large part of the movement of opinion in polls. However, they go on to note that opinions on none of these issues changed very much. The main influence of heightened media coverage, as repeatedly documented by communication research, is to move an issue or a problem into the public sphere, followed by an increase of concern and discussion, which then puts pressure on government to do something. When the press plays up crime, this helps create the impression of a crime wave, just as vivid details about a disaster drive home its magnitude. Serious discussion about the nature, causes, and consequence in the media of what came to be called "child abuse" facilitated its recognition as a social problem calling for intervention (Nelson 1984).

Once there is concern, the issue itself undergoes change. The actions of the principal actors involved in Watergate, as reported by the media, moved the debate away from Nixon's involvement in the illegal break-in into Democratic headquarters to his obstruction of justice, abuse of power, and contempt of Congress—the three counts on which the House Judiciary Committee voted for impeachment (Lang and Lang 1983). A similar kind of shift occurred in the "sex" scandal surrounding Bill Clinton. What began as the exposé of an improper affair with a White House intern was progressively redefined into whether the president had perjured himself and whether this felony, if committed by a president in connection with a civil suit later found without merit, rose to the level of an impeachable offense.

Media power is nevertheless limited. First of all, opinions, once they have crystalized, respond more to actual changes in circumstances than to

media messages. Second, the potential influence of the media varies according what they report about. It is obviously greater when they expose widespread corruption in high places, report on a mishap in foreign relations, or highlight something that hardly anyone would ever know about unless alerted by the news media. Far less dependent on media recognition are inflation, shortages, a severe economic downturn, and other events. Neither they nor specific grievances anchored in shared group experience will go away for mere lack of mention. Third, media managers are less than fully independent. They have to accommodate other actors intent on publicizing only those issues (or aspects of issues) that work in their favor. While competition among the various practitioners of the highly developed art of news management introduce some balance, public discussion, and (indirectly) public opinion do respond to media strategies not so much aimed at persuasion as bent upon seizing the right issues.

#### PUBLIC OPINION AND POLICY

As to the effects of opinion change, leaders of major institutions, whether elected or not, have proved distinctly sensitive to trends in public opinion. The measure most consistently repeated over the most years is the presidential approval ratings. Most presidents have experienced a gradual slide during their terms in office, which Mueller (1973) attributes to a coalition of all the minorities that, over the years, will have been antagonized by the many decisions a president as the chief executive is forced to make. At least two American presidents have been driven from office by clear evidence of a loss of public support. Lyndon Johnson, with the failure of his Vietnam policy glaringly evident to all, took himself out of the race for reelection, even though as an incumbent president he would have been assured renomination as the standard bearer of his party (Schandler 1977). Public outrage during Watergate forced Richard Nixon to make several concessions and, ultimately, to resign after the Supreme Court forced the release of tapes with the incriminating evidence that made his impeachment and subsequent removal from office a near certainty (Lang and Lang 1983). Twenty-five years later, a president's high approval, despite his publicly acknowledged wrongdoing, caused Republicans in both houses of Congress, who wanted to

hold him responsible, to frame the issue in the most narrow legal terms lest their moves backfire.

Congress, knowing that public opinion responds to events, sometimes even to the turn of a debate, often moves cautiously on potentially divisive issues. When Franklin Roosevelt tried to pack the Supreme Court and Nixon resisted the full exposure of Watergate, legislators were inclined to wait, watching to see which way the public tilted. And when it came to civil rights, major legislation was passed only after mass demonstration and media attention to discriminatory practices had created public concern and support for the principle had reached or exceeded the two-thirds mark. Laws subsequent to the initial pathbreaking legislation could then be enacted without direct pressure from below. But none of these things could have been achieved without effective political leadership responsive to the just demands of an obvious minority, which Johnson, still enjoying the honeymoon bequeathed an incoming president, was able to provide.

The accumulating evidence points to a basic congruence between public opinion on basic issues and legislative action (cf. Monroe 1998; Page and Shapiro 1983). Even the decisions of a judiciary with lifelong tenure are not entirely insulated from the political cross-currents that affect representative institutions. The U.S. Supreme Court, the most august of judicial bodies, as Marshall (1989) concludes, has been an essentially majoritarian institution. Of 142 decisions from the mid-1930s to the mid-1980s for which there existed corresponding opinion data, over four-fifths have been consistent with preferences expressed in polls. The linkage, strongest in times of crisis, reflects more the court's sensitivity to legislative and executive concerns that incorporate public opinion than direct to public pressure on the judges.

It is on the constitutional rights of dissident minorities that the Supreme Court has most consistently set itself against majority opinion. Other countermajority opinions have either articulated a rising trend, strengthened by the voice of the court, or been modified by later decisions that, in an apparent response to public opinion, carved out exceptions and introduced qualifications to the broad rule laid down in the original case. Public opinion continues to play a role in how

these court decisions are implemented. Following a decade of increasingly liberal attitudes toward abortion, the issue seemed settled once and for all by *Roe v. Wade*. Though the massive campaign proliferators have mounted since has had little impact nationally, states vary in their laws and practices. The states in which opposition to abortion has been strong have generally adopted more restrictive policies, resulting in lower abortion rates, than states in which the weight of public opinion was to make it a matter of choice (Wetstein 1996).

### THE ERA OF POLLS

Elites have always been under some constraints, even in dealing with the more obscure issues typically resolved through specialized networks. Sheer prudence nevertheless commands that appearances be managed. This is done today, more than ever before, on the basis of increasingly accurate information from polls. Experts on polling have become members of the advisory staff not only of candidates for major offices but also of national party organizations. They also work in the White House and are consulted by presidents about proprietary polls which, unlike those sponsored by the media, especially during election season, the public rarely sees. There is little persuasive evidence so far about how the current high volume of polling activity has affected the political process.

One perhaps not very obvious consequence is the dissemination of a view on public opinion as statistics on a set of questions instead of a process to determine of what people will ultimately settle for. The public, too, is somewhat suspicious of polls. A large proportion of those queried on the subject believe that polls have effects and, particularly, that the majorities they record generate bandwagons even though these same people vehemently deny that the polls would have any effect on themselves. Nor has the variety of split-ballot polls and experiments, in which only one of two matched samples is informed about of the majority viewpoint, turned up differences large enough to support what many apparently fear. Noelle-Neumann (1984), in a more elaborate formulation, has shifted the emphasis away from direct bandwagons (i.e., people rushing to join the majority) to more complex sequences of events. She reasons that

people who perceive themselves as being in the minority, even if falsely, are reluctant to speak out against the dominant opinion, thereby creating a spiral of silence through which the apparent majority gains extra strength. Such spirals have no doubt occurred, for example during the repressive era remembered as McCarthyism. But to generalize from this and other instances is to overlook situations in which committed minorities have raised their voices to compensate for what they lack in number. Polls do function as a corrective by reporting on opinions that, although underrepresented in the forums which most public discourse takes place, are too important to ignore.

Polls have other significant consequences insofar as those who rely on them react strategically. They certainly influence the choice of candidates, the platform and program on which a candidate or party stands for election, and the policies an administration pursues. Candidates who can demonstrate electability attract the financial support, endorsements, and media coverage necessary for an effective campaign. Citizens, too, vote strategically whenever they abandon a likely loser in favor of a minimally acceptable second choice. This is what gives a front-runner in the polls the momentum to increase his or her distance from the pack of competitors early in the primary season (Bartels 1988). A strong showing can also have a potential downside: A victory by less than the anticipated margin, even in a small state with low turnout and therefore difficult to predict, has sometimes been read as a defeat. Hence, candidates often work to lower expectations regardless of what their own polls may show.

Polls on the concerns and issue preferences of the electorate allow campaign managers to maximize their candidates' appeal. Political actors who act rationally, rather than ideologically, will moderate the more doctrinaire positions cherished by their loyal followers, who have no other place to go. According to the spatial model of politics first outlined by Downs (1957), parties and candidates are impelled to move toward the political center to reach out to the still uncommitted. More and more do the modern catch-all parties, competing for an ever-larger share of the electorate, rely on general promises and finely honed advertisements. They differentiate themselves more by the image the party seeks to project than by any clearly defined policies. Practices of this sort are bound to

feed voter cynicism about the political system. One would also expect adverse effect on traditional party loyalties.

Republican leaders who heralded their success in 1994, an off-year election in which they took over both houses of Congress, as an endorsement of the party's Contract with America, quickly learned about the danger of being taken in by their own campaign slogans. At least their managers must have known that, as late as two weeks before the election, 65 percent of the public polled by the Gallup Organization had not yet even heard of the contract and that, according to another poll just days before the election, only 7 percent of those interviewed said they would be more likely to vote for the Republican candidate for Congress if he or she supported the contract. Acting as if they had a clear mandate, the Republicans in Congress made this their legislative agenda and, when embroiled in a budget dispute with the Democratic president, forced three closedowns of government. President Clinton's approval ratings rose sharply, and Republicans, who had misread the election results, lost credibility. By the same token, Democratic gains in 1998 plus polls showing clear majorities opposed to impeachment apparently misled the Clinton White House to underestimate the determination of House Republicans to bring down the president. Electoral results are not carbon copies of the public mood.

There is no question that elites are now in a better position than ever before to anticipate the public reaction to whatever they may do. Accurate information is all the more vital when the loosening of ideological bonds has forced an elected leader to engage in a continuous campaign for popularity. Such a leader has good reason to adopt the majority point of view on issues about which the public feels strongly and to save his political capital for other less salient issues, on which views have not yet crystalized. This is where Geer (1996) locates the opportunities for effective leadership. Informed by polls, a president now is in a better position to identify these issues than were presidents in previous eras, several of whom exercised what Geer calls "leadership by mistake." But although a democratic leader must accede to at least some of the wishes of his followers, failure to lead when possible would amount to more than a missed opportunity; it would be a serious mistake, because it cedes the territory to the opposition,

which presumably has access to the same information and no hesitation to exploit it.

Still, what counts in the long run are the results. A leadership too sensitive to what a majority prefers at a given moment may sidestep a problem that cries for remedies the public is not yet prepared to approve. The absence of serious policy debates becomes an open invitation to topple a leader with personal attacks on his character and by grasping at every sign of a possible scandal. Other complications stem from the fact that governing majorities are coalitions of sometimes rather diverse interests, many of them organized and in possession of expertise beyond the comprehension of most citizens. Things of overriding importance to one interest may be anathema to all the others. The voice of the people comes across most clearly when the multitude is truly stirred by an all too apparent failure of policy or misbehavior in high places. It is on these occasions that the power of the public as a "reserve force," wisely or otherwise, is felt most directly.

## REFERENCES

- Bartels, Larry M. 1988 *Presidential Primaries and the Dynamics of Public Choice*. Princeton, N.J.: Princeton University Press.
- Blumer, Herbert 1972 "Outline of Collective Behavior." In Robert R. Evans, ed., *Readings in Collective Behavior*. Chicago: Rand McNally.
- Bogart, Leo 1972 *Polls and the Awareness of Public Opinion*. New York: John Wiley.
- Buchanan, James M., and Roger D. Congleton 1998 *Politics by Principle: Toward Nondiscriminatory Democracy*. New York: Cambridge University Press.
- Crespi, Irving 1988 *Pre-Election Polling: Sources of Accuracy and Error*. New York: Russell Sage.
- Downs, Anthony 1957 *An Economic Theory of Democracy*. New York: Harper.
- Ginsberg, Benjamin 1986 *The Captive Public: How Mass Opinion Promotes State Power*. New York: Basic Books.
- Lang, Gladys Engel, and Kurt Lang 1983 *The Battle for Public Opinion: The President, the Press, and the Polls During Watergate*. New York: Columbia University Press.
- Lippmann, Walter 1922 *Public Opinion*. New York: Macmillan.
- Mannheim, Karl 1940 *Man and Society in an Age of Reconstruction*. New York: Harcourt, Brace.
- Marshall, Thomas 1989 *Public Opinion and the Supreme Court*. Boston: Unwin Hyman.
- Monroe, Alan D. 1998 "Public Opinion and Public Policy, 1980-1993." *Public Opinion Quarterly* 62:6-28.
- Mosteller, Frederick, Herbert Hyman, Philip J. McCarthy, Eli S. Marks, and David B. Truman 1949 *The Preelection Polls of 1948*, Bulletin 68. New York: Social Science Research Council.
- Mueller, John E. 1973 *War, Presidents, and Public Opinion*. New York: John Wiley.
- Nelson, Barbara 1984 *Making an Issue of Child Abuse: Agenda Setting for Social Problems*. Chicago: University of Chicago Press.
- Neuman, W. Russell 1986 *The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate*. Cambridge, Mass.: Harvard University Press.
- Noelle-Neumann, Elisabeth 1984 *The Spiral of Silence*. Chicago: University of Chicago Press.
- Page, Benjamin I., and Robert Y. Shapiro 1983 "The Effects of Public Opinion on Policy." *American Political Science Review* 77:175-190.
- 1992 *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.
- Renshaw, Stanley A. (ed.) 1977 *A Handbook of Political Socialization*. New York: Free Press.
- Ross, Edward A. 1901 *Social Control: A Survey of the Social Foundations of Order*. New York: Macmillan.
- Schandler, Herbert J. 1977 *The Unmaking of a President: Lyndon B. Johnson and Vietnam*. Princeton N. J.: Princeton University Press.
- Schattschneider, E. E. 1960 *The Semisovereign People: A Realist's View of Democracy in America*. Hinsdale, Ill.: Dryden Press.
- Schuman, Howard, and Lawrence Bobo 1985 *Racial Attitudes: Trends and Interpretations*. Cambridge, Mass.: Harvard University Press.
- Schumpeter, Joseph A. 1942 *Capitalism, Socialism, and Democracy*. New York: Harper and Row.
- Sigel, Roberta S. (ed.) 1989 *Political Learning in Adulthood: A Sourcebook of Theory and Research*. Chicago: University of Chicago Press.
- Tilly, Charles 1978 *From Mobilization to Revolution*. New York: Random House.
- Tönnies, Ferdinand 1922 *Kritik der öffentlichen Meinung*. Berlin: Springer.
- Wetstein, Matthew E. 1996 *Abortion Rates in the United States: The Influence of Opinion and Policy*. Albany: State University of New York Press.

which presumably has access to the same information and no hesitation to exploit it.

Still, what counts in the long run are the results. A leadership too sensitive to what a majority prefers at a given moment may sidestep a problem that cries for remedies the public is not yet prepared to approve. The absence of serious policy debates becomes an open invitation to topple a leader with personal attacks on his character and by grasping at every sign of a possible scandal. Other complications stem from the fact that governing majorities are coalitions of sometimes rather diverse interests, many of them organized and in possession of expertise beyond the comprehension of most citizens. Things of overriding importance to one interest may be anathema to all the others. The voice of the people comes across most clearly when the multitude is truly stirred by an all too apparent failure of policy or misbehavior in high places. It is on these occasions that the power of the public as a "reserve force," wisely or otherwise, is felt most directly.

## REFERENCES

- Bartels, Larry M. 1988 *Presidential Primaries and the Dynamics of Public Choice*. Princeton, N.J.: Princeton University Press.
- Blumer, Herbert 1972 "Outline of Collective Behavior." In Robert R. Evans, ed., *Readings in Collective Behavior*. Chicago: Rand McNally.
- Bogart, Leo 1972 *Polls and the Awareness of Public Opinion*. New York: John Wiley.
- Buchanan, James M., and Roger D. Congleton 1998 *Politics by Principle: Toward Nondiscriminatory Democracy*. New York: Cambridge University Press.
- Crespi, Irving 1988 *Pre-Election Polling: Sources of Accuracy and Error*. New York: Russell Sage.
- Downs, Anthony 1957 *An Economic Theory of Democracy*. New York: Harper.
- Ginsberg, Benjamin 1986 *The Captive Public: How Mass Opinion Promotes State Power*. New York: Basic Books.
- Lang, Gladys Engel, and Kurt Lang 1983 *The Battle for Public Opinion: The President, the Press, and the Polls During Watergate*. New York: Columbia University Press.
- Lippmann, Walter 1922 *Public Opinion*. New York: Macmillan.
- Mannheim, Karl 1940 *Man and Society in an Age of Reconstruction*. New York: Harcourt, Brace.
- Marshall, Thomas 1989 *Public Opinion and the Supreme Court*. Boston: Unwin Hyman.
- Monroe, Alan D. 1998 "Public Opinion and Public Policy, 1980–1993." *Public Opinion Quarterly* 62:6–28.
- Mosteller, Frederick, Herbert Hyman, Philip J. McCarthy, Eli S. Marks, and David B. Truman 1949 *The Preelection Polls of 1948*, Bulletin 68. New York: Social Science Research Council.
- Mueller, John E. 1973 *War, Presidents, and Public Opinion*. New York: John Wiley.
- Nelson, Barbara 1984 *Making an Issue of Child Abuse: Agenda Setting for Social Problems*. Chicago: University of Chicago Press.
- Neuman, W. Russell 1986 *The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate*. Cambridge, Mass.: Harvard University Press.
- Noelle-Neumann, Elisabeth 1984 *The Spiral of Silence*. Chicago: University of Chicago Press.
- Page, Benjamin I., and Robert Y. Shapiro 1983 "The Effects of Public Opinion on Policy." *American Political Science Review* 77:175–190.
- 1992 *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.
- Renshaw, Stanley A. (ed.) 1977 *A Handbook of Political Socialization*. New York: Free Press.
- Ross, Edward A. 1901 *Social Control: A Survey of the Social Foundations of Order*. New York: Macmillan.
- Schandler, Herbert J. 1977 *The Unmaking of a President: Lyndon B. Johnson and Vietnam*. Princeton N. J.: Princeton University Press.
- Schattschneider, E. E. 1960 *The Semisovereign People: A Realist's View of Democracy in America*. Hinsdale, Ill.: Dryden Press.
- Schuman, Howard, and Lawrence Bobo 1985 *Racial Attitudes: Trends and Interpretations*. Cambridge, Mass.: Harvard University Press.
- Schumpeter, Joseph A. 1942 *Capitalism, Socialism, and Democracy*. New York: Harper and Row.
- Sigel, Roberta S. (ed.) 1989 *Political Learning in Adulthood: A Sourcebook of Theory and Research*. Chicago: University of Chicago Press.
- Tilly, Charles 1978 *From Mobilization to Revolution*. New York: Random House.
- Tönnies, Ferdinand 1922 *Kritik der öffentlichen Meinung*. Berlin: Springer.
- Wetstein, Matthew E. 1996 *Abortion Rates in the United States: The Influence of Opinion and Policy*. Albany: State University of New York Press.

Zaller, John 1992 *The Nature and Origin of Mass Opinion*.  
New York: Cambridge University Press.

KURT LANG

## PUBLIC POLICY ANALYSIS

Public policy analysis is a large, sprawling intellectual enterprise involving numerous academic disciplines, private research organizations, and governmental agencies each sharing a common concern with the formulation, implementation, or consequences of public policy decisions. There are approximately thirty journals published in the English language alone and nearly twenty professional associations that are devoted more or less exclusively to policy analysis. Departments, centers, and institutes dealing in whole or in part with policy analysis can be found at over forty American universities.

As currently practiced, policy analysis involves contributions from the entire gamut of scientific disciplines. Much present-day public policy analysis is undertaken by scholars from the various applied physical and biological sciences (for example, environmental impact studies, technology assessments, seismic risk analyses, and the like). The focus here, however, is on public policy analysis as it is conducted within the social and behavioral sciences, principally economics, political science, and sociology.

The diversity of research work conducted under the rubric of "public policy analysis," even when restricted to the social science component, is perhaps the distinguishing characteristic of the subject; in the space available here we can do little more than indicate the range of topics and approaches with which policy analysts are concerned. Rogers (1989) has developed a typology of public policy research that is useful for this purpose; the following is adapted from his discussion.

### PROBLEM DEFINITION OR NEEDS ASSESSMENT

Public policy usually addresses real or sensed problems, and a great deal of public policy analysis is therefore devoted to defining or clarifying problems and assessing needs. What are the health care needs of a particular neighborhood? What are the

housing or nutritional needs of the nation's poverty population? What social services do homeless persons require? It is obvious that the development and formulation of public policy will be enhanced when underlying needs have been adequately described and analyzed. There is a large literature on the theory and practice of problem definition and needs assessment; students seeking additional information will find Johnson and colleagues (1987) invaluable.

### VALUE EXPLORATION OR CLARIFICATION

Given a demonstrated need, any number of policies might be developed to address it. Which policies, goals, or outcomes are most desirable? If an area is found to have unmet health needs, is it better to open freestanding clinics or to provide subsidized health insurance? Are the housing needs of the poor best addressed through public housing projects or through housing vouchers that can be used in lieu of rent? Should our policies with respect to the homeless attempt to ameliorate the conditions of a homeless existence, or prevent people from becoming homeless in the first place?

Assessing the relative desirability of policy options is only rarely an empirical matter; such decisions are more often ethical or ideological. MacRae (1985) stresses the unavoidable role of values in the process of policy analysis and the ensuing conflicts for the policy analyst. He identifies four principal "end values" widely shared throughout American society and against which policy decisions can be compared: economic benefit, subjective well-being, equity, and social integration. Sadly, policies that maximize equity may not maximize net economic benefit; those that enhance social integration may destroy subjective well-being. Thus, public policy analysis is not an arena for those who wish to pursue "value-neutral" science nor is it one for the morally or ideologically faint of heart.

### CONCEPTUAL DEVELOPMENT

Much work in the area of public policy analysis consists of developing conceptual schemes or typologies that help sort out various kinds of policies or analyses of policies (such as the typology we are presently using). Nagel (1984) and Dubnick



and Bardes (1983) review numerous conceptual schemes for typifying policies and policy analyses, with useful suggestions for synthesis; the former is an especially good overview of the field as a whole.

### **POLICY DESCRIPTION**

Adequate description of public policy is essential for proper evaluation and understanding, but many public policies prove frustratingly complex, especially as delivered in the field. "Poverty policy" in the United States consists of a vast congeries of federal, state, and local programs each focused on different aspects of the poverty problem (income, employment, housing, nutrition) or on different segments of the poverty population (women, children, women with children, the disabled, the elderly). The same can obviously be said of housing policy, tax policy, environmental policy, health policy, and on through a very long list. Even a single element of poverty policy such as Temporary Assistance to Needy Families (TANF) has different eligibility requirements, administrative procedures, and payment levels in each of the fifty states. Thus, accurate policy description is by no means a straightforward task. Outstanding examples of policy description, both focused on poverty policy, are Haveman (1977) and Levitan (1985).

### **POLICY FORMULATION**

Social science has a role to play in the formulation of policy as well as its description or evaluation. Most of the issues that policy attempts to address have been the focus of a great deal of basic social science research: poverty, ill health, homelessness, crime, violence, and so on. Although the once-obligatory discussion of "policy implications" of basic research has abated in recent years, few social scientists who work on policy-relevant issues can resist the urge to comment on the possible implications of the results for policy formulation. Much more work of this sort needs to be done, as it is evident that many policies are formulated and enacted in utter disregard for the extant state of knowledge about the topic. Indeed, Peter Rossi has hypothesized that the major reason social programs fail is that they are typically designed by amateurs who are largely innocent of social science theory, concepts, and results. Various job programs, mental health interventions, and crime reduction policies represent obvious cases in point.

### **METHODOLOGICAL RESEARCH**

Unlike much basic disciplinary research in the social and behavioral sciences, whose results are largely inconsequential except to a handful of specialists, the results of policy studies will often influence peoples' lives and well-being and the cost of being wrong can run into millions or billions of dollars. Thus issues of internal and external validity, errors of measurement and specification, proper statistical modeling, and the like are more than methodological niceties to the policy analyst; they are worrisome, ever-present and potentially consequential threats to the accuracy of one's conclusions and to the policy decisions that ensue. A technical error in a journal article can be corrected in a simple retraction; an equivalent error in a policy analysis might result in wrong-headed or counterproductive policies being pursued.

Much of the literature on public policy analysis, and especially on impact evaluation (see below), is therefore mainly methodological in character; indeed, many recent innovations in research procedure have been developed by scholars working on applied, as opposed to basic, problems. There are many texts available on the methodology of public policy analysis. Rossi and colleagues (1998) provide a comprehensive overview; Judd and Kenny (1981) are highly recommended for the more advanced student.

### **POLICY EXPLANATION**

Much public policy analysis undertaken by political scientists focuses on the processes by which policy is made at federal, state, and local levels. Classic examples are Marmor's analysis of the passage of Medicare (1970) and Moynihan's study of the ill-fated Family Assistance Plan proposed early in the Nixon administration but never enacted (1973).

Explanations of how public policy is made are invariably replete with the "dirty linen" of the political process: competing and often warring constituencies, equally legitimate but contradictory objectives and values, vote trading, compromises and deals, political posturing by key actors, intrusions by lobbying, advocacy and special interest groups, manipulation of public sentiment and understanding—in short, the "booming, buzzing

confusion” of a fractious, pluralistic political system. For those whose understanding of such matters does not extend much beyond the obligatory high school civics lesson in “how a bill becomes a law,” the policy explanation literature is a revelation.

### POLITICAL INTELLIGENCE OR PUBLIC OPINION

In a democratic society, public opinion is supposed to “count” in the policy formation process. Sometimes it does; often it does not. Policy analysis thus sometimes involves plumbing the depths and sources of support or opposition to various policy initiatives, and in a larger sense, explicating the process by which policy becomes legitimated.

There is no easy answer to the question whether (or under what conditions) public opinion dictates the direction of public policy. It is evident that policy makers are sensitive to public opinion; many presidents, for example, are morbidly fascinated by their standing in the polls (e.g., Sussman 1988). It is equally evident, however, that many policies with strong majority support are never enacted into law. An interesting study of the effects of public opinion on policy formation is Verba and Nie (1975).

### EVALUATION RESEARCH

The ultimate analytic question to be asked about any public policy is whether it produced (or will produce) its intended effects (or any effects, whether intended or not). The search for bottom-line effects—impact assessment—is one of two major activities subsumed under the rubric of evaluation research. The other is so-called process evaluation, discussed below under “Implementation Analysis.”

There are many formidable barriers to be overcome in deciding whether a policy or program has produced its intended (or any) effects. First, the notion of “intended effects” presupposes clearly defined and articulated program goals, but many policies are enacted without a clear statement of the goals to be achieved. Thus, many texts in evaluation research recommend an assessment of the “evaluability” of the program prior to initiating the evaluation itself. A second barrier is the often-pronounced difference between the program-as-designed and the program-as-delivered. This

is the issue of program implementation, discussed below.

The most troublesome methodological issue in evaluation research lies in establishing the *ceteris paribus* (or “all else equal”) condition, or in other words, in estimating what might have happened in the absence of the program to be evaluated. In an era of declining birthrates, any fertility reduction program will appear to be successful; in an era of declining crime rates, any crime reduction program will appear to be successful. How, then, can one differentiate between program effects and things that would have happened anyway owing to exogenous conditions? (Students of logic will see the problem here as the *post hoc, ergo propter hoc* fallacy.)

Because of this *ceteris paribus* problem, many evaluations are designed as experiments or quasi-experiments. In the former case, subjects are randomly assigned to various treatment and control conditions, and outcomes are monitored. Randomization in essence “initializes” all the starting conditions to the same values (except for the vagaries of chance). In the recent history of evaluation research, the various negative income tax experiments (see Rossi and Lyall 1976) are the best-known examples of large-scale field experiments of this general sort. *Quasi-experiments* are any of a number of research designs that do not involve randomization but use other methods to establish the *ceteris paribus* condition; the definitive statement on quasi-experiments is Cook and Campbell (1979).

Nowhere is the trade-off between internal and external validity more vexing than in the design of program evaluations. Evaluation designs with high internal validity, such as randomized experiments, are excellent in detecting program effects but the experimental conditions may not generalize to real-world settings. Thus, one telling critique of the Negative Income Tax (NIT) experiments is that participants knew from the beginning that the program would end in three (or in some cases five) years, so the labor-force response may have been very different than it would have been if negative income taxation became a permanent element of national income policy. Likewise, as the research setting comes to more closely mimic real-world conditions (that is, as it develops high external

validity), the ability to detect real effects often declines.

A final problem in doing evaluation research is that most policies or programs are relatively small interventions intended to address rather large, complex social issues. The poverty rate, to illustrate, is a complex function of the rate of employment, trends in the world economy, prevailing wage rates, the provisions of the social welfare system, and a host of additional macrostructural factors. Any given antipoverty program, in contrast, will be a relatively small-scale intervention focused on one or a few components of the larger problem, often restricted to one or a few segments of the population. Often, the overall effects of the various large-scale, macrostructural factors will completely swamp the program effects—not because the program effects were not present or meritorious but because they are very small relative to exogenous effects.

The literature on the theory and practice of evaluation research is expansive; students seeking additional information will find themselves well served by Chambers and colleagues (1992), and by Rossi and colleagues (1998).

### OUTCOME ANALYSIS

Assuming that a program has been adequately evaluated and an effect documented, one can then analyze that effect (or outcome) to determine whether it was worth the money and effort necessary to produce it. Outcome analysis thus examines the cost effectiveness or cost beneficiality of a given policy, program, or intervention.

Cost-benefit and cost-effectiveness analysis are intrinsically complex, technically demanding subjects. One complication lies in assessing the so-called opportunity costs. A dollar spent in one way is a dollar no longer available to use in some other way. Investing the dollar in any particular intervention thus means that one has lost the “opportunity” to invest that dollar in something that may have been far more beneficial.

A second complication is in the “accounting perspective” one chooses to assess benefits and costs. Consider the Food Stamp program. A recipient receives a benefit (a coupon that can be redeemed for food) at no cost; from the accounting

perspective of that recipient, the benefit-cost ratio is thus infinite. The Food Stamp program is administered by the United States Department of Agriculture (USDA). From the USDA perspective, the benefit of the program presumably lies in the contribution it makes to relieving hunger and malnutrition in the population; the cost lies in whatever it takes to administer the program, redeem the coupons once submitted by food outlets, etc. Accounted against the USDA perspective, the benefit-cost ratio will be very different, and it will be different again when accounted against the perspective of society as a whole. The latter accounting, of course, requires asking what it is worth to us as a nation to provide food to those who might otherwise have to go without, clearly a moral question more than an empirical or analytic one.

This last example illustrates another thorny problem in doing cost-benefit analyses, namely, the incommensurability of benefits and costs. The dollar costs of most programs or policies can be reasonably well estimated. (The dollar costs are usually not the only costs. There may also be ethical or political costs that cannot be translated into dollars and cents but that are, nonetheless, real. Let us ignore the nondollar costs, however.) Unfortunately, the benefits of most interventions cannot be readily expressed in dollars; they are expressed, rather, in less tangible (but equally real) terms: lives saved, improvements in the quality of life, reductions of hunger, and the like. If the outcome cannot be converted to a dollar value, then a strict comparison to the dollar costs cannot be made and a true benefit-cost ratio cannot be calculated.

Cost effectiveness analysis, in contrast, compares the benefits of one program (expressed in any unit) at one cost to the benefits of another program (expressed in the same unit) at a different cost. Thus, a program that spends \$10,000 to save one life is more cost effective than another program that spends \$20,000 to save one life. Whether either program is cost beneficial, however, cannot be determined unless one is willing to assign a dollar value to a human life.

Many texts by economists deal at length with these and related complexities; accessible overviews include Levin (1975) and Yates (1996).

**IMPLEMENTATION ANALYSIS**

“Much is the slippage between the spoon and the mouth.” A program as it is delivered in the field is rarely identical to the program as designed in the policy making process; sometimes, there is only a superficial resemblance. Since slippage between design and implementation might provide one explanation for the failure to achieve significant program effects, implementation analysis is an essential component of all capable policy evaluations.

There are many reasons why programs-as-delivered differ from programs-as-designed: technical impossibility, bureaucratic inertia, unanticipated conditions, exogenous influences. An elegantly designed policy experiment can fail at the point of randomization if program personnel let their own sentiments about “worthy” and “unworthy” clients override the randomizing process. Many educational policy initiatives are subverted because teachers persist in their same old ways despite the program admonition to do things differently. Welfare reform will mean little if caseworkers continue to apply the same standards and procedures as in the past. More generally, the real world finds ways to impinge in unexpected and often unwanted ways on any policy initiative; failure to anticipate these impingements has caused many a policy experiment to fail.

Loftin and McDowell (1981) provide a classic example of the utility of implementation analysis in their evaluation of the effects of the Detroit mandatory sentencing law. The policy-as-designed required a mandatory two-year “add on” to the prison sentence of any person convicted of a felony involving a firearm. Contrary to expectation, the rate of firearms crime did not decline after the law was enacted. Implementation analysis provided the reason. Judges, well aware of the overcrowded conditions in the state’s prisons, were loath to increase average prison sentences. Yet, state law required that two years be added to the charge. To resolve the dilemma, judges in firearms cases would begin by reducing the main sentence by two or so years and then adding the mandated two-year add-on, so that the overall sentence remained about the same even as the judges remained in technical compliance with policy. A more thorough discussion of the implementation problem can be found in Chambers and colleagues (1992, chap. 1).

**UTILIZATION**

A consistent frustration expressed throughout the literature is that policy analysis seems only rarely to have any impact on actual policy. Utilization is an ongoing problem in the field of evaluation research. A more detailed treatment of the utilization problem can be found in Chambers and colleagues (1992, chapter 1), Shadish and colleagues (1991, chapters 6, 7), and Weiss (1988). For examples of ways in which evaluation can impact practice, see articles by Gueron, Lipsey, and Wholey in *New Directions for Evaluation* (1997).

Many reasons for nonutilization have been identified. One of the most important is timeliness. Good research takes time, whereas policy decisions are often made quickly, well before the results of the analysis are in. The negative income tax experiments mentioned earlier were stimulated in substantial part by a Nixon administration proposal for a modified negative income tax to replace the then-current welfare system. The shortest of the experiments ran for three years; several ran for five years; none were completed by the time the Nixon proposal was killed mainly on political grounds.

A second factor in the nonutilization of policy studies is that research is seldom unequivocal. Even the best-designed and best-executed policy researches will be accompanied by numerous caveats, conditions, and qualifications that strictly limit the safe policy inferences one may draw from them. Policy makers, of course, prefer simple declarative conclusions; policy research rarely allows one to make such statements.

Finally, even under the most favorable conditions, the scientific results of policy analyses are but one among many inputs into the policy-making process. There are, in addition, normative, economic, political, ethical, pragmatic, and ideological inputs that must be accommodated. In the process of accommodation, the influence of scientific research is often obscured to the point where it can no longer be recognized. It should not be inferred from this that policy analysis is not utilized, only that the research results are but one voice in the cacophony of the policy-making process.

Weiss has written extensively on the utilization problem and ways in which evaluation can be

used effectively to change policy. She argues that “in its ideal form, evaluation is conducted for a client who has decisions to make and who looks to the evaluation for answers on which to base his decisions” (1972, p. 6). This is often not the case, however, as evaluation results seldom influence important decisions regarding programs and policies. Weiss’s general conclusion regarding utilization is that evaluation results affect public policy by serving as the impetus for public discourse and debate that form social policy, rather than through extensive program reform or termination.

## REFERENCES

- Chambers, K., K. R. Wedel, and M. K. Rodwell 1992 *Evaluating Social Programs*. Boston: Allyn and Bacon.
- Cook, Thomas, and Donald Campbell 1979 *Quasi-Experimentation*. Chicago: Rand MacNally.
- Dubnick, Melvin, and Barbara Bardes 1983 *Thinking about Public Policy: A Problem Solving Approach*. New York: John Wiley.
- Gueron, Judith M. 1997 “Learning about Welfare Reform: Lessons from State-Based Evaluations.” *New Directions for Evaluation* 76:79–94. (Edited by D. Rog and D. Fournier)
- Haveman, Robert 1977 *A Decade of Federal Antipoverty Programs: Achievements, Failures, and Lessons*. New York: Academic.
- Johnson, D., L. Meiller, L. Miller, and G. Summers 1987 *Needs Assessment: Theory and Methods*. Ames: Iowa State University Press.
- Judd, Charles, and David Kenny 1981 *Estimating the Effects of Social Interventions*. New York: Cambridge University Press.
- Levin, Henry 1975 “Cost-Effectiveness Analysis in Evaluation Research.” In M. Guttentag and E. Struening, eds., *Handbook of Evaluation Research*. Newbury Park, Calif.: Sage.
- Levitan, Sar 1985 *Programs in Aid of the Poor*. Baltimore, Md.: Johns Hopkins University Press.
- Lipsey, Mark W. 1997 “What Can You Build with Thousands of Bricks? Musings on the Cumulation of Knowledge in Program Evaluation.” *New Directions for Evaluation* 76:7–24. (Edited by D. Rog and D. Fournier.)
- Loftin, Colin, and David McDowell 1981 “One with a Gun Gets You Two: Mandatory Sentencing and Firearms Violence in Detroit.” *Annals of the American Academy of Political and Social Science* 455:150–168.
- MacRae, Duncan 1985 *Policy Indicators: Links between Social Science and Public Debate*. Chapel Hill, N.C.: University of North Carolina Press.
- Marmor, Theodore 1970 *The Politics of Medicare*. New York: Aldine.
- Moynihan, Daniel 1973 *The Politics of a Guaranteed Annual Income: The Nixon Administration and the Family Assistance Plan*. New York: Vintage.
- Nagel, Stuart 1984 *Contemporary Public Policy Analysis*. Birmingham: University of Alabama Press.
- Rogers, James 1989 “Social Science Disciplines and Policy Research: The Case of Political Science.” *Policy Studies Review* 9:13–28.
- Rossi, Peter, Howard Freeman, and Mark Lipsey 1998 *Evaluation: A Systematic Approach*, 6th ed. Newbury Park, Calif.: Sage.
- Rossi, Peter, and Kathryn Lyall 1976 *Reforming Public Welfare*. New York: Russell Sage.
- Shadish, William R., Jr., T. D. Cook, and L. C. Leviton 1991 *Foundations of Program Evaluation*. Newbury Park, Calif.: Sage.
- Sussman, Barry 1988 *What Americans Really Think and Why Our Politicians Pay No Attention*. New York: Pantheon.
- Verba, Sidney, and Norman Nie 1975 *Participation in America: Political Democracy and Social Equality*. New York: Harper and Row.
- Weiss, Carol H. 1972 *Evaluation Research: Methods for Assessing Program Effectiveness*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1988 “Evaluation for Decisions: Is Anybody There? Does Anybody Care?” *Evaluation Practice* 9:15–28.
- Wholey, Joseph S. 1997 “Clarifying Goals, Reporting Results.” *New Directions for Evaluation* 76:95–106. (Edited by D. Rog and D. Fournier.)
- Yates, Brian T. 1996 *Analyzing Costs, Procedures, Processes, and Outcomes in Human Services*. Thousand Oaks, Calif.: Sage.

JAMES D. WRIGHT

# Encyclopedia of Sociology

---

Second Edition

# Encyclopedia of Sociology

---

Second Edition

VOLUME 4

**Edgar F. Borgatta**

*Editor-in-Chief*

*University of Washington, Seattle*

**Rhonda J. V. Montgomery**

*Managing Editor*

*University of Kansas, Lawrence*

**Macmillan Reference USA**

*an imprint of the Gale Group*

*New York • Detroit • San Francisco • London • Boston • Woodbridge, CT*

# Q

---

## QUALITATIVE METHODS

The term *qualitative methods* refers to a variety of research techniques and procedures associated with the goal of trying to understand the complexities of the social world in which we live and how we go about thinking, acting, and making meaning in our lives. These research practices, which emphasize getting close to participants and trying to understand how they (and we) view the world, include, among others, participant observation, interviews, life histories, and focus groups; autoethnographic, phenomenological, narrative, and most ethnomethodological and feminist approaches; particular forms of documentary, content, discourse, and conversational analysis research; and some action research.

Qualitative researchers may be placed along a broad continuum ranging from an orientation akin to positivist science to one more akin to art and literature. In between is a vast middle ground where elements of both orientations are present. Moving along the qualitative continuum from science to art and literature, one finds practitioners who see social life as something out there to be discovered independently of the researcher, those who view social life as something constructed through interaction and engagement with the world, and those who focus more closely on the person describing social life and the modes and practices of description (see Crotty 1998; Denzin 1997). Across the continuum, the focus changes from studying others who are assumed to be uniquely separate from the researcher, to examining interactions between the researcher and others, to

including the positionality, politics, and story of the researcher who interacts with others.

Currently qualitative work enjoys a burgeoning interest across social science disciplines including anthropology, sociology, communication, education, social work, and nursing. The result is a growing sense of a qualitative community unconstrained by disciplinary boundaries. As this community grows and forms its identity, the spectrum of possibilities broadens, creating new alternatives for qualitative research and, in the process, raising vexing and controversial questions (see Denzin and Lincoln 1995; Snow and Morrill 1995). Given the interpretive turn (Rabinow and Sullivan 1987) in social science, more and more researchers are applying art-based criteria to their work; at the same time, new computer programs, such as NVivo, allow for more systematic and rigorous coding of qualitative data. We view these differences and ensuing conversations as strengthening qualitative research and representing a coming of age for qualitative work (Bochner and Ellis in press).

We organize our discussion of qualitative methods according to three ideal types representing points on the qualitative continuum from the science/causal pole through the middle ground to the artful/interpretive pole. These categories are intended as a useful means of dividing the territory, but they are also arbitrary and should not be taken as a literal map of the field. Rather, we encourage readers to envision a wide expanse of methodological approaches and to view the boundaries we have constructed as permeable, flexible, and fleeting, with many qualitative researchers



likely to position themselves in more than one category. As the authors, both of us have engaged in middle ground and artful/interpretive qualitative work; our current allegiance lies primarily in the artful/interpretive region of the continuum.

### QUALITATIVE RESEARCH AS SCIENCE

At one end of the qualitative spectrum, researchers approach qualitative research as an extension of quantitative inquiry. Their goal is to produce propositional knowledge about human behavior generalizable to specific populations. They see truth as something "out there" to be discovered and explained by the researcher. Positioning themselves as neutral, disinterested parties in the research process, they want to be able to control and predict subsequent behavior of samples within the population being investigated. These researchers follow the building-block, foundational model of scientific knowledge, viewing inquiry as a linear progression, with each new discovery adding on to available explanations. Qualitative researchers in this tradition express many of the same concerns as their quantitative counterparts, including an interest in random sampling, reliability, validity, and ethical issues. The language they use to present and discuss results applies many of the familiar lines of the hypothetico-deductive model.

Random sampling of the studied population (regardless of specific research tools to be used) provides assurance that the qualitative researcher has obtained a representative and unbiased sample of research subjects (Lindlof 1995, p. 121). Since researchers want to claim that their findings can be generalized to the selected population, it is critical that the demographic characteristics of the sample match those of the population defined in the study. For example, Lowry and Towles (1989), in their study of the portrayal of sex and its consequences in afternoon soap operas, randomly sampled episodes of soap operas from each TV network in order to be able to draw conclusions about soap opera content in general.

Using an approach similar to that of quantitative research, qualitative researchers in this tradition examine variables that relate to specific behaviors, traits, or characteristics that are narrowly defined in as specific a manner as possible. They manipulate the independent variables and measure the outcome of the experiment in terms of

dependent variables, those defined behaviors, traits, or characteristics thought to exist in relationship to the independent variables. Before they conduct experiments, researchers form a hypothesis about the relationship between the variables. Data interpretation then centers on determining whether the hypotheses are negated by the results; researchers do not generally examine data for themes or issues unrelated to the predetermined focuses of the study. Based on a review of relevant literature, Chavez (1985), for example, hypothesized that writers treat men and women differently in comic strips. She then collected a sample of comic strips, coded the gender, settings, and activities of the characters, and concluded that her hypothesis was supported.

This process of hypothesis formation and testing often is less formal, however, even among those striving to maintain a scientific approach to their research, and it can take many different forms. In the study of the realism of aggression on television by Potter et al. (1995), for example, the authors laid out a set of premises about what would constitute a realistic (similar to the real world levels of violence in numbers and context) portrayal of aggression. They then collected a sample of television programming, coded the various acts of aggression, and compared the numbers and types of aggression in their sample to the premises they had developed.

Even those researchers who do not convert their data into numerical form often go to great lengths to assure that their findings are valid and reliable. For researchers at this end of the qualitative continuum, validity means that the concepts they examine are those intended to be examined and not confounded or overlapping with other concepts not intended to be included. A study is reliable if researchers find, or could expect to find, the same, or very similar, results when they conduct the study again. For example, Waitzkin (1990) provides criteria to establish reliability and validity of qualitative data: (1) discourse should be selected through a sampling procedure, preferably a randomized technique; (2) recordings of sampled discourse should be available for review by other observers; (3) standardized rules of transcription should be used; (4) the reliability of transcription should be assessed by multiple observers; (5) procedures of interpretation should be decided in advance, should be validated in relation to theory,

and should address both content and structure of texts; (6) the reliability of applying interpretive procedures should be assessed by multiple observers; (7) a summary and excerpts from transcripts should accompany the interpretation, but full transcripts should also be available for review; and (8) texts and interpretations should convey the variability of content and structure across sampled texts (pp. 480–482).

Waitzkin's criteria (1990) emphasize three specific concerns associated with reliability and validity from a scientific perspective. First, the collective decisions and interpretations of multiple researchers would be closer to an "objective" reality than a presumably more biased perspective of a single individual. Second, it is important to consider the whole body of available data as the basis for interpretation to avoid making general statements that reflect only a subset of the data; the emphasis is on what is common throughout the data, not on that which is unusual, distinctive, or unique. Third, written transcripts must be publicly available for verification.

Qualitative researchers working in this tradition use a system of coding to categorize videotaped or observed behaviors, written responses to survey questions, verbal responses to an interviewer, or other data (Wimmer and Dominick 1997). Once labeled, the observed behaviors can be counted, sorted, and analyzed statistically. Coding schema can be standardized typologies that are used by other researchers, or they can be developed in light of a specific research question or body of data. Chavez (1985), mentioned above, developed a typology of settings and activities (e.g., child care or working in an office) for the cartoon characters based on what she found in her comic strip data set. To aid in the analysis of transcript data, specialized computer software programs, such as NUD\*IST, Ethnograph, or NVivo (a new program that integrates text, image, sound, and video data sets), are available. A critical component to coding is establishing intercoder agreement; that is, a measure to ensure that the coding schema can be consistently applied to the same set of data by different coders and the same or very similar results obtained (Wimmer and Dominick 1997).

Ethical issues at this end of the continuum, similar to those in quantitative research, focus on methodological procedures, in particular honesty

and thoroughness in data collection and analysis. Authors often elaborately spell out their research procedures in their publications, making the procedures and data available for scrutiny in order to justify claims or conclusions they draw. In science-oriented qualitative research, authors stay behind the scenes, portraying themselves as trustworthy and credible through their disembodied discussion of methods without showing in their texts their own involvement or self-interest.

Those engaging in scientific approaches to qualitative research usually adhere closely to the writing style used by quantitative researchers. A passive voice shadows the presence of the author and obscures the "I" of the researcher (Gergen 1994). Statements such as "It was found that . . ." and "The data revealed that . . ." reinforce the notion of neutral authors who have discovered preexisting ideas, and who, without contaminating the material with their own perspectives, then pass it along for readers to receive as knowledge. Of course, researchers who see their work as scientific often acknowledge that the author is not a blank slate without values and beliefs, but their use of a disinterested, passive voice remains a sign of how important they view the ideal of distance and objectivity, even if it is not fully attainable.

#### MIDDLE-GROUND APPROACHES TO RESEARCH

Between science and art, one finds a sprawling middle ground of qualitative researchers, who seek to analyze events, find patterns, and create models from their data (Neuman 1997). Here, researchers do not adhere rigidly to the rules of empiricism; but they are not likely to experiment with narrative, poetic, or literary forms of writing either. In the middle, researchers seek some combination of scientific rigor and artistic imagination. How these two goals intersect differs for various researchers and often connects to the author's specific location relative to art and science on the qualitative continuum.

Middle-ground researchers use a variety of methodologies to gather data for analysis, including unstructured or semistructured interviewing (Fontana and Frey 1994; Mishler 1986), focus groups (Kitzinger 1994), participant observation or fieldwork (Lincoln and Guba 1985; Lofland and Lofland 1995), textual analysis (Reinharz 1992),

and analysis of narrative (M. M. Gergen 1992; Riessman 1993). While there are many ways to go about selecting a sample, those in the middle ground of qualitative research often use purposeful sampling (Miles and Huberman 1984), in which they try to obtain cases that are information rich, or a “snowball” approach (Reinharz 1992), in which they ask current participants to identify other possible participants.

One of the most useful and widely applied strategies associated with the middle ground is an approach, developed by Glaser and Strauss (1967), called *grounded theory*. In this approach, researchers emphasize the generation of categories and theory from systematic coding and analysis of qualitative data (often transcripts) (Charmaz 1990; Glaser and Strauss 1967; Janesick 1994). This method of inductive reasoning differs from traditional social science in which researchers use previously established theory and test it deductively with data to see whether the theory can be sustained.

Methodological concerns are very important to grounded theory researchers, who often hold to the belief that if you apply a valid and systematic methodological approach, you’ll get closer to an accurate representation of what’s actually going on. In analyzing the data, some adhere rigidly to formal steps in grounded theory research—data notes, sort and classify, open coding, axial coding, selective coding, with memo writing occurring throughout the process (Charmaz 1990; Neuman 1997; Strauss and Corbin 1994). The more the researcher adheres to systematic analysis, the greater the likelihood of using computer programs to assist in coding. Other theorists, in the middle of the continuum, who think of themselves as grounded theorists, “eyeball” the data in less systematic ways, perhaps not even completely transcribing or coding data into categories. Yet they too seek patterns among the data they have collected, though they view the process as more subjective and intuitive than scientific and objective.

While some middle-ground researchers may place less emphasis on scientific precision, they usually adhere to criteria or guidelines concerning the processes of data analysis, though these rules may vary widely (see, for example, Charmaz 1997; Glaser 1978; Strauss and Corbin 1997). Tompkins (1994), for example, refers to representativeness,

consistency (for public and private texts), and recalcitrance (sanction by research participants or a similar group) as standards for evaluating data used in qualitative research (see also Fitch 1994). Working closer to the interpretive pole, Lather (1986) argues that validity in openly ideological research can be established through four guidelines: *triangulation* of multiple data sources, methods, and theoretical perspective; *assessment* of construct validity through use of systematized reflexivity between the daily experiences of people and the theoretical constructs being developed; *establishment* of face validity through sharing analysis with respondents and refining analysis with their input; and *determination* of catalytic validity, that is the potential for bringing about positive change and social transformation (p. 67). These different criteria are similar insofar as they provide standards for bridging researcher and participant perspectives, so that findings reflect the meanings of the people whose lives were examined.

Ethical issues for those in the middle group focus on research practices such as covert research, deception, informed consent, anonymity, confidentiality, and revealing knowledge about the less powerful. These issues then lead to ethical questions about what should be studied, how, and by whom (see Lofland and Lofland 1995).

Most middle-ground researchers note the positionality of participants, such as race, class, and sexual orientation, in order to avoid obscuring these factors. For example, Ellingson and Buzzanell (1999) studied a group of white, middle-class, heterosexual breast cancer survivors in a small Midwestern city. They acknowledged that these demographic characteristics impacted the results of the study; a more racially mixed group, or a group composed of lesbians, for instance, most likely would have produced a different set of findings.

Just as those in the middle ground acknowledge the positionality of participants, they also sometimes acknowledge their standpoint, personal background, politics, and interests in the topic (Collins 1991). Insofar as they see knowledge as “constructed” rather than discovered, middle-ground researchers discuss their personal perspectives or political commitments as an acknowledgment that all knowledge is generated from a

specific social position and reflects the perspectives of the researchers involved. Researchers in the middle ground may try to decrease the power disparity between themselves and the people they study (DeVault 1990; Ellingson and Buzzanell 1999). They generally refer to those in the study as research participants or informants rather than subjects, indicating a degree of respect for the people whose lives are being studied (Reinharz 1992).

In the middle ground, researchers study a variety of topics, including complex issues that are difficult or impossible to address with quantitative methodology, such as understanding hierarchy in groups (Whyte [1943] 1993) or awareness contexts in death (Glaser and Strauss 1964). Some seek to make visible previously invisible aspects of the lives of women and other groups underrepresented in traditional social scientific research such as ethnic and racial minorities, gays and lesbians, and people with disabilities (Spitzack and Carter 1989). Others examine groups that are hidden, unknown, or inaccessible to most people, such as mushroom gatherers (Fine 1992) or white supremacists (Mitchell 1998). As the topics get more complex and oriented toward meanings, subjectivity, and emotionality, it becomes more difficult to invoke older, more traditional, systematic "scientific methods" and apply them.

Writers in this tradition alter some of the conventions of scientific writing. They may include standpoint statements within the introductory section of articles, indicating their personal interest in the topic. They may use a conventional format but preface the article with vignettes or include excerpts from participant narratives in the discussion of findings or in an appendix to add texture and authenticity to the work. To acknowledge their presence in the work, authors may write in the first person. Nevertheless, researchers in this tradition usually privilege theory generation, typicality, and generalization to a wider world over evocative storytelling, concrete experience, and multiple perspectives that include participants' voices and interpretations. They tend to write realist tales in an authorial, omnipotent voice. Snippets of fieldwork data then represent participants' stories, primarily valued for illustrating general concepts, patterns, and themes (see Van Maanen 1988).

## RESEARCH AS ARTISTIC ENDEAVOR

During the last two decades, many qualitative researchers have moved toward an emphasis on the artistic aspects of qualitative work (Wolcott 1995). Working from an orientation that blends the practices and emphases of social science with the aesthetic sensibility and expressive forms of art, these researchers seek to tell stories that show experience as lived in all its bodily, cognitive, emotional, and spiritual aspects. The goal is to practice an artful, poetic, and empathic social science in which readers can keep in their minds and feel in their bodies the complexities of concrete moments of lived experience. These writers want readers to be able to put themselves in the place of others, within a culture of experience that enlarges their social awareness and empathy. Their goals include: evoking emotional experience in readers (Ellis 1997); giving voice to stories and groups of people traditionally left out of social scientific inquiry (DeVault 1990); producing writing of high literary/artistic quality (Richardson in press); and improving readers', participants', and authors' lives (see Denzin 1997; Fine 1994).

According to Bochner, Ellis, and their colleagues (Bochner 1994; Bochner et al. 1998; Ellis 1997), the interpretive, narrative, autoethnographic project has the following distinguishing features: the author usually writes in the first person, making herself or himself the object of research (Jackson 1989; Tedlock 1991); the narrative text focuses on generalization within a single case extended over time (Geertz 1973); the text is presented as a story replete with a narrator, characterization, and plot line, akin to forms of writing associated with the novel or biography; the story often discloses hidden details of private life and highlights emotional experience; the ebb and flow of relationship experience is depicted in an episodic form that dramatizes the motion of connected lives across the curve of time; a reflexive connection exists between the lives of participants and researchers that must be explored; and the relationships between writers and readers of the texts is one of involvement and participation.

Rather than believing in the presence of an external, unconstructed truth, researchers on this end of the continuum embrace narrative truth (Spence 1982), which means that the experience as described is believable, lifelike, and possible.

Through narrative we learn to understand the meanings and significance of the past as incomplete, tentative, and revisable according to contingencies of present life circumstances (Crites 1971). In this research, authors are concerned about issues of validity, reliability, and generalizability, but these issues may be redefined to meet the goals of the research.

Ellis and Bochner (in press), for example, define *validity* to mean that the work resonates with readers and evokes in them a feeling that the experience has verisimilitude. A story is valid if it stimulates readers to enter the experience described or to feel and think about their own, whether in memory, anticipated, or lived. Validity might be judged by whether it offers readers assistance in communicating with others different from themselves, or a way to improve the lives of participants and readers or even the author's own. Since writers always create their personal narrative from a situated location, trying to make their present, imagined future, and remembered past cohere, orthodox *reliability* does not exist in narrative research. But reliability checks are still important. Researchers often take their work back to participants and give them a chance to comment, add materials, change their minds, and offer their interpretations. Since we all participate in a limited number of cultures and institutions, lives are typical and generalizable, as well as particular. A story's *generalizability* is constantly being tested by readers as they determine whether the story speaks to them about their own experiences or about the experiences of others they know. Likewise, does it tell them about people or lives about which they are unfamiliar? Does a work bring "felt" news from one world to another and provide opportunities for the reader to have vicarious experience of the things told (Stake 1994)?

Interpretive research reflects the messiness of lived experience and emotions. Unlike researchers at the scientific end and in the middle ground, artful or interpretive researchers do not look for common denominators and generalities, but instead examine experience that is unique, particular, moving, and possible. They embrace their own subjectivity not to acknowledge bias but to celebrate positionality and their particular construction of the world. While some of these writers employ traditional analysis in their work, examining stories for concepts, patterns, and themes,

others argue for the importance of thinking *with* a story, not just about a story. Thinking *with* a story means to allow yourself to resonate with the story, reflect on it, become a part of it (see Frank 1995b). Others argue that theory is embedded in the story (Ellis 1995b), that all good stories make a theoretical point.

Arguments about methods are not nearly as prevalent here as in the other two groups. Methods articles are much more likely to emphasize flexibility and emergence than to offer strict rules for how to proceed (Ellis et al. 1997). One of the most discussed methodological issues is whether researchers should attempt to follow the rules of traditional ethnographic methods or whether narrative research should be approached more like writing fiction. One's position on that question most likely intersects with one's stance regarding the role of criteria in evaluating narrative writing. In the former, traditional ethnographic criteria might be more commonplace; in the latter, narratives might be judged by their usefulness, the compassion they encourage, and the dialogue they promote (Ellis and Bochner in press). Seeking to position interpretive texts as an intersection of social science and art form, Richardson (in press) discusses five criteria she uses to evaluate interpretive ethnography: (1) *Substantive contribution*: Does the work contribute to an understanding of the text? (2) *Aesthetic merit*: Is the text artistically shaped, complex, and satisfying? (3) *Reflexivity*: Does the writer reflect on the production of the text so that the reader can make judgments about the point of view employed in the text? (4) *Impactfulness*: Does this work affect me and move me to respond or act? (5) *Expression of Reality*: Does this work seem a credible account of the real?

Rather than method and criteria, most articles about interpretive/artful ethnography grapple with issues of writing. For narrative researchers, writing (the form) is inseparable from the process of data interpretation (the content). As Richardson (in press) phrases it, writing in interpretive qualitative research is not a "mopping-up" activity at the end of a research project but an integral part of data analysis; authors write to find out what they have to say about their data and their experiences. The voice of the author is as central to the text as the voices of those "being studied." Writers experiment with approaches and writing conventions to

dislodge assumptions about science and incorporate the artistic into representations of data.

Interpretive research embraces a range of processes and approaches, including biographical method (Denzin 1989), observation of participation (Tedlock 1991), ethnography (Van Maanen 1988), autoethnography (Ellingson 1998; Reed-Danahay 1997), interactive interviewing (Ellis et al. 1997), systematic sociological introspection (Ellis 1991), co-constructed methods (Bochner and Ellis 1992), personal experience methods (Clandinin and Connelly 1994), narrative analysis (Bochner 1994), and feminist methods (Reinharz 1992).

Examples of creative representation include layered accounts (Ronai 1995), ethnographic fiction (Angrosino 1998), personal essays (Krieger 1991), impressionist tales (Van Maanen 1988), co-constructed narratives (Bochner and Ellis 1992), poetic representation of data (Austin 1996; Glesne 1997; Richardson 1992, 1994), writing stories (Richardson 1997), ethnographic performance or ethnodrama (Jones 1998; Mieniczakowski 1996), and polyvocal texts (Lather and Smithies 1997).

A particularly controversial narrative writing practice and form is autoethnography, which is an autobiographical genre of writing and research that displays multiple layers of consciousness as it connects the personal to the cultural. Several autoethnographic genres currently exist side by side (Ellis and Bochner in press): (1) Reflexive ethnographies focus on a culture or subculture, but authors use their own experiences in the culture reflexively to bend back on self and look more deeply at self-other interactions. Reflexive ethnographers ideally use all their senses, their bodies, moment, feeling, and their whole being to learn about the other (Jackson 1989). (2) In native ethnographies, researchers who are natives of cultures that have been marginalized or exoticized by others write about and interpret their own cultures for others. (3) In personal narratives, social scientists take on the dual identities of academic and personal selves to tell autobiographical stories about some aspect of their daily life.

In all these forms of qualitative writing, narrative researchers use fiction-writing techniques such as dramatic recall, dialogue, flashback, strong imagery, scene setting, character development, interior monologue, suspense, and a dramatic plot line that is developed through the specific actions

of specific characters with specific bodies doing specific things. They ask readers to relive the events with the writer and then to reflect on their emotional responses and life experiences, and on the moral questions and concerns that arise. Yet, the work differs from fiction in that the writing and publishing conventions used arise out of social science traditions, and in that the work often, though not always, has more of an overt analytic purpose and more of an analytic frame than fiction has.

Ethical concerns include matters of how we go about doing our research and what we owe those who become characters in and readers of our stories (Ellis 1995a). How do we deal with issues of confidentiality when including people, such as family members, who can be easily identified? What do we do if characters in our stories disagree with our interpretations or want us to omit material about them? How do we include multiple voices and perspectives within our stories? What is the role of the traditional authorial voice of the author? How do we stay true to our participants yet not deceive our readers (Josselson 1996; see also Chase 1996)? Is there value in working from an ethic of care, empathy, and responsibility, rather than informed consent, and is that ever possible in the world of research (Collins 1991; Denzin 1997)? How do we make our projects therapeutic for ourselves as well as our participants and readers? What do we want the world to be? How can we contribute to making it that way and, in the process, become better human beings?

## CONCLUSION

Qualitative methods is a rich and varied set of approaches to research. Journals such as *Qualitative Inquiry*, *Qualitative Sociology*, *Symbolic Interaction*, and the *Journal of Contemporary Ethnography*, as well as a number of research annuals such as *Studies in Symbolic Interaction* and *Cultural Studies: A Research Annual*, along with subfield-specific journals, such as *Qualitative Health Research*, showcase examples of qualitative research in sociology and provide forums for discussion of methodological issues. The joy (and sometimes the frustration!) of qualitative methods is the promotion and valuing of a wide spectrum of methods, none of

which should be viewed as the only way to conduct research in sociology. It might be more comforting if there was one set of rules to follow, but that comfort would come with the tragic price of close-mindedness, silencing of voices, and narrowing of vision. We agree with Rorty (1982), who says we ought to learn to live with differences without feeling we have to resolve them.

REFERENCES

- Angrosino, M. V. 1998 *Opportunity House: Ethnographic Stories of Mental Retardation*. Walnut Creek, Calif.: AltaMira.
- Austin, D. A. 1996 "Kaleidoscope: The Same and Different." In C. Ellis and A. P. Bochner, eds., *Composing Ethnography: Alternative Forms of Qualitative Writing*. Walnut Creek, Calif.: AltaMira.
- Bochner, A. P. 1994 "Perspectives on Inquiry II: Theories and Stories." In M. Knapp and G. R. Miller, eds., *The Handbook of Interpersonal Communication*, 2nd ed. Thousand Oaks, Calif.: Sage.
- , and C. Ellis 1992 "Personal Narrative as a Social Approach to Interpersonal Communication." *Communication Theory* 2:165–172.
- in press "Which Way to Turn?" *Journal of Contemporary Ethnography*.
- , and L. Tillmann-Healy 1998 "Mucking Around Looking for Truth." In B. Montgomery and L. Baxter, eds., *Dialectical Approaches to Studying Personal Relationships*. Mahwah, N.J.: Lawrence Erlbaum.
- Chase, S. 1996. "Personal Vulnerability and Interpretive Authority in Narrative Research." In R. Josselson, ed., *Ethics and Process in The Narrative Study of Lives*, vol. 4. Thousand Oaks, Calif.: Sage.
- Charmaz, K. 1990 "'Discovering' Chronic Illness: Using Grounded Theory." *Sociology of Health and Illness* 30:1161–1172.
- Chavez, D. 1985 "Perpetuation of Gender Inequality: A Content Analysis of Comic Strips." *Sex Roles* 13: 93–102.
- Clandinin, D. J., and F. M. Connelly 1994 "Personal Experience Methods." In N. K. Denzin and Y. S. Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Collins, P. H. 1991 *Black Feminist Thought*. New York: Routledge.
- Crites, S. 1971 "The Narrative Quality of Experience." *Journal of the American Academy of Religion* 39:291–311.
- Crotty, M. 1998 *The Foundations of Social Research*. London: Sage.
- Denzin, N. 1989 *Interpretive Biography*. Newbury Park, Calif.: Sage.
- 1997 *Interpretive Ethnography: Ethnographic Practices for the 21st Century*. Thousand Oaks, Calif.: Sage.
- , and Y. Lincoln 1995 "Transforming Qualitative Research Methods: Is It a Revolution?" *Journal of Contemporary Ethnography* 24:349–358.
- DeVault, M. L. 1990 "Talking and Listening from Women's Standpoint: Feminist Strategies for Interviewing and Analysis." *Social Problems* 37:96–116.
- Ellingson, L. L. 1998 "'Then You Know How I Feel': Empathy, Identification, and Reflexivity in Fieldwork." *Qualitative Inquiry* 4:492–514.
- , and P. M. Buzzanell 1999 "Listening to Women's Narratives of Breast Cancer Treatment: A Feminist Approach to Patient Satisfaction with Physician-Patient Communication." *Health Communication* 11:153–183.
- Ellis, C. 1991 "Sociological Introspection and Emotional Experience." *Symbolic Interaction* 14:23–50.
- 1995a "Emotional and Ethical Quagmires in Returning to the Field." *Journal of Contemporary Ethnography* 24:711–713.
- 1995b *Final Negotiations: A Story of Love, Loss, and Chronic Illness*. Philadelphia: Temple University Press.
- 1997 "Evocative Autoethnography: Writing Emotionally about Our Lives." In W. Tierney and Y. Lincoln, eds., *Representation and the Text: Reframing the Narrative Voice*. Albany, N.Y.: SUNY Press.
- , and A. Bochner in press "Autoethnography, Personal Narrative, Reflexivity: Researcher as Subject." In N. K. Denzin and Y. S. Lincoln, eds., *Handbook of Qualitative Research*, 2nd ed. Thousand Oaks, Calif.: Sage.
- Ellis, C., K. E. Kiesinger, and L. M. Tillmann-Healy 1997 "Interactive Interviewing: Talking about Emotional Experience." In R. Hertz, ed., *Reflexivity and Voice*. Thousand Oaks, Calif.: Sage.
- Fine, G. 1992 "Wild Life: Authenticity and the Human Experience of 'Natural' Places." In C. Ellis and M. Flaherty, eds., *Investigating Subjectivity: Research on Lived Experience*. Newbury Park, Calif.: Sage.
- Fine, M. 1994 "Working the Hyphens: Reinventing Self and Other in Qualitative Research." In N. K. Denzin and Y. S. Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Fitch, K. L. 1994 "Criteria for Evidence in Qualitative Research." *Western Journal of Communication* 58:32–38.
- Fontana, A., and J. Frey 1994 "Interviewing: The Art of Science." In N. K. Denzin and Y. S. Lincoln, eds.,

- Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Frank, A. 1995 *The Wounded Storyteller: Body, Illness, and Ethics*. Chicago: University of Chicago Press.
- Geertz, C. 1973 *The Interpretation of Cultures: Selected Essays*. New York: Basic Books.
- Gergen, K. J. 1994 *Realities and Relationships: Soundings in Social Construction*. Cambridge: Harvard University Press.
- Gergen, M. M. 1992 "Life Stories: Pieces of a Dream." In G. C. Rosenwald and R. L. Ochberg, eds., *Storied Lives: The Cultural Politics of Self-Understanding*. New Haven, Conn.: Yale University Press.
- Glaser, B. 1978 *Theoretical Sensitivity*. Mill Valley, Calif.: Sociological Press.
- , and B. Strauss 1964 *Awareness of Dying*. Chicago: Aldine.
- 1967 *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Glesne, C. 1997 "That Rare Feeling: Re-Presenting Research through Poetic Transcription." *Qualitative Inquiry* 3:202–221.
- Jackson, M. 1989 *Paths toward a Clearing: Radical Empiricism and Ethnographic Inquiry*. Bloomington: Indiana University Press.
- Janesick, V. J. 1994 "The Dance of Qualitative Research Design: Metaphor, Methodology, and Meaning." In N. K. Denzin and Y. S. Lincoln, eds., *Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Jones, S. H. 1998 *Kaleidoscope Notes: Writing Women's Music and Organizational Culture*. Walnut Creek, Calif.: AltaMira.
- Josselson, R. 1996. "On Writing Other People's Lives: Self-Analytic Reflections of a Narrative Researcher." In R. Josselson, ed., *Ethics and Process in the Narrative Study of Lives*, vol. 4. Thousand Oaks, Calif.: Sage.
- Kitzinger, J. 1994 "The Methodology of Focus Groups: The Importance of Interaction between Research Participants." *Sociology of Health and Illness* 16:103–121.
- Krieger, S. 1991 *Social Science and the Self: Personal Essays on an Art Form*. New Brunswick, N.J.: Rutgers University Press.
- Lather, P. 1986 "Research as Praxis." *Harvard Educational Review* 56:257–277.
- , and C. Smithies 1997 *Troubling the Angels: Women Living with HIV/AIDS*. Boulder, Colo.: Westview.
- Lincoln, Y. S., and E. Guba 1985 *Naturalistic Inquiry*. Beverly Hills, Calif.: Sage.
- Lindlof, T. R. 1995 *Qualitative Communication Research Methods*. Thousand Oaks, Calif.: Sage.
- Lofland, J., and L. H. Lofland 1995 *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*, 3rd ed. Belmont, Calif.: Wadsworth.
- Lowry, D. T., and D. E. Towles 1989 "Soap Opera Portrayals of Sex, Contraception, and Sexually Transmitted Diseases." *Journal of Communication* 39:76–83.
- Mienczakowski, J. 1996 "An Ethnographic Act: The Construction of Consensual Theatre." In C. Ellis and A. P. Bochner, eds., *Composing Ethnography: Alternative Forms of Qualitative Writing*. Walnut Creek, Calif.: AltaMira.
- Miles, M., and A. M. Huberman 1984 *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills, Calif.: Sage.
- Mishler, E. G. 1986 *Research Interviewing: Context and Narrative*. Cambridge, Mass.: Harvard University Press.
- Mitchell, R. 1998 "Thirteen Stars and Chocolate Cake: Reasoning with Apocalypse," Paper presented to the Society for the Study of Symbolic Interaction, San Francisco.
- Neuman, W. L. 1997 *Social Research Methods: Qualitative and Quantitative Approaches*, 3rd ed. Boston: Allyn and Bacon.
- Potter, W. J., M. W. Vaughan, R. Warren, K. Howley, A. Land, and J. C. Hagemeyer 1995 "How Real Is the Portrayal of Aggression in Television Entertainment Programming?" *Journal of Broadcasting and Electronic Media* 39:494–516.
- Rabinow, P., and W. Sullivan 1987 *Interpretive Social Science: A Second Look*. Berkeley: University of California Press.
- Reed-Danahay, D. 1997 *Auto-Ethnography: Rewriting the Self and the Social*. Oxford, England: Berg.
- Reinharz, S. 1992 *Feminist Methods in Social Research*. New York: Oxford University Press.
- Richardson, L. 1992 "The Consequences of Poetic Representation." In C. Ellis and M. G. Flaherty, eds., *Investigating Subjectivity: Research on Lived Experience*. Newbury Park, Calif.: Sage.
- 1994 "Nine Poems." *Journal of Contemporary Ethnography* 23:3–13.
- 1997 *Fields of Play: Constructing an Academic Life*. New Brunswick, N.J.: Rutgers University Press.
- in press "Writing: A Method of Inquiry." In N. Denzin and Y. Lincoln, eds., *Handbook of Qualitative Research*, 2nd ed. Thousand Oaks, Calif.: Sage.



- Riessman, C. K. 1993 *Narrative Analysis. Qualitative Research Methods*, vol. 30. Newbury Park, Calif.: Sage Publications.
- Ronai, C. R. 1995 "Multiple Reflections of Child Sex Abuse: An Argument for a Layered Account." *Journal of Contemporary Ethnography* 23:395-426.
- Rorty, R. 1982 *Consequences of Pragmatism (Essays 1972-1980)*. Minneapolis: University of Minnesota Press.
- Snow, D., and C. Morrill 1995 "A Revolutionary Handbook or a Handbook for Revolution?" *Journal of Contemporary Ethnography* 24:341-348.
- Spence, D. 1982 *Narrative Truth and Historical Truth*. New York: Norton.
- Spitzack, C., and K. Carter 1989 *Doing Research on Women's Communication: Perspectives on Theory and Method*. Norwood, N.J.: Ablex.
- Stake, R. 1994 "Case Studies." In N. K. Denzin and Y. S. Lincoln, eds., *The Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Strauss, A., and J. Corbin 1994 "Grounded Theory Methodology: An Overview." In N. K. Denzin and Y. S. Lincoln, eds., *The Handbook of Qualitative Research*. Thousand Oaks, Calif.: Sage.
- Tedlock, B. 1991 "From Participant Observation to the Observation of Participation: The Emergence of Narrative Ethnography." *Journal of Anthropological Research* 41:69-94.
- Tompkins, P. K. 1994 "Principles of Rigor for Assessing Evidence in 'Qualitative' Communication Research." *Western Journal of Communication* 58:44-50.
- Van Maanen, J. 1988 *Tales of the Field: On Writing Ethnography*. Chicago: University of Chicago Press.
- Waitzkin, H. 1990 "On Studying the Discourse of Medical Encounters: A Critique of Quantitative and Qualitative Methods and a Proposal for Reasonable Compromise." *Medical Care* 28:473-487.
- Whyte, W. F. (1943) 1993 *Street Corner Society: The Social Structure of an Italian Shum*, 5th ed., Chicago: University of Chicago Press.
- Wilkinson, S. 1998 "Focus Groups in Feminist Research: Power, Interaction and the Co-Construction of Meaning." *Women's Studies International Forum* 21:111-125.
- Wimmer, R. D., and J. Dominick 1997 *Mass Media Research: An Introduction*, 5th ed. Belmont, Calif.: Wadsworth.
- Wolcott, H. 1995 *The Art of Fieldwork*. Walnut Creek, Calif.: AltaMira.

CAROLYN ELLIS  
LAURA ELLINGSON

## QUALITATIVE MODELS

Qualitative models describe structure and metamorphoses among things or events or among properties of things or events. Sociologists have several ways of formulating qualitative models.

Qualitative modeling based on *logic* involves the following ideas. Propositions are simple sentences such as "All humans are mortal" and "A dictator is a human." Propositions can be true or false, and negation of a proposition transforms truth into falsity, or falsity into truth. Compound statements are formed when two or more propositions are placed in disjunction or conjunction, signified in English by the words *or* (or *nor*) and *and* (or *but*). Compound statements are true if all their component propositions are true, and compound statements are false if all their component propositions are false. Disjunction of true and false propositions yields a compound statement that is true, whereas conjunction of true and false propositions yields a compound statement that is false. These definitions are sufficient for logical analyses, but a supplementary definition is useful: the conditional "P implies Q," or "If P, then Q," means that whenever proposition P is true, proposition Q is true also, but when P is false, Q may be either true or false.

Set theory corresponds closely with logic, to the point that logic formulations can be interpreted in terms of sets, and information about the existence of elements in sets and subsets can be interpreted in terms of logic. Logic also can be translated to Boolean algebra (which operates as does ordinary algebra except that there are only two numbers, 0 and 1, and  $1 + 1 = 1$ ), so any formulation in terms of logic can be transformed to an algebraic problem and processed mathematically.

Logic models have been used to define sociological constructs. Balzer (1990), for example, employed logic plus some additional mathematical ideas in order to construct a comprehensive definition of social institution. Logic models also can be used to compare competing sociological theories. Hannan (1998), for example, formalized different "stories" about how organizational age relates to organizational demise, and he used a computer program for automated deduction to

prove that various empirical observations can be derived from different theoretical assumptions.

Znaniecki (1934) systematized *analytic induction* as a method for deriving logic models from statements known to be true as a result of sociological research. For example (alluding to a study by Becker [1953] that applied the method), field research might have disclosed a set of fourteen males who are marijuana users, all of whom were taught to enjoy the drug; a set of three females who use marijuana though they were never taught to enjoy it; and a set of six males who were taught how to enjoy marijuana, but who do not use it. Implicitly it is understood that other people were never taught to enjoy marijuana and do not use it. From this information one might conclude that for males like the ones who were studied, using marijuana implies being taught to enjoy the drug. Robinson's critique of analytic induction (1951) led to a hiatus in the development of logic models in sociology until modeling difficulties were understood better.

Ragin (1988) developed a method for constructing logic models from cross-sectional data. Empirically valid propositions about all cases in a population are conjoined into a complex compound statement, transformed into Boolean algebra format, and processed by a computer program. The result is a reduced compound statement that is empirically true for the cases and the propositions studied. The approach differs from statistical analysis of multifold tables in ignoring count information (other than whether a cell in a table has zero cases or more than zero cases), and in describing data patterns in terms of logic statements rather than in terms of the effects of variables and their interactions.

Abell (1987) and Heise (1989) developed a logic model approach for *event sequence analyses*. Logic models for sequences do not predict what will happen next but instead offer developmental accounts indicating what events must have preceded a focal event. A narrative of events is elicited from a culturally competent consultant who also defines prerequisites of the events in terms of other events within the happening. Since prerequisites define implication relations, a logic model is obtained that accounts for sequencing of events within the happening and that can be tested as a possible explanation of event sequencing in other happenings. Routines that appear to have little

surface similarity may be accountable by abstract events in a logic model; for instance, Corsaro and Heise (1990) showed that an abstract model accounted for observed play routines among children in two different cultures. Abell (1987) suggested that abstraction involves homomorphic reduction: That is, abstract events categorize concrete events that have identical logical relations with respect to events outside the category. Abbott (1995) reviewed logic models and other approaches to sequence analysis.

Careers are sequences in which the events are status transformations. Heise's logic model analysis of careers (1990) emphasized that individuals' sequences of status transformations are generated in limited patterns from institutional taxonomies of roles. *Guttman scaling* can be employed as a means of analyzing individual experiences in order to infer logic models that generate career sequences (e.g., see Wanderer 1984). Abbott and Hrycak (1990) applied *optimal matching* techniques to the problem of comparing career sequences, with the similarity of two sequences being measured as the minimum number of transformations required to change one sequence into the other; clusters of similar sequences discovered from the similarity measures are identified as genres of career patterns.

A *formal grammar* defines sequences of symbols that are acceptable in a language, being "essentially a deductive system of axioms and rules of inference, which generates the sentences of a language as its theorems" (Partee et al. 1990, p. 437). A grammar, like a logic model, is explanatory rather than predictive, interpreting why a sequence was constructed as it was or why a sequence is deviant in the sense of being unprincipled. Grammars have been applied for modeling episodes of social interaction, viewing sequences of social events as symbolic strings that are, or are not, legitimate within a language of action provided by a social institution (Skvoretz and Fararo 1980; Skvoretz 1984). The grammatical perspective on institutionalized action can be reformulated as a *production system* model in which a hierarchy of if-then rules defines how particular conditions instigate particular actions (Axten and Fararo 1977; Fararo and Skvoretz 1984).

*Case frame grammar* (Dirven and Radden 1987) deals with how syntactic position within a set of

symbols designates function. For example, syntactic positioning in a sentence can designate an event's agent, action, object, instrument, product, beneficiary, and location (e.g., "The locksmith cut the blank with a grinder into a key for the customer in his shop"). Heise and Durig (1997) adapted case frame grammar to define an *event frame* for theoretical and empirical studies of social routines. The case-grammar perspective also informed Heise's (1979) symbolic interactionist modeling of social interaction by providing an agent-action-object-location framework for analyzing social events. Guttman's *facet mapping sentences* (see Shye 1978) implicitly employ a case grammar framework for analyzing a conceptual domain in terms of sets of concepts that fit into different syntactic slots and thereby generate a large number of propositions related to the domain. For example, Grimshaw (1989) developed a complex mapping sentence that suggested how different kinds of ambiguities arise in conversation and are resolved as a function of a variety of factors.

The mathematics of *abstract groups* provide a means for modeling some deterministic systems. Suppose a few different situations exist, and combining any two situations establishes another one of the situations; the result of a string of combinations can be computed by combining adjacent situations two at a time in any order. Also suppose that any situation can be reproduced by combining it with one particular situation, and this identity situation can be obtained from any other situation through a single combination. Then the set of situations and the scheme for combining them together constitute a group, and the group describes a completely deterministic system of transformations. Kosaka (1989) suggested a possible application of abstract groups by modeling the aesthetic theory of a Japanese philosopher in which there are sixty-four defined transformations, such as "yabo" (rusticity) combines with "hade" (flamboyance) to produce "iki" (chic urbanity).

A classic sociological application of groups involved kinship. Classificatory kinship systems (which are common in aboriginal cultures) put every pair of people in a society into a kinship relationship that may have little relation to genetic closeness, and each person implicitly is in a societywide kinship class that determines relationships with others. White (1963) showed through

mathematical analysis that classificatory rules regarding marriage and parentage generate clans of people who are in the same kinship situation and that the resulting classificatory kinship system operates as an abstract group; then he demonstrated that existing kinship systems accord with analytic results.

Models of social networks sometimes employ the notion of *semigroup*—a set of situations and a scheme for combining them (i.e., a group without an identity situation). For example, Breiger and Pattison (1986) examined economic and marriage relations among elite families in fifteenth-century Florence and showed that each family's relations to other families constituted a semigroup that was part of the overall semigroup of family relations in the city; they were able to identify the allies and enemies of the famous Medici family from the structure of family relationships. Social network research, a sophisticated area of qualitative modeling in sociology, employs other algebraic and graph-theoretic notions as well (Marsden and Laumann 1984; Wasserman and Faust 1994).

In general, qualitative models describe systematic structures and processes, and developing qualitative models aids in interpreting nebulous phenomena. Creating and manipulating qualitative models confronts researchers with technical challenges, but software providing computer assistance is lessening the difficulties.

#### REFERENCES

- Abbot, Andrew 1995 "Sequence Analysis: New Methods for Old Ideas." *Annual Review of Sociology* 21:93–113.
- , and Alexandra Hrycak 1990 "Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96:144–185.
- Abell, Peter 1987 *The Syntax of Social Life: The Theory and Method of Comparative Narratives*. New York: Oxford University Press.
- Axten, N., and Thomas J. Fararo 1977 "The Information Processing Representation of Institutionalized Social Action." In P. Krishnan, ed., *Mathematical Models of Sociology*. Keele, United Kingdom: University of Keele Press.
- Balzer, Wolfgang 1990 "A Basic Model for Social Institutions." *Journal of Mathematical Sociology* 16:1–29.
- Becker, Howard S. 1953 "Becoming a Marijuana User." *American Journal of Sociology* 59:235–243.

- Breiger, Ronald L., and Philippa E. Pattison 1986 "Cumulated Social Roles: The Duality of Persons and Their Algebras." *Social Networks* 8:215-256.
- Corsaro, William, and D. Heise 1990 "Event Structure Models from Ethnographic Data." In C. Clogg, ed., *Sociological Methodology: 1990*. Cambridge, Mass.: Basil Blackwell.
- Dirven, René, and Günter Radden (eds.) 1987 *Fillmore's Case Grammar: A Reader*. Heidelberg, Germany: Julius Groos Verlag.
- Fararo, Thomas J., and John Skvoretz 1984 "Institutions as Production Systems." *Journal of Mathematical Sociology* 10:117-182.
- Grimshaw, Allen D. 1989 *Collegial Discourse: Professional Conversation among Peers*. Norwood, N.J.: Ablex.
- Hannan, Michael T. 1998 "Rethinking Age Dependence in Organizational Mortality: Logical Formalizations." *American Journal of Sociology* 104:126-164.
- Heise, D. R. 1979 *Understanding Events: Affect and the Construction of Social Action*. New York: Cambridge University Press.
- 1989 "Modeling Event Structures." *Journal of Mathematical Sociology* 14:139-169.
- 1990 "Careers, Career Trajectories, and the Self." In J. Rodin, C. Schooler, and K. W. Schaie, eds., *Self-Directedness: Cause and Effects Throughout the Life Course*. New York: Lawrence Erlbaum.
- , and Alex Durig 1997 "A Frame for Organizational Actions and Macroactions." *Journal of Mathematical Sociology* 22:95-123.
- Kosaka, Kenji 1989 "An Algebraic Reinterpretation of IKI NO KOZO (Structure of IKI)." *Journal of Mathematical Sociology* 14:293-304.
- Marsden, Peter V., and Edward O. Laumann 1984 "Mathematical Ideas in Social Structure Analysis." *Journal of Mathematical Sociology* 10:271-294.
- Partee, B. H., A. ter Meulen, and R. E. Wall 1990 *Mathematical Methods in Linguistics*. Boston: Kluwer Academic.
- Ragin, Charles C. 1988. *Between Complexity and Generality: The Logic of Qualitative Comparison*. Berkeley: University of California Press.
- Robinson, W. S. 1951 "The Logical Structure of Analytic Induction." *American Sociological Review* 16:812-818.
- Shye, S. (ed.) 1978. *Theory Construction and Data Analysis in the Behavioral Sciences: A Volume in Honor of Louis Guttman*. San Francisco: Jossey-Bass.
- Skvoretz, John 1984 "Languages and Grammars of Action and Interaction: Some Further Results." *Behavioral Science* 29:81-97.
- , and Thomas J. Fararo 1980 "Languages and Grammars of Action and Interaction: A Contribution to the Formal Theory of Action." *Behavioral Science* 25:9-22.
- Wanderer, J. J. 1984 "Scaling Delinquent Careers Over Time." *Criminology* 22:83-95.
- Wasserman, Stanley, and Katherine Faust 1994 *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
- White, Harrison C. 1963 *An Anatomy of Kinship: Mathematical Models for Structures of Cumulated Roles*. Englewood Cliffs, N.J.: Prentice Hall.
- Znaniecki, Florian 1934 *The Method of Sociology*. New York: Farrar and Rinehart.

DAVID R. HEISE  
ALEX DURIG

## QUALITY OF LIFE

Although the concept of quality of life (QL) is not new, quality of life as an area of research and scholarship dates back only to the 1960s. Schuessler and Fisher (1985) noted that President Dwight Eisenhower's 1960 Commission on National Goals and Bauer's book on social indicators (1966) are often credited as providing the impetus for the development of QL as an area of research. Campbell (1981) suggested that the 1960s were favorable times for the development of QL research because of the emergence then of a belief that people must examine the quality of their lives and must do so in an environment that goes beyond providing material goods to foster individual happiness. Campbell quotes President Lyndon Johnson, who stated in 1964:

*The task of the Great Society is to ensure our people the environment, the capacities, and the social structures which will give them a meaningful chance to pursue their individual happiness. Thus the Great Society is concerned not with how much, but with how good—not with the quantity of goods but with the quality of their lives.* (Campbell 1981, p. 4)

Schuessler and Fisher (1985) note that the Russell Sage Foundation promoted QL and research on social indicators in the 1960s and 1970s and that the Institute for Social Research at the University of Michigan and the National Opinion Research Center at the University of Chicago have

conducted QL research since the late 1960s. Despite the high volume of QL research during the 1960s and 1970s, it was not until 1979 that “quality of life” became an index entry in *Sociological Abstracts*.

The emerging QL research in the 1970s provided a departure from previous work that focused on objective indicators, primarily economic in nature, of individual well-being. The book *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*, published by Campbell and colleagues in 1976, particularly promoted the use of subjective or psychological indicators of well-being. The work reported was founded on the conviction that the relationship between objective and subjective well-being indicators was weak and poorly understood. Moreover, the rising affluence of the post-World War II era had been accompanied by steady increases in social problems afflicting American society as well as other Western societies.

The year 1976 also saw the publication of another major work focusing on subjective indicators of well-being. *Social Indicators of Well-Being: Americans' Perceptions of Life Quality* by Andrews and Withy (1976) reported findings from interviews with representative samples of more than 5,000 Americans. The interviews focused on satisfaction with the quality of various life domains. A more recent volume, titled *Research on the Quality of Life* and edited by Frank Andrews (1986), brought together a variety of papers originating at a symposium honoring the memory of Angus Campbell, one of the founders of the Institute for Social Research. Although this volume included important papers on cross-national differences in life satisfaction and papers on African-Americans and Hispanics, a number of the papers had no direct relationship to QL research. Rockwell (1989) noted that a useful focus of the field was lost in this volume, the focus on subjective indicators of the quality of life. Andrews also noted that support for large-scale, wide-ranging surveys had become increasingly difficult in the 1980s in the United States, resulting in a lack of replication of the national surveys conducted in the previous decade by the Institute for Social Research.

Parallel to the large national surveys of subjective well-being during the 1970s, there was a proliferation of studies focusing on the subjective well-being of the elderly. In a useful article, Larson

(1978) reviewed three decades of research that focused on the psychological well-being of older people. Perhaps no other area of research in the emerging field of gerontology had received as much attention during the 1960s and 1970s as the area of life satisfaction, morale, mental health, and psychological well-being in general. Much of this research was spurred by the lively debate over the merits of disengagement theory (proposed by Cumming and Henry 1961) and activity theory (identified with various authors, including Havighurst et al. 1968; Maddox 1968, 1970) in predicting “successful aging.” Gerontological work in the 1980s showed a marked decline in the number of articles predicting life satisfaction and morale and an increase in articles focusing on specific dimensions of psychological well-being, such as depression and psychological distress, positive and negative affect (Lawton 1996), as well as articles focusing on the prediction of physical health outcomes (Markides 1989).

An exception to the general decline of sociological studies of QL in the 1980s was a study by Thomas and Hughes of racial differences in QL in the United States (1986), in which they found significantly lower subjective well-being among African-Americans than among whites over the period 1972–1984. In their recent extension of their work to 1996 using data from the General Social Survey, Hughes and Thomas (1998) found that African-Americans continue to have a lower subjective QL than whites, as expressed in terms of happiness, life satisfaction, marital happiness, and self-rated health. This was in contrast to other recent work that challenged the notion that African-Americans had lower subjective well-being. However, much of this work had examined such indicators of QL as psychiatric disorders, including depression (Kessler et al. 1994; Williams et al. 1992). Moreover, one analysis did not find that race magnified the negative effect of socioeconomic status on psychiatric disorders (Williams et al. 1992). Hughes and Thomas conclude that their findings suggest that their measures of subjective QL capture separate dimensions of life rather than measures of psychiatric disorder and that African Americans in all social classes express lower QL of “social life experience” (1998, p. 792). This is an example of how different measures of QL can produce substantially different results.

The relative decline in research on the subjective QL of Americans in general, as well as on the subjective well-being of the elderly during the 1980s, was accompanied by a marked increase in QL research in medicine, which continued to accelerate during the 1990s, both in North America and in Europe. This development has included the publication of the massive *Quality of Life and Pharmacoeconomics in Clinical Trials* (Spilker 1996), the journal *Quality of Life Research*, the *Quality of Life Newsletter*, and the establishment of the *International Society for Quality of Life Research*. In medicine, as in the social sciences, the field of QL is conceptually weak. As Leplège and Hunt (1997) recently argued, a problem has been the overwhelming emphasis of the medical model on function and health status at the expense of attention to social and psychological aspects of QL as expressed by patients themselves.

Within medicine, there has been particular interest in studying the quality of life of cancer patients. Before 1970, cancer research focused almost exclusively on survival and life extension. With extended survival from cancer becoming the rule, research has given increasing attention to the quality of life of the surviving patients afflicted with cancer or patients treated for cancer. In 1987, for example, a volume entitled *The Quality of Life of Cancer Patients* was published. The volume, edited by Aaronson and Beckman (1987), contains papers from researchers in a number of European countries as well as the United States. More recently, Gotay and Muraoka (1998) reviewed thirty-four studies published in English language journals from 1980 to 1998 on the quality of life of long-term survivors of adult-onset cancers.

Another parallel to this work has been the focus on active life expectancy. The work has gone beyond predicting extension of life in general to investigating the extent to which recent extensions of life expectancy have been accompanied by extensions of "active" life. A recent variant of this work is the concept of health expectancy or the proportion of life expectancy that consists of healthy years (Olshansky and Wilkins 1998).

#### DEFINITIONS OF QUALITY OF LIFE

As seen in the previous section, there has been a movement in recent decades away from objective,

quantitative research and toward subjective, qualitative assessments of QL in sociology and other fields. Even within these broad approaches to QL, there appears to be little agreement about an appropriate definition of QL.

Some writings include under QL research the social indicators movement. Land (1971) noted that in the early years of the movement, the most popular definition of social indicators was given in *Toward a Social Report*:

*A social indicator . . . may be defined to be a statistic of direct normative interest which facilitates concise, comprehensive and balanced judgements about the condition of a major aspect of a society. It is in all cases a direct measure of welfare and is subject to the interpretation that, if it changes, in the "right" direction, while other things remain equal, things have gotten better, or people are "better off." Thus statistics on the number of doctors or policemen could not be social indicators whereas figures on health or crime rates could be. (U.S. DHEW 1969, p. 97)*

Land criticized the above definition and proposed a broader one that treats social indicators as both "outputs" and "inputs" in "a sociological model of a social system or some segment thereof" (1971, p. 324). Thus, for example, the number of doctors is essential to understanding the health of the population, as are other factors. Land's definition has been largely accepted by the social indicators movement (Mukherjee 1989, p. 53).

This article gives only limited attention to social indicators, because a separate article is devoted to the topic. Yet the term "social indicators" is often used interchangeably with "quality of life," at least with respect to what Mukherjee calls "need-based" quality of life research (1989, p. 49). Moreover, the journal *Social Indicators Research* is subtitled *An International Journal of Quality of Life Measurement*.

In his book *The Quality of Life Valuation in Social Research*, Mukherjee notes that QL researchers employ several dichotomies, such as "quantity" and "quality," "behavior" and "perception," and "objective" and "subjective" indicators. He argues:

*Economists and planners . . . are almost exclusively concerned with behavioral research*

*on the basis of quantitative variables to improve the quality of life of the people. In that context, they ignore qualitative variations in the appraisal of a better quality of life or treat these variations as introducing a classificatory . . . distinction in the field of enquiry. They also equate the individual-wise subjective perception of reality to a group-wise "objective" perception by experts. Their appraisal of social reality in this manner leads them to formulate what the people need in order to improve their quality of life. (1989, pp. 37-38).*

The dependent variables of this research tend to be items or scales measuring satisfaction or happiness. Milbrath, for example, argues: "I have come to the conclusion that the only defensible definition of quality of life is a general feeling of happiness" (1978, p. 36). Even though such global evaluations have been common, much of the research has focused on describing and explaining satisfactions with various life "domains" such as work, family, and housing.

In discussing subjective indicators of QL, Land notes the difficulties in relating them to objective indicators. He notes, for example, that while income levels tend to be associated with satisfaction and happiness within given countries and at given times, "higher per capita levels of national income do not produce higher average levels of national satisfaction over time or cross sectionally" (1983, p. 5). He goes on to suggest that from the standpoint of generating theory of social change, it is not clear that satisfaction indexes provide an unambiguous criterion for the formulation of public policy.

According to O'Boyle (1997), a more focused definition of QL favored by those in the health field has its origins in the original definition of health by the World Health Organization (WHO) that emphasized the presence of social, mental, and physical well-being instead of only focusing on the absence of disease. The WHO Quality of Life Group offered the following definition:

*Quality of life is defined as the individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by a person's physical health, psychological*

*state, and level of independence and their relationships to salient features of their environment. (WHOQoL Group 1993, p. 5)*

### MEASURING QUALITY OF LIFE

The broadest and most commonly employed distinction in measures of QL is between objective and subjective measures. Among the former are indicators such as per capita income, average calorie consumption, percent of adult illiteracy, quality of air, average daily temperature, crime rates, life expectancy, and a myriad of other indicators that are best seen as causes of quality of life.

Any one of the above has shortcomings. For example, gross national product (GNP) per capita has been acknowledged to suffer from many well-known limitations, including that it may not capture the spending power of the masses but rather that of a small minority (Mukherjee 1989, p. 42). To overcome the limitations of single indicators, researchers have proposed a number of composite indexes, such as the Physical Quality of Life Index (PQLI; see Morris 1977), which includes, among other variables, life expectancy at birth, infant mortality, and literacy. The purpose of the PQLI is to rank countries by physical well-being. Yet it has limitations, as its proponent acknowledges, including that "it is based on the assumption that the needs and desires of individuals initially and at the most basic level are for larger life expectancy, reduced illness, and greater opportunity" (Morris 1977, p 147).

Another composite index of objective indicators of QL is the Index of Social Progress (ISP) proposed originally by Estes (1984) and revised more recently by the same author (Estes 1988). The latest version (ISP83) consists of thirty-six indicators and is divided into ten subindexes covering "education, health status, status of women, defense effort, economic, demographic, political participation, cultural diversity and welfare effort" (Estes 1988, p. 1). A number of equally important indicators (e.g., crime rates, suicide rates, maldistribution of wealth) were not included because reliable data were not available on the 124 nations studied.

There has also been a lively interest in developing indexes consisting of objective indicators to rank quality of life of American cities on the basis

of such domains of QL as economic, environmental, health, education, social, and political. Such rankings of cities elicit national attention and often surprise individuals about how high or low their community ranks. Rankings also do not often correlate with each other. For example, Berger and colleagues (1987) found that their revealed-preference rankings had a correlation of  $-0.075$  with those proposed by Boyer and Savageau (1981) and a correlation of  $0.048$  with Liu's rankings (1976).

There have been numerous subjective measures of QL, with most relating to happiness or life satisfaction. Some measures are global in the sense that they aim at capturing happiness or satisfaction with life as a whole, while others pertain to happiness or satisfaction with certain life domains. The studies by Andrews and Withy (1976) and by Campbell and colleagues (1976) include measures of both domain-specific and global life satisfaction and employ the former as predictors of the latter. In general, they find that the best predictors of global satisfaction are marriage and family life, leisure activities, work and finances, housing, the community, and friendships.

Well before these landmark studies, W. Wilson (1967) reviewed prior literature on subjective well-being and concluded that the "happy person emerges as a young, healthy, well-educated, well-paid, extroverted, optimistic, worry-free, religious, married person with high self-esteem, high job morale, modest aspirations, of either sex and of a wide range of intelligence" (1967, p. 294). He also concluded that little progress had been made in understanding happiness since the time of the Greek philosophers. Diener (1984) noted that between W. Wilson's 1967 article and 1984, over seven hundred studies on subjective well-being had been published. In general, Wilson's conclusions regarding predictors of well-being appeared to be supported by the literature, including that little theoretical progress had been made in the field since the ancient Greeks.

This voluminous literature on subjective well-being has employed a variety of single-item and multiple-item measures of happiness and life satisfaction. Among the best-known single-item measures are: Cantril's "self-anchoring ladder" (1965), which asks respondents to place themselves on a nine-rung ladder ranging from "best possible for

you" to "worst possible for you"; Gurin and colleagues's item (1960), "Taking all things together, how would you say things are these days?" with possible response choices being "very happy," "pretty happy," and "not too happy"; and Andrews and Withy's item (1976), "How do you feel about how happy you are?" with seven choices ranging from "delighted" to "terrible."

A problem with single-item measures is that because internal reliability estimates cannot be computed, the only way of assessing their reliability is through temporal correlation, which makes it difficult to separate measurement error from true change. However, convergence with other measures of well-being has suggested that these single-item measures enjoy moderate levels of validity. They do suffer from other limitations, however, such as positive skewness, acquiescence, and inability to capture the various dimensions of well-being (Diener 1984).

There have also been a variety of multi-item scales employed. Some of the best-known general scales include: the Affect Balance Scale (Bradburn 1969), which consists of ten items capturing positive and negative well-being. Respondents are asked whether in the past few weeks they felt: "particularly excited or interested in something," "so restless you couldn't sit long in a chair," "proud because someone complimented you on something you had done," "very lonely or remote from other people," "pleased about having accomplished something," "bored," "on top of the world," "depressed or very unhappy," "that things were going your way," and "upset because someone criticized you." Summing the positive items provides a positive score and summing the negative ones a negative score. An "affect balance score" is obtained by subtracting the negative score from the positive score. The two subscales have been found to be relatively independent of each other and are sometimes used as different scales of positive and negative affect.

Another multi-item scale is Campbell and colleagues' Index of General Affect (1976), which asks respondents to describe their present lives using semantic differential scales (miserable-enjoyable, hard-easy, boring-interesting, useless-worthy, lonely-friendly, discouraging-hopeful, empty-full, disappointing-rewarding, and doesn't give me a chance-brings out the best in me).



Although happiness and satisfaction are often used interchangeably, many writers believe they are distinct measures of well-being. George, for example, suggests that "happiness refers to an affective assessment of quality of life," while "life satisfaction refers to an assessment of the overall conditions of life, as derived from a comparison of one's aspirations to one's actual achievements" (1981, p. 351). Campbell and colleagues (1976) prefer satisfaction measures over happiness measures because they are more sensitive to intervention. While happiness tends to be transitory and volatile, life satisfaction changes gradually and systematically in response to changing life conditions (see also Stull 1987). Satisfaction scales have been particularly popular in gerontology.

### QUALITY OF LIFE IN THE ELDERLY

An area in which lively interest has been shown in subjective indicators of QL has been the field of gerontology. As mentioned earlier, use of subjective measures of well-being was particularly high during the 1960s and 1970s, when social gerontologists were occupied with assessing the merits of disengagement and activity theories. In the late 1970s, Larson (1978) reviewed three decades of research and concluded that the most consistent predictors of subjective well-being are self-reports of health.

Although gerontological studies have employed general well-being measures (e.g., the Affect Balance Scale), they have also employed scales specifically developed for use with older people. The two best known are the Life Satisfaction Index A (Neugarten et al. 1961) and the Philadelphia Geriatric Morale Scale (Lawton 1975). The Life Satisfaction Index A consists of twenty items, with which respondents indicate agreement or disagreement. A combined life satisfaction score is obtained by summing scores on all twenty items. Twelve items are positive (e.g., "I am just as happy as when I was younger," "I expect some interesting and pleasant things to happen to me in the future," "As I look back on my life, I am fairly well satisfied") and eight items are negative (e.g., "When I think back over my life, I didn't get most of the important things I wanted," "Most of the things I do are boring and monotonous," "Compared to other people, I get down in the dumps too often").

Because the index covers a variety of areas, including happiness, satisfaction, and "activation level" (see Cherlin and Reeder 1975), the combined score confounds separate dimensions of well-being (Stull 1987).

The Philadelphia Geriatric Center Morale Scale (PGCMS) originally consisted of twenty-two items (Lawton 1972), and the revised version consisted of seventeen items (Lawton 1975). Like the Life Satisfaction Index, the PGCMS consists of positive items (e.g., "I am as happy now as I was when I was younger," "As I get older things are better than I thought they would be," "I have as much pep as I did last year") and negative items (e.g., "Things keep getting worse as I get older," "I sometimes feel life is not worth living," "I sometimes worry so much I can't sleep"). Factor analyses have produced three dimensions: agitation, attitude toward own aging, and lonely dissatisfaction. The scale has problems similar to those of the Life Satisfaction Index, such as the confounding of satisfaction and happiness. The two scales are in many ways similar (in fact, they share some items) and have been found to be highly intercorrelated ( $r = 0.76$ ; see Lohman 1977).

Liang (1985) attempted to integrate the Life Satisfaction Index A and the Affect Balance Scale by selecting seven items from the former and eight from the latter. His analysis yielded four factors (congruence, happiness, positive affect, and negative affect) that correspond to dimensions of well-being discussed by Lawton (1983). However, Liang acknowledged a gap between the operationalization of well-being and its theoretical operationalization: "Most instruments were developed with only a general conceptual definition, and the sampling of the item domain is usually based on intuition, experience, and empirical experimentation" (Liang 1985, p. 553).

After reviewing the voluminous literature on subjective well-being among the elderly, Gubrium and Lynott (1983) concluded that it was time to "rethink life satisfaction" in old age. One of their key concerns was that the dominant measures employed tended to dwell on the earlier years of people's lives and have less relevance for their current circumstances. In addition, "current measures do not allow for co-equal dialogue between subject and researcher about the content of items and responses" (Gubrium and Lynott 1983, p. 37).

Possibly because of these and other conceptual and methodological problems in subjective well-being measures, we have seen a substantial decline in published studies in major journals aiming at predicting life satisfaction, morale, and related concepts during the 1980s and 1990s. Social gerontologists have instead concentrated on predicting more narrow dimensions of well-being, such as psychological distress and depression (Ferraro and Su 1999; Lawton et al. 1999; Lawton 1996), and are increasingly employing a life-course perspective that involves examination of main and interactive effects of stress, social support, coping styles, and related factors (e.g., George 1989). Measures of depression and psychological distress are being employed more frequently, perhaps because they are perceived as more amenable to intervention than are measures of life satisfaction and morale.

A general conclusion of the field of subjective well-being in the elderly is that most elderly people, particularly those who have reasonably good health and finances, and are socially engaged, report relatively high levels of well-being. According to some literature, elderly people may even report higher levels of subjective QL than do people in younger age groups (Lawton 1996).

#### **STUDIES OF QUALITY OF LIFE IN MEDICINE**

Perhaps the most activity in the area of quality of life is currently found in medicine, much of it conducted by behavioral and social scientists. Interest in QL after medical treatments is based on the realization that chronic diseases cannot be cured, and, therefore, the goal of much therapy becomes to limit the effects of illness so that patients may live productive, comfortable, and satisfying lives. Traditionally, success of medical treatment was evaluated in terms of lengthening lives and post-treatment complications. However, there has been a realization that medical and surgical treatments may extend survival but often reduce quality of life (Eisman 1981).

Hollandsworth (1988) reviewed studies evaluating the impact of medical treatment on QL during the period 1980 to 1984 and compared his results with those of studies conducted during 1975 to 1979 (Najman and Levine 1981). Hollandsworth's comparison (1988) revealed a

marked increase between the two time periods in both quantity and quality of studies. Although recent studies tended to be more sophisticated, the majority nevertheless relied on convenience samples. One marked improvement in the recent research is the increase in use of subjective measures of quality of life, with 60 percent of the recent studies employing at least one such measure, compared to only around 10 percent in the earlier period.

Another interesting outcome of Hollandsworth's analysis (1988) was the increase over time in the proportion of studies that do not report favorable outcomes. Studies published in the late 1970s were almost unanimous in claiming favorable outcomes of treatment, but this optimism must be tempered by the many methodological limitations of these studies (Najman and Levine 1981). Of the more sophisticated studies published from 1980 to 1984, almost one-half reported either negative outcomes or at least mixed results. In fact, it appeared that the probability of reporting negative outcomes (or lack of positive results) tended to be correlated with the methodological sophistication of the studies (Hollandsworth 1988).

The impact of a variety of medical treatments have been examined, including cardiovascular therapies (e.g., Jenkins et al. 1983; Wenger et al. 1984), end-stage renal disease (e.g., Evans et al. 1985) and chronic obstructive pulmonary disease (e.g., McSweeney et al. 1982). However, by far the most frequently studied area is that relating to outcomes of cancer treatment (see Aaronson 1989; Aaronson and Beckman 1987; Cella and Cherin 1988). Aaronson noted that while "there is no universally accepted definition of the quality of life concept, in oncology it is most often used to describe such aspects of health status as physical symptoms, daily activity level, psychological well-being, and social functioning" (1989, p. 69). This increasing use of subjective QL indicators is becoming an integral part of evaluation in clinical cancer research, but a major challenge facing researchers is the development of measures capturing all dimensions of QL while meeting rigorous standards of reliability and validity (Aaronson 1989).

In a recent review, Gotay and Muraoka (1998) observed that recent studies of cancer survivors have increasingly been using standardized instruments relying primarily on self-reports to assess

QL. They identified thirty-four studies published in English-language journals during 1990 to 1998 that focused on the QL of patients surviving five or more years. A variety of standardized instruments were used measuring a variety of aspects of QL. As in other studies of patients, a popular instrument has been the Short-Form 36 Health Survey (SF-36) (Ware and Sherbourne 1992). The SF-36 is a 36-item short form of the Rand Corporation's Medical Outcomes Study. It was designed to be a general measure of health status that can be used in community studies as well as in studies of patients. It consists of multi-item scales measuring eight dimensions of health: general perceptions of health, physical functioning, activity limitations due to physical health problems, bodily pain, social functioning, psychological well-being, activity limitations due to emotional problems, and energy/vitality.

During the 1990s the SF-36 has become the instrument of choice in studies of health-related QL in both community surveys and studies of patients. Like other abbreviated instruments of general health status and QL, the SF-36 has been criticized on a number of grounds, including that it covers some areas only superficially (McDowell and Newell 1996, p. 454).

In their recent critical assessment of QL measurement in medicine, Leplège and Hunt (1997) have criticized the field for relying on instruments (like the SF-36) that have physical, emotional, and social functioning components. While these, according to the authors, are measuring health status, it is not clear that they are measuring QL. This overwhelming emphasis on function, they argue, ignores the person's perspective on the meaning and importance of functions and roles under study. They go on to cite evidence that many physically disabled people do indeed consider their QL to be high despite severe limitations (Leplège and Hunt 1997).

Leplège and Hunt (1997) are also critical of the economic model of QL as expressed in the notion of Quality Adjusted Life Years (QALYs). The concept has evolved over the years as a tool for health policy. Typically, the utility value of specific health status measures, including function and symptoms, during a given period of time is combined with survival data. This perspective usually

assumes that rational human beings, if given a choice, would prefer a shorter but relatively healthy life, to a longer life with serious discomfort and handicap. Leplège and Hunt argue that:

*the methods used for the valuation of health states do not encompass the fact that the same people value the same state differently at different times, while different people have different preferences that become meaningless if aggregated. The concept of utility . . . operatively addresses what people think they might (or should) do under artificial circumstances and not what they actually do in the real world. (1997, p. 48)*

An overview of the literature on medical treatment outcomes does indeed reveal increasing use of subjective QL measures. As in the broader QL field, many studies tend to employ single-item global indicators capturing life satisfaction or happiness. However, an increasing number of studies are employing the general multiple-item measures discussed earlier, such as the Affect Balance Scale and the Life Satisfaction Index Z. Other scales capturing more specific and narrow dimensions of QL include measures of mood, anxiety, self-concept, and depression (Lawton 1996) as well as more comprehensive instruments that capture physical, emotional, and social functioning, such as the McMaster Health Index Questionnaire (Chambers et al. 1982) and the SF-36. A general conclusion of much of the research is that most patients, including cancer patients, demonstrate an incredible capacity to cope with and adapt to the challenges of life-threatening disease and disability. A welcome recent addition to the field has been the development of a variety of disease-specific QL measures (Bowling 1995).

## CONCLUSION

This brief and selective overview of the field of quality of life indicates a variety of perspectives employed within sociology and in related fields. In fact, it may be said that there is more interest in QL outside the mainstream of sociology, as, for example, in the area of medical treatment. While much of the pioneer work and many of the large-scale national studies in the 1970s were conducted by sociologists, research on quality of life remains

very much outside the mainstream of sociology. For example, Schuessler and Fisher's review (1985) uncovered only one article (Gerson 1976) explicitly on quality of life, which was published in the *American Sociological Review* way back in 1976. More recently, we noted the work of Thomas and Hughes published in 1986 and 1998.

This overview also reveals some patterns and trends in QL research in the last four decades. First, there have been two broad approaches, one focusing on objective indicators and one focusing on subjective indicators. Related dichotomies noted by Mukherjee (1989) include quantity versus quality and behavior versus perception. It is clear that there has been a trend away from relying simply on objective indicators and toward relying increasingly on people's subjective reports about the quality of their lives. Objective measures have been the domain primarily of the social indicator movement, with subjective approaches to QL increasingly perceived as the domain of QL research.

Within the subjective QL approach, we also see a trend away from single-item indicators capturing global happiness and life satisfaction and toward multiple-item scales such as the Affect Balance Scale and the Life Satisfaction Index Z. At the same time, there have been attempts to measure subjective quality of life in specific life domains, and there has been continuing interest by sociologists, economists, and others (including popular magazines) in ranking urban areas according to a variety of objective QL indicators.

During the 1960s and 1970s a great deal of subjective QL research was conducted by social and behavioral gerontologists, who used measures of life satisfaction and morale as indicators of successful aging. For a number of reasons, gerontologists began abandoning research on life satisfaction and morale in favor of measures more amenable to intervention, such as measures of psychological distress, depression, and physical health function. Perhaps the most exciting research on QL currently being conducted is in the area of medical treatment outcomes, particularly cancer treatment. In this field, as in others, there is considerable disagreement about what constitutes quality of life and how it should be measured.

It is becoming increasingly difficult to obtain funding to conduct large-scale national surveys of

subjective quality of life such as those conducted during the 1970s. The future of QL research is uncertain, at least as a broad unified field of inquiry. Studies ranking urban areas are likely to continue, because of the immediate and broad appeal they elicit. It is also safe to predict that the concept of quality of life will continue to have some appeal in social gerontology. The most exciting work may well take place in the area of medical intervention outcomes. Sociologists and other behavioral scientists are increasingly conducting research related to medicine, and much of this research relates to quality of life. It is becoming apparent that medical interventions (as well as other factors) are enabling us to live longer, but it is not clear that the added years of life are "quality" years. There will be increasing interest in finding ways to improve the quality of the added years, and sociologists have an opportunity and a responsibility to help find ways of accomplishing that.

(SEE ALSO: *Social Indicators*)

REFERENCES

- Aaronson, Neil K. 1989 "Quality of Life: What Is It? How Should It Be Measured?" *Oncology* 2:69-74.
- , and J. H. Beckman (eds.) 1987 *The Quality of Life of Cancer Patients*. New York: Raven.
- Andrews, Frank M. (ed.) 1986 *Research on the Quality of Life*. Ann Arbor: University of Michigan, Institute for Social Research.
- Andrews, Frank M., and Stephen B. Withey 1976 *Social Indicators of Well-Being: Americans' Perceptions of Life Quality*. New York: Plenum.
- Bauer, Raymond A. (ed.) 1966 *Social Indicators*. Cambridge: Massachusetts Institute of Technology Press.
- Berger, Mark C., Glenn C. Blomquist, and Werner Waldner 1987 "A Revealed-Preference Ranking of Quality of Life for Metropolitan Areas." *Social Science Quarterly* 68:761-778.
- Bowling, Ann 1995 *Measuring Disease*. Buckingham, U.K.: Open University Press.
- Boyer, Rick, and D. Savageau 1981 *Places Rated Almanac*. Chicago: Rand McNally.
- Bradburn, Norman 1969 *The Structure of Psychological Well-Being*. Chicago: Aldine.
- Campbell, Angus 1981 *The Sense of Well-Being in America: Recent Patterns and Trends*. New York: McGraw-Hill.

- , Phillip Converse, and Willard L. Rogers 1976 *The Quality of American Life: Perceptions, Evaluations, and Satisfactions*. New York: Russell Sage Foundation.
- Cantril, Hadley 1965 *The Pattern of Human Concerns*. New Brunswick, N.J.: Rutgers University Press.
- Cella, David F., and E. A. Cherin 1988 "Quality of Life during and after Cancer Treatment." *Comprehensive Therapy* 14:69–75.
- Chambers, Larry W., Lorry A. MacDonald, Peter Tugwell, W. Watson Buchanan, and Gunnar Kraag 1982 "The McMaster Health Index Questionnaire as a Measure of Quality of Life for Patients with Rheumatoid Disease." *Journal of Rheumatology* 9:780–784.
- Cherlin, Andy, and Leo G. Reeder 1975 "The Dimensions of Psychological Well-Being: A Critical Review." *Sociological Methods and Research* 4:189–214.
- Cumming, Elaine, and William E. Henry 1961 *Growing Old: The Process of Disengagement*. New York: Basic.
- Diener, Ed 1984 "Subjective Well-Being." *Psychological Bulletin* 95:542–575.
- Eisman, B. 1981 "The Second Dimension." *Archives of Surgery* 116:11–13.
- Estes, Richard J. 1984 *The Social Progress of Nations*. New York: Praeger.
- 1988 *Trends in World Social Development: The Social Progress of Nations, 1970–1987*. New York: Praeger.
- Evans, R. W., D. L. Manninen, L. P. Garrison, L. G. Hart, C. R. Blagg, R. A. Gutman, A. R. Hull, and E. G. Lowrie 1985 "The Quality of Life of Patients with End-State Renal Disease." *New England Journal of Medicine* 312:553–559.
- Ferraro, Kenneth F., and Ya-ping Su 1999 "Financial Strain, Social Relations, and Psychological Distress among Older People: A Cross-Cultural Analysis." *Journal of Gerontology: Social Sciences* 54B:S3–S15.
- George, Linda K. 1981 "Subjective Well-Being: Conceptual and Methodological Issues." *Annual Review of Gerontology and Geriatrics* 2:345–382.
- 1989 "Stress, Social Support, and Depression over the Life Course." In Kyriakos S. Markides and Cary L. Cooper, eds., *Aging, Stress and Health*. Chichester, England: John Wiley.
- Gerson, Elihu M. 1976 "On 'Quality of Life'." *American Sociological Review* 41:793–806.
- Gotay, Carolyn C. and Miles Y. Muraoka 1998 "Quality of Life in Long-Term Survivors of Adult-Onset Cancers." *Journal of the National Cancer Institute* 90:656–667.
- Gubrium, Jaber F., and Robert J. Lynott 1983 "Rethinking Life Satisfaction." *Human Organization* 42:30–38.
- Gurin, Gerald, J. Veroff, and S. Feld 1960 *Americans View Their Mental Health*. New York: Basic.
- Havighurst, Robert J., Bernice Neugarten, and Sheldon S. Tobin 1968 "Disengagement and Patterns of Aging." In Bernice L. Neugarten, ed., *Middle Age and Aging*. Chicago: University of Chicago Press.
- Hollandsworth, James G. 1988 "Evaluating the Impact of Medical Treatment on the Quality of Life: A Five-Year Update." *Social Science and Medicine* 26:425–434.
- Hughes, Michael, and Melvin E. Thomas 1998 "The Continuing Significance of Race Revisited: A Study of Race, Class and Quality of Life in America, 1972 to 1996." *American Sociological Review* 63:785–795.
- Jenkins, C. David, Babette A. Stanton, J. A. Savageau, P. Denlinger, and M. D. Klein 1983 "Coronary Artery Bypass Surgery: Physical, Psychological, Social, and Economic Outcomes Six Months Later." *Journal of the American Medical Association* 250:782–788.
- Kessler, Ronald C., Katherine A. McGonagle, Shanyang Zhao, Christopher B. Nelson, Michael Hughes, Suzann Eshleman, Hans-Ulrich Wittchen, and Kenneth S. Kendler 1994 "Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States." *Archives of General Psychiatry* 51:8–19.
- Land, Kenneth C. 1971 "On the Definition of Social Indicators." *American Sociologist* 6:322–325.
- 1983 "Social Indicators." *Annual Review of Sociology* 9:1–26.
- Larson, Reed 1978 "Thirty Years of Research on Subjective Well-Being of Older Americans." *Journal of Gerontology* 33:109–125.
- Lawton, M. Powell 1972 "The Dimensions of Morale." In Donald Kent, Robert Kastenbaum, and Sylvia Sherwood, eds., *Research, Planning, and Action for the Elderly*. New York: Behavioral Publications.
- 1975 "The Philadelphia Geriatric Center Morale Scale" A Revision." *Journal of Gerontology* 30:85–89.
- 1983 "The Varieties of Well-Being." *Experimental Aging Research* 9:65–72.
- 1996 "Quality of Life and Affect in Later Life." In C. Magai and S. McFadden, eds. *Handbook of Emotion, Adult Development, and Aging*. Orlando, Fla.: Academic.
- , Laraine Winter, Morton H. Kleban, and Katy Ruckdeschel 1999 "Affect and Quality of Life: Objective and Subjective." *Journal of Aging and Health* 11:169–198.
- Leplège, Alain, and Sonia Hunt 1999 "The Problem of Quality of Life in Medicine" *Journal of the American Medical Association* 278:47–50.

- Liang, Jersey 1985 "A Structural Integration of the Affect Balance Scale and the Life Satisfaction Index A." *Journal of Gerontology* 40:552-561.
- Liu, Ben-Cheih 1976 *Quality of Life Indicators in U.S. Metropolitan Areas*. New York: Praeger.
- Lohman, Nancy 1977 "Correlations of Life Satisfaction, Morale, and Adjustment Measures." *Journal of Gerontology* 32:73-75.
- McDowell, Ian, and Claire Newell 1996 *Measuring Health: A Guide to Rating Scales and Questionnaires*, 2nd ed. New York: Oxford University Press.
- McSweeney, A. J., I. Grant, R. K. Heaton, K. Adams, and R. M. Timms 1982 "Life Quality of Patients with Chronic Obstructive Pulmonary Disease." *Archives of Internal Medicine* 142:473-478.
- Maddox, George L. 1968 "Persistence of Life Styles among the Elderly: A Longitudinal Study of Patterns of Social Activity in Relation to Life Satisfaction." In Bernice L. Neurgarten, ed., *Middle Age and Aging*. Chicago: University of Chicago Press.
- 1970 "Themes and Issues in Sociological Theories of Aging." *Human Development* 13:17-27.
- Markides, Kyriakos S. 1989 "Aging, Gender, Race/Ethnicity, Class, and Health: A Conceptual Overview." In K. S. Markides, ed., *Aging and Health: Perspectives on Gender, Race, Ethnicity, and Class*. Newbury Park, Calif.: Sage.
- Milbrath, L. W. 1978 "Indicators of Environmental Quality." In *Indicators of Environmental Quality and Quality of Life*. UNESCO Reports and Papers in the Social Sciences, No. 38. Paris: UNESCO.
- Morris, David M. 1977 "A Physical Quality of Life Index (PQLI)." In J. W. Sewell, ed., *The United States and the World Development Agenda 1977*. New York: Praeger.
- Mukherjee, Ramkrishna 1989 *The Quality of Life Valuation in Social Research*. New Delhi: Sage.
- Najman, J. M., and Sol Levine 1981 "Evaluating the Impact of Medical Care and Technologies on the Quality of Life: A Review and Critique." *Social Science and Medicine* 15(F):107-115.
- Neurgarten, Bernice L., Robert J. Havighurst, and Sheldon S. Tobin 1961 "The Measurement of Life Satisfaction." *Journal of Gerontology* 16:134-143.
- O'Boyle, Ciaran A. 1997 "Measuring the Quality of Later Life." *Philosophical Transactions of the Royal Society of London B*. 352:1871-1879.
- Olshansky, S. Jay, and Russell Wilkins (eds.) 1998 "Policy Implications of the Measures and Trends in Health Expectancy." *Journal of Aging and Health* 30(2).
- Rockwell, R. C. 1989 "Review of Research on Quality of Life." F. M. Andrews, ed., *Social Forces* 67:824-826.
- Schuessler, Karl F., and G. A. Fisher 1985 "Quality of Life Research in Sociology." *Annual Reviews of Sociology* 11:129-149.
- Spilker, B. (ed.) 1996 *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia: Lippincott-Raven.
- Stull, Donald E. 1987 "Conceptualization and Measurement of Well-Being: Implications for Policy Evaluation." In Edgar F. Borgatta and Rhonda J. V. Montgomery, eds., *Critical Issues in Aging Policy*. Newbury Park, Calif.: Sage.
- Thomas, Melvin E., and Michael Hughes 1986 "The Continuing Significance of Race: A Study of Race, Class, and Quality of Life in America, 1972-1985." *American Sociological Review* 51:830-841.
- U.S. Department of Health, Education, and Welfare 1969 *Toward a Social Report*. Washington, D.C.: U.S. Government Printing Office.
- Ware, John E., and C. D. Sherbourne 1992 "The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual Framework and Item Selection." *Medical Care* 30:473-483.
- Wenger, Nanette K., Margaret E. Mattson, Curt D. Furberg, and Jack Elinson (eds.) 1984 *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*. New York: Le Jacq.
- WHO QoL Group 1993 *Measuring Quality of Life: The Development of the World Health Organization Quality of Life Instrument*. Geneva: World Health Organization.
- Williams, David R., David T. Takeuchi, and Russell K. Adair 1992 "Socioeconomic Status and Psychiatric Disorder among Blacks and Whites." *Social Forces* 71:179-194.
- Wilson, Warner 1967 "Correlates of Avowed Happiness." *Psychological Bulletin* 67:294-306.

KYRIAKOS S. MARKIDES

## QUASI-EXPERIMENTAL RESEARCH DESIGNS

The goal of most social scientific research is to explain the causes of human behavior in its myriad forms. Researchers generally attempt to do this by uncovering causal associations among variables. For example, researchers may be interested in whether a causal relationship exists between income and happiness. One might expect a positive

association between these two variables. That is, an increase in income, the independent variable, produces an increase in happiness, the dependent variable. Unfortunately, observing a positive correlation between these two variables does not prove that income causes happiness. In order to make a valid causal inference, three conditions must be present: (1) there must be an association between the variables (e.g., income and happiness); (2) the variable that is the presumed cause (e.g., income) must precede the effect (e.g., happiness) in time; and (3) the association between the two variables cannot be explained by the influence of some other variable (e.g., education) that may be related to both of them. The purpose of any research design is to construct a circumstance within which a researcher can achieve these three conditions and thus make valid causal inferences.

*Experimental designs* are one of the most efficient ways to accomplish this goal of making valid causal inferences. Four characteristics are especially desirable in designing experiments. First, researchers *manipulate* the independent variable. That is, they actively modify persons' environment (e.g., provide some people with money they otherwise would not have received)—as contrasted with passively observing the existing, “natural” environment (e.g., simply measuring the amount of income persons normally make). Second, researchers have complete control over *when* the independent variable is manipulated (e.g., when persons receive supplementary income). Third, researchers have complete control over *what* they manipulate. That is, they can specify the exact content of the different “treatment conditions” (different levels) of the independent variable to which subjects are exposed (e.g., how much supplementary income persons receive and the manner in which they receive it). Fourth, researchers have complete control over *who* is assigned to which treatment condition (e.g., who receives the supplementary income and who does not, or who receives higher versus lower amounts of income).

Of these four characteristics important in designing experiments, only manipulation of the independent variable and control over who receives the treatments are essential to classify a study as a true experimental design.

Control over who receives treatment conditions is especially powerful in enhancing valid

causal inference when researchers use the technique of *random assignment*. For example, in evaluating the effect of income on happiness, investigators might randomly assign individuals who are below the poverty level to treatment groups receiving varying levels of supplementary income (e.g., none versus \$1,000).

Table 1 (a) illustrates this example. It depicts an experimental design in which subjects are randomly assigned to one of two groups. At time 1, researchers manipulate the independent variable ( $X$ ): Each subject in group 1 receives \$1,000 in supplementary income. Conversely, no subjects in group 2 receive any supplementary income. At time 2, researchers observe (measure) the average level of happiness ( $O$ ) for group 1 versus group 2. The diagram  $X \rightarrow O$  indicates an expected increase in happiness when supplementary income increases. That is, the average happiness score should be higher for group 1 than for group 2.

By assigning each subject to a particular treatment condition based on a coin flip or some other random procedure, experimental designs ensure that each subject has an equal chance of appearing in any one of the treatment conditions (e.g., at any level of supplementary income). Therefore, as a result of random assignment, the different treatment groups depicted in Table 1 (a) should be approximately equivalent in all characteristics (average education, average physical health, average religiosity, etc.) except their exposure to different levels of the independent variable (i.e., different levels of supplementary income). Consequently, even though there is a large number of other variables (e.g., education, physical health, and religiosity) that might affect happiness, none of these variables can serve as a plausible alternative explanation for why the higher-income group has higher average happiness than does the lower-income group.

For example, due to random assignment, physically healthy versus unhealthy persons should be approximately equally distributed between the higher versus lower supplementary income treatment groups. Hence, a critic could not make a plausible argument that the treatment group receiving the higher amount of supplementary income (i.e., group 1) also has better health, and it is the better health and not the greater income that is producing the higher levels of happiness in that treatment

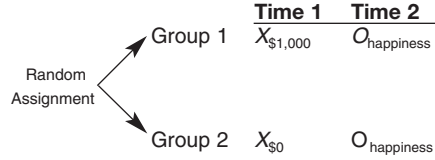
**Types of Research Designs**

**Causal inference for each design:**



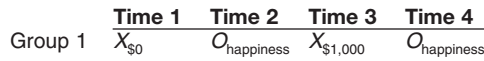
**RANDOMIZED EXPERIMENTAL DESIGN:**

a. Posttest-only control group

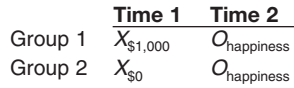


**PREEXPERIMENTAL DESIGNS:**

b. One-group pretest-posttest

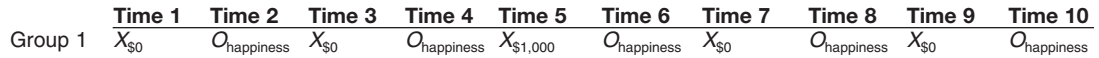


c. Static-group comparison

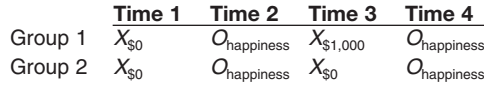


**QUASI-EXPERIMENTAL DESIGNS:**

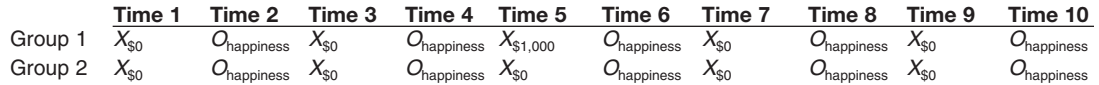
d. Time series



e. Nonequivalent control group



f. Multiple time series

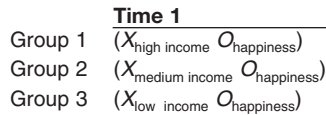


g. Regression-discontinuity

(See Figure 1.)

**NONEXPERIMENTAL DESIGNS:**

h. Passive static group comparison<sup>1</sup>



i. Panel (passive nonequivalent control group)<sup>2</sup>

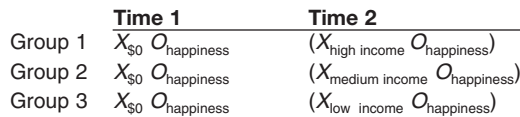


Table 1



Table 1, continued

NOTE: *O* = Observed effect on the dependent variable

*X* = Independent variable (the cause)

<sup>1</sup>Parentheses around the independent and the dependent variable (*X* and *O*) indicates that the measure of *X* and *O* (income and happiness) occur at the same time.

<sup>2</sup>The panel design depends on statistical controls to make groups equivalent in income (and happiness) at time 1.

group. Indeed, this same logic applies no matter what other causal variables a critic might substitute for physical health as an alternative explanation for why additional income is associated with greater happiness.

In sum, strong causal inferences are possible where social scientists manipulate the independent variable and retain great control over when treatments occur, what treatments occur, and, especially, who receives the different treatments. But there are times when investigators, typically in “field” (i.e., nonlaboratory, natural, or real-world) settings, are interested in the effects of an intervention but cannot do randomized experiments. More specifically, there are times when researchers in naturally occurring settings can manipulate the independent variable and exercise at least some control over when the manipulation occurs and what it includes. But these same field researchers may have less control over who receives the treatment conditions. In other words, there are many real-world settings in which *random assignment is not possible*.

Where randomized experiments are not possible, a large number of potential threats to valid causal inference can occur. Under these less-than-optimal field conditions, investigators may resort to a number of alternative research designs that help reduce at least some of the threats to making valid causal inferences. These alternative procedures are collectively referred to as *quasi-experimental designs*. (See also Campbell and Stanley 1963; Cook and Campbell 1979; Cook and Shaddish 1994; Shaddish et al. in preparation.)

None of these designs is as powerful as a randomized experiment in establishing causal relationships, but some of the designs are able to overcome the absence of random assignment such that they approximate the power of randomized experiments. Conversely, where the designs are particularly weak in establishing causal relationships, Campbell and Stanley (1963) have described

them as *preexperimental designs*. Furthermore, social scientists describe as *nonexperimental designs* those studies in which the researcher can only measure (observe) rather than manipulate the independent variable. As we shall see, however, one type of nonexperimental design—the “panel”—may surpass preexperimental designs and approach the power of some quasi-experimental designs in overcoming threats to valid causal inference.

Below we describe common threats to “internal validity” (i.e., the making of valid causal inferences) in field settings, the conditions under which such threats are likely to occur, and representative research designs and strategies used to combat the threats. Later we briefly examine threats to “external validity,” “construct validity,” and “statistical conclusion validity,” and strategies used to reduce these threats. As we shall see, whereas randomized experiments are the exemplary design for enhancing internal validity (causal inference), they often suffer in comparison to other research designs with regard to external validity (generalizability across persons, places, and times) and construct validity (whether one is measuring and manipulating what one intended).

### THREATS TO INTERNAL VALIDITY

Where researchers are unable to assign subjects to treatment conditions randomly, a large number of threats to internal validity (causal inference) can occur. These potential threats include effects due to history, maturation, testing, instrumentation, regression to the mean, selection, mortality, and reverse causal order. (See Cook and Campbell 1979, and Shaddish et al. in preparation for more elaborate lists.)

Research designs vary greatly in *how many* and *which* of these potential threats are likely to occur—that is, are likely to serve as plausible alternative explanations for an apparent causal relationship between an independent and a dependent variable. As an example of a weak (preexperimental)

research design in which most of the various threats to internal validity are plausible, consider the “one-group pretest-posttest design” (Campbell and Stanley 1963). Furthermore, assume that researchers have adapted this design to study the effect of income on happiness. As depicted in Table 1 (b), investigators observe the happiness ( $O_{\text{happiness}}$ ) of persons at time 2 following a period (during time 1) in which subjects (all below the poverty line) receive no supplementary income ( $X_{\$0}$ ). Subsequently, subjects receive a \$1,000 “gift” ( $X_{\$1,000}$ ) at time 3, and their happiness is remeasured ( $O_{\text{happiness}}$ ) at time 4.

The investigators find that posttest happiness (i.e., time 4  $O_{\text{happiness}}$ ) is indeed substantially higher than pretest happiness (i.e., time 2  $O_{\text{happiness}}$ ). Accordingly, an increase in supplementary income is associated with an increase in happiness. But is this association due to supplementary income’s *causing* an increase in happiness? Or is the association due to some alternative explanation?

Given this weak, preexperimental research design, there are a number of threats to internal validity that serve as plausible alternative explanations for increases in happiness other than the \$1,000 gift. These plausible threats include effects due to history, maturation, testing, instrumentation, and regression to the mean, with less likely or logically impossible threats due to selection, mortality, and reverse causal order.

*History* effects refer to some specific event that exerts a causal influence on the dependent variable, and that occurs at roughly the same time as the manipulation of the independent variable. For instance, during the period between the pretest (time 2) and posttest (time 4) measure of happiness as outlined in Table 1 (b), Congress may have declared a national holiday. This event could have the effect of elevating everyone’s happiness. Consequently, even if the \$1,000 gift had no effect on happiness, researchers would observe an increase in happiness from the pretest to posttest measure. In other words, the effects of the \$1,000 gifts are totally confounded with the effects of the holiday, and both remain reasonable explanations for the change in happiness from time 2 to time 4. That is, a plausible rival explanation for the increase in happiness with an increase in income is that the holiday and not the additional income made people happier.

*Maturation* effects are changes in subjects that result simply from the passage of time (e.g., growing hungrier, growing older, growing more tired). Simply put, “people change.” To continue with our current example using a weak, preexperimental research design, assume that individuals, as they grow older, increase in happiness owing to their improved styles of coping, increasing acceptance of adverse life events, or the like. If such developmental changes appear tenable, then maturation becomes a plausible rival explanation for why subjects’ happiness increased after receiving the \$1,000 gift. That is, subjects would have displayed an increase in happiness over time even if they had *not* received the \$1,000 gift.

*Testing* effects are the influences of taking a pretest on subsequent tests. In the current study of income and happiness, pretest measures of happiness allow participants to become familiar with the measures’ content in a way that may have “carryover” effects on later measures of happiness. That is, familiarity with the pretest may make salient certain issues that would not be salient had subjects not been exposed to the pretest. Consequently, it is possible that exposure to the pretest could cause participants to ponder these suddenly salient issues and therefore change their opinions of themselves. For example, people may come to see themselves as happier than they otherwise would have perceived themselves. Consequently, posttest happiness scores would be higher than pretest scores, and this difference need not be due to the \$1,000 gift.

*Instrumentation* effects are a validity threat that occurs as the result of changes in the way that a variable is measured. For instance, in evaluating the effect of income on happiness, researchers may make pretest assessments with one type of happiness measure. Then, perhaps to take advantage of a newly released measure of happiness, researchers might use a different happiness measure on the posttest. Unless the two measures have exactly parallel forms, however, scores on the pretests and posttests are likely to differ. Accordingly, any observed increase in happiness may be due to the differing tests and not to the \$1,000 gift.

*Regression to the mean* is especially likely to occur whenever two conditions are present in combination: (1) researchers select subjects who have extreme scores on a pretest measure of the

dependent variable, and (2) the dependent variable is less than perfectly measured (i.e., is less than totally reliable owing to random measurement error). It is a principle of statistics that individuals who score either especially high or low on an imperfectly measured pretest are most likely to have more moderate scores (i.e., regress toward their respective mean) on the posttest. In the social sciences, almost all variables (e.g., happiness) are less than perfectly reliable. Hence, whenever social scientists assign subjects to treatment conditions based on high or low pretest scores, regression to the mean is likely to occur. For example, researchers may believe that those persons who are most unhappy will benefit most from a \$1,000 gift. Therefore, only persons with low pretest scores are allowed into the study. However, low scorers on the pretest are likely to have higher happiness scores on the posttest simply as a result of remeasurement. Under such circumstances, regression to the mean remains a plausible rival explanation for any observed increase in happiness following the \$1,000 gift.

*Selection* effects are processes that result in different kinds of subjects being assigned to one treatment group as compared to another. If these differences (e.g., sex) affect the dependent variable (e.g., happiness), then selection effects serve as a rival explanation for the assumed effect of the hypothesized causal variable (e.g., income). Because there is not a second group in the one-group pretest-posttest design illustrated here, the design is *not* subject to validity threats due to selection. That is, because the same group receives all treatment conditions (e.g., no gift versus a \$1,000 gift), the characteristics of subjects (e.g., the number of females versus the number of males) remain constant across treatment conditions. Thus, even if females tended to be happier than males, this effect could not explain why an increase in happiness occurred after subjects received the \$1,000 gift.

*Mortality* effects refer to the greater loss of participants (e.g., due to death or disinterest) in one treatment group compared to another. For instance, in the study of the effects of income on happiness, the most unhappy people are more likely than other subjects to drop out of the study before its completion. Because these dropouts appear in the pretest but not the posttest, the average level of happiness will increase. That is, an increase in happiness would occur even if the

supplementary income had no effect whatsoever. Mortality is *not*, however, a plausible alternative explanation in the current example of a study using the one-group pretest-posttest design. Researchers can simply exclude from the study any subjects who appear in the pretest but not the posttest measure of happiness.

*Reverse causal order* effects are validity threats due to ambiguity about the direction of a causal relationship; that is, does  $X$  cause  $O$ , or does  $O$  cause  $X$ ? The one-group pretest-posttest design is *not* subject to this internal validity threat. The manipulation of the independent variable (giving the \$1,000 gift) clearly precedes observation of the dependent variable (degree of happiness). In general, where research designs manipulate rather than measure the independent variable, they greatly reduce the threat of reverse causal order.

As an overview, the reader should note that the various threats to internal validity, where plausible, violate the last two of three conditions necessary for establishing a valid causal inference. Recall that the three conditions are: (1) an association between two variables is present; (2) the presumed cause must precede the presumed effect in time; and (3) the association between the two variables cannot be explained by the influence of a “third” variable that may be related to both of them.

Only the violation of the *first* condition is *not* covered by the list of specific threats to internal validity. (But see the later discussion of threats to statistical conclusion validity.) Reverse causal order is a threat to internal validity that violates the *second* condition of causal inference. Furthermore, history, maturation, testing, instrumentation, regression to the mean, selection, and mortality are all threats to internal validity that one can broadly describe as the potential influence of a “third” variable—that is, threats that violate the *third* condition of causal inference. That is, each of these threats represents a specific type of third variable that affects the dependent variable and coincides with the manipulation of the independent variable. In other words, the third variable is *related* to both the independent and dependent variable. Because the third variable affects the dependent variable at the same time that the independent variable is manipulated, it will *appear* that the independent variable causes a change in the dependent variable. But in fact this apparent causal

relation is a *spurious* (i.e., noncausal) by-product of the third variable's influence.

As an illustration, recall how validity threats due to history can produce a spurious correlation between income and happiness. In the example used earlier, Congress declared a national holiday that increased subjects' happiness and coincided with subjects receiving a \$1,000 gift. Hence, the occurrence of a national holiday represents a "third" variable that is related both to income and happiness, and makes it appear (falsely) that income increases happiness.

Research, in its broadest sense, can be viewed as an investigator's attempt to convince the scientific community that a claimed causal relationship between two variables really exists. Clearly, the presence of one or more threats to internal validity challenges the researcher's claim. That is, the more likely a validity threat seems, the less convincing is the investigator's claim.

When confronted with the specific threats to internal validity in field settings, investigators can attempt to modify their research design to control one or more of these threats. The fact that a specific threat is *possible* for a given research design, however, does not mean it is *plausible*. Implausible threats do little to reduce the persuasiveness of researchers' claims. Therefore, the specific design researchers' use should be determined in large part by the specific threats to validity that are considered most plausible.

Furthermore, as noted earlier, each research design has a given number of possible threats to internal validity, and some designs have more possible threats than do other designs. But only a certain number of these threats will be plausible for *the specific set of variables under study*. That is, different sets of independent and dependent variables will carry different threats to internal validity. Thus, researchers may select weaker designs where the plausible threats for a given set of variables are relatively few and not among the possible threats for the given design. Campbell and Stanley (1963) note, for example, that the natural sciences can often use the one-group pretest-posttest design despite its long list of *possible* threats to internal validity. Given the carefully controlled laboratory conditions and focus on variables measuring nonhuman phenomena, *plausible* threats to internal validity are low.

The next section examines some common quasi-experimental designs and plausible threats to internal validity created by a given design. The discussion continues to use the concrete example of studying the relationship between income and happiness. Examples using a different set of variables might, of course, either reduce or increase the number of plausible threats for a given design.

### QUASI-EXPERIMENTAL DESIGNS

When researchers have the opportunity to make more than a single pretest and posttest, some form of *time series design* becomes possible. Table 1 (*d*) illustrates the structure of this design. The *O*'s designate a series of observations (measures of happiness) on the same individuals (group 1) over time. The table shows that subjects receive no supplementary income ( $X_{\$0}$ ) through the first two (times 2 and 4) observational periods. Then at time 5 subjects receive the \$1,000 gift ( $X_{\$1,000}$ ). Their subsequent level of happiness is then observed at three additional points (times 6, 8, and 10).

This quasi-experimental design has a number of advantages over the single-group pretest-posttest (preexperimental) design. For instance, by examining the trend yielded by multiple observations prior to providing the \$1,000 gift, it is possible to rule out validity threats due to maturation, testing, and regression to the mean. In contrast, instrumentation could still be a threat to validity, if researchers changed the way they measured happiness—especially for changes occurring just before or after giving the \$1,000 gift. Moreover, artifacts due to history remain uncontrolled in the time series design. For example, it is still possible that some positive event in the broader environment could occur at about the same time as the giving of the \$1,000 gift. Such an event would naturally serve as a plausible alternative explanation for why happiness increased after the treatment manipulation.

In addition to eliminating some threats to internal validity found in the one-group pretest-posttest design, the time series design provides measures of how long a treatment effect will occur. That is, the multiple observations (*O*'s) following the \$1,000 gift allow researchers to assess how long happiness will remain elevated after the treatment manipulation.

In some circumstances, the time series design may not be possible, owing to constraints of time or money. In such cases, other quasi-experimental designs may be more appropriate. Consequently, as an alternative strategy for dealing with some of the threats to internal validity posed by the single-group pretest-posttest (preexperimental) design, researchers may add one or more *comparison groups*.

The simplest multigroup design is the *static-group comparison* (Campbell and Stanley 1963). Table 1 (c) provides an illustration of this design. Here observations are taken from two different groups ( $G_1$  and  $G_2$ ) at the same point in time. The underlying assumption is that the two groups differ only in the treatment condition (a \$1,000 gift versus no gift) they receive prior to the measure of happiness. In many instances, this is not a safe assumption to make.

The static-group comparison design does reduce some potential threats to internal validity found in the single-group pretest-posttest design; namely, history, testing, instrumentation, and regression to the mean. That is, each of these threats should have *equal* effects on the two experimental groups. Thus, these threats cannot explain why experimental groups differ in posttest happiness.

Conversely, the static-group comparison design adds other potential threats—selection, reverse causal order, and mortality effects—not found in the single-group pretest-posttest design. Indeed, these threats are often so serious that Stanley and Campbell (1963) refer to the static-group comparison, like the single-group pretest-posttest, as a “pre-experimental” design.

*Selection* effects are generally the most plausible threats to internal validity in the static-group comparison design. That is, in the absence of random assignment, the treatment groups are likely to differ in the type of people they include. For example, researchers might assign poverty-level subjects to the \$1,000 gift versus no gift treatment groups based on physical health criteria. Subjects in poor health would receive the supplementary income; subjects in better health would not. Note, however, that poor health is likely to reduce happiness, and that less healthy—and therefore less happy—people appear in the \$1,000 treatment condition. Hence, it is possible that this selection effect based on physical health could *obscure* the increase in happiness due to the

supplementary income. In other words, even if the \$1,000 gift does have a positive effect on happiness, researchers might make a false causal inference; namely, that supplementary income has *no* effect on happiness.

This result illustrates the point that threats to internal validity are not always ones that refute a claim that a causal effect occurred. Threats to internal validity can also occur that refute a claim that a causal effect did *not* happen. In other words, threats to internal validity concern possible false-negative findings as well as false-positive findings.

The preceding example showed how false-negative findings can result due to selection effects in the static-group comparison. *False-positive* findings can, of course, also occur due to selection effects in this design. Consider, for instance, a situation in which researchers assign subjects to treatment conditions based on contacts with a particular governmental agency that serves the poor. Say that the first twenty subjects who contact this agency on a specific day receive the \$1,000 gift, and the next twenty contacts serve as the no-gift comparison group. Furthermore, assume that the first twenty subjects that call have extroverted personalities that made them call early in the morning. In contrast, the next twenty subjects are less extroverted and thus call later in the day. If extroverted personality also produces higher levels of happiness, then the group receiving the \$1,000 gift would be happier than the no-gift comparison group even before the treatment manipulation. Accordingly, even if supplementary income has no effect on happiness, it will appear that the \$1,000 gift increased happiness. In other words, extroverted personality is a “third” variable that has a positive causal effect on both level of supplementary income and happiness. That is, the more extroversion, the more supplementary income; and the more extroversion, the more happiness. These causal relationships therefore make it appear that there is a positive, causal relationship between supplementary income and happiness; but in fact this latter correlation is *spurious*.

*Reverse causal order* effects are another potential threat to internal validity when researchers use the static-group comparison design. Indeed, reverse causal order effects are really just a special case of selection effects. More specifically,

reverse causal order effects will occur whenever the dependent variable is also the “third” variable that determines who is assigned to which treatment groups.

By substituting happiness for extroversion as the “third” variable in the preceding example, one can demonstrate how this reverse causal order effect could occur. Recall that subjects who contacted the government agency first were the most extroverted. Assume now, instead, that the earliest callers were happier people than those who called later (because unhappy people are more likely to delay completing tasks). Under these conditions, then, prior levels of happiness comprise a “third” variable that has a positive causal effect on both level of supplementary income and subsequent happiness. That is, those subjects who are initially happier are more likely to receive supplementary income; and those subjects who are initially happier are more likely to experience subsequent (posttest) happiness. These causal relationships hence make it appear that there is a positive, causal association between supplementary income and happiness. In fact, however, this correlation is spurious. Indeed, it is not supplementary income that determines happiness; it is happiness that determines supplementary income.

*Mortality* is another possible threat to internal validity in the static-group comparison design. Even if the treatment groups have essentially identical characteristics before the manipulation of the independent variable (i.e., no selection effects), differences between the groups can occur as a consequence of people dropping out of the study. That is, by the time researchers take posttest measures of the dependent variable, the treatment groups may no longer be the same.

For example, in the study of income and happiness, perhaps some individuals in the no-gift group hear that others are receiving a \$1,000 gift. Assume that among those people, the most likely to drop out are those who have a “sour” disposition, that is, those who are likely to be the most unhappy members of the group in general. Consequently, the no-gift comparison group will display a higher posttest measure of happiness than the group would have if all members had remained in the study. Thus, even if the \$1,000 gift increases happiness, the effect may be obscured by the corresponding, “artificial” increase in happiness

in the no-gift comparison group. In other words, mortality effects may lead researchers to make a false causal inference; namely, that there isn’t a causal relationship between two variables, when in fact there is.

One of the most common quasi-experimental designs is the *nonequivalent control group design*. This design is an elaboration of the static-group comparison design. The former is a stronger design than the latter, however, because researchers administer pretests on all groups prior to manipulating the independent variable. Table 1 (*e*) illustrates this design.

A major advantage of the pretests is that they allow researchers to *detect the presence* of selection effects. Specifically, by comparing pretest scores for the different treatment groups before the manipulation of treatment conditions, it is possible to discern whether the groups are initially different. If the groups differ at the time of the pretest, any observed differences at the posttest may simply be a reflection of these preexisting differences.

For instance, in the income and happiness study, if the group receiving the \$1,000 gift is happier than the no-gift comparison group at the time of the pretest, it would not be surprising for this group to be happier at posttest, even if supplementary income had no causal effect. The point is that the nonequivalent control group design, unlike the static-group comparison design, can test whether this difference is present. If there is no difference, then researchers can safely argue that selection effects are not a threat to internal validity in their study.

The inclusion of pretest scores also permits the nonequivalent control group design to detect the presence or absence of other threats to internal validity not possible using the static-group comparison design—namely, mortality and reverse causal order. Recall that threats due to reverse causal order are a special subset of selection effects. Thus, the ability of the nonequivalent control group design to detect selection effects means it should also detect reverse causal order effects. Selection effects occur as a consequence of differences in pretest measures of the dependent variable. Therefore, in the present example, differences between groups in pretest happiness would

indicate the possibility of reverse causal order effects. In other words, the amount of pretest happiness determined the amount of supplementary income subjects received (\$1,000 gift versus no gift), rather than the converse, that the amount of supplementary income determined the amount of posttest happiness.

Furthermore, the pretest scores of the nonequivalent control group design also allow assessment of mortality effects. Regardless of which subjects drop out of which treatment condition, the researcher can examine the pretest scores for the remaining subjects to ensure that the different treatment groups have equivalent initial scores (e.g., on happiness).

In sum, the nonequivalent control group design is able to reduce all the threats to internal validity noted up to this point. Unfortunately, it is unable to detect one threat to internal validity not previously covered—*selection by maturation interactions*. (For a more complete list of interactions with selection, see Cook and Campbell 1979, and Shadish et al. in preparation.) This threat occurs whenever the various treatment groups are maturing—growing more experienced, tired, bored, and so forth—at different rates.

For example, consider a situation in which the pretest happiness of the group receiving no gift is as high as the group receiving the \$1,000 gift. Moreover, the pretest measures occur when both groups are in stimulating environments, in contrast to the boring environments for the posttest measures. Assume now that there is a greater proportion of people who become bored easily in the no-gift group as compared to the \$1,000-gift group. That is, there is a selection effect operating that results in different kinds of people in one group compared to the other. But this difference doesn't manifest itself until a nonstimulating environment triggers the maturational process that generates increasingly higher levels of boredom. The differential rates at which boredom occurs in the two groups result in higher levels of boredom and corresponding unhappiness in the no-gift as compared to the \$1,000-gift group. In other words, the group receiving the \$1,000 gift will display higher levels of posttest happiness than the no-gift

group, even if supplementary income has no effect on happiness.

The *multiple time series design* incorporates aspects of both the nonequivalent control group and the time series designs. Table 1 (f) illustrates the results of this combination. By extending the number of pretest and posttest observations found in the nonequivalent control group design, the multiple time series design can detect selection-maturation interactions. For instance, if differential reactions to boredom explain why the group receiving the \$1,000 gift has higher happiness than the no-gift group, then we should expect to see these differences in at least some of the additional pretest measures (assuming that some of these additional group comparisons occur in nonstimulating environments). We would also expect the differential reaction to boredom to manifest itself in the additional posttest measures of the multiple time series design. That is, whenever researchers take posttest measures in stimulating environments, they should observe no group differences. Conversely, whenever researchers take posttest measures in nonstimulating environments, they should observe higher happiness among the group receiving the \$1,000 gift.

Furthermore, by adding a second group to the original, single-group time series, the multiple time series reduces the threat of history that is a major problem with the single-group design. Events (e.g., national holidays) that coincide with the manipulation of the independent variable (e.g., \$1,000 gift versus no gift) should have equal impacts on each group in the analysis.

By incorporating multiple groups with pretests and posttests, the multiple time series and nonequivalent control group designs can be effective at reducing a long list of internal validity threats; but in order to actually reduce a number of these threats, researchers must assume that the different treatment groups are functionally equivalent prior to manipulating the independent variable. Pretest scores allow researchers to detect, at least in part, whether this condition of equivalence is present; but what if the groups are not initially equivalent? Under these conditions, researchers may attempt

to equate the groups through “matching” or other, “statistical” adjustments or controls (e.g., analysis of covariance). However, matching is never an acceptable technique for making groups initially equivalent (see Nunnally 1975; Kidder and Judd 1986). And statistical controls are a better but still less-than-desirable procedure for equating groups at the pretest (see Lord 1967; Dwyer 1983; Rogosa 1985).

In sum, an overview of pre- and quasi-experimental designs using multiple groups indicates the importance of establishing the equivalence of the groups through pretest measures. Further, researchers should try to obtain as many additional observations as possible both before and after manipulating the treatments. When groups are nonequivalent at the outset, it is extremely difficult to discern whether treatment manipulations have a causal effect.

In certain field settings, however, ethical considerations may *mandate* that groups be nonequivalent at the outset. That is, researchers must assign subjects to certain treatment conditions based on who is most “needy” or “deserving.” If the dependent variable (e.g., happiness) is associated with the criterion (e.g., physical health) that determines who is most needy or deserving, then the experimental groups will not display pretest equivalence (e.g., the people with the worst health and hence lowest pretest happiness must be assigned to the group receiving the \$1,000 gift).

Fortunately, the *regression-discontinuity design* (Thistlethwaite and Campbell 1960; Cook and Campbell 1979) often allows researchers to make relatively unambiguous interpretation of treatment effects, even where groups are not initially equivalent. Indeed, evidence indicates that this design, when properly implemented, is *equivalent* to a *randomized experiment* in its ability to rule out threats to internal validity (Cook and Shadish 1994; Shadish et al. in preparation).

To continue with the example of income and happiness, researchers may feel compelled to give the \$1,000 gift to those individuals with the poorest health. The investigators would therefore begin by developing a scale of “need” in which

participants below a certain level of physical health receive the gift and those above this cutting point do not. This physical health scale constitutes the “pseudo”-pretest necessary for the regression-discontinuity design. The usual posttest measures of the dependent variable—happiness—would follow the manipulation of the no-gift versus the \$1,000-gift treatment conditions. Researchers would then regress posttest happiness measures on “pretest” measures of physical health. This regression analysis would include the calculation of separate regression lines for (1) those subjects receiving the \$1,000 gift and (2) those subjects not receiving it.

Figure 1 provides an illustration of the results using the regression-discontinuity design. (The structure of the design does not appear in Table 1 due to its relative complexity.) If the \$1,000 gift has a discernible impact on happiness, a “discontinuity” should appear in the regression lines at the cutting point for “good” versus “poor” health. An essential requirement for the regression-discontinuity design is a clear cutting point that defines an unambiguous criterion (e.g., degree of physical health) by which researchers can assign subjects to the treatment conditions. It is the clarity of the decision criterion, not its content, that is important.

An interesting characteristic of the regression-discontinuity design is that it works even if the decision criterion has no effect on the outcome of interest (e.g., happiness). Indeed, as the variable that forms the decision criterion approaches a condition in which it is totally unrelated to the outcome, the decision criterion becomes the functional equivalent of *assigning subjects randomly* to treatment conditions (Campbell 1975). Even where the criterion is strongly related to the outcome, the regression-discontinuity design, when properly implemented, can still approximate a randomized experiment in reducing threats to internal validity.

There are, however, several threats to internal validity that can occur in using the regression-discontinuity design (hence the use of the above qualifier: “when properly implemented”). One threat emerges when the relationship between the pseudo-pretest measure (e.g., physical health) and



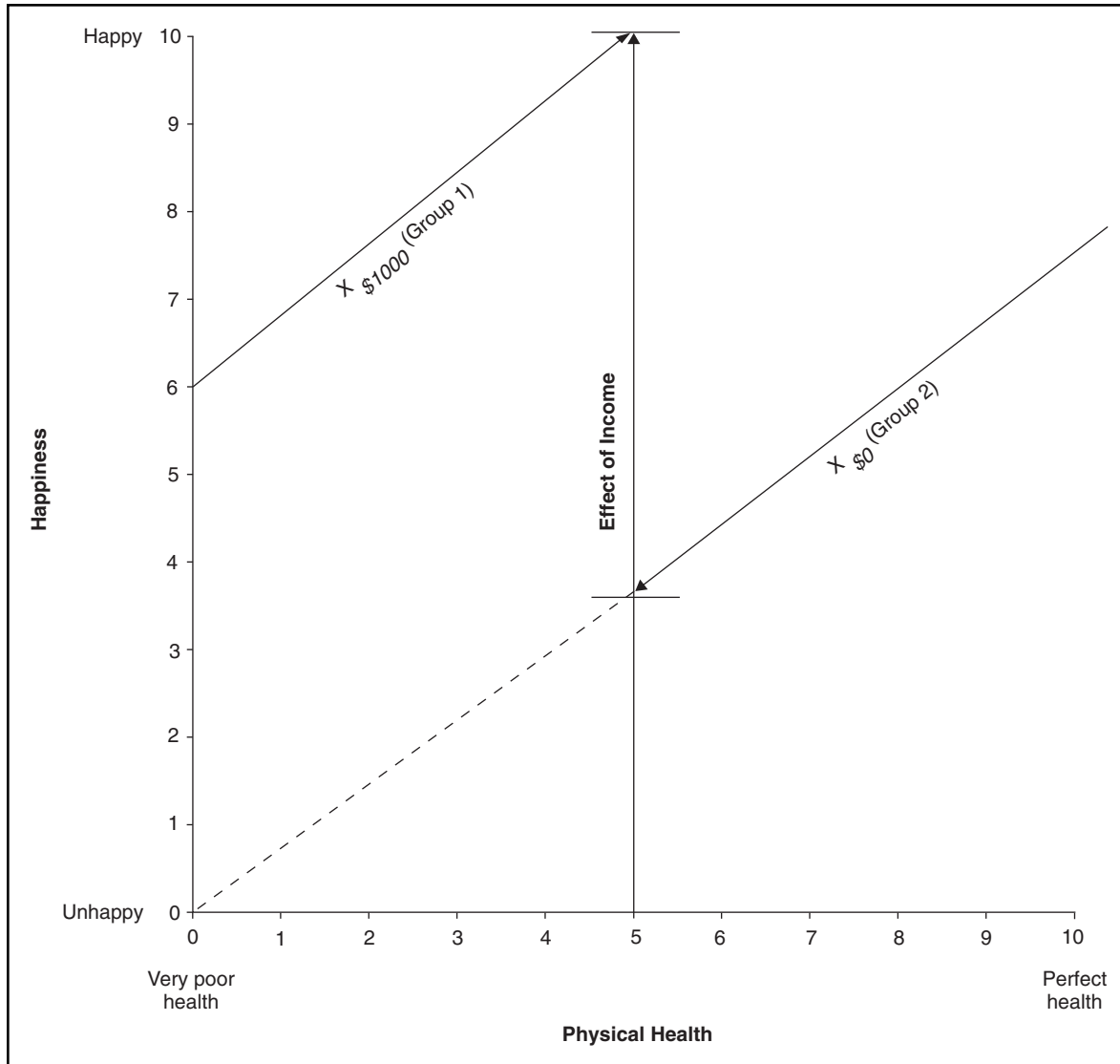


Figure 1. Regression-Discontinuity Design Showing Treatment Effect of Income on Happiness (Dashed line indicates projected happiness scores for group 1 if the \$1,000 treatment had no effect on happiness.)

the posttest measure (e.g., happiness) does not form a linear pattern. In fact, a curvilinear relationship near the cutting point may be indistinguishable from the discontinuity between the separate regression lines. Moreover, another threat to internal validity arises if the investigators do not strictly adhere to the decision criterion (e.g., if they feel sorry for someone who is close to qualifying for the \$1,000 and thus gives that person the “benefit of the doubt”). Additionally, if only a few people receive a given treatment condition (e.g., if

only a few \$1,000 awards can be made, for financial reasons), the location of the regression line may be difficult to estimate with any degree of accuracy for that particular treatment condition. Accordingly, researchers should include relatively large numbers of subjects for all treatment conditions.

In summary, quasi-experimental designs allow researchers to maintain at least some control over how and when they manipulate the independent variable, but researchers lose much control over who receives specific treatment conditions (i.e.,

the designs do not permit random assignment). Quasi-experimental designs differ in how closely they approximate the power of randomized experiments to make strong causal inferences. As a general rule, the more observations quasi-experimental designs add (i.e., the more *O*'s, as depicted in the diagrams of Table 1), the more the designs are able to reduce threats to internal validity.

### NONEXPERIMENTAL DESIGNS

In contrast to quasi-experimental designs, *nonexperimental designs do not manipulate* the independent variable. Thus, researchers have no control over who falls into which category of the independent variable when there is a change from one to another category of the independent variable, or what content the various levels of the independent variable will contain. Rather than serving as an agent that actively changes (manipulates) the independent variable, the researcher must be content to passively observe (measure) the independent variable as it naturally occurs. Hence, some social scientists also refer to nonexperimental designs as *passive-observational designs* (Cook and Campbell 1979).

When researchers can only measure rather than manipulate the independent variable, threats to internal validity increase greatly. That is, reverse causal order effects are much more likely to occur. There is also a much greater probability that some "third" variable has a causal effect on both the independent and the dependent variables, therefore producing a spurious relationship between the latter two variables.

To illustrate these points, consider the most widely used research design among sociologists; namely, the static-group comparison design with measured rather than manipulated independent variables—or what we will refer to here as the *passive static-group comparison*. Table 1 depicts the structure of this nonexperimental, cross-sectional design. Researchers generally combine this design with a "survey" format, in which subjects self-report their scores on the independent and dependent variables (e.g., report their current income and happiness).

Note that the structure of this nonexperimental design is basically the same as the static-group

comparison design found in Table 1 (c). To better capture the various levels of naturally occurring income, however, the diagram in Table 1 expands the number of categories for income from two *manipulated* categories (\$1,000 gift versus no gift) to three *measured* categories (high, medium, and low personal income). Furthermore, whereas the original static-group design manipulated income *before* measuring happiness, the passive static-group design measures both personal income and happiness at the *same time* (i.e., when subjects respond to the survey). Consequently, the temporal ordering of the independent and dependent variable is often uncertain.

Indeed, because researchers do not manipulate the independent variable, and because they measure the independent and dependent variables at the same time, the threat of reverse causal order effects is particularly strong in the passive static-group comparison design. In the present example, it is quite possible that the amount of happiness a person displays is a determinant of how much money that person will subsequently make. That is, happiness causes income rather than income causes happiness. What is even more likely is that both causal sequences occur.

Additionally, the passive static-group comparison design is also especially susceptible to the threat that "third" variables will produce a spurious relationship between the independent and dependent variables. For example, it is likely that subjects who fall into different income groupings (high, medium, and low) also differ on a large number of other (selection effect) variables. That is, the different income groups are almost certainly nonequivalent with regard to characteristics of subjects in addition to income. One would expect, for instance, that the higher-income groups have more education, and that education is associated with greater happiness. In other words, there is a causal link between education and income, and between education and happiness. More specifically, higher education should produce greater income, and higher education should also produce greater happiness. Hence, education could produce a spurious association between income and happiness.

As noted earlier, researchers can attempt to equate the various income groups with regard to

third variables by making statistical adjustments (i.e., controlling for the effects of the third variables). But this practice is fraught with difficulties (again, see Lord 1967; Dwyer 1983; Rogosa 1985).

It is especially sobering to realize that a design as weak as the passive static-group comparison is so widely used in sociology. Note, too, that this design is a substantially weaker version of the static-group comparison design that Campbell and Stanley (1963) considered so weak that they labeled it “preexperimental.” Fortunately, there are other nonexperimental, longitudinal designs that have more power to make valid causal inferences. Most popular among these designs is the *panel design*. (For additional examples of passive longitudinal designs, see Rogosa 1988; Eye 1990.)

Table 1 depicts the structure of this longitudinal survey design. It is very similar to the nonequivalent control group design in Table 1 (*e*). It differs from the quasi-experimental design, however, because the independent variable is measured rather than manipulated, and the independent and dependent variable are measured at the same time.

In its simplest, and weaker, two-wave form (shown in Table 1), the panel design can address at least some threats due to reverse causal order and third variables associated with the independent and dependent variable. (This ability to reduce threats to internal validity is strengthened where investigators include three and preferably four or more waves of measures.) The most powerful versions of the panel design include data analysis using structural equation modeling (SEM) with multiple indicators (see Kessler and Greenberg 1981; Dwyer 1983; and for a more general introduction to SEM, see Kline 1998). Simultaneous equations involve statistical adjustments for reverse causal order and causally linked third variables. Thus, the standard admonishments noted earlier about using statistical control techniques apply here too.

Finally, Cook and his associates (Cook and Shadish 1994; Shadish et al. in preparation) have noted a real *advantage* of nonexperimental designs over randomized experiments. Specifically, experimental designs lend themselves to studying causal linkages (i.e., “descriptions” of causal conditions) rather than the *processes* accounting

for these linkages (i.e., explanations of causal effects). In contrast, nonexperimental designs lend themselves to using causal (path) modeling to study “process”—that is, intervening variables that “mediate” the effect of the independent on the dependent variable. The field of causal modeling using nonexperimental designs has developed at a tremendous pace, dominated by the very flexible and sophisticated data analytic procedure of SEM (Kline 1998).

### RANDOMIZED EXPERIMENTS REVISITED

The great problems that nonexperimental designs encounter in making causal inferences help illustrate the increased power researchers obtain when they move from these passive-observational to quasi-experimental designs. But no matter how well a quasi-experimental design controls threats to internal validity, there is no quasi-experimental design that can match the ability of randomized experiments to make strong causal inferences. Indeed, threats to internal validity are greatly reduced when researchers are able to randomly assign subjects to treatment groups. Therefore, the value of this simple procedure cannot be overstated.

Consider the simplest and most widely used of all randomized experiments, the *posttest-only control group design* (Campbell and Stanley 1963), as depicted in Table 1 (*a*). Note that it is similar in structure to the very weak, preexperimental, static-group comparison design in Table 1 (*c*). These two designs differ primarily with regard to whether they do or do not use *random assignment*.

The addition of random assignment buys enormous power to make valid causal inferences. With this procedure, and the manipulation of an independent variable that temporally precedes observation of the dependent variable, reverse causal order effects are impossible. Likewise, with the addition of random assignment, other threats to internal validity present in the static-group comparison design dissipate. Specifically, selection effects are no longer a major threat to internal validity. That is, selection factors—different kinds of people—should appear approximately evenly distributed between categories of the independent variable (e.g., \$1,000 gift versus no gift). In

other words, the different groups forming the treatment conditions should be roughly equivalent prior to the treatment manipulation. Further, given this equivalence, threats due to selection-maturation interactions are also reduced.

Conversely, given that pretest measures of the dependent variable (e.g., happiness) are absent, mortality effects remain a potential threat to internal validity in the posttest-only control group design. Of course, for mortality effects to occur, different kinds of subjects have to drop out of one experimental group as compared to another. For example, in the study of income on happiness, if certain kinds of subjects (say, those who are unhappy types in general) realize that they are not in the group receiving the \$1,000 gift they may refuse to continue. This situation could make it appear (falsely) that people receiving the \$1,000 gift are less happy than those not receiving the gift.

Naturally, the probability that *any* subjects drop out is an increasing function of how much time passes between manipulation of treatment conditions and posttest measures of the dependent variable. The elapsed time is generally quite short for many studies using the posttest-only control group design. Consequently, in many cases, mortality effects are not a plausible threat.

In sum, one may conclude that random assignment removes, at least in large part, all of the major threats to internal validity. (But see Cook and Campbell 1979, Cook and Shadish 1994, and Shadish et. al. in preparation for some additional qualifications.)

Two more points are noteworthy with respect to random assignment. First, it is important to realize that this procedure *does not ensure* that third variables that might influence the outcome will be evenly distributed between groups in any particular experiment. For instance, random assignment does not ensure that the average level of education will be the same for the group receiving the \$1,000 gift as for the group receiving no gift. Rather, random assignment allows researchers to *calculate the probability* that third variables (such as education) are a plausible alternative explanation for an apparent causal relationship (e.g., between supplementary income and happiness). Researchers are generally willing to accept that a causal relationship between two variables is real if the relationship could have occurred by chance—that is, due

to the *coincidental* operation of third variables—less than one time out of twenty.

Some researchers add pretests to the posttest-only control group design in order to evaluate the “success” of the random assignment procedure, and to add “statistical power.” According to Nunnally (1975), however, the use of a pretest is generally not worth the attendant risks. That is, the pretest may sensitize subjects to the treatment conditions (i.e., create a treatment-testing interaction). In other words, the effect of the independent variable may not occur in other situations where there are no pretest measures. Thus, any gain from pretest information is likely to be offset by this threat to “construct validity.” (For an example of a treatment-testing interaction, see the section on threats to construct validity, below).

Second, random *assignment* of subjects is different from random *selection* of subjects. Random assignment means that a subject has an equal probability of entry into each treatment condition. Random selection refers to the probability of entry into the study as a whole. The former issue bears on internal validity (i.e., whether observed outcomes are unambiguously due to the treatment manipulation); the latter is an issue of external validity (i.e., the extent to which the results of the study are generalizable).

### THREATS TO EXTERNAL VALIDITY

Whereas internal validity refers to whether or not a treatment is effective, *external validity* refers to the conditions under which the treatment will be effective. That is, to what extent will the (internally valid) causal results of a given study apply to different people and places?

One type of threat to external validity occurs when certain treatments are likely to be most effective on certain kinds of people. For example, researchers might find that a \$1,000 gift has a strong effect on happiness among a sample of young adults. Conversely, a study of extremely elderly persons might find that the \$1,000 has no effect on happiness (say, because very old people are in general less stimulated by their environment than are younger age groups). Cook and Campbell (1979) label this external validity threat an *interaction between selection factors and treatment*.

Researchers sometimes mistakenly assume that they can overcome this threat to external validity by *randomly selecting* persons from the population across which they wish to generalize research findings. Random samples do not, however, provide appropriate tests of whether a given cause-effect finding applies to different kinds of people. Obtaining a random sample of, say, the U.S. population would not, for instance, reproduce the above (hypothetical) finding that a \$1,000 gift increases happiness among younger but not older people. Combining younger and older persons in a representative sample would only lead to an averaging or “blending” of the strong effect for youths with the no-effect result for the elderly. In fact, the resulting finding of a “moderate” effect would not be an accurate statement of income’s effect on happiness for either the younger or older population.

Including younger and older persons in a random sample would only increase external validity if the researchers *knew* to provide separate analyses for young and old—among an infinite variety of *possible* human characteristics that researchers might choose to do subsample analyses on (e.g., males and females). But if researchers suspected that the treatment might interact with age, then they could simply make sure that their nonrandomly selected, “convenience” sample contained sufficient numbers of both youths and elderly to do separate analyses on each age group.

Additionally, threats to external validity occur because certain treatments work best in certain settings. Giving \$1,000 to a person at a shopping mall may increase their happiness substantially compared to the same gift given to someone stranded on a desert island with nowhere to spend the money. Cook and Campbell (1979) label this external validity threat an *interaction between treatment and setting*. Given that quasi-experimental designs are most often located in “real-life,” field settings, they are somewhat less susceptible to this threat than are randomized experiments—which most often occur in “artificial,” laboratory settings.

Note that threats to external validity are concerned with restricting cause-effect relationships to particular persons or places. Therefore, the best procedure for reducing these restrictions is to *replicate* the findings on different persons and in different settings—either within a single study or

across a series of studies (Cook and Campbell 1979; Cook and Shadish 1994).

### THREATS TO CONSTRUCT VALIDITY

*Construct validity* refers to the accuracy with which researchers manipulate or measure the construct intended rather than something else (Cook and Campbell 1979; and for updates, see Cook and Shadish 1994; and Shadish et al. in preparation). Thus, for example, investigators might establish that their manipulation of a variable labeled “income” does indeed have a causal effect on their measure of the outcome labeled “happiness.” That is, the researchers have avoided plausible threats to internal validity and, consequently, have presented a convincing claim for a cause-and-effect relationship. Critics might question, however, whether the manipulation labeled “income” and the measure labeled “happiness” do in fact represent the concepts that the investigators claim they have manipulated and measured, respectively.

For instance, in providing supplementary income to selected subjects, researchers might also have manipulated, say, the perception that the researchers really are concerned about the welfare of the subjects. It may be subjects’ perceptions of this “caring attitude,” and not an increase in “economic well-being,” that produced the effect the \$1,000 gift had on happiness. In other words, investigators were manipulating a dimension in addition to the economic dimension they intended to manipulate.

Likewise, in asking subjects to answer a questionnaire that purportedly measures “happiness,” researchers may not be measuring happiness but rather the degree to which subjects will respond in socially desirable ways (e.g., some subjects will respond honestly to questions asking how depressed they are, and other subjects will hide their depression).

Cook and Campbell (1979) provide an extensive list of threats to construct validity. The description of these threats is rather abstract and complicated. Hence, the following discussion includes only a few concrete examples of potential threats. For a more complete list and discussion of these threats, the interested reader should consult the original article by Cook and Campbell, as well

as the update to their book (Shadish et al. in preparation) and other volumes on the construct validity of measures and manipulations (e.g., Costner 1971; Nunnally and Bernstein 1994).

One type of threat to construct validity occurs in research designs that use pretests (e.g., the nonequivalent control group design). Cook and Campbell (1979) label this threat an *interaction of testing and treatment*. This threat occurs when something about the pretest makes participants more receptive to the treatment manipulation. For example, in the study of income and happiness, the pretest may make salient to participants that “they don’t have much to be happy about.” This realization may, in turn, make subjects more appreciative and thus especially happy when they later receive a \$1,000 gift. In contrast, the \$1,000 gift might have had little or no causal impact on happiness if subjects were not so sensitized, that is, were not exposed to a pretest. Accordingly, it is the combination of the pretest and \$1,000 gift that produces an increase in happiness. Neither condition alone is sufficient to create the causal effect. Consequently, researchers who use pretests must be cautious in claiming that their findings would apply to situations in which pretests are not present. Because quasi-experimental designs are dependent on pretest observations to overcome threats to internal validity (i.e., to establish the initial equivalence of the experimental groups), researchers cannot safely eliminate these measures. Thus, to enhance the construct validity of the manipulation, researchers should strive to use as unobtrusive measures as possible (e.g., have trained observers or other people familiar with a given subject secretly record the subject’s level of happiness).

Another set of potential threats to construct validity concerns what Campbell and Stanley (1963) describe as *reactive arrangements*. Cook and Campbell (1979) have subsequently provided more specific labels for these threats: *hypothesis-guessing within experimental conditions*, *evaluation apprehension*, and *experimenter expectancies* (see also Rosenthal and Rosnow 1969). Threats due to reactive arrangements result as a consequence of the participants’ knowing they are in a study, and therefore behaving in a way that they might not in more natural circumstances. With regard to this phenomena, Orne (1962) used the term “demand characteristics” to refer to the totality of cues that affect a

subject’s response in a research setting in the sense that certain characteristics “demand” certain behaviors. For instance, subjects receiving the \$1,000 gift may guess the hypothesis of the study when they are subsequently asked to respond to questions about their state of happiness. Realizing that the study may be an attempt to show that supplementary income increases happiness, participants may try to be “good subjects” and confirm the experimental hypothesis by providing high scores on the happiness questionnaire. In other words, the treatment manipulation did in fact produce an increase in the assumed *measure* of “happiness,” but the measure was actually capturing participants’ willingness to be “good subjects.”

A classic example of reactive arrangements is the Hawthorne effect (see Lang 1992 for a more comprehensive review). The Hawthorne effect was named for a series of studies conducted between 1924 and 1932 at the Hawthorne Works of the Western Electric Company near Chicago (Mayo 1933; Roethlisberger and Dickson 1939). Researchers attempted to determine, among other things, the effects of illumination on worker productivity. The results were puzzling. There was no clear relationship between illumination and worker performance. Every change, even seemingly adverse changes in which normal lighting was reduced by 50 to 70 percent resulted in increased productivity. In addition, productivity often remained high even after workers were returned to their original working conditions. Even more confusing was the fact that not all the studies reported increased productivity. In some studies, depending upon the factors being manipulated, the effect was even reversed, with workers apparently deliberately reducing their output.

The three most common explanations for the Hawthorne effect are: (1) subjects in studies respond to special attention; (2) awareness of being in a study affects subjects’ performance; and (3) subjects react to the novelty of certain aspects of the research procedures (Lang 1992). Subsequent research has not supported any of these explanations conclusively (Adair et al. 1989). Nor is there widespread evidence of the Hawthorne effect in either experimental or field settings (Cook and Campbell 1979). What is clear, however, is that employees, within organizations, are part of social systems that can affect behavior in research settings. Indeed, the Hawthorne studies provided

impetus to the development of the new field of “organizational behavior,” which has strong links to sociology.

No widely accepted model of the processes involved in subject reactivity presently exists. But to reduce threats to construct validity due to reactive arrangements, researchers may attempt, where feasible, to disguise the experimental hypothesis, use unobtrusive measures and manipulations, and keep both the subject and the person administering the treatments “blind” to who is receiving what treatment conditions. These disguises are generally easier to accomplish in the naturally occurring environments of quasi-experimental designs than in the artificial settings of laboratory experiments. Finally, there are additional, sophisticated structural equation modeling procedures for discerning where reactive arrangements may be present in a study, and for making “statistical adjustments” to correct for the bias that these threats would otherwise introduce (Blalock 1985).

#### THREATS TO STATISTICAL CONCLUSION VALIDITY

Before researchers can establish whether an independent variable has a causal effect on the dependent variable, they must first establish whether an *association* between the two variables does or does not exist. *Statistical conclusion validity* refers to the accuracy with which one makes inferences about an association between two variables—without concern for whether the association is causal or not (Cook and Campbell 1979; Shadish et al. in preparation). The reader will recall that an association between two variables is the first of three conditions necessary to make a valid causal inference. Thus, statistical conclusion validity is closely linked to internal validity. To put it another way, statistical conclusion validity is a necessary but not sufficient condition for internal validity.

Threats to statistical conclusion validity concern either one of two types of errors: (1) inferring an association where one does not exist (described as a “Type I error,” or (2) inferring no association where one does exist (described as a “Type II error”). Researchers ability to avoid Type II errors depends on the power of a research design to uncover even weak associations, that is, the power

to avoid making the mistake of claiming an association is due to “chance” (is statistically insignificant) when in fact the association really exists. Type II errors are more likely to occur the lower the probability level that researchers set for accepting an association as being statistically significant; the smaller the sample size researchers use; the less reliable their measures and manipulations; and the more random error introduced by (1) extraneous factors in the research setting that affect the dependent variable, and (2) variations among subjects on extraneous factors that affect the dependent variable (Cook and Campbell 1979).

Investigators can reduce Type II errors (false claims of no association) by: (1) setting a higher probability level for accepting an association as being statistically significant (e.g.,  $p.05$  instead of  $p.01$ ); (2) increasing the sample size; (3) correcting for unreliability of measures and manipulations (see Costner 1971); (4) selecting measures that have greater reliability (e.g., using a ten-item composite measure of happiness instead of a single-item measure); (5) making treatment manipulations as consistent as possible across occasions of manipulation (e.g., giving each subject the \$1,000 gift in precisely the same manner); (6) isolating subjects from extraneous (outside) influences; and (7) controlling for the influence of extraneous subject characteristics (e.g., gender, race, physical health) suspected to impact the dependent variable (Cook and Campbell 1979).

Type I errors (inferring an association where one does not exist) are more likely the higher the probability level that researchers set for accepting an association as being statistically significant, and the more associations a researcher examines in a given study. The latter error occurs because the more associations one includes in a study, the more associations one should find that are statistically significant “just by chance alone.” For example, given 100 associations and a probability level of  $.05$ , one should on the average find 5 associations that are statistically significant due to chance.

Researchers can reduce threats of making Type I errors by setting a lower probability level for statistical significance, particularly when examining many associations between variables. Of course, decreasing Type I errors increases the risk of Type II errors. Hence, one should set lower probability

levels in conjunction with obtaining reasonably large samples—the latter strategy to offset the risk of Type II errors.

Research designs vary greatly in their ability to implement strategies for reducing threats to statistical conclusion validity. For example, very large sample sizes (say, 500 subjects or more) are generally much easier to obtain for nonexperimental designs than for quasi-experimental or experimental designs. Moreover, experimental designs generally occur in laboratory rather than naturally occurring settings. Thus, it is easier for these designs to control for extraneous factors of the setting (i.e., random influences of the environment). Additionally, experimental designs are generally better able than quasi-experimental designs to standardize the conditions under which treatment manipulations occur.

### SUMMARY AND CONCLUSIONS

Quasi-experimental designs offer valuable tools to sociologists conducting field research. This article has reviewed various threats that researchers must overcome when using such designs. In addition, to provide a context in which to evaluate the relative power of quasi-experimental designs to make valid causal inferences, this article also reviewed examples of experimental and nonexperimental designs.

It is important to note that the quasi-experimental designs described here are merely illustrative; they are representative of the types of research designs that sociologists might use in field settings. These designs are not, however, exhaustive of the wide variety of quasi-experimental designs possible. (See Campbell and Stanley 1963, Cook and Campbell 1979, and Shaddish et al. in preparation, for more extensive reviews.) In fact, great flexibility is one of the appealing features of quasi-experimental designs. It is possible literally to combine bits and pieces from different standard designs in order to evaluate validity threats in highly specific or unusual situations. This process highlights the appropriate role of research design as a *tool* in which the specific research topic dictates what design investigators should use. Unfortunately, investigators too often use a less appropriate design for a specific research topic simply because they are most familiar with that design. When thoughtfully constructed, however, quasi-experimental designs can provide researchers with

the tools they need to explore the wide array of important topics in sociological study.

### REFERENCES

- Adair, J., D. Sharpe, and C. Huynh 1989 "Hawthorne Control Procedures in Educational Experiments: A Reconsideration of Their Use and Effectiveness." *Review of Educational Research* 59:215–228.
- Blalock, H. M. (ed.) 1985 *Causal Model in Panel and Experimental Designs*. New York: Aldine.
- Campbell, D. T. 1975 "Reforms as Experiments." In M. Guttentag and E. Struening, eds., *Handbook of Evaluation Research*, vol. 1. Beverly Hills, Calif.: Sage.
- , and J. C. Stanley 1963. "Experimental and Quasi-Experimental Designs for Research on Teaching." In N. I. Gage, ed., *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Cook, T. D., and D. T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Cook, T. D., and W. R. Shaddish 1994 "Social Experiments: Some Developments over the Past Fifteen Years." In L. W. Porter and M. Rosenzweig, eds., *Annual Review of Psychology* 45:545–580.
- Costner, H. L. 1971 "Utilizing Casual Models to Discover Flaws in Experiments." *Sociometry* 34:398–410.
- Dwyer, J. H. 1983 *Statistical Models for the Social and Behavioral Sciences*. New York: Oxford University Press.
- Eye, A. (ed.) 1990 *Statistical Methods in Longitudinal Research*, vols. I and II. San Diego, Calif.: Academic.
- Kessler, R. C., and D. G. Greenberg 1981 *Linear Panel Analysis: Models of Quantitative Change*. New York: Academic.
- Kidder, L. H., and J. M. Judd 1986 *Research Methods in Social Relations*, 5th ed. New York: Holt, Rhinehart, and Winston.
- Kline, R. B. 1998 *Principles and Practice of Structural Equation Modeling*. New York: Guilford.
- Lang, E. 1992 "Hawthorne Effect." In E. Borgatta and M. Borgatta, eds., *Encyclopedia of Sociology*, 1st ed. New York: Macmillan.
- Lord, F. M. 1967 "A Paradox in the Interpretation of Group Comparisons." *Psychological Bulletin* 68:304–305.
- Mayo, E. 1933 *The Human Problems of an Industrial Civilization*. New York: Macmillan.
- Nunnally, J. 1975 "The Study of Change in Evaluation Research: Principles Concerning Measurement, Experimental Design, and Analysis." In M. Guttentag and R. Struening, eds., *Handbook of Evaluation Research*. Beverly Hills, Calif.: Sage.



- Nunnally, J. C., and I. H. Bernstein 1994 *Psychometric Theory*, 3rd ed. New York: McGraw-Hill.
- Orne, M. 1962 "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications." *American Psychologist* 17:776-783.
- Roethlisberger, F., and W. Dickson 1939 *Management and the Worker*. New York: John Wiley.
- Rogosa, D. 1985 "Analysis of Reciprocal Effects." In T. Husen and N. Postlethwaite, eds., *International Encyclopedia of Education*. London: Pergamon Press.
- 1988 "Myths about Longitudinal Research." In K. W. Schaie, R. T. Campbell, W. Meredith, and S. C. Rawlings, eds., *Methodological Issues in Aging Research*. New York: Springer.
- Rosenthal, R., and R. L. Rosnow 1969 *Artifacts in Behavioral Research*. New York: Academic.
- Shadish, W. R., T. D. Cook, and D. T. Campbell in preparation "Experimental and Quasi-Experimental Designs for Generalized Causal Inference." Boston: Houghton-Mifflin
- Thistlethwaite, D. L., and D. T. Campbell 1960 "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment." *Journal of Educational Psychology* 51:309-317.

KYLE KERCHER  
KARL KOSLOSKI

# R

---

## RACE

The study of race and race relations has long been a central concern of sociologists. The assignment of individuals to racial categories profoundly affects the quality and even the length of their lives. These assignments are ostensibly made on the basis of biological criteria, such as skin color, hair texture, and facial features. Yet the biological meaning of race is so unclear that many social and natural scientists argue that *race*, as a biological phenomenon, does not exist. Others take the position that while different races exist, extensive interbreeding in many societies has produced large numbers of people of mixed ancestry. The assignment of these people to racial categories depends on social, rather than on biological, criteria. Thus the social consequences of biologically inherited traits is the fundamental issue of the sociological study of race.

### BIOLOGICAL CONCEPTIONS OF RACE

While the terms *race* and *ethnicity* are often used interchangeably, social scientists assign them distinct meanings. Scholars differ on the precise definition of ethnicity, but these definitions usually include some or all of the following criteria. First, ethnic groups are extended kinship groups, although kinship may be defined loosely, as based on a common homeland rather than common ancestry. Second, coethnics share a distinctive culture, marked by differences ranging from language and religion to styles of dress or cooking. A

distinctive culture need not be a matter of everyday practice, however. It may be primarily symbolic, as when a group's traditional language is no longer widely used, or its religious observance is confined to holidays. Third, an ethnic group shares a common history, in which key events such as immigration, colonization, and the like form a sense of collective memory. Finally, an ethnic group is marked by self-consciousness, in that its members see themselves as a people, and are seen as such by others (Cornell and Hartmann 1998).

For most of human history, ethnic groups living in close proximity did not differ significantly in physical appearance. Thus the observable biological differences associated with *race* were not used to distinguish friend from foe, and interracial antagonisms were virtually unknown. The rapid, long-distance migration required to bring members of different racial groups together is a comparatively recent phenomenon that was accelerated by trade and the large-scale European exploration and colonial expansion of the sixteenth through the nineteenth centuries (van den Berghe 1981). It was also during this period that Western science assumed a central role in the attempt to understand the natural and social worlds. Thus, as Europeans became aware of peoples who differed from them in culture and appearance, the concept of race entered the popular and scientific vocabularies as a means of classifying previously unknown groups.

Not content merely to classify people into racial groups, nineteenth- and early-twentieth-century scientists attempted to sort these groups into

a hierarchy. Darwin's theory of evolution, which holds that species are engaged in a struggle for existence in which only the fittest will survive, was gaining widespread acceptance during this period. Herbert Spencer, William Graham Sumner, and other early social theorists extended this evolutionary argument, suggesting that different social groups, including races, were at different states of evolution; the more advanced groups were destined to dominate groups less "fit." This idea, called *social Darwinism* (which Darwin himself did not support), provided justification for European imperialism and for America's treatment of its racial minorities.

Building on the notion that some races were at a more advanced stage of evolution than others, a number of scientists tried to measure differences between the races, especially in the area of intelligence. The first intelligence test was developed by Alfred Binet and Theodore Simon in 1905. Modified versions of this test were administered to approximately one million American soldiers in World War I, and the results were used to argue that there were large, genetically determined differences in intelligence between blacks and whites. Such a conclusion implied that blacks could not benefit from education to the extent that whites could; these findings were then used as a justification for the inferior education made available to blacks.

Binet himself rejected the notion that intelligence was a fixed quantity determined by heredity, or that intelligence could be measured with the kind of precision claimed by other intelligence testers, especially in the United States. Furthermore, other scholars demonstrated that the early tests were heavily biased against members of certain class, ethnic, and racial groups, including blacks. While the *average* scores of blacks have tended to fall below the average scores of whites, greater variation occurs within each group than between the two groups; that is, many blacks outscore many whites.

Charles Murray and Richard Herrnstein argue in *The Bell Curve* (1994) that much of the gap between black and white average scores can be attributed to heredity, rather than to environmental influences. These writers use an extensive array of studies on race and intelligence to support their claim. Yet their methods and conclusions have

been roundly attacked by leading scholars in the field, some of whom contend that intelligence is multidimensional, and cannot therefore be summarized in a single test score. Others point out that since intelligence tests measure academic achievement rather than innate potential, the impoverished background and substandard education of some African Americans offers a reasonable explanation for their lower average scores. Research has repeatedly failed to demonstrate that racial groups differ in terms of their innate capacity for learning. Today, therefore, the vast majority of social scientists reject the idea that any one race is superior in intelligence or any other ability, at least to the extent that such abilities are determined by heredity. (For interesting accounts of the race-intelligence controversy, see Gould 1981; Fraser 1995.)

Controversy continues also on the subject of race itself. In the nineteenth century the concept was defined quite loosely, and the idea was widely held that people of similar appearance but different nationalities constituted different races. As recently as World War II it was not uncommon to hear people speak of the "British race," the "Jewish race," and so on. Some physical anthropologists narrowed the concept to include three main racial groups: the Negroid, Caucasoid, and Mongoloid, or black, white, and Asian races.

Others argue that human populations have always exhibited some degree of interbreeding, and that this has greatly increased during the last few centuries, producing large groups of people who defy such racial classification. "Pure races" have probably never existed, and certainly do not exist now. According to this thesis, race is a cultural myth, a label that has no biological basis but is attached to groups in order to buttress invidious social distinctions. (For an interesting discussion on the biological meaning of race, see Begley 1995; the following section owes much to this work.)

### IS RACE A MYTH?

At the most basic level, biologists sort organisms into species. A species is essentially a breeding boundary, in that it defines the limits within which interbreeding can occur. Thus golden retrievers can be bred with standard poodles, but not with pigs or cats. By this fairly straightforward criterion, all humans, regardless of appearance, belong to the same species. The difficulty arises when one

attempts to identify subspecies, the technical equivalent of races. In some species, this is relatively simple because their distinctive traits are “concordant.” That is, the same subgroups are produced using any of a number of traits: a golden retriever can be distinguished from a standard poodle on the basis of fur color and texture, ear shape, or body type. Among humans, however, identifying subspecies is not so simple, because the traits associated with human subpopulations are nonconcordant. In short, using one trait will result in one set of “racial” categories, while another trait will produce an entirely different set.

Consider the traits commonly used to divide humans into the three conventional races. These traits include skin color, hair color and texture, and facial features. Asians are usually identified primarily by the epicanthic eye fold, yet if this criterion were applied consistently, the San (Bushmen) of South Africa would be considered Asian. And while skin color certainly helps distinguish Swedes from the Masai of East Africa, it also distinguishes Swedes from Turks, both of whom are considered “white,” and the Masai from the San, whose olive complexion more closely resembles the Turk’s than the much darker Masai’s.

Humans are visual creatures, so that in categorizing others, we fixate on differences of appearance. But these criteria are biologically arbitrary; other, less obvious traits associated with human subpopulations might just as easily be used. A common anatomic trait among Asians is front teeth that are “scooped out,” or shovel shaped, in the back. Yet Swedes and Native Americans also share this trait, so we could divide the species into one race with shovel-shaped incisors, and one without. Considering biochemistry, some peoples produce lactase (an enzyme that aids milk digestion) into adulthood, while others do not. A “lactase-positive race” would include northern and central Europeans, Arabs, northern Indians, and many West Africans, while other Africans, Native Americans, southern Europeans, and Southeast Asians would be in the “lactase-negative race.” Genetics multiplies the possibilities even further: antimalarial genes, including but not limited to the sickle cell gene, could be used to distinguish a “malaria-resistant race” (in which Greeks and Italians would

be grouped with Southeast Asians, New Guineans, and tropical Africans) from the “malaria-susceptible race” (which would place Scandinavians with the Xhosa of South Africa). Because these various traits are nonconcordant, classifying the human species on the basis of one will produce a set of “races” entirely different from the set based on another trait.

Biologically speaking, then, all such classification schemes are both arbitrary and meaningless. The genetic variation contained *within* any identifiable human subpopulation, including the conventional “races,” is vastly greater than the variation *between* populations. That is, any two Asian people are likely to have less in common than either has with a randomly chosen white person. To put it in slightly different terms, Harvard biologist Richard Lewontin once observed that if a holocaust wiped out everyone on earth except a small tribe deep in the New Guinea forest, almost all the genetic variation found in the four (now five) billion people currently on earth would be preserved (cited in Gould 1981). Grouping people into racial categories tells us nothing about how biologically related they are. It tells us only that we perceive them to share some trait that we humans have chosen to consider important.

If racial categories tell us nothing meaningful about biology, however, they tell us a great deal about history, as we can see from another argument against the traditional view of race: that individuals’ racial identification can change as they move from one society to another. Americans are accustomed to thinking of black and white as two separate categories, and assigning people with any African ancestry to the former category. This is the “hypodescent rule,” in which the offspring of a mixed union are assigned to the lower-ranked group. In Brazil, however, black and white are poles on a continuum, and individuals can be placed at any point on that continuum, depending on their facial features, skin color, and hair texture. Even siblings need not share the same identity, which also to some extent depends on social class: the expression “money bleaches” reflects the fact that upward mobility can move a person’s racial assignment closer to the white end of the

continuum (van den Berghe 1967). Thus, a black American who is light skinned and well to do may find that in Brazil he is not considered “black” at all, and may even be labeled “white.” Should he go to South Africa instead, our light-skinned black American would be neither black nor white, but “coloured,” the term that denotes a person of mixed ancestry in that society.

Finally, consider this consequence of the hypodescent rule: in America, a white woman can give birth to a black child, but a black woman cannot give birth to a white child. This convention is biologically nonsensical and arbitrary; it can only be understood historically. In the United States, hypodescent was carried to the extreme of the “one-drop rule,” in which one drop of African blood was enough to designate a person as black. This practice evolved out of a desire to maximize the profitable slave population, and to maintain the “purity” of the white race. Clearly, the racial categories commonly used in America do not reflect an underlying biological reality, but rather the more grim chapters of our history. This point has important implications, as we can see by returning to the debate over the relationship between race and intelligence. If the precise nature and meaning of intelligence remains unclear, and if race itself has no biological significance at all, then how are we to interpret a statistical association between “race” and “intelligence”? It becomes little more than a mathematical exercise, yielding information of dubious value.

### SOCIAL CONCEPTIONS OF RACE

While race may lack biological significance, it does have tremendous social significance. Sociologist W. I. Thomas’s famous dictum is certainly true of race: “If men define situations as real, they are real in their consequences” (quoted in Coser 1977, p. 521). Racial distinctions are meaningful because we *attach* meaning to them, and the consequences vary from prejudice and discrimination to slavery and genocide. Since people believe that racial differences are real and important, and behave accordingly, those differences become real and important. Hitler, for example, believed that Jews

constituted a distinct and inferior race, and the consequences of his belief were very real for millions of Jews. Thus the major questions confronting sociologists who study race relations concern the social consequences of racial categorization. To what degree are different racial and ethnic groups incorporated into the larger society? How can we account for variations in the economic, political, legal, and social statuses of different groups?

American sociologists have found their own society to be a natural laboratory for the study of these issues. The United States has a wide variety of racial and ethnic groups, and some of these have been more successful in American society than others. Within any group there is substantial variation in economic achievement, and the success of “model minorities” is often exaggerated. Still, considered as groups, Jews and the Japanese have been more successful in America, in material terms, than have blacks and Mexicans. One explanation for these differences that has found some acceptance both within and outside scientific circles is that the cultures and values of these groups differ. Some groups’ values are believed to be more conducive to success than others. Jews, for example, have traditionally been seen as valuing scholarship and business acumen; as a result they have worked hard in the face of discrimination, educated their children, and pulled themselves up from poverty. African Americans, by contrast, allegedly lacked these values; the result is their continued concentration in the poor and working classes.

Most sociologists reject this argument, which Stephen Steinberg (1981) refers to as the “ethnic myth.” Steinberg argues that this line of reasoning is simply a new form of social Darwinism, in which the fittest *cultures* survive. A closer look at the experiences of immigrants in America (including African Americans) reveals that not all immigrant groups start at the bottom; some groups arrive with the skills necessary to compete in the American labor market while others do not. Furthermore, the skills possessed by some groups are in high demand in the United States, while other groups find fewer opportunities. Thus Steinberg argues that the success of an immigrant group

depends on the occupational structure of its country of origin, the group's place in that structure, and the occupational structure of the new country.

Steinberg uses the case of American Jews to support his argument. In terms of education, occupation, and income, Jews have been highly successful. Thirty-six percent of the adult Jewish population had graduated from college in 1971, compared to 11 percent of non-Jews. Seventy percent of Jews were in business or the professions, compared with roughly a third of non-Jews. The median family income of Jews in 1971 was \$14,000, approximately 30 percent more than the average American family. Again, it is possible to overstate Jewish success, since many Jews are still poor or working class; middle-class Jews are concentrated in small business and the professions, and are nearly absent from corporate hierarchies. Furthermore, Jews have experienced a great deal of economic and social discrimination. Nevertheless, when compared with other ethnic and racial groups in America, they have been quite successful.

This success, Steinberg argues, is attributable in part to the origins of Jewish immigrants, most of whom came from Russia and eastern Europe, and arrived in the United States in the late nineteenth and early twentieth centuries. Since Jews in eastern Europe could not own land, they tended to live in cities; even those who lived in rural areas were mostly merchants and traders, rather than farmers. The urban concentration and above-average literacy rates of Jews affected their occupational distribution: 70 percent of Russian Jews worked as artisans or in manufacturing or commerce in 1897; even unskilled Jews worked in industrial occupations. Sixty-seven percent of Jewish immigrants who arrived in America between 1899 and 1910 were skilled workers, compared to 49 percent of English immigrants, 15 percent of Italians, and 6 percent of Poles.

Furthermore, Jewish immigrants were disproportionately represented in the garment industry, which was growing at two to three times the rate of other American industries. Jobs in the garment industry were better paid than other industrial jobs, and Jews, with their higher skill level, tended to have the better-paid jobs within the industry.

The garment industry also offered unusual opportunities for individual entrepreneurship, since little capital was required to start a small clothing business.

In sum, Jewish immigrants did well in America because they brought industrial skills to an industrializing country. Although the majority of Jewish immigrants arrived with little money and encountered widespread discrimination, American industry could not afford to ignore them completely. Steinberg concludes that while a case can be made that Jews have traditionally valued educational and occupational achievement, and that this contributed to their success, Jews do not hold a monopoly on these values. Furthermore, if they had encountered an occupational structure that offered no hope for the fulfillment of these aspirations, Jews would have scaled their goals down accordingly.

The inability of other racial and ethnic groups to match the success achieved by Jewish Americans has also been attributed to the cultures and values of those groups. Glazer and Moynihan (1970), for example, blame the persistent poverty of blacks on "the home and family and community. . . . It is there that the heritage of two hundred years of slavery and a hundred years of discrimination is concentrated; and it is there that we find the serious obstacles to the ability to make use of a free educational system to advance into higher occupations and to eliminate the massive social problems that afflict colored Americans and the city" (pp. 49, 50). Yet, as Gutman (1976) has shown, the black family emerged from slavery relatively strong and began to exhibit signs of instability only when blacks became concentrated in urban ghettos. Furthermore, for generations after emancipation, blacks faced extreme educational and employment discrimination; the notion that a free educational system provided a smooth path to the higher occupations is simply inconsistent with blacks' experience in America.

Most sociologists tend, like Steinberg, to locate the cause of African Americans' poverty relative to white immigrant groups in the structure of opportunity that awaited them after slavery. The South was an economically backward region where

blacks remained tied to the land and subject to conditions that were in many cases worse than those they had known under slavery. The vast majority of white immigrants settled in the North, where industry provided jobs and taxpaying workers provided schools. The more agricultural South had fewer educational opportunities to offer blacks or whites. Immediately after the Civil War, when they were provided access to education, blacks flocked to southern schools. This opportunity was short lived, however, since the scarcity of educational resources made it advantageous for whites to appropriate the blacks' share for themselves, a temptation they did not resist.

By the time large numbers of blacks migrated north, the industrial expansion that had provided so many jobs for immigrants was on the wane. Moreover, the newly freed slaves did not have industrial skills and were barred from industrial occupations. Given the generations of social, economic, political, and legal discrimination that followed, and the fact that blacks did take advantage of the opportunities that presented themselves, it is unnecessary to call on "inferior values" to explain the difference in achievement between African Americans and white immigrants. (For an interesting comparison of the struggle of blacks in the postbellum South and the North to that of black South Africans, see Frederickson 1981; for a comparison of the conditions faced by U.S. blacks and white immigrants, and the effects of these differences on each group's success, see Lieberman 1980.)

### CONCLUSION

Ever since Darwin proposed that the evolutionary process of natural selection ensures that only the fittest species survive, social science has been bedeviled by the notion that some human groups, especially races, are more biologically or culturally fit than others. This extension of Darwin's principle to competition for survival *within* the human species, especially when applied to industrial or postindustrial societies, cannot withstand close scrutiny. While human subpopulations have evolved certain traits such as malaria resistance and the

retention of lactase into adulthood as adaptations to environmental conditions, these physical traits do not sort our species into consistent categories, and they are hardly relevant to performance in today's school or workplace.

Furthermore, cultural differences between groups can be identified, and these differences may have economic consequences, but they are more likely to reflect a group's historical experiences than the value its members attach to economic success. Thus, the current trend in sociology is to explain differences in the success of racial and ethnic groups in terms of the economic and political resources possessed by those groups, and by the groups with whom they are in competition and conflict.

One reason for the longevity of the biological and cultural forms of social Darwinism may be that for many years most natural and social scientists have been white, and middle class to upper class. While the objective search for truth is the goal of the scientific enterprise, race is an emotionally and ideologically loaded concept, and even the most sincere humanitarians have been led to faulty conclusions by their own biases. An important prospect for the advancement of the scientific study of race, then, is the recruitment of new scholars with a wide diversity of ethnic, racial, and national backgrounds. This increasing diversity will help to broaden the exchange of ideas so necessary to scientific inquiry, and will yield an understanding of race that is more balanced and less subject to bias than it has been in the past.

### REFERENCES

- Begley, Sharon 1995 "Three Is Not Enough." *Newsweek* (February 13):67-69.
- Cornell, Stephen, and Douglas Hartmann 1998 *Ethnicity and Race*. Thousand Oaks, Calif.: Pine Forge.
- Coser, Lewis A. 1977 *Masters of Sociological Thought*, 2nd ed. New York: Harcourt Brace Jovanovich.
- Fraser, Steven (ed.) 1995 *The Bell Curve Wars*. New York: Basic Books.
- Frederickson, George M. 1981 *White Supremacy*. New York: Oxford University Press.

- Glazer, Nathan, and Daniel Patrick Moynihan 1970 *Beyond the Melting Pot*. Cambridge: Massachusetts Institute of Technology Press.
- Gould, Stephen Jay 1981 *The Mismeasure of Man*. New York: Norton.
- Gutman, Herbert 1976 *The Black Family in Slavery and Freedom*. New York: Pantheon.
- Herrnstein, Richard J., and Charles Murray 1994 *The Bell Curve*. New York: Free Press.
- Lieberson, Stanley 1980 *A Piece of the Pie*. Berkeley: University of California Press.
- Steinberg, Stephen 1981 *The Ethnic Myth*. New York: Atheneum.
- van den Berghe, Pierre L. 1967 *Race and Racism*. New York: John Wiley.
- 1981 *The Ethnic Phenomenon*. New York: Praeger.

SUSAN R. PITCHFORD

## RACISM

See Discrimination; Prejudice; Race; Segregation and Desegregation.

## RAPE

See Sexual Violence and Exploitation.

## RATIONAL CHOICE THEORY

Rational Choice theory is typically seen as the use of economic reasoning in contexts that were traditionally the concern of disciplines other than economics, especially of political science, sociology, and anthropology. If we take a more nuanced historical view, however, we might as soon see mainstream economics as the stepchild of the kind of reasoning about larger social institutions, norms, behaviors, and so forth that was central to the Scottish Enlightenment in the works of David Hume, Adam Smith, and many others. The genius of these thinkers was to make sense of such institutions, norms, and so forth as the products

of individuals acting from their own private incentives. Their concern was that of James Coleman (1990), to explain macrophenomena from microchoices. Through most of the past two centuries, economists increasingly focused such reasoning on explaining the nature and working of the market, for example on prices and conditions for an equilibrium of supply and demand, and the earlier concern with broader sociological issues faded. The great classical economists, such as Smith, Alfred Marshall, and Vilfredo Pareto, were interested not only in the market but also in broader social institutions and practices. Much of their work can readily be counted among the great contributions to sociology in their eras.

Contemporary rational choice theory represents a resurgence of such earlier efforts. The efflorescence of such theory in our time has followed on the development of game theory by John von Neumann and Oskar Morgenstern, Kenneth Arrow's demonstration that individual preferences do not aggregate into analogous collective preferences, and Anthony Downs's analysis of democratic participation. The largest bodies of contemporary work are those on the study of group behavior, under the rubric of Mancur Olson's logic of collective action, and on political participation, which in part can be seen as merely a special case of collective action. Perhaps the fastest-growing area of inquiry today is in the analysis of institutions, much of it focused on historical institutions, as in the work of Douglass North.

A cognate area is social exchange theory, which had its contemporary origins well before the works of von Neumann, Arrow, and Downs. Its origins were also distinctively sociological, as in the work of George Homans, and anthropological.

A seemingly cognate area is economic sociology, although much of economic sociology is like the bulk of voting studies in that it takes the form of simple correlations of behaviors with personal characteristics—gender, age, ethnicity, education, occupation, nationality, religion, and so on. Much of economic sociology might therefore be counted as behavioralist. Rational choice analysis at least implicitly assumes intentionality. Hence, in effect,



rational choice theory is contrary to the behavioralism of much of sociology, especially political sociology, in the mid-twentieth century. The point of behavioralism is to avoid the use of mental or intentional explanations of behavior, to treat the mind as virtually a black box. Rational choice theory imputes preferences and intentions to actors. In part, of course, these might derive from or be explained by such sociological characteristics as race, gender, or religion. Although there is a large field of behavioral economics, economics in general was not centrally influenced by the behavioral movement just because it is largely about preferences and intentions. Indeed, much of the work on behavioral economics is directed primarily at establishing the nature and content of preferences and preference functions. It gives measures of the preferences that might go into intentionalist accounts of behavior.

Although much of rational choice theory explains behavior as a response to interests, such theory need not be so narrowly restricted. In its more general form, it explains behavior as the product of preferences, which can cover virtually anything from values to interests. For example, holding all else constant, you might prefer a higher to a lower income. But you might prefer a lower income with peace to a higher income in a state of war, even if your safety and livelihood are not at risk in the war. In many contexts, however, interests seem to be adequate to explain behavior, often because other values are not at stake in the behavior to be explained or because they are substantially less important than interests. Rational choice theory is commonly most compelling in contexts in which interests are predominant. In part, this is because interests can often be more systematically imputed to relevant actors than can other values, although this is not always true. For example, in a given population, particular religious values might regularly trump concern with interests to some substantial degree in some aspects of life. Even when we might suppose other values are very important, however, we might also suppose that analyzing the force of interests gives us a clear baseline for then coming to understand the import or weight of these other values in explaining behavior.

The value theory of rational choice theory is essentially the utility theory that has been developed over the past few centuries by economists and others. It is sometimes asserted that this is an empty value theory and that we can put almost anything we wish into utility functions. For example, I can put your pleasures or various normative commitments in my own utility function. While this is technically correct, most of the major results of rational choice analysis turn on the use of utility functions that are about as spare as we can imagine. They include nothing more than interests, which are conceived to be resources, such as money and time. The results surveyed here virtually all depend only on such simple utility functions. Or occasionally, in a somewhat fuller version, they require inclusion of some of the pleasure one gets from various consumptions, as in the account below of the norm of conformity to neighborhood tastes and manners.

Rational choice theory has been applied to so many diverse issues that a full survey of its characteristic results would be exhausting. I will therefore take up several applications, some of them especially important both in establishing rational choice theory and in recasting the nature of major problems that had already long been the focus of much research. I will first discuss the main methodological or fundamental theories behind rational choice theory. In applying such theory, I will begin with the grandest of sociological issues: the problem of social order. Then I will take up two major areas of research that got the contemporary field of rational choice theory under way: the study of group action and the corollary study of political participation. The first of these is historically a major focus of sociological research, while the second has naturally been the special domain of political science. Then I will take up three efforts that show the breadth of the approach. These are the analyses of institutions, norms, and functional explanation.

### **GAME THEORY**

The methods of rational choice theory are essentially the methods of economics, including standard equilibrium analysis, price theory, econometrics, and game theory. Game theory is less a theory

<b>Game Theory</b>		
<b>GAME 1. PURE CONFLICT</b>		
	<b>COLUMN</b>	
<b>Row</b>	<b>I</b>	<b>II</b>
I	1,2	2,1
II	2,1	1,2
<b>GAME 2. PURE COORDINATION</b>		
	<b>COLUMN</b>	
<b>Row</b>	<b>I</b>	<b>II</b>
I	1,1	2,2
II	3,3	1,1
<b>GAME 3: PRISONER'S DILEMMA OR EXCHANGE</b>		
	<b>COLUMN</b>	
<b>Row</b>	<b>Cooperate</b>	<b>Defect</b>
Cooperate	2,2	4,1
Defect	1,4	3,3

Table 1

than a format for presenting the array of choices and outcomes that face two or more actors whose outcomes are determined by the joint choices or actions of all of them. Games can be represented in many forms. In the matrix form, each player has a set of choices or strategies, and outcomes are determined by the intersection of the strategy choices of all players. Games in which two players have two strategies each are called two-by-two games. Such games can be pure coordination, pure conflict, or a mixture of these two, as represented in Games 1–3, in Table 1. In each of these games, Row has two strategy choices available and Column also has two choices. When both have chosen their strategies, an outcome is determined. The payoffs in each outcome are given in ordinal terms. 1 is the best payoff, 2 next best, and so forth. The first payoff in each outcome is to Row and the second payoff is to Column.

In the pure conflict game, one player can be better off only if the other is worse off. In the pure coordination game, both players achieve their best payoff together. Mixed games are commonly called

*mixed-motive games*. There are many different types. The one represented here is the prisoner’s dilemma, which is surely the most studied of all simple games, probably because it represents ordinary exchange and is therefore ubiquitous and central in social interaction. In the prisoner’s dilemma both players can be made better off together in the move from (3,3) to (2,2), so that the game has a strong element of coordination. But Row is made better off while Column is made worse off in the move to (1,4) from any other outcome, so that the game also has a strong element of conflict.

Unfortunately, the (3,3) outcome of a prisoner’s dilemma is an equilibrium in the sense that we cannot move to the Pareto superior (2,2) outcome through individually beneficial or neutral moves, because my change of strategy from defecting to cooperating while you continue with your strategy of defecting makes me worse off. To move to the (2,2) outcome requires joint action. The prisoner’s dilemma is the only one of the seventy-eight ordinarily distinct two-by-two games that has a Pareto-dominated equilibrium. Its solution therefore commonly requires incentives from outside a single play of the game. The incentive can be from the benefits of cooperative play in an iterated series of plays of the game or from external enforcement by other parties, as under a legally binding contract.

### SOCIAL ORDER

There are three grand schools on social order. One of these is the conflict school associated with Plato’s Thrasymachus, Karl Marx’s class theory, Ralf Dahrendorf, and many others. Another is the shared-values school of John Locke, Émile Durkheim, and Talcott Parsons. A third is the school of exchange theory associated with Adam Smith, George Homans, and others. These can be characterized by the three classes of games represented above. Because there can be all these—and many other—classes of interactions in society, all these theories are partially right about social order. There is a fourth plausible account of much of the order we see, an account that fits the coordination game. We do not coordinate only because we share values; we can coordinate merely to stay out of

each other's way while we pursue our different values. (Indeed, in some sense we can share values—we both want the same thing—in such a way as to have severe conflict.)

One of the simplest of social coordinations is the coordination on driving that makes traffic flow much faster and with less mishap. In North America, we all drive to the right. That is merely one of two possible conventions that could work equally well. The other convention is that we all drive to the left; this convention is followed in the United Kingdom and many nations of the British Commonwealth. It is merely coordination that resolves the traffic problem. But when you and I coordinate on this convention, we do so not because we share any substantial values. We might each be utterly self-serving. The only value we share is to keep others out of our way as much as possible. This is characteristic of much of social order in contemporary liberal societies, in which rampant individualism and great diversity of values might seem to lead to great and disruptive conflict. Instead, it commonly gets channeled in ways that avoid conflict but that could hardly be called cooperative in the sense in which you and I might cooperate in building a house.

Coordination without confluence of values makes the problem of social order seem relatively simple, as it must be for many activities in which we spontaneously achieve order without either the imposition of power by authorities, as is required for the conflict school, or the relatively deliberate cooperation of the exchange school. While such coordination is not the original discovery of rational choice theory or game theory, it is made far more perspicuous by these because these give it a structure of motivation that makes sense of it in many contexts, just as they make the other major schools of social order clearer. Hume presented a clear account of coordination and convention in many social contexts in which stabilizing expectations is fundamentally important to social order. Game theory provides a framework for characterizing the interactions that we must govern if we are to achieve order. Most forcefully, perhaps, game theory suggests why we cannot ground an account of social order in merely one of the traditional

schools, because it displays the greater complexity of the forms of interaction that must be ordered.

### GROUPS AND COLLECTIVE ACTION

There is a tradition, still alive today, in which it is supposed that if the members of a group share an interest in some result, they will act to provide themselves that result. For example, Karl Marx's theory of revolution requires this simplistic assumption coupled with his account of class consciousness. Against this tradition, Olson (1965) argued that, if our group interest requires for its achievement that each (or at least many) of us make a personally costly contribution to its achievement, then we commonly do not have individual incentive to act for the collective good. Each of us bears the full cost of our own contribution but receives only a minuscule part of the small piece of the larger collective good provided by that contribution. Typically, therefore, the collective benefit to me of my contribution to our collective provision will be less than the cost to me of that contribution. Hence, if I am narrowly self-interested, I will not make a contribution to the collective good but will hope merely to free ride on the provision that results from the contributions of others. If all other members of my group have my structure of interests, none of us will contribute and our collective good will not be provided. If we could vote to have ourselves compelled to contribute, we might all vote to do so. But if we must voluntarily contribute, none of us might do so. This is the logic of collective action.

Olson used standard micro-economic analysis to demonstrate this logic. He modeled the problem as an instance of Paul Samuelson's theory of public goods, in which the efficient price of access to the public good would be zero even though, at that price, there would be no voluntary supply of the good. Olson's logic can as well be demonstrated game theoretically as an instance of a large-number prisoner's dilemma (Hardin 1982). As in the discussion of game theory above, individually motivated action would not lead us out of the dismal equilibrium of no cooperation, even though the Pareto superior outcome in which all would

contribute might be enormously superior to the status quo equilibrium in which all defect.

A very large literature has been directed at explaining the collective action and the achievement of collective benefits that we see despite this logic. For example, we apparently see a great deal of collective action in the form of social movements that sometimes entail great individual costs and even severe individual risks. Much of this literature supposes that people are motivated by commitments beyond self-interest, such as social and moral commitments, but much of it supposes that there are incentives apart from the direct logic of collective action. For example, there may be specific personal benefits corollary to the collective benefit. Alternatively, our group might be provided its collective good but not through spontaneous individual actions. Rather, an entrepreneur might see the possibility of making a career out of leadership in providing our group its collective benefit. Such an entrepreneur might especially arise if our group's good could be provided by government without requiring our voluntary cooperation.

#### **POLITICAL PARTICIPATION AND DEMOCRACY**

There are two main lines of theory on political participation in a democratic system. One of these began with Kenneth Arrow's ([1951] 1963) impossibility theorem. That theorem essentially says that collective preferences cannot be modeled simply on individual preferences. We might each individually have well-ordered preferences over all the choices we face, and yet we might collectively not have such well-ordered preferences. Indeed, we can generally expect not to have such collective preferences over any complex realm of choice, such as we often face in normal politics over the large array of policies at issue. This has far broader implications than merely that democratic choice may have problems. It is an important and broadly interesting instance of the fact that the imputation of various characteristics of individuals, such as their pattern of preferences, to groups composed of individuals is a fallacy of composition.

The typical implication of Arrow's theorem goes back to the Marquis de Condorcet and to Lewis Carroll. It is that our collective preferences may cycle through some set of possibilities. For example, in majority votes we may collectively prefer A to B, B to C, and C to A. If our collective preferences were as well behaved as our individual preferences, the fact that we prefer A to B and B to C would entail that we prefer A to C. In majority votes, the majority who prefer A to B can be different from that who prefers C to A. For example, my preferences may be  $A > B > C$ ; yours may be  $B > C > A$ ; and a third person's may be  $C > A > B$ . These preferences yield the cycle above if the three of us vote by majority. Many institutional devices, such as legislative practices of opposing new laws against each other before opposing the winner of such a series against the status quo, tend to block any evidence of cyclic preferences; but such a device gives a strong conservative bias to collective choice.

The other main line of rational choice theory of democracy began with Joseph Schumpeter and was developed extensively by Downs (1957). There are two major classes of claims. First, Downs supposed that two parties or candidates in an election face an electorate that is divided along a left-right dimension. If the voters have a normal distribution about some central tendency on this dimension, the two candidates will want to place themselves at the peak of that normal distribution. Hence, the two candidates will tend to have quite similar positions. Second, he supposed each voter faces what is de facto a logic of collective action on whether to vote. Suppose there are some costs involved in casting a vote—waiting on line, traveling in foul weather, and paying a fee for registration. It follows that individual voters should see the election of candidate A over candidate B as essentially a collective good to be provided at individual cost to themselves. It is therefore subject to the perverse logic of collective action, and we should expect that many voters would not vote unless they have motivations that go beyond their own interests. Furthermore, if it is not in my interest even to vote, it is unlikely to be in my interest to learn enough about the issues to vote intelligently.

## INSTITUTIONS

The rational choice analysis of institutions has roots in ancient accounts of the rise of civilization. Among many such accounts in the era of the Scottish Enlightenment is Smith's theory (1978) of the stages of development of society from very primitive, to pastoral, to more nearly modern society in his *Lectures on Jurisprudence* (these lectures were not published in Smith's time and have played little role in the development of such analyses since then). Smith's account turns very clearly on the incentives that individuals have to submit to various forms of social order and, eventually, government. His later analysis of the wealth of nations is itself a theory of one of the grandest of all institutions: the economy of a modern commercial society. Smith argued that the wealth of the nation is a function of the efforts of individuals to do well for themselves.

More recent work has gone in diverse directions. The two main directions are the microanalysis of institutions and the behavior of individuals within them and the more nearly macroanalysis of why certain institutional forms arise and prevail. The microanalysis is applied to a wide variety of institutions, most of them of relatively small scale, such as committee structures and formal organizations of many kinds. The macroanalysis is often broadly historical and is directed at accounting for the rise of economic and other institutions. For example, there is extensive work on the rise of devices for handling trade across a broad array of cultures in the absence of any centralized governmental authorities.

Much of the institutional analysis builds on accounts of transaction costs. According to the so-called Coase theorem, due to Ronald Coase, if there were no transaction costs, property rights assignments would have no effect on overall production. Introducing transaction costs can distort production substantially. Firms sometimes internalize functions for which transaction costs would be high if they had to deal with outside suppliers for those functions, and they externalize functions for which markets work well to reduce transaction costs so that competitive suppliers can drive down

production costs. Attention to the structure of transaction costs therefore can explain much of an economic organization's structure. Attention to changes in markets over time can also explain the evolution of such organizations' structures. While most transaction cost analysis has so far been done by economists, including economic historians, it is also increasingly done by sociologists and others.

North (1990, p. 131) argues that the use of standard neoclassical economic methodologies exacts a heavy price in our effort to understand institutions. Because such understanding must of necessity be developmental, it must include stories of how the institutions came to be what they are. Neoclassical price theory is concerned with allocations at a specific moment in time under particular institutional arrangements. Game theory lends itself more readily to developmental stories, but to some extent we lack the methodology for putting such stories into order.

A clear implication of the transaction cost analysis of institutions is that, once in place, institutions influence incentives and interests, so that one cannot simply take institutions as dependent on rational choice. They are, additionally, shapers of rational choice. This is conspicuously true, of course, for such institutions as those of government and law, part of whose function is specifically to give people incentive to behave in some ways rather than others. It is true far more generally of essentially all institutions that have significant value to us, even institutions whose purpose might be seen as merely to produce certain goods or services. This means that the actual set of institutions we have and the set of individual behaviors we see are partly determined by the order and the era in which institutions have developed. Some part of what is commonly referred to as culture is merely the happenstance dependency of such historical developmental patterns.

## NORMS

Much of the study of norms has been psychological or even psychoanalytic, and perhaps most of it has assumed that the motive for following norms is

essentially normative or otherwise not self-interested. Efforts to explain the rise of norms, however, are often forced to take account of how the interests of at least some people are to act on and to enforce various norms. One way to characterize the problem of creating and maintaining a norm is as a problem of collective action. It would seemingly be in the interest of almost all of us if a certain norm prevailed, but it is in the interest of almost none of us actually to abide by the norm unless there is some sanctioning system to keep us in line. In some theories, norms are morally or psychologically internalized, so that the sanctioning system is internal to the actor. Such theories require a substantial account of just how the internalization works. No doubt, there is some internalization of norms, but many norms must still depend heavily on external sanctioning, either for them to work at all or for them to work very well. If that were not so, we could dispense with institutionalized law.

Against accounts that require external sanctions, it is sometimes supposed that sanctioning has costs, so that sanctioning a violator of a norm runs against the interests of those who would like to see the norm upheld. For many norms, this is apparently not true. For norms of exclusion, I may actually prefer to shun you if you violate our norm. Hence, I sanction you and benefit from doing so. For example, if you do not follow neighborhood norms of using a relevant slang or dressing in certain ways, I might actively prefer not to associate with you, because your behavior makes me uncomfortable. My reticence and that of others in our neighborhood affects you and damages the pleasures you might get from associating with us or even shuts you out of such association.

For universalistic norms the problem is more complex. Some of these, such as the norm of promise keeping or truth telling, are enforced between dyads or small numbers of participants. In these cases, it may commonly be my interest to sanction you by not cooperating with you on other matters if you break your promise to me on some current matter. Hence, these norms are like norms of exclusion in that they can also be backed by sanctions that serve the interest of the sanctioners.

For universalistic norms that govern large-number interactions that are essentially instances of collective action, there may be no sanctioning device that serves the sanctioner's interest, and such norms are, not surprisingly, relatively weak in comparison to dyadic universalistic norms and norms of exclusion (Hardin 1995).

### FUNCTIONAL EXPLANATION

An example of the ways in which rational choice theory is applied to apparently contrary approaches is recent work on functional explanation. This work does not contribute to the older school of functionalism or structural functionalism, as represented in much of twentieth-century anthropology or in the sociology of Parsons and many others. Rather, it reconstructs functional accounts in terms of individual incentives, as did Robert Merton in his effort to be very clear about the logic of functional argument. Oddly, the most important contribution to this new work was intended as a dismissal of functional analysis. Jon Elster (1979) argued that, if the form of functional explanation is properly spelled out, then very few supposedly functional accounts fit that form.

Pared down to its essentials, Elster's account is as follows:

*An institution or a behavioral pattern X is explained by its function F for group G if and only if:*

1. F is an *effect* of X;
2. F is *beneficial* for G;
3. F maintains X by a causal *feedback* loop passing through G.

Many groups that benefit from some behavior on the part of their members induce that behavior through incentives that they give to their members.

As an example of a functional account of a major institution, return to the problem of social order, which is typically governed in substantial part by a legal system. A common view of much of law is that its function is to coordinate us or to facilitate our interactions. Hence in a functional explanation of law, F is coordination, X is law, and

G is our law-governed populace. The feedback loop passing through the populace is that our coordination by law enables us to coordinate to produce still further law to coordinate us still further. Hence, law is functional for us. But we would coordinate in such ways only because it is in our interest to coordinate. Hence, we can explain a major, pervasive, and seemingly all important social institution in functional terms as grounded in the rational choices of the actors.

### CURRENT DIRECTIONS

Perhaps the easiest assessment of where the field of rational choice will go and where it should go is to extrapolate from current trends. Clearly, institutional work looms large for the near term and normative work seems likely to become more important. In both of these developments, one might hope and even urge that research in rational choice take the findings of other approaches seriously. Doing this would entail two quite different programs. The first and simpler program is merely to make extensive use of findings from other approaches. The second and intellectually more challenging program is to attempt to show the complementary relation of various other approaches to rational choice theory—or, alternatively, to demonstrate their incompatibility, which is often asserted but seldom shown. Sometimes, this might even entail showing that some other approach is, at least in some applications, equivalent to rational choice analysis. For example, work that has unpacked the logic of functional explanation often reveals the rationality of actors involved in replicating some supposedly irrational or extrarational behavioral pattern.

Perhaps the greatest challenge to rational choice theory is to fit it to vast bodies of behavioral research that does not focus on individual incentives and intentions. Part of the task here would be to impute incentives and intentions to relevant actors, perhaps by analogy from other studies and contexts. Another part of the task, as in the rational choice analysis of institutions, functional explanation, and norms, is to restructure the problems in ways that make their choice structures clear.

Unfortunately, rational choice theory is less well developed in sociology than in economics and even political science, in part because it is embattled. Despite the heyday of exchange theory in anthropology earlier in the twentieth century, rational choice theory is almost entirely absent from that discipline.

Finally, just because rational choice theory focuses on the incentives for microchoices that produce macroeffects, it is particularly suited to policy analysis. Empirical work on incentive systems and how they work can be put to use in designing policies to change behavior in productive ways. This is little more than common sense in many contexts, but a resolute focus on the relation of microincentives to macroresults is an especially natural part of rational choice analyses.

(SEE ALSO: *Social Exchange Theory*)

### REFERENCES

- Arrow, Kenneth J. (1951) 1963 *Social Choice and Individual Values*, 2nd ed. New Haven, Conn.: Yale University Press.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Cook, Karen S., and Richard Emerson 1978 "Power, Equity, and Commitment in Exchange Networks." *American Sociological Review* 43:721–739.
- Downs, Anthony 1957 *An Economic Theory of Democracy*. New York: Harper.
- Elster, Jon 1979 *Ulysses and the Sirens*. Cambridge: Cambridge University Press.
- Hardin, Russell 1982 *Collective Action*. Johns Hopkins University Press.
- 1995 *One for All: The Logic of Group Conflict*. Princeton, N.J.: Princeton University Press.
- Hirschman, Albert O. 1970 *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, Mass.: Harvard University Press.
- Homans, George Caspar (1961) 1974 *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace Jovanovich.
- Mueller, Dennis C. (ed.) 1997 *Perspectives on Public Choice: A Handbook*. New York: Cambridge University Press.

North, Douglass C. 1990 *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Olson, Mancur, Jr. 1965 *The Logic of Collective Action*. Cambridge, Mass.: Harvard University Press.

Smith, Adam 1978 *Lectures on Jurisprudence* R. L. Meek, D. D. Raphael, and P. G. Stein, eds. Oxford: Oxford University Press. (From notes of Smith's lectures in 1762–1964.)

Ullmann-Margalit, Edna 1977 *The Emergence of Norms*. Oxford: Oxford University Press.

RUSSELL HARDIN

## REFERENCE GROUP THEORY

See Role Theory; Role Theory: Foundations, Extensions and Applications; Self-concept.

## REFUGEES

See Genocide; International Migration.

## RELIABILITY

The reliability of a measured variable has two components, consistency and stability. *Consistency* is the degree to which two measures of the same concept provide the same assessment at the same time; consistency is based on “cross-sectional” research. *Stability* is the degree to which a measure of a concept remains unchanged across time; stability is based on “longitudinal” research. Let us illustrate consistency and stability on the measurement of height.

### THE MEASUREMENT OF HEIGHT

As an example, we often measure how tall people are. Height, how tall someone is, is a measure of distance. In order to measure distance, we establish an arbitrary standard. A common arbitrary standard for measuring distance is the “yardstick.”

The yardstick is 36 inches long, and is broken down into feet, inches, and fractions of inches. Another common measuring rod is the “meterstick.” The meterstick is 100 centimeters long, and is broken down into decimeters, centimeters, and millimeters. If we know how tall someone is in inches, we can calculate how tall he or she is in centimeters, and vice versa. For example, rounding to two decimal places, a 70-inch-tall person is 177.80 centimeters tall (1 inch = 2.54 centimeters;  $70 \times 2.54 = 177.80$ ). Conversely, rounding to two decimal places, if we know that someone is 160 centimeters tall, we also know that that person is 62.40 inches tall (1 centimeter = 0.39 inches;  $160 \times 0.39 = 62.40$ ).

Indeed, the yardstick and the meterstick are highly consistent. With reasonable attention to proper measurement protocol, the correlation between height as measured by the yardstick and height as measured by the meterstick across a sample with sufficient variation in height would be very high. For all intents and purposes, the yardstick and the meterstick are interchangeable; the researcher need not establish their consistency. This leads to the *principle of consistency*:

*If two measures of the same concept are perfectly consistent, they provide identical results. When this is so, the use of multiple measures is needlessly repetitive.*

In this situation, the researcher need only use either the yardstick or the meterstick; using both sticks provides no additional information.

When babies are born, they are usually 18–24 inches “tall.” Parents (and developmental researchers) often measure how tall babies and children are as time passes. This over-time height measurement is a stability assessment. Ordinarily, children grow a rough average of 3 inches per year for 15 years, resulting in most 15-year-olds being between 63 and 69 inches tall. Then female height growth stops while male height growth continues. By the time females are 18 years old, they average about 66 inches tall, while males of the same age average about 72 inches. Their heights then remain roughly stable for the remainder of their adult lives. This leads to the *principle of stability*:



*A measure of a concept is perfectly stable when it provides identical results at different points in time. When this is so, repeated measurement over time is needlessly repetitive.*

Height measurement from year to year provides useful information for children but not for adults. This is because the children grow taller with the passage of time, but adults do not. Elderly people who suffer from osteoporosis (loss of bone density) will become shorter, but this decline in height is slight, compared to their growth when they were children.

Let us now turn to a discussion of how the principles of consistency and stability apply to the measurement of sociological concepts. We will first discuss the protocols of good sociological measurement. Then we will discuss the implications of these protocols for the assessment of the consistency of self-esteem and the stability of alienation.

### PROTOCOLS OF GOOD MEASUREMENT

**The Measurement of Self-Esteem.** Researchers often attempt to measure how well people feel about themselves. Many decades ago, Charles Horton Cooley (1902) and George Herbert Mead (1934) theorized about the concept “self-esteem.” In offering the “looking glass self,” Cooley assumed that people develop a definition of themselves by evaluating what they believe others think of them. Mead differentiated between what a person actually is and what that person believes about himself or herself.

Rosenberg (1965) wished to measure self-esteem as conceptualized by Cooley and Mead. He did so by creating ten questionnaire items, each of which he believed would provide an empirical measure of the concept of self-esteem. His measurement attempt will be discussed in detail later in this paper. For now, let us assume that each of these items positively but imperfectly represents self-esteem. The positive representation implies that the concept “self-esteem” has a positive causal effect on each item. The imperfectness of the representation implies that there are other factors

that also cause each item. Under this condition, none of these ten different measures of self-esteem was nearly as consistent as the yardstick and the meterstick. That is, the correlations among these ten questionnaire items were far from perfect. When this is so, the use of multiple measures is more consistent than the use of any single measure alone. Thus, the *converse principle of consistency*:

*If multiple measures of the same concept provide imperfect assessments of the same concept, then the use of multiple measures is more consistent than the use of any single measure alone.*

Commonly, items presumed to measure sociological concepts do so imperfectly. Therefore, sociological researchers often turn to the use of multiple items in social surveys as indexes to represent concepts. Combining their substantive knowledge of the literature with their clinical knowledge of people who exhibit various aspects of the concept, these researchers design items to represent each of these aspects. Then researchers are faced with the tasks of evaluating the consistency and stability of the items as measures of their respective concepts. In order to do this, good researchers employ a set of protocols of good measurement. We now turn to a discussion of these good measurement protocols.

Good measurement of sociological concepts satisfies the following criteria:

- Clear definitions of concepts.
- Multiple items.
- Clear definitions of items.
- Strong positive interitem correlation.
- Score construction.
- Known groups validity.
- Construct validity.
- Consistency.
- Stability.

These protocols represent the state of the art not only in sociology (Mueller 1997), but also in a

wide variety of other scientific disciplines. A computer search of the literature revealed more than five hundred articles citing these protocols in the 1990s, including: political science (Jackman and Miller 1996); psychology (Hendrix and Schumm 1990); nursing research (Budd et al. 1997); the family (Grandin and Lupri 1997); sports and leisure (Riemer and Chelladurai 1998); computer information systems (McTavish 1997); management (Szulanski 1996); gerontology (Wright 1991); genetics (Tambs et al. 1995); social work (Moran et al. 1995); higher education (Aguirre et al. 1993); market research (Lam and Woo 1997); and preventive medicine (Saunders et al. 1997).

Let us briefly discuss each of these protocols in turn. Then we will focus the attention of the remainder of this paper on the two major focuses of reliability, consistency and stability.

*Clear Definitions of Concepts.* Good measurement protocol requires that each concept be clearly defined and clearly differentiated from every other concept. Good measurement protocol can document that an ambiguous concept is, indeed, ambiguous. Moreover, such protocol may suggest points of theoretical clarification. However, there is no substitute for clear theoretical thinking augmented by a thorough knowledge of the literature and a clinical immersion in the domain of content.

*Multiple Items.* Good measurement protocol requires that each aspect of a concept be assessed using multiple items. A single item, taken alone, suffers from measurement error. That is, the item is, in part, a representation of its respective concept. However, this same item may be a representation of other concepts, of systematic measurement error, and of random content. These other contents are called "error"; they reduce the degree to which the item accurately represents the concept it is designed to measure empirically. The rationale for the use of multiple items revolves around minimizing this measurement inaccuracy. That is, all items designed to measure a concept contain inaccuracies. If a single item is used to measure the concept, the researcher is, in essence, stuck with the specific inaccuracies of the single item. However, if multiple items designed to measure the same concept are used, the inaccuracies of

one item may be offset by different inaccuracies of the other items.

*Clear Definitions of Items.* Good measurement protocol requires that each item be designed to measure one and only one concept. The response categories should be constructed so that the higher the code of the response category, the more positive the respondent's attitude on that concept.

*Strong Positive Interitem Correlation.* When multiple items are designed to measure a single variable, the standard of the research industry has long been that the items should be coded in such a way that the higher the score, the more positive the empirical representation on the concept. Good measurement protocol requires strong positive intercorrelations among items designed to measure a concept. Ordinarily, these intercorrelations are presented in a correlation matrix. A visual inspection of the correlation matrix will be revealing. An item that correlates strongly (e.g.,  $r > .4$ ) with other items will generally emerge as a strong contributor to the reliability of the resulting score; an item that has a zero correlation with other items will not add to the reliability of the score; and an item that inversely correlates with other items (assuming that it has been coded such that the higher the score on the item, the higher the measure of the concept) will detract from the reliability of the score.

To the author's knowledge, the sole exception to this principle was articulated by Curtis and Jackson (1962, p. 199) who argued that "two equally valid indicators of the same concept may. . . be strongly related to one another, or they may be totally unrelated (or negatively related)." The difficulty with the Curtis and Jackson position is that it effectively invalidates the most powerful empirical argument that can be made for multiple items representing a single dimension—that of the equivalence established using convergence. Instead, the author would argue that if two items are unrelated or negatively related to one another, either they represent different dimensions, or they are reflecting a method artifact or both. For a more detailed discussion of this matter, see Zeller and Carmines (1980, p. 77–136) or Carmines and Zeller (1979).

Factor analysis is the major statistical technique designed to describe a matrix of item intercorrelatedness. As such, factor analysis enables researchers to (1) describe a large number of items in terms of a small number of factors and (2) select those items which best represent each of the identified concepts (see Bohrnstedt 1970, p. 96; and Zeller and Carmines 1980, p. 19–46). Items that have high factor loadings on a factor that represents a concept are then retained. These items are then used to construct a score to represent the concept.

In evaluating the degree to which a large set of items represents a small number of theoretical concepts, the application of factor analytic techniques is as much an art as it is a science. This is because there are numerous ambiguities in the measurement setting. The researcher defines one or more concepts and explores the degree to which the factors coincide with the hypothesized concepts. For each item, the researcher wants to know the degree to which it is a function of the concept it was designed to measure, other concepts, method artifacts, and random error.

*Score Construction.* Once the number of factors and which items define which factors has been established, the researcher needs to create scores. One score should be created to represent each concept empirically for each respondent. If the items defining a concept have roughly equal variances, the simplest way to create a score is to sum the items defining the concept. In practice, researchers can tolerate moderate variation in the item variances. For example, if the item variances for a set of Likert items range from, say, .8 to 1.4, summing the items seems to make the most sense. However, if the variation in the items is severe (say from .5 to 2.5), then the researcher should first create standardized scores using the following formula:  $z = (\text{score} - \text{mean}) / \text{standard deviation}$ . The standardized scores have equal variances (i.e., 1); the sum of these standardized scores will create each desired score.

*Known Groups Validity.* Once scores have been constructed, comparisons of scores between groups known to be high and low on the dimensions of the

concept should be made. Known groups validity is established if groups known to be high on the concept have substantially higher scores than groups known to be low on the concept.

*Construct Validity.* Construct validity is intimately related to theory testing. Construct validity involves (1) specifying theoretical relationships among concepts, (2) assessing empirical relationships among scores, and (3) interpreting how the evidence clarifies the validity of any particular measure. For more information on this concept, see Carmines and Zeller (1979, pp. 22–26).

*Consistency.* Good measurement protocol requires that the consistency among items designed to measure a concept should be strong. This means that the correlation between any two items designed to measure the same concept should positively and strongly correlate. We will apply the principle of consistency to Rosenberg's attempt to consistently measure the concept of self-esteem.

*Stability.* Good measurement protocol requires that, if a concept does not change over time, a score designed to measure that concept also does not change over time. The trick of consistency is that the researcher ordinarily does not know whether there is a change in the value of the concept over time. We will apply the principle of stability to the attempt by R. A. Zeller, A. G. Neal, and H. T. Groat (1980) to stably measure the concept of alienation.

When these protocols of good measurement are not followed, the researcher increases the risk of torpedoing the best of conceptual schemes and sentencing them to the intellectual trash heap, whether they belong there or not. High-tech statistical tools, such as structural equation modeling (SEM), make requirements that, by definition, are not present in the measurement development situation (Bollen 1989; Bollen and Long 1993; Hayduk 1987; and Hoyle 1995). That is, SEM requires both strong theory and strong measurement a priori. Indeed, SEM demands that the researcher know beforehand:

- How many factors there are.
- Which items represent which factors.

But these are precisely the major questions that the researcher wants to answer! The end result of

the factor analysis should be that the researcher has inferred how many factors are represented by the items, and which items define which factors.

We now turn to a discussion of the consistency of self-esteem.

### CONSISTENCY OF SELF-ESTEEM

Good measurement protocol requires that the consistency be strong among items designed to measure each dimension of a concept. This means that the correlation between any two items designed to measure the same concept should positively and strongly correlate. Often different measures of the same concept have relatively modest positive intercorrelations.

In constructing the self-esteem scale, Rosenberg created ten items using the response categories “Never true,” “Seldom true,” “Sometimes true,” “Often true,” and “Almost always true.” Five of these were positive items; these items made a positive statement about self-esteem. The other five were negative items; these items made a negative statement about self-esteem. The response categories for the positive items were assigned the values 1 through 5, respectively, such that the higher the score, the higher that respondent’s self-esteem was inferred to be. These positively stated items were:

- “I feel that I have a number of good qualities.”
- “I feel that I’m a person of worth, at least on an equal place with others.”
- “I take a positive attitude toward myself.”
- “I am able to do things as well as most other people.”
- “On the whole, I am satisfied with myself.”

For the five negatively phrased items, a higher score indicated a lower self-esteem. These items had the same response categories as above, but the assignment of values was reversed. That is, the negatively stated items were assigned the values 5

through 1 respectively. That is, a “Never true” response to the item “I wish I could have more respect for myself” was assigned a 5 and an “Almost always true” response to that item was assigned a 1. These five negatively stated items were:

- “I wish I could have more respect for myself.”
- “I feel I do not have much to be proud of.”
- “I certainly feel useless at times.”
- “All in all, I’m inclined to feel that I am a failure.”
- “At times I think I am no good at all.”

Given the reverse scoring for these items, a higher score indicated higher self-esteem. In order to create a self-esteem scale, the scores were summed into a value that ranged from 10 representing the lowest measured self-esteem possible to 50 for the highest possible measured self-esteem.

How consistent are these items? We suggest the following as a *consistency rule of thumb* for a variable to be used in sociological research:

- If  $r$  is above .8, the score is highly consistent.
- If  $r$  is between .6 and .8, the score is modestly consistent.
- If  $r$  is less than .6, the score may not be used in research.

In the author’s research (Zeller and Carmines 1976), interitem correlations among the ten Rosenberg items designed to measure self-esteem ranged from a low of .05 to a high of .58 with a mean  $r$  of .32. These intercorrelations do not meet this rule of thumb. When properly analyzed, however, they will. We now turn to a discussion of the strategy for this analysis that will address this criterion of consistency.

*Split-Half Consistency.* The “split-half” approach to estimating the consistency of items designed to measure a concept is to divide the items into two subscores and calculate the correlation between those subscores. For example, the ten items can be

divided into two subscores of five items each. The resulting split-half correlation between the two subscores provides an estimate of consistency. If the average interitem correlation equals .3, a score created by summing the responses to the ten items into two five-item subscores would have a split-half correlation of .68.

However, it is well known that, given items whose intercorrelations are equal (i.e.,  $r = .3$ ), the greater the number of items, the higher the consistency of a score resulting from those items. Thus, a ten-item score will have more consistency than a five-item score when both scores are made up of items that intercorrelate .3. The split-half reliability correlation, however, does not represent the ten-item score, it is two subscores made up of five items each. Therefore, this split-half correlation will be lower than the actual consistency of the ten-item score.

Two researchers, Spearman (1910) and Brown (1910), independently recognized and solved this statistical estimation problem. Specifically, they noted that the split-half reliability correlation can be adjusted to project what the consistency of a ten-item score would have been if it had been calculated on the basis of two ten-item subscores instead of two five-item subscores. They shared attribution for this solution and called the result the *Spearman-Brown Prophecy*. It is presented in formula (1):

$$r_{xx''} = 2r_{xx'} / (1 + r_{xx'})$$

where

$r_{xx''}$  is the Spearman-Brown Prophecy formula. (1)

$r_{xx'}$  is the split-half correlation coefficient.

Using the example from above, we can see that the Spearman-Brown Prophecy formula projects the consistency of the entire ten-item scale using formula (1) as follows:

$$r_{xx''} = 2r_{xx'} / (1 + r_{xx'}) = (2)(.68) / (1 + .68) = 1.36 / 1.68 = .81$$

This .81 is an unbiased estimate of the consistency of the total score. Applying the above rule of

thumb, such a scale is quite consistent and can be used in sociological research.

In actual research, intercorrelations among score items vary substantially. In the self-esteem example, item intercorrelations varied from .05 to .58. Moreover, the researcher must decide which items to assign to which subscales. One assignment of items to subscales will produce a different reliability estimate than another. When this occurs, the split-half reliability correlation between the two subscales is beset with the *problem of equivalence*: Which items are assigned to which subscales? We now turn to a way of handling variations in intercorrelations among items.

*Equivalence Consistency* The researcher could assign the even-numbered items to one subscore and the odd-numbered items to the other; or items 1, 2, 3, 4, and 5 to one subscore and 6, 7, 8, 9, and 10 to the other; or items 1, 4, 5, 8, and 10 to one score and 2, 3, 6, 7, and 9 to the other. There are many combinations of assignments that could be made. Which one should the researcher use?

Lee Cronbach (1951) solved this dilemma by creating *Cronbach's Alpha*. Cronbach's Alpha uses the average of all possible split-half reliability correlations that are Spearman-Brown projected to the number of items in the score. This is presented in formula (2):

$$\alpha_{xx} = N(\text{Mean } r_{xx'}) / [1 + \text{Mean } r_{xx'}(N - 1)]$$

where

$\alpha_{xx}$  is Cronbach's Alpha. (2)

$N$  is the number of items.

Mean  $r_{xx'}$  is the mean interitem correlation.

Applying formula (2) to the ten-item score designed to measure self-esteem where the mean interitem correlation is .3, we get:

$$\alpha_{xx} = N(\text{Mean } r_{xx'}) / [1 + \text{Mean } r_{xx'}(N - 1)] = (10)(.3) / [1 + (.3)(9)] = 3 / 3.7 = .81$$

Thus, Cronbach's Alpha produces the same value that we obtained when we calculated a split-half correlation and applied formula (1), the Spearman-Brown Prophecy formula. This occurred because

all the items were, we assumed, equally correlated with each other.

Both the number of items and the average interitem correlations influence the value of Cronbach's Alpha as follows:

- As the number of equally intercorrelated items increases, Cronbach's Alpha increases.
- As the average intercorrelation among the same number of items increases, Cronbach's Alpha increases.

We now turn to the implications of these two patterns:

*Number of Items* The researcher often faces the question "How many items should I use to measure a concept?" The oversimplified answer to this question is, "More!" The more equally intercorrelated items a researcher uses to measure a concept, the higher the reliability will be.

The trick, of course, is that the items must be equally intercorrelated. In most research situations, items designed to measure a concept are not equally correlated. Some items will intercorrelate more strongly with the set of items than others. When this occurs, the researcher's judgment must be employed to decide how much of a reduction in interitem correlation offsets the increase in the number of items in the score. At a minimum, the researcher does not want to add an item which decreases the Cronbach's Alpha consistency of a scale. Standard computer software provides an option which allows the researcher to examine the Cronbach's Alpha if any item is removed from the score. When the alpha with the item removed is higher than the alpha when that item is included, there is consistency justification for the removal of that item from the scale.

This question can be posed in terms of how many items the researcher needs to meet specific alpha reliability thresholds given the mean interitem correlations. Table 1 addresses this concern. In Table 1, three alpha reliability thresholds (.7, .8, and .9) and eight mean interitem correlations (.1 through .8) are specified. We then solved formula

**Sample Size Needed for Various Alphas with Various Mean Correlations**

Mean <i>r</i>	CRONBACH'S ALPHA		
	.7	.8	.9
.1	21	36	81
.2	10	16	36
.3	6	10	21
.4	4	6	14
.5	3	4	9
.6	2	3	6
.7	1	2	4
.8	1	1	3

Table 1

(2) algebraically for the sample size needed to achieve each threshold, given each mean interitem correlation using formula (3):

$$N = \frac{[\alpha_{xx}(1 - \text{Mean } r_{xx})]}{[(\text{Mean } r_{xx})(1 - \alpha_{xx})]} \quad (3)$$

Using formula (3), the body of Table 1 presents the number of items needed for each Cronbach's Alpha threshold for each mean interitem correlation. For example, if the mean item intercorrelation is .2, sixteen items will be needed in order to achieve a Cronbach's Alpha of .8.

An examination of Table 1 reveals that when the mean interitem correlation is equal to .5, only three items are needed for an alpha of .7, four items for an Alpha of .8, and nine items for an alpha of .9. If the mean interitem correlation is .3, six, ten, and twenty-one items are needed for alphas of .7, .8, and .9, respectively. Moreover, if the mean interitem correlation is .1, twenty-one, thirty-six, and eighty-one items are needed for Alphas of .7, .8, and .9, respectively.

Thus, weak interitem correlations can be used to achieve consistency thresholds when many items are used. This is what ordinarily occurs in academic achievement tests. An exam of eighty-one items with a mean interitem correlation of .1 reaches the highly consistent .9 alpha; and an exam of only

thirty-six items with a mean interitem correlation of .1 is a reasonably consistent .8 alpha. At the same time, strong interitem correlations reach these thresholds with a small number of items. This harkens back to the observation that if two measures correlate strongly, the researcher merely picks the most convenient measure and uses it with little concern for consistency reliability.

However, the law of diminishing returns suggests that at some point, additional items with the same average intercorrelation with other items will not provide sufficient value in exchange for the effort to be expended to include those additional items. When the number of items is small, an additional equally correlated item adds substantial enough value to the reliability of the score to warrant the effort needed to include it.

Table 2 presents Cronbach's Alphas for various numbers of items with various mean interitem correlations. An examination of Table 2 illustrates the law of diminishing returns. When the mean interitem correlation is .9, the alpha is .98 with five items; adding additional items does not, indeed cannot, increase the consistency much. This is so because the maximum consistency is a perfect 1.0. When the mean interitem correlation is .3, the alpha of .68 with five items is only marginally consistent. However, the alpha increases to an acceptable .81 when ten items are used and to a highly consistent .9 when twenty items are used. Finally, the alpha for five items with a mean interitem correlation of .1 is .37. In order for a score made up of such items to be adequately consistent, the number of such items must be increased substantially.

Cronbach's Alpha can be calculated using formula (2) above. Standard statistical computer software packages can also be used for this purpose. However, care must be taken in using these packages to assure that all the items and only the items that define a specific score be included in the calculations. Indeed, the attentive researcher will want to produce the Cronbach's Alpha by hand, using formula (2), and by computer. When these two measures are identical, the researcher can take comfort that both are likely to have been done

**Cronbach's Alpha for Various Numbers of Items with Various Mean Correlations**

Mean <i>r</i>	NUMBER OF ITEMS				
	5	10	20	30	50
.1	.37	.53	.69	.77	.850
.2	.56	.71	.83	.88	.930
.3	.68	.81	.90	.93	.960
.4	.77	.87	.93	.95	.970
.5	.83	.91	.95	.97	.980
.6	.88	.94	.97	.98	.990
.7	.92	.96	.98	.99	.990
.8	.95	.98	.987	.992	.995
.9	.98	.99	.994	.996	.998

Table 2

properly. As a postscript on this discussion, we note that the Cronbach's Alpha consistency of Rosenberg's ten-item self-esteem score calculated on the data presented in Zeller and Carmines (1980, p. 92) was equal to a reasonably consistent .83. More advanced procedures which take into account which items are more highly correlated with the total score, such as theta and omega, have been omitted from this discussion. For a discussion of theta and omega, see Carmines and Zeller (1979, pp. 60-62) or Zeller and Carmines (1980, pp. 60-63). Rosenberg's self-esteem scale continues to attract academic attention (e.g., GrayLittle et al. 1997).

## STABILITY OF ALIENATION

**The Measurement of Alienation.** The concept of alienation is one of the major "unit ideas" of sociology (Nisbet 1966). But the concept is so imbued with different meanings that some have come to question its usefulness as a sociological concept (Lee 1972). Seeman (1959) believed that the conceptual confusion surrounding the study of alienation can be addressed by construing it as multidimensional. Neil and Rettig (1967) have operationalized Seeman's original conceptualizations. Following the protocols of good measurement



described above, Neal and Groat (1974) theoretically defined and empirically confirmed powerlessness, normlessness, meaninglessness, and social isolation as the four dimensions of alienation. Specifically, they constructed items designed to measure each of the four dimensions of alienation, gathered data, conducted factor analyses, noted that the observed factor structure coincided with the conceptual dimensions, created factor-based scores, and conducted substantive analyses.

R. A. Zeller, A. G. Neal, and H. T. Groat (1980) conducted a consistency and stability analysis. Data on the same sample in 1963 and 1971 revealed that reliabilities ranged from .64 to .83 in 1963 and from .65 to .88 in 1971. The authors needed accurate consistency estimates because they wished to minimize the correction for attenuation. Correction for attenuation will be discussed shortly. Zeller and colleagues wished to describe the amount of stability in the dimensions of alienation over the turbulent years between 1963 and 1971. Specifically, they wished to assess the degree to which those who had high levels of alienation in 1963 would also have high levels of alienation in 1971. In order to do so, they created scores for each of the four dimensions of alienation in both 1963 and 1971. For each score, the correlation between the 1963 and the 1971 value represented the “stability” of that dimension over that time period. High correlations would suggest substantial stability in which respondents were alienated over that time period; low correlations would suggest substantial change.

**Correction for Attenuation Due to Measurement Inconsistency.** In order to assess the stability of the dimensions of alienation over time, Zeller et al. (1980) calculated correlation coefficients between the scale scores for each dimension of alienation. They found stability coefficients ranging from .40 for normlessness to .53 for social isolation. It would appear that there was substantial stability over the eight years under investigation. Before we jump to any conclusions, however, we must consider that measurement inconsistency attenuates (i.e., reduces) the observed correlation from what it would have been if each concept had been perfectly measured at each time point. That

is, they needed to correct their stability correlations for measurement inconsistency. Formula (4) presents the correction for attenuation:

$$r_{xtyt} = r_{xy} / \text{SQRT}(r_{xx}r_{yy})$$

where

$r_{xtyt}$  is the correlation over time corrected for attenuation.

$r_{xy}$  is the observed correlation between X and Y.

$r_{xx}$  is the Cronbach's alpha for X.

$r_{yy}$  is the Cronbach's alpha for Y.

$\text{SQRT}(r_{xx}r_{yy})$  is the square root of the product of the alphas at the two points in time.

Let us apply the correction for attenuation to Zeller and his colleagues' meaninglessness score. The meaninglessness stability correlation = .52; meaninglessness had an omega consistency of .64 in 1963 and of .65 in 1971. Substituting these estimates into formula (4), we get:

$$\begin{aligned} r_{xtyt} &= r_{xy} / \text{SQRT}(r_{xx}r_{yy}) = .52 / \text{SQRT}(.64 * .65) \\ &= .52 / .64 = .81 \end{aligned}$$

Similar analyses were conducted on the other dimensions of alienation.

This analysis led Zeller and colleagues (1980, pp. 1202–1203) to conclude that their data “indicate substantial stability in the dimensions of alienation over an eight-year period.” They believe that their longitudinal data “have provided evidence to suggest that operationalizing dimensions of alienation is not only feasible, but may be accomplished with a high degree of confidence in the (consistency) reliability of the measuring instruments. The obtained stability of alienation scores over a long period of time lends credence to the search for the causal, antecedent conditions.”

**Method Artifacts in Longitudinal Research.** There are several method artifacts that can artificially attenuate or inflate the estimation of stability. As noted above, score inconsistency attenuates the stability estimate. Memory tends to inflate the stability estimate. That is, if, at time 2, respondents remember what they answered at time 1 and wish



to present themselves as being stable in their answers, they will make the same response to the item at time 2 that they made at time 1. We do not believe that this “memory effect” operated to any great degree in the analysis by Zeller and colleagues, because we doubt that respondents would remember their specific response to a specific questionnaire item for eight years. However, when the interval between time 1 and time 2 is relatively short, memory becomes a problem.

A conventional wisdom in stability research is that the interval of time that elapses between time 1 and time 2 should be long enough that respondents will not remember their specific answers to specific items, but short enough that very little change (i.e., instability) takes place in the interim. We believe, on the contrary, that part of what we wish to estimate in stability research is how much change actually takes place. Given our perspective, it does not matter how much time elapses between time 1 and time 2.

Still, the threat of artifactual deflations and inflations to the stability estimate is real. Consider the effect of item-specific variance. The respondent may answer an item in a “stable” fashion over time not because of the stability of the concept it measures, but because of some idiosyncratic nuance of the item. Idiosyncratic nuances of items unrelated to the concept the item is designed to measure are systematic, not random, error. As such, idiosyncratic nuances of items threaten to inflate the stability estimate. We now turn to a statistically advanced discussion of the identification and removal of item specific variance from stability estimation. This section requires a working knowledge of path analysis as described in Asher [(1976), 1983].

**COMBINING CONSISTENCY AND STABILITY INTO A MEASUREMENT MODEL**

The path model presented in Figure 1 combines consistency and stability into a measurement path model. In this measurement model,  $X_1$  and  $X_2$  represent the value of the concept at time 1 and time 2;  $P_{21}$  is the theoretical causal path from  $X_1$  on  $X_2$ , it represents stability, the theoretical effect of  $X_1$  on  $X_2$ . This and the other effects in this model

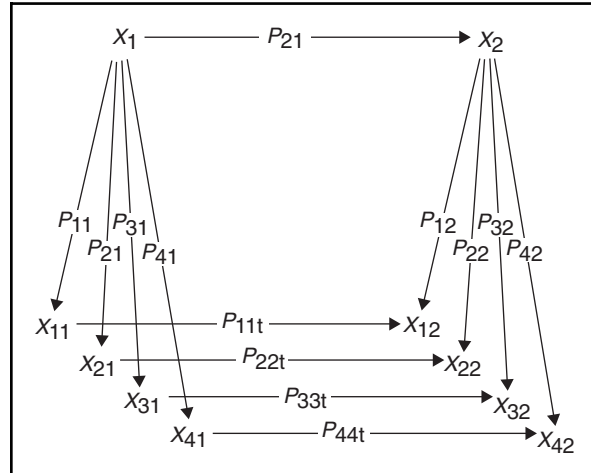


Figure 1. Consistency and Stability in a Measurement Model

can be thought of as path coefficients. The  $x_{ij}$  represent the observations; specifically,  $x_{21}$  is item 2 at time 1;  $x_{32}$  is item 3 at time 2. The  $p_{ij}$  are the epistemic correlations, the effect of the concept on each respective measure; specifically,  $p_{21}$  is effect of  $X_1$  on item 2 at time 1;  $p_{32}$  is the effect of  $X_2$  on item 3 at time 2. The  $p_{ijt}$  are the item specific effects over time, the effects of an item at time 1 on that same item at time 2; specifically,  $p_{11}$  is effect of  $x_{11}$  on  $x_{12}$  over and above the effect mediated through the concept. For a more complete discussion of epistemic correlations, see Blalock (1969).

Figure 2 presents this consistency and stability measurement model where the stability effect is  $P_{21} = .8$ , the epistemic correlation are  $p_{ij} = .7$ , and the item-specific effects are  $p_{ijt} = .3$ . These are approximately the effect parameters for the meaningfulness measurement over time in Zeller and colleagues (1980).

Table 3 presents the correlation matrix that results from applying the rules of path analysis [Asher (1976) 1983] to Figure 2. Specifically, within the measurement model, the correlation between  $x_{11}$  is  $x_{21}$  is equal to the product of the path from  $X_1$  to  $x_{11}$  times the path from  $X_1$  to  $x_{21}$ . That is,  $r = (p_{11})(p_{21}) = (.7)(.7) = .49$ . In the same way, all the time 1 measures intercorrelate .49; all the time 2 measures correlate .49.

The correlation between the time 1 and time 2 items must differentiate between the correlations

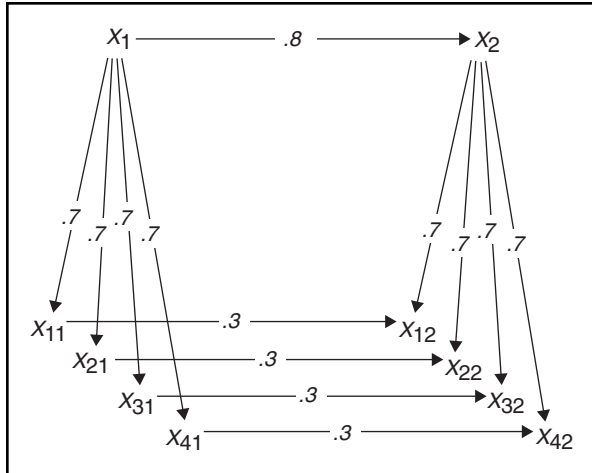


Figure 2. Hypothetical Consistency and Stability Data in a Measurement Model

between different items over time and the correlations between the same item over time. Let us first address correlations between different items over time.

The correlation between  $x_{11}$  (item 1 at time 1) and  $x_{22}$  (item 2 at time 2) is equal to the product of the path from  $X_1$  to  $x_{11}$  times the path from  $X_1$  to  $X_2$  times the path from  $X_1$  to  $x_{22}$ . That is,  $r = (p_{11})(P_{21})(p_{22}) = (.7)(.8)(.7) = .39$ . In the same way, all the time 1-time 2 correlations among different items are .39.

The correlation between  $x_{11}$  (item 1 at time 1) and  $x_{12}$  (item 1 at time 2) is equal to the product of the path from  $X_1$  to  $x_{11}$  times the path from  $X_1$  to  $X_2$  times the path from  $X_1$  to  $x_{12}$  plus  $p_{11t}$ . That is,  $r = (p_{11})(P_{21})(p_{12}) + p_{11t} = (.7)(.8)(.7) + .3 = .69$ . In the same way, all the time 1-time 2 correlations for the same item are .69.

Using formula 2, we can solve for the Cronbach's alpha, at both time 1 and time 2, as follows:

$$\begin{aligned} \alpha_{xx} &= N(\text{Mean } r_{xx})/[1 + \text{Mean } r_{xx}(N - 1)] \\ &= (4)(.49)/[1 + (.49)(3)] = 1.96 / 2.47 \\ &= .79 \end{aligned}$$

Using the criteria described above, the score is modestly, and almost highly, consistent.

The correlation between two scores can be calculated using the correlation matrix with the following formula:

$$r = e / \text{SQRT}(a + 2b) \text{SQRT}(c + 2d)$$

where

$r$  is the stability correlation between a score at time 1 and time 2.

$a$  is the number of items in the time 1 score.

$b$  is the sum of the correlations at time 1. (5)

$c$  is the number of items in the time 2 score.

$d$  is the sum of the correlations at time 2.

$e$  is the sum of the intercorrelations between the measures at times 1 and 2.

Applying formula 5 to the data in Table 3, we get:

$$a = 4$$

$$b = (6)(.49) = 2.94$$

$$c = 4$$

$$d = (6)(.49) = 2.94$$

$$e = (12)(.392) + (4)(.692) = 7.472$$

and

$$\begin{aligned} r &= e / \text{SQRT}(a + 2b) \text{SQRT}(c + 2d) = 7.472 / \\ &\text{SQRT}[4 + (2)(2.94)]^2 \\ &= 7.472 / 9.98 = .756 \end{aligned}$$

Correcting this correlation for attenuation using formula 4, we get:

$$\begin{aligned} r_{xyt} &= r_{xy} / \text{SQRT}(r_{xx} r_{yy}) = .76 / \\ &\text{SQRT}(.79)^2 = .96 \end{aligned}$$

Thus, the stability coefficient is .96. But we specified this stability coefficient to be .8 in Figure 2! What is wrong with our procedures? Why did we overstate the stability of the model? We overstated the model's stability because we included the item specific effects as stability effects. That is, the observed correlations between the same item at time 1 and time 2 represented the effects of both stability and item-specific variance. We need to remove the item specific variance from our estimation of the stability coefficient.

We can estimate the item-specific effects by subtracting the mean of the correlations of different items at time 1 compared to time 2 (mean  $r = .39$ ) from the mean of the correlations of the same item at time 1 compared to time 2 (mean  $r = .69$ ). Then we use only the variance that is not item-specific in the same item correlations across time

**Correlation Matrix among Four Measures at Two Points in Time**

Item	$x_{11}$	$x_{21}$	$x_{31}$	$x_{41}$	$x_{12}$	$x_{22}$	$x_{32}$	$x_{42}$
$x_{11}$	—	.49	.49	.49	<u>.692</u>	.392	.392	.392
$x_{21}$		—	.49	.49	.392	<u>.692</u>	.392	.392
$x_{31}$			—	.49	.392	.392	<u>.692</u>	.392
$x_{41}$				—	.392	.392	.392	<u>.692</u>
$x_{12}$					—	.490	.490	.490
$x_{22}$						—	.490	.490
$x_{32}$							—	.490
$x_{42}$								—

Table 3

( $r = .69 - .30 = .39$ ) as our estimate of what these correlations would have been if there had been no item-specific variance.

We now reapply formula 5 to the adjusted data in Table 3 to get:

$$a = 4$$

$$b = (6)(.49) = 2.94$$

$$c = 4$$

$$d = (6)(.49) = 2.94$$

$$e = (12)(.39) + (4)(.39) = 6.24$$

and

$$r = e / \text{SQRT}(a + 2b) \text{SQRT}(c + 2d) = 6.24 /$$

$$\text{SQRT}[4 + (2)(2.94)]^2$$

$$= 6.24 / 9.98 = .63$$

Correcting this correlation for attenuation using formula 4, we get:

$$r_{x1yt} = r_{xy} / \text{SQRT}(r_{xx} r_{yy}) = .63 / \text{SQRT}(.79)^2 = .80$$

Thus, the stability coefficient corrected for the removal of item specific variance is .80; despite rounding, this is equal to the .8 which was the stability coefficient specified in Figure 2.

Estimating the stability of concepts measured by scores across time is complex. A simple correlation between a measure at time 1 and the same

measure at time 2 is subject to a variety of influences. First, this stability coefficient is attenuated due to inconsistency of the item. We can address this by using multiple measures. The multiple measures allows us to estimate the consistency of the score and to correct the stability coefficient for the attenuation that occurs because the score is not perfectly consistent.

Second, the stability coefficient is artificially inflated because of item-specific variance. We can address this by estimating the size of the item-specific variance and removing it from the correlation matrix. Then we can correlate the score at time 1 and time 2 on the correlation matrix (with the item specific variance having been removed). This correlation, corrected for attenuation, is an unbiased estimate of the actual stability. For the past twenty years, the author has been searching in vain for someone who will solve for the model presented in Figure 2 from the correlation matrix presented in Table 3 using SEM techniques. Many have claimed to be able to do so, but so far, to my knowledge, no one has succeeded in doing so.

### CONCLUSION

Thirty years ago, Hauser (1969, pp. 127-128) noted that “it is inadequate measurement, more than inadequate concept or hypothesis, that has plagued social researchers and prevented fuller explanations of the variances with which they are confounded.” We have come a long way since then.

The scientific community has given greater attention to the measurement properties of the variables we use. Our capacity to conduct numerous alternative data analyses on large and well-documented data sets has been substantially enhanced. At the same time, nature is jealous of her secrets and there are many land mines buried along the paths we need to follow (or blaze) in order to make sense of our social scene. Moreover, there are many who seek shortcuts to sociological knowledge. Hence, we continue to address the challenges of establishing the consistency and stability of our measures.

## REFERENCES

- Aguirre, A., R. Martinez, and A. Hernandez 1993 "Majority and Minority Faculty Perceptions in Academe." *Research in Higher Education* 34(3):371-385.
- Asher, H. B. (1976) 1983 *Causal Modeling*. Beverly Hills, Calif.: Sage.
- Blalock, H. M. 1969 "Multiple Indicators and the Causal Approach to Measurement Error." *American Journal of Sociology* 75:264-272.
- Bohrnstedt, G. W. 1970 "Reliability and Validity Assessment in Attitude Measurement." In G. F. Summers, ed., *Attitude Measurement*. Chicago: Rand McNally.
- Bollen, K. 1989 *Structural Equations with Latent Variables*. New York: John Wiley.
- , and S. Long 1993 *Testing Structural Models*. Newbury Park, Calif.: Sage.
- Brown, W. 1910 "Some Experimental Results in the Correlation of Mental Abilities." *British Journal of Psychology* 3:269-322.
- Budd, K. W., D. RossAlaolmolki, and R. A. Zeller 1997 "Psychometric Analysis of the Self-Coherence Survey." *Archives Psychiatric Nursing* 11(5):276-281.
- Carmines, E. G., and R. A. Zeller 1979 *Reliability and Validity Assessment*. Beverly Hills, Calif.: Sage.
- Cooley, C. H. 1902 *Human Nature and the Social Order*. New York: Scribner's.
- Cronbach, L. J. 1951 "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16:297-334.
- Curtis, R. F., and E. F. Jackson 1962 "Multiple Indicators in Survey Research." *American Journal of Sociology* 68:195-204.
- Fredman, L., M. P. Daly, and A. M. Lazur 1995 "Burden among White and Black Caregivers to Elderly Adults." *Journals of Gerontology Series-B-Psychological Sciences and Social Sciences* 50(2):S110-S118.
- Grandin, E., and E. Lupri 1997 "Intimate Violence in Canada and the United States." *Journal of Family Violence* 12(4):417-443.
- GrayLittle, B. V. S. Williams, and T. D. Hancock 1997 "An Item Response Theory Analysis of the Rosenberg Self-Esteem Scale." *Personality and Social Psychology Bulletin* May 23(5):443-451.
- Hauser, P. 1969 "Comments on Coleman's Paper." Pp. 122-128 in R. Bierstedt, ed., *A Design for Sociology: Scope, Objectives, and Methods*. Philadelphia: American Academy of Political and Social Science.
- Hayduk, L. 1987 *Structural Equation Modeling with LISREL*. Baltimore, Md.: Johns Hopkins University Press.
- Hendrix, C., and W. Schumm 1990 "Reliability and Validity of Abusive Violence Scale." *Psychological Reports* 66(3):1251-1258.
- Hoyle, R. 1995 *Structural Equation Modeling: Concepts, Issues, and Applications*. Newbury Park, Calif.: Sage.
- Jackman, R. W., and R. A. Miller 1996 "A Renaissance of Political Culture?" *American Journal of Political Science* 40(3):632-659.
- Lam, S. S., and K. S. Woo 1997 "Measuring Service Quality." *Journal of Marketing Research* 39(2):381-396.
- Lee, A. M. 1972 "An Obituary for Alienation." *Social Problems* 20(1): 121-127.
- McTavish, D. G. 1997 "Scale Validity—A Computer Content Analysis Approach." *Social Science Computer Review* 15(4):379-393.
- Mead, G. H. 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Moran, J. R., D. Frans, and P. A. Gibson 1995 "A Comparison of Beginning MSW and MBA Students on Their Aptitudes for Human-Service Management." *Journal of Social Work Education* 31(1):95-105.
- Mueller, C. 1997 "International Press Coverage of East German Protest Events, 1989." *American Sociological Review* 62(5):820-832.
- Neal, A. G., and H. T. Groat 1974 "Social Class Correlates of Stability and Change in Levels of Alienation: A Longitudinal Study." *Sociological Quarterly* 15(4):548-558.
- , and S. Rettig 1967 "On the Multidimensionality of Alienation." *American Sociological Review* 32(1):54-64.
- Nisbet, R. A. 1966 *The Sociological Tradition*. New York: Basic Books.
- Riemer, H. A., and P. Chelladurai 1998 "Development of the Athlete Satisfaction Questionnaire." *Journal of Sport and Exercise Psychology* 20(2):127-156.

- Rosenberg, M. 1965 *Society and the Adolescent Self Image*. Princeton, N.J.: Princeton University Press.
- Seeman, M. 1959 "On the Meaning of Alienation." *American Sociological Review* 24:772-782.
- Spearman, C. 1910 "Correlation Calculated from Faulty Data." *British Journal of Psychology* 3:271-295.
- Szulanski, G. 1996 "Exploring Internal Stickiness." *Strategic Management Journal* 17:27-43.
- Tambs, D., J. R. Harris, and P. Magnus 1995 "Sex-Specific Causal Factors and Effects of Common Environment for Symptoms of Anxiety and Depression in Twins." *Behavioral Genetics* 25(1):33-44.
- Wright, L. K. 1991 "The Impact of Alzheimers Disease on the Marital Relationship." *Gerontologist* 31(2):224-237.
- Zeller, R. A., and E. G. Carmines 1980 *Measurement in the Social Sciences: The Link between Theory and Data*. New York: Cambridge University Press.
- 1976 "Factor Scaling, External Consistency, and the Measurement of Theoretical Constructs." *Political Methodology* 3:215-252.
- Zeller, R. A., A. G. Neal, and H. T. Groat 1980 "On the Reliability and Stability of Alienation Measures: A Longitudinal Analysis." *Social Forces* 58:1195-1204.

RICHARD A. ZELLER

## RELIGION, POLITICS, AND WAR

The Peace of Westphalia (1648) marked the end of the Thirty Years' War and the beginning of the modern European state system. The development and evolution of the principles laid out in the Westphalian treaties made the territorially defined independent sovereign state the dominant political unit for managing and governing populations. Those principles also led to recognition of the state as the primary unit for interaction (including the "interaction" of war) between territorially bounded populations. By the end of the twentieth century, the state system had encompassed the entire globe.

An understanding of the relationship between religion, politics, and war begins with an analysis of regimes (Swanson 1967). A state governs a population by means of a regime. A regime is responsible for maintaining peace and securing justice within the territorial bounds of a state. A regime acts by exercising its own autonomous powers and by implementing and enforcing laws.

Regimes take many forms, including (but not limited to) absolute monarchy (France under Louis XVI), personality-centered dictatorship (Hitler's Germany), government by cabinet embedded in a constitutional monarchy (the United Kingdom), an executive presidential system linked through a division of powers to legislative bodies (United States of America), and rule by a single party (People's Republic of China).

The autonomous powers of a regime are, in principle, unlimited, and they include the legitimate use of force. The powers of a regime and a regime's performance can be (and usually are) constrained and directed by laws and limitations such as traditional rights, even where the regime is a monarchy or dictatorship. In that regard, a major question in the analysis of the state pertains to the relationship between a regime and a society's political system or body politic. Can interests (including religious interests) be legitimately organized and expressed within a political system in ways that effectively influence a regime's actions? Answers to this question can be framed within answers to another fundamental question. What is the direct relation between a regime and the religious institutions of a society, the relation that is not mediated through an aggregation of interests in civil society and their expression in a political system? This question defines the church-state problematic.

### CHURCH AND STATE

Although the categories "church" and "state" are products of the history and experience of the West, they can be applied to the analysis of religion, regimes, and politics across the contemporary world (see Martin 1978). Two ideal-type models and their variants specify the range of relations between religion and the state. The *interpenetration model* assumes a high degree of church-state (or religion-regime) congruence and the effective unity of religious and political action. Important variants of this model include theocracy and caesaropapism. Where theocracy exists, the authority of religion is decisive for the state. In the case of caesaropapism the tables are turned and the state dominates religion.

The *separation model* posits institutional autonomy for both religion and the state. Variants include the two-powers and the strict separation

subtypes. The two-powers subtype recognizes the distinct jurisdictional domains of the state and religion within a framework of church-religion-state-regime accommodation that may include church establishment (the official recognition and support of a church by a state) but also may encompass mutual suspicion, struggle, and cross-institutional interference. Strict separation eliminates all traditional, legal, and organizational bonds and most forms of accommodation between the state and religion. Each institution is not only autonomous but also “out of bounds” with respect to the other.

Contemporary examples of the interpenetration and separation models include the Islamic Republic of Iran (theocratic), Russia (caesaro-papist), Poland (two powers), and the United States of America (strict separation). The 1979 revolution in Iran led by Ayatollah Ruhollah Khomeini imposed Twelver Shi'a Islamic law and tradition on the Iranian state. This move invigorated the Iranian masses and was hailed as progressive in many parts of the Muslim world. It also led to tension between the development of Iran as a modern state and the application of Islamic rules, principles, and laws as the authentic seal of the revolution (Arjomand 1993). The lawmaking power of the state dominates the law finding power of the traditional Shi'a jurists in Iran. Nevertheless the country can be classified as a theocracy because the resolution of disputes is constitutionally in the hands of those who have been trained in Islamic religious jurisprudence. As well, lay citizens may serve in lower offices including the Majlis (parliament), but only clerics (religious jurists) can occupy high political, administrative, or judiciary positions in Iran.

The reforms of Peter the Great, czar of Russia from 1682 to 1725, transformed the administration of the Russian Orthodox Church into an office of the state and, thereby, instituted a caesaro-papist regime. All religious activities were governed by regulations promulgated by the czar. The church was the state's instrument for controlling and educating the population. It remained so until the Bolshevik Revolution in 1917 and the foundation of the Soviet state. Under Communist Party rule, with Marxism as the official ideology of the regime, Russian caesaro-papism became a negative, eradicating, secularizing force. Church property was confiscated. The church was sheared of all

nonreligious functions. Religious practice was strictly controlled.

After the collapse of the Soviet state (1989–1991), a liberal constitution guaranteed freedom of expression and practice for all religions. No provision was made for an established church. Subsequently, however, the Russian Orthodox Church was recognized by the state as the preeminent church in Russia, thus formally linking the Russian state to the Russian church once again. In the former Soviet Union, the autocracy of the Communist Party replaced the autocracy of the Russian czar. In both cases there was a caesaro-papist relation between a regime and religion, a relation that persists in the recognition of the preeminence of the Russian Orthodox Church by the liberalized, post-Soviet Russian state.

Unlike the Russian Orthodox Church, the Roman Catholic Church in Poland has never acceded to nor successfully been forced into a caesaro-papist relation with the state. The two-powers variant of the separation model has always prevailed, even during the period of Communist rule (1948–1989).

In the early modern period, Poland did not develop a state controlled by an absolute prince. A protodemocratic republic of nobles resisted state autocracy and, thereby, prevented control of the church by a powerful, centralized regime. In the nineteenth century, Polish national identity was fused with Polish Roman Catholicism. Thus, the stage was set for the church to serve as the defender of the interests of the nation when those interests were at risk during the period of Communist rule.

Identified with the Polish nation, the Polish church stood against the Polish state in a two-powers relationship during most of the Cold War. The opposition of the church in Poland to the Communist regime made an important contribution to the downfall of communism in the Eastern bloc of European states. As one of the most powerful actors in the civil society of post-Communist Poland (having long organized itself as a hierarchic weapon in defense of its own interests), the Polish Roman Catholic Church now faces the challenge of dealing with the tendency “in victory” to become a quasitheocratic actor, a role that is inconsistent with the Vatican's position that national churches should be institutions with interests and

not would-be hegemony in contemporary liberal democratic societies (Casanova 1994).

Where Poland exemplifies the two-powers variant of the separation model, the United States is the primary example of strict separation. Constitutionally guaranteed free religious practice and a constitutional provision forbidding the establishment or state sponsorship of any religion underwrite the autonomy and differentiation of religion and the state in a historically exceptional pattern. Unlike the case of Russia, for example, where the post-Communist state recognized the Russian Orthodox Church as the preeminent national religious body, no part (legislative, executive, or judicial) of any American government (local, state, or federal) can grant legal privileges to a religious organization. By the same token, American governments can neither limit nor curtail religious practice, as happened in the former Soviet Union, nor can they pass and enforce laws that impose religious norms and rules on a population, as is the case in Iran.

Unlike Poland, where the post-Communist liberal constitution has not prevented the government from granting funds to the Roman Catholic Church in support of church-sponsored education, no church or religious body could successfully petition any government in America for material support. On the other hand, American churches and religious organizations are not taxed by governments unless they are, in effect, business organizations within the meaning and application of the law. Thus, although the state neither interferes with religious institutions nor provides recognition and material support for them, there is a pattern of passive accommodation in America whereby the state allows religious organizations to operate as untaxed nonprofit corporate bodies. Where strict separation holds, the state is not anticlerical, nor is it a direct, overt benefactor of religion.

The examples adduced above do not exhaust the range of historical detail and local variance that will be discovered where the fit of the interpenetration or separation models to a particular state is assessed. The models, themselves, however, do exhaust the range of possibilities for the regime-religion relation and, thus, can be used to classify any state and organize the history of its intercourse with religion.

## RELIGION, REGIMES, AND POLITICAL SYSTEMS

The analysis of the church-state relation focuses on links between the governing center of a society (a regime) and religious groups and institutions in a society. The analysis of religion and political systems, on the other hand, addresses questions pertaining to the organization and expression of interests in a society. Variation in the form of political systems and the relations between political systems and regimes affects the organization of interests in a society and the impact that interests including religious interests can have on regimes.

Do most of the members of a society have a role in determining who controls the society's regime? Do most of the members of a society have a role in selecting those who make the society's laws? For most of human history, the answer to both questions has been no. Kings, queens, emperors, military despots, dictators, patriarchs, oligarchs, leaders of single-party states, religious autocrats, petty tyrants, and the like—those who attain office without the broad-based consent of the people whom they dominate—were the rule, not the exception, until well into the twentieth century. Today, the democratic election of state leaders and lawmakers is a global norm, although there are many departures from it in practice (see Meyer et al. 1997).

The extent to which religious interests in a democratic society are organized and expressed in the society's political system is conditioned on the regime-religion (church-state) relation. As well, the organization of democratic political systems varies in ways that are independent of the regime-religion relation and that can affect the political expression of religious interests. Also, the properties of a society's population can have an impact on the political expression of religious interests. Finally, a state and its society may face external conditions that influence the internal relation between religion and politics.

Canada, the United States, and Israel, for example, are all modern democratic societies, but there are significant differences between them regarding religion and politics—differences that reflect variations in the regime-religion (church-state) relation, the organization of political systems, the populations, and external conditions. An examination of these variations illustrates the

range of linkages between religion and politics in contemporary democracies.

A substantial majority of the population of Israel is Jewish. Where religious matters are politically salient in Israel, controversies and conflicts often reflect sharp differences between secular and observant as well as Zionist and non-Zionist Israeli Jews. Most Israelis do not identify themselves as religiously observant in a strict sense, although only about 20 percent of the population eschew all religious observance. The state was founded (in 1948) and continuously governed for many years by a social-democratic party (Labour) whose members were for the most part secularized Jews of eastern European origin. Immigration added a significant component of observant oriental Jews to the Israeli population. Notwithstanding the secular origins of the state and the persistent secular orientation of many Israelis, religion is highly salient in Israeli politics today (Liebman 1993).

The State of Israel does not have a written constitution. Thus, the relation between religion and the state is not formally defined in a constitutional sense. No legislative or administrative acts either separate religion and the state or establish a particular variant of Judaism as the religion of the state. This means, among other things, that the resources of the state can be allocated on an ad hoc basis to religious organizations, including schools and welfare organizations that are controlled by religious groups. In this circumstance, religious groups have formed political parties to obtain state support for their organizations and to secure political goals linked to religious ideologies.

Representation in the Israeli parliament (Knesset) is based on proportional voting for parties. Proportional voting enables small parties to win seats. This feature of the Israeli electoral system plus the absence of legal or constitutional provisions either specifying or restricting the use of state resources in support of religion have encouraged the development of religious political parties in Israel. In recent years both large secular parties (socialist Labour and Likud) and smaller religious parties have received enough votes to claim seats in the parliament, but no single party has secured enough seats (a majority) to form a government. In that circumstance a government is formed by bargaining between parties. The leader of the party with the most seats negotiates with other parties to

form a majority coalition. Cabinet positions are the important bargaining chips in the process.

By entering into agreements with large secular parties when they need votes to form a government, some religious parties in Israel have secured cabinet portfolios. The electoral successes of some religious parties have enabled them to pursue interests and goals that in some cases are at considerable variance with the sensibilities and inclinations of secular Israelis, especially with regard to territorial settlements and peace, but also in matters pertaining to religious observance and the public presence of religion in Israeli life.

Israel, then, is a qualified variant of the theocratic model of religion-regime relations in the sense that there are no constitutional restrictions on the implementation of politically salient derivations that could flow from the ideology of a governing party. No religious party in Israel has succeeded, so far, in becoming a governing party, but were that to happen, there is no constitutional limitation that would, in principle, prevent a religious governing party from imposing its prescriptions and practices on the population of Israel.

Both Israel and Canada are parliamentary democracies, but the presence of religion in the contemporary political life of Canada is subdued in comparison with its visibility in the Israeli public arena. Reasons for this difference include the form of the constitutionalization of the church-state relationship in Canada, the organization of the Canadian political system, the secularization of cleavages or political fault lines in Canada, and the religious composition of the Canadian population.

The constitution of Canada underwrites a "soft" version of the caesaro-papist subtype of the interpenetration model of church-state relations. There is both a guarantee of free religious practice and the recognition of a form of religious establishment whereby public funds flow in support of primary and secondary schools affiliated with some churches. As well, hospitals and other health care organizations owned by religious groups are publicly funded.

Where education and health care are constitutionalized provincial responsibilities, there are variations between provinces regarding patterns and levels of funding for church-related schools. Also,



provinces are free to expand support beyond constitutionally required levels or to withdraw support in some circumstances. Thus, until 1999 Roman Catholic secondary schools in Ontario were not funded at the same level as public high schools. Following a provincial referendum in Newfoundland, church-related school boards were abolished in 1999, although public funds continue to support religious instruction in all schools in the province.

Constitutional and legal provisions pertaining to the church-state relation in Canada circumscribe the possibilities for political action by religious actors. The state will not provide material support to any religion for the asking, although it does provide funds in support of the educational and health care activities of designated churches, of which the Roman Catholic Church is the most notable example. This pattern has roots in the accommodation of Roman Catholic francophone Canada and Protestant anglophone Canada and the political balancing act that has kept Canada together since the founding of the Canadian state in 1867. Essentially, federal governments act in ways that are designed to ensure continuing accommodation between divisive political actors. The form that the constitutionalization of the church-state relation has taken in Canada enables governments to control the fissiparous tendencies of religiously oriented political actors through quasi-establishment (the funding of religiously affiliated schools and health care facilities) and caesaro-papist policies that, in effect, limit the public, political expression of divisive religious views (Simpson 1985).

The organization of the Canadian political system also encourages the containment of religious expression as a form of political action. The electoral rule of “first past the post” (a plurality of votes) makes it unlikely that the candidates of small parties will win seats in the federal House of Commons. This is unlike, say, the case in Israel, where the proportion of votes obtained by a party provides small parties with an opportunity for representation in the Knesset. Views on issues including religious views that lack a broad base in the Canadian population are not likely to be represented in government.

A Canadian government is formed by the appointment of a cabinet of ministers (who are the

government) by the leader of the party with a majority of seats in the House of Commons. (Not every member of parliament is a member of a government.) Once formed, a government (as is the case in any parliamentary system) can work its will without bargaining with other parliamentary parties—assuming that it is a majority government, meaning that it has a majority of seats in the House of Commons. Typically, matters related to the country’s economy and the unity of Canadian confederation dominate federal elections in Canada and, thus, rise to the level of government policy making, where they obscure other issues, including those that have a base in sectarian religious sensibilities.

The pattern of church-state articulation and the mechanics of the electoral system work against a simple, direct relation between religion and politics in contemporary Canada. As well, there has been shift in the last forty years in the nature of the cleavage that divides Canada into the nation of Quebec and the rest of Canada, a shift that has had implications for the tie between religion and politics. The rapid institutional secularization of Quebec in the 1960s and parallel movements in Anglophone Canada transformed the religion-language duality of Canada into a culture-language duality. Culture burst the bounds of religion and church control (see Borduas 1948). By the 1970s, Canada was officially defined as a bilingual and bicultural society, an understanding that was constitutionalized in 1982 along with provisions that, in effect, recognized rights pertaining to cultural maintenance and expression for all groups in Canada.

By moving the “game” of Canada to the politics of language, culture, and national expression and away from the politics of the religion-language link that was associated with Anglo dominance and the power of the Roman Catholic clergy in Quebec (a pattern that had been in place since British ascendance in the 1760s), the secularization of Canada in the post-World War II period gradually pushed religion out of the central place it had once occupied in the Canadian political system (Simpson 1988). While religion is no longer in a commanding position, its voice is not dead in the contemporary Canadian public arena. The Roman Catholic Church, mainline Protestant denominations, and Jewish organizations provide a public conscience in matters related to economic and social justice, and they engage the political system

with pursuant representations that are underwritten by the pattern of elite accommodation that typifies the Canadian political system (Simpson and Macleod 1985).

In recent years, a number of highly charged issues pertaining to abortion, homosexuality, pornography, prayer in the schools, and extramarital sex have been thrust into the American political arena. While some of these matters have public visibility in Canada, they have not generated the same level of public attention and scrutiny that they have in the United States, nor have they been a source of sustained conflict and serious divisiveness within the political arenas of Canada. Why?

The sociomoral issues that have been politicized in America have a base in the moral practices of conservative, sectarian Protestants as well as those of Roman Catholics to some extent. The reincorporation of sectarian Protestants into the American presidential electoral party system in the late 1970s brought sociomoral issues into the American public domain with force (Simpson 1983). A comparison of the religious demography of the Canadian and American populations suggests why these issues, once they were defined as public concerns, have had more prominence in America than in Canada.

About 45 percent of Americans are conservative Protestant Christians, whereas less than 10 percent of the Canadian population falls into the same category. Mainline Christian churches (including Roman Catholic) account for about 65 percent of the Canadian population and 40 percent of the U.S. population (Simpson 1988; Kosmin and Lachman 1993). There are far fewer sectarian Protestants in Canada than in the United States. Canada, then, lacks the population base for generating, supporting, and sustaining a politics of sociomoral issues. In the United States, on the other hand, a substantial portion of the population resonates with sociomoral issues that provide a public presence for everyday moral sensibilities anchored in conservative Protestant beliefs and practices.

More than religious demography underwrites the presence and tenacity of sociomoral issues in the American political system. Once thematized as politically relevant, issues have staying power that can be traced to the organization of the American system of government and, for religiously anchored

sociomoral issues, to the constitutionalization of free religious practice and no establishment of religion. Regarding the latter, practices in civil society that are approved by many but have been legally undone because they are not constitutionally valid (for example, prayer in public schools) may be a source of political agitation. A sense of unresolvable grievance arises where unfettered religious practice in the circumstance of no establishment leads to voluntaristic integrity, strength, and, even, militance in pursuit of outcomes that can only be achieved by violating the rule (no establishment) that underwrites and sustains the integrity and strength of religious practice in the first place. Unresolvable grievances are the stuff of primordial political mobilization. A politician signals that she or he is "onside" and wishes that something could be done but that it cannot because it is unconstitutional. Where the principle of no establishment is constitutionally moot, as in Israel or Canada, for example, a sense of unresolved grievance pertaining to issues linked to religion is less likely to persist, since a problem can be solved, at least in principle, by political action that entangles religion and the state.

Issues including politically relevant sociomoral issues also persist because there are three centers of political authority in the American form of government: the president, the Senate, and the House of Representatives. Each center is elected separately, and law can be made only where there is agreement among them. In this circumstance, politicized issues have lives in multiple electoral jurisdictions that encompass local, state, and national publics.

Although the majority party organizes the business of the American House of Representatives and the Senate, it does not form a government as a majority party does in a parliamentary system (for example, Canada or Israel). A government in a parliamentary system does not need the agreement of other centers of political authority to make law since a parliamentary majority guarantees that the political will of the majority party leader and her or his cabinet can become law without external advice or consent. In parliamentary systems a politicized issue tends to have a "half-life" that is as long as the attention span of the leader of the government and her or his cabinet of ministers. Where the electorate, itself, is not attuned to issues by virtue of its demographic

characteristics—as is the case in Canada regarding religiously linked sociomoral issues—it is unlikely that issues will rise to the level of serious consideration by a government, and even if they do, a government may not pay much attention to them. In the U.S. system, on the other hand, there is a tighter fit between what is on the public mind and the exercise of political authority. As long as an issue persists in the media, it has a guaranteed hearing in multiple centers of political power at the federal level.

Variations in the links between religion and politics in modern democratic systems have been illustrated above with an analysis of Israel, Canada, and the United States of America. The analysis could be extended to other democratic states and, as well, to political systems that are not democratic—for example, to the People's Republic of China where single party, top-down government has imposed a strict caesaro-papist pattern of control on religion. In any case, an analysis will examine the social forms and properties that inscribe the details of the relationship between religion and politics: the religion-regime linkage, the organization of a state's political system, and the religious demography of a state's population.

### RELIGION AND WAR

The state system that emerged from the treaties ending the Thirty Years' War in 1648 led to the contemporary global international system. In this system, a state and its regime are the elementary units of interaction. Diplomacy, trade, and war are the fundamental forms of international behavior.

By the time of World War I (1914–1918), the international system was a mix of nation-states and imperial powers with nation-state cores. Western nation-states and imperial powers controlled nearly all the inhabited world. World War I dissolved the Ottoman Empire and the Austro-Hungarian Empire. World War II (1939–1945) led to the rapid decline and breakup of the British Empire. It also led to the bipolar Cold War that pitted the West against the Soviet Union and its client states. The bipolar system came to an end in 1989.

The outcome of each of the major wars in the twentieth century, including the Cold War, led to the decline or collapse of imperial and international structures. Each such decline or collapse, in

turn, led to the formation of nation-states. Typically (but not always), states were constructed from the pieces of an imperial or a neo-imperial structure (the bipolar system of the Cold War). Thus, Turkey emerged from the Ottoman Empire and Hungary from the Austro-Hungarian Empire in the wake of World War I. India, Pakistan, Burma, and Indonesia were devolutions from the British Empire and the Dutch Empire, which were broken up following World War II. The end of the Cold War led to the formation or revival of nation-states that had been parts of the Russian empire constituted as the Soviet Union and the client states of the Warsaw Pact. They include (but are not limited to) the Baltic states (Estonia, Latvia, Lithuania), Ukraine, Georgia, Kazakhstan, and Azerbaijan.

The history of state formation in the twentieth century demonstrates a clear tendency for a nation—that is, a people who share territory, language, and religion—to form or seek to form a state in which the leash of imperial control or regional control with imperial-like features slackens or disappears. When the Cold War global structure collapsed, for example, so did the forces that were the basis for the integration of the former Yugoslavia, a regional, quasi-imperial power from the immediate post-World War II period to the end of the 1980s.

War and state formation in the former Yugoslavia followed the path of ethnoreligious differences. Slovenia (Roman Catholic) and Macedonia (Orthodox) were more or less peaceful devolutions. Serbs (Orthodox) and Croats (Roman Catholic) fought before the state of Croatia was secured. Serbs, Croats, and Bosnians (Muslim) fought in Bosnia-Herzegovina prior to a Western-imposed settlement. Having receded to the Orthodox provinces of Serbia and Montenegro, Serb-dominated Yugoslavia attempted to drive out the ethnic Albanian (Muslim) population from the region of Kosovo (a part of Serbia) before the North Atlantic Treaty Organization (NATO) intervened in 1999.

Beyond the former Yugoslavia there are other contemporary armed conflicts that are also rooted in ethnoreligious differences. These include sporadic border fighting between India and Pakistan, and armed civil conflicts of one variety or another—guerrilla war, terrorism, state military suppression—in the Philippines, Northern Ireland, Indonesia, and Sri Lanka. In the Middle East,

religiously oriented militant factions of Muslims and Jews have complicated the pursuit of peace. The successful testing of nuclear weapons by both India and Pakistan has fortified the division of the world along ethnoreligious lines.

Huntington (1996) argues that contemporary patterns of conflict in the world provide evidence that the ideologically based political divisions of the Cold War have been replaced by a set of differences that are grounded in the congruence of religion, language, and territory. These differences support, encourage, and sustain action that is rooted in collective and individual identities. At the global level, identities are embedded in civilizations, and civilizations, according to Huntington, are the key to understanding international order and disorder in the twenty-first century.

Analysis of the conflicts in the former Yugoslavia provides a sense of the developing pattern of global tension as envisioned by Huntington. Ethnoreligious divisions within collapsing Yugoslavia were replicated in the religious identities of the sources of support for the divisions in the outside world. Orthodox Serbia received economic, military, political, and diplomatic support from Orthodox Russia and Germany and, in particular, the heavily Roman Catholic state of Bavaria supported Roman Catholic Croatia. Bosnians and ethnic Albanians received sympathy and aid from sources in the Muslim world.

Jumping to the global level, Huntington argues that divisions in the world will increasingly congeal along civilizational lines. The central civilizational players (and their core states) in the intercivilizational global system of the twenty-first century, according to Huntington, include Orthodox civilization (Russia), Confucian civilization (China), Hindu civilization (India), Japanese civilization (Japan), Islamic civilization (no core state), and Western civilization (a consortium of core states consisting of the United Kingdom, France, Germany, and the United States). International conflicts and wars will tend to occur along civilizational fault lines with core states leading and/or abating intercivilizational hostilities.

If Huntington is right, conflict and war at the intercivilizational level in the twenty-first century will resemble the European wars of religion (Roman Catholics versus Protestants)—the so-called Thirty Years War—that ended with the Peace of

Westphalia in 1648. In each case, ethnoreligious or cultural identities have been the basis for the cross-state attribution of lethal patterns of “us versus them.” Whereas the modern state emerged from the seventeenth-century European wars of religion as the dominant political unit for governing and managing populations, it remains to be seen whether new forms of political organization will arise from the intercivilizational tumults of the twenty-first century.

## REFERENCES

- Arjomand, Said Amir 1993 “Shi’ite Jurisprudence and Constitution Making in the Islamic Republic of Iran.” In M. E. Marty and R. S. Appleby, eds., *Fundamentalisms and the State: Remaking Politics, Economics, and Militance*. Chicago: University of Chicago Press.
- Borduas, Paul-Emile 1948 *Refus Global*. Montreal: Mithra-Mythe Editeur.
- Casanova, José 1994 *Public Religions in the Modern World*. Chicago: University of Chicago Press.
- Huntington, Samuel P. 1996 *The Clash of Civilizations and the Remaking of World Order*. New York: Simon and Schuster.
- Kosmin, Barry A., and Seymour P. Lachman 1993 *One Nation under God: Religion in Contemporary American Society*. New York: Crown.
- Liebman, Charles S. 1993 “Jewish Fundamentalism and the Israeli Polity.” In M. E. Marty and R. S. Appleby, eds., *Fundamentalisms and the State: Remaking Politics, Economics, and Militance*. Chicago: University of Chicago Press.
- Martin, David 1978 *A General Theory of Secularization*. New York: Harper and Row.
- Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez 1997 “World Society and the Nation-State.” *American Journal of Sociology* 103:144–181.
- Simpson, John H. 1983 “Moral Issues and Status Politics.” In R. C. Liebman and R. Wuthnow, eds., *The New Christian Right*. New York: Aldine.
- 1985 “Federal Regulation and Religious Broadcasting in Canada and the United States: A Comparative Sociological Analysis.” In William Westfall, Louis Rousseau, Fernand Harvey, and John Simpson, eds., *Religion/Culture: Comparative Canadian Studies*. Ottawa: Association for Canadian Studies.
- 1988 “Religion and the Churches.” In J. Curtis and L. Tepperman, eds., *Understanding Canadian Society*. Toronto: McGraw-Hill Ryerson.

—and Henry G. MacLeod 1985 “The Politics of Morality in Canada.” In R. Stark, ed., *Religious Movements: Genesis, Exodus, and Numbers*. New York: Paragon House.

Swanson, Guy E. 1967 *Religion and Regime*. Ann Arbor: University of Michigan Press.

JOHN H. SIMPSON

## RELIGIOUS FUNDAMENTALISM

See Religious Movements; Sociology of Religion.

## RELIGIOUS MOVEMENTS

Most people do not perceive religious beliefs as changing very much over time. Religions, after all, are engaged in the propagation of eternal truths. While it is understandable that religious organizations may change, that process is perceived to occur only very slowly. Indeed, changes in beliefs are perceived to occur so slowly that adherents and leaders alike hardly notice.

Contrary to popular perceptions, both religious beliefs and religious organizations are dynamic, ever-changing phenomena. Changes in the beliefs and organizational structure of religions reflect adaptations, accommodations, and innovations to continuously changing cultural, political, and economic environments. Without more or less constant change, religions would become irrelevant to their environments, and would simply become defunct. Indeed, this has been the fate of many religions in the history of humankind.

Changes in both beliefs and organizational structure may occur as the result of actions taken by leaders empowered to make such changes, but even greater change occurs as the result of *religious movements*. How religious movements effect change, both within and outside religious organizations, is the subject of this essay. The essay unfolds in three parts. First, the classic scholarly literature about religious movements is examined within the framework of a simple typology that pivots on the origins and target of change sought by movements. Second, the phenomenon of religious fundamentalism is examined as an exemplar of religious

movements. Selection of fundamentalism is appropriate because it is one of the most important religious movements of the twentieth century. In the final part of the essay, the emerging theoretical work on religious movements is linked to some practical implications of religious movements and the future of religion in human cultures.

## UNDERSTANDING RELIGIOUS MOVEMENTS

Religious movements may be understood as a subcategory of *social movements*—that is, *organized efforts to cause or prevent change*. There are three discrete types or categories of religious movements. First, *endogenous religious movements* constitute efforts to change the internal character of the religion. Second, *exogenous religious movements* attempt to alter the environment in which the religion resides. Third, *generative religious movements* seek to introduce new religions into the culture or environment.

Religions consist of beliefs, symbols, practices, and organizations. *Endogenous movements* seek to change one or more of these aspects of a religion. Some endogenous movements have had monumental impact on both history and culture—for example, the great schism that split Christianity into Western Catholicism and Eastern Orthodoxy in the eleventh century; and the Reformation, which developed in the sixteenth century and split Protestantism from Roman Catholicism. Other movements, while important to the participants, have been of little cultural significance.

Endogenous movements frequently result in a schism—the division of the religious organization into two or more independent parts. Protestantism has been particularly schism-prone. Melton (1996) has identified more than 750 Protestant groups in North America alone. New religious groups formed through the process of schism are known as *sects*. *Sectarian movements* tend to be led by laity, or lower-echelon clergy. Many religious movements result in reform rather than schism. *Reform* is a more likely outcome when movements are initiated by the religious leaders, or when the religious hierarchy responds to and coopts grass roots demands for change. Through the centuries, the Roman Catholic Church has been particularly

effective in coopting and reincorporating movements into the church. We see them today as religious orders.

Pope John XXIII called the Second Vatican Council (1962–1965) in response to strong internal pressures to modernize the Roman Catholic Church and improve relations with other faiths. The Council produced many wide-sweeping changes in the Catholic Church. In addition, it spawned many other religious movements within the Catholic Church (e.g., liberation theology, the movement for women’s ordination, and movements for greater lay participation).

A second important hierarchically initiated movement of the twentieth century was the Protestant *ecumenical movement*. After several centuries of denominational proliferation, mostly occurring as the result of schism, the second half of the twentieth century has witnessed a powerful ecumenical movement that has resulted in the union of diverse Protestant traditions.

*Exogenous movements* constitute a second general type of religious movement. They are concerned with changing some aspect of the environment in which a religious organization exists. All religious organizations bring four imperatives to the environments in which they exist: (1) survival, (2) economic viability, (3) status, and (4) ideology (Hadden 1980). As long as these interests are secure, the religious organization exists in equilibrium or harmony with its environment. This is the normal relationship between religions and the culture. In sociological literature, these groups are identified as *churches*, or *denominations*.

When a religious group’s interests are threatened, or the leadership seeks to enhance or expand interests, religious movements may ensue. Often, exogenous religious movements are indistinguishable from social movements. Indeed, they are frequently pursued in coalition with secular social movement organizations.

In addition to legitimating religious movements with transcendental principles, religious leaders are often enlisted by secular social movement leaders to legitimate their movements. As a general proposition, religious leaders are specialists in the legitimization of social movement causes.

In the second half of the twentieth century, liberal Protestantism has forged coalitions with

virtually every liberal cause on the scene. Evangelical (conservative) Protestantism, on the other hand, has coalesced with economically and socially conservative causes. In both instances, religious organizations engage in movement activity to promote some element of their ideology. At the same time, in doing so, they hope to enhance their status. Much of the exogenous social activism of evangelical Christian groups during the past quarter-century has been grounded in the presupposition that a morally corrupt society threatens the survival of culture itself.

The very essence of religious organizations is that they carry cultural values, ideals, customs, and laws that claim transcendental character. When religious leaders engage in exogenous religious movements, they almost always draw on these transcendental principles to legitimate their cause. The claim that movement objectives are part of a divine plan, or that God is on the side of the movement, may serve as a powerful motivation for adherents of the faith to participate. Witness, for example, the many occasions during the 1980s when Islamic leaders exhorted their followers to engage in *jihad* (holy war).

The civil rights movement in the United States was substantially a religious movement. It was led by black ministers, activities were organized in black churches, funds were raised in liberal white churches, white clergy bolstered the troops who marched and picketed, and idealistic black and white youth—motivated by their religious traditions—participated in civil rights campaigns and projects.

The strength of the movement came not only from the broad base of religious participation but also from the ability of the leaders to legitimate the movement constantly in terms of sacred religious principles. For example, civil rights leaders repeatedly quoted the passage in the Declaration of Independence that acknowledges the role of the Creator in human rights: “We hold these Truths to be self-evident, that all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the Pursuit of Happiness. . . .”

The Solidarity labor movement in Poland, which was the first major social movement that led to the collapse of communism in eastern Europe, sought and received legitimacy from the Catholic

Church. Not only in Poland but also throughout eastern Europe, religious traditions were deeply involved in the movement for liberation from communism.

Not all exogenous religious movements are movements of liberation. Around the globe religious groups call on divine providence to help them in struggles against other religions, ethnic rivals, and unsympathetic governments.

There are literally hundreds of these movements around the world in various stages of ascendancy or abatement. In predominantly Hindu India, Muslims in the northern province of Kashmir seek independence or union with Pakistan, while Sikhs in the nearby province of Punjab have for many years waged a bloody confrontation with the government for independence. In Sri Lanka, just off India's southern shores, Tamils, members of a Hindu sect, seek an independent state in a nation that is predominantly Buddhist. In Northern Ireland, Protestants and Catholics have experienced periodic conflict since Protestant settlers arrived in the middle of the seventeenth century, but since 1968 the two rivals have been locked in a high level of tension punctuated with intermittent outbursts of violence.

The third type of religious movement is *generative*—a deliberate effort to produce a new religious movement. Either new religions are introduced to a culture externally by missionaries, or they are products of innovation (invention) by members of the culture. Whereas schismatic movements produce variations on an existing religion within a culture, new religions are novel to the host culture. Sociological scholars refer to these new religions as *cults*.

New religions are not necessarily newly created. Hare Krishnas, adorned in saffron robes and chanting on street corners, first appeared in the United States during the mid-1960s. The Krishnas brought with them Hindu beliefs and practices that were clearly novel to North America, but that had been first practiced in India in the sixteenth century. In contrast, the Reverend Sun Myung Moon, a Korean and founder of the Unification Church, created a religion that involved a blending of significantly reconstructed Christian beliefs with important elements of Eastern religions. In still another example, L. Ron Hubbard, a science fiction writer, published a book in 1950 titled

*Dianetics*, which outlined psychotherapeutic or mental health techniques. The book became a best-seller, and in 1954 Hubbard founded the Church of Scientology.

In these three groups we have examples of, first, the importation of an old religion based on sacred texts of Hinduism (Hare Krishna); second, a newly created religion based on reported revelation from the God of the monotheistic traditions of Judaism and Christianity (Unificationism); and third, an indigenous religion based on techniques of modern psychotherapy (Scientology). All are new and novel to North American culture.

The late 1960s and early 1970s produced a flurry of new religious movements in the United States. The youth counterculture of the 1960s provided a receptive environment for new religions. Equally important, the repeal of the Oriental Exclusion Acts in 1965 paved the way for many Eastern gurus to come to the United States as missionaries of their faiths. While not nearly as extensive, this activity can be compared to that of the Christian missionaries, who flocked to Africa and Asia, during the late nineteenth and early twentieth centuries, to seek converts to their faith.

This period of rapid cult formation was not particularly unique. The nineteenth century, for example, produced a large number of cult and sectarian movements in the United States. Christian Science, Mormonism, Seventh-Day Adventism, Jehovah's Witnesses, and Theosophy are but a few examples of groups that emerged in that time frame and that remain viable in the late twentieth century.

Significant social science literature exists on all three types of religious movements: endogenous, exogenous, and generative. The focus of inquiry has shifted significantly over time, and the discipline of the investigators has influenced the selection of questions addressed.

During the formative years of sociology much attention was devoted to discerning how new religions arise and evolve. This interest was motivated by the legacy of early sociological writing on the *church-sect* typology by Max Weber and Ernst Troeltsch. Sects develop as a result of dissent within churches, when the dissenters break away and form sects. Over time, they institutionalize and gradually become more like the churches they

earlier broke from, even as new sects are being formed. Sects tend to recruit disproportionately from the “disinherited,” or the economically deprived classes. Further, sectarian groups tend to socialize their members to the dominant middle-class values of society.

Until the mid-1960s, much of the sociological literature focused on movements that have here been identified as endogenous. Much of this literature concerned questions relating to the formation of religious movements and could be classified as consisting of theories of: (1) deprivation (socioeconomic and other), (2) social dislocation, and (3) socioeconomic change.

Historical work has focused on exogenous and generative movements. Norman Cohn’s monumental work *The Pursuit of the Millennium* concludes that revolutionary millenarian movements during the eleventh and sixteenth centuries drew their strength from groups on the margin of society. Cohn uses the term *marginal* to describe persons who are not just poor but who also have no “recognized place in society [or] regular institutionalized methods of voicing their grievances or pressing their claims” (1970, p. 282). Historical literature supports much sociological work that finds religious groups emerging on the fringe of society.

Anthropological literature tends to focus on generative movements. The question that dominates their inquiry was inherited from the evolutionary agenda of Social Darwinism in the late nineteenth century: What are the origins of religion? New religions, they conclude, emerge during periods of rapid social change, disorganization, and dislocation. In anthropological literature, this cultural strain is most often identified as the result of the invasion of an indigenous culture by a militarily advanced culture—the typical pattern of conquest and colonization by European cultures from the late fifteenth century forward.

The new religions are variously identified as “cargo cults,” “messianic movements,” “nativistic movements,” and “revitalization movements.” Anthropological literature postulates that new religions emerge as a means of dealing with cultural stress. La Barre (1972) generalizes from the scores of ethnographic studies of anthropologists to locate the origins of all religions in cultural crisis.

Lanternari, surveying anthropological and historical literature on new religions that emerged as a result of intercultural conflict, concludes that these religions “tend to seek salvation by immediate action through militant struggle or through direct and determined opposition to the foreign forces” (1965, p. 247).

Psychological literature has been much less concerned with religious movements. Following the logic of Sigmund Freud’s cultural bias against religion, many psychologists have identified the leaders of religious movements as psychopathological and their followers as psychologically defective. This literature has not been particularly productive of insights about religious movements.

The ferment of generative religious movements in the wake of the youth counterculture of the late 1960s stimulated a tremendous volume of sociological inquiry. In terms of sheer volume, research and theorizing about “new religious movements” eclipsed all other subtopics of inquiry in the social scientific study of religion during the 1970s and 1980s (Bromley and Hadden 1993). These studies examined (1) the organizational development of new religions, (2) the structural and social-psychological dynamics of affiliation and disaffiliation, and (3) the persistence of intragroup conflict between new and established religious traditions.

In addition, the 1980s saw significant theoretical developments in the conceptualization of the role of religious movements in sustaining religion in human cultures. By the early 1990s, this literature was beginning to be recognized as a distinct departure from, and a challenge to, the prevailing model of religion that dominated the social sciences for most of the twentieth century. Indeed, from the inception of the social sciences, scholars worked within an intellectual framework that viewed religion as incompatible with the modern world that is dominated by science and reason.

In 1993 Stephen Warner published an article that proclaimed a “new paradigm” in the sociology of religion. The key theoretical ideas in Warner’s argument were most closely identified with the work of Rodney Stark and his colleagues. Warner showed how support for the new paradigm has been mounting in much social research, but especially in the work of young social historians. We will return to the “new paradigm” in the



final section of this essay and examine how it has altered our understanding of the role of religious movements in human culture. Further, we will see how the new paradigm restructures the central task of the sociology of religion as the study of religious movements.

We turn next to an examination of *fundamentalism*, a religious movement that spans most of the twentieth century. Several reasons are offered for exploring fundamentalism in some depth. First, fundamentalism is one of the most important movements of the twentieth century and, thus, deserves to be examined in its own right. Second, the widespread cultural prejudice against this conservative manifestation of religion has spilled over into scholarly literature, with the result that little empirical understanding of fundamentalism developed until near the end of the century. Third, the foundation for a much better understanding of fundamentalism is now in place. We shall briefly discuss the literature that constitutes this foundational work, and point to some research tasks that are required before a mature scholarly understanding of fundamentalism can be achieved.

### RELIGIOUS FUNDAMENTALISM AS EXEMPLAR

Fundamentalism is an important religious movement that dates from the early twentieth century. Fundamentalism may be understood as protest against the quest of liberal Protestant scholars to resolve the apparent contradictions between religious knowledge and scientific discoveries. Such an effort, it was believed, could lead only to Christianity's capitulation to the ontological superiority of science as a path to truth.

From the early years of fundamentalism, most persons who have stood outside the movement have seen it through stereotypical lenses—provided substantially by the mass media—and have failed to grasp the complexity, nuances, and implications of the movement. Thus, almost from the beginning of the movement, fundamentalism has been a concept associated with religions that are perceived to be backward and potentially dangerous.

Fundamentalism has four distinct meanings. The first three meanings were in place by the end of the first quarter of the twentieth century; the

fourth does not appear until the last quarter of the century. The different meanings are often intermingled in both mass media and scholarly usage, with the result of considerable confusion and misunderstanding of the phenomenon. Significant theological and historical literatures are available on fundamentalism, but neither a theoretical nor an empirical sociological literature is well developed. This section of the essay of religious movements first identifies the four distinct meanings of fundamentalism and locates each in historical context. It then turns to a discussion of how a seriously flawed construct might be employed to better understand the phenomenon, and especially its utility for comparative sociological research.

First, fundamentalism refers to a Christian *theological movement* that experienced its greatest strength in the first quarter of the twentieth century. This movement was concerned with defending the faith against an internal movement seeking to make changes to accommodate Protestant Christianity to the modern world. As a theological movement, fundamentalism sought to purge the teachings of “modernism” from churches and theological schools. Modernist teachings had emerged during the late nineteenth century as a means of accommodating Christian doctrine to the evidence and teachings of science.

The most basic teaching of fundamentalism is that the scriptures are *inerrant*—that is, literally true. Closely related is the doctrine of *millenarianism*, which prophesies the imminent return of Christ. In the early days of the fundamentalist movement, these theological battles were waged in the leading theological seminaries of the nation (e.g., Princeton Theological Seminary).

An important development in this struggle was the publication, between 1910 and 1915, of a series of twelve books that sought to defend and reaffirm “fundamental” Christian principles in the face of the teachings of liberal scholars that did not believe the Bible should be understood as literal truth. Leading scholars from the United States and England contributed articles to the books, titled *The Fundamentals*, which were published by two wealthy Christian brothers, Lyman and Milton Steward. Copies of *The Fundamentals* were distributed gratis to over a quarter-million clergy members, seminary students, and teachers

throughout the United States. These tracts provided the inspiration for the name of the movement. The term “fundamentalism,” however, was not coined until 1920 by Curtis Lee Laws, a Baptist newspaper editor.

Defense of the faith against the encroachment of modernist theological teachings was at the core of the movement. But the *holiness movement* profoundly influenced fundamentalism, which was just as concerned with *correct behavior* as fundamentalism was with *correct belief*. The personal piety and renunciation of “worldly” vices of the holiness movement was combined with the combative spirit of theological fundamentalism to produce a *political fundamentalism*, the second distinct meaning of fundamentalism.

Fundamentalism as a political movement has had several phases. The first wave of political fundamentalism was a short-lived but vigorous conservative movement with several agendas, including temperance and anticommunism. The critical and ultimately fatal crusade of the fundamentalist movement occurred in the arena of public education policy. Charles Darwin’s theory of evolution, which had gained popularity among scientists and teachers, was clearly incompatible with a literal reading of the Bible. Among other incompatible passages in the Bible, the Genesis story of creation states that the earth and all that dwells therein were created in six days. The fundamentalists launched a campaign to prohibit the teaching of Darwinism in public schools, a campaign that initially met with considerable success in several states.

The struggle came to a climax in 1925 in one of the most celebrated trials of the twentieth century. John Scopes, a substitute biology teacher, was charged with violating a Tennessee state law that prohibited the teaching of evolution. Dubbed the “Monkey Trial,” this epochal event drew two of America’s greatest trial lawyers to the tiny town of Dayton, Tennessee. Speaking for the prosecution was William Jennings Bryan, brilliant orator and presidential nominee of the Democratic Party on three occasions. Bryan was the unchallenged leader of the fundamentalist political movement. Clarence Darrow, a bitter foe of organized religion, defended Scopes. Darrow gained prominence as a defender of labor unions and in litigation against

monopolistic corporations. He was believed by many to be the outstanding trial lawyer in the nation.

The highlight of the trial came when, in a surprise move, Darrow called Bryan as a witness. While Bryan claimed to have been a scholar of the Bible for fifty years, his ability to defend some of the finer points and implications of fundamentalist theology proved wanting. In the end, his testimony was a debacle for the prosecution. George Marsden described the scene thus:

*Bryan did not know how Eve could be created from Adam’s rib, where Cain got his wife, or where the great fish came from that swallowed Jonah. He said that he had never contemplated what would happen if the earth stopped its rotation so that the sun would “stand still.”*  
(Marsden 1980, p. 186)

The trial was quintessentially a confrontation between the emerging modern world and the forces of tradition. The drama was played out on the turf of traditionalism—a sleepy, small town in Tennessee—but journalists who were sages of the new modern order communicated the trial to the world. The fundamentalists were portrayed as fossilized relics from an era long past. Darrow himself portrayed the trial as a struggle of modern liberal culture against “bigots and ignoramuses” (Marsden 1980, p. 187).

John Scopes was convicted, but that fact seemed inconsequential. The forces of modernity and tradition had met face to face, and modernity triumphed. William Jennings Bryan collapsed and died a few days after the trial without ever leaving Dayton, Tennessee. In popular myth, Bryan died of a broken heart. Bryan had planned a national campaign to compel schools across the nation to teach evolution as a theory, not a scientific fact. There was no one else of Bryan’s stature to pick up the cause. The first wave of political fundamentalism died when Bryan was unable to defend its theological underpinnings.

The third distinct meaning of fundamentalism emerges from a melding of the theological and political movements to create a *popular caricature* of small-town Americans as culturally unenlightened religious fanatics. Journalists H. L. Mencken and Sinclair Lewis, writing in the second and third decades of this century, set the tone and style of a

genre of lambasting literature that subsequent generations of writers have admired and sought to emulate. Fundamentalists were portrayed as backwater fools preyed on by hypocritical evangelists, and as a withering species destined to disappear from the modern world. They could be meddling and annoying, but they were not viewed as politically dangerous. In time they would most certainly die off, even in the rural hinterlands of America.

The Scopes trial clearly marked the demise of the first wave of political fundamentalism. Fundamentalists in the major Protestant denominations lost ground to the modernists on the theological front. But biblical fundamentalism did not so much wane as pass from high public visibility. A number of leaders bolted from mainline Protestant churches and formed new denominations and seminaries. But out of the limelight of the press and mainstream culture, fundamentalism did not wither as had been forecast. Rather, it continued to grow, particularly in the Midwest and the South (Carpen-ter 1997).

Fundamentalists also were schism-prone, so that none of the scores of groups developed into large new denominations, such as occurred with the Baptists and Methodists in the nineteenth century (Finke and Stark 1992). This, too, served to diminish the fundamentalists' visibility. An important development occurred during the 1940s when fundamentalism effectively divided into two camps. The first, which was more insular and combative toward the larger culture, joined with Carl McIntire to create the American Council of Christian Churches (ACCC). McIntire was militantly antimodernist, and he viewed the ACCC as an instrument for doing battle with the liberal Federal Council of Churches (FCC), which in 1950 became the National Council of Churches.

A second and larger contingent of fundamentalists, who were neither with the militant McIntire contingent nor the modernist tradition aligned with the FCC, came together in 1942 to found the National Association of Evangelicals (NAE). Theologically the NAE might have considered themselves neofundamentalists, but they recognized the negative cultural stereotype associated with fundamentalism. The use of the term "evangelical" was a reappropriation of a term that most

Protestant groups used to describe themselves before the modernist-fundamentalist schism. Some of the leaders of the NAE later admitted that the name "evangelical" was a strategy to escape the negative cultural stereotypes against fundamentalism. The concept of evangelicalism was simply more respectable.

Publicly NAE leaders stressed their desire to emphasize the positive aspects of their beliefs in contrast to the highly negative and combative posture of the ACCC toward both theological modernism and political liberalism. The label "evangelical" has served well those millions of Christians whose theological beliefs are hardly discernible from those identified as fundamentalists. Billy Graham, perhaps the most respected religious leader of the second half of the twentieth century, is considered an evangelical. Theologically speaking, his basic beliefs are virtually indistinguishable from those of fundamentalist Jerry Falwell, or from those of the leadership of the Southern Baptist Convention, which staged a takeover of the Southern Baptist denomination during the 1980s.

While fundamentalism continues to be defined in terms of assent to biblical literalism, fundamentalism in the United States is highly varied in terms of the social organization, nuances of belief, and social class background of adherents. To the general public, however, fundamentalists are known in terms of the caricature that is the legacy of the Scopes trial debacle—people who are narrow-minded, bigoted toward persons different from their own kind, obscurantist, sectarian, and hostile to the modern world. The mass media dredge up enough examples of people exhibiting these traits to keep the stereotype alive.

From the 1930s forward there have been periodic flurries of right-wing political activity led by preachers and laypersons who have been labeled fundamentalists. During the Depression, William Dudley Pelley, Gerald B. Winrod, and Gerald L. K. Smith led movements that blended religion with anti-Semitism. For many decades from the 1940s forward, Carl McIntire was a strident anti-Catholic propagandist; Frederick C. Schwarz, Billy James Hargis, and Edgar Bundy were among the most visible anticommunist crusaders of the post-World War II era.

Liberal political pundits and scholars have always viewed fundamentalist groups with mixed

feelings. Some have unequivocally looked on them with great alarm, and that sense of alarm has always been greatest during periods when fundamentalist movements were highly visible. Outside periods of high visibility, the general consensus of scholars is that the fundamentalist right embodies doctrines and attracts an element that is on the fringe of the mainstream of American politics. While perhaps repugnant to the liberal ethos, they have not been widely perceived as a serious threat to democratic institutions.

This perception vacillated yet once more in 1979 when Jerry Falwell, a fundamentalist television preacher, founded a political organization named the Moral Majority. Initially the media paid little attention, but when Falwell and his fellow right-wing fundamentalists organized to help elect presidential candidate Ronald Reagan, interest picked up. When Reagan reached out and embraced the fundamentalists, attention escalated. Following the Reagan victory, along with the defeat of several ranking senators and congressmen, Falwell claimed responsibility. Pollster Lou Harris agreed that the fundamentalist vote was the margin of victory for Reagan. Postelection analysis did not support this claim (Hadden and Swann 1981), but the high media profile of Falwell and his fellow televangelists gave fundamentalism its highest public profile since the 1920s.

This wave of concern about the political power of fundamentalists might have blown over quickly were it not for the timing of the development with the rise of the Islamic imam Ayatollah Khomeini. Khomeini led a revolution that deposed the shah of Iran. Shortly thereafter, his followers held sixty-two Americans hostage for fourteen months. Political analysts concerned with the power of the fundamentalists in America were soon comparing the religious right in America with the followers of Ayatollah Khomeini and other radical Muslim factions in the Middle East. From these comparisons was born the concept of *Islamic fundamentalism*. This linkage was quickly followed by the labeling of selected politically active religious groups around the world as "fundamentalist."

Thus was born the concept of *global fundamentalism*, the fourth distinct meaning of fundamentalism. During the 1980s the idea of global fundamentalism became widely accepted by the mass media and scholars alike. But like previous uses of

the term, global fundamentalism has suffered from lack of systematic conceptualization and consistent application. The global application of the concept thus has many of the same underlying presuppositions of the popular caricature of fundamentalism in U.S. Protestantism.

"Global fundamentalism," thus, is an uncomplimentary epithet for religious groups that are viewed as out of sync with the modern world. Fundamentalism, whether the American variety or of some other culture and faith, is characterized by blind adherence to a religious dogma or leader, and by zealous rejection of the modern world. It is also widely assumed that fundamentalists are contemptuous of democratic institutions. The concept is not applied to religious movements that are perceived to be on the side of human betterment. Thus, *liberation theology*, which is global in character, is not considered to be fundamentalist in spite of the fact that it bears some considerable resemblance to movements that have been characterized as fundamentalist.

Inconsistencies in the application of the concept fundamentalism are readily apparent. This was nowhere so evident as the failure to apply the concept to Afghan Muslim guerrillas who fought the Soviet army to a standoff during the 1980s. Both theologically and politically the Afghan rebels were unmistakably Islamic fundamentalists, but they were almost never so identified in the Western press. Rather, these Afghans are almost always referred to as the *mujaheddin*, usually with positive references such as "courageous," "brave," and "freedom fighting." But seldom did anyone mention that *mujaheddin* means, literally, one who fights a *jihad* or holy war. This and other instances of inconsistent application suggest that the concept of fundamentalism is reserved for religious zealots who are disapproved.

In sum, popular use of the concept of global fundamentalism has tended to connote the same stereotypical content that the term conveys when it is applied to Protestant fundamentalists in the United States. Given this history, it might be argued that fundamentalism has not been a concept of great utility for sociological analysis and, thus, its use should be discouraged. In support of this argument, it can be said that the social scientists have done little work to define, conceptualize, or measure fundamentalism.

To propose that fundamentalism should be abandoned as a social science concept may be premature. The introduction of the idea of global fundamentalism has served to focus rather considerable attention on the phenomenon—whatever it may be. The suggestion that the phenomenon might exist across cultures and world religions invites comparative analysis that was lacking when the concept was restricted to American Protestantism.

By the late 1980s serious comparative analysis of fundamentalism had begun. One of the first things learned from comparative inquiry is that fundamentalism cannot be explained away as part of a broader conservative cultural resistance to innovation. Bruce Lawrence's research on Islam led him to the conclusion that fundamentalism is a product of modernity and, thus, is a phenomenon that did not exist prior to modernity.

To view fundamentalism as merely an unenlightened backwater resistance to innovation is to give it a misplaced emphasis. Lawrence argues that fundamentalism is a product of modernity, and "because modernity is global, so is fundamentalism" (Lawrence 1989, p. 3).

Shupe and Hadden (1989) similarly argue that "fundamentalism is a truly modern phenomenon—modern in the sense that the movement is always seeking original solutions to new, pressing problems" (p. 112). Further, the solutions they propose are new. Secularization, the cognitive counterpart to modernization, has progressively sought to compartmentalize religion from, and defined it as irrelevant to, other institutional spheres. Fundamentalism acknowledges that religion has lost authority in the secular world. Further, it perceives secular values to be seriously at variance with the sacred tradition it proclaims.

Fundamentalism may be seen as a movement that seeks to reintegrate religion into the mainstream of culture. Thus conceived, *fundamentalism* may be defined as *a proclamation of reclaimed authority of a sacred tradition that is to be reinstated as an antidote for a society that has strayed from its cultural moorings*. Sociologically speaking, fundamentalism involves (1) a refutation of the radical differentiation of the sacred and the secular that has evolved with modernization and (2) a plan to dedifferentiate this institutional bifurcation and thus bring religion back to center stage as an important factor of interest in public policy decisions.

Fundamentalism is clearly an assault on the cognitive components of modernization. Insofar as the process of modernization is not globally uniform, the development of fundamentalism may be expected to manifest a different character in different cultures. Thus conceived, the varieties of fundamentalism can be examined without the baggage of presuppositions that assume it is necessarily a regressive force in culture.

So conceived, fundamentalism is not antimodern. Fundamentalists, for example, are typically not against the use of modern technology, but rather certain applications of technology viewed to be at variance with the faith. Fundamentalists have proved themselves to be particularly adept at utilizing technology, particular communications technology, to their own ends. From the invention of radio to the development of syndicated television broadcasting, fundamentalists have dominated the use of the airwaves for religious broadcasting in the United States. They have also succeeded in developing a significant global presence. In terms of sheer volume, the four major international religious broadcasting organizations transmit more hours per week in more languages than the BBC, Radio Moscow, and the Voice of America together (Hadden 1990, p. 162).

The American Academy of Arts and Sciences launched the most ambitious comparative study of fundamentalism to date in 1987 with a substantial grant from the John D. and Catherine T. MacArthur Foundation. Over a period of five years, the Fundamentalism Project, led by historian Martin E. Marty, brought together scholars of religion from around the world to prepare studies of groups that have been identified as fundamentalist.

This project was important for several reasons. First, it both encouraged a large number of scholars to study the phenomenon seriously, and provided resources for them to do so. Second, by bringing these scholars together to critique one another's work, the project significantly leavened the individual and collective intellectual products. Third, the monumental five-volume work published by the University of Chicago Press, with over one hundred research papers, constitutes an enormous repository of information about fundamentalist and fundamentalist-like groups in every major faith tradition around the world. (See Marty and Appleby 1991, 1993a, 1993b, 1994, 1995.)

The one major fault of this otherwise marvelous inquiry into the nature and scope of fundamentalism is that the leadership early made a strategic decision not to define the subject matter. From the onset, project leaders spoke of *family resemblances*. If a group shared some family resemblances, it was an appropriate group for investigation. This strategy served to avoid long debates that would most certainly have ensued among scholars from many disciplines, cultures, and faith traditions. But in the end, the family resemblances became a proxy for a definition. The proxy, in turn, became a typology that suffers the same conceptual flaw as the *church-sect* typology—the indiscriminate mixing of correlates with attributes in a definition (Stark and Bainbridge 1979).

A serious result of this study design flaw is an inability to differentiate between fundamentalism and that which is not fundamentalism. The implications of the flaw become evident when, for example, one seeks to differentiate between fundamentalist and nationalist movements. This criticism notwithstanding, the Fundamentalism Project has given an enormous boost to an understanding of the nature, origins, and scope of a phenomenon that has been substantially shrouded in misunderstanding for almost a century. Clearly, this is a fine example of scholarly inquiry into the study of religious movements moving in the right direction. Further, the raw data in the form of five volumes of published papers is a valuable resource for others to use in advancing understanding of fundamentalism.

#### EMERGING THEORETICAL UNDERSTANDING OF RELIGIOUS MOVEMENTS

During the last two decades of the twentieth century, the study of religious movement became the cutting edge in the development of sociological theory about religion. To understand how and why this occurred, it necessary to back track just a bit to explore the development of the sociology of religion during the century.

For the better part of the twentieth century, the sociology of religion was presaged by the classic writings by the founding generation of social science: Emile Durkheim, Sigmund Freud, William James, Bronislaw Malinowski, Karl Marx,

Max Weber, and a few other notables. By midcentury, the advent of survey research added a new dimension of interest in studying religion. Much of this work could be characterized as applied, not very theoretical, and conducted by nonacademic scholars. In the academy, scholars of religions studied the classics with an eye toward fine-tuning the founding fathers' brilliance, or aligning their own work with classic writings.

In 1973 Charles Y. Glock and Phillip Hammond, working under the imprimatur of the Society for the Scientific Study of Religion, published a volume of essays entitled *Beyond the Classics?* The question mark in the title of this stock-taking volume reveals the negative conclusion—social scientists had not moved much beyond the classic statements about religion.

Insofar as their perceptions of religion were concerned, the classic scholars held highly variable views, but they tended toward a common assessment of the future of religion. The prospects of religion against the rising tide of rational thought, science, and modernity were highly precarious. *Secularization theory* was a template laid over the work of the classicists to explain the fate of religion. In a phrase, secularization is the process whereby human cultures and institutions are loosed from ecclesiastical dominance.

From the perspective of secularization theory, the concept *religious movement* almost has the quality of an oxymoron. Many scholars who studied modernization during the second third of the century viewed religious movements as a kind of “residual noise” that had no relevance to the main currents of social change. Anthropologists frequently viewed religion as a prime mover in the formation of “revitalization movements,” but the evidence of their case studies pointed to less than felicitous conclusions for aborigines; that is, in the end, modern cultures triumph over “primitive cultures.” And modern culture is secular. If religion is moving toward extinction, the study of this phenomenon is less interesting than if it were perceived to have a buoyant quality in the face of dynamic new ways of understanding the world and the human condition.

Beginning in the 1960s, several factors would serve to point intellectual thought in the sociology of religion away from the presuppositions of secularization theory. First, those who studied the

civil rights movement, and the human rights movements that quickly followed, could hardly ignore the role of religion both in leadership roles and as a force that legitimated movement activity. Second, the youth counterculture of the late 1960s began with young people's "freaking out" on drugs and then came to a conclusion with many of those same youths' "freaking out" on Jesus, or some guru from Asia. Third, the power of religion to energize large communities of people became increasingly evident in the late 1970s and early 1980s with the coincidence of a renewed fundamentalist political movement in the United States, and the politicized Muslim youth in Iran. Once scholars began to focus their attention beyond the United States, the presence of religious movements on every continent and in every faith tradition became increasingly clear.

Each of these developments drew the attention of scholars who found resources to investigate the phenomenon. In differing ways, each of these developments called attention to the incongruity of religion as a dynamic and vital force in culture and the inherited legacy of secularization theory that had long ago sentenced religion to extinction. If religion is dead, or about to slip into a comatose condition, why are religious movements thriving everywhere scholars turn?

In the discipline of sociology, more persons gravitated to the study of new religious (generative) movements than any of the other movements. Many scholars began their inquiries much as anthropologists might pursue ethnography, but soon they were engaged in comparative work, and looking at the history of new religions in the United States during the nineteenth century. Eventually the scholars of new religious movements would see the discordant implications of their work for secularization theory.

As noted in the conclusion to the first part of this essay, Warner has argued that the sociology of religion is in the midst of a process of ferment that he characterizes as a "paradigm shift." His argument is framed in Thomas Kuhn's classic statement, *The Structure of Scientific Revolutions* (1962). The accumulation of anomalous findings to the prevailing theory has mounted to the point of producing an intellectual crisis. Defenders of the old paradigm rally to shore up their position, while

new evidence accumulates that simply cannot be incorporated.

The tone of Warner's unrelenting presentation of evidence in support of a new paradigm appears early in his forty-page essay:

*The emerging paradigm begins with the theoretical reflection on a fact of U.S. history highly inconvenient to secularization theory: the proportion of the population enrolled in churches grew hugely throughout the 19th century and the first half of the 20th century, which, by any measures, were times of rapid modernization.* (1993, pp. 1048-1049)

Warner draws heavily on historians who collectively have accumulated enormous evidence that demonstrates American religious history simply does not fit the secularization model. Warner finds an explanation for this in the historical fact of *disestablishment*, that is, the separation of church and state in the American Constitution. By protecting the institutionalization of religion, but protecting no religion in particular, the Constitution creates an environment in which religions can compete and, thus, thrive and grow, or atrophy and die. Sociologically speaking, the Constitution legitimated a social structure in which *pluralism* is not only permitted but also encouraged. Whereas religious monopolies discourage the expansion of religion, pluralism promotes the growth of religion.

Warner is cautious in not making his claim for a paradigm shift beyond the U.S. boundaries, but his limiting of the argument to the United States seems clearly to be overly cautious (Hadden, 1995). The role of disestablishment in promoting the growth of religion is important in the case of the United States, and Warner's argument can be fully incorporated into a larger framework of a global paradigm.

Over the past two decades, the seminal work of Rodney Stark and his colleagues has galvanized the study of religious movements and contributed almost immeasurably to the development of the new paradigm that Warner invisions. As the senior scholar in the enterprise, Stark has surrounded himself with a group of exceptionally able collaborators. Together they have developed a theory, identified a research agenda, pursued that research agenda vigorously, and inspired others who

have joined the task of building a new model for understanding religion.

For almost a decade, Stark's principal collaborator was William Sims Bainbridge. Stark and Bainbridge's work together began in 1979 with an article that bore a modest subtitle, "Preliminary Concepts for a Theory of Religious Movements." In this paper they identified and defined key concepts for the study of religious movements. They then proceeded to fill in details with more than twenty published papers over the next five years. These papers were published together in 1985 under the title *The Future of Religion*. Two years later Stark and Bainbridge published *A Theory of Religion* (1987), a work of monumental importance. *A Theory of Religion* goes beyond religious movements to provide a comprehensive theory that accounts for the origins, dynamics, and persistence of religions in human cultures. Beginning with six axioms, these authors deductively create more than a hundred definitions and nearly three hundred and fifty propositions.

The theory has sweeping implications for an understanding of religion in the modern world. First, the theory deduces why religion must necessarily be ubiquitous in human cultures. Second, religious institutions are human constructions and, thus, subject to constant change. Third, religious movements are the principal mechanism through which religious change occurs. Hence, *if religious are constantly in flux, and religious movements constitute the principal mechanism for change, it follows that the core focus of inquiry for social science is the study of religious movements.*

The theory is clearly inspired by *rational choice theory*. Stark's collaborations with Roger Finke, a sociologist, and Lawrence Iannaccone, an economist, are especially rich with concepts and imagery from economics. They speak of a *religious economy* as embracing all the religious activity of a given culture, and they work comfortably with concepts like *markets, niches, supply-side religion, monopoly, and competition*. Still, both Stark and Bainbridge vigorously object to the criticism that the paradigm is "nothing more" than a wholesale borrowing from economics and rational choice theory. (See Bainbridge 1997, pp. 350–356; Stark and Finke 2000).

As of the writing of this essay, the new paradigm lacked a name and, hence, the reference to

the "new paradigm." Rational choice theory has emerged as a significant development within American sociology. The American Sociological Association now has a rational choice section, and there is a journal entitled *Rationality and Society*. This perspective has also been referred to as the "theory of religious economy," but this too implies a narrower perspective than either Stark or Bainbridge wishes to communicate. In 1994 Lawrence Young, of Brigham Young University, convened a conference on "Rational Choice Theory and Religion," at which the principal contributors to the "new paradigm" were participants. Stark's own reservations to the name "rational choice theory" are communicated in his opening essay to the proceedings, entitled "Bringing Theory Back In." His concluding remark captures his commitment to follow theory wherever it may lead to fruitful insights: "My goal is to bring real theories into sociology, not to found a new theoretical sect."

Rejecting the names "rational choice theory" and "theory of religious economy" would appear to be a wise strategy, at least for the present time. Many scholars who are presently studying religious movements would also reject the identification of their work with either of these names. Still, a very large proportion of the emerging scholarship about religious movements is informed by and also adds credence to the emerging new paradigm. For a more detailed examination of this literature, and an annotated bibliography of religious movements studies that are informed by the new paradigm, see Hadden's Religious Movements Homepage on the Internet, <http://www.religiousmovements.org> (1999). To locate these materials, go to the front page and search on "new paradigm bibliography."

## REFERENCES

- Bainbridge, William S. 1997 *The Sociology of Religious Movements*. New York: Routledge.
- Bromley, David G., and Jeffrey K. Hadden, eds. 1993 *Handbook of Cults and Sects in America*, 2 vols. Greenwich, Conn.: JAI.
- Carpenter, Joel A. 1997 *Revive Us Again: The Reawakening of American Fundamentalism*. New York: Oxford University Press.
- Cohn, Norman 1970 *The Pursuit of the Millennium*. New York: Oxford University Press.



- Finke, Roger, and Rodney Stark 1992 *The Churching of America: 1776–1990*. New Brunswick, N.J.: Rutgers University Press.
- Glock, Charles Y., and Phillip E. Hammond, eds. 1973 *Beyond the Classics?* New York: Harper and Row.
- Hadden, Jeffrey K. 1980 "Religion and the Construction of Social Problems." *Sociological Analysis* 41:99–108.
- 1990 "Precursors to the Globalization of American Televangelism." *Social Compass* 37:161–167.
- 1995 "Religion and the Quest for Meaning and Order: Old Paradigms, New Realities." *Sociological Focus* 28:1(February):83–100.
- 1999 "Annotated Bibliography of 'New Paradigm' Research in the Sociology of Religion." Religious Movements Homepage. University of Virginia. <http://www.religiousmovements.org>. Search on "new paradigm bibliography" to access.
- , and Charles E. Swann 1981 *Prime Time Preachers*. Reading, Mass.: Addison-Wesley.
- Kuhn, Thomas S. 1962 *The Structures of Scientific Revolutions*. Chicago: University of Chicago Press.
- La Barre, Weston 1972 *The Ghost Dance*. New York: Delta.
- Lanternari, Vittorio 1965 *The Religions of the Oppressed*. New York: Mentor Books.
- Lawrence, Bruce B. 1989 *Defenders of God*. San Francisco: Harper and Row.
- Marsden, George M. 1980 *Fundamentalism and American Culture*. New York: Oxford University Press.
- Marty, Martin E., and R. Scott Appleby, eds. 1991 *Fundamentalism Observed. The Fundamentalism Project*, vol. 1. Chicago: University of Chicago Press.
- , eds. 1993a *Fundamentalisms and Society. The Fundamentalism Project*, vol. 2. Chicago: University of Chicago Press.
- , eds. 1993b *Fundamentalisms and the State. The Fundamentalism Project*, vol. 3. Chicago: University of Chicago Press.
- , eds. 1994 *Accounting for Fundamentalisms. The Fundamentalism Project*, vol. 4. Chicago: University of Chicago Press.
- , eds. 1995 *Fundamentalism Comprehended. The Fundamentalism Project*, vol. 5. Chicago: University of Chicago Press.
- Melton, J. Gordon 1996 *The Encyclopedia of American Religions*, 5th ed. Detroit, Mich.: Gale Research.
- Robbins, Thomas 1988 *Cults, Converts and Charisma*. Newbury Park, Calif.: Sage.
- Shupe, Anson, and Jeffrey K. Hadden 1989 "Is There Such a Thing as Global Fundamentalism?" In Jeffrey K. Hadden and Anson Shupe, eds., *Secularization and Fundamentalism Reconsidered. Religion and the Political Order*, vol. 3. New York: Paragon House.
- Smith, Christian 1998 *American Evangelicalism: Embattled and Thriving*. Chicago: University of Chicago Press.
- Stark, Rodney 1987 "How New Religions Succeed: A Theoretical Model." In David G. Bromley and Phillip E. Hammond, eds., *The Future of New Religious Movements*. Macon, Ga.: Mercer University Press.
- , and William Sims Bainbridge 1979 "Of Churches, Sects, and Cults: Preliminary Concepts for a Theory of Religious Movements." *Journal for the Scientific Study of Religion* 18:2(June):117–131.
- , 1985 *The Future of Religion*. Berkeley: University of California Press.
- , (1987) 1996 *A Theory of Religion*. New Brunswick, N.J.: Rutgers University Press.
- , and Roger Finke 2000 *The Human Side of Religion*. Berkeley: University of California Press.
- Wallace, Anthony F. C. 1966 *Religion: An Anthropological View*. New York: Random House.
- Warner, Stephen 1993 "Work in Progress Toward a New Paradigm for the Sociological Study of Religion in the United States." *American Journal of Sociology* 98(March):1044–1093.
- Young, Lawrence A. (ed.) 1997 *Rational Choice Theory and Religion*. New York: Routledge.

JEFFREY K. HADDEN

## RELIGIOUS ORGANIZATIONS

The social organization of religion in the United States is diverse and complex. Most religious organizations are local churches (congregations, parishes, synagogues) tied to national religious bodies (usually referred to as denominations). The *Yearbook of American and Canadian Churches* lists 189 denominations in the United States with a total of almost 360,000 churches. The membership reported by these churches equals almost 58 percent of the U.S. population (Lindner 1999). The largest denominations are the Roman Catholic Church (61,207,914 members in 1996), the Southern Baptist Convention (15,891,514 members in 1997), and The United Methodist Church (8,496,047 members in 1996). Most denominations are quite small. In all, the twenty-one denominations with membership in excess of one million members account

for more than 140,000,000 members—about 91 percent of all church members in the United States. In contrast, the seventy-seven denominations with fewer than 25,000 members account for about 586,000 members, fewer than one-half of 1 percent of church members (figures calculated from information in Lindner 1999).

Typically, local churches hold worship services at least once a week and also have educational activities, especially for children and youths. Most churches organize various groups within the church to accomplish particular tasks (for example, missions, evangelism, or community action), or for the association of persons with common interests (such as women, youths, or senior citizens). Women's groups are especially active. Many churches serve as community centers providing space for meetings of all sorts of neighborhood and community organizations. Ammerman's recent work (1997) shows the crucial role churches play in their communities. "Not only are they linked to other parts of the community through the multiple memberships and loyalties of their members, but they are also linked *as organizations* to larger organizational networks" (p. 360) that pursue community-based goals.

Local churches usually have a pastor (or a priest or rabbi) and a lay governing board. There is great variation from denomination to denomination on the authority of lay boards, and, within denominations, there is variation from church to church in informal power. Research has shown that control by inner circles of informal leaders is likely to emerge when formal mechanisms of control and official leaders are not performing effectively (Houglund and Wood 1979).

The degree of control of the denomination over the local church depends in large part upon the polity, or political structure, of the denomination. Students of religious organizations place denominations in three categories according to polity. *Congregational* polity is the weakest. In this polity the local church owns its own property and hires and fires its own pastor. In contrast, in a *hierarchical* (often *episcopal*) polity the national (or regional) body holds title to the property and controls the placement of pastors. An in-between category is often called *presbyterial*. There are a number of correlates of polity. For example, denominations with strong polities were more active

supporters of the civil rights movement and more aggressively pressed for the integration of their churches (Wood 1981).

Denominational polities continue to evolve as they adapt to changing social environments (Dykstra and Hudnut-Beumler 1992). Recent years have seen a growth in the influence of local congregations—with a concomitant gain in local control—even within the most hierarchical polities. Given the communications revolution, which allows the almost instant dissemination of information throughout the church, this trend toward local control is congruent with Michels' argument (1962) that leaders' control of information is a principal basis of hierarchical control. A major study being conducted by researchers at Hartford Seminary, "Organizing Religious Work for the 21st Century," will provide massive data for the understanding of how denominational polities are adapting to their changing environments.

Though the organization of Jewish synagogues is similar to that of many Protestant churches in the United States, the Jewish perspective on religious organization is somewhat different. In 1987 the officials of the congregational organizations of the Orthodox, Conservative, and Reform branches of Judaism reported 3,750,000 persons associated with their synagogues and temples. However, there are approximately six million Jews, who are seen as an ethnic, social, and religious community (Jacquet 1989, pp. 243–244). Daniel Elazar stresses that Jews see no meaningful line of separation between "churchly" purposes and other communal need, and hence Jewish organizations are not neatly divided into religious and nonreligious ones. "It is *not simply* association with a *synagogue* that enables a Jew to become part of the organized Jewish community. Affiliation with *any of a whole range of organizations*, ranging from clearly philanthropic groups to 'secularist' cultural societies, offers the same option" (Elazar 1980, p. 133). Elazar argues that local Jewish federations for welfare, educational, and cultural activities should be seen as religious organizations (p. 133).

#### RELIGIOUS ORGANIZATIONS IN SOCIOLOGICAL CONTEXT

Religious organizations provide patterns for the interaction of religious individuals. Social forces

change these patterns, but in turn, the collective action of religious people influences society. Sociologists looking at religious organizations have been interested especially in their importance as plausibility structures that foster specific beliefs and values (Berger 1967) and as structures of action that mobilize people to seek social change.

Until the 1970s the sociological approach to religious organizations was guided primarily by the church-sect typology. This theoretical framework helped to explain the number and variety of religious bodies and differences in their behaviors by reference to the social class of their adherents. Max Weber distinguished between a church, a continuously operating rational, compulsory association that claims a monopolistic authority, and a sect, “a voluntary association [that] admits only persons with specific religious qualifications” (Weber 1978, p. 56). “One becomes a member of the church by birth . . . [but a] sect . . . makes membership conditional upon a contractual entry into some particular congregation” (p. 456). Weber’s student, Ernst Troeltsch (1961), developed a typology from these concepts and some variation of the church-sect typology has been used repeatedly in studying U.S. religious organizations.

In the Weberian tradition, H. Richard Niebuhr stressed the sociological sources of sect formation and the way in which social forces tended to turn sects into churches. He argued that sects originate “in the religious revolts of the poor, of those who were without effective representation in church or state” (1954, p. 19) and who employed a democratic, associational pattern in pursuing their dissent because it was the only way open to them. Niebuhr observed that the pure sectarian character of organization seldom lasts more than one generation. As children are born to the voluntary members of the first generation,

*the sect must take on the character of an educational and disciplinary institution, with the purpose of bringing the new generation into conformity with ideals and customs which have become traditional. Rarely does a second generation hold the convictions it has inherited with a fervor equal to that of its fathers, who fashioned these convictions in the heat of conflict and at the risk of martyrdom. As generation succeeds generation, the isolation of*

*the community from the world becomes more difficult. Furthermore, wealth frequently increases when the sect subjects itself to the discipline of asceticism in work and expenditure; with the increase of wealth the possibilities for culture also become more numerous and involvement in the economic life of the nation as a whole can less easily be limited. (Niebuhr 1954, pp. 19–20).*

Nancy Ammerman’s work continues the research tradition that relates the evolution of churches to social class backgrounds. Ammerman traces the rise of fundamentalism in the Southern Baptist Convention to the erosion of cultural support for traditional beliefs. She finds that fundamentalism decreases with increased levels of education and with increased levels of income. But “many at the edges of this transition are likely to respond by embracing fundamentalist beliefs more vigorously than ever (Ammerman 1986, p. 487).

According to James Beckford, “The question of the degree to which any particular organisation was church-like or sect-like was taken seriously for what it implied about that organisation’s capacity to survive in the modern world” (1984, p. 85). The church-sect theorizing was dominated by considerations of rationalization and compromise. Beckford detected a shift in the focus of sociologists studying religious organizations in the 1970s toward “the capacity of religious organisations to foster a sense of personal authenticity, conviction and self-identity” (p. 85). The 1970s saw a great many studies about recruitment and mobilization by religious organizations. Many of these studies focused on the growth and decline of traditional organizations, but many others dealt with religious movements that were new, or at least new upon the U.S. scene. Beckford refers to a number of authors who have found that cultlike formations are appropriate to an age marked by rationalization, bureaucratization, and privatization. That is, small groups of people cultivating esoteric religion in private are flexible and adaptable to the conditions of highly mobile and rapidly changing societies. Some of these scholars have linked cults’ ability to inspire and mobilize their members to their distinctive forms of organization.

In recent years more emphasis is placed on applying general organization theory to religious

organizations. Many recent studies of religious organizations are characterized by an open-systems approach, which views organizations as adaptive organisms in a threatening environment (Scherer 1980). The questions of adaptability to the modern world and of inspiration and mobilization of followers come together in studies of the Roman Catholic Church. John Seidler and Katherine Meyer (1989) examine that denomination's accommodations to the modern world, many of which involve important structural changes, such as priest's councils, and other changes that allowed both priests and lay people to have more say in the operation of the church.

A relatively new theoretical perspective within the sociology of organizations and social movements—resource mobilization—has illuminated much of the current scene of new religious movements. Bromley and Shupe did a detailed resource mobilization analysis of the Unification Church. They argue that one key element in the church's establishment in the United States was the development of mobile fund-raising teams (1979).

### CURRENT ISSUES

A more varied theoretical approach to religious organizations has allowed scholars to focus on different kinds of issues. A major concern has been the decline of the liberal mainline denominations and the significance of that decline (Roof and McKinney 1987; Hoge and Roozen 1979). The liberal mainline churches in the United States share with other churches in a vast mobilization of voluntary time and money in activities caring for individuals such as the poor, the sick, and the elderly. Churches are particularly effective at such mobilization because they instill philanthropic values and present information and opportunities for philanthropic activities in face-to-face settings such as worship services and Sunday School classes. The liberal churches have played the additional role of implementing socially liberal policies, that is, policies designed to change the structure of society so that, for example, income as well as opportunities for individual achievement are more widely distributed throughout society. The liberal social agenda also includes sharp criticism of the U.S. government's role as promoter of U.S. business interests abroad. Mobilizing individuals and

groups to press for the acceptance and implementation of a liberal social agenda may be these churches' most significant contribution to U.S. society (Wood 1990).

Social and cultural changes in the United States in the last three decades have led to important changes in most traditional denominations and to the decline of membership and resources in many of them. At the same time, a new form of religious organization—the megachurch—has spread rapidly. These large (usually having at least 2,000 members), multiservice congregations are rarely affiliated with any denomination. However, megachurches are often associated with one another in networks that provide some of the services that denominations provide their member churches. In the twenty-first century, megachurches and their networks continue to provide significant competition for denominations. Yet these churches are themselves not immune to social and cultural change. Already there are reports that some of the leading megachurches are finding it necessary to make major changes in their programs to attract post-baby boomers (Jorstad 1999).

Another issue related to denominational polity is the role of women in the ministry. Chaves (1997) argues convincingly that the great variability in denominations' approval of women's ordination can be explained as responses to the denominations' significant social environments.

In an era of rapid social and cultural change, religious organizations play a crucial role in the process of consensus formation in our society. Amitai Etzioni (1968) argues that a healthy society is one in which the relationship between citizens and national leaders is mediated by a large network of groups and organizations where multiple perspectives are reduced toward consensus. The effect of any direct appeal by national leaders or by mass media campaigns to individual citizens is determined largely by the multiple membership of the citizens in groups and organizations. This mediation protects against mass emotional manipulation. At the national level the many "legislatures" within the major religious bodies in this country are of enormous importance in shaping the working consensus that enables both the formulation and the implementation of national policies. The representative selection procedures for

national meetings and the deliberative consensus formation processes typical of the major denominations are an important contribution to informed public opinion in U.S. society.

At the local level, congregations provide forums in which differing opinions can be expressed in a context of mutual respect. David Knoke and James Wood (1981) show that a wide variety of nonreligious social influence associations did not attract people with views as diverse as those in the church. They suggest that "in most of these organizations, policy-dissatisfied members probably do not feel the social pressure to remain in the organization comparable to that felt by dissatisfied church members" (p. 103). Churches' multiple goals and their emotional and value ties provide holding power that keeps members with different views together in the same church. Voluntary associations in which individuals can debate the critical issues face to face encourage individuals to act out their selfless values rather than their selfish interests, and provide a bulwark against the manipulation of the public by computer-generated direct mailing and mass media campaigns for a particular group's vested interest in ideology, money, or power.

Wood (2000) argues that the consensus formation process described above is contributing to the resolution of the issue of homosexuality, one of the most controversial social issues in church and society today.

#### **OTHER RELIGIOUS ORGANIZATIONS**

Robert Wuthnow (1988) has described the rise of numerous special-purpose organizations that are rooted in religion, drawing legitimation and resources from the more traditional religious organizations but with the objective of achieving a quite specific purpose.

These organizations provide new options for religious people in addition to participation in local churches. A wide variety of purposes are pursued, including the advancement of nuclear disarmament and meeting the spiritual needs of senior citizens. Wuthnow suggests that "as far as the society as a whole is concerned, these organizations may be the ones that increasingly define the public role of American religion. Rather than

religion's weight being felt through the pressure of denominations, it may be exercised through the more focused efforts of the hundreds of special-purpose groups now in operation" (1988, p. 121). Though these special-purpose groups are in many ways a revitalizing influence on traditional religious organizations (denominations and local churches), they may also have important sociological implications. For example, while the traditional organizations have often held together people of diverse social backgrounds, special-purpose groups may have a tendency toward homogeneity.

There are also a number of important umbrella organizations, such as the National Council of Churches and the National Association of Evangelicals, that facilitate the cooperation of sets of denominations. The National Council of Churches was particularly important in mobilizing a segment of the church population into the civil rights movement (Wood 1972). There has also been a growth of community councils of churches.

#### **NEW RELIGIOUS ORGANIZATIONS**

In recent years a number of the religious organizations have been in the news are unrelated either to the Judaeo-Christian heritage or to immigrant groups. They draw their adherents largely from the broad center of the U.S. middle class. Robert Ellwood and Harry Partin (1988) discuss more than forty of these groups. None of them are very large and in most of them most of their members remain affiliated for less than a year. Perhaps their greatest importance from the sociological perspective is that they introduce new organizational models into the U.S. scene.

#### **THE FUTURE**

New immigrant people are bringing their religions with them to the United States. Islam in particular is growing rapidly. People in the United States may have to start thinking of themselves as a Protestant, Catholic, Jewish, *Islamic* nation. According to one source, in 1973 there were fifteen or twenty local centers of Muslim worship in the United States; by 1980 these centers were reported in all of the three hundred largest cities in the United States. Two million adherents were reported in 1980; at the end of the century most authorities

estimate that there are between three and four million Muslims in the United States (some estimates go as high as six million). Most Islamic organizations in the United States are local centers (variously called Islamic Societies, Islamic Centers, or Muslim Mosques). Each of these organizations provides a place of worship and a place for other religious, social, and educational activities. Islam does not have an organized hierarchy, but several regional and national groups help to coordinate the work of local groups and promote unity among them (Jacquet 1989). If, as Elazar contends, many Jewish organizations in addition to the synagogue play a religious role, in Islam it appears that the religious centers play many roles in addition to the religious one. Perhaps this is always the case with the churches of recent immigrants.

Stark and Bainbridge (1985) say that traditionally organized religion may decline drastically as more and more people pursue individualistic “careers” of going from one self-enhancement group to another. If they are correct, any societal influence of religious organizations would be felt more through influence on individuals than through collective action of large religious bodies. However, there is much evidence that the traditional structure of religious organization in the United States will persist in the twenty-first century.

## REFERENCES

- Ammerman, Nancy T. 1986 “The New South and the New Baptists.” *The Christian Century* (May 14):486–488.
- 1997 *Congregation and Community*. New Brunswick, N.J.: Rutgers University Press.
- Beckford, James A. 1984 “Religious Organisation: A Survey of Some Recent Publications.” *Archives de Sciences Sociales Des Religions* 57:83–102.
- Berger, Peter L. Sector 1967. *The Sacred Canopy* Garden City, N.Y.: Doubleday.
- Bromley, David G., and Anson D. Shupe, Jr. 1979 “Moonies” in America: *Cult, Church, and Crusade*. Beverly Hills, Calif.: Sage.
- Chaves, Mark 1997 *Ordaining Women: Culture and Conflict in Religious Organizations*. Cambridge, Mass.: Harvard University Press.
- Dykstra, Craig, and James Hudnut-Beumler 1992 “The National Organizational Structures of Protestant Denominations: An Invitation to a Conversation.” In Milton J. Coalter, John M. Mulder, and Louis B. Weeks, eds., *The Organizational Revolution: Presbyterians and American Denominationalism*. Louisville, Ky.: Westminster/John Knox.
- Elazar, Daniel J. 1980 “Patterns of Jewish Organization in the United States.” In Ross P. Scherer, ed., *American Denominational Organization*. Pasadena, Calif.: William Carey Library.
- Ellwood, Robert S., and Harry B. Partin 1988 *Religious and Spiritual Groups in Modern America*. Englewood Cliffs, N.J.: Prentice Hall.
- Etzioni, Amitai 1968 *The Active Society*. New York: Free Press.
- Hoge, Dean R., and David A. Roozen 1979 *Understanding Church Growth and Decline: 1950–1978*. New York: Pilgrim.
- Hougland, James G., Jr., and James R. Wood 1979 “‘Inner Circles’ in Local Churches: An Application of Thompson’s Theory.” *Sociological Analysis* 40:226–239.
- Jacquet, Constant H., Jr. (ed.) 1973, 1980, 1989 *Yearbook of American and Canadian Churches*. Nashville, Tenn.: Abingdon.
- Jorstad, Erling 1999 “Megachurch Regrouping after Growth Lull.” *Religion Watch* 14:1.
- Knoke, David, and James R. Wood 1981 *Organized for Action: Commitment in Voluntary Associations*. New Brunswick, N.J.: Rutgers University Press.
- Lindner, Eileen W., ed. 1999 *Yearbook of American and Canadian Churches*. Nashville, Tenn.: Abingdon.
- Michels, Robert 1962 *Political Parties*. New York: Free Press.
- Niebuhr, Gustav 1998 “American Religion at the Millennium’s End.” *Word and World* 18(Winter):5–13.
- Niebuhr, H. Richard 1954 *The Social Sources of Denominationalism*. Hamden, Conn.: Shoe String.
- Roof, Wade Clark, and William McKinney 1987 *American Mainline Religion: Its Changing Shape and Future*. New Brunswick, N.J.: Rutgers University Press.
- Scherer, Ross P. 1980 *American Denominational Organization*. Pasadena, Calif.: William Carey Library.
- Seidler, John, and Katherine Meyer 1989 *Conflict and Change in the Catholic Church*. New Brunswick, N.J.: Rutgers University Press.
- Stark, Rodney, and William S. Bainbridge 1985 *The Future of Religion: Secularization, Revival, and Cult Formation*. Berkeley, Calif.: University of California Press.
- Thumma, Scott 1996 “The Kingdom, the Power, and the Glory: Megachurches in Modern American Society.” Ph.D. diss., Emory University.

- Troeltsch, Ernst 1961 *The Social Teachings of the Christian Churches*. New York: Harper and Row.
- Weber, Max 1978 *Economy and Society*, vol. 1. Berkeley: University of California Press.
- Wood, James R. 1972 "Unanticipated Consequences of Organizational Coalitions: Ecumenical Cooperation and Civil Rights Policy." *Social Forces* 50:512-521.
- 1981 *Leadership in Voluntary Organizations: The Controversy over Social Action in Protestant Churches*. New Brunswick, N.J.: Rutgers University Press.
- 1990 "Liberal Protestant Social Action in a Period of Decline." In Robert Wuthnow and Virginia A. Hodgkinson, eds., *Faith and Philanthropy in America: Exploring the Role of Religion in America's Voluntary Sector*. San Francisco: Jossey-Bass.
- 2000 *Where the Spirit Leads: The Evolving Views of United Methodists on Homosexuality*. Nashville, Tenn.: Abingdon.
- Wuthnow, Robert 1988 *The Restructuring of American Religion: Society and Faith since World War II*. Princeton, N.J.: Princeton University Press.

JAMES R. WOOD

## RELIGIOUS ORIENTATIONS

Sociologists generally conceive of religion as a system of symbols that evokes a sense of holistic or transcendent meaning (Bellah 1970, p. 16; Geertz 1973, pp. 90-125). This definition reflects sociology's claim that symbols are essential to the human capacity to experience and interpret reality (Berger and Luckmann 1966). Symbols are acts, objects, utterances, or events that stand for something—that is, that give meaning to something by connecting it to something else. Symbols give order and meaning to life. Without them, life would be experienced as senseless and chaotic. Indeed, research suggests that individuals are able to experience and understand only those aspects of their worlds for which they have symbols (Farb 1973).

Sociologists' emphasis on holistic or transcendent meaning as the defining feature of religion arises from their view that meaning is always contextual (Langer 1951). The meaning of a particular word depends on the other words that form its immediate context. For example, the word "courts"

means one thing if it appears with the word "tennis," but something different when the word "justice" or, the word "dating" is present. Similarly, in their daily lives people give meaning to their activities by associating them with various frames of reference. Hitting a tennis ball has meaning, for example, because it is associated with the rules of the game of tennis. Each frame of reference, moreover, has meaning because it can be placed within a more encompassing symbolic context (tennis, say, within the context of physical exercise and health). But if each symbolic framework requires a broader framework to have meaning, then, some form of holistic or transcendent symbol system that embraces all of life must be present. These are what sociologists call religious orientations or religious systems (Berger 1967; Roberts 1984).

The questions that typically invoke religious symbols involve the quest to make life itself meaningful. Such questions arise at the extremities of human existence: Where did I come from? Why am I here? What happens when I die? These questions, framed at the individual level, may also be asked about the collectivity to which one belongs or about humanity in general: How did our tribe originate? Where is humanity headed? Other questions focus on the absolutes or landmarks that make life recognizable in its most basic sense: What is beauty? What is truth? How can we know truth? What is essential about the human condition? There are also questions that arise because the events they deal with make no sense to us on the surface: Why must I die? Why is there suffering in the world? What is the reason for evil?

Transcendent symbol systems address these questions at a variety of levels. Elaborate philosophical and theological doctrines sometimes supply rational answers that satisfy canons of logic and empirical evidence. In daily life these questions are more likely to be addressed through narratives, proverbs, maxims, and ikonic representations rich with experiential connotations. Religious orientations are likely to be structured less by abstract deductive reasoning than by parables that raise questions but leave open precise answers, by personal stories that link experience with wider realities, and by creeds and images that have acquired meaning through long histories of interpretation in human communities (Greeley 1982, pp. 53-70).

Like other symbol systems, religious orientations are thought to be the products of social interaction. Although the role of such factors as divine revelation cannot be ruled out, sociologists focus on the ways in which symbols come to have meaning through the interaction of individuals and groups in human communities. Sometimes these communities invent collective symbols to articulate powerful experiences they may have undergone. More commonly, communities borrow symbols available within their cultural traditions, but then adapt these symbols to their own use, giving them new meanings and interpretations. Communities also underwrite the plausibility of religions belief systems (Berger 1967, p. 45). They do so by providing evidence that such beliefs are not the product of individual imaginations alone, by encouraging the public expression of beliefs, and by creating occasions on which beliefs may be enacted and predictions fulfilled. Without the ongoing interaction of people in communities, it is doubtful whether belief systems could long be sustained. Research has also demonstrated that personal religious orientations are more likely to have behavioral consequences if these orientations are supported by communities of like-minded individuals (Roof 1978).

In defining religion as a symbol system that deals with ultimate questions, sociologists assume that humans have the capacity to question their experience and a desire to make sense of their worlds. Whether all people pursue this desire with equal intensity is more doubtful. It is possible, for example, to explain a plane crash by observing that a rivet came loose. It is also possible to let the incident raise questions about the meaning of pain, the frailty of human existence, or the meaning and purpose of one's own life. How much the quest for holistic meaning and transcendence enters into people's lives is, therefore, a matter of variation. Studies indicate that most people say they have thought about the meaning and purpose of life, but individuals vary in the extent to which they have been troubled by this issue. They also vary in the amount of explicit attention they have devoted to it and in their views about the possibility of arriving at definite answers (Stark and Glock 1968, p. 77). Agnosticism, for example, is a religious orientation that grants the importance of ultimate questions about meaning and purpose

but denies the possibility of finding answers to these questions.

The kinds of symbols that come into play in relation to such questions are also matters of variation. While all such symbol systems may perform functionally similar roles, it is useful to distinguish them substantively. These substantive distinctions are usually the basis on which religious orientations are delineated in popular discourse. At the broadest level, sociologists distinguish theistic meaning systems, which recognize the existence of a God or divine being, from atheistic systems, which do not acknowledge a divine being (Glock and Stark 1965, pp. 3–17). Christianity is an example of the former; Marxism, of the latter. Insofar as it addresses the same higher-order questions about the meaning of life, Marxism would be considered functionally similar to Christianity. But this does not mean that Marxism necessarily functions this way. Just as one might study Marxism to derive economic principles, so one might study Christianity simply as an example of literature. In neither case would it be appropriate to say that a religious orientation is at work. Only as they function to evoke holistic meaning and transcendence do symbol systems become religious orientations.

The distinction between theistic and atheistic meaning systems is useful when the relevant concept is the presence or absence of a divine entity. But this distinction may be less useful in other contexts. For example, contemporary discussions in theology and in science sometimes distinguish religious orientations on the basis of whether they posit a reality that is humanly knowable or ultimately mysterious, whether reality is empirical or includes a supraempirical dimension, or whether being implies something that is not being itself but the ground of being. In these debates the boundary between varieties of ultimate meaning systems is often ambiguous.

In contemporary societies religious orientations are often distinguished in popular belief according to the dominant force or power that people perceive as governing their lives (Wuthnow 1976). Some people may conceive of this force as God; others, as luck or fate. Natural or human causes may also be considered dominant; for example, the force of heredity, of scientific law, society, or individual willpower. Whether a part of



elaborate philosophical systems or simple pieces of folk wisdom, such understandings help people to make sense of their lives by identifying the causal agents that control human events.

Sociologists have insisted that religious orientations become important to the study of human behavior insofar as these orientations are internalized as part of the individual's worldview. A worldview can be defined as a person's guiding outlook on life. The essential aspects of a religious orientation are the person's beliefs and assumptions about the meaning of life and such matters as the existence and nature of God, goodness and evil, life beyond death, truth, and the human condition. These beliefs and assumptions help the individual make sense of life cognitively. They also have an emotional dimension, perhaps including a feeling of awe, reverence, fear, or peace, comfort, and security. In addition, they are regarded as behavioral predispositions that lead to various actions, each as participation in worship, prayer, or ethical decisions (Spilka et al. 1985).

The importance of religious orientations for ethical decisions has been of longstanding interest to sociologists. In the classical work of Max Weber (1963), religious orientations were conceived of as symbolic frameworks that made sense of the world, in part, by providing explanations for the existence of evil (also known as theodicies). Some religious orientations, for example, explained evil as a struggle between God and the devil, others saw evil as part of a cycle of regeneration and renewal, while still others attributed evil to the workings of an all-powerful but inscrutable deity. The implications for ethical action derived from the prescriptions for salvation implied by these different conceptions of evil. In one tradition, for example, people might be expected to pray and meditate in order to escape from the cycle of evil and regeneration; in another tradition, they might be expected to do good deeds as a way of siding with the forces of good against those of evil.

Much of the research by sociologists on religious orientations has dealt with their subjective aspects (Blasi and Cuneo 1986). Assuming that the important feature of symbolism is its meaning, researchers have tried to discover what religious symbols mean to individuals. Efforts have been made to tap the deeper predispositions presumed

to underlie such religious expressions as prayer and worship, to say how deeply implanted the religious impulse is, and to classify varieties of religious outlooks and experiences.

Recent developments in sociological theory have resulted in some rethinking of this emphasis on subjective religiosity. Current research is beginning to focus more on the observable manifestations of religious symbolism itself, rather than claiming to know what lies beneath the surface in the subjective consciousness of the individual (Wuthnow 1987). Discourse, language, gesture, and ritual have become more important in their own right (Tipton 1982). The contrast between this and the earlier approach can be illustrated by comparing two statements: "I believe God exists" and "God speaks to us through the Word." A subjective approach would treat both statements as manifestations of some inner conviction of the individual. The more recent approach would pay closer attention to the language itself, noting, for example, the more personalized style of the first statement and the collective reference contained in the second.

The value of the more recent approach is that it recognizes the public or social dimension of religious orientations. Observers may not know what goes on in the dark recesses of the believer's soul. But if that person tells a story, or participates in worship, the researcher can then study the observable manifestations of that person's faith.

To account for variations in religious orientations, sociologists usually look at the social conditions to which people are exposed. They assume that most people do not make up their own religions from scratch. Rather, they borrow from the various symbol systems that are available in their environment. The most significant borrowing occurs in early childhood. Family is thus an important factor, and it, in turn, is influenced by broader conditions such as social class, levels of education, race and ethnicity, and exposure to regional subcultures.

A generation ago, sociologists often held the view that scientific generalizations could be made about the relationships between social factors and religious orientations. For example, much work was inspired by the hypothesis that theistic religious orientations were more common among persons with lower levels of education than among

persons in better-educated social strata. Another common hypothesis suggested that religious orientations were likely to be associated with various kinds of social deprivation, since the deprived would presumably seek solace in otherworldly beliefs. Empirical studies have found some support for such hypotheses, but the ability to make generalizations has remained limited. Different relationships seem to be present in different communities and in different time periods.

More attention has turned in recent years, therefore, toward describing the rich and complex processes by which religious orientations and social environments intermingle. In one setting people without college educations may turn to religious views that shield them from the uncertainties of science and other modern ideas. In another setting people with high levels of education may also turn to religion, but do so in a way that combines ideas from science and Scripture or that focuses on the therapeutic needs of people working in the professions. In both settings, religious orientations provide answers to ultimate questions. But the composition of these orientations reflects ideas present in the different social settings.

An earlier generation of social theorists also sought to explain the variations in religious orientations in ways that often reduced them to little more than the by-products of social or psychological needs. Sociologists following in the tradition of Karl Marx, for example, regarded religion merely as a reflection of class struggles, while some following Émile Durkheim viewed it as a reflection of the corporate authority of society (Swanson 1960, 1967). The reductionism in these approaches consisted not only of regarding social structure as more basic than religion but also of implying that religion would gradually disappear as people became more aware of its origins (Fenton 1970). Recent work is decidedly less reductionistic in its assumptions about religion. It still assumes that religion fulfills human needs and that it is influenced by social conditions, but regards religion as a more active contributor to human experience and considers its future more viable.

In addition to the more general social conditions that may influence the religious orientations of individuals, sociologists have also been particularly interested in the institutions that devote specific energies to the promulgation of religious

orientations. These institutions supply the resources needed for religious orientations to be perpetuated. Leadership, producers of religious knowledge, specialists in the dissemination of such knowledge, organizational skills, physical facilities, and financial resources are all required for religious orientations to be maintained over time. Religious institutions must compete with other institutions, such as governments, businesses, and families, for these resources.

In most modern societies competition is also present among the adherents of various religious orientations (Wuthnow 1988a). When such competition has been recognized either governmentally or culturally, we say that a condition of religious pluralism exists (Silk 1988). Pluralism often becomes a kind of religious orientation itself, imposing norms of civility and tolerance on particularistic religious traditions. When multiple religious orientations are forced to compete with one another, the plausibility of any one such tradition may be diminished as a result of believers' seeing others who hold views different from their own. At the same time, pluralism appears to contribute to the overall vitality of religious orientations in a society by encouraging competition among them for adherents and by giving believers more options from which to choose (Christiano 1987).

It has been common in the past for individuals to choose one particular religious orientation with which to identify. Often these orientations have been defined by religious institutions, such as the Roman Catholic church, or by denominational organizations, such as the Presbyterian or Methodist churches (Greeley 1972). Increasingly, however, it appears that individuals in modern societies are exposed to a variety of religious institutions and orientations. As a result, they may pick and choose particular elements from several different faiths and traditions. Their religious orientation therefore takes on a more personalized character (Bellah et al. 1985, pp. 219-249; Roof and McKinney 1987, pp. 40-71).

Although some individuals work out highly coherent religious orientations that have internal consistency and integrity, it appears that the more common result of living in religiously pluralistic settings is a form of personalized eclecticism. People become heteroglossic—that is, they gain the capacity to speak with many religious voices. Their

religious orientations may not provide a guiding philosophy of life that maintains an orderly view of the world. Rather, religious orientations become tool kits (Swidler 1987) assembled from a variety of personal experiences, social contacts, books, sermons, and other cultural repertoires, and from which the individual is able to draw as he or she is confronted with the challenges of life.

At present, research studies indicate that large proportions of the population in societies like the United States hold theistic religious orientations (Wuthnow 1988a). In other societies where religious institutions have had fewer resources in the past, such orientations are less common. In all societies, though, theistic orientations are confronted by the humanistic orientations promulgated by secular institutions. The outcome appears to involve a balance between pressures to adapt, on the one hand, and tendencies by religious adherents to resist these pressures, on the other hand (Hammond 1985; Beckford 1989). Much of the struggle depends on the ability of religious leaders to articulate visions that grow out of particular confessional traditions in ways that appeal to the universalistic norms governing wider social audiences.

Although religious orientations have become more diverse and eclectic as a result of cultural contact and mass communication, evidence also suggests that in some societies a basic polarization has emerged between those whose orientation involves traditionalistic, fundamentalistic, or conservative norms, on one side, and those whose orientation involves progressive, modernistic, or liberal norms, on the other side (Wuthnow 1988a). Conservatives are characterized by adherence to the authority of traditional scriptural texts, while liberals emphasize more the relativity of these texts and the need for reason and experience in interpreting them. Liberal religious orientations have been nurtured by relativistic views in higher education, in the professions, and in the mass media in market-oriented societies, but conservative orientations have grown as well, not only in reaction to liberalism, but also as a result of conservatives gaining educational or political advantages and seizing on opportunities created by the ill effects of rapid societal change (Ammerman 1987; Hunter 1987). Whereas earlier discussions predicted the demise of fundamentalist religious orientations, current studies are more concerned

with the ongoing tensions between fundamentalist and more liberal or more humanistic religious orientations.

Research on religious orientations continues to be produced. Fundamentalism and evangelicalism have been examined comparatively and through quantitative studies (Marty and Appleby 1994; Shibley 1996; Smith 1998), and conflicts between fundamentalist or orthodox views and those of liberals or progressives have been examined (Hunter 1991; Wolfe 1997). Other important currents in recent research include studies emphasizing eclecticism, individualism, and instability in contemporary spiritual orientations (Roof 1993), the apparent shift in popular thinking from beliefs rooted in organized religion toward orientations emphasizing personal spirituality (Roof 1999), and the rising importance of spiritual practices in response to uncertainties about formal religious beliefs (Hall 1997; Wuthnow 1998).

(SEE ALSO: *Social Philosophy*; *Sociology of Religion*)

## REFERENCES

- Ammerman, Nancy Tatom 1987 *Bible Believers: Fundamentalists in the Modern World*. New Brunswick, N.J.: Rutgers University Press.
- Beckford, James A. 1989 *Religion and Advanced Industrial Society*. London: Unwin Hyman.
- Bellah, Robert N. 1970 *Beyond Belief: Essays on Religion in a Post-Traditional World*. New York: Harper and Row.
- , Richard Madsen, William M. Sullivan, Ann Swidler, and Steven M. Tipton 1985 *Habits of the Heart: Individualism and Commitment in American Life*. Berkeley: University of California Press.
- Berger, Peter L. 1967 *The Sacred Canopy: Elements of a Sociological Theory of Religion*. Garden City, N.Y.: Doubleday.
- , and Thomas Luckmann 1966 *The Social Construction of Reality; A Treatise in the Sociology of Knowledge*. Garden City, N.Y.: Doubleday.
- Blasi, Anthony J., and Michael W. Cuneo 1986 *Issues in the Sociology of Religion: A Bibliography*. New York: Garland.
- Christiano, Kevin J. 1987 *Religious Diversity and Social Change: American Cities, 1890–1906*. Cambridge, England: Cambridge University Press.
- Farb, Peter 1973 *Word Play*. New York: Bantam.

- Fenton, John Y. 1970. "Reductionism in the Study of Religion." *Soundings* 53:61-76.
- Geertz, Clifford 1973 *The Interpretation of Cultures*. New York: Harper and Row.
- Glock, Charles Y., and Rodney Stark 1965 *Religion and Society in Tension*. Chicago: Rand McNally.
- Greeley, Andrew M. 1972 *The Denominational Society: A Sociological Approach to Religion in America*. Glenview, Ill.: Scott Foresman.
- 1982 *Religion: A Secular Theory*. New York: Free Press.
- Hall, David D., ed. 1997 *Lived Religion in America: Towards a History of Practice*. Princeton, N.J.: Princeton University Press.
- Hammond, Phillip E., ed. 1985 *The Sacred in a Secular Age*. Berkeley: University of California Press.
- Hunter, James Davison 1987 *Evangelicalism: The Coming Generation*. Chicago: University of Chicago Press.
- 1991 *Culture Wars: The Struggle to Define America*. New York: Basic.
- Langer, Susanne K. 1951 *Philosophy in a New Key*. New York: Mentor.
- Marty, Martin E., and R. Scott Appleby, eds. 1994 *Accounting for Fundamentalisms: The Dynamic Character of Movements*. Chicago: University of Chicago Press.
- Roberts, Keith A. 1984 *Religion in Sociological Perspective*. Belmont, Calif.: Wadsworth.
- Roof, Wade Clark 1978 *Community and Commitment: Religious Plausibility in a Liberal Protestant Church*. New York: Elsevier.
- 1993 *A Generation of Seekers: The Spiritual Journeys of the Baby Boom Generation*. San Francisco: Harper San Francisco.
- 1999 *Spiritual Marketplace: Baby Boomers and the Remaking of American Religion*. Princeton, N.J.: Princeton University Press.
- , and William McKinney 1987 *American Mainline Religion: Its Changing Shape and Future*. New Brunswick, N.J.: Rutgers University Press.
- Sibley, Mark A. 1996 *Resurgent Evangelicalism in the United States: Mapping Cultural Change since 1970*. Columbia: University of South Carolina Press.
- Silk, Mark 1988 *Spiritual Politics: Religion and America since World War II*. New York: Simon and Schuster.
- Smith, Christian 1998 *American Evangelicalism: Embattled and Thriving*. Chicago: University of Chicago Press.
- Spilka, Bernard, Ralph W. Hood, Jr., and Richard L. Gorsuch 1985 *The Psychology of Religion: An Empirical Approach*. Englewood Cliffs, N.J.: Prentice-Hall.
- Stark, Rodney, and Charles Y. Glock 1968 *American Piety: The Nature of Religious Commitment*. Berkeley: University of California Press.
- Swanson, Guy E. 1960 *The Birth of the Gods: The Origin of Primitive Beliefs*. Ann Arbor: University of Michigan Press.
- 1967 *Religion and Regime: A Sociological Account of the Reformation*. Ann Arbor: University of Michigan Press.
- Swidler, Ann 1987 "Culture in Action: Symbols and Strategies." *American Sociological Review* 51:273-286.
- Tipton, Steven M. 1982 *Getting Saved from the Sixties: Moral Meaning in Conversion and Cultural Change*. Berkeley: University of California Press.
- Weber, Max 1963 *The Sociology of Religion*. Boston: Beacon.
- Wolfe, Alan 1997 *One Nation After All: What Americans Really Think about God, Country, Family, Racism, Welfare, Immigration, Homosexuality, Work, the Right, the Left and Each Other*. Boston: Viking.
- Wuthnow, Robert 1976 *The Consciousness Reformation*. Berkeley: University of California Press.
- 1987 *Meaning and Moral Order: Explorations in Cultural Analysis*. Berkeley: University of California Press.
- 1988b "Sociology of Religion." In Neil J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- 1988a *The Restructuring of American Religion: Society and Faith since World War II*. Princeton, N.J.: Princeton University Press.
- 1998 *After Heaven: Spirituality in America since the 1950s*. Berkeley: University of California Press

ROBERT WUTHNOW

## REMARRIAGE

Remarriages have become almost as common as first marriages in contemporary America. In fact, in recent years, almost half of all marriages involved at least one spouse who had been married previously (U.S. Bureau of the Census 1998). Remarriage rates in the United States (the number of people remarrying each year per 1,000 persons divorced or widowed) increased during the 1960s, declined precipitously across the 1970s, and have continued to decline throughout the 1980s and 1990s, although at a much slower rate (Sweet and Bumpass 1988; U.S. Bureau of the Census 1998).

Contemporary remarriages are more likely to follow divorce, in contrast to earlier centuries, in which remarriage typically followed the death of a spouse. For example, Demos (1970) reported that in the Plymouth Colony, approximately 40 percent of men and 26 percent of women over the age of 50 had been married more than once, due primarily to the death of the first spouse. In contrast, throughout the twentieth century, remarriages following the death of a spouse have been increasingly outnumbered by remarriages following divorce, owing to the combination of the dramatic decrease in the mortality rate in the first few decades of the century and the increase in the divorce rate. In fact, by 1990, remarriages in which both parties had been divorced occurred twenty times more often than remarriages in which both parties had been widowed (Clarke 1995).

While remarriage rates are related to divorce rates, they do not always follow the same trend. For example, the rise in remarriage rates in the 1960s accompanied the rise in divorce rates, but the decline in remarriage rates in the 1970s occurred at a time when the divorce rates were still growing. This apparent incongruence in the 1970s and early 1980s can be explained to a great extent by an increase in cohabitation among formerly marrieds across the same period (Bumpass et al. 1991).

### WHO REMARRIES?

Rates of remarriage vary substantially by age, gender, race and ethnicity, and marital status (i.e., divorced or widowed).

*Age* at the time of termination of the first marriage is clearly the best predictor of remarriage—particularly for women—with younger people remarrying at much higher rates. For example, almost 90 percent of women under the age of 25 when their first marriage ends remarry, while fewer than one-third of women over the age of 40 at the time of termination remarry (Bumpass et al. 1990).

Remarriage rates also differ by *race and ethnicity*. White non-Hispanic women are substantially more likely to remarry than are black women, and somewhat more likely than Hispanic women (U.S. Bureau of the Census 1998). White non-Hispanic

women also remarry much more quickly than do either black or Hispanic women. Based on 1988 data, 35 percent of white women had remarried within two years, compared to only 16 percent of black women and 17 percent of Hispanic women (U.S. Bureau of the Census 1998). Further, black and Hispanic women are less likely to cohabitate following divorce than are white women (U.S. Bureau of the Census 1998). Thus, the variations in remarriage appear to reflect a general tendency for white non-Hispanic women to be more likely than black or Hispanic women to establish new households following divorce.

*Gender* is a particularly important factor to consider when discussing remarriage, not only because remarriage rates vary by gender, but because gender interacts with several other factors that predict remarriage. Overall, three-quarters of divorced men and two-thirds of divorced women in the United States remarry (Cherlin and Furstenberg 1994). However, gender differences in likelihood of remarriage increase substantially across the life course. For formerly marrieds under 30 years of age, there is little difference in men's and women's likelihood of remarriage; however, by age 35, men are much more likely to remarry, and they are increasingly more likely in each successive age group. For example, among those aged 30–34, the rate of remarriage per 1,000 in the formerly married population is 138 for women, and 178 for men. In the age group 50–54, the gender discrepancy increases to a rate of 25 per 1,000 for women, and 75 per 1,000 for men. By age 65 and over, the rate of remarriage rate for women is only 2 per 1,000 formerly marrieds in the population, compared to 15 per 1,000 for men (Clarke 1995). Thus, women's likelihood of remarriage not only decreases considerably across the life course, but decreases particularly markedly compared to that of men.

Various explanations for the overall lower remarriage rates for women than men have been developed. The most commonly cited explanation focuses on the limited “marriage market,” or field of eligibles, for women who experience the termination of their marriage through divorce or death. First, there are fewer men than women, and this discrepancy increases with age (U.S. Bureau of the Census 1998). Also, women tend to marry men who are older than themselves, further limiting

the pool of eligibles for women who are themselves older.

The effect of the *presence of children* on likelihood of remarriage also varies considerably by gender. In particular, the presence of children has much greater effects on the likelihood of remarriage for women than for men. For example, a recent study in New York State found that only 45 percent of divorced women with children remarried, compared to 67 percent of divorced men with children. Further, the likelihood of remarriage for women, but not men, declined as the number of children from the previous marriage increased (Buckle et al. 1996). About half of the women with one child remarried, but only about a quarter of the women with four children remarried. In contrast, about two-thirds of men with either one or four children remarried. It is possible to speculate that the lower rate of remarriage for women with children, particular those with multiple children, might be a function of the fact that women with several children are likely to be older than their childless counterparts; however, even when controlling on age (i.e., comparing women within the same age groups), the rate of remarriage is about one-quarter lower among mothers than among childless women (Bumpass et al. 1990).

Buckle and colleagues (1996) found that the presence of children also affected men's, but not women's, choice of a new partner. Divorced men who were childless were four times more likely to marry women who had not been married previously; in contrast, divorced men with children were almost twice as likely to marry a woman who had been married before. While divorced women were almost twice as likely to marry men who had previously been married than those who had not, this choice was not affected by whether the women had children.

Recent work by Sweeney (1997) suggests that the *age of children* must also be considered when examining the effect of children on their mothers' likelihood of remarriage. She found that preschoolers reduced women's likelihood of remarriage, while school-age children and teenagers had no effect, and children 18 years or older were associated with an *increased* likelihood of remarriage; these findings are similar to those reported by Koo and Suchindran almost twenty years earlier (1980),

suggesting little change across time in the effects of this factor on remarriage.

#### MATE SELECTION AMONG THOSE WHO REMARRY

One of the most striking patterns in first marriages is the tendency for individuals to marry others with similar social characteristics, such as age, educational attainment, religion, and socioeconomic background—a pattern known as homogamy. Although there is also a tendency toward homogamy in remarriage, the degree of similarity has been less than in first marriages, although this trend appears to be changing.

The differences in status similarity between first marrieds and remarrieds can be seen by examining partners' age discrepancies. Individuals who remarry tend to select mates from a wider field of eligibles compared to first marrieds, resulting in greater age difference between spouses in remarriage; however, the discrepancy appears to have decreased over the past two decades. Throughout the past thirty-five years, the groom in a first marriage has been, on average, two years older than the bride. Between the early 1960s and the mid-1970s, remarried grooms were, on average, four years older than their brides; however, since 1980, the discrepancy has declined to about three years (Clarke 1995). Thus, while there continues to be less age homogamy in remarriages than first marriages, the difference has become muted.

Individuals who are formerly married also appear to be increasingly likely to select a partner who shares the same marital status. For example, by 1990, divorced women were almost twice as likely to marry men who were also divorced as to marry men who had never been married (Clarke 1995). However, as in the case of age, homogamy of marital status was greater than in earlier years; in fact, the rate of divorced women marrying divorced men almost doubled from 1970 to 1990.

Socioeconomic factors appear to have somewhat different effects on patterns of first marriages and subsequent marriages—particularly for women. The women most likely to enter first marriages are those with greater socioeconomic prospects (cf. Goldscheider and Waite 1986; Lichter et al. 1992). However, the patterns are more complicated for remarriage. Haskey (1987) reported

that women with higher occupational prestige were *less* likely to remarry; however, more recent research by Sweeney (1997) suggests that the effect of women's occupational prestige on remarriage may be more complex. She found that, among women under the age of 25 at the time of separation, those with higher occupational prestige were less likely to remarry; however, among women separating at age 45, higher occupational prestige was associated with an increased likelihood of remarriage. Thus, it appears that the middle-aged women who are the least able to support themselves are also the women least likely to be able to achieve economic stability through remarriage, further increasing the risk of poverty for women in this age group.

For men, it is less clear whether socioeconomic prospects have differential effects on entry into first marriages and remarriages. As in the case of women, greater economic prospects increase men's likelihood of entering into first marriages (cf. Goldscheider and Waite 1986; Lichter et al. 1992); however, evidence on economic factors and remarriage are less consistent. While Haskey (1987) found that men with higher occupational prestige were more likely to marry, Sweeney (1997) found no effects for any dimension of men's socioeconomic prospects, including educational attainment, job status, occupational aspirations, and work commitment.

Taken together, these findings suggest that patterns of remarriage may be becoming more similar to patterns of first marriages in terms of some demographic characteristics, such as age and previous marital status, but not necessarily in terms of socioeconomic prospects.

#### DIFFERENCES BETWEEN FIRST MARRIAGES AND REMARRIAGES

**Marital Quality.** While the popular press has debated the relative happiness of first marriages and remarriages, scholarly research on this issue has failed to find important differences. The most comprehensive review of this research to date (Vemer et al. 1989) found that first marrieds report only slightly greater marital satisfaction than do remarrieds. Elizabeth Vemer and her colleagues suggest that even this small difference may be

accounted for by the fact that most studies combine data from individuals who had married twice with those who had married more than twice, and there is evidence that individuals in the latter group are less happy in general. Vemer and her colleagues also found that remarried men were more satisfied with their relationships than were remarried women; however, the differences were very small and paralleled the differences found between women's and men's satisfaction with first marriages. The absence of notable differences in marital quality between first marriages and remarriages has also been found in more recent investigations of this issue (cf. McDonald and DeMaris 1995). Thus, taken together, the findings indicate little difference in satisfaction between first marriages and remarriages for either men or women.

**Division of Household Labor.** The literature suggests that the division of household labor also differs for couples in first and subsequent marriages. Ishii-Kuntz and Coltrane (1992) found that husbands in remarriages contributed more to household tasks such as cooking, meal cleanup, shopping, laundry, and housecleaning than did men in first marriages, particularly when couples had only their own biological children. In families in which there were stepchildren, women performed a greater actual number of hours of household labor, but the *proportion* that they contributed relative to their spouses was still smaller than that of women in first marriages. Ishii-Kuntz and Coltrane suggest that this is because the "incomplete institutionalization" of remarriage leads to a reduction in gender traditionalism.

However, other studies suggest that whether the woman has been married previously may have more effect on the division of household labor than whether her husband has been married before. Funder (1986) found that women in second marriages perceived that household labor was shared more equally than in their first marriages; however, there was no such trend in men's perceptions. Further, Sullivan (1997) reported that women who had been previously married contributed a smaller proportion of household labor, but that men who were formerly married contributed no differently than did men in first marriages.

**Children in Remarriages.** The large majority of remarriages involve children. Not only do more

than half of formerly married individuals bring children to their new marriages (Buckle et al. 1996), but the rate of childbearing in remarriages is relatively high. Approximately half of all women who enter remarriage under 35 years of age bear at least one child during that marriage, generally within the first two years (Wineberg 1990).

One question that is often raised in both the popular and the scholarly literature is the effect of the presence of children on marital quality. The presence of children has generally been found to have a negative effect on parents' marital quality; however, the effects of children specifically on *remarried* couples is less clear. While White and Booth (1985) found that the presence of stepchildren was associated with somewhat lower marital quality, Martin and Bumpass (1989) found no effect, and both Albrecht and colleagues (1983) and Kurdek (1989) reported that the presence of stepchildren was weakly but *positively* associated with marital quality. The findings regarding mutual children are also inconsistent. Ganong and Coleman (1988) found that mutual children had no effect on marital quality, but Albrecht and colleagues (1983) found a weak but positive effect. Thus, it is unclear how the presence of either mutual children or stepchildren affects marital quality among remarrieds.

Another question that is often the focus of research on stepfamilies (or blended families as they are increasingly labeled) concerns the effect of the remarriage on children. These effects are of considerable importance, given that about 15 percent of all children live in blended families, and it is estimated that between one-third and one-half of today's young people will become stepsons or stepdaughters at some point (Glick 1989; U.S. Bureau of the Census 1997). In the majority of these cases, the children in stepfamilies live with a biological mother and a stepfather.

Remarriage presents several challenges to the family. Members must make numerous decisions and adjustments to living arrangements and family relationships. In addition, parenting approaches and activities may change following the remarriage. Further, remarriage creates a complex set of relationships between former spouses, between stepparents and stepchildren, between step- and half-siblings and extended kin. These relationships are

often ill defined and lack the support of social expectations and norms. The ambiguous nature of these relationships can be seen when members of blended families are asked whom they consider to be family. While only 10 percent of children fail to mention a biological parent when asked to define their family, over 30 percent do not mention a stepparent (Furstenberg and Cherlin 1991).

Some studies have found that the bonds between stepparents and stepchildren are somewhat less warm, more conflictual, and less enduring than those between biological parents and children, while others have found substantial levels of closeness between stepparents and stepchildren (Kurdek and Fine 1993; White 1994). Such variation is not surprising, considering the wide variety of roles that stepparents adopt within the family. Some stepparents adopt a parental role, including the formation and maintenance of close emotional bonds with their stepchildren. Others take on the activities and roles of more distant relatives such as aunts or uncles, and still others act like adult boarders in the home, having little involvement with or showing little affection for the children (Arendell 1997).

Several factors appear to influence the nature of the stepparent-stepchild relationship. In a 1994 review, Cherlin and Furstenberg concluded that the primary factor influencing the character of the stepparent-stepchild relationship is the effort made by the stepparent to forge kinlike relations. Further, it appears easier to be a stepfather than a stepmother, in part because children seem more willing to accept substitute fathers than substitute mothers.

Other factors also affect the quality of the stepparent-stepchild relationship. For example, the younger the child at the time of the remarriage, the more likely he or she is to come to view the stepparent as a "real" parent. The more frequent the contact between the nonresidential parent and the child, the less likely the child is to develop a parentlike relationship with the stepparent (Arendell 1997).

Following remarriage, children often must adapt to the presence of step-siblings and half-siblings, as well as to stepparents. Such siblings may present additional challenges for the child's adjustment. For example, children may be asked



to share space and material resources, as well as the attention of their biological parent. Children may also find that their positions in the household hierarchy have changed with the addition of step-siblings; for example, a child may suddenly find that he or she has lost the status of being the “oldest” or the “baby” of the family. Despite these changes, most step-siblings eventually adjust and form close relationships. In fact, even as adults, step-siblings often maintain contact, albeit on a less frequent basis than full siblings (White and Reidman 1992).

Research on blended families provides a mixed picture with regard to the degree to which stepfamilies differ from intact families in terms of children’s adjustment (for reviews, see Amato and Booth 1997; Amato and Keith 1991a, 1991b; Hetherington et al. 1998). Children in stepfamilies appear to be at greater risk than children in intact families. For example, children in divorced and remarried families are more likely than children in intact families to have academic problems; to be less socially responsible; to have lower levels of self-esteem; to be withdrawn; to be less happy; to have trouble concentrating; and to have troubled relationships with parents, siblings, and peers. Adolescents may experience the same negative outcomes; they may also be more likely to drop out of school, to become sexually active at an earlier age, to have children out of wedlock, to be involved in delinquent activities, to abuse drugs and/or alcohol, and to be unemployed. Further, adult offspring of divorced and remarried families have been shown to be less satisfied with their lives, to have higher levels of marital instability, and to have lower levels of socioeconomic attainment.

Fortunately, despite their increased risk, the majority of children from divorced and stepfamilies do not experience these problems, and some studies suggest that many of the detrimental consequences of divorce or remarriage on children are temporary (c.f., Chase-Lansdale and Hetherington 1990; Emery and Forehand 1994; Hetherington 1993).

It is also important to note that the magnitude of differences in well-being and behavioral outcomes linked to family structure is reduced when the well-being and adjustment of the child prior to divorce and remarriage are taken into consideration. Further, there is a tendency for studies of

children’s adjustment following divorce and remarriage to combine different categories of stepfamilies together. This is particularly problematic due to the set of children that experience multiple transitions in family status. Approximately 10 percent of children experience at least two divorces by their custodial parent before they reach the age of 16 (Furstenberg 1988). The inclusion of children of multiple divorces with those of parents who have divorced and remarried a single time may artificially inflate the risks for children associated with remarriage, since these children are at greatest risk for negative outcomes. In fact, Kurdek (1994) argues that it is the children of the multiple-divorce group—not the stepfamily group—that are most at risk for negative outcomes when compared to children of intact families.

Last, it is essential to point out that remarriage may help to compensate for some of the negative consequences of divorce. For example, remarriage improves the financial well-being of children and their divorced mothers. Only 8 percent of children in mother-stepfather households live below the poverty line, compared to 49 percent of children in single-mother households (Cherlin and Furstenberg 1994). Remarriage also provides an additional adult in the household, thus adding more opportunities for interaction between children and adults, taking some of the burden off the custodial parent, and providing another role model for the child. These factors may help to explain why some studies have found that children in blended families are at less risk for negative outcomes than are children in single-parent homes (Amato and Keith 1991b).

In sum, although there are special concerns and greater complexity in family relationships, many blended families manage to form relationships that are close, loving, and lasting, and to function as effective family units, with the same variation in relationship quality found in more traditional family forms.

**Stability of Remarriages.** Clearly, remarriages and first marriages differ in complexity. The remarried couple must develop ways of interacting with the former spouse; with children from the former marriage (regardless of whether the children are minors or adults); and, in some cases, with both the extended kin and the new partner of the

former spouse. In addition, the emotional history of the relationship with the first spouse, positive and negative, may carry over to influence the new relationship, regardless of whether the first marriage was terminated by death or divorce. Further, material possessions and financial considerations emanating from the first marriage often have a significant impact on second marriages.

This complexity, and perhaps the attendant problems, might help to explain the slightly higher rate of divorce among remarried couples. Whereas about half of first remarriages end in divorce, approximately 60 percent of remarriages do so (Ihinger-Tallman and Pasley 1987).

There are several explanations for the higher rate of divorce among the remarried. Cherlin (1978) has suggested that there is an "incomplete institutionalization" of remarriage. Remarriages, according to Cherlin, are more difficult than first marriages due to the absence of guidelines in language, law, and custom for remarried couples.

Another interpretation of the higher rate of divorce among the remarried is offered by Furstenberg and Spanier (1984), who identify the predisposition to divorce of the remarried as the key explanatory factor. According to this perspective, since remarried individuals have already demonstrated their willingness to leave an unsatisfactory marriage, they will be willing to do so again if dissatisfied with the current relationship. Therefore, it may not be the quality of the marital relationship that precipitates divorce but the propensity to leave an unsatisfactory relationship.

Last, Martin and Bumpass (1989) have suggested that another explanation for the higher rate of divorce among the remarried is that these individuals have sociodemographic characteristics that increased the likelihood that their first marriage would end in divorce. These characteristics, such as low educational level, income instability, and parental divorce, are then carried into the remarriage, also increasing its instability. Further, Martin and Bumpass suggest that individuals who remarry are disproportionately likely to have married as teenagers, which may indicate differences in personality or experiences that might make it more difficult to maintain a successful marriage. Consistently with Martin and Bumpass's argument, there is evidence that men who are socially disadvantaged

remarry more quickly than their counterparts who are more advantaged (Monk-Turner and White 1995).

In sum, while the literature on remarriage emphasizes differences between first marriages and remarriages, it is important to reiterate that, overall, there is substantially more similarity than difference between the two. The patterns of mate selection, marital quality, and marital stability of individuals who remarry do not differ markedly from the patterns of individuals marrying for the first time.

## CONCLUSION

In conclusion, married life in America today often involves a sequence of marriage, divorce, and remarriage, sometimes followed by a subsequent divorce. Many sociologists suggest that the high rate of remarriage followed by a subsequent divorce in America society indicates a strong commitment to married life, albeit of a slightly different form. These scholars argue that contemporary Americans are not rejecting marriage, they are only rejecting specific relationships that became unsatisfying. Taken together, the literature on remarriage and stepfamilies suggests that the individuals who choose to follow this pattern have approximately as great a likelihood of finding a satisfying and stable relationship as do those who marry for the first time.

For further discussions and reviews of work regarding remarriage, see Bumpass and colleagues (1990), Ganong and Coleman (1994), and Vemer and colleagues (1989). For a further discussion of literature on the effects of divorce and remarriage on children, see Amato and Booth (1997), Amato and Keith (1991a, 1991b), Cherlin and Furstenberg (1994), and Hetherington and colleagues (1998).

## REFERENCES

- Albrecht, Stan L., Howard M. Bahr, and Kristen L. Goodman 1983 *Divorce and Remarriage: Problems, Adaptations and Adjustments*. Westport, Conn.: Greenwood.
- Amato, Paul R., and Alan Booth 1997 *A Generation at Risk: Growing Up in an Era of Family Upheaval*. Cambridge: Harvard University Press.
- Amato, Paul R., and Bruce Keith 1991a "Parental Divorce and Adult Well-Being: A Meta-Analysis." *Journal of Marriage and the Family* 53:43-58.

- 1991b "Parental Divorce and the Well-Being of Children: A Meta-Analysis." *Psychological Bulletin* 11:26-p46.
- Arendell, Terry (ed.) 1997 *Contemporary Parenting Challenge, and Issues*. Thousand Oaks, Calif.: Sage.
- 1983 "Divorce and Remarriage." *Contemporary Parenting: Challenges and Issues*. Thousand Oaks, Calif.: Sage.
- Bolgar, R., H. Sweig-Frank, and J. Paris 1995 "Childhood Antecedents of Interpersonal Problems in Young Adult Children of Divorce." *Journal of the Academy of Child and Adolescent Psychiatry* 34:143-150.
- Buckle, Leslie, Gordon G. Gallup, Jr., and Zachary A. Rodd 1996 "Marriage as a Reproductive Contract: Patterns of Marriage, Divorce, and Remarriage." *Ethology and Sociobiology* 17:363-377.
- Bumpass, Larry, James Sweet, and Teresa Castro Martin 1990 "Changing Patterns of Remarriage." *Journal of Marriage and the Family* 52:747-756.
- Bumpass, Larry, James Sweet, and Andrew Cherlin 1991 "The Role of Cohabitation in Declining Rates of Marriage." *Journal of Marriage and the Family* 53:913-927.
- Chase-Lansdale, P. Lindsay, and E. Mavis Hetherington 1990 "The Impact of Divorce on Life-Span Developmental: Short and Long Term Effects." In D. Featherman and R. Lerner, eds., *Life Span Development and Behavior*, vol. 10. Hillsdale, N.J.: Lawrence Erlbaum.
- Cherlin, Andrew J. 1978 "Remarriage as an Incomplete Institution." *American Journal of Sociology* 84:634-650.
- , and Frank F. Furstenberg 1994 "Stepfamilies in the United States: A Reconsideration." *Annual Review of Sociology* 20:359-381.
- Clarke, Sally C. 1995 "Advance Report of Final Marriage Statistics 1989 and 1990." *Monthly Vital Statistics Report* 43 (July 14). Washington, D.C.: U.S. Department of Health and Human Services.
- Demos, John 1970 *A Little Commonwealth: Family Life in Plymouth Colony*. London: Oxford University Press.
- Emery, Robert E., and R. Forehand 1994 "Parental Divorce and Children's Well-Being: A Focus on Resilience." In R. J. Haggerty, L. R. Sherrod, N. Garnezy, and M. Rutter, eds., *Stress, Risk, and Resilience in Children and Adolescents*. Cambridge, England: Cambridge University Press.
- Funder, K. 1986 "His and Her Divorce." In Peter McDonald, ed., *Setting Up: Property and Income Distribution on Divorce in Australia*. Sidney: Australian Institute of Family Studies, Prentice Hall.
- Furstenberg, Frank F. 1988 "Child Care after Divorce and Remarriage." In E. Mavis Hetherington and J. D. Arasteh, eds., *Impact of Divorce, Single Parenting, and Stepparenting on Children*. Hillsdale, N.J.: Erlbaum.
- , and Andrew J. Cherlin 1991 *Divided Families*. Cambridge, Mass.: Harvard University Press.
- Furstenberg, Frank F., and Graham B. Spanier 1984 "The Risk of Dissolution in Remarriage: An Examination of Cherlin's Hypothesis of Incomplete Institutionalization." *Family Relations* 33:433-441.
- Ganong, Lawrence H., and Marilyn Coleman 1988 "Do Mutual Children Cement Bonds in Stepfamilies?" *Journal of Marriage and the Family* 50:687-698.
- 1994 *Remarried Family Relationships*. Thousand Oaks, Calif.: Sage.
- Glick, Paul C. 1989 "Remarried Families, Stepfamilies, and Stepchildren: A Brief Demographic Analysis." *Family Relations* 38:7-26.
- Goldscheider, Frances Kobrin, and Linda J. Waite 1986 "Sex Differences in the Entry into Marriage." *American Journal of Sociology* 92:91-109.
- Haskey, J. 1987 "Social Class Differentials in Remarriage after Divorce: Results from a Forward Linkage Study." *Population Trends* 47:34-42.
- Hetherington, E. Mavis 1993 "An Overview of the Virginia Longitudinal Study of Divorce and Remarriage with a Focus on Early Adolescence." *Journal of Family Psychology* 7:39-56.
- , Margaret Bridges, and Glendessa M. Insabela 1998 "What Matters? What Does Not? Five Perspectives on the Association between Marital Transitions and Children's Adjustment." *American Psychologist* 53:167-184.
- Ihinger-Tallman, Marilyn, and Kay Pasley 1987 *Remarriage*. Thousand Oaks, Calif.: Sage.
- Ishii-Kuntz, Masako, and Scott Coltrane 1992 "Remarriage, Stepparenting, and Household Labor." *Journal of Family Issues* 13:215-233.
- Koo, Helen P., and C. M. Suchindran 1980 "Effects of Children on Women's Remarriage Prospects." *Journal of Family Issues* 1:497-515.
- Kurdek, Lawrence A. 1989 "Relationship Quality for Newly Married Husbands and Wives: Marital History, Stepchildren, and Individual-Difference Predictors." *Journal of Marriage and the Family* 51:1053-1064.
- 1994 "Remarriages and Stepfamilies Are Not Inherently Problematic." In Alan Booth and Judy Dunn, eds., *Stepfamilies: Who Benefits? Who Does Not?* Hillsdale, N.J.: Lawrence Erlbaum.

- , and Mark A. Fine 1993 “The Relation between Family Structure and Young Adolescents’ Appraisals of Family Climate and Parenting Behavior.” *Journal of Family Issues* 14:279–290.
- Lichter, Daniel T., Diane K. McLaughlin, George Kephart, and David J. Landry 1992 “Race and the Retreat from Marriage: A Shortage of Marriageable Men?” *American Sociological Review* 57:781–799.
- MacDonald, William L., and Alfred DeMaris 1995 “Remarriage, Stepchildren, and Marital Conflict: Challenges to the Incomplete Institutionalization Hypothesis.” *Journal of Marriage and the Family* 57:387–398.
- Martin, Teresa Castro, and Larry Bumpass 1989 “Recent Trends in Marital Disruption.” *Demography* 26:37–51.
- Monk-Turner, Elizabeth, and Garland White 1995 “Factors Shaping the Probability of Divorce and Early Remarriage among Young Men.” *International Journal of Contemporary Sociology* 32:97–105.
- Sullivan, Oriel 1997 “The Division of Housework among ‘Remarried’ Couples.” *Journal of Family Issues* 18:205–223.
- Sweeney, Megan M. 1997 “Remarriage of Women and Men after Divorce.” *Journal of Family Issues* 18:479–502.
- Sweet, James A., and Larry L. Bumpass 1988 *American Families and Households*. New York: Russell Sage Foundation.
- U.S. Bureau of the Census 1997 *Statistical Abstract of the United States, 1997*. Washington, D.C.: Government Printing Office.
- 1998 *Statistical Abstract of the United States, 1998*. Washington, D.C.: Government Printing Office.
- Vemer, Elizabeth, Marilyn Coleman, Lawrence H. Ganong, and Harris Cooper 1989 “Marital Satisfaction in Remarriage: A Meta-Analysis.” *Journal of Marriage and the Family* 51:713–725.
- White, Lynn K. 1994 “Coresidence and Leaving Home: Young Adults and Their Parents.” *Annual Review of Sociology* 20:81–102.
- , and Alan Booth 1985 “The Quality and Stability of Remarriages: The Role of Stepchildren.” *American Sociological Review* 50:689–698.
- White, Lynn K., and Agnes Riedman 1992 “When the Brady Bunch Grows Up: Step/Half and Full-Sibling Relationships in Adulthood.” *Journal of Marriage and the Family* 54:197–208.
- Wilson, Barbara Foley 1989 “Remarriages and Subsequent Divorces” *National Vital Statistics*, series 21, no. 45. Hyattsville, Md.: National Center for Health Statistics.

- Wineberg, Howard 1990 “Childbearing after Remarriage.” *Journal of Marriage and the Family* 52:31–38.

J. JILL SUITOR  
SHIRLEY A. KEETON

## REPLICATION

Philosophers have long identified replication as an important facilitator of scientific progress. Several terms have been used to denote the ability to assess past work through replication, including “intersubjective testability,” “reliability,” and “verifiability by repetition.” Authors of scientific papers typically describe the methods and materials they used in their research so that, at least hypothetically, others can repeat the work and reproduce the reported results. Successful replication of their own and others’ work gives researchers confidence in its validity and reassures them about the fruitfulness of the general line of inquiry they are following. In contrast, inability to replicate one’s own or others’ results casts doubt upon the validity of the previous work. Critics argue that because sociologists infrequently attempt to replicate findings, they are both less able to identify valid lines of inquiry and more likely to follow spurious ones.

One can identify a continuum ranging from exact to weakly approximate replication. The former, also called repetition, consists of attempts to use the same materials and procedures as previous research to determine whether the same results can be obtained. Approximate replication, on the other hand, consists of using some but not all of the conditions of a previous study. By systematically varying research conditions in a series of approximate replications, it may be possible to determine the precise nature of a previous study’s results and the extent to which they also hold for different populations and situations (Aronson et al. 1998). Researchers usually value successful approximate replication more than successful exact replication because the latter contributes less to existing knowledge.

In the natural sciences, experimentalists are usually expected to carry out successful exact replications of their own work before submitting it for

publication. This reduces the likelihood of reporting spurious results and of misleading one's colleagues. Exact replications of others' research are often difficult and costly to execute, however, and natural scientists rarely attempt them except in cases where the original work is theoretically important or has high potential practical value, or where there is suspicion of fraud. Another disincentive for carrying out exact replications of already published work is that such work is usually difficult to publish. This is true not only because little new knowledge results from an exact replication, but also because the meaning of a failure to replicate exactly is often ambiguous. Failures can indicate that the original work was flawed, but they may also be due to inadequate specification of research procedures, the existence of a stochastic element in the production of results, or errors in the replication itself (Harry M. Collins 1985). By contrast, approximate replications, especially those involving the modification of research instruments and their application to new areas of inquiry, are common in the natural sciences, and this has led some to identify them as constituting a central element of "rapid-discovery, high-consensus science" (Randall Collins 1994).

Many hold that social scientists' opportunities to carry out replications, especially exact replications, are severely limited. This is partly because social scientists often use nonexperimental research techniques that are difficult to repeat exactly. In addition, changing social and historical contexts can influence studies' results. As a result, failures to obtain the same results as reported by previous studies are even more ambiguous in the social sciences than in the natural sciences (Schuman and Presser 1981). This ambiguity may account for social scientists' continued interest in concepts and theories stemming from studies whose results have repeatedly failed to be replicated (e.g., sex differences in fear of success and patterns of moral development).

Nevertheless, critics have long argued that behavioral scientists need to attempt more replications of previous research because their dependence on statistical inference produces many spurious reports of "statistically significant" results. Statistical inference allows researchers only to reject or fail to reject a null hypothesis. Each of these

two outcomes is subject to error due to the probabilistic nature of statistical hypothesis testing; sometimes researchers reject null hypotheses that are actually true (type one error), and sometimes they fail to reject null hypotheses that are actually false (type two error). However, failure to reject a null hypothesis does not justify accepting it, and studies that do not yield rejections therefore are often judged as contributing little. As a result, scholarly journals tend to publish only papers that report the rejection of null hypotheses, some of which are the result of type one errors (Sterling 1959). Furthermore, to ensure that they will be able to reject null hypotheses, researchers sometimes use inappropriate analytic procedures that maximize their chances of obtaining statistically significant results (Selvin and Stuart 1966), increasing the likelihood that published findings are due to type one errors. To counteract these patterns, some have argued that behavioral science editors should set aside space in their journals for the publication of replication attempts, and to publish studies that fail to replicate earlier results even when the replications themselves fail to reject null hypotheses.

Despite the calls for increased replication, behavioral science journals publish few papers reporting replication attempts. In an early examination of this issue, Sterling (1959) reported that among 362 articles in psychology journals, 97 percent of those reporting a test of significance rejected the null hypothesis, but that none was an explicit replication. Ironically, many have replicated Sterling's results (cf. Dickersin 1990; Gaston 1979; Reid et al. 1981). These studies probably underestimate the prevalence of replication, because they do not count papers reporting a set of experiments that comprise both an original result and one or more approximate replications of it. By not encouraging more replication, however, behavioral science journals may foster elaborate and vacuous theorizing at the expense of identifying factual puzzles that deserve theoretical analysis (Cook and Campbell 1979, p. 25).

Although the traditional view of replication entails the collection of new data—including data on additional cases or additional measures—statisticians and social scientists have suggested alternative replication strategies. One is to build replication into a study from the start. For example, a researcher can draw a sample large enough to allow its random partition into two subsamples.

Data from one subsample can then be used to check conclusions drawn on the basis of analyses of data from the other. Another approach, requiring the intensive use of computing resources, is to draw multiple random subsamples from already collected data and then use these subsamples to crossvalidate results (Finifter 1972). This general strategy, which includes such techniques as “jackknifing” and “bootstrapping,” is also used to assess sampling variances for complex sampling designs (see *Sampling Procedures*). Still another elaboration of the basic idea of replication is the general approach called *meta-analysis*. Here the analyst treats previous studies on a topic or relationship as a sample of approximate replications. By statistically analyzing whether and how studies’ results vary, one can determine how generalizable a finding is and the extent to which differences in study design account for variation in results (Hunter and Schmidt 1990). Finally, replication may also be fostered by the increased availability of already-collected data sets stemming from the establishment of data depositories, and funding agency requirements that data from supported projects be made accessible to other researchers. Access to previously collected data makes it possible to carry out both exact replications of previous analyses and approximate replications that alter the analytic procedures used by the original researcher.

(SEE ALSO: *Sampling Procedures*)

#### REFERENCES

- Aronson, Elliot, Timothy D. Wilson, and Marilyn B. Brewer 1998 “Experimentation in Social Psychology.” In Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *Handbook of Social Psychology*, 4th ed. Boston: McGraw-Hill.
- Collins, Harry M. 1985 *Changing Order: Replication and Induction in Scientific Practice*. London: Sage.
- Collins, Randall 1994 “Why the Social Sciences Won’t Become High-Consensus, Rapid-Discovery Science.” *Sociological Forum* 9:155–177.
- Cook, Thomas D., and Donald T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Studies*. Chicago: Rand-McNally.
- Dickersin, Kay 1990 “The Existence of Publication Bias and Risk Factors for Its Occurrence.” *Journal of the American Medical Association* 263:1385–1389.
- Finifter, Bernard M. 1972 “The Generation of Confidence: Evaluating Research Findings by Random Subsample Replication.” In Herbert L. Costner, ed., *Sociological Methodology 1972*. San Francisco: Jossey-Bass.
- Gaston, Jerry 1979 “The Big Three and the Status of Sociology.” *Contemporary Sociology* 8:789–793.
- Hunter, John E., and Frank L. Schmidt 1990 *Methods of Meta-Analysis*. Newbury Park, Calif.: Sage.
- Reid, L. H., L. C. Soley, and R. D. Rimmer 1981 “Replications in Advertising Research: 1977, 1978, 1979.” *Journal of Advertising* 10:3–13.
- Schuman, Howard, and Stanley Presser 1981 “Mysteries of Replication and Non-Replication.” In Howard Schuman and Stanley Presser, eds. *Questions and Answers in Attitude Surveys*. New York: Academic Press.
- Selvin, Hannan C., and Alan Stuart 1966 “Data-Dredging Procedures in Survey Analysis.” *American Statistician* 20:20–23.
- Sterling, Theodore D. 1959 “Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance—Or Visa Versa.” *Journal of the American Statistical Association* 54:30–34.

LOWELL L. HARGENS

## RESEARCH FUNDING IN SOCIOLOGY

### IMPORTANCE OF EXTERNAL FUNDING

Jesse Unruh, a well-known twentieth-century politician in California, is alleged to have said, “Money is the mother’s milk of politics.” Observers at the century’s close may have made similarly accurate comments about the importance of money in the conduct of research. Natural scientists (including those in fields ranging from basic biomedical science to space exploration) have traditionally depended on funds beyond the internal resources of the organizations at which they are employed. Major discoveries in the natural sciences would be unthinkable without funding for particle accelerators, field expeditions, and clinical research staffs. Social scientists and particularly sociologists may be moving toward similar dependence on funding from beyond the boundaries of their universities and institutes. Such resources appear crucial for truly significant investigations, at least those which require original empirical data.

Sponsorship citation in the leading sociology journals in 1997 and 1998 provides an indication of the importance of external research support to

sociologists. Volume 62 (1997) of the *American Sociological Review* (*ASR*) contains 55 articles. In 31 of these articles (56 percent), the authors cite sources of funding outside their home organizations. An additional 3 authors, all university based, cite internal sources of support (such as university research committees), which presumably grant awards through a competitive proposal process. *American Journal of Sociology* (*AJS*), volume 103 (1997 and 1998), contains 27 articles; 9 of these cite funding sources, 6 (22 percent) of which are clearly outside agencies or internal units distributing funds obtained from outside. At century's end, external research funding appeared to be a frequent if not ubiquitous part of the research process in sociology.

Review of research published a generation ago presents a similar picture. Volume 42 of *ASR* (1977) contains 59 articles, 23 (39 percent) of which cite outside funding sources. An additional 4 authors cite funding sources internal to their institutions. Volume 83 of the *AJS* (1977 and 1978) includes 38 articles, of which 24 (63 percent) cite funding from sources outside the authors' universities, institutes, or agencies. An additional 3 *AJS* articles cite internal sources.

### SOURCES OF FUNDING

Sources named in the above-cited *AJS* and *ASR* volumes provide an indication of where sociologists obtain research funds. Among the articles published in 1997 and 1998 that cite external funding, a plurality (17 citations) named the National Science Foundation (NSF) as the sole or primary source of support. The second most frequently cited source was the National Institute of Child and Human Development (NICHD), with three citations, followed by the National Institute on Aging (NIA), with two citations. One article cited the National Institute of Mental Health (NIMH) and the remainder acknowledged a variety of private funders, such as the Ford, Rockefeller, and Guggenheim foundations.

Among 1977 and 1978 journal articles that cited external funding, the National Science Foundation (NSF) received the most frequent acknowledgement (16 citations), followed by NIMH (eight citations), the Office of Economic Opportunity (four citations), the Ford Foundation (four citations), the National Institute of Education (two

citations), and the Russell Sage Foundation (two citations). An assortment of federal agencies and private foundation received one citation each, including the Social Security Administration (SSA), the Department of Labor (DOL), the National Endowment for the Humanities (NEH), the National Institute for Drug Abuse (NIDA), and the Grant and Guggenheim foundations.

Comparison of funding patterns in *ASR* and *AJS* across the decades carries some risk. These journals do not necessarily represent the most important work of general sociology and may not reflect patterns of support (or its absence) in subdisciplines, such as medical sociology and sociology of religion. Conventions and habits prevailing among authors regarding citation of funding sources may have changed between the 1970s and the 1990s, but inspection of acknowledgements in the articles cited above suggests that external funding of major sociological research remained as common a phenomenon in the late 1990s, when 43 percent of *ASR* and *AJS* articles cited external funding, as it was in the late 1970s, when 47 percent of *ASR* and *AJS* articles acknowledged such support. Funding of the research leading to these articles depended somewhat more on public agency support in the late 1970s than in the late 1990s. About 80 percent of the above-referenced 1977 and 1978 *ASR* and *AJS* articles which acknowledged external funding cited public sources and 20 percent cited private sources. About 71 percent of comparable *ASR* and *AJS* articles in 1997 and 1998 acknowledged public funding, 29 percent citing private sources.

### STRUCTURE AND OPERATION OF FUNDING ORGANIZATIONS

Public and private sources of research funding comprise different social worlds. Representatives of each evaluate research proposals according to different rules and criteria. These differences derive from the distinct institutional surroundings of public agencies and private organizations.

**Public Funding Organizations.** Procedures followed by the U.S. Public Health Service (PHS) illustrate the processes by which public research funding takes place and the organizational components which facilitate these processes. The PHS, a subunit of the Department of Health and Human

Services (DHHS), houses a large number of agencies making research grants. These include several of the sources cited in the above-mentioned *AJS* and *ASR* volumes, such as the National Institute of Mental Health (NIMH), the National Institute of Aging (NIA) and the National Institute of Child Health and Disease (NICHD). Other PHS subunits such the National Cancer Institute (NCI) and the Agency for Health Care Policy and Research (AHCPR) have funded significant sociological work, principally in the fields of organizational and medical sociology. It appears likely that other governmental funding organizations, such as NSF, operate in a manner similar to these PHS units.

The institutional background and functioning of AHCPR provide a paradigm applicable to all federal funding organizations. Federal legislation established AHCPR through the Omnibus Budget Reconciliation Act of 1989 (AHCPR 1992). The legislation authorizes the secretary of DHHS to undertake "research, demonstration, and evaluation activities" regarding delivery of health services in the United States. The AHCPR administrator acts as operational head of the agency and derives his or her authority via delegation from the secretary of DHHS. By statute, the administrator is advised by the National Advisory Council for Health Care Policy, Research, and Evaluation. The council recommends priority areas for research funding. In addition, it provides "secondary review" of grant proposals. All PHS awarding organizations have a similar administrative structure and advisory council.

The council's composition is mandated to include seventeen public (nonfederal) members who have voting power. These are primarily scientifically qualified individuals. The council also includes health practitioners; individuals drawn from business, ethics, law, and policy; and consumer representatives. In addition, the council includes seven federal officials with voting power, appointed to serve *ex officio*.

Of key importance in the process of research funding by AHCPR and analogous groups is the Initial Review Group (IRG), sometimes referenced as the "scientific review group" or "study section." As mandated by statute, the IRG advises the secretary of DHHS on the scientific and technical merit of research grant applications within AHCPR's areas of responsibility. The AHCPR administrator

officially invites individuals selected for IRG membership to join. Selection is made with assistance from the Office of Scientific Review (OSR), the administrative body within the agency with responsibility for organizing the review of applications for grant support and reporting IRG findings to the AHCPR administrator. Normally, IRG members are appointed to overlapping four-year terms.

Official criteria for IRG membership include scientific expertise in areas of concern to the agency, often indicated by a history of receiving funds under the agency's jurisdiction. IRG members cannot be employees of the federal government. In 1992, AHCPR operated four IRGs, each covering a different specialty area or agency concern. Agencies such as AHCPR occasionally appoint special panels to evaluate proposals obtained in response to Requests for Applications (RFAs) issued to solicit interest in new or unusual areas of scientific concern.

The IRG meets three times each year, a few months after regularly scheduled deadlines for submission of grant applications. An official of the Office of Scientific Review, known as the Scientific Review administrator, in consultation with the IRG chair, assigns committee members the task of reviewing each application received by the agency. For each application, the Scientific Review administrator appoints a primary reviewer and at least two secondary reviewers.

At IRG meetings, primary and secondary reviewers present evaluations of the scientific merit of the proposals which they have been assigned. Discussion then takes place among all members of the IRG. The chair then requests a recommendation from the primary reviewer to assign a "priority score" to the proposal, defer discussion, or remove the proposal from further consideration. Priority scores are assigned ranging from 1.0 (most meritorious) to 5.0 (least meritorious). Following the reviewers' reports and ensuing discussion, all voting members of the IRG formulate priority scores according to their own judgment and submit them to the chair. Composite priority scores and summaries of the IRG's findings regarding scientific merit are reported to the council and the grant applicant.

The agency establishes a "pay line" indicating the priority score below which proposals are likely



to receive funding. In the late 1990s, pay lines varied between priority scores of 1.80 and 2.20, criteria stringent in comparison with the early 1970s, when pay lines were often in the 300s. Strictly speaking, the findings of the IRG are advisory to the council and ultimately to the secretary. Rarely, the council or a high official may elect to fund a proposal "out of priority" due to extraordinary merit or urgent public need.

Government research organizations such as AHCPR review all grant applications that are complete and received on time. In a mid-1992 funding cycle, IRGs (including all those of AHCPR, the National Institute of Health [NIH], and related agencies) reviewed 8,017 applications (NIH Division of Research Grants 1993). Three such cycles occur annually. IRGs usually award priority scores resulting in funding of between 10 and 20 percent of the applications they review. Structure and procedures at the National Science Foundation appear somewhat simpler but similar in form to those of the PHS. Proposals are reviewed by technical staff (an NSF program officer) with the aid of outside reviewers. Final award decisions are made at the senior management level.

**Private Funding Organizations.** If review of funding sources cited in *ASR* and *AJS* provides a valid indication, most nongovernment funding for sociological research comes from private foundations. None of the journal articles acknowledge funding from corporations, which are major sources of research support in some fields, such as engineering and pharmacy.

A private foundation is a nongovernmental, nonprofit organization with funds from a permanent portfolio of investments known as an endowment. Typically, an endowment originates from a single source, such as an individual, a family, or a corporation. Legally, foundations are chartered to maintain or aid social, educational, religious, or other charitable activities serving the common good. They are owned by trustees who hold the foundation's assets "in trust" for the people of the jurisdiction in which they are chartered. Governance is carried out by a board of directors. Trustees usually sit on boards of directors, but directors are not always trustees. Foundations of significant size employ staff to recommend policy, evaluate applications for funding, and perform day-to-day administrative tasks.

Like federal agencies, private foundations identify areas of interest and disseminate this information to the grant-seeking community. New interests and areas of emphasis emerge periodically through initiation by staff or discussion among directors and trustees. Large foundations such as Robert Wood Johnson and W. K. Kellogg seem to change their areas of emphasis approximately every decade or upon accession of a new president. Foundations often develop initiatives articulating interests focused on specific concerns and issue program announcements and RFAs in these areas.

The process of evaluating grant requests appears considerably less formal in private foundations than it is in government. Typically, foundations advise potential grant applicants to submit short letters of interest as a first step. Typically, lower-level staff read these letters, screening them for conformity with the foundation's interests, credibility of the prospective applicant, and such nonstandard criteria for consideration as the foundation might maintain. Letters that pass this screening process are transmitted to higher-level staff, which may ask the applicant for a detailed proposal. Some foundations assemble review panels and hire outside consultants to evaluate proposals of a technical nature. More typically, though, foundation officials discuss full-scale proposals among themselves and formulate recommendations to the governing board, which makes final award decisions.

The percentage of letters of interest which result in eventual funding is quite low. Most letters of inquiry generate a notice of rejection via form letter. Likelihood of funding appears roughly comparable with federal sources. In 1999, the Robert Wood Johnson Foundation's Web site advised potential applicants that their chances of success were 1 in 20. In the 1990s, this foundation operated an investigator program in health policy. Each year, this program received over 400 letters of inquiry and awarded no more than 10 grants.

Some features of the private foundation's informality appear advantageous to applicants. Federal grant application forms are extremely long and detailed, requiring significant effort for completion. Foundation forms tend to be simpler. Many foundations require no standard form. Generally, though, federal agencies provide larger

amounts of money to grantees than private foundations. In the late 1900s, for example, few Ford Foundation grants exceeded \$100,000. Several AHCPR awards approached \$1 million. Generous overhead recovery rates provided by federal grants are attractive to university officials. Many private foundations severely restrict overhead payments or allow no such funds in grantee budgets.

### THE POLITICS OF RESEARCH FUNDING

The bureaucratic structure of public funding organizations clearly aims at promoting accountability and fairness. Legally, much of the grant award process is visible to the public. All successful grant applications may be obtained through the Freedom of Information Act. IRG deliberations, the venue in which key funding decisions take place, tend to be conscientiously conducted discussions in the manner of serious graduate-level seminars.

Key features of the operation of agencies such as AHCPR, though, raise questions about the degree to which awards are made strictly according to scientific merit. Deliberations of the Scientific Review Administrator and the agency administrator which result in IRG membership selection are subject to the same uncertainty as all dyadic interactions. Disciplinary bias, partiality to particular questions and approaches, and personal ties may affect IRG membership appointments. IRG membership, of course, strongly affects the agency's funding pattern.

Assignment of primary review responsibilities by the Scientific Review administrator and the IRG chair can also have a profound effect on funding decisions. Positively or negatively biased assignment of the primary reviewer responsibility can determine favorable versus unfavorable outcome. Although all IRG members are expected to read all applications, only the primary and secondary reviewers are expected to do so in detail. IRG members look to these individuals for guidance in their determination of priority scores. Given the competitiveness of funding, most applications have no chance of receiving priority scores below the pay line without strong support by their primary reviewers. In a sense, the primary reviewer must function as an advocate for the applicant. Officials

prejudiced against a particular investigator or approach could ensure selection of a primary reviewer who shared their negative inclination.

Advocacy plays a role of similar importance in private foundation funding. Communication with a foundation official is a virtual necessity prior to submission of a letter of interest. Established relations with these officials is necessary to avoid the screening process that eliminates most grant-seekers and to promote favorable action at each step in the decision-making process.

### REFERENCES

- National Institute of Health, Division of Research Grants 1993 *Grant Application Statistics*. NIH Peer Review Notes, (February), p. 8.
- Agency for Health Care Policy and Research 1992 *Orientation Handbook for Members of AHCPR Review Groups*. Rockville, Md.: Agency for Health Care Policy and Research.

HOWARD P. GREENWALD

## RESEARCH METHODS

See Case Studies; Comparative-Historical Sociology; Ethnomethodology; Ethnography; Evaluation Research; Qualitative Methods; Survey Research; Statistical Methods.

## RETIREMENT

Retirement is primarily a twentieth-century phenomenon that developed through a convergence of public and private employment policies, a restructuring of the life span relative to work activity, and a redefinition of the terms of monetary compensation for work performed. It may be tempting to view retirement as the "natural" development of a social institution matched to the needs of older people experiencing declines in capacity; but the invention of a distinctive nonemployment status called *retirement* was not simply a response to human aging. Rather, in reconciling a transformed economy to an aging population with an increasing amount of surplus labor, an explicit policy of job distribution was produced. Retirement policies incorporated age as a characteristic that served

as both a qualifying and an exclusionary principle for work and income. The fact that these policies were age-based can be linked to the social production of age as a predictor of individual capacity and potential, a production that had ideological roots in the science and larger culture of the time.

### HISTORICAL DEVELOPMENTS

Whereas *retirement contracts* existed in both Europe and colonial America, Plakans (1989) argues that preindustrial retirement was a gradual transition. The head of a household transferred legal title to an heir in exchange for some combination of monetary payments, material provisions, and services as stipulated by the aged person or couple. These contracts were typical of agrarian economies in which land was the main factor in production; they represented the final step in a long and sometimes elaborate process of property transfer. These “stepping-down” practices were therefore most immediately linked to inheritance patterns; they could be used to ensure that family control of the land was maintained (Sorensen 1989).

Between 1790 and 1820, American legislatures introduced policies of mandatory retirement for certain categories of public officials. By the late 1800s, the majority of businesses still had no formal policies of fixed-age retirement. Instead, informal policies eliminated older workers from the labor force (Fischer 1977). This decline in the demand for older workers can be linked to changes in the structure of American capitalism. During the late 1800s the structure of American capitalism began to change from small-producer, competitive capitalism to large-scale corporate capitalism (Sklar 1988). Part of this reconstruction involved attempts to rationalize age relations in the workplace, a process that was embedded in a more general disenchantment with older workers and a devaluation of their skills. Indeed, the employment rates for men aged 65 and older showed a steady decline during this period, from 80.6 percent in 1870 to 60.2 percent in 1920 (Graebner 1980). According to Graebner’s analysis, retirement became the impersonal and egalitarian method adopted by both public and private employers for dealing with superannuated workers. It allowed employers to routinize the dismissal of older workers, thereby restructuring the age composition of their workforces in a way they believed

would enhance efficiency, a belief supported by the principles of scientific management. Pension plans legitimized this process and, at the same time, served as an effective labor control device.

The first pension plan (1875) is credited to the American Express Company, but benefits were restricted to permanently incapacitated workers who were at least 65 years old with a minimum of 20 years of service (Schulz 1976). In 1920, the first general piece of retirement legislation, the Civil Service Retirement Act, provided pension coverage for federal civilian employees. One year later, the Revenue Act of 1921 encouraged businesses to implement private plans by exempting both the income of pension and profit-sharing trusts and the employer contributions to these trusts from income tax liability. Nevertheless, coverage remained concentrated in a few industries, and 85 percent of the workforce continued to be without pension coverage.

By the 1930s, the problem of superannuated workers was coupled with the more general problem of managing surplus labor. The changing technology of the workplace helped transform the labor process. A subsequent increase in worker productivity and the growing recognition of the cyclical crises inherent in industrial capitalism broadened the concern beyond that of simple superannuation to that of job distribution and consumption capacity.

The Depression of the 1930s greatly exacerbated the growing problem of old-age poverty and unemployment. By 1935 unemployment rates among those 65 and older were well over 50 percent. Even those with pension benefits did not escape poverty; trade union plans were collapsing, and state and local governments were reducing or discontinuing pension payments (Olsen 1982). Legislative proposals for alleviating some of these problems included the Townsend Plan and the Lundeen Bill. The Townsend Plan proposed a flat \$200 monthly pension for older Americans; recipients had to spend the pension within thirty days. The Lundeen Bill proposed benefits at the level of prevailing local wages for all unemployed workers aged 18 and older (including the elderly) until suitable employment was located. Neither of these plans was directly related to a retirement transition. The Townsend Plan granted equal benefits to all nonemployed persons over age 60. The Lundeen

Bill focused more on job creation for workers of all ages than on limiting labor supply through age exclusions.

In 1934, President Franklin Roosevelt appointed the Committee on Economic Security (CES) to develop legislation to address the problems of old-age poverty and unemployment. The Social Security Act of 1935 offered a solution that based benefits on the level of workers' contributions to a general trust fund. Upon their retirement, covered workers (primarily those in manufacturing) could draw retirement benefits, assuming they met the age and work eligibility requirements.

For the CES, retirement referred to complete withdrawal from the labor force. As stated by Barbara Armstrong, an original member of the CES, "[r]etirement means that you've stopped working for pay." According to Armstrong, the option facing the Roosevelt administration pitted older workers against younger workers (Graebner 1980, p. 186). Retirement would reduce unemployment by moving older workers out of the labor force, allowing industries characterized by surplus labor to transfer jobs from older to younger workers. The federal government could facilitate this process by shifting the focus of older workers' financial dependency from the wage contract to a federal income maintenance program. In that sense, the Social Security Act of 1935 established what was primarily a program of old-age relief; its limited coverage and low benefit levels precluded its serving as an effective instrument of retirement. However, in establishing a measure of income support for retired older workers, the act reinforced the view that in the competition for jobs, age was a legitimate criterion, and youth had the "higher claim" (Graebner 1980). Ironically, the mobilization for World War II created job opportunities for older workers, as it did for women. Even though these opportunities proved temporary, they challenged the connection between retirement and superannuation, a connection that was asserted more emphatically when the supply of labor exceeded the demand.

During the next several decades, considerable growth in private pension plans occurred; coverage increased from 4 million workers in the late 1930s to 10 million workers in 1950 and 20 million workers in 1960. The expansion was spurred by a

number of factors including the desire of firms to encourage loyalty and reduce turnover, favorable tax treatment, and the 1949 Supreme Court decision to uphold a National Labor Relations Board (NLRB) ruling that pensions were appropriate issues of negotiation through collective bargaining (Schulz 1976). During this same period, Social Security coverage was extended to more workers, and Congress continued to raise benefits in response to changes in the cost of living, although real benefit levels remained relatively low (Derthick 1979).

#### **DECLINING RATES OF LABOR-FORCE PARTICIPATION**

Early research on retirement was centrally concerned with the question of voluntarism in the retirement transition, as well as with the financial, social, and psychological consequences of leaving the labor force. Even though the expansion of both government and employer-based pensions had improved the economic situation of older people in retirement, poverty rates among the elderly were still high. By 1960, 35.2 percent of persons aged 65 and older were below the poverty line, compared with 22.4 percent of the general population. Poverty was the norm for older white women and older African Americans (U.S. Bureau of the Census 1987).

During the 1950s, the Social Security Administration began studying the characteristics of newly entitled beneficiaries. Initial reports stated that early retirement occurred primarily because of poor health and difficulties finding and keeping jobs; the availability of Social Security retirement benefits played a secondary role (Wentworth 1945; Stecker 1955). Although these studies relied on beneficiary-reported reasons for retirement, a measurement strategy that was criticized because of its susceptibility to social desirability bias, the findings cannot be totally discounted. Retirement in the 1950s was not a financially attractive status for many older workers. Given that retirement income programs offered "fixed" benefits (benefits that remained nominally the same but, with inflation, declined in real terms), the financial security of middle- and working-class retirees was in jeopardy.

During the 1950s, retirement became more common. Insurance companies led the way in

developing “retirement preparation” programs, as researchers attempted to define strategies for “successful aging.” In an era of postwar prosperity, retirement came to be viewed as a “period of potential enjoyment and creative experience which accrues as a social reward for a life-time of labor” (Donahue et al. 1960, p. 361). Researchers investigating the effect of retirement on life satisfaction found that “retirement does not cause a sudden deterioration in psychological health as [had] been asserted by other writers” (Streib and Schneider 1971, p. 161). Rather than rejecting retirement, advocacy groups for the elderly lobbied for improved conditions of retirement, including more generous pension benefits. Mandatory retirement had not yet become an issue. Instead, the trend was in the direction of earlier retirement, that is, before the age of 65. In 1956 women and in 1962 men were allowed to retire at age 62 by accepting actuarially reduced Social Security benefits.

During the mid-1960s, in the context of Lyndon Johnson’s War on Poverty, the issue of old-age poverty was again addressed. In the Older Americans Act of 1965, Congress established a “new social contract for the aged” (Achenbaum 1983, p. 85) by specifying a series of objectives that, if met, would significantly improve the quality of life enjoyed by older people. Among these objectives was an “equal opportunity to the full and free enjoyment of . . . an adequate income in retirement in accordance with the American standard of living . . . [and] retirement in health, honor, and dignity” (U.S. Department of Health, Education, and Welfare 1976, p. 2).

Richard Nixon’s presidency inaugurated the era of modern retirement. Whereas previous amendments to the Social Security Act had brought more and more workers into the system, they had not significantly improved the level of retirement benefits (Munnell 1977). The presumption that Social Security benefits should serve as retirement income supplements rather than as the primary source of income had not been challenged. But the persistently high rates of old-age poverty lent credence to the charge that benefits were inadequate. During the decade following passage of the Older Americans Act, benefits were increased five times and indexed to changes in the consumer price index. Both the “real” level of benefits and

the replacement rate of benefits to previous earnings were improved. Enhanced retirement benefits allowed workers to maintain their standard of living across the retirement transition and helped redefine retirement as a legitimate nonwork status that average-income workers could afford to enter voluntarily. During the 1970s, employer-sponsored pensions were also being reorganized. The 1974 passage of the Employee Retirement Income Security Act (ERISA) regularized vesting plans and provided workers with some protection against benefit loss. Private sector initiatives aimed at inducing early retirement were also increasingly common (Barfield and Morgan 1969). Until the 1970s, workers choosing early retirement virtually always accepted reduced benefits. During the 1970s, however, employers began to offer early retirement incentive plans. Not only did these plans pay benefits at younger ages, but the present value of the benefits often exceeded that of normal retirement benefits.

The parallel changes in labor-force participation rates and in poverty rates among the elderly are noteworthy. In the latter part of this century, labor-force participation rates at older ages declined significantly. During the 1970s, rates for men aged 55–64 dropped from 83 percent (1970) to 75.6 percent (1975) to 72.1 percent (1980) (U.S. Department of Labor 1983). In addition, at the beginning of the decade, 24.6 percent of those aged 65 and older were living below the poverty line, twice the 12.1 percent that characterized the general population. By the end of the decade, the poverty rate among the elderly had dropped to 15.2 percent, compared to an overall poverty rate of 11.7.

As Figure 1 illustrates, the changes in labor-force participation rates (derived from the Current Population Survey) differ by both age and gender. Rates for men and women aged 65 and older appear to have stabilized and perhaps marginally increased from their lowest point. Rates for men aged 55–59 and 60–64 show overall declines with recent hints of a slight upturn. In contrast, rates for women in these age ranges have steadily increased during the last quarter-century, with rates for women in their late fifties surpassing rates for men in their early sixties during the last decade.

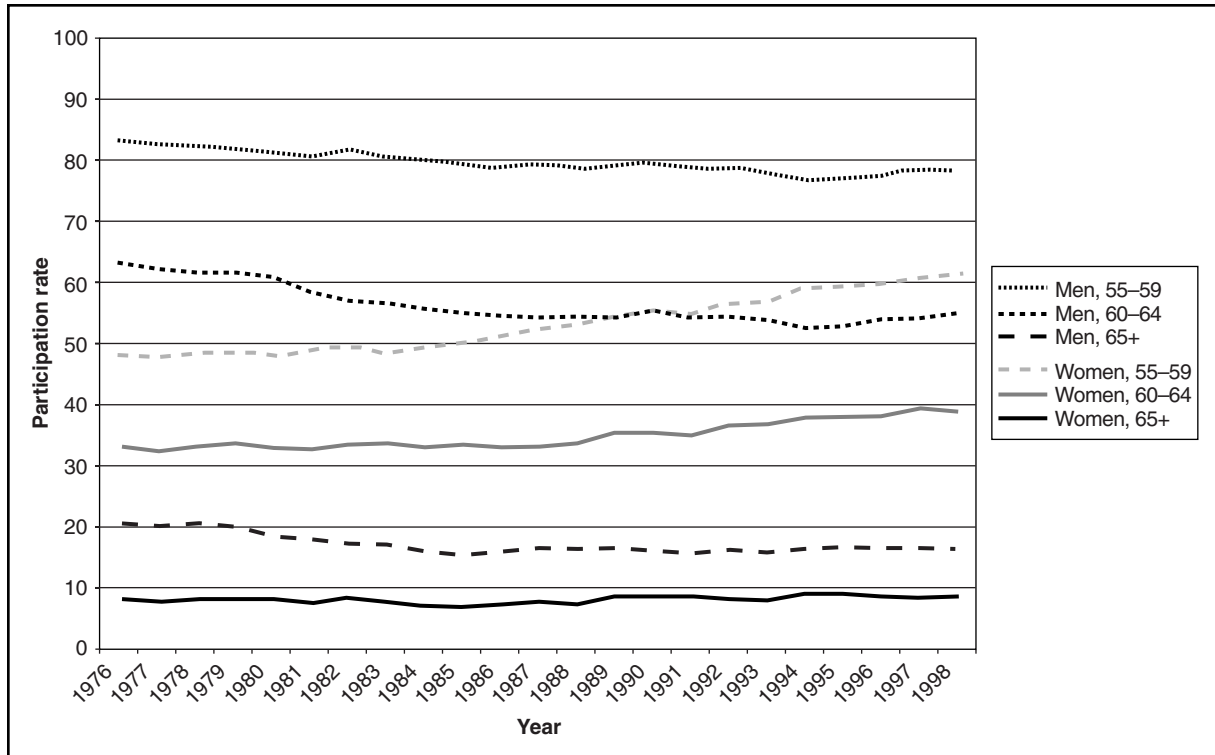


Figure 1. Labor-Force Participation of Men and Women in the United States, 1976–1998

### DETERMINANTS OF RETIREMENT

Most current research examining retirement is of one of two types: national longitudinal studies that address the behavior of workers across a wide variety of occupations and industries (e.g., Quinn and Burkhauser, 1994; Hayward et al. 1989) and studies of specific firms that assess the retirement behavior of workers who share substantial commonality in workplace features (e.g., Hardy et al. 1996; Stock and Wise 1990). The former design maximizes variation in both work context and individual characteristics, using statistical controls to assess the relative impact of individual and job characteristics on retirement transitions. Because of the sampling strategy, results from these studies can be generalized to the national population of older workers. The second design—the case study—limits observations to workers who share a particular work context and attempts to explain variation in workers' responses to a common decisional matrix.

In both these designs, individualized models of retirement have been dominant, although the

relative emphasis of “sociological” versus “economic” models may differ. Retirement models proposed by economists have emphasized the financial considerations involved in exchanging one income flow (e.g., wages and salary) for alternative sources of income (e.g., pensions and income from savings and investments). They have proposed measurement strategies based on the calculation of the present discounted values of the various income streams to attempt to disentangle relative effects. Sociological studies of retirement frequently focus either on the social psychological consequences of retirement or on the importance of occupational structure in shaping behavioral contingencies. Both opportunities and constraints are unequally distributed across workers. Whereas unionized manufacturing jobs may protect older workers through seniority systems, they also provide early retirement incentives through employer-based pension plans. Older workers are also vulnerable to plant closings and job dislocations that accompany mergers, downsizing, and cutbacks. This theoretical framework views retirement transitions as a career characteristic, with

late-career behaviors being at least partially contingent on earlier career opportunities. In addition, retirement behavior is embedded in more general macroeconomic conditions. Rates of unemployment, changes in both government and employer pension programs, and the age structure of both the population and the labor force are implicated. Combining insights from both disciplines leads to models in which the financial trade-off is captured by the relative effects of current earnings and the present value of pension benefits; human capital (and, indirectly, the probability of alternative employment) through education, health status, skill level, and age; characteristics of career through years of seniority, history of recent layoffs, and overtime work; and family situation through marital status and the presence of children in the home.

### MEASURING RETIREMENT

As the nature of retirement transitions changed, the question of measurement became more difficult. Armstrong's definition of "no longer working for pay" was being replaced by a variety of definitions oriented toward transitions out of career employment or full-time work, changes in major income sources (the receipt of Social Security or employer pension benefits), or changes in identity structures (e.g., through self-identification as a retiree). As definitions of retirement shifted toward receipt of pensions and the inclusion of part-time workers, the concerns of government turned toward the escalating cost of retirement and the advisability of delaying it. It became important to distinguish among exiting the labor force through "early" retirement, "regular" retirement, or disability, since these distinctions have implications for income replacement as well as labor-force reentry. Professionals, managers, and salespeople tend to delay retirement, whereas skilled and semiskilled blue-collar workers move more rapidly into retirement; clerical workers move more quickly into both retirement and disability statuses; and service workers experience relatively high rates of disability and death out of employment (Hayward et al. 1989). In addition, reentry into the labor force has become more common, with estimates of as many as one-third of retirees becoming reemployed, often within one to two years of their retirements. In short, the heterogeneity of what it means to be retired has increased considerably: it encompasses a broader age range;

it involves diversity of income sources; and it allows for some level of postretirement employment.

### ISSUES OF GENDER AND RACE

The development of retirement policy has been primarily oriented around the work careers of men, predominantly white men. The original Social Security program excluded industries in which women and blacks were concentrated. Although later amendments eventually covered these categories of workers, the benefit structure continued to reward long and continuous attachment to the labor force and to penalize workers for extended or frequent work interruptions. The temporal organization of women's lives relative to work and family, paid and unpaid labor, put women "off schedule" in accumulating claims to retirement income.

Spousal and survivors' benefits were designed to support couples and (primarily) widows during their later years. Research on women's retirement often focused on unmarried women and found that the determinants of retirement for women were similar to the determinants that had been identified for men. Although unmarried women and men differed in occupational locations, wages, health, and access to employer-sponsored pensions, these determinants appeared to sort unmarried women into retirees and workers in much the same way as they sorted men.

The pattern of women's labor-force participation has changed in recent decades, and more women—particularly more married women—are in the labor market. In fact, the trends in rates of labor-force participation for older women reflect both the increasing employment rates for successive cohorts of women and the tendency for more recent cohorts of older working women to retire at younger ages. The increase in dual-earner couples suggests that retirement decisions may be interdependent, with age differences, relative earnings, and the relative health of spouses figuring into joint decisions about careers, retirement, and postretirement employment.

Work and income disadvantages that are experienced at earlier stages of the life course cast a shadow on retirement transitions among minority group members. Lower earnings, lower job status, and discontinuity in labor-force attachment all

undermine the financial platform for retirement. Work histories characterized by frequent spells of unemployment, illness, or temporary disability are linked to lower average retirement benefits and lower rates of savings and asset accumulation. In addition, African Americans are less likely to be married than whites and therefore more likely to be limited to their individual earnings and retirement resources. African Americans are more likely to exit work through disability and also more likely to continue to work intermittently after retirement. Gibson (1987) argues that disability can be used as another pathway to retirement for older African Americans, one that offers financial advantages. Because work histories can appear sporadic in old age as well as youth, establishing the timing of retirement as an event also can be difficult (Gibson 1987).

### INTERNATIONAL COMPARISONS

Industrialization, economic development, demographic shifts, and politics are implicated in international comparisons of both the prevalence and the financing of retirement. Pampel (1985) reports that, among advanced industrial nations, a pattern of low labor-force participation among aged males is related to the level of industrialization and population aging. Cross-national comparisons of employment-to-population ratios demonstrate a continued decline in labor-force participation between 1970 and 1990 for men aged 55 and older. This decline is not, however, consistent across all age groups. Employment rates for men aged 55–59 have not shown the same proportional decrease as rates for men aged 60–64, or for those aged 65 and older. Compared to other Organization for Economic Cooperation and Development (OECD) countries (see Figure 2), Canada, Finland, Japan, Sweden, and the United States have relatively high rates of labor-force participation for men and women aged 65 and older.

Despite an overall downward trend in average age of retirement, nations continue to differ in the patterns of labor-force exits. Early retirement in European countries such as France, the Netherlands, and Germany remains the norm, with one-half to three-quarters of 60–64-year-old men out of the labor market (Guillemard and Rein 1993). Whereas early retirement in the United States was

primarily financed through early retirement incentive programs (ERIPs) offered by private firms, more severe problems of unemployment in Europe fueled early retirement through expanded eligibility for state programs. Among some countries of western Europe and North America, disability programs also can operate as pseudo-retirement programs that allow workers to exit prior to normal retirement age. In contrast to the pattern of western Europe, Canada, and the United States, labor-force participation rates in Japan for men in all three age groups have been more resilient. Rates in 1990 remained relatively high, with more than one-third of men aged 65 and older participating in the labor force (Quinn and Burkhauser 1994). Cross-national comparisons of women's retirement patterns are more complicated. Whereas some countries show little change since 1970 (e.g., Canada, Australia, Italy, Japan, and the United States), others show patterns of labor-force withdrawal that parallel the trends for older men (e.g., Finland, France, West Germany, Spain, and the United Kingdom).

When comparing rates of labor-force participation, it is important to take into account national differences in census procedures, the definition of the labor force, and the kinds of activities that constitute "work." Because countries differ in the pathways workers take to retirement, using pension receipt as an indicator is also flawed. In Germany, for example, disability benefits and intermediate unemployment benefits also provide access to early retirement. Comparisons based on rates of labor-force participation confound country differences in full-time versus part-time employment, complete versus partial retirement, and unemployment and employment. In addition, cross-sectional figures do not allow a comparison of rates of withdrawal or reentry that would allow us to distinguish relatively stable from volatile labor markets.

Within each nation, policy development and social dynamics exert an important influence on retirement behavior. Recent debate in the United States and other countries has centered on the financial burdens of supporting a growing population of retirees and the desirability of reversing the trend toward early retirement. In 1983 the United States amended the Social Security Act to legislate a gradual increase in the age of full entitlement from 65 to 66 by 2009 and to 67 by 2027. Germany



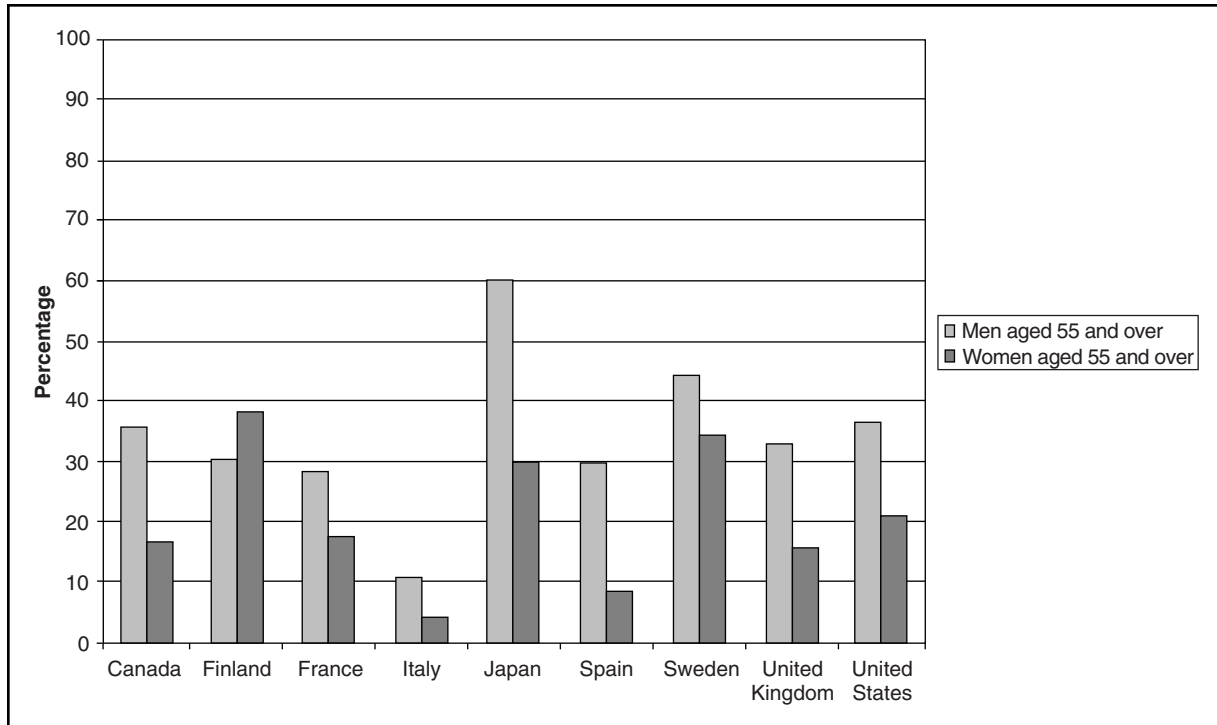


Figure 2. *The Employment-to-Population Ratio for Men and Women Aged 55 and Over in OECD Countries, 1990*

(1992 Pension Reform Act) and Italy have made similar policy changes. Japan tried to raise the retirement age from 60 to 65 but failed in their initial attempt. In 1990 Sweden succeeded in blocking the early retirement pathway through the disability fund, but in 1992 failed to abolish the partial retirement pension system. Even in France, where “old age is seen as a time of life when work is illegitimate (Guillemard 1983, p. 88) and where the legal retirement age was lowered to 60 in 1982, it is likely that the retirement age will be raised. This type of political reaction to demographic aging is one way of trying to shift the cost of early retirement from government programs to firms and individuals.

Other workplace policies linked to demographic aging, such as flexible retirement systems, are also being considered in countries like Germany, France, and Great Britain. Countries with lower rates of early exit, such as Japan and Sweden, provide an alternative approach to labor-market withdrawal which may alleviate the pressures of demographic aging. These countries structure retirement as a gradual process involving lowering

wages, reducing hours, and reassigning workers. To some extent this model is already in place in the United States through “bridge jobs,” in Sweden through partial retirement pensions, and in Japan through wage reduction arrangements.

#### DIRECTIONS FOR RESEARCH

Sociologists are also interested in the social context in which retirement decisions are made and retirement policies are developed. Family contexts, work contexts, economic contexts, and historical contexts all provide important frames of reference in which these behaviors are negotiated. To date, retirement research has refined both the measurement of concepts and the complexity of the behavioral models. Many of these refinements have involved the economic dimensions of retirement. The financial trade-off between pensions and wages, the changes in accumulated pension wealth, and the age-earnings profiles of different occupations have been captured in current models. What models of retirement continue to lack, however, is sensitivity to the social frames of reference (e.g., the shop floor, the office, the firm, the

family, and the community). Studies that have addressed some of these issues (e.g., Hardy and Hazelrigg 1999) suggest, for example, that firm level features may be implicated in both the rate and the determinants of early retirement.

Retirement behavior may be primarily motivated by financial considerations. But given a threshold of financial security, perhaps the unfolding of the retirement process also involves the culture of the workplace, family dynamics, and societal values. Our first task is to theorize the social aspects of these decisions so that we can develop hypotheses. To test these hypotheses, a different data collection strategy is required—one that samples a sufficient number of observations of individual workers within sufficient numbers of organizational or, more generally, cultural contexts that can themselves be measured in terms of their salient characteristics. What has become clear is that retirement decisions are shaped by individual preferences, but that these preferences are shaped by the opportunities and constraints that workers encounter.

## REFERENCES

- Achenbaum, W. Andrew 1983 *Shades of Gray: Old Age, American Values, and Federal Policies since 1920*. Boston: Little, Brown.
- Atchley, Robert 1982 "Retirement: Leaving the World of Work." *Annals of the Academy of Political and Social Science* 464:120–131.
- Barfield, Richard E., and James N. Morgan 1969 *Early Retirement: The Decision and the Experience*. Ann Arbor: University of Michigan Press.
- Derthick, Martha 1979 *Policymaking for Social Security*. Washington, D.C.: Brookings Institution.
- Donahue, Wilma, Harold Orbach, and Otto Pollak 1960 "Retirement: The Emerging Social Pattern." In Clark Tibbitts, ed., *Handbook of Social Gerontology*. Chicago: University of Chicago Press.
- Fischer, David Hackett 1977 *Growing Old in America*. New York: Oxford University Press.
- Gibson, Rose 1987 "Reconceptualizing Retirement for Black Americans." *Gerontologist* 27(6):691–698.
- Graebner, William 1980 *A History of Retirement: The Meaning and Function of an American Institution 1885–1978*. New Haven, Conn.: Yale University Press.
- Guillemard, Anne-Marie 1983 "The Making of Old Age Policy in France: Points of Debate, Issues at Stake, Underlying Social Relations." In Anne-Marie Guillemard, ed., *Old Age and the Welfare State*. Beverly Hills, Calif.: Sage.
- , and Martin Rein 1993 "Comparative Patterns of Retirement: Retirement Trends in Developed Societies." *Annual Review of Sociology* 19:469–503.
- Hardy, Melissa A., and Lawrence Hazelrigg 1999 "A Multilevel Analysis of Early Retirement Decisions among Auto Workers in Plants with Different Futures." *Research on Aging* 21(2):275–303.
- , and Jill Quadagno 1996 *Ending a Career in the Auto Industry*. New York: Plenum.
- Hayward, Mark D., Melissa A. Hardy, and William Grady 1989 "Labor Force Withdrawal Patterns among Older Men in the United States." *Social Science Quarterly* 70(2):425–448.
- Munnell, Alicia H. 1977 *The Future of Social Security*. Washington, D.C.: Brookings Institution.
- Olsen, Laura Katz 1982 *The Political Economy of Aging: The State, Private Power, and Social Welfare*. New York: Columbia University Press.
- Pampel, Fred 1985 "Determinants of Labor Force Participation Rates of Aged Males in Developed and Developing Nations, 1965–1975." In Zena Blau, ed., *Current Perspectives on Aging and the Life Cycle*. Greenwich, Conn.: JAI Press.
- Plakans, Andrejs 1989 "Stepping Down in Former Times: A Comparative Assessment of Retirement in Traditional Europe." In David Kertzer and K. Warner Schaie, eds., *Age Structuring in Comparative Perspective*. Hillsdale, N.J.: Erlbaum.
- Quinn, Joseph F., and Richard V. Burkhauser 1994 "Retirement and Labor Force Behavior of the Elderly." In Linda Martin and Samuel H. Preston, eds., *Demography of Aging*. Washington, D.C.: National Academy Press.
- Schulz, James H. 1976 *The Economics of Aging*. Belmont, Calif.: Wadsworth.
- Sklar, Martin J. 1988 *The Corporate Reconstruction of American Capitalism, 1890–1916*. New York: Cambridge University Press.
- Sorensen, Aage B. 1989 "Old Age, Retirement, and Inheritance." In David Kertzer and K. Warner Schaie, eds., *Age Structuring in Comparative Perspective*. Hillsdale, N.J.: Erlbaum.
- Stecker, Margaret 1955 "Why Do Beneficiaries Retire? Who among Them Return to Work?" *Social Security Bulletin* 18:3.
- Stock, James H., and David A. Wise. 1990. "The Pension Inducement to Retire." In David A. Wise, ed., *Issues*

*in the Economics of Aging*. Chicago: University of Chicago Press.

Streib, Gordon F., and Clement Schneider 1971 *Retirement in American Society*. Ithaca, N.Y.: Cornell University Press.

U.S. Bureau of the Census 1987 *Statistical Abstract of the United States 1986*. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Health, Education, and Welfare 1976 *Older Americans' Act of 1965, as Amended and Related Acts, Bicentennial Compilation, March 1976*. Washington, D.C.: Office of Human Development, Administration on Aging.

U.S. Department of Labor, Bureau of Labor Statistics 1983 *Current Population Reports*, ser. P-60, no. 148. Washington, D.C.: U.S. Government Printing Office.

Wentworth, Edna C. 1945 "Why Beneficiaries Retire." *Social Security Bulletin* 8:16-20.

MELISSA A. HARDY  
KIM SHUEY

## REVOLUTIONS

*Revolutions* are rapid, fundamental transformations of a society's socioeconomic and political structures (Huntington 1968). Social revolutions differ from other forms of social transformation, such as rebellions, coups d'état, and political revolutions. *Rebellions* involve the revolt of society's subordinate classes—peasants, artisans, workers—but do not produce enduring structural changes. *Coups d'état* forcibly replace the leadership of states but do not fundamentally alter state structures. *Political revolutions* transform state structures but leave social structures largely intact. What is distinctive to social revolutions is that basic changes in social structures and political structures occur in a mutually reinforcing fashion (Skocpol 1979). A social revolution is more than a change in the state. It is a change in the state of an entire society.

Recent sociological work on revolutions recognizes their importance in the making of the modern world order and the opportunities revolutions offer for building theories of social and political change. These opportunities were most emphatically embraced by Marx, who placed the study and the making of revolution at the center of his lifework. Virtually all theories of revolution

since Marx share his concern with three separate yet interrelated phenomena: (1) the social conditions that lead to revolution or its absence, (2) the character of participation in revolutions, and (3) the outcomes of revolutions (see Tucker 1978). This review examines Marx's theories in light of significant contemporary analyses of revolution, in order to evaluate how well his theories have stood the test of time and to consider how much of his legacy may endure in future sociological work.

First, Marx understood modern revolutions to be by-products of economic advance. Revolutionary situations emerged when contradictions between the forces of production (how the means of existence are produced) and the relations of production (how labor's product is distributed) within an existing mode of production reached their limits. For Marx, the coming of the 1789 French revolution lay in the irresolvable contradiction between feudal restrictions on land, labor, and credit and emerging capitalist arrangements advanced by an ascending bourgeoisie. Revolutions brought the resolution of these contradictions by serving as bridges between successive modes of production, enabling the ascent of capitalism over feudalism and later the replacement of capitalism by socialism.

Second, Marx held that revolutions were accomplished through class struggle. In revolutionary situations, conflict intensified between the existing dominant class and the economically ascendant class. Under feudalism, class conflict pitted the aristocracy against the ascendant bourgeoisie. Under capitalism, the differing situations of segments of society determined their revolutionary tendencies. Some classes, such as the petite bourgeoisie, would become stakeholders in capitalism and allies of the dominant bourgeoisie. Others, such as the peasantry, that did not fully participate in wage labor and lacked solidarity, would stay on the sidelines. The industrial proletariat would be the midwife of socialist revolution, for wage labor's concentration in cities would generate solidarity and collective consciousness of the proletariat's exploitation by the bourgeoisie. Class consciousness was a necessary (though not a sufficient) condition for revolution.

Third, Marx believed that revolutions so thoroughly transformed class relations that they put in place new conditions enabling further economic

advance. Revolutions were locomotives of history that brought in train new structures of state administration, property holding, and political ideology. Reinforcing the fundamental changes in class relations, these transformations culminated the transition from one mode of production to another.

In sum, Marx's theory identified the conditions that spawn revolutionary situations, the classes that would make revolutions, and the outcomes of revolutions. How well has Marx's analysis served later generations of scholars and revolutionaries? Many of the sociopolitical transformations since his death were unanticipated by Marx. In light of these events, contemporary sociologists have reconsidered his thinking.

Consider first the social conditions making for revolution or its absence. Social revolutions are rare. Modern capitalist societies have experienced severe economic crises that intensified class conflict and gave the appearance of revolutionary situations, but they have skirted actual revolution. Modern capitalist economies have great staying power, and reform rather than revolution is the rule in advanced nations. Indeed, the great revolutions of France, Russia, and China, and those in Third World societies such as Cuba, Vietnam, and Nicaragua, have occurred in predominantly agrarian economies where capitalist relations of production were only moderately developed. The 1917 Russian Revolution stands as proof of Lenin's claim that revolution was possible in Russia despite its failure to develop a fully capitalist economy in the manner of the western European states Marx saw as the likely candidates for socialist revolution.

Rather than growing from contradictions between the forces and the relations of production, revolutionary situations arise in political crises occasioned by international competitive pressures. States with relatively backward economies are vulnerable to military defeats, as occurred in Russia in 1917, and to financial crises like that of 1789 France after a century of costly struggle with Britain and Continental Powers. Nation-states that are disadvantaged within the international states system are most at risk to externally induced crises and to the possibility of social revolution.

A state's vulnerability to crisis, however, depends fundamentally on its autonomy, that is,

the extent of its dependence on elites, whether nobles, landlords, or religious authorities. State managers are forever caught in a vice between their obligation to increase revenues to meet international competitive challenges and the resistance of angry elites to resource extraction. States that are weakly bureaucratized, where elites control high offices and key military posts, cannot act autonomously. When powerful elites paralyze the state's resource accumulation, severe crises occur, as in the English, French, and Chinese revolutions (Goldstone 1986).

However, externally induced crises may initiate an "elite revolution," as occurred in Japan's 1868 Meiji restoration and Atatürk's 1919 revolution in Turkey. In such regimes, a bureaucratic elite of civil and military officials emerged that lacked large landholdings or ties to merchants and landlords. In the face of Western military threats to national sovereignty and, consequently, to their own power and status, these elites seized control of the state apparatus. With the aim of resolving economic and military difficulties, they transformed existing sociopolitical structures through land reform, leveling of status distinctions, and rapid industrialization (Trimberger 1978). These transformative episodes, sometimes called "revolutions from above," are distinguished by the absence of popular revolts "from below."

Neopatrimonial regimes are highly vulnerable to revolutionary overthrow (Eisenstadt 1978). Examples include pre-1911 Mexico, pre-1959 Cuba, and pre-1979 Nicaragua. Centered on the personal decisions of a dictator, each of these regimes operated through extensive patronage networks rather than through a bureaucratized civil service and a professionalized military. Their exclusionary politics made reform nearly impossible and, in the event of withdrawal of support by stronger states, invited challenges that might overwhelm corrupt armed forces.

In contrast, revolution is unlikely in open, participatory regimes typical of modern capitalist democracies. By enfranchising new groups, these systems incorporate potential challengers. By redistributing wealth and opportunity—or appearing to do so through open markets and meritocracy—they are able to mute class antagonisms.

In sum, contradictions between the state apparatus and a society's dominant classes have been

crucial to the onset of revolutionary situations. State bureaucrats are caught in the cross pressures between meeting the challenges of the international states system and yielding to the competing claims of elites. Consequently, state structures differ in their vulnerability to political crises.

Consider next the character of participation in revolutions. Critics of Marx's voluntarist theory assert that no successful social revolution has been made by a self-consciously revolutionary movement. They allow that revolutionaries do guide the course of revolutions, but assert that they do not create revolutionary situations which emerge from externally induced political crises. These critics also offer a different reading of the roles of social classes in revolution. Urban workers, favored by Marx's theory, played important parts in revolutions. However, their numbers were small and their protests lacked the impact of uprisings by peasants who composed the vast bulk of producers in agrarian societies. Indeed, peasant revolts provided the "dynamite" to bring down old regimes when political crises immobilized armies and upset food supplies and distribution (Moore 1966). Peasant revolts against the landed upper classes made it impossible to continue existing agrarian class relations and thereby reduced the prospects for liberal reforms or counterrevolution. It was the conjunction of political crisis and peasant insurrection that brought about the fundamental transformations identified with social revolutions in France, Russia, and China (Skocpol 1979).

But peasants do not act alone. A key difference between Old Regime revolutions and revolutions in the Third World is the importance of coalitions in the latter (Tilly 1978). Peasants in France and Russia lived in solitary and relatively autonomous village communities that afforded them the solidarity and tactical space for revolt (Wolf 1969). In the twentieth century, professional revolutionaries provided leadership and ideologies that cemented dispersed local groups with disparate interests into potent national movements. The success of their efforts depended in part on the breadth of the coalition they were able to realize among peasants, landless and migrant laborers, rural artisans, and sometimes landlords (Goodwin and Skocpol 1989).

Add to that the importance of urban groups which played crucial parts in making Third World

revolutions. In Cuba and Nicaragua, students, professionals, clerics, and merchants joined workers and peasants in coalitions that toppled dictatorial regimes. Similarly, the 1979 overthrow of the Shah of Iran resulted from the mobilization of a broad coalition of merchants and workers, students and professionals, that met little resistance from powerful military forces (Farhi 1990).

Finally, revolutionary leaderships have not come from the ranks of an ascendant bourgeoisie or a risen proletariat. Rather, marginal political elites were most likely to consolidate power, for they were both skilled in the running of state organizations and tied by identity and livelihood to the aggrandizement of national welfare and prestige. Their role is clearest in revolutions from above but is no less prominent in social revolutions.

Consider, last, the outcomes of revolutions. For Marx, revolutions were bridges between successive modes of production. Bourgeois revolutions marked the transition from feudalism to capitalism; socialist revolutions opened the way for the transition from capitalism to communism, history's final stage. Through the dictatorship of the proletariat and the abolition of private property, socialist revolution would bring the end of class struggle and the disappearance of state power.

Revolutions did transform class structures, economic arrangements, and political institutions, however, in all successful social revolutions, the transformations were accomplished by new state structures. The state did *not* wither away. Moreover, the new states were more centralized and bureaucratized than their predecessors. For example, liberal parliamentary regimes appeared in the early phases of the 1789 French and 1917 Russian revolutions. In the face of threats to sovereignty from abroad and counterrevolutionary threats at home, these regimes gave way to centralized governments which rationalized the machinery of state for national defense and internal control. Liberal parliamentary regimes were similarly short-lived during revolutionary episodes in post-World War II Europe, Asia, Africa, and Latin America.

Indeed, what changed most in revolutions was the mode of social control of the lower strata as regimes centralized political power. All social revolutions ended with the consolidation of new mass-mobilizing state organizations through which peasants and urban workers were for the first time

directly incorporated into national economies and polities. While they aimed for independence from colonial powers, national liberation movements in Asia and Africa shared with these newly formed revolutionary states the project of state-centered mass mobilization.

Last, and vastly more important than the ways that individual revolutions changed sociopolitical institutions where they occurred, is the collective legacy of four centuries of revolution for the modern world. The legacy is clear in ways that revolutions transformed the workings of the world order and altered our sense of human prospects and limits. Through new political institutions, revolutions created our vocabulary of citizenship and its opposite, the machinery of tyranny. Through new structures of labor and investment, revolutions boosted agricultural and industrial output. By breaking old alliances and creating new ones, they altered the international states system and the shape of world markets. With wars and depressions, revolutions rightfully became the major markers of world history. Past revolutions are a benchmark from which we reckon humankind's advance or descent. In all these ways, revolutions helped shape the structures and consciousness we know as modernity.

For a century, sociologists who study revolutions have pursued a project of revision. Informed by the study of recent sociopolitical transformations and by new interpretations of past revolutions, their work questions and qualifies much of Marx's analysis. It considers national revolutions in the context of the world economy and the international states system. It places the relations between states and social classes in a new light, and it examines how social transformations in the recent or distant past weigh on the course of revolutionary events. On balance, it largely affirms and continues the main thrust of Marx's perspective, notably his focus on actors embedded in concrete organizational settings within historically specific circumstances. This perspective—rather than a focus on personality, collective mentality, or system dysfunction—distinguishes the sociology of revolution. It owes much to Marx's legacy.

What future sociologists will take from Marx's legacy is an open question. Are revolutions in the coming century likely to differ in their causes, dynamics, or outcomes from the revolutions of

centuries past? Will revolutions become unlikely as global prosperity shrinks the now yawning gap between rich and poor? Is it possible that new state structures, either more responsive or more repressive than those we now know, will bring revolutions to an end? These are questions for the next generation of theorists.

Their search for answers will surely cover old ground. There is no "Y2K problem" for revolutionary theory. Changing the calendar does not change the world. The familiar masterplots that made modern world history—notably, the globalization of markets, communications, and culture, the international states system, and persistent inequalities in power and resources—suggest the shape of revolutions to come.

Indeed, for many sociologists, the past is prologue. For example, in an analysis of revolution and rebellion in Europe, China, and the Middle East from 1500 to 1850, Goldstone (1991) links population growth and state breakdown. He identifies a process in which population growth overwhelms the administrative capacities of agrarian states, bringing inflation and fiscal crisis, intralite conflicts, popular unrest, and delegitimation of traditional regimes and their policies. Starting in roughly 1850, the underlying conditions changed when new forms of investment and infrastructure that accompanied urbanization and industrialization enabled modern states to manage population growth and prices. In much of the contemporary Third World, however, urbanization and industrialization did not strengthen the hand of state managers, but had the opposite effect. There, population growth and price pressures—often the unintended consequences of failed development schemes—continued to court state crises.

Nevertheless, Goldstone and other analysts assert revolutions will be rare. This is so because revolutions are not the result of a single cause, such as population pressure, but arise only when several causal elements, such as state crisis, elite withdrawal, and mass mobilization, converge in a rare revolutionary conjuncture. The worker masses of Sergei Eisenstein's films and massed peasants of Hollywood spectacles give a misleading impression of revolutionary causation. There have been many popular uprisings—some of them large and quite violent—which did not bring revolutionary transformations.

Lachmann (1997) calls attention to the pivotal role of elites. He develops an explanation for revolutionary outbreak and success in terms of varying combinations of elite unity or conflict and mass participation or quiescence. Lachmann finds that single, unified elites are immune to mass revolutionary challenges (as happened in Hungary in 1956 and Czechoslovakia in 1968). Conversely, mass mobilization succeeds at moments of heightened elite conflict (as in 1789 France). These conditions are self-reinforcing. By creating opportunities and potential alliances, elite conflict encourages nonelite mobilization (sometimes accidentally, sometimes intentionally), which, in turn, can alter the system of political domination and economic exploitation through which elites rule and live.

The study of mass-elite relations signals a turn toward understanding the importance of agency and ideology in making revolutions. The turn follows a generation of structural theories which spotlighted the crises through which revolutionary situations come, but left us unenlightened as to how revolutionary outcomes are realized in the actions of millions of men and women. For example, Markoff (1996) shows how in 1789 France, the abolition of feudalism was not foreordained by the state's fiscal crisis, but a contingent outcome of the interplay of waves of rural insurrection, Parisian mobs, and legislative concessions. Through their dialogue with the "people," revolutionary legislators created a form of discourse that defined the revolution as a watershed moment between the old order and the new. Markoff shows that words, as well as deeds make revolutions. While the French revolution was more than talk, talk became the frame in which action was interpreted and given form.

Revolutionary outcomes, though, depend on more than the unfolding of actions and the shaping of ideologies by the makers of revolution. History shows that revolutionary outcomes are constrained by global circumstances. Contrasting the outcomes of the Mexican (1910–1917) and Bolivian (1952) revolutions, Eckstein (1976) shows the impact of U.S. foreign investment and military assistance on class structure, party formation, and national well-being. While both modernized their economies, Mexico had greater economic diversification, productivity growth, and political stability than Bolivia. Their revolutions did not alter their

standing relative to stronger states, leaving Bolivia with its weaker postrevolutionary state (and important tin mines) open to U.S. foreign aid and military assistance. Subsidized and schooled by the United States, the armed forces played the decisive role in the 1964 overthrow of the Bolivian regime that the United States had supported in 1952.

Revolutions will not disappear as long as nation-states remain the dominant form of world political organization. While the globalization of markets, communications, and culture seems to augur a new world, we live with the legacy of the capitalist world order. Ours is still a world of states, albeit one increasingly structured by regional alliances, multinational corporations, and international bodies like the United Nations. What Marx recognized in the late nineteenth century will likely hold in the century ahead: that persistent inequalities in power and resources between masses and elites and among the world's nation-states will shape revolutionary prospects and processes.

(SEE ALSO: *Marxist Sociology*)

REFERENCES

- Eckstein, Susan 1976 *The Impact of Revolution: A Comparative Analysis of Mexico and Bolivia*. London: Sage.
- Eisenstadt, S. N. 1978 *Revolution and the Transformation of Societies*. New York: Free Press.
- Farhi, Farideh 1990 *States and Urban-Based Revolutions: Iran and Nicaragua*. Urbana: University of Illinois Press.
- Goldstone, Jack A. 1986 "The Comparative and Historical Study of Revolutions." In J. A. Goldstone, ed., *Revolutions*. San Diego; Calif.: Harcourt Brace Jovanovich.
- 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.
- Goodwin, Jeff, and Theda Skocpol 1989 "Explaining Revolutions in the Contemporary Third World." *Politics and Society* 17:489–509.
- Huntington, Samuel P. 1968 *Political Order in Changing Societies*. New Haven, Conn.: Yale University Press.
- Lachmann, Richard 1997 "Agents of Revolution: Elite Conflicts and Mass Mobilization from the Medici to Yeltsin." In John Foran, ed., *Theorizing Revolutions*. New York: Routledge.
- Markoff, John 1996 *The Abolition of Feudalism*. University Park: The Pennsylvania State University Press.

- Moore, Barrington, Jr. 1966 *Social Origins of Dictatorship and Democracy*. Boston: Beacon.
- Skocpol, Theda 1979 *States and Social Revolutions*. New York: Cambridge University Press.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.
- Trimberger, Ellen Kay 1978 *Revolution from Above: Military Bureaucrats and Development in Japan, Turkey, Egypt, and Peru*. New Brunswick; N.J.: Transaction.
- Tucker, Robert (ed.) 1978 *The Marx-Engels Reader*. New York: Norton.
- Wolf, Eric R. 1969 *Peasant Wars of the Twentieth Century*. New York: Harper and Row.

ROBERT C. LIEBMAN

## ROLE CONFLICT

See Role Theory; Role Theory: Foundations, Extensions and Applications; Stress.

## ROLE MODELS

See Gender; Role Theory; Role Theory: Foundations, Extensions and Applications; Socialization; Symbolic Interaction Theory.

## ROLE THEORY

Role theory concerns the tendency for human behaviors to form characteristic patterns that may be predicted if one knows the social context in which those behaviors appear. It explains those behavior patterns, (or *roles*) by assuming that persons within a context appear as members of recognized social identities (or *positions*) and that they and others hold ideas (*expectations*) about behaviors in that setting. Its vocabulary and concerns are popular among both social scientists and practitioners, and role concepts have generated both theory and a good deal of research. Nevertheless, conflicts have arisen about the use of role terms and the focus of role theory, and different versions of the theory have appeared among groups of authors who seem to be unaware of alternative versions. Role theory has been weakened by association with controversial theories in sociology, as well.

## HISTORY, DIFFERENTIATION, AND CONFUSION

Role theory arose when social scientists took seriously the insight that social life could be compared with the theater, in which actors played predictable "rôles." This insight was pursued independently by three major contributors in the early 1930s with somewhat different agendas. For Ralph Linton (an anthropologist), role theory was a means for analyzing social systems, and roles were conceived as "the dynamic aspects" of societally recognized social positions (or "statuses"). In contrast, George Herbert Mead (a social philosopher) viewed roles as the coping strategies that individuals evolve as they interact with other persons, and spoke of the need for understanding others' perspectives ("role taking") as a requisite for effective social interaction. And Jacob Moreno (a psychologist) saw roles as the habitual, sometimes harmful, tactics that are adopted by persons within primary relationships, and argued that imitative behavior ("role playing") was a useful strategy for learning new roles.

Additional insights for role theory were generated by other early authors, particularly Muzafer Sherif's studies of the effects of social norms; Talcott Parsons's functionalist theory, which stressed the importance of norms, consensus, sanctioning, and socialization; Robert Merton's analyses of role structures and processes; the works of Neal Gross, Robert Kahn, and their colleagues, which discussed role conflict and applied role concepts to organizations; Everett Hughes's papers on occupational roles; Theodore Newcomb's text for social psychology, which made extensive use of role concepts; and (in Europe) the seminal monographs of Michael Banton, Anne-Marie Rocheblave, and Ragnar Rommetveit, as well as Ralf Dahrendorf's essay "Homo Sociologicus."

The contrasting insights of these early contributors affected many subsequent writers, and various traditions of role theory have since appeared. Unfortunately, advocates for (or critics of) these differing traditions often write as if they are unaware of other versions. In addition, advocates may propose inconsistent uses for terms, or contrasting definitions for concepts, that are basic in role theory. To illustrate, for some authors the term "role" refers only to the concept of social position, for others it designates the behaviors



characteristic of social position members, and for still others it denotes shared expectations held for the behaviors of position members. Such inconsistent uses pose problems for the unwary reader.

Also, role theorists may disagree about substantive issues. For example, some authors use role concepts to describe the social system, whereas others apply it to the conduct of the individual. Again, some writers assume that roles are always tied to functions, whereas others conceive roles as behaviors: that conform to expectations, that are directed towards other in the system, that are volitional, that validate the actor's status, or that project a self-image. Such differences in stance have reflected both accidents of intellectual history and the fact that role theorists have wrestled with differing social system forms.

Despite these differences, role theorists tend to share a basic vocabulary, an interest in the fact that human behavior is contextually differentiated and is associated with the social position of the actor, and the assumption that behavior is generated (in part) by expectations that are held by the actor and others. This means that much of role theory presumes a thoughtful, phenomenally aware participant, and role researchers tend to adopt methods that call for the observing of roles and for asking respondents to report about their own or others' expectations. Moreover, it also means that role theory may be contrasted with alternative theoretical positions that give stronger emphasis to unconscious motives or behavior-inducing forces of which the actor may be unaware (such as mechanisms that are not obvious but that serve to maintain structured inequalities of power, wealth, or status).

### FUNCTIONALIST ROLE THEORY

One early perspective in role theory reflected functionalism. Functionalist thought arose from the contributions of Talcott Parsons and was, at one time, the dominant orientation in American sociology. This theory made use of role concepts, and some authors continue, today, to write as if role theory was or is largely an attempt to formalize functionalism.

Functionalist theory was concerned with the problem of explaining social order. Stable but differentiated behaviors were thought to persist

within social systems because they accomplished functions and because actors in those systems shared expectations for behaviors. Such consensual expectations (or "roles") constituted norms for conduct, and actor conformity to norms was induced either because others in the system imposed sanctions on the actor or because the actor internalized them. In addition, those in the system were thought to be aware of the norms they held and could be counted on to teach them to (i.e., to socialize) neophytes as the latter entered the system.

Functionalist thought has been under attack since the 1950s, and many of its basic assumptions have been challenged. Critics have pointed out that persisting behaviors may or may not be functional for social systems, that norms for conduct are often in conflict, that actor conformity need not be generated by norms alone but can also reflect other modes of thought (such as beliefs or preferences), that norms might or might not be supported by explicit sanctions, that norms internalized by the actor may be at odds with those supported by external forces, and that processes of socialization are problematic. Above all, critics have noted that social systems are not the static entities that functionalist thought portrayed, and that human conduct often responds to power and conflicts of interest in ways that were ignored by functionalists. As a result of these attacks, interest in functionalist role theory has declined, although it is still possible to find writers who advocate (e.g., Bates and Harvey 1975) or denounce (Connell 1979) role theory as if it were merely a gloss for functionalism.

### ROLE CONFLICT AND ORGANIZATIONAL ANALYSIS

Interest in organizational role theory began with the works of Neal Gross, Robert Kahn, and their associates, which questioned the assumption that consensual norms were required for social stability. Instead, these writers suggested that formal organizations were often characterized by *role conflict* (i.e., opposing norms that were held for actors by powerful others), that such conflicts posed problems for both the actors and the organizations in which they appeared, and that strategies for coping with or "resolving" role conflict could

be studied. These insights stimulated both texts that applied role concepts to organizational analysis and many studies of role conflict and role conflict resolution in organizational contexts (see, for example, van de Vliert 1979; Van Sell et al. 1981; Fisher and Gitelson 1983).

In addition, the concept of role conflict has proven attractive to scholars who wanted to conceptualize or study problems that are faced by disempowered persons, particularly married women who must cope with the opposing demands of the workplace, home maintenance, and support for their husbands (Stryker and Macke 1978; Lopata 1980; Skinner 1980). Unfortunately (for the argument), evidence suggests that role conflicts are not always shunned by disempowered persons (see Sales et al. 1980) and that “resolving” those conflicts does not necessarily lead to empowerment.

Despite these problems, research on role conflict within the organization continues actively, and some proponents of the organizational perspective have recently turned their attention to the events of role transition—that is, to phenomena associated with entry into or departure from a role (see Allen and van de Vliert 1984; Ebaugh 1988).

### THE STRUCTURAL PERSPECTIVE

Another use of role concepts has appeared among structuralists and network theorists. This third perspective reflects the early contributions of anthropologists such as S. F. Nadel and Michael Banton, sociologists such as Marion Levy, and social psychologists ranging from Dorwin Cartwright and Frank Harary to Oscar Oeser. As a rule, structuralists concern themselves with the logical implications of ways for organizing social systems (conceived as social positions and roles) and eschew any discussion of norms or other expectation concepts.

To date, much of the work in structural role theory has been expressed in formal, mathematical terms (see Burt 1982; Winship and Mandel 1983). This means that it has had greater appeal for scholars who are mathematically trained. It also constitutes one form of network analysis (although other network perspectives have appeared that do not use role concepts).

### ROLE THEORY AMONG SYMBOLIC INTERACTIONISTS

Interest in role theory has also appeared among symbolic interactionists who were influenced not only by George Herbert Mead but also by Everett Hughes, Irving Goffman, and other influential figures. In general, symbolic interactionists think of a role as a line of action that is pursued by the individual within a given context. Roles are affected by various forces, including preexisting norms applying to the social position of the actor, beliefs and attitudes that the actor holds, the actor’s conception and portrayal of self, and the “definition of the situation” that evolves as the actor and others interact. Roles need not have common elements, but they are likely to become quite similar among actors who face common problems in similar circumstances.

These concepts have been applied by symbolic interactionists to a host of interesting concerns (see, for example, Scheibe 1979; Gordon and Gordon 1982; Ickes and Knowles 1982; Stryker and Serpe 1982; Zurcher 1983; Hare 1985), and a continuing and useful contribution has flowed from Ralph Turner’s interest in the internal dynamics of roles and the fact that roles tend to evolve over time (1979, 1990).

Unfortunately, some persons within this perspective have also been guilty of tunnel vision and have produced reviews in which role theory is portrayed largely as an extension of symbolic interactionist thought (see Heiss 1981; Stryker and Statham 1985). In addition, symbolic interactionism has attracted its share of criticism—among other things, for its tendencies to use fuzzy definitions, recite cant, and ignore structural constraints that affect behaviors—and some of these criticisms have tended to rub off on role theory.

### COGNITIVE PERSPECTIVES IN ROLE THEORY

Empirical research in role theory has been carried out by cognitive social psychologists representing several traditions (see Biddle 1986, for a general review). Some of this work has focused on role playing, some of it has concerned the impact of group norms, some of it has studied the effects of anticipatory role expectations, and some of it has examined role taking.

In addition, cognitive social psychologists have studied conformity to many forms of expectations, including instrumental norms, moral norms, norms attributed to others, self-fulfilling prophecies, beliefs about the self (such as those induced by identity projection or labeling), beliefs about others, and preferences or “attitudes.” These studies suggest that roles are often generated by two or more modes of expectational thought, and several models have also appeared from cognitive theorists reflecting this insight (see, for example, Bank et al. 1985).

Unfortunately, much of this effort ignores expectations for social positions and concentrates, instead, on expectations for individual actors. Cognitive role theory also tends to ignore the implications of its findings for structural analysis, and thus appears to be atheoretical from a sociological perspective. However, Bruce Biddle (1979) has authored a broad vision for role theory that uses information from cognitive research to build models for social system analysis.

#### RECENT TRENDS IN ROLE THEORY

Four recent trends in the development of role theory should be noted. First, although the term “role” continues to appear in most textbooks for basic courses in sociology and social psychology, it normally does not appear by itself as a major concept but rather is likely to surface in chapters on such topics as “the self,” “groups,” “institutions,” and “role taking.” In contrast, extensive discussions of roles and related concepts may be found in texts for various types of advanced courses for these fields. To illustrate, consider recent texts for courses on group dynamics. In the latest edition of his highly successful work, Donelson Forsyth (1999) devotes an entire chapter to “norms,” “roles,” and related issues, and in her new text, Joann Keyton (1999) focuses a major chapter on “group member roles,” “group norms,” and associated materials. As a rule, portrayals of role theory in such sources is straightforward: “roles” are deemed to refer to specific patterns of behavior that are associated with individuals or recognized identities; “norms” are shared expectations for conduct that may apply to all persons in the group or only to certain identities (such as “leaders”); and related concepts such as “socialization” and “role conflict” appear frequently.

Second, many authors continue to employ role concepts for discussing social relations within a specific institution or for portraying the lives of those who share an occupational identity. For example, a substantial literature has now appeared concerned with “the role of the school principal,” and a useful summary of this work may be found in a recent review by Ronald Heck and Philip Hallinger (1999). In another example, Biddle (1997) provides an extensive overview of recent research on “the role of the school teacher.” Again, much of this applied work makes clear use of concepts from role theory, with the “role” term normally used to refer to differentiated behaviors, whereas notions about behaviors that are thought to be appropriate for roles are normally termed “norms” or “role expectations.”

Third, for at least a generation, authors who have written about differences between the conduct, problems, or outlooks of men and women have used role theory as a vehicle for interpreting their findings, and this interest continues. To illustrate, for years a key journal that publishes studies concerned with gender and its problems has borne the title *Sex Roles*, but recently a particularly strong advocate for using role theory to interpret evidence about gender differences in behavior has appeared in the person of Alice Eagly (1987, 1995). Eagly asserts that such differences appear as a result of structural forces in societies—hence may differ among countries—but are sustained and reproduced because men and women develop role-appropriate expectations for those behaviors. Given the earlier, pioneering studies of Margaret Mead, such assertions would seem unexceptionable, and yet they have touched off a storm of criticism from evolutionary psychologists who prefer to believe that gender differences in conduct are hard wired and culturally universal, and have arisen from the mechanisms of Darwinian selection. (See, for example, Archer [1996].) Unfortunately, in her 1987 book on the subject, Eagly did not make clear that her argument involved only one version of role theory, and it has seemingly not occurred to her evolutionary critics that there might be other versions of the role story that would also bear on their concerns. So, in criticizing her, they have made foolish assertions about “the scope of social role theory,” and have condemned it for assumed stances that most role theorists would not advocate.

Fourth and last, every few years interesting works are published by authors who have apparently just discovered some version of role theory and are intrigued with its potential for generating insights or resolving problems in cognate fields. A good example of this type of work appears in a recent article by James Montgomery (1998). Montgomery begins by noting that, in a widely cited work, Granovetter (1985) had argued that economic action is embedded in social relationships and that rational choice theorists have subsequently explored this insight through research on prisoner's dilemma games in which long-term interaction is thought to be governed by general assumptions about "calculative trust." Empirical support for this thesis has been weak, and—drawing on work by James March (1994)—Montgomery argues that a stronger case can be made for assuming that, when engaged in long-term interaction, persons make assumptions about the social identities which they and others have assumed, and that these identities are associated with shared expectations about behaviors that are appropriate in the relationship. To illustrate, Montgomery suggests that expectations are far different when one assumes the other to be a "profit-maximizing 'businessperson'" than when the other is assumed to be a "nonstrategic 'friend.'"

Montgomery's arguments are well wrought, and their implications are spelled out through techniques of formal logic. Moreover, Montgomery points out how his arguments relate to recent work on various cognate concerns such as identity processes, artificial intelligence, situation theory, and cognitive psychology. So far so good, but (like too many recent converts) Montgomery seems not to be familiar with the bulk of work in the role field, and this leads him to make foolish errors. To illustrate, he refers to social identities as "roles" and shared expectations about behaviors as "rules"—idiosyncratic uses that will surely confuse readers. Worse, he seems not to be familiar with prior work by role theorists on his topic, including major works within the structural role theory tradition; with Ralph Linton's writings on the evolution of roles; and with the fact that much of his argument was actually made *forty* years ago by John Thibaut and Harold Kelley (1959). It does not help work in any field if scholars are unwilling to familiarize themselves with prior work on their subject, and one wonders how role theory is to make progress

in the future if even its advocates are unwilling to do their homework.

### ROLE THEORY AND THE FUTURE

As the foregoing examples suggest, role theory is currently weakened by terminological and conceptual confusion, diffuse effort, and the narrow visions of some of its proponents and critics. Nevertheless, role theory concerns central issues for sociology and social psychology, and assumptions about social positions, role behaviors, and expectations for human conduct appear widely in current social thought. Role theory will prosper as ways are found to discuss these issues with clarity, consistency, and breadth of vision.

#### REFERENCES

- Allen, Vernon L., and Evert van de Vliert, eds. 1984 *Role Transitions: Explorations and Explanations*. New York: Plenum.
- Archer, John 1996 "Sex Differences in Social Behavior: Are the Social Role and Evolutionary Explanations Compatible?" *American Psychologist* 51:909-917.
- Bank, Barbara J., Bruce J. Biddle, Don S. Anderson, Ragnar Hauge, Daphne M. Keats, John A. Keats, Marjorie M. Marlin, and Simone Valantin 1985 "Comparative Research on the Social Determinants of Adolescent Drinking." *Social Psychology Quarterly* 48(2):164-177.
- Bates, Frederick L., and Clyde C. Harvey 1975 *The Structure of Social Systems*. New York: John Wiley.
- Biddle, Bruce J. 1979 *Role Theory: Expectations, Identities, and Behaviors*. New York: Academic Press.
- 1986 "Recent Developments in Role Theory." In R. H. Turner, ed., *Annual Review of Sociology*, vol. 12. Palo Alto, Calif.: Annual Reviews.
- 1997 "Recent Research on the Role of the Teacher." In B. J. Biddle, T. L. Good, and I. F. Goodson, eds., *International Handbook of Teachers and Teaching*. Dordrecht, The Netherlands: Kluwer.
- Burt, Ronald S. 1982 *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. New York: Academic.
- Connell, Robert W. 1979 "The Concept of 'Role' and What to Do with It." *Australian and New Zealand Journal of Sociology* 15:7-17. Reprinted in R. W. Connell, ed., *Which Way Is Up? Essays on Sex, Class and Culture*, chap. 10. Sydney: Allen and Unwin, 1983.

- Eagly, Alice H. 1987 *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, N.J.: Lawrence Erlbaum.
- 1995 “The Science and Politics of Comparing Women and Men.” *American Psychologist* 50:145–158.
- Ebaugh, Helen Rose Fuchs 1988 *Becoming an EX: The Process of Role Exit*. Chicago: University of Chicago Press.
- Fisher, Cynthia D., and Richard Gitelson 1983 “A Meta-Analysis of the Correlates of Role Conflict and Ambiguity.” *Journal of Applied Psychology* 68:320–333.
- Forsyth, Donelson R. 1999 *Group Dynamics*, 3rd ed. Belmont, Calif.: Wadsworth.
- Gordon, Chad, and Paddy Gordon 1982 “Changing Roles, Goals, and Self-Conceptions: Process and Results in a Program for Women’s Employment.” In W. Ickes and E. S. Knowles, eds., *Personality, Roles, and Social Behavior*. New York: Springer-Verlag.
- Granovetter, Mark 1985 “Economic Action and Social Structure: The Problem of Embeddedness.” *American Journal of Sociology* 91:481–510.
- Hare, A. Paul 1985 *Social Interaction as Drama: Applications from Conflict Resolution*. Beverly Hills, Calif.: Sage.
- Heck, Ronald H., and Philip Hallinger 1999 “Next Generation Methods for the Study of Leadership and School Improvement.” In J. Murphy and K. S. Louis, eds., *Handbook of Research on Educational Administration*, 2nd ed. San Francisco: Jossey-Bass.
- Heiss, Jerrold 1981 “Social Roles.” In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Ickes, William, and Eric S. Knowles, eds. 1982 *Personality, Roles, and Social Behavior*. New York: Springer-Verlag.
- Keyton, Joann 1999 *Group Communication: Process and Analysis*. Mountain View, Calif. Mayfield.
- Lopata, Helena, ed. 1980 *Research in the Interweave of Social Roles: Women and Men—A Research Annual*, vol. 3. Greenwich, Conn.: JAI.
- March, James G. 1994 *A Primer on Decision Making*. New York: Free Press.
- Montgomery, James D. 1998 “Toward a Role-Theoretic Conception of Embeddedness.” *American Journal of Sociology* 104(1):92–125.
- Sales, Esther, Barbara K. Shore, and Floyd Bolitho 1980 “When Mothers Return to School: A Study of Women Completing an MSW Program.” *Journal of Education for Social Work* 16(1):57–65.
- Scheibe, Karl E. 1979 *Mirrors, Masks, Lies, and Secrets: The Limits of Human Predictability*. New York: Praeger.
- Skinner, Denise A. 1980 “Dual-Career Family Stress and Coping: A Literature Review.” *Family Relations* 29:473–481.
- Stryker, Sheldon, and Anne Statham Macke 1978 “Status Inconsistency and Role Conflict.” In R. Turner, J. Coleman, and R. Fox, eds., *Annual Review of Sociology*, vol. 4. Palo Alto, Calif.: Annual Reviews.
- , and Richard T. Serpe 1982 “Commitment, Identity Salience, and Role Behavior: Theory and Research Example.” In W. Ickes and E. S. Knorales, eds., *Personality, Roles, and Social Behavior*. New York: Springer-Verlag.
- , and Anne Statham 1985 “Symbolic Interaction and Role Theory.” In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Thibaut, John W., and Harold H. Kelley 1959 *The Social Psychology of Groups*. New York: John Wiley.
- Turner, Ralph 1979 “Strategy for Developing an Integrated Role Theory.” *Humboldt Journal of Social Relations* 7:123–139.
- 1990 “Role Change.” In W. R. Scott, ed., *Annual Review of Sociology*, vol. 16. Palo Alto, Calif.: Annual Reviews.
- van de Vliert, Evert 1979 “Gedrag in Rolconfliktsituaties: 20 Jaar Onderzoek Rond Een Theorie” (Behavior in Role Conflict Situations: 10 Years of Research About a Theory). *Nederlands Tijdschrift voor de Psychologie* 34:125–146.
- Van Sell, Mary, Arthur P. Brief, and Randall S. Schuler 1981 “Role Conflict and Role Ambiguity: Integration of the Literature and Directions for Future Research.” *Human Relations* 34:43–71.
- Winship, Christopher, and M. Mandel 1983 “Roles and Positions: A Critique and Extension of the Blockmodeling Approach.” Pp. 314–344 in S. Leinhardt, ed., *Sociological Methodology 1983–1984*. San Francisco: Jossey-Bass.
- Zurcher, Louis A., Jr. 1983 *Social Roles: Conformity, Conflict, and Creativity*. Beverly Hills, Calif.: Sage.

BRUCE J. BIDDLE

## ROLE THEORY: FOUNDATIONS, EXTENSIONS, AND APPLICATIONS

Role theory provides conceptual elements and dynamic relations across the social sciences. Indeed, the notion of *role* has become something of a “meta-construct” that has been adapted to the

scholarly focus and methodological predilections of fields such as sociology, psychology, anthropology, and management, to name just a few. Such broad application, while suiting testimony to the importance of role constructs in social theory, has led to some conceptual confusion, formulatic imprecision, and sharply diverging interpretations. Nevertheless, there remains a great deal of consensus about the integral nature of roles in the operation of social systems and the behavior of individuals.

Fundamentally, roles are organized behavioral patterns and expectations that attend a given position (hierarchical, functional, or social) or that accompany a specific situation. That is, roles encapsulate and invoke the accepted repertoire of individual conduct associated with a specific position or extant circumstance. In this way, roles provide behavioral guidelines, prescriptions, or boundaries in the form of *expectations*. These expectations can be formally assigned and explicitly stated—as in the case of occupational job descriptions—or informally assumed and tacit—as in the case of one who plays the “facilitator” role in a friendship clique. Additionally, by evoking behavioral expectations, roles affect how individuals cognitively frame, interpret, and process physical or social stimuli, and thus they further condition emotional responses. There is some controversy as to whether individuals are fully cognizant of the roles they play, but that is incidental to the underlying assumption that roles influence behavior, and thus are powerful predictors of individual action and key to understanding social systems.

This essay is not intended to provide a comprehensive review of role theory nor to propose new theoretical formulations. Rather, this essay will offer a framework for organizing role theory that hinges on *levels of analysis* and the particular *phenomenon of focus*.

There are two primary levels of analysis relevant to role theory. The first emphasizes how roles operate within and through *social systems*, such as societies or groups. The second level is concerned with how roles influence, or are influenced by, the *individuals* who inhabit them. This is essentially a classic *macro versus micro* distinction, the former being characteristic of sociological and anthropological inquiry, the latter of management and psychological inquiry (though there is, of course, some crossover). The phenomenon of focus refers

to the particular object of inquiry within each level of analysis. For instance, a researcher in the social systems tradition may focus on nations, ethnic heritage, or group cohesion, whereas a researcher in the individual tradition may focus on self-conceptions, cognitions, or conflict. The phenomena of focus vary widely within each level of analysis, and are discussed under subheadings.

### SOCIAL SYSTEMS

The underlying assumption of role theory at the broadest level is that social systems—particularly societies, cultures, organizations, groups and families—are organized and operate through roles. Hence, roles function dynamically to structure the interaction of participants so as to maintain, defend, alter, innovate, or advance the purpose of social systems. In this way, roles become the primary linkage between the social system and the individual, and are designed to communicate the expectations of the larger concern to the particular actor. Roles, then, can be viewed as indispensable mechanisms that embody the values of the social system.

**Societies and Stasis.** One of the earliest uses of role theory in social science involved the proposal that societies, like organisms, have differentiated parts that function interdependently to allow the whole to operate. In any given society, those parts would include institutions like the state or the church, each of which carry out defined obligations that reflect the priorities of that society. However, institutions in and of themselves do not execute the role. To accomplish their purposes, institutions convey that responsibility to individuals through socialization and inculcation, who in turn are responsible for enacting them. Hence, roles become the primary theoretical construct for explaining *social stability*. That is, roles *function* in a manner conducive to social order and stasis. The term “function” is important here, as *functionalism* was the name given to the major school of thought at the time (Parsons 1951).

The chief concern of functionalism was how societies decided upon, designed, communicated, and enforced roles. This concern opened up a series of issues that have occupied sociological role theory, such as: which parties designate a role, the rationale for the privilege or status assigned given roles, the mechanisms by which the social system

inculcates roles, and how to ensure a role is faithfully enacted (see Biddle 1986).

**Culture and Change.** Role theory has found its way into the study of cultures primarily through anthropology. Here, the dramatic, theatrical flavor of roles is clearly evidenced. The basic thrust is that all cultures have forms of ritual, ceremony, and pageantry that encompass symbolic societal roles which in turn play crucial social functions. Unlike sociologists, who see such institutions and their prescribed roles as maintaining stasis and order, anthropologists, notably Victor Turner (1986), argue that the purpose of such social drama is *change*. Specifically, Turner contends that, whenever individuals act in accordance with social scripts (i.e., roles), the possibility exists for “liminality”: a lodgment in time and circumstance where individuals depart from proscribed patterns and initiate new ones. The very idea of roles is to trigger or generate novelty and creativity by stepping out from that which is expected, and thus bring new meaning to the dynamic represented in the social drama. It is the tension between norms and expectations and the stability they imply, versus the necessity for change for survival’s sake, that animates the alteration of roles, which is viewed as the engine of cultural development.

**Organizations and Performance.** Whereas the emphasis is respectively on stability or change when societies or cultures are the phenomenon of focus, when organizations are the focus, the emphasis is squarely on *performance* (typically operationalized as productivity, or the difference between inputs and outputs, or costs and profits). Role theory finds its way into management at the macro level with research concerned with organizational design. The major concern is the proper way to arrange an organization for optimal performance, which constitutes a structure through which the organization is managed. Principles involved in organization design include differentiation, integration, centralization, complexity, and formalization. But a key element in the erection of an structure is the formal designation of roles that organizational actors are assigned to play (see Hall 1991).

The roles that individuals assume in organizations are typically assigned based on expertise and previous experience. That is, an individual is specifically trained or has the background to execute

the relevant duties; he or she is prepared to fill a role. But beyond possessing the requisite skills, organizational roles are designed to place individuals into the particular structure of the organization. This is accomplished primarily through two formal mechanisms and one informal mechanism. The first is the job description, which is a detailed documentation of all duties and responsibilities. The job description, then, effectively posits expectations and sets strict behavioral boundaries. The second is the reporting relationship, which describes the hierarchical order of the organization, and thus dictates channels for approval and communication. The third, and informal, mechanism by which individual conduct is guided is the organizational culture. In this case, culture refers to the organizations climate as well as its tacit mores and traditions.

From the perspective of research in organizational design, the question is the relationship between structure and performance. For instance, in industries where there is a high rate of change, research suggests that looser structures, with fewer specifications for job descriptions and more open channels of communication, tend to perform better. Suffice it to say, nowhere are roles more formally communicated, monitored, and controlled than in the management of firm performance.

**Groups and Functionality.** Another area of inquiry where role concepts play a major part is groups. Defined as two or more interdependent individuals who have come together to achieve an objective, groups can include formal work teams, friendship cliques, and even families (though family relations is often treated as an independent, free-standing field of inquiry). The conceptual elements in group research are not fundamentally different from those involved in the study of societies, cultures, or organizations. That is, to accomplish its purpose—whether that purpose be completing an organizationally assigned task or comradeship—group members must function in a complimentary manner. That *functioning*, then, is typically arranged around roles that members are assigned or assume.

The role concepts most frequently employed in group research are role identity (the attitudes and behaviors consistent with a role), role perception (an individual’s view of how to behavior in a

given situation), role expectations (other's beliefs of how one should act in a given situation), and role conflict (the contradiction of two role's expectations). These concepts are then used to predict various group dynamics—such as conformity, status relations, and cohesiveness—and outcomes—such as performance, satisfaction, creativity, and efficiency (for a review, see Goodman et al. 1987).

### INDIVIDUALS

Whether examining societies, cultures, organizations, or groups, roles are enacted by individuals. The term “enacted” is important here, since it belies the theatrical, dramaturgical roots of role theory (Simmel 1920). Moreno (1934) for instance, stressed the importance of *role playing* as a natural act of children in learning about themselves and their world, and an important aid for education as well as therapy. Perhaps the most memorable proposition is Goffman's powerful theatrical analysis (1959). Goffman's basic premise, not unlike that of Shakespeare's “all the world's a stage,” is that *all* human behavior is acted, with some allowance for the nature or type of audience. Although varying in the degree of their theoretical commitments, these thinkers underscore the central place that the metaphor of *stage acting* commands in role theory, especially as an explanatory and illustrative aid in understanding individual behavior.

**Identity and Interaction.** There is no question that individual identity—the self-conception and personality of the individual—is impacted by the society in which individuals live, the family into which they are born, the community in which they were raised, and the people with whom they associate. Identity is surely a complex, interwoven interaction of the person and his or her situation. And that roles exert a strong influence on individual identity is equally obvious in individuals' descriptions of themselves, which invariably involve roles (e.g., daughter, husband, student, lawyer). Individuals, then, show a marked propensity to understand themselves through the roles they have assumed.

The study of roles in identity formation was largely sparked by a school of thought known as *symbolic interactionism*. According to this perspective, identity evolves through the dynamic process of a communicating society. Here, society is not a

static structure that dictates roles and thus identity. Rather, it is built through interaction heavy in symbolic communication. Therefore, society is continually formed and reformed through the *reciprocal* influence of individuals taking into account one another's characteristics, and the symbolic meanings that emerge as they interact. Accordingly, neither society nor the individual ontologically precedes the other.

Traditional role theory (especially that which employees social systems as the level of analysis) and symbolic interactionism diverge on the precedence of the relationship between society, individuals, and roles. Traditional role theory assumes that roles are defined by society, which in turn logically determines identity. Symbolic interactionism, on the other hand, views roles as emerging from symbolic communication in a reciprocal relationship between the society and the individual. Here, individuals are credited with being active, creative, mindful, and volitional in their identity.

Symbolic interactionism is grounded in the philosophy of the American pragmatists (e.g., W. James, J. Dewey, and C. S. Pierce) and subsequent social scientists like G. H. Mead, C. H. Cooley, and E. Goffman. The basic premise is that the self emerges through symbolic interactions with socially recognized categories and the roles that correspond with those categories. Because roles *emerge* in relationship to others and society, the self does as well. The self is the way in which individuals understand themselves in relation to others (see Stryker and Statham 1985).

A practical implication of this is that how individuals think of themselves depends, to a great extent, on the social roles that they play. This is nicely captured by W. James: “Many a youth who is demure enough before his parents and teachers, swears and swaggers like a pirate among his tough young friends” (1890, p. 294). Also implicit is James's assertion that individuals have many selves and many social identities: “a man has as many social selves as there are individuals who recognize him and carry an image of him in their mind” (1890, p. 294). Thus, individuals can be said to have many linked selves as opposed to one united self.

The active and emergent nature of the self and identity is indicative of the work of those in the symbolic interactionism tradition. Individuals choose selves upon which to stake a claim and abandon



others which did not prove adaptive or failed to garner positive feedback. Thus, self-esteem is directly tied to the choice of selves to maintain or dismiss. In addition, roles and selves are not merely foisted upon individuals, but rather the options available open the opportunity for the exploration of *possible* selves. Recent papers, derived from James's formulations, contemporary theories of evolution, and performative dynamics, have proposed a model of the exploration and construction of possible selves (Bailey and Ford 1994; Yost et al. 1992).

**Cognition and Schematic Processing.** Roles affect individual perceptions, determinations and judgments of people, events, and causal relations through *schematic processing*. A schema is a highly ordered cognitive structure composed of knowledge, beliefs, and feelings about persons, objects, and events. Schemas, then, are mental frameworks that coherently organize memory and associations that in turn facilitate the efficient processing of information. Although there are many types of schemas—such as event schemas (e.g., the script individuals follow when dining at a restaurant) or person schemas (e.g., the knowledge, feelings, and expectations an individual has about another)—role schemas are those that organize proper behavioral patterns according to position or situation.

The notion of role schema is central to the role construct inasmuch as roles are behavioral guidelines. From a cognitive perspective, the question is how role schemas influence individual information processing. This influence occurs in both directions; that is, as *observer* and as *actor*. Research demonstrates that, when observing another, the activation of a role schema influences attention, memory, and attribution. For instance, when observing an elderly person, individuals tend to notice, recall, and render causal explanations that are consistent with an age-based role schema (e.g., the older gentleman crinkled his nose because he disapproved of the loud music). In this way, role schemas provide observers with a richly interconnected network of information by which they can categorize and thus interpret the behavior of others. Of course, as a means for comprehending others, role schemas sacrifice accuracy for the sake of efficiency, as is the case with stereotypes. As an actor, role schemas refer to the mental representations of the expectations that attend a role. Similarly, individuals access and process information

more quickly when it is related to the role they are occupying at the moment (see Fiske and Taylor 1991).

**Transition and Alteration.** Research on role transition acknowledges that individuals develop and move from one role to another in the course of their lives. Hence, role transition refers to the movement from one role to another, and specifically how individuals adapt to the transition. For instance, a promotion from staff programmer to project supervisor requires learning new duties and expectations, but also altering attitudes toward others. The same holds true for transitions from son or daughter to parent, from student to employee, and from child to adult. Such role transitions, then, challenge individuals to reconceptualize their notion of themselves, their relations to others, and their opinions and attitudes toward domain-relevant objects and events. Role transition has been examined in the management arena, with emphasis on how to facilitate the transition in order to improve performance, and in the psychological counseling arena, with attention to assuaging the emotional distress that often accompanies such periods of adjustment.

Role change can be defined as an *alteration* in the consensual understanding of the behavioral patterns of an established role. This is not a transition from one role to another, but rather a change in the expectations and boundaries of an established role. The assumption here is that roles are not static entities, but must *evolve* in order to adequately address the demands of the cultural milieu, economic conditions, or social situation (see Turner 1990).

There are three fundamental ways in which roles can change. First, roles can change according to shifting societal priorities or cultural patterns. For instance, gender roles have gone through considerable alteration as attitudes toward equal rights, access to career opportunities, and traditional obligations have been reconsidered and reconfigured in society. Second, roles can change because of formal dictates from authority. For instance, one's job responsibilities could be expanded quantitatively (e.g., supervising more people) or qualitatively (e.g., involving an entirely different skill set). Third, and perhaps the most interesting, roles can be changed by the individual who inhabits the role. For instance, individuals

may, because of either personal preferences or attitudes, redefine a "director" role to be less about planning and monitoring and more about mentoring and directing.

### CONCLUSION

Role theory has come full circle. Early formulations, especially those of Parsons (1951), Moreno (1934), and Goffman (1959), have recently gained considerable currency. For instance, functionalism has proved useful as an analytic framework for describing alterations in emerging democracies. Moreno's emphasis on role playing has found its way into pedagogy in the form of classroom exercises to illustrate concepts and executive workshops for skill development, as well as a fruitful method for therapeutic intervention. And Goffman's reliance on stage acting has influenced current thinking on identity and even research methodology. What this suggests is that role theorists are acutely aware of their theoretical heritage and progenitors, and are willing to mine the past in order to better understand the present.

Roles change as broad conditions shift. Political, economic, and technological factors are especially volatile, each in its own way altering the social system in which individuals reside and the manner in which they understand themselves. Although role theory has not been as intensely researched in last decade—a victim of academic fashion—it continues to provide an intellectual and structural foundation for fields across the social sciences. Moreover, because the late twentieth century is marked as much by change as anything else, social conditions are changing at a dizzying pace. No theoretical construct is more suited to examine the impact of such changes on the social system and the individual than role theory.

### REFERENCES

- Bailey, J. R., and C. M. Ford 1994 "Of Methods and Metaphors: Theater and Self-Exploration in the Laboratory." *Journal of Applied Behavioral Science* 30:381-396.
- Biddle, B. J. 1986 "Recent Developments in Role Theory." In R. H. Turner and J. F. Short, eds., *Annual Review of Sociology*, vol. 12. Palo Alto, Calif.: Annual Reviews.
- Fiske, S. T., and S. E. Taylor 1991 *Social Cognition*. New York: Random House.
- Goffman, E. 1959 *The Presentation of Self in Everyday Life*. New York: Doubleday.
- Goodman, P. S., E. Ravlin, and M. Schminke 1987 "Understanding Groups in Organizations." In L. L. Cummings and B. M. Staw, eds., *Research in Organizational Behavior*. Greenwich, Conn.: JAI.
- Hall, R. H. 1991 *Organizations: Structures, Processes, and Outcomes*. Englewood Cliffs, N.J.: Prentice Hall.
- James, W. 1890 *Principles of Psychology*, vol. 1. New York: Dover.
- Moreno, J. L. 1934 *Who Shall Survive?* Washington, D.C.: Nervous and Mental Disorders Publishers.
- Parsons, T. 1951 *The Social System*. Glencoe, Ill.: Free Press.
- Simmel, G. 1920 "Sur philosophie des schauspielers." *Logos* 1:339-362.
- Stryker, S., and A. Statham 1985 "Symbolic Interaction and Role Theory." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*. New York: Random House.
- Turner, R. H. 1990 *Role Change*. In W. R. Scott and J. Blake, eds., *Annual Review of Sociology*, vol. 16. Palo Alto, Calif.: Annual Reviews.
- Turner, V. 1986 *The Anthropology of Performance*. New York: PAJ.
- Yost, J. H., M. J. Strube, and J. R. Bailey 1992 "The Construction of the Self: An Evolutionary View." *Current Psychology: Research and Review* 11:110-121.

JAMES R. BAILEY  
JOHN H. YOST

## RURAL SOCIOLOGY

*Rural sociology* is the study of social organization and social processes that are characteristic of geographical localities where population size is relatively small and density is low (Warner 1974). Thus, rural sociology can be defined as the sociology of rural society. Since rural societies do not exist in isolation, rural sociology also addresses the relation of rural society to the larger society. Therefore, it deals also with spatial organization and the processes that produce spatial allocations of population and human activities (Newby 1980; Newby and Buttel 1980).

There is a temptation to equate rural sociology with American rural sociology because the latter is most thoroughly institutionalized and there

are more practitioners in the United States than anywhere else in the world. While rural sociology, in its institutionalized form, originated in America, it has flourished in other regions of the world, especially since the end of World War II. No doubt this is due in large part to the “modernization” efforts in the many nations that gained independence since 1950. Outside North America, sociological investigations of rural society often are referred to as peasant studies, development studies, or village studies rather than rural sociology (Newby 1980). Moreover, some aspects of rural sociological analysis are closely related to other social science disciplines, such as settlement patterns with human geography, family and kinship systems with social anthropology, and land tenure and farming systems with agricultural and land economics.

### ROOTS IN SOCIAL THOUGHT

Although the subject matter of rural sociology has been of keen interest to social thinkers for centuries, its treatment by the major nineteenth-century classical theorists led to a polarization that continues today (Duncan 1954; Hofstee 1963; LeFebvre 1953; Mendras 1969). Two points of view, both deeply embedded in the social thought and literature of Western culture, and both quite limiting if not erroneous, have predominated. The first tradition, an image drawn from the Arcadia of Greek mythology, has been the glorification of village life for the supposed pastoral virtue of its people. The second tradition has been that of the Enlightenment and modern Western rationalism, which viewed the technological and organizational character of urban industrial forces as being superior to the alleged backwardness of rural areas.

These two traditions were ultimately embraced in major nineteenth-century social theories (Nisbet 1966). Some theorists, typified by Emile Durkheim and by Karl Marx to a lesser extent, viewed the urban industrial complex as the center of a new civilization emerging from the social transformations of the industrial revolution. Rural society, in this perspective, was regarded as a residual of preindustrial society and increasingly to be relegated to a secondary status. Other theorists, such as Toennies [1887] (1957) and early-twentieth-century interpreters of Toennies (e.g., Sorokin

and Zimmerman 1929), viewed the emergent cities of industrial capitalism as monuments to the degradation of civilization. Both points of view are deeply imbedded in the social thought of Western culture and continue to shape the perspectives of rural sociology as a scientific enterprise.

### RURAL SOCIOLOGY IN AMERICA

The roots of rural sociology in America lie in the social and political turmoil associated with America's version of the Industrial Revolution, which followed the Civil War. As industrial capitalism made its great surge, urban America was on the move, quickly surpassing earlier achievements of European nations—yet in the midst of obviously rising affluence there existed a paradoxical injustice of poverty and inequality, especially in rural areas (Goodwyn 1978). William Jennings Bryan was defeated in 1896 as the Populist Party candidate for president, but the political unrest in the countryside continued to be a source of concern to urban industrialists, who depended on farmers to provide a stable supply of cheap food for the growing army of industrial workers.

The Country Life Movement emerged at the turn of the century as an urban-sponsored alternative to the radical economic proposals of the rural Populists (Bowers 1974; Danbom 1979; Swanson 1972). It was a social, cultural, and moral reform movement that adopted the view that rural society was backward, lagging behind the evolution of an advanced urban society. The problems of rural people were viewed as stemming from a lack of organization, failures of rural social institutions, inadequate infrastructures, and technological backwardness, rather than from the failures of the industrial capitalist system, as the Populists claimed.

In 1908 President Theodore Roosevelt gave legitimacy to the reform movement by appointing the Commission on Country Life. Spurred by the President's Commission and the Country Life Movement, Congress in 1914 passed the Smith-Lever Act, which created the Cooperative Extension Service to modernize rural America (Hooks and Flinn 1981). In 1925 Congress passed the Purnell Act, which provided colleges of agriculture and agricultural experiment stations with funds to support rural sociological research. Shortly thereafter, departments of rural sociology began to emerge

within universities, often separated from departments of sociology (Sewell 1965). The institutionalization of rural sociology was given further impetus in 1936, when rural sociologists established their own journal, *Rural Sociology*, and during the following year, when they divorced themselves from the American Sociological Society (now the American Sociological Association) by forming the Rural Sociological Society. During the Depression, rural sociology received substantial support for research regarding the socioeconomic status of farm families and the effectiveness of various New Deal federal programs (Larson et al. 1992).

Because of its historical roots, rural sociology has been an active participant in two conflicting social policies derived from the opposing views of rural society in social thought. The institutional separation of rural sociology from sociology, its organizational location in colleges of agriculture, and its functional integration with cooperative extension have given American rural sociology a strong attachment to technologically driven modernization. For many of its institutional sponsors, whose primary goal has been the technological advancement of agriculture, the predominant justification for supporting rural sociology research has been its presumed ability to enhance the process of modernization of rural society.

Two important consequences have followed from this sponsorship. First, the research agenda of rural sociology has been significantly influenced by politicians and administrators of colleges of agriculture and agricultural experiment stations. Thus, American rural sociological research has tended to be driven primarily by the need to be "useful" in solving practical problems involved in transforming rural society. Second, theoretical development within rural sociology has atrophied. Theoretical work that may contradict the prevailing social policy dogma and thereby threaten its financial and institutional support has been particularly uncommon. Thus, the practice of American rural sociology has been part of an explicit social policy of transforming rural society (Newby 1980).

The opposing cultural theme portrays rural society as a way of life that is superior to existence in the cities and threatened by urban industrial capitalism (Sorokin and Zimmerman 1929). It has protagonists within rural sociology and in society

for whom the problem is how to preserve the wholesome qualities of rural society against the encroachments of urban industrial capitalism (e.g., how to avoid community disintegration, loss of local autonomy, the collapse of the family farm, the decline of the traditional rural way of life, degradation of the rural landscape, and depletion of nonrenewable natural resources). These Jeffersonian values of community, individualism, family entrepreneurship, and grass-roots democracy inspire private and public sponsorship of many rural sociological endeavors (Gilbert 1982). Thus, American rural sociology has been significantly involved in two explicit and conflicting social policies. First, it has contributed to positivistic social science by providing the basic descriptive information about rural populations, institutions, and social processes that have guided the development of programs to transform rural society. Second, it has served those committed to preserving selected elements of rural society, a practice that often is perceived by agricultural administrators and proponents of technological innovations as creating barriers to progress.

### MAJOR RESEARCH TOPICS

Within the context of these conflicting and vacillating social policy orientations, rural sociology in America has generated a substantial body of research. Some research topics have emerged principally in response to the social policy of transforming rural society and have followed the paradigm of positivism. Other topics are associated more clearly with the preservationist policy orientation and the paradigm of critical sociology. While the alignment of social policy and scientific paradigms is not perfect, there is a clear pattern of association. Both sets of orientations have existed within rural sociology since its inception, with the modernization-positivism orientation clearly dominating the research enterprise until recently.

#### **Modernization-Positivism—Oriented Research.**

One of the primary concerns of the Commission on Country Life was the lack of complete and accurate information about the conditions of life in rural America. Thus, study of the rural population was one of the first research topics to emerge (Brunner 1957). Initially research was devoted primarily to description of the rural population,

not only in an effort to provide more accurate counts of people but also to report on their characteristics in greater detail and to describe demographic processes more accurately. Population studies continue to be extremely important in providing the basic descriptive information about the rural population that is needed to guide the development of programs to transform rural society (Fuguitt et al. 1989; Garkovich 1989).

To the extent that rural population studies depart from purely demographic analyses and venture into sociological investigations, they are usually guided by the systemic perspective of human ecology (Hawley 1950, 1986). In this more sophisticated systems model, population size and density are treated as interdependent with the environment, the level of technology, and the social organization of a locality. It is presumed that population size and density will expand to the maximum possible within the constraints imposed by the other components of the system, especially the technology of transportation and communication. While the perspective offers promise of merging social and spatial analysis, the results have been only partially successful. Rural population studies cum human ecology have yet to integrate the social and spatial levels of reality.

As more information about rural populations became available, comparisons with urban populations became possible, and there followed a prolific production of research to examine the belief that population size and density set the conditions of social action and social organization. This was a fundamental premise of the romanticists among the classical sociological writers noted earlier, and it was translated sociologically into the "rural-urban continuum" of Sorokin and Zimmerman (1929) and later the "folk-urban continuum" of Redfield (1947). The evidence that there are universal differences in the cultural and social characteristics that may be derived from differences in population size and density has not been convincing (Pahl [1966] 1970). Thus, while comparisons are drawn between rural and urban populations, the causality argument associated with the rural-urban continuum has been discarded by most rural sociologists.

Rural sociologists have conducted hundreds of community studies that serve as a major source

of information for the design of community development programs (Bell and Newby 1972; Summers 1986, Luloff and Swanson 1990; Wilkinson 1991). From Galpin's pioneering study in 1915 until the mid-1960s, the study of community was almost synonymous with rural sociology in the United States. By that time the rural-urban continuum, which was the chief frame of reference for many investigators, was falling into disrepute (Pahl [1966] 1970). Their studies were being criticized for their impressionistic methodologies and their excessively descriptive nature (Colin Bell and Newby 1972). Moreover, proponents of the mass-society thesis argued that communities had been eclipsed by the forces of urbanization, bureaucratization, and centralization (Stein 1964; Vidich and Bensman 1958; Warren 1963). Community was alleged to be no longer a meaningful locus of social decision making. It was presumed that the increased presence of extralocal forces in the community (vertical integration) had destroyed the horizontal integration of communities and rendered small rural communities powerless in the face of broad and powerful forces of mass society. Although the tradition of holistic community studies has not returned to its former status, evidence clearly supports the argument that increased vertical integration does not necessarily destroy horizontal integration (Richards 1978; Summers 1986). Rather, it is more consistent with the empirical data to view local autonomy as a variable, and the impact of changes in vertical integration as varying according to a complex matrix of variables characterizing the external agent and the community.

In 1897, W. E. B. DuBois began a series of analyses of economic conditions among rural black groups and their relation to agriculture (DuBois 1898, 1901, 1904). Indeed, as we will stress later, since the turn of the century rural sociologists have been studying the "sociology of agriculture," although that expression did not come into use until the 1970s (Buttel et al. 1990). Land tenure and types of farming enterprises were studied to understand the relations of farming and agriculture-based businesses to the conditions of rural living. The methodology of these studies was often that of the community survey, and consequently there was much overlap with population and community studies. Most of these studies were descriptive in nature, and they generated taxonomies of

farming enterprises, which provided further refinements of farm family and farming community characteristics. The resulting social maps of farming communities provided detailed information to guide modernization programs, especially those of the Cooperative Extension Service, with its offices in virtually every rural county in the United States.

Although technological innovations have been occurring in agriculture for centuries, technological change was revolutionized with the introduction of hybrid corn (Ryan and Gross 1943). With this innovation, adoption and diffusion research became a new research field led by rural sociology. The first research focused on identifying which farmers had the highest rates of adoption of hybrid corn and how the adoption process was diffused to other farmers. Soon the research encompassed other innovations and spread to other countries with the modernization era at the end of World War II (Rogers 1995; Fliegel 1993). The basic processes of adoption and diffusion are now reasonably well understood, and training programs based on this knowledge are being implemented worldwide in areas of human behavior that reach well beyond farming practices to include health and nutrition, resource conservation, business management, and many other areas.

#### **Preservationist–Critically Oriented Research.**

By the 1960s there was a strong and growing disillusionment with the societal consequences of positivistic social science and the absence of a structuralist perspective (Newby and Buttel 1980). It was claimed that theory and research had become uncoupled, with theory being excessively abstract and research exhibiting a mindless empiricism. Several rural sociologists involved in international development research offered challenges to the Western development orthodoxy by claiming that modernization was serving the interests of the powerful and wealthy rather than improving the social and economic well-being of peasants and poor people (Havens 1972; Havens and Flinn 1975; Thiesenhusen 1978). In North America and Europe similar claims were being expressed in the environmental, civil rights, and other social justice movements of the late 1960s. The emerging research topics in rural sociology manifest the intellectual ferment of a more critical perspective on existing public policies, especially

in relation to established institutions of agricultural and rural research and programs claiming to improve rural communities and institutions. The emergent critical perspective incorporates a diversity of theoretical views that recognize the active role of the state in public policy and argue that it is subject to the influences within society of powerful interest groups that often are formed along the lines of class, race, ethnicity, or gender. Thus, the contemporary theoretical debates within rural sociology draw heavily on neo-Marxist and neo-Weberian orientations, with the result that rural sociology and sociology are closer intellectual partners today than at any time in the past fifty years. Although virtually all facets of rural social organization and processes are subjected to the emergent critical perspective, some areas have received more attention than others.

As we note below, the most distinctive feature of this “new rural sociology” (Newby and Buttel 1980) was the prominence of Marxist and neo-Marxist interpretations of the social differentiation of agriculture. This critical new rural sociology was applied most extensively to understanding the paradox of the growth of large-scale capitalist agriculture accompanied by the persistence of the small-scale family or subfamily farm. Efforts to understand this duality of agricultural structure has led to sharp debates about the barriers to capitalist transformation of agriculture, the role of small-scale and part-time farms in a functionally integrated capitalist industrial and agricultural system, and the role of the state in promoting capitalist agriculture. These critical perspectives have also been directed to understanding the social significance of the research apparatus of the land grant university system itself, particularly as to whether land-grant agricultural science has essentially served as a state policy that has helped to underwrite the growth of large-scale capitalist agriculture (Busch et al. 1991; Goodman and Redclift 1991; Kloppenburg 1988).

Until the late 1960s and early 1970s, rural sociology’s contribution to the sociology of development was confined largely to adoption-diffusion research related to new agricultural technologies in Third World countries (Hoogvelt 1997; Toye 1987; Webster 1990). Shortly thereafter, following on the growing disillusion with adoption-diffusion research, development-related inquiry in rural sociology shifted dramatically. Much of the

impetus behind criticism of the adoption-diffusion approach came from rural social scientists who did Third World research and who became acutely aware of its shortcomings as a vehicle for understanding agricultural change in the developing countries (Havens and Flinn 1975; George 1976; Lipton 1977; Flora 1990). The theoretical ferment in rural sociology in the 1970s and 1980s was to a large extent derived from new concepts in the sociology of development, such as the "development of underdevelopment," "dependent development," "core-periphery relations," and "capitalist world-system," which were developed as critiques of modernizationism as applied to the developing world.

The "post-diffusion" phase of the sociology of development has led to a far more diversified program of rural sociological research on development processes in the Third World. Although rural sociologists who do sociology of development research tend, not surprisingly, to give particular stress to agricultural development and its environmental implications, increasingly rural sociologists in the United States and other advanced countries do research on development processes that is often indistinguishable from that conducted by scholars who are not identified as rural sociologists. Also, as noted earlier, in many developing countries rural sociology is virtually synonymous with sociology of development, development studies, peasant studies, village studies, and so on.

To a certain extent this emerging research area overlaps the political economy of agriculture with its emphasis on technological change and its effects of distribution of ownership and control of resources, as well as equity in the distribution of benefits of new technologies (Field and Burch 1988). There is the additional concern with the depletion and pollution of nonrenewable resources (Schnaiberg and Gould 1994; Bell 1998). Social and economic impact assessment has emerged as a research activity that often is characterized by its critical perspective (Freudenburg 1986). A comprehensive theory has not yet emerged that links technological change in natural resource industries to the full range of its ramifications for the environment, its socioeconomic impacts, and its associations with industrial structures. However, the magnitude of its potential impacts and the

associated public concern suggests that this area of research has a viable future (Freudenburg 1992).

Since the 1920s, agriculture and natural-resource-based industries have been declining as sources of employment; the rate of decline accelerated dramatically after World War II. For a brief period during the 1970s manufacturing was a major source of employment growth in rural areas as industries sought cheaper land, lower taxes, and a nonunion labor force willing to work for lower wages and fewer benefits. Although this process continues, service industries have emerged as the major source of employment growth (Brown et al. 1988). These shifting labor demands have been accompanied by high unemployment in rural areas and a growth of temporary and part-time work, with resulting loss of wages and increasing levels of poverty. Rural labor market analysis has emerged as a new research area in rural sociology as a consequence (Summers et al. 1990). Much of the research is devoted to describing more precisely the nature and extent of rural unemployment and underemployment. However, the theoretical interpretations generally are sensitive to the linkages of rural labor markets to broader issues of economic restructuring. While labor demand-oriented and human capital explanations persist, there are attempts to understand the functioning of rural labor markets within the context of capitalist market institutions in a manner that is reminiscent of institutional labor economics.

Gender studies are not new to rural sociology; the role of women in farming has been a subject of research for at least a quarter-century (Haney and Knowles 1988). However, the past decade has witnessed the emergency of theoretical and empirical studies that attempt to explain how the institutions of capitalism, patriarchy, and the domestic ideology influence the work roles of men and women. A major focus of these recent studies has been the nature and extent of farm women's involvement in farm, household, and off-farm work. The rich descriptive detail of gender-based allocations of labor is being integrated into more comprehensive theoretical interpretations of structural changes in both agricultural and nonagricultural industries (Beneria 1985; Leon and Deere 1987; Sachs 1996).

For the past twenty-five years the United States has pursued a variety of programs and policies

intended to alleviate poverty (Sanderfur and Tienda 1988; Snipp 1989; Wilson 1987). In spite of these efforts, poverty persists at rates that are higher in rural areas than in urban areas, and the difference is increasing. Moreover, rural poverty is disproportionately concentrated in minority populations. Within the critical perspective it is argued that past and present institutional barriers limit the access of minority populations to the means of economic well-being. Persons of working age are disproportionately handicapped by deficiencies of human capital and discriminatory practices in the labor market. Moreover, these failures have produced a generation of elderly persons who are denied access to important public insurance programs such as Social Security because they were excluded from the labor market in years past or were employed in industries that were not covered by such programs. Thus, the state is called into question for its poor performance in developing and implementing adequate public policies, a failure that is alleged to benefit the interests of the wealthy and powerful classes of society (Summers 1991).

#### **AGRICULTURAL CHANGE AND THE SOCIOLOGY OF AGRICULTURE**

As social scientists and historians have begun to reflect on the momentous and often convulsive changes that have occurred during the twentieth century, many have noted that the most far-reaching social change of the century has arguably been the rapid decline of peasantries and of farm life, particularly since World War II (Hobsbawm 1994). The "depeasantization" of the advanced industrial countries has proceeded the farthest, but the very rapid decline of the peasant societies along with massive streams of rural-to-urban migrants that is now occurring in the developing world is, if anything, more stark (Araghi 1995).

The manner in which rural sociologists have conceptualized the processes and the significance of social-structural changes in agriculture has involved not only debate between the two overarching theoretical positions that have long characterized rural sociology, but also political and ideological positions on agriculture in society at large. Thus, on one hand, theories in the sociology of agriculture tend to fall within either the modernizationist

tradition (e.g., adoption-diffusion) or the critically oriented tradition (e.g., Lenin's and Kautsky's theories of capitalism and rural differentiation; see Goodman and Redclift 1982) discussed earlier. Over and above the differences and debates across theoretical traditions are changing sociopolitical views about agriculture and food.

We noted earlier that, from the beginnings of rural sociology around the turn of the twentieth century, agriculture was one of its most central subject matters. But what was considered interesting or important about agriculture has changed dramatically over time. Early rural sociology was largely focused on the sociology of agricultural communities. Rural sociology was later dominated by the adoption and diffusion of agricultural innovations. While these two traditions differed in their views of what it was about agriculture that was most worthy of study, both were modernizationist perspectives that tended to see the decline of family farming and restructuring of agriculture as being natural components of rural (and overall social) development.

The term "sociology of agriculture" can be best understood as a movement among rural sociologists in the mid- to late 1970s in reaction to two related but distinct components of modernizationism. The sociology of agriculture was, in the first instance, a reaction against rural sociological theories which, at least implicitly, accepted the inevitability and desirability of the demise and destruction of peasantries in the developing world and family farming in the industrial world. What Newby and Buttel (1980) meant by the notion of the "new rural sociology" was that a more adequate rural sociology required a more critical theoretical view about how and why farmers and other rural people were witnessing disintegration of their ways of life. The second defining feature of the new rural sociology was that it sought to take seriously the growing public and social movement concerns about the loss of family farms, the problems faced by agricultural communities, and the role of land-grant universities and public research.

The pattern of farm structural change that has occurred in the United States is not entirely typical of that of the rest of the industrial world, but the past century of changes in the American structure of agriculture typify the theoretical and broader



social issues at stake in the sociology of agriculture. In 1940, there were about 7 million American farms, home to about 30 million people (or about 25 percent of the U.S. population). By the end of the century there were only about 1.8 million farms, and the farm population (which numbered a little less than 7 million people) was less than 2 percent of the U.S. total. Even more striking is that fact that the last Census of Agriculture in the twentieth century (the 1997 Census) showed that a mere 26,000 farms with gross annual sales of \$1,000,000 or more (representing only 1.4 percent of the total number) accounted for about 42 percent of gross farm sales; by contrast, less than twenty years earlier, farms with gross annual farm sales of \$200,000 or more represented 3.3 percent of farms and about 44 percent of total sales (according to the 1978 Census of Agriculture). Thus, American farming has become increasingly concentrated. U.S. farm structure has also become highly dualistic; a handful of very large farms account for the bulk of output, while roughly 1.2 million small, "subfamily" (mostly part-time) farms account for the bulk of the farm population but very little of the output. In between, the middle stratum of farms—the prototypical medium-sized full-time family farm—has declined in numbers and percent of farm sales as the dualism of agriculture has been continued apace. Despite the rapid restructuring of agriculture from 1940 to the end of the century, nearly 95 percent of American farms continue to be family-proprietor or partnership farms. Thus, family farming—even if many of the largest and smallest family operations bear little resemblance to the traditional notion of a family farm—has persisted in the midst of otherwise convulsive change in agriculture and rural America.

The new rural sociology of agriculture in the late 1970s and early 1980s therefore ironically had two very different problematics—the decline/differentiation and the persistence of family farming—on which to focus its research. The new rural sociology drew on three major early-twentieth-century classical theories in focusing on these two problematics. V. I. Lenin tended to be the principal classical antecedent of theories of rural class differentiation (e.g., de Janvry 1980; Friedland et al. 1981; Havens et al. 1986), which tended to foresee agriculture undergoing differentiation into capital and labor, in much the same way that had

occurred in nonfarm industry. A. V. Chayanov and K. Kautsky were most influential in the work of scholars who sought to explain the persistence of family farming. Theories of the persistence of family farming generally explain the phenomenon in terms of the obstacles or the forms of resistance that exist to the development of capitalist agriculture (e.g., how the seasonal-biological nature of agriculture makes farming unattractive for large-scale investments [Mann 1990] or how independent commodity producers exhibit different rationalities [Mooney 1988] or enjoy certain advantages over capitalist producers [Friedmann 1978]).

While the agrarian differentiation/persistence debate dominated the sociology of agriculture through the early 1990s, the sociology of agriculture has made two significant shifts—toward studies of farming styles, on one hand, and the globalization of agriculture, on the other—over the last decade. The most recent versions of the sociology of agriculture have been partly a response to the current era of "globalization," trade liberalization, hypermobility of financial capital, World Bank–International Monetary Fund (IMF) imposition of structural adjustment reforms on the developing world, the rapid industrialization of certain sectors of farming (especially livestock and fresh fruits and vegetables), and the remarkable pace of concentration in the agricultural inputs and agro-food industries. It also became apparent to many scholars that most of the theories that dominated the sociology of agriculture in the 1970s through the early 1990s had two possible weaknesses: First, "new rural sociology" theories tended to be somewhat economic and deterministic. Second, these "new rural sociology" theories tended to locate the dynamics of agricultural change largely, if not entirely, within agriculture itself. These theories tended to give short shrift to the off-farm components of agro-food systems and to the global political-economic environment of agriculture.

The second generation of the sociology of agriculture can be understood as being a response to these two shortcomings of new rural sociology theories as well as to the intellectual and policy challenges posed by globalization. The first response has been the "farming styles" research tradition, and is often referred to as the "Wageningen School" approach because two of its most prominent researchers (van der Ploeg 1992; Long 1992)

are located at Wageningen University in the Netherlands. The Wageningen School perspective is a neo-Weberian or "actor-oriented" approach which stresses how diverse rural cultures interacting within diverse national economies and natural environments tend to give rise to diverse "farming styles." Thus, it is argued that there are multiple sources of diversity in farming structures, technologies, rationalities, and practices that serve to obviate the otherwise powerful political-economic processes of globalization and homogenization.

The second, and most influential, new approach in the sociology of agriculture—the agro-industrial globalization tradition—reflects a conviction that chief among the factors propelling agricultural change are matters such as national political-economic processes, the world economy, and geopolitics which lie outside of the realm of agriculture per se (see Friedmann 1982; Friedmann and McMichael 1989). Many scholars working within this new agro-industrial globalization tradition have emphasized the growing ascendancy on the part of agribusiness multinationals as post-World War II protectionist institutions and regulations have been dismantled. Studies in this genre emphasize how private firms are increasingly assuming the standard-setting and regulatory functions formerly undertaken primarily by governments, and how large corporations are playing a growing role in shaping the structure and performance of agro-commodity chains (e.g., Bonanno et al. 1994). Other scholars stress how the emerging structure of the new world food order reflects the growing role of monetary instability and Third World debt. Monetary disorder and debt have created the political-economic conditions for the liberalization of agricultural trade through international regimes such as the World Trade Organization (WTO) and the North American Free Trade Agreement (NAFTA). For example, by serving to justify structural adjustment policies which require that development countries adopt agro-food export policies and reduce food subsidies in order to repay their loans, monetary disorder and debt have been crucial factors in the late twentieth century restructuring of food systems (Friedmann and McMichael 1989; McMichael 1994). These new global food systems contrast sharply with the postwar national-type food order in which world nations tended to control their own agricultural systems and food

supplies through national food regulations and domestic agricultural policies.

## CONCLUSION

These emergent research topics have not displaced those of an earlier period of rural sociology; they coexist. In doing so, rural sociology continues to serve two conflicting social policy agendas that reflect divergent views of rural society. The field has not escaped its origins in the social thought of nineteenth-century Europe. It does appear to be renewing its intellectual kinship with sociological theory.

The future of rural sociology as a research domain and as an intellectual endeavor appears to be very promising. Only a decade ago some observers were predicting its demise on the grounds that agriculture was declining as a source of employment and urbanization was continuing on a worldwide scale. However, predictions of the death of rural sociology seem to have been premature. The majority of the world population still lives in rural areas, and agriculture still plays a major role in the economies of most nations of the world. The globalization of food systems remains one of the most critical determinants of human well-being (Goodman and Watts 1997). The ending of the Cold War and the opening of the Eastern Bloc to greater scientific and intellectual exchanges create a vast new market for rural sociology, since all of these nations are predominately rural in composition. Finally, rural sociologists are expanding the scope of their work to include a much broader array of social phenomena and accepting the challenge of building the theoretical and empirical bridges between rural and urban aspects of society.

The growth of rural sociology professional associations is further evidence of its good health. In addition to the Rural Sociological Society, which was created in 1937, there are now the International Rural Sociological Association and independent associations in all the world's regions. Membership in all these associations is increasing; the Rural Sociological Society remains over 1,000, and annual meeting attendance has been in excess of 500 for most years in the 1990s.

(SEE ALSO: *Agricultural Innovation; Community; Human Ecology and the Environment; Population*)

## REFERENCES

- Araghi, Farshad 1995 "Global Depeasantization, 1945–1995." *Sociological Quarterly* 36 (Spring):337–368.
- Bell, Colin, and Howard Newby 1972 *Community Studies: An Introduction to the Sociology of the Local Community*. New York: Praeger.
- Bell, Michael 1998 *An Invitation to Environmental Sociology*. Thousand Oaks, Calif.: Pine Forge.
- Beneria, Lourdes 1985 *Women and Development: Sexual Division of Labor in Rural Societies*. New York: Praeger.
- Bonanno, Alessandro, Lawrence Busch, William H. Friedland, Lourdes Gouveia, and Enzo Mingione (eds.) 1994 *From Columbus to ConAgra*. Lawrence: University Press of Kansas.
- Bowers, W. L. 1974 *The Country Life Movement in America, 1900–1920*. Port Washington, N.Y.: Kennikat.
- Brown, David L., and Kenneth Deavers, eds. 1988 *Rural Economic Development in the 1980s*. Rural Development Research Report No. 69. Washington: U.S. Department of Agriculture, Economic Research Service.
- Brunner, Edmund des. 1957 *The Growth of a Science: A Half-Century of Rural Sociological Research in the United States*. New York: Harper and Row.
- Busch, Lawrence, William B. Lacy, Jeffrey Burkhardt, and Laura R. Lacy 1991 *Plants, Power and Profit*. Oxford: Basil Blackwell.
- Buttel, Frederick H., Olaf F. Larson, and Gilbert W. Gillespie, Jr. 1990 *The Sociology of Agriculture*. Westport, Conn.: Greenwood.
- Danbom, D. B. 1979 *The Resisted Revolution: Urban America and the Industrialization of Agriculture, 1900–1930*. Ames: Iowa State University Press.
- de Janvry, Alain 1980 "Social Differentiation in Agriculture and the Ideology of Neopopulism." In F. H. Buttel and H. Newby, eds., *The Rural Sociology of the Advanced Societies*. Montclair, N.J.: Allanheld, Osmun.
- DuBois, W. E. B. 1898 "The Negroes of Farmville, Virginia: A Social Study." *Bulletin of the Department of Labor* 3(14):1–38.
- 1901 "The Negro Landholders of Georgia." *Bulletin of the Department of Labor* 6(35):647–777.
- 1904 "The Negro Farmer." In U.S. Bureau of the Census, *Negroes in the United States*. Washington: U.S. Government Printing Office.
- Duncan, Otis Durant 1954 "Rural Sociology Coming of Age." *Rural Sociology* 19(1):1–12.
- Field, Donald R., and William R. Burch, Jr. 1988 *Rural Sociology and the Environment*. Westport, Conn.: Greenwood.
- Fliegel, Frederick C. 1993 *Diffusion Research in Rural Sociology*. Westport, Conn.: Greenwood.
- Flora, Cornelia Butler 1990 "Rural Peoples in a Global Economy." *Rural Sociology* 55(Summer):157–177.
- Freudenburg, William R. 1986 "Social Impact Assessment." *Annual Review of Sociology* 12:451–478.
- 1992 "Addictive Economies: Extractive Industries and Vulnerable Localities in a Changing World Economy." *Rural Sociology* 57(Fall):305–332.
- Friedland, William H., Amy Barton, and Robert J. Thomas 1981 *Manufacturing Green Gold*. New York: Cambridge University Press.
- Friedmann, Harriet 1978 "World Market, State, and Family Farm: The Social Basis of Household Production in an Era of Wage Labour." *Comparative Studies in Society and History* 20(4):545–586.
- 1982 "The Political Economy of Food: The Rise and Fall of the Postwar International Food Order." *American Journal of Sociology* 88(Supplement):S248–S286.
- , and Philip McMichael 1989 "Agriculture and the State System: The Rise and Decline of National Agricultures, 1870 to the Present." *Sociologia Ruralis* 29(2):93–117.
- Fuguitt, Glenn V., David L. Brown, and Calvin L. Beale 1989 *Rural and Small-Town America*. New York: Russell Sage Foundation.
- Galpin, Charles J. 1915 *The Social Anatomy of an Agricultural Community*. Madison: Wisconsin Agricultural Experiment Station.
- Garkovich, Lorraine 1989 *Population and Community in Rural America*. New York: Praeger.
- George, Susan 1976 *How the Other Half Dies*. New York: Penguin.
- Gilbert, Jess 1982 "Rural Theory: The Grounding of Rural Sociology." *Rural Sociology* 47(Winter):609–633.
- Goodman, David, and Michael Redclift 1982 *From Peasant to Proletarian*. Oxford: Basil Blackwell.
- 1991 *Refashioning Nature*. London: Routledge.
- , and Michael Watts (eds.) 1997 *Globalising Food*. New York: Routledge.
- Goodwyn, Lawrence 1978 *The Populist Movement: A Short History of the Agrarian Revolt in America*. New York: Oxford University Press.
- Haney, Wava G., and Jane B. Knowles (eds.) 1988 *Women and Farming: Changing Roles, Changing Structures*. Boulder, Colo.: Westview.
- Havens, A. Eugene 1972 "Methodological Issues in the Study of Development." *Sociologia Ruralis* 12(3–4):252–273.

- , and William L. Flinn 1975 "Green Revolution Technology and Community Development: The Limits of Action Programs." *Economic Development and Cultural Change* 23(April):469–481.
- , Gregory Hooks, Patrick H. Mooney, and Max J. Pfeffer, eds. 1986 *Studies in the Transformation of U.S. Agriculture*. Boulder, Colo.: Westview.
- Hawley, Amos H. 1950 *Human Ecology*. New York: Ronald.
- 1986 *Human Ecology: A Theoretical Essay*. Chicago: University of Chicago Press.
- Hobsbawm, Eric 1994 *The Age of Extremes*. New York: Pantheon.
- Hofstee, E. W. 1963 "Rural Sociology in Europe." *Rural Sociology* 28(3):329–341.
- Hoogvelt, Ankie M. M. 1997 *Globalization and the Postcolonial World*. Baltimore, Md.: Johns Hopkins University Press.
- Hooks, Gregory M., and William L. Flinn 1981 "The Country Life Commission and Early Rural Sociology." *Rural Sociologist* 1(March):95–100.
- Kloppenborg, Jack, Jr. 1988 *First the Seed*. New York: Cambridge University Press.
- Larson, Olaf F., Edward O. Moe, and Julie Zimmerman (eds.) 1992 *Sociology in Government*. Boulder, Colo.: Westview.
- LeFebvre, Henry 1953 "Perspectives de la Sociologie Rurale." *Cahiers Internationaux de Sociologie (International Journal of Sociology)* 14:122–140.
- Leon, Magdalena, and Carmen Diana Deere (eds.) 1987 *Women and Rural Policy: Feminist Perspectives*. Boulder, Colo.: Westview.
- Lipton, Michael 1977 *Why Poor People Stay Poor*. London: Temple Smith.
- Long, Norman 1992 "From Paradigm Lost to Paradigm Regained: The Case for an Actor-Oriented Sociology of Development." In N. Long and A. Long, eds., *Battlefields of Knowledge*. London: Routledge.
- Luloff, A. E., and Louis E. Swanson, eds. 1990 *American Rural Communities*. Boulder, Colo.: Westview.
- McMichael, Philip (ed.) 1994 *The Global Restructuring of Agro-Food Systems*. Ithaca, N.Y.: Cornell University Press.
- Mann, Susan Archer 1990 *Agrarian Capitalism in Theory and Practice*. Chapel Hill: University of North Carolina Press.
- Mendras, Henri 1969 *Rural Sociology in France*. Paris: Mouton.
- Mooney, Patrick H. 1988 *My Own Boss*. Boulder, Colo.: Westview.
- Newby, Howard 1980 "Rural Sociology—A Trend Report." *Current Sociology* 28(1):1–141.
- , and Frederick H. Buttel 1980 "Toward a Critical Rural Sociology." In F. H. Buttel and H. Newby, eds., *The Rural Sociology of the Advanced Societies*. Montclair, N.J.: Allanheld, Osmun.
- Nisbet, Robert 1966 *The Sociological Tradition*. New York: Basic.
- Pahl, Ray E. (1966) 1970 "The Rural-Urban Continuum." *Sociologia Ruralis* 6(3–4):299–327. Reprinted in R. E. Pahl, ed., *Readings in Urban Sociology*. Oxford: Pergamon.
- Redfield, Robert 1947 "The Folk Society." *American Journal of Sociology* 52(3):293–308.
- Richards, R. O. 1978 "Urbanization of Rural Areas." In David Street and Associates, eds., *Handbook of Contemporary Urban Life*. San Francisco: Jossey-Bass.
- Rogers, Everett M. 1995 *Diffusion of Innovations*, 4th ed. New York: Free Press.
- Ryan, Bryce, and Neal C. Gross 1943 "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology* 8:15–24.
- Sachs, Carolyn 1996 *Gendered Fields*. Boulder, Colo.: Westview.
- Sandefur, Gary D., and Marta Tienda (eds.) 1988 *Divided Opportunities: Minorities, Poverty, and Social Policy*. New York: Plenum.
- Schnaiberg, Allan, and Kenneth Gould 1994 *Environment and Society*. New York: St. Martin's.
- Sewell, William H. 1965 "Rural Sociological Research, 1936–1965." *Rural Sociology* 30(December):428–451.
- Snipp, C. Matthew 1989 *American Indians: The First of This Land*. New York: Russell Sage Foundation.
- Sorokin, P. A., and Carl C. Zimmerman 1929 *Principles of Rural-Urban Sociology*. New York: Henry Holt.
- Stein, Maurice 1964 *The Eclipse of Community*. New York: Harper and Row.
- Summers, Gene F. 1986 "Rural Community Development." *Annual Review of Sociology* 12:347–371.
- 1991 "Minorities in Rural Society." *Rural Sociology* 56:177–188.
- , Francine Horton, and Christina Gringeri 1990 "Rural Labour-Market Changes in the United States." In Terry Marsden, Philip Lowe, and Sarah Whatmore, eds., *Rural Restructuring: Global Processes and Their Responses*. London: David Fulton.
- Swanson, R. M. 1972 The country life movement, 1900–1940. Ph.D. diss. University of Minnesota.

- Thiesenhusen, William C. 1978 "Reaching the Rural Poor and Poorest: A Goal Unmet." In H. Newby, ed., *International Perspectives in Rural Sociology*. New York: John Wiley.
- Toennies, Ferdinand (1887) 1957 *Community and Society: Gemeinschaft und Gesellschaft*. New York: Harper and Row.
- Toye, John 1987 *Dilemmas of Development*. New York: Basil Blackwell.
- van der Ploeg, Jan Douwe 1992 "The Reconstitution of Locality: Technology and Labour in Modern Agriculture." In T. Marsden, P. Lowe, and S. Whatmore, eds., *Labour and Locality*. London: David Fulton.
- Vidich, Arthur J., and Joseph Bensman 1958 *Small Town in Mass Society: Class, Power and Religion in a Rural Community*. Princeton, N.J.: Princeton University Press.
- Warner, W. Keith 1974 "Rural Society in a Post-Industrial Age." *Rural Sociology* 39(3):306-317.
- Warren, Roland 1963 *The Community America*. Chicago: Rand McNally.
- Webster, Andrew 1990 *Introduction to the Sociology of Development*, 2nd ed. Atlantic Highlands, N.J.: Humanities Press International.
- Wilkinson, Kenneth 1991 *The Community in Rural Society*. Westport, Conn.: Greenwood.
- Wilson, William Julius 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

GENE F. SUMMERS  
FREDERICK H. BUTTEL

# S

---

## SAMPLE SELECTION BIAS

In a linear regression model, sample selection bias occurs when data on the dependent variable are missing nonrandomly, conditional on the independent variables. For example, if a researcher uses ordinary least squares (OLS) to estimate a regression model in which large values of the dependent variable are underrepresented in a sample, estimates of slope coefficients typically will be biased.

Hausman and Wise (1977) studied the problem of estimating the effect of education on income in a sample of persons with incomes below \$15,000. This is known as a *truncated sample* and is an example of explicit selection on the dependent variable. This is shown in Figure 1, where individuals are sampled at three education levels: low (L), middle (M), and high (H). In the figure, sample truncation leads to an estimate of the effect of schooling that is biased downward from the true regression line as a result of the \$15,000 ceiling on the dependent variable. In a variety of special conditions (Winship and Mare 1992), selection biases coefficients downward. In general, however, selection may bias estimated effects in either direction.

A sample that is restricted on the dependent variable is effectively selected on the error of the regression equation; at any value of  $X$ , observations with sufficiently large positive errors are eliminated from the sample. As is shown in Figure 1, as the independent variable increases, the ex-

pected value of the error becomes increasingly negative, making these two elements negatively correlated. Because this contradicts the standard assumption of OLS that the error and the independent variables are not correlated, OLS estimates become biased.

A different type of explicit selection occurs when the sample includes persons with incomes of \$15,000 or more but all that is known about those persons is their educational attainment and that their incomes are \$15,000 or more. When the dependent variable is outside a known bound but the exact value of the variable is unknown, the sample is *censored*. If these persons' incomes are coded as \$15,000, OLS estimates are biased and inconsistent for the same reasons that obtain in the truncated sample.

A third type of selection that leads to bias occurs when censoring or truncation is a stochastic function of the dependent variable. This is termed *implicit selection*. In the income example, individuals with high incomes may be less likely to provide information on their incomes than are individuals with low incomes. As is shown below, OLS estimates also are biased when there is implicit selection.

Yet another type of selection occurs when there is *selection on the measured independent variable(s)*. For example, the sample may be selected on educational attainment alone. If persons with high levels of schooling are omitted from the model, an OLS estimate of the effect for persons with lower levels of education on income is unbi-

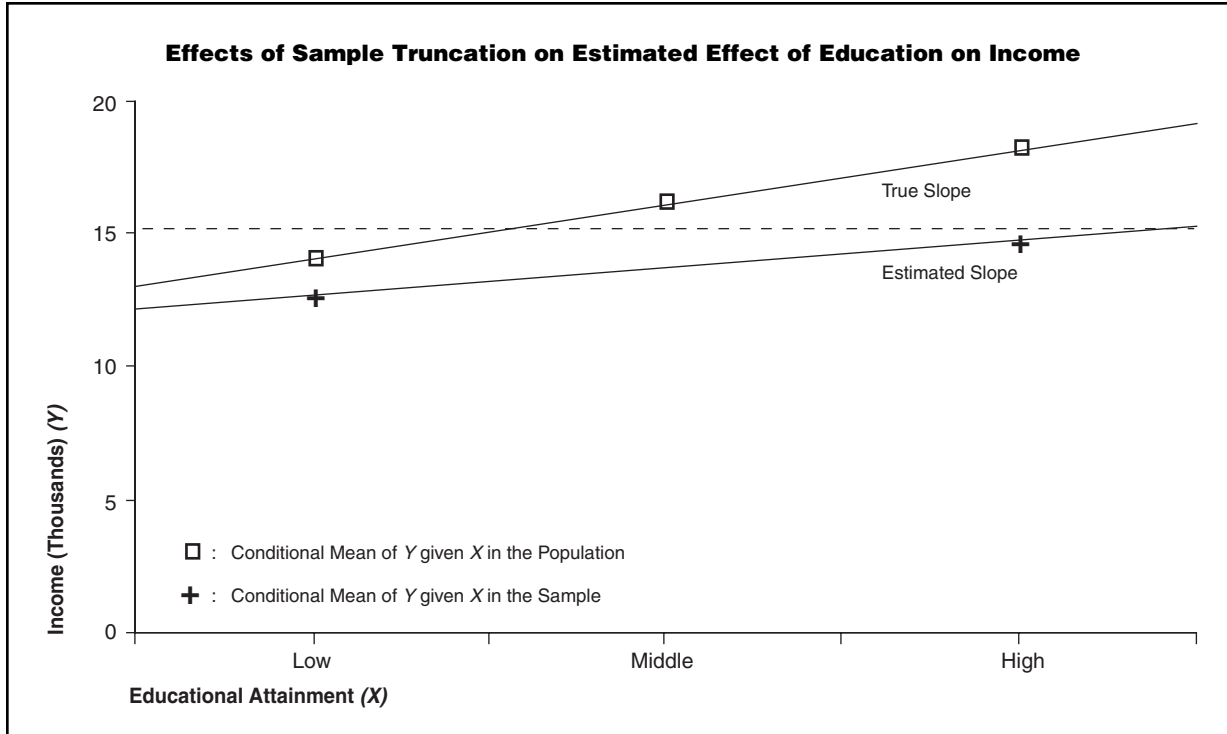


Figure 1

ased if schooling has a constant linear effect throughout its range. Because the conditional expectation of the dependent variable (or, equivalently, the error term) at each level of the independent variable is not affected by a sample restriction on the independent variable, when a model is specified properly OLS estimates are unbiased and consistent (DuMouchel and Duncan 1983).

**EXAMPLES**

Sample selection can occur because of the way data have been collected or, as the examples below illustrate, may be a fundamental aspect of particular social processes. In econometrics, where most of the basic research on selection bias has been done, many of the applications have been to labor economics. Many studies by sociologists that deal with selection problems have been done in the cognate area of social stratification. Problems of selection bias, however, pervade sociology, and attempts to grapple with them appear in the sociology of education, family sociology, criminology, the sociology of law, social networks, and other areas.

**Trends in Employment of Out-of-School Youths.** Mare and Winship (1984) investigate employment trends from the 1960s to the 1980s for young black and white men who are out of school. Many factors affect these trends, but a key problem in interpreting the trends is that they are influenced by the selectivity of the out-of-school population. Over time, that selectivity changes because the proportion of the population that is out of school decreases, especially among blacks. Because persons who stay in school longer have better average employment prospects than do persons who drop out, the employment rates of nonstudents are lower than they would be if employment and school enrollment were independent. Observed employment patterns are biased because the probabilities of employment and leaving school are dependent. Other things being equal, as enrollment increases, employment rates for out-of-school young persons decrease as a result of the compositional change in this pool of individuals. To understand the employment trends of out-of-school persons, therefore, one must analyze jointly the trends in employment and school

enrollment. The increasing propensity of young blacks to remain in school explains some of the growing gap in the employment rates between blacks and whites.

**Selection Bias and the Disposition of Criminal Cases.** A central focus in the analysis of crime and punishment involves the determinants of differences in the treatment of persons in contact with the criminal justice system, for example, the differential severity of punishment of blacks and whites (Peterson and Hagan 1984). There is a high degree of selectivity in regard to persons who are convicted of crimes. Among those who commit crimes, only a portion are arrested; of those arrested, only a portion are prosecuted; of those prosecuted, only a portion are convicted; and among those convicted, only a portion are sent to prison. Common unobserved factors may affect the continuation from one stage of this process to the next. Indeed, the stages may be jointly determined inasmuch as legal officials may be mindful of the likely outcomes later in the process when they dispose cases. The chances that a person will be punished if arrested, for example, may affect the eagerness of police to arrest suspects. Analyses of the severity of sentencing that focus on persons already convicted of crimes may be subject to selection bias and should take account of the process through which persons are convicted (Hagan and Parker 1985; Peterson and Hagan 1984; Zatz and Hagan 1985).

**Scholastic Aptitude Tests and Success in College.** Manski and Wise (1983) investigate the determinants of graduation from college, including the capacity of the Scholastic Aptitude Test (SAT) to predict individuals' probabilities of graduation. Studies based on samples of students in colleges find that the SAT has little predictive power, yet those studies may be biased because of the selective stages between taking the SAT and attending college. Some students who take the SAT do not apply to college, some apply but are not admitted, some are admitted but do not attend, and those who attend are sorted among the colleges to which they have been admitted. Each stage of selection is nonrandom and is affected by characteristics of students and schools that are unknown to the analyst. When one jointly considers the stages of selection in the college attendance decision, along with the probability that a student

will graduate from college, one finds that the SAT is a strong predictor of college graduation.

**Women's Socioeconomic Achievement.** Analyses of the earnings and other socioeconomic achievements of women are potentially affected by nonrandom selection of women into the labor market. The rewards that women expect from working affect their propensity to enter the labor force. Outcomes such as earnings and occupational status therefore are jointly determined with labor force participation, and analyses that ignore the process of labor force participation are potentially subject to selection bias. Many studies in economics (Gronau 1974; Heckman 1974) and sociology (Fligstein and Wolf 1978; Hagan 1990; England et al. 1988) use models that simultaneously represent women's labor force participation and the market rewards that women receive.

**Analysis of Occupational Mobility from Nineteenth-Century Censuses.** Nineteenth-century decennial census data for cities provide a means of comparing nineteenth- and twentieth-century regimes of occupational mobility in the United States. Although one can analyze mobility by linking the records of successive censuses, such linkage is possible only for persons who remain in the same city and keep the same name over the decade. Persons who die, emigrate, or change their names are excluded. Because mortality and migration covary with socioeconomic success, the process of mobility and the way in which observations are selected for the analysis are jointly determined. Analyses that model mobility and sample selection jointly offer the possibility of avoiding selection bias (Hardy 1989).

## MODELS OF SELECTION

Berk (1983) provides an introduction to selection models; Winship and Mare (1992) provide a review of the literature before 1992. We start by discussing the censored regression, or tobit, model. We forgo discussion of the very closely related truncated regression model (Hausman and Wise 1977).

**Tobit Model.** The censored regression, or tobit, model is appropriate when the dependent variable is censored at an upper or lower bound as an artifact of how the data are collected (Tobin



1958; Maddala 1983). For censoring at a lower bound, the model is

$$Y_{1i}^* = X_i\beta + \varepsilon_i \quad (1)$$

$$Y_{1i} = Y_{1i}^* \quad \text{if } Y_{1i}^* > 0 \quad (2)$$

$$Y_{1i} = 0 \quad \text{if } Y_{1i}^* \leq 0 \quad (3)$$

where for the  $i$ th observation,  $Y_{1i}^*$  is an unobserved continuous latent variable,  $Y_{1i}$  is the observed variable,  $X_i$  is a vector of values on the independent variables,  $\varepsilon_i$  is the error, and  $\beta$  is a vector of coefficients. We assume that  $\varepsilon_i$  is not correlated with  $X_i$  and is independently and identically distributed. The model can be generalized by replacing the threshold zero in equations (2) and (3) with a known nonzero constant. The censoring point also may vary across observations, leading to a model that is formally equivalent to models for survival analysis (Kalbfleisch and Prentice 1980).

**Standard Sample Selection Model.** A generalization of the tobit model involves specifying that a second variable  $Y_{2i}^*$  affects whether  $Y_{1i}$  is observed. That is, retain the basic model in equation (1) but replace equations (2) and (3) with

$$Y_{1i} = Y_{1i}^* \quad \text{if } Y_{2i}^* > 0 \quad (4)$$

$$Y_{1i} = 0 \quad \text{if } Y_{2i}^* \leq 0 \quad (5)$$

Variants of this model depend on how  $Y_{2i}$  is specified. Commonly,  $Y_{2i}^*$  is determined by a binary regression model:

$$Y_{2i}^* = Z_i\alpha + U_i \quad (6)$$

$$Y_{2i} = 1 \quad \text{if } Y_{2i}^* > 0 \quad (7)$$

$$Y_{2i} = 0 \quad \text{if } Y_{2i}^* \leq 0 \quad (8)$$

where  $Y_{2i}^*$  is a latent continuous variable. The classic example is a model for the wages and employment of women where  $Y_{1i}$  is the observed wage,  $Y_{2i}$  is a dummy variable indicating whether a woman works, and  $Y_{2i}^*$  indexes a woman's propensity to work (Gronau 1974). In a variant of this model,  $Y_{2i}$  is hours of work and equations (6) through (8) are a tobit model (Heckman 1974). In both variants,  $Y_{1i}^*$  is observed only for women with positive hours of work. One can modify the model by assuming,

for example, that  $Y_{1i}$  is dichotomous. If  $\varepsilon_i$  and  $v_i$  follow a bivariate normal distribution, this leads to a bivariate probit selection model.

Estimation of equation (1) using OLS will lead to biased estimates. When  $Y_{2i}^* > 0$ ,

$$\begin{aligned} Y_{1i} &= X_i\beta + E[\varepsilon_i | Y_{2i}^* > 0] + \eta_i \\ &= X_i\beta + E[\varepsilon_i | U_i - Z_i\alpha > 0] + \eta_i \end{aligned} \quad (9)$$

The OLS regression of  $Y_{1i}$  on  $X_i$  is biased and inconsistent if  $\varepsilon_i$  is correlated with  $v_i - Z_i\alpha$ , which occurs if  $\varepsilon_i$  is correlated with  $v_i$  or  $Z_i$  or both. If the variables in  $Z_i$  are included in  $X_i$ ,  $\varepsilon_i$  and  $Z_i$  are not correlated by assumption. If, however,  $Z_i$  contains additional variables,  $\varepsilon_i$  and  $Z_i$  may be correlated. When  $\sigma_{\varepsilon v} = 0$ , selection depends only on the observed variables in  $Z_i$ , not those in  $X_i$ . In this case, selection can be dealt with either by conditioning on the additional  $Z$ 's or by using propensity score methods (Rosenbaum and Rubin 1983).

Equation (9) shows how selectivity bias may be interpreted as an omitted variable bias (Heckman 1979). The term  $E[\varepsilon_i | Y_{2i}^* > 0]$  can be thought of as an omitted variable that is correlated with  $X_i$  and affects  $Y_{1i}$ . Its omission leads to biased and inconsistent OLS estimates of  $\beta$ .

**Nonrandom Treatment Assignment.** A model intimately related to the standard selection model that is not formally presented here is used when individuals are assigned nonrandomly to some treatment in an experiment. In this case, there are essentially two selection problems. For individuals not receiving the treatment, information on what their outcomes would have been if they had received treatment is "missing." Similarly, for individuals receiving the treatment, we do not know what their outcomes would have been if they had not received the treatment. Heckman (1978) explicitly analyzes the relationship between the nonrandom assignment problem and selection. Winship and Morgan (1999) review the vast literature that has appeared on this question in the last two decades.

## ESTIMATORS

A large number of estimators have been proposed for selection models. Until recently, all these estimators made strong assumptions about the distribution of errors. Two general classes of meth-

ods—maximum likelihood and nonlinear least squares—typically assume bivariate normality of  $\varepsilon_i$  and  $v_i$ . The most popular method is that of Heckman (1979), known as the lambda method, which assumes only that  $v_i$  in equation (6) is normally distributed and  $E[\varepsilon_i|v_i]$  is linear.

For a number of years, there has been concern about the sensitivity of the Heckman estimator to these normality and linearity assumptions. Because maximum likelihood and nonlinear least squares make even stronger assumptions, they are typically more efficient but even less robust to violations of distributional assumptions. The main concern of the literature since the early 1980s has been the search for alternatives to the Heckman estimator that do not depend on normality and linearity assumptions.

**Heckman’s Estimator.** The Heckman estimator involves (1) estimating the selection model (equations [6] through [8]), (2) calculating the expected error,  $v_i = E[v_i|v_i > -Z_i\alpha]$ , for each observation using the estimated  $\alpha$ , and (3) using the estimated error as a regressor in equation (1). We can rewrite equation (9) as

$$Y_{1i} = X_i\beta + E(\varepsilon_i|U_i > -Z_i\alpha) + \eta_i \tag{10}$$

If  $\varepsilon_i$  and  $v_i$  are bivariate normal and  $\text{Var}(v_i) = 1$ , then  $E(\varepsilon_i|v_i) = \sigma_{\varepsilon v} v_i$  and

$$E(\varepsilon_i|U_i > -Z_i\alpha) = \sigma_{\varepsilon v} \phi(-Z_i\alpha) / [1 - \Phi(-Z_i\alpha)] = \sigma_{\varepsilon v} \lambda(-Z_i\alpha) \tag{11}$$

where  $\phi$  and  $\Phi$  are the standardized normal density and distribution functions, respectively. The ratio  $\lambda(-Z_i\alpha)$  is the inverse Mills’s ratio. Substituting equation (11) into equation (10), we get

$$Y_{1i} = X_i\beta + \sigma_{\varepsilon v} \lambda(-Z_i\alpha) + \eta_i \tag{12}$$

where  $\eta_i$  is not correlated with both  $X_i$  and  $\lambda(-Z_i\alpha)$ . Equation (12) can be estimated by OLS but is preferably estimated by weighted least squares since its error term is heteroskedastic (Heckman 1979).

The precision of the estimates in equation (12) is sensitive to the variance of  $\lambda$  and collinearity between  $X$  and  $\lambda$ . The variance of  $\lambda$  is determined by how effectively the probit equation at the first stage predicts who is selected into the sample. The better the equation predicts, the greater the vari-

ance of  $\lambda$  is and the more precise the estimates will be. Collinearity will be determined in part by the overlap in variables between  $X$  and  $Z$ . If  $X$  and  $Z$  are identical, the model is identified only because  $\lambda$  is nonlinear. Since it is seldom possible to justify the form of  $\lambda$  on substantive grounds, successful use of the method usually requires that at least one variable in  $Z$  not be included in  $X$ . Even in this case,  $X$  and  $\lambda(-Z_i\alpha)$  may be highly collinear, leading to imprecise estimates.

**Robustness of Heckman’s Estimator.** Because of the sensitivity of Heckman’s estimator to model specification, researchers have focused on the robustness of the estimator to violations of its several assumptions. Estimation of equations (6) through (8) as a probit model assumes that the errors  $v_i$  are homoskedastic. When this assumption is violated, the Heckman procedure yields inconsistent estimates, though procedures are available to correct for heteroskedasticity (Hurd 1979). The assumed bivariate normality of  $v_i$  and  $\varepsilon_i$  in the selection model is needed in two places. First, normality of  $v_i$  is needed for consistent estimation of  $\alpha$  in the probit model. Second, the normality assumption implies a particular nonlinear relationship for the effect of  $Z_{i\alpha}$  on  $Y_{2i}$  through  $\lambda$ . If the expectation of  $\varepsilon_i$  conditional on  $v_i$  is not linear and/or  $v_i$  is not normal,  $\lambda$  misspecifies the relationship between  $Z_{i\alpha}$  and  $Y_{2i}$  and the model may yield biased results.

Several studies have analytically investigated the bias in the single-equation (tobit) model when the error is not normally distributed. In a model with only an intercept—that is, a model for the mean of a censored distribution—when errors are not normally distributed, the normality assumption leads to substantial bias. This result holds even when the true distribution is close to normal (for example, the logistic) (Goldberger 1983). When the normality assumption is wrong, moreover, maximum likelihood estimates may be worse than estimates that simply use the observed sample mean. For samples that are 75 percent complete, bias from the normality assumption is minimal; in samples that are 50 percent complete, bias is substantial in the truncated case but not in the censored case; and in samples that are less than 50 percent complete, bias is substantial in almost all cases (Arabmazar and Schmidt 1982).

The fact that estimation of the mean is sensitive to distributional misspecification suggests that

the Heckman estimator may not be robust and raises the question of how often such problems arise in practice. In addition, even when normality holds, the Heckman estimator may not improve the mean square error of OLS estimates of slope coefficients in small samples (50 or less) (Stolzenberg and Relles 1990). This appears to parallel the standard result that when the effect of a variable is measured imprecisely, inclusion of the variable may enlarge the mean square error of the other parameters in the model (Leamer 1983).

No empirical work that the authors know of directly examines the sensitivity of Heckman's method for a standard selection model. Work by LaLonde (1986) using the nonrandom assignment treatment model suggests that in specific circumstances the Heckman method can inadequately adjust for unobserved differences between the treatment and control groups.

**Extensions of the Heckman Estimator.** There are two main issues in estimating equation (12). The first is correctly estimating the probability for each individual that he or she will be selected. As it has been formulated above, this means first correctly specifying both the linear function  $Z\alpha$  and second specifying the correct, typically nonlinear relationship between the probability of selection and  $Z\alpha$ . The second issue is the problem of what nonlinear function should be chosen for  $\lambda$ . When bivariate normality of errors holds,  $\lambda$  is the inverse Mills's ratio. When this assumption does not hold, inconsistent estimates may result. Moreover, since  $X_i$  and  $Z_i$  are often highly collinear, estimates of  $\beta$  in equation 12 may quite be sensitive to misspecification of  $\lambda$ .

The first problem is handled rather easily. In the situation where one has a very large sample and there are multiple individuals with the same  $Z$ , the simplest approach is to estimate the probability of selection nonparameterically by directly estimating the probability of being selected for individuals with each vector of  $Z$ 's from the observed frequencies. With smaller samples, kernel estimation methods are available. These methods also consist of estimating probabilities directly by grouping individuals with "similar"  $Z$ 's and directly calculating the probability of selection from weighted frequencies. Variants of this approach involve different definitions of similarity and/or weight (Hardle 1990). In both methods, the problem of

estimating how the probability of selection depends on  $Z$  is bypassed.

Semiparametric methods are also available. These methods are useful if their underlying assumptions are correct, since they generally produce more efficient estimates than does a fully nonparametric approach. These methods include Manski's maximum score method (1975), nonparametric maximum likelihood estimation (Cosslett 1983), weighted average derivatives (Stoker 1986; Powell et al. 1989), spline methods, and series approximations (Hardle 1990).

The problem in the second stage is to deal with the fact that one generally has no a priori knowledge of the correct functional form for  $\lambda$ . Since  $\lambda$  is simply a monotonic function of the probability of being selected, this is equivalent to asking what nonlinear transformation of the selection probability should be entered into equation (12). A variety of approaches are available here. One approach is to approximate  $\lambda$  through a series expansion (Newey 1990; Lee 1982) or by means of step functions (Cosslett 1991). An alternative is to control for  $\lambda$  by using differencing or fixed effect methods (Heckman et al. 1998). The essential idea is to control for the probability of selection by implicitly including a series of dummy variables in equation (1), with each dummy variable being used to indicate a set of individuals with the same probability of selection or, equivalently, the same  $\lambda$ . Generally, this will produce significantly larger standard errors of the slope estimates. This is appropriate, however, since selection increases uncertainty. With small samples, these methods can be generalized through kernel estimation (Powell 1987; Ahn and Powell 1990). Newey et al. (1990) apply a variety of methods to an empirical problem.

## CONCLUSION

Selection problems bedevil much social science research. First and foremost, it is important for investigators to recognize that there is a selection problem and that it is likely to affect their estimates. Unfortunately, there is no panacea for selection bias. Various estimators have been proposed, and it is important for researchers to investigate the range of estimates produced by different methods. In most cases this range will be considerably broader than the confidence intervals for the

OLS estimates. Selection bias introduces greater uncertainty into estimates. New methods for correcting for selection bias have been proposed that may provide a more powerful means for adjusting for selection. This needs to be determined by future research.

REFERENCES

- Ahn, H., and J. J. Powell 1990 *Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism*. Madison: Department of Economics, University of Wisconsin.
- Arabmazar, A., and P. Schmidt 1982 "An Investigation of the Robustness of the Tobit Estimator to Non-Normality." *Econometrica* 50:1055-1063.
- Berk, R. A. 1983 "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48:386-398.
- Cosslett, S. R. 1983 "Distribution-Free Maximum Likelihood Estimator of the Binary Choice Model." *Econometrica* 51:765-781.
- 1991 "Semiparametric Estimation of a Regression Model with Sampling Selectivity." In W. A. Barnett, J. Powell, and G. Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge: Cambridge University Press.
- DuMouchel, W. H., and G. J. Duncan 1983. "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples." *Journal of the American Statistical Association* 78:535-543.
- England, P., G. Farkas, B. Kilbourne, and T. Dou 1988 "Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects." *American Sociological Review* 53:544-558.
- Fligstein, N. D., and W. Wolf 1978 "Sex Similarities in Occupational Status Attainment: Are the Results Due to the Restriction of the Sample to Employed Women?" *Social Science Research* 7:197-212.
- Goldberger, A. S. 1983 "Abnormal Selection Bias." In S. Karlin, T. Amemiya, and L. A. Goodman, eds., *Studies in Econometrics, Time Series, and Multivariate Statistics*. New York: Academic Press.
- Gronau, R. 1974 "Wage Comparisons—Selectivity Bias." *Journal of Political Economy* 82:1119-1143.
- Hagan, J. 1990 "The Gender Stratification of Income Inequality among Lawyers." *Social Forces* 68:835-855.
- , and P. Parker, 1985. "White-Collar Crime and Punishment." *American Sociological Review* 50:302-316.
- Hardle, W. 1990 *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Hardy, M. A. 1989 "Estimating Selection Effects in Occupational Mobility in a 19th-Century City." *American Sociological Review* 54:834-843.
- Hausman, J. A., and D. A. Wise, 1977 "Social Experimentation, Truncated Distributions, and Efficient Estimation." *Econometrica* 45:919-938.
- Heckman, J. J. 1974 "Shadow Prices, Market Wages and Labor Supply." *Econometrica* 42:679-694.
- 1978 "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46:931-959.
- 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153-161.
- , H. Ichimura, S. Smith, and P. Todd 1998 "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66:1017-1099.
- Hurd, M. 1979 "Estimation in Truncated Samples." *Journal of Econometrics* 11:247-58.
- Kalbfleisch, J. D., and R. L. Prentice 1980 *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Lalonde, R. J. 1986 "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review* 76:604-620.
- Leamer, E. E. 1983 "Model Choice and Specification Analysis." In Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics, vol. 1*. Amsterdam: North-Holland.
- Lee, L. F. 1982 "Some Approaches to the Correction of Selectivity Bias." *Review of Economic Studies* 49:355-372.
- Maddala, G. S. 1983 *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- Manski, C. F., and D. A. Wise 1983 *College Choice in America*. Cambridge, Mass.: Harvard University Press.
- Mare, R. D., and C. Winship 1984 "The Paradox of Lessening Racial Inequality and Joblessness among Black Youth: Enrollment, Enlistment, and Employment, 1964-1981." *American Sociological Review* 49:39-55.
- Newey, W. K. 1990 "Two-Step Series Estimation of Sample Selection Models." Paper presented at the 1988 European meeting of the Econometric Society.
- , J. L. Powell, and J. R. Walker 1990 "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review* 80:324-328.
- Peterson, R. and J. Hagan 1984 "Changing Conceptions of Race: Towards an Account of Anomalous Findings of Sentencing Research." *American Sociological Review* 49:56-70.
- Powell, J. L. 1987 *Semiparametric Estimation of Bivariate Latent Variable Models*. Working Paper No. 8704.

- Madison: Social Systems Research Institute, University of Wisconsin.
- , J. H. Stock, and T. M. Stoker 1989 “Semiparametric Estimation of Index Coefficients.” *Econometrica* 57:1403–1430.
- Rosenbaum, P., and D. B. Rubin 1983 “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70:41–55.
- Stoker, T. M. 1986 “Consistent Estimates of Scaled Coefficients.” *Econometrica* 54:1461–1481.
- Stolzenberg, R. M., and D. A. Relles 1990 “Theory Testing in a World of Constrained Research Design: The Significance of Heckman’s Censored Sampling Bias Correction for Nonexperimental Research.” *Sociological Methods Research* 18:395–415.
- Tobin, J. 1958 “Estimation of Relationships for Limited Dependent Variables.” *Econometrica* 26:24–36.
- Winship, C., and R. D. Mare 1992 “Models for Sample Selection Bias.” *Annual Review of Sociology* 18:327–350.
- , and S. Morgan 1999 “The Estimation of Causal Effects from Observational Data.” *Annual Review of Sociology*. 25:657–704.
- Zatz, M. S., and J. Hagan 1985 “Crime, Time, and Punishment: An Exploration of Selection Bias in Sentencing Research.” *Journal of Quantitative Criminology* 1:103–126.

CHRISTOPHER WINSHIP  
ROBERT D. MARE

## SAMPLING PROCEDURES

The analysis of data from samples constitutes a major proportion of contemporary research in the social sciences. For example, researchers use sample data from the U.S. population to estimate, with specified levels of confidence and precision, quantities such as average household size, the proportion of Americans who are unemployed during a given month, and the correlation between educational attainment and annual earnings among members of the labor force. Sample-based estimates are called *sample statistics*, while the corresponding population values are called *population parameters*. The most common reason for sampling is to obtain information about population parameters more cheaply and quickly than would be possible by using a complete census of a population. Sampling also is used sometimes when it is not feasible to carry out a complete census. For example,

except perhaps in a few nations with population registers, attempts to carry out a census in large countries invariably fail to enumerate everyone, and those who are missed tend to differ systematically from those who are enumerated (Choldin 1994). In these cases sampling may be the only way to obtain accurate information about population characteristics.

Researchers can use sample statistics to make inferences about population parameters because the laws of probability show that under specified conditions, sample statistics are unbiased estimators of population parameters. For example, if one used the same procedure to draw repeated samples from a population and determine the proportion of females in each sample, the average value of the observed sample proportions would equal the actual proportion of females in the population. The laws of probability also show that one can use data from a single sample to estimate how much a sample statistic will vary across many samples drawn from the population. Knowing the variability of a sample statistic (in statistical parlance, its *sampling variance*) in turn makes it possible to draw conclusions about the corresponding population parameter even though one has data for only a single sample. The key condition that must be met for these conditions to hold is that the data must come from a *probability sample*, a sample in which each case in the population (cases may be individuals, households, cities, days, etc.) has a known and nonzero probability of being selected into the sample. To produce this type of sample, a sampling technique must employ a random selection mechanism, one in which only the laws of chance determine which cases are included in the sample.

Researchers sometimes use *nonprobability samples*, such as convenience samples, quota samples, and snowball samples, to avoid the costs in time and money of probability sampling. *Convenience samples*, which consist of cases chosen simply because they are readily available (such as pedestrians passing by a street corner and volunteers from a classroom), sometimes are used in exploratory research and in the development of questionnaires or interview protocols. Opinion polls sometimes use *quota samples*, in which interviewers are assigned certain types of people to interview (e.g., a white female over age 50 living on a farm or a black male aged 20 to 30 living in a central city) but

are free to select specific individuals fitting these criteria to interview, to gauge political opinions cheaply. *Snowball samples* are generated when investigators ask known members of a group to identify other members and subsequently ask those who are so identified to identify still other members. This procedure takes advantage of the fact that the members of certain small and hard to locate groups (for example, biochemists studying a particular enzyme) often know one another. None of these nonprobability sampling procedures ensures that all the cases in the target population have a known and nonzero probability of being included in a sample, however (Kalton 1983, pp. 90–93). As a result, one cannot be confident that these procedures will provide unbiased estimates of population parameters or make statistical inferences from the samples they yield. Unless cost and time constraints are severe, researchers seeking to estimate population parameters therefore nearly always use procedures that yield probability samples.

One should distinguish between the representativeness of a sample and whether it was drawn by using probability sampling procedures. Although probability samples have a decidedly better track record in regard to representativeness, not all probability samples are representative and not all nonprobability samples are unrepresentative. Political polls based on quota samples, for example, often produce results that come very close to the subsequent vote. However, there is usually no reason to believe that a nonprobability sampling procedure that has been successful in the past will continue to yield representative results. In contrast, probability sampling procedures are likely to produce representative samples in the future because they are based on a random selection procedure.

Sampling theory, a branch of statistical theory, covers a variety of techniques for drawing probability samples. Many considerations can influence the choice of a sampling procedure for a given project, including feasibility, time constraints, characteristics of the population to be studied, desired accuracy, and cost. Simple sampling procedures are often sufficient for studying small, accessible, and relatively homogeneous populations, but researchers typically must use more complicated procedures to study large and heterogeneous populations. Using complicated procedures requires

consultation with a sampling specialist at a survey organization (the University of Illinois Survey Research Laboratory provides a list of these organizations on its Web site: [www.srl.uic.edu](http://www.srl.uic.edu)).

Any study in which a probability sample will be drawn must begin by defining the population of interest: the target population. The purpose of the study restricts the definition of the target population but rarely specifies it completely. For example, a study of characteristics of U.S. families obviously will define the population as consisting of families, but it will be necessary to define precisely what counts as a family as well as decide how to treat various cases from which it may be difficult to collect data (such as the families of U.S. citizens who live overseas). Sudman (1976, pp. 11–14) discusses general issues involved in defining target populations.

The next step in probability sampling is to construct a *sampling frame* that identifies and locates the cases in the target population so that they can be sampled. The most basic type of sampling frame is a list of the cases in the target population. Such lists are often unavailable, however, and so researchers usually must construct an alternative. For example, to draw a sample of U.S. public high schools, a researcher might begin with a list of U.S. census tracts, select a sample of those tracts, and then consult maps that indicate the locations of public high schools in the selected tracts. Here the sampling frame would consist of the list of census tracts and their corresponding maps.

A perfect sampling frame includes all the cases in the target population, no inappropriate cases, and no duplications. Most sampling frames are imperfect, however, with failure to include all the cases in the target population being the most serious type of *coverage error*. For example, telephone-number sampling frames, such as those employed in random-digit dialing procedures, do not cover people without a telephone, and sampling frames that are based on dwelling units do not cover homeless people. Undercoverage errors bias sample statistics, with the extent of the bias being positively related to (1) the proportion of the target population not covered by the sampling frame and (2) the magnitude of the difference between those covered and those not covered. Sampling experts have developed many methods to reduce coverage errors, including the use of

multiple frames, multiplicity techniques, and postsurvey adjustments. Kish (1995, pp. 53–59, 384–439) and Groves (1989, pp. 81–132) provide helpful discussions of sampling-frame problems and possible solutions.

### BASIC PROBABILITY SAMPLING PROCEDURES

Characteristics of one's sampling frame influence the specific sampling procedure appropriate for producing a probability sample. For example, some sampling procedures require that the sampling frame list all the cases in the population, while others do not. In addition, sampling procedures often are combined in situations where the sampling frame is complex. In all situations, however, the key element required for producing a probability sample is the use of a formally random procedure for selecting cases into the sample.

**Simple Random Sampling.** Simple random sampling (SRS) is the most elementary probability sampling procedure and serves as a benchmark for the evaluation of other procedures. To use SRS, one's sampling frame must list all the cases in the population. Usually the researcher assigns a unique identification number to each entry in the list and then generates random numbers by using a random number table or a computer program that produces random numbers. If a random number matches one of the identification numbers in the list, the researcher adds the indicated case to the sample (unless it has already been selected). This procedure is followed until it produces the desired sample size. It is important that only the randomly generated numbers determine the sample's composition; this condition ensures that the sampling procedure will be unbiased and that the chosen cases will constitute a probability sample.

With SRS, all cases in the sampling frame have an equal chance of being selected into the sample. In addition, for a sample of size  $n$ , all possible combinations of  $n$  different cases in the sampling frame have an equal chance of constituting the sample. The formulas for standard errors found in nearly all statistics textbooks and those used in statistical programs for computers assume that SRS generated the sample data. Most studies of human populations use sampling procedures that are less efficient than SRS, however, and using SRS

formulas in these instances underestimates the sampling variances of the statistics. As a consequence, researchers frequently conclude that differences or effects are statistically significant when they should not do so, or they may report misleadingly small confidence intervals.

**Systematic-Simple Random Sampling.** When a sampling frame contains many cases or the size of the prospective sample is large, researchers often decide to economize by setting a sampling interval and, after a random start, using that interval to choose the cases for the sample. For example, suppose a researcher wanted to select a sample of  $n$  cases from a population of size  $N$  and  $n/N = 1/25$ . To use systematic simple random sampling (SSRS), the researcher would draw a random number,  $r$ , between 1 and 25 and, starting with the  $r$ th case, select every twenty-fifth case in the sampling frame (for more complicated examples, see Kalton 1983, p. 17). This procedure gives all the cases in the frame an equal probability of being chosen for the sample but, unlike SRS, does not give all combinations of cases equal probabilities of selection. In the above example there are only 25 possible combinations of cases that could constitute the resulting sample (for example, cases 105 and 106 could never be in the same sample).

When the order of the cases in the sampling frame is random with respect to the variables of interest in a study, this property of SSRS is inconsequential, but when the frame is cyclically ordered, the results of SSRS can differ significantly from those of SRS. For example, suppose one wished to sample starting players on college basketball teams to determine their average height and had a sampling frame ordered by team and, within each team, by position. Since there are five starting players on each team, a sampling interval of any multiple of 5 would yield a sample composed of players who all play the same position. There would be a 1 in 5 chance that these players would all be centers (usually the tallest players) and a 2 in 5 chance that they would all be guards (usually the shortest). Thus, in this instance the sampling variation of the players' mean height would be substantially greater than the variation that SRS would produce. However, there are also situations in which stratified random sampling SSRS is equivalent to (StRS) (see below) and yields samples that have smaller sampling variances than those from SRS (Kish 1995, pp. 113–23). In prac-

tice, most lists of entire populations have orderings, often alphabetical, that are essentially random with respect to the purposes of a study, and lists with potential problems usually are obvious or are quickly recognized. Thus, in most applications SSRS is essentially equivalent to SRS (Sudman 1976, pp. 56–57).

**Stratified Random Sampling.** When a sampling frame consists of a list of all the cases in a population and also contains additional information about each case, researchers may use StRS. For example, a list of people also might indicate the sex of each person. A researcher can take advantage of this additional information by grouping individuals of each sex into a sublist (called a *stratum*) and then sampling, using SRS or SSRS, from each stratum. One can use either the same sampling fraction for each stratum, in which case the procedure is called *proportionate* StRS, or different fractions for different strata (*disproportionate* StRS). In either case one usually attempts to use the additional information contained in the sampling frame to produce a sample that will be more efficient than one derived from other sampling procedures (i.e., it will need fewer cases to produce a sample with a given precision for estimating a population parameter).

Efficiency is commonly measured by a sampling procedure's *design effect*, the ratio of the sampling variance of a statistic based on that procedure to the sampling variance of the same statistic derived from an SRS with the same number of cases (Kalton 1983, pp. 21–24). The efficiency of proportionate StRS is directly related to the correlation between the variable used to stratify the sampling frame and the variable or variables being studied. Thus, if one wished to determine the mean individual income of a population of Americans, proportionate StRS based on sex would produce a more efficient sample than would SRS and would have a design effect smaller than unity, because sex is correlated with income. In the limiting case in which the stratifying variable is perfectly correlated with the variable or variables being studied—for example, if each woman earned \$15,000 per year and each man earned \$25,000—proportionate StRS would always yield a sample mean exactly equal to the population mean. By contrast, if sex were completely uncorrelated with income, proportionate StRS would be no more efficient than SRS, and the design effect of StRS

would equal unity. In practice it is usually difficult to obtain sampling frames that contain information about potential stratifying variables that are substantially correlated with the variables being studied, especially when the cases are individuals. As a result, the gains in efficiency produced by proportionate StRS are often modest.

Proportionate StRS often yields small sample sizes for strata that consist of small proportions of a population. Thus, when researchers want to estimate parameters for the individual strata in a population, they sometimes employ disproportionate StRS to ensure that there will be enough cases from each stratum in the overall sample. A second reason for using disproportionate StRS is to design an optimal sample, one that produces the most precise estimates for a given cost, when there are differences between the strata in terms of (1) the cost of sampling and obtaining data, (2) the variability of the variables under study, or (3) prior knowledge about the variables under study. Sudman (1976, pp. 107–130) discusses and gives examples of each of these situations. The benefits of disproportionate StRS may be hard to attain when one wants to draw a multipurpose sample with observations on many variables, however, because the optimal procedures for the different variables may conflict. In addition, although proportionate StRS cannot have a design effect greater than unity, the design effects for disproportionate StRS can be larger than unity, meaning that disproportionate StRS can produce samples that are less efficient than those derived from SRS (Kalton 1983, pp. 20–26).

**Cluster Sampling.** All the sampling procedures discussed above require that the researcher have a sampling frame that lists the cases in the target population. Unfortunately, such sampling frames rarely exist, especially for human populations defined by area of residence. One can still draw a probability sample, however, if the population can be organized in terms of a grouping principle and each case can be assigned to one of the groups (called *clusters*). For example, dwellings in cities are located in blocks defined by streets. Even if a list of dwellings does not exist, it is possible to draw a probability sample by constructing a sampling frame that consists of a listing of the blocks, drawing a random sample of the blocks, and then collecting data on the dwellings in the chosen blocks.



This procedure, which is called cluster sampling (CS), is also advantageous when one wishes to use face-to-face interviewing to survey geographically dispersed populations of individuals. In this case CS is less costly because it allows the survey to concentrate interviewers in a small number of locations, thus lowering traveling costs. However, CS usually produces samples that have larger sampling variances than those drawn from SRS. The efficiency of CS is inversely related to (1) the extent to which clusters are internally homogeneous and differ from each other and (2) the number of cases sampled from each cluster. CS is maximally efficient when a population can be divided into clusters that are identical, because each cluster will then be a microcosm of the population as a whole. When clusters are internally homogeneous and differ sharply from each other, as tends to be true for human populations clustered by area of residence, CS is considerably less efficient than SRS (Kalton 1983, pp. 30–33). In this situation, researchers usually attempt to select only a few cases from each of many clusters, but that strategy eliminates the cost savings of CS.

**Multistage Sampling.** Researchers who want to collect data through face-to-face interviews with a probability sample of people living in a certain area, such as the United States, a state, or even a city, usually combine elements of the procedures discussed above in a multistage sampling procedure. For example, to draw a probability sample of U.S. adults, one might begin by obtaining a list of counties and parishes in the United States and collecting data on several characteristics of those units (region, average household income, etc.). These variables can be used to group the units, called *primary sampling units*, into strata so that one can use StRS. In addition, one would obtain estimates of the number of residents in each unit so that they could be sampled with probabilities proportional to their estimated population sizes (Sudman 1976, pp. 134–50). After selecting a sample of counties in this fashion, the researcher might proceed to draw a series of nested cluster samples. For example, one could divide each selected county into subareas (perhaps townships or other area-based governmental divisions) and then select a cluster sample from these units, with probabilities once again proportional to estimated population size. Next the researcher might divide each of the selected units into subareas (perhaps on the order

of the U.S. Bureau of the Census's "blocks") and draw a cluster sample of them. For each chosen block, the researcher might obtain a list of dwelling units and draw another cluster sample. Finally, from each chosen dwelling unit the researcher would choose, according to a specified procedure (Kish 1995, pp. 396–404), an individual to be interviewed. It is crucial that the selection procedure at each stage of the sampling process be based on a formally random selection procedure. For more detailed discussions and examples of the selection of multistage sampling procedures, see Kish (1995, pp. 301–383), Moser and Kalton (1972, pp. 188–210), and Sudman (1976, pp. 131–170). Multistage sampling usually requires considerable resources and expertise, and those who wish to draw such samples should contact a survey organization. Studies of the *design effects* of multistage samples, such as those carried out by the University of Michigan's Survey Research Center, show that they usually vary from 1.0 to 2.0, with values around 1.5 being common (Kish 1995, p. 581). A design effect of 1.5 means that the standard error of a statistic is twenty-two percent larger than estimated by standard statistics programs, which assume simple random sampling. There is also variation across kinds of statistics, with univariate statistics, such as the mean, often having larger design effects than do bivariate statistics, such as regression coefficients (Groves 1989, pp. 291–292). Unfortunately, estimating standard errors for a multistage sample is usually a complicated task, and this complexity, combined with the fact that popular statistics programs for computers use only SRS formulas, has led most researchers to ignore the problem, producing many spurious "statistically significant" findings.

## RECENT ADVANCES

Sampling practitioners have made considerable progress in developing techniques for drawing probability samples of rare or elusive populations for which there are no lists and for which conventional multistage sampling procedures would produce sufficient cases only at an exorbitant cost. Sudman et al. (1988) review procedures for screening clusters to determine those that contain concentrations of a rare population's members and also discuss how multiplicity sampling procedures and capture-recapture methods can be applied to this problem. Researchers also have begun to use

multiplicity sampling of individuals to draw probability samples of businesses and other social organizations to which individuals belong; Sudman et al. (1988) outline the general strategy involved, and Parcel et al. (1991) provide an informative discussion and an example. This approach also can produce "linked micro-macro samples" that facilitate contextual analyses.

Recent developments in statistical theory and computer software promise to make the calculation of standard errors for statistics based on multistage samples much easier. One approach to overcoming these difficulties is to use a computer program to draw many subsamples from an existing sample and then derive an overall estimate of a standard error from the many estimates given by the subsamples. There are several versions of this general approach, including "bootstrapping," "jackknife replication," and "cross-validation" (Hinkley 1983). A second approach is to develop computer statistical packages that incorporate information about the sampling design of a study (Wolter 1985, pp. 393–412, contains a list of such programs). The increased availability of such programs should produce greater recognition of the need to take a study's sampling procedure into account in analyzing the data the study yields.

There is now greater recognition that sampling error is just one of many types of error to which studies of human populations are subject. Nonsampling errors, including nonresponse error, interviewer error, and measurement error, also affect the accuracy of surveys. Groves (1989) comprehensively discusses both sampling and nonsampling errors and argues that minimizing one type can increase the other. Thus, decisions about sampling procedures need to take into account likely sources and magnitudes of nonsampling errors.

#### REFERENCES

- Choldin, Harvey M. 1994 *Looking for the Last Percent: The Controversy over Census Undercounts*. New Brunswick, N.J.: Rutgers University Press.
- Groves, Robert M. 1989 *Survey Errors and Survey Costs*. New York: Wiley.
- Hinkley, David 1983 "Jackknife Methods." In Samuel Kotz and Norman L. Johnson, eds., *Encyclopedia of Statistical Sciences*, vol. 4. New York: Wiley.
- Kalton, Graham 1983 *Introduction to Survey Sampling*. Newbury Park, Calif.: Sage.
- Kish, Leslie 1995 *Survey Sampling*. New York: Wiley.
- Moser, Claus A., and Graham Kalton 1972 *Survey Methods in Social Investigation*, 2nd ed. New York: Basic.
- Parcel, Toby L., Robert L. Kaufman, and Leanne Jolly 1991 "Going Up the Ladder: Multiplicity Sampling to Create Linked Macro-to-Micro Organizational Samples." In Peter V. Marsden, ed., *Sociological Methodology*, vol. 21. Oxford, UK: Basil Blackwell.
- Sudman, Seymour 1976 *Applied Sampling*. New York: Academic Press.
- , Monroe G. Sirken, and Charles D. Cowan 1988 "Sampling Rare and Elusive Populations." *Science* 240:991–996.
- Wolter, Kirk M 1985 *Introduction to Variance Estimation*. New York: Springer-Verlag.

LOWELL L. HARGENS

## SCALES AND SCORES

See Factor Analysis; Measurement.

## SCANDINAVIAN SOCIOLOGY

Scandinavian sociology emerged in its modern form as an academic discipline just after World War II, although its roots go back considerably further. In Helsinki, the sociologist, ethnologist, and philosopher Edvard A Westermarck (1862–1939) lectured on sociology in 1890; in Göteborg, Gustaf Fredrik Steffen (1864–1929) became a professor of economics and sociology in 1903.

In 1850, the Norwegian clergyman Eilert Sundt (1817–1875) published a study of Norwegian tramps and the lowest stratum of the rural population. Between 1850 and 1869, when he became a vicar, Sundt received state support for his demographic and sociological studies of Norwegian manners and customs, poverty, and living conditions. In demography he is remembered for "Sundt's law," which states that irregularities in the age distribution at a given time generate similar irregularities in the next generation (Ramsøy 1998).

Westermarck held chairs in applied philosophy until 1930, and between 1907 and 1930 he also had a chair in sociology in London. He belonged to the small group of leading European sociologists and philosophers in the early part of the century. His best-known works are studies of the

history of marriage (first volume published in 1891) and of the origin and development of moral ideas (1906–1908, 1932).

Between 1907 and 1913 the statistician A Gustav Sundbärg (1857–1914) directed “Emigration-sutredningen,” an official investigation of Swedish emigration, which was considered one of the most serious problems in Sweden at that time (Sundbärg 1913). The final report was presented in 1913, and between 1908 and 1912 twenty appendices by Sundbärg himself; Nils R Wohlin (1881–1948), secretary of the investigation; and others were published. The investigation contains a wealth of statistical information of great interest.

### EARLY SOCIOLOGY

In Scandinavia in the 1940s, sociology was confronted with an established discipline of demography, a reliable and accessible population registration system, a positivistic philosophy, and a reformist policy in need of empirical studies of societal problems. Until the late 1960s, Scandinavian sociologists engaged in empirical studies, mostly quantitative, of social inequality, social mobility and the educational system, work conditions, problems of physical planning and social epidemiology, alcohol problems, and delinquency. The works of Allardt in Helsinki (1965), Carlsson in Lund (1958), Rokkan (1921–1979) in Bergen (Rokkan and Lipset 1967), and Segerstedt (1908–1999) in Uppsala (1955) are representative of early mainstream sociology, and those of Aubert (1922–1988) in Oslo (1965) had a qualitative approach. There was a strong American influence from Scandinavians who had studied at American universities, from journals and textbooks, and from visiting Americans; Norway in particular received a series of Fulbright scholars. Scandinavian sociology became strongly empirical, technical, sociopsychological, and survey-oriented. There was less interest in functionalism, Talcott Parsons, and the classics than in survey analysis and social exchange theory.

In the late 1960s and throughout the 1970s, new sociological voices were heard in Scandinavia, perhaps more in Denmark and Sweden than in Finland and Norway. They can be characterized as aggressively political and Marxist, antipositivistic, antifunctionalistic, and antiquantitative. In general, the “new” sociology was more theoretical and sociophilosophical and less empirical than the main-

stream versions. Conflict theories, critical theory, symbolic interactionism, and the labeling perspective came to the fore, as did socioanthropological and hermeneutic methods.

### INSTITUTIONAL STRUCTURE

In Denmark, the German refugee Theodor Geiger (1891–1952) was appointed a professor of sociology at Aarhus University in 1938. In 1955, Kaare Svalastoga (1914–1997), a Norwegian historian with a sociological education from the University of Washington in Seattle, became professor of sociology at Copenhagen, where a graduate program in sociology was started in 1958. In the 1960s and 1970s, additional chairs were created at Copenhagen and Aalborg and in sociological subdisciplines at universities in Copenhagen, Aarhus, Aalborg, and Roskilde as well as at the Handelshøjskolen (School of Economics and Business Administration) in Copenhagen. Mostly as a result of problems at the two institutes at the University of Copenhagen, Danish sociology was in a state of crisis after the late 1960s. In 1986–1987 the government closed down the Copenhagen institutes, and by 1994 they had been reorganized into a new institute with two chairs. The institute offers (as did Aalborg after 1997) a three-year undergraduate program leading to a B.A., followed by a two-year master’s program. At present there are only eight chairs in sociology in Copenhagen, Aalborg, Roskilde, and Handelshøjskolen. Outside the universities, the most important institute in sociology is the Socialforskningsinstituttet (Institute for Social Research). With Henning Friis as director from its start in 1958 until 1979, the institute has carried out many social investigations using its own field organization, for instance, the 1976 and 1986 Danish Welfare Surveys and the report on the 1992 follow-up of the 1968 youth study (Hansen 1995).

In Finland, two of Westermarck’s ethnosociological students held chairs in sociology in Helsinki and Turku in the 1920s. In the mid-1940s, modern-type sociologists got chairs, such as the criminologist Veli Verkko (1893–1955) in Helsinki in 1946. Heikki Waris (1901–1989), professor of social policy from 1948 in Helsinki, also should be mentioned here. Sociology chairs were first established at Helsinki, Turku, Åbo Academy, and Tampere. At present, there are more than twenty

professors of sociology and its subdisciplines at eleven universities and colleges: eight chairs at Helsinki University, four at Tampere, and two each at Jyväskylä and Turku, plus one at Åbo Akademi, also in Turku. Undergraduate studies lead to a master's degree after five years. There also are important research institutes outside the universities. The National Research and Development Centre for Welfare and Health (STAKES), under the Ministry of Social and Health Care, was established in 1992. Within STAKES there is a Social Research Unit for Alcohol Studies that conducts research on four principal areas, including narcotics. The research directors are Hannu Uusitalo and Jussi Simpura. There also is the Finnish Foundation for Alcohol Studies, which was established in 1950, with Klaus Mäkelä as its leading researcher.

In Norway, Sverre Holm filled the first sociology chair in 1949 in Oslo, and chairs were added in Bergen, Trondheim, and Tromsø in the 1960s. At present there are some twenty-five professorships. A basic undergraduate program takes four years, after which a two-year program leads to a "cand.polit." degree. Several institutes are important in sociological research and education: Institutt for Samfunnsforskning (Institute for Social Research [ISF]) since 1950, the newly established NOVA (Norwegian Institute for Research on Childhood and Adolescence, Welfare, and Aging), FAFO (Trade Union Movement's Research Foundation), and Statistisk Sentralbyrå (Statistical Central Bureau [SSB]). Institutt for Anvendt Sosialforskning (Institute for Applied Social Research [INAS]), which was established in 1968, is now included in NOVA. Natalie R Ramsøy was its research director until 1981.

In Sweden, the educator and sociologist E. H. Thörnberg (1873–1961) never held a university position. Gunnar Myrdal (1898–1987) could claim to be the first modern sociologist, but Torgny T. Segerstedt (1908–1999) became the first professor of sociology by converting his chair in applied philosophy to sociology at Uppsala in 1947. Sociology is now taught at Uppsala, Stockholm, Lund, Göteborg, Umeå, and Linköping as well as in several colleges, of which those at Karlstad, Växjö, and Örebro have become universities. There are about thirty sociology professors at the universities. A bachelor's program in sociology is scheduled for three years, a master degree for a fourth year, and a "licentiat" degree for a fifth

year. Outside the universities, Arbetslivsinstitutet (the National Institute for Working Life); Folkhälsoinstitutet (the National Institute of Public Health), established in 1992; the National Council for Crime Prevention (BRÅ), starting in 1974; and Statistics Sweden are important in Swedish sociology. The Institute for Social Research ([SOFI] 1972, beginning in with five sociology chairs) and the National Center for Social Research on Alcohol and Drugs (since 1999), with Robin Room as research director, are parts of Stockholm University. Since 1992 BRÅ has published *Studies on Crime and Crime Prevention*, an international bianual journal.

There is considerable interaction among the Scandinavian sociological communities: Professors hold chairs in neighboring countries, such as the leading Finnish alcohol researcher Kettil Bruun (1924–1985) who had a chair in Stockholm in 1982–1984 and Swedish sociologists from Lund who crossed the sound to Copenhagen; comparative studies of the Scandinavian countries are conducted; and there have been joint comparative projects with one editor from each of four countries.

The Scandinavian Sociological Association has some 2,500 members. Approximate memberships are 550 in Denmark, 50 in Iceland, 600 in Finland (in 1995), 700 in Norway, and 500 in Sweden. Since 1955 the association has published *Acta Sociologica*, a refereed quarterly journal. For the first twelve years Torben Agersnap of Handelshøjskolen in Copenhagen was the editor, and since then editorship has rotated among the countries, with three-year periods since 1985. Each Scandinavian country has a national sociological association. The Finnish Westermarck Society, founded in 1940, is the oldest and includes social anthropologists. There are also national journals: *Sociologia* (since 1990) in Denmark, *Sosiologia* in Finland, *Tidskrift for Samfunnsforskning* (since 1960) and *Sociologisk Tidskrift* in Norway, and *Sociologisk forskning* (since 1964) in Sweden.

Most foreign contacts are still with the United States, although interaction with European, especially British, French, and German, sociologists tends to be more frequent now. Polish, Hungarian, and Estonian contacts are important to Finnish sociology. Comparative studies including non-Scandinavian European or OECD countries (Erikson and Goldthorpe 1992; Korpi and Palme 1998)

are conducted frequently as a heritage from Stein Rokkan in political sociology. Comparisons often have been made with the United States and Canada, as occurred when Norway and Sweden were included in Tim Smeeding's and Lee Rainwater's so-called "Luxembourg Income Study" around 1980. Furthermore, several Scandinavian sociologists have been visiting researchers at the European Sociological Institute in Florence, Italy. The Danish welfare researcher Esping-Andersen (1990) has a chair at the University of Trent in Italy. Several Scandinavian sociologists have spent most or parts of their careers abroad, especially in the United States, for example, Aage Bøttger Sørensen from Denmark; the Norwegians Stein Rokkan, Jon Elster, and Trond Pedersen; and the Swedes Bo Andersson and Hans L. Zetterberg.

### PRESENT SOCIOLOGY

Essays on Scandinavian sociology can be found in Bertilsson and Therborn (1997). New conceptions in Finnish sociology in the early 1990s and their relation to earlier structural-cultural traditions are discussed by Alapuro (1995), while Martinussen (1993) deals with present-day Norwegian sociology. Allardt et al. (1988) offer an official evaluation of Swedish sociology. Gunnlaugsson and Bjarnason (1994) describe Icelandic sociology.

Even after the wave of New Left sociology subsided toward the end of the 1970s, Scandinavian sociology remained diversified, with a continued interest in Marxism and the classics as well as in social exchange theory and a broad spectrum of data analysis. On balance, the focus now is somewhat more on theory and on macrosociology. Interest in theories of organizations (Ahrne 1994), economic sociology (Swedberg 1990), and, especially after Coleman's book (1990), rational-choice approaches (Hedström and Swedberg 1998) has been increasing. The statistical analytic orientation has held its own, at least partly because of the rapid development of personal computers and their programs, which permit more adequate analyses. At the same time, the old interest in qualitative methods has remained, especially in connection with the growing field of gender studies.

In Norway, Østerberg (1986, 1988) writes in the hermeneutic and humanistic essay tradition, Yngvar Løchen (1931–1998) carried out action research in medical sociology with relevance to the

lives of sociologists, and Elster is a well-known philosopher and social scientist (1989), favorably disposed to Marxism (1985) but closer to the rational-choice approach. In Sweden, Göran Therborn is an internationally respected Marxist and the social psychologist and essayist Johan Asplund has a prestigious position in the discipline; Antti Eskola is Finland's best known social psychologist.

A cluster of overlapping core areas concerning forms of inequality has remained central to Scandinavian sociology. Studies concern gender inequality as well as inequality among social classes and other groupings in regard to political, economic, educational, and social resources, which can be seen as constituting inequality of welfare in a broad sense; these studies present indicators of various welfare dimensions. Research institutes such as the Institute for Social Research in Copenhagen, INAS (now in NOVA) and ISF in Oslo, and SOFI in Stockholm have been important, as STAKES in Helsinki will be. One of the tasks of SOFI has been to follow up an officially commissioned 1968 study of low-income categories. The study was carried out by Sten Johansson, then at Uppsala, and his team, which published several reports in 1970–1971, creating considerable political commotion. SOFI has continued this study as a panel, *Levnadsnivåundersökningen* (Level of Living Study [LNU]), with new surveys in 1974, 1981, and 1991. A volume edited by Erikson and Åberg (1987) provided a partial summary. Comprehensive welfare studies (level-of-living studies) have been carried out in all Scandinavian countries since the 1970s (Erikson et al. 1987; Hansen et al. 1993). Statistics Sweden has conducted annual level-of-living surveys since 1974 (Vogel 1988). Studies by Gudmund Hernes in Norway have been influential both politically and sociologically. Hernes was the leading researcher in the first Norwegian level-of-living survey and in an official Norwegian 1972–1982 project on power in society (Hernes 1975, 1978, also involving exchange theory). The task force gave its final report in 1982 (Hernes 1982). The official Swedish study of societal power, ordered after the success of the Norwegian study, was run mainly by political scientists, historians, and economists, although Åberg (1990) was asked to repeat in modern form a community study from the 1950s (Segerstedt and Lundquist 1952, 1955). Here the different roles of Norwegian and Swed-

ish sociology may reflect differences in public opinion: Norwegian sociologists have easy access to the mass media and the political elite. The criminologist Nils Christie has become a well-known participant in public discussion of social policy issues, and Hernes was the Labor Party's minister of education and research.

Social stratification and social mobility have long been an area of strong interest in Scandinavian sociology (Carlsson 1958; Svalastoga 1959, 1965). Later works in the field include those of Alestalo (1986), Erikson (1987), Knudsen (1988), and Erikson and Goldthorpe (1992). More recently Erikson and Jonsson reported an officially commissioned investigation of social selection to higher education (Erikson and Jonsson 1993, 1996).

In political sociology Stein Rokkan's (1921–1979) far-reaching project (Rokkan and Lipset 1967; Rokkan et al. 1970; Rokkan 1987) was left unfinished. Other important works in this field include Korpi (1989), Therborn (1986, 1995), Martinussen (1988), Alapuro (1988), and Mjøset (1992).

Gender studies are an exceptionally strong field in Norway and fairly strong in Scandinavia generally. In the late 1960s, Holter (1970) analyzed gender roles and their impact on work life, political behavior, and education. Several Norwegian gender studies, among others those by Helga Hernes and Kari Waerness, were published in the series "Kvinnerens levkår og livsløp" (Women's Level of Living and Life Course). In Sweden Edmund Dahlström and Rita Liljeström and in Finland Elina Haavio-Mannila pioneered the field.

Outside the sociological core area of inequality there are neighboring fields and sociological specialties. The Finnish demographer Tapani Valkonen is known for skillful context analyses and epidemiological studies. To revitalize Swedish demography, which had stagnated in the 1970s, the Norwegian demometrician Jan M. Hoem was called in from Copenhagen in 1983.

In the lively sociological subfield of literature and mass media, Karl-Erik Rosengren is the leading Swedish researcher. To some extent through east European contacts, the life-course narrative as a research instrument has been developed into a Finnish specialty, mostly by J.P. Roos.

Deviance is another subfield. Usually it is not considered a sociological core area, although it has long-term credentials as a central field both from the classics and from early Scandinavian sociology. The study of deviance has remained an important part of Scandinavian sociology in terms of both applied and basic research. Since the Finnish State Alcohol Monopoly established its research institute in 1950, alcohol studies have been a strong Finnish field. Bruun (Bruun and Frånberg 1985) was its longtime director, and Mäkelä (1996) was his successor. Also Skog (1985) in Oslo and Köhlhorn (Köhlhorn and Björ 1998) and Norström (1988) in Stockholm are well known in the field of alcohol and drug studies. Mathiesen's penology (1987) is a part of Norwegian sociology, just as studies of crime and crime prevention have been a part of Swedish sociology since the official 1956 juvenile delinquency project (Carlsson 1972) and early BRÅ projects.

Gunnlaugsson and Bjarnason (1994) claim that the four fields of welfare research, stratification research, women's studies, and cultural studies do not capture the structure of Icelandic sociology, which centers on two broad themes: social conditions and social and cultural problems associated with the development of Icelandic society. These problems essentially concern crime, in particular incidental violence by strangers, alcohol abuse, and the perceived threat of drug abuse, that is, deviance.

Finally, access to an extensive and reliable population registration system has made good sampling frames available to surveys, and governmental microdata have been helpful in longitudinal data sets in Scandinavian behavioral projects. However, since the mid-1970s, statistically oriented behavioral researchers, especially in Sweden, have had problems with privacy-protecting data legislation. According to the pioneering Swedish Data Act of 1973, running data on identified persons by computer requires a permit from a Data Inspection Board (DI). The DI expanded the informed-consent condition for the permit to include the computer use of identified governmental microdata accessed according to the century-old right-of-access principle. Although this condition has been waived in some cases, it has led to frequent and well-publicized controversies between the DI and social researchers, with the media usually supporting the DI. In Sweden, the media debated privacy

issues mostly in connection with statistical data sets, whereas administrative files largely went unnoticed. Since the conflict came to a head in 1986 over the deidentification of the data set of Project Metropolitan, a longitudinal study of a cohort of Stockholmers, the tensions have smoothed out and access to register data has been made easier. The rich governmental microfiles, including the censuses, remain assets to Scandinavian sociology. In 1998, a new Swedish Personal Data Protection Act in line with EU regulations and modern Internet use was substituted for the 1973 Data Act.

## REFERENCES

- Ahrne, Göran 1994 *Social Organizations*. Newbury Park CA: Sage.
- Alapuro, Risto 1995 "Structure and Culture: Finnish Sociology, 1990-94" *Acta Sociologica* 38(2):167-180.
- Alestalo, Matti 1986 *Structural Change, Classes, and the State*. Helsinki: Research Group for Comparative Sociology.
- Åberg, Rune, ed. 1990 *Industrisamhälle i omvandling* (Industrial Communities in Transition). Stockholm: Carlssons.
- Allardt, Erik 1965 *Samhällsstruktur och sociala spänningar* (Societal Structure and Social Tension). Tampere: Söderströms. Also in Finnish.
- , Sverre Lysgaard, and Aage Bøttger Sørensen (1988) *Sociologin i Sverige* (Sociology in Sweden). Stockholm: HSFR.
- Aubert, Vilhelm 1965 *The Hidden Society*. Totowa, N.J.: Bedminster.
- Bertilsson, Margareta, and Göran Therborn, eds. 1997 *From a Doll's House to the Welfare State*. Copenhagen: ISA.
- Bruun, Kjetil, and Per Frånberg, eds. 1985 *Den svenska supen* (The Swedish Snaps). Stockholm: Prisma.
- Carlsson, N. Gösta 1958 *Social Mobility and Class Structure*. Lund: Gleerup.
- 1972 *Unga lagöverträdare II* (Young Law Violators II). Stockholm: SOU 1972:76
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Belknap.
- Elster, Jon 1985 *Making Sense of Marx*. Cambridge, UK: Cambridge University Press.
- 1989 *Nuts and Bolts for the Social Sciences*. Cambridge, UK: Cambridge University Press.
- Erikson, Robert 1987 "The Long Arm of the Origin." In Ulla Bergryd and Carl-Gunnar Janson, eds., *Sociological Miscellany*. Stockholm: Department of Sociology, University of Stockholm.
- , and Rune Åberg, (eds.) 1987 *Welfare in Transition*. Oxford, UK: Clarendon Press.
- , and John Goldthorpe 1992 *The Constant Flux: A Study of Class Mobility in Industrial Society*. Oxford, UK: Clarendon.
- , Erik Jørgen Hansen, Stein Ringen, and Hannu Uusitalo, eds. 1987 *The Scandinavian Model: Welfare State and Welfare Research*. Armonk, N.Y.: Sharpe.
- , and Jan O. Jonsson 1993 *Ursprung och utbildning* (Social Origin and Education). Stockholm: SOU 1993:85.
- , and — (eds.) 1996 *Can Education Be Equalized?* Oxford, UK: Westview Press.
- Esping-Andersen, Gøsta 1990 *The Three Worlds of Welfare Capitalism*. Cambridge, UK: Cambridge University Press.
- Gunnlaugsson, Helgi, and Thoroddur Bjarnason 1994 "Establishing a Discipline: The Impact of Society on the Development of Icelandic Sociology." *Acta Sociologica* 1994(3):303-312.
- Hansen, Erik Jørgen 1995 *En generation blev voksen* (A Generation Grew Up) Copenhagen: Socialforskningsinstituttet, Report 95:8.
- , Stein Ringen, Hannu Uusitalo, and Robert Erikson, eds. 1993 *Welfare Trends in the Scandinavian Countries*. Armonk, N.Y.: Sharpe.
- Hedström, Peter and Richard Swedberg 1998 *Social Mechanisms*. Cambridge: Cambridge University Press.
- Hernes, Germund 1975 *Makt og avmakt* (Power and Powerlessness). Bergen, Norway: Universitetsforlaget.
- 1978 *Forhandlingsøkonomi og blandingsadministration* (Transaction Economy and Mixed Administration). Bergen, Norway: Universitetsforlaget.
- 1982 *Maktutredningen: slutrapport* (The Power Study: Final Report). Oslo: NOU 1982:3.
- , and Willy Martinussen 1980 *Demokrati og politiske resurser* (Democracy and Political Resources). Oslo: NOU 1980:7.
- Holter, Harriet 1970 *Sex Roles and Social Structure*. Oslo: Norwegian University Press.
- Knudsen, Knud 1988 "Class Identification in Norway." *Acta Sociologica* 31:69-79.
- Korpi, Walter 1989 "Power, Politics, and State Autonomy in the Development of Social Citizenship." *American Sociological Review* 54:309-328.
- , and Joakim Palme 1998 "The Paradox of Redistribution and Strategies of Equality." *American Sociological Review* 63:661-687.

- Kühlhorn, Eckart, and Jill Björ, eds. 1998 *Svenska alkoholvanor i förändring* (Changing Swedish Drinking Habits). Stockholm: Sober.
- Martinussen, Willy 1988 *Solidaritetsens grenser* (Limits of Solidarity). Oslo: Universitetsforlaget.
- 1993 “A Snapshot of Present-Day Norwegian Sociology,” *Acta Sociologia*, 36, 305–392.
- Mathiesen, Thomas 1987 *Kan fengsel forsvares?* (Can Prisons Be Justified?) Oslo: PAX; sec. ed. 1995.
- Mjøset, Lars ed. 1992 *Contributions to the Comparative Study of Development*. Oslo: ISF, Report 92:2.
- Norström, Thor 1988 “Alcohol and Suicide in Scandinavia.” *British Journal of Addiction* 83:553–559.
- Østerberg, Dag 1986 *Fortolkende sosiologi* (Interpretive Sociology). Oslo: Universitetsforlaget.
- 1988 *Metasociology*. Oslo: Norwegian University Press.
- Ramsøy, Natalie R. 1998 “Eilert Sundt.”
- Rokkan, Stein 1987 *Stat, nasjon, klasse* (State, Nation, Class). In B. Hagtvet, ed., Oslo: Universitetsforlaget.
- , Angus Campbell, Per Torsvik, and Henry Valen 1970 *Citizens, Elections, Parties*. Oslo: Norwegian University Press.
- , and Seymour M. Lipset, eds. 1967 *Party System and Voter Alignments*. New York: Free Press.
- Segerstedt, Torgny 1955 “Kriminalsociologi” (Sociology of Crime). In Ivar Agge, ed., *Kriminologi* (Criminology). Stockholm: Wahlström and Widstrand.
- , and Agne Lundqvist 1952 and 1955 *Människan i industrisamhället. del I och II* (Man in Industrial Society. Parts I and II). Stockholm: SNS.
- Skog, Ole-Jørgen 1985 “The Collectivity of Drinking Cultures.” *British Journal of Addiction* 81:365–379.
- Sundbärg, Gustav 1913 *Emigrationsutredningen* (Investigation on Emigration). Stockholm: Offentligt betänkande. Twenty appendices 1908–1912.
- Svalastoga, Kaare 1959 *Prestige, Class, and Mobility*. Copenhagen: Gyldendal.
- 1965 *Social Differentiation*. New York: McKay.
- Swedberg, Richard 1990 *Economics and Sociology*. Princeton: Princeton University Press.
- Therborn, Göran 1986 *Why Are Some People More Unemployed Than Others?* London: Verso.
- 1995 *European Modernity and Beyond*. Newbury Park, Calif.: Sage.
- Vogel, Joachim 1988 *Inequality in Sweden*. Stockholm: SCB.
- Westermarck, Edvard A. 1891 *History of Human Marriage*, 3 vols. London: Macmillan.

- 1906–1908 *The Origin and Development of Moral Ideas*. London: Macmillan.
- 1932 *Ethical Relativity*. New York: Harcourt; paperback ed. 1960 by Littlefield.

CARL-GUNNAR JANSON

## SCIENCE

Sociologists of science study the social organization of science, the relationships between science and other social institutions, social influences on the content of scientific knowledge, and public policy regarding science. The definition of the term “science” is problematic. Science can refer to a changing body of shared knowledge about nature or to the methods used to obtain that knowledge; in that form, science has existed for millennia. Research on “indigenous scientific knowledge” is reviewed in Watson-Verran and Turnbull (1995). Sociologists of science are more likely to define science in institutional terms, and most research in that area studies those who work in differentiated social institutions. The “demarcation” problem of distinguishing between science and nonscience persists. Gieryn (1995, 1998) argues that scientists and their advocates continually engage in contested “boundary work” to demarcate science. He discusses the rhetorical and organizational devices used in those contests; thus, scientists are likely to emphasize the disinterested search for knowledge in their attempts to distinguish science from technology and stress the utility of scientific knowledge in their attempts to distinguish it from religion. Gieryn argues against the notion that there are “essential” features of science that determine the outcome of those contests; these “essential features” are instead “provisional and contextual results of successful boundary-work” (1995, p. 406).

Unless the production of knowledge about the empirical world is delegated to relatively autonomous specialists, knowledge accumulates at a slow pace. When beliefs about nature are closely linked to major social institutions, institutional rigidity tends to produce cognitive rigidity (Znaniecki 1940; Parsons 1951, pp. 326–348). There were communities of relatively autonomous specialists in several great civilizations in the past, but most failed to produce stable institutions. Modern science dates from seventeenth-century Europe. Europe-



ans at that time believed in a deity that gave laws to nature as well as to people and expected to discover those laws. Seventeenth-century Europeans could build on the basis of a science produced by the medieval schoolmen. With the rise of capitalism, intellectual elites developed a strong interest in using new knowledge to improve material conditions and enrich themselves. Merton, the leading founder of the sociology of science, argued in 1938 (1970) that in addition to these conditions, Puritanism contributed to the scientific ethos through its emphasis on work in the everyday world, rationalism and empiricism, openness to free inquiry, and desire to glorify God by describing His creation. (This still-controversial thesis is reviewed in a symposium in *Isis* 1988. For a general review of theories about the scientific revolution, see Cohen 1994.)

A distinctive normative ethos was institutionalized in modern science. Merton (1973, chap. 13) identified four salient norms: (1) "Universalism" requires that scientific contributions be evaluated according to general impersonal criteria without regard to "irrelevant" characteristics of the contributors such as their race, religion, and nationality. It also requires that scientists be rewarded according to their scientific contributions without regard for those irrelevant criteria. (2) "Communism" requires that knowledge be shared, not kept secret. Thus, the only way a scientist can claim a discovery as "his" or "hers" is to make it known to others. In this regard, modern scientists differ from Renaissance mathematicians and magicians, who were often quite secretive. (3) "Disinterestedness" refers to the injunction that the procedures and products of science not be appropriated for private gain. This need not imply altruism, although scientists often are driven to discover as an end in itself, but in addition, situations usually are structured so that it is in a scientist's career interest to act in a disinterested manner. (4) "Organized skepticism" permits and encourages challenges to knowledge claims. Science tends to be unlike many other areas of social life, in which conformity in matters of belief is demanded as a sign of loyalty.

Merton's essay on the normative ethos of science, first published in 1942, has drawn fruitful criticism. While Merton argued that the scientific ethos was functional for the advancement of knowledge, Mitroff (1974) argued that scientists could invoke "counter-norms," for example, could fail

to be skeptical about their own theories, and this could be equally functional in some situations. Mulkay (1980) invoked ethnomethodological ideas to make an argument of general significance: "We should not assume that *any* norm can have a single literal meaning independent of the contexts in which it is applied. . . . Scientists *must* engage in inferential and interpretive work on norms. They are likely to do this *after* their actions, in order to construct acceptable accounts of their behavior. The norms don't determine behavior."

Ambiguity involving the norm of universalism was present at the birth of modern science: Which characteristics of those who advance knowledge claims are relevant or irrelevant? Shapin (1994) argues that in England only the testimony of "gentlemen" was accepted as valid, and not all gentlemen; those who rejected the empiricism of men such as Francis Bacon and Robert Boyle and accepted the arguments from first principles of men such as Thomas Hobbes were excluded from the scientific community (Shapin and Shaffer 1985).

Scientists in the seventeenth and eighteenth centuries were usually amateurs, such as Robert Boyle and Benjamin Franklin, or the intellectual servants of amateurs. In the later eighteenth and nineteenth centuries, science was professionalized. Scientists received formal education in universities; found full-time employment, often in the universities; formed self-governing associations; and developed the modern scientific journal and other means of communication. A case study of the process for the newly emerging discipline of geology in the 1830s is presented in Rudwick (1985), where it is linked with the conduct of intense disputes about geological history; the "professionals" in those disputes got no special respect. The more general process of professionalization is described by Ben-David (1984, 1991), who notes the importance of national differences in the organization of science. Ben-David shows that there was more competition among universities in Germany and the United States in the late nineteenth and twentieth centuries than there was in those in Britain and France and claims that this partly accounted for the greater productivity of science in the first two countries. Other organizational characteristics of American science also help account for its superior productivity in the past half century: Science is not highly centralized, the competitive units are large enough and heterogeneous

enough to provide a critical mass of closely interacting scientists, and senior scientists tend to have less authority over younger scientists than they have elsewhere. (For statistics on national science productivity, see U.S. National Science Board 1996.)

### SOCIAL STRATIFICATION IN SCIENCE

Competition remains intense among organizations that engage in basic research in the United States, particularly universities. Organizational prestige is central; as is usually true when it is difficult to measure organizational outputs directly, social comparisons become more important. Periodic surveys of faculty members have been used to rate the prestige of research departments. While outputs are difficult to measure, departments with high prestige are more successful in obtaining research resources and have higher rates of research productivity.

Competition is also intense among individual scientists, who compete for recognition from their peers for being the first to make valued discoveries (Merton 1973, chaps. 14 and 15; Hagstrom 1965, 1974). Competition may lead to secretive behavior and premature publication; it also may encourage scientists to move into new areas where there is less competition. A common consequence of competition is simultaneous or multiply independent discovery (see Zuckerman 1988 and the references cited there). The frequency of such events shows the extent to which science is a social product. When apparently simultaneous discoveries occur, those involved often engage in priority disputes; they are often ambivalent in those disputes, torn between a desire for the recognition due to originality and the demand for humility, the recognition of their dependence on the work of others.

There is a great degree of inequality in the research productivity of scientists. The chances that a scientist will publish as many as  $n$  papers is  $1/n^2$ ; in other words, about 6 percent of all scientists produce 50 percent of all published papers (Price 1986). This inequality is even greater if one looks at the distribution of citations of the work of scientists. With some reservations, the number of citations can be taken as a measure of the quality of scientific work; frequently cited papers are usually those which other scientists have found useful in their own research. If  $c$  is the number of citations, the chances that the work of a scientist will have  $c$

citations is proportional to  $1/c^3$ ; that is, about 3 percent of all scientists receive 50 percent of all the citations of scientific papers.

Most of the variation in scientific productivity can be explained in terms of individual characteristics of scientists, such as years required to earn a doctorate, and characteristics of their employers, especially departmental prestige. While men have been more productive than women (the difference has been declining), that difference is almost entirely the result of differences in background and employer characteristics (Xie and Shauman 1998). In the United States (more than in most countries), there is considerable mobility of scientists among organizations. High research productivity predicts mobility to institutions of higher prestige and to a higher rank, but employment in a high-prestige organization in turn causes higher productivity (Allison and Long 1990). In general, American universities tend to conform to universalistic norms in making appointments and promotions (Cole and Cole 1973). There is an apparent exception to this in the early phases of careers, when productivity is difficult to assess; the prestige of a scientist's doctoral department is strongly correlated with the prestige of the initial employer.

Inequality of productivity increases over the careers of scientists (Allison et al. 1982) as a manifestation of Merton's (1973) "Matthew effect": "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath." Initially productive scientists obtain more and better resources for research, their work is more visible to others, and they are more likely to interact with other highly productive scientists.

### WORK GROUPS, SPECIALTIES, AND DISCIPLINES

Scientific research is a nonroutine activity; outcomes and problems cannot be predicted, and it is difficult to plan research. As organization theories lead one to expect in such situations, scientific work tends to be done in small groups with few hierarchical levels and a small degree of control by supervisors (Hagstrom 1976). Most basic research in universities is done by groups of four to nine graduate students and technicians led by one to a few professors. Over the course of time, faculty

members have found it increasingly desirable to collaborate with their peers, and most publications are multiply authored. Some aspects of research can be routinized, and the extent to which this can be done varies among disciplines; for example, work is more readily routinized in chemistry than it is in mathematics (Hargens 1975). Thus, work groups are smaller in mathematics than in chemistry. Chemists can delegate tasks to assistants, whereas mathematicians cannot; while the number of assistants does not explain much of the variation in the productivity of mathematicians, it does so in regard to the productivity of chemists. In other areas of science, major changes in research methods have led to what is called big science, which is epitomized by high-energy physics. Despite the use of labor-saving devices, the work groups at high-energy particle laboratories can be very large, with well over 150 scientists. Such groups have a greater division of labor, a broader span of supervisory control, and greater centralization of decision making.

These work groups ordinarily are embedded in larger organizations such as universities and governmental or industrial establishments. They also are likely to be linked informally with other groups working on the same or related research problems in other establishments. These loosely linked and loosely bounded sets of work groups can be called "specialties" or, more evocatively, "invisible colleges" (Price 1986). Groups in a specialty simultaneously compete with one another and make contributions to one another's research. The number of groups in a specialty worldwide (there is a great deal of international collaboration) is ordinarily small, perhaps 50 on the average and seldom over 100, although specialties with over 500 groups exist. Scientists spend much of their time communicating with one another: writing papers, reviewing papers by others, attending meetings of scientific societies, and informally (Nelson and Pollock 1970).

The public nature of science tends to inhibit deviant behavior, but some deviance is to be expected. The extent of research fraud, such as forging and trimming data, is difficult to ascertain, as is the case in white-collar crime generally. Evidence and theories about such deviance are summarized in Zuckerman (1988). Fraud is most likely to occur when researchers are under pressure to get results (such as postdoctoral fellows and

nontenured faculty members) and when it is less likely to be detected (as in collaborative research with workers from different disciplines, where one is unable to evaluate the work of another); both conditions are especially likely to exist in experimental research in the biomedical sciences. Of courses, scientists with a high reputation also have engaged in research fraud; the case of the psychologist Cyril Burt is discussed in Gieryn (1998).

Scientific specialties usually exist within disciplines represented by their own university departments and scientific societies, but interdisciplinary research is common. The growth of an interdisciplinary area can lead to the differentiation of disciplines, and so the number of scientific disciplines has grown (Hagstrom 1965). The different scientific disciplines differ greatly in the degree of consensus about theories and methods; one indicator of this is variation in the rejection rates of manuscripts submitted to scientific journals, which is high in fields such as sociology and low in fields such as physics (Hargens 1975, 1988). Variations in consensus can affect the careers of scientists by affecting judgments of the merits of the work of individuals; it is easier to achieve early success in disciplines with a high degree of consensus. Disciplines also vary in the degree to which the work of scientists depends on and contributes to the work of others in their disciplines. This interdependence is related to Durkheim's concept of "organic solidarity." It is lower in mathematics than it is in the empirical sciences, as is indicated by fewer references in and citations of papers written by mathematicians, and it can be experienced as a problem by mathematicians (Hagstrom 1965).

#### THE SOCIOLOGY OF SCIENTIFIC KNOWLEDGE

*The Structure of Scientific Revolutions* by the historian Kuhn (1970), first published in 1962, strongly influenced the sociology of science. Kuhn made a distinction between normal and revolutionary science. Normal science is a puzzle-solving activity governed by paradigms. A paradigm consists of shared commitments by a group of scientists to a set of values, presuppositions about nature, methods of research, symbolic generalizations such as Newton's laws, and exemplars such as particular experiments. In normal science, researchers force nature into agreement with the paradigm; appar-

ently disconfirming evidence does not weaken commitment to the paradigm. Normally scientists are successful in explaining away apparently disconfirming evidence, but persistent critical anomalies can trigger a scientific revolution. In a successful revolution, one paradigm is succeeded by another paradigm with quite different presuppositions and exemplars. Kuhn (1970) argued that the contending paradigms in revolutionary situations are “incommensurable”; the choice between them is not and cannot be determined by evidence and formal decision rules alone. Kuhn illustrated his argument with evidence from major revolutions ranging from the Copernican Revolution of the sixteenth century to the revolutions that overthrew Newtonian physics in the twentieth century as well as smaller revolutions that affected the work of smaller sets of scientists.

The sociologists who developed the sociology of scientific knowledge, initially largely British, advanced radical arguments far beyond those of Kuhn. Not only are paradigms, or theories, “underdetermined” by data, theories are largely or entirely socially constructed. In Harry Collins’s words, “the natural world has a small or non-existent role in the construction of scientific knowledge. . . . [N]othing outside the courses of linguistics, conceptual and social behaviour can affect the outcome of these arguments” (quoted in Gieryn 1982). The constructivists have done a number of detailed case studies of changes in the content of science to support their claims. Their early work is summarized in Collins (1983), who shows how “data” were insufficient for resolving conflicts about an allegedly new type of laser. Others have studied cases such as disputes about gravity waves, the construction of quarks, and the history of statistics and genetics in the early twentieth century. In an ethnographic study of a laboratory that investigated neurohormones, Latour and Woolgar (1979) describe how facts were socially constructed. For example, initial claims constitute conjectures, and lower-order factual statements are qualified by the names of those making the claims. However, when they are successfully constructed, these qualifications are dropped and the facts are taken for granted, perhaps embedded in laboratory equipment or algorithms. Related work by other sociologists has involved detailed analyses of scientific discourse. Gilbert and Mulkay (1984) studied biochemists who did research on the process of

oxidative phosphorylation. Those authors showed that the sober prose of the scientific papers, where evidence and argument lead to conclusions, was contradicted by the informal discourse of the same scientists, who were partly aware that evidence and argument would be insufficient to persuade their opponents.

The constructivist position naturally leads to a relativistic position: If theories are social constructs, they could equally well be different. From his detailed study of the ways in which physicists constructed quarks in the period 1964–1974, Andrew Pickering (1984) concluded that “there is no obligation upon anyone framing a view of the world to take account of what twentieth century physics has to say. The particle physicists of the late nineteen-seventies were themselves quite happy to abandon most of the phenomenal world and much of the explanatory framework which they had constructed in the previous decade. There is no reason for outsiders to show the *present* HEP worldview any more respect.” This relativism leads constructivists to challenge the conventional demarcation between science and nonscience or pseudoscience. Thus, an article reporting a study of parapsychologists was titled “The Construction of the Paranormal: Nothing Unscientific Is Happening.”

These extreme claims have elicited much controversy. Representative criticisms by sociologists can be found in Gieryn (1982) and Amsterdamska (1990). Some natural scientists have argued that constructivism, along with several other “postmodern” schools of thought in the social sciences and humanities, represents a dangerous form of antisocialism; see Gieryn (1998) for a discussion of these “science wars.” Nevertheless, persuasive evidence has been produced about the importance of social factors in changing scientific knowledge. Stewart (1990) studied the recent revolution most widely known to the general public: plate tectonics in the 1950s and 1960s. He found strong resistance to the revolution. Earth scientists who had invested heavily in earlier perspectives were most likely to resist plate tectonics. Usually conversion to the new paradigm was gradual, sealed when scientists saw the relevance of the paradigm for their own research, but Stewart found some whose acceptance of plate tectonics came as the kind of “gestalt switch” described by Kuhn (1970).

In the conflicts accompanying the revolution, scientists on both sides deviated from conventional norms and used coercive methods to advance their positions and resist their opponents. Such intense conflict does not always accompany revolutions; in the one in physics that produced quarks, there was little acrimony or duress (Pickering 1984). In the earth sciences and physics, interests internal to the scientific disciplines affected the reception of theories. External interests also can have significant effects. Desmond (1989) shows how the interests of social classes interacted with religion in affecting the reception of Lamarckian ideas about evolution in England in the 1840s; the participants in the disputes were aware of the ideological implications of biological theories. Feminist sociologists of science have shown how gender interests have influenced perceptions of nature and the formulation of biological theories. See Keller (1995) for a review of some examples.

#### APPLIED RESEARCH AND DEVELOPMENT

The preceding discussion has concerned mostly basic research oriented primarily toward the advancement of knowledge. However, most research is done to advance other goals: corporate profits, weaponry, health, and human welfare. Of the 2 to 4 percent of their gross national products that advanced industrial countries devote to research and development (R&D), less than 10 percent is devoted to basic research (U.S. National Science Board 1996). Of the remainder, much is devoted to defense, particularly in the United States, where a substantial majority of federal R&D expenditures are devoted to that use.

Independent inventors are still an important source of innovations, but most applied scientists and engineers are salaried employees of corporations and mission-oriented government agencies. Such employees lack most of the autonomy of basic scientists. University-trained scientists are likely to chafe under this loss of autonomy, but successful applied research organizations have developed procedures for harmonizing their scientists' desires for autonomy with an organization's desire for useful knowledge (Kornhauser 1962). Engineers are important in translating knowledge into products and processes. Engineers are more pragmatic than scientists and are committed less

to paradigms and more to physical objects (when a scientist moves, he or she is likely to pack his or her journals first; when an engineer moves, she or he packs her or his catalogues). While scientists tend to seek autonomy in organizations, engineers tend to seek power; it is usually necessary to control organizational resources to do successful engineering work.

One of the conflicts that can occur between scientists and their industrial employers concerns communications. Scientists want to communicate their discoveries to their colleagues to gain recognition; their employers want to profit from the discoveries, and that may require keeping them secret. The patent system can provide an accommodative mechanism: Discoveries are made public, but those who wish to use the discoveries for practical purposes must pay royalties to the patent holder. The patent system represents one aspect of the commodification of knowledge. Marxist theories imply that in capitalist social formations, goods and services are produced for sale as commodities, not for use, and this is increasingly the case for scientific knowledge. Kloppenburg (1988) has applied Marxist thought effectively in his history of plant breeding. There were and are inherent problems in making seeds into a commodity, since seeds tend to reproduce themselves; they can be both objects of consumption and part of the means of production. Until recently, seeds seldom were produced as commodities; new varieties were exchanged among farmers or distributed to them by government agencies at little cost, and the farmers would then grow their own seeds. This changed first with the development of hybrid corn, where farmers could not use the corn they produced as seed and instead bought new seed from the seed companies each season. This process has since been extended to other crops. In addition, consistent with Marxist thought, the seed industry has become increasingly centralized and concentrated, with fewer and larger firms dominating it. Those firms also expand into world markets, acquiring germ plasm in third world countries and selling seeds as commodities in those countries. The development of biotechnology has increasingly taken this form. Rapid developments in this area blur the distinction between basic and applied research. The emerging pattern seems to be one in which research that cannot be used to generate a profit is done in universities and gov-

environmental agencies, usually at public expense, while research that can be used for profit is done in corporations.

Modern science has led to massive changes in the lives of people in all countries, and it has the potential for further changes. For example, it has made major contributions to economic growth (Mansfield 1991). However, not all these changes have been beneficial, and not all beneficial changes are allocated equitably. While polls show high support for science in general, there are intense public controversies in many areas, from the use of animals in biomedical research, to global warming, to military technologies (Nelkin 1995). Sometimes research and development efforts can achieve considerable autonomy. MacKenzie (1993) shows how those who developed the inertial navigation system for submarine-launched missiles successfully “black-boxed” their efforts so that political officials would not interfere. The navigation technology could have had seriously destabilizing effects in the cold war, without any deliberation by elected officials. The autonomy of engineers sometimes achieve does not imply autonomy for scientists. Thus, while oceanographers have made major discoveries in the past forty years, their expensive research has been driven largely by the interests of the U.S. Navy and their autonomy has been constrained by its interests (Mukerji 1990). Attempts have been made to develop more democratic means for developing science policy. Collingridge (1980) argues for an approach of “disjointed incrementalism”: Since problems are rarely foreseen, policy making should be fragmented rather than centralized and will often be remedial; since it is not feasible to investigate all solutions, analysis and evaluation should be serial and incremental. Democratic governments have developed organizations to mediate between science and governmental institutions. These organizations can be nongovernmental, such as the National Academy of Sciences–National Research Council; part of the legislative branch, such as the Office of Technology Assessment of the U.S. Congress; or part of the executive branch, such as the Environmental Protection Agency. (For a description of these efforts in the United States and the difficulties they face, see Bimber and Guston 1995; Cozzens and Woodhouse 1995.) The growth and rapid change of science-based technologies present difficult problems in regard to support and

control. Knowledge about the organization of science and its relationships with other institutions can help in dealing with those problems.

## REFERENCES

- Allison, Paul D., and J. Scott Long 1990 “Departmental Effects on Scientific Productivity.” *American Sociological Review* 55:469–78.
- , J. Scott Long, and Tad K. Krauze 1982 “Cumulative Advantage and Inequality in Science.” *American Sociological Review* 47:615–625.
- Amsterdamska, Olga 1990 “Review of *Science in Action* by Bruno Latour.” *Science, Technology, and Human Values* 15:495–504.
- Ben-David, Joseph 1984 *The Scientist’s Role in Society*. Chicago: University of Chicago Press.
- 1991 *Scientific Growth: Essays on the Social Organization and Ethos of Science*. Berkeley: University of California Press.
- Bimber, Bruce, and David H. Guston 1995 “Politics by the Same Means: Government and Science in the United States.” In Sheila Jasanoff, G. Markle, J. C. Petersen, and T. Pinch eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif. Sage.
- Cohen, H. Floris 1994 *The Scientific Revolution: A Historiographical Inquiry*. Chicago: University of Chicago Press.
- Cole, Jonathan R., and Stephen Cole 1973 *Social Stratification in Science*. Chicago: University of Chicago Press.
- Collingridge, David 1980 *The Social Control of Technology*. London: Frances Pinter.
- Collins, Harry M. 1983 “The Sociology of Scientific Knowledge: Studies of Contemporary Science.” In *Annual Review of Sociology*, vol. 9. R. H. Turner and J. F. Short, Jr., eds. Palo Alto, Calif. Annual Review.
- Cozzens, Susan E., and Edward J. Woodhouse 1995 “Science, Government, and the Politics of Knowledge.” In Sheila Jasanoff, G. Markle, J. C. Petersen, and T. Pinch, eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif.: Sage.
- Desmond, Adrian 1989 *The Politics of Evolution*. Chicago: University of Chicago Press.
- Gieryn, Thomas F. 1982 “Relativist/Constructivist Programmes in the Sociology of Science: Redundance and Retreat.” *Social Studies of Science* 12:279–297.
- 1995 “Boundaries of Science.” In Sheila Jasanoff, G. Markle, J. C. Petersen, and T. Pinch eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif.: Sage.

- 1998 *Cultural Boundaries of Science*. Chicago: University of Chicago Press.
- Gilbert, G. Nigel, and Michael Mulkey 1984 *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge, UK: Cambridge University Press.
- Hagstrom, Warren O. 1965 *The Scientific Community*. New York: Basic Books.
- 1974 "Competition in Science." *American Sociological Review* 39:1–18.
- 1976 "The Production of Culture in Science." *American Behavioral Scientist* 19:753–768.
- Hargens, Lowell L. 1975 *Patterns of Scientific Research: A Comparative Analysis of Research in Three Scientific Fields*. Washington, D.C.: American Sociological Association.
- 1988 "Scholarly Consensus and Journal Rejection Rates." *American Sociological Review* 53:139–151.
- Isis 1988 Symposium on the Fiftieth Anniversary of *Science, Technology and Society*, vol. 79, pp. 571–605.
- Keller, Evelyn Fox 1995 "The Origin, History, and Politics of the Subject Called 'Gender and Science.'" In Sheila Jasanoff, G. Markle, J. C. Petersen, and T. Pinch eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif.: Sage.
- Kloppenborg, Jack R., Jr. 1988 *First the Seed: The Political Economy of Plant Biotechnology 1492–2000*. New York: Cambridge University Press.
- Kornhauser, William 1962 *Scientists in Industry: Conflict and Accommodation*. Berkeley: University of California Press.
- Kuhn, Thomas S. (1962) 1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Latour, Bruno, and Steve Woolgar 1979. *Laboratory Life: The Social Construction of Scientific Facts*. Beverly Hills, Calif.: Sage.
- MacKenzie, Donald A. 1993 *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. Cambridge, Mass.: MIT Press.
- Mansfield, Edwin 1991 "Academic Research and Industrial Innovation." *Research Policy* 20:1–12.
- Merton, Robert K. (1938) 1970. *Science, Technology and Society in Seventeenth Century England*. New York: Howard Fertig.
- 1973 *The Sociology of Science*. Chicago: University of Chicago Press.
- Mitroff, Ian 1974 "Norms and Counternorms in a Select Group of Scientists: A Case Study in the Ambivalence of Scientists." *American Sociological Review* 39:579–595.
- Mukerji, Chandra 1990 *A Fragile Power: Scientists and the State*. Princeton N.J.: Princeton University Press.
- Mulkey, Michael J. 1980 "Interpretation and the Use of Rules: The Case of the Norms of Science." In T. F. Gieryn, ed., *Science and Social Structure: A Festschrift for Robert K. Merton*. New York: New York Academy of Sciences.
- Nelkin, Dorothy 1995 "Scientific Controversies: The Dynamics of Public Disputes in the United States." In Sheila Jasanoff G. Markle, J. C. Petersen, and T. Pinch eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif.: Sage.
- Nelson, Carnot E., and Donald K. Pollock, eds. 1970 *Communication among Scientists and Engineers*. Lexington, Mass.: Heath.
- Parsons, Talcott 1951 *The Social System*. Glencoe, Ill.: Free Press.
- Pickering, Andrew 1984 *Constructing Quarks: A Sociological History of Particle Physics*. Chicago: University of Chicago Press.
- Price, Derek J. de S. (1963) 1986 *Little Science, Big Science . . . and Beyond*. New York: Columbia University Press.
- Rudwick, Martin J. S. 1985 *The Great Devonian Controversy: The Shaping of Scientific Knowledge among Gentlemanly Specialists*. Chicago: University of Chicago Press.
- Shapin, Steven 1994 *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago: University of Chicago Press.
- , and Simon Shaffer 1985 *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton N.J.: Princeton University Press.
- Stewart, John A. 1990 *Drifting Continents and Colliding Paradigms: Perspectives on the Geoscience Revolution*. Bloomington: Indiana University Press.
- U.S. National Science Board 1985 *Science and Engineering Indicators*. Washington, D.C.: National Science Foundation.
- , 1996 *Science and Engineering Indicators*. Washington, D.C.: National Science Foundation.
- Watson-Verran, Helen, and David Turnbull 1995 "Science and Other Indigenous Knowledge Systems." In Sheila Jasanoff, G. Markle, J. C. Peterson, and T. Pinch eds., *Handbook of Science and Technology Studies*. Thousand Oaks, Calif.: Sage.
- Xie, Yu, and Kimberlee A. Shauman 1998 "Sex Differences in Research Productivity." *American Sociological Review* 63:847–870.
- Znaniecki, Florian 1940 *The Social Role of the Man of Knowledge*. New York: Columbia University Press.

Zuckerman, Harriet 1988. "The Sociology of Science."  
In Neil J. Smelser, ed., *Handbook of Sociology*. Newbury  
Park, Calif.: Sage.

WARREN HAGSTROM

## SCIENTIFIC EXPLANATION

Science and scientific knowledge achieved high status in twentieth-century Western societies, yet there continues to be disagreement among scientists and those who study science (historians, philosophers, and sociologists of science) about the meaning of scientific explanation. Indeed, the use of the word "explanation" has been the subject of heated debate (Keat and Urry 1982).

One way to make sense of science is to "reconstruct" the logic scientists use to produce scientific knowledge. The reconstructed logic of science differs from what scientists actually do when they engage in research. The research process is seldom as clear, logical, and straightforward as the reconstructed logic presented in this article makes it appear. For a long time, the most popular reconstruction of the logic of the scientific process was the "hypothetico-deductive" model. In this model, "the scientist, by a combination of careful observation, shrewd guesses, and scientific intuition arrives at a set of postulates governing the phenomena in which he is interested; from these he deduces observable consequences; he then tests these consequences by experiment, and so confirms or disconfirms the postulates, replacing them, where necessary, by others, and so continuing" (Kaplan 1964, pp. 9-10; see also Braithwaite 1968; Nagel 1961). The description of scientific explanation presented here is broadly consistent with this model as it is used in the social sciences.

Scientific explanations can be contrasted to other, nonscientific types of explanation (Babbie 1989; Kerlinger 1973; Cohen and Nagel 1934). Some explanations obtain their validity because they are offered by someone in authority, for example, a police officer, the president, or parents. Validity also may rest on tradition. For instance, the correct way to do a folk dance is the way it has always been danced, handed down over the generations. This knowledge is not obtained by going through textbooks or conducting experiments but is stored in the memories and beliefs of

individuals. Another way of knowing is a priori, or intuitive, knowledge. This knowledge is based on things that "stand to reason," or seem to be obvious, but are not necessarily based on experience. People tend to cling strongly to intuitive knowledge even if the "facts" do not match their experience. Situations that contrast with strongly held beliefs are explained away as unique occurrences that will not happen again. For example, it "stands to reason" that if you are nice to other people, they will be nice to you.

The scientific method is a way of obtaining information, or knowledge, about the world. Theoretically, the same knowledge will be obtained by everybody who asks the same questions and uses the same investigative method. Scientific explanation uses theories, deductive and inductive logic, and empirical observation to determine what is true and what is false. Unlike authoritarian, traditional, or intuitive explanations, scientific knowledge is always supposed to be open to challenge and continual correction.

A theory is a hypothetical explanation for an observation or a question such as Why is the sky blue? or Why do victims of child abuse often grow up to be perpetrators? Scientists develop and test theories by using deductive logic, trying to show that empirical observations are instances of more general laws. Scientific theories are hypothetical explanations that state the possible relationships among scientific concepts. Theories consist of "a set of interrelated constructs (concepts), definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting the phenomena" (Kerlinger 1973, p. 9). Theories also are used by scientists to interpret, criticize, and bring together established laws, often modifying them to fit unanticipated data. They also guide the enterprise of making new and more powerful generalizations (Kaplan 1964, p. 295).

Scientific theories generally take the form of "If X happens, then Y will happen." For instance, Karl Marx's theory of surplus value suggests that as the level of surplus value in a capitalist society increases, so will inequality. This is an attempt to determine causal relations, so that theories not only predict what will happen in the world but also explain why it happens.



In general, scientific explanations are derived using nomothetic methods, which have the goal of making generalizations or of establishing universal laws. The experiment is perhaps the best known nomothetic method. Scientific theories try to generalize, or predict, beyond the specific data that support them to other similar situations. In contrast, some forms of the social sciences and humanities use idiographic methods, which are designed to provide knowledge of one particular event which may be unique. The best known idiographic method may be the case study. For example, both social scientists and historians investigate wars. A social scientist tries to explain what is common to all wars, possibly so that she or he can develop a general theory of intersocietal conflict. In contrast, a historian studies individual wars and tries to chronicle and explain the events and conditions that cause a specific war, and is generally not interested in a scientific theory of what may be common to all wars.

It seems that there is a paradox here: Scientific explanations are the best explanations that can be offered for an event, yet scientific theories are always open to correction by a better explanation or theory. What counts as a “better” explanation or theory has been the subject of debate in the philosophy of science. Some people believe that the better theories are those which can explain anomalies that previous theories could not. In other words, the new, “better” theory can explain everything the old theory could but also can explain some things that it left unexplained. There are many debates among philosophers of science about how to judge the “goodness” of a theory. They all admit that theories can never be confirmed definitively by any amount of observational material. The possibility always exists of finding an event that does not fit the theory, thus falsifying it. However, some theories have so much observational evidence on their side that they are said to be well confirmed, and the possibility of finding observations that falsify them is considered negligible.

However, the philosopher of science Popper said that while one can never absolutely *confirm* theories, one can definitively *falsify* them (1959). In other words, it is possible to find definite events that disconfirm, or falsify, a theory. However, other philosophers argue that this is not necessarily true, because it is always an open question whether it is the theory that is wrong or one of the

assumptions that is not tested when the theory is tested.

A famous example of this problem of falsification is provided by the philosopher of science Carl Hempel in his historical examination of the work of the Hungarian physician Semmelweiss (1966). Semmelweiss was concerned with the high rates of maternal mortality during childbirth. He theorized that those deaths resulted from blood poisoning, which was caused by infectious matter carried on the physician’s hands. Physicians were examining women right after performing dissections in the autopsy room. Semmelweiss’s hypothesis led him to believe that if the infectious matter was removed before the women were examined, the death rates would drop. To test this, he had doctors wash their hands in a solution of chlorinated lime *after* performing dissections and then examine women who had just given birth. As he predicted, the mortality rates fell as this procedure was practiced, providing evidence *confirming* his theory.

However, if the mortality rates had *not* fallen, that would not necessarily have meant that the theory was wrong. It could have meant that one of the unexamined assumptions, such as that chlorinated lime destroys infectious matter, was wrong. Thus, the theory would have been true but the experiment would not have provided evidence to confirm it because one of its untested assumptions was incorrect. Thus, falsification is a double-edged sword: When a theory is not confirmed, it is necessary to determine whether it is the thing that is being manipulated experimentally (the hand washing in chlorinated lime) that is the causal factor or whether one of the assumptions underlying the experiment is faulty (if it turned out that chlorinated lime did not kill infectious matter) (Hempel 1966, pp. 3–6). Scientists have to be careful not to give up on a theory too soon, even if early results appear to falsify it, because many major scientific achievements would not have occurred if they had been quickly abandoned (Swinburne 1964).

Whether philosophers of science hold to the confirmationist view or the falsificationist view of testing scientific theories, they agree on two things. The first is that scientific theories are universal statements about regular, contingent relationships in nature; the second is that the observations used to evaluate scientific theories provide an objective

foundation for science (Keat and Urry, 1982, p. 16). One of the goals of science is to develop and test theories, although some scientists believe that science proceeds inductively, purely by amassing facts and building theories from the amassed data.

Scientific laws fall broadly into two types: deterministic laws and stochastic (probabilistic) laws. For deterministic laws, if the scientist knows the initial conditions and the forces acting on a system and those factors do not change, the state of the system can be determined for all times and places. Deterministic laws are the ideal of the Newtonian, or mechanistic, model of science. In this model, it is assumed that causes precede effects and that changes come only from the immediately preceding or present state, never from future states. It is assumed that if two systems are identical and are subject to the same initial conditions and forces, they will reach the same end point in the same way. Deterministic laws assume that it is possible to make a complete separation between the system and the environment and that the properties of the system arise from its smallest parts. The smallest parts of a system are those about which nothing can be determined except their location and direction. There is nothing in the parts themselves that influences the system, and all changes in the state of the system come from the forces acting on it. Deterministic laws are based on the assumption that the universe is regular and that connections between events are independent of time and space. The idea with a scientific explanation is that all other things being equal (*ceteris paribus*), identical circumstances lead to identical results.

Stochastic laws are expressed in terms of probability. For large or complex systems, it is not possible to identify precisely what state the system will be in at any given time but only to assess the probability of its being in a certain state. Quantum physics, chemistry, card games, and lotteries utilize stochastic laws. Those laws are stated in terms of probability over time and apply to classes of events rather than to specific instances. Most relationships in the social sciences are stated in stochastic terms because individual behavior is very difficult to predict. The use of probability does not mean that events are viewed as random, or uncaused, simply that the behavior of the individual elements of a system cannot be predicted with perfect accuracy.

Scientific theories are systematically linked to existing knowledge that is derived from other generally accepted theories. Each scientist builds on the work of other scientists, using tested theories to develop new theories. The scientific method is dedicated to changing theories, and scientific knowledge progresses through the challenge and revision of theories.

Often a new theory is preferred not because it is based on facts (data) that are different from those on which the old theory was based but because it provides a more comprehensive explanation of existing data. For example, Newton's theory of the solar system superseded Kepler's explanation of planetary motion because Newton's theory included the theory of gravity (which predicted a gravitational attraction between all physical bodies in the universe) as well as the laws of motion. The two theories together provided many circumstances that could "test" the theory because they *predicted* not only where planets should be in relation to each other at given times but also phenomena such as falling apples and swinging pendulums. Newton's theory was more comprehensive and more economical, and although it provided more opportunities for falsification than did Kepler's (which made it more vulnerable), it also resisted falsification better and became the accepted scientific explanation (Chalmers 1982).

The premises, or propositions, in a scientific theory must lead logically to the conclusions. Scientific explanations show that the facts, or data, can be deduced from the general theory. Theories are tested by comparing what deduction says "should" hold if the theory is true with the state of affairs in the world (observations). The purpose of a theory is to describe, explain, and predict observations.

The classic example of deductive logic is the familiar syllogism "All men are mortal; Socrates is a man; therefore, Socrates is mortal." Deductive conclusions include only the information included in the propositions. Thus, deductive reasoning can be logically correct but empirically incorrect. If a theory is based on empirically false premises, it probably will result in empirically false conclusions. A scientific test of the truth of the conclusions requires a comparison of the statements in the conclusion with actual states of affairs in the "real" world.

Scientific explanations and theories are usually quite complex and thus often require more information than can be included in a deductively valid argument. Sometimes it is necessary to know that a conclusion is probably true, or at least justified, even if it does not follow logically from a set of premises and arguments (Giere 1984). Thus, there is a need for inductive logic, which is based on particular instances (facts or observations) and moves to general theories (laws).

Many sociologists and other scientists believe that scientific knowledge is produced mainly by induction (Glaser and Strauss 1967). For example, after one has observed many politicians, a theory might postulate that most politicians are crooked. Although this theory is based on many observations, its proof, or verification, would require observing every politician past, present, and future. Falsifying the theory would require finding a substantial number of politicians who were not crooked. The absolute and final verification of scientific theories is not possible. However, it should be possible to “falsify” any scientific theory by finding events or classes of events that do not support it (Stinchcombe 1987; Popper 1959).

Because inductive arguments are always subject to falsification, they are stated in terms of probabilities. Good inductive arguments have a high probability associated with their being true. This high probability comes from a large number of similar observations over time and in different circumstances. For example, although it is not absolutely certain that if someone in North America becomes a medical doctor, he or she will earn a high income, the evidence provided by observing doctors in many places and many times shows that a high probability can be assigned to the assertion that medical doctors earn high incomes.

Inductive arguments are not truth-preserving. Even with true premises, an inductive argument can have a false conclusion because the conclusions of inductive arguments generally contain more information or make wider generalizations than do the premises (Giere 1984). Science requires both deductive and inductive methods to progress. This progress is circular: Theories are developed and tested, and new data give rise to new theories, which then are tested (Wallace 1971).

Several steps are involved in testing scientific theories. Theories first must be expressed in both

abstract, verbal terms and concrete, operationalized terms. Concepts and constructs are rich, complex, abstract descriptions of the entity to be measured or studied. Concepts have nominal definitions (they are defined by using other words) and are specifically developed for scientific purposes. A variable is operationally defined to allow the measurement of one specific aspect of a concept. Operationalization is a set of instructions for how a researcher is going to measure the concepts and test the theory. These instructions should allow events and individuals to be classified unambiguously and should be precise enough that the same results will be achieved by anyone who uses them (Blalock 1979).

For example, one theory posits that the relationship between “anxiety” and test performance is curvilinear. This theory predicts that very little anxiety leads to poor performance on tests (as measured by grades), a medium amount of anxiety improves test performance, and very high anxiety causes poor test performance. If it were drawn on a graph, this curve would be an upside-down U. To test the theory, both anxiety and test performance must be measured as variables expressed in empirical terms. For an observation to be empirical means that it is, or hypothetically could be, experienced or observed in a way that can be measured in the same manner by others in the same circumstances.

As a concept, anxiety encompasses many different things. The measurement theory must specify whether anxiety will be measured as feelings, such as being tense, worried, or “uptight,” or as physical reactions, such as shortness of breath, heart palpitations, or sweaty palms. The researcher may decide to measure anxiety by asking subjects how worried or tense they felt before an examination. Racing hearts, sweating palms, and upset stomachs are part of the concept, but they are excluded from the operationalization. The researcher must decide whether this is or is not a valid (measures what it purports to measure) and reliable (obtains the same results on repeated tests) measure of anxiety, in part by comparing the results of the research to other research on anxiety and test performance. It is also necessary to strike a balance between the scope of the concept (the different things it refers to) and precision. The wider the scope of a concept, the more it can be generalized to other conditions and the fewer

conditions are required to construct a theory, making it more parsimonious. However, if the scope of a concept is too wide, the concept loses precision and becomes meaningless.

Scientific explanation involves the accurate and precise measurement of phenomena. Measurement is the assignment of symbols, usually numbers, to the properties of objects or events (Stevens 1951). The need for precise measurement has led to an emphasis on quantification. Some sociologists feel that some qualities and events that people experience defy quantification, arguing that numbers can never express the meaning that people's behavior holds for them. However, mathematics is only a language, based on deductive logic, that expresses relationships symbolically. Assigning numbers to human experiences forces a researcher to be precise even when the concepts, such as "anxiety" and "job satisfaction," are fuzzy.

Another important aspect of scientific explanations is that they attempt to be "objective." In science this term has two broad meanings. First, it means that observers agree about what they have observed. For example, a group of scientists observing the behavior of objects when they are dropped would agree that they saw the objects "fall" to the ground. For this observation to be objective, (1) there must be an agreed-on method for producing it (dropping an object), (2) it must be replicable (more than one object is released, and they all "fall"), and (3) the same results must occur regardless of who performs the operation and where it is performed (objects must behave the same way for all observers anywhere in the world). Scientific operations must be expressed clearly enough that other people can repeat the procedures. Only when all these conditions are met is it possible to say that an observation is objective. This form of objectivity is called "intersubjectivity" and it is crucial to scientific explanations.

The second use of the word "objective" in science means that scientific explanations are not based on the values, opinions, attitudes, or beliefs of the researcher. In other words, scientific explanations are "value-free." A researcher's values and interests may influence what kinds of things she or he chooses to study (i.e., why one person becomes a nuclear physicist and another becomes a sociologist), but once the problem for study is chosen the

scientist's personal values and opinions do not influence the type of knowledge produced. The value-free nature of science is the goal of freeing scientific explanations from the influence of any individual or group's biases and opinions.

The relationships in a theory state how abstract constructs are to be linked so that antecedent properties or conditions can be used to explain consequent ones. An antecedent condition may be seen as either necessary or sufficient to cause or produce a consequent condition. For example, higher social status may be seen as *sufficient* to increase the probability that farmers will adopt new farming techniques (innovation). It also could be argued that awareness and resources are necessary conditions for innovation. Without both, innovation is unlikely (Gartrell and Gartrell 1979).

Relationships may be asymmetrical (the antecedent produces the effect) or symmetrical (both cause each other): Frustration may cause aggression, and aggression may cause frustration. Relationships may be direct, or positive (an increase in knowledge causes an increase in innovation), or negative (an increase in stress leads to a decrease in psychological well-being). They may be described as monotonic, linear, or curvilinear. Sociologists often assume that relationships are linear, partly because this is the simplest form of a relationship.

Relationships between variables are expressed by using a wide variety of mathematical theories, each of which has its own "language." Algebra and calculus use the concepts of "greater than," "less than," and "equal to." Set theory talks about things being "included in," and graph theory uses "connectedness" or "adjacency between." Markov chains attempt to identify a connectedness in time or a transition between states, and symbolic logic uses the terms "union" and "intersection" to talk about relationships.

Scientific explanation is also very explicit about the units to which relationships between propositions refer. Sociologists refer to a host of collectivities (cultures, social systems, organizations, communities), relationships (world systems, families), and parts of collectivities (social positions, roles). There is strength in this diversity of subject matter but also potential weakness in failing explicitly to define the unit of analysis. Some properties cannot be attributed to all units of analysis. For example, "income" is a concept that can apply to an individ-

ual or a group (e.g., “average” income), but “inequality” is always a property of an aggregate. The “ecological fallacy” (Hannan 1970) involves the incorrect attribution of properties of aggregates to individuals. Aggregation is not a matter of simple addition, and some relationships between subunits (homogeneity, complexity, inequality) have complicated aggregation algorithms. Care must be taken in switching units of reference from social collectivities to individuals. For example, communities with high divorce rates also may have high homicide rates, but this does not necessarily imply that divorced people kill one another or are more likely to be homicide victims or perpetrators.

To test theories, the relationships among concepts are stated as hypotheses, linking variables in an operationalized form. Since the existence of a relationship cannot be proved conclusively, a scientist instead tries to show that there is no relationship between the variables by testing hypotheses that are stated in the “null” form. In the test performance and anxiety example, a null hypothesis would state, “There is no curvilinear relationship between the number of correct responses on tests and the reported level of worry and tension.” If this hypothesis was rejected, that is, found to be highly unlikely, the researcher would have evidence to support the alternative hypothesis suggested by the theory: There is a curvilinear relationship between the variables.

Social scientists use a variety of methods to study human behavior, including experiments, surveys, participant observation, and unobtrusive measures. In essence, experiments try to identify causal sequences by determining the effect of an independent variable (the stimulus) on a dependent variable. Experiments require stringent conditions that often are difficult to fulfill with human beings, sometimes for ethical reasons but more often because there is a wide variation in individual responses to the same stimulus (Babbie 1989; Kerlinger 1973; Cook and Campbell 1979).

Social scientists have developed other research methods, such as surveys and field research, which allow them to produce scientific knowledge without resorting to experimental manipulation. Statistical analysis of survey data allows social scientists to examine complex problems in large populations by statistically controlling several variables that represent competing explanations (Blalock

1964). The distinctive characteristic of survey research is that the subjects of the study tell the scientist about themselves.

Social scientists also use qualitative methods such as participant observation to conduct research in the “field” where phenomena actually occur. Field research focuses on the empirical richness and complexity of the whole subject in order to understand what is subjectively meaningful. Participant observation proceeds inductively rather than deductively. The researcher observes and participates in order to understand (subjectively) and then attempts to externalize the observations by constructing categories of responses, or theory. In contrast to other research designs, participant observation deliberately does not attempt to control conditions; the researcher strives to obtain an unbiased picture of how the subjects see things in their natural setting (Whyte 1961). The emphasis is on the richness of subjects’ understanding of events and on subjectivity rather than on objectivity. Theory developed from this type of research is called grounded theory (Glaser and Strauss 1967). Unobtrusive methods such as content analysis focus on the study of artifacts (newspapers, homes), partly to overcome reactivity by subjects and biases on the part of the researcher.

#### **CRITIQUES OF THE HYPOTHETICO- DEDUCTIVE MODEL OF SCIENCE**

In the 1930s and 1940s, the dominant view of science was “radical positivism,” which viewed science as a process based only on inductive generalizations and empirical verification. Abstract theoretical concepts that could not be observed were considered literally meaningless. The revision of positivism in the 1950s (logical empiricism) recognized the importance of abstract concepts and theories but continued to insist that all scientific statements be subject to empirical falsification. In short, the empiricists persisted in their belief that “facts” were purely objective entities and that what was viewed as a fact did not depend on theory or theoretical concepts. However, theories play as large a role in scientific change and knowledge production as do empirical observations. In part, this internal confusion laid the groundwork for a wide range of critiques of both positivism and empiricism (Alexander 1982; Bernstein 1976).

Reconstructed logic suggests that scientific knowledge can be accounted for by following formal rules of logic. The progress of knowledge is such that unscientific or prescientific explanations for phenomena are replaced successively by scientific explanations, which are ever closer approximations to the “truth.” It stresses that the knowledge produced by the scientific method is objective and value-free, corresponding to states of the world as it really is, not as it is seen by a particular group of people in a particular social and historical location.

However, the “facts” on which scientific explanations are based are not independent of “point of view” (Polanyi 1958; Hanson 1972). All scientific data are theoretically informed. What is “fact” and what is “theory” are what is convenient to the focus of scientific attention at a particular time. Because science is a social and cultural activity, it is grounded in an everyday, taken-for-granted reality. Scientists can perceive “facts” only in a particular social and cultural context. Observations take place in a cultural milieu that literally affects what the observer perceives, not just how it is interpreted. The totally objective, theory-free observation aspired to in science is not possible; to “see” something is always to see it “as” something. For example, to observe the medical “facts” in an x-ray, a physician must first learn what parts of the picture to ignore. The “fact” that objects “fall” to the ground is a fact only in a social context in which gravity is an accepted explanation for the behavior of falling objects. Scientific facts are constructed and developed through situated human labor; they do not have an independent, objective existence of their own (Fleck 1979).

Most twentieth-century philosophers of science have assumed that there is something called the scientific method that applies equally to all sciences and that sciences can be judged by their ability to adhere to that method. This is called the “unity of the sciences” model. However, the philosophy of science has ignored the actual behavior of scientists, concentrating instead on reconstructing the logic of science. The result has been an idealized and unrealistic picture of how scientific knowledge is produced. When the actual practice of scientists is observed, it is apparent that in different sciences, scientists reason in a wide variety of modes.

These different modes of reasoning were hidden by the philosophical approach of viewing scientific knowledge as resulting from the simple application of scientific logic to problems. Scientific knowledge is better seen as the outcome of an active, work-oriented process than as an uninvolved description of a “passive” natural world. This means that scientific knowledge production consists largely in activities in which scientists make decisions about how to proceed in different circumstances. This does *not* imply that scientific knowledge is “made up” and thus completely relative but instead, by looking at scientific practice rather than only scientific logic, that the view of science has shifted from science as a “representation” of nature to science as “action” or “work” (Knorr Cetina 1981).

The most definitive research into how the various sciences produce knowledge differently is represented by the work of the sociologist of science Knorr Cetina (1999). Knorr Cetina has examined the practical activity and “cultures of knowing” of two very different sciences: molecular biology and high-energy physics. She has focused on the “concrete, mundane, everyday practices of inquiring and concluding through which participants establish, for themselves and for others, knowledge claims” (1991, p. 108). Her research shows that what counts as “scientific method” differs radically between these two sciences. In other words, the cultural structure of scientific methodology varies from science to science (1991, p. 107).

Knorr Cetina demonstrates that the epistemic culture in a molecular biology laboratory is such that molecular biologists have to become “repositories of unconscious experience” and individual scientists have to develop an embodied sense of a reasonable response to different situations (1992, p. 119). A practicing molecular biologist literally *becomes* a measurement instrument. These scientists become highly skilled at seeing things others cannot see, and their bodies learn to perform delicate operations in loading gels and manipulating DNA that cannot be taught, only learned through experience. In their scientific work, individual molecular biologists often have to *guess* what procedure is best in a given situation. For this reason, the sense of what counts as a successful procedure depends heavily on an individual’s experience and the predictive ability “which indi-

viduals must somehow synthesize from features of their previous experience, and which remains implicit, embodied, and encapsulated within the person" (1992, p. 121). What counts as a successful procedure or as proper scientific method is implicit: It is a blend of the individual's experience and the culture in the laboratory. Knorr Cetina calls this kind of reasoning "biographical" because "it is sustained by a scientist's biographical archive and the store of his or her professional experience" (1991, p. 115).

In contrast to the highly individual and personalized culture of knowing in a molecular biology laboratory, high-energy physics laboratories are very different kinds of epistemic spaces. Their organization is best compared to that of a superorganism, such as highly organized colonies of bees, ants, or termites. High-energy physics involves more circularities and contingencies than does molecular biology; its experiments are long term and "supra-individual."

In high-energy physics (HEP) experiments, the work of producing knowledge is detached from the individual scientist and shifted to the group. These experiments can involve from 200 to 2,000 individuals from 200 different institutions around the world, all focused on a common goal, for up to twenty years (Knorr Cetina 1999, p. 160). Authorship belongs to the experiment as a whole; individual scientists feel that they are representatives of the whole, and there is a sense of collective responsibility among them. (Knorr Cetina 1995). Unlike the highly trained body and eyes of a molecular biologist, data interpretation in HEP is done not by individual scientists but by computers. In fact, individual scientists literally cannot run experiments. HEP experiments are huge, they take many years to run, and each experiment seeds new generations of experiments. High-energy physicists do not think in terms of individual achievements in months but of group successes over years and decades.

In HEP, forming a consensus about what counts as adequate scientific knowledge and the proper application of scientific method is very much a group process. In molecular biology, the group is involved in terms of the culture of the laboratory but each individual scientist is a highly skilled measuring instrument that makes most procedural decisions on his or her own. Thus, by examining

the organization of the laboratories and the working practices of the scientists in these two domains, Knorr Cetina has challenged the philosophical assumption of a unitary scientific method.

Science is now widely regarded as a social activity rather than an application of logic to nature. It is seen as an interplay between practical activity, empirical observations, and broad theoretical "paradigms" (Kuhn 1970; Fleck 1979). Paradigms dictate the valid questions for research as well as the range of possible answers and can be so powerful that contradictory data (anomalies) are explained away under the assumption that they can be brought into the theory at a later time. Confronted by contradictory empirical evidence that cannot be ignored, the adherents of a theory often develop ad hoc hypotheses and residual categories to account for anomalies. Thus, they encompass or explain observations that contradict their theories and often cling to those theories in dogmatic fashion. The reconstructed logic of science leads one to believe that theories would be rejected under those conditions.

However, sociological research has shown that "the data" do not and cannot speak for themselves and decide between competing scientific theories. Sometimes a theory wins out over its competitors because its survival is in the best interests of a group or researcher (Woolgar 1981; Shapin 1979). For example, when high-energy particle physicists were searching for the subatomic particles now known as quarks, two competing explanations were advanced: the "charm" and "color" theories. Both models were consistent with the data. The ultimate success of the charm model occurred because more people had an interest in seeing it succeed. Charm theorists were more successful in relating their theory to an existing body of practice and interests. The color theory was never empirically refuted but eventually "died" because its proponents were reduced to talking to themselves (Pickering 1982).

Part of the problem is that the decision about whether certain experiments or observations are critical to the proof or falsification of a theory is possible only after the fact, not before, and the possibility always exists that an experiment failed because it was not performed competently. It is difficult to establish the criteria for determining whether an experiment has been successful. To

know whether the experimental apparatus, the theory, and the competence of the researcher have combined to produce a successful experiment, it is necessary to know beforehand what the correct outcome is. However, the definition of a competently performed experiment is having a successful outcome, leading to the “experimenter’s regress” (Collins 1985).

The replication of results is an essential criterion for the stability of scientific knowledge, but scientific inquiry requires a high degree of tacit or personal knowledge (Polanyi 1958). This knowledge is by nature invisible, but its importance is strongly denied by a scientific community that bases its claims to validity on the potential for replication. Scientific developments often cannot be replicated unless there is direct, personal contact between the original researcher and the people attempting to do the replication. Few replications are possible using published results and procedures, and successful replication often rests on the original researcher’s tacit knowledge, which is not easily transferable (Collins 1985). To complicate matters, science reserves its highest rewards for original research rather than replication. As a consequence, there is little glory and less funding for replication, and the “replicability” requirement is reduced to demonstrating the possibility of replication.

Feminists have added their voice to critiques of science and the scientific method. The most successful feminist critiques of science are those identified as “feminist empiricist,” which attempt to restructure “bad science” and provide a more objective, gender-free knowledge (Harding 1986). Feminists have pointed out some androcentric (male-centered) categories in science and have identified the patriarchal social organization of “science as an institution.” Haraway has argued that there is no purely “objective” stance that can be taken; knowledge is always a “view” from somewhere (Haraway 1988). The concept of power based on gender has become a permanent category of analysis in feminist approaches (Smith 1987; Connel 1983).

By differentiating between “good science” and “bad science,” feminist empiricists strive to separate the wheat from the chaff by eradicating gender biases in the scientific process. The ultimate goal is to provide more objective, value-free knowl-

edge (Harding 1987). At the very least, feminist approaches often attempt to show the hidden biases in many scientific theories. The argument is that some types of knowledge are true only for certain social groups and do not reflect the experience of women, homosexuals, and many ethnic and racial groups, or other groups on the margins of society (Haraway 1988).

This perspective has had some success in the social sciences, perhaps because its revisions provide results that are intuitively appealing. By including categories that often are ignored, oppressed, and invisible to traditional sociology, feminist research gives a voice to what were previously “non-questions” under the mainstream, or as feminists call it, the male-stream model of science (Vickers 1982). For example, feminist research suggests that many women do not make a yes-or-no decision about having children but instead leave it to luck or time to decide. This type of decision-making behavior has implications for fertility and deserves the same theoretical status as the yes and no categories. However, a male-stream model of science that assumed that fertility decisions were the outcome of a series of rational cost-benefit analyses was blind to this conceptualization (Currie 1988).

It is ironic that while feminist empiricist criticisms of “bad” science aspire to strengthen science, they ultimately subvert the understandings of science they attempt to reinforce: “If the concepts of nature, of dispassionate, value free, objective inquiry, and of transcendental knowledge are androcentric, white, bourgeois, and Western, then no amount of more rigorous adherence to scientific method will eliminate such bias, for the methods themselves reproduce the perspectives generated by these hierarchies and thus distort our understandings” (Harding 1987, p. 291).

Another critique of science comes from the hermeneutic, or interpretive, perspective, which takes issue with the positivist assumption that the concepts, categories, and methods used to describe the physical world are applicable to human behavior. Human studies proponents insist that the universal categories and objective arguments required for prediction and explanation in the natural sciences cannot be achieved in the social sciences. The proper subject matter of the social sciences is the internal, or subjective, meanings of



human behavior that guide human action. Because these meanings are nonempirical and subjective rather than objective, they cannot meet the requirements for scientific explanation. Therefore, the goal of the social sciences is to understand rather than predict and explain human behavior (Hughes 1961; Habermas 1971; Gadamer 1976). Validation of interpretations is one of the biggest problems with the hermeneutic position because no firm ground exists from which to judge the validity of different interpretations of meaning and behavior. Hermeneutic explanations are ultimately subjective and in their extreme form focus solely on the explanation of individual, unique events (Alexander 1982).

The value-free nature of scientific knowledge also has been challenged by critical theory, which suggests that scientific knowledge is knowledge that is one-sided and specifically oriented to the domination and control of nature. This “interest” in domination and control does not lie in the application of scientific knowledge but is intrinsic to the knowledge itself. In contrast, communicative knowledge is knowledge that is oriented to reaching understanding and achieving human emancipation (Habermas 1984).

### CONCLUSION

Although scientific explanation has been the subject of many critiques, it is still the most methodical, reliable form of knowledge. It is ironic that while the natural sciences are becoming less positivistic and are beginning to recognize nonempirical, subjective, and cultural influences on scientific knowledge, the social sciences continue to emphasize the refinement of methodology and measurement in an attempt to become more positivistic (Alexander 1982). The result is that in sociology, theoretical inquiry is increasingly divorced from empirical research. Paradoxically, this schism may be a source of strength if the two sides can learn to communicate. Sociology may be in a unique position to integrate critiques of the scientific model with ongoing empirical research, perhaps producing a hybrid that is neither relativistic nor positivistic.

(SEE ALSO: *Causal Inference Models; Epistemology; Experiments; Measurement; Metatheory; Positivism; Quasi-Experimental Research Design; Reliability; Validity; Statistical Inference*)

### REFERENCES

- Alexander, Jeffrey C. 1982 *Positivism, Presuppositions and Current Controversies*. Berkeley: University of California Press.
- Babbie, Earl 1989 *The Practice of Social Research*, 5th ed. Belmont, Calif: Wadsworth.
- Bernstein, Richard J. 1976 *The Restructuring of Political and Social Theory*. Philadelphia: University of Pennsylvania Press.
- Blalock, Hubert M. 1964 *Causal Inferences in Non-Experimental Research*. Chapel Hill: University of North Carolina Press.
- 1979 *Social Statistics*, rev. 2nd ed. New York: McGraw-Hill.
- Braithwaite, Richard Bevan 1968 *Scientific Explanation*. London: Cambridge University Press.
- Chalmers, A. F. 1982 *What Is This Thing Called Science?* Milton Keynes, UK: Open University Press.
- Cohen, M., and E. Nagel 1934 *An Introduction to Logic and Scientific Method*. New York: Harcourt.
- Collins, H. M. 1985 *Changing Order. Replication and Induction in Scientific Practice*. Beverly Hills, Calif: Sage.
- \*Collins, H. M., and Pinch, T. 1998 *The Golem: What You Should Know about Science*. New York: Cambridge University Press.
- Connell, R. W. 1983 *Which Way Is Up?* Boston: Allen and Unwin.
- Cook, Thomas D., and Donald T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Currie, Dawn 1988 “Re-Thinking What We Do and How We Do It: A Study of Reproductive Decisions.” *Canadian Review of Sociology and Anthropology* 25:231–253.
- Fleck, Ludwik 1979 *Genesis and Development of a Scientific Fact*. Chicago: University of Chicago Press (originally published in German in 1935).
- Gadamer, Hans-Georg 1976 *Philosophical Hermeneutics*. Berkeley: University of California Press.
- Gartrell, John W., and C. David Gartrell 1979 “Status, Knowledge, and Innovation: Risk and Uncertainty in Agrarian India.” *Rural Sociology* 44:73–94.
- Giere, Ronald N. 1984 *Understanding Scientific Reasoning*, 2nd ed. New York: Holt, Rinehart and Winston.
- Glaser, Barney G., and Anselm L. Strauss 1967 *The Discovery of Grounded Theory*. Chicago: Aldine.
- Habermas, Jurgen 1971 *Knowledge and Human Interests*. Boston: Beacon Press.
- 1984 *The Theory of Communicative Action*. Boston: Beacon Press.

- Hannan, Michael T. 1970 *Problems of Aggregation and Disaggregation in Sociological Research*. Chapel Hill: University of North Carolina Press.
- Hanson, Norbert R. 1972 *Patterns of Discovery*. Cambridge, UK: Cambridge University Press.
- Haraway, Donna 1988 "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14:575-609.
- Harding, Sandra 1986 *The Science Question in Feminism*. Ithaca, N.Y.: Cornell University Press.
- 1987 "The Instability of the Analytical Categories of Feminist Theory." In Sandra Harding and Jean F. O'Barr, eds., *Sex and Scientific Inquiry*. Chicago: University of Chicago Press.
- Hempel, Carl 1966 *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice-Hall.
- Hughes, Stuart 1961 *Consciousness and Society*. New York: Vintage.
- Kaplan, Abraham 1964 *The Conduct of Inquiry: Methodology for Behavioral Science*. San Francisco: Chandler.
- Keat, R., and J. Urry 1982 *Social Theory as Science*. London: Routledge and Kegan Paul.
- Kerlinger, Fred N. 1973 *Foundations of Behavioral Research*, 2nd ed. New York: Holt, Rinehart and Winston.
- Knorr Cetina, Karin 1981 *The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science*. Oxford, UK: Pergamon Press.
- 1991 "Epistemic Cultures: Forms of Reason in Science" *History of Political Economy*, 23(1):105-122.
- 1992 "The Couch, the Cathedral, and the Laboratory: On the Relationship between Experiment and Laboratory in Science." In Andrew Pickering, ed., *Science as Practice and Culture*. Chicago: University of Chicago Press.
- 1995 "How Superorganisms Change: Consensus Formation and the Social Ontology of High-Energy Physics Experiments." *Social Studies of Science* 25:119-47.
- 1999 *Epistemic Cultures. How the Sciences Make Knowledge*. Cambridge and London: Harvard University Press.
- Kuhn, Thomas 1970 *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- \*Latour, Bruno 1987 *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, Mass.: Harvard University Press.
- , and Steve Woolgar 1986 *Laboratory Life*. Princeton, N.J.: Princeton University Press.
- Nagel, Ernest 1961 *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace and World.
- Pickering, Andrew 1982 "Interests and Analogies." In B. Barnes and D. Edge, eds., *Science in Context*. Milton Keynes, UK: Open University Press.
- Polanyi, Michael 1958 *Personal Knowledge*. Chicago: University of Chicago Press.
- Popper, Karl R. 1959 *The Logic of Scientific Discovery*. London: Hutchinson.
- Shapin, Steven 1979 "The Politics of Observation: Cerebral Anatomy and Social Interests in the Edinburgh Phrenology Disputes." In Roy Wallis, ed., *On the Margins of Science: The Social Construction of Rejected Knowledge*. Monograph 27, 1979 *Sociological Review*.
- Smith, Dorothy 1987 *The Everyday World as Problematic*. Boston: Northeastern University Press.
- Stevens, S. S. 1951 "Mathematics, Measurement, and Psychophysics." In S. S. Stevens, ed., *Handbook of Experimental Psychology*. New York: Wiley.
- Stinchcombe, Arthur L. 1987 *Constructing Social Theories*. Chicago: University of Chicago Press.
- Swinburne, R. G. 1964 "Falsifiability of Scientific Theories." *Mind*, 73:434-436.
- Vickers, Jill 1982 "Memoirs of an Ontological Exile: The Methodological Rebellions of Feminist Research." In Angela Miles and Geraldine Finn, eds., *Feminism in Canada: From Pressure to Politics*. Montreal: Black Rose.
- Wallace, Walter L. 1971 *The Logic of Science in Sociology*. Chicago: Aldine.
- Whyte, William Foote 1961 *Street Corner Society: The Social Structure of an Italian Slum*, 2nd ed. Chicago: University of Chicago Press.
- \*Woolgar, Steve 1988 *Science, the Very Idea*. London, New York: Tavistock.
- Woolgar, Steve 1981 "Interests and Explanation in the Social Study of Science." *Social Studies of Science* 11:365-394.

LINDA DERKSEN  
JOHN GARTRELL

## SCIENTIFIC METHOD

See Epistemology; Positivism; Scientific Explanation.

## SECONDARY DATA ANALYSIS AND DATA ARCHIVES

The creation and growth of publicly accessible data archives (or data banks) have revolutionized

the way sociologists conduct research. These resources have made possible a variety of secondary analyses, often utilizing the data in ways never anticipated by their creators. Traditionally, secondary data analysis involves the use of an available data resource by researchers to study a problem different from the one treated in the original analysis. For example, a researcher might have conducted a survey of workers' reactions to technological change and analyzed those data to evaluate whether the workers welcomed or resisted such change in the workplace. As a matter of secondary interest, the researcher collects data on workers' perceptions of the internal labor-market structures of their firms. She then lends those data to a colleague who studies the determinants of (workers' perceptions of) job-ladder length and complexity in order to understand workers' views on prospects for upward mobility in their places of employment. The latter investigation is a secondary analysis.

More recently, however, the definition of a secondary analysis has expanded as more data sets have been explicitly constructed with multiple purposes and multiple users in mind. The creators, or principal investigators, exercise control over the content of a data set but are responsive to a variety of constituencies that are likely to use that resource. The creators may undertake analyses of the data, addressing questions of intellectual interest to themselves while simultaneously releasing the data to the public or depositing the data resource in an archive. Data archives are depositories where data produced by a number of investigators are available for secondary analyses. The data bank generally takes responsibility for providing documentation on the data sets and other information needed for their use. The term also refers more generally to any source of data already produced that an investigator may uncover in the course of an investigation, such as government or business records housed in libraries. For example, the U.S. government archives thousands of government documents yearly in libraries around the world. The data in those documents cover a wide variety of topics and are often useful in sociological investigations. It remains the responsibility of the analyst to configure the data in a way that is useful to his or her investigation. This entry illustrates these expanded opportunities by describing one key data archive and indicating the extent and

breadth of data resources that this and other archives include. It then describes the process of conducting secondary analyses from resources such as these.

### DATA ARCHIVES AND DATA SOURCES

One of the most important data archives for social scientists is the Interuniversity Consortium for Political and Social Research (ICPSR) at the University of Michigan, Ann Arbor. The ICPSR publishes an annual *Guide to Resources and Services* (much of this description was taken from the 1996–1997 volume). Additional information is available at the ICPSR Web site ([www.icpsr.umich.edu](http://www.icpsr.umich.edu)). The consortium was founded in 1962 as a partnership between the Survey Research Center at the University of Michigan and twenty-one U.S. universities. In 1997 the holdings included over 3,500 titles, some of them capturing several panels of data on the same respondents or several waves of data involving comparable information. These titles are available to researchers at member institutions. The consortium charges fees on a sliding scale to academic institutions for membership privileges; researchers whose institutions are not members can obtain data for a fee. In 1997, over four hundred institutions in the United States, Canada, and countries throughout the world were members. While ICPSR originated as a service to political analysts, it currently serves a broad spectrum of the social sciences, including economics, sociology, geography, psychology, and history as well, and its data resources have been used by researchers in education, social work, foreign policy, criminal justice, and urban affairs.

Although ICPSR provides training in research and statistical methods and helps members in the effective use of computing resources, its central function is the archiving, processing, and distribution of machine-readable data of interest to social scientists. Although data capturing elements of the U.S. political process are well represented in its holdings, data are available on consumer attitudes, educational processes and attainment, health care utilization, social indicators of the quality of American life, employment conditions, workers' views on technology, and criminal behavior. The data come from over 130 countries, include both contemporary and historical censuses, and are not confined to the individual level but also provide

information on the characteristics of nations and organizational attributes. ICPSR actively seeks out high-quality data sets, and the user fees finance additional data acquisition as well as other operations. It also encourages investigators to deposit their data holdings in the archives to make them available to researchers for secondary analyses. Researchers whose data production efforts are funded by federal agencies such as the National Science Foundation are required to make their data publicly available after their grants have expired, and ICPSR is a logical depository for many data sets produced in the social sciences.

ICPSR maintains over ninety serial data holdings, including the earlier waves of the National Longitudinal Surveys of Labor Market Experience (NLS) (discussed below), the Survey of Income and Program Participation, the General Social Surveys, National Crime Surveys, the Panel Study of Income Dynamics, the Detroit Area Studies, the U.S. Census of Population and Housing, and the American National Elections Studies. These serial holdings include longitudinal surveys (in which the same respondents are interviewed repeatedly over time) such as the NLS and the Panel Study of Income Dynamics. These resources are particularly useful in determining the impact of earlier life events on later life outcomes, since the causal orders of all events measured on the data sets are clearly indicated. The holdings also include sets of cross-sectional studies conducted at regular intervals, such as the Detroit Area Studies and the General Social Surveys (GSS). These studies contain different cross sections from the same populations over time and are useful in charting trends in the attitudes of the respective populations over time, assuming that the same questions are repeated. Sources, such as the GSS, that ask the same questions over several years allow the researcher to pool samples across those years and obtain larger numbers of cases that are useful in multivariate analyses.

To illustrate one data set, consider the National Longitudinal Surveys of Labor Market Experience. These surveys are produced by the Center for Human Resource Research (CHRR) at Ohio State University. The CHRR produces a yearly *NLS Handbook*, and much of the following information regarding the NLS was taken from the 1998 *NLS Handbook*. These surveys began in 1966 with a study of older men aged 45–59 and a

survey of young men aged 14–24, continued in 1967 with a survey of mature women aged 30–44, and were followed up with a survey of young women aged 14–24 in 1968. In 1979, CHRR began a survey of over 12,000 youths aged 14–22, known as the NLSY79. In 1997, CHRR surveyed a new cohort of over 9,000 youths aged 12–16, called the NLSY97, and is continuing with yearly surveys of this cohort. The six major surveys contain a wealth of data on labor-force experience (e.g., labor-force and employment status, work history, and earnings) as well as investment in education and training, marital status, household composition and fertility, background material on respondents' parents, work-related attitudes, health, alcohol and drug use, and region of residence.

Each of these cohorts has been followed at varying intervals since the surveys' inception. For example, the Young Women were surveyed nineteen times between 1968 and 1997. The NLSY79 respondents were surveyed every year until 1994, when surveys in even-numbered years began. The Older Men were surveyed every year until 1983, and they or their widows were resurveyed in 1990. Data production for the Older Men and Young Men is complete; data production for the Mature Women and Young Women is ongoing biennially. In 1986 the NLS added a survey of the children of the NLSY79 cohort's women; that described the social, cognitive, and physiological development of those children and, given the longitudinal nature of the data on the mothers, allows an explanation of these child outcomes in terms of maternal background and current maternal characteristics. Surveys of the children occur in even-numbered years; this accumulated longitudinal database on child outcomes allows important inferences regarding the process of child development, with the numbers of children surveyed far exceeding those in most other sources. This additional resource has expanded NLSY79's usefulness to other disciplines, including psychology, and to other researchers interested in child development.

The NLS data sets are produced with the cooperation of CHRR, NORC (formerly the National Opinion Research Center) at the University of Chicago, and the U.S. Bureau of the Census. For example, for NLSY79, the CHRR takes responsibility for questionnaire construction, documentation, and data dissemination, while NORC has handled sample design, fieldwork, and data reduc-

tion. The Census Bureau has handled sample design, fieldwork, and data reduction for the four original cohorts. All data are available on CD-ROM from CHRR. Waves of data prior to 1993 are also available from ICPSR, as was noted above.

Social scientists from several disciplines, including sociology, economics, and industrial relations, have found the NLS to be a critical resource for the study of earnings and income attainment, human capital investment, job searches, fertility, racial and sex discrimination, and the determinants of labor supply. Inferences from these studies have been useful in regard to theory as well as policy formation. Other topics the data resource can usefully inform include family structure and processes, child outcomes, and aging processes. The CHRR estimates that by 1998 over 3,000 articles, books, working papers, and dissertations were produced using the NLS data. The 1998 *NLS Handbook* provides a wealth of detail regarding the designs of the surveys, survey procedures, variables, and CD availability. It also describes the extensive documentation available on the NLS data sets and lists references to key Web sites, including one that contains NLS publications. This handbook is indispensable for any researcher considering a secondary analysis using NLS data. The CHRR at Ohio State University disseminates the data and provides documentation and assistance to users with questions about the data sets. This summary gives a glimpse of the tremendous potential for secondary analyses of NLS data; this potential is multiplied many times over when one considers the number of other data sets available to researchers.

Because of the increase in resources devoted to survey research in sociology and related social sciences, the ICPSR holdings containing surveys of individuals have grown rapidly. However, ICPSR also archives data produced at varying levels of aggregation, thus facilitating secondary analyses in which the theoretically appropriate units of analysis are countries or organizations. For example, ICPSR archives the World Tables of Economic and Social Indicators, 1950–1992, provided by the World Bank. These data contain economic and social indicators from 183 countries, with the indicators including measures such as gross national product, value of imports and exports, gross national savings, value added across major industrial categories, net direct foreign investment, public

long-term debt, international reserves excluding gold, and gold holdings at the London market price. Demographic and social variables include population, total fertility rate, crude birthrate, percentage of the labor force in agriculture, percentage of the labor force that is female, and primary and secondary school enrollment rates. An older data set, also from the World Bank, contains similar measures from 1950 to 1981 as well as additional indicators not included in the data set covering the 1950–1992 period. Because these are also longitudinal data sets, there is the potential for pooling across time variation in these measures across the countries so that cross-sectional and longitudinal variations can be studied simultaneously.

ICPSR also maintains a small number of holdings useful for studying organizational processes. For example, a 1972 study of industrial location decisions obtained from the Economic Behavior Program of the Survey Research Center at the University of Michigan surveyed 173 industrial plants in Detroit, Chicago, and Atlanta. The interviewees were organizational informants such as president, vice president, general manager, and public relations director. The items included reasons for the location of the plant and the advantages and disadvantages of a location; other constructs measured included duration of plant operations, levels of sales and production, production problems, and plans for future expansion.

More recent arguments, however, have suggested that although sociology has invested considerably in surveys of individuals, it has invested insufficiently in surveys of organizations (Freeman 1986; see also Parcel et al. 1991). Kalleberg et al. (1996) present results from the National Study of Organizations, a National Science Foundation-sponsored study of a representative cross section of organizations that addresses their structures, contexts, and personnel practices. Although they demonstrate the utility of this design for addressing some questions regarding organizational functioning, these data cannot address issues of organizational change. A possible solution would be to produce a longitudinal database of organizations. The characteristics of a representative sample of organizations would be produced across time, analogous to the panel data sets of individual characteristics described above. Such a resource would enable researchers to study processes of organiza-

tional change with models that allow a clear causal ordering of variables. This type of resource also would permit analyses of pooled cross sections. Most important, the resource would allow organizational theories to be subjected to tests based on a representative sample of organizations, in contrast to the purposive samples that are used more frequently. To date, the resources have not been sufficient to approach the panel design suggested above. Clearly, the capacity to conduct secondary analyses at the organizational level is in its infancy relative to studies of individual-level processes and phenomena.

Finally, ICPSR also archives a variety of data sets that make possible historical analyses of social, economic, and political processes. For example, it archives the Annual Time Series Statistics for the United States, 1929–1968, which includes 280 variables for most of that period, although only 127 variables are available for the period 1947–1968. Available data include population characteristics, measures of political characteristics of the U.S. Congress, business and consumer expenditures, and expenditures by various federal government departments. ICPSR also archives Political Systems Performance Data for France, Sweden, and the United States, 1950–1965, in which the central constructs measured include size of public debt, gross national product (GNP), energy consumption, income tax rates, birthrates and death rates, labor force and unemployment, voting behavior, urbanization, and agricultural growth. Each of these historical data sources makes possible time series analyses of the macro-level phenomena they measure.

Additional major archives include the Roper Center for Public Opinion Research at the University of Connecticut and the Lewis Harris Data Center at the University of North Carolina at Chapel Hill. Kiecolt and Nathan (1985) provide additional information on the major archives, and Stewart (1984) outlines the extensive holdings in U.S. Government Document Depositories, especially the products of the U.S. Bureau of the Census. Other important archives include several in Europe with which ICPSR maintains a relationship, such as the Norwegian Social Science Data Services, the Australian Social Science Data Archives, and the Zentralarchiv für empirische Sozialforschung (ZA) at the University of Cologne. There is the potential for member institutions to

obtain from ICPSR data contained in those local archives as well. The International Social Survey Program (ISSP) has worked toward coordinating survey research internationally by asking common questions cross-nationally in given years, facilitating cross cultural analyses of social phenomena. For example, in 1990 social surveys in Austria, West Germany, Great Britain, Hungary, Ireland, Israel, Italy, the Netherlands, and Norway all included questions on work, including the consequences of unemployment, union activities, working conditions, and preferred job characteristics. A comparable module in 1987 focused on social inequality in Australia, Austria, West Germany, Great Britain, Hungary, Italy, and the United States. The 1993 module focused on nature, the environment, recycling, and the role of science in solving environmental problems. Data from the ISSP are available from ICPSR.

#### **THE NATURE OF SECONDARY DATA ANALYSIS**

The key advantage of secondary data analysis is also the key disadvantage: The researcher gains access to a wealth of information, usually far in excess of what he or she could have produced with individual resources, but in exchange must accept the myriad operational decisions that the investigators who produced the data have made. On the positive side, the researcher frequently is able to take advantage of a national sample of respondents or data produced on national populations when individual resources would have supported only local primary data production. The numbers of cases available in secondary resources often far outstrip the sample sizes individual investigators could have afforded to produce; these large sample sizes enhance the precision of parameter estimates and allow forms of multivariate analyses that smaller sample sizes preclude. A secondary analyst also can take advantage of the significant expertise concentrated in the large survey organizations that produce data sets for secondary analysis. This collective expertise usually exceeds that of any single investigator. Despite these advantages, the researcher must carefully match the requirements of the research project to the characteristics of the data set. When the match is close, the use of secondary data will enhance the research effort by making use of existing resources and taking advantage of the time, money, and expertise of others

devoted to data production. If the match is poor, the research project will fail because the data will not address the questions posed.

Because many secondary analyses are conducted on survey data, effective use of secondary survey sources frequently depends on knowledge of sample design, question wording, questionnaire construction, and measurement. Ideally, the researcher conceptualizes precisely what he or she wishes to do with the data in the analysis, since analytic requirements must be met by existing data. If the research questions posed are longitudinal in nature, the researcher must be sure that the survey questions are measured at time points that mirror the researcher's assumptions of causal order.

The researcher also must be certain that the survey samples all the respondents relevant to the problem. For example, analyses of racial differences in socioeconomic outcomes must use data sets in which racial minorities are oversampled to ensure adequate numbers of cases for analysis. The researcher also must be certain that a data set contains sufficient cases for the analysis she or he intends to perform. Kiecolt and Nathan (1985) stress the challenges for trend and cross-cultural studies that result from changes in sampling procedures over time. For example, suppose a researcher wants to ascertain whether more people support a voucher system for public education in 2000 compared with 1990. Changes in the sampling frame over the decade may introduce variations into survey responses that would not otherwise exist. These variations can be in either direction, and hypotheses regarding their direction are a function of the nature of sampling changes. Gallup surveys have increased their coverage of noninstitutionalized civilian adult populations over time, with the result that there has been an artifactual decrease in the levels of education they report (Kiecolt and Nathan 1985, pp. 62–63), since the later surveys have progressively included groups with lower levels of schooling. Sampling changes also can occur over time because of changes in geographic boundaries. Cities change boundaries owing to annexation of areas, and Metropolitan Statistical Areas (MSAs, formerly Standard Metropolitan Statistical Areas [SMSAs]) are created over time as increased numbers of counties meet the population and economic criteria for defining MSAs.

The most common problem in conducting secondary analyses, however, occurs in the questionnaire coverage of items needed to construct appropriate measures. It is likely that the original survey was constructed with one purpose and asked adequate numbers and forms of questions regarding the constructs central to that problem but gave only cursory attention to other items. A secondary researcher must evaluate carefully whether the questions that involve his or her area of central interest are adequate for measurement and for analytic tasks. The biggest fear of a secondary researcher is that some variables needed for proper model specification have been omitted. Omitted variables pose potentially severe problems of misspecification in estimating the parameters of the variables that are included in the models. In these cases the researcher must decide whether an adequate proxy (or substitute) variable exists on the data set, whether the research problem can be reformulated so that omission of that construct is less critical, or whether the initially chosen data set is unsuitable and another must be sought. Researchers can also purchase time on major social surveys such as the GSS administered by NORC. This strategy enables researchers with adequate financial resources to be certain that the questions needed to investigate the issues of interest to them will be included in a national survey. This strategy mixes primary data production with secondary analysis of a multipurpose data set. The entire data resource then becomes available to other secondary analysts.

Other challenges for secondary analysts occur as a function of the particular form of secondary analysis used. For example, Kiecolt and Nathan (1985) note that survey researchers who produce series of cross sections of data that are useful in studying trends may "improve" the wording of questions over time. In regard to the problem of voucher systems in public education, the researcher may observe increased percentages of survey respondents favoring this option over the period covered by the surveys but still may have difficulty eliminating the possibility that question wording in the later survey or surveys may have encouraged a more positive response. Such changes also can occur if the wording of the question remains the same over time but the nature of the response categories changes. Secondary analysts who conduct cross-cultural comparisons must be sensitive

to the fact that the same question can mean different things in different cultures, thus interfering with their ability to compare the same social phenomenon cross-culturally.

Dale, et al. (1988) note that in-depth studies of specific populations may be most realistic with national samples that provide sufficient cases for analyses of the subgroups while allowing the researcher to place those data within a broader empirical context. It is also possible that surveys produced by different survey organizations will produce different results even when question wording, response categories, and sampling procedures remain the same (Kiecolt and Nathan 1985, p. 67). A secondary analyst must be certain that the survey organization or individual responsible for producing the data set exercised appropriate care in constructing the data resource. As was noted above, detailed familiarity with the documentation describing the data set production procedures is essential, as is a codebook indicating frequencies on categorical variables, appropriate ranges for continuous variables, and codes for missing data.

There is often an interactive nature to the process of conducting a secondary data analysis. While the researcher's theoretical interests may be reasonably well formulated when he or she identifies a useful data set, the variables present in the data resource may suggest additional empirical opportunities of theoretical interest that the researcher had not previously considered. Also, familiarity with data resources can facilitate the formulation of empirical investigations that otherwise might not be initiated. Once a researcher is familiar with the features of a particular secondary source, accessing additional variables for the analysis of a related problem may require less investment than would accessing a new data resource. However, there is general agreement that data availability should never dictate the nature of a research question. Although it is legitimate for a researcher to use his or her awareness of data resources to recognize that analyses of problems of long-standing interest are now empirically possible, "data dredging" has a deservedly negative connotation and does not result in the advancement of social science. Hyman's (1972) classic treatment of secondary analyses of survey data richly chronicles the experiences of a number of sociologists as they interactively considered the matching of theoretic

cal interests and data availability in formulating and conducting secondary analyses. He also describes a number of ways in which secondary analysts can configure existing data to test hypotheses.

Recent developments in technology have streamlined several steps in secondary analyses that formerly were time-consuming and labor-intensive. Many secondary data sets are now available on CD-ROM (compact disk-read only memory); the NLS data discussed above are only one example. With many computers having attached CD readers, analysts can read the disks and extract from them the variables and cases they wish to study. Often the disks also contain searching devices that enable researchers to locate variables of interest easily. These "search engines" simultaneously enable analysts to select a sample and obtain the variables needed on each case. These capabilities totally bypass older technologies involving nine-track tapes containing data. In tape-based technologies, analysts had to write original computer programs to extract the needed variables and cases. A typical analyst no longer depends on a centralized computing facility for storing, mounting, and reading magnetic tapes.

The next steps in secondary analysis differ only slightly from the steps that investigators who produce primary data undertake. In both cases, data must be cleaned to remove coding errors that might result in erroneous findings. Similarly, both investigators need to address problems with missing data. The primary data producer is close enough to the actual data production not only to identify such problems but also to resolve many of them appropriately. For example, if the researcher is studying a single organization and notes that a respondent has failed to report his or her earnings, the researcher, knowing the respondent's occupation, may be able to obtain data from the organization that approximates that respondent's earnings closely. The secondary analyst would not have access to the original organization but might approximate the missing data by searching for other respondents who reported the same occupation but who also reported earnings. Variations on this theme involve the imputation of missing data by using mathematical functions of observed data to derive reasonable inferences about values that are missing (Little and Rubin 1987, 1990; Jinn and Sedransk 1989).



Both types of investigator have to be familiar with the descriptive properties of their data. For a primary investigator, observing distributions of respective variables as well as their central tendencies should be an outgrowth of data production itself. A secondary analyst has less familiarity with the data someone else produces but is under the same obligation to become familiar with the descriptive properties of the data in a detailed way. For both researchers, good decisions involving measurement of variables and model specification for multivariate analyses depend on knowledge of the descriptive properties of the data.

Within the respective multipurpose data sets, research traditions often arise from the sometimes unique suitability of certain resources for addressing given problems. These traditions derive from the fact that several investigators have access to the data simultaneously, a feature that distinguishes secondary data analysis from analyses undertaken by different primary investigators, each of whom has a unique data set. For example, in the late 1980s and into the 1990s, the NLSY79 with Mother and Child Supplements was virtually unique in combining a large sample size, longitudinal data on maternal familial and work histories, observed child outcomes, and oversamplings of racial minorities. Problems tracing the impact of maternal events on child outcomes are addressable with this data resource in a way that they were not with other resources. Investigators with an interest in these issues use the data and exchange information regarding strategies for measuring constructs and data analysis and then exchange their findings. Over time, bodies of findings emerge from common data sources where the findings are contributed by a number of secondary investigators, although the particular problems, theoretical frameworks, and empirical strategies represented in each one may differ markedly. As was suggested above, multipurpose data sets frequently allow secondary analyses by researchers from several disciplines. The products of these investigations bear the stamps of their respective disciplines. In addition, the NLSY79 with Mother and Child Supplements has served as a model for the Michigan Panel Study of Income Dynamics (PSID) in its 1997 Child Development Supplement on the PSID respondents. This new data resource, which combines longitudinal data on parents and developmental assessments of children from birth to age

12, will enable replication of key findings produced with the NLSY79 child data set as well as the production of new findings. For example, both data sets contain age-appropriate cognitive assessments for children, permitting findings produced with the NLSY79 child data set to be replicated with the PSID Child Development Supplement. The PSID, however, contains data on how children spend their time. These variables should allow researchers to understand the effects of children's time use on several developmental outcomes, something that the NLSY79 child data do not permit.

The wealth of secondary data sources also permits investigators to use more than one data source to pursue a particular line of inquiry. No single data set is perfect. Researchers can analyze several data sets, all with key measures but each with unique strengths, to check interpretations of findings and evaluate alternative explanations. McLanahan and Sandefur (1994) use this approach in their study of the effects of single parenthood on the offspring's academic success and social adjustment. Their data sources include the NLSY79, the PSID, and the High School and Beyond Study. The result is a stronger set of findings than those which could have been produced with any one of those sources.

Another model for conducting secondary research is suggested by researchers who use census data produced by the U.S. Department of Commerce. Census holdings cover not only information on the general U.S. population but also data on businesses, housing units, governments, and agricultural enterprises. Researchers who use these sources singly or in combination must be familiar with the questionnaires used to produce the data and with the relevant features of sample coverage. While some census data are available on machine-readable tape, other data exist only in printed form. In these cases, the researcher must configure the needed data into a form suitable for analyses, in many cases a rectangular file in which cases form row entries and variables form column entries. Data produced on cities from the County and City Data Books, for example, allow a variety of analyses that involve the relationships among urban social and economic characteristics. In these analyses, the unit of analysis is probably an aggregate unit such as a county or city, illustrating the applicability of secondary analysis to problems

conceptualized at a level of aggregation higher than that of the individual.

Another advantage of secondary analyses is the potential for those most interested in a particular set of findings to replicate them by using the same data and to introduce additional variables or alternative operationalizations as a method for evaluating the robustness of the first secondary investigator's findings. A classic example is Beck et al.'s 1978 investigation of differences in earnings attainment processes by economic sector. Hauser's (1980) reanalysis of those data suggested that most of the differences in sectoral earnings reported in the original study were a function of coding decisions for low-earnings respondents, since the differences disappeared when the code for low earnings was changed. Despite this criticism, the impact of the original investigation has been enormous, with many additional investigators exploring the structure and implications of economic sectors. The point, of course, is that such debate is more likely to occur when researchers have access to common data sets, although gracious investigators often lend their data resources to interested critics. Hauser (1980) acknowledges that Beck et al. shared their original data, although he could have obtained the original data set from ICPSR.

Secondary data sets can be augmented with additional data to enrich the data resource and allow the derivation of additional theoretical and empirical insights. Contextual analysis, or the investigation of whether social context influences social outcomes, is a key example. Parcel and Mueller (1983) used the 1975 and 1976 panels from the PSID to study racial and sex differences in earnings attainment. To evaluate the impact of occupational, industrial, and local labor-market conditions on workers' earnings, they augmented the PSID data with archival data from U.S. Census and Dictionary of Occupational Titles sources that were based on the occupations, industries, and local markets of respective PSID respondents. Illustrative contextual indicators included occupational complexity, industrial profitability, and local-market manufacturing-sector productivity. Analyses then suggested how these contextual, as well as individual-level, indicators affected workers' earnings differently depending on ascriptive statuses. Computer software is now available to correct for problems in estimating models that use contextual data.

The potential for many sociologists to use secondary analysis to conduct studies of theoretical and practical importance probably has contributed to a change in productivity standards in sociology, particularly in certain subfields. The fact that certain issues can be addressed by using existing data can result in enormous savings in time relative to the time that would be required if primary data had to be produced. Research-oriented departments either implicitly or explicitly take this into account in assigning rewards such as salaries, tenure, and promotion. The potential for secondary analyses thus may create pressures toward increased scientific productivity; whether these pressures work generally for the good of social science or against it may be a matter of debate.

It is undeniable that progress in addressing some of the most important problems in social science has been facilitated greatly by the existence of multipurpose data sets and secondary resources. It is also true that the resources needed to produce and disseminate these data are considerable and that the existence and continuation of these resources are vulnerable to changes in political climate and priorities when those priorities influence resource allocation. It is critical that such decisions on resource allocation, particularly those made at the level of the federal government, recognize the important role that secondary resources have played in furthering both basic social science and applications informing social policy.

(SEE ALSO: *Census, Social Indicators, Survey Research*)

#### REFERENCES

- Beck, E. M., Patrick Horan, and Charles W. Tolbert II 1978 "Stratification in a Dual Economy: A Sectoral Model of Earnings Determination." *American Sociological Review* 43:704-720.
- Dale, Angela, Sara Arber, and Michael Proctor 1988 *Doing Secondary Analysis*. London: Unwin Hyman.
- Freeman, John 1986 "Data Quality and the Development of Organizational Social Science: An Editorial Essay." *Administrative Science Quarterly* 31:298-303.
- Guide to Resources and Services, 1996-1997*. Ann Arbor, Mich.: Interuniversity Consortium for Political and Social Research, University of Michigan.
- Hauser, Robert 1980 "Comment on 'Stratification in a Dual Economy.'" *American Sociological Review* 45:702-712.

- Hyman, Herbert H. 1972 *Secondary Analysis of Sample Surveys: Principles, Procedures, and Potentialities*. New York: Wiley.
- Jinn, J. H. and J. Sedransk 1989 "Effect on Secondary Data Analysis of Common Imputation Methods." *Sociological Methodology* 19:213–241.
- Kalleberg, Arne L., David Knoke, Peter V. Marsden, and Joe L. Spaeth 1996 *Organizations in America: Analyzing Their Structures and Human Resource Practices*. Thousand Oaks, Calif.: Sage.
- Kiecolt, K. Jill, and Laura E. Nathan 1985 "Secondary Analysis of Survey Data." Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-053. Beverly Hills, Calif.: Sage.
- Little, Roderick J. A. and Donald B. Rubin 1987 *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.
- 1990 "The Analysis of Social Science Data with Missing Values." Pp. 375–409 in John Fox and J. Scott Long, eds., *Modern Methods of Data Analysis*. Newbury Park, Calif.: Sage Publications.
- McLanahan, Sara, and Gary Sandefeur 1994 *Growing Up with a Single Parent: What Hurts, What Helps*. Cambridge, Mass.: Harvard University Press.
- NLS Handbook 1998*. Columbus, Ohio: Center for Human Resource Research, Ohio State University.
- Parcel, Toby L., and Charles W. Mueller 1983 *Ascription and Labor Markets: Race and Sex Differences in Earnings*. New York: Academic Press.
- , Robert L. Kaufman, and Leeann Jolly 1991 "Going Up the Ladder: Multiplicity Sampling to Create Linked Macro-Micro Organizational Samples." *Sociological Methodology, 1991*. 21:43–80.
- Stewart, David W. 1984 *Secondary Research: Information Sources and Methods*. Beverly Hills, Calif.: Sage.
- U.S. Department of Commerce [various years] *County and City Data Book*. Washington, D.C.: U.S. Government Printing Office.

TOBY L. PARCEL

## SECULARIZATION

Secularization is the process by which the sacred gives way to the secular, whether in matters of personal faith, institutional practice, and political power. It involves a transition in which things once revered become ordinary, the sanctified becomes mundane, and the otherworldly loses its prefix. Whereas the term "secularity" refers to a state of

sacredlessness and "secularism" is the ideology devoted to that state, secularization is a historical dynamic that may occur gradually or suddenly and may be replaceable (if not reversible).

The concept of secularization has been both an organizing theme and a source of contention among scholars of religion since the beginning of the European "Enlightenment" in the seventeenth century. One might expect an increasing consensus on a matter so long on the scholarly agenda, but discord has crescendoed in recent years. Secularization has taken on different meanings in different camps. It matters whether the reference is to religion's displacement, decline, or change; to the sacred at the level of the individual, the institution, the community, or the culture; or to a pattern that is long term, linear, and inevitable or short term, cyclical, and contingent.

The object of this essay is to disentangle both the issues and the combatants. After describing the early protagonists and more recent sociological proponents of secularization, this article considers recent arguments against their theses. In the face of a seemingly intractable conflict, it is important to describe the issues in dispute. This will lead to a consideration of secularization and sacralization as opposite phenomena that actually are more mutually linked than mutually exclusive.

### EARLY AND RECENT CONCEPTIONS OF SECULARIZATION

Any conception of the sacred is likely to engender skeptical—if often marginal—detractors. While both the process and the thesis of secularization have precursors early in Western history, it was the Enlightenment that provided their first codification.

The term "secularization" dates back to France in the mid-seventeenth century. The first high priest of this antichurch was the French bourgeois intellectual Voltaire (1694–1778). A professed "deist" whose belief in impersonal forces stood in sharp contrast to "theistic" conceptions of a personal God, Voltaire railed against the Catholicism's superstitions and ecclesiastical trappings (Voltaire 1756). However, Voltaire was not the most materialist figure of his day and he was distinguished more by the expression of his views than by their substance, including his sense that the end of religion was near, possibly in his life-

time. The main thrust of his views was shared by many Europeans and Americans, including Benjamin Franklin and Thomas Jefferson.

The prophets of secularization soon multiplied. By the second half of the nineteenth century, they included the father or at least namer of “sociology,” the French positivist Comte (1852). Comte’s conception of a future that belonged more to the social sciences than to religion was shared by Britain’s Spencer (1874), whose sales rivaled those of Dickens. Marx ([1844] 1963) envisioned a denarcotized future once the masses learned the real secret of their misery, substituted class consciousness for false consciousness, and exchanged otherworldly sighs for this-worldly action.

Weber and Durkheim continued the tradition in the first two decades of the twentieth century. Both provided key statements about the importance of religion: Weber’s “Protestant ethic” as a precondition of capitalism and Durkheim’s conception of religion as the latent worship of society. However, neither was personally religious, and both envisioned a secularized future without predicting it directly.

For Weber ([1905] 1993), secularization was an implication of the “rationalization” that was uniquely characteristic of the West. He was ambivalent about the results. On the one hand, he appreciated its cultural underpinnings of everything from capitalism and bureaucracy to architecture and music. On the other hand, he wrote in the tradition of German historiography and a concern for the spirit of every age. Weber lamented a dark side of rationality that would lead to secularized disenchantment. Toward the end of *The Protestant Ethic and the Spirit of Capitalism*, he commends the cynical sentiment:

*Specialists without spirit, sensualists without heart; this nullity imagines that it has attained a level of civilization never before achieved.*  
([1905] 1993, p. 182)

Durkheim worked selectively within the tradition of Comte and French positivism and therefore was more positive about secularization. Durkheim (1961) was optimistic about a secular morality and an autonomous ethic for society. Although religious beliefs would be displaced by science, the sense of society as a sacred collectivity would remain. On the eve of World War I, he

described France as undergoing a period of “moral mediocrity,” but he was certain that it would soon be revitalized through a sense of “collective effervescence” and sacred renewal, possibly independently of conventional religion (Durkheim 1912).

By the middle of the twentieth century, secularization had become one of the master motifs of the social sciences. It was at least implicit in major transitional distinctions such as Durkheim’s “mechanical versus organic solidarity” Toennies’s ([1887] 1957) “Gemeinschaft” versus “Gesellschaft” societies, and Redfield’s (1953) “folk” versus “urban” cultures. At the same time, prophecies had given way to theories as sociology began to develop more nuanced versions of secularization. The 1960s produced a bumper crop of new works, among the most influential of which were Berger’s *The Sacred Canopy* (1967) and Wilson’s *Religion in Secular Society* (1966).

Berger dealt with both the rise and the decline of religion. Having described religion’s importance as a source of meaning for a cosmos that is often inchoate, he then noted factors involved in religion’s erosion. These included privatization, pluralism, and a new religious marketplace, all of which contributed to a secularization he defined as “the process by which sectors of society and culture are removed from the domination of religious institutions and symbols” (1967, p. 107). Berger did not place all the blame for the decline of religion on external factors. Liberal clergy and theologians were often ahead of the process in diluting religion to avoid conflicts with a secular society.

If Berger’s conception of secularization suggests society pulling away from a still-religious core, Wilson conveys a scenario in which religion recedes to the margins and suffers a diminution of influence. For Wilson, secularization is “the process whereby religious institutions, actions, and consciousness lose their social significance” (1966, p. xiv). However, Wilson was aware of a profound difference between the declining influence of the established churches and the surging growth of sectarian movements (Wilson 1998): As society becomes more complex, all its institutions become more differentiated from each other and have more autonomy but less influence. However, the process does not occur equally, and traditional institutions such as religion are more

affected by these changes. Often seen as part of a larger process of “modernization,” differentiation has been a prominent theme among functionalists such as Parsons (1977) and Luhmann (1982) and neofunctionalists such as Bell (1976) and Habermas (1988).

Differentiation takes different forms and exacts different tolls. The Belgian scholar Dobbelaere (1981) draws a parallel between secularization and the French term “laicization,” which Durkheim and others used to denote a loss of priestly control, with a consequent decanonization of religion. While developing the concept for European settings, Dobbelaere draws two sets of distinctions: between the processes of differentiation, decline, and change (1981) and between the levels of the individual, the organization, and the society (Dobbelaere 1985).

By this time, secularization had become a major priority for social scientists examining religion. In analyzing the United States, Fenn (1979) stresses that secularization involves a blurring rather than a sharpening of the boundaries between the sacred and the secular; more recently, Fenn refers to secularization as the “domestication of charisma” (1993). Meanwhile, the concept is at least a subtheme of Bellah et al. (1985) in a work that depicts the community’s losing struggle with individualism, perhaps the ultimate form of differentiation at the personal level.

Roof and McKinney (1987) describe a similar pattern as a “new voluntarism” that has displaced old denominational loyalties. Similarly, Wuthnow (1988) notes how other forces of differentiation have shifted religious action away from the denominations and congregations and in the direction of “special-purpose groups” whose single-issue agendas are often more a reflection of political morality than of religious doctrine or theology. Wuthnow also describes a differentiation between America’s liberal and conservative “civil religions” and the rise of a third national faith in the form of secular technology.

Finally, Chaves (1993) documents the emergence of differentiated “dual structures” within denominations. This duality represents a split between declining “religious” authority and increasing secular “agency” authority. This formulation is consistent with other traditions of organizational analysis in religion, including the classic

distinction between “sects” and “churches” and the process by which the purity of sects is compromised by their transformation into accommodating churches.

### SECULARIZATION MYTHOLOGIZED

Originally, the detractors of secularization were defenders of the faith. More recently, they have portrayed themselves as critics of a very different faith, which they have played a large role in constructing. Recent years have seen the attribution of a full-blown “secularization thesis” that is not so much a series of questions for investigation as a definitive answer with all the qualities of an epochal narrative. Here the older eighteenth-century prophetic vision of secularization has been substituted for more recent and less sweeping versions. Secularization is presented as a tenuous article of faith that is suspended between two mythical points. The first point involves the fiction of a deeply and universally religious past; the second involves the conceit of a religionless present and future (Stark 1992). Thus, secularization has been recast as a sweeping saga that serves as a sort of antisacred doctrine, in its own right—though it is important to bear in mind that it is the critics of secularization who have both popularized this version and savaged it.

The British anthropologist Douglas (1982) was among the first to chastise proponents of secularization for imagining a mythical past against which the present inevitably comes up short. In her case, the past involved those simple, undifferentiated societies studied by anthropologists but used by others as convenient foils. Thus, even here religious piety and participation are not always deep or universal. If these societies are the beginnings of the neoevolutionary process of modernization, their religion has inconvenient similarities with the religion of complex societies toward the end of the process.

Stark (1998) elaborates this point for early Western societies. To the extent that a secularizing trend depends on a contrast with a pious ancient and medieval Europe, Stark cites evidence suggesting that this past is also mythical. Once one looks beyond the public displays of ecclesiastical officialdom, the masses appear to be antichurch, if not antireligious. Attitudes toward organized faith were conspicuous for their alienation, corruption, and

raucousness. Many “Christian” nations founded in the late middle ages were only inches deep as surface monopolies atop an impious base.

What of the myth of religious demise? Martin (1969) was among the first to find religion in the midst of putative nonreligion, in this case in “highly secularized” Great Britain. In fact, Martin called for dropping the term “secularization” because of the confusion it had elicited, though ten years later he adopted the semantic fashion by publishing *A General Theory of Secularization* (1978).

Stark has also been a relentless critic of the second myth, and he has had company. His book with Finke, *The Churching of America* (Finke and Stark 1992), uses actual and reconstructed church membership data to argue that the real “winners” over the past two centuries have been conservative churches while liberal (and more secular) churches have been the “losers.” Critics note that the work is not without problems; for example, its thesis refers to rates of growth and decline rather than absolute size, and it assumes that membership is a reliable measure of general religiosity over time (Demerath 1992).

Many other scholars have noted the continued vitality of religion in America. Warner’s “new paradigm” (1993) provides a systematic description of how the American case may differ from the European scene that spawned secularization theory. Meanwhile, Stark has taken his methods and “market” model of religion abroad. He and the economist Iannaccone (1994) developed a nonmonopolistic, “supply-side” interpretation of European religion, arguing that its death and secularization have been greatly exaggerated. This argument has had both supporters (Davie 1994; Swatos 1997) and detractors (Bruce 1995; Dobbelaere 1993; Wilson 1998).

Meanwhile, the dispute over secularization is not restricted to the West. In fact, the Western version of the debate is comparatively innocuous because it is confined largely to scholars removed from political conflicts and because the politics of religion has generally been laid to rest except in a few cases, such as the tragic violence in Northern Ireland and the anticlimactic decision of Sweden to sever state ties with the Lutheran Church as of 2000. Once one leaves the West, however (Demerath 2000), assessments of secularization and secularity have become volatile public issues exacerbated by

the ideological conflict between forthright pro- and antisecularists.

Moving from Poland and eastern Europe through the remains of the Soviet Union to Afghanistan, from the Balkans through Turkey and into Iran, from Algeria through Egypt to Israel, from Pakistan through India to Sri Lanka, and from Indonesia through China to Japan, one sees countries whose national identities are being defined by a prolonged conflict over secularization (Juergensmeyer 1993). In each case, the struggle involves less one religious group versus another than religion generally versus secular alternatives.

In addition to what might be termed a “bottom-up” process of seeping secularization, there are instances of a “top-down” coercive scenario, and the two are not mutually exclusive. The former Soviet Union, Turkey, and China illustrate the latter process through political systems headed by Lenin, Ataturk, and Mao Tse-Tung and their followers, respectively. This structurally imposed secularization had cultural effects as specifically defined state rituals became common alternatives to traditional religious ceremonies. However, in all these countries, traditional religion remains in evidence in the private sphere and occasionally bursts into the public arena.

Although there are examples of externally coerced secularization (e.g., the U.S. insistence on Japan’s abolishing “State Shinto” after World War II), secularization generally takes a far less direct form. Consider India as a case in point. Over the centuries, the south Asian subcontinent has given the world Hinduism, Buddhism, Jainism, and Sikhism, but from the early sixteenth century until the mid-twentieth century, it was dominated by outside rulers representing first Islam in the Moghul period and then Christianity under the British “raj.” When independence was won in 1947, the partitioning of Pakistan and India created two states, one Muslim and the other dominantly Hindu. The religious resorting involved a massive cross-migration as long-time residents of each area moved to the other so that they could live among their coreligionists. The violence that ensued is estimated to have left from 250,000 to 500,000 people dead.

Religious conflict has continued in both areas, but in each case, it is not simply one religion against another but also religion versus secularity.

After independence, India instituted a national government that followed the Western model of a secular and thus religiously neutral state. However, after a half century, a series of violent conflicts between Hindus, Muslims, and Sikhs have left a cloud over India's state secularity. In the 1990s, the dominant and secularist Congress Party lost its voting plurality to the Hindu nationalist Bharatiya Janata Party (BJP). Momentum is gathering on behalf of a Hindu state that would reflect the country's Hindu majority. Nor is the movement confined to right-wing religious zealots. A number of India's most prominent intellectuals have entered the fray and produced a series of strident exchanges (Nandy 1990; Madan 1998; Beteille 1994). For many people, the commitment to state secularity has ebbed; paradoxically, secularism has been secularized.

### FINDING A MUDDLED GROUND

Today it is common to hear that secularization has been categorically "disproved" and that anyone who still uses the term is more of an ideological antediluvian than an *au courant* scholar. Yet one must be wary of throwing out the baby with a bathwater both drawn and drained by the critics themselves. And certainly one must always be suspicious of prophets who predict the vindication of their own ideology. Most of the early visionaries of secularization and a disproportionate number of the theorists who have followed have been personally nonreligious, if not necessarily antireligious. At the same time, the ranks of the antiseccularizationists have included a number of theorists with personal religious loyalties. Although a scholarly discipline should provide methods to avoid or transcend these biases, history indicates otherwise.

A full review of the empirical literature on the secularization debate is beyond the scope of this article, and it is not feasible to conduct an investigation that would constitute a critical test. However, this is not an issue that can be settled empirically. Statistical arguments will be irrelevant until a series of pressing ideological and conceptual issues are confronted.

One must decide what to test before deciding how to test it. Because the two great myths attributed to the secularizationists by their critics are

by their nature overblown, they are not hard to puncture. Debating the matter at such mythical levels lends an all-or-nothing quality to the dispute: Insofar as the thesis fails to document a shift from all to nothing, it is suspect. However, no recent secularization theorists stake their claim in those terms.

It is not difficult to refute the first myth of secular dynamics concerning a seamless and universal religiosity in tribal settings and in the historical past. However, for the past to be more religious, it is not necessary for it to be either consistently or totally so. For a society to have been dominated by religion as political power, it need not have been more religious at the level of the individual and *vice versa*. Even at that level, of the individual, the past may be more religious in terms of personal piety and belief without necessarily being more religious in terms of formal institutional participation. Also, to say that one group or society's past was more religious than its present is not necessarily to say that another's must be the same. Finally, there are multiple pasts, none of which need be linear in their linkages.

Meanwhile, the second myth of secular dynamics is even easier for critics of secularization to deflate. The notion of religion's actual death and disappearance has shifted from the sublime to the ridiculous, especially in the formulations of eighteenth- and nineteenth-century figures, some of whom foresaw the end in their own lifetimes (Stark 1998). Somehow religion survived the eighteenth and nineteenth centuries, not to mention the twentieth. Again, however, it is not clear that this is either necessary or sufficient to disprove a more nuanced conception of secularization. Today it is common to reject the concept of secularization simply because religion persists, but mere persistence masks a host of questions concerning religion's changing terms and circumstances.

The "secularization thesis" with a mythical beginning and a mythical end is erroneous, but it is a largely noninstructive error akin to "denying all climatology and the particular hypothesis of global warming because we have not yet been burned to a crisp and the nights do, after all, still get cooler" (Demerath 1998b, p. 9).

Clearly secularization as a textured social process remains a fruitful concept. In fact, once the focus shifts to a less extreme version, the consen-

sus widens considerably. Consider two recent remarks from arch critic Stark:

*This refers to a decline in the social power of once-dominant religious institutions whereby other social institutions, especially political and educational institutions, have escaped from prior religious domination. If this were all that secularization means, there would be nothing to argue about. Everyone must agree that, in contemporary Europe, for example, Catholic bishops have less political power than they once possessed, and the same is true of Lutheran and Anglican bishops. . . . Nor are primary aspects of public life any longer suffused with religious symbols, rhetoric, or ritual. (1998, pp. 4–5)*

*Of course, religion changes. Of course, there is more religious participation and even greater belief in the supernatural at some times and places than in others, just as religious organizations have more secular power in some times and places than in others. Of course, doctrines change—Aquinas was not Augustine, and both would find heresy in the work of Avery Dulles. But change does not equate with decline. (1998, p. 29)*

These statements greatly narrow the gap between secularization’s advocates and one key antagonist. For many of the former, Stark’s first passage suggests a battlefield conversion, though it is not a new position for him (Stark and Bainbridge 1985). While the second remark is correct in that change and declension are not identical, the implied invitation to deconstruct the two should be welcomed.

There is little question that secularization has come to connote decline. Whether in its long-range mythical or short-term process form, secularization posits some variant of religious erosion, if not extinction. However, all these versions represent a myopic and one-sided perspective compared to the alternative that follows.

#### PARADOXES OF SECULARIZATION AND SACRALIZATION

At a time when work on secularization might be expected to yield a consensually validated paradigm (Tschannen 1991), it is far closer to producing a new set of divisive paradoxes. Much of this

conflict results from the terms at issue. Both “secular” and “sacred” are mutually referential in that each makes a statement about the other. To be secular is to be nonsacred; to be sacred is to transcend and transform the secular. The same is true when one shifts from semantics to social processes. Just as an object must have been sacred for it be subsequently secularized, it must have been secular for it to be subsequently “sacralized.” Just as secularization marks a decline of the sacred, sacralization denotes an increase in the sacred in one form or another and at one level or another.

However, linking the processes of secularization and sacralization can have paradoxical results. The following eight propositions can serve as examples:

1. Religious revivals and “great awakenings” require previous eras of religious decline and secular “naps.” American religious history has been charted in terms of its eighteenth-, nineteenth-, and possibly twentieth-century awakenings (McLoughlin 1978), but an opposite focus has equal merit (May 1949; Erikson 1966; Turner 1985). It is the combination of the two that establishes the basic rhythm of a country’s religious history.
2. Modernization may lead to both secularization and sacralization. The grand narrative of the secularization thesis is that religion beats a steady and linear retreat in the face of mounting modernization. There is considerable truth to this but also some half-truth. This is what Berger referred to in recanting some of his earlier writing on secularization (Berger 1997). Modernization often leads to forms of secularization, but those often spark a sacralizing response—one that uses the means of modernity to protest the ends of modernity. This characterizes “fundamentalisms” everywhere, whether in the original Christian version in the United States or in the Islamic and Hindu variants around the global girdle of religious extremism. As was noted earlier, many countries demonstrate religion’s continuing presence, but these countries also bear witness to the incursions of secularity as a perceived threat to religious interests. If either religion or secularity



were fully dominant in these settings, the conflicts would be obviated.

3. The rise of a vital “religious marketplace” is evidence of both secularization and sacralization. An increase in religious competition often reflects the decline of religion’s structural monopolies and/or cultural hegemonies. Religious dominations once taken for granted are now subject to doubt and dismissal, yet the new consumer’s mentality may involve more stained-glass window shopping than long-term buying (actually joining a church). The debate over changing patterns of religiosity turns on this point, as does a current dispute over the significance of religious “switching” in the United States (Demerath and Yang 1998a).
4. Because movements that go against the societal grain often create more friction than do trends that go with it, one must be careful not to mistake the sacred exceptions for the secular rule. It is tempting to interpret the flames of a small religious movement as being more important than the smoking embers of its larger and more secularized context. In the same spirit, one must be wary of confusing growth rates with size. Both have their place, but even small, conservative religious movements with high growth rates may be marginal to the larger population and culture. As an example, see the “winners” and “losers” cited by Finke and Stark (1992).
5. Sacred manifestations may reflect secular forces, and vice versa. The relationship between any form of behavior and the motivations behind it is problematic. Standard indicators of religiosity such as civil religious loyalty, church membership, church attendance, and religious belief are all subject to myriad interpretations, not all of which are unambiguously sacred (Demerath 1998a; Haddaway et al. 1993). It may be more the case that the civil is religious than that the religious is civil: Church membership and attendance reflect a variety of sacred and secular meanings that vary across a population and across time, and affirming a religious belief may be less a matter of cognitive conviction than of cultural affiliation and continuity. Even the various “fundamentalist” movements may not be as uniformly or fanatically “religious” as they are often portrayed. Many of their members have a predominantly secular agenda that religion legitimizes (Demerath 2000). Similarly, a withdrawal from conventional religious frameworks may coexist with a more privatized faith (see the “little voice” of the pseudonymous Sheila Larson in Bellah et al. 1985, p. 221). Finally, there are any number of conventionally secular commitments that take on sacred valences for their devotees (see below).
6. Moderate secularization can be a prophylactic against ultimate secularization. Changing social conditions require changing forms of the sacred. Hence, some degree of secularization may serve as a form of sacred adaptation. This has been a tactical assumption in the trajectory of liberal Protestantism over the last century as pastors and theologians have made concessions to their secularizing adherents (Berger 1967; Demerath 1992). This tactic has been challenged by advocates of strict doctrine and strict churches (Kelley 1972; Iannaccone 1994), but cleaving to strictness may have cost the churches far more defections than has the alternative.
7. Secularization and sacralization are engaged in a dialectical oscillation in which each is contingent on and responsive to the other. The presence of one does not necessarily involve the absence of the other. As was noted above, a secularization that goes too far is likely to elicit a sacralizing reaction. Similarly, sacralizing may exceed the bounds of pertinence, propriety, credibility, or convenience in a complex social context. Thus, lapsing and laicization of various sorts result in a secularizing adjustment. Without suggesting that secularization is always balanced by a corresponding sacralization to create a religious equilibrium, one can say that this mutual responsiveness is an important

reason why secularization, like a sense of the sacred itself, will always be with us.

8. Focusing on the fate of old forms of religion may deflect attention from new forms of the sacred. An obsession with secularization in the past may preclude an analysis of sacralization in the present and future. Just as conventional religion may not necessarily be sacred, new sources of the sacred are not necessarily religious. Today one hears a good deal of talk about a growing distinction between religion and spirituality and about profound sacred commitments in everything from socialism to sex. Just because they have attained cliché status does not mean that these concepts should be jettisoned as possibilities for deeper investigation.

These eight propositions lead to a series of issues beyond the scope of this article: Does every individual need a sense of a sacred commitment and a regimen that is self-consciously maintained and ritually reinforced? Does every collectivity and society require something similar that is shared among its members? If the answers to these questions are affirmative, what is the relation between the sacredness required and conventional religion on the one hand and more secular sources on the other? To what extent can the sacred reside in high and low culture, moral and ethical convictions, and movements on behalf of political causes, personal identities, and nationalist ambitions? Is it possible to investigate these matters without falling into tautology and teleology? Precisely because these questions are so old, it is time for freshly conceptualized and newly researched answers.

The alternation of secularization and sacralization is a crucial historical dynamic not just for religion but for culture as a whole. Secularization without sacralization is a nearly defining characteristic of putative *postmodernity*, with its loss of grand narratives and collective bearings. At the other extreme, sacralization without secularization is a similarly defining characteristic of stereotypic *premodernity*, where the sacred is static and unchallenged. However, it is in historical *modernity* that secularization and sacralization play off each other in both producing and responding to change. Whether causes or effects, these are critical processes in the world of time as opposed to timelessness.

## SUMMARY

The importance of any scholarly issue is revealed in the debates it engenders. By this standard, secularization qualifies as very important indeed. As a matter that seems to defy either empirical or ideological consensus, it has become a kind of Gordian knot for social scientific scholarship on religion.

Clearly, it is possible to construct versions of secularization that are either outrageous or reasonable. It matters greatly how the concept is deployed. For some, it is a prophecy of religious demise, whether a tragic jeremiad or a triumphant anticipation. For others, it is a set of historically and sociologically specified processes that move less linearly and with less certainty through time. For still others, secularization converges with sacralization to form a stream of constantly shifting conceptions and locations of the sacred. Whichever option is at issue, the stakes are high, and the sight of scholars impaled upon them is not uncommon.

## REFERENCES

- Bell, Daniel 1976 *The Cultural Contradictions of Capitalism*. New York: Basic Books.
- Bellah, Robert N., Richard Madsen, William M. Sullivan, Ann Swidler, and Steven M. Tipton 1985 *Habits of the Heart*. Berkeley: University of California Press.
- Berger, Peter L. 1967 *The Sacred Canopy*. Garden City, N.Y.: Doubleday.
- 1997 “Epistemological Modesty: An Interview with Peter Berger.” *Christian Century* 114:972–978.
- Beteille, Andre 1994 “Secularism and the Intellectuals.” *Economic and Political Weekly* 29(10):559–566.
- Bruce, Steve 1995 “The Truth about Religion in Britain,” *Journal of the Scientific Study of Religion* 34:417–430.
- Chaves, Mark 1993 “Denominations as Dual Structures: An Organizational Analysis.” *Sociology of Religion* 54:147–169.
- Comte, August 1852 (1891) *The Catechism of Positive Religion*, 3rd ed. London: Routledge.
- Davie, Grace 1994 *Religion in Britain since 1945: Believing without Behaving*. Oxford, UK: Blackwell.
- 1998a “Excepting Exceptionalism: American Religion in Comparative Relief.” *The Annals* 558:28–39.
- 1998b “Secularization Disproved or Displaced.” In Rudy Laermans, Bryan Wilson, and Jaak Billiet,

- eds., *Secularization and Social Integration*. Leuven: Belgium: Leuven University Press.
- Demerath, N. J. III 1992 "Snatching Defeat from Victory in the Decline of Liberal Protestantism: Culture versus Structure in Institutional Analysis." in P. D. Demerath, T. Schmitt, and R. H. Williams, eds., *Sacred Companies*. New York: Oxford University Press.
- 2000 *Crossing the Gods: Religion, Violence, Politics and the State across the World*. New York: Rutgers University Press.
- Dobbelaere, Karel 1981 "Secularization: A Multi-Dimensional Concept." *Current Sociology* 20:1-216.
- 1985 "Secularization Theories and Sociological Paradigms." *Sociological Analysis* 46(4):377-387.
- 1993 "Church Involvement and Secularization: Making Sense of the European Case." In Eileen Barker, James A. Beckford, and Karel Dobbelaere, eds., *Secularization, Rationalism, and Sectarianism*. Oxford, UK: Clarendon.
- Douglas, Mary 1982 "The Effects of Modernization on Religious Change." In Douglas and Tipton and Steven M. Tipton, eds., *Religion in America: Spirituality in a Secular Age*. Boston: Beacon.
- Durkheim, Emile 1912 (1995) *The Elementary Forms of the Religious Life*, trans. Karen Fields. New York: Free Press.
- 1961 *Moral Education*. New York: Free Press.
- Erikson, Kai 1966 *Wayward Puritans*. New York: Wiley.
- Fenn, Richard 1979 *Toward a Theory of Secularization*. Provo, Utah: Society for the Scientific Study of Religion.
- 1993 "Crowds, Time, and the Essence of Society." In Eileen Barker, James A. Beckford, and Karel Dobbelaere, eds., *Secularization, Rationalism and Sectarianism*, Oxford, UK: Clarendon.
- Finke Roger, and Rodney Stark 1992 *The Churching of America, 1776-1990: Winners and Losers in Our Religious Economy*. New Brunswick, N.J.: Rutgers University Press.
- Habermas, Jurgen 1988 *The Theory of Communicative Action*, 2 vols. Cambridge, UK: Polity Press.
- Haddaway, C. Kirk, Penny Long Marler, and Mark Chaves 1993 "What the Polls Don't Show: A Closer Look at U.S. Church Attendance." *American Sociological Review* 58:741-752.
- Iannaccone, Laurence R. 1994 "Why Strict Churches Are Strong." *American Journal of Sociology* 99:1180-1211.
- Juergensmeyer, Mark 1993 *The New Cold War: Religious Nationalism Confronts the Secular State*. Berkeley: University of California Press.
- Kelley, Dean 1972 *Why Conservative Churches Are Growing*. Macon, Ga.: Mercer University Press.
- Luhmann, Niklas 1982 *The Differentiation of Society*. New York: Columbia University Press.
- Madan, T. N. 1998 *Modern Myths, Locked Minds*. Delhi: Oxford University Press.
- Martin, David 1969 *The Religious and the Secular*. New York: Schocken.
- 1978 *A General Theory of Secularization*. New York: Harper & Row.
- Marx, Karl (1844) 1963 "Contribution to the Critique of Hegel's Philosophy of the Right." In Thomas B. Bottomore, ed., *Early Writings*. New York: McGraw-Hill.
- May, Henry F. 1949 *Protestant Churches in Industrial America*. New York: Harper.
- McLoughlin, W. G. 1978 *Revivals, Awakenings, and Religious Change*. Chicago: University of Chicago Press.
- Nandy, Ashish 1990 "The Politics of Secularism and the Recovery of Religious Tolerance." In Veena Das, ed., *Mirrors of Violence*. Delhi, India: Oxford University Press.
- Parsons, Talcott 1977 *The Evolution of Societies*. Englewood Cliffs, N.J.: Prentice-Hall.
- Redfield, Robert 1953 *The Primitive World and Its Transformations*. Ithaca, N.Y.: Cornell University Press.
- Roof, Wade Clark, and William McKinney 1987 *American Mainline Religion*. New Brunswick, N.J.: Rutgers University Press.
- Spencer, Herbert 1874 1915 *Essays Scientific, Political, Speculative*, 3 vols. New York: Appleton.
- Stark, Rodney 1992 "Sociology of Religion." In *Encyclopedia of Sociology*. New York: MacMillan.
- 1998 "Secularization R.I.P." Paper presented at Meetings of the Society for the Scientific Study of Religion. *Sociology of Religion*.
- , and William H. Bainbridge, Jr. 1985 *The Future of Religion*. Berkeley: University of California Press.
- , and Laurence R. Iannaccone 1994 "A Supply-Side Interpretation of the 'Secularization' of Europe." *Journal for the Scientific Study of Religion* 33:230-252.
- Toennies, Ferdinand (1887) 1957 *Community and Society*, trans. and ed. Charles Loomis. East Lansing: Michigan State University Press.
- Tschannen, Oliver 1991 "The Secularization Paradigm: A Systematization." *Journal of the Scientific Study of Religion* 30:395-415.
- Turner, James 1985 *Without God, Without Creed*. Baltimore: Johns Hopkins University Press.

Voltaire, Francois-Marie Arouet 1756 (1963) *Essai sur les Moers et l'Esprit des Nations*, 2 vols. Paris: Garnier.

Warner, Steven 1993 "Work in Progress toward a New Paradigm for the Sociological Study of Religion in the U.S." *American Journal of Sociology* 98:1044–1093.

Weber, Max (1905) 1993 *The Protestant Ethic and the Spirit of Capitalism*, trans T. Parsons, intro. Randall Collins. Los Angeles: Roxbury.

Wilson, Bryan 1966 *Religion in Secular Society*. London: Penguin.

— 1998 "The Secularization Thesis: Criticisms and Rebuttals." In Rudy Laermans, Bryan Wilson, and Jaak Billiet, eds., *Secularization and Social Integration*. Leuven, Belgium: Leuven University Press.

Wuthnow, Robert 1988 *The Restructuring of American Religion*. Princeton, N.J.: Princeton University Press.

N. J. DEMERATH III

## SEGREGATION AND DESEGREGATION

In the early years of the American colonies and the new republic of the United States, segregation was not only impractical but undesirable. To benefit from slavery, slave masters had to manage and control slaves; therefore, they had to work with them. Not all slaves were field hands or agricultural workers; some were domestic servants, and so the slave master and mistress had to share their private quarters with slaves. Thus, many white Americans, especially Southerners in the pre-Civil War South, accepted daily, intimate, personal, primary face-to-face contact with slaves as a necessity. They insisted, however, that all such contacts reflect proper social distance: slaves were always to be subservient, behavioral assimilation was allowed only to a point, and slaves were supposed to know the dominant-group culture, use it appropriately, and always recognize that they were not the equals of their masters. Although structural assimilation occurred at a primary level, it was not among equals.

With the Emancipation Proclamation of 1863 and the ratification of the Thirteenth Amendment in 1865, some Americans seriously considered the idea of separating blacks and whites. As some blacks emigrated to poor urban areas in the South and as their numbers increased, some whites recognized that blacks were becoming a threat to the hard-won victories of higher-priced white labor

(Bonacich 1972). They recognized that the former mechanisms of deference and social distance would no longer allow whites to maintain the subordination of black men and women, and so they insisted on a system of separation. It was not enough to separate residentially; it was necessary to establish a caste system that would deny blacks equal access to most jobs, social and governmental services, schools and colleges, public accommodations, and the right to vote.

In both the South and the North, segregation was practiced long before it became embodied in law. It was a Supreme Court decision, however, that in 1896 established segregation as the law of the land. It was through the medium of statutes, therefore, that domination was ultimately exercised. In other words, it was the polity, not the economy, that suppressed the competition of black urban laborers and that established the shift from paternalistic to competitive race relations (Scott 1977; van den Berghe 1967).

Segregationist laws were passed as early as 1875 in Tennessee; they rapidly advanced throughout the South, and by the 1880s blacks were not only separated on all modes of transportation (Franklin 1947). However, the Civil Rights Act of 1875, which guaranteed black Americans all the privileges and rights of citizenship, was an impediment to the policy of segregation. Consequently, the impediment was removed in 1883, when the Supreme Court declared the Civil Rights Act of 1875 unconstitutional. Soon after that decision, black Americans were banned from most Southern venues, from hotels and other places of public accommodation—restaurants, theaters, and places of public amusement. The process of limiting opportunities for blacks continued, and by 1885 most Southern states had enacted laws requiring separate schools for blacks and whites. Finally, on May 18, 1896, the Supreme Court in the infamous *Plessy v. Ferguson* decision made segregation the law of the land (Kromkowski 1991). Although the North and the South were elated, the implication of the decision and the way it was to be implemented would be considerably different in the two regions. As a result, the consequences and effects of segregation in the South would be different from those in the North.

If segregation had not legitimated the rights of Southern whites to degrade and control blacks,

blacks might have seen opportunities for independent growth in segregation. Segregation in the South meant biracialism, and biracialism meant the creation of black institutions that were to some extent administered and controlled by blacks. Although most blacks in the South worked for whites, they did not have to depend on them for all their basic services: They had separate schools, hospitals, and churches. Most blacks in the South became sharecroppers, working rented land. The land meant debt for the sharecropper, but it also meant a certain amount of daily independence. It is conceivable, therefore, that under a more positive set of circumstances blacks could have focused on the "equal requirement" of the *Plessy* "separate but equal" decision. However, because segregation became the detested symbol of injustice, Southern blacks insisted on destroying it.

As blacks struggled against segregation, they were beaten and murdered. Law enforcement participated in those affronts either by refusing to protect black people or by becoming the perpetrators of violence. Such actions reinforced the view of Southern blacks that segregation was the symbol of black inferiority. As blacks struggled to defend themselves, they learned that sheriffs and law enforcement officials, mayors, governors, the FBI, the federal government, the attorney general of the United States, and even the president participated in one way or another in the maintenance of a system of segregation that declared black people inferior and denied them equal access to the labor market and to educational opportunity.

Although Southern blacks were eventually successful in destroying the system of segregation in the South, blacks in the North, where the *Plessy* decision had been implemented differently, often failed. Because the major problem in the North was not segregation, the strategies of Southern blacks were inappropriate for the problems of Northern blacks and those who moved north. Desegregationist strategies were designed for problems such as residential segregation but not for problems such as poverty and differential access to occupational opportunities. This is why the Southern Christian Leadership Conference left the urban slums of Chicago in 1965, where the real problems were, and attacked the issue of segregated housing in Cicero, Illinois, which for blacks at that time was insignificant.

Although Southern whites insisted on black inferiority, one should not assume that they therefore wanted to dispose of blacks. They needed blacks for at least two reasons: to establish their alleged superiority and to exploit black labor. Blacks had been their slaves, had worked their fields, had stabilized and maintained their households, and had been a source of wealth and sometimes pleasure. Many Southern whites had even developed a degree of affection for blacks.

Northern whites were quite different in this regard. Some knew the value of black Americans, but their major goal was to make certain that blacks and whites remained apart. A biracial system was not required because occupational and economic discrimination kept blacks and whites apart. When and where necessary, whites would use restrictive real estate practices to keep the races separate. Whites in the North wanted blacks to stay completely to themselves unless there was some need for their labor. With the exception of hiring black women, whites did not really want to make competitive use of black labor. It seems that Northern whites wanted blacks to disappear, and so they pretended that they did not exist.

In the South, segregationist policies eventually led to a biracial system that produced unanticipated consequences. It actually laid the groundwork for the development of a black middle class composed of clergy, college administrators and professors, medical doctors, journalists, schoolteachers, artisans, and skilled craftspeople, all of whom had learned to be independent in their respective institutional settings. They were the decision makers and leaders of their people. They would train the new teachers, the new professionals, and even a new business elite. Their protégés would become the new entrepreneurs and open businesses of various kinds—barbershops, beauty shops, grocery stores, restaurants, and nightclubs. They would establish black banks, publish black newspapers, and establish professional societies. Many of the college graduates would become ministers and establish their own churches. In time, all these professionals would combine their resources and expertise and, using their two institutional bases, the school and the church, lead a struggle against the very system that made their existence possible: the system of segregation. In the South segregation did not mean separation only. It meant the right of whites to degrade blacks and treat

blacks unjustly, but mostly it meant the right to keep blacks in an inferior position by denying them equal access and equal opportunity.

Eventually the black church, a product of segregation and discrimination, would become the institutional base for the fight against segregation and discrimination. Not only did the black church provide the leadership, it also provided the following. However, since black churches had existed for decades and their congregations had been ready for change for decades, why did the "movement" take until 1955 to start? A critical component is the size of the black middle and skilled-working classes. In the middle to late 1950s, those two classes constituted approximately 39 percent of the black community, a larger percentage than ever before. World War II had been a major period of opportunity for African Americans, and as a result, they garnered more resources and consequently expected more from the system. In short, they experienced a revolution of rising expectations. They had become intolerant of abuse, the various forms of discrimination they had experienced, and insults to their dignity. They were in need of a social movement.

#### DESEGREGATION: THE CIVIL-RIGHTS MOVEMENT

The impetus for the civil-rights movement, the movement to desegregate the South, actually began before Mrs. Rosa Parks's heroic refusal in 1955 to give up her bus seat to a white person. The initial stimulus was the May 17, 1954, decision of the Supreme Court in *Brown v. Board of Education* (1954) that the 1896 *Plessy* decision was unconstitutional. Black soldiers returning from World War II and the burgeoning black middle class praised the decision and proclaimed that the *Brown* decision must usher in a new social order.

No sooner had the decision been made, however, than the nation was shocked by the grisly murder of a young teenager, Emmett Till, in Sumner, Mississippi. That murder dramatized the fact that no change in the law would change the customs of Southern whites, and the case demonstrated how the circumstances of blacks in the South were radically different from those of blacks in the North. According to Emmett Till's uncle, Emmett had been bragging to some black youngsters outside a rural store. He claimed to have

white friends, even white girlfriends, in Chicago and showed photographs of his friends. Emmett had just arrived in Sumner and was trying to impress those young boys to gain their friendship. One of the boys apparently said to Emmett, "I bet you won't go into that store and say something to that white lady." Till accepted the challenge, went in, purchased some candy, and in leaving said, "Bye, baby." Late the same night, two or more white men knocked at the door of Emmett's grandfather, Mose Wright, and took the boy away in a car. When Emmett Till was found, he had been mutilated and beaten beyond recognition, with a bullet hole through his temple. The picture of Emmett Till's disfigured body was published in *Jet* magazine by Johnson Publications, a black publishing firm, and black people throughout the nation saw the picture. Till's mother insisted on an open casket. Two men were charged with the murder, but both were found not guilty. Black people recognized that a change in the law was not enough. More had to be done.

Emmett Till was a Northern urban kid who had grown up and apparently gone to school with some liberal whites, and although the commingling of whites and blacks in the North could lead to violence, in some circles it was tolerated. Because the issue in the North was residential separation, it was easy for a black person to find himself in a predominantly black school, though generally there were at least a few white students. More important, however, was the fact that the overwhelming majority of the teachers were white (Jones 1985, p. 180). Those teachers and other professionals usually lived outside the school districts in which they taught. Although they insisted that black schoolchildren obey them, they did not insist that blacks be subservient and inferior. As teachers, they were proud of their successful black students. Northern blacks thus developed self-esteem, a sense of "somebodyness," a belief that they were the equals of others. That attitude was reinforced in black urban enclaves. In the South, however, every contact a black person had with a white person required a demonstration of black inferiority and even fear. The idea of being equal to whites was generally unthinkable, that is, if the idea was to be put into action. Northern blacks were always warned by their relatives when they went to the South that the rules were different there, that not obeying them could place every-

body in jeopardy and could even lead to the loss of life.

Emmett Till was a tough urban kid, not unlike many of the gang members of the 1990s, and the fact that he was not afraid of his captors and refused to stop fighting back made them angrier. He obviously did not know that what he did in the North could get him killed in the South. He had not been warned, or he did not heed the warning.

Emmett Till's murder and the injustice of the final verdict produced mounting frustration. Thus, on December 1, 1955, Mrs. Rosa Parks told a bus driver who asked her to give her seat to a white person, which was the law, that she would not. This galvanized the entire black population of Montgomery, Alabama. The black community organized a bus boycott, and soon the buses were empty. The leadership was surprised (Raines 1977). Black people were fed up. They had always been angered by such demands and customs, but as Christians they had been taught to accept them and hope for change. Now, however, former soldiers and their families who had been patriotic and had sacrificed during World War II had become intolerant. Segregation did not mean biracialism to them. Instead it meant abuse and insult. A social movement had started.

Soon a brilliant young black Baptist minister would join the movement, and even though he was only twenty-six years of age, he would become the leader. That leader, Martin Luther King, Jr., defined the enemy as segregation. Segregation, King insisted, "scars the soul of the segregated. . . . It not only harms one physically, but injures one spiritually." It is a system, asserted King, that "forever stares the segregated in the face saying you are less than, you are not equal to." Segregation denies a human being the right to express his or her true essence; therefore, it must be destroyed. King declared that nonviolence would be the movement's strategy and philosophy. Nevertheless, violence erupted immediately. Whites were resisting, but the Montgomery Improvement Association won its victory when the Supreme Court declared segregated busing unconstitutional. King and his leadership cadre immediately set about the task of desegregating other public facilities in Montgomery. The movement had begun, and from that point on other struggles would erupt sponta-

neously across the South, all of them devoted to desegregation.

As African-American college students observed the activities of Dr. King and his organization, the Southern Christian Leadership Conference (SCLC), they agreed to continue the process of desegregation. Dr. King was desegregating downtown department stores in Montgomery; they would desegregate lunch counters. It was the custom in the South not to serve blacks at lunch counters in the various dime stores, especially the Woolworth's chain. On November 1, 1960, four students from the local black college took seats at the lunch counter in Greensboro, North Carolina. They asked to be served, and when the management refused, they resolved to stay. After a day or two, violence broke out. A group of young white toughs and some older adults began to pull them out of their seats and beat them. The police were called in, but they refused to arrest the perpetrators of the violence. Instead they arrested the victims, those who were involved peacefully in what became known as sit-ins. As a result of the police actions, Southern blacks noted again that not only were the citizens of the South opposed to their rights, so were public officials. Segregation had to be destroyed "lock, stock, and barrel, top to bottom, left to right" (Carmichael 1971) because it also corrupted public officials and officers of the law whose sworn duty it was to protect the citizenry. From this point on segregation was the enemy, and going to jail to end it became a badge of honor.

The issue of segregation on buses involving interstate travel remained a problem even after the Montgomery victory. Therefore, it was not long before groups of Freedom Riders were mobilized to test the Supreme Court decision's relevance to interstate travel. The Freedom Riders included blacks and whites, a fact that should not be forgotten. The Freedom Rides began in May 1961 and were immediately confronted with violence. Buses were bombed. Freedom Riders were beaten unmercifully at several destinations, and some were permanently disabled. The perpetrators were indiscriminate: they beat blacks and whites. Their hatred seemed greater for whites—"nigger lovers," they were called then. The Freedom Riders expected to be protected by the FBI, but J. Edgar Hoover, the director, made it clear that his agency had no intention of protecting those agitators. The failure of the federal govern-

ment to uphold the law in this instance finally communicated to black people and some whites that the existence of segregation had corrupted not just local public officials but even officials of the federal government. The fight had to begin at the top.

The next major chapter in the effort to desegregate the South took place in Albany, Georgia, in 1961. Failing in their desegregation efforts there, King and the SCLC launched a new project to protest segregated lunch counters in downtown Birmingham, Alabama. King was jailed. While in jail, he wrote his philosophically brilliant "Letter from a Birmingham Jail." Although Birmingham's white business leaders agreed on a desegregation plan, King's motel was still bombed. Medgar Evers was shot to death in neighboring Jackson, Mississippi, and four young children were murdered in the bombing of the Sixteenth Street Baptist Church in Birmingham. Blacks learned that even if they could get local public officials and businessmen to change segregationist policies, some Southern whites, perhaps even the majority, would not accept change. They also learned that among the majority there were those who were willing to use violence. Blacks had to have protection from another source.

In 1964, the Student Nonviolent Coordinating Committee (SNCC) began its Freedom Summers Project in Mississippi. Mississippi was considered by blacks the most dangerous state in the South, and it lived up to its reputation. On Sunday, August 4, 1964, Mississippi claimed the lives of James Chaney, Michael Swerner, and Andrew Goodman—the latter two were white. All three were members of SNCC's Freedom Summer Project. Their only offense was that they had volunteered to teach black youth, work with the rural poor, and register blacks to vote. If it was not apparent during the Freedom Rides, it was now apparent that Southern whites would kill anybody, whites included, who opposed their way of life.

Blacks now had a growing collection of concerned Northern whites. Swerner's wife commented that it was unfortunate, but apparently whites had to die before other, complacent whites would listen. The parents of the two young white students, Swerner and Goodman, talked about the martyrdom of their children. They were proud but grief-stricken. They insisted that the monstrous

evil of segregation must be destroyed. Black members of the Congress of Racial Equality (CORE) and SNCC were furious. Some of them had been personal friends of James Chaney, who was black. They blamed the governor of the state and the federal government for what happened in Philadelphia, Mississippi, during the summer of 1964.

As a result of those murders, SNCC and SCLC mobilized a march to Montgomery. Near the end of their march, however, they were attacked by mounted sheriff's officers wielding clubs. Men and women, as well as young adults and children, were beaten.

In summary, the central focus of black struggle in the South from 1955 to 1965 was desegregation. Blacks insisted on desegregating public transportation facilities, public eating establishments, public water fountains, public bathrooms, and public institutions of higher education. As a result of the violence they experienced, black Southerners learned that desegregation required more than protests, it required changes in the law at the national level. A civil-rights bill was required. Certainly a change in the law was required, but even that was not enough. In order for changes to be implemented, government officials had to demonstrate a willingness to protect and defend the rights of African Americans.

It was not long after Selma that Watts, an urban ethnic enclave near Los Angeles, exploded, beginning a series of race riots that developed spontaneously throughout the latter half of the 1960s. Stores were torched and looted. Surveying the destruction in Watts, Dr. King and SCLC decided that it was time to take their movement north. What they were not aware of was that their desegregation strategies would not solve the problems of Northern blacks, because the central problem for that group was *not* segregation. To understand this, it is necessary to contrast the evolution of the black middle class in the South with that in the North.

#### DESEGREGATION VERSUS INTEGRATION

A biracial system similar to that in the South never surfaced in the American North. As a result, blacks there depended almost completely on whites for employment. Northern whites, furthermore, had not come to depend on black labor, with the



possible exception of domestic labor. Domestic labor, however, did not produce wealth; it was a symbol of surplus wealth. In addition, Northern whites who wanted to remain physically (residentially) separated from blacks did not feel any need to employ them, with the exception of menial labor jobs. With the influx of European immigrants, Northern whites preferred to hire the sons and daughters of Europe rather than the emancipated slaves of the South (Blauner 1972; Jones 1985). Indeed, from the turn of the century to the beginning of World War II, Northern blacks never established a foothold in the manufacturing industries of the North (Jones 1985). According to Blauner, even in ancillary industries such as meatpacking where blacks initially gained a foothold because of the unhealthy working conditions they were actually displaced by European immigrants during the 1930s.

Given their background, the problem for the black middle class in the North was different from that for the black middle class in the South, and the leadership of the civil-rights movement knew it. At one point, Dr. King said that “the struggles of the past decade were not national in scope; they were Southern; they were specifically designed to change life in the South” (1968, p. 70). Northern blacks had only been segregated (de facto) residentially. Otherwise they could ride public transportation and eat at many of the major restaurants, although it was understood that some owners would discourage blacks from coming by being discourteous. The major concern of Northern middle-class blacks, therefore, was not formal desegregation but discrimination and unequal access. They insisted that they should get the same quality of goods or service for their money. Their major concern was reflected in their insistence on greater job opportunities. They rejected the idea of caste barriers in employment, and they insisted that promotions be tied fairly to evaluation, irrespective of race. They rejected job ceilings and the idea of determining job status on the basis of race. These kinds of problems could not be solved by civil rights marches. They could not be solved simply by desegregation or changing the law. Such changes would help, perhaps, but what was required was to get the federal government to establish civil-rights policies that would declare such acts as violations of the law and then, even more important, connect those policies to some kind of

enforcement device so that private corporations and governmental agencies would comply with the law. This is exactly what the Civil Rights Act of 1964 in combination with affirmative action, did.

### **THE FAILURE OF INTEGRATION: THE URBAN POOR**

Soon after the Civil Rights Act was passed by Congress and signed by President Lyndon B. Johnson, Executive Orders 11246 and 11375 were issued. Those orders led to the policy of affirmative action (Black 1981). Affirmative action policies essentially required that all city, state, and federal agencies, as well as any private corporation that contracted with the federal government, make every reasonable attempt to increase the proportion of minority workers in their workforces. Affirmative action was to be a device to address the effects of past discrimination. It did not take long to realize, however, that mostly middle-class blacks were benefiting from affirmative-action policies (Wilson 1987). The reason for this was twofold. First, middle-class blacks were the only ones who had competitive resources (such as skills they had acquired from higher education), owned businesses, or had parents who as a result of their professional status (doctors, dentists, ministers, etc.) were able to provide a competitive advantage for their children. Second, the American economy underwent structural changes that created more opportunities for professional, technical, human service, and clerical staff. As these opportunities increased, affirmative-action policies increased the likelihood that some of those jobs would go to black Americans. It was not long, however, before it also became apparent that neither affirmative-action policies nor the structural shift in the economy would aid black Americans who were poor and unskilled. In fact, as the economy shifted from a majority of manufacturing industries to a majority of service industries, a segmented labor market developed. A segmented labor market generated differential rates of mobility for differing class segments of the same group (Wilson 1978; 1981; 1987).

It is not surprising that when Mayor Richard Daley of Chicago and Martin Luther King, Jr., met early in 1966 and King complained about the slum housing of poor blacks in that city, Daley responded, “How do you expect me to solve the problems of

poverty and joblessness overnight?" King had no answer. He would quickly leave Chicago and the North after unsuccessful attempts both to help the impoverished and to desegregate Cicero, Illinois. It is to Dr. King's credit, however, that he recognized that the problems of the poor had not been solved and that a Poor People's Campaign was required.

Oblivious to the needs of the poor in the black community, Northern blacks who had turned a desegregationist movement into an integrationist movement (those Sowell [1984] incorrectly labels as people with a civil-rights vision) pursued integration with a vengeance. When the Civil Rights Act of 1964 became the law of the land, affirmative action was to be its guiding policy and equal opportunity and equal access the measure of fairness. It was not long, however, before civil-rights advocates recognized that something was amiss not only for the Northern black poor but also for the middle class. For example, although data from the 1970 census showed that black male college graduates in 1969 received a slightly higher average income than did comparable whites, other data demonstrated that the majority of black college students did not graduate from college. In fact, when Fleming (1985) researched this issue and compared the performance of black colleges with limited resources to that of predominantly white urban universities with considerably more resources that attracted black students with higher SAT scores, she found that the black colleges produced more intellectual and psychosocial development among black students than did the white colleges. Further, she found that typically white colleges produced "academic deterioration" among black students and concluded that better facilities and more institutional resources do not necessarily translate into a higher-quality college or university education (Fleming 1985, p. 186). She added that similar findings were reported in desegregated or so-called integrated public schools (Knowles 1962).

The fact is that whether or not schools are integrated, the situation confronting black children in most Northern and Southern public schools is catastrophic. Indeed, for the most part integration has failed black children. Once they enter school, they fall quickly behind their white counterparts on most measures of intelligence and scholastic achievement (Coleman 1966; Denton

1981). In fact, the longer black children remain in school, the further they fall behind. Denton (1981) reports that compared to white children, black children are three times as likely to be labeled mentally retarded, twice as likely to be suspended for discipline and attendance problems, and twice as likely to drop out of high school (White 1984, pp. 102-103). Black students who remain in school on average are two to three years below grade level in the basics—reading, writing, and arithmetic.

Educational integration consequently has often led to less growth and, even worse, the actual deterioration of the academic potential of black students in institutions of higher education. In those situations where deterioration does not actually occur, stagnation does (Black 1981).

These kinds of problems compound in later life such that black students have only "half as much . . . chance as a white child of finishing college and becoming a professional person," twice as much chance of being unemployed, and a one in ten chance of getting in trouble with the law (and, if these students are young males, a one in four chance of involvement with the criminal justice system); finally, as they age, black students have a life expectancy that is five years shorter than that of white adults (White 1984, p. 103).

Without an adequate education, black males become less employable, less marriageable, and more criminal.

Wilson (1981, 1987) examined the combined indicators of unemployment rates, labor-force participation rates, employment-population ratios, and work experience and concluded that not only do these indicators reveal a disturbing picture of joblessness, they also indicate that a growing percentage of young black males are not marriageable, that is, cannot contribute to the support of a family. Examining rates of teenage pregnancy; crime and violence, especially homicide; and increases in substance abuse, Wilson argues that many of these young men are more likely to become predators than responsible workers.

Further, according to Wilson, poverty has compounded in black urban ethnic enclaves. He demonstrates that there has been a significant increase in what he refers to as extreme poverty areas (i.e., areas with a poverty rate of at least 40 percent) in the black urban ethnic enclave. Wilson contrasts

the growth of these areas with low-poverty areas (census tracts with a poverty rate of at least 20 percent) and high-poverty areas (with a poverty rate of at least 30 percent). The number of extreme poverty areas, he emphasizes, increased by a staggering 161 percent.

Wilson also demonstrates that the black community is losing its vertical class integration. Black middle-class and stable working-class families are choosing to live in the suburbs, and as they do, the institutions they used to staff, support, and nourish decline in number and importance.

The eventual demise of ethnic enclaves in urban areas has been experienced by all ethnic groups in America; for Europeans the process has taken from four to six generations. For blacks the process began in the late 1960s and early 1970s. Of course there was resistance to residential integration, as mirrored in the hostility that exploded in Cicero, Illinois, in the 1960s, and it continued in the 1990s. Although blacks residing in white suburbs are often racially harassed, residential integration is gaining momentum even as some whites move out of suburbia to exurbia or back to the city.

It should be noted, however, that the process of ethnic enclave decline for blacks is fundamentally different from that for European ethnics. Europeans settled in urban areas at a time when urban job opportunities were increasingly plentiful. Most of the jobs were in manufacturing and did not require skilled labor. And since Europeans were preferred over blacks, the sheer numbers of jobs allowed them to lift a whole mass of people out of squalor. As economic stability increased, European ethnics began the process of preparing themselves for increased mobility within the American occupational structure. To do this, education was critical—educational institutions are essentially preparatory institutions. In sum, the occupational and economic success of European ethnics required a stable economic base first, education second, and occupational success third (Greeley 1976). The circumstances of black Americans (a sizable segment of whom were denied stable employment opportunities in the North) were totally different, particularly in urban areas and especially in the North.

European ethnics were preferred over black laborers. Consequently, while European ethnics

were reaping the benefits of full employment, blacks were denied equal access to the labor market, undermining their ability to establish a stable economic base. And for those who would come later, after manufacturing jobs actually began to diminish because of the restructuring of the economy, there would be nothing but long-term unemployment. These groups would eventually form the black underclass as one generation of unemployed workers would quickly give rise to another. European ethnics were described by the sociologists of the 1920s and 1930s as socially disorganized (Thomas and Znaniecki 1927). Their communities were plagued by crime, delinquency, gangs, prostitution, and filth, but the availability of employment opportunities in the 1940s and 1950s allowed many to “lift themselves up by their own bootstraps.”

The jobs that are available to blacks now because of the growth in the service sector of the American economy are either jobs that do not pay enough for a person to support a family or require considerable education and training, and so black urban ethnic enclaves are likely to undergo a different kind of transformation than did the European ethnic enclaves of the early 1900s. The middle class will be increasingly siphoned off from such enclaves, leaving behind a large residue of the most despondent and dependent, the most impoverished, the most violent, and the most criminal elements. Without new institutions to play the role of surrogate parents, without some kind of mandatory civilian social service corps, blacks in those communities may become a permanent underclass. A residue was also left behind by European ethnics, but it was much smaller and therefore much less problematic. As the black middle class leaves, it leaves its ethnic community devoid of the leadership or resources needed to regain its health. And as the numbers of female-headed families increase, the middle class will eventually have left the majority of black people behind. Integration, then, has undermined the health and the integrity of the black community.

The counterposition is now being proffered by many people, organizations, and school systems throughout the United States. This can be seen in the proliferation of segregated black programs where black youngsters are being taught only by black teachers. In this context, race clearly

is the critical issue. Gender however, has also become an issue. In many of these schools, black males insist that only they can do the job. Since black women have had to bear the burden of rearing children alone for so long, there is no doubt that they can use some help. One critical problem remains for people of this persuasion, however, and that is the continuing trend of black middle-class and stable working-class flight. Can the black community stem the tide? It is not suggested here that the black middle class can solve the problem alone but rather that it must provide the leadership, as it did in the segregated black institutions of the South, and that government must pay for it. Can the exodus be diminished?

Possibly, the passage of anti-affirmative action legislation and the increased reliance on standardized testing in higher education may alert the black middle class that the opportunities for their children are diminishing. Already they are starting to send their children to historically black colleges and universities in record numbers. This is a major shift in black higher education, and the number of available admissions is limited. As their children's opportunities decrease, maybe they will come to see that in America the opportunities of black Americans will always be dependent upon the amount of pressure that blacks as a people can bring to bear on the system.

In the last few months of 1998, several anti-affirmative action programs were passed—Proposition 209 in California and Initiative 200 in Washington, for example. These initiatives were passed despite the demonstrated benefits of affirmative action for the broader society.

(SEE ALSO: *Apartheid; Discrimination; Equality of Opportunity; Ethnicity; Prejudice; Race; Segregation and Desegregation; Slavery and Involuntary Servitude*)

## REFERENCES

- Black, Albert N., Jr. 1977 "Racism Has Not Declined, It Has Just Changed Form." *Umoja* 1, no. 3 (fall).
- 1981 "Affirmative Action and the Black Academic Situation." *Western Journal of Black Studies* 5:87–94.
- Bowen, William, G., and Derek Bok 1998 *The Shape of the River: Long-Term Consequences of Considering Race in*

- College and University Admissions*. Princeton N.J.: Princeton University Press.
- Blauner, Robert 1972 *Racial Oppression in America*. New York: Harper and Row.
- Bonacich, Edna 1972 "A Theory of Ethnic Antagonism: The Split Labor Market." *American Sociological Review* 37 (October): 547–559.
- Coleman, James S. 1966 *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.
- Denton, Herbert 1981 "Future Still Bleak for Black Children, Lobbying Group's Statistics Show." *Los Angeles Times*, January 14(1A):4.
- Fleming, Jacqueline 1985 *Blacks in College*. San Francisco: Jossey-Bass.
- Franklin, John Hope 1947 *From Slavery to Freedom: A History of Negro Americans*, 4th ed. New York: Knopf.
- Greeley, Andrew M. 1976 "The Ethnic Miracle." *The Public Interest* 45: 20–36.
- Jones, Jacqueline 1985 *Labor of Love, Labor of Sorrow: Black Women, Work and the Family, From Slavery to the Present*. New York: Vintage.
- 1998 *American Work: Four Centuries of Black and White Labor*. New York: Norton.
- King, Martin Luther, Jr. 1968 *Where Do We Go From Here: Chaos or Community?* Boston: Beacon Press.
- Knowles, L. W. 1962 "Part 1, Kentucky." In United States Commission on Civil Rights, *Civil Rights USA: Public Schools, Southern States*. Washington, D.C.: U.S. Government Printing Office.
- Krombowski, John A., ed. 1991 *Annual Edition: Race and Ethnic Relations, 91 to 92*. Guilford, Conn.: Dushkin.
- Massey, Donald, and Nancy Denton 1993 *American Apartheid: Segregation and the Making of the Under-class*. Cambridge, Mass.: Harvard University Press.
- Raines, Howell 1977 *My Soul Is Rested*. New York: Penguin.
- Scott, Joseph W. 1977 "Afro-Americans as a Political Class: Towards Conceptual Clarity." *Sociological Focus* 10 (4):383–395.
- Sowell, Thomas 1984 *Civil Rights: Rhetoric or Reality?* New York: Morrow.
- Thomas, William I., and Florian Znaniecki 1927 *The Polish Peasant in Europe and America*. New York: Alfred A. Knopf.
- Van den Berghe, Pierre L. 1967 *Race and Racism: A Comparative Perspective*. New York: Wiley.
- White, Joseph 1984 *The Psychology of Blacks: An Afro-American Perspective*. Englewood Cliffs, N.J.: Prentice-Hall.

- Wilson, William J. 1978 "The Declining Significance of Race: Revisited but Not Revised." *Society*, vol. 15.
- 1981 "The Black Community in the 1980s: Questions of Race, Class, and Public Policy." *Annals of the American Academy of Political and Social Science*, vol. 454.
- 1987 *The Truly Disadvantaged: The Inner City, Underclass, and Public Policy*. Chicago: University of Chicago Press.
- 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Vintage.

ALBERT WESLEY BLACK, JR.

## SEGREGATION INDICES

Residential segregation has been a prominent topic in sociology since Burgess (1928) published his landmark study more than seventy years ago, and for almost as long, sociologists have argued about how to measure it. The debate has ebbed and flowed, and for a time the issue seemed to have been settled. In 1955, Duncan and Duncan published a landmark article (Duncan and Duncan 1955) demonstrating that there was little information in any of the prevailing indices that had not been captured already by the index of dissimilarity. For twenty years afterward, that measure was employed as the standard index of residential segregation.

This *Pax Duncanae* came to an abrupt end in 1976 with the publication of a critique of the dissimilarity index by Cortese and colleagues, ushering in a period of debate that has not ended (Cortese et al. 1976). Over the ensuing decade, a variety of old indices were reintroduced and new ones were invented, yielding a multiplicity of candidates. In an effort to bring some order to the field, Massey and Denton (1988) undertook a systematic analysis of twenty segregation indices they had identified from a review of the literature. They argued that segregation is not a unidimensional construct but encompasses five distinct dimensions of spatial variation. No single dimension is intrinsically more "correct" than any other; each reflects a different facet of the spatial distribution of social groups.

The five dimensions they identified are evenness, exposure, clustering, concentration, and cen-

tralization. To verify that conceptualization, Massey and Denton (1988) carried out a factor analysis of indices computed from 1980 census data for U.S. metropolitan areas. Their results showed that each index correlated with one of five factors corresponding to the dimensions they postulated. On theoretical, empirical, and practical grounds, they selected a single "best" indicator for each dimension of segregation. The dimensional structure of segregation and Massey and Denton's (1988) selection of indices have been reaffirmed using 1990 census data (Massey et al. 1996).

The first dimension of segregation is evenness, which refers to the unequal distribution of social groups across areal units of an urban area. A minority group is segregated if it is unevenly spread across neighborhoods. Evenness is not measured in an absolute sense but is scaled relative to another group. It is maximized when all areal units have the same relative number of minority and majority members as the city as a whole and is minimized when minority and majority members have no areas in common.

The index of dissimilarity quantifies the degree of departure from an even residential distribution. It computes the number of minority group members who would have to change neighborhoods to achieve an even distribution and expresses that quantity as a proportion of the number that would have to change areas under conditions of maximum unevenness. The index varies between zero and one, and for any two groups X and Y it is computed as:

$$D = .5 \sum_{i=1}^n \left| \frac{x_i}{X} - \frac{y_i}{Y} \right| \quad (1)$$

where  $x_i$  and  $y_i$  are the number of group X and group Y members in areal unit  $i$  and  $X$  and  $Y$  are the number of group X and group Y members in the city as a whole, which is subdivided into  $n$  areal units.

Among its properties, the index is inflated by random factors when the number of minority group members is small relative to the number of areal units (Cortese et al. 1976). It is also insensitive to the redistribution of minority group members among areal units with minority proportions above or below the city's minority proportion

(James and Taeuber 1985; White 1986). Only transfers of minority members from areas where they are overrepresented (above the city's minority proportion) to areas where they are underrepresented (below the minority proportion) affect the value of the index.

The property means that the dissimilarity index fails the "transfers principle," which requires that segregation be lowered whenever minority members move to areas where they constitute a smaller proportion of the population. This and other problems led James and Taeuber (1985) to recommend using another measure of evenness, the Atkinson index (Atkinson 1970). Massey and Denton (1988), however, pointed out that the Atkinson index and dissimilarity indices are highly correlated and generally yield the same substantive conclusions. Moreover, the Atkinson index is actually a family of indices, each of which gives a slightly different result, creating problems of comparability. Given that  $D$  has been the standard index for more than thirty years, that its use has led to a large body of findings, and that that index is easy to compute and interpret, Massey and Denton (1988) recommended using it to measure evenness in most cases.

White (1986) points out, however, that another index may be preferred in measuring segregation between multiple groups, since the dissimilarity index is cumbersome to compute and interpret when the number of groups exceeds two. Thus, if one wants to generate an overall measure of segregation between ten ethnic groups, separate dissimilarity indices will have to be computed between all possible pairs of groups and averaged to get a single measure. An alternative index is Theil's (1972) entropy index, which yields a single comprehensive measure of ethnic segregation. The entropy index also can be expanded to measure segregation across two or more variables simultaneously (e.g. ethnicity and income) and can be decomposed into portions attributable to each of the variables and their interaction (see White 1986).

The second dimension of segregation is exposure, which refers to the degree of potential contact between groups within the neighborhoods of a city. Exposure indices measure the extent to which groups must physically confront one another because they share a residential area. For

any city, the degree of minority exposure to the majority is defined as the likelihood of having a neighborhood in common. Rather than measuring segregation as a departure from an abstract ideal of "evenness," however, exposure indices get at the *experience* of segregation from the viewpoint of the average person.

Although indices of exposure and evenness are correlated empirically, they are conceptually distinct because the former depend on the relative size of the groups that are being compared, while the latter do not. Minority group members can be evenly distributed among the residential areas of a city but at the same time experience little exposure to majority group members if they constitute a relatively large share of the population of the city. Conversely, if they constitute a small proportion of the city's population, minority group members tend to experience high levels of exposure to the majority regardless of the level of evenness. Exposure indices take explicit account of such compositional effects in determining the degree of segregation between groups.

The importance of exposure was noted early by Bell (1954), who introduced several indices. However, with the establishment of the *Pax Duncanae* in 1955, sentiment coalesced around the dissimilarity index and exposure was largely forgotten until Lieberson reintroduced the  $P^*$  index in the early 1980s (Lieberson 1980, 1981). This index has two basic variants. The interaction index ( ${}_xP^*_y$ ) measures the probability that members of group X share a neighborhood with members of group Y, and the isolation index ( ${}_xP^*_x$ ) measures the probability that group X members share an area with each other.

The interaction index is computed as the minority-weighted average of each neighborhood's majority proportion:

$${}_xP^*_y = \sum_{i=1}^n \left[ \frac{x_i}{X} \right] \left[ \frac{y_i}{t_i} \right] \quad (2)$$

where  $x_i$ ,  $y_i$ , and  $t_i$  are the numbers of group X members, group Y members, and the total population of unit  $i$ , respectively, and  $X$  represents the number of group X members citywide. The isolation index is computed as the minority-weighted average of each neighborhood's minority proportion:

$${}_xP^*_x = \sum_{i=1}^n \left[ \frac{x_i}{X} \right] \left[ \frac{x_i}{t_i} \right] \quad (3)$$

Both indices vary between zero and one and give the probability that a randomly drawn group X member shares a neighborhood with a member of group Y (in the case of  ${}_xP^*_y$ ) or with another group X member (in the case of  ${}_xP^*_x$ ). Values of  ${}_yP^*_x$  and  ${}_yP^*_y$  can be computed analogously from equations (2) and (3) by switching the  $x$  and  $y$  subscripts. When there are only two groups, the isolation and interaction indices sum to one, so that  ${}_xP^*_y + {}_xP^*_x = 1.0$  and  ${}_yP^*_x + {}_yP^*_y = 1.0$ . The interaction indices are also asymmetrical; only when group X and group Y constitute the same proportion of the population does  ${}_xP^*_y$  equal  ${}_yP^*_x$ .

$P^*$  indices can be standardized to control for population composition and eliminate the asymmetry (Bell 1954; White 1986). Standardizing the isolation index yields the well-known correlation ratio, or  $\eta^2$  (White 1986). Stearns and Logan (1986) argue that  $\eta^2$  constitutes an independent dimension of segregation, but Massey and Denton (1988) hold that it straddles two dimensions. Since it is derived from  $P^*$ ,  $\eta^2$  displays some properties associated with an exposure measure, but standardization also gives it the qualities of an evenness index. Massey and Denton (1988) demonstrate this duality empirically and argue that it is better to use  $D$  and  $P^*$  as separate measures of evenness and exposure. Nonetheless, Jargowsky (1996) has shown that one version of  $\eta^2$  yields a better and more concise measure of segregation when one wishes to measure segregation between multiple groups simultaneously (e.g., between income categories).

The third dimension of segregation is clustering, or the extent to which areas inhabited by minority group members adjoin one another in space. A high degree of clustering implies a residential structure in which minority areas are arranged contiguously, creating one large enclave, whereas a low level of clustering means that minority areas are widely scattered around the urban environment, like a checkerboard.

The index of clustering recommended by Massey and Denton (1988) is White's (1983) index of spatial proximity, SP. It is constructed by calculating the average distance between members of

the same group and the average distance between members of different groups and then computing a weighted average of those quantities. The average distance, or proximity, between group X members is

$$P_{xx} = \sum_{i=1}^n \sum_{j=1}^n \frac{x_i x_j c_{ij}}{X^2} \quad (4)$$

and the average proximity between members of group X and group Y is

$$P_{xy} = \sum_{i=1}^n \sum_{j=1}^n \frac{x_i y_j c_{ij}}{XY} \quad (5)$$

where  $Y$  is the number of group Y members city-wide,  $x_i$  and  $y_j$  are the numbers of group X and group Y members in units  $i$  and  $j$ , and  $c_{ij}$  is a distance function between these two areas, defined here as a negative exponential:  $c_{ij} = \exp(-d_{ij})$ . The term  $d_{ij}$  represents the linear distance between the centroids of units  $i$  and  $j$ , and  $d_{ij}$  is estimated as  $(.6a_i)^{.5}$ , where  $a_i$  is the area of the spatial unit. Use of the negative exponential implicitly assumes that the likelihood of interaction declines rapidly as the distance between people increases.

Average proximities also may be calculated among group Y members ( $P_{yy}$ ) and among all members of the population ( $P_u$ ) by analogy with equation (4). White's SP index (1983) represents the average of intragroup proximities,  $P_{xx}/P_u$  and  $P_{yy}/P_u$ , weighted by the fraction of each group in the population:

$$SP = \frac{XP_{xx} + YP_{yy}}{TP_u} \quad (6)$$

SP equals one when there is no differential clustering between group X and group Y and is greater than one when group X members live nearer to each other than they do to group Y members. In practice, SP can be converted to a zero-to-one scale by taking the quantity  $SP-1$  (Massey and Denton 1988). White (1984) also has proposed a more complex standardization by taking  $f(d_{ij})=d_{ij}^2$ , which yields a statistic equivalent to the proportion of spatial variance explained.

Jakubs (1981) and Morgan (1983a, 1983b) have proposed that  $D$  and  $P^*$  be adjusted to incorporate the effects of clustering. Massey and Denton



(1988) argue against this procedure because it confounds two different dimensions of segregation. They maintain that it is better to measure clustering directly as a separate dimension than to try to adjust other measures to reflect it.

The fourth dimension of segregation is centralization, or the degree to which a group is located near the center of an urban area. In the postwar period, African-Americans became increasingly isolated in older central cities as whites gravitated to the suburbs. Centralization is measured by an index that reflects the degree to which a group is spatially distributed close to or far away from the central business district (CBD). It compares a group's distribution around the CBD to the distribution of land area around the CBD by using a formula adapted from Duncan (1957):

$$CE = \left( \sum_{i=1}^n X_{i-1} A_i \right) - \left( \sum_{i=1}^n X_i A_{i-1} \right) \quad (7)$$

where the  $n$  areal units are ordered by increasing distance from the CBD and  $X_i$  and  $A_i$  are the respective cumulative proportions of group X members and land area in unit  $i$ .

In most circumstances, the centralization index varies between plus one and minus one, with positive values indicating a tendency for group X members to reside close to the city center and negative values indicating a tendency for them to live in outlying areas. A score of zero means that the group has a uniform distribution throughout the metropolitan area. The index states the proportion of group X members who would have to change their area of residence to achieve a uniform distribution around the CBD.

The last dimension of segregation is concentration, or the relative amount of physical space occupied by a minority group in the urban environment. Concentration is a relevant dimension of segregation because discrimination restricts minorities to a small set of neighborhoods that together account for a small share of the urban environment. The index of concentration takes the average amount of physical space occupied by group X relative to group Y and compares that quantity to the ratio that would obtain if group X were maximally concentrated and group Y were maximally dispersed:

$$CO = \frac{\sum_{i=1}^n \frac{x_i a_i}{X} \sum_{i=1}^n \frac{y_i a_i}{Y} - 1}{\sum_{i=1}^{n_1} \frac{t_i a_i}{T_1} \sum_{i=n_2}^n \frac{t_i a_i}{T_2} - 1} \quad (8)$$

where areal units are ordered by geographic size from smallest to largest,  $a_i$  is the land area of unit  $i$ , and the two numbers  $n_1$  and  $n_2$  refer to different points in the rank ordering of areal units from smallest to largest:  $n_1$  is the rank of the unit where the cumulative total population of units equals the total minority population of the city, summing from the smallest unit up, and  $n_2$  is the rank of the areal unit where the cumulative total population of units equals the majority population, totaling from the largest unit down.  $T_1$  equals the total population of areal units from 1 to  $n_1$ , and  $T_2$  equals the total population of areal units from  $n_2$  to  $n$ . As before,  $t_i$  refers to the total population of unit  $i$  and  $X$  is the number of group X members in the city.

In most circumstances, the resulting index varies from minus one to plus one; a score of zero means that the two groups are equally concentrated in urban space, and a score of minus one means that group Y's concentration exceeds group X's to the maximum extent possible; a score of positive one means the converse. In certain circumstances, however, Egan et al. (1998) demonstrate that whenever the number of group X members is very small and the areas in which they live are very large, the index becomes unbounded in the negative direction. Thus, caution should be used in measuring the concentration of groups with very few members.

Which of these five indices of segregation is chosen for a particular application depends on the purpose of the study. All are valid measures, and arguments about which one is "correct" or "best" are meaningless, since they measure different facets of segregation.  $D$  provides an overall measure of evenness that is highly comparable with prior work, widely understood, readily interpretable, and independent of population composition.  $P^*$  captures the degree of inter- and intragroup contact likely to be experienced by members of different groups and directly incorporates the effect of population composition. Most recent work has relied most heavily on these two segregation mea-



sures (Massey and Denton 1993; Frey and Farley 1994, 1996; Massey and Hajnal 1995; Peach 1998).

Neither  $D$  nor  $P^*$  is inherently spatial, however, and each may be applied to study nongeographic forms of segregation, such as segregation between men and women across occupations (see Jacobs 1989). The remaining three dimensions are relevant whenever it is important to know about the physical location of a group in space. If the extent to which group members cluster is important, SP should be computed; if it is important to know how close to the city center a group has settled, CE may be calculated; and if the amount of physical space occupied by a group is relevant, CO is the appropriate index.

The most comprehensive understanding of residential segregation is achieved, however, when all five indices are examined simultaneously. That multidimensional approach yields a fuller picture of segregation than can be achieved by using any single index alone. Thus, Massey and Denton (1989) found that blacks in certain U.S. cities were highly segregated on all five dimensions simultaneously, a pattern they called "hypersegregation." Denton (1994) has shown that this pattern not only persisted to 1990 but extended to other metropolitan areas. By relying primarily on the index of dissimilarity, prior work overlooked this unique aspect of black urban life and understated the severity of black segregation in U.S. cities.

## REFERENCES

- Atkinson, A. B. 1970 "On the Measurement of Inequality." *Journal of Economic Theory* 2:244-263.
- Bell, Wendell 1954 "A Probability Model for the Measurement of Ecological Segregation." *Social Forces* 32:357-364.
- Burgess, Ernest W. 1928 "Residential Segregation in American Cities." *Annals of the American Academy of Political and Social Science* 14:105-115.
- Cortese, Charles F., R. Frank Falk, and Jack C. Cohen 1976 "Further Considerations on the Methodological Analysis of Segregation Indices." *American Sociological Review* 41:630-637.
- Denton, Nancy A. 1994 "Are African Americans Still Hypersegregated?" In Robert D. Bullard, J. Eugene Grigsby III, and Charles Lee, eds., *Residential Apartheid: The American Legacy*. Los Angeles: CAAS Publications, University of California.
- Duncan, Otis D. 1957 "The Measurement of Population Distribution." *Population Studies* 11:27-45.
- , and Beverly Duncan 1955 "A Methodological Analysis of Segregation Indices." *American Sociological Review* 20:210-217.
- Eagan, Karla L., Douglas L. Anderton, and Eleanor Weber 1998 "Relative Spatial Concentration among Minorities: Addressing Errors in Measurement." *Social Forces* 76:1115-1121.
- Frey, William H., and Reynolds Farley 1994 "Changes in the Segregation of Whites from Blacks during the 1980s: Small Steps toward a More Integrated Society." *American Sociological Review* 59:23-45.
- 1996 "Latino, Asian, and Black Segregation in U.S. Metropolitan Areas: Are Multiethnic Metros Different?" *Demography* 33:35-50.
- Jacobs, Jerry A. 1989 *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, Calif.: Stanford University Press.
- Jakubs, John F. 1981 "A Distance-Based Segregation Index." *Journal of Socio-Economic Planning Sciences* 15:129-136.
- James, David R., and Karl E. Taeuber 1985 "Measures of Segregation." In Nancy B. Tuma, ed., *Sociological Methodology 1985*. San Francisco: Jossey-Bass.
- Jargowsky, Paul A. 1996 "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." *American Sociological Review* 61:984-998.
- Lieberson, Stanley 1980 *A Piece of the Pie: Blacks and White Immigrants since 1880*. Berkeley and Los Angeles: University of California Press.
- 1981 "An Asymmetrical Approach to Segregation." In Ceri Peach, Vaughn Robinson, and Susan Smith, eds., *Ethnic Segregation in Cities*. London: Croom Helm.
- Massey, Douglas S., and Nancy A. Denton 1988 "The Dimensions of Residential Segregation." *Social Forces* 67:281-315.
- 1989 "Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions." *Demography* 26:373-393.
- 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Massey, Douglas S., and Zoltan Hajnal 1995 "The Changing Geographic Structure of Black-White Segregation in the United States." *Social Science Quarterly* 76:527-542.
- Massey, Douglas S., Michael J. White, and Voon-Chin Phua 1996 "The Dimensions of Segregation Revisited." *Sociological Methods and Research* 25:172-206.

- Morgan, Barrie S. 1983a "An Alternate Approach to the Development of a Distance-Based Measure of Racial Segregation." *American Journal of Sociology* 88:1237-1249.
- 1983b "A Distance-Decay Interaction Index to Measure Residential Segregation." *Area* 15:211-216.
- Peach, Ceri 1998 "South Asian and Caribbean Ethnic Minority Housing Choice in Britain." *Urban Studies* 35:1657-1680.
- Stearns, Linda B., and John R. Logan 1986 "Measuring Segregation: Three Dimensions, Three Measures." *Urban Affairs Quarterly* 22:124-150.
- Theil, Henri 1972 *Statistical Decomposition Analysis*. Amsterdam: North Holland.
- White, Michael J. 1983 "The Measurement of Spatial Segregation." *American Journal of Sociology* 88:1008-1019.
- 1984 "Reply to Mitra." *American Journal of Sociology* 90:189-191.
- 1986 "Segregation and Diversity: Measures in Population Distribution." *Population Index* 52:198-221.

DOUGLAS S. MASSEY

## SELF-CONCEPT

The self is the central concept used to represent the individual in sociological social psychology. The importance of the self reflects the influence of symbolic interactionism in sociology. In the last twenty years social psychologists trained in psychology also have developed a strong interest in the self as their emphasis has shifted from behaviorism to cognitive theories (Baumeister, 1998).

### STABLE SELF-CONCEPTS

The social psychological conception of the self is based on the idea that people are reflexive, responding to themselves just as they respond to other "objects." Since reflexive thinking requires language, it is assumed that infants and nonhuman animals lack a self-concept. However, there is some evidence that chimpanzees are aware of what they look like, since they notice markings on their faces (Gallup 1977). This self-recognition suggests that some animals and prelinguistic humans have a rudimentary sense of self but that it lacks meaning or content.

Some sociologists, particularly those with a philosophical and qualitative orientation, view the

self as a process involving people's internal conversations. Those with a more positivistic and quantitative orientation emphasize more stable aspects of the self. From their point of view, the self-concept refers to all the ways in which people describe themselves. These linguistic descriptions refer to the way people think they are, not to their actual personal characteristics.

People describe themselves in many different ways. One way to find out about the content of the self-concept is to ask respondents to answer the question "Who am I?" Studies of responses to this question reveal that people often think of themselves in terms of their roles (or role identities). For example, people often describe themselves on the basis of their sex, age, race, and occupation. Stryker (1968) suggests that these and other roles are organized in a hierarchy according to their salience for a person. The salience of a role is based in part on the extent to which adequate performance of that role affects relationships with "significant others." Salience is also a function of how distinctive a role is (McGuire and Padawer-Singer 1976). For example, a female is more likely to mention her gender in describing herself if she is in a group of males.

People also describe themselves in terms of personal attributes, such as "lazy," "smart," and "attractive." In contrast to roles—which usually are described with nouns—these self-concepts are more likely to be defined by adjectives. They often reflect individuals' conceptions of their abilities or performance in different roles. For example, on a questionnaire children can be asked, "How smart in school do you think you are, among the smartest, above average, average, or below average?" Respondents sometimes try to be objective in answering this type of question and to place themselves according to the criteria they think the researcher is using. Sometimes they report more subjective feelings about where they stand in accordance with their own standards. For example, professional athletes may be dissatisfied with their level of play even if they think they are better than most people. Other personal attributes involve self-attributed traits such as "aggressive" or "nice." Also included here are the ways in which people characterize their beliefs and attitudes. For example, people may conceive of themselves as prejudiced or not independently of whether they are prejudiced by an objective standard.

While individuals think of themselves in terms of specific roles and specific evaluations of their personal attributes, they also have a more general opinion of themselves (Gecas and Burke, 1995). This global evaluation Brown, 1993 called *self-esteem* and is measured by statements such as "I feel I do not have much to be proud of" and "At times I think I am no good at all" (Rosenberg 1965). The global nature of self-esteem is indicated by the tendency of individuals to describe themselves as consistently positive or negative on different personal attributes. However, self-esteem also has different dimensions, such as self-efficacy and self-worth (Gecas 1982). Self-esteem, like depression and anxiety, usually is considered an aspect of mental health. Research using longitudinal data has shown that self-esteem affects and is affected by depression among adolescents (Rosenberg et al. 1989).

There is considerable evidence that people are motivated to enhance their self-esteem. For example, respondents tend to give inflated evaluations of themselves on anonymous questionnaires. In addition, subjects in experiments are more likely to explain their successes in terms of internal attributes, such as effort and ability, while attributing their failures to external factors, such as task difficulty (Bradley 1978).

### **SITUATIONAL SELF-IMAGES**

Some self-statements are more temporary than those described above, involving the roles or personal attributes people use to describe themselves in particular situations. For example, a woman may think of herself as a "teacher" when she is talking to her students and as "foolish" when she has made a mistake. If repeated, these situational images may become stable as people come to believe them. Emotions also can be considered temporary self-concepts if one thinks of them as statements about how people say they feel rather than as a physiological process. Thinking about emotions in this way leads to the examination of how emotions are affected by social processes.

Some scholars focus on the presentation of situational self-images to others (Goffman 1959). Borrowing language from the theater, they view behavior as a performance displayed in front of an audience, a form of self-presentation or impression

management. This approach presents a challenge to those who attempt to measure self-concepts, since it suggests that responses on questionnaires reflect self-presentation rather than privately held beliefs. Researchers try to minimize this problem by using carefully worded questions and guaranteeing anonymity.

Situational self-images often are studied in laboratory experiments. For example, subjects may be asked to respond after receiving false feedback about themselves. To determine whether a self-description involves impression management, the privacy of subjects' responses may be manipulated. When behavior in front of an audience is different from behavior performed in private, this suggests that the behavior reflects impression management rather than privately held beliefs. This type of research also can tell researchers something about how behavior is affected by subjects' awareness that they are being studied. Some behaviors in experimental settings have been shown to result from subjects doing what they think is expected of them (Orne 1962).

Self-presentation behavior is particularly likely to occur when people have done something that is apt to gain disapproval from an audience. When people find themselves in these "predicaments," they are embarrassed and engage in various forms of "facework" to avoid a negative image. Frequently, people give excuses and justifications in an attempt to explain their behavior and avoid condemnation from others. Research shows that subjects are more likely to use self-presentation tactics when they are dependent on the audience for rewards. An important role of self-presentation in conformity, altruism, aggression, and other behaviors has been demonstrated. For example, self-presentation processes are important in explaining the behavior of bullies and the tendency for people to retaliate when attacked (Tedeschi and Felson 1994).

### **DEVELOPMENT OF SELF-APPRAISALS**

Three processes have been used to explain why people have favorable or unfavorable opinions about themselves: (1) attribution, (2) comparison, and (3) reflected appraisal. The first two processes have been emphasized by psychological social psychologists, while the third has been the focus of

sociological social psychologists, particularly those sympathetic toward symbolic interactionism.

According to attribution theory, people learn about themselves and others in similar ways. Individuals base judgments about themselves on observations of their own behavior just as they base judgments about others on their observations of those people's behavior (Bem 1972). These judgments are socially influenced, since beliefs about the association of behaviors and personal attributes are learned from others. In judging their abilities, for example, people rely in part on observations of their performances on tasks they believe reflect those abilities. Thus, children who get high grades tend to attribute more ability to themselves. Individuals are likely to attribute a high level of ability to themselves when there is a consistent pattern of success (Kelley 1967).

When people view their behavior as being caused by external forces, they treat it as uninformative about themselves. However, when they view their behavior as being internally caused, there is likely to be some change in their self-appraisals. For example, research shows that external rewards sometimes can reduce the motivation of children to do things they have enjoyed in the past, such as playing with magic markers (Deci and Ryan 1980). If they are rewarded for playing with magic markers, they tend to lose interest when they are no longer rewarded because they attribute their behavior to the reward rather than to their intrinsic motivation. The external reward can decrease their interest in the behavior because it affects their judgments about why they did it. More generally, there is evidence that people's behaviors can affect their attitudes, just as their attitudes can affect their behaviors (Liska et al. 1984). For example, a person may decide that she likes an activity because she observes herself voluntarily engaging in that activity.

Comparison processes are also important factors in the development of self-appraisals. They affect the standards people use in evaluating their behavior. For example, students may think a B is a good grade or a poor grade depending on the standard they use. Standards are a function of two types of comparisons. A *temporal comparison* is a comparison of present performance and past performance. People are likely to judge their recent performances more harshly if they have been suc-

cessful in the past. A *social comparison* is a comparison of one's own behavior to the behavior of others. The more successful the others are, the higher the standard is and the more negative the self-appraisal is. Thus, subjects are more negative in describing themselves when there is another person with very positive qualities present than they are when that person has negative qualities (Morse and Gergen 1970). This implies that self-appraisals tend to be more favorable if one is a "big fish in a small pond." For example, research shows that high school students tend to have more negative self-appraisals of their academic ability if their schoolmates are bright (Felson and Reed 1986). However, sometimes the performance of others has a positive effect on self-appraisal. This occurs when people "borrow status" from successful others with whom they are associated and "bask in reflected glory" (Cialdini et al. 1976).

Festinger (1954) suggested that social comparison processes result from the desire to gain accurate appraisals of one's abilities and to find out whether one's opinions are correct. When objective information is not available, people compare themselves to others. Further, Festinger suggested that people usually choose similar others for comparison because the behavior of those persons provides the most information. Some research has examined the hypothesis that people evaluate their abilities by comparing themselves to others who are similar to themselves on attributes (other than ability) that are related to performance. For example, comparisons with people who have engaged in a similar effort will be the most informative. Similarly, if a boy believes that gender is related to athletic performance, he will compare himself to other boys in order to decide how much athletic ability he has.

According to the reflected appraisal process, people come to see themselves as others see them, or at least as they think others see them. This notion of the "looking-glass self" focuses on how individuals think they appear to others (Cooley 1902). According to Mead (1934), this helps explain the initial formation of self in young children. Mead suggested that when children role-play, they respond to themselves when they play the role of others. This role-taking process leads them to see themselves as objects. Later, the appraisals of significant others shape the specific content of people's self-concepts. The appraisals

of others are accurately perceived and then are incorporated into the self-concept. Significant others may have special expertise or may be parents or close friends, but those who influence one aspect of the self-concept do not necessarily influence other aspects.

Experimental research suggests that subjects' self-appraisals are affected by the false feedback they receive from others. Survey research—which examines correlations between self-appraisals, the appraisals of significant others, and a person's perception of those appraisals—suggests that the appraisals of significant others are not perceived very accurately (Schrauger and Schoeneman 1979). Apparently, rules of politeness limit the amount of open communication—particularly criticism—making it difficult for people to find out what others think of them (Felson 1980). When feedback is given, it tends to involve specific comments about behavior rather than global evaluations. When praise is given, it often is not believed. As a result, people usually have only vague, general impressions of what others think of them and self-appraisals tend to be idiosyncratic and idealized (Felson 1989). While others are in some agreement about a person, that person does not share in the consensus. In addition, ambiguous feedback allows people to think more favorably about themselves and thus protect their self-esteem.

This discussion also applies to global self-esteem. Educators and parents may overemphasize the importance of praise in the development of self-esteem in children. While there is evidence that parents' praise and other supportive behavior affects the self-esteem of children (Felson and Zielinski 1989), successful performance in activities that children value may be more important.

There are other processes that increase the correspondence between people's appraisals of themselves and the appraisals of others. First, in some instances, people have access to the same information others have. For example, children's self-appraisals of their ability and their friends' appraisals of them correspond, because both are affected by the children's grades. Second, some other people can influence self-concepts if they have control over formal evaluations. For example, evidence shows that teachers influence self-concepts because they assign grades, but that is usually the extent of their influence in this area.

The discussion above has focused on the interpersonal environment. Social-demographic characteristics also affect self-appraisal. Social class, for example, has been shown to affect the self-esteem of adults but not that of children (Rosenberg and Pearlin 1978). Blacks and whites, by contrast, have similar levels of self-esteem (Porter and Washington, 1993; Wylie, 1979). A key element here appears to be whether people associate with others who are like themselves. The self-esteem of minority group members is likely to be lower in more heterogeneous settings where invidious comparisons are made and where members of higher-status groups may act in prejudicial ways.

### CONSEQUENCES OF SELF-CONCEPTS

The way individuals think of themselves has an important impact on how they behave. Thus, people who think of themselves in terms of particular role identities tend to act in ways that are consistent with those identities. For example, a man who identifies himself as a father will engage in the behaviors he associates with being a father. These roles provide links between the individual and society. Individuals are plugged into the social structure through the roles that are mapped onto selves. In other words, role performance reflects the way people think about themselves. Of course, people vary in terms of the importance they attach to different roles. When people must decide between roles, they tend to choose the role more salient to them (Stryker 1968). For example, the choice between doing work and playing with children on a Sunday afternoon may reflect the relative salience of family and occupational roles.

Success and failure frequently are attributed to variations in self-confidence. Self-appraisals and performance are certainly correlated, but this does not necessarily mean that the former causes the latter. Longitudinal studies, which attempt to disentangle these causal relationships, suggest that students' global self-esteem does not affect their academic performance. However, there is evidence that specific self-appraisals of ability affect performance. Longitudinal analyses of high school students suggest that self-appraisals of academic ability have an effect on grades (Felson 1984). Self-appraisals of ability affect performance through two processes: effort and test anxiety. Those who are self-confident about their ability are likely to

work harder because they think effort will bring success. In addition, they are less anxious when they are tested, and so nervousness does not interfere with their performance. However, the effect of self-appraisal on performance probably is not as strong as people think. The effect of grades on self-appraisal is much stronger, suggesting that success is more likely to lead to self-confidence than self-confidence is to lead to success.

Causal interpretation has been problematic in the study of the effects of self-concept on other behaviors as well. While self-esteem and various self-appraisals have been shown to correlate with behavior, it is difficult to show that the self-concepts cause the behaviors. Relatively few studies have attempted to sort out these relationships. Exceptions include longitudinal studies that suggest that low self-esteem increases delinquency among adolescents (e.g., Kaplan 1980; Rosenberg et al. 1989). In general, criminologists today are more likely to attribute criminal behavior to low self-control than to low self-esteem.

An interesting experimental method for examining the effects of self-concept on behavior has been suggested by Duval and Wicklund (1972). They suggest that since much human behavior is automatic or habitual, people do not always think about themselves before they engage in a behavior. These authors argue that self-concepts affect behavior when attention is directed toward the self rather than toward the environment, a condition they call "objective self-awareness." Objective self-awareness is likely to occur when people are in unfamiliar surroundings, when there are disruptions in social interaction, and when people find themselves in a minority. Mirrors are commonly used in experiments to create objective self-awareness. These studies show that subjects are more likely to engage in behavior that is consistent with their self-standards when they are facing a mirror (e.g., Beaman et al. 1979). In addition, there are individual differences: Some people are more chronically focused on themselves as objects. The behavior of such people is more likely to be consistent with their self-appraisals and internalized standards.

A number of researchers have examined the role of self-concepts in resisting change. For example, research suggests that people are motivated to

reaffirm self-concepts when they are challenged (Swann 1984). Markus (1977) considers the generalizations people make about themselves as "self-schemas" that affect the way they process information. Self-schemas usually refer to personality traits (e.g., "independent" and "generous") that people attribute to themselves on the basis of past actions. Once formed, they affect the information people attend to and remember and how quickly they process it. For example, people are more likely to learn and recall information that is associated with their self-schemas. In other words, self-schemas act like filters, guiding the processing of incoming information. Thus, self-schemas have a conservative function because they lead people to focus on information that is consistent with their views of themselves.

### SUMMARY

The determinants and consequences of the self have become central concerns for both sociologically and psychologically trained social psychologists. Self-concepts depend on the way individuals think they are viewed by others, on individuals' observations of their behavior, and on the standards individuals use to judge that behavior. These judgments in turn depend on the performance (for comparison) and appraisals of others. Self-concepts have consequences in that they affect which roles are performed and how successfully they are performed. They also affect conformity and deviance and the management of impressions. Finally, they are important in their own right as indicators of mental health.

### REFERENCES

- Baumeister, Roy 1998 "The Self." In D. T. Gilbert, S. T. Fiske & G. Lindzey, eds., *Handbook of Social Psychology*, 4th ed. New York: McGraw-Hill.
- Beaman, A. L., B. Klentz, E. Diener, and S. Svanum 1979 "Objective Self-Awareness and Transgression in Children: A Field Study." *Journal of Personality and Social Psychology* 37:1835-1846.
- Bem, Daryl 1972 "Self-Perception Theory." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 6. New York: Academic Press.
- Bradley, G. W. 1978 "Self-Serving Biases in the Attribution Process: A Reexamination of the Fact or Fiction

- Question." *Journal of Personality and Social Psychology* 36:56-71.
- Brown, J. D. 1993 "Self-esteem and Self-evaluation: Feeling is Believing." In J. Suls, ed., *Psychological Perspectives on the Self*. Hillsdale, NJ: Erlbaum.
- Cialdini, R. B., R. J. Borden, A. Thorne, M. R. Walker, S. Freeman, and L. R. Sloan 1976 "Basking in Reflected Glory: Three (Football) Field Studies." *Journal of Personality and Social Psychology* 34:366-374.
- Cooley, Charles H. 1902 *Human Nature and the Social Order*. New York: Scribner's.
- Deci, E. L., and R. M. Ryan 1980 "The Empirical Exploration of Intrinsic Motivational Processes." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 13. New York: Academic Press.
- Duval, S., and R. A. Wicklund 1972 *A Theory of Objective Self-Awareness*. New York: Academic Press.
- Felson, Richard B. 1980 "Communication Barriers and the Reflected Appraisal Process." *Social Psychology Quarterly*, 43:223-233.
- Felson, Richard B. 1984 "The Effects of Self-appraisals of Ability on Academic Performance." *Journal of Personality and Social Psychology* 47:944-952.
- 1989 "Parents and the Reflected Appraisal Process: A Longitudinal Analysis." *Journal of Personality and Social Psychology* 56:965-971.
- , and Mark Reed 1986 "The Effect of Parents on the Self-Appraisals of Children." *Social Psychology Quarterly*, 49:302-308.
- , and Mary Zielinski. 1989 "Children's Self-Esteem and Parental Support." *Journal of Marriage and the Family* 51:727-735.
- Festinger, Leon 1954 "A Theory of Social Comparison Processes." *Human Relations* 7:117-140.
- Gallup, G. G., Jr. 1977 "Self-Recognition in Primates: A Comparative Approach to the Bidirectional Properties of Consciousness." *American Psychologist* 32:329-338.
- Gecas, Viktor 1982 "The Self-Concept." *Annual Review of Sociology* 8:1-33.
- Gecas, Viktor, and Peter J. Burke 1995 "Self and Identity." In K. S. Cook, G. A. Fine and J. S. House, eds., *Sociological Perspectives on Social Psychology* Needham Heights, Mass.: Allyn & Bacon.
- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. New York: Doubleday.
- Kaplan, Howard B. 1980 *Deviant Behavior in Defense of Self*. New York: Academic Press.
- Kelley, H. H. 1967 "Attribution Theory in Social Psychology." In D. Levine, ed., *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- Liska, A., R. Felson, M. Chamlin, and W. Baccaglini 1984 "Estimating Attitude-Behavior Relations within a Theoretical Specification." *Social Psychology Quarterly* 47:15-23.
- Markus, H. 1977 "Self-Schemata and Processing Information about the Self." *Journal of Personality and Social Psychology* 35:63-78.
- McGuire, William J., and A. Padawer-Singer 1976 "Trait Salience in the Spontaneous Self-Concept." *Journal of Personality and Social Psychology* 33:743-754.
- Mead, G. H. 1934 *Mind, Self and Society*. Chicago: University of Chicago Press.
- Morse, S., and K. J. Gergen 1970 "Social Comparison, Self-Consistency, and the Concept of Self." *Journal of Personality and Social Psychology* 16:148-156.
- Orne, Martin T. 1962 "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications." *American Psychologist* 17:776-783.
- Porter, J. R., and R. E. Washington 1993 "Minority Identity and Self-esteem." *Annual Review of Sociology* 19:139-161.
- Rosenberg, M. 1965 *Society and the Adolescent Self-Image*. Princeton, N.J.: Princeton University Press.
- , and Leonard Pearlman 1979 "Social Class and Self-Esteem among Children and Adults." *American Journal of Sociology* 84:53-87.
- , Carmi Schooler, and Carrie Schoenback 1989 "Self-Esteem and Adolescent Problems." *American Sociological Review* 54:1004-1018.
- Schrauger, J. S., and T. J. Schoeneman 1979 Symbolic Interactionist View of Self: Through the Looking Glass Darkly. *Psychological Bulletin*, 86:549-573.
- Stryker, Sheldon 1968 "Identity Salience and Role Performance: The Relevance of Symbolic Interaction Theory for Family Research." *Journal of Marriage and the Family* 30:558-564.
- Swann, W. B. 1984 "Self-Verification: Bringing Social Reality into Harmony with the Self." In J. Suls and A. G. Greenwald, eds., *Psychological Perspective on the Self*, vol. 2. Hillsdale, N.J.: Erlbaum.
- Tedeschi, James T. and Richard B. Felson 1994 *Violence, Aggression, & Coercive Actions*. Washington, D.C.: APA Books.
- Wylie, Ruth 1979 *The Self-Concept: Revised Edition*, vol. 2: *Theory and Research on Selected Topics*. Lincoln: University of Nebraska Press.

RICHARD B. FELSON

## SELF-ESTEEM

Self-esteem is a concept that has been used to explain a vast array of emotional, motivational, and behavioral phenomena. Most Americans believe intuitively that low self-esteem is undesirable; indeed, the link between low self-esteem and depression, shyness, loneliness, and alienation supports the general idea that low self-esteem is an aversive state. The view that self-esteem is a vital component of mental health is also evident in the popular media and in educational policy. Low self-esteem has been viewed as the root cause of societal problems ranging from drug abuse to teenage pregnancy to poor school performance. A number of educational and therapeutic programs have been developed to solve these problems by increasing self-esteem. Self-esteem is one of the most frequently examined constructs in sociology and psychology, with more than 15,000 research articles referring to it over the past thirty years. This entry reviews the research that has focused on the conceptual and functional basis of self-esteem.

Self-esteem is defined as the evaluative component of the self-concept, the extent to which people view themselves as likable and worthy as opposed to unlikable and unworthy. As a self-reflexive attitude, self-esteem is composed of cognitive and affective components. Self-esteem is related to personal beliefs about skills, abilities, and future outcomes as well as the strategies people use to gain self-knowledge. However, the personal experience of self-esteem is more emotional than rational. Some people dislike themselves in spite of objective evidence suggesting that they should feel very good about themselves. Many successful doctors, lawyers, professors, and entrepreneurs are filled with self-loathing despite their objective career success.

The term "self-esteem" sometimes is used interchangeably with terms such as "self-confidence," "self-efficacy," and even "self-concept," but such usage is inaccurate and should be discouraged. Self-confidence and self-efficacy refer to the belief that one can attain specific outcomes. Although people with high self-esteem often are self-confident, evaluative reactions to personal outcomes vary greatly, and it is possible for people to be confident about attaining a goal without feeling good about themselves in the process. The term "self-concept" refers to the components of self-

knowledge and includes things such as name, race, ethnicity, gender, occupation, likes and dislikes, and personality traits. As such, self-concept refers to cognitive beliefs and other forms of self-relevant knowledge (Felson 1992). Although self-esteem clearly is influenced by the contents of the self-concept, they are not the same thing.

### STRUCTURE AND MEASUREMENT OF SELF-ESTEEM

An important issue in the literature on self-esteem is whether self-esteem is best conceptualized as a unitary global trait or a multidimensional trait with independent subcomponents. An example of a multidimensional trait model is Tatarodi and Swann's (1995) differentiation between self-liking and self-competence. From this perspective, it is possible for people to like themselves generally but view themselves as not particularly efficacious at various tasks. Conversely, it is possible for people to view themselves as generally competent but not really like themselves. Mismatches between self-liking and self-competence lead to biases in the interpretation of social and performance feedback that confirm the level of self-liking. For instance, those who are high in self-liking but low in self-competence perceive negative feedback more positively than do those who are low in self-liking but high in self-competence.

Global self-esteem is best conceptualized as a hierarchical construct with three major components: *performance* self-esteem, *social* self-esteem, and *physical* self-esteem. Each component can be broken down into progressively smaller subcomponents. Performance self-esteem refers to one's sense of general competence and includes intellectual ability, school performance, self-regulatory capacities, self-confidence, efficacy, and agency. People who are high in performance self-esteem believe that they are smart and capable. As will be discussed below, personal beliefs about performance are poorly related to objective outcomes. Social self-esteem refers to how people believe they are perceived by others. It is perception rather than reality that is critical here. If people believe that others, especially significant others, value and respect them, they experience high social self-esteem even if others truly dislike them or hold them in contempt. The influence of these reflected appraisals on self-esteem is an integral part



of Cooley's (1902) "looking-glass self" and has been implicated in the development of self-esteem by sociological theorists such as George Herbert Mead and Stanley Rosenberg. People who are low in social self-esteem often experience social anxiety and are high in public self-consciousness. They are highly attentive to their public images and worry about how others view them. Physical self-esteem refers to how people view their physical bodies and includes things such as athletic skills, physical attractiveness, and body image as well as physical stigmas and feelings about race and ethnicity.

How are these subcomponents of self-esteem related to global self-esteem? James (1892) proposed that global self-esteem is the summation of specific components of self-esteem, each of which is weighted by its importance to the self-concept. In other words, people have high self-esteem to the extent that they feel good about the things that matter to them. Not being good at tennis is irrelevant to the self-concept of a nonathlete, and doing poorly in school may have little impact on inner-city youth who do not identify with mainstream values (Steele 1997). Pelham (1995) and Marsh (1995) debated the value of global versus specific component models. Pelham's research generally supports the Jamesian view that the centrality of self-views is an important predictor of the emotional response to the self (i.e., one's feelings of self-esteem), whereas Marsh claims that domain importance does not have a strong impact on self-esteem. Although the jury is still out on this issue, the concept of domain importance is a central feature of most theories of self-esteem.

In terms of measurement, most research uses global measures of self-esteem, since this is viewed as having the greatest theoretical importance (Baumeister 1998). The most widely used measure of global self-esteem is the Rosenberg (1965) scale, which consists of ten general statements such as "On the whole, I am satisfied with myself," "I certainly feel useless at times," and "I take a positive attitude toward myself." Unfortunately, scores on this scale tend to be tightly clustered around its mean, limiting its predictive value. A review of self-esteem measures conducted by Blascovich and Tomaka (1991) recommended the Revised Feelings of Inadequacy Scale (Fleming and Courtney 1984), which is a modified version of the Janis and Field (1959) scale. This scale has five factors: social

confidence, school abilities, self-regard, physical appearance, and physical ability. The total score on this scale is widely used as a measure of global self-esteem.

Another issue in the measurement and definition of self-esteem is whether it is best conceptualized as a stable personality *trait* or a context-specific *state*. Most theories of self-esteem view it as a relatively stable trait: If one has high self-esteem today, one probably will have high self-esteem tomorrow. Around this stable baseline, however, there are fluctuations; although people generally may feel good about themselves, there are times when they may experience self-doubt and even dislike themselves. In terms of research, the selection of trait or state measures of self-esteem depends on whether one is interested in predicting long-term outcomes or in the immediate effects associated with feelings about the self. Obviously, measures of state self-esteem are more useful for the latter group. The State Self-Esteem Scale (Heatherton and Polivy 1991) is a commonly used measure that has been shown to be sensitive to laboratory manipulations of self-esteem. This scale measures context-specific feelings related to performance, social, and physical self-esteem. Measures of trait and state self-esteem are highly correlated. However, frequent fluctuations in state self-esteem have been found to be associated with increased sensitivity to and reliance on social evaluations, increased concern about how one views the self, and even anger and hostility (Kernis 1993). In general, those with a fragile sense of self-esteem respond extremely favorably to positive feedback and extremely defensively to negative feedback.

#### SOURCES AND FUNCTIONS OF SELF-ESTEEM

A central issue in self-esteem pertains to its source. Research in psychology and sociology has focused on the role of early childhood experiences, especially in terms of parental treatment. Harter has incorporated the Jamesian and Cooley views of the development of self-esteem into a general model of self-esteem development (Harter 1993). Harter proposes that reflected appraisals about important dimensions affect the development of self-esteem but that specific domains are closely linked to potential audiences. According to her theory, parents are particularly concerned about behav-

ioral conduct and school performance, and therefore, children's beliefs about how their parents view them on these dimensions influence self-esteem. Children who do well scholastically and behave in accordance with parental expectations believe that their parents support and love them. However, parents have less impact than do peers on self-perceptions related to physical appearance, athletic ability, and peer likability. To obtain the support of one's peers, children believe they have to be attractive, athletic, and likable. Failure to obtain support from either parents or peers can lead to feelings of hopelessness, depression, and poor global self-esteem. Harter's model offers a significant advance over earlier developmental theories by integrating importance and social support from a domain-specific perspective.

Further evidence for socialization processes can be found when one considers the influence of gender differences. A number of studies suggest that boys and girls diverge in their primary sources of self-esteem, with girls being more influenced by relationships and boys being more influenced by objective success. Stein et al. (1992) examined participants in an eight-year longitudinal study of adolescent growth and development. During adolescence, an agentic orientation predicted heightened self-esteem for males but not for females, whereas a communal orientation predicted heightened self-esteem for females but not for males. The possibility that males and females differ in terms of what constitutes the self-concept was also addressed by Josephs et al. (1992). In a series of studies, men and women were given false feedback indicating that they had deficits either on a performance dimension (e.g., competition, individual thinking) or on a social dimension (e.g., nurturance, interpersonal integration). Consistent with predictions, men high in self-esteem enhanced their estimates of being able to engage successfully in future performance behaviors, whereas women high in self-esteem enhanced their estimates of being able to engage successfully in future social behaviors. The authors of this entry recently compared the experiences of boys and girls in a summer tennis camp designed to increase self-esteem (Hebl et al. 1999). Scores on a children's version of the state self-esteem scale showed that both boys and girls had increases in overall self-esteem during the tennis camp, but whereas boys gained self-esteem primarily in the performance self-esteem

domain, girls gained self-esteem primarily in the social self-esteem domain. In each case it can be seen that boys gain self-esteem from getting ahead whereas girls gain self-esteem from getting along.

From a completely different perspective, some researchers have begun to explore the possibility that self-esteem is determined more by biology than by socialization. Although direct evidence is minimal, there is circumstantial evidence that some components of self-esteem are based in biology. Twin studies have suggested that self-esteem is moderately heritable, with estimates ranging from 30 to 50 percent (Kendler et al. 1998). In addition, traits known to be associated with self-esteem, such as extraversion and neuroticism, have long been known to have a genetic component. Kramer (1993) argues that self-esteem is rooted in activity of the serotonergic neurotransmitter system. He notes that pharmacological treatments that increase the activity of serotonin are associated with an increased sense of self-confidence and self-esteem. However, there have not been any systematic or rigorous tests of this hypothesis. The possibility that self-esteem has a biological component remains an important empirical issue.

Some theorists have portrayed self-esteem as a mechanism that has evolved through adaptation to promote survival of the species. Accordingly, self-esteem is viewed as a force that promotes feelings of confidence and competence that may lead to superior performance across a broad range of activities. Interestingly, this perspective can be used to explain gender differences in the major sources of self-esteem. Throughout human evolutionary history, males were of value to the group primarily through their role as hunters and protectors whereas women gathered food and nurtured offspring. Hence, being good at tasks closely associated with ancestral sex roles may be associated with increased feelings of self-esteem. However, because theories of evolution and socialization predict the same gender pattern for self-esteem, it is impossible to clarify which perspective is correct or whether they are equally correct or incorrect. Baumeister (1998) has noted that simple evolutionary accounts of self-esteem are difficult to accept because of the rather negligible benefits associated with self-esteem and the possibility that high self-esteem may promote overconfidence and excessive risk taking.

A novel and important functional account of self-esteem has been proposed by Leary and his colleagues. Leary begins with the assumption that humans have a fundamental need to belong that is rooted in evolutionary history (Baumeister and Leary 1995). For most of human evolution, survival and reproduction depended on affiliation with a group. Those who belonged to social groups were more likely to survive and reproduce than were those who were excluded from groups and left to survive on their own. According to Leary, self-esteem functions as a monitor of the likelihood of social exclusion. When people behave in ways that increase the likelihood that they will be rejected, they experience a reduction in state self-esteem. Thus, self-esteem serves as a monitor, or sociometer, of social acceptance and/or rejection. At the trait level, those with high self-esteem have sociometers that indicate a low probability of rejection and therefore do not worry about how they are perceived by others. By contrast, those with low self-esteem have sociometers that indicate the imminent possibility of rejection and therefore are highly motivated to manage their public impressions.

#### **CONSEQUENCES OF HAVING HIGH OR LOW SELF-ESTEEM**

Self-esteem has both cognitive and affective components. Accordingly, a number of researchers have examined the cognitive and affective reactions of those with high and low self-esteem. The overall view suggests that people process information in a way that confirms and supports their chronic self-views. People with high self-esteem actively defend their positive self-views, whereas those with low self-esteem appear to be less able to do so. This section reviews research that has examined differences between individuals with high and low self-esteem.

Self-esteem differences have been reported for a wide range of intrapsychic phenomena, including emotional reactions, cognitive processes, and motivational states. There are obvious differences in how individuals with high and low self-esteem feel about themselves; the positivity and negativity of self-feelings are of course central to self-esteem. For instance, people with low self-esteem are more likely to report being depressed and anxious than are those with high self-esteem. These differences appear to be more subjective

than objective. Researchers have used diary studies to examine whether high and low self-esteem people differ in their daily moods and emotions (Campbell et al. 1991). Compared with individuals with high self-esteem, individuals with low self-esteem judged the events in their lives more negatively and as having a greater impact on their moods. However, when outside judges read participants' diaries, they could not distinguish between the events experienced by participants with high and low self-esteem. Thus, similar circumstances are perceived and experienced differently as a function of a person's self-esteem level. In terms of specific emotional states, there are no differences in how high and low self-esteem individuals experience impersonal emotions (e.g., happiness), but there are differences in how they experience self-relevant emotions (e.g., pride and shame). People with high self-esteem are more likely to report pride, whereas those with low self-esteem are more likely to report shame. Once again, this pattern is independent of actual events in the lives of people with high and low self-esteem.

A robust finding in social psychological research is that everyone feels good after receiving positive feedback regardless of one's self-esteem level. People with low self-esteem like to hear good things about themselves just as much as do people with high self-esteem, and both groups hope to be successful in life. However, people with high self-esteem are much more likely to believe positive feedback. People with low self-esteem are distrustful of overly positive feedback because it contradicts what they believe to be true about themselves. Swann (Swann et al. 1987) argues that people with low self-esteem are attracted to negative information because it validates and confirms their negative self-views. Swann likens the conflict between an emotional preference for positivity and a cognitive preference for negativity to being caught in the cross fire between two warring factions.

A consistent theme in the literature on self-esteem is that self-esteem involves a cognitive bias in processing evaluative and social information. In a world filled with ambiguities and uncertainties, people selectively construct their own reality through biased encoding, retrieval, and interpretation of life events. Research on information-processing styles shows that high self-esteem is associated with cognitive strategies aimed at en-

hancing self-appraisals and thinking of oneself in the most positive way. These objectives are accomplished by means of deeper encoding and more frequent retrieval of positive self-knowledge coupled with an avoidance of negative self-relevant information. That is, people with high self-esteem pay attention to information that says good things about them but ignore information that challenges a positive self-view. By contrast, the processing style associated with low self-esteem is one of self-consciousness and rumination. Individuals with low self-esteem focus on their own thoughts and feelings, often dwelling on negative life events. They are vigilant for information that confirms a negative self-view and ruminate on past failures, embarrassments, and setbacks in a nonproductive fashion.

Biased information processing helps people maintain their existing self-views. For example, an individual with high self-esteem may create a personal definition of what it means to be a "good student" that includes areas in which he or she excels while downplaying the importance of areas of personal deficiency. People with high self-esteem also believe that their talents are unique and special but that their weaknesses are common and trivial. As a result of their selective processing of evaluative information, people with high self-esteem are more adept at defending their self-esteem from external threats. Thus, people with high self-esteem discredit sources of negative feedback while readily accepting positive feedback. These people are also more likely to show a self-serving bias, which refers to the tendency of individuals to take personal credit for success but to blame failure on external circumstances. Some studies in fact find a reversal of the self-serving bias in individuals with low self-esteem such that they credit success to the environment (e.g., an easy task or luck) and blame themselves for failure. People with low self-esteem appear to be generally less able to put a positive spin on negative personal information. For instance, after receiving a negative evaluation, individuals with low self-esteem are more likely to dwell on their weaknesses while individuals with high self-esteem recruit thoughts about their strengths. Differential responses to feedback appear to be an automatic consequence of self-esteem that does not require effort or conscious initiation (Dodgson and Wood 1998).

Baumeister and Tice (Baumeister et al. 1989) have proposed that the basic distinction between

high self-esteem and low self-esteem is motivational. People with high self-esteem are concerned primarily with self-enhancement, whereas people with low self-esteem are concerned primarily with self-protection. The self-enhancement motive emphasizes feeling good about oneself with the aim of increasing one's self-esteem. Thus, people with high self-esteem look for areas in which they can excel and stand out. When they fail at a task, they set higher goals so that they can prove they possess exceptional skills. By contrast, people with low self-esteem are concerned with avoiding humiliation, embarrassment, and rejection. Low self-esteem individuals' self-protective orientation guards against feeling even worse about themselves. Thus, when they fail at a task, they set more modest goals so that they do not lose further esteem through failure.

Self-esteem is known to be relevant to interpersonal behavior. For instance, high and low self-esteem people differ in their perceptions of interpersonal acceptance or rejection. A study in which participants were told that they had been rejected or accepted by their peers (Nezlek et al. 1997) showed that individuals with low self-esteem perceived peer inclusion and exclusion accurately, corresponding to experimental feedback. High self-esteem individuals, however, always perceived inclusion, even when they had been told they had been personally rejected. In general, people with high self-esteem believe that others admire, like, and respect them, whereas people with low self-esteem do not feel that others provide them with adequate support. This pattern fits in well with the reflected appraisal model of self-esteem that was discussed earlier. Because they are concerned with how others view them and are uncertain about their own beliefs, people with low self-esteem are especially yielding; they change their minds and behaviors to conform to the beliefs and opinions of others. Conversely, people with high self-esteem are confident about their opinions and tend not to be influenced by others. Indeed, they often view their own ideas and beliefs as being superior to those of others.

The currently available evidence paints the cognitive, affective, and social worlds of those with high and low self-esteem to be quite different. But do people truly differ as a function of self-esteem? When people are interviewed, there appear to be great differences between those with high and low

self-esteem. When describing themselves, high self-esteem people say they are physically attractive, intelligent, socially skilled, outgoing, upbeat, optimistic, and satisfied with the state of their lives. Low self-esteem people depict themselves in a much less positive light and describe themselves by using more neutral and negative self-aspects than do those with high self-esteem. Unfortunately, the use of self-reports is problematic because they are confounded by one's level of self-esteem. High self-esteem people generally like and believe favorable things about themselves, and it is not surprising that they rate themselves highly on positive personality traits. However, the extent to which high self-esteem individuals actually possess and exhibit these positive traits—and the extent to which they do not possess or exhibit negative traits—is not well known.

The few studies that have compared claims by high and low self-esteem people to objective standards have not found differences between self-esteem groups. For instance, although people with high self-esteem think of themselves as more attractive than do people with low self-esteem, both groups are seen as equally attractive by others (Diener et al. 1995). Ratings of intelligence show the same pattern: High self-esteem people claim to be more intelligent, but intelligence tests show no differences as a function of self-esteem (Gabriel et al. 1994). Similarly, self-report data indicate that high self-esteem individuals are more likable than are low self-esteem individuals. However, likability ratings of high and low self-esteem people by interaction partners show no relation to self-esteem (Brockner and Lloyd 1986). Overall, most researchers have concluded that there are few differences in objective outcomes between those high and low in self-esteem (Baumeister 1998). A review of the literature by the California Task Force to Promote Self-Esteem and Personal and Social Responsibility (1990) conceded that “the associations between self-esteem and its expected consequences are mixed, insignificant, or absent” (Mecca et al. 1989, p. 15). Self-esteem appears to be related to *subjective* rather than *objective* life outcomes.

It is even possible that high self-esteem is associated with negative outcomes in some contexts. For instance, although high self-esteem typically is associated with superior self-regulation, some evidence suggests that high self-esteem may

interfere with self-regulation when self-esteem is threatened. Relative to when they succeed, failure elicits higher goals and increased persistence—even on unsolvable tasks—in those with high self-esteem (McFarlin et al. 1984). Baumeister et al. (1993) demonstrated that ego threats sabotage self-regulation among those with high self-esteem. In this research, high and low self-esteem participants chose performance contingencies in which greater payoffs were associated with loftier and riskier personal goals. In the control condition, those with high self-esteem set appropriate goals and showed superior self-regulation. However, after ego threat (being told to play it safe if they “didn't have what it takes”), high self-esteem participants set inappropriate, risky goals and ended up with smaller monetary rewards than did participants with low self-esteem. Under threat, participants with high self-esteem also were significantly more likely to choke under pressure (i.e., to show performance decrements under conditions where superior performance is important) than were participants with low self-esteem. These findings suggest that people with high self-esteem are prone to self-regulatory failure in certain situations.

Similarly, extremely positive self-appraisals have been linked to poor interpersonal outcomes, especially when such self-appraisals are challenged or discredited. Baumeister et al. (1996) examined the literature linking self-esteem to interpersonal violence. In contrast to widely held assumptions that low self-esteem is associated with violent actions, they found a consistent pattern in which those who thought highly of themselves but encountered a threat or challenge to their positive self-views were more likely to act in hostile and violent ways. Similarly, Kernis and colleagues (Kernis 1993) have demonstrated that when it is unstable, high self-esteem is associated with increased hostility and aggression. They found that those with unstable high self-esteem are likely to respond to ego threats with self-aggrandizement and defensiveness (Kernis et al. 1997).

Other evidence suggests that those with highly positive self-views may exhibit poor interpersonal skills. Colvin et al. (1995) examined individuals with apparently inflated self-views, as indicated by the difference between self and other ratings. They found that those individuals were viewed by others as hostile and unlikable. In addition, during a structured and highly charged debate, such indi-

viduals engaged in a variety of negatively evaluated interpersonal behaviors, such as bragging, interrupting their partners, and showing an overall lack of genuine concern for their partners. Once again, these patterns are most likely to occur when high self-esteem individuals feel personally challenged. Schlenker et al. (1990) exposed high and low self-esteem subjects to contexts in which they were motivated to make a positive impression on a critical or a supportive audience. They found that high self-esteem subjects became egotistical when the evaluative pressures were greatest. They concluded that "people with high self-esteem become more boastful as the social stakes increase" (p. 891). Similar findings have been reported by Schneider and Turkat (1975), who found that high self-esteem subjects who also expressed high needs for social approval presented themselves much more positively after negative feedback than they did after positive feedback. Perhaps ironically, high self-esteem subjects who receive negative feedback about their intellectual abilities claim to have especially good social skills. However, the available evidence does not support their claims.

### SUMMARY

Having high self-esteem confers a number of benefits to those who possess it: Such people feel good about themselves, are able to cope effectively with challenges and negative feedback, and live in a social world in which they believe that people value and respect them. Although there may be some negative consequences associated with having extremely high self-esteem, most people with high self-esteem lead happy and productive lives. People with low self-esteem see the world through a more negative filter, and their general dislike for themselves influences their perceptions of everything around them. It is striking that the objective record does not validate the subjective experiences of those with high and low self-esteem. The interesting unanswered questions about self-esteem are related to what allows people to hold such positive or negative views of themselves in spite of objective evidence. For instance, it is possible that self-esteem is rooted in neurochemistry and therefore is not sensitive to contextual influence, but this has not been established. It is equally plausible that self-esteem is a cognitive style that develops through early socialization experiences, but this perspective also requires more conclusive

evidence. In any case, the subjective experience of having high or low self-esteem plays an important role in how people interpret the world around them.

### REFERENCES

- Baumeister, R. F. 1998 "The Self." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *Handbook of Social Psychology* 4th ed. Boston: McGraw-Hill.
- , T. F. Heatherton, and D. M. Tice 1993 "When Ego Threats Lead to Self-Regulation Failure: Negative Consequences of High Self-esteem." *Journal of Personality and Social Psychology* 64:141-156.
- , and M. R. Leary 1995 "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation." *Psychological Bulletin* 117:497-529.
- , L. Smart, and J. M. Boden 1996 "Relation of Threatened Egotism to Violence and Aggression: The Dark Side of High Self-Esteem." *Psychological Review* 103:5-33.
- , D. M. Tice, and D. G. Hutton 1989 Self-Presentational Motivations and Personality Differences in Self-Esteem." *Journal of Personality* 57:547-579.
- Blascovich, J., and J. Tomaka 1991 "Measures of Self-Esteem." In J. P. Robinson, P. R. Shaver, and L. S. Wrightsman, eds., *Measures of Personality and Social Psychological Attitudes*, vol. 1. San Diego, Calif.: Academic Press.
- Brockner, J., and K. Lloyd 1986 "Self-Esteem and Likability: Separating Fact from Fantasy." *Journal of Research in Personality* 20:496-508.
- Campbell, J. D., B. Chew, and L. S. Scratchley 1991 "Cognitive and Emotional Reactions to Daily Events: The Effects of Self-Esteem and Self-Complexity." *Journal of Personality* 59:473-505.
- Colvin, C. R., J. Block, and D. C. Funder 1995 "Overly Positive Self-Evaluations and Personality: Negative Implications for Mental Health." *Journal of Personality and Social Psychology* 68:1152-1162.
- Cooley, C. H. 1902 *Human Nature and the Social Order*. New York: Scribner's.
- Diener, E., B. Wolsic, and F. Fujita 1995 "Physical Attractiveness and Subjective Well-Being." *Journal of Personality and Social Psychology* 69:120-129.
- Dodgson, P. G., and J. V. Wood 1998 "Self-Esteem and the Cognitive Accessibility of Strengths and Weaknesses after Failure." *Journal of Personality and Social Psychology* 75:178-197.
- Felson, R. B. 1992 "Self-Concept." In E. Borgatta, ed., *Encyclopedia of Sociology*, New York: Macmillan.

- Fleming, J. S., and B. E. Courtney 1984 "The Dimensionality of Self-Esteem: Hierarchical Facet Model for Revised Measurement Scales." *Journal of Personality and Social Psychology* 46:404-421.
- Gabriel, M. T., J. W. Critelli, and J. S. Ee 1994 "Narcissistic Illusions in Self-Evaluations of Intelligence and Attractiveness." *Journal of Personality* 62:143-155.
- Harter, S. 1993 "Causes and Consequences of Low Self-Esteem in Children and Adolescents." In R. Baumeister, ed., *Self-Esteem: The Puzzle of Low Self-Regard*. New York: Plenum.
- Heatherton, T. F., and J. Polivy 1991 "Development and Validation of a Scale for Measuring State Self-Esteem." *Journal of Personality and Social Psychology* 60:895-910.
- Hebl, M., T. F. Heatherton, and C. Lloyd 1999 "Differential Bases of Boys' and Girls' State Self-Esteem." Unpublished manuscript. Dartmouth College.
- James, W. 1892 *Psychology: The Briefer Course*. New York: Henry Holt.
- Janis, I. L., and P. Field 1959 "Sex Differences and Personality Factors Related to Persuasibility." In C. Hovland and I. Janis, eds., *Personality and Persuasibility*. New Haven, Conn.: Yale University Press.
- Josephs, R. A., H. R. Markus, and R. W. Tafarodi 1992 "Gender and Self-Esteem." *Journal of Personality and Social Psychology* 63:391-402.
- Kendler, K. S., C. O. Gardner, and C. A. Prescott 1998 "A Population-Based Twin Study of Self-Esteem and Gender." *Psychological Medicine* 28:1403-1409.
- Kernis, M. H. 1993 "The Roles of Stability and Level of Self-Esteem in Psychological Functioning." In R. Baumeister, ed., *Self-Esteem: The Puzzle of Low Self-Regard*. New York: Plenum.
- , K. D. Greenier, C. E. Herlocker, C. R. Whisenunt, and T. A. Abend, 1997 "Self-Perceptions of Reactions to Doing Well or Poorly: The Roles of Stability and Level of Self-Esteem." *Personality and Individual Differences* 22:845-854.
- Kramer, P. D. 1993 *Listening to Prozac*. New York: Viking.
- Marsh, H. W. 1995 "A Jamesian Model of Self-Investment and Self-Esteem: Comment on Pelham." *Journal of Personality and Social Psychology* 69:1151-1160.
- McFarlin, D. B., R. F. Baumeister, and J. Blascovich 1984. "On Knowing When to Quit: Task Failure, Self-Esteem, Advice, and Nonproductive Persistence." *Journal of Personality* 52:138-155.
- Mecca, A. M., N. J. Smeiser, and J. Vasconcellos, eds. 1989 *The Social Importance of Self-Esteem*. Berkeley: University of California Press.
- Nezlek, J. B., R. M. Kowalski, M. R. Leary, T. Belvins, and S. Holgate 1997 "Personality Moderators of Reactions to Interpersonal Reactions: Depression and Trait Self-Esteem." *Personality and Social Psychology Bulletin* 23:1235-1244.
- Pelham, B. W. 1995 "Self-Investment and Self-Esteem: Evidence for a Jamesian Model of Self-Worth." *Journal of Personality and Social Psychology* 69:1141-1150.
- Rosenberg, M. 1965 *Society and the Adolescent Self-Image*. Princeton, N. J.: Princeton University Press.
- Schlenker, B. R., M. F. Weigold, and J. R. Hallam 1990 "Self-Serving Attributions in Social Context: Effects of Self-Esteem and Social Pressure." *Journal of Personality and Social Psychology* 58:855-863.
- Schneider, D. J., and D. Turkat 1975 "Self-Presentation Following Success or Failure: Defensive Self-Esteem Models." *Journal of Personality* 43:127-135.
- Steele, C. 1997 "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist* 52:613-629.
- Stein, J. A., M. D. Newcomb, and P. M. Bentler 1992 "The Effect of Agency and Communality on Self-Esteem: Gender Differences in Longitudinal Data." *Sex Roles* 26:465-483.
- Swann, W. B., Jr. 1996 *Self-Traps: The Elusive Quest for Higher Self-Esteem*. New York: W. H. Freeman.
- , J. J. Griffin, Jr., S. C. Predmore, and B. Gaines 1987 "The Cognitive-Affective Crossfire: When Self-Consistency Confronts Self-Enhancement." *Journal of Personality and Social Psychology* 52:881-889.
- Tafarodi, R., and W. B. Swann, Jr. 1995 "Self-Liking and Self-Competence as Dimensions of Global Self-Esteem: Initial Validation of a Measure." *Journal of Personality Assessment* 65:322-342.

## SENTIMENTS

To define the affective state, four terms can be used: passion, state of mind, emotion, and sentiments. The term "passion" is linked with the philosophical and literary tradition and designates a violent tension that the individual sustains for a certain duration. States of mind or moods are affective states of low intensity that are durable and pervasive, lack an immediately perceptible cause, and can influence initially neutral events. The term "emotion" indicates an intense affective state of short duration with a precise external or inner cause, a clear cognitive content, and the

ability to reorient attention. Most scholars agree that emotion is a psychological construct with several components: (1) cognitive: finalized by the stimulus caused by emotion, (2) physiological: from the participation of the neurovegetative system, (3) expressive: linked with movement, (4) motivational: linked with intentions and the tendency to act or react, and (5) subjective: consisting of the sentiment felt by the individual.

Sentiments are more enduring than are emotions (e.g., hatred compared with a momentary explosion of anger) and are more cognitively structured. Sentiments may last seconds (embarrassment) or months (mourning) and may be more or less intense, conscious or unconscious, and controllable or outside one's control. When an individual names a sentiment by using a certain terms he or she refers to different elements: affection, cognitive contents or structures of evaluation, awareness of the level of readiness for action, and awareness of the body. Sentiments pervade daily life and are found in several artistic forms: music, poetry, literature, and painting.

In the philosophical tradition, the affective dimension appeared as a perturbation of human behavior that was considered the result of reasoning. Plato was the first to define passions as "diseases of the spirit." For the Stoics, passion was a disease that gets hold of the entire spirit at the expense of reason: Crisippo considered it an unbridled agitation, an unstoppable force that makes the ego leave the self. For Aristotle, passions act as a destabilizer of the rational powers of the individual. An ambivalent attitude toward passions developed at the beginning of the modern age. They were execrated during the age of illuminism, exalted during the romantic era, condemned during the period of critical rationalism and rehabilitated in the postmodern age.

#### **THE INATTENTION OF THE FOUNDERS**

For a long time sociology, evidently influenced by the hegemonic philosophical paradigm, excluded the study of the affective dimension from its theoretical concerns. Affects thus were relegated to anthropology and psychology. The founders of sociology saw the role of sentiments in social life as rather marginal. An interest in sentiments can be

found in Durkheim, who saw them as agents and factors of cohesion in the formation of solidarity and morality. In his view, sentiments are learned and internalized during rituals and collective ceremonies through the sharing of emotions.

In Pareto's terminology, *residui* are sentiments or the expressions of sentiments inscribed in human nature and *derivazioni* are the conceptual systems of justification with which individuals disguise their passions or give an appearance of rationality to propositions or behaviors that are not rational. People rarely behave in a logical way but always want to make their fellow people believe that their conduct is logical. The main characteristic of human nature is that it lets itself be guided by sentiments and puts forward pseudological justifications for sentimental attitudes. According to Pareto, nonlogical is not necessarily equivalent to illogical: Emotions follow some principles. The logic of sentiments Pareto identified can be summarized in five points. First, while reason is analytic, sentiment is synthetic. Second, sentiment follows justificational and persuasive principles. Third, sentiments are inaccurate, indefinite, and indeterminate. Fourth, sentiments may be ambivalent and conflicting. Fifth, sentiments can take on extreme, absolute characteristics that prevent learning from reality.

Weber begins by distinguishing four types of action: rational action in relation to an aim, rational action in relation to a value, affective or emotional action, and traditional action. The action Weber calls affective results from the state of mind or mood of an individual: the slap a mother gives to her unbearable child or the fist of a soccer player who loses his temper. Action, that is, is defined with reference not to an aim or system of values but to the emotional reaction of an agent who finds himself or herself in certain circumstances. The distinctive characteristic of the world in which people live is rationalization: a firm is rational, as is the bureaucratic management of a state, while scientific action is a combination of a rational action in relation to an aim and a rational action in relation to a value, which is truth.

The importance of the emotional factor in the formation of charisma and the religious spirit is mentioned explicitly by Weber. Perhaps the spirit of capitalism is seen by Weber as deriving from the



pressure of strong and pervasive emotions such as anxiety, desperation, and fear on which Calvinist doctrine was founded. Weber does not ignore affection, but he considers it a parasite of reason, a noise producer, a disturber in the ascent to rationality.

### SOME PRECURSORS

Sentiments are central in the work of Simmel, who thoroughly analyzed their role and formation, especially in his essay on love and his reflections on sociability. For Simmel, emotional reality is the foundation of an individual's experience and social interactions. Individuals, he maintains, enter into relationships with one another through the sentiments. Social relations produce other sentiments that therefore are always connected to interactions. The Italian protosociologist Ferrero studied fear and tried to read the historical existence of civilizations in the light of such sentiments. For Ferrero, fear is the spirit of the universe. People, like all creatures, are molded by fear: To overcome fear, people arm themselves, but because they are armed, they cannot help fearing each other. Consequently, people become the most frightening and the most frightened beings because they are the only ones able to manufacture deadly and thus frightful instruments of offense. Moreover, fantasy creates imaginary dangers of all kinds. The human condition therefore is characterized by a permanent dialogue with fear. Even sacrificial rituals are interpreted as expedients to exorcise the fear that every person has of her or his fellow creatures, who are seen as potential bearers of death. Thus, the salute is a technique of mutual reassurance, a ceremony through which one manifests the absence of aggressive intentions toward others. The main instrument for defeating the original fear is power: the institutionalization of the monopoly of violence. This solution transforms people into subject by taking away their freedom. As a result of the monopoly of violence among those who exercise command, peace is guaranteed, but its price is very high. The fear of war of all against all is eliminated, but it is replaced by fear of power.

However, Ferrero emphasizes that those in power fear the subjects over whom they exercise command. No government is certain of the total

obedience of its citizens, and a revolt can break out among the most compliant subjects. In this way, feedback develops between the use of force and fear: The more those in power fear their subjects, the more they resort to instruments of repression, the more they are repressive, and the more they feel fear. For the usurper of power there is not only an external threat, as Plato, Senofonte, Machiavelli, and Montesquieu showed, but also an internal torment: fear of one's own illegitimacy. In *Community and Society*, Toennies, in theorizing about the fundamental category of "community," emphasized the sentiment of belonging. He recognized in sentiments a relative autonomy when he assumed that they were a guide for action and that the fundamental directions of human action depend largely on inner conditions (dispositions) and to a lesser extent on external conditions (circumstances). Veblen based his concept of the wealthy class on a complex mix of sentiments (superiority and inferiority, emulation and imitation) and showed their influence on collective behavior.

Scheler showed, through a phenomenology of resentment, the tight link between the sphere of individual sentiments and that of collective behavior. For Scheler, the forming of resentment and its mode of expression have an individual component tied to temperament and a social component: "The way and the measure in which resentment takes shape in entire groups and in individuals is linked in the first place to the disposition of the human material at issue, in the second place to the structure of society in which this lives." Therefore, there is constant interaction between individual attitudes and collective mentality. The manifestation of resentment is linked to social inequalities and the inability to heal an offense, true or presumed, through vengeance or rebellion. The spreading of resentment creates the psychological premises that become linked with the structural premises, leading to the explosion of mass movements.

For Scheler, "the formal structure of the expression of resentment is always the same: A is approved of, supported and praised not for his intrinsic qualities, but with the intention—that remains unsaid in words—to deny, devalue and reprove B." The refusal of wealth is nothing but a repressed desire for wealth among the poor, who despise what they want but cannot have. Like

Tocqueville, Scheler sees in democracy the fuel of resentment. He maintains: "It will possess a maximum charge of resentment a society in which political rights and others of almost equal kind, jointly with a formal social parity publicly recognized, go together with great differences of power, of effective possession of assets and effective cultural formation."

In the 1960s, MacIver explored the sentiment of attachment to a community, referring to the complex of memories, traditions, customs, and institutions shared by the members of a community. This sentiment of community, which is acquired through the socialization process, has three main elements: (1) the sentiment of we, (2) the sentiment of role, and (3) the sentiment of dependence. The first leads to the identification of a person with the others; the second is associated with the functions individual members perform in the community and expresses the way in which the individual normally realizes his or her belonging to the entire community. The sentiment of dependence expresses the dependence of a person on the community, which is considered a necessary condition for his or her existence, lessening his or her isolation. In modern society, an individual belongs not only to a single community but also to specific associations, as social life is subject to the process of differentiation and specialization. Moreover, the sentiment of belonging to the community seems to persist only in rural areas; in urban areas, it is replaced by attachment to other, less inclusive and more sectorial groups. In rural areas, the sentiment of attachment to the community remains strong because of people's longer periods of residence. An enlarged sentiment of community is the sentiment of nationality, which is produced by historical circumstances and supported by common psychological factors.

Helen Lynd was the first to recognize the importance of shame in the complex dynamics of identity growth. According to Lynd, the values of a culture do not stimulate the experience and communication of emotions such as shame, joy, wonder, and love. However, once trust has been created—allowing one to reveal oneself and thus look for the ways in which to communicate shame—the risk of being exposed may become an experience of abandonment, self-revelation, and intimacy with others. Once accepted, shame leads a person to a

greater awareness and to an enrichment of his or her personality.

### CONTEMPORARY APPROACHES

Sociologists' interest in emotions and sentiments can be traced to the mid-1970s. In that period, one of the foundations of sociology—the idea of a rational social actor guided in his or her conduct, motivation, and choices by a utilitarian and instrumental logic while sentiments play a residual and socially insignificant role—was questioned. This model started showing its abstractness and rigid partiality and its inadequacy in explaining the complexity of individual and collective action, which depends on how people feel and manage their emotions. Thus, emotions started losing their residual and even disturbing function and were recognized as integral parts of social action.

This crack in the rational actor construction went together with a new cultural climate characterized by attention to the self and to the world of private and interpersonal relations, a desire for authenticity, and the reevaluation of sentiments. Emotions and sentiments became objects of knowledge, attention, and communication both in daily life and in sociological reflection. In women and youth subcultures especially, the expression of one's sentiments was encouraged. Advertising saw the positive value of the affective dimension and used it in its campaigns by linking it to products potentially able to stimulate emotions, such as sports, film, television, cars, and even food. The reevaluation of the affective dimension also was stimulated by the increase in the number of magazines dealing with psychoanalytic issues and the growth of university courses on psychological subjects: In many textbooks, at least one chapter is devoted to the dynamics of sentiments and emotions. A widespread representation of emotions has resulted from cinematic and television fiction.

In the growing social sciences literature on sentiments and emotions, five approaches can be singled out: (1) sociohistorical, (2) positivist, (3) functionalist, (4) conflict, (5) interactionist, and (6) socialconstructivist.

**The Sociohistorical Approach.** According to Elias, the phenomenon of civilization is a matter of controlling emotions and natural impulses. From

the documents he collected, it can be inferred that anger, which now seems psychopathic, was in medieval times considered normal, and that the desire to destroy was openly confessed, while it now is admitted with shame or used to attract attention.

For Elias, emotions and their expressive forms are strongly correlated with the social contexts in which they manifest themselves. Every social structure, he maintains, corresponds to a structure of emotions and sentiments, and their inhibition, repression, or free expression depends on their being functional to different social systems.

According to Elias's theory of the process of civilization, daily life in the Middle Ages was characterized by abuse and violence. People, lived in a permanent state of insecurity and fear. The situation began to change when a stronger territorial power prevailed on the weaker ones and the monopoly of the state's legal violence gradually was established. Violence, reserved to specialized bodies, began to be excluded from other areas of life, and some calm, protected zones started to form. Within those areas, "good manners" developed and eventually replaced violence and abuse in interpersonal relations. Starting in the higher ranks, individuals began to abandon spontaneity and impetuosity and learned to dominate themselves, control their impulses and passions, and regulate aggressiveness.

For Elias, institutionalized sports channel and often divert people's natural emotions. Hunting, for example, often has been recognized as a substitute for war. In their free time, people can carry out activities that incite emotions closely tied to those of normal life.

**The Positivist Approach.** According to Kemper's positivist approach, people have phylogenetically inherited a set of primary emotions—fear, anger, joy, and depression—that serve evolutionary and adaptive needs and are sensitive to certain contingent environmental situations. For human beings, the main contingent environmental situations are of a social nature; therefore, social vicissitudes determine emotions unless society intervenes normatively with different demands.

Kemper states that emotions are universally elicited by two fundamental dimensions of social relationships: power and status. Possessing ade-

quate power produces feelings of security, possessing excessive power produces feelings of guilt, and possessing inadequate power produces anxiety. Fear develops when the power of two interacting individuals is uneven.

Anger emerges from interactions in which an expected, habitual status is denied or withdrawn by an actor considered responsible for the status reduction. Depression is tied to a loss of status, and the person considers himself or herself responsible for the loss. Individuals feel shame if they perceive that their status is too high; in the opposite case, they feel depression. Satisfaction results from interactions in which power is not felt as threatening and the status is expected and desired.

The positivist approach to the study of emotions attempts to (1) treat emotion as a variable, (2) undertake studies that embrace various cultures, (3) develop the history of particular emotions, and (4) examine the link between emotions and sociological variables such as social class, gender, and ethnicity.

**The Functionalist Approach.** In the functionalist perspective, innate emotional behavior may have developed in the course of evolution because of its value in being functional to adaptation. The purpose of emotions is to face emergency situations by estimating the importance of the events in order to organize appropriate action. First, emotions prepare the body to react to hostile or dangerous situations, predisposing the organism for the emergency and thus leading to rebalance and adaptation. As Darwin noted, fear is an emotion common to various species, as it is conducive to survival, aids in adaptation, and is linked to a safeguarding mechanism that is evident in the first phases of life.

Emotions also have the function of interpersonal signaling. That is, they produce the effect of communicating the state of the organism to the outside world. The variety and specificity of emotions reflect the flexibility of the forms of adaptation to the environment. Therefore, for example, fear helps avert danger, sadness is a call for help, and happiness is a phase of recovery after the attainment of a goal.

**The Conflict Approach.** The conflict theory of emotion can be found in the works of Coser and

Collins, according to whom the crucial determinant of emotion is membership in competing groups and classes. Emotion, Coser says, is a resource that can be mobilized, directed, and exploited in conflicts over power. Moreover, it can be used to create emotional loyalty and solidarity. Human biological nature supplies emotional energy, but the differentiation and management of emotion are achieved by competing groups that fight for control over the means of emotional reproduction, that is, the resources for assembling the ritualistic arousal of emotions in support of one's group. According to conflict theory, persons not only manage their own emotions but also to stimulate, suppress, or transform emotions in other people. This also happens at the collective level. For example, hostility perceived as coming from external groups increases in-group solidarity, and emotion is intensified by the presence of a large number of people in an expressive ritual.

For Collins, emotional energy is an essential part of the model of chain ritual interaction, and he maintains that the basis of every interaction is a minimum feeling of positivity toward others. Thus, a person accepted in a conversation receives from that experience not only an increase of positive emotional energy but also additional emotional resources (trust, warmth, enthusiasm) with which to negotiate in the next interaction.

**The Interactionist Approach.** In the interactionist approach, emotions essentially result from social interaction. They differ according to the social and cultural worlds one belongs to and thus continuously change.

Symbolic interactionists draw a distinction between biological emotions and social sentiments. An emotion is fixed in the human organism, is experienced concurrently with bodily change, and consists of a fixed configuration of bodily sensations and gestures in response to simple, standard stimuli. In contrast, a sentiment originates through cultural definition and social interaction and continues over time in enduring social relationships. The expressive gestures and internal sensations of a sentiment are culturally defined and can be altered according to the situation and on a broader cultural scale, since sentiments are not fixed in human biological nature. In the course of interaction, differences among sentiments are developed,

and sentiments are redefined continuously through encounters and relationships.

Goffman has examined the system of rules that dominate interaction and the efforts made by individuals to be in tune with the emotional states of a situation. This control of impressions essentially is a strategy for avoiding embarrassment or shame and is inspired by pride and the desire to make a positive impression.

According to Goffman, embarrassment is linked to the impression of people transmitted to others and is a phenomenon that blocks interactions among individuals and makes them unable to interact. The attitudes deriving from this process violate behavioral standards shared with others and cause inappropriate behaviors. Embarrassment, however, is a normal part of everyday life, since it is a manifestation of adaptation to social organization. One structures in several ways other people's perception of oneself according to the various roles one takes on. What creates embarrassment and uneasiness for an individual is the diversity of multiple interactive situations, which imply different and often discordant rules of behavior. Goffman also describes the process of emotion management: By managing the impressions others have of them, people facilitate their goals and attract the responses they need. Often, however, expression management involves inauthentic sentiments to avoid situationally inappropriate emotional gestures, such as laughter at a funeral or a forced smile on meeting an enemy. Emotion management is guided by implicit and shared social rules about appropriate or inappropriate displays of sentiment and emotion. These emotional rules constitute a normative framework of expectations that people take into account.

**The Social Constructivist Approach.** For constructionist theorists, emotions are not natural answers but experiential and expressive models determined by the sociocultural context; they are learned answers that have neither natural contents nor functions and are tied to the conservation of the individual and the species. From the constructionist point of view, emotions are complex syndromes whose meaning and function can be understood only by referring to the social system of which they are part.

According to constructionists, emotions are functional, as they are socially constructed and pre-

scribed to maintain and support a certain system of values. Therefore, a community will try to promote emotional manifestations that are functional to the maintenance of a socially founded moral order and will try to penalize and remove emotions that are in opposition to the moral order. Examples of how an emotion with a negative connotation, such as hatred, can be promoted and prescribed by a community include Nazi Germany and the Ku Klux Klan. In both cases, hatred became the accepted emotional behavior, as its negative side was socially nationalized. Does the ethics of success on which contemporary Western culture is founded not use in a functional way the emotion of envy as a motivational force for the individual?

### SOCIAL IMPACT OF SENTIMENTS

Imagining a passionate individual in total isolation is impossible: Passions require sharing, participation, and a common horizon of values and rules. A person perceives his or her passions through those he or she is able to stimulate in other people and decodes them on the basis of their reactions. There is no passion without a mental or a real relational context.

“Others” are necessary interlocutors for the expression of sentiments that involve a mutual exchange: Often children wait to be in the presence of an obliging listener before venting their emotions. Moreover, some places are considered socially more suitable than others for expressing certain emotions. For example, aggressiveness is more often shown in a stadium than in a church.

Research on large social groups shows that the distribution of specific emotions (pain, jealousy, love) varies from one social group to another, while studies conducted at the micro level indicate that the experience and expression of an emotion, as well as attempts to control it, are influenced by socialization and the prevailing situational factors.

A society is certainly founded on some sociostructural presuppositions, but it cannot function if it does not also rest on sentiments, beliefs, and obligations. These factors act as a social glue, as an “a priori” that, without being the object of codification or knowledge from individuals is nonetheless necessary for society to function.

One of these psychosocial elements is trust, which represents a precontractual requirement of the fulfillment of social exchanges. As Simmel maintained, society would disintegrate in the absence of trust among men.

It is not trust alone that supports social order, but trust is one of the psychosocial elements that play a fundamental role in maintaining the social system, even if structural theories neglect it. Trust is a powerful cultural resource, a precondition for a full use of other resources, such as entrepreneurship, citizenship, and legality. It is an intangible resource, like the spirit of belonging or consent.

Different social objects can be invested with trust, and so there can be various types of trust:

1. Generalized trust gives people an ontological security about the system.
2. Segmental trust involves entire institutional segments of society, such as the economy and justice.
3. Technological trust involves the expert systems such as transportation and computer science networks.
4. Organizational trust refers to real organizations such as the army and the university.
5. Commercial trust refers to goods with a specific trademark or produced in certain countries.
6. Positional trust is invested in particular social roles, such as priests and doctors.
7. Personal trust can refer to virtues revealed by public or private persons.

The social objects in which trust is placed can be found in one’s own society or in foreign societies. From this fact arises the distinction between inner trust and outer trust. One speaks of focused trust when trust is placed in a particular category of objects; one speaks of diffuse trust when one metaphorically refers to the trustful or distrustful atmosphere that pervades the society. This climate can be incorporated in a culture as well as in a normative perspective. For example, a large part of Italian fiscal legislation appears to be based on distrust of the honesty of the contributors. Trust in a specific object can be justified in an indirect way,

when its reliability depends on the trust that is placed in bodies of supervision and control. In contrast, distrust feeds conflict and social atomization, therefore becoming a powerful destructive force.

Once distrust has been initiated, it soon becomes impossible to know if it was justified, since it has the ability to fulfill itself, to generate a reality that is coherent with it. Observing events that refute past fiduciary relations, for example, finding out that certain scientific data have been counterfeited so that one can publish spectacular results, can cause an unexpected loss of trust. The absence of trust involves an excessive increase in vigilance and the creation of more complex systems of control.

The slump in institutional trust is reflected in a generalized loss of trust in monetary means of exchange, the authorities' legitimacy, the credibility of the political system, and the effectiveness of specialized spheres such as scholastic and religious institutions. A similar collapse of institutional trust can lead to a loss of trust in individuals.

In the postcommunist societies of eastern Europe, there is a widespread lack of trust that leads to a spreading syndrome of diffidence. The Polish sociologist Sztompka has cited some indicators. The strongest behavioral indicator of generalized distrust in the society to which one belongs is choosing to emigrate. This is the obvious form of escape people choose when living conditions become unbearable and signs of improvement cannot be seen. Another form of escape is withdrawing from public life and taking refuge in primary groups. Here there is a horizontal trust that compensates for the lack of vertical trust in institutions. Another indicator of public distrust is the number of protest demonstrations. Yet another is reluctance to think about the long-term future. In the economic field, an indicator of distrust is neglecting investment forms and spending a lot on consumer goods, especially foreign-made goods. The opening and dissemination of casino chains are also important indicators of distrust. Finally, in the postcommunist societies of eastern Europe, trust is undermined because of normative disorganization and the high level of expectations after the "glorious" revolution of 1989.

Some sentiments may derive from the social order, such as loneliness and historical contingen-

cies that generate fear. Some demographic transformations, such as the lengthening of life, and other social changes, such as breaking the bond of marriage or choosing to remain single, amplify the sentiment of loneliness. Feeling lonely has two sides: It is a sad sentiment connected with loss, refusal, and isolation, but it is also a source of serenity, as it means staying with oneself. This develops the inner world and fosters creativity and the birth of the new.

A large amount of the research on loneliness has established only simple correlations between loneliness and other variables. On the whole, it seems that loneliness is more common among women than men, poor people than rich, and adolescents than elderly people. Data on attitudes and sentiments generally indicate that a person who feels lonely also feels distrust and hostility, is rarely attracted by other people, and expects to be rejected. A large number of studies have found positive correlations between loneliness and feelings of impotence and inadequacy, a tendency to self-devaluation, and self-deprecation. A gap can be found in research on loneliness concerning the etiology of the phenomenon. Sociologists are not in a position to establish which factors cause loneliness and which are facilitating conditions. Sociologists do not know, for example, the relationship between objective causes (mournings, transfers, separations) and subjective causes (introversion, shyness, low self-esteem).

In regard to fear, it is necessary to assume that in the past, some events connected to the adversities of nature generated emotional reactions that were stronger than those of today. The historian Delumeau showed how in past centuries, primary concrete fears (of war, scarcity, plague, tyranny, and earthquake) were grafted onto secondary fears that resulted from cultural processes and movements that singled out dangers and adversaries (the heretic, the Hebrew, the witch, the demon, the vampire, the plague spreader) on which it was easier to unload the anguish caused by real but frightening and unmanageable phenomena.

In modern society, fear of machines has sometimes developed. In England at the beginning of the nineteenth century, the Luddites were organized bands of laborers who destroyed textile machines, which caused layoffs of many craftsmen.

Fear of the dominion of the machine is found in many science fiction, utopian, and dystopian novels in this century. In *Erewhon* by Samuel Butler, there are no machines; in Spengler, the machine is devilish and is seen as the dissolver of Western civilization. In the *Brave New World* by Aldous Huxley and *1984* by George Orwell, pervasive technology supplants people's individuality. Similarly dark scenes are accompanying the current computer science phase that features the presence of apocalyptic Cassandras afraid of the effect of computers on the human brain.

In the second half of the 1980s, after the Chernobyl accident, a strong fear of nuclear energy developed, while the 1990s were characterized by the fear of ecological catastrophes tied to the greenhouse effect or the perforation of the ozone layer. The modern age has brought new threats in the form of drugs and the acquired immune deficiency syndrome (AIDS). Sometimes a social fear may be born, the perception, that is, of a "barbarization" of social life, of a progressive loss of freedom, the feeling of being a pawn in the mosaic of daily life, impotent before the great bureaucratic and occult powers. Another source of fear is food, which in Europe was manifested after the Chernobyl accident and during the mad cow crisis. Today a great fear is that of being killed by someone whom one does not know and who has no reason to direct his anger against an individual. Bodyguards represent a new defense against this fear, while the fear of burglars, especially in cities and urban areas, is causing a rush to purchase electronic systems for defense for the home. In Europe, the fear of a demographic decline accompanies the fear of immigration and of being submerged by other ethnic groups.

### EMOTIONS AND CULTURE

A common question among those who study sentiments and emotions is whether these affective states are culturally universal or vary from culture to culture. There are two schools of thought. According to the universalistic school, emotions are always the same in all cultures, while for the relativistic school, they vary in time and among different cultures. Cross-cultural studies have shown differences among cultures concerning the language, the meaning, the objects, and the appraisal

of emotions: One therefore would have to speak of ethnotheories of emotions.

One of the distinctions made by anthropologists is based on the prevalence of sentiments of shame versus guilt. In regard to the fact that a certain culture attributes a preponderant importance to the former or the latter emotion, one speaks of "cultures of shame" and "cultures of guilt": cultures of shame regulate individual behavior through external sanctions, while cultures of guilt are based on interiorized sanctions; that is, guilt is caused by the knowledge of having transgressed important moral or social norms. The cultures of guilt would therefore be those of Western societies, characterized by the Judeo-Christian tradition or the Protestant ethic.

In cultures that systematically resort to humiliation as an educational method, the sentiment of shame is present in massive doses. This happens, for example, in the Japanese culture, which is defined by anthropologists as "the culture of shame." Anthropological data on some cultures of New Guinea show that shame may have a paralyzing effect. In New Guinea, people who have not made a good impression shelter under the porch of the house, cover their faces in chalk powder and wait for the others members to come and playfully reproach them. After a few days, such a person washes his or her face and reenters the group.

The geography of jealousy also shows some variation. In fact, anthropological research shows that jealousy is more widespread and accepted in cultures that are based on private property, authorize sex only within marriage, and grant the status of a responsible adult only to married persons.

In some cultures there are social classes in which a quarrel is unthinkable and an accusation of another person makes clear one's own victimized state. Examples are ritual sounds among the Australian aboriginals, accusing the object of one's anger of witchcraft, and the icy silence of the English bourgeois. In some cultures, such as the Korean, a person who offends is treated with intensified respectful manners, that is, with an increased formality of words and intonation. It is said that the Tasaday did not have in their language a word to express anger. Among the Ibo in Nigeria, anger also seems to be completely absent. A culture, such as the Tahitian, in which social life

is little differentiated, produces few emotions; in this case, one speaks of hypocognition or underidentification.

### EMOTION NORMS

Emotion norms indicate the expected range, intensity, duration, and target of a specific emotion in specific situations. The principles that govern emotion are of four types:

1. Rules of appraisal that define how a situation is perceived and evaluated
2. Rules of behavior that establish how an emotion has to be expressed
3. Prognostic rules that indicate the correct duration of an emotion
4. Rules of attribution that legitimate an emotion in regard to the social system

The rules of expression defined by a given culture and subgroup influence how much the emotional state is expressed in self-description, behavior and bodily expression. Emotion rules prescribe the depth and duration of feelings and provide a measure of how much a certain feeling is unusual, crazy, unsuitable, or normal in a certain social context.

Feeling and expression rules are an integral part of what has been called "emotion culture." Moreover, emotion culture includes beliefs concerning sentiments as well as concepts regarding the way in which one should lend attention, codify, evaluate, control, and express feelings. Emotion culture is reflected in films, religion, psychiatric theories, and the law. Many systems of law mention, for example, two sentiments, widespread among the population, that are worth of being safeguarded: decency and honor.

**Decency.** The sentiment of decency concerns the intimate sphere of one's personality, the area a person thinks should be private and confidential. Analyses carried out in the social sciences reveal extreme variability in the sphere of decency: In different epochs and cultures, different actions and behaviors are considered as damaging this feeling, whereas in other contexts they appear totally legitimate. However, the sentiment of decency is present in all ages and cultures, and

through it, a peculiar relationship occurs between subjective perceptions and objective images of confidentiality, especially in the sexual sphere. Reflecting on the characteristics of the sentiment of decency in a society makes it possible to individuate the prevailing values of a social group concerning the protection of confidentiality but also to explore in depth the reasons for behaviors that do not fall within the sphere of the individual. For example, the meticulous and detailed Starr report on President Clinton disseminated by electronic mail all over the world is an example of the absence of decency in the American judicial system.

**Honor.** Honor is the pretension and the right to be proud, but it also means that society acknowledges this individual pretension. The sentiment of honor spurs honored behavior. Honor is the social acknowledgment of those who fulfill the duties of their status.

In contemporary society, there are a great number of ritual situations in which specific persons are given, while alive, special honors or special offices, degrees, or prizes. After death, these persons are honored with symbolic actions: A recent example was the garlands of flowers that were laid on Mother Teresa's coffin.

A call to honor is made when a citizen sees his or her name and respect wounded by his or her fellow citizens. Civil and penal laws protect honor legally. There are professional groups that supervise the observance of specific norms by means of a code of honor. In this epoch, honor brings to mind retrograde and fragmented elements. In past societies, honor was not only a particular type of sentiment linked to occasional events and one that operated as a cultural superstructure: The conceptions of honor were more than simple accessories of life; they went beyond abstract values one referred to only occasionally. For centuries, they determined the style of life of wide layers of people in premodern societies, orienting their social and private conduct.

Through honor, the individual identifies as member of a group, which in turn confers a status, that is, a dignity and a value in the social relations occurring within and outside the group to which one belongs. The functional cost an individual has to face consists of control of morals through endorsement mechanisms whose most qualifying phase is the formation and application of honor



codes. Acts of violence in defense of the honor of one's family were widely tolerated by the law until relatively recent times.

The Mediterranean culture of honor has been characterized as a phenomenon of archaic societies that have not reached a high level of civilization. With modernization, it was maintained, this dark phase would disappear. Instead, the culture of honor has succeeded in preserving its importance among immigrants in the industrialized metropolis.

The fact that the code of honor has not lost its valence, at least in the Mediterranean area, makes one question whether the linear concept of modernization is truly valid or whether a concept based on a permanent dialectic between persistence and change is more valid.

The regulation of emotion comes largely from social sources that consist partly in the interests that others have and in the norms of social interaction and partly in the expression and feelings rules of society: Boys should not cry, and mourning should not be shown too much, at least in the Nordic countries. Expression rules are specific to a given culture or social role: English people, for example, are little inclined to express how they feel, and the clergy behave with dignity. In some cultures, jealousy cannot be felt, while in others, it must be; in some cultures, an offense to honor is strongly felt, while in others, the honor concept is almost nonexistent. The rules are often explicit to the point of being codified in books of ethics and good manners. In the Western culture, the ethic of control of emotional impulses prevails.

Anger also is subject to social rules. For every class, there are prescriptive rules that indicate the choices that have to be made and proscriptive rules that indicate those to avoid. For example, one can show anger for a behavior that attacks one's honor, freedom, or property (prescriptive rule), but anger cannot exceed the limits of what is necessary to correct the situation (proscriptive rule).

#### SENTIMENTS IN THE POSTMODERN AGE

In the contemporary world, several indications seem to point to a situation in which the consumer society is increasingly inviting people to be passionate but people are distanced from passion. The threshold of emotional involvement has be-

come high to protect people from a useless daily mobilization: In post-modern culture, emotions and social action seem more and more separated.

The process of "blaiseization" that Simmel saw at the beginning of the century thus has accelerated. However, certain episodes occur, like Princess Diana's death, that are experienced with a sentiment of pain that is almost cosmic, so that Pareto's *residuo* of "displaying sentiments with exterior acts" seems still present, although only latently.

The relevance of the affective dimension also is evidenced within the religious and family domain. In religion, several scholars maintain that once people get rid of ritualistic and pretentious display and once reciprocal aggressiveness has disappeared, the prevalent function of religion is to provide affective support of people against the uncertainties of the world. The family which has lost some of its fundamental functions, such as economic and socialization, has maintained and perhaps strengthened its function of affective support both horizontal and vertical.

Within the job world concerns about feelings of alienation during the assembly line period, have been replaced by concerns about the burnout phenomenon, which is a feeling of powerlessness and indifference, which affects especially the health, helping, and education professions.

According to some management scholars, in the next decades a resource to contrast these negative phenomena could be the development of emotional intelligence, which is the ability to effectively perceive, realize, and apply the strength and perspicacity that emotions—meant as sources of energy, information, relations and influence—provide human beings with.

In the 1990s, different authors tried to individuate the prevalence of some sentiments in society or in parts of it. At the beginning of the twenty-first century a sentiment that seems widespread is narcissism, understood as self-complacency and desire to be admired. Some indicators of this feeling are the cult of the body, abandoning every emotional involvement, the proliferation of interpersonal relationships governed by appearance, and a flight from the social. Post-modern society thus seems to be pervaded with a dominant passion: love for the self, which is translated into

vanity and desire for power, wish to control anxiety over self-approval, self-interest, and egoism. Such syndrome, however, seems to co-exist with an altruistic dimension, normally rather latent but that grows in emergencies. The prevailing sentiments seem however to be linked to the spreading of post-materialistic values, such as love and friendship. Along the continuum localism—cosmopolitan, following Huntington, it seems that we are moving towards an attachment for intermediate cultural areas, essentially configured according to the religious tradition.

Contrary to the expectations raised by the fall of the Berlin wall, in various parts of the world (Burundi, Algeria, Bosnia, Kosovo) during the 90s there were carnages caused also by feelings of interethnic hatred.

Another sentiment, felt especially by younger generations, is a worry about the future that generates anxiety and is capable of determining behaviors that are aggressive and antisocial.

Lately mass media scholars complain about the spreading, especially in television and movies, of a certain provocative impudence, together with the weakening of feelings of shame. Moreover, some psychologists have noticed an increase in interpersonal aggressiveness that parallels the increase in the time devoted to city mobility.

In a time of increasing globalization of markets the McDonaldization of tastes, the process of conforming seems to proceed also in the affective sphere, with the consequent weakening of passions and the misting over of the ability to suffer but also to feel joy, indignation, and pity. All these worries had been expressed by the antiutopians, by Riesman, and by the supporters of the critical theory of society. Leading to the valorization of sentiments as bulwark against any attempt of hegemonic oppression. In postmodernity, it thus seems appropriate to place *Homo sentiens* near *Homo faber* and *Homo sapiens*.

## REFERENCES

- Alberoni, Francesco 1991 *Gli Invidiosi*. Milan: Garzanti.  
 Cooper, Robert K., and Ayman Sawaf 1999 *Il Fattore Emozione*. Milano: Sperling and Kupfer.  
 D'Urso, Valentina, and Rosanna Trentin 1998 *Introduzione alla Psicologia delle Emozioni*. Bari: Laterza.

- Elias, Norbert 1978 *The Civilizing Process: The History of Manners*. Oxford, UK: Basil Blackwell.  
 Gambetta, Diego 1988 *Trust: Making and Breaking Cooperative Relations*. Oxford, UK: Basil Blackwell.  
 Goffman, Erving 1967 *Interaction Ritual*. Garden City, N.Y.: Doubleday.  
 Hancock, Barry W. 1986 *Loneliness: Symptoms and Social Causes*. Lanham: University Press of America.  
 Hochschild, Arlie Russel 1985 *The Managed Heart*. Berkeley: University of California Press.  
 Lewis, Michael 1992 *Shame: The Exposed Self*. New York: Free Press.  
 Misztal, Barbara 1996 *Trust in Modern Societies*. Cambridge, UK: Polity Press.  
 Turnaturi, Gabriella 1995 *La Sociologia delle Emozioni*. Milan: Anabasi.  
 Van Sommers, Peter 1988 *Jealousy*. London: Penguin Books.  
 Zingerle, Arnold ed. 1991 "Sociologia dell'Onore." *Annali di Sociologia* 7.

BERNARDO CATTARINUSI

## SEX DIFFERENCES

"Sex differences" is the label used in describing variations between women and men. The term "sex" reflects the division of women and men into two groups on the basis of their unique biological features; the term "differences" derives from the tradition of differential psychology, in which distinct groups of people (defined by natural categories such as sex or constructed categories such as socioeconomic class) are compared in terms of an outcome.

Both terms have been criticized. To many people, references to sex differences imply a biological determinism that ignores the role of socialization and context. The more contemporary term "gender" directs attention to the social meanings assigned to the categories of male and female. With either term, however, it is not necessary to assume a particular causal factor for an observed pattern of variation. A problem with the term "differences" is that it suggests that *dissimilarities* are the norm and *similarities* are the exception. More appropriately, one makes comparisons between groups to see whether there are similarities or differences.

Regardless of the rubric used to characterize the investigation, comparisons between women and men pervade social science literature. Thousands of studies have analyzed sex-related patterns in physical performance, cognitive abilities, personality traits, moral reasoning, social interaction, occupational choice, sexual behavior, attitudes, aggression, depression, self-esteem, leadership, non-verbal behavior, self-disclosure, intelligence, life satisfaction, workplace achievement, and almost every other domain of human activity. To learn how these behaviors emerge, developmental psychologists explore the ways in which specific socialization practices contribute to observed differences between girls and boys and, by extension, women and men. Sociologists study structural features that shape the roles of women and men in organizational settings, family units, and labor markets. All these analyses contribute to the understanding of gender.

Not all of these topics can be discussed in this entry, and this review will not deal with sexual behavior or the biological basis of sex-related differences. Instead, this article will examine some of the main theoretical accounts of sex differences and then turn to 1) sex differences in four domains that were highlighted in early reviews (verbal, quantitative, and spatial ability and aggression), 2) differences in mathematics and science achievement, 3) other analyses of sex differences, and 4) contextual influences on sex-related patterns of behavior.

### THEORETICAL ACCOUNTS OF SEX DIFFERENCES

Two predominant and contrasting theoretical accounts of sex differences are evolutionary approaches and a variety of social roles, expectations, and/or status perspectives (Eagly and Wood 1999). Evolutionary theory predicts that “the sexes will differ in precisely those domains in which women and men have faced different sorts of adaptive problems” in evolutionary history (Buss 1995, p. 164). These domains primarily are those relevant to reproduction, such as mate selection and sexual behavior (Feingold 1992), but also may include physical skills, the commission of violent crimes (Daly and Wilson 1988), and thresholds for physical risk taking (Wilson and Daly 1985). The evolutionary perspective suggests that in domains

irrelevant to sexual selection, the sexes should and do tend to be psychologically similar.

The social roles and/or expectations perspective focuses on the relationship between male gender and high status in American and other societies. Sex differences in behavior may emerge because men are expected to be more competent and authoritative than women, and many social situations are structured to support this outcome (Foschi 1998; Geis 1993; Ridgeway and Diekema 1992). Eagly’s (1987) social role theory further suggests that the unequal distribution of men and women in social roles (e.g., homemaker versus paid employee) contributes to both biased perceptions of the sexes and the development of skills and behavior that fit those roles.

The literature on self-fulfilling prophecy (Geis 1993; Rosenthal and Jacobson 1968) is consistent with this perspective. Expectations lead people to behave in ways that make the expected outcome more likely to occur. Gender-stereotyped beliefs of parents, for example, predict students’ performance in mathematics courses: Parents who believe that girls are inferior in mathematical ability are more likely to have daughters who do poorly in mathematics (Eccles 1985). Deaux and Major’s (1987) interactionist perspective further details how contextual factors determine whether and how gender-related expectations are translated into gendered behavior.

Other accounts of sex differences exist, including biological approaches (emphasizing hormones, brain structure, and genetics), developmental and/or learning approaches (e.g., social learning theory and gender schema theory) (Jacklin and Reynolds 1993), and constructionist perspectives (which describe “doing gender” rather than “having gender”) (West and Zimmerman 1987). The basic theme of “biology or environment” pervades these perspectives. As is discussed below, any theoretical account faces the challenge of accounting for the size and pattern of sex differences across domains and the susceptibility of those differences to contextual moderation.

### WELL-ESTABLISHED SEX DIFFERENCES?

In discussing sex differences, one can begin with Maccoby and Jacklin’s (1974) landmark book *The Psychology of Sex Differences*. In this volume, the

authors review the literature on sex differences and highlight several that are “fairly well established” (p. 351): Females tend to have greater verbal ability than males, and males tend to surpass females in quantitative ability, spatial ability, and aggression.

After the appearance of this narrative review, the statistical technique of “meta-analysis” was developed, enabling researchers to quantitatively combine results across many studies. Rather than rely on subjective impressions of the literature or a “count” of significant and nonsignificant effects, meta-analysis is based on calculation of an effect size ( $d$ ) that reflects the mean sex difference in pooled standard deviation units in a given study.  $D$ 's are combined across studies to compute an average effect size. As a rough guide, Cohen (1977) suggests that  $d$ 's of .20 are small, those of .50 are medium, and those of .80 are large in magnitude. (To take a familiar example of a physical sex difference, the effect size for U.S. male-female differences in height is very large at  $d = 1.93$ .)

Meta-analyses of the four “established” sex differences have reached somewhat different conclusions than did Maccoby and Jacklin's (1974) narrative review. For example, females and males appear to be comparable in verbal ability in general and in most specific types of verbal ability, such as vocabulary, verbal analogies, and reading comprehension (overall  $d = .11$ ) (Hyde and Linn 1988); positive  $d$ 's indicate a relative female advantage. This effect was fairly stable across age despite Maccoby and Jacklin's (1974) conclusion that the onset of sex differences in verbal ability occurs at around age 11. A possible exception to the smallness of sex differences is verbal fluency (e.g., speed and accuracy of speech production and the production of sentences meeting meaning requirements), where females more clearly outperform males ( $d = .33$ ) (Hyde and Linn 1988). It is also the case that males are more prevalent at the low end of the verbal abilities distribution. For example, there are three to four times as many male stutterers as female stutterers (Skinner and Shelton 1985), and severe dyslexia is about 10 times more common in males than in females (Sutaria 1985; see also Halpern 1992).

In regard to quantitative ability, some meta-analyses have reported modest differences between males and females in performance on tests of

mathematical skill ( $d = -.36$  and  $-.43$ ) (Feingold 1988; Hyde 1981). A more recent study, however, found a very small overall effect in the direction of female superiority ( $d = .05$ ), although differences favoring men emerged in high school ( $d = -.29$ ) and college samples ( $d = -.32$ ) (Hyde et al. 1990). Sex differences also tend to be greater among selected samples of mathematically talented youth tested on the mathematics section of the Scholastic Aptitude Test (Benbow 1988). Overall, the average scores of 12- and 13-year-old boys are higher than those of girls of the same age in these samples; the variability in the boys' scores is also greater. As a result, in the upper 3 percent of the distribution defined as mathematically precocious, boys outnumber girls in ratios that sometimes are dramatic (Benbow and Lubinski 1993).

Spatial ability has a possible (though still unclear) link to mathematical aptitude and scientific and engineering achievement. Meta-analyses of sex differences in visual-spatial ability have shown a variety of effect sizes, depending on the specific type of task or skill assessed. For example, on tasks of mental rotation, in which one is asked to visualize the rotation of a three-dimensional object, large sex differences favoring males have been found on one specific type of test (the Shepard-Metzler test;  $d = -.94$ ) but not on others ( $d = -.26$ ) (Linn and Peterson 1986). Tasks involving spatial perception (e.g., determining the true vertical plane when one is seated in a tilted chair) also indicate relatively smaller sex effects ( $d = -.44$ ), and those involving spatial visualization (e.g., finding a simple shape in a complex pattern of shapes) show virtually no sex difference ( $d = -.13$ ) (Linn and Peterson 1986; see also Masters and Sanders 1993; Voyer et al. 1995). Thus, sex differences in this domain are quite task-specific.

The only social behavior highlighted in Maccoby and Jacklin (1974) was aggression. Two meta-analyses of sex differences in aggression that were published a decade apart produced comparable, small effect sizes that indicated less evidence of aggression among females than among males ( $d = -.24$  and  $d = -.29$  in Eagly and Steffen 1986 and Bettencourt and Miller 1996, respectively). Bettencourt and Miller (1996) also identified an important moderating condition of this sex effect: Unprovoked men tend to be more aggressive than are unprovoked women ( $d = -.33$ ), but under conditions of provocation, this sex effect is much

smaller ( $d = -.17$ ). A sense of the relatively small size of these effects can be gleaned by comparing them to the effect of provocation on aggression ( $d = .76$ ). Although sex differences in aggression emerge, they do not seem to be as strong or straightforward as previously was thought.

### MATHEMATICS AND SCIENCE ACHIEVEMENT

Although meta-analyses indicate relatively small sex differences in quantitative ability, an important related question is whether females and males differ in real-world achievements in quantitative domains such as mathematics, science, and engineering.

A report on women's representation in science and engineering by the National Science Foundation (1999) describes several significant findings. First, with regard to secondary education, male and female students are similar in their completion of high school mathematics and science courses (about 17 percent of both sexes have taken trigonometry, and about 25 percent have taken physics). Furthermore, average mathematics scores for females and males in the eighth and twelfth grades are not significantly different, although science scores among twelfth-graders are slightly higher for males than for females (based on the 1996 National Assessment of Educational Progress mathematics and science assessment).

On the mathematics portion of the Scholastic Assessment Test (SAT), men tend to score about 35 points higher than do women, a gap that remains even for students who reported having taken calculus and physics courses in high school. However, this difference appears to be due in part to the fact that a larger number of women than men who take the SAT are from lower-income families. Because parental income is related to SAT scores, the greater proportion of low-income women may reduce the overall female average (National Science Foundation 1999). It is unclear whether this income factor also accounts for findings regarding men's and women's performance on the 1996 College Board Advanced Placement (AP) tests. These data indicated a male advantage over females on AP tests in physics, computer science, chemistry, and calculus ( $d$  from  $-.26$  to  $-.52$ ) as well as an overrepresentation of males in the upper tail and an underrepresentation in the lower

tail of the distribution of test scores (Stumpf and Stanley 1998). Interestingly, across twenty-nine different AP tests, the size of the effect favoring males was highly correlated with the percentage of males taking a given test ( $r = .71$ ). That is, the greater proportion of male to female high school students who choose to take a given AP test, the greater the male advantage in performance. Students taking AP tests are college-bound and self-select test subjects on the basis of their levels of preparation and expertise. Thus, this finding demonstrates that males feel more competent than females in the domains where they outperform females and that sex differences are apparent even among this highly self-selected group of "prepared" male and female students.

The National Science Foundation report (1999) details larger differences in female versus male achievement in regard to the attainment of advanced degrees and employment. Although women received 46 percent of science and engineering bachelor's degrees in 1995, they earned only 31 percent of the total science and engineering doctoral degrees in that year (up from 26 percent in 1985). The only science field in which women received the majority of doctoral degrees (64 percent) was psychology.

With regard to employment, women represent 46 percent of the total U.S. labor force but only 22 percent of scientists and engineers. Representation differs dramatically across fields: More than half of psychologists and 47 percent of sociologists are women, but women account for only 12 percent of physicists and 9 percent of engineers (National Science Foundation 1999). Among those with academic employment, women with science training are more likely than are men to work in elementary or secondary schools and two-year colleges. In four-year colleges, women are less likely than men to be tenured (35 percent versus 59 percent) or full professors (24 percent versus 49 percent). Men also are more likely than women to be managers, and across fields, full-time male scientists and engineers earn more than do females (overall median salary of \$52,000 versus \$42,000). Some, but not all, of these differences in employment placement may be due to differences in age, and women in science and engineering tend to be younger than men. Similarly, salary differences are due in part to age and field differences, since men are more likely than women to enter the high-

paying fields of computer science and engineering. Nonetheless, differences in academic rank remain after one controls for age, and with increasing age, the gap in salaries between female and male scientists and engineers widens (National Science Foundation 1999)

### OTHER SEX DIFFERENCES AND SIMILARITIES

It is impossible to provide a thorough overview of other research on sex differences in this brief article. In addition to the thousands of original empirical studies, over 180 meta-analyses have been published since Hall's (1978) initial meta-analytic study of sex differences in the decoding of nonverbal cues (where  $d = .40$ ).

In a review of the "science and politics of comparing women and men," Eagly (1995) summarizes meta-analytic research by suggesting that the sizes of sex differences vary considerably across domains. The largest differences ( $d > .80$ ) tend to be found for (1) some physical abilities (e.g., throwing distance, speed, and accuracy), (2) the Shepard-Metzler mental rotation task described above, (3) social behaviors such as facial expressiveness, (4) sexual behaviors such as the frequency of masturbation and attitudes toward casual sex, and (5) nurturant personality traits (tender-mindedness) (Feingold 1994). Most other examined attributes, however, tend to support sex differences that are small to moderate (or negligible) in size. For example, the mean effect size is  $-.14$  for studies of sex differences in self-esteem (Major et al. 1999),  $.21$  for fine eye-motor coordination (Thomas and French 1985),  $.02$  for perceived leadership effectiveness (Eagly et al. 1995),  $.26$  for negative attitudes toward homosexuality (Whitley and Kite 1995), and  $.07$  for competitiveness in negotiation (Walters et al. 1998). Ashmore (1990) summarizes his review of the meta-analytic sex differences literature by noting that "relatively large sex differences are demonstrated for physical abilities and for body use and positioning; more modest differences are shown in abilities and social behaviors; and many negligible sex differences are sprinkled across all domains" (p. 500).

Whether one views observed sex differences as "meaningful" depends on one's opinion and perspective. On the one hand, as Eagly (1995) points out, even a large effect size translates into

substantial overlap between the sexes: A  $d$  of around  $.80$  indicates about a 53 percent overlap in female-male distributions. In this manner, findings of difference may mask substantial similarity between the sexes. Furthermore, many sex differences are quite small in magnitude compared with other important psychological phenomena (e.g., the provocation-aggression link of  $.76$  described above and the effect of group pressure on conformity of  $d = 1.06$ ) (Bond and Smith 1996).

On the other hand, a number of other sex differences are comparable in magnitude to psychological effects that typically are interpreted as both theoretically and socially meaningful. For example, the effect of exposure to media violence on aggression is  $d = .27$  (Wood et al. 1991), and the effect of having a type A personality style on systolic blood pressure is  $d = .33$  (Lyness 1993). It is also the case that sex differences based on general population samples may be smaller than the differences that appear when one looks at selected populations or those in the "tails" of frequency distributions. The greater representation of males among the mathematically precocious, the verbally challenged, and the most violently criminal, for example, suggests that for some highly salient important outcomes, sex differences may be more striking than one might assume based on observations of "average" people encountered in everyday experience.

Some have questioned the wisdom of overreliance on meta-analysis, as this quantitative approach cannot overcome the shortcomings or biases in the original set of studies on which the data summary is based. For example, in Eagly and Crowley's (1986) meta-analysis of sex differences in helping behavior, the authors note that their findings are based primarily on studies that involve short-term helping of strangers ("heroic helping"), a type of helping that is more consistent with the male role than the female role. Any finding of sex differences must be taken in light of this fact and not generalized across all helping situations.

Others have illustrated the limitations of meta-analysis for making causal inferences (Knight et al. 1996). For example, a number of meta-analyses of sex differences have indicated a "year of publication" effect: Studies published relatively recently suggest smaller sex differences in domains such as

aggression; verbal, cognitive, and mathematical abilities; helping behavior; influenceability; leadership; and sexuality. Although some have interpreted these findings to mean that sex differences are disappearing (an interpretation more supportive of an environmental than a biological account), changes in research methodology over time may provide an alternative explanation. Indeed, in revisiting a meta-analysis of sex differences in aggression, Knight et al. (1996) found that statistically controlling for method characteristics (e.g., experimental or not, observational or self-report measures, physical or verbal aggression) removed the year of publication effect; small sex differences in aggression appear to have remained stable over time.

The entire enterprise of studying sex differences has its critics, and a number of relevant issues have been questioned and debated in several recent venues (*American Psychologist* 1995; *Feminism and Psychology* 1994; *Journal of Social Issues* 1997). Criticisms center on the tendency in this work to treat females and males as belonging to polar, global categories that ignore the other social characteristics of individuals (race, class, age, sexual orientation), invoke unsophisticated nature versus nurture explanations, ignore social context, and imply that a difference means a deficit. Despite these criticisms, the research literature continues to grow, and there appears to be a trend “toward a more contextualized version of gender-related behavior” (Deaux and LaFrance 1998, p. 816).

#### CONTEXTUAL INFLUENCES ON SEX-RELATED PATTERNS

As is indicated in many of the patterns of sex difference described above, researchers often find important moderators of the sex differences that emerge. For example, experimental work has demonstrated that the way in which a task is framed can affect performance outcomes. Thus, when the “spatial nature” of a spatial reasoning task was emphasized, men outperformed women, but when this aspect of the task was ignored, no sex difference emerged (Sharps et al. 1993). Similar findings have appeared with regard to mathematics performance, as predicted by Steele’s (1997) theory of “stereotype threat.” This perspective suggests that women’s (and minorities) weaker mathe-

matics performance may be based in part on the threat of being judged consistently with negative group stereotypes. When that threat is lifted, performance differences should disappear. In one study, when a mathematics test was described as producing sex differences, women performed worse than did similarly qualified men; when the test was described as not producing sex differences, women and men did not differ in performance (Spencer et al. 1999). Other testing characteristics also may affect the size of sex differences. For example, structured tests that use a free-response format tend to produce smaller sex differences than do those which use a multiple-choice format (Kimball 1989).

The contextual setting also can play a role in the size of sex differences. A case in point is the analysis of leadership effectiveness, where military settings promote sex differences of moderate size ( $d = -.42$ ) but other organizational contexts produce effect sizes of a much smaller magnitude or reversed direction, ranging from  $-.07$  to  $.15$  (Eagly et al. 1995). An analysis of gender and self-esteem indicated the importance of considering age and ethnicity in discussions of sex differences (Major et al. 1999). No sex difference in self-esteem was apparent before adolescence ( $d = -.01$ ), but very modest differences began to emerge around ages 11 through 13 ( $d = -.12$ ). Similarly, sex differences in self-esteem are apparent among whites ( $d = -.20$ ) but virtually nonexistent among North American minority groups, especially African-Americans ( $d = .03$ ). Finally, many abilities addressed in studies of sex differences are easily modified by experience. For example, spatial skills improve with exposure, and specific training programs designed to improve spatial skills are equally effective for women and men (Baenninger and Newcombe 1989).

Thus, task characteristics, familiarity, and social context can create sex differences or make them disappear. As Deaux and Major (1987) suggest, the basic behavioral repertoires of women and men are quite similar (e.g., both women and men know how to be aggressive, how to be helpful, how to smile). What men and women actually do is determined less by differential abilities than by the context in which they act. Norms, expectations, the actions of others, and the actor’s goals and objectives may all combine to produce sex differences in behavior in some circumstances.

No doubt people will continue to ask how or why women and men differ, but the answers will never be simple and the explanations will not fall squarely on the side of biology or that of environment. Evolutionary and social role models ultimately may complement each other in their relative emphases on distal (distant) versus proximal (immediate) causal factors. The causal direction of some reasoning may need to be examined further. For example, observed differences between the sexes cannot be used as a simple explanation for broader gender roles. Instead, accepted roles may channel men and women into different patterns of behavior. Whatever the patterns observed, most sex differences will continue to reflect a gendered environment and be subject to change as social factors shift over place and time.

## REFERENCES

- American Psychologist* 1995 Current Issues. 50:145–171.
- Ashmore, R. D. 1990 “Sex, Gender, and the Individual.” In L. A. Pervin, ed., *Handbook of Personality: Theory and Research*. New York: Guilford Press.
- Baenninger, M., and N. Newcombe 1989 “The Role of Experience in Spatial Test Performance: A Meta-Analysis.” *Sex Roles* 20:327–344.
- Benbow, C. P. 1988 “Sex Differences in Mathematical Reasoning Ability in Intellectually Talented Preadolescents: Their Nature, Effects, and Possible Causes.” *Behavioral and Brain Sciences* 11:169–232.
- , and D. Lubinski 1993 “Consequences of Gender Differences in Mathematical Reasoning Ability and Some Biological Linkages.” In M. Haug, R. E. Whalen, C. Aron, and K. L. Olsen, eds., *The Development of Sex Differences and Similarities in Behaviour*. London: Kluwer Academic.
- Bettencourt, B. A., and N. Miller 1996 “Gender Differences in Aggression as a Function of Provocation: A Meta-Analysis.” *Psychological Bulletin* 119:422–447.
- Bond, R., and P. B. Smith 1996 “Culture and Conformity: A Meta-Analysis of Studies Using Asch’s (1952b, 1956) Line Judgment Task.” *Psychological Bulletin* 119:111–137.
- Buss, D. M. 1995 “Psychological Sex Differences: Origins through Sexual Selection.” *American Psychologist* 50:164–168.
- Cohen, J. 1977 *Statistical Power Analysis for the Behavioral Sciences*. San Diego, Calif.: Academic Press.
- Daly, M., and M. Wilson 1988 *Homicide*. New York: Aldine de Gruyter.
- Deaux, K., and M. LaFrance 1998 “Gender.” In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, 4th ed. Boston: McGraw-Hill.
- , and B. Major 1987 “Putting Gender into Context: An Interactive Model of Gender Related-Behavior.” *Psychological Review* 94:369–389.
- Eagly, A. E. 1987 *Sex Differences in Social Behavior: A Social Role Interpretation*. Hillsdale, N.J.: Erlbaum.
- 1995 “The Science and Politics of Comparing Women and Men.” *American Psychologist* 50:145–158.
- , and M. Crowley 1986 “Gender and Helping Behavior: A Meta-Analytic Review of the Social Psychological Literature.” *Psychological Bulletin* 100:283–308.
- , S. J. Karau, and M. Makhijani 1995 “Gender and the Effectiveness of Leaders: A Meta-Analysis.” *Psychological Bulletin* 117:125–145.
- , and V. J. Steffen 1986 “Gender and Aggressive Behavior: A Meta-Analytic Review of the Social Psychological Literature.” *Psychological Bulletin* 100:309–330.
- , and W. Wood 1999 “The Origins of Sex Differences in Human Behavior: Evolved Dispositions versus Social Roles.” *American Psychologist* 54:408–423.
- Eccles, J. E. 1985 “Sex Differences in Achievement Patterns.” In T. B. Sonderegger, ed., *Psychology and Gender*. Nebraska Symposium on Motivation, 1984. Lincoln: University of Nebraska Press.
- Feingold, A. 1988 “Cognitive Gender Differences Are Disappearing” *American Psychologist* 43:95–103.
- 1992 “Gender Differences in Mate Selection Preferences: A Test of the Parental Investment Model.” *Psychological Bulletin* 112:125–139.
- 1994 “Gender Differences in Personality: A Meta-Analysis.” *Psychological Bulletin* 116:429–456.
- Feminism and Psychology* 1994 “Special Feature: Should Psychologists Study Sex Differences?” 4:507–546.
- Foschi, M. 1998 “Double Standards: Types, Conditions, and Consequences.” *Advances in Group Processes* 15:59–80.
- Geis, F. L. 1993 “Self-Fulfilling Prophecies: A Social Psychological View of Gender.” In A. E. Beall and R. J. Sternberg, eds., *The Psychology of Gender*. New York: Guilford Press.
- Hall, J. A. 1978 “Gender Effects in Decoding Nonverbal Cues.” *Psychological Bulletin* 85:845–857.
- Halpern, D. F. 1992 *Sex Differences in Cognitive Abilities*, 2nd ed. Hillsdale, N.J.: Erlbaum.
- Hyde, J. S. 1981 “How Large Are Cognitive Gender Differences? A Meta-Analysis Using  $w^2$  and  $d$ .” *American Psychologist* 36:892–901.



- , E. Fennema, and S. J. Lamon 1990 "Gender Differences in Mathematics Performance: A Meta-Analysis." *Psychological Bulletin* 107:139–155.
- , and M. C. Linn 1988 "Gender Differences in Verbal Ability: A Meta-Analysis." *Psychological Bulletin* 10:53–69.
- Jacklin, C. N., and C. Reynolds 1993 "Gender and Childhood Socialization." In A. E. Beall and R. J. Sternberg, eds., *The Psychology of Gender*. New York: Guilford Press.
- Journal of Social Issues* 1997 "The Significance of Gender: Theory and Research about Difference." 53:213–408.
- Kimball, M. M. 1989 "A New Perspective on Women's Math Achievement." *Psychological Bulletin* 105:198–214.
- Knight, G. P., R. A. Fabes, and D. A. Higgins 1996 "Concerns about Drawing Causal Inferences from Meta-Analyses: An Example in the Study of Gender Differences in Aggression." *Psychological Bulletin* 119:410–421.
- Linn, M. C., and A. C. Peterson 1986 "A Meta-Analysis of Gender Differences in Spatial Ability: Implications for Mathematics and Science Achievement." In J. S. Hyde and M. C. Linn, eds., *The Psychology of Gender: Advances through Meta-Analysis*. Baltimore: Johns Hopkins University Press.
- Lyness, S. A. 1993 "Predictors of Differences between Type A and Type B Individuals in Heart Rate and Blood Pressure Reactivity." *Psychological Bulletin* 114:266–295.
- Maccoby, E. E., and C. N. Jacklin 1974 *The Psychology of Sex Differences*. Stanford, Calif.: Stanford University Press.
- Major, B., L. Barr, J. Zubek, and S. H. Babey 1999 "Gender and Self-Esteem: A Meta-Analysis." In W. B. Swann, J. H. Langlois, and L. A. Gilbert, eds., *Sexism and Stereotypes in Modern Society*. Washington, D.C.: American Psychological Association.
- Masters, M. S., and B. Sanders 1993 "Is the Gender Difference in Mental Rotation Disappearing?" *Behavior Genetics* 23:337–341.
- National Science Foundation 1999 *Women, Minorities, and Persons with Physical Disabilities in Science and Engineering: 1998*. NSF Publication No. 99-87, February 1999.
- Ridgeway, C. L., and D. Diekema 1992 "Are Gender Differences Status Differences?" In C. L. Ridgeway, ed., *Gender, Interaction, and Inequality*. New York: Springer-Verlag.
- Rosenthal, R., and L. Jacobson 1968 *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development*. New York: Holt, Rinehart, & Winston.
- Sharps, M. J., A. L. Walton, and J. L. Price 1993 "Gender and Task in the Determination of Spatial Cognitive Performance." *Psychology of Women Quarterly* 17:71–83.
- Skinner, P. H., and R. L. Shelton 1985 *Speech, Language, and Hearing: Normal Processes and Disorders*, 2nd ed. New York: Wiley.
- Spencer, S. J., C. M. Steele, and D. M. Quinn 1999 "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology* 35:4–28.
- Steele, C. M. 1997 "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance" *American Psychologist* 52:613–629.
- Stumpf, H., and J. C. Stanley 1998 "Stability and Change in Gender-Related Differences on the College Board Advanced Placement and Achievement Tests." *Current Directions in Psychological Science* 7:192–196.
- Sutaria, S. D. 1985 *Specific Learning Disabilities: Nature and Needs*. Springfield, Ill.: Charles C Thomas.
- Thomas, J. R., and K. E. French 1985 "Gender Differences across Age in Motor Performance: A Meta-Analysis." *Psychological Bulletin* 98:260–282.
- Voyer, D., S. Voyer, and M. P. Bryden 1995 "Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables." *Psychological Bulletin* 117:250–270.
- Walters, A. E., A. F. Stuhlmacher, and L. L. Meyer 1998 "Gender and Negotiator Competitiveness: A Meta-Analysis." *Organizational Behavior and Human Decision Processes* 76:1–29.
- West, C., and D. H. Zimmerman 1987 "Doing Gender." *Gender and Society* 1:125–151.
- Whitley, B. E., and M. E. Kite 1995 "Sex Differences in Attitudes toward Homosexuality: A Comment on Oliver and Hyde (1993)." *Psychological Bulletin* 117:146–154.
- Wilson, M., and M. Daly 1985 "Competitiveness, Risk Taking, and Violence: The Young Male Syndrome." *Ethology and Sociobiology* 6:59–73.
- Wood, W., F. Y. Wong, and J. G. Chachere 1991 "Effects of Media Violence on Viewers' Aggression in Unconstrained Social Interaction." *Psychological Bulletin* 109:371–383.

MONICA BIERNAT  
KAY DEAUX

## SEX-ROLE MODELS

See Gender; Femininity/Masculinity; Role Theory; Socialization.

## SEXISM

*See* Feminist Theory; Gender.

## SEXUAL BEHAVIOR IN MARRIAGE AND CLOSE RELATIONSHIPS

The scientific study of sexuality is not limited to a single discipline but involves scholars from a variety of fields, including sociology, psychology, and biology. The multidisciplinary nature of sexuality research lends itself to various and often contradictory theoretical interpretations of sexual phenomena. Furthermore, individuals have various perspectives on the purpose of sexual behavior and engage in a variety of sexual behaviors. For example, Reiss (1960) asserts that individuals take one of three general approaches to sexuality. Some individuals have a "procreational orientation," believing that the purpose of sexual intercourse is reproduction; others have what Reiss (1960) refers to as a "relational orientation," believing that sexual intercourse is for the purpose of expressing emotional attachment to a partner; and still others have a "recreational orientation," viewing the purpose of sexual behavior as enjoyment. Other researchers (e.g., Peplau et al. 1977) have constructed similar typologies. These perspectives or standards are likely to overlap; for example, couples who engage in sexual intercourse in order to reproduce also express their love and affection for each other and may enjoy themselves in the process. However, Reiss (1960) argues that individuals have one of these orientations as a primary explanation for engaging in sexual behavior. This article focuses on sexuality within close relationships. It begins by presenting some of the theoretical perspectives on sexuality and then discusses recent research findings on sexuality in close relationships.

### THEORETICAL FRAMEWORKS

As was stated above, numerous theories have been developed to explain why certain individuals engage in sexual activities with certain other individuals and the context in which sexual behaviors occur. These theories generally fall on either side of an essentialist versus social constructionist dichotomy (DeLamater and Hyde 1998).

**Essentialism.** According to the essentialist perspective, certain behaviors are biologically driven and thus are natural and ubiquitous. An example of essentialist theories would be evolutionary theories such as sociobiology and sexual strategies theory (Buss 1998; DeLamater and Hyde 1998). According to those theories, which often are applied to gain an understanding of mate selection, certain contemporary sexual behaviors have emerged over time in response to environmental conditions. For example, both men and women are thought to strive to maximize their reproductive success (Oliver and Hyde 1993). Men are theorized to want to pass their genes to successive generations; therefore, they wish to impregnate as many women as possible to increase the likelihood that they will have offspring. Women also want to maximize their reproductive success, but they do so by striving to secure a long-term partner who will provide for their children financially (Chodorow 1978). Because men would not be passing along their genetic heritage by caring for another man's child, a woman's sexual exclusivity to one man has been socially desirable. These theories correspond to Freudian theories in psychology ("anatomy is destiny") and structural functionalist theories in sociology, which argue that traditional heterosexual marriage is the only appropriate forum in which sexual behaviors should occur, as the integrity of families and societies can be maintained only by that institution (see Parsons and Bales 1955).

**Social Constructionism.** Social constructionist theories minimize the influence of biological drives in explaining sexuality, instead focusing on how sexuality, like other phenomena, is socially constructed (DeLamater and Hyde 1998). For example, instead of arguing that men are biologically driven to be sexually promiscuous while women are biologically driven to be sexually exclusive, social constructionists argue that men and women engage in specific sexual behaviors as a result of the cultural messages they receive through socialization. Examples of theories within this framework are symbolic interactionism, social learning, social exchange, and conflict theories (Longmore 1998; Oliver and Hyde 1993; Sprecher 1998).

According to symbolic interactionism, individuals create meaning through their interactions with others; that is, they learn their sexual identities by communicating with others. Indeed, individuals learn and follow sexual scripts (Laumann

et al. 1994; Longmore 1998; Oliver and Hyde 1993) that are specific to their culture, gender, race, social class, and so forth. These scripts are fluid; individuals frequently modify them to adapt to their environments. According to network theory (Laumann et al. 1994), sexual relationships are embedded in larger social networks that influence intimate dyads (Sprecher et al. in press). More specifically, network members influence who pairs up with whom, what behaviors (sexual and otherwise) should be encouraged and/or tolerated in an intimate relationship, and whether dissolution should occur.

To explain sexual behavior, social exchange theories postulate that the members of a dyad exchange resources with each other in an attempt to maximize their rewards and minimize their costs (Sprecher 1998). Examples of social exchange theories are equity theory (individuals reward their partners or withhold rewards in an attempt to influence the balance of equity in the relationship; Sprecher 1998) and choice theory (individuals pursue goals on the basis of the resources available to them, such as energy, physical attractiveness, and money; Laumann et al. 1994).

In explaining sexuality, conflict theories emphasize the power struggle inherent in intimate relationships. According to this perspective, men sexually dominate and exploit women in order to achieve their own goals, while women, lacking power, service men's sexual needs (Weis 1998b). While men may be better positioned to dominate and exploit because of their greater average physical strength, the power struggle is social and is based on the system of patriarchy (see Richardson 1996 for a discussion of this issue).

**Summary.** Despite these theoretical perspectives, sexuality research has been accused of being atheoretical (Weis 1998a, 1998b). Weis argues that theories of sexuality are still in an early stage compared to the theoretical development of other scholarly fields because of a lack of sexuality research that tests hypotheses (below, it will be seen that much sexuality research is descriptive) and because few strong connections have been established between theory and research. Christopher and Sprecher (in press) concur; in their decadelong review of sexuality in close relationships, they note the limited progress made in theories of sexuality despite the considerable increase in research in-

terest in this topic. They call for more theoretical advancements in future research on sexuality.

The following sections present some of the research findings concerning sexuality in close relationships. Several large-scale studies have been conducted within the last decade (most notably the National Health and Social Life Survey in addition to the second wave of the National Survey of Families and Households and several waves of the General Social Survey, which include measures tapping sexual behavior) despite the politicized nature of conducting studies on sexual behavior. These studies have greatly increased sociological understanding of sexuality in committed relationships. The discussion here focuses on three main concerns: (1) descriptive information on sexual behavior in close relationships (e.g., frequency of intercourse, number of partners), (2) the character of extramarital/extradynamic sex, and (3) qualities of homosexual close relationships.

#### **DESCRIPTIVE CHARACTERISTICS OF SEXUAL BEHAVIOR**

As was stated above, much sexuality research is descriptive. Thus, there is considerable knowledge of how frequently individuals engage in sexual behaviors, the types of behaviors in which they engage, how many sexual partners individuals have had over the course of their lives, the prevalence and character of extramarital or extradynamic sex, and the extent of sexual and overall life satisfaction, in addition to numerous other behavioral and attitudinal characteristics. This section focuses on the frequency of sex, the number of partners, and sexual and overall life satisfaction.

The state of descriptive knowledge of sexual behavior has increased considerably over the last several years as a result of the implementation of large-scale national surveys that use rigorous sampling methods. In particular, the National Health and Social Life Survey, which is based on a national sample of the noninstitutionalized population between ages 18 and 59, has provided extensive information on the sex lives of U.S. citizens (Laumann et al. 1994). The design of this survey is much more rigorous than that of the famous Kinsey studies (Kinsey et al. 1948, 1953), which relied largely on volunteers, calling into question the extent to which their data represented the general population.

**Frequency of Sex.** Contrary to the popular opinion that married couples engage in sex less than anyone else does, Laumann et al. (1994) found that only 1.3 percent of married men and 3.0 percent of married women did not engage in sex in the past year (respondents defined the meaning of engaging in sex, which may or may not have included vaginal intercourse). Twenty-two percent of the never-married/noncohabiting men and 30.2 percent of the never-married/noncohabiting women did not engage in sex in the past year. These figures are similar to those for divorced, separated, or widowed individuals who were not cohabiting (23.8 percent of the men and 34.3 percent of the women, respectively). However, all the cohabiting men and all the divorced/separated/widowed cohabiting women had engaged in sex at least a few times in the past year (1.4 percent of the never-married cohabiting women reported not engaging in sexual activity in the past year). According to Laumann et al. (1994), then, while cohabitators on average are more likely to engage in sex than are respondents in any other heterosexual category, sexual activity occurs in nearly all marriages.

According to Laumann et al. (1994), cohabitators are also more likely than are other heterosexuals to engage in frequent sex. More specifically, 18.6 percent of never-married cohabiting men (11.1 percent of divorced/separated/widowed cohabiting men) and 16.7 percent of never-married cohabiting women (11.3% of divorced/separated/widowed cohabiting women) engaged in sex at least four times a week, compared with 7.3 percent of married men and 6.6 percent of married women. Blumstein and Schwartz (1983), who conducted face-to-face interviews and collected questionnaire data from a convenience sample of heterosexual married couples, heterosexual cohabiting couples, and gay and lesbian cohabiting couples, found more substantial differences. They report that while 45 percent of couples in short marriages (two years or less) engage in sex at least three times a week or more, the corresponding figure for short-term cohabitators is 61 percent. This pattern persists when one compares marital and cohabiting unions of two to ten years (27 percent of the marrieds and 38 percent of the cohabitators engage in sex at least three times a week) (Blumstein and Schwartz 1983). Similarly, Rao and DeMaris (1995), using National Survey of Families and Households

data, found that the mean monthly frequency of sexual intercourse among cohabitators was approximately 1.3 times the mean frequency of the legally married.

As was noted above, those who are not currently married and are not cohabiting are more likely than anyone else not to have engaged in sex in the last year (Laumann et al. 1994). However, never-married/noncohabiting men and women who do engage in sex are slightly more likely than are married men and women, respectively, to have sex four or more times a week (7.6 percent of the noncohabiting men versus 7.3 percent of the married men and 7.0 percent of the noncohabiting women versus 6.6 percent of the married women). The divorced/separated/widowed respondents who are not cohabiting are less likely than are married couples to engage in sex at least four times a week (4.6 percent of the noncohabiting men and 3.7 percent of the noncohabiting women engage in sex this frequently).

Blumstein and Schwartz (1983) also found that gay men engaged in sexual behaviors more frequently than did anyone else in their sample (67 percent of gay men in short-term relationships of less than two years reported engaging in sex three or more times a week, compared with 32 percent of gay men in unions of two to ten years). Lesbians engaged in sex less frequently than did anyone else in the sample (33 percent of lesbians in relationships of less than two years engaged in sex at least three times a week; only 7 percent of lesbians in unions of two to ten years engaged in sex that frequently). However, Laumann et al. (1994) found no statistically significant difference in the frequency of sex among gay men versus all men. Statistically significant differences also did not appear among the women (however, the sample size of lesbians was very small).

In summary, cohabitators engage in more sex than do other heterosexuals, but marital sex is not the anomaly that popular opinion asserts (in other words, that sex rarely occurs in marriage; see Blumstein and Schwartz [1983] for a discussion of this issue). However, sexual frequency does decline with increasing age and duration of marriage (see Donnelly 1993; Marsiglio and Donnelly 1991; Call et al. 1995). Findings concerning the frequency of sex among homosexual men and women

compared with their heterosexual counterparts are mixed.

**Number of Sexual Partners.** Individuals vary in terms of their number of lifetime (since age 18) sexual partners as a function of gender and relationship status. First, women are much more likely than men to have had only one sex partner since age 18 (31.5 percent of women report one partner compared with 19.5 percent of men) (Laumann et al. 1994). Men are more likely to report having had at least eleven sexual partners since age 18 (32.9 percent of men and 9.2 percent of women reporting having had at least eleven sexual partners).

Married individuals are less likely than are individuals in any other relationship status to have had numerous sex partners since age 18 (Laumann et al. 1994). More specifically, 37.1 percent of married respondents have had only one sex partner since age 18, compared with 24.6 percent of the never-married cohabitators and none of the divorced/widowed/separated cohabitators (who have all had more than one sexual partner). Nearly 15 percent of the never-married/noncohabitators have had only one sex partner (12.3 percent have had no sex partners), and 11.1 percent of the divorced/widowed/separated noncohabitators have had only one sex partner (0.2 percent have had no sex partners). The percentage of respondents who have had two to four sex partners since age 18 is quite similar across all relationship statuses. However, married respondents are consistently less likely than are respondents in all other relationship statuses (with the exception of never-married cohabitators) to have at least five or more sex partners since age 18. Those in other relationship statuses are quite similar in terms of the percentages who have had at least twenty-one sex partners. Thus, while cohabitators are more likely to engage in sexual behaviors and do so more frequently than other heterosexuals do, they do not differ appreciably from heterosexuals in all other nonmarital relationship statuses in the likelihood of having a high number of sexual partners.

**Sexual Satisfaction.** Sexual frequency is associated with overall perceptions of the quality of one's sex life (Blumstein and Schwartz 1983). Among those who engage in sex at least three times a week, the vast majority are satisfied with the quality of their sex lives (89 percent of both husbands and wives, 87 percent of male cohabitators,

88 percent of female cohabitators, 85 percent of gay men, and 95 percent of lesbians report satisfaction). The percentages of respondents who are satisfied with the quality of their sex lives decreases linearly with decreasing sexual frequency in all types of unions. For example, among those engaging in sex once a month or less, only 32 percent of husbands and wives are satisfied with the quality of their sex lives, compared with 4 percent of male cohabitators, 30 percent of female cohabitators, 26 percent of gay men, and 37 percent of lesbians. Other research has confirmed these findings (e.g., Call et al. 1995; Donnelly 1993).

The number of sexual partners a respondent has within the last year also influences physical pleasure and emotional satisfaction (Laumann et al. 1994). Married and cohabiting respondents who have had only one sexual partner in the last year are identical in terms of the the percentage reporting that they are extremely or very physically pleased with their relationships (87.4 percent and 84.4 percent, respectively). However, married and cohabiting respondents with more than one partner in the last year are less physically pleased by their primary partners (61.2 percent and 74.5 percent, respectively). The majority of those who are neither spouses nor cohabitators are also extremely or very physically pleased with their primary relationships regardless of whether they have one or more partners (78.2 percent with one partner are extremely or very physically pleased, compared with 77.9 percent of those with more than one partner).

Spouses with only one sexual partner in the last year are more likely than are respondents in the other union types to report that they are extremely or very emotionally satisfied with their relationships (84.8 percent; 75.6 percent of the cohabitators and 71.0 percent of those who are neither married nor cohabiting report that they are extremely or very emotionally satisfied). However, only 56.7 percent of marrieds with more than one partner report being extremely or very emotionally satisfied with their primary partners (the corresponding figures for cohabitators and those who are neither married nor cohabiting are 57.9 percent and 61.7 percent, respectively). In short, these numbers indicate that most individuals with one partner are quite physically pleased and emotionally satisfied with their relationships. Unfortu-

nately, statistics are not provided comparing homosexual relationships to heterosexual relationships.

**Overall Satisfaction with Life.** As previous researchers have argued (e.g., Cupach and Comstock 1990; Edwards and Booth 1994; Greeley 1991; Lawrance and Byers 1995), sexual satisfaction is associated with overall relationship satisfaction, which is associated with overall satisfaction with life. Men and women are very similar in terms of their self-reports of overall life satisfaction (Laumann et al. 1994). There are, however, differences by marital status, sexual frequency, and number of partners. First, married respondents are more likely than are the never married, the divorced, the widowed, and the separated to state that they are extremely or very happy overall (67.5 percent, compared with 51.9 percent of never-marrieds, the next highest percentage) (Laumann et al. 1994). Similarly, they are less likely than are those in other statuses to report that they are fairly unhappy or unhappy with life most times (8.7 percent versus 15.2 percent of the never-married, the next lowest percentage in the sample). Unfortunately, Laumann et al. (1994) did not compare cohabitators, nonmarried/noncohabitators, or gays and lesbians to the legally married on this measure. However, they report that 62.5 percent of heterosexual men and 59.2 percent of heterosexual women are extremely or very happy, compared with only 47.1 percent of homosexual men and 45.6 percent of lesbians.

Respondents with only one sexual partner in the last year are happier overall (63.4 percent report being extremely or very happy) than are those with no sexual partners in the last year (40.7 percent) and those with more than one sexual partner (44.9 percent of those with two to four sexual partners and 47.2 percent of those with five or more sexual partners) (Laumann et al. 1994). Furthermore, the frequency of sex is associated with overall happiness in a generally linear fashion: The proportion of respondents stating that they are extremely or very happy increases with increasing frequency of sex in the past year unless the respondents engage in sex at least four times a week; the proportion reporting that they are extremely or very happy decreases slightly at this point. Again, Laumann et al. (1994) do not report distinctions among the respondents based on union status (married, cohabiting, homosexual, etc.).

**Summary.** These results clearly indicate that sex in marriage is not boring or non-existent for most couples and that unattached (not married and not cohabiting) individuals are not having all the fun. Indeed, because frequency of sex is associated with satisfaction with one's sex life and with overall feelings of happiness, one would expect married individuals to exhibit high levels of well-being in this regard, since most have steady access to a sexual partner. Also, variety of sexual experience as a function of having numerous partners is not associated with increasing feelings of happiness; instead, exclusive attachment to one sexual partner is associated with high levels of well-being for most respondents. These results raise the question: Why do some individuals engage in extramarital/extradyadic sex?

#### EXTRAMARITAL/EXTRADYADIC SEX

**Behavioral Incidence.** Kinsey et al. (1948, 1953) estimated that approximately half of married men have engaged in extramarital intercourse; the corresponding proportion for married women is approximately one-fourth. More recent figures that are based on more rigorous sampling methods indicate that on average, approximately 25 percent of men and 15 percent of women have experienced extramarital sex (Laumann et al. 1994). In additional analyses conducted by Laumann et al. using 1991 General Social Survey data, 21.7 percent of men and 13.4 percent of women (both between the ages of 18 to 59, consistent with their own sample) reported having extramarital sexual experience. These figures are consistent with those of Wiederman (1997), who, using data from the General Social Survey, found that 22.7 percent of men and 11.6 percent of women have engaged in extramarital sex. In their analysis of the National AIDS Behavioral Survey, Choi et al. (1994) found that 2.9 percent of men in the national sample (4.1 percent of men in the urban sample) and 1.5 percent of women in the national sample (1.0 percent of women in the urban sample) had engaged in extramarital sexual activity within the last twelve months (these percentages would increase with longer periods of exposure to opportunities for extramarital sex). Laumann et al. (1994) found that 3.8 percent of the married respondents had engaged in extramarital sex within the last twelve months (they did not differentiate between men and women among marrieds).

Blumstein and Schwartz (1983) found that 26 percent of husbands and 21 percent of wives had been nonmonogamous in their current relationships. According to those authors, 29 percent of nonmonogamous husbands had only one extramarital partner, while 42 percent reported having two to five extramarital partners. The corresponding figures for wives are 43 percent and 40 percent, respectively. Blumstein and Schwartz (1983) also found that 33 percent of male cohabitators and 30 percent of female cohabitators had engaged in extradyadic sex. Thirty-six percent of nonmonogamous male cohabitators had only one extradyadic partner (49 percent reported having had two to five partners). Forty-four percent of nonmonogamous female cohabitators reported having had only one extradyadic partner (41 percent reported having had two to five partners). Furthermore, 82 percent of gay men have engaged in extradyadic sex; among them, only 7 percent reported having had only one extradyadic partner (43 percent reported having had at least twenty). Blumstein and Schwartz (1983) report that 28 percent of lesbians were nonmonogamous; 53 percent had only one extradyadic partner, while 42 percent had two to five.

Unfortunately, Laumann et al. (1994) did not present the percentages of those in nonmarital/noncohabiting relationships who are sexually nonexclusive. However, Forste and Tanfer (1996) found that 18 percent of dating women between ages 20 and 37 had engaged in a least one instance of extradyadic sex.

As these results indicate, when respondents are asked if they have engaged in extramarital/extradyadic sex, the majority (with the exception of gay men) report sexual exclusivity with their spouses/partners. Also, at young ages (those under age 40), there does not appear to be a gender difference in sexual nonexclusivity among married respondents (Wiederman 1997). Over the age of 40, however, men are more likely to have engaged in extramarital sex. For example, 29.3 percent of men in their forties have engaged in extramarital sex, compared with 19.3 percent of women in their forties (Wiederman 1997). The greatest gender difference was found among respondents in their sixties, among whom 34 percent of men but only 7.6 percent of women reported having engaged in extramarital sex. It appears that the extramarital sexual behavior of women is becoming

more similar to that of men as younger cohorts age.

**Attitudes.** According to Sprecher and McKinney (1993), much research has been conducted on attitudes toward extramarital sex. Generally, the measures employed in these studies assess either normative attitudes (those concerning the acceptability of extramarital sex in general) or personal standards (those assessing the acceptability of extramarital sex for oneself). Most research on normative standards has found that the majority of Americans disapprove of sexual relations with someone other than a person's spouse (Davis 1980; Greeley 1991). For example, Laumann et al. (1994) found that 77.2 percent of their (unweighted) sample believe that extramarital sex is "always wrong." Similarly, using General Social Survey data and restricting the sample to those between the ages of 18 and 59, they found that 74.3 percent believe that extramarital sex is "always wrong." However, other studies measuring personal standards have found that among younger men (under age 40), 70 percent envision themselves as having extramarital affairs at some point (Pietropinto and Simenaur 1977). Furthermore, Atwater (1982) reports that current predictions assert that as young married women age, approximately 50 percent will engage in extramarital sex. Of course, there is considerable difference between actual behavior, attitudes toward behavior for oneself versus others, and predictions concerning the future, as these numbers indicate. It may be expected that statistics on actual behavior are the most accurate in determining the prevalence of extramarital sex.

**Explanations for Extramarital/Extradyadic Sex.** Why do individuals engage in extramarital/extradyadic sex? It appears that the answer to this question depends in part on one's gender and relationship status. Extradyadic sex appears to be part of the culture of male homosexuality (Blumstein and Schwartz 1983). Indeed, it may be part of the socialization experience of most young men, who may enjoy high status among their peers for being sexually nonexclusive, while women are encouraged to be sexually exclusive (Peplau et al. 1977; Rubin 1990). Also, cohabitators in general (both men and women) tend to be much more liberal than are married couples on a variety of measures, such as premarital sex, abortion, and divorce (Blair 1994; Denmark et al. 1985; Macklin 1983a, 1983b). These liberal attitudes may correspond with more

liberal attitudes and behaviors regarding sexuality, including extradyadic sexual behaviors.

Atwater (1982) conducted an in-depth study of married women who engaged in extramarital affairs with single men. She argues that married men and women engage in extramarital affairs for different reasons. Married women who have such affairs report that their self-esteem and confidence increase as a result of an affair; they report feeling more powerful, independent, and resourceful. Married men, in contrast, typically become involved in extramarital affairs in response to unsatisfying marital sex or the belief that their wives (or any one woman, for that matter) could not possibly satisfy all their sexual needs (Meyers and Leggitt 1975; Yablonsky 1979).

In a study of justifications for extramarital sex, Glass and Wright (1992), found that men and women differ in their approval of specific justifications. They found that men are more likely to support sexual justifications for extramarital sex, including engaging in extramarital sex for the purposes of enjoyment, curiosity, excitement, and novelty. Women, in comparison, are more likely to support love justifications (getting love and affection and falling in love) and emotional justifications (intellectual sharing, understanding, companionship, enhanced self-esteem, and respect for extramarital sex). The data collection methods employed in this study were not rigorous (questionnaires were handed out on the street to be completed and mailed back, with a response rate of 36 percent), and only respondents' attitudes were assessed, not actual justifications for their own extramarital sexual experience. However, this research suggests that men and women think differently about extramarital/extradyadic sex (as they do about other forms of sexuality, as evidenced by the sexual double standard; see Sprecher and McKinney 1993). Future research should continue to explore the use of justifications by men and women for engaging in extramarital/extradyadic sex.

**The Character of Extramarital/Extradyadic Relationships.** What are extramarital/extradyadic relationships like? The perspectives of those involved probably differ. Such relationships have been explored in depth only among legally married spouses involved with single individuals. Thus, the discussion here applies only to married cou-

ples, although the findings presented may apply to cohabitators as well.

In an extensive analysis of the relationships between married men and single women, Richardson (1985, 1988) discusses the power play inherent in an extramarital (and, by extension, an extradyadic) relationship. Popular opinion would suggest that the unmarried partner is the one with the power because at any moment she could reveal the affair to her married lover's wife. However, the partner who is married typically holds most of the power. Being married typically requires that the extramarital relationship be maintained in secret and thus in privacy. Often the home of the unmarried partner becomes "their" home because it provides the only safe setting in which the partners can come together. Furthermore, the married partner decides when and how much time the couple will spend together. The man is more likely to have competing obligations (notably a spouse and children), and so the unmarried partner often makes herself constantly available to her married lover whenever he can find time to spend with her. As a result, the single woman often constructs her entire life around her married lover, as she cannot obtain social support for her relationship from other family members and friends (because the relationship is maintained in secret). By being so dependent on her married lover, the single woman empowers him while giving up any control she formerly had.

Also, the unmarried partner has much to lose by revealing the affair to her married lover's wife (Richardson 1985). She may lose the relationship altogether, since the married partner's spouse would in most situations call for an immediate end to the relationship. Also, the betrayed spouse could destroy the reputation of the other woman by accusing her of being a home wrecker, causing emotional distress to the betrayed spouse and her children, causing financial problems, and so on. Furthermore, if the other woman works with her married lover, she may be labeled as "sleeping her way to the top," which could mean career and financial problems. This does not mean, however, that the single woman never reveals to the wife that her husband is having an extramarital affair. Block (1978) reports that some of these women (and men, in addition to the married lovers) intentionally plant evidence of an affair in an attempt to force a marital separation. However, it is not the



case that the single woman could only benefit from such a revelation.

Popular opinion also states that the “wife is always the last to know” about her husband’s extramarital liaisons, but it is more accurate to say that the wife is always the last to acknowledge an affair. Indeed, numerous researchers (e.g., Atwater 1982; Block 1978; Framo et al., 1975; Richardson 1985; Yablonsky 1979) have discussed the lengths to which a monogamous spouse will go to pretend that the affair does not exist, some bordering on the absurd. Atwater (1982) refers to such feigned ignorance of an affair as a “‘pretense’ context” (p. 86). Most wives pretend not to know about the affair because admitting knowledge of it would force them to feel compelled to respond to such a transgression (Framo et al. 1975). These women appear to have been lulled into a sense of complacency. They do not want to believe that their comfortable (although possibly dull) marriages may be in danger. Therefore, feigning ignorance may be a strategy to maintain the marriage (Richardson 1985). This may be especially important to wives who feel they have few alternatives to the marriage. Women in midlife, who have minimal chances of successfully competing with other women to find a long-term heterosexual relationship, and women who have been financially dependent on their husbands for many years may feel they have no choice but to stay with their husbands.

Similarly, a single woman interested in a permanent relationship with a married man (not all of these women are interested in permanent relationships, as that typically would entail performing housework and other services for the man, something their wives do instead; Richardson 1985) often denies the existence of the wife in order to more easily engage in her fantasies of permanence (Richardson 1985; Yablonsky 1979). Thus, feigned ignorance allows both women to indulge in their preferred fictions. The result, of course, is that the married man has the implicit permission of his wife to engage in the extramarital affair, while the single woman does not place undue pressure on him to divorce his wife.

Sometimes the evidence of an extramarital affair cannot be ignored. In these cases, there are several possible outcomes. One may be the immediate end of the extramarital affair (Framo et al.

1975). Another may be the eventual end of the marriage. Even if the marriage ends, however, it is unlikely that the newly divorced husband will marry the other woman even if he leaves his wife for her (Richardson 1985). A third possible outcome is that the married couple will arrive at an understanding of the husband’s infidelity. For example, upon the discovery of the husband’s extramarital liaison, the couple may engage in a tremendous and ugly conflict, followed by the husband promising to never stray again and the couple maintaining their relationship. However, few wives place much faith in a husband’s new claims to fidelity (Ziskin and Ziskin 1973). Other couples construct an arrangement in which husbands are permitted to engage in extramarital sex, presumably because of their greater sexual need (Yablonsky 1979), but wives are expected to remain monogamous (Ziskin and Ziskin 1973). The wives typically accept this arrangement with resignation: They believe that they will not be able to stop their husbands from engaging in extramarital sex, and as long as the husband continues to remain married to the wife, does not allow himself to develop a deep emotional attachment to any of the other women with whom he is involved, and does not bring home a sexually transmitted disease, the stability of the marriage is not threatened (Block 1978; Yablonsky 1979; Ziskin and Ziskin 1973). A few wives may believe that they benefit from this situation, as they are granted the ability to pursue their own interests (rather than responding to the husband’s needs), while the husband has some of his needs met by other women (Moultrup 1990). Finally, some spouses construct a new agreement in which both spouses are permitted to engage in extramarital sex (Myers and Leggitt 1975; Yablonsky 1979).

**Summary.** These results indicate that extramarital/extradyadic sex is not a majority experience. Indeed, with the exception of gay men, most individuals in committed relationships are sexually monogamous. These results also indicate that single women involved in extramarital affairs are not necessarily home wreckers or are looking to “steal away” another woman’s husband. Furthermore, the married men involved in these affairs are not the hapless victims of single women’s feminine wiles, powerless to reject any and all sexual advances. Finally, the monogamous husbands and wives left at home while their spouses rendezvous with single lovers are not necessarily

the fools that others make them out to be but are often making a conscious choice to ignore a spouse's infidelity because of a lack of attractive alternatives to the current relationship.

As this discussion illustrates, little research has been conducted on extradyadic affairs among cohabitators, gay men, and lesbians, with the exception being the collection of statistics on the incidence of such experiences, as was discussed above. While heterosexual cohabitators may experience the same conditions and responses to their own or their partners' infidelity as do the legally married, future research should specifically address the experience of extradyadic sex among them in addition to infidelity among homosexual couples.

As has been suggested throughout this article, the relationships of homosexuals may be both similar to and different from those of heterosexuals in important respects (as an example, recall the very high relative rates of nonmonogamy among gay men, while the rates of nonmonogamy are similar among lesbians and heterosexual wives). Indeed, the relationships of gay men and lesbians are similar to and different from each other in important ways as well. The following section presents the results of research indicating the incidence and prevalence of homosexuality for both men and women as well as how those relationships may be similar to and different from the relationships of heterosexuals.

### HOMOSEXUALITY

**Behavioral Incidence.** In discussing statistics on the prevalence of homosexual experience among men and women, one must be extremely cautious. Even at the end of the twentieth century, homosexuality was still a decidedly stigmatized status. This can be seen in attitudes toward homosexuality: In their analysis of General Social Survey data, Davis and Smith (1987) found that 75 percent of adults in the United States believe that sexual relations between two adults of the same sex are "always wrong." Only 12 percent of adults believe that sexual relations between two same-sex partners are "not wrong at all." This stigma probably results in some individuals claiming no homosexual experience when such experience has occurred or continues to occur for those individuals. Thus, one may suspect that reported statistics on the

incidence and prevalence of homosexuality are inaccurately low.

Furthermore, homosexuality may be measured in a number of ways. For example, homosexuality may be defined as having same-gender sex partners, as expressing homosexual desires, or defining oneself as homosexual, bisexual, heterosexual, and so on (Laumann et al. 1994). Such variability leads Laumann et al. to conclude that there is "unambiguous evidence that no single number can be used to provide an accurate and valid characterization of the incidence and prevalence of homosexuality in the population at large" (1994, p. 301).

Laumann et al. (1994) report different percentages of the incidence and prevalence of homosexuality depending on how homosexual experience is defined. For example, they found that 0.6 percent of men and 0.2 percent of women have exclusive homosexual experience; that is, these men and women have never engaged in sexual activity with a person of the opposite sex. However, when respondents were asked if they had had any same-gender sex partners since the age of 18, 4.9 percent of the men and 4.1 percent of the women reported having had same-gender sex partners. Furthermore, 9.1 percent of men and 4.3 percent of women reported having engaged in at least one sexual practice with a same-gender partner since puberty (these practices include oral and anal sex). Also, 2.0 percent of men and 0.9 percent of women define themselves as homosexual (0.8 percent of men and 0.5 percent of women define themselves as bisexual; the remainder define themselves as heterosexual). In short, as these statistics illustrate, accurately determining the incidence and prevalence of homosexuality is both a political issue and a methodological issue.

#### **The Character of Homosexual Relationships.**

Today most individuals maintain their intimate relationships in ways that differ from the traditional model (i.e., a breadwinning father and a stay-at-home mother). These various methods of maintaining relationships typically are referred to by social scientists and the larger public as "alternatives" rather than as legitimate family relationships (Boswell 1994; Scanzoni et al. 1989). Defining these relationships as alternatives to the standard (the traditional nuclear family) implies that such

relationships are somehow not as legitimate or valid as the standard (Scanzoni et al. 1989).

A closer examination of these relationships, however, indicates that dyads that do not conform to the traditional standard are similar in certain important respects to traditional heterosexual relationships, particularly in the dyads' search for emotional intimacy and permanence (Scanzoni et al. 1989). More specifically, both heterosexual and homosexual individuals often seek potential partners who are similar to themselves in important ways (Murray 1996). For example, in her analysis of the relationships of homosexual dyads, Sherman (1992) interviewed numerous couples who specifically stated that they chose their current partners on the basis of shared behaviors or interests. For example, John and Reid, a couple that had been together for seventeen years, discussed how

*when we first got together, we talked a great deal about life goals and how we felt about relationships. . . . We decided early on that we wouldn't allow anything to come between us. We've really worked at that. We've made sacrifices about where we wanted to live and what we did professionally to help the other out at different points (p. 60).*

However, both men agreed that while their relationship comes before the pursuit of individual interests, "we both sense the importance of being individuals and helping the other toward self-actualization" (p. 60). Thus, this couple struck a balance between pursuing their individual autonomy and maintaining their relationship on the basis of their similar values.

Similarly, in a study of eighty-four lesbian couples, Weber (1998) interviewed numerous couples who emphasized the importance of shared interests in forming and maintaining their relationships. A 46-year-old educator and a 43-year-old registered nurse who had been cohabiting for two years reminisced about how they came together: "We met at a Democratic fund-raiser, so through our work there we gained respect for each other overall and realized that we are intellectual and political equals. We also noted that we are peers on an educational and vocational level" (p. 57).

It appears that similarity in relevant characteristics attracts potential long-term partners to each other regardless of whether the couple consists of

heterosexual partners, gay men, or lesbians. These long-term couples share a strong sense of commitment and the expectation of permanence based on similar values, making their relationships similar.

However, the relationships of both gay men and lesbians differ in some important respects from those of heterosexual couples, mainly because of the lack of social support homosexual couples experience as a result of their stigmatized status. More specifically, legal and religious institutions for the most part do not acknowledge the legitimacy of homosexual relationships. While in some cities homosexual couples may obtain domestic partnership certificates that publicly acknowledge their relationships, they still may not legally marry in the United States (Wisensale and Heckart 1993; Worsnop 1992). Also, while some representatives of mainstream religious organizations are willing to perform commitment ceremonies (see Sherman 1992 for a discussion of couples who engaged in these ceremonies), the official position of Catholicism, Judaism, and mainstream Protestantism is not to acknowledge such relationships in a religious sense.

Homosexuals confront numerous issues that do not affect heterosexuals. For example, homosexuals often must come to terms with a sexual identity that is stigmatized by the society in which they live. They also must decide whether they will "come out" to family members, friends, and coworkers, recognizing that doing so may jeopardize their relationships and employment. Heterosexuals do not "come out" with their sexual identity, as it is assumed that they are straight, and such an identity is encouraged and valued. Also, heterosexuals are rarely concerned that their sexual identity may result in social ostracism. One may expect that such concerns have an impact on the relationships of homosexual couples, something that is not experienced by heterosexual couples.

Homosexual couples often are shunned by family members and heterosexual friends when they reveal their sexual identity and introduce their partners to others. As a result, friends who do support the couple become defined by that couple as family (Nardi 1992). Those friends may be more important to the support of homosexual relationships than are the friends and family members of heterosexual couples, who also enjoy societywide support for their relationships, as was noted above.

Homosexuals are clearly aware of their second-class status in U.S. society. As a 35-year-old Department of Defense worker explained (Weber 1998, p. 50): "I am a valid and vital human being. I am a taxpayer, a property owner, a veteran, a professional, and also happen to be a lesbian. It is one aspect of who I am, yet it is the only aspect by which I am judged." A 36-year-old school counselor shared her experiences with the pressure associated with her socially ascribed second-class status (Weber 1998, pp. 51-52):

*I love my child. I work. I pay taxes. I follow the Ten Commandments. I have friends and relatives that love me and I love them. I would help anyone in their time of need. I live next door to you in a clean house with a manicured yard. I live with my spouse, my child, and two dogs. I'm so normal that I'm boring. Yet, while I call her my spouse, I cannot legally marry the love of my life. While I love and support and care for my child, she could be taken away from me at the drop of a hat. While I have a responsible job at which I am very good, I could get fired without recourse. While I pay taxes, I cannot claim "head of household" or file jointly with my partner. While I follow the Ten Commandments, many churches will not allow me to attend their services.*

*My family and friends are the most important things in my life, yet if many of my friends knew what my family included, they would cease to be my friends. How can this be? Because I am a lesbian, and according to the laws of this country and the moral judgments of most of you, I am a pervert that should not be allowed to exist.*

In discussing both lesbian and gay male relationships together and comparing them to heterosexual relationships, it should not be assumed that homosexual relationships are similar regardless of the gender of the partners (indeed, it should not be assumed that all heterosexual relationships, all lesbian relationships, or all gay male relationships are similar). As was noted above, gay men and lesbians differ significantly with regard to sexual frequency, number of partners, and other characteristics. However, lesbians and gay men suffer from the same stigmatized status as homosexuals. Future research should continue to explore the various coping mechanisms employed by homosexual couples in dealing with this stigma as well as

the ways in which their relationships are affected by it.

**Summary.** As this discussion illustrates, homosexual relationships are similar in a variety of ways to heterosexual relationships, particularly with regard to the degree of commitment in long-term relationships and the expectation for permanence. Although, as was noted earlier, the majority of gay men are sexually nonexclusive, one should not interpret this to mean that among gay men involved in long-term unions, few are committed to their partners as evidenced by sexual nonexclusivity. Indeed, there are heterosexual couples who engage in nonmonogamy yet remain committed to their primary unions (see the above discussion of extramarital sex). Monogamy should not be confused with commitment to one's relationship, and nonmonogamy should not be considered an indication of a lack of commitment.

Homosexual relationships do, however, differ in important respects from heterosexual relationships as a result of the second-class status of their unions. One may suspect that this stigma has an impact on the relationship dynamics of homosexual unions. Future research is needed to understand more fully how societal conditions affect these intimate relationships.

## CONCLUSIONS

This article has demonstrated that even in examining only close, committed relationships, knowledge of sexual behavior is largely descriptive. Researchers need to explore in much greater depth the sexual dynamics of these relationships. More specifically, there is a need for a better understanding of how couples negotiate their sexual behaviors. Who decides how frequently a couple will engage in sexual activity and in what behaviors they will engage in and when? What is the negotiation process? Also, there is virtually no understanding of how the larger society affects the sex lives of individuals and couples, with the exception of theories of gender socialization in explaining differences between men and women. Of course, asking these questions is a political as well as a methodological endeavor that requires a strong financial and philosophical commitment on a societal basis. While the knowledge of sexual behavior in close relationships has progressed considerably the last few years, a greater understanding is

needed of how individuals negotiate and manage their sex lives in the context of committed relationships rather than simply understanding what individuals *do* sexually. (Note: The authors wish to thank Susan Sprecher for her assistance on this article.)

(SEE ALSO: *Courtship; Sexual Behavior Patterns; Sexual Orientation*)

## REFERENCES

- Atwater, L. 1982 *The Extramarital Connection: Sex, Intimacy, and Identity*. New York: Irvington.
- Blair, S. L. 1994 "Marriage and Cohabitation: Distinctions and Similarities across the Division of Household Labor." *Family Perspective* 28:31-52.
- Block, J. D. 1978 *The Other Man, the Other Woman: Understanding and Coping with Extramarital Affairs*. New York: Grosset & Dunlap.
- Boswell, J. 1994 *Same Sex Unions in Premodern Europe*. New York: Morrow.
- Blumstein, P., and P. Schwartz 1983 *American Couples*. New York: Morrow.
- Buss, D. M. 1998 "Sexual Strategies Theory: Historical Origins and Current Status." *Journal of Sex Research* 35:19-31.
- Call, V., S. Sprecher, and P. Schwartz 1995 "The Incidence and Frequency of Marital Sex in a National Sample." *Journal of Marriage and the Family* 57:639-652.
- Chodorow, N. 1978 *The Reproduction of Mothering*. Berkeley: University of California Press.
- Choi, K., J. A. Catania, and M. M. Dolcini 1994 "Extramarital Sex and HIV Risk Behavior among U.S. Adults: Results from the National AIDS Behavioral Survey." *American Journal of Public Health* 84:2003-2007.
- Christopher, F. S., and S. Sprecher In press "Sexuality in Marriage, Dating, and Other Relationships: A Decade Review." *Journal of Marriage and the Family*.
- Cupach, W. R., and J. Comstock 1990 "Satisfaction with Sexual Communication in Marriage: Links to Sexual Satisfaction and Dyadic Adjustment." *Journal of Social and Personal Relationships* 7:179-186.
- Davis, J. A. 1980 *General Social Surveys, 1972-1988. Cumulative Data*. Chicago: National Opinion Research Center.
- , and T. W. Smith 1987 *General Social Surveys, 1972-1988*. Chicago: National Opinion Research Center.
- DeLamater, J. D., and J. S. Hyde 1998 "Essentialism vs. Social Constructionism in the Study of Human Sexuality." *Journal of Sex Research* 35:10-18.
- Denmark, F. L., J. S. Shaw, and S. D. Cialli 1985 "The Relationship among Sex Roles, Living Arrangements, and the Division of Household Responsibilities." *Sex Roles* 12:617-625.
- Donnelly, D. A. 1993 "Sexually Inactive Marriages." *Journal of Sex Research* 30:171-179.
- Edwards, J. N. and A. Booth, 1994 "Sexuality, Marriage, and Well-Being: the Middle Years." In A. S. Rossi, ed., *Sexuality Across the Life Course*. Chicago: University of Chicago Press.
- Forste, R., and K. Tanfer 1996 "Sexual Exclusivity among Dating, Cohabiting and Married Women." *Journal of Marriage and the Family* 58:33-47.
- Framo, J. L., Y. A. Cohen, H. A. Otto, C. Symonds, A. Ellis, L. G. Smith, J. R. Smith, and L. Salzman 1975 "How Does an Affair Affect a Marriage? 7 Views." In L. Gross, ed., *Sexual Issues in Marriage: A Contemporary Perspective*. New York: Spectrum.
- Glass, S. P., and T. L. Wright 1992 "Justifications for Extramarital Relationships: The Association between Attitudes, Behaviors, and Gender." *Journal of Sex Research* 29:361-387.
- Greeley, A. M. 1991 *Faithful Attraction: Discovering Intimacy, Love, and Fidelity in American Marriage*. New York: Doherty.
- Kinsey, A. C., W. B. Pomeroy, and C. E. Martin 1948 *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- , and P. H. Gebhard 1953 *Sexual Behavior in the Human Female*. Philadelphia: Saunders.
- Laumann, E. O., J. H. Gagnon, R. T. Michael, and S. Michaels 1994 *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
- Lawrance, K., and E. S. Byers 1995 "Sexual Satisfaction in Long-Term Heterosexual Relationships: The Interpersonal Exchange Model of Sexual Satisfaction." *Personal Relationships* 2:267-285.
- Longmore, M. A. 1998 "Symbolic Interactionism and the Study of Sexuality." *Journal of Sex Research* 35:44-57.
- Macklin, E. D. 1983a. "Nonmarital Heterosexual Cohabitation." In A. S. Skolnick and J. H. Skolnick, eds., *Family in Transition*, 4th ed. Boston: Little, Brown.
- 1983b "Nonmarital Heterosexual Cohabitation: An Overview." In E. D. Macklin and R. H. Rubin, eds., *Contemporary Families and Alternative Lifestyles*. Beverly Hills, Calif.: Sage.
- Marsiglio, W., and D. Donnelly 1991 "Sexual Relations in Later Life: A National Study of Married Persons." *Journal of Gerontology: Social Sciences* 46:S338-344.

- Meyers, L., and H. Leggitt 1975 "A Positive View of Adultery." In L. Gross, ed., *Sexual Issues in Marriage: A Contemporary Perspective*. New York: Spectrum.
- Moultrup, D. J. 1990 *Husbands, Wives, and Lovers: The Emotional System of the Extramarital Affair*. New York: Guilford Press.
- Murray, S. O. 1996 *American Gay*. Chicago: University of Chicago Press.
- Nardi, P. M. 1992 "That's What Friends Are for: Friends as Family in the Gay and Lesbian Community." In K. Plummer, ed., *Modern Homosexualities: Fragments of Lesbian and Gay Experience*. London: Routledge.
- Oliver, M. B., and J. S. Hyde 1993. "Gender Differences in Sexuality: A Meta-analysis." *Psychological Bulletin* 114:29-51.
- Parsons, T., and R. F. Bales 1955 *Family, Socialization, and Interaction Process*. Glencoe, Ill.: Free Press.
- Peplau, L. A., Z. Rubin, and C. T. Hill 1977 "Sexual Intimacy in Dating Relationships." *Journal of Social Issues* 33:86-109.
- Pietropinto, A., and J. Simenaur 1977 *Beyond the Male Myth*. New York: New American Library.
- Rao, K. V., and A. DeMaris 1995 "Coital Frequency among Married and Cohabiting Couples in the United States." *Journal of Biosocial Science* 27:135-150.
- Reiss, I. L. 1960 *Premarital Sexual Standards in America*. New York: Free Press.
- Richardson, D. 1996 *Theorising Heterosexuality: Telling It Straight*. Buckingham, UK: Open University Press.
- Richardson, L. 1985 *The New Other Woman: Contemporary Single Women in Affairs with Married Men*. New York: Free Press.
- 1988 "Secrecy and Status: The Social Construction of Forbidden Relationships." *American Sociological Review* 53:209-219.
- Rubin, L. B. 1990 *Erotic Wars: What Happened to the Sexual Revolution?* New York: HarperCollins.
- Scanzoni, J., K. Polonko, J. Teachman, and L. Thompson 1989 *The Sexual Bond: Rethinking Families and Close Relationships*. Newbury Park, Calif.: Sage.
- Sherman, S. 1992 *Lesbian and Gay Marriage*. Philadelphia: Temple University Press.
- Sprecher, S. 1998 "Social Exchange Theories and Sexuality." *Journal of Sex Research* 35:32-43.
- , and K. McKinney 1993 *Sexuality*. Newbury Park, Calif.: Sage.
- Sprecher, S., D. Felmlee, T. L. Orbuch, and M. C. Willetts In press "Social Networks and Change in Personal Relationships." In A. Vangelisti, H. Reis, and M. A. Fitzpatrick, eds., *Advances in Personal Relationships*, vol. 2: *Stability and Change in Relationship Behavior*. Cambridge.
- Weber, J. C. 1998 *Lesbian Dyads as Families*, Ph.D. dissertation. University of Florida.
- Weis, D. L. 1998a "Conclusion: The State of Sexual Theory." *Journal of Sex Research* 35:100-114.
- 1998b "The Use of Theory in Sexuality Research." *Journal of Sex Research* 35:1-9.
- Wiederman, M. W. 1997 "Extramarital Sex: Prevalence and Correlates in a National Survey." *Journal of Sex Research* 34:167-174.
- Wisensale, S. K., and K. E. Heckart 1993 "Domestic Partnerships: A Concept Paper and Policy Discussion." *Family Relations* 42:199-204.
- Worsnop, R. L. 1992 "Domestic Partners." *CQ Researcher* 2:762-783.
- Yablonsky, L. 1979 *The Extra-Sex Factor: Why over Half of America's Men Play Around*. New York: Times Books.
- Ziskin, J., and M. Ziskin 1973 *The Extra-Marital Sex Contract*. Los Angeles: Nash.

MARION C. WILLETTS  
JANIS C. WEBER

## SEXUAL BEHAVIOR PATTERNS

In the face of significant political and methodological obstacles, social science researchers have continued to advance the understanding of human sexuality. Clearly, the political climate surrounding sex research has improved since Kinsey and his colleagues conducted their pioneering studies on male and female sexuality in the late 1940s and early 1950s (Kinsey et al. 1948, 1953), but the politics of sex research continue to impede progress in this area. One prominent research team, for example, was forced to abandon its efforts to secure federal funding and turn to private foundations to support a landmark study on human sexuality in the general population (Laumann et al. 1994b).

Despite the efforts of some conservative politicians, policymakers have become more willing to fund research on sexuality and related issues. This funding pattern is documented by the numerous large-scale national surveys that were supported by federal monies during the late 1980s and 1990s. These studies dealt extensively with sex and re-

lated issues (e.g., National AIDS Behavioral Survey, National Surveys of Adolescent Males, National Surveys of Men) or included a small battery of sex-related questions in surveys dealing primarily with other topics (e.g., ADD Health Surveys, National Household Survey of Drug Use, National Surveys of Families and Households, National Survey of Labor Market Experience–Youth, Youth Risk Behavior Survey). Funding agencies such as the Social Science Research Council Sexuality Research Fellowship Program are expanding studies of sexuality by offering crucial support for both quantitative and qualitative research projects about sexuality (Di Mauro 1997). Efforts also are being made to expand international studies of sexual behavior (Parker 1997). Most observers agree that the human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS) epidemic has provided the major impetus for this turnaround.

While a great deal has been learned about sexual behavior in recent years, especially among adolescent and young adult men, much of this research has focused on sexuality within a “social problems” context. This research typically has dealt with issues associated with teenage sexuality and pregnancy, HIV/AIDS, coercive sexuality, or another form of sexuality that has attracted a “deviant” label. As a result, relatively little attention has been devoted to studying the expression of sexuality in noncontroversial, everyday life circumstances.

Sex research took an impressive step forward with the publication of *The Social Organization of Sexuality: Sexual Practices in the United States* (Laumann et al. 1994a) and *Sex in America* (Michael et al. 1994), the less technical version of this study prepared for the general public. These publications were based on data drawn from the National Health and Social Life Survey (NHSL) that included face-to-face interviews and supplemental self-administered questionnaires given to 3,432 respondents aged 18–59 who were selected randomly from the noninstitutionalized population. The NHSL represents the most comprehensive sex survey to date that uses a probability sample of the general noninstitutionalized U.S. population.

Conducting sex surveys or other types of research in this area continues to be difficult because sex researchers, compared with researchers in

most other areas, deal with particularly sensitive and personal topics and therefore face serious difficulties with issues related to response bias, sample representativeness, measurement, and ethics. With these methodological issues in mind (Bancroft 1997; Bentler and Abramson 1980; Jayne 1986; Kelley 1986) and given this article’s space limitations, this review assesses the available research on sexual *behavior* in the United States. The brevity of this article precludes a review of all the literature on individuals’ subjective perceptions of sexuality. Notably, much of the research on gay, lesbian, and bisexual experiences has revolved around identity, etiology, community, and AIDS (Risman and Schwartz 1988). While this entry discusses same-sex sexual behavior patterns, the reader should consult “Sexual Orientation” in this encyclopedia for a summary of literature about gay, lesbian, bisexual, and transsexual/transvestite identities and experiences. Since much of the literature on sexuality subsumes behavior under identity, this discussion uses the terms “heterosexual,” “homosexual,” and “bisexual” as adjectives rather than nouns and uses the terms “gay,” “lesbian,” and “bisexual” to refer to self-identity as reported in research studies (Risman and Schwartz 1988).

This discussion takes into account five basic and interrelated features of sexuality: (1) patterns of behavior, emphasizing gender, race, and sexual orientation, (2) the varied meaning of sexuality at distinctive periods throughout the life course, (3) social control aspects, particularly as they relate to prostitution, (4) the consensual and coercive contexts within which sex occurs, and (5) the relationship between the HIV epidemic and sexual behavior. While this entry reviews primarily social science literature, a number of studies have been published in the popular press, some of which have received considerable attention from the lay population, that are based on self-selected samples of persons who returned magazine surveys (Kelley 1986 provides a review of these works) or volunteered for qualitative interview studies (Rubin 1990).

### SEXUALITY IN THE LIFE COURSE

**Childhood Sexual Behavior.** Given the Western view of children as asexual beings, little social science research has been conducted on childhood sexuality in the United States; most of the

limited research has occurred within northern Europe. Data that are available typically rely on adults' recollections of their childhood experiences or the reports of parents about their children. While this research is methodologically limited, it does offer a glimpse of children's sexual behavior and play.

Anthropological research has shown clearly that while individuals in Western cultures tend to view children as asexual, children are seen as being capable of sexual activity in many nonindustrialized, non-Western countries (Ford and Beach 1951). Parents in non-Western countries sometimes tolerate and even encourage their children to pursue heterosexual behaviors (including intercourse), homosexual behaviors (e.g., fellatio), or both. In some societies mothers masturbate their children to soothe them.

The available research indicates that prepubertal boys are much more likely to experience intense sexual interest and masturbate and to do so at younger ages than their female counterparts. Boys also appear to participate more frequently throughout childhood in both heterosexual and homosexual play activities that have sexual overtones (e.g., "doctor and nurse"). These activities typically include an element of exhibition, exploration, and experimentation. In a unique longitudinal study of children in California, researchers found that about 48 percent of mothers reported that their children had engaged in interactive sex play (77 percent when masturbation was included) before age 6 and that exposure to such sex play was not related to children's long-term adjustment at ages 17 to 18 (Okami et al. 1997).

Other researchers focusing on heterosexual and lesbian women in Brazil, Peru, the Philippines, and the United States found similarities across cultures in regard to memories of childhood sexual behavior. Self-identified lesbians were found to be more sexually active as children and reported earlier contact than did heterosexual women. Lesbians also were more interested in girls than heterosexual women were in boys, even though lesbians reported more early attractions to men than did heterosexual women (Whitam et al. 1998). Another study of gay and bisexual male youth ranging in age from 17 to 23 also depends on memories of childhood to understand early childhood same-sex attractions. The results of this

study show that in most cases, same-sex attractions began in childhood and were given sexual meaning at the onset of puberty (Savin-Williams et al. 1996). Both studies suggest that same-sex attractions are felt in childhood and are acted on in adolescence.

**Adolescent and Young Adult Sexual Behavior.** In contrast to childhood sexuality, an expansive body of literature on adolescent heterosexual behavior has emerged during the past two decades, and much more extensive data are now available on young males (Moore et al. 1995, 1998; Sonenstein et al. 1997). Much of this research has used one of several national data sets to document and examine rates and trends for age at first intercourse and sexual activity patterns, with particular attention given to racial patterns.

The bulk of the evidence suggests that there was a sizable increase in the rate of sexual activity among teenage females in the 1970s and 1980s, although that increase appears to have leveled off in the 1990s. In fact, among 15- to 19-year-old females, while the proportion that had ever had intercourse increased from 29 percent in 1970 to 55 percent in 1990, it declined slightly to 50 percent in 1995. About 77 percent of 19-year-old females have had sex at least once. The main reason teenage females remain virgins is that having sex would violate their religious or moral values. While rates among African-American and white females have converged over the past two decades, black females are still more likely to report being sexually active than are their white counterparts. Whereas 90 percent of black females have had intercourse by age 19, about 75 percent of whites and Hispanics have had coitus.

Sexual activity rates among comparably aged males have increased since the 1970s and have always been higher than those of females. A 1998 study compared three separate national samples of males aged 17-19 living in metropolitan areas in 1979, 1988, and 1995 and found that the percentage of respondents who had ever had heterosexual intercourse shifted from 66 percent, to 76 percent, to 68 percent (Ku et al. 1998). However, these changes were restricted to nonblack males because the rates for blacks increased from 1979 to 1988 and then remained basically constant between 1988 and 1995. A similar curvilinear pattern



for the overall sample (but not for blacks) was observed when researchers examined the proportion of teenage males who had had intercourse in the four-week period before the interview. Meanwhile, the average number of times young males engaged in sexual intercourse during the previous twelve-month period increased from fourteen in 1979 to seventeen in 1988 to twenty-one in 1995. While black males did not experience a change in frequency between 1979 and 1988, they reported a significant increase from 1988 to 1995 (thirteen versus twenty-four acts per year).

Attempts to explain why youth initiate sex have focused primarily on the direct and indirect influence of sociodemographic, social psychological, and biological factors. Most of the research designs, especially those used in studying females, are not ideal for concluding a causal relationship between beliefs, attitudes, or values and sexual behavior because these variables tend to be measured simultaneously. Despite this methodological shortcoming, a number of factors appear to be related positively (when controlling for numerous variables) to the probability of individuals engaging in heterosexual intercourse at a young age: being black, living in a poverty area, having weak religious beliefs, attending a segregated school (for blacks), attending an integrated school (for whites), lower parental education, having a mother who was sexually active at a young age, living in a single-parent household, having more siblings, and having a low level of academic achievement.

In addition to these social variables, a small number of researchers have used cross-sectional and longitudinal designs to study the relationship between adolescent hormones and heterosexual behavior (Udry 1988). While this research has produced mixed results, hormones and biological markers associated with puberty appear to be related to adolescents' sexuality (both attitudes and behavior). Consequently, a growing number of social scientists are advocating the development of biosocial models that take into account the complex interrelationship among the pubertal process, sexual identity development, sexual behavior, and societal norms. Collecting saliva and blood samples has become an acceptable, and increasingly expected, component of national data collection efforts that target adolescent sexuality and fertility.

The discussion of social and biological factors leading toward the initiation of sexual behavior is also present in studies of same-sex attractions, desires, and behaviors. Similar to heterosexual onset of active sexuality, gay, lesbian, and bisexual youth report the development of sexual interest during puberty. Gagnon (1977) explains that young men typically self-identify as gay and have homosexual experiences at an earlier age than do young women. Jay and Young (1979) report awareness of same-sex attraction for young men developing at the median age of 13 or 14 and for young women at age 18. Cohen and Savin-Williams (1996, pp. 120-121) summarize the results of ten different studies of the initiation of sexual behavior (mostly among young men) and find that the average age of reported first experience of homosexual sex is 15 for young men, with young women reporting an average age of 16. Importantly, same-sex sexual behavior is not always correlated with the development of a gay, lesbian, or bisexual identity. Studies show that youth will participate in homosexual activities without later "coming out" as gay, lesbian, or bisexual. In particular, as a result of cultural understandings of gender and sexuality, Chicano men who participate in same-sex anal intercourse do not consider it "gay sex" unless one is in the subordinant (receptor) role (Almaguer 1993; Alonso and Koreck 1993; Carrier 1989). Research in this area supports the supposition that social and cultural factors are important in the initiation of same-sex sexual behavior and its attendant meaning.

Research on adolescent homosexual and bisexual behavior is complicated by the stigma associated with a gay, lesbian, or bisexual identity, and so many youths do not self-report same-sex activities (Savin-Williams et al. 1996). While many self-identified gay, lesbian, and bisexual adolescents report desiring steady, loving same-sex relationships, many maintain heterosexual relationships and fear revealing close same-sex friendship because of the stigma associated with being "gay" (Hetrick and Martin 1987). Thus, studies of sexual behavior among gay, lesbian, and bisexual adolescents have documented heterosexual behavior, although with a lesser frequency than is the case with heterosexual counterparts. Interestingly, lesbian and bisexual female adolescents tend to report more heterosexual experiences than do gay and bisexual male adolescents (Herdt and Boxer

1993). Research on the frequency of sexual activity among gay, lesbian, and bisexual young adults indicates that for some populations, entry into college allows an increase in the frequency of sexual experiences (Evans and D'Augelli 1996).

While most research has focused on heterosexual intercourse, a number of studies of adolescent and young adult populations have examined issues related to the number of sexual partners, frequency of sexual intercourse, other types of sex acts, and the sequencing of petting behaviors (Laumann et al. 1994a; Moore et al. 1995; Warren et al. 1998). The majority of teenagers and those in their twenties tend to have one sexual partner at a time, a form of serial monogamy. One study found that 72 percent of women aged 18–29 reported having only one partner in the previous year, and another found that only about 10 percent of females have two or more partners in a three-month period. Among the sexually experienced, about 77 percent of 15- to 19-year-old females and 85 percent of those aged 20–24 report having sex more than once a month. Among sexually experienced teenage males aged 15–19, 54 percent report having had no more than one partner in the previous year, 80 percent report having had two partners, and 6 percent report having had five or more.

Researchers also have shown that the prevalence of oral sex has grown tremendously in this century. Kinsey's data revealed that very few college women born between 1910 and 1935 performed fellatio (11 percent) or received cunnilingus (12 percent). More recent studies in California and North Carolina suggest that between one-third and one-half of adolescents aged 15–18 have engaged in oral sex (Hass 1979; Newcomer and Udry 1985), while nonrepresentative studies of college students in the United States and Canada indicate that between 32 and 86 percent of females have administered oral sex and between 44 and 68 percent say they have received it (Herold and Way 1983; Young 1980). Furthermore, Kinsey's data suggested that oral sex was primarily experienced only among those who also had experienced coitus (only 5 percent of male and female virgins reported performing it), but more recent research indicates that a sizable minority of youth are experiencing oral sex while they are technically virgins.

**Single-Adult Sexual Behavior.** Because researchers have not clearly and consistently distin-

guished between young adult (18- to 24-year-olds) and adult sexual behavior in presenting their findings, it is difficult to present a clear-cut review of "adult" sexual behavior among single persons. Research has shown that rising divorce rates and postponement of marriage for heterosexual men and women have increased the population of single adults (Blumstein and Schwartz 1983). The inability of gay and lesbian partners to marry legally also contributes to the growing population of "single adults" because some persons in committed relationships are categorized as single. This growing population of single adults and the typical transition from experimentation and dating to long-term, committed relationships bring into focus a significant type of adult sexual behavior. While these demographic patterns are noteworthy, a large proportion of individuals still experience a significant proportion of their adult sexual histories within committed relationships or marriage.

For single adults, it is known that individuals today move through the sequence from first kiss to first intercourse much more quickly than did older cohorts of a similar age. Many of these first sexual experiences with a new partner occur within an arrangement that is perceived in some ways to be a relationship. Although men are more likely to find recreational sex outside a relationship acceptable, both men and women prefer to have sex within an ongoing romantic relationship. Thus, contrary to stereotypical images, most adults have sex infrequently when they are not in a relationship. Indeed, national data based on a sample of 18- to 59-year-olds who report being sexually active in the past year indicate that without controlling for age, 48 percent of men and 54 percent of women who have never been married and are not currently cohabiting have had sex only a few times or not at all in the past year. These figures varied only slightly for men (46 percent) and women (58 percent) who were divorced, separated, or widowed but were not married or cohabiting. The mean monthly frequency of sex (vaginal, oral, or anal) was 5.6 and 5.3 for men and women who never married and were not currently cohabiting and 5.4 and 5.1 for men and women who were divorced, separated, or widowed. Only 8 to 9 percent of men and women who had never married but were currently cohabiting—persons who could be perceived technically as being single—reported similarly low levels of sexual activity: a

few times a year or not at all. Meanwhile, among never married, noncohabiting persons, 29 percent of men and 21 percent of women reported giving oral sex to their partners during the most recent sexual experience. Among the entire sample, regardless of whether the respondents had sex in the past year, 67 percent of men and 70 percent of women reported that they had engaged in active oral sex at some point during their lives.

One of the revelations of Kinsey's early research was the extent to which people had same-sex sexual experiences without identifying themselves as gay, lesbian or bisexual. Since Kinsey's research, this finding has been reproduced in a variety of different contexts. Laumann et al.'s comprehensive sex study (1994a) illustrates how the interaction of behavior, attraction, and sexual identity complicates the measurement of the number of people who could be understood to be gay, lesbian, or bisexual. Ultimately, Laumann et al. (1994a) conclude that there is a core group of people who define themselves as "homosexual or bisexual," have same-gender partners, *and* express homosexual desires. Yet there are also a number of men and women who have adult same-sex sexual experiences or desires but do not identify themselves as "homosexual or bisexual." Additionally, variables (e.g., place of residence and level of education) are found to influence the number of homosexual experiences. Because of the complexity of measurement, Laumann et al. (1994a) conclude that there is no single answer to the prevalence of homosexuality. Studies of adult sexual experiences among men of color in the United States and men of color in other cultures also have illustrated the extent to which heterosexual adults engage in same-sex sexual behavior: Research has identified a large population of men who have sex with men (MSM) yet do not self-identify as gay or bisexual (Manalansan 1996). Research since Kinsey's original attempt to develop a comprehensive (although nonrepresentative) understanding of sexual behavior in the United States has shown that some adults participate in homosexual behavior without self-identifying as such.

Despite the complexity of sexual identity and behavior, a great deal of research on same-sex adult behavior patterns has been comparative, creating a binary heterosexual-homosexual comparison. This dualistic approach is so prevalent that many sexual activities are understood as being

in the domain of only heterosexual behavior or only homosexual behavior. For example, tranvestitism (cross-dressing) has been defined in the psychological literature as primarily the domain of heterosexual men, yet more recent research suggests that men with cross-dressing habits behave as homosexual, bisexual, and asexual as well as heterosexual, suggesting a complexity to adult sexual behavior patterns that is not fully understood (Bullough and Bullough 1997). In a significant study comparing same-sex sexual behavior patterns with opposite-sex behavior patterns, Masters and Johnson (1979) found some general trends of similarity for adult "heterosexuals" and "homosexuals" (both men and women). Overall, their research shows that "heterosexuals" and "homosexuals" have similar fantasy patterns and physiological responses to sexual stimuli. In other words, response to sexual stimuli is not conditioned by a particular sexual identity. Some differences between heterosexuals and nonheterosexuals emerge when women's experiences are highlighted. A study of 70 self-categorized heterosexual, bisexual, or homosexual women showed that bisexual and lesbian women were significantly more likely than were heterosexual women to describe their orgasms as "strong." In addition, bisexual and lesbian women put more emphasis on oral and manual sexuality, while heterosexual women put more emphasis on intercourse as a source of sexual response (Bressler and Lavender 1986). Overall, comparative studies suggest a similarity of sexual response and fantasy for both heterosexual and homosexual behaviors but some distinctions in terms of actual sexual activities participated in by self-identified gay, lesbian, or bisexual individuals.

Another arena of research that compares heterosexual behavior to gay and lesbian behavior is the study of committed relationships. A great deal of research illustrates the prevalence of stereotypes about gay, lesbian, and bisexual people being promiscuous and not involved in long-term, committed relationships. However, in one of the few large-scale studies of relationships (including heterosexual, gay, and lesbian couples), Blumstein and Schwartz (1983) found that many lesbians and gay men establish lifelong partnerships. Research also shows that in established couples, gay men and lesbians report as much satisfaction (measured in numerous ways) as heterosexuals do (Peplau 1982). Specific research about sexual be-

havior among gay and lesbian couples has shown some interesting trends. A notable finding in Blumstein and Schwartz's (1983) research is that the frequency of sex in lesbian couples is low compared to that in other couples. Among lesbian couples who had been together two to ten years, more than 25 percent reported having sex once a month or less. This finding has been explained in terms of women's sexual socialization, different definitions of sexuality, and age of couples (Blumstein and Schwartz 1983; Johnson 1990). By contrast, Blumstein and Schwarz (1983) report that gay male couples report higher sexual frequency than do married couples (of one to ten years). This often is explained in terms of men's socialization to value sexuality. Research on sexual frequency in committed couples illustrates the significance of gender and thus the diversity of relationships and attendant sexual behavior among self-identified gay men and lesbians.

Research about the diversity of sexual behaviors among gay, lesbian, and bisexual adults expanded in the 1990s. In general, gender is understood to account for variation of sexual behavior in the gay, lesbian, and bisexual population: Men having sex with men are understood to do so differently (in terms of frequency and activities) than women having sex with women. Variation also has been found within sexual identity categories. For example, it has been found that individuals involved in bisexual behavior have a variety of relationship types (Blumstein and Schwartz 1976). Meanwhile, distinctions in regard to adult same-sex sexual behavior in terms of race are not well understood. In a national-level study of more than 700 coupled, homosexually active African-American men and women, most of the respondents reported satisfying sex lives with their current partners and variation of sexual frequency among couples but no systematic difference between women and men (Peplau et al. 1997). The varieties of sexual behavior within gay, lesbian, and bisexual categories and couples are just being discovered.

The politics of sexuality research in the era of HIV/AIDS has expanded the understanding of sexual behavior among adults, but only in particular arenas. For example, national-level research on sexual behaviors among bisexual men and women and lesbians remains limited (Doll 1997). In addition, the way in which studies of male homosexuality have been conducted differentially operationalizes

"homosexual" as an identity, so that populations of men who have sex with men (MSM) sometimes are included and at other times are excluded (Carballo-Diequez 1997; Sandfort 1997). Clearly, studies of sexuality expanded dramatically in the 1990s, yet this growth pattern has been shaped by concerns associated with an era of HIV/AIDS.

**Sexual Behavior among Elderly Persons.** Although the proportion of the U.S. population over age 50 continues to grow, research addressing the relationship between aging and sexuality, in particular the sexual behavior of the elderly population (65 and older) is quite limited because of its frequent use of small nonrepresentative samples, its cross-sectional research designs, and its narrow, youth-oriented definition of sexuality as coitus. Consequently, generalizations are difficult to make, and most research on elderly persons' sexual behavior deals with the physiological and psychological aspects of this phenomenon. The Viagra revolution undoubtedly will prompt researchers to devote more attention to older persons' sexuality (Butler 1998).

Two of the more frequently cited, though perhaps dated, studies of aging and sexuality issues are the Starr-Weiner Report (1981) and the second Duke Longitudinal Study (George and Weiler 1981), neither of which was based on random sampling techniques. The former study included 800 sociodemographically diverse participants in senior centers, while the latter included a panel design of men and women health insurance program participants who were 46 to 71 years of age at the first observation period in 1969 and were followed for six years ( $n = 348$  for those enrolled in all four data collection points). Seventy-five percent of the respondents in the first study reported that sex felt as good as or better than it did when they were younger. The results from an analysis restricted to the 278 married respondents who had been retained throughout the Duke study revealed that patterns of sexual interest and activity remained fairly stable over time, men reported higher levels of sexual interest and activity than did their female age peers, and younger cohorts of respondents reported higher levels of sexual interest and activity.

Compared to earlier studies (e.g., the Kinsey reports), the Starr-Weider and Duke studies as well as more recent ones have found higher levels

of sexual activity among older persons. Accordingly, Riporetelia-Muller (1989, p. 214) concluded that "for those elderly who remain sexually active and have a regular partner, the rate of decline is not as great as formerly believed." However, many older persons do not remain sexually active. Using data from the second Duke study, George and Weiler (1981) found that among those who were at least 56 years of age at the first observation date, 21 percent of men and 39 percent of women reported six years later that they had abstained from sexual relations throughout the study or were currently inactive.

In a study using data from a nationally representative household sample (Marsiglio and Donnelly 1991), about 53 percent of all married persons 60 years of age and older reported having sex in the past month, with 65 percent of those 60 to 65 years old being sexually active compared to 44 percent of those 66 or older. Among those who had been sexually active during the past month, the overall mean frequency for sexual relations was 4.3 times. In a multivariate context, persons were most likely to have had sex during the past month if they were younger, had a higher sense of self-worth and competency, and were married to a spouse who self-reported his or her health status as favorable. Surprisingly, when an interaction term was used to compare husbands and wives, no significant differences were observed in the way the partner's health status was related to sexual behavior, although other research has found that both husbands and wives report that males' attitude or physical condition tends to be the principal reason why they have curtailed or ceased to have sexual relations. Meanwhile, other research indicates that being widowed is the most frequently cited reason for not being sexually active among older women overall. While data on persons who are institutionalized are scarce, it appears that their sexual activity levels are low.

Being without a spouse does not necessarily mean that older persons are sexually inactive. Brecher (1984) reported, for example, that among unmarried persons 60 years of age and older, about 75 percent of men and 50 percent of women were sexually active. This is reinforced by Starr and Weiner's (1981) finding that 70 percent of their respondents over 60 were sexually active, although only 47 percent were married. Finally, masturbation is an option used by some elderly

persons to express their sexuality, presumably in a nonsocial setting. In a few studies, about one-third of women and slightly less than one-half of men over 70 report masturbating.

Just as stereotypes about elderly people and sexuality abound, so do stereotypes about the sexuality of gay, lesbian, and bisexual elders. However, research suggests that gays and lesbians offer a model of successful aging (Berger 1996). In terms of actual sexual behaviors, many of the gay men interviewed for this study reported that their frequency of sex had changed as they aged, supporting the idea that age is associated with a lower frequency of sexual relations (Berger 1996). Similarly, another study based on self-report data obtained from men between 40 and 77 years old found that the majority of these self-identified gay men were currently sexually active, although sexual interest declined somewhat with age. The majority reported no change in their enjoyment of sex from their younger years to the present (Pope and Schulz 1990). The myth of sexual and emotional isolation for lesbians is also challenged by a series of in-depth interviews with 20 women over age 50 who self-identified as lesbian. This research found that lesbians continue to be sexually active and tend to seek out other older women as partners (Raphael and Robinson 1992). Like other research on aging and sexuality, research on self-identified gay and lesbian elders shows that they have continued interest in expressing themselves sexually.

### CONSENT AND COERCION

In recent decades, researchers have increasingly focused on "rape and other forms of sexual coercion" (Muehlenhard 1994, p. 143). A few observers have reacted by asserting that the data assessing sexual coerciveness among adults who are known to one another are seriously flawed and thus exaggerate their substantive and policy significance (Roiphe 1993; Muehlenhard et al. 1994). However, most social scientists and clinicians consider sexual coerciveness to be both serious and significant and have begun to scrutinize the conceptual and measurement problems that underlie its use.

Perhaps the most significant historical distinction in this realm was between stranger versus nonstranger, or unknown versus known. Rape

occurred most convincingly—to the minds of jurors, for example—when a man who was totally unknown to a woman executed penile penetration. If the man had been known to the woman in any way, her charge of forcible sex could be undermined, depending on how well she knew the man. If, for example, the man was her husband, she could not claim rape at all because consenting to be married carried the obligation to give her body to him unreservedly.

Koss and her colleagues were among the first to expand the discourse on sexual coercion beyond its conventional historical understanding. Using a national survey of college students, they reported that 38 percent of female college students reported sexual victimization that met the legal criterion for rape or attempted rape and almost 8 percent of males admitted to behaviors that could be classified as raped or attempted rape with at least one woman since the fourteenth birthday (Koss et al. 1985, 1988). Their conclusions were supported by other studies (Mosher and Anderson 1986; Rapaport and Burkhart 1984). Using data from the 1987 wave of the National Survey of Children, Moore et al. (1989) found that about 7 percent of U.S. adults aged 18–22 (females being more likely than males) confirmed that they had sex against their will or had been raped on at least one occasion. Other researchers have reported that between 10 and 12 percent of women report having been raped by dates and between 8 and 14 percent of wives report that their husbands have sexually assaulted or raped them (Finkelhor and Yilo 1985; Russell and Howell 1982).

Estimating the prevalence of the various forms of coercive sexuality is exceedingly difficult because the Uniform Crime Reports are widely believed to underreport the true rate of sex crimes and anonymous, self-report surveys vary widely in sampling techniques and findings. Nevertheless, Grauerholz and Solomon's (1989) review of research in this area suggests that a large proportion of the U.S. population has experienced or will experience coercive sexual relations as a victim, a perpetrator, or both. A comparison of sexual coercion among university students in the United States and Sweden found higher rates in the United States (Lottes and Weinberg 1996). Researchers' estimates of the pervasiveness of incest in the United States may vary the most. Whereas some researchers have observed that about 1 percent of

U.S. females have been incest victims (Kempe and Kempe 1984), others have reported much higher figures. For example, Russell (1984) found that 16 percent of her large household sample of women 18 years of age and older in San Francisco had experienced incest before age 18, and 20 percent of Finkelhor's (1979) predominantly white, middle-class New England college student sample who were raised primarily in nonmetropolitan areas reported having been an incest victim. The 1998 National Violence Against Women survey of 8,000 women estimated that some 17.7 million women in the United States, nearly 18 percent, have been raped or have been the victim of attempted rape. Nearly half the victims were assaulted before their seventeenth birthday. Some three-quarters of those saying they had been raped or assaulted as adults reported that the perpetrator was a current or former husband, a cohabiting partner, or a date.

Not only is sexual coercion an issue for many women, there is evidence to suggest that it is pertinent for some boys and men as well (Finkelhor 1979; Moore et al. 1989). A study of 115 sexually assaulted men found that the majority were assaulted at age 15 or younger, about half were assaulted more than once, and a majority knew their assailants. Among the 115 surveyed, 100 were assaulted by at least one man, 7 by a man and a woman, and 8 by women (King and Wollett 1997). Further, much of the research about sexual assault on men assumes that it is primarily heterosexual men who are the perpetrators of sexual assault on other men, doing so to express power and control. However, in a study of 930 "homosexually-active males" in England and Wales, about 28 percent reported having been sexually assaulted or having had sex against their will at some point in their lives. Moreover, 33 percent of the respondents reported being forced into sexual activity by men with whom they had previously had or were currently having consensual sexual activity. This research suggests that sexual assault in the gay, lesbian, and bisexual community needs to be researched further (Hickson et al. 1994). Some studies (Struckman-Johnson and Struckman-Johnson 1994) indicate that some men report that they were pressured or coerced by women into a sexual experience. Other studies focus on men both in and outside of prison who coerce and/or physically force other men into having sex (King and Wollett 1997; Struckman-Johnson et al. 1996).

In an effort to clarify matters of coercion and consent in the case of both sexes, O'Sullivan and Allgeier (1998) proposed a series of points along a continuum. First, there is sexual activity among persons known to each other that is the result of actual or threatened restraint, aggression, or force. It can be done by men against women and by men against men. Logically, it could involve women against women, but empirically that appears to be quite rare even in a prison situation (Struckman-Johnson et al. 1996), although Renzetti's (1992) research on partner abuse among lesbian couples suggests that lesbian couples experience a cycle of violence that can include sexual assault. Nevertheless, the general lack of empirical evidence about coercive sex between women could indicate either lower rates of occurrence or gaps in the research. Coercive sex also could be initiated by women against men, but again, this seems to be rare and is associated with the male having become intoxicated (Struckman-Johnson and Struckman-Johnson 1994). O'Sullivan and Allgeier (1998, p. 234) summarize the first category as "coercive interactions in which a person [submits] to . . . sexual activity but does so under duress."

Their second category of sexual activity occurs as the result of "seduction interactions in which one partner is clearly resistant," [and a third is] "'token resistance' interactions in which a person expresses nonconsent but is willing and intends to engage in the sexual activity." Their final category is "willing participation in an unwanted sexual activity." This refers to "situations in which a person freely consents to sexual activity . . . without experiencing a concomitant desire for the initiated sexual activity." They review literature suggesting the validity of these four categories and present their own research in support of the last one. Future research, they imply, should utilize these and/or other constructs to sort out differences along the consent-coercion continuum.

With evidence that coercive sexuality is prevalent, the question remains: Why do many men and some women coerce others to have sex with them? Some theorists argue that males are more likely to engage in various forms of coercive sexuality if they have strong ties to a peer group that supports sexually aggressive behavior. Compared to their less stereotypically masculine counterparts, men studied in the research noted above who possess traditionally masculine personality characteristics

and hold rigid views of gender stereotypes are more likely to report that they have used physical force and threats to have sex and probably would use physical force to obtain sex if they could be assured that they would not be prosecuted. Other theorists have argued that a small percentage of women may facilitate different forms of coercive sexuality by playing sexually receptive or seductive roles. Not surprisingly, it is common for sexual assailants to believe, or at least report, that their victims were willing participants who enjoyed themselves while being sexually assaulted, even though these perceptions are clearly inconsistent with the victims' accounts (Scully and Marolla 1984).

While the several forms of coercive sexuality share a number of themes, such as the objectification of women, individual factors and circumstances may be significant in accounting for why particular types of coercive sexuality occur. One of the important factors that seem to distinguish the typical stranger or acquaintance rapist from the "average" date rapist is the former's greater tendency to have been sexually or physically abused by his parents or others. Date rapes tend to involve partners who knew one another and had established at least a modicum of interpersonal trust by making a commitment to spend time together. The dynamic nature of the interaction episodes that typify date rapes and the fact that at least one of the persons often has been influenced by drugs, alcohol, or both can obscure the participants' intentions and behavior. Furthermore, the more individuals' sexual scripting is influenced by traditional gender socialization, the more likely it is that coercive sexuality will occur because of sexual miscommunication and males' reliance on coping strategies that emphasize dominance and aggression. Many men assume that women will offer token resistance to their sexual advances to create an impression that they are not sexually "promiscuous" (Check and Malamuth 1985). One study of 610 female undergraduates revealed that almost 40 percent had engaged in this type of token resistance at least once (Muehlenhard and Hollabaugh 1988). While these patterns should not be used to justify date rape, it is not surprising that some men distort the consensual petting that generally precedes date rape as a woman's way of acknowledging her willingness to engage in more intimate forms of sexual interaction, even if this means that in some cases men will be required to pursue it

forcefully. Finally, while there are many factors related to fathers' incestuous behavior, one of the more frequently noted arguments underscores the common pattern in which a father pursues sexual and emotional intimacy with his female children (usually a series of episodes over time with the oldest female child) to compensate for his unfulfilling relationship with the adult female, who generally has withdrawn from her roles as mother and wife (partner).

### **SEXUAL RISK-TAKING BEHAVIOR**

Despite the well-documented overall increase in condom use in the 1990s, most research on risk-taking behavior indicates that a high percentage of people still regularly place themselves at risk of getting pregnant and/or contracting a sexually transmitted disease (STD) and/or HIV/AIDS. These patterns persist despite increased efforts at safe-sex education programs and services. Some trends are becoming increasingly evident. Data from face-to-face interviews with adult men and women aged 18–59 in the 1996 National Household Survey on Drug Abuse showed that 19 percent of the respondents used a condom during their last sexual experience (vaginal, oral, or anal sex, with no distinction being made in terms of the gender of the partner) when the event occurred within a relationship, while 62 percent did so when sex occurred outside an ongoing relationship (Anderson et al. 1999). Ninety-five percent of the respondents indicated that their last sexual experience occurred within a relationship. These data also showed that once relationship status was statistically controlled, individuals with high-risk sex and drug use profiles did not use condoms at a higher rate than did their counterparts who were not classified as being at an increased risk for HIV infection. Only about 22 percent of high-risk individuals used a condom at last intercourse within an ongoing relationship. Thus, both the casual and steady partners of high-risk individuals continue to place themselves at risk.

Meanwhile, a study of college students highlights gender differences in which men are found to be engaged in more risk-taking behaviors relevant to partner choice and sexual practices (Poppen 1995). Another study indicates that college women tend to put themselves at risk for sexually transmitted diseases by rarely using condoms and only

minimally discussing sexual history with a partner (Sheahan et al. 1994). Further, differences in behavior based on sexual identity are empirically supported. Interview data from gay and bisexual men suggest different behavior patterns in regard to "safe sex." These data suggest that gay men were more likely to have had a steady male partner and to have engaged in unprotected anal sex than were bisexual men (Stokes et al. 1997). Other studies indicate that young gay men's reports of having unprotected anal intercourse were more common when they knew their partners. Also, increased involvement in the gay community was related to higher levels of risk-taking behavior (Meyer and Dean 1995). Bisexual men use condoms inconsistently with male and female partners, seldom disclose their bisexuality to female partners, and are more likely than men who participate exclusively in same-sex sexual relations to report risk behaviors associated with HIV. The four factors found to raise HIV risk for bisexual men were (1) male prostitution, (2) injection drug use, (3) sexual identity exploration, and (4) culturally specific gender roles and norms such as those characterizing some African-American and Hispanic communities (Doll and Beeker 1996).

### **PROSTITUTION**

Sociological research on prostitution emerged from studies of crime and delinquency, with an emphasis on theories of innate criminal drives. Contemporary sociological research on prostitution more often focuses on actual behaviors. In fact, the commercial sex industry has become a renewed area of research interest and funding as it has been identified as a site of potentially high transmission of HIV. Most research on prostitution still typically relies on self-report data with small-scale, situationally specific convenience samples. Thus, Kinsey's nonrepresentative data, which are more than forty-five years old, remain the best available comprehensive data on commercial sexual behavior. The available data on commercial sexual behavior identifies a variety of types: heterosexual prostitution of women (street prostitutes, brothels, etc.), heterosexual prostitution of men (e.g., escort services), homosexual prostitution of men (hustlers), homosexual prostitution of women, transvestite/transsexual prostitution, and the global/tourist sex industry (Brock and Thistlethwaite 1996; Perkins and Bennett 1985). Among these



types of prostitution, some trends are increasingly visible.

The most historically well known form of prostitution is the heterosexual prostitution of women. Feminists offer a variety of interpretations for why women enter into prostitution, ranging from exploitation to legitimate work (Jenness 1990). Gagnon (1977) suggests that the six most common reasons men visit prostitutes are (1) sex without negotiation, (2) sexual involvement without commitment, (3) sex for eroticism and variety, (4) a form of socializing, (5) sex away from home, and (6) sex without rejection. Some sociologists believe that prostitution serves a useful societal function by providing men with a convenient sexual outlet, which in turn minimizes the numbers of sexual transgressions against "respectable" women. Other sociologists view female prostitution as an extreme form of sexism.

Rather than addressing this debate, let us note some empirical trends. Kinsey reported that before World War II, between 60 and 70 percent of adult men had visited a prostitute and about 15 to 20 percent used them regularly. However, many commentators have indicated that rates of prostitution, a form of commercialized sex that in the past provided men with sexual opportunities in a less sexually open society, have decreased drastically since World War II. Notably, current research on sexual behavior in the context of prostitution shows the preponderance of incidents of sexual violence and the increased use of condoms. For example, Miller's (1993) study of 16 women incarcerated for prostitution found that almost all had experienced some form of sexual assault and other violent crimes. Benson and Mathews (1995) surveyed vice squads, women working as prostitutes on the streets and in brothels, and clients and resident groups in England, finding that most street prostitutes began working in their teens and have been victims of repeated sexual and physical attacks, while the majority of clients have regular partners or are married and tend to be middle-class and middle-aged. A study in the Netherlands of 559 male clients of female prostitutes found that 14 percent of the clients do not use condoms (De Graaf et al. 1997). Typically, men who are less educated, have more commercial sex contacts, and have more contacts with prostitutes are the least likely to use condoms. It is not known if similar trends are characteristic of the heterosex-

ual prostitution of men or the homosexual prostitution of women because there is little empirical research. Thus, it is difficult to determine whether similar patterns of sexual violence and safe sex are present across gender and sexuality lines.

The concern about the transmission of HIV/AIDS has, however, stimulated new and more comprehensive research about the homosexual prostitution of men and transvestite/transsexual prostitution. For example, Browne and Minichiello (1996) illustrate that many early studies of homosexual prostitution of men (from the 1960s to the 1980s) focused on biopsychological concerns, creating the need for future studies to focus on the behaviors and attendant risk for HIV infection of both the prostitute and the client. A study of 211 male street prostitutes and 15 male customers further makes this point (Morse et al. 1992). Morse et al. (1992) found that despite knowledge of HIV infection and its transmission, customers engage in high-risk sexual and drug use behaviors with prostitutes. However, another study contradicts these findings. Waldorf and Lauderback (1992) interviewed 552 men who solicit clients in public places (hustlers) and men who solicit by telephone and advertisements (call men) and found that 90 percent of the respondents used condoms in the past year and 75 percent used them in the past week. Similar studies of risk behavior are emerging for the transvestite/transsexual prostitution population. For example, Boles and Elifson (1994) interviewed 52 transvestite prostitutes and found that those who are socially isolated tend to be at more risk of HIV infection and those networked with nontransvestite male prostitutes tend to have a lower risk of HIV infection. Sociological research on men involved in homosexual prostitution and transvestites involved in prostitution suggests that sexual behaviors have been altered in an era of HIV/AIDS. However, little research about sexual behavior patterns outside the context of risk behaviors has been reported.

Another area of research in prostitution concerns the growing global sex industry. Sex tourism has long existed, but increasing globalization of travel and Internet advertising have sparked research in this area. Sociologists often explain the global sex industry as an outgrowth of "third world" poverty and the inequity of a global economy. Sex tourism does not involve only heterosex-

ual prostitution of women; homosexual prostitution of men is also prevalent. Empirical studies of sexual behavior in the global sex industry remain minimal as it is illegal in many countries and most research is still done on European or American prostitution (Brock and Thistlewaite 1996).

Sociological studies of the varieties of prostitution in the United States and around the world point to a more general trend of contemporary empirical research focusing on sexual risk behaviors. Many of these studies indicate that the study of sexuality is complicated by both politics and the complexities of research methodology. However, the emergence of HIV/AIDS also reminds the sociological community of the importance of studies of sexual behavior and patterns.

### CONCLUSION

This review has documented the extensive efforts of social scientists in recent decades to enhance the understanding of human sexual behavior. Indeed knowledge of sexual behavior has increased dramatically since the days of Kinsey's early studies. This review also serves as a reminder, however, that knowledge about human sexual behavior, in some areas more than in others, remains limited because sex research is a function of both political and ethical decisions. For example, knowledge about homosexual experiences will continue to be less reliable and complete than it is for heterosexual behavior because the political commitment and considerable resources needed to secure nationally representative samples of gays, lesbians, and bisexuals are missing. Thus, just as social scientists can be held accountable for any shortcomings associated with the prevailing theoretical approaches to sexuality issues, the larger society and its institutionalized, politicized mechanisms for providing research support are responsible for impeding the research community's efforts to understand sexual behavior. Because sexual activity tends to be a highly private social experience, social scientists' incremental advances in documenting and explaining it are linked to the lay population's commitment to this type of research endeavor.

(SEE ALSO: *Courtship, Sexual Behavior, and Marriage; Sexual Orientation; Sexual Violence and Abuse*)

### REFERENCES

- Almaguer, Tomas 1993 "Chicano Men: A Cartography of Homosexual Identity and Behavior." In Henry Abelove, Michele Aina Barale, and David M. Halperin, eds., *The Lesbian and Gay Studies Reader*. New York: Routledge.
- Alonso, Ana Maria, and Maria Teresa Koreck 1993 "Silences: 'Hispanics,' AIDS, and Sexual Practices." In Henry Abelove, Michele Aina Barale, and David M. Halperin, eds., *The Lesbian and Gay Studies Reader*. New York: Routledge.
- Anderson, John E., Ronald Wilson, Lynda Doll, Stephen Jones, and Peggy Barker 1999 "Condom Use and HIV Risk Behaviors among U.S. Adults: Data from a National Survey." *Family Planning Perspectives* 31:24-28.
- Bancroft, John 1997 *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- Benson, Catherine, and Roger Matthews 1995 "Street Prostitution: Ten Facts in Search of a Policy." *International Journal of the Sociology of Law* 23(4):385-415.
- Bentler, P. M., and P. R. Abramson 1980 "Methodological Issues in Sex Research: An Overview." In R. Green and J. Weiner, eds., *Methodological Issues in Sex Research*. Rockville, Md.: National Institute of Mental Health. U.S. Department of Health and Human Services.
- Berger, Raymond M. 1996 *Gay and Grey: The Older Homosexual Man*. New York: Haworth Press.
- Blumstein, Philip W., and Pepper Schwartz 1976 "Bisexuality in Women." *Archives of Sexual Behavior* 5(2):171-181.
- 1983 *American Couples: Money, Work, Sex*. New York: Morrow.
- Boles, Jacqueline, and Kirk W. Elifson 1994 "The Social Organization of Transvestite Prostitution and AIDS." *Social Science and Medicine* 39(1):85-93.
- Brecher, Edward 1984 *Love, Sex, and Aging: A Consumer Union Report*. Mount Vernon, N.Y.: Consumers Union.
- Bressler, Lauren C., and Abraham D. Lavender 1986 "Sexual Fulfillment of Heterosexual, Bisexual, and Homosexual Women." *Journal of Homosexuality* 12(3-4):109-122.
- Brock, Rita Nakashima, and Susan Brooks Thistlewaite 1996 *Casting Stones: Prostitution and Liberation in Asia and the United States*. Minneapolis: Fortress Press.
- Browne, Jan, and Victor Minichiello 1996 "Research Directions in Male Sex Work." *Journal of Homosexuality* 31(4):29-56.
- Bullough, Bonnie, and Vern Bullough 1997 "Are Transvestites Necessarily Heterosexual?" *Archives of Sexual Behavior* 26(1):1-12.

- Butler, Robert N. 1998 "The Viagra Revolution." *Geriatrics* 53(10):8-9.
- Caraballo-Dieiguez, Alex 1997 "Sexual Research with Latino Men Who Have Sex with Men." In John Bancroft, ed., *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- Carrier, Joseph M. 1989 "Gay Liberation and Coming Out in Mexico." *Journal of Homosexuality* 17:225-252.
- Check, J. V. P., and N. M. Malamuth 1983 "Sex Role Stereotyping and Reactions to Depictions of Stranger vs. Acquaintance Rape." *Journal of Personality and Social Psychology* 45:344-356.
- Cohen, Kenneth M., and Ritch C. Savin-Williams 1996 "Developmental Perspectives on Coming Out to Self and Others." In Ritch C. Savin-Williams and Kenneth M. Cohen, eds., *The Lives of Lesbians, Gays, and Bisexuals: Children to Adults*. New York: Harcourt Brace.
- De Graaf, Ron, and Gertjan Van Zessen, Ine Vanwesenbeeck, Cees J. Straver, and Jan H. Visser 1997 "Condom Use by Dutch Men with Commercial Heterosexual Contacts: Determinants and Considerations." *AIDS Education and Prevention* 9(5):411-423.
- Di Mauro, Diane 1997 "Sexuality Research in the United States." In John Bancroft, ed., *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- Doll, Lynda S. 1997 "Sexual Behavior Research: Studying Bisexual Men and Women and Lesbians." In John Bancroft, ed., *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- , and Carolyn Beeker 1996 "Male Bisexual Behavior and HIV Risk in the United States: Synthesis of Research with Implications for Behavioral Interventions." *AIDS Education and Prevention* 8(3):205-225.
- Evans, Nancy J., and Anthony R. D'Augelli 1996 "Lesbians, Gay Men and Bisexual People in College." In Ritch C. Savin-Williams and Kenneth M. Cohen, eds., *The Lives of Lesbians, Gays and Bisexuals*. New York: Harcourt Brace.
- Finkelhor, D. 1979 *Sexually Victimized Children*. New York: Free Press.
- , and K. Yllo 1985 *License to Rape*. New York: Holt, Rinehart and Winston.
- Ford, Clellan, and F. Beach 1951 *Patterns of Sexual Behavior*. New York: Harper.
- Gagnon, John 1977 *Human Sexualities*. Glenview, Ill.: Scott, Foresman.
- George, L. K., and S. J. Weiler 1981 "Sexuality in Middle and Late Life." *Archives of General Psychiatry* 38:919-923.
- Grauerholz, Elizabeth, and Jennifer Crew Solomon 1989 "Sexual Coercion: Power and Violence." In Kathleen McKinney and Susan Sprecher, eds., *Human Sexuality: The Societal and Interpersonal Context*. Norwood, N.J.: Ablex.
- Hass, W. 1979 *Teenage Sexuality*. New York: Macmillan.
- Herd, Gilbert, and Andrew Boxer 1993 *Children of Horizons: How Gay and Lesbian Teens Are Leading a New Way Out of the Closet*. Boston: Beacon.
- Herold, W., and L. Way 1983 "Oral-Genital Sexual Behavior in a Sample of University Females." *The Journal of Sex Research* 19:327-338.
- Hetrick, Eric, and A. D. Martin 1987 "Developmental Issues and Their Resolution for Gay and Lesbian Adolescents." *Journal of Homosexuality* 14:25-44.
- Hickson, Ford C. I., Peter M. Davies, Andrew J. Hunt, Peter Weatherburn, Thomas J. McManus, and P. M. Anthony 1994 "Gay Men as Victims of Nonconsensual Sex." *Archives of Sexual Behavior* 23(3):281-294.
- Jay, Karla, and Allen Young, eds., 1979 *The Gay Report: Lesbians and Gay Men Speak Out about Sexual Experiences and Lifestyles*. New York: Simon and Schuster.
- Jayne, C. E. 1986 "Methodology on Sex Research in 1986: An Editor's Commentary." *The Journal of Sex Research* 22:1-5.
- Jeness, Valerie 1990 "From Sex as Sin to Sex as Work: COYOTE and the Reorganization of Prostitution as a Social Problem." *Social Problems* 37(3):403-420.
- Johnson, Susan E. 1990 *Staying Power: Long Term Lesbian Couples*. Tallahassee, Fla.: Naiad Press.
- Kelley, K. 1986 "Integrating Sex Research." In D. Byrne and K. Kelly, eds., *Alternative Approaches to the Study of Sexual Behavior*. Hillsdale, N.J.: Erlbaum.
- Kempe, R. S., and H. Kempe 1984 *The Common Secret: Sexual Abuse of Children and Adolescents*. New York: Freeman.
- King, Michael, and Earnest Woollett 1997 "Sexually Assaulted Males: 115 Men Consulting a Counseling Service." *Archives of Sexual Behavior* 26(6):579-588.
- Kinsey, Alfred C., Wardell Pomeroy, and Clyde Martin 1948 *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- , ——, Clyde Martin, and Paul Gebhard 1953 *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- Koss, Mary, Kenneth E. Leonard, Dana A. Beezley, and Cheryl J. Oros 1985 "Nonstranger Sexual Aggression: A Discriminant Analysis of the Psychological Characteristics of Undetected Offenders." *Sex Roles* 12:981-992.
- , Thomas E. Dinero, Cynthia A. Seibel, and Susan L. Cox 1988 "Stranger and Acquaintance Rape: Are

- There Differences in the Victim's Experience?" *Psychology of Women Quarterly* 12:1-23.
- Ku, Leighton, Freya Sonenstein, Laura D. Lindberg, Carolyn H. Bradnere, Scott Boggess, and Joseph H. Pleck 1998 "Understanding Changes in Sexual Activity among Young Metropolitan Men: 1979-1995." *Family Planning Perspectives* 30:256-262.
- Laumann, Edward O., John Gagnon, Robert T. Michael, and Stuart Michaels 1994a *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
- Laumann, Edward O., Robert T. Michael, and John Gagnon 1994b "A Political History of the National Sex Survey of Adults." *Family Planning Perspectives* 26:34-38.
- Lottes, Ilsa L., and Martin S. Weinberg 1996 "Sexual Coercion among University Students: A Comparison of the United States and Sweden." *The Journal of Sex Research* 34:67-76.
- Manalansan, Martin F. 1996 "Double Minorities: Latino, Black and Asian Men Who Have Sex With Men." In Ritch C. Savin-Williams and Kenneth M. Cohen, eds., *The Lives of Lesbians, Gays and Bisexuals*. New York: Harcourt Brace.
- Marsiglio, William, and Denise Donnelly 1991 "Sexual Intercourse in Later Life: A National Study of Married Persons." *Journal of Gerontology* 46:S338-344.
- Masters, William H., and Virginia E. Johnson 1979 *Homosexuality in Perspective*. Boston: Little, Brown.
- Meyer, Ilan H., and Laura Dean 1995 "Patterns of Sexual Behavior and Risk Taking among Young New York City Gay Men." *AIDS Education and Prevention* 7:13-23.
- Michael, Robert T., John H. Gagnon, Edward O. Laumann, and Gina Kolata 1994 *Sex in America*. Boston: Little, Brown.
- Miller, Jody 1993 "'Your Life is on the Line Every Night You're on the Streets': Victimization and the Resistance among Street Prostitutes." *Humanity and Society* 17(4):422-446.
- Moore, Kristin Anderson, Anee K. Dirscoll, and Laura Duberstein Lindberg 1998 *A Statistical Portrait of Adolescent Sex, Contraception, and Childbearing*. Washington, D.C.: National Campaign to Prevent Teen Pregnancy.
- , Brent C. Miller, Dana Gleib, and Donna Ruane Morrison 1995 *Adolescent Sex, Contraception, and Childbearing: A Review of Recent Research*. Washington, D.C.: Child Trends.
- , Christine Winquist Nord, and James L. Peterson 1989 "Nonvoluntary Sexual Activity among Adolescents." *Family Planning Perspectives* 21:110-114.
- Morse, Edward V., Patricia M. Simon, Paul M. Balson, and Howard J. Osofsky 1992 "Sexual Behavior Patterns of Customers of Male Street Prostitutes." *Archives of Sexual Behavior* 21(4):347-357.
- Mosher, D. L., and R. D. Anderson 1986 "Macho Personality, Sexual Aggression, and Reactions to Guided Imagery of Realistic Rape." *Journal of Research in Personality* 20:77-94.
- Muehlenhard, Charlene L. 1994 "Controversy about Rape Research and Activities." *Journal of Sex Research* 31:143.
- , and L. C. Hollabaugh 1988 "Do Women Sometimes Say No When They Mean Yes?: Prevalence and Correlates of Women's Token Resistance to Sex." *Journal of Research in Personality and Social Psychology* 54:872-879.
- , Susie C. Sympson, Joi L. Phelps, and Barrie J. Highby 1994 "Are Rape Statistics Exaggerated? A Response to Criticism of Contemporary Rape Research." *Journal of Sex Research* 31:144-146.
- Newcomer, S., and J. Udry 1985 "Oral Sex in an Adolescent Population." *Archives of Sexual Behavior* 14:41-46.
- Okami, P., R. Olmstead, and P. R. Abramson 1997 "Sexual Experiences in Early Childhood: 18-Year Longitudinal Data from the UCLA Family Lifestyles Project." *The Journal of Sex Research* 34:339-347.
- O'Sullivan, Lucia, and Elizabeth Rice Allgeier 1998 "Feigning Sexual Desire: Consenting to Unwanted Sexual Activity in Heterosexual Dating Relationships." *The Journal of Sex Research*, 35:234-243.
- Parker, Richard G. 1997. "International Perspectives on Sexuality Research." In John Bancroft, ed., *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- Peplau, Latitia Anne 1982 "Research on Homosexual Couples: An Overview." *Journal of Homosexuality* 8(2):3-7.
- , Susan D. Cochran, and Vicki M. Mays 1997 "A National Survey of the Intimate Relationships of African-American Lesbians and Gay Men." In Beverly Green, ed., *Ethnic and Cultural Diversity among Lesbians and Gay Men*. Thousand Oaks, Calif.: Sage.
- Perkins, Roberta, and Gary Bennett 1985 *Being a Prostitute: Prostitute Women and Prostitute Men*. London: George Allen & Unwin.
- Pope, Mark, and Richard Schulz 1990 "Sexual Attitudes and Behavior in Midlife and Aging Homosexual Males." *Journal of Homosexuality* 20(3-4):169-177.
- Poppen, Paul J., 1995 "Gender and Patterns of Sexual Risk Taking in College Students." *Sex Roles* 32(7-8):545-555.

- Rapaport, K., and B. R. Burkhart 1984 "Personality and Attitudinal Correlates of Sexual Coercive College Males." *Journal of Abnormal Personality* 93:216-221.
- Raphael, Sharon M., and Mina K. Robinson 1992 "The Older Lesbian: Love Relationships and Friendship Patterns." In Wayne R. Dynes and Stephen Donaldson, eds., *Sociology of Homosexuality*. New York: Garland.
- Renzetti, Claire M. 1992 *Violent Betrayal: Partner Abuse in Lesbian Relationships*. Newbury Park, Calif.: Sage.
- Riportella-Muller, Roberta 1989 "Sexuality in the Elderly: A Review." In Kathleen McKinney and Susan Sprecher, eds., *Human Sexuality: The Sexual and Interpersonal Context*. Norwood, N.J.: Ablex.
- Risman, Barbara J., and Pepper Schwartz 1988. "Sociological Research on Male and Female Homosexuality." *Annual Review of Sociology* 14:125-147.
- Roiphe, Katie 1993 *The Morning After: Sex, Fear, and Feminism*. Boston: Little, Brown.
- Rubin, Lillian 1990 *Erotic Wars: What Happened to the Sexual Revolution*. New York: HarperPerennial.
- Russell, D. E. H. 1984 *Sexual Exploitation*. Beverly Hills, Calif.: Sage.
- , and N. Howell 1982 *Rape in Marriage*. New York: Macmillan.
- Sandfort, Theo G. M. 1997 "Sampling Male Homosexuality." In John Bancroft, ed., *Researching Sexual Behavior: Methodological Issues*. Bloomington: Indiana University Press.
- Savin-Williams, Ritch C., and Kenneth M. Cohen 1996 "Memories of Childhood and Early Adolescent Sexual Feelings among Gay and Bisexual Boys: A Narrative Approach." In Ritch C. Savin-Williams and Kenneth M. Cohen, eds., *The Lives of Lesbians, Gays, and Bisexuals: Children to Adults*. New York: Harcourt Brace.
- Scully, D., and J. Marolla 1984 "Convicted Rapists' Vocabulary of Motive: Excuses and Justification." *Social Problems* 31:530-544.
- Sheahan, Sharon L., Stephen Joel Coons, John P. Seabolt, Lance Churchill, and Dale Thomas 1994 "Sexual Behavior, Communication, and Chlamydial Infections among College Women." *Health Care for Women International* 15(4):275-286.
- Sonenstein, Freya, Kellie Stewart, Laura Duberstein Lindberg, Marta Pernas, and Sean Williams 1997 *Involving Males in Preventing Teen Pregnancy: A Guide for Program Planners*. Washington, D.C.: Urban Institute.
- Starr, Bernard D., and Marcella B. Weiner 1981 *The Starr-Weiner Report on Sex and Sexuality in the Mature Years*. Briarcliff Manor, N.Y.: Stein & Day.
- Stokes, Joseph P., Peter Vanable, and David J. McKirnan 1997 "Comparing Gay and Bisexual Men on Sexual Behavior, Condom Use and Psychosocial Variables Related to HIV/AIDS." *Archives of Sexual Behavior* 26(4):383-397.
- Struckman-Johnson, Cindy, and David Struckman-Johnson 1994 "Men Pressured and Forced into Sexual Experience" *Archives of Sexual Behavior* 23:93-114.
- , David Struckman-Johnson, Lila Rucker, Kurt Bumby, and Stephan Donaldson 1996 "Sexual Coercion Reported by Men and Women in Prison." *Journal of Sex Research* 33:67-76.
- Udry, J. Richard 1988 "Biological Predispositions and Social Control in Adolescent Sexual Behavior." *American Sociological Review* 53:709-722.
- Waldorf, Dan, and David Lauderback 1992 "The Condom Use of Male Sex Workers in San Francisco." *AIDS & Public Policy Journal* 7(2):108-119.
- Warren, Charles W., John S. Santelli, Sherry A. Everett, Laura Kann, Janet L. Collins, Carol Cassell, Leo Morris, and Lloyd J. Kolbe 1998 "Sexual Behavior among U.S. High School Students, 1990-1995." *Family Planning Perspectives* 30(4):170-172, 200.
- Whitam, Frederick L., Christopher Daskalos, Curt Sobolewski, and Peter Padilla 1998 "The Emergence of Lesbian Sexuality and Identity Cross-Culturally: Brazil, Peru, the Phillipines, and the United States." *Archives of Sexual Behavior* 27(1): 31-56.
- Young, M. 1980 "Attitudes and Behaviors of College Students Relative to Oral-Genital Sexuality." *Archives of Sexual Behavior* 9:61-67.

WILLIAM MARSIGLIO  
JOHN H. SCANZONI  
KENDAL L. BROAD

## SEXUAL INEQUALITY

See Discrimination; Sex Differences; Social Stratification.

## SEXUAL ORIENTATION

As Nietzsche noted, "[O]nly that which has no history is definable" ([1887] 1968, p. 516). This observation is clearly supported by the dramatic changes in the ways in which sexual orientation has been conceptualized over the last quarter century and, in particular, the last decade. As a result, the concept of sexual orientation may be difficult to define with any assurance of general agree-

ment. It is currently mired, and surely will continue to be mired, in conflicting interpretations of the history of the behaviors that are assumed to be the expression of specific sexual orientations. The question of sexual orientation remains a conceptual battleground where many of the most critical issues regarding the nature of human sexuality, if not the human condition itself, are debated.

Sexual orientation generally can be described as the integration of the ways in which individuals experience the intersection of sexual desires and available sexual social roles. For some people, this intersection is experienced happily as an unproblematic confluence of personal and social expectations. For others, it is experienced as a persistent conflict. For still others, issues of sexual orientation are experienced as an occasion for experimentation, compromise, and sometimes change in how they see themselves, how they present themselves to others, and how different segments of social life respond to such outcomes.

Sexual orientation is also part of the conceptual apparatus of contemporary scientific and popular discourse; it has become a way in which people recognize and “explain” sexual behavior. It is as if establishing an individual’s sexual orientation, however inaccurately, were enough to explain most of what has to be known about that individual’s sexuality. As a result, the discourses of sexual orientation often become a point of contact with the discourses of age, gender, morality, and law.

#### SEXUAL ORIENTATION AND GENDER

Following Freud’s distinction between the “object” (the “who”) and the “aim” (the “what”) of sexual desire ([1905] 1953), current conceptions of sexual orientation can be said to focus primarily on the nature of the object defined in narrow terms of gender. This almost exclusive distinction derives from the dimorphic nature of the human species, that is, two genders giving rise to three possible categories—homosexual, heterosexual, and bisexual—although within each of these categories there is a wide range of variations in both sexual and nonsexual attributes of individuals and there are many aspects of sexual preference that are shared across these categories. Among such aspects of desire would be the other’s age, race, social class, and ethnic status; the nature of the emotional bond; and the conventions of physical

beauty. Important differences regarding sexual aims, such as sadomasochism, pedophilia, hebophilia (sexual attraction to postpubescent minors), and transvestism, are most often subsumed within each of these gender-based categories. Most often, they become adjectives modifying the label “homosexual” or “heterosexual.”

The continuing significance of gender may reflect the fact that within modern Western societies, gender is possibly the last fully pervasive aspect of identity that provides cohesion among the increasingly complex components of multiple social roles. Gender serves this role in a social context of continuing change because of its seeming permanence and seemingly *ascriptive* character. As a result, the gender of the object of one’s desires continues to dominate the meaning of sexual orientation almost to the exclusion of all the other attributes of potential partners that contribute to or preclude sexual interest or excitement.

This emphasis on the gender of the object of desire may be a culturally specific development. Some, for example, have argued that in other cultural or historical contexts, gender may be less significant in defining categories of legitimate sexual access than are other social distinctions. Thus, the acceptability of same-gender sexual contacts among males in ancient Greece was contingent on differences in age (mature adult versus youth) and social status (free citizen versus slave). Respect for those distinctions in social status required that there be no direct reciprocity, that the “active” role (the seeking of sexual pleasure) and the “passive” role (the providing of sexual pleasure) remain respectful of social status (Halperin 1989). Men engaging in such behavior were viewed as conventional so long as those rules were maintained. Such examples indicate that not all persons engaging in sexual acts experience their participation as erotic or experience those activities in the context of what might be termed sexual excitement. By the same token, they also indicate that not all motives for engaging in specific sexual acts derive from intrinsically sexual motives.

Through much of the twentieth century, the question of sexual orientation would not have appeared problematic. In a range of theoretical positions, from Freud’s assumption of an inherent bisexuality ([1905] 1953) to those postulating an exclusive heterosexuality, sexual orientation was

taken as being so firmly rooted in the “natural” process of human psychosexual development that it was treated as a transcultural phenomenon (Simon 1996). This was true for heterosexuality, which often was, and for many people still is, viewed as being phylogenetically programmed as a requirement of species survival (Symons 1979; Wilson 1978). Homosexuality and bisexuality were viewed as a disturbance of “normal” development (Freud [1905] 1953), an inherited decadence (Ellis 1937), a gender-discordant development (Krafft-Ebing [1896] 1965), later as a normal but minor genetic variant (Kinsey et al. 1948), and more recently as a sociocultural construction (see below). This social constructionist position has been extended to the treatment of heterosexuality as well (Katz 1995; Richardson 1996).

Explanations of sexual orientation currently might be described as a continuum anchored at one polar position by the assumption of an entirely biological or phylogenetic source (essentialism) and at the other polar position by sources reflecting the adaptation of specific individuals within given sociocultural settings (constructionism) and still more recently as the reflection of the sociocultural construction of gender as a binary phenomenon (queer theory).

**Essentialist Perspectives.** The extreme essentialist position leads to a view of sexual orientation (as gender preference) that is potentially present in all human populations, varying only in its manifest expression as a result of differing qualities of encouragement or repression (Gladue 1987; Boswell 1983; Whitam 1983; LaVay 1996). Other biologically oriented explanations link biological developments with experiential adaptations. Typically, those approaches link variations in phenomena such as prenatal hormonal chemistry with critical but often unpredictable postnatal experiences in the shaping of sexual orientation (Money 1988).

The essentialist end of the conceptual continuum assumes that at some basic level of character or personality, there are objective, constitutional sources of sexual orientation (Green 1988). It is almost as if such approaches viewed different categories of sexual orientation as different species or subspecies, as if all those included within a specific category of sexual interactions shared a common origin. A commitment to such permanent distinctions is often evident in the use of a concept such

as *latent homosexuality*, which implies that even when such differences fail to be manifested or are manifested late in life, this orientation is viewed as the “real” one.

**Constructionist Perspectives.** At the other end of this continuum are constructionists, who view sexual orientation as the product of specific historical contingencies, as something to be acquired or perhaps even “an accomplishment” (Stoller 1985a). Most of those holding this position reject the idea of a sexual drive or at best see such a drive as an unformed potential that is largely dependent on experience to give it power and directionality. “Every culture has a distinctive cultural configuration with its own ‘anthropological’ assumptions in the sexual area. The empirical relativity of these configurations, their immense variety, and luxurious inventiveness, indicate that they are products of man’s own socio-cultural formations rather than a biologically fixed human nature” (Berger and Luckman 1966, p. 49).

For most constructionists, sexual orientation is a reflection of the more general practices of a time and place and is expressive of social power (Foucault 1978; Weeks 1985; Padgug 1979; Greenberg 1988; Halperin 1989). Others would add concern for the specific contexts of interaction and the management of identities and social roles (Simon and Gagnon 1967; McIntosh 1976; Plummer 1975; Ponse 1968; Weinberg 1983), and still others would add concern for the experiences that constitute primary socialization (Gagnon and Simon 1973; Stoller 1985b; Simon and Gagnon 1986; Mitchell 1988).

From a constructionist perspective, the concept of sexual orientation itself is viewed as an aspect of the very cultural practices that sustain the differential evaluations of the sexual behaviors the concept purports to explain. The focusing of attention on something that can be called sexual orientation is seen as signifying an importance to be assigned to the sexual that may not be intrinsic to it but may derive from the evolved meanings and uses that constitute the sexual in specific sociohistorical contexts.

Whereas essentialists tend to view the sexual as a biological constant that presses on evolving social conventions, constructionists view the sexual as the product of the individual’s contingent response to the experiencing of social conven-

tions. For essentialists, the sexual might be said to develop from the inside out, whereas for constructionists, the sexual, like most other social practices, is learned from the outside in. A middle ground is taken by many who view sexual behavior as the outcome of a dialectical relationship between biology and culture (Erikson 1950).

### CURRENT CONCEPTS

If only in recognition of the enormous diversity of sexual practices in different cultural and historical settings despite the relative stability of human physiology, almost everyone who has approached the study of human sexuality admits the need for some degree of sociological explanation of specific patterns of sexual interaction and the significance accorded to them (Gregersen 1983). The question of homosexuality was the dominant issue in most discussions of sexual orientation until relatively recently. Heterosexuality, insofar as it was viewed as doing what came naturally, seemingly required no "explanation" unless it was expressed in unconventional ways. Instead, it was homosexuality that was viewed as problematic, if not pathological, and whose explanation was more urgent. The medicalization of same-gender sexual preference, which preceded the initial public use of terms such as "homosexuality" and "heterosexuality" in 1880 (Herzer 1985), involved the "disease" model of seeking a specific cause as well as a mode of prevention and possible cure. This implicitly homophobic commitment persists in some of the scientific community's considerations of homosexuality (Irving 1990).

The acceptance of homosexuality as an alternative life-style, however, did not necessarily require the abandonment of a concern for explaining its appearance; it merely made it more obvious that heterosexuality cannot be taken for granted but requires explanation (Katz 1990). One characteristic of the modern Western condition is that it has made sexual orientation and the closely related issue of sexual identity problematic. The question, What will I be when I grow up? is asked of an ever-growing number of dimensions of life, including the sexual, and is asked with increasing uncertainty regarding the possible answers.

**Heterosexuality.** Heterosexuality, defined as *cross-gender sexual intercourse*, has been a preference in all societies, though not necessarily an exclusive

preference in all societies. Nor does the universality of this preference establish the full range of definitions of with whom, when, where, or in what manner it should occur. Thus, outside of incest taboos involving immediate family members and a variable list of other close relatives, different cultures and periods of history have defined legitimate and illegitimate sexual contacts in dramatically contrasting ways (Bullough [1978] 1980). These differences involve not only what might be called the mechanics of sexual acts, that is, matters of relationship, time, place, costume, sequence of gestures, and positions, but also the determinants of their relative significance.

The potential reproductive consequences of heterosexual genital intercourse inevitably led to a linking of the desire for sex with a conscious or unconscious desire for reproduction. This view has been criticized as resting on the questionable assumption of a biologically rooted commitment to species survival (Beach 1956). Valid or not, such views constitute a cultural legacy that gives credence to many current norms regarding sexual acts, norms that enhance the social regulation of reproduction in the name of an assumed natural mandate.

More specifically, expectations regarding gender and family, influenced by many aspects of social life, generally have shaped the social meaning of sexual acts. Current language for describing cross-gender sexual contacts explicitly assumes a relationship to the family—marital, premarital, postmarital, and extramarital sex—and implicitly evaluates behaviors in terms of their "distance" from location within the family.

Similarly, genital intercourse still is viewed commonly as the ultimate or purest form of sexual exchange, as the "fulfillment of nature's intent." As a result, it continues to serve as the measure of the "normality" of alternative forms of sexual contact. This was reflected in the historical, but declining, practice of criminalizing not only sexual acts occurring outside of marriage but also those involving oral or anal contact or in viewing masturbation as pathogenic when practiced by the young and symptomatic when practiced by adults, although recent research (Lauman et al. 1994) indicates that masturbation occurs in North America among significant segments of postpubertal individuals at all stages of the life course.



Many of the conventions surrounding gender expectations also directly reinforced the “scripting,” or construction, of heterosexuality. This involves presenting images of the sexual that both naturalize and normalize evolved Western heterosexual practices, making them appear unquestionably proper. The labels “active” and “passive,” terms that had applications in many domains of social life, virtually became synonymous with “masculine” and “feminine,” respectively. Even physical positions in sexual intercourse—“Who is on top?”—have often had to pay homage to prevailing patterns of social domination.

The nineteenth century witnessed the elaboration of images of the female as fragile, domestic, nurturant, receptive, and either only minimally sexual or capable of insatiable lusts. These images of femininity were complemented by images of the male as strong, given to exploratory curiosity, possessively protective, and aggressively lustful. Although applied diffusely, these implicit norms were not always applied equally. The restraint and fragility of the female found a common application in the parlors and bedrooms of the urban middle class and rural gentry but was applied far less in the fields, factories, servants’ quarters, and brothels of the day.

While the images of heterosexuality reinforced patterns of family life and gender differentiation, it is equally appropriate to speak of the ways in which patterns of family life and gender differentiation reinforced prevailing concepts of the “naturalness” of heterosexuality. This same gender-based division of labor within the family was taken for granted by mid-twentieth century sociological theorists (Parsons and Bales 1955), as it was inscribed in the most widely held views regarding “normal” human development (Erikson 1950).

From the late nineteenth century on, concepts of the family became substantially more voluntary and egalitarian. However, those modifications further empowered the heterosexual scenario, which now plays an even more important role in the creation of marital bonding and the preservation of the nuclear family. Heterosexuality, given the assumption of its powers as a basic drive, simultaneously became a nearly constant threat to and vital aspect of family life. This in turn gave rise to various methods, both formal and informal, of restricting nonmarital expressions of sexual activity.

The emphasis placed on the heterosexual scenario led in turn to a greater emphasis on the subjective aspects of one’s sexual orientation. Faith in the mute logic of “nature’s” intent gave way to concern for the fashioning and maintenance of individual desire. Women increasingly were expected not only to be receptive but to desire as well as to be desirable. Men increasingly were expected to use the sexual to affirm their masculinity not only by their ability to find sexual pleasure but also by their ability to provide pleasure to their partners. Heterosexual preference continued to be taken for granted while heterosexual competence was being placed on the agenda in new and unanticipated ways.

In recent years, evident trends have called into question many of these practices, challenging many earlier basic expectations regarding family and gender. The conjugal family is no longer the exclusive social address for heterosexuality. Premarital sex has become *statistically* normal at all social levels, and it approaches becoming *attitudinally* normative. Moreover, the age at which sexual intercourse first occurs has declined, particularly for females. By age 18, over half are no longer virgins, which is more than double the proportion of nonvirgins reported two generations ago. This suggests that most of what occurs by way of sexual activity among adolescents and young adults can be described as pre-premarital, as much of this early sexual behavior occurs outside the context of family-forming courtship, where much of the premarital experience of older generations took place.

Similarly, at the premarital and postmarital stages, there has been increasing acceptance of nonmarital cohabitation in the sense that it tends to be more openly acknowledged with little anticipation of social rejection or stigmatization. While the number of middle- to upper-middle-class females who have deliberately borne children without marriage or an acknowledged male partner is not great, the fact that this practice has achieved considerable visibility and implicit legitimacy is significant.

Reflecting the diffusion of feminist values, support for women with regard to sexual interest, sexual activity, and especially sexual competence, with the latter measured by the capacity to achieve orgasm, has visibly increased (Ehrenreich et al. 1986). As a result, gender stereotypes with regard

to sexual behavior have experienced changes that for the most part have served to blur many of the gender distinctions that previously appeared to give heterosexuality its distinctive complementarity.

Specific behaviors, such as oral sex, that once were associated with devalued sexual actors, homosexuals, and prostitutes in recent years have become a conventional part of the heterosexual script. This is particularly true at higher social class levels, where oral sex tends to occur regularly, often substituting for genital intercourse (Gagnon and Simon 1987; Blumstein and Schwartz 1983; Simon et al. 1990).

Heterosexuality remains the dominant erotic imagery of Western societies. However, changing concerns for reproduction, continuing changes in the organization of family life, and the constraints describing gender presentations indicate that present trends toward a pluralization of the ways in which heterosexuality is experienced and the contexts within which it is expressed will continue into the imaginable future.

**Homosexuality.** Same-gender sexual interactions have been reported in a sufficient number of social settings to suggest that they fall within the normal range of human behaviors (Ford and Beach 1951; Gregersen 1983). As Kinsey and associates (1948) observed, "The homosexual has been a significant part of human sexual activity ever since the dawn of history, primarily because it is an expression of capacities that are basic in the human animal" (p. 666). This essentialist view implies that a predisposition to same-gender sexual acts is an immutable fact of nature like gender and race and as such is totally independent of personal preference and societal values (Green 1988).

The fact that same-gender sexual involvements fail to be reported or occur as atypical behaviors in sufficient numbers suggests that there is little about them on which to predicate a universal or singular explanation. Where homosexual behavior occurs, the specific forms it takes, and the kinds of sexual acts and the relations within which they occur, as with most aspects of heterosexuality, vary so much that a full understanding must be sought in terms of the contingent features of specific social contexts. In other words, apparent uniformity of acts, such as members of the same gender engaging in sexual acts, allows one to assume very little, if any, uniformity of actors, their development, their mo-

tives, or the social and personal meanings of their behavior. When constructionists assert that the homosexual is an invention of the modern world, they are not suggesting that same-gender contacts were unknown in earlier periods of Western history or in other cultural settings. What they do suggest is that the processes that constitute the behavior, that give it meaning, and that transform otherwise identical forms of "behavior" into different forms of evaluated "conduct" may be of a fundamentally different character.

The variety of meanings given to same-gender "sexual" contacts is as wide as that given to cross-gender contacts. "Sexual" is placed in quotation marks as a reminder that while genital contact and orgasm may be present, in many instances the behavior is not necessarily experienced as sexual in the contemporary Western sense of that word. Such same-gender contacts range from those which are incidental to religious rites or rites of puberty, to those specific to certain statuses that may be temporary and that are not in themselves significant aspects of the individual's social identity, to those in which same-gender contacts are defined as permanent features of the individual's character.

An example of age-specific sexual contact can be found among the Sambians of New Guinea. Male children at about age 6 are removed to the men's hut, where they ingest semen, a practice that is viewed as necessary for full masculine development, by engaging in fellation with older, unmarried fellow villagers. At puberty, such males enter the role of semen donor by making their penises available to their younger fellow villagers. During early adulthood, they enter arranged marriages and are expected to practice heterosexual sex exclusively for the remainder of their lives. Observers report a nearly universal absence of fixation with regard to the activities of earlier stages or a reversal of age roles (Herdt 1981; Stoller 1985a).

This of course stands in dramatic contrast with the modern Western experience, in which the imagery of the behavior is associated with powerful meanings whose very invocation is often capable of exciting intense emotional responses of all kinds. Thus, negative images promote strong feelings of homophobia and at times cause "homosexual panic" in which the fear of being or becoming homosexual generates highly charged nonrational responses. At the same time, the possibility of

same-gender sexual contacts often generates responses sufficiently strong to allow many individuals to experience and accept themselves as being homosexual despite the homophobic character of their immediate social settings (Bell and Weinberg 1978; Weinberg 1983).

In the examples of both the Sambians and the contemporary Western experience, the biological processes associated with arousal and orgasm are undoubtedly the same. What vary are the meanings and the representations that occasion arousal. As Beach noted, "Human sexual arousal is subject to extensive modification as a result of experience. Sexual values may become attached to a wide variety of biologically inappropriate stimulus objects or partners" (1956, p. 27).

Patterns of homosexual behavior, like those of heterosexual behavior, have manifested persistent change. While same-gender sexual contact was known in premodern Europe and was severely sanctioned, often treated as a capital offense, it was not viewed as being the behavior of a different kind of person but as a moral failing, a sin, to which all might be vulnerable (Bray 1982). Some have argued that a conception of homosexuality as a sexual orientation involving a distinct kind of person was a correlate of many of the changing patterns and values associated with the emergence of urban, industrial capitalism (Adam 1978; Hocquenghem 1978; Foucault 1978).

Within the category of male homosexuality, different styles of homosexual activity predominated in different periods of history and different social settings. If the concept of homosexuality is to have any meaning, such variations suggest that modern forms of homosexuality reflect an eroticization of gender, not a fixation on a specific form of sexual activity. In other words, it is the gender of the participants that generates and sustains sexual interest and only secondarily the specific form of sexual activity (Gagnon 1990; Simon et al. 1990).

The significance of gender in considerations of homosexuality has marked much of its recent history. Initial nineteenth-century views defined and implicitly explained homosexuality as an inversion of gender. Lesbians often were viewed as "men trapped in women's bodies," and gay men the reverse. Consistent with this, a common designation was "invert." Despite this early view, more

recent research indicates that in many respects lesbians and gay men tend in their sexual development and subsequent behavior to approximate their genders. This suggests that sexual development tends to follow gender socialization: Gender roles and gender role expectations influence sexuality more often than sexuality prompts changes in gender identity (Gagnon and Simon 1973; Blumstein and Schwartz 1983).

Change in sexual patterns has been a critical aspect of recent social history. Whereas heterosexual practice might be described as being increasingly privatized and dissociated from the major institutions of society, homosexual practice has moved from the margins of society to sharing the central stage. Whereas the family becomes less and less the exclusive legitimate context for heterosexual activities, the appearance and survival of bonded relationships among homosexuals, particularly gay men, has visibly increased. Whereas the larger community appears increasingly anomic, gay communities (which once were limited to bars, discreet networks of friends, and, for gay men, locations for anonymous sexual contacts) now rival even the most solidary of ethnic groups. There is a flowering of recreational, religious, welfare, political, and other affinity groups and organizations as well as of areas of residential dominance (Epstein 1987; Escoffier 1998; Levine 1998).

Homosexuality remains negatively valued, remains stigmatized. Discrimination in employment and housing, instances of "gay bashing," and criminalization of same-gender sexual activity in some jurisdictions speak directly to continued homophobic practices and fears. However, on the whole, the 1960s, 1970s, and 1980s witnessed increasing acceptance of both homosexuality and the homosexual. Even the identification of gay men with transmission of the HIV virus, which initially was associated with an incipient moral panic and occasioned expressions of antihomosexual attitudes, became an occasion for sympathetic representation in the major public media and broadened understanding of gay men, their life-styles, and the many roles they play in and contributions they make to the larger society.

Currently, possibly owing to the heightened visibility of lesbians and gay men, the homogenizing of identities has been called into question. What once was viewed as a singular phenome-

non is now more generally seen as pluralized. It now encompasses different developmental histories, affording different ways of incorporating a homoerotic commitment within a specific life history. The same is true for the development of a heterosexual orientation (Murray 1984; Stoller 1985a). As a consequence, the major questions that previously dominated issues of homosexuality (What is *the* cause? How many are there?) have become incurably cloudy as claims for a "biology of identity" conflict with what has been called the "politics of identity." Same-gender sexual interactions remain a characteristic of a minority of persons (Lauman et al. 1994), but the actual number cannot be established without resolving the question of what a homosexual is. Reflecting the complexity of these questions, Lauman et al. (1994) distinguish between three critical dimensions: same-gender sexual *behavior*, same-gender sexual *attraction* or *desire*, and *identity* as a homosexual. They found that among the 10.1 percent of the men in their survey who reported *any* adult same-gender sexuality, only 24 percent reported positive responses on all three dimensions and among the 8.6 percent of the women reporting *any* adult same-gender sexuality, only 15 percent were positive on all three dimensions. (p. 298)

What is clear is that there may be many more reasons for developing a homosexual orientation than there are ways of giving it expression. Aspects of development such as variations in the development of a gender identity may be significant in the development of homosexual orientations for some individuals (Harry 1982; Green 1987), but these aspects may have to be reconsidered as society modifies its more general beliefs and practices regarding gender identity. For example, the question must be asked, On what basis should "effeminacy" in male children be treated as symptomatic of some pathology any more than comparable displays of "effeminacy" in female children? Or the reverse, regarding what is commonly referred to as "tomboyishness" among young females?

Individuals with a marked homosexual preference appear in virtually all social contexts: different types of community settings; at different class levels; in all racial, ethnic, and religious categories; and from all manner of family backgrounds (Gebhard and Johnson 1979; Bell et al. 1981). There are differences between such categories, but in the absence of unbiased and comprehensive data, it is

difficult to determine with any confidence whether significant effects are associated with possible differentials. There is reason to suspect that such statistics can provide only an approximation of current populations and a poor guide to future developments, developments that depend more on society's conceptions and uses of sex and gender, which appear to be in continuing transition.

Whether rooted in biology, social experience, or some combination of these elements, homosexuality as a concept, as a class of persons, and as social groups will persist into the twenty-first century. However, factors such as the greater visibility of representations in the media that challenge prior negative stereotypes and folk psychologies, the depathologizing of homosexuality by medical and social science communities, and its greater acceptance by conventional major institutions may be in the process of transforming what was once a closeted, isolated group into one that is different but is seen as neither abnormal nor threatening. Indeed, one may be observing a process of the *normalization* of homosexuality. Some, in effect, have raised the question of the disappearance of homosexuality through its assimilation into mainstream social life as minor variant. (Bech 1997; Seidman et al. 1999). Alternative arguments suggest that this ultimate assimilation is an unrealistic and undesirable possibility based on conceptions of gender (queer theory) that are ethnocentrically biased in their narrow emphasis on the North American and Western European experiences (Murray 1996).

**Bisexuality.** Bisexuality is a complex concept and, like homosexuality and heterosexuality, has become more complex in recent years as the framing concepts of gender have become less arbitrarily complementary and distinct (Butler 1993; Weinberg et al. 1994). Bisexuality can refer to behavior (those who have had both homosexual and heterosexual experience), psychic response (those capable of being erotically aroused by both homosexual and heterosexual imagery), and either social labeling or self-labeling. Substantial numbers of people have had, if only incidentally, both homosexual and heterosexual experiences while retaining a firm self-identity as being one or the other (Lauman et al. 1994). Even larger numbers have or can be assumed to have experienced sexual arousal in association with both heteroerotic and homoerotic imagery (Kinsey et al. 1948, 1953;

Bell and Weinberg 1978; Bell et al. 1981). The mere experience of having sex with members of both genders may not be sufficient to justify the application of the term "bisexual." What can be called "situational same-gender, sexual contacts," such as those which occur in single-gender penal institutions, may represent little more than conventionally styled heterosexual orientations expressed in restrictive circumstances (Gagnon and Simon 1973).

Relatively few people conceive of themselves as bisexual or can be labeled as such, particularly if the concept is defined as an attraction to both genders and an attraction for the sexual behaviors commonly attributed to both genders. However, bisexual identification probably has increased as an expression of an increase among younger cohorts of "open gender schemas" (Weinberg et al. 1994).

Many people whose sexual histories involve interaction with both genders still see themselves as being either homosexual or heterosexual in orientation. This may be a reflection of the fact that outside of relatively few "bisexual support groups," until recently neither heterosexual nor homosexual social worlds appeared to accept or validate such an identity. The very concept of bisexuality, when used to refer to a specific type of person, was viewed with skepticism (Tripp 1987). Having bisexual interests often was viewed as a mask or apology for an underlying orientation. Undoubtedly, for some people the bisexual label served as a transitional phase in the complicated task of identity transformation. Although a large number of the psychotherapeutic communities accept bisexuality as a distinct type of psychosexual development (Hill 1989), even among those who identify themselves as bisexual there are some who tend to have patterns of sexual behavior that are "amazingly diverse and that [their] day to day life roles are greatly different from one another . . . [and] it is clear that people come to bisexuality in an incredibly diverse number of ways" (Blumstein and Schwartz 1976, p. 180).

Bisexuality as denoting a special orientation tends to be a recent conceptualization, that reflects the increased recognition of gender as a crystallization of erotic responses that are not necessarily coded by the logic of an excluding complementarity. Prior images of bisexuality re-

flected the assumed differences of masculinity and femininity such as the persistently masculinized dominant sexual actor and the individual who could switch between stereotypical presentations of gender. Increased recognition of a bisexual possibility follows the recognition of the possible absence of complementarity, that is, with each participant providing what is absent in and desired by the other, within many heterosexual and homosexual relationships and the calling into question an implicit complementarity within existing conceptions of gender (Garber 1995).

**Transvestism and Transsexuality.** These two concepts do not represent discrete categories so much as a continuum describing the degree to which an individual biologically of one gender desires and enacts the identities or aspects of the identities of the other (Feinbloom 1976). For an unknown number of people this is limited to using the clothing of the other gender to elicit sexual excitement, with little more being directly involved. For most, however, more is involved; for most it involves adopting and enacting, if only for an audience of oneself, aspects of the identity and selected roles of the other gender, not merely cross-dressing but cross-gendering. However, for transvestites, cross-dressing is temporary, and they do not abandon their primary gender identity; they play at being the other (Newton 1979).

At the other end of the continuum is the transsexual who ideally seeks to adopt permanently the gender, costumes, and roles of the other gender (Green 1974). While an absolute realization of this aspiration is impossible, combined modern surgical and pharmacological techniques and permissive bureaucracies (the former cosmetically "redesigning" the body, while the latter allow for a redesigning of one's identifying credentials) have brought about the possibility of coming close to allowing some people to more fully realize their aspiration to live their lives, as fully as possible, in the costumes and roles of the other gender (Bolin 1988; Lothstein 1983).

The desirability of supporting transsexuality remains a matter of continuing contention that involves issues of mental health and gender. Several medical centers that once maintained programs of "surgical gender reassignment" have suspended those programs after reporting results

that were too mixed to justify their continuation. Additionally, feminists have criticized such programs as catering to the desire to enact some of the most extreme forms of gender stereotypes (Irving 1990).

Midway between these extremes, between the erotic fetishizing of the clothing of the opposite gender and the desire to become the opposite gender, are those who prefer the costumes and behavior roles of the opposite gender without wanting to or needing to abandon their own initial gender or genitalia. This ranges from those who deliberately blur costumes and the coding or semiotics of gestures to obscure distinctions—women who have masculinized or men who have feminized their presentation of self—to those who experience a continuing conflict between a “masculine self” and a “feminine self,” feeling that each of these components of a divided self requires its own costumes, vocabulary of gestures, and social space. (Bullough et al. 1997)

Again, as is true for most forms of stigmatized behavior, estimates of how many individuals are involved in such practices are virtually impossible to determine with any accuracy. Across this continuum of cross-gendering, both males and females can be observed. Most researchers speculate that more males than females are involved, generalizing the apparent tendency for significantly more males than females to be involved in various kinds of sexual deviance.

This speculation is made additionally plausible by the manifest tendency for violations of gender by men to generate more nervousness and be more heavily sanctioned than are comparable violations by females. It is possible for many females to mask their transvestic desires through the broader range of fashion available to women. For example, female cross-dressing in film and literature often involves the beginnings of romantic investment, while male cross-dressing is almost entirely restricted to the comic mode.

These two concepts, transvestism and transsexuality, perhaps more than any other, speak to the powers of gender and its multiple correlates. They speak as well to the complex relationship between gender and sexuality. For relatively few people are gender presentations altered to facilitate a specific sexual aim; more often the sexual is

organized to facilitate desired gender effects. Little that is manifestly sexual appears in the cross-gendering of some people, as in the case of many male heterosexual transvestites. In the case of the transsexual, surgical procedures often diminish orgasmic capacity. However, confirmation is often one of the major motives for engaging in sexual behavior and a major source of its capacity to gratify, a capacity that may go well beyond the narrow physicalist emphasis on orgasm.

However, conflict between the sexual as genital involvement and the sexual as gender confirmation has been seen by some as an expression of the application of a socially constructed arbitrary binary system that coercively mutes existing heterogeneities of desires and identity themes. This recently articulated perspective views *transgendering* as a moment of potential liberation from an arbitrary binary gender system, allowing individuals to give fuller expression to the totality of their eroticized and noneroticized desires (Stone 1991).

## CONCLUSIONS

Sexual orientation is a complex construct rather than a simple thing. While it tends to identify individuals in terms of commitment to similar sexual preferences, it also has the capacity to mask differences among those who appear to otherwise share identical orientations. This is not surprising. Sexual behaviors, like many other aspects of human experience that are linked in critical ways to biology, are also historical and subject to change and as such reflect the very connections of the sexual ultimately to the total fabric of social life.

At the same time, concepts of sexual orientation are aspects of the cultural apparatus of a time and place and are used to explain the behaviors of others as well as one's own behavior. As such, they have the capacity to influence the very behaviors they appear merely to describe. Thus, to view the sexual in isolation from the continuing dynamic of social life, which until recently has largely been its fate, is to run the risk of unself-consciously transforming the science of social life into an oppressive disciplinary instrument of social life.

(SEE ALSO: *Alternative Life-Styles; Heterosexual Behavior Patterns; Sexual Behavior and Marriage*)

## REFERENCES

- Adam, Barry 1978 "Capitalism, the Family, and Gay People." *Sociologists Gay Caucus Working Papers I*.
- Beach, Frank 1956 "Characteristics of Masculine Sex Drive." In Marshall R. Jones, ed., *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- Bech, Henning 1997 *When Men Meet: Homosexuality and Modernity*. Chicago: University of Chicago Press.
- Bell, Allen C., and Martin S. Weinberg 1978 *Homosexualities: A Study of Diversity among Men and Women*. New York: Simon & Schuster.
- , ———, and Susan Hammersmith 1981 *Sexual Preference*. Bloomington: Indiana University Press.
- Berger, Peter, and Thomas Luckman 1966 *The Social Construction of Reality*. Garden City, N.Y.: Doubleday.
- Blumstein, Phillip W., and Pepper Schwartz 1976 "Bisexuality in Women." *Archives of Sexual Behavior* 5(2):171-181.
- 1983 *The American Couple*. New York: Morrow.
- Bolin, Anne 1988 *In Search of Eve: Transsexual Rites of Passage*. Parsagr. South Hadley, Mass: Bergin and Garvey.
- Boswell, John 1983 "Revolutions, Universals, and Sexual Categories." *Salamundi* 58-59:89-113.
- Bray, Alan 1982 *Homosexuality in Renaissance England*. London: Gay Men's Press.
- Bullough, Bonnie, Vern Bullough, and James Elias 1997 *Gender Blending*. Amherst, N.Y.: Prometheus.
- Bullough, Vern L. (1978) 1980 *Sexual Variance in Society and History*. Chicago: University of Chicago Press.
- Butler, Judith 1993 *Bodies That Matter: On the Discursive Limits of "Sex"*. London: Routledge.
- Ehrenreich, Barbara, Elizabeth Hess, and Gloria Jacobs 1986 *The Feminization of Sex*. New York: Doubleday, Anchor.
- Ellis, Havelock 1937 *Studies in the Psychology of Sex*, vol. 2. New York: Random House.
- Epstein, Steven 1987 "Gay Politics, Gay Identities: The Limits of Social Construction." *Socialist Review* 93/94:9-53.
- Erikson, Erik 1950 *Childhood and Society*. New York: Norton.
- Escoffier, Jeffrey 1998 *American Homo: Community and Perversity*. Berkeley: University of California Press.
- Feinbloom, Deborah 1976 *Transvestites and Transsexuals: Mixed Reviews*. New York: Delacorte.
- Ford, Clellan, and Frank A. Beach 1951 *Patterns of Sexual Behavior*. New York: Harper.
- Foucault, Michel 1978 *The History of Sexuality*, vol. I: *An Introduction*. New York: Pantheon.
- Freud, Sigmund (1905) 1953 "Three Essays on the Theory of Sexuality." *Standard Edition*, vol. 7. London: Hogarth.
- Gagnon, John H. 1990 "Gender Preferences in Erotic Relations, the Kinsey Scale and Sexual Scripts." In David McWhorter, Stefanie Sanders, and June Reinisch, eds., *Heterosexuality, Homosexuality, and the Kinsey Scale*. New York: Oxford University Press.
- , and William Simon 1973 *Sexual Conduct: The Social Sources of Human Sexuality*. Chicago: Aldine.
- 1987 "The Scripting of Oral Genital Contacts." *Archives of Sexual Behavior* 16(1):1-25.
- Garber, Marjorie 1995 *Vice Versa: Bisexuality and the Eroticism of Everyday Life*. New York: Simon and Schuster.
- Gebhard, Paul H., and Alan B. Johnson 1979 *The Kinsey Data: Marginal Tabulation of the 1938-1963 Interviews Conducted by the Institute for Sex Research*. Philadelphia: Saunders.
- Gladue, Brian A. 1987 "Psychobiological Contributions." In Louis Diamant, ed., *Male and Female Homosexuality: Psychological Approaches*. New York: Hemisphere.
- Green, Richard 1974 *Sexual Identity Conflict in Children and Adults*. New York: Basic Books.
- 1987 *The "Sissy Boy Syndrome" and the Development of Homosexuality*. New Haven, Conn.: Yale University Press.
- 1988 "The Immutability of (Homo)Sexual Orientation: Behavioral Science Implications for a Constitutional (Legal) Analysis." *Journal of Psychiatry and Law* Winter, pp. 537-557.
- Greenberg, David F. 1988 *The Construction of Homosexuality*. Chicago: University of Chicago Press.
- Gregersen, Edgar 1983 *Sexual Practices: The Story of Human Sexuality*. New York: Franklin Watts.
- Halperin, David M. 1989 *One Hundred Years of Homosexuality and Other Essays on Greek Love*. New York: Routledge, Chapman, and Hall.
- Harry, Joseph 1982 *Gay Children Grow Up*. New York: Preager.
- Herd, Gilbert H. 1981 *Guardians of the Flute*. New York: McGraw-Hill.
- Herzer, Manfred 1985 "Kertbeny and the Nameless Love." *Journal of Homosexuality* 12(1):1-23.
- Hill, Ivan, ed. 1989 *The Bisexual Spouse: Different Dimensions in Human Sexuality*. New York: Harper Prenal.
- Hocquenghem, Guy 1978 *Homosexual Desire*, Danniella Dangoor, trans. London: Allison and Busby.

- Irvine, Janice M. 1990 *Disorders of Desire: Sex and Gender in Modern American Sexology*. Philadelphia: Temple University Press.
- Katz, Jonathan M. 1990 "The Invention of Heterosexuality." *Socialist Review* 90/91:7-34.
- 1995 *The Invention of Heterosexuality*. New York: Dutton.
- Kinsey, Alfred C., Wardell B. Pomeroy, Clyde E. Martin, and Paul H. Gebhard 1948 *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- 1953 *Sexual Behavior in the Human Female*. Philadelphia: Saunders.
- Krafft-Ebing, Richard von (1886) 1965 *Psychopathia Sexualis*, trans. by Harry E. Wedeck. New York: Putnam's.
- Lauman, Edward O., John H. Gagnon, Robert T. Michael, and Stuart Michaels 1994 *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
- LaVay, Simon 1996 *Queer Science*. Cambridge, Mass.: MIT Press.
- Levine, Martin P. 1998 *Gay Macho: The Life and Death of the Gay Clone*. Michael S. Kimmel (ed.). New York: New York University Press.
- Lothstein, Leslie 1983 *Female to Male Transsexualism: Historical, Clinical, and Theoretical Issues*. Boston: Routledge and Kegan Paul.
- McIntosh, Mary 1976 "The Homosexual Role." *Social Problems* 16(2):182-192.
- Mitchell, Stephan A. 1988 *Relational Concepts in Psychoanalysis*. Cambridge, Mass.: Harvard University Press.
- Money, John 1988 *Gay, Straight, and In-Between: The Sexology of Erotic Orientations*. New York: Oxford University Press.
- Murray, Stephen O. 1984 *Social Theory, Homosexual Reality*. New York: Gay Academic Union.
- 1996 *American Gay*. Chicago: University of Chicago Press.
- Newton, Esther 1979 *Mother Camp: Female Impersonators in America*. Chicago: University of Chicago Press.
- Nietzsche, Frederic (1887) 1968 *On the Genealogy of Morals*. In Walter Kaufmann trans., *The Basic Writings of Nietzsche*. New York: Modern Library.
- Padgug, Robert A. 1979 "Sexual Matters: On Conceptualizing Sexuality in History." *Radical History Review* 20:3-33.
- Parsons, Talcott, and Robert F. Bales 1955 *Family Socialization and Interaction Process*. New York: Free Press.
- Plummer, Kenneth 1975 *Sexual Stigma: An Interactionist Approach*. London: Routledge and Kegan Paul.
- Ponse, Barbara 1978 *Identity in the Lesbian World*. Westport, Conn.: Greenwood.
- Richardson, Diane, ed. 1996 *Theorizing Heterosexuality*. Buckingham, U.K.: Open University Press.
- Seidman, Steven, C. Meeks, and F. Trachten 1999 "Beyond the Closet: The Changing Meaning of Homosexuality in the United States." *Sexualities* 2(1):11-34.
- Simon, William 1996 *Postmodern Sexualities*. London: Routledge.
- , and John H. Gagnon 1967 "Homosexuality: The Formulation of a Sociological Perspective." *Journal of Health and Human Behavior* 2(1):9-37.
- 1986 "Sexual Scripts: Permanence and Change." *Archives of Sexual Behavior* 15(2):97-120.
- , D. Kraft, and H. Kaplan 1990 "Oral Sex: A Critical Overview." In June Reinisch, Bruce Vollmer, and Michael Goldstein, eds., *AIDS and Sex: A Biomedical and Behavioral Approach*. New York: Oxford University Press.
- Stoller, Robert J. 1985a *Observing the Erotic Imagination*. New Haven, Conn.: Yale University Press.
- 1985b *Presentations of Gender*. New Haven, Conn.: Yale University Press.
- Stone, Sandy 1991 "The Empire Strikes Back: A Posttranssexual Manifesto." In Julia Epstein and Kristina Straub, eds., *Body Guards: The Cultural Politics of Gender Ambiguity*. New York: Routledge.
- Symons, Donald 1979 *The Evolution of Human Sexuality*. New York: Oxford University Press.
- Tripp, Clarence 1987 *The Homosexual Matrix*, 2nd ed. New York: New American Library.
- Weeks, Jeffrey 1985 *Sexuality and Its Discontents: Meanings, Myths, and Modern Sexualities*. London: Routledge and Kegan Paul.
- Weinberg, Martin S., Colin J. Williams, and Douglas W. Pryor 1994 *Dual Attractions: Understanding Bisexuality*. New York: Oxford University Press.
- Weinberg, Thomas S. 1983 *Gay Men, Gay Identities*. New York: Irvington.
- Whitam Frederic L. 1983 "Culturally Invariable Properties of Male Homosexuals." *Archives of Sexual Behavior* 12:207-222.
- Wilson, Edward O. 1978 *On Human Nature*. Cambridge, Mass.: Harvard University Press.

WILLIAM SIMON



## SEXUAL VIOLENCE AND EXPLOITATION

Sexual violence and exploitation are social problems that were relatively neglected as research topics until the late 1960s. This article focuses primarily on the different categories of rape, sexual harassment, child sexual abuse, and related issues.

### RAPE AS A FORM OF SEXUAL VIOLENCE

Rape was not conceptualized as a major problem in the United States until the late 1960s; this awareness accompanied the resurrected women's movement and the establishment of the National Center for the Prevention and Control of Rape. After this awakening of interest, enough information was generated to document the fact that there were explanations for rape that went beyond the biological and psychological. Society and its institutions, laws, and attitudes were seen as contributing greatly to the problem. In the 1970s, rape was clearly defined as a social problem.

The works of three feminists contributed to advancing awareness of rape or sexual violence against women: Millet's *Sexual Politics* (1972), Griffin's "Rape: The All American Crime" (1971), and Brownmiller's *Against Our Will: Men, Women, and Rape* (1975). Millet argued that rape is linked to the concept of patriarchy, in which men use power and coercion to control women's sexuality. Nowhere is this better articulated than in the "crimes of honor" concept or social norm still evident in many Middle Eastern countries. Griffin's article focused on the nature of rape. She argued that rape is not a sexual act but a violent political act and that the threat of being raped controls women socially. This social control is present even among very young females. In response to the survey questions, "Would you rather be a man or woman?" and "Why?" a 10-year old girl replied that she would rather be a man, because women get killed and raped (Renzetti and Curran 1995, p. 338). Research indicates that rape is the crime women fear most, and this fear is compounded by the possibility of getting pregnant and contracting AIDS. Brownmiller's book is a historical account of rape that expands on the ideas of Millet and Griffin. The writings of these women clearly demonstrated that rape is more a sociological than a psychological phenomenon.

The definition of rape that served as the basis for most rape laws is grounded in English common law: "the unlawful carnal knowledge of a woman by force and against her will." Implicit in this idea is the assumption that the assailant is a man. In the 1970s, many states redefined rape laws to more comprehensively describe the behaviors that constituted rape and define the age of consent. By the end of the decade, forty-one states had passed some form of rape shield laws that limited the use in court of victims' prior sexual conduct with persons other than the offender (Green 1988, pp. 16-40). Many states have abandoned the concept of rape in favor of the more gender-neutral concept of sexual assault.

The Federal Bureau of Investigation (FBI) indicates that a rape happens every five minutes in the United States. American women are eight times more likely to be raped than are European women and twenty-six times more likely than Japanese women. In the 1980s, the rate of rape in the United States rose four times faster than did the total crime rate. While the number of rapes appears to have fallen over the last few years, rape is still the most frequently committed and least reported violent crime. A total of 97,464 forcible rapes were reported to law enforcement bodies in 1995, the lowest number since 1989. Based on this statistic, it is estimated that in 1995, 72 of every 100,000 females were reported rape victims (U.S. Department of Justice 1995). However, a study funded by the U.S. Department of Health and Human Services found that 683,000 women had been raped in 1990, a figure more than five times as high as the 130,260 reported to the police in that year (Johnston 1992).

Victimization studies such as the National Crime Survey (NCS) and other research projects were initiated to get better estimates of the prevalence of rape. These studies show that the amount of rape is greatly underestimated, although prevalence rates vary from study to study and are hard to reconcile because of variations in research design, sampling, and geographic location.

Rape is described in a number of ways. Sometimes it is discussed in terms of the number of

offenders per victim: the single-offender, two-offender, or multiple, group, or gang rape. Victimization statistics for these forms of rape are, respectively, 81 percent, 10 percent, and 8 percent (Koss and Harvey 1987, p. 10). Rape also is classified as stranger rape, in which the victim and offender have no relationship to one another, or acquaintance rape, which includes date rape and rape between individuals who knew each other before the assault. According to the U.S. Department of Justice, about 55 percent of rapes are acquaintance rapes; the younger the victim is, the more likely it is that she knew the rapist. Acquaintance rapes are especially common on college campuses. Warshaw (1988, chap. 1), citing statistics from the *Ms.* magazine Project on Campus Assault, reports fewer stranger (16 percent) and group rapes (15 percent), with the vast majority of incidents (95 percent) being individual assaults that involve acquaintances or dates (84 percent). The NCS's rape statistics indicate that 27 percent of rapes involve multiple offenders. While acquaintance rapes are more common, women raped by strangers are ten times more likely to report the incident (Renzetti and Curran 1995, p. 341).

Many researchers have attempted to classify rapists into types. Some typologies suggest as few as two or three types (e.g., Groth 1979), while others describe five (e.g., Rada 1978). A review of this literature leads to the conclusion that the typologies are tied closely to the developer's theoretical orientation and educational background.

Rape legally occurs when a person uses force or the threat of force to engage in sexual intercourse (vaginal, oral, or anal) with another person. This seems like a straightforward definition that may lead to the conclusion that rape cases are easy to prosecute. However, one of the most common defenses in rape cases involves the issue of consent. For example, if the victim was drinking, how can she be sure she did not consent? Many rapists have learned that the "alcohol excuse" can help them escape rape charges. A more insidious drug, rohypnol, is being used on college campuses and in bars and nightclubs to facilitate raping and getting away with it. It is dropped into a potential victim's drink. The symptoms are dizziness,

disorientation, lack of coordination, and passing out. The victim is then raped but has no memory of the sexual assault. The question again is posed, If she cannot remember, how does she know she did not consent?

According to the Center for Women Policy Studies (1991, pp. 3-4), fewer than 40 percent of reported rapes result in charges against the offenders and only 3 percent result in convictions. Many researchers have commented on the comparatively light sentences associated with rape convictions. There appear to be several reasons for this, including the tendency to blame the victim rather than the perpetrator: "She should have been more cautious or more sensible or not been walking alone at night." Closely associated with blaming the victim is the notion of "victim precipitation," or the idea that the victim did something to provoke the rape. This can be tied to how the victim was dressed, her profession (e.g., a prostitute), or where she was when the rape happened (e.g., in a fraternity member's room or a male's hotel room, as in the Mike Tyson rape case). Many researchers who have studied rape and the criminal justice system argue that women who are rape victims are "double victims." They are victimized by both the rapist and the criminal justice system.

#### OTHER TYPES OF RAPE AS SEXUAL VIOLENCE

**Marital Rape.** Most discussions of rape focus on date rape or acquaintance rape and stranger rape. A third type of rape among intimates that has received increased attention since the 1980s is marital rape. This type of rape occurs when the victim and offender are spouses or are living in a spouselike arrangement. Although this was thought to be rare, Russell (1984, p. 59) found that 8 percent of ever-married women reported being raped by their husbands. Some researchers believe that marital rape is more common than all other types combined. The 1980s brought about many changes, as spousal immunity laws, which historically prevented husbands from being charged with raping their wives because a wife was viewed as a husband's property, were challenged.

In 1977, Oregon repealed the marital exemption to its rape statute, and in 1979, James K. Chretien became the first person in the United States to be convicted of marital rape. Currently all states prohibit forced sex between a husband and a wife, although most include numerous spousal exemptions. As of 1990, seventeen states and the District of Columbia allowed the prosecution of husbands for raping their wives without exemptions. In twenty-six states, husbands are exempt from prosecution in certain circumstances, such as if only force is employed (Wallace 1999, p. 509) but there is no additional violence such as the threat of using a weapon. In eight states (Kentucky, Missouri, New Mexico, North Carolina, Oklahoma, South Carolina, South Dakota, and Utah), a husband cannot be prosecuted for raping his wife unless they are living apart, are legally separated, or have filed for divorce. In five states, exemptions are extended to unmarried cohabiting couples, and there is an exemption for dating relationships in Delaware.

**Males as Victims of Sexual Violence.** Males appear to be victimized both in and out of prison by heterosexual and homosexual males and by females. Currently, men are thought to make up less than 10 percent of all rape victims (Renzetti and Curran 1995, p. 39.) In *Rape in Prison* (1975), Scacco emphasized that rape in prisons is not an act of homosexuality. Rather, it is an extension of the traditional male sex role and the patriarchal culture in which a single male or group of males seeks to dominate another person through a violent act. In the absence of females, those who are powerful, dominant conquerors are viewed as the “real” or “masculine” men and those who are raped are relegated to a lower status. As Scacco puts it, the victims of rape become “a punk, a queen, or a female.” They have, in the words of the rapists, “had their manhood taken” (1975, p. 52). Scacco’s interpretation of male rape in prison fits into most feminist theories of rape: Rape is an act of power and control rather than a sexual act. While the prevalence rates for the rape of males in prisons appear to vary from one institution to another, it seems safe to conclude that these rates are higher in prisons than outside prisons, where women are available targets.

**Rape Outside Prisons.** Outside the prison environment, most people think rape happens only to women. Scacco (1975) argues, however, that rape would not exist in prison if it did not exist first on the streets. A survey of literature on males raping males indicates that the overwhelming majority of rapists are heterosexual both in and out of prison. However, as more is learned about sexual assaults of males, rape is found to be part of gay men’s relationships as well as involving gay men raping women and other men. If, as most rape theorists agree, rape is a crime of power and not sex, there is no reason to believe that some homosexual males do not commit rape. Russell (1984, p. 74) argues that the insignificant number of females raping males (less than 1 percent) has received more attention than has rape by homosexual men. She notes that this politically controversial issue has resulted in an avoidance of even discussing the issue (1984, p. 74).

While there is no agreement about the prevalence of males being raped, there is a consensus that men get raped by other men and women, that homosexual and heterosexual males are rape victims and offenders, that rape happens in all parts of society and not just in prisons, and that men are less likely than women to report rape. It is known that male victims share some of the traumas of female victims as well as having special concerns. Those added concerns include issues of masculinity and/or sexuality, medical procedures, reporting to the police, telling others, and finding resources and support.

**Rape and War.** Brownmiller’s (1975, chap. 3) historical account of rape provides evidence that rape has always been a part of war. This applies whether one analyzes biblical accounts or the raping of women by American soldiers during the Vietnam War.

There is much documentation that the rape of women and even children is not, as some military and government officials argue, an unfortunate but inevitable part of war. Instead, Brownmiller argues, it is a planned part of war strategy that has included not only rape during the conquest of a village, town, or city but women being captured

and used as sexual slaves. She contends that it is the winning side or the conquerors in battle who do the raping. Rape is not only a way of measuring victory but also the way males assert their masculinity. Consistent with the idea that women are property, women are seen as tangible rewards of war.

Rape during war is considered a criminal act under international law, punishable by imprisonment or death, but wartime rape continues. This violence against women generated public attention as a human rights issue when, in 1993, the U.S. and European media reported the systematic rape, sexual enslavement, torture, and murder of Bosnian Muslim women and children by Serbian military forces in the former Yugoslavia. The estimate of the number of rapes in 1993 was 20,000 (Riding 1993). In 1999, the rape of women and children by the Serb military continued. Investigations of these atrocities conclude, as did Brownmiller, that rape is used as a weapon of war. They argue that it is a central part of the Serbian ethnic cleansing campaign. Bosnian and Kosovar men are murdered, while the women are raped with the objective of impregnating them to produce offspring with Serbian genetic characteristics. It is also psychological warfare, with the intent to demoralize and terrorize Muslims and drive them from their homes. European investigators reported that many women and children died during sadistic and brutal rapes. Brownmiller's historical account of rape and war indicates that an additional consequence of rape and impregnation is suicide and infanticide.

Rape during war often takes the form of group or gang rape. There are documented cases of gang rape by police and military personnel in numerous countries, including Haiti, Honduras, El Salvador, and Iran (Sontag 1993). There are also recent accounts of Algerian women being raped by Islamic radicals. Muslim women in Bosnia, Kosovo, and Algeria who survive rape face psychological problems associated not only with rape but with living in a culture that values sexual purity, by which the male's honor as well as the family's honor is defined.

The motivations for rape and violence against women in war are similar to those for rape be-

tween strangers, acquaintances, and intimates in everyday society and prisons. It is an attempt to socially control and harm individuals who are viewed as having lower or undesirable statuses or are viewed as the enemy. In war, the rapes occur on a larger scale.

In the 1990s, the United Nations attempted to develop policies to eliminate all forms of discrimination against women. However, as long as rape of women and children remains an open or hidden part of war strategy, this violent act will continue.

### EXPLANATIONS FOR RAPE

Explanations for rape come from a variety of disciplines, although many explanations have been dominated by the psychiatric perspective and the medical model (Koss and Leonard 1984, pp. 213–232). The basic assumption underlying this thinking is that rapists are psychologically sick and that rape is sexually motivated. Hence, many explanations focus on developing profiles of rapists (e.g., Groth 1975), leading to a body of research whose validity is questioned because it is based primarily on nonrepresentative clinical or prison populations.

Sociological explanations of rape focus on the social dynamics of society that promote, treat, or proscribe such behaviors and include sociocultural, sociohistorical, and sociopsychological analyses. While these types of explanations lagged behind those of psychology and biology, they offer an alternative to individual-based theories.

Sociocultural explanations of rape are supported by large cross-cultural studies of tribal societies (Sanday 1981) and large studies in nonindustrialized nations. All these studies draw similar conclusions. Rape-prone societies endorse the macho personality and a fundamental belief in the inferiority of females, a belief system that incorporates an acceptance of physical aggression, a high amount of risk taking, and a casual attitude toward sex. Rape also is related to a culture's socioeconomic structure.

Among industrialized societies, the United States has high rape rates. Baron and Straus's

(1989, p. 180) state-level analyses show that rape rates vary by region and state within the United States and that those rates are positively related to the amount of social disorganization, sex magazine circulation (pornography), and gender inequality in a state or region. This adds weight to earlier research that ties rape-specific myths and attitudes to a larger attitudinal construct supportive of sex role stereotyping, violence against women, and adversarial sexual beliefs.

Sociocultural and sociopsychological explanations of rape support the sociohistorical analysis of Brownmiller (1975). While Brownmiller has been criticized for exaggerating the notion of male intimidation, sociological and anthropological research indicates that rape is a socially created phenomenon that is much more likely to occur in cultures that support patriarchy and violence. Sociocultural explanations also may explain why the majority of offenders are male irrespective of whether the victims are male or female, adult or child.

Regardless of the varied theoretical perspectives, most researchers agree that multifactorial (psychological and sociological factors), compared with single-factor, explanations are better predictors of rape. Russell (1984, p. 111) notes that most rape theories could fit into Finkelhor's (1984) four-factor explanation of child sexual abuse, which is discussed in a later section.

### SEXUAL HARASSMENT

Sexual harassment is a form of violence against both males and females, but the overwhelming majority of victims are women and most of the perpetrators are men. While the exact prevalence seems to vary from one study to another, it is safe to conclude that sexual harassment in the workplace is widespread. While sexual harassment can occur in other social institutions, such as colleges and schools where professors and teachers use their positions of power to obtain sexual favors from students, most sexual harassment research has focused on the workplace.

Sexual harassment policies, laws, and rulings generally have relied on Title VII of the Civil

Rights Act of 1964, which prohibits employment discrimination on the basis of race, color, religion, sex, pregnancy, or national origin. In 1980, the Equal Employment Opportunity Commission (EEOC) adopted guidelines that specifically address discrimination on the basis of sex. While these guidelines were not binding, they were used by courts in making decisions in sexual harassment cases. In 1993, the Supreme Court further articulated the meaning of sexual harassment in *Harris v. Forklift Systems, Inc.* Basically, the case focused on the "nature" of sexual harassment and ruled that there was no requirement that the severity of the harassing conduct cause the victim psychological or physical harm. The Court held that victims alleging sexual harassment met their burden of proof if they proved that the environment was *perceived* as hostile or abusive. The circumstances for making such a judgment could include the frequency of the alleged harassment, whether it was physically threatening or humiliating, and the extent to which it interfered with an employee's work performance.

Sexual harassment may take two forms that are not mutually exclusive: quid pro quo harassment and hostile environment harassment. The former type occurs when the victim is placed in a situation where an employer or supervisor uses his or her position to request sexual favors as a basis for continued employment or job benefits. The victim is put in the position of choosing between providing sex and losing employment or benefits. Hostile environment harassment happens when victims are exposed to a series of unwelcome sexual acts that result in psychological harm or humiliation. The difference between the two types of harassment is that the quid pro quo form focuses on harassment that is tied to economic disadvantage, while the hostile environment type focuses on psychological harm.

Tangri et al. (1982) offer several models to explain sexual harassment: the biological, organizational, and sociocultural models. The biological model assumes that sexual behavior in the workplace is an extension of human sexuality or strong sex drives. The organizational model focuses on the situation in which the workplace offers oppor-

tunities for sexual aggression. This model focuses on power that derives from formal roles in an organization. The sociocultural model is derived from theories of patriarchy in which men are viewed as dominant in Western culture and are the gatekeepers of political and economic power. This explanation is the most congruent with a feminist view of sexual harassment.

Two prominent sexual harassment cases have captured the public's attention. The first involved Anita Hill and Clarence Thomas. Hill, a law professor, testified to the U.S. Congress that Thomas, who had been nominated to the Supreme Court, had sexually harassed her. The second was the case of *Paula Jones v. William J. Clinton*. Jones brought charges against President Clinton, alleging that when he was governor of Arkansas and she was an employee of that state, he asked her to engage in sexual activities. Her reluctance, she argued, led to the loss of job benefits. Both cases demonstrated the difficulties of proving sexual harassment cases. They also sent messages that bringing charges against very powerful males can lead to additional reputational, economic, and psychological victimization.

Simply having agencies and courts establish procedures and policies to process sexual harassment will not terminate this behavior. The public must send a clear message that it stands behind sexual harassment laws and that everyone, regardless of social status, must be held accountable. Accounts of sexual harassment, rape, or sexual exploitation by President Clinton of Gennifer Flowers, Paula Jones, Monica Lewinsky, Kathleen Willey, Juanita Broaddrick, and others raise questions about the public's concern about these issues as well as the issue of a double standard. The public seems willing to support or excuse some offenders, such as President Clinton, while punishing others.

### **CHILD SEXUAL ABUSE**

Since the mid-1970s, there has been increased interest in understanding child sexual abuse. This concern led to the founding of the National Cen-

ter on Child Abuse and Neglect and the passage of the 1977 Protection of Children against Sexual Exploitation Act. A review of the child sex abuse research (Finkelhor et al. 1986) demonstrates many similarities with what is known about rape. Paralleling Brownmiller's (1975) historical analysis of rape, for example, is Rush's (1980) analysis of child sexual abuse. Like Brownmiller, Rush elevates child sexual abuse to the level of a social problem by tracing its roots to patriarchal societies and their social institutions, belief systems, and myths. However, there are groups that do not view child-adult sex as a problem, such as the North American Man Boy Love Association; in fact, they advocate it.

In general, child sexual abuse can be defined as sexual exploitation of or sexual activities with a child by an adult, in which the child's health or welfare may be harmed or threatened. While all states consider child sexual abuse a crime, definitions vary from state to state and some states leave the interpretation to the courts. There is no real agreement among social scientists about the true extent of child sexual abuse, but all concede that the number of reported cases is increasing and that the actual numbers are underestimated, particularly for males. A comparison of national findings with less comprehensive research studies reveals conflicting estimates of the incidence of child sexual abuse. Commonly cited estimates suggest that one in four girls and one in six to seven boys will be sexually abused by the time they reach age 18.

Studies do not support a general social characteristics profile of the victims or offenders, although females tend to be at somewhat greater risk of victimization than males and males are overwhelmingly the offenders. The ages of greatest risk for victimization appear to be between 4 and 9. Most offenders are either known by or related to their victims and are not old. Only about 10 percent of child sexual abusers are strangers, although the Internet could change this statistic.

The impact of being a victim of child sexual abuse has been widely studied. It generally is agreed that the amount of trauma experienced varies with the type of abuse, how the offender is related to

the child, how long the abuse lasted, how sexually intrusive the abuse was, the age when the abuse began, the reactions of others to the disclosure of the abuse, and the child's personality.

As an ever-increasing number of child sexual abuse cases have moved into the courts, particularly in cases involving divorce and custody, concern for balancing the needs of the legal system against those of children has arisen. Placing children in the courtroom subjects them to many of the same problems faced by rape victims: victim blaming and other courtroom-produced traumas. Although still in the early stages, reforms are being implemented and suggested, with one area of legislative reform focusing on increasing convictions and community notification once sex offenders are released. The latter was largely an outgrowth of Megan's Law, which was named after Megan Kanka, a 7-year-old who was molested and killed in 1994 by a released sex offender who was living in her neighborhood without residents' knowledge of his history of sexually abusing children. By 1995, forty-three states had adopted laws requiring offenders to register with a law enforcement agency when they were released.

The late 1980s and early 1990s ushered in a new area of child sexual abuse. Children 12 years of age and younger have sexually abused other children. These young offenders have predatory patterns that are very similar to those of adult and adolescent molesters. Some, but not all, are reacting to their own victimization. An interesting finding from this research is that females tend to be about as likely to be perpetrators as are males and are equally aggressive in their sexual acts. Questions have been raised about why females are not found in larger numbers among the adolescent and adult sex offender populations. The family system appears to be the learning and training ground for these youthful (some as young as 2 years old) perpetrators (Araji 1997).

Studies of child sexual abuse do not offer a consistent link between social class, race, and ethnic group, although the lower socioeconomic classes tend to be overrepresented, as are, in the case of young perpetrators, single-parent families. It is too

early to make accurate cross-national comparisons of the extent of child sexual abuse. Some research has been conducted in other countries and Finkelhor (1994) suggests that, as in the United States, the problem is widespread. There have been no cross-cultural studies of children who sexually abuse other children.

### **EXPLANATIONS OF CHILD SEXUAL ABUSE**

The development of child sexual abuse explanations has been similar to that of explanations for rape, as most early studies focused on offenders' psychopathological or biological motivations or attempted to develop profiles of child molesters, victims, and families. As with rape, most explanations offered only single-factor explanations, an approach criticized by Araji Finkelhor (1986, chap. 34). As an alternative, those authors proposed a four-factor explanation of child sexual abuse that includes the categories of emotional congruence, sexual arousal, blockage, and disinhibition. Respectively, these factors incorporate explanations of why an adult would have an emotional need to relate to a child, could be sexually aroused by a child, would not have alternative sources of gratification, and would not be deterred from such an interest by normal prohibitions (Araji and Finkelhor 1986, p. 117). The model includes both individual (e.g., arrested emotional development) and socio-cultural (pornography) factors. As was noted above, Russell (1984) believes that this model could be adapted to explain rape.

With respect to explaining youthful perpetrators' sexual aggression, Araji (1997) found much of the same theoretical history noted above. She proposes systems theory as a necessary explanatory and remedial guide.

### **INCEST**

Most researchers and practitioners consider incest different from child sexual abuse, with incest viewed as intrafamilial and child sexual abuse viewed as extrafamilial sexual relations. While some argue that the two types of abuse have much in common,

one of the most significant differences is that the victims of incest are always betrayed by someone who has been charged with loving and protecting them. Much of the sexual abuse committed by children 12 years of age and younger is intrafamilial. Such experiences can have extremely destructive results, but this varies with the factors discussed above.

The most common type of incest researched and written about is between father and daughter (Herman and Hirschman 1981), although the most common type may be brother-sister. As stepfamilies are formed, the probability for incestuous relationships involving stepfathers increases. Estimates of incest range from a high of 38 percent of females in Russell's (1984) study to about 25 percent in several other survey studies (e.g., Finkelhor 1979). However, the same research problems surrounding rape and child sexual abuse apply to incest literature: underestimating the extent of the crime, not accurately defining it, and theory and sampling problems.

#### **TREATMENT AND PREVENTION**

A considerable body of literature on the prevention and treatment of sexual violence and abuse has emerged from applied disciplines such as social work and clinical psychology as well as from community and feminist organizations. An extensive discussion is beyond the scope of this article, and so only a few generalizations are noted. First, as treatment programs are informed by a variety of theories of rape and sexual abuse, it is not surprising that there is an array of treatment programs. With respect to offenders, most programs focus on punishment and/or rehabilitation (Groth 1979; Conte and Berliner 1981), and many are based on the psychiatric or medical model. In cases of child sexual abuse and incest, programs typically focus on helping the victims and their families get through the crisis. In the area of rape, programs have been aimed primarily at treating victims and families for the rape trauma syndrome and have been directed at females. Treatment programs have also been developed to deal with the short- and long-term effects of abuse (Brown and Finkelhor 1986) and can be either clinical or community-based. Some

are beginning to focus on sexual and ethnic differences. At present, however, there remain debates among professionals and practitioners about what the appropriate treatments are, which ones work, and how well.

Stemming from myths and callous societal attitudes toward sexual aggression and the extensive involvement of feminists, prevention programs have followed a victim advocacy model (Araji 1989, chap. 17). Under this concept, potential victims are taught in various ways how to protect themselves from becoming victims. Some newer programs are aimed at men and promote the development of a nurturant rather than a macho male image.

#### **SUMMARY AND CONCLUSIONS**

Extensive public and professional attention to sexual violence and exploitation has come about only since 1960 and has been primarily a grassroots movement. As a result, a large body of cross-disciplinary literature has developed and various reforms have taken place. While no consensus exists about the scope of these problems, the rates at which they are increasing or decreasing, definitions, explanations, social policies, and solutions, sociologists, particularly feminists, have concluded that sexual violence and exploitation will not be reduced or eliminated until societies stop supporting and/or promoting aggression, violence, exploitation, and inequality.

However, the prognosis that sexual violence and exploitation involving women and children will soon end is not convincing. As long as societal members, many of them women, are willing to turn a blind eye to predatory behaviors toward women on the part of the president of the United States and surveys of college, high school, and elementary students report large percentages (31 to 87 percent) of students believing that it is acceptable for males to engage in forced sex or rape of females under certain circumstances (White and Humphrey 1991) or if they do not get caught (Malamuth 1981), one must conclude that the fight against sexual violence and exploitation of women and children is still in the beginning stages.



## REFERENCES

- Araji, S. 1989 "The Effects of Advocates on Prevention." In N. C. Barker, ed., *Child Abuse and Neglect: An Interdisciplinary Method of Treatment*. Dubuque, Iowa: Kendall/Hunt.
- 1997 *Sexually Aggressive Children: Coming to Understand Them*. Thousand Oaks, Calif.: Sage.
- , and D. Finkelhor 1986 "Abusers: A Review of the Research." In D. Finkelhor, S. Araji, L. Baron, A. Brown, S. Doyle Peters, and G. E. Wyatt, eds., *A Sourcebook on Child Sexual Abuse*. Beverly Hills, Calif.: Sage.
- Baron, L., and M. Straus 1989 *Four Theories of Rape in American Society*. New Haven, Conn.: Yale University Press.
- Brown, A., and D. Finkelhor 1986 "Impact of Child Sexual Abuse: A Review of the Research." *Psychological Bulletin* 99:16-77.
- Brownmiller, S. 1975 *Against Our Will: Men, Women, and Rape*. New York: Simon & Schuster.
- Center for Women Policy Studies 1991 *More Harm Than Help: The Ramifications of Mandatory HIV Testing of Rapists*. Washington, D.C.: Center for Women Policy Studies.
- Conte, J. R., and L. Berliner 1981 "Prosecution of the Offender in Cases of Sexual Assault against Children." *Victimology* 6:102-109.
- Finkelhor, D. 1994 "The international epidemiology of child sexual abuse." *Child Abuse & Neglect*, 18, 409-417.
- 1979 *Sexually Victimized Children*. New York: Free Press.
- 1984 *Child Sexual Abuse: New Theory and Research*. New York: Free Press.
- , S. Araji, L. Baron, A. Brown, S. Doyle Peters, and G. E. Wyatt 1986 *A Sourcebook on Child Sexual Abuse*. Beverly Hills, Calif.: Sage.
- Green, W. M. 1988 *Rape*. Lexington, Mass.: Lexington Books.
- Griffin, S. 1971 "Rape: The All-American Crime." *Ramparts*, September.
- Groth, N. 1979 *Men Who Rape*. New York: Plenum.
- Herman, J. L., and L. Hirschman 1981 *Father-Daughter Incest*. Cambridge, Mass.: Harvard University Press.
- Johnston, D. 1992 "Survey Shows Number of Rapes Far Higher Than Official Figures." *New York Times* April 24, p. A14.
- Koss, M., and M. Harvey 1987 *The Rape Victim: Clinical and Community Approaches to Treatment*. Lexington, Mass.: Stephan Green Press.
- Koss, M. P., and K. E. Leonard 1984 "Sexually Aggressive Men: A Review of Empirical Findings." In N. Malamuth and E. Donnerstein, eds., *Pornography and Sexual Aggression*. New York: Academic Press.
- Malamuth, N. M. 1981 "Rape Proclivity among Males." *Journal of Social Issues* 37:138-157.
- Millett, K. (1970). *Sexual Politics*. New York: Doubleday.
- Rada, R. T., Ed. 1978 *Clinical Aspects of the Rapist*. New York: Grune & Stratton.
- Renzetti, C. M., and D. J. Curran 1995 *Women, Men and Society*. Boston: Allyn and Bacon.
- Riding, A. 1993 "European Inquiry Says Serbs' Forces Have Raped 20,000," *New York Times*, January 1, p. 4.
- Rush, F. 1980 *The Best Kept Secret: Sexual Abuse of Children*. New York: McGraw-Hill.
- Russell, D. 1984 *Sexual Exploitation, Rape, Child Sexual Abuse, and Workplace Harassment*. Beverly Hills, Calif.: Sage.
- Sanday, P. R. 1981 "The Socio-cultural Context of Rape: A Cross-cultural Study." *Journal of Social Issues* 37:5-27.
- Scacco, A. M., Jr. 1975 *Rape in Prison*. Springfield, Ill.: Charles C Thomas
- Sontag, D. 1993 "Women Asking U.S. Asylum Expand Definition of Abuse." *New York Times* September, 27 pp. A1, A13.
- Tangri, S., M. Burt, and L. Johnson 1982 "Sexual Harassment at Work: Three Explanatory Models." *Journal of Social Issues* 38: 33-54.
- U.S. Department of Justice, Federal Bureau of Investigation 1995 *Crime in the United States: Uniform Crime Reports, 1995*. Washington, D.C.: U.S. Government Printing Office.
- Wallace, H. 1999 *Family Violence: Legal, Medical and Social Perspectives*. Boston: Allyn and Bacon.
- Warshaw, R. 1988 *I Never Called It Rape*. New York: Harper & Row.
- White, J. W., and J. A. Humphrey 1991 "Young People's Attitudes toward Acquaintance Rape." In A. Parrot, ed., *Acquaintance Rape: The Hidden Crime*. New York: Wiley.

SHARON K. ARAJI

## SEXUALITY

*See* Sexual Behavior Patterns; Sex Differences; Sexual Orientation.

## SEXUALLY TRANSMITTED DISEASES

Until the 1980s, social science research on sexually transmitted diseases (STDs) focused primarily on the history of various pestilences, the epidemiology of those diseases, and the description of mass disasters (Brandt 1985; Aral and Holmes 1999). The disease that was most commonly researched was syphilis, long the best identified and most feared STD. Historians and anthropologists wrote numerous treatises on its origin and the social consequences of its introduction into isolated, tribal, or third world societies (Wood 1978; Hart 1978). In the early 1980s, the consequences of other STDs were studied, especially as sequelae of prostitution (Kalm 1985; Poherat et al. 1981).

When awareness of the “sexual revolution” finally induced social scientists and epidemiologists to think about the effects of STDs on less sexually active populations than prostitutes and their clients, the literature turned to the newly sexually active: vulnerable teenagers (Washington et al. 1985) and other young people who engaged in premarital sex (O’Reilly and Aral 1985). The exponential intensification of the discussion of and social science research on STDs, however, came only after the medical community’s horrified acknowledgment that the newest STD to become epidemiologically important, the acquired immune deficiency syndrome (AIDS), was also the deadliest and that social, not just biological, information was essential in order to combat it.

When AIDS first began to be discussed in 1981, the medical community was already alarmed at its mystery and virulence (Aral and Holmes 1999). Unhappily, it took years before more focused measures, such as those taken by fund-raising organizations and institutes devoted to AIDS research, were initiated. Journalists, most notably Randy Shilts (1987), have persuasively ar-

gued that the lack of a strong reaction from the start was due to the fact that the early victims were gay men, not American Legionnaires or Girl Scouts. Today the basic facts of AIDS have been well disseminated. Almost everyone knows, for example, that in the United States gay men are disproportionately affected, as are people who mix blood while exchanging hypodermic needles. However, in the early 1980s, information was abysmally inadequate and myth and rumor educated more people than did social or medical research.

With the accuracy of hindsight, it is evident that sociologists should have looked at sociocultural histories of recent sexual behavior and used that information to help study disease transmission. But no one was proactive, and it took years for a pertinent literature to emerge. The exceptions were a small group of social researchers at the Centers for Disease Control (CDC), whose work was restricted to STD-related topics, and a few epidemiologists studying the social location of this disease among gay American men (Aral and Holmes 1999) and among heterosexuals in Africa. Otherwise, the analysis of AIDS remained mostly ghettoized in the medical literature until the mid-1980s. Until the publication in 1994 of the results from a large national survey of adult sexuality in the United States (Laumann et al. 1994), the best available data on sexual behavior were the Kinsey studies from the 1940s and early 1950s.

Finally, the combination of organized gay activism and public alarm created the kind of political pressure that made more money available and launched a flood of AIDS research. Indeed, interest in sex research in general, previously an area treated like a poor relation, received more credibility, though not enough to allow the funding of a nationally representative study of sexual behaviors. A large national study slated to be funded by the National Institutes of Health was stopped in 1990 after Senator Jesse Helms helped persuade the U.S. Senate that Americans should not be exposed to such questions. Nonetheless, research on sexuality, especially on STDs, found funds and larger and more diverse professional audiences. Even before the prevalence of AIDS among gay men was understood or publicized, some research

described how great numbers of anonymous sexual contacts in gay bars, baths, and parks happened and how such activity set the stage for infection (Darrow 1979; Ross 1984; Klovdahl 1985). By the 1990s research attention had become focused on how sexual cultures, such as a "gay lifestyle," increased exposure to the human immunodeficiency virus (HIV). Later, the issue was reframed to focus on behaviors that place people at risk (rather than on the groups to which they belong) and the social and psychological factors that influence sexual decision making. This avenue of research affirms that sex is a social behavior that must be studied in its social context.

Current AIDS literature centers on two main issues: (1) who is at risk and why and (2) risk-reduction factors, including education.

#### **WHO IN THE UNITED STATES IS AT RISK AND WHY**

Most researchers include in their lists of risk factors number of partners, sex of partners, intravenous (IV) drug use or an IV-drug-using partner (Ehrhardt et al. 1995), frequency of intercourse (Aral and Cates 1989), use of condoms (Morrison et al. 1995), contact with commercial sex workers (Plummer et al. 1999), and sex with bisexual men (Doll and Ostrow 1999). The last factor has been of increasing interest, since there seems to be more bisexuality than mainstream research acknowledged previously and because this is an obvious bridge between high- and low-HIV-rate populations. Information about bisexuality has become important in efforts to understand AIDS transmission. For example, a study of lesbian women, a group usually thought of as low-risk, showed that not only had 81 percent of these women had sex with men, at least one-third of their male partners had had sex with other men. Women with bisexual male partners were also more likely to have had anal sex, an activity thought to be an especially efficient mode of HIV transmission (Padian et al. 1987). Bisexuality among men constitutes a greater risk factor than was previously recognized because these men may make regular forays into the gay male world unknown to their female partners.

This may be more likely to occur among married couples or in some minority communities, where the behavior itself may necessitate the utmost secrecy and may even be defined by the participants as "not homosexual" and therefore not risky (Carrier 1985; Blumstein and Schwartz 1977; Humphreys, 1975).

Other, less obvious risks include the possibility of deviousness or outright lies from a partner. In a poll conducted by Cochran and Mays (1990), 196 men and 226 women aged 18–25 completed an anonymous questionnaire on sexual strategies. The findings indicated that a significant number of both men and women had told a lie in order to have sex. Men lied more frequently than did women, but both sexes were actively and passively willing to deceive a date.

Sociodemographic characteristics also have been studied as risk factors. Age is one such factor. Among U.S. teenagers the rate of AIDS is low overall (less than 1 percent of AIDS cases); however, such data underestimate the risk because of the long incubation period from HIV infection to the development of AIDS (ten or more years). Moreover, AIDS rates are higher among more vulnerable subpopulations of adolescents, such as runaway and homeless youth, STD clinic populations, and young people in the juvenile justice system. Such youths are more likely to engage in activities, such as drug use and "survival sex," that put them at high risk of HIV infection. Particularly worrisome is the high rate of other STDs among teenagers, especially young women, because having an STD is thought to enhance the probability of becoming infected with HIV after exposure. Unfortunately, although condom use has increased among teenagers, adolescents still do not use condoms consistently (Sonenstein et al. 1998.). Teenagers in the United States have higher rates of nonmarital pregnancies than do their counterparts in any other industrialized nation despite the fact that they are no more sexually active than are teens in other countries. These facts suggest that teenagers in the United States are not adequately protecting themselves against the outcomes of their sexuality. This may be the result of a lack of comfort with the idea of teenage sexuality in a nation that sends mixed messages to its youth.

In general, better-educated persons are more likely to use condoms. College students, however, are surprisingly casual about condom use. In a study by Reinish et al. (1990), less than two-thirds of the students studied had used a condom in the previous year, less than one-third had used a condom the last time they had had vaginal or anal intercourse, and only half had ever used contraceptive methods that also prevent STD transmission. Those in exclusive sexual relationships reported the highest levels of intercourse. This finding prompted concern among the researchers that while it seems that being in a committed sexual relationship lowers the overall risk of HIV infection by reducing the number of partners, risk may be increased because of frequency of relations unless partners use condoms or have accurate information about each other's sexual and drug use histories. Even well-educated college students tend to use criteria irrelevant to AIDS transmission, such as "just knowing" a partner is safe, as a means of determining when condom use is necessary (Civic, in press). A set of qualitative research notes indicates that condom use may decrease, even among populations that most need to use them, for socioemotional reasons. Kane's (1990) population of women with HIV-probable partners in the drug culture refused, as an act of solidarity, to use condoms. These women felt that using a condom would indicate their awareness and condemnation of the partner's addiction, alienating him and harming the relationship.

In the United States, not all racial and ethnic groups have been equally affected by the AIDS epidemic. Although European-Americans initially were the group with the highest rate of AIDs, this is no longer true: AIDS rates among racial and ethnic minority groups have increased substantially since the beginning of the epidemic. Data from the CDC show that in 1990 whites accounted for over half the AIDS cases (56 percent in January 1990), but they now account for fewer than half the cases (46 percent in January 1998). African-Americans in impoverished inner cities have been especially hard hit. In January 1990, 27 percent of AIDS cases occurred among African-Americans, but by January 1998, this figure had increased to 35 percent.

The infection rate also has been increasing among Americans of Hispanic origin (18 percent of cases in January 1998).

One of the newest groups in the United States to receive research attention is women. Since AIDS surfaced in this country among gay men, the lack of attention to women might seem reasonable until one remembers that as partners of bisexual men or drug users, as drug users themselves, or as inhabitants of countries where AIDS is not a "gay disease," women always were at risk. AIDS has been increasing steadily among women in the United States. CDC figures show that in January 1989 women represented only 9 percent of AIDS cases, but by January of 1998 this figure rose to 16 percent, a 44 percent increase in less than a decade. Women also have substantially higher rates of chlamydia—the most prevalent bacterial STD—as well as genital herpes. Unfortunately, STDs are more likely to be asymptomatic in women, thus causing delayed treatment, increased complications such as infertility, and increased vulnerability to HIV infection.

Initially, IV drug use was the primary mode of AIDS transmission among U.S. women, but today, heterosexual intercourse is the primary mode. In late 1989, 52 percent of AIDS cases among women were attributable to IV drug use, whereas in June 1998, only 23 percent of cases were thought to be IV drug use-related. Most women with AIDS are of childbearing age. This increases the chances of transmission from an infected pregnant woman to her unborn baby as well as increasing the number of children, often in single-parent families, who are forced to cope with the prolonged illness and then loss of a mother. The effects of AIDS on children have only recently begun to be studied. Studies show that between 13 and 35 percent of babies born to HIV-positive women are HIV-infected, a figure that has been rapidly declining since the beginning of the AIDS epidemic, largely as a result of the advent of newer antiviral drugs. Many pregnant HIV-infected women, even with knowledge of the risk, choose not to abort (*Proceedings NIMH/NIDA Conference on Women and AIDS* 1989). AIDS concerns attend other reproductive issues as well, such as the safety of artificial insemi-

nation. While the American Fertility Association has guidelines that exclude high-risk men from donating sperm, methods for testing for HIV seem to be inconsistent, and private physicians may not test at all (Campbell 1990).

The use of alcohol and drugs has been shown to be related to unsafe sex practices, although there have been contradictory findings (Leigh and Stall 1993 provide for a review). Fullilove and Fullilove (1989), for example, found that 62 percent of the 222 black inner-city teenagers they studied used crack, and 51 percent of the users said they combined crack use with sex. Forty-one percent of the teenagers surveyed had had at least one STD; those who used crack with sex had a significantly higher rate of STD infection. The correlation between crack use, sex, and STD transmission has been found by many other investigators (Aral and Holmes 1999). Although crack use has declined substantially since the late 1980s, the use of other drugs (e.g., methamphetamines) that also appear to be related to unsafe sex has risen. Leigh (1990) found that patterns of drug use and the effects of drugs on behavior differ among groups: Gay men were less affected by drinking and engaged in more risk taking when using cocaine, whereas heterosexual risk taking was predicted largely by total frequency of sex, with only a small amount of the variance explained by having partners who used drugs or alcohol. A note of caution is warranted in interpreting the findings of studies of the relationship between sexual practices and substance use. Most of this research that shows a relationship between substance use and risky sex has not demonstrated that substance use *causes* risky sex, because researchers rarely have information on which came first. Moreover, it is conceivable that a third factor, such as a propensity to take risks, is responsible for both substance use and sexual risk taking (Leigh and Stall 1993). Clearly, more research is needed.

Research has increasingly concentrated on infection resulting from exchange of blood caused by mutual use of needles during intravenous drug use. In a paper by Freeman et al. (1987), a comparison of gay males and IV drug users showed that peer support helped create safer sexual practices

for gays while lack of social organization reduced IV drug users' chance of self-protection. Among the drug users, 95 percent were well aware of their exposure and 68 percent knew that needle sharing could transmit AIDS. Some individual attempts at decreasing the use of potentially contaminated needles had been made, but the authors felt that the only way to reduce risk in this population was to create organizations for needle dispensation that eventually could create a culture of mutual protection. Opinion has changed from considering IV drug users uneducable to recognizing substantial successes in changing their drug-taking practices to include more self-protective habits. Still, ethnography shows that needle sharing between addicted partners is seen as an intimate and bonding behavior, and this makes change more difficult. Moreover, despite their demonstrated effectiveness, needle-exchange programs are under attack by conservatives who claim that they encourage drug use, even though there is no evidence to support that contention.

Although sexual behavior certainly has biological underpinning, it is clear from the research that sexual behavior and the spread of STDs are socially driven. Sociocultural factors, not biology, determine not only with whom one has sex, whether and when one is likely to have sex, and the specific forms of sexual expression in which one engages but even which persons, activities, and things are experienced as erotic. This fact is eminently evident in the early studies of the Kinsey group, Ford and Beach's (1951) cross-cultural studies, and more recent research (e.g., Laumann et al. 1994).

#### **RISK REDUCTION AND EDUCATION IN THE UNITED STATES**

It is difficult to discuss risk without touching on risk reduction. Indeed, a growing body of research literature reports the results of studies investigating specific risk-reduction strategies, curricula, and behavior modification that targets specific populations—such as gay men, minority group members, teenagers, mothers, and drug users—with what is hoped to be a useful approach.

The most encouraging findings indicate that it is possible to change risky sexual behavior. This is

a significant discovery, since many personal habits, such as drinking and overeating, are notoriously resistant to sustained modification. In the 1980s, researchers in San Francisco found that gay men there had reduced their numbers of partners and frequency of sex and increased their use of safer sex practices (McKusick et al. 1985). However, since much of this change was attributed to the extraordinary social power of organized gay groups in that city, the question arose whether such dramatic changes could occur among gay men who are less embedded in gay communities (Fisher 1988). However, change has occurred across the United States, indicating both the strength of educational and social control efforts among gay activists and, perhaps, the great motivation for change that exists when suffering and death are not only possible but probable (Martin 1987; Siegal and Glassman 1989; Roffman et al. 1998).

This conclusion is not self-evident because change in gay male circles has not been complete. Changes in behavior are associated with proximity to populations with high incidences of the disease: The more distance, the less change in behavior (Fox et al. 1987). Furthermore, even in densely infected areas, a significant minority of HIV-infected persons seem to continue engaging in risky sexual practices (Kelly et al. 1998). There is even evidence that for some men, fear wears off and unprotected sex practices increase (Martin 1987), whereas for others, initial changes toward safer sex are difficult to maintain in the long run (Kelly et al. 1998).

Naturally, there is a great deal of pressure on researchers to find out what helps all kinds of people protect themselves from AIDS. Depressingly but predictably, it has been found education alone is inadequate to induce behavior change. For example, Calabrese et al. (1987) reported that for gay men outside big cities, attendance at a safe-sex lecture, reading a safe-sex brochure, HIV antibody testing, advice from a physician, and counseling were all inadequate. Other sex education efforts have had a limited effect, often for a limited period. For example, researchers assessed the impact of a ten-week university course on human sexuality and AIDS-related behavior. While stu-

dents who had taken the course possessed more information about actual risk, worried about AIDS more, and asked sexual partners more questions relating to AIDS than did a control group, they did not increase their use of condoms or other contraceptives, decrease the number of sexual partners, or spend a longer time getting to know a prospective partner (Baldwin et al. 1990). Studies have consistently shown that sex education for adolescents is necessary but not sufficient to produce behavioral changes (Kirby and Coyle 1997).

The disappointing results of sex education have led researchers to search for more viable strategies. Because self-esteem, confidence, and ego strength have been hypothesized to help individuals protect themselves from others' as well as their own desires, a number of researchers have looked for ways to bolster these characteristics (Becher 1988). To date, the most promising programs appear to be ones that are conceptually based and provide information and skill building, enhance motivation and normative support for change, and are ethnically and culturally sensitive (Fisher and Fisher 1992 provide a review).

No one believes that any single approach is appropriate for all audiences. Increasingly, this literature has been investigating separate strategies for different groups. Students of race, ethnicity, and gender understand not only that various groups use language differently but also that reality is filtered through culture. This sociological truism has benefited research and education among at-risk populations. An example is understanding how gender differences affect health behavior. Campbell (1990) noted the limitations of educational programs aimed at women, especially non-IV drug users, who resist feeling at risk. She found that the partners of IV drug users are unlikely to be assertive and to insist on safer sex. Most of these women are already in subordinate, if not abusive, situations, and their vulnerability and passivity have to be addressed before progress can be made. This fact is underscored by findings from a study in which Beadnell et al. (in press) report that a strong predictor of the extent of at-risk women's attendance at a multi session AIDS prevention intervention is whether they are in an abusive

relationship. Campbell (1990) warns that there are special considerations about condom use among minority group women, since minority group men may reject condoms more resolutely than do white men. Campbell also reminds the reader that in educating commercial sex workers both for their own safety and for that of others, one needs to take into account the dual issues of their gender and their profession. She argues against overreliance on women as the safety net in sexual relations. To date, little AIDS research has focused on heterosexual men.

Among working and lower-income African-American women, gender issues often make safe-sex guidelines seem impossibly theoretical. Unemployment has put African-American men in transient relations with African-American women, and partners are unlikely to engage in the kind of cooperative communication many safe-sex guidelines assume (Fullilove, et al. 1994). Fullilove et al. also highlight black women's and teenagers' increased vulnerability to disease because of relatively high rates of nonmonogamy among potential partners. Those authors feel that individual strategies are unlikely to be as powerful as a "reknitting of community connections" for the evolution of protective norms. Social disorganization further complicates the problem by giving less and less accurate information to African-American and Hispanic populations.

Even designing messages for minorities or finding community outlets for their dissemination does not begin to handle the difficulty of reaching and influencing at-risk persons. Target audiences for prevention are not necessarily self-identified. For example, almost no AIDS-prevention research has been conducted with lesbian women who may have occasional intercourse with bisexual partners and do not consider themselves at risk. Latino men who occasionally visit gay bars and have anal intercourse often do not use condoms with their wives in part because, as the "*activo*," they do not see themselves as homosexual or as having participated in a homosexual act and therefore do not perceive themselves as being at risk (Magana 1990). Similarly, because virginity before marriage is highly valued in Latino culture, young unmarried Latinas

may engage in higher-risk anal sex in order to remain "technical virgins." Given the high value placed on motherhood among Latinas, using condoms during vaginal sex can imply that a man's intentions are less than honorable by eliminating the possibility of motherhood. Almost nothing is known about the sexual practices of Asian-Americans, perhaps because they have been depicted as the "model" minority. However, some data suggest that Asian-American youths are less knowledgeable about AIDS than are other groups (Wells et al. 1995), and a substantial minority of Asian-American adolescents have been shown to engage in unprotected sex (Schuster et al. 1998). Issues of identification, culture, and gender relations bedevil both researchers and health workers.

Despite the optimism of biomedical researchers, there is still no cure for AIDS or an effective vaccine against contracting HIV. Therefore, behavioral change remains the best hope for controlling the epidemic. Sociological and psychological theories have generated a number of programs that appear to be effective in reducing an individual's risk of contracting HIV. Most of these programs have focused on changing individuals' skills, knowledge, attitudes, and behaviors. However, sexual intercourse is inherently dyadic, and partner issues will need greater attention for such programs to achieve maximum effect, especially for women, who often do not have control over their sexuality. Moreover, the advent of new antiviral drugs and protease inhibitors has transformed the character of the AIDS epidemic to the public at large and the researchers as well as to those with HIV infection. Many people in the United States now view AIDS as a manageable chronic illness rather than a death sentence. Persons with AIDS are living longer, mortality rates have decreased, and those infected with HIV do not appear to be developing AIDS at the same rate as was the case before the use of these new drugs. In response to this change, the focus of AIDS research has changed: Researchers are increasingly studying how to prevent or reduce adverse outcomes, such as depression and relapse to unsafe behaviors, and enhance the quality of life among those living with HIV and AIDS (Kelly et al. 1998 provide a cogent discus-

sion). However, the new drugs are not a panacea. For some, they do not work at all; for others, they appear to work initially but then gradually become ineffective; and for far too many, the enormous cost of the drugs is beyond their means. The latter fact is even more tragic when one considers that in the United States most new cases of HIV infection are occurring among society's most vulnerable populations.

### THE INTERNATIONAL SCENE

Data on STDs, including AIDS, from other countries are inadequate because many countries, especially third world nations, do not have sophisticated systems to monitor these diseases. The World Health Organization (WHO) estimates that the current annual prevalence of STDs is about 333 million cases worldwide, with about 5.8 million persons being newly infected with HIV in 1997. In 1999 there were nearly 31 million persons living with HIV or AIDS worldwide. If the current trend continues, WHO estimates that more than 40 million persons will be infected with HIV by the year 2000. To date, over eleven million people worldwide have died of AIDS, and in 1998 alone, about 2.3 million deaths were attributable to AIDS, 46 percent of which occurred among women. The vast majority of cases of STDs worldwide, including AIDS, have occurred in developing countries, and the growth of the AIDS epidemic has been fastest in sub-Saharan Africa, which is thought to have two-thirds of the world's population of persons infected with HIV. Although rates of infection are lower in Asia, where the epidemic started later, the number infected there is estimated to be quite large. In North America, western Europe, Australia and New Zealand, and a few third world countries that have sound economies (e.g., Costa Rica, Thailand), rates of several STDs (e.g., syphilis and gonorrhea) have been declining steadily over the past decade, but they have been skyrocketing (or being reported better) in regions such as China and Russia and remain high in Africa and Asia.

Why is the AIDS epidemic so heavily concentrated in the developing world? Decosas (1996) argues that differences in sexual behaviors alone

cannot explain this fact. He argues that the high incidence of untreated STDs that increase vulnerability to HIV infection, cultural factors such as age differences between male and female sexual partners (older persons typically have more sexual partners and therefore an increased likelihood of exposure to HIV), and demographic factors such as large-scale labor migration and refugee movements have all contributed to the difference. All these factors are related directly or indirectly to poverty and gender inequality. Young women in third world countries, for example, often have limited access to education or vocational training and may have few alternatives for economic survival other than having sex with men. Studies in Africa suggest that women in traditional relationships typically do not engage in risky sexual practices, yet they are increasingly being infected with HIV (Way et al. 1999). Even women who report having only one sexual partner in their lives and who believe that they are in monogamous relationships are found to be infected with HIV. In many regions of Africa, however, men not infrequently engage in extramarital sexual relations, often with higher-risk partners such as commercial sex workers (Way et al. 1999).

Today, the AIDS epidemic worldwide is spread primarily through heterosexual contact. As more women of childbearing age become infected, perinatal transmission from mother to infant is increasing. In urban areas of Uganda, Zambia, Malawi, and parts of Southern Africa, for example, HIV infection among pregnant women has been increasing rapidly; more than one-fifth, and in some areas up to 40 percent, of these women have been found to be HIV-positive (Way et al. 1999). There are important differences across the globe in the main mode of AIDS transmission and who has been infected as well as substantial differences in these factors within regions. In China, for example, there appears to be two epidemics: one among IV drug users in the mountainous regions and the southwestern areas and the other in the more prosperous eastern coastal areas, where commercial sex is reemerging in response to the growing gap between rich and poor. Ironically, AIDS was virtually unknown in China before that country it



opened its borders. In Africa, the highest HIV prevalence rates occur in urban as opposed to rural areas. In Mexico and Latin America, the epidemic is concentrated among the poorest and most marginalized members of society, as is becoming increasingly true of the United States. In eastern Europe, IV drug use accounts for the majority of new infections. Not only are there important differences in the major modes of AIDS transmission worldwide, the modes of transmission within regions have changed over time (Way et al. 1999). This is evident in the United States, where the AIDS epidemic first appeared among gay men, then among IV drug users, and now among racial and ethnic minorities and women. In Latin America, AIDS was first concentrated among IV drug users and gay men, but heterosexual sex is playing an increasing role in its transmission. The epidemic in Asia is moving rapidly from high-risk populations to the general population, largely as a result of heterosexual transmission. In developing nations particularly, new AIDS cases are occurring among younger persons.

The rate of STDs worldwide is alarming because although the most common nonviral STDs (syphilis, gonorrhea, chlamydia, and trichomoniasis) are curable, they are often initially asymptomatic. This is especially so for women, who also are more likely than men to become infected if exposed. Moreover, the increased mobility of populations, accompanied by urbanization, poverty, sexual exploitation of women, and changes in sexual behaviors, places an increasing proportion of the world's population at risk. Because younger people make up a much larger proportion of the population in developing nations than they do in industrialized ones and because younger persons today have more sexual partners than was previously the case, one can expect to see a steeper rise in the rate of AIDS in these countries.

Uganda was among the first nations to respond to the AIDS epidemic, making a strong effort to prevent the spread of the disease that appears to be having an effect, as is evident in the declining proportion of persons infected with HIV (Asiimwe-Okiror et al. 1997). This decline is consistent with behavioral studies that show increased

condom use, delay in sexual initiation, and a decline in the number of sex partners. Similarly, in Zaire, condom use among commercial sex workers dramatically increased from zero to 68 percent after a three-year condom promotion program (Adler 1998). The result was a marked decline in STDs. In Thailand, which is believed to have the best documented AIDS epidemic among developing countries, an aggressive and sustained national campaign to reduce HIV infection rates was instigated once authorities recognized that there was an epidemic. The campaign focused on increasing condom use, promoting respect for women, discouraging men from having sex with commercial sex workers, and increasing opportunities for education and employment to keep younger women from becoming commercial sex workers. The effects of this campaign have demonstrated the success of a concerted national effort. The use of condoms among commercial sex workers rose from 14 percent in 1989 to over 90% by 1994, rates of STDs declined, and the rate of new HIV infections has been declining, especially among commercial sex workers and their clients (Rojanapithayakorn and Hanenberg 1996). The success of this campaign has been attributed to several factors, including use of the existing infrastructure, a focus on a limited goal (improving condom use among commercial sex workers rather than trying to eliminate prostitution), widespread advertising, and a systematic means of monitoring the epidemic. Although less dramatic, local programs in other third world countries appear to be having some impact, as they have had in the United States. In Uganda, for example, a longitudinal study in an urban area showed a substantial increase in condom use, delay of sexual initiation, and a decrease in casual sex among adolescents over a period of two years. These changes were accompanied by a 40 percent decline in HIV seroprevalence among pregnant women who attended antenatal clinics (Asiimwe-Okiror et al. 1997). Similarly, a study of urban factory workers in Tanzania showed a decline over a two-year period in the number of sexual partners and casual sex, although condom use did not increase substantially (Ng'weshemi et al. 1996). A behavioral risk reduction program targeting truck-

ers in Kenya has shown significant declines in extramarital sex and sex with commercial sex workers, although condom use has not changed. These behavioral changes were accompanied by significant declines in STDs such as gonorrhea. In Bolivia, an HIV prevention intervention aimed at commercial sex workers appears to have increased condom use and resulted in lower STD rates in this high-risk group (Levine et al. 1998). In Jamaica, researchers have concluded that a comprehensive HIV/STD control program has resulted in reduced rates of STDs and increased condom use, although a significant minority continue to have unprotected sex in high-risk situations (Figueroa et al. 1998). Officials in some countries, such as India, deny that AIDS is a problem in their regions, and few efforts have been launched to stem the possibility of a widespread AIDS epidemic in those areas.

### CONCLUSIONS

Studies have demonstrated that the psychosocial and economic impact of AIDS extends far beyond the persons infected (Carael et al. 1998). Families have been shown to suffer from stress, economic hardship, stigma, and discrimination when one of their members has AIDS. Women, especially in third world countries, are in structurally less powerful positions than men and often have little or no control over their own sexual safety. These socioeconomic effects of AIDS, along with increasing mortality from AIDS among those in the most productive years of their lives, are expected to increase the gap between rich and poor and contribute to the feminization of poverty (Decosas 1996). Moreover, an infrequently mentioned consequence of AIDS in developing nations is its effect on children, who are being orphaned at alarming rates; WHO estimates that over eight million children in the world have been orphaned because of AIDS. In some areas of the world, AIDS may be changing the demography. It has been estimated that in areas hardest hit by the epidemic, life expectancy has declined by as much as seventeen years and mortality rates for children under age 5 have increased by 74 percent (Stover and Way 1998). Demographers project that by the year

2005, there will be 13 to 59 million fewer people and a 27 percent reduction in life expectancy in countries with the most severe epidemic compared with what have been the case if AIDS had not hit those regions (Stover and Way 1998). Mortality rates from AIDS are expected to continue to increase, and the gap in the death rates from AIDS in developing and developed countries is expected to grow wider. Finally, most people in third world countries cannot afford or do not have access to the new drugs that have prolonged lives of persons living with HIV or AIDS in the United States; thus, mortality rates there are not likely to drop. This may mean, as Way et al. (1999, p. 90) predict, that "the worst is yet to come."

The emergence of AIDS as a social issue not only has revitalized interest in the social context and consequences of STDs, it has caused medical research to understand more fully how disease can never be studied effectively apart from social conditions or without adequate information about relevant social actors. Still in its infancy is a fuller consideration of institutional and public responses to STDs, for example, how public policy gets made and by whom (Volinn 1989) or why some communities respond with compassion, others with fear, and others not at all. The social construction of disease, and that of STDs in particular, is an important and understudied area of social science research.

### REFERENCES

- Adler, M. 1998 "Strategies for the Prevention and Treatment of Sexually Transmitted Diseases." *International Journal of STD & AIDS* 9:8-10.
- Aral, S., and W. Cates 1989 "The Multiple Dimensions of Sexual Behavior as Risk Factors for Sexually Transmitted Disease: The Sexually Experienced Are Not Necessarily Sexually Active." *Sexually Transmitted Diseases* 16:173-177.
- , and K. Holmes 1999 "Social and Behavioral Determinants of the Epidemiology of STDs: Industrialized and Developing Countries." In K. Holmes, P. F. Starling, P. A. Mardh, S. Lemon, W. Stamm, P. Piot, and J. Wasserheit, eds., *Sexually Transmitted Diseases*, 3rd ed. New York: McGraw-Hill.
- Asiimwe-Okiror, G., A. A. Opio, J. Musinguzi, E. Madraa, G. Tembo, and M. Carael. 1997 "Change in Sex-

- ual Behaviour and Decline in HIV Infection among Young Pregnant Women in Urban Uganda. *AIDS* 11:1757-1763.
- Baldwin, J. I., S. Whitely, and J. D. Baldwin 1990 "Changing AIDS and Fertility Related Behavior: The Effectiveness of Sexual Education." *Journal of Sex Research* 27:245-262.
- Beadnell, B., S. A. Baker, D. M. Morrison, and K. Knox (in press) "HIV/STD Risk Factors for Women with Violent Male Partners. *Sex Roles*.
- Becher, M. H. 1988 "AIDS and Behavior Change." *Public Health Review* 16:1-11.
- Blumstein, P., and P. Schwarz 1976 "Male Bisexuality." *Urban Life*.5:339-358.
- Brandt, A. M. 1985 *No Magic Bullet*. New York: Oxford University Press.
- Calabrese, L. H., B. Harris, and K. Easley 1987 "Analysis of Variables Impacting on Safe Sexual Behavior among Homosexual Men in an Area of Low Incidence for AIDS." Paper presented at the Third International Conference on AIDS, Washington, D.C.
- Campbell, C. A. 1990 "Women and AIDS." *Social Science and Medicine* 30:407-415.
- Carael, M., B. Schwartlander, and D. Zewdie 1998 Preface to *AIDS* 12(suppl 1):S1-S2.
- Carrier, J. M. 1985 "Mexican Male Bisexuality." *Journal of Homosexuality* 1:75-85.
- Civic, D. in press "College Students' Reasons for Non-Use of Condoms within Dating Relationships." *Journal of Sex and Marital Therapy*.
- Cochran, S. D., and V. M. Mays 1990 "Sex, Lies and HIV." *New England Journal of Medicine* 322:774-775.
- Darrow, William 1979 "Sexually Transmitted Diseases in Gay Men. An Insider's View." *Sexually Transmitted Diseases* 6:278-280.
- Decosas, J. 1996 "HIV and Development." *AIDS* 10(suppl 3):S69-S74.
- Doll, L. S., and D. G. Ostrow 1999 "Homosexual and Bisexual Behavior." In K. Holmes, P. F. Starling, P. A. Mardh, S. Lemon, W. Stamm, P. Piot, and J. Wasserheit, eds., *Sexually Transmitted Diseases*, 3rd ed. New York: McGraw-Hill.
- Ehrhardt, A. A., C. Noretlinger, H. F. L. Meyer-Bahlburg, T. M. Exner, R. S. Gruen, S. L. Yingling, J. M. Gorman, W. El-Sadr, and S. J. Sorrell 1995 "Sexual Risk Behavior among Women with Injected Drug use Histories." *Journal of Psychology and Human Sexuality* 7:99-19.
- Figueroa, J. P., A. R. Brathwaite, M. Wedderburn, E. Ward, K. Lewis-Bell, J. A. Amon, Y. Williams, and E. Williams 1998. "Is HIV/STD Control in Jamaica Making a Difference?" *AIDS* 12(suppl 2):S89-S98.
- Fisher, J. D. 1988 "Possible Effects of Reference Group Based Social Influence on AIDS-Risk Behavior and AIDS Prevention." *American Psychologist* 49:914-920.
- , and W. A. Fisher 1992 "Changing AIDS-Risk Behavior." *Psychological Bulletin* 111:455-474.
- Ford, C. S., and F. A. Beach 1951 *Patterns of Sexual Behavior*. New York: Harper.
- Fox, R., D. Ostrow, R. Valdiserri, B. Van Rader, and B. F. Pall 1987 "Changes in Sexual Activities among Participants in the Multi Center AIDS Cohort Study." Paper presented at the Third International Conference on AIDS, Washington, D.C.
- Freeman, S. R., D. C. Des Jarlais, J. L. Sotheran, J. Garber, H. Cohen, and D. Smith 1987 "AIDS and Self-Organization among Intravenous Drug Users." *International Journal of Addictions* 23:201-219.
- Fullilove, M. T., and R. E. Fullilove 1989 "Intersecting Epidemics: Black Teen Crack Use and Sexually Transmitted Disease." *Journal of American Women's Association* 44:146-153.
- , R. E. Fullilove, K. Hayes, and S. Gross 1994 "Black Women and AIDS Prevention: A View toward Understanding the Gender Rules." *Journal of Sex Research* 27(1):47-64.
- Hart, G. 1978 "Social and Psychological Aspects of Venereal Disease in Papua New Guinea." *British Journal of Venereal Disease* 54:215-217.
- Humphreys, L. 1975. *Tearoom Trade: Impersonal Sex in Public Places*. Chicago: Aldine.
- Kalm, F. 1985 "The Two Faces of Antillean Prostitution." *Archives of Sexual Behavior* 14:203-217.
- Kane, S. 1990 "AIDS, Addiction and Condom Use: Sources of Sexual Risk for Heterosexual Women." *Journal of Sex Research* 27:427-444.
- Kelly, J. A., L. L. Otto-Salaj, K. J. Sikkema, S. D. Pinkerton, and F. R. Bloom 1998 "Implications of HIV Treatment Advances for Behavioral Research on AIDS: Protease Inhibitors and New Challenges in HIV Secondary Prevention." *Health Psychology* 17:310-319.
- Kirby, D., and K. Coyle 1997 "School-Based Programs to Reduce Sexual Risk-Taking Behavior." *Children and Youth Services Review* 19:415-436.

- Klovdahl, Alden S. 1985 "Social Networks and the Spread of Infectious Diseases: The AIDS Example." *Social Science Medicine* 21:1203-1216.
- Laumann, E. O., J. H. Gagnon, R. T. Michael, and S. Michaels 1994 *The Social Organization of Sexuality*. Chicago: University of Chicago Press.
- Leigh, B. C. 1990 "Sex and Drugs." *Journal of Sex Research* 27:199-213.
- , and R. D. Stall 1993 "Substance Use and Risky Sexual Behavior for Exposure to HIV: Issues in Methodology, Interpretation, and Prevention." *American Psychologist* 48:1035-1045.
- Levine, W. C., R. Revollo, V. Kaune, J. Vega, F. Tinajeros, M. Garnica, M. Estenssoro, J. S. Lewis, G. Higuera, R. Zurita, L. Wright-DeAguero, R. Pareja, P. Miranda, R. L. Ransom, A. A. Zaidi, M. L. Melgar, and J. N. Kuritsky 1998. "Decline in Sexually Transmitted Disease Prevalence in Female Bolivian Sex Workers: Impact of an HIV Prevention Project." *AIDS* 12:1899-1906.
- Magana, Raul 1990 "Bisexuality among Hispanics." Paper Presented at *CDC Workshop on Bisexuality and AIDS, American Institutes for Research*.
- Martin, J. L. 1987 "The Impact of AIDS on Gay Male Sexual Behavioral Patterns in New York City." *American Journal of Public Health* 77:578-584.
- McKusick, L., J. A. Wiley, T. J. Coates, R. Stall, G. Saika, S. Morin, K. Charles, W. Horstman, and M. A. Conant 1985 "Reported Changes in the Sexual Behavior of Men at Risk for AIDS in San Francisco 1982-84: The AIDS Behavioral Research Project." *Public Health Report* 100:622-629.
- Morrison, D. M., M. R. Gillmore, and S. A. Baker 1995 "Determinants of Condom Use among High-Risk Heterosexual Adults: A Test of the Theory of Reasoned Action." *Journal of Applied Social Psychology* 25:651-676.
- Ng'weshemi, J. Z. L., J. T. Boerma, R. Pool, L. Barongo, K. Senkoro, M. Maswe, R. Isingo, D. Schapink, S. Nnko, and M. W. Borgdorff 1996 "Changes in Males' Sexual Behaviour in Response to the AIDS Epidemic: Evidence from a Cohort Study in Urban Tanzania." *AIDS* 10:1415-1420.
- O'Reilly, K. R., and S. O. Aral 1985 "Adolescence and Sexual Behavior: Trends and Implications for STDs." *Journal of Adolescent Health Care* 6:262-270.
- Padian, N., L. Marquis, D. P. Francis, R. E. Anderson, G. W. Rutherford, P. M. O'Malley, and W. Winkelstein 1987 "Male to Female Transmission of Human Immunodeficiency Virus." *Journal of the American Medical Association* 258:788-790.
- Plummer, F. A., R. A. Coutinho, E. N. Ngugi, and S. Moses 1999 "Sex Workers and Their Clients in the Epidemiology and Control of Sexually Transmitted Diseases." In K. Holmes, P. F. Starling, P. A. Mardh, S. Lemon, W. Stamm, P. Piot, and J. Wasserheit, eds., *Sexually Transmitted Diseases*, 3rd ed. New York: McGraw-Hill.
- Poherat, J. J., R. Rothenberg, and D. C. Bross 1981 "Gonorrhea in Street Prostitutes: Epidemiology and Legal Implications." *Sexually Transmitted Diseases* 8:241-244.
- Proceedings NIMH/NIDA Conference on Women and AIDS: Promoting Health Behavior* 1989 Washington, D.C.: American Psychiatric Press.
- Reinish, J. M., C. A. Hill, S. A. Sanders, and M. Ziemba-Davis 1990 "Sexual Behavior among Heterosexual College Students" *Focus* 5:3.
- Roffman, R. A., R. S. Stephens, L. Curtin, J. R. Gordon, J. N. Craver, M. Stern, B. Beadnell, and L. Downey 1998 "Relapse Prevention as an Interventive Model for HIV Risk Reduction in Gay and Bisexual Men." *AIDS Education and Prevention* 10:1-18.
- Rojanapithayakorn, W., and R. Hanenberg 1996 "The 100% Condom Program in Thailand." *AIDS* 10:1-7.
- Ross, M. W. 1984 "Sexually Transmitted Diseases in Homosexual Men: A Study of Four Societies." *British Journal of Venereal Disease* 60:52-66.
- Schuster, M. R. Bell, G. Nakajima, and D. Kanouse 1998. "The Sexual Practices of Asian and Pacific Islander High School Students." *Journal of Adolescent Health* 23:221-231.
- Shilts, R. 1987 *And the Band Played On*. New York: St. Martin's.
- Siegal, K., and M. Glassman 1989 "Individual and Aggregate Level Change in Sexual Behavior among Gay Men at Risk for AIDS." *Archives of Sexual Behavior* 18:335-348.
- Sonenstein, F. L., L. Ku, L. D. Lindberg, C. F. Turner, and J. H. Pleck 1998. "Changes in Sexual Behavior and Condom Use among Teenaged Males: 1988 to 1995." *American Journal of Public Health* 88:956-959.
- Stover, J., and P. Way 1998. "Projecting the Impact of AIDS on Mortality." 1998 *AIDS* 12(suppl 1):S29-S39.
- Volinn, I. 1989 "Issues of Definitions and Their Implication: AIDS and Leprosy." *Social Science and Medicine* 20:1157-1162.

- Washington, A. E., R. L. Sweet, and M. A. B. Shafer 1985 "Pelvic Inflammatory Disease and Its Sequelae in Adolescents." *Journal of Adolescent Health Care* 6:298–310.
- Way, P. O., B. Schwartlander, and P. Piot 1999. "The Global Epidemiology of HIV and AIDS." In K. Holmes, P. F. Starling, P. A. Mardh, S. Lemon, W. Stamm, P. Piot, and J. Wasserheit, eds., *Sexually Transmitted Diseases*, 3rd ed. New York: McGraw-Hill.
- Wells, E. A., M. J. Hoppe, E. E. Simpson, M. R. Gillmore, D. M. Morrison, and A. Wilsdon 1995. "Misconceptions about AIDS among Children Who Can Identify the Major Routes of HIV Transmission." *Journal of Pediatric Psychology* 20:671–686.
- Wood, C. S. 1978 "Syphilis in Anthropological Perspective." *Social Science and Medicine* 12:47–55.

PEPPER SCHWARTZ  
MARY ROGERS GILLMORE

## SLAVERY AND INVOLUNTARY SERVITUDE

Many observers view slavery and freedom as polar opposites, but both slave and free wage labor systems rely on compulsion. Slave systems depend ultimately on physical coercion to force slaves to work for masters, although cultural, ideological, and economic pressures typically augment physical force. Wage labor systems, by contrast, depend on workers being free "in the double sense" (Marx [1867] 1967, pp. 168–169): Not only must workers be free to seek employment and choose among potential employers, they also must be free of all other means of subsistence that would allow voluntary withdrawal from the labor market. In the absence of subsistence alternatives, economic necessity compels "free" workers to exchange labor services for wages. Although wage labor systems depend primarily on labor-market processes to supply employers with workers, physical coercion often supplements those processes, especially during periods of economic decline. Cultural expectations and ideological appeals also reinforce market mechanisms. Nevertheless, large-scale labor systems are maintained primarily by a mixture of physical and economic coercion that varies with the availability of subsistence alternatives.

The way in which the constellation of physical and economic coercion and subsistence alternatives is determined by the power of contending groups as well as historically specific cultural and ideological factors has been of great interest to social scientists. Perhaps the simplest and most durable statement of the causes of slavery is a conjecture known as the Nieboer–Domar hypothesis (Nieboer 1900; Domar 1970; Engerman 1986a; see Patterson 1977b for a critique), which links slavery to an abundance of arable land combined with a shortage of labor. The way in which slavery differs from other forms of involuntary servitude is explained in the next section. The Nieboer–Domar hypothesis is then amended to provide a provisional explanation for the worldwide trend away from slavery and toward freedom in large-scale labor systems over the last several hundred years. Finally, the Nieboer–Domar hypothesis is reevaluated in light of current patterns of slavery and involuntary servitude around the world.

### SLAVERY AND OTHER FORMS OF INVOLUNTARY SERVITUDE

Patterson (1982, p. 13) argues that slavery is defined by three conditions. First, slaves suffer perpetual domination that ultimately is enforced by violence. The permanent subjugation of slaves is predicated on the capacity of masters to coerce them physically. Second, slaves suffer natal alienation, or the severance of all family ties and the nullification of all claims of birth. They inherit no protection or privilege from their ancestors, and they cannot convey protection or privilege to their descendants. Third, slaves are denied honor, whereas masters are socially exalted. This condition appears to be derivative rather than definitive of slavery because all hierarchical social systems develop legitimating ideologies that elevate elites and denigrate those at lower levels. The first two conditions, which distinguish slavery from other forms of involuntary servitude, constitute the working definition used in this article.

In chattel slave systems, slaves are movable property owned by masters and exchanged through market processes. Because some societies con-

structed elaborate slave systems without well-developed notions of property and property rights, property relationships cannot be an essential defining element of slavery (Patterson 1982; 1977a). Nevertheless, property relations and economic processes had important effects on slavery and other forms of unfree labor in the Americas, Europe, and Africa in the period after the fifteenth century, which is the major focus of this analysis.

An unfree laborer cannot voluntarily terminate service to a master once the servile relationship has been established. Slavery maximizes the subordination of servant to master. Other servile workers, such as indentured and contract laborers, debt servants, peons, and pawns, are less dominated than slaves are and do not suffer natal alienation. Pawns, for example, were offered by their families in return for loans. Pawns maintained kinship ties to their original families, a situation which gave them some protection, and were freed once the loans were repaid. Indentured servants agreed to be bound to a master for a specific term, such as seven years, in exchange for a benefit such as passage to America or release from prison (Morris 1946; Smith 1947; Morgan 1975). Contract laborers also were bound for specified terms but could not be sold against their will to other masters, as was the case with indentured servants. Debt servitude consists of labor service obligations that are not reduced by the amount of work performed (Morris 1946; Sawyer 1986). Peons are tied to land as debt servants and owe labor services to a landlord. Serfs are not debt servants, but they are tied to land and perform labor services on their lords' estates. The right to labor services enjoyed by European feudal lords was vested in their political authority rather than in land ownership, although serfs were reduced to slaves in all but name in some instances (e.g., Russia in the nineteenth century) (Kolchin 1987).

Indentured servants and contract laborers may agree to the initial terms of their servitude, but they cannot willingly end it during its term once it begins. Usually some form of coercion, such as poverty, debt, or impending imprisonment, was necessary to force people to agree to terms of contractual servitude or pawnship. By contrast,

the status of the slave, serf, peon, and debt servant typically was inherited or imposed on workers against their will.

### SLAVERY, THE LAND-LABOR RATIO, AND THE STATE

In its simplest form, the Nieboer-Domar hypothesis states that abundant free land makes it impossible for free workers and nonworking landowners to coexist. If free land is available and laborers can desert landowners whenever they choose, landowners will be unable to keep enough workers to maintain their status as nonworkers. If landlords can compel workers to perform labor services despite the availability of free land, landlords become labor lords and workers are not free. By contrast, scarce land combined with an abundant labor supply drives wages down, making wage laborers less expensive than slaves and other servile workers. When they are denied access to land, hunger forces workers to labor for wages and wage labor systems displace slave labor systems.

This model appears to be deficient in at least four ways. First, as Domar recognized, political factors determine the degree of freedom enjoyed by workers. Chief among those factors is the extent to which the state protects the interests of landowners when they conflict with those of laborers. Large-scale slave labor systems cannot exist without states that defend the power of slave masters to control and utilize the labor of slaves. A powerful state is essential for protecting slave masters against slave rebellions, capturing runaways, and enforcing slave discipline. State power is required for the enslavement of new supplies of slaves. If the state is responsive to the demands of workers or if workers can voluntarily withdraw their labor services, unfree labor systems cannot be maintained.

Second, the model presumes that slave masters exploit slaves in response to economic incentives, but slaves and other unfree laborers often provided military, administrative, domestic, and sexual services largely unrelated to economic activities (Roberts and Miers 1988; Patterson 1982).

The Nieboer–Domar hypothesis therefore does not apply to societies that employ slaves and other servile workers in noncommercial or minor economic roles (Lovejoy 1983; Finley 1968). It also does not apply to states that use race, religion, gender, or other status criteria to restrict the freedom of workers for noneconomic reasons (James 1988).

Third, the key issue from an employer’s perspective is not simply the ratio of land to labor but the relative costs and benefits of different forms of labor that can be profitably employed using existing capital (including land). A more general version of the Nieboer–Domar model compares the stock of available capital to the availability of different forms of labor at prevailing prices. Thus, labor scarcity means the scarcity of labor at prices that allow it to be employed profitably.

Fourth, the simple version of the Nieboer–Domar hypothesis ignores the organizational capacities of workers and capitalists’ ability to adopt labor-saving innovations. If workers demand concessions that threaten profits or engage in strikes and other production disruptions, capitalists experience “labor shortages” that stem not from insufficient numbers but from the organized resistance of the workers who are present (Miles 1987). Faced with such disruptions, capitalists with sufficient capital may adopt labor-saving innovations if they are available. When capitalists are unable to adopt those innovations, they may resort to coercive strategies to curb workers’ market-based demands (Paige 1975). This case contradicts the Nieboer–Domar hypothesis, which assumes that high ratios of labor to capital (or land) make coercive labor control strategies unnecessary.

#### UNFREE LABOR IN THE AMERICAS

From the fifteenth through the nineteenth centuries, Europe, Africa, and the Americas were closely linked by flows of people and commodities (Lovejoy 1983; Eltis 1987). The colonization of the Americas by strong European states provided vast, lightly populated lands for commercial exploitation. Expanding markets in Europe for sugar, cotton,

tobacco, coffee, and other commodities stimulated the demand for greater supplies of servile labor to work the plantations and mines of the Americas. Weak states in large areas of sub-Saharan Africa left large populations vulnerable to armed predation by stronger states that supplied the expanding markets for slaves.

Estimates of the numbers of bondsmen and slaves transported to the Americas are subject to sizable errors because of the paucity and unreliability of existing records, but relative magnitudes are thought to be reasonable (see Table 1). Differences in the sources of servile labor produced different racial compositions across American regions. Slaves from Africa outnumbered arrivals from Europe nearly four to one before 1820, and most were bound for sugarcane plantations in Brazil and the West Indies. British North America was atypical because its early immigrants were predominantly white indentured servants from Britain, Ireland, and Germany; perhaps two-thirds of the white immigrants who arrived before the American Revolution were bonded servants (Smith 1947, p. 336). Before being displaced by African slaves, white bondsmen were the principal source of labor in the plantation regions of all British colonies, including those in the Caribbean (Engerman 1986a; Galenson 1981).

Indentured servitude was the principal method of defraying the costs of supplying the colonies with workers. British laws and customs regulating master–servant relationships were modified significantly to fit American circumstances (Galenson 1981). Because of the high costs of transatlantic passage, longer periods of service were required, typically four to seven years rather than one year or less in England. English servants could not be sold against their will to another master, but that practice was sanctioned in colonial laws and customs because European servants could not negotiate terms with prospective masters before immigrating to America. Finally, opportunities for escape were much greater in America. Consequently, elaborate state enforcement mechanisms were implemented to discourage runaways and to catch, punish, and return those who did. Most indentured servants were transported to plantation re-

**Immigration to and Populations of Regions in the Americas (in thousands)**

<b>TOTAL IMMIGRATION TO THE AMERICAS UP TO AROUND 1820</b>				
	<b>African</b>	<b>European</b>	<b>Total</b>	<b>% African</b>
United States	550	651	1,201	46
Continental Spanish America	1,072	750	1,822	59
Brazil and the West Indies	<u>6,777</u>	<u>964</u>	<u>7,741</u>	<u>88</u>
Total	8,399	2,365	10,764	78
<b>TOTAL POPULATION AROUND 1650</b>				
	<b>Native Americans and Mestizos</b>	<b>Europeans</b>	<b>Blacks and Mulattos</b>	<b>Total</b>
North America	860 (86%)	120 (12%)	22 (2%)	1,002 (100%)
Continental Spanish America (excluding Peru)	8,773 (90%)	575 (6%)	437 (4%)	9,785 (100%)
Brazil, the West Indies, and the Guyanas	<u>843 (51%)</u>	<u>154 (9%)</u>	<u>667 (40%)</u>	<u>1,664 (100%)</u>
Total	10,476 (84%)	849 (7%)	1,126 (9%)	12,451 (100%)
<b>TOTAL POPULATION AROUND 1825</b>				
	<b>Native Americans and Mestizos</b>	<b>Europeans</b>	<b>Blacks and Mulattos</b>	<b>Total</b>
North America	423 (4%)	9,126 (80%)	1,920 (17%)	11,469 (100%)
Continental Spanish America (excluding Peru)	12,660 (79%)	2,937 (18%)	387 (2%)	15,984 (100%)
Brazil, the West Indies, and the Guyanas	<u>381 (5%)</u>	<u>1,412 (20%)</u>	<u>5,247 (75%)</u>	<u>7,040 (100%)</u>
Total	13,464 (39%)	13,475 (39%)	7,554 (22%)	34,493 (100%)

*Table 1*

SOURCES: Immigration rates are adapted from Eltis (1983, p. 278). Population figures are adapted from Slicher Van Bath (1986, p. 21), in which the West Indies include the Spanish islands but exclude the Bahamas.

gions because plantation labor produced greater returns than did any other economic activity in the Americas (Galenson 1981). Employers in areas such as New England could afford few or no servants because they specialized in trades with lower labor productivity and lower profit margins.

White servile labor was replaced by black slavery throughout the Americas between 1600 and 1800. Racial prejudice encouraged the shift but probably was not decisive (Morgan 1975). First, the limited supply of indentured servants could not satisfy the demand for servile labor, whereas the supply of African slaves was almost completely elastic. Improving economic conditions in Britain and state restrictions on the emigration of British

servants reduced the numbers seeking passage to America, causing the price of servants to increase. As the price of servants exceeded the price of slaves, first for unskilled and later for skilled workers, slaves came to be preferred to bonded servants (Galenson 1981). Second, Africans were more resistant to the diseases of the tropics, where the most important export crops were grown (Eltis 1983).

Third, slaves could be compelled to comply with the labor-intensive plantation work regime that developed (Fogel 1989). Slaves were more efficient and profitable than free or indentured workers in sugar, cotton, coffee, rice, and tobacco agriculture because the work required by those crops could be performed efficiently by slave work



gangs. Work gangs were organized according to specialized tasks, and slaves were assigned to particular gangs according to their skills and capacities. The work was performed under close supervision to maintain work intensity and quality. Slave masters often used brutal violence to enforce discipline, but naked force might have been used less than once was thought. Slave masters experimented with different mixtures of positive and negative incentives, to encourage slaves to maximize their output (Fogel 1989). Thus, slave plantations anticipated the discipline of workers in the great factories of industrial capitalism, where assembly lines regulate the rhythms and intensity of work.

Forced migration from Africa greatly exceeded all migration from Europe as sugar production became the greatest consumer of servile labor in the Americas. High death rates and a preference for male slaves in the sugar-producing regions led to net population declines among blacks and mulattoes (compare immigration numbers to population sizes in Table 1), but the proportion of blacks in the British West Indies increased from 25 to 91 percent between 1650 and 1770 (Fogel 1989, p. 30). By the 1820s, the proportion of blacks and mulattoes in Brazil, the Guyanas, and the West Indies reached 75 percent (Table 1).

British North America was an exception to this pattern as both black and white populations had high rates of natural increase. Almost all major slave societies were unable to maintain the size of slave populations without continuous replenishment from outside sources. By contrast, the slave population in the United States multiplied because of unusually high fertility rates and low mortality rates (see Table 1 and Fogel 1989).

Political factors also encouraged the transition from white servitude to black slavery (Engerman 1986b; Galenson 1981). As British citizens, indentured servants retained state-protected natal rights that their masters were obliged to respect. For example, masters could beat servants and slaves to enforce work discipline, but colonial courts protected servants against unfair punishment (Smith 1947). Importantly, Europeans could choose the place of their servitude, and most refused trans-

portation to the plantation regions from the eighteenth century on. African slaves could not avoid the plantation regions and were citizens of no state in Africa or America that could or would defend their interests.

Because Spain conquered the continental regions with the largest Native American populations (Table 1), it had less need of African slaves. Instead, Spanish colonists installed a coercive labor system patterned on Spanish feudalism that forced natives to work part-time on colonial estates although slavery was still preferred in the mines (Slicher Van Bath 1986; Kloosterboer 1960). Unfree labor markets and compulsory labor endured for 400 years, eventually evolving into debt servitude in the nineteenth century. Native Americans and mestizos accounted for nearly 80 percent of the population of continental Spanish America by 1825 but were almost annihilated in the West Indies (Table 1).

Nowhere in the Americas was slavery in danger of withering away economically at the time when it was abolished (Eltis 1987). Furthermore, with the principal exception of Haiti in 1804, slave rebellions were not successful in conquering slave masters and transforming a slave system into a wage labor system. Paradoxically, Britain played the dominant role in abolishing slavery and the transatlantic slave trade even though it controlled half the transatlantic commerce in slaves and half of the world's exports in sugar and coffee, which were produced primarily on slave plantations (Eltis 1987). Britain outlawed the slave trade in 1808 and freed the slaves in its West Indian colonies in 1833 over the strenuous objections of slave owners. The United States prohibited the importation of slaves after 1808, and civil war led to abolition in 1865. By the 1870s, all the major European and American maritime and commercial powers had acquiesced to British pressure and outlawed the slave trade. Brazil, the last state in the Americas to abolish slavery, did so in 1888.

The land-labor ratio strongly affected planters' responses to abolition. In places where ex-slaves could find no alternative to plantation work, such as Barbados and Antigua, the transition to

free labor was rapid, and plantation production did not decline appreciably (Boogaart and Emmer 1986). In places where land or alternative employment was available, such as Jamaica and Trinidad, the ex-slaves abandoned the plantations, and plantation productivity declined (Engerman 1985). In response, planters implemented a variety of servile labor systems with mixed results. A second wave of indentured servants was imported chiefly from Asia, especially China and India, which more than compensated for the labor shortages induced by abolition in some cases, such as Mauritius and British Guiana (Engerman 1985, 1986b). China and colonial India eventually banned the recruitment of servants because of objections to employers' poor treatment of servants, and Brazil was never able to gain access to Asian indentured laborers (Boogaart and Emmer 1986).

In areas where planters retained a degree of political power, such as the West Indies and Brazil, vagrancy statutes and other compulsory labor schemes forced workers to accept wages below free market levels (Kloosterboer 1960; Huggins 1985). Indentured labor and other forms of involuntary servitude were banned in the United States in 1865 by the Thirteenth Amendment to the U.S. Constitution, but planters regained substantial influence over black workers through their control of racially discriminatory state and local governmental institutions (James 1988). Blacks were disfranchised by 1900, making them vulnerable to racial segregation, physical coercion, and economic discrimination. The extent to which racial discrimination interfered with free labor markets in the South is controversial (Wright 1986). Nevertheless, the most determined resistance to the civil rights movement of the 1960s occurred in the plantation regions (James 1988). The success of that movement led to increased protection of the citizenship rights of blacks and doomed widespread coercive labor control practices. The transition to capital-intensive agricultural practices was rapid during that period.

#### UNFREE LABOR IN AFRICA AND ASIA

Slavery was an indigenous institution in Africa and Arabia for centuries before Europeans entered

the African slave trade (Thornton 1998). While approximately 9.9 million Africans were transported to the Americas before the Atlantic slave trade was suppressed (Fogel 1989), an additional 5.2 million African slaves were transported across the Sahara, the Red Sea, and the Indian Ocean into the Islamic world between 1500 and 1900. Moreover, perhaps 6.4 million more were exported to Islamic societies between A.D. 650 and 1500 ("a rough approximation," Lovejoy 1983, p. 24). Many thousands more were enslaved in African societies in that period (Thornton 1998).

Whereas chattel slavery in the Americas was predicated on profit making, African slavery typically did not have a narrowly economic basis. African slaves were menial servants and field workers, but they also were concubines, surrogate kin, soldiers, commercial agents, and candidates for human sacrifice (Roberts and Miers 1988, p. 5). Female slaves were especially valued because women performed most agricultural and domestic work. African societies were based on kinship relations in which all individuals were linked in a complex network of dependency. Because power in kinship systems depends on the size of social groups, slave masters could increase their power by obtaining more slaves. Furthermore, slaves were immune to the appeals of their masters' rivals within kin groups because they had no kinship ties that mediated their subordination to their masters. Large numbers of persons were enslaved as a result of military victories in wars between African kingdoms and societies.

African slave masters also responded to economic incentives. An increasing number of slaves were provided to the Atlantic slave trade as the demand for slaves in the Americas increased. Thornton (1998, p. 125) concludes that African participation in the slave trade was voluntary because European slavers did not have the economic or political power to force African leaders to sell slaves. The established African practices of holding and trading slaves made it possible for African states to respond to the increasing European demand so long as the prices paid were attractive.

Islamic slavery also differed from chattel slavery in important ways. Islamic law prohibited the

enslavement of Muslims but permitted the enslavement of people born to slave parents or captured for the purpose of conversion to Islam (Gordon 1989). Concubines could not be sold if they bore a child to a master, and the child could not be enslaved. Allowing slaves to purchase their freedom brought honor to former masters. Manumitting slaves was also meritorious and could atone for certain sins and offenses.

Islamic slaves typically were employed as household servants, domestic workers, concubines, and to a lesser extent soldiers. Female slaves typically brought higher prices than did males because the heads of patriarchal Muslim families prized female slaves for assignment to sexual and domestic roles in their households. Slave eunuchs performed special tasks in large households and usually brought higher prices than did female slaves. Consequently, pre-twentieth-century slave traders castrated large numbers of African slave boys in crude operations that killed up to 90 percent of them (Gordon 1989, pp. 91–97). However, Islamic slave masters also responded to economic incentives as did their American counterparts when market opportunities arose. During the nineteenth century, over 750,000 slaves were transported to the clove plantations on Zanzibar and other locations on the east coast of Africa, for example (Cooper 1977; Lovejoy 1983, p. 151).

British diplomatic and military pressure finally led to the suppression of the Islamic and African slave trades as it did with the transatlantic traffic. In 1890, all the European powers agreed to suppress slave trading and slave raiding and to assist ex-slaves, a commitment that legitimated the conquest of Africa in the eyes of European citizens. However, European colonial administrators were reluctant abolitionists (Roberts and Miers 1988). Inadequate military and administrative power, fear of economic and political disruptions, and unfamiliarity with African customs delayed the process.

Colonial governments outlawed slavery almost everywhere in sub-Saharan Africa by the 1930s, but involuntary servitude persisted. Roberts and Miers (1989, pp. 42–47) identify three factors that

retarded the emergence of free labor markets in Africa. The first two were responses to abundant land and scarce labor. First, colonial states conscripted natives, imposed labor levies that local chiefs had to fill, and implemented other compulsory labor mechanisms to maintain a supply of cheap labor for European employers and administrators. Second, many Africans had access to land or livestock and were unwilling to work for wages. Colonial states tried to reduce the attractiveness of nonwage occupations by, for example, raising taxes above what peasant agriculturalists and pastoralists could pay and prohibiting Africans from growing lucrative cash crops. In settler colonies such as South Africa, native Africans were pushed off the land and confined to strictly regulated labor markets by pass laws. Third, Africans resorted widely to pawnship after abolition.

The reluctance of colonial administrators and the power of postcolonial states allowed slavery to survive in some nations in north Africa and the Arabian peninsula well into the twentieth century. In 1926, the League of Nations codified its opposition to slavery by adopting the Convention to Suppress the Slave Trade and Slavery, which defined slavery as the ownership of another person. Gradually, the remaining slave states abolished slavery officially: Ethiopia in 1942, Saudi Arabia in 1962, Muscat and Oman in 1970, and Mauritania for the third time in 1980. Nevertheless, reports of slavery persisted. Saudi Arabia allegedly failed to free some 250,000 slaves in the late 1960s; an estimated 100,000 chattel slaves existed in Saharan regions of Mauritania in 1980, although many were freed by 1984; and nomadic tribesmen allegedly held 250,000 slaves in the Sahelian districts of Mali in 1984 (Gordon 1989, pp. 232–234; United Nations 1984, pp. 18–19; United Nations 1988, p. 197; Sawyer 1986, p. 14).

In 1956, the United Nations increased the international attack on slavery and involuntary servitude by adopting the Supplementary Convention on the Abolition of Slavery, the Slave Trade, and Institutions and Practices Similar to Slavery. In addition to outlawing slavery, the Supplementary Convention pledged signatory nations to suppress debt bondage, serfdom, the pawning of chil-

dren, and servile marriage (forcing women to marry in exchange for payments to their family members or assigning wives, after the death of their husbands, to others as an inheritance). Progress has been slow. For example, India outlawed bondage in 1976, but a survey found more than 2.5 million bonded workers in 1978; only 163,000 had been freed by 1985 (Sawyer 1986, pp. 124–134). Debt servitude has been reported since 1970 among landless peasants in India and Nepal and among Native American rubber collectors in the Peruvian Amazon. As late as 1986, the Dominican Republic used its army to round up Haitian immigrants for forced work on sugar plantations during the harvest season (Plant 1987).

#### **PATTERNS OF SLAVERY AND UNFREE LABOR SINCE 1990**

Large-scale systems of slavery and involuntary servitude can be maintained only if slave owners and labor lords can use physical coercion to maintain labor discipline. Hence, large-scale systems of unfree labor depend on state institutions that deny citizenship rights to unfree workers and augment the power of dominant classes to coerce their workers physically. Today no nation officially protects the rights of employers to reduce their workers to slavery or involuntary servitude. Virtually all members of the United Nations have ratified the Supplementary Convention on the Abolition of Slavery, the Slave Trade, and Institutions and Practices Similar to Slavery (see United Nations [1957] 1999 for the current list of ratifying nations). Ratification of the Supplementary Convention officially commits a nation to the elimination of slavery and involuntary servitude within its borders and obligates it to cooperate with other nations in suppressing those practices. Although it took 200 years, the international antislavery campaigns led by politically powerful nations with wage labor markets were successful. Large-scale systems of slavery and involuntary servitude supported and protected by complementary state institutions no longer exist.

The expansion of capitalism and increasing world population displaced large numbers of peo-

ple from subsistence agriculture and other means of support in many regions. Great disparities between rich and poor nations drive people across state boundaries in search of jobs and improved living conditions. State power plays a crucial role in shaping migration and molding the relationship between capital and labor, but states with expanding economies now prevent the entrance of many willing workers rather than compelling the entrance of the unwilling. The whip of unemployment and poverty replaces the slave master's lash as free labor replaces slave labor.

However, slavery and forced labor persist and are widespread in some areas. Anti-Slavery International, the world's oldest human rights organization, estimates that over 200 million people, about 3 percent of the world's population, labor in some form of bondage. Table 2 provides examples of some existing systems of unfree labor. Because reliable information is difficult or impossible to obtain in some cases, the examples in Table 2 should not be considered exhaustive or the most egregious. The best available information suggests that slavery and involuntary servitude occur with the greatest frequency in nations that are ravaged by civil war or have weak states that are unwilling or unable to suppress coercive labor practices.

Somalia, Sudan, and Uganda provide examples of how civil war places defenseless people at the mercy of powerful military groups. A United Nations special rapporteur confirmed that armed militia groups abducted people in southern Sudan for use as forced laborers or for sale as slaves. Prisoners were subjected to beatings, electric shock, exposure to the sun for long periods, pouring of cold water on the naked body, rape and the threat of rape, sleep deprivation, and the refusal of food and medical treatment. Sudanese government security forces and allied militias as well as insurgent groups were guilty of conscripting children and forcing them to fight as soldiers (United Nations 1997). The civil war in Sudan has disrupted agricultural production to the extent that some impoverished parents give or sell their children to others to prevent their starvation (Finnigan 1999).

A large and active chattel slavery market is in operation in Sudan. Its magnitude is not known,

SLAVERY AND INVOLUNTARY SERVITUDE

**Examples of Slavery and Involuntary Servitude in 1998**

COUNTRY	Servitude Type	Sector of Employment	Legal Status of the Servitude	Estimates of the Scope of Involuntary Servitude
Bangladesh	Bonded labor; forced labor; forced child labor; forced prostitution	Garment industry; forced prostitution	Prohibited by constitution	10,000 to 29,000 child prostitutes; bonded labor is not widespread; trafficking in women and children is widespread
Benin	Bonded labor; forced child labor	Agriculture; domestic service	Prohibited by labor codes	Some poor parents indenture their children
Brazil	Forced labor; bonded labor; forced child labor	Agriculture; sugar industry; mining industries	Prohibited by constitution	The government conducted more than 400 raids between 1995 and 1997 that freed more than 130,000 forced laborers; forced workers numbered 1.3 million in 1992 (Sutton 1994)
Burma	Forced labor; forced child labor;	Irrigation, transportation, and tourism services, and military service	No law prohibits forced or bonded labor by children	Government use of forced labor is widespread; bonded labor is not practiced
Cambodia	Forced and bonded labor; child prostitution	Military; wood-processing, rubber, and brick industries; prostitution	Labor law prohibits forced or bonded labor, including children	Prostitution and trafficking in children is widespread; bonded labor occurs but is not widespread
Cameroon	Slavery; forced and bonded child labor; contracting prison labor to private employers	Agriculture	Labor code does not protect children	The slavery still practiced in northern Cameroon is primarily enslavement of Kirdi by Fulani, a Muslim group that conquered the Kirdi 200 years ago
China	Contracting prison labor to private employers; forced prostitution	Manufacturing; agriculture; mines	Government prohibits export of goods made by prisoners and prohibits forced labor by children	Kidnapping and sale of women and children for prostitution is a problem recognized by the government; prohibitions against private use of prison labor have not been enforced effectively
Ethiopia	Sexual bondage	Children, especially girls, used as prostitutes	Prohibited by constitution and criminal code	Large-scale use of children as prostitutes; children are kidnapped and sold for about \$36; some poor parents sell their children

Table 2

SOURCE: United States Department of State (1999) unless otherwise noted.

but international human rights groups estimate that chattel slaves number in the tens of thousands and that the market may extend as far as Saudi Arabia and the Gulf states (Finnigan 1999, p. 71). Anti-Slavery International (1999a) reports that more than 2,700 slaves were freed during the first four months of 1999 in return for over \$100,000 in payments. Finnigan (1999) photographed a Sudanese Arab slaver who hoped to sell over 130 indi-

viduals to a Christian antislavery organization for \$50 each. If that plan failed, he decided that he might return the former slaves to their families for a much lower price. Many others were not so fortunate. Slaves in Sudan are subject to severe punishment; are stripped of their cultural, religious, and personal identities; and can become the property of another person for life, traded and inherited, branded and bred.

SLAVERY AND INVOLUNTARY SERVITUDE

Continued

COUNTRY	Servitude Type	Sector of Employment	Legal Status of the Servitude	Estimates of the Scope of Involuntary Servitude
India	Bonded labor; forced labor; indentured and bonded child labor; forced prostitution; contracting prison labor to private employers and brothels	Carpet industry; prostitution; domestic servants; gemstone, glass, footwear, textiles, silk, and fireworks industries	Prohibited by constitution and criminal code	5 million bonded laborers; 300,000 forced child laborers in the carpet industry alone; frequent reports of sale of children and women for forced domestic service and prostitution; some poor parents sell their children
Indonesia	Bonded labor; bonded child labor; forced prostitution	Fishing industry; prostitution	Law prohibits forced labor	Several thousand bonded child laborers; hundreds of bonded adults; widespread forced prostitution; 1,500 child prostitutes in one province
Mauritania	Slavery and vestiges of slavery; unpaid labor; forced child labor	Agriculture; shepherds and herdsmen	Slavery officially abolished several times, most recently in 1980	Slavery in the form of forced and involuntary servitude reportedly persists in some isolated areas; psychological, tribal, and religious bonds continue to tie former slaves to their prior masters; forced child labor is now rare
Morocco	Adoptive servitude of children	Domestic servants	Forced and compulsory labor prohibited by statute in 1957	The adoption of young girls for domestic service is socially accepted, and the government does little to discourage the practice
Nepal	Bonded labor; child labor; trafficking in women and children for sex work	Agriculture; forced prostitution; carpet industry	Prohibited by constitution	100,000 bonded labors in one region; 40,000 bonded child laborers; 5,000 to 7,000 girls forcibly transported to India as prostitutes annually; forced prostitution is widespread; forced child labor in carpet industry has been greatly reduced to international pressure
Nigeria	Child sexual slavery; forced labor	Domestic servants; prostitution	Prohibited by constitution but children are not protected	Child slavery rings operated between Nigeria and neighboring countries; also see Effah (1996); forced labor is now rare

Table 2, continued

Some slavery and slavery-like practices in India, Benin, Mauritania, Morocco, and Pakistan are supported by local customs and traditions that have long histories (Table 2). Mauritania and India have attempted to eliminate these practices, whereas other states are more reluctant or unable to act decisively. Debt bondage is widespread in the rural areas of India, but involuntary servitude has been adapted to the economic opportunities provided

by the operation of global markets. Just as European demand for sugar and coffee drove the Atlantic slave trade in the seventeenth and eighteenth centuries, employers of servile labor in poor countries find opportunities to supply consumer demands for cheap goods in economically advanced countries. For example, India, Pakistan, and Nepal employed as many as one million servile child workers in the hand-knotted carpet industry

SLAVERY AND INVOLUNTARY SERVITUDE

Continued

COUNTRY	Servitude Type	Sector of Employment	Legal Status of the Servitude	Estimates of the Scope of Involuntary Servitude
Pakistan	Bonded labor; bonded child labor; forced prostitution of children	Carpet, fish, glass and brick industries; agriculture; construction industry; prostitution	Forced labor is specifically prohibited by law	Bonded labor is widespread; landlords held 4,500 bonded laborers in one region; poor parents traditionally sell children to rich landlords as permanent bond servants in exchange for money or land; significant numbers of women and children are forced to work as prostitutes.
Somalia	Forced labor; child labor	Agriculture; military service	Civil war has eliminated any effective governmental system for the protection of human rights	Widespread vulnerability to forced labor in the service of warlords; in contrast to previous years, there were no reports of the use of forced labor by multinational fruit export firms
Sudan	Slavery including enslavement of prisoners of war, forced labor, and child labor	Agriculture; domestic service; concubines; trafficking in slaves; military service	Law prohibits forced or compulsory labor, but civil war has eliminated effective protection of human rights in many areas	Widespread use of slavery and forced labor by government troops and their allies; government forces and insurgents forcibly conscript children; international religious organizations paid \$50 each to purchase the freedom of slaves captured by raiding parties (Finnigan 1999)
Thailand	Trafficking of women and children for sex work; bonded child laborers	Prostitution, especially in response to tourists' demand; agriculture	Constitution prohibits forced labor but does not protect children	Forced prostitution is often protected by government officials who profit from it; child prostitutes number 20,000 to 40,000.
Togo	Forced labor; indentured servitude	Domestic servants	No law addresses forced or compulsory labor by adults or children	International trafficking in children, especially girls, for use as indentured servants or slaves; the government attempts to suppress the practice.
Uganda	Bonded labor; forced child labor; contracting prison labor to private employers	Agriculture; domestic service; concubines; soldiers	Prohibited by law	Government forces were too weak to defend rights of citizens; an insurgent militia group abducted 3000 Ugandan children and forced them to become soldiers or sexual slaves

Table 2, continued

in 1994 (United States Department of Labor 1995). The United States imported \$329 million worth of hand-knotted carpets in 1996, a large proportion of which were produced in those countries (United States Department of Labor 1997). Children in India have been kidnapped and transported from their villages to face years of forced labor in the

carpet industry. Forced child labor is characterized by long hours, threats of violence, dangerous conditions, little or no pay, and poor or nonexistent health care (United States Department of Labor 1994, 1998; Anti-Slavery International 1999b). Other commercial products produced by bonded and forced child workers for export are brassware,

silk cloth and silk garments, and stone and glass products (United States Department of Labor 1997).

Thailand, Nepal, Indonesia, and Cambodia are examples of weak states or states crippled by corruption that facilitate the persistence of forced labor and forced sexual prostitution. The flourishing Asian sex trade involves the transport of women and children across borders to work in brothels for foreign and domestic customers. Victims are lured by false promises of decent employment, kidnapped or sold by family members, or reduced to debt bondage by poverty. In Thailand, debt bondage sometimes continues from generation to generation as a result of very low wages, high interest charges, and fraudulent debts that cannot be repaid (U.S. Department of Labor 1994, 1996, 1998). Because of the spread of AIDS, brothel customers increasingly prefer very young girls who are supposedly disease-free. Consequently, brothel owners purchase or kidnap young girls from the surrounding countries to supply the demand. Girls from Burma, Cambodia, China, Laos, and Vietnam can be found in the brothels of Thailand (United States Department of Labor 1995).

The Asian sex trade is huge. UNICEF estimates that 100,000 child prostitutes are employed in Thailand alone, but the number could be more than 200,000 (U.S. Department of Labor 1995). Patterns vary from country to country, but the brothels typically are patronized by local customers and to a lesser extent by foreigners, including tourists, businessmen, and military personnel from the United States and Europe. Travel agencies arrange sex tours that include accommodations and a choice of escorts (United States Department of Labor 1995, 1996). As is the case with many handmade and manufactured products, the Asian sex trade and the exploitation of servile workers are intimately connected to the global economy.

In contrast to Britain's use of the navy to suppress the Atlantic slave trade in the nineteenth century, current international efforts to suppress slavery and involuntary servitude are weak. One enforcement tactic is to expose governments that do not suppress servile labor practices to the condemnation of world opinion. For example, the United Nations Commission on Human Rights investigates patterns of human rights violations within countries, including slavery and servitude, and disseminates its reports widely. The Interna-

tional Labor Organization, a special agency of the United Nations, formulates international labor standards and monitors the compliance efforts of governments (United Nations 1991). A wide variety of nongovernmental organizations are involved in the publicity campaigns against human bondage. Perhaps the most famous is Anti-Slavery International, which promotes the eradication of slavery and slavery-like practices by supporting the victims of those practices and through the collection and dissemination of information on specific cases. The negative publicity created by the public exposure of servile labor practices can diminish servile labor practices, but governments often do not cooperate because negative publicity alone is a feeble enforcement tool.

Economic pressure may be more effective than the investigative reporting and publicity efforts of the United Nations and other governmental and nongovernmental organizations. For example, the annual reports of the U.S. Department of State on Human Rights Practices are provided to the U.S. Congress to assist it in formulating foreign aid policies (Table 2 is based on the 1999 report). Countries are encouraged to suppress slavery and unfree labor practices to continue receiving aid from the United States and international organizations. Pressure from consumer organizations is effective in some cases. For example, a number of product-labeling programs were created to reassure buyers that products imported into affluent countries were not made by children. A nongovernmental organization in India initiated the RUGMARK program in 1994, which encourages carpet manufactures to stop using child labor by providing a labeling service that certifies that products are not made by children (United States Department of Labor 1997). By mid-1997, RUGMARK inspections had found over 1,000 illegal child workers in the carpet industry. Other labeling programs cover leather footwear, soccer balls, and the tea industry (United States Department of Labor 1997).

The historical decline in the land-labor ratio did not produce the abolition of slavery and involuntary servitude. The simple version of the Nieboer-Domar hypothesis is inadequate. From the middle of the nineteenth century until the present, political factors have played a decisive role in breaking the link between the availability of land and unfree labor. Rather than defending slavery when land



was plentiful, Britain used its political power to abolish slavery and suppress the slave trade even though it was not in its economic interest to do so.

By the late twentieth century, most nations had officially prohibited slavery and slavery-like practices, but both persist. As the modified Niebor-Domar hypothesis predicts, some employers will use physical coercion to drive down the cost of labor when opportunities arise. Weak or corrupt state institutions cannot or will not defend the rights of those who are most vulnerable to coercion. Civil wars destroy the ability of states to maintain order and subject citizens to the deprivations of warring militias. In some areas of the world, cultural support for servile labor and grinding poverty combine to make it difficult to eliminate some forms of slavery and forced labor. In all cases, the typical victims of these practices are the weakest and most vulnerable groups: women, children, migrant workers, low-status class or caste groups, and racial, ethnic, and religious minorities (Table 2).

All nations regulate the passage of individuals across their borders and assign superior rights and privileges to citizens compared to noncitizens. In advanced capitalist democracies with ostensibly free labor markets, the state-enforced distinction between citizen and noncitizen is a key mechanism in maintaining dual labor markets that disproportionately relegate noncitizens to the lowest-paying jobs (e.g., Thomas 1985; Miles 1987; Cohen 1987). Typically, noncitizen "guest workers" are less likely to enjoy state protection and more vulnerable to discrimination. Because the demand for cheap labor often can be satisfied by choosing among citizens and noncitizens who have no other labor market alternatives, democratic states can regulate noncitizens' access to domestic labor markets rather than forcibly import unfree workers from foreign lands.

However, many states are not liberal democracies. Thousands were confined for political reasons in forced labor camps during the Stalin era in the Soviet Union. Nazi Germany forced Jews and other minorities into slavery where they were to be "worked to death" (Sawyer 1986). Blacks were disfranchised and rigidly segregated in the southern United States for much of the twentieth century, making them vulnerable to coercive labor practices. The Republic of South Africa's now

abolished policy of apartheid denied citizenship status to indigenous blacks and exposed them to forced labor practices. Since 1988, the military government of Burma has engaged in systematic human rights abuses, including the imposition of forced labor on large segments of the population for military purposes and for the construction and development of infrastructure (Bureau of International Labor Affairs 1998).

The international condemnation of slavery and involuntary servitude represents a great victory for those who support and defend human rights. The use of forced and other forms of servile labor has not been eliminated, but it has been widely branded as criminal activity. As a consequence, those who would employ servile labor must risk prosecution or search for opportunities in countries with weak or corrupt political institutions that cannot or will not suppress slavery and involuntary servitude. Although the struggle to eradicate servile labor practices has not been won, nations and international human rights organizations appear to be more concerned about the loss of life and other human rights violations that accompany civil wars and international conflicts. For example, NATO recently intervened to stop the murder and forced displacement of thousands of ethnic Albanians in Kosovo. A similar intervention into the affairs of another nation to eliminate slavery or involuntary servitude is unlikely.

#### REFERENCES

- Anti-Slavery International 1999a "Slavery in Sudan—Appeal for Releases." News Release available from the Anti-Slavery International website ([www.charitynet.org/~asi/](http://www.charitynet.org/~asi/)).
- 1999b. "What Is Modern Slavery?" Available from the Anti-Slavery International website ([www.charitynet.org/~asi/](http://www.charitynet.org/~asi/))
- Boogaart, Ernst Van Den, and P. C. Emmer 1986 "Colonialism and Migration: An Overview." In P. C. Emmer, ed., *Colonialism and Migration*. Dordrecht, Netherlands: Martinus Nijhoff.
- Bureau of International Labor Affairs 1998 "Report on Labor Practices in Burma." Washington, D.C.: U.S. Department of Labor.
- Cohen, Robin. 1987. *The New Helots*. Brookfield, Vt.: Gower.
- Cooper, Frederick 1977 *Plantation Slavery on the East Coast of Africa*. New Haven, Conn.: Yale University Press.

- Domar, Evesy D. 1970 "The Causes of Slavery or Serfdom: A Hypothesis." *Journal of Economic History* 30:18–31.
- Effah, Josephine 1996 *Modernised Slavery: Child Trade in Nigeria*. Lagos, Nigeria: Constitutional Rights Project.
- Eltis, David 1983 "Free and Coerced Transatlantic Migrations: Some Comparisons." *American Historical Review* 88:251–280.
- 1987 *Economic Growth and the Ending of the Transatlantic Slave Trade*. New York: Oxford University Press.
- Engerman, Stanley L. 1985 "Economic Change and Contract Labour in the British Caribbean." In D. Richardson, ed., *Abolition and Its Aftermath*. London: Frank Cass.
- 1986a "Slavery and Emancipation in Comparative Perspective: A Look at Some Recent Debates." *Journal of Economic History* 46:317–339.
- 1986b "Servants to Slaves to Servants: Contract Labour and European Expansion." In P. C. Emmer, ed., *Colonialism and Migration*. Dordrecht, Netherlands: Martinus Nijhoff.
- Finley, M. I. 1968 "Slavery." In David Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 14, pp. 307–313. New York: Macmillan.
- Finnigan, William 1999 "The Invisible War." *The New Yorker*, January 25, pp. 50–73.
- Fogel, Robert W. 1989 *Without Consent or Contract*. New York: Norton.
- Galenson, David W. 1981 *White Servitude in Colonial America*. New York: Cambridge University Press.
- and Stanley L. Engerman 1974 *Time on the Cross: The Economics of Negro Slavery*, 2 vols. Boston: Little, Brown.
- Gordon, Murray 1989 *Slavery in the Arab World*. New York: New Amsterdam Books.
- Huggins, Martha K. 1985 *From Slavery to Vagrancy in Brazil*. New Brunswick, N.J.: Rutgers University Press.
- James, David R. 1988 "The Transformation of the Southern Racial State: Class and Race Determinants of Local-State Structures." *American Sociological Review* 53:191–208.
- Kloosterboer, W. 1960. *Involuntary Labour since the Abolition of Slavery*. Leiden, Netherlands: E. J. Brill.
- Kolchin, Peter 1987 *Unfree Labor: American Slavery and Russian Serfdom*. Cambridge, Mass.: Harvard University Press.
- Lovejoy, Paul E. 1983 *Transformations in Slavery*. Cambridge, UK: Cambridge University Press.
- Marx, Karl. (1867) 1967 *Capital*, vol. 1. New York: International Publishers.
- Miles, Robert 1987 *Capitalism and Unfree Labour*. London: Tavistock.
- Morgan, Edmund S. 1975 *American Slavery, American Freedom*. New York: Norton.
- Morris, Richard B. 1946 *Government and Labor in Early America*. New York: Columbia University Press.
- Nieboer, Herman J. 1900 *Slavery as an Industrial System*. The Hague, Netherlands: Martinus Nijhoff.
- Paige, Jeffery 1975 *Agrarian Revolution*. New York: Free Press.
- Patterson, Orlando. 1977a "Slavery." *Annual Review of Sociology* 3:407–449.
- 1977b "The Structural Origins of Slavery: A Critique of the Nieboer Domar Hypothesis." *Annals of the New York Academy of Sciences* 292:12–34.
- 1982 *Slavery and Social Death*. Cambridge, Mass.: Harvard University Press.
- Plant, Roger 1987 *Sugar and Modern Slavery*. London: Zed.
- Roberts, Richard, and Suzanne Miers 1988 "The End of Slavery in Africa." In S. Miers and R. Roberts, eds., *The End of Slavery in Africa*. Madison: University of Wisconsin Press.
- Sawyer, Roger 1986 *Slavery in the Twentieth Century*. London: Routledge and Kegan Paul.
- Slicher Van Bath, B. H. 1986 "The Absence of White Contract Labour in Spanish America during the Colonial Period." In P. C. Emmer, ed., *Colonialism and Migration*. Dordrecht, Netherlands: Martinus Nijhoff.
- Smith, Abbot Emerson 1947 *Colonists in Bondage*. Chapel Hill: University of North Carolina Press.
- Sutton, Alison 1994 *Slavery in Brazil*. London: Anti-Slavery International.
- Thomas, Robert J. 1985 *Citizenship, Gender, and Work*. Berkeley: University of California Press.
- Thornton, John 1998 *Africa and Africans in the Making of the Atlantic World, 1400–1800*, 2nd ed. Cambridge, UK: Cambridge University Press.
- United Nations 1984 "Slavery: Report Prepared by B. Whitaker, Special Rapporteur of the Sub-Commission on Prevention of Discrimination and Protection of Minorities." New York.
- 1988 "United Nations Action in the Field of Human Rights." New York.
- 1991 "Contemporary Forms of Slavery." New York: Center for Human Rights.

- 1997 “Situation of Human Rights in the Sudan.” New York: Commission on Human Rights.
- (1957) 1999 “The Supplementary Convention on the Abolition of Slavery, the Slave Trade, and Institutions and Practices Similar to Slavery,” including a current list of ratifying nations as well as all the major documents and treaties concerning human rights. Available on the United Nations website, [www.unhchr.ch/html/intlinst.htm](http://www.unhchr.ch/html/intlinst.htm).
- United States Department of Labor 1994 *By the Sweat and Toil of Children vol. 1: The Use of Child Labor in American Imports*. Washington, D.C.: Bureau of International Labor Affairs.
- 1995 *By the Sweat and Toil of Children, vol. II: The Use of Child Labor in U.S. Agricultural Imports & Forced and Bonded Labor*. Washington, D.C.: Bureau of International Labor Affairs.
- 1996 *Forced Labor: The Prostitution of Children*. Washington, D.C.: Bureau of International Labor Affairs.
- 1997 *By the Sweat and Toil of Children, vol. IV: Consumer Labels and Child Labor*. Washington, D.C.: Bureau of International Labor Affairs.
- 1998 *By the Sweat and Toil of Children, vol. V: Efforts to Eliminate Child Labor*. Washington, D.C.: Bureau of International Labor Affairs.
- United States Department of State 1999 *Country Reports on Human Rights Practices for 1998*. Washington, D.C.: U.S. Government Printing Office. Available on the State Department website, [www.state.gov/www/global/human\\_rights/1998\\_hrp\\_report](http://www.state.gov/www/global/human_rights/1998_hrp_report).
- Wright, Gavin 1986 *Old South, New South*. New York: Basic Books.

DAVID R. JAMES  
SARA HEILIGER

## SMALL GROUPS

In sociology, the concept “group” implies more than simply an aggregate of individuals. Additional elements involved are (1) structure—interaction patterned in terms of statuses and roles, (2) history—some frequency and regularity of interaction over time, (3) interdependence—some degree of members’ mutual reliance on each other for needed or valued material and nonmaterial resources, and (4) common identity—grounded in shared meanings, values, experiences, and goals. Frequently there is a group product, not necessar-

ily of a material nature, which is the outcome or consequence of collective effort and interaction.

These elements are dimensional in that groups possess and manifest them to a greater or lesser degree. At one extreme, family groups typically have well-established and enduring structures, share extensive histories, encompass a wide range of activities, exert a broad scope of influence, and provide the basis of individual identity. At the other extreme, ad hoc work groups (and groups studied in laboratory experiments) may be assembled to perform specific tasks of very limited duration with little or no relevance for or influence on the members outside a clearly defined situation and range of activity. McGrath (1984) developed a comprehensive typology of groups in terms of origin, scope of activity, task, duration, and interaction.

Groups are regarded as small if meaningful and direct face-to-face interaction can take place among all members. The number of members usually is thought of as ranging from two to twenty, with three to seven common in many laboratory studies of groups.

### PRIMARY AND SECONDARY GROUPS

Cooley (1909) identified a fundamental type of small group that is characterized by intimate association and cooperation, which he regarded as the basic building block of society. Cooley called groups of this sort “primary groups” and held them to be forms of association found everywhere. Primary groups work on the individual to form and develop the social nature of the person. “This nature consists of certain primary social sentiments and attitudes, such as consciousness of one’s self in relation to others, love of approbation, resentment of censure, emulation, and a sense of social right and wrong formed by the standards of a group” (1909, p. 32).

Membership and participation in primary groups are valued and rewarding for their own sake. The groups typically are long-lasting. Members interact as “whole persons” rather than merely in terms of specialized, partial roles. Primary groups are basic sources of socioemotional support and gratification, and participation in them is considered essential for a person’s psychological and emotional well-being. Some (the family, the neigh-

borhood peer group) are also primary in the sense that they are settings for early childhood socialization and personality development.

In contrast are groups formed and maintained to accomplish a task, to which people belong for extrinsic purposes (because they are paid or to achieve an external goal). These “secondary groups” are characterized by limited, instrumental relationships. They may be relatively short-term, and their range of activity is restricted. Affective ties and other “irrational” personal influences are intended to be minimized or eliminated.

It has been widely observed, however, that primary relationships develop pervasively within secondary groups and organizations. In a synthesis of observations and research findings, Homans (1950) attempted to identify universal variables of group behavior. He sought to develop a general theoretical scheme that would permit an understanding of groups as diverse as an industrial work unit, a street-corner gang, and a Polynesian family. Homans approached the small group as a system in which activity, interaction, and sentiment are interrelated. He concluded that interaction among group members increases their liking for one another and that they tend to express their friendship in an increasing range of activities and to interact more frequently. Affective elements emerge in virtually all ongoing groups and may enhance or interfere with the purposes for which a group was established. Soldiers are motivated to fight and workers are motivated to increase or restrict work output by loyalty to their friends and the norms of the immediate group.

#### **BASES AND DEVELOPMENT OF SMALL GROUP RESEARCH**

Sociological interest in small groups has several bases, including (1) the perception of small groups as fundamental, universal social units on which all larger organizational structures depend, (2) a concern with the description and understanding of particular small groups both for their own importance and as a source of observations from which hypotheses and general theories can be developed, and (3) the usefulness of the laboratory group as a research context in which to study the characteristics of the group as the unit of interest and as a setting for the investigation of social influence on individual cognition and behavior.

Foundations for small group research may be seen in nineteenth-century sociological thought, such as Emile Durkheim’s analyses of the development of social structures, specialization and task differentiation, and the bases of social cohesion and Georg Simmel’s work on the importance of group size and coalition formation. Early in the twentieth century Charles H. Cooley and George Herbert Mead stressed the social construction of the self through interaction within immediate group settings.

In the 1930s and 1940s, Jacob L. Moreno developed a systematic approach to the understanding and charting of group structure and Muzafer Sherif conducted key studies of group influence and conformity. William Foote Whyte’s field study of a street-corner gang demonstrated the existence and importance of group norms and structure in an urban milieu generally thought to lack social organization. Of major importance was Kurt Lewin’s work, which provided direction and inspiration for the postwar generation of social psychologists. Lewin combined principles of Gestalt psychology and concepts from the physical sciences to develop field theory in social psychology as a basis for the study of group dynamics. Interested in both theoretical and applied aspects of group interaction, in 1945 he established the first organization devoted to research on group dynamics. The widely utilized sensitivity-training group method originated serendipitously in sessions Lewin organized in 1946.

The period from the end of World War II to the early 1960s produced burgeoning activity in small group research (Hare et al. 1965). The pervasiveness of Lewin’s ideas was evident in the growth of group dynamics as an area of research and theoretical development. Cartwright and Zander’s important compilation, *Group Dynamics* (1968), first published in 1953, presented a theoretical overview and numerous influential studies of cohesiveness, group pressures and standards, individual motives and group goals, leadership and group performance, and the structural properties of groups.

Substantial work with a different orientation reflected concerns with functional needs that groups must meet in order to survive and with the relationship of those functions to dimensions of interpersonal behavior and personality traits. At

the same time, influences from anthropology, economics, and behavioral psychology were being melded in a view of social interaction as an exchange of resources, a perspective applied to the analysis of interdependence, cooperation and competition, and interpersonal relationship (Homans 1950, 1974; Thibaut and Kelley 1959). During those years small group research shared the methodological advances that were occurring throughout the social sciences, developing an increasing sophistication in research design, measurement, and analysis. The excitement, optimism, and productivity of the field led some to define social psychology as the study of small groups.

Small group research since the 1960s has not been as prominent, prolific, or influential as it was during the immediate postwar years, when social psychology was virtually dominated by the small groups "movement" (Borgatta 1981). The production of studies is steady, if moderate compared to the enthusiasm of the peak period, and some significant attempts have been made to organize and integrate the diverse body of work and theory that has accumulated (Hare 1982; McGrath 1984; Foschi and Lawler 1994). There is renewed interest in conceptualizing groups as entities with distinctive properties that cannot be understood in terms of reductionist individual psychology (Turner 1987). Many aspects and procedures of group process and dynamics are commonly utilized in applied settings, while practical concerns with group productivity, efficiency, and success are widespread (Hare et al. 1992; Forsyth 1999).

#### APPROACHES TO SMALL GROUP RESEARCH

Small group studies are characterized by a wide variety of research techniques and theoretical and practical concerns. Research methods vary in regard to the types of groups and circumstances studied—whether "natural" or contrived for research purposes—and in the intrusiveness of research procedures. Some investigators are concerned with properties of the group itself as the unit of interest, while others use the small group setting as context for exploring individual behavior. Although laboratory studies have predominated, the research techniques employed include direct observation of groups in natural as well as controlled settings; the use of structured observa-

tional systems to code communication or other aspects of behavior; the use of checklists, questionnaires, or interviews to elicit ratings, choices, opinions, or attitudes from group members; and field experimentation.

Laboratory studies have marked advantages in terms of the control and manipulation of variables in the precision of observation and measurement. The procedures employed normally permit replication of observation under controlled conditions. The experimental method is regarded as superior to others for rigorously testing causal hypotheses. Fundamental technical issues are whether relevant variables can be brought into laboratory situations and whether a meaningful range of variation can be achieved.

Criticisms of laboratory research center on the artificiality of the setting and the short-term nature of most studies. Representativeness of subject groups and thus the generalization of the findings also are questioned. Concerns for protecting the rights and well-being of human subjects have led to procedural safeguards that now inhibit or prevent practices that were typical of some well-known earlier studies.

The technical advantages of laboratory procedures, the desire to emulate the natural sciences in developing theory based on experimental evidence, and the compatibility of laboratory methods with the academic environment within which most researchers work all have contributed to the proliferation of laboratory studies that constitute much of small group research.

Direct observation of group behavior under basically uncontrolled ("natural") conditions may be coupled with the investigator's more or less active participation in the affairs of the group. Such research can employ structured systems for coding behavior and interaction patterns that are used by uninvolved "objective" observers, as when a children's play group is studied by adults. A more informal ethnographic approach was employed by Goffman (1964) in collecting the information that illustrated his characterization of human interaction as an elaborate sequence of symbolic presentations of self and groups as collaborating teams of performers.

Participant observation is a procedure in which the researcher acts as part of a (usually natural)

group to understand a situation from within, as members of the group define and experience it. Group members may know that the observer is an outsider who is there for his or her own purposes or may be led or allowed to believe that the observer is simply another “genuine” group member. In either case the observer’s status influences and constrains both the kinds and amount of information available and the opportunities for recording information. The observer also has some influence on the situations and processes being studied, thus producing outcomes different from those which would have occurred in his or her absence. The use of multiple observers increases opportunities for observation while also increasing the effect of the research on the group’s behavior (Festinger et al. 1964). For these reasons, reliability and validity are particularly problematic issues in using this technique.

Participant observation is regarded as useful primarily for descriptive and exploratory research and for generating or illustrating, as opposed to testing, theory. It is favored by those who want to understand the meanings of situations and actions generated and maintained by groups in their natural, everyday environments.

An important naturalistic study was conducted in the late 1930s by Whyte (1955), who studied a street-corner gang as a participant observer over a period of three and a half years. (The appendix to his monograph provides an informative discussion of practical and ethical issues in participant observation.) Whyte gained access to the gang through his association with its leader, and his view is from the top of the social structure. He described the recurrent patterns of relationships among members, group values and codes of behavior, the existence of implicit exchange relationships, territorial behavior, and the nature and functions of gang leadership. His observations of the ways in which members’ social rankings in the group affected their performance in athletic competition suggested a program of experimental studies of diffuse status characteristics: an exploration of the manner in which “logically” irrelevant social rank affects the amount of influence an individual has on others in activities ranging from pedestrian behavior to the making of perceptual judgments.

Sociometry, a seminal form of network analysis developed by Moreno (1953), is a technique for

eliciting and representing the patterns and structure of choices and liking among group members. While the most common procedure is for researchers to ask group members who they like, dislike, would prefer to work with, or would like to “be like,” ratings also can be based on direct observations of members’ behavior. The information can be represented as a sociogram showing individuals as circles and choices as arrows between the circles: The diagram depicts group structure in terms of affective relations. Indices of liking or disliking can be computed for each member, and ratings can be organized in a matrix format. The density and patterning of choices may be taken as indicators of group cohesiveness. In practical applications, sociometric data are used to restructure groups on the basis of members’ mutual choices.

**Interaction Process Research.** A prominent research concern has been the description and analysis of group interaction processes, focusing primarily on communication. The approaches employed have ranged from purely formal examination of the amount of communication sent and received by each member of the group to extremely detailed analyses of linguistic and paralinguistic material, including posture, gestures, and inflection.

The widely used system for Interaction Process Analysis (IPA) developed by Bales (1950, 1970) involves a set of twelve categories for coding units (acts) of communication. The categories reflect Bales’s conclusion that all groups confront two domains of concerns: instrumental concerns related to the task the group must accomplish and expressive concerns associated with the socioemotional needs and interrelationships of the group members. Both sets of concerns operate continuously and must be dealt with if a group is to succeed and survive, and there is a virtually constant conflict between them. The set of categories is used by observers to code types of active and passive task-related acts and positive and negative socioemotional acts, as they are generated by group members in the course of interaction.

Numerous studies using the IPA system have sought to document the patterns or “phase movements” of instrumental and expressive communication as groups try to establish the equilibrium necessary to operate. Interaction process scores

have been related to personality characteristics and to peer assessments and self-assessments (Borgatta 1962). Attention also has been paid to the roles of particular group members in exercising task leadership or socioemotional leadership.

The division of group leadership into instrumental and expressive functions proved compatible with accepted notions of "typical" male and female personal attributes and with a conceptualization of the family (at least in the Western world) as a small group with the father as task leader and the mother as socioemotional specialist. However, recent research comparing "natural" families with ad hoc laboratory groups indicates that the instrumental versus expressive specialization found in the laboratory seldom holds for groups in natural settings. There is greater diversity of behavior and less gender-linked stereotypical conduct in longer-lasting groups that cover a greater scope of activities (McGrath 1984).

The IPA system has been criticized on both theoretical and operational grounds, and numerous revisions and alternatives have been proposed. Bales and his colleagues developed an elaborated observational system, SYMLOG (Bales and Cohen 1979), that models personal space in three dimensions: dominant-submissive, friendly-unfriendly, and instrumentally controlled-emotionally expressive. Group interaction is observed and members' behaviors are coded on each dimension by outside observers or by the group members themselves. On the basis of combinations of multiple observations, each individual is located within the three-dimensional space and the positions of all group members are charted. The resulting diagram and indices based on the scores indicate the degree to which members are perceived as acting in a similar fashion. Interest in the SYMLOG technique is substantial, and it is utilized in many studies of group structure and performance.

### GROUP COHESIVENESS

The understanding of what holds a social unit together, a central issue in sociology, also has been central in small group analysis. Cohesion—the sum of the forces that bind members to the group—was viewed by Lewin and other Gestaltists as a property or characteristic of the group itself, a sort of force field analogous to a magnetic or gravitational field. However, the assessment of cohesion

usually depends on observations of the attitudes and behaviors of the individual group members: their self-reported attraction to the group, their feeling of being accepted by the group, similarity in expressions of sentiment, how regularly they attend group meetings, how prompt or tardy they are, or how responsible they are in performing actions that benefit the group. Members also may be asked to describe the unity of the group (Evans and Jarvis 1986; Bollen and Hoyle 1990). Although Steiner (1972) suggested that "A true test of a group's cohesion would entail observation of its members' reaction to disruptive influences," he rejected this procedure on technical and ethical grounds (1972, p. 161).

The bases of cohesion include (1) rewards available within and through the group, (2) the congruence between individual goals and group goals, (3) the attraction and/or liking of members for each other, (4) the importance of the group as a source or ground of the individual's identity and self-perception and his or her internalization of group culture and values, and (5) in psychoanalytic group theory, the members' identification with and attraction to the group leader and "the alignment between particular individual superego formation and its corresponding punitive group structure" (Kellerman 1981, p. 11).

Although high cohesiveness often is taken as indicating a "healthy" group, its effect is to heighten members' susceptibility to influences in the group. Thus, group productivity, for example, may be increased or decreased depending on the nature of the predominant influences. A positive association between group cohesion and performance was found by Evans and Dion (1991) in a review of previous studies, but the relationship was modest.

Major importance in the study of cohesiveness has been placed on interpersonal attraction and interdependence, emphasizing the exchange of emotional and affective resources. Work by Tajfel (1981) and Turner (1987) supports, alternatively, an emphasis on social identity and self-categorization. A concept of cohesion based on interpersonal liking that is not mediated by shared group membership and depersonalized attraction to the group is held to be inadequate. Group membership and the resultant self-categorization occur prior to interaction and the emergence of interdependence, cooperation, influence, and cohesion.

Hogg (1987, 1992) advocates research that will produce answers about group solidarity and social identity rather than about interpersonal relationships.

Self-categorization in a most elemental form has been demonstrated in “minimal group” experiments (Tajfel 1981). Subjects are divided into two groups, sometimes presumably on the basis of an arbitrary and unimportant criterion and sometimes in an obviously random manner. The participants do not interact within or between groups during the experiment. Given the task of dividing a sum of money between two persons about whom they know nothing except their group membership, subjects show a marked bias in favor of members of their own group.

Interdependence in a most elemental form has been realized in experiments with the “minimal social situation” (Sidowski 1957). Two subjects, each of whom controls resources that may reward or punish the other and each of whom depends primarily on the other’s behavior as a source of reward or punishment, learn to exchange rewards despite being completely unaware of the nature of the situation.

Thibaut and Kelley (1959) identified two criteria individuals use in evaluating the rewards available within a particular situation: a usual, expected level of reward to which the person feels entitled, called the “comparison level,” and the person’s perceived best level of reward available outside the situation, called the “comparison level for alternatives.” An individual’s satisfaction with his or her group membership and participation depends on the relationship of rewards available within the group to his or her comparison level, while the likelihood that one will stay in or leave a group depends on the comparison level for alternatives.

Although the value and availability of rewards are usually emphasized in assessing the attractiveness of a group, Leon Festinger has pointed out the persistence of loyalty to “lost causes” and the effect that insufficient reward, or even aversive experiences, can have in strengthening members’ positive attitudes. In one experiment (Aronson and Mills 1959), potential group members who were subjected to a severe initiation expressed greater liking for the group than did those who had a mild initiation. And while an equitable and balanced exchange of rewarding outcomes is considered important in sustaining interpersonal rela-

tionships and participants’ satisfaction with them, Kelley and Thibaut (1978) noted that problematic situations provide particular opportunities. Attributions about a partner’s personality and motivations and self-presentations that encode messages of commitment and concern for the other person are facilitated when behavior cannot be explained simply in terms of “rational” self-interest. Such attributions and encodings strengthen affective ties and promote interdependence of the characteristics and attitudes displayed in the relationship.

### GROUP INFLUENCE

**Social Facilitation and Inhibition.** In a study credited as the first social psychological experiment (1897), Triplett measured the average time his subjects took to wind 150 turns on a fishing reel, working both alone and in competition with one another. Subjects working in competition wound the reels faster than did those working alone. Numerous subsequent experiments (including some with nonhuman subjects) have supported and modified these results. It was found that the mere presence of other persons (as observers or coactors, whether or not they were competitors) facilitated well-learned responses but that the presence of others interfered with the acquisition of new responses. This “audience effect” thus facilitates performance but inhibits learning. Various explanations of social facilitation and inhibition have been proposed, generally incorporating the idea that the presence of others increases motivational arousal. Such arousal is a basic feature of the group environment (Zajonc 1966).

**Conformity.** Similarities of values, attitudes, beliefs, perceptions, and behavior are a ubiquitous and virtually defining feature of group existence. These similarities can facilitate coordination of goal-directed activity, motivate the members, provide sources of psychological security and emotional reward, reinforce members’ identification with the group, and increase cohesiveness. They also may prevent reasoned consideration of alternatives to group decisions and the potential consequences of group actions, reduce flexibility in adapting to new circumstances, and inhibit change in general. Closed circles of conformity in cohesive groups that are isolated from dissenting view-



points, producing “groupthink” (Janis 1982), have been implicated in producing military blunders, fascist atrocities, government scandals, and space shuttle disasters. Conformity (to modeled indifference or uncertainty) is a factor in the failure of bystanders to help others in emergencies.

The amount of conformity in a group may be seen as a characteristic of the collectivity. Experimental studies, however, usually have been concerned with effects on the individual. Considered from this viewpoint, conformity is defined as a change in an individual’s attitudes, beliefs, or behavior in the direction of a group norm. It is an example of social control resulting from peer influence (as distinct from, for example, obedience to a constituted authority) (Milgram 1974). Two types of conformity have been identified: belief (or informational) conformity and behavioral (or normative) conformity. Both types are increased by strong group cohesiveness.

Belief conformity involves an internalized and lasting change grounded in an individual’s dependence on social sources of information and guidance. Once they are internalized, the group’s standards and perceptions are constantly carried with the individual and constitute an ongoing element of social control.

Sherif (1935) asked individual subjects to judge the apparent movement of a pinpoint of light in an otherwise totally dark room. Under these conditions the light, which in fact was stationary, appeared to most people to move. Different individuals perceived different amounts of movement. Assembled in small groups viewing the light together, the subjects began to agree on the amount of movement they perceived: A group norm emerged in an ambiguous situation. After the group interaction, subjects were asked to view the light, again in isolation. They continued to see the amount of movement agreed on by the group rather than the amount they originally perceived individually. The group’s perceptions apparently had been internalized.

The strength of belief conformity varies with the ambiguity and unfamiliarity of the situation, the individual’s trust in the credibility of the group, the individual’s attraction to and identification with the group, and the individual’s prior experience and confidence.

Behavioral conformity is grounded in the potential rewards and punishments dispensed by the group and in the individual’s previous experience with the consequences of conformity and nonconformity. The consequences of agreeing with others’ judgments and opinions, emulating others’ behaviors, and following the customs of a group are usually pleasant, while disagreement and deviancy generally lead to unpleasant effects. Group members who hold deviant opinions typically receive, at first, greater than normal amounts of communication in an attempt to influence them to conform. If these efforts fail, they are likely to be isolated or rejected, depending on the severity of the deviance. Monitoring of behavior is necessary if reward or punishment is to depend on its occurrence; thus this type of influence is effective only if and when an individual’s actions are known to the group.

Experiments conducted by Asch (1951) demonstrated behavioral conformity. The subjects engaged in a perceptual estimation task that required them to pick out lines of the same length printed on boards that were presented side by side. The boards were presented in pairs, and the judgment of each pair constituted one experimental trial. In a typical experiment there was only one genuine subject; the other participants were employed by Asch, and their judgments were prearranged. After a number of trials in which correct judgments were given, the confederate “subjects” began stating unanimous wrong judgments. The genuine subjects conformed to a substantial extent by expressing judgments that agreed with those of the group. When removed from the group or allowed to state judgments in private, the real subjects did not persist in making these errors. Their conformity occurred only when it was witnessed by the other group members.

The strength of behavioral conformity varies with the size and unanimity of the group, the importance of the group to the individual, and the disclosure of relevant judgments or behaviors to the group.

Belief and behavioral conformity can be distinguished analytically (and empirically under some laboratory conditions), but in natural situations they operate in conjunction. The group member not only is rewarded for conforming but also depends on others as models for behavior and

guides for judgments and opinions. While it is common to think of beliefs and attitudes as existing before the behaviors that reflect them, a large body of research indicates that people come to believe the opinions they express. "Mere" behavioral conformity can lead to internalization.

Conformity effects usually are thought to reflect the majority influence in a group, but evidence shows that a determined minority can prevail. Minority influence seems especially relevant in regard to internalization (Moscovici 1980).

**Group Polarization.** Early theories of "group contagion" and the madness of crowds notwithstanding, a general assumption has been that conformity processes within a group operate to bring extreme opinions and judgments in toward the center of the range of opinions and judgments. However, a body of research has contradicted the notion that group actions are more moderate than those of individuals.

The experimental procedure called for individual subjects to evaluate each of twelve "choice dilemmas," situations in which a person was asked to choose between a highly desirable risky alternative and a less desirable but certain alternative. The subjects were instructed to indicate for each dilemma, the lowest probability of success they would accept in recommending that the desirable risky alternative be chosen. Probabilities were averaged for each subject over all dilemmas to generate a "riskiness" score for that person. Small groups of subjects were then formed and instructed to discuss each situation, reach a group decision, and indicate the group riskiness score for the dilemma. A group's scores were averaged over the twelve situations, and that value was compared to the mean of the individual scores of the group members.

Initial research that employed the choice dilemmas procedure found a significant "risky shift" in the group decisions compared to the mean of the individual scores. Numerous experiments and further analyses followed that extended and qualified those findings (Cartwright 1973). Certain kinds of choice dilemma scenarios produced risky shifts, while others produced conservative shifts or showed no significant difference. Shifts tended to move in the direction of the initial inclinations of the group: The interaction resulted in a collective outcome more extreme than might have been predicted on

the basis of the individual positions, but the individual positions forecast the nature of the shift.

Group polarization, as the effect is now called, has been theoretically interpreted in terms of risk as a cultural value, the persuasive influence of "risky" individuals, and the diffusion of responsibility in group action. However, the effect can be explained as being due to the normative and informational influences involved in conformity processes (McGrath 1984).

### GROUP INTERACTION AND PERFORMANCE

Group performance in terms of problem solving, productivity, or effectiveness is a subject of both practical and theoretical concern that has generated numerous studies and a large body of theory. Productivity may refer to the quality of a group product, the efficiency of output per unit time or progress toward a group goal, or the realization of group potential. The establishment of an appropriate basis of evaluation is often problematic, and expected outcomes depend heavily on the type of task undertaken. When groups fall short of what (from some standpoint) it is felt they should accomplish, the failure often is attributed to "process losses" resulting from problems in interaction.

Steiner (1972) distinguished between tasks that require a coordinated division of effort, which he labeled "divisible," and those with a single outcome or product, which he called "unitary." Disjunctive unitary tasks are those which can be accomplished successfully by one individual alone. In such cases the group should be as "good" as the best member. Conjunctive unitary tasks require all the members to contribute successfully; in these tasks, the group can be only as good as the worst member. Tasks in which members' contributions are simply summed to produce the group outcome are called additive, and group performance should depend on the "average" member. Numerous laboratory studies of ad hoc groups performing a wide range of judgment tasks have been conducted. Overall, the results indicate that groups seldom do as well as the best member but usually do better than the average member.

Field studies of industrial workers in natural settings illustrate how influence processes in the group can regulate behavior. Work groups de-

velop norms with respect to what they, not the company, regard as an appropriate day's output. While pay, potential promotion, and retention or termination may be controlled by the employer, the immediate group controls powerful social rewards and sanctions that are brought to bear on a day-to-day basis. Those who exceed the group's production norm ("rate busters") and those who fail to produce an acceptable amount or attain an acceptable standard of quality are subjected to group pressure ranging from "kidding" and mild criticism to serious harassment. Since group cohesiveness increases conformity, some companies find it desirable to move workers frequently and attempt in other ways to inhibit the formation of interpersonal ties and identification with the group. Other organizations attempt to mobilize small group processes to support their goals.

Successful performance requires that a group have the necessary resources (material resources and members' skills, knowledge, and competencies) and time needed to accomplish its tasks. In addition, issues of coordination and motivation arise. When it confronts a disjunctive unitary task, the group simply must assure that the "best" member has the opportunity, recognition, and authorization to function and is motivated to do so. The only coordination needed may be to prevent interference from other members. For other types of tasks the quality, sequence, and articulation of many or all members' contributions are important (Miller and Hamblin 1963).

Allocation of opportunity to participate and evaluation of members' actions constitute elements of the status structures of groups. The explanation of how interaction inequalities in task groups are developed and maintained is the concern of expectation states theory (Foschi 1997; Foschi and Lawler 1994).

Group members hold expectations about the nature, quality, and value of each other's performances. Their expectations influence the quality of those performances and affect the evaluation of performances after they occur; they are in this sense self-fulfilling prophecies.

Though expectations may derive from firsthand task experience within the group, they also are based on "external" status characteristics of the members. Diffuse status characteristics such as age, race, gender, or perceived social rank may

influence expectations whether or not they are objectively relevant to the group's task and goals. Inequalities in participation, evaluation of performance outputs, and influence over the group's decisions reflect inequalities in status characteristics that members bring to the group. These inequalities tend to be maintained within the group regardless of their pertinence. Evaluations of performance output depend on previous evaluations, and expectations that arise out of group interaction influence subsequent interaction to produce their own confirmation (Berger et al. 1972, 1980). Thus the degree of influence exerted by group members and the impact of their contributions to the group's effort may not be highly correlated with their task-related competence and abilities.

Processes of influence and conformity may degrade performance quality. Majorities generate social pressure whether or not they are competent. Techniques have been devised to control these effects by regulating the kind of interaction that can take place. Some procedures, such as Multi-Attribute Utility Analysis, require the clear identification of task elements and their accomplishment in specified sequences. Group members can interact freely but must adhere to the task stages. Other approaches impose rules for communication in decision-making processes.

Many studies have been concerned with evaluating the effects of different patterns of interaction, primarily communication, on performance. "Brainstorming," a group interaction technique in which members generate as many ideas as possible within a given time period without evaluation or criticism, is intended to overcome inhibiting social influence processes while taking advantage of those which stimulate creativity. However, research indicates that a brainstorming group is generally less effective in producing ideas than are the same number of individuals working alone. The Delphi Method and the Nominal Group Technique are two approaches to regulating the combination of individual effort with group feedback or interaction to reduce the deleterious effects of social relations within the group, conformity, and personalized conflict. In the Delphi Method individuals, without communicating with each other, make judgments that are combined into a group "product." The results are made known to the participants, who then make another round of judgments. This procedure is repeated until a final

group judgment is attained. The Nominal Group Technique, which is used for developing plans or ideas or for choosing a correct or best solution, begins by having individuals work separately to generate plans, ideas, or judgments. The group then collectively lists and evaluates the material that was produced individually. These methods have advocates, but the desirability of some of their results is questionable and the time and cost involved in their utilization may be significant (McGrath 1984).

A substantial body of research has compared the relative effectiveness of structured networks of communication available to members of problem-solving groups. Typically, groups of three to five persons were required to combine information distributed across the individual members, communicating only through channels provided by the experimenters. Various networks of communication channels have been investigated to see how they affect a group's efficiency and the members' satisfaction. The networks differ in terms of how centralized or open they are. The most centralized network compels all messages to flow through one position, while the most open permits direct communication among all the members.

The conclusions from this research are that centralized networks are most efficient in dealing with simple tasks but that group members tend to be dissatisfied, except for the person occupying the central position. In more complex tasks the advantages of centralization are lost. Burgess (1969) suggested that the network experiments were basically flawed in failing to provide meaningful consequences for group performance and in studying groups only for brief periods, while they were learning to use the networks. His research demonstrated that when subjects had enough time to learn to use the channels provided, and received rewards based on performance, the type of network made no difference. Given time and motivation, groups adapted efficiently to overcome the structural constraints.

A type of process loss observed in both physical and cognitive tasks is the reduction in effort people put into group performance compared with the effort they make when working individually. This effect, called social loafing, has been observed in numerous cultures and is related to group size: As groups get larger, individual effort

tends to diminish. Social impact theory explains social loafing in terms of a conflict between a person's sense of responsibility and his or her feeling that inaction is the safest or least costly course of behavior. Diffusion of personal responsibility occurs in group situations and reduces the blame for inaction. Also, a lack of individual identification and the absence of evaluation by others become more likely as group size increases. Thus both rewards for effort and punishment for lack of effort become less certain and less consistent. While laziness and "goldbricking" often occur in individual situations, those behaviors can be concealed more easily in a crowd. Nonetheless, members who identify with and value a group and hold strongly to its norms will exert effort on behalf of the group (Hogg 1992). As was noted above, there is a positive association between group cohesion and productivity in groups with norms that support good performance. This relationship is reciprocal: Group success tends to increase cohesion.

The social loafing effect is confounded with problems of coordination, since both increase as the number of persons involved in a task gets larger. In addition, members' impatience and/or frustration with coordination problems can undermine their motivation and sense of responsibility, exacerbating social loafing.

Although researchers have paid much attention to process loss, there also are many significant process gains in group interaction. Social facilitation, stimulation, learning, socioemotional support, development and reinforcement of identity, and even conformity processes can enhance creativity, productivity, and effectiveness. While questions of individual versus group superiority may be provocative, most human endeavor occurs in group contexts and requires group effort.

The recognition that productivity can be affected substantially by the functions and quality of leadership in a group has stimulated much research and theorizing about leadership styles and effectiveness. An early experimental program conducted by Lewin et al. (1939) systematically varied the behavior of adult leaders of clubs of 11-year-old boys engaged in craft work and recreational activity. The leaders were trained to enact democratic, authoritarian (autocratic), and laissez-faire styles of leadership, and those styles were experienced by each club for several weeks. The resulting

changes in the boys' task performance, social relationships and interaction, and motivation as well as some aspects of group "climate" were intensively documented and analyzed. The results of the study, while complex, generally favored the democratic leadership style both for producing increased motivation and originality and for fostering more mutual friendliness and group-mindedness among the boys. Group members preferred the democratic leader to either the autocratic or the laissez-faire leader.

A contingency model of leadership effectiveness has been developed by Fiedler (1981), whose concepts of task-motivated versus relationship-motivated style recall Bales's identification of instrumental and socioemotional leadership functions. A leader's effectiveness results from the combination of style and situation: Task leaders are most effective in situations that are either highly favorable or unfavorable, while relationship leaders are most effective in middle-range situations. The contingency model, though supported by a large body of research, is questioned by some who feel that effective leaders are those who deal with both task and relationship elements of group situations.

### COOPERATION AND COMPETITION IN GROUPS

Two different orientations are evident in research on competition and cooperation within groups. In one approach cooperation and competition are treated as imposed external conditions that influence the quality of group interaction and task performance. Alternatively, cooperation and competition have been studied as dependent behaviors that are affected by reward and risk contingencies, the availability of communication, and other situational factors.

Numerous studies have compared the productivity and efficiency of groups working under cooperative conditions (defined as working for group goals) and competitive conditions (defined as working for individual goals). The concept of cooperation in early research usually specified only mutual dependency of outcomes, with little attention paid to the interdependency of the members' task activities. The findings indicated that the efficiency of work under competition was greater than that under cooperation for tasks that did not require coordination of effort. Some research indicated

that cooperative groups worked together more frequently and were more highly coordinated.

Analysis of research focusing on the nature of tasks used as criteria in comparing cooperative and competitive reward structures points to the importance of "means interdependence," the degree to which group members are reliant on one another (Schmitt 1981, 1998). When tasks are simple, requiring no division of labor or sharing of information or resources, the advantage of cooperative contingencies seems to hold. However, cooperative contingencies are typically superior for tasks high in means interdependence involving distribution of effort, coordination of responses, or information sharing. Also, the long-term consequences of different reward structures may be substantial in natural groups, involving issues of morale and sustained member motivation that seldom arise in relatively brief laboratory studies. Such consequences can depend on the way in which competitive payoffs are determined; the possibility that some group members may become perpetual "losers" while others are constant winners will affect the efforts of all the members.

In a number of cases an additional element of competition between groups has been found to increase the productivity of internally cooperative groups. Turner (1987) observed that competition (for mutual distinctiveness) can develop between groups even in the absence of conflicts of interest. This striving to enhance positive social identity increases group cohesiveness and solidarity, making cooperation more likely.

Laboratory research treating cooperation as a dependent effect has focused on the participants' choice of cooperative rather than competitive behaviors and the distribution and coordination of responses. The effects of threat and communication were investigated in a well-known "trucking game" study (Deutsch and Krauss 1962). Two subjects could cooperate by taking turns using a "short route" to reach a destination and thus make money. Cooperation was reduced when one subject could block the route with a gate ("unilateral threat") and was extremely rare when both subjects had gates ("bilateral threat"). Communication between subjects did not increase cooperation under the threat conditions.

Communication sometimes has been found to increase cooperation in some of the many studies

using the "Prisoner's Dilemma." In this situation, two participants benefit moderately if both choose to cooperate and lose substantially if both "defect." If either one chooses to cooperate while the other defects, the cooperator suffers a very large loss and the defector's outcome is highly favorable. Thus, cooperation involves risk while defection implies motives of self-protection, exploitation, or both. The structure of outcomes is paradoxical: The rational choices of each individual lead to poor collective consequences.

The rates of cooperation observed in these studies are low. The Prisoner's Dilemma epitomizes the class of situations called social traps, in which individual (usually short-term) "rational" self-interest conflicts with the (usually longer-term) well-being of the group, leading to collective irrationality (Kollock 1998).

Inequity of outcomes and the presence of risk have been found to reduce cooperation across a wide range of experimental research (Marwell and Schmitt 1975). Beneficial effects of communication were dependent on the timing of its availability and the pattern of behavior that had occurred before communication took place.

Studies of cooperation and competition have addressed problems of motivation and coordination, issues of equity, the effects of short-term and long-term consequences, and the relationship of individual outcomes to collective outcomes. The analysis of these topics is a notable feature of recent small group research, particularly as concern with social traps and dilemmas resonates with the environmental and social issues facing contemporary society.

## REFERENCES

- Aronson, Elliot, and Judson Mills 1959 "The Effect of Severity of Initiation on Liking for a Group." *Journal of Abnormal and Social Psychology* 59:177-181.
- Asch, Solomon E. 1951 "Effects of Group Pressure upon the Modification and Distortion of Judgments." In H. Guetzkow, ed., *Groups, Leadership, and Men*. Pittsburgh: Carnegie.
- Bales, Robert F. 1950 *Interaction Process Analysis: A Method for the Study of Small Groups*. Cambridge, Mass.: Addison-Wesley.
- 1970 *Personality and Interpersonal Behavior*. New York: Holt, Rinehart and Winston.
- , and Stephen P. Cohen 1979 *SYMLOG: A System for the Multiple Level Observation of Groups*. New York: Free Press.
- Berger, Joseph, Bernard P. Cohen, and Morris Zelditch, Jr. 1972 "Status Characteristics and Social Interaction." *American Sociological Review* 37:241-255.
- , Susan J. Rosenholtz, and Morris Zelditch, Jr. 1980 "Status Organizing Processes." *Annual Review of Sociology* 6:479-508.
- Bollen, Kenneth A., and Rick H. Hoyle 1990 "Perceived Cohesiveness: A Conceptual and Empirical Examination." *Social Forces* 69:479-504.
- Borgatta, Edgar F. 1962 "A Systematic Study of Interaction Process Scores, Peer and Self-Assessments, Personality and Other Variables." *Genetic Psychology Monographs* 65:269-290.
- 1981 "The Small Groups Movement: Historical Notes." *American Behavioral Scientist* 24:607-618.
- Burgess, Robert L. 1969 "Communication Networks: An Experimental Evaluation." In R. L. Burgess and D. Bushell, Jr., eds., *Behavioral Sociology: The Experimental Analysis of Social Process*. New York: Columbia University Press.
- Cartwright, Dorwin 1973 "Determinants of Scientific Progress: The Case of Research on the Risky Shift." *American Psychologist* 28:222-231.
- , and Alvin Zander, eds., 1968 *Group Dynamics: Research and Theory*, 3rd ed. New York: Harper & Row.
- Cooley, Charles H. 1909 *Social Organization*. New York: Charles Scribner's Sons.
- Deutsch, Morton, and Robert M. Krauss 1962 "Studies of Interpersonal Bargaining." *Journal of Conflict Resolution* 6:52-76.
- Evans, Charles R., and Kenneth L. Dion 1991 "Group Cohesion and Performance: A Meta-Analysis." *Small Group Research* 22(2):175-186.
- Evans, Nancy J., and Paul A. Jarvis 1986 "The Group Attitude Scale: A Measure of Attraction to Group." *Small Group Behavior* 17:203-216.
- Festinger, Leon, Henry W. Riecken, and Stanley Schachter 1964 *When Prophecy Fails*. New York: Harper & Row.
- Fiedler, Fred E. 1981 "Leadership Effectiveness." *American Behavioral Scientist* 24:619-632.
- Forsyth, Donelson R. 1999 *Group Dynamics*, 3rd ed. Belmont, Calif.: Brooks/Cole-Wadsworth.
- Foschi, Martha 1997 "On Scope Conditions." *Small Group Research* 28(4):535-555.
- , and Edward J. Lawler, eds. 1994 *Group Processes: Sociological Analyses*. Chicago: Nelson-Hall.

- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, N.Y.: Doubleday.
- Hare, A. Paul 1982 *Creativity in Small Groups*. Beverly Hills, Calif.: Sage.
- , Herbert H. Blumberg, Martin F. Davies, and M. Valerie Kent 1992 *Small Group Research: A Handbook*. Norwood, N.J.: Ablex.
- , Edgar F. Borgatta, and Robert F. Bales, eds. 1965 *Small Groups: Studies in Social Interaction*, rev. ed. New York: Knopf.
- Hogg, Michael A. 1987 "Social Identity and Group Cohesiveness." In J. C. Turner, ed., *Rediscovering the Social Group*. Oxford, UK: Basil Blackwell.
- , 1992 *The Social Psychology of Group Cohesiveness: From Attraction to Social Identity*. New York: New York University Press.
- Homans, George C. 1950 *The Human Group*. New York: Harcourt, Brace.
- 1974 *Social Behavior: Its Elementary Forms*, rev. ed. New York: Harcourt Brace Jovanovich.
- Janis, Irving L. 1982 *Groupthink: Psychological Studies of Policy Decisions and Fiascos*, 2nd ed. Boston: Houghton-Mifflin.
- Kellerman, Henry 1981 "The Deep Structures of Group Cohesion." In H. Kellerman, ed., *Group Cohesion: Theoretical and Clinical Perspectives*. New York: Grune & Stratton.
- Kelley, Harold H., and John W. Thibaut 1978 *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.
- Kollock, Peter 1998 "Social Dilemmas: The Anatomy of Cooperation." *Annual Review of Sociology* 24:183–214.
- Lewin, Kurt, Ronald Lippit, and Ralph K. White 1939 "Patterns of Aggressive Behavior in Experimentally Created 'Social Environments.'" *Journal of Social Psychology* 10:271–299.
- Marwell, Gerald, and David R. Schmitt 1975 *Cooperation: An Experimental Analysis*. New York: Academic Press.
- McGrath, Joseph E. 1984 *Groups: Interaction and Performance*. Englewood Cliffs, N.J.: Prentice-Hall.
- Milgram, Stanley 1974 *Obedience to Authority*. New York: Harper & Row.
- Miller, L. Keith, and Robert L. Hamblin 1963 "Interdependence, Differential Rewarding, and Productivity." *American Sociological Review* 28:768–778.
- Moreno, Jacob L. 1953 *Who Shall Survive?* rev. ed. Beacon, N.Y.: Beacon House.
- Moscovici, Serge 1980 "Toward a Theory of Conversion Behavior." In L. Berkowitz, ed., *Advances In Experimental Social Psychology*, vol. 13. New York: Academic Press.
- Ofshe, Richard A., ed. 1973 *Interpersonal Behavior in Small Groups*. Englewood Cliffs, N.J.: Prentice-Hall.
- Olmstead, Michael, and A. Paul Hare 1978 *The Small Group*, 2nd ed. New York: Random House.
- Schmitt, David R. 1981 "Performance under Cooperation or Competition." *American Behavioral Scientist* 24:649–679.
- , 1998 "Effects of Reward Distribution and Performance Feedback on Competitive Responding." *Journal of the Experimental Analysis of Behavior* 69:263–273.
- Sherif, Muzafer 1935 "A Study of Some Social Factors in Perception." *Archives of Psychology* 27, no. 187.
- Sidowski, Joseph B. 1957 "Reward and Punishment in a Minimal Social Situation." *Journal of Experimental Psychology* 54:318–326.
- Steiner, Ivan D. 1972 *Group Process and Productivity*. New York: Academic Press.
- Tajfel, Henri 1981 *Human Groups and Social Categories*. Cambridge; UK: Cambridge University Press.
- Thibaut, John W., and Harold H. Kelley 1959 *The Social Psychology of Groups*. New York: Wiley.
- Triplett, N. 1897 "The Dynamogenic Factors in Pacemaking and Competition." *American Journal of Psychology* 9:507–533.
- Turner, John C. 1987 *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, UK: Basil Blackwell.
- Whyte, William F. 1955 *Street Corner Society: The Social Structure of an Italian Slum*, 2nd ed. Chicago: University of Chicago Press.
- Zajonc, Robert B. 1966 *Social Psychology: An Experimental Approach*. Belmont, Calif.: Brooks/Cole.

ROBERT W. SHOTOLA

## SOCIAL AND POLITICAL ELITES

At one level, elites can be defined simply as persons who hold dominant positions in major institutions or are recognized leaders in art, education, business, and other fields of achievement. Such individuals exist in all societies, but beyond this mundane observation, social scientists are interested in why particular individuals attain positions of status and power. Does achievement reflect superior talent, or is it a product of social or

cultural advantage? Why are some achievements valued over others? How does the distribution of elite positions in society reflect the particular social structures in which they exist? These questions are the focus of much research on stratification and social inequality.

In the social sciences, the concept of elites refers to a more specific issue as well: the concentration of societal power—especially political power—in the hands of a few. At the heart of theoretical debates and empirical research on elites is the famous assertion of Mosca (1939, p. 50): “In all societies . . . two classes of people appear—a class that rules and a class that is ruled.” One can distinguish the conception of “functional elites” in a variety of institutional contexts from that of a “ruling” or “political” elite that in some sense wields societal-level power. Then the key questions concern the existence and nature of this dominant group. Is power over the major institutions of society highly concentrated, or is it broadly dispersed as “pluralists” claim? If a cohesive ruling elite exists, then who is in it and what is the basis of its power? What is the extent of its power in relation to the nonelite “masses”? Does this societal elite exercise power responsibly in the interests of society as a whole, or do elites maximize their own interests against those of subordinate groups?

### CLASSICAL ELITE THEORY

Social thought on elites goes back at least to Plato and Aristotle, but contemporary debates usually begin with the “neo-Machiavellians” Pareto (1935), Mosca (1939), and Michels ([1915] 1959). Reacting to the turmoil of European society in the early twentieth century, each developed arguments supporting the inevitability of elite rule in opposition to classical democratic theory, Marxian class analysis, and socialist political movements. For Pareto, elites in general were those holding leadership positions in business, politics, education, and other areas of accomplishment. Those individuals could be distinguished from the rest of “nonelite” society. He further distinguished between the “governing elite”—the segment of the elite with broad political power—and the nongoverning elite. His best known statements concerned the former group. Though famous for his work in mathematical economics, Pareto believed that most human behavior was nonrational, the expression of

deep-seated “sentiments” and their observable manifestations, or “residues.” These motivational orientations led to behaviors that were then “explained” through our post hoc rationalizations, or “derivations” (1935, chap. IX). For Pareto, the governing elites were those with dominant talents or leadership skills derived primarily from superior individual attributes. Borrowing from Machiavelli, he distinguished two ideal types of political leaders on the basis of their dominant personal qualities and motivations (“residues”). “Lions” appealed to the conservative instincts that were most common in the masses, relying on tradition, strength, and coercion to rule. “Foxes” were more innovative leaders who relied on cunning, new ideas, and manipulation. Both types were necessary, but Pareto tended to see a cyclical pattern of rule in societies in which “foxes” dominated in periods of upheaval and transition, only to be displaced by “lions” after the restoration of social order (1935, chap. XII).

Pareto also noted that individuals in positions of power often attempt to maintain their privileged positions by closing off access for others. This risks social disruption by shutting off avenues of achievement and power to other talented individuals, who then mobilize to affect change. The “circulation of elites” refers to the process by which the ruling class is renewed periodically by superior individuals from other ranks. For Pareto, obstacles to elite circulation often resulted in the stagnation of the ruling class. Closed aristocracies and caste-like systems fostered tension, conflict, and eventually social change.

Like Pareto, Mosca began with the assertion that elite rule is an empirical fact in all societies. Although he also noted the superior individual attributes of the “ruling class,” his analysis was considerably more sociological than that of Pareto. Mosca emphasized the organizational advantages of the ruling elite in that they represented a relatively cohesive and easily organized minority against the disorganized masses (Mosca 1939, p. 53). He also discussed the role of the “subelite,” a technocratic stratum of managers, intellectuals, and bureaucrats that was increasingly important for elite rule in modern societies (1939, pp. 404–409; see also Marger 1987, p. 54). Mosca’s conception of social change and the circulation of elites was also more sociological. Social, economic, and technological



changes often generated new opportunities and called forth new talents, bringing new elites into prominence. Mosca agreed with Pareto that closed systems of rule threatened social stability, since a stagnant elite impeded adaptation to change.

In *Political Parties* ([1915] 1959), Michels traced the necessity of elite rule in modern societies to the imperatives of complex organization. His classic study analyzed the German Social Democratic Party, but his arguments have been applied to a variety of organizational contexts. Influenced by Weber's ([1921] 1968) work on politics and bureaucracy, Michels's most famous conclusion is summarized in his "Iron Law of Oligarchy," the argument that large-scale organizations necessarily concentrate power in the hands of a few at the top. Once in power, leaders in organizations such as labor unions and political parties act to preserve their positions. Those who rise from lower levels in the organization are co-opted in a process that preserves the structure of power. The resources available to institutional leaders and their relative unity of interest and perspective give them numerous advantages in maintaining their power over the unorganized rank and file. Over time, leaders develop similar interests and intraelite attachments that reflect their elevated position and separate them from the masses. For their part, Michels saw the masses contributing to elite rule through their general apathy and acquiescence. With his focus on organizational factors, Michels has been very influential in the development of contemporary elite approaches to power (see Marger 1987, pp. 56–58; Burton and Higley 1987).

### C. WRIGHT MILLS AND THE ELITE-PLURALIST DEBATE

Among elite theorists there is an important distinction between those who see the concentration of power as inevitable or desirable and those who do not. The former group includes the classical elite theorists and those who have extended their ideas (see Field and Higley 1980; Burton and Higley 1987). In contrast, "critical" or "radical" elite theorists recognize the concentration of power in society but argue that this condition is neither inevitable nor desirable. Unlike the classical theorists who emphasized mass apathy or incompetence, critical elite theorists argue that elite domination

is maintained through the manipulation and exploitation of nonelites.

The most influential representative of the critical elite perspective is Mills (1956). Mills, Hunter (1953), and other critical elite theorists developed their work in response to the dominance of "pluralist" studies of political power in the United States. Pluralism, as represented in the work of Dahl (1956), Truman (1951), Riesman (1950), and others, held that power in modern democratic societies was widely dispersed and that those in decision-making positions were subject to significant mass pressures (through electoral or other processes) or the countervailing power of other institutional elites or organized interest groups. For Mills, the notion of a pluralist balance of power between competing interest groups was a romantic ideal rather than a description of political reality in the United States. He acknowledged the activities of labor unions, farm groups, professional associations, and other organized interest groups but argued that those groups operated mainly at the secondary, local, and "middle levels" of power. The power to make decisions of national and international scope rested with a "power elite" of individuals in top positions of authority in major corporations, the executive branch of government, and the military. Congress was consigned to the middle levels of power, along with most of the interest groups studied by pluralist social scientists. Mills traced the historical consolidation of the power elite to the growth of the federal government in the 1930s and especially during World War II, as industrial production was coordinated with military needs through the government. That institutional alignment was strengthened in the Cold War years as the state expanded its commitment to national security, social welfare, and the direction of economic policy. By the 1950s there was a significant shift in power from Congress to the executive branch, reflecting an expansion of government that required a complex information-gathering and administrative capacity. Congress lacked the resources and coherence required for modern state administration.

Mills argued that most of the members of the power elite had similar values and interests, which reflected their similar backgrounds, common schools, shared membership in elite social clubs, and informal social interaction. He also empha-

sized the continuous professional interaction between these institutional leaders and the frequent exchange of top personnel between major corporations, the military, and the executive branch of government. Another factor contributing to the relative homogeneity of the power elite was their common experience at the apex of bureaucratic institutions. The skills, status, and even personality type required for success were similar in each sphere, reflecting their similarity of organizational structure (Mills 1956, p. 15).

The other side of Mills's conception of the power elite was that of mass society. The same social processes that had concentrated political power had created a society of increasingly fragmented individuals whose lives and interests were shaped for them from above. Information filtered selectively through bureaucratized institutions of mass education and the mass media, which became more susceptible to elite manipulation as they became more centralized. The media emphasized entertainment and consumption over information and critique. Educational institutions had developed into sites of large-scale vocational training rather than havens for the development of critical thought and an informed citizenry necessary for democratic politics (Mills 1956, chap. 13).

Mills's work became the touchstone for debates about the structure of power in the United States that have continued to this day. Pluralists argue that he exaggerated the unity of functional elites and neglected the influence of the electoral process and interest group competition. From the other direction, neo-Marxist and other class-theoretical analysts have been critical of the Millsian model for not acknowledging the extent to which political power is shaped by dominant economic interests (see the debates collected in Domhoff and Ballard 1968). A key question in these disputes concerns the degree of elite cohesion. How much consensus (or competition) between elites is required to support an elite (or pluralist) model? What is the extent of elite competition? Is there a hierarchy of elites, with a ruling class or "power elite" on top, or a "polyarchy" (Dahl 1971) of diverse institutional powers? Elite theorists acknowledge that individuals with different skills and constituencies hold leadership positions in a variety of institutions such as prestigious universities, private foundations, major civic organizations,

and the media (see Dye 1995). Pluralists view these institutions as relatively autonomous sources of societal influence. Although one may identify "strategic elites," or influential leaders, in a variety of fields (Keller 1963), they see no overall cohesion or uniform coordination of policy within a single ruling group. However, those defending an elite perspective argue that disagreements over particular interests occur within a general elite consensus on basic ideology and acceptable policy. Developing Mills's arguments, elite theorists have studied a variety of coordinating mechanisms that foster elite cohesion, such as private school ties, social networks, shared membership in policy planning organizations, and the general recruitment process in which future leaders are instilled with attitudes conducive to maintaining the existing structure of power (see Prewitt and Stone 1973; Marger 1987; Bottomore 1993; Dye 1995; Domhoff 1998). Some who work in this tradition go further than Mills in emphasizing the prominence of *class* interests and corporate power over the political process and other institutions in capitalist societies (Miliband 1969; Useem 1983; Domhoff 1990, 1998). Indeed, the distinction between "elite" and "class" analysis disappears in many such works (on the similarities and differences, see Marger 1987). From this perspective, prestigious Ivy League universities may harbor intellectuals critical of the existing power structure, that but only those academics with "acceptable" views are selected as advisers to political elites in turn must maintain acceptable levels of business confidence and campaign finance to remain in power.

In a similar vein, all parties agree that in a modern democratic system, the "elite," however defined, must pay some attention to the "masses." The question is, How much attention must be paid, and how do public preferences impose themselves on elites? Pluralists hold that the public has a significant influence on elite decision making through voting, public opinion, and the threat of social protest. From a different starting point, some class-based analysts note the role of working class mobilization or the effects of other nonelite social movements, such as the civil rights movement, that force changes in the polity and society (Piven and Cloward 1977). By contrast, those who emphasize elite power tend to leave little room for the influence of nonelites in promoting major social change. Change is viewed as the result of

elite mobilization, intraelite conflict, or the circulation of elites. Mills, for example, viewed the major societal decisions in the United States as the product of elite decision making, while more specific, localized issues were more likely to be negotiated at the “middle levels” of power. Domhoff (1998) considers major policy formation processes by looking at the “agenda-setting” power of elites, noting that while pluralistic interest group competition does occur on specific issues, the general parameters of public discourse and public policy are set in advance and behind the scenes through organizations such as policy planning groups and presidential task forces that bring elites together to build a consensus on major policy issues before specific proposals enter the formal legislative process. Other elite theorists point out that even in periods of mass mobilization over policy issues, the power of elites over the public agenda allows public sentiment to be deflected or diffused by temporary measures or by redirecting public attention to peripheral issues (Prewitt and Stone 1973, pp. 107–108).

#### **MODERNIZATION, MERITOCRACY, AND ELITE RECRUITMENT**

The dominant view of elites in the 1950s developed out of the structural functionalism of Parsons (1940; 1951) and the “end-of-ideology” arguments that appeared around that time (see Waxman 1968). This perspective, which is popular again today, holds that with the emergence of modern industrial societies and liberal democracy, elites increasingly represent a stratum of talented individuals filling important positions of leadership in dominant institutions (see especially Keller 1963; see also Mannheim 1940; Aron 1950). Variations on this theme point to a “New Class” of “knowledge” workers in “postindustrial” managerial and information-based professions and a proliferation of new institutional elites that transcend the old hierarchies of caste and class (Keller 1963; Bell 1974). From this perspective, modern elites are functionally necessary in a society of complex organizations and increasingly specialized occupations. Echoing Davis and Moore’s (1945) functionalist theory of stratification, status and material rewards are seen to reflect the high skill and social responsibility required for those positions. The legitimacy of functional elites is supported to the

extent that relatively equal opportunities to attain those positions are available to all talented and motivated individuals.

The validity of this “meritocracy” model of power is directly related to the issue of elite recruitment and the extent to which positions of power are open to nonelites. Once again, at one level there is general agreement among all parties on the relative openness of modern societies in comparison to traditional systems in which elite “recruitment” often was based on birth. In contemporary societies, differentiation fostered a proliferation of institutional elites requiring specific talents and skills in a variety of fields (see Keller 1963). However, beyond this empirical fact, the questions of contention are: (1) How much openness is there? and (2) Does it matter?

The first question has been the subject of much research in stratification and will be dealt with only briefly here. The meritocracy model assumes equal opportunity for individuals, but considerable research has challenged this assumption. For example, if elite positions are based on merit, educational institutions must provide avenues for mobility and equal opportunity for talented individuals from nonelite backgrounds, but a basic criticism of functionalist theories of stratification is that existing structures of inequality create barriers to nonelite achievement (see Tumin 1953). Beyond the obvious inequality of economic resources and formal educational institutions, the work of Bourdieu and others (Bourdieu 1984; Bourdieu and Passeron 1977; DiMaggio and Mohr 1985) has shown how the unequal distribution of “cultural capital” among groups in different locations in the class structure contributes to the reproduction of inequality in a variety of subtle ways. This research also points to the difficulty in assessing differences in “talent” among individuals or groups, since indicators such as “intelligence,” cultural appreciation, and political knowledge may reflect a preexisting distribution of cultural resources.

A more fundamental question regarding the openness of elite recruitment is: Does it matter? First, if the concern is the overall structure of power, as it was for Mills and most elite theorists, the success of a few upwardly mobile individuals from the lower strata does not affect the analysis: Power still may be concentrated in a few in key

positions. Second, elite recruitment from the lower ranks does not necessarily affect the *content* of elite decision making, given the selection process involved in the rise of “talented” individuals into elite positions. Most analysts agree that in modern societies, attainment of elite positions often requires a degree of talent, effort, and achievement, but elite theorists argue that those who make it to the top are selected for specific orientations that are compatible with existing structures of power. Those from privileged backgrounds, with access to economic, social, or cultural capital, have a definite advantage, but it is possible for nonelites people who possess the right attitudes and skills to rise into positions of power. This maintains the existing structure of power while providing legitimating examples of individual success. Limited avenues for mobility also provide a mechanism for the co-optation of promising leaders from below, as Pareto and Mosca would recognize (see the discussions of elite recruitment in Prewitt and Stone 1973; Marger 1987; Bottomore 1993; Dye 1995; Domhoff 1998).

#### ELITES AND DEMOCRACY

There have been a number of modifications of both “elite” and “pluralist” theory that have brought the two closer together. The work of Weber, Michels, and others influenced later theorists who viewed the concentration of power in modern institutions as necessary. From this vantage point, the issue was not whether broad “democratic” participation in political, economic, and other institutions was possible (it was not) but whether the interests of nonelites could be preserved in the face of modern bureaucratic organization. *Democratic elitism* represented a refinement of pluralist assumptions that redefined democracy in a manner congruent with elite rule. Given the inevitable concentration of power in modern societies, the central problem became: What legitimates elite rule or preserves elite “accountability”? The traditional answer of conservative elitists had been the “virtue,” “character,” or inward convictions of elite leaders in comparison to the selfish and undisciplined masses. In this tradition, with adherents from Plato to Pareto, elites have a stronger commitment to the “public interest” than do the “people” (Prewitt and Stone 1973, pp. 188–196). The meritocracy model represents a contempo-

rary variation of this viewpoint: Modern institutions require skilled leadership, and this means that institutional elites are increasingly likely to be selected on the basis of superior talent.

Another source of public accountability important for “democratic elitists” is elite competition for electoral support. Political elites must compete for votes in formal democracies, and this acts as a broad restraint on their actions. However, once in office, elite decision makers are relatively free to act as they see fit as long as their actions remain within acceptable limits. In a well-known formulation by Schumpeter (1942), elite rule is preserved both by superior talent and by the general mass apathy that he saw as functional for political rule. For Schumpeter (1942, pp. 269–296) and other conservative advocates of democratic elitism, the efficiency of modern representative governments depends on the “people” selecting their leaders and then leaving them alone. Note that the definition of “democracy” has been transformed from an emphasis on maximum public participation in political life to an assertion of the functional necessity of *non*participation. Far from government “by the people,” democracy is now defined as a procedure for the selection of political elites. This underscores the difficulty in weighing empirical claims concerning “democratic” representation made by competing theories, given the radically divergent definitions of the key concept. Classical theories of democracy emphasized the importance of political participation as an end in itself, one that was necessary for the *creation* of political citizens capable of democratic self-rule (Pateman 1970). Critics of conservative elitism such as Bottomore (1993, p. 95) wonder if “a person can live in a condition of complete and unalterable subordination for much of the time, and yet acquire the habits of responsible choice and self government which political democracy calls for.” This issue is muddier for the fact that many *critical* elite theorists are ambivalent about the possibilities for participatory democracy in modern society. If elite rule is undesirable, it would seem necessary to provide an alternative. Mills held up participatory democracy as an ideal from which to judge the contemporary United States in his concept of “publics” (1956, pp. 302–304), but he was not very clear about how that ideal could be implemented in modern “mass society.” Other critical elite theorists seem to have accepted

the classical argument for the inevitability of elite rule in modern, bureaucratically organized societies (see Prewitt and Stone 1973; Marger 1987; Burton and Higley 1987). For these critics, the only option to a “democracy” of mass apathy is one in which political institutions and decision-making elites are as open as possible to public scrutiny by a truly informed electorate.

### LIMITATIONS OF THE ELITE PARADIGM

The elite paradigm—and by extension the “elitism-pluralism” debate as it usually is formulated—focuses on the leaders of large-scale organizations and organized institutions. Power is based on command over organizational resources; the elites are those in positions of organized power. While there are other ways to conceptualize elites, this is the dominant model in the social sciences today (see Marger 1987; Burton and Higley 1987; Dye 1995). Most scholars agree that power is concentrated in such organizations. The disagreements occur over whether there is a unified “ruling elite” above and beyond these multiple institutional elites that characterize all modern societies.

This controversy has led to much fruitful research and theoretical debate, but the “elitist-pluralist” framework is less adequate for dealing with other dimensions of societal power. For example, many of the social and cultural processes involved in the reproduction of class, gender, and racial inequality cannot be encompassed within an organizational paradigm. This includes socialization processes and everyday practices within the family, school, and workplace that reproduce the hegemony of a dominant culture. Further, it is possible to map the formal leadership structure of educational institutions, research foundations, and media organizations without explaining the *content* of their decisions. One might ask: Elite power, yes, but power for what? With its emphasis on the power of individuals within organizations, the elite paradigm neglects many structural and cultural forces that constrain those organizations and the elites within them. For example, how do global economic conditions and the imperative of “business confidence” constrain the decisions of political and economic elites? How are the ideologies and cultural practices that govern gender relations reproduced in the boardroom or the executive

mansion? These questions are significant, because without them it is difficult to explain why elites make the decisions they do or why some societal interests are better represented than others are in the decision-making process. It is necessary to consider elites and the organizations they command in their larger social and cultural context.

This issue was highlighted many years ago in debates over the “managerial revolution thesis.” This was an argument that modern corporations are different from traditional capitalist enterprises because of their separation of ownership from management. Managers were seen to have aims different from those of capitalists, reflecting their organizational position. They were more interested in long-term growth, stability, labor peace, and good community relations—good management—and less concerned with profit maximization (Berle and Means 1932; Burnham 1941). The simple but fundamental weakness in such an argument was that managers—the “elites” who wield organizational power in modern corporations—were still constrained by the imperatives of the market and capital accumulation. Time has shown that their ability to act “managerially” reflected a brief postwar period of U.S. dominance in the world economy. Global competition has since required that corporate elites act more like representatives of capital.

The same might be said about political elites as well, which brings one back to the issue of the relationship between political power and class interest. Recent debates in political sociology over the degree to which state institutions, governing officials, and policy intellectuals are “relatively autonomous” from the constraints of class interest or other societal pressures have again brought into focus the relationship between political power and economic interest (see Skocpol 1985; Jessop 1990). Class theorists have argued that the decisions of political elites are shaped not only by the superior resources of a dominant class but also by the structural constraints on the state in a market economy (see Block 1977; Lindbloom 1977). Others have traced a clear class bias and pro-capital selectivity inherent in the very institutions of modern states and the dominant political discourse (Jessop 1990). Parallel arguments have been made by feminists who hold that patriarchal domination is embedded in the very structure of the state (e.g.,

MacKinnon 1989). These lines of inquiry do not negate the importance of research on elites, but they lead one to ask questions about the larger social forces that shape their decisions.

## REFERENCES

- Aron, Raymond 1950 "Social Structure and the Ruling Class." *British Journal of Sociology* 1:126-144.
- Bell, Daniel 1974 *The Coming of Post-Industrial Society*. New York: Basic Books.
- Berle, A. A., and G. C. Means 1932 *The Modern Corporation and Private Property*. New York: Macmillan.
- Block, Fred 1977 "The Ruling Class Does Not Rule." *Socialist Revolution* 7(3):6-28.
- Bottomore Tom 1993. *Elites and Society*. London: Routledge.
- Bourdieu, Pierre 1984 *Distinction: A Social Critique of the Judgment of Taste*. Cambridge, Mass.: Harvard University Press.
- , and J. C. Passeron 1977 *Reproduction in Education, Society, and Culture*. Beverly Hills, Calif.: Sage.
- Burnham, James 1941 *The Managerial Revolution*. New York: Day.
- Burton, Michael G., and John Higley 1987 "Invitation to Elite Theory." In G. William Domhoff and Thomas Dye, eds., *Power Elites and Organizations*. Newbury Park, Calif: Sage Publications pp. 219-238.
- Dahl, Robert A. 1956 *A Preface to Democratic Theory*. Chicago: University of Chicago Press.
- 1971 *Polyarchy*. New Haven, Conn.: Yale University Press.
- Davis, Kingsley, and Wilbert Moore, 1945 "Some Principles of Stratification." *American Sociological Review* 10(2): 242-249.
- DiMaggio, Paul, and John Mohr 1985 "Cultural Capital, Educational Attainment, and Marital Selection." *American Journal of Sociology* 90: 1231-1261.
- Domhoff, G. William 1990 *The Power Elite and the State*. New York: Aldine de Gruyter.
- 1998 *Who Rules America? Power and Politics in the Year 2000*. Mountain View, Calif.: Mayfield.
- and Hoyt B. Ballard, eds. 1968 *C. Wright Mills and the Power Elite*. Boston: Beacon Press.
- Dye, Thomas 1995 *Who's Running America? The Clinton Years*. Englewood Cliffs, N.J.: Prentice Hall.
- Field, G. Lowell, and John Higley 1980 *Elitism*. London: Routledge and Kegan Paul.
- Hunter, Floyd 1953 *Community Power Structure*. Chapel Hill: University of North Carolina Press.
- Jessop, Bob 1990 *State Theory: Putting Capitalist States in their Place*. University Park: Penn State University Press.
- Keller, Suzanne 1963 *Beyond the Ruling Class: Strategic Elites in Modern Society*. New York: Random House.
- Lindbloom, Charles 1977 *Politics and Markets*. New York: Basic Books.
- MacKinnon, Catherine 1989 *Toward a Feminist Theory of the State*. Cambridge, Mass.: Harvard University Press.
- Mannheim, Karl 1940 *Man and Society in an Age of Reconstruction*. London: Kegan Paul.
- Marger, Martin N. 1987 *Elites and Masses: An Introduction to Political Sociology*. Belmont, Calif.: Wadsworth.
- Michels, Robert (1915) 1959 *Political Parties*. New York: Dover.
- Miliband, Ralph 1969 *The State in Capitalist Society*. New York: Basic Books.
- Mills, C. Wright 1956 *The Power Elite*. New York: Oxford University Press.
- Mosca, Gaetano 1939 *The Ruling Class*. New York: McGraw-Hill.
- Pareto, Vilfredo 1935 *The Mind and Society*. New York: Harcourt Brace.
- Parsons, Talcott 1940 "An Analytical Approach to the Theory of Social Stratification." *American Journal of Sociology* 45:841-862
- 1951 *The Social System*. Glencoe, Ill.: Free Press.
- Pateman, Carole 1970 *Participation and Democratic Theory*. Cambridge, UK: Cambridge University Press.
- Piven, Frances, and Richard Cloward 1977 *Poor People's Movements*. New York: Random House.
- Prewitt, Kenneth, and Alan Stone 1973 *The Ruling Elites*. New York: Harper & Row.
- Riesman, David 1950 *The Lonely Crowd*. New Haven, Conn.: Yale University Press.
- Schumpeter, Joseph A. 1942 *Capitalism, Socialism, and Democracy*. New York: Harper Colophon.
- Skocpol, Theda 1985 "Bringing the State Back In: Strategies of Analysis in Current Research." In Peter Evans, Dietrich Reuschmeyer, and Theda Skocpol, eds. *Bringing the State Back In*. Cambridge, UK: Cambridge University Press.
- Truman, David B. 1951 *The Governmental Process*. New York: Random House.
- Tumin, Melvin, 1953 "Some Principles of Stratification: A Critical Analysis." *American Sociological Review* 18: 387-393.
- Useem, Michael 1983 *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and the U.K.* New York: Oxford University Press.

Waxman, Chaim I. ed. 1968 *The End of Ideology Debate*. New York: Clarion.

Weber, Max, (1921) 1968. *Economy and Society*, vols. Berkeley: University of California Press.

PATRICK AKARD

## SOCIAL BELONGING

### THE DIMENSIONS OF HUMAN INVOLVEMENT

It is necessary to distinguish four different dimensions or states in the involvement of individuals in the context of human relations: territorial location, ecological participation, social belonging, and cultural conformity (Pollini 1990) (Figure 1). *Territorial location*, as Weber showed in his famous sociological analysis of the medieval European city (Weber 1921) does not involve any form of social relation among the individuals of a population in a particular territorial area. This dimension was subsequently defined by Parsons as one of the three primary relational criteria, with the other two being biological position and temporal location (Parsons 1959, pp. 89–96).

Unlike territorial location, *ecological participation* involves some sort of reciprocal relationality among the individual members of a human population, whether settled in the same territorial area or not. To use the terminology of human and social ecology in reference to nonsymbolic social relations, recalling Mead's well-known distinction (Mead 1934), ecological participation involves a specific form of interdependence among individuals ("symbiosis") (Park 1936, 1939) that is distinctly different from social interaction (Quinn 1939). For Parsons, the ecological system is "a state of mutually oriented interdependence of a plurality of actors who are not integrated by bonds of solidarity to form a collectivity but who are objects to one another" (1959, p. 93). Thus, instrumentally, the customers of a commercial firm, the participants in a market, and the antagonists in a struggle, and expressively, a network of purely personal friendships and the inhabitants of a neighborhood or district in a modern metropolis are paradigmatic examples of the dimension of participation in networks of ecological interaction or

in purely ecological systems. Parsons defined the state of ecological participation as a secondary relational criterion.

*Social belonging* refers to the state in which an individual, by assuming a role, is characterized by inclusion in the social collectivity, which is exclusively a *Gemeinschaft*, according to Weber (Weber 1922, 136), and which is a *Gemeinschaft* (an organization or association), according to Parsons (Parsons 1959, p. 100). In this frame of reference, the dimension of social belonging relates to any form of social collectivity, whether predominantly expressive (nonrational in Weber's terms) or predominantly instrumental. Strictly speaking, the status of belonging concerns only the *symbolic* dimension of human and social relations and interactions (Durkheim 1912; Pareto 1916; Weber 1921, 1922; Mead 1934; Park 1939; Parsons 1959; Merton 1963; Shils 1975). Parsons defines it as a secondary relational criterion.

*Cultural conformity* is symbolic in character. This dimension differs from social belonging in that it involves the sharing by individuals of value systems and therefore of attitudes of "consensus" as defined by Weber (Weber 1913) as well as, though not necessarily, conformism (Parsons 1959). The distinction between social belonging and cultural conformity demonstrates that belonging to a collectivity can be compatible with the exercise of internal opposition; thus, social membership does not exclude the possibility of disagreement, especially in regard to value orientations.

The distinction between social belonging and cultural conformity also has been drawn by Robert K. Merton, who expressly asserts the noncoincidence between membership groups and reference groups. The latter groups constitute a focus of reference toward which a certain degree of positive orientation is shown rather than being an already-established social bond that is manifest in the interactions among the individual members of a group (the membership group).

On the basis of the distinction between ecological participation and social belonging—both of which are secondary relational criteria, according to Parsons—it is possible to use the findings of human ecology and sociological analysis to differentiate between attachment to the community and belonging to the *Gemeinschaft*. Whereas attach-

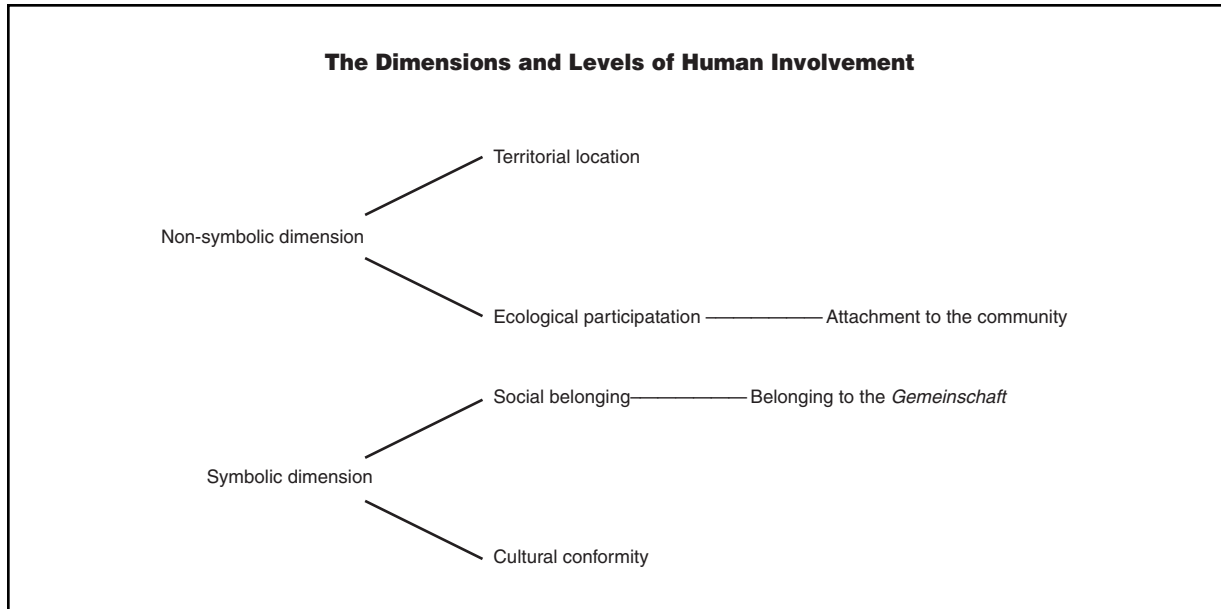


Figure 1

SOURCE: Pollini 1990, p. 188.

ment to the community involves the ecological concept of community defined as “a) a population, territorially organized; b) more or less completely rooted in the soil it occupies; c) its individual units living in a relationship of mutual interdependence that is symbiotic rather than societal” (Park 1936, p. 148), the belonging to the *Gemeinschaft* concerns that sociological concept of *Gemeinschaft* as defined by Toennies (1887), Weber (*Vergemeinschaftung*) (1922), and Parsons (1959), although for Parsons as well as for some others, social belonging concerns not *Gemeinschaft* alone but any social collectivity and the social collectivity *qua talis*.

**THE STRUCTURE OF SOCIAL BELONGING**

The distinction between attachment to the ecological community and belonging to the social collectivity (particularly the *Gemeinschaft*) introduces the fundamental question of the structure of social belonging and the relations among its main components, which from an analytic and multidimensional perspective include attachment. Using Parsons’s scheme of reference, together with the contributions of other sociologists, the structure of social belonging can be described by

starting from the relations among the four chief components that define it as such: *attachment, loyalty, solidarity,* and the *sense of affinity* or *we-feeling* (Figure 2).

*Attachment* is a form of investment or “cathexis” (from the Freudian term *Besetzung*, denoting the relationship between emotional energy and an object) in a social object (the collectivity in this case), where “cathexis” refers to “the significance of ego’s relation to the object or objects in question for the gratification-deprivation balance of his personality” (Parsons 1959, p. 17). Attachment involves an “orientation to alter in which the paramount focus of cathective-evaluative significance is in alter’s attitudes” (Parsons 1959, p. 213), where “the relation to alter is the source, not merely of discrete, unorganized, ad hoc gratifications for ego, but of an organized *system* of gratifications which include expectations of the future continuance and development of alter’s gratificatory significance” (Parsons 1959, p. 77).

When attachment is organized into a symbolic pattern, particularly a pattern of expressive symbols whose meaning’s shared between ego and alter, become values—in other words, when they serve as a criterion or standard for selection (or an



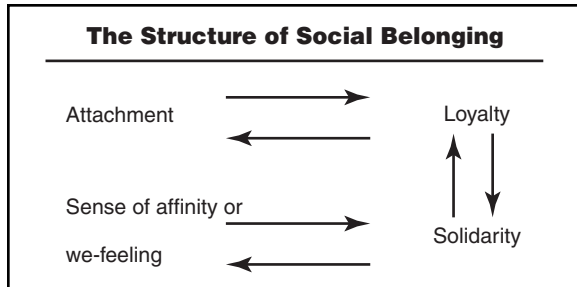


Figure 2

appreciative criterion in this case, which concerns expressive symbolism) among the alternatives of orientation that are intrinsically available in the situation—*loyalty* arises (Parsons 1959, p. 77). In the case of social belonging, loyalty defines the relation between ego as a subject and the collectivity as social object of which the ego is a member. Besides being the social object of attachment, the social collectivity thus becomes the object of loyalty as well. This raises the question of the trust the collectivity requires and the individual grants.

Along with attachment and loyalty, social belonging involves the *solidarity* of the collectivity. Solidarity, which “involves going a step beyond ‘loyalty,’” is defined by Parsons as “the institutionalized integration of collectivity,” and it is distinguished from loyalty because it entails that “collectivity-orientation converts this ‘propensity’ into an institutionalized obligation of the role-expectation. Then whether the actor ‘feels like it’ or not, he is obligated to act in certain ways and risks the application of negative sanctions if he does not” (Parsons 1959, p. 98).

The final component that defines the structure of social belonging is what has been called the “sense of affinity” (Shils 1975) or the “we-feeling” (MacIver and page 1949, p. 5ff). (Weber treated belonging in terms of *Zusammengehörigkeit*, or “subjective feeling of the parties, whether affectual or traditional, that they belong together”) (Weber 1922, p. 136). Although this component can be considered the final outcome of attachment, loyalty and solidarity, it also can be viewed as the component that controls and legitimates the others and therefore performs the function of pattern maintenance in the system of social belonging. It may include, at least in part, two of the factors that

Merton states constitute a collectivity as a social group: people’s definition of themselves as “members” of the group and definition by others as “belonging to the group,” with the others, including fellow members and nonmembers (Merton 1963).

In short, social belonging is constituted by the relations of interdependence among the dimensions of attachment, loyalty, solidarity, and sense of affinity, according to paths that extend from attachment to a sense of affinity or we-feeling and back, passing through the intermediate components of loyalty and solidarity.

From the point of view of the collectivity as a social system, belonging is the dimension that can be called a “residue,” to use Pareto’s term. A residue is the relatively constant symbolic-social element that can be deduced from the symbolic-linguistic expressions (or nonlogicoexperimental theories) associated with the nonlogical actions of associated individuals and that performs a function of “persistence of aggregates” maintain the equilibrium of the system. This equilibrium is characterized by the interdependence relations among the residues of various classes, genera, and specie, and between these and the system’s other “internal’ elements,” such as derivations, interests and social heterogeneity, and which is the circulation among the parts (Pareto 1916; Pollini 1987).

In regards to attachment, Shils, adopting a concrete rather than an analytic perspective, has drawn up a typology of four kinds of attachment that can be compared with the notions presented here. Shils’s first type of attachment is the *primordial attachment* that arises among individuals by virtue of “particularist existential connections” (such as the biological bond of kinship) and stable sexual relations or the sharing of a territorial area. It can be compared with the community attachment of human ecology and, more generally, the cathexis in a broad sense involved in ecological participation.

*Personal attachment* and *civil attachment* operate at different levels of the process of social structuring. They are distinctive of the social belonging defined by the “emerging” components of loyalty, solidarity, and the we-feeling.

*Sacred attachment* is grounded in beliefs and therefore also in notions of truth, justice, good-

ness, and beauty. It mainly but not exclusively includes cultural conformity and a consensus on beliefs, although it gives rise to a social community that Shils views as a community of *believers* (Shils 1961).

### SOCIAL BELONGING AND ITS RELATIONS WITH OTHER COMPONENTS OF HUMAN ACTION

This discussion of the structure of social belonging and its main components has identified a number of elements by which it is constituted and conditioned. By adding further elements of fundamental importance, we may outline a complete frame of reference which social belonging involves interrelations among the following subsystems or “complexes”: the ecological complex of *territorial location* and *ecological interaction*, the mental complex of the *identity* of the personality, the social complex of the *solidarity* of the collectivity, and the cultural complex of *expressive and evaluative symbolism* (Figure 3).

Starting from territorial location and ecological interaction, central importance is assumed by the relationship between the *identity* of the personality and the *solidarity* of the collectivity, both of which stand in relation to the complex of *expressive and evaluative symbolism*. It is the latter factor in particular that, through internalization and institutionalization, characterizes personal identity and collective solidarity (Durkheim 1912), of which the personal identity involves the process by which the symbolic complex is acknowledged and the collective solidarity later involves the process by which it is represented.

The process by which the social collectivity relates to the individual person can be called the process of *inclusion*, while the mental process by which a person comes to be inducted in a collectivity may be called the mechanism of *identification*, or the mechanism by which a person learns “to play a role complementary to those of other members in accord with the pattern of values governing the collectivity” (Parsons 1958, p. 91). In other words, identification is “motivational ‘acceptance’—at levels of ‘deep’ motivational ‘commitment’—of membership in collective systems” (Parsons 1970, p. 356).

From the point of view of the personality, the multiple social belongings or even belonging to

multiple collectivities or social circles so distinctive of the individual condition today are inevitable components of an identity (Parsons 1968, p. 21), to the point where the perception of personal individuality is determined by membership in a collectivity or social circle (first sociological *a priori*) (Simmel 1890, 1908). However, just as the individual as a member of society (a social person) is determined not wholly by the fact that she/he is a member of society but also by the fact that she or he is “not socialized” (second sociological *a priori*) (Simmel 1890, 1908), identity marks out “the individual autonomy relative to *any* role and collectivity membership” (Parsons 1968, p. 20).

Simmel defined the relations between individual identity and belonging to social circles on the basis of the following principles:

1. A positive correlation exists between the development of personal identity and the widening of the social circle of belonging.
2. There is a positive correlation between the increased extension of the social circle and the centrifugal tendency of the individual toward other circles.
3. The individual belongs to an intermediate social circle that fosters the individualization of identity even in very large communities.
4. The determinacy of personal identity increases in direct proportion to belonging to other social circles.
5. The determinacy of individual identity is positively correlated with the dispersion and diffusion of the multiple social circles of belonging rather than with their overlapping, concentration and coincidence.
6. The shared belonging of several individuals to the same social circle may not be incompatible with their single and distinct belonging to other competing and conflicting circles.
7. The modern form of belonging displays voluntary and autonomous, rather than coercive and heteronomous, belonging to a social circle or circles.
8. In the modern age, the totalizing and globalizing nature of belonging to a single and all-encompassing social circle tends

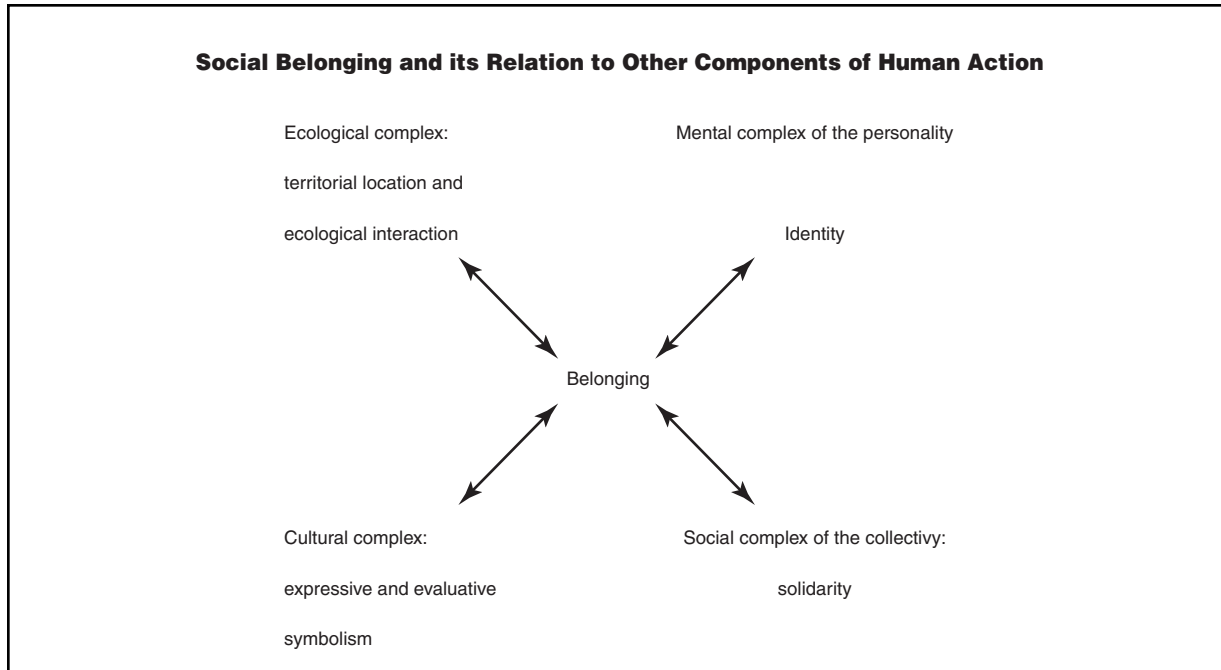


Figure 3

to give way to the functional feature of belonging to multiple social circles (Simmel 1890, 1908).

A person's role as member of the collectivity, which is acquired by means of the mechanism of identification by which social belongings come to constitute inevitable components of personal identity, entails acknowledgment and internalization by the individual personality of the symbolic complex. This is both the foundation of the social collectivity and its representation. According to Durkheim, for whom the social group is defined by the symbolic complex and especially by the totemic symbol (Durkheim 1912), and according to Parsons, for whom expressive symbolism involves not only individual members or units (in that it is shared by each of them) but the entire collectivity constituted as a social object by symbolic social interaction, membership in the social collectivity is not expressed and represented only by the symbolic complex and by distinct symbols and emblems. More important, it is reinforced, developed, and augmented by them, in particular by participation in specific symbolic actions and rituals such as ceremonies, celebrations, gatherings, and meetings and by the projection of shared

value sentiments into individual members, especially those who assume the role of chief or leader and thus symbolically embody shared value patterns (Parsons 1959, pp. 395–399). Thus, ritual participation does not only express and manifest belonging to the collectivity and group; it also strengthens and develops this belonging, in particular the component of it denominated we-feeling or sense of affinity.

#### MEMBERSHIP AND NONMEMBERSHIP GROUPS

On the basis of the theory of the reference group, Merton has examined the limiting condition of nonmembership, which he defines as a positive orientation toward groups that, although not belonged to, are nevertheless reference groups. It is thus possible to identify a diversified set of nonmembership features that depend on the one hand on the nonmember's attitudes toward membership and on the other hand on possession of the qualities necessary for membership established by the group. Merton uses these two distinct dimensions to draw up a typology of a nonmembership group that includes a variety of forms and condi-

tions of nonmembership that may or may not eventually give rise to membership. These forms range from the “antagonistic nonmember” (out-group) to the “candidate for membership” through the intermediate types “potential member,” “autonomous nonmember,” “marginal man,” and “detached nonmember.” Merton thus overcomes the membership-nonmembership dichotomy by proposing gradations of nonmembership and investigating the various types of membership that cannot be defined by lack of membership.

A brief discussion is required of the type of nonmembership Merton calls “marginal man,” following Stonequist (1937) and Park (1928). This is the individual who, although he or she aspires to join a particular group, does not fulfil the requirements established by the group to do so. In other words, the marginal man is a person who, by simultaneously participating in two cultures, tends to occupy not the center but the margins of both. As a consequence, he is not fully accepted by either culture. More precisely, the “marginal man” is the individual who seeks to abandon his membership group, leaving its institutionalized value patterns and norms behind, but is unable to belong to the new group to which he aspires even though he has already absorbed its values and norms to some extent. This process by which the “marginal man” absorbs the values and groups of the group to which he aspires but does not yet belong is called “anticipatory socialization” by Merton (1963).

The “marginal man” also may be likened to the “stranger” described by Simmel (1908, 1964) and Schutz (1944).

### **BELONGING: INCLUSION AND PARTICIPATION**

Belonging to a collectivity or social group characterized by solidarity through the assumption of some sort of role within it by an individual is brought about by two concomitant processes: *inclusion* and *participation*. By means of the process of inclusion, the social collectivity constantly retains or acquires individuals within its relational ambit, thus responding to the problem of integration by eliminating possible exclusion. The intensity of the inclusion process by which individuals

become full-fledged members of the collectivity may vary greatly with the features of the collectivity, primarily according to the criteria that must be fulfilled to join it. These criteria are defined by the collectivity itself and can be characterized as universalism or particularism (Levy 1952).

At the micro level it is possible to identify a number of ways in which the inclusion process comes about: physical and mental coercion, monetary and symbolic remuneration, persuasion, and co-option. At the macro-sociological level, the ways in which individuals are included in collectivities have been variously defined, mainly in the context of membership in the national community (i.e., citizenship) (Turner 1993; van Steenberg 1994) and in relation to immigration (Alexander 1980; Bauböck 1994; Miles and Thränhardt 1995). With no claim to exhaustiveness, one can point out some of these forms of inclusion: adjustment (Eisenstadt 1954), incorporation (Smith 1974; Horowitz 1975; Portes and Boeroc 1989; Schmitter Heisler 1992), absorption (Eisenstadt 1954), and assimilation (Park 1913; Gordon 1964; Horowitz 1975; Geazer 1993).

While inclusion is the process that moves from the collectivity to the individual as object, participation is the process that moves from the individual as subject to the collectivity.

The etymology of the term “participation” has a twofold meaning by which it can denote either “taking part in” something or “being part of” something. It is therefore advisable to distinguish the concept of participation from that of belonging by defining “participation” only in the former sense, in other words, on the basis of the action of “taking part in,” and defining “belonging” as “being part of.” However, it is possible to identify a relationship between belonging and participation, since participation denotes the action of “taking part in” the collectivity, of which one “becomes part” by virtue of the process of inclusion activated by the collectivity.

Besides operating “transitively”—through the relationship a person establishes with the collectivity as a symbolic-social object of attachment and loyalty—participation operates “intransitively” through the relationship the person establishes with himself or herself while participating. By undertaking the action of “taking part in” the

collectivity, a person exerts effects on himself or herself, generally in the direction of change in the elements, aspects, and traits of the psychic structure of his or her personality. Participation thus not only helps reinforce belonging but changes and develops the various components of the individual unit.

From the transitive point of view or from that of the individual vis-à-vis the collectivity, Smelser's theory of collective behaviour (Smelser 1963) and Merton's modes of individual adaptation (Merton 1963) provide the basis for the identification of at least four dimensions of participation: *diffuse* participation, *technical decision-making* participation, *reformatory* participation, and *revolutionary* participation.

Diffuse participation is the dimension of participation, of adaptive character, that accepts the institutionalized means and ends established by the collectivity and attempts to ensure its continued functioning. Technical decision-making participation is the dimension of participation that, while accepting the ends of the collectivity of membership, attempts to change its institutionalized means, which are deemed unable to achieve the goals established. Innovative-reformatory participation accepts the collectivity's means or at any rate is unconcerned by them and attempts instead to change its ends and goals. Revolutionary participation endeavours to change the institutionalized means and ends of the collectivity with a view to creating a "collectivity of the future" that differs from the "collectivity of the present."

## REFERENCES

- Alexander, Jeffrey C. 1980 "Core Solidarity, Ethnic Outgroups and Structural Differentiation: Toward a Multidimensional Model of Inclusion in Modern Societies." In Jacques Dofny and Akinsola Akiwowo, eds., *National and Ethnic Movements*. Los Angeles and London: Sage.
- Bauböck, Rainer 1994 *From aliens to citizens: redefining the status of immigrants in Europe*. Alolershot: Avebury
- Durkheim, Emile 1912 *The Elementary Forms of Religious Life*. London: Allen & Unwin, reissued 1976.
- Eisenstadt, Shmuel N. 1954 *The Absorption of Immigrants*. London: Routledge & Kegan Paul.
- Glazer, Nathan 1993 "Is Assimilation Dead?" *Annals of the American Academy of Political and Social Science*, vol. 530, nov., pp. 122-136
- Gordon, Milton M. 1964 *Assimilation in American Life. The Role of Race, Religion and National Origins*. New York: Oxford University Press.
- Horowitz, Donald L. 1975 "Ethnic Identity." In Nathan Glazer, and Daniel P. Moynihan, eds., *Beyond the Melting Pot*. Cambridge, Mass: Harvard University Press and MIT Press, originally published 1963.
- Levy, Marion J., Jr. 1952 *The Structure of Society*. Princeton, N.J.: Princeton University Press.
- MacIver, Robert M., and Charles H. Page 1949 *Society: An Introductory Analysis*. New York: Holt, Rinehart & Winston, reissued 1961.
- Mead, George H. 1934 *Mind, Self and Society: From the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press, reissued 1959.
- Merton, Robert K. 1963 *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Miles, Robert, and Dietrich Thränhardt 1995 *Migration and european integration: the dynamics of inclusion and exclusion* London: Pinter
- Pareto, Vilfredo 1916 *Trattato di sociologia generale*. Florence: G. Barbèra.
- Park, Robert E. 1913 "Racial Assimilation in Secondary Groups." *American Sociological Review* VIII:66-83.
- 1928 "Human Migration and Marginal Man." *American Journal of Sociology* XXXIII(6):881-893.
- (1936) "Human Ecology." In *Human Communities: The City and Human Ecology*. New York: Free Press, reissued 1952.
- 1939 "Symbiosis and Socialization: A Frame of Reference for the Study of Society." *American Journal of Sociology* XLV(1):1-25.
- Parsons, Talcott 1958 "Social Structure and the Development of Personality: Freud's Contribution to the Integration of Psychology and Sociology." In *Social Structure and Personality*. New York: Free Press, reissued 1964.
- 1959 *The Social System*. Glencoe, Ill.: Free Press, first edition 1951.
- 1968 "The Position of Identity in the General Theory of Action." In Chad Gordon and Kenneth J. Gergen, eds., *The Self in Social Interaction*, vol. 1. New York: Wiley.
- 1970 "Equality and Inequality in Modern Societies, or Social Stratification Revisited." In Talcott Parsons, *Social Systems and the Evolution of Action Theory*, New York: Free Press, pp. 321-380
- Pollini, Gabriele 1987 *Appartenenza e identità: Analisi sociologica dei modelli di appartenenza sociale* (Belong-

- ing and Identity: Sociological Analysis of Models of Social Belonging). Milan: Franco Angeli.
- 1990 “Appartenenza socio-territoriale e mutamento culturale” (Socioterritorial Belonging and Cultural Change). In Vincenzo Cesareo, ed. (a cura di), *La cultura dell'Italia contemporanea*. Turin: Fondazione G. Agnelli.
- , 1999 “The socio-territorial belonging.” In Renzo Gubert, ed., *The Territorial Belonging between Ecology and Culture*. Trento, Italy: University of Trento Press.
- Portes, Alejandro, and József Boerococz 1989 “Contemporary Immigration: Theoretical Perspectives on Its Determinants and Modes of Incorporation.” *International Migration Review* 23(3): 606–630.
- Quinn, James A. 1939 “Ecological versus Social Interaction.” *Sociology and Social Research*. 28:565–570.
- Schmitter Heisler, Barbara 1992 “The Future of Immigrant Incorporation: Which Models? Which Concepts?” *International Migration Review* XXVI(2): 623–645
- Schutz, Alfred 1944 “The Stranger.” *American Journal of Sociology* XLIX(6):499–507.
- Shils, Edward A. 1961 “Society: The Idea and Its Sources.” *Revue Internationale de Philosophie* 15:93–114.
- 1975 *Center and Periphery: Essays in Macrosociology*. Chicago: University of Chicago Press.
- Simmel, Georg 1890 *Ueber soziale Differenzierung: Sociologische und psychologische Untersuchungen*. Leipzig: Dunker & Humblot.
- 1908 *Soziologie: Untersuchungen ueber die Formen der Vergesellschaftung*. Berlin: Duncker & Humblot, reissued 1983.
- 1964 “The Stranger.” In Kurt H. Wolff, ed., *The Sociology of Georg Simmel*. New York: Free Press.
- Smelser, Neil J. 1963 *Theory of Collective Behavior*. New York: Macmillan.
- Smith, M. G. 1974 *Corporations and Society*. London: Duckworth.
- Stonequist, Everett V. 1937 *The Marginal Man*. New York: Scribner's.
- Toennies, Ferdinand 1887 *Community and Society*. New Brunswick, N.J.: Transaction, reissued 1988.
- Turner, Bryan S. 1993 ed., *Citizenship and Social Theory*, London: Sage.
- van Steenberg, Bart 1994 ed., *The Condition of Citizenship* London: Sage.
- Weber, Max 1913 “Ueber einige Kategorien der verstehenden Soziologie.” In *Gesammelte Aufsätze zur Wissenschaftslehre*. Tuebingen: Mohr, 2d ed. 1951.

— 1921 *The City*. New York: Free Press, reissued 1958.

— 1922 *The Theory of Social and Economic Organization*. New York: Free Press, reissued 1966.

GABRIELE POLLINI

## SOCIAL CAPITAL

Social capital is a form of capital that exists within relationships among individuals. According to Bourdieu (1986), capital is accumulated labor that can be appropriated by individuals or groups for their exclusive use to further their interests and increase their capital holdings. By drawing on the social capital resources in their relationships, individuals can further their own goals. For example, the larger the social network an individual has while looking for a job, the more resources—through company contacts, information, and higher-status references—that individual can draw on. Greater availability of resources increases the likelihood that the individual will find a job better than will an individual with fewer social network resources. It is also likely that a well-connected individual will find employment sooner. Simply put, social capital is “an elegant term to call attention to the possible individual and family benefits of sociability” (Portes and Landolt 1996, p. 94).

Some scholars have remarked on the “plethora of capitals” recently appearing in social and economic theories that refer to virtually all aspects of social life as a form of capital (Baron and Hannan 1994). The notion of physical capital, as embodied in machines, tools, and equipment, has been extended by economists to include human capital. Just as investments can be made to improve physical capital—newer and better tools—human capital can be increased by enlarging an individual's skills or knowledge base. Social capital, then, is created when relationships are used to enable actions by individuals to further their own interests. According to Coleman (1988), each actor has control over and interests in certain resources and events. Social capital constitutes a specific kind of resource that is available to an actor. Unlike other resources, social capital is based on reciprocity and thus comes with the expectation that obligations will be repaid as requested by other individuals in the network. Social capital

does not have a set rate of exchange, and payments often are made according to need rather than in standard purchase terms as is done with equipment and education.

The theoretical foundation of the concept of social capital is still in a nascent phase, and there is much debate about its definition, creation, and utility as well as its role in public policy and modernization strategies. Advanced most notably in the work of the French sociologist Bourdieu and further developed by Coleman, the notion of social capital has been put forth as a conceptual tool to bridge two divergent theories of social action: economism and semiologism (Bourdieu 1986). Economism reduces all social exchanges to economic transactions in which independent social actors pursue their own self-interest with little attention to social context. In this view, social action is based on the principle of maximizing utility. In contrast, semiologism reduces social exchanges to communicative acts by socialized actors and downplays the impact of economic factors. In this theoretical orientation, social action is governed by social norms, rules, and obligations. The notion of social capital mediates these divergent orientations by acknowledging both the self-interest of actors and the influence of the social and economic context.

As social capital refers to the aspects of the social structure that are of value to social actors as resources that can be mobilized in pursuit of their interests, it is defined by its function. Social capital is not located in the actors themselves (as with human capital) or in any physical implements of production (as with physical capital). Instead, it is located in the relationships and personal networks between and among social actors. Social capital, then, appears in a variety of forms that have two common elements: (1) Social capital appears as an aspect of social structures, and (2) actors are able to use social capital as a resource to achieve their goals within the social structure.

Coleman (1988, 1990) conceives of social capital as a social structural resource that is a capital asset for the individual. It is productive, making it possible to achieve certain goals that cannot be achieved in its absence, and it is constituted within social organizations often as a by-product of activities undertaken for other purposes. For Coleman,

the value of the concept lies primarily in accounting for different outcomes of individual efforts and illuminating how resources can be combined with other resources to create system-level differences. The concept of social capital is most useful in qualitative analyses of social systems and studies relying on qualitative indicators.

### FORMS OF SOCIAL CAPITAL

Coleman (1990) identifies six forms of social capital: obligations and expectations, information potential, norms and effective sanctions, authority relations, appropriable social organizations, and intentional organizations.

**Obligations and Expectations.** A social system that relies heavily on reciprocal actions creates obligations and expectations on the part of its participants. Each “favor” is expected to be repaid, and those who can provide “favors” are expected to do so when requested. This form of exchange engenders social capital for a group member who has done many favors without collecting reciprocal favors in return. These unreciprocated favors create obligations that allow the favor-granting member to request aid from those who are obligated to her or him. These unpaid obligations accrue in the form of social capital that the member can use.

This form of social capital has two critical elements: the mutual trust within the social system and the extent of outstanding obligations. Without trust that obligations will be reciprocated, there is no incentive to accrue social capital. For social capital to have value, there must be trust that the resources will be there to be drawn on when needed. Furthermore, it is the extent of outstanding obligations that denotes the amount of social capital an individual can draw on. Indeed, the overall number of outstanding obligations within a system can be a measure of its interconnectedness as members are obligated to one another. This connectedness also increases the resources available to each member.

**Information Potential.** By interacting with informed members, individuals can increase their knowledge without having to obtain the information directly, whether by reading the newspaper or by interpreting research findings. A member also may become privy to specialized information—

such as unadvertised business opportunities—through informal information exchange. Useful information can be the impetus for action that furthers the individual's on goals and can be a beneficial commodity.

**Norms and Effective Sanctions.** Within a social system, norms can support and provide rewards for specific behaviors. Norms that encourage the subjugation of self-interest to the needs of the community are an especially powerful form of social capital. By promoting certain activities, norms by nature constrain other activities. Criminal pursuits are an obvious example of activities that communities want to constrain. Less obvious examples would be that the promotion of athletic activities constrains the time available for other interests and that norms that promote conformity constrain innovation. Effective “norms can constitute a powerful form of social capital” (Coleman 1990, p. 311).

**Authority Relations.** Within groups organized to address a specific issue, a leader often is chosen and given the right to make decisions and speak for the group. Thus, the members of the group transfer the “rights of control” to one individual, who then has access to an extensive network of social capital that can be directed toward a specific goal. When the rights of control are located in one individual, the social capital of all the members is amplified. Examples of this form of social capital can be found in political action groups, business cartels, and grassroots organizations.

**Appropriable Social Organizations.** Social organizations usually are created to address a specific issue, and after that issue is resolved, the organization often continues to exist through a redefinition of its goals. Thus, an organization that was developed for one purpose can be appropriated for another purpose. This constitutes a form of social capital available for use.

**Intentional Organizations.** This form of social capital occurs when individuals join together to create an organization that will benefit them directly. An example of this can be a joint business venture or a voluntary association that produces a public good, such as a Parents and Teachers Association (PTA) chapter. This form of social capital advances the interests of those who invest in it. Additionally, it can create two by-products as social capital: a public good that benefits others who

did not invest directly and a social organization that can be appropriated for other purposes.

These six forms of social capital have certain properties that distinguish them from other assets. Social capital's value lies in its use and cannot be exchanged easily. Additionally, it is not the private property of those who benefit from it. Rather, it is most often a by-product of other intentioned actions. Individuals who invest in the creation of the necessary social structures (norms, reciprocal obligations, etc.) are not the primary beneficiaries of the social capital that is generated. Instead, the social capital profits all those who are part of the social structure. Thus, it is not necessarily in an individual's self-interest to bring social capital into being. If another form of assistance that does not incur an obligation, such as a government program, is available or if individuals can meet their needs through self-sufficiency, they may choose to utilize these resources rather than those which accrue social capital through reciprocity. Using such forms of assistance or sufficiency does not increase the general pool of social capital in the community.

Coleman (1990) identifies factors that can increase or diminish social capital. A high degree of *closure* in a social network strengthens its ability to engender norms and effective sanctions. Closure is also important in achieving mutual trust among members. Conversely, lower closure reduces the effective norms and trust that would lessen the social capital engendered. The *stability* of a social structure affects the development or destruction of social capital. Individual mobility can threaten the stability of an organization, which in turn threatens its social capital. Organizations and social relations that are dependent on specific individuals are less stable than are those which rely on positions that can be filled by various individuals. Those organized around positions have more stability and thus a more steady supply of social capital on which the members may draw. *Ideology* can create social capital by influencing individuals to act in the interest of the whole rather than in their own interests. Religious ideologies are an example of this factor. Alternatively, ideologies of self-sufficiency or individualism can hinder the generation of social capital. Other factors Coleman mentions include affluence, availability of official aid such as governmental support, and other factors that may make individuals less de-



pendent on each other. Additionally, social capital must be maintained, as it depreciates over time and must be renewed.

### REFINING SOCIAL CAPITAL THEORY

Portes and Sensenbrenner (1993) have extended Coleman's basic theory of social capital, claiming that it contains two specific shortcomings. First, they call for a more defined discussion of the forms of social capital and how they are developed. Second, they question Coleman's optimistic instrumentalist orientation which focuses on the positive side of social capital. Furthermore, Portes and Sensenbrenner reidentify the sociological origins of notions of social capital by grounding their contribution in the works of classical social theorists. They begin by redefining social capital as "those expectations for action within a collectivity that affect the economic goals and goal-seeking behavior of its members" (1993, p.1323). This definition differs from Coleman's emphasis on social structures that can facilitate individual actions.

Portes and Sensenbrenner (1993) outline four different types of social capital, with each one corresponding to a classical tradition. From Parsons and Durkheim, they define *value introjection* as the first source of social capital because it promotes behaviors based on morals and values rather than on self-interest. From Durkheim's work, they conceive of economic transactions as reflections of an underlying moral order and contracts as reframing existing norms or values rather than creating new rules. Their second source of social capital—*reciprocity transactions*—comes from the work of Simmel and focuses on the dynamics of group membership. Reciprocity transactions are the obligations and expectations, backed by norms of reciprocity, that emerge through social networks of exchange. The third form of social capital is *bounded solidarity*, developed from Marx and Engels's writings on situational circumstances that lead to principled group action. This type of social capital is created when individuals join together in response for an adverse situation. The fourth form of social capital is drawn from Weber's distinction between formal and substantive rationality in market transactions; *enforceable trust*. Enforceable trust refers to the different mechanisms that formal institutions and particularistic group settings use to engender members' disciplined compliance with group norms

and expectations. Formal institutions use legal, rational mechanisms, whereas particularistic groups utilize substantive, social means.

**Acknowledging the Negative Effects.** Portes and Sensenbrenner also point out the negative effects of social capital in direct contrast to Coleman's more positive position. The same social structures "that give rise to appropriable resources for individual use can also constrain action or even derail it from its original goals" (1993, p. 1338). The costs of solidarity can be obligations and interconnections within the community. The costs of community conformity can constrain individual freedom. Obligations to the community can inhibit attempts to succeed in a broader network with a richer array of rewards. Individuals may come to see these costs of developing and maintaining social capital as being too high and not in their best interest.

In spite of its inchoate state and the ongoing debates, Coleman's optimistic discussion of social capital has become popular with policymakers who see it as the key to solving a variety of social problems. These policy proponents identify social capital as the features of social organizations (networks, norms, and social trust) that enable cooperative efforts for mutual benefit. Social capital is favored for its ability to promote and maintain voluntary associations that allow individuals to work together to resolve collective difficulties. In direct response, Portes and Landolt (1996) warn of the "downside" of social capital by making three specific criticisms of this use of social capital. First, social capital is discussed as the property of groups (communities), not individuals. Second, no distinction is made between the ability to access resources and the quality of those resources: Having networks is not enough; the networks need to have sufficient resources of value to make a difference. Third, policy proponents focus exclusively on the positive benefits of high levels of social capital and ignore the possibility of negative consequences.

Portes and Landolt (1996) specifically identify the less desirable possibilities of developing social capital. Social networks can promote "public bads" just as easily as public goods. Social capital can contribute to discrimination, restriction of individual freedom and creativity, lack of economic opportunity, and overwhelming obligations. Strong

voluntary associations, communities, and social networks that maintain high levels of solidarity often do so by excluding outsiders. Thus, noninsiders are disadvantaged within those groups. Additionally, high social capital is contingent on a high degree of conformity within the group, and non-conformists can be ostracized. This greatly impinges on personal freedom and expression. It also can result in a great deal of power for those in leadership positions in the group. Mafia-type power structures are an example of this.

Tight social networks also can undermine entrepreneurial activity. Successful business owners often are expected to help others, and this can affect their ability to maintain their businesses. Portes and Landolt (1996) identify further “downward leveling pressures” that can be consequences of social capital. The pressure to conform to group norms in order to access group resources (which may be perceived as the only resources available) can keep an individual from attempting to enter the mainstream and find a way up from poverty. Portes and Landolt use the examples of prostitution rings and youth gangs. The network norms function to keep individuals within the familiar group culture. Any attempt by a member to achieve something outside the network may be seen as a threat to group solidarity and is discouraged.

### SOCIAL CAPITAL AND SOCIOLOGY

Within the discipline, sociologists recognize the need to conduct empirical investigations as an important component of theory building. The concept of social capital has been advanced in many diverse subfields of sociology. Sociologists have applied it to the macro issues of modernization, economic development or lack of it, networks, and organizations. Others have studied the empirical implications of social capital for families and youth behavior problems, schooling and education, community life, work and organizations, democracy and governance, and collective action (see Woolcock 1998 for an overview).

As a theoretical concept, social capital holds great promise for furthering the sociological understanding of social action. There is still much to learn; the perspective needs to be grounded in established bodies of empirical research before it can be translated into optimistic public policies. Its greatest promise, Woolcock (1998, p.188) points

out, “is that it provides a credible point of entry for sociopolitical issues into a comprehensive multi- and interdisciplinary approach to some of the most pressing issues of our time.” Social capital may be seen as a common theoretical language that can allow historians, political scientists, anthropologists, economists, sociologists, and policymakers to work together in an open and constructive manner.

### REFERENCES

- Baron, James, and Michael Hannan 1994 “The Impact of Economics on Contemporary Sociology.” *Journal of Economic Literature*, 32:1111-1146.
- Bourdieu, Pierre 1986 “The Forms of Capital.” In John G. Richardson, ed., *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood Press.
- Coleman, James 1988 “Social Capital in the Creation of Human Capital.” *American Journal of Sociology*, 94: S95-S120.
- 1990 *Foundations of Social Theory*, Cambridge, Mass. Belknap Press of Harvard University Press.
- Portes, Alejandro, and Patricia Landolt 1996 “The Downside of Social Capital” *American Prospect*, 26:18-21, 94.
- , and Julia Sensenbrenner 1993 “Embeddedness and Immigration: Notes on the Social Determinants of Economic Action.” *American Journal of Sociology*, 98(6):1320-1350.
- Woolcock, Michael 1998 “Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework” *Theory and Society*, 27(2):151-208.

TRACY X. KARNER

### SOCIAL CHANGE

Social change is ubiquitous. Although earlier social scientists often treated stability as normal and significant social change as an exceptional process that required a special explanation, scholars now expect to see change at all times and in all social organizations. Much of this type of change is continuous; it occurs in small increments and reveals long-term patterns such as growth. Discontinuous changes, however, are more common than has been assumed. From the perspective of individual organizations, these changes are relatively common and often result in sharp departures from

previous states such as when corporations are created, merged, or terminated. From the perspective of larger populations of such organizations, relatively few discontinuous changes result in comparably sharp departures from long-term patterns and trends. Even revolutions that result in dramatic changes of political and legal institutions generally do not transform all of society equally. Some previous patterns continue; others are restored.

Cumulative social change must be distinguished from recurrent fluctuations and the processual aspect of all social life. Both sociologists and historians study the latter by focusing on those dynamic processes through which the social lives of particular individuals and groups may change even though overall patterns remain relatively constant. Marriages and divorces are major changes in social relationships, but a society may have a roughly constant marriage or divorce rate for long periods. Similarly, markets involve a continuous flow of changes in regard to who possess money or goods, who stands in the position of creditor or debtor, who is unemployed or employed, and so forth. These specific changes, however, generally do not alter the nature of the markets. Researchers both study the form of particular transactions and develop models to describe the dynamics of large-scale statistical aggregations of such processes (see "Social Dynamics.")

As Bourdieu (1977, 1990) and Giddens (1986) suggest, it is necessary to see human social life as always being structured, but incompletely so. "Structuration," to use their term, is as much a process of change as a reflection of stability. Indeed, the existence of stable social patterns over long periods requires at least as much explanation as does social change. This situation has led to renewed attention to social reproduction, or the ways in which social patterns are re-created in social action. This contrasts with earlier views of continuity as a matter of inertia or simple endurance. Some continuity in the social order is achieved intentionally by actors with enough power to resist changes desired by others; rulers thus maintain their rule by force. Much social reproduction, however, works at a less consciously intentional level and is based on the ways in which people learn to think and act rather than on overt, material force. Bourdieu and Passeron (1977), for example, follow Weber in studying the ways in which

ingrained, habitual ways of deciding what new action fits an individual's situation work without conscious intention to reproduce overall social patterns. A pattern of inequality in educational attainment that is understood officially as meritocratic and is genuinely intended by teachers to be so thus may be reproduced in part because students from nonelite backgrounds unconsciously lower their expectations for themselves, expecting elites to do better. Teachers may unconsciously do the same thing. When decisions are to be made, such as whether to go to university, or which university to choose, elite students and their families are more likely to have the confidence and knowledge to invest in options with a higher long-term payoff.

To understand social change, thus, it is necessary also to understand what produces social continuity. It would be a mistake to explain social change always in terms of a new factor that intervenes in an otherwise stable situation. Rather, social change commonly is produced by the same factors that produce continuity. These factors may change in quantity or quality or in relation to each other.

Sometimes, however, specific processes of social life undergo long-term transformations. These transformations in the nature, organization, or outcomes of the processes are what is usually studied under the label "social change." Social life always depends, for example, on the processes of birth and death that reproduce populations through generations. These rates (adjusted for the age of a population) may be in equilibrium for long periods, resulting in little change in the overall size of a population. Alternatively birthrates may exceed death rates most of the time, resulting in gradual population growth, but periodic disasters such as war, famine, and pestilence may cut the population back. In this case, the population may show little or no cumulative growth, but instead exhibit a dynamic equilibrium in which every period of gradual increase is offset by one of rapid decline. Approximations to these two patterns characterize most of world history. Population growth generally has been quite slow, although periodic declines have not offset all the increases. In the last three hundred years, however, a new phenomenon has been noted. As societies industrialize and generally grow richer and change the daily lives of their members, they undergo a "fertility transition." First, improvements in nutrition, sanitation,

and health can allow people to live longer. This results in population growth that can be very rapid if the improvements are introduced together rather than gradually developing over a long period. After a time lag, this encourages people to have fewer children because more of the children they do have survive. As fertility rates (birthrates standardized by the number of women of child-bearing age) also drop, a new equilibrium may be reached; population growth will slow or stop. This is a cumulative transition, because after it, the typical rates of birth and death are much lower even though the population may be much larger. A variety of other changes may follow from or be influenced by this process. For example, family life may change with declining numbers of children, parents' (especially mothers') lives are likely to change as fewer of their years are devoted to bearing and raising children, and childhood deaths may become rarities rather than common experiences.

Social history is given its shape by such cumulative social changes. Many of these changes are quite basic, such as the creation of the modern state; others are more minor, such as the invention and spread of the handshake as a form of greeting. Most, such as the development of team sports, fast-food restaurants, and the international, academic conference, lie in the broad area in between. Thus, cumulative social changes may take place on a variety of different scales, from the patterns of small group life through institutions such as the business corporation or church to overall societal arrangements. Significant changes tend to have widespread repercussions, however, and so it is rare for one part of social life to change dramatically without changing other parts.

While certain important changes, such as an increasing population, are basically linear, others are discontinuous. There are two senses of discontinuity. The first is abruptness, such as the dramatic contraction of the European population in the wake of plague and other calamities of the fourteenth century and the occurrence of the Russian Revolution after centuries of tsarist rule and failed revolts. Second, some social changes alter not just the values of variables but the relationship of variables to each other. Thus, for much of history the military power and wealth of a ruler was based directly on the number of his or her subjects; growing populations meant an increasing total

product from which to extract tribute, taxes, and military service. With the transformation first of agriculture and then of industrial production in the early capitalist era or just before it, this relationship was in many cases upset. Increasingly, from the sixteenth through eighteenth centuries, for example, the heads of Scottish clans found that a small population raising sheep could produce more wealth than could a large one farming; their attempt to maximize this advantage contributed to the migration of Scots to Ireland and America. This process was of course linked also to growing demand for wool and the development of the industrial production of textiles. Those factors in turn involved new divisions of social labor and increased long-distance trade. At the same time, the development of industrial production and related weapons technologies reduced the military advantages of large population size by contrast to epochs when wars generally were won by the largest armies; indeed, population may be inversely related to power if it impedes industrialization.

This case provides an example of how shifts in the relationships of certain variables can alter not only overall social patterns but broad cultural orientations to social change. Along with industrialization (and other dimensions of modern social life) has come a continuous process of technological and social innovation. As Weber (1922) emphasized, this process is at odds with a traditional orientation to social life. Traditionalism implies an expectation of continuity and respect for the ways in which things have always been done. Constant innovation is linked to the pursuit of more efficient ways to do things and an expectation of continuous change. Leaders of China, long thought the absolute size of armies would be decisive in conflict. They were shocked when both Japan and Western powers were able to win victories in the nineteenth century mainly on the basis of superior technology rather than superior size. This helped produce not only the collapse of a specific imperial dynasty but a crisis in a whole pattern of traditionalism. Instead of assuming that the best lessons for military strategy lay in the teachings of the past, some leaders recognized that they needed to look for new ways in which to make the country strong. This produced a tension between trying to preserve cultural identity by continuing to do things the same way and trying to achieve technological and other gains by innovating. This tension is

common in societies that have undergone broad patterns of social change in the modern era. In China, after the death of Mao Zedong leaders decided that strengthening the country and improving people's lives depended on technological advancement and economic development. Recognizing both that large armies would not win wars against enemies with technologically advanced weapons and that rapid population growth would make it difficult to educate the whole population and produce rapid economic growth, Deng Xiaoping and other leaders introduced policies to reduce population growth rates. They also decided that they needed to liberalize the economy and encourage private business because state-owned enterprises could not innovate rapidly enough. On the one hand, they encourage innovation in economy and technology, and on the other hand, they resist change in politics and culture. Although perhaps contradictory, these two responses have been typical of leaders in societies undergoing the process of modernization. Although it is impossible to prevent major changes in technology and the economy from having an impact on politics and culture, it is possible to shape what those impacts will be.

Sociologists generally have taken three approaches to studying cumulative social changes. The first is to look for generalizable patterns in how all sorts of changes occur, the second is to seek an explanation for the whole overall pattern of history, and the third is to analyze historically specific processes of change.

Following the first of these approaches, sociologists have looked for characteristic phases through which any social innovation must pass, such as skepticism, experimentation, early diffusion among leaders, and later general acceptance. Ogburn ([1922] 1950) was a pioneer in this sort of research, examining topics such as the characteristic "lag" between cultural innovations and widespread adjustments to them or exploitation of their potential. In regard to the fertility transition, when improved health care and nutrition make it possible for nearly all children to survive to adulthood, it takes a generation or two before parents stop having large families as "insurance policies" to provide for support in their old age. Earlier researchers often hoped to find general laws that would explain the duration of such lags and account for other features of all processes of social change. Contemporary sociologists tend to place

much more emphasis on differences among various kinds of social change and their settings; accordingly, their generalizations are more specific. Researchers may limit their studies to the patterns of innovation among business organizations, for example, recognizing that those organizations may act quite differently from others. They also may ask questions such as, Why do innovations gain acceptance more rapidly in formal organizations (e.g., businesses) than in informal, primary groups (e.g., families), or what sorts of organizations are more likely to innovate? The changes may be very specific, such as the introduction of new technologies of production, or very general, such as the Industrial Revolution as a whole (Smelser 1958). The key distinguishing feature of these sorts of studies is that they regard changes as individual units of roughly similar sorts and aim to produce generalizations about them.

The second major sociological approach to cumulative change—seeking an explanation for the whole pattern of cumulation—was long the province of philosophies of history that culminated in the sweeping syntheses of the nineteenth century. Sociology was born partly out of the attempt to understand the rise of science, industry, and urban society. These and related transitions were conceptualized in frameworks that emphasized shifts from tradition to modernity, feudalism to capitalism, and monarchy to republicanism or democracy. As Sztompka (1993) points out, three basic visions were developed, each of which has left a mark on sociology and continues to be influential in research: cycles, evolutionary progress, and historical materialism. The roots of the cyclical vision stretch back to antiquity. The image of the human life cycle, from birth and infancy to old age and death, for example, was used to conceptualize the rise and fall of whole societies and of imperial dynasties that were thought to be vigorous in youth and feeble in old age. Few scientific sociologists have regarded such images as more than metaphors, but they have been influential among writers attempting to generalize about the course of history (e.g., Spengler [1918] 1939; Toynbee 1934–1961). A number of sociologists, however, have studied more specific cyclical patterns. Pareto ([1916] 1980) analyzed what he called the circulation of elites, a pattern in which specific groups rose into and then fell from social dominance. Sorokin (1937) analyzed cultural cy-

cles, especially the oscillating dominance of ideational (spiritual, intellectual) and sensate (sensual, materialist) orientations. More recently, sociologists have identified cycles in social movements and collective action (Tilly 1989; Tarrow 1998; Traugott, 1995).

Both historical materialism and evolutionism are indebted to another ancient idea, that of progress. Here the idea is that social change tends to produce a pattern of improvements in human life as measured in relationship to a standard of evaluation. In this regard, sociological evolutionism has commonly differed from evolutionary theory in biology, which has been less focused on the overall direction of change and normative evaluation. The great nineteenth-century evolutionary thinkers Comte ([1830–1842, 1851–1854] 1975) and Spencer (1893) conceptualized history as progress through a series of stages. Comte based his analysis on what he saw as improvements in social knowledge through theological, metaphysical, and positive stages. Spencer, who was also an originator of evolutionary theory in biology, had a much more complex and sophisticated theory, focusing on the way structures developed to meet functional imperatives and gaining direction from the idea that “incoherent homogeneity” progressively gives way to “coherent heterogeneity” through the process of structural differentiation. Spencer (1893) addressed particularly the transition from military to industrial societies, which he saw as basic to modernity. Durkheim (1893) developed a similar analysis in his description of the movement from mechanical to organic solidarity.

[<sup>1</sup> Schrecker (1991) has analyzed a pattern in which something similar to Spencer’s two stages alternated cyclically in Chinese society rather than forming the basis for a single evolutionary trend. Periods of increasing industrialization and commercialization (*fengjian*) were followed by eras in which agriculture and military prowess figured more prominently (*junxian*). Schrecker (1991) suggests that this intriguing combination of evolutionary and cyclical theories initially was developed by classical Chinese scholars, although it was recast after the importation of Spencerian evolutionary theory.]

Historical materialists, starting with Marx (1863), also analyzed stages in historical development (such as feudalism and capitalism), but with

three crucial differences from other evolutionary theories. First, Marx and his followers argued that material factors, especially the mode of production, shape the rest of society and that change is driven largely by improvements in the capacity for material production. Second, following a dialectical logic, Marxists emphasized the internal contradictions within each stage of development. Capitalism, for example, generated tremendous increases in productivity but distributed the resulting wealth so unequally that it was prone to economic crises and social revolutions. Rather than a simple, incremental progress, thus, Marxists saw evolution as taking place in discontinuous breaks marked by clashes and struggles. Third, most versions of Marxist theory gave greater emphasis to human agency or ability consciously to shape the direction of social change than was typical of evolutionary theory. The question of the extent to which evolution can be directed consciously has, however, recently come to the fore of non-Marxist evolutionary theory as well, as in in the work of the sociobiologist Wilson (Wilson and Wilson 1999).

The most important contemporary theories of social evolution attempt to generate not only overall descriptions of stages but causal explanations for social change. Lenski, for example, has argued that increases in technological capacity (including information processing as well as material production and distribution) account for most of the major changes in human social organization (Lenski et al. 1994). In his synthesis, Lenski arranges the major forms of human societies in a hierarchy based on their technological capacity and shows how other features, such as their typical patterns of religion, law, government, class inequality, and relations between the sexes, are rooted in those technological differences. In support of the idea that there is an overall evolutionary pattern, Lenski et al. (1994) point to the tendency of social change to move only in one direction. Thus, there are many cases of agricultural states being transformed into industrial societies but very few (if any) examples of the reverse. Of course, Lenski acknowledges that human evolution is not completely irreversible; he notes, however, not only that cases of reversal are relatively few but that they commonly result from an external cataclysm. Similarly, Lenski indicates that the direction of human social evolution is not strictly dictated from the start but only channeled in certain direc-

tions. There is room for human ingenuity to determine the shape of the future through a wide range of potential differences in invention and innovation. There are a number of other important versions of the evolutionary approach to cumulative social change. Some stress different material factors, such as human adaptation to ecological constraints (Harris 1979; White 1949); others stress culture and other patterns of thought more than material conditions (Parsons 1968; Habermas 1978).

Adherents to the third major approach to cumulative social change argue that there can be no single evolutionary explanation for all the important transitions in human history. They also stress differences as well as analogies among particular instances of specific sorts of change (Stinchcombe 1978). These historians and historical sociologists emphasize the importance of dealing adequately with particular changes by locating them in their historical and cultural contexts and distinguishing them through comparison (Abrams 1982; Skocpol 1984; Calhoun 1995, 1998). Weber was an important pioneer of this approach. A prominent variety of Marxism has stressed the view that Marx's mature analysis of capitalism emphasizes historical specificity rather than the use of the same categories to explain all of history (Postone 1993). Historical sociologists have argued that a particular sort of transformation, such as the development of the capacity for industrial production, may result from different causes and have different implications on different occasions. The original Industrial Revolution in eighteenth- and nineteenth-century Britain thus developed with no advance model and without competition from established industrial powers. Countries that are industrializing today are influenced by both models and competition from existing industrial countries, along with influences from multinational corporations. The development of the modern world system thus fundamentally altered the conditions of future social changes, making it misleading to lump together cases of early and late industrialization for the purpose of generalization. Similarly, prerequisites for industrial production may be supplied by different institutional formations; one should compare not just institutions but different responses to similar problems.

Accident and disorder also have played crucial roles in the development of the modern world system. Wallerstein (1974–1988) shows the cen-

trality of historical conjunctures and contingencies: the partially random relationships between different sorts of events (on historical accidents, see also Simmel 1977; Boudon 1986). For example, the outcome of military battles between Spain (an old-fashioned empire) and Britain (the key industrial-capitalist pioneer) were not foregone conclusions. There was room for bravery, weather, strategy, and a variety of other factors to play a role. However, certain key British victories, notably in the sixteenth century, helped make not only British history but world history different by creating the conditions for the modern world system to take the shape it did. Against evolutionary explanation, historical sociologists also argue that different factors explain different transformations. Thus, no amount of study of the factors that brought about the rise of capitalism and industrial production can provide the necessary insight into the decline of the Roman Empire and the eventual development of feudalism in Europe or the consolidation of China's very different regions into the world's most enduring empire and most populous state. These different kinds of events have their own different sorts of causes.

Predictably, some sociologists seek ways to combine some of the benefits of each type of approach to explaining cumulative social change. Historical sociologists who emphasize the singularity of specific transformations can learn from comparisons among such changes and achieve at least partial generalizations about them. Thus, different factors are involved in every social revolution, yet certain key elements seem to be present, such as crises (financial as well as political) in a government's capacity to rule (Skocpol 1979; Goldstone 1991). This recognition encourages one to focus on structural factors that may help create potentially revolutionary situations as well as the ideologies and actions of specific revolutionaries. Similarly, even though a variety of specific factors may determine the transition to capitalism or industrialization in every instance, some version of a fertility transition seems to play a role in nearly all cases. Although evolutionary theory is widely rejected by historical sociologists, some look to evolutionary arguments for suggestions about what factors might be important. Thus, Lenski's emphasis on technology and Marx's focus on the relationship of production and class struggle can provide foci for research, and that research can help deter-

mine whether those factors are equally important in all societal transformations and whether they work the same way in each one. More radically, evolutionary socioglogy might follow biology in focusing less on the selection of whole populations (societies) for success or failure and look instead at the selection of specific social practices (e.g., the bearing of large numbers of children) for reproduction or disappearance. Such an evolutionary theory might provide insight into how practices become more or less common, following biology in looking for mechanisms of reproduction and inheritance, the initiation of new practices (mutation), and the clustering of practices in interacting groups (speciation) as well as selection. It would, however, necessarily give up the capacity to offer a single explanation for all the major transitions in human social history, which is one of the attractions of evolutionary theory to its adherents.

Certain basic challenges are particularly important in the study of cumulative social change today. In addition to working out a satisfactory relationship among the three main approaches, perhaps the most important challenge is to distinguish social changes that are basic from those which are ephemeral or less momentous. Sociologists, like historians and other scholars, need to be able to characterize broad patterns of social arrangements. This is what sociologists do when they speak of “modernity” or “industrial society.” Such characterizations involve at least implicit theoretical claims about the crucial factors that distinguish these eras or forms. In the case of complex, large-scale societal processes, these factors are hard to pin down. How much industrial capacity does a society need to have before one can call it industrial? How low must employment in its increasingly automated industries become before one can call it postindustrial? Is current social and economic globalization the continuation of a long-standing trend or part of a fundamental transformation? Although settling such questions is difficult, debating them is crucial, for sociologists cannot grasp the historical contexts of the phenomena they study if they limit themselves to studying particulars or seeking generalizations from them without attempting to understand the differences among historical epochs (however hard to define sharply) and cultures (however much they may shade into each other with contact). Particularly because of the many current contentions that

humanity stands on the edge of a new age—postmodern, postindustrial, or something else—researchers and theorists need to give strong answers to the question of what it means to claim that one epoch ends and another begins (Calhoun 1999).

Many prominent social theorists have treated all of modernity as a continuous era and stressed its distinction from previous (or anticipated future) forms of social organization. Durkheim (1893) argued that a new, more complex division of labor is central to a dichotomous distinction of modern (organically solidary) from premodern (mechanically solidary) society. Weber (1922) saw Western rationalization of action and relationships as basic and as continuing without rupture through the whole modern era. Marx (1863) saw the transition from feudalism to capitalism as basic but held that no change in modernity could be considered fundamental unless it overthrew the processes of private capital accumulation and the commodification of labor. Recent Marxists thus argue that the social and economic changes of the last several decades mark a new phase within capitalism but not a break with it (Mandel 1974; Wallerstein 1974–1988; Harvey 1989). Many sociologists would add a claim about the centrality of increasing state power as a basic, continuous process of modernity (e.g., Tilly 1990; Mann 1986–1993). More generally, Habermas (1984–1988) has stressed the split between a life world in which everyday interactions are organized on the basis of mutual agreement and an increasingly prominent systemic integration through the impersonal relationships of money and power outside the reach of linguistically mediated cooperative understanding. Common to all these positions is the notion that there is a general *process* (not just a static set of attributes) common to all forms of modernity. Some claim to discern a causal explanation; others only point to the trends, suggesting that those trends may have several causes but that there is no single “prime mover” that can explain an overall pattern of evolution. All would agree that no really basic social change can be said to have occurred until the fundamental processes they identify have ended, been reversed, or changed their relationship to other variables. Obviously, a great deal depends on what processes are considered fundamental.

Rather than stressing the common processes that organize all forms of modernity, some scholars have followed Marx (and recent structuralist



theory) in pointing to the disjunctures between relatively stable periods. Foucault (1973), for example, emphasized basic transformations in the way knowledge is constituted and an order is ascribed to the world of things, people, and ideas. Renaissance culture was characterized by an emphasis on resemblances among the manifold different elements of God's single, unified creation. Knowledge of fields as diverse to modern eyes as biology, aesthetics, theology, and astronomy was thought to be unified by the matching of similar characteristics, with those in each field serving as visible signs of counterparts in the others. The "classical" modernity of the seventeenth and early eighteenth centuries marked a radical break by treating the sign as fundamentally distinct from the thing it signified, noting, for example, that words have only arbitrary relationships to the objects they name. The study of representation thus replaced that of resemblances. In the late eighteenth and early nineteenth centuries, another rupture came with the development of the modern ideas of classification according to hidden, underlying causes rather than superficial resemblances and an examination of human beings as the basic source of systems of representation. Only this last period could give rise to the "human sciences"—psychology, sociology, and so forth—as they are known today. Similarly, Foucault (1977) argued that the modern individual is a distinctive form of person or self, produced by an intensification of disciplining power and surveillance. Where most theories of social change emphasize processes, Foucault's "archaeology of knowledge" emphasizes the internal coherence of relatively stable cultural configurations and the ruptures between them.

Foucault's work has been taken as support for the claim (which was not his own) that the modern era has ended. Theories of "postmodernity" commonly argue that at some point the modern era gave way to a successor, though some scholars (e.g., Lyotard 1977) have indicated, against the implications of the label "postmodern," that they mean not a simple historical succession but a recurrent internal challenge to the dominant "modernist" patterns (see Lash 1990; Seidman 1995; Harvey 1989; Calhoun 1995). Generally, they hold that where modernity was rigid, linear, and focused on universality, postmodernity is flexible, fluidly multidirectional, and focused on differ-

ence. Some postmodernist theories emphasize the impact of new production technologies (especially computer-assisted flexible automation), while others are more exclusively cultural. The label "postmodernity" often is applied rather casually to point to interesting features of the present period without clearly indicating why they should be taken as revealing a basic discontinuous shift between eras.

At stake in debates over the periodization of social change is not just the labeling of eras but the analysis of what factors are most fundamentally constitutive of social organization. Should ecology and politics be seen as determinative over, equal to, or derivative of the economy? Is demography or technological capacity prior to the other? What gives capitalism, feudalism, a kinship system, or any other social order its temporary and relative stability? Such questions must be approached not just in terms of manifest influence at any single point in time or during specific events but also in terms of the way particular factors figure in long-term processes of cumulative social change.

## REFERENCES

- Abrams, Philip 1982 *Historical Sociology*. Ithaca, N.Y.: Cornell University Press.
- Boudon, Raymond 1986 *Theories of Social Change*. Cambridge, UK: Cambridge University Press.
- Bourdieu, Pierre 1977 *Outline of a Theory of Practice*. Cambridge, UK: Cambridge University Press.
- , 1990 *The Logic of Practice*. Stanford, Calif.: Stanford University Press.
- , and Jean-Claude Passeron 1977 *Reproduction in Education, Culture, and Society*. London: Sage
- Calhoun, Craig 1995 *Critical Social Theory: Culture, History, and the Challenge of Difference*. Cambridge, Mass.: Blackwell.
- 1998 "Explanation in Historical Sociology: Narrative, General Theory, and Historically Specific Theory." *American Journal of Sociology* 108:846–871.
- 1999 "Nationalism, Social Change, and Historical Sociology," In F. Engelstad and R. Kalleberg, eds., *Social Change and Historical Sociology*. Oslo: Scandinavian University Press.
- Comte, Auguste (1830–1842, 1851–1854) 1975 *Auguste Comte and Positivism: The Essential Writings*, edited by G. Lenzer. New York: Harper.
- Durkheim, Emile 1893 *The Division of Labor in Society*. New York: Free Press.

- Foucault, Michel 1973 *The Order of Things: An Archaeology of the Human Sciences*. New York: Random House.
- 1977 *Discipline and Punish*. New York: Pantheon.
- Giddens, Anthony 1986 *The Constitution of Society*. Berkeley: University of California Press.
- Goldstone, Jack 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.
- Habermas, Jürgen 1978 *Communication and the Evolution of Society*. Boston: Beacon.
- 1984–1988 *The Theory of Communicative Action*, 2 vols. Boston: Beacon.
- Harris, Marvin 1979 *Cultural Materialism*. New York: Vintage.
- Harvey, David 1989 *The Postmodern Condition*. Oxford, UK: Oxford University Press.
- Lash, Scott 1990 *Postmodern Sociology*. London: Routledge.
- Lenski, Gerhard, Jean Lenski, and Patrick Nolan 1994 *Human Societies*, 7th ed. New York: McGraw-Hill.
- Lyotard, Jean-François 1977 *The Postmodern Condition*. Minneapolis: University of Minnesota Press.
- Mandel, Ernst 1974 *Late Capitalism*. London: Verso.
- Mann, Michael 1986–1993 *The Sources of Social Power*. Cambridge, UK: Cambridge University Press.
- Marx, Karl 1863 *Capital*, vol. 1. New York: Viking.
- Ogburn, W. F. (1922) 1950 *Social Change with Respect to Culture and Original Nature*. New York: Viking.
- Pareto, Wilfredo (1916) 1980 *Compendium of General Sociology*. Minneapolis: University of Minnesota Press.
- Parsons, Talcott 1968 *The Evolution of Societies*. New York: Free Press.
- Postone, Moishe 1993 *Time, Labor and Social Domination*. New York: Cambridge University Press.
- Schrecker, John 1991 *The Chinese Revolution in Historical Perspective*. New York: Praeger.
- Seidman, Steven 1995 *The Postmodern Turn: New Perspectives on Social Theory*. New York: Cambridge University Press.
- Simmel, Georg 1977 *The Problem of the Philosophy of History*. New York: Free Press.
- Skocpol, Theda 1979 *States and Social Revolutions*. New York: Cambridge University Press.
- , ed. 1984 *Vision and Method in Historical Sociology*. New York: Cambridge University Press.
- Smelser, Neil J. 1958 *Social Change in the Industrial Revolution*. London, Routledge and Kegan Paul.
- Sorokin, Pitirim 1937 *Social and Cultural Dynamics*, 4 vols. New York: American Book Company.
- Spencer, Herbert 1893 *Principles of Sociology*, 3 vols. London: Williams and Norgate.
- Spengler, Oswald (1918) 1939 *The Decline of the West*. New York: Knopf.
- Stinchcombe, Arthur 1978 *Theoretical Methods in Social History*. New York: Academic Press.
- Sztompka, Piotr 1993 *The Sociology of Social Change*. Oxford: Blackwell.
- Tarrow, Sidney 1998 *Power in Movement: Social Movements and Contentious Politics*. Cambridge, UK: Cambridge University Press.
- Tilly, Charles 1989 *The Contentious French*. Cambridge, Mass.: Harvard University Press.
- 1990 *Coercion, Capital, and European States, AD 990–1990*. Oxford, UK: Blackwell.
- Toynbee, Arnold 1934–1961 *A Study of History*, 12 vols. Oxford, UK: Oxford University Press.
- Traugott, Mark, ed. 1995 *Repertoires and Cycles in Collective Action*. Durham, N.C.: Duke University Press.
- Wallerstein, Immanuel 1974–1988 *The Modern World System*, 3 vols. San Diego: Academic Press.
- Weber, Max 1922 *Economy and Society*. Berkeley: University of California Press.
- White, Leslie 1949 *The Science of Culture*. New York: Grove Press.
- Wilson, Edward O., and Edmund O. Wilson 1999 *Consilience*. New York: Random House.

CRAIG CALHOUN

## SOCIAL COMPARISON PROCESSES

How do people come to understand themselves? A response to this age-old question involves what has been labeled everyone's "second favorite theory" (Goethals 1986): social comparison. The original formulation of social comparison theory (Festinger 1954) demonstrated how, in the absence of objective standards, individuals use other people to fulfill their informational needs to evaluate their own opinions and abilities. The process of social comparison underlies social evaluation (Pettigrew 1967) and relates to reference group processes (e.g., Hyman and Singer 1968), which in turn are critical to understanding diverse sociological issues pertaining, for example, to identity development, justice, interpersonal and intergroup rela-

tionships, and group decision making. Thus, the “second favorite” status of social comparison theory reflects the preference of researchers for particular theories about each of these topics, which nonetheless promote the centrality and breadth of social comparison processes in sociological pursuits. To explain the multifaceted role of social comparisons, this article first describes the theory and its elaborations and then shifts to a sampling of the extensive applications of that theory.

### SOCIAL COMPARISON THEORY

For nearly fifty years, social comparison theory has shifted between the categories of “lost and found” (Goethals 1986): The theory flourishes for a while and then lies dormant. Suls (1977) outlines the first twenty years of social comparison research, beginning with its inception in 1954 by the psychologist Festinger and then describing its theoretical decline while applications to affiliation (Schachter 1959), emotions (Schachter and Singer 1962), and justice (Adams 1965) emerged; its momentary revival in a 1966 issue of the *Journal of Experimental Social Psychology*; and its second, more enduring revival in 1977 in the form of a landmark volume of collected essays (Suls and Miller, 1977). During its fourth decade, the resurgence of social comparisons research (Wood 1989) anchored it firmly in the “found” category. Almost fifty years after its inception, its contribution to numerous substantive areas appears to be unrivaled (Suls and Wills 1991).

Festinger (1954) incorporates his observations regarding research on aspiration levels and social pressures into the premises of his formal theory of social comparison. First, individuals are driven to evaluate their abilities and opinions. That drive increases with the importance of the ability or opinion, its relevance to immediate behavior, the relevance of the group to the ability or opinion, and the individual’s attraction to the group. These factors also increase the pressure toward uniformity with relevant others. Second, people first attempt to make these evaluations through objective, nonsocial means, but if those means are unavailable, they are likely to compare themselves with others. Third, individuals are likely to choose for comparison someone close to the opinion or ability in question (the “similarity hypothesis”). Festinger notes that social comparisons may be

based on the similarity of attributes related to the dimension under evaluation (the “related attributes hypothesis”). As a rationale for the preference for similar comparison others, he argues that comparisons with divergent others produce imprecise and unstable evaluations. Comparisons with moderately different others (those within tolerable limits of discrepancy), however, produce changes in individuals’ evaluations of their own or the others’ abilities or opinions. These changes ensure uniformity in the group and reinforce stable and precise evaluations.

Although Festinger (1954) treats abilities and opinions similarly in most respects, he notes a major distinction that may influence the consequences of comparisons. In the case of abilities, the cultural value of doing “better and better” encourages individuals to make upward comparisons. Typically, because of the pressure toward uniformity, the upward drive is limited to comparisons with those who are slightly better. This unidirectional drive upward, however, inspires competition among group members that may inhibit the emergence of social uniformity. In contrast, no upward drive characterizes comparisons of opinions. In light of the pressure toward uniformity, the absence of an upward drive and the greater flexibility of opinions (compared with the nonsocial constraints on changes in abilities) suggest that a state of social quiescence is more likely to emerge in terms of the evaluation of opinions.

Despite the greater likelihood of opinion uniformity, Festinger (1954) notes that if potential comparison others have highly discrepant opinions, the cessation of the comparison process may result in hostility toward or derogation of those others. These negative reactions stem from the belief that opinion discrepancy means that an individual’s opinions are incorrect. In contrast, no negative implications characterize discrepancy in abilities, which are more independent of relative value orientations, indicating different forms of “correctness.”

As is suggested by his emphasis on group uniformity, one of Festinger’s (1954) major concerns was to describe the role of social comparison processes in group formation and maintenance as well as social structure. Presumably, people who have similar abilities and opinions group together; as a result of their distinguishing themselves from

others, segments form in society. These two implications of social comparison processes are represented in research on group decision making and intergroup relations (see below). Although Festinger's perspective suggests a wide range of implications, his theory is not definitive; in fact, Arrowood (1986) has labeled the 1954 version a "masterpiece of ambiguity." Points of ambiguity involve theoretical issues regarding the nature of social comparison per se, the motivations that underlie comparison choices, and the choice of a comparison other (especially the meaning of similarity).

In Festinger's (1957) formulation, the nature of the social comparison is oblique, referring only to nonobjective information regarding abilities and opinions. Others have extended the domain of social comparisons to include emotions (Schachter and Singer 1962), outcomes (Adams 1965), health (Buunk and Gibbons 1997), relationships (VanYperen and Buunk 1994), and traits (Thornton and Arrowood 1966, but see Suls 1986). Nonobjective information may include personal comparisons (see Masters and Keil 1987). These comparisons involve information about the self and draw attention to time as an important factor in determining the nature of comparisons. Temporal comparisons (Albert 1977) involve a "now versus then" dimension, meaning that a person compares pieces of information about an ability or opinion at different points in time.

Arrowood (1986) describes several forms of nonobjective (ability) information used in social comparison studies, for example, the presentation of two ratings (the evaluator's and that of an unidentified other), the display of a distribution of ratings that includes the evaluator's, and the presentation of the evaluator's rating along with the identifying characteristics of potential comparison others and their ratings. The latter two factors allow the development of an estimate about others in general, with the last one also allowing an assessment of similarity. Compared with objective information, relative standing generally exerts greater influence except when the desirability of objective information is great (Klein 1997). Not all methods, however, truly capture social comparison. To address this problem, Wood (1996) stresses the need for consistency between definitions and measures and the need to be wary of alternative interpretations. Insofar as information presenta-

tion stimulates the comparison process, its effects may depend on the motivations that drive the comparison.

The most explicit but general, motivation included in Festinger's (1954) perspective is that people need information. Later work, however, demonstrates that the motivations underlying social comparisons are far more extensive. Twenty years ago, Fazio (1979) specified two types of information motivations that underlie self-evaluations: (1) the "construction" motivation, referring to a person's desire to obtain information he or she lacks, and (2) the "validation" motivation, representing the use of information to determine the valid source (e.g., the person or the entity) of a person's judgment about an entity. Fazio argues that individuals appear to be motivated by construction when they lack information and by validations when they have sufficient information. Validation, however, is distinct from a third motivation: self (or ego) enhancement. Festinger (1954) hinted at the possibility of self-enhancement when he posited that in evaluating abilities, individuals are likely to make upward comparisons; by noting "how close" an individual is to a superior performer, that individual enhances the evaluation of his or her own ability. In addition to self-evaluation and self-enhancement goals, one study (Hegeson and Mickelson 1995) lists four other motives for social comparison: common bond, self-improvement, altruism, and self-destruction.

A recent trend in distinguishing among motivations is to focus on the role of self-esteem in social comparisons. Generally, individuals low in self-esteem rely more on social comparison information to meet goals of accuracy, self-enhancement, and self-improvement than do those high in self-esteem (Wayment and Taylor 1995). However, insofar as people with low self-esteem are more oriented to self-protection, they are likely to seek social comparison information that is "safe" (i.e., that carries little risk of humiliation) after receiving success feedback (Wood et al. 1994). In effect, low-esteem individuals are seizing a safe form of self-enhancement. Threats to self-esteem, especially among those who have high self-esteem, seem to stimulate more egocentric contrasts in judgments of others (Beauregard and Dunning 1998), suggesting that individuals tailor their evaluations of others to affirm their own self-worth.

That tailoring of evaluations coincides with other evidence that challenges Festinger's assumption that individuals are rational and accurate in information processing. Wood (1989) demonstrates that people are not unbiased evaluators of information about themselves and potential comparison others, and Hoorens's (1995) study captures some of those biases. Subjects comparing themselves to another person show more unrealistic optimism and illusory superiority regarding future positive events and traits than do subjects who compare another person to themselves. In other words, people perceive themselves more favorably than they perceive others. Motivations, coupled with perceptual processes, are critical in determining the choice of comparison others.

In an extensive review of the choice of comparison others, Gruder (1977) examines the conflict between evaluation and enhancement motivations. He concludes that the evaluation motivation is important in new situations and that under such conditions, individuals are likely to choose others who are similar to themselves in terms of the ability or opinion at issue. Enhancement motivations arise when enhancement is feasible. For example, in the absence of threats to self-esteem, individuals are likely to make upward, self-enhancing comparisons that provide useful information for self-improvement (Collins 1996; Blanton et al. 1999). If an upward comparison is unavoidable, individuals may deflect a threat to their self-esteem by perceptually distorting through exaggeration the performance of the other person (Alicke et al. 1997). When self-esteem is unavoidably threatened, people are more likely to make downward (defensive) comparisons (e.g., Hakmiller 1966; Taylor and Lobel 1989). In effect, to protect their self-esteem, individuals choose dissimilar others (i.e., inferior performers) for their comparisons. Researchers are examining a number of additional factors that may affect responses to upward and downward comparisons (e.g., Aspinwall and Taylor 1993). Whether individuals choose similar or dissimilar others for comparison depends in part on the nature of the dimension under scrutiny, the context (including the characteristics of potential comparison others), and the importance of the enhancement goal (Wood 1989).

Questions about the choice of comparison others typically raise the issue of the meaning of similarity. Festinger (1984) offers different yet

potentially complementary definitions of similarity: closeness of ratings on abilities or opinions and attributes related to the evaluation dimension. Goethals and Darley (1977) note that similarity in the first sense is paradoxical: "[P]resumably the comparison is made in order to find out what the other's opinion or score is, yet prior knowledge of the similarity of his score or opinion is assumed as the basis for comparison" (p. 265). Those authors advocate the interpretation of similarity on the basis of "related attributes." One is likely to choose others for comparison who *should be* close to one's own ability or opinion by virtue of their standing on characteristics related to the evaluative dimension. Wheeler et al. (1982) review the extensive support for the related attributes hypothesis. Two trends, however, qualify that support. Kulik and Gump (1997) show that related attribute information has an effect only in the absence of information on relative performances. And Wood (1989) indicates that people choose comparisons with those with similar characteristics regardless of whether the attribute relates to the dimension under scrutiny; for example, people of the same sex are more likely to compare themselves to each other even if sex is unrelated to the ability that is being compared.

Concern with related attributes similarity and its evaluative consequences provides the basis for one of the main elaborations of social comparison theory: Goethals and Darley's (1977) attributional approach. Their approach applies Kelley's (1973) attributional concepts of discounting and augmentation (see Howard in this volume) to assess the certainty of one's standing on an ability or opinion. With regard to abilities, attribution logic focuses on the configuration of possible comparison others and one's own ability level. For example, a person who compares himself or herself to an advantaged other expects his or her performance to be worse. The implications for an ability evaluation are ambiguous, however, because other plausible causes (the superior related attributes) allow the discounting of low ability. In contrast, if an individual performed as well as or better than the advantaged other, he or she overcame inhibitory causes (the inferior related attributes), and this augments a claim to higher ability. In general, conclusive ability evaluations are likely when a person compares herself or himself to others with similar attributes. The role of attributions in evalu-

ating opinions is more complicated because the unexpected cases (disagreement with similar others and agreement with dissimilar others) result in more useful information for validating an opinion than do the expected cases (agreement with similar others and disagreement with dissimilar others). More recent research has concentrated less on attributions than on other cognitive processes (e.g., the use of heuristics and memory) (see Masters and Keil 1987) and perceived control (see Major et al. 1991) to better understand social comparisons.

Complementing the intra- and interpersonal level focus of psychologists are sociological extensions to the level of intergroup relations. Tajfel and Turner (1979) build on Festinger's (1934) assumption that a major consequence of social comparisons to similar others is the development of groups or, in their terminology, social categories. They argue that social comparisons between categories, complemented by individuals' needs for a positive group identity (i.e., the self-enhancement motive), are likely to stimulate in-group bias. People tend to emphasize the positive characteristics of their group while derogating those of other groups. Evidence of such bias is particularly strong in competitive situations. Intergroup discrimination and conflict are the potential consequences of these intergroup comparisons.

Intergroup comparison processes form the main thrust of current research and raise issues of motivations, identity, and choice of comparison. For example, results from Rothgerber and Worchel (1997) confirm Tajfel and Turner's (1979) expectations about perceptions of groups resulting from intergroup comparisons. Disadvantaged in-group members harmed and saw as more homogeneous a disadvantaged out-group whose performance was similar to or better than that of the in-group. The effects of group status and performance, however, may be conditioned by the extent to which actors identify with their groups. High identifiers are likely to demonstrate group solidarity when there are threats to group identity in the form of a superior-status out-group comparison (Spears, Doosje, and Ellemers 1997). And Major et al. (1993) show how upward and downward group comparisons influence self-evaluations: People who compared unfavorably with in-group members reported lower self-esteem and more depressed af-

fect than did those who compared unfavorably with out-group members.

Gartrell (1987) also emphasizes the connection between the individual and the group. Rather than relying on motives for a positive social identity, his analysis concentrates on networks—relations among concrete entities—to highlight an often overlooked aspect of comparison choice: the social context in which individuals make comparisons. The examination of a person's social network relations is a way to understand more clearly not only the selection of relevant comparisons but also whether people seek comparisons actively or passively accept those which are readily present in the network. In addition, network analyses may inform how comparison choices affect the network and influence the ties among a person's multiple networks.

#### **SOCIAL COMPARISON AND SOCIAL BEHAVIOR: APPLICATIONS OF THE THEORY**

This review of developments in social comparison theory raises issues at the individual, group, and intergroup levels. Similarly, the extensive applications of the theory cross levels of analysis.

Festinger's (1954) assumption that individuals are driven to evaluate themselves implies a concern with self-knowledge that underlies self-concept and identity formation as well as self-esteem. A number of studies examine the role of social comparisons in shaping individual identity through the life course. For example, Chafel (1988) concludes that during childhood, achievement identities reflect first autonomous self-generated norms, then social comparison norms, and finally an integration of the two. Another study (Young and Ferguson 1979) demonstrates that across all grade levels (grades 5, 7, 9, and 12), parents are the comparison choice for the evaluation of moral issues whereas peers typically serve as a reference point for social issues, especially among older students. Moreover, it appears that social comparisons remain a major source of self-evaluation throughout a person's life (Robinson-Whelen and Kiecolt-Glaser 1997).

Just as the choice of a comparison other provides the basis for self-knowledge, comparisons of opinions within a group affect group dynamics. A

large body of research examines choice shift or group polarization, in which the group voices a more extreme opinion than would be expected on the basis of initial individual opinions (see Myers and Lamm 1976). The social comparison explanation of polarization states that insofar as people want their own opinions to remain distinct from those of others, exposure to others' opinions stimulates shifts in stances to retain that uniqueness. Consequently, the subsequent group opinion is more extreme.

Concerns with justice potentially stimulate social comparisons at all levels. To assess justice, people evaluate how their rewards or outcomes stack up against what they have earned in the past (internal comparison), what another individual like them earns (local or egoistic comparison), what members of their group typically earn (referential comparison), and what their group earns compared to another group (intergroup or fraternalistic comparison). The type of comparison and the specific person or group chosen define whether an individual is likely to perceive himself or herself or the group as being unfairly treated. Justice obtains when outcomes (or the ratio of outcomes to inputs) are equal across the comparison. Equity theory (Adams 1965; "Equity" in this volume) focuses on local comparisons as the basis for individual reactions to an imbalance in outcome-input ratios, while relative deprivation theory (focusing on outcomes only) attempts to explain when an individual or group will feel deprived and opt for collective action to redress that deprivation (see Masters and Smith 1987; Olson et al. 1986).

Social comparisons underlying justice processes have been extended in two ways beyond the traditional concerns that were noted above. First, researchers have begun to explore the role of social comparisons in understanding procedural justice and its relationship to distributive justice. There is, however, contradictory evidence about whether internal self-referents (Van den Bos et al. 1998) or social referents exert greater influence on that relationship (Ambrose et al. 1991). A second emphasis is more applied: To what extent do comparisons contribute to an understanding of perceptions of fair levels of pay? For example, although women earn lower wages than men do, they do not necessarily perceive this inequality as unfair. It appears that women's perceptions stem

from their comparisons with other women rather than with men in general or even with men in their same occupations, resulting in lower pay entitlements (e.g., Moore 1991; Demarais and Curtis 1997).

Although justice concerns have been applied to close relationships, the role of social comparisons in those relationships is a relatively new area of inquiry. VanYperen and Buunk (1994) review research on individuals' decisions to compare themselves to their partners or to their reference groups and the implications of those decisions. For example, people with egalitarian sex-role beliefs are likely to use their partners as a source of comparisons and in doing so enhance their relationship satisfaction; in contrast, individuals with a more traditional sex-role orientation use reference group comparisons to ensure satisfaction. Highlighting the role of motivations in the research reviewed above, Beach et al. (1998) find that reactions to comparisons between individuals in a close relationship support a self-evaluation maintenance model. Comparisons made by intimate couples also appear to parallel those which are evident in intergroup relationships: Subjects have more positive beliefs and fewer negative ones about their own relationships than they do about other relationships (Lange and Rusbult 1995)

A growing body of research examines applications of social comparison theory to health concerns (see Buunk and Gibbons 1997); as in research on close relationships, the application confirms and extends existing theoretical ideas. For example, Buunk and Ybema (1997) introduce an identification-contrast model to explain the choice of comparisons among individuals coping with stress. They suggest that individuals attempt to maintain their self-esteem by identifying upward and by finding ways, such as downward comparisons, to feel that they are doing better than others are. Affleck and Tennen (1991) also stress that downward comparisons are a means for victims of illness to find meaning in their plight and mitigate threats to their self-esteem. These authors also raise important issues involving the role of temporal comparisons in coping with illness and the need to recognize the distinctions between the comparison process and the comparison conclusion. Further extending the domain of social comparison theory, Croyle (1992) attempts to incorporate comparison concerns into stan-

dard models of stress and coping and the self-regulation of illness behavior.

### CONCLUSIONS AND FUTURE DIRECTIONS

The applications attest to the pivotal role of social comparisons in explaining diverse phenomena. They also reiterate a number of theoretical and empirical issues. Future research should examine conditions affecting the motivations that underlie social comparison processes. Such work also would entail the conditions, such as uncertainty, that stimulate comparisons. The establishment of social contexts activating various motivations will allow a clearer assessment of the cognitive processes that influence social comparisons. In turn, this knowledge may enhance understanding of the choice among comparison others and allow an assessment of the extent to which choices vary with nonmotivational factors. Finally, as much of the applied work suggests, additional studies may reveal the range of consequences of social comparisons for individuals, relationships, and groups. Specific theories about identity formation, group decision making, justice, intimate relationships, and health may be "favorites," but their explanatory success depends on everyone's second favorite theory: social comparison.

### REFERENCES

- Adams, J. Stacy 1965 "Inequity in Social Exchange." *Advances in Experimental Social Psychology* 2:267-299.
- Affleck, Glenn, and Howard Tennen 1991 "Social Comparison and Coping with Major Medical Problems." In Jerry Suls and Thomas Ashby Wills, eds., *Social Comparison: Contemporary Theory and Research*. Hillsdale, N.J.: Erlbaum.
- Albert, S. 1977 "Temporal Comparison Theory." *Psychological Review* 84:485-503.
- Alicke, Mark D., Frank M. LoSchiavo, and Jennifer Zerbst 1997 "The Person Who Outperforms Me Is a Genius: Maintaining Perceived Competence in Upward Social Comparison." *Journal of Personality and Social Psychology* 73:781-89.
- Ambrose, Maureen L, Lynn K. Harland, and Carol T. Kulik 1991 "Influence of Social Comparisons on Perceptions of Organizational Fairness." *Journal of Applied Psychology* 76:239-246.
- Arrowood, A. John 1986 "Comments on 'Social Comparison Theory: Psychology from the Lost and Found.'" *Personality and Social Psychology Bulletin* 12:279-281.
- Aspinwall, Lisa G., and Shelley E. Taylor 1993 "Effects of Social Comparison Direction, Threat, and Self-Esteem on Affect, Self Evaluation, and Expected Success." *Journal of Personality and Social Psychology* 64:708-722.
- Beach, Steven R. H., Abraham Tesser, and Frank D. Fincham 1998 "Pleasure and Pain in Doing Well, Together: An Investigation of Performance-Related Affect in Close Relationships." *Journal of Personality and Social Psychology* 74:923-938.
- Beauregard, Keith S., and David Dunning 1998 "Turning Up the Contrast: Self-Enhancement Motives Prompt Egocentric Contrast Effects in Social Judgments." *Journal of Personality and Social Psychology* 74:606-621.
- Blanton, Hart, Bram P. Buunk, Frederick X. Gibbons, and Hans Kuyper 1999 "When Better-Than-Others Compare Upward: Choice of Comparison and Comparative Evaluation as Independent Predictors of Academic Performance." *Journal of Personality and Social Psychology* 76:420-430.
- Buunk, Bram P., and Frederick X. Gibbons 1997 *Health, Coping, and Well-Being: Perspectives from Social Comparison Theory*. Mahwah, N.J.: Erlbaum.
- Buunk, Bram P., and Jan F. Ybema, 1997. "Social Comparisons and Occupational Stress: The Identification-Contrast Model." In Bram P. Buunk and Frederick X. Gibbons, eds., *Health, Coping and Well-Being: Perspective from Social Comparison Theory*, pp. 359-388. Mahwah, NJ: Erlbaum.
- Chafel, Judith A. 1988 "Social Comparisons by Children: Analysis of Research on Sex Differences." *Sex Roles* 18:461-487.
- Collins, Rebecca L. 1996 "For Better or Worse: The Impact of Upward Social Comparison on Self-Evaluations." *Psychological Bulletin* 119:51-69.
- Croyle, Robert T. 1992 "Appraisals of Health Treats: Cognition, Motivation, and Social Comparison." *Cognitive Therapy and Research* 16:165-182.
- Demarais, Serge, and James Curtis 1997 "Gender Differences in Pay Histories and Views on Pay Entitlement among University Students." *Sex Roles* 37:623-642.
- Fazio, Russel H. 1979 "Motives for Social Comparison: The Construction-Validation Destination." *Journal of Personality and Social Psychology* 37:1683-1698.
- Festinger, Leon 1954 "A Theory of Social Comparison." *Human Relations* 14:48-64.
- Gartrell, C. David 1987 "Network Approaches to Social Evaluations." *Annual Review of Sociology* 13:49-66.



- Goethals, George R. 1986 "Social Comparison Theory: Psychology from the Lost and Found." *Personality and Social Psychology Bulletin* 12:261-278.
- , and John M. Darley 1977 "Social Comparison Theory: An Attributional Approach." In J. Suls and R. L. Miller, eds., *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington, D.C.: Hemisphere.
- Gruder, Charles L. 1977 "Choice of Comparison Persons in Evaluations of Oneself." In J. Suls and R. L. Miller, eds., *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington, D.C.: Hemisphere.
- Hakmiller, K. 1966 "Threat as a Determinant of Downward Comparison." *Journal of Experimental Social Psychology* 1(suppl.):32-39.
- Hegeson, Vicki S., and Kristin D. Mickelson 1995 "Motives for Social Comparison." *Personality and Social Psychology Bulletin* 21:1200-1209.
- Hoorens, Vera 1995 "Self-Favoring Biases, Self-Presentation, and the Self-Other Asymmetry in Social Comparison." *Journal of Personality* 63:793-817.
- Hyman, Herbert H., and Eleanor Singer 1968 *Readings in Reference Group Theory and Research*. New York: Free Press.
- Kelley, Harold H. 1973. "The Process of Causal Attribution." *American Psychologist* 28:107-128.
- Klein, William M. 1997 "Objective Standards Are Not Enough: Affective, Self-Evaluative, and Behavioral Responses to Social Comparison Information." *Journal of Personality and Social Psychology* 72:763-774.
- Kulik, James A., and Brooks B. Gump 1997 "Affective Reactions to Social Comparison: The Effects of Relative Performance and Related Attributes Information about Another Person." *Personality and Social Psychology Bulletin* 23:452-468.
- Lange, Paul A. M. van, and Caryl E. Rusbult 1995 "My Relationship Is Better Than—and Not as Bad as—Yours Is: The Perception of Superiority in Close Relationships." *Personality and Social Psychology Bulletin* 21:32-44.
- Major, Brenda, Anne M. Sciacchitano, and Jennifer Crocker 1993 "In-Group Versus Out-Group Comparisons and Self-Esteem." *Personality and Social Psychology Bulletin* 19:711-721.
- , Maria Testa, and Wayne H. Bylsma 1991 Responses to Upward and Downward Social Comparisons: The Impact of Esteem-Relevance and Perceived Control." In Jerry Suls and Thomas Ashby Wills, eds., *Social Comparison: Contemporary Theory and Research*. Hillsdale, N.J.: Erlbaum.
- Masters, John C., and Linda J. Keil 1987 "Generic Comparison Processes in Human Judgment and Behavior." In J. C. Masters and W. P. Smith, eds., *Social Comparison, Social Justice, and Relative Deprivation*. Hillsdale, N.J.: Erlbaum.
- , and William P. Smith, eds. 1987 *Social Comparison, Social Justice, and Relative Deprivation*. Hillsdale, N.J.: Erlbaum.
- Moore, Dahlia 1991 "Entitlement and Justice Evaluations: Who Get More and Why." *Social Psychology Quarterly* 54:208-223.
- Myers, D. G., and David Lamm 1976 "The Group Polarization Phenomena." *Psychological Bulletin* 83:602-627.
- Olson, James M., C. Peter Herman, and Mark P. Zanna, eds. 1986 *Relative Deprivation and Social Comparison*. The Ontario Symposium, vol. 4. Hillsdale, N.J.: Erlbaum.
- Pettigrew, Thomas F. 1967 "Social Evaluation Theory." In D. Levine, ed., *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- Robinson-Whelen, Susan, and Janice Kiecolt-Glaser 1997 "The Importance of Social versus Temporal Appraisals among Older Adults." *Journal of Applied Social Psychology* 27:959-966.
- Rothgerber, Hank, and Stephen Worchel 1997 "The View from Below: Intergroup Relations from the Perspective of the Disadvantaged Group." *Journal of Personality and Social Psychology* 73:1191-1205.
- Schachter, Stanley 1959 *The Psychology of Affiliation*. Stanford, Calif.: Stanford University Press.
- , and Jerome E. Singer 1962 "Cognitive, Social, and Physiological Determinants of Emotional State." *Psychological Review* 69:379-399.
- Spears, Russell, Bertjan Doozeje, and Naomi Ellemers 1997 "Self-Stereotyping in the Face of Threats to Group Status and Distinctiveness: The Role of Group Identification." *Personality and Social Psychology Bulletin* 23:538-553.
- Suls, Jerry M. 1977 "Social Comparison Theory and Research: An Overview from 1954." In J. Suls and R. L. Miller, eds., *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington, D.C.: Hemisphere.
- 1986. "Notes on the Occasion of Social Comparison Theory's Thirtieth Birthday." *Personality and Social Psychology Bulletin* 12:289-296.
- , and Richard L. Miller, eds. 1977 *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington, D.C.: Hemisphere.

- , and Thomas Ashby Wills 1991 *Social Comparison: Contemporary Theory and Research*. Hillsdale, N.J.: Erlbaum.
- Tajfel, Henri, and John C. Turner 1979 "An Integrative Theory of Intergroup Conflict." In W. G. Austin and S. Worchel, eds., *The Social Psychology of Intergroup Relations*. Monterey, Calif.: Brooks/Cole.
- Taylor, Shelley E., and Marci Lobel 1989 "Social Comparison Activity under Threat: Downward Evaluation and Upward Contacts." *Psychological Review* 96:569–575.
- Thorton, D., and A. John Arrowood 1966 "Self-Evaluation, Self-Enhancement, and the Locus of Social Comparison." *Journal of Experimental Social Psychology*. 1(suppl.):40–48.
- Van den Bos, Kees, Henk A. Wilke, Allan E. Lind, and Riel Vermunt 1998 "Evaluating Outcomes by Means of the Fair Process Effect: Evidence for Different Processes in Fairness and Satisfaction Judgments." *Journal of Personality and Social Psychology* 74:1493–1503.
- VanYperen, Nico W., and Bram P. Buunk 1994 "Social Comparison and Social Exchange in Marital Relationships." In Melvin J. Lerner and Gerold Mikula, eds., *Entitlement and the Affectional Bond: Justice in Close Relationships*. New York: Plenum.
- Wayment, Heidi A., and Shelley E. Taylor 1995 "Self-Evaluation Processes: Motives, Information Use, and Self-esteem." *Journal of Personality* 63:729–757.
- Wheeler, L., R. Koestner, and R.E. Driver 1982 "Related Attributes in the Choice of Comparison Other: It's There, but It Isn't All There Is." *Journal of Experimental Social Psychology* 18:489–500.
- Wood, Joanne 1989 "Theory and Research Concerning Social Comparisons of Personal Attributes." *Psychological Bulletin* 106:231–248.
- 1996. "What Is Social Comparison and How Should We Study It?" *Personality and Social Psychology Bulletin* 22:520–537.
- , Maria Giordano-Beech, and Kathryn L. Taylor 1994 "Strategies of Social Comparison among People with Low Self-Esteem: Self-Protection and Self-Enhancement." *Journal of Personality and Social Psychology* 67:713–731.
- Young, James W., and Lucy Rau Ferguson 1979 "Developmental Changes through Adolescence in the Spontaneous Nomination of Reference Groups as a Function of Decision Content." *Journal of Youth and Adolescence* 8:239–252.

KAREN A. HEGTVEDT

## SOCIAL CONTROL

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

The study of social control has been an integral part of sociology since its inception. Originally, the concept was defined as any structure, process, relationship, or act that contributes to the social order. Indeed, to some extent, the study of social order and social control were indistinguishable. This conceptual problem was particularly evident in the early Chicago perspective in which the concepts *social disorganization*, *social control*, and *deviance* were not distinguished. Deviance was thought to be the consequence of lack of social control and was often used to measure the presence of social control. Within the structural functionalism of the late 1940s and 1960s, the study of social control was allocated to the sidelines. It dealt with residual problems of deviance in a social system assumed to be generally integrated and well functioning. By the early 1960s society was, again, assumed to be considerably less orderly and integrated and, again, the concept of social control rose to the forefront. Studies examined both the causes and the consequences of social control. Thus, by the mid-1960s the intellectual ground had been laid for renewed scholarly interest in the study of social control. This chapter reviews the study of social control from that time through the 1970s and 1980s.

A consensus is now emerging that distinguishes social control from the social order it is meant to explain and that distinguishes among social-control processes. One basic distinction is among processes of internal control and external control. The former refers to a process whereby people adhere to social norms because they believe in them, feeling good, self-righteous, and proud when they do and feeling bad, self-critical, and guilty when they do not. This process has recently been termed *socialization*. External control refers to a social process whereby people conform to norms or rules because they are rewarded with status, prestige, money, and freedom when they do and are punished with the loss of them when they do not. This process has sometimes been termed *coercive*, *external*, or just *social control*.

Reflecting contemporary usage, this chapter emphasizes social control as external or coercive control. Research is organized, first, and foremost by whether social control is studied as an independent or dependent variable and, second, by whether it is studied at the micro level (the study of individuals) or the macro level (the study of cities, states, regions, and countries).

### SOCIAL CONTROL AS AN INDEPENDENT VARIABLE

Social-control theories assume that norm violations can frequently be so pleasurable and profitable that many, if not most, people are motivated to violate them. Thus, it is not necessary to study deviant motives; rather it is necessary to study what constrains or controls most people from acting on their deviance motives most of the time. Studies of social control as an independent variable focus on the relative effectiveness of social relationships and arrangements in constraining behavior to social norms and laws. Three general areas have developed. One examines the effectiveness of social ties (bonds, relationships, attachments) to conventional institutions in constraining people from acting on deviant motives. The second examines the effectiveness of macro structures and processes in providing the foundation for these ties. The third examines the effectiveness of the criminal justice system in constraining people from violating the law.

Drawing on a long tradition of work, Hirschi (1969) published an influential formulation of micro social-control theory. He states that the relationship between people and conventional society consists of four bonds: belief, attachment, commitment, and involvement. *Belief* refers to the extent to which conventional norms are internalized (another term for internal control). *Commitment* refers to the extent to which people's social rewards are tied to conformity; the more people have to lose upon being socially identified as norm violators, the lower their likelihood of violating the social norms. *Attachment* refers to people's sensitivity to the opinions of others; the more people are concerned with the respect and status afforded them by others, the more they are subject to social control. *Involvement* refers to the amount of time people spend on conventional activities; the more

people are involved in conventional activities, the less time they have left for deviant activities.

This theory has inspired considerable research on juvenile delinquency. Studies from the 1960s to the 1980s (e.g., Kornhauser 1978; and Matsueda 1982) show that, as attachment to parents and school increase, delinquency decreases. These studies, however, do not show the casual order underlying the relationship between social attachment and delinquency. The theory assumes that low attachment leads to high delinquency; yet high delinquency could very well lead to low attachment. Trying to unravel these causal processes, Liska and Reed (1985) show that low parent attachment leads to high delinquency and that high delinquency leads to low school attachment.

From the 1920s onward, sociologists at the University of Chicago have been interested in the ecological distribution of deviance. Their studies of delinquency, mental illness, and suicide, for example, show that deviance tends to center in cities, particularly in the area where residential and business activity intermesh. They argued that the ecological conditions that disrupt traditional social-control processes are accentuated in these areas, and, when social-control processes weaken, deviance occurs. Industrialization creates a need for the concentration of labor, thereby increasing population size and density through migration and immigration. Both industrialization and urbanization lead to value and norm conflicts, social mobility, cultural change, and weak primary ties. These social conditions, in turn, disrupt internal and external processes of social control. The internal process is weakened because people are unlikely to accept normative standards as right and proper when they experience value and norm conflicts and social change. The external process is weakened because people are unlikely to constrain their behavior to conventional norms when social support for unconventional behavior is readily visible and primary ties to family and conventional friends are weak. In small towns, for example, people may conform even though they may not accept the moral standard because their deviance is easily visible to family and conventional friends.

Perhaps the major problem with this line of research was the failure to measure the disruptive processes directly and the tendency to infer them

from either remote causes such as industrialization and urbanization or more immediate causes such as the social, racial, and class composition of areas. Unable to solve this problem, the theory withered from the 1950s through the 1970s.

During the 1980s a group of young sociologists reexamined the theory to understand the renewed disorder of cities. Directly addressing the problem of measuring the processes that disrupt social control, Sampson and Groves (1989) show how community structural characteristics (such as racial, class, and ethnic composition; residential mobility; and divorce rate) affect crime by weakening ties to conventional institutions. Contrary to the early Chicagoans, Bursik (1986) shows that the ecological distribution of crime is no longer stable over time, that crime rates actually influence community characteristics, and that changes in these structures influence social deviance.

A third body of research examines the effectiveness of the criminal justice system in controlling crime. The underlying theory (deterrence) ignores inner controls and emphasizes punishment as the means of social control, particularly state-administered punishment. It assumes that people are rational and that crime is the result of calculating the costs and benefits of law violations; therefore, it assumes that, the higher the costs of crime, the lower the level of crime. As state-administered punishment is a significant cost of crime, it follows that the higher the level of such punishment, the lower the level of crime.

Two types of deterrence processes have been studied: general and specific. *General* refers to a process by which the punishment of some law violators provides information about the costs of crime to those unpunished (the general public), thereby reducing their law violations. *Specific* refers to a process by which punishment reduces the future law violations of those punished. Research focuses on three dimensions of punishment: severity, certainty, and celerity. *Severity* refers to the harshness of punishment, such as the length of incarceration; *certainty* refers to the probability of punishment, such as the likelihood of being arrested; and *celerity* refers to the swiftness of punishment. In sum, deterrence theory predicts that crime is lowest when punishment is severe, certain, and swift.

The political climate of the 1980s stimulated considerable interest in this theory, leading to hundreds of studies (Cook 1980). Yet, after all this research, it is still difficult to find any firm evidence for either specific or general deterrence. Regarding general deterrence, which has generated the bulk of the research, there is little consistent evidence of a severity effect. There is somewhat more evidence for a certainty effect, although its strength and duration remain unclear. For example, some studies suggest that, as certainty of punishment increases, crime rates decrease but that the decrease does not occur until certainty reaches about 30 percent, which is infrequently reached (Tittle and Rowe 1974). Some studies suggest that the certainty effect only occurs for crimes about which people have the opportunity to think and calculate, like property crimes, but not for violent crimes. And some studies of drunken driving suggest that the certainty effect occurs only if high certainty is well publicized, and even this effect is short lived (Ross 1984).

In sum, assuming that people are generally motivated to deviate, researchers have tried to understand how people are constrained from acting on their motives. Contemporary studies of social control focus on three areas: the interpersonal relationships that constrain people from acting on their motives, the macro structures and processes that provide the social foundation for these relationships, and the criminal justice system as a source of legal constraints.

#### **SOCIAL CONTROL AS A DEPENDENT VARIABLE**

During the 1960s sociologists began to question the assumption of normative consensus and stability and, thus, by implication the viability of the theories built on them. Without clear and stable reference points from which to judge behavior, deviance is difficult to define. Many sociologists came to define it in terms of visible social efforts to control it. Deviance is thus defined as that behavior that society controls, and deviants are defined as those people whom society controls. Research shifts from studying social control as a cause of deviance to studying the causes of social control.

Micro-level studies examine the social processes by which acts and people are defined, labeled, and treated as deviants by family, friends,

the public, and formal agencies of social control such as the criminal justice and mental health systems. Drawing on labeling and conflict theories, many sociologists argue that social control is directed against those who are least able to resist (the disadvantaged and the unfortunate) and that social-control agencies are used by the powerful to control the behavior of others.

During the 1960s, research reported that resources and power (as indicated by class, ethnicity, and race) significantly affect defining, treating, and controlling people as criminals (Black and Reiss 1970), such as arresting, prosecuting, and sentencing them. Unfortunately, these studies do not adjust for the effects of legal considerations, such as seriousness and frequency of offense, which are related to social resources. Without examining the effects of both legal and social resource variables in the same analysis it is difficult to isolate the effects of one from those of the other.

During the 1970s studies addressed this issue. The results are inconsistent, some studies showing race and class effects and some showing no such effects (Cohen and Kleugel 1978).

During the 1980s research tried to resolve these inconsistencies. One group of researchers tried to show that the effect of resources depends on the stage of the criminal justice process (e.g., arrest, prosecution) and the characteristics of the local community. Some stages may be more sensitive to social status and social power than are others, and some communities may be more sensitive to status and power than are others. Dannefer and Schutt (1982) report more racial discrimination at the arrest stage than at other stages, arguing that police have more discretion than do other decision makers, and they report more racial discrimination when the percentage of nonwhites is high, arguing that a high percentage of nonwhites is threatening to authorities.

Macro studies of social control examine the level and the form of social control across such units as cities, states, regions, and countries. They study why one form of control (physical pain) occurs at one time and place and another form (incarceration) occurs at another time and place.

Since the 1970s, conflict theory has provided the major stimulus for this research. It assumes that social control is more likely when the ruling

class or the authorities perceive their interests to be threatened. Threat is thought to be associated with the presence of disruptive acts (crime, civil disorders, social movements) and problematic people (the unemployed, minorities, the urban lower class). The theory assumes that, as disruptive acts and problematic people increase, authorities expand the capacity for social-control bureaucracies and pressure existing bureaucracies to expand the level of control.

Research has focused on the expansion and contraction of three such bureaucracies: the criminal justice system, the mental health system, and the welfare system. It generally suggests that the expansion of the criminal justice system is not necessarily a response to crime, that the expansion of the mental health system is not necessarily a response to mental health, and that the expansion of the welfare system is not necessarily a response to economic need. Rather the expansion and contraction of all three are responses by authorities to the acts and people deemed threatening to their interests.

Studies of the criminal justice system have examined the expansion of the police force in the late 1960s and 1970s, as an indicator of the potential for social control, and the expansion of the prison population in the 1980s, as an indicator of the actual level of control. Liska, Lawrence, and Benson (1981) report that, while the size of the police force is sensitive to the crime rate, it may be even more sensitive to the level of civil disorders, the relative size and segregation of the minority population, and the level of economic inequality. Studies of the prison population and admission rates show that, while these rates, too, are sensitive to the crime rate, they are equally sensitive to the size of problematic or threatening populations such as the unemployed. Studies in England, Canada, and the United States show a substantial relationship between the prison admission rate and the unemployment rate, adjusting for the crime rate (Berk et al. 1981; Inverarity and McCarthy 1988).

Some historical studies (Foucault 1965) assert that mental asylums emerged in the seventeenth century as another social mechanism for controlling the poor urban masses. During the twentieth century the population of mental asylums in the United States continually increased, reaching about

550,000 by the mid-1950s, while the prison population, in comparison, was less than 200,000 at the time. The mental health system seemed to be taking over the role of the criminal justice system in controlling problematic populations. However, since that time the trends for both bureaucracies have reversed. The mental asylum population has decreased from 500,000 to 150,000 and the prison population has increased from 200,000 to 300,000. These trend reversals have stimulated research to examine the extent to which the two bureaucracies are functional alternatives for controlling threatening or problematic populations and acts. Some research (Steadman 1979) studies how various threatening populations that in the past might have been admitted directly into asylums are now first processed in the criminal justice system. Then some of them remain in local jails and others, through various mechanisms such as pleas of *Incompetent to Stand Trial* and *Not Guilty by Reason of Insanity* are channeled into asylums.

Welfare is frequently conceptualized as a form of social control. Piven and Cloward (1971) have stimulated considerable controversy by arguing that the welfare expansion in the United States during the mid- and late 1960s was a response to the urban riots of that period, an attempt to control an economically deprived and threatening population. Various studies provide some support for this thesis. Schram and Turbett (1983) report that the riots affected welfare in two stages. Riots during the mid-1960s prodded the federal government to liberalize welfare policies generally; these policies were then more likely to be implemented in the late 1960s by those states experiencing the most rioting.

In sum, the 1970s and 1980s evidenced a research effort to explain the expansion and contraction of bureaucracies of social control, not so much as responses to crime, mental illness, or economic need, but as responses by authorities to control acts and populations deemed threatening to their interests.

The study of social control has come a long way since its inception at the birth of sociology, at which time it was vaguely defined and not distinguishable from the concept *social order*. Contemporary usage distinguishes the sources of social order from the order itself. The concept *socialization* has come to refer to internal sources of con-

trol, and the concept *social control* has come to refer to external sources of control, the processes whereby people conform to social norms because they are rewarded when they do and punished when they do not. Studying social control as an independent variable, a body of research examines the relative effects of interpersonal relations, social institutions, and formal agencies in constraining social behavior. Studying social control as a dependent variable, another body of research examines how social resources influence social control and how the aggregate amount and form of control varies over time and place.

(SEE ALSO: *Crime Theories of; Criminal Sanctions; Criminology; Deviance Theories; Juvenile Delinquency, Theories of; Law and Society; Sociology of Law*)

## REFERENCES

- Ben-Yehuda, Nachman 1997 "Political Assassination Events as a Cross-Cultural Form of Alternative Justice." *International Journal of Comparative Sociology* 38:1-2.
- Berk, Richard A., David Rauma, Sheldon L. Messinger, and Thomas F. Cooley 1981 "A Test of the Stability of the Punishment Hypothesis: The Case of California, 1851-1970." *American Sociological Review* 46:805-829.
- Black, Donald J., and Albert J. Reiss 1970 "Police Control of Juveniles." *American Sociological Review* 35:63-77.
- Blomberg, Thomas G. and Stanley Cohen (eds.) 1994 *Punishment and Social Control: Essays in Honor of Sheldon Messinger*. New York: Aldine de Gruyter.
- Bridges, George S. and Martha A. Myers (eds.) 1994 *Inequality, Crime and Social Control*. Boulder, Colo.: Westview Press.
- Bursik, Robert Jr., 1986 "Ecological Stability and the Dynamics of Delinquency." In Albert J. Reiss and Michael Tonry, eds., *Communities and Crime*, pp. 35-66. Chicago: University of Chicago Press.
- Cohen, Lawrence E., and James R. Kluegel 1978 "Determinants of Juvenile Court Dispositions." *American Sociological Review* 43:162-177.
- Cook, Philip J. 1980 "Research in Criminal Deterrence: Laying the Groundwork for the Second Decade." In Norval Morris and Michael Tonry, eds., *Crime and Justice: An Annual Review of Research*, pp. 211-268. Chicago: University of Chicago Press.
- Dannefer, Dale, and Russell K. Schutt 1982 "Race and Juvenile Justice Processing in Court and Police Agencies." *American Journal of Sociology* 87:1113-1132.

- Foucault, M. 1965 *Madness and Civilization*. New York: Vintage.
- Guthrie, Douglas J. 1994 "From Cultures of Violence to Social Control: An Analysis of Violent Crime in U.S. Counties with Implications for Social Policy." *Berkeley Journal of Sociology* 39:67-99.
- Hirschi, Travis 1969 *Causes of Delinquency*. Berkeley: University of California Press.
- Horwitz, Allan V. 1990 *The Logic of Social Control*. New Brunswick N.J.: Rutgers University Press.
- Inverarity, James, and Daniel McCarthy 1988 "Punishment and Social Structure Revisited: Unemployment and Imprisonment in the United States, 1948-1984." *Sociological Quarterly* 29:263-279.
- Johnson, W. Wesley 1996 "Transcarceration and Social Control Policy: The 1980s and Beyond." *Crime and Delinquency* 42:114-126.
- Katovich, Michael A. 1996 "Cooperative Bases of Control: Toward an Interactionist Conceptualization." *Social Science Journal* 33:257-271.
- Kornhauser, Ruth 1978 *Social Sources of Delinquency*. Chicago: University of Chicago Press.
- Lacombe, Dany 1996 "Reforming Foucault: A Critique of the Social Control Thesis." *British Journal of Sociology* 47:332-352.
- Liska, Allen E., Joseph J. Lawrence, and Michael Benson 1981 "Perspectives on the Legal Order: The Capacity for Social Control." *American Journal of Sociology* 87:412-426.
- Liska, Allen E., and Mark Reed 1985 "Ties to Conventional Institutions and Delinquency: Estimating Reciprocal Effects." *American Sociological Review* 50:547-560.
- 1997 "Modeling the Relationships between Macro Forms of Social Control." *Annual Review of Sociology* 23:39-61.
- Lyon, David and Elia Zureik (eds.) 1996 *Computers, Surveillance and Privacy*. Minneapolis: University of Minnesota.
- Matsueda, Ross L. 1982 "Testing Control Theory and Differential Association: A Causal Modeling Approach." *American Sociological Review* 47:489-504.
- McConville, Michael and Chester L. Mirsky 1990 "Understanding Defense of the Poor in State Courts: The Sociolegal Context of Nonadversarial Advocacy." *Studies in Law, Politics, and Society* 10:217-242.
- Ross, H. Lawrence 1984 "Social Control through Deterrence: Drinking and Driving Laws." In Ralph H. Turner and James F. Short, Jr., eds., *Annual Review of Sociology*, Palo Alto, Calif.: Annual Reviews.
- Piven, Frances Fox, and Richard A. Cloward 1971 *Regulating the Poor: The Functions of Public Welfare*. New York: Vintage.
- Sampson, Robert J., and W. B. Groves 1989 "Community Structure and Crime: Testing Social Disorganization Theory." *American Journal of Sociology* 94:774-802.
- Schram, Stanford F., and J. Patrick Turbett 1983 "Civil Disorder and the Welfare Expansion: A Two-Step Process." *American Sociological Review* 48:408-414.
- Staples, William G. 1997 *The Culture of Surveillance: Discipline and Social Control in the United States*. New York: St. Martin's Press.
- Steadman, Hank J. 1979 *Beating a Rap*. Chicago: University of Chicago Press.
- Stenson, Kevin and David Cowell 1991 *The Politics of Crime Control*. Newbury Park, Calif.: Sage Publications.
- Tittle, Charles R., and Allan R. Rowe 1974 "Certainty of Arrest and Crime Rates: A Further Test of the Deterrence Hypothesis." *Social Forces* 52:455-562.

ALLEN E. LISKA

## SOCIAL DYNAMICS

The term "social dynamics" is used in a wide variety of contexts that vary in level from the societal to the individual and in approach from qualitative (verbal) to quantitative (mathematical). For example, on the societal level, one can point to Sorokin's ([1937-1941] 1957) qualitative approach in *Social and Cultural Dynamics*. At the other extreme, though also at the global level, there are works such as Forrester's (1971) mathematical, computer-oriented approach in *World Dynamics* and the statistical, empirical approaches found in Ramirez et al.'s (1997) study of the adoption of women's suffrage throughout the world and Frank et al.'s (1997) research on the spread and development of a world environmental regime. On the individual level, examples of qualitative approaches include Hareven's (1982) *Family Time and Industrial Time* and the relevant chapters in Bertaux and Thompson's *Pathways to Social Class* (1997). Also on the individual level, there are mathematical approaches such as White's *Chains of Opportunity* (1970) and statistical, empirical approaches such as Zhou et al.'s (1996, 1997) studies of stratification dynamics in China. Studies that combine qualitative and quantitative approaches are rare. A classic example is Elder's *Children of the Great*

*Depression* (1974). Because of the great diversity in substance and approach, one cannot identify a single line of cumulative research on social dynamics. Instead, there are distinct, loosely related developments that arise in several contexts.

This article has five main sections. The first describes the three main sociological contexts for studies of social dynamics and summarizes their contributions to cumulative sociological research. Since the term “social dynamics” invariably implies a focus on change over time in a social entity, it is closely related to the term “social change.” Some of the key differences between the two terms are discussed in the second section. The third section summarizes reasons for studies of social dynamics in general and the formulation of dynamic models in particular. The fourth section explains the fundamental differences between dynamic models and other types of models. The fifth section reviews the main variations in the types of dynamic models that sociologists have used.

### MAIN CONTEXTS

In one context, social scientists refer to the *dynamics of a phenomenon*, meaning that they focus on how it changes over time. In this traditional usage, the emphasis is primarily on a substantive social phenomenon, and research progress depends on acquiring a deeper theoretical understanding and expanding empirical knowledge about that phenomenon. Topics vary, for example, from “group dynamics” (social interactions among the members of a small group over time) to the “dynamics of development” (change from a traditional rural society to a modern urban industrial society and then to a postindustrial society that belongs to a global system). It is hard to identify substantive commonalities across disparate topic areas except ones of the most abstract sort, for example, that social change is universal but varies in speed. Despite the limited number of substantive generalizations about social dynamics, the study of social dynamics is theoretically and methodologically helpful for reasons that are summarized in the third section.

In a second context that is more typical in recent work, researchers refer to a *dynamic model of a phenomenon*, meaning that their goal is to formulate, test, or explore the consequences of a set of mathematical assumptions or a computer algo-

rithm that is intended to mimic the behavior of the phenomenon of interest. For example, researchers may use a model of population growth and decline in a society; a model of foundings, reorganizations, divestments, mergers, and failures in businesses or other organizations; or a model of the diffusion of an innovation through a population (e.g., the adoption of a new social policy by governments or a new contraceptive by women). Despite the substantive diversity, the formal properties of dynamic models of different phenomena are often similar. This similarity has fostered cumulative progress in studies of social dynamics because a model developed for one topic may be transferable to another topic after only minor modifications of its formal properties. For example, the notion that growth rates are “density-dependent” (depend on population size) arose first in dynamic models of population growth, with the main rationale being that a growing population increases competition among members of the system and depletes environmental resources, eventually leading to a lower rate of population growth. Later this notion was applied to explorations of dynamic models of the formation and survival of unions, businesses, and other kinds of organizations and how those processes depend on the structure of competition (Hannan and Carroll 1992; Carroll and Hannan 2000).

In a third context, authors use a *dynamic analysis of empirical data on a phenomenon*, meaning some form of temporal (longitudinal) analysis of data pertaining to different points in time. Since dynamic analyses are based on dynamic models, work done in the second and third contexts has close parallels. Typically, however, a focus on dynamic models implies a greater emphasis on the model itself, whereas a focus on dynamic analyses indicates a greater stress on the problems of estimating and testing the model as well as the resulting substantive empirical findings. Advance made in methods for the dynamic analysis of one social phenomenon often can be used in dynamic analyses of other phenomena. This also has facilitated cumulative progress in research on social dynamics. For example, Tuma et al.’s (1979) proposal for *dynamic analysis of event histories*, which originally was applied to data on marriage formation and dissolution, subsequently has been applied to dynamic analysis of data on occupational and geographic mobility, organizational mergers and fail-



ures, changes in political regimes, the adoption of governmental policies, and many other social phenomena.

### SOCIAL DYNAMICS VERSUS SOCIAL CHANGE

Although the terms “social dynamics” and “social change” both indicate a focus on change over time, they are used in different circumstances. Social dynamics has a more precise meaning.

First, social dynamics usually presumes change within a social *system*. That system may consist of similar entities (e.g., members of a family, families in a neighborhood, nations in the world) or disparate entities (e.g., different types of actors in a political or economic system) or various attributes of a single social entity (e.g., an individual’s education, occupational prestige, and income or a business firm’s age, size, and structure). The system usually is regarded as bounded, allowing the rest of the world to be ignored for purposes of explanation.

Whether the system consists of actors or variables, the term “system” presumes interdependence and typically involves feedback. Thus, action by one entity in the system leads to counteraction by another entity. For example, managers of a firm may counter a strike by workers by acquiescing to the workers’ demands, outwaiting them, or hiring nonunion laborers. Alternatively, change in one variable in the system leads to an opposing or reinforcing change in one or more other variables. For example, an increase in educational level is followed by an increase in prestige and then an increase in income. Changes resulting from interdependent forces and feedback effects within the system are called *endogenous* changes.

There also may be *exogenous* changes, that is, unexplained (perhaps random) changes that influence change within the system under study but whose causes originate outside that system. For example, in analyses of interaction between a husband and wife, changes in the economy and society in which the couple lives usually are treated as exogenous changes that affect the couple’s behavior, but the societal-level changes themselves are not explained.

Because of interdependent forces and feedback effects as well as possible exogenous changes,

social dynamics typically implies a concern with *complex changes*. Simple linear changes or straightforward extrapolations of previous trends are rarely of primary interest.

Second, social dynamics connotes social changes that have a *regular pattern*. That pattern may be one of growth (e.g., economic expansion, growth of a population), decline (e.g., rural depopulation, the extinction of a cultural trait), cyclical change (e.g., boom and bust in the business cycle), a distinctive but nonetheless recurring transition (e.g., ethnic succession in neighborhoods, societal modernization, the demographic transition from high mortality and fertility to low mortality and fertility), or simply a drift in a particular direction (e.g., the slow but accelerating spread of a social belief or practice through a population).

Third, social dynamics usually implies a degree of *predictability*: Social change not only can be comprehended in terms of post hoc reasons but also can be explicitly modeled. The model, whether it consists of verbal statements or mathematical equations or computer instructions, involves a set of assumptions or propositions that permit fundamental patterns of change to be deduced. In contrast, although a unique historical event may foster social change, its uniqueness makes successful prediction impossible. One challenge in studies of social dynamics is therefore to convert phenomena that are unique on one level to ones that are representative and therefore predictable on another level. Thus, what some regard as a unique historical event, others see as an example of a regular pattern of change. For example, to a historian, the Russian Revolution of 1917 is a unique event, whereas a sociologist may regard it as exemplifying a response to changes in underlying social conditions. Thus, while recognizing many distinctive factors, Skocpol (1979) argues that similar patterns of causes underlie the dramatic political and social transformations that historians call the French, Russian, and Chinese revolutions.

Fourth, the term “social dynamics” is used more commonly than is the term “social change” when regularity in patterns of change is associated with some kind of *equilibrium* (steady state or homeostasis), that is, when feedback effects are such that small deviations from equilibrium lead to compensating effects that cause equilibrium to be restored. For example, in the United States, the

distribution of family income (the share of total income received by different families) was remarkably stable throughout the twentieth century despite tremendous growth in population and economic output and social upheavals such as the civil rights and women's liberation movements. This stability suggests that the process governing the allocation of family income was nearly in equilibrium. The term "social change," especially change seen as part of a unique historical process, usually is associated with change from one distinctive situation to another, very different situation. It implies the antithesis of social equilibrium. The way in which the social status of women and minorities has changed during the twentieth century exemplifies social disequilibrium.

Studies of social dynamics do not necessarily assume the existence of an equilibrium. This point is made clear by studies of the dynamics of economic growth, which often envision a process of never-ending expansion and improvement. Similarly, some dynamic processes imply not a steady-state condition but continual oscillation between conditions. A classic sociological example is Pareto's (1935) analysis of the circulation of elites.

Fifth, the term "social dynamics" is almost always used in situations in which there is an interest in the *process* of change: the step-by-step sequence of causes and effects and the way in which intermediary changes unfold. It is rarely used when only a simple before-after comparison of the condition of the system is the object of interest. Instead, when authors use the term "social dynamics," there is usually a sense that the details and sequencing of changes are important because changes are contingent: If the sequence had been interrupted or altered at an intermediary point, the final outcome might have been different. For example, models of social protest often recognize that the state's response to protests may range from peaceful conciliation to violent suppression. The nature of the state's response is an important contingency because it affects the likelihood, timing, and character of future protests.

Sometimes the sequence of changes occurs on the level of the system as a whole rather than on the level of individual members. For example, in a simple model of population growth, individual-level changes are very elementary: birth followed by death, with the timing of the two being the only

question. On the population level, the addition and loss of individuals over time represent a sequence of changes even though on the individual level there may be few, if any intermediary changes and thus little sense of a sequence of causes and effects.

### REASONS FOR STUDYING SOCIAL DYNAMICS

What motivates sociological interest in social dynamics in general and dynamic models in particular? The most potent reason is the long-standing interest of sociologists in social change, coupled with an increasing recognition that a tremendous amount of scientific leverage can be gained from identifying regularities in patterns of change and then formulating sociological theories that explain them, that is, from studying social dynamics and not just unique historical events. Leverage comes not only from the increased richness of theories of social dynamics but also from greater methodological power in discriminating among competing explanations, as is indicated in more detail below.

As observers of the great social transformations of the nineteenth century, the founders of modern sociology (e.g., Marx, Spencer, and Weber) were keenly interested in social change. However, in the middle of the twentieth century, when structural functionalism and Parsonian thought were dominant, social change was regarded as a minor subfield of sociology. Interest in social change was renewed after a reawakened recognition of social conflict and the concomitant criticism of the assumption of social equilibrium in structural functional theories.

An accelerating pace of global change has added to this interest (Sassen 1988; Boli and Thomas 1997). Rapid growth of the world's population; high levels of mobility of people, goods, and capital between and within nations; the transformation of agricultural societies to industrial and postindustrial societies; social upheavals ranging from strikes, to social protests, to revolutions, to wars; the creation of new organizational forms (e.g., holding companies, multinational corporations, international organizations); fundamental transformations of political regimes, including the failure of communist governments; steady increases in the number and volume of new technological innovations; depletion of natural resources; ex-

tion of plant and animal species; and changes in climate induced by human activities are only a few of the world-level changes that make it virtually impossible for sociologists who study large-scale social systems not to be interested in social change.

Although some scholars view many of these changes as historically unique, the concrete social and economic problems that result from them motivate attempts to find regular patterns and predict future changes, in short, to develop dynamic models of societal and global changes. Consider the massive changes in eastern European nations that began in the late 1980s. From the late 1940s to the late 1980s, those nations were governed by totalitarian polities and had command-type socialist economies. Now most of them appear to be headed in the direction of market-type capitalist economies and democratic polities. The intellectual challenge, as well as a major problem for policymakers, is to develop a theory of the transition from one to the other, that is, a theory of the dynamics of the social change that is expected to occur. The fact that no satisfactory theory existed when the transition began was apparent to the general public as well as to social scientists. It points to the practical as well as scholarly value of studying social dynamics.

Sociologists who study micro-level phenomena (individuals and families) also cannot ignore social change. The life course of individuals in modern societies has a typical sequence of activities associated with aging (e.g., birth, day care, school, work, marriage, child rearing, retirement, death) that commands considerable attention by sociologists. Historical changes in family patterns (e.g., increases in premarital cohabitation, delays in marriage, changes in husband–wife roles, increases in divorce, baby booms and baby busts, increasing institutionalization of the elderly) also put social change at the forefront of the attention of sociologists who study the family. These subjects are perhaps more easily viewed in terms of social dynamics than are ones pertaining to global and societal changes because similar patterns across individuals and families are more readily apparent.

Sociologists who study behavior in small groups were among the earliest to express an interest in social dynamics. This interest received a major boost from Bales's *Interaction Process Analysis* (1950).

Game theorists, who attempt to explain the moves and countermoves of actors in highly structured situations, also exemplify a concern with social dynamics in small groups, though they, much more than Bales and his intellectual descendants, concentrate on formal models and deemphasize hypothesis testing and empirical results (see Shubik 1982).

There are also metatheoretical and methodological reasons for studying social dynamics even when the primary intellectual concern is with statics, that is, with relationships among actors or variables at a single point in time.

First, studies of relationships at a single point in time implicitly assume a steady state or equilibrium. Otherwise relationships at a given time point must be transitory and in the process of changing, a situation that would degrade their potential contribution to enduring sociological knowledge. A steady state may or may not exist. If it does not exist, one needs to study social dynamics to understand relationships at a point in time. If a steady state does exist, much can be learned by studying social dynamics that cannot be learned easily by studying relationships at a single point in time. For one thing, two theories may imply the same relationship among variables at a given point in time but imply different time paths of change. In that case, a study of social dynamics can differentiate between them, whereas a study of the steady state cannot. For another thing, a theory of relationships at a point in time is invariably the special case of one or more theories of change over time, and the latter theories almost always have a richer set of implications than do the former. This means that in general there are more ways to test theories of social dynamics than to test theories of social statics.

### THE NATURE OF DYNAMIC MODELS

As was noted earlier, developments in dynamic models (and derivative developments in methods of dynamic analysis) are the major commonality in sociological studies of social dynamics. To understand the main features of dynamic models, it is important to differentiate them from other types of models.

The most basic distinction is between *static* and *dynamic* models. Static models describe rela-

tionships among social actors in a system or among the attributes of a social entity at a given point in time. As was noted earlier, they implicitly assume a steady state or equilibrium, a phenomenon that is about as common in nature as a vacuum. In contrast, dynamic models describe the process or sequence of changes among actors in a social system or among the attributes of a social system.

Dynamic models also can be contrasted with *comparative static models*, which are especially common in economic analyses and analyses of social experiments. Although both deal with change over time, they differ in an important way. The process of change leading from conditions at the earlier time point to conditions at the later time point is fundamental to a dynamic model. In contrast, the change process is ignored in a comparative static model, which resembles a black box that relates conditions at one point in time to conditions at a later time.

To illustrate this distinction, consider alternative ways of explaining a son's occupational prestige. In a comparative static model, the son's prestige may be related to his father's socioeconomic status and his own education without any attention being paid to the mechanisms and processes that lead from those background conditions to the son's condition as an adult. In a dynamic model, the father's socioeconomic status and the son's education may be seen as giving access to certain entry-level jobs, which in turn provide opportunities for further career mobility, leading to jobs with varying levels of prestige. In a dynamic model, the timing and sequence of job shifts are of concern, not just the son's initial condition (i.e., his father's social status and his own education). In summary, dynamic models are used to explain not only why the later condition of a phenomenon differs from its earlier condition but also how a sequence of changes leads from one condition to the other.

### TYPES OF DYNAMIC MODELS

Different types of dynamic models are distinguished on the basis of a variety of formal properties. One basic distinction seems to be whether the components of the system are social actors or the attributes of a social entity. In the former case, dynamic models of the behavior of social actors are developed: Actor A does X, in response actor B

does Y, then actor A does Z, and so on. Although much of game theory is not concerned with dynamic models, some of it formulates precisely these kinds of models. In the latter kind of dynamic model, values of variables describing the social entity are related to one another. Ecological theories of organizational survival utilize these models, for example, relating the degree of environmental variability and the degree of specialization of various types of social organizations in the environment to the survival of these types (Hannan and Freeman 1987). The distinction between systems of actors and systems of variables is not as important as it may seem at first because the behaviors of actors usually can be translated into variables.

A more basic and important distinction is whether time is discrete or continuous. Most empirical phenomena can change at any moment, and this leads one to expect that time should be treated as continuous in most dynamic models. In fact, time more often is treated as discrete for two main reasons. First, the empirical data used to test a dynamic model usually measure time at only a few discrete points. Some researchers then find it convenient to build a dynamic model of the data rather than model the underlying social process. Second, some researchers consider discrete-time models to be simpler and believe that little information is lost from approximating truly continuous-time processes with discrete-time models. If the discrete time points in the data are sufficiently numerous and close together, the approximation is almost always satisfactory. If they are not, important intermediary steps in the process are likely to be ignored, possibly resulting in misleading conclusions. Whether discrete-time models are simpler than continuous-time models is less clear. To some extent, it is a matter of a researcher's taste and training.

Another key distinction concerns whether the variables that describe the social system are discrete, metric, or a mixture of the two. Discrete variables have a finite set of values; for example, political regimes may be categorized into a small number of basic types. Metric variables have a continuum of values; for example, a person's income and occupational prestige usually are treated as continuous variables. In fact, in both instances, the number of values is finite but is so large that treating the values as continuous may be conve-

nient and is often fairly realistic. Age is a continuous variable, but measurements of it are always discrete (e.g., to the nearest year, month, or day).

A distinction also may be made in the way in which variables in a system change over time. By their nature, discrete variables can change only in jumps. For example, there may be a sudden change from a military political regime to a multiparty government. Metric variables often are regarded as changing gradually. For example, a firm's profits may be treated as shifting upward or downward by small increments. In fact, metric variables also may change in jumps. For example, income may fall from a high value to nearly zero when a family's main breadwinners lose their jobs.

Another important distinction is whether the change process is treated as deterministic or stochastic (having a random component). There is a broad consensus that stochastic models are almost always more realistic. Few social changes occur in a strictly determined fashion, and those which do change deterministically are rarely sociologically interesting. Nevertheless, deterministic models can be useful when the solution of realistic stochastic models presents severe technical problems. Those formidable problems tend to occur when there is a high degree of interdependence in the social system (e.g., in models of the diffusion of an innovation) and when both time and outcomes are treated as continuous.

Whether time is treated as discrete or continuous, models of changes in discrete variables are invariably stochastic because changes by jumps almost dictate reference to probabilities. By contrast, continuous-time models of changes in metric variables are typically deterministic.

Some progress has been made in developing empirically estimable stochastic models of heterogeneity in the spread of a social practice in the presence of interdependent influences in a social system (Strang and Tuma 1993). Davis and Greve (1997) provide an intriguing application of this modeling approach to the use of "poison pills" and "golden parachutes" among firms.

Formal dynamic models are of two main types. In one type, the model consists of a set of mathematical equations that relate some elements of the system to other elements. In the other type, the model consists of a set of computer instructions

that relate inputs of various variables and/or actors at one time to outputs at a later time. The computer instructions in fact represent mathematical equations that are so complex that they cannot be solved in practice without the aid of a computer. Still, it is convenient to think of computer models as very complicated mathematical models.

A clear introduction to both discrete-time and continuous-time deterministic models of metric variables can be found in Baumol's *Economic Dynamics* (1951), which also introduces several economic theories of potential interest to sociologists, including theories of wages and profits in firms and economic growth. For a discussion of deterministic models of change in metric variables, see Doreian and Hummon (1976).

Two of the early classic discussions of stochastic models of change in discrete variables are Coleman's *Introduction to Mathematical Sociology* (1964) and Bartholomew's *Stochastic Models for Social Processes* (1973). Tuma and Hannan's *Social Dynamics: Models and Methods* (1984) discusses both deterministic and stochastic models of change in metric variables in continuous time as well as continuous-time stochastic models of change in discrete variables. This work also discusses metatheoretical and methodological reasons for studying social dynamics, applies dynamic models to a variety of different sociological problems, and provides an extensive bibliography pertaining to models and methods used in studying social dynamics. It also contains a comprehensive introduction to event history analysis.

## CONCLUSION

Sociologists, whether studying whole societies or small groups, have had a long-standing and far-reaching interest in social change. Traditional approaches focused on specific substantive phenomena and, especially in macro-level studies, often stressed unique historical occurrences rather than common dimensions underlying patterns of change. Recent studies of social dynamics usually focus on what is regular and predictable about social change and the social mechanisms that generate a sequence of contingent changes. Often they embed ideas about change in dynamic models and test them in dynamic analyses of over-time data. This

approach has been especially valuable in fostering cumulative research.

(SEE ALSO: *Cohort Perspectives; Diffusion Theories; Game Theory and Strategic Interaction; Life Course; Life Histories and Narrative; Longitudinal Research; Paradigms and Models; Social Change; Social Forecasting*)

## REFERENCES

- Bales, Robert F. 1950 *Interaction Process Analysis: A Method for the Study of Small Groups*. Cambridge, Mass.: Addison-Wesley.
- Bartholomew, David J. 1973 *Stochastic Models for Social Processes*. New York: Wiley.
- Baumol, William J. 1951 *Economic Dynamics: An Introduction*. New York: Macmillan.
- Bertaux, Daniel, and Paul Thompson eds. 1997 *Pathways to Social Class: A Qualitative Approach to Social Mobility*. Oxford, UK: Clarendon Press.
- Boli, John, and George M. Thomas 1997 "World Culture in the World Polity: A Century of International Non-Governmental Organization." *American Sociological Review* 62:171–190.
- Carroll, Glen R., and Michael T. Hannan 2000 *The Demography of Corporations and Industries*. Princeton, N.J.: Princeton University Press.
- Coleman, James S. 1964 *Introduction to Mathematical Sociology*. New York: Free Press.
- Davis, Gerald F., and Henrich R. Greve 1997 "Corporate Elite Networks and Governance Changes in the 1980s." *American Journal of Sociology* 103:1–37.
- Doreian, Patrick, and Norman P. Hummon 1976 *Modeling Social Processes*. Amsterdam: Elsevier.
- Elder, Glen H. 1974 *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press.
- Forrester, Jay W. 1971 *World Dynamics*. Cambridge, Mass.: Wright-Allen.
- Frank, David J., Ann Hironaka, John W. Meyer, Evan Schofer, and Nancy Brandon Tuma 1997 "The Structuring of a World Environmental Regime, 1870–1990." *International Organization* 51:623–651.
- Hannan, Michael T., and Glenn R. Carroll 1992 *Dynamics of Organizational Populations*. New York: Oxford University Press.
- , and John Freeman 1989 *Organizational Ecology*. Cambridge, Mass.: Harvard University Press.
- Hareven, Tamara 1982 *Family Time and Industrial Time: The Relationship between the Family and Work in a New England Industrial Community*. Cambridge, UK: Cambridge University Press.
- Pareto, Vilfredo 1935 *The Mind and Society*, vol. III, Arthur Livingston, ed.; Andrew Bongiorno and Arthur Livingston, trans. New York: Harcourt, Brace.
- Ramirez, Francisco O., Yasemin Soysal, and Suzanne Shanahan 1997 "The Changing Logic of Political Citizenship: Cross-National Acquisition of Women's Suffrage Rights, 1890 to 1990." *American Sociological Review* 62(5):735–745.
- Sassen, Saskia 1988 *The Mobility of Labor and Capital: A Study in International Investment and Labor Flow*. Cambridge, UK: Cambridge University Press.
- Shubik, Martin 1982 *Game Theory in the Social Sciences: Concepts and Solutions*. Cambridge, Mass.: MIT Press.
- Skocpol, Theda 1979 *States and Social Revolutions in France, Russia, and China*. Cambridge, UK: Cambridge University Press.
- Sorokin, Pitirim A. (1937–1941) 1957 *Social and Cultural Dynamics: A Study of Change in Major Systems of Art, Truth, Ethics, Law and Social Relationships*, revised and abridged version. Boston: Porter Sargent.
- Strang, David G., and Nancy Brandon Tuma 1993 Spatial and Temporal Heterogeneity in Diffusion." *American Journal of Sociology* 99:614–639.
- Tuma, Nancy Brandon, and Michael T. Hannan 1984 *Social Dynamics: Models and Methods*. Orlando, Fla.: Academic Press.
- , ———, and Lyle P. Groeneveld 1979 "Dynamic Analysis of Event Histories." *American Journal of Sociology* 84:820–854.
- White, Harrison C. 1970 *Chains of Opportunity: Systems Models of Mobility in Organizations*. Cambridge, Mass.: Harvard University Press.
- Zhou, Xueguang, Nancy Brandon Tuma, and Phyllis Moen 1996 "Stratification Dynamics under State Socialism: The Case of Urban China, 1949–1993." *Social Forces* 74:759–796.
- 1997 "Institutional Change and Job Shift Patterns in Urban China, 1949–1994." *American Sociological Review* 62 339–365.

NANCY BRANDON TUMA

## SOCIAL EXCHANGE THEORY

Social exchange theory is a major theoretical perspective in sociology. Within this framework, social behavior is viewed primarily in terms of the pursuit of rewards and the avoidance of punish-

ment and other forms of cost. Individuals engage in interaction to meet their needs. The basic unit of analysis is the relationship between actors. Thus, exchange theorists view social relations and the social structures generated by the ties that bind people in different forms of association as the central object of sociological inquiry. Major topics of study within this tradition of research include the nature and effects of the interconnections among actors and the distribution of power within exchange structures. Power and status relations among actors in different types of social structures are considered key forces in determining the nature of structural change over time. The major exchange theorists all have treated power, structural sources of power, and the dynamics of power use as primary in their theoretical formulations.

Social exchange theory derives from several distinct lines of theoretical work in the social sciences, including social behaviorism, utilitarianism, and functionalism (Turner 1986). Major proponents of the social exchange perspective within sociology include Homans (1961, 1974), Blau (1964, 1987), and Emerson (1962, 1972a, 1972b). Within psychology, the work of Thibaut and Kelley (1959; Kelley and Thibaut 1978) bears a strong resemblance to social exchange theory in its emphasis on the interdependence of actors and the social implications of different forms of interdependence. Anthropologists such as Malinowski (1922), Mauss (1925), Schneider (1974), and Levi-Strauss (1949, 1969) have contributed in different ways to the emergence of this theoretical perspective (see Ekeh 1974). In addition, the foundation of microeconomics has much in common with some variants of social exchange theory (Heath 1976). This affinity is clearest in Blau's *Exchange and Power in Social Life* (1964) and in subsequent theoretical developments (e.g., Cook and Emerson 1978; Coleman 1972, 1990). The breadth of the intellectual heritage of social exchange theory accounts in part for its continued significance in the social sciences.

Homans's well-known essay "Social Behavior as Exchange" (1958) clarified the nature of this theoretical orientation and introduced it into mainstream sociology. An elaboration of Homans's perspective was published in *Social Behavior: Its Elementary Forms* (revised in 1974). An important distinguishing feature of Homans's work was its reliance on the language and propositions of behavioral psychology. The use of operant psychol-

ogy as the behavioral basis of exchange theory created much of the early controversy surrounding the utility of this perspective for sociologists. In particular, the corresponding claim made by Homans that laws of social behavior could be "reduced to" the basic underlying principles of psychological behaviorism generated much debate (e.g., Deutsch 1964). According to Homans, "The general propositions we shall use in explanation are psychological in two senses: they refer to the actions of individuals and they have . . . been formulated and tested by psychologists" (1974, p. 12). However, Homans explicitly took as the major theoretical task the explanation of social phenomena. This emphasis on social behavior and the social structures generated and altered by human social interaction has sustained the influence of social exchange theory in sociology. In this regard, Homans viewed the line drawn between psychology and sociology as fundamentally arbitrary.

The initial theoretical formulation developed by Homans (1961) and revised in 1974 included five main propositions, all of which have to do with the fact that behavior is a function of its payoffs: the consequent rewards and punishments. The first proposition is the "success proposition," which states that the more frequently an activity is rewarded, the greater is the likelihood of its performance. Behavior that generates positive consequences for the individual is likely to be repeated. The second proposition, the "stimulus proposition," stipulates that similar environmental or situational circumstances will stimulate behavior that has been rewarded on similar occasions in the past. This allows for the generalization of behavioral responses to "new" situations. The "value proposition" specifies that the more valuable the result of an action is to the actor, the more likely that action is to be performed. This proposition is qualified by the "deprivation-satiation" proposition, which introduces the general ideal of diminishing marginal utility. According to this proposition, the more often a person has recently received a particular reward for an action, the less valuable is an additional unit of that reward. Thus, some rewards become less effective over time in eliciting specific actions, though this is less true for generalized rewards such as money and affection and for anything for which satiation is less likely to occur except at extreme levels. The fifth theoretical proposition in Homans's basic framework specifies the

conditions under which persons react emotionally to different reward situations. This proposition has two parts. People who do not receive what they anticipate are expected to become angry and behave aggressively, based on the original Miller and Dollard (1941) "frustration-aggression" hypothesis (see Homans 1974, p. 37). People who receive more than they expect or do not receive anticipated punishments will be happy and will behave approvingly. This system of propositions forms the original core set of ideas of what has come to be called social exchange theory.

Homans's (1961, 1974) uses this set of theoretical ideas to explain phenomena such as the exercise of power and authority, cooperation, conformity and competition, structures of sentiment and interaction, status and influence, satisfaction and productivity, leadership, distributive justice, and the emergence of stratification. He addressed these social phenomena primarily in terms of the nature of the interpersonal relations involved. Furthermore, he emphasized "elementary" forms of behavior, or what he referred to as the "subinstitutional" level of analysis. "We gain our fullest understanding of the elementary features of social behavior by observing the interactions between members of small, informal groups," Homans (1974, p. 356) argued. By studying such forms of behavior, he hoped to illuminate the elementary, informal subinstitutional bases of more complex forms of social behavior that often are more formal and institutionalized. What he bequeathed to modern-day sociology, besides his particular form of theorizing, was an emphasis on the microfoundations of social structures and social change.

Whereas Homans focused on elementary forms of behavior and the subinstitutional level of analysis, Blau (1964, 1987) moved beyond the micro level to the institutional level, dealing with authority and power, conflict, and change in the context of institutionalized systems of exchange. In disagreement with Homans's reductionistic strategy, Blau (1987, p. ix) claims that his "theory is rooted in the peculiarly social nature of exchange, which implies that it cannot be reduced to or derived from psychological principles that govern the motives of individuals, as Homans aims to do." In distinct contrast to Homans's reductionism, Blau assumed that social structures has "emergent" properties that cannot be explained by character-

istics or processes that involve only the subunits. Thus, Blau parted company from Homans in two major ways. First, Blau's framework was not based on principles of behavioral psychology; instead, he introduced microeconomic reasoning into the analysis of distinctly social exchange. Second, he explicitly introduced the notion of emergent processes into his theoretical treatise, not only rejecting reductionism but also expanding the theory to extend far beyond its original subinstitutional base.

Blau (1964) developed a general framework for analyzing macro structures and processes based on an extension of his micro-level theory of social exchange processes. Drawing on Simmel's understanding of social life, he explains the general structure of social associations rooted in psychological processes, such as attraction, approval, reciprocation, and rational conduct. Group formation, cohesion, and social integration as well as processes of opposition, conflict, and dissolution are explained in terms of social exchange processes. These forms of social association generated by exchange processes come to constitute very complex social structures (and substructures) over time. These more complex social structures are then examined by Blau as they are created and changed by power processes and the dynamics of legitimation and political opposition. Common values mediate and make possible indirect exchanges and thus the coordination of action in large collectivities. According to Blau, they also "legitimate the social order." Throughout this major work, Blau contrasts and compares social exchange processes in simple structures with those in more complex social structures and institutions. The major social forces he analyzes include differentiation, integration, organization, and opposition that sets up the dialectic necessary for the explanation of structural change.

The strategy of building a theory of macro structure and processes on an explicitly micro-level theory was a distinguishing feature of Blau's (1964) original work, which also became the focus of a major stream of theoretical work in sociology on the "micro-macro link" in the 1980s and 1990s. Ironically, Blau (1986) himself challenged the utility of his approach in his subsequent writings (Blau 1987), fueling the debate further. In his introduction to the second printing of his book on exchange and power (1986), he argues that microsociological and macrosociological theories



“require different approaches and conceptual schemes though their distinct perspectives enrich each other” (1986, p. xv). This theoretical debate will not be over soon since it lies at the heart of the nature of sociological analysis and relates to broad issues of the primacy of particular units and levels of analysis as well as to complex metatheoretical and methodological issues.

Blau (1964) and subsequently Emerson (1962, 1972, 1972b) made power the central focus of analysis. Blau treated power, authority, opposition, and legitimation as key topics in his discussion of macro structures and the dynamics of structural change. Emerson’s (1962) theory of power-dependence relations was partially incorporated into Blau’s (1964) treatment of power imbalance and the conditions of social independence. For Emerson (1962), these strategies were power-balancing mechanisms. The central proposition in Emerson’s (1962) article classic was that power, defined in relational terms, is a function of the dependence of one actor on another. In a two-party exchange relation, the power of one party (A) over another party (B) is a function of the dependence of B on A. Dependence is a function of the value one actor places on the resources (or valued behavior) mediated by the other actor and the availability of those resources from alternative sources. The greater the availability of these resources from other actors (or alternative sources), the lower one actor’s dependence on another. Two features of this approach to power are important: (1) It treats power as relational (a feature of a social relation, not simply a property of an actor), and (2) it treats power as potential power; that is, it may or may not be exercised. This relational conception of power became the basis for most subsequent work on exchange and power.

Emerson (1972a, 1972b) expanded his treatment of power and dependence to form a more extensive exchange theory of social relations. In many ways, his work combined the approaches of Homans (1961) and Blau (1964). In the original formulation, Emerson (1972a) adopted the language and principles of behavioral psychology to form a theory of social relations. However, he quickly moved beyond behavioral principles to the formation of more complex propositions regarding the emergence of various kinds of social structures. Here the theory picks up the Simmelian focus of Blau’s work as well as the concern with

emergent properties and complex social structures. Emerson (1972b), like Blau (1964, 1986), viewed the major task of exchange theory as the creation of a framework in which the primary dependent variable is social structure and structural change. The major task was eminently sociological, not psychological, even though all three theorists explicitly incorporated into their thinking notions about the psychology of actors. Emerson and Cook’s subsequent work (e.g., Cook and Emerson 1978) adopted a more cognitive perspective on the actors involved in social exchange. Molm’s (1981, 1987) earlier work extended the original behavioral underpinnings of the theory.

Exchange theory, though originally dyadic in focus, has been extended to apply to the analysis of exchange networks. Both Homans and Blau recognized the ubiquity of social networks and different forms of social association, but Emerson (1972b) made networks and corporate groups a central focus of his theoretical formulation. The definition of exchange relations as being “connected” in various ways to form network structures was the key to this development in the theory. Emerson defined two major types of connections between exchange relations: negative connections and positive connections. Two relations are negatively connected if the magnitude or frequency of exchange in one is negatively correlated with the magnitude or frequency of exchange in the other. In essence, the two relations are strictly alternatives. If a supplier gets parts in an exchange with one vendor, he or she does not need to get the same parts from another vendor. Negatively connected relations are thus competitive in nature. In contrast, when two relations are positively connected, exchange in one relation enhances exchange in the other. For example, the resources one party gets in exchange with one supplier can be used to obtain needed goods from another supplier. In this case, a positive connection exists and the two exchange relations are positively correlated. Such exchange relations are more cooperative than competitive in nature and form the basis for some types of division of labor and specialization within exchange networks. Subsequent theorists such as Willer (1987), Markovsky et al. (1988), Bonacich (1986), and Yamaguchi (1996) have developed other ways of classifying types of exchange connections. Some of this work is discussed below in the discussion of alternative perspectives.

A key concept in Emerson's exchange theory of power is the idea that exchange relations can be balanced or imbalanced. A power inequality results from an imbalance in power relations between two or more actors. An exchange relation is balanced if both parties are equally dependent on each other for exchange (or resources of value). If they are equally dependent, they have equal power. The central idea that power is based on dependence allows for the specification of ways in which dependencies are altered so that they affect the balance of power in the exchange relation and in networks of exchange relations.

Emerson postulated four power-balancing mechanisms to explain some of the ways in which exchange relations and the networks they form change either to maintain and preserve existing structural arrangements and distributions of power or to alter them. Coalition formation is one of the mechanisms by which power-disadvantaged actors in less powerful network positions can gain power through the collective advantage gained through cooperative action. Not all coalitions are power-balancing, however. In subsequent work, Emerson addressed the kinds of coalitions that form between powerful actors (sometimes referred to as collusion) or between powerful actors and a subset of the less powerful actors (a divide-and-conquer strategy).

Division of labor, or specialization within a network, may operate as a power-balancing mechanism, since it can result in changes in the distribution of power in a network through modifications in the distribution of resources and the nature of the structural arrangements. For example, two suppliers of the same resource who have been competitors may decide to specialize and offer different services in a way that makes them no longer competitive with each other in a particular network. Network extension also can alter the balance of power in a network as new exchange partners become available. In addition, when other strategies are not available, actors can devalue what they obtain from a more powerful actor as a way to reduce their dependence on the relationship. This strategy may be a precursor to an exit from the relation in many instances. Various theorists have continued this line of work, specifying the principles that predict the distribution of power in different exchange structures and the

processes that modify it (e.g., Cook et al. 1983, 1986; Bonacich 1986; Yamaguchi 1996).

Other extensions of the exchange theory originally developed by Emerson have focused on the links between structure and process and on other bases of power. In a major research program that extended over a ten-year period, Molm (1997) investigated the role of coercive power in social exchange. Emerson's work and that of most of the exchange theorists had focused almost exclusively on reward power, or the control over positively valued goods and services. Coercive power is the ability to control negative events (e.g., to withhold rewards) or to inflict punishment on another in an exchange relation. Unlike reward power, coercive power is used less often in exchange relations, especially by those in power-advantaged positions, who seem to understand that it may be viewed as unjustified in many circumstances. The fear of retaliation is also a deterrent to the use of coercive power. The use of coercive power is more costly since it imposes losses on the exchange partner in addition to the opportunity costs involved. Molm's (1989, 1997) major accomplishment was to expand the treatment of power in the classic power-dependence formulation to include forms of coercion. Since exchange relations often involve control over both things people value and things people wish to avoid, this is a significant extension of the theory.

Alternative theoretical formulations have been developed for investigating power processes in exchange networks. They include the "elementary theory" developed by Willer and his collaborators (e.g., Willer and Anderson 1981; Markovsky et al. 1988), Friedkin's (1992) "expected value model" of social exchange, and game theory, which has been applied to the analysis of exchange networks by Bienenstock and Bonacich (1992). While some of these formulations have an affinity with the original power-dependence framework developed by Emerson (1972a, 1972b), most have explored other bases of power. For example, Willer and his collaborators have developed a different terminology for specifying the nature of the relations in an exchange network. They define three types of relations: null (no connection), inclusion (when someone has to be involved in an exchange for it to take place) and exclusion (when someone may be involved in an exchange but is in competition with others and thus may be excluded from the

exchange at any time). These theorists go on to develop different principles for the distribution of power in networks characterized by different types of relations. Exclusion is viewed as the main determinant of power. The ability to exclude others from exchange is thus the key source of power in this theory.

Bienenstock and Bonacich (1992, 1997) analyze exchange networks by using a game theory perspective. They attempt to understand the efforts of actors to maximize certain well-defined interests by adopting strategies that can be analyzed usefully with the tools of game theory. Based on different solution concepts (e.g., the core, the kernel), they make predictions concerning the outcomes of the exchanges in various types of network structures. In addition, this application of game theory provides predictions about the role of information in exchange processes. Building on the early contributions of Blau, Coleman, Emerson, and Cook, Yamaguchi (1996) works out a rational choice model for predicting the distribution of power and the effects of network centrality in what he terms substitutable and complementary exchange relations.

Further developments in the theory of exchange include the formulation of explicit propositions concerning the use of power in different types of exchange network structures and the specification of some of the determinants of power use. These factors include concern over the fairness of the distribution of outcomes, the commitments that emerge between actors (e.g., Lawler and Yoon 1996), the formation of coalitions, particular strategies of action, and whether the power is reward power or punishment power. More recent developments focus more on methodologies for specifying the distribution of power in complex network structures (see, for example, Markovsky's work). Interest in this topic is in part driven by the potential for synthesizing exchange theoretic conceptions of power with network models of social structure (see Cook 1987; Cook and Whitmeyer 1992). Another arena of current theoretical and empirical work is the specification of dynamic models of power use and structural change that include a more sophisticated model of the actors involved and the strategies they adopt in their attempts to obtain resources and services that are of value to them. These general theoretic

and empirical efforts will be important if exchange theory is to fulfill its promise of providing an approach to linking micro-level theories of action and interaction with macro-level explanations of structure and processes of social change, an agenda that was originally set by Homans, Blau, and Emerson.

The application of exchange theory to understanding a variety of social phenomena has grown over the last two decades. Early applications focused on the explanation of the initiation and termination of social relations in work settings and families and then in the domain of romantic relationships and dating. Topics of interest to researchers included the conception of fairness in social exchange relations and its link to relational satisfaction and dissolution, the use of power in social relations based on control of both rewards and costs, and the abuse of power as well as the role of coalitions in altering the balance of power among actors in a network of individuals or organizations. Beyond the application to family and work settings, exchange theory has been applied in many different contexts to the study of organizations and interorganizational relations. Since organizations typically require resources from other entities much of their time is devoted to the strategic management of those dependencies. The resource dependence perspective (Pfeffer and Salancik 1978) in the field of organizations represents a straightforward application of exchange reasoning to the strategic actions of organizations and their subunits (e.g., at the divisional level). The developing field of economic sociology is now drawing to some extent on ideas derived from exchange theory to explain the emergence of network forms of organization and the nature of the power processes that emerge in those networks. Network effects on labor practices, informal influence among organizations, the organization of business groups, and the formation of international linkages that cross traditional national boundaries of economic and productive activity are central topics of inquiry in economic sociology. Some of these efforts involve understanding the effects of network location on outcomes and the various strategies actors use to enhance their bargaining power and influence. These efforts derive in part from the power-dependence reasoning first introduced by Emerson and Blau into exchange theorizing.

Other applications of exchange theory include broader efforts to investigate the balance of power in the health care industry, the strategic role of insurance companies in an era of managed care, and the response of physicians to the loss of power and autonomy. Several researchers have attempted to analyze the nature of physician referrals in network exchange terms and to characterize the nature of physician-patient interaction as an exchange relation in which power is asymmetrical (or imbalanced) and trust plays a key role in "balancing" that power differential. The patient must place his or her fate in the hands of a more competent, more informed actor and trust that the physician will do no harm. Future applications of the exchange model of interaction and of network exchange in other domains will help clarify and extend the underlying theoretical framework.

## REFERENCES

- Blau, P. M. 1964 *Exchange and Power in Social Life*. New York: Wiley, 2d printing, 1986. New Brunswick: N.J.: Transaction.
- 1987 "Microprocess and Macrostructure." In K. S. Cook, ed., *Social Exchange Theory*. Newbury Park, Calif.: Sage
- Bienenstock, Elisa I., and Phillip Bonacich 1992 "The Core as a Solution to Negatively Connected Exchange Networks." *Social Networks* 14:231-243.
- 1997 "Network Exchange as a Cooperative Game." *Rationality and Society* 9:937-965.
- Bonacich, P. 1986 "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92:1170-1182.
- Coleman, J. S. 1972 "Systems of Social Exchange." *Journal of Mathematical Sociology* 2:145-163.
- 1990 *The Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press
- Cook, Gillmore and Yamaguchi 1986 "Point and line vulnerability as bases for predicting the distribution of power in exchange networks: Reply to Willer." *American Journal of Sociology* 92:445-448.
- Cook, K. S., ed. 1987 *Social Exchange Theory*. Newbury Park, Calif.: Sage.
- , and R. M. Emerson 1978 "Power, Equity, and Commitment in Exchange Networks." *American Sociological Review* 43:721-739.
- , —, M. R. Gillmore, and T. Yamagishi 1983 "The Distribution of Power in Exchange Networks: Theory and Experimental Results." *American Journal of Sociology* 89:275-305.
- Cook and Whitmeyer 1992 "Two Approaches to Social Structure: Exchange Theory and Network Analysis." *Annual Review of Sociology*. 18:109-127
- Deutsch, M. 1964 "Homans in the Skinner Box." *Sociological Inquiry* 34:156-165.
- Ekeh, P. P. 1974 *Social Exchange Theory: The Two Traditions*. Cambridge, Mass.: Harvard University Press.
- Emerson, R. M. 1962 "Power-Dependence Relations." *American Sociological Review* 27:31-40.
- 1972a "Exchange Theory, Part I: A Psychological Basis for Social Exchange." In J. Berger, M. Zelditch, and B. Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.
- 1972b "Exchange Theory, Part II: Exchange Relations and Networks." In J. Berger, M. Zelditch, and B. Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.
- Friedkin, Noah E. 1992 "An Expected Value Model of Social Power: Predictions for Selected Exchange Networks." *Social Networks* 14:213-229.
- Heath, A. 1976. *Rational Choice and Social Exchange: A Critique of Exchange Theory*. Cambridge, Mass.: Cambridge University Press.
- Homans, G. C. 1958 "Social Behavior as Exchange." *American Journal of Sociology* 62:597-606.
- 1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace, and World.
- 1974 *Social Behavior: Its Elementary Forms*, 2nd ed. New York: Harcourt, Brace, and World.
- Kelley, H. H., and J. Thibaut 1978 *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.
- Lawler, Edward J., and Jeongkoo Yoon 1996 "Commitment in Exchange Relations: A Test of a Theory of Relational Cohesion." *American Sociological Review* 61:89-108.
- Levi-Strauss, C. 1949 *Les Structures Elementaires de la Parents*. Paris: Presses Universitaires de France.
- 1969 *The Elementary Structures of Kinship*. Boston: Beacon.
- Malinowski, B. 1922 *Argonauts of the Western Pacific*. London: Routledge and Kegan Paul.
- Markovsky, B., D. Willer, and T. Patton 1988 "Power Relations in Exchange Networks." *American Sociological Review* 53:220-236.
- Mauss, M. 1925 *Essai sur le don in Sociologie et Anthropologie*. Paris: Presses Universitaires de France. Translated into English by Ian Cunnison as *The Gift*. New York: Free Press, 1954.
- Miller, N. E. and J. Dollard 1941 *Social Learning and Imitation*. New Haven, Conn.: Yale University Press.

- Molm, L. D. 1981. "The Conversion of Power Imbalance to Power Use." *Social Psychology Quarterly* 44:151-163.
- 1987 "Power-Dependence Theory: Power Processes and Negative Outcomes." In E. J. Lawler and B. Markovsky, eds., *Advances in Group Processes*, vol. 4. Greenwich, Conn.: JAI Press.
- 1989 "Punishment Power: A Balancing Process in Power-Dependence Relations." *American Journal of Sociology* 94 (6):1392-1418.
- Molm, Linda D. 1997 *Coercive Power in Social Exchange*. Cambridge, UK: Cambridge University Press.
- Pfeffer, Jeffrey, and Gerald R. Salancik 1978 *The External Control of Organizations: A Resource Dependence Perspective*. New York: Harper and Row.
- Schneider, H. K. 1974 *Economic Man: The Anthropology of Economics*. New York: Free Press.
- Thibaut, J., and H. H. Kelley 1959 *The Social Psychology of Groups*. New York: Wiley.
- Turner, J. H. 1986 *The Structure of Sociological Theory*, 4th ed. Chicago: Dorsey Press.
- Willer, David 1987 *Theory and Experimental Investigation of Social Structures*. New York: Bordon and Breach.
- Willer and Anderson 1981 Willer, David and Bo Anderson, eds. 1981. *Networks, Exchange and Coercion*. New York: Elsevier/Greenwood
- Yamaguchi, K. 1996 "Power in Networks of Substitutable and Complementary Exchange Relations: A Rational-Choice Model and an Analysis of Power Centralization." *American Sociological Review* 61:308-322.

KAREN S. COOK

## SOCIAL FORECASTING

Forecasting has been important in sociological thought. Early European sociologists argued that societies progress through inevitable historical stages; those theories helped sociologists predict all societies' futures. Early American sociologists adopted the pragmatists' rule that a science proves it "works" by predicting future events (Schuessler 1971). Sociologists, however, have only recently adopted methods appropriate for those early goals. The review in this article of the delayed development of social forecasting includes (1) three sociologists' conceptual uses of forecasting and some reasons their suggestions were not followed, (2) qualitative and quantitative methods of forecast-

ing, and (3) recent indications of increased interest in forecasting.

### FORECASTING TRADITIONS

Sociologists have contributed several social forecasting concepts that were historically significant enough to become traditional orientations in the analysis of the future. William F. Ogburn "held that in the modern world technological inventions commonly come first and social effects later. By reason of this lag, it is possible, he argued, to anticipate the future and plan for its eventualities" (Schuessler 1971, p. 309). For example, new possibilities came into conflict with family values when the invention of effective birth control gave women new choices. Ogburn's contribution was to suggest that cultural lags are inevitable but that the period of disruption they cause can be shortened (Reiss 1986).

Merton (1949) challenged Ogburn's idea that the effects of inventions can be easily anticipated. Each invention has an apparent goal, or manifest function, that it is hoped it will perform in society. Each change, however, also contains the possibility of performing a number of *latent functions*. These are unanticipated side effects that often are not desired and sometimes are dangerous. The institutions of society are closely intertwined, and an invention in one area can cause shocks throughout the system. The automobile is an example. Its manifest function of changing transportation has been fulfilled, but at the cost of serious ecological and sociological changes.

Merton's (1949) second warning was that social forecasting is unique because it tries to predict the behavior of humans, who change their minds. The *self-fulfilling prophecy* is a forecast that makes people aware of real or imagined new opportunities or dangers to be avoided. Merton demonstrated that false forecasts can have powerful effects if they gain public acceptance. For example, a sound bank can be destroyed by a run on its funds caused by a prediction of failure. Henshel's more inclusive concept—the self-altering prediction—shows that forecasts can be self-defeating as well as self-fulfilling. W. I. Thomas's theorem, "If men define situations as real, they are real in their consequences," applies particularly to the definitions societies make of the future (Henshel 1978, p. 100).

Moore challenged sociologists to go beyond safe prophecy based on orderly trends and attack the difficult problem of “how to handle sharp changes in the magnitude of change, and sharp (or at least clear) changes in direction” (Moore 1964, p. 332). There are four types of *discontinuous societal change*: (1) Some societies are changed drastically by an *exogenous variable*, an idea or value from another society. Modern Japan is an example. (2) A society’s rate of development can increase spontaneously, creating an abundance of new ideas. This is an exponential acceleration, a *change in the rate of change*. (3) Moore attributes *changes in the direction of change* to the existence of a dialectic of values in each society’s apparent trend. For example, a society may appear to be profit-oriented and ecologically exploitative, but there also exists a counterset of values that stress harmony with each other and with nature. If a shift in such basic value emphases could be predicted, many other associated forecasts could be made. (4) Finally, Moore recognizes that there are pure *emergents*, inventions such as money and writing, that cannot be thought of as parts of trends.

Moore drew a methodological moral from these complexities: “One must somehow move from discrete necessary conditions to cumulative and sufficient ones” (Moore 1964, p. 334). That is, the search for the one trend or causal variable that drives societal change should be abandoned. The summation and particularly the interaction of many component developments create events.

In 1966 Moore asked sociologists to put aside value-free scientific rules and attempt to construct *preferable futures* that might help “mankind survive for the next twenty years” (Moore 1966, p. 270). Moore was confronting what he felt to be the main reason why forecasting was done so infrequently. It is professionally permissible for sociologists to examine social change both currently and retrospectively, but making a forecast leaves one liable to being labeled a utopian (Winthrop 1968, p. 136). Utopian thinking is in disrepute because past advocates allowed their values to cloud their constructions. However, images of the future provide goals and determine how people plan and therefore how they behave in the present. Moore sought utopias that would perform a necessary social planning function by constructing alternative directions for human purpose.

### WHY SOCIAL FORECASTING HAS DEVELOPED SLOWLY

Sociologists’ basic methodological orientations preclude an interest in forecasting. Sociologists analyze society’s static interconnections and concentrate on the social structures that persist. They have not developed skill in isolating the sequences of dynamic social behavior (Moore 1966). They are better at categorizing and typing people than at predicting how individuals might change from one type to another.

Many sociologists feel that not enough is known to predict future events. They point to economists and demographers and ask, If they are failing with their more quantifiable data, how can complex social changes be anticipated? One school of thought sees sociology as a qualitative art form that will never be a statistically modeled science. Critical sociologists object on moral grounds. They feel that society requires essential restructuring before positive change can be effected. Since most forecasting is based on models of the current structure, they feel that it sanctions unjust social arrangements (Henshel 1982).

### JUDGMENTAL AND QUALITATIVE FORECASTING METHODS

The futurists (Bell 1997; Kurian and Molitor 1996) see “the challenge being not just to forecast what the future will be, but to make it what it ought to be” (Enzer 1984, p. 202). The actual future is too complex to be predefined, but possible futures can be constructed that can be instructive. In addition, secondary forecasts can be made that estimate the effects of policy actions on the original course of development (Colquhoun 1996). The pace of change is considered too rapid to be captured by traditional methods reliance on a careful quantitative reconstruction of the past. This justifies the use of experts’ opinions, and futurists’ methods are ways of systematizing those judgments (Allen 1978, p. 79).

A discontinuous social change usually is preceded by a “substantial restructuring of basic tenets and beliefs” (Holroyd 1978, p. 37). Such paradigm shifts are revolutionary, such as the rejection of the earth as the center of the universe. They appear in fields of knowledge in which one system

of thought seems to be in control but is unable to solve important problems. Holroyd, for example, predicts a paradigm shift in economics because its current theories are unable to deal with essential problems such as scarcity of natural resources. Futurists anticipate shifts by compiling lists of crucial issues in the institutions of society. When the gap between current and desired conditions is large, that area is monitored closely for discontinuous change (Holroyd 1978, p. 38).

*Cross-impact matrices* are constructed by listing all possible future events in the problem area under study (Allen 1978, pp. 132–145). Each event is recorded as a row and a column in a square matrix. This allows the explicit examination of every intersection of events when one asks: What is the probability that the first will occur if it has been preceded by the other? The probabilities of occurrence can be derived from available data but are often judgments. Cross-impact analysis is a systematic way of heeding Merton's warning about not overlooking possibly damaging latent consequences. It is a tool for spotting crucial turning points or originating novel viewpoints by examining the intersections of change at which experts' judgments conflict.

*Delphi surveys* constitute an ingenious method for allowing the interaction of expert judgments while avoiding the contamination of social status or damage to reputations because of radical or mistaken pronouncements (Henshel 1982). In a series of survey rounds, everyone sees the distribution of others' responses without knowing the proponents' identities. A composite forecast emerges as anonymous modifications are made at each round.

After a review of forecasting methods, Ascher (1978) chose *scenarios* as one of only two methods he could recommend. A scenario is "a hypothetical sequence of events constructed for the purpose of focusing attention on causal processes and decision points" (Herman Kahn, quoted in Wilson 1978, p. 225). It is a story, but a complex one based on all available data and usually constructed after a cross-impact analysis has isolated possible turning points. Usually, two or three related scenarios are constructed to illustrate alternative futures that could be determined by particular decisions.

It is not surprising that an expert's decision process can be made explicit. What is surprising is

that in many studies the systematic model of an expert often forecasts better than the person does (Armstrong 1978). In *bootstrapping*, the forecaster's individualized decision procedures become the "bootstraps" by which a systematized procedure is "lifted" into an orderly routine. Such a model can be made deductively through interviews that isolate and formalize the decision rules or inductively by starting with a series of past forecasts and attempting to infer the rules that accounted for the differences between them.

Metaforecasting (Makridakis 1988) represents an essential summary of these considerations and a bridge to more quantitative methods. It combines judgmental and statistical estimates. It attempts to include historical and social information to overcome the tendency to ignore or overreact to changes in established patterns or relationships.

### SOCIAL DEMOGRAPHY

Demography is the most established form of social forecasting, and its methods and record can be found elsewhere (Henshel 1982). This article will discuss only two elements from its continuing development: a method that has had wide influence and what can be learned from its frequent failures to predict future population sizes.

A cohort is an aggregate of individuals of similar age who therefore experience events during the same time period (Reiss 1986, p. 47). *Cohort analysis* was first used by Norman Ryder to study the changing fertility behaviors of women born during the same five-year periods. Since that time, cohorts have been used in the study of many areas of social change to differentiate the changes that are result from individuals maturing through the stages of life from those caused by powerful societal events or value shifts.

Demographers failed to anticipate the post-war baby boom and the onset of its decline. These errors were due to *assumption drag*, "the continued use of assumptions long after their validity has been contradicted by the data" (Ascher 1978, p. 53). Henshel (1982) says that demographers probably ignored these turning points because they simply talked to each other too much. They reassured each other that their assumptions and their extrapolations from past trends would soon reas-

sert themselves in the data. Recognition of this error of developing an isolated club of forecasters has helped economists and will help sociologists avoid a similar regimentation of estimates.

The mix of assumptions and actual data varies widely in *simulation models*. The most useful models test a set of explicit assumptions so that no interactions between variables are overlooked. Models have contributed the idea of the feedback loop as an important caution against unidirectional thinking. This common system characteristic occurs when an effect reaches a sensitive level and begins a reaction that modifies its own cause (Simmons 1973, p. 195). Often, however, the mix of assumptions and facts in simulations leans too heavily toward judgments. So-called black-box modeling (McLean 1978), in which equations are hidden, can produce output that is plausible and provocative but also unrealistic. The creator of the *Limits to Growth* study admitted that “in *World Dynamics* . . . there is no attempt to incorporate formal data. . . . All relationships are intuitive” (Simmons 1973, p. 208). That study extrapolated what have come to be seen as extreme assumptions of geometric growth unchecked by social adaptation. Its dramatic predictions of imminent shortages had a wide but unwarranted impact (Cole et al. 1973). A comment on those failed predictions and their popularity at the time of their publication sets the context in which all “modeled” forecasts should be received: “The apparent detached neutrality of a computer model is as illusory as it is persuasive. Any model of a social system necessarily involves assumptions about the workings of that system, and these assumptions are necessarily colored by the attitudes and values of the individuals or groups concerned. . . . [C]omputer models should be regarded as an integral part of political debate. . . . The model is the message” (Freeman 1973, p. 7).

#### PRAGMATIC STATISTICAL ANALYSIS OF TIME SERIES

Attention has shifted to techniques that are less concerned with demonstrating the effects of assumed patterns. Time series are records of observations through time. Traditional *time series analysis* projects “future values of a variable based entirely on the past and present observations of that vari-

able” (Levine et al. 1999). It involves isolating the trend inside the many “noisy” or seasonal factors that may obscure it. The techniques have been well developed, are taught in undergraduate management statistics courses, and have been adapted for spreadsheet software available on most computers. The problem, however, is how much faith one can put in the idea that “people do what they usually do.” Time series projections are essential first steps in discovering patterns of behavior of aggregates of people over time. Such patterns often persist, but some shock (invention, immigration, social redefinition such as “the sixties,” or adjustment of tradition such as decreasing sexism) may cause disruption. In recognition of these sociological disruptions, time series are being explored from the viewpoint that any variable may be uniquely complex and subject to sudden change.

*Time series regressions* uncover structural relationships involved in the history of two or more variables. Before the relationship can be assessed, sources of error must be isolated and controlled. The most important of these errors are (1) the overall trend of change that would obscure any specific interrelationship and (2) the autocorrelation effect of internal dependence of an observation on previous observations. If a relationship seems to explain the data series’ movements, it is tested with ex-post forecasts that can be verified within the range of available data. If these succeed, “ex-ante-forecasts can be used to provide educated guesses about the path of the variables into the blind future” (Ostrom 1990, p. 77).

*Autoregressive moving average (ARIMA)* models predict a variable’s current status by using a combination of its previous observations and mathematically approximated random shocks. The goal is to find a pattern that fits the immediate data, not to understand relationships. ARIMA models are useful in interrupted *time series analysis*, in which the impact of a policy or another intervention can be examined by seeing how different the variable’s patterns are before and after the intervention (McDowall et al. 1980). Autoregressive models have a limitation important for social forecasting, in which historical data are relatively scarce. “Because ARIMA models must be identified from the data to be modeled, relatively long time series are required” (McCleary and Hay 1980, p. 20). Fifty observations are recommended.



*Exponential smoothing* is widely used and is as reliable as more complicated methods (Gardner 1985). In its simplest form, the next period's forecast is based on the current forecast plus a portion of the error it made. That is, the difference between the current time period's forecast and the actual value is weighted and used to adjust the next period's expected value. The higher the value of the weight used is, the more the error adjustment contributes and the more quickly the model will respond to changes. Exponential smoothing is used in early detection of curvilinear changes, when the rate of change speeds or slows (Gardner 1987).

### FUTURE TRENDS

Forecasting is being done. It is central in business and government planning. Even though many of these forecasts' essential variables are social or are found in social contexts (such as family decisions to move, build, and purchase or the development of social problems), economists have become society's designated forecasters (Henshel 1982; Stimson and Stimson 1976). Sociologists will not change this imbalance easily, but there are some indications that forecasting may finally become part of everyday sociological work.

Assumptions that a particular cycle or curve is the natural or underlying process of all change have been abandoned, and pragmatic methods are now widespread. It is also accepted that a forecast is developed only to be monitored for possible discontinuities. Trend extrapolations rarely are done without accompanying methods for describing the expected deviations.

Two forecasting methods are particularly promising because they allow sociologists to build on traditional skills. Componential or segmentation forecasting (Armstrong 1978) recognizes that an aggregate forecast can be improved by combining forecasts made on the population's component social groups. Sociologists are best able to distinguish the groups that should be treated separately. Pooled time series analysis (Sayrs 1989) combines cross-sectional descriptions such as one-time surveys. Sociologists are expert at describing interconnections in the structures of organizations or societies, and now they have the opportunity to study these social arrangements over time.

Society has recognized the wisdom of the early concern about anticipating the latent effects of social and technological inventions. Progress no longer seems inevitable. The popular question now is, Can someone assure us that a new element will not be as destructive as past changes?

Sociologists seem to be uniquely suited to help forecasting become more plausible because their working assumptions counter the weaknesses of current methods. The idea that technological innovation or economic cycles drive social change has produced today's mechanistic, ultrarational, anti-individualistic models that assume that the population is homogeneous (Dublin 1992). All these weaknesses are naturally contradicted when sociologists expand their vision of a population to include the cultural diversity of the social contexts that produce, accept or reject, and always modify the effects of technological and economic circumstances.

The future acceptance of forecasting also depends on sociologists' ability to improve the preparation and presentation of forecasts by using their traditional strengths. Forecasts will be accepted by policymakers and the public only when the quasi-theories they hold about the future are specifically addressed and proved false. Sociologists know this better than other social scientists do; they often are called on to dispel labels and popular theories that are so entrenched that they make any new attempt at explanation seem a "fool's experiment" to the forecaster's audience. They also are used to the idea of various and multiple causes acting in a situation and therefore are skilled at isolating "unanticipated consequences."

Forecasts will improve and become more plausible when they place less importance on traditional scientific formulations. A forecast is not a hypothesis. Hypotheses must be made in advance of the behavior they are meant to predict to assure a full and objective test of the theories that produced them. Forecasts demand monitoring of predictions and adaptation of forecasts to circumstances. A forecast is as good as its ability to anticipate and allow the inclusion of changing social forces. That is, its main function is not to make an accurate prediction of future events but to isolate and interrelate the many factors in the current situation that may be causally powerful. Understanding

the current social situation's complexity is the most important factor.

REFERENCES

Allen, T. Harrell 1978 *New Methods in Social Science Research: Policy Sciences and Futures Research*. New York: Praeger.

Armstrong, J. Scott 1978 *Long-Range Forecasting: From Crystal Ball To Computer*. New York: Wiley.

Ascher, William 1978 *Forecasting: An Appraisal for Policymakers and Planners*. Baltimore: Johns Hopkins University Press.

Bell, Wendell 1997 *Foundations of Futures Studies*, vols. I and II. New Brunswick, N.J.: Transaction.

Cole, H. S. D., Christopher Freeman, Marie Jahoda, and K. L. R. Pavitt, eds. 1973 *Models of Doom: A Critique of the Limits to Growth*. New York: Universe.

Colquhoun, Robert 1996 "The Art of Social Conjecture: Remembering Bertrand de Jouvenel." *History of the Human Sciences* 9(1): 27-42

Dublin, Max 1992 *FutureHype: The Tyranny of Prophecy*. New York: Plume.

Enzer, Selwyn 1984 "Anticipating the Unpredictable." *Technological Forecasting and Social Change* 26:201-204.

Freeman, Christopher 1973 "Malthus with a Computer" In H. S. D. Cole, Christopher Freeman, Marie Jahoda, and K. L. R. Pavitt eds., *Models of Doom: A Critique of the Limits to Growth*. New York: Universe.

Gardner, Everette S., Jr. 1985 "Exponential Smoothing: The State of the Art." *Journal of Forecasting* 4:1-28.

— 1987 "Short-Range Forecasting." *LOTUS* 3:54-58.

Henshel, Richard L. 1978 "Self-Altering Predictions." In Jib Fowles, ed., *Handbook of Futures Research*. Westport, Conn.: Greenwood Press.

— 1982 "Sociology and Social Forecasting." In Ralph H. Turner and James F. Short, eds., *Annual Review of Sociology*, vol. 8. Palo Alto, Calif.: Annual Reviews.

Holroyd, P. 1978 "Change and Discontinuity: Forecasting for the 1980s." *Futures* 10:31-43.

Kurian, George T., and Graham Molitor 1996 *Encyclopedia of the Future*, vols. I and II. New York: Macmillan.

Levine, David M., Mark L. Berenson, and David Stephan 1999 *Statistics for Managers Using Microsoft Excel*. Upper Saddle River N.J.: Prentice-Hall

Makridakis, Spyros 1988 "Metaforecasting: Ways of Improving Accuracy and Usefulness." *International Journal of Forecasting* 4(3):467-91.

McCleary, Richard, and Richard A. Hay, Jr. 1980 *Applied Time Series Analysis for the Social Sciences*. Beverly Hills, Calif.: Sage.

McDowall, David, Richard McCleary, Errol E. Meidinger, and Richard A. Hay, Jr. 1980 *Interrupted Time*. Newbury Park, Calif.: Sage.

McLean, J. Michael 1978 "Simulation Modeling." In Jib Fowles, ed., *Handbook of Futures Research*. Westport, Conn.: Greenwood Press.

Merton, Robert K. 1949 *Social Theory and Social Structure*, rev. ed. Glencoe, Ill.: Free Press.

Moore, Wilbert E. 1964 "Predicting Discontinuities in Social Change." *American Sociological Review* 29:331-338.

— 1966 "The Utility of Utopias." *American Sociological Review* 31:756-772.

Ostrom, Charles W., Jr. 1990 *Time Series Analysis. Regression Techniques*, 2nd ed. Newbury Park, Calif.: Sage.

Reiss, Albert J., Jr. 1986 "Measuring Social Change." In Neil J. Smelser and Dean R. Gerstein, eds., *Behavioral and Social Science: Fifty Years of Discovery*. Washington, D.C.: National Academy Press.

Sayrs, Lois W. 1989 *Pooled Time Series Analysis*. Beverly Hills, Calif.: Sage.

Schuessler, Karl 1971 "Continuities in Social Prediction." In H. L. Costner, ed., *Sociological Methodology, 1971*. San Francisco: Jossey-Bass.

Simmons, Harvey 1973 "System Dynamics and Technocracy." In H. S. D. Cole, Christopher Freeman, Marie Jahoda, and K. L. R. Pavitt eds., *Models of Doom: A Critique of the Limits to Growth*. New York: Universe.

Stimson, John, and Ardyth Stimson 1976 "Sociologists Should Be Put to Work as Forecasters." *American Sociologist* 11:49-56.

Wilson, Ian H. 1978 "Scenarios." In B. Fowles, ed. *Handbook of Futures Research*. Westport, Conn.: Greenwood Press.

Winthrop, Henry 1968 "The Sociologist and the Study of the Future." *American Sociologist* 3:136-145.

JOHN STIMSON

**SOCIAL GERONTOLOGY**

See Aging and the Life Course; Cohort Perspectives; Filial Responsibility; Intergenerational Relations; Intergenerational Resource Transfers; Long-Term Care, Long Term Care Facilities; Retirement; Widowhood.

## SOCIAL IMITATION

See Behaviorism; Socialization; Social Psychology.

## SOCIAL INDICATORS

Social indicators are statistical time series that are “used to monitor the social system, helping to identify changes and to guide intervention to alter the course of social change” (Ferriss 1988, p. 601). Examples are unemployment rates, crime rates, estimates of life expectancy, health status indices such as the average number of “healthy” days (or days without activity limitations) in the past month for a specific population, school enrollment rates, average achievement scores on a standardized test, rates of voting in elections, and measures of subjective well-being such as satisfaction with life as a whole.

### HISTORICAL DEVELOPMENTS

**Social Indicators in the 1960s.** The term “social indicators” was given its initial meaning in an attempt by the American Academy of Arts and Sciences for the National Aeronautics and Space Administration in the early 1960s to detect and anticipate the nature and magnitude of the second-order consequences of the space program for American society (Land 1983, p. 2; Noll and Zapf 1994, p. 1). Frustrated by the lack of sufficient data to detect such effects and the absence of a systematic conceptual framework and methodology for analysis, some in the project attempted to develop a system of social indicators—statistics, statistical series, and other forms of evidence—with which to detect and anticipate social change and to evaluate specific programs and determine their impact. The results of this part of the project were published in a volume (Bauer 1966) called *Social Indicators*.

The appearance of this volume was not an isolated event. Several other influential publications commented on the lack of a system for charting social change and advocated that the U.S. government establish a “system of social accounts” that would facilitate a cost-benefit analysis of more than the market-related aspects of society already

indexed by the National Income and Product Accounts (National Commission on Technology, Automation and Economic Progress 1966; Sheldon and Moore 1968). The need for social indicators also was emphasized by the publication of *Toward a Social Report* on the last day of the Johnson administration in 1969. The report was conceived of as a prototypical counterpart to the annual economic reports of the president, and each of its chapters addressed major issues in an area of social concern (health and illness; social mobility; the physical environment; income and poverty; public order and safety; learning, science, and art; and participation and alienation) and provided an assessment of the current conditions. In addition, the document firmly linked social indicators to the idea of systematic reporting on social issues for the purpose of public enlightenment.

Generally speaking, the sharp interest in social indicators in the 1960s grew out of the movement toward the collection and organization of national social, economic, and demographic data that began in Western societies in the seventeenth and eighteenth centuries and accelerated in the twentieth century (Carley 1981, pp. 14–15). The work of the sociologist William F. Ogburn and his collaborators at the University of Chicago in the 1930s and 1940s on the theory and measurement of social change is more proximate and sociologically germane (Land 1975). As chairman of President Herbert Hoover’s Research Committee on Social Trends, Ogburn supervised the production of the two-volume *Recent Social Trends* (1933), a pathbreaking contribution to social reporting. Ogburn’s ideas about the measurement of social change influenced several of his students—notably Albert D. Biderman, Otis Dudley Duncan, Albert J. Reiss, Jr., and Eleanor Bernert Sheldon—who played major roles in the emergence and development of the field of social indicators in the 1960s and 1970s.

**Social Indicators in the 1970s and 1980s.** At the end of the 1960s, the enthusiasm for social indicators was sufficiently strong and broad-based for Duncan (1969, p. 1) to write of the existence of a social indicators movement. In the early 1970s, this led to, among other things, the establishment in 1972, with National Science Foundation support, of the Social Science Research Council Center for Coordination of Research on Social Indica-

tors in Washington, D.C.; the publication of several major efforts to define and develop a methodology for the measurement of indicators of subjective well-being (Campbell and Converse 1972; Andrews and Withey 1976; Campbell et al. 1976); the commencement of a federal government series of comprehensive social indicators books of charts, tables, and limited analyses (U.S. Department of Commerce 1974, 1978, 1980); the initiation of several continuing data series based on periodic sample surveys of the national population (such as the annual National Opinion Research Center's [NORC] General Social Survey and the annual National Crime Survey of the Bureau of Justice Statistics); the publication in 1974 of the first volume of the international journal *Social Indicators Research*; and the spread of social indicators and/or social reporting to numerous other nations and international agencies, such as the United Nations and the Organization for Economic Cooperation and Development.

Social indicators activities slowed in the 1980s as funding cuts and nonrenewals led to the closing of the Center for Coordination of Research on Social Indicators; the discontinuation of related work at several international agencies; the termination of government-sponsored social indicators reports in some countries, including the United States; and the reduction of statistical efforts to monitor various aspects of society. Several explanations have been cited for this slowdown (Andrews 1989; Bulmer 1989; Innes 1989; Johnston 1989; Rockwell 1987). Certainly, politics and the state of national economies in the early 1980s are among the most salient proximate causes. Administrations that came to power in the United States and elsewhere based decisions more on a "conservative ideology" and less on current social data than had been the case earlier. Also, faltering economies that produced large government budget deficits provided an incentive to make funding cuts. In addition to these immediate factors, there was a perceived lack of demonstrated usefulness of social indicators in public policymaking that was due in part to an overly simplistic view of how and under what conditions knowledge influences policy; this topic is treated more fully below in the discussion of current uses of social indicators. Before that, a more detailed discussion of types of indicators and their measurement and organization into accounting systems is necessary.

### THREE TYPES OF SOCIAL INDICATORS

**Criterion Indicators.** On the basis of the premise that social indicators should relate directly to social policymaking considerations, an early definition by the economist Mancur Olson, the principal author of *Toward a Social Report*, characterized a social indicator as a "statistic of direct normative interest which facilitates concise, comprehensive and balanced judgments about the condition of major aspects of a society" (U.S. Department of Health, Education, and Welfare 1969, p. 97). Olson stated that such an indicator is in all cases a direct measure of welfare and is subject to the interpretation that if it changes in the "right" direction while other things remain equal, things have gotten better or people are better off. Accordingly, by this definition, statistics on the number of doctors or police officers could not be social indicators, whereas figures on health or crime rates could be.

In the language of policy analysis (Fox 1974, pp. 120–123), social indicators are "target" or "output" or "outcome" or "end-value" variables toward changes in which a public policy (program or project) is directed. This use of social indicators requires (Land 1983, p. 4) that (1) society agree about what needs improving, (2) it be possible to decide unambiguously what "getting better" means, and (3) it be meaningful to aggregate the indicators to the level of aggregation at which the policy is defined.

In recognition of the fact that other meanings have been attached to the term "social indicators," the tendency among recent authors is to use a somewhat different terminology for the class of indicators identified by Olson. For instance, Land (1983, p. 4) termed this the class of "normative welfare indicators." Building on the Olson approach, MacRae (1985, p. 5) defined "policy indicators" as "measures of those variables that are to be included in a broadly policy-relevant system of public statistics." With a meaning similar to that of MacRae, Ferriss (1989, p. 416) used the felicitous term "criterion indicators."

**Life Satisfaction and/or Happiness Indicators.** Another class of social indicators has its roots in the work of Campbell and Converse in the early 1970s. In *The Human Meaning of Social Change* (1972), they argued that the direct monitoring of key social-psychological states (attitudes, expectations, feelings, aspirations, and values) in the popula-

tion is necessary for an understanding of social change and the quality of life. In this approach, social indicators are used to measure psychological satisfaction, happiness, and life fulfillment by employing survey research instruments that ascertain the subjective reality in which people live. The result may be termed “life satisfaction,” “subjective well-being,” or “happiness indicators.”

The Campbell-Converse approach led to two major methodological studies in the 1970s (Andrews and Withey 1976; Campbell et al. 1976) and a subsequent edited volume (Andrews 1986) exploring the utility of various survey and analytic techniques for mapping individuals’ feelings of satisfaction with numerous aspects (“domains”) of their experiences. These studies examine domains ranging from the highly specific (house, family, etc.) to the global (life as a whole). A large number of other studies and applications of these concepts and techniques have appeared over the past three decades (for reviews, see Diener 1994; Diener et al. Smith 1999; and Veenhoven 1996) and continue to appear; one or more studies of subjective well-being indicators can be found in almost every issue of *Social Indicators Research*. Research on the related concept of happiness as an index of well-being was surveyed by Veenhoven (1984).

The principle that the link between objective conditions and subjective well-being (defined in terms of responses to sample survey or interview questions about happiness or satisfaction with life as a whole) is sometimes paradoxical; therefore, the idea that subjective as well objective states should be monitored is well established in the social indicators literature. However, numerous studies of the measurement and psychodynamics of subjective well-being over the last three decades have led to a better understanding of this construct (Cummins 1995, 1998). While research continues and the debates have not been settled, it appears that this construct may have both *traitlike* (i.e., a durable psychological condition that differs among individuals and contributes to stability over time and consistency across situations) and *statelike* (i.e., a condition that is reactive to situational differences) properties (Stones et al. 1995; Veenhoven 1994, 1998).

With respect to the statelike properties of subjective well-being, Davis (1984) used an accumulated sample from several years of NORC Gen-

eral Social Surveys to document the responsiveness of happiness with life as a whole to (1) “new money” (recent changes in the respondents’ financial status compared with the current income level), (2) “an old man/lady” (being married or having an intimate living partner), and (3) “two’s company” (a household size of two compared to living alone or in a family of three or more). Many other studies have found additional factors that are more or less strongly associated with variations in subjective well-being, but the relevance of intimate living conditions and/or family status almost always is replicated. The connection of subjective well-being to income levels has been an intriguing problem for social indicators researchers since Easterlin’s (1973) finding that income differences between nations predict national differences in happiness but that the association of happiness with income within countries is much weaker (for a review of this research literature, see Ahuvia and Friedman 1998). However, Davis’s finding of a positive relationship between “new money” or recent income changes and happiness has been replicated by Saris (1998), using data from a panel study conducted in Russia in the period 1993–1995.

**Descriptive Social Indicators.** Building on Ogburn’s legacy of research on social trends, a third approach to social indicators focuses on social measurements and analyses designed to improve the understanding of what the main features of society are, how they interrelate, and how these features and their relationships change (Sheldon and Parke 1975, p. 696). This produces *descriptive social indicators*—indices of the state of society and the changes taking place within it. Although descriptive social indicators may be more or less directly (causally) related to the well-being goals of public policies or programs and thus include policy or criterion indicators, they are not limited to such uses. For instance, in the area of health, descriptive indicators may include preventive indicators such as the percentage of the population that does not smoke cigarettes as well as criterion indicators such as the number of days of activity limitations in the past month and an index of self-reported satisfaction with health. Ferriss (1990) published a compilation of descriptive indicators for the United States at the end of the 1980s; regularly published national social indicator compilations for other nations also contain numerous examples.

The various statistical forms descriptive social indicators can take are described by Land (1983, p. 6). These forms can be ordered by degree of abstraction from those which require only one or two data series and little processing (e.g., an age-specific death rate) to those which involve more complicated processing into a single summary index (e.g., years of life expectancy at a given age and years of active or disability-free life expectancy at a given age). Descriptive social indicators can be formulated at any of these levels of abstraction. Moreover, as described in Juster and Land (1981), these indicators can, at least in principle, be organized into demographic- or time-budget-based systems of social accounts.

#### THE ENLIGHTENMENT FUNCTION: MONITORING, SOCIAL REPORTING, AND FORECASTING

The social indicators movement was motivated by the principle that it is important to *monitor changes over time* in a broad range of social phenomena that extend beyond the traditional economic indicators and include *indicators of quality of life* (Andrews 1989, p. 401; Noll and Zapf 1994, p. 5). Many organized actors in contemporary society—including government agencies, organizations and activists interested in social change programs, scholars, and marketing researchers interested in market development and product innovations—monitor indicators in which they have a vested interest and want to see increase or decline (Ferriss 1988, p. 603).

A second principle that has been part of the social indicators movement from the outset (Biderman 1970; Land 1996) is that a critically important role of social indicators in contemporary democratic societies is *public enlightenment through social reporting*. In brief, modern democracies require social reporting to describe social trends, explain why an indicator series behaves as it does and how this knowledge affects interpretation, and highlight important relationships among series (Parke and Seidman 1978, p. 15).

It also is important to document the consequences that are reasonably attributable to changes in a series. This includes the systematic use of social indicators to *forecast trends in and/or turning points in social conditions* (Land 1983, p. 21). The

area of projection or forecasting is filled with uncertainties. Techniques range from the naive extrapolation of recent trends to future scenario construction to complicated model building with regression, time series, or stochastic process techniques. Moreover, there appear to be intrinsic limits to the accuracy of forecasts in large-scale natural and social systems (Land and Schneider 1987). However, demands for the anticipation of the future (at a minimum, a description of “what will happen if present trends continue”), foresight and forward thinking in the public and private sectors, and the assessment of critical trends (Gore 1990) appear to be an intrinsic part of contemporary postindustrial societies. Thus, it is prudent to expect that “anticipation” will become an increasingly important part of the enlightenment function of social indicators.

#### Social Reporting at the Turn of the Century.

As the decade of the 1990s unfolded, the model of a comprehensive national social report in the tradition pioneered by Ogburn and Olson clearly faltered in the United States, at least in the sense of federal government sponsorship and/or production. However, the key ideas of monitoring, reporting, and forecasting were evident to a greater or lesser extent in the production of continuing periodic subject-matter-specific publications by various federal agencies, including *Science Indicators* (published by the National Science Foundation), *The Condition of Education* (published by the Department of Education), the *Report to the Nation on Crime and Justice* (published by the Department of Justice), and numerous Census Bureau publications. Special topics involving groups of federal agencies also receive attention from time to time. For instance, in 1997 the Federal Interagency Forum on Child and Family Statistics began the annual publication *America's Children: Key National Indicators of Well-Being*. In addition, numerous private research organizations, policy institutes, and scholars continue to produce reports, monographs, and books interpreting social trends and developments in several areas of social concern.

In contrast to the situation in the United States, comprehensive social reports and social indicators compendiums continue to be published periodically in several other countries. Examples are the *Social Trends* series published annually since 1970 by the United Kingdom's Central Statistical Office, the *Datenreport* series published bien-

nially since 1983 by the Federal Republic of Germany, the *Social and Cultural Report* published biennially by the Social and Cultural Planning Office of the Netherlands, and *Australian Social Trends* published annually by the Australian Bureau of Statistics. Citations and summary reviews of these and other social indicators and social reports publications can be found in the quarterly newsletter and review of social reports *SINET: Social Indicators Network News* (see the World Wide Web home page: <http://www.soc.duke.edu/dept/sinet/index.html>).

The difference between the organization of social indicators and reporting work in the United States and that in other countries is in part attributable to the lack of a central statistical office responsible for the coordination of all government statistical activities in the United States. More generally, it is indicative of the fact that despite the invention of the ideas of social indicators and comprehensive social reporting in the United States, this nation has lagged in their institutionalization (Johnston 1989). Whether a new round of legislative efforts (e.g., then-Senator Albert Gore, Jr.'s, proposed Critical Trends Assessment Act [Gore 1990]) will create the necessary institutional base remains to be seen. Perhaps marking a turning point and indicative of things to come is Public Law 100-297, enacted April 28, 1988, which requires an annual education indicators report to the president and Congress.

**Quality of Life as a Unifying Concept.** Another development became vividly apparent in the 1990s (Land 1996): the widespread political, popular, and theoretical appeal of the "quality-of-life" (QOL) concept. As was noted above, this concept emerged and became part of the social indicators movement in the late 1960s and early 1970s as doubts were raised in highly developed Western industrial societies about economic growth as the major goal of societal progress (Noll and Zapf 1994, pp. 1-2). The "social costs" of economic growth were cited, and there was increasing doubt about whether "more" should be equated with "better." The QOL concept that resulted from this discussion was posed as an alternative to the increasingly questionable concept of the affluent society and entered discussions of social policy and politics as a new but more complex multidimensional goal. As a goal of social and economic

policy, QOL encompasses all (or at least many) domains of life and subsumes, in addition to individual material and immaterial well-being, collective values such as freedom, justice, and the guarantee of natural conditions of life for present and future generations. The political use of the idea of QOL is paralleled in the private sector by the widespread use and popularity of numerous rankings—based on weighted scales of multiple domains of well-being—of the "best" places to live, work, do business, play, and so on, whether they are cities, states, regions, or nations.

The theoretical appeal of the QOL concept as an integrating notion in the social sciences and related disciplines is due in part to the perceived importance of measuring individuals' subjective assessments of their satisfaction with various life domains and with life as a whole, as was reviewed above. For instance, QOL has become a concept that bridges the discipline of marketing research and strategic business policy with social indicators. Marketing is an importance social force—with far-reaching direct and indirect impacts on the prevailing QOL in a society—through consumer satisfaction (Samli 1987; Sirgy and Samli 1995) and its impact on satisfaction with life as a whole. The intersection of marketing research with social indicators through the QOL concept led to the organization in the mid-1990s of the International Society for Quality-of-Life Studies (for information about the society and its activities, see its Web homepage: <http://www.cob.vt.edu/market/isqols/>). Sociologists who want to become more involved in the field of social indicators should participate in this international and interdisciplinary society.

**Summary Indices of the Quality of Life.** As the twenty-first century approaches, it is evident that the field of social indicators is entering a new era of the construction of summary social indicators. Often these indices attempt to summarize indicators (objective and/or subjective) of a number of domains of life into a single index of the quality of life. They thus attempt to answer one of the original questions that motivated the social indicators movement: How are we doing overall in terms of the quality of life? With respect to our past? With respect to other comparable units (e.g., cities, states, regions, nations)? Many pioneers of the social indicators movement in the 1960s and 1970s backed away from the development of sum-

many indices to concentrate on conducting basic research on social indicators and the measurement of the quality of life and the development of a richer social database. With the tremendous increase in the quality of social data available for many societies today compared to two or three decades in the past, a new generation of social indicators researchers has returned to the task of summary index construction. Some examples are (1) at the level of the broadest possible comparisons of nations with respect to the overall quality of life, the Human Development Index (United Nations Development Programme 1993), Diener's (1995) Value-Based Index of National Quality of Life, and Estes's (1988, 1998) Index of Social Progress and (2) at the level of comparisons at the national level over time in the United States, the *American Demographics Index of Well-Being* (Kacapyr 1996), *The Fordham Index of Social Health* (Miringoff 1996), and the *Genuine Progress Indicator* (Redefining Progress 1995). The field of social indicators probably will see several decades of such index construction and competition among various indices, with a corresponding need for careful assessments to determine which indices have substantive validity for which populations in the assessment of the quality of life and its changes over time and across social space.

**THE POLICY ANALYSIS FUNCTION:  
POLICY GUIDANCE AND DIRECTED  
SOCIAL CHANGE**

Policy analysts distinguish various ways of guiding or affecting public policy, including *problem definition*, *policy choice and evaluation of alternatives*, and *program monitoring* (MacRae 1985, pp. 20–29). The social reporting–public enlightenment approach to social indicators centers on the use of social indicators in problem definition and the framing of the terms of policy discourse. Indeed, studies of the actual use of social indicators suggest that this is precisely the manner in which they have affected public action (Innes 1989).

However, policy analysts from Olson to MacRae always have hoped for more from social indicators: the shaping of public policy and planing through the policy choice process. At a minimum, this requires the identification of key variables that determine criterion indicators and changes in them

(i.e., causal knowledge). More generally, it requires the construction of elaborate causal models and forecasting equations (often in the form of a “computer model”) that can be used to simulate “what would happen if” under a variety of scenarios involving policies and actions. An example of this is the development of the National Cancer Institute model for the control and reduction of the incidence of cancer in the United States to the year 2000 (Greenwald and Sondik 1986). Various policy and action scenarios involving prevention, education, screening, and treatment and their implications for cancer mortality were simulated and estimated with this computer model. These simulations led to a decision to allocate funds to prevention, education, screening, and treatment, and their implications for cancer mortality were simulated and estimated with this computer model. These simulations led to a decision to allocate funds to a prevention program rather than to additional clinical treatment.

At a more discursive level, the following *model for directed social change* has emerged in policy uses of social indicators in areas such as health, education, and the welfare of children and youth in the United States (Ferriss 1998):

1. *Identify trends in criterion indicators*, the direction or rate of change of which should be changed.
2. *Gather intelligence* from experiments, field research, or theory that suggests what should be done to bring about the desired change.
3. *Launch a decentralized program to effect change in specific criterion indicators* by specific amounts, to be attained by a target date.
4. *Monitor progress* by periodically assessing trends on the specific indicators, modifying strategies as needed.
5. As initial goals are reached, *set new goals* for continued progress.

Many more applications of social indicators to policy choice and evaluation are likely to appear in the future. In particular, such applications probably will occur in three areas. The first is the additional development of well-grounded, theoretically informed, and policy-relevant indicators



and models for national and/or regional-level analyses in fields such as health, education, crime, and science (Bulmer 1989). In such applications, the phenomena to be included are definable and delimited and the limitations of the data on which the indicators are based are known. The second is the use of social indicators in the field of social impact assessment (Finsterbusch 1980; Land 1982), which has arisen as part of environmental impact assessment legislation and attempts to anticipate the social effects of large-scale public projects (e.g., dams, highways, nuclear waste disposal facilities) as well as to assess damage from both natural and human-made disasters (e.g., earthquakes, oil spills, nuclear plant accidents). This application of social indicators in impact assessments brings the field back full circle to its point of origination in the American Academy's effort of the 1960s. Finally, the many times series of indicators now available will increasingly be used by sociologists to assess theories, hypotheses, and models of social change, thus bringing social indicators data to bear on core issues in sociology.

(SEE ALSO: *Attitudes; Longitudinal Research; Public Opinion; Quality of Life; Social Change; Social Forecasting*)

## REFERENCES

- Ahuvia, Aaron C., and Douglas C. Friedman 1998 "Income, Consumption, and Subjective Well-Being: Toward a Composite Macromarketing Model." *Journal of Macromarketing* 18:153-168.
- Andrews, Frank M., ed. 1986 *Research on the Quality of Life*. Ann Arbor, Mich.: Institute for Social Research.
- 1989 "The Evolution of a Movement." *Journal of Public Policy* 9:401-405.
- , and Stephen B. Withey 1976 *Social Indicators of Well-Being: Americans' Perceptions of Life Quality*. New York: Plenum.
- Bauer, Raymond A., ed. 1966 *Social Indicators*. Cambridge, Mass.: MIT Press.
- Biderman, Albert D. 1970 "Information, Intelligence, Enlightened Public Policy: Functions and Organization of Societal Feedback." *Policy Sciences* 1:217-230.
- Bulmer, Martin 1989 "Problems of Theory and Measurement." *Journal of Public Policy* 9:407-412.
- Campbell, Angus, and Philip E. Converse 1972 *The Human Meaning of Social Change*. New York: Russell Sage Foundation.
- , —, and Willard L. Rodgers 1976 *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*. New York: Russell Sage Foundation.
- Carley, Michael 1981 *Social Measurement and Social Indicators: Issues of Policy and Theory*. London: Allen and Unwin.
- Cummins, Robert A. 1995 "On the Trail of the Gold Standard for Subjective Well-Being." *Social Indicators Research* 35:170-200.
- 1998 "The Second Approximation to an International Standard for Life Satisfaction." *Social Indicators Research* 43:307-334.
- Davis, James A. 1984 "New Money, An Old Man/Lady and 'Two's Company': Subjective Welfare in the NORC General Social Survey." *Social Indicators Research* 15:319-351.
- Diener, Ed 1994 "Assessing Subjective Well-Being: Progress and Opportunities." *Social Indicators Research* 31:103-157.
- 1995 "A Value-Based Index for Measuring National Quality of Life." *Social Indicators Research* 36:107-127.
- , Eunkook M. Suh, Richard E. Lucas, and Heidi L. Smith 1999 "Subjective Well-Being: Three Decades of Progress." *Psychological Bulletin* 110: forthcoming.
- Duncan, Otis Dudley 1969 *Toward Social Reporting: Next Steps*. New York: Russell Sage Foundation.
- Easterlin, Richard 1973 "Does Money Buy Happiness?" *The Public Interest* 30:3-10.
- Estes, Richard J. 1988 *Trends in World Development*. New York: Praeger.
- 1998 "Social Development Trends in Transition Economies. In K. R. Hope, Sr., ed., *Challenges of Transformation and Transition from Centrally Planned to Market Economies*. Nagoya, Japan: United Nations Centre for Regional Development.
- Ferriss, Abbott L. 1988 "The Uses of Social Indicators." *Social Forces* 66:601-617.
- 1989 "Whatever Happened, Indeed!" *Journal of Public Policy* 9:413-417.
- 1990 "The Quality of Life in the United States." *SINET: Social Indicators Network News* 21:1-8.
- 1998 "Slow Progress toward Education 2000 Goals." *SINET: Social Indicators Network News* 53:11-12.
- Finsterbusch, Kurt 1980 *Understanding Social Impacts: Assessing the Effects of Public Projects*. Beverly Hills Calif.: Sage.
- Fox, Karl A. 1974 *Social Indicators and Social Theory: Elements of an Operational System*. New York: Wiley-Interscience.

- Gore, Albert, Jr., 1990 "The Critical Trends Assessment Act: Futurizing the United States Government." *The Futurist* 24:22-28.
- Greenwald, Peter, and Edward J. Sondik, eds. 1986 *Cancer Control Objectives for the Nation: 1985-2000*, NCI Monographs 2. Washington, D.C.: U.S. Government Printing Office.
- Innes, Judith Eleanor 1989 "Disappointment and Legacies of Social Indicators." *Journal of Public Policy* 9:429-432.
- Johnston, Denis F. 1989 "Some Reflections on the United States." *Journal of Public Policy* 9:433-436.
- Juster, F. Thomas, and Kenneth C. Land, eds. 1981 *Social Accounting Systems: Essays on the State of the Art*. New York: Academic Press.
- Kacapyr, Elia 1996 "The Well-Being Index." *American Demographics* 18:32-43.
- Land, Kenneth C. 1975 "Theories, Models and Indicators of Social Change." *International Social Science Journal* 27:7-37.
- 1982 "Ex Ante and Ex Post Assessment of the Social Consequences of Public Projects and Policies." *Contemporary Sociology* 11:512-514.
- 1983 "Social Indicators." *Annual Review of Sociology* 9:1-26.
- 1996 "Social Indicators and the Quality-of-Life: Where Do We Stand in the Mid-1990s?" *SINET: Social Indicators Network News* 45:5-8.
- , and Stephen H. Schneider 1987 "Forecasting in the Social and Natural Sciences: An Overview and Statement of Isomorphisms." In K. C. Land and S. H. Schneider, eds., *Forecasting in the Social and Natural Sciences*. Boston: Reidel.
- MacRae, Duncan, Jr., 1985 *Policy Indicators: Links between Social Science and Public Policy*. Chapel Hill: University of North Carolina Press.
- Miringoff, Marc L. 1996 *1996 Index of Social Health*. Tarrytown, N.Y.: Fordham Institute for Innovation in Social Policy.
- National Commission on Technology, Automation and Economic Progress 1966 *Technology and the American Economy*, vol. 1. Washington, D.C.: U.S. Government Printing Office.
- Noll, Heinz-Herbert, and Wolfgang Zapf 1994 "Social Indicators Research: Societal Monitoring and Social Reporting." In I. Borg and P. P. Mohler, eds., *Trends and Perspectives in Empirical Social Research*. New York: Walter de Gruyter.
- Parke, Robert, and David Seidman 1978 "Social Indicators and Social Reporting." *Annals of the American Academy of Political and Social Science* 435:1-22.
- President's Research Committee on Social Trends 1933 *Recent Trends in the United States*. New York: McGraw-Hill.
- Redefining Progress 1995 *The Genuine Progress Indicator: Summary of Data and Methodology*. San Francisco: Redefining Progress.
- Rockwell, Richard C. 1987 "Prospect for Social Reporting in the United States: A Receding Horizon." In Jesse R. Pitts and Henri Mendras, eds., *The Tocqueville Review*, vol. 8. Charlottesville: University Press of Virginia.
- Samli, A. Coskun 1987 *Marketing and the Quality-of-Life Interface*. Westport, Conn.: Quorum.
- Saris, Willem E. 1998 "The Strength of the Causal Relationship between Living Conditions and Satisfaction." Paper presented at the fourteenth World Congress of Sociology Meeting, Montreal, Canada, July 26-August 1, 1998.
- Sheldon, Eleanor B., and Wilbert E. Moore, eds. 1968 *Indicators of Social Change: Concepts and Measurements*. New York: Russell Sage Foundation.
- , and Robert Parke 1975 "Social Indicator." *Science* 188:693-699.
- Sirgy, M. Joseph, and A. Coskun Samli, eds. 1995 *New Dimensions in Marketing/Quality-of-Life Research*. Westport, Conn.: Quorum.
- Stones, M. J., T. Hadjistavropoulos, J. Tuuko, and A. Kozma 1995 "Happiness Has Traitlike and Stalike Properties." *Social Indicators Research* 36:129-144.
- United Nations Development Programme 1993 *Human Development Report 1993*. New York: Oxford University Press.
- U.S. Department of Commerce 1974 *Social Indicators, 1973*. Washington, D.C.: U.S. Government Printing Office.
- 1978 *Social Indicators, 1977*. Washington, D.C.: U.S. Government Printing Office.
- 1980 *Social Indicators, III*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Health, Education, and Welfare 1969 *Toward a Social Report*. Washington, D.C.: U.S. Government Printing Office.
- Veerhoven, Ruut 1984 *Conditions of Happiness*. Boston: Reidel.
- 1994 "Is Happiness a Trait? Tests of the Theory That a Better Society Does Not Make People Any Happier." *Social Indicators Research* 33:101-160.
- 1996 "Developments in Satisfaction Research." *Social Indicators Research* 37:1-46.

— 1998 “Two State-Trait Discussions on Happiness: A Reply to Stones et al.” *Social Indicators Research* 43:211–225.

KENNETH C. LAND

## SOCIAL INEQUALITY

Social inequality refers to the graduated dimensions (Blau 1977), vertical classifications (Ossowski 1963; Schwartz 1981) and bounded categories (Tilly 1998), or hierarchical relations (Burt 1982) by which human populations at varying levels of aggregation are differentiated. This concept is among the oldest and most diversely defined in sociology, extending back at least as far as Plato's conception of the republic and developed subsequently in the social theories of Marx [1859] 1976–1978, Mosca (1939), Weber [1947] 1978, Simmel (1896), Sorokin (1941), Eisenstadt (1971), Merton (1968), and others. The construct often is used interchangeably with related (though relatively more specific) concepts such as social class, social stratification, socioeconomic status, power, privilege, cumulative advantage, dependence, and dominance. It is relevant for the study of social systems that range in size from the dyad (Simmel 1896) to the modern world system (Wallerstein 1974).

### SOCIAL INEQUALITY AS A GRADUATED DIMENSION

When social inequality is conceptualized as a graduated dimension, it is treated as a distributional phenomenon. Here the approach is to define inequality in terms of the distribution of socially valued attributes such as education, income, information, health, and influence in a population. However, distributional phenomena can be examined from one of two very different assumptions. The first assumption views inequality as being an outcome of or generated by the underlying distribution of valued traits among individuals. In this sense, it refers to “regular differences in power, goods, services, and privileges among defined sets” of actors (Granovetter and Tilly 1988). The second assumption views inequality strictly as a system-level property with individual-level differences that are defined as derivative rather than generative (Blau 1977). Distributions such as the size of the system and its total volume of resources are exam-

ined as higher levels of aggregation, with the goal of determining the overall level of inequality (oligarchy) across systems and without reference to individual differences (e.g., Lenski 1966; Mayhew 1973; Mayhew and Schollaert 1980).

Both approaches operationalize inequality along criteria that usually are measurable at the level of individual actors (persons, races, gender categories, organizations, nation-states) in a system. Early applications of the first assumption can be found in Pareto's ([1897] 1980) examinations of income distributions and the circulation of elites. Pareto proposed that economic and political inequality emerged from the distribution and redistribution of “congenital abilities” that were valued within social systems. Sorokin (1941) proposed similar arguments to explain social and cultural processes of mobility and inequality.

Among the most influential and controversial conceptualizations of inequality as a graduated dimension emerging from individual differences was Davis and Moore's (1945) functionalist statement of the principles of stratification. Those authors argued that social inequality results from the differential distribution of societal rewards to individuals on the basis of their relative achievement of ranked social positions. This achievement process, with its implications for social mobility, was formally specified by Blau and Duncan (1967), who established that educational attainment mediated the process of intergenerational social mobility among men. Those researchers defined social inequality as socioeconomic status based on the economic and prestige rewards accorded to achieved occupational positions in American society. The strong parallel between this model of inequality and the neoclassical model of human capital (see Becker 1964) is well established (Wright 1978).

The most prominent distributional theories of inequality, however, are founded on macrosocial views of the division of labor, the rationalization of authority, and the distribution of social and economic rewards in industrial societies. Weber's ([1947] 1978) theory of economic organization proposed that capitalist systems of property, power, and prestige developed out of the conjunction of changing systems of economic exchange (money economies) and accounting (double-entry bookkeeping) with rationalized systems of social con-

trol (rational-legal authority). Thus, social inequality in industrial society developed along economic and political dimensions to produce the multidimensional bases of inequality: class, status, and party. Lenski's (1966) comparative study of the evolution of inequality attempted to test Weber's rationalization thesis that inequality evolves necessarily (functionally) with increasing differentiation in the direction of systems of privilege based on rational authority and away from socially illegitimate systems of force or economic dominance.

Accordingly, distributional inequality can be concerned with more than the single dimension of individual socioeconomic outcomes. It also addresses macrosocial patterns of inequality (Eisenstadt 1971). According to Blau (1977), the parameters of social structure include inequality and heterogeneity—or graduated and nominal dimensions, respectively—which intersect to constrain and differentiate individuals' opportunities as well as their motivations and outcomes. The intersection of graduated and nominal parameters creates diverse systems or populations with differing distributional properties that cannot be reduced to an original individual source. Blau's distributional theory is "macrosocial in the sense that the 'cases' are populations or communities and the 'variables' measure some aspect (a rate or a distributional property) of these populations" (Skvoretz and Fararo 1986, p. 30). Following this approach, indicators of inequality can be defined in terms such as Lorenz curves (e.g., Gini indices, Theil coefficients), social welfare functions, or similar distributional properties (see Allison 1978; Wolfson 1997).

The emergence of the new global economy over the last two decades of the twentieth century has been associated with what has been characterized as a "surge" in wage and household income inequality (Gottschalk and Smeeding 1997) and a "winner-take-all income market" (Levy 1998) in advanced industrial countries. By and large, the growth of very high incomes in some sectors and the stagnation of wages in selected labor markets have produced a widening distribution of income. Distributional measures of economic inequality such as Gini and Theil coefficients reveal growing inequality among employed workers across advanced industrial societies with some of the highest inequality observed in the United States.

Economic inequality also may intersect with the nominal category of race, for example, and produce more diverse outcomes than traditional functional or neoclassical economic theories would predict. Examinations of patterns of interracial/interethnic marriage, for example, indicate that the association between occupational achievement and race is mediated by the extent of interracial/interethnic marriage in a community (see Blum 1984; Blau et al. 1982). This treatment of inequality, which is based on notions of dispersion and association, departs from the simple reduction of unequal outcomes to individual attributes and embeds the process in extended distributional contexts.

Other distributional approaches introduce constructs to explain inequality at levels above individual attributes, although individuals usually remain the units of analysis. Spatial and temporal contexts, for example, define and constrain distributions of individual outcomes. The examination of occupational mobility within organizational or labor-market contexts attempts to nest the process of inequality in the workplace within organizational and occupational boundaries. The availability of occupational positions within a system is seen as being independent of the motivations and other attributes of workers. White's (1970) influential notion of "vacancy chains" exemplifies this approach with its argument that job vacancies produce opportunity structures for individual mobility and define the mobility chances, and thus distributional outcomes, of individuals. Vacancy-chain models have been particularly useful for examining closed opportunity systems, such as internal labor markets (Sorensen 1977).

Distributions of individuals in systems of inequality also are influenced by temporal factors. Merton's (1968) provocative discussion of the "Matthew effect" in scientific career systems argues that over time, initial inequalities in a system bias distributional outcomes in favor of initial advantage. Formal extensions and applications of Merton's notion of accumulative advantage have been applied across contexts (Cole and Cole 1973; Allison and Stewart 1974) to establish patterns of temporal regulation of distributional outcomes over and above the attributes of individuals over time. The cumulative advantage hypothesis has received considerable attention in research on the relationship between age and inequality within cohorts (Dannefer

1987; O'Rand and Henretta 1999). Succeeding cohorts of U.S. populations display growing inequality across the age span, with higher coefficients of inequality at older ages within cohorts and (in recent decades) increased inequality among the aged in successive cohorts (Crystal 1995).

### SOCIAL INEQUALITY AS A VERTICAL CLASSIFICATION OR BOUNDED CATEGORY

When social inequality is conceptualized as a vertical classification system, it is treated as an oppositional phenomenon. Here the approach is to define inequality in terms of "the relative position in a matrix of oppositions" (Schwartz 1981, p. 94) of social categories that determine relations of dominance, such as class, race, and gender. Vertical classifications grow out of antagonistic and contradictory interests in the relations of "objective" positions in the social division of labor, not out of the dispersed motivations and interests of individuals. Dominance and subordination emerge from the objective opposition of social categories. Dichotomous, binary, and polar conceptions of inequality (e.g., ruler-ruled, rich-poor, white-black, masculine-feminine) generally are informed by an oppositional framework. Some researchers have argued that this approach to inequality may be the most ancient in human social consciousness (Ossowski 1963; Schwartz 1981).

Class theories that follow Marxian frameworks dominate this approach (Braverman 1974; Wright 1985). Marx's theory of class proposes that class relations in capitalist systems are inevitably in conflict. Since all value is ultimately produced by labor, all (capitalist) profit must be at the expense of labor. The objective positions of the owning class (bourgeoisie) and the laboring class (proletariat) therefore are necessarily antagonistic. Advanced capitalist systems sustain the exploitation of labor through rationalized job-definition systems and the degradation of work (Braverman 1974). Wright (1978) has argued, furthermore, that in advanced capitalist societies, the elaborate differentiation of functions originally embodied in entrepreneurial capitalism into many different categories has not overcome the fundamental oppositional inequality of its origins; contradictory class positions continue to exist as a result of the underlying structure of capitalist relations.

Oppositional frameworks lend themselves to the examination of classlike relations such as those observable in race- and gender-centered systems of inequality. Oppositional approaches to the examination of race inequality can be traced to Myrdal's (1944) pioneering analysis of racial exploitation in the U.S. context. These approaches argue that race is an invariant principle of vertical classification that is masked by ideologies of economic progress and attainment (Pinkney 1984). Debates regarding the inevitability of racial opposition as the basis of inequality center on the substitutability of race and class as categories in the recent history of U.S. inequality. Wilson (1980) has proposed the controversial argument that class inequality has superseded race inequality as the basis of cross-race differences in economic and social outcomes.

Theories of gender inequality extend back to Mill's libertarian essay on the subjection of women (Mill 1859) and Engels's Marxian analysis two decades later (Engels [1884] 1942) of the relationship between private property and the stratification of family (gender) roles. However, contemporary feminist theories provide the strongest argument for gender inequality as an oppositional, vertical classification system. The sex/gender system, it is argued, subordinates women in patriarchal relations that exist over and above class relations (Jaggar 1984), since male dominance over women's productive and reproductive roles predates the emergence of capitalism (Harding 1983). This system of inequality leads inevitably to a conflict of interests and to the emergence of competing ideologies.

Since the notion of dominance is central to vertical-classification approaches to inequality, these approaches are readily applied to the analysis of large-scale systems of inequality, such as the state (Skocpol 1979) and the modern world system (Wallerstein 1974). Mechanisms of domination extend beyond class (or classlike) interests and are observable in the historical relations of nation-states (Reddy 1987) and multistate sectors of the modern world system (Wallerstein 1974). Asymmetrical relations of exchange and dependence between states and geopolitical state sectors create relations of dominance, which define global inequalities. Those inequalities can be formulated as distributional phenomena by following a functional framework; however, the historical analysis

of dominance systems lends itself more readily to oppositional analysis. The classification of the world system into core and periphery sectors that resulted from historically contingent factors introduces notions of centrality and dominance that suggest more than an underlying distribution of resources (Wallerstein 1974).

Tilly's (1998) statement on "durable inequalities" argues that persistent inequalities based on exploitation, opportunity hoarding, adaptation, and emulation largely take the form of bounded (usually dichotomous) categories (male-female, slave-owner, citizen-foreigner, white-black, etc.) that are resilient and readily generalizable across time and social systems. Relationships of inequality persist because participants in paired categories adapt to and participate in the perpetuation of those arrangements.

### SOCIAL INEQUALITY AS HIERARCHICAL RELATIONS

When social inequality is conceptualized as hierarchical relations, it is treated as a system of interactions or interdependencies characterized by relative symmetry (equality) and asymmetry (inequality) among relations. Here the approach usually is to define the form of social relations rather than the attributes of individuals in those relations and to account for patterns of unequal relations without referring to oppositions. Inequality or dominance stems from positions in hierarchical relations, not from the a priori possession or control of resources or power by individuals, groups, or categories (Marsden 1983). This relational approach to inequality can be traced to Simmel (1896), whose studies of the structures of superordination-subordination by persons, groups, and principles continue to inform research on hierarchical relations and social networks in modern life (Coleman 1982).

Because social relationships have formal properties such as connectedness, transitivity, reciprocity, and multiplexity, they are measurable units of analysis in the study of social inequality within populations at all levels, from siblings to communities to transnational trading systems (Lin and Marsden 1982). These social units make up complex configurations of social relations within which distinctive positions of relative equivalence or cen-

trality can be revealed (Burt 1982). Thus, in their study of coalitions and elite structures in the German community of Altneustadt, Laumann and Pappi (1976) determined the relational bases of influence between natives and newcomers by using network techniques that emphasized associational patterns rather than personal attributes. Patterns of social distance and connectedness among corporate actors, not the preexisting distribution of resources, defined the influence process in that community.

A study by Granovetter (1974) of the job-search process clearly demonstrates the relative utility of relational over distributional approaches to inequality. Granovetter demonstrates that weak ties, rather than strong ties, in a community prevail in a successful job search. The "strength of weak ties" hypothesis (related to Simmel's *tertius gaudens*, or the third who enjoys) provides the counterintuitive argument that weaker (secondary) social contacts increase individuals' access to jobs more than stronger (primary) ties do. These ties operate independently of the attributes of individual job seekers.

The "strength of weak ties" phenomenon can be extended beyond the job-search process to examine structures of relational inequality in different contexts. Studies of interlocking directorates and informational brokerage systems, for example, demonstrate that loosely coupled relational systems of different forms produce different systems of social inequality (Burt 1982). The network of ties constitutes a social-constraint context within which actors are "captured." Burt's (1983) study of corporate philanthropy as a cooptive relation is a specific example of the relational bases of inequality in a market context. Using Internal Revenue Service data on firm expenditures for advertising and philanthropy, Burt demonstrates that firm philanthropy co-opts the household sector by legitimizing the firm to the public as a protector and by improving the ability of specific classes to purchase the firm's products (Burt 1983, p. 424). The strength of this approach is that advertising, which is more blatantly co-optive, does not escape public suspicion, whereas philanthropy does so more easily. Firms in an economic sector perform unequally as a result of their relative co-optive relations with the public, and the public has a co-optive relationship as consumers in that context.

Finally, it should be mentioned that despite the rationale provided above for the bulk of sociological research on relational inequality, the relational approach has been used to examine the importance of individual resources for social inequality. Indeed, early experimental efforts to study small group processes of inequality demonstrated that both individual resources and social relations can create systems of inequality, whether measured as leadership processes or as communication networks (Thibaut and Kelley 1959). More recently, studies of what Burt (1982) has termed "ego-centered" networks examine network position itself as an individual resource with implications for social inequality.

### APPROACHES TO SOCIAL INEQUALITY

The three major approaches to the study of social inequality outlined above have different implications for theory as well as for method. The distributional approach that examines social inequality as a graduated dimension depends primarily on sample data and can be directed toward individual as well as structural explanations of inequality. The oppositional approach to vertical classifications and bounded categories may use sample data but has tended to adopt historical and qualitative approaches to study the institutionalization of dominance in various forms, such as class, race, and gender, as well as other forms of domination/subordination. The relational approach, which provides a direct method for examining the social context of inequality, may use sample or case data to map the configurations of the relations of inequality with implications for explanation at both the individual and structural levels.

(SEE ALSO: *Equality of Opportunity*; *Social Stratification*)

#### REFERENCES

- Allison, Paul D. 1978 "Measures of Inequality." *American Sociological Review* 43:865-880.
- , and John A. Stewart 1974 "Productivity Differences among Scientists: Evidence for Accumulative Advantage." *American Sociological Review* 99:596-606.
- Becker, Gary S. 1964 *Human Capital*. Chicago: University of Chicago Press.

- Blau, Peter M. 1977 *Inequality and Heterogeneity*. New York: Free Press.
- , Terry C. Blum, and Joseph E. Schwartz 1982 "Heterogeneity and Inter marriage." *American Sociological Review* 47:45-62.
- , and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Blum, Terry C. 1984 "Racial Inequality and Salience: An Examination of Blau's Theory of Social Structure." *Social Forces* 62:607-617.
- Braverman, Harry 1974 "Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century." *Monthly Review* 26:1-134.
- Burt, Ronald S. 1982 *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. New York: Academic Press.
- , 1983 "Corporate Philanthropy as a Cooptive Relation." *Social Forces* 62:419-449.
- Cole, Jonathan R., and Stephen Cole 1973 *Social Stratification in Science*. Chicago: University of Chicago Press.
- Coleman, James S. 1982 *The Asymmetric Society*. Syracuse, N.Y.: Syracuse University Press.
- Crystal, Stephen 1995 "Economic Status of the Elderly." In Robert H. Binstock and Linda K. George, eds., *Handbook of Aging and the Social Sciences*, 4th ed. New York: Academic Press.
- Dannefer, Dale 1987. "Aging As Intracohort Differentiation: Accentuation, The Matthew Effect, and the Life Course." *Sociological Forum* 2: 211-236.
- Davis, Kingsley, and Wilbert E. Moore 1945 "Some Principles of Stratification." *American Sociological Review* 10:242-249.
- Eisenstadt, S. M. 1971 *Social Differentiation and Stratification*. Glencoe, Ill.: Scott Foresman.
- Engels, Frederick (1884) 1942 *The Origin of the Family, Private Property and the State*. New York: International Publishers.
- Granovetter, Mark 1974 *Getting a Job*. Cambridge, Mass.: Harvard University Press.
- , and Charles Tilly 1988 "Inequality and Labor Processes." In Neil J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Gottschalk, Peter, and Timothy Smeeding 1997 "Cross-National Comparisons of Earnings and Income Inequality." *Journal of Economic Literature* 35:633-687.
- Harding, Sandra C. 1983 *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*. Dordrecht, Holland: D. Reidel.

- Jaggar, Allison M. 1984 *Feminist Frameworks: Alternative Theoretical Accounts of the Relations between Men and Women*, 2d ed. New York: Academic Press.
- Laumann, Edward O., and Franz U. Poppi 1976 *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: Academic Press.
- Lenski, Gerhard E. 1966 *Power and Privilege*. New York: McGraw-Hill.
- Levy, Frank 1998 *The New Dollars and Dreams: American Incomes and Economic Change*. New York: Russell Sage.
- Lin, Nan, and Peter V. Marsden, eds. 1982 *Social Structure and Network Analysis*. Beverly Hills, Calif.: Sage
- Marsden, Peter V. 1983 "Restricted Access in Networks and Models of Power." *American Journal of Sociology* 88:686–717.
- Marx, Karl (1859) 1976–1978 *Capital: A Critique of Political Economy*. Harmondsworth, UK: Penguin.
- Mayhew, Bruce H. 1973 "System Size and Ruling Elites." *American Sociological Review* 38:468–475.
- , and Paul T. Schollaert 1980 "The Concentration of Wealth: A Sociological Model." *Sociological Focus* 13:1–35.
- Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159:56–63.
- Mill, John Stuart (1859) 1970 *The Subjection of Women*. New York: Source Book Press.
- Mosca, G. 1939 *The Ruling Class*. New York: McGraw-Hill.
- Myrdal, Gunnar 1944 *An American Dilemma: The Negro Problem and Modern Democracy*. New York: Harper and Brothers.
- O'Rand, Angela M., and John C. Henretta 1999 *Age and Inequality: Diverse Pathways through Later Life*. Boulder, Colo.: Westview.
- Ossowski, Stanislaw 1963 *Class Structure in the Social Consciousness*. New York: Free Press.
- Pareto, Vifredo (1897) 1980 *Compendium of General Sociology*. Minneapolis: University of Minnesota Press.
- Pinkney, Alphonso 1984 *The Myth of Black Progress*. New York: Cambridge University Press.
- Reddy, William 1987 *Money and Liberty in Modern Europe*. Cambridge, UK: Cambridge University Press.
- Schwartz, Barry 1981 *Vertical Classification*. Chicago: University of Chicago Press.
- Simmel, Georg 1896 "Superiority and Subordination as Subject Matters of Sociology," trans. A. Small. *American Journal of Sociology* 2:167–189, 392–415.
- Skocpol, Theda 1979 *States of Social Revolutions*. New York: Cambridge University Press.
- Skvoretz, John, and Thomas J. Fararo 1986 "Inequality and Association: A Biased Net Theory." *Current Perspectives in Social Theory* 7:29–50.
- Sorensen, Aage 1977 "The Structure of Inequality and the Process of Attainment." *American Sociological Review* 42:965–978.
- Sorokin, Pitirim A. 1941 *Social and Cultural Dynamics*. New York: Bedminster.
- Thibaut, John W., and Harold H. Kelley 1959 *The Social Psychology of Groups*. New York: Wiley.
- Tilly, Charles 1998 *Durable Inequality*. Berkeley: University of California Press.
- Wallerstein, Immanuel Maurice 1974 *The Modern World System*. New York: Academic Press.
- Weber, Max (1947) 1978 *Economy and Society*, vols. I and II. Berkeley: University of California Press.
- White, Harrison C. 1970 *Chains of Opportunity*. Cambridge, Mass.: Harvard University Press.
- Wilson, William J. 1980 *The Declining Significance of Race: Blacks and Changing American Institutions*. Chicago: University of Chicago Press.
- Wolfson, Michael C. 1997 "Divergent Inequalities: Theory and Empirical Results." *Review of Income and Wealth* 43:401–421.
- Wright, Erik Olin 1978 *Class, Crisis, and the State*. London: NLB.
- , 1985 *Classes*. London: Verso.

ANGELA M. O'RAND

## SOCIAL JUSTICE

Justice is a basic element of social life. It is a central moral standard in human affairs that involves the necessity of "assuring that each person receives what she or he is due" (Cohen 1986, p. 1). Distributive justice is in the eye of the beholder, and debate usually surrounds the questions of what each person is due and what principles and procedures should be used to decide this. These differences not only occur among persons and social categories within a society but also vary across time and across cultures. A range of competing principles—rights or entitlements, equality of outcomes, equality of opportunity, equity or proportionality of rewards, and the satisfaction of basic needs—are prevalent standards of justice in most realms of



social existence. These principles compete for recognition and application in human affairs, and there is considerable interest in this subject among philosophers, social and behavioral scientists, and others; substantial effort has been invested in the understanding of justice in human life.

Questions of justice, or fairness, arise in virtually all aspects of social life, and the topic of social justice covers a vast array of subjects. The social goods that are of concern in questions of distributive justice include a wide array of things that people want, usually referred to as *primary goods*, including basic freedoms, political enfranchisement, power, authority, status, income and wealth, education and employment opportunities, housing, and health care. In most discussions, it is assumed that those things are scarce, although in some cases that is clearly not always the case; for example, there should be an unlimited supply of basic freedoms in a well-ordered democracy. However, even if social goods are abundant, they have inherent satisfaction value—they are things that bring both extrinsic and intrinsic satisfaction to the individual—and their distribution is governed in part by principles of justice.

A major form of justice that concerns sociologists as well as legal scholars is *criminal or legal justice*, but justice issues pervade many other types of social relationships as well. Justice issues arise very often when inequalities of outcomes exist, but equalities of outcomes also raise questions in regard to justice. In *Nicomachean Ethics*, Aristotle (1953) stated: “For if the persons are not equal, they will not have equal shares; it is when equals possess or are allotted unequal shares, or persons not equal, equal shares that quarrels and complaints arise.” Thus, in certain circumstances, equality of outcomes may be perceived as unjust while inequality of results may be seen as perfectly just.

The experimental literature in social psychology indicates that when persons perceive “inequitable” inequalities, they frequently experience cognitive tensions and a drive to reduce those tensions by changing their judgments about relative investments and contributions or changing their values in regard to the importance of reward-relevant criteria. Research also shows that when experimental subjects in task-oriented settings have well-defined expectations linked to objective indica-

tors and contributions and investments, they find reward inequalities more acceptable (Brickman 1977; Cook 1975). Inequalities in the distribution of rewards also result from power differentials, and those in disadvantaged positions are more likely to view such inequalities as unfair (Cook and Hegtvedt 1986; Molm 1991).

There are several overlapping spheres of equality/inequality in which justice issues are especially important. These spheres concern the *legal, political, economic, and social* realms of existence and cover a broad range of human social behavior. They can be summarized as follows: (1) *legal justice*: the application of laws and procedures to individuals and organizations through a system of rules laid down or established, whether by custom or through the will of the state, for which penalties exist for disobedience, (2) *political justice*: other aspects of social life in which issues of dependence/independence arise from interdependence and from the extent of power and influence held by actors and other parties to the relationship, (3) *economic justice*: the distribution of the material outcomes of existence, where the economic well-being of parties to the relationship is at issue, including access to basic needs and shelter, and (4) *social justice*: the realm of status, respect, and the sense of worth given and received in social interaction or in relation to society.

Because questions of justice in society are so pervasive, most social sciences claim to understand the ways in which human societies deal with them. Thus, the literature on justice is massive and is perceived differently from a variety of perspectives. Extensive scholarship exists in philosophy (Buchanan and Mathieu 1986), anthropology (Nader and Sursock 1986), economics (Boulding 1981; Solo and Anderson 1981; Worland 1986), psychology (Deutsch 1975, 1986; Folger 1984; Furby 1986; Greenberg and Cohen 1982; Mukula 1980; Messick and Cook 1983), and political science (Barry 1981; DiQuattro 1986; Elster 1989; Hochschild 1981; Rae 1981). Justice is also a prominent theme in many traditions within sociology (Alwin 1987; Hamilton and Rauma 1995; Hegtvedt and Markovsky 1995; Kluegel et al. 1995; Markovsky 1985; Rytina 1986; Jasso 1980; Jasso and Wegener 1997), but the issues of justice have been studied primarily by those working in the tradition of social psychology; this article reflects that emphasis.

### PERCEIVING JUSTICE

As was noted above, justice sentiments occur with respect to what each person receives relative to what he or she is expected to receive in regard to an important social good in which expectations may be governed by the application of a principle of justice. People's expectations and perceptions of justice focus on a wide range of phenomena, such as sentiments about the fairness of social exchange and contracts, fairness in interpersonal relationships, and the treatment of themselves and others by a social group or by society as a whole. Justice concerns thus represent a ubiquitous aspect of social life at many different levels and in many different spheres.

Expectations are formed not only by reward recipients but by others as well, including both people who are involved in the relationship and observers. Justice sentiments thus derive from comparisons of what is received with what one believes should be received, that is, a comparison of the real with the ideal in a particular context. Those evaluations of differences between these two entities or quantities engage human faculties of perception, cognition, and emotion. Even if one knows little else about justice evaluations, one knows that they are subjective, and justice almost always refers to "justice in the eyes of the observer" (Walster et al. 1973). Of course, these facts complicate the application of justice principles because, even setting aside the issue of which principle of justice to invoke in a particular situation, actors may not agree on what is real, that is, what the true outcomes are. However, to paraphrase a theme in interactionist sociology (from W. I. Thomas), *what is real in the perceptions of humans is real in its consequences*. Thus, if individuals perceive the social mechanisms for allocating scarce social resources and/or rewards as just, presumably the resulting distribution of outcomes also will be perceived as just.

From a social psychological viewpoint, then, a focus on justice is a focus on *beliefs* about inequality and *perceptions* of justice. Justice perceptions are pervasive in interpersonal interaction and exchange as well as in the nature of the relationship of the individual to the larger social collective in the macro-social realm. One way to simplify the vastness of this sociological terrain is to separate ques-

tions of justice and its evaluation posed at the micro-social level from those which occur at the macro level of society or the state (Brickman et al. 1981; Markovsky 1985; Hegtvedt and Markovsky 1995). Micro justice concerns justice evaluations at the individual level with regard to a person's immediate circumstances. Macro justice involves the evaluation of justice above the individual level, for groups or for society as a whole. Clearly, justice evaluations take place across several domains at different levels and focus on the abstract or the concrete. Within each level, one also may distinguish between beliefs about inequality and perceptions of justice at an abstract level, such as how things should be distributed in a fair world, and the evaluation of justice in the "real world," that is, assessments of how fair things are in reality (Kluegel and Smith 1981). In most cases, these various levels and degrees of abstraction cannot be easily separated, for example, in the evaluation of the fairness of child support payments (Schaeffer 1990), the fairness of child custody resolutions (Elster 1989), and the fairness of economic rewards (Sennett and Cobb 1972), but it is useful to draw many of these conceptual distinctions for research purposes.

### PHILOSOPHICAL ROOTS

The historical roots of Western conceptions of justice lie in classical philosophy, the Judeo-Christian religious traditions, and the theoretical and ideological underpinnings of legal, economic, and political arrangements. Aristotle's *Nicomachean Ethics* (Book V) provides the classical formulation of the problem of justice. Aristotle's work sought to clarify principles of distributive and retributive justice and formulate the rules for the regulation of social exchange. The Aristotelian logic of justice in market relationships stressed the *proportionality* principle with regard to expected reward given considerations of merit and indicated that most contracts and purchases comply with the going rate of exchange (Cohen and Greenberg 1982, p. 4). Of course, it is not always easy for parties to agree on the "going rate." Worland suggests a crucial dilemma is revealed in Aristotle's work involving competition between justice principles:

*It becomes clear that the reference to two different kinds of justice and two different*

*rules of proportions poses a crucial dilemma. If commodities sell at their fair price—or at a price that reflects the “fair” rate of exchange—then how can society guarantee that exchange at such prices will also provide society’s participants with an income proportionate to their relative “merit” or their standing in the community? How is the rule requiring distribution of common goods in proportion to “merit” to be reconciled with the rule requiring “reciprocal proportionate equality” in the contractual, private exchange of commodities? (1986, p. 48)*

Eighteenth- and nineteenth-century philosophers and social theorists resolved this dilemma in a number of competing ways. Marx’s labor theory of value (see Buchanan and Mathieu 1986, p. 12) suggested that the need for principles of justice provided evidence that social institutions would be restructured, abolishing the market system and private goods. Although the notion of communal ownership of the means of production can be traced back to early Greek philosophy and before, this has not been viewed as the most adequate solution to Aristotle’s dilemma. Indeed, communism poses its own dilemmas, and the recent unpopularity of socialism as a political and economic system may provide some evidence of this (see Alwin et al. 1996, p. 124). Markets exist even under state socialism, and from the point of view of studying justice, the existence of a market for social exchange, as well as its comprehension and acceptance, seems to represent an important component in understanding the fairness of social interaction (Thiabaut and Kelley 1959; Lane 1986, p. 384).

Adam Smith, who is considered the philosophical father of modern capitalism, posed the Aristotelian dilemma differently. He observed that the modern economic system had become specialized in the sense that the production, distribution, and exchange of social goods had evolved to a stage in which economic activity was highly differentiated as an institution (Worland 1986, p. 50). Smith believed the “natural” rules by which markets developed and the moral issues of justice were resolved could be determined empirically. What is just, then, was clearly a question of what individual actors considered just not only at the subjective level but also at the macro-social level. The opera-

tion of principles of supply and demand to set a going rate was seen as providing the answer to the moral question of justice. Thus, Smith propounded an early example of a principle that later was articulated in the social psychology literature: the principle that “what is” determines “what ought to be” (see Homans [1961] 1974, p. 250; Heider 1958, p. 235). This, however, falls short of resolving the Aristotelian dilemma because it gives too great a role to existential factors in defining justice.

As Worland (1986, p. 57) suggests, Marx’s theory of surplus value, in which Marx identified the real sources of profit and nonwage income as exploitation of working people by the capitalist class, did much to clarify what can now be seen as a condemnation of capitalist economies in Aristotelian terms. The twentieth-century response to Marx’s critique of capitalism in economics is known as marginal utility theory, which explains the issue of injustice in terms of market imperfections and the existence of monopolistic forms of capitalism instead of real free enterprise. According to Worland (1986, p. 81), the neoclassical response is able only to isolate and clarify “the rules that a market society needs in order to comply with the Aristotelian moral imperative.” It does not answer the deeper question of moral psychology concerning “whether such a society would be able to achieve the social consensus necessary for the practical implementation of the rules.”

The perspective offered by neoclassical economics on moral issues of justice often is seen more as a justification for social inequalities than as a “natural law” of justice. It is reminiscent of Cohen and Greenberg’s (1982) suggestion that each social and economic system evolves its own unique concept of justice. For neoclassical economics, the idea that there may be merit in nonproductive activities or considerations is foreign. By contrast, contemporary political philosophers have attempted to address the broader question of merit. Rawls (1971) evaluates the Aristotelian imperative from a nonmaterial, rational perspective. Rawls’s principles of justice are as follows (quoted from Buchanan and Mathieu 1986, p. 27):

1. *The principle of greatest equal liberty: Each person is to have an equal right to the most extensive system of equal basic liberties compatible with a similar system of liberty for all.*

2. *The principle of equality of fair opportunity: Offices and positions are to be open to all under conditions of equality of fair opportunity—persons with similar abilities and skills are to have equal access to offices and positions.*
3. *The difference principle: Social and economic institutions are to be arranged so as to benefit maximally the worst off.*

According to Rawls (1971), these principles are ordered in terms of their primacy; that is, if principles conflict, the one listed first takes precedence. These principles obviously refer not only to social contracts and exchanges but also to the basic structure of society, including legal, political, economic, and social institutions. In most societies, that includes charters and constitutions, the means of production, competitive markets, the family, the legal system of laws and procedures, and so forth. The basic structure of society is said to specify how and by what principles society should distribute primary goods: basic liberties, powers, authority, opportunity, income, and wealth. The principle of greatest equal liberty specifies Rawls's theory of how basic liberties are distributed, the principle of equality of fair opportunities regulates the distribution of life chances or prospects in the domains of power and authority, and the difference principle governs the distribution of income and wealth (see Buchanan and Mathieu 1986, p. 28).

It is beyond the scope of this article to review Rawls's theory in full or summarize the critical comment that followed, except to note that his theory is utopian in the sense that it describes an ideal society run on the basis of just principles. Rawls does not spell out how one moves from states of injustice such as those found in contemporary society to the ideal situation. Still, no single book has generated more discussion and critique than his *A Theory of Justice*. A prominent contemporary critique is Nozick's (1974) entitlement theory, which departs radically from Rawls by stressing a libertarian view that a person is entitled to the ownership of a social good if it was acquired through just principles. Walzer (1983) argues that it is not possible to write a general theory of justice in the abstract without attempting to assay the "substantive ways of life" of different cultures. To Walzer, "justice" is a relative term, and no abstract

theory of the ideal society can really work because justice does not exist in an abstract theoretical sense but only in terms of social meanings ordained by a particular way of life.

### DISTRIBUTIVE VERSUS PROCEDURAL JUSTICE

Distributive justice issues arise when one considers two sets of questions: Who gets what, and how? and Who should get what, and how? Some writers have suggested that the distinction between the concepts of *procedural justice* and *distributive justice* is critical to a complete understanding of the ways in which people evaluate justice (see Cohen 1986; Hegtveldt and Markovsky 1995; Thibaut et al. 1975; Tyler 1984, 1986). Procedural justice refers to the mechanisms or decision rules by which reward allocations of social goods are made, while distributive justice is concerned with the resulting allocation. These are two aspects of the same process and are clearly related, but conceptually they purport to refer to two distinct features of justice. Typically, distributive justice issues are thought of in terms of the comparison of the rewards received by a person or a group with a standard of fairness or deservedness, whereas procedural justice issues refer to the "mechanics" of the system that regulates the process of distribution.

It is useful to distinguish three components of the distributive justice process: (1) the principles for the allocation of goods, (2) the system that governs the application of those allocative principles, and (3) the resulting distribution. While separable in this sense, these are all components of what is best thought of in terms of distributive justice, which is the overriding concept. This view is in agreement with the work of Deutsch (1986, p. 35), who states that "procedural justice is a key aspect of distributive justice," not something that is necessarily separate or separable. Procedural justice matters come to the forefront, suggests Deutsch, and arouse complaints of injustice more often than do the principles of justice, primarily because justice principles are often taken for granted, whereas procedural matters are not.

From this formulation, one can see that these three components—principles, procedures, and distributive outcomes—are likely to be confounded and confused in social life. If there is consensus on evaluative principles and if a clear and just set of

procedures can be said to exist to implement those values, distributive justice presumably will follow. Procedural justice thus is an important component in the evaluation of distributive justice, and in the pure case the evaluation of distributive justice is not problematic. If, however, there is no consensus on the principles for allocating rewards in a social group or society or if there is a consensus but procedures are seen as ineffective or corrupted, distributive justice will be called into question. In such situations, the evaluation of distributive justice focuses on the evaluative principles that should be used, the application of the principles, or both. Therefore, the principles of allocation and the procedural aspects of allocation may be intrinsically inseparable, and it may not be clear whether unjust outcomes result from the “wrong” principle being used or from misapplication of the “right” principle.

In summary, while it seems useful to distinguish procedural justice from other components of the distributive process, it is not at all clear that procedural and distributive justice are really different forms of justice. Instead, they are different aspects of a common process. Clearly, distributive justice issues arise when persons perceive the allocative mechanisms to be unjust or perceive imperfections in the application of just mechanisms to real life. It can be seen, then, that overall evaluations of justice at the micro-social level or the macro-social level may be influenced by perceptions of the fairness of procedural or distributive justice issues. However, this distinction ultimately becomes the empirical issue of whether and under what sets of conditions distributive versus procedural injustice is perceived by the actors involved.

### EQUITY THEORY

Much current research on distributive justice traces its theoretical roots to Homans's (1976, p. 249) discussion of the rule of distributive justice, which hypothesizes that unless persons' perceived inputs (contributions, investments, resources, etc.) are equal, some inequality of outcomes or rewards can be expected, and that rewards generally are expected to be allocated in proportion to inputs (see also Heider 1958, p. 288). Responding to what they saw as the need for a general theory of social behavior, Walster and her colleagues (see Walster

et al. 1973, 1978; Berkowitz and Walster 1976) formulated equity theory to integrate the insights from a variety of social psychological theories, including reinforcement theory, cognitive consistency theory, psychoanalytic theory, and exchange theory (Walster et al. 1978, p. 2). Building on the work of Homans, Lerner, and others, equity theory as formulated by Walster et al. (1978, p. 6) contains four basic propositions:

*Proposition I:* Individuals will try to maximize their outcomes.

*Proposition IIA:* Groups can maximize their collective reward by evolving accepted systems for equitably apportioning resources among their members. Thus, groups will evolve such systems of equity and attempt to induce their members to accept and adhere to those systems.

*Proposition IIB:* Groups generally will reward members who treat others equitably and punish members who treat others inequitably.

*Proposition III:* When individuals find themselves participating in inequitable relationships, they will become distressed. The more inequitable the relationship, the more distress individuals will feel.

*Proposition IV:* Individuals who discover they are in an inequitable relationship will attempt to eliminate their distress by restoring equity. The greater the inequity that exists, the more distress they will feel and the harder they will try to restore equity.

To define justice, equity theory distinguishes *inputs* and *outcomes*, both of which are expressed in the same units, say, dollars, points, or another unit. Assuming that inputs are positive (for simplicity), the equity principle states that under conditions of justice, there is an equality of relative gains. For two actors (persons, groups, nations, etc.) engaged in social exchange, equity is said to exist when the ratio of profits (outcomes minus inputs) to inputs is the same for both actors involved in the exchange.

While it is clear that rational self-interest is a strong motive for behavior in many types of situations and may be a safe assumption in competitive situations, there are other motivations for behavior (Becker 1996; Elster 1984, 1989). This can be

seen as one of the unnecessarily restrictive assumptions of equity theory; the prevalence of altruism, cooperation, and other forms of prosocial behavior strongly questions this basic assumption. There is plenty of evidence that competitive behavior in mixed-motive games such as the Prisoner's Dilemma can evolve into stable cooperation under certain conditions, such as the anticipation of future interaction (Axelrod 1984). In relationships where intimacy and identification are present, the importance of self-interest as a motive for behavior is considerably lessened (Austin 1977). As was noted earlier, Rawls's (1971) theory of justice, for example, suggests that one principle that underlies the moral judgments made in human society is the principle that social and economic arrangements are established to maximally benefit those who are less advantaged with respect to desirable social goods. Thus, one can see that equity theory as stated by Walster et al. (1978) probably overstates the need to specify a type of utilitarian-based logic for defining justice.

Implicit in this theoretical statement and the experimental literature on which it is based are the assumptions of cognitive dissonance theory (Festinger 1959). Thus, an important way to resolve dissonance is to adjust one's beliefs to conform to reality. However, while the theory's allowance for changes in beliefs is an important feature, there are limits to what people will believe. Elster comments as follows on the concept of "belief adjustment":

*Dissonance reduction can also take the form of belief adjustment. Workers who take jobs in unsafe industries alter their estimated probabilities of accidents. As a result, when safety equipment becomes available, they may choose not to purchase it. Here, as in other cases, misformation of private beliefs (or preferences) creates a case for government intervention. . . . Belief-oriented dissonance reduction is a form of wishful thinking [but] . . . acting on beliefs formed in this way can be disastrous and is likely to force a change in beliefs. When action is not called for, the wishful beliefs can be more stable. The "just-world" theory, for instance, suggests that people adjust their beliefs about guilt and responsibility so as to preserve their belief that the world is fundamentally just. The best-known example is the "blame the victim" syndrome. . . . (Elster 1989, pp. 22–23)*

Presumably, many beliefs, including beliefs about inequality and justice, achieve a high degree of stability relatively early in adulthood and may not be subject to alteration (Alwin 1994). Expectations may not be altered easily if they are rooted in firmly held beliefs about the legitimacy of inequality.

### FROM EQUITY TO DISTRIBUTIVE JUSTICE

In the early 1970s, Deutsch (1975) began to question the proportionality, or equity, principle implicit in much theorizing about justice. He, among others, argued that equity is only one of several principles used to evaluate the justice of outcomes in social life. Principles of justice are used as a basis of "judging individual persons and in judging the basic structure of societies" (Cohen 1986, p. 1). At least five competing principles—equality, equality of opportunity, equity, rights or entitlement, and need—are applied routinely in most realms of social existence. Other principles, such as chance, are really devices for ensuring that a principle of justice—namely, equality of opportunity—is procedurally adhered to (see the discussion of lotteries in Elster 1989, pp. 62–122). While these principles may more or less exhaust the possible criteria for evaluating justice, perhaps the equity principle has been seen as most relevant in many cases. For example, in the work of Homans ([1961] 1974), the proportionality principle is given a prominent place in the discussion of distributive justice, and it is clearly present in many discussions of equity theory (see Adams 1965; Berkowitz and Walster 1976; Cook and Hegtvædt 1983).

However, with time, theoretical work and empirical research have cast the problem of equity into the broader framework of distributive justice, in which justice may be evaluated by one principle or a combination of many different principles. Contemporary interest in the study of distributive justice can be traced to the early articulation of relative deprivation theory, particularly in the work of Stouffer, Merton, and Homans (see Williams 1975). In the last several decades, issues of distributive justice have moved from a primary focus on the psychological and behavioral consequences of patterns of social reward distributions in small group settings (see reviews by Adams 1965; Berkowitz and Walster 1976; Crosby 1976; Cohen and Greenberg 1982; Hegtvædt and Markovsky 1995) to a more recent focus among sociologists on

studying principles of justice evaluation (Alves and Rossi 1978; Jasso and Rossi 1977; Jasso 1978, 1980; Jasso and Wegener 1997).

### Justice Evaluation and Social Comparison

Current theories of distributive justice implicitly or explicitly specify *social comparison* as the basis for justice evaluation and the reference standard persons use in evaluating the fairness of social rewards. The following discussion summarizes the author's theory of justice evaluation and social comparison (Alwin 1987). As in most contemporary theories, it is assumed that in evaluating the fairness of a particular reward allocation, persons compare themselves with others (see Berger et al. 1972; Pettigrew 1967; Williams 1975; Gartrell 1982). For example, Homans's rule of distributive justice often is stated as follows (see Walster et al. 1978):

$$\text{Justice} = [\text{P's Reward} / \text{P's Inputs}] - [\text{O's Reward} / \text{O's Inputs}] = 0 \quad (1)$$

Here P and O are two persons in a *local exchange* in which P is assumed to evaluate his or her comparison ratio or rewards to inputs against that of another individual, O. If P perceives that O's comparison ratio is equal to his or her own, then he or she will perceive a state of justice. If this is not the case, some degree of injustice is presumed to exist, and P will perceive that he or she is underrewarded or overrewarded relative to O.

The term "inputs" is used in equation (1) to represent the general reward-relevant characteristics of individuals that are involved in making assessments of the fairness or justice of rewards. These may be contributions, investments, resources, or global status characteristics (see Cook 1975; Cook and Yamagishi 1983). This formulation ignores the concept of costs and their effects on rewards and inputs. In other words, inputs and rewards are thought of as positive quantities. For a somewhat different formulation involving negative inputs and rewards, see Walster et al. (1978).

Several investigators (e.g., Blau 1971; Berger et al. 1972; Jasso 1978) have pointed out that in viewing this identity (equation [1]) objectively, it is impossible to determine which parties are over- or underrewarded when perfect justice does not prevail. Moreover, formulations of justice evaluation as local comparisons cannot cope with the possibility that from the perspective of a more general

referential standard, both P and O may be unjustly rewarded even though their comparison ratios may be equal. This view in no way denies that people make local comparisons of such ratios. The point is that when people make local comparisons, referential standards existing outside the local situation typically are invoked to evaluate fairness.

Berger et al. (1972, p. 122) criticize this formulation, arguing that distributive justice issues arise only in the presence of a stable frame of reference and that justice evaluations are inherently made on the basis of reference to generalized individuals rather than specific others. Using this observation as a basis for reconceptualizing the classical exchange-based conception of justice, Berger et al. (1972) formulate a "theory of status value" that formalizes the process by which persons evaluate the fairness of rewards. They formulate the process in terms of *referential standards*: frames of reference that contain existing information regarding the characteristics and rewards of generalized others. The referential structure formalized by Berger et al. (1972) consists of information about the relationship between levels of characteristics possessed by general classes of persons and the associated levels of social reward. According to this theory, through social exchange persons develop normative expectations about the reward levels typically associated with general classes of individuals, and when persons perceive their reward-relevant characteristics to be similar to those of a particular general class of individuals, they come to expect that their reward levels also are similar. As a consequence of these beliefs about "what is," normative expectations are formed about reward levels that persons can legitimately claim (Berger et al. 1972, p. 139). This conclusion is consistent with Homans's ([1961] 1974) and Heider's (1958) observations that the "ought" is determined in the long run by the "is" and with recent sociological theorizing that argues that social inequalities are often the major basis of their own legitimation (Sennett and Cobb 1972; Della Fave 1980; Stolte 1983).

The Berger et al. (1972) *status-value formulation*, however, is limited in its consideration of the process by which a person selects a referential comparison standard among the many possibilities. Referential comparisons may be based on relatively small groups of persons such as one's coworkers or large classes of persons such as occu-



pational categories. This may appear to be a flaw in the status-value theory; however, it also may be viewed as an asset in that it allows flexibility in specifying the role of various types of referential standards in justice evaluation processes.

**Referential Comparisons.** While near consensus exists in the social psychological literature that persons use referential comparisons to evaluate how satisfactory their outcomes are, Gartrell (1982) observes that little is known about the origin and visibility of comparative frames of reference. In his own research, for example, Gartrell finds that information on wage rates is often invisible, and such information for persons in other jobs frequently originates from concrete, personal references rather than from knowledge of rates for broad social categories. Moreover, the awareness of wage comparisons frequently appears to relate to somewhat idiosyncratic factors, such as informal social contacts (see Walster et al. 1978). This is consistent with other studies of relative deprivation and status comparison, in which persons are found to rely heavily on information from their own social circles (see Runciman 1966; Rainwater 1974; Coleman and Rainwater 1978). Thus, it may be difficult to specify the origin or basis of a person's referential comparison with objective accuracy. Some experimental research (Major and Forcey 1985) suggests that subjects are most interested in same-sex and same-job wage comparisons.

Moreover, the individual's subjective judgment regarding the "fairness" of a given reward outcome may be a more relevant concept. Jasso's (1978, 1980) theory of distributive justice introduces the term "just reward" to refer to the reward level individuals expect on the basis of referential comparisons conceived in general terms. Jasso formulates the status-value model of Berger et al. (1972) for justice evaluation as follows (1978, p. 1402):

$$\text{Justice} = \text{P's Actual Reward} - \text{P's Just Reward} = 0 \quad (2)$$

That is, for a person to determine the justice or fairness of his or her reward, the actual level of reward is simply compared with the reward expected on the basis of existential (or other) criteria.

This formulation is more general than the one given by Berger et al. (1972) because it permits a

wide range of reference group comparisons and because both existential and nonexistential criteria may be used in the calculus of the "just reward" (see Jasso 1978, 1980). As Blau (1971, pp. 58-59) points out in his criticism of Homans ([1961] 1974), "not all existing practices reflect justice; some are unjust by prevailing moral standards, and the fact that they are expected to continue to exist does not make them just." Thus, there are both existential and nonexistential standards of justice, and either or both may be combined to determine the "just reward."

This "comparison difference" formulation of distributive justice makes the nature of over- or underreward clear, unlike the formulation used by most equity theorists (e.g., Walster et al. 1978). Moreover, in contrast with the classical formulation given in equation (1), both quantities in the equation are expressed in the same units: units of reward. Further, this formulation (equation [2]) satisfies the notion that the individual's subjective judgment regarding the "expected" or "deserved" reward may be the most relevant concept. Finally, the distinction between kinds and degrees of injustice can be measured on a scale that starts at zero, where "perfect justice" occurs. Southwood (1978, p. 1157) has proposed a model very similar to Jasso's model involving what he calls "subtractive interaction" that is intended to estimate the effects of departures of actual reward from expected (or just) reward. This model simply involves estimating the effects of the quantity  $x_1 - x_2$ , where  $x_1$  represents the actual reward and  $x_2$  represents the expected reward.

**Reformulating the Justice Evaluation Model.**

Using Jasso's concept of the just reward, it is possible to reformulate the classical exchange-based conception in a way that permits the measurement of the direction and magnitude of departures from justice. First, it is necessary to recast Homan's ([1961] 1974) "rule of distributive justice" given in equation (1) as an equivalence of the ratio of P's and O's rewards to the ratio of their inputs as follows (see Patchen 1961; Adams 1965; Homans [1961] 1974, 1976):

$$\text{Justice} = \text{P's Actual Reward} - \text{P's Just Reward} = 0 \quad (3)$$

Here the comparison ratios are different from those in the classical statement (Walster et al.



1978). They now involve terms in common units: reward units on the one hand and units of input on the other hand. The present formulation of the classical model has two desirable properties: It is intuitively simpler to have the numerator and denominator of such ratios in the same units, and this formulation fits with the psychological mechanism often assumed in justice evaluation: Persons expect their inputs (contributions, investments, resources, or general status characteristics) to be in constant proportion to the rewards they associate with a standard of comparison.

If the Berger et al. (1972) theory of status value is correct is stating that a person, P, uses referential structures that specify levels of reward-relevant characteristics that are similar, indeed equivalent, to his or her own (see also Pettigrew 1967; Williams 1975), it is possible to equate P's and O's inputs on the right-hand side of equation (3), setting the second term on the right at unity, as follows:

$$\text{Justice} = [\text{P's Reward} / \text{O's Reward}] - 1 = 0 \quad (4)$$

If one generalizes the concept of "O's reward" to be the same as Jasso's "just reward," the justice evaluation process devolves to a comparison of P's actual and expected/just rewards:

$$\text{Justice} = [\text{P's Actual Reward} / \text{P's Expected Reward}] - 1 = 0 \quad (5)$$

An examination of this expression indicates that the classical formulation restated in this way permits the distinction between kinds and degrees of injustice measured on a scale that starts at zero, where perfect justice occurs. These units may conveniently be thought of as justice units because when the comparison exceeds zero, overreward occurs, and when it is less than zero, P is said to be underrewarded.

Note the convergence of the reformulation of the classical model given here with that proposed by Jasso (1978), which was derived empirically from the analysis of vignette data. The natural logarithm of equation (5), which is derived from classical exchange and status value theories, equals the formulation for justice evaluation proposed by Jasso. I have derived theoretically a principle of justice evaluation that is equivalent to Jasso's (1978, 1980) empirically derived "Universal Law of Jus-

tice Evaluation." This formulation can be used as a basis for defining departures from justice.

Jasso argues, however, that the simple ratio of actual to just rewards does not capture the justice evaluation phenomenon precisely: It does not account for the fact that positive departures from justice (overreward) are not equivalent to negative departures (underreward), and "this appears to violate the human experience that deficiency is felt to be more unjust than a comparable excess" (Jasso 1978, p. 1403). In other words, the injustice created by an actual reward above the just reward  $k$  is not equivalent to the injustice created by an underreward of the same magnitude. Jasso (1978, p. 1415) resolves the problem by proposing the natural logarithm of the comparison ratio (i.e., actual reward/just reward). Such a formulation assumes that an overreward of  $k$  times is equal in the magnitude of injustice to an underreward of  $1/k$  times. Using satisfaction with material well-being, Alwin (1987) empirically examined Jasso's hypothesis that the effects of underreward are more potent than the effects of overreward. While support was found for the importance of the sense of injustice in the prediction of material satisfaction, the hypothesis that the extent of satisfaction depends on measured departures from justice was not supported.

### BELIEFS ABOUT INEQUALITY

As was mentioned above, distributive justice issues arise in response to two sets of questions: (1) the realm of the ideal, that is, who should get what, and how, and (2) the real, or who gets what, and how. As was noted, these issues may be phrased with regard to the individual, the micro-justice level, or groups of individuals or the whole society (i.e., macro-justice). Regardless of the justice principles one espouses, behavior and sentiment are conditioned by how the system works or is perceived to work. One of the important aspirations of social science is the understanding of who gets what and how, and it was the major focus of several important works on social stratification and mobility in the 1960s and 1970s (e.g. Blau, and Duncan 1967; Hauser and Featherman 1977; Featherman and Hauser 1978; Jencks et al. 1972, 1979; Levy 1988; Moynihan 1968; Sewell and Hauser 1975; Wright and Perrone 1977). As a result of that

work, a great deal more is known about the ways in which race, gender, and class affect inequalities and the manner in which families and educational institutions operate to promote or deny access to opportunities for socioeconomic advancement. However, the study of social stratification and mobility is an arcane field of inquiry, and there is not always agreement among social scientists about the main factors that create socioeconomic inequalities (see Herrnstein and Murray 1994 and the debate it provoked, e.g., Fischer et al. 1996).

Sociologists often assume that there is a causal linkage between the structural conditions of society (e.g., the standard of living and economic inequality) and public beliefs and sentiments regarding the acceptability of those conditions, but little is known about the nature of that linkage. There is a growing body of empirical data regarding trends in income distribution and income inequality, much of which indicates growing economic inequality and hardship for segments of the population (e.g., Duncan and Rodgers 1991; Levy 1988; Levy and Murnane 1992; Thurow 1987), but much less is known about subjective interpretations of economic conditions. Thus, while social scientists attempt to understand the inner workings of the stratification system, there is much that remains to be understood, and it is important to realize that a person's beliefs about sources of inequality affect evaluations of justice as much as if not more than objective conditions do (see Kluegel and Smith 1981, 1986; Robinson and Bell 1978; Kluegel et al. 1995).

### SOCIAL JUSTICE RESEARCH

This article has argued that one can distinguish three components of the distributive justice process with regard to any primary good: (1) the principles for the allocation of goods, (2) the system that governs the application of those allocative principles, and (3) the resulting distribution. Justice sentiments derive from comparisons of what is received with what one believes should be received, that is, a comparison of the real with the ideal in a particular context.

According to Jasso and Wegener, empirical justice analysis has four major objectives:

*(i) to obtain numerical approximations of the quantities and relations identified by jus-*

*tice theory; (ii) to gauge the extent of interindividual and intergroup variation in the quantities and relations; (iii) to explain their etiology, including the effects of social structure and of the observer's position in the stratification structure; and (iv) to assess their behavioral and social consequences* (1997, p. 393).

Three fundamental quantities pertain to justice: the actual condition, the just condition, and the justice evaluation. While justice evaluations involve the comparison of the other two quantities, it is not clear that individuals actually quantify justice in the way theoretical formulations suggest, and it remains to be seen whether most people "calculate" more than a general "sense of justice" from this comparison.

To summarize current research on social justice, one would have to focus on a wide range of distributional issues with respect to the primary social goods of distributed in society, including basic freedoms, political rights, power, authority, status, income and wealth, education and employment opportunities, housing, health care, and the pursuit of happiness. Below, this article briefly mentions five areas in which a consideration of justice theory is relevant to sociological understanding: (1) income inequality and the welfare state, (2) discrimination and affirmative action, (3) gender, work, and comparable worth, (4) divorce, child custody, and child support, and (5) intergenerational relations.

#### **Income Inequality and the Welfare State.**

One of the central preoccupations of sociologists who study distributive justice has been the economic realm, with an explicit focus on wages or earnings (e.g. Alves and Rossi 1978; Gartrell 1982; Gartrell and Paille 1997; Jasso 1978, 1999; Jasso and Rossi 1977; Patchen 1961; Randall and Mueller 1995; Robinson and Bell 1978). Research in the United States shows convincingly that individualistic attributions for poverty and wealth predominate (Kluegel and Smith 1986). Such attributions are found across the spectrum of socioeconomic positions, and among lower socioeconomic status (SES) groups these beliefs are held concurrently with structural explanations. At the same time, Americans show considerable antagonism toward any form of systemwide redistribution of income

beyond current welfare assistance to children, the disabled, and the indigent and current forms of social security for senior members of the population.

Whereas Americans hold equality as the standard of justice in the political realm, inequality is the standard in the economic realm. A recent multinational comparison of a U.S. sample with comparable data from other Western nations showed that in assessing the deservingness of their own earnings, “what ought to be” is strongly linked to “what is” (Alwin et al. 1996). Thus, existential considerations play a strong role in the development of judgments about levels of deserved income, although there was considerable variation in the magnitude of those linkages. The weakest effects, as predicted on the basis of perceived system legitimacy, were in the postcommunist economies of eastern and central Europe and the former republics of the Soviet Union. The strongest effects were in the Western capitalist democracies. Indeed, the linkage between job desserts and job income was so strong in those countries (Germany, Great Britain, the Netherlands, the United States, and Japan) that it seemed hardly possible that any factors other than current income levels could contribute to variation in perceptions of justice. By contrast, the level of perceived family need played a much stronger role in evaluations of the justice of earnings in the eastern European countries than it did in the West (see Alwin et al. 1996, pp. 123–128).

**Discrimination and Affirmative Action.** Discrimination by dominant groups against ethnic minorities and women has been a significant concern of those interested in equality and justice. This set of issues can be addressed easily within the framework of social justice as an instance of the lack of congruence between justice principles such as equality of opportunity and the actual workings of society. Although considerable progress has been made in establishing constitutional prohibitions against discrimination on the grounds of race, ethnic origin, and sex in employment practices, education, and public accommodations, reality has lagged behind those ideals. Substantial inequalities among racial groups persist despite substantial opinion that racial discrimination is no longer a problem in American society (Blauner 1989). Considerable research has focused on the consequences of discrimination.

The concept of *affirmative action* has been used in the United States and other countries to refer to social policies that go beyond prohibitions against discriminatory practices that deprive minorities of their rights and aim social policy toward remedying the effects of past discrimination. Affirmative action represents an effort to restore equity to social groups, rather than to individuals, by targeting women and minorities for educational opportunities, jobs, promotion, government contracts, and other arenas where past discrimination has been documented. Affirmative action policies have been controversial because they appear to represent a form of *reverse discrimination* inasmuch as they violate the principle of equal opportunity by giving preferential treatment on the basis of race and national origin. These policies have faced a number of legal challenges that are likely to continue as long as they are perceived to be unjust by some members of society. Regardless of how one views these policies, they have placed increased numbers of women and minorities in good jobs and selective educational institutions, but they may have increased tensions over these matters.

**Gender, Work, and Comparable Worth.** An important application of the social justice framework has been the examination of equity in the job rewards of men and women. In the United States, as in other countries, there are legislative guarantees to the right of equal pay for equal work, such as Title VII of the Civil Rights Act of 1964. The concept of *comparable worth* was designed to go beyond this mandate to describe a situation of equal pay for work of equal value, that is, work that requires comparable levels of effort, skills, and responsibility. Rather than being a remedy for past discrimination, as in the case of affirmative action, this approach is aimed at developing policies that will guarantee equal treatment for men and women. While this idea seems to be catching on, there are still substantial gaps in the workplace authority and earnings of men and women throughout the world (see Wright et al., 1995; Jasso and Wegener, 1999), and researchers face a “paradox of the contented female worker” (Crosby 1982). Generally speaking, women have jobs with lower pay, less autonomy, and less authority than jobs held by men, but their level of earnings satisfaction is no different from that of men. Mueller and Wallace (1996) suggest that when levels of perceived justice are taken into account, substantial amounts of the

difference between men and women in their satisfaction with earnings can be explained. What remains to be understood are the factors that contribute to differences between men and women in perceptions of what is just.

#### **Divorce, Child Custody, and Child Support.**

Marital disruption is a conspicuous feature of the contemporary family that raises issues of justice not only for the people involved but for the public as well. Apart from justice matters arising from the division of property and assets in a divorce, when children are involved, the settlement becomes more complicated. Child custody disputes and child support issues present unique problems for justice analysis because in the first case the primary goods are indivisible and in the second case costs (or negative rewards) are being allocated. Procedures for resolving both sets of issues are governed by state and federal statutes or guidelines as well as by local customs and have varied considerably across time and culture. Historically, courts have employed justice principles involving parental entitlements to various degrees, but in recent times, the principle has evolved that the best interests of the child ought to be the sole or major consideration in custody decisions. As Elster (1989, p. 126) puts it, "although the child may to some extent and for some purposes be considered a consumption good for the parents, he is also and predominantly a person in his own right" who has an interest in the allocation. Elster argues, however, against the principle of the best interests of the child and suggests that in contemporary society, when joint custody is not feasible, three options present themselves: a presumption in favor of the mother, a presumption in favor of the primary caretaker (usually the mother), and tossing a coin.

Child support payments are vastly more determinate, and those judgments are considerably less Solomonic. Indeed, courts generally have allowed the participants, on the advice of counsel, to negotiate the nature of the awards. In child support situations, decision making occurs in two steps. There is a preliminary determination of an amount to be divided based on the needs of the child and then a decision on how that amount should be divided based on the relative resources of the parents. As Schaeffer (1980, p. 158) puts it, "beliefs about child support awards differ in important ways from other beliefs about justice . . . [because] allocating child support obligations involves allo-

cating not rewards, but responsibilities expressed as contributions." Schaeffer (1980) analyzed beliefs about the fairness of child support judgments by using a factorial survey involving vignettes. She found that child support awards are allocated according to a "proportional contribution-variable need" system in which parents' contributions are proportional to their resources. She concluded that preferences for child support awards embody a "modified version of 'from each according to their abilities, to each according to their needs'" (1980, p. 172).

**Intergenerational Relations.** Relationships among generations are one of the most important structural features in all societies. Some authors have argued that the prevailing social contract between generations in Western society regarding expectations, obligations, and well-being is changing (Bengtson and Achenbaum 1993). This is due to the changing demography of age but also to shifting cultural understandings of age differences and issues of equity. Bengtson (1993, p. 4) argues that "we have reached a cultural watershed concerning the implicit understanding of rights and obligations between age groups and generations in human societies. Never before have so many elders lived so long; never before have so relatively fewer members of younger age groups lined up behind them in the succession of generations." He argues further that as a consequence, "we are faced with new and historically unique dilemmas of family life and social policy agendas regarding the expectable life course and the succession of generations."

These observations raise a number of questions regarding the nature of intergenerational conflict and the linkage between age and economic expectations, economic performance, and evaluations of material well-being. The concept of justice has been applied to the study of intergenerational relations (see Norris 1987), but little work has explicitly linked social psychological theory regarding equity or justice evaluation to this range of issues. Moreover, researchers often have settled for conclusions based on conjecture or weak and inappropriate kinds of evidence. There is a wide range of intergenerational issues to which the justice framework can be applied, including not only the material well-being of older age groups but also the sense of obligation that adult children have about the support of their elderly parents as

well as public concern about the future health care and social security systems that will support the elderly in the future.

(SEE ALSO: *Affirmative Action; Comparable Worth; Decision-Making Theory and Research; Gender; Interpersonal Power; Poverty; Social Psychology; Utopian Analysis and Design*)

## REFERENCES

- Adams, J. Stacy 1965 "Inequity in Social Exchange." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 9. New York: Academic Press.
- Alves, Wayne M., and Peter H. Rossi 1978 "Who Should Get What? Fairness Judgments of the Distribution of Earnings." *American Journal of Sociology* 84:541-564.
- Alwin, Duane F. 1987 "Distributive Justice and Satisfaction with Material Well-Being." *American Sociological Review* 52:83-95.
- 1994 "Aging, Personality and Social Change: The Stability of Individual Differences Across the Life-Span." In D. L. Featherman, R. M. Lerner, and M. Permuter, eds., *Life-Span Development and Behavior*, vol. 12. Hillsdale N.J.: Erlbaum.
- , Galin Gornev, and Ludmilla Khakhulina 1996 "Comparative Referential Structures, System Legitimacy, and Justice Sentiments: An International Comparison." In *Social Justice and Political Change: Public Opinion in Capitalist and Post-Communist States*. New York: Aldine de Gruyter.
- Aristotle 1953 *Nicomachean Ethics*, trans. J. A. K. Thompson. London: George Allen and Unwin.
- Austin, William 1977 "Equity Theory and Social Comparison Processes." In J. M. Suls and R. L. Miller, eds., *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington D.C.: Hemisphere.
- Axelrod, Robert 1984 *The Evolution of Cooperation*. New York: Basic Books.
- Barry, Brian 1981 "Social Science and Distributive Justice." In Robert A. Solo and Charles W. Anderson, eds., *Value Judgment and Income Distribution*. New York: Praeger.
- Becker, Gary S. 1996 *Accounting for Tastes*. Cambridge, Mass.: Harvard University Press.
- Bengtson, Vern L. 1993 "Is the 'Contract across Generations' Changing? Effects of Population Aging on Obligations and Expectations across Age Groups." In V. L. Bengtson and W. A. Achenbaum, eds., *The Changing Contract across Generations*. New York: Aldine de Gruyter.
- , and W. Andrew Achenbaum 1993 *The Changing Contract across Generations*. New York: Aldine de Gruyter.
- Berger, Joseph, Maurice Zelditch, Bo Anderson, and Bernard P. Cohen 1972 "Structural Aspects of Distributive Justice: A Status-Value Formulation." In J. Berger, M. Zelditch, and B. Anderson, eds., *Sociological Theories in Progress*, vol. 2. Boston: Houghton-Mifflin.
- Berkowitz, Leonard, and Elaine Walster, eds. 1976 *Equity Theory: Toward a General Theory of Social Interaction. Advances in Experimental Social Psychology*, vol. 9. New York: Academic Press.
- Blau, Peter M. 1971 "Justice in Social Exchange." In H. Turk and R. L. Simpson, eds., *Institutions and Social Exchange: The Sociologies of Talcott Parsons and George C. Homans*. New York: Bobbs-Merrill.
- , and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Blauner, Robert 1989 *Black Lives, White Lives*. Berkeley: University of California Press.
- Boulding, Kenneth E. 1981 "Allocation and Distribution: The Quarrelsome Twins." In R. A. Solo and C. W. Anderson, eds., *Value Judgment and Income Distribution*. New York: Praeger.
- Brickman, Philip 1977 "Preferences of Inequality." *Sociometry* 40:303-310.
- , et al. 1981 "Micro-Justice and Macro-Justice." In M. J. Lerner and S. C. Lerner, eds., *The Justice Motive in Social Behavior*. New York: Plenum.
- Buchanan, Allen, and Deborah Mathieu 1986 "Philosophy and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- Cohen, Ronald L., ed. 1986. *Justice: Views from the Social Sciences*. New York: Plenum.
- Cook, Karen S. 1975 "Expectations, Evaluations and Equity." *American Sociological Review* 40:372-388.
- , and Karen A. Hegtvedt 1983 "Distributive Justice, Equity, and Equality." *Annual Review of Sociology* 9:217-241.
- , and ——— 1986 "Justice and Power: An Exchange Analysis." In H. W. Bierhoff, R. L. Cohen, and J. Greenberg, eds., *Justice in Social Relations*. New York: Plenum.
- , and Toshio Yamagishi 1983 "Social Determinants of Equity Judgments: The Problem of Multidimensional Input." In D. M. Messick and K. A. Cook, eds., *Equity Theory: Psychological and Sociological Perspectives*. New York: Praeger.
- Crosby, Faye 1976 "A Model of Egoistical Relative Deprivation." *Psychological Review* 83:85-113.

- 1982 *Relative Deprivation and Working Women*. New York: Oxford University Press.
- Davis, James A. 1959 "A Formal Interpretation of the Theory of Relative Deprivation." *Sociometry* 22:280-296.
- Della Fave, L. Richard 1980 "The Meek Shall Not Inherit the Earth: Self-Evaluation and the Legitimacy of Social Stratification." *American Sociological Review* 45:955-971.
- Deutsch, Morton 1975 "Equity, Equality and Need: What Determines Which Value Will Be Used as the Basis of Distributive Justice." *Journal of Social Issues* 31:137-149.
- 1986 *Distributive Justice: A Social Psychological Perspective*. New Haven Conn.: Yale University Press.
- DiQuattro, Arthur 1986 "Political Studies and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- Duncan, Greg, and Willard Rodgers 1991 "Has Children's Poverty Become More Persistent?" *American Sociological Review* 56:538-550.
- Elster, Jon 1984 *Ulysses and the Sirens: Studies in Rationality and Irrationality*, rev. ed. Cambridge, UK: Cambridge University Press.
- 1989 *Solomonic Judgements: Studies in the Limitations of Rationality*. New York: Cambridge University Press.
- Farkas, Arthur J., and Norman H. Anderson 1979 "Multidimensional Input in Equity Theory." *Journal of Personality and Social Psychology* 37:879-896.
- Featherman, David L., and Robert M. Hauser 1978 *Opportunity and Change*. New York: Academic Press.
- Festinger, Leon 1959 *A Theory of Cognitive Dissonance*. Stanford, Calif.: Stanford University Press.
- Fischer, Claude S., et al. 1996 *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, N.J.: Princeton University Press.
- Folger, Robert 1984 *The Sense of Injustice: Social Psychological Perspectives*. New York: Plenum.
- Furby, Lita 1986 "Psychology and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- Gartrell, C. David 1982 "On the Visibility of Wage Referents." *Canadian Journal of Sociology* 7:117-143.
- , and Bernard E. Paille 1997 "Wage Cuts and the Fairness of Pay in a Worker-Owned Plywood Cooperative." *Social Psychology Quarterly* 60:103-117.
- Greenberg, Jerald, and Ronald L. Cohen 1982 *Equity and Justice in Social Behavior*. New York: Academic Press.
- Hamilton, V. Lee, and David Rauma 1995 "Social Psychology of Deviance and the Law." In Karen Cook, Gary Fine, and James House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Hauser, Robert M., and David L. Featherman 1977 *The Process of Stratification: Trends and Analyses*. New York: Academic Press.
- Hegtvedt, Karen A., and Barry Markovsky 1995 "Justice and Injustice." In Karen Cook, Gary Fine, and James House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Heider, Fritz 1958 *The Psychology of Interpersonal Relations*. New York: Wiley.
- Herrnstein, Richard J., and Charles Murray 1994 *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Hochschild, Jennifer L. 1981 *What's Fair? American Beliefs about Distributive Justice*. Cambridge, Mass.: Harvard University Press.
- Homans, George C. (1961) 1974 *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace Jovanovich.
- 1976 "Commentary." In L. Berkowitz and E. Walster, eds., *Equity Theory: Toward a General Theory of Social Interaction*. New York: Academic Press.
- Jasso, Guillermina 1978 "On the Justice of Earnings: A New Specification of the Justice Evaluation Function." *American Journal of Sociology* 83:1398-1419.
- 1980 "A New Theory of Distributive Justice." *American Sociological Review* 45:3-32
- 1999 "How Much Injustice Is There in the World? Two New Justice Indexes." *American Sociological Review* 64:133-168.
- , and Peter H. Rossi 1977 "Distributive Justice and Earned Income." *American Sociological Review* 42:639-651
- , and Bernd Wegener 1997 "Methods for Empirical Justice Analysis: Part 1. Framework, Models, and Quantities." *Social Justice Research* 10:393-430.
- 1999 "Gender and Country Differences in the Sense of Justice." *International Journal of Comparative Sociology* 40:94-116.
- Jencks, Christopher S., et al. 1972 *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- 1979 *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic Books.
- Kluegel, James R., David S. Mason, and Bernd Wegener 1995 *Social Justice and Political Change: Public Opinion in Capitalist and Post-Communist States*. New York: Aldine de Gruyter.
- , and Eliot R. Smith 1981 "Beliefs about Stratification." *Annual Review of Sociology* 7:29-56.

- 1986 *Beliefs about Inequality: Americans' Views of What Is and What Ought to Be*. New York: Aldine de Gruyter.
- Lane, Robert 1986 "Market Justice, Political Justice." *American Political Science Review* 80:383-402.
- Levy, Frank 1988 *Dollars and Dreams: The Changing American Income Distribution*. New York: Norton.
- , and Richard J. Murnane 1992 "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations." *Journal of the Economic Literature* 30:1333-1381.
- Major, Brenda, and Blythe Forcey 1985 "Social Comparisons and Pay Evaluations: Preferences for Same-Sex and Same-Job Wage Comparisons." *Journal of Experimental Social Psychology* 21:393-405.
- Markovsky, Barry 1985 "Toward a Multilevel Distributive Justice Theory." *American Sociological Review* 50:822-839.
- Messick, David M., and Karen A. Cook 1983 *Equity Theory: Psychological and Sociological Perspectives*. New York: Praeger.
- Mikula, Gerald 1980 *Justice and Social Interaction: Experimental and Theoretical Contributions from Psychological Research*. New York: Springer-Verlag.
- Molm, Linda D. 1991 "Affect and Social Exchange: Satisfaction in Power-Dependence Relations." *American Sociological Review* 56:475-493.
- Moynihan, Daniel Patrick 1968 *On Understanding Poverty*. New York: Basic Books.
- Mueller, Charles W., and Jean E. Wallace 1996 "Justice and the Paradox of the Contented Female Worker." *Social Psychology Quarterly* 59:338-349.
- Nader, Laura, and Andree Surssock 1986 "Anthropology and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- Norris, Joan E. 1987 "Justice and Intergenerational Relations: An Introduction." *Social Justice Research* 1:393-403.
- Nozick, Robert 1974 *Anarchy, State, and Utopia*. New York: Basic Books.
- Patchen, Martin 1961 *The Choice of Wage Comparisons*. Englewood Cliffs, N.J.: Prentice-Hall.
- Pettigrew, Thomas F. 1967 "Social Evaluation Theory: Convergences and Applications." In D. Levine, ed., *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- Rae, Douglas 1981 *Equalities*. Cambridge, Mass.: Harvard University Press.
- Rainwater, Lee 1974 *What Money Buys*. New York: Basic Books.
- Randall, Christina S., and Charles W. Mueller 1995 "Extensions of Justice Theory: Justice Evaluations and Employee Reactions in a Natural Setting." *Social Psychology Quarterly* 58:178-194.
- Rawls, John 1971 *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Robinson, Robert V., and Wendell Bell 1978 "Equality, Success and Social Justice in England and the United States." *American Sociological Review* 43:125-143.
- Runciman, W. 1966 *Relative Deprivation and Social Justice*. Berkeley: University of California Press.
- Rytina, Steve 1986 "Sociology and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- Schaeffer, Nora Cate 1990 "Principles of Justice in Judgments about Child Support." *Social Forces* 69:157-179.
- Sennett, Richard, and Jonathan Cobb 1972 *The Hidden Injuries of Class*. New York: Vintage Books.
- Sewell, William H., and Robert M. Hauser 1975 *Education, Occupation, and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Solo, Robert A., and Charles W. Anderson 1981 *Value Judgment and Income Distribution*. New York: Praeger.
- Southwood, Kenneth E. 1978 "Substantive Theory and Statistical Interaction: Five Models." *American Journal of Sociology* 83:1154-1203.
- Stolte, John F. 1983. "The Legitimation of Structural Inequality." *American Sociological Review* 48:331-342.
- Thibaut, John W., and Harold H. Kelley 1959 *The Social Psychology of Groups*. New York: Wiley.
- , et al. 1975 *Procedural Justice: A Psychological Analysis*. Hillsdale, N.J.: Erlbaum.
- 1987 "A Surge in Inequality." *Scientific American* 256:30-37.
- Tyler, Tom R. 1984 "Justice in the Political Arena." In R. Folger, ed., *The Sense of Injustice: Social Psychological Perspectives*. New York: Plenum.
- 1986 "The Psychology of Leadership Evaluation." In H. W. Bierhoff, R. L. Cohen, and J. Greenberg, eds., *Justice in Social Relations*. New York: Plenum.
- Walster, Elaine, Ellen Berscheid, and G. William Walster 1973 "New Directions in Equity Research." *Journal of Personality and Social Psychology* 25:151-176.
- Walster, Elaine G., William Walster, and Ellen Berscheid 1978 *Equity: Theory and Research*. Boston: Allyn and Bacon.
- Walzer, Michael 1983 *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.

- Williams, Robin M., Jr. 1975 "Relative Deprivation." In L. A. Coser, ed., *The Idea of Social Structure: Papers in Honor of Robert K. Merton*. New York: Harcourt Brace Jovanovich.
- Worland, Stephen T. 1986 "Economics and Justice." In R. L. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum.
- , and Luca Perrone 1977 "Marxist Class Categories and Income Inequality." *American Sociological Review* 42:32–55.
- Wright, Eric Olin, Janeen Baxter, and Gunn Elisabeth Birkelund 1995 "The Gender Gap in Workplace Authority: A Cross-National Study." *American Sociological Review* 60:407–435.

DUANE F. ALWIN

## SOCIAL MOBILITY

The term "social mobility" describes the nature and amount of change in social position over time. In principle, this change can be defined for any social entity. Thus, one can study the "collective mobility" of classes, ethnic groups, or entire nations in terms of, for example, average health status, literacy, education, or gross domestic product per capita. More commonly, the term is used in connection with the movement of individuals or families. However, even though social mobility typically is defined with respect to micro units of society, the pattern of mobility across those units generally is considered a core characteristic of a society's social structure, and the study of this mobility generally is recognized as a fundamental area of macro-level sociology.

Social mobility typically is conceptualized in terms of the quantity of movement and the distribution of its direction and distance. The different rates that together constitute the mobility structure of a society is highly complex, however, for several reasons. First, societies have more than one dimension along which mobility can occur. Thus, one can speak of occupational mobility, social class mobility, educational mobility, job mobility, income mobility, wealth mobility, and so on. In principle, one also can use the term "social mobility" to describe movement among nonhierarchical social statuses, such as religious affiliation mobility and geographic mobility or mobility across categories that describe attitudes, belief systems, life

styles, and the like. The dominant use of the term in the literature, however, concerns mobility along a social hierarchy that defines a dimension of social inequality in a society. Second, even with respect to a single hierarchy, the mobility structure is not easy to summarize. A different rate of mobility can be calculated with respect to each combination of origin and destination position along the social hierarchy in question. Empirically, it may be possible to summarize this collection of rates accurately in terms of a function of the social distance between origin and destination or in terms of specific relationships between the origin and destination categories. In general, however, an accurate summary cannot be expressed in terms of a single number. Thus, for each social hierarchy, there is not a single rate of social mobility but a core set of rates that, taken together, can be termed the structure of mobility with respect to the particular hierarchical dimension.

Social mobility is an important issue in sociology for several reasons. For one thing, it is relevant to social equity. Philosophical and moral evaluations of social inequality often depend not only on the level of inequality in a society but also on the extent to which individuals or families can leave disadvantaged states during their lifetimes or across generations. Social mobility is also an important explanatory factor in social theory. The basic stratification variables affect a wide variety of social outcomes and behaviors, but these effects accumulate over time; social mobility therefore affects outcomes by changing the states and durations of these key explanatory variables. The societal rate of mobility also may have macro-level consequences. An early conjecture in this area appears in the work of Werner Sombart, who argued that the failure of early twentieth century socialist parties in the United States stemmed in part from the high rate of American social mobility, which prevented the formation of strong class identification.

The longest-standing tradition in sociological mobility research concerns mobility in occupational groupings or social classes. Much of this work has used "mobility tables" (cross-classifications of origin by destination position) to study "intergenerational mobility," that is, the extent to which the social position of adults differs from that of their parents. Another large body of work has focused on "intragenerational mobility," or the mobility experienced by individuals or families



over the course of their adult lives. Because male labor force participation generally has been higher and more persistent than female participation and because of the somewhat controversial presumption that the status of a family derives from the status of the male breadwinner, for many years these studies focused on intergenerational mobility between fathers and sons, although more recent literature has examined the structure of mobility for women as well.

An important question in intergenerational mobility research is whether overall rates of intergenerational social mobility differ by country. Earlier in the century, scholars hypothesized that the United States had especially high rates of mobility, and some argued that those rates were a consequence of the American meritocratic value system. More recently, it became clear that the primary factors in cross-national differences in mobility rates are structural, not cultural. Differences in so-called structural mobility across countries arise from the extent to which the distribution of positions for sons or daughters differs from the distribution of positions for their fathers. Changes in this distribution across generations (as well as more subtle factors such as class differences in fertility, death rates, and migration rates) necessarily produce intergenerational social mobility. Countries whose occupational distribution is changing rapidly (high rates of structural change) therefore have greater levels of mobility than do countries whose occupational distribution is changing slowly.

Not all social mobility occurs as a result of structural change. The component of social mobility that occurs beyond the amount produced by structural change is typically called circulation mobility, exchange mobility, or relative mobility. The Featherman, Jones, Hauser (FJH) hypothesis of the mid-1970s asserts that cross-national and historical differences in social mobility are accounted for almost completely by differences in levels of structural mobility. According to the strong form of this hypothesis, once structural mobility is taken into account, the pattern of relative mobility chances is invariant over time and across countries. This pattern has three principal features: (1) relatively high immobility at the top and bottom of the hierarchy, (2) higher levels of short-range mobility than long-range mobility (moves from the top to the bottom or from the bottom to the top of the hierarchy are especially rare), and (3) a rela-

tively small impact of origins on destinations in the middle of the hierarchy.

More recent research has determined that even though the weak form of the FJH hypothesis (overall mobility differences are due largely to differences in structural mobility) is supported by the data, the strong form (invariance of relative mobility chances) appears to be false. However, further progress on this issue has been elusive. In particular, the question of whether cross-national differences in relative mobility chances are the subject of such complex national historical differences that they are idiosyncratic or whether they are the product of a more parsimonious set of structural forces (e.g., the extent to which the political system is democratic, the level of modernization, and the level of social inequality) remains to be answered.

Another continuing challenge in mobility research concerns conceptualization and measurement of the component of mobility that is due to structural change. The specification of this causal force in terms of differences in the distribution of positions of fathers and their adult children is problematic for subtle but important reasons. Such an identification assumes that the observed destination distribution is caused by forces (such as technological change) that are not affected by (and therefore are a legitimate cause of) the observed mobility process. This amounts to assuming that the observed destination distribution constitutes a rigid supply constraint, a set of preexisting empty vacancies that are filled by the movement of sample members with respect to their origin positions. This assumption is never perfectly true. If the "supply constraint" is not rigid (and it is unlikely to be so), the observed distribution of destination positions (which by definition represents a summing up of the mobility outcomes for a particular statistical sample) is a consequence of the mobility process as well as of the "structural forces" that constrain the character of this destination distribution. It therefore cannot be taken to be a pure cause of social mobility. The logic of structural mobility becomes especially problematic when subgroups of the population are studied in this fashion. For example, if the distribution of women's occupations shifts toward high-status occupations relative to the total occupational distribution, it is problematic to argue that the relative improvement of women's destinations is a "cause" of

women's higher levels of social mobility as opposed to being a consequence of that mobility. This problem, which is sometimes referred to as the "reflection" problem, has not had a satisfactory solution.

Although structural change is a major part of the explanation for overall levels of intergenerational social mobility in a society, it cannot explain differences in the likelihood that particular individuals will be upwardly or downwardly mobile. The prevailing pattern of circulation mobility that was noted above (relatively high levels of immobility at the top and bottom, the predominance of short-range over long-range mobility, etc.) implies that class of origin is a significant predictor of the types of mobility that do occur. However, an explanation for destination positions that relied solely on the status of origin would be unsatisfactory in two respects: First, the predictive power of social origins by itself is relatively weak; second, the explanation does not indicate how and why social origins matter.

Efforts to redress these deficiencies stem largely from the publication of Blau and Duncan's *The American Occupational Structure* (1967). A major goal of that work was to understand whether the educational system operated primarily as a device that transmitted the status of parents to their children or as an engine of social mobility that freed children from the effects of the status of their parents. To accomplish that goal, Blau and Duncan developed what has come to be known as the status attainment model. Their approach to the study of mobility assumed a dominant metric to social hierarchy: the socioeconomic status of occupations. Their research showed that, at least for men (Blau and Duncan did not study the mobility of women), education was a more important determinant of a son's adult socioeconomic status than were his socioeconomic origins. Furthermore, while educational attainment was strongly influenced by socioeconomic origin, most of the individual-level variation in educational attainment was not explained by socioeconomic origin. Those authors also showed that most of the effect of socioeconomic origins on outcomes was indirect, derived from the effect of those origins on education. Finally, the effect of education on occupational attainment regardless of social background was much larger in the United States than was the direct effect of father's socioeconomic background

(regardless of the son's education). These findings led many to interpret Blau and Duncan's research to mean that the United States more closely approximated an "achievement" than an "ascription" society, although others pointed to the still large (even if not decisive) disadvantage arising from low socioeconomic origin along with the disadvantages associated with being a first-generation immigrant, an African-American, or a woman as constituting important qualifications to such a generalization.

The Blau and Duncan approach essentially divided the intergenerational mobility process into three segments. The first segment concerned the process of educational attainment, the second concerned the transition from school to work, and the third concerned the "intragenerational" mobility that occurs over the working life. Leaving aside the powerful but difficult to specify force of structural change, this division may offer the best possibility for understanding the mechanisms that lie behind intergenerational social mobility as well as identifying possible policy interventions and shedding light on three processes that have great importance in their own right. Each of these processes calls attention to specific institutions (in particular, the educational system and the labor market) that facilitate, limit, or channel social mobility. The focus on how institutional forces constrain the impact of individual resources on individual outcomes sometimes is referred to as the "fourth generation" of social mobility research (with early mobility studies being the first generation, the status attainment tradition being the second, and statistically sophisticated analyses of mobility tables being the third).

A large body of literature has grown around each of these components of the intergenerational mobility process. With respect to education, scholars have conceptualized the educational career as a set of transitions to successively higher grades and have asked whether family background has the same influence at each grade level of this transition process. Results for the United States and several other countries suggest that the effects of family background decline at higher-grade transitions, though these findings are controversial. Assuming that the decline is real, some scholars have argued that the historical raising of the minimum school-leaving age should have reduced the impact of family of origin on outcomes over time.

Again, however, while there is some evidence that the effects of family background have declined during the twentieth century and that these declines are caused by the expansion of education, empirical studies have failed to confirm this conjecture decisively.

A second major focus in the literature concerns the reasons why socioeconomic background is associated with educational performance. It has been appreciated since Sewell and associates developed the "Wisconsin model" in the early 1970s that there is a social psychological component to mobility in which family status is related to parental expectations for the child. In combination with grades in school, peer group influences, and teachers' expectations, this shapes a student's educational and occupational aspirations. More recent work has reconceptualized these family advantages or disadvantages in terms of cultural resources ("cultural capital"), which sometimes are specified as a family's participation in "high-cultural" activities (exposure to art museums, opera, theater, dance, etc.); in other studies, they are defined more broadly (and vaguely) as encompassing all the cultural advantages a family may possess that affect a child's ability to do well in school. Other recent literature focuses on "social capital," which sometimes is interpreted to mean the level and quality of interaction parents have with their children and at other times is interpreted to refer to the resources embedded in the parents' social networks that could in principle influence a child's outcomes. A third, rather controversial focus of attention in recent years concerns possible links between socioeconomic status and genes and the extent to which intergenerational correlations among status variables (particularly educational outcomes) indicate the presence of a genetic force. A fourth focus concerns the specific consequences of low income on children's development and later socioeconomic outcomes. A fifth focus concerns the extent to which the characteristics of schools, neighborhoods, and communities can mute or exaggerate the impact of family characteristics on educational outcomes.

The second mobility component is the transition from school to work. A large body of literature focuses specifically on aspects of this transition, including variation in the extent to which the diplomas, degrees, and advanced degrees provided by schools are linked by law or custom to

specific occupational careers; the extent to which credentials are standardized in a country; the extent to which the supply of those credentials is controlled by schools in light of estimated demand; and the extent to which students who graduate with these diplomas or degrees are provided with knowledge of the relevant job market. Many policy concerns in the United States focus on those who leave school before the tertiary level and the extent to which they are provided with a mix of academic and vocational skills and credentials that is valuable on the job market. Vocational education in particular is organized quite differently across industrialized societies, and in recent years comparative research on this transition has accelerated.

The third component concerns intragenerational mobility over the life course. This research has taken different forms. The Blau and Duncan approach largely emphasized the mean or typical pattern of life-course development as a function of education, first job, and father's occupation. In this form, the question of mobility is reduced to a question about the average status "return" to the resources an individual possesses on first entering the labor market. Although this approach is informative about the typical level of status advancement during the work career as a function of origin conditions, it suffers from two deficiencies: First, it does not explain how education and the first job lead to the current job; second, it does not provide an explanation of the frequency or consequences of deviations from the typical amount of status advancement during the work career.

An understanding of the full distribution of outcomes (i.e., both upward and downward career mobility) is made possible through the use of the "mobility table" approach that has been applied to the study of intergenerational mobility between the status of the father and the status of the son or daughter. The prevalent approach in recent sociology, however, has been more institutional. One line of work has focused on structural linkages between jobs in particular occupational or organizational labor markets. This work has addressed the implications of entering these "internal labor markets" for subsequent career advancement, with an important subset of it directed at questions about whether these institutional mechanisms reproduce, enhance, or mute racial or gender differences. Because these job linkages generally are not

expressed in terms of abstract hierarchical measures such as class and socioeconomic status, studies of these organized labor markets frequently have turned away from the earlier focus on class or status and toward more concrete reward variables such as earnings and job level within an organizational hierarchy. Jobs outside of organized hierarchies that lacked other forms of institutional protection (such as professional licensing requirements) were characterized as “open” or (if low-quality) “secondary” labor market jobs. For several years, sociologists hoped that a parsimonious set of labor market “boundaries” could be operationalized and used to explain labor market outcomes in terms of labor market segment early in one’s career. The promise of this “segmented labor market” approach to career mobility has faded, however. It is now recognized that the boundary between unstable, low-paying jobs in what once was commonly referred to as the “secondary labor market” and “internal labor market” jobs is by no means impermeable, especially in the early years of the adult life course. The segmented labor market approach has been undermined further by the appreciation of the numerically high levels of mobility (including involuntary mobility) out of corporate jobs, often as a result of plant closings and corporate restructuring. Literature on “displaced workers” that has developed largely in labor economics rather than sociology has attempted to quantify the short-term and medium-term career consequences of job displacement (the literature shows only transitory effects on employment but more durable effects on earnings). It can be assumed that job displacement is a principal mechanism by which structural change produces short-term and longer-term intragenerational (and ultimately intergenerational) occupational mobility. However, sociologists have only begun to explore the connection between job displacement and the structural mobility observed in mobility tables.

A separate body of literature has addressed the intragenerational mobility of people who at one time or another in their lives are poor. Aside from questions about the intergenerational transmission of poverty, much of this literature has focused on whether poverty is a permanent or transitory status. It has been recognized that most poverty in the United States is transient, although an important fraction of the poor remain poor for

long periods, while many who escape poverty have a relatively high probability of returning to poverty in the future. Much of this literature addresses the factors that influence rates of entry into and exit from poverty.

Poverty studies use a measure of income, especially income in relation to needs, rather than class or status as the basic measure of position. They typically make the family the relevant unit of measurement because it is family income, not individual earnings or status, that most directly determines poverty status. They also direct attention to the facts that income mobility is a household, not an individual-level, concept; that income mobility can be generated by labor market events involving one’s partner as well as oneself; that public transfers can be an important source of income and can play a significant role in determining levels of income mobility; and that changes in household composition (including marriage, cohabitation, and union dissolution) can strongly influence income mobility.

More recent mobility literature has focused as much attention on instability as on stability (or stable “career advancement”) over time. This emphasis raises important questions about an important presupposition underlying the sociological framework for mobility studies: that it was meaningful to conceptualize the socioeconomic status of the family of origin as a stable point and the “current” status of the adult son or daughter as a “realized” socioeconomic status that could be compared with the point of origin. Early studies were forced by the limitations of data to use parental status at a single point in time (e.g., the point at which the respondent was 16 years old) as the measurement of family status over the duration of childhood. The growing availability of multigenerational panel data that provide information about the possibly changing status of the family of origin during childhood has made it possible for scholars to study how temporal variations in the status of parents affect the process of intergenerational transmission. These more extensive data on the lives of parents and their grown-up children are allowing scholars to study intergenerational and intragenerational mobility with respect to statuses such as income, wealth, and poverty, which are perhaps more volatile than are occupational status and class position.

Questions about racial and/or ethnic or gender differences in mobility have largely been subordinated to gender and racial and/or ethnic inequality and changes in levels in inequality over time. In other words, the focus has been more on the *collective* mobility of these groups with respect to white males than on whether the structure of individual-level mobility within groups defined by race or gender is different from the structure of mobility for white males. This literature has perhaps paid more attention to economic outcomes than to class or status outcomes. Research on gender inequality in particular has avoided the use of status metrics or broad occupational groups, which understate the gender inequality that is visible in earnings. The issue of mobility still plays an important role in this literature because of the possible role of mobility processes in explaining how gender or racial and/or ethnic inequality comes about. In the case of women, the literature has focused on why women take a different mix of academic subjects than men do and on how gender differences in the transition from school to work and in career mobility produce differences in the average earnings of women and men over the life course. In this literature, the pattern and quantity of work experience and the different distribution of men and women across jobs and occupations (“sex segregation”) have been the issues of greatest interest. In contrast, the role of specifically intergenerational processes and their possible impact on gender inequality or on the well-documented declining level of gender inequality in earnings has received less attention.

The issue of race is in some respects parallel to that of gender but has its own unique features. Women and men grow up in the same families, while whites and nonwhites grow up in different families, and these differences involve socioeconomic factors as well as race or ethnicity per se. Furthermore, racial and/or ethnic segregation by educational major, job, or occupation has not been as extreme as gender segregation in recent years. However, just as research has shown that the effects of race on income declined over the post-war years (though the trend has stalled and perhaps reversed since about 1980), it also has shown that the direct effects of race on socioeconomic attainment have declined, at least through the late 1980s. African-Americans experience no disadvantage at all in educational attainment because of

race per se (though they experience a disadvantage stemming from their lower average socioeconomic origins). Specifically race-based intergenerational factors still may affect the levels of black-white inequality in the next generation, but they probably operate through the quality (as opposed to quantity) of schooling, and these effects are not well understood.

Comparative analyses of mobility that go beyond the mobility table approach described above are complicated by substantive differences in institutional structures across nations and differences in the measurement of key variables. Nonetheless, progress is being made. In perhaps the most notable application of the original Blau and Duncan model to comparative analysis, Treiman and Yip used the ratio of the net effect of education on occupational attainment to the direct effect of the father’s socioeconomic background (this might be conceptualized as a ratio of “achievement” to “ascription”) to compare the process of attainment in different countries. This ratio varies from a relatively high level in industrialized societies (particularly in Scandinavia) to a low level in less industrialized societies (with India having the lowest value in their study). Scholars also are focusing on comparative studies of topics such as the transition from school to work, job mobility, earnings mobility, sex segregation, and family dynamics. Many of these studies are being carried out with newly available panel data on demographic and socioeconomic outcomes that are being collected in many industrialized societies. These new sources of data are complex, and it will take several years before a broad-based comparative literature that uses them becomes available. The direction and pace of research, however, are encouraging. Social mobility is likely to retain its vitality as well as its centrality in sociology for the foreseeable future.

## REFERENCES

- Blau, Peter, and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Corcoran, Mary 1995 “Rags to Rags: Poverty and Mobility in the United States.” *Annual Review of Sociology* 21:237–267.
- Duncan, Greg J., and Jeanne Brooks-Gunn, eds. 1997 *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

- Erikson, Robert, and John H. Goldthorpe 1992 *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford, UK: Clarendon Press.
- Fischer, Claude S. 1996 *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, N.J.: Princeton University Press.
- Ganzeboom, Harry B. G., Donald J. Treiman, and Wout C. Ultee 1991 "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277-302.
- Grusky, David B. 1994 *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colo.: Westview.
- Jacobs, Jerry A. 1996 "Gender Inequality and Higher Education." *Annual Review of Sociology* 22:153-185.
- Kerckhoff, Alan 1995 "Institutional Arrangements and Stratification Processes in Industrial Societies." *Annual Review of Sociology*. 21:323-347.
- Reskin, Barbara 1993 "Sex Segregation in the Workplace." *Annual Review of Sociology* 19:241-269.
- Rosenbaum, James E., T. Kariya, R. Settersten, and T. Maier 1990 "Market and Network Theories of the Transition from High School to Work: Their Application to Industrialized Societies." *Annual Review of Sociology* 16:263-299.
- Rosenfeld, Rachel 1992 "Job Mobility and Career Processes." *Annual Review of Sociology* 18:39-61.
- Shavit, Yossi, and Hans-Peter Blossfeld 1993 *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview.
- , and Walter Müller, eds. 1997 *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford, UK: Oxford University Press.
- Sørensen, Annemette 1994 "Women, Family, and Class." *Annual Review of Sociology* 20:27-47.

THOMAS A. DiPRETE

## SOCIAL MOVEMENTS

Social movements can be described most simply as collective attempts to promote or resist change in a society or group. The degree of change advocated and the level at which changes are pursued vary across all types of social movements, whether religious, political, or student. Some movements clamor for sweeping, revolutionary transformations, whereas others pursue specific moderate reforms. The level at which changes are sought

varies from global and national alterations of social structures to attitudinal, spiritual, and lifestyle changes.

### TYPES OF MOVEMENTS

Revolutionary movements such as the Bolshevik, Palestinian, Islamic jihad, and Irish Republican movements seek fundamental structural changes. These movements pursue radical changes in a society's basic institutions or, in some cases, major changes in the world order. Because these groups challenge the legitimacy of extant authorities, powerful elites typically use every means, including violence, to repress revolutionary movements.

Reform movements, in contrast, attempt to modify structural relations without threatening existing institutions. Consequently, while some elites oppose any reforms, they are usually more tolerant of reform movements than they are of revolutionary ones. Some reform movements, such as the peace, women's, and environmental movements, are general in scope (Blumer 1946) and often blend a plethora of political and lifestyle objectives. Peace movements, for example, not only pursue a variety of political objectives (e.g., preventing and stopping wars, opposing specific weapons, promoting disarmament, changing foreign policy, establishing conflict resolution institutions) but also strive to persuade individuals to change their attitudes and live more peaceful everyday lives.

Other reform movements, such as the anti-abortion, women's temperance, and anti-drunken-driving movements, focus on specific issues. Although specific reform movements are considerably narrower in scope than general reform movements are, they also may organize around both political and lifestyle objectives (Staggenborg 1987).

Still other reform movements, such as various self-help, human potential, and New Age movements, focus almost exclusively on lifestyle and identity issues. In contrast to other movements, these movements tend to disregard social structural issues. Instead, they concentrate on changing individuals.

Finally, social movements frequently generate organized opposition in the form of countermovements. Countermovements attempt to prevent revolutionary or reform movements from securing the

changes they promote. As a result of their counterreformist tendencies, most countermovements (e.g., the antibusing, McCarthyist, stop-ERA, and Moral Majority movements) are conservative (Lo 1982); that is, they attempt to preserve extant institutions, cultural practices, and lifestyles.

Regardless of the particular type of social movement and the scope and level of change it advocates or opposes, all movements share certain common characteristics that are of interest to social scientists. First, all movements emerge under a specific, complex set of historical, cultural, and structural conditions. Second, as a movement emerges, a variety of participation issues arise, including recruiting new members, building commitment, and sustaining participation. Third, every movement is organized to some degree. The most visible manifestations of movements are their organizations and their strategies and tactics. Third, by virtue of its existence, every social movement has some consequences, however minimal. Although researchers frequently are concerned with the extent to which movements affect social change, definitive answers to this question have proved illusory.

### MOVEMENT EMERGENCE

Social scientists have devoted considerable attention to the factors associated with the emergence of social movements. Early theory and research asserted that movements arise when societies undergo structural strain, such as during times of rapid social change (Smelser 1962). These “breakdown theories” posit that “large structural rearrangements in societies—such as urbanization and industrialization” lead to the dissolution of social controls and heighten “the impulse toward antisocial behavior” (Tilly et al. 1975, p. 4). Hence, these systemic “breakdowns” were said to cause an increase in strikes, violent collective action, and social movements.

Later social movement scholars criticized breakdown theories on empirical and theoretical grounds. Rather than viewing movement emergence and participation as aberrations, scholars now view them as “simply ‘politics by other means,’ often the *only* means open to relatively powerless challenging groups” (McAdam 1988, pp. 127–128).

However, if these groups are powerless, it is important to understand the conditions that affect the likelihood that they will mobilize. To do so researchers first turned to the structural factors that are conducive to the emergence of social movements.

One macro structural factor concerns the “structure of political opportunities” (Eisinger 1973). Movements emerge when there is a “receptivity or vulnerability of the political system to organized protest” (McAdam et al. 1988, p. 699). Researchers exploring the U.S. civil rights movement, for instance, conclude that that movement’s emergence was facilitated by a series of interrelated changes in the structure of political opportunities. Those changes included the decline of cotton markets, African-American migration to the North, the expansion of the black vote, and the electoral shift to the Democratic party (McAdam 1982).

Additionally, researchers have identified the absence of repression as a related macro structural factor. Social movements sometimes are spared a violent or otherwise repressive response from the authorities not only during times of breakdown or regime crisis but also during periods of expanding political opportunities such as times of state building. For example, as the former Soviet Union took strides between 1985 and 1989 to open discourse and other political opportunities (“perestroika”), there was a dramatic increase in protest activities (Zdravomyslova 1996). The result is what researchers have termed a cycle of contention: “the phase of heightened conflict across the social system” (Tarrow 1998, p. 142).

During times of increased movement activity, the authorities can, and sometimes do, take an active stance toward challenging groups. However, while their initial attempts to repress movements often fan the flames of discontent and fuel further protest activities, research suggests that the relationship between collective action and repression is bell-shaped (Tilly 1978). If the authorities later respond by increasing the severity of the repression, as occurred when the Chinese authorities ordered tanks and troops into Tiananmen Square to fire on student demonstrators, the cost of collective action usually becomes too high for movements to continue their challenges.

A nation-state can vary in structural factors such as institutional strength, access of challengers

to polity membership and/or decision making, and configurations of power (Kriesi 1996; Rucht 1996; Ferree and Gamson 1999). The political environment, or context structure, of a nation-state then can influence movements and movement emergence. For instance, France after 1981 had a strong, exclusive government and a political party system with large inter- and intraparty divisions. In this case, movement emergence and growth depended in part on the left's support of solidarity movements to gain political advantage. In contrast, Switzerland had a weak but inclusive government in the 1980s, allowing for the growth of diverse groups with multiple areas of focus (Kriesi 1996).

Many contemporary social movements are affected increasingly by globalization, or the creation and intensification of "worldwide social relations which link distinct localities in such a way that local happenings are shaped by events occurring miles away and vice versa" (Giddens 1990, p. 64). Technological revolutions in communications and transportation as well as economic, cultural, and political developments have increased global interdependence and consciousness of the global whole. Globalization thus spawns similarities in movement mobilizations and context structures across different nations (della Porta and Kriesi 1999). These similarities often lead to the cross-national diffusion of values and beliefs, facilitated in part by direct and indirect links between similar movements in various countries. German students studying in the United States in the early 1960s, for example, drew on the American student movement and later on access to networks they found during their stay in the United States in mobilizing their own student movement in Germany (McAdam and Rucht 1993).

Operating in much the same way as cross-national linkages, preexisting organizations in a nation-state serve as communication networks for the discontented members of a population (Freeman 1973). In fact, they can aid or inhibit the spread of information from cross-national linkages. More important, they provide a base for mobilizing the resources needed to sustain a movement. Churches, for example, were important indigenous organizations that contributed to the emergence of the contemporary peace, civil rights, and Moral Majority movements.

Several European scholars contend that state intervention into private domains of life has generated new social movements. According to this perspective, various structural changes in Western industrialized societies, especially changes in the system of production, led the state to seek control over previously private domains. Consequently, private domains such as sexual relations, biological identity, birth and death, illness and aging, and one's relationship to nature "have entered the realm of 'public' conflict" (Melucci 1980, p. 219). New social movements (e.g., the women's, gay rights, euthanasia, and environmental movements) emerged to reclaim those areas from the state.

The foregoing analysis indicates that numerous structural factors are crucial to providing an opportunity for the emergence of social movements, yet those factors alone cannot account for the rise of a particular movement. Why is it that when the structural conditions appear to be ripe for the emergence of a particular movement, frequently no movement appears? To address this question, some researchers have begun to investigate both the cultural and the micro interactional factors associated with the emergence of social movements. What is most important here are the reasons people take action in the first place: their grievances.

As a commonly shared stock of knowledge and activity, culture plays a role in movement emergence by giving individuals a sense of commonality with others. Specifically, culture can aid in the creation of a common identity that sets itself apart from that of other groups. Furthermore, when groups feel subservient in a society, they may create forms of activities and beliefs that express opposition to the dominant culture. This opposition can result in interests and needs that conflict with the dominant culture and in the development of grievances. In the 1980s, for example, the naming of the *Québécois* nationalist movement in Canada set that group apart from the rest of Canadian society by emphasizing the common culture shared by the group's members and symbolized a collective desire for political empowerment (Jenson 1995).

Culture is important in the emergence and cross-national diffusion of social movements, then, because it provides a common way in which to view the world and ways to express that worldview. Indeed, movements, such as the Japanese-based



Nichiren Shoshu/Sokagakkai Buddhist group may actively attempt to spread their worldviews and the cultural actions and artifacts to which they are attached to facilitate movement emergence in other countries (Snow and Benford 1999).

Worldviews are also important because they provide a yardstick by which to evaluate events. When an event fails to measure up against that yardstick, people experience a moral shock that can lead to movement emergence (Jasper 1997). However, although a single event may generate several movements with similar goals, culture can play a role in amplifying the moral shock that leads to collective action. Protests against the Gulf War emerged throughout the world more quickly than they had during previous wars, and there initially appeared to be similarity in the information received by those movements, along with the timing and stances taken by them. For example, the slogan “No Blood for Oil” was used by movements globally (Koopmans 1999). However, while they all were against the Gulf War, the peace movements in various countries mobilized differently on the basis of culturally filtered considerations. French peace movements, for instance, were against any coalition with the Americans. By contrast, the German peace movement was against any potential use of the German military on foreign soil.

The bulk of micro interactional research focuses on individuals’ processes of interpreting grievances. These processes refer to the means by which people collectively arrive at similar definitions of a situation or “interpretive frames” regarding social changes they support or oppose (Snow et al. 1986). Aggrieved but previously unmobilized people must revise the manner in which they look at a problematic condition or aspect of life; social arrangements must come to be seen as “unjust and mutable” (Piven and Cloward 1977, p. 12). This process of cognitive liberation typically involves an attributional shift from blaming oneself to blaming the system for particular problems (McAdam 1982). The expression of these understandings—the definitions of problems, protagonists, antagonists, ideas for change, and reasons for action—constitutes a movement’s collective action frame.

As a cycle of contention continues, movement emergence is influenced indirectly by the move-

ments that emerged early in the cycle. In addition to the structural influences, cultural influences, and need for cognitive liberation mentioned above, latecomers in a cycle must align their collective action frame with a master frame (Snow and Benford 1992). Early collective action frames may gain in both attention and popularity; they may resonate with the audience for which they are intended. In the minds of individuals, these frames are translated into generic codes that indicate how both audiences and movements should understand and react to events. Later attempts to extend the frames by adding further diagnoses, prognoses, and rationales for action may be met with resistance, constraining the emergence of a new movement.

In sum, social movements are most likely to emerge when the structural conditions for mobilization are ripe, cultural contexts provide a common worldview and set of activities to be applied to the situation, the collective interpretation of grievances produces cognitive liberation, and, if necessary, collective action frames are aligned to at least a minimal degree with a master frame.

### MOVEMENT PARTICIPATION

Closely related to the issue of movement emergence are questions regarding movement participation: Who joins and why? What conditions affect the likelihood of participating? How do movements build membership and sustain participation? Initial attempts to address questions about movement participation were influenced by breakdown theories. Movement participation was viewed as an irrational response to social structural strains. The factors regarded as key determinants of movement participation ranged from alienation and social isolation to status strains and relative deprivation. Each of these approaches suggested that some sort of psychological malaise or personality defect predispose some individuals to react to structural strains by participating in social movements.

The outburst of collective action and the proliferation of social movements in the 1960s led many social scientists to reconsider the assumptions of breakdown theory. Some theorists redefined movement participation as a rational choice. According to this perspective, people take part in

social movement activities only when they perceive that the anticipated benefits outweigh the expected costs of participation (Klandermans 1984). Research on the conditions affecting cost-benefit participation decisions indicates that this is a complex process that involves numerous structural and social psychological factors (Snow and Oliver 1995).

Social networks also have a crucial effect on differential recruitment to social movements. Movements tend to recruit the majority of their new members from the networks of existing members (Snow et al. 1980). A person typically decides to attend her or his first movement function because a friend, coworker, or relative invited her or him. Those outside such networks are less likely to be aware of the existence of specific movement groups. They also are less likely to attend a movement function if they are not sure there will be others present whom they know.

Overlapping networks in a community increase the probability that an individual will participate in a social movement. For instance, during the 1871 Paris uprisings, persons who were in Parisian National Guard units drawn from their own neighborhoods were more likely to defect and join the communal revolution. Participation increased further when adjacent neighborhoods had similar overlapping networks. In short, the interaction and intricacy of multiple networks increased the likelihood of social movement participation (Gould 1991).

While having social ties to people who are movement participants increases the likelihood of movement participation, other social ties can diminish that probability. Social ties in the form of family and career attachments can constrain movement participation in a number of ways. For one thing, these competing commitments may result in role conflict. The demands of being a movement participant and the demands of being a parent or employee may be incompatible. Married persons who have parental responsibilities as well as full-time jobs may not have sufficient discretionary time to participate in social movements (McCarthy and Zald 1973). Furthermore, spouses and employers can be displeased by a person's participation in a social movement.

To justify their movement participation to themselves and others, participants develop vo-

cabularies of motives. These are rationales that offer compelling reasons for their participation, particularly when their actions are called into question by employers, family members, or friends. Movement participants socially construct these vocabularies of motive as they interact with one another. Activists in turn employ these rationales to encourage sympathizers and adherents to take action on behalf of movement goals.

Vocabularies of motives not only facilitate recruitment to movements but also serve as commitment-building mechanisms. They help participants justify to themselves making sacrifices for a cause. The more sacrifices the participants make, the more costly leaving the movement seems to be. As they relinquish old attachments in favor of new ones, their commitment grows deeper (Kanter 1968). Research indicates that contrary to popular myths regarding participation in new religious movements and cults, these conversion and commitment-building processes are typically voluntary (Snow and Machalek 1984).

Closely related to social ties and vocabularies of motives is the concept of collective identity. Usually based on shared values, beliefs, and personal identities, a collective identity refers to the qualities and characteristics attributed to a group by the members of that group (Hunt 1991). Movement actors develop this sense of "weness" or "groupness" in the course of participating in social movement activities. Participants who have made an emotional commitment to a movement "communicate, influence each other, negotiate and make decisions" (Melucci 1995, p. 45). During these interactions with others, participants continually create and re-create consensus on a movement's goals, strategies, and sites of activity.

Movements generate their collective identity in part by articulating the ways in which movement goals and interests appear to be aligned with the beliefs and values of potential supporters. In the case of the 1989 Chinese Democracy Movement, students framed their goals in ways that were consistent with traditional Chinese cultural narrations. Drawing on Confucian principles, communist ideology, and the rhetoric of nationalism, the students fashioned collective action frames and protest tactics that were concerned with patriotism and the way in which it was defined and drama-

tized. Through their words and deeds, the student demonstrators conveyed their deep sense of responsibility to their country and willingness to offer themselves in sacrifice for the greater good. Those framings tended to resonate well among the general population. Consequently, the Chinese Democracy Movement spread rapidly from students to ordinary citizens (Zuo and Benford 1995).

The emotional component is also important, since it can be changed through persuasion and thus can become a powerful motivation for initial involvement in a movement. Animal rights groups, for example, use pictures of stabbed bulls, starved dogs, clubbed baby seals, and cats with electrodes implanted in their skulls to create anger and draw recruits. Emotions also can help sustain participation. For instance, the songs and dances of Diablo Canyon nuclear power plant protestors in California generated a degree of bonding that helped sustain activity at the protest site (Jasper 1997).

By presenting images of movement participation, the mass media play an important role in the recruitment of individuals. Mass media allow the public to form a response to a potential social problem quickly because events are covered instantaneously. The media provide movements with a larger audience and at times are the only resource people have in constructing the meaning of an event. Although media representations of reality are filtered through people's experiences, the media serve as gatekeepers of information and thus exercise considerable influence on the framing of social problems and thus social movement recruitment.

Social movements do not necessarily rely exclusively on traditional media, however; sometimes movement activists devise their own means of communicating with their target audiences. In the Chinese Democracy Movement, the "illegitimate" status of the student protests precluded student activists from accessing major state-controlled media outlets such as newspapers, television, and radio. During the "crisis," the state even cut off student telephone and telegram services in most Beijing universities. To cope with communication problems, student leaders devised a number of creative means of communication, including protest notices and posters reporting the latest movement decisions and suggestions on campus

building walls and bulletin boards, bicycles to relay strategic and tactical information between campuses, pirate radio broadcasts, E-mail, fax machines, and on-campus speeches and press conferences, to mobilize additional support (Zuo and Benford 1995).

Taken together, research on social movements reveals that participation factors, motives, and experiences are diverse. No single explanation can account for movement participation. Instead, a confluence of factors affect the decision to participate. Similarly, there are a variety of ways in which individuals may participate, ranging from those which require little commitment of time, such as signing a petition and writing a letter to a political official, to those requiring extensive commitment, such as coordinating national campaigns and committing acts of civil disobedience.

#### MOVEMENT ORGANIZATIONS

The activities of movements and their participants are coordinated by social movement organizations (SMOs). These organizations vary in a number of ways. An important way in which they differ relates to their origins. Some SMOs are organized at the grassroots level by people directly affected by a particular social problem. For example, a woman whose child had been killed by a drunken driver founded Mothers Against Drunk Driving (MADD). Other SMOs are established and sustained by powerful elites from the top downward. Before the disintegration of the Soviet bloc, the state orchestrated official "peace movements" in several communist countries.

Social movement organizations also vary in terms of the level or levels at which a group operates. Some SMOs operate at only a local level, focusing their attention solely on community issues. Others operate primarily at a state, provincial, or regional level, mobilizing around issues that affect that jurisdiction. Still others mobilize at a national level, often attempting to affect national decision making, policies, and legislation. Finally, some SMOs operate in several countries simultaneously. Many of those transnational SMOs focus on issues of human rights (Smith 1995).

Finally, SMOs vary in terms of how they are structurally linked to one another. At one end of

the continuum, SMOs are formally linked to a central authority, usually a national or international SMO. Local Amnesty International chapters, for example, must follow specific guidelines and rules dictated by headquarters. At the other end of the continuum, SMOs are relatively autonomous, not answering to any central authority beyond their own group. Between those two poles, the social movement sector yields a variety organizational arrangements, including loosely federated clusters of SMOs, ad hoc coalitions, and more permanent coalitions.

Most general reform movements spawn numerous SMOs. For instance, by 1984, the U.S. nuclear disarmament movement included some 3,000 independent SMOs as well as another 1,000 local chapters of national organizations. Specific reform movements, by contrast, tend to generate fewer SMOs. Regardless, SMOs are formal groups that can be thought of as the command posts of movements. They acquire and deploy resources, mobilize adherents, and plan movement strategy.

Resource mobilization theorists were among the first to emphasize the importance of SMOs in performing these functions. In particular, they point out that in the absence of an organization, it is difficult for movements to acquire the resources needed to sustain their challenges (Tilly 1978). Contemporary movements require money for advertising, printing, postage, lobbying, staff, and the like.

Other resource mobilization theorists have suggested that studying SMOs reveals how the macro and micro levels are reciprocally linked (McAdam et al. 1988). For example, the resource level of a society affects the resources available to SMOs, which in turn affect recruitment efforts (McCarthy and Zald 1977). In times of economic prosperity, such as the 1960s, the entire social movement sector expands because there are more discretionary resources available for movements in those periods. In this illustration, the macro level (a society's surplus resources), mediated by SMOs, affects the micro level (individual participation).

However, many movements also try to affect the macro level from below, with SMOs again playing a mediating role. Individuals with similar grievances get together in an informal, small group

setting, what McAdam (1988) refers to as a "micro-mobilization context." Sometimes the participants in those ad hoc meetings decide to establish a more formal, enduring organization (i.e., an SMO) to act on their collective grievances. The SMO in turn devises a strategy aimed at changing the system in some way. Occasionally, SMOs succeed in bringing about macro-level changes.

The strategies and tactics a movement employs in pursuit of its objectives typically are selected or devised by SMOs. A movement strategy refers to the broad organizing plans for the acquisition and use of resources toward achieving movement goals. For instance, as was suggested above, movements may pursue social change by devising strategies aimed at changing structural arrangements, strategies aimed at changing people, or both. Similarly, movements may choose between legal and illegal strategies and between violent and nonviolent strategies.

Tactics refer to the specific techniques movements employ to carry out their strategies. Teach-ins, sit-ins, marches, rallies, strikes, and mass mailings are only a few of the tactics contemporary reform movements typically utilize. There appears to be considerable tactical borrowing across the political spectrum. Conservative movements of the 1980s and 1990s in the United States, such as the Moral Majority and the antiabortion movements, for example, employed many of the tactics originally developed by the civil rights and New Left movements of the 1960s.

Tilly made a similar observation regarding eighteenth-century American revolutionary movement tactics. He accounted for tactical similarities across movements and SMOs by noting that every place and time has limited "repertoires of collective action" that are well defined but limited compared with the various theoretically available tactical options. These "standard forms are learned, limited in scope, slowly changing, and peculiarly adapted to their settings" (Tilly 1979, p. 131).

While tactical diffusion across movements and SMOs occurs, a division of tactical labor also commonly arises within movements. Each SMO tends to develop its own specific tactical preferences and expertise. These specializations arise as a consequence of cooperation and competition among the various SMOs that constitute a movement. By

refining and employing specialized tactics, an SMO is able to carve out a niche the movement that distinguishes it from other movement organizations.

Once an SMO establishes an organizational identity, it can build a stable resource base. Some SMOs have been so successful in that regard that they have survived the decline of a movement. Research on the women's movement indicates that such "abeyance organizations" provide continuity from one cycle of movement activity to the next (Taylor 1989) by sustaining activist interaction and commitment during periods when the opportunity structures are unfavorable to mass mobilization. In sum mary, SMOs contribute stability to what is otherwise a fluid, emergent phenomenon.

### MOVEMENT OUTCOMES

What effects, if any, do social movements have on social change? This crucial question is not as easy to answer as might be assumed. Because of the difficulties associated with studying a large sample of movements, most researchers study movements one at a time. Although these case studies provide researchers with rich, detailed data on specific movements, they are not helpful in making generalizations. However, even in case studies, the question of the effects of a particular movement is difficult to answer. First, the logic employed is counterfactual (Moore 1978). That is, in evaluating the effects of a particular movement, researchers have to speculate about what the outcome would have been if that movement had *not* existed. Second, the effects of movements are not always immediate and apparent. Some movements, such as the civil rights and women's movements, produce rippling effects that gradually engulf societal institutions, sometimes generating effects several decades after a movement's most intense period of agitation has ended.

To evaluate the outcomes of a movement, researchers examine its explicit and implicit goals, the direction of those goals, and the intended and unintended outcomes of attempting to reach those goals. For example, while the women's movement has ostensibly been geared toward enacting and changing policy, there is also an underlying goal of raising the consciousness of society concerning

women's issues (e.g., women's health, reproductive rights, violence against women, employment). Movement success may be evaluated with respect to a variety of dimensions, including a movement's ability to mobilize people to act, the diffusion of ideas across many cultures or countries, changes in a specific culture and individual sensibilities, and social policy changes. However, while a movement may succeed in some areas, it may fail in others. On the one hand, the women's movement could be considered successful in that it mobilized women both in the United States and elsewhere to take part in the struggle for rights and brought issues such as sexual harassment and unequal occupational status into the open. On the other hand, it has failed to persuade Congress to enact the Equal Rights Amendment, and despite popular myths to the contrary, women still suffer gross inequities in the workplace and at home.

In general, research suggests that movements seem to be more effective in producing cultural than structural change. The enduring legacy of the movements of the 1960s, for example, appears to be cultural. These cultural changes are reflected in attitudinal shifts regarding women and minorities, fashion trends (e.g., blue jeans), popular music, hedonistic lifestyles (e.g., the proliferation of illicit drugs), and the like. By contrast, these movements have had negligible success in terms of structural changes. While civil rights legislation helped dismantle caste restrictions and nearly equalized voting rights in the South, African-Americans continue to suffer "grinding poverty" and "persistent institutional discrimination in jobs, housing, and education" (McAdam 1982, p. 234). Women have realized even fewer structural gains. Finally, the sweeping changes in the economic, political, and educational institutions advocated by student activists never came to pass.

Social movements have been able to affect the sensibilities of both localized and broad publics. Researchers have argued that part of the function of social movements is to make use of and spread the knowledge created in various institutions (Eyerman and Jamison 1991). In the 1950s, a local grassroots environmental movement in Minamata, Japan, for instance, made use of rallies, disruptive protests, legal action, and increasingly favorable media attention during a cycle of contention to bring attention to what medical authorities called

“Minamata disease,” a methyl mercury poisoning caused by local industrial waste that contaminated the fish the local residents ate. Knowledge of how industries affect the inhabitants of the areas in which they operate became a matter of national attention by the late 1960s. However, the goal of obtaining local control over pollution eventually failed in the face of the nation’s need to sustain economic growth through increased national influence on industry after the mid-1970s (Almeida and Stearns 1998). Nevertheless, success may be judged on the basis of the movement’s ability to open lines of communication with the public and its use of those lines to spread new knowledge as it attempts to affect social change.

It should not be inferred, however, that movements always fail to achieve their structural objectives. The mid-nineteenth-century abolitionist movement succeeded in abolishing slavery. Over a century later, the global movement against apartheid in South Africa yielded dramatic successes. In the late 1980s and early 1990s, grassroots movements radically transformed the totalitarian political structures of a number of eastern European countries into more democratic states. Similarly, the transnational social movement organization Greenpeace has aided in the creation, enforcement, and increasing support for international policies controlling the trade in toxic waste. In short, although movements occasionally achieve dramatic outcomes, social structures initially tend to be more resistant than cultures to the revolutionary or reform efforts of social movements.

Gamson (1990) is one of the few researchers who have attempted to identify systematically the conditions under which social movements are likely to achieve their objectives. He traced the activities of a representative sample of fifty-three “challenging groups,” SMOs that emerged in the United States between 1800 and 1945. Gamson measured the relative success or failure of those SMOs in terms of whether they (1) gained new advantages and/or (2) gained acceptance from their antagonists. He found that thirty-one (58 percent) of them gained new advantages or acceptance while twenty (38 percent) gained both.

One of Gamson’s strongest findings pertained to the degree of change advocated. Movement groups that sought to displace extant elites rarely succeeded. Gamson reported that the SMOs most

likely to succeed exhibited the following characteristics: selective incentives for participants (some form of inducement, including rewards and punishments, to participate); unruly tactics (e.g., strikes, violence), especially when the target was relatively weak; bureaucratic, centralized organizational structures; and the absence of factional splits in a group. Although Gamson’s research has been criticized as being too simplistic, it identifies several factors that affect the outcomes of social movements.

In examining the effectiveness of a social movement, it is important to see that success in changing worldviews can be linked to success in changing social structures. Movement activists devote considerable time to the task of transforming the ways in which people view or frame a social issue or domain of life: their interpretive frames (Snow et al. 1986). If a movement’s framing efforts are successful either locally or on a global scale, a general shift in public opinion can occur, as has been the case for the movement against drunk driving (Gusfield 1981). Drivers who once were thought of as foolish or careless have been redefined as “killer drunks.” Subsequently, the movement has found it relatively easy to secure legislation raising minimum drinking ages and increasing the penalties for driving under the influence of alcohol. Although favorable public opinion is not a sufficient condition for social change to occur, it can lead to advantageous changes in the opportunity structure as well as the availability of resources.

Social movements may not always succeed in achieving their goals. Movements have, however, played a significant role in changing the way their members understand their world, the way others understand their world, and most societal reforms, revolutions, and changes in the world order.

#### REFERENCES

- Almeida, Paul, and Linda Brewster Stearns 1998 “Political Opportunities and Grassroots Environmental Movements: The Case of Minamata.” *Social Problems* 45:37–60.
- Blumer, Herbert 1946 “Collective Behavior.” In A. M. Lee, ed., *A New Outline of the Principles of Sociology*. New York: Barnes and Noble.
- della Porta, Donatella, and Hanspeter Kriesi 1999 “Social Movements in a Globalizing World: An Introduction.” In D. McAdam, J. D. McCarthy, and M. N. Zald, eds., *Comparative Perspectives on Social Movements:*

- Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, UK: Cambridge University Press.
- Eisinger, Peter K. 1973 "The Conditions of Protest Behavior in American Cities." *American Political Science Review* 67:11-28.
- Eyerman, Ron, and Andrew Jamison 1991 *Social Movements: A Cognitive Approach*. University Park: Pennsylvania State University Press.
- Ferree, Myra Marx, and William A. Gamson 1999 "The Gendering of Abortion Discourse: Assessing Global Feminist Influence in the United States and Germany." In D. McAdam, J. D. McCarthy, and M. N. Zald, eds., *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, UK: Cambridge University Press.
- Freeman, Jo 1973 "The Origin of the Women's Liberation Movement." *American Journal of Sociology* 78:792-811.
- Gamson, William A. 1990 *The Strategy of Social Protest*, 2nd ed. Belmont, Calif.: Wadsworth.
- Giddens, Anthony 1990 *The Consequences of Modernity*. Cambridge, UK: Polity Press.
- Gould, Roger V. 1991 "Multiple Networks and Mobilization in Paris Commune, 1871." *American Sociological Review* 56:716-729.
- Gusfield, Joseph R. 1981 *The Culture of Public Problems: Drinking-Driving and the Symbolic Order*. Chicago: University of Chicago Press.
- Hunt, Scott A. 1991 "Constructing Collective Identity in a Peace Movement Organization." Ph.D. dissertation, Department of Sociology, University of Nebraska, Lincoln.
- Jasper, James 1997 *The Art of Moral Protest: Culture, Biography, and Creativity in Social Movements*. Chicago: University of Chicago Press.
- Jenson, Jane 1995 "What's in a Name? Nationalist Movements and Public Discourse." In H. Johnston and B. Klandermans, eds., *Social Movements and Culture*. Minneapolis: University of Minnesota Press.
- Kanter, Rosabeth M. 1968 "Commitment and Social Organization: A Study of Commitment Mechanisms in Utopian Communities." *American Sociological Review* 33:499-517.
- Klandermans, Bert 1984 "Mobilization and Participation: Social-Psychological Expansions of Resource Mobilization Theory." *American Sociological Review* 49:583-600.
- Koopmans, Ruud 1999 "A Comparison of Protests against the Gulf War in Germany, France, and the Netherlands." In D. della Porta, H. Kriesi, and D. Rucht, eds., *Social Movements in a Globalizing World*. London: Macmillan.
- Kriesi, Hanspeter 1996 "The Organizational Structure of New Social Movements in a Political Context." In D. McAdam, J. D. McCarthy, and M. N. Zald, eds., *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, UK: Cambridge University Press.
- Lo, Clarence Y. 1982 "Counter-movements and Conservative Movements in the Contemporary U.S." *Annual Review of Sociology* 8:107-134.
- McAdam, Doug 1982. *Political Process and the Development of Black Insurgency, 1930-1970*. Chicago: University of Chicago Press.
- 1988. "Micromobilization Contexts and Recruitment to Activism." *International Social Movement Research* 1:125-154.
- , John D. McCarthy, and Mayer N. Zald 1988 "Social Movements." In N. Smelser, ed., *Handbook of Sociology*. Newbury Park, Calif.: Sage.
- , and Dieter Rucht 1993 "Cross-National Diffusion of Social Movement Ideas." *Annals of the American Academy of Political and Social Science* 528:56-74.
- McCarthy, John D., and Mayer N. Zald 1973 *The Trend of Social Movements in America: Professionalization and Resource Mobilization*. Morristown, N.J.: General Learning Press.
- 1977 "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82:1212-1241.
- Melucci, Alberto 1980 "The New Social Movements: A Theoretical Approach." *Social Science Information* 19:199-226.
- 1995 "The Process of Collective Identity." In H. Johnston and B. Klandermans, eds., *Social Movements and Culture*. Minneapolis: University of Minnesota Press.
- Moore, Barrington, Jr. 1978 *Injustice: The Social Bases of Obedience and Revolt*. White Plains, N.Y.: Sharpe.
- Piven, Frances Fox, and Richard A. Cloward 1977 *Poor Peoples' Movements: Why They Succeed, How They Fail*. New York: Vintage.
- Rucht, Dieter 1996 "The Impact of National Contexts on Social Movement Structures: A Cross-Movement and Cross-National Comparison." In D. McAdam, J. D. McCarthy, and M. N. Zald, eds., *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, UK: Cambridge University Press.
- Smelser, Neil J. 1962 *Theory of Collective Behavior*. New York: Free Press.

- Smith, Jackie 1995 "Transnational Political Processes and the Human Rights Movement." *Research in Social Movements, Conflict and Change* 18:187-221.
- Snow, David A., and Robert D. Benford 1992 "Master Frames and Cycles of Protest." In A. D. Morris and C. M. Mueller, eds., *Frontiers in Social Movement Theory*. New Haven, Conn.: Yale University Press.
- , and — 1999 "Alternative Types of Cross-National Diffusion in the Social Movement Arena." In D. della Porta, H. Kriesi, and D. Rucht, eds., *Social Movements in a Globalizing World*. London: Macmillan.
- , and Richard Machalek 1984 "The Sociology of Conversion." *Annual Review of Sociology* 10:167-180.
- , and Pamela E. Oliver 1995 "Social Movements and Collective Behavior: Social Psychological Dimensions and Considerations." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- , E. Burke Rochford, Jr., Steven K. Worden, and Robert D. Benford 1986 "Frame Alignment Processes, Micromobilization, and Movement Participation." *American Sociological Review* 51:464-481.
- , Louis A. Zurcher, Jr., and Sheldon Eklund-Olson 1980 "Social Networks and Social Movements: A Microstructural Approach to Differential Recruitment." *American Sociological Review* 45:787-801.
- Staggenborg, Suzanne 1987 "Life-Style Preferences and Social Movement Recruitment: Illustrations from the Abortion Conflict." *Social Science Quarterly* 68:779-797.
- Tarrow, Sidney 1998 *Power in Movement: Social Movements and Contentious Politics*, 2nd ed. Cambridge, UK: Cambridge University Press.
- Taylor, Verta 1989 "Social Movement Continuity: The Women's Movement in Abeyance." *American Sociological Review* 54:761-775.
- Tilly, Charles 1978 *From Mobilization to Revolution*. Reading, Mass. Addison-Wesley.
- 1979 "Repertoires of Contention in America and Britain, 1750-1830." In M. N. Zald and J. D. McCarthy, eds., *The Dynamics of Social Movements*. Cambridge, Mass.: Winthrop.
- , Louise Tilly, and Richard Tilly 1975 *The Rebelious Century: 1830-1930*. Cambridge, Mass.: Harvard University Press.
- Zdravomyslova, Elena 1996 "Opportunities and Framing in the Transition to Democracy: The Case of Russia." In D. McAdam, J. D. McCarthy, and M. N. Zald, eds., *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, UK: Cambridge University Press.
- Zuo, Jiping, and Robert D. Benford 1995 "Mobilization Processes and the 1989 Chinese Democracy Movement." *The Sociological Quarterly* 36:131-156.

ROBERT D. BEAFORD  
TIMOTHY B. GONGAWARE  
DANNY L. VALADEZ

## SOCIAL NETWORKS

Social networks—structures of relationships linking social actors—are omnipresent in contemporary society. People often obtain information about such things as job opportunities, housing, and medical care through interpersonal contacts rather than from formal sources such as the mass media. Networks provide emotional support in times of crisis as well as instrumental aid such as help with household tasks. Identities are constituted by locations in networks; opinions are formed and decisions are made in light of information and conformity pressures that flow through network linkages. Also, social networks are important channels through which both infectious diseases and innovations are diffused.

Ties among individuals in social networks give rise to important larger-scale social patterns. Levels of socioeconomic or ethnoreligious segregation in a society, for example, reflect the degree to which social ties such as marriage and friendship are confined to sets of persons with a common social status or heritage. Such characteristics become salient as markers of differentiation to the degree that they serve as bases for the formation of intimate social relationships. Networks that link individuals to supraindividual units such as work organizations and voluntary associations may serve as modes of integration or separation. Thus, the study of networks contributes to the linking of micro and macro levels of analysis in sociology.

Many macro-level social phenomena can be understood as networks. Increasingly, systems of production for goods (such as automobiles) and services (such as care for the severely mentally ill) are located in multiorganizational fields consisting of separate but interdependent units linked by contingent cooperation rather than in self-contained units administered through elaborate formalized sets of rules. Innovations in corporate strategy and governance diffuse through overlap-



ping boards of directors and other interorganizational structures. Patterns of consensus and cleavage in community and national politics are shaped by alliances and conflicts in networks of governmental agencies, interest groups, and party organizations. In international relations, network ties among nation-states and international organizations define geopolitical alignments.

Emphasis on social networks grew as a result of substantive observations about contemporary society. Many early twentieth-century observers posited that large-scale transformations associated with industrialization—especially urbanization, bureaucratization, and the development of mass media—led to a “mass society” of atomized individuals in which formal, special-purpose ties supplanted diffuse interpersonal relations.

Several lines of research, however, pointed to the continuing vitality of social ties. Industrial sociologists found that informal structures were crucial to the day-to-day functioning of work organization. Indeed, workplace social networks came to be seen as a solution to the inflexibility and excess formalization of bureaucracies and as important incentives for (or impediments to) the performance of individual employees. Urban sociologists found that friendship, neighboring, and informal assistance remained prominent in large cities, although technological innovations in transportation and communication reduced the extent to which the formation and maintenance of those social ties were constrained by spatial considerations. Rural-to-urban and international migrants were not rootless citizens in a normless society but tended to settle in districts populated by persons from their places of origin. In contrast to the expectations of theories predicting protest and activism among those marginal and peripheral to society, researchers found that those active in social movements tended to be drawn from among the persons best integrated into communities, and social ties proved to be important channels through which new members were recruited to social movements.

A social network perspective highlights the interdependence among social actors. This extends beyond the competitive interdependence of actors competing for shares of a stock of scarce resources to encompass obligations and commitments that accumulate as a result of past social

interaction. Individual action is embedded in, and therefore continually affected by, preexisting ties between specific actors (Granovetter 1985). In some theories, individuals are viewed as largely passive recipients of environmental pressure; in this structural emphasis, social networks constitute constraints that limit an actor’s discretion. An alternative view makes more room for individual agency, viewing networks as structures of opportunity or social resources. Assuming a context of constrained voluntarism, it treats individuals as proactive, self-interested agents who use networks to manipulate outcomes to their advantage (Haines 1988; Emirbayer and Goodwin 1994).

Social scientists have used the term “social network” metaphorically for some time. Beginning in the 1970s, however, scholarly attention to the analytic development of the social network approach increased. It is now seen as a distinct specialty within several social science disciplines, especially sociology and anthropology, and has many adherents in professional schools, particularly those of business administration and public health. Mathematicians and statisticians have participated in this work, especially the development of novel techniques for studying relations between interdependent social units. These methods are distinct from the conventional ones used to study relations between variables within presumably autonomous units. The theoretical assumptions of formal network models are often implicit (Granovetter 1979), and a distinct “network theory” has not developed. Contemporary studies of social networks instead draw on diverse sociological and social psychological theories.

## PRECURSORS

Some foundations of a methodology for studying networks of social relations were laid in Moreno’s *Who Shall Survive?* (1934). Moreno coined the term “sociometry” to refer to methods for describing group structures and individual positions within them. His work focused on the affinities and disaffinities of individuals for one another, and his “sociometric test” accordingly stressed affective choices and rejections. These network data were mapped in “sociograms” in which persons were located at different points, with lines indicating the connections between them. Such graphic representations are also common in contemporary

network analysis (Figure 1). On the basis of their locations in sociometric networks of affect, individuals were classified as attractive, isolated, rejected, and so forth.

Jennings collaborated with Moreno in developing sociometric methods; her (1943) work reported studies of attractiveness and emotional expansiveness. Sociometry as practiced by Moreno and Jennings had an applied component: They attempted to use sociometric measurements as a basis for rearranging groups to enhance both group functioning and individual creativity.

Displays such as sociograms are extremely useful as visualization devices but do not facilitate formal analysis. Forsyth and Katz (1946) and Luce and Perry (1949) were among the first to attempt to surmount the limitations of sociograms by representing networks in the form of matrices. They argued that this would reduce the subjectivity of statements about sociometric structure and allow objective identification of chains, groups, and cliques in network data. In later studies, the application of graph theory (Harary et al. 1965) to the analysis of networks has extended those early efforts.

A second set of influences on the development of social network analyses has emanated from the fieldwork of social anthropologists in complex societies. Those analysts observed that the categorical concepts of a structural-functional approach were insufficient for the study of societies in which not all behavior was regulated by “corporate groups”—institutions of kinship, community, or work. Barnes (1954) is credited with the first use of the term “social network” to refer to a set of existing social relationships as distinct from cultural prescriptions about the construction of such ties. In its initial formulations, the concept was used to refer to informal or extrainstitutional links, but it was soon noted that formalized relations and groups also could be analyzed as networks of interactions.

Classic studies in the anthropological tradition display the same variability in theoretical orientations seen in present-day work: Bott (1957) treated social networks as sources of norms prescribing an appropriate allocation of tasks between spouses, while Boissevain (1974) stressed the potential use of networks by maneuvering, self-interested entrepreneurs.

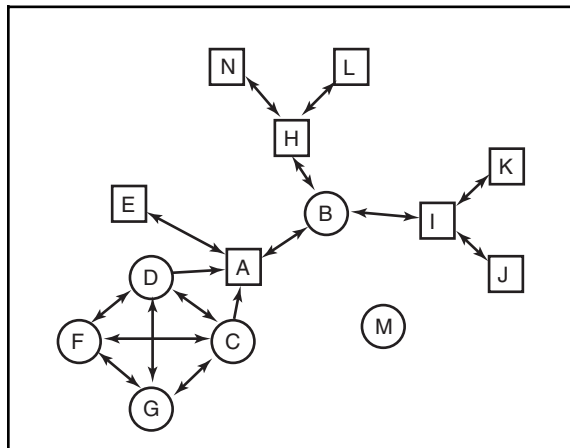


Figure 1. A Social Network Diagram

### STRUCTURAL PROPERTIES OF NETWORKS

Social networks are studied from many standpoints, including the sociocentric and egocentric perspectives. Researchers with a sociocentric orientation examine complete networks of actors and relations, studying the global properties of a network and characterizing the position of any given actor by reference to all the others. Egocentric network analysis takes the perspective of individual actors and the “personal networks” surrounding them, focusing on the local structure of each actor’s interpersonal environment.

Figure 1 depicts a complete network with fourteen actors, with each one represented as a circle or square. Lines connecting those points indicate relationships between pairs of actors, such as communication ties and expressions of affect. An actor’s “first-order” egocentric network consists of the other actors to which it is linked and the relationships among them; actor G’s egocentric network includes C, D, and F, while that of H consists of B, L, and N. The “second-order” zone of a focal actor’s egocentric network includes those to which the actor is linked via one intermediary; the second-order zone of G’s network consists of actor A, while that of H includes both A and I.

Arrows indicate directionality; for example, actors A and E are involved in a *reciprocal* relationship, while A and C are linked by an *asymmetric* tie in which the relationship flows in only one direction. *Indirect* relationships join actors to one an-

other through intermediaries: for instance, H and I are indirectly linked through B.

The *density* of a network reflects the overall intensity of connectedness among actors. In Figure 1, sixteen of the ninety-one distinct pairs of actors have direct relationships, and so the network has a density of 18 percent. There are, however, wide variations in the density of the egocentric networks surrounding actors: G is in a very dense or closely knit locality, while H is in a sparse or loosely knit region.

Diagrams of more complicated networks might use lines of different thicknesses to represent relationships of varying intensity or different types of lines to show distinct relationships, such as providing emotional support versus giving instrumental aid. Different types of actors can be represented by different symbols; the actors in circles in Figure 1 might be men, while those in squares might be women. When a relationship typically joins actors that have the same attributes or statuses, as in Figure 1, it is said to display *homophily*.

Social ties, especially those involving positive sentiment, often create *transitive* configurations. In Figure 1, for example, transitivity implies that by virtue of the links between C and D and between C and F, there will also be a tie between D and F. Strong tendencies toward transitivity create closure and tend to fragment a group into distinct *cliques*: The mutually interconnected set of four actors (C, D, F, and G) provides an illustration. Relationships that do not conform to the transitivity principle are important sources of integration between otherwise separate parts of a social structure (Granovetter 1973): in Figure 1, the I-B and A-B ties exemplify *bridges* of this kind.

Actors may be *central* in a network because they are involved in many relationships, are in a controlling position between other pairs of actors, or are comparatively close to others (Freeman 1979). Sociometry referred to central actors as “sociometric stars.” In Figure 1, actors A and B occupy relatively central locations. Those such as E, J, and L are *peripheral* within the network. Actor M is not involved in any relationships and is said to be an *isolate*.

Two actors are said to occupy the same *position* in a social network when they have profiles of

relationships to other actors that are identical in a particular way (Borgatti and Everett 1992); these actors are therefore substitutable for each other from an observer’s standpoint. Two actors having the same relationships to the same others are *structurally equivalent*. In Figure 1, actors J and K both have a reciprocal relationship with actor I but no direct relationship with each other or any of the other actors. J and K are thus structurally equivalent, as are three other pairs: C and D, F and G, and L and N. Actors are said to be *role-equivalent* when they have the same types of relationships to the same types of other actors. Actors L and K are role-equivalent, but not structurally equivalent; although they occupy similarly peripheral locations in the system of relations, L’s direct link is to H, while K’s is to I.

Additional patterns often appear in networks that involve more than one type of social relationship. *Exchange* patterns arise when a flow of one type in one direction is directly or indirectly reciprocated by a flow of a second type. A *multiplex* pattern occurs when the relationship between two actors consists of two or more distinct strands, such as kinship and emotional support; a *uniplex* relationship has only one strand.

Concepts and methods for identifying subgroups within networks have drawn much attention. The two dominant approaches focus on cohesion and equivalence as grouping principles. Techniques that emphasize cohesion locate subsets of densely interconnected actors; fully connected cliques are a limiting case. Blockmodel analysis (White et al. 1976) and related positional methods group equivalent actors together. This yields considerable flexibility. Social positions can be defined by the common ties of actors to outsiders rather than by actors’ links to one another; for example, those in “broker” roles occupy mediating locations between other positions but are not necessarily linked to other brokers. Moreover, positional approaches are not confined to a single type of tie; they can identify subgroups on the basis of patterns in multistranded relations rather than focusing on a single type of tie extracted from its context, as approaches resting on cohesion generally must. Thus, positional analyses of international trade relations may examine flows of raw materials and flows of processed goods simultaneously.

### THE STRUCTURING OF NETWORKS

Rational choice explanations of the formation of network ties are based on exchange theory: Social relations form when actors depend on one another for resources. Related behaviorist accounts stress a reinforcement history. An exemplar is Blau's (1955) description of the exchange of advice for expressions of deference among coworkers in a bureaucracy. The terms of exchange and thus the relative power of the actors involved depend on the number of alternative actors who control different resources, the extent of unity among those with a given resource, and the parties' relative interests in outcomes controlled by others (Cook 1982; Burt 1980).

Exchange-theoretic reasoning provides a basis for some commonly observed micro-level network patterns. To maintain autonomy and avoid power disadvantages, actors should avoid asymmetric relations in favor of reciprocal ones. Multiplex ties in which relations of solidarity overlie instrumental exchange links can protect people against exploitation by actors in positions of power. These ideas have played an important role in resource dependence theories of interorganizational relations. Such theories suggest, for example, that interdependent organizations tend to form network ties such as long-term contracts, joint ventures, interlocking directors, and mergers (Pfeffer 1987).

Some attempts to explain tendencies toward homophily are preference-driven accounts in which people actively seek out similar associates. From this viewpoint, communication is easier if people share implicit premises regarding interaction and trustworthiness in the face of uncertainty is enhanced if partners can assume that they have shared interests and predispositions.

Other lines of theorizing about the sources of homophily stress the structure of opportunities for association. Feld (1981) observes that most relationships arise within "foci" of association such as families, neighborhoods, workplaces, and voluntary associations. When segregating processes create foci composed of persons with similar attributes, they create systematic biases toward homophily. Blau's macrosociological theory (Blau and Schwartz 1984) postulates that networks vary with opportunities for association. Blau shows, without any assumptions about preferences, that intergroup relations—the converse of homophilous ones—

are more likely for small than for large groups, when there is great inequality or heterogeneity in a population instead of little, and when different characteristics (e.g., socioeconomic status, race, ethnicity) that structure the formation of social ties are intersecting (or uncorrelated) rather than consolidated. The implications of these factors for intergroup relations depend on the degree to which they penetrate into substructures such as Feld's foci: Distributional effects on intergroup association are strongest when heterogeneity, inequality, and intersection lie within rather than between substructures.

Transitivity has been most intensively studied for relations of positive sentiment; there the theoretical case for it rests on balance theories that posit pressures toward cognitive consistency (Davis and Leinhardt 1972). Tendencies toward closure also may result from increased opportunities for contact resulting from the copresence of actors in a triad, or an actor in contact with two others may facilitate or serve as a guarantor for a venture involving the other two. Principles of expectation states theory predict transitivity in dominance relations (Fararo and Skvoretz 1986).

### CONSEQUENCES OF NETWORKS

A network perspective lends itself to the construction of theories at multiple levels. Contextual theories that examine the effects of an actor's position in a network on achievement, well-being, and other individual-level outcomes are common, and extensive research literatures have developed around some of them. Network entrepreneurship, diffusion and influence, social support, and power are among the most prominent themes in such works. Group-level theorizing about the properties of complete networks is less typical, although there have been important efforts in this direction. As Coleman (1990) stresses, group-level theorizing is a demanding enterprise.

Granovetter's (1973) discussion of "weak ties" has been a fruitful source of ideas for network analysts. Although Granovetter defines tie strength in terms of dyadic content (intimacy, intensity, exchange of services and time commitments), the "strength of weak ties" arises from their location within a network structure. Weak ties are less subject than strong ones to the transitivity pressures that induce closure and thus are more likely

to be bridges that join distinct subgroups, thereby serving as channels for integration and diffusion. Different contextual theories positing network effects may stress the virtues of either strong or weak ties.

The purposive use of networks as individual-level “social capital” is a prominent theme in writing about network effects. Much research has been done on how networks facilitate or impede instrumental actions, particularly job seeking. Granovetter (1995) stressed the informational advantages of wide-ranging networks composed of many weak ties. He reasoned that such networks are likely to connect actors to diverse information sources that provide novel information and access to powerful others; by contrast, persons in a densely connected clique are apt to have similar social standing and know similar things. Other writing in this vein emphasizes the content rather than the form of networks. Lin (1990) contends that networks composed of highly ranked contacts are the most advantageous to an actor. Weak ties may facilitate access to social resources, but the aid such contacts can provide reflects their power and influence rather than the type of channel that links them to the actor.

Burt (1992), developed influential ideas about network entrepreneurship. Like Granovetter, Burt stressed the benefits of loosely knit networks for obtaining information quickly. Structural holes refer to the absence of connections between an actor’s contacts in a sparse network; thus, in Figure 1 there are holes separating actor B’s contacts (A, I, and H). Beyond providing actors with the ability to acquire information sooner than competitors can, Burt argues that networks rich in structural holes confer control benefits. Actors at the middle of networks that bridge many holes gain autonomy and leverage over others by virtue of their unique interstitial positions. They are able to place their contacts in competition with one another, avoid excessive dependence, and negotiate favorable bargains. Empirical applications in diverse settings lend support to these ideas: Structural holes have proved advantageous in studies of competition for promotions and bonuses in firms and studies of comparative profit margins in manufacturing industries.

Studies of the role of networks in diffusion and influence processes echo debates over cohe-

sion and equivalence as bases for defining network subgroups (Marsden and Friedkin 1993). This work posits that networks provide actors with a basis for constructing reference groups and social comparisons (Erickson 1988). Approaches that stress the socializing potential of cohesive ties contend that people look to close associates for guidance toward appropriate attitudes or conduct in conditions of uncertainty. Network closure thus is viewed as a source of locally defined norms, as persons in dense networks respond to consistent conformity pressures from their strong ties.

Burt (1987) suggests an important alternative process of diffusion or influence, arguing that in seeking normative direction, actors look not to their close contacts but to their competitors. They engage in a process of role taking, examining the views held or the actions taken by structural peers who occupy similar positions within a system of relations. Conformity to norms here is less a matter of social pressure from the environment than the result of an actor’s mimicry of others (DiMaggio and Powell 1983).

An extensive research literature on the sociology of health and illness draws a connection between social networks, exposure to stressors, the availability of social support, and well-being (House et al. 1988; Thoits 1995). Social networks are “structural” elements that may facilitate access to supportive contacts but also may expose someone to additional stressors, conflicts, and demands; hence, the quality as well as the structure of ties within networks must be considered. A “direct effect” view states that receiving social support enhances health and well-being in all conditions. The “buffering hypothesis” suggests a conditional effect in which those with supportive networks have less severe responses to stressful life events.

Several mechanisms have been suggested to account for the way in which aspects of social networks translate into social support and measures of physical and mental well-being. Some of these suggestions have a Durkheimian emphasis, reasoning that people integrated into dense networks—that is, strongly tied to a number of partners who are strongly tied to each other—have better-defined social identities, stronger senses of internal control, and more positive self-evaluations, which in turn may lead to the use of more effective coping strategies under stress. Collins

(1988) suggests that network density indicates that an individual is integrated into the “interaction rituals” of a solidary group, a situation that produces moral sentiments and energies that enhance well-being.

It is relatively well established that the availability of a strong tie or confidant has health-promoting effects (Thoits 1995). Theories stressing the availability of support resources imply that the size of a personal network is linked to physical and psychological health. Large, diverse networks are thought to facilitate adjustment to change. Contacts in one’s network may offer companionship as well as both instrumental and emotional assistance; they can both provide health information and focus attention on it. Social contacts also may be sources of regulation and control, providing social feedback on one’s behavior and performance, discouraging harmful behaviors such as substance abuse, and encouraging beneficial ones such as adherence to treatment protocols. It is increasingly recognized that providing social support can be burdensome and that stress as well as aid may emanate from relationships within social networks.

Network theorists have given substantial attention to the connection between an actor’s location in a social network and social standing or prestige. Freeman’s (1979) discussion of the conceptual foundations of centrality measures focused on processes in communication networks. He observed that distinct measures are sensitive to communication activity, the capacity to control the communications of others, and the capacity to avoid the controlling actions of others. Many empirical studies have documented an association between centrality and manifestations of power or influence.

Exchange-theoretic approaches observe that actors with predominantly favorable exchange ratios with others are central within networks of dependency relations and hence acquire power (Cook 1982). No universal principle leads to that connection, however; instead, it depends on a particular form of exchange. For Cook this is “positively connected” or “productive” exchange in which actors must combine diverse resources to be successful; hence, the use of one exchange relation tends to encourage the use of others, and advantages emerge for those in intermediary “bro-

ker” positions. Willer (1992) terms these “flow” networks. For Coleman (1990), the requisite conditions include resource transferability or fungibility: For systems of social exchange to develop fully, actors must be able to transfer not only control over resources but also the right to further transfer such control.

Under other conditions of exchange, network positions have different consequences for power. Among these are what Cook (1982) calls “negatively connected” exchange networks and Willer (1992) terms “restricted” exchange networks. In these networks, resources do not flow through positions and the use of one relation rules out or makes less likely the use of others. Matching systems such as marriage and dating markets exemplify negatively connected or restricted networks. In these conditions no special advantages accrue to an actor in a central position; instead, power is concentrated in actors who can exclude others from exchanges (Markovsky et al. 1988). The capacity to exclude others can depend both on the nuances of a network’s structure and on the incentives or rules governing exchange in any given situation. Depending on such subtleties, there may be a nonlinear association between centrality and power; a network also may decompose into small substructures as some potential exchange relations fall into disuse.

Theorizing about network effects at the level of aggregates is less extensive than is that about networks as contexts for individual action. One line of work examines the implications of structural biases toward homophily or transitivity for the spread of diseases and innovations (Morris 1993). Another considers how the centralization or dispersion of networks shapes differences in system performance or effectiveness (Laumann and Knoke 1987). Still another deals with networks and collective action: Social density has been viewed as an infrastructure that allows latent “interest groups” to overcome social dilemmas (Marwell et al. 1988). Granovetter (1973) warns against excessive closure, however: Once formed, a coalition or “collective actor” may lack the social connections required for ultimate political success.

Some theorists have developed the notion of social capital as a collective rather than an individual property. Coleman (1990) suggests that closed social networks can create trust and enforce strong

norms that facilitate collective action. Actors in densely interconnected systems can expect to encounter one another frequently in the future, and a reputation for trustworthiness therefore becomes valuable. Moreover, frequent communication among densely linked actors means that information about a failure to honor obligations diffuses quickly and that sanctions can be applied rapidly. Coleman asserts that strong norms of trust are essential to the creation of "social credit": aid or assistance contributed in exchange for future compensation. Preexisting social networks thus are valuable as sources of social capital to the extent that they are appropriable for new purposes.

Different varieties of network theory make contrasting observations about network density. Density can integrate an actor into a subculture, provide a well-defined social identity, create and enforce norms, and promote the production of collective goods it simultaneously may subject one to conformity pressures and limit the diversity of one's affiliations. Reconciling these seemingly conflicting effects of density represents a challenge in the contemporary study of networks. Understanding the manner in which networks channel and block transactions among the elements of society will remain a crucial and intriguing task for twenty-first-century sociology.

#### FURTHER LITERATURE

A number of review articles and collections treat aspects of network analysis in greater depth. Mitchell (1974) provides a useful review of the anthropological approach. Wellman (1983) reviews basic principles from a sociological perspective. Review articles on substantive applications appear in Wasserman and Galaskiewicz (1994). The journal *Social Networks* (1978–present), edited by Linton Freeman, presents methodological advances and substantive studies. Marsden (1990) surveys the literature on measurement, and Wasserman and Faust (1994) provide a comprehensive review of analytic methods.

#### REFERENCES

- Barnes, J. A. 1954 "Class and Committees in a Norwegian Island Parish." *Human Relations* 7:39–58.
- Blau, Peter M. 1955 *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press.

- , and Joseph E. Schwartz 1984 *Crosscutting Social Circles*. New York: Academic Press.
- Boissevain, Jeremy 1974 *Friends of Friends*. New York: St. Martin's.
- Borgatti, Stephen P., and Martin G. Everett 1992 "Notions of Position in Social Network Analysis." In Peter V. Marsden, ed., *Sociological Methodology 1992*. Oxford, UK: Blackwell.
- Bott, Elizabeth 1957 *Family and Social Network* London: Tavistock.
- Burt, Ronald S. 1980 "Autonomy in a Social Topology." *American Journal of Sociology* 85:892–925.
- 1987 "Social Contagion and Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology* 92:1287–1335.
- 1992 *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Collins, Randall 1988 *Theoretical Sociology*. New York: Harcourt Brace Jovanovich.
- Cook, Karen S. 1982 "Network Structures from an Exchange Perspective." In Peter V. Marsden and Nan Lin, eds., *Social Structure and Network Analysis*. Beverly Hills, Calif.: Sage.
- Davis, James A., and Samuel Leinhardt 1972 "The Structure of Positive Interpersonal Relations in Small Groups." In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress*. Boston: Houghton Mifflin.
- DiMaggio, Paul, and Walter W. Powell 1983 "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–160.
- Emirbayer, Mustafa, and Jeff Goodwin 1994 "Network Analysis, Culture, and the Problem of Agency." *American Journal of Sociology* 99:1411–1454.
- Erickson, Bonnie H. 1988 "The Relational Basis of Attitudes." In Barry Wellman and S. D. Berkowitz, eds., *Social Structures*. New York: Cambridge University Press.
- Fararo, Thomas J., and John Skvoretz 1986 "E-State Structuralism: A Theoretical Method." *American Sociological Review* 51:591–602.
- Feld, Scott L. 1981 "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–1035.
- Forsyth, Elaine, and Leo Katz 1946 "A Matrix Approach to the Analysis of Sociometric Data: Preliminary Report." *Sociometry* 9:340–347.

- Freeman, Linton C. 1979 "Centrality in Social Networks I: Conceptual Clarification." *Social Networks* 1:215–239.
- Granovetter, Mark S. 1973 "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–1380.
- 1979 "The Theory-Gap in Social Network Analysis." In Paul W. Holland and Samuel Leinhardt, eds., *Perspectives on Social Network Research*. New York: Academic Press.
- 1985 "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481–510.
- 1995 *Getting a Job: A Study of Contacts and Careers*, 2d ed. Chicago: University of Chicago Press.
- Haines, Valerie A. 1988 "Social Network Analysis, Structuration Theory and the Holism-Individualism Debate." *Social Networks* 10:157–182.
- Harary, Frank, Robert Z. Norman, and Dorwin C. Cartwright 1965 *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: Wiley.
- House, J. S., D. Umberson, and K. R. Landis 1988 "Structures and Processes of Social Support." *Annual Review of Sociology* 14:293–318.
- Jennings, Helen Hall 1943 *Leadership and Isolation*. New York: Longmans, Green.
- Laumann, Edward O., and David Knoke 1987 *The Organizational State*. Madison: University of Wisconsin Press.
- Lin, Nan 1990 "Social Resources and Social Mobility." In Ronald L. Breiger, ed., *Social Mobility and Social Structure*. New York: Cambridge University Press.
- Luce, R. Duncan, and Albert D. Perry 1949 "A Method of Matrix Analysis of Group Structure." *Psychometrika* 14:95–116.
- Markovsky, Barry, David Willer, and Travis Patton 1988 "Power Relations in Exchange Networks." *American Sociological Review* 53:220–236.
- Marsden, Peter V. 1990 "Network Data and Measurement." *Annual Review of Sociology* 16:435–463.
- , and Noah E. Friedkin 1993 "Network Studies of Social Influence." *Sociological Methods and Research* 22:127–151.
- Marwell, Gerald, Pamela E. Oliver, and Ralph Prahl 1988 "Social Networks and Collective Action: A Theory of the Critical Mass III." *American Journal of Sociology* 94:502–534.
- Mitchell, J. Clyde 1974 "Social Networks." *Annual Review of Anthropology* 3:279–299.
- Moreno, J. L. 1934 *Who Shall Survive?* Washington, D.C.: Nervous and Mental Disease Publishing.
- Morris, Martina 1993 "Epidemiology and Social Networks: Modeling Structured Diffusion." *Sociological Methods and Research* 22:99–126.
- Pfeffer, Jeffrey 1987 "A Resource Dependence Perspective on Intercorporate Relations." In Mark S. Mizruchi and Michael Schwartz, eds., *Intercorporate Relations*. New York: Cambridge University Press.
- Thoits, Peggy A. 1995 "Stress, Coping, and Social Support Processes: Where Are We? What Next?" *Journal of Health and Social Behavior*, extra issue, pp. 53–79.
- Wasserman, Stanley, and Katherine Faust 1994 *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
- , and Joseph Galaskiewicz, eds. 1994 *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*. Thousand Oaks, Calif.: Sage.
- Wellman, Barry 1983 "Network Analysis: Some Basic Principles." *Sociological Theory* 1:155–199.
- White, Harrison C., Scott A. Boorman, and Ronald L. Breiger 1976 "Social Structure from Multiple Networks I: Blockmodels of Roles and Positions." *American Journal of Sociology* 81:730–780.
- Willer, David 1992 "Predicting Power in Exchange Networks." *Social Networks* 14:187–211.

PETER V. MARSDEN

## SOCIAL NORMS

See Compliance and Conformity; Deviance Theories; Socialization; Social Control; Social Values and Norms.

## SOCIAL ORGANIZATION

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

It is necessary to highlight this central area selectively, considering the breadth and variety in its specification, and to emphasize current concerns, relying on past overviews to detail their origins (e.g., by Znaniecki 1945; Gerth and Mills 1953; Faris 1964; Eisenstadt 1968; Parsons 1968; Udy 1968; Smelser 1988).

Social organization is nonrandom *pattern* within human populations that comprise society by shar-



ing the main aspects of a common existence over time as well as nonrandom *patterning*, the human and interhuman activities through which patterns are formed, retained, altered, or replaced. These twin aspects of social organization had been considered *structure*, relatively stable patterns of interrelations among persons or other social units, and *process*, the manner in which the patterns are produced, reproduced, or transformed (see, e.g., Faris 1964). The distinction is blurred to the extent that interrelations vary in *degrees* of regularity, uniformity, and permanence in the rhythms of coexistence, contact, or avoidance of which they consist (Williams 1976). In short, structure can also be viewed as patterned process among human agents (e.g., Blumer 1969, pp. 78–89; Giddens 1979, pp. 49–95; Coleman 1990, esp. pp. 1–44).

At issue is not what is patterned or how, but simply the extent to which there is any pattern or patterning at all. The antithesis of social organization is not opposition or discord. Conflict and other aspects of tension or unrest can, for example, exhibit regularity and uniformity as readily as can union, harmony, and tranquillity. Rather it is randomness, consisting of chaos, formlessness, and idiosyncratic human behavior (Blau 1975) and is called social disorganization. Yet patterned *occurrence* of disorganization is no contradiction.

Social organization is characterized by *interdependence*—that is, what occurs among certain components has, to varying degree, consequences for some or all of the other components and their relations with one another. These consequences can range from loss, even annihilation, to survival and other types of gain. Subsumed are regulation and stability as well as replacement and transformation.

The socially organized units or sets of units are generally activities or actors, individual or plural, that affect one another more immediately—even if simply by coexisting or by their sheer numbers—than do other activities or actors. The former are therefore distinguished (to varying extents) from an environment that might include those other units.

The units considered vary in their distinguishability, modifiability, and permanence. For some purposes they have been defined as concrete entities, such as persons or countries, or as activities by these entities, such as acts of persuasion or conquest. For other purposes the units have been

defined abstractly as only certain aspects of concrete entities or of their activities, called roles or functions (e.g., Hawley 1986, pp. 31–32), that signify position or participation in a particular aspect of collective living, or as complexes of these entities or activities. Examples are worker, labor, industry in production; official or bureau, directive, central administration in governance; judge or court, adjudication, court system in jurisprudence; supplicant or temple, prayer, denomination in religion; friend or friendship group, attachment, solidary web in emotional bonding; and lecturer or college, teaching, school system in social learning. It will be noted from these examples that units can be sets or combinations of other units. They are called substructures when they constitute broad components of social organization not detailed as to their composition. Examples of *societal* substructures are political state, economy, or moral community.

Structure subsumes both form and content: form generally in the senses of numbers, sizes, shapes, assemblage, connections among units, and directions of flow (say, of resources or persuasion); content in the sense of type of unit, substructure, relationship, or process. Clearly this distinction, though convenient for exposition, is not absolute. Units have been assigned to types on the basis of their forms. And there is form when the same everyday dramas can be “performed” in virtually any setting, whether work or nonwork (Goffman 1959).

## FORM

In conventional usage, form implies arrangement in space (not necessarily physical space) and in time. It also implies relationships among elements, in this case the units, against a background—the environment—which is conceived as external to structure (Hawley 1986, pp. 10–44; Smelser 1988). The environment may contain units with which concrete portions of the structure have relationships. Thus environment’s separation from structure is often abstract, especially where environment’s nonphysical attributes are concerned.

Form can be classified by the processes that occur among the units. Among pure types that are widely investigated are markets, characterized by competition and exchange; arenas, characterized by interunit struggle and alliance; collectivities,

characterized by cooperation in joint activity—that is, acting in concert (Parsons and Smelser [1956] 1965, p. 15), even if coerced; and aggregates, characterized by the absence of relations between the units.

These forms occur with a variety of contents. For example, markets not only process goods, services, and labor; they can also process social resources, such as information, intimacy, commitment, information, influence, and prestige. Arenas need not only be political; they can also exist within, say, the family. Collectivities may be found in any or all aspects of human living. And labor forces, electorates, viewerships, school enrollments, and populations of organizations (Aldrich and Marsden 1988) may all be conceived to constitute aggregates. Crosscutting the degrees to which these pure types are approximated are the variables of segmentation, stratification, specialization, scale, and endurance.

**Segmentation.** Segmentation is division into or banding together of like units (Parsons and Smelser [1956] 1965, p. 256; Parsons [1966] 1977, p. 25; Luhmann [1977] 1982, pp. 231–235, 242–245; Wallace 1988). Tribal societies, for example, can, in their inception, be considered markets constituted by exchange of spouses among relatively small family groups—the segments—that avoid internal marriages (Habermas [1981] 1987, pp. 161–162). In the case of collectivities, segmentation—called categoric organization (Hawley 1986, pp. 70–73)—enables each to accomplish more than could be accomplished separately, as, say, in the union of family groups in hunting and gathering societies (Duncan 1964; Lenski 1975; Hawley 1986, pp. 34–35), in a union of persons pursuing the same occupation, or in the replacement of one national corporation with several regional units. Categoric organization can also replace, at least in part, what would otherwise be mutually destructive competition or conflict or can otherwise create a more predictable environment for the participating unit. Examples are umbrella units that form among populations of special collectivities facing unpredictability in “turbulent environments” (cited by Scott 1987, pp. 122–123), even if they serve as no more than clearinghouses for information. All things being equal, ethnic pluralism—that is, subdivision into collectivities having different histories and life-styles—is another illustration of segmentation.

**Stratification.** For present purposes, stratification means the ranking of units or sets of units in their capacities to affect the existence and activity of other units or sets by controlling resources. This occurs through manipulation or struggle, as a result of competition or exchange, or voluntarily. Stratification and segmentation can co-occur. Where agriculture is the main economic activity, for example, it has been observed that rank tends to be associated primarily with size of landholding, if any; and where land ownership is also associated with other ranking criteria, as in earlier India, it is likely to serve as a basis either for strata (layers) or for collectivities of same-rank kinship units (Landecker 1981, pp. 33–34, 97). More complex social organization can consist of stratification among units that are themselves stratified. For example, it has been shown that the nations of the world are both internally stratified and stratified vis-à-vis one another (Wallerstein 1979), although effects of the one structure on the other are variable (see, e.g., Evans and Stephens 1988).

**Specialization.** In its pure sense, specialization refers to composition of unlike units that only taken together can accomplish all that is deemed significant. As a division of labor (Smith [1789] 1976; Durkheim [1902] 1964; Rueschemeyer 1986), specialization is characterized by greater interdependence—facilitative *or* inhibitory—than in the cases of simple segmentation or stratification. Segments can be added or lost with little effect; specialties cannot. Specialization ranges from little more than by age and sex in hunting and gathering societies, possibly with part-time political and religious leadership, to thousands of occupational specialties and nonoccupational roles in industrial and postindustrial societies (Lenski 1975, 1979; Hawley 1986, pp. 31–37, 64–67).

Generally also stratified, this form of social organization has been called corporate in distinction to categoric (Hawley 1986, pp. 68–69). The specialized units can be individuals or certain of their roles, as in a family, a small commune, or a small business enterprise; or they can themselves be segmented, stratified, or specialized internally.

**Scale.** Frequently associated with other aspects of social organization is its scale, variously specified as, for example, the number of units encompassed; the number of levels at which units are nested into progressively more comprehensive

units; or lengths of chains in the modification and flow of materials and services, information, influence, or command.

Contrasted with today's enormous urban settlements and nation-states are the unspecialized and unnested bands of twenty-five to forty or so hunters and gatherers (Lenski 1979). Specific examples of nesting—often accompanying or accompanied by stratification and one type of chain—are world system, country, province, and locality; Catholic Church, archdiocese, diocese, parish, and priest or parishioner; corporation, division, department, and job. Nesting is bypassed when, for example, transnational corporations become disassociated from local, provincial, or national jurisdictions; when religious ties crosscut nation-states; or when staff activities or occupational associations occur irrespective of level (Turk 1977, pp. 923–224; Tilly 1984, p. 136; Hawley 1986, p. 104).

The more activities become organized into large, specialized, but unnested units at higher levels of social organization, according to one body of theory, the less likely are the constituent units to be interconnected (references cited by Turk 1977, p. 65; McAdam et al. 1988). Each of the aggregates that result, sometimes the set of all such aggregates, is called a mass. An example is loss of relationships based on common residence to the extent that the local community is penetrated by specialized large-scale nonlocal collectivities called organizations (e.g., Turk 1977, pp. 65–66, 208–209).

**Endurance.** An important basis of classifying structure—or, for that matter, any complex unit or relationship that comprises structure—is by the extent to which it predates and outlives, or is otherwise independent of, specific units. Factory workers and managers are relatively replaceable, but a marriage does not substitute a new spouse for one that has been lost without becoming a different marriage. A market can be independent of particular producers and consumers, but a partnership or international bloc is not. Moreover, general features of factory, family, market, partnership, or bloc—or the means of generating these—tend to predate and survive as prototypes any of their specific instances (see, e.g., Jackson 1990).

**Combinations.** Hybrid forms, some of them complex, occur when segmentation, stratification,

specialization, and nesting are considered with respect to one another or with respect to markets, arenas, collectivities, and aggregates. Further, units can themselves be composed of other units in ways other than by nesting, or they can themselves be patterned. A few illustrations follow.

Organizations are specialized collectivities. Defined as complexes of more or less cooperative relations directed toward more or less specific objectives, these units have been said to occur in every known society as “the major vehicle through which concentrated goal-directed effort takes place” (Udy 1979). They, the aggregates they comprise, and relations among them—including organizations of organizations—are considered to be primary units of social organization, at least in industrial and postindustrial communities and nations (see, e.g., Turk 1977, 1985; Skocpol 1979; Perrow 1986; Evans and Stephens 1988; Perrucci and Potter 1980; Coleman 1990).

Markets and arenas generally affect and are affected by collectivities—including ones that they nest or in which they are nested, that they constrain or by which they are constrained—shaping units so they can compete, exchange, struggle, or ally themselves with one another (see, e.g., Stinchcombe 1986; Coleman 1990, esp. pp. 266–321, 371–396, 689). Indeed, organizations and other kinds of collectivities can affect the conditions under which other organizations are formed in substitution for markets (Williamson 1990). Clearly, positions within the organizational substitutes may be filled in turn by labor markets (Stinchcombe 1986; Granovetter and Tilly 1988).

Sizes of aggregates can affect social organization (see citations by Eisenstadt 1968). Sheer numbers make it impossible for each unit to have relations with each other one. This affects the probabilities of positions, networks, or organizations that channel and mediate social relationships (research stimulated by Simmel 1908, pp. 55–56). More recent work shows how specific interconnections, such as marriage or crowd behavior under conditions of threat, can be affected by aggregates, by their relative sizes and other properties, and by relations among these properties (e.g., Blau 1987, Coleman 1990).

*Network analysis* has added precision to the measurement of form—say, of stratification or of degrees of interconnectedness—by detailing the

connections (links) among units and the patterns that these provide (see, e.g., Cook 1977; Leinhardt 1977; Burt 1982; Turner 1986, pp. 287–305). Its techniques are uniquely suited to the chains, clusters, and sequences of exchange, cooperation, alliance, or command over which goods, services, money, information, or influence flow. Associated with scale, for example, can be the number of points between origin of flow and its completion. Network formulations have also proven especially useful in identifying relations among organizations that affect concerted action locally (Turk 1977; Galaskiewicz 1989) and ones that affect it nationally (Laumann and Knoke 1989).

**Complexity.** The very complexity of social organization—the number and variety of units, levels, and interconnections—is itself an aspect of form. Though admittedly crude (Luhmann [1977] 1982, pp. 232–233; Tilly 1984, pp. 48–50; Rueschemeyer 1986, p. 168), this variable can be used to account for other aspects of social organization (e.g., Lenski 1075, 1979; Turk 1977; Habermas [1981] 1987, pp. 153–197; Luhmann [1977] 1982, pp. 229–254; Rueschemeyer 1986).

## CONTENT

Classification of units—including complex units—is necessary for similarity, stratification, or specialization to signify more than simply differences in form. Among many, two bases of classification stand out. Sometimes applied jointly, the one emphasizes objective consequences—positive, negative, or neutral, and varying from 0 in degree—the other communicated, remembered, or recorded meanings and rules. Examples of the two bases follow.

**Consequences.** Substructure has been classified according to its consequences for stability and change in overall structure (e.g., Marx and Engels [1846] 1970; Marx [1859] 1971; Parsons [1966, 1971] 1977; Luhmann [1977] 1982; Habermas [1981] 1987; Hawley 1986). Among these, adaptation to the environment and of units to one another have been stressed (see, e.g., Duncan 1964; Lenski 1975, 1979; Parsons [1966, 1971] 1977; Habermas [1981] 1987; Luhmann [1974–1977] 1982; Hawley 1986), quite likely because socially organized life has been observed to be the major

adaptive means available to primates (e.g., by Lenski 1975).

Thus, the distinction is often made between (1) economic substructure affecting environmental adaptation for the generation and distribution of general resources (e.g., gross national product, homelessness) and (2) political substructure affecting the generation and distribution of general capacity to mobilize resources for concerted action, including action by opposing collectivities (e.g., national efforts, party campaigns, uprisings). Further distinction involves (3) substructure affecting the generation and distribution of general bonds or schisms that provide harmony or discord among units (e.g., community cohesion, solitary antagonism; see Parsons and Smelser [1956] 1965, pp. 48–49; Parsons [1966] 1977, pp. 135–140; Hawley 1986, p. 66; Coleman 1990, pp. 91–116, 175–196, 517–527) and (4) substructure affecting the generation, distribution, and maintenance of participation in structure: recruitment of units (including but not limited to procreation) and their training, allocation, motivation, and retention—in short, the populating and regulating of social structure.

**Meanings and Rules.** The classification of content can also rest on disputed or common meanings, understandings, purposes, or binding rules (including law) that are communicated about environment and structure. Communication can involve all kinds of participants or only certain ones and can be modified depending on the context (Goffman 1959). The products of communication vary in their permanence through repetition, recording, or recall; in their breadth of dissemination and acceptance; and in their association with sanction—that is, support by enforcement or other incentives to comply. They can, but often do not, coincide with the structure's objective consequences (Habermas [1981] 1987, esp. pp. 153–197) but can affect it.

Most theories of social organization allow for the effects on stability and change of sanctioned agreement, or of oppositions among sanctioned agreements, calling the product institutional or cultural. They differ in terms of the importance the institutional component is said to have for structure.

Near the one extreme, institutional rules are viewed as “higher level” determinants of social

organization (Parsons [1971] 1977, pp. 234–236), not only of relations among units but also of the units themselves (Meyer et al. 1987; Coleman 1990, pp. 43–44, 325–70). Here the rules governing specific structure are considered to be products of more inclusive structure, such as political state or church (Znaniński 1945; Turk 1977, pp. 210, 215–221), or to be elements of a world “culture pool” (Moore 1988; Meyer et al. 1987). Near the other extreme, meanings (“ideas”) are viewed primarily as by-products of the material relations of production (Marx and Engels [1846] 1970, pp. 57–60) or in terms of their significance for organized adaptive processes vis-à-vis the environment (e.g., Duncan 1964; Lenski 1975). Meaning, if the concept is employed at all, is restricted to the acting unit’s purposive rationality: its adoption, within the limits of error and imperfect knowledge or skill, of means that are appropriate to specified outcomes (see, e.g., Hawley 1986, pp. 6–7).

A second source of variability is the degree to which the institutional can be seen as analytically distinct from structure. Some consider the institutional to be an aspect of the environments of substructures and other units (Parsons 1968; Meyer and Rowan 1978; Hawley 1986, p. 79; Meyer et al. 1987; Coleman 1990, pp. 43–44). Others see it as relatively inextricable from structure (e.g., Giddens 1979, pp. 49–85), specifically where structure is viewed as formed and modified through interpretive interaction between persons or groups, in which interpretations are frequently but not always shared (Blumer 1969, pp. 86–88).

The third issue regards the degree to which structure is “spelled out” by the institutional. There can be precise rules that govern even the minutiae, as on an assembly line or in religious ritual. There are also general principles that reduce the number of structural alternatives without determining structure precisely (Parsons [1971] 1977, pp. 193–194), as in such ideals as freedom, rationality, retribution, obedience, protest, solidarity, revolution, contract, and property. Relatedly, there can be shared common-sense reasoning (Collins 1988, pp. 273–291 on Garfinkel) on the basis of which “sense” is made of structure. Or there can be broad myths that serve to provide accounts of social organization (Meyer and Rowan 1978, Goffman 1959) on the basis of which, rather than on the basis of performance, structure is justified.

*Agency*, as do most theories incorporating meaning, rests on the general idea of interest—variously called purpose, intention, motive, or goal and variably emerging during a course of action—and on the availability of action alternatives (e.g., Giddens 1979, pp. 55–56). Agency is the extent to which purposive action by and interaction among the units affect social organization. There is little need to consider what individual units contribute to organization or why (Hawley 1986, pp. 6–7). When structure, environment, or rule are viewed as an absolute constraint or as providing only limited choice (see, e.g., Blau 1987), or where social organization results from natural selection (Lenski 1975, 1979).

This is not the case where structure and culture simply set loose conditions for action and interaction. Important here for structural and cultural stability and change are the processes (1) by which interests are pursued under the influence and constraint of other actors (Homans 1975); (2) by which the conditions of action are interpreted through social interaction (Blumer 1969); (3) by which the acting unit monitors and adjusts the components of action, even intention, throughout (Giddens 1979, pp. 53–59); and (4) by which individual and collective actors choose means of implementing their interests (Coleman 1990). These processes are significant both to the reproduction and to the transformation of structure and institutional rule.

*Classifications of meaning* are numerous. Yet they frequently rest on one or more of the following variables (polar approximations in parentheses): (1) specificity: the scope of the relationship, from specific to diffuse content (e.g., organizations or special markets vs. unspecialized collectivities or conflict arenas); (2) universalism: the extent to which relationships hold for all units belonging to a category or only for particular ones from that category (e.g., upholding sovereignty of any nation or opposition toward all governments vs. a treaty or a declaration of war); (3) neutrality: the extent to which relationships are means to ends vs. ends in themselves (e.g., banking transactions or job competition vs. flag-raising or flag-burning ceremonies); and (4) performance: the extent to which relationships are based on what units do rather than on what they are (e.g., production relations or industrial conflict vs. aristocracy or racial conflict).

Concrete as well as abstract structures have been classified according to various combinations of values of these variables, leading to widely used typologies. More generally, any variables used to define content and form serve the analysis of overall structure by describing its various aspects or its different parts.

### FORM AND CONTENT

Structure can vary in the extent to which substructures—such as the four whose consequences were noted—are abstractions, involving the same concrete units rather than different ones. At the low end of this continuum, the units comprise a single unstratified and unspecialized collectivity, approximated by a tribal society or a commune, in which each act affirms the totality (Durkheim [1902] 1964; Luhmann [1974] 1982, 1987, pp. 153–197), having every kind of consequence for it. In it myth tends to blur distinctions among the objective, social, and subjective worlds and between society and its natural surroundings (Habermas [1981] 1987, pp. 158–159), and there is comprehensive and detailed regulation of activity. The units not only resemble one another but also tend not to have separate identities. In short, relations tend to be particular to the given collectivity, not universal; affective, not neutral; ascriptive, not performance-based; and diffuse, not specific.

In examples of this noncomplex instance, economic and political organization tend to be extensions of family, extended family, religious group, and common territory, which overlap to considerable degrees. Commitment to any one aspect of social life tends to be supported and sanctioned within all of the others (Habermas [1981] 1987, pp. 156–157), and there are relatively few conflicting constraints, structural or cultural (Blumer 1969, pp. 87–88).

Compliance to one cluster of rules and understandings is approximated, with utility, sanction, attachment, and/or commitment as its basis or bases. Here change has been attributed, in the main, to changes in the environment or in ways of coping with it (e.g., Hawley 1986, pp. 15–18) or to ubiquitous “tension” between rules of action and the situation of the acting unit (e.g., Parsons and Smelser [1956] 1965, esp. pp. 50–51). There is little institutional provision for change.

With greater complexity, however, activities can be removed from these primordial units and assumed by large-scale economic, political, and other kinds of organizations (e.g., Lenski 1979; Coleman 1990, pp. 584–585). Varied organizational purpose as well as interaction among organizations can constitute bases of change that are themselves institutionalized.

**Domination.** Stratification implies domination—that is, setting the conditions of existence by certain units for other units through disposition over key resources; over the generation and selection of often self-serving meanings and rules (see, e.g., Landecker 1981); and over the means of securing compliance. Such disproportion is one of the most widely considered sources of strain, hence of change through conflict (see, e.g., the modifications of Marxian theory by Dahrendorf 1959 and Skocpol 1979).

Considered under various names (see, e.g., Marx and Engels [1846] 1970; Parsons and Smelser [1956] 1965; Dahrendorf 1959; Duncan 1964; Hawley 1986; Giddens 1985), domination includes disposition over means of coercion, material inducement, social support, or rules of command (adapted from Weber [1920] 1978, pp. 53–54; Hawley 1986, pp. 33–37). Domination reflects power, the capacity to affect action and its outcomes (Giddens 1979, pp. 88–94). Suggested by Weber ([1920] 1978, p. 53) as the probability that the acting unit can carry out its will despite resistance, power has also been defined as the structure’s capacity to mobilize resources in effecting outcomes through concerted action, such as the production of sustenance (or other ways of supporting units) and environmental control (Parsons and Smelser [1956] 1965, p. 48; Hawley 1986, pp. 36–37). This capacity is conceived to be distributed in different ways, depending on the structure under consideration (Hawley 1986, pp. 74–77), serving political relations as money serves economic relations (Parsons 1975).

Domination can effect both form and content. For example, domination can cause segmentation to give way to stratification, frequently as a result of conquest, as when multicomunal societies become kingdoms or empires (Lenski 1979) or numerous petty sovereignties are gathered under nation-states (Tilly 1984, p. 48). By setting the conditions for competition and exchange or more

directly by affecting concerted action that favors certain specialties or even by imposition, domination can also make specialization possible or influence its nature and degree (Habermas [1981] 1987, pp. 161–163; Hawley 1986, pp. 64–67, 91–95; Rueschemeyer 1986).

The resources on which domination is based need not only be material. They can be social—for example, active support, institutionalization by organizations and movements that generate and implement ideology or legality, and the absence of opposition. Generalized control over all manner of resources has been called hegemony, and political processes have been characterized as struggles for hegemony (Wallerstein 1979, 1984; Tuchman 1988). Domination has been conceived, in its extreme form, as rendering certain alternatives invisible (Giddens 1985, pp. 8–10) through taken-for-granted opposition by powerful units (Polsby 1980, pp. 189–218), lack of relevant language (Parsons 1975), absence of relevant substructures, agendas set by the mass media of communication (Tuchman 1988), and uncontested legitimacy—that is, common understandings as to what is valid or binding. The elementary stratified society is, by definition, hegemonic, since it is unspecialized in terms of the control of various types of resources.

Domination can be by one or more substructures over the others. History shows, for example, cases of kinship-based, religious, military, economic, and political domination (e.g., Tilly 1984; Evans and Stephens 1988). The hypothesis has been suggested that the greater the specialization, the less stable is domination through fusion of, say, political with economic activities (Parsons and Smelser [1956] 1965, p. 83). The expected trend is for substructures to become concretely separate. For example, underground markets arise as responses to shortages and bottlenecks in socialist societies, which are characterized by political domination of economic activity, or by the “informally” organized demands of their workers (Jones 1984).

**Pluralism and Plasticity.** The degree to which meanings are single or plural can be affected by segmentation or specialization. Different collectivities, constituted, say, on the basis of different fundamental beliefs, different descent, or different economic circumstances (Landecker 1981), coexist segmentally either by loose agreement or by coercive regulation, each with separate rules and un-

derstandings (Tenbruck 1989), or struggle with one another over which ones shall prevail (e.g., Landecker 1981, pp. 136–169; Wallace 1988).

The greater the specialization among units, all things being equal, the greater also is the plurality of interests, according to some models. A recurrent theme is that this form of pluralism affects the probabilities of different degrees of involvement by given units and of different alignments among units from one issue to the next. Such differential participation can have a negative effect on the probability of broad or intense conflict (Dahrendorf 1959, pp. 215–231; Polsby 1980, pp. 84–97, 122–138; Turk 1977, pp. 97–103, 1985; McAdam et al. 1988).

Specialized forms, including organizations, are by definition relatively indifferent to the activities or aspirations of supporting, component, or utilizing units in other social settings (Luhmann [1975] 1982, pp. 78–79; Labovitz and Hagedorn 1977, pp. 12–15). Entire areas of indifference are seen to result, for example, from “gaps” left between interests served by organizations (Luhmann [1975] 1982, pp. 79–80, 87, [1977] 1982, p. 237). The intrusion of, say, race and gender in contemporary labor markets suggests, however, that organizational indifference is a matter of degree (see, e.g., Stinchcombe 1986; Granovetter and Tilly 1988).

Accompanying segmentation or specialization, according to several theories, is plasticity: the probability of loose and variable connections among segments or among organizations subsumed by substructures like the four singled out. Here overall organization is limited to compatibility, falling short of the pursuit of unified or concerted outcomes across substructures (see, e.g., Luhmann [1975] 1982, pp. 78–79). Change is endemic in the absence of overall structure, save for markets and conflict arenas and rules that govern these (Luhmann [1977] 1982, pp. 238–242). It tends to occur as accommodation (1) through exchange, say, between segments or between political and economic substructures (Parsons and Smelser [1956] 1965), or (2) through new forms and modification of older ones in responses to changes elsewhere, as in the case of the family’s loss of economic activity in the United States but its growing importance in providing incentive for such activity (Schumpeter, cited by Suttles and Janowitz 1979).

The greater the segmentation or specialization, the more general the accommodative meanings and rules that encompass social organization overall—for example, the idea of tolerance and its enforcement, or of universal civility (Parsons [1971] 1977, pp. 182–193; Hawley 1986, p. 66). Another example is the idea of freedom in classic liberal society, implemented as economic *laissez faire*, religious autonomy, voluntary rather than arranged marriage, political competition for electoral support, and stratum membership on the basis of achievement rather than family (Gerth and Mills 1964, pp. 354–357).

There is disagreement about the extent of hegemony under conditions of specialization, even where, say, political or economic organizations overshadow other organizations in control over resources. At the one extreme incumbents of dominant positions within the various specialized organizations have been conceived as constituting a single elite capable of joint domination (e.g., Mills 1956) or as being generally dominant to the extent that they hold positions in multiple organizations (Perrucci and Pilisuk 1970). Relatedly, large organizations have been observed to divorce themselves from the interests they were established to represent—as, say, those of capital and labor—and strike bargains with one another (Evans and Stephens 1988). These conceptions of unified domination have partly been verified and partly refuted (e.g., by Lieberson 1971; Mizruchi 1982; Johnsen and Mintz 1989).

Specialized as they are, according to another partly verified view, the same set of organizations can facilitate one another in certain respects and be in mutual struggle in others, therefore resisting direct domination by any one or a few of their number (citations in Turk 1985). Under these conditions policy that is the basis for binding domination is formed through action by those masses or by those nonpermanent coalitions of public agencies and private organizations that are concerned with any particular matter (Turk 1977, pp. 136–205; Galaskiewicz 1989; Laumann and Knoke 1989).

Crosscutting the issue is the question of whether, given specialization, either an organized elite or an interorganizational coalition has the capacity for concerted action that transcends the substructures. Specialization can mean that, at the

most, even the most powerful organization is dominant only with respect to one or two issue areas (Luhmann [1975] 1982, pp. 76–89; Polsby 1980, pp. 122–128). However, even within organizations the normal decision process has been defined as organized anarchy by certain investigators (e.g., citations by Scott 1987, pp. 277–282; Meyer and Rowan 1978). Carried to an extreme, the question is one of the extent to which coordination by domination is haphazard, whether actual coordination might not result from the “invisible hand” of a market (Smith [1789] 1976, Vol. 1, pp. 477–478) or of an arena of conflict.

**Standardization.** Mitigating the diversifying possibilities of segmentation and specialization that have been noted are tendencies for units to become alike in certain aspects of form and content—that is, isomorphic (Hawley 1986, pp. 66–70; Udy 1979; Kerr 1983, pp. 85–89). This can be because units require similar internal arrangements for purposes of connection with one another (Hawley 1986, p. 70), or model themselves after other units in the preservation of competitive effectiveness (suggested by Aldrich and Marsden 1988, citing DiMaggio and Powell). It can also be through institutionalization by drawing upon a common “culture pool” available, say, to the countries of the world (Moore 1988) or through, say, political or religious imposition within a given society (Znaniecki 1945; Landecker 1981, p. 136) or by either (Hawley 1986, p. 66; Meyer et al. 1987).

Not only can there be standardization of form, there can also be standardization of process. For reasons of predictability and economy of effort, among others, the joint reproduction of habitual patterns has been considered to lie at the heart of social life (e.g., Berger and Luckmann 1966). This not only accounts in part for compliance with meanings and rules that standardize feeling, thinking, and acting in such relatively unspecialized settings as tribal societies, it also accounts for standardized, partly area-specific media and codes that symbolize and routinize all manner of activities and products where there is more specialization. Examples are language (Habermas [1981] 1987, pp. 56–57), money, property, prestige, influence, power, legality, administrative principles, criteria of truth (Parsons 1975, Habermas [1981] 1987, pp. 153–197, 367–373; Luhmann [1974] 1982, pp. 168–170), and credentialed expertise (Collins 1988, pp. 174–184; Luhmann [1976] 1982,



pp. 303–331; Bauman 1989). These standardize and objectify aspects of markets, conflict arenas, collectivities, or aggregates. Unless it fails noticeably, it is believed, routine tends not to be questioned, especially where relevant knowledge is not widely pursued and alternatives are not at hand.

**Collective Agency.** Mainly based on exchange, conquest, or revolution, change is seldom institutionalized in segmental or stratified society. Pluralism generally includes provisions for categorical organization in the form of mass action, say through referenda, or in the form of movement and interest organizations, which not only participate in the coalitions that seek to dominate given matters but also define matters for action by the mass. These institutionalized means of structural and cultural change through collective agency are thought to occur where pluralism means lack of overall institutional detail (e.g., Gusfield 1979). Recent investigations have not only examined their causes and effects but also their forms and ways of acting (Zurcher and Snow 1981; McAdam et al. 1988). Their efforts can be toward increasing the material and social resources controlled by given categories of units, as in the case of gay rights, or they can be directed toward broad structural changes or changes in meaning that occur for their participants and for nonparticipants alike, as in the case of civil liberties (Gusfield 1979). Like revolutions, movements have been considered to be processes that can begin by effecting transitional social organization and end with new institutionalized structure (Alberoni [1981] 1984).

**Commitment and Trust.** Commitment of units to one another and to their common structure is a widely recognized influence. The “we” that characterizes collectivities (Cooley 1902, 1916) or the “consciousness” that causes support of existing social organization or generates struggle within it (Marx and Engels [1846] 1970; Durkheim [1902] 1964) is seen as having either of two effects. It can be direct, producing commitment to the given structure, or indirect by habitualizing the commitment and trust that serves participation in a variety of settings, even ones that are specific and neutral (also see Coleman 1990, p. 297).

Increasing specialization and other forms of organizational complexity and increasing scale have been viewed as negative influences on commitment, even on commitment to disputing factions

or to revolutionary movements seeking structural change. Commitment to large-scale organizations comprising specialized substructure is generally considered less than to other kinds of collectivities because of anonymity among the constituent units, which are likely to have the characteristics of a mass, and because only part of the constituent is involved.

With complexity and scale, participation in one substructure is less contingent on participation in others, incurring separate, possibly conflicting obligations; indeed, everyday interactions that produce common meaning have been considered divorced from structure and rules (Luhmann [1974–1977] 1982, 1987; Habermas [1981] 1987, pp. 117, 153–197), and even everyday life is penetrated by the actions and generalized media controlled by organizations (e.g., legal and monetary), reducing its potential for social integration (Habermas [1981] 1987, pp. 267, 330–331, 367–373). Under these conditions the influence of trust on social organization is less likely (Coleman 1990, pp. 300–321), and benefits are more likely sought without corresponding contributions (Coleman 1990, pp. 650–655).

The greater the specialization and accompanying standardization, it has been observed, the more of the population can be included in whatever substructures comprise various areas of social life (Parsons [1966] 1977)—that is, the more universalistic, for example, are criteria for suffrage and military service, access to public facilities, mass education, and employment opportunity. One reason follows. The more social structure consists of aggregates of organizations whose concerted activities are narrowly focused on specialized consequences—meaning specific and neutral orientation toward performance—that tend to be unranked, the less relevant these aggregates are to one another as criteria of exclusion (Luhmann [1977] 1982, pp. 236–238). Examples can be found in references to “customer,” “patient,” “student,” or “defendant,” independently of the beneficiary’s other social attributes. At its extreme it can even lead to “the gall bladder in bed 27.” Aided by standardized media such as money, the result of inclusion can be the diminution of commitment and trust through the transformation of agents into commodities for exchange (e.g., Marx [1859] 1971, pp. 78–84), or their more general removal

from meaningful communicative interaction (Habermas [1981] 1987, e.g., p. 343).

The other side of inclusion is regulation. The more central that large, specialized organizations are to social structure, it has also been claimed, the less social organization depends on commitment and the greater is the shift from mutual trust to trust in expertise. The state, for example, is said to require less legitimacy as its expert-driven political technology provides, say, surveillance, "correction," welfare supervision, "medicalization," or "psychiatrization" (Bauman 1989). Indeed, the electorate's growing cynicism about government (e.g., Institute for Social Research 1979) has had little apparent effect on political structure in the United States.

(SEE ALSO: *Organizational Structure*; *Social Dynamics*; *Social Network Theory*; *Social Structure*)

## REFERENCES

- Ahrne, Goran 1994 *Social Organizations: Interaction Inside, Outside and Between Organizations*. Beverly Hills, Calif.: Sage.
- Alberoni, Francesco (1981) 1984 *Movement and Institution*, trans. P.C.A. Delmoro. New York: Columbia University Press.
- Aldrich, Howard E., and Peter V. Marsden 1988 "Environments and Organizations." In N.J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Alvarado, Felix 1996 "Concerning Postmodernity and Organizations in the Third World: Opening a Debate and Suggestions for a Research Agenda." *Organization Science* 7:667-681.
- Bauman, Zygmunt 1989 "Legislators and Interpreters: Culture as Ideology of Intellectuals." In H. Haferkamp, ed., *Social Structure and Culture*. New York: Walter de Gruyter.
- Berger, Peter L., and Thomas Luckmann 1966 *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, N.Y.: Doubleday.
- Blau, Judith R. 1996 "Organizations as Overlapping Jurisdictions: Restoring Reason in Organizational Accounts." *Administrative Science Quarterly* 41:172-179.
- Blau, Peter M. 1975 "Introduction: Parallels and Contrasts in Structural Inquiries." In P. M. Blau, ed., *Approaches to the Study of Social Structure*. New York: Free Press.
- 1987 "Contrasting Theoretical Perspectives." In J. C. Alexander, B. Giesen, R. Mönch, and N. J. Smelser, eds., *The Micro-Macro Link*. Berkeley: University of California Press.
- Blumer, Herbert 1969 *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, N.J.: Prentice-Hall.
- Brass, Daniel J. 1992 "Power in Organizations: A Social Network Perspective." *Research in Politics and Society* 4:295-323.
- Burt, Ronald S. 1982 *Toward a Structural Theory of Action: Network Models of Social Structure, Perception, and Action*. New York: Academic Press.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Collins, Randall 1988 *Theoretical Sociology*. New York: Harcourt Brace Jovanovich.
- Cook, Karen S. 1977 "Exchange and Power in Networks of Interorganizational Relations." *Sociological Quarterly* 18:62-82.
- Cooley, Charles Horton 1902 *Human Nature and the Social Order*. New York: Scribners.
- 1916 *Social Organization: A Study of the Larger Mind*. New York: Scribners.
- Dahrendorf, Ralf 1959 *Class and Class Conflict in Industrial Society*. Stanford, Calif.: Stanford University Press.
- Duncan, Otis Dudley 1964 "Social Organization and the Ecosystem." In R. E. L. Faris, ed., *Handbook of Sociology*. Chicago: Rand McNally.
- Durkheim, Emile (1902) 1964 *The Division of Labor in Society*, trans. George Simpson. New York: Free Press.
- Edelman, Lauren B. and Mark C. Suchman 1997 "The Legal Environments of Organizations." *Annual Review of Sociology* 23:479-515.
- Eisenstadt, Shmuel N. 1968 "Social Institutions, I: The Concept." In D. L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan.
- Evans, Peter B., and John D. Stephens 1988 "Development and the World Economy." In N.J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Faris, Robert E. L. 1964 "Social Organization (Sociology)." In J. Gold and W. L. Kolb, eds., *A Dictionary of the Social Sciences*. New York: Free Press.
- Galaskiewicz, Joseph 1989 "Interorganizational Networks Mobilizing Action at the Metropolitan Level." In R. Perrucci and H. R. Potter, eds., *Networks of Power: Organizational Actors at the National, Corporate, and Community Levels*. New York: Aldine de Gruyter.
- Gerth, Hans, and C. Wright Mills (1953) 1964 *Character and Social Structure: The Psychology of Social Institutions*. New York: Harcourt, Brace and World.
- Gherardi, Silvia 1995 *Gender, Symbolism and Organizational Cultures*. Beverly Hills, Calif.: Sage.

- Giddens, Anthony 1979 *Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis*. Berkeley: University of California Press.
- 1985 *The Nation-State and Violence*. Berkeley: University of California Press.
- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, N.Y.: Doubleday.
- Granovetter, Mark, and Charles Tilly 1988 "Inequality and Labor Processes." In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Gusfield, Joseph 1979 "The Modernity of Social Movements: Public Roles and Private Parts." In A. H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- Habermas, Jurgen (1981) 1987 *The Theory of Communicative Action*. Vol. 2: *Lifeworld and System: A Critique of Functionalist Reason*, trans. Thomas McCarthy. Boston: Beacon Press.
- Hawley, Amos H. 1986 *Human Ecology: A Theoretical Essay*. Chicago: University of Chicago Press.
- Homans, George C. 1975 "What Do We Mean by Social 'Structure'?" In P. M. Blau, ed., *Approaches to the Study of Social Structure*. New York: Free Press.
- Institute for Social Research 1979 "Deepening Distrust of Political Leaders Is Jarring Public's Faith in Institutions." *ISR Newsletter* 7:4-5.
- Jackson, John E. 1990 "Institutions in American Society: An Overview." In J. E. Jackson, ed., *Institutions in American Society: Essays in Market, Political, and Social Organizations*. Ann Arbor: University of Michigan Press.
- Johnsen, Eugene, and Beth Mintz 1989 "Organizational versus Class Components of Director Networks." In R. Perrucci and H. R. Potter, eds., *Networks of Power: Organizational Actors at the National, Corporate, and Community Levels*. New York: Aldine de Gruyter.
- Jones, A. Anthony 1984 "Models of Socialist Development." In Lenski, Gerhard, ed., *Current Issues and Research in Macrosociology*. Leiden, Neth.: E. J. Brill.
- Kerr, Clark 1983 *The Future of Industrial Societies: Convergence or Continuing Diversity?* Cambridge, Mass.: Harvard University Press.
- Labovitz, Sanford, and Robert Hagedorn 1977 "Social Norms." In Sanford Labovitz, *An Introduction to Sociological Concepts*. New York: Wiley.
- Landecker, Werner S. 1981 *Class Crystallization*. New Brunswick, N.J.: Rutgers University Press.
- Laumann, Edward O., and David Knoke 1989 "Policy Networks of the Organizational State: Collective Action in the National Energy and Health Domains." In R. Perrucci and H. R. Potter, eds., *Networks of Power: Organizational Actors at the National, Corporate, and Community Levels*. New York: Aldine de Gruyter.
- Leinhardt, Samuel (ed.) 1977 *Social Networks: A Developing Paradigm*. New York: Academic Press.
- Lenski, Gerhard E. 1975 "Social Structure in Evolutionary Perspective." In P. M. Blau, ed., *Approaches to the Study of Social Structure*. New York: Free Press.
- 1979 "Directions and Continuities in Societal Growth." In A. H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- Levy, Marion J., Jr. 1996 *Modernization and the Structure of Societies*. Vol. 2: *The Organizational Contexts of Societies*. New Brunswick, N.J.: Transaction.
- Lieberson, Stanley 1971 "An Empirical Study of Military-Industrial Linkages." *American Sociological Review* 76:562-585.
- Luhmann, Niklas (1974-1977) 1982 *The Differentiation of Society*, parts trans. S. Holmes and C. Larmore. New York: Columbia University Press.
- 1987 "The Evolutionary Differentiation Between Society and Interaction." In J. C. Alexander, B. Giesen, R. Münch, and N. J. Smelser, eds., *The Micro-Macro Link*. Berkeley: University of California Press.
- Marx, Karl (1859) 1971 *The Grundrisse*, D. McLellan, ed. and trans. New York: Harper and Row.
- Marx, Karl, and Frederick [sic] Engels (1846) 1970 *The German Ideology*, ed. C. J. Arthur, trans. W. Lough, C. Dutt, and C. P. Magill. New York: International Publishers.
- McAdam, Doug, John D. McCarthy, and Mayer N. Zald 1988 "Social Movements." In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- McKinlay, Alan and Ken Starkey, eds., 1998 *Foucault, Management and Organization Theory*. Beverly Hills, Calif.: Sage.
- Meyer, John W., John Boli, and George M. Thomas 1987 "Ontology and Rationalization in the Western Cultural Account." In G. M. Thomas, J. W. Meyer, F. O. Ramirez, and J. Bali, eds., *Institutional Structure: Constituting State, Society, and the Individual*. Beverly Hills, Calif.: Sage.
- Meyer, John W., and Brian Rowan 1978 "The Structure of Educational Organizations." In M. Meyer and associates. *Environments and Organizations*. San Francisco: Jossey-Bass.
- Mills, C. Wright 1956 *The Power Elite*. New York: Oxford University Press.
- Mizruchi, Mark S. 1982 *The American Corporate Network, 1904-1974*. Beverly Hills, Calif.: Sage.
- Monge, Peter R. 1995 "Theoretical and Analytical Issues in Studying Organizational Processes." In G. P.

- Huber, and A. H. Van de Ven, eds. *Longitudinal Field Research Methods: Studying Processes of Organizational Change*. Beverly Hills, Calif.: Sage.
- Moore, Wilbert E. 1988 "Social Change." In E. F. Borgatta and K. S. Cook, eds., *The Future of Sociology*. Beverly Hills, Calif.: Sage.
- Orlikowski, Wanda J. 1992 "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3:398-427.
- Parsons, Talcott 1968 "Social Systems." In D. L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan.
- 1975 "The Present Status of 'Structural-Functional' Theory in Sociology." In L. A. Coser, ed., *The Idea of Social Structure: Papers in Honor of Robert K. Merton*. New York: Harcourt Brace Jovanovich.
- (1966, 1971) 1977 *The Evolution of Societies*, ed. Jackson Toby. Englewood Cliffs, N.J.: Prentice-Hall.
- Parsons, Talcott, and Neil J. Smelser (1956) 1965 *Economy and Society: A Study in the Integration of Economic and Social Theory*. New York: Free Press.
- Perrow, Charles 1991 "A Society of Organizations." *Theory and Society* 20:725-762.
- Perrow, Charles 1992 "Organisational Theorists in a Society of Organisations." *International Sociology* 7:371-380.
- 1986 *Complex Organizations: A Critical Essay*, 3rd ed. New York: Random House.
- Perrucci, Robert, and Marc Pilisuk 1970 "Leaders and Ruling Elites: The Interorganizational Bases of Community Power." *American Sociological Review* 35:1040-1057.
- Perrucci, Robert, and Harry R. Potter 1989 "The Collective Actor in Organizational Analysis." In R. Perrucci and H. R. Potter, eds., *Networks of Power: Organizational Actors at the National, Corporate, and Community Levels*. New York: Aldine de Gruyter.
- Polsby, Nelson W. 1980 *Community Power and Political Theory: A Further Look at Problems of Evidence and Inference*, 2nd ed. New Haven, Conn.: Yale University Press.
- Rueschemeyer, Dietrich 1986 *Power and the Division of Labor*. Stanford, Calif.: Stanford University Press.
- Scott, W. Richard and Soren Christensen 1995 *The Institutional Construction of Organizations: International and Longitudinal Studies*. Beverly Hills, Calif.: Sage.
- Scott, W. Richard 1987 *Organizations: Rational, Natural, and Open Systems*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Simmel, Georg 1908 *Soziologie: Untersuchungen über die Formen der Vergesellschaftung*. Leipzig: Duncker und Humblot.
- Skocpol, Theda 1979 *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. New York: Cambridge University Press.
- Slappendel, Carol 1996 "Perspectives on Innovation in Organizations." *Organization Studies* 17:107-129.
- Smelser, Neil J. 1988 "Social Structure." In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Smith, Adam [1789] 1976 *The Wealth of Nations*. Chicago: University of Chicago Press.
- Stinchcombe, Arthur L. 1990 *Information and Organizations*. Berkeley: University of California Press.
- 1986 "Economic Sociology: Rationality and Subjectivity." In U. Himmelstrand, ed., *Sociology: From Crisis to Science?*, vol. 1, *The Sociology of Structure and Action*. Beverly Hills, Calif.: Sage.
- Strang, David and Sarah A. Soule 1998 "Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills." *Annual Review of Sociology* 24:265-290.
- Suttles, Gerald, and Morris Janowitz 1979 "Metropolitan Growth and Democratic Participation." In A. H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- Tenbruck, Friedrich H. 1989 "The Cultural Foundations of Society." In H. Haferkamp, ed., *Social Structure and Culture*. New York: Walter de Gruyter.
- Tilly, Charles 1984 *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- 1998 *Durable Inequality*. Berkeley: University of California Press.
- Tuchman, Gaye 1988 "Mass Media Institutions." In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Turk, Herman 1977 *Organizations in Modern Life: Cities and Other Large Networks*. San Francisco: Jossey-Bass.
- 1985 "Macrosociology and Interorganizational Relations: Theory, Strategies, and Bibliography." *Sociology and Social Research* 69:487-500. (Reprint with corrected typography is available.)
- Turner, Jonathan H. 1986 *The Structure of Sociological Theory*, 4th ed. Chicago: Dorsey Press.
- 1990 "Emile Durkheim's Theory of Social Organization." *Social Forces* 68:1089-1103.
- Udy, Stanley, H., Jr. 1968 "Social Structure: Social Structural Analysis." In D. L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan.

- 1979 “Societal Growth and Organizational Complexity.” In A. H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- Wallace, Walter L. 1988 “Toward a Disciplinary Matrix in Sociology.” In N. J. Smelser, ed., *Handbook of Sociology*. Beverly Hills, Calif.: Sage.
- Wallerstein, Immanuel 1979 “World Networks and the Politics of the World Economy.” In A. H. Hawley, ed., *Societal Growth: Processes and Implications*. New York: Free Press.
- 1984 “The Three Instances of Hegemony in the History of the Capitalist World Economy.” In G. Lenski, ed., *Current Issues and Research in Macrosociology*. Leiden, Neth.: E. J. Brill.
- Weber, Max (1920) 1978 *Economy and Society: An Outline of Interpretive Sociology*, G. Roth and C. Wittich, eds.; E. Fischhoff, H. Gerth, A. M. Henderson, C. W. Mills, T. Parsons, M. Rheinstein, G. Roth, and C. Wittich, trans. Berkeley: University of California Press.
- Wharton, Amy S. 1994 “Structure and Process: Theory and Research on Organizational Stratification.” *Current Perspectives in Social Theory* supplement 1:119–148.
- Williams, Raymond 1976 *Keywords: A Vocabulary of Culture and Society*. New York: Oxford University Press.
- Williamson, Oliver E. 1990 “Chester Barnard and the Incipient Science of Organization.” In O. E. Williamson, ed., *Organization Theory: From Chester Barnard to the Present and Beyond*. New York: Oxford University Press.
- Znaniecki, Florian 1945 “Social Organization and Institutions.” In G. Gurwitsch and W. E. Moore, eds., *Twentieth Century Sociology*. New York: Philosophical Library.
- Zurcher, Louis A., and David A. Snow 1981 “Collective Behavior: Social Movements.” In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.

HERMAN TURK

## SOCIAL PERCEPTION

Social perception theories and investigations deal with the nature, causes, and consequences of perceptions of social entities, including one’s self, other individuals, social categories, and aggregates or groups to which one may or may not belong. The content of a perception can be virtually any property. Individual attributes may include personality traits, behavioral dispositions, physical characteristics, and ability evaluations. Group attributes can include properties such as

size, cohesiveness, cultural traits, stratification patterns, network patterns, legitimacy, and historical elements. With some notable exceptions, however, the field of social perception traditionally has emphasized the micro side, focusing on individual inferences regarding one individual or a very small number of other individuals.

Social perception is best viewed as an umbrella label that covers a range of loosely related and usually loosely formulated theoretical conjectures and associated research. Today “social cognition” may be the more popular label, subsuming theory and research indexed under numerous other headings: person perception, social judgment, social representation, schema theory, reference group theory, impression formation, attribution theory, and more. Little of this work will be discussed here, although most of it is easy to access. Reference lists in books, chapters, and articles under the various headings tend to intersect rather than be isolated. Review articles have appeared with regularity, and so it is relatively easy to locate the seminal, general, or esoteric references one seeks.

This article provides an introduction and selective overview of the social perception area, with additional attention to some threads that have been or could be of particular interest within sociology. First there is a brief discussion of perception in general, followed by sections that divide the field into three major realms: self-perception, person perception, and group perception.

## PERCEPTION

Social perception is only one manifestation of the general phenomenon of human perception. All perceptions begin with energy-producing events, either inside people or from the environment. Each of the senses operates as a “transducer,” encoding a particular form of energy (e.g., radiant, kinetic, chemical) into neurological signals that are carried to the brain as complex, parallel streams of bioelectrical impulses. In the brain, these streams of information are filtered and transformed through several stages, producing dynamic neural representations almost instantaneously. Depending on anatomic factors, prior experiences, and the nature of the signals, these representations may or may not reach the level of conscious awareness.

When sensations survive this elaborate preprocessing and exceed sensory thresholds, however, they break through into a person's conscious awareness, appearing as coherent and meaningful perceptions: hunger pangs, one's reflection in the mirror, a smile from a friend. These perceptions seem to capture all the essential properties of the events that instigated them.

Roughly speaking, then, people acquire energetic impulses from internal and external environments that, in turn, impinge on the sensory apparatuses as sensations and are transformed by the brain into perceptions. This suggests a close relationship between the perceptions that are formed and the subsequent actions taken on their behalf. For instance, on the basis of perceptions of the personal qualities of others (and perceptions of others' perceptions of those qualities), people make judgments about those qualities (e.g., good or bad); on the basis of those judgments, people formulate intentions about how they will behave toward others (e.g., plan to engage with them or avoid them); on the basis of those intentions, the actions of others, the prevailing context, and so on, people enact their impressions and intentions in social interactions.

Three qualities of perceptions bear further elaboration: *Structure*, *stability*, and *meaning* are definitive subjective properties (Schneider et al. 1979). A fourth quality—*accuracy*—is best understood as an objective property of perceptions, at least in principle. Accuracy and bias in social perception are addressed later in this article.

**Structure.** Humans experience the world as structured. Rather than seeming chaotic and unpredictable, elements and events generally appear to correspond to one another in patterned ways. Things seem to happen for reasons. Much of this patterning is imposed, however, and one person's perceptions may be very different from those of others, even under identical conditions. This is especially relevant in regard to the interpretations people impose on complex phenomena. For instance, people tend not to be aware of how differences in the expectations they bring to a situation color their perceptions. People cannot take in information on everything around them and those expectations direct a person to attend selectively to the available stimuli in a situation. This biasing of attention can have a tremendous influence on

interpretations of the situation. Moreover, each person may impose a unique subjective structure on the same objective reality. Every sports fan has experienced the perception that compared to the opponent, his or her favored team is consistently the victim of "bad calls" by the officials. Supporters of the opponent generally disagree, and rarely does one perceive the opposition to have been treated unfairly by officials. What fans actually "see" are slices of reality unconsciously chosen to conform with their beliefs and expectations.

**Stability.** Different sports fans may see different things, but if pressed, they probably could identify broad areas of the game they agree about and take for granted. They would agree that there were no sudden changes in the sport they were watching; they would profess not to have seen players on the field dematerializing in one place and rematerializing elsewhere; the ball appears to stay the same size and shape even as it moves nearer to and farther from the fans' points of view. In general, while observers may disagree on some points, most of what is observed has an underlying sense of continuity. Indeed, among the myriad sensations to which one might attend in a given situation, the bias is toward those which engender a sense of stability—a sense of the temporal endurance of these patterned sensations.

**Meaning.** If structure and stability were the only properties of experience, the world would appear merely as successions of discrete, insignificant objects and events, each with no particular import transcending the moment. In contrast, most perceptions seem meaningful. That is, perceptions are conceived of as threads in a larger fabric. Through their interconnections and the patterns they form, they seem to have significance, purpose, causes, and consequences beyond their own existence. With cognitive development comes the ability to recognize and select impressions and events that are significant in terms of the information they convey. As will be discussed below, meaningfulness and significance do not imply accuracy. Perceptions—especially social perceptions—are imperfect representations that can be highly misleading.

Among the variety of ways in which one could organize the social perception literature, one of the simplest and most useful proceeds from the individual perceiver, to perceptions of other individuals, and finally to group perceptions. These

categories define the three sections below, which are followed by some remarks about future directions.

### SELF-PERCEPTION

Self-perception is social perception with the self as the object. Through introspection and information from others, people develop beliefs about their many qualities: personality, physical appearance, behavioral tendencies, moral stature, athletic prowess, and the like. "Self-concept" is the general term for the system of beliefs about the self. Although introspection is a source of self-knowledge (Andersen and Williams 1985), mounting evidence suggests that it is not the predominant source that people used to believe it is and that it is generally biased and inaccurate (Nisbett and Wilson 1977; Wilson et al. 1981).

One major branch of self-perception research focuses on inaccuracies in self-knowledge, and a second on how information from others shapes the self-concept. An example of work in the first branch is a review by Greenwald (1980) of evidence of three types of self-conceptual biases: (1) *egocentricity*: the anchoring of judgments, recollections, thought experiments, and attributions about others with reference to the self, (2) *benefactance*: the tendency to perceive the self as generally efficacious, and (3) *cognitive conservatism*: a resistance to cognitive change. Bem's influential self-perception theory (1972) asserts that in conditions of uncertainty, people use their own behavior as a guide to inferences about their inner selves. Later approaches to the self-concept focus on structures such as category systems, conceptual networks, and complex schemas that can represent explicit connections and nonconnections among elements of the self-concept (Greenwald and Pratkanis [1984] provide a review).

The early insights of Cooley (1964[1902]) and Mead (1934) still guide sociological theory and research on the social origins of the self-concept, the second branch mentioned above. Cooley described the "looking-glass self" as the use of others' appraisals as mirrors of the "true" self. Mead noted that the images people form of themselves are greatly affected by how they imagine significant others would respond to and evaluate them. Social comparison theory (Festinger 1954; Suls and Miller 1977) deals with, among other issues, the question of whom one refers to when seeking

comparative self-knowledge and the effects of the various available social referents on one's self-concept and behavior.

Under the rubric of self-perception also are found topics such as self-efficacy, self-evaluation, self-esteem, and self-identity. Self-efficacy is the perception of one's competence with respect to specific tasks (Bandura 1986; Cervone and Peake 1986). Self-esteem is the extent to which one thinks positively about oneself (Rosenberg 1979). The concept of self-evaluation, when distinguished from efficacy and esteem, has been used in theory and research on how the characteristics of evaluators affect self-evaluations in specific, collective task situations (Webster and Sobieszek 1974). There are two major approaches to self-identity: identity theory (Stryker 1980; Burke 1991) and social identity theory (Hogg and Abrams 1988; Hogg et al. 1995). Although there are many shared concepts in these approaches, identity theory is distinguished by a greater emphasis on the performance of social roles as the source of self-definitions; in contrast, social identity theory emphasizes the ways in which self-categorizations hinge on salient properties of the groups with which individuals align themselves.

The more strongly a person's identity is tied to a particular social role or category, the greater is the extent to which that individual empathizes with other occupants of that role or category. It is as if the boundary between self and others became blurred, and the individual empathizes with similar others to whom good things and bad things happen. For instance, an experiment by Markovsky (1985) subtly emphasized self-identification versus group identification and created unjust reward allocations to both individuals and groups. The subjects responded more strongly to the type of injustice that corresponded to their identification.

### PERSON PERCEPTION

The core of social perception theory and research addresses how people formulate impressions about the inner qualities and outward behaviors of other individuals. The focal points for this work include the properties of the people who are perceived and the characteristics of the situations in which a perception is developed, the logic by which basic sensations are integrated to form complex social perceptions, and the way in which perceptions, once formed, are affected by new information.

Attribution theories are concerned with how people form inferences about the causes of others' behaviors. The basic question in these approaches concerns the conditions under which another person's behavior is attributed to an internal disposition or to aspects of the situation in which it occurred (Jones et al. 1989). The so-called fundamental attribution error is the pervasive tendency for observers to underestimate the impact of situational factors on others' behavior (Ross 1977). In fact, people tend to make personal attributions for others' behavior and situational attributions for their own (Jones and Nisbett 1972). Gilbert (1989) has modified this question in a fruitful way by asserting that personal attributions occur automatically; situational attributions occur only as the possible result of an effortful search for additional information (Gilbert et al. 1988).

Although *schemas* could be discussed under the "self-perception" and "group perception" headings, most often they are invoked in theory and research on person perception. Schemas are organized structures of cognitions pertaining to social objects such as the self, other persons, groups, roles, and events (Taylor and Crocker 1981). Thus, one's schema for "college professor" may include beliefs such as "intelligent" and "scattered," negative attitudes such as "inaccessible" and "too political," and expectations for behaviors such as "lecturing" and "conducting research." Schemas have a variety of effects on social perception. For instance, they induce people to attend to certain social and situational features, influence people's judgments by inducing particular expectations for the consequences of their actions, and affect how people recollect social events by making some pieces of information more salient than others. Schemas also transcend individuals by becoming cultural elements that can be communicated among group members or from parent to child.

Other approaches to person perception focus on the integration of bits of information associated with particular others. Information integration theory (Anderson 1981) provides rigorous mathematical models of how an observer employs weighted combinations of another individual's traits to form an overall impression. Social applications of psychophysics (Stevens 1975; Lodge 1981) apply a magnitude scaling technology first developed for expressing judgments of physical properties (e.g., weight, brightness, numerosity, sound pres-

sure, saltiness) to the quantification and validation of judgments of personal or social properties (e.g., competence, fairness, attractiveness). Status characteristics theory (Berger et al. 1985) explains the emergence of status and influence hierarchies in collective, task-performing groups on the basis of individuals' relative standings on combinations of salient characteristics that can order interaction whether or not they are explicitly relevant to a task.

The accuracy of social perceptions was an early research focus but languished for years because of conceptual and methodological roadblocks (Cronbach 1955; Zebrowitz 1990; Fiske 1993). One problem is that determining accuracy requires the existence of a criterion against which a social perception is judged. Often, however, there is no assurance that the criterion is accurate because it may be arbitrary, subjective, or biased. Research in this area has seen a resurgence in recent years, however, partly as a result of approaches like Kenny and Albright's (1987) social relations model. That approach measures the accuracy of judgments of a given characteristic by using multiple observers and targets, permitting the researcher to control for observers' response sets, targets' attributes, and other aspects of the relationships between observers and targets. A mere recent trend is to attempt to disentangle the combined effects of observers' expectancies and targets' characteristics, specifying the conditions under which either set of factors predominates in determining social perceptions.

In a related vein, attributional and social perceptual biases constitute a vast field of inquiry. In recent years, a number of universal human perceptual inclinations have been cataloged that are capable of generating perceptual biases. Many perceptions depend on the ability to gauge one's relevant behaviors and characteristics, yet people often have difficulty assessing their own qualities and properties in an absolute way (Bem 1972). Preconceived notions powerfully influence subsequent perceptions by inducing selective perceptions. Once an idea is accepted, falsifying information is discounted and verifying information is accepted uncritically. People not only are subject to such errors of perception, they also underestimate the degree to which this is so. They are overconfident in their judgments; employ useless, distracting, and unrepresentative information contained in anecdotes; and infer illusory covariations among



social characteristics. In recent years, cognitive and social psychologists have begun to identify and systematically examine these and other types of social perceptual biases. (For some examples, see Taylor and Fiske 1978; Taylor et al. 1978; Nisbett and Ross 1980; Kahneman et al. 1982; and Goldstein and Hogarth 1997.)

### GROUP PERCEPTION

Two sets of approaches to group perception predominate: those concerned with reference group choices and effects, and those addressing social categorization processes. A reference group is a set of individuals whose standing or perspective is taken into account by an actor in selecting a course of action or making a judgment about a specific issue (Farmer 1992).

Research on reference group phenomena represents one of the first and longest-lived attempts in sociology and social psychology to understand how individuals orient themselves to groups, which groups they choose, and the consequences of their choices. (See the early work of Newcomb [1943] and Merton and Rossi [1968] and the more recent review by Singer [1981].) Among the functions of reference groups are providing sources for normative information and offering bases for social comparisons (Gecas 1982). Normative information dictates ostensibly correct and incorrect courses of action and positive and negative values. For example, people may adopt as their own the expressed values of respected members of the community or may assert a position opposite to that held by a disrespected group. In a similar way, social comparisons with reference groups provide bases for evaluating one's beliefs, actions, and accomplishments. For instance, without making reference to the set of people with incomes comparable to one's own, it is impossible to gauge one's level of generosity in donating money to charitable organizations. Three hundred dollars donated in a year may seem high until one discovers that the average donation of people in one's income bracket is ten times that amount.

Although virtually everyone makes use of reference groups, *which* reference groups one selects for one's comparisons and what consequences follow from those selections are more complex issues. Reference group choices have been shown in both natural and experimental settings to be

influenced by numerous factors, including attitude similarity, structural inducements, and normative prescriptions. The consequences of referential comparisons that have been studied include the treatment of social deviants and the emergence of negative social evaluations, changes in self-esteem, and feelings of relative deprivation, gratification, or inequity. Although a good deal of interesting research and many theoretical conjectures have been associated with this area of inquiry, as Singer (1981) noted, there is no reference group theory per se, and the explanatory promise of this area remains unfulfilled. However, many of the research lines spawned by interest in reference groups remain active.

The reference group literature takes as given the existence of groups and the issue of which people are and are not members. Social categorization approaches (Tajfel 1981; Wilder 1986; Abrams and Hogg 1999) are closely related to the social identification literature noted earlier and view the perception of membership versus nonmembership as problematic. In general, people say that they detest being categorized and avoid categorizing others. However, social categorization is a manifestation of a perceptual process that is fundamental to survival. Everyone does it, consciously or not. By learning to recognize and categorize elements of their environments, humans are able to distinguish nutriment from poison and ally from adversary.

Despite its indispensability, the categorization process has side effects in the social realm. The most important and robust of these effects is the tendency for people to overestimate differences between groups and underestimate differences among group members. "They" appear uniform, but "we" are individuals (Quattrone 1986). This phenomenon lies at the heart of stereotyping: the overgeneralization of perceived group attributes (Stangor and Lange 1994). Once formed, stereotypes are maintained by virtue of the types of perceptual biases previously noted, such as forming illusory correlations and relying on anecdotes. A classic finding in research on social identity (Tajfel 1982; Turner 1987) demonstrated that arbitrary we-they distinctions created by random assignments to groups in a laboratory setting were sufficient to produce in-group favoritism and a variety of negative attributions regarding the out-group.

Discrimination—the differential treatment of others solely on the basis of their group memberships—and prejudice—negative attitudes toward certain groups and their members—are common behavioral manifestations of perceptual stereotyping. In American society and in the social and behavioral sciences, gender- and race-based forms of discrimination and prejudice have received the most attention (Eagly 1987; Dovidio and Gaertner 1986, respectively), although the list of bases for discrimination is probably as long as the list of conceivable social characteristics.

### FUTURE DIRECTIONS

Social perception theory and research embrace multiple levels of analysis: cognitive processing, individual and interpersonal behavior, perceptions of groups, and group behavior. The social perception theories that may hold the greatest promise for the future are those amenable to integrating explicit formulations developed within these different levels of analysis. Undoubtedly, much social perception research in the near term will be business as usual, identifying new theoretical contingencies and empirical patterns. However, social and behavioral scientists are developing new approaches to modeling social and social psychological phenomena that may prove fruitful in social perception research.

Burt (1982) integrated a psychophysical model of human perception with explicit models of social network structure. The result is a conceptualization of social groupings at any scale in which network members (1) receive information about certain properties of others (e.g., resource holdings, attitudes), (2) take into account structural information about those others (e.g., the patterns of their social relations and of their relations' relations), (3) evaluate and combine the information received, and (4) make self-referential comparisons involving the information obtained from the network (e.g., relative resource holdings). The models show precisely how structural configurations of social relationships, in combination with individually based social perception and comparison processes, can theoretically account for a far broader class of phenomena than can either individual-level theories that do not consider structures or structural theories that do not consider individuals. Unfortunately, this formulation has

not inspired a corresponding program of research, and the potential contributions of this innovative approach have not been tapped.

Significant progress has been made, however, using network models of a different sort. Within a broader class of approaches known as complexity theory (Eve et al. 1997), neural network models (Read and Miller 1998) and related alternatives (Carley and Svoboda 1996; Macy and Skvoretz 1998; Gilbert and Conte 1995) are beginning to account for social perception phenomena using parallel distributed processing models. This is a type of computer simulation in which numerous interconnected elements (e.g., neurons or agents) repeatedly receive, process, and respond to information from their environments, which may consist largely of similar elements. For example, using this approach, Smith and DeCoster (1998) devised a unified computational model that accomplishes several feats: It learns the social characteristics it “perceives” in individual cases and recognizes those characteristics from partial cues, learns stereotypes from exposure to multiple cases and recognizes those stereotypes from partial cues, and develops novel concepts from old ones. Although so-called connectionist approaches are relatively new, findings such as these bode well for further investigations.

There is no lack of good ideas in the social perception field, and this area may well play a central role in future attempts to integrate micro and macro sociology. Lacking, however, are concerted, programmatic efforts to develop and test explicit and rigorous social perception theories. Some exceptions were noted above. For the most part, however, the absence of explicitness and rigor has resulted in a minimal level of competition among different approaches, virtually no critical testing between formulations, and few time-tested conceptual and methodological refinements. However, this area remains attractive to a large number of psychologists and sociologists, in part because of its many unanswered questions and the ubiquity of its phenomena.

### REFERENCES

- Abrams, Dominic, and Michael A. Hogg, eds. 1999. *Social Identity and Social Cognition: An Introduction*. Oxford, UK: Blackwell.

- Andersen, Susan M., and Marirosa, Williams 1985 "Cognitive/Affective Reactions in the Improvement of Self-Esteem: When Thoughts and Feelings Make a Difference." *Journal of Personality and Social Psychology* 49:1086-1097.
- Anderson, Norman H. 1981 *Foundations of Information Integration Theory*. New York: Academic Press.
- Bandura, Albert 1986 *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice-Hall.
- Bem, Daryl 1972 "Self-Perception Theory." In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 6. New York: Academic Press.
- Berger, Joseph, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch, Jr. 1985 "Expectation-States Theory." In Joseph Berger and Morris Zelditch, Jr., eds., *Status, Rewards, and Influence*. San Francisco: Jossey-Bass.
- Burke, Peter J. 1991 "Identity Processes and Social Stress." *American Sociological Review* 56:836-849.
- Burt, Ronald S. 1982 *Toward a Structural Theory of Action*. New York: Academic Press.
- Carley, Kathleen, and David M. Svoboda 1996 "Modeling Organizational Adaptation as a Simulated Annealing Process." *Sociological Methods and Research* 25(1):138-168.
- Cervone, Daniel, and Philip K. Peake 1986 "Anchoring, Efficacy, and Action: The Influence of Judgmental Heuristics on Self-Efficacy Judgments and Behavior." *Journal of Personality and Social Psychology* 50:492-501.
- Cronbach, Lee J. 1955 "Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity.'" *Psychological Bulletin* 52:177-193.
- Cooley, Charles Horton 1964 *Human Nature and the Social Order*. New York: Schocken [originally published 1902].
- Dovidio, John F., and Samuel L. Gaertner, eds. 1986. *Prejudice, Discrimination, and Racism: Theory and Research*. Orlando, Fla.: Academic Press.
- Eagly, Alice H. 1987 *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, N.J.: Erlbaum.
- Eve, Raymond A., Sara Horsfall, and Mary E. Lee, eds. 1997 *Chaos, Complexity and Sociology: Myths, Models, and Theories*. Thousand Oaks, Calif.: Sage.
- Farmer, Yvette 1992 "Reference Group Theory." In Edgar F. Borgatta and Marie L. Borgatta, Eds., *Encyclopedia of Sociology*. New York: Macmillan.
- Festinger, Leon 1954 "A Theory of Social Comparison Processes." *Human Relations* 7:117-140.
- Fiske, Susan T. 1993 "Social Cognition and Social Perception." *Annual Review of Psychology* 44:155-194.
- Gecas, Viktor 1982 "The Self-Concept." *Annual Review of Sociology* 8:1-33.
- Gilbert, Daniel T. 1989 "Thinking Lightly about Others: Automatic Components of the Social Inference Process." In James S. Uleman and John A. Bargh, eds., *Unintended Thought: Limits of Awareness, Intention, and Control*. New York: Guilford.
- , Brett W. Pelham, and Douglas S. Krull 1988 "On Cognitive Busyness: When Person Perceivers Meet Persons Perceived." *Journal of Personality and Social Psychology* 54:733-740.
- Gilbert, Nigel, and Rosaria Conte, eds. 1995 *Artificial Societies: The Computer Simulation of Social Life*. London: UCL Press.
- Goldstein, William M., and Robin M. Hogarth, eds. 1997 *Research on Judgment and Decision Making*. New York: Cambridge University Press.
- Greenwald, Anthony G. 1980 "The Totalitarian Ego: Fabrication and Revision of Personal History." *American Psychologist* 35:603-618.
- , and Anthony R. Pratkanis 1984 "The Self." In Robert S. Wyer, Jr., and Thomas K. Srull, eds., *Handbook of Social Cognition*, vol. 3. Hillsdale, N.J.: Erlbaum.
- Hogg, Michael A., and Dominic Abrams 1988 *Social Identifications*. London: Routledge.
- , Deborah J. Terry, and Katherine M. White 1995 "A Tale of Two Theories: A Critical Comparison of Identity Theory with Social Identity Theory." *Social Psychology Quarterly* 58:255-269.
- Jones, Edward E., A. Flammer, A. Grob, R. Luthi, J. Rubin, and G. Fletcher 1989 "Attribution Theory." In J. Forgas and M. Innes, eds., *Recent Advances in Social Psychology: An International Perspective*. Amsterdam: North-Holland.
- , and Richard E. Nisbett 1972 "The Actor and the Observer: Divergent Perceptions of Causality." In Edward E. Jones, David E. Kanouse, Harold H. Kelley, Richard E. Nisbett, Stuart Valins, and Bernard Weiner, eds., *Attribution: Perceiving the Causes of Behavior*. Morristown, N.J.: General Learning Press.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. 1982 *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Kenny, David A., and Linda Albright 1987 "Accuracy in Interpersonal Perception: A Social Relations Analysis." *Psychological Bulletin* 102:390-402.
- Lodge, Milton 1981 *Magnitude Scaling*. Beverly Hills, Calif.: Sage.
- Macy, Michael W., and John Skvoretz 1998 "The Evolution of Trust and Cooperation between Strangers: A

- Computational Model." *American Sociological Review* 63:638–660.
- Markovsky, Barry 1985 "Toward a Multilevel Distributive Justice Theory." *American Sociological Review* 50:822–839.
- Mead, George Herbert 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Merton, Robert K., and Alice S. Rossi 1968 "Contributions to the Theory of Reference Group Behavior." In Robert K. Merton, *Social Theory and Social Structure*, enlarged ed. New York: Free Press.
- Newcomb, Theodore M. 1943 *Personality and Social Change*. New York: Holt, Rinehart, and Winston.
- Nisbett, Lee, and Timothy DeCamp Wilson 1977 "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84:231–259.
- Nisbett, Richard E., and Lee Ross 1980 *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, N.J.: Prentice-Hall.
- Quattrone, George A. 1986 "On the Perception of a Groups' Variability." In Stephen Worchel and William G. Austin, eds., *Psychology of Intergroup Relations*, 2nd ed. Chicago: Nelson-Hall.
- Read, Stephen J., and Lynn C. Miller, eds. 1998 *Connectionist Models of Social Reasoning and Social Behavior*. Mahwah, N.J.: Erlbaum.
- Rosenberg, Morris 1979 *Conceiving of the Self*. New York: Basic Books.
- Ross, Lee 1977 "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process." In Leonard Berkowitz, ed. *Advances in Experimental Social Psychology*, vol. 10. New York: Academic Press.
- Schneider, David J., Albert H. Hastorf, and Phoebe C. Ellsworth 1979 *Person Perception*, 2nd ed. Reading, Mass.: Addison-Wesley.
- Singer, Eleanor 1981 "Reference Groups and Social Evaluations." In Morris Rosenberg and Ralph H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Smith, Eliot R., and James DeCoster 1998 "Person Perception and Stereotyping: Simulation Using Distributed Representations in a Recurrent Connectionist Network." In Stephen J. Read and Lynn C. Miller, eds., *Connectionist Models of Social Reasoning and Social Behavior*. Mahwah, N.J.: Erlbaum.
- Stangor, Charles, and J. E. Lange 1994 "Mental Representations of Social Groups: Advances in Understanding Stereotypes and Stereotyping." In Mark P. Zanna, ed., *Advances in Experimental Social Psychology*, vol. 26. San Diego: Academic Press.
- Stevens, S. S. 1975 *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. Beverly Hills, Calif.: Sage.
- Stryker, Sheldon 1980 *Symbolic Interactionism: A Social Structural Version*. Menlo Park, Calif.: Benjamin Cummings.
- Suls, Jerry M., and Richard L. Miller, eds. 1977 *Social Comparison Processes*. Washington, D.C.: Halsted-Wiley.
- Tajfel, Henri 1981 *Human Groups and Social Categories*. Cambridge, UK: Cambridge University Press
- 1982 *Social Identity and Intergroup Relations*. London: Cambridge University Press.
- Taylor, Shelley E., and J. Crocker 1981 "Schematic Bases of Social Information Processing." In E. T. Higgins, C. P. Herman, and M. P. Zanna, eds., *Social Cognition: The Ontario Symposium*. Hillsdale, N.J.: Erlbaum.
- , and Susan T. Fiske 1978 "Salience, Attention, and Attribution: Top of the Head Phenomena." In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 11. New York: Academic Press.
- , Susan T. Fiske, Nancy L. Etcoff, and Audrey J. Ruderman 1978 "The Categorical and Contextual Bases of Person Memory and Stereotyping." *Journal of Personality and Social Psychology* 36:778–793.
- Turner, John C. 1987 *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, UK: Basil Blackwell.
- Webster, Murray, and Barbara Sobieszek 1974 *Sources of Self Evaluation*. New York: Wiley.
- Wilder, David A. 1986 "Social Categorization: Implications for Creation and Reduction of Intergroup Bias." Pp. 291–355 In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 19. New York: Academic Press.
- Wilson, Timothy D., Jay G. Hull, and Jim Johnson 1981 "Awareness and Self-Perception: Verbal Reports on Internal States." *Journal of Personality and Social Psychology* 40:53–71.
- Zebrowitz, Leslie A. 1990 *Social Perception*. Pacific Grove, Calif.: Brooks/Cole.

BARRY MARKOVSKY

## SOCIAL PHILOSOPHY

American sociology, with the significant exception of symbolic interactionism, generally has turned to Europe for a philosophical grounding. The years after World War II, when Talcott Parsons at Harvard and Paul Lazarsfeld and Robert Merton

at Columbia dominated the field, were no exception. Parsons had studied in Germany and had translated Weber; Lazarsfeld brought to the United States an Austrian philosophical heritage; Merton was a specialist in the philosophy and sociology of science. Under the guidance of these three men, American sociology was strongly influenced by the philosophers of the Vienna Circle, especially Carnap and Popper, and to a lesser degree by British logical positivism.

The common ingredient in this philosophical heritage was the notion that there were few obstacles to the creation of a science of human behavior modeled on the natural sciences. A great deal of attention was paid to hypothesis testing, criteria for evidence, and the nature of statistical proof. Most important, this way of thinking assumed that there was a social reality that would prove to have the same nature as physical reality. A predictive science of human behavior was possible because the social world was assumed to work in lawlike ways. The problems of the social sciences lay in discovering reality, not in the nature of reality.

This assumption would be shaken from two different directions. On the one hand, in many cases sociologists would lose their faith in the natural science model as the most appropriate way to think about society. In part as a result of the political upheavals of the 1960s, adherence to the patient accumulation of facts to verify what Merton called "middle level" theories about the world was difficult to maintain. On the other hand, philosophers increasingly came to question the epistemology associated with the natural science model. Kuhn's thesis that scientific insight resulted from a new paradigm had a stunning effect on social science, while Polanyi argued for a more "artistic" and "sociological" conception of scientific inquiry (Kuhn 1970; Polanyi 1968). As epistemological skepticism invaded the natural sciences, its implications for the social sciences became even more serious.

One of the first consequences of this increasing skepticism was the discovery in America that the European philosophical heritage was far broader than it first had appeared. Weber, for example, had been influenced strongly by Nietzsche, yet Parsons's interpretation of Weber downplayed the significance of heroism, irrationality, and cultural pessimism and presented Weber as an American

pluralist. The two philosophers who perhaps had the greatest influence in Europe during and after Parsons's visits to Germany were Heidegger and Husserl, and neither played a significant role in Parsons's outlook. The late Wittgenstein was discovered to be quite different from the earlier one, who had been convinced that pure logic would make philosophy unnecessary. French existentialism would become prominent in the 1960s, yet its dominance would not last, and its influence on American sociology was minimal. What had been a selective and partial reading of European philosophy by American sociologists could no longer last.

Even during the years when Parsons was the leading American sociologist, alternatives had existed. Husserl's phenomenology was one. As brought to the United States and applied to sociology in the work of Schutz (1967), phenomenology argued for the importance of the "life-world," the everyday events out of which people's understanding of the world around them becomes possible. Schutz was unable to convince Parsons of the importance of phenomenology (Schutz 1978), but he did influence one of Parsons's most brilliant students, Garfinkel (1967). Ethnomethodology became the most important alternative to structural functionalism in the 1960s and 1970s. The legacy of phenomenology also could be seen in the work of other sociologists, such as Berger and Luckman (1967). Other alternative traditions, such as the influence of Wittgenstein, also existed during this period, especially in Great Britain. Finally, among Marxists, the traditions of the Frankfurt School and the legacy of Lukacs constituted an important basis for social science theorizing (Arato and Gebhart 1977).

As the wide variety of ideas associated with European philosophy became increasingly known to sociologists, confidence in the natural sciences as a model for the social sciences gave way to far more nuanced approaches. No theorist in contemporary sociology was more aware than Habermas (1987) of the necessity of incorporating insights from a wide variety of philosophical traditions. Not only did Habermas bring to the task of sociological theorizing his background in Frankfurt School Marxism, plus a far more realistic reading of Weber than the one offered by Parsons, he also grounded his work in British linguistic philosophy, the Schutzian life-world, and American pragmatism. To read Habermas was to learn about

those whom Habermas had read, bringing wide swaths of European philosophy to the attention of many American readers. Habermas was not alone in trying to synthesize philosophical traditions with the social sciences. Bourdieu (1984) in France was equally influential in sociology; his combination of theoretical insight, empirical investigation, and concern with how knowledge is produced made him a major figure in American sociology (Bourdieu and Waquant 1992; Lamont and Fournier 1992).

Habermas in particular argued for the importance of modernity, for the possibility of using rationality as a standard by which communicative utterances could be judged, thus making possible a social order held together by norms that lay outside the purely subjective preferences of the individuals who constituted that social order. This belief in reason would lead others to criticize Habermas as excessively rationalistic and therefore too close to the assumptions of a nonproblematic reality that had guided Parsons (Lyotard 1984). Postmodernism would become the most radical challenge imaginable to the earlier Parsons's faith in a scientific model for the social sciences.

With postmodernism, the philosophers once ignored by American social science became the most important ones to read. Foucault (1971), under the influence of Nietzsche, argued that knowledge is the product of a general *episteme* that in itself is not a reflection of a reality in the world but the product of a particular historical period and its self-understanding. Lyotard (1984), borrowing from Wittgenstein, viewed science as a "language game" preoccupied with strategy and tactics, anything but a dispassionate and objective search after truth. Derrida (1978), under the influence of Heidegger, argued for the indeterminacy of concepts such as truth, justice, and morality. Philosophers could not seek fixed universals; they were engaged in the practice of rhetoric, defending or attacking contingent, local, and socially constructed practices that defined truth, justice, and morality in the interests of certain groups and against the interests of others.

Postmodernism has not had the impact in sociology that it has had in literary criticism, history, and law. Numerous sociologists continue to study the empirical world by testing various hypotheses on the basis of evidence collected through

surveys, demographic data, and other essentially quantitative methods. However, the postmodern challenge to normal science is likely to prove to be significant. Postmodernism is the culmination of all the challenges to the Parsonian consensus that once existed in the field; it represents what seems to be an end point to the process of questioning the existence of a nonproblematic social reality that can be understood by an observer standing outside that reality.

The question raised by the varieties of epistemological skepticism currently prevalent in the humanities is whether any science, let alone a social science, is possible. Just as some scientific fields seek to reduce the laws of one field, such as biology, to those of another, say, biochemistry, postmodernism argues that all fields of inquiry can be reduced to the study of rhetoric. In a perhaps unintended fashion, the implication of this argument is a hegemonic one: The rhetorical methods associated with literary criticism will become the model for all inquiry, just as science was once understood to be.

The rhetorical issues involved in social science theorizing have been analyzed in the case of economics by McCloskey (1985). In addition, some work in the sociology of science has made it clear that scientists have strategies by which they present themselves to the world and hope to gain acceptance (for an example, see Latour [1987]). However, does it follow that there is no grounding for social science, no Durkheimian facts in the world to which one can appeal to resolve disputes about knowledge? There is, of course, no way to answer such a large question satisfactorily in a brief article, but it is possible to offer the hope that both science and rhetoric can play a role in the sociological enterprise.

Sociology began, as Lepenies (1988) has argued, "between science and literature." Its most important theorists were attracted to positivistic understandings of knowledge, but they were also essayists dealing with some of the most significant issues in the moral philosophy of their time. One could argue that this "ambivalence" between fields has consistently characterized sociology at its best (Merton 1976). Merton, for example, though engaged in what he called the "systematics" of theory, was also a historian of science, wrote with

reference to the entire Western tradition of philosophy and literature, and was very much part of the milieu of the New York intellectuals of his day (Merton 1967, pp. 1–37).

The fate of sociology may lie in finding a balance between a scientific grounding in fact and an appreciation of how rhetorical strategies affect the ways in which scholars argue about facts. Unlike earlier sociologists, sociologists today cannot be confident that there will exist a one-to-one relationship between social reality and its representations; we must always be skeptical about the likelihood that the indicators we use actually measure the real world. But this does not mean that sociologists should give up collecting evidence, testing hypotheses, or trying to establish facts. It means only that the truths uncovered by these methods are not universals that exist for all time but are contingent on historical periodization and location.

From this perspective, the question becomes one of the length of the historical periods and the width of the locations by which we can judge the truths we discover. If we could determine that it is possible to discover regularities in, for example, the way liberal democracies have organized themselves over the past 200 years, that would be a significant and important accomplishment. It would mean neither that we had discovered an unchanging reality nor that our discoveries are simply part of a rhetorical strategy. A sociology that modeled itself neither on physics nor on literary reading, but combined elements of both would be a sociology more chastened than Parsonianism but more hopeful than postmodernism.

(SEE ALSO: *Epistemology; Human Nature; Marxist Sociology; Postmodernism; Pragmatism*)

## REFERENCES

- Arato, Andrew, and Eike Gerhardt, eds. 1977 *The Frankfurt School Reader*. New York: Urizen Books.
- Berger, Peter, and Thomas Luckmann 1967 *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, N.Y.: Doubleday.
- Bourdieu, Pierre. 1984 *Distinction: A Social Critique of the Judgement of Taste*. Richard Nice, trans. Cambridge, Mass.: Harvard University Press.
- , and Lois Wacquant 1992 *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.
- Derrida, Jacques 1978 *Writing and Difference*. Alan Bass, trans. Chicago: University of Chicago Press.
- Foucault, Michel 1971 *The Order of Things: An Archeology of the Human Sciences*. New York: Pantheon.
- Garfinkel, Harold 1967 *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.
- Habermas, Jurgen 1987 *Lifeworld and System: A Critique of Functionalist Reason*, vol. 2 of *The Theory of Communicative Action*. Thomas McCarthy, trans. Boston: Beacon Press.
- Kuhn, Thomas S. 1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lamont, Michele, and Marcel Fournier 1992 *Cultivating Differences: Symbolic Boundaries and the Making of Inequality*. Chicago: University of Chicago Press.
- Latour, Bruno 1987 *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, Mass.: Harvard University Press.
- Lepenes, Wolf 1988 *Between Literature and Science: The Rise of Sociology*. R. J. Hollingdale, trans. Cambridge, UK: Cambridge University Press.
- Lyotard, Jean-François 1984 *The Post-Modern Condition: A Report on Knowledge*. Geoff Bennington and Brian Massumi, trans. Minneapolis: University of Minnesota Press.
- McCloskey, Donald 1985 *The Rhetoric of Economics*. Madison: University of Wisconsin Press.
- Merton, Robert K. 1967 *On Theoretical Sociology: Five Ways, Old and New*. New York: Free Press.
- 1976 *Sociological Ambivalence and Other Ways*. New York: Free Press.
- Polanyi, Michael 1968 *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: University of Chicago Press.
- Schutz, Alfred 1967 *The Phenomenology of the Social World*. George Walsh and Frederick Lehnert, trans. Evanston, Ill.: Northwestern University Press.
- 1978 *The Theory of Social Action: The Correspondence of Alfred Schutz and Talcott Parsons*. Richard Grathoff, ed. Bloomington: Indiana University Press.

ALAN WOLFE

## SOCIAL POWER

See Social and Political Elites; Social Control.

## SOCIAL PRESTIGE

See Occupational Prestige; Status Attainment; Social Mobility.

## SOCIAL PROBLEMS

The discipline of sociology was born during a century of rapid social change attributable largely to the Industrial Revolution. Social theorists in nineteenth-century Europe devoted much of their attention to the institutional consequences of the erosion of the old social structure. American sociologists in the late nineteenth and early twentieth centuries, especially at the University of Chicago, added an intellectual orientation derived from the political idealism and meliorist pragmatism then fashionable in this country. "Social problems" generally were understood to be conditions that disrupted peaceful social life (e.g., crime) or produced obvious human misery (e.g., poverty) and that could be eliminated or alleviated by means of enlightened social policy and effective social engineering. Nearly all the conditions envisioned as social problems were associated with burgeoning cities and the dispossessed immigrants (foreign and domestic) they attracted. Few scholars doubted that these problems could be objectively diagnosed and treated, much like a malady in the human body.

By midcentury, however, American sociology was influenced increasingly by the functionalism and positivism of some of the earlier European intellectuals, whose major works had only recently been translated into English. Under this influence, sociologists assumed a more detached and value-neutral posture toward traditional social problems. While still acknowledging certain social conditions as problematic, most sociologists limited their responsibilities to scientific analysis of problems and their causes, leaving meliorist activism to politicians, social workers, and interest-group partisans. A split within the American Sociological Association between the more detached and the more activist visions of the discipline led to the formation in 1952 of the Society for the Study of Social Problems, which generally favors a more actively meliorist role for sociology. Since its inception, that society has published the quarterly *Social Problems*, the journal that most defines the

content and direction of social problems theory and research in American sociology.

## THE TRADITIONAL PARADIGM

Historically, the reigning paradigm in the sociology of social problems has properly been called an *objectivist* one. From this perspective, social problems are conditions that (1) are in some sense undesirable to any trained and objective observer and (2) are in one way or another amenable to alteration. Interest-group differences are acknowledged in regard to how the problems should be defined and corrected, but it is taken for granted that the problems themselves exist as objective realities in the form of structural arrangements, material conditions, institutional processes, interactional patterns, and the like. In much of the literature on social problems, especially textbooks, there is an implicit analogy to medical diagnosis; that is, a social problem is to society as a disease is to the body, an objective reality apart from popular opinion. Often conventional medical terms are even used, such as *pathology* (or *social pathology*), *epidemiology*, and *etiology*, not only for social problems that might be quasi-medical in nature (such as alcoholism and mental illness) but even for those which are entirely societal (such as poverty and juvenile delinquency). From this point of view, social problems require the investigation of diagnostic experts (social scientists) who can objectively discern the nature of a problem and the most effective way of addressing it and thus enhancing the welfare of people and society.

These diagnostic experts also come from different schools of thought, as is often the case in other disciplines. Two theoretical orientations have tended to dominate the diagnosis and analysis of social problems by sociologists who work within the traditional paradigm. The first is functionalism in one version or another. From a functionalist perspective, society is seen as a more or less organic entity made up of interdependent parts (i.e., institutions). Breakdowns occasionally occur, but the whole generally succeeds in maintaining its natural state of equilibrium. These breakdowns (also called *dysfunctions* or *social disorganization*) are social problems; once they are fixed, the social system can return to a normal state of functioning. This perspective has been criticized by activists as being informed by the conservative assumption



that the existing social system is acceptable as it is, with only meliorative adjustments needed to deal with the occasional breakdown.

The second school of thought within the traditional paradigm, one generally preferred by activists and derived mainly from the Marxist heritage, is a more radical theoretical perspective sometimes called *critical theory*. From this perspective, social problems are the inevitable and endemic characteristics of a capitalist system. Efforts to deal with social problems in such a system can never produce more than temporary palliation, or, worse still, only co-op discontent and protest in the class interests of the dominant and oppressive establishment. Therefore, only the drastic overhaul or total overthrow of the capitalist system can produce a society free of social problems. This radical perspective shares with its more conservative functionalist counterpart the premise that social problems can in principle be objectively identified, diagnosed, analyzed, and corrected.

The belief that social problems have objective bases, which is the foundation of the traditional paradigm, continues to inform public policy at local, state, and federal levels of government in the United States and elsewhere. Accordingly, public funding for research on and amelioration of conditions officially designated as social problems has helped promote the objectivist paradigm in sociology. In the war on poverty of the 1960s, the war on drugs of the 1990s, and the war on violence at the beginning of the twenty-first century, billions of dollars in government support have gone into research on the prevention and/or correction of the social problems deemed most prevalent and consequential in a given era. Both the objectivist paradigm and the professional careers of its proponents, including sociologists with an applied orientation, have been strengthened in the process.

### THE EMERGENT PARADIGM

While the objectivist approach continues to dominate textbooks and policy-oriented sociology, the academic literature on social problems in recent decades, including much that has appeared in *Social Problems*, has provided a robust alternative theoretical paradigm. Generally called *subjectivist*, in contrast to its traditional counterpart, this paradigm derives from different epistemological pre-

ises. With roots in phenomenology and symbolic interactionism rather than in the positivism of the objectivist paradigm, the subjectivist approach to social problems in its original formulation is akin to what is sometimes called the *labeling* perspective in the study of crime and deviance. Intimations of this subjectivist paradigm can be found in some nineteenth-century European theoretical literature and even in the work of some American sociologists earlier in the twentieth century (Fuller 1937; Fuller and Myers 1941a, 1941b; Mills 1943; Waller 1936). The emergence of this paradigm into the mainstream of the discipline, however, has occurred mainly since the 1960s in the work of Becker (1966), Blumer (1971), and the team of Kitsuse and Spector (Kitsuse and Spector 1973; Spector and Kitsuse 1973, 1977). More recently, the paradigm has been promoted in the works of Best (1989, 1990) and the team of Holstein and Miller (1993; Miller and Holstein 1993a).

In the simplest terms, the subjectivist paradigm holds that a social problem lies in the eye of the beholder, not in objective reality. Certain social *conditions* may be real (e.g., deviance, inequality, depletion of the ozone layer, and natural disasters), but to term them *problems* requires an evaluative judgment not inherent in the condition itself; thus, there is a distinction between a (real or imagined) social condition and its acquisition of the standing as a recognized social problem. In this formulation, social problems are best seen as projections of collective sentiments and representations rather than as mirrors of objective conditions (Best 1989; Holstein and Miller 1993; Mauss et al. 1975; Miller and Holstein 1993a; Spector and Kitsuse 1977). Collective sentiments and representations reflect judgments emanating from the political, economic, cultural, moral, religious, and other interests of the persons who allege the existence of a problem. Accordingly, the subjectivist paradigm denies the analogy to medical diagnosis on the grounds that there are no generally accepted scientific standards in sociology, as there are in medicine, for judging what is “pathological.” For example, in the study of societies there is no counterpart to the medical standard that “normal” human body temperature is 98.6 degrees Fahrenheit (and that significant departures from that figure are symptomatic of pathology). Instead, the judgment that a social condition is undesirable or problematic in any sense is entirely

relative to cultural (or subcultural), generational, and many other social variables.

As evidence of relativity and variability in the definition of the term "social problems," subjectivists make at least two historical observations. First, social conditions once regarded as serious problems in the United States are no longer so regarded, whether or not they are still extant. One example is witchcraft, which now is considered never to have existed "in reality" but nevertheless was recognized as a major social problem in seventeenth-century Massachusetts and elsewhere. Other, more recent examples include miscegenation, which gave rise to a powerful eugenics movement early in the twentieth century; prostitution, which is now legal in some jurisdictions, and illegal in others but nowhere regarded as the national social problem that it was at one time; and homosexuality, which once rated attention, along with prostitution, in most social problems textbooks and even a listing as a disorder in the official diagnostic manual of psychiatry but now often is considered a legitimate lifestyle and is protected by civil rights legislation. In none of these examples has there been any evidence of a reduction in the *actual incidence* of the condition, and so the erosion in their status as "real" social problems can be attributed only to changes in public *perception*.

The second observation, related to the first, is that no consistent relationship is apparent between the waxing and waning of given social *conditions*, on the one hand, and the official or public designation of those conditions as problems, on the other hand. For example, the official and professional recognition of racial discrimination as a social problem in the United States came not while the Jim Crow regime was at its worst (before World War II) but only after the lot of African-Americans had begun to improve significantly. Similarly, it is only since the 1980s that so-called hate crimes (crimes motivated by bias based on race, religion, ethnicity, sexual orientation, gender, etc.) have entered into national discourse despite the fact that this type of intergroup violence is as old as humankind. In the opposite direction, the war on poverty of the 1960s was effectively dropped from the national agenda of social problems with no change in the actual incidence of poverty. Similarly, the national war on drugs was officially declared after more than a decade of *decline* in the incidence of alcohol and

drug use by both adults and youths in the United States.

This common and paradoxical lack of correspondence between *objective* social *conditions* and the ebb and flow of social *problems* means that the two phenomena can vary independently. From the subjectivist viewpoint, the traditional objectivist focus has made the mistake of studying not social problems but only certain social conditions. In evaluating these conditions as problems, objectivists have not assumed the positivist, value-neutral stance of science, as they have supposed; instead, they have unquestioningly accepted the transitory evaluative definitions of interest groups, government funding agencies, and a fickle public. If the objectivists have been studying the wrong thing, what *should* the focus be for sociologists of social problems?

From the subjectivist viewpoint, the sociological study of social problems, properly understood, is the study of the process by which putative social conditions come to be defined as social problems by governments and publics (Spector and Kitsuse 1977). In this formulation, social problems are reconceptualized in terms of discourse, interactions, and institutional practices by which only some social conditions are defined as social problems and given a place in the social problems marketplace (Best 1990; Hilgartner and Bosk 1988). The social conditions themselves can be left to sociologists with other specialties: Criminologists can and should study crime; family sociologists can and should study divorce, spousal conflict, and child-rearing stresses; environmental sociologists can and should study the ways in which the physical environment and the organization of social life interact; specialists in stratification can and should study inequality; and so on. None of this other work, however important, constitutes the sociology of social problems per se, for social problems do not originate in social conditions. They originate in *claims-making activities* of interest groups and partisans who undertake to gain political acceptance for their perceptions of certain conditions as "problems," after which those perceptions reflect the evolution of social problems construction (Spector and Kitsuse 1977; Jenness 1993). It is those claims-making activities, then, that should constitute the focus of study in the sociology of social problems. Put another way, the study of social problems is a study of the social construction of reality (Berger and Luckman 1966).

Thus, subjectivists who study social problems in this way also are called *social constructionists* (or simply *constructionists*).

To understand how social problems are constructed, Kitsuse, Spector, and their disciples focus primarily on claims-making activities. These activities consist of the usual tactics of interest groups in asserting claims and grievances, including boycotts, demonstrations, and lawsuits, as well as the strategic use of terms, labels, semantics, and other rhetorical devices, to win political and public support for definitions of certain putative social conditions as problems (Spector and Kitsuse 1977, pp. 9–21, 72–79; Schneider 1985, pp. 213–218, 224). The particular arena within which claims-making activities occur can be as large as American society or as small as a college campus, a corporation, or even a professional society. Whatever the arena, it is claims-making activities or “social problems work” that constitutes the appropriate focus of research in social problems (Miller and Holstein 1989).

While the subjectivist or social constructionist paradigm is relatively recent as a theoretical alternative to the positivist or objectivist tradition in the study of social problems, it is actually cognate to several other well-established lines of inquiry in sociology. One of these fields is *cultural analysis*, represented especially in the work of Douglas (1966; Douglas and Wildavsky 1982). Broadly defined, cultural analysis is the study of the production and distribution of culture (including popular culture). As applied to social problems, cultural analysis includes the assessment of risk as a cultural product, much like a religious ideology (Stallings 1990). Miller and Holstein (1989) see a parallel here to the early Durkheimian concept of *collective representations*, which they understand as a cultural product of social problems work (see also Gusfield 1981).

A second theoretical link can be made between the social constructionist approach to social problems and *conflict theory* (Pong 1989). Inspired partly by Marxism, conflict theory, like critical theory, rejects the functionalist assumption that the natural condition of society is a state of equilibrium among interdependent parts (institutions). Instead, it emphasizes the naturally conflicting tendencies in society among economic, political, religious, and other interest groups (Eitzen and

Zinn 1988; Skolnick and Currie 1985). While the objectivist tendency in the way conflict theory defines social problems does not accord well with the social constructionist paradigm, the claims-making activities that are so important in that paradigm are classical examples of the political struggle and agitation that constitute the typical focus of conflict theorists and, more recently, of the critical theorists working in the deconstructionist (Michaelowski 1993; Pfohl 1993), feminist (Gordon 1993), postmodernist (Agger 1993), and poststructuralist (Miller 1993) traditions. These recent critical challenges “represent promising opportunities to reconsider assumptions, distinctions, and categories on which social constructionism and social problems theory rests” (Miller and Holstein 1993c, p. 536).

The social constructionist perspective converges perhaps even more closely with a third important preoccupation of sociology: *social movement theory*. The claims-making activities that are the principal focus of the constructionist approach are also the classical tactics and strategies of social movement activists. This would include both rhetorical products, such as ideology and propaganda, and mobilizing activities, such as demonstrations, agitation, and political organization (Hilgartner and Bosk 1988; Mauss et al. 1975; Mauss 1989). The “resource mobilization” approach to social movements that has gained currency in recent years seems especially close to the claims-making focus of the subjectivist or constructionist perspective on social problems (Zald and McCarthy 1987; Turner 1981).

At least in part because of the many ways in which the subjectivist approach to social problems has converged with other literatures in the social sciences, it has come to dominate scholarly approaches to social problems theory. Indeed, by the mid-1980s Schneider, a former editor of *Social Problems*, concluded in the *Annual Review of Sociology* that “over the last two decades the constructionist approach to social problems has constituted the only serious and sustained recent discussion of social problems theory” (Schneider 1985, p. 210). By the mid-1990s, it had produced “a torrent of empirical studies” (Miller and Holstein 1993b, p. 3) that focus on diverse social conditions and activities, including alcohol and driving (Gusfield 1981), hyperactivity (Conrad 1975), child abuse (Best 1990), AIDS (Albert 1989), alcoholism

(Schneider 1978), cigarette smoking (Markle and Troyer 1979), crime (Fishman 1978), rape (Rose 1977), homosexuality (Spector 1977), hate crime (Jenness and Broad 1997), premenstrual syndrome (Rittenhouse 1992), chemical contamination (Aronoff and Gunter 1992), drugs (Orcutt and Turner 1993), satanism (Richardson et al. 1991), prostitution (Jenness 1993), and even earthquakes (Stallings 1995). True to the constructionist perspective, this work has been sustained by research questions and procedures focused not on social conditions but on the full range of political and claims-making activities that cause these activities to be understood as social problems.

As demonstrated by Mauss (Mauss et al. 1975; 1989) these are also the activities typically associated with social movements. Accordingly, an issue among subjectivists and their allies is the question of whether the constructionist approach to social problems is equivalent to the study of social movements. The 1985 and 1989 annual meetings of the Society for the Study of Social Problems featured the theme of social problems as social movements. At those meetings and in subsequent publications, proponents of the subjectivist perspective have not agreed on whether the study of social problems should be subsumed by social movement theory or, alternatively, should be approached as a theoretically distinct field of study. Some would say that the dominant way in which social conditions come to be seen as social problems is through the work of social movements (Gerhards and Rucht 1992; Mauss 1989; Troyer 1989; Holstein and Miller 1993; Miller and Holstein 1993a). As early as 1975, Mauss (p. 38) presented the case for considering "social *problems* as simply a special kind of *movement*." This case rests largely on the proposition that the characteristics of social problems are typically those of social movements and that social problems are always outcomes of social movements. In this formulation, the study of social movements and the study of social problems are rendered compatible through an examination of the genesis of social problems movements, the organization and mobilization of social movements, the natural history of social movements, and the decline and legacy of a social problem. More recently, Bash (1995, pp. xiii-xiv) argued "that what is addressed as the Social Movement, in the one instance, and what is targeted as a host of social problems, in the other, may not reflect *distinctive*

sociohistorical phenomena at all." In contrast, Troyer (1989) has noted the similarities *and* differences between the definitional approach, the standard structural approach, and the more recent resource mobilization approach to social movements (Zald and McCarthy 1987). More recently, Klandermans (1992, p. 77) has observed that "many situations that could be considered a social problem never become an issue, even though they may be no less troublesome than situations that do become a rallying point. Further, a social problem does not inevitably generate a social movement" (he does not cite any examples).

A more vigorous debate among subjectivists began with the provocative critique offered by Woolgar and Pawluch (1985) of the "ontological gerrymandering" they attributed to the constructionist analysis of social problems. A successful constructionist explanation for social problems, they argued, "depends on making problematic the truth status of certain states of affairs selected for analysis and the explanation, while backgrounding or minimizing the possibility that the same problems apply to assumptions upon which the analysis rests" (Woolgar and Pawluch 1985, p. 216). Since constructionist analysis is itself a claims-making activity, these authors observe, it involves selective relativism and theoretical inconsistency: Constructionists invoke subjectivist assumptions about the phenomena under study but at the same time present their claims as objective facts.

In an attempt to rescue the constructionist paradigm from this apparent inconsistency, Ibarra and Kitsuse (1993, p. 26) have proposed replacing the term "putative condition" with the term "condition-category" as the appropriate focus of study. Mauss (1989, pp. 36-37, notes 1 and 2), however, implicitly rejects the entire issue raised by Woolgar and Pawluch by insisting that constructionists need not question the objective reality of "putative conditions" any more than that of any of the other elements in the construction of social problems but should focus only on the process by which those conditions come to be defined as problems. Yet if the issue of "ontological gerrymandering" has stalled the theoretical development of the constructionist perspective, as is feared by Ibarra and Kitsuse (1993), at least that argument has required of subjectivists more awareness of their own potential biases.

The position advocated by Mauss might be called “strict constructionist,” an example of which can be found in the work of Jennes and Broad (1997) on the anti-hate crime movement in the United States. In this work, the authors do not attempt to evaluate the accuracy or “objective” reality of the claims made by the women’s movement or gay/lesbian movement in the efforts of those movements to define violence as a social problem. In contrast, a somewhat nuanced approach, that might be called a “contextual constructionist approach” includes the evaluation of claims by interest groups as an important part of the analysis of the development of a social problem. For example, Best (1993), in his work on threats to children, relies on statistical and other data to assess the claims he describes. Calling a statement a “claim” does not discredit it, according to Best, who calls for social scientists to “worry a little less about how we know what we know, and worry a little more about what, if anything, we do know about the construction of social problems” (Best 1993, p. 144).

#### THE CURRENT INTELLECTUAL STATE OF AFFAIRS

In effect, two paradigms for understanding and explaining social problems have historically defined the sociological study of social problems. The traditional objectivist paradigm, which dominates textbooks and anchors policy-relevant research, evaluates certain social conditions as problems by definition. In this conceptualization, there is an implicit analogy to a cancer or another disease in the human body. That is, a social problem is understood as an objectively real and undesirable condition in society that should be diagnosed and treated by experts (especially sociologists). The more recently developed subjectivist paradigm, in contrast, denies that any social condition can be defined objectively as a social problem, for such a definition depends on the evaluation and claims-making activities of interest groups. Those groups organize and mobilize social movements in an attempt to gain general political acceptance of certain “undesirable” conditions as “social problems.”

Despite the growing prominence of the subjectivist approach in academic debates, at the level of social policy and undergraduate teaching the objectivist paradigm is likely to remain dominant.

One reason for this dominance in the public policy arena is the objectivist bias in the political culture of North America, in which public opinion has traditionally taken for granted the amenability of social ills to objective diagnosis and the responsibility of the political system to ameliorate those ills. As for undergraduate teaching, the lack of intellectual sophistication typically found at the freshman and sophomore levels of college education, where most courses in social problems are taught, makes teaching the subjectivist position difficult. The phenomenological epistemology of the subjectivist paradigm may be too abstract for most college students. Thus, the overwhelming majority of social problems textbooks continue to be organized around one or both versions of the traditional objectivist paradigm insofar as they assume that select social conditions—most frequently things such as crime, mental illness, poverty, environmental degradation, and suicide—are inherently social problems. It is rare to find a textbook that is organized around subjectivist concerns (but see Best 1995).

The emergent subjectivist paradigm is likely, however, to continue to provide the basis for the most novel and creative scholarly and professional literature on social problems. This is the case because its epistemology and methodology leave much room for development, in comparison with objectivism and positivism. Some scholars have even attempted to reconcile the objectivist and the subjectivist paradigms (e.g., Jones et al. 1989), but such attempts are most often made in textbooks, which are notorious for their sacrificing of theoretical focus in favor of eclecticism. Ultimately, the two paradigms probably cannot be reconciled, since they proceed from different epistemologies and study different topics. Objectivism focuses on certain social conditions, such as inequality and deviant behavior, taking for granted the evaluation of those conditions as problems. Subjectivism focuses on the *political players and processes* by which those conditions come to be defined as problems.

#### REFERENCES

- Agger, Ben 1993 “The Problem with Social Problems: From Social Constructionism to Critical Theory.” In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.

- Albert, Edward 1989 "AIDS and the Press: The Creation and Transformation of a Social Problem." In Joel Best, ed., *Images of Issues: Typifying Contemporary Social Problems*. Hawthorne, N.Y.: Aldine de Gruyter.
- Aronoff, Marilyn, and Valerie Gunter 1992 "Defining Disaster: Local Constructions for Recovery in the Aftermath of Chemical Contamination." *Social Problems* 39:345-365.
- Bash, Harry 1995 *Social Problems and Social Movements*. Atlantic Highlands, N.J.: Humanities Press.
- Becker, Howard 1966 *Social Problems: A Modern Approach*. New York: Wiley.
- Berger, Peter L., and Thomas Luckmann 1966 *The Social Construction of Reality*. New York: Doubleday.
- Best, Joel 1989 "Afterword." In Joel Best, ed., *Images of Issues*. New York: Aldine de Gruyter.
- 1990 *Threatened Children: Rhetoric and Concern about Child Victims*. Chicago: University of Chicago Press.
- 1993 "But Seriously Folks: The Limitations of the Strict Constructionist Interpretation of Social Problems." In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- 1995 *Images and Issues: Typifying Contemporary Social Problems*. New York: Aldine de Gruyter.
- Blumer, Herbert 1971 "Social Problems as Collective Behavior." *Social Problems* 18:298-306.
- Conrad, Peter 1975 "The Discovery of Hyperkinesis: Notes on the Medicalization of Deviant Behavior." *Social Problems* 23:12-21.
- Douglas, Mary L. 1966 *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. New York: Praeger.
- , and Aaron Wildavsky 1982 *Risk and Culture*. Berkeley: University of California Press.
- Eitzen, D. Stanley, and Maxine B. Zinn 1988 *Social Problems*, 4th ed. Boston: Allyn and Bacon.
- Fishman, Mark 1978 "Crime Waves as Ideology." *Social Problems* 25:531-543.
- Fuller, Richard C. 1937 "Sociological Theory and Social Problems." *Social Forces* 4:496-502.
- , and Richard R. Myers 1941a "Some Aspects of a Theory of Social Problems." *American Sociological Review* 6:24-32.
- , and Richard R. Myers 1941b "The Natural History of a Social Problem." *American Sociological Review* 6:320-328.
- Gerhards, Jürgen, and Dieter Rucht 1992 "Mesomobilization: Organizing and Framing in Two Protest Campaigns in West Germany." *American Journal of Sociology* 98:555-595.
- Gordon, Avery 1993 "Twenty-Two Theses on Social Constructionism: A Feminist Response to Ibarra and Kitsuse's 'Proposal for the Study of Social Problems.'" In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Gusfield, Joseph R. 1981 *The Culture of Public Problems: Drinking-Driving and the Symbolic Order*. Chicago: University of Chicago Press.
- Hilgartner, Stephen, and Charles L. Bosk 1988 "The Rise and Fall of Social Problems: A Public Arenas Model." *American Journal of Sociology* 94:53-78.
- Holstein, James A., and Gale Miller, eds. 1993 *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Ibarra, Peter R., and John I. Kitsuse 1993 "Vernacular Constituents of Moral Discourse: An Interactionist Proposal for the Study of Social Problems." In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Jenness, Valerie 1993 *Making It Work: The Prostitutes' Rights Movement in Perspective*. New York: Aldine de Gruyter.
- , and Kendal Broad 1997 *Hate Crimes: New Social Movements and the Politics of Violence*. New York: Aldine de Gruyter.
- Jones, Brian J., Joseph A. McFalls, Jr., and Bernard J. Gallagher III 1989 "Toward a Unified Model for Social Problems Theory." *Journal for the Theory of Social Behavior* 19:337-356.
- Kitsuse, John I., and Malcolm Spector 1973 "Toward a Sociology of Social Problems: Social Conditions, Value Judgments, and Social Problems." *Social Problems* 20:407-419.
- Klandermans, Bert 1992 "The Social Construction of Protest and Multiorganizational Fields." In Aldon D. Morris and Carol McBurn Miller, eds., *Frontiers in Social Movement Theory*. New Haven, Conn.: Yale University Press.
- Markle, Gerald E., and Ronald Troyer 1979 "Smoke Gets in Your Eyes: Cigarette Smoking as Deviant Behavior." *Social Problems* 26:611-625.
- Mauss, Armand L. 1989 "Beyond the Illusion of Social Problems Theory." In James A. Holstein and Gale Miller, eds., *Perspectives on Social Problems*, vol. 1. Greenwich, Conn.: JAI Press.
- , and associates 1975 *Social Problems as Social Movements*. Philadelphia: Lippincott.

- Michaelowski, Raymond 1993 "(De)Construction, Postmodernism, and Social Problems: Facts, Fiction, and Fantasies at the 'End of History.'" In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Miller, Gale 1993 "New Challenges to Social Constructionism: Alternative Perspectives on Social Problems Theory" In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter
- , and James A. Holstein 1989 "On the Sociology of Social Problems." In James A. Holstein and Gale Miller, eds., *Perspectives on Social Problems*, vol. 1 Greenwich, Conn.: JAI Press.
- 1993a *Constructionist Controversies: Issues in Social Problems Theory*. New York: Aldine de Gruyter.
- 1993b "Constructing Social Problems: Context and Legacy." In Gale Miller and James A. Holstein, eds., *Constructionist Controversies: Issues in Social Problems Theory*. New York: Aldine de Gruyter.
- 1993c "Social Constructionism and Its Critics: Assessing Recent Challenges." In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Mills, C. Wright 1943 "The Professional Ideology of Social Pathologists." *American Journal of Sociology* 49:165–180.
- Orcutt, James, and J. Blake Turner 1993 "Shocking Numbers and Graphic Accounts: Quantified Images of Drug Problems in the Print Media." *Social Problems* 40:190–206.
- Pfohl, Stephen 1993 "Revenge of the Parasites: Feeding Off the Ruins of Sociological (De)Construction." In James A. Holstein and Gale Miller, eds., *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: Aldine de Gruyter.
- Pong, Raymond W. 1989 "Social Problems as a Conflict Process." In James A. Holstein and Gale Miller, eds., *Perspectives on Social Problems*, vol. 1. Greenwich, Conn.: JAI Press.
- Richardson, James T., Joel Best, and David Bromley 1991 "Satanism as a Social Problem." In James T. Richardson, Joel Best, and David Bromley, eds., *The Satanism Scare*. Hawthorne, N.Y.: Aldine de Gruyter.
- Rittenhouse, C. Amanda 1992 "The Emergence of Premenstrual Syndrome as a Social Problem." *Social Problems* 38:412–425.
- Rose, V. M. 1977 "Rape as a Social Problem: A By-product of the Feminist Movement." *Social Problems* 25:75–89.
- Schneider, Joseph 1978 "Deviant Drinking as Disease: Alcoholism as Social Accomplishment." *Social Problems* 25:361–372.
- 1985 "Social Problems Theory: The Constructionist View." *Annual Review of Sociology* 11:209–229.
- Skolnick, Jerome H., and Elliott Currie, eds. 1985 *Crisis in American Institutions*, 6th ed. Boston: Little, Brown.
- Spector, Malcolm 1977 "Legitimizing Homosexuality." *Society* 14:20–24.
- , and John I. Kitsuse 1973 "Social Problems: A Reformulation." *Social Problems* 21:145–159.
- 1977 *Constructing Social Problems*. Menlo Park, Calif.: Cummings.
- Stallings, Robert A. 1990. "Media Discourse and the Social Construction of Risk." *Social Problems* 37:80–95.
- 1995 *Promoting Risk: Constructing the Earthquake Threat*. New York: Aldine de Gruyter.
- Troyer, Ronald J. 1989 "Are Social Problems and Social Movements the Same Thing?" In James A. Holstein and Gale Miller, eds., *Perspectives on Social Problems*. vol. A. Greenwich, Conn.: JAI Press.
- Turner, Ralph H. 1981 "Collective Behavior and Resource Mobilization as Approaches to Social Movements: Issues and Continuities." In Louis Kreisberg, ed., *Research in Social Movements, Conflicts, and Change*, vol. 4 Greenwich, Conn.: JAI Press.
- Waller, Willard 1936 "Social Problems and the Mores." *American Sociological Review* 1:922–933.
- Woolgar, Steve, and Dorothy Pawluch 1985 "Ontological Gerrymandering." *Social Problems* 32:214–227.
- Zald, Mayer N., and John D. McCarthy 1987 *Social Movements in an Organizational Society: Collected Essays*. New Brunswick, N.J.: Transaction.

ARMAND L. MAUSS  
VALERIE JENNESS

## SOCIAL PSYCHOLOGY

Social psychology is the study of individual behavior and psychological structures and processes as both outcomes of and influences on interpersonal relationships, the functioning of groups and other collective forms, and culturally define macrosocial structures and processes. Social psychologists vary in the theoretical orientations and methods they use, the conceptual distinctions they draw, and the substantive causal linkages they study. Much of the variability in these areas is accounted for by the

academic tradition in which a social psychologist has been trained.

Contemporary social psychology has intellectual roots in both psychology and sociology. Psychological social psychologists are guided by social learning theory as well as by orientations such as exchange and role theories. For the most part, their methods consist of laboratory and field experiments, and data analysis is accomplished with quantitative techniques. They discriminate between individual behavior and psychological structures and processes and interpersonal settings. The primary interest of psychological social psychology is the influence of the perceived social environment on individual cognitive, affective, and behavioral responses (Gilbert et al. 1998).

Contemporary sociological social psychology encompasses two major perspectives: symbolic interactionism and personality and social structure. Within symbolic interactionism, other distinctions are drawn according to the degree to which the proponents emphasize consistencies in human behavior as opposed to creative and emergent aspects of behavior, the influence of social structure in placing constraints on social interaction through which concepts of the self and others are formed, and the relative merits of qualitative and quantitative research methods. Considering these perspectives together, sociological social psychologists are influenced most frequently by symbolic interactionism, role theory, and exchange theory. They employ a range of research methods, including social surveys, unstructured interviews, observational techniques, and archival research methods; laboratory and field experiments also are used on occasion. Data analysis is accomplished with both qualitative and quantitative techniques. Distinctions are drawn between individual behaviors, psychological structures, groups and other interpersonal systems, and culturally defined macrosocial structures and processes. Sociological social psychologists focus on the reciprocal causal influences between individual psychological structures and macrosocial structures and processes or those between psychological processes and ongoing interpersonal systems (Cook et al. 1995; Michener and DeLamater 1994).

Implicit in the explanatory constructs sociologists use in investigating patterns of social behavior are individual-level psychological constructs.

The concept of culture, for example, frequently is defined in terms of shared normative expectations that are learned and transmitted in the course of social interaction. This definition implies subjective probabilities, evaluative judgments, and processes of symbolic communication through which normative expectations are transmitted and shared. Further, the substantive referents of culture relate to individual-level phenomena such as systems of values, beliefs, and perceptual orientations. In short, definitions of sociological explanatory constructs and the substantive referents of those constructs tend to be abstractions from individual-level psychological responses and systems, including those relating to cognition, affect, and goal orientation. A full understanding of most, if not all, sociological constructs depends on comprehension of the psychological responses and systems that the sociological constructs connote and from which the sociological concepts can be generalized.

The current state of social psychology is best understood through a description of the range of theoretical orientations and research methods used, the conceptual distinctions that are drawn, and the causal linkages that are investigated by representatives of the two social psychological traditions.

## THEORY AND METHOD

Social psychologists use a broad range of theoretical perspectives and research methods to study the reciprocal causal linkages between individual-level and social-level variables.

**Theoretical Perspectives.** Among the more frequently used theoretical perspectives are symbolic interactionism, role theory, exchange theory, and social learning theory.

*Symbolic interactionism.* From this perspective, people are perceived as acting toward others on the basis of the meaning those others and their behaviors have for the actors. Those meanings are derived and modified during social interaction in which people communicate with one another through the use of shared symbols. Symbolic interactionism encompasses the notion that people's ability to respond to themselves as objects permits them to communicate to themselves, through the use of symbols, the meanings that are given to people and objects by the persons who



perceive them. Thus, people interpret the world to themselves and respond according to that interpretation. The interpretation of a situation occurs in the course of ongoing social interaction. In short, persons become objects to themselves, interact with themselves, and interpret to themselves ongoing events and objects in the environment.

Proponents of symbolic interactionism vary in the extent to which they focus on the influence of a stable social structure on these processes. Those who deny the significance of a social structure concentrate on the process of cognitive interpretation and the creative construction of behavior that grows out of a person's interpretation of the ongoing interactive situation. Appropriate to this emphasis, empirical investigations employ observation and in-depth interviewing to the exclusion of experimental and quantitative, nonexperimental methods (the Chicago School). Derivatives of this approach to symbolic interactionism include the dramaturgical school, in which the metaphor of the theater is used to study how people create impressions of themselves during face-to-face interaction, and ethnomethodology, in which theoretical perspective students study the implicit rules governing interaction in particular situation to understand how people construct reality through social interaction.

For those who focus on the significance of social structures to symbolic interaction, the meanings of the behaviors in social interaction depend on the relevance of those behaviors for the social-identity-related standards by which people evaluate themselves. Individuals interact within a framework that defines the social identities of the interacting parties and the normative expectations that are applicable to each identity as it relates to the other identities in that situation. The behaviors that have the most meaning are relevant for highly placed standards in a person's hierarchy of values. The more a behavior of a person or the others with whom he or she is interacting validates or contradicts the social identity (male, father) that is important to that person, the more meaningful that behavior will be to him or her. To the extent that the behaviors of others toward a person signify evaluatively significant aspects of the self, it is important to anticipate responses from others. The others whose responses are more likely to signify evaluatively relevant information about the self are significant others (Charon 1998; Couse

1977; Kaplan 1986; Michener and DeLamater 1994; Stephan and Stephan 1990).

*Role theory.* From this perspective, human social behavior is viewed in the context of people playing roles (that is, conforming to normative expectations) that apply to people who occupy various social positions and interact with people in complementary social positions. As individuals change from one social position to another in the course of a day, they play different roles (as a father, for example, and then as an employer). The roles individuals play also change as they interact with people in different positions (a professor interacting with a colleague, with the dean, and with a student). As people shift roles, they also change the ways in which they view the world, the attitudes they hold toward different phenomena, and their behaviors. Although people identify more with some roles than with others, their ability to play their preferred roles is limited by the contradictory demands made on them by the other roles they are called on to play (Biddle 1986; Turner, 1990).

*Exchange theory.* This perspective is relevant to the investigation of the conditions under which individuals enter into and maintain stable relationships. One is most likely to do this when the rewards gained from the relationship are perceived as high, the costs are low, and the reward-cost differential is favorable compared with the perceived alternatives. Rewards (power, prestige, material goods) and costs (interpersonal hostility, great expenditures of money, long hours of work) are defined by personal values. Attraction to relationships is also a function of the extent to which the participants perceive each other as receiving outcomes (rewards) that are appropriate to their inputs (costs). In the absence of such equity, the participants adjust their behavior or way of thinking in an attempt to restore the fact or appearance of equity in a relationship (Molm and Cook 1995; Stephan and Stephan 1990).

*Social learning theory.* This orientation addresses how individuals learn new responses that are appropriate in various social situations. The primary processes through which social learning occurs include conditioning, by which one acquires new responses through reinforcement (that is, the association of rewards and punishments with particular behaviors), and imitation, by which one

observes the reinforcement elicited by another person's behavior (Bandura 1986; Taylor 1998).

**Methods.** Social psychological research employs a variety of methods, including social surveys, naturalistic observation, experiments, and analysis of archival data. Social surveys may be conducted by personal or telephone interviews or by self-administered questionnaires. For the most part, naturalistic observation involves observing ongoing activity in everyday settings (that is, field studies); in participant observation, the investigator plays an active role in the interaction. Experimental research involves the manipulation of independent variables to assess their effects on outcomes. Subjects are assigned at random to the independent conditions. The experiments may be conducted in the laboratory or in natural settings; in the latter case, the experimenter has less control over theoretically irrelevant variables but the experimental conditions are more realistic for the subjects. Archival research involves the use of existing data to test hypotheses. In some instances, the data can be used exactly as they appear, as with some statistical data. In other instances, such as newspaper stories, the data must be converted into another form, for example, for use in content analysis, which involves categorizing and counting particular occurrences (Cook et al. 1995; Gilbert et al. 1998; Michener and DeLamater 1994).

### CONCEPTUAL DISCRIMINATIONS

The pursuit of the goals of social psychology by scientists from psychological and sociological traditions has entailed the differentiation between concepts at the individual level and the social level.

**Individual-Level Concepts.** Social psychologists have focused on dynamic psychological structures, intrapsychic responses, and individual behaviors as outcomes of or influences on social structures and processes.

*Psychological structures.* At the individual level, psychological structures have been represented as dynamic organizations of dispositions to respond at the intrapsychic level or the behavioral level. More inclusive concepts, such as the personality, reflect the organization of psychological dispositions in terms of a structure of relatively stable cognitive, evaluative, affective, and behavioral tendencies. The concept of the person has been under-

stood in terms of a structure of predispositions to respond at the intrapsychic or behavioral level that are organized around a hierarchically related system of situationally defined social identities. The self has been treated as an inclusive structure of dispositions to respond reflexively at the cognitive, evaluative, affective, and behavioral levels. Less inclusive structures refer to organizations of particular psychological dispositions, such as personal value systems, treated as the hierarchy of situationally applicable criteria for self-evaluation; the structure of attitudes or generalized evaluative responses; and the system of concepts and schemas (structures of related concepts) a person uses to order stimuli (Kaplan 1986).

These structures are treated as components that are related to one another in a stable dynamic equilibrium and at the same time as having the potential to change. The structures of predispositions, when stimulated by internal or external cues, respond at the intrapsychic level or the behavioral level. The predispositions are inferred from the observed behaviors and self-reports of intrapsychic responses to recurrent stimuli in particular situations.

*Intrapsychic responses.* These are cognitive (including awareness and conceptual structuring), evaluative, and affective (or emotional) responses to contemporary stimuli, including one's own or others' behaviors in particular situational contexts. The current situation may stimulate one to attend to particular aspects of oneself, classify others in terms of group-membership concepts, attribute others' failures to external rather than internal causes, evaluate oneself as a failure, or experience attraction to other people. Intrapsychic responses are inferred from one's perceptible behaviors or self-reports of percepts, beliefs, attitudes, and feelings relating to the current situation.

*Behavior.* Individual behavior refers to the class of responses that are perceptible to others as well as to oneself. Behavior is distinguished from intrapsychic responses and the stable organization of dispositions to respond (person or personality) that are perceptible only to the self. Behavior includes purposive or unintended communications about oneself or others, helping and hurtful responses, affiliation and disaffiliation with other individuals or groups, conformity to or deviation from one's own or others' expectations, coopera-

tion and competition, positive and negative sanctioning of one's own or others' behaviors, and the myriad other perceptible responses one may make to oneself, others, or other aspects of one's environment (Kaplan 1986; Michener and DeLamater 1994; Stephan and Stephan 1990). Behavior is conceptualized as the outcome of socially influenced psychological structures and intrapsychic processes and as influencing social-level variables.

**Social-Level Concepts.** These concepts include interpersonal systems and culturally defined macrosocial structures.

*Interpersonal systems.* Interpersonal systems are defined as those in which two or more individuals interact with each other or otherwise influence each other over a brief or extended period. The interaction or mutual influence is governed by shared normative expectations that define appropriate behavior for individuals who occupy complementary or common social positions in the course of the interaction or mutual influence. The shared expectations may exist before participation in the interpersonal system and reflect the common culturally defined macrosocial structure or may be refined or emerge during the ongoing social interaction or mutual influence in response to the unique characteristics of the interacting individuals or other situational demands. The social positions a person occupies and the interpersonal systems in which a person participates as a consequence may be given at birth or may be adopted later in life according to stage in the life cycle and current situational demands. Interpersonal systems include interpersonal relationships, groups, and collective forms.

Interpersonal relationships are those in which two individuals have an ongoing interaction that is governed by their shared normative expectations. These expectations are derived from social definitions that delineate appropriate behavior for people occupying the social positions that characterize the individuals and emerge in the course of the ongoing social interaction. For example, a married couple's shared expectations depend on a common understanding of the obligations and rights of a husband and a wife in relation to each other, and the same is true of friends; in addition, in the course of social interaction, specific evaluative expectations regarding what each person in the relationship will and should do in various

circumstances develop. Individuals may interact with one another in the capacity of having the same status (such as group member or friend) or complementary statuses (such as husband and wife) or in the capacity of representing conflicting or cooperating groups. Relationships develop through predictable stages. Intimate relationships develop from the awareness of available partners, to contact with those who are thought to be desirable, to various stages of emotional involvement. The accompanying increases in emotional involvement represent increases in self-disclosure, trust, and mutual dependence (Berscheid and Reis 1998; Michener and DeLamater 1994).

A group consists of a number of individuals in ongoing interaction who share a set of normative expectations that govern the behavior of the members in relation to one another. Normative expectations may refer uniformly to all group members as they interact with one another and with nongroup members or to different individuals in their various social relational contexts. Individuals share an identity as members of a group as well as common goals; these goals may include the personal satisfaction gained from the intrinsically or instrumentally satisfying intragroup relationship or from a group identity that evokes favorable responses from extragroup systems. Group members may share norms from the outset and refine or change their expectations over time, or the norms may emerge in the course of member interaction. Groups include friendship networks, work groups, schools, families, voluntary associations, and other naturally occurring or purposively formed ad hoc associations. Groups vary in size, stability, the degree to which interaction among the members is regulated by preexisting role definitions, and complexity of role differentiation as well as the extent to which a group is embedded in more inclusive groups. Groups also vary according to whether the gratifications achieved from participation in a group are intrinsic to the social relationships and are diffuse as opposed to instrumental to the achievement of other ends and delimited.

Over the course of time, groups develop structures characterized by status hierarchies and functional role differentiation, or those structures may be predefined for new members. Status hierarchies reflect the values placed by group members on positions within the group. The individuals who occupy those positions are more or less es-

teemed depending on the valuation (status) of a position. Individuals who have higher-status positions and are consequently more highly esteemed ordinarily receive greater rewards (as these are defined by group members) and exercise greater influence over group decisions. In formal groups, functional differentiation is indicated by the formal role definitions associated with the various social positions that make up a group. In informal groups, over time some individuals come to be expected to perform certain functions, such as leading the group toward solving a problem (the task leader) or accepting responsibility for relieving tensions and maintaining group solidarity (the social-emotional leader) (Levine and Moreland 1998; Ridgeway and Walker 1995).

Collective forms include publics, audiences, crowds, and social movements. Collective forms are characterized by the mutual influence of individuals in responding cognitively, affectively, and behaviorally to a common focus. Individuals are undifferentiated according to social position: They share the social position defined by their common attention to an idea, person, object, or behavior. The common stimulus, previously learned dispositions to respond to that stimulus, and mutual influences through social contagion, social observation, and emergent norms that govern mood, action, and imagery lead to collective behaviors. Collective behaviors by large numbers of individuals who are not physically proximate in response to mass media and interpersonal stimulation include mass expressions of attitudes (public opinion), attraction (fads, fashions, crazes), and anxiety (panics). Crowd behaviors are collective responses by large numbers of physically proximate individuals who are influenced by social contagion, observation, and the resultant emergent norms. Social movements are expressions of dispositions to behave similarly with regard to a social issue (Michener and DeLamater 1994).

*Culturally defined macrosocial structures.* The inclusive sociocultural structure provides shared meanings and defines relationships among individuals depending on their social positions or identities in a situation. The social structure is made up of the stable relationships between social positions or identities that are culturally defined in terms of the rights and obligations people who occupy one position have in interacting with people who occupy another position. In the course of

the socialization process, individuals learn the rights and obligations that apply to those who occupy the various social positions, and those rights and obligations constitute the role that defines a social position. The inclusive social structure is a system consisting of components that are related to one another in a relatively stable dynamic equilibrium but may change over time as changes in structural positions and their role definitions become prevalent in interpersonal settings throughout the society. The culturally defined inclusive macrosocial structure encompasses systems of stratification, social differentiation according to race or ethnicity, and major social institutions as well as other consensually defined social structures (Kerckhoff 1995).

### CAUSAL RELATIONSHIPS

Within a social psychological framework, a person's psychological structure, intrapsychic responses, and behaviors are viewed in terms of the profound influence exerted on that person by his or her past and continuing participation in interpersonal and social systems. In turn, the person behaves in ways that have consequences for the interpersonal systems and social structures in which he or she participates. Implicit in this framework is a general causal model. Social structural arrangements define systems of shared meanings that in turn define the role expectations that govern behavior in interpersonal systems. A person is born into functioning interpersonal systems and throughout the life cycle participates in other interpersonal systems that together reflect culturally defined macrosocial structures and processes. In the course of a person's life in the context of dynamically evolving interlocking interpersonal systems, biogenetically given capabilities are actualized; the person learns to view the world through a system of concepts, internalizes needs, symbolizes those needs as values, accepts social identities, and develops emotional cognitive and behavioral dispositions to respond. These relatively stable psychological structures are stimulated by contemporary social situations that have symbolic significance for the individual and thus evoke predictable personal responses. Over time, the same social situations stimulate personal change.

The development of language skills, along with a person's experiences as the object of others'

responses to him or her in the course of the socialization process, influences the development of a person's tendencies to become aware of, conceive of, evaluate, and have feelings about herself or himself as well as dispositions to behave in ways that are motivated by the need to protect or enhance the self. The nature of a person's responses to herself or himself are influenced by past and present social experiences. Those responses in turn influence the relationships and groups in which a person participates and indirectly influence the more inclusive social system, thus intervening between social influences on the person and her or his influence on interpersonal systems and the culturally defined social structure (Corsaro and Eder 1995; Fiske et al. 1998; Elder and O'Rand 1995; Kaplan 1995; Kerckhoff 1995; Krauss and Chiu 1998; Maynard and Whalen 1995; Miller-Loessi 1995).

The substantive concerns of social psychological theory and research reflect detailed consideration of these general processes. These concerns address (1) the influence of culturally defined macrosocial structures and processes or interpersonal systems on psychological structures, intrapsychic responses, and individual behaviors or (2) the influence of psychological structures, intrapsychic responses, and individual behaviors on interpersonal systems and culturally defined macrosocial structures and processes.

#### **Social Influences on Psychological Structures.**

Substantive concerns with social influences on individual psychological structures, intrapsychic responses, and behaviors have focused on long-term social structural influences through socialization processes and contemporary interpersonal influences in interpersonal settings.

*Social structural effects.* Social structural arrangements define the content, effectiveness, and style of the socialization experience and thus influence a person's psychological structures. Individuals occupy social positions by being born into them or achieving them later in life. Each social position is defined in terms of role expectations that specify appropriate behavior for people who occupy that position in the context of particular relationships. As a result of occupying positions, people become part of interpersonal systems that consist of themselves and those who occupy complementary positions. In these relationships and groups people

become socialized. Socialization is the lifelong process through which an individual learns and becomes motivated to conform to the norms defining the social roles that are played or might be played in the future that individual and those with whom she or he interacts. Socialization occurs in a variety of social contexts, including the family, school, play groups, and work groups, through the experience of rewards and punishment consequent to performing behaviors, observation of the consequences of behaviors for others, direct and intended instruction by others, and self-reinforcement. The acquisition of language skills permits one to be rewarded and punished through the use of symbolic responses, communicate with others about the appropriateness of different responses, and reinforce responses through the process of becoming an object to oneself and disapproving or approving of one's past or anticipated behaviors. The cognitive structures used in coding and processing information about one's own behavior and the hierarchy of self-evaluative criteria also are learned in the course of socialization.

The content of role definitions and the centrality of particular roles for a person's identity structure depend on stage in the life cycle, role definitions associated with other social positions, and the historical era. The roles that are most central to a person's identity and contribute most to self-esteem depend on that person's position in the social structure, including age and gender. During a particular historical period, for example, men may base their self-esteem more on success in the occupational sphere whereas women in the same stage of life base theirs on adequate performance of family roles.

The effectiveness of the socialization process is influenced by more or less invariant developmental stages of cognitive and emotional development in interaction with the varying demands made on the individual at various stages in life as well as by discrepancies between the demands made on a person and the resources that would permit her or him to meet those demands (Corsaro and Eder 1995; Elder and O'Rand 1995; Miller-Loessi 1995).

The social structure affects the style of the socialization process as well. Higher-socioeconomic-status parents are more likely than are lower-class parents to base rewards and punishment on a

child's intentions than on actual behavior and to rely on reasoning and the induction of shame and guilt rather than physical punishment. As a family becomes larger, parents are more likely to exercise autocratic parenting styles, while children elicit less attention from parents and develop more independence (Michener and DeLamater 1994).

The end result of the social-structure-influenced socialization process is the development of psychological structures that are stimulated by social-identity-related situations or are evoked more generally in the course of social interaction. Depending on whether persons are born into male or female social positions, they develop different achievement orientations and evaluate themselves in accordance with the success their in approximating the standards of achievement they set for themselves. Individuals in higher socioeconomic classes tend to value a sense of accomplishment and family security more highly than do those with lower socioeconomic status, who tend to put more emphasis on a comfortable life and hope of salvation. More specifically, individuals who are born into a higher social class are more likely to be socialized to value educational achievement and aspire to higher levels of education. Those who achieve at higher levels in school are more likely to interact with others who respond to them in ways that reinforce academically oriented self-images and values that reflect an achievement orientation. Individuals whose occupational status involves self-direction tend to develop a high valuation of responsibility, curiosity, and good sense, while those in occupational positions characterized by close supervision, routine activities, and low levels of complexity in work tasks tend to develop a high valuation of conformity (Heiss 1981; House 1981; Kohn et al. 1983; Rokeach 1973; Sewell and Hauser 1980).

In general, in the course of socialization people become disposed to identify others in their environment, anticipate their responses, imagine aspects of themselves as eliciting those responses, behave in ways calculated to elicit those responses, and value the responses of others and the aspects of the self that elicit those responses. Radical resocialization, by which an individual unlearns lifelong patterns and learns new attitudes, values, and behaviors, may occur in circumstances in which the agents of socialization have uniform and

total control over the individual's outcomes, as in some psychiatric hospitals, penal institutions, traditional military academies, and prisoner-of-war camps (Goffman 1961).

*Interpersonal effects.* The contemporary interpersonal context stimulates self-conceptions and self-evaluative, affective, and behavioral responses. Each social situation provides participants with physical cues that allow them to make inferences about the social identities of the other participants, the role expectations each person holds of the others, and the perceived causes of the behaviors of the interacting parties. These conceptions regarding the situated identities are in part responses to the demand characteristics of the situation and in part the outcomes of the need of one party to project a particular social identity on the other person so that the first party can play a desired complementary role (Alexander and Wiley 1981). The situational context provides symbolic cues that specify the relevance of particular traits, behaviors, or experiences for one's current situation; it also provides a basis for comparing one's characteristics with those of other people.

The current social situation defines the relevance of some self-evaluative standards rather than others. The presence of other people, cameras, or mirrors makes people more self-aware and thus stimulates their disposition to evaluate themselves. Certain responses of others (sanctions), in addition to constituting intrinsically value-relevant responses, communicate to persons the degree to which they have approximated self-evaluative standards. In the early stages of development of a group, individuals may be assigned higher-status positions on the basis of status characteristics (such as those relating to age, sex, ethnicity, and physical attractiveness) that have evaluative significance in the more inclusive society. Although these characteristics may not be relevant to the ability to perform the functions for which the group exists, the high valuation of these status characteristics may lead to the assignment of individuals to high-status positions in the group through a process of status generalization (Berger et al. 1989; Ridgeway and Walker 1995). As a consequence of culturally defined preconceptions regarding the merits of various status characteristics, persons with those characteristics are expected to perform better on a group task.

With regard to affective responses, social stimuli evoke physiological reactions that are labeled as specific emotional states, depending on the cues provided by social circumstances (Kelley and Michela 1980). In turn, individuals who label an experience as a particular emotional state selectively perceive bodily sensations as cues that validate that experience (Leventhal 1980; Pennebaker 1980). Social stimuli that evoke psychological distress include contexts in which a person is unable to fulfill role requirements because of the absence of personal and interpersonal resources and the presence of situational barriers to fulfilling the obligations associated with the social positions that person occupies. Other such stimuli are represented by intrinsically distressing aspects of social positions. People may be distressed not only because they cannot do their jobs well but also because of the absence of meaning that their jobs have for them and because of other noxious circumstances correlated with the position (time pressures, noise, lack of autonomy, conflicting expectations) (Kaplan 1996).

Situational contexts define expectations regarding appropriate and otherwise attractive behavior and thus stimulate behavior that anticipates fulfillment of the expectations and achievement of the goals (including avoidance of noxious states such as social rejection). In collective forms of interpersonal systems, emergent norms govern actions as well as moods and imagery for publics and crowds and so lead to mass behaviors such as crazes and panics and crowd behaviors such as rioting. The motivation for participation in social movements is influenced by expectations regarding the value and likelihood of the success of a movement (Michener and DeLamater 1994).

In interpersonal relationships as well, shared expectations govern attraction to others, helping behavior, and aggressive behavior. An individual tends to be attracted to those with whom interaction is facilitated, those who are characterized by socially appropriate and desirable traits (including physical attributes), those who share tastes with and are otherwise like that individual, those who manifest liking for him or her, and in general those who may be expected to occasion rewarding outcomes. Helping behavior is evoked by situational demand characteristics, such as role definitions that define helping behavior as appropriate for people who occupy particular social positions, or

by interpersonal expectations that helping behavior by the other person should be reciprocated. The likelihood of conforming to these situational demand characteristics increases when a person perceives that the rewards for doing so will be forthcoming (including personal satisfaction in helping others, a sense of fulfillment in doing what one is called on to do, and approval by others) and that failure to do so will bring negative sanctions (social disapproval, a self-evaluation of having failed to do the right thing). Conformity to demand characteristics that require helping behavior may be impeded if a person perceives that it would involve costs, such as hindering the achievement of other goals. The awareness of potential rewards or costs for engaging or failing to engage in helping behavior is facilitated by situational characteristics such as the presence of observers and circumstances that produce self-awareness. The need to help others is increased by experiences that evoke negative self-evaluations. The resulting negative self-feelings motivate helping behavior as a way of improving one's self-evaluation (Michener and DeLamater 1994).

Aggressive behavior may arise in response to situational demand characteristics such as perceiving oneself as playing a role that requires aggressive behavior either as a response to intentional aggressive behavior directed toward one by others or simply as a communication of an aggressive stance. Reinforcement by rewards increases the frequency or continuity of aggressive patterns. Rewarding outcomes of aggression include the related rewards of social approval, an improved position in the prestige hierarchy of the group, self-approval, and material gain. Individuals are inhibited from engaging in aggressive behavior when they perceive the act as contrary to normatively proscribed roles or otherwise anticipate adverse consequences of the behavior. These inhibiting effects may be obviated by the reduced self-awareness that results from being part of a crowd, for example, or by the administration of psychotropic drugs (Bandura 1973; Baron 1977; Kaplan 1972; Singer 1971).

In group contexts, role definitions and influences on self-awareness affect individual behavior. The assignment of individuals to higher-status positions in a group and concomitant expectations of higher levels of performance or of the adoption of particular functional roles frequently motivate peo-

ple to conform to those expectations or lead to the provision of resources that permit them to do so (Berger et al. 1989). When socially induced self-awareness causes people to attend to public aspects of themselves, they tend to be responsive to group influences. When self-awareness causes them to attend to their personal standards, they tend to direct their behavior to conform with those values even when they conflict with group standards. Thus, exposure to cameras induces public self-awareness and increases social conformity, while exposure to mirrors evokes private self-awareness and an increase in self-direction (Scheier and Carver 1983).

The effects of social stimuli on deviant, as opposed to conforming, behavior have been addressed from a variety of theoretical frameworks, including structured strain theory, differential association and deviant subculture theories, control theory, self-theory, and the labeling perspective. Attempts to integrate or elaborate any of these approaches encompass the following ideas (Gibbs 1981; Hollander 1975; Kaplan 1984; Messner et al. 1989; Moscovici 1985). First, individuals who experience rejection and failure in conventional social groups lose their motivation to conform to conventional norms and are motivated to deviate from those norms. At the same time, these individuals are disposed to seek alternative deviant patterns to attain or restore feelings of self-worth. Second, individuals who participate in groups that endorse behaviors that are defined as deviant in other groups (whether because they seek alternative deviant patterns through which they can improve self-worth or because of long-term identification with a deviant subculture) positively value the "deviant" patterns and are provided with opportunities and the resources to engage in the deviant behavior. Third, individuals with the motivation and opportunity to engage in deviant behavior are deterred by anticipated negative responses from groups that define that behavior as deviant and to which they remain emotionally bonded. Individuals tend to conform to the normative expectations of a group to the extent that they are made aware of the deviant nature of their behavior or attributes, are attracted to the group and therefore are highly vulnerable to the sanctions the group may administer for deviant behavior, are prevented from leaving the group and so freeing themselves from vulnerability to the group's

negative sanctions, identify with the group and thus adopt its normative standards, and internalize the normative standards and regard conformity as intrinsically valuable. Fourth, individuals who evoke negative social sanctions in response to initial deviance continue or increase the level of deviant behavior as a result of the effects of the negative social sanctions on increased alienation from the conventional group, increased association with deviant peers, and increased motivation to justify the initial deviance by more highly evaluating a deviant act. Continuity or escalation of deviant behavior also is likely to occur if motives that ordinarily inhibit the performance of deviant acts are weakened and if a person perceives an association between the deviant behavior and satisfaction of preexisting needs (including the need to enhance one's self-esteem).

#### **Psychological Influences on Social Systems.**

The consequences of socially influenced psychological processes may be observed at the interpersonal level and at the more inclusive, culturally defined macrosocial-structure level (Kaplan 1986).

*Interpersonal systems.* Intrapsychic responses and behaviors influence interpersonal systems in a wide variety of ways. Among the more salient consequences are those relating to the stability and functioning of groups, intragroup influences, group membership, and intergroup relationships.

Individuals affect both the *stability* and the *functioning* of their groups through their behavior. The stability of a group is enhanced to the extent that individuals conform to the expectations other people have of them and thus validate those expectations. An individual contributes to group functioning by playing the roles other people expect her or him to play in the group, permitting others to play their complementary roles. Conformity is influenced by the need for self-approval and the approval of others when the criteria for approval are the group standards. If personal and group standards reflect the value of scholarship, individuals may study hard; if approximation to the standards of a particular social identity (such as male) is a salient basis for self-evaluation, people strive to conform to what they perceive as the role expectations associated with that position. More generally, persons may evaluate themselves in terms of conforming to others' expectations. A salient value may be to evoke approving responses from



others. To that end, a person may behave in number of ways, including conforming to others' expectations and presenting oneself to others in ways calculated to evoke approving responses. However, a person will strive to conform to group standards in order to approximate self-values only to the extent that success or failure is attributed to the degree of personal effort rather than to circumstances (Kaplan 1995). Conformity is also an outcome of the need for others' approval. In a group in which the members are highly attracted to the group, conformity to group norms, including those related to productivity, is high. In such cohesive groups, members have greater power over one another than they do in groups where the individuals are less attracted to and dependent on the group (Cartwright and Zander 1968; Cialdini and Frost 1998; Hare 1976). The need for others' approval is reflected in the use of disclaimers and excuses to mitigate others' responses to personal behaviors (Hewitt and Stokes 1975; Karp and Ybels 1986; Spencer 1987). A perceived threat to the group increases members' attraction to the group and conformity to shared norms while decreasing tolerance of deviance.

Interpersonal influence occurs through the use of both overt and covert behaviors. Overt methods of persuasion include the use of information or arguments, the offering of rewards, and the threat of punishment. Covert attempts to influence others are reflected in self-presentation in order to create the impression of oneself as likable or in other ways to manipulate the impression others have of one. Attempts at persuasion are more or less effective depending on the characteristics of the source of the communication, the message itself, and the target of the communication. For example, communications are more persuasive if they come from a number of independent sources each of which is perceived to be expert, trustworthy, or otherwise attractive to the target of the communication, than they are if they occur in mutually exclusive circumstances. The effectiveness of threats and promises in influencing others depends on the magnitude and certainty of the proffered rewards and punishments. When the parties involved in the influence process all have the capacity to reward or punish one another, changes in opinions or behaviors are influenced by bargaining and negotiation processes. Among the possible outcomes, depending

on a number of circumstances, are mutual influence, escalation of conflict, accommodation of one person to the demands of the other, and failure of the parties to agree (Michener and DeLamater 1994). In the course of group interaction, individuals develop more extreme attitudes than they held as individuals. This may be due to the pooling of arguments, which adds new reasons for the initially held attitude, or to the social support provided by other group members, which permits the person to be more extreme in his or her opinions with less fear of group rejection (Brandstatter et al. 1982).

Persons are motivated to present themselves in ways that evoke desired responses from others. This is accomplished through a variety of tactics. A significant feature of self-presentation is an individual's social identity in a situation. By projecting a particular identity, the individual effectively imposes complementary identities on others; if the other people perform the roles associated with those identities, they in effect endorse the identity that the individual wishes to project. This imposition of social identities on others (altercasting) has the desirable effect of affirming the social identity the individual wishes to project. For example, by complying with a reason's demands or following her or his lead, the others affirm that person's position of authority or leadership. In addition, people's favorable responses are intrinsically valued, other rewarding outcomes are contingent on them, and they indicate to a person that her or his public image reflects her or his personal ideals. Self-presentation may be used to create false as well as true images of oneself. The creation of false images (impression management) also is used to evoke responses from others that serve one's personal needs. Tactics involving the false presentation of self include pretenses that one admires other individuals or share their opinions and presenting oneself as if one had admirable qualities that one does not in fact possess (Baumeister 1982; Tedeschi 1981).

The attraction and maintenance of group membership are a function of the perception by members that group participation is intrinsically desirable or instrumental to the achievement of shared or individually defined goals (Evans and Jarvis 1980). Relationships, as well as larger groups, grow and become resistant to dissolution as the partners become increasingly dependent on each other

for need satisfaction, which may lie in the relationship itself or in the role the partner plays in facilitating the satisfaction of other needs outside the relationship (that is, by providing social support). Primary relationships dissolve to the extent that the costs come to exceed the rewards—whether in absolute terms or relative to the cost-benefit ratio that may be obtained from alternative relationships—and to the extent that the costs of remaining in the relationship outweigh the costs (including social disapproval) associated with terminating it (Cialdini and Frost 1998; Kelley and Thibaut 1978; Kerckhoff 1974). Among the costs are perceptions of inequity. Group members tend to compare the relationship between their own contributions to the group and the rewards they receive with other members' contributions and rewards. Judgments of inequity are made when members perceive rewards to be out of proportion to contributions. Judgments that inequitable states exist stimulate responses to reduce the inequity or at least the perception of the inequity. The inability to redress or tolerate inequitable relationships may lead to eschewing membership in a group (Walster et al. 1978). In general, people select group memberships, when they have a choice, and maintain them in accordance with their value in facilitating self-approving responses. People maintain relationships by whose standards they may evaluate themselves positively and tend not to associate with groups by whose standards they would be considered failures (Kaplan 1986).

The nature of the responses *that groups evoke from nonmembers* is influenced by the nonmembers' perceptions, evaluations, and feelings toward themselves and others. Negative emotions, such as anger, and consequent aggressive behavior may be directed toward groups when individuals' interests cannot be served except at the cost of frustration of the objectives of another group or when individuals associate the other group with past experiences of failure. Among the benefits persons may experience at the cost of the other group's outcomes is increased self-esteem. Aggressive behavior directed toward others deflects anger that might have been directed toward oneself. When the basis of one's feelings of accomplishment are judged relative to the achievements of another group, aggressive behaviors that lead to the failure or destruction of the other group enhance feelings of pride in one's own group. Stronger

levels of identification with one's group or social category increase the need to enhance one's group identity at the cost of adverse outcomes for other groups (Bobo 1983; LeVine and Campbell 1972; Worchel and Austin 1986). The tendency to devalue others as they deviate from one's own group's standards increases the justification for negative attitudes and hostile actions toward the other group. The need to justify aggressive attitudes toward another group also frequently leads to biased perceptions that reinforce or validate preexisting attitudes toward that group. Reversal of the process is impeded by the decreases in communication that accompany negative attitudes toward that group. Frequent experiences of observing aggressive responses desensitize a person to the effects of these responses and establish a normative judgment that they are within the expected range of responses.

Other individuals or groups may be the objects of helping behavior, depending on the actor's intrapsychic responses. Negative affect (particularly negative self-feelings) decreases helping behavior by focusing attention inward and away from the plight of other individuals. Thus, some individuals are less likely to empathize with or even be aware of others' needs. At the same time, distressful self-feelings motivate an individual to behave in ways that will earn self-approval. Helping behavior may serve this function by fulfilling others' expectation that helping behavior be offered, conforming to role definitions of helping behavior as appropriate for particular social identities, and conforming to self-values regarding altruistic behavior and thus compensating for feelings of rejection and failure (Dovidio 1984).

*Macrosocial structures*. Psychological structures, intrapsychic responses, and behaviors influence the substance of the social structure at any given time and social change over time. Dimensions of personality that reflect evaluative standards affect the positions an individual has in the social structure. High value placed on educational attainment and achievement orientation lead ultimately to educational achievement and high occupational status. Similarly, studies of the relationship between the occupational structure and personality suggest that workers may be selected into jobs because of the fit between their personality characteristics and the requirements of the work situation (Kerckhoff 1989). Individuals who value self-

direction select occupations that permit the exercise of self-direction, that is, ones that involve less routine, more complex tasks, and low levels of supervision. Persons who place a high value on conformity tend to opt for occupations that are closely supervised, routine, and noncomplex.

The effects of persons as products of past socialization experiences and as stimulated by contemporary social situations on interpersonal social systems and the more inclusive social structure are mediated by the responses of those persons to themselves. An individual influences the current and future functioning of interpersonal systems by becoming self-aware and conceiving of the self in particular ways, evaluating the self as more or less closely approximating personal standards, and experiencing self-feelings that stimulate self-protective and self-enhancing responses, some of which directly and indirectly affect the functioning of the interpersonal or social systems in which that individual participates.

If a person fails to behave in ways that meet self-imposed demands, that person will experience negative self-feelings that motivate him or her to behave in ways that will reduce the self-rejecting feelings. If the person identifies the self-rejecting experiences with particular social identities, she or he may reject the group and define it as a negative reference group, overidentify with the group and reevaluate formerly denigrated attributes as desirable ones, or project undesirable characteristics onto other groups or social categories and act with hostility toward them. Negative self-feelings also may lead to reduced levels of socioeconomic aspirations, occupational change, withdrawal from political participation or association with political activism, and changes in patterns of religious affiliation and participation (Kaplan 1986; Rosenberg and Kaplan 1982). If the circumstances that hinder a person from behaving in ways that earn self-approval and the self-protective responses they stimulate are widespread, the inclusive social structure will be affected. The person's responses directly affect interpersonal systems, that is, individuals who interact in the context of social relationships and groups that are governed by shared situation-specific, identity-specific, or person-specific expectations. If the individual is motivated to withdraw from or otherwise disrupt the functioning of the interpersonal systems in which she or he participates, the function-

ing of the other individuals will be similarly disrupted, since others' performance is contingent on the individual's conformity to their expectations. However, the functioning of the interpersonal system will be facilitated if the individual is motivated to conform to the normative expectations that the participants in the interaction situation view as applicable to the person in that particular situational context. If the disposition to deviate from normative expectations is prevalent, disruptions of social relationships will be widespread and the social structure will be less resistant to changes in patterns of response over time. While widespread conformity to shared expectations in particular social contexts has stabilizing influences on the broader social structure, widespread innovation or deviation from them influences the development of new social structural arrangements and definitions.

(SEE ALSO: *Affect Control Theory and Impression Formation; Aggression; Attitudes; Attribution Theory; Behaviorism; Cognitive Consistency Theories; Collective Behavior; Decision-Making Theory and Research; Extreme Influence; Field Theory; Game Theory and Strategic Interaction; Identity Theory; Intelligence; Interpersonal Attraction; Personality and Social Structure; Personality Theory; Persuasion; Prejudice; Role Theory; Self-Concept; Small Groups; Social Perception; Socialization; Symbolic Interaction Theory*)

## REFERENCES

- Alexander, C. Norman, Jr., and Mary Glenn Wiley 1981 "Situating Activity and Identity Formation." In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*, New York: Basic Books.
- Bandura, Albert 1973 *Aggression: A Social Learning Analysis*. Englewood Cliffs, N.J.: Prentice-Hall.
- . 1986. *The Social Foundations of Thought and Action*. Englewood Cliffs, N.J.: Prentice-Hall.
- Baron, Robert A. 1977 *Human Aggression*. New York: Plenum.
- Baumeister, Roy F. 1982 "A Self-Presentational View of Social Phenomena." *Psychological Bulletin* 91:3–26.
- Berger, Joseph, Morris Zelditch, Jr., and Bo Anderson, eds. 1989 *Sociological Theories in Progress*. Newbury Park, Calif.: Sage.
- Berscheid, Ellen, and Harry T. Reis 1998 "Attraction and Close Relationships." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed. New York: McGraw-Hill.

- Biddle, Bruce J. 1986 "Recent Developments in Role Theory." In A. Inkeles, J. Coleman, and N. Smelser, eds., *Annual Review of Sociology*, vol. 12. Palo Alto, Calif.: Annual Reviews.
- Bobo, Lawrence 1983 "Whites' Opposition to Busing: Symbolic Racism or Realistic Group Conflict?" *Journal of Personality and Social Psychology* 45:1196-1210.
- Brandstatter, Hermann, James H. Davis, and Gisela Stocker-Kreichgauer, eds. 1982 *Group Decision Making*. New York: Academic Press.
- Cartwright, Dorwin, and Alvin Zander 1968 *Group Dynamics*. New York: Harper & Row.
- Charon, Joel M. 1998 *Symbolic Interactionism: An Introduction, an Interpretation, an Integration*, 6th ed. Upper Saddle River, N.J.: Prentice Hall.
- Cialdini, Robert B., and Melanie R. Frost 1998 "Social Influence: Social Norms, Conformity, and Compliance." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed. New York: McGraw-Hill.
- Cook, Karen S., Gary Alan Fine, and James S. House, eds. 1995 *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Corsaro, William A., and Donna Eder 1995 "Development and Socialization of Children and Adolescents." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Dovidio, J. H. 1984 "Helping Behavior and Altruism: An Empirical and Conceptual Overview." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*, vol. 17. New York: Academic Press.
- Elder, Glen H. Jr., and Angela M. O'Rand 1995 "Adult Lives in a Changing Society." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*, Needham Heights, Mass.: Allyn and Bacon.
- Evans, N. J., and Patricia A. Jarvis 1980 "Group Cohesion, a Review and Reevaluation." *Small Group Behavior* 11:359-370.
- Fiske, Alan P., Shinobu Kitayama, Hazel R. Markus, and Richard E. Nisbett 1998 "The Cultural Matrix of Social Psychology." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed. New York: McGraw-Hill.
- Gibbs, Jack P. 1981 *Norms, Deviance, and Social Control*. New York: Elsevier.
- Gilbert, Daniel T., Susan T. Fiske, and Gardner Lindzey, eds. 1998 *The Handbook of Social Psychology*, vols. 1 and 2, 4th ed. New York: McGraw-Hill.
- Goffman, Erving 1961 *Asylums*. Garden City, N.Y.: Anchor.
- Hare, A. Paul 1976 *Handbook of Small Group Research*. New York: Free Press.
- Heiss, Jerold 1981 "Social Roles." In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Hewitt, John P., and R. Stokes 1975 "Disclaimers." *American Sociological Review* 40:1-11.
- Hollander, Edwin P. 1975 "Independence, Conformity and Civil Liberties: Some Implications from Social Psychological Research." *Journal of Social Issues* 31:55-67.
- 1981. "Social Structure and Personality." In M. Rosenberg and R. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- Kaplan, Howard B. 1972 "Toward a General Theory of Psychosocial Deviance: The Case of Aggressive Behavior." *Social Science and Medicine* 6:593-617.
- 1984 *Patterns of Juvenile Delinquency*. Beverly Hills, Calif.: Sage.
- 1986 *Social Psychology of Self-Referent Behavior*. New York: Plenum Press.
- 1995 "Drugs, Crime, and Other Deviant Adaptations." In H. B. Kaplan, eds., *Drugs, Crime, and Other Deviant Adaptations: Longitudinal Studies*. New York: Plenum Press.
- ed. 1996 *Psychosocial Stress: Trends in Theory and Research*. New York: Academic Press.
- Karp, David Allen, and William C. Ybels 1986 *Sociology and Everyday Life*. Itasca, Ill.: Peacock.
- Kelley, Harold H., and J. Michela 1980 "Attribution Theory and Research." *Annual Review of Psychology* 31:457-501.
- , and John W. Thibaut 1978 *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.
- Kerckhoff, Alan C. 1974 "The Social Context of Interpersonal Attraction." In T. Huston, ed., *Foundations of Interpersonal Attraction*. New York: Academic Press.
- 1989 "On the Social Psychology of Social Mobility Processes." *Social Forces* 68:17-25.
- 1995 "Social Stratification and Mobility Processes: Interaction between Individuals and Social Structures." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Kohn, Melvin L., and Carmi Schooler, with the collaboration of Joanne Miller, K. Miller, S. Schoenbach, and R. Schoenberg 1983 *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, N.J.: Ablex.

- Krauss, Robert M., and Chi-Yue Chiu 1998 "Language and Social Behavior." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed. New York: The McGraw-Hill.
- Leventhal, Howard 1980 "Toward a Comprehensive Theory of Emotion." In L. Berkowitz, ed., *Advances in Experimental Social Psychology*. vol. 13. New York: Academic Press.
- Levine, John M. and Richard L. Moreland 1998 "Small Groups." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*, vol. 2, 4th ed. New York: McGraw-Hill.
- Levine, Robert A., and Donald T. Campbell 1972 *Ethnocentrism: Theories of Conflict, Ethnic Attitudes and Group Behavior*. New York: Wiley.
- Maynard, Douglas W., and Marilyn R. Whalen 1995 "Language, Action, and Social Interaction." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Messner, Steven F., Marvin D. Krohn, and Allen E. Liska 1989 *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*. Albany: State University of New York Press.
- Michener, H. Andrew and John D. DeLamater 1994 *Social Psychology*, 3rd ed. New York: Harcourt Brace College Publishers.
- Miller-Loessi, Karen 1995 "Comparative Social Psychology." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Molm, Linda D., and Karen S. Cook 1995 "Social Exchange and Exchange Networks." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Moscovici, Serge 1985 "Social Influence and Conformity." In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, vol. 2, 3rd ed. Reading, Mass.: Addison-Wesley.
- Pennebaker, James W. 1980 "Self-Perception of Emotion and Internal Sensation." In D. W. Wegner and R. R. Vallacher, eds., *The Self in Social Psychology*. New York: Oxford University Press.
- Ridgeway, Cecilia L. and Henry A. Walker 1995 "Status Structures." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Needham Heights, Mass.: Allyn and Bacon.
- Rokeach, Milton 1973 *The Nature of Human Values*. New York: Free Press.
- Rosenberg, Morris, and Howard B. Kaplan, eds. 1982 *The Social Psychology of the Self-Concept*. Arlington Heights, Ill.: Harlan Davidson.
- Scheier, Michael F., and Charles Carver 1983 "Two Sides of the Self: One for You and One for Me." In J. Suls and A. G. Greenwald, eds., *Psychological Perspectives on the Self*, vol. 2. Hillsdale, N.J.: Erlbaum.
- Sewell, William H., and Robert M. Hauser 1980 "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." *Research in Sociology of Education and Socialization* 1:59-99.
- Singer, Jerome L., ed. 1971 *The Control of Aggression and Violence: Cognitive and Physiological Factors*. New York: Academic Press.
- Spencer, J. William 1987 "Self-Work in Social Interaction: Negotiating Role-Identities." *Social Psychology Quarterly* 50:131-142.
- Stephan, Cookie White, and Walter G. Stephan 1990 *Two Social Psychologies*, 2nd ed. Belmont, Calif.: Wadsworth.
- Taylor, Shelley E. 1998. "The Social Being in Social Psychology." In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*. vol. 1, 4th ed. New York: McGraw-Hill.
- Tedeschi, James T. ed. 1981 *Impression Management Theory and Social Psychological Research*. New York: Academic Press.
- Turner, Ralph H. 1990 "Role Change." In W. R. Scott and J. Blake, eds., *Annual Review of Sociology*, vol. 16. Palo Alto, Calif.: Annual Reviews.
- Walster (Hatfield), Elaine, William Wolster, and Ellen Berscheid, eds. 1978 *Equity: Theory and Research*. Boston: Allyn and Bacon.
- Worchel, Stephen, and William G. Austin, eds. 1986 *The Social Psychology of Intergroup Relations*. Monterey, Calif.: Brooks/Cole.
- Zillman, Dolf 1979 *Hostility and Agression*. Hillsdale, N.J.: Erlbaum.

HOWARD B. KAPLAN

## SOCIAL PSYCHOLOGY OF STATUS ALLOCATION

The processes of status allocation are among the most important phenomena of societal stratification structures, along with the ways in which those structures vary over time and between societies and the causes and consequences of those variations (Haller 2000). Within these partly stable

structures, these processes include the formation of individuals' status aspirations, the effect of those aspirations on attained statuses, and the causes of status aspirations (Haller 1982), including those describing a person's social origins (Alwin 1989). One's social status is defined in terms of both ascribed and achieved characteristics. Unless the value or meaning of ascribed characteristics changes, it is only through change in one's achieved characteristics that one's status in society can change. In modern society, occupation, income, and education are the most common achieved characteristics that are studied within the framework of status attainment.

People tend to see themselves and others as occupying positions along hierarchical continua, with evaluations being associated with a person's location in the hierarchy. One's plan, however vague, to strive for a particular place in a status hierarchy is defined as one's level of aspiration. As a mediator of achieved characteristics, and perhaps of ascribed characteristics as well, one's levels of aspiration affect one's levels of attainment (Haller 1982).

Status attainment research is concerned with the process by which people come to occupy their positions in life, some higher and some lower. Through social policy and social practice, the social system provides the opportunity structure for status attainment. The interface between sociology's study of stratification and psychology's study of individual motivation and achievement is dealt with by the field of the social psychology of status attainment.

### **HISTORY OF THE THEORETICAL FIELD**

From the 1960s to the early 1980s, the theory of status allocation developed vigorously, increasing its theoretical comprehensiveness while maintaining its conceptual parsimony and explanatory power (Sewell et al. 1969, 1970; Haller and Woelfel 1972; Haller and Portes 1973; Haller 1982). These developments were matched by similar improvements in methods for testing the hypotheses of the theory (Duncan 1968; Duncan et al. 1968; Van De Geer 1971; Hauser et al. 1983). The social psychological theory of status attainment is both coherent and comprehensive by current standards of social science theory. The theory emerged from two traditions: one sociological, focusing on social

stratification research, and the other from social psychology, focusing on the self-concept and its formation.

The modern study of social stratification began with the work of Sorokin (1927) and was continued by Svalastoga (1965), Duncan (1968), and Haller (1970), among others. Stratification theorists typically posit at least four classes of fundamental status variables, or status content dimensions. Those content dimensions reflect the societal rewards an individual receives and the means to obtain those rewards. Thus, they also may be seen as dimensions of power. Content dimensions are conceptualized as political status, economic status, social status, and informational status (Haller 2000). In status attainment research, educational attainment, occupational prestige, and income are examined as the main variables by which one can measure a person's positions on the content dimensions.

The social psychological tradition that influenced status attainment is based primarily on work by Mead (1934). Lewin (1939) and Heider (1958) influenced this tradition as well. This tradition suggests that status in open societies is earned as opposed to being bestowed by a person's lineage. Before persons assume their eventual statuses, they knowingly or unknowingly develop a level of aspiration for educational status, occupational status, income, and political influence. One's level of aspiration is formed in three ways: by the modeling of others who are present, by self-reflection, and by adopting the status expectations held for a person by others. One can model the behavior of another person whom one knows only through mediated communication (e.g., by learning about another person's behavior by reading about it in books, newspapers, or magazines or by observing that behavior on television or in movies). One also can learn about the status expectations that others hold for oneself through the media. There is evidence, however, that most sources of influence are not transmitted in this way but instead by the direct and indirect effects of "significant others" (Haller and Woelfel 1969, 1972; Haller et al. 1969).

Once formed, aspirations are difficult to change, and they guide the decisions one makes about life. Consequently, one's level of aspiration is a significant determinant of one's level of attainment (Haller 1982).

Since the early 1980s, much work on status attainment has not been systematic but instead directed by policy considerations (Haller 1982). Interestingly, in the last two decades, research on status attainment has grown more in other countries than it has in the United States. In the United States, approximately the same number of studies were conducted on status attainment in each year over the last two decades. However, studies in other countries have more than doubled, and for the first time since the inception of the field, more research is being conducted outside of the United States than within it.

### STATUS ATTAINMENT MODELS

A complete model of the process of status attainment does not exist. However, many significant contributions to what is known about status attainment have come from just a few perspectives. The work of Blau and Duncan (1967) was significant for its careful operationalization of concepts and presentation of formal models subject to statistical analysis. Those authors presented an important model concerned with the effects of parents' status, one's own education, and one's first job, although their model had four flaws. First, it lacked indicators for wealth and power, two important status content dimensions. Second, it contained a relatively primitive theory of the mechanisms of status attainment. Third, it lacked a comprehensive set of exogenous variables, which were limited to the father's occupational and educational statuses. Fourth, much variance in educational and occupational statuses was not explained.

A more complete model of status attainment was introduced informally in 1967; a test of it appeared in Sewell et al. (1969). The conceptual system it embodied, however imperfectly, became known as the Wisconsin model. Path analyses emphasized the social psychological and social structural antecedents of educational and occupational attainment. This early form of the model assumes that all relationships between the key variables are linear and that social psychological variables mediate the process of status attainment. Status aspiration for education and occupation was found to be a powerful mediator of status attainment.

An important contribution of this research was the elaboration, both theoretical and meth-

odological, of the concept of the significant other's influence. One's set of significant others is often larger than one's referent group of parents and peers. This set was defined to include all those who serve as definers and models. *Definers* are those who communicate their expectations, *whereas* models are those who provide illustrations of their statuses and related behaviors; an individual can serve in both capacities. The Wisconsin model changed the conception of social influence used in the study of status allocation from a list of individuals to a set of social processes by which the individuals in one's environment help determine one's status destination.

Sewell et al. (1969) found that the social structural and psychological factors of socioeconomic status and mental ability affect the academic performance of youths and that significant others have great influence in the status attainment process. However, the sample on which their original model was tested was limited, and this was a drawback for evaluating the internal and external validity of their model. The original sample consisted of 929 Wisconsin students who completed a survey as well as a follow-up group of males studied in 1964 whose fathers had been farmers in 1957 (Sewell et al. 1969).

The original form of the Wisconsin model was modified slightly in a study by Sewell et al. (1970), who proposed small changes that would make it applicable to boys with different residential backgrounds. As in the 1969 versions, one's ability and one's significant others were shown to affect one's educational and occupational aspirations. In turn, those aspirations affected one's educational and occupational status attainment (Sewell et al. 1970). Most significant was the finding that the Sewell et al. (1969) model could be used with minor modifications for young men from a variety of backgrounds.

The Wisconsin model was refined in two significant ways by Haller and Portes (1973). First, the model was modified by clarifying and completing the set of content dimensions of status. Second, it was shown that each content dimension of status also is manifested in both of two social psychological isomorphs (status "mirror" images): a status aspiration variable of each focal person and a corresponding variable describing the status orientation levels each of the focal person's significant others expects of him or her (definers) or

illustrates to him or her (models). Third, the Wisconsin model incorporated structural dimensions of status. Two critical structural dimensions are status dispersion and status crystallization. Status attainment models were held to work best when status dispersion is wide, which means that inequality is great; there is little to learn about attainment in fairly homogeneous status systems. Status crystallization is the degree of correlation among status content dimensions (e.g., how one's wealth corresponds to one's power). When crystallization is high, a status attainment model is relatively simplified: It is as if the ultimate endogenous variable were an unobserved variable with many correlated indicators. In complex societies with moderate to high crystallization, status attainment models must be more complex, because the isomorphs of each status content dimension must be treated as separate endogenous variables in the model.

The Wisconsin model was retested in 1983 with more recently developed estimation procedures. It was found to be even more effective in explaining the process of educational and occupational attainment than it previously had been thought to be (Hauser et al. 1983). However, the model was not tested as a whole because not all the variables in the model were measured.

**Concerns with Current Status Attainment Models.** Research on status attainment models has not advanced much since the early 1980s, and flaws in the models of that time remain. More recent research has focused mostly on status inheritance, but the inheritance models are incomplete. None of these models has included measures drawn from the power dimension; indeed, none has seriously attempted to cover the entire range of variables implied by each of the four general status content dimensions (Haller 2000). In the models that have been tested, about 25 to 35 percent of the variance in attainment can be explained by parental status variables. The more nearly complete models explain much more than this.

The social psychological variables used in models today are virtually the same ones used twenty-five years ago. However, the world has changed drastically in the last twenty-five years, particularly with the advent of and pervasive use of communication technology. It is conceivable that mediated sources of influence play a greater role in status

attainment process, but it is impossible to know whether this is true. More generally, it is quite possible that the relationship among status attainment variables has changed, and there may be additional variables to consider.

**Concerns with Current Status Attainment Research.** The initial fifteen years of research from 1967 through 1982 were significant in terms of formulating theory, defining variables, and enlarging the subject populations to which the attainment model applied. Haller (1982) described four avenues of study that would advance status attainment research greatly.

First, there has not been a longitudinal study that followed a cohort sample and observed the influence of the variables in the model on status attainments beyond mid-career. For example, what influences the selection of a second occupation? In addition, how do one's high school occupational, educational, and income aspirations affect one's position and income thirty years or more after those aspirations are formed? Second, most earlier attainment research was restricted to high school students and their first positions after school. The development of status orientations among young children needs to be explored. Third, it is important to understand how status definers emerge in young children. Fourth, the mechanisms that activate the status attainment process have to be explicated.

Above all, the full panoply of variables specified in Haller (1982) should be tested on today's youthful cohorts as they move through life.

## THE STATE OF KNOWLEDGE

In status attainment research, the key variables that are feasible today are educational attainment, occupational prestige, and income. What is known about those variables?

**Educational Status Attainment in the United States.** *Young children and early teens.* The first longitudinal study of educational attainment conducted with young children spanned the period from fall 1985 to spring 1987. Achievement in the first grade was studied for minority groups, including African-Americans and Hispanics. Parental involvement, mobility, and motivation all had a direct influence on first-graders' outcomes. Interestingly,



at least two of the three variables—parental involvement and mobility—are under the direct control of the parents. Additionally, the cognitive readiness of children entering kindergarten was found to have an indirect effect on first-graders' outcomes. All the variables examined had a significant direct or indirect effects on those outcomes, including motivation, peer environment, parental involvement, readiness, and mobility (Reynolds 1989).

The second longitudinal study of educational attainment spanned a one-year period from fall 1987 to fall 1988, using a national sample of 3,116 youths. Data from this study were used to assess science achievement, attitude toward mathematics, and science and mathematics achievement in grades seven and eight. For young teenagers, prior achievement played a large mediating role in future mathematics and science achievement, but classroom context and parental involvement influenced their achievement as well (Reynolds 1991). Science achievement was directly affected by prior achievement, peer environment, and the amount and quality of instruction; mathematics achievement was most strongly influenced by prior achievement and the home environment (Reynolds and Walberg 1991, 1992).

*Teenagers.* Intelligence is a factor in predicting educational and occupational aspirations, but other variables appear to have a greater effect. There is a relatively small effect of parents' statuses on their children's but a relatively large effect of status aspirations and significant others' expectations on one's educational and occupational attainment. Thus, if one has significant others who expect one to go to college, it is more likely that one believes that going to college is possible and one is more likely to go.

Using annual data collected since 1975 for the Monitoring the Future survey, Morgan (1998) found that the educational aspirations of high school seniors increased between the late 1970s and the early 1990s. This effect was greater for white students than for African-American students. Morgan (1998) viewed aspirations as part of a cognitive process of continually calculating the costs and benefits of one's educational aspirations; this variable may influence the effects of significant others. The finding is consistent with the results of many early studies (Haller and Miller, 1963, Hypothesis 4, pp. 31, 41–45, 96), and is an expression of the

well-known Zeigarnik effect (Zeigarnik 1927, Lewin 1951).

The Wisconsin model from the 1970s worked well for the group of white males but not as well for others. With the addition of identity theory (Burke 1989) to the Wisconsin model, that model was found to work not only for white males but for African-American and white females as well. Identity theory suggests that one's identity originates from one's social interactions with others. When one's identity is established, one acts in ways that maintain and confirm that identity. White males and females and African-American females constructed an academic identity that directly influenced their college plans. In other words, if they constructed the meaning of going to college in terms of job-related reasons ("Going to college will help me get the kind of job I want"), they were more likely to go to college. However, the model's predictive ability for African-American males has been minimal. Preliminary research has indicated that one of the reasons for this may be that in that group there is insufficient correspondence between the constructs used in the modified Wisconsin model and the processes associated with African-American male attainment (Burke 1989; Burke and Hoelter 1988).

**Education Outside the United States.** *Australia.* In Australia, education is required for those aged 6–15. Thus, required education ends after the tenth year of school; however, students may elect to stay in the system for two more years. Students who left school after the tenth year earned less than did those who stayed the additional two years. Importantly, socioeconomic background, type of school attended, and career orientations appeared to be unrelated to the decision to leave school after the tenth year (Saha 1985).

Australian research sought to determine whether attending a private Catholic high school influenced attendance at college and the receipt of a college degree. Approximately two-thirds of those from public schools and two-thirds of those from private schools eventually obtain a college degree. However, only approximately 58 percent of those who receive a secondary degree from a Catholic school obtain a college degree (Williams and Carpenter 1990). Thus, receiving a private education results in the same chance of obtaining a college degree as does receiving a public school educa-

tion, but receiving a Catholic school education gives one a relatively lesser chance to obtain a college degree.

Young adult unemployment in Australia varied between 27 percent and 52 percent from the mid-1960s to the early 1980s. One might assume that most youths were encouraged to obtain further education, but that was not always the case. A revision of the Ajzen–Fishbein model of attitude–behavior relations (Carpenter et al. 1989) was used to assess youths’ intentions toward entering the workforce immediately after high school; the influence of economic conditions also was considered in this model. Parental and peer influence played a powerful role in molding a youth’s intentions; however, the youth’s decision to transform the intention into action was mediated by his or her self-perception of past academic performance.

*Greece.* A Greek study (Kostakis 1992) examined information and occupational demands in terms of the specific sources Greek students might use to influence their decisions about the future. How and from whom one gains information was considered a socially determined process. An individual’s significant others, consisting of friends and relatives, appeared to be the most important source of information for all groups; however, significant others were more available to higher-status individuals. Considered as an information source for vocational occupations, schooling was relied on by lower-status youths, rural youths, and girls.

*Israel.* In many studies, teachers were viewed as significant others, but their influence on status attainment did not appear to be large. However, little was known about a teacher’s long-term influence on status attainment. A national representative sample of 834 Israeli adults aged 21–65 was studied to ascertain the effect of the influence of a former teacher (Enoch et al. 1992). The group was divided into two cohort groups: older (ages 40–65) and younger (ages 21–39). Perceived teacher influence was found to be a determining factor in respect to occupation only for the older group. Furthermore, it was found that the Oriental (Sephardic) older group perceived teacher influence as being greater than did any of the other groups in terms of occupational attainment, whereas the older Ashkenazic group identified perceived

teacher influence as a significant variable in educational attainment.

**Occupational Status Attainment in the United States: The Role of Gender.** Typical status attainment models account for more variance in male occupational attainment than in female occupational attainment. A possible reason is that typical occupational status attainment models view occupations as discrete categorical variables as opposed to preferences along a continuum. When occupational titles were measured as a continuous variable, as was done by the Wisconsin researchers, it was found that a student’s gender and a family’s socioeconomic status were related to occupational choice. Additionally, significant others’ expectations for a student played a role in determining aspirations. Significant others’ expectations appeared to be affected most by gender as opposed to aptitude or ability. Thus, males and females may have their aspirations influenced by significant others who seem to choose traditional gender occupations for them (Saltiel 1988).

After the 1960s and 1970s, occupational status attainment research in the United States focused on gender differences in attainment, perhaps because changes in gender roles accelerated at that time. More women have chosen to pursue higher education and enter male-dominated careers than ever before. How will the increasing diversity of occupations open to women affect women’s aspirations and eventual occupational attainment?

There were no significant differences in levels of occupational aspiration between boys and girls and in different high school grades in the early 1960s (Haller et al. 1974). However, little research examining gender differences was conducted in the 1970s and the early to middle 1980s to determine whether aspirations affected occupational attainment.

Consistent with the Wisconsin model, women in the 1980s who pursued male-dominated careers such as business, engineering, and law were found to be subjected to a network of influences, as opposed to a single influence. Parents’ educational level was acknowledged as an influence in typical status attainment research, but until the late 1980s, the specific effect of parents’ possession of a college degree on their children was not considered. For both African-American and white women, parental income was found to have a

significant indirect effect on women's educational attainment. Furthermore, for white women only, the father's and mother's education proved to have significant indirect effects (Gruca et al. 1988).

Women recently have been receiving approximately 50 percent of the bachelor's degrees in life sciences and mathematics, but they are significantly underrepresented in science and mathematics occupations (National Research Council 1991). If women are attaining initial degrees in the same numbers as men, why are they not represented equally in the fields to which those degrees lead?

For women, significant factors associated with persistence in scientific and mathematical careers after college include receiving encouragement from teachers and parents, particularly the mother (Rayman and Brett 1995). Those who stay in science and mathematics careers after graduation do not necessarily believe that their current occupation is compatible with family life; however, the majority of these women have not been affected by family needs. Those who changed careers from a mathematics- or science-related occupation to one not oriented in those directions were more likely to believe that family need plays a role in occupational attainment. Over 50 percent of this group had taken time off from work, refused promotions, reduced their work schedules to part-time, or changed location because of family need. Grades in science and mathematics courses did not significantly affect attitude and achievement in science and mathematics. Additionally, self-esteem and perceived self-confidence did not play a role in deciding who stayed with scientific and mathematical careers and who did not (Rayman and Brett 1995).

**Occupational Status Attainment outside the United States.** *Canada.* The Canadian Mobility Study (Boyd et al. 1995) was directly influenced by the 1962 Occupational Change in a Generation Study in the United States. Using data from the early 1970s, this study showed that great inequalities exist in income, assets, and educational attainment between genders and among those with various ethnic origins (Porter 1995, p. 61). Preliminary research indicated that motherhood, as opposed to or in addition to being married, was the most significant variable determining the occupational status attainment of native-born Canadian

women (Boyd 1995a, p. 284). This finding makes sense, as motherhood, as opposed to marriage, generally requires time off from work. Some women choose to stay at home to raise a child; even if this time off from work is brief, by affecting continuity of employment, it affects one's advancement potential.

Interestingly, occupation and status seem to be consistent from generation to generation (McRoberts 1995, p. 98). One of the reasons for this is that the advantages of background often are passed on to children. For example, wealthy parents are more likely to have received higher education and are better able to afford to have their children receive higher education. Although it is not impossible for a child from a lower-income family to attend an institution of higher education, it is not as likely. This conclusion comes from a study of Canadian-born males in 1973 whose occupations were compared with those of their fathers.

Also of interest is the role of immigrants in status attainment levels in Canada. Native-born Canadian men have an average occupational status lower than that of American-born, German-born, and United Kingdom-born male immigrants and an average occupational status higher than that of immigrants from Poland, Italy, Greece, and Portugal (Boyd 1995b, p. 440). Much of the inequality between the Canadian-born and non-Canadian-born men results from differences in family origin and education.

Similarly, non-Canadian-born women tend to have an average occupational status lower than that of Canadian-born women, but non-Canadian status seems to have less of an impact on female immigrants from the United States and the United Kingdom. As in the United States, females have a lower average occupational status than do males (Boyd 1995b, p. 441).

*Taiwan.* Taiwan consists primarily of three ethnic groups: the aborigines, the Taiwanese, and the mainlanders. Although the Taiwanese account for slightly more than 85 percent of the population, the mainlanders, who account for approximately 12 percent, hold the political power. A study involving 3,924 men from the three ethnic groups determined that the mainlanders had an average occupational status higher than that of the Taiwanese or the aborigines (Tsai 1992).

The father's occupation was found to be the determining factor for first occupation among aborigines over 35 years old. However, for those under 35 years of age, residence and educational attainment were found to be significant influences on the first occupation. This result was different for the Taiwanese and the mainlanders, for whom the most important determinant of first occupation was their level of educational attainment (Tsai 1992).

*Israel.* In Israel, a better education does not necessarily predict better occupational attainment (Semyonov and Yuchtman-Yaar 1992). It once was believed that as Arabs became increasingly integrated into the general Jewish population in Israel, educational attainment and status attainment would become more equal between those two groups. From 1972 through 1983, the Arab population increased its average educational attainment level; however, its average occupational attainment declined. Market discrimination was estimated to account for 6.5 percent of the occupational gap between Jews and Arabs in the highest age group (ages 54–65), but its effect increased to nearly 25 percent in the youngest age group (ages 25–36). Clearly, there are social variables at work here that are not included in traditional status attainment models.

**Political and Economic Status.** Little research has been conducted in the areas of political and economic status. The link of economic status to the prior generation is much weaker than are educational and occupational links that generation. There are a number of reasons why economic status is much more difficult to study than are the educational or occupational variables. Many studies rely on participants to report information for their parents, and although occupational and educational attainments generally are known to family members, specific income information, particularly over a life span, is not. Also, the fluctuating rate of inflation makes it difficult to compare incomes directly across generations. Additionally, an economy's supplies and demands vacillate and ultimately determine an occupation's worth at a given moment. Thus, although the prestige of occupations may not change much, the income associated with those positions may change a great deal, in part as a result of market forces. Finally, the range of incomes today is greater than ever before. Chief executive officers, entertainment

performers, and professional athletes command high incomes. With more income "outliers" today, reliably measuring income and incorporating it into status attainment models are difficult. As a result of these factors, less is known about income status attainment than about educational or occupational attainment.

Similarly, little is known about political status attainment. Political status originally was defined as influence, authority, coercion, and power. Unlike occupational and educational achievement, which have been relatively well defined, there is little agreement on a person's political status.

In a study involving sports teams across cultures, age, experience, and performance were deemed to be the most significant factors in defining status in Canada and India (Jacob and Carron 1996). Both cultures gave more importance to achieved sources of status, such as experience and performance level, than to ascribed status, such as religion, race, and parental occupation. Surprisingly, age was found to be a significant determinant of status, apparently because it serves as an indicator of experience.

## CONCLUSION

Some analysts, including Breiger (1995) and Ganzeboom et al. (1991), believe that theory formulation has become very narrow in social attainment research. However, there are several directions future research can take. First, status allocation research can increasingly feature the systematic incorporation of societal factors considered from the perspective of the individual. For example, advances in network analysis will allow measurement of extended networks and the influences of their members (Wasserman and Faust 1994). Similarly, simulation can provide a method to test models of social influence posited in status allocation models, allowing the investigation of the stability, equilibrium, rate of change, and other qualitative features of status dynamics in a social group (see Gilbert and Doran 1994; Jacobsen and Bronson 1995; Latané 1996). Finally, more extensive measurement of the multidimensional features of occupations and the related variables will allow the creation of models of greater complexity, for example, reflecting nonprestige or nonhierarchical features of status allocation (see Woelfel and Fink 1980).

This article makes it evident that researchers need to focus anew on conceptual clarity and theoretical parsimony. In the future, it is important that new research be executed with variables that include all that have been specified as crucial, along with the causal lines that were so specified, as elements of the theory referred to as the Wisconsin model (Haller 1982). As implied here, this should be done for both males and females in different decades and in societies with differing stratification structures.

## REFERENCES

- Alwin, D. F. 1989 "William H. Sewell: Recipient of the 1988 Cooley-Mead Award." *Social Psychology Quarterly* 52(2):85-87.
- Blau, P. M., and O. D. Duncan, 1967 *The American Occupational Structure*. New York: Wiley.
- Boyd, M. 1995a "Educational and Occupational Attainments of Native-Born Canadian Men and Women." In M. Boyd, J. Goyder, F. E. Jones, H. A. McRoberts, P. C. Pineo, and J. Porter, eds., *Ascription and Achievement: Studies in Mobility and Status Attainment in Canada*. Ottawa, Canada: Carleton University Press.
- 1995b "Immigration and Occupational Attainment in Canada." In M. Boyd, J. Goyder, F. E. Jones, H. A. McRoberts, P. C. Pineo, and J. Porter, eds., *Ascription and Achievement: Studies in Mobility and Status Attainment in Canada*. Ottawa, Canada: Carleton University Press.
- , J. Goyder, F. E. Jones, H. A. McRoberts, P. C. Pineo, and J. Porter 1995 *Ascription and Achievement: Studies in Mobility and Status Attainment in Canada*. Ottawa, Canada: Carleton University Press.
- Breiger, R. L. 1995 "Social-Structure and the Phenomenology of Attainment." *Annual Review of Sociology* 21:115-136.
- Burke, P. J. 1989 "Academic Identity and Race Differences in Educational Aspirations." *Social Science Research* 18:136-150.
- , and J. W. Hoelter 1988 "Identity and Sex-Related Differences in Educational and Occupational Aspirations Formation." *Social Science Research* 17:29-47.
- Carpenter, P. G., J. A. Gleishman, and J. S. Western 1989 "Job Intentions and Job Attainment: Young People's Career Beginnings." *Australian Journal of Education* 33(3):299-319.
- Duncan, O. D. 1968 "Social Stratification and Mobility: Problems in the Measurement of Trend." In E. B. Sheldon and W. E. Moore, eds., *Indicators of Social Change*. New York: Russell Sage Foundation.
- Enoch, Y. R. Shapira, and A. Yogeve 1992 "Teachers as Significant Others in the Status Attainment of Israeli Adults." In A. Yogeve, ed., *International Perspectives on Education and Society*, vol. 2. Greenwich, Conn.: JAI Press.
- Ganzeboom, H. B. G., D. J. Treiman, and W. C. Ultee 1991 "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277-302.
- Gilbert, N., and J. Doran, eds. 1994 *Simulating Society: The Computer Simulation of Social Phenomena*. London: UCL Press.
- Gruca, J. M., C. A. Ethington, and E. T. Pascarella 1988 "Intergenerational Effects of College Graduation on Career Sex Atypicality in Women." *Research in Higher Education* 29(2):99-124.
- Haller, A. O. 1970 "Changes in the Structure of Status Systems." *Rural Sociology* 35:469-487.
- 1982 "Reflections on the Social Psychology of Status Attainment." In R. M. Hauser, D. Mechanic, A. O. Haller, and T. S. Hauser, eds., *Social Structure and Behavior*. New York: Academic Press.
- 2000 "Societal Stratification." In E. A. Borgatta and Rhonda V. J. Montgomery, eds., *Encyclopedia of Sociology*, 2nd ed. New York: Macmillan.
- Haller, A. O., and Miller, I. W. 1971 *The Occupational Aspiration Scale*. Cambridge, MA: Schenkman.
- , L. B. Otto, R. F. Meier, and G. W. Ohlendorf 1974 "Level of Occupational Aspiration: An Empirical Analysis." *American Sociological Review* 39:113-121.
- , and A. Portes 1973 "Status Attainment Processes." *Sociology of Education* 46:51-91.
- , and J. Woelfel 1969 "Identifying Significant Others and Measuring Their Expectations for a Person." *Revue Internationale de Sociologie* 5:395-429.
- 1972 "Significant Others and their Expectations: Concepts and Instruments to Measure Interpersonal Influence on Status Aspirations." *Rural Sociology* 37:591-622.
- , —, and E. L. Fink 1969 *The Wisconsin Significant Other Battery*. Arlington Heights, Va.: Educational Resources Information Document Center.
- Hauser, R. M., S. Tsai, and W. H. Sewell 1983 "A Model of Stratification with Response Error in Social and Psychological Variables." *Sociology of Education* 56:20-46.
- Heider, F. 1958 *The Psychology of Interpersonal Relations*. New York: Wiley.

- Jacob, C. S., and A. V. Carron 1996 "Sources of Status in Sport Teams." *International Journal of Sport Psychology* 27(4):369-382.
- Jacobsen, C., and R. Bronson 1995 "Computer Simulations and Empirical Testing of Sociological Theory." *Sociological Methods and Research* 23:479-506.
- Kostakis, A. 1992 "Social Determinations of the Use of Information as It Relates to Occupational Choices of Greek Youth." In A. Yougev, ed., *International Perspectives on Education and Society*, vol. 2. Greenwich, Conn.: JAI Press.
- Latané, B. 1996 "Dynamic Social Impact Theory: The Creation of Culture by Communication." *Journal of Communication* 46(4):13-25.
- Lewin, K. 1939 "Field Theory and Experiment in Social Psychology." *American Journal of Sociology* 44:868-897.
- Lewin, K. 1951 *Field Theory in Social Science*. D. Cartwright, ed. New York: Harper.
- McRoberts, H. A. 1995 "Mobility and Attainment in Canada: The Effects of Origin." In M. Boyd, J. Goyder, F. E. Jones, H. A. McRoberts, P. C. Pineo, and J. Porter, eds., *Ascription and Achievement: Studies in Mobility and Status Attainment in Canada*. Ottawa, Canada: Carleton University Press.
- Mead, G. H. 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Morgan, S. L. 1998 "Adolescent Educational Expectations." *Rationality and Society* 10:131-162.
- National Research Council 1991 *Women in Science and Engineering: Increasing Their Numbers in the 1990's*. Washington, D.C.: National Academy Press.
- Porter, J. 1995 "Canada: The Societal Content of Occupational Allocation." In M. Boyd, J. Goyder, F. E. Jones, H. A. McRoberts, P. C. Pineo, and J. Porter, eds., *Ascription and Achievement: Studies in Mobility and Status Attainment in Canada*. Ottawa, Canada: Carleton University Press.
- Rayman, P., and B. Brett 1995 "Women Science Majors—What Makes a Difference in Persistence after Graduation?" *Journal of Higher Education* 66(4):388-414.
- Reynolds, A. J. 1989 "A Structural Model of First Grade Outcomes for an Urban, Low Socioeconomic Status, Minority Population." *Journal of Educational Psychology* 81(4):594-603.
- 1991 "The Middle Schooling Process: Influences on Science and Mathematics Achievement from the Longitudinal Study of American Youth." *Adolescence* 26(101):133-158.
- , and H. J. Walberg, 1991 "A Structural Model of Science Achievement." *Journal of Educational Psychology* 83(1):97-107.
- 1992 "A Process Model of Mathematics Achievement and Attitude." *Journal for Research in Mathematics Education* 23(4):306-328.
- Saha, L. J. 1985 "The Legitimacy of Early School Leaving: Occupational Orientations, Vocational Training Plans, and Educational Attainment among Urban Australian Youth." *Sociology of Education* 58:228-240.
- Saltiel, J. 1988 "The Wisconsin Model of Status Attainment and the Occupational Choice Process." *Work and Occupations* 15(3):334-355.
- Semyonov, M., and E. Yuchtman-Yaar 1992 "Ethnicity, Education, and Occupational Inequality: Jews and Arabs in Israel." In A. Yougev, ed., *International Perspectives on Education and Society*, vol. 2. Greenwich, Conn.: JAI Press.
- Sewell, W. H., A. O. Haller, and G. W. Ohlendorf 1970 "The Educational and Early Occupational Status Attainment Process: Replication and Revision." *American Sociological Review* 35(16):1014-1027.
- , ——, and A. P. Portes 1969 "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34(1):82-92.
- Sorokin, P. A. 1927 *Social Mobility*. New York: Harper and Row.
- Svalastoga, K. 1965 *Social Differentiation*. New York: David McKay.
- Tsai, S. 1992 "Social Change and Status Attainment in Taiwan: Comparisons of Ethnic Groups." In A. Yougev, ed., *International Perspectives on Education and Society*, vol. 2. Greenwich, Conn.: JAI Press.
- Van de Geer, J. P. (1971). *Introduction to Multivariate Analysis for the Social Sciences*. San Francisco: W. H. Freeman.
- Wasserman, S., and K. Faust 1994 *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.
- Williams, T., and P. G. Carpenter 1990 "Private Schooling and Public Education." *Australian Journal of Education* 34(1):3-24.
- Woelfel, J. D., and E. L. Fink 1980 *The Measurement of Communication Processes: Galileo Theory and Method*. New York: Academic Press.
- Zeigarnik, B. 1927, Über das behalten von erledigten und unerledigten Handlungen. *Psychologisch Forschung* 9:1-85.

ARCHIBALD O. HALLER  
EDWARD L. FINK  
LAURA JANUSIK

## SOCIAL RESOURCES THEORY

**NOTE:** Although the following article has not been revised for this edition of the *Encyclopedia*, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

This article introduces the theory of social resources (Lin 1982, 1983). It describes the fundamental propositions of the theory and reviews empirical research programs and results pertaining to the theory. It concludes with a discussion of some issues regarding extensions and modifications of the theory.

Resources are goods, material as well as symbolic, that can be accessed and used in social actions. Of particular interest are the valued resources—resources consensually considered as important for maintaining and improving individuals' chances of survival as they interact with the external environment. In general, valued resources are identified with indicators of class, status, and power in most societies. In the following discussion, resources refer to valued resources.

Resources can be classified in two categories: personal resources and social resources. *Personal resources* are resources belonging to an individual; they include such ascribed and achieved characteristics as gender, race, age, religion, education, occupation, and income as well as familial resources. These resources are in the possession of the individual and at the disposal of the individual. *Social resources*, on the other hand, are resources embedded in one's social network and social ties. These are the resources in the possession of the other individuals to whom ego has either direct or indirect ties. A friend's car, for example, may be ego's social resources. Ego may borrow it for use and return it to the friend. Ego does not possess the car, and accesses and uses it only if the friend is willing to lend it. The friend retains the ownership. Similarly, a friend's social, economic, or political position may be seen as ego's social resources. Ego may seek the friend's help in exercising that resource in order for ego to obtain or achieve a specific goal.

Much of sociological research focuses on personal resources. While social network analysis has been a long-standing research tradition in sociology and psychology, attention had been given to

the structure and patterns of ties and relations. Only recently, in the past two decades, sociologists and anthropologists have explored the theoretical significance of the resources brought to bear in the context of social networks and social ties. The theory of social resources makes explicit the assumption that resources embedded in social connections play important roles in the interaction between social structure and individuals. More specifically, the theory explores how individuals access and use social resources to maintain or promote self-interests in a social structure that consists of social positions hierarchically related and organized in terms of valued resources. It has been argued that social resources are accessed and mobilized in a variety of actions by an individual to achieve instrumental and/or expressive goals.

Two terms need some clarifications here. I assume that a social structure consists of different levels, each of which can include a set of structurally equivalent positions. They are equivalent primarily on the basis of levels of similar valued resources, and secondarily, similar life-styles, attitudes, and other cultural and psychological factors. For the purposes here, the terms, "levels" and "positions," are used interchangeably. Also, status attainment is assumed to refer to the voluntary aspect of social mobility. Involuntary social mobility, due to job dissatisfaction, lack of alternatives, or other "pushing" or forced factors, is excluded from consideration. As Granovetter (1986) pointed out, voluntary social mobility generally results in wage growth. Likewise, it is argued that voluntary social mobility accounts for the majority of occurrences in status attainment.

### THEORY OF SOCIAL RESOURCES AND SOCIAL ACTIONS

Attention in this article will be given to the theory of social resources as it is applied to the context of instrumental actions. Instrumental actions are a class of actions motivated by the intent to gain valued resources (e.g., seeking a better occupational position). In contrast, expressive actions are a class of actions motivated by the intent to maintain valued resources (e.g., seeking to maintain a marital relationship). Social resources have broad implications for both types of social actions (Lin 1986). However, for the present discussion, social resources will be considered in the perspective of

instrumental actions only. To carry the discussion at a more concrete level, attention will be given to the status attainment process, which can be seen as a typical process focusing on an instrumental goal. In the following material, the propositions of the theory of social resources will be presented in the specific framework of the status attainment process, to illuminate clearly and concretely the theoretical implications in a specific research tradition.

I have specified three hypotheses (Lin 1982): the social resources hypothesis, the strength-of-position hypothesis, and the strength-of-ties hypothesis. The *social resources hypothesis*, the primary proposition of the theory, states that *access to and use of better social resources leads to more successful instrumental action*. In the case of status attainment, it predicts that job-seekers are more likely to find a better job (in terms of prestige, power, and/or income) when they are able to contact a source with better resources (in terms of occupation, industry, income, etc.).

The other two hypotheses identify factors that determine the likelihood of access to and use of better social resources. The *strength-of-position hypothesis* stipulates that *the level of original position is positively associated with access to and use of social resources*. For the process of status attainment, it suggests that the original social position of a job-seeker is positively related to the likelihood of contacting a source of better resources. Position of origin can be represented by characteristics of ego's parents or previous jobs.

The *strength-of-ties hypothesis* proposes that *use of weaker ties is positively related to access to and use of social resources*. For status attainment, it states that there is a positive relationship between the use of weaker ties and the likelihood of contacting a source of better resources. For the formulation of the strength of weak ties argument, see Granovetter (1973, 1974).

Thus, the theory contains one proposition postulating the effect of social resources and two propositions postulating causes of social resources. The strength-of-position hypothesis implies an inheritance effect. A given position of origin in the hierarchical structure in part decides how well one may get access to better social resources embedded in the social structure. It is a *structural* factor and independent of individuals in the structure, although individuals may benefit. On the other

hand, the strength-of-ties hypothesis suggests the need for *individual action*. Normal interactions are dictated by the homophily principle, the tendency to engage in interaction with others of similar characteristics and life-styles. Going beyond the routine set of frequent interactants and seeking out weaker ties represent action choices beyond most of the normative expectations of the macrostructure (see Granovetter 1973, 1974).

It is true that the beginning of a job search often is unplanned. Many job leads become available through casual occasions (e.g., parties) and through interactions with casual acquaintances. It is not necessarily the case that a job search always begins with the individual actively seeking out contacts for this purpose. However, this does not negate the basic premise that individuals are situated at different levels of positions in the structure and have, therefore, access to "casual" occasions involving participants of certain types and amounts of resources, including social resources. In fact, it has been empirically demonstrated (Campbell, Marsden, and Hulbert 1986; Lin and Dumin 1986) that higher-level positions have greater access to more diverse and heterogeneous levels of positions in the hierarchical structure than lower-level positions, therefore having greater command of social resources. Thus, it can be expected that "casual" occasions for the higher-level positions are structurally richer in job and other types of information and influence. Such structural advantage, deducible from the pyramidal assumption of the theory, has distinct effect when a job search is eventually launched by the individual. *In relative terms, the strength of position should have stronger effects on social resources than the strength of ties*. This statement recognizes the significance of structural constraints everywhere in the social structure. In empirical systems, both factors are expected to operate, even though their relative effects may vary.

#### EMPIRICAL RESEARCH AND THEORETICAL EXTENSIONS

Research programs examining the theory of social resources in the context of socioeconomic attainment have been carried out in North America (Ensel 1979; Lin, Ensel, and Vaughn 1981; Lin, Vaughn, and Ensel 1981; Marsden and Hulbert 1988), in West Germany and the Netherlands (Flap and De Graaf 1988; De Graaf and Flap 1988;



Sprengers, Tazelaar and Flap 1988; Boxman, Flap, and De Graaf 1989; Wegener 1991), in Taiwan (Sun and Hsiong 1988), and in China (Lin and Bian 1990). Thus far, evidence strongly supports two of the three hypotheses: the social-resource hypothesis and the strength-of-positions hypothesis. Those with better origins tend to find sources for better resources in job-seeking, while contacting a source of better resources increases the likelihood of finding a better job. These relations hold even after the usual status attainment variables (e.g., education and first-job status) are taken into account. These results, as Marsden and Hulbert showed, are not biased by the fact that only those contacting interpersonal sources in job-seeking are selected for study.

However, evidence is equivocal on the strength-of-(weak) ties hypothesis. For example, Lin and associates have found evidence that weaker ties linked job-seekers to contacts with better resources, whereas Marsden and Hulbert (1988) did not. The different findings may be due to the interaction between the two exogenous variables: the strength of position and the strength of ties. Lin and others have found that the advantage of using weaker ties over the use of stronger ties decreases as the position of origin approaches the top of the levels. Lin, Ensel, and Vaughn (1981) hypothesized a ceiling effect for weak ties. At the top of the hierarchical structure there is no advantage to using weak ties, since such ties are likely to lead to inferior positions and therefore inferior resources. They did not anticipate similar ineffectiveness of weaker ties toward the bottom of the structure. Marsden and Hulbert (1988), however, also found that those with the lowest origins did not benefit more from contacts with weaker ties in gaining access to better resources than from contacts with stronger ties. One speculation is that those at the lower positions have more restricted range of contacts (Campbell, Marsden and Hulbert 1986; Lin and Dumin 1986), rendering the weaker ties accessible less effective. Thus, a nonlinear relationship (interaction) between strength of ties and social resources may be involved (Wegener 1991).

Another elaboration concerns the distinction between two types of social resources: network resources and contact resources. Network resources refer to resources embedded in one's ongoing social networks and ties. In this conceptualization, the researcher is interested in identifying the on-

going social ties, and from these identified ties, exploring resources they have. These resources are seen as social resources to ego (Campbell, Marsden, and Hulbert 1986; Lin and Dumin 1986; Boxman and Flap 1990). Contact resources, on the other hand, refer to resources associated specifically with a tie or ties accessed and mobilized in a particular action. For example, the researcher is interested in identifying the contact ego used in a particular job-seeking situation and specifying the social resources in terms of what resources the contact possessed (Lin, Ensel, and Vaughn 1981; Marsden and Hulbert 1988; Sun and Hsiong 1988; Lin and Bian 1990). Recent research (Lai, Leung, and Lin 1990) shows that network resources and contact resources are two conceptually distinctive and causally related components of social resources. Network resources, reflecting resources in ego's social network, contribute to the access of contact resources in the context of a particular action (e.g., seeking a job). Each in turn contributes to the ultimate success of the action (e.g., getting a high-status occupation).

#### FURTHER RESEARCH ISSUE

Some theoretical and methodological issues remain in the extension and application of the social resources theory.

One issue concerns the cost of social resources. Unlike personal resources, which ego may use and dispose of relatively free of constraints, social resources are "borrowed" from one's social ties. Thus, there should be a cost attached to such access. In most cases there is an implied obligation of reciprocity—that is, ego is committed to offer his or her resources as social resources to the alter from whom resources have been borrowed. The problem arises when ego and the alter do not occupy similar social positions, thus possessing dissimilar resources. In the case of ego seeking help from the alter, in fact, the better the social position the alter occupies, the more effective it provides social resources to ego. It is conceivable that ego possesses other resources, which may provide to be useful to the alter in the reciprocity process. For example, a banker (ego) may seek political influence from a politician (the alter), who in turn may secure financial benefit with ego's help. Fair exchange of different valued resources occurs. There will also be situations where ego

with inferior resources gains as a result of help from an alter with superior resources (e.g., a graduate student getting a desirable job with the help of a professor), the reciprocity becomes more intricate. One way of reciprocity requires quantity in compensation of quality (e.g., willingness to put more effort into a research or writing collaboration). Another form of reciprocity requires efforts to increase the value of the alter's resources (e.g., citations to the professor's work in one's publications). Variations in such reciprocal uses of social resources and, therefore, in cost deserve further conceptualization and research.

Another area worthy of research attention is the use of social resources for expressive actions. It has been hypothesized that, in contrast to instrumental actions, expressive actions would be more effective if ego and the alter share similar traits and experiences. The argument is that homophily (sharing similar characteristics and life-styles) increases the likelihood of the alter understanding the emotional stress experienced by ego (Lin 1986). Thus, the expectation is that strong ties, rather than weak ties, may provide the more desirable social resources for expressive actions. However, reality is much less tidy than this conceptualization. In some expressive actions (e.g., seeking support in time of a divorce), both emotional and instrumental support are needed. Further complicating the situation is that often the strong ties (e.g., spouse) are the sources of stress, and expressive actions must by definition be provided by either weaker ties or surrogate strong ties (e.g., relatives or a friend or professional helper) (Lin and Westcott forthcoming). Much more conceptual and empirical work is needed to tease out these issues.

Finally, there is the intriguing question of whether the theory of social resources can help conceptualizing the interplays between social structure and social action. I argue that the theory of social resources makes two kinds of contributions toward an understanding of social structure and social action (Lin 1990a, 1990b). First, research on social resources has offered the plausibility that under structural constraints, individual choices (in terms of social ties and social contacts) may yield different and meaningful consequences. It has been shown that given two individuals with similar personal resources (including original social positions), they might experience different outcomes in instrumental actions, depending on social re-

sources they access and use. To an extent such different access is dictated by structural constraints. As mentioned earlier, original position affects the range of social ties in the social hierarchy and therefore the likelihood of accessing better social resources. However, after such structural constraints have been taken into account, there is evidence that some flexibility remains in the choice of social ties and use of social resources, and such choice and use yield meaningful and different results.

Second, much of past research on social structure as well as social resources has assumed that social structure has a priori existence and imposes constraints within which individuals conduct meaningful actions. The theoretical possibility that individual actions and choices may constitute fundamental driving forces in the formation and functioning of social structures has gained currency in sociology (Coleman 1986, 1988, 1990). Social resources, it is argued, may also contribute to this theoretical formulation.

One may assume that individuals strive to gain resources for the promotion and maintenance of one's survival and well-being. Personal resources may be preferred to social resources in this striving, since the former incur less cost and are more manipulatable. However, the speed of cumulation may differ for the two types of resources. Acquisition and cumulation of personal resources may be additive. On the other hand, acquisition and cumulation of social resources may be exponential, in that once a social tie is established, not only the tie's personal resources become social resources to ego, but the tie's social resources (through its ties) also become social resources. Thus, social ties, through their networking patterns and dynamics, accelerate one's social resources. While social resources come at a cost, as discussed earlier, it is to the benefit of ego to acquire as much social resources as possible. Thus, social resources constitute the fundamental motivation to networking in the promotion and maintenance of one's self-interest and well-being. Such networking constitutes the elementary blocks in the emergence of social structure. Subsequently, the management and manipulation of the constructed and extended network that contains increasingly heterogeneous participants with varying demands for secondary resources (e.g., quality of life considerations) dictate the development of hierarchical positions and

role expectations, which in turn reduce the range of possible individual action choices. Further theoretical work along these lines promises to contribute to the current interest and debate in the interrelationships between social actions and social structure.

(SEE ALSO: *Exchange Theory*; *Social Network Theory*; *Social Support*)

## REFERENCES

- Barbieri, Paolo 1997 "The Hidden Treasure. A Map of the Social Capital in a Metropolitan Area" (Il Tesoro nascosto. La mappa del capitale sociale in un'area metropolitana) *Rassegna Italiana di Sociologia* 38:343-370.
- Beggs, John J., and Jeanne S. Hurlbert 1997 "The Social Context of Men's and Women's Job Search Ties: Membership in Voluntary Organization, Social Resources, and Job Search Outcomes." *Sociological Perspectives* 40:601-622.
- Boxman, E. A. W., Hendrik D. Flap, and P.M. De Graaf 1990 "Social Capital, Human Capital, and Income Attainment: the Impact of Social Capital and Human Capital on the Income Attainment of Dutch Managers in 1986." Paper presented at the European Conference on Social Network Analysis, June, Groningen, Netherlands.
- Boxman, E. A. W., and Hendrik D. Flap 1990 "Social Capital and Occupational Chances." Paper presented at the International Sociological Association, XIIth World Congress of Sociology, July, Madrid.
- Breiger, Ronald L. 1990 *Social Mobility and Social Structure*. New York: Cambridge University Press.
- Campbell, Karen E., Peter V. Marsden and Jeanne S. Hurlbert, 1986 "Social Resources and Socioeconomic Status," *Social Networks*, 8(1):97-116.
- Coleman, James S 1986 *Individual Interests and Collective Action*. Cambridge: Cambridge University Press.
- 1988. "Social Capital in the Creation of Human Capital," *American Journal of Sociology* 94(Supplement):S95-S120.
- 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- De Graaf, Nan Dirk, and Hendrik Derk Flap 1988 "With a Little Help from My Friends," *Social Forces* 67-2:452-472.
- Ensel, Walter M 1979 "Sex, Social Ties, and Status Attainment." Ph.D. diss., Department of Sociology, State University of New York at Albany.
- Flap, Hendrik D. and Nan Dirk DeGraaf 1988 "Social Capital and Attained Occupational Status," *Netherlands Journal of Sociology*.
- Green, Gary P., Leann M. Tigges, and Irene Browne 1995 "Social Resources, Job Search, and Poverty in Atlanta." *Research in Community Sociology* 5:161-182.
- Granovetter, Mark 1973 "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-1380.
- 1974 *Getting a Job*. Cambridge, Mass.: Harvard University Press.
- 1982 "The Strength of Weak Ties: A Network Theory Revisited." In Peter V. Marsden and Nan Lin, eds., *Social Structure and Network Analysis*. Beverly Hills: Sage.
- Kulik, Liat 1997 "Anticipated Dependence: A Determinant in an Integrative Model of Power Relations among Elderly Couples." *Journal of Aging Studies* 11:363-377.
- Lai, Gina Wan-foon, Shu-yin Leung, and Nan Lin 1990 "Network Resources, Contact Resources, and Status Attainment: Structural and Action Effects of Social Resources." Paper presented at the annual meeting of the American Sociological Association, August, Washington, D.C.
- 1998. "Network Resources, Contact Resources, and Status Attainment." *Social Networks* 20:159-178.
- Lin, Nan 1982 "Social Resources and Instrumental Action." In Peter V. Marsden and Nan Lin, eds., *Social Structure and Network Analysis*. Beverly Hills: Sage.
- 1983 "Social Resources and Social Actions: A Progress Report." *Connections* 6:10-16.
- 1986 "Conceptualizing Social Support." In Nan Lin, Alfred Dean, and Walter M. Ensel, eds., *Social Support, Life Events, and Depression*. Orlando, Fla.: Academic Press.
- 1990a "Social Resources and Social Mobility: A Structural Theory of Status Attainment," In Ronald Breiger, ed., *Social Mobility and Social Structure*. New York: Cambridge University Press.
- 1990b "Social Resources and the Emergence of Social Structure," Paper presented at the XII World Congress of Sociology, July, Madrid.
- and Yan-jie Bian 1990 "Getting Ahead in Urban China: Differential Effects of Social Connections (Guanxi)." Paper presented at the Sunbelt Network Conference, February, San Diego.
- , and Mary Dumin 1986 "Access to Occupations through Social Ties." *Social Networks* 8:365-385.
- , Walter M. Ensel, and John C. Vaughn 1981 "Social Resources and Strength of Ties: Structural

- Factors in Occupational Status Attainment." *American Sociological Review* 46:393-405.
- , John C. Vaughn, and Walter Ensel 1981 "Social Resources and Occupational Status Attainment." *Social Forces* 59:1163-1181.
- , and Jeanne Westcott (Forthcoming) "Marital Engagement/Disengagement, Social Networks, and Mental Health." In John Eckenrode, ed., *The Social Context of Stress and Coping*. New York: Plenum.
- Marsden, Peter, and Jeanne Hulbert 1988 "Social Resources and Mobility Outcomes: A Replication and Extension," *Social Forces* 66(4):1038-1059.
- Sprengers, Maarten, Frits Tazelaar, and Hendrik D. Flap 1988 "Social Resources, Situational Constraints, and Reemployment." *Netherlands Journal of Sociology* 24.
- Sun, Chingshan, and Ruimei Hsiong 1988 *Social Resources and Social Mobility*. Taiwan: Tunghai University.
- Tausig, Mark 1990 "Microsocial Implications of the Macrostructural Distribution of Social Resources." *Sociological Focus* 23:333-340.
- Tigges, Leann M., Irene Browne, and Gary P. Green 1998 "Social Isolation of the Urban Poor: Race, Class, and Neighborhood Effects on Social Resources." *Sociological Quarterly* 39:53-77.
- Wegener, Bernd 1991a "Job Mobility and Social Ties: Social Resources, Prior Job and Status Attainment." *American Sociological Review* 56(February)1-12.
- 1991b "Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment." *American Sociological Review* 56:60-71.

NAN LIN

## SOCIAL SECURITY SYSTEMS

In the United States, Social Security refers to a set of programs, including old-age, survivors, and disability insurance, for the elderly and their dependents. This particular use of the term "social security" relates as much to the special and delimited character of the welfare state in the United States as it does to the generally accepted meaning of the term. For organizations such as the International Labour Office and the International Social Security Association and for scholars concerned with comparative studies, the term refers to a wider variety of programs. For instance, in its volume *Social Security throughout the World*, the Social Security Administration states:

*The term "social security" in the context of this report refers to programs established by government statutes which insure individuals against interruption or loss of earning power, and for certain special expenditures arising from marriage, birth, or death. (1985, p. ix)*

The concept of social protection that underlies this definition includes unemployment programs to cover involuntary temporary loss of work, sickness programs to cover loss of income from sickness and the cost of medical care, disability or occupational injury programs to cover physical limitations on working, family allowances to cover loss of economic status from the addition of members to the family, and social assistance to cover circumstances such as family disruption that cause income to fall below specified levels. Protection of earning power from loss of work or health conditions associated with old age also remains crucial. Still, many of the other programs, more common in the advanced welfare systems of western European nations than in the United States, must be considered part of social security systems.

The relative size of programs devoted to the elderly perhaps warrants the special attention paid to old age in social security systems. Among all expenditures for education and social security programs, those for old-age pensions represent the largest component, averaging 36 percent across high-income nations in 1985 (Organization for Economic Cooperation and Development 1988, Tables 1 and 3). The next largest component, 22 percent, is devoted to health care, which also disproportionately benefits the elderly. Furthermore, the growth rate of programs for the aged has exceeded that for other programs, and in the future, expenditures for those programs will account for an even greater proportion of the total. Spending for unemployment, family allowances, and social assistance represents a relatively small part of social security programs. As Myles (1984) notes, the welfare state is primarily a welfare state for the elderly.

The need for collective protection for the aged or others stems from the existence of economic insecurity. Loss of earning power as a result of poor health, old age, or unemployment remains a possibility for nearly all the participants in a market economy but is uncertain enough to make it difficult to predict loss of income or future

savings potential. Traditional protection against such risk in preindustrial societies developed informally through the family. Under ideal circumstances, children and relatives could support parents who were unable to provide for themselves or wanted to step down from their economic role of provider. In premodern societies, social security thus took the form of an intergenerational contract, based on norms of filial piety and parental control over wealth, between children or other relatives and parents (Simmons 1960). Never a guaranteed source of protection, however, other family members became an even less reliable source of support with the decrease in family size, the increase in mobility, and the industrialization of labor that accompanied the demographic and industrial transitions. With the development of large-scale corporate capitalism in the late nineteenth and early twentieth centuries, the risks of forced retirement and unemployment grew. Systems of social security collectivized and formalized the relationship between young workers and elderly, unemployed, or disabled nonworkers. Workers would contribute support to certain categories of non workers in return for the expectation that they would be covered if they became unable to work. The state has always played a crucial role in this collectivized contract by making participation in the system compulsory for most workers. Voluntary programs of saving for unexpected contingencies are insufficient because many people are not rational in saving for events that may not occur or occur only in the far future. Private compulsory systems industries, unions, and businesses similarly face problems of incomplete coverage, financial insolvency, and job movement. In contrast, collectivizing social security provides for reliable funding, and it is easier to predict events for a group than it is for individuals.

Most nations provide for more than social security alone. The broader welfare state in capitalist societies also supports education, retraining, full employment, business regulation, price supports, infrastructure, and legal rights. In the former socialist societies of Eastern Europe, social security systems involved broader social protection through guaranteed employment; subsidized food, housing, and energy prices; and the reduced importance of market performance as the criterion for economic support. Recent market-oriented reforms in eastern Europe may expand the

emphasis on social security systems as they are typically and more narrowly defined in capitalist societies.

If motives of social protection are common to social security systems, the coverage of the population and the distribution of benefits vary widely. Benefits may be distributed on the basis of at least four criteria, each of which may be emphasized or deemphasized in particular systems. First, citizenship entitlement provides basic benefits—usually in the form of flat-rate cash payments—to individuals or families as a right of citizenship regardless of work history, contributions, or income. Second, employment-related criteria base eligibility on wage or payroll contributions made before the contingency that causes earnings to cease. As a form of public or social insurance, these benefits reinforce market criteria of income determination. Third, need-based criteria provide benefits by comparing resources with a standard that typically is based on subsistence needs. Means-tested or social assistance programs target benefits at the most needy, usually those not covered by citizenship or insurance programs. Fourth, entitlement sometimes is granted on the basis of marital or family status, usually to women and homemakers or families with young children.

To a large extent, nations mix their degrees of reliance on the different criteria. Nations that began with universal systems added earnings-based supplements (e.g., Sweden), and those which originally enacted earnings-based benefits have added universal benefits (e.g., Great Britain) or some form of minimum benefit (e.g., the United States). Similar claims have been made about the mix of public and private systems. To limit inequality, nations that traditionally relied on private systems (e.g., the United States) have increasingly expanded public system benefits, while nations that traditionally relied on public systems (e.g., West Germany) have increasingly expanded private benefits for high-income workers who want a higher return on their contributions.

Some argue that social citizenship remains the most important component of social protection because security is not complete until the state grants alternative means to economic welfare to the market (Esping-Andersen 1989; Korpi 1989). Because meager means-tested benefits are structured to avert work disincentive effects, they fail to

emancipate individuals from dependence on the market. Because social insurance benefits stem from labor-based contributions—that is, qualification based on previous contributions defines the right to receive benefits—they maintain links to the market. Because family benefits depend on the qualification of others by virtue of need or contribution, they also fail to detach distribution from the market mechanism. A definition of social security would thus require decommodifying labor or insulating workers from dependence on the market for economic support. According to Esping-Anderson, in decommodifying welfare states,

*citizens can freely, and without potential losses of job, income or general welfare, opt out of work under conditions when they themselves consider it necessary for reasons of health, family, age or even educational self-improvement; when, in short, they deem it necessary for participating in the social community.* (1989, p. 22)

Few, if any, nations meet the high standards defined by citizenship rights or decommodification. Nearly all nations rely at least partially on earnings-related benefits to supplement universal benefits; flat-rate benefits available to all are too expensive to provide generously for all elderly persons. Still, the trend is toward expanded social rights. Recent efforts to gain the right to protection from economic insecurity follow efforts in previous centuries to gain civil rights such as freedom of speech and equality before the courts and the political right to universal voting (Marshall 1964). This process highlights the dynamic meaning of social security and the continuing evolution of its definition.

Political debate over how far governments should extend definitions of social security to include citizenship rights reflects larger tensions between the relative roles in the market and the state in public policy (Myles 1984). On the one hand, inequality in earnings and contributions during work life means that the market retains a strong influence on social security benefits and the financial circumstances of nonworkers. On the other hand, equality of participation in the democratic political system provides impetus for equality in benefits unrelated to the market. The underlying dynamics of market and democracy—differentiation versus equality—both show in varying degrees in the benefit structures of different systems and

debates over the definition of what social security should provide.

### HISTORICAL BACKGROUND

A formal social security system was slow to come to the United States. The first public social security system (although limited in coverage and generosity) emerged in Bismarckian Germany in 1889 and was followed by systems in Denmark in 1891, New Zealand in 1898, Austria in 1906, and Australia and Great Britain in 1908 (Social Security Administration 1985). Legislation at the national level was not passed in the United States until 1935, and the first old-age pension was not paid until 1940. In part, the expansion of disability benefits to Civil War veterans (even if they had not been injured or seen combat) in 1890 provided a de facto pension system for northern whites but did not promote the implementation of a more general national pension system for nonveterans (Orloff and Skocpol 1984).

The reasons why Civil War pensions did not lead to a more comprehensive social security system have been examined extensively. The historical persistence of individualist, laissez-faire values obstructed public support for public programs (Rimlinger 1971). Big business preferred private negotiations with labor, and small business wanted to avoid the cost of social security provisions. Relative to a powerful business community, weak, decentralized labor unions were unable to agree on a common approach or push redistributive public programs as they did in several European nations (Stephens 1979). Relatedly, the United States did not have a socialist or social democratic party committed to labor goals, because regional, ethnic, and racial divisions split clearly defined class interests in support of social legislation. Southern congressional representatives, who wanted to maintain cheap agricultural (particularly black) labor in their region, used their power in a committee-dominated federal government to block legislation (Quadagno 1988). Finally, the lack of a professional civil service bureaucracy to administer the program and the existence of often corrupt patronage politics at the local level might have limited public support for a large public social security system (Skocpol and Ikenberry 1983). All these forces played a role in blocking attempts in the first several decades of the twentieth century to

expand protection beyond the veterans' pension and partial state-based programs for mothers' pensions or industrial accident insurance.

The impetus for the passage of old-age and unemployment social security came from the Great Depression. The rapidly expanding costs of private pensions and a crisis of capitalist growth lessened opposition of big business to federal pension legislation and a more general role of the government in the capitalist economy (Jenkins and Brents 1989). Southern congressmen were persuaded to support legislation that excluded agricultural and domestic workers and insisted that means-tested levels for old-age assistance be set at the state level; both factors would limit the disruption of the low-wage southern economy. Popular demands in the early 1930s by several hundred thousand supporters of the Townsend movement for a federal government pension for every citizen over age 60 may have hastened enactment (Williamson et al. 1982). Ultimately, the goal of reducing unemployment by removing older workers from the labor force and supporting at least temporarily those who were unemployed proved crucial in passing the initial legislation in 1935 (Schulz 1988; Graebner 1980).

The original 1935 Social Security Act mandated only limited coverage and benefit levels for old-age retirement. Only 60 percent of the workforce was covered: Agricultural, domestic, and self-employed workers; military personnel; federal, state, and local employees; and employees of nonprofit, tax-exempt organizations were all excluded. Moreover, benefit levels were quite low: Policymakers intended not to replace work income fully or assure the maintenance of workers' preretirement standard of living but instead to supplement private sources of retirement income with minimal public benefits (Achenbaum 1986). Social Security benefits alone would hardly meet what would be considered poverty levels in many states at the time (Quadagno 1984).

The initial structure of the social security system, along with the incremental changes made in the following decades, was for the most part market-conforming. Early debates about the degree to which the program should redistribute income across classes were settled in favor of those who wanted to maintain the connection between contributions and benefits (Cates 1983). Funding from

general revenues was rejected in favor of contribution-based financing, reinforcing the view of the system as an insurance system. Flat-rate benefits were rejected as unsuitable for a nation with such regional and social heterogeneity; instead, benefits would reflect preretirement income levels. A cap placed on taxable wages, which ostensibly concentrated both contributions and benefits for ordinary middle-income wage workers, introduced some regressiveness into the formula. The major exception to this strategy was that benefits for low-wage workers were higher relative to contributions than were those for high-wage workers (Myers 1981). Also, provisions for unemployment benefits, aid to dependent children, and relief for the blind targeted modest benefits for needy groups (Achenbaum 1989). The system thus began as and remains a mixture of social insurance based on contributions and social adequacy based on social need (Munnell 1977).

Expansion of the system began before the first benefits were paid out and continued for several more decades. In 1939, dependents and survivors were made eligible for benefits. Coverage was extended in the 1950s to include most self-employed, domestic, and agricultural workers, and the participation of state and local employees was made elective (federal employees kept their own system until 1984). In 1956, actuarially reduced benefits were made available at ages 62–64 for women, and in 1961 the same option for early retirement was made available to men, an option now exercised by a majority of new beneficiaries. Also in the 1950s, benefits equal to those for retirees were added for disabled persons aged 50–64 and later for disabled workers of all ages. In 1965, Medicare for the elderly and Medicaid for the poor were added to provide protection against the high costs of medical care. Benefit and contribution levels also rose with extensions of coverage and disability. Ad hoc adjustments to benefit levels, which well exceeded inflation (Tomasson 1984), were common until 1972, when benefits were linked to yearly increases in the consumer price index. Payroll taxes and the maximum taxable wage also increased.

The growth of benefits and coverage nonetheless proceeded more quickly than did that of contributions, and by the late 1970s this situation resulted in funding problems. The concept of the accumulation of a reserve was replaced quickly by

a pay-as-you-go system in which current workers paid for current retirees (with enough of a surplus to cover year-to-year fluctuations). In the early years of the system, the ratio of one retiree to 120 workers made this system of funding workable. By the 1970s, the ratio of retirees to workers was one to five. Combined with increasingly high benefit levels, the growing dependency ratio resulted in payments that exceeded contributions. Amendments in 1977 “deliberated” benefits for the first time by, among other things, freezing minimum benefits and making the earnings test more stringent (Tomasson 1984). Far from sufficient to deal with the implications of higher benefits and an older age structure, these changes only delayed a more serious restructuring. A \$17 billion deficit in 1983 made further deliberalization necessary. In 1981, a Reagan administration proposal to lower benefits, change the retirement age, delay cost-of-living increases, and reduce family benefits for dependents and survivors was met with nearly universal opposition. To move the negotiations out of the public eye, where painful and politically unpopular choices could be agreed on, a bipartisan commission was appointed to develop proposals to deal with both short-term and long-term funding problems (Light 1985). The commission offered a compromise plan that was quickly passed by Congress and signed by President Reagan.

To summarize a complex 1983 amendment, a number of major changes were made in the direction the system was to take compared with previous decades. For the first time, Social Security benefits above specified levels were to be taxed. The age of eligibility for full retirement benefits would be extended gradually to 67 beginning in 1999, and payroll taxes would be increased along with the maximum taxable wage base. All these changes have had the desired effect: Contributions now exceed benefits paid. The long-run projection is that the surplus accrued during the next thirty years probably will balance the expected deficit when large baby boom cohorts reach retirement age (Social Security Administration 1989). The surplus, however, is by law used to purchase Treasury bonds, which fund deficits in general revenue spending. Since the bonds will have to be paid off by taxpayers through general income taxes later on, funding problems will not disappear.

The cumulative changes in the system now result in the coverage of over 90 percent of work-

ers, who qualify for benefits by accumulating forty quarters, or ten years, of covered employment. Besides the basic benefit, a minimum benefit is available for those with long-term covered employment at low wages, a dependent’s benefit at 50 percent of the spouse’s benefits is available to spouses, and a survivor’s benefit is available at 100 percent of the deceased spouse’s benefits. Supplemental Security Income (SSI) provides cash assistance—unrelated to contributions and funded from general revenues—for needy aged, disabled, and blind persons who meet the means test. Among the elderly, 38 percent of all income comes from Social Security, and a majority of elderly persons depend on Social Security for more than half their income (Sherman 1987).

The position of the U.S. Social Security system relative to those of other nations depends on how generosity is measured. As a percentage of gross domestic product (GDP), the U.S. systems ranks quite low. Considering pensions alone, however, a measure of the benefit of a new retiree as a percentage of the wage of the average manufacturing worker ranks the United States higher. The United States falls slightly below average for single workers and slightly above average for married workers (Aldrich 1982). Part of the discrepancy stems from the concentration of public spending on pensions in the United States to the neglect of other programs. The family allowance spending and free health care for the nonaged that are common in other advanced industrial democracies are absent altogether in the United States except for need-based public assistance such as Aid to Families with Dependent Children and Medicaid. The United States provides well for those whose contributions during their work lives are high—the average retiree, in other words—but spends less in the aggregate for those who are not covered. Finally, the low percentage of the aged in the United States relative to other advanced industrial nations makes it possible to replace an above-average proportion of preretirement wages while spending a below-average fraction of GDP.

### COMPARATIVE PERSPECTIVES

Many developing nations have begun to implement more formal social security systems, primarily for the benefit of urban workers and civil



servants, but few of those countries have the economic resources needed to provide more than minimal coverage or protection from economic contingencies (Midgley 1984). Comparative studies have concentrated on the historical emergence and current policies of mature welfare states in advanced industrial nations.

Among the high-income democracies, substantial variation exists in spending levels and the structure of benefit distribution. Including pension, health care, occupational injury, unemployment, family allowance, public assistance, and related programs for civil servants and veterans, mean spending as a percentage of GDP in 1980 was 19 percent (International Labour Organization 1985). Nations that spend the most include Sweden (31.2 percent), the Netherlands (27.6 percent), Denmark (26.2 percent), and France (25.5 percent), and the nations that spend the least are Japan (9.8 percent), Italy (11.3 percent), Australia (11.6 percent), and the United States (12.2 percent). As was discussed above, countries also vary in the extent to which they rely on universal benefits relative to insurance or need-based benefits. According to Esping-Andersen (1989), Sweden and Norway in particular have the most equalizing social security programs; Finland, Denmark, Belgium, and the Netherlands also structure benefits on the basis of citizenship rights. The English-speaking nations and Switzerland tend to base their systems most on market-related criteria.

A comparison of the maximum and minimum benefit levels of pensions during the 1980s further illustrates important intercountry differences. In the United States, the difference between the maximum and minimum benefit is \$9,900; in West Germany, it is \$11,000 (Social Security Administration 1985). These figures contrast with those for nations with primarily flat-rate systems, such as Canada (\$500), Denmark (\$1,300), and the Netherlands (\$0). Nations also differ in the frequency of adjustment for the cost of living, the ages of eligibility for early or normal retirement, the degree of retirement required for the receipt of benefits on reaching retirement age (i.e., the existence of a retirement test), and the wage ceiling for social security taxation. Scales summarizing national differences on all these dimensions provide an overview of the divergence in pensions (Day 1978; Myles 1984).

In the 1950s and 1960s, scholars predicted convergence in social security systems as advanced industrial technology spread: The standardizing effects of technology would reduce preexisting cultural and political differences among the economically developed nations. The need for a recently trained, highly educated, and geographically mobile labor force in industrial economies would make older workers superfluous to the production process. Without means of employment, the elderly would depend on government programs for economic support. In this functionalist framework, the state meets the needs of business for a differentiated labor force while simultaneously meeting the financial needs of surplus workers unable to find employment (Wilensky 1975). Hence, retirement and social security grew rapidly among all developed nations, especially in the decades after World War II.

Similar convergence in social security systems is predicted by neo-Marxist theories of monopoly capitalism. Here the focus is on the requirements of the capitalist mode of production and the power of the capitalist elite. State-sponsored insurance subsidizes the costs of the production of capital, and state-sponsored social assistance helps maintain the legitimacy of the political and economic system in the face of discontent among the superfluous population (O'Connor 1973). The standardizing force is therefore the needs of increasingly monopolized capital to maintain high profit and investment, but the consequence is still the expansion of the state in similar forms among advanced industrial nations. Partisan democratic politics play a minimal role in either the industrialist or the capitalist logic.

The fact that in contrast to the predictions of convergence theories, expenditure levels have continued to diverge across nations over the last several decades has led more recently to a number of political explanations of variation in social security. The most common explanations focus on the differential political power of labor unions across the advanced industrial democracies. In places where labor is centralized and unions have high membership, labor gains power in negotiation with capital and also can contribute to the election of socialist, social democratic, and labor parties that represent its interests. As a result, social legislation decreasing the scope of the market and emphasizing distribution based on political power

emerges in areas where labor is strong and leftist parties have ruled for significant periods. In places where labor is weaker and more fragmented, rightist parties are more powerful and market-reinforcing programs with low benefits are common. Relatedly, the emergence of corporatist bargaining structures in which officially designated representatives of labor and capital negotiate economic policy with state managers has emerged in some nations—usually small nations with a strong political representation of labor. The corporatist bargain has been for labor to hold down wage demands in return for full employment and generous, redistributive welfare spending (Goldthorpe 1984).

Other theories agree with the importance of political forces in generating divergence but focus on the political activity of the aged as well as on classes (Pampel and Williamson 1989). Even among the advanced industrial nations, substantial differences in the percentage of the aged exist and appear to be related to welfare spending through both demographic and political channels. Given the same benefit level in 1980 as in 1960, aging of the population can account for only some of the observed increase in pension spending. However, the size of benefit increases over time correlates closely with the size of the elderly population. Beyond demographic effects, then, the elderly appear, at least in some countries, to be an influential political interest group in supporting higher pension and health care spending.

Others have emphasized the role of the state in divergent social security policies. Beginning with the assumption that public policies cannot be reduced to the demands and preferences of any single social group, state-based theories have examined how the structure of relatively autonomous state agencies and the interests of state managers can shape the way in which demands are expressed and translated into legislation. Qualitative studies have identified, within specific historical and national contexts, the state characteristics important for particular policy outcomes. The quantitative literature, however, has had less success relating state characteristics such as size and centralization to measures of social security spending or citizenship rights.

Any resolution of the theoretical debates and mixed empirical results will come from synthetic efforts at theory building and statistical analysis.

Class, status-demographic, political, productive, and state factors all may prove important for understanding social security system development once theories and models more clearly specify how one set of factors varies with the levels of the others. Efforts to estimate nonlinear, interactive models are under way and should prove crucial for future research (Hicks et al. 1989; Pampel et al. 1990).

## CONSEQUENCES

The huge literature on the consequences of social security spending for social equality and social behavior is beyond the scope of this article. Controversy exists not so much on whether spending has an effect but on the kinds of social phenomena it most affects.

One view is that social security spending directly reduces economic inequality without substantially changing social behavior such as labor force participation, living arrangements, and savings. The major evidence in favor of redistributive consequences comes from studies that subtract transfers from total income and compare inequality with and without those transfers (Smeeding, et al. 1988). In the United States and a number of European nations, pretransfer inequality and poverty are higher than they are for posttransfer income distribution. According to the results of this methodology, expenditures for pensions are particularly egalitarian. However, advocates of this view have been less willing to accept the claim that transfers promote inequality by providing incentives to leave the labor force, in other words, by inducing behavior that indirectly contributes to higher rates of poverty and inequality. Implicitly, unemployment and low income are seen as the result of discrimination and lack of opportunities, situations that do not change with the receipt of benefits.

Other views weigh the behavioral responses to transfers as important relative to the redistributive consequences. If transfers induce labor force and living arrangement changes that make pretransfer income distribution less egalitarian than it would be if transfers were not present, the evidence of redistribution cited above would have to be seen as flawed (Danziger et al. 1981). For instance, pensions have the largest effect in reducing pretransfer inequality but also induce voluntary retirement

that lowers earnings relative to what they would be without pensions or retirement benefits. Similarly, transfers increase an individual's ability to afford independent living arrangements, and this makes pretransfer income figures misleading.

Trends in poverty and inequality do not provide unambiguous evidence for either view. Certainly, the absolute income of the elderly in the United States has risen with the growth of Social Security benefits. As Social Security benefits rose dramatically in the last several decades, poverty among the aged declined from 35 percent (compared with 22 percent for the general population) in 1960 to 12 percent (compared with 13 percent for the general population) by 1987 (U.S. Bureau of the Census 1989). However, the improved economic position of the elderly also stems from the fact that recent cohorts entering old age have been better off financially and more likely to have accumulated private pensions and savings to support themselves than were previous cohorts. For overall income inequality, the trend shows little change (at least until 1980) despite the massive growth of transfers (Levy 1987). Either transfers were not redistributive or pretransfer inequality increased. Perhaps household changes, in part an indirect response to transfers, balanced the direct effects of transfers on inequality (Treas 1983). After 1980, inequality grew, but again, it is difficult to separate the effects of changes in the occupational structure from changes in real Social Security benefits for the poor and unemployed.

Comparative evidence on the relationship between social security spending and inequality across advanced industrial nations is also mixed (compare Pampel and Williamson 1989 with Esping-Andersen 1985). Nations with high spending levels and benefit structures based on citizenship rights, such as those in Scandinavia, have always had lower levels of income inequality among both the aged and the general population. However, it is difficult to establish a causal association between those levels and social security benefits across nations that differ in so many other social and economic characteristics.

Given the mixed empirical evidence, views on the redistributive consequences of the welfare state reflect theoretical assumptions about the determinants of the levels and structure of social security spending. Neo-Marxist theories of monopoly capi-

talism, which assume that high inequality is an inherent and necessary feature of advanced capitalism, argue that social security systems help maintain that structure rather than change it. Interest-group and neopluralist theories see middle-class, politically powerful groups as the primary recipients of most spending, which limits the extent of redistribution to the poor. Other theories claim the opposite. In industrialism theories, spending is directed at the surplus workers who are most in need. In social democratic theories, spending is directed to the working class and the poor represented by leftist parties and unions. Still others claim that the state and institutional context shapes the ability of spending to reduce inequality. As in the study of the determinants of spending, interactive or contextual studies probably will be needed to make sense of the comparative experience.

## ISSUES

A number of issues or problems face policymakers who deal with social security systems. A few of these issues are reviewed briefly below. Some apply especially or primarily to the United States, while others apply to all advanced industrial nations and third world nations.

First, concern has been expressed over the inequitable treatment of women in earnings-related social security systems. When receipt of benefits for women in old age depends on the benefits of their spouses, high rates of marital breakup and widowhood make reliance on this source of financial security risky. When receipt of benefits of women in old age depends on wage contributions, discontinuous labor force participation during the childbearing years penalizes women. Universal benefits provide some support for older women, but other policy options are emerging to deal more directly with the gender-based problems. Some nations give social security contribution credits to women who leave the labor force to raise children or split the earned credits of a couple equally between the spouses. Classification of welfare state regimes needs to consider gender as a component of social rights (Orloff 1993).

Second, the improved economic position of the elderly, declining poverty rates, and higher public benefits in the 1970s and the 1980s stand in contrast to the declining real level of benefits and

increasing poverty among children in the United States. The improved position of the elderly relative to children may stem from the increasing size and political power of elderly cohorts compared to the smaller cohorts of children (Preston 1984). The fact that benefits for children take the form of means-tested social assistance—a type of program that receives weak public support relative to pensions because it is not shared by large parts of the population—also contributes to this inequality. Other nations that have family allowance systems that provide cash benefits to all or nearly all parents have experienced little concern over generational equity.

Third, after decades of expansion, policymakers must face problems in balancing continued demands for more spending with limits on taxation. On the one hand, with the problems of support that still exist among vulnerable groups such as the oldest old, minority group members, and widowed women, more spending is needed. Increasingly expensive health and long-term care for the elderly and disabled add to the cost of social security systems. Despite the cost, support for pension and health care continues to be strong (Coughlin 1980). On the other hand, critics have argued that the rising costs of social security contribute to inflation and unemployment by reducing savings and productivity. Those who are more sociologically oriented suggest that high expenditures tend to weaken community and family bonds, which ultimately are the source of protection for those in need (Glazer 1988). Certainly, concern with high tax levels has led politicians to attempt to control spending and reduce taxes in nearly all advanced industrial democracies. Balancing these goals without resorting to deficit spending will remain the task of governments in the decades to come.

Fourth, concern over population aging relates to debates about controlling the cost of social security. The difficulties in meeting funding demands are likely to worsen in the next century with the entrance of large baby boom cohorts into old age. In part, this is a problem of declining fertility, which reduces the size of younger, working cohorts relative to older, retired cohorts. In the recent past, when the relative sizes of working and elderly cohorts were reversed, social security recipients were treated generously, receiving bene-

fits worth five to six times their contributions and the accrued interest (Wolff 1987). Future retiring cohorts are not likely to experience such high returns on their contributions and are sometimes skeptical of receiving any at all. Still, funding problems for aging populations are not insurmountable. Many European nations, whose fertility levels fell faster than those of the United States, already have aged populations as large as 17 percent of the total—levels that will not be reached for thirty years in the United States. Through appropriate political and economic policies, the United States can meet the needs of its elderly population (Aaron et al. 1989).

Finally, these issues are emerging as important in third-world nations. Although the percentage of the aged in those nations is small and social security remains primarily a family rather than a state responsibility, that situation can change quickly. Rapid declines in fertility sharply increase the percentage of the aged, make family care for the elderly difficult, and generate demands for public support. With scarce resources, the state may risk being overwhelmed by these demands. Public understanding of the process of building social security systems in those nations remains meager.

### RECENT TRENDS

If they follow the trends of recent years, the first decades of the twenty-first century will see demands for both stability and change in social security programs in the United States. In terms of stability, major changes in the system have been difficult to legislate. In the late 1980s, Congress passed legislation to add coverage for catastrophic health care and prescriptions to Medicare by increasing taxes on the affluent elderly. The funding mechanism avoided cross-generational taxes on workers but concentrated the costs of the new benefits on a relatively small part of the aged population that already had supplemental private health care coverage. The vocal opposition of those paying the higher taxes for the new provisions led Congress to rescind the legislation shortly after the new program began. Concentrating the costs of expanded social security programs on a small group of beneficiaries did not prove successful; sharing the costs among persons of all ages

and among all the recipients appears crucial to the success of any changes in social security programs.

Another major social security initiative to improve health care in the early 1990s also failed. After entering office in 1993, President William Clinton proposed a form of national health care that represented the most substantial change in American social security protection since the 1960s. The proposal did not advance a single-payer model of national health care such as those used in Canada, the United Kingdom, and many other European nations but aimed to provide universal coverage through a complex system of public and private health care delivery. Reflecting in part the suspicions of citizens of a huge change in government's role in health care as well as resistance to the change from the health and medical care professions, Congress rejected the proposed legislation.

Just as efforts to expand social security programs have failed, so have efforts to cut benefits substantially. Proposals by Republicans to control Medicare, Medicaid, and Social Security costs have met with acute resistance that has essentially blocked legislation. These examples indicate the desires of citizens for stability in most social insurance programs. Social assistance programs have, however, undergone major changes. With bipartisan political support, Aid to Families with Dependent Children has been renamed Temporary Assistance to Needy Families and now contains a work requirement for the continued receipt of benefits. The lack of change in social insurance programs and the major change in social assistance programs reflect the broad-based support for the former relative to the latter. Means-tested programs have never had the public approval enjoyed by insurance programs (Marmor et al. 1990).

In terms of demands for change, much concern remains about the long-term future of social security programs even as citizens resist short-term changes. As the baby boom generation approaches retirement age and the expected future deficit in the Social Security Fund comes closer, worries about funding have reemerged. Some economists predict serious problems. Thurow (1996, p. 46) states, "Already the needs and demands of the elderly have shaken the social welfare state, causing it for all practical purposes to go

broke." Even if it is not broke, Social Security probably will provide returns on contributions to future generations of retirees that do not reach today's high levels. According to Kotlikoff (1992), those age 25 in 1989 will pay \$193,000 more in taxes than they will receive in government benefits over their lifetimes; in contrast, those age 75 in 1989 will receive \$42,00 more in benefits than they pay in taxes. Despite the uncertain assumptions about the future contained in these forecasts, they present a disconcerting picture. As a result of these sorts of claims, polls show that young workers doubt that Social Security will even exist when they retire (Kingson and Berkowitz 1993, p. 87).

Given future funding concerns, recent federal budget surpluses have generated a desire to "save" Social Security. Ironically, the surplus results in large part from Social Security revenues that exceed current payments and thus mask deficit spending in other parts of the government, yet the surplus has produced debate about how to proceed. Some want to return the surplus to taxpayers in the form of tax cuts that, under the assumptions of supply-side economics, will generate economic growth and make it easier to support the large baby boom population of retirees in the decades to come. Others want to use the surplus to invest in Social Security by paying off current government debt that in future decades would aggravate the problem of funding retirement benefits. No legislation has passed yet, and policies change quickly with economic and political circumstances. The consensus seems to be to use part of the federal surplus for debt reduction toward the goal of Social Security solvency. However, unexpected military costs, such as those for the Serbian-Kosovar conflict, can reduce the surplus and eliminate its use for Social Security.

An alternative approach to improving Social Security solvency in the twenty-first century is to privatize contributions. An extreme version of privatization would follow the lead of Chile in allowing individual workers to invest their contributions in private accounts that fund their own retirement. The shift from a pay-as-you-go program to a funded program would involve an enormous change in the nature of old-age benefits that would create numerous risks (Williamson 1997). A less extreme version recommended by a recent commission on Social Security would allow the

government to invest some contributions in stock funds. The stunning upward movement in the stock market in the mid-1990s brought enormous returns to those with private pension investments and highlighted the low returns provided by the current system of using surplus contributions to buy government bonds. Controversy over government involvement in the stock market has, however, slowed action on the commission's recommendations.

Economists have taken the initiative in making policy recommendations for social security, while sociologists have aimed more to defend the current system against attacks. More than sociologists, economists tend to view the low rates of return on old-age contributions and the work disincentive effects of social assistance with alarm. Sociologists, in contrast, highlight the threats of privatization to social equality and universalism in public benefits (Minkler and Estes 1991), the government's role in social protection (Quadagno 1996), and the widespread sense of generational solidarity that citizens share in their attitude toward Social Security (Bengtson and Achenbaum 1993). Their contribution will continue to come from studies of the consequences of varied social security policies across the high-income nations for social equality (Esping-Andersen 1990; Korpi and Palme 1998), generational relations (Cohen 1993; Marmor et al. 1994; Myles and Quadagno 1991; Pampel 1994), and economic well-being (Rainwater and Rein 1993; Crystal and Waehrer 1996).

(SEE ALSO: *Government Regulation; Public Policy Analysis; Retirement; Social Gerontology*)

## REFERENCES

- Aaron, Henry J., Barry P. Bosworth, and Gary Burtless 1989 *Can America Afford to Grow Old? Paying for Social Security*. Washington D.C.: Brookings Institution.
- Achenbaum, W. Andrew 1986 *Social Security: Visions and Revisions*. Cambridge, UK: Cambridge University Press.
- 1989 "Public Pensions as Intergenerational Transfers in the United States." In Paul Johnson, Christoph Conrad, and David Thomson, eds., *Workers versus Pensioners: Intergenerational Justice in an Aging World*. Manchester, UK: Manchester University Press.
- Aldrich, Jonathan 1982 "The Earnings Replacement Rate of Old-Age Benefits in Twelve Countries, 1969–80." *Social Security Bulletin* 44(12):3–11.
- Bengtson, Vern L., and W. Andrew Achenbaum, eds. 1993 *The Changing Contract across Generations*. New York: Aldine de Gruyter.
- Cates, Jerry R. 1983 *Insuring Inequality: Administrative Leadership in Social Security, 1935–53*. Ann Arbor: University of Michigan Press.
- Cohen, Lee M., ed. 1993 *Justice across Generations: What Does it Mean?* Washington DC: Public Policy Institute, American Association of Retired Persons.
- Coughlin, Richard M. 1980 *Ideology, Public Opinion, and Welfare Policy*. Berkeley: Institute of International Studies, University of California.
- Crystal, Steven, and K. Waehrer, 1996 "Later-Life Economic Inequality in Longitudinal Perspective." *Journal of Gerontology: Social Sciences* 51B:S307–S318.
- Danziger, Sheldon, Robert H. Haverman, and Robert Plotnick 1981 "How Income Transfer Programs Affect Work, Savings, and the Income Distribution: A Critical Review." *Journal of Economic Literature* 19:975–1028.
- Day, Lincoln 1978 "Government Pensions for the Aged in Nineteen Industrialized Countries." *Comparative Studies in Sociology* 1:217–234.
- Esping-Andersen, Gosta 1985 "Power and Distributional Regimes." *Politics and Society* 14:222–255.
- 1989 "The Three Political Economies of the Welfare State." *Canadian Review of Sociology and Anthropology* 26:10–35.
- 1990 *The Three Worlds of Welfare Capitalism*. Princeton, N.J. Princeton University Press.
- Glazer, Nathan 1988 *The Limits of Social Policy*. Cambridge, Mass.: Harvard University Press.
- Goldthorpe, John H. 1984 "The End of Convergence: Corporatist and Dualist Tendencies in Modern Western Societies." In John H. Goldthorpe, ed., *Order and Conflict in Contemporary Capitalism*. Oxford, UK: Clarendon.
- Graebner, William 1980 *A History of Retirement*. New Haven, Conn: Yale University Press.
- Hicks, Alexander, Duane Swank, and Martin Ambuhl 1989 "Welfare Expansion Revisited: Policy Routines and Their Mediation by Party, Class, and Crisis, 1959–1982." *European Journal of Political Research* 4:401–430.
- International Labour Organization 1985 *The Cost of Social Security*. Geneva: International Labour Organization.
- Jenkins, J. Craig, and Barbara G. Brents 1989 "Social Protest, Hegemonic Competition, and Social Reform: A Political Struggle Interpretation of the Ori-

- gins of the American Welfare State." *American Sociological Review* 54:891-909.
- Kingson, Eric R., and Edward D. Berkowitz 1993 *Social Security and Medicare: A Policy Primer*. Westport, Conn.: Auburn House.
- Korpi, Walter 1989 "Power, Politics, and State Autonomy in the Development of Social Citizenship: Social Rights during Sickness in Eighteen Countries since 1930." *American Sociological Review* 54:309-328.
- , and Joachim Palme 1998 "The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries." *American Sociological Review* 63:661-687.
- Kotlikoff, Laurence J. 1992 *Generational Accounting: Knowing Who Pays, and When, for What We Spend*. New York: Free Press.
- Levy, Frank 1987 *Dollars and Dreams: The Changing American Income Distribution*. New York: Russell Sage Foundation.
- Light, Paul 1985 *Artful Work: The Politics of Social Security Reform*. New York: Random House.
- Marmor, Theodore R., Jerry L. Mashaw, and Philip L. Harvey 1990 *America's Misunderstood Welfare State: Persistent Myths, Enduring Realities*. New York: Basic Books.
- , Timothy M. Smeeding, and Vernon L. Greene, eds. 1994 *Economic Security and Intergenerational Justice: A Look at North America*. Washington D.C.: Urban Institute.
- Marshall, T. H. 1964 *Class, Citizenship, and Social Development*. Chicago: University of Chicago Press.
- Midgley, James 1984 *Social Security, Inequality, and the Third World*. New York: Wiley.
- Minkler, Meredith, and Carroll L. Estes, eds. 1991 *Critical Perspectives on Aging: The Political and Moral Economy of Growing Old*. Amityville N.Y.: Baywood.
- Munnell, Alice H. 1977 *The Future of Social Security*. Washington D.C.: Brookings Institution.
- Myers, Robert J. 1981 *Social Security*, Rev. ed. Homewood, Ill.: Irwin.
- Myles, John 1984 *Old Age in the Welfare State: The Political Economy of Public Pensions*. Boston: Little, Brown.
- , and Jill Quadagno, eds. 1991. *States, Labor Markets, and the Future of Old Age Policy*. Philadelphia: Temple University Press.
- O'Connor, James 1973 *The Fiscal Crisis of State*. New York: St. Martin's.
- Organization for Economic Cooperation and Development 1988 *The Future of Social Protection*. OECD Social Policy Studies No. 6. Paris: Organization for Economic Cooperation and Development.
- Orloff, Ann Shola 1993 "Gender and the Social Rights of Citizenship: The Comparative Analysis of Gender Relations and Welfare States." *American Sociological Review* 58:303-328.
- , and Theda Skocpol 1984 "Why Not Equal Protection? Explaining the Politics of Public Social Spending in Britain, 1900-1911, and the United States, 1880-1920s." *American Sociological Review* 49:725-750.
- Pampel, Fred C. 1994 "Population Aging, Class Context, and Age Inequality in Public Spending." *American Journal of Sociology* 100:153-195.
- , and John B. Williamson 1989 *Age, Class, Politics, and the Welfare State*. Cambridge, UK: Cambridge University Press.
- , —, and Robin Stryker 1990 "Class Context and Pension Response to Demographic Structure." *Social Problems* 37:535-550.
- Preston, Samuel H. 1984 "Children and the Elderly: Divergent Paths for America's Dependents." *Demography* 21:435-457.
- Quadagno, Jill S. 1984 "Welfare Capitalism and the Social Security Act of 1935." *American Sociological Review* 49:632-647.
- 1988 *The Transformation of Old Age Security: Class and Politics in the American Welfare State*. Chicago: University of Chicago Press.
- 1996 "Social Security and the Myth of the Entitlement Crisis." *The Gerontologist* 36:391-399.
- Rainwater, Lee, and Martin Rein 1993 "Comparing Economic Well-Being of Older Men in Six Countries." In A. B. Atkinson and Martin Rein, eds., *Age, Work and Social Security*. New York: St. Martins.
- Rimlinger, Gaston 1971 *Welfare Policy and Industrialization in Europe, America, and Russia*. Toronto: Wiley.
- Schulz, James 1988 *The Economics of Aging*. Dover, Mass.: Auburn House.
- Sherman, Sally R. 1987 "Fast Facts and Figures about Social Security." *Social Security Bulletin* 50(5):5-25
- Simmons, Leo 1960 "Aging in Preindustrial Societies." In Clark Tibbits, ed., *Handbook of Social Gerontology*. Chicago: University of Chicago Press.
- Skocpol, Theda, and John Ikenberry 1983 "The Political Formation of the American Welfare State in Historical and Comparative Perspective." *Comparative Social Research* 6:87-148.
- Smeeding, Timothy, Barbara Boyle Torrey, and Martin Rein 1988 "Patterns of Income and Poverty: The Economic Status of Children and the Elderly in Eight Countries." In John L. Palmer, Timothy Smeeding,

- and Barbara Boyle Torrey, eds., *The Vulnerable*. Washington D.C.: Brookings Institution.
- Stephens, John D. 1979 *The Transformation from Capitalism to Socialism*. London: Macmillan.
- Social Security Administration 1985 *Social Security throughout the World*. Washington D.C. U.S. Government Printing Office.
- 1989 “Actuarial Status of the OASI and DI Trust Funds.” *Social Security Bulletin* 52(6):2–7.
- Thurow, Lester C. 1996 *The Future of Capitalism*. New York: William Morrow.
- Tomasson, Richard F. 1984 “Government Old Age Pensions under Affluence and Austerity: West Germany, Sweden, the Netherlands, and the United States.” *Research in Social Problems and Public Policy* 3:217–272.
- Treas, Judith 1983 “Trickle-Down or Transfers? Postwar Determinants of Family Income Inequality.” *American Sociological Review* 48:546–559.
- U.S. Bureau of the Census 1989 *Statistical Abstract*. Washington D.C.: U.S. Government Printing Office.
- Wilensky, Harold 1975 *The Welfare State and Equality*. Berkeley: University of California Press.
- Williamson, John B. 1997 “A Critique of the Case for Privatizing Social Security.” *The Gerontologist* 37:561–571.
- , Linda Evans, and Lawrence Powell 1982 *The Politics of Aging*. Springfield, Ill.: Charles C Thomas.
- Wolff, Nancy 1987 *Income Redistribution and the Social Security Program*. Ann Arbor: University of Michigan Press.

FRED C. PAMPEL

## SOCIAL STRATIFICATION

In all complex societies, the total stock of valued resources is distributed unequally, with the most privileged individuals and families receiving a disproportionate share of power, prestige, and other valued resources. The term “stratification system” refers to the constellation of social institutions that generate observed inequalities of this sort. The key components of such systems are (1) the institutional processes that define certain types of goods as valuable and desirable, (2) the rules of allocation that distribute those goods across various positions or occupations (e.g., doctor, farmer, “housewife”), and (3) the mobility mechanisms that link individuals to positions and generate

unequal control over valued resources. The inequality of modern systems is thus produced by two conceptually distinct types of “matching” processes: The jobs, occupations, and social roles in society are first matched to “reward packages” of unequal value, and the individual members of society then are allocated to the positions defined and rewarded in that manner.

There are, of course, many types of rewards that come to be attached to social roles (see Table 1). The very complexity of modern reward systems arguably suggests a multidimensional approach to understanding stratification in which analysts specify the distribution of each of the valued goods listed in Table 1. Although some scholars have advocated a multidimensional approach of this sort, most have opted to characterize stratification systems in terms of discrete classes or strata whose members are similarly advantaged or disadvantaged with respect to various assets (e.g., property and prestige) that are deemed fundamental. In the most extreme versions of this approach, the resulting classes are assumed to be real entities that predate the distribution of rewards, and many scholars therefore refer to the “effects” of class on the rewards that class members control (see the following section for details).

The goal of stratification research has thus devolved to describing the structure of these social classes and specifying the processes by which they are generated and maintained. The following types of questions are central to the field:

1. What are the major forms of class inequality in human history? Is such inequality an inevitable feature of human life?
2. How many social classes are there? What are the principal “fault lines” or social cleavages that define the class structure? Are those cleavages strengthening or weakening with the transition to advanced industrialism?
3. How frequently do individuals cross occupational or class boundaries? Are educational degrees, social contacts, and “individual luck” important forces in matching individuals to jobs and class positions? What other types of social or institutional forces underlie occupational attainment and allocation?



**Types of Assets, Resources, and Valued Goods Underlying Stratification Systems**

Asset Group	Selected Examples	Relevant
1. Economic	Ownership of land, farms, factories, professional practices, businesses, liquid assets, humans (i.e., slaves), labor power (e.g., serfs)	Karl Marx, Erik Wright
2. Political	Household authority (e.g., head of household); workplace authority (e.g., manager); party and societal authority (e.g., legislator); charismatic leader	Max Weber, Ralf Dahrendorf
3. Cultural	High-status consumption practices; "good manners"; privileged lifestyle	Pierre Bourdieu, Paul DiMaggio
4. Social	Access to high-status social networks, social ties, associations and clubs, union memberships	W. Lloyd Warner, James Coleman
5. Honorific	Prestige; "good reputation"; fame; deference and derogation; ethnic and religious purity	Edward Shils, Donald Treiman
6. Civil	Rights of property, contract, franchise, and membership in elective assemblies; freedom of association and speech	T H. Marshall, Rogers Brubaker
7. Human	Skills; expertise; on-the-job training; experience; formal education; knowledge	Kaare Svalastoga, Gary Becker

Table 1

4. What types of social processes and state policies maintain or alter racial, ethnic, and sex discrimination in labor markets? Have these forms of discrimination been weakened or strengthened in the transition to advanced industrialism?
5. Will stratification systems take on new and distinctive forms in the future? Are the stratification systems of modern societies gradually shedding their distinctive features and converging toward a common (i.e., postindustrial) regime?

These questions all adopt a critical orientation to human stratification systems that is distinctively modern in its underpinnings. For the greater part of history, the existing stratification order was regarded as an immutable feature of society, and the implicit objective of commentators was to explain or justify that order in terms of religious or quasi-religious doctrines. During with the Enlightenment, critical "rhetoric of equality" gradually emerged and took hold, and the civil and legal advantages of the aristocracy and other privileged status groupings were accordingly challenged. After these advantages were largely eliminated in the eighteenth and nineteenth centuries, that egalitarian ideal was extended and recast to encompass not only such civil assets voting rights but also economic assets in the form of land, property, and the means of production. In its most radical form,

this economic egalitarianism led to Marxist interpretations of human history, and it ultimately provided the intellectual underpinnings for socialism. While much of stratification theory has been formulated in reaction against these early forms of Marxist scholarship, the field shares with Marxism a distinctively modern (i.e., Enlightenment) orientation that is based on the premise that individuals are "ultimately morally equal" (see Meyer 1994, p. 733; see also Tawney 1931). This premise implies that issues of inequality are critical in evaluating the legitimacy of modern social systems.

### BASIC CONCEPTS

The five questions outlined above cannot be addressed adequately without first defining some of the core concepts in the field. The following definitions are especially relevant:

1. The degree of *inequality* in a given reward or asset (e.g., civil rights) depends on its dispersion or concentration across the individuals in the population. Although many scholars attempt to capture the overall level of societal inequality in a single parameter, such attempts obviously are compromised insofar as some types of rewards are distributed more equally than others are.

2. The *rigidity* of a stratification system is characterized by the continuity over time in the social standing of its members. If the current wealth, power, or prestige of individuals can be predicted accurately on the basis of their prior statuses or those of their parents, then there is much class reproduction and the stratification system is accordingly said to be rigid.
3. The process of stratification is *ascriptive* to the extent that traits present at birth (e.g., sex, race, ethnicity, parental wealth, nationality) influence the subsequent social standing of individuals. In modern societies, ascription of all kinds usually is seen as undesirable or “discriminatory,” and much state policy is therefore directed toward fashioning a stratification system in which individuals acquire resources solely by means of their own achievements.
4. The degree of *status crystallization* is characterized by the correlations among the assets listed in Table 1. If these correlations are strong, the same individuals (i.e., the “upper class”) will consistently appear at the top of all status hierarchies, while other individuals (i.e., the “lower class”) will consistently appear at the bottom of the stratification system.

These four variables can be used to characterize differences across societies in the underlying structure of stratification. As the discussion below reveals, there is great cross-societal variability not only in the types of inequality that serve as the dominant stratifying forces but also in the extent of such inequality and the processes by which it is generated, maintained, and reduced.

### FORMS OF STRATIFICATION

It is useful to begin with the purely descriptive task of classifying the various types of stratification systems that have appeared in past and present societies. Although the staple of modern classification efforts has been the tripartite distinction between class, caste, and estate (e.g., Svalastoga 1965), there is also a long tradition of Marxian typological work that introduces the additional categories of primitive communism, slave society, and socialism (Marx [1939] 1971; Wright 1985). As is shown in

Table 2, these conventional approaches are largely but not entirely complementary, and it is therefore possible to fashion a hybrid classification that incorporates most of the standard distinctions (see Kerbo 1991 for related work).

For each of the stratification forms listed in Table 2, it is conventionally assumed that certain types of assets emerge as the dominant stratifying forces (see column 2) constitute the major axis around which social classes, strata, or status groupings are organized (see column 3). If this assumption holds, the rigidity of stratification systems can be indexed by the amount of class persistence (see column 5) and the degree of crystallization can be indexed by the correlation between class membership and each of the assets listed in Table 1 (see column 6). The final column in Table 2 rests on the further assumption that stratification systems have reasonably coherent ideologies that legitimate the rules and criteria by which individuals are allocated to positions in the class structure (see column 7).

The first panel in Table 2 pertains to the “primitive” tribal systems that dominated human society from the beginning of human evolution until the Neolithic revolution of 10,000 years ago. Although tribal societies assumed various forms, the total size of the distributable surplus was in all cases limited, and this cap on the surplus placed corresponding limits on the overall level of economic inequality. Also, customs such as gift exchange, food sharing, and exogamy were practiced commonly in tribal societies and had some redistributive effects. In fact, many observers (e.g., Marx [1939] 1971) treated these societies as examples of “primitive communism,” since the means of production (e.g., tools and land) were owned collectively and other types of property were distributed evenly among tribal members. This does not mean that a perfect equality prevailed; after all, the more powerful medicine men (i.e., shamans) often secured a disproportionate share of resources, and the tribal chief could exert considerable influence on political decisions. However, these residual forms of power and privilege were never inherited directly and typically were not allocated in accordance with such simple ascriptive traits as ethnicity, race, or clan. The main pathway to political office or high status and prestige was through superior skills in hunting, magic, or leadership (see Lenski 1966 for further details). While

SOCIAL STRATIFICATION

<b>Basic Parameters of Stratification for Eight Ideal-Typical Systems</b>						
<b>System</b>	<b>Principal Assets</b>	<b>Major Strata or Classes</b>	<b>Inequality</b>	<b>Rigidity</b>	<b>Crystallization</b>	<b>Justifying Ideology</b>
<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>	<b>(7)</b>
A. Hunting and gathering society						
1. Tribalism	Human (hunting and magic skills)	Chiefs, shamans, and other tribe members	Low	Low	High	Meritocratic selection
B. Horticultural and agrarian society						
2. Asiatic mode	Political (i.e., incumbency of state office)	Officeholders and peasants	High	Medium	High	Tradition and religious doctrine
3. Feudalism	Economic (land and labor power)	Nobility, clergy, and commoners	High	Medium-high	High	Tradition and Roman Catholic doctrine
4. Slavery	Economic (human property)	Slave owners, slaves, "free men"	High	Medium-high	High	Doctrine of natural and social inferiority (of slaves)
5. Caste society	Honorific and cultural (ethnic purity and "pure" lifestyles)	Castes and subcastes	High	High	High	Tradition and Hindu religious doctrine
C. Industrial society						
6. Class system	Economic (means of production)	Capitalists and workers	Medium-high	Medium	High	Classical liberalism
7. State socialism	Political (party and workplace authority)	Managers and managed	Low-medium	Low-medium	High	Marxism and Leninism
8. "Advanced" industrialism	Human (i.e., education, expertise)	Skill-based occupational groupings	Medium	Low-medium	Medium	Classical liberalism

Table 2

meritocratic forms of allocation often are seen as prototypically modern, they were present in an incipient form at very early stages of societal development.

With the emergence of agrarian forms of production, the economic surplus became large enough to support more complex systems of stratification. The "Asiatic mode," which some commentators regard as a precursor of advanced agrarianism, is characterized by (1) the absence of strong legal institutions recognizing private property rights (with village life taking on a correspondingly communal character), (2) a state elite that extracts the surplus agricultural production through rents or taxes and expends it on "defense, opulent living, and the construction of public works" (Shaw 1978, p. 127), and (3) a constant flux in elite personnel

resulting from "wars of dynastic succession and wars of conquest by nomadic warrior tribes" (O'Leary 1989, p. 18). This mode thus provides the conventional example of how a "dictatorship of officialdom" can flourish in the absence of private property and a well-developed proprietary class. The parallel with modern socialism looms so large that various scholars have suggested that Marx downplayed the Asian case for fear of exposing it as a "parable for socialism" (Gouldner 1980, pp. 324-352).

Whereas the institution of private property was underdeveloped in the East, the ruling class under Western feudalism was very much a propertied one. The distinctive feature of feudalism was the institution of personal bondage; that is, the nobility not only owned large estates, farms, or

manors but also held legal title to the labor power of its serfs (Table 2, line B3). If a serf fled to the city, this was considered a form of theft: the serf was stealing the portion of his or her labor power owned by the lord (Wright 1985, p. 78). As such, the statuses of serf and slave differ only in degree, with slavery constituting the “limiting case” in which workers lose all control over their own labor power (line B4). At the same time, it would be a mistake to reify this distinction, since the history of agrarian Europe reveals “almost infinite gradations of subordination” (Bloch 1961, p. 256) that blur the conventional dividing lines between slavery, serfdom, and freedom. While the slavery of Roman society provides the best example of complete subordination, some slaves in the early feudal period were bestowed with “rights” of real consequence (e.g., the right to sell surplus product), and some nominally free men were obliged to provide rents or services to a manorial lord. The social classes that emerged under European agrarianism thus were structured in quite diverse ways, but in all cases rights of property ownership were firmly established and the life chances of individuals were defined largely by their control over property in its differing forms. Unlike the ideal-typical Asiatic case, the nation-state was peripheral to the feudal stratification system, since the means of production (i.e., land and humans) were controlled by a proprietary class that emerged independently of the state.

The historical record shows that agrarian stratification systems were not always based on strictly hereditary forms of inequality (Table 2, panel B, column 5). The case of European feudalism is especially instructive in this regard, since it suggests that stratification systems often become more rigid as the underlying institutional forms mature and take shape (Mosca 1939; Kelley 1981). Although it is well known that feudalism after the twelfth century (i.e., “classical feudalism”) was characterized by a “rigid stratification of social classes” (Bloch 1961, p. 325), the feudal structure appears to have been more permeable in the period before the institutionalization of the manorial system and the associated transformation of the nobility into a legal class. In this transitional period, access to the nobility was not legally restricted to the offspring of nobility and marriage across classes or estates was not prohibited, at least not formally. The case of ancient Greece provides

a complementary example of a relatively open agrarian society. As Finley (1960) and others have noted, the condition of slavery was heritable under Greek law, yet manumission (the freeing of slaves) was so common that the slave class had to be replenished constantly with new captives secured through war or piracy.

The most extreme examples of hereditary closure are found in caste societies (Table 2, line B5). While some scholars have argued that American slavery had “caste-like features” (Berreman 1981), it is Hindu India which clearly provides the defining case of caste organization. The Indian caste system is based on (1) a hierarchy of status groupings (i.e., castes) that are ranked by ethnic purity, wealth, and access to goods or services, (2) a corresponding set of “closure rules” that forbid all forms of intercaste marriage or mobility and thus make caste membership both hereditary and permanent, (3) a high degree of physical and occupational segregation enforced by elaborate rules and rituals governing intercaste contact, and (4) a justifying ideology (Hinduism) that successfully induces the population to regard such extreme forms of inequality as legitimate and appropriate. What makes this system distinctive is not only its well-developed closure rules but also the fundamentally honorific (and noneconomic) character of the underlying social hierarchy. As is indicated in Table 2, the castes of India are ranked on a continuum of ethnic and ritual purity, with the highest positions in the system reserved for castes that prohibit behaviors that are seen as dishonorable or “polluting.” In some circumstances, castes that acquired political and economic power eventually advanced in the status hierarchy, yet they usually did so after mimicking the behaviors and lifestyles of higher castes.

The defining feature of the industrial era (Table 2, panel C) has been the emergence of egalitarian ideologies and the consequent “delegitimation” of the extreme forms of stratification found in caste, feudal, and slave systems. This can be seen in the European revolutions of the eighteenth and nineteenth centuries that pitted the egalitarian ideals of the Enlightenment against the privileges of rank and the political power of the nobility. In the end, these struggles eliminated the last residue of feudal privilege, but they also made new types of inequality and stratification possible. Under the class system that ultimately emerged (line C6), the

estates of the feudal era were replaced by purely economic groups (i.e., “classes”) and the old closure rules based on hereditary principles were supplanted by formally meritocratic processes. The resulting classes were neither legal entities nor closed status groupings; consequently, the emergent class-based inequalities could be represented and justified as the natural outcome of economic competition among individuals with differing abilities, motivation, or moral character (i.e., “classical liberalism”). This class structure had such a clear “economic base” (Kerbo 1991, p. 23) that Marx ([1894] 1972) of course defined classes in terms of their relationship to the means of economic production. The precise contours of the industrial class structure are nonetheless a matter of continuing debate; for example, a simple (“vulgar”) Marxian model focuses on the cleavage between capitalists and workers, whereas more refined Marxian and neo-Marxian models identify additional intervening or “contradictory” classes (Wright 1985) and other (non-Marxian) approaches represent the class structure as a continuous gradation of socioeconomic status or “monetary wealth and income” (Mayer and Buckley 1970, p. 15).

Regardless of the relative merits of these models, the ideology underlying the socialist revolutions of the nineteenth and twentieth centuries was explicitly Marxist. The intellectual heritage of these revolutions and their legitimating ideologies ultimately can be traced to the Enlightenment; that is, the egalitarianism of the Enlightenment was still very much in force, but now it was deployed against the economic power of the capitalist class rather than against the status and honorific privileges of the nobility. The evidence from eastern Europe and elsewhere suggests that these egalitarian ideals were only partially realized. In the immediate postrevolutionary period, factories and farms were collectivized or socialized and fiscal and economic reforms were instituted expressly to reduce income inequality and wage differentials among manual and nonmanual workers. Although these egalitarian policies subsequently were weakened or reversed through the reform efforts of Stalin and others, this does not mean that inequality on the scale of prerevolutionary society was ever reestablished among rank-and-file workers. At the same time, it has long been argued that the socialization of productive forces did not have the intended effect of empowering workers,

since the capitalist class was replaced by a “new class” of party officials and managers who continued to control the means of production and allocate the resulting social surplus. This class has been variously identified with intellectuals or the intelligentsia (e.g., Gouldner 1979), bureaucrats or managers (e.g., Rizzi 1985), and party officials or appointees (e.g., Djilas 1965). Whatever the formulation adopted, the assumption is that the working class ultimately lost out in contemporary socialist revolutions, just as it did in the so-called bourgeois revolutions of the eighteenth and nineteenth centuries.

Whereas the means of production were socialized in the revolutions in eastern Europe and the former Soviet Union, the capitalist class remained largely intact throughout the process of industrialization in the West. The old propertied class may, however, be weakening in the West and the East alike as a postindustrial service economy diffuses and technical expertise emerges as a “new form of property” (Berg 1973, p. 183). It follows that human and cultural capital may be replacing economic capital as the principal stratifying force in advanced industrial society (Table 2, line C8). According to Gouldner (1979) and others (e.g., Galbraith 1967), a dominant class of cultural elites is emerging in the West, much as the transition to state socialism allegedly generated a new class of intellectuals in the East. This does not mean that all theorists of advanced industrialism posit a grand divide between the cultural elite and a working mass. In fact, some commentators (e.g., Dahrendorf 1959, pp. 48–57) have argued that skill-based cleavages are crystallizing throughout the occupational structure, resulting in a continuous gradation or hierarchy of socioeconomic classes. In nearly all models of advanced industrial society, it is further assumed that education is the principal mechanism by which individuals are sorted into such classes, and educational institutions thus serve in this context to “license” human capital and convert it to cultural currency.

### SOURCES OF STRATIFICATION

Although the preceding sketch indicates that a wide range of stratification systems emerged over the course of human history, it remains unclear whether some form of stratification or inequality is an inevitable feature of human society. In ad-

addressing this question, it is useful to consider the functionalist theory of Davis and Moore (1945), which is the best-known attempt to understand “the universal necessity which calls forth stratification in any system” (p. 242). The starting point for any functionalist approach is the premise that all societies must devise some means to motivate the best workers to fill the most important and difficult occupations. This “motivational problem” can be addressed in a variety of ways, but the simplest solution may be to construct a hierarchy of rewards (e.g., prestige, property, power) that privileges the incumbents of functionally significant positions. As noted by Davis and Moore (1945, p. 243), this amounts to setting up a system of institutionalized inequality (i.e., a “stratification system”), with the occupational structure serving as a conduit through which unequal rewards and perquisites are disbursed. The stratification system therefore may be seen as an “unconsciously evolved device by which societies insure that the important positions are conscientiously filled by the most qualified persons” (Davis and Moore 1945, p. 243). Under the Davis–Moore formulation, it is claimed that some form of inequality is needed to allocate labor efficiently, but no effort is made to specify how much inequality is sufficient for this purpose. The extreme forms of stratification found in existing societies may well exceed the “minimum . . . necessary to maintain a complex division of labor” (Wrong 1959, p. 774).

The Davis–Moore hypothesis has come under criticism from several quarters. The prevailing view among postwar commentators is that the original hypothesis cannot adequately account for inequalities in “stabilized societies where statuses are ascribed” (Wesolowski 1962, p. 31). Whenever vacancies in the occupational structure are allocated on purely hereditary grounds, there is no need to attend to the “motivational problems” that Davis and Moore (1945) emphasized, and one cannot reasonably argue that the reward system is serving its putative function of matching qualified workers to important positions. It must be recognized, however, that a purely hereditary system is rarely achieved in practice; in fact, even in the most rigid caste societies, talented and qualified individuals typically have some opportunities for upward mobility. Under the Davis–Moore formulation (1945), this slow trickle of mobility is regarded as so essential to the functioning of the

social system that elaborate systems of inequality have evidently been devised to ensure that the trickle continues. Although the Davis–Moore hypothesis therefore can be used to explain stratification in societies with some mobility, the original hypothesis becomes wholly untenable in societies with complete closure (if such societies could be found).

The functionalist approach also has been criticized for neglecting the “power element” in stratification systems. It has long been argued that Davis and Moore (1945) failed “to observe that incumbents [of functionally important positions] have the power not only to insist on payment of expected rewards but to demand even larger ones” (Wrong 1959, p. 774). In this regard, the stratification system may be seen as self-reproducing: The holders of important positions can use their power to influence the distribution of resources and preserve or extend their own privileges. It would be difficult, for instance, to account fully for the advantages of feudal lords without referring to their ability to enforce their claims through moral, legal, and economic sanctions. The distribution of rewards thus reflects not only the “latent needs” of the larger society but also the balance of power among competing groups and their members.

Whereas the early debates addressed conceptual issues of this kind, subsequent researchers shifted their emphasis to constructing “critical tests” of the Davis–Moore hypothesis. This research effort continued throughout the 1970s, with some commentators reporting evidence consistent with functionalist theorizing (e.g., Cullen and Novick 1979) and others providing less sympathetic assessments (e.g., Broom and Cushing 1977). The 1980s was a period of relative quiescence, but Lenski (1994) recently reopened the debate by suggesting that “many of the internal, systemic problems of Marxist societies were the result of inadequate motivational arrangements” (p. 57). That is, Lenski argues that the socialist commitment to wage leveling made it difficult to recruit and motivate highly skilled workers, while the “visible hand” of the socialist economy could never be calibrated to mimic adequately the natural incentive of capitalist profit taking. These results lead Lenski to conclude that “successful incentive systems involve . . . motivating the best qualified people to seek the most important positions” (p. 59). It remains to be seen whether this reading of

the socialist “experiments in destratification” (Lenski 1978) will generate a new round of functionalist theorizing and debate.

### THE STRUCTURE OF MODERN STRATIFICATION

The recent history of stratification theory is in large part a history of debates about the contours of class, status, and prestige hierarchies in advanced industrial societies. These debates may appear to be nothing more than academic infighting, but among the participants they are treated as a “necessary prelude to the conduct of political strategy” (Parkin 1979, p. 16). For instance, considerable energy has been devoted to drawing the correct dividing line between the working class and the bourgeoisie, since the task of identifying the oppressed class is seen as a prerequisite to devising a political strategy that might appeal to it. In such mapmaking efforts, political and intellectual goals are often conflated and debates in the field are accordingly infused with more than the usual amount of scholarly contention. While these debates are complex and wide-ranging, it will suffice here to distinguish between four major schools of thought.

**Marxists and Post-Marxists.** The debates within the Marxist and neo-Marxist camps have been especially contentious not only as a result of such political motivations but also because the discussion of class in *Capital* (Marx [1894] 1972) is too fragmentary and unsystematic to adjudicate between competing interpretations. At the end of the third volume of *Capital*, one finds the famous fragment on “the classes” (Marx [1894] 1972, pp. 862–863), but this discussion breaks off at the point where Marx appeared to be ready to advance a formal definition of the term. It is clear, nonetheless, that his abstract model of capitalism was resolutely dichotomous, with the conflict between capitalists and workers constituting the driving force behind further social development. This simple two-class model should be viewed as an ideal type designed to capture the developmental tendencies of capitalism; after all, whenever Marx carried out concrete analyses of existing capitalist systems, he acknowledged that the class structure was complicated by the persistence of transitional classes (i.e., landowners), quasi-class groupings (e.g., peasants), and class fragments (e.g., the lumpen

proletariat). It was only with the progressive maturation of capitalism that Marx expected these complications to disappear as the “centrifugal forces of class struggle and crisis flung all *dritte Personen* [third persons] to one camp or the other” (Parkin 1979, p. 16).

The recent history of modern capitalism suggests that the class structure has not evolved in such a precise and tidy fashion. As Dahrendorf (1959) points out, the old middle class of artisans and shopkeepers has declined in relative size, yet a new middle class of managers, professionals, and nonmanual workers has expanded to occupy the vacated space. The last fifty years of neo-Marxist theorizing can be seen as the intellectual fallout from this development, with some commentators attempting to minimize its implications and others putting forward a revised mapping of the class structure that explicitly accommodates the new middle class. In the former camp, the principal tendency is to claim that the lower sectors of the new middle class are in the process of being proletarianized, since “capital subjects [nonmanual labor] . . . to the forms of rationalization characteristic of the capitalist mode of production” (Braverman 1974, p. 408). This line of reasoning suggests that the working class may gradually expand in relative size and therefore regain its earlier power.

At the other end of the continuum, Poulantzas (1974) has argued that most members of the new intermediate stratum fall outside the working class proper, since they are not exploited in the classical Marxian sense (i.e., surplus value is not extracted). This approach may have the merit of keeping the working class conceptually pure, but it reduces its size to “pygmy proportions” (see Parkin 1979, p. 19), and hence dashes the hopes of those who see workers as a viable political force in advanced industrial society. There is, then, much interest in developing class models that fall between the extremes advocated by Braverman (1974) and Poulantzas (1974). For example, the neo-Marxist model proposed by Wright (1978) describes an American working class that is acceptably large (approximately 46 percent of the labor force), yet the class mappings in this model still pay tribute to the various cleavages and divisions among workers who sell their labor power. That is, professionals are placed in a distinct “semi-autonomous class” by virtue of their control over the work process, while upper-level supervisors are located in a “mana-

gerial class” by virtue of their authority over workers (Wright 1978). The dividing lines proposed in this model thus rest on concepts (e.g., autonomy and authority relations) that once were purely the province of Weberian or neo-Weberian sociology. As Parkin (1979) puts it, “inside every neo-Marxist there seems to be a Weberian struggling to get out” (p. 25).

These early class models, which once were quite popular, have been superseded by various second-generation models that rely more explicitly on the concept of exploitation. In effect, Roemer (1988) and others (Wright 1997; Sørensen 1996) have redefined exploitation as the extraction of “rent,” which refers to the excess earnings that are secured by limiting access to positions and thus artificially restricting the supply of qualified labor. If an approach of this sort is adopted, one can test for skill-based exploitation by calculating whether the cumulated lifetime earnings of skilled laborers exceed those of unskilled laborers by an amount larger than the implied training costs (e.g., school tuition and forgone earnings). In a perfectly competitive market, labor will flow to the most rewarding occupations, equalizing the lifetime earnings of workers and eliminating exploitative returns. However, when opportunities are limited by the imposition of restrictions on entry (e.g., qualifying examinations), the equilibrating flow of labor is disrupted and the potential for exploitation emerges. By implication, the working class can no longer be viewed as a wholly cohesive and unitary force, as some workers presumably have an interest in preserving and extending the institutional mechanisms (e.g., schools) that allow them to reap exploitative returns.

**Weberians and Post-Weberians** The rise of the “new middle class” has proved less problematic for scholars working within a Weberian framework. Indeed, the class model advanced by Weber suggests a multiplicity of class cleavages, since it equates the economic class of workers with their “market situation” in the competition for jobs and valued goods (Weber [1922] 1968, pp. 926–40). In this formulation, the class of skilled workers is privileged because its incumbents are in high demand on the labor market and because its economic power can be parlayed into high wages, job security, and an advantaged position in commodity markets (Weber [1922] 1968, pp. 927–928). At the same time, the stratification system is compli-

cated by the existence of “status groupings,” which Weber saw as forms of social affiliation that can compete, coexist, or overlap with class-based groupings. Although an economic class is merely an aggregate of individuals in a similar market situation, a status grouping is defined as a community of individuals who share a style of life and interact as status equals (e.g., the nobility, an ethnic caste). In some circumstances, the boundaries of a status grouping are determined by purely economic criteria, yet Weber notes that “status honor normally stands in sharp opposition to the pretensions of sheer property” (Weber [1922] 1968, p. 932).

The Weberian approach has been elaborated and extended by sociologists attempting to understand the “American form” of stratification. In the postwar decades, American sociologists typically dismissed the Marxist model of class as overly simplistic and one-dimensional, whereas they celebrated the Weberian model as properly distinguishing between the numerous variables Marx had conflated in his approach. These scholars often disaggregated the stratification dimensions identified by Weber into a multiplicity of variables (e.g., income, education, ethnicity) and then showed that the correlations between those variables were weak enough to generate various forms of “status inconsistency” (e.g., a poorly educated millionaire). The resulting picture suggested a “pluralistic model” of stratification; that is, the class system was represented as intrinsically multidimensional, with a host of cross-cutting affiliations producing a complex patchwork of internal class cleavages. While one critic has remarked that the multidimensionalists provided a “sociological portrait of America as drawn by Norman Rockwell” (Parkin 1979, p. 604), some post-Weberians also emphasized the “seamy side” of pluralism. In fact, Lenski (1954) and others (e.g., Lipset 1959) have argued that modern stratification systems can be seen as breeding grounds for personal stress and political radicalism, since individuals with contradictory statuses may feel relatively deprived and thus support “movements designed to alter the political status quo” (Lenski 1966, p. 88). This line of research died out in the early-1970s under the force of negative and inconclusive findings (Jackson and Curtis 1972). Although there has been a resurgence of theorizing about issues of status disparity and relative deprivation (e.g., Baron 1994), much of this work focuses on the generic proper-



ties of all “post modern” stratification systems rather than the allegedly exceptional features of the American case.

In recent years, the standard multidimensionalist interpretation of “Class, Status, and Party” (Weber 1946, pp. 180–95) has fallen into disfavor, and an alternative version of neo-Weberian stratification theory has gradually taken shape. This revised reading of Weber draws on the concept of social closure as defined and discussed in the essay “Open and Closed Relationships” (Weber [1922] 1968, pp. 43–46; 341–348). By social closure, Weber is referring to the processes by which groups devise and enforce rules of membership, typically with the objective of improving the position [of the group] by monopolistic tactics” (Weber [1922] 1968, p. 43). While Weber does not directly link this discussion with his other contributions to stratification theory, later commentators pointed out that social classes and status groupings are generated by simple exclusionary processes operating at the macro-structural level (e.g., Giddens 1973). Under modern industrialism, there are obviously no formal sanctions that prevent labor from crossing class boundaries, yet various institutional forces (e.g., private property, union shops) are quite effective in limiting the amount of class mobility over the life course and between generations. These exclusionary mechanisms not only “maximize claims to rewards and opportunities” among the members of closed classes (Parkin 1979, p. 44), but also provide the demographic continuity needed to generate distinctive class cultures and “reproduce common life experience over the generations” (Giddens 1973, p. 107). As is noted by Giddens (1973, pp. 107–12), barriers of this sort are not the only source of “class structuration,” yet they clearly play a contributing role in the formation of identifiable classes under modern industrialism. This revisionist interpretation of Weber has reoriented the discipline toward examining the causes and sources of class formation rather than the potentially fragmenting effects of cross-cutting affiliations and cleavages.

**Gradational Status Groupings** The theorists discussed above have all proceeded by mapping individuals or families into mutually exclusive and exhaustive categories (“classes”). As this review indicates, there continues to be much debate about the location of the boundaries separating these categories, yet the shared assumption is that fun-

damental class boundaries of some kind are present, if only in a latent or incipient form. By contrast, the implicit claim underlying gradational approaches is that such “dividing lines” are largely the construction of overzealous sociologists and that the underlying structure of modern stratification can be more closely approximated with gradational measures of income, status, or prestige. The standard concepts of class action and consciousness are similarly discarded; that is, whereas most categorical models are based on the (realist) assumption that the constituent categories are “structures of interest that provide the basis for collective action” (Wright 1979, p. 7), gradational models usually are represented as taxonomic or statistical classifications of purely heuristic interest.

This approach has been pursued in various ways, but typically not by operationalizing social standing in terms of income alone. It does not follow that distinctions of income are sociologically uninteresting; in fact, if one is intent on assessing the “market situation” of workers (Weber [1922] 1968), there is much to recommend a direct measurement of their income and wealth. The preferred approach has nonetheless been to define classes as “groups of persons who are members of effective kinship units which, as units, are approximately equally valued” (Parsons 1954, p. 77). This formulation was first operationalized in postwar community studies (e.g., Warner 1949) by constructing broadly defined categories of reputational equals (“upper upper class,” “upper middle class,” etc.). However, when the disciplinary focus shifted to the national stratification system, the measure of choice soon became occupational scales of prestige (e.g., Treiman 1977), socioeconomic status (e.g., Blau and Duncan 1967), or global “success in the labor market” (Jencks et al. 1988). Although there is much debate about the usefulness of such scales, they continue to serve as standard measures of class background in sociological research of all kinds.

## GENERATING STRATIFICATION

The language of stratification theory makes a sharp distinction between the distribution of social rewards (e.g., the income distribution) and the distribution of opportunities for securing those rewards. As sociologists have noted, it is the latter

distribution that governs popular judgments about the legitimacy of stratification: the typical American, for example, is willing to tolerate substantial inequalities in power, wealth, or prestige if the opportunities for securing those social goods are distributed equally across all individuals. Whatever the wisdom of this popular logic, stratification researchers have long explored its factual underpinnings by describing and explaining the structure of mobility chances.

The study of social mobility is, then, a major sociological industry. The relevant literature is vast, yet much of this work can be classified into one of three traditions of scholarship.

1. The conventional starting point has been to analyze bivariate “mobility tables” formed by cross-classifying the occupational origins and destinations of individuals. These tables can be used to estimate the densities of occupational inheritance, describe patterns of mobility and exchange between occupations, and map the social distances between classes and their constituent occupations. Moreover, when comparable mobility tables are assembled for several countries, it becomes possible to address long-standing debates about the underlying contours of cross-national variation in stratification systems (e.g., Erikson and Goldthorpe 1992).
2. It is a sociological truism that Blau and Duncan (1967) and their colleagues (e.g., Sewell et al. 1969) revolutionized the field with their formal “path models” of stratification. These models were intended to represent, if only partially, the process by which background advantages can be converted into socioeconomic status through the mediating variables of schooling, aspirations, and parental encouragement. Under formulations of this kind, the main sociological objective was to show that socioeconomic outcomes are structured not only by “native ability” and family origins but also by various intervening variables (e.g., schooling) that are themselves only partly determined by origins, race and gender, and other ascriptive forces (Blau and Duncan 1967, pp. 199–205). This line of research, which

had fallen out of favor by the mid-1980s, has been rediscovered and revived as stratification scholars react to the now-fashionable argument (i.e., Herrnstein and Murray 1994) that inherited intelligence is increasingly determinative of stratification outcomes.

3. These “status attainment” models have been criticized for failing to attend to the social structural constraints that operate on the stratification process independently of individual-level traits. The structuralist accounts that ultimately emerged from these critiques amounted in most cases to refurbished versions of the dual economy and market segmentation models that were introduced and popularized several decades ago by institutional economists. When these models were redeployed by sociologists in the early 1980s, the usual objective was to demonstrate that women and minorities were disadvantaged not only because of deficient investments in human capital (e.g., inadequate schooling and experience) but also by their consignment to “secondary” labor markets that on average paid lower wages and offered fewer opportunities for promotion or advancement.

These three approaches to stratification analysis typically are implemented with quantitative models of the most sophisticated sort. In a classic critique, Coser (1975) suggested that stratification researchers were so entranced by quantitative models of mobility, attainment, and dissimination that “the methodological tail was wagging the substantive dog” (p. 652). This latter argument can no longer be taken exclusively in the intended pejorative sense because new models and methods have raised important substantive questions that previously had been overlooked. In this sense, the development of structural equation, log-linear, and event history models are properly viewed as watershed events in the history of mobility research.

### ASCRIPTIVE PROCESSES

The forces of race and gender have long been relegated to the sociological sidelines by class theorists of both Marxist and non-Marxist persuasions. In most versions of class-analytic theory,

status groups are treated as secondary forms of affiliation, whereas class-based ties are seen as more fundamental and decisive determinants of social and political action. Although Race and gender have not been ignored altogether in such treatments, they typically are represented as vestiges of traditional loyalties that will wither away under the rationalizing influence of socialism, industrialism, or modernization.

The first step in the intellectual breakdown of this model was the fashioning of a multidimensional approach to stratification. Whereas many class theorists gave theoretical or conceptual priority to the economic dimension of stratification, the early multidimensionalists emphasized that social behavior could be understood only by taking into account all status group memberships (e.g., race, gender) and the complex ways in which they interact with one another and with class outcomes. The class-analytic approach was further undermined by the apparent reemergence of racial, ethnic, and nationalist conflicts in the late postwar period. Far from withering away under the force of industrialism, the bonds of race and ethnicity seemed to be alive and well: The modern world was witnessing a “sudden increase in tendencies by people in many countries and many circumstances to insist on the significance of their group distinctiveness” (Glazer and Moynihan 1975, p. 3). This resurgence of status politics continues today. In one last several decades, ethnic and regional solidarities have intensified with the decline of conventional class politics in central Europe and elsewhere, and gender-based affiliations and loyalties have strengthened as feminist movements diffuse throughout much of the modern world.

This turn of events has led some commentators to proclaim that the factors of race, ethnicity, and gender are now the driving force behind the evolution of stratification systems. In one such formulation, Glazer and Moynihan (1975) conclude that “property relations [formerly] obscured ethnic ones” (p. 16), but now it is “property that begins to seem derivative, and ethnicity that becomes a more fundamental source of stratification” (p. 17). The analogous position favored by some feminists is that “men’s dominance over women is the cornerstone on which all other oppression (class, age, race) rests” (Hartmann 1981, p. 12; see also Firestone 1972). This formulation

begs the question of timing; that is, if the forces of gender or ethnicity are truly primordial, it is natural to ask why they began expressing themselves with real vigor in more recent history. In addressing this issue, Bell (1975) suggests that a trade-off exists between class-based and ethnic forms of solidarity, with the latter strengthening whenever the former weaken. As the conflict between labor and capital is institutionalized (via “trade unionism”), Bell argues, class-based affiliations typically lose their affective content and workers must turn to racial or ethnic ties to provide them with a renewed sense of identification and commitment. It could be argued that gender politics often fill the same “moral vacuum” that this decline in class politics has allegedly generated.

It may be misleading to treat the competition between ascriptive and class-based forces as a sociological horse race that only one of the two principles can ultimately win. In a pluralist society of the American kind, workers can choose an identity appropriate to the situational context; a modern-day worker may behave as “an industrial laborer in the morning, a black in the afternoon, and an American in the evening” (Parkin 1979, p. 34). Although this situational model of status has not been adopted widely in contemporary research, there is some evidence of renewed interest in conceptualizing the diverse affiliations of individuals and the “multiple oppressions” (see Wright 1989, pp. 5–6) that those affiliations engender. It is now fashionable, for example, to assume that the major status groupings in contemporary stratification systems are defined by the intersection of ethnic, gender, and class affiliations (e.g., black working-class women, white middle-class men). The theoretical framework for this approach is not always well articulated, but the implicit claim seems to be that these subgroupings shape the “life chances and experiences” of individuals (Ransford and Miller 1983, p. 46) and thus define the social settings in which subcultures typically emerge. The obvious effect is to invert the traditional post-Weberian perspective on status groupings; that is, whereas orthodox multidimensionalists described the stress experienced by individuals with inconsistent statuses (e.g., poorly educated doctors), the new multidimensionalists emphasize the shared interests and cultures generated within commonly encountered status sets (e.g., black working-class women).

The sociological study of gender, race, and ethnicity has thus burgeoned. As is noted by Lieberman (1994, p. 649), there has been a certain faddishness in the types of research topics that scholars of gender and race have chosen for study, with the resulting body of literature having a correspondingly haphazard and scattered feel. The following research questions have nonetheless emerged as relatively central ones in the field:

1. How are class relations affected by ascriptive forms of stratification? Can capitalists exploit ethnic antagonisms and patriarchy to their advantage? Do male and majority group workers also benefit from stratification by gender and race?
2. What accounts for variability across time and space in ethnic conflict and solidarity? Will ethnic loyalties weaken as modernization diffuses across ethnically diverse populations? Does modernization instead produce a "cultural division of labor" that strengthens communal ties by making ethnicity the principal arbiter of life chances? Is ethnic conflict further intensified when ethnic groups compete for the same niche in the occupational structure?
3. What are the generative forces underlying ethnic, racial, and gender differentials in income and other socioeconomic outcomes? Do those differentials proceed from supply-side variability in the occupational aspirations or the human capital workers bring to the market? Alternatively, are they produced by demand-side forces such as market segmentation, statistical or institutional discrimination, and the seemingly irrational tastes and preferences of employers?
4. Is the underlying structure of ascriptive stratification changing with the transition to advanced industrialism? Does the "logic" of industrialism require universalistic personnel practices and consequent declines in overt discrimination? Can this logic be reconciled with the rise of a modern ghetto underclass, the persistence of massive segregation by sex and race, and the emergence of new forms of poverty and hardship among women and recent immigrants?

These questions make it clear that ethnic, racial, and gender inequalities often are classed together and treated as analytically equivalent forms of ascription. Although Parsons (1951) and others (e.g., Tilly 1998) have emphasized the shared features of "communal ties," such ties can be maintained (or subverted) in very different ways. It has long been argued, for example, that some forms of inequality can be rendered more palatable by the practice of pooling resources (e.g., income) across family members. As Lieberman (1994) points out, the family operates to bind males and females together in a single unit of consumption, whereas extrafamilial institutions (schools, labor markets, etc.) must be relied on to provide the same integrative functions for ethnic groups. If these functions are left wholly unfilled, one might expect ethnic separatist and nationalist movements to emerge. The same "nationalist" option is obviously less viable for single-sex groups; indeed, barring any revolutionary changes in family structure or kinship relations, it seems unlikely that separatist solutions will ever garner much support among men or women. These considerations may account for the absence of a well-developed literature on overt conflict between single-sex groups.

### CONCLUSIONS

In recent years, criticisms of the class-analytic framework have escalated, with many scholars arguing that the concept of class is "ceasing to do any useful work for sociology" (Pahl 1989, p. 710). Although such postmodern accounts have taken many forms, most proceed from the assumption that social classes no longer definitively structure lifestyles and life chances and that "new theories, perhaps more cultural than structural, [are] in order" (Davis 1982, p. 585). In accounts of this sort, the labor movement is represented as a fading enterprise rooted in the old conflicts of the workplace and industrial capitalism, whereas new social movements (e.g., environmentalism) are assumed to provide a more appealing call for collective action by virtue of their emphasis on issues of lifestyle, personal identity, and normative change.

This argument has not been subjected to convincing empirical tests and may prove to be premature. However, even if lifestyles and life chances are truly "decoupling" from economic class, this should not be misunderstood as a more general

decline in stratification per se. The massive facts of economic, political, and honorific inequality will still be operative even if conventional models of class ultimately are found deficient in characterizing the postmodern condition. As is well known, some forms of inequality have increased in recent years (e.g., income inequality), while others show no signs of disappearing or withering away (e.g., political inequality).

This persistence and in some cases deepening of inequality is coupled with the continuing diffusion of antistratification values and a correspondingly heightened sensitivity to all things unequal. As egalitarianism spreads, the postmodern public becomes heavily involved in monitoring and exposing illegitimate (i.e., nonmeritocratic) forms of stratification, and even small departures from equality are increasingly viewed as problematic and intolerable (Meyer 1994). Moreover, because stratification systems are deeply institutionalized, there is good reason to anticipate that demands for egalitarian change will outpace actual changes in stratification practices. These dynamics imply that issues of stratification will continue to generate discord and conflict even in the unlikely event of a long-term trend toward diminishing inequality.

## REFERENCES

- Baron, James N. 1994 "Reflections on Recent Generations of Mobility Research." In David B. Grusk, ed., *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colo.: Westview.
- Bell, Daniel 1975 "Ethnicity and Social Change." In Nathan Glazer and Daniel P. Moynihan, eds., *Ethnicity: Theory and Experience*. Cambridge, Mass.: Harvard University Press.
- Berg, Ivar 1973 *Education and Jobs: The Great Training Robbery*. Harmondsworth, UK: Penguin.
- Bererman, Gerald 1981 *Caste and Other Inequities*. Delhi: Manohar.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Bloch, Marc 1961 *Feudal Society*. London: Routledge and Kegan Paul.
- Braverman, Harry 1974 *Labor and Monopoly Capital*. New York and London: Monthly Review Press.
- Broom, Leonard, and Robert G. Cushing 1977 "A Modest Test of an Immodest Theory: The Functional Theory of Stratification." *American Sociological Review* 42:157-169.
- Coser, Lewis A. 1975 "Presidential Address: Two Methods in Search of a Substance." *American Sociological Review* 40:691-700.
- Cullen, John B., and Shelley M. Novick 1979 "The Davis-Moore Theory of Stratification: A Further Examination and Extension." *American Journal of Sociology* 84:1424-1437.
- Dahrendorf, Ralf 1959 *Class and Class Conflict in Industrial Society*. Stanford, Calif.: Stanford University Press.
- Davis, James 1982 "Achievement Variables and Class Cultures: Family, Schooling, Job, and Forty-Nine Dependent Variables in the Cumulative GSS." *American Sociological Review* 47:69-86.
- Davis, Kingsley, and Wilbert E. Moore 1945 "Some Principles of Stratification." *American Sociological Review* 10:242-249.
- Djilas, Milovan 1965 *The New Class*. New York: Praeger.
- Erikson, Robert, and John H. Goldthorpe 1992 *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford, UK: Clarendon Press.
- Finley, Moses J. 1960 *Slavery in Classical Antiquity*. Cambridge, UK: W. Heffer and Sons.
- Firestone, Shulamith 1972 *The Dialectic of Sex*. New York: Bantam.
- Galbraith, John K. 1967 *The New Industrial State*. Boston: Houghton Mifflin.
- Giddens, Anthony 1973 *The Class Structure of the Advanced Societies*. London: Hutchinson.
- Glazer, Nathan, and Daniel P. Moynihan 1975 "Introduction." In Nathan Glazer and Daniel P. Moynihan, eds., *Ethnicity: Theory and Experience*. Cambridge, Mass.: Harvard University Press.
- Gouldner, Alvin 1979 *The Future of Intellectuals and the Rise of the New Class*. New York: Seabury.
- 1980 *The Two Marxisms: Contradictions and Anomalies in the Development of Theory*. New York: Seabury Press.
- Hartmann, Heidi 1981 "The Unhappy Marriage of Marxism and Feminism: Towards a More Progressive Union." In Lydia Sargent, ed., *Women and Revolution*. Boston: South End Press.
- Herrnstein, Richard J., and Charles Murray 1994 *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Jackson, Elton F., and Richard F. Curtis 1972 "Effects of Vertical Mobility and Status Inconsistency: A Body of Negative Evidence." *American Sociological Review* 37:701-13.
- Jencks, Christopher, Lauri Perman, and Lee Rainwater 1988 "What Is a Good Job? A New Measure of Labor-

- Market Success." *American Journal of Sociology* 93:1322–1357.
- Kelley, Jonathan 1981 *Revolution and the Rebirth of Inequality*. Berkeley: University of California Press.
- Kerbo, Harold R. 1991 *Social Stratification and Inequality: Class Conflict in Historical and Comparative Perspective*. New York: McGraw-Hill.
- Lenski, Gerhard E. 1954 "Status Crystallization: A Non-Vertical Dimension of Social Status." *American Sociological Review* 19:405–413.
- 1966 *Power and Privilege*. New York: McGraw-Hill.
- 1978 "Marxist Experiments in Destratification: An Appraisal." *Social Forces* 57:364–383.
- 1994 "New Light on Old Issues: The Relevance of 'Really Existing Socialist Societies' for Stratification Theory." In David B. Grusky, ed., *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colo.: Westview.
- Liebersohn, Stanley 1994 "Understanding Ascriptive Stratification: Some Issues and Principles." In David B. Grusky, ed., *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colo.: Westview Press.
- Lipset, Seymour M. 1959 *Political Man: The Social Bases of Politics*. Baltimore: Johns Hopkins University Press.
- Marx, Karl (1893) 1971 *The Grundrisse*. New York: Harper & Row.
- (1894) 1972 *Capital*, 3 vols. London: Lawrence and Wishart.
- Mayer, Kurt B., and Walter Buckley 1970 *Class and Society*. New York: Random House.
- Meyer, John W. 1994 "The Evolution of Modern Stratification Systems." In David B. Grusky, ed., *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colo.: Westview Press.
- Mosca, Gaetano 1939 *The Ruling Class*. New York: McGraw-Hill.
- O'Leary, Brendan 1989 *The Asiatic Mode of Production*. Oxford, UK: Basil Blackwell.
- Pahl, R. E. 1989 "Is the Emperor Naked? Some Questions on the Adequacy of Sociological Theory in Urban and Regional Research." *International Journal of Urban and Regional Research* 13:709–720.
- Parkin, Frank 1979 *Marxism and Class Theory: A Bourgeois Critique*. New York: Columbia University Press.
- Parsons, Talcott 1951 *The Social System*. Glencoe, Ill.: Free Press.
- 1954 *Essays in Sociological Theory*. Glencoe, Ill.: Free Press.
- Poulantzas, Nicos 1974 *Classes in Contemporary Capitalism*. London: Verso.
- Ransford, H. Edward, and Jon Miller 1983 "Race, Sex, and Feminist Outlooks." *American Sociological Review* 48:46–59.
- Rizzi, Bruno 1985 *The Bureaucratization of the World*. London: Tavistock.
- Roemer, John 1988 *Free to Lose*. Cambridge, Mass.: Harvard University Press.
- Sewell, William H., Archibald O. Haller, and Alejandro Portes 1969 "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34:82–92.
- Shaw, William H. 1978 *Marx's Theory of History*. Stanford, Calif.: Stanford University Press.
- Sørensen, Aage B. 1996 "The Structural Basis of Social Inequality." *American Journal of Sociology* 101:1333–1365.
- Svalastoga, Kaare 1965 *Social Differentiation*. New York: D. McKay.
- Tawney, R. H. 1931 *Equality*. London: George Allen and Unwin.
- Tilly, Charles 1998 *Durable Inequality*. Berkeley: University of California Press.
- Treiman, Donald J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
- Warner, W. Lloyd. 1949 *Social Class in America*. Chicago: Science Research Associates.
- Weber, Max (1922) 1968 *Economy and Society*. Berkeley: University of California Press.
- 1946 *From Max Weber: Essays in Sociology*, ed. and transl. by Hans Gerth and C. Wright Mills. New York: Oxford University Press.
- Wesolowski, Włodzimierz 1962 "Some Notes on the Functional Theory of Stratification." *Polish Sociological Bulletin* 3–4:28–38.
- Wright, Erik O. 1978 *Class, Crisis, and the State*. London: New Left Books.
- 1979 *Class Structure and Income Determination*. New York: Academic Press.
- 1985 *Classes*. London: Verso.
- 1989 *The Debate on Classes*. London: Verso.
- 1997 *Class Counts: Comparative Studies in Class Analysis*. Cambridge, UK: Cambridge University Press.
- Wrong, Dennis H. 1959 "The Functional Theory of Stratification: Some Neglected Considerations." *American Sociological Review* 24:772–782.

DAVID B. GRUSKY

## SOCIAL STRUCTURE

“Social structure” is a general term for any collective social circumstance that cannot be altered by isolated actions and thus is fixed or given for the individual. It thus provides a context, environment, or fixed backdrop for action. The size of organizations, the distribution of activities in space, shared language, and the distribution of wealth all might be regarded as social structural circumstances that set limits on feasible activities for individuals.

Social structure is objective in the sense that it is the same for everyone and is beyond the capacity for alteration by any individual. Accordingly, social structure often is spoken of in the singular and as a thing apart, as if there were only one from whose effects no one can escape. This usage masks disagreement about the exact extension of the term but reflects the intention of authors to highlight abstract patterns as an inflexible collective circumstance to which individuals must adapt.

Social structure, or the weaker structural regularities, arises because of the prevalence of social routine. Many social patterns change very slowly either through unmotivated inertia, through willful efforts to renew or reproduce them, or as a collective consequence of individual efforts undertaken for independent reasons. An image or picture, such as a map colored by the linguistic practices of the inhabitants of geographic areas, will lose accuracy slowly and often would remain largely accurate after a century or more. Such substantial durability, along with the accompanying slow continuous change, suggests the possibility of regularities or even scientific laws governing the phenomena that underlie the description.

Routines endure and structural regularities persist for at least three general reasons. Social life is subject to physical constraints such as distance. Thus, most people live close to where they work or do both at one place. For related reasons, many persons maintain stable residences. Furthermore, many people need or desire the company or cooperation of representative social types, such as those who share their religious convictions, or particular work skills. Accordingly, one can associate social attributes with geographic maps. This was a central activity of the Chicago school of sociology (Park and Burgess 1924) and gave rise to the

perspective of human ecology (Hawley 1986). The specialization of social types and activities in space is subject to powerful incentives that induce similarity in the face of turnover among individuals. For example, ethnic concentrations result in specialized facilities, such as food shops, that attract replacements that conserve the ethnic character. Such patterns often persist beyond the lifetimes of the people who initiated them.

A second source of routine is limited learning capacity or the complexity of many social activities. Linguistic rules, moral codes, and work skills illustrate social capacities whose acquisition requires considerable time and effort. This socialization often requires extended exposure to others who know the routines well, especially when the delicate skills of interpretation are involved.

The difficulties of acquiring capacity can confound individual wills. Bernstein (1975) described how linguistic conventions acquired in the home reflect the conditions of adult work and render individuals unsuited for occupations that are not similar to those of their parents. In the same way, a New Yorker who wished to speak in Latin would have to make a huge investment in learning a novel linguistic code. However, this would not undo the investment in English by other New Yorkers, and thus Latin would be impractical for directing taxi drivers. Similar reasons impel the adoption of the abrasive social style of New Yorkers by newcomers. The general principle is that most people must adapt to many surrounding ways of doing things because those ways change so slowly.

A third source of structural regularity is laws governing averages. An example is the suicide rates studied by Durkheim. People commit suicide for a variety of personal motives and the act is never repeated by anyone, yet the frequency of the act is fairly stable over time and thus is stably different among different populations. This is the case because variable causes tend to average into stable totals whenever many instances are drawn from constant underlying conditions. Many of the rates that result are sufficiently stable to sustain plans and projections, which in turn can be embedded in routines, even though the underlying activity is very complex in its detailed causation.

The several sources of stable routines underlie the properties that frequently are associated with proposed structural regularities. Structural regu-

larities often are depicted as abstract, enduring, and operative across a large scale of units. These attributes reflect genesis, for many structural regularities ultimately stem from the long historical process of imposing routines that made a large scale feasible. For example, Tilly (1975) shows how the modern European state resulted from parallel decisions by state makers forced by military competition to pursue centralization by reordering the established routines of ordinary people. The history of collective contention (Tilly 1986) can be seen as the efforts of the victims to defend older patterns against intrusions by state agents such as tax collectors and against the vast reorganization of work and fortune implicit in the expansion of capitalism.

Mann (1986) argues that large-scale cooperation rests on enduring patterns of power. For example, shared religious ideology is a form of power because it makes people subject to claims on their activity. Rapid religious change is not infrequently the result of conquest. Once it has been established, religion is often compulsory. Coercion aside, religious conformity can provide insurance against the risks and pains of social isolation. Other large-scale patterns, such as the division of labor, are maintained in the face of considerable shifting by persons among different roles. This often is implemented by powerful actors who are motivated to induce (or coerce) approximate substitutes to fill in for those who withdraw (or die). In such terms, abstract stability, a large scale, and consequent duration often can be seen to be sustained through underlying causal regularities. In human terms, the reproduction of social structure consists of a myriad of modest efforts that sum to a stable result.

This interdependence underlies the transcendence of abstract structural regularity over individual will. Generally, one cannot learn more than locally applicable routines and must rely on others for critical needs. Thus, one assumes that the staff members in the emergency room will not all take the day off. This frees accountants from the necessity of acquiring medical skill to meet their own needs. As a result, the details of actual routines are known only locally, and the only possible knowledge of the overall pattern is coarse or abstract. Even accountants cannot count up the details they must count on. A further implication is that the alternatives to enacting the routines with which

one is familiar are often limited. It requires time on a historical scale to construct such a pattern. That history has happened, and if the conditions that made narrow, specialized learning practical suddenly came unglued most people would be in a terrible fix.

A special case of social routines consists of those worked out with others, who then are often hard to replace. Replacement is generally more troublesome as the duration is longer, such as many kinship bonds, which can be effectively irreplaceable for many adults. These social relations can be mapped as social networks describing the pattern of the links that surround individuals. The analysis of such patterns is not infrequently (or unreasonably) called "structural analysis," though it hardly exhausts the term.

Elaborate routines, especially social ties, are subject to pressures toward isomorphy, which is defined as a common anatomy or structure. An example is the formation of families, which are different in detail but share common features partly in response to common problems that must be solved within a shared environment. Goode (1970) analyzes the sources and consequences of such regularities. In modern societies, assumptions about such features often are written into administrative procedures such as tax codes, which provide further impetus for individuals to adopt a variant of the pattern defined as normal. Changing and varied individual desires are often in conflict with those pressures to cooperate in the reproduction of the supposedly "normal" pattern.

Emergent properties that apply to wholes but not to parts often are attributed to social structure. Some properties, such as size distributions and complexity, do not have direct individual analogues. Others arise because the net result of many partially independent actions can be different from the intentions of individuals. Thus, markets with many participants can experience crashes in value when many people try to sell in anticipation that others are about to do so, producing a result that no one desires. Kindleberger (1980) describes the recurrence of such crises. Routines are executed by fallible humans and are only locally adapted, somewhat independent, and imperfectly flexible. Many properties of the resulting averages or combinations do not follow from the components in any simple sense.



As the preceding analysis suggests, structural visions are various. One unifying theme is an appeal to abstract, extraindividual patterns that change slowly or not at all. A second unifying theme is that those regularities cause or condition many of the choices and behaviors of individuals. A final common theme is less unifying than divisive. Some structural visions are accompanied by claims of centrality. A particular array of simple elements is proclaimed, often on metatheoretical or philosophic grounds, to be the central deep structure whose inevitable unfolding underlies a vast array of surface appearances. Such comprehensive views have inspired competing, incompatible schools of thought on whose behalf a claim is sometimes made to *the* structural vision of society or the human condition.

Most of these structural visions are comprehensive worldviews that require detailed study in their own right. Among the most prominent are those of Marx and Freud, but there have been structuralist movements in nearly every field of social studies. Nearly all proceed from some highly abstract characterization of the human mind, laws of thought, or the human condition. All of social or mental life is viewed as a manifestation of the reproduction of such elements, often unfolding dialectically. This is presented as the inevitable underpinnings of individual or collective biographies. Piaget (1970) has provided an unusually concise description of an interdisciplinary structuralism based on mathematical progress; this description parallels his more famous theory of discontinuous advancement in human cognitive development. Originators and their descendants often delight in such subtle and insightful reductions of familiar patterns to the chosen central supports.

The term “structure” is most commonly employed in sociology without these all-encompassing ambitions. In empirical sociology, especially quantitative studies based on random samples of persons, the term is invoked for varied efforts to use the larger and often more durable features of social life as explanatory factors for individual conduct and outcomes. The most common contrast is with individual-level causes, including attitudes and aspirations. Sometimes attributes such as race, gender, and class are labeled structural to imply that the underlying mechanism is an exter-

nal force imposed on individuals independently of their wills.

The reasoning behind this is not always explicit, but the usage is justifiable. Generally, the factors labeled structural are alternatives among a differentiated array of possibilities to which individuals are confined for substantial periods. “Structure” then refers to the differentiating average conditions in which people live their lives. At least implicitly, such differences correspond to differences in the routines employed to adapt to local conditions as well as to resources that render routines practical. Classifying people by indicators of the local conditions that surround them reflects the opportunities they have for association and hence for processes such as influence, cooperation, and victimization. Some characterizations also correspond to labels, most notably race and gender, and broadly indicate common tendencies in routines of others to which one is likely to be exposed. Such differences are quite stable, impersonal, and hard to evade. Taken together, these differences in conditions contribute to differences in average responses or individual behaviors.

There is some confusion about the nature of such structural causation, which often is framed as an alternative explanation to individual choice. Persuasive force often comes from stories in which the predominant outcome is made to feel inevitable. This is at odds with the normal empirical result of a difference in tendency or proportion. Rules that hold without exception are rare. This should be expected. Structural abstractions mask much detail that varies. The implicit reference is to averages over multiple executions of complex routines. To take an obvious example, racial discrimination involving job applicants is not invariant but occurs often enough to lead to considerable differences.

Structural causes are not literally the antithesis of individual choice. More precisely, they reflect patterns over which individuals have limited control. The binding force of structural regularity is intrinsically probabilistic. People almost invariably have options, and exceptions to regularities are somewhere in reach. However, established structure—ultimately routines acquired over time, bonds developed to particular others, and the meshed ways of doing that result—exerts a frictional tug. Friction is implicit in the pain of for-

gone routines and the time required to work out new ones. Such pains may be amplified when those who benefit from regularities exert their power to maintain them. On any large scale, the path of least resistance consists of acting today nearly the same way as one acted yesterday. By no means does this rule out individual exceptions, resistance, or willful alterations to parts of the overall web, but friction is cumulative. For example, the rupture and replacement of one bond are quite different from the rupturing of all bonds at once. Similarly, any single person may change jobs, although in practice only to a very limited range of alternatives, yet if all jobs were randomly reshuffled one day, nearly all would go undone, for every job would be subject to the incompetence of the "first day on the job." In summary, the frictional forces of social structure do not rule out rare and/or modest exceptions but generally ensure that wholesale, simultaneous exceptions are rare to the vanishing point.

Empirical applications generally draw on fragments of social structure that are taken as conditioning factors for particular outcomes. The larger challenge is to translate the impersonal, durable complexity of stable differences in condition into a formal calculus, or a theory of social structure. Parsons's (1951) extensive analysis of the logic of social systems was an early and seminal attempt. His student Merton, under the banner of "theories of the middle range," provided a more easily applied set of general tools for structural analysis. Several of Merton's students, including Boudon, Blau, and Coleman, further developed formal calculi for social structure that benefit from the use of mathematical tools.

Parsons's complex system begins with the conditions for stabilizing interaction or, in current terms, meshing routines. Parsons characterizes the routines that govern choice as extended chains of logic linking means to ends. At their most abstract, those chains are anchored in ultimate ends, or values. Durable stability results from consensus on the values that are installed in individuals by more or less extended socialization.

In Parsons's view, the logical chains governing decision making are morally potent norms, or rules governing social conduct. The durable web that shapes individual choice is therefore the complex of norms animated by the anchoring ultimate

values. Parsons imposes on this a logical calculus of the different functions necessary for ensuring that the pattern is resistant to shocks that draw it away from equilibrium. A concomitant of this theme of differentiation is complementary specialization in distinct but interdependent expectations bundled into the social roles enacted by different players.

Parsons's calculus of the functional necessities of meshing differentiated normative specifications proved widely compelling but difficult to apply. His presentation is notoriously hard to read. Applications of the scheme usually consisted of classifying normative elements into taxonomies delimiting functional contributions. These qualitative operations were by no means mechanical or easily communicated as a stable procedure that would steer different investigators to identical results. This rendered moot the possibility of generating conclusions from initial conditions through the application of formal tools. In a similar way, while many were inclined to agree that Parsons's system illuminated how a social system governed by a logic over normative rules might work, it was less than evident that concrete social systems had such logical coherence.

Merton's (1968) "theories of the middle range" provided a more readily applicable set of tools. Like Parsons, Merton proposed that the enduring regularities that make up social structure are normatively defined. However, instead of attempting to calculate over extended normative webs, he drew attention to the implications of positions. Thus, he emphasized that roles place individuals in relations with concrete others or that membership in groups, both present and anticipated, provides reference points for calculating comparisons of expectations and outcomes. Unlike Parsons's more elaborate concerns, Merton's lent themselves to the construction and interpretation of surveys and other manageable research projects.

Merton did not assume, as Parsons did, that norms and roles can be divined from an overarching logic. More frequently, he treated contrasting norms as empirical counterparts of lay distinctions among different roles or group memberships. This can be viewed as a central motivation for the common use of the structural concepts outlined above. However, Merton more often used factual (or readily inferred) norms grounded in different

stable positions to highlight dilemmas. Concrete people could be understood as facing practical problems of resolving competing and often contrary normative standards. This strategy of framing the practical problem as the resolution of contrary expectations frequently leads to insight into choices that at first seem senseless or even self-defeating.

Merton's analyses rested on qualitative inferences, often turning on the meaning of norms. One of Merton's students, Boudon (1982), provides formulations in which social structure refers to numerically definite distributions so that the implications of such extraindividual constraints emerge from formal calculations. For example, he posits an array of young persons committed to personal advancement who make investments in education. However, when all do what is individually sensible, the collective result illustrates Merton's unintended consequences. If there is a fixed and therefore scarce supply of desired positions that will go to those who have the most education, many of those who invest will discover that their efforts are frustrated by the simultaneous striving of others. Boudon provides many illustrations of the perverse effects that can obtain when individual motive operates against a backdrop of a fixed system of positions.

Another of Merton's students, Blau (1977), presents a deductive structural theory based on the notion that social structure consists of arrays of positions, which he calls parameters. Blau divides differentiation into two types: among unranked or nominal categories such as religion and among continuous arrays of ranked positions that differ in their amounts of a scarce and valued resource. The distribution of individuals over positions gives rise to numerical properties of whole social structures, including the heterogeneity of nominal differences, inequality among ranks, and consolidation intersection, or the degree of correlation independence of positions on separate dimensions.

Blau's concept of social structure leads to differences in the sizes of collections of individuals occupying different positions. Size in turn strongly conditions the rate of interaction, or social association. More differentiated structures result in higher rates of intergroup association, and Blau argues that this leads to the successful meshing of routines, or social integration. The intersection of

different dimensions, which results in even smaller subgroups defined by multiple positions, also enhances social integration. Conversely, the consolidation of dimensions, homogeneity rather than heterogeneity, diminishes rates of intergroup contact and hence hinders social integration. Inequality emerges as a special case that illustrates Blau's taste for paradoxical results. Greater inequality leads to smaller strata and fosters intergroup relations, but those relations often take the form of interpersonal conflict, including crime (Blau and Blau 1982).

Blau's notions are particularly suitable for research application because his notion of structure more or less directly corresponds to widely used operationalizations such as gender, race, ethnicity, religion, occupational rank, and wealth. Of course, these are social constructs and in some final analysis are defined by norms and other ideal elements. At the same time, they are for most people most of the time subject to slow or even no change. This sustains the usefulness of a numerical calculus that rests on the notion that size is an objective, impersonal, and durable reality.

Coleman (1990) provides one of the most ambitious attempts to specify social structure as a mathematically tractable map of interdependence. He posits actors with rights of control over their own actions and over tangible things desired by others or resources. His actors maximize the achievement of their desires by exchanging their control in return for that which others control. The result in general is an equilibrium in which initial control in conjunction with the desires of others produces differential power. Within this apparatus, Coleman is able to provide a rigorous analysis of the emergence of larger-scale phenomena, including groups, norms, and corporate actors.

Although all these accounts lie along a single path of intellectual descent, there is a major divide with respect to the elemental nature of social structure. For Parsons, it is an interdependent complex of norms. Unfortunately, there does not exist at present any way to formalize or calculate the mutual implications in a web of symbolic elements. Later analysts who have gone much farther in rendering complexity calculable have done so from "hard" assumptions that take social structure from the outset as a set of objective positions (with objective properties) so that size,

distribution, rates of exchange, and so forth, can be treated mathematically.

More recent treatments have built on the rich, although eclectic, tradition of taking relations or social ties that are knitted into networks as fundamental. One point of departure is Granovetter's (1973) observation that weak ties are surprisingly efficacious in securing resources, notably access to better jobs. Weak ties are most likely to form bridges between clusters of interconnected and thus redundant strong ties. Burt (1992) generalizes this, suggesting that "structural holes," or gaps spanned by positions whose ties unite the otherwise disconnected, are a potent source of advantage. He was able to display supporting evidence from contexts as diverse as executives competing for promotion and sectors of an industrial economy. Burt's concepts are derivative in the best sense; that is, they are a conceptual refinement that moves on to novel terrain, building on what he can take as an established view of social structure.

Tilly (1998) has distilled from network concerns a potent challenge to much received thinking about stratification. He proposes that "durable inequality" reverberates from underlying schemas governing how networks are formed. Categorical divisions such as race, gender, nationality, and citizenship are embedded in widely shared, deeply learned propensities for action, or routines. Recurrent organizational problems, such as assigning work and dividing rewards, are most easily and durably resolved when they are consonant with widely shared assumptions about categorical differences. Somewhat like Burt, Tilly focuses less on origins and more on implications. He examines how relational considerations secure inequalities through persistent configurations of exploitation and resource hoarding that are diffused by emulation and ultimately underpinned by adaptation. In this view, social structure is not globally coherent or uniform but is, somewhat like DNA, a complex melange constructed from varying combinations of a few very simple elements.

A noteworthy gap here is that the proponents of formal theory (and those proposing building blocks) tend to posit or assume "hard" properties, giving limited attention to how or why the hypothesized elementary patterns emerged or became predominant. This leaves open issues of variability and interpretive options (or meaning)

that others see as fundamental. Indeed, some authors believe that human judgment is distinctive and that no mechanical analogue or simulation of human society (Habermas 1987) or human cognition, (Penrose 1989) will ever be possible.

In summary, there are no widely accepted sets of notions that capture all the properties that have been seen as fundamental to the concept of social structure. The huge catalogue of demonstrated effects of structural regularities cannot be organized in a tidy way. Enthusiasm for the different attempts to represent the concept in compact terms varies widely. Sufficiently close attention to the details of competing claims could convince one that no shared subject is really at issue. As in the analysis of social structure itself, it is necessary to carefully select the right degree of abstraction and appropriate pattern of highlighting to discern any common pattern in the competing pictures, but there is nevertheless a pattern to be found.

## REFERENCES

- Bernstein, Basil 1975 *Class, Codes, and Control*. New York: Schocken.
- Blau, Peter 1977 *Inequality and Heterogeneity*. New York: Academic Press.
- Blau, Judith, and Peter Blau 1982 "The Cost of Inequality: Metropolitan Structure and Violent Crime." *American Sociological Review* 47:114-129.
- Boudon, Raymond 1982 *The Unintended Consequences of Social Action*. New York: St. Martin's.
- Burt, Ronald S. 1992 *Structural Holes*. Cambridge, Mass.: Harvard University Press.
- Coleman, James S. 1990 *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Goode, William J. 1970 *World Revolution in Family Patterns*. New York: Free Press.
- Granovetter, Mark 1973 "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-1380.
- Habermas, Jurgen 1987 *The Theory of Communicative Action*, vol. 2: *Lifeworld and System: A Critique of Functionalist Reason*. Boston: Beacon Press.
- Hawley, Amos 1986 *Human Ecology*. Chicago: University of Chicago Press.
- Kindleberger, Charles 1980 *Panics, Manias, and Crashes*. New York: Basic Books.
- Mann, Michael 1986 *The Sources of Social Power*, vol. 1: *A History of Social Power from the Beginning to A.D. 1760*. New York: Cambridge University Press.

- Merton, Robert 1968 *Social Theory and Social Structure*. New York: Free Press.
- Park, Robert, and Ernest Burgess 1924 *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.
- Parsons, Talcott 1951 *The Social System*. New York: Free Press.
- Penrose, Roger 1989 *The Emperor's New Mind*. New York: Oxford University Press.
- Piaget, Jean 1970 *Structuralism*. New York: Basic Books.
- Tilly, Charles 1975 *The Formation of Nation States in Western Europe*. Princeton, N.J.: Princeton University Press.
- 1986 *The Contentious French*. Cambridge, Mass.: Harvard University Press.
- 1998 *Durable Inequality*. Berkeley: University of California Press.

STEVEN L. RYTINA

## SOCIAL VALUES AND NORMS

Values and norms are evaluative beliefs that synthesize affective and cognitive elements to orient people to the world in which they live. Their evaluative element makes them unlike existential beliefs, which focus primarily on matters of truth or falsehood, correctness or incorrectness. Their cognitive element makes them unlike motives that can derive from emotions or psychological drives. Values and norms involve cognitive beliefs of approval or disapproval. Although they tend to persist through time and therefore foster continuity in society and human personality, they also are susceptible to change (Moss and Susman 1980; Alwin 1994).

The evaluative criteria represented in values and norms influence the behavior of subject units at multiple levels (e.g., individuals, organizations, and societies) as well as judgments about the behavior of others, which also can influence behavior. For example, values and norms affect the evaluation of individuals as suitable marriage partners and in that way influence marital behavior. Values and norms also affect evaluation of the governing policies and practices of societies and thus have an impact on diplomatic relations and the policies of one society's government toward other societies.

## CONCEPT OF A VALUE

A value is a belief about the desirability of a mode, means, or end of action (Kluckhohn 1951; Schwartz and Bilsky 1987). It indicates the degree to which something is regarded as good versus bad. A value tends to be general rather than specific, transcending particular types of action and situations. As a general evaluative criterion, it is used to assess specific behaviors in specific situations.

The evaluative criteria represented by values derive from conceptions of morality, aesthetics, and achievement. That is, a mode, means, or end of action can be regarded as good or bad for moral, aesthetic, or cognitive reasons and often for a combination of those reasons (Kluckhohn 1951; Parsons and Shils 1951). For example, being considerate of others may be valued positively (i.e., be viewed as desirable or good) for moral reasons, neatness may be valued positively for aesthetic reasons, and intelligence may be valued positively for cognitive reasons. Since the distinguishing characteristic of a value is evaluation as good or bad, a value that has a cognitive basis is a function of cognitive appraisal based on competency and achievement rather than on scientific or utilitarian grounds. For example, the choice of steel rather than iron to construct a building is a decision based on scientific or utilitarian criteria rather than on values.

The concept of a value must be differentiated from other concepts that appear to be similar. One of those concepts is a preference. A value may be thought of as a type of preference, but not all preferences are values. The distinctive characteristic of a value is that it is based on a belief about what is desirable rather than on mere liking. A preference for an equitable rather than inequitable distribution of rewards is a value, but a preference for vanilla rather than chocolate ice cream is not.

The concept of a value also bears some similarity to the concept of an attitude. Some analysts have suggested that a value is a type of attitude (Fishbein and Ajzen 1975; Glenn 1980), but there are differences between the two concepts. An attitude refers to an organization of several beliefs around a specific object or situation, whereas a value refers to a single belief of a specific kind: a belief about desirability that is based in conceptions of morality, aesthetics, or achievement and transcends specific behaviors and situations. Be-

cause of its generality, a value occupies a more central and hierarchically important place in human personality and cognitive structure than does an attitude. It is a determinant of attitudes as well as behavior. Thus, evaluations of numerous attitude objects and situations are based on a relatively small number of values. Not all attitudes, however, derive from values. For example, an attitude toward skiing may be based on the extent to which that sport is found to be enjoyable rather than on a value. The concept of a value also differs from the concept of an interest in much the same way that it differs from the concept of an attitude, since an interest is a type of attitude that results in the directing of one's attention and action toward a focal object or situation. As is true of attitudes more broadly, some interests derive from values but others do not.

The concept of a value also can be distinguished from the related concept of a motive. The basic property of a motive is the ability to induce valences (incentives) that may be positive or negative. A value has a motive property, involving a predisposition to act in a certain way, because it affects the evaluation of the expected consequences of an action and therefore the choice among possible alternatives; however, it is a less person-centered concept than a motive, which also encompasses emotions and drives. A value is a particular type of motive involving a belief about the desirability of an action that derives from an evaluation of that action's expected consequences in a situation. A value is a distinctively human motive, unlike motives that operate at both the human and the infrahuman levels.

A value also differs from a need. Although both function as motives because of their ability to induce valences, a need is distinctive in being a requirement for the continued performance of an activity and the attainment of other valued outcomes (Emerson 1987). Some needs have a biological basis; others are psychological, often deriving from the persistent frustration of important goals. Although a value may arise from a need, becoming a cognitive transformation of that need, not all needs are transformed into values and not all values derive from needs. Needs also may derive from the structure of a situation, having a social or economic basis rather than a person-centered biological or psychological basis. For example, a need for income may cause an actor to

behave in ways that conflict with his or her values. A need differs from a value in that the continued functioning of the actor and the acquisitions of other valued outcomes are contingent on its being met. A need also differs from a value in that it implies a deficit that imposes a requirement, whereas a value implies motivation that is based on a belief about desirability.

Finally, a value can be differentiated from a goal. A value sometimes is thought of as a goal because goals are selected on the basis of values. However, some values focus on modes of action that are personal attributes, such as intelligence, rather than ends of action, or goals. Values are not goals of behavior. They are evaluative criteria that are used to select goals and appraise the implications of action.

### CONCEPT OF A NORM

Like a value, a norm is an evaluative belief. Whereas a value is a belief about the *desirability* of behavior, a norm is a belief about the *acceptability* of behavior (Gibbs 1965; Marini 1984). A norm indicates the degree to which a behavior is regarded as right versus wrong, allowable versus unallowable. It is an evaluative criterion that specifies a rule of behavior, indicating what a behavior *ought* to be or *ought not* to be. A *prescriptive* norm indicates what should be done, and a *proscriptive* norm indicates what should not be done. Because a norm is a behavioral rule, it produces a feeling of obligation. A value, in contrast, produces a feeling of desirability, of attraction or repulsion.

A norm also differs from a value in its degree of specificity. A norm is less general than a value because it indicates what should or should not be done in particular behavioral contexts. Whereas a value is a general evaluative criterion that transcends particular types of action and situations, a norm is linked directly to particular types of action and situations. For example, there may be a norm proscribing the killing of other human beings that is generally applicable except in situations such as war, self-defense, capital punishment, and euthanasia. Situational variability of this type sometimes is referred to as the *conditionality* of a norm. A norm, like a value, is generally applicable to the *types* of action and situations on which it focuses, but it is less general than a value because it is less

likely to transcend particular types of action and situations.

Because norms often derive from values, they have their basis in conceptions of morality, aesthetics, and achievement and often in a combination of those conceptions. The basis of a norm tends to affect its strength, or the importance attached to it. For example, a norm based in morality that differentiates right from wrong is likely to be considered more important than a norm based in aesthetics that differentiates the appropriate from the inappropriate, for example, in matters of dress or etiquette. A norm, however, differs from a custom in much the same way that a value differs from a preference. A norm involves an evaluation of what an actor *should* do, whereas a custom involves an expectation of what an actor *will* do. It may be expected, for example, that people will drink coffee, but it is usually a matter of indifference whether they do. Drinking coffee is therefore a custom, not a norm; it is not based on a belief about what people ought to do.

### THE STRUCTURE OF VALUES AND NORMS

Multiple values and norms are organized and linked in the cultures of human social systems and also are linked when they are internalized by individuals. Cultural “value orientations” organize and link values and norms to existential beliefs in general views that also might be called worldviews or ideologies (Kluckhohn 1951). They are sets of linked propositions embracing evaluative and existential elements that describe preferred or obligatory states. Values and norms are linked to and buttressed by existential beliefs about human nature, the human condition, interpersonal relations, the functioning of social organizations and societies, and the nature of the world. Since existential beliefs focus on what is true versus untrue, they are to some degree empirically based and verifiable.

In most of the early conceptual and theoretical work on values, values and norms were not differentiated clearly. Later, particularly as attempts to measure values and norms were made, the two concepts were routinely considered distinct, and studies focusing on them have been carried out separately since that time. As a result, the relationship between values and norms rarely has been analyzed theoretically or empirically.

Values and norms are closely related because values usually provide the justification for norms. As beliefs about what is desirable and undesirable, values often are associated with normative beliefs that require or preclude certain behavior, establishing boundaries to indicate what is acceptable versus unacceptable. For example, the positive value attached to human safety and security is supported by norms that proscribe doing harm to other persons and their property. Not all values are supported by norms, however. Displaying personal competence in a variety of ways is positively valued, but norms do not always require it. Similarly, not all norms support values. For example, norms in regard to dress and etiquette can be quite arbitrary. Their existence may support values, but the specific rules of behavior they establish may not.

Many cultural value orientations organize and link the values and norms that operate as evaluative criteria in human social systems. These orientations are learned and internalized by individuals in unique ways that vary with an individual’s personal characteristics and social history and the interaction between the two. Cultural value orientations and internalized individual value orientations are more comprehensive systems of values and norms than those activated as influences on particular types of behavior. The latent structure of values and norms that characterizes a social system or an individual can be thought of as a map or blueprint (Rokeach 1973). Only a portion of the map or blueprint that is immediately relevant to the behavioral choices being made is consulted, and the rest is ignored temporarily. Different subsets of values and norms that make up different portions of the map or blueprint are activated when different types of behavioral choices are made. For example, the values and norms relevant in choosing a mate differ from those relevant in deciding how to allocate one’s time among various activities.

**The Object Unit.** A characteristic of values and norms that is important for understanding their structure is the type of *object unit* to which they pertain, such as an individual, an organization, or a society. Values and norms establish what is desirable or acceptable for particular types of object units. For example, physical and psychological health are positively valued ends of action for individuals, and norms that proscribe or prescribe action to maintain or promote health gov-

ern individual action. Democracy, distributive justice, and world peace are positively valued ends of action for societies, and norms, usually in the form of laws, proscribe and prescribe certain actions on the part of a society's institutions in support of those values. Individuals may value democracy, justice, and peace, but these are societal values, not individual values, since they pertain to the characteristics of societies, not to those of individuals. Differentiating values by their object units is important in conceptualizing and measuring values relevant to the explanation of behavior because correspondence between the actor, or subject unit, and the object unit determines the extent to which behavior by the actor is relevant to achieving a particular end. Individuals differentiate between personal and societal values because they do not have direct influence over social values, thus distinguishing their beliefs on the basis of whether they think those beliefs will lead to action (Braithwaite and Law 1985).

**The Basis of Evaluation.** As evaluative criteria, values and norms have the ability to induce valences (incentives). They affect evaluation of the behavior of others and involve a predisposition to act in a certain way because they affect the evaluation of the expected consequences of action. The evaluation that occurs on the basis of values and norms derives from two structural properties: the polarity, or directionality, of the value or norm and the standard of comparison that is used.

*Polarity.* The polarity of a value or norm is the direction of its valence, or motive force, which may be positive or negative. In the case of a value, something that is evaluated as desirable will have a positive valence, whereas something that is evaluated as undesirable will have a negative valence. In the case of a norm, something that should be done will have a positive valence, whereas something that should not be done will have a negative valence.

*Standard of Comparison.* A value or norm also is characterized by a standard, or level, of aspiration or expectation. This evaluative standard is a reference point with respect to which a behavior and its consequences are evaluated. A subject unit's own action and that of others, as well as the ends that result or may result from action, are evaluated on the basis of whether they are above or below an evaluative standard.

In the case of a value, the evaluative standard determines the neutral point on the value scale at or above which a behavior or its consequences will be evaluated as desirable and below which a behavior or its consequences will be evaluated as undesirable. In both economics and psychology, it has been recognized that there is a utility, or value, function that should be considered nonlinear (Marini [1992] provides a discussion of these developments), and there is empirical evidence that it generally is appropriate to assume the existence of a reference point on a utility, or value, scale. This reference point plays a critical role in producing a nonlinear relationship between the value scale and the objective continuum of behavior and its consequences. It has been observed that value functions change significantly at a certain point, which is often, although not always, zero. In the prospect theory of Kahneman and Tversky (1979), outcomes are expressed as positive or negative deviations from a neutral reference outcome that is assigned a value of zero. Kahneman and Tversky propose an S-shaped value function that is concave above the reference point and convex below it but less steep above than below. This function specifies that the effect of a marginal change decreases with the distance from the reference point in either direction but that the response to outcomes below the reference point is more extreme than is the response to outcomes above it. The asymmetry of the value function suggests a stronger aversion to what is evaluated as undesirable, an asymmetry that is consistent with an empirically observed aversion to loss.

In the case of a norm, the evaluative standard is set by what is defined to be acceptable versus unacceptable. It is a level of expectation that is determined by the specific behaviors that are regarded as right versus wrong, appropriate versus inappropriate. An important difference between a value and a norm is that whereas there is a continuous, nonlinear relationship between a value scale and the objective continuum of behavior or its consequences above the neutral point set by the evaluative standard, this relationship is not expected between the scale of evaluation based on a normative criterion and the objective continuum of behavior. Because a normative standard establishes a boundary of acceptability or requirement that applies to all those covered by the norm, compliance with a normative expectation is not



evaluated as a continuous variable on the basis of variation in behavior above the reference point set by the normative expectation. However, violation of a normative standard is evaluated as a continuous variable on the basis of variation in behavior below the reference point set by the standard. Negative deviations from the standard are likely to be evaluated in much the same way as are negative evaluations from the reference point on a value scale, which is convex below the reference point. Because of the strong aversion to what is evaluated as being below the reference standard, behavior that violates a normative standard is likely to be eliminated from consideration as an option.

The level of aspiration or expectation that operates as an evaluative standard for an actor is socially determined to a large degree. It is a "comparison level" learned from others whom the actor takes as referents. As a result of variation in the characteristics of actors, the social environments to which they are exposed, and the interaction between those two factors, the evaluative standards associated with values and norms vary across actors. Even among actors in the same social environment, the evaluative standard is specific to the actor, although there may be a high degree of consensus about it in a social group.

The evaluative standards associated with values and norms are subject to change in an individual actor. An important source of change is experience that affects the level of ability, knowledge, or accomplishment of an actor. For example, the evaluative standard for achievement values is affected by an actor's level of achievement. There is evidence that people tend to raise their value standards with success and lower them with failure. Thus, as a worker learns a job, that worker's ability to perform the job increases, as does the worker's evaluative standard. A level of ability that once was aspired to and evaluated as "extremely good" may, after increases in the worker's ability, come to be viewed as "mediocre" and below the worker's current evaluative standard for expected performance. Experience also may affect the evaluative standard for norms. For example, there is evidence that the experience of divorce changes normative beliefs about divorce in the direction of increasing its acceptability (Thornton 1985). Another source of change in the evaluative standards associated with the values and norms of an actor is an increase in knowledge of the world that alters

the existential beliefs connected with values and norms.

The evaluative standards associated with values and norms vary not only among actors and over time for the same actor but also with the characteristics of other actors whose behavior is the object of evaluation. These characteristics may differentiate among actors or among the circumstances of the same actor at different times. For example, the value standard used by an adult to evaluate a child's knowledge will vary for children who have completed different amounts of schooling, such as an elementary school student, a high school student, or a college student: The amount of knowledge evaluated as "very good" for an elementary school student will differ from that evaluated as "very good" for a student at a more advanced stage of schooling. Different value standards will be applied to different students and to the same student at different stages of schooling. Similarly, in a work organization, the value standard used to evaluate performance may vary for different categories of workers: Those with more experience may be evaluated according to a higher standard. Again, these different standards may be applied to different workers who are in different categories or to the same worker as he or she progresses from one category to another.

Like a value standard, a normative standard may vary with the characteristics of other actors whose behavior is an object of evaluation. However, there is a difference between a value and a norm in this regard. Because a value is a continuous variable, variation in the value standard with the characteristics of the other actors whose behavior is being evaluated need not have implications for whether the value applies to those actors. In contrast, because a norm is a discrete variable that differentiates what is acceptable from what is unacceptable, variation in the evaluative standard of a norm with the characteristics of other actors whose behavior is being evaluated determines whether the norm applies to other actors with particular characteristics. This variability—that is, variability in whether a value or norm *applies* based on the characteristics of the actors being evaluated—is a dimension of the importance of a value or norm and is labeled its conditionality.

**Dimensions of Importance.** It is commonly recognized that values and norms differ in their

priority, or importance, and that those differences are another aspect of the structure of values and norms. Differences in priority produce a structure that is to some degree hierarchical. Recognition that not all values are of equal importance has led to the use of ranking procedures to measure values (Allport et al. 1960; Rokeach 1973). These procedures have been criticized for forcing respondents to represent their values in a ranked order that does not allow for the possibility that some values may be of equal importance (Alwin and Krosnick 1985; Braithwaite and Law 1985). Although there is a hierarchy among values, there may be sets of values that occupy the same position in the hierarchy. The priority of a value or norm not only has implications for its influence on behavior but also may have implications for the probability that it will change, since values and norms of high priority have been argued to be less likely to change than are those of low priority.

The priority, or importance, of a value or norm can be assessed on a number of dimensions: (1) strength, or intensity, (2) centrality, (3) range, (4) conditionally, and (5) intent. Although these dimensions are conceptually different, they are likely to overlap empirically to a considerable degree. The extent to which and ways in which they overlap in reflecting the importance of a value or norm are not known.

*Strength.* The *strength* of a value or norm can be defined as the maximum strength of the force field it can induce. The strength of the valence reflects its hierarchical position in the latent map or blueprint that characterizes the structure of values and norms for a social system or an individual. Although the strength of a value or norm is likely to display considerable stability, it is also subject to change. At the level of the social system, it may change as a result of long-term changes in social organization and aspects of culture as well as precipitating events. As the social system changes, socializing influences on individuals change. Changes in the values and norms of individuals occur both over the life course (Glenn 1980; Alwin 1994) and as a result of differences between those who are born and move through life in different historical periods. The motivational force of a value at a particular time, however, is not necessarily the maximum strength of its latent force field, because attaining a valued outcome may reduce the subjective utility of additional units of that outcome as a

result of diminishing marginal utility, or satiation. In the case of either a value or a norm, whether one attains an outcome also may alter the maximum strength of its latent force field. For example, if attainment is problematic, the importance of a value or norm may decline as a way of reducing cognitive dissonance.

*Centrality.* The *centrality* of a value or norm can be defined as the number and variety of behaviors or ends to which it applies. Because a central value or norm contributes more than does a peripheral one to the coherent organization and functioning of the total system, the disappearance of a central value or norm would make a greater difference to the total system than would the disappearance of a peripheral value or norm. A central value or norm is more resistant to change than is a peripheral value or norm; however, if change occurs, the more central the value or norm changed, the more widespread its repercussions (Rokeach 1973, 1985).

For individuals and even for social groups, concern and responsibility for the well-being of others is a central value that pertains to a large number and variety of specific behaviors and ends. It is supported by a central proscriptive norm that one should not harm others and a central prescriptive norm that one should help others, particularly if they are in need. These norms pertain to a large number and variety of specific behaviors. In contrast, excitement and adventure are more peripheral values, affecting a smaller number and variety of specific behaviors and ends. In connection with these values, peripheral norms govern the carrying out of specific types of activities that may be sources of excitement and adventure, such as the rules governing sports and potentially dangerous recreational activities.

For individuals, life values that pertain to the overall ends, or goals, of life along with the norms that support them tend to be more central than are the values and norms that pertain to particular life domains or social roles. Part of the reason for this is that life values affect whether particular life domains or social roles are entered into and the amounts of time and energy a person spends in different domains and roles. They also affect an individual's domain- and role-specific values and norms. For example, life values include things such as attaining a high material standard of living, having meaningful family relationships and friend-

ships, making the world a better place, and having a good time. Life values of this type are among the factors that influence entry into various life domains and roles, the activities in those domains and roles, and how much investment is made in each one (e.g., marriage, parenthood, employment, friendships, leisure activities and hobbies, community activities, religion). Values and norms pertaining to each of the domains and roles are to some degree a function of overall life values. For example, if an individual places a higher priority on making the world a better place than on material well-being, that individual's employment values will place a higher priority on the possible influence and significance of the work performed than on the earnings derived from the work. Similarly, if an individual places a higher priority on meaningful relationships than on material well-being, marital values will place a higher priority on love and mutual respect than on the shared material standard of living.

*Range.* The *range* of a value or norm can be defined as the number and variety of actors of a particular type of object unit (e.g., individuals, organizations, and societies) to which it applies. Whereas the dimension of centrality focuses on the characteristics of action and its ends (i.e., the number and variety of behaviors or ends to which a value or norm applies), the dimension of range focuses on the characteristics of actors (i.e., the number and variety of individuals or larger social units to which a value or norm applies). The characteristics of actors used to define the range of a value or norm tend to be ascriptive or group-defining characteristics of individuals or larger social units. In the case of individuals, these are characteristics such as age, sex, nationality, race, and ethnicity. A value or norm with a broad range applies to all actors of a particular type of object unit, whereas a value or norm with a narrow range applies to a very restricted category of actors of that type. For example, concern about and responsibility for the well-being of others is a value with a broad range that applies universally to individuals throughout the world. In contrast, wisdom is a value with a narrower range because although it applies throughout the world, it applies primarily to people of older ages. Similarly, the norm against incest has a broad range because it applies universally to individuals throughout the world. In contrast, the norm prescribing paid employment has a

narrower range because it applies primarily to men in particular age categories.

*Conditionality.* The *conditionality* of a value or norm can be defined as the number and variety of situations to which it applies. Whereas the dimension of centrality focuses on the characteristics of action or its ends and the dimension of range focuses on the characteristics of actors, the dimension of conditionality focuses on the characteristics of situations, including a situation's actors. When conditionality pertains to the characteristics of a situation's actors, it usually refers to emergent or potentially changing characteristics of actors that define the situation rather than to ascriptive characteristics that define membership in social groups. Although values are less tied to specific types of situations than norms are, both values and norms vary in the degree to which they are conditioned on the characteristics of situations. For example, some values pertaining to modes of conduct, such as courtesy, cleanliness, and honesty, are applicable across most situations. Others are applicable in many fewer situations or may even be bipolar, with the polarity of the value being conditional on the situation. For example, aggressiveness is positively valued in some types of competitive situations, such as warfare and sports, but negatively valued in some types of cooperative situations, such as conversation and child rearing.

The conditionality of a value or norm is evident when a given subject actor who is evaluating a given type of action or end of action makes different evaluations in different types of situations, that is, when the evaluation varies with the characteristics of the situation. For example, friendliness is valued positively, but it is a value characterized by some conditionality, since it is valued negatively when exhibited toward strangers in dangerous environments. Killing other human beings is normatively proscribed in almost all situations, but the norm has some conditionality because killing is not proscribed in warfare, self-defense, capital punishment, and euthanasia. In capital punishment and some types of warfare, killing actually is proscribed. Abortion is believed by some people to be normatively proscribed, and whether it is normatively proscribed often depends on the characteristics of the situation, including how conception occurred, whether the mother's health is in danger, and whether the mother can care for

the child. Opposition to abortion is therefore a norm of higher conditionality than is the proscription against killing other human beings. The conditionality of a value or norm is defined by the number and variety of situations to which it applies *consistently*, that is, with the same polarity. A value or norm that has the same polarity across many and varied types of situations is a value or norm of low conditionality and therefore of high priority. A value or norm that has the same polarity in only a few similar types of situations is a value or norm of high conditionality and low priority.

*Intent.* Whether a value applies to a mode, means, or end of action has been labeled its intent (Kluckhohn 1951). Mode values pertain to the manner or style in which an action is carried out and refer to both the action and the actor. They pertain to qualities manifested *in* the act, and if such qualities are observed consistently over time for a type of action or for an actor, they are applied not just to a single instance of action but to a type of action or to an actor more generally. Adjectives such as “intelligent,” “independent,” “creative,” “responsible,” “kind,” and “generous” describe mode values. Instrumental values focus on necessary means to other ends. They refer to action that constitutes the means or from which the means are derived. For example, a job and the earnings it provides may be viewed as means to other ends such as acquiring the material resources necessary to sustain life. Goal values, in contrast, pertain to self-sufficient, or autonomous, ends of action. They are not subordinate to other values and are what an actor values most. Some analysts have argued that they can be defined as what an actor desires without limit. They focus on sources of intrinsic satisfaction or happiness but are distinguished from pleasures, which, except when elevated to become goal values, are satisfactions that are enjoyed incidentally and along the way. Pleasures are not necessarily based on beliefs about desirability, since they can be based on mere liking.

A norm may apply to a mode or means of action but not to an end of action. By requiring or prohibiting a way of acting or a type of action, norms limit the modes and means used in accomplishing ends. For example, the values of honesty and fairness govern modes and means of accomplishing ends, and associated with these values are norms that require honest and fair action.

Values and norms cannot always be identified as falling into a single category of intent. For some types of action, mode values and norms and instrumental or goal values and norms overlap; choosing an action as a means or to directly achieve an end actually defines the mode of action. For example, accomplishing a task by a means that shows concern for others defines a mode of acting that is kind, considerate, polite, and caring. Choosing to accomplish a task by honest means defines a mode of acting honestly. Acting to achieve an end that benefits others defines a mode of acting that is caring, giving, and generous. Mode values and norms and instrumental or goal values and norms do not always overlap, however. A given mode may be applied to a variety of means and ends, and choosing a means or acting to achieve an end does not necessarily imply or define a mode. For example, for modes that reflect ability or competence, as described by adjectives such as “intelligent,” “creative,” “efficient,” “courageous,” “organized,” and “self-reliant,” there may be no necessary connection or only a limited one between the values reflected in the mode and the values reflected in the acts undertaken as means or ends.

Differentiating between instrumental values and goal values is difficult because the two types are interdependent. Their relationship is not just one of sequence, since achieving particular ends may require the use of certain means (Kluckhohn 1951; Fallding 1965). Differentiating between instrumental values and goal values also requires reflection by the actor. An important concern of moral philosophy has been identifying the end or ends of action that ultimately bring satisfaction to human beings, that is, that have genuine, intrinsic value (Lovejoy 1950). The focus has been on identifying important goal values and distinguishing them from less important instrumental values. This means-end distinction is not as well developed in the category systems of all cultures as it is in Western culture (Kluckhohn 1951), and even among persons exposed to Western culture, it is not developed equally or similarly in all actors. Not all actors make the distinction or make it in the same way. What are instrumental values to some actors are goal values to others.

When mode, instrumental, and goal values are separable, they can all affect behavior. Sometimes they point to identical actions, and sometimes they do not. Similarly, when mode and

instrumental norms are separable, both can affect behavior. Among values that can pertain to either means or ends, the distinction between instrumental and goal values is a dimension of importance, with goal values being of higher priority than instrumental values (Fallding 1965; Braithwaite and Law 1985). However, values that can pertain only to a mode or means are not necessarily of lower priority than are values that can pertain to ends.

**Interrelationships.** Because social structure, as defined both organizationally and culturally, links sets of values and norms, there are patterned relationships among the sets of values and norms held by actors. These relationships can be seen as being influenced by conceptual domain, dimensions of importance, behavioral context, and interdependence.

*Conceptual Domain.* Values and norms that are conceptually similar are thought of as falling within the same conceptual domain, and a conceptual domain is identified by the observation of strong empirical relationships among sets of values or norms. Domains that are conceptually distinct also can have relationships to one another. Compatible domains are positively related, and contradictory domains are negatively related. Empirical research provides some evidence of the existence of conceptual domains of values and norms and the relationships among them. For example, in Western societies, a value domain emphasizing pleasure, comfort, and enjoyment has a negative relationship to a prosocial value domain that emphasizes concern and responsibility for others. Similarly, a value domain emphasizing the extrinsic attainment of power, money, and position has a negative relationship to the prosocial value domain (Schwartz and Bilsky 1987). Values appear to be organized along at least three broad dimensions: (1) emphasis on the self versus others, (2) emphasis on achievement versus pleasure, and (3) emphasis on the external versus the internal. Although there has been less research on the pattern of interrelationships among norms, evidence indicates that norms fall into conceptual domains. Norms pertaining to honesty, for example, are conceptually separable from norms pertaining to personal freedom in family matters, sexuality, and mortality.

*Dimensions of Importance.* Interrelationships among values and norms also are affected by dimensions of importance, since these dimensions

affect their application across object units, social institutions, social roles, and behavioral contexts. Dimensions of importance such as centrality, range, and conditionality are linked to variability in application across object units, social institutions, and social roles. Values and norms that have high importance because they are broadly applicable are more likely to be interrelated than are values and norms that have low importance, which apply more narrowly. Values and norms that apply narrowly are related to each other and to values and norms that apply more broadly only under the conditions in which they apply.

*Behavioral Context.* Interrelationships among values and norms are influenced not only by conceptual domains and dimensions of importance but also by the behavioral contexts to which they apply. Values and norms that are relevant to the same or related behavioral contexts tend to be interrelated. For example, the values and norms that play a role in interpersonal relationships differ in some respects from those which play a role in educational and occupational performance. The value of concern for others and the norms that support it are of high priority in interpersonal relationships but can be of low priority in the performance of educational and occupational tasks.

*Interdependence.* Socially structured or otherwise necessary links among modes, means, and ends of action are a source of interdependence among values and norms. Mode values and norms and instrumental or goal values and norms can overlap, and instrumental and goal values are interdependent when achieving particular ends requires the use of certain means. This interdependence constrains the extent to which the relative priority of values can affect action. For example, attaining a less highly valued means cannot be forgone to attain a more highly valued end if the end cannot be attained without the means.

## THE ORIGIN OF VALUES AND NORMS

Multiple values and norms are organized and linked in the cultures of human social systems, which are linked when they are internalized by human actors or institutionalized by corporate actors. *Social* values and norms, in contrast to *personal*, or *internalized*, values and norms refer to the values and norms of a social unit that encompasses more than one person. These may refer to the officially stated

or otherwise institutionalized values and norms of an organization or society, or to the collective, or shared, values and norms of the individuals who constitute a social unit such as an informal reference group, a formal organization, a society, or a societal subgroup defined by a shared characteristic. When a social value or norm refers to a collective property of the members of a social unit, it may be held with varying degrees of consensus by those who constitute that unit (Rossi and Berk 1985). An important difference between formal organizations and informal social groups or geographically defined social units is that formal organizations usually come into being for a specific purpose and are dedicated to particular types of activity and to achieving particular ends. As a result, their objectives are both narrower and more varied than those of other social units.

**The Social Origin of Personal Values and Norms.** The values and norms of individual persons derive from the social environments to which they are exposed. Through socialization, individuals become aware of and internalize social values and norms, which then become important internal determinants of action. An individual's internalized values and norms reflect the values and norms of the society and the various subgroups and organizations within that society to which that individual is exposed, particularly, although not exclusively, in the early stages of the life course. Once social values and norms are internalized, they can direct the behavior of individuals irrespective of external influences. Internalized values and norms are a source of *self-expectations* and a basis of *self-evaluation*, with the subjective response to an outcome ensuing from the self-concept. Adherence to self-expectations enhances self-esteem, producing a sense of pride and other favorable self-evaluations. Violation of self-expectations reduces self-esteem, producing guilt, self-depreciation, and other negative self-evaluations. To preserve a sense of self-worth and avoid negative self-evaluations, individuals try to behave in accordance with their internalized values and norms. Sociologists tend to see internalized values and norms as an important influence on human behavior, and this makes them see the social values and norms of society as governing and constraining the choices individuals make. Social values and norms also affect behavior because they are internalized by significant others and thus affect an

actor's perception of other people's expectations. To the extent that actors are motivated to comply with what they perceive the views of others to be, social values and norms become a source of external pressure that exerts an influence that is independent of an individual's internalized values and norms.

Although change in personal values and norms occurs over the life course, there is some evidence that levels of stability are relatively high (Moss and Susman 1980; Sears 1983; Alwin 1994). It has been argued that values and norms that are more closely tied to the self-concept and considered more important are more resistant to change (Rokeach 1973; Glenn 1980). Those values and norms may undergo less change because they are internalized through conditioning-like processes that begin early in life and are strongly linked to existential beliefs. They tend to be tied to shared mental models that are used to construct reality and become embedded central elements of cognitive organization with a strong affective basis. Some types of values, norms, and attitudes (for example, political attitudes) are quite malleable into early adulthood and then become relatively stable. After this "impressionable," or "formative," period when change is greatest, they are relatively stable in midlife, and this stability either persists or declines in the later years (Alwin et al. 1991; Alwin 1994).

The pattern of life-course change and stability described above has been argued to be due to a number of influences. One is the process of biological and psychological maturation with age, which is most rapid in the early stages of life. As functional capacity develops, influences at that time have the advantage of primacy, and when they are consistent over a period of years, affective "mass" is built up. Nevertheless, some types of values, norms, and attitudes remain malleable into early adulthood, and strong pressure to change or weak earlier socialization can lead to resocialization in late adolescence or early adulthood (Sears 1981; Alwin et al. 1991). It is likely that change declines after early adulthood in part because individuals tend to act on previously formed values, norms, and attitudes as they seek new information and experiences. This selective structuring of new inputs enhances consistency over time, since new inputs tend to reinforce rather than call into question earlier ones.

Another influence on life-course change and stability in values and norms is change in social experiences and roles over the life course (Wells and Stryker 1988; Elder and Caspi 1990). These changes are extensive during the transitional years of early adulthood and may increase after retirement. They represent opportunities for change because they bring the individual into contact with new individuals, reference groups, and situations, and change in values and norms is likely to occur through both interaction with others and adaptation to situations. Role change can produce change as a role occupant engages in new behaviors, is exposed to new circumstances and information, and learns the norms governing role behavior. After early adulthood, a decline in the number of changes in social experiences and roles leads to greater stability in values and norms.

**Sources of Change in Social Values and Norms.** Change in social values and norms occurs through a variety of processes. One influence is historical change in the conditions of life that occurs through technological innovation, alterations in economic and social organization, and change in cultural ideas and forms. Historical change by definition involves “period effects,” but because those effects tend to be experienced differently by different birth cohorts (i.e., those at different ages when a historical change occurs), the influence of historical change on social values and norms occurs to some degree through a process of cohort succession.

Change in social values and norms also occurs through change in the social values and norms of subgroups of social units. This change can be of several types. First, change in the presence and size of subgroups with different values and norms produces change in the collective values and norms of the group. For example, the presence of new immigrant groups with different values and norms or a change in the relative size of groups with different values and norms affects the values and norms of the collective unit. Second, change in the degree of similarity or difference in the values and norms of subgroups can produce change in overall values and norms. On the one hand, acculturation through intergroup contact and similar experiences will reduce the distinctiveness of subcultural groups; on the other hand, segregation and increasing divergence in the life experiences of subgroups will widen their cultural distinctiveness.

Third, some subcultural groups may be more subject to particular period influences than others are, and this differential responsiveness can increase or decrease differences in values and norms among subgroups.

Another source of change in social values and norms is change in exposure to social organizations that exert distinct socializing influences. For example, exposure to religious, educational, or work organizations may produce differences in values and norms between those with such exposure and those without it. The extent to which exposure to different organizational environments is likely to affect personal values and norms depends on the distinctiveness of those environments, which also is subject to change. Thus, social values and norms are affected by both changes in the exposure of the population to different organizations and changes in what is socialized by those organizations.

#### **THE ROLE OF VALUES AND NORMS IN EXPLAINING BEHAVIOR**

The ways in which values and norms influence behavior must be understood in a larger explanatory framework, and models of purposive action in all the social sciences provide that framework (Marini 1992). These models rest on the assumption that actors are purposive, acting in ways that tend to produce beneficial results. Although the models of purposive action that have emerged in various social sciences differ in the nature of the assumptions made about purposive action, they share the basic proposition that people are motivated to achieve pleasure and avoid pain and that this motivation leads them to act in ways that, at least within the limits of the information they possess and their ability to predict the future, can be expected to yield greater reward than cost. If reward and cost are defined subjectively and individuals are assumed to act in the service of subjective goals, this proposition links subjective utility, or value, to action. In sociology, a model of purposive action assumes the existence of actors who may be either persons or corporate actors. The usefulness of these models in sociology hinges on making appropriate connections between the characteristics of social systems and the behavior of actors (the macro–micro connection) and between

the behavior of actors and the systemic outcomes that emerge from the combined actions of multiple actors (the micro–macro connection).

In a model of purposive action, an individual actor (person or corporate actor) is assumed to make choices among alternative actions structured by the social system. Choices among those actions are based on the outcomes expected to ensue from those actions, to which the actor attaches some utility, or value, and which the actor expects with some probability. The choices of the actor are governed by beliefs of three types: (1) the perceived alternatives for action available, (2) the perceived consequences expected to result from each alternative, and (3) the perceived probabilities with which those consequences are expected to result. The choices of the actor also are governed by the actor's preferences, or the subjective utility (rewards and costs) of the consequences expected to result from each alternative. Values and norms are among the preferences of an actor that influence action. As evaluative beliefs that synthesize affective and cognitive elements, they affect the utility of the outcomes expected to ensue from an action. Action often results not from a conscious weighing of the expected future benefits of alternatives but from a less deliberate response to internalized or institutionalized values and norms (Emerson 1987). The actor's finite resources—the human, cultural, social, and material capital available to the actor that enables or precludes action—operate as influences on the choices made by the actor.

The component of a model of purposive action that makes the macro–micro connection links the characteristics of the social system to the behavior of actors and models the effects of social structure (both organizational and cultural) on the beliefs and preferences of actors as well as on the available alternatives for action and actors' resources. In this component of the model, characteristics of the micro model are taken as problematic and to be explained. These characteristics include: (1) the beliefs and preferences on the basis of which an actor makes choices, (2) the alternatives available to an actor, and (3) the resources available to an actor. A third component of a model of purposive action makes the micro–macro connection, linking the behavior of individual actors to the systemic outcomes that emerge

from the combined actions of multiple actors. This link may occur through a simple mechanism such as aggregation, but it is more likely that outcomes emerge through a complex interaction in which the whole is not just the sum of its parts. The action, or behavior, of the system is usually an emergent consequence of the interdependent actions of the actors that compose it.

## REFERENCES

- Allport, Gordon W., Philip E. Vernon, and Gardner Lindzey 1960 *Manual: Study of Values*, 3rd ed. Boston: Houghton Mifflin.
- Alwin, Duane F. 1994 "Aging, Personality, and Social Change: The Stability of Individual Differences Cover the Adult Life Span." In D. I. Featherman, R. M. Lerner, and M. Perlmutter, eds., *Life Span Development and Behavior*, vol. 12. Hillsdale, N.J.: Erlbaum.
- , Ronald L. Cohen, and Theodore M. Newcomb 1991 *Political Attitudes over the Life-Span: The Bennington Women after Fifty Years*. Madison: University of Wisconsin Press.
- , and Jon A. Krosnick. 1985. "The Measurement of Values in Surveys: A Comparison of Ratings and Rankings." *Public Opinion Quarterly* 49:535–552.
- Braithwaite, V. A., and H. G. Law 1985 "Structure of Human Values: Testing the Adequacy of the Rokeach Value Survey." *Journal of Personality and Social Psychology* 49:250–263.
- Elder, Glen H., and Avshalom Caspi 1990 "Studying Lives in a Changing Society: Sociological and Personological Explorations." In A. I. Rabin, R. A. Zucker, R. A. Emmons, and S. Frank, eds., *Studying Persons and Lives*. New York: Springer.
- Emerson, Richard M. 1987 "Toward a Theory of Value in Social Exchange." In K. S. Cook, eds., *Social Exchange Theory*. Newbury Park, Calif.: Sage.
- Fallding, Harold 1965 "A Proposal for the Empirical Study of Values." *American Sociological Review* 30:223–233.
- Fishbein, Martin and Icek Ajzen 1975 *Belief, Attitude, Intention and Behavior*. Reading, Mass.: Addison-Wesley.
- Gibbs, Jack P. 1965 "Norms: The Problem of Definition and Classification." *American Journal of Sociology* 70:586–594.
- Glenn, Noval D. 1980 "Values, Attitudes and Beliefs." In O. G. Brim, Jr., and J. Kagan, eds., *Constancy and Change in Human Development*. Cambridge, Mass.: Harvard University Press.



Kahneman, Daniel, and Amos Tversky 1979 "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47:263-291.

Kluckhohn, Clyde 1951 "Values and Value-Orientations in the Theory of Action: An Exploration in Definition and Classification." In T. Parsons and E. A. Shils, eds., *Toward a General Theory of Action*. New York: Harper & Row.

Lovejoy, Arthur O. 1950 "Terminal and Adjectival Values." *The Journal of Philosophy* 47:593-608.

Marini, Margaret Mooney 1984 "Age and Sequencing Norms in the Transition to Adulthood." *Social Forces* 63:229-244.

—, 1992 "The Role of Models of Purposive Action in Sociology." In J. S. Coleman and T. J. Fararo, eds., *Rational Choice Theory: Advocacy and Critique*. Newbury Park, Calif.: Sage.

Moss, Howard A., and Elizabeth J. Susman 1980 "Longitudinal Study of Personality Development." In O. G. Brim, Jr., and J. Kagan, eds., *Constancy and Change in Human Development*. Cambridge, Mass: Harvard University Press.

Parsons, Talcott, and Edward A. Shils 1951 "Values, Motives, and Systems of Action." In T. Parsons and E. A. Shils, eds., *Toward a General Theory of Action*. New York: Harper & Row.

Rokeach, Milton 1973 *The Nature of Human Values*. New York: Free Press.

— 1985 "Inducing Change and Stability in Belief Systems and Personality Structures." *Journal of Social Issues* 41:153-171.

Rossi, Peter H., and Richard A. Berk 1985 "Varieties of Normative Consensus." *American Sociological Review* 50:333-347.

Schwartz, Shalom H., and Wolfgang Bilsky 1987 "Toward a Universal Psychological Structure of Human Values." *Journal of Personality and Social Psychology* 53:550-562.

Sears, David O. 1981 "Life-Stage Effects on Attitude Change, Especially among the Elderly." In S. B. Kiesler, J. N. Morgan, and V. K. Oppenheimer, eds., *Aging: Social Change*. New York: Academic Press.

— 1983 "The Persistence of Early Political Predispositions: The Roles of Attitude Object and Life Stage." In L. Wheeler, eds., *Review of Personality and Social Psychology*, vol. 4. Beverly Hills, Calif.: Sage.

Thornton, Arland 1985 "Changing Attitudes toward Separation and Divorce: Causes and Consequences." *American Journal of Sociology* 90:856-872.

Wells, L. E., and Sheldon Stryker 1988 "Stability and Change in Self over the Life Course." In P. B. Baltes,

D. L. Featherman, and R. M. Lerner, eds., *Life-Span Development and Behavior*. Hillsdale, N.J.: Erlbaum.

MARGARET MOONEY MARINI

## SOCIAL WORK

Social work has been defined as being "concerned with the interactions between people and their social environment which affect the ability of people to accomplish life tasks, alleviate distress, and realize their aspirations and values. The purpose of social work therefore is to (1) enhance the problem-solving and coping capacities of people, (2) link people with systems that provide them with resources, services, and opportunities, (3) promote the effective and humane operation of these systems, and (4) contribute to the development and improvement of social policy" (Pincus and Minahan 1973, p. 9). A key difference between social work and sociology lies in the emphasis placed on intervention in social work. A social worker expects to be actively involved in the amelioration of social problems, while a sociologist typically focuses on understanding the nature and extent of social issues. Social workers establish a helping relationship with a client system (individual, family, small group, community), using their assessment skills and knowledge of helping resources to identify alternatives that may improve a situation.

## PROFESSIONAL ROOTS

Professional social work is historically tied to the emergence of social welfare as a social institution. Social welfare as it has come to be known, can be traced to society's numerous attempts to accommodate changes in economic and social relationships over time. The beginning of institutionalized social welfare is frequently ascribed to the English Poor Law of 1601. As the most critical part of modern social welfare's foundation, the Elizabethan poor laws were characterized by the articulation and promulgation of the principle of public responsibility and obligation for the economic well-being of the people. However, "the Poor Laws in England and in American communities were not primarily concerned with poverty and how to eliminate it. Instead, they were concerned with pauperism and the potential claims on community

funds, the danger that paupers might get by without working” (Dolgoff and Feldstein 1984, p. 80). This continuing tension between public obligation and social control is one of several dualities that characterize the context of professional social work practice. Institutionalized social welfare is the environment in which the profession of social work has developed. The history of social welfare is paralleled by and enmeshed with the increasing professionalism of those who administer social welfare programs.

Early social work was characterized by two streams of activity: social reform and direct assistance to individuals and families. The practice of friendly visiting and the development of both the Charity Organizations Societies and settlement houses illustrate both types of effort. Representatives of Charity Organization Societies, the so-called friendly visitors, engaged in social investigation and moral susasion improve the lives of the poor. The thrust of those encounters was to place responsibility on the persons or families for their economic and social status, what is known now as “blaming the victim.” The work of the Charity Organization Societies formed the origins of the social work method later known as social casework.

Residents of settlement houses, Jane Adams included, were friendly visitors who came to stay. A group of middle-class or upper-class individuals moved into residence in a poor area in an effort to study neighborhood conditions firsthand and work with neighborhood residents on solving neighborhood problems. While some settlement house efforts focused on assimilation, later programs focused on improving conditions in immigrant communities. In cities across the nation, settlement houses helped acculturate vast numbers of immigrants in the early part of the twentieth century. Settlement house activities emphasized teaching English, health practices, occupational skills, and environmental changes through cooperative efforts. Settlement house staff developed social group work, community organization, social action, and environmental change efforts. Furthermore, settlement house workers were active in the legislative arena, gathering and promulgating facts in order to influence social policy and legislation.

An early and continuing cleavage in the profession has its origins in differing explanations of social dysfunction. Some early social workers es-

poused the theory of the social causation of social problems and sought governmental actions to meet needs as well as developing coalitions for reform and institutional change. The educational foundation came from sociology, economics, and political science. Others emphasized individual causation of social problems, promoting an individually focused therapeutic approach to helping. These social workers identified the need to draw on psychological theory but emphasized the individual interacting with a social environment. These two primary orientations would feed the development of professional social work and provide the basis for conflict within the practice community and in professional social work education

### PROFESSIONALIZATION

An issue throughout the development of professional social work has been the nature of its professional status. In 1915 Abraham Flexner critiqued the professional status of social work at the National Conference of Charities and Corrections. Although Flexner criticized social work as lacking a specific skill for a specific function, he also recognized its professional spirit. The ideal-type model of a profession has been the conception against which social work has measured itself through much of its history. Greenwood’s (1957) analysis examined the extent to which social work possessed five classic traits of a profession: systematic theory, authority, community sanction, an ethical code, and a professional culture. Characterizing social work as a less-developed profession, Greenwood concluded that it possessed these attributes to a moderate extent. The predominant direction of the field, however, has been to continue its professional development along all five dimensions. The recent emphasis on building the empirical base of practice coupled with more stringent licensure requirements by states are indicators of the continued progression of social work toward greater professional status. It would be incorrect to assume, however, that this direction is embraced by the profession as a whole. For those whose dominant professional identification is with the field’s social action tradition, increasing professionalization means being co-opted. Achieving the public acceptance accorded to a profession can distance social workers from their constituencies and limit confrontational strategies that are central to advocacy for the oppressed.

In the 1920s the practice of social work emerged in so-called fields of practice or settings: family and child welfare and medical, psychiatric, and school social work. Social workers defined their central problems and responsibilities as being characteristic of their particular fields. The concept of method also emerged during this period. Method developed first around casework and later in relation to both group work and community organization. Methods were based on selected theories of human behavior drawn from psychology and sociology. Setting referred to the organizational context within which services were delivered.

This combination of method and field of practice or setting fragmented professional social work, slowing the development of an integrated theoretical base for practice across methods and settings. Social casework theory and method developed to a large extent in isolation from group work and community organization. The curricula for professional social work education followed the same pattern, with separate tracks for each method. It took until the 1970s for the development of a conceptual approach based on the essential components of professional practice regardless of where a social worker was employed. Pincus and Minahan (1970) articulated a conceptual framework for generalist practice, that is, for social work service delivery across practice settings. This approach encompassed three major components: the *social systems* in relation to which a social worker carries out his or her role, the stages of *planned change* or problem-solving processes, and interactional and analytic *skills* for data collection, analysis, and intervention.

### VALUES, ETHICS, AND THE BUREAUCRATIC CONTEXT

Since social work as a profession is concerned with social change and the improvement of the conditions in which people live, its orientation cannot be value-free or purely theoretical. A defining characteristic of social work practice is a fundamental commitment to knowledge, skills, and a core set of professional values to enhance the well-being of people and ameliorate environmental conditions that affect people adversely. Among the values and principles that guide professional practice are respect for individual worth, dignity, the right to self-determination, and active partici-

pation in the helping process; helping clients obtain needed resources; demonstrating respect for and acceptance of the characteristics of diverse populations; a commitment to the promotion of social change to achieve social and economic justice; an understanding of the dynamics of oppression and discrimination, along with attention to populations at risk; and a holistic view of the interactions between people and the complex environment in which they live. These values are embedded in the *Code of Ethics* of the National Association of Social Workers (1994). The code focuses on the conduct and comportment of a social worker as well as ethical responsibilities to clients, colleagues, employers, the profession, and society.

A distinguishing characteristic of social work is that the majority of its practitioners are employed by a variety of public and private social welfare agencies. Some social workers are employed by agencies that are sanctioned to function as agents of social control, while others have the authority to determine eligibility for benefits and services. The bureaucratic environment, however manifested, dramatically shapes the practice of social workers. The process of professional socialization is designed to instill a culture, a set of values and expectations, that may conflict with the work environment.

Professionals' autonomy can be circumscribed by organizational commitments, policies, and procedures. In these circumstances, just whose agent is the professional social worker: the agency's, the client's, the community's, or his or her own as an autonomous professional? In an organizational context, what form can a social worker's social action efforts take? How far can an employed social worker go in challenging an agency's priorities, policies, and procedures before his or her services are no longer desired? How long does it take before a professional social worker starts to identify more as an agency employee than as an autonomous professional? Given the range of practice settings and the variety of roles of social workers, there are no easy answers to these questions. These realities can produce a conservatizing effect on social work, limiting many workers' willingness or ability to take risks as autonomous professionals in the name of social justice and reform. In these circumstances, one can see how theories of individual causation can prevail over

explanations that invoke the influence of larger social forces in the creation and amelioration of social problems. This tension, with its roots in the origins of the profession, continues, as demonstrated by the overwhelming preference of students and professionals for work with individuals and families, mostly in the psychotherapeutic model.

### THE KNOWLEDGE BASE AND EMPIRICALLY BASED PRACTICE

The creation of a systematic body of theory has been under development from the early days of the profession. Richmond's *Social Diagnosis* (1917) organized the contemporary theory and method of social work and formulated a data collection approach designed to serve as the foundation for diagnosis. Richmond organized and analyzed the naturalistic observations she made while working with individuals and families. Her work is the origin of psychosocial history taking and treatment plan development and perhaps the core of social casework practice methods. Richmond's contribution to the organization of what eventually would become social casework practice is legendary, forming the bedrock of clinical social work. Her approach, later to be known as empirically based practice, represents one of the two major streams of knowledge and theory development in social work. The other major focus has relied heavily on the application of social science (primarily sociological and psychological) theory to the explanation of social problems and the development of interventions to ameliorate those problems.

The breadth of social work practice (encompassing work with individuals, families, groups, and communities and including social work program administration, public policy development, and social planning) provides a rich and continually changing field for exploratory, descriptive, and explanatory empirical efforts. The early 1970s was a benchmark in the development of the profession's knowledge base. Along with the massive investment in social programs of the 1960s came the realization that good intentions and humane values are not enough. Funders focused increasingly on outcomes. Attention was shifted to the development of empirically based justifications for programs, services, and budgets. Program evaluation became the dominant focus of much of

social work research during this period, including methodology, design, outcomes, and professional accountability. This direction came to be known as the practice effectiveness movement. As articulated by Fischer, the question became, "Is Case-work Effective?" (Fischer 1973; Fischer and Hudson 1976).

During this period, a study of the effects of adult protective services by Blenkner et al. (1971) at the Benjamin Rose Institute in Cleveland created a furor. An early social experiment, this demonstration program, which employed skilled caseworkers, was reported to be associated with more negative effects than was the control program, which employed less highly trained workers. After one year, the findings were alarming. The experimental group manifested higher death rates, higher utilization of protective services, higher rates of institutionalization, a nonsignificant increase in contentment, and a nonsignificant decrease in symptoms of emotional disturbance. The authors concluded that the "effect of more skilled social workers on the clients was to 'overdose' them with help. This led to more concrete assistance, including institutionalization, which in turn was responsible for the higher death rate. . . . More highly trained social workers were apparently more lethal" (Tobin 1978). These findings could not demonstrate the effectiveness of professional social work intervention, illustrate accountability, or be used to justify program expenditures.

This study and the controversy it generated shifted attention from program description to research design, sampling, and data analysis. Investigators (Berger and Piliavin 1976; Fischer and Hudson 1976) reanalyzed the data in an attempt to discover alternative explanations for the findings. Berger and Piliavin argued that although randomization had been used to assign clients to groups, the experimental group was older and more mentally and physically impaired than were the controls. Fischer and Hudson (1976) challenged the sample size used in Berger and Piliavin's regression analysis and demonstrated that age, mental status, and physical status, although separate variables, produced an additive effect. The nature of the debate had shifted: Methodological issues had become the basis of discussion. Values and good intentions alone would no longer be sufficient grounds for justifying programs or demonstrating professional accountability.

### THE PRACTITIONER-RESEARCHER

Concern with the outcomes of social work interventions led to the concept of the social worker as both practitioner and researcher. From this perspective, social workers are seen as having the opportunity and responsibility to develop methods and skills from an empirical base, from the experience provided in their own practice to develop, test, and refine practice innovations. Embedded in this movement toward practitioner-based empirical practice was the notion that evaluation and research were too critical to leave in the hands of a group of research "specialists." Perhaps more fundamental is the belief that social work research is too important to leave in the hands of those who are not social workers: "It is the practicing professional who encounters and struggles with current issues and who is most sensitive to the critical knowledge gaps in the field. Thus social workers are in the best position to formulate and conduct the needed research and evaluation and they must be committed to acquiring the understanding required to direct the helping effort" (Grinnell 1996, p. 5)

These developments coincided with the expansion of doctoral education in social work. While past doctoral preparation often focused on the development of advanced clinical skills, contemporary training at the doctoral level is almost exclusively research-based, designed to provide students with the skills needed to contribute to the empirically anchored knowledge base of the profession. As a result, a cohort of social work researchers has been trained over the last twenty years, and this group has developed a body of theory and knowledge that has been generated directly as social work research. Social work no longer defers to sociology for the methodological sophistication to evaluate its programs and practice outcomes.

### SOCIOLOGY AND SOCIAL WORK

Over time the link between social work and sociology has been strong, although the two fields have grown increasingly distant. There can be little doubt, however, regarding the importance of sociological theory and research for the development of the knowledge and theoretical base of social work practice. For example, social stratification, conflict theory, deviance, organizational

theory, community development and dynamics, family studies, occupational sociology, criminology, and life-span theories are only a few areas of sociological theory development and research that have informed and directly influenced both the theory and the practice of social work. Landmark social program evaluation studies were undertaken by sociologists, some of whom were members of faculties of social work, in the late 1950s and 1960s (Meyer and Borgatta 1959; Meyer et al. 1965).

Clearly, social work and sociology are related, although there are fundamental differences. Sociologists study and analyze social organizations and institutions. The emphasis has been on theory development, primarily through positivistic approaches, focusing on measurement and design issues. Although the development of grounded theory (Glaser and Strauss 1967) has been a major conceptual contribution in sociology, it has not been the dominant influence. Although there are reform-minded, "radical" sociologists, they are a minority. Sociologists are interested in understanding the "why" of human interaction. Sociology observes; it maintains a detached posture.

In contrast, social workers attempt to apply theories of social organization and interaction to improve social functioning. Social workers go beyond understanding social problems in their efforts to improve social functioning; social work intervenes. The goal is engendering progressive social change, improving social conditions, creating more humane delivery systems, and problem solving with individuals, families, groups, communities, and organizations and in public policy. Social workers develop and implement interventions in the form of programs, policies, and services in the context of public funding and demands for professional accountability. The orientation is toward outcomes, cost-effectiveness, and cost-benefit analyses.

There has been and continues to be tension in the relationship. Heraud notes that "the social worker may be able to participate actively in policy making through social science research; there is considerable need for research related to both intended and unintended consequences of social policy . . . there is considerable need in the initial stages of such research for intuition and speculation. Instead of only using the sociologist at this stage, who may be a distant figure, the social

worker may have an important role to play" (1970, p. 287). Several years earlier, Halmos noted that social workers could function "as an intelligence agent of the sociologist and of the policy maker, and a trusty pilot of the sociological researcher" (1961, p. 9). Although these attitudes may be antiquated, elements of such elitism remain, particularly in sociology's limited interest in applied social research.

Some attention has been paid to the development of so-called applied sociology. While the main body of sociological thought focuses on exploratory, descriptive, and explanatory theory; modeling; and empirical testing, "applied sociology" briefly emerged in response to the increasing interest in social program evaluation and the limited supply of trained methodologists who could design and execute well-formulated evaluative studies. Thus, applied sociology could provide an alternative, public-policy-oriented career path for sociologists, since the preferred, higher-status university-based employment opportunities were limited.

Over the last twenty five years, however, social work researchers have become key players in the design, implementation, and analysis of applied social research, particularly through their involvement in federally funded demonstration projects. During this period, there has been a proliferation of journals of social work, including research journals (*Social Work Research and Abstracts, Research on Social Work Practice*), as well as a range of specialty journals (*Gerontological Social Work, Health and Social Work, Child Welfare, School Social Work*), which provide publication outlets for researchers and practitioners.

At one time, social work education occurred within the social sciences, frequently attached to sociology. More recently, social work has emerged as an independent professional discipline, forming alliances with a variety of other professions, such as law, education, business, and nursing. Increasing numbers of pragmatic students have been attracted to social work because of the ability of graduates to find employment.

The undergraduate degree (BSW) offers a generalist foundation that is built on a set of social science prerequisites. The graduate degree (MSW), the terminal educational degree for the profession, is based on specialized courses that offer advanced theoretical content in fields of practice

and methodological approaches. The Council on Social Work Education (CSWE) has exercised a substantial influence in setting standards for social work education. Periodic accreditation reviews by the council assure uniformity and consistency in the required content. Particular attention has been paid to including content on minorities and oppressed populations. Accreditation by the CSWE is essential for the credibility of any social work education program in the United States.

Social work is an evolving profession, with its form and emphasis changing in response to the societal context within which social workers practice:

*"Most social workers feel that although there are critical problems and pressures, numerous opportunities are available for the social work profession to move ahead on a sound basis, strengthening current delivery of services and innovating services that have been practically untouched to date. . . . Once thought of as a basket-on-the arm assistance for the poor, it is now a discipline, scientific in method and artful in manner, that takes remedial action on problems in several areas of society"*  
(Skidmore et al. 1997 pp. 376 to 3).

#### REFERENCES

- Berger, Raymond, and Irving Piliavin 1976 "The Effect of Casework: A Research Note." *Social Work* 21:205-208.
- Blenkner, Margaret, Martin Bloom, and Margaret Nielson 1971 "A Research and Demonstration Project of Protective Services." *Social Casework* 52:483-499.
- Dolgoff, Ralph, and Donald Feldstein 1984 *Understanding Social Welfare*, 2nd ed. New York: Longman.
- Fischer, Joel 1973 "Is Casework Effective? A Review." *Social Work* 18:5-20.
- , and Walter Hudson 1976 "An Effect of Casework? Back to the Drawing Board." *Points and Viewpoints*. *Social Work* 21:347-349.
- Glaser, Barney, and Anselm Strauss 1967 *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.
- Greenwood, Ernest 1957 "Attributes of a Profession." *Social Work* 2:45-55.
- Grinnell, Richard 1996 *Social Work Research and Evaluation*. Itasca, Ill.: Peacock.
- Halmos, Peter 1961 "Problems Arising in the Teaching of Sociology to Social Workers." *International Social Service Review* 8:122-130.

- Heraud, Brian 1970 *Sociology and Social Work*. New York: Pergamon.
- Meyer, Henry, and Edgar Borgatta 1959 *An Experiment in Mental Health Rehabilitation: Evaluating a Social Agency Program*. New York: Russell Sage Foundation.
- , —, and Wyatt Jones 1965 *Girls at Vocational High: An Experiment in Social Work Intervention*. New York: Russell Sage Foundation. National Association of Social Workers 1994 *Code of Ethics*. Washington, D.C.: National Association of Social Workers.
- Pincus, Allen, and Anne Minahan 1970 "Toward a Model for Teaching a Basic First Year Course in Methods for Social Work Practice." In Lillian Ripple, ed., *Teaching Social Work Practice*. New York: Council on Social Work Education.
- 1973 *Social Work Practice: Model and Method*. Itasca, Ill.: Peacock.
- Richmond, Mary 1917 *Social Diagnosis*. New York: Russell Sage Foundation.
- Skidmore, Rex, Milton Thackeray, and William Farley 1997 *Introduction to Social Work*, 7th ed. Boston: Allyn and Bacon.
- Tobin, Sheldon 1978 "Old People." In Henry Maas, ed., *Social Service Research: Reviews of Studies*. Washington, D.C.: National Association of Social Workers.

CAROL D. AUSTIN  
ROBERT W. McCLELLAND

## SOCIAL-AREA ANALYSIS

*See* Cities.

## SOCIALISM AND COMMUNISM

What is socialism? According to a Hungarian joke made during the "gentle revolution" of 1989, it is the "longest and most painful road from capitalism to capitalism" (Garton Ash 1990). Although this biting definition was fashionably cynical about Soviet-type societies in the wake of their fall, it provides no substantive insights into one of the major social organizational forms of modern history.

The origins of socialism are obscure. Intellectual historians have traced its beginnings to the religious utopias of the Old Testament (Laidler 1968), the principles of Mosaic law (Gray 1963), the anti-individualism of the radical sects that emerged after the French Revolution (Lichtheim

1969), and the publication of the *Communist Manifesto* (Sweezy 1983). As well as can be determined, the term made its first appearance in Italian print in 1803, although its meaning at that time differed somewhat from the current interpretations (Cole 1959). For this reason, the origin of the term usually is attributed to the *London Co-Operative Magazine*, where it was used to designate followers of Robert Owen (Nutti 1981). The first French usage followed shortly thereafter when, in 1832, a French periodical, *Le Globe*, used it to characterize the writings of Saint-Simon (Bell 1968; Kolakowski 1978).

Despite its complicated origins, by 1840 the concept was used commonly throughout Europe and was making its way across the Atlantic to the United States. By the early 1920s, the Soviet Union had already claimed "socialism" as its overall organizing principle; ironically, at that time, over 260 definitions of the term were available in the social scientific literature (Griffith 1924), rendering its meaning somewhat ambiguous. Since then, further transformations of the concept have appeared; for instance, scholars now differentiate among Chinese socialism, corporatist socialism, democratic socialism, radical socialism, and Russian socialism.

The common core of socialist ideas is hard to define. To be sure, all socialists were critical of the competitive and unequal nature of capitalist society, and without fail, they championed a more egalitarian and just future. At the same time, their visions of the organization of a socialist future were sufficiently diverse to render a single definition of the term practically impossible. It is frequently assumed, for example, that all socialists wanted to establish communal ownership, yet many were content with the centralization of resources in the hands of the state (e.g., Bernstein 1961) and others actually protested the abolition of private property (e.g., Saint-Simon 1964). Battles also were waged over the role of the state: Some believed that centrally managed administrative organs would become superfluous under a socialist regime (Proudhon 1966), while others regarded those organs as essential for the management of community affairs (e.g., Cabet 1975). Many argued that the freedom of the individual must be guaranteed at all costs even under socialism (e.g., Fourier 1971), while others were willing to impose limitations on such freedom in the name of equality and

efficient production (Mao 1971). Finally, some believed that socialism could be realized through gradual reforms (Bernstein 1961), while others thought that it was possible only through a major revolution (Lenin 1971).

Because of the nontrivial nature of these differences, a single definition of socialism is likely to conceal more than it illuminates. For this reason, it is more productive to highlight features of the concept by examining separately some of the best known schools of socialist thought.

### THE IDEA OF SOCIALISM

In the view of utopian socialists, socialism was a romantic vision whose purpose was not necessarily to be realized but to serve as an ideal against which the evils of capitalism could be compared. The specific content of this vision varied from author to author, but two central themes can be identified.

The ideal of community was the first of those themes. From Fourier to Cabet, through Owen and Saint-Simon, all utopian theorists championed a new social order organized around small communities. In most sketches of socialism, this vision was realized in an agrarian setting (e.g., Cabet 1975), although some required advanced industrial development (e.g., Saint-Simon 1964). In either case, however, it was assumed that those communities would be based on fellowship, harmony, and altruism—virtues that utopian theorists favored on moral grounds over bourgeois individualism.

Nostalgia for the past is the second common theme in utopian socialist thought. It frequently appeared in utopian novels and usually assumed one of two forms. In some versions, the protagonists in those novels were returned to a romanticized preindustrialism, while in others, they returned to an even more distant past, such as the Middle Ages (e.g., Morris 1970). Despite such variation in the settings of those novels, the message they sought to convey was more or less the same: In the transition to industrial capitalism, people abandoned the “golden age” of social harmony and replaced it with a fragmented and competitive social order that is unable to provide for the full satisfaction of human needs.

In the hands of scientific socialists, the idea of socialism represented more than just an attractive

dream (Marx and Engels 1968). Karl Marx, for example, considered it a historically possible future for capitalism, as he assumed that the internal contradictions of capitalism would create some of the preconditions for socialism. According to his theory of historical materialism, the demands made by capitalist development will create increasingly grave crises for the ruling class. He maintained that with the mechanization of production and the concentration of capital in the hands of a few, there will be greater polarization in terms of class inequalities and an increase in the degree of exploitation of the working class. As capitalism enters its advanced stage, the condition of the working class will deteriorate and the struggle over the quality of its existence will intensify. At first, the war between the “two hostile camps” of capitalist society (the bourgeoisie and the proletariat) will be waged within the boundaries of particular nation-states. However, as capitalism expands into new markets internationally, workers across the world will be forced to unite in their effort to overthrow capitalist society. Socialism, according to Marx, will emerge out of this final instance of class struggle.

It is ironic that the “father of socialism” never provided a detailed blueprint for his model of the future. It is evident from a number of passages, however, that Marx envisioned two stages in the evolution of socialism. In the lower stage (which he referred to as *socialism*, or the “dictatorship of the proletariat”), he foreshadowed major improvements in the human condition. He predicted, for example, that private property would be abolished, the forces of production would be nationalized and placed in the hands of the state, rights of inheritance would be eliminated, universal suffrage would be introduced, state representatives would be elected from among the working people, and education would become accessible to all. At the same time, because Marx expected this to be a transitional stage, he believed that some elements of capitalist society would continue to prevail. Specifically, he mentioned that income inequalities would continue to exist in the lower stage because workers would still be paid according to the amount of work they contributed to the social good.

At some point, according to Marx, this transitional phase in the development of human history would evolve into the higher stage of socialism, a



stage that he often referred to as *communism*, or the “realm of freedom.” Under communism, work would no longer be an obligation but a free and creative activity, alienation would be transcended, the production process would be under the direct control of the producers, and rewards would be distributed in accordance with the principle of “to each according to his need” rather than “to each according to his ability.”

Scientific socialism gained considerable popularity among French, German, and British socialists during the nineteenth century. Many agreed with Marx’s assessment of bourgeois society and were attracted to his vision of the future. As the century progressed, however, and the Marxist scenario still appeared to be far away, some began to raise questions about the continued relevance of scientific socialism in the modern age. The main protagonist in this debate was Eduard Bernstein, a leading advocate of democratic socialism.

Bernstein and his followers called into question various elements of scientific socialism, but they were especially concerned about Marx’s predictions concerning the development of industrial capitalism. On the basis of new empirical evidence, Bernstein (1961) noted that the standard of living at the turn of the century was improving rather than deteriorating, class inequalities were far from polarized, and the ownership of capital, rather than being concentrated in the hands of a few, was becoming diversified. In addition, he observed that general strikes were becoming less common and socialist parties were gaining considerable strength in the political organization of the state. In light of those findings, Bernstein called for a revision of the Marxist program and offered a new interpretation of socialism.

According to Bernstein, democracy was the most important feature of socialist society. He discouraged his confederates from describing socialism as a “dictatorship of the proletariat” and recommended that they acknowledge its fundamentally pluralist character. Of course, for Bernstein, the significance of democracy was not simply that it guaranteed the representation of minority rights under socialism; it was also that it assured a peaceful transition from capitalism through a series of parliamentary reforms. For many later socialists, this emphasis on reform came to represent the essence of democratic socialism; it was this

idea, in fact, that earned the “revisionist” label for this school of socialist thought.

Needless to say, Bernstein was not the only theorist to revise Marx’s ideas on socialism. In the early part of the twentieth century, Vladimir Ilyich Lenin (1971) also amended the concept by adding to it several new notions, some of which were derived from his experience with political organization in tsarist Russia. Taken together, these propositions constitute Russian socialism, also known as Bolshevik theory.

The best known contribution of this school of thought to socialist theory is the idea of the “vanguard party.” According to Lenin, Marx was unduly optimistic in his belief that the proletariat could develop the necessary class consciousness to overthrow capitalism. If left to their own devices, Lenin claimed, workers would defend only their immediate (i.e., economic or trade union) interests and would not know how to translate them into revolutionary action. To assist them in this task, he suggested that a vanguard party of intellectuals must be formed, the task of which would be to develop a revolutionary theory, “go among the masses,” and politically educate the proletariat. From the point of view of Bolshevik theory, therefore, the success of the socialist revolution depends not on the political maturity of the working class but on the strength of the vanguard party.

A second feature of Russian socialism that sets it apart from the Marxist scheme is grounded in its claim that the prospects of a proletarian revolution can arise not only in advanced industrial societies but also in precapitalist economic formations. Given the importance of the vanguard party in Lenin’s version of socialism, this idea makes perfect sense: As long as a country is equipped with a group of willing, dedicated, and professional revolutionaries, it should be able to make the transition to socialism without the benefits of advanced technology or without having passed through the capitalist stage.

Last but not least, Lenin took from Marx the idea that socialism will come in two stages. In terms of his scheme, however, the lower stage (the “dictatorship of the proletariat”) would not be a brief transitional period but would require a whole epoch in human history. During this time, the bourgeois state would be “smashed,” the class rule of the proletariat would be institutionalized, and

opponents of the socialist regime would be suppressed by the “special coercive force” of the proletarian state. The higher state of socialism (“communism”) would be realized once the socialist state had “withered away” and democracy had become a “force of habit.”

Russian socialism constitutes one of many indigenous graftings of the socialist vision. Another well-known attempt in this direction was made by Mao Zedong (1971), who accommodated the idea of socialism to the conditions of a peasant country. Those revisions led to the emergence of what is known as Chinese socialism or Maoism.

Unlike most interpretations of socialism, Mao’s is famous for its glorification of the peasantry. Earlier socialists, among them Marx and Lenin, were skeptical about the revolutionary potential of agricultural laborers. For the most part, they regarded them as inherently petty bourgeois and, consequently, as unlikely allies of the proletariat. Mao argued, however, that in a peasant country such as China, traditionally conceived paths to socialism are not viable because they require the mass mobilization of something that his type of country does not have: an industrial proletariat. He insisted therefore that the socialist revolution in China was a peasant revolution and had no reservations about organizing agricultural workers into a revolutionary force.

Another trademark of Chinese socialism is its lack of confidence in the guaranteed future of socialism. According to Mao’s writings, socialist victories are not everlasting; even as the dust from the revolution begins to settle, old inequalities can resurface and new ones may emerge. For this reason, the work of revolutionaries is never complete: They must be constantly on guard against opposition and be prepared to wage a permanent revolution.

### THE REALITY OF SOCIALISM

During the nineteenth century, a number of communities were established to attempt the realization of the socialist vision, including Etienne Cabet’s Icaria in Illinois, Charles Fourier’s Brook Farm in Massachusetts, William Lane’s New Australia in Paraguay, and Robert Owen’s New Harmony in Indiana. In nearly all these cases, an attempt was made to isolate a small group of dedicated social-

ists from the rest of society and create a model environment for efficient production and egalitarian social exchange. The documented history of these communities suggests that they experienced varying amounts of success (Ross 1935). Some attracted a large number of followers (e.g., Icaria) and prospered for more than a decade (e.g., Brook Farm). Others were fraught with hardships from the beginning (e.g., New Australia), and some collapsed within a few years (e.g., New Harmony). In the end, however, all the utopian experiments failed: They suffered from lack of preparation and meager financial support, harsh living environments and a dearth of agricultural skills, heterogeneous membership, and a lack of long-term commitment to the socialist vision. The individuals who flocked to those communities were sufficiently adventuresome to embark on a project to build a new world but were not prepared for the trials of pioneering.

Experiments with socialism in the twentieth century were more successful and longer-lasting than their utopian counterparts. After the Russian Revolution, 1917–1923, the Soviet Union was the first country to call itself socialist. By the middle of the century, however, there were regimes in Europe, Asia, Latin America, Africa, and the Near East modeling themselves after the Soviet scheme (Hollander 1983). At the risk of oversimplifying, the following traits may be identified as the most important features of those “actually existing” (Bahro 1978) socialist societies: (1) They were characterized by a common ownership of the means of production and distribution. (2) Their economic activities were centrally planned by the state, and market forces played little or no role in the allocation of their resources. (3) One party ruled their political life and legitimated itself by reference to some version of Marxism and Leninism. (4) That party dominated their political culture with a unitary ideology and directed all their executive, legislative, and judiciary powers.

In their purest form, Soviet-type societies have secured a number of major achievements. Within decades of the revolution, they industrialized their outmoded economies (Berend and Ránki 1974), guaranteed full employment and attained price stability (Nove 1989), incorporated women into the labor force (Rueschemeyer and Szelényi 1989), developed their natural resources and advanced science and technology (Nutti 1981), strengthened

their military power (Starr 1988), and improved their educational, health care, and welfare systems (Ferge 1979). Along with those changes, socialist societies made a strong commitment to reducing income, educational, and occupational differentials after World War II (Szelényi 1998). Empirically, a number of studies have shown that those formally egalitarian policies have had impressive results: In nearly all these countries, inequalities in income have decreased (Matthews 1972; Walder 1989), educational opportunities have expanded (Lane 1976), and distinctions of prestige between manual and nonmanual occupations have narrowed (Parkin 1971; Giddens 1973). Policies also were implemented by socialist states to reduce the intergenerational transmission of social inequalities: Inheritance of wealth was eliminated, and quotas were imposed on educational and occupational recruitment to favor children from the working class and from peasant families (Simkus and Andorka 1982; Szelényi and Aschaffenburg 1993). Perhaps in part as a result of these changes, socialist societies carved out for themselves a position of considerable importance in the world system in the twentieth century. In the 1960s, for example, the Soviet Union competed directly with the United States in space exploration, the race for military power, and the development of science, technology, athletics, and the arts.

The economic and social miracles achieved by these countries in the years after World War II could not be sustained, however. By the early 1970s, centrally managed economies began to exhibit multiple signs of strain. Bureaucratic blunders on the part of state officials resulted in poor investment decisions (Nove 1983b), frequent bottlenecks created breakdowns in production (Bauer 1978), chronic shortages of consumer items provoked anger and dissatisfaction among the citizens (Kornai 1986), and curious managerial techniques (in the form of bribing, hoarding, and informal networking) had to be developed to mitigate the ineffective relationship between economic units and the state (Stark 1986).

Problems with central management, of course, were not restricted to the economic sphere. With a growing number of empirical studies during the 1970s (see Hollander 1983), the social and political consequences of Soviet-type planning became evident, although most scholars continued to be impressed by the initially positive outcome of

egalitarian state policies in socialist societies. At the same time, they soon began to realize that the quotas introduced after World War II were often applied inconsistently and in almost all circumstances disturbingly short-lived (Szelényi 1998). It is clear from these studies that the initial attempts to “build socialism” soon were overturned by a “second stage” in socialist development (Kelley and Klein 1986) that was marked by the crystallization of inequalities and the emergence of new privileges (Ossowski 1963; Nove 1983a). By the 1970s, many of those societies began to demonstrate substantial inequalities in their prestige hierarchies (Inkeles 1966), patterns of social mobility (Connor 1979), opportunities for educational attainment (Simkus and Andorka 1982), and distribution of monetary and nonmonetary rewards (Szelényi 1976; Walder 1986).

The political inequalities that characterized Soviet-type societies during their heyday are well documented in the literature. Many studies have shown, for example, that Communist Party functionaries and the so-called *nomenklatura* elite enjoyed definite social, political, and economic advantages: They attended party schools, shopped at special stores, vacationed at the most desirable holiday resorts, and had better access to decision-making posts (Szelényi 1987). In addition to those privileges, they were more likely to receive state-subsidized housing, purchase a car or vacation home, eat meat several times a week, and participate in cultural activities. Such differences in the allocation of goods and resources have led many to conclude that the political sphere was central to the stratification system of socialist societies (Goldthorpe 1966; Bauman 1974). Indeed, some scholars have suggested that the political elite may well have constituted a New (dominant) Class in socialist regimes (Djilas 1957; Konrád and Szelényi 1979).

In light of these problems as well as the apparent failure of the egalitarian experiment, socialist states made a number of attempts to reform their ailing economies. Yugoslavia began this trend by introducing a new economic program that combined free market principles with workers’ self-management; in 1949, Yugoslav leaders abandoned central planning, tied wages to the financial success of firms, and liberalized foreign trade (Sirc 1979). Hungary followed suit in 1968 by introducing its own version of market socialism (Hare et al.

1981), and China joined the trend in the late 1970s with similar economic reforms (Nee 1989).

Partial reprivatization, however, was not the only way for centrally managed economies to embark on the road to recovery. East Germany, for example, refused to combine planning with market reforms and chose to strengthen the operation of its central management (Szelényi 1989). In an effort to “scienticize” economic planning, East German leaders purchased state-of-the-art computers and sophisticated econometric programs to model the behavior of thousands of firms and anticipate the needs of millions of consumers. Cuba also refrained from market reforms in the late 1960s (Leogrande 1981). Hoping to prevent the restoration of capitalism in his country, Fidel Castro argued against the implementation of profit incentives to motivate workers. Instead, he introduced a rigorous political education program, the main purpose of which was to convince workers that they needed to expend maximum effort at work not for personal financial benefit but out of a moral commitment to socialism.

Despite those efforts to revitalize their economies, socialist societies were unable to recover from their experiences with overcentralization. Paradoxically, perhaps, reform plans were applied inconsistently, market rules were not followed rigorously, and the state continued its paternalistic practice of bailing out unsuccessful firms. Meanwhile, political opposition to those regimes continued to grow: Peasants asked for market reforms (Lewis 1979), workers demanded a say in management (Pravda 1979; Kennedy 1991), and intellectuals called for expanded political democracy and protection of their civil rights (Harman 1983). In the spring of 1989, many of those conflicts came to a head as a “gentle revolution” began to unfold in those countries. With a few exceptions, Soviet-type societies formally accepted the principles of multiparty democracy and announced their intention to move in the direction of a market economy.

### THE LEGACIES OF SOCIALISM

If attempts to establish the socialist vision during the twentieth century were fraught with social and economic problems, efforts to undo the structure of existing socialist societies have proved equally challenging. Perhaps the biggest task facing postcommunist societies is to conquer the eco-

nomie legacies of socialism and make the transition to capitalism without the assistance of a capitalist class (Eyal et al. 1998). In this sense, the postcommunist revolution in Central Europe resembles the Russian Revolution. In 1917, a group of intellectuals constituted themselves as a political class in a peasant country to lead a “proletarian revolution” without a proletariat but with the express purpose of creating a proletariat. In 1989, a fraction of the intelligentsia seized power in Central Europe and sought to lead a “bourgeois revolution” without a bourgeoisie but with the objective of creating a bourgeoisie (Szelényi et al. 1995).

Needless to say, this objective was not an easy one. In all formerly socialist countries, the economic infrastructure was poorly developed and arguably deteriorating, the industrial firms of classical socialism were too large to be privatized easily, and the transition to a postindustrial service economy had not progressed very far (Böröcz and Róna-Tas 1995; Volgyes 1995). The distinctive feature of the transition is that despite such seeming homogeneity in the conditions of origin, there was great heterogeneity in the pathways to capitalism. For example, the East German model is one of centrally managed privatization in the context, of course, of West German “colonization.” By contrast, the Czech reformers acquiesced entirely to the “invisible hand” of capitalism, by which all workers were granted vouchers that could be redeemed for shares in any company. Finally, in Hungary, the transformation is best described as a form of “political capitalism” (Hankiss 1990; Staniszkis 1991), by which former communist bureaucrats used their political position to accumulate wealth and buy state companies. Where these privatization strategies will lead remains unclear, but one thing is certain: There is no single plan for designing capitalism, just as there was no simple blueprint for establishing socialism (Stark 1992).

Although most discussions of the transition to postcommunism have focused on the economic legacies of socialism, the political legacies are no less problematic as some form of successful marketization is sought. There are two political legacies of particular interest here. First, after forty years of communist rule and rampant political deception, the reigning view among East European workers involved considerable cynicism toward political elites, and such deep-seated cynicism

could not be overcome immediately even when new leaders were vying for power (Kovrig 1995). This cynicism undermined popular support for long-term sacrifices of the sort that all marketization strategies would necessarily entail. Second, the concept of marketization was not completely endorsed by the general population, as there was a long heritage of support for state paternalism in which basic needs, such as health care, education, and a living wage, were guaranteed (Szelényi et al. 1996). There was also widespread concern that marketization would increase inequality to levels that were unacceptably high. It has to borne in mind, then, that the transition to a market economy was undertaken simultaneously with a transition to political democracy. Democratic regimes, for all their possible virtues, are not necessarily well suited for revolutionary economic transformations and the popular sacrifice that such transformations typically imply (O'Donnell and Schmitter 1986).

Finally, the emergence of postcommunism is further complicated by the social legacies of communism. Most notably, the transition to the high-unemployment economy of postcommunism created special problems of legitimacy for a new market society, since Central Europeans had come to expect full employment from the state (Moskoff 1994). Similarly, one of the great successes of communism was its low levels of income inequality (relative to the capitalist alternative), and consequently, the sudden and visible increases in inequality in the postcommunist world were not readily accepted. It was all the more problematic that the prime beneficiaries of this growing wealth were in some instances the former communist elites themselves (Róna-Tas 1994; Fodor et al. 1995; Szelényi and Szelényi 1995). For all its economic failings and political repression, actually existing socialism was at least partially consistent with the original vision of social egalitarianism, and one cannot expect such success to be relinquished without a struggle.

### THE FUTURE OF SOCIALISM

The question remains: Can the *idea* of socialism survive the *reality* of the past eighty years? For some, the answer to this question is in the negative, as the failings of socialism are so dramatic that the concept of socialism is inextricably associated with its particular realization, thus rendering it

effectively dead for all of history (Jowitt 1992). This is, then, a peculiar form of path dependency in which the possibly premature turn to socialism in the early twentieth century proved in the end to be its historical downfall. As a fallback position, one might argue that while socialism is perhaps dead in all the countries that experienced its "grief and shame" (Djilas 1998), it might nonetheless surface anew in countries that never underwent this premature experiment. Is there, in other words, a viable base for socialism in the Western world? The standard postmaterialist position on this score is that the base for socialism was at its strongest in the early twentieth century but has since dissipated with the decline in the size of the working class, the weakening of trade unions, and the associated rise of interest politics focusing on issues such as the environment, nuclear war, and gender politics (Inglehart 1983; Piven 1992). The implication is that socialism is dead not because of its tarnished history but because there is no longer a substantial base of working-class supporters.

This line of reasoning, for all its appeal, is not easily reconciled with the continuing support for social democratic policies and communist political leaders in formerly socialist countries. In many formerly communist societies, the initially extreme anticommunist sentiment weakened quickly, and the Communist Party was returned to power in the "second round" elections (Szelényi et al. 1996). Moreover, public opinion polls in those countries consistently reveal that the general population remains supportive of fundamentally social democratic policies even while disavowing support for highly repressive forms of communism of the sort that characterized Soviet-type societies. Under this formulation, a more mature civil society is in formation that probably will pursue a "Swedish form" of social democracy that maintains some elements of classical socialism (i.e., economic egalitarianism) yet abandons others (i.e., political inequality and repression).

### REFERENCES

- Bahro, Rudolf 1978 *The Alternative in Eastern Europe*. London: New Left Books.
- Bauer, T. 1978 "Investment Cycles in Planned Economies." *Acta Oeconomica* 21(3):243-260.
- Bauman, Zygmunt 1974 "Officialdom and Class: Basis of Inequality in Socialist Society." In Frank Parkin,

- ed., *The Social Analysis of Class Structure*. London: Tavistock.
- Bell, Daniel. 1968. "Socialism." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*, vol. 14. New York: Macmillan and Free Press.
- Berend, Iván T., and György Ránki 1974 *Economic Development in East-Central Europe in the Nineteenth and Twentieth Centuries*. New York: Columbia University Press.
- Bernstein, Eduard 1961 *Evolutionary Socialism*, Edith C. Harvey, trans. New York: Schocken.
- Böröcz, József, and Ákos Róna-Tas 1995 "Small Leap Forward: Emergence of New Economic Elites." *Theory and Society* 24:751–781.
- Cabet, Etienne 1975 *History and Constitution of the Icarian Community*, Thomas Teakle, trans. New York: AMS Press.
- Cole, George D. H. 1959 *A History of Socialist Thought*, 5 vols. London: Macmillan.
- Connor, Walter D. 1979 *Socialism, Politics, and Equality*. New York: Columbia University Press.
- Djilas, Milovan 1957 *The New Class: An Analysis of the Communist System*. New York: Harcourt Brace Jovanovich.
- 1998. *The Fall of the New Class: A History of Communism's Self-Destruction*. New York: Knopf.
- Eyal, Gil, Iván Szelényi, and Eleonor Townsley 1998 *Making Capitalism Without Capitalists: The New Ruling Elites in Eastern Europe*. London: Verso.
- Ferge, Zsuzsa 1979 *A Society in the Making*. New York: M. E. Sharpe.
- Fodor, Éva, Edmund Wnuk-Lipinski, and Natasha Yershova 1995 "The New Political and Cultural Elites in Hungary, Poland, and Russia." *Theory and Society* 24:783–800.
- Fourier, Charles 1971 *Harmonian Man: Selected Writings of Charles Fourier*, Mark Poster, ed.; Susan Hanson, trans. Garden City, N.Y.: Doubleday.
- Garton Ash, Timothy 1990 *The Uses of Adversity: Essays on the Fate of Central Europe*. New York: Vintage.
- Giddens, Anthony 1973 *The Class Structure of the Advanced Societies*. London: Hutchinson.
- Goldthorpe, John H. 1966 "Social Stratification in Industrial Society." In Reinhard Bendix and Seymour Martin Lipset, eds., *Class, Status, and Power*. New York: Free Press.
- Gray, Alexander 1963 *The Socialist Tradition: Moses to Lenin*. London: Longmans.
- Griffith, D. F. 1924 *What Is Socialism? A Symposium*. London: Richards.
- Hankiss, Elemér 1990 *East European Alternatives*. Oxford, UK: Clarendon Press.
- Hare, P., P. Radice, and Nigel Swain 1981 *Hungary: A Decade of Economic Reform*. London: Allen and Unwin.
- Harman, Chris 1983 *Class Struggles in Eastern Europe, 1945–1983*. London: Pluto Press.
- Hollander, Paul 1983 *The Many Faces of Socialism*. New Brunswick, N.J.: Transaction.
- Inglehart, Ronald 1983 "The Persistence of Materialist and Postmaterialist Value Orientations." *European Journal of Political Research* 11:81–91.
- Inkeles, Alex 1966 "Social Stratification and Mobility in the Soviet Union." In Reinhard Bendix and Seymour Martin Lipset, eds., *Class, Status, and Power*. New York: Free Press.
- Jowitt, Ken 1992 *New World Disorder*. Berkeley: University of California Press.
- Kelley, Jonathan, and Herbert S. Klein 1986 "Revolution and the Rebirth of Inequality: Stratification in Post-Revolutionary Society." In Jack A. Goldstone, ed., *Revolutions*. San Diego, Calif.: Harcourt Brace Jovanovich.
- Kennedy, Michael D. 1991 *Professionals, Power, and Solidarity in Poland*. Cambridge, UK: Cambridge University Press.
- Kolakowski, Leszek 1978 *Main Currents of Marxism*, 3 vols., P. S. Falla, trans. Oxford, UK: Oxford University Press.
- Konrád, György, and Iván Szelényi 1979 *The Intellectuals on the Road to Class Power*, Andrew Arato and Richard E. Allen, trans. New York: Harcourt Brace Jovanovich.
- Kornai, János 1986 *Contradictions and Dilemmas: Studies on the Socialist Economy and Society*. I. Lukacs, J. Parti, B. McLean, and G. Hajdú, trans. Cambridge, Mass.: MIT Press.
- Kovrig, Bennett 1995 "Marginality Reinforced." In Zoltan Barany and Ivan Volgyes, eds., *The Legacies of Communism in Eastern Europe*. Baltimore: Johns Hopkins University Press.
- Laidler, Harry W. 1968 *History of Socialism*. New York: Crowell.
- Lane, David 1976 *The Socialist Industrial State: Towards a Political Sociology of State Socialism*. London: Allen and Unwin.
- Lenin, V. I. 1971 *Selected Works*. New York: International Publishers.
- Leogrande, William 1981 "Republic of China." In Bogdan Szajkowski, ed., *Marxist Governments*, vol. 2. New York: St. Martin's.
- Lewis, Paul G. 1979 "Potential Sources of Opposition in the East European Peasantry." In Rudolf Tökés, ed., *Opposition in Eastern Europe*. Baltimore: Johns Hopkins University Press.

- Lichtheim, George 1969 *The Origins of Socialism*. New York: Praeger.
- Mao Zedong 1971 *Selected Readings from the Works of Mao Zedong*. Peking: Foreign Language Press.
- Marx, Karl, and Frederick Engels 1968 *Selected Works*. Moscow: Progress Publishers.
- Matthews, Mervyn 1972 *Class and Society in Soviet Russia*. New York: Walker.
- Morris, William 1970 *News from Nowhere*. London: Routledge and Kegan Paul.
- Moskoff, William 1994 "Unemployment in the former Soviet Union." In James R. Millar and Sharon L. Wolchik, eds., *The Social Legacies of Communism*. Washington, DC: Woodrow Wilson Center Press.
- Nee, Victor 1989 "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54: 663–681.
- Nove, Alec 1983a "The Class Nature of the Soviet Union Revisited." *Soviet Studies* 3:298–312.
- 1983b *The Economics of Feasible Socialism*. London: Allen and Unwin.
- 1989 *An Economic History of the USSR*, 2nd ed. Harmondsworth, UK: Penguin.
- Nuti, Domenico Mario 1981 "Socialism on Earth." *Cambridge Journal of Economics* 5:391–403.
- O'Donnell, Guillermo, and Philippe C. Schmitter 1986 *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore: Johns Hopkins University Press.
- Ossowski, Stanislaw 1963 *Class Structure in the Social Consciousness*. London: Routledge and Kegan Paul.
- Parkin, Frank 1971 *Class Inequality and Political Order*. New York: Praeger.
- Piven, Frances Fox 1992 "The Decline of Labor Parties: An Overview." In F. F. Piven, ed., *Labor Parties in Postindustrial Societies*. New York: Oxford University Press.
- Pravda, Alex 1979 "Industrial Workers: Patterns of Dissent, Opposition, and Accommodation." In Rudolf Tökés, ed., *Opposition in Eastern Europe*. Baltimore: Johns Hopkins University Press.
- Proudhon, Pierre Joseph 1966 *What Is Property? An Inquiry into the Principle of Right and of Government*, B. R. Tucker, trans. New York: H. Fertig.
- Róna-Tas, Ákos 1994 "The First Shall Be Last? Entrepreneurship and Communist Cadres in the Transition from Socialism." *American Journal of Sociology* 10:40–69.
- Ross, Lloyd 1935 *William Lane and the Australian Labor Movement*. Sydney, Australia: Forward Press.
- Rueschemeyer, Marilyn, and Szonja Szelényi 1989 "Socialist Transformation and Gender Inequality: Women in the GDR and Hungary." In David Childs, Thomas A. Baylis, and Marilyn Rueschemeyer, eds., *East Germany in Comparative Perspective*. London: Routledge.
- Saint-Simon, Claude H. 1964 *Social Organization, the Science of Man, and Other Writing*, Felix M. H. Markham, ed. and trans. London: Harper and Row.
- Simkus, Albert A., and Rudolf Andorka 1982 "Inequalities in Educational Attainment in Hungary, 1923–1973." *American Sociological Review* 47:740–751.
- Sirc, Ljubo 1979 *The Yugoslav Economy under Self-Management*. London: Macmillan.
- Staniszki, Jadwiga 1991 *The Dynamics of Breakthrough in Eastern Europe*. Berkeley: University of California Press.
- Stark, David 1986 "Rethinking Internal Labor Markets: New Insights from a Comparative Perspective." *American Sociological Review* 51:492–504.
- 1992 "Path Dependence and Privatization: Strategies in East Central Europe." *East European Politics and Societies* 6:17–54.
- Starr, Richard F. 1988 *Communist Regimes in Eastern Europe*, fifth ed., Stanford, Calif.: Hoover Institution.
- Swezy, Paul M. 1983 "Socialism." In Tom Bottomore, ed., *A Dictionary of Marxist Thought*. Cambridge, Mass.: Harvard University Press.
- Szelényi, Iván 1976 "The Housing System and Social Structure in Hungary." In Bernard Lewis Faber, ed., *The Social Structure of Eastern Europe*. New York: Praeger.
- 1989 "Eastern Europe in an Epoch of Transition: Toward a Socialist Mixed Economy?" In Victor Nee and David Stark, eds., *Remaking the Economic Institutions of Socialism*. Stanford, Calif.: Stanford University Press.
- , and Szonja Szelényi 1995 "Circulation or Reproduction of Elites during Post-Communist Transformation in Eastern Europe: Introduction." *Theory and Society* 24:615–638.
- Szelényi, Szonja 1987 "Social Inequality and Party Membership: Patterns of Recruitment into the Hungarian Socialist Workers' Party." *American Sociological Review* 52:559–573.
- 1998 *Equality by Design: The Grand Experiment in De-stratification in Socialist Hungary*. Stanford, Calif.: Stanford University Press.
- , and Karen E. Aschaffenburg 1993 "Inequalities in Educational Opportunity in Hungary." In Yossi Shavit and Hans-Peter Blossfeld, eds., *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview.

- , Iván Szelényi, and Imre Kovách 1995 “The Making of the Hungarian Postcommunist Elite: Circulation in Politics, Reproduction in the Economy.” *Theory and Society* 24:697–722.
- , —, and Winifred R. Poster 1996 “Interests and Symbols in Post-Communist Political Culture: The Case of Hungary.” *American Sociological Review* 61:466–477.
- Volgyes, Ivan 1995 “The Economic Legacies of Communism.” In Zoltan Barany and Ivan Volgyes, eds., *The Legacies of Communism in Eastern Europe*. Baltimore: Johns Hopkins University Press.
- Walder, Andrew G. 1986 *Communist Neo-Traditionalism: Work and Authority in Chinese Industry*. Berkeley: University of California Press.
- 1989 “Social Change in Post-Revolutionary China.” *American Review of Sociology* 15:405–424.

SZONJA SZELÉNYI

## SOCIALIZATION

Socialization has had diverse meanings in the social sciences, partly because a number of disciplines claim it as a central process. In its most common and general usage, the term “socialization” refers to the process of interaction through which an individual (a novice) acquires the norms, values, beliefs, attitudes, and language characteristic of his or her group. In the course of acquiring these cultural elements, the individual self and personality are created and shaped. Socialization therefore addresses two important problems in social life: societal continuity from one generation to the next and human development.

Different disciplines have emphasized different aspects of this process. Anthropologists tend to view socialization primarily as cultural transmission from one generation to the next, sometimes substituting the term “enculturation” for socialization (Herskovits 1948). Anthropological interest in socialization or enculturation coincided with the emergence of the “culture and personality” orientation of the late 1920s and 1930s, when the works of Mead (1928), Benedict (1934), and Malinowski (1927) focused on cultural practices affecting child rearing, value transmission, and personality development and helped shape the anthropological approach to socialization. Much

of the work in the culture and personality field was influenced by psychoanalytic theory. Contemporary cultural anthropology is guided less by psychoanalytic theory and more by social constructionist theories (such as symbolic interactionism), which view socialization as a collective and interpretive process of reality construction involving the reproduction of culture. This orientation has been shaped largely by the work of Geertz (1973), whose influence is also evident in sociological work on socialization, such as that of Corsaro and Eder (1995).

Psychologists are less likely to emphasize the transmission of culture and more likely to emphasize various aspects of individual development (Goslin 1969). There is considerable diversity within psychology in regard to the aspect of socialization studied. For developmental psychologists, particularly those influenced by Piaget (1926), socialization is largely a matter of cognitive development, which typically is viewed as a combination of social influence and maturation. For behavioral psychologists, socialization is synonymous with learning patterns of behavior. For clinical psychologists and personality theorists, it is viewed as the establishment of character traits, usually within the context of early childhood experiences. The subfield of child development is most closely associated with the topic of socialization within psychology, where socialization is largely equated with child rearing (Clausen [1968] provides a historical overview of socialization in these disciplines).

Political science has shown some interest in socialization, but in a limited sense. Its studies have not gone much beyond political socialization: the process by which political attitudes and orientations are formed. However, a different and more esoteric use of the term occasionally appears in this literature: socialization as “collectivization,” that is, the transformation of capitalism to socialism and/or communism.

Within sociology, there have been two main orientations toward socialization. One views socialization primarily as the learning of social roles. From this perspective, individuals become integrated members of society by learning and internalizing the relevant roles and statuses of the groups to which they belong (Brim 1966). This view has been present in some form from the beginnings of sociology as a discipline but has



been most closely associated with structural-functional perspectives.

The other, more prevalent sociological orientation views socialization mainly as self-concept formation. The development of self and identity in the context of intimate and reciprocal relations is considered the core of socialization. This view is closely associated with the symbolic interactionist perspective, a synthesis of various strands of pragmatism, behaviorism, and idealism that emerged in the 1920s and 1930s in the writings of a number of scholars at the University of Chicago, especially Mead (1934). In Mead's writings, the self is a reflexive, thoroughly social phenomenon that develops through language or symbolic interaction. Language enables the development of role-taking, by which the individual is able to view himself or herself from the perspective of another person. This becomes the basis for selfhood and the interpenetration of self and society. Mead and other symbolic interactionists have argued that self and society are two sides of the same coin. The basis for their assertion is that the content of self-conceptions (e.g., identities) reflects the aspects of the social process with which the individual is involved through the internalization of role identities, values, and meanings. This internalization in turn reproduces society. From the interactionist perspective, both self and society depend on the same process of social interaction by which "realities" are created and constantly negotiated (Gecas 1982).

For contemporary interactionists as well, socialization is distinguished from other types of learning and other forms of social influence by its relevance for self-conceptions, that is, for people's thoughts and feelings about themselves. As such, socialization is not merely the process of learning rules or norms or behavior patterns; it is a matter of learning these things only to the extent to which they become part of the way people think of themselves. The mark of successful socialization is the transformation of social control into self-control. This is accomplished largely through the development of identities, the various labels and characteristics attributed to the self. Commitment to identities (such as son, mother, professor, honest person) is a source of motivation for individuals to act in accordance with the values and norms implied by those identities (Foote 1951; Stryker

1980; Gecas 1986). The focus on identity also emphasizes the membership component of socialization: To be socialized is to belong to a social group.

Socialization as identity formation occurs through a number of more specific processes associated with self-concept development: reflected appraisals, social comparisons, self-attributions, and identification (Gecas and Burke [1995] and Rosenberg [1979] discuss these processes). Reflected appraisals, based on Cooley's (1902) "looking-glass self" metaphor, refer to people's perceptions of how others see and evaluate them. To some extent people come to see themselves as they think others (particularly significant others) see them. People also develop conceptions of themselves with regard to specific attributes by comparing themselves to others (social comparisons) and making self-inferences from observing their own actions and their consequences (self-attributions). Particularly important to socialization as identity formation is the process of identification. Initially used by Sigmund Freud, this concept refers to the child's emotional attachment to the parent and desire to be like the parent; as a consequence, the child internalizes and adopts the parent's values, beliefs, and other characteristics. Among other things, through identification with the parent, the child becomes more receptive to parental influence.

Identification also is used to refer to the imputation or ascription of identities. Here the focus is on the establishment of identities in social interaction, which is an important aspect of defining situations and constructing realities. This also has important socializing consequences, as much of the literature on labeling, stereotyping, and expectancy effects attests.

#### CONTENT AND CONTEXTS OF SOCIALIZATION

Much research on socialization has been concerned with identifying the aspects of the socializee's development that are affected by particular agents and contexts of socialization and through particular processes. The focus has been primarily on the family context, in which the initial or *primary* socialization of the individual takes place. Studies of child rearing in "normal" as well as "abnormal" situations (e.g., institutionalized children, "closet children" feral, children") have identified a num-

ber of conditions that must be present for primary socialization to take place, that is, for the child to become a person. These conditions include the use of symbolic interaction (language) in the context of an intimate, nurturant relationship between an adult and a child. These conditions are necessary for the initial sense of self to emerge and for normal cognitive and even physical development to take place. The claim that the family (in some form) is a universal feature of human societies is based in large part on this important socialization function.

Parental support continues to be important in the socialization of offspring through childhood, adolescence, and beyond. It is one of the most robust variables in the literature on child rearing. Parental support has been found to be positively related to a child's cognitive development, moral behavior, conformity to adult standards, self-esteem, academic achievement, and social competence. Conversely, lack of parental support is associated with negative socialization outcomes for children and adolescents: low self-esteem, delinquency, deviance, drug use, and various other problem behaviors (Rollins and Thomas 1979; Peterson and Hann 1999).

Parental control is almost as prominent as support in the socialization literature. "Control" refers to the degree to and the manner in which parents attempt to place constraints on a child's behavior. Other terms used for this dimension of parenting are punishment, discipline, restrictiveness, permissiveness, protectiveness, supervision, strictness, and monitoring. Parental control is a more complicated variable than is parental support. It is necessary to distinguish different types or styles of control because they frequently have opposite socialization consequences. An important distinction is that between "authoritarian" and "authoritative" control (Baumrind 1978) or "coercion" and "induction" (Rollins and Thomas 1979). Authoritarian or coercive control (control based on force, threat, or physical punishment) is associated with negative or unfavorable socialization outcomes, whereas authoritative or inductive control (control based on reason and explanation) has positive outcomes.

The most powerful models of parental influence in the socialization of children are those which combine the dimensions of support and

control. Parents are most effective as agents of socialization when they express a high level of support and exercise inductive control. In these conditions, children are most likely to identify with their parents, internalize parental values and expectations, use parents as their models, and become receptive to attempts at parental influence. Conversely, low parental support and reliance on coercive control are associated with unfavorable socialization outcomes (for reviews of this literature, see Peterson and Rollins 1987; Maccoby and Martin 1983; and Rollins and Thomas 1979).

Parental support and control cover much of the ground in the research on child rearing but not all of it. Other important socialization variables here are extent of parental involvement with the child (e.g., time spent), level of performance expectations, extent to which political or religious beliefs and value systems are taught to the child by the parent, and various characteristics of the parent, such as patience, tolerance, honesty, integrity, competence, and age and sex (of parent and child). Many factors affect the process and outcomes of family socialization.

Much of the socialization that takes place in the family involves learning appropriate role behavior associated with the various family positions. For the child, the most significant of these behaviors involve sex and age roles. Through processes of reinforcement from parents and others, identification with various role models, and parental admonitions and instructions, a child is socialized into the behavioral expectations associated with these roles. Of the two, sex roles have received more of the research attention on role learning in the family (Block [1983] provides a review). This research suggests that sex-role socialization is extensive (usually starting at birth with differential treatment of male and female infants), pervasive (various agents and contexts of socialization), and consequential for a wide range of other individual and social outcomes. A prominent theme in much contemporary research on sex-role socialization is that the differential treatment that emphasizes "masculine" characteristics for boys and "feminine" characteristics for girls is detrimental to the development of both girls and boys and to the relationship between the sexes (Bem [1974] discusses the virtues of androgyny). This research reflects the ethos of equality between the sexes in most modern societies.

Most studies of socialization within the family assume a unidirectional influence from parent to child. Parents typically are viewed as agents of socialization (part of the job description of a parent), and children as objects of socialization. Given the disparities in power, status, and competence between parent and child, it is justifiably assumed that the direction of influence is mainly from parent to child. However, it has become increasingly evident that socialization is a reciprocal process, with children influencing parents as well. Over the past few decades, the thinking with regard to socialization processes has shifted from unidirectional to bidirectional and reciprocal models (Corsaro and Eder 1995; Gecas 1981). For example, in considering the association between parental punishment and a child's deviant behavior, which is one of the most consistent findings in socialization research, it can be argued that the child's behavior is both a consequence and a cause of the parental behavior. That is, a child's aggressive or deviant behavior may *elicit* more punitive parental behavior as well as being affected by the parental behavior. Socialization increasingly is viewed as reciprocal, even though the degree of influence is typically not equal.

Besides parents and other adult kin, siblings serve as agents of socialization within the family context. As family size increases, more of the socialization of the younger children is taken on by their older siblings, either by default or because the parents delegate this responsibility to the older children. Some have argued that this puts younger children in large families at a disadvantage with regard to cognitive development, since they have relatively less contact with the most competent and committed family members, the parents (Zajonc 1976). However, these findings, based mostly on cross-sectional data, have not gone unchallenged (Galbraith 1982; Blake 1989).

An increasingly pervasive agent of socialization in contemporary families is television. Children spend more time watching television than at any other activity except school and sleep (Bronfenbrenner 1970). The purpose of most television programs children watch is typically not to socialize or educate but to entertain and sell products. However, a good deal of unintended socialization is likely to occur, from shaping conceptions of reality (e.g., sex roles and ethnic stereotypes) to styles of behavior and tastes. In general, television

is perceived as having a negative influence on children, with the exception of a few educational programs on public television. Much of the concern has focused on the extensive violence and sexual themes and situations in television programs. Bandura et al.'s (1963) work on modeling has persuasively shown that exposure to aggressive behavior tends to increase aggression in the viewer. Along with its undesirable consequences for child socialization, Bronfenbrenner (1970) observes that television is detrimental to child development with regard to the behavior it prevents, that is, the human interaction that is forgone in the course of being a passive viewer. The role of television as an agent of socialization in families seems to be increasing by default as the amount of contact between parents and their children decreases. Various social forces (such as increasing numbers of working mothers, single-parent families, dual-career families, and the professionalization of child care) have decreased the amount of parent-child interaction and thus parents' role as a socializing agent. This vacuum has been filled increasingly by the child's involvement with television and with peers. For children in American society, television, peer groups, and school are increasingly important agents of socialization.

Like the family, the school is an institution whose mandate is to socialize children. The school's mission, however, is more narrowly defined than is that of the family and is concerned primarily with the formal instruction and the development of children's cognitive skills. In this sense, the school context is less involved in *primary* socialization (i.e., the development of basic values, beliefs, motivations, and conceptions of the self) and more involved in *secondary* socialization (i.e., the development of knowledge and skills). This is not a very precise distinction, however. In the course of the socialization experienced in school, things other than skills and knowledge also are learned, such as norms, values, attitudes, and various aspects of a child's personality and self-concept. Much more is typically learned in school than what is explicitly taught.

Many activities associated with school (specifically in the classroom) have implications for a child's self-concept (Hewitt 1998). For example, one of the most important activities involves evaluation of the student's performance by the teacher:

performance on tests, class reports, presentations, assignments, and the like. Success in these activities, based on one's own efforts, is good for self-esteem and builds confidence in one's abilities. However, failure is not, and *public* failure is even worse. School provides numerous opportunities for public failure as well as public success. One of the consequences of performance evaluations may be the categorization or "labeling" of students, by teachers as well as others, as "smart," "dumb," "slow learner," "underachiever," and so on. Negative as well as positive labels affect the way in which others respond to a person and, through their responses, reinforce and shape that person in the labeled direction. This process is called "expectancy effects" (Jones 1977) or "self-fulfilling prophecy" (Merton 1957). Rosenthal and Jacobson (1968) found that teachers' expectations of students, even when based on erroneous information, had a significant effect on how students developed over the course of the school year: When the teacher was led to believe that a student would be a "slow learner," that student was more likely to do poorly in class. Labeling and expectancy effects occur in most socialization contexts and have important consequences for self-concept development.

However, students, like other socializees, are not passive recipients of the pressures they experience. Covington and Beery (1976) propose that two fundamentally different motivation patterns emerge in schools as a result of these pressures: One is oriented toward striving for success, and the other toward avoiding failure. Failure-avoiding strategies (such as nonparticipation, withdrawal, procrastination, and putting off work assignments until too late) are attempts to disassociate one's performance from one's ability and worth. Failure then can be attributed to lack of effort or to external circumstances (less damaging attributions for the self), not to lack of ability (a more damaging attribution). This is a form of role distancing, the separation of the self from the behavior required of a role occupant; it is also an obstacle to school achievement. As Covington and Beery (1976) point out, failure-avoiding strategies are self-defeating: In their attempts to avoid feelings of failure, these students increase the probability of actual failure. For some students this is one of the unintended and undesirable consequences of classroom socialization. In the process of socializing students toward achievement and mastery (desir-

able outcomes), pressures are generated that may result in undesirable adaptations.

The third most important context for the socialization of children and adolescents is the peer group. In terms of structure and function, the peer group is a very different context from family and school. Unlike those two contexts, it is not the "job" of peers to socialize each other, even though a great deal of childhood socialization occurs in this context, some of it in reaction against the socialization experienced in the family and school.

There are several important features of the peer group as a context of socialization. Most important, it is a *voluntary* association, and for most children it is the first. This permits greater freedom of choice regarding associations in the group. A second important feature is that association is between status equals. Consequently, interaction is more likely to be based on egalitarian norms. Status distinctions emerge, of course, but are more likely to be based on achievement and negotiation. However, the basic relationship within peer groups is not hierarchical; rather, it is the friendship bond, based on equality, mutual tolerance, and concern. Third, the peer group is an arena for the exercise of independence from adult control. As such, it is often the context for the development of values, norms, and behavior in opposition to those of adults (such as the subcultures described by Coleman [1961] and in much of the literature on juvenile delinquency). Fourth, children's peer groups, in contemporary American society at least, typically are segregated by sex and differ in organizational patterns: Girls' peer groups tend to feature closely knit and egalitarian friendships, whereas boys' peer groups tend to be loosely knit, larger groups with clear status hierarchies. An important socialization consequence of intensive association with same-sex peers and involvement in sex-typed activities is that this strongly reinforces identification and belongingness with members of the same sex and contributes to the development of stereotypical attitudes toward members of the opposite sex. Not only sex-role identity but also much of sexual socialization during childhood occur in the context of peer rather than parent-child associations, since parents are much less interested in discussing sexual matters with their children than are the children's peers (Fine 1987; Corsaro and Eder 1990). Peers provide an alternative reference group for children as well as

an alternative source of self-esteem and identity. For these reasons, attachment to peers may be even stronger than attachment to family, especially for adolescents.

The socialization experienced by adults generally falls in the category of secondary socialization, building on the socialization experiences of childhood. Much of this is role-specific (Brim 1968), that is, learning the knowledge and skills required for the performance of specific adult roles, such as occupation, marriage, and parenthood. As individuals become committed to the roles they play, they come to identify themselves and think of themselves in terms of these role-identities (Stryker 1980).

Since work is a dominant activity and setting for most adult men and women, much of adult socialization involves either preparation for an occupation or career (which usually takes place specialized schools or training programs such as law school, medical school, and college) or on-the-job training. The work setting can have a substantial socializing effect on workers, affecting more than just their knowledge and skills. Kohn and Schooler (1983) have shown how certain occupational conditions affect the development of a worker's values and personality. Specifically, they found that work that is routine, closely supervised, and relatively uncomplicated gives rise to values of conformity, whereas work that is complex and encourages self-direction increases the value workers place on independence and autonomy. Kanter (1977) found that the nature of work relations, particularly the structure of opportunity on the job, affects workers' attitudes and behaviors as a consequence of their adaptations to the work situation. Workers' adaptations to their work situations do not necessarily lead to commitment to the job or self-investment in terms of the occupational role. On the contrary, a prevalent theme in much of the sociological literature on work and workers (especially that with a Marxist perspective) deals with the alienating consequences of work in capitalist societies.

Many other contexts have socializing consequences for adults: family, political and religious organizations, recreational settings, and voluntary associations. The socialization that takes place in these contexts can be considered "developmental" (Wheeler 1966) because it builds on previous

socialization and is a continuation and expansion of past socialization experiences. *Resocialization* refers to socialization experiences that represent a more radical change in the person. Resocialization contexts (e.g., mental hospitals, some prisons, reform schools, therapy groups, political indoctrination camps, religious conversion settings) have as their explicit goal the transformation of the individual. An important feature of resocialization is the replacement of one's previous set of beliefs, values, and especially conceptions of the self with a new set grounded in the socializing group's ideology or world view. This has been described as a process of death and rebirth of the self (Lifton 1963). Typically, this is accomplished through intense small group interaction in which the physical and symbolic environments are highly controlled by the agents of socialization. It is an experience that usually involves considerable stress for the socializee.

#### **SOCIALIZATION OVER THE LIFE COURSE**

Socialization is a lifelong process of change. Even though the socialization experienced in the family is in some ways the most consequential, individuals typically have important socializing experiences throughout their lives. A central theme in the life-course literature is the degree of continuity and consistency in personality as an individual moves through the life course. Positions on this issue range from the claim that personality is shaped largely during early childhood (most evident in psychoanalytic theories) to the claim that people are thoroughly malleable, changing across situations and throughout their lives (characteristic of constructivist theories). The majority appear to argue for an intermediate position, maintaining that the "core" personality or self-concept develops in early socialization experiences, while various other characteristics are added to self through the acquisition of new roles, identities, and socializing experiences (Brim 1966). For example, Clausen (1993) found that the development of "planful competence" during the childhood and early adolescent years affected the life course of adults many years later, resulting in individual histories of cumulative advantage or disadvantage. The previous discussion suggests how contexts of socialization, which are typically age-graded, can contribute to the development of different aspects of

individuals associated with different ages and stages of life.

Some important socializing experiences and changes are keyed to developmental or maturational considerations: There are differences between the concerns and capabilities of children and those of adolescents, those of young adults and those of people in middle age, and those of the middle-aged and those of older persons. Erikson's (1959) developmental scheme, building on the Freudian theory of psychosexual development but extending it beyond childhood, emphasizes the different developmental tasks associated with different stages of life. The challenges or developmental tasks proposed by Erikson are (1) trust versus mistrust, (2) autonomy versus shame, (3) initiative versus guilt, (4) industry versus inferiority, (5) identity versus identity confusion, (6) intimacy versus isolation, (7) generativity versus self-absorption, and (8) integrity versus despair. Most of the socialization research guided by Erikson's formulations has focused on stage 5, adolescence, and the developmental task highlighting identity concerns. In modern society adolescence has long been considered a time when self-concept concerns increase in prominence. Physiological changes and changes in social circumstances (e.g., high school, dating, career considerations) contribute to an increase in self-awareness and concern about how one is viewed by others. Research by developmental psychologists generally has found that good family relations (those high in parental support, communication, involvement, and inductive control) facilitate the development of ego identity in adolescence (Gecas and Seff [1990] and Steinmetz 1999 provide reviews). By adolescence, however, the influence of parents is substantially less than it was during childhood. Increasingly, other agents and contexts of socialization become important to the adolescent: peers, school, friends, coaches, and so forth. The adolescent's struggles with identity are worked out in a number of competing arenas.

Identity concerns are not limited to adolescence, of course. If one considers socialization a lifelong process of self-concept formation, matters of identity are important at various stages of the individual's development. Identity concerns are most likely to be accentuated during periods of transition, particularly those involving entrance into or exit from social statuses and roles. Some of

these role transitions are institutionalized and highly ritualized. The rites of passage in various cultures marking the transition from childhood to adulthood can be elaborate and dramatic. Sometimes this involves acquiring a new name as well as a new status (as in many of the Plains Indian cultures). In contemporary Western societies, these status passages may be less dramatic but still quite consequential for the person: getting a driver's license, high school graduation, marriage, divorce, the first full-time job, retirement, widowhood. In general, each major transition initiates a new socializing experience or situation that has implications for the individual's self-concept.

Some theorists have focused on transitions in adulthood (Levinson 1978; Levinson and Crumpler 1996) and examined the circumstances that can lead to a "midlife crisis" an acute reexamination of the self. Evidence that such a crisis typically occurs at midlife is sparse. The adult years, especially in the later stages, are still relatively neglected by life-course scholars compared to studies of childhood and adolescence, but this may be changing. As longevity continues to increase in modern society, so will concern with socializing experiences in the later stages of life. We may be expanding some stages (e.g., postretirement and widowhood) as well as creating new ones, such as the "nursing home stage," as the life span increases.

In considering socialization over the life course, it is necessary to take history and culture into account. Not just the content of socialization during various "stages" of life but also the stages themselves vary with the culture and the historical context. For example, adolescence as an identifiable stage of life is a relatively recent historical construct in Western societies, closely associated with the extension of formal education to high school (Gecas and Seff 1990). Even childhood, as Aries (1960) documents, is not universally considered a distinct stage of life. The modern conception of childhood as an identifiably distinct stage emerged during the European Renaissance, partly as a consequence of the emergence of parochial schools. More recently, Elder (1974) has shown a historical consciousness in his life-course analyses by examining how specific historical events (e.g., the Great Depression; World War II) differentially affected two cohorts of children and their families. Whether there are eight stages of life, four stages,

or seven stages (as Shakespeare observed) depends on the society and one's analytic purposes.

### CONTEMPORARY ISSUES AND THEMES IN SOCIALIZATION RESEARCH

For much of its history, the concept of socialization has been heavily imbued with the notion of adaptation and conformity of the individual to societal expectations. The past few decades, however, have seen a marked shift to a more active view of the self, with an emphasis on self-socialization. Renewed interest in the self-concept as a source of motivation (Gecas 1986) and an agent in its environment has contributed to this shift, as has the increased interest in adult socialization (Levenson & Crumpler 1996). Even in studies of parent-child interaction, the child (even the infant) is increasingly viewed as an active partner in his or her socialization (Rheingold 1969). In short, the outcomes of socialization (whether conceptualized as values, self-conceptions, behavior patterns, or beliefs) are increasingly viewed as the products of reciprocal and negotiated interactions between agent and socializee.

A concern with social structure and its effects on the process and outcomes of socialization is still the hallmark of the sociological orientation to socialization, from social class influences (Gecas 1979) to the effects of family structure. Changes in family structure over the past few decades have increased interest in the effects of single-parent families, reconstituted families, and day care on child socialization (McLanahan and Sandefur 1994). Some of these changes have negative consequences for child socialization, some are benign, and some are ambiguous. It is evident that family trends and their consequences for child socialization will continue to generate a great deal of passionate debate in public and academic spheres because of their policy and value implications (Popenoe [1993] and Skolnick [1991] present contrasting views on the implications of these trends).

While most of socialization research has involved mainstream American populations, interest in cultural and subcultural variations in socialization experiences has been increasing. Socialization scholars have become sensitive to the criticism that our theories of socialization and self-conception are ethnocentric, reflecting a parochial, Western

perspective. This cultural self-consciousness has generated some interesting comparative studies, especially comparisons of socialization experiences in individualistic cultures (such as most Western societies) with those in collectivistic cultures (such as most Asian and many African cultures). In general, the self in collectivistic cultures is experienced much more in relational terms, that is, as interdependent, socially situated, and lacking definition outside the group context. By contrast, in individualistic cultures the self is more likely to be experienced as autonomous and unique and much less as a part of the social context. The socialization experiences in the former cultures tend to emphasize the primacy of the group (e.g., one's family, clan, or society) over the individual, whereas in the later the emphasis is on the development of independence, individual uniqueness, the ego, and autonomy (Markus and Kitayama 1991; Neisser and Jopling 1997). Ethnic subcultures within American society, such as Mexican-Americans and Native Americans, also reflect a more collectivist ethos than does mainstream culture, with consequences for socialization patterns. Within American society the cultural influences of ethnic minorities on socialization patterns typically interact with and may be confounded by social class influences. For example, studies of African-American families and peer groups may reflect adaptations to economic deprivation as much as they reflect distinct subcultural elements.

What effect will rapidly evolving computer technology have on the socialization of children and adults? The ubiquity of computers in classrooms, homes, workplaces, and recreation centers is evident in American society. Children are becoming involved with this technology at a very early age; even preschoolers are becoming "computer-literate." Unlike television, computers are an interactive technology that is likely to engage and develop the user's cognitive and motor skills. Thus, there is reason to hope that the influence of computers on children's development will be beneficial or at least benign, as some scholars have suggested (Turkle 1984). Computers and information technologies are rapidly transforming modern societies, affecting most aspects of people's lives. Surely they also are affecting the processes and outcomes of socialization, but the nature of these processes and outcomes remains to be studied by socialization scholars.

In traditional, relatively stable societies (which are increasingly rare), socialization is relatively routine and unproblematic. By contrast, in modern societies characterized by rapid social and cultural change, the socialization of children and adults is increasingly problematic and more likely to be contentious but also more interesting.

REFERENCES

- Aries, Phillip 1960 *Centuries of Childhood: A Social History of Family Life*. New York: Random House.
- Bandura, Albert, Dorothea Ross, and Sheila A. Ross 1963 "Imitation of Film-Mediated Aggressive Models." *Journal of Abnormal and Social Psychology* 66:3-11.
- Baumrind, Diana 1978 "Parental Disciplinary Patterns and Social Competence in Children." *Youth and Society* 9:239-276.
- Bem, Sandra L. 1974 "The Measurement of Psychological Androgyny." *Journal of Consulting and Clinical Psychology* 42:155-162.
- Benedict, Ruth 1934 *Patterns of Culture*. Boston: Houghton Mifflin.
- Blake, Judith 1989 *Family Size and Achievement*. Berkeley: University of California Press.
- Block, Jean H. 1983 "Differential Premises Arising from Differential Socialization of the Sexes: Some Conjectures." *Child Development* 54:1335-1354.
- Brim, Orville G., Jr. 1966 "Socialization through the Life Cycle." In O. G. Brim, Jr., and S. Wheeler, eds., *Socialization after Childhood: Two Essays*. New York: Wiley.
- 1968 "Adult Socialization." In J. A. Clausen, ed., *Socialization and Society*. Boston: Little, Brown.
- Bronfenbrenner, Urie. 1970. *Two Worlds of Childhood: U.S. and U.S.S.R.* New York: Russell Sage Foundation.
- Clausen, John A. 1968 "A Historical and Comparative View of Socialization Theory and Research." In J. A. Clausen, ed., *Socialization and Society*. Boston: Little, Brown.
- 1993 *American Lives*. New York: Free Press.
- Coleman, James S. 1961 *The Adolescent Society*. New York: Free Press.
- Cooley, Charles H. 1902 (1964) *Human Nature and the Social Order*. New York: Scribner's.
- Corsaro, William A., and Donna Eder 1990 "Children's Peer Cultures." *Annual Review of Sociology* 16:197-220.
- 1995 "Development and Socialization of Children and Adolescents." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Covington, Martin V., and Richard G. Beery 1976 *Self-Worth and School Learning*. New York: Holt, Rinehart and Winston.
- Elder, Glen H., Jr. 1974 *Children of the Great Depression*. Chicago: University of Chicago Press.
- Erikson, Erik H. 1959 *Identity and the Life Cycle*. New York: International Universities Press.
- Fine, Gary A. 1987 *With the Boys: Little League Baseball and Preadolescent Culture*. Chicago: University of Chicago Press.
- Foote, Nelson N. 1951. "Identification as the Basis for a Theory of Motivation." *American Sociological Review* 16:14-21.
- Galbraith, Robert C. 1982 "Sibling Spacing and Intellectual Development: A Closer Look at the Confluence Models." *Developmental Psychology* 18:151-173.
- Gecas, Viktor 1979 "The Influence of Social Class on Socialization." In W. R. Burr, R. Hill, F. I. Nye, and I. L. Reiss, eds., *Contemporary Theories about the Family*, vol. 1. New York: Free Press.
- 1981 "Contexts of Socialization." In M. Rosenberg and R. H. Turner, eds., *Social Psychology: Sociological Perspectives*. New York: Basic Books.
- 1982 "The Self-Concept." In R. H. Turner and J. F. Short, Jr., eds., *Annual Review of Sociology*. vol. 8. Palo Alto, Calif.: Annual Reviews.
- 1986 "The Motivational Significance of Self-Concept for Socialization Theory." In E. J. Lawler, ed., *Advances in Group Processes*. vol. 3. Greenwich, Conn.: JAI Press.
- , and Peter J. Burke 1995 "Self and Identity." In K. S. Cook, G. A. Fine, and J. S. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon
- , and Monica Seff 1990 "Families and Adolescents: 1980s Decade Review." *Journal of Marriage and the Family* 52:941-958.
- Geertz, Clifford 1973 *The Interpretation of Cultures*. New York: Basic Books.
- Goslin, David A. 1969 *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- Herskovits, Melville J. 1948 *Man and His Works: The Science of Cultural Anthropology*. New York: Knopf.
- Hewitt, John P. 1998 *The Myth of Self-Esteem*. New York: St. Martin's Press.
- Jones, R. A. 1977 *Self-Fulfilling Prophecies: Social Psychological and Physiological Effects of Expectancies*. New York: Wiley.



- Kanter, Rosabeth M. 1977 *Men and Women of the Corporation*. New York: Basic Books.
- Kohn, Melvin L., and Carmi Schooler 1983 *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, N.J.: Ablex.
- Levenson, Michael and Cheryl Crumpler. 1996 "The Models of Adult Development," *Human Development* 39:135-149.
- Levinson, Daniel J. 1978 *The Seasons of a Man's Life*. New York: Knopf.
- Lifton, Robert J. 1963 *Thought Reform and the Psychology of Totalism: A Study of Brainwashing in China*. New York: Norton.
- Maccoby, Eleanor E., and John A. Martin 1983 "Socialization in the Context of the Family: Parent-Child Interaction." In P. H. Mussen, ed., *Handbook of Child Psychology*, vol. 4. New York: Wiley.
- Malinowski, Bronislaw 1927 (1953) *Sex and Repression in Savage Society*, London: Routledge.
- Markus, Hazel R., and Shinobu Kitayama 1991 "Culture and the Self: Implications for Cognition, Emotion, and Motivation." *Psychological Review* 98:224-253.
- McLanahan, Sara S. and G. Sandefur. 1994 *Growing Up with a Single Parent*. Cambridge, MA: Harvard University Press.
- Mead, George H. 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Mead, Margaret 1928 *Coming of Age in Samoa*. New York: Morrow.
- Merton, Robert K. 1957 *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Neisser, Ulric and David A. Jopling, eds. 1997 *The Conceptual Self in Context*. New York: Cambridge University Press.
- Peterson, Gary W. and Della Hann 1999 "Socializing Children and Parents in Families." In M. B. Sussman, S. K. Steinmetz, and G. W. Peterson, eds., *Handbook of Marriage and the Family*. 2nd Edition. New York: Plenum.
- Piaget, Jean 1926 *The Language and Thought of the Child*. London: Kegan Paul.
- Popenoe, David 1993 "American Family Decline, 1960-1990: A Review and Appraisal." *Journal of Marriage and the Family* 55:527-555.
- Rheingold, Harriet L. 1969 "The Social and Socializing Infant." In D. A. Goslin, ed., *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- Rollins, Boyd C., and Darwin L. Thomas 1979 "Parental Support, Power, and Control Techniques in the Socialization of Children." In W. R. Burr, R. Hill, F. I. Nye, and I. L. Reiss, eds., *Contemporary Theories about the Family*, vol. 1. New York: Free Press.
- Rosenberg, Morris 1979 *Conceiving the Self*. New York: Basic Books.
- Rosenthal, Robert and L. Jacobson 1968 *Pygmalion in the Classroom*. New York: Holt, Rinehart, and Winston.
- Skolnick, Arlene 1991 *Embattled Paradise: The American Family in an Age of Uncertainty*. New York: Basic Books.
- Steinmetz, Suzanne K. 1999 "Adolescence in Contemporary Families." In M. B. Sussman, S. K. Steinmetz, and G. W. Peterson, eds., *Handbook of Marriage and the Family*, 2nd ed. New York: Plenum.
- Stryker, Sheldon 1980 *Symbolic Interactionism: A Social Structural Version*. Menlo Park, Calif.: Benjamin/Cummings.
- Turkle, Sherry 1984 *The Second Self: Computers and the Human Spirit*. New York: Simon and Schuster.
- Wheeler, Stanton 1966 "The Structure of Formally Organized Socialization Settings." In O. G. Brim, Jr., and S. Wheeler, eds., *Socialization after Childhood: Two Essays*. New York: Wiley.
- Zajonc, Robert B. 1976 "Family Configuration and Intelligence." *Science* 192:227-236.

VIKTOR GECAS

## SOCIAL-POLICY ANALYSIS

See Public policy analysis.

## SOCIETAL STRATIFICATION

Societal stratification phenomena are the relatively enduring, hierarchically ordered relationships of power among the units of which society is composed. The smallest units are adults, gainfully employed men and/or women, nuclear families, or sometimes extended families or households. Such units are ordered from highest to lowest in terms of power: political power, acquisitional power, the power of prestige, and the power of informational standing. Everybody experiences stratification every day, although a person often notices it only in the sense that some people seem better or worse off than he or she is. Social thinkers, powerful people, and revolutionaries have always been especially concerned with stratification.

Secure knowledge of the varying forms stratification structures may take is important because of the effects those structures have on many aspects of human experience, such as people's dreams of a better life, efforts to improve their situations, strivings for success, fear of failure, sympathy for the less fortunate, envy of others' good fortune, and even feelings about revolution.

A complete understanding of stratification requires several kinds of knowledge: first, what stratification structures consist of and how they vary; second, the individual and collective consequences of the different states of those structures; and third, the factors that make stratification structures change. This article reviews current thinking on the first of these elements.

### HISTORY: CLASSICAL THEORY

Two different lines of thought inform modern theory on societal stratification. One is *classical* theory; concerned with political power and privilege, it employs historical evidence. The other is the *empirical tradition*, which deals with systematic data on stratification as it exists contemporarily. Present-day theory of the behavior of stratification phenomena can be traced to Karl Marx's challenge to the manufacturing and financial elites of his day. Behind his concerns and those of the working class for which he was Europe's chief spokesman for many years lay the great economic and political upheavals of the eighteenth and nineteenth centuries.

The American and French revolutions and their aftermath culminated in legislation that made adults in many countries equal before the law. The related wave of emancipation of slaves and serfs in Europe and the Americas was also part of the intellectual environment of that day. Of more direct relevance to Marx's thinking was the rise of trade and the factory system, along with the growth of cities and the expansion of wealth. Marx saw urban populations dividing into two opposed classes. The capitalist class employed the workers; owned the workplaces, machines, and tools; and had ready access to large amounts of money for investment. The capitalists were opposed by their employees, the working class, who had nothing to offer but their time and energy. In Marx's view,

these two classes differ in terms of power and privilege: power because capitalists give orders that workers must accept, privilege because capitalists take the surplus (whatever is left after paying the cost of production) for themselves and their investments, leaving for workers only the wages that the market for labor forces capitalists to pay. Actually, Marx was interested in how these classes came into being and the conflicting interests they expressed. He did not write specifically on societal stratification as it is understood today.

Later writers on stratification, attempting to elucidate or contradict Marx, spelled out more complex sets of stratification dimensions. Weber (1946, 1947) saw power as the general factor basic to the enduring inequalities referred to as stratification. Sometimes, like Marx, he used categories whose underlying dimensions had to be elucidated by others. Party, class, and status groups were his key concepts. When these concepts are dimensionalized (reconstituted as variables), "party" is seen to be legitimate political influence, "class" is seen to express a hierarchical order of economic status, and the variable underlying "status groups" is seen to be their hierarchical order according to the degree of social honor. In other writings, Weber saw education as a stratification variable. In still others, he often wrote about authority, or legitimate superordinate and subordinate relations of power. Weber said nothing about how people are distributed in these dimensions or, of course, about how and why such distributions vary.

More thoroughly and precisely than Marx or Weber, Sorokin (1927) crafted the bases of modern theories of societal stratification. He distinguished political stratification, economic stratification, and occupational stratification. The first is a dimension of political power, and the second a dimension of the power of income and wealth. He left the dimensionality of occupational status unclear, sometimes implying that it was authority, sometimes privilege, and sometimes intelligence. Much of Sorokin's theory of societal stratification remains intact. First, he noted that all societies are stratified to some degree, a position widely accepted today. Second, empirical researchers continue to refine and elucidate his concepts of occupational status and occupational mobility. Third, in this connection he asked why occupational stratification exists and concluded that organized com-

munal life requires mechanisms and people to coordinate essential activities and that such coordination demands and rewards unusual ability. This view, now called *the functionalist hypothesis*, has been elaborated and disputed ever since. Fourth, he held that the degree of stratification varies from society to society and over time within given society: Stratification, he said, is in “ceaseless fluctuation.” Sorokin specified several ways in which stratification structures may vary. The whole structure may rise or fall; the top may rise or fall, changing the degree of inequality; and the “profile,” or the shape of the distribution, may vary. Similarly, the rate of individual upward or downward mobility may vary, and whole strata may rise or fall.

Sorokin thus presented a theory that specified (1) the general dimensions by which people are stratified within a society, (2) some ways in which the distributions of people on those dimensions may vary, and (3) why stratification exists. Also, he held such structures to be in ceaseless change.

The latest work in the classical tradition is that of Lenski (1966). His key dimensions are power, privilege, and prestige, in that order of importance. Beyond this, Lenski offers three main ideas. First, both functional theory and conflict theory, its opposite, are partly right. Society’s needs demand coordination, implying the existence of strata based on power or authority and implying a degree of consent on the part of many of those whose activities are organized by others. However, conflict results from that control: Authority is often abused and, even when it is not, may be misunderstood. Second, inequalities are mostly those of power, with inequalities of privilege and prestige following mostly as consequences of them. Third, the degree of inequality, which is seen as a single phenomenon encompassing the rate of mobility and the distance between strata, increased with the growing comprehensiveness and complexity of society until the Industrial Revolution, after which it declined. According to Lenski, the main forces driving change in the degree of inequality are the size of the surplus of production and, undergirding this, the march of technological efficiency.

Lenski is clearly in the classical tradition in his concern with power and privilege and dependence on historical evidence. To some extent, he echoes

Sorokin’s concern with variations in stratification structures through an emphasis on the degree of inequality. He provides a compelling treatment of the issue of conflict versus societal necessity in regard to the existence of stratification. He uses historical evidence effectively and systematically to mark variations of inequality in agrarian and horticultural societies. However, Lenski’s emphasis on two main, all-encompassing aspects of stratification—power (his key criterion variable) and inequality (used to denote the way in which power and its concomitants are apportioned)—forces too many separately varying stratification phenomena into too few molds. This problem becomes critical in industrial societies, where stratification dimensions vary independently of one another.

### HISTORY: THE EMPIRICAL TRADITION

As has been noted, this tradition of research on stratification is concerned with the here and now. This line of research has developed excellence in the measurement of the hierarchical positions of small demographic units within larger stratification structures. Although newer than the classical tradition, it has a long history. Several more or less independent status-measurement devices were formulated in the 1920s and 1930s. Most were concerned with either the prestige of the breadwinner’s occupation or the quality of the home. They tended to share certain assumptions: that stratification consists of a single hierarchy, in the early days usually called *social class*; that one or two different scales are sufficient to test hypotheses concerning social class; that social class positions can be distinguished by direct observation and/or interviews with someone who knows the status holders; that routines can be devised that allow one to assign valid and reliable numerical scores to each status holder on each of the scales used to measure social class; that the unit to be scored is the household, which can be one person or several persons living in a single home; and that it is the whole unit that is to be scored, whether with data on the home or data on the head of the household. Many of these devices became obsolete because they had to be recalibrated for each new community or type of community to which they were applied. Those that survived—education and occupational status—did so because they provide

comparable scores across large populations, such as nations.

Of the two main survivors, educational attainment is easy to measure: the exact number of school years successfully completed from none through sixteen, seventeen, eighteen, and so on. Measurement of occupational status is another matter. Two systems are currently in use. Occupational prestige ratings assume that each person in a given occupation shares the prestige most people attribute to that occupation. Occupational prestige scales have been constructed for many countries (Treiman 1977). Occupational socioeconomic status indices (SEI) are scales that use education and income to measure the status level of each occupation and then attribute to an individual the resulting score of her or his occupation. In the United States, Treiman's prestige scale and the SEI provide highly correlated occupational scores (Featherman and Hauser 1978).

Regardless of the original intent of such scales—to measure positions in what once was believed to be the only stratification hierarchy—the two variables educational attainment and occupational status are also appropriate for use with the classical theorists' multidimensional view of stratification.

### A SYNTHESIS

The current synthesis was carried out by stratification theorists who were both sensitive to the concerns of classical theorists with power and privilege and steeped in the empirical tradition. Thus, they brought the classical theorists' concern with political power, economic power, and social honor (Weber), including Sorokin's occupational status and Lenski's prestige, together with the empiricists' concern with education and occupational status (overlapping Weber, Sorokin, and Lenski) and with quantitative measurement and analysis.

Svalastoga's *Social Differentiation* (1965) appears to be the first statement of the synthesis. Svalastoga indicates the centrality of four dimensions of status: political, economic, social (mostly occupational), and informational (mostly educational). He calls attention to structural variations through his "parameters": the degree of inequality, the correlation among dimensions, and the degree of permeability (intergenerational circulation mo-

bility or movement up and down the hierarchies). Duncan (1968) both accepted and clarified Svalastoga's synthesis. His list of "scales of reward or status" provides a good outline of the large number of variables that should be measured to achieve a full-scale determination of people's levels on each status dimension. Also, he divides three of Svalastoga's four dimensions into two categories each. He, like Svalastoga, then lists three ways in which the structure of stratification variable may vary. The first is the degree of inequality. The second is called "rigidity of inequality" or "status crystallization," which is the same as Svalastoga's "correlation." The third is "rigidity of stratification," which is Svalastoga's "permeability" turned upside down.

Like Sorokin's and others' positions, Haller's (1970) statement of the synthesis assumes that stratification to one degree or another exists in all societies at all times. Revised slightly in this article, this form of the synthesis holds that there are two classes of dimensions of stratification. The first are "content"— or *power*—dimensions, after Weber: the capability of a given unit to elicit from others behavior promoted by the first unit, with such power having been routinized by coercion or consent. Agreeing conceptually but not always terminologically with the classical writers, this expression of the synthesis posits political power, economic power, and the power of prestige as universal dimensions of power. For civilized societies, it adds the power of years of formal education.

This position thus posits legitimized political influence (including authority) as the dimension underlying Weber's "party," Sorokin's "political stratification," Lenski's "power," and Svalastoga's "political status." It posits Weber's "class," Sorokin's "economic stratification," Lenski's "privilege," and Svalastoga's economic status as referring to the same set of hierarchical phenomena: access to goods and services—the *economic* dimension of a stratification structure. From Weber, it takes the variable of social honor; from Sorokin and modern occupational status researchers, that of occupational stratification; from Svalastoga, that of social status; and from Lenski, that of prestige. From the empirical tradition, it takes the measurement of occupational power ("status"). All these elements are seen as referring to a third homogeneous set of hierarchical phe-

nomena: the power of respect or deference attributed to a unit because of that unit's participation in a social category (such as an occupation) that has a specific level of evaluation by a society—the *prestige* dimension of societal stratification. As has been indicated, from Svalastoga and Duncan, with much support from the empiricists and also some from Weber, it takes informational power as a content dimension of a stratification structure, with education as its main indicator.

At the general level, each power dimension is of course presumed to be applicable in some form to all human societies as far back as human communal life can be traced. It is the exact expression of each dimension and the relationship among the dimension that vary across time and place. For entire contemporary societies, the main expressions of each dimension seem to be the following: for the political power dimension, *political power*, a variable researchers cannot yet measure despite its centrality in classical theory; for the economic dimension, *income* (occasionally wealth), a variable of concern to those in the empirical tradition; for the *prestige dimension*, *occupational status* in either of its two main forms of occupational prestige ratings (Treiman 1977) and occupational socio-economic index scores (Featherman and Hauser 1978); and for the informational power dimension, *educational attainment level* in terms of years of formal schooling successfully completed. Thus, in recent years it has become apparent that for today's societies, the main variables of the empirical tradition have central places among the content dimensions of the classical tradition. Income, occupational status, and education are the theoretically defensible variables most readily available to measure three of the four classical content dimensions.

Like Sorokin's, Svalastoga's, and Duncan's, Haller's formulation of the synthesis specifies several structural dimensions, with each one held to be applicable to every appropriate measure of each content dimension. The three structural dimensions of Svalastoga and Duncan are included: degree of inequality, status crystallization, and degree of status inheritance. Two others from Sorokin are included, although they are modified to fit today's understanding. One is the general level or central tendency, and the other is a division of Sorokin's concept of profile into two con-

cepts: mode structure and skewness. Although calculated from data on small units, each structural dimension applies to the society as a whole. Although logically they are partly dependent on one another, each one makes a unique contribution to an understanding of stratification. Each appears to be amenable to statistical description. Each is applicable to every indicator of the standing of every small unit (say, family) in the society. Valid measures of each content dimension taken at one point in time on a generalizable sample of the population of small units of that society would provide a complete description of the stratification structure of that society at that time. Successive measures would provide a complete description of the evolution of that society's stratification structure over time, thus providing a general idea of the variations in the degree of stratification in that society. Each applies to comparisons over time or among societies.

**General Level.** As Sorokin realized, the levels of structural dimensions may rise and fall as wholes. That is, the average economic, political, prestige, and informational standing of small units changes over time. These rises and falls may be seen in changes in the central tendency—say, the arithmetic mean or the median value—of the standing of small units. The rises and falls of the central tendency of any one of these dimensions do not necessarily follow the same pattern as those of another. Average economic, prestige, and informational power may increase, for example, while average political influence falls. This could happen in a society where a development-oriented dictatorship reduces citizen political participation while increasing levels of income, raising prestige by upgrading the occupational structure, and increasing access to education. Indeed, the economic, prestige, and educational levels of the populations of the more developed democracies have increased almost consistently since World War II, though this may not always be said for dictatorships. Also, raising the level of the occupational structure is exactly what some researchers mean by *upward structural mobility*, the case in which almost everyone is carried upward by changes in the economy that eliminate low-skill jobs while adding specialized jobs.

**Degree of Inequality.** The distances among the small units of a society may increase or de-

crease over time. This, so to speak, stretches the positions on the power dimensions apart or squeezes them together. The statistical term for this is the *degree of dispersion*. A number of measures of dispersion exist, such as the standard deviation (or its square, the variance), the range, the semi-interquartile (or quintile, decile, etc.) range, the share distributions, and the Gini, Theil, and Kuznets coefficients. There are two basic types of inequality: absolute and relative. Absolute conceptions assume that the metric on which the degree of inequality is measured is fixed so that as, say, real income per capita grows, the dollar difference between the mean of the top 10 percent and the mean of the bottom 10 percent of the small units may increase while each is rising above its previous level, with the slope of the top rising faster than is the slope of the bottom. For income, a proper description of these phenomena would be "changes in the *size distribution* of income." Absolute inequality and its changes are sometimes published. Much more often published are the *share distributions* of income. For any society at any time, share distributions take the total amount of, say, income as a constant 100 percent (or 1.00) and determine the degree to which the whole amount, regardless of its absolute size, is evenly or unevenly divided among the population. These distributions include the percentage of all income held by the top  $X$  percent and the bottom  $Y$  percent of the population. Or, as in the case of the Gini, Theil, and Kuznets coefficients, they use values ranging from 1.0 to zero, in which 1.00 is the maximum degree of inequality and zero is complete equality. Viewed at one point in time in a single society, measures of relative inequality are useful, but for comparison among societies or across time in the same society, they may be misleading. In fact, for many years the share distribution measures of the income of the American people remained essentially unchanged while the size distribution inequality increased dramatically (U.S. Department of Commerce 1980). This was the case because real per capita income was increasing rapidly. The greater the degree of inequality, the greater the degree of stratification.

**Crystallization.** It has long been recognized that a stratification structure may tend toward or away from monolithicity, in which the different power dimensions merge into a single hierarchy or tend to be in partially separate hierarchies. At one extreme, the position of a small unit on any

one of the dimensions can be found by knowing its position on any other dimension. In other words, if the four content dimensions are perfectly correlated, those in lofty positions on one dimension also will be in lofty positions on all the other dimensions, while those in humble positions on one will be in similarly low positions on any other. At the opposite extreme, a unit's position on a given content dimension is irrelevant to its position on any other. In the real world, any two or three might be highly interrelated, all might be moderately intercorrelated, and so forth. For obvious reasons, Svalastoga called this structural dimension "correlation." Others have called it "status crystallization." Note that crystallization levels and forms may be summarized better by a method called *factor analysis* than by the correlations themselves. Factor analysis can show which sets of content variables tend to vary together in a population and which do not. It also can help determine which are the dominant dimensions and which are of lesser importance in a given stratification structure. For example, it appears that the former Soviet stratification structure was dominated by the political dimension; the American, by the economic dimension. Factor analysis of the correlations of the content dimensions could indicate whether these beliefs are true. The greater the degree of crystallization, the greater the degree of stratification.

**Status Inheritance.** Status inheritance is the degree to which people's level on a given content dimension is controlled by that of their parents. It is exactly the obverse of circulation mobility: A high degree of power position inheritance implies a low degree of circulation mobility. The basic statistical summary of this phenomenon is either the correlation coefficient ( $r$ ) or the coefficient of determination ( $r^2$ ) of the power dimension positions of offspring and their parents. (The  $r^2$  tells how much one variable is determined by the others.) The greater the degree of status inheritance, the greater the degree of stratification.

**Sorokin's Profile.** Every variable has a so-called distribution, a shape that appears when the number of scores (the *frequency*) is plotted against the value of scores. Much statistical theory today assumes that real-world distributions conform to certain mathematical shapes. The bell-shaped "normal" curve is the one most often used. For distri-

bution of income, the “log normal” curve, with which the distribution of the natural logarithm of the individual amounts forms a normal curve, is often employed. Stratification researchers often take it for granted that the distributions of power dimensions are either normal or log normal, but there is no sociological reason to assume this. The shape of the distribution of a content dimension is precisely what Sorokin meant by his term “profile.” Lacking the data and concepts to proceed further, he simply called the real-world shapes of these distributions their profiles. Today we can see that there are two aspects of each profile: mode structure and skewness.

In strikingly underdeveloped societies, almost everyone is concentrated at the very lowest values of economic, political, occupational, and educational power: extremely poor, utterly uninfluential, of low prestige, and illiterate. Above those people, their “betters” are arranged in rank order, with a wide range in which the few people who are above the bottom dwindle up the line to a handful of individuals of lofty standing. Each such distribution would have a very low mode (or distinct cluster of cases) and median (where half of the cases are higher and half are lower) and a higher arithmetic mean, with a sharply skewed tail. In somewhat more developed societies, such distributions, instead of yielding bell-shaped or log normal curves, might show multiple modes, with many people concentrated around a fairly low point, quite a few concentrated around a point a bit higher, a few concentrated toward the top, and after that a sharp skewing up to the very few at the top. The consequences of such forms for the lives of the people involved are no doubt great. For example, if in a certain society almost everyone is destitute, the few who are more or less well to do are highly visible. Even if the wealthy were really not far above the others, everybody would think of that society as being highly stratified. If in another society people are bunched together at several points along a hierarchy, thus forming multiple modes, or discrete classes, those in each mode might come to consider themselves members of a special class in opposition to those concentrated at another mode. Thus, the exact forms of profiles are essential to a description of a society’s stratification structure. Theoretically, these forms have substantial consequences for many stratification-dependent behaviors.

**Profile: Mode Structure.** Mode structure refers to the number, size, and location of distinct modes on the distribution of each content variable. In polymodal structures, the more pronounced the modes, the greater the degree of stratification.

**Profile: Skewness.** Several statistical devices exist to mark the degree of skewness. The greater the level of skewness, the greater the degree of stratification.

## ILLUSTRATIONS

Data by which to measure and compare stratification structures are exceedingly difficult to obtain. A complete description at a specific point in time requires well-measured, valid indicators of four power dimensions, one or more for each dimension. For each indicator, several measurements must be made: The average level, the degree of absolute and relative inequality, the degree and factor-analytic forms of the crystallization of the whole set of indicators of the power dimension, the degree of power position inheritance, and the distributions of each one must be plotted to indicate its mode structure and measure its skewness. Describing such an overall structure requires the construction of 24 or more different indicators of structural dimensions. These indicators have to be based on representative, societywide samples large enough to permit the recording of small differences, as in the case of the few people at the upper end of a skewed distribution. The study of variations in the structure of stratification demands that comparable measurements be taken on the same variable at different times and in different places. In itself, the requirement of comparability is extremely severe when one is making comparisons among societies with different cultures or over long periods of time within the same society.

Exploratory work of this sort has been conducted on data provided by Brazil. The data were collected on a national probability sample of households in 1973 and are available for all employed men and women in the households sampled. These people are the “small units” of the descriptive analysis presented below. Brazil is a particularly good place in which to conduct such exploratory research for two reasons: It is a large country

whose regions are markedly different from each other in terms of development, and it has only one language and culture. The first factor makes it feasible to test for structural variations of stratification associated with development levels, treating regions as societies; the second eases the problem of comparability.

As was indicated earlier, it is not currently feasible to obtain measures of the political power dimension in Brazil or anywhere else. However, there is widespread agreement that income is a proper measure of the economic status dimension, that occupational status instruments based on the average education and income of each occupation are proper measures of the prestige dimension, and that education is a similarly appropriate measure of the informational status dimension. These data are available for some of the parameters that would have to be assessed to obtain a complete description of the regional-development variations of the Brazilian stratification structure in 1973.

Here one is comparing sharply different development regions. The stratification structures of three of Brazil's socioeconomic development (SED) macroregions in 1970 were delineated by obtaining multiple-item, factor-weighted SED scores on that nation's 360 official continental microregions and plotting their levels on the map of Brazil (Haller 1983). This showed the following five macroregions: the Developed South (median SED = 78 on a scale of zero to 100), the South's Developing Periphery (median SED = 54), the Undeveloped Amazonia (median SED = 32.5), the Unevenly Developed Northeast (median SED = 31), and the Underdeveloped Middle North (median SED = 13).

Obviously, this article cannot reproduce each one of the structural dimensions for each SED macroregion for men and for women. Instead, it provides a few key illustrations for three of the regions: the Developed, the Developing, and the Underdeveloped.

Variables routinely used as indicators were formulated to measure three of the four stratification content dimensions: education in years successfully completed, occupational status scores (composed of canonically weighted scores based on the education and income of each occupation), and annual income in 1973 U.S. dollars.

The illustrations are based on regularly employed men and women 15 to 65 years of age. All such persons who lived in the three regions under comparison and were part of the sample have been included. The numbers of sample members vary sharply by region and by sex. The Developed South is much more populous than the other two regions, and about three times more men than women are employed. The largest of the six gender-by-region subsamples thus consists of men in the South: over 40,000 (see Table 1). The smallest consists of women in the South's Developing Periphery: over 2,500.

Let us begin with the profiles (see Figure 1), graphs that have been sketched to show the shape of the stratification structures for men and for women as they appear in the three regions. There are two reasons for paying close attention to these curves. First, they show power relations among the people. The presence of multiple modes shows the existence of discrete and potentially opposed classes. Both the mode structure and the marked skewing indicate a high degree of stratification for each sample. Second, the fact that these distributions diverge sharply from normal or log normal curves shows that the numbers, that is, the data presented in Tables 1 and 2, are at best approximate because the shapes of the distributions affect their meaning.

The curves show the following:

1. Multiple modes are exhibited by both men and women in 11 of the 12 graphs pertaining to the developed and developing regions. The exception is distribution of women's income in the developing region.
2. For the two most developed regions, comparable curves show just about the same mode structure. In these regions, education tends to be bi- or trimodal and occupational status tends to be at least trimodal. Among men and among women in the developing region, income also exhibits multiple modes. In the underdeveloped region, the shape of the curves is markedly different from that of the others. The curves in this region show a heavy concentration of both men and women at the bottom of each indicator variable,



SOCIETAL STRATIFICATION

**Illustrative Variations of Brazilian Regional Stratification Structures by Development,  
Employed Persons Age 15-65, 1973.**

STRATIFICATION CONTENT VARIABLE	REGION					
	Men			Women		
	Developed	Developing	Underdeveloped	Developed	Developing	Underdeveloped
<i>Education</i>						
General level (average)	4.9	4.2	1.7	5.3	5.1	1.6
Absolute inequality (Standard deviation)	3.9	3.8	2.3	4.3	4.5	2.7
<i>Occupational Status</i>						
General level (average)	19.4	16.8	6.7	20.3	21.1	8.6
Absolute inequality (Standard deviation)	18.9	18.0	10.8	19.7	20.4	14.6
Circulation mobility (1-r <sup>2</sup> )	0.72	0.79	0.85	0.69	0.75	0.63
<i>Income, Annual</i>						
General level (average)	1,800	1,423	536	891	610	264
Absolute inequality (Standard deviation)	2,670	2,330	903	1,132	864	400
Number of Persons	1,578	7,686	5,841	15,711	2,581	2,777

*Table 1*

NOTE: *Education* is in estimated years. *Occupational status* is in canonical socioeconomic status units (0-100); *circulation mobility* is intergenerational. *Income* is in U.S. dollars.

- though some of the region's six graphs show the formation of small second and sometimes third modes at high status levels. The apparent conclusion is that the underdeveloped area exhibits a relatively high degree of equality at the very bottom of the Brazilian stratification structure. This is precisely the opposite of the thinking among many observers of Brazil, who believe that inequality is greater in the underdeveloped region (perhaps because of the glaring visibility of the tiny stratum at the top).
3. Each curve shows a high degree of skewness. That is, the highest positions are held by a tiny proportion of the people, and on the whole, as the tail of the distribution lengthens, the higher the level, the tinier the percentage of the people.
  4. In every case, the main modes are the one or two at the bottom, where most people tend to be concentrated.
  5. Almost every graph shows a tendency for one or two smaller modes to appear toward the middle of the distribution. For education, this occurs at around grade 12. For occupational status it is about 50 units, or the level of office clerks, primary school teachers, and the like. For income, it is about \$2,000 to \$3,000 per year, or a monthly wage between \$160 and \$250.
  6. There may be a tiny mode near the top of the educational and occupational status distribution in the more developed regions.
  7. For occupational status and income, women are more concentrated toward the bottom than are men.

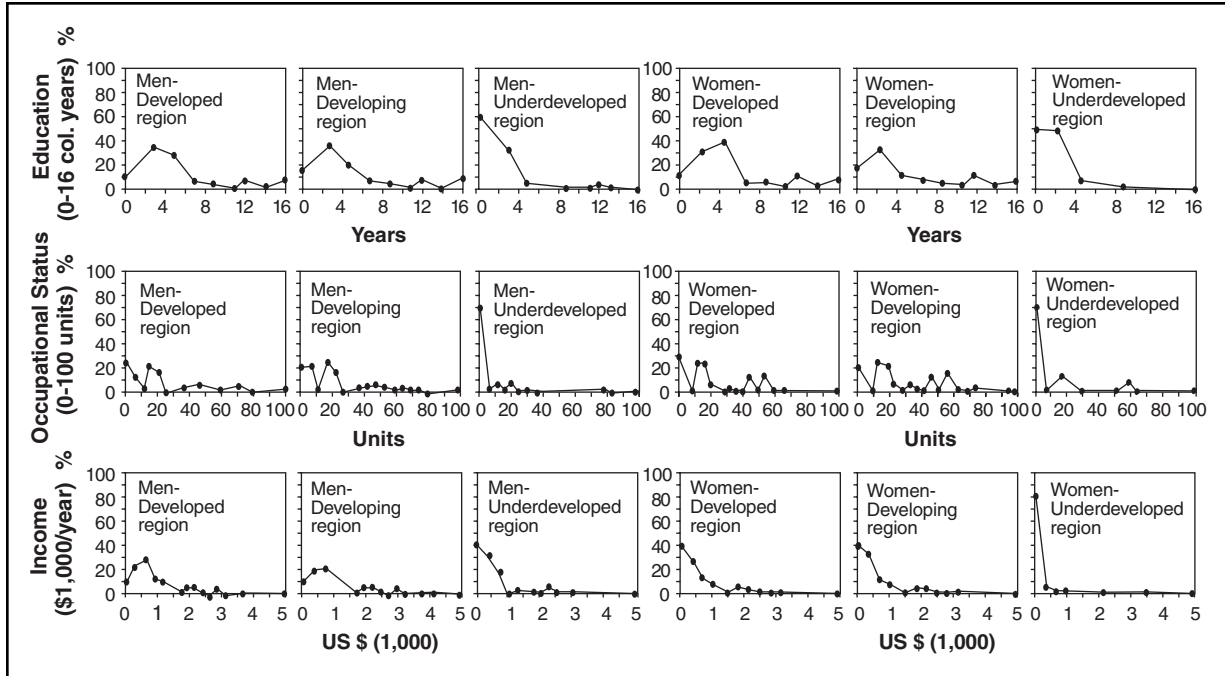


Figure 1. Illustrative Variations of Brazilian Regional Stratification Profiles, by Development, Employed Persons Age 15–65, 1973

SOURCE: See Table 1 for definitions and sample sizes.

8. Clearly, the main regional variations in profile are between the two more developed regions and the underdeveloped region.
9. In terms of mode structure, the more developed areas seem more stratified than does the underdeveloped area.
10. In terms of skewness, it appears that the underdeveloped area is more highly stratified.

Data on the general levels and absolute inequality levels of the three content dimensions are presented in Table 1. For occupational status, the degree of circulation mobility also is presented. The general level rises with development for all three variables, except for the occupational levels of women in the developing region, whose status is slightly higher than that of women in the developed region. Again, with two exceptions among women in the developing region, the higher the level of development, the greater the degree of absolute inequality. Echoing what was gleaned from the graphs, the general level and the absolute

inequality levels of the underdeveloped region are markedly lower than those in the other areas. Finally, among men, the higher the level of development, the lower the degree of circulation mobility. Women show no trend in this regard.

Evidence regarding structural crystallization is presented in Table 2. Among men, the higher the level of development, the higher the degree of crystallization. Among women, the same trend may be present in the data, but with one small inconsistency.

### CONCLUSION

This article has attempted to describe the contemporary synthesis of classical and empirical traditions of sociological thought concerning societal stratification, with special emphasis on what may be learned about ways to describe variations in stratification structures. It also presents some illustrations showing how indicators of some of the structural dimensions vary across development regions among employed men and women in

**Illustrative Variations in Structural Crystallization among Brazilian Development Regions, Employed Persons Age 15-65, 1973 (Correlation Coefficients).**

STRATIFICATION CONTENT VARIABLE	REGION					
	Men			Women		
	Developed	Developing	Underdeveloped	Developed	Developing	Underdeveloped
Education by occupational status	0.52	0.51	0.35	0.65	0.67	0.52
Education by income	.27	.18	.16	.23	.20	.20
Occupational status by income	.23	.16	.13	.23	.17	.16

Table 2

Brazil. In general, these indicators show that the more highly developed a region is, the more stratified it is.

Measuring stratification variations among societies is an immense task because of the number of variables that must be studied and because of differences in culture, language, and social organization among peoples. Still, at both the individual and societal levels, the effects of structural differences of stratification are among the most perplexing of this age and perhaps of all ages. For this reason, understanding how and why stratification structures vary and specifying the consequences of such differences are worth the considerable effort required.

REFERENCES

Duncan, Otis Dudley 1968 "Social Stratification and Mobility: Problems in the Measurement of Trend." In Eleanor Bernert Sheldon and Wilbert E. Moore, eds., *Indicators of Social Change: Concepts and Measurements*. New York: Russell Sage Foundation.

Featherman, David L., and Robert M. Hauser 1978 *Opportunity and Change*. New York: Academic Press.

Ganzeboom, Harry B. G., Donald J. Treiman, and Wout C. Ultee 1991 "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277-302.

Grusky, David B. 1994 *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, Colorado: Westview Press.

Haller, Archibald O. 1970 "Changes in the Structure of Status Systems." *Rural Sociology* 35:469-487.

— 1983 *The Socioeconomic Macroregions of Brazil*. Lanham, Md.: Bernam-Unipub.

Lenski, Gerhardt 1966 *Power and Privilege: A Theory of Social Stratification*. New York: McGraw-Hill.

Sorokin, Pitrim A. 1927 *Social Mobility*. New York: Harper.

Svalastoga, Kaare 1965 *Social Differentiation*. New York: David McKay.

Treiman, Donald J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

U.S. Department of Commerce 1980 *Social Indicators II*. Washington, D.C.: U.S. Government Printing Office.

Weber, Max 1946 "Class, Status, and Party." In Hans Gerth and C. Wright Mills, eds., *From Max Weber*. New York: Oxford University Press.

— 1947 "Social Stratification and Class Structure." In A. M. Henderson and T. Parsons, eds., *The Theory of Social and Economic Organization*. New York: Free Press.

ARCHIBALD O. HALLER

**SOCIETY AND TECHNOLOGICAL RISKS**

If there is an organizing theme in sociology, it is social order: what it looks like, how to think about the various forms it takes, and how to explain it. Conversely, what happens when social order breaks down? What changes are wrought in how people see the world, and most important, what is altered in how they relate to one another when social order goes awry? The study of risk, danger, and catastrophe is a special case of the larger field of social breakdown.

Sociologists have long been interested in phenomena that harm people and what people value. Until recently, most of this work concentrated on harm from natural events such as earthquakes, floods, and tornadoes, but many researchers now write about “technical” or “technological” risks. In some ways the distinction between natural and technological risks or disasters is not helpful: There is no objective difference between a death caused by a fire and a death caused by an airplane crash. Yet in other ways those who have been fascinated by how modern technologies fail people have asked a broader set of questions than they could have if they did *not* see a difference between natural and technological risks. They have asked new questions about the functions of expertise and science in modern society, the roles of power and authority in the creation of danger, and the capacity of people to build systems they cannot control.

In this encyclopedia of sociology, risk and danger are treated mainly as a sociological problem, but this is not necessarily the case. Scholars writing about these issues come from economics, geography, psychology (Mellers et al. 1998), anthropology (Oliver-Smith 1996), and even engineering (Starr 1995) and physics. This is basically a good thing: Too much sociology is self-referential and inbred, and truly interdisciplinary work creates considerable intellectual, if not professional, excitement. No one can write about technological risks in an interesting way without reading and thinking in interdisciplinary terms.

Scholars concerned with technological risks have addressed a wide variety of topics that range from how individuals think about risks to how nation-states develop strategies to mitigate threats from failures of high technology. Some scholars even write about risks that might be faced by societies far in the future. Toxic threats have drawn particularly close scrutiny from scholars, and there are important sociological studies of Love Canal, Three Mile Island, Chernobyl, Bhopal, the *Challenger*, nuclear waste, and nuclear weapons. One reason for this is that toxic risks invert the way natural disasters do damage. Rather than assaulting people from the outside, as do other calamities, toxic hazards assault bodies from within. Toxic injuries also have no definable end, and so their victims can never know when they are safe from further damage. The point here is that the

meaning of toxic threats is fundamentally different from that of natural disasters (Couch and Kroll-Smith 1985; Erikson 1990, 1994). The disruption in social order thus can be largely internal, with psychological and emotional suffering caused by the breakdown of external social systems (Sorokin 1968).

In general, the sociology of risk is concerned with researching and explaining how interactions between technology and modes of social organization create hazards or the potential for hazards (Clarke and Short 1993). A hazard can be an actual threat to people’s lives (toxic chemical contamination, for example) or the *perception* that there is a threat. Indeed, many analysts focus on risk perception: what people think is dangerous and why they think what they do (Freudenburg 1988). The word “technology” refers to the social and mechanical tools people use to accomplish something, such as the design of a nuclear power plant and the vocabularies used by experts when they talk about effectively evacuating an urban area after a major radiation release from a nuclear power plant. “Modes of social organization” refers to both social structure (e.g., hierarchies of power) and culture (e.g., the degree of legitimacy granted to experts). In the twenty-first century society will continue its march toward social and technical complexity. One expression of this complexity is a capacity to create machines and institutional arrangements that are at once grand and terrifying. With these developments, it seems, publics are increasingly aware of the potentially devastating consequences of system failures even as they enjoy the cornucopia engendered by modern social organization and modern technology.

This is an opportune place to identify an area of research that will be increasingly important for both intellectual and policy reasons. A lot of work in psychology and economics, which echoes the concerns of political and economic elites, concerns public perception of risk. Much of that work has shown that the general public does not make decisions in accordance with a hyperrational calculus in which its preferences and values are always consistent and, more to the point, agree with those of trained scientific researchers. Consonant with the concern with public irrationality is the notion that people panic when faced with risks they do not understand. It is easy to find this idea

in media reports of high-level politicians' remarks after a large accident: Politicians worry that people will overreact in very unproductive ways. The image is one of people trampling each other to make for the exits of a burning building or escape from a sniper's random rifle shots. Translated to perception of risk regarding accidents and disasters, the image becomes one of individuals pursuing their self-interest to the exclusion of those of their neighbors and communities: to get out of Love Canal, to run away from Three Mile Island, or to flee a burning airplane that has barely managed to land.

In fact, research indicates that people *rarely* panic even when it might be rational to do so. I have reviewed scores of cases of natural and technological disasters—trains fall over and release toxic chemicals that endanger a town, earthquakes shake a city to its core, fires threaten to level an entire neighborhood—and have found very few instances of uncontrolled flight at the expense of others. After the Chernobyl catastrophe in 1986 there was some panic, though that response might have been highly sensible. The U.S. firebombing of Tokyo in World War II also elicited some cases of panic. With exceptions of that sort, it is hard to find widespread panic after any type of disaster. Even events such as the fire at the Beverly Hills Supper Club and the stampede at the Who concert, which are commonly thought of as examples of panic, were not (Johnson 1987). Rather than panic, the modal reaction is one of terror, followed by stunned reflection or sometimes anomie and ending with a fairly orderly response (e.g., reconstruction or evacuation). Even in the horrors chronicled by the U.S. Strategic Bombing Survey, cities burn, bodies explode, houses fall down, and still people do not panic (Janis 1951; Hersey 1985).

One way to classify research on risk is in terms of micro and macro perspectives. Both micro and macro studies have made important contributions to an understanding of the connections between risk, technology, and society. Micro-level research, generally speaking, is concerned with the personal, political, and social dilemmas posed by technology and activities that threaten the quality of people's lives. Macro-level work on risk does not deny the importance of micro-oriented research but asks different questions and seeks answers to those questions at an institutional level of analysis.

As some of the examples below illustrate, much macro work emphasizes the importance of the institutional context within which decisions about risk are made. Sociologists of risk are keen to distinguish between public and private decisions. Some people make choices that affect mainly themselves, while those in positions of authority make choices that have important implications for others. This is only one among many ways in which the sociology of risk is concerned with issues of power and the distribution of hazards and benefits.

### THE MICRO LEVEL

As noted above, a substantial body of work has demonstrated that the public overestimates threats that are dramatic (e.g., from airplane accidents), particularly violent (e.g., from handguns), and potentially catastrophic (e.g., from nuclear power plants). Similarly, people tend to underestimate more prosaic chronic threats such as those from botulism and asthma. Several explanations for this phenomenon have been proposed; the one that is most convincing focuses on the mechanisms through which information about risks is channeled to people (Kahneman et al. 1982). Specifically, the media—especially newspapers and television—are more likely to feature dramatic, violent, or catastrophic calamities than less sensational threats. One reason the media find such risks more interesting is that they are easier to cover and hence more easily fit into tight deadlines. Covering prosaic risks is also more time-consuming than covering short, dramatic accidents. Thus, there are several good structural reasons why the media pay attention to high-drama risks and neglect low-drama risks. Scholars are able to explain why the public has biased estimates of risk by focusing on the structural connections between people and the media, specifically on the constraints that lead the media to be biased about certain types of information.

Another example at the micro level of analysis is found in the work of Heimer (1988, 1992; Heimer and Staffen 1998), who analyzed how information is used and transmitted in intensive care units for infants. Her study is cast at a micro level of analysis in the sense that one of her concerns is how parents think about information regarding terribly sick babies. However, like all

good sociological studies, Heimer's connects what parents think with the social contexts in which they find themselves. For example, one of her findings is that when hospital personnel transmit information to parents about their infants, that process is structured to protect the hospital from lawsuits and only secondarily to apprise parents of the condition of their children. Thus, Heimer describes, much as a psychologist might, how parents think but also demonstrates that how they think is contingent on the organizational needs of hospitals.

### THE MACRO LEVEL

Macro-level work on risk includes research on how professionals influence the behavior of regulatory agencies, how organizations blunder and break down (Vaughan 1999), how social movements arise to push issues into the public debate, and how national cultures influence which risks are considered acceptable (Douglas 1985). Many macro theorists are deeply concerned with how the institutional structure of society makes some risks more likely than others to gain political and intellectual attention (Clarke 1988). Consider motor vehicle risks. Nearly 40,000 people are killed on U.S. highways every year, and most people would agree that that is an appalling mortality rate. Although it is a commonplace that half those deaths are alcohol-related, it is not really known how much of the carnage is in fact due to alcohol (Gusfield 1981). Nevertheless, one probably can reasonably assume that a significant proportion is caused by drunk drivers (even 10 percent would be 4,000 people). Most people would agree that 4,000 deaths per year is grounds for concern, yet it also is known that high rates of fatal traffic accidents are associated with speeding, wrong turns, and improper passing. Objectively, there is no difference between a death caused by someone making a wrong turn and one caused by a drunk driver, yet most cultures say the two deaths are very different. In the United States there is even a small social movement galvanized around the issue of drunk drivers, an example of which is the organization called Mothers Against Drunk Driving. Why is there no organization called Mothers Against Improper Passers? A sociological answer is that most cultures frown on using drugs to alter one's degree of self-control and that a person who does so is defined as morally decadent and lacking social

responsibility. Thus the opprobrium unleashed on drunken drivers has less to do with the objective magnitude of the problem than with the apparent danger such drivers represent to the cultural value of self-control.

Another example of macro work on risk, this time concerning organizations and symbols, is the oil spill from the *Exxon-Valdez* tanker in March 1989. At the time, the Exxon spill was the worst ever to occur in U.S. waters, leaking at least eleven million gallons into Prince William Sound and the Gulf of Alaska. The spill caused massive loss of wildlife, and although no people died, it did disrupt social relationships, create a political crisis, and reorient debates about the safety of oil transportation systems in the United States. From a sociological point of view, one of the most interesting things about the spill is how corporations and regulatory agencies plan for large oil spills. Sound research shows that not much can be done about large amounts of spilled oil (Clarke 1990), yet organizations continue to create elaborate plans for what they will do to contain large spills and how they will clean the oil from beaches and shorelines. They do this even though there has never been a case of successful containment or recovery on the open seas. Organizations create plans that will never work because such plans are master metaphors for taming the wild, subjugating uncertainty, and proclaiming expertise. In modern societies, expert knowledge and rational organization are of paramount importance, and there now seems to be an institutionalized incapacity to admit that some things may be beyond people's control (Clarke 1999).

Another example is the 1984 tragedy in Bhopal, India. At least 2,600 people died when a complex accident in a Union Carbide plant released toxic chemicals into the environment. At the time, Bhopal was the worst single industrial accident in history (the nuclear meltdown at Chernobyl in 1986 eventually will lead to more deaths). The Bhopal tragedy was certainly an organizational failure, as studies have documented (Shrivastava 1987). However, what was most interesting about the Bhopal accident was that the risk created by the Union Carbide chemical plant had become institutionalized to the point where very few, if any, of the key players, were worried about a potential catastrophe. The poor people who lived next to the plant

seemed to have accepted official assurances that they were safe and in any case had little choice in the matter. Government officials—the ones assuring those who lived near the plant of their safety—seemed to have accepted the disastrous potential of the plant as part of the price of having a large corporation in their country. For their part, corporate officials and experts seemed to have given insufficient thought to the possibility of killing several thousand Indians. One reason the Bhopal disaster is sociologically interesting is the degree to which groups and organizations come to accept risk as part of their everyday lives. The same observation might be made of automobile driving, nuclear power plants, and lead-contaminated water pipes (even brass pipes contain about 7 percent lead).

Anyone reading about how social organization and technology can break down must wonder whether it must always be so. Some people believe that it must, and although those scholars may be wrong (Clarke 1993), they cannot be ignored. High-reliability organizations (HROs), as they are called, are said to be possible. These organizations are alleged to be so safe that they are error-free. The claim is not that these organizations cannot fail but that they never do so in an important way (Roberts 1993). Somehow these organizations (a U.S. nuclear aircraft carrier is a good example) are able to maintain a strict hierarchy while enabling people low in the hierarchy to intervene in the functioning of the organization to prevent failure (contradicting how most organizations work). Moreover, rather than cover up their mistakes as most organizations do, HROs try hard to learn from theirs. Finally, these organizations have a lot of redundancy built into them, preventing small errors from escalating into complete system failure.

High-reliability theory is animated by a disagreement with what is called normal accident theory (NAT) (Perrow 1984). NAT views the complexity of high-technology systems as problematic. The components in complex systems, say, a nuclear power plant, can fail in ways that no one could have anticipated and that no one understands when a catastrophe unfolds. From this view, rather than safety, redundancies can add technical complexity and lead to the formation of political interest groups, both of which can interfere with safe operations (Sagan 1993). Rather than honest learning, NAT stresses that managers and experts

often engage in *symbolic* representations of safety, trying to convince the public, regulators, and other organizations that they are on top of potential problems (Clarke and Perrow 1996). An important contribution of NAT lies in locating the source of risk in organizations per se. Its structural emphasis draws attention away from easy and familiar explanations such as human error.

The contrast between NAT and HRO theory goes beyond their assessments of the inevitability of organizational failure, for the two schools of thought exemplify the concern with social order and disorder that I mentioned above. High-reliability theory is optimistic about human perfectibility, highlighting society's tendency to create and maintain order; normal accident theory is pessimistic about human perfectibility, fundamentally viewing failure and disorder as inherent in the human condition.

An important work that emphasizes the imperfections of organization is Vaughan's book on the *Challenger* accident (1996). Vaughan argues that what looked like a highly risky decision—to send the *Challenger* up that day—was in fact normal given the routines and expectations that organized the thoughts of the officials and experts involved in that choice. An important reason for this “normalization of danger” was the high production pressures that the decision makers faced. It was not a matter of people deliberately taking chances they knew were unreasonable because of clear, external, imposing pressures. It was a more subtle process by which the very definition of “reasonable” shifted in a way that did not contravene those pressures. Vaughan's view stresses the commonality of error in all organizations. Some organizations make computer chips, some make space flights, but all fail. Vaughan details the mechanisms that produced a certain worldview of danger and safety that prevailed at NASA, and in so doing she connects micro and macro, structure and culture.

Every year the natural environment seems to be more polluted than it was the preceding year. Why? A commonsense explanation might claim that people do not care enough about the environment, perhaps attributing callous attitudes and personal greed to politicians and corporations. Such an explanation would focus on the motives of individual managers and politicians but from a

sociological point of view would miss the all-important institutions in which such decision makers function. Sociologists know that those who occupy top positions in government and corporate organizations are not without intelligence and good sense. These decision makers may be individually quite concerned about environmental degradation, but because of their structural locations, they are subject to pressures that may be at odds with environmental health and welfare. These pressures originate in specific social structures that create institutional interests that may be contrary to individual preferences (Clarke 1988). A corporate executive seeks (and *must* seek) to remain in business, if necessary at the expense of others' well-being or the environment. A similar explanation accounts for why Ford's president, Lee Iacocca, in the 1970s marketed Pintos that had propensity to explode and burn. In other words, market institutions are arranged so that it is sensible for any individual or organization to force negative externalities on society. For its part, one of the key functions of government is to maintain a political and economic environment that is favorable to business. Thus, an explanation that centers on the institutional constraints and incentives that shape decisions about pollution can account for the behavior of organizations, experts, and officials far better than can an explanation that focuses on their personal characteristics.

### FUTURE DIRECTIONS

Future developments in the sociology of risk probably will revolve around issues of social conflict: its bases, meaning, and role in spurring social change. Society—and sociology—will be confronted with fundamental dilemmas in the twenty-first century. Society will have to deal with issues of environmental justice and the likelihood that pollution and risk are unequally distributed. Modernity brings both fruits and poisons. In particular, people must come to grips with what may be the primary dilemma of modern times: How can an industrial and democratic system that yields such a high standard of living also be involved in the creation of terrible hazards? Answering this question will require a recognition that many of the most frightening threats—nuclear meltdowns near large cities, toxic leachate in water tables, ozone destruction, explosions of liquefied natural gas from supertankers, failure to contain nuclear waste—almost seem

beyond control. It may be the case that before society can better control political and technological systems, people must admit that some aspects of the technical world are not within human control.

### REFERENCES

- Clarke, Lee 1988 "Explaining Choices among Technological Risks." *Social Problems* 35(1):501-514.
- 1990. "Oil Spill Fantasies." *Atlantic Monthly*, November, pp. 65-77.
- 1993 "Drs. Pangloss and Strangelove Meet Organizational Theory: High Reliability Organizations and Nuclear Weapons Accidents." *Sociological Forum* 8(4):675-689.
- 1999 *Mission Improbable: Using Fantasy Documents to Tame Disaster*. Chicago: University of Chicago Press.
- , and Charles Perrow 1996 "Prosaic Organizational Failure." *American Behavioral Scientist* 39 (8):1040-1056.
- , and James F. Short, Jr. 1993 "Social Organization and Risk: Some Current Controversies." *Annual Review of Sociology* 19:375-399.
- Couch, Stephen R., and J. Stephen Kroll-Smith 1985 "The Chronic Technical Disaster." *Social Science Quarterly* 66(3):564-575.
- Douglas, Mary 1985 *Risk Acceptability According to the Social Sciences*. New York: Russell Sage Foundation.
- Erikson, Kai 1990 "Toxic Reckoning: Business Faces a New Kind of Fear." *Harvard Business Review* 90(1):118-126.
- 1994 *A New Species of Trouble: Explorations in Disaster, Trauma, and Community*. New York: Norton.
- Freudenburg, William R. 1988 "Perceived Risk, Real Risk: Social Science and the Art of Probabilistic Risk Assessment." *Science*. 242:44-49.
- Gusfield, Joseph 1981 *The Culture of Public Problems: Drinking-Driving and the Symbolic Order*. Chicago: University of Chicago Press.
- Heimer, Carol A. 1988 "Social Structure, Psychology, and the Estimation of Risk." *Annual Review of Sociology* 14:491-519.
- 1992 "Your Baby's Fine, Just Fine: Certification Procedures, Meetings, and the Supply of Information in Neonatal Intensive Case Units." In James F. Short, Jr., and Lee Clarke, eds., *Organizations, Uncertainties, and Risk*.
- , and Lisa R. Staffen 1998 *For the Sake of the Children: The Social Organization of Responsibility in the Hospital and the Home*. Chicago: University of Chicago Press.



- Hersey, John 1985 *Hiroshima*. New York: Knopf.
- Janis, Irving Lester 1951 *Air War And Emotional Stress*. New York: McGraw-Hill.
- Johnson, Norris R. 1987 "Panic and the Breakdown of Social Order: Popular Myth, Social Theory, Empirical Evidence." *Sociological Focus* 20(3):171-183.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky 1982 *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Mellers, B. A., A. Schwartz, and A. D. J. Cooke 1998 "Judgment and Decision Making." *Annual Review of Psychology* 49:447-477.
- Oliver-Smith, Anthony 1996 "Anthropological Perspectives on Hazards and Disasters." *Annual Review of Anthropology* 25:303-328.
- Perrow, Charles 1984 *Normal Accidents: Living with High Risk Technologies*. New York: Basic Books.
- Roberts, Karlene H., ed. 1993 *New Challenges to Understanding Organizations*. New York: Macmillan.
- Sagan, Scott 1993 *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons*. Princeton, N.J.: Princeton University Press.
- Shrivastava, Paul 1987 *Bhopal: Anatomy of a Crisis*. Cambridge: Ballinger.
- Sorokin, Pitirim A. 1968 [1942] *Man and Society in Calamity*. New York: Greenwood Press.
- Starr, Chauncey 1995 "A Personal History: Technology to Energy Strategy." *Annual Review of Energy and the Environment* 20:31-44.
- Vaughan, Diane 1996 *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago: University of Chicago Press.
- 1999 "The Dark Side of Organizations." *Annual Review of Sociology* 25:271-305

LEE CLARKE

## SOCIOBIOLOGY, HUMAN

### THE DARWINIAN SETTING

Sociobiology is the term used to describe a relatively recent stage in the continuing development of evolutionary biology. It systematically brings the study of social behavior under the umbrella of the Synthetic Theory (or the Modern Synthesis) that, starting in the 1920s, arose from the marriage of Darwinian theory and Mendelian, or genetic, science (Huxley 1942). The most challenging aspect of the new elaboration concerns a decisive

step into human behavior. Sociologists (and social scientists in general) have not responded with enthusiasm; the old anthropocentrism with its extreme stress on culture and socialization (environmentalism) is still dominant. Resistance, however, is slowly breaking down, and one may speak of a human sociobiology taking the form of "evolutionary anthropology," "evolutionary psychology," "evolutionary sociology", and so forth. It can even be stated that human sociobiology may represent the beginning of the long-desired synthesis of the social sciences. At the same time, the introduction of cultural parameters into evolutionary explanation may further enrich the modern synthesis.

The term "sociobiology" harks back to the mid-1940s, and the evolutionary study of behavior began to develop rapidly only in the 1960s. However, its roots can be traced back to Darwin's ([1859] 1958) theory of evolution by natural selection, still the cornerstone of evolutionary science. From today's perspective, Darwin's theory can be conveniently stated as follows:

1. The rate of reproduction in populations, such as species, tends to be faster than the growth of the resources needed to sustain all their members.
2. As a result, populations experience a real or potential scarcity of resources.
3. This scarcity stimulates the "struggle for existence": competition of various kinds both within and between populations.
4. Some individuals in any given population are more successful than others in the struggle and thus are more likely to survive long enough to reproduce.
5. This differential reproductive success is in the last analysis the result of "variations" between individuals: genetic differences. That is, some variations are better suited (adapted) than others for the competition.
6. The better adapted these variations are, the more likely they are to be inherited by one's descendants.
7. This generational preservation of favorable variations and the concomitant elimination of unfavorable ones are together referred to as "natural selection."

This theory either contains or fosters the basic elements of the theoretical program of sociobiology. Specifically, (1) it establishes natural selection as the basic mechanism in the evolution of behavior, (2) its stress on competition calls attention to the perennial question regarding the role of altruism and selfishness in social life, (3) its focus on heredity suggests that aspects of human behavior, including culture, may be at least partly the result of natural selection and thus of forces long at work in human evolution, (4) perhaps more important, it guides inquiry toward a systematic theory of human nature. The crucial question of this theory is: In the course of human evolutionary time, what *innate* behavioral tendencies (predispositions, psychological adaptations, epigenetic rules, etc.) have been forged by natural selection and environmental pressures acting on genetic matter? These adaptations are likely to be implicated to some degree in socialization processes, thus proposing a more complete explanation in social science.

#### ELEMENTS OF SOCIOBIOLOGICAL THEORY

The modern synthesis stimulated scientific activity in general and, through such disciplines as entomology, primatology, and ethology, paved the way to the evolutionary study of behavior. Then in 1975 a seminal work proclaimed the advent of a "new synthesis" (Wilson 1975). The master stroke of sociobiology is the revival of "the struggle for existence," that is, the *behavioral* aspect of evolution on which natural selection was clearly based in the work of Darwin. The ultimate consequence of this struggle, or competition for resources, is natural selection: the differential contribution of offspring to future generations. This differential is a rough measure of what Darwin, following the sociologist Herbert Spencer, called "fitness." Accordingly, we are already in a position to state that *in the last analysis* individuals (or simply organisms) may be productively viewed as being in competition with one another for reproductive success, or genetic fitness. (Without the behavioral component, the statement would read, classical Darwinian logic: Organisms are the descendants of the more reproductively successful organisms.)

**The Maximization Principle.** This Darwinian-Mendelian idea is rendered more formally by what is considered the general law or principle of

sociobiology. Sometimes referred to as the maximization (or fitness) principle, this law states that, while organisms engage in all sorts of behaviors, in the last analysis they tend to behave in ways that maximize their inclusive fitness, or the chance of conveying their genotype (genetic makeup) to future generations. This is a probability statement: That is, some individuals are more successful than others in the reproductive competition. This idea of variability reflects the logic of natural selection and accommodates many facts that may seem strange or contradictory of evolutionary theory. A case in point is the parental abuse of children. Clearly, if organisms differ in the degree to which they behave adaptively one effective way to show this is for them to differ in the way they treat their children. However, this statement is heuristically a bit "lazy": It stimulates research insufficiently. Like many other laws in science, it would be more useful if it took the form of a contingent proposition. A first approximation in this direction is available and the provisos proposed are "creature comforts," "self-deception," and "autonomization of behavior" (the tendency of means, such as wealth, to become ends in themselves), all of which appear to condition negatively the maximization tendency (Lopreato 1989; for incisive analyses of this argument, see Crippen 1994; Maryanski 1998, pp. 11-16; Maryanski and Turner 1998, pp. 128-131).

The use of the adjective "inclusive" is intended to underscore the fact, better understood in the post-Darwinian period (Hamilton 1964), that the fitness of organisms is measured in terms of their relatedness both to their offspring and to other members of their genetic kin, whom they typically favor in many fundamental ways over nonkin. Hence, family life issues are central foci of sociobiology. Also, the maximization principle does not fit neatly into the mold of individual experiences. Human beings are not overtly obsessed with the enhancement of their fitness, and some, as was noted above, actually behave maladaptively. It is necessary to keep in mind, therefore, that the principle does not assume consciousness of the fitness consequences of behavior.

The fitness principle performs various functions. The crucial one is to logically structure established discoveries and thus stimulate cumulative, systematic knowledge. Central in this undertaking is the discovery of the mental rules that may be said to constitute human nature. Are people,

for example, constituted to facilitate the persistent prejudices and ethnic affiliations that periodically flare up in bloody conflicts across the globe? Surely, people grow up with such prejudices and absorb them, but why is this learning universal? Is it possible that it is an effect of evolutionary forces? To learn any given behavior, one must have a capacity for it in the brain. Try as humans might, they never can be socialized to behave like foxes, or like any other animal. Aristotle was correct in his metaphor of the oak acorn: If an acorn becomes anything, it can only become an oak, never, say, a fig tree.

**The Law of Altruism.** As fitness theory develops, the principle surrounds itself with auxiliary statements of theory that facilitate the quest for a theory of human nature. In what follows, this article touches on a few of these. Let us start by noting that central to sociobiological reasoning is the metaphor of “the selfish gene” (Dawkins 1976), which recalls the so-called Hobbesian question of order: For whose benefit does the individual behave? Selfishness is such a prevalent concept in evolutionary science that according to Wilson (1975, p. 3), the question of its opposite, what the sociologist Auguste Comte termed “altruism,” constitutes “the central theoretical problem of sociobiology.” Acts that benefit others are commonly observed in all behavioral disciplines. How to explain them? The social and moral disciplines have been of little help, largely because of the ambiguity of their concepts. What seems altruistic to scholar X is viewed as selfish by colleague Y.

Sociobiologists have taken a major step toward the solution of this problem. Countless observations coupled with the logic of the fitness principle have led to the position that altruism strictly viewed refers to genetically “self-destructive behavior performed for the benefit of others” (Wilson 1975, p. 578). In short, genuine altruism reduces the benefactor’s fitness; hence, if it arises in a given population, natural selection may be expected to wipe it out fairly quickly. If, for example, Mary is driven by her genes to do good for John at the expense of her own reproductive interest (e.g., cohabit with him and then be abandoned childless at an age when her chances of marriage and/or reproduction are greatly reduced), her altruistic genes will not be represented in the next generation (unless they are conveyed by her blood kin). What is it, then, that people call altru-

ism? Typically, it is either favoritism toward kin (*nepotistic favoritism*) or favoritism accompanied by the expectation of reciprocation (*reciprocal altruism*). Both types benefit the “altruist,” sometimes with interest. Evolutionists’ venture into the topic of altruism has produced rich harvests, particularly the discovery of *kin selection and inclusive fitness* (Hamilton 1964) and the theory of *reciprocal altruism* (Trivers 1971), which together yield the law of altruism.

*Kin selection and inclusive fitness* are dramatically illustrated by the study of eusocial insects, such as ants. Approximately three-quarters of these animals are female (“workers”), and very few reproduce. It would be a mistake, however, to consider them genuine altruists. Workers are so named because they are very diligent in catering to the needs of the queen (typically their reproductive mother) and her prodigious brood. Furthermore, given their peculiar reproductive system (*haplodiploidy*, whereby females have both parents while males, hatched from unfertilized eggs, have only a mother), workers are more closely related to the future generations than are their counterparts in diploid species such as mammals. As a result, failure to reproduce results in little or no loss in fitness. In short, eusocial insects have evolved according to kin selection and inclusive fitness. Indeed, this strategy is widespread among social animals. In humans, this fact is underscored by last wills and testaments, according to which people rarely bequeath their (fitness-enhancing) resources to anyone except blood kin (Clignet 1992).

*Kin Selection and Ethnic Conflict.* The familism inherent in kin selection has numerous expressions. One is related to the widespread phenomenon of ethnic identification and the recurrent cases of ethnic violence that often take everyone but the participants by surprise. More than eighty years ago, Pareto applied the logic of kin selection to explain the formation of persistent groups such as ethnic groups. They are “natural formations,” he argued (1916, section 1022), “growing up about a nucleus which is generally the family, with appendages of one sort or another, and the permanence of such groups in time engenders or strengthens certain sentiments that, in their turn, render the groups more compact, more stable, better able to endure.” This evolutionary perspective on ethnicity and ethnocentrism had been foreshadowed by Sumner (1906) and was subsequently approximated

by several sociologists (e.g., Park and Burgess 1921; Gordon 1964). More recently, van den Berghe (1981) produced a thoroughly evolutionary theory of ethnicity and ethnic conflict (see also Lopreato and Crippen 1999, chap. 9).

Sociologists continue to debate the causes of ethnic phenomena, typically focusing on cultural factors, such as differences in language or religion, that are specific to given times and places. Are such factors relevant? Very probably, but ethnic phenomena are persistent and universal; across the globe there are numberless mixtures of peoples who have an awfully hard time living together in peace. Any universal phenomenon requires first and foremost a universal explanation.

In brief, during nearly all of human evolutionary history people lived in small bands of about twenty-five to fifty individuals, often surrounded by neighbors who coveted their resources. Warfare or the threat of it was frequent. Benedict (1934, pp. 7–8) described “primitive man”: “From the beginning he was a provincial who raised the barriers high. Whether it was a question of choosing a wife or taking a head, the first and important distinction was between his own group and those beyond the pale. His own group, and all its ways of behaving, was unique.” Intense internal solidarity was a precondition of survival. To practice it was also to practice kin selection and kin favoritism. Since the modern human brain evolved in such circumstances, it can be concluded that the tendency to identify with one’s own “clan”—to distinguish between “us” and “them”—is alive and well in human society. Ethnicity is an extension of the family. Nepotistic favoritism is wired in the brain, and at the *ultimate* or general level it and the kin selection to which it is inextricably associated are the cause of the persistence of ethnic identification and recurrent ethnic conflict. “Who am I?” “Who are mine?” These are enduring whispers in the human psyche.

Of course, it is difficult to answer such questions in the megasociety, and that is one reason why most of the time people live in a reasonably peaceful relationship with their neighbors. Still, people are attentive to markers of “weness” such as a common name, common historical experiences, distinctive cuisine and artistic expressions, and often a common language, in historical time if not now (van den Berghe 1981).

*Reciprocal altruism.* Reciprocal altruism refers to the fact that if and when people engage in actions that benefit others, they do so with the expectation, conscious or not, that the others will repay, especially if they are not related. This is the object of a much-tested and growing theory first stated by Trivers (1971). According to the basics of this theory, the evolution of reciprocal altruism was facilitated by three broad conditions: (1) repeated situations in which the value of altruism to beneficiaries was, in terms of fitness, greater than the cost incurred by the benefactors, (2) membership in a small group featuring little or no migration, thus enhancing the chances of reciprocity, (3) fairly equal ability between pairs of individuals to engage in mutual help. Trivers proceeds to argue that a system of reciprocal altruism is subject to “cheating”: Some individuals do not reciprocate the benefits they receive. Indeed, an underlying assumption of Trivers’s theory is that givers are motivated to receive more than they give. Accordingly, as reciprocal altruism was evolving, another set of adaptations, what Trivers terms a “psychological system,” was arising fairly in step with it; their function was to regulate cheating. They include emotions such as friendship, sympathy, trust, suspicion, and moralistic aggression, along with hypocrisy and feelings of guilt.

Combining the logic of kin selection and reciprocal altruism, it is possible to state a *law of altruism* as follows: *In keeping with the fitness principle, social organisms have evolved to favor others (1) in direct proportion to their degree of genetic relatedness to them and (2) to the extent that the benefit they derive from doing good to others is, in terms of fitness, equal to or greater than the cost of their altruism.*

The law is relevant to various human phenomena, and casts light on a number of puzzles. For instance, social scientists have noted that the exchange of gifts is a universal institution in human society (Mauss [1925] 1954). Moreover, a version of this tendency termed “potlatch” is in varying degrees “a universal mode of culture” (Lévi-Strauss [1949] 1969). In an extreme form of potlatch practiced by the “Indians” of Vancouver and Alaska, one gives with a view to crushing another and thus gaining “privileges, titles, rank, authority, and prestige” (Lévi-Strauss [1949] 1969, chap. 5). Specifically, gifts are given to a competitor with the shared understanding that the recipient will reciprocate with interest after a reasonable interval.

When such an obligation cannot be met, the recipient loses status, titles, and so forth. Social theorists have been eager to downplay the individual's selfish undercurrent and the social conflict it engenders. The law of altruism conversely predicts selfishness, competition, and precariousness of status in one's group.

Nevertheless, the extreme stress on selfishness cannot go unchallenged for the human species. The only known attempt of this sort avoids the genetic trap (altruists are by definition "selected out"). Using a biocultural perspective at the core of which is the evolution of *the idea of the soul* and *self-deception*, it concludes that behaviors intended to save the soul mimic fitness-enhancing behaviors (Lopreato 1984, pp. 207-235). Catholic nuns and Buddhist monks, for example, practice celibacy and other forms of ascetic behavior in view of immortal ends. Typically, they contribute little or nothing to the fitness-enhancing resources of their blood relatives. Yet their type manages somehow to "reproduce" itself. Genuine or "ascetic" altruism may be rare in human society, but it is a cultural fact.

**The Law of Anisogamy.** Human psychological adaptations may be divided into two major types. One is specieswide, referring to innate tendencies that in varying degrees cut across gender. Examples include tendencies toward kin favoritism, ethnic identification, reciprocal altruism, and cheating. The other major class accounts for the fact that throughout human society there have been some remarkable differences, as well as similarities, along sex lines (Trivers 1972; Kimura 1992; Lopreato and Crippen 1999). The differences are suggested by a basic diversity in physioanatomy. *Anisogamy*, the name given to it, refers to the difference in size and structure between male and female sex cells (gametes). Male gametes (sperm) are minuscule and contribute only genes to reproduction. They are produced in huge quantities almost continuously after the onset of sexual maturity. The reproductive potential of males is therefore huge, and some men have fathered thousands of children (Betzig 1986). By contrast, female gametes are much larger, are nutritious, and are produced *in utero* once in a lifetime. Then, beginning at menarche, they are released, typically one at a time, about once a month, so that on average women produce some 300 to 400 mature eggs in a lifetime, only a very small number of which are

likely to result in offspring. Women are constituted to bear the cost of pregnancy, nursing, and much of the protection and guardianship, at least during the offspring's tender years. In short, each child represents a huge reproductive investment for the typical woman.

By contrast, males make a very small investment. They do not get pregnant or suffer nausea; nor do they risk their life at the birth of a child and for months or years afterward. If we consider such other facts as abandonment, divorce, and the refusal or failure to provide child support, the level of paternal investment is on average truly puny. There is no intention here to condemn men or glorify women. It is a matter of trying to grasp certain facts in order to understand certain others. Males and females have evolved under the pressure of significantly different, though partly complementary, reproductive strategies, and much of their behavior is an effect of this fact. It is now time to state what was earlier termed the *law of anisogamy* (Lopreato 1992, p. 1998): *The two sexes are endowed with differing reproductive strategies, and their behaviors reflect that difference in direct proportion to their relevance to it.* The closer one gets to the fundamental activities of life (sexual behavior, family life, and among endless others the conditions that recall the division of labor in the clan, the type of society in which the human species spent 99.5 percent of its history), the more likely one is to observe the effects of anisogamy. The basic implications of anisogamy have been drawn by Trivers (1972) in a seminal paper on "relative parental investment."

*Differential Parental Investment and Sexual Selection.* The law of anisogamy contains a number of corollaries. The two major ones noted briefly here are very closely related. Differential parental investment (DPI) states in effect that females make both a greater initial parental investment and greater subsequent parental investments than do males, so that their behavior is more finely adjusted to the well-being and reproductive success of the offspring. Supporting facts are legion. They are epitomized by the following widely noted findings, among others: On average females are more cautious than males in their sexual activity, and they tend to prefer as mates men who are in fact, or show promise of becoming, relatively rich in the resources needed to raise healthy and reproductively viable offspring. Mating has always been far riskier for women, and this fact is deeply rooted in the

brain. As a consequence, it tends to express itself even in times and places where mating need not have reproductive consequences. For millions of years, and even today in much of the world, an unwanted pregnancy or a pregnancy with a partner who will contribute little or nothing to the well-being of the offspring may consume a large portion of a woman's reproductive potential. In summary, women go for quality; men, for the allure of quantity.

A study of 10,047 individuals living in various countries on six continents strongly suggests that culture has a notable influence on mate preferences and that the two sexes agree on some of the basic requirements of a good mate, such as honesty and dependability. However, the findings also reveal some marked universal differences that are predictable from the law of anisogamy. For instance, females are significantly more likely than males to emphasize ambition and industriousness. Conversely, males more than females prefer mates who are physically attractive, younger than themselves, and at the peak of their reproductive value even if they are of lower socioeconomic status (Buss 1989).

The sexual selection (SS) corollary may be stated in Trivers's (1972, p. 140) words as follows: "Individuals of the sex investing less will compete among themselves to breed with members of the sex investing more, since an individual of the former can increase its reproductive success by investing successfully in the offspring of several members of the limiting sex."

The idea of sexual selection is a remarkable example of great ideas that emerge out of creative confusion. Darwin understood that in the final analysis the measure of survival is reproductive success. Nevertheless, he tended to focus on survival as *longevity*. As a result, certain observations, both behavioral and physical, confronted him with a special challenge. Why the great horns, the displays, the mimicry, the special weapons, "the instrumental music," or, among other male characteristics, the huge tail of the peacock? Such unusual features tend to attract predators and thus reduce longevity. Darwin (1859, but especially 1871) concluded that if such "secondary sexual characters" enhance the bearers' ability to reproduce, they are likely to be favored by natural selection even if they act negatively on longevity. However, because

such characters were conspicuous in sexual competition, they suggested to him the label "sexual selection." There is some debate over the meaning of this term (see Mayr 1972 for a review), but it is safe to say that it refers not to a type of selection but instead to a major cause of natural selection. "After all," as Dobzhansky et al. (1977, p. 118) noted, "Darwinian fitness is reproductive fitness," whatever the cause. Sexual selection, or *competition for mates*, refers most explicitly to the struggle for genetic survival. It further suggests that much animal behavior and appearance are adapted not so much to the problem of daily survival as to the job of securing adequate mates.

Viewed as competition, sexual selection has become a valuable tool of research, especially in view of certain distinctions suggested by Darwin himself. In current language there are two major types of sexual selection. One, often termed "intrasexual selection," subsumes a female-female competition and a male-male one. The other refers to a form of male-male competition, too (the competition "to charm the females"), but Darwin viewed it as an effect of the females' response to "charm." Accordingly, it has come to be known as *female choice* (or "intersexual selection"). It is evident that differential parental investment and sexual selection are closely linked properties of anisogamy. In fact, one may combine the logic of DPI and SS to state what may be termed the DPI-SS corollary as follows: *Given anisogamy, females have been selected to engage in choosy behavior, while males have been selected to specialize in agonistic behavior.*

Female choosiness in human beings takes many forms. This article has mentioned the tendency to prefer resource-rich mates. As a group, even in a highly developed society like the United States, women pay less attention to looks, have and claim to want fewer sexual partners, are less likely to have their first sexual experience with a stranger, are more likely to expect a commitment before engaging in sex, and among many other differences, are less likely to cheat on their mates, whether husbands, cohabitators, or boyfriends (Laumann et al. 1994).

Competitiveness, too, takes many forms. One is reflected in the ancient and still fairly common practice of polygyny, especially if one considers the greater tendency of divorced men to remarry and have further children (Betzig 1986; Lenski

and Lenski 1987). Another form is expressed in violent behavior. Killing, for example, is throughout the world a largely male behavior and is concentrated among young men during the peak years of their reproductive life. Prominent among the motives for homicide are sexual jealousy and rivalry as well as dominance contests in various contexts (Daly and Wilson 1988). Death by trauma (murder and accidents) accounts for a relatively high percentage of male mortality and part of the lower life expectancy of men (Verbrugge 1989). Of course, there are many other forms of competitiveness, and women are not immune to the tendency to practice them.

*Of Sex Roles.* Sociological practice is almost entirely environmentalist and often clashes with an evolutionary perspective. For example, according to “feminist” authors, especially so-called *gender feminists*, men and women are born with identical potentials. The idea that, given anisogamy, the brain has been neuroendocrinally gendered (Kimura 1992) is extraneous if not altogether offensive to them. They argue conversely that society is ruled by a system of *patriarchy*, and socialization thus proceeds to produce differences that are detrimental to women. Girls are socialized to be passive and subordinate, while boys are trained to strive for success and dominance (Lerner 1986, p. 29). In short, so-called sex roles are the effect of culturally prescribed discrimination and must be explained in terms of cultural causes only. Some writers go so far as to argue that physiology is irrelevant or that “human physiology is *socially* constructed and gendered” (Lorber 1994, p. 46; emphasis added).

Sociobiologists do not deny that once an arrangement such as patriarchy is in place, the channels of socialization tend to develop in view of its mandates. However, science does not merely assert facts; it seeks to explain them. As Hrdy (1997, pp. 7–8; italics in the original) notes, “an evolutionary perspective pushes the search for patriarchy’s origins back . . . by millions of years by asking an additional question: *Why* should males seek to control females?” That is, why patriarchy in the first place? Accordingly, it less superficially identifies sexual selection “rather than male desire for power as the engine driving the system” of patriarchy.

To understand patriarchal phenomena and have more than a wishful chance of bringing effective cultural forces to bear on them, one must

begin by answering Hrdy’s question. In the process, a very unpleasant irony will be uncovered. The dynamics that produced patriarchy include female complicity with domineering males. Male dominance has been achieved at least in part because of female preference (female choice) for dominant males. Indeed, males dominate females by dominating other males with female help. Thus, the pickle that many people find so distasteful “turns out to have been seasoned with only a sprinkle of culture at best, although once culture arose it made it even more tartish. It is also true that our female ancestors had a hand in the preparation” (Lopreato and Crippen 1999). Moreover, women throughout the world continue to support patriarchy through their persistent tendency to favor dominant males.

## CONCLUSION

In order to advance, would-be sciences need to discover the value of a number of time-tested techniques. These techniques include especially the use of remote concepts and logical ways to operationalize them, the nomothetic derivation and testing of hypotheses, and the logical structuring of the hypotheses in a body of systematic, cumulative knowledge that facilitates further research and discovery. The history of science strongly suggests that to accomplish even these minimal feats, it is necessary to either discover one general principle or to borrow it in full or modified form from a cognate and more advanced science. At the start of the third millennium there is still no general principle in sociology, general in the sense that it would contain the logic for a large number of derivative statements linking discovery and explanation across the institutional framework. The fitness principle and the theoretical tools surrounding it constitute an invitation from sociology’s most proximate natural science to embrace the fact that the human brain represents a tenacious link to a past that in part is still present. The potential payoff is likely to be far greater than even the most sanguine evolutionists can imagine. It may suffice to consider that just as Newtonian laws eliminated the old prejudice of geocentrism, thus freeing the mind to behold previously inconceivable wonders of nature, sociobiology offers a human perspective from a distance, thus freeing the mind from the still-oppressive assumptions of

*anthropocentrism* and *temporecentrism*, benighting corollaries of geocentrism.

According to many critics of sociobiology, since certain behaviors are “natural,” they are not subject to cultural intervention. This is an error that sociobiologists do not commit. Because they are evolutionists, their theorizing is subject to a systemic perspective: “Phenotype” (any feature of anatomy, physiology, or behavior) is a result of the interaction between genotype and environment, including culture. It is essential to have knowledge of both. To change the world, as many scholars are inclined to do, with knowledge (always imperfect) of one to the exclusion of the other is not only obscurantism; it is poor, perhaps dangerous, engineering as well.

REFERENCES

Benedict, R. 1934 *Patterns of Culture*. Boston: Houghton Mifflin.

Betzig, L. L. 1986 *Despotism and Differential Reproduction: A Darwinian View of History*. New York: Aldine.

Buss, D. M. 1989 “Sex Differences in Human Mate Preferences: Evolutionary Hypotheses Tested in 37 Cultures.” *Behavioral and Brain Sciences* 12:1–49.

Clignet, R. 1992 *Death, Deeds, and Descendants*. New York: Aldine.

Crippen, T. 1994 “Toward a Neo-Darwinian Sociology: Its Nomological Principles and Some Illustrative Applications.” *Sociological Perspectives* 37:309–335.

Daly, M., and M. Wilson 1988 *Homicide*. New York: Aldine.

Darwin, C. (1859) 1958 *The Origin of Species*. New York: Mentor.

— 1871 *The Descent of Man and Selection in Relation to Sex*. New York: Appleton.

Dawkins, R. 1976 *The Selfish Gene*. Oxford: Oxford University Press.

Dobzhansky, T., F. J. Ayala, G. L. Stebbins, and J. W. Valentine 1977 *Evolution*. San Francisco: Freeman.

Gordon, M. M. 1964 *Assimilation in American Life*. New York: Oxford University Press.

Hamilton, W. D. 1964 “The Genetical Theory of Social Behaviour: I and II.” *Journal of Theoretical Biology* 7:1–52.

Hrdy, S. B. 1997 “Raising Darwin’s Consciousness: Female Sexuality and the Prehominid Origin of Patriarchy.” *Human Nature* 8:1–49.

Huxley, J. 1942 *Evolution: The Modern Synthesis*. New York: Harper.

Kimura, D. 1992 “Sex Differences in the Brain.” *Scientific American* 268:119–125.

Laumann, E., R. Michael, S. Michaels, and J. Gagnon 1994 *The Social Organization of Sexuality*. Chicago: University of Chicago Press.

Lenski, G., and J. Lenski 1987 *Human Societies*. New York: McGraw-Hill.

Lerner, G. 1986 *The Creation of Patriarchy*. Oxford: Oxford University Press.

Levi-Stauss, C. (1949) 1969 *The Elementary Structures of Kinship*. Boston: Beacon.

Lopreato, J. 1984 *Human Nature and Biocultural Evolution*. London: Unwin Hyman.

— 1989 “The Maximization Principle: A Cause in Search of Conditions.” In R. W. Bell and N. J. Bell, eds., *Sociobiology and the Social Sciences*. Lubbock: Texas Tech University Press.

— 1992 “Sociobiology.” In E. F. Borgatta, ed., *Encyclopedia of Sociology*. New York: Macmillan.

—, and T. Crippen 1999 *Crisis in Sociology: The Need for Darwin*. Rutgers, N.J.: Transaction.

Lorber, J. 1994 *The Paradoxes of Gender*. New Haven, Conn.: Yale University Press.

Mauss, M. (1925) 1954 *The Gift*. London: Cohen and West.

Maryanski, A. R. 1998 “Evolutionary Sociology.” *Advances in Human Ecology* 7:1–56.

—, and J. H. Turner 1998 “New Evolutionary Theories.” In J. H. Turner, ed., *The Structure of Sociological Theory*, 6th ed. New York: Wadsworth.

Mayr, E. 1972 “Sexual Selection and Natural Selection.” In B. G. Campbell, ed., *Sexual Selection and the Descent of Man*. Chicago: Aldine.

Pareto, V. (1916) 1963 *A Treatise on General Sociology*. New York: Dover.

Park, R. E., and E.W. Burgess 1921 *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.

Sumner, W. G. 1906 *Folkways*. New York: Ginn.

Trivers, R. L. 1971 “The Evolution of Reciprocal Altruism.” *Quarterly Review of Biology* 46:35–47.

— 1972 “Parental Investment and Sexual Selection.” In B. H. Campbell, ed., *Sexual Selection and the Descent of Man, 1871–1971*. Chicago: Aldine.

van den Berghe, P. L. 1981 *The Ethnic Phenomenon*. New York: Elsevier.

Verbrugge, L. M. 1989 “The Twain Meet: Empirical Explanations of Sex Differences in Health and Mortality.” *Journal of Health and Social Behavior* 30:282–304.



Wilson, E. O. 1975 *Sociobiology: The New Synthesis*. Cambridge, Mass.: Harvard University Press.

JOSEPH LOPREATO

## SOCIOCULTURAL ANTHROPOLOGY

In the United States, anthropology usually is considered to consist of four subdisciplines, or “subfields”: archaeology (describing and understanding past human behavior by examining material remains), physical or biological anthropology (describing the evolution and modern physical variation of the human species), anthropological linguistics, and sociocultural anthropology. Most university departments of anthropology have faculty in three or four of these subdisciplines. Sociocultural anthropology often is called simply cultural anthropology in the United States, although a few academic programs use the term “social anthropology,” the common designation in Europe. Some anthropologists identify applied anthropology as a fifth subfield, while others consider it part of sociocultural anthropology.

Anthropology is defined as the study of human commonalities and differences and expressly includes the entire temporal and geographic range of humankind in its scope. The database of the discipline is large, including prehistoric populations as well as every variety of contemporary society. In distinguishing itself from other social sciences, anthropology emphasizes the holistic, comparative, culture-centered, and fieldwork-dependent nature of the discipline.

In Europe, social anthropology is more closely allied with economics, history, and political philosophy than it is with physical anthropology and archaeology, which often are taught in separate programs. As social anthropology evolved in Europe, it came to be associated with studies of the economy, ecology, polity, kinship patterns, and social organization of non-Western peoples, particularly in colonial Africa and Asia. The European approach to theory was associated with sociological (especially functionalist) and, more recently, historical approaches. In the United States, where research focused initially on Native Americans and was strongly influenced by the particularistic descriptive approach of Franz Boas’s ethnogra-

phy, anthropology came to be associated with culture, that “complex whole” (in Edward Tylor’s words) encompassing customs, language, material culture, social order, philosophy, arts, and so on. European social anthropologists have not failed to address culture and Americans have not neglected social structure, yet the difference in terminology distinguishes an emphasis on social relations from an emphasis on shared meaning and behavior.

The heart of sociocultural anthropology is ethnography, the written description of a culture group. Ethnography has undergone many changes since it began with field reports by missionaries and colonial officials. The pace of change has increased since the 1960s, as recognition of global links has become standard, other social scientists have adopted ethnographic methods, and postmodernism has imposed stricter self-reflective criteria on writers. The methodological partner of ethnography is ethnology, the comparative study of societies. In its first decades, anthropology established the ideal that a complete ethnographic record of the world’s cultures would allow comparative studies that would lead to generalizations about the evolution and functioning of all societies. Cross-cultural studies continue to be one of the distinctive contributions of anthropology to the social sciences.

## HISTORY

Anthropology and sociology share common origins in the nineteenth-century European search for a science of society. Sociocultural anthropology and sociology also share a theoretical history in the ongoing struggle between the desire for a generalizing, rule-seeking science and that for a humanistic reflection of particular lives. Throughout the twentieth century, academic specialization and differences in research topics, geographic focus, and methodological emphasis separated the two disciplines. In the last several decades, globalization has fostered a partial reconvergence of methods and subjects, though not of worldviews, ethos, or academic bureaucracies.

Sociocultural anthropology often is contrasted with sociology: It is said that anthropologists study small-scale societies, assume that those societies are self-sufficient, and are usually outsiders (politically, ethnically, and economically) to the groups they study. These generalizations are partly true.

The methods of sociocultural anthropology have emphasized the usefulness of seeking “the large in the small” by becoming intimately acquainted with a single band, village, tribe, island, or neighborhood, and anthropology’s early link to colonialism and its base of support in Europe, Japan, China, and the United States has privileged wealthy outsiders as observers of peasants, tribal peoples, and marginalized groups. However, anthropology has always kept the larger picture in mind, and for every study of an “isolated” population, there are ethnographies that reveal links at the regional, national, and global levels. The affiliation of sociocultural anthropology with archaeology and paleoanthropology ensures that the long term and the large scale are never far from sight. Ethnographies of industrialized societies, ranging from ethnic minorities to corporate cultures, begin with the microcosm but connect to larger questions. Sociology has been associated from its beginnings with studies of modernization and globalization in Western societies. In the postwar world, anthropologists became of necessity students of these processes in the same small communities that had been their prewar subjects of study. Anthropologists have sought ways to encompass urban life, regional processes, and global economic and political transformations in their work, leading them to develop skills in quantitative social research as well as their traditional qualitative methods.

Developments in method and theory in the twentieth century have led to a widely perceived split between sociocultural anthropologists who seek a “natural science of society” and those who emphasize anthropology’s humanistic role as an interpreter of cultural worlds. These differences are reflected in the distinction between “emic” and “etic” strategies. Based on the linguistic concept of the phoneme, emic work calls for the researcher to understand the “inside” view, focus on meaning and interpretation, and “grasp the native’s point of view . . . to realize *his* vision of *his* world,” in Bronislaw Malinowski’s words. A good ethnography enables readers to understand the motives, meanings, and emotions of a different cultural world. The etic (from “phonetic”) approach seeks generalizations beyond the internal cultural worlds of actors, applying social science concepts to the particulars of a culture and often using cross-cultural comparisons to test hypotheses. A good ethnography presents data that can be

compared with other cases. In recent years, the writing of ethnography has self-consciously struggled to develop a style that can evoke the sensibility of a culture while including descriptive information in a format that allows cross-cultural comparisons.

Sociocultural anthropology begins with description and usually intends that description (ethnography) to be a prelude to cross-cultural comparison that will lead to generalizations about types of societies or even about human universals. At the same time, anthropologists are as likely as other social scientists to be influenced by fashions in theory.

### THEORY

The nineteenth-century origins of anthropology, like those of sociology, are rooted in the expanding inquiry into the nature of human society that characterizes the nineteenth and early twentieth centuries, but anthropology’s roots also involve the questions of biological and social evolutionism characteristic of the era, as epitomized in the work of Charles Darwin and Herbert Spencer. Anthropology and sociology share origins in the foundational work of Durkheim, Weber, and Marx. However, cultural anthropology adds to its pantheon of ancestors Tylor, Morgan, and Frazer; it is in the work of these three men that one can see how anthropology was set on a different trajectory. The American Lewis Henry Morgan (*Ancient Society*, 1877) and the British Edward Burnett Tylor (*Primitive Culture*, 1871) and James Frazer (*The Golden Bough*, 1890) are counted among the founders of anthropology because they sought to establish general laws of human society through the comparative study of historical and contemporary peoples. Tylor, Morgan, and Frazer were unilineal evolutionists who believed that universal stages of evolution could be identified in the transition from simple to complex societies and that modern peoples could be ranked in this evolutionary scale. These two strands—the belief that comparison can produce scientific generalizations and the search for evolutionary processes—continue to characterize anthropology, though the racist evolutionism of these early approaches was discarded as anthropology was established as a discipline in the 1920s and 1930s.

While the work of the nineteenth-century social theorists presaged both anthropology and sociology, by the turn of the century, each field was established in separate academic departments and increasingly distinct research programs. In the United States, anthropology as a scholarly project emerged through the work of scholars drawn to the task of reconstructing Native American cultures and languages, especially under the auspices of the Bureau of American Ethnology and the formative political, administrative, and scientific work of Franz Boas. Boas responded to the prevailing ideas of unilineal evolutionism with a theory that came to be called historical particularism, rejecting broad generalizations about stages of evolution in favor of detailed studies of the environmental context and historical development of particular societies. Boas also trained the first generation of professional anthropologists in the United States, and his students, such as Alfred L. Kroeber, Robert Lowie, and Edward Sapir, pioneered new theories that could replace unilineal evolutionism. Sapir's and Benjamin Whorf's work on links between language and culture, Margaret Mead's on enculturation and psychological anthropology, Ruth Benedict's on ethos, Zora Neale Hurston's on folklore, and Kroeber's on the superorganic all fostered decades of theoretical development that pushed American anthropology in distinctive directions. Field studies with Native Americans and other North American minorities honed the skills of the first generations of American anthropologists in linguistic work, informant interviews, life histories, and historical reconstruction and established the holistic style of American anthropology, integrating archaeology, linguistics, and physical anthropology with the study of society and culture.

While Boas's students filled library shelves with detailed and impressive ethnographies, a new theoretical orientation developed in Great Britain that would have a great impact on the culture-centered world of American anthropology. This was functionalism, and its key proponents in anthropology were Bronislaw Malinowski (psychological functionalism) and A. R. Radcliffe-Brown (structural functionalism). The period of interest in the ways in which cultural institutions maintain social order—which affected the United States when Radcliffe-Brown and Malinowski spent time at American departments of anthropology in the

1930s—marks the point at which most texts officially distinguish British social anthropology from American cultural anthropology. Radcliffe-Brown countered Boasian particularism with an emphasis on the search for general laws of society and stimulated a generation of European and American students to do the same. British social anthropologists turned their analytic focus on the study of persons and relations in persisting social structures and pushed themselves and their students to develop the close observation, incisive analysis, and careful record keeping that marked the coming of age of long-term participant observation as a research method. Functionalist studies took place in the context of colonialism, with the limitations and power imbalance that that implies, yet remain impressive for the quality of detail and their capacity to integrate descriptions of political, economic, and kinship relations. Many ethnographic classics were produced by British social anthropologists of that era (e.g., Malinowski's *Argonauts of the Western Pacific* in 1922 and Evans-Pritchard's *Witchcraft, Oracles and Magic among the Azande* in 1937 and *The Nuer*, 1940) and their students, including Raymond Firth, Meyer Fortes, Audrey Richards, Lucy Mair, Edmund Leach, Max Gluckman, and Fred Eggan.

While American anthropologists added the study of social structure and function to their repertoire, they did not abandon their interest in historical developments, language, personality, and ethos and retained a “four-fields” orientation in the training of graduate students. While some social anthropologists found the idea of culture impossibly vague, American anthropologists reveled in the complexity of the concept, with Kroeber and Kluckhohn assembling a compendium of more than 150 definitions of “culture.” Stimulated by the challenge of British social anthropology, the work of Kroeber, Mead, Benedict, and Sapir from the 1920s through the 1950s explored culture as a distinct level of analysis and a way to grasp the distinctive ethos and worldview of each culture, along with the active role of the individual's acts and words in shaping a culture.

In the 1940s and 1950s, the influence of materialist approaches in the social sciences, while limited by the anticommunism in American public life (explicitly Marxist approaches did not appear until the 1970s), was manifested in a new set of evolutionary and generalizing approaches in Ameri-

can anthropology. The work of Julian Steward and Leslie White laid the groundwork for a new approach to studies of adaptation and cultural change. White argued for an evolutionary scheme in which culture (the uniquely human capacity to manipulate symbols), as the superorganic human adaptive mechanism, develops through evolutionary stages marked by the increasing ability of human groups to capture energy through technological systems. Steward worked on a smaller scale, arguing for the analysis of structural similarities among cultures at a regional level, which can be understood by recognizing the hierarchical relations among three “levels of sociocultural integration”: techno-economics (infrastructure), sociopolitical organization, and ideology (superstructure). Steward’s scheme allowed anthropologists to catalogue cultures as structural types and encouraged the study of change over time in a “multilineal evolutionary” process that he contrasted with White’s more abstract global stages.

Materialist studies continued to develop and to shape archaeology as well as cultural anthropology. Marshall Sahlins and Elman R. Service merged White’s and Steward’s approaches in a neoevolutionist theory that encouraged both archaeologists and materialist-oriented sociocultural anthropologists to consider the regional and large-scale classification and development of societies. Marvin Harris, Eleanor Burke Leacock, and Morton Fried attempted to explain cultural diversity and change in the context of the causal primacy of production and reproduction. In the 1960s and 1970s, the new field of cultural ecology developed a “neofunctionalist” approach that allowed scientists to include cultural and social aspects of human behavior in natural science research. Roy Rappaport’s 1967 *Pigs for the Ancestors* began with an effort to measure the energy intake and outflow of a highland New Guinea population; the 1984 edition included a lengthy discussion of criticisms of neofunctionalist theory and the applicability of adaptive and evolutionary concepts to human groups.

In France, Claude Levi-Strauss was developing ideas that would transform the world of social science through structuralism, which emerged in the 1960s and 1970s as a totalizing theory aiming at uncovering the common structures of the human mind. Structuralism, which was influenced by the linguistics theories of Saussure and Jakobson, treated the products of culture as symbolic sys-

tems and examined the formal patterns of those systems in order to envision discern universal structures and cognitive patterns of the human mind. Structuralism was applied to myths, kinship, relations to art, and every other aspect of culture. The work of Levi-Strauss, Edmund Leach, and other structuralists drew sharp rebuttals from theorists who sought explanations of human diversity in material and social conditions rather than in mental templates. Although the abstractness of structuralism eventually limited its interest to students of culture, it continues to be a useful technique, particularly in the analysis of the symbolic products of culture.

Ethnoscience, which emerged in the 1950s, also examined the mental categories underlying cultural products. Drawing heavily on linguistic theory and methodology, ethnoscience tried to develop fieldwork methods sufficiently rigorous to delineate the mental models that generate words and behavior and, in its emphasis on the emic approach, insisted on the necessity of fully accessing the native understanding of cultural domains. As ethnoscience faded in importance in the 1970s, it was succeeded by cognitive anthropology, the cross-cultural study of cognition.

Structuralism, ethnoscience, and responses to materialist neoevolutionist theory stimulated the emergence of symbolic anthropology and cultural analysis in the 1960s and 1970s, and this in turn led to the “interpretive turn” that has continued in cultural anthropology through the rest of the century. Again, linguistics proved influential, as David Schneider, Clifford Geertz, and Victor Turner explored new ways to study the cultural construction of meaning and the public representation of meaning in cultural elements. Most symbolic anthropologists focus on the description and interpretation of particular cultural cases, emphasizing the ethnographer’s role in explicating cultural events or products, though a few symbolic anthropologists, such as Mary Douglas, have sought general models of symbol systems. Symbolic anthropology shifted in the 1980s toward interpretive anthropology, which in turn generated a decade of reflection on the writing of ethnography, seeking modes of representation that would represent the worldview, internal logic, and emotional sensibility of a culture. Emerging from interpretive approaches have been experiments in ethnography, renewed interest in life histories, and extensive

critiques of an etic-oriented ethnography that relies on the authoritative voice of an “outside” observer and author. The 1980s also saw a new interest in history, spurred in part by the work of French scholars such as Braudel, Bourdieu, and Foucault and also playing a part in drawing some sociocultural anthropologists toward humanistic approaches.

American cultural anthropology has always taken an interest in evolutionary questions, and in the 1970s, the biologist E. O. Wilson used sociobiology to challenge social scientists to study the role of natural selection in human behavior. Anthropologists’ immediate response was to criticize sociobiology as sociologically naive, culture-bound, and potentially racist and sexist. In the longer term, however, this challenge renewed anthropologists’ interest in the holistic approach to culture, stimulating new approaches to the flexible and complex linkage of genetic inheritance and cultural malleability. Archaeologists, physical anthropologists, and cultural anthropologists share an interest in these long-term questions, which now are studied as “human behavioral ecology.”

### ORGANIZATION

While anthropological theory has participated in many of the trends in the social sciences in this century, anthropologists most often speak of themselves in terms of the topics they study and the geographic areas in which they are expert. A cultural anthropologist might say that she studies “gender issues in the Middle East,” “political hierarchy in Polynesia,” or “hunter-gatherer ecology in the Arctic,” with the implication that her theoretical school is a less useful category or that one might include several different theoretical or methodological approaches to one’s topic.

A review of textbooks in anthropology and courses offered in larger departments provides an indication of the overlap and the difference in range between sociological and anthropological topics. Traditional topics in anthropology include the categories of sociopolitical life (political anthropology, the anthropology of religion, social organization, patterns of subsistence, economic anthropology), cross-cultural approaches to all social science topics (ethnicity and identity, psychological anthropology, urban anthropology, ethnohistory, gender), theoretical approaches (sym-

bolic anthropology, cultural ecology), applied topics (legal anthropology, developmental anthropology, culture change, medical anthropology, education and culture), and topics reflecting the persistent holism of the anthropological enterprise (language and culture, genetics and behavior).

Anthropologists’ regional focus traditionally has been small-scale non-Western societies, but this has changed dramatically in the last fifty years. While sociologists and other social scientists have become more active in non-Western contexts (particularly economic development and modernization), anthropologists have become more active in studying Western societies, using their traditional skills of small-community ethnography, cultural models, and comparison in these situations. However, as part of their postgraduate training, most American and European anthropologists do a lengthy period of participant observation research in a small-scale society, usually a foraging band or a tribal or peasant society.

One stimulus to anthropologists’ willingness to become wholeheartedly involved in the study of Western, industrialized, and mass societies has been the growth in applied work. While sociology was committed to researching public policy issues from its beginning, anthropology has only intermittently taken on research directed at social problems and policy issues. Beginning with government work during World War II and the postwar Fox and Vicos projects in applied anthropology and as a result of globalization and limited academic job opportunities for anthropologists, there has been an increase in putting anthropological concepts and methods to the service of immediate outcomes rather than academic research. The greatest demand for applied anthropology is in economic and social development, medical anthropology, the anthropology of education, and international business.

### METHODS

Anthropology was born in the theories of “armchair anthropologists” who based their theories about the evolution of human beliefs and societies on the reports of colonial officials, missionaries, and merchants. Since that time, the commitment of researchers such as Boas, Mead, and Malinowski to detailed, long-term field studies has generated

the impulse that has sustained generations of anthropologists in an effort to produce detailed, fine-grained, firsthand descriptions of the world's cultures. Cultural anthropology has long held that long-term participant observation, including mastery of local languages, is the best way to produce valid ethnographic description. Participant observation is the source of anthropology's ethnographic database and the foundation on which controlled cross-cultural comparison is built.

The work of field research and the writing of ethnography have received much attention in recent decades. Participant observation is now an umbrella term for a research project that, while it extends over the long term (usually at least a year) and relies on the use of the local language, key informants, and living "close to the ground" with the people being studied, is likely to include a range of additional research techniques. Sociocultural anthropologists also are trained in kinship analysis, unstructured and structured interviews, questionnaires, scales, taxonomies, and direct and unobtrusive observation. In the past decade, there has been a growing expectation that researchers will combine qualitative and quantitative research methods, increasing both the validity and the reliability of ethnographic work. Applied anthropology has generated its own methods, some of them shaped by the time and cash restraints of nonacademic research, such as rapid rural assessment, participatory appraisal, and decision-tree modeling.

Cross-cultural comparison has been a goal of anthropology from the start. The first armchair anthropologists used sometimes unreliable secondhand information to generate categories and stages of social evolution, but researchers soon employed more scientific methods. Archaeologists' work on regional and chronological linkages encouraged ethnologists to trace the development, distribution, and diffusion of culture traits (especially in the United States, with Boas's encouragement). British social anthropologists and the neoevolutionists urged the use of regional and global comparisons to generate models of structural stability and change. George P. Murdock greatly facilitated large-scale comparison when he created the Human Relations Area Files, the physical form of the great database of human cultures anthropology had long sought. Cross-cultural studies in anthropology have allowed anthropologists to generate and test midlevel hypotheses about

cultural patterns and allowed social scientists to test the broader validity of hypotheses generated in Western contexts.

### CURRENT ISSUES

In surveying the history of anthropological theory, one often notices the persistent tension between materialist and idealist ways of studying culture. In the current environment, after a decade of postmodern critiques, this tension has actually split a few academic departments, severing archaeology and biological anthropology from cultural anthropology, or "scientific" from "humanistic" approaches. Research specialization and job-market pressures also interfere with the holistic four-fields approach that American anthropologists have long considered their hallmark. In addition, sociocultural anthropology has been pressed by the inroads of literary criticism, cultural studies, ethnic studies, and other related fields into its traditional preserve. Like other social sciences, anthropology feels that it is living through a "crisis" that represents both a point in a repeated cycle of theoretical change and a response to national and global contexts.

However, the end of the twentieth century has seen a wider range of research and applied work than had ever been done previously (see recent issues of *American Anthropologist*, *American Ethnologist*, *Current Anthropology*, and *Human Organization*). Current work in anthropology includes traditional detailed ethnographies that aim to increase the descriptive database of the world's cultures, problem-focused fieldwork aimed at elucidating theoretical puzzles, reflexive ethnography that attempts to find a moral and artistic center from which to write, analyses of organizations and evaluations of programs intended to guide policy decisions, and hypothesis-testing data crunching. The long-standing distinction between materialist and idealist approaches continues as interpretive, postmodern anthropology seeks new ways to do the job it has been critiquing for a decade and as ecological, evolutionist, and materialist approaches argue with renewed vigor for a scientific discipline.

Sociocultural anthropology and sociology share modern interests in agency; power; the relative role of social structures and individual action in culture change; the intersections of ethnicity, class,

and gender; and the historical shaping of modern institutions and cultural representations. In all its interests, ongoing input from archaeology, biological anthropology, and linguistics has given sociocultural anthropology a uniquely broad and deep perspective on the human condition, and its stream of theory is fed from these other sources of knowledge about the human condition. In describing the commonalities that unite cultural anthropology, Rob Borofsky speaks of anthropologists' shared ethics: a desire to publicize "human commonalities" (especially in countering racism), the valuing of cultural diversity, and the use of cultural differences "as a form of cultural critique" of the anthropologist's home culture and in general of industrial mass society. Despite an explosion of variation in what sociocultural anthropologists do, anthropologists' holistic and comparative worldview remains distinctive.

REFERENCES

Bernard, H. Russell 1994 *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. Thousand Oaks, Calif.: Sage.

Borofsky, Robert, ed. 1994 *Assessing Cultural Anthropology*. New York: McGraw-Hill.

Harris, Marvin 1968 *The Rise of Anthropological Theory: A History of Theories of Culture*. New York: Crowell.

Hatch, Elvin 1973 *Theories of Man and Culture*. New York: Columbia University Press.

Kaplan, David, and Robert A. Manners 1972 *Culture Theory*. Prospect Heights, Ill.: Waveland Press.

Kuper, Adam 1983 *Anthropology and Anthropologists: The Modern British School*. London: Routledge and Kegan Paul.

Kuznar, Lawrence 1997 *Reclaiming a Scientific Anthropology*. Walnut Creek, Calif.: AltaMira Press.

Marcus, George E., and Michael M. J. Fischer 1986 *Anthropology as Cultural Critique: An Experimental Moment in the Human Sciences*. Chicago: University of Chicago Press.

McGee, R. Jon, and Richard L. Warms 2000 *Anthropological Theory: An Introductory History*. 2nd ed. Mountain View, Calif.: Mayfield.

Moore, Jerry D. 1997 *Visions of Culture: An Introduction to Anthropological Theories and Theorists*. Walnut Creek, Calif.: AltaMira Press.

LIN POYER

**SOCIOLINGUISTICS**

When Brown and Gilman published their classic work on pronouns of power and solidarity (1960; see also 1989), no one characterized that paper as a major contribution to "sociolinguistics." When Gumperz and Hymes published their updated *Directions in Sociolinguistics* in 1986 (the 1972 edition was based on a 1966 publication of the American Anthropological Association), they were providing a paradigmatic definition of recognizable enterprise; that book included contributions by many of the founders. A two-part survey of sociolinguistics written in 1973 (Grimshaw 1973b, 1974a) noted that more had been published on sociolinguistic topics in the early 1970s than in all previous years. That review commented on about fifty new titles; only a few sociologists (particularly Basil Bernstein and Joshua Fishman, each with several volumes) were represented. In the three decades since that time, interest in language *in use* (micro sociolinguistics) has continued to grow exponentially; while that interest still is not seen as part of mainstream sociology, it is moving in that direction (Lemert 1979). Interest in more macro dimensions of the sociology of language—for instance, language conflict, language maintenance, and language spread and decline—also has grown, though much more slowly.

**SOME ACTIVITIES AND SOME LABELS**

At least a dozen specialties investigate some aspect of language: its origins, structure, invariant and variant features, acquisition, use in social contexts, change, spread, and death, and so on. Among those specialties, there are at least five whose practitioners do not consider themselves sociolinguists or sociologists of language and whose research seldom is incorporated directly into sociolinguistics/sociology of language (SL/SOL) investigations:

1. Formal linguistics that focuses on languages as autonomous systems and investigates how those systems work independently of human and/or social agency. This activity often is referred to as "autonomous linguistics" and occasionally as "nonhyphenated linguistics."
2. Anthropological linguistics devoted to a "description" (writing of grammars and

dictionaries and audio and phonemic recording of phonological systems) of languages in specific, usually nonmodern societies.

3. Psycholinguistics, which covers a wide range of topics, including the acoustics of perception, cognitive constraints on the complexity of clausal embedding, theories of innateness and learning in language acquisition, and the physical location of language functions in the brain.
4. Social psychology of language (from psychological social psychology), wide-ranging specialty that includes research on message characteristics and influence, self-disclosure, relationships between personality and speech, and relationships among body movements, speech, and "meaning."
5. Conversation analysis/ethnomethodology (CA), an approach that views talk in much the same way formal linguists view language: as a system that is syntactically organized and has structure that can be discerned independently of the social attributes of participants in particular talk.

CA has identified devices such as "preinvitations" and "preclosings" as well as ways of constructing accusations without accusing anyone explicitly (Atkinson and Drew 1979); workers in the field are interested in how these devices are used in the course of the immediate talk, not in how they might be directed to more complex goals of conversational participants. Whalen (1991) notes that CA "examines talk as an object in its own right, as a fundamental type of social action, rather than primarily as a resource for documenting other social processes." None of the five activities listed above deals with language primarily as social resource.

In contrast, another handful of specialties focuses on the social dimensions of language/talk as interactional resource, a component of individual and group identity, and a social object. The ethnography of speaking and ethnolinguistics, like the anthropological practices from which they take their names, focus on the diversity of available linguistic resources and the uses to which those resources are put in individual speech communities and in human society at large, respectively.

There is a strong comparative dimension to these arenas of investigation.

Sociolinguistics manifests a different kind of comparative orientation. The micro variety usually focuses on interactional accomplishment through the medium of language in use in social contexts: (1) comparisons of means and ends, including attention both to how individual ends can be accomplished by different means (ways of talking) and to how different outcomes may simultaneously result (intentionally or otherwise) from the production of same or very similar bits of talk, and (2) comparisons of the different resources available to different participants in talk. The sociology of language, as the macro variety of sociolinguistics often is called (Grimshaw 1987a), tends to focus on distributional studies, such as the distribution of language varieties across individual repertoires and the distribution of repertoires across social aggregates, categories, and groups (nations or classes, genders or age groups, and families or friendship networks, respectively). At the most macro level, this implies studies of language maintenance, supersession and change, conflict, and so on.

The sociological social psychology of language is oriented to group effects on individual behaviors, including the acquisition of social-cultural competence through the medium of talk, the role of talk in the acquisition and organization of evaluative orientations, and uses of talk/written language in social control. At some point, the last activity shades off into symbolic interactionism; this boundary is not explored here. Finally, specialized studies of proxemics (social and interpersonal spacing) and kinesics (body movement, the organization of facial features, gesture, posture) have been done from both sociological and psychological perspectives (Hall 1966, 1974; Kendon, [1977] 1990).

### SOME QUESTIONS OF ORIENTATION

Since later sections of this article illustrate how sociological theory can be enriched by empirical SL/SOL research in specific substantive areas, the comments here are limited to four questions of general orientation in theoretical work in SL/SOL: (1) What are causal and other relations between language/speech and other social behavior(s)? (2) Are grammars of social interaction pos-



sible, and is there a universal grammar? (3) What is the relevance of a micro–macro distinction for understanding the importance of language/speech in social life, and now are the two levels articulated in social behavior? (4) Is theoretical advance and/or understanding best sought by focusing on social processes or on specific substantive arenas of social behavior?

**Causal Directionality/Covariation/Cotemporality/Mutuality.** As in other varieties of social behavior, SL/SOL theory and research must deal with complex problems of cause and effect. There are four principal perspectives on the causal relationship between social structure and language (see Grimshaw 1974b; Hymes 1966):

1. That which sees language as fundamental (or as source, cause, independent variable, or set of independent variables), a position consonant both with an extreme Whorfian position (language *determines* how people think) and the commonsense observation that people sometimes do not know what is going on until they hear other people talking
2. That which sees social structure as a determinant or an independent variable or set of such variables, position consonant with people’s awareness that they talk differently in different situations, with different interlocutors, and depending on the nature of their interactional goals
3. That which sees neither as prior to the other, with both being seen as co-occurring and codetermining
4. That which sees both as being determined by a third factor, whether innate features of the human mind—the view of Cartesian linguistics (Chomsky 1966, 1968)—*Weltanschauung*, or the intrinsic demands of an ordered universe

Most SL/SOL correlational studies focus on how the location of individuals or groups in the social structure is *reflected* in speech and/or other language behavior, as in the case of regional or class dialects, or *determines* it, as in the case of the selection of a language variety in different situations and with different conversational partners (Blom and Gumperz 1972) or of pronominal forms or

other names (for a review of some of this literature, see Grimshaw 1980a). A smaller but substantial number of correlational studies attempt to discover how language use (spoken and written) is associated with interactional outcomes as varied as providing or not providing a requested favor, succeeding or not succeeding in school, and deciding whether to go to war (for a review, see Grimshaw 1981; for illustrations of claims about language use and the risks of war, see Chilton 1985; Wertsch and Mehan 1988). Although closer scrutiny often reveals that *ways of talking* are themselves resources that are differentially available to interactants with different social origins, some language resources appear to be available throughout social structures. Ways of talking in turn have been shown to have effects independent of structural relations.

Figure 1 is a simplified schematic representation of a mutual-embeddedness perspective. It is also a schematic showing how the processes of cultural reproduction would operate in a world without change. Bernstein (1975), Bourdieu (Bourdieu and Passeron 1977), Cicourel (1980a, 1980b, 1981), Collins (1981a, 1981b), and Habermas (1984–1987) all address the question of cultural reproduction and questions of change. All take essentially mutuality perspectives. All accord central importance to language in the reproduction process. Collins explicates ways in which language is simultaneously a resource in interaction and a source of change. Only Bernstein and Cicourel actually collect data on language in use, and only Cicourel directly investigates talk. None of these scholars would strongly disagree with this characterization; each would wish to “complete” the chart by incorporating neglected features (see Bernstein’s diagram of the process [1975, p. 24], with its foregrounding of different transmission agencies, such as family and education; modes of social control; specific speech varieties; and context-dependent and -independent meanings).

In the mid-1960s, Fischer (1965, 1966) published perhaps the strongest version of the mutual-embeddedness position and, from the disciplinary perspective of sociology, perhaps the most esoterically documented. (The papers are reviewed extensively in Grimshaw 1974b.) Fischer argued nothing less than phonological and syntactic differences between two related but mutually unintelligible languages (Trukese and Ponapean, separated for

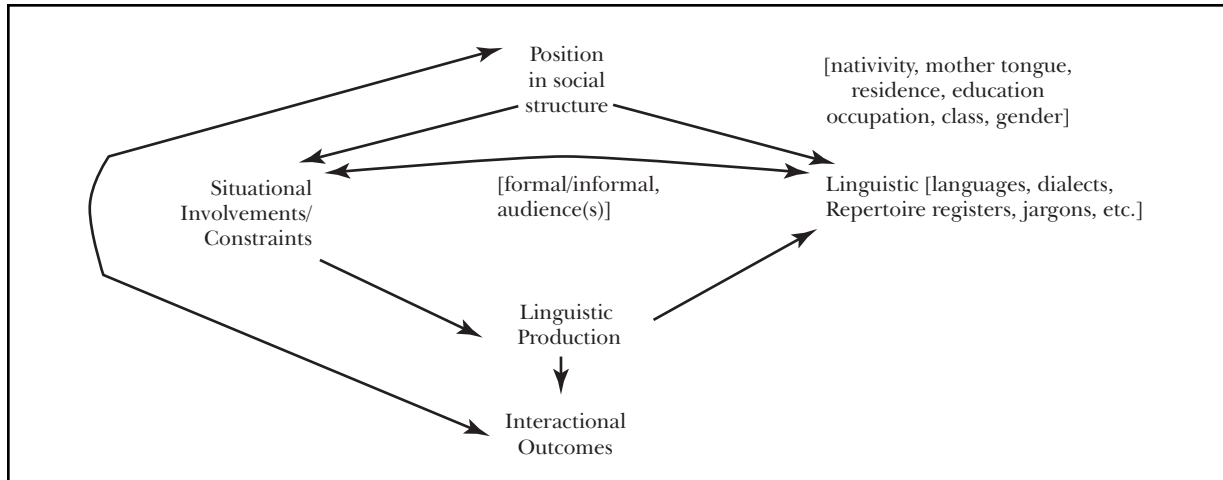


Figure 1. Sketch of a mutual-embeddedness perspective (cultural reproduction without change)

about eight centuries) are isomorphic to differences in the social structures of the two societies:

*As societies become more complex and social roles become more differentiated, the realized meaning of words in particular contexts becomes less important than the common or basic meaning. Speakers are forced to assume a greater cognitive gap between themselves and their listeners. At the same time, the basic meaning of the items of the lexicon tends to become more abstract and attenuated, since speakers have less need for words which can express much meaning in compact form to listeners who are conceived as being much like the self; they have more need, instead, for words which can be used in many different contexts with many different listeners who are conceived of as being very different from the self and from each other. (Fischer 1966, p. 178)*

The mutuality perspective is a richly suggestive one.

**Grammars of Social Interaction/A Grammar of Social Interaction.** Linguists write grammars; that is, they describe and write “rules” for phonological and syntactic systems for individual languages. They are also committed to the goal of writing a grammar of language, that is, identifying in grammars of individual languages features and/or rules that hold for all languages: a universal grammar. They distinguish between absolute and quantitative universals (i.e., between features

of all languages that can be explained on theoretical grounds as required constituents and features that occur in all or most languages, such as terms for female derived from that for male [this kind of feature is known as marking], but for which no theoretically principled basis can be identified) and between weaker and stronger claims of universality (i.e., between a claim that all languages contain certain elements [nouns, verbs, prepositions] and a claim that those elements appear in the same order in utterances in every language).

An interest in the intrinsic ordering of the universe and a concern to avoid repeating old errors and rediscovering the already known are central in linguists’ interest in both the regularities in individual languages and universal rules. Sociologists have similar concerns in seeking to discover the rules of interactional grammars for specific societies or groups and in seeking social interactional universals and the role of language in use in both grammars and the grammar. Although there are greetings in most, if not all, societies (this is more a quantitative than an absolute universal, and there are societies in which greeting is the marked case and nongreeting the unmarked), how they are done, to whom, and to what purpose may vary considerably (Firth 1972; Goffman 1971; Ibrahim et al. 1976; Kendon and Ferber 1973).

Similarly, there must be a need for information everywhere, but questions are not the appropriate manner for obtaining information in every

society (see Goody 1978; sources cited in Grimshaw 1969). Again, it seems likely that interpersonal relations of power and affect and considerations of valence and cost are everywhere involved in requesting behavior (Brown and Levinson 1987; Grimshaw 1989); their relative importance and the consequent variety of modes of requesting behaviors vary considerably.

Ways of talking are everywhere critical resources in interaction; very little is known, however, about what features of *language in use in social contexts* may be universal or, for that matter, about which rules *within* speech communities (or social groups) are variant and which are invariant (Labov 1968; Grimshaw 1973a). Indeed, some sociologists find the notion of rule misleading on grounds that expectations and behaviors are always under negotiation (Berger and Luckmann 1966). Notions of rules and exceptions vary across disciplines (Edgerton 1985; Labov 1968; Grimshaw 1973a, 1981).

**Micro–Macro, Conversation, and Interaction/ Official Languages and Language Policies, SL/ SOL.** These distinctions, along with the familiar polarities of social psychology and social organization—or qualitative and quantitative methods—often appear in discussions of sociological interest in language and language in use. Three sorts of questions can be asked in this arena:

1. What are sociology’s interests in what goes on in conversation/interaction, and what does a specific focus on talk teach sociologists that other modes of study do not?
2. What are sociology’s interests in looking at language as an individual social attribute that, aggregated, has supraindividual importance in ways similar to ethnic group, class, religion, and other categorical attributes? What are sociology’s interests in questions about how language is linked to life chances, why and how it becomes a focus of positive and negative attitudes, and how languages spread, change, contract, and die?
3. How do the things that go on in individual conversations on the micro level get articulated with, and aggregated into, processes of change in languages themselves, in their prestige, in policies re-

garding their use, and so on, on the macro level?

As was suggested above, the micro–macro question is closely related both to those about mutual embeddedness and to those about cultural reproduction. The dimension added by asking the articulation question is that of social change: If socializing/cultural transmission agencies operate to *reproduce* values, attitudes, behaviors, and so on, in new generations, how does change occur? A perspective offered by Collins (1981a, 1981b) is that participants bring to *everyday* conversations interactional resources that are enhanced or reduced in the course of interaction and that modest changes in interactional resources ultimately eventuate in changes in institutions and cultural systems—and languages. Related formulations are cited in Grimshaw (1987b). The macro–micro questions constrain one to think deeply both about processes of change and about how people try to get co-conversationalists to agree with them or to do what they want those them to do.

**Substance or Process: “Top-Down” or “Bottom-Up.”** Two additional questions about the construction and use of theory have methodological as well as theory-building implications. The first question is whether when sociologists study the uses of language in specific contexts such as educational, military, or medical institutions, they are interested primarily in understanding (1) the institutions themselves, (2) social processes such as negotiation or socialization or, more broadly, conflict or cooperation, (3) a specific kind of situated interaction, such as an interview, as a representative of a species of situation, or (4) how talk works in interaction. There are, of course, no pure cases. The second question has been put by Cicourel (1980a) as a distinction between “top-down” and “bottom-up” theorizing. By “top-down,” Cicourel means approaching corpora of talk with sets of conceptual notions ranging from the generality of “cultural reproduction” or “role” and/or “conflict” to the specificity of “role conflict” or different “footings” in talk (Goffman 1981). By “bottom-up,” he refers to researchers immersing themselves in their data and identifying regularities, then validating that identification, then discovering regularities in relations between previously observed regularities, and so on. All the investigators whose work is mentioned in this

article—indeed, all sociologists—would like to believe that they let their data guide them to theory construction, and all are to some extent guided in their work by prior theoretical constructions.

### DATA, DATA, EVERYWHERE

While sociologists sometimes are intimidated by the complex structure of formal linguistic theory, they may be equally envious of the easy access of linguists to their data, either in their own intuitions about the languages they speak or in the bath of talk and writing in which all people live (compare the ways of studying phonology in, for example, Chomsky and Halle 1968; W. Labov 1980). Students of SL/SOL share the advantage that many of the data in which they are interested are fairly accessible and, with modern technology, fairly easy to collect (denial of access and questions of ethics aside). They share the disadvantage that many of the sociological questions they want to study—matters as varied as (1) attitudes about different speech varieties, (2) the impact of stratification on the acquisition of those attitudes, and (3) the ways in which phonological variation *affects* stratification—can be considerably more difficult to identify, conceptualize, and measure. Just as there are fundamental questions about theoretical orientation in SL/SOL, there are fundamental questions about methods.

This brief discussion can comment on only a few of these methodological questions: (1) What constitutes optimal data for SL and SOL, or micro and macro, research? (2) How can the optimal data best be collected? (3) What need is there for modes of work, such as comprehensive discourse analysis (CDA), that differ from more familiar modes of sociological investigation? (4) What are the roles of collaborative and comparative studies in SL/SOL research?

**What Constitutes Optimal Types of Data for SL/SOL?** Labov (1972a) remarked that linguists work variously in library, bush, closet, laboratory, and street, where they collect and/or produce data that can be labeled texts, elicitation, intuitions, experimental results, and observations, respectively. Among the many sorts of data that can be useful in the investigation of SL/SOL issues, those associated with the following four “how” questions are central:

1. *How* do people actually talk/write? This has two dimensions: (a) What varieties of language (spoken and written are assumed in the following discussion) do individual members of speech communities control? (b) How do individuals employ their language resources in social interaction? The optimal data for such studies are extended texts.
2. *How* are language varieties and patterns of use distributed across categories of age, class, gender, occupation, nationality, religious affiliation, and residence? The optimal data here are a combination of sampled texts and observations.
3. *How* do members of social groups learn about language and its appropriate use, and how do they learn second (and higher-order) languages? The optimal data here are experimental results and observations and, to a lesser extent, texts.
4. *How* do people feel about language; that is, what are the attitudes of individuals and groups toward language varieties, repertoires, language change, and literacy? Data that have been employed in addressing these questions have included all five of the varieties listed by Labov (1972a), and each has proved useful.

### What Are the Criteria for Optimal Data?

There is no such thing as a “verbatim” record without electronic recording, and optimal records of conversation include both high-fidelity audio recording and possibly multiple sound-image recordings (for discussions of sound-image recording, including some of the controversies about such data collection, see Feld and Williams 1975; Grimshaw 1982, 1989). When one is working with written texts, optimal data include photographic copies of handwritten originals as well as printed versions. Whatever texts and observations are collected and used as data, however, those materials are valuable only to the extent that contexts of both “situation” and “text” (i.e., embedding talk and written material) are provided (the distinction is Halliday’s following Malinowski). Two excellent articulations of the importance of context that suggest different boundaries for what must be taken into account are those of Corsaro (1981, 1985) and Cicourel (esp. 1994).

People are often skeptical of claims about what talk is actually like until they see it transcribed; they then are skeptical that the transcription is accurate until they hear electronically recorded audio while reading a transcript. Investigators who work with texts, elicitation, and observations must always take into account the effects of monitoring in most varieties of SL/SOL data; Labov (1972a) has referred to the Observer's Paradox; that is, "we want to observe how people talk when they are not being observed." One also should keep in mind, however, Sol Worth's observation that all behavior, however carefully monitored, is "natural" (personal communication). Labov has developed elicitation techniques that have the advantage of generating different levels of self-consciousness of, and thus monitoring of, talk (see, e.g., 1972b).

Other concerns with data in SL/SOL research are, like those just reviewed, quite similar to those in sociological research in general. Self-report data on language varieties employed by oneself or one's family or of uses of literacy are notoriously unreliable, and definitions and measurements of individual attributes such as literacy and bilingual fluency are often inconsistent.

**The Need for Methods Specific to SL/SOL.** Many of the data employed in SL/SOL research are the same as or very similar to those employed in other arenas of sociology, as are the methods employed in analysis. This similarity may be least evident in the case of the activities labeled "conversation analysis" (see Whalen 1991) and "comprehensive discourse analysis" (Labov and Fanshel 1977). Labov and Fanshel realize that the goal of comprehensiveness is chimerical: their pioneering study demonstrated the importance of such aspects of talk as prosodic and paralinguistic features. Lexical, syntactic, and even phonological selection are deeply involved in what is "actually said" (i.e., interactionally intended) in talk (for a discussion of the process of "disambiguation" of text, see Grimshaw 1987c). Sociologists have employed CDA and adaptations of it to ask more specifically sociological questions. Other students have developed similarly fine-grained approaches to written texts (Silverman and Torode 1980). Perhaps sociologists are now aware that questions about language as language have sociological significance and that talk and writing are no longer just media that contain answers to other questions.

**Collaborative, Comparative, and Corroborative Research on SL/SOL.** While SL and SOL research and publications have increased tremendously in the last few decades, their literatures continue to be diffuse. There have been few replications. Most research has been on English, and much of the material on other languages is published in English. While much of the early activity in SL was interdisciplinary, there have been few truly interdisciplinary studies (for a discussion of problems with such projects, see Grimshaw, Feld et al. 1994) or parallel studies of shared data (see, however, Chafe 1980; Dorval 1990; Grimshaw, Burke, et al. 1994). There have been few explicitly comparative studies in which the same or collaborating investigators have simultaneously studied the "same" phenomenon in different speech communities (see, however, Watson-Gegeo and White 1990) or in different institutional contexts in same societies (see, however, Grimshaw 1990). There is reason to believe that all three of these important kinds of research are on the increase. Scholars all over the world are trying out SL formulations largely generated in the United States and Europe in their own societies, more and more work is being done on related SL phenomena in societies and speech communities where earlier work followed more traditional courses in linguistics and anthropological linguistics, and researchers from an increasingly wide range of disciplinary backgrounds are finding in SL/SOL data and theory materials to use in addressing their own questions.

**An Additional, Residual, Neglected Question.** Claims about the relative validity, reliability, and general worth of quantitative and qualitative research appear in SL/SOL, as they do in most areas of sociological work. While the modes of work are loosely associated with the micro-macro distinction, there are representations of both modes in both arenas.

#### SL/SOL AS A RESOURCE IN SOCIOLOGICAL THEORY BUILDING

A diminishing number of sociologists remain unfamiliar with SL/SOL and therefore unaware of both substantive findings and theoretical developments that could be helpful in their work. It is possible here to mention only a few contributions to the understanding of (1) substantive areas, (2) social processes, and (3) relations among language,

literature, identity, and so on. Let us begin with two instances of sociologically relevant contributions by the linguist best known to sociologists (possibly excepting Chomsky), William Labov, and then turn to research by sociologists and other nonlinguists that focuses from the outset on identifiably sociological concerns.

**Suggestive Empirical Findings.** Studies of classrooms, courtrooms, and clinics have generated findings that sometimes have resulted in changes in pedagogic, legal, and medical practice as well as contributed to our theoretical understanding of SL/SOL. Studies of a multitude of other settings, ranging from street, to dinner table, to backyard party, to workplace, also have produced important theoretical insights. Two of these findings come from Labov's quantitative studies of language use in urban areas of the eastern United States (New York City and Philadelphia); the first has important implications for understanding social stratification, the second for understanding ethnic (and possibly class) relations, and both for understanding processes of social change.

In studies of dialects associated with class, Labov and others have shown that linguistically insecure informants, more often than not women with aspirations for upward social mobility (or concerns about slipping), often hypercorrect their phonological production in the direction of what they perceive as prestige variants, thus producing the prestige variant with higher frequencies than do those at social levels above them—and sometimes produce it inappropriately (Labov gives as an example, “Hi, say, that’s hawfully good of you”). Awareness of this phenomenon should alert sociologists to look for analogues in other behavioral arenas (there is a family relationship to anticipatory socialization; the roots of the labeled behaviors may differ quite considerably); Labov (1972b, 1986) has pointed out important implications of this research for studies of linguistic (and, one may add, social) change.

Labov's second finding is that the speech of urban blacks who have contact with whites continues to be modified in the direction of the grammar of the dominant group, while blacks in the increasingly segregated inner cities speak ever more divergent language varieties. This divergence co-occurs with concomitant differentiation in incomes and educational achievement, heightening

social distance and probably enhancing intergroup hostility (see Williams et al. 1964). Labov comments, “The linguistic situation correlates with the formation of what has been called a ‘permanent underclass’” (1986, p. 278).

**Some Little Concepts and Some Central Processes.** Studies of actual talk that occurs in the course of everyday interaction have generated both (1) concepts that allow one to taxonomically identify previously unspecified regularities in that interaction (in a manner similar to Goffman's labeling of, among other things, front and back regions, side involvements, and the more specifically talk-related “footing”) and (2) new understandings of the working of what might be called “master social processes,” such as conflict and socialization. Instances of the former (all from Grimshaw 1989) are the identification of hyperinvolvement (a phenomenon in which interactants are so deeply involved in the ongoing that they miss the things they intend to monitor), defects of nerve (a situation in which interactants know how to do something but are reluctant to do it because they are concerned that it may generate injury to self or another party), and phenomena such as topic avoidance, topic exploitation, and topic truncation (the last occurring when it becomes obvious to an interactant that interactional goals are not going to be accomplished).

An instance of this is Corsaro's (1985) demonstration of the processes involved in children's learning how to recognize and construct the cultures and social structures in which they find themselves. Another is Grimshaw's (1990) distillation of the reported findings of a number of individual studies of conflict talk into propositions about the conflict process (see below).

*Discourse in organizations:* What can be learned from a language-oriented approach to social structure and social behavior depends on researchers' choices of an “entering wedge” (the term is John Useem's), that is, the selection of units of analysis, particulars of language in use to be attended, and questions to be asked. The paragraphs below will address things that have been learned from studies of language in use in (1) public bureaucracies in Sweden, focusing on *narratives* in the public domain, (2) organizations (primarily in the United States), employing a *conversation analytic* method, (3) a specific event within a university (a disserta-

tion defense) examined by an interdisciplinary group employing a range of approaches and asking a number of different questions, and (4) a sampling of more narrowly focused studies of language in use.

*Focus on narrative:* Czarniawska is a student of management and organization who disclaims linguistic or sociolinguistic competences language in use is nonetheless central to her analyses of public administration in Sweden. In *Narrating the Organization* (1997), Czarniawska demonstrates for the study of organizations, a central descriptive and analytic role of stories/narratives and a dramaturgical perspective. Czarniawska asserts that stories rule people's lives and constitute the basis for the construction of society (p. 5); she seems to believe that the centrality of narrative as the main source of knowledge "in the practice of organizing . . . is not likely to generate much opposition" (pp. 5–6). I find this surprising if true. Her project is to apply the narrative perspective (for which she credits Jean-Francois Lyotard) to elucidate continuities and change in Swedish public institutions and/or organizations at various governmental levels. The argument is dense, and the examples unfamiliar; the demonstration is persuasive both for the organizations studied and for the application of this perspective to other arenas of social life.

Czarniawska conceptualizes organizational life as stories, and organizational theories as ways of reading stories (pp. 26–29). She invokes Burke and Goffman in identifying drama and autobiography as special kinds of narratives (p. 32), noting (following Merelman) how drama, for example, can simultaneously or alternatively generate catharsis, personification, identification, and suspense (p. 36). She observes that stories not only can be vehicles for identity claiming by their tellers but also can be contested by other stories, unsuccessfully performed, and turned into serials (and sagas); importantly, all descriptions favor the theories of their tellers (p. 71).

In her exposition of how she studied Swedish public organizations, Czarniawska reviews a number of methodological conundrums related to the logic of inquiry of interpretive studies, including (1) *interruption* of texts (p. 92, per Silverman and Torode 1980), (2) the advantages of "outsiderness" (p. 62), (3) issues of case studies versus "window" studies (pp. 64 ff), (4) the identification of action

"nets" in organizational "fields" (p. 66), and (5) the place of notions of institutionalization and norms in studies focusing on narratives (pp. 68 ff.).

Focusing on two specialized fields—municipal administration and social insurance—she shows how *stories*, *themes*, and *serials* can be employed to elucidate the role of "good" and "bad" friction in social change, how new and old ways of acting have been integrated, and how new processes of "companyization" and "computerization" change the workplaces of individuals as well as the larger bureaucratic landscape. Narratives are central to people's lives, and Czarniawska has shown the necessity of attending to them.

**Conversation(al) analysis as a sociological method:** CA is treated elsewhere in this encyclopedia; conversational analysis (CAs) do not consider themselves sociolinguists or sociologists of language. For many years, the author has told CAs that their work is highly original, exciting, and of great potential value to sociology and that that potential will be achieved only when they integrate CA methods and concepts into more traditional sociology, simultaneously showing how traditional sociological concepts and perspectives ranging from status and role, to social structure, to socioeconomic status (SES), to self-esteem could help in interpreting CA findings.

Increasing numbers of researchers across the social sciences and humanities have come to value CA as an approach to everyday talk; until recently, CA-trained sociologists did not undertake to *demonstrate* the value of talk as data for studying fundamental sociological questions such as how social organization is constituted, reproduced, and modified—and how members contribute to that constitution, reproduction, and modification through talk—in what may appear to be mundane and unremarkable interactions. (CA methods are increasingly being employed in sociological analyses. Atkinson and Drew [1979] on Court proceedings, Maynard [1984] on plea bargaining, and Goodwin [1990] on black children's play groups are impressive examples. These studies do not as directly as the work discussed here foreground the epistemological issues implied by the CA posture described above [see, for example, Boden pp. 214–215].)

Boden (1994) provides a demonstration that many readers will find convincing. Using audio-recorded talk from telephone calls and meetings

of varying levels of formality collected in organizations ranging from a travel agency and a local television station to hospitals and a university administrative department to the Oval Office, Boden shares her understanding of the sometimes extraordinarily delicate but analytically identifiable ways in which talk is employed to “inform, amuse, update, gossip, review, reassess, reason, instruct, revise, argue, debate, contest, and actually *constitute* the moments, myths and, through time, the very *structuring* of [the] organization” (p. 8). The argument is often dense, and for readers unfamiliar with CA, it may be difficult to follow. The way in which an accountant comes to see that physicians in different departments may differently view policy change that could improve a hospital’s overall revenue position but reduce “their” money (p. 58 ff) is a case in point.

Boden’s goal is to use her collected talk to undertake two quite different but complementary projects:

*(1) to examine a range of more specific aspects of the organization of talk in the work settings that make up the business day; and (2) to discover through those materials how an apparently fragmentary process of information gathering, transmission, and very local assimilation is transformed into the goals, agendas, and decisions of organizations (p. 107).*

In the course of pursuing her projects, Boden shows how members of organizations can at the same time account for their behaviors in terms of a “rational actor” model and be unaware of how actual decision making is accomplished incrementally in fragments of unremembered and individually unremarkable chat rather than through a focused weighing of “rational” considerations. Boden simultaneously shows how concurrent and articulated employ of the previously segregated conceptual apparatuses of general sociology and CA (e.g., adjacency organization, agenda, bracketing, placement, sequence [centrally and critically], turn and so on) is mutually enhancing.

Boden argues that stages of (1) collection of actual talk, (2) identification of sequentiality in that talk, and (3) discovery in the talk and its sequentiality of the fundamental stuff and fundamentals of organization, (4) allow and/or contribute to sociological theory at levels of considerable abstractness (p. 206 ff). One may find in Boden’s

study a convincing *demonstration* of Collins’s “microfoundations of macrosociology” perspective.

Boden goes a step further and anticipates the result in offices of the future of today’s incremental changes (p. 209 ff); they will be places with more talk (and reduced opportunities for reflection), sped-up expectations of productivity, increased scope of action for all personnel, a flattened hierarchy, and real downsizing. Careful study of Boden’s book allows one to see how (1) she came to make such a projection and (2) one can evaluate a novel and productive application of CA to central sociological issues.

**Collaborative work across disciplines:** As early as the 1970s, concern was expressed about the increasingly disparate and noncumulative character of work on language in use in social contexts (Grimshaw 1973b, 1974a). In response to this concern, the Committee on Sociolinguistics of the Social Science Research Council initiated in the 1970s the Multiple Analysis Project (MAP), in which representatives of different disciplines and different theoretical and methodological approaches agreed to undertake analyses of a shared corpus of data, in this case a ten-minute fragment of a doctoral dissertation defense (Grimshaw 1989; Grimshaw, Burke, et al. 1994). The author has (Grimshaw 1994a) attempted to assess the success of this project in achieving its goals of theoretical cumulation, testing the comparative strengths and weaknesses of analytic approaches and methods, and an illumination of a shared corpus of data beyond that available from one or two studies by individual researchers. While it is not possible to conclude that one or another of the analytic modes employed is more comprehensive, a reading of the studies together conveys a sense of the complexity of language use in talk in its several contexts beyond the richness of most studies done from single analytic perspectives (but not all; see Labov and Fanshel [1977], who synthesized a variety of perspectives in their pioneering study of Rhoda and her therapist).

The anthropologists, linguists, and sociologists who investigated the dissertation segment variously focused on laminations of context (Cicourel 1994), formulaic talk (Wong Fillmore 1994), clause structure (Halliday 1994), humor (Fillmore 1994), [Who didn’t find funny that which sociologists in the defense thought was.], speech acts, cohesive



devices (Burke 1994; Hasan 1994), prosodic features and contextualization cues (Cook-Gumperz and Gumperz 1994), proforms (Grimshaw 1994b), and many other elements of the talk to study an equally wide range of behavioral outcomes. The outcomes included the reproduction and/or reshaping of social structure; control of the course of talk (and, by extension, interaction); creation of social relations; social and sociolinguistic constraints on discourse and interaction; negotiation of meaning; the notion of fluency, humor, laughter, and short-term social reorganization; membership-affiliation-identification processes and the structuring of group boundaries; and “face work” and political maneuvering. Each of these varieties of behavioral outcome is autonomously a sociological matter. In more macro directions, they are also deeply implicated in the search for answers to the following five interrelated questions; the answers to which require attending to particularities of language in use:

1. How is social structure generated, sustained, reproduced, and changed?
2. What is the role of social interaction/talk in the creation, realization, maintenance, and so on, of social structure?
3. How does interaction *work*? What permits it to continue and even flourish in social environments of competitiveness and aggressiveness?
4. What is the role of “meaning” in driving interaction and shaping social relations? How is “meaning” socially constrained and sociolinguistically/linguistically signaled?
5. What resources—social, sociolinguistic, and linguistic—are available to interactants for signaling meaning, sustaining interaction, and creating social relations?

What is going on, in short, when someone orders another person rather than asking or cajoling, or when someone says, “Don’t include *me* in that you!” or when an interactant says, “Along those lines . . .” and then challenges something said by an interlocutor/conversational partner?

**Other convergences:** Ochs, Schegloff, et al. (1996), another volume of studies of language in use, provides additional evidence of the convergence of CA (and linguistic) concerns with those of

both sociolinguists and more traditional sociologists. The book is innovative in that the collaboration it reports is among pioneers in articulating work in formal syntax, CA, paralinguistics (e.g., voice quality, intonation, and tempo), kinesics (e.g., facial expression, gesture, and “body language”), and proxemics (interpersonal distancing and arrangements of “things” in space) in investigating interactional accomplishment. It illustrates some of the range of autonomously syntactic (e.g., clausal organization and reorganization, employ of inflection and particles, movement of nouns or verbs), prosodic and paralinguistic, pragmatic, and other devices available in different languages and cultures (English, Finnish, Japanese, Kaluli) for the management of turns (including shifts of rights to the floor, acceptance or rejection of interruptions, and simultaneous speech), making credible claims, and making cautious characterizations of those not present or accomplishing a joint response to disruption. Fox et al. (1996) claim that English clause beginnings are richer than Japanese with information about “how clause is likely to continue,” thus providing potential interrupters with valuable information. They also observe that repair (by the speaker or another person) extends the possibilities for how an utterance can be completed in *any language* (p. 220). They are talking about *syntactic resources for interaction*. Some of these resources and/or devices are both familiar and fun in a Goffmanesque manner, as when the reader is introduced to and immediately resonates to notions such as “trailoffs” and “rush-throughs” (Schegloff) and “outlastings” (Lerner 1996).

The book is informative because in addition to demonstrating the interinfluence of syntax and interaction, it provides windows to sociological understandings revealed when seldom exploited perspectives are applied to concerns at the core of sociology, including issues of symbolic interaction, the socialization of neophytes, and social change. Beyond this, it allows similarly (and seminal) novel views of specialties such as the sociology of science and that of occupations. Particularly instructive, for example, are Ochs, et al. (1996), on how scientists construct indeterminate referential identities—sometimes in the process blurring the distinction between themselves and the physical world under their scrutiny—and how meaning is built through routine interpretive activity involving talk, gesture, and graphic representation. Also

engrossing is Goodwin's (1996) rich integration of different channels of behavior (syntactic production, intonation, body movement, display of awareness of a world beyond the immediate and ongoing) in a dynamic reconceptualization of Goffman's notion of participation framework as it operates in situations of disruption (of initially unknown magnitude) of routines—in an airport control tower.

Sociologists can profit from this book because of sociologically relevant questions posed, answers given, and because of passing observations. Lerner's (1996) thoughtful and suggestive work on how and why interlocutors complete utterances of speakers without being asked initially suggests as functions of such completion (1) agreement, (2) preemption of disagreement, (3) collaboration, and (4) heckling (p. 244). Lerner later argues that an early opportunistic completion may be intended to initiate or sustain a special alignment with a speaker such as affiliation (pp. 263–264). Sorjonen (1996) suggests that repeats and the Finnish particles *niin* and *joo* variously function as (1) interrogatives, (2) exclamations, (3) requests for confirmation (p. 279 ff.), (4) challenges, and (5) expressions of ritualized disbelief. There are obvious parallels to English. She also has some observations on sweetening recommendations of others when the recommender suspects that an unwanted invitation or request may be forthcoming. Schieffelin (1996, p. 442 ff.) shows how the invention and introduction of an evidential construction to refer to printed religious material, translatable as “known from this source/not known before,” not only has granted authority to written text when there is no basis in fact for doing so but also has been associated with the introduction of higher status for a new role of interpreter of Christianity in a society where prior stratification rested on different bases (and, not incidentally, also to a lowering of the status of women in a previously more egalitarian society).

**Social conflict as process: conflict talk as language in use in social context:** Early studies of intragroup conflict were largely experimental (often involving researcher-instigated disputes in dyads), usually nonattentive to particulars of subjects' talk, and, in part because of these two features, likely to overestimate the proportion of disputes that are in some way “resolved” (Corsaro and Rizzo 1990; Goodwin 1996). These writers and others have suggested that many pioneer stu-

dents of conflict talk (and social conflict more generally; for an early modern commentary, see Bernard 1950) were concerned with the disruptive consequences of disputation and thus tended to underestimate more positively valued outcomes, such as the creation of social organization and socialization of conflict participants (long ago identified by Simmel and others). In recent years researchers have, with great profit, turned increasingly to texts of actual disputes.

Students of processes of social conflict have more frequently than most other social scientists sought to formalize the regularities they have discovered in this phenomenon in propositions (see Coser 1956; Mack and Snyder 1957; Williams 1947; Williams et al. 1964). Taking into account the interaction of the sociological variables (affect, power, valence) and considerations of continua such as intensity, hostility, and violence and matters of external threat and internal cohesion as manifested in a range of studies of conflict talk, the author has formulated preliminary propositions. Space limitations constrain discussion of various sorts of propositions or of how sets of propositions permit the forecasting of patterns of conflict talk. Consider, however, the following:

Many disputes include instances of assignment of blame or responsibility (see Fillmore 1971). A discourse rule for this behavior might look like the following:

1. *Rule for assigning blame (responsibility).* If A asserts that B should and could have performed a behavior  $X_1$  but willfully did not or that B should and could have avoided performing a behavior  $X_2$  but nonetheless wilfully performed it, A is heard as blaming B for the nonoccurrence or occurrence of  $X_1$  or  $X_2$ , respectively (Grimshaw, 1992 p. 312).

The influence of power on the availability of aggressive, uncompromising, and sometimes hostile modes of talk in conflict is similar to its influence and constraint on other selection of ways of talking:

2. *Ceteris paribus*, selection of more “confrontational” modes of conflict talk (e.g., threats or insults and increased amplitude or physiological rage displays, threatening kinesic posture, or gestures)

is directly related to increasing relative power (Grimshaw 1992, p. 315).

It is interesting to note that in talk where the parties are proxy military representatives of superpowers discussing matters of very high valence, confrontational modes generally are avoided (Grimshaw 1992). A growing literature on conflict talk demonstrates that the behavior of interest is simultaneously immensely complicated and has a rich potential for new understandings both of social conflict itself and of discourse more generally. This is true even though records of critically important conflict-talk events on the group and international levels have not been available for study (Grimshaw 1992).

#### **Language, Writing, Literacy, and Literature.**

Recent events have demonstrated the continuing importance of what Geertz (1963) labeled "primordial sentiments" and shown that feelings about language are central among those sentiments. The range of sociological and sociologically relevant ways in which both language in general and writing and literacy in particular permeate and/or pervade human cultural and social structures and relations, as well as conceptions of identity and of self on the individual level, defies easy summary or description. People go to court in defense of their mother tongue; people have also fought in the streets and burned themselves alive over language issues. Becoming literate in any language can be primarily an instrumental acquisition; in some instances, it can have profound effects on both individual personalities and social organization (see, particularly, J. Goody 1987). The "invention" and development of national languages can have reverberating effects through previously atomized collectivities (Anderson 1990); when printed material becomes available, it can have critical impacts both on change in general (Eisenstein 1979) and on the development of national communities and identities (Anderson 1983).

Sociologists of literature have shown how national literatures can reveal cultural and social values (e.g., Moore 1971); sociologists who study both contemporary life and that in past times are becoming increasingly aware of the rich data in personal documents from journals to correspondence. It is even possible to hear the question, "Who wants citizens to be literate, and to what

ends?" (the implication is that social control may be as much a goal as is the enrichment of individual lives [see Kress and Hodge 1979]). Related interests have drawn a number of investigators to study of *how* written materials affect their readers, a question that has been addressed both by methods that project the "interruption" or "interrogation" of written and/or spoken texts (Silverman and Torode 1980) and by those of psycholinguistics or cognitive science.

#### **Language and personal and social identity:**

This introduction to matters of language in use in social contexts should not be closed without mention of a dimension of social life increasingly recognized by sociological social psychologists as well as sociolinguists. This section includes brief reviews of two studies that focus on this *use* of language in identity matters and closes with a listing of suggestive but previously unexamined questions.

*Constitution of morally relevant categories of people:* T. Labov (1980) has been concerned with specifying how ascriptions of morality are made in conversational discourse and about whom (i.e., which persons and collections of persons) they are made. She has concluded that a task prior to the location of evaluation and obligation in talk is the specification of how people are located in talk and how morally relevant categories of people are constituted.

According to Labov (1980), all types of "collections of people" (a term "used to designate any plurality of people which can be referred to in talk or systematically inferred from the talk") are potentially relevant in moral matters; since there is a potentially infinite number of such collections, it is imperative to develop procedures for reducing the number to be examined in any given investigation, that is, to discover principled bases for classifying collections. She does this by first developing discovery procedures for locating collections and then gathering the collections into sets bounded by common identifying dimensions.

Labov (1980) observes that references in talk to collections of people often occur in the form of common nouns, proper names, and pronouns. What people are often unaware of, she argues, are collections of "hidden people," that is, "those collections of people not immediately evident in

the surface talk, but systematically retrievable.” Collections are hidden in five ways:

1. In references to social organization(s), incidents, and specific categories (e.g., in the case of academic settings, belonging-to-department-people, participating-in-defense-people, and faculty people).
2. Social characterizations, that is, collections defined by verbs of activity or specific attributes (e.g., candidate-attacking-people or candidate-defending-people or identification by gender, academic rank, features of personal appearance, voice quality, etc.) She (p. 135) notes that some activities, such as guessing, telling, and thinking, since they are features of all people, are nondiscriminating and thus analytically without value.
3. Ellipsed individuals or collections (i.e., where there is shared knowledge of past characterizations).
4. Collections concealed in references to time or place.
5. Plurals hidden in singulars.

Collections can conveniently be labeled as “feature plus people,” for instance, “doing research on language and identity people” or “reading encyclopedia articles people,” to make “*explicit in a standardized way what features of the people are being considered*” (emphasis added). It is precisely such explicitness that is needed for the specification of identities and their boundaries.

Labov realizes that the identification of collections used by individual speakers is not sufficient to permit an understanding of how discussion of moral matters is accomplished in talk, and she continues by raising several critical questions. The general question is, “How do analysts (and interlocutors) know that coconversationalists are talking about a ‘same’ collection?” There is no easy answer to this question; collections to which reference is being made change in the course of a single speaker’s utterance, overlap across utterances, shift across utterances because the original speaker’s identification was unclear or because a hearer-become-speaker misheard or deliberately misunderstood (Grimshaw 1980b) the original identification, are layered, subsumed, expanded (Labov’s

notion of layering is loosely akin to both the linguistic and Goffman’s [1981] more specialized uses of the term “embedding”), and so on. Resolution of these complexities is a requirement for coherent and cohesive discourse. Labov proposes the “notion of ‘category consensus’ for a situation where the relevance of a given collection of people is shown interactional support” and “which occurs as co-interactants ratify the use of specific collections of people.” Category consensus is not always achieved; like other varieties of consensus, it is often the subject of challenge, negotiation, and metadiscussion. This is, of course, what the study of identity(ies) is about.

*Exploitation of referential ambiguity in pronominal usage:* A complication is introduced by the use of definite or indefinite articles such as “reading-the-Borgatta-encyclopedia people” versus “reading-an-encyclopedia people.” This introduces the possibility of the use of referential ambiguity in language as an interactional resource. The question of how collectivities (both categories and groups) are constituted and bounded and how that boundedness may be explicitly or implicitly signaled in spoken or written discourse provides a venue for additional demonstration of the value for sociology of examination of language in use. Personal and other pronouns are a useful resource in boundary work; their referential ambiguity also provides a resource exploitable for including and excluding both those present and those absent from relevant social collectivities (Grimshaw 1994a).

The fact that there are times when hearers or readers don’t know to what person or set of persons reference is being made can be weighted with social implication when it is not clear, for example, *who* is being scolded or praised, positively or negatively or neutrally characterized, or invited or rejected. Uncertainty can persist even in the presence of apparently disambiguating specifications such as “you all,” “all of you,” “the *n* of us” (when the collection address includes *n*-plus persons), and “the four of them.” Hearer-readers ordinarily are able to make inferences that are correct or close enough that they can sustain conversation (or reading) without continuously finding it necessary to stop to resolve ambiguities. It is also true, of course, that ambiguities may go unrecognized, be recognized but not resolved, or even be intentionally exploited. Unresolved ambiguities can be in-

consequential, but they may occasionally have delayed consequences of considerable importance (e.g., when “uninvited” guests turn up or unintended “insults” are repaid with interest).

Most readers will be far more familiar with the language–identity connection in which people participate whenever they talk with others. Except under extreme circumstances (ongoing or pending disaster) or service situations in which interactants are treated more as part of the scene than as other humans, the first thing people do when a partner in interaction speaks (in person, on the telephone, or in writing) is to “place” that person in terms of background (in which one usually includes age, education, ethnicity and national origin, gender, occupation, regional provenance, social class, and, depending on situation, other achieved and ascribed attributes). This sensitivity to the link between how people speak and who they are is further demonstrated by the ways in which the speech of others is imitated in “poshing up” (Goffman 1979) and in the production, with varying degrees of friendliness, condescension, and hostility, of “mock” Spanish or black English or other real or imagined languages (Hill).

Note, for example, the insertion of foreign words and phrases (insertion of not-currently-in-use-code speech: (1) foreign words and phrases, e.g., *āp kē bād*, *buenos dias*, *je ne sais pas*, *obrigado*, *paz*, *СПАСІБО*, *was gibts*) (2) technical terms and phrases, such as “deep structure,” “diglossia,” “dope,” “gigabyte,” “identity,” “S and M,” “sole-noid,” and (3) phonological variants and regional dialect lexical items or, more comprehensively, code *switches* in which a different language, dialect, or register is employed for an extensive stretch of talk. How are such insertions and switches to be interpreted? Readers will be able to construct scenarios in which the following are or are intended to be conveyed: “I am one of you,” “I am not one of you, but I am attuned and sympathetic to you,” “I and those of my auditors who understand what I have just said are different (superior to?) from those who did not understand,” “I and those who understand my metaphorical use of a variant are different from (superior to) those who processes it nonmetaphorically” (e.g., “humorous” employ of socially disvalued variants).

Questions of language in use and matters of identity (and thus of stratification, life chances,

social conflict, and so on) are inextricably interrelated and intertwined; neither can be fully comprehended without attention to the other.

### APPLIED SOCIOLINGUISTICS, SOCIAL AMELIORATION, AND THEORY BUILDING

The increased interest in SL/SOL has been accompanied by and contributed to by growing public exposure to and interest in language as a social problem. Consider in the last decade of the twentieth century in the United States alone issues of free speech (“hate crime” versus political correctness, arrests for public “cursing”) in both public discourse and on private computers, the “English as official language” movement and accompanying disputes about “rights” to non-English ballots or other government documents, and the public hue and cry about “Ebonics.” Public concern about propriety in language use is not a new phenomenon (see Kamensky 1997).

Work on “real” problems in a variety of institutional areas has benefited from a growing body of theory, to which it has in turn contributed. Again a distinction can be made between micro and macro concerns. Micro sociolinguistic research has been done on how communication fails in classrooms, courtrooms, and clinics; macro studies have examined how the speaking of socially disvalued language varieties is associated with educational failure, differential treatment in the judicial system, and unsuccessful interaction with medical services delivery systems. Ameliorative programs have ranged from bilingualism in education to the English as official language movement, from the provision of interpreters in the courtroom to attempts to simplify legal language, and from attempts to teach prospective doctors to become better interviewers and listeners to trying to get doctors to use less technical language. Bitter controversies have raged over how the ways children talk are related to educational success and failure; the Ebonics dispute is one instance among many (see Labov 1982). Some investigators have argued that some language varieties are not suited for abstract, critical, logical, and propositional thought; others, that the success and failure of persons who speak in different ways are determined by the political preferences about language varieties of gatekeepers such as teachers and employers.

Recent years have seen the development of the role of the “language scientist” as an expert witness (e.g., Rieber and Stewart 1990); SL considerations are sometimes deeply involved in such testimony. Many of these programs and much of this work initially grew out of concerns with language varieties associated with ethnicity (in the United States, different varieties of Spanish and, particularly, Black Vernacular English (BVE)). There has also been more explicit attention paid to problems of communication across classes, age groups, and, particularly, gender; Tannen’s (1990) book on gender differences in talk spent many months on best-seller lists.

On a more explicitly macro level, language planning and language policy have become more visible arenas of government activity in both rich countries, which must deal with visiting or immigrant workers who speak unfamiliar languages, and poor countries, which must make decisions about which competing languages are going to receive official status and support or about which orthography to employ for previously unwritten languages. (The latter is a decision that is likely to have political as well as economic implications.) They must in some cases decide whether high literacy (often seen as an index of modernism) will ultimately contribute to their economies (or other values) as much as or more than would other investments (on outcomes of increases in literacy in industrial [izing] and less developed countries, respectively, see Graff 1979; Goody 1987). Both rich and poor countries must deal with native multilingualism; they have done it with varying success in Belgium, Canada, Finland, India, Indonesia, the former Soviet Union, Spain, Switzerland, and a number of countries in Africa (see McRae 1983, 1986, 1997, for excellent studies on Switzerland, Belgium, and Finland, respectively).

## REFERENCES

- Anderson, Benedict R. O’G 1983 *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- 1990 *Language and Power: Exploring Political Cultures in Indonesia*. Ithaca, N.Y.: Cornell University Press.
- Atkinson, J. Maxwell, and Paul Drew 1979 *Order in Court: The Organisation of Verbal Interaction in Judicial Settings*. Atlantic Highlands, N.J.: Humanities Press.
- Berger, Peter L., and Thomas Luckmann 1966 *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, N.Y.: Doubleday.
- Bernard, Jessie 1950 “Where is the Modern Sociology of Conflict?” *American Journal of Sociology* 56:11–16.
- Bernstein, Basil 1975 *Class, Codes, and Control*. vol. 3: *Towards a Theory of Educational Transmission*. London: Routledge and Kegan Paul.
- Blom, Jan-Petter, and John J. Gumperz 1972 “Social Meaning in Linguistic Structure(s): Code-Switching in Norway.” In J. J. Gumperz and D. Hymes, eds., *New Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston.
- Boden, Deirdre 1994 *The Business of Talk: Organizations in Action*. Cambridge, UK: Polity.
- Bourdieu, Pierre, and Jean-Claude Passeron 1977 *Reproduction in Education, Society, and Culture*, trans. Richard Nice. London: Sage.
- Brown, Penelope, and Steven C. Levinson 1978 “Universals in Language Usage: Politeness Phenomena.” In E. Goody, ed., *Questions and Politeness: Strategies in Social Interaction*. Cambridge, UK: Cambridge University Press.
- Brown, Roger, and Albert Gilman 1960 “The Pronouns of Solidarity.” In T. A. Sebeok, ed., *Style in Language*. Cambridge, Mass.: MIT Press.
- 1989 “Politeness Theory and Shakespeare’s Four Major Tragedies.” *Language in Society* 18:159–212.
- Burke, Peter J. 1994 “Segmentation and Control of a Dissertation defense.” In A.D. Grimshaw et al., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Chafe, Wallace, ed. 1980 *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, N.J.: Ablex.
- Chilton, Paul, ed. 1985 *Language and the Nuclear Arms Debate: Newspeak Today*. London: Frances Pinter.
- Chomsky, Noam 1966 *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper and Row.
- 1968 *Language and Mind*. New York: Harcourt, Brace and World.
- , and Morris Halle 1968 *The Sound Pattern of English*. New York: Harper and Row.
- Cicourel, Aaron V. 1980a “Three Models of Discourse Analysis: The Role of Social Structure.” *Discourse Processes* 2:101–131.
- 1980b “Language and Social Action: Philosophical and Empirical Issues.” *Sociological Inquiry* 40:1–30.

- 1981 “Notes on the Integration of Micro- and Macro-Levels of Analysis.” In K. Knorr-Cetina and A. V. Cicourel, eds., *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro-Sociologies*. Boston: Routledge and Kegan Paul.
- 1994 “Theoretical and Methodological Suggestions for Using Discourse to Recreate Aspects of Social Structure.” In A. D. Grimshaw et al., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Collins, Randall 1981a “On the Microfoundations of Macrosociology.” *American Journal of Sociology* 86:984–1014.
- 1981b “Micro-Translation as a Theory-Building Strategy.” In K. Knorr-Cetina and A. V. Cicourel, eds., *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro-Sociologies*. Boston: Routledge and Kegan Paul.
- Cook-Gumperz, Jenny, and John J. Gumperz 1994 “The Politics of a Conversation: Conversational Inference in Discussion” In A. D. Grimshaw, et al., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Corsaro, William A. 1981 “Communicative Processes in Studies of Social Organization: Sociological Approaches to Discourse Analysis.” *Text* 1:5–63.
- 1985 *Friendship and Peer Culture in the Early Years*. Norwood, N.J.: Ablex.
- , and Thomas A. Rizzo 1990 “Disputes in the Peer Culture of American and Italian Nursery School Children.” In A. D. Grimshaw, ed., *Conflict Talk: Sociolinguistic Investigations of Arguments in Conversations*. Cambridge, UK: Cambridge University Press.
- Coser, Lewis A 1956 *The Functions of Social Conflict*. Glencoe, Ill.: Free Press.
- Czarniawska, Barbara 1997 *Narrating the Organization: Dramas of Institutional Identity*. Chicago: University of Chicago Press.
- Dorval, Bruce, ed. 1990 *Conversational Organization and Its Development*. Norwood, N.J.: Ablex.
- Edgerton, Robert B. 1985 *Rules, Exceptions, and Social Order*. Berkeley: University of California Press.
- Eisenstein, Elizabeth L. 1979 *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early Modern Europe*. Cambridge, UK: Cambridge University Press.
- Feld, Steven, and Carroll Williams 1975 “Toward a Researchable Film Language.” *Studies in the Anthropology of Visual Communication* 2:25–32.
- Fillmore, Charles J. 1971 “Verbs of Judging: An Exercise in Semantic Description.” In C. J. Fillmore and D. T. Langendoen, *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston.
- 1994 “Humor in Academic Discourse.” In A. D. Grimshaw et al., eds., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Firth, Raymond 1972 “Verbal and Bodily Rituals of Greeting and Parting.” In J. S. La Fontaine, ed., *The Interpretation of Ritual: Essays in Honour of A. L. Richards*. London: Tavistock.
- Fischer, John L. 1965 “The Stylistic Significance of Consonantal Sandhi in Trukese and Ponapean.” *American Anthropologist* 67:1495–1502.
- 1966 “Syntax and Social Structure: Truk and Ponape.” In W. Bright, ed., *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964*. The Hague: Mouton.
- Fox, Barbara A., Makoto Hayashi, and Robert Jaspersen 1996 “Resources and Repair: A Cross-Linguistic Study of Syntax and Repair.” In E. Ochs, et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Geertz, Clifford 1963 “The Integrative Revolution: Primal Sentiments and Civil Politics in the New States.” In C. Geertz, ed., *Old Societies and New States: The Quest for Modernity in Asia and Africa*. New York: Free Press.
- Goffman, Erving 1971 *Relations in Public: Microstudies of the Public Order*. New York: Basic Books.
- 1981 *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goody, Esther N. 1978 “Towards a Theory of Questions.” In E. N. Goody, ed., *Questions and Politeness: Strategies in Social Interaction*. Cambridge, UK: Cambridge University Press.
- Goody, Jack 1987 *The Interface between the Oral and the Written*. Cambridge, UK: Cambridge University Press.
- Goodwin, Charles 1996 “Transparent Vision.” In E. Ochs et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Goodwin, Marjorie Harness 1990 *He-Said-She-Said: Talk as Social Organization among Black Children*. Bloomington: Indiana University Press.
- Graff, Harvey J. 1979 *The Literacy Myth: Literacy and Social Structure in the Nineteenth-Century City*. New York: Academic Press.
- Grimshaw, Allen D. 1969 “Language as Obstacle and as Data in Sociological Research.” *Items* 23:17–21.
- 1973a “Rules in Linguistic, Social and Sociolinguistic Systems and Possibilities for a Unified Theory.” In R. S. Shuy, ed., *Twenty-Third Annual Round Table*.

- Monograph Series on Language and Linguistics (1972). Washington, D.C.: Georgetown University Press.
- 1973b “On Language in Society: Part I.” *Contemporary Sociology* 2:575–585.
- 1974a “On Language in Society: Part II.” *Contemporary Sociology* 3:3–11.
- 1974b “Sociolinguistics.” In I. de Sola Pool and W. Schramm, eds., *Handbook of Communication*. Chicago: Rand McNally.
- 1980a “Social Interactional and Sociolinguistic Rules.” *Social Forces* 58:789–810.
- 1980b “Mishearing, Misunderstandings and Other Nonsuccesses in Talk: A Plea for Redress of Speaker-Oriented Bias.” *Sociological Inquiry* 40:31–74.
- 1981 “Talk and Social Control.” In M. Rosenberg et al., eds., *Sociological Perspectives on Social Psychology*. New York: Basic Books.
- , ed. 1982 *Sound-Image Records in Social Interaction Research*. Special issue of *Sociological Methods and Research* 11.
- 1987a “Sociolinguistics versus Sociology of Language: Tempest in a Teapot or Profound Academic Conundrum?” In U. Ammon, N. Dittmar, and K. J. Mattheier, eds., *Sociolinguistics: An International Handbook of the Science of Language and Society*, vol. 1. Berlin: Walter de Gruyter.
- 1987b “Micro-Macrolevels.” In U. Ammon, N. Dittmar, and K. J. Mattheier, eds., *Sociolinguistics: An International Handbook of the Science of Language and Society*, vol. 1. Berlin: Walter de Gruyter.
- 1987c “Disambiguating Discourse: Members’ Skill and Analysts’ Problem.” *Social Psychology Quarterly* 50:186–204.
- 1989 *Collegial Discourse: Professional Conversation among Peers*. Norwood, N.J.: Ablex.
- , ed. 1990 *Conflict Talk: Sociolinguistic Investigations of Arguments in Conversations*. Cambridge, UK: Cambridge University Press.
- 1992 “Research on the Discourse of International Negotiations: A Path to Understanding International Conflict Processes?” *Sociological Forum* 7:87–119.
- 1994a “Referential Ambiguity in Pronominal Inclusion: Social and Linguistic Boundary Marking” In A.D. Grimshaw et al., eds., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- 1994b “What We Have Learned: Some Research Conclusions and Some Conclusions about Research.” In A. Grimshaw et al., eds., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- , Peter J. Burke, Aaron V. Cicourel, Jenny Cook-Gumperz, Steven Feld, Charles J. Fillmore, Lily Wong Fillmore, John J. Gumperz, Michael A. K. Halliday, Ruqaiya Hasan, and David Jenness, eds., 1994 *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- , Steven Feld, and David Jenness 1994 “The MAP: An Ethnographic History of the Project and a Description of the Data.” In A. Grimshaw et al., eds., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Gumperz, John J., and Dell Hymes, eds. (1972 [1966]) 1986 *Directions in Sociolinguistics: The Ethnography of Communication*. Oxford, UK: Basil Blackwell.
- Habermas, Jürgen 1981–1984 *The Theory of Communicative Action*, trans. Thomas McCarthy. Boston: Beacon Press.
- Hall, Edward T. 1966 *The Hidden Dimension*. Garden City, N.Y.: Doubleday.
- 1974 *Handbook for Proxemic Research*. Washington, D.C.: Society for the Anthropology of Visual Communication.
- Halliday, Michael A. K. 1994 “So You Say ‘Pass’ . . . Thank You Three Muchly.” In A. D. Grimshaw et al., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Hasan, Ruqaiya 1994 “Situation and the Definition of Genres.” In A. D. Grimshaw et al., *What’s Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.
- Hill, Jane H. 1999 “Language, Race, and White Public Space.” *American Anthropologist* 100: 680–689.
- Hymes, Dell 1966 “Two Types of Linguistic Relativity (with Examples from Amerindian Ethnography).” In W. Bright, ed., *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964*. The Hague: Mouton.
- Ibrahim ag Youssouf, Allen D. Grimshaw, and Charles S. Bird 1976 “Greetings in the Desert.” *American Ethnologist* 3:797–824.
- Kamensky, Jane 1997 *Governing the Tongue: The Politics of Speech in Early New England*. New York: Oxford University Press.
- Kendon, Adam 1977 1990 *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge, UK: Cambridge University Press.
- , and Andrew Ferber 1973 “A Description of Some Human Greetings. In R. P. Michael and J. H. Crooks, eds., *Comparative Ecology and Behavior of Primates*. London: Academic Press.



- Kress, Gunther, and Robert Hodge 1979 *Language as Ideology*. London: Routledge and Kegan Paul.
- Labov, Teresa G. 1980 *The Communication of Morality: Cooperation and Commitment in a Food Cooperative*. Unpublished dissertation. New York: Columbia University.
- Labov, William 1968 "A Proposed Program for Research and Training in the Study of Language in Its Social and Cultural Settings." New York: Columbia University (mimeograph).
- 1972a "Some Principles of Linguistic Methodology." *Language in Society* 1:97-120.
- 1972b *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- , ed. 1980 *Locating Language in Time and Space*. New York: Academic Press.
- 1982 "Objectivity and Commitment in Linguistic Science: The case of the Black English Trial in Ann Arbor." *Language in Society* 11:165-201.
- 1986 "Language Structure and Social Structure." In S. Lindenberg, J. S. Coleman, and S. Nowak, eds., *Approaches to Social Theory*. New York: Russell Sage Foundation.
- , and David Fanshel 1977 *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.
- Lemert, Charles C. 1979 *Sociology and the Twilight of Man: Homocentrism and Discourse in Sociological Theory*. Carbondale: Southern Illinois University Press.
- Lerner, Gene H. 1996 "On the 'Semi-Permeable' Character of Grammatical Units in Conversation: Conditional Entry into the Turn Space of Another Speaker." In E. Ochs et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Mack, Raymond W., and Richard C. Snyder 1957 "The Analysis of Social Conflict—Toward an Overview and Synthesis." *Journal of Conflict Resolution* 1:212-248.
- Maynard, Douglas W. 1984. *Inside Plea Bargaining: The Language of Negotiation*. New York: Plenum.
- McRae, Kenneth D. 1983 *Conflict and Compromise in Multilingual Societies: Switzerland*. Waterloo, Ontario, Canada: Wilfred Laurier University Press.
- 1986 *Conflict and Compromise in Multilingual Societies: Belgium*. Waterloo, Ontario, Canada: Wilfred Laurier University Press.
- 1997 *Conflict and Compromise in Multilingual Societies: Finland*. Waterloo, Ontario, Canada: Wilfred Laurier University Press.
- Moore, T. Inglis 1971 *Social Patterns in Australian Literature*. Berkeley: University of California Press.
- Ochs, Elinor, Patrick Gonzales, and Sally Jacoby 1996 "When I Come Down I'm in the Domain State': Grammar and Graphic Representation in the Interpretive Activity of Physicists. In E. Ochs et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- , Emmanuel A. Schegloff, and Sandra A. Thompson, eds. 1996 *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Rieber, Robert W., and William A. Stewart, eds. 1990 *The Language Scientist as Expert in the Legal Setting: Issues in Forensic Linguistics*. Vol. 606 of *Annals of the New York Academy of Sciences*.
- Schegloff, Emanuel A. 1996 "Turn Organization: One Intersection of Grammar and Interaction." In E. Ochs, et al. eds., *Interaction and Grammar*. Cambridge, U.K.: Cambridge University Press.
- Schieffelin, Bambi B. 1996 "Creating Evidence: Making Sense of Written Words in Bosavi." In E. Ochs et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Silverman, David, and Brian Torode 1980 *The Material Word: Some Theories of Language and Its Limits*. London: Routledge and Kegan Paul.
- Sorjonen, Marja-Lenna 1996 "On Repeats and Responses in Finnish Conversations." In E. Ochs et al., eds., *Interaction and Grammar*. Cambridge, UK: Cambridge University Press.
- Tannen, Deborah 1990 *You Just Don't Understand: Men and Women in Conversation*. New York: Morrow.
- Watson-Gegeo, Karen A., and Geoffrey M. White, eds. 1990 *Disentangling: Conflict Discourse in Pacific Societies*. Stanford, Calif: Stanford University Press.
- Wertsch, James, and Hugh Mehan, eds. 1988 *Discourse of the Nuclear Arms Debate*. Special issue of *Multilingua: Journal of Cross-Cultural and Interlanguage Communication* 7.
- Whalen, Jack 1991 "Conversation Analysis." In E. Borgatta and M. Borgatta, eds., *Encyclopedia of Sociology*, vol. 4. New York: Macmillan.
- Williams, Robin M., Jr. 1947 *The Reduction of Intergroup Tensions*. New York: Social Science Research Council.
- , John P. Dean, and Edward A. Suchman 1964 *Strangers Next Door: Ethnic Relations in American Communities*. Englewood Cliffs, N.J.: Prentice-Hall.
- Wong Fillmore, Lily 1994 "The Role and Function of Formulaic Speech in Conversation." In A. D. Grimshaw et al., eds., *What's Going on Here? Complementary Studies of Professional Talk*. Norwood, N.J.: Ablex.

ALLEN D. GRIMSHAW

# Encyclopedia of Sociology

---

Second Edition

# Encyclopedia of Sociology

---

Second Edition

VOLUME 5

**Edgar F. Borgatta**

*Editor-in-Chief*

*University of Washington, Seattle*

**Rhonda J. V. Montgomery**

*Managing Editor*

*University of Kansas, Lawrence*

**Macmillan Reference USA**

*an imprint of the Gale Group*

*New York • Detroit • San Francisco • London • Boston • Woodbridge, CT*

## SOCIOLOGY AMONG THE SOCIAL SCIENCES

The relationship between sociology and the other social sciences is in reality a relationship between sectors of different disciplines, not between whole disciplines. Sociology is one of the most open disciplines toward other disciplines. This openness is manifested in the citation patterns in academic publications, which allows one to measure the degree of coherence of a discipline, the relationship between specialties within a discipline, and the interactions among disciplines. If specialists in a subdiscipline tend to cite mostly or exclusively specialists in the same subdiscipline and if relatively few authors cite outside their subdiscipline, as a whole the discipline has a low degree of internal coherence. In this case, the real loci of research are the specialties. If, by contrast, a significant proportion of authors cross the borders of their specialties, the discipline as a whole can be considered an integrated territory.

As can be seen in the analytic and alphabetical indices of most compendiums and textbooks, sociology has a weak core. The fragmentation of the discipline into isolated specialties can be seen in most sociological treatises: "We divide up the discipline into a number of topics, each the subject of a chapter. These chapters are minimally integrated" (Calhoun 1992, p. 185). Theoretical sociology is presented as a subfield disconnected from substantive domains: "General sociology has been relegated primarily to introductory textbooks and to a lesser extent to a sort of social theory that most practicing sociologists use but little in their work" (Calhoun 1992, p. 185). For instance, in Smelser (1988), the twenty-two chapters represent autonomous specialties that are only weakly related to each other. Few of the 3,200 authors cited in that work are mentioned in more than one specialty (Dogan 1997). This lack of a consensus among sociologists has been emphasized in a symposium devoted to that book (Calhoun and Land 1989).

In most general works of sociology published in the last two decades, the most frequently cited authors are ancestors, not contemporary sociologists. With exceptions such as Parsons, Merton, Lazarsfeld, and Mills, few mentors belong to the immediately previous generation. Nowadays, sociologists, in their pattern of references, are like children elevated by their grandparents. This cult

of the ancestors is surprising, because "following advances in the division of labor and specialization, the works of the classics ceased to be directly useful to an average sociologist. To do correct research in a specialized branch of sociology one does not in fact have to read the works—bulky, often abstruse, and semi-philosophical in nature—written by Marx and Spencer, Simmel and Weber, Mead and Znaniecki. To do such research it suffices to master, on the basis of a possibly recent handbook, the standard techniques and the current theories of the middle range" (Szacki 1982, p. 360).

The fragmentation of sociology can be explained in part by the absence of consensus on a dominant, integrative theory or widely accepted paradigm. If a consensus could be reached among sociologists, it would be that sociology has a small, soft, and old core, that sociology is not a centripetal discipline, and that it expands in all directions.

There is very little communication between the fifty specialized domains recognized by the International Sociological Association (ISA) and between the thirty sectors of the American Sociological Association. If cooperation among the specialized fields is weak or absent, a vivid traffic can be observed between each specialized sociological domain and across disciplinary borders: the specialized group of scholars belonging formally to other disciplines, particularly specialties rooted in social psychology, social demography, social anthropology, social history, social geography, social ecology, some branches of political science, political economy, and sociolinguistics. A double phenomenon appears in the sociological literature of the last two decades: a division of the discipline into noncommunicating specialties and an opening of the disciplinary frontiers to specialties from different disciplines.

Bridges are built over the disciplinary borders. The circulation on these bridges is almost as important as the circulation along the internal arteries of formal sociology. The importance of this "foreign" trade can be measured. In a study covering four decades from 1936 to 1975, it was found that sociologists cited articles in sociology journals 58 percent of the time; political scientists cited scholars from their own discipline only 41 percent of the time; anthropologists referred 51 percent of the time to their colleagues; psychologists referred

73 percent of the time to their own kin, and 79 percent of the economists did the same (Rigney and Barnes 1980, p. 117). These figures indicate that in each social science a significant proportion of theoretical, methodological, and substantive communication has been with other disciplines, with the most open discipline being sociology and the most autonomous being economics.

In an analysis of journals identified as belonging to sociology and economics, there was a significant shift from sociology to “interdisciplinary sociology” and from economics to “interdisciplinary economics” between 1972 and 1987. The criterion for interdisciplinarity was the proportion of cited references in the journals of the respective disciplines (Crane and Small 1992, p. 204–205). An analysis by those authors in terms of clusters of references shows a clear increase in interdisciplinary relationships.

In addition to this the crossing of disciplinary borders, another important trend in the last fifteen years has been the multiplication of new hybrid journals that cross disciplines and specialties. More than 300 hybrid journals in English that concern sociology directly or indirectly have been established in this period, along with many others in French and German. Most of these new journals have a limited circulation and are addressed to readers in highly specialized subfields.

#### FROM SPECIALIZATION TO FRAGMENTATION TO HYBRIDIZATION

The fragmentation of disciplines is generated by an inevitable and growing process of specialization. All the sciences experience such specialization. As a discipline grows, its practitioners generally become increasingly specialized and inevitably neglect other areas of the discipline. The division of physics into physics and astronomy and the division of chemistry into organic chemistry and physical chemistry are examples in the natural sciences. In the social sciences, what was originally the study of law divided into law and political science; anthropology split into physical anthropology and cultural anthropology; and psychology broke up into psychology, social psychology, psychotherapy, and psychiatry.

Each formal discipline gradually becomes too large and unmanageable for empirical research.

No theory or conceptual framework can encompass the entire territory of sociology. Talcott Parsons was the last one to attempt such a unification, but his ambition was unrealistic (Johnston 1997). Contemporary theories are influential only within their subdisciplines. The process of fragmentation and specialization eventually is followed by a process of recombination of the specialties into new hybrid domains. These recombinations correspond to the logic of multiple and concatenated causality in the social sciences.

The more renowned new hybrid domains hoist their own flags, for instance, political sociology, which is a fusion of sectors from both of its parent disciplines; social psychology, which is already autonomous; political economy, which detaches large sectors from economics and political science and smaller sectors from sociology; and historical sociology, which has revived on both sides of the Atlantic. None of these four subfields were mentioned three decades ago by Smelser (1967). This absence shows the changes that have occurred since then.

It is pointless to lament the fragmentation of sociology or any other social science, because the interaction between specialties in different disciplines is beneficial. All social sciences, sociology in particular, have grown in depth and breadth through exchanges with cognate specialties in other disciplines. What some scholars perceive as dispersion is in reality an expansion of knowledge and an inevitable trend.

In the history of social sciences, the progression from fragmentation to specialization to hybridization has taken one of the following six forms:

1. Division in two parts, or bifurcation. The history of the sciences is a long chain of divisions. One of the oldest, going back to Aristotle, is the separation of philosophy and political theory. One of the most recent is the divorce of cognitive science from traditional psychology.
2. Changing the boundaries of formal disciplines. The growth of specialties at the interstices between disciplines has as a consequence the shrinking of the borders of the parent disciplines. When social psychology became independent, psychol-

ogy lost an enormous territory. One of the borders of economics retracted when political economy was emancipated. Anthropology has seen its frontiers retract as a result of modernization, industrialization, and urbanization; consequently, urban studies expanded. The margins of political science are in perpetual change.

3. Migration of individual scholars from one formal discipline to another or to a new territory. The founders of sociology have moved away from philosophy, such as Durkheim; from history, such as Weber; or from economics, such as Pareto (see Dogan and Pahre 1990).
4. The convergence of two domains in a new hybrid field. One of the most recent examples in medical sciences is the intermingling of fragments of cardiology with fragments of pneumonology. The nomenclature of social sciences is full of such hybrid fields.
5. Outgrowth from the mother discipline for pragmatic reasons, to the point of joining another formal discipline. For instance, sociology of medicine, the most populated sociological subdiscipline, is today located more often in hospitals than in departments of sociology; it has become a problem-solving subdiscipline.
6. Borrowing from neighboring disciplines and exchanging concepts, theories, methods, practices, tools, and substance. This borrowing and lending process is an important route of hybridization. All the social sciences share concepts, theories, and methods. The contribution of sociology to this shared repository is impressive. Sociology has devised and exported many more concepts to neighboring disciplines than have borrowed from it (Dogan 1996). Most theories formulated in a discipline sooner or later spread to other disciplines. The diffusion of theories across disciplinary borders is one of the arguments that could be invoked by those who advocate more interdisciplinary strategies in the social sciences. The borrowing and lending of methods among disciplines have itineraries different from those for the

spread of concepts and theories. As contributions of methodology in social sciences, the most productive in strategies and techniques of research were until recently psychology, econometrics, social psychology, and statistics. For concepts and theories, the most creative disciplines are sociology, political science, economics, anthropology, and philosophy.

A distinction must be made between interdisciplinary amalgamation and hybridization through recombination of specialties belonging to different disciplines. A “unified sociology” existed only in the early phase of sociological development. Hybridization of specialties came later, after the maturation of the process of the internal fragmentation of disciplines. The word “interdisciplinary” is misleading when used to describe contemporary trends, because today only specialties overlap, not entire disciplines. The word “hybridization” may seem to be imported from biology, but it has been used by social scientists such as Piaget and Lazarsfeld.

Sociometric studies show that many specialists are more in touch with colleagues in other disciplines than with colleagues in their own disciplines. The “invisible college” described by Robert Merton, Diana Crane, and other sociologists of science is an eminent multispeciality institution because it ensures communication not only from one university to another and across all national borders but also between specialists attached administratively to different disciplines. The networks of cross-disciplinary influence are obliterating the old classification of the social sciences.

#### **RECOMBINATION OF SOCIOLOGICAL SPECIALTIES WITH SPECIALTIES IN OTHER SOCIAL SCIENCES**

Sociology has exchanged concepts, theories, methods, practices, and substance most intensively with three other disciplines: political science, history, and economics. The analysis here will focus first on these three disciplines. The well-known dominion of social psychology can be surveyed briefly. The relationships between sociology and social geography have long been difficult and poor. What happens in the absence of intermingling? Other specialties intervene in the empty space, as in case of ecological geography. I have to forgo the

overlapping areas between sociology and social anthropology, social demography, ethnology and sociolinguistics, but the comments on the process of fragmentation of disciplines, on multiplication of specialties, and recombination of the specialties in new hybrid fields are also applicable to them.

**Relations with Political Science.** A double phenomenon can be observed in the relationship between sociology and political science. First, there is weak communication within each of these two disciplines among the multiple specialized fields, an impermeability between the specialized research subfields that belong formally to the same discipline. The disciplines appear like watertight compartments in large ships (Dogan 1997). Typically, there is relatively little scholarly exchange between a student of the American Congress and a specialist in Middle Eastern politics, between a political philosopher and an expert in statistical analysis, and between an Africanist and an expert on welfare states. However most of these scholars are likely to have relationships with cognate specialties in neighboring disciplines. The diversity of methodological schools contributes to the fragmentation of each discipline.

Second, across disciplinary frontiers there is a vivid traffic between special fields or subfields belonging to one discipline and similar or cognate fields in the other discipline. A convincing way to show the importance of these cross-disciplinary bridges is to rank on two columns the fifty research committees of the ISA and the forty committees of the International Political Science Association (IPSA). For each area of research in one discipline there is a homologue in the other discipline: religion, ethnicity, generations, gender, mass communication, elites, socialization, crime, social inequality and so on. To these interminglings should be added theoretical and methodological pairs: All major schools and sects are represented in both disciplines from rationalists to Marxists and from qualitative methods to proponents of quantification.

The relationships between sociology and political science can be observed by counting the proportion of authors belonging to a discipline who cite articles from other disciplines. Such an analysis of footnotes in major journals shows the trade across disciplinary frontiers and the changes in trade routes over time. In terms of import-export balance, political science has borrowed

from economics, sociology, and social psychology, and has exported mostly to sociology.

There has been a change in the cross-fertilization of political science. In the 1950s and 1960s, sociology was the major lender to political science, making important contributions such as group theory, political socialization, social cleavages, and systems theory. In the 1970s and 1980s, economics was the major cross-fertilizer of political science, especially with theories of public goods and collective action, game theory, social choice, and international trade theory. Psychology has been a constant exporter to political science and sociology, but at a lower level. In the 1960s, its major contributions came from personality theory and the study of values.

One domain of sociology—political sociology—and one domain of political science—comparative politics—have privileged relations, in some cases achieving a real fusion. In the history of comparative research, there was a privileged moment of cooperation and convergence between political sociology and comparative politics in the 1960s. Between 1958 and 1972, three dozen important books and articles were published that shared three characteristics: comparison by quantification, hybridization, and cumulative knowledge. That combination had never previously been achieved in the history of sociology and political science (Dogan 1994, p. 39). This privileged moment also marks a break with European classical comparisons in the sociological style of Tocqueville, Mill, Marx, Weber, and Pareto.

The alarm over the parochial state of comparative politics after the subjugation of all social sciences during the period of totalitarianism in Europe (Scheuch 1992) and before their renaissance in the United States was raised by Roy Macridis in 1955. At the same time (1954), the Statistical Bureau of the United Nations started to publish “social statistics” on demographic variables, income, standards of living, social mobility, sanitary conditions, nutrition, housing, education, work, and criminality. These sources facilitated the encounter between political sociology and comparative politics.

In 1957, *Reports on the World Social Situation* began to be published by the Department of Economic and Social Affairs of the United Nations. The chapters in these publications in 1961 and

1963 on “the interrelations of social and economic development and the problem of balance” and on “social-economic patterns” are contributions that can be read profitably today by sociologists interested on developmental theories. Lipset’s *Political Man* (1959) borrowed from all the social sciences. A year later, Deutsch produced his “manifesto” (Deutsch 1960), followed by a seminal article (Deutsch 1962). Both articles dealt with comparative indicators. The following year an important article by Cutright (1963) was published that appears in retrospect to have been prophetic. In the same year, Arthur Banks and Robert Textor published *A Cross-Polity Survey*, in which the majority of the fifty-seven variables are of direct interest to sociologists. Shortly afterward, the *World Handbook of Political and Social Indicators* by B. Russett et al. discussed seventy-five variables, the majority with sociological significance. In *Comparative Politics* by G. Almond and G. Bingham Powell (1966), several social sciences, particularly sociology and social anthropology, are seen in the background. From that moment on, the field of international comparisons became bifurcated, with both trends being related to political sociology. One road continues with quantitative research, in which contributors constantly use nonpolitical factors in their analyses of the correlates of democracy and transition to democracy. An important contribution comes again from the Development Program of the United Nations, the *Human Development Report* (1990 and after). In this publication, gross national product (GNP) per capita is replaced by a new indicator: purchasing power parity (PPP).

The other road gave priority to sectoral comparisons, for instance, the eight volumes on development published by the Princeton University Press, where politics is most of the time a dependent variable explained by social economic and cultural factors. About a thousand books and articles appear in a selected bibliography of sectoral comparisons published during the last three decades. About half of their authors belong administratively to political science, a quarter belong to sociology, and a quarter are hybrids scholars.

Comparative political sociology does not consist only in cross national analysis. It is also a cross-disciplinary endeavor, because in comparative research one is crossing units (nations) and variables (numerical or nominal). The variables are usually more numerous than the units. The relations be-

tween variables are often more important for theoretical explanations than are discoveries of analogies and differences between nations. In comparative political sociology, there is not a single major book that attempts to explain politics strictly by reference to political variables. Of course, the amount of hybridization varies with the subject and the ability of the author to omit what should be implicitly admitted.

More than 200 contemporary European and American scholars have held a joint appointment in the departments of sociology and political science or have moved from one to the other. Some comparativists cannot be locked in only one of these two disciplines.

**Historical Sociology and Social History.** History is the most heterogeneous discipline in the social sciences, dispersed in time and space. It is divided into a nomothetic part and an ideographic part. The dispute over the role and borders of history, which in France goes back to Durkheim, Simiand (1903), and Seignobos, does not seem to have ended. Three generations later, history has been excluded from the social sciences under the authority of an international institution: It is not numbered among the nomothetic sciences covered in UNESCO’s *Main Trends in the Social and Human Sciences*. Historians do not appear to have reacted vigorously to this affront. Indeed some have come to terms with it: “The progress of history in the last fifty years is the result of a series of marriages: with economics, then with demography, even with geography . . . with ethnology, sociology and psychoanalysis. When all is said and done, the new history sees itself as something like an auxiliary science of the other social sciences” (Chaunu 1979, p. 5). This is clearly not the opinion of the other French historians (*Annales* 1989, p. 1323), who are resolutely committed to interdisciplinarity: “History will progress only in the context of interdisciplinarity.”

As long as the focus is on the long time span and the comparative approach, there is agreement between Durkheim and Braudel. At a distance of sixty years, using different words, they say much the same thing: “History can be a science only in so far as it compares, and there can be no explanation without comparison . . . Once it starts comparing, history becomes indistinct from sociology” (Durkheim in the first issue of *L’Année Sociologique*



1898). Braudel is just as accommodating: "Where the long time span is concerned, the point is not simply that history and sociology tie in with each other and support each other but rather that they merge into one" (Braudel 1962, p. 93). However, this refers only to the part of history that compares while considering the long time span; other fields of history have very little to do with sociology. Similarly, many sociologists do not need to have recourse to history to resolve the problems with which they are concerned. Durkheim and Braudel would have been more explicit if instead of considering their disciplines as a whole, they had referred clearly to their common territory, which is now called comparative social history or historical sociology. Once it is accepted that history and sociology overlap only in certain areas, the long territorial dispute between history and sociology will become a thing of the past. However, this is only one sector of history brought face to face with a sector of another discipline. Exchanges with economics have thus generated economic history, which is of interest only to enough historians and economists, to provide material for several major journals. Each human activity has its historian, who, in order to perform his or her task, has to hunt in other people's lands.

On the other side of the Atlantic, as soon as their disciplines had begun to fragment, innovative historians and sociologists reached out to one another. Frederick Jackson Turner's study of the American frontier was a marriage of sociology and history with the benediction of geography. Later, sociologists such as Bellah (*Tokugawa Religion*, 1959) and Lipset (*The First New Nation*, 1963) were joined by a new generation of historians, represented by Charles Tilly's *The Vendee* (1964), Barrington Moore's *Social Origins of Dictatorship and Democracy* (1966). This interweaving of sociology and history continues to the present day.

"Most sociologists and historians have no clear understanding of what historical sociology really is" (Aronson 1969, p. 294). Unlike in economics, political science, or linguistics, the distinction is not based on subject matter. Many have attempted to clarify the differences between the two disciplines, leaving no two authors in agreement (see Boudon 1979; Lipset and Hofstadter 1968; Tilly 1981). The reason for the lack of consensus is clear: The remarkable diversity of the historical sociologies, to say nothing that of their parent

disciplines, makes any unidimensional characterization of the issue unsatisfactory.

The comparative method is a very useful way to unify general statements of causality of historical events. One of the first to take this path was the French school in the journal *Annales*, which developed an approach to social history that was both sociological and comparative. Marc Bloch was one of the most influential figures in the development of this school both in his programmatic statement *Pour une histoire comparée des sociétés européennes* (1928) and in *La société féodale* (1939–1940). For some historians it is impossible to assess the validity of any causal interpretation on the basis of a single case, making a comparative approach absolutely necessary for useful explanation (Cahnman and Boskoff 1964, p. 7). Comparative history overcomes the fragmentation of specialized (and especially national) history. Examining similar causal processes in two or more specific contexts can illuminate the nature of the causal forces at work and improve one's understanding of the events being studied.

The dialogue between the specific and the general is an important issue explored by many who discuss historical sociology. Along these lines, Burke (1980) isolates two different aspects of the contributions history can make to sociology, one negative and one positive. The negative contribution entails picking away at the edifice constructed by others by showing how a theory does not fit one's society. This entails tests that are hazardous for any theory, but the theories that survive are proved to be of greater value. The positive contribution involves working out from the general to the particular in order to construct a revised general theory. This task is especially valuable because a sociologist's generalizations often appear vacuous to a historian. Historians have invalidated many of the theories of sociologists and political scientists.

When posed in this fashion, the social sciences' insistence on generalization can be helpful for historians. In the words of a sociologist turned historian, "Whatever else they do, the social sciences serve as a giant warehouse of causal theories and concepts involving causal analogies; the problem is to pick one's way through the junk to the solid merchandise" (Tilly 1981, p. 12). When one

finds solid ground, a simple application of sociological theory to historical problems can be innovative.

Knapp (1984) suggests that historians can help overcome the inattention to context of most social theory. He argues that one of the major problems in sociological theory is the implicit or explicit *ceteris paribus* (all else being equal) clause. Since all other factors are never the same in the real world, such theories are repeatedly disconfirmed and often appear vacuous: "When sociologists (or political scientists, economists, or anthropologists) decide that concern with theory absolves them from concern with history, their product will not only be irrelevant historically, it will not even be adequate as theory" (Knapp 1984, p. 34). When theories are opened up to allow variation in the *ceteris paribus* clause, they can be applied to specific historical contexts. Historians who are most familiar with the peculiarities of "their" period or country have much to add to social theory in this type of research.

Contrary to what is generally believed, historical sociology sometimes is not based on quantified research. Nonetheless, quantification is so ubiquitous in most social sciences that it is easy for historians to misunderstand the nature of the field. As Tilly points out, "In field after field, the leading edge of the change was some form of quantification. Because of that uniformity, many nonquantitative historians mistook the prow for the whole ship" (Tilly 1981, p. 34). Quantified data are for most sociologists what primary sources are for historians: Some historians cannot resist quoting diaries, and some sociologists cannot resist quantifying. Both kinds of evidence have advantages and disadvantages, and each discipline can gain from making greater use of the kind of evidence most useful to the other.

In addition to a difference in method, history and sociology often are distinguished by their conceptual inventories. There are a number of sociological concepts historians can use to their advantage, such as structure, function, social role, kinship, socialization, deviance, social class and stratification, social mobility, modernization, patrons and clients, and factions. The breadth of this list makes it clear that there is much room for hybridization of subfields across the disciplinary boundaries. For instance the concept of "development" is central in several social sciences (Riggs 1984).

**Relations with Eclectic Economics.** To discuss the relationship between sociology and economics, it is necessary to distinguish several varieties of economists: econometricians, monodisciplinary monetarist theorists, landless theorists, and eclectic transgressors of borders (a fifth variety, economic historians, has been expelled from the field). The first two varieties have well-known physiognomies. Landless theorists (Rose 1991) are economists who believe that they do not have to deal with nation-states and tend to reduce all countries to a single model. They travel at the level of landless economies. One may assume that the first three varieties are outstanding contributors to scientific knowledge, since so many of them have been awarded Nobel prizes, but here only the last variety has good relations with the other social sciences.

Eclectic economists denounce the reductionism advocated by other economists. Four decades ago, Hayek wrote that "nobody can be a great economist—and I am even tempted to add that the economist who is only an economist is likely to become a nuisance if not a positive danger" (Hayek 1956, p. 463). For the Nobel prize laureate Buchanan, "it becomes increasingly clear that the channels of effective communication do not extend throughout the discipline that we variously call 'economics' and that some 'economists' are able to communicate far more effectively with some scholars in the noneconomic disciplines than with those presumably within their own professional category" (Buchanan 1966, p. 181). Another Nobel prize laureate asked: "Why should economics be interdisciplinary? The answer is, presumably, because otherwise it will make mistakes; the neglect of all but the narrowly economic interactions will lead to false conclusions that could be avoided" (Solow 1970, p. 101). Many economists state that "it is necessary to reduce the use of the clause *ceteris paribus*, to adopt an interdisciplinary approach, that is to say to open economics to multidimensionality" (Bartoli 1991, p. 490).

Economics is also divided, but to a lesser degree than the other social sciences. It has maintained some coherence but has had to pay a high price for this by considerably reducing its field. At one time, economics reached a fork in the path: It could have chosen intellectual expansion and the penetration of other disciplines at the cost of heterogeneity and diversification and at the risk of

dispersal (a risk taken by sociology and by political science); it chose instead to remain true to itself, thereby forfeiting vast territories. Many economists consider that the choice of purity, methodological rigor, and hermetic terminology was the right choice.

It is thus clear that self-sufficiency eventually leads to a shrinking of borders, but this does not mean general impoverishment, since the lands abandoned by economists were soon cultivated by others. Those lands now have their own departments, research centers, and professional schools (management, political economy, development science). The position of economics in the constellation of the social sciences today might have been more dominant if so many economists had not withdrawn into monodisciplinary.

This situation is surprising in that “few classical sociologists have failed to assign a central place in their theories to the relationship between economy and society: from Marx and Weber to Schumpeter, Polanyi, Parsons and Smelser” (Martinelli and Smelser 1990).

If many economists have locked themselves in an ivory tower and allowed whole areas to escape from their scrutiny, other economists have advocated an “imperialistic expansion of economics into the traditional domains of sociology, political science, anthropology, law and social biology” (Hirschleifer 1985, p. 53; Radnitzky and Bernholz 1986). Several of these economists are famous scholars, including several Nobel laureates. A kind of manifesto has been published in *The American Economic Review*:

*It is ultimately impossible to carve off a distinct territory for economics, bordering upon but separated from other social disciplines. Economics interpenetrates them all, and is reciprocally penetrated by them. There is only one social science. What gives economics its imperialist invasive power is that our analytical categories are truly universal in applicability. Thus economics really does constitute the universal grammar of social science. But there is a flip side to this. While scientific work in anthropology and sociology and political science and the like will become increasingly indistinguishable from economics, economists will reciprocally have to become aware of their functions. Ultimately, good economics will also*

*have to be good anthropology and sociology and political science and psychology. (Hirschleifer 1985, p. 53)*

This view is anachronistic, but many outstanding economists have succeeded not only in exporting their knowledge to other disciplines but also in invading them with their methods and theories and achieving innovative research. Arrow’s *Social Choice and Individual Values* (1951) led mathematically trained economists to apply game theory to a variety of social conflict situations. Several works made such applications, including Anthony Down’s *An Economic Theory of Democracy* (1957), Duncan Black’s *The Theory of Committees and Elections* (1958), Buchanan and Tullock’s *The Calculus of Consent* (1962), Riker’s *The Theory of Political Coalitions* (1962), and Olson’s *The Logic of Collective Action* (1965). Since then, many social scientists have borrowed ideas and techniques from economists and applied them to the analysis of various processes and situations. The economists were the first in the field because they had a longer tradition of mathematical training and used more abstract and thus more widely applicable concepts. The other social sciences had learned statistics in order to handle the interpretation of their empirical data but were much slower to learn advanced mathematics. In a number of important graduate schools, economists hold joint appointments with other social science departments.

Some economists continue to spread the application of their analytic techniques to outside fields. Becker wrote a book on discrimination and prejudice and in *Treatise on the Family* (1981), applied economic analysis to topics such as the incidence of marriage, divorce, and childbearing. He was awarded the Nobel Prize in Economic Sciences in 1992 for his work applying economics to different areas of human behavior, particularly the family, a traditional stronghold of demography. Gordon Tullock’s *The Economics of Non-Human Societies* deals with ants, termites, bees, mole rats, sponges, and slime molds. Many similar examples could be given (Szenberg 1992): “The fields to which the economic approach or perspective has been applied over the last thirty or forty years include politics, sociology, ethnology, law, biology, psychology” (Radnitzky and Bernholz 1986). An examination of recent issues of journals of economic literature shows that some economists explore a wide range of issues. Among these

eclectic economists a few who work in another discipline and then return immediately to the home discipline. Intriligator (1991) has presented in a schematic way the patterns of cross-fertilization among the behavioral sciences by identifying concepts and theories developed in economics and adopted by others. He traces in terms of input-output the itinerary of social choice theory, structural models, decision theory, organization theory, bounded rationality, utility theory, game theory, the concept of balance of power, and anomie.

The interactions between economics and political science are deeper than those between economics and sociology. Many economists are better known in political science than in economics, particularly in the domain of political economy. In *A New Handbook of Political Science* (Goodin and Klingemann 1996), the new economic sociology receives great attention, but it is not clear how it is different from the older political economy. This work should be confronted with *Handbook of Economic Sociology* (Smelser and Swedberg 1994). For instance, Offe describes the "asymmetry" between the two disciplines: "Political economists do have an economic theory of institutions and tend to disregard this demarcation line separating spheres. Sociologists have perhaps only the rudiments of a sociological theory of what is going on in markets and firms, while the most ambitious argument that sociologists do have to offer effectively demonstrates that non-economic spheres of society are not only constituted in different ways than the economy, but that the economy itself depends on non-economic spheres" (Offe 1992, p. 687).

**Social Psychology.** Most sociologists are not involved in the kind of research that interests most psychologists, and vice versa. For the majority of sociologists and the majority of psychologists, their respective territories are clearly separated. Nevertheless, between the two disciplines there is a condominium, social psychology, inhabited by hybrid scholars, some of whom have begun their scientific activity in one of the two disciplines while others started as "hybrids." In addition, for many sociologists who are not social psychologists, psychology is the nearest and most important disciplinary neighbor. What Inkeles wrote three decades ago is still valid: "It would not be at all difficult to assemble a set of fifty or one hundred recent articles in social psychology, chosen half from the psychological and half from the sociological jour-

nals, which would be so much alike that no one, judging without knowledge of source or author, could with any precision discriminate those written by professional sociologists from those written by psychologists. Several considerations follow from this simple fact. Clearly, the two disciplines cannot be defined in terms of what psychologists and sociologists respectively do, since they so often do the same thing" (Inkeles 1970, p. 404).

The growth of social psychology during the last two generations makes Durkheim's arguments in favor of the supremacy of sociology over psychology irrelevant, along with the old debate about the individual-society dichotomy: "The claim to a principled distinction of sociology from psychology based on the distinction of individual from society is challenged by the substantial attention that at least some sociologists pay to individuals, by difficulties in describing psychology as the study of individuals, and by difficulties in the very conceptual distinction of individual from society" (Calhoun 1992, p. 175). At the early stages of the discipline's postwar history, psychology had been the most cited cognate discipline by sociologists, but during the last two decades, it was partly overtaken by political science and economics. Meanwhile social psychology has become an autonomous discipline.

**Relations with Ecological Geography.** As a reaction against the exaggerations of the sociologist Huntington (1924), who was criticized by Pitirim Sorokin in 1928, an entire generation of American sociologists was dissuaded from taking geographic factors into consideration. Even today, most sociologists and geographers ignore each other.

Until recently, sociologists neglected environmental and climatic factors, but many prominent hybrid scholars did not remain silent. Lewis noted that: "it is important to identify the reasons why tropical countries have lagged during the last two hundred years in the process of modern economic growth" (Lewis 1955, p. 53). Galbraith wrote: "If one marks off a belt a couple of thousand miles in width encircling the earth at the equator one finds within it no developed countries . . . Everywhere the standard of living is low and the span of human life is short" (Galbraith 1957, p. 39-41). The book published by Kamarck, director of the Economic Development Institute of the World Bank, challenges the common perception of tropical areas.

Trypanosomiasis, carried by the tsetse fly, prevented much of Africa from progressing beyond the subsistence level: "For centuries, by killing transport animals, it abetted the isolation of Tropical Africa from the rest of the world and the isolation of the various African peoples from one another" (Kamarck 1976, p. 38). An area of Africa larger than the United States thus had been denied to cattle (Kamarck 1976, p. 39). Agricultural production in the humid tropics is limited by the condition of the soil, which has become laterite (Kamarck 1976, p. 25). Surveys by the World Health Organization and the World Food Organization estimated that parasitic worms infected over one billion people throughout the tropics and subtropics. Hookworm disease, characterized by anemia, weakness, and fever, infected 500 million in those areas (Kamarck 1976, p. 75).

These ecological factors are confirmed by a considerable amount of research in tropical areas during the last several decades by geologists, geographers, biologists, zoologists, botanists, agronomists, epidemiologists, parasitologists, climatologists, experts of the World Bank and several agencies of the United Nations, and hybrid scientists well versed in tropical agriculture, the exploitation of minerals, and the sanitary conditions in those countries. The situation has improved, according to dozens of reports prepared by international organizations. To explain the economic underdevelopment of tropical Africa and other tropical areas, natural sciences and demography are brought into the picture. Dependency theory may be of some help for Latin America and eastern Europe, though much less so for tropical Africa.

The literature on the ecological parameters of the tropics can be contrasted with the literature on the transfer of flora and fauna from one temperate zone to another. For instance, Crosby's *Ecological Imperialism: The Biological Expansion of Europe 900-1900* (1986), casts new light on the building of American power.

This is an example of what can happen when a discipline neglects an important topic. The vacuum left by the absence of sociological studies of this geographic-ecological-economic issue has been filled by eclectic economists and hybrid ecologists.

Sociologists and geographers have met not in vast "interdisciplinary" work but in a series of individual fields such as urban studies. In the

history of this hybrid in the United States, important work came from sociologists in the subfields of "human ecology," geographers influenced by sociologists, and scholars in both disciplines working on spatial statistics. Once a hybrid, urban studies is now a department at many large universities in Europe and the United States.

Urban studies as a quasi-discipline includes subfields that overlap specialties in sociology, geography, and anthropology. It also encompasses architecture, which covers engineering (building design and methods), the natural sciences (climatology, energy conservation), the social sciences (social-physical research), the humanities (history of architecture), and some hybrids of its own (urban planning). Some architects today are well versed in engineering, the natural sciences, the social sciences, and the humanities, as well as urban planning.

Urban studies also has been influenced by economics and economic geography. This hybrid has made its major contribution in the area of location theories for agricultural, industrial, and commercial activities. Communication seems to be much better with geographers and even sociologists than it is with economists, partly because the inductive nature of much of this work makes it difficult to integrate into deductive economic theory.

Other sociologists have drawn from sectors of geography in conjunction with history and economy. Rokkan (1995) has suggested a conceptual framework for comparative political analysis. He weaves together Parsonian pattern variables, the sequence of various kinds of "crises," and the typically Scandinavian notion of center-periphery relations into a geographic schema built around the main Hansa-Rhine-Italy trade routes, the notion of a country's distance from Rome, and whether a state faces seaward or is landbound. This schema is very suggestive not only because it can clarify the different political outcomes in the states of modern Europe but also because it can help one understand why many once-powerful states have disappeared, such as Scotland, Wales, Brittany, Bohemia, Bavaria, and Aragon.

Today, geography's breadth can be seen in the multiplication of hybrid subfields. The discipline now encompasses the subfields of human geography, cultural geography, biogeography, geomorphology, climatology, medical geography, economic

geography, political geography, urban geography, environmental science, regional geography, and cartography. Each subfield relates directly to specialties outside the discipline. Different interests have favored closer contacts sometimes with one field and sometimes with another. These outside fields have made some of geography's most important advances.

As a result of all these trends, there is an incredible fragmentation that has made geography span large areas in both the natural and social sciences, with a general tendency to drift from the former to the latter. From studying habitats, geographers have turned to studying societies. Many traditional geographers have become social scientists.

As in other disciplines, interaction has kept geography on the move. Many geographers have developed their method and have penetrated other disciplines to such a degree that they have become specialists in another discipline (geology, hydrology or ethnology) or one sector of another discipline. Such emigration leaves the old core of the discipline empty. At a symposium on the social sciences in Paris in 1982, a geographer asked, "With the progress of the other social sciences, what remains proper to geography? A residual part, or a boring nomenclature? . . . Does geography still have its own domain, or is it a relic . . . of an old division of labor? Has geography an identity and, if so, of what is it made?" (Brunet 1982, pp. 383, 402). As is true for the other social sciences, its identity can be found in hybrid specialties, not in disciplinary unity.

### CONCLUSION

The contemporary social sciences have experienced three major trends: rapid expansion, fragmentation of formal disciplines by increasing specialization, and recombination of specialties in new hybrid domains. The social sciences have expanded enormously over the last four decades. During the years 1956–1960, the number of citations in the Social Science Citation Index (SSCI) for all social sciences amounted to 2,400,000. Thirty years later, in the years 1986–1990 the number of articles cited in this thesaurus rose to about 18,000,000, increasing by a factor of 7.5 (SSCI 1994, pp. 61–63).

It is difficult to evaluate the number of articles rooted in sociology or relevant for sociologists even if one can locate the origin of the articles and adopt criteria for what is relevant and what is not. The main difficulty comes from the ambiguity and arbitrariness of the borders of these disciplines. Between one-quarter and one-third of the articles cited by sociologists in the last few decades were written by economists, political scientists, psychologists, historians, geographers, and other social scientists.

In 1994, the SSCI contained almost two million citations involving 400,000 authors from fifteen disciplines and from many countries, an average of five citations per author. Among those citations, between 5 and 8 percent referred to articles written by sociologists. Obviously, no one can master the entire spectrum of sociology. There are no paradigms in the discipline, only partial and contested theories and moving borders. One can succeed in finding one's way in the bibliographical labyrinth because the scientific patrimony is structured in sectors, subdisciplines, areas, fields, subfields, specialties, topics, and niches in spite of the fact that the borders are blurred. This increasing specialization within sociology is the main route of scientific advancement. Some scholars recommend an interdisciplinary approach. Just as some seem to believe that the social sciences can be neatly categorized, many others persist in pursuing interdisciplinarity. That recommendation is not realistic because it overlooks an essential phenomenon in the history of science: specialization through a process of fragmentation.

To understand scientific creativity, another phenomenon is even more important than the expansion of the scientific literature and the increase in specialization: the recombination of specialties into new hybrid domains, a phenomenon called the hybridization of scientific knowledge.

A hybrid scholar is a specialist who crosses the borders of her or his home discipline by integrating into her or his research factors, variables, theories, concepts, methods, and substance generated in other disciplines. Different disciplines may proceed from different foci to examine the same phenomenon. This multidisciplinary implies a division of territories between disciplines. In contrast, hybridization implies an overlapping of segments of disciplines, a recombination of knowl-

edge in new and specialized fields. Innovation in each discipline depends largely on exchanges with other fields belonging to other disciplines. At the highest levels, most researchers belong to a hybrid subdiscipline. Alternatively, they may belong to a hybrid field or subfield.

An innovative recombination is a blending of fragments of sciences. When old fields grow, they accumulate such masses of material that they split up. Each fragment of the discipline then confronts the fragments of other fields across disciplinary boundaries, losing contact with its siblings in the old discipline. A specialist in urbanization has less in common with a sociologist studying elite recruitment than he or she does with a geographer doing research on the distribution of cities, who in turn has more in common with a colleague in economics analyzing urban income inequality.

Most hybrid specialties and domains recognize their genealogical roots: political economy, social psychology, social geography, historical sociology, genetic demography, psycholinguistics, political anthropology, social ecology, biogeography, and many others. The hybrid specialties branch out in turn, giving rise, to an even larger member of hybrids (Dogan and Pahre 1990, pp. 63–76).

Among the ISA research committees and study groups, about half focus on hybrid specialties. The number of sociologists who work across disciplinary borders is so high that there is more communication between various fields of sociology and their cognates outside the discipline than there is between fields within sociology.

One can find in the literature of each social science, with the possible exception of linguistics and econometrics, complaints about the “lack of core”: “The substantive core of the discipline may have dissolved” (Halliday and Janowitz 1992, p. 3). Dozens of similar testimonies could be collected. If so many scholars formulate the same diagnosis, that means that most disciplines are facing a problem of self-identity. However, if one considers that the real world cannot be cut into disciplinary pieces, this issue of disciplinary identity may appear fallacious.

It is difficult or impossible to inquire into the large social phenomena within a strictly monodisciplinary framework. Only by taking a position at the crossroads of many branches of

knowledge can one explain the impact of technological advancement on structural unemployment in western Europe, the proliferation of giant cities in the third world, the economic decline of the United Kingdom and the economic growth of Japan, or how a child learns to speak. Whenever a question of such magnitude is raised, one finds oneself at the intersection of numerous disciplines and specialties. All major issues cross the formal borders of disciplines: war and peace, generational change, the freedom–equality nexus, individualism in advanced societies, and fundamentalism in traditional societies. Most specialists are not located in the so-called core of a discipline. They are in the outer rings, in contact with specialists from other disciplines. They borrow and lend at the frontiers; they are hybrid scholars. The notion of hybridization does not mean “two whole disciplines in a single skull” but a recombination of two or several domains of knowledge originating from different disciplines.

Most classical sociologists were interdisciplinary generalists, but in recent times, cross-disciplinary advancements have been achieved not by generalists but by hybrid specialists. The hybrid specialist today may be in reality a “marginal” scholar in each of the disciplines from which he or she borrows, including his or her original discipline, but such a specialist becomes central to the intersection of two or several disciplines (Dogan 1999).

Today most social scientists admit that the best alternative to the difficulty of experimentation in their disciplines is the comparative method, which is one of the few ways to validate or falsify generalizations in the “soft” sciences. The comparative method is the key to circulation among sciences.

Comparative sociologists and comparative political scientists have developed methods to a greater extent than have workers in other social sciences. One of them wrote: “There is no noncomparative sociological theory. All scientific analyses are a subset of the general set entitled comparative analysis . . . any generalized statement involving variables implies a comparison” (Levy 1970, p. 100).

Major social phenomena cannot be explained in a strictly monodisciplinary framework or in the absence of a comparative perspective. It is only by taking up a position at the crossroad of various branches of knowledge and simultaneously adopt-

ing comparative perspective that social scientists can advance knowledge. The intersections of hybrid specialties and comparative approaches are privileged sites in the social sciences.

REFERENCES

- Aronson, Sidney 1969 "Obstacles to a Rapprochement between History and Sociology: A Sociologist's View." In Muzafer Sherif and Caroline Sherif, eds., *Interdisciplinary Relationships in the Social Sciences*. Chicago: Adline.
- Aymard, Maurice 1988 "Histoire et Sociologie." In H. Mendras and M. Verret, eds., *Les Champs de la Sociologie Française*. Paris: Colin.
- Bartoli, H. 1991 *L'Économie Unidimensionnelle*. Paris: Economica.
- Boudon, Raymond 1979 "Sociologie et Histoire: L'analyse Sociologique du Singulier." In *La Logique du Social*. Paris: Hachette.
- Braudel, Fernand 1962 "Histoire et Sociologie." In G. Gurvitch, ed., *Traité de Sociologie*. Paris: PUF.
- Brunet, Roger 1982 "La Géographie." In *Les Sciences de l'Homme et de la Société en France: La Documentation Française*.
- Buchanan, James 1966 "Economics and Its Scientific Neighbors." In Sherman R. Krupp, ed., *The Structure of Economic Science*. Englewood Cliffs, N.J.: Prentice-Hall.
- Burke, Peter 1980 *Sociology and History*. London: Allen & Unwin.
- Cahman, Werner, and Alvin Boskoff (eds.) 1964 *Sociology and History and Research*. Free Press.
- Calhoun, Craig 1992 "Sociology, Other Disciplines and the Project of a General Understanding of Social Life." In T. C. Halliday and M. Janowitz, eds., *Sociology and Its Publics*. Chicago: University of Chicago Press.
- , and K. C. Land 1989 "Symposium: Smelser's Handbook: An Assessment." *Contemporary Sociology* 18:475–513.
- Chaunu, Pierre 1979 "Interview." *Le Courrier du CNRS* 33:5.
- Collins, Randall 1986 "Is 1980s Sociology in the Doll-drums?" *American Journal of Sociology* 91 (6):1336–1355.
- Crane, Diana, and Henry Small 1992 "American Sociology since the Seventies: The Emerging Identity Crisis in the Discipline." In T. C. Halliday and M. Janowitz, eds., *Sociology and Its Publics*. Chicago: University of Chicago Press.
- Dogan, Mattei 1994 "Fragmentation of the Social Sciences and Recombination of Specialties." *International Social Science Journal* 139:27–42.
- 1996 "Political Science and the Other Social Sciences." In R. Goodin and H. D. Klingemann, eds., *A New Handbook of Political Science*. Oxford, U.K.: Oxford University Press.
- 1997 "The New Social Sciences: Crack in the Disciplinary Walls." *International Social Science Journal*. 153:429–443.
- 1999 "Marginality." In M. Runco and S. Pritzker, eds., *Encyclopedia of Creativity*. New York: Academic Press.
- , and Robert Pahre 1990 *Creative Marginality, Innovation at the Intersections of Social Sciences*. Boulder, Colo.: Westview.
- Galbraith, J. K. 1957 "Conditions for Economic Change in Underdeveloped Countries." *Journal of Farm Economics* 33:255–269.
- Gay, David, and Alan R. Waters 1983 "The Interrelationships of Economics with the Social Sciences." *Social Science Journal* 20(3):1–8.
- Goodin, Robert E., and Hans-Dieter Klingemann 1996 *A New Handbook of Political Science*. Oxford, U.K.: Oxford University Press.
- Halliday, Terence, and Morris Janowitz (eds.) 1992 *Sociology and Its Publics*. Chicago: University of Chicago Press.
- Hayek, F. A. 1956 "The Dilemma of Specialization." In Leonard White, ed., *The State of the Social Sciences*. Chicago: University of Chicago Press.
- Hirshleifer, Jack 1985 "The Expanding Domain of Economics." *American Economic Review* 72 (6):53–68.
- Huntington, Ellsworth 1924 *Civilization and Climate*. New Haven, Conn.: Yale University Press.
- Inkeles, Alex 1970 "Sociological Theory in Relation to Social Psychological Variables." In J. C. McKinney and E. A. Tiryakian, eds., *Theoretical Sociology*. New York: Appleton.
- 1983 "The Sociological Contribution to Advances in the Social Sciences." *Social Science Journal* 20(3):27–44.
- Intriligator, Michael 1985 "Independence among the Behavioral Sciences." *Structural Changes and Economic Dynamics* 1:1–9.
- Johnston, Barry U. 1997 "Dominant Intellectual Tradition and the Coherence or Disintegration of the Discipline?" Paper presented at the International Institute of Sociology, Cologne, Germany.
- Kamarck, Andrew M. 1976 *The Tropics and Economic Development*. Johns Hopkins University Press for the World Bank.
- Klein, Julie Thompson 1996 *Crossing Boundaries*. Charlottesville: University of Virginia Press.



- Knapp, Peter 1984 "Can Social Theory Escape from History: View of History in Social Science." *History and Theory* 23(1):34–52.
- Kohn, Melvin, L. (ed.) 1989 *Cross-National Research in Sociology*. London: Sage.
- Levy, Marion, J. 1970 "Scientific Analysis is a Subset of Comparative Analysis." In John C. McKinney and Edward A. Tiryakian, eds., *Theoretical Sociology*. New York: Appleton.
- Lewis, W. A. 1955 *The Theory of Economic Growth*. Allen Unwin.
- Lipset, S. M., and R. Hofstadter 1968 *Sociology and History: Methods*. New York: Basic Books.
- Martinelli, A., and N. Smelser 1990 "Economic Sociology, Historical Trends and Analytic Issues." *Current Sociology* 38 (2):1–49.
- McKinney, John C., and Edward A. Tyriakian eds. 1970 *Theoretical Sociology*. New York: Appleton.
- Messer-Davidow, E., D. R. Schumway, and D. J. Sylvan, eds. 1993 *Knowledge: Historical and Critical Studies in Disciplinarity*. Charlottesville: University of Virginia Press.
- Offe, Claus 1992 "Political Economy: Sociological Perspectives." In Robert B. Goodin and Hans-Dieter Klingemann, eds., *A New Handbook of Political Sciences*. Oxford, U.K.: Oxford University Press.
- Radnitzky, G., and P. Bernholz 1986 *Economic Imperialism: The Economic Approach Outside the Traditional Areas of Economics*. New York: Paragon.
- Riggs, Fred W. 1984 "Development." In G. Sartori, ed., *Social Science Concepts: A Systematic Analysis*. Sage.
- Rigney, D., and D. Barnes 1980 "Patterns of Interdisciplinary Citation in the Social Sciences." *Social Science Quarterly* 114–127.
- Rokkan, Stein 1995 Special issue dedicated to Rokkan's geoeconomic model. *Revue Internationale de Politique Comparée* 2 (1):5–170.
- Rose, Richard 1991 "Institutionalizing Professional Political Science in Europe." *Political Studies* 39 (3):446–462.
- Scheuch, Erwin, K. 1992 "German Sociology" In E. F. Borgatta and M. L. Borgatta, eds., *Encyclopedia of Sociology*. New York: MacMillan.
- Simiand, François 1903 "Méthode historique et science sociale." *Revue de Synthèse Historique*.
- Smelser, Neil 1967 "Sociology and the Other Social Sciences." In P. Lazarsfeld, H. Sewell, and H. L. Wilensky, eds., *The Uses of Sociology*. New York: Basic Books.
- ed. 1988 *Handbook of Sociology*. London: Sage.
- Smelser, Neil, and Richard Swedberg eds. 1994 *Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press.
- Social Sciences Citations Index, SSCI, Philadelphia Institute for Scientific Information.
- Solow, Robert 1970 "Science and Ideology in Economics." *The Public Interest* 21:94–107.
- Szacki, Jerzy 1982 "The History of Sociology and Substantive Sociological Theories." In T. Bottomore, S. Nowak, and M. Sokolowska, eds., *Sociology: The State of the Art*. London: Sage.
- Szenberg, Michael 1992 *Eminent Economists: Their Life Philosophies*. Cambridge, U.K.: Cambridge University Press.
- Tilly, Charles 1981 *As Sociology Meets History*. New York: Academic Press.
- 1984 *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- Turner, R. H. 1991 "The Many Faces of American Sociology: A Discipline in Search of Identity." In D. Easton and C. Schelling, eds., *Divided Knowledges: Across Disciplines, Across Cultures*. Newbury Park, Calif.: Sage.

MATTEI DOGAN

## SOCIOLOGY OF EDUCATION

In the broadest perspective, education refers to all efforts to impart knowledge and shape values; hence, it has essentially the same meaning as socialization. However, when sociologists speak of education, they generally use a more specific meaning: the deliberate process, outside the family, by which societies transmit knowledge, values, and norms to prepare young people for adult roles (and, to a lesser extent, prepare adults for new roles). This process acquires institutional status when these activities make instruction the central defining purpose, are differentiated from other social realms, and involve defined roles of teacher and learner (Clark 1968). Schools exemplify this type of institutionalization.

The central insight of the sociology of education is that schools are socially embedded institutions that are crucially shaped by their social environment and crucially shape it. The field encompasses both micro- and macro-sociological concerns in diverse subfields such as stratification, economic development, socialization and the family, organi-

zations, culture, and the sociology of knowledge. To understand modern society, it is essential to understand the role of education. Not only is education a primary agent of socialization and allocation, modern societies have developed formidable ideologies that suggest that education *should* have this defining impact (Meyer 1977).

Durkheim (1977) was the intellectual pioneer in this field, tracing the historical connections between the form and content of schools and larger social forces such as the rise of the bourgeoisie and the trend toward individualism. Largely because the field focuses so intensively on stratification-related issues (e.g., the impact of family background on educational attainment), the larger issues raised by Marx and Weber are readily evident in current scholarship. However, as Dreeben's (1994) historical account indicates, the direct contribution of the discipline's founders to the development of the sociology of education in the United States was minimal; indeed, even the foremost early American sociologists in the field did not decisively shape its development.

In *The Sociology of Teaching*, Waller (1932) examined teaching as an occupational role and school organization as a mechanism of social control. He emphasized the role of the school in the conflict-ridden socialization of the young as well as the interpersonal and organizational mechanisms that furthered students' acceptance of the normative order. Although now recognized as a classic, Waller's analysis stimulated little work for several decades.

Although less focused on education per se, Sorokin (1927) portrayed schools as a key channel of mobility with their own distinctive form of social testing. He argued that increasing opportunities for schooling would stratify the society, not level it. However, Blau and Duncan's (1967) paradigm-setting study of status attainment (see below) did not refer to Sorokin's analysis of education despite their appreciation of his larger concern for the significance of social mobility. Warner's and Hollingshead's community studies considered education integral to community social organization, especially through its connection to the stratification system, but their influence, like that of Waller and Sorokin, was more a matter of suggesting general ideas than of establishing a cumulative research tradition.

As a subfield within the sociological discipline, the sociology of education has been propelled largely by a host of practical, policy-related issues that emerged with the development of the mass educational system. Essentially, research has focused on whether education has delivered on its promise of creating more rational, culturally adapted, and productive individuals and, by extension, a "better" society. The field was particularly energized by the egalitarian concerns of the 1960s: How "fair" is the distribution of opportunity in schools and in the larger society, and how can disparities be reduced? These questions continue to animate the field.

### THEORETICAL DEBATES

Much research, even the most policy-oriented, has been grounded, often implicitly, in more general analytic perspectives on the role of education in modern society. The two main orientations are functionalism and conflict theory, though other, less encompassing perspectives also have shaped the field significantly.

**Functionalism.** In the functionalist view, schools serve the presumed needs of a social order committed to rationality, meritocracy, and democracy. They provide individuals with the necessary cognitive skills and cultural outlook to be successful workers and citizens (Parsons 1959; Dreeben 1968) and provide society with an efficient, fair way of sorting and selecting "talents" so that the most capable can assume the most responsible positions (Clark 1962). Complementing this sociological work is human capital theory in economics, which contends that investment in education enhances individual productivity and aggregate economic growth (Schultz 1961). The criticism in the 1980s that poor schooling had contributed significantly to America's decline in the international economy reflects a popular version of this theoretical orientation.

However, in the 1970s, both the increasing prominence of critical political forces and the accumulated weight of research spurred a theoretical challenge. Important parts of the empirical base of functionalism were questioned: that schools taught productive skills, that mass education had ushered in a meritocratic social order, and that education had furthered social equality. A number of conflict-oriented approaches emerged.

**Neo-Marxist Theory.** Neo-Marxist scholars have provided the most thorough challenge to the functionalist position. For all the diversity within this conflict theory, the main point is that the organization of schools largely reflects the dictates of the corporate-capitalist economy. In the most noted formulation, Bowles and Gintis (1976) argue that education must fulfill the needs of capitalism: efficiently allocating differently socialized individuals to appropriate slots in the corporate hierarchy, transferring privilege from generation to generation, and accomplishing both while maintaining a semblance of legitimacy. Thus, the changing demands of capitalist production and the power of capitalist elites determine the nature of the educational system.

More recent neo-Marxist scholarship (Willis 1981) emphasized that schools are not only agents of social reproduction but also important sites of resistance to the capitalist order. Many neo-Marxists also have emphasized the “relative autonomy” of the state from economic forces and, correspondingly, the partial responsiveness of schools to demands from subordinate groups (Carnoy and Levin 1985). Other scholars in this general critical tradition have turned in “post-Marxist” directions, emphasizing inequities related to gender and race along with class, but the common, defining point remains that educational inequities reflect and perpetuate the inequities of capitalist society and that oppressed groups have an objective interest in fundamental social transformation (Aronowitz and Giroux 1985). This newer critical approach has developed with relatively little connection to mainstream approaches (i.e., positivistic, often reform-oriented research) despite some similarities in concerns (e.g., student disruptions and challenges to authority in schools) (Davies 1995).

Obviously, neo-Marxists do not share the essentially benign vision of the social order in functionalist thought, but both perspectives view the organization of schooling as “intimately connected with the changing character of work and the larger process of industrialization in modern society” (Hurn 1993, p. 86). These competing perspectives are rooted in similar logical forms of causal argument: To explain educational organization and change, functionalists invoke the “needs” of the society, while neo-Marxists invoke the “needs” of the capitalist order for the same purpose. Critics contend that both perspectives posit

an overly tight, rational link between schools and the economy and concomitantly downplay the institutional autonomy as schools as well as the complexity of political struggles over education (Kingston 1986).

**Status Conflict.** Arising out of the Weberian tradition, the status conflict approach emphasizes the attempts of various groups—primarily defined by ethnicity, race, and class—to use education as a mechanism to win or maintain privilege (Collins 1979). The evolving structure of the educational system reflects the outcomes of these struggles as groups attempt to control the system for their own benefit. With varying success, status groups use education both to build group cohesion and to restrict entry to desired positions to those certified by “their” schools. However, as lower-status groups seek social mobility by acquiring more educational credentials, enrollments may expand beyond what is technically necessary. In this view, then, the educational system is not necessarily functional to capitalist interests or other imputed system needs.

Consistent with this view, a primary effect of schools, especially at the elite level, is to provide *cultural capital*, of which educational credentials are the main markers (Bourdieu and Passeron 1977). This form of capital refers to the personal style, social outlooks and values, and aesthetic tastes that make a person suitable for socially valued positions. (The point of comparison is *human capital*, an individual’s productive, technical skills.) In this perspective, education is rewarded because occupational gatekeepers value particular forms of cultural capital, and thus education is a key mechanism of class and status reproduction.

**The Interpretative Tradition.** Sociologists in the interpretative tradition view schools as places where meaning is socially constructed through everyday interactions. This tradition incorporates the general orientations of phenomenology, symbolic interactionism, and ethnomethodology. Accordingly, micro-level concerns predominate—for example, what do teachers expect their students to learn, and how do those expectations condition their conduct in class?—and research tends to rely on qualitative techniques. This tradition is unified by a general sense of what kinds of questions to ask (and how to ask them) rather than a set of related

theoretical propositions or a body of accumulated findings.

### EMPIRICAL STUDIES

The highly selective review of empirical studies that follows focuses on the two key questions in contemporary American sociology of education: (1) How is education involved in the distribution of life chances? (2) How are family status and school characteristics connected to educational attainment and/or academic achievement? With few exceptions, analyses of education in other countries are not considered. The field is dominated by American research, and American sociologists have engaged in relatively little comparative research. Baker (1994) speculates that this lack of a comparative research tradition in the United States reflects both a belief in American "exceptionalism" (for instance, an extreme emphasis on mass access) and a strong focus on micro-level issues that do not necessarily call for comparative research designs.

**Schooling and Life Chances.** Throughout the twentieth century in all industrial countries, there has been a dramatic upgrading in the occupational structure *and* a dramatic expansion in educational systems. Ever more jobs have come to require academic qualifications, a process that usually is interpreted as being driven by the rationalism and universalism of modernization. In this functionalist perspective, academic skills are presumed to be technically required and meritocratically rewarded, transforming the stratification system so that individual achievements rather than ascriptive characteristics determine life chances.

This interpretation has been subject to empirical test at two levels: (1) the *individual level*—to what extent, absolutely and relatively, does education affect economic attainment? and (2) the *macro level*—to what extent have educational expansion and the increasing significance of schools for occupational attainment increased overall equality of opportunity?

At the first level, as part of the general analysis of *status attainment*, researchers have concentrated on measuring the connection between individuals' schooling and their economic position. Building on Blau and Duncan's (1967) work, researchers have repeatedly documented in multivariate mod-

els that education (measured in years of schooling and degree completion) has by far the largest independent impact on adult attainment (Featherman and Hauser 1978; Jencks et al. 1979). By comparison, the net direct effects of family status (usually measured in terms of parental education and occupation) are modest. Indeed, among the college-educated in recent years, higher family status confers no extra advantage at all (Hout 1988).

Earlier in life, however, family status is substantially related to educational attainment. The total effect (direct and indirect) of family status on occupational attainment is therefore substantial, though its impact is mediated very largely through educational attainment. In effect, then, education plays a double-sided role in the stratification process. Education is the great equalizer: It confers largely similar benefits to all regardless of family origins. However, it is also the great reproducer: Higher-status families transmit their position across generations largely through the educational attainment of their children.

The strong connection between schooling and occupational attainment is open to diverse interpretations. Most prominently, human capital theory suggests that education enhances productivity, and because people are paid in accordance with their marginal productivity, the well educated enjoy greater prospects. In favor of this interpretation is the fact that schooling is demonstrably linked to the enhancement of academic competencies (Fischer et al. 1996) and that basic academic skills are substantially correlated with job performance in a wide variety of settings (Hunter 1986).

By contrast, credentials theory portrays the educational institution as a sorting device in which individuals are slotted to particular positions in the occupational hierarchy on the basis of academic credentials, often with little regard for their individual productive capacities. The fact that possessing specific credentials (especially a college degree) has positive career effects, net of both years of schooling and measured academic ability, provides indirect support for this view. That is, there appears to be a "sheepskin effect," so that employers value the degree per se, although people with degrees may have unmeasured productive capacities or dispositions that account for their success (Jencks et al. 1979). Moreover, the credentialist argument is strengthened by the fact

that in some elite segments of the labor market, employers primarily recruit graduates of certain prestigious programs and make little effort to discern differences in the academic-based skills of those included in the restricted applicant pool (Kingston and Clawson 1990).

Both views seem to have some merit; indeed, they may be partially complementary. Employers may generally use educational attainment as a low-cost, rough proxy for productive skill, and for certain positions they may favor holders of particular degrees because of their presumed cultural dispositions and the prestige that their presence lends the organization. The relative explanatory power of the human capital and credentialist perspectives may vary across segments of the labor market.

At the macro level, it might be expected that the great expansion of access to education has reduced the impact of family origins on educational attainment, increasing equality of opportunity, but that has proved to be more the exception than the rule. A rigorous thirteen-country comparative study identified two patterns: greater equalization among socioeconomic strata in the Netherlands and Sweden and virtual stability in the rest, including the United States (Shavit and Blossfeld 1993), where the strata have largely maintained their relative positions as average attainment has increased. Thus, the impact of educational policies designed to promote equality appears minimal; even in Sweden and the Netherlands, the trend toward equalization emerged before reforms were introduced.

**Socioeconomic Status and Achievement.** Given the centrality of educational attainment in the general attainment process, researchers have focused on the substantial relationship between socioeconomic status and educational attainment. (This relationship appears to be stronger in highly developed societies than in developing societies.) The best predictor of educational attainment is academic achievement (i.e., higher grades and test scores); the school system consistently rewards academic performance and in that sense is meritocratic. Regardless of academic performance, children from socially advantaged families have somewhat disproportionate success in moving through the educational system, but the main

reason higher-status students have this success is that they achieve better in schools.

The question here is, Why do higher-status students achieve better in schools? Clearly, there is no simple answer. Research has pointed to the following family-related factors, among others:

1. *Material resources.* Richer families can purchase the materials (e.g., books) and experiences that foster intellectual development.
2. *Parental expectations and/or encouragement.* Well-educated parents more actively stress the importance of academic achievement, and their own success through schooling encourages their children to accept that value.
3. *Direct parental involvement in home learning activities.* Higher-status parents are more willing and able to teach academic lessons at home and help with homework.
4. *Verbal and analytic stimulation.* In higher-status families, interactions between parents and children are more likely to promote verbal sophistication and reasoning.
5. *Family structure and parenting style.* The presence of two parents and parenting styles involving warm interactions favor academic achievement, and both factors are related to socioeconomic status (SES).
6. *Parental involvement in schools.* Higher-status parents are better able to interact effectively with teachers and administrators to secure favorable treatment and understand expectations.
7. *Cultural "fit" with schools.* The cultural styles of higher-status students are more compatible with the prevailing norms and values in schools.
8. *Social capital.* Initially Coleman's (1988) idea, this refers to the extent and nature of the connections between parents and children as well as the connections with other family and community members. By providing informational, emotional, and other resources, these connections facilitate adaptations to the demands of schools.

9. *Social context.* Higher-status families are likely to live in communities where other families promote achievement and their children's peers are committed to academic achievement.
10. *Genetic advantage.* Early IQ is related to SES, and intelligence is related to academic performance.

Individually, none of these factors seems to account for a large part of the overall relationship between SES and academic achievement, nor is the relative significance of these factors clear, yet the very length of the list suggests the complexity of the issue. Higher-status students are not all similarly advantaged by each of these factors, and lower-status students are not all similarly disadvantaged by each one. The substantial aggregate relationship between SES and achievement undoubtedly reflects complex interactions among the many home-related contributing causes. As is more thoroughly discussed below, the mediating impact of school resources and practices is much less consequential.

**The Racial Gap.** The black-white disparity in academic performance remains large despite some notable reductions in recent years, and it is economically significant. A number of researchers have shown that for younger cohorts, the racial disparity in earnings is accounted for very largely by differences in basic academic skills as measured by scores on tests such as Armed Forces Qualifications Test (Farkas 1996).

Why this gap persists is unclear, partly because until recently, sociologists and other social scientists were wary of addressing such a politically explosive issue. Most relevant for the discussion here is the fact this gap cannot be explained by blacks' lesser school resources (see "School Effects," below). Largely drawing on the work of scholars in related fields, the sociological consensus appears to be that the racial disparity does not reflect a *group*-based difference in genetic potential (Jencks and Phillips 1998). (At the individual level, there is undoubtedly some genetic component to IQ among people of all races.) Moreover, this gap cannot be attributed largely to racial differences in economic advantage: Socioeconomic status explains only about a third of it. However, a broader index of family environment, including parental practices, may account for up to two-

thirds of the gap (Phillips et al. 1998). A complete explanation probably will involve many of the factors previously noted in the discussion of the relation between SES and achievement but also include the distinctive cultural barriers that "involuntary minorities" face in many societies (Ogbu 1978) as well as subtle interactional processes within schools.

Racial disparities in educational attainment have declined dramatically. High school graduation rates are now virtually the same, and the remaining disparity in college attendance reflects blacks' lower economic resources, not a distinctive racial barrier.

**School Effects.** The governmental report *Equality of Educational Opportunity* (Coleman et al. 1966) strongly challenged conventional wisdom about the connections among economic status, schools, and achievement. In doing so, it fundamentally shaped the agenda for further research in this area.

Attempting to identify the characteristics of schools that improve learning, the so-called Coleman Report documented two key points. First, there is a weak relationship between social status and school quality as measured by indicators such as expenditure per pupil, teachers' experience, and class size despite considerable racial segregation. Second, these measures of school quality have very little *overall* effect on school achievement (scores on standardized tests) independent of students' family background. The Coleman Report also showed, however, that school effects were notably larger for black and Hispanic students than they were for whites and Asians. Among the school effects, the racial composition of schools was the most critical: Blacks did somewhat better in integrated schools.

Later research largely validated the main conclusions of the Coleman Report, but also modified them, often by considering more subtle aspects of school quality. For instance, some school resources, including expenditures, seem to enhance achievement, but the predominance of home factors on achievement remains undisputed. In regard to another between-schools effect, Coleman argued for the educational superiority of Catholic schools, an advantage he attributed to their communal caring spirit and high academic expectations for all students. The Coleman Report did not consider such cultural matters or specific educational prac-

tices. Much of the post-Coleman Report research focused on within-school effects because gross between-school effects appeared to be relatively minor.

Ability grouping in elementary schools and tracking in high schools have attracted attention, largely as a source of inequalities of academic performance. The premise of these practices is that students differ substantially in academic ability and will learn more if taught with students of similar ability. Although many different practices are grouped under the term "tracking," students in the "top" groups generally receive a more demanding education, with higher expectations, more sophisticated content, and a quicker pace, and are disproportionately from advantaged families. The obvious but not fully settled issue is whether schools "discriminate" in favor of the socially advantaged in making placements. At the high school level, controlling for measures of prior achievement (themselves affected by family factors), higher SES seems to enhance one's chances modestly, though achievement factors are predominant in placement. Blacks are somewhat favored in the process if one controls for prior achievement. At the elementary school level, research is less consistent, though one study indicates that neither test scores nor family background predicts early reading group placement (Pallas et al. 1994).

Another important but not fully settled issue is whether students in certain ability groups or tracks learn more because of their placement. Gamoran (1992) shows that the effects of tracking are conditioned substantially by the characteristics of the tracking system (for example, how much mobility between tracks is allowed) and subject matter. However, by way of gross summary, higher track placement per se generally seems to have a modestly beneficial impact on achievement and also seems to increase students' educational aspirations and self-esteem. However, to exemplify the important exceptions to this generalization, it appears that within-class grouping for elementary school mathematics may help both low and high groups.

**Teacher Expectations.** It is commonly supposed that differences in teachers' expectations explain at least some of the racial and socioeconomic disparities in academic achievement. The

claim is that a self-fulfilling prophecy is at work: Teachers expect less from socially disadvantaged students and treat them accordingly, and therefore these students perform less well in school. Rosenthal and Jacobson's (1968) small-scale experimental study provided the initial impetus for this argument, but follow-up studies in real classrooms suggest that teachers' expectations have little or no effects on later performance.

If the standard for fairness is race neutrality in light of past academic performance, there is little evidence of racial bias in teachers' expectations, but some limited evidence suggests that teachers' beliefs are more consequential for blacks than for whites (Ferguson 1998). More generally, research has not established that socially discriminatory practices in schools significantly explain the link between family and/or racial status and achievement.

**Contextual Effects.** Not only do students come to school with different backgrounds that affect learning, schools provide students with different social environments that are importantly shaped by the economic and racial composition of the student body. Because peers are so influential in children's and adolescents' lives, the obvious question is whether the social composition of a school affects individual learning beyond the effects attributable to an individual's status characteristics. This issue has had practical significance in light of ongoing public debates about the impact of racial desegregation initiatives.

Evidence about the impact of social context on learning is mixed, but in any case the impact is not large. To the extent that the SES of a student body is consequential, this appears to result from the connection between SES and a positive academic climate in a school. Greater racial integration generally seems to promote black student achievement slightly, but the benefits are more pronounced for black students when they actually have classroom contact with white students rather than just attending a formerly integrated school.

More recent research suggests an important cautionary note about whether integration "works." Entwistle and Alexander (1992), for example, show that on a yearlong basis, in the early grades black students in integrated schools had better reading comprehension than did black students in segregated schools. However, the apparent advantage of integrated schools totally reflects the fact that

black students at integrated schools improved more during the summer than did black students at segregated schools. During the school year black students did slightly better in segregated schools. This analysis exemplifies the increasing recognition that a simple conclusion about integration—works versus does not work—is inadequate.

**Learning through the Year.** As should be evident, a major issue in the sociology of education is separating the effects of the home from the effects of the school. The perplexing finding is that racial and class disparities in achievement in the early grades become substantially greater as students progress through school. Critics have seized on this finding to indict schools for discriminatory practices that exacerbate social inequality.

However, so-called summer learning research suggests a different interpretation (Alexander and Entwistle 1995; Gamoran 1995). Examining the same students' test scores at the beginning and ending of each of several school years, researchers have shown that (1) despite initial disparities, advantaged and disadvantaged groups have roughly similar gains in achievement during the school year but that (2) advantaged students continue to improve during the summer while disadvantaged students stagnate or decline. As the effects of this process accumulate over the years, initial disparities become ever larger. The important implication is that schools neither reduce nor add to the inequalities that are rooted in homes. Schools in effect passively reproduce existing inequalities.

**Enhancing Performance.** Although crude measures of school resources (e.g., teacher certification levels) appear at most to be weakly related to school achievement, a burgeoning and increasingly sophisticated line of research finds that effective schools can be identified. These schools are marked by strong leadership committed to academically focused goals and order, high academic demands, and frequent practice of academic skills. This research also directs attention to the benefits of an overall communal culture and classroom interactions that stress cooperative efforts between students and teachers (Lee and Croninger 1994). What appears critical is how resources are organizationally applied.

**Macro-Level Effects.** This article has focused on the experiences of individuals: how education affects life chances and how personal character-

istics and school experiences affect learning. The unit of analysis, in other words, is the individual. Research in the field much less commonly takes the society as the unit of analysis: How do societal features shape the nature of the educational system? How do the features of this system affect other societal arrangements? An important example of macroanalysis is the generally limited impact of increasing educational access on equality of opportunity (see "Schooling and Life Chances," above). Perhaps the most studied macro-level topic is the relationship between educational expansion and economic growth.

If the individual economic benefits of education are clear, the impact of educational expansion on economic growth is less certain. The orthodox view in economics is that educational expansion promotes growth. This view follows from human capital theory: People with more schooling get higher pay *because* they are more productive, and if more people get more schooling, they will produce more and get paid more, with the aggregate effect being economic growth. Many sociologists are at least partially skeptical of this idea. Undoubtedly, more educated workers get paid more, but the positive (private) rate of return they enjoy reflects greater productivity only if it is assumed that the labor market is perfectly competitive and in equilibrium. This assumption is at least partly problematic given socially discriminatory employment practices, internal labor markets with seniority rules and restricted job mobility, professional and union restrictions of labor supply, and public sector employment with politically determined pay structures.

Allocation theory—which also is called the credentialing perspective—offers an alternative explanation of the link between education and economic rewards. In brief, employers assume that the more educated, as a group, are *relatively* desirable people to hire (for reasons that may or may not reflect their individual productive capacities); and in turn, how people are ranked in the educational hierarchy becomes linked to how they are ranked in the hierarchy of the *existing* job structure. Educational expansion, then, does not necessarily promote economic growth; it only affects who gets which of the already existing jobs. To the extent that credentialing processes are operative, it is impossible to infer aggregate effects on growth from individual-level data on income.



Given the ambiguous implications of individual income data, the best way to examine the issue is through aggregate, *national*-level studies of how education affects economic growth. The accumulated weight of this research undercuts claims about the large *universal* benefits of more education of all types. Benavot (1992), for example, establishes the following for a large sample of developed and poor countries in the period 1913–1985: Throughout the period, the expansion of primary education promoted growth; the expansion of secondary education had more modest impact, and only during times of worldwide prosperity; and tertiary education tended to *retard* growth at all times. In the United States, moreover, tertiary enrollments have never stimulated growth (Walters and Rubinson, 1983).

However, even if more education is not a universal economic “fix,” in certain circumstances particular types of education may stimulate growth in specific sectors. Reviewing single-country times series studies that use an aggregate production function model, Rubinson and Fuller (1992) conclude that education had the greatest beneficial impact when it created the kinds of skills that were suited to an economy’s sectoral mix and technological demands. However, a good fit between the educational system and the economy is by no means certain because educational expansion and the actual educational content of schools are so often driven by political processes, not technological demands.

Even if the actual economic impact of education is often less than is commonly supposed, the widespread belief in the general modernizing benefits of education is central to an ideology that permeates the entire world. Indeed, in Meyer’s institutionalist perspective (Meyer and Hannan 1979; Meyer 1977), the quest to appear modern has induced later-developing societies to mimic the educational practices of the early modernizers so that many school structures, rituals, and formal curricular contents are remarkably similar throughout the world. In turn, this institutionalized similarity means that on a global basis, certain types of knowledge become defined as relatively significant, the elite and mass positions become defined and legitimated by educational certification, and assumptions about a national culture rest on the existence of mass education. Nevertheless, if education is associated at the individual level with

certain democratic values, educational expansion per se does not appear to contribute to the emergence of democratic regimes or state power.

### THE REFORMIST PROJECT

Policy debates about education have often been contentious, fueled by larger ideological and political struggles. In conservative times, schools have been pressed to emphasize discipline and social and/or intellectual sorting; conversely, in more liberal times, issues of equality and inclusion have come to the fore. The apparent result is cyclical, pendulum-like swings in policy between, say, an emphasis on common core requirements and highly differentiated curricula.

While differences at the rhetorical level have sometimes been sharp, actual changes in practice in much of the twentieth century have been relatively minor. This reflects the institutionalization of the school, meaning that there is a widespread collective sense of what a “real” school is like (Tyack and Cuban 1995). This institutionalization rests on popular legitimization and the recurrent practices of school administrators and teachers. Concrete practices such as the division of knowledge into particular subject areas, the spatial organization of classrooms, and the separation of students into age-based grades are all part of the “real” school. Educational practices that depart from this pattern have had limited acceptance, for example, open classrooms in the 1970s. The lesson for current reformers is that policies that modify institutionalized practices, not fundamentally challenge them, are more likely to be successful and that the political support of in-the-school educators is critical for success.

Indeed, much policy-oriented research has had a mildly reformist bent, primarily concerned with making existing schools “work better.” That has largely — and narrowly — meant producing students with higher scores on standardized tests in the basic academic subject areas. Critics have questioned both the validity of these tests and the desirability of evaluating school “success” in these limited terms alone. Proponents contend that scores on these tests have considerable predictive validity for later school and occupational performance and that their standardized results permit rigorous comparisons across groups and school settings.

The welter of policy-related studies is impossible to summarize here (and the distinction between sociological research and educational research is hardly sharp), but two general types of contributions from sociologists stand out. The first is essentially a debunking contribution: Sociologists have shown what does *not* work despite fervent beliefs to the contrary. The previously discussed Coleman Report is the most prominent example, undercutting the liberal faith of the 1960s that differences in school resources substantially account for racial and socioeconomic differences in academic achievement.

The second contribution is essentially methodological, alerting policymakers to the fact that many apparent school effects may largely or even totally reflect selection biases. That is, if groups of students are subject to different educational practices, are any differences in their performance attributable to the educational practices per se, or are different sorts of students subject to different practices, thus accounting for the association between practice and performance? In recent years, controversies about the efficacy of private and Catholic schools, related to larger debates about school choice plans, have centrally involved the issue of selection bias. In the most sophisticated study, Bryk et al. (1993) demonstrate net positive effects of Catholic schools on academic achievement and show that the gap in achievement between white and minority students is reduced in Catholic schools.

Even with the most sophisticated multilevel, multivariate statistical models, however, sociologists cannot make firm *causal* claims by analyzing survey data. However, by ruling out many potential sources of spuriousness, these analyses can suggest interventions that are likely to have a positive effect. True experiments, which involve the actual manipulation of the treatment and/or practice, are rare. In a state-sponsored experiment in Tennessee, starting in kindergarten, students were randomly assigned to varying sized classes (with and without a teacher's aide). The results showed that students, especially minority students, benefited academically from small classes (thirteen to seventeen students) and that the benefits persisted even when the students later moved to larger classes (Finn and Achilles 1990). Prior nonexperimental analyses had shown, across the

range of class size in existing schools, that class size had very little or no effect.

Now that it is accepted that schools can make a difference in learning despite the great significance of family-based factors, the research agenda probably will focus on specifying the conditions in which particular school practices are most effective. This will involve analyzing inside-school practices as well as the links between families and schools and between schools and the workplace.

REFERENCES

- Alexander, Karl, and Doris Entwistle 1995 "Schools and Children at Risk." In Allan Booth and Judith Dunn, eds., *Family-School Links: How Do They Affect Educational Outcomes?* Hillsdale, N.J.: Erlbaum.
- Aronowitz, Stanley, and Henry Giroux 1985 *Education under Siege: The Conservative, Liberal and Radical Debate over Schooling*. South Hadley, Mass.: Bergin and Garvey.
- Baker, David 1994 "In Comparative Isolation: Why Comparative Research Has So Little Influence on American Sociology of Education." *Research in Sociology of Education and Socialization* 10:53-70.
- Benavot, Aaron 1992 "Educational Expansion and Economic Growth in the Modern World, 1913-1985." In Bruce Fuller and Richard Rubinson, eds., *The Political Construction of Education*. New York: Praeger.
- Blau, Peter, and Otis D. Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Bourdieu, Pierre, and Jean-Claude Passeron 1977 *Reproduction in Society, Culture, and Education*. Beverly Hills, Calif.: Sage.
- Bowles, Samuel, and Herbert Gintis 1976 *Schooling in Capitalist America*. New York: Basic Books.
- Bryk, Anthony, Valerie Lee, and Peter Holland 1993 *Catholic Schools and the Common Good*. Cambridge, Mass.: Harvard University Press.
- Carnoy, Martin, and Henry Levin 1985 *Schooling and Work in the Democratic State*. Stanford, Calif.: Stanford University Press.
- Clark, Burton 1962 *Educating the Expert Society*. San Francisco: Chandler.
- 1968. "The Study of Educational Systems." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- Coleman, James 1988 "Social Capital and the Creation of Human Capital." *American Journal of Sociology* 94(Supplement):S95-S120.

- Ernst Campbell, Carol Hobson, James McPartland, Alexander Mood, Frederick Weinfeld, and Robert York 1966 *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.
- Collins, Randall 1979 *The Credential Society*. New York: Academic Press.
- Davies, Scott 1995 "Leaps of Faith: Shifting Currents in Critical Sociology of Education." *American Journal of Sociology* 100(6):1448–1478.
- Dreeben, Robert 1968 *On What Is Learned in School*. Reading, Mass.: Addison-Wesley.
- 1994 "The Sociology of Education: Its Development in the United States." *Research in Sociology of Education and Socialization* 10:7–52.
- Durkheim, Émile 1977 *The Evolution of Educational Thought*, trans. Peter Collins. London: Routledge and Kegan Paul.
- Entwisle, Doris, and Karl Alexander 1992 "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School." *American Sociological Review* 59:446–460.
- Farkas, George 1996 *Human Capital or Cultural Capital?* New York: Aldine de Gruyter.
- Featherman, David, and Robert Hauser 1978 *Opportunity and Change*. New York: Academic Press.
- Ferguson, Ronald 1998 "Teachers' Perceptions and Expectations and the Black-White Test Score Gap." In Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press.
- Finn, Jeremy and Charles Achilles 1990 "Answers and Questions about Class Size: A Statewide Experiment." *American Educational Research Journal* 27:557–577.
- Fischer, Claude, Michael Hout, Martin Sanchez Jankowski, Samuel Lucas, Ann Swidler, and Kim Voss 1996 *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, N.J.: Princeton University Press.
- Gamoran, Adam 1992 "The Variable Effects of High School Tracking." *American Sociological Review* 57:812–828.
- 1995 "Effects of Schooling on Children and Families." In Allan Booth and Judith Dunn, eds., *Family-School Links: How Do They Affect Educational Outcomes*. Hillsdale, N.J.: Erlbaum.
- Hout, Michael 1998 "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358–1400.
- Hunter, John 1986 "Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance." *Journal of Vocational Behavior* 29(3):340–362.
- Hurn, Christopher 1993 *The Limits and Possibilities of Schooling*, 3rd ed. Boston: Allyn and Bacon.
- Jencks, Christopher, and Meredith Phillips, eds. 1998 *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press.
- et al. 1979 *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic Books.
- Kingston, Paul 1986 "Theory at Risk: Accounting for the Excellence Movement." *Sociological Forum* 1:632–656.
- and James Clawson 1990 "Getting on the Fast Track: Recruitment at an Elite Business School." In Paul Kingston and Lionel Lewis, eds., *The High Status Track: Studies of Elite Schools and Stratification*. Albany, N.Y.: State University of New York Press.
- Lee, Valerie, and Robert Croninger 1994 "The Relative Importance of Home and School for Middle-Grade Students." *American Journal of Education* 102:286–329.
- Meyer, John 1977 "The Effects of Education as an Institution." *American Journal of Sociology* 83:55–77.
- , and Michael Hannan, eds. 1979 *National and Political Change, 1950–1970*. Chicago: University of Chicago Press.
- Ogbu, John 1978 *Minority Education and Caste: The American System in Cross-Cultural Perspective*. New York: Academic Press.
- Pallas, Aaron, Doris Entwisle, Karl Alexander, and M. Francis Stluka 1994 "Ability Group Effects: Instructional, Social or Institutional?" *Sociology of Education* 67:27–46.
- Parsons, Talcott 1959 "The School Class as a Social System." *Harvard Educational Review* 29:297–308.
- Phillips, Meredith, Jeanne Brooks-Gunn, Greg Duncan, Pamela Klebanov, and Jonathan Crane 1998 "Family Background, Parenting Practices, and the Black-White Test Score Gap." In Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press.
- Rosenthal, Robert, and Lenore Jacobson 1968 *Pygmalion in the Classroom*. New York: Holt, Rinehart, and Winston.
- Rubinson, Richard, and Bruce Fuller 1992 "Specifying the Effects of Education on National Economic Growth" In B. Fuller and R. Rubinson, eds., *The Political Construction of Education*. New York: Praeger.
- Schultz, Theodore 1961 "Investment in Human Capital." *American Economic Review* 51:1–17.
- Shavit, Yossi and Hans-Peter Blossfeld, eds., 1993 *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview Press.
- Sorokin, Pitirim 1927 *Social Mobility*. Glencoe, Ill.: Free Press.

Tyack, David, and Larry Cuban 1995 *Tinkering Toward Utopia: A Century of Public School Reform*. Cambridge, Mass.: Harvard University Press.

Waller, Willard 1932 *The Sociology of Teaching*. New York: Wiley.

Walters, Pamela, and Richard Rubinson 1983 "Educational Expansion and Economic Output in the United States." *American Sociological Review* 48:480-493.

Willis, Paul 1981 *Learning to Labor*. New York: Columbia University Press.

PAUL W. KINGSTON

## SOCIOLOGY OF ISLAM

Like Christianity and Judaism, Islam is an Abrahamic religion based on prophecy, prophethood, and the revealed text. It began in sixth-century Arabia and spread rapidly to regions outside the Arabian peninsula. A hundred years after Mohammed had declared it a prophetic religion, Islam had spread to almost all the regions of the known civilized world. This early political success and the idea that the divine message for the proper ordering of society is complete and final account for the social pervasiveness of this religion. The first factor inhibits the handing over of spheres of life to nonreligious authority, and the second makes it difficult to offer rival versions of the blueprint. This social pervasiveness makes Islam especially interesting in the sociology of religion (Gellner 1983, p.2).

Islam is the second largest religion, with an estimated 1.2 billion adherents, constituting about 20 percent of the world population in 1998. Approximately 900 million Muslims live in forty-five Muslim-majority countries. Table 1 provides a sociodemographic profile of Muslim countries included in the *World Development Report* published annually by the World Bank. In terms of size, the Islamic world constitutes a significant part of humanity and therefore warrants a sociologically informed understanding and analysis of its religious, social, and political trends. The following topics will be covered in this article: social, ideological, and economic factors in the origins of Islam; Islam and the rise of the modern West; Islam, Muslim society, and social theory; Islam and fundamentalism; the Islamic state; gender issues

in Muslim societies; and, Muslim minorities in the West.

### SOCIAL FACTORS IN THE ORIGINS OF ISLAM

Social science scholarship in the twentieth century has been influenced by three dominant intellectual traditions: Marxism, Weberian, and functionalism. Their influence has shaped the analytic approach to historical events, resulting in an increasing focus on the relationship between social and economic factors and historical events. The study of Islam and Muslim societies often reflects these influences.

One strand of scholarship has focused on the analysis of various factors in the origins and early development of Islam. A discussion of the economic and social aspects of the origins of Islam provides a test case for a closer investigation of the wider issues raised by the dominant paradigms in sociology. A number of historical studies have dealt with this issue primarily in terms of the diffusion of Jewish and Christian teaching in pre-Islamic Arabia that laid the foundation for the rise of Islam (Torrey 1933; Bell 1926; Kroeber 1948). The aim of these and similar studies has been to identify and understand how certain ideas and cultural elements utilized by Islam derived from preexisting religions or to point to the existence of elements analogous to Islam in other religious traditions in the same general area.

Another scholarly tradition has approached the analysis of the early development of Islam in terms of sociological and anthropological concepts and traces the origins of Islam primarily to the change in social organization in pre-Islamic Meccan society caused by the spread of trade. Wolf (1951) provides an overview of these studies and shows that the tendencies Mohammed brought to fruition were prominent in pre-Islamic Arabia. The spread of commerce and rapid urban development had caused the emergence of classlike groupings from the preceding network of kin relations. This also contributed to the emergence of a divine being specifically linked to the regulation of nonkin relations as the chief deity. These changes created a disjunction between the ideological basis of social organization and the functional social reality and thus spawned disruption and conflict. Islam arose as a moderating religious-

**Sociodemographic Profile of Selected Muslim Countries**

COUNTRY	Population Millions (1997)	Urban Population (% of total)	GNP Per Capita 1997 \$	Life Expectancy (males and females)	Adult Illiteracy Rate (15 years and (males and females)
Indonesia	200	37	1110	63/67	10/22
Pakistan	137	35	490	62/65	50/76
Bangladesh	124	19	270	57/59	51/74
Nigeria	118	41	260	51/55	33/53
Turkey	64	72	3130	62/65	8/28
Iran	63	60	2190	67/68	25/44
Egypt	60	45	1180	64/67	36/61
Sudan	27	25	125(e)	52/55	45/68
Algeria	29	57	1490	68/72	26/51
Morocco	28	53	1250	64/68	41/53
Uzbekistan	24	42	1010	66/72	—
Afghanistan	22	—	—	43/44	55/86
Malaysia	21	55	4680	70/74	11/22
Saudi Arabia	20	84	6790	66/71	29/50
Yemen	16	35	270	54/54	—
Kazakhstan	16	60	1340	60/70	—
Syria	15	53	1150	66/71	14/44
Mali	10	28	260	48/52	61/77
Tunisia	9	63	2090	68/71	21/45
Niger	10	19	200	44/49	79/93
Senegal	9	45	550	49/52	57/77
Guinea	7	31	570	46/47	50/78
Libya	5	—	5100	62/65	14/41
Jordan	4	73	1570	69/72	7/21
Lebanon	4	88	3350	68/71	10/20
Mauritania	2	54	450	52/55	50/74
United Arab Emirates	3	85	17360	74/76	21/20
Oman	2	79	4950	69/73	—
Kuwait	2	—	19420	76/76	20/27
Albania	3	38	750	69/75	—

Table 1

SOURCE: World Bank: *World Development Report 1998/99 and 1997*. New York, Oxford University Press. UNDP, *Human Development Report 1996*. New York; Oxford University Press.

ethical social movement under these social conditions. According to Wolf:

*The religious revolution associated with the name of Mohammed permitted the establishment of an incipient state structure. It replaced allegiance to the kinship unit with allegiance to a state structure, an allegiance phrased in religious terms. It limited the disruptive exercise of kin-based mechanisms of blood feud. It put an end to the extension of ritual kin ties to serve as links between tribes. It based itself instead on the armed force of the faithful as the core of a social order which included both believers and unbelievers. It evolved a rudimentary judicial authority, patterned after the role of the pre-Islamic soothsayer, but possessed of new significance. The limitation of the blood feud permitted war to emerge as a special prerogative of the state power. The state taxed both Muslims and non-Muslims, in ways patterned after pre-Islamic models but to new ends. Finally, it located the center of the state in urban settlements, surrounding the town with a set of religious symbols that served functionally to increase its prestige and role.* (1951, pp. 352–353)

In his historical studies of early Islam, Watt (1954, 1955, 1962a, 1962b) also analyzed the economic, social, and ideological aspects of the origins of Islam. His analysis of the economic situation in pre-Islamic Arabia shows that the economic transition from a nomadic to a mercantile economy had resulted in social upheaval and general malaise. He also found a close affinity between the ideology of Islam and the situation that prevailed in early seventh-century Mecca. However, his analysis led him to question the nature and direction of the relationship between Islamic doctrines and the social and economic conditions of pre-Islamic Meccan society. Are doctrines causally dependent on the social order in such a way that they can be deduced from it? Or is the ideology of Islam a creative factor that made a contribution to the course of events? Watt argues that there was nothing inevitable about the development of a world religion from the economic and social circumstances of early seventh-century Mecca. The malaise of the times might have been alleviated without achieving anything of more than transient and local importance. He argues that the formulation of Islamic ideology was a creative response to the

situation, not an automatic result of interacting factors.

According to Watt, the creative response of Islamic ideology is reflected in key foundational Koranic ideas such as *Ummah* and *Rasul*. Like other Koranic ideas, these ideas can be connected to earlier Jewish and Christian conceptions as well as to pre-Islamic Arabian ideas, but the Koranic conceptions had a unique new and creative dimension that made them especially relevant to the contemporary Arabian situation. Mere repetitions of current ideas in the Koran would have rendered those ideas devoid of creative novelty, whereas sheer novelty would have made them unintelligible. What the Koran does is take the familiar conceptions and transmute them into something new and original (Watt 1954, p. 172). In this synthesis, the old images are to some extent transformed but retain their power to release the energy of the human psyche. From this perspective, the Koranic conceptions and images of *Ummah* and *Rasul* took on new meanings that were a combination of the old conceptions and additional meanings conferred by the Koran, which was thus able to release the energies of the older images and inaugurate a vigorous new religion. This energy was directed, among other things, toward the establishment of the Islamic state and the unification of Arabia (Watt 1954, pp. 173–4).

Debate about the social factors in the origins of Islam continues (Engineer 1990; Crone 1996). However, it is evident that under the influence of dominant theoretical paradigms in sociology, this debate has provided new insights into the role of social, economic, and cultural factors in shaping the ideology of Islam and the early development of Islamic social formations.

### THE “SOCIAL PROJECT” OF ISLAM

Another recent development has been a revival of interest in the “social project” of Islam. The most significant contributions have come from the work of Rahman (1982, 1989), who claims, “A central aim of the Koran is to establish a viable social order on earth that will be just and ethically based” (1989, p. 37). This aim was declared against the backdrop of an Arabian society characterized by polytheism, exploitation of the poor, general neglect of social responsibility, degradation of morals,

injustice toward women and the less powerful, and tribalism. The Koran and the genesis of the Muslim community occurred in the light of history and against the social historical background. The Koranic response to specific conditions is the product of a "coherent philosophy" and "attitude toward life" that Rahman calls "the intellectual tradition" of Islam. This tradition was subverted and undermined by an emphasis on literalist interpretations of the Koran by Ulema Islamic scholars. The Islamic scholarship molded by Ulema came to emphasize "minimal Islam", focusing on the "five pillars," and negative and punitive Islam. Islamic scholarship thus became rigid, fossilized, and largely removed from the intellectual tradition of the Koran. Rahman argues that the intellectual tradition of the Koran requires that Koranic thought be dependent on a factual and proper study of social conditions in order to develop Islamic social norms for reforming society (Rahman 1982).

#### ISLAM AND THE RISE OF THE MODERN WEST

An important strand of historical scholarship has focused on the relationship between Islam and the rise of the modern West. This question was the focus of *Mohammed and Charlemagne* (Pirenne 1939). According to Pirenne, for centuries after the political collapse of the Roman Empire, the economic and social life of western Europe continued to move exclusively to the rhythm of the ancient world. The civilization of Romania had long outlived the Roman Empire in the West. It survived because the economic life based on the Mediterranean had continued to thrive. It was only after the Arab-Muslim conquests of the eastern and southern Mediterranean in the seventh century A.D. that this Mediterranean-wide economy was disrupted by the Islamic conquest. The Arab-Muslim war fleets closed the Mediterranean to shipping in the later seventh century.

Deprived of its Mediterranean-wide horizons, civilized western Europe closed in on itself, and the under-Romanized world of northern Gaul and Germany gained prominence. The Mediterranean Roman Empire in the West was replaced by a western Europe dominated by a northern Frankish aristocracy that gave rise to a society in which wealth was restricted to land. Its rulers, deprived of the wealth generated by trade, had to reward

their followers with grants of land, and thus feudalism was born. The empire of Charlemagne, a northern Germanic empire inconceivable in any previous century, marked the beginning of the Middle Ages. Pirenne shows that by breaking the unity of the Mediterranean, the conquest made by the Arab-Muslim war fleets ruptured Romano-Byzantine economic and cultural domination over western Europe, which was forced to rely on its own material and cultural resources. From this analysis Pirenne draws his famous observation: "It is therefore strictly correct to say that without Mohammed Charlemagne would have been inconceivable" (Pirenne 1939, p. 234).

The Pirenne thesis linked great historical events that have occupied the attention of historians for a long time: the demise of the classical world centered on the Mediterranean and the rise of the empire of Charlemagne. Pirenne demonstrated that these two events, which are central to the rise of the modern West, are linked to the rise of Islam and its expansion to the Mediterranean. Pirenne's well-documented generalizations have attracted praise as well as criticism from historians who are often wary of broad generalizations (see Hodges and Whitehouse 1983).

While Pirenne's thesis attempts to link the rise and development of Islam to the rise of the modern West, paradoxically, equally influential hypothesis postulates instead a "clash of civilizations." This hypothesis, advanced by Samuel Huntington (1993), holds that whereas in the pre-Cold War era military and political conflicts occurred within the Western civilizations, after the end of the Cold War the conflict moved out of its Western phase and its centerpiece became the interaction between the West and non-Western civilizations and among non-Western civilizations. According to Huntington, future conflicts will occur along the fault lines that separate those civilizations. Globalization tends to heighten civilizational identity, and as a result, civilizational differences are difficult to reconcile and override political and economic factors.

Huntington (1993) postulates that the greatest threat of conflict for the West comes from religious fundamentalism, especially Islamic fundamentalism. He sees Islamic fundamentalism as arising from the failures of Muslim countries to achieve political and economic development of

their masses. This failure is exacerbated by the demographic structure of the Muslim world, especially the large bulge in the middle of the age pyramid (the youth). Huntington suggests that Muslim countries have a historical propensity toward violence. The domination and hegemony of the West, he claims, will force an alliance between the Confucianist and Islamic civilizations, and that alliance will challenge Western interests, values, and power, resulting in a civilizational clash. Huntington postulates that civilizational conflict will replace ideological and other forms of conflicts in the future. The outcome of this change is that civilizational conflicts will become more intense, violent, and sustained. This thesis was criticized as a new form of Orientalism. Other criticisms have centered on Huntington's assumption of civilizational unity as well as his assumptions about the basis of alliances between Confucianist and Islamic civilizations (Ajami 1993; Ahluwalia and Mayer 1994).

#### ISLAM, MUSLIM SOCIETY, AND SOCIAL THEORY

**Ibn Khaldun and the political sociology of Muslim society.** The sociology of Islam primarily refers to the empirical study of Muslim societies. In this respect, it has occupied an important place in the theoretical discourse of a number of theorists from Ibn Khaldun and Weber to Gellner. It is beyond the scope of this article to provide an exhaustive overview of how Muslim society and Islam have been treated in social theory (for studies of the Islamic revolution in Iran, see Shariati 1979; Fischer 1980; Arjonaud 1988). This section will provide a general overview of the subject in the works of four social theorists: Ibn Khaldun, Weber, Gellner, and Geertz.

Ibn Khaldun, an Arab historian and sociologist (1332–1406), is perhaps the most notable theorist of Muslim society. In the prolegomena (introduction) to his monumental work on universal history, he conceived and formulated the most comprehensive synthesis in the human sciences ever achieved by a Muslim thinker. In the prolegomena, among other topics, he probably provided the first modern outline of sociological principles. He defined sociology as “the study of human society in its different forms, the nature and characteristics of each of these forms, and the

laws governing its development” (Khaldun 1992, p. 7). The basic sociological principles he enunciates are as follows:

1. Social phenomena seem to obey laws that while not as absolute as those governing natural phenomena, are sufficiently constant to cause social events to follow regular, well-defined patterns and sequences.
2. These laws operate on masses and cannot be influenced significantly by isolated individuals.
3. Sociological laws can be discovered only by gathering many facts and observing circumstances and sequences through historical records and the observation of present events.
4. Societies are not static. Social forms change and evolve as a result of contact and interaction between different people and classes, population changes, and economic inequality.
5. Sociological laws are not a reflection only of biological impulses or physical factors but also of social forces.

He then applied these principles to the analysis of Muslim societies (Khaldun 1992, p. 8–9).

The core of Ibn Khaldun's sociology is his concept of *Asabiyya* (social solidarity). For Khaldun, society is natural and necessary, since isolated individuals can neither defend themselves against powerful enemies nor satisfy their economic wants. However, individual aggressiveness would make social life impossible unless it was curbed by some sanction. This sanction may be provided by a powerful individual imposing his will on the rest or by social solidarity. The need for a common authority generates the state, which is to society as form is to matter and is inseparable from it. Ibn Khaldun traces the origin of social solidarity to blood and kinship ties. Nevertheless, social solidarity is shaped by the nature and character of social organization. In this lies the genius of his theory of Muslim social formations and circulation of the elite.

The nature of tribal life generates the strongest form of social solidarity and social cohesion, producing social, political, and civic virtues that



characterize tribespeople. For Khaldun, leadership exists only through superiority, and superiority only through group feeling. Domination and authority are the rewards for social cohesion. Only those with superior social cohesion succeed in becoming rulers, but a civilization (state-society) consists of tribes and cities. The division of labor is the essence of urban life. It is the key to cities' capacity to supply economic and cultural services that tribespeople are unable to provide for themselves because the tribal ethos spurns specialization. Civilization needs cities to provide economic wealth, which is achieved through specialization and a complex division of labor. Specialization, however, is inherently incompatible with social cohesion and the martial spirit. There emerges a need to provide a new basis for social bonds, and religion becomes the most powerful force in holding together a sedentary people. Scripturalistic and puritanical religion has as a natural affinity with urban life. The combination of religious and tribal solidarity is formidable, and to it Ibn Khaldun attributes the rapid and sweeping conquests of the Muslim Arabs in the seventh century.

The dialectic between the tribe and the city forms the basis of the model of circulation of the elite in society. The Khaldunian model rests on the distinction and contrast between the tribe and the city. Zubiada (1995) has provided a succinct summary of this model:

*Dynasties which have conquered the city and its wealth do so with the militant vigor of their nomadic stock, and the solidarity (asabiyya) of their kinship bonds. In time, the rulers become settled and accustomed to the comforts and luxuries of the city, the branches of their kin develop factional interests and competition over wealth and power which saps solidarity. The cost of their expanding retinue and luxury spending leads to an intensification of the taxation burden on the urban populations and their growing discontent. The growing weakness of the rulers encourages aspiring tribal dynasties, lusting for the city, to organize military campaigns which ultimately topple the rulers and replace them, only to repeat the cycle. (1995, p. 154; see also Turner 1999)*

Ibn Khaldun's sociological generalizations about the Muslim social formations of his time can be summarized in the following statements:

1. Nomadic tribes conquer sedentary societies because of their greater cohesiveness.
2. The combination of tribal solidarity and a puritanical scripturalistic urban religion is overwhelming.
3. Conquest tends to be followed by luxury and softening, which lead to decay and annihilation of the ruling dynasty.

These three statements describe the rise and fall of many historical Muslim social formations in the Middle East and North Africa.

**Weber, Islam, and Capitalism.** Weber's theoretical interest in and interpretation of Islam is related to his exploration of the affinity of faith and modern socioeconomic organizations. Through a comparative study of world religions, Weber formulated *The Protestant Ethic and Spirit of Capitalism* (Weber 1958). In his analysis, Weber demonstrated an elective affinity between certain types of religious ideas and particular types of economic activity. He hypothesized a nexus between Protestant religious beliefs and the development of modern capitalism and used his study of comparative religion to show why modern capitalism could not have emerged in other societies, including Islamic society. Weber saw Islam as a prophetic, this-worldly, salvationist religion with strong connections with other Abrahamic religions and regarded it as a useful test case of his thesis.

Weber argued that rational formal law, autonomous cities, an independent bourgeois class, and political stability were totally absent in Islamic society because of prebendal feudalism and the domination of patrimonial bureaucracy. He also argued that a hedonistic spirit and an accommodating Koranic ethic could not produce salvation anxiety and that asceticism was blocked by two important social groups: the warrior group that was the social carrier of Islam and the Sufi brotherhoods that developed mystical religiosity.

Weber's characterization of Islam has been criticized as "factually wrong" (Turner 1974b, p. 238). Gellner (1983) describes Weber's notion about the affinity between the bourgeoisie style of life and religious sobriety and asceticism as "a piece of Judaeo-Protestant ethnocentrism" (Gellner 1983, p. 78). Gellner also challenges Weber's contention that the institutional preconditions of modern capitalism were not restricted to the West but

that it was the ideological element (i.e., the Protestant ethic) that provides the crucial differentia, the extra spark that, in conjunction with the required structural preconditions, explains the miracle. According to Gellner, "the differentiae of Islam seem institutional rather than ideological. Ideological parallels to Christianity can be found, but they operate in a contrasted institutional milieu" (1983, p. 6).

**Gellner's Theory of Muslim Society.** Gellner made some of the most significant contributions to the sociology of Islam over the past three decades. Building on David Hume, Ibn Khaldun, Marshal Hodgson, and others, he provides a model of Muslim society that aspires to a general interpretation of all past and present Muslim societies. In *Muslim Society* (1983) and other writings (Gellner 1969, 1992, 1994), Gellner identifies unvarying features of Muslim societies that make them susceptible to sociological analysis. Building on the work of Ibn Khaldun, he postulates a dialectic between city and tribe, each with its own form of religion. The central and perhaps most important feature of Islam, according to Gellner, is that it was internally divided into the high Islam of scholars and the folk (low) Islam of the people. High Islam is primarily urban, and folk Islam is primarily tribal and rural. Although the boundaries between the two were not sharp but gradual and ambiguous, they nevertheless projected a distinctive tradition.

High Islam is carried by urban scholars recruited largely from the trading bourgeois classes and reflecting the natural tastes and values of urban middle classes. Those values include order, rule observance, sobriety, and learning, along with an aversion to superstition, hysteria, and emotional excess. High Islam stresses the severely monotheistic and nomocratic nature of Islam, is mindful of the prohibition of claims of mediation between God and the individual, and generally is oriented toward puritanism and scripturalism. Folk Islam is superstitious and mediationist. It stresses magic more than learning and ecstasy more than rule observance. Rustics encounter writing mainly in the form of amulets and manipulative magic. Far from avoiding mediation, folk Islam is centered on it. Its most characteristic institution is the saint cult, in which the saint is more often living rather than dead. This form of faith generally is known in the literature as religious brotherhoods

or Sufi orders. Urban religion is Weberian (textual and puritanical), and rural and tribal religion is Durkheimian.

Each religious tradition has a place in the social structure. Saint cults are prominent in the tribal or rural countryside and provide invaluable services in rural conditions: mediating between groups, facilitating trade and exchanges, and providing symbolism that allows illiterate rustics believers to identify enthusiastically with a scriptural religion. The folk Islamic tradition, through its ecstatic rituals, provides the poor with an escape from their miserable conditions. High Islam provides the urban population, and to some extent the whole society, with its charter and constitution entrenched by the sacred texts, which can mobilize resistance against an unjust state. The two systems often coexisted in an amiable symbiosis, but a tension remained that would surface from time to time in the form of a puritan revivalist movement to transform folk Islam in the image of high Islam. Gellner argues that in the traditional order Islam may be described as a permanent or recurrent, but ever-reversed, Reformation. In each cycle, the revivalist puritan impulse would in the end yield to the contrary social requirements (Gellner 1994).

Under modern conditions, the pattern of interaction between the two religious traditions has been transformed. The centralization of political power and the ability of the state to rule effectively with modern technology and control over the military and the economy have undermined the social basis of folk Islam. Puritanism and scripturalism have become symbols of urban sophistication and modernity. According to Gellner, this constitutes the basic mechanism of the massive transfer of loyalty from folk Islam to a scripturalist, fundamentalist variant of Islam: "This is the essence of the cultural history of Islam of the last hundred years. What had once been a minority accomplishment or privilege, a form of the faith practised by a cultural elite, has come to define society as a whole" (Gellner 1994, p. 22).

In short, conditions of modernity (mass literacy, urbanization, modern education, and technology) have reinforced the power of scripturalist, puritanical urban Islam and its challenge to secular power; this explains the current rise of Islamic revivalist and fundamentalist movements. The validity of Gellner's model of Muslim society has

been challenged in the historical and the modern contexts. It has been criticized for ignoring the different meanings and roles of concepts and entities such as Ulema in different historical contexts and in different societies and instead treating them as sociological or political constants. Modern Islamism, critics argue, is a political ideology and is distinct from anything in Muslim history, which in recent years has become a dominant idiom for the expression of various and sometimes contradictory interests, aspirations, and frustrations (Zubiada 1995). However, even his critics agree that Gellner's model of Muslim society is the most ambitious attempt in modern sociology to identify the internal religious dynamics that play a significant and in certain conditions, critical role in determining the political character and socioreligious trajectories of Muslim societies.

**Geertz and the Islamization process.** Like Gellner, Geertz has made significant contributions to the sociology of Islam through his anthropological studies, in this case of religious life in Indonesia and Morocco. His work illustrates the modes of incorporation of Islam into already existing and well-developed cultures and shows how those incorporations manifest themselves in the different Islamic traditions that over time come to characterize them. Geertz shows that in the socio-cultural and ecological setting of Morocco, the "cultural center" of Islam was developed not in the great cities but in the mobile, aggressive, fluid, and fragmented world of tribes on the periphery. It was out of the tribes that the forming impulses of Islamic civilization in Morocco came and stamped their mentality on future developments. "Islam in Barbary was—and, to a fair extent still is, basically the Islam of saint worship and moral severity, magical power and aggressive piety, and this for all practical purposes is as true in the alleys of Fez and Marrackech as in the expanses of the Atlas or the Sahara" (Geertz 1968, p. 9).

In the tropical heartland of Indonesia, with its productive peasant society and Indic cultural heritage, once Islam was incorporated, it found a distinctive cultural and religious expression. In Indonesia, Islam did not construct a civilization but appropriated it. The Javanese social structure was shaped by a centralized state and a productive and industrious peasantry. The social structure was highly differentiated and developed, and when Islam came, its expression was influenced pro-

foundly by the context. The Indonesian Islamic tradition was malleable, tentative, syncretic, and multivocal. In Morocco and other Middle Eastern societies, Islam was a powerful force for cultural homogeneity, moral consensus and standardization of fundamental beliefs and values. In Indonesia, Islam was a powerful force for cultural diversification and sharply variant and even incompatible worldviews and values.

The gentry, which was acculturated to Indic ritualism and pantheism, developed a subjectivist and illuminationist approach to Islam. The peasantry absorbed Islamic concepts and practices into its folk religion and developed a distinctive contemplative tradition. The trading classes were exposed to Arabian Islam and, because of their greater exposure to the Meccan pilgrimage, cultivated a doctrinal religious tradition. Islam in Indonesia therefore developed as a syncretic and multivocal religious tradition whose expression differed from one sector of the society to another (Geertz 1960, 1968).

Geertz's work, like Gellner's, provides a framework for explaining the diversity of religious traditions in Muslim societies and indeed the existence of religious diversity in all religions. As Geertz observes, "Religious faith, even when it is fed from a common source, is as much a particularising force as a generalizing one, and indeed whatever universality a given religious tradition manages to attain arises from its ability to engage a widening set of individual, even idiosyncratic, conceptions of life and yet somehow sustain and elaborate them all" (Geertz 1968, p. 14). The purpose of this account is to illustrate that Islam occupies an important place in theoretical discourse on modern sociology. As empirical and comparative study of Muslim societies develops, it will provide more opportunities to test and refine some of the existing theoretical propositions as well as develop new ones (see Arjomand 1988; Fischer 1980; Beyer 1994; Irfani 1983).

**Islam and Fundamentalism.** Fundamentalism emerged in all the major world religions in the last quarter of the twentieth century and gained prominence and influence in the 1990s (Marty and Appleby 1991, 1992, 1993). It is defined as "a distinctive tendency—a habit of mind and a pattern of behaviour—found within modern religious communities and embodied in certain rep-

representative individuals and movements. Fundamentalism is, in other words, a religious way of being that manifests itself as a strategy by which beleaguered believers attempt to preserve their distinctive identity as a people or group" (Martin and Appleby 1992, p. 34). Feeling that this identity is at risk, fundamentalists try to fortify it by means of a selective retrieval of doctrines, beliefs, and practices from a sacred past as well as modern times. This renewed religious identity becomes the exclusive and absolute basis for a re-created political and social order. While there are differences between fundamentalist movements in general, their endeavor to establish a "new" political and social order always relies on charismatic and authoritarian leadership. These movements also feature a disciplined inner core of elites and organizations as well as a large population of sympathizers who may be called on in times of need. Fundamentalists often follow a rigorous sociomoral code and have clear strategies to achieve their goals.

Religious fundamentalism is a growing and important part of social change in Muslim countries. Its main goal is to establish the Sharia (Islamic law) as the explicit, comprehensive, and exclusive legal basis of society (Marty and Appleby 1991, 1992; Beinin and Stork 1997; Esposito 1983). Hardly a day passes without a reference to Islamic fundamentalism in the international media. All Muslim societies are affected by it, although there are large differences among them in terms of its presence and power. Is Islamic fundamentalism the inevitable destiny of all Muslim countries, or is it only a part of larger process of social change? Are there certain social, economic, historical, and other preconditions that predispose some Muslim countries more than others to Islamic fundamentalism? Are there different types of Islamic fundamentalism? These and related questions have been posed and explored by several contributors to the Fundamentalism Project of the American Academy of Arts and Sciences (Marty and Appleby 1991). There are three competing theories of Islamic fundamentalism: Watt's (1988) "crisis of self-image," Gellner's (1983) "pattern of distribution of dominant religious traditions," and the "modernization and religious purification" theory advanced by a number of social scientists (Tamney 1980; Hassan 1985; Yap 1980; Rahman 1982).

**Crisis of Self-Image.** Distilling insights from his works on the history and sociology of Islam,

Watt (1988) has proposed that the principal root of Islamic fundamentalism is the domination of the traditional "Islamic world view" and the corresponding "self-image of Islam" in the thinking of Islamic intellectuals and great masses of ordinary Muslims. According to Watt,

*the important distinction is between those Muslims who fully accept the traditional world view and want to maintain it intact and those who see that it needs to be corrected in some respects. The former group are fundamentalists . . . while the latter group will be referred to as Liberals.* (1988, p. 2)

Among both groups, many different political movements and attitudes can be found. The Ulema (religious scholars), who are the primary bearers and transmitters of the traditional worldview, are mostly reactionary in the sense that they tend to oppose reforms. Other Islamic intellectuals subscribe to a variety of reformist elements and sometimes are very critical of the Ulema, but the reforms they are interested in are mostly social and political and leave the traditional worldview of Islam unchanged. Watt then identifies important aspects of the traditional worldview: (1) the unchanging static world that is predicated on the complete absence of the idea of development, (2) the finality of Islam, (3) the self-sufficiency of Islam (Watt sees this reflected in the Muslim's conception of knowledge; when a Muslim thinks of knowledge, it is primarily "knowledge for living," whereas when a Westerner thinks of knowledge, it is mainly "knowledge for power"), (4) Islam in history (the widespread belief that Islam will ultimately be triumphant in changing the whole world into *dar-al-Islam* (the sphere of Islam), and (5) the idealization of Muhammed and early Islam, which renders critical and historically objective scholarship highly problematic in the Muslim consciousness and deviation from (1988) idealized and romanticized notions as a heresy and "unthinkable." According to Watt "These features of the Islamic worldview and the corresponding self-image are the basis of Islamic fundamentalism. The support for fundamentalism is embedded in the consciousness, which fully accepts the traditional worldview and wants to maintain it intact."

**Patterns of Distribution of Dominant Religious Tradition.** Building on the sociological and historical analyses of Muslim society of Ibn Khaldun

(1958), Weber (1964), Hume (1976), Hodgson (1975), and others, Gellner has advanced a theory of Muslim social formation that is based on his conceptualization of "two strands of Islam." One strand is characterized by "scripturalist puritanism" and represented by the Ulema. This is the Islam of the "fundamentalists." The other strand is characterized by a "hierarchical ecstatic mediationist style and is represented by the 'Saints.'" These two strands have evolved historically as representing two major social structural features of Muslim society: the city and the countryside. Gellner combines these strands of Islam with the political orientation of the elites and proposes a model of Muslim social formations. If one contrasts fundamentalism with laxity along one dimension and social radicalism with traditionalism along another, according to Gellner, one gets four types of Muslim societies or social formations.

The old-style puritanism prevails in areas where a traditional elite survives but is still fairly close to its origin in an Ibn-Khaldunian swing of the pendulum that brought it to power in a fusion of religious enthusiasm and tribal aggression. The new-style puritanism with its elective affinity for social radicalism prevails in areas where colonialism destroyed old elites and a new one elite came from below rather than from the outer wilderness (Gellner 1983, p. 89). An elaboration of Gellner's typology of Islamic social formations is shown in Figure 1.

**Modernization and Religious Purification.**

This theory holds that religious fundamentalism is one of the consequences of the modernization process. Building on studies by Mol (1972) and Folliet (1955), Tamney (1980) proposed that one way in which modern people are different from traditional people is that they practice purer religious styles. The relationship between modernization and religious purity can take two forms. In its general sense, purification is the opposite of syncretism: It is the elimination of religious elements originating in a traditional religion. Purification means the differentiation of religious traditions at the personality level, so that the individual's religious lifestyle reflects one style of tradition. If being modern means that people are more conscious about the history and the internal structures of various religions, modern people can realize the inconsistencies in a syncretic lifestyle, feel uneasy or even insincere, and seek to purify their lives by

deliberately eliminating elements from religious traditions other than their own. Using this conceptualization, Tamney hypothesizes that modernization is associated with religious purification. His empirical examination of this hypothesis in Indonesia tends to support his theory. Studies by Hassan (1984, 1985a, 1985b) and Irfani (1983) provide some support for this theory.

**Islamic Militancy: A New Paradigm?** Using the current religious, social, and political conditions of Muslim countries as a kind of "natural experiment," the author is conducting a multicountry study to examine the three competing theories of Islamic fundamentalism outlined in the preceding section. Over 4,400 mostly highly educated Muslim respondents have been surveyed. The empirical evidence shows that the heartlands of the Islamic world, from Indonesia to Egypt, are undergoing a religious renaissance. A large majority of the respondents were devoutly religious. If the term "fundamentalism" is defined to mean a high degree of devotional religiosity, these heartlands are becoming fundamentalist (Hassan 1999d). What are the implications of this for Islamic radicalism? Does this mean increasing support for the militant Islamic movements that are agitating to establish their versions of the Islamic state? Would this increase militancy against the groups or countries they regard as enemies of Islam?

Religious devotion appears to be associated with a decline in the support for militant Islamic movements. A large majority of Muslims do not belong to radical Islamic group. In fact, most of the respondents approved of moderate political leaders who are leading political and social movements for democratic and tolerant societies and political cultures. The declining support for radical and militant movements is paradoxically further radicalizing these movements and transforming them into more violent and secretive organizations. The nature and ruthlessness of violence reflect their desire to gain public attention and are symptomatic of their desperation.

The new form of violence is different from the earlier form that was carried out by organizations often with tacit support from political structures. The new militancy appears to be fueled by a sense of desperation and humiliation caused by globalization and the increasing economic, cultural, technological, and military hegemony of the West.

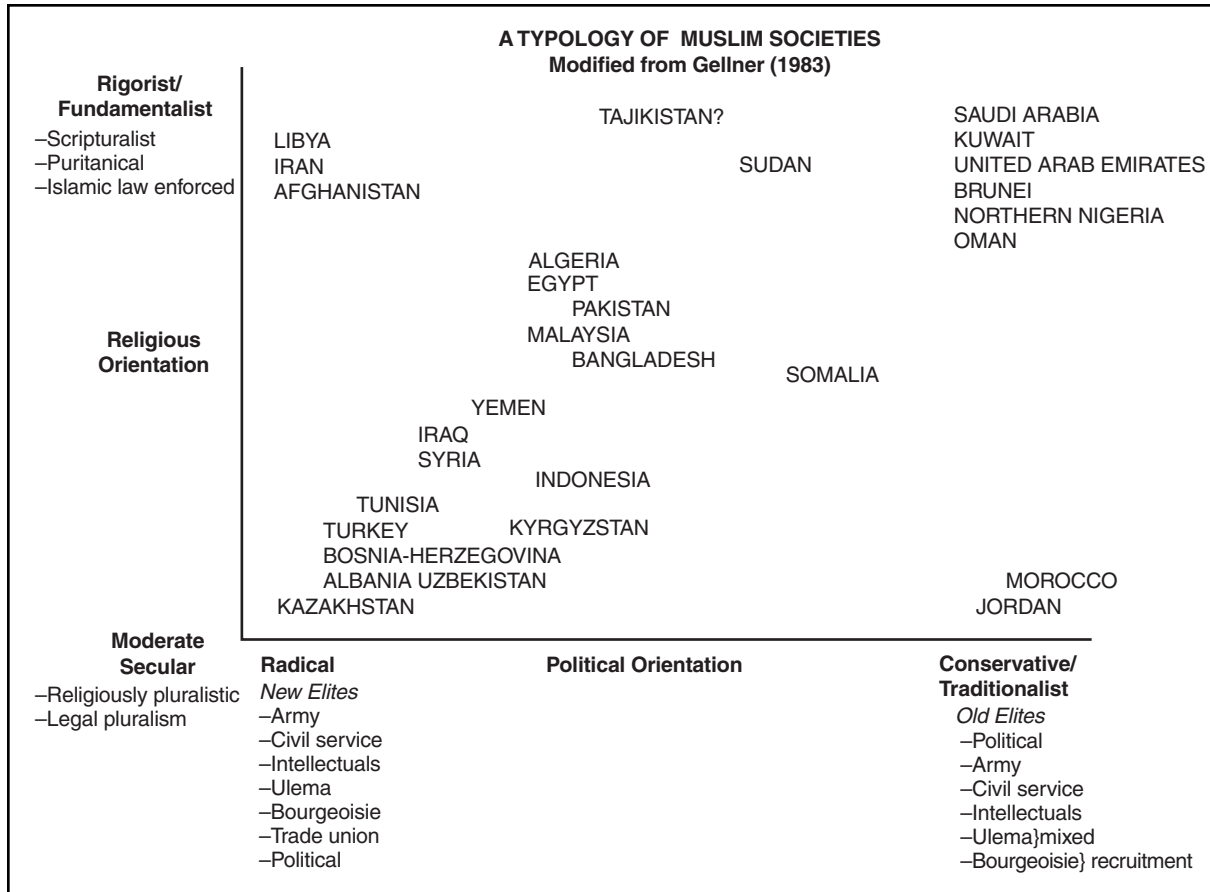


Figure 1

This pattern represents a kind of paradigm shift in the nature, causes, and targets of terrorism carried out by the new militant groups. The old form of militancy attempted to establish the legitimacy of political goals; the new form is guided by religious fanaticism, destruction, and revenge. The old form of militancy identified enemies. The new enemies are ephemeral global conspiracies.

A majority of the respondents regard major Western countries as anti-Islamic. The primary reason for this attitude is not religion, but the perceived indifference and inaction of Western countries toward protecting the Muslim populations of Bosnia-Herzegovina, Palestine, and Chechnya destruction. These views are widely held among the elites. The study provides new insights into the dynamics of the new Islamic militancy. It shows that contrary to the general belief, increasing religiosity in Muslim countries is associated with

political liberalization and diminishing support for militant Islamic groups. The impact of these developments is making the militant movements highly secretive and more violent.

The globalization process is creating a social and cultural hiatus that is affecting the nature and organization of Islamic militancy. The new militancy is not motivated by attitudes toward colonialism and struggles to win the hearts and minds of Muslim populations. Instead, it is fueled by a sense of powerlessness, revenge, and religious fanaticism. The enemy is ephemeral global conspiracies. How Muslim countries and the international community respond to these new developments will have a profound impact on the nature and activities of the new militancy. The solution would require more open and stronger political structures in Muslim countries to legally and politically pursue solutions to the problems posed by the new

militancy. It also will require a change in the attitude that increasing religiosity increases support for militancy, when it actually diminishes support for it.

### THE ISLAMIC STATE

The relationship between politics and religion in Muslim societies has been a focus of debate among scholars of Islam for most of this century. A commonly stated view of many Western and Muslim scholars is that Islam is not only a religion but also a blueprint for social order and therefore encompasses all domains of life, including law and the state (Maududi 1960, Lewis 1993; Huntington 1993; Rahman 1982; Watt 1988; Pipes 1981; Esposito 1995; Weber 1978; Turner 1974a; Gellner 1983). This view is reinforced by the fact that Islam does not have a church institution, although it does have the institutions of the Ulema, who act as the guardians of the interpretations of the sacred texts, and the Iman Masjid (leaders of the mosques), who lead the mandatory daily prayers in mosques. It is further argued that this characterization sets Islamic societies apart from Western societies built on the separation of state and religious institutions.

After reviewing the evidence on the separation of state and religion in Islamic history, Lapidus (1996) concludes that the history of the Muslim world reveals two main institutional configurations. The undifferentiated state-religion configuration characterized a small number of Middle Eastern societies. This configuration was characteristic of lineage or tribal societies. The historical norm for agrourban Islamic societies was an institutional configuration that recognized the division between the state and religious spheres:

*Despite the common statement (and the Muslim ideal) that the institutions of state and religion are unified, and that Islam is a total way of life which defines political as well as social and family matters, most Muslim societies did not conform to this ideal, but were built around separate institutions of state and religion. (Lapidus 1996, p. 24)*

Keddie (1994, p. 463) has described the supposed near identity of religion and the state in Islam more as a "pious myth than reality for most of Islamic history." Similar views of Islamic history

have been advanced by others (Zubiada 1989; Sadowski 1997; Ayubi 1991; Sivan 1985).

Historical scholarship indicates that the institutional configurations of Islamic societies can be classified into two types: (1) "differentiated social formations" (societies in which religion and the state occupy different spaces) and (2) "undifferentiated social formations" (societies in which religion and the state are integrated). While a majority of Islamic societies have been and are differentiated social formations, a small but significant number have been and are societies that can be classified as undifferentiated social formations. A common label used in contemporary discourse to refer to undifferentiated Muslim social formations is "the Islamic state."

The empirical evidence shows that religious institutions and religious elites tend to enjoy greater public trust and legitimacy in differentiated compared to undifferentiated Muslim societies. The underlying dynamics that appear to produce this pattern are related to the functional and performance roles of religious institutions (Luhmann 1982; Beyer 1994) and the ability of religious institutions to mobilize public resistance against an authoritarian state that has a deficit of legitimacy in the public mind (Hassan 1999a and 1999b).

### GENDER ISSUES IN MUSLIM SOCIETIES

For many Islamic and Western scholars of Islam, the status, role and position of women are important distinguishing features of Muslim societies that, set them apart from their Western counterparts. Many people in the West regard the status of women in Muslim society as symptomatic of their oppression in Islam (Esposito, 1995, p. 5). It is further argued that gender relations in Islam have been shaped primarily by their Arabian origins. While Islam has borne the marks of its Arabian origin throughout its history, in regard to the position held by women in his community, Mohammed was able to introduce profound changes (Levy 1972; Rahman 1966; Ali 1970).

Islam was instrumental in introducing wide-ranging legal-religious enactments to improve the status and position of women in Arabian society and protect them from male excesses. There are numerous Koranic injunctions to give effect to these changes (Ali 1970, pp. 55-59). These injunc-

tions brought about significant improvements in the status of women in a wide range of public and private spheres, but most important, they gave women a full-fledged personality (Rahman 1966).

However, selective literal, noncontextual, and ahistorical interpretations of sacred texts by Islamic scholars over time have shaped the average Muslim's conservative views and attitudes toward women. One of the major dilemmas faced by the nationalist leaders who spearheaded independence movements from Indonesia to Pakistan and Egypt was "woman issue." Their problem was how to respond to the questions raised by women about their role, status, and function in the new independent states. This generated highly emotional and divisive debates between the Islamic scholars and the nationalist leaders that centered on the issues of marriage and family law and the role and status of women in a modern independent Muslim state (see Esposito 1982; Haddad and Esposito 1998).

Notwithstanding strong resistance from Islamicists in several countries, the new nationalist leaders were able to overcome centuries of resistance and introduce modest changes in family and marriage laws. Those changes were introduced within an Islamic framework that did not expressly violate the appropriate Koranic injunctions and Sunnah (Anderson 1976). Those reforms have been criticized and opposed by a majority of Islamic Ulema and their followers, who regard them as violations of Islamic law and commandments as codified in classical Islamic legal texts as well as thinly veiled attempts to find an Islamic justification for an essentially Western approach to issues of interpersonal relations (Haeri 1993; Esposito 1982). This debate between nationalists and Islamicists continues and according to some evidence is becoming an important part of the political agenda of Islamic fundamentalists (Hardacre 1993; Haeri 1993).

**Attitudes toward Veiling and Patriarchy:** Veiling and seclusion of women and patriarchy have been important features of Islamic societies. In recent years they have attracted much criticism from Muslim and Western feminist scholars. The tradition and custom of veiling in Islam can be attributed to Islamic history, Islamic texts, and the privileged position of males and their control and dominance of positions of power and authority in Muslim society. Veiling and seclusion of women

and their role and function in society also are intertwined with the management of sexuality in Islam (Levy 1972).

Islam recognizes sexual desire as a natural endowment of the human body and enjoins its followers to satisfy and even enjoy sexual needs, providing a framework for doing so enunciated in the sacred texts. Unlike Christianity, Islam does not sanction or idealize celibacy. Over the centuries, the interpretations of sacred texts by the Ulema have led to the development of an institutional framework for the management and satisfaction of human sexuality through the imposition of control over women. As women are seen not only as sexual beings but also as the embodiment of sex, the social framework that has evolved has come to view the woman's body as pudendal. This conceptualization has led to the development and observance of strict dress codes for women, including veiling and seclusion, to prevent them from displaying their bodily charm and beauty (Haeri 1993; Hardacre 1993; Levy 1972).

Other features of the institutional framework arose out of the fact that women were made the principal actors responsible for preserving the sanctity of the family and reproduction. This led to strict injunctions on the types of roles they could play in the public sphere. Strong social and cultural traditions evolved that placed serious obstacles in the way of women seeking to succeed in public roles. Men, in contrast, were assigned all the public roles as providers, protectors, and arbiters, and this reinforced their power in the domestic domain as well. Patriarchal family structures thus became more functionally suitable to the perpetuation of the institutional framework for the satisfaction and management of the family.

That institutional framework and its accompanying normative requirements as they apply to gender roles, dress codes, veiling and seclusion, and patriarchy are by and large universally accepted in Muslim societies, although their observance varies with economic conditions. For most ordinary Muslims, this practice is in keeping with the supremacy of the male over female postulated by the Koran. However, the vagueness of these edicts has given the Ulema greater authority to interpret them as local custom demands. Some Ulema even appear to have invented tradition to bolster their interpretations which may in fact



conflict with Koranic statements (Levy 1972; Rahman 1982; Mernissi 1989; Rugh 1984).

As a result of internal and external pressures, the governments of most Islamic countries have initiated reforms to improve the quality of citizenship accorded to Muslim women. These reforms have sought to remove some of the obstacles that have prevented gender equality. While varying in scope and intensity from country to country, these reforms have been initiated in most Muslim countries. Some of the reforms have been successful, and, in some countries, such as Iran and Pakistan, the pendulum has swung to more traditionalist views that have gained favor with the current ruling elites. In general, the reforms are having a positive effect, although obstacles still exist. Those obstacles will continue until the rigid attitudes of the Ulema change or lose significance for the general body of Muslims as a result of the decline of their religious authority.

The empirical evidence about attitudes toward veiling, seclusion, and patriarchy indicates that those attitudes are an outcome of complex processes, including the prevalent social, economic, and political conditions in the country that mediate between the traditional Islamic norms and their practice in the local milieu. The material conditions of the country influence the shaping of attitudes toward these issues more strongly than does traditional Islamic ideology. The empirical evidence also indicates that in Muslim societies where men have experienced greater status loss relative to women as a result of public policies aimed at improving the quality of female citizenship, they appear to have compensated for that loss by developing more conservative attitudes toward women, including veiling, seclusion of women, and patriarchy. The evidence also suggests that paradoxically, Muslim societies, that are more successful in providing women with institutional equality may be more successful in generating more positive attitudes toward traditional Islamic values of patriarchy, veiling, and segregation among women (Hassan 1999c).

### MUSLIM MINORITIES

The London-based Institute of Muslim Minority Affairs estimates that about three hundred million Muslims live in one hundred forty-nine non-Muslim states. The institute publishes a biannual jour-

nal devoted to studies of Muslim minorities in different countries. With over 100 million Muslims, India is the home of the largest Muslim minority. Over the past fifty years, international labor migration and political upheavals have resulted in increasing Muslim settlement in Europe, Australia, and North America. It is estimated that about 20 million Muslims now reside in Europe. Most of them arrived as immigrants to meet the labor needs of booming west European economies, and their numbers are likely to increase in the future. (Nielsen 1995; Castles 1989).

The Muslim presence in west European countries has raised challenges to both Muslim and European traditions. The evidence shows that on the whole, Muslim communities in European countries are making cultural, social, and religious adjustments to secure their position in society. The host societies are responding by promoting cultural pluralism and containing racism and xenophobic attitudes among some segments of their populations. The cultural interactions between European and Muslim communities is shaping a distinctive European Muslim identity among second-generation Muslims (Nielsen 1995; Gerholm and Lithman 1988).

The Muslim presence in North America, especially in the United States, has been increasing gradually. While there are no reliable statistics on the exact size of the Muslim population in the United States, it is estimated to range from 2 million to 6 million. The American Muslim Council estimates the figure to be around 5.2 million (Duran 1997). Whatever the size, it is a well-established fact that Islam is an important feature of the American religious milieu. Its adherents include American Muslims, who are predominantly of African origin, and more recent immigrants from all over the Muslim world. The most widely known American Muslim group is the Nation of Islam (Gardell 1994). Most of the immigrant Muslims come from south Asian and Middle Eastern countries. They have arrived as skilled and unskilled laborers, students, and refugees from political developments in Muslim countries. The largest concentrations of Muslims are in the metropolitan areas of Los Angeles, New York, and Chicago. Recent sociological studies have focused on acculturation and the tensions Muslim communities experience in adjusting to life in America and how those communities are responding to the compet-

ing demands of belonging to a universal Islamic community (*Ummah*) and maintaining their ethnic or national identity. The evidence shows that the majority of American Muslims are not active participants in the organized religious life of mosque or Islamic center, but continue to identify themselves as Muslims in a social setting characterized by prejudice and misunderstanding (Haddad and Smith 1994; Haddad and Lummis 1987). Similar findings have been reported in studies of Muslim communities in Australia (Hassan 1991; Bouma 1994).

The social pervasiveness of Islam in the modern world and the sociopolitical and religious trajectories of contemporary Muslim societies raise important sociological questions. This article has identified some of the questions and issues that make the sociology of Islam a challenging field of social inquiry. Empirical studies of Muslim societies can be a rich source for evaluating the validity of some of the major propositions of social theory that have been formulated in the context of increasingly secular social settings of modern European and North American countries. Through systematic and comparative studies of Muslim societies, modern sociological scholarship can lay the foundations for a more informed understanding of the social reality of the Muslim world.

REFERENCES

Ahluwalia, P., and P. L. Mayer, 1994 "Clash of Civilisations—or Balderdash of Scholars." *Asian Studies Review* 18:129-30.

Ajami Fouad 1993 "The Summarising." *Foreign Affairs* 72(4):2-9.

Ali, Chiragh 1970 "The Position of Women in Aziz Ahmad and E. von Grunebaum, eds., *Muslim Self-Statement in India and Pakistan 1857-1968*. Weisbaden, Germany: Otto Harrassowitz.

Anderson, J. N. D. 1976 *Law Reform in Muslim World*. London: Athlone Press.

Arjomand, Said 1988 *The Turban and the Crown: The Islamic Revolution in Iran*. New York: Oxford University Press.

Ayubi, Nazih N. 1991 *Political Islam: Religion and Politics in the Arab World*. London: Routledge.

Beinin, J., and J. Stork, eds. 1997 *Political Islam, Berkeley*: University of California Press.

Bell, Richard 1926 *The Origins of Islam in Its Christian Environment*. London: Macmillan.

Beyer, Peter 1994 *Religion and Globalization*. London: Sage.

Bouma, Gary D. 1994 *Mosques and Muslim Settlement in Australia*. Canberra, Australia: AGPS.

Castles, S. 1989 *Here for Good: Western Europe's New Ethnic Minorities*. London: Pluto Press.

Crone, Patricia 1996 "The Rise of Islam in the World." In Francis Robinson, ed., *The Cambridge Illustrated History of the Islamic World*. Cambridge, UK: Cambridge University Press.

Duran, Khalid 1997 "Demographic Characteristics of the American Muslim Community." *Islamic Studies* 36(1): 57-76.

Engineer, Asghar Ali 1980 *The Origins and Development of Islam*. Kuala Lumpur: Ikraq.

Esposito, John L. 1982 *Women in Muslim Family Law*. Syracuse, N.Y.: Syracuse University Press.

—, ed. 1983 *Voices of Resurgent Islam*. New York: Oxford University Press.

— 1995 *The Islamic Threat: Myth or Reality?* New York: Oxford University Press.

Fischer, Michael 1980 *Iran: From Religious Dispute to Revolution*. Cambridge, Mass.: Harvard University Press.

Folliet, Joseph 1955 "The Effects of City Life upon Spiritual Life." In R. Fisher, ed., *The Metropolis in Modern Life*. New York: Doubleday.

Gardell, Mattias 1994 "The Sun of Islam Will Rise in the West: Minister Farrakhan and the Nation of Islam in Latter Days." In Y. Haddad and J. Smith, eds., *Muslim Communities of North America*. New York: State University of New York Press.

Geertz, Clifford 1960 *The Religion of Java*. Chicago: University of Chicago Press.

— 1968 *Islam Observed: Religious Developments in Morocco and Indonesia*. Chicago: University of Chicago Press.

Gellner, Ernest 1969 *Saints of the Atlas*. London: Weidenfeld & Nicholson.

—, 1983 *Muslim Society*. Cambridge, UK: Cambridge University Press.

— 1992 *Postmodernism, Reason and Religion*. London: Routledge.

— 1994 *Conditions of Liberty*. London: Penguin Books.

Gerholm, T. and Y. G. Lithman, eds. 1988 *The New Islamic Presence in Europe*. London: Mansell.

Haddad, Yvonne Y. and John L. Esposito (eds.) 1998 *Islam, Gender and Social Change*. New York: Oxford University Press.

- and Jane I. Smith (eds.) 1994 *Muslim Communities in North America*. New York: State University of New York Press.
- Haddad, Yvonne Y. and A.T. Lummis 1987. *Islamic Values in the United States*. New York: Oxford University Press.
- Haeri, Shahla, 1993 "Women and Fundamentalism in Iran and Pakistan." In Martin E. Marty and R. Scott Appleby, eds., *Fundamentalisms and Society*. Chicago: University of Chicago Press.
- Hardacre, Helen 1993 "The Impact of Fundamentalisms on Women, the Family and Interpersonal Relations" in Martin E. Marty and R. Scott Appleby, eds. *Fundamentalisms and Society*. Chicago: University of Chicago Press.
- Hassan, Riaz 1984 "Iran's Islamic Revolutionaries: Before and After the Revolution." *Third World Quarterly*, 6(3):675-686
- 1985a *Islam Dari Konservatisme Sampai Fundamentalisme*. Jakarta: Rajawali Press.
- 1985 "Islamization: An Analysis of Religious, Political and Social Change in Pakistan." *Middle Eastern Studies*, 21(3):263:283.
- 1987 "Pirs and Politics: Religion, Society and the State in Pakistan." *Asian Survey* 26(5): 552-565
- 1991 "The Muslim Minority-Majority Relations in Australian Society: A Sociological Analysis." *Journal of Institute of Muslim Minority Affairs* 12(2):285-306.
- 1999a "Faithlines: Religion, Society and the State in Indonesia and Pakistan." *Islamic Studies* 38(1):45-62.
- 1999b "Faithlines: Social Structure and Religiosity in Muslim Societies." Working Paper, G. E. Von Grunebaum Center for Near Eastern Studies, University of California of Los Angeles.
- 1999c. "Attitudes towards Veiling and Patriarchy in Four Muslim Societies." Working Paper, G. E. Von Grunebaum Center For Near Eastern Studies. University of California of Los Angeles.
- 1999d "The Islamic Ummah: Myth or Reality?" unpublished paper. Department of Sociology, Flinders University of South Australia, Adelaide.
- Hodges, Richard, and David Whitehouse, 1983 *Mohammed, Charlemagne and the Origins of Europe*. Ithaca N.Y.: Cornell University Press.
- Hodgson, Marshall G. 1975 *The Venture of Islam*. Chicago: University of Chicago Press.
- Hume, David 1976 *The Natural History of Religion*. Oxford: Oxford University Press.
- Huntington, Samuel P. 1993 *The Clash of Civilizations and the Remaking of World Order*. New York: Simon & Schuster.
- Irfani, Suroosh 1983 *Revolutionary Islam in Iran*. London: Zed.
- Keddie, Nikki R. 1994 "The Revolt of Islam, 1700 to 1993: Comparative Considerations and Relations to Imperialism." *Comparative Studies in Society and History* 36(3): 463-487.
- Khaldun, Ibn 1958 *The Muqaddamah*. London: Rosenthal.
- 1992 *An Arab Philosophy of History*, trans. and arranged by Charles Issawi. Cairo: American University of Cairo Press.
- Kroeber, Alfred 1948 *Anthropology*. New York: Harcourt Brace.
- Lapidus, Ira M. 1996 "State and Religion in Islamic Society." *Past and Present* 151:3-27.
- Levy, Reuben 1972 *The Social Structure of Islam*. Cambridge, UK: Cambridge University Press.
- Lewis B. 1993 *Islam and the West*. New York: Oxford University Press.
- Luhmann, Niklas 1982 *The Differentiation of Society*, trans. Stephen Holmas and Charles Lamore. New York: Columbia University Press.
- Marty, Martin E. and R. Scott Appleby, eds. 1991 *Fundamentalism Observed*. Chicago: University of Chicago Press.
- 1993 *Fundamentalism and Society*. Chicago: University of Chicago Press.
- 1992 *The Glory and the Power: The Fundamentalist Challenge to the Modern World*. Boston: Beacon Press.
- Maududi, A. A. 1960 *The Islamic Law and Constitution*. Lahore, Pakistan: Islamic Publications.
- Mernissi, Fatimah 1989 *Women and Islam*, trans. by Mary Jo Lakeland. Oxford, UK: Basil Blackwell.
- Mol, J. J. 1972 *Western Religion*. The Hague: Mouton.
- Nielsen, Jorgen 1995 *Muslims in Western Europe*. Edinburgh, Scotland: Edinburgh University Press.
- Pipes, Daniel, 1981. *Slaves, Soldiers and Islam: The Genesis of a Military System*. New Haven, Conn.: Yale University Press.
- Pirenne, Henry 1939 *Mohammed and Charlemagne*. London: Barnes and Noble.
- Rahman, F. 1966 *Islam*. Chicago: University of Chicago Press.
- 1982 *Islam and Modernity*. Chicago: University of Chicago Press.—1989 *Major Themes of the Quran*. Minneapolis: Bibliotheca Islamica.
- Rugh, Andrea 1984 *Reveal and Conceal: Dress in Contemporary Egypt*. Syracuse, N.Y.: Syracuse University Press.

- Sadowski, Yahya 1997 "The New Orientalism and the Democracy Debate." In J. Beinlin and J. Stork, eds., *Political Islam*. Berkeley: University of California Press.
- Shariati, Ali 1979 *On the Sociology of Islam*. Berkeley, Calif.: Mizan Press.
- Sivan, Emmanuel 1985 *Radical Islam: Medieval Theology and Modern Politics*. New Haven, Conn.: Yale University Press.
- Tamney, Joseph 1980 "Modernization and Religious Purification: Islam in Indonesia." *Review of Religious Research* 22(2): 208–218
- Torrey, Charles 1933 *The Jewish Foundation of Islam*. New York: Jewish Institute.
- Turner, Bryan 1974 *Weber and Islam*. London: Routledge & Kegan Paul.
- 1974 "Islam, Capitalism and the Weber Thesis." *British Journal of Sociology*. 25:230–243.
- 1999 "The Sociology of Islamic Social Structure." *Sociology* 33:1.
- Watt, W. Montgomery 1954 "Economic and Social Aspects of the Origin of Islam." *Islamic Quarterly* 1:90–103.
- 1955 "Ideal Factors in the Origin of Islam." *Islamic Quarterly*. 2:160–74.
- 1962a *Mohammed at Mecca*. Oxford: Oxford University Press.
- 1962b *Mohammed at Medina*. Oxford: Oxford University Press.
- (1988) *Islamic Fundamentalism and Modernity*. London: Routledge.
- Weber Max, 1958 *The Protestant Ethic and the Spirit of Capitalism*, trans. Talcott Parsons. New York: Scribner.
- 1964 *The Sociology of Religion*, trans. E. Fischoff. Boston: Beacon.
- 1978 *Economy and Society*. Geunther Ross and Claus Wittich, eds. Berkeley: University of California Press.
- Wolf, Eric R. 1951 "The Social Organization of Mecca and the Origins of Islam." *Southwestern Journal of Anthropology* 7(4): 329–356.
- Yap, M. E. 1980 "Contemporary Islamic Revival." *Asian Affairs Journal of the Royal Society for Asian Affairs* 11(2):178–195.
- Zubadia, Sami 1989 *Islam, the People and the State*. London: Routledge.
- 1995 "Is There a Muslim Society? Ernest Gellner's Sociology of Islam." *Economy and Society* 24(2): 151–188.

RIAZ HASSAN

## THE SOCIOLOGY OF KNOWLEDGE

The sociology of knowledge as a subdiscipline in sociology deals with the social and group origins of ideas. In its brief history as a field of study, it has included the entire ideational realm (knowledge, ideas, theories, and mentalities), in an attempt to comprehend how that realm is related to particular social and political forces and how the mental life of a group of people arises within the context of the groups and institutions in which those people live and act. More recently, its subject matter has included not only a society's authoritative ideas and formal knowledges but also those which operate in the realm of everyday life: informal knowledges.

The term "sociology of knowledge" (*Wissenssoziologie*) was first used in 1924 and 1925 by Scheler (1874–1928) (Scheler [1924] 1980, 1992) and Mannheim (1893–1947) (Mannheim [1924] 1952). From its inception, it described a field of inquiry closely linked to problems of European philosophy and historicism. In several important respects, this is an accurate description, for the sociology of knowledge reflected the nineteenth-century German philosophical interest in problems surrounding relativism that were linked to the legacies of Karl Marx, Friedrich Nietzsche, and the historicists, whose cultural philosophy of worldviews (*Weltanschauungsphilosophie*) was influential in German social science from the 1890s to the 1930s. Each of these developments was concerned in different ways with the determinate relationship between thought and society, between knowledge and social structure. For Scheler and Mannheim, *Wissenssoziologie* would serve as an empirical and historical method for resolving the intense conflicts of ideologies in Weimar Germany that followed political and social revolutions of the nineteenth and early twentieth centuries and produced warring groups whose battles were manifestly ideational and grounded in conflicting worldviews. Sociology of knowledge would provide a method, outlined in early statements by Scheler and Mannheim, for unmasking the assumptions of political ideologies and indicating their truth content. However much Scheler and Mannheim differed about the nature of truth within relativism, both agreed that truths do not exist apart from historical and social processes. As mem-

bers of a postwar generation of European intellectuals, they also shared a sense that they were witnessing the gradual disappearance of epistemology and its replacement by the sociology of knowledge as a foundational discipline for all philosophy. As participants in this historical process, they also believed, as did their contemporaries, that intellectuals play a vital role in thought and politics.

The excitement and urgency with which the framers of *Wissenssoziologie* approached the study of the social origins of ideas has been replaced by a widespread acceptance of their premises concerning the social origins of ideas, ideologies, and worldviews. To borrow Weber's term, the sociology of knowledge was "routinized" into the established structures and practices of modern social science. Many of the positions advanced by Scheler, Mannheim, and other early writers in this field (e.g., in the United States by Florian Znaniecki, C. Wright Mills, and Edward Shils) operate today as working propositions for a range of social scientists as well as for specialists in other disciplines, including the subfields of the history of ideas, social psychology, social studies of science, feminist theory, and cultural studies. Even the urgency, expressed by Mannheim, surrounding the problem of relativism as a "contemporary predicament" has been transformed into a commonplace fact. Today, this is certainly the case in the academic world, whereas in the past, the sociology of knowledge provided the occasion for intense controversies about the postulate of the essential "sociality of mind" (Child 1941).

#### THE SOCIOLOGY OF KNOWLEDGE: TWO APPROACHES

Partly because of the diffusion of the idea of the social nature of knowledge, the sociology of knowledge has been described as an approach or subdiscipline that has no unified field, but only a series of theoretical works and research agendas. Despite this characterization, the subdiscipline of the sociology of knowledge is a recognized field of endeavor that continues to draw new generations of sociologists. Therefore, one may speak of two ways of introducing the sociology of knowledge: The *broad approach* identifies a range of works in sociology and social theory that examine the social nature of mind and knowledge; the *particular ap-*

*proach* includes the works of specialists the field identified as the sociology of knowledge. Several leading contributors to the sociology of knowledge have provided similar schemes for delineating its subject matter, pointing out that the sociology of knowledge includes both a broad field and a narrow field of studies and that both fields contribute to the sociology of mental and cognitive structures (Remmling 1973; Curtis and Petras 1970; Berger and Luckmann 1966).

The broad approach incorporates a number of works that deal with the relationship of mental life (cognition, consciousness, collective ideas, etc.) and social life (groups, institutions, communities, entire societies). The broad approach treats the sociology of knowledge as a "frame of reference," not a "definite body of theory in its own right" (Curtis and Petras 1970, p.1). Accordingly, the sociology of knowledge is a broad tradition of inquiry, a handing down of key texts and theories, such as theories of the "social determination" of ideas, theories of ideology, the relationship of "real" and "ideal" factors, and the notion of *Weltanschauung*, that are closely linked to the history of sociology. The social theories of Karl Marx, Emile Durkheim, Max Weber, Georg Simmel, and others are studied as classic statements on the relationship of mind, knowledge, and society. This broad approach to the sociology of knowledge has provided not only the basic materials from which particular treatises in the sociology of knowledge have been written (e.g., Stark [1958] 1991; Berger and Luckmann 1966; Gurvitch 1971) but also the materials that have been incorporated into commentaries and edited collections on the sociology of knowledge (Remmling 1973; Curtis and Petras 1970).

In this general sense, the sociology of knowledge is understood as a field that systematizes the leading propositions of the modern social sciences about the social nature of mind. Furthermore, like sociology, the sociology of knowledge constitutes a tradition of inquiry that reflects and shapes the development of modernity. That is, sociology offers a theory of the human mind that is compatible with "our time": The sociology of knowledge "appears as a revision of our way of . . . looking at ourselves and the world. . . . It 'defines' a new 'situation'" (Wolff 1953, p. 618; Wolff cf.1959). Linked as it is to the modernization of conscious-

ness, the sociology of knowledge, broadly conceived, has several distinct national traditions, each focusing on themes characteristic of its own modern intellectual legacies. Therefore, one can speak of French, German, and American traditions of the sociology of knowledge whose roots are based in a Durkheimian “structuralist” legacy, a Marxist or Mannheimian theory of ideology, and a pragmatist theory of mind such as that offered by John Dewey and George Herbert Mead. Each of these national intellectual traditions reflects its particular national and cultural legacies in nineteenth- and twentieth-century modernity; each legacy also can be seen as complementing the sociologies of knowledge of other modern nations, cultures, and civilizations.

A second way of defining the sociology of knowledge considers the field as a particular body of work and examines its origins, development, and future prospects. This approach begins with the original statements of Scheler and Mannheim and proceeds to the later principal works and arguments. The approach also examines major substantive statements made by sociologists identified with its precise subject matter. One of the merits of this approach is that it allows for a critical view of the substantive work in the sociology of knowledge over time and, in keeping with the field’s presuppositions about the existential determination of thought (*Seinsgebundenheit*), opens the question of how social theories of knowledge are themselves subject to change and revision over time. In this sense, the sociology of knowledge offers a metatheory through which sociology can examine how its leading concepts and theories arise in response to particular social and political situations. For example, Marx’s theory of ideology is closely implicated in particular historical conditions of the industrial capitalist order, and its validity is dependent on particular conditions of social and economic organization, such as the separation and autonomy of economic forces in the social order.

Integral to the sociology of knowledge is a *relative theory of knowledge* from which its own concepts and theories are not excepted. Its methods are critical in the classical sense of the word, for it offers a continuous criticism of what it studies, including its own forms of knowledge and criteria of judgment. With this view in mind, a brief history of its statements and theories offers

more than a recounting of its nature and scope. It also draws attention to the reflexive features of all sociological inquiries, particularly the fact that sociology is part of the social reality it studies in that its changing concepts and insights develop out of and address particular social worlds. Sociological theories are neither external nor formal. The brief history of the subfield of the sociology of knowledge that follows is intended to be both a recounting of the leading ideas in this field and a reflexive statement about the social foundations of its theories and presuppositions. The implications that can be drawn from this inquiry are taken up in the conclusion.

### A BRIEF SUBSTANTIVE HISTORY

Since its inception in the writings of Scheler and Mannheim, the sociology of knowledge has identified a number of precise ways in which knowledge is socially determined. Scheler’s original essays ([1924] 1980) identifying the field of study provoked commentary and debate. His concept of a society’s “relatively natural *Weltanschauung*” is still central to cultural sociology, as are his propositions concerning the origins of the modern worldview and its scientific ethos. However, Scheler’s importance would be felt decades later (Bershady 1992). It was Mannheim’s formulation of the discipline in *Ideology and Utopia* (German edition 1929, English edition 1936) that originally defined the subject matter of the field and continued to do so for years to come. Those who proposed different sociologies of knowledge after its publication defined their positions relative to Mannheim’s arguments concerning ideology, utopia, and relationism.

Mannheim’s treatise begins with a review and critique of the prevailing and authoritative Marxist theories of ideology (the “particular theory of ideology”) and proceeds toward a theory of ideology in the broader sense: the mental structure in its totality as it appears in different currents of thought and across different social groups. This total conception of ideology examines thought on the structural level, allowing the same object to take on different (group) aspects. This understanding of ideology refers to a person’s, group’s, or society’s entire way of conceiving things as it is provided by particular historical and social settings. The “total conception of ideology” defines

the subject matter of the sociology of knowledge. Like ideologies, “utopias” arise out of particular social and political conditions, but they are distinguished by their opposition to the prevailing order. Utopias are the embodiment of “wish-images” in collective actions that shatter and transform social worlds partially or entirely. Both concepts form part of Mannheim’s theoretical apparatus for a critical but nonevaluative treatment of “ideology” that supersedes the sociohistorical determinism and relativism of Marxism while moving toward a “relationist” notion of truth. The enterprise of the sociology of knowledge examines how collective actions and ideas (ideologies and utopias) emerge out of and are “determined” by the multiple social contexts and positions of their proponents. From an analysis of the various and competing social positions of ideologists and utopians, a kind of “truth” emerges that is grounded in the conditions of intellectual objectivity and detachment from the social conditions that more directly determine ideas. *Ideology and Utopia* established the criteria for a valid knowledge, albeit a *relational* knowledge, of sociohistorical processes. More important, it raised the problems surrounding the historicity of thought and did this within the newly emerging academic discourse of sociology. In the process, this work gave legitimacy to a new set of methodological problems involving the problems of objectivity and truth for the sciences and the humanities.

Despite the many criticisms of *Ideology and Utopia* (particularly Mannheim’s attempt to avoid the pitfalls of historical relativism), the work received wide attention and appreciation inside and outside the social sciences, where the problems posed by relativism continued to attract the attention of workers in both the sciences and the humanities. While reviews of the work focused on its failure to overcome relativism and Mannheim’s excessive reliance on the Marxist conception of ideology, Mannheim’s book provoked discussion and commentary for years (Hughes 1958, p. 420).

In the decades after its publication, *Ideology and Utopia* engaged the leading American social theorists of the period: Merton and Parsons. Merton’s two chapters on the sociology of knowledge in his major work (1957) attempt to integrate the social theory of knowledge with his own “structural-functional” theory and demonstrate how other

theorists (Marx, Weber, Freud, Durkheim) belong to the broader tradition of the social determination of ideas. In essays by Parsons (1959, 1961), Mannheim’s work is criticized and integrated into the approach with which Parsons is identified: the “general theory of action.” The contributions of Merton and Parsons were significant, principally with respect to the prevailing functionalist theories of “action,” not with respect to advancing the sociology of knowledge. In fact, it could be said that the projects outlined by Scheler and Mannheim, particularly their historical and cultural emphasis, did not conform to the program of formal or “general theory” outlined by Parsons for the functional study of all societies or to Merton’s proposal for theories of the “middle range.”

This was not the case for their contemporary Stark ([1958] 1991), whose sociology of knowledge attempted to clarify the principal themes in the study of the social determination of ideas and advance its arguments beyond the Marx–Mannheim tradition and its theory of ideology. An émigré for more than half his life and a scholar educated at the universities of Hamburg, Prague, London, and Geneva, Stark was accustomed to move within many mental, linguistic, and moral frameworks. When it is confronted by an almost dizzying array of viewpoints, social existence loses its taken-for-granted quality. As Remmling (1973) has observed, the relationship between social existence and knowledge, which has been the preoccupation of the sociology of knowledge, has always been that of marginal figures and outsiders. This and other traits Stark shared with Scheler and Mannheim. Stark regarded *Wissenssoziologie* as an indispensable method for understanding both the truth of ideas and the history of ideas; truths do not exist apart from the historical and social process. The traditions of German cultural sociology and *Wissenssoziologie* contain the ideals and conventions in which Stark’s sociology of knowledge becomes most intelligible. He brought to it judgments concerning the “real” and the “ideal” from Weber’s and Simmel’s sociology. From Scheler’s works in particular, he would find ways of returning to the problem of how to find truths or “ideal values” in the realm of relative social realities or “existential facts.”

Stark wrote *The Sociology of Knowledge* to clarify the principal themes of writers, especially sociolo-

gists, who had addressed the problem of the social element in thinking. He also intended it to serve as an introduction to the field that would prepare the way for a detailed and comprehensive history of the sociology of knowledge and its most significant ideas, including the theories of ideology of Marx and Mannheim, the philosophical speculations of the neo-Kantians Heinrich Rickert and Max Weber, and the views of the German phenomenological school of the 1920s, especially Scheler. Each of these ideas was vitally important for his project, but Stark's strongest affinity was with Scheler's struggle to reconcile the antithetical claims of idealism and materialism and his view of the sociology of knowledge as the foundation for a knowledge of eternal values.

Stark's sociology of knowledge is directed primarily toward the study of the precise ways in which human experience, through the mediation of knowledges, takes on a conscious and communicable shape. Eventually, Stark intends to direct this inquiry to the problem of truth, a synthesis of the different styles of thought and their limited truths. For either one of these intentions to be realized, he insists that the theory of ideology can have no place within the bounds of the sociology of knowledge. The idea that social influences enter mental life in the form of lies, self-deceptions, and distortions in thinking and are due to class positions and interests has dominated the Marx-Mannheim tradition and its theory of ideology. Stark's contention, shared by many contemporary writers, is that the sociology of knowledge is concerned with the "social determination of knowledge" (a term with a precise meaning of its own), not with the problem of ideology. In fact, this distinction is an indispensable precondition of the sociology of knowledge. It is intended to direct attention to the study of the extent to which all mental life is grounded in conditions that are ineluctably social and historical; it grants to "social determination" a depth that the theory of ideology does not permit, since that theory deals only with errors and misperceptions (Stark [1958] 1991, pp. 50-55). Even more important, in Stark's view, the theory of social determination is entirely compatible with the theory of truth, whereas the theory of ideology is concerned principally with the social conditions of error or false consciousness. While the theory of ideology will always play a vital role in sociology and the history of ideas, it must

be relegated to a status outside the principal concerns of the sociology of knowledge.

Stark's project involves building bridges between opposing positions (Mannheim's theory of ideology and Scheler's theory of social determination), what in scholastic philosophy was called a *concordantia discordantium canonum*, a reconciliation of opposing positions of thought. Whether these two traditions are indeed contradictory is not of consequence in grasping Stark's argument in *The Sociology of Knowledge*. Stark's willingness to explore with frankness and according to his own convictions the kind of epistemology that he thought consistent with cultural sociology brought him to a social theory of knowledge that was compatible with both *Verstehen* (lit., understanding) sociology and social phenomenology. This theory also dismissed the relevance of either a simple historical materialist theory or a positivist one. The outcome is a theory of social determination that has moved away from Marx and Mannheim and in the direction of a cultural sociology, one that is consistent with contemporary sociology's interest in the broad range of cultural studies.

Less than a decade after Stark's work appeared, Berger and Luckmann's *The Social Construction of Reality* (1966) advanced a sociology of knowledge that was compatible with the view of sociology as a humanistic discipline and the notion that "human reality" is a "socially constructed reality" (p. 189). These authors broadened the field to include all types of knowledge, including the knowledge of everyday life: "[T]he sociology of knowledge must concern itself with whatever passes for 'knowledge' in a society, regardless of the ultimate validity or invalidity (by whatever criteria) of such 'knowledge'" (p. 3). More important, their treatise asked that the sociology of knowledge address how, in the domain of the quotidian, knowledge constitutes social reality, redirecting the traditional theory of social determination of ideas by social realities. What Berger and Luckmann proposed was that knowledge and reality (by which they always mean social reality) exist in a reciprocal or dialectical relationship of mutual constitution. As many have argued, this work placed the sociology of knowledge on an entirely new footing whose focus is the broad range of signifying systems that form and communicate the realm of social realities. Since its publication, the idea of a "con-



structed reality” has summarized a number of concerns of contemporary writers in the sciences and humanities that may be best described as the problem of meaning and the use of philosophical, literary, and historical approaches to study the social construction of meaning.

The methodological implications of this change in sociology and the sociology of knowledge are noteworthy, since interest in the problem of meaning is linked to a methodological framework that is neither causal nor explanatory (the attitudes expressed by Mannheim’s theory of “social determination”) but *semiotic*. The semiotic study of culture is directed toward the study of the symbolic and signifying systems through which a social order is communicated and reproduced. These signifying systems and social practices make up a culture and its structures of meaning. According to Geertz (1973, p. 5), one of the principal semiotic theorists, the analysis of knowledge and culture is “not an experimental science in search of law but an interpretive one in search of meaning.”

In addition to being a proponent of an interpretive method in the social sciences in the 1960s, Geertz, in his essay “Ideology as a Cultural System” (1964), explicitly criticized the sociology of knowledge (Mannheim and Stark explicitly), arguing that the entire enterprise identified with the twin problems of “ideology” and “truth” should be reformulated as the “sociology of meaning.” For Geertz, the sociology of knowledge remained lodged in an older set of presuppositions (principally its use of “ideology”) that prevented it from moving toward a nonevaluative understanding of “culture.” This and other criticisms at the time effectively redirected sociology in the period of the late 1960s toward what has been described as its “postpositivist” phase (Rabinow and Sullivan 1979, 1987). Ironically, this most recent period of sociology, with its rejection of the classical concerns of the sociology of knowledge, has been described by Robertson (1992) as a “*general* sociology-of-knowledge turn” characterized by a focus on the ideational features of the social world or a resurgence of interest in cultural forms more generally. Put simply, contemporary sociology’s turn away from the classical problems and perspectives of the sociology of knowledge occurred precisely at the time when “culture,” “knowledge,” and “language” became central to sociology.

What has been called “the new sociology of knowledge” (Swidler and Arditì 1994; McCarthy 1996) can be seen as part of this larger movement in the social sciences generally, distinguished by a turn away from materialist theories or theories of social structure and in the direction of semiotic theories that focus on the ways in which a society’s multifarious meanings are communicated and reproduced. Hall (1980) has described the theoretical significance of this cultural turn in social science: Its problematic has become closely identified with the problem of the autonomy of cultural practices. The paradigm for studying the range of cultural practices has come largely from structuralist theories (Althusser, Levi-Strauss, Barthes). Language is the theoretical and empirical model, one that is neither positivist nor reductionist but interpretive rather than causal.

#### THE NEW SOCIOLOGY OF KNOWLEDGE

These arguments and others have been made by recent commentators in what has been called the “new sociology of knowledge.” In the case of Swidler and Arditì (1994), the new approach examines how specific kinds of social organizations (e.g., the media through which knowledge is preserved, organized, and transmitted) order knowledges rather than examining social locations and group interests. These scholars also examine, in light of new theories of social power and practice (Michel Foucault and Pierre Bourdieu), how knowledges maintain social hierarchies and how techniques of power are simultaneously and historically linked to discursive forms (knowledges). They also argue that newer theories of power, gender, and knowledge depart from the economic, class, and institutional focus of the classical sociology of knowledge.

McCarthy’s (1996) theoretical treatise traces changes in three broad national traditions in the sociology of knowledge (German, French, American), delineating the precise ways in which the classical traditions identified with these three national intellectual legacies have moved to models that are linguistically based. McCarthy also points to feminist theories as important contributions to the sociology of knowledge, particularly works in the sociology of science by feminists such as Smith (1987, 1990). These and other changes in sociology are examined against changes in the social location of knowledge and culture today, particu-

larly the predominant role played by systems of mass communication and information technologies. These changes in turn have produced a contemporary culture that is more globally aware, reflexive, and attuned to the operations of culture itself.

### CONCLUSION

The brief history of the sociology of knowledge from Mannheim to contemporary sociology lends itself to the type of interpretive scheme that originated with the classical sociology of knowledge, for its principal argument has worked against any formal understanding of either theory or science. Changes in the structures and organizations of social worlds have been functionally related to collective “standpoints” and “perspectives.” Sociologists have witnessed the shift, since midcentury, from “social structure” to “culture” as authoritative schemes for describing and interpreting how social knowledges are “socially determined.” Clearly, this intellectual shift registers changes in the social landscape of late modernity, where the configurations known to sociology as “economy,” “culture,” and “social structure” have undergone changes.

Proponents of the new sociology of knowledge and others have documented changes in the industrial societies of the last half century that correspond to the newer “cultural” theories. Neither Swidler and Arditì nor McCarthy claims that the sociology of knowledge as a subfield of sociology has been superseded by newer work in sociology and cultural studies. However, they note that the new sociology of knowledge is not yet a unified field, and their proposals for what constitutes this diffuse field constitute an argument reminiscent of those of the proponents of a broad or diffuse sociology of knowledge discussed above. It would seem that the more “cultural” sociology becomes, the more likely it is that the sociology of knowledge will be seen as a broadly inclusive set of studies rather than a subfield with a distinct subject matter. The subject matter of the sociology of knowledge has undergone significant change: What began as the study of competing and conflicting collective ideas and ideologies has become something more cultural and diffuse. Its subject matter today is both more differentiated and more diffuse and includes the study of “informal knowl-

edge,” the knowledge of everyday life. What began as the study of conflicting ideologies has become the study of the unspoken understandings of everyday life, what L. Wirth described in his preface to *Ideology and Utopia* as “the most elemental and important facts about a society . . . those that are seldom debated and generally regarded as settled” (1936, p. xxiii). Today these understandings have become the subject matter of sociology and are no longer generally regarded as “settled.”

### REFERENCES

- Berger, Peter, and Thomas Luckmann 1996 *The Social Construction of Reality*. New York: Doubleday.
- Bershady, Harold J. 1992 Introduction to Max Scheler, *On Feeling, Knowing, and Valuing*. Chicago: University of Chicago Press.
- Child, Arthur 1941 “The Theoretical Possibility of the Sociology of Knowledge.” *Ethics* 51:392–418.
- Curtis, James E, and John W. Petras 1970 *The Sociology of Knowledge*. New York: Praeger.
- Geertz, Clifford. 1964 “Ideology as a Cultural System.” In David E. Apter, ed., *Ideology and Discontent*. New York: Free Press.
- 1973 *The Interpretation of Cultures*. New York, Basic Books.
- Gurvitch, Georges 1971 *The Social Framework of Knowledge*, trans. M. A. Thompson and K. A. Thompson, introductory essay by K. A. Thompson. New York: Harper.
- Hall, Stuart. 1980. “Cultural Studies and the Centre: Some Problematics and Problems.” In *Culture, Media, Language: Working Papers in Cultural Studies*. London: Hutchinson.
- Hughes, H. Stuart 1958 *Consciousness and Society: The Reorientation of European Social Thought 1890–1930*. New York: Vintage.
- Mannheim, Karl (1924) 1952 “Historicism.” In P. Kecskerneti, ed., *Essays on the Sociology of Knowledge*. New York: Harcourt, Brace & World.
- (1929) 1936 *Ideology and Utopia*. New York: Harcourt, Brace & World.
- McCarthy, E. Doyle 1996 *Knowledge as Culture: The New Sociology of Knowledge*. New York and London: Routledge.
- Merton, Robert K. 1957 *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Parsons, Talcott 1959 “An Approach to the Sociology of Knowledge.” *Transactions of the Fourth World Congress of Sociology* IV:25–49. Louvain: International Sociological Association.

- 1961 “Culture and the Social System.” In T. Parsons, E. Shils, K. P. Naegle, and J. R. Pitts, et al., eds., *Theories of Society*, vol. II. New York: Free Press.
- Rabinow, Paul, and William M. Sullivan 1979 *Interpretive Social Science*. Berkeley: University of California Press.
- 1987 *Interpretive Social Science: A Second Look*. Berkeley: University of California Press.
- Remmling, Gunter 1967 *Road to Suspicion: A Study of Modern Mentality and the Sociology of Knowledge*. New York: Appleton-Century-Crofts.
- 1973 *Towards the Sociology of Knowledge*. London: Routledge & Kegan Paul.
- Robertson, Roland 1992 “Cultural Relativity and Social Theory: Werner Stark’s *Sociology of Knowledge* Revisited.” In E. Leonard, H. Strasser, and K. Westhues, eds., *In Search of Community: Essays in Memory of Werner Stark 1909–1985*. New York: Fordham University Press.
- Scheler, Max (1924) 1980 *Problems of a Sociology of Knowledge*, trans. M. S. Frings, ed. and with an introduction by K. W. Stikkers. London: Routledge & Kegan Paul.
- 1992 *On Feeling, Knowing, and Valuing*. Chicago: University of Chicago Press.
- Smith, Dorothy E. 1987 *The Everyday World as Problematic*. Boston: Northeastern University Press.
- 1990 *The Conceptual Practices of Power: A Feminist Sociology of Knowledge*. Boston: Northeastern University Press.
- Stark, Werner (1958) 1991 *The Sociology of Knowledge*, with a new introduction by E. Doyle McCarthy. New Brunswick, N.J.: Transaction.
- Swidler, Ann, and Jorge Arditi 1994 “The New Sociology of Knowledge.” *Annual Review of Sociology* 20:305–329.
- Wirth, Louis 1936 “Preface” in Karl Mannheim, *Ideology and Utopia*. New York: Harcourt, Brace & World.
- Wolff, Kurt H. 1953 “A Preliminary Inquiry into the Sociology of Knowledge from the Standpoint of the Study of Man.” In *Scritti di Sociologia e Politica in Onore di Luigi Sturzo*. Bologna: Nicola Zanichelli III.
- 1959 “The Sociology of Knowledge and Sociological Theory.” In L. Gross, ed., *Symposium on Sociological Theory*. New York: Harper & Row.

E. DOYLE MCCARTHY

## SOCIOLOGY OF LAW

In 1963, the U.S. Supreme Court ruled that in all felony trials the accused must be provided with

legal counsel. The case of *Gideon v. Wainwright* (372 U.S. 355, 1963) was widely celebrated as a David and Goliath story of the triumph of the rule of law: An indigent defendant’s handwritten petition had persuaded all nine justices of the Supreme Court to provide a nationwide right to counsel (Lewis 1964). Shortly after Gideon’s victory, Blumberg (1967) published an empirical case study describing the actual work of criminal defense attorneys. That study suggested that Gideon’s case had little relevance to the 90 percent of felony convictions that the prosecution wins not in a courtroom trial but through informal plea bargaining. Moreover, the attorneys to whom the poor were now constitutionally entitled, Blumberg contended, had over the years mutated from trial advocates into bureaucratic cogs whose primary function was to assist the state in processing legal files efficiently.

Blumberg’s deconstruction of the legal myth of the centrality of criminal trials and adversarial counsel exemplifies two central features of the sociology of law. First, that field challenges *legal formalism*, the philosophy that the law stands above social life, develops according to its own internal logic, and autonomously constrains or facilitates social interaction. A sociology of law becomes essential once the law’s dependence on its social organizational context is recognized (for a defense of legal formalism as a research agenda, see Watson 1985). Blumberg tried to show that the right to legal representation is contingent on the economics of legal services and the networks of dependency that link judges, prosecutors, and defense attorneys in ways that undermine the abstract legal model of the adversarial contest. Second, Blumberg’s case rests on observations of legal practice rather than interpretation of the texts of cases and legislation, the stock-in-trade of conventional legal scholarship. As empirical evidence continued to accumulate, Blumberg’s (1967) conclusions about the origins, causes, and consequences of plea bargaining were qualified or supplanted; later research suggests that plea bargains may be even more adversarial than trials ever were (Feeley 1997), that the relationship between caseload pressure and plea bargaining is complex (Holmes et al. 1992), and that the real role of the courtroom trial may be independent of its frequency of occurrence because out-of-court negotiations are conducted “in the shadow of the law” (Mnookin and Kornhauser

1979). Blumberg's study and the later work it inspired illustrate how the sociology of law examines empirical evidence to understand how law is created, enforced, and manipulated in the context of social organization.

### SOCIOLOGICAL VERSUS JURISPRUDENTIAL PERSPECTIVES ON LAW

The discipline of sociology does not hold a monopoly on efforts to unveil the connections between law and society. In the twentieth century, Roscoe Pound, Jerome Frank, and other legal scholars abandoned legal formalism and created new ways to understand the differences between the "law in the books" and the "law in practice" (for a concise overview of both developments, see Hunt 1978). Since the late 1970s, the critical legal studies movement and its variants have emerged as a major competitor to legal formalism in legal research and education (Kelman 1987). For example, Freeman (1978) examines how major Supreme Court decisions on civil rights have shifted the bases for legally defining discrimination from the consequences for the victims to the intentions of the perpetrators. Freeman shows how the law's emphasis on the actor's intention constrains the principle of equal protection and perpetuates inequality. While his conclusions are radical, his method is identical to the legal formalists' practice of textual interpretation (Trubek 1984). Critical legal studies' *doctrinal analysis*—its reliance on interpretation of constitutions, statutes, and judgments—has more affinity with literary criticism than with sociological methodologies based on the observation of events. For an example of this distinction, compare Klare (1978) with Wallace et al. (1988). Sociology of law is distinguished more by its methods than by its theories or subject matter.

### SOCIAL ORIGINS OF LAWS

A substantial number of historical case studies (e.g., Hall 1952) have traced the social origins of substantive and procedural law. Sociology enters these investigations with a broader comparative agenda, formulating and assessing general theories of the origin of law. Chambliss's (1964) analysis of six centuries of vagrancy laws as a ruling-class

manipulation of criminal law to control labor was a pioneer of contemporary efforts to pursue this line of investigation. Hagan (1980) provides a representative overview of the subsequent sociological analyses of the origins of alcohol and drug prohibition, sexual psychopathology and prostitution laws, and probation. Humphries and Greenberg (1981) produced one of the few sociological efforts showing the relationships among disparate legal changes and linking those changes to their social bases. They explain the diffusion of juvenile courts, probation, parole, and indeterminate sentences in terms of the shift in the political domination of corporate versus competitive capital during the Progressive era. An alternative approach to the study of the creation and diffusion of legal innovation looks to cultural transmission and organizational linkages rather than to underlying economic or social transformations. Grattet et al. (1998) show, for example, how the diffusion of hate crime legislation appears to be influenced by interstate processes of diffusion rather than by local conditions of the economy and society. Soule and Zylan (1997) similarly explore structural and diffusion factors in the reform of Aid to Families with Dependent Children (AFDC) eligibility rules. In terms of both theory and method, the sociology of law offers a rich body of work that reveals the social foundations of change in the law.

### SOCIAL STRATIFICATION OF LAW

The most prominent aspect of social structure in sociological investigations of law is stratification. In his early essay *On the Jewish Question*, Marx examined how a legal system that made all litigants equal before the law left them unequal in economic resources and social relationships. Much current research has been devoted to finding new evidence showing how formal legal equality reproduces social hierarchies. Galanter (1974) points out how the organizational properties of the legal system reinforce and in some instances generate inequality. Apart from the extralegal resources they bring to the dispute, repeat players (corporations and career criminals), for example, gain knowledge, skills that are not available, and networks denied to one-shot players. Feeley (1979) found that in a misdemeanor court "the process is the punishment": For the poor, the costs of conviction were minor compared to the costs imposed by

the pretrial stages of the process. Shapiro (1990) developed similar insights into the way in which the rules of evidence and organizational priorities of law enforcement bureaucracies create class differences in the punishment of white-collar crime. These studies go beyond the populist notion that the law is like a cobweb that catches the small flies but lets the large bugs go free. Individual resources matter, but sociological research shown how organizational and institutional contexts shape the manner in which equality *before* the law results in inequality *after* the law.

The largest body of research in the field has been devoted to the examination of discrimination in sentencing in criminal courts. Disparities in the type and duration of sanctions vary markedly by class, race, and/or ethnicity, and gender. For example, with 5 percent of the general population, young African-American males account for nearly half the admissions to state prisons. The initial research problem was to determine the extent to which such disparities represent differential involvement in the kinds of crime that lead to more severe sentences or reflect biases in discretionary decision making in the legal system (for a succinct overview of this research, see Walker et al. 1996). The sociologically relevant discoveries of this research include covariation in the extent of discriminatory decision making with social location (see Myers and Talarico 1987).

### **SOCIAL IMPACT OF LEGAL CHANGE**

*Brown v. Board of Education* (1954) is perhaps the most celebrated Supreme Court decision of the century. It marked the end of over half a century of the Court's acceptance of legalized racial segregation as being consistent with the constitutional requirement for equal protection under the law. It is usually the case one associates with the conviction that law—Supreme Court decisions, in particular—powerfully shapes social change. Less widely recognized is the fact that in the decade after *Brown*, racial segregation in public schools remained virtually unchanged. The sharpest challenge to conventional conceptions of the social impact of law is Rosenberg's (1991) study of the effect of Supreme Court decisions on school desegregation, abortion, reapportionment, and criminal procedure.

While current controversy centers on Rosenberg's thesis, several other research programs address the conditions under which legal reforms engineer social change. Burstein (1985), for example, specifies the contingencies that influenced the impact of civil rights legislation on the economic position of minorities. Horney and Spohn (1991) examine the impact of rape reform laws in six jurisdictions on several indicators of prosecution. The measurable impact of legal reform proved to be limited, because of the response of local court organizations to externally imposed change. Heimer (1995) illustrates that similar complications appear when legal changes are imposed on hospital work groups. Organizational responses occasionally facilitate rather than inhibit change. Edelman et al. (1992), for example, found that personnel departments tend to exaggerate the legal risks of noncompliance in equal-opportunity cases as a way to enhance their power within the corporation.

### **ORGANIZATIONAL CONTEXTS OF LEGAL PROCEDURES**

A public defender explained to Sudnow (1965) that to work in such an office, one has to know the law—and the ropes. Learning about the organizational ropes of courts, police departments, and law offices has been the objective of a large body of contemporary research in the sociology of law.

Albonetti (1987) utilizes organizational theories to explain variation in the decisions of prosecutors to drop cases or reduce charges; apart from the legal evidence, prosecutors' decisions are shaped by extralegal factors that govern their uncertainty about winning a case at trial. Ofshe and Leo (1997) investigate the coercive persuasion that continues to occur in post-Miranda police interrogations. Police investigators generally follow the letter of the Miranda rules while continuing to practice forms of coercive persuasion that induce most suspects to waive their rights and confess.

Many discoveries about procedure turn on the emergence of informal organizational rules and relationships. Sudnow (1965) found that plea bargains were forged in a common currency of offense seriousness that existed apart from the penal code's definitions of crimes and punishments.

Emerson (1969) showed how the legally relevant aspects of a juvenile's offense and career are organizationally transformed into judgments of character, which then become the real bases for determining verdicts and imposing sentences. This work suggests that due process is a variable whose appearance and effects are shaped by organizational contexts (see Dobbin et al. 1988).

### THE ROLE OF GENERAL THEORY IN THE SOCIOLOGY OF LAW

The sociology of law can be distinguished from economics, psychology, and other social science enterprises that have law as their subject matter principally in terms of its integration of its investigations with general theories of social structure. The role of general theory becomes apparent, for example, in comparisons of Japanese and U.S. legal systems that "explain away Japan by attributing every finding to 'Japanese uniqueness' [rather than] treat Japan as a point on a universal continuum" (Miyazawa 1987, p. 239). The case for engaging in the search for such universal continua is made by Black (1976, 1997).

Much current research, however, continues to be guided by one or a combination of the four general theories that initially defined the field. Bentham's utilitarian philosophy underlies rational choice theories of the behavior of law. Studies of deterrence at both individual and organizational levels of analysis continue to pursue this line of theorizing (see Vaughn 1998 for a summary and critique of organizational analysis). Alternatively, the sociological theories of Marx, Durkheim, and Weber articulate properties of social organization that shape and constrain the choices of persons and firms (for an overview, see Garland 1990). Work in the sociology of law thus not only illuminates the institution of law in unique ways but contributes more fundamentally to basic knowledge about human social organization.

#### REFERENCES

Albonetti, Celesta A. 1987 "Prosecutorial Discretion: The Effects of Uncertainty." *Law and Society Review* 21:291-313.

Black, Donald 1976 *The Behavior of Law*. New York: Academic Press.

— 1997 *The Social Structure of Right and Wrong*. Revised edition. San Diego: Academic Press.

Blumberg, Abraham S. 1967 "The Practice of Law as a Confidence Game: Organization Cooptation of a Profession." *Law and Society Review* 1:15-39.

Burstein, Paul 1985 *Discrimination, Jobs and Politics: The Struggle for Equal Employment Opportunity in the U.S. Since the New Deal*. Chicago: University of Chicago Press.

Chambliss, William J. 1964 "A Sociological Analysis of the Law of Vagrancy." *Social Problems* 12:67-77.

Dobbin, Frank R., Lauren Edelman, John W. Myer, W. Richard Scott, and Ann Swidler 1988 "The Expansion of Due Process in Organizations." In Lynne G. Zucker, ed., *Institutional Patterns and Organizations: Culture and Environment*. Cambridge, Mass.: Ballinger.

Edelman, Lauren B., Steven E. Abraham, and Howard S. Erlanger 1992 "Professional Construction of Law: The Inflated Threat of Wrongful Discharge." *Law and Society Review* 26:47-83.

Emerson, Robert M. 1969 *Judging Delinquents: Context and Process in Juvenile Court*. Chicago: Aldine.

Feeley, Malcolm M. 1979 *The Process Is the Punishment: Handling Cases in a Lower Court*. New York: Russell Sage Foundation.

— 1997 "Legal Complexity and the Transformation of the Criminal Process: The Origins of Plea Bargaining." *Israel Law Review* 31:183-222.

Freeman, Alan D. 1978 "Legitimizing Racial Discrimination through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine." *Minnesota Law Review* 62:1049-1119.

Galanter, Marc 1974 "Why the 'Haves' Come Out Ahead." *Law and Society Review* 9:95-160.

Garland, David 1990 *Punishment and Modern Society: A Study in Social Theory*. Chicago: University of Chicago Press.

Grattet, Ryken, Valerie Jenness, and Theodore R. Curry 1998 "The Homogenization and Differentiation of Hate Crime Law in the United States, 1978 to 1995: Innovation and Diffusion in the Criminalization of Bigotry." *American Sociological Review* 63:286-307.

Hagan, John 1980 "The Legislation of Crime and Delinquency: A Review of Theory, Method, and Research." *Law and Society Review* 14:603-628.

Hall, Jerome 1952 *Theft, Law and Society*, 2nd ed. Indianapolis: Bobbs-Merrill.

Heimer, Carol A. 1995 "Explaining Variation in the Impact of Law: Organizations, Institutions, and Professions." *Studies in Law, Politics and Society* 15:29-59.

- Holmes, Malcom D., Howard C. Daudistel, and William A. Taggart 1992 "Plea Bargaining Policy and State District Court Caseloads: An Interrupted Time Series Analysis." *Law and Society Review* 26:139-159.
- Horney, Julie, and Cassia Spohn 1991 "Rape Law Reform and Instrumental Change in Six Urban Jurisdictions." *Law and Society Review* 25:117-154.
- Humphries, Drew, and David Greenberg 1981 "The Dialectics of Crime Control." In David Greenberg, ed., *Crime and Capitalism* Palo Alto, Calif.: Mayfield.
- Hunt, Alan 1978 *The Sociological Movement in Law*. Philadelphia: Temple University Press.
- Kelman, Mark 1987 *A Guide to Critical Legal Studies*. Cambridge, Mass.: Harvard University Press.
- Klare, Karl E. 1978 "Judicial Deradicalization of the Wagner Act and the Origins of Modern Legal Consciousness, 1937-1941." *Minnesota Law Review* 62:265-339.
- Lewis, Anthony 1964 *Gideon's Trumpet*. New York: Random House.
- Miyazawa, Setsuo 1987 "Taking Kawashima Seriously: A Review of Japanese Research on Japanese Legal Consciousness and Disputing Behavior." *Law and Society Review* 21:219-241.
- Mnookin, Robert H., and Lewis Kornhauser 1979 "Bargaining in the Shadow of the Law: The Case of Divorce." *Yale Law Journal* 88:950-997.
- Myers, Martha, and Susette M. Talarico 1987 *The Social Contexts of Criminal Sentencing*. New York: Springer-Verlag.
- Ofshe, Richard J., and Richard A. Leo 1997 "The Social Psychology of Police Interrogation: The Theory and Classification of True and False Confessions." *Studies in Law, Politics and Society* 16:189-251.
- Rosenberg, Gerald N. 1991 *The Hollow Hope: Can Courts Bring about Social Change?* Chicago: University of Chicago Press.
- Shapiro, Susan 1990 "Collaring the Crime, Not the Criminal: Reconsidering the Concept of White Collar Crime." *American Sociological Review* 55:346-365.
- Soule, Sarah A., and Yvonne Zylan 1997 "Runaway Train?: The Diffusion of State-Level Reform in ADC/AFDC Eligibility Requirements, 1950-1967." *American Journal of Sociology* 103:733-762.
- Sudnow, David 1965 "Normal Crimes: Sociological Features of the Penal Code in a Public Defender's Office." *Social Problems* 12:255-276.
- Trubek, David 1984 "Where the Action Is: Critical Legal Studies and Empiricism." *Stanford Law Review* 36:575-622.
- Vaughn, Dianne 1998 "Rational Choice, Situated Action, and the Social Control of Organizations." *Law and Society Review* 32:23-61.
- Walker, Samuel, Cassia Spohn, and Miriam DeLorme 1996 *The Color of Justice: Race, Ethnicity and Crime in America*. Belmont, Calif.: Wadsworth.
- Wallace, Michael, Beth A. Rubin, and Brian T. Smith 1988 "American Labor Law: Its Impact on Working-class Militancy, 1901-1980." *Social Science History* 12:1-29.
- Watson, Alan 1985 *The Evolution of Law*. Baltimore: Johns Hopkins University Press.

JAMES M. INVERARITY

## SOCIOLOGY OF RELIGION

An important intellectual shift has taken place in the social scientific study of religion as many of its longest held theoretical positions, passed down from the founders of the field, have been overturned. These changes have been so dramatic and far-reaching that Warner (1993, p. 1044) identified them "as a paradigm shift in progress," an assessment that since that time "has been spectacularly fulfilled" (Greeley 1996, p. 1).

Typically, the emergence of a new paradigm rests on both an empirical basis and a theoretical basis. Over the past thirty years, there has been an explosion of research on religious topics and a substantial number of new facts have accumulated. The bulk of these discoveries have turned out to be inconsistent with the old paradigm. In response to the growing incompatibility between fact and traditional theory, new theories have been constructed to interpret the empirical literature.

There are five major points of dispute between the old and new paradigms. In this article, each one is described, followed by a brief summary of the pertinent evidence. Finally, additional recent trends in the field are noted.

### RELIGION IS HARMFUL

For nearly three centuries, social scientists condemned religion as harmful to the individual because it impedes rational thought and harmful to society because it sanctifies tyrants (Stark 1999b). The premise that religion is irrational and psychologically harmful has taken many forms, all of them notable for the open contempt and an-

tagonism they express toward faith. Thus, as Freud explained on one page of his psychoanalytic exposé of faith, *The Future of an Illusion* (1961 [1927], p. 88), religion is an “illusion,” a “sweet—or bitter-sweet—poison,” a “neurosis,” an “intoxicant,” and “childishness to be overcome.” More recently, Carroll (1987, p. 491) claimed that praying the Rosary is “a disguised gratification of repressed anal-erotic desires,” a substitute for playing “with one’s feces.” In a similar fashion, Ostow (1990, p. 113) asserted that evangelical Protestantism is a matter of regression “to the state of mind of the child who resists differentiation from its mother. The messiah and the group itself represent the returning mother.”

In rejecting assertions that religion is rooted in irrationality, proponents of the new paradigm cite a growing literature that finds religion to be a reliable source of better mental and even physical health (Ellison 1991, 1993; Idler and Kasl 1997; Levin 1996; Pargament and Park 1995). Two literature reviews published in 1987, pointed to the positive health effects of religious involvement regardless of the age, sex, race, ethnicity, or nationality of the population being studied (Jarvis and Northcutt 1987; Levin and Schiller 1987). In a more recent review, Levin (1996, p. 850) found that that relationship still holds and suggests that these results point to a “protective epidemiologic effect of religiosity.”

In the field of gerontology, of research on religion and aging has grown so rapidly that a new journal (*Journal of Religious Gerontology*) has emerged and older journals have devoted special issues or sections to discussions of the topic. Krause (1997, p. S291) summarized the literature: “[A]n impressive body of research indicates that elderly people who are involved in religion tend to enjoy better physical and mental health than older adults who are not as religious.”

Not only is religion associated with better mental and physical health, all the current theorizing about religion accepts the rational choice principle as its first axiom (Gill 1998; Greeley 1995; Iannaccone 1990, 1995a, 1995b; Miller 1995; Sherkat 1997; Stark 1996a, 1996b, 1999a; Stark and Bainbridge 1980 [1987], 1996; Stark and Finke 2000; Stark and Iannaccone 1993, 1994). Most of these scholars do not employ the “thin” version of

rational choice currently used in economics (Iannaccone 1995b) but the “thick” version (Ferejohn 1991), similar to what Weber ([1922] 1993, p. 1) had in mind when he wrote that

*religiously or magically motivated behavior is relatively rational behavior . . . It follows rules of experience. . . . Thus, religious and magical behavior or thinking must not be set apart from the range of everyday purposive conduct . . .*

What about the harmful *social* effects of religion as it sustains the powerful and dispenses false consciousness to the exploited and debased? Engels (Marx and Engels 64 195, p. 316) claimed that early Christianity “first appeared as a religion of slaves and emancipated slaves, of poor people deprived of all rights, of peoples subjugated and dispersed by Rome.” Does it not follow that *religion appeals most strongly to the lower classes?*

While the old paradigm identified religion as the opiate of the people, the new paradigm notes that religion also often is the “amphetamines” of the people in that religion animated many medieval peasant and artisan rebellions (Cohn 1961), generated repeated uprisings among the native peoples of Africa and North America against European encroachment (Wilson 1975), and recently served as a major center of mobilization against tyranny in eastern Europe (Echikson 1990). The notion that religion primarily serves to compensate the deprived and dispossessed has become untenable. The consensus among scholars rejects as “imaginary history” Engels’s notion that the early Christian movement was rooted in proletarian suffering. The facts force the conclusion that Christianity’s greatest early appeal was to the privileged classes (Stark 1996a). In similar fashion, since the early 1940s many researchers have attempted to connect religiousness to social class, but their findings have been weak and inconsistent (Stark and Finke 2000). Consequently, the need for new theorizing on the role of religion in the political affairs of nations has been recognized (Gill 1998).

## RELIGION IS DOOMED IN MODERN TIMES

As the social sciences emerged in the wake of the Enlightenment, the leading figures eagerly pro-



claimed the demise of religion. Toqueville wrote in his famous early nineteenth-century study, *Democracy in America* ([1840] 1956, vol. II, p. 319):

*The philosophers of the eighteenth century explained in a very simple manner the gradual decay of religious faith. Religious zeal, said they, must necessarily fail the more generally liberty is established and knowledge diffused.*

This came to be known as the *secularization thesis*: In response to modernization, “religious institutions, actions, and consciousness, [will] lose their social significance” (Wilson 1982, p. 149). Toqueville was virtually alone in rejecting the secularization thesis; perhaps no other social scientific prediction enjoyed such nearly universal acceptance for so long. Thus, the anthropologist Wallace (1966, p. 265) wrote in an undergraduate textbook:

*The evolutionary future of religion is extinction. Belief in supernatural beings and supernatural forces that affect nature without obeying nature’s laws will erode and become only an interesting historical memory . . . Belief in supernatural powers is doomed to die out, all over the world, as the result of the increasing adequacy and diffusion of scientific knowledge.*

In the late 1990s, the secularization thesis has been buried under a mountain of contrary facts (Bossy 1985; Duffy 1992; Greeley 1989, 1995, 1996; Murray 1972; Stark 1999c). The primary empirical basis for claims of ongoing secularization has been the very low rates of religious participation in contemporary European nations, where weekly rates of church attendance often are below 5 percent. However, the overwhelming weight of historical research shows that these low rates do not represent a decline. Church attendance always was extremely low in those nations, and it is not clear that they ever were effectively Christianized (Greeley 1995; Stark 1999c). Furthermore, in those nations the overwhelming majority express firm belief in the supernatural, pray, and describe themselves as religious. It is perverse to describe a nation as highly secularized (as those committed to the old paradigm still do) when two-thirds or more of its residents say they are “religious persons” and fewer than 5 percent say they are atheists. The interesting question thus does not con-

cern secularization but is, “[W]hy are these societies of believing non-belongers?” as Davie (1994) has expressed it. What is it about the churches in those nations that prevents them from mobilizing participation?

Looking to the world as a whole, there is no consistent relationship between religious participation and modernization. Indeed, the very few significant, long-term declines in religious participation that have been seen in the world are greatly outnumbered by remarkable increases (Stark and Finke 2000). What needs to be explained, therefore, is not religious decline but *variation*. Finally, the spread of science cannot cause secularization, because science and religion are unrelated. Scientists are as religious as anyone else, and the more scientific their fields, the more religious are American academics (Stark et al. 1996).

#### ONE LAST SPASM

The twin propositions that religious behavior is rooted in irrationality and that religion must soon yield to secularization have been dealt a blow by the finding that the more liberal (or secularized) a religious body becomes, the more rapidly it loses members, while denominations that sustain more vigorous and traditional theologies have prospered (Finke and Stark 1992; Iannaccone 1994; Kelley 1972; Stark and Finke 2000). How can it be that the “fundamentalists” grow while the liberals lose out? Proponents of the old paradigm have invoked the notion that this is but one final, dying spasm of piety. They claim that the expansion of evangelical Protestant churches in the United States (and presumably in Latin America, where they are experiencing explosive growth) is a frantic “flight from modernity,” that people who feel threatened by the erosion of traditional morality are flocking to religious havens (Berger 1967; Hunter 1987, 1983). Berger (1967, p. 11) described American evangelical Protestant churches as follows: “They are like besieged fortresses, and their mood tends toward a militancy that only superficially covers an underlying sense of panic.” Nearly thirty years later Thurow (1996, p. 232) explained, “Those who lose out economically or who cannot stand the economic uncertainty of not knowing what it takes to succeed in the new era ahead retreat into religious fundamentalism.”

A fatal problem with this explanation is that, as was noted above, the relationship between social class and religiousness is weak and inconsistent. Conservative churches actually include a fair share of highly educated, successful, and sophisticated people who display no apparent fears of modernity (Smith 1998, 2000; Stark and Finke 2000; Woodberry and Smith 1998).

The new paradigm has no difficulty explaining the growth of evangelical churches because it does not confuse price with value. As Iannaccone (1994, 1992) has demonstrated, "strict" churches—those which require more from their members—are a better value because they offer far more in the way of rewards, both worldly and otherworldly. In this sense, to opt for a more traditional religious affiliation is to make the more rational choice, in that it yields a greater ratio of rewards over costs.

#### IDEALISTIC HUMBUG

Generations of social scientists have embraced the notion that religion is a dependent variable and that whatever appears to be a religious effect is ultimately merely a mask for something more basic, something "material."

Although social scientists in most other areas of study have long acknowledged the truism that if people define something as real, it can have real consequences, this concession usually has been denied in the area of religion. Instead, there has been a general willingness to agree with Marx that any attempt to explain "reality" by reference to an unreality such as religion is "idealistic humbug." Rather, one must explain religion by reference to "realities" such as "the mode of production." That is, one "does not explain practice from the idea but explains the formation of ideas from material practice" (Marx [1845] 1998, p. 61). As Marx's collaborator Engels explained, "All religion . . . is nothing but the fantastic reflection in men's minds of those external forces which control their daily life . . . the economic conditions . . . the means of production" (Marx and Engels 1964, pp. 147-148).

These views did not originate with Marx; they have been nearly universal among social scientists for close to three centuries (Stark 1999b). Even Weber, having attributed the rise of capitalism to

the "Protestant ethic," traced the source of that ethic to material conditions (including the rise of the bourgeoisie, population growth, and colonialism), thus limiting Calvinist doctrines to being at most a *proximate* rather than a *fundamental* cause of capitalism. Even so, Weber has been bitterly criticized for affording religion *any* causal role. Emile Durkheim and his functionalist heirs dismissed religious belief as an insignificant epiphenomenon, regarding ritual as the only active religious ingredient and as being only a proxy for a more basic factor, social solidarity (Stark and Bainbridge 1997).

The new paradigm is committed to the proposition that people often act from religious motives and that in many cases no more fundamental or material cause can be found. Four historical examples reveal the conflict between paradigms on this central issue.

**Crusading for Land and Loot.** For centuries, historians believed that the Crusades to the Holy Land were motivated by faith, that tens of thousands of European nobles and knights marched to the Holy Land to rescue it from Muslim "desecration." However, by the end of the nineteenth century social scientists had penetrated those appearances to discover that the crusaders really went in pursuit of land and loot. Having summarized the many economic problems facing Europe in the eleventh century, including the population pressures and land shortages that were said to beset the knightly class, Mayer (1972, pp. 22-25) stressed the "lust for booty" and the "hunger for loot" that motivated the crusaders: "Obviously the crusade acted as a kind of safety valve for a knightly class which was constantly growing in numbers." He went on to emphasize the need to recognize "the social and economic situation of a class which looked upon the crusade as a way of solving its material problems."

Although there is extensive evidence that the crusaders truly believed they were going for purely religious reasons, this material can be ignored because there exists a definitive refutation of the materialist position. In 1063, thirty-two years before Urban II called for the First Crusade to the Holy Land, Pope Alexander II, backed by the evangelical efforts of the monks of Cluny, attempted to organize a Crusade to reclaim Moorish Spain.

Here, very close at hand, lay great wealth and an abundance of fertile land, and the Pope had declared that all who fought for the Cross in Spain were entitled not only to absolution for their sins but to all the wealth and “lands they conquered from the infidel” (Runciman 1951, vol. 1, p. 91). However, hardly anyone responded, and little or nothing was achieved. The materialist interpretation of the Crusades fails when faced with the fact that crusaders were not lured to nearby Spain in pursuit of rich and relatively easy pickings, while soon afterward tens of thousands set off for the dry wastes of faraway Palestine and did so again and again. Why did they do that rather than go to Spain? Because Spain was not the Holy Land. Jesus had not walked the streets of Toledo or been crucified in Seville.

**Heresy and Class Struggle.** Beginning in the eleventh century and lasting through the sixteenth, Europe was swept by mass heretical movements—Waldensians, Cathars (called Albigensians in southern France), Hussites, and many others—culminating in the Reformation. Tens of thousands died on behalf of their religious beliefs, but maybe not.

Many historians possessed of an excessive sociological imagination have claimed that these great heretical movements were not primarily about doctrines and morals, if indeed religious factors were of any significance at all. Instead, they argue, the religious aspect of these movements masked their real basis, which was of course class struggle. Engels (Marx and Engels 1964, pp. 97–123) identified some of these movements, including the Albigensians, as urban heresies in that they represented the class interests of the town bourgeoisie against those of the feudal elites of church and state. But most of the heretical movements were, according to Engels, based on the proletariat, which demanded restoration of the equality and communalism of early Christianity (Engels and many other Marxists have claimed that the early Christians briefly achieved true communism). Engels granted that these class struggles were characterized by religious and mystical rhetoric but dismissed this as false consciousness. Following Engels, many Marxist historians have “exposed” the materialism behind the claims of religious dissent. Thus, in 1936 the Italian historian Antonino de Stefano claimed, “At bottom, the economic argument must have constituted, more than any

dogmatic or religious discussions, the principle motive of the preaching of heresy” (quoted in Russell 1965, p. 231). Even many historians not committed to orthodox Marxism have detected materialism behind medieval dissent. For example, the non-Marxist historian Cohn (1961, p. xiii) reduced medieval heresies to “the desire of the poor to improve the material conditions of their lives,” which “became transfused with phantasies of a new Paradise.”

It is not necessary for proponents of the new paradigm to deny that class conflicts existed in medieval times or to suppose that people participating in heresy never paid any heed to their material interests to reaffirm that religion lay at the heart of these conflicts. If their primary concerns had been worldly, surely most heretics would have recanted when that was the only way out. It was, after all, only their religious notions they had to give up, not their material longings. However, large numbers of them chose death instead. Moreover, these movements drew participants from all levels of the class system. The Albigensians, for example, enlisted not only the bourgeoisie but, in contradiction to Engels, most of the nobility as well as the clergy of southern France and indeed the “masses” (Costen 1997; Lambert 1992; Mundy 1985). Finally, the claim that the majority of the participants in any given heresy consisted of peasants and the poor is lacking in force, even in the instances in which it might be true. Almost everyone in medieval Europe was poor and a peasant. Gauged against this standard, it seems likely that the “proletarian masses” were quite underrepresented in most of these movements (Lambert 1992).

**Medieval Jewish Messianic Movements.** For Jews the messiah has yet to come, but again and again over the centuries, groups of Jews have hailed his arrival. An early episode resulted, of course, in Christianity, but it would not be an exaggeration to say that hundreds of other messianic movements have occurred in Jewish communities over the past two millennia, and such movements were especially common in the European diaspora during medieval times (Cohen 1967; Lenowitz 1998; Sharot 1982).

In a sophisticated analysis of these religious movements, Sharot (1982, p. 18) noted the huge literature that stresses that messianic movements are

*responses to the disruption of social and cultural patterns . . . [produced by] a disaster such as an epidemic, famine, war, or massacre. Following a disaster, persons feel vulnerable, confused, full of anxiety, and they turn to millennial beliefs in order to account for otherwise meaningless events. They interpret the disaster as a prelude to the millennium; thus their deepest despair gives way to the greatest hope.*

Although some messianic Jewish movements did erupt after a disaster, as he worked his way through all the better-known cases, Sharot (1982, pp. 65–66) was forced to agree with Cohen's (1967) earlier study that many movements seemed to come out of nowhere in the sense that they arose during periods of relative quiet and therefore that "disaster was not a necessary condition of a messianic outburst."

Sharot made this concession very reluctantly, and often seems to forget it. Nevertheless, his scrupulous accounts of specific incidents frequently show that a movement was the direct result of religious rather than secular influences. In many cases, an episode began with an individual or small group poring over the *Kabbalah* (a collection of Jewish mystical writings) out of purely personal motives and then "discovering" that the millennium was at hand. Thereafter, they shared this knowledge with others, who in turn assisted in arousing a mass following. In other instances, someone became convinced that he was the messiah and was able to convince his family and friends (Stark 1999d).

One can of course argue that Jews in medieval Europe were always victims and hence always ripe for millenarian solutions. However, constants cannot explain variations, and in as many cases as not, nothing special was going on to cause a movement to arise then rather than at some other time except for direct religious influences in the form of people advocating a new religious message or circumstance.

Of course, people often do turn to religion in times of trouble and crisis, but the new paradigm rejects the claim that crises are a necessary condition for religious innovations and recognizes that religious phenomena can be caused by other religious phenomena.

**The Mystical 1960s.** A huge literature attributes the "explosive growth" of new religious movements in the United States in the late 1960s and early 1970s to profound social causes. Particular attention has been given to uncovering the secular causes of the special appeal of Eastern faiths for Americans in that period. Cox (1983, p. 42) blamed "the most deteriorated, decadent phase of consumer capitalism," charging that converts to Eastern faiths had "been maddened by consumer culture" (p. 40). Serious journals published equally hysterical explanations. As Robbins summarized (1988, p. 60), each of these analyses identified one or more "acute and distinctively modern dislocation which is said to be producing some mode of alienation, anomie or deprivation to which Americans are responding." With a fine grasp of the essentials, Barker (1986, p. 338) commented that "those who have read some of the sociological literature could well be at a loss to understand why *all* young adults are not members [of new religious movements], so all-encompassing are some of the explanations."

In fact, there was no growth, explosive or otherwise, of new religious movements in this era (Melton 1988; Finke and Stark 1992); the rate of new movement formation was constant from 1950 through 1990. As for the brief increase in the proportion of Eastern faiths among new American movements, capitalism had nothing to do with it. Rather, in 1965 the elimination of exclusionary rules against Asian immigration made it possible for the first time for authentic Eastern and Indian religious leaders to seek American followers directly. Consequently, there was an increase in the number of Eastern religious organizations, but the number of actual converts was minuscule. Even so, these movements were the result of *religious* efforts, of face-to-face recruitment activities motivated by the religious convictions of missionizing gurus.

#### THE EVILS OF PLURALISM

More than three centuries ago, early scholars of comparative religion assumed that by publicizing the beliefs of the world's many faiths, they could advance the cause of atheism, that by virtue of their competing claims, each religion would refute the others (Preus 1987). This view has led to the

claim that faith is a very fragile thing that cannot survive challenge; hence, *pluralism*—the existence of several competing religious bodies in a society—is said to be incompatible with strong religiosity. Durkheim ([1897] 1951, p. 159) asserted that when multiple religious groups compete, religion becomes open to question, dispute, and doubt and thus “the less it dominates lives.” Eventually these views were formulated into elegant sociology by Berger (1967, 1979), who repeatedly argued that pluralism inevitably destroys the plausibility of all religions and only where one faith prevails can there exist a “sacred canopy” that is able to inspire universal confidence and assent.

These notions are mistaken, having been taken over uncritically from the justifications given by European state churches for their monopolies. It is indicative of the undue respect given European social science that American sociologists accepted this view, since religious competition is an obvious basis for the extraordinary levels of religious participation in the United States, in contrast to the religious apathy prevalent in societies with a monopoly church. Indeed, the positive role of competition is obvious in American history. In 1776, when most American colonies were dominated by a state-supported church, about one person in five belonged to any church. After the Revolution, the onset of vigorous religious competition eventually resulted in about two-thirds of Americans belonging to a church (Finke and Stark 1992)

To fully appreciate the power of pluralism, it was necessary to cease treating religion as primarily a psychological phenomenon and take a more sociological view, an approach that also has been characteristic of the new paradigm. The concept of a *religious economy* (Stark 1985) made it possible to adopt an overall perspective on the religious activities in a society and examine the interplay among religious groups. This analysis quickly revealed that the main impact of religious competition on individuals is not confusion or the corrosion of faiths but to present the individual with vigorously offered choices. As Adam Smith pointed out more than two centuries ago, monopoly religions are as subject to laziness and inefficiency as are monopoly business firms. Thus, it is axiomatic in the new paradigm that religious competition strengthens religion because as firms vie for supporters, they tend to specialize their ap-

peals, with the overall result that a higher proportion of the population will be enrolled. As of 1999 there had been more than twenty-five published studies based on many different societies and different eras, offering overwhelming support for this view (Finke and Stark 1998; Stark and Finke 2000).

### A FOCUS ON RELIGION

Despite emphasizing that religion does have effects, the new paradigm is not limited to that perspective. Rather, in addition to a sociology of religious effects, the new paradigm has promulgated a sociology of religion *per se*.

For a long time sociologists interested in religion attempted to justify their topic by demonstrating its importance to those who specialized in one of the more of the “secular” areas of the field. Thus, some sociologists devoted studies to demonstrating religious effects on political behavior such as voting and opinions on current issues. Others sought to convince demographers that religion was crucial to fertility studies. This trend has been enshrined in textbooks on the sociology of religion, all of which have consisted almost entirely of chapters on “religion and family,” “religion and economics,” “religion and prejudice,” and so on.

However, having become part of a relatively large and well-established specialty, sociologists in this area have become sufficiently confident to make religion the real center of study rather than trying to draw legitimacy from its connections to other topics. Consequently, there has been renewed attention to what religion is as well as what it does (Boyer 1994; Greeley 1995; Guthrie 1996; Stark 1999a). There also is much new work on religious and mystical experiences (Hood 1997; Howell 1997; Neitz and Spickard 1990; Stark 1999d). Other scholars have focused not on the causes or consequences of prayer but on its nature and practice (Poloma and Gallup 1991; Swatos 1987). Also, increasing attention is being paid to images of God (Barrett 1998; Greeley 1995; Stark forthcoming).

In addition, there is an impressive new literature on religious socialization (Ellison and Sherkat 1993a, 1993b; Granqvist 1998; Kirkpatrick and Shaver 1990; Smith 1998), on denominational switching (Musick and Wilson 1995; Perrin et al.

1997; Sherkat and Wilson 1995), and on conversion (Hall 1998; Rambo 1993; Stark and Finke 2000). Amid all this activity, the case study literature is blooming as never before (Davidman 1991; Goldman 1999; Heelas 1996; Lang and Ragvald 1993; Lawson 1995, 1996, 1998; Neitz 1987; Poloma 1989; Washington 1995).

REFERENCES

- Barker, Eileen 1986 "Religious Movements: Cult and Anti-Cult Since Jonestown." *Annual Review of Sociology* 12:329-346.
- Barrett, Justin L. 1998 "Cognitive Constraints on Hindu Concepts of the Divine." *Journal for the Scientific Study of Religion* 37:608-619.
- Berger, Peter 1967 *The Sacred Canopy*. New York: Doubleday.
- 1979 *The Heretical Imperative: Contemporary Possibilities of Religious Affiliation*. New York: Doubleday.
- Bossy, John 1985 *Christianity in the West: 1400-1700*. New York: Oxford University Press.
- Boyer, Pascal 1994 *The Naturalness of Religious Ideas: A Cognitive Theory of Religion*. Berkeley: University of California Press.
- Carroll, Michael P. 1987 "Praying the Rosary: The Anal-Erotic Origins of a Popular Catholic Devotion." *Journal for the Scientific Study of Religion* 26:486-498.
- Cohen, Gershon D. 1967. "Messianic Postures of Ashkenazim and Sephardim (Pior to Sabbatai Zevi)." In Max Kreutzberger, ed., *Studies of the Leo Baeck Institute*. New York:
- Cohn, Norman 1961 *The Pursuit of the Millennium*. New York: Harper and Row.
- Costen, Michael 1997 *The Cathars and the Albigensian Crusade*. Manchester, UK: Manchester: University Press.
- Cox, Harvey 1983 "Interview." In Steven J. Gelberg, ed., *Hare Krishna, Hare Krishna*. New York: Grove Press.
- Davidman, Lynn 1991 *Tradition in a Rootless World: Women Turn to Orthodox Judaism*. Berkeley: University of California Press.
- Davie, Grace. 1994. *Religion in Britain Since 1945: Believing without Belonging*. Oxford:Blackwell.
- Duffy, Eamon. 1992. *Stripping of the Altars*. New Haven, Conn.: Yale University Press.
- Durkheim, Emile (1897) 1951 *Suicide*. Glencoe, Ill.: Free Press.
- Echikson, William 1990 *Lighting the Night: Revolution in Eastern Europe*. New York: Morrow.
- Ellison, Christopher G. 1991 "Religious Involvement and Subjective Well-Being." *Journal of Health and Social Behavior* 32:80-99.
- 1993. "Religion, the Life Stress Paradigm, and the Study of Depression." In Jeffrey S. Levin, ed., *Religious Factors in Aging and Health: Theoretical Foundations and Methodological Frontiers*. Newbury Park, Calif.: Sage
- Ferejohn, John A. 1991 "Rationality and Interpretation: Parliamentary Elections in Early Stuart England." In Kristen Renwick Monroe, ed., *The Economic Approach to Politics: A Critical Reassessment of the Theory of Rational Action*. New York: HarperCollins.
- 1992 *The Churching of America, 1776-1990: Winners and Losers in our Religious Economy*. New Brunswick, N. J.: Rutgers University Press.
- Finke, Roger, and Rodney Stark 1998 "Religious Choice and Competition." *American Sociological Review*. 63:761-766.
- Freud, Sigmund (1927) 1961 *The Future of an Illusion*. Garden City, N.Y.: Doubleday.
- Gill, Anthony J. 1998 *Rendering Unto Caesar: The Roman Catholic Church and the State in Latin America*. Chicago: University of Chicago Press.
- Goldman, Marion 1999 *Passionate Journeys: Why Successful Women Joined a Cult*.
- Granqvist, Pehr 1998 "Religiousness and Perceived Childhood Attachment: On the Question of Compensation or Correspondence." *Journal for the Scientific Study of Religion*. 37:350-367.
- Greeley, Andrew M. 1989 *Religious Change in America*. Cambridge, Mass.: Harvard University Press.
- 1995 *Religion as Poetry*. New Brunswick, N.J.: Transaction.
- 1996 "The New American Paradigm: A Modest Critique." Paper read at the German Sociological Association annual meetings, Cologne.
- Guthrie, Stewart Elliott 1996 "Religion: What Is It?" *Journal for the Scientific Study of Religion*. 35:412-419.
- Heelas, Paul 1996 *The New Age Movement*. Oxford: Blackwell.
- Hood, Ralph W., Jr. 1997 "The Empirical Study of Mysticism." In Bernard Spilka and Daniel N. McIntosh, eds., *The Psychology of Religion: Theoretical Approaches*. Boulder, Colo. Westview Press.
- Howell, Julia Day 1997 "ASC Induction Techniques, Spiritual Experiences, and Commitment to New Religious Movements." *Sociology of Religion*. 58:141-164.
- Hunter, James Davison 1983 *American Evangelicalism: Conservative Religion and the Quandary of Modernity*. New Brunswick, N.J.: Rutgers University Press.

- 1987 *Evangelicalism: The Coming Generation*. Chicago: University of Chicago Press.
- Iannaccone, Laurence R. 1990 "Religious Practice: A Human Capital Approach." *Journal for the Scientific Study of Religion*. 29:297-314.
- 1992 "Sacrifice and Stigma: Reducing Free-Rising in Cults, Communes, and other Collectives." *Journal of Political Economy*. 100(2):271-292.
- 1994 "Why Strict Churches Are Strong." *American Journal of Sociology*. 99:1180-1211.
- 1995a "Risk, Rationality, and Religious Portfolios." *Economic Inquiry*. 33:285-295.
- 1995b "Voodoo Economics? Reviewing the Rational Choice Approach to Religion." *Journal for the Scientific Study of Religion*. 34:76-89.
- Idler, Ellen L. and Stanislav V. Kasl 1997b "Religion among Disabled and Nondisabled Persons: II. Attendance at Religious Services as a Predictor of the Course of Disability." *Journal of Gerontology* 52B(6):S306-316.
- Jarvis, G. K. and H. C. Northcutt 1987 "Religion Differences in Morbidity and Mortality." *Social Sciences and Medicine* 25:813-824.
- Kelley, Dean M. 1972 *Why Conservative Churches Are Growing*. New York: Harper and Row.
- Kirkpatrick, Lee A., and Philip R. Shaver 1990 "Attachment Theory and Religion: Childhood Attachments, Religious Beliefs, and Conversion." *Journal for the Scientific Study of Religion*. 29:315-334.
- Krause, Neal 1997 "Religion, Aging, and Health: Current Status and Future Prospects." *Journal of Gerontology* 52B(6):S291-293.
- Lambert, Malcolm 1992 *Medieval Heresy*. Oxford, UK: Basil Blackwell.
- Lang, Graeme, and Lars Ragvald 1993 *The Rise of a Refugee God: Hong Kong's Wong Tai Sin*. Hong Kong: Oxford University Press.
- Lawson, Ronald 1995 "Sect-State Relations: Accounting for the Differing Trajectories of Seventh-Day Adventists and Jehovah's Witnesses."
- 1996 "Church and State at Home and Abroad: The Evolution of Seventh-Day Adventist Relations with Governments." *Journal of the American Academy of Religion* 64:279-311.
- 1998 "From American Church to Immigrant Church: The Changing Face of Seventh-Day Adventism in Metropolitan New York." *Sociology of Religion* 59:329-351.
- Lenowitz, Harris 1998 *The Jewish Messiahs: From Galilee to Crown Heights*. New York: Oxford University Press.
- Levin, Jeffrey S. 1996 "How Religion Influences Morbidity and Health." *Social Science and Medicine* 43(5):849-864.
- , and P. L. Schiller 1987 "Is There a Religious Factor in Health?" *Journal of Religion Health* 26:9-36.
- Marx, Karl, with Frederick Engels (1845) 1998 *The German Ideology*. Amherst, N.Y. Prometheus.
- , and Friedrich Engels 1964 *On Religion*. New York: Schocken.
- Mayer, Hans Eberhard 1972 *The Crusades*. London: Oxford University Press.
- Miller, Alan 1995 "A Rational Choice Model of Religious Behavior in Japan." *Journal for the Scientific Study of Religion* 34:234-244.
- Mundy, John Hine 1985 *The Repression of Catharism and Toulouse*. Toronto: Pontifical Institute of Mediaeval Studies.
- Murray, Alexander 1972 "Piety and Impiety in Thirteenth-Century Italy." *Studies in Church History*. 8:83-106.
- Musick, Marc, and John Wilson 1995 "Religious Switching for Marriage Reasons." *Sociology of Religion* 56:257-270.
- Neitz, Mary Jo 1987 *Charisma and Community: A Study of Religious Commitment within the Charismatic Renewal*. New Brunswick, N.J.: Transaction.
- , and James V. Spickard 1990 "Steps toward a Sociology of Religious Experience: The Theories of Mihai Csikszentmihayi and Alfred Schutz." *Sociological Analysis* 51:15-33.
- Ostow, Mortimer 1990 "The Fundamentalist Phenomenon: A Psychological Perspective." In Norman J. Coher, ed., *The Fundamentalist Phenomenon: A View from Within, a Response from Without*. Grand Rapids, Mich.: Eerdmans.
- Pargament, K. I. and C. L. Park 1995 "Merely a Defense?: The Variety of Religious Ends and Means." *Journal of Social Issues*, 51(2):13-32.
- Perrin, Robin D., Paul Kennedy, and Donald Miller 1997 "Examining the Sources of Conservative Church Growth: Where Are the New Evangelical Movements Getting their Numbers?" *Journal for the Scientific Study of Religion* 36:71-80.
- Poloma, Margaret M. 1989 *The Assemblies of God at the Crossroads*. Knoxville: University of Tennessee Press.
- , and George H. Gallup, Jr. 1991 *Varieties of Prayer: A Survey Report*. Philadelphia: Trinity Press International.
- Preus, J. Samuel 1987 *Explaining Religion: Criticism and Theory from Bodin to Freud*. New Haven, Conn.: Yale University Press.

- Rambo, Lewish R. 1993 *Understanding Religious Conversion*. New Haven, Conn.: Yale University Press.
- Robbins, Thomas 1988 *Cults, Converts and Charisma: The Sociology of Religious Movements*. Beverly Hills, Calif.: Sage.
- Runciman, Steven 1951 *A History of the Crusades* (3 vols.). Cambridge, UK: Cambridge University Press.
- Russell, Jeffrey Burton 1965 *Dissent and Reform in the Early Middle Ages*. Berkeley: University of California Press.
- Sharot, Stephen 1982 *Messianism, Mysticism, and Magic: A Sociological Analysis of Jewish Religious Movements*. Chapel Hill: University of North Carolina Press.
- Sherkat, Darren E. 1997 "Embedding Religious Choices: Preferences and Social Constraints into Rational Choice Theories of Religious Behavior." In Lawrence A. Young, ed., *Rational Choice Theory and Religion: Summary and Assessment*. New York: Routledge.
- Smith, Christian. 2000. *Christian America? What Evangelicals Really Want*. Berkeley: University of California Press.
- , with Michael Emerson, Sally Gallagher, Paul Kennedy, and David Sikkink 1998 *American Evangelism: Embattled and Thriving*. Chicago: University of Chicago Press.
- Stark, Rodney 1985 "From Church-Sect to Religious Economies" In Phillip E. Hammond, ed., *The Sacred in a Post-Secular Age*. Berkeley: University of California Press.
- 1996a *The Rise of Christianity: A Sociologist Reconsiders History*. Princeton, N.J. Princeton University Press.
- 1996b "Why Religious Movements Succeed or Fail: A Revised General Model." *Journal of Contemporary Religion* 11:133–146.
- 1999a "The Micro Foundations of Religion: A Revised Theory." *Sociological Theory* 17:
- 1999b "Atheism, Faith, and the Social Scientific Study of Religion." *Journal of Contemporary Religion*: 14:41–62.
- 1999c "Secularization, R.I.P." *Sociology of Religion* 60:
- 1999d "A Theory of Revelations," *Journal for the Scientific Study of Religion* 37:
- Forthcoming. *Gods: Their Social and Historical Power*.
- , and William Sims Bainbridge 1980 "Towards a Theory of Religion: Religious Commitment," *Journal for the Scientific Study of Religion*, 19:114–128.
- (1987) 1996 *A Theory of Religion*, Republished ed. New Brunswick, N.J.: Rutgers University Press.
- 1997 *Religion, Deviance, and Social Control*. New York: Routledge.
- , and Roger Finke 2000 *The Human Side of Religion: A Social Science Paradigm*. Berkeley: University of California Press.
- , and Laurence R. Iannaccone 1993 "Rational Choice Propositions about Religious Movements." In David G. Bromley and Jeffrey K. Hadden, ed., *Religion and the Social Order, (vol. 3–A): Handbook on Cults and Sects in America*. Greenwich, Conn.: JAI Press.
- , 1994 "A Supply-Side Reinterpretation of the "Secularization" of Europe." *Journal for the Scientific Study of Religion*. 33:230–252.
- , Laurence R. Iannaccone, and Roger Finke 1996 "Religion, Science and Rationality." *American Economic Review* (papers and proceedings):433–437.
- Swatos, William H., Jr. 1987 "The Power of Prayer: Observations and Possibilities." In W. H. Swatos, ed., *Religious Sociology: Interfaces and Boundaries*. New York: Greenwood Press.
- Thurow, Lester 1996 *The Future of Capitalism*. New York: Morrow.
- Tocqueville, Alexis de. (1840) 1956 *Democracy in America* (2 vols). New York: Vintage.
- Wallace, Anthony F. C. 1966 *Religion: An Anthropological View*. New York: Random House.
- Warner, R. Stephen 1993 "Work in Progress towards a New Paradigm for the Sociological Study of Religion in the United States." *American Journal of Sociology* 98:1044–1093.
- Washington, Peter 1995 *Madame Blavatsky's Baboon*. New York: Schoecken.
- Weber, Max (1922) 1993 *The Sociology of Religion*. Boston: Beacon.
- Wilson, Bryan 1975 *Magic and the Millennium*. Frogmore, UK: Paladin.
- 1982 *Religion in Sociological Perspective*. Oxford: UK: Oxford University Press.
- Woodberry, Robert D., and Christian S. Smith 1998 "Fundamentalism et al.: Conservative Protestants in America." *Annual Review of Sociology*. 22:25–56.

RODNEY STARK

## SOCIOMETRY

See Social Networks; Social Psychology.



## SOUTHEAST ASIA STUDIES

Southeast Asia consists of the ten countries that lie between the Indian subcontinent and China. On the mainland of Southeast Asia are Myanmar (Burma), Thailand, Laos, Cambodia, and Vietnam. Insular Southeast Asia includes Indonesia, the Philippines, Brunei, Malaysia, and Singapore. While most of Malaysia (Peninsular Malaysia) is on the mainland, that country usually is considered part of insular Southeast Asia because the Malay population (the majority ethnic population) shares a common language and religion with much of the Indonesian population. The city-state of Singapore (on an island connected by a mile-long causeway to Peninsular Malaysia) was historically part of Malaysia, but because of its unique ethnic composition (three-quarters of the population is of Chinese origin), it is more similar to East Asia than to Southeast Asia.

While there are some common geographic and cultural features, diversity is the hallmark of the region. Incredible indigenous cultural variation has been overlaid by centuries of contact, trade, migration, and cultural exchange from within the region, from other parts of Asia, and for the past five hundred years from Europe (for general overviews of the region, see Osborne 1985; Wertheim 1968). The common characteristic of mainland Southeast Asia is Buddhism, although there are very significant variations across and within countries: Islam is the majority religion in Indonesia, Brunei, and Malaysia, while Christianity is the major religion in the Philippines. The lowlands of both mainland and insular Southeast Asia tend to be densely settled, and wet (irrigated) rice agriculture is the predominant feature of the countryside. Rural areas are knitted together with small- and medium-sized market towns. The major metropolitan areas of the region (Jakarta, Bangkok, Singapore, Manila, Rangoon, Kuala Lumpur, Ho Chi Minh City) are typically port cities or are located along major rivers. Many of these towns and cities have significant Chinese minorities (often intermarried with the local population) that play an important role in commerce. Every country has remote highland and mountainous regions that often are populated by ethnic minorities.

In terms of land area, population size, and cultural and linguistic diversity, Southeast Asia is comparable to Europe (excluding the former Soviet Union). By the year 2000, the population of

Southeast Asia will exceed 500 million, about 8 percent of the world's total. Indonesia is the fifth most populous country in the world, while the oil-rich sultanate of Brunei (on the island of Borneo) is one of the smallest. The other large countries of the region—Thailand, Vietnam, and the Philippines—are more populous than all European countries except for the former Soviet Union and Germany. The sea (South China Sea and Indian and Pacific Oceans) surrounds much of the region, especially the immense Indonesian and Filipino archipelagoes. While the sea can be a barrier, the ocean and the rivers of the region are avenues that have fostered local and long-distance trade throughout history. Moreover, the ease of movement throughout the region seems to have shaped cultures that easily absorbed new ideas and immigrants and have been tolerant of diversity.

### HISTORY

The contemporary political divisions of the region are largely a product of European imperialism, especially of the nineteenth century. Before European intervention, there were great regional civilizations, both agrarian states and maritime empires that waxed and waned over the millennium. The remains of the temple complexes of Angkor (Cambodia) and Pagan (Burma) rival the architectural achievements of any premodern world civilization. Early Western observers of the city of Melaka (a fifteenth-century maritime empire centered on the west coast of the Malayan peninsula) described it as more magnificent than any contemporary European city. These early polities were founded on intensive rice cultivation with complex irrigation systems, the dominance of regional and long-distance trade, or both. The region also has been deeply influenced by contacts with the great civilizations of India and China. The cultural influences from outside have invariably been transformed into distinctive local forms in different Southeast Asian contexts. Because relatively few written records have survived the tropical environment of Southeast Asia, historical research relies heavily on archeological investigations, epigraphs, and records from other world regions, especially Chinese sources.

European influence began in the sixteenth century with the appearance of Portuguese and Spanish naval forces, followed by the arrival of the

Dutch in the seventeenth century and then by that of the British and French. In the early centuries of contact, European powers were able to dominate the seas and thus limit the expansion of Southeast Asian polities, but they rarely penetrated very far inland from their coastal trading cities. All Southeast Asia was transformed, however, in the nineteenth century as the Industrial Revolution in the West stimulated demand for mineral and agricultural products around the globe. New economic organizations of plantations, mines, and markets led to large-scale migration of people and capital to frontier areas and to the cities of Southeast Asia. There was an accompanying flurry of imperialist wars to grab land, people, and potential resources. In a series of expansions, the British conquered the area of present-day Myanmar (Burma) and Malaysia, the Dutch completed their conquest of the East Indies (now Indonesia), and the French took the areas that formed their Indochina empire (present-day Vietnam, Cambodia, and Laos). At the turn of the twentieth century, the United States defeated nationalist forces to take control of the Philippines just as the Spanish Empire was crumbling. Siam (Thailand) was the only indigenous Southeast Asian state to escape the grip of colonialism.

The political history of the region has not been stable. As Western countries moved toward more democratic social and political institutions over the first decades of the twentieth century, the colonists (British, Dutch, American, and French) constructed authoritarian dependencies in the tropics that were based on export economies and racial ideologies. Although there were stirrings of nationalist sentiment in the first half of the twentieth century, it was only after World War II that the nationalist forces were strong enough and the international environment favorable enough to bring political independence to the region. The critical turning point was the Japanese conquest and occupation of Southeast Asia from 1942 to 1945, which permanently shattered the myth of European superiority. The colonial powers returned after World War II, but they encountered popular nationalist movements that demanded the end of colonialism.

Independence was negotiated peacefully by the Americans in the Philippines and the British in Burma and Malaya, but nationalist forces had to wage wars of independence against the Dutch in

Indonesia (1945–1950) and the French in Vietnam (1945–1954). The interplay of nationalist struggles, class conflicts, and East-West cold war rivalry had a marked influence on political developments in the region. In almost every country there were radical and communist movements that held the allegiance of significant sectors of the population. In several cases, communist parties were part of the nationalist movement but left (or were driven out of) the political arena as domestic and international tensions escalated. Vietnam was unique in that the nationalist movement was led by Communists. After the French were defeated in 1954 and agreed to grant independence to Vietnam, the United States intervened to set up a non-Communist Vietnamese state in the southern region of the country. After another twenty years of war and a million casualties, Vietnam was finally united as an independent state in 1975. Since 1975, however, political tension between the socialist states of Vietnam, Cambodia, and Laos and the other countries in the region has been the dominant feature of international relations there.

Domestic political developments within individual countries in the region have been no less dramatic. Governments have oscillated between authoritarian and democratic forms, with no linear trend. Behind the headlines of military coups, regional wars for autonomy, and “managed” elections have been complex political struggles among various contending groups defined by class, region, ethnicity, and kinship. These struggles have ranged from civil war to fairly open elections. Large-scale violence is not the norm, but massacres in Indonesia, Cambodia, and East Timor have been among the worst of such episodes in modern times. Popular civil protests against ruling elites in the Philippines and Burma had significant domestic and international reverberations. Neither academic scholarship nor political reporting has offered generalizations about or convincing interpretations of the postwar political change in Southeast Asia.

Many of the countries in Southeast Asia have experienced remarkable socioeconomic modernization in the postindependence era. This is most evident for the ASEAN (Association of Southeast Asian Nations) countries of Thailand, Malaysia, Singapore, Indonesia, the Philippines, and Brunei. All indicators of socioeconomic development (gross national product, educational levels, occupational

structure, infant mortality) suggest that Southeast Asia has successfully narrowed the gap with the first world, while other regions of the third world have fallen farther behind. The reasons for the success of some countries and the economic stagnation in other countries are a matter of dispute. The East Asian model of state-sponsored export industrialization is widely discussed in policy and academic circles, but the parallels between East Asian and Southeast Asian economic development strategies are still a matter of considerable uncertainty. Few scholarly studies have examined the causes and consequences of the economic modernization of Southeast Asia.

### THE STATUS OF WOMEN

Several theoretical concepts and empirical generalizations have arisen from studies of Southeast Asian societies that have relevance far beyond the region. Empirically, the most common cultural characteristic across the region is the relatively high status of women in Southeast Asian societies, especially compared with East Asia and South Asia. While women still face many social and cultural obstacles in Southeast Asia, the situation appears much different from that in the patriarchal societies of other Asian societies and the traditional female domesticity of many Western societies. While there are a few matrilineal societies in the region, Southeast Asian kinship systems are typically bilateral, with equal importance attached to the husband's and wife's families. The patrilocal custom of an obligatory residence of a newly married couple with or near the groom's family is largely absent in Southeast Asia. The residence of young couples after marriage seems to be largely a matter of choice or is dependent on relative economic opportunities. There is no strong sex preference for children in Southeast Asia, with both girl and boy children seen as desirable.

The relatively positive status of women was evident in earlier times. Reid (1988, pp. 146-172) reports that early European observers were struck by the active role of women in economic and political affairs in Southeast Asia. Traditional folklore also suggests that women play an active role in courtship and that female sexual expectations were as important as men's. Perhaps most unusual was the custom (reported in the fifteenth and sixteenth centuries) of inserting spurs or balls in male

genitals to enhance the sexual pleasure of women (Reid 1988, pp. 148-151).

At present, women seem to be well represented in schools, universities, and employment in all modern sectors of the economy in almost every country in Southeast Asia. There is only a modest scholarly literature on the higher status of women in Southeast Asia (Van Esterik 1982), and few efforts have been made to explain the links between the traditional roles of women as productive workers in the rural rice economy and their relative ease of entry into the modern sector. Demographic research has revealed very rapid declines in fertility in several Southeast Asian countries, particularly Singapore, Thailand, Malaysia, and Indonesia. If the current pace of decline continues, replacement-level fertility (two children per woman) should be reached in the near future (Hirschman and Guest 1990).

### AGRICULTURAL INVOLUTION

Scholarship on Southeast Asia often has reached beyond the boundaries of the region to influence debates over social science concepts, theory, and models. Perhaps most influential have been the books and articles on Indonesia by the anthropologist Geertz. His evocative concepts of the "theatre state," "thick description," and "agricultural involution" have stimulated debate and research in several social science disciplines, including sociology. His model of agricultural involution (Geertz 1968) has been one of the most provocative developments in scholarship on Indonesia over the last generation.

A strikingly bold thesis, agricultural involution is an attempt to explain how Java became one of the most densely settled populations in the world within a traditional agricultural economy. To address this question, Geertz presents an ecological interpretation of the evolution (involution) of Javanese social structure in the face of rapid population growth and Dutch colonialism within the constraints (and possibilities) of a wet rice economy. The colonial system prevented industrialization and the development of an indigenous entrepreneurial class. The traditional rice economy, however, could absorb a larger population because additional labor inputs in the maintenance of irrigation facilities, water control, weeding, and harvesting yielded marginal increments in

rice production. Over the decades, this refinement of traditional production technology (involution) led to an increasing rigidification of traditional Javanese culture, thus discouraging innovation and any efforts at social change and reinforcing the structural limits of the colonial system. Even after independence, when structural limits were lifted, the legacy of the past, as reflected in Javanese culture, remained.

Geertz's thesis remains highly controversial, and many of its components have been confronted with negative evidence (for a review of the debate, see White 1983; and Geertz 1984). For example, Geertz deemphasized social class divisions with his interpretation of "shared poverty" as the traditional social strategy. Most research has shown significant inequality of landholding and other socioeconomic dimensions in Javanese villages, although it is not clear if inequality is permanently perpetuated between families across generations. Even accepting many of the criticisms, agricultural involution is a seminal sociological model that should generate empirical research on the historical development of Asian societies.

### THE MORAL ECONOMY

A classic question in social science involves the causes of revolution or rebellion. Neither Marxian theory, which emphasizes exploitation, nor relative deprivation theory seems to be a satisfactory model to explain the occurrence of revolutions or rebellions. The most sophisticated sociological theory of peasant rebellion is based on historical materials from Burma and Vietnam by the political scientist J Scott (1976) in *The Moral Economy of the Peasant: Subsistence and Rebellion in Southeast Asia*. Scott argues that peasants rebel only when their normative expectations of a minimum subsistence level are not met. These conditions are more likely to occur when capitalist market relations and colonial states erode traditional societies and the reciprocal obligations of peasants and their patrons.

Scott's thesis has been criticized and hotly debated (Popkin 1979; Keyes 1983). One criticism is that Scott believes that peasants prefer traditional societies and are not responsive to economic opportunity. Scott acknowledges that peasants can be quite innovative and individualistic as long as their minimum subsistence is not at risk.

This debate, however, does not really address the central theoretical contribution of Scott's thesis about the specification of the causes of peasant rebellion.

In a more recent study based on fieldwork in a rural Malaysian village, Scott (1985) examines how class antagonisms are displayed in everyday life. Given that rebellion is a very rare event in most societies, Scott calls attention to political, social, and linguistic behaviors that reveal the depth of descensus and potential social conflict but do not risk violent reaction from the state and powerful elites. In these two books and related publications, Scott has provided original interpretations of peasant political behavior in Southeast Asia and set a research agenda for scholars of other world regions and, more generally, the development of social theory.

### CONCLUSION

Scholarship on Southeast Asia, whether in sociology or in other disciplines, has tended to focus on individual countries rather than on the region. Different languages (colonial and indigenous) as well as variations in religious traditions and political and economic systems have reinforced the image of a heterogeneous collection of countries that is labeled a region largely by default. There is tremendous political, economic, and sociocultural diversity in the region; many of these differences, however, are a product of the colonial era and its legacy. The similarity of family systems and the status of women throughout Southeast Asia suggest some common historical and cultural roots for the region. There may well be other social and cultural parallels across Southeast Asia that will be revealed as more comparative research is undertaken (Wolters 1982).

Many indicators of development in Southeast Asia, including very low levels of mortality and almost universal secondary schooling, are approaching the prevailing standards of developed countries. Assuming that current socioeconomic trends continue, several countries in the region probably will follow Japan, Korea, and Taiwan along the path of development in the early decades of the twenty-first century. The study of these processes of modernization and the accompanying changes in politics, family structure, ethnic relations, and

other social spheres should make Southeast Asia an extraordinarily interesting sociological laboratory.

Evolutionary—and sometimes revolutionary—social change continued throughout much of Southeast Asia in the 1990s. After the collapse of the Soviet Union, the socialist countries in the region, including Vietnam, Cambodia, and Laos, moved rapidly toward more market-driven economies. Several political regimes that appeared to be stable for long periods have been transformed. The “people power” popular protests that ended the Marcos regime in the Philippines in the 1980s was echoed by the peaceful transition of power from a military regime in Thailand in the early 1990s and by the ending of the Suharto regime in Indonesia in 1998.

For much of the 1990s, most of Southeast Asia experienced rapid economic growth and the major question was the emerging role of the new middle class (McVey 1992; Girling 1996). This trend was halted in late 1997 by the “Asian economic crisis” that hit the region and affected Thailand, Malaysia, and Indonesia in particular. Both the causes of this crisis and its consequences are currently the subject of much debate. The change of regime in Indonesia and political protests in Malaysia may be the most visible long-term impact may be more profound.

Scholarship inevitably lags behind current events. Several important publications, including the second volume of Reid’s (1990, 1995) *Southeast Asia in the Age of Commerce 1450–1680* and a much expanded version of Wolters’s classic *History, Culture, and Region in Southeast Asian Perspectives 1999*, offer a new understanding of the history of the premodern era. Although the definition of Southeast Asia as a region sometimes has been considered arbitrary, historical studies show common cultural, political, and social forms in many places throughout the region.

One of the defining features of the region has been the relatively easy absorption of peoples, ideas, and cultural practices from elsewhere. In the twentieth century, assimilation into Southeast Asian societies became more difficult with the creation of political and social barriers. These issues are illuminated with considerable insight in Chirot and Reid’s (1997) edited collection that compares the experience of the Chinese in Southeast Asia with that of the Jews in central Europe.

Research on Southeast Asia over the last decade also has been influenced by Anderson’s (1991) *Imagined Communities*, a book originally published in the 1980s. Although Anderson is a specialist on Southeast Asia, his book on the development of nationalism provides comparisons from across the world.

## REFERENCES

- Anderson, Benedict 1991 *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, rev. ed. London: Verso.
- Chirot, Daniel, and Anthony Reid, eds. 1997 *Essential Outsiders: Chinese and Jews in the Modern Transformation of Southeast Asia and Central Europe*. Seattle: University of Washington Press.
- Geertz, Clifford 1968 *Agricultural Involvement: The Processes of Ecological Change in Indonesia*. Berkeley: University of California Press.
- 1984 “Culture and Social Change.” *Man* 19:511–532.
- Girling, John 1996 *Interpreting Development: Capitalism, Democracy, and the Middle Class in Thailand*. Ithaca, N.Y.: Southeast Asia Program, Cornell University.
- Hirschman, Charles, and Philip Guest 1990 “The Emerging Demographic Transitions of Southeast Asia.” *Population and Development Review* 16:121–152.
- Keyes, Charles F., ed. 1983 “Peasant Strategies in Asian Societies: Moral or Rational Economic Approaches—A Symposium.” *Journal of Asian Studies* 42:753–868.
- McVey, Ruth, ed. 1992 *Southeast Asian Capitalists*. Ithaca, N.Y.: Southeast Asia Program, Cornell University.
- Osborne, Milton 1985 *Southeast Asia: An Illustrated Introductory History*. Sydney: Allen and Unwin.
- Popkin, Samuel L. 1979 *The Rational Peasant*. Berkeley: University of California Press.
- Reid, Anthony 1988 *Southeast Asia in the Age of Commerce, 1450–1680*, vol. 1: *The Lands below the Winds*. New Haven, Conn.: Yale University Press.
- 1990 *Southeast Asia in the Age of Commerce 1450–1680: The Lands below the Winds*. New Haven, Conn.: Yale University Press.
- 1995 *Southeast Asia in the Age of Commerce 1450–1680: Expansion and Crisis*. New Haven, Conn.: Yale University Press.
- Scott, James C. 1976 *The Moral Economy of the Peasant: Subsistence and Rebellion in Southeast Asia*. New Haven, Conn.: Yale University Press.
- 1985 *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven, Conn.: Yale University Press.

- Van Esterik, Penny 1982 *Women of Southeast Asia*. Dekalb: Center for Southeast Asian Studies, Northern Illinois University.
- Wertheim, W. F. 1968 "Southeast Asia." In David Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- White, Benjamin 1983 "Agricultural Involvement and Its Critics: Twenty Years After." *Bulletin of Concerned Asian Scholars* 15:18-41.
- Wolters, O. W. 1982 *History, Culture, and Region in Southeast Asian Perspectives*. Singapore: Institute of Southeast Asian Studies.
- 1999 *History, Culture, and Region in Southeast Asian Perspectives*, rev. ed. Ithaca, N.Y.: Southeast Asia Program, Cornell University.

CHARLES HIRSCHMAN

## SOVIET AND POST-SOVIET SOCIOLOGY

In prerevolutionary Russia, sociology occupied a marginal position. The state universities offered no instruction in the field, but there was a solid intellectual tradition of historical and theoretical sociology (Maxim Kovalevsky, Nikolai Mikhailovsky, Evgeny de Roberty), the sociology of law (Leon Petraizky, Pitirim Sorokin), and the sociology of social problems (living conditions of industrial workers and peasants, public health, crime and prostitution in the cities). Beginning in the 1860s, the provincial intelligentsia initiated a kind of social movement, *Zemskaja statistika* (Statistics for Local Administration). Since official governmental statistics were unreliable, local statisticians made systematic surveys of households, daily life and public health conditions, and the reading preferences of the population (N. A. Rubakin). A modern system of sampling was elaborated by the statistician A. A. Chuprov for those surveys; K. M. Takhtarev introduced the concept of statistical sociological methods in social research.

In 1916, the Russian Sociological Society was founded, along with the "Sociological Institute," where M. M. Kovalevsky, K. M. Takhtarev, N. I. Kareev, and P. A. Sorokin gave lectures. Western sociological classics by Auguste Comte, Herbert Spencer, Emile Durkheim, Gabriel Tarde, Gustave Le Bon, Georg Simmel, Lester Ward, and others

were available in Russian translations. Most important European sociological papers were immediately translated in the series *New Ideas in Sociology*. There was also a well-developed ethnography and a literary genre of sociological journalism.

The Bolshevik Revolution provided strong stimulus to sociological reflection and empirical social research. In the Soviet government decree "About the Socialist Academy of the Social Sciences," drafted in May 1918, Lenin (1962, p. 372) stressed the need "to organize a series of social researches" and called it "one of the most urgent tasks of the day." However, the Bolsheviks tolerated research only from Marxist and procommunist positions. In the early postrevolutionary years, censorship was relatively weak or inefficient. For example, Sorokin not only established the first sociological laboratory in Petrograd University but also succeeded in publishing (illegally) his two-volume *System of Sociology* (Sorokin 1920), for which he was awarded a doctorate in April 1922. He also conducted important empirical investigations on mass starvation in the districts of Samara and Saratov and examined its influence on various aspects of social life and human behavior.

However, this liberalism or negligence on the part of the authorities was short-lived. In autumn 1922, a group of leading Russian intellectuals, including Sorokin and other prominent social philosophers, was expelled from the country, ending non-Marxist sociology in Soviet Russia.

The tightening ideological control proved detrimental to socialist and Marxist social research as well. Nevertheless, the 1920s was a fruitful period both in empirical research and in theoretical-methodological work. The most important theoretical contributions were in the field of economic sociology (A. V. Chajanov, N. D. Kondratjev). There were also interesting studies on the social organization of labor, the budgeting of time in work and leisure activities (S. G. Strumilin), population dynamics, rural and urban ways of life (A. I. Todorsky, V. E. Kabo), marriage and sexual behavior, social psychology (V. M. Bekhterev), social medicine, and other topics. All this research was finished by the early 1930s.

The Stalinist totalitarian system was incompatible with any kind of social criticism, problem-oriented thinking, or empirical research. Most creative original thinkers were liquidated, and their

books were prohibited. Sociology was declared “bourgeois pseudo-science.” Official social statistics were kept secret or falsified. Empirical research that relied on questionnaires, participant observation, and similar methods was forbidden. All social theory was reduced to the official dogmatic version of historical materialism, which had very little in common with genuine Marxist dialectics. Practically no firsthand information about Western sociology was available.

The revival of sociology in the Soviet Union began during the Khrushchev’s era in the late 1950s. It was initiated by a group of young philosophers and economists with a liberal political orientation. This intellectual initiative received support from reformist and technocratically oriented people in the party and state leadership. The first organizational step in this direction was the establishment in 1958 of the Soviet Sociological Association (SSA). The primary aim of this move was to facilitate participation in international sociological congresses by Soviet ideological bureaucrats in administrative academic positions. Gradually, thanks to personal efforts of Gennady Ossipov, among others, the SSA became a sort of organizational center for the emerging discipline.

To avoid conflicts with the dominant ideology, it was unanimously agreed that the only acceptable “scientific” general sociological theory was Marxist historical materialism but that it should be supplemented by “concrete social research” and eventually some middle-range theories. In 1960 Ossipov organized in the Institute of Philosophy of the USSR Academy of Sciences in Moscow a small unit for research on the new forms of work and daily life. This unit later was transformed into the Department of Concrete Social Research. At about the same time, Vladimir Iadov organized, within the philosophical faculty of Leningrad State University, the Laboratory of Concrete Social Research, which was dedicated to the study of job orientation and workers’ personalities. At the Novosibirsk Institute of Industrial Economics and Organization, Vladimir Shubkin developed a unit for studies of youth issues, including high school children’s professional orientations and social mobility, and Tatiana Zaslavskaiia initiated the fields of economic and rural sociology. Sociology research units appeared under various names at the universities of Sverdlovsk and Tartu (Estonia). In 1968, the independent Institute of Concrete So-

cial Research of the USSR Academy of Sciences was established in Moscow, headed by the eminent economist and vice-president of the USSR Academy of Sciences A. M. Rumiantsev.

According to Shlapentokh (1987), 1965–1972 were the golden years of Soviet sociology. Important original research was done on workers’ attitudes toward their jobs and on the interrelationship of work and personality (Iadov et al. 1970), professional orientations of youth, rural sociology and population migrations (Zaslavskaiia 1970; Zaslavskaiia and Ryvkina 1980; Arutiunian 1971), public opinion and mass media (Grushin 1967; Shlapentokh 1970), industrial sociology (Shkaratan 1978), marriage and the family (Kharchev 1964), personality (Kon 1967), leisure (Gordon and Klopov 1972), political institutions (F. M. Burlatsky, A. A. Galkin), and other topics. At the same time, research on the history of sociology had begun, and a dialogue with Western theoretical ideas instead of a blunt ideological denunciation of everything “non-Marxist” was initiated (Andreeva 1965; Kon Zamoshkin 1966). In theoretical terms, structural functionalism, symbolic interactionism, and C. Wright Mills’s “new sociology” were of particular interest to Soviet sociologists. The American Sociological Association aided these developments by arranging to send professional books and journals to the Soviet Union. In the 1960s, a few Western sociological books and textbooks, beginning with *Modern Sociological Theory in Continuity and Change* edited by H. Becker and A. Boskoff, were translated and published in Russian.

The social and intellectual situation of Soviet sociology was very uncertain. It was completely dependent on the official ideology and the goodwill of party authorities. Even a hint of social criticism was deemed dangerous, and such work could be published only if it was formulated in the ESOPs language. The Institute of Concrete Social Research was under constant attack. Especially devastating and venomous was an attack on Levada’s *Lectures on Sociology* (1969); soon after the attack, Levada was dismissed from Moscow University and deprived of a professorial title. In 1972, the liberal head of the Institute, A. M. Rumiantsev, was replaced by the reactionary Mikhail Rutkevich, who had initiated an ideological campaign against “Western influences.” As a result of his policies, the most prominent and qualified scholars were forced to leave the institute.

Until 1986, Soviet sociology was in bad shape, but the process of its institutionalization continued. It was a period of extensive growth of sociological units. Many new laboratories and departments of applied social research in the universities and sociological and social psychological laboratories in the big industrial plants had been established. Industrial sociologists (the most numerous and active group in the SSA) studied motivation to work, trends in the workforce, the efficiency of different forms of labor organization, in-group relations between workers and employers, and systems of management. The managers, who pretended to be "progressive," elaborated and reported to the party authorities "the plans of social developments" based on sociological studies (later, some of these industrial sociologists were able to consult the new post-soviet businessmen).

In 1972, the Institute of Concrete Social Research was renamed the Institute for Sociological Research. In 1974, the first professional journal, *Sotsiologicheskie Issledovania (Sociological Research)*, was inaugurated (the first editor in chief was Anatoly Kharchev). SSA membership grew continuously. In the late 1980s, the SSA had about 8,500 individual and 300 collective members and twenty-one regional branches. The technical and statistical level of sociological research in the 1970s and 1980s improved considerably. Some new sociological subdisciplines emerged. At its apogee, before the collapse of the Soviet Union, the SSA had thirty-eight specialized sections, including twelve research committees, directly connected with the respective International Sociological Association (ISA) committees. The geography of sociological research centers has also expanded.

The general intellectual and theoretical level of Soviet sociology was, with few exceptions, inadequate. Relatively free theoretical reflection was limited to the marginal fields of social psychology, anthropology, and history. Most sociological research was done on the micro level and involved separate industrial plants, without any attempt at broad theoretical generalization. Publications of a more general character were mostly apologies for the so-called real socialism. Sociological theories were divided between historical materialism and dogmatic ideological scholasticism, "the theory of scientific communism." Attempts to narrow the gap between sociological statements and social realities were ruthlessly punished by the authori-

ties. The Leningrad sociological school, perhaps the best in the country, was decimated by the local party leadership in the mid-1980s. Zaslavskaja was in serious trouble when her report, which was highly critical of the prospects for economic reforms without parallel political changes, was published in the West. The public image of sociology had changed dramatically: In the 1960s, the new discipline was associated in the public's mind with social criticism and progressive economic reforms, and in the late 1970s, industrial sociologists sometimes were represented in the mass media as sly manipulators helping plant managers play down workers' discontent.

Perestroika and glasnost drastically changed the place of sociology in Soviet society. Mikhail Gorbachev and his team claimed that they needed an objective social science for information and advice, and the majority of Soviet sociologists were, from the beginning, strong supporters of reforms. In 1986, Zaslavskaja was elected president of the SSA. In 1987, a special resolution of the Communist Party Central Committee acknowledged that sociology was an important scientific discipline. In 1988, the Institute of Sociological Research was transformed into the Institute of Sociology, and V. Iadov was appointed its director. Sociologists (for example, Galina Starovoitova) took an active part in political life not only as advisers to the government but as deputies of central and local soviets and, after 1991, the post-Soviet parliaments of independent states. There were no longer official restrictions on the topics suitable for sociological research, and the publication of results became much easier. Some newspapers introduced regular sociological columns.

However, the relationship between sociology and political power is always problematic. On the one hand, neither Gorbachev nor Boris Yeltsin really needed or followed sociological advice. Very often, they did the opposite of what they have been advised to do. For example, Gorbachev's catastrophic antialcohol campaign, which was the first irreparable blow to the state budget and created the first wave of organized crime, was initiated despite strong and unanimous objections from social scientists. While making his fatal decisions about the Chechen war, Yeltsin completely ignored professional opinions. These experiences made sociologists more critical of the regime.



On the other hand, sociologists have been neither intellectually nor morally ready for new social responsibilities. The lack of a sociological imagination and their predominantly functionalist or empiricist mentality made them more comfortable with post hoc explanations of events than with responsible and reliable predictions. Social scientists are always more sure about what should not be done than about what to do, and Soviet sociology had never had a unified professional body.

By 1991 but especially after 1993, there was a deep political and intellectual schism in the former Soviet sociology. The majority of its founders remained faithful to liberal, democratic, and pro-Western ideas. However, liberal politicians, they often did not know how to apply those general principles to particular Russian, Ukrainian, or other situations. On the contrary, the former "scientific communists," who declared themselves sociologists or politologists after 1991 and who hold now many if not most university chairs, proclaim their fidelity to Marxism-Leninism, often with a strong flavor of Russian nationalism, traditionalism, and religious orthodoxy. The gap between these two wings is irreconcilable, and that gap has many organizational, ideological, and educational implications.

In the 1990s, there were essential changes in the institutional structure of sociological communities in all the post-Soviet states as well as in areas of research. To replace the SSA, several national, republican sociological associations have been formed. Sometimes there are more than one sociological association in the same country. Alongside the national Sociological Association of Russia (Russian Sociological Society), which is a collective member of the ISA, Ossipov organized an alternative Association of Sociologists and Demographers; he also initiated the split in the Institute of Sociology (IS) of RAS and created in the framework of RAS a new Institute of Social and Political Problems (ISPP), that became one of the main intellectual centers of communist and nationalist opposition to reforms. The coexistence of the two centers is by no means peaceful.

The main research projects of the IS include the theory and history of the discipline, quantitative and qualitative methodology, social stratification, sociocultural processes in Russia in the context of global social and economic changes, changes

in personality, social identities and new forms of solidarities, economic and political elites, environmental studies, family and gender, social organizations, and social conflicts. The IS has an affiliation in St. Petersburg (director Serguei Golod). The IS is also combining research with teaching undergraduates and postgraduate students. The European University in St. Petersburg (rector Boris Firsov), has departments of history, political sciences, and sociology.

Fundamental sociological research is also being done in other academic institutions and universities, such as those in Novosibirsk (rural and regional sociology), Samara (sociology of labor), and Niznii Novgorod (stratification and regional studies). Research on interethnic relationships and conflicts is concentrated in the Institute of Ethnology and Anthropology of the RAS; population and gender studies are conducted in the Institute for Social-Economic Studies of Population, and so on. Many sociological groups and centers are moving from one academic institute to another or becoming fully independent, especially if they can make money by doing applied research.

Public opinion and market surveys centers became independent enterprises, some of which were united in the Russian Guild of Pollsters and Marketing Researchers. The All-Russian Public Opinion Research Center (directed by Yuri Levada) is a leading national center for public opinion polls; among many others, the Independent Public Opinion Research Service *Vox Populi* (VP), founded by Boris Grushin, and *Obshechestvennoe mnenie* (the Foundation of Public Opinion polls) are the most visible. Many sociologists are working as political image makers, speechwriters, economic consultants, and so on.

Sociology is now an institutionalized discipline in Baltic states, Armenia, Georgia, Kazakhstan, Uzbekistan, Belarus, and Ukraine. Especially visible progress in research and teaching sociology has occurred in Estonia and Ukraine. In the Soviet Union, Estonia was one of the few places where Western traditions of sociology were known and maintained. Since 1991, the main focus of sociological research in Estonia has been the empirical description and theoretical interpretation of the rapid social changes taking place in all spheres of society. The main traditional branches of Estonian sociology were social structure and stratification

(M. Titma, E. Saar); family and living conditions (Narusk 1995); the environment (M. Heidmets, Y. Kruusvall); urban sociology (M. Pavelson, K. Paadam), the mass media; youth; and education (P. Kenkmann). New situations have stimulated theoretical analyses of transitional processes (Lauristin and Vihalemm 1997) and explorations of new areas of research, such as the integration of the Russophone minority in Estonian society, poverty and social deprivation, political sociology, and public opinion research. In the second half of the 1990s, the dominant theoretical paradigm of social research in Estonia shifted from traditional structural functionalism to social constructivism. The main centers of sociological research in Estonia are the University of Tartu, the Pedagogical University of Tallinn, and the Institute for International and Social Studies in Tallinn.

In 1991, the Institute of Sociology of the Academy of Sciences of Ukraine and the first independent research center, the Kiev International Institute of Sociology, were founded. Together with the universities of Kiev, Kharkiv, Lviv, and Odessa, these Institutes have become centers of the development of sociological science in Ukraine. The basic topics of studies are social transformations and change (E. Golorakha, V. Khmelko, O. Kutsenko, E. Yakuba), economic and political sociology (I. Bekestina, N. Panina), ethnosociology, (N. Chernysh, M. Shulga, and B. Yertukh) sociology of mass consciousness (N. Kostenko, V. Ossorskiy, I. Popora), social psychology, relationships between social structures and personality under conditions of radical social change, the sociology of the Chernobyl catastrophe, and gender studies. In 1992, the Sociological Association of Ukraine was reorganized as an independent national association. Since 1993, the preparation of sociologists, using the programs and textbooks of Western universities, began at the oldest university in eastern Europe, Kiev-Mohyla Academy (founded in 1632). The academic journal *Sociology: Theory, Methods, Marketing* began to be issued in Ukrainian (1998) and Russian (1999).

The main problem confronting Russian sociology is the shortage of money and professional personnel. Until 1989 in the Soviet Union, there was practically no undergraduate sociological education; only a few courses in applied (mainly industrial) sociology were offered. Now sociological departments and schools have established in Mos-

cow, Saint Petersburg, Novosibirsk, Ekaterinburg, and some other state universities, and there are about two hundred departments of sociology and political science in other colleges. The Russian Ministry of Higher Education issued the "State Standard" in sociology, which prescribed teaching the discipline as a multitheoretical one, not merely Marxist-oriented. Up-to-date methods of teaching sociology are provided by new educational centers: the Moscow School of Social and Economic Sciences, the European University in Saint Petersburg, the Faculty of Sociology of the Academic Institute of Sociology, and the High School of Economics in Moscow. According to the official statistics, in 1998 more than 6,600 university students studied sociology as their main subject. The discipline is taught also in many high schools and lyceums. With the financial support of different foundations (George Soros is the leading donor), sociological classics, world-recognized modern authors (P. Bourdieu, Z. Bauman, A. Giddens, Y. Habermas, and many others) and teaching materials (handbooks and readers) have been published. New professional journals, including *The Russian Public Opinion Monitor* (edited by T. Zaslavskaja and Y. Levada), *Sociological Journal* (edited by G. Batygin), *Sociology-4M: Methodology, Methods, Mathematical Models* (edited by V. Iadov); *The World of Russia* (edited by O. Shkaratan), have been published. In Russia, the Baltic states, and Ukraine, there are summer schools and advanced courses in theory and subdisciplines of sociology for young teachers and postgraduates where internationally renowned scholars lecture. The exchange of graduate students in sociology between post-Soviet, U.S., and west European universities is growing rapidly. Prominent Western sociologists are invited regularly to give lectures and seminars at Russian and other independent state universities and vice versa.

Post-Soviet sociology is now ideologically and organizationally open and interested in international contacts and exchanges on all levels. There are many joint research projects with American, Canadian, German, French, Finnish, Japanese, and other scholars. Most of these projects are related to current political attitudes and value orientations, ethnic relations and regional studies, stratification, personality studies, social minorities, organizational culture, and modernization. The annual international symposia "Where Is Russia going?" are organized by the Independent Mos-

cow School of Economics and Political Sciences (T. Shanin and T. Zaslavskaja).

High-level studies are being conducted on the problems of the economic and political elites (Kryshtanovskaja 1997), environmental sociology (Yanitsky 1993), gender and life stories (Semenova and Foteeva 1996), political sociology (Zdravomyslova 1998), and the sociology of culture (Ionin 1996). Some of these projects are the result of academic international cooperation, while others are financed by charity funds the State foundation for humanities (John D. and Catherine T. MacArthur Foundation, Open Society Institute, Ford Foundation, and others), and voluntary associations.

The prospects for the development of post-Soviet sociology depend on the fate of economic and democratic transformations. The gigantic social experiment unfolding in the post-Soviet region needs creative support from the social sciences. It is a powerful stimulus for sociological imagination and theory construction. Today sociologists in these countries are overburdened by the need to search for immediate practical solutions to urgent political and economic issues and have no time for quiet theoretical reflection. The most important sociological contributions to reforms are still the public opinion polls and information about current social processes. The next step seems to be the emergence of a sociology of social problems interpreted not only in the specific national contexts but in the context of the global problems of civilization as well. This, may lead to the revival of historical and comparative macrosociology and produce new theoretical insights. All this will be feasible, however, only as the result of intensive international and interdisciplinary intellectual cooperation.

## REFERENCES

- Andreeva, G. M. 1965 *Sovremennaia Bourzhuznaia Empiricheskaia Sotsiologia*. Moscow: Mysl'.
- Arutiunian, I. V. 1971 *Struktura Sel'skogo Naselenia SSSR*. Moscow: Mysl'.
- Batygin, G. S., and I. F. Deviatko 1994 "Russian Sociology: Its Origins and Current Trends." In R. P. Mohan and A. S. Wilke, eds., *International Handbook of Contemporary Developments in Sociology*. Westport, Conn.: Greenwood.
- Gordon, L. A., and E. V. Klopov 1972 *Chelovek Posle Raboty*. Moscow: Nauka.
- Grushin, B. A. 1967 *Mnenia o Mire i Mir Mnenij*. Moscow: Politizdat.
- Iadov, V. A., V. Rozhin, and A. Zdravomyslov, 1970 *Man and His Work*. White Plains, N.Y.: International Arts and Sciences Press.
- , eds. 1998 *Sotsiologia v Rossii*. Moscow: Institut Sotsiologii Press.
- Ionin, L. 1996 *Russische Metamorphosen: Aufsätze zu Politik: Alltag und Kultur*. Berlin: Berliner Debatte.
- Kohn, M., K. Slomczynski, K. Janicka, V. Khmelko, B. Mach, V. Paniotto, W. Zaborowski, R. Guttierrez, and C. Heyman 1997 "Social Structure and Personality under Conditions of Radical Social Change: A Comparative Analysis of Poland and Ukraine." *American Sociological Review* 62:614–638.
- Kon, I. S. 1967 *Sotsiologia Lichnosti*. Moscow: Politizdat.
- Kryshtanovskaia, O. V. 1997 "The Emerging Russian Elite: Old & New; The Aftermath of 'Real Existing Socialism'" in Jacques Heshz Johannes, ed., *Eastern Europe*. vol. 1. London: Macmillan.
- Lauristin, M, and P. Vihalemm, eds. 1997 *Return to the Western World: Cultural and Political Perspectives on the Estonian Post-Communist Transition*. Tartu, Estonia: Tartu University Press.
- Lenin, V. I. 1962 *O Sotsialisticheskoi Akademii Obshchestvennykh Nauk*. Vol. 36 of *Polnoe Sobranie Sochinenij*. Moscow: Politizdat.
- Levada, I. A. 1969 *Lektsii po Sotsiologii*, Vols. 1 and 2. Moscow: IKSI AN SSSSR.
- Lewada, Y. 1992 *Die Sovietmensen 1989–1991: Soziogram eines Zehrfall*. Berlin: Argon Verlag
- Narusk, A., ed. 1995 *Every-Day Life and Radical Social Change in Estonia*. Tallin, Estonia: Institute of International and Social Studies.
- Moskvichev, P. N. 1997 *Sotsiologia i Vlast: Dokumenty 1953–1968*. Moscow: Academia Press.
- Semenova, V., and E. Foteeva E. 1996 *Sudby Ludei. Rossia XX vek. Biografii Semei Kak Objekt Sotsiologicheskogo Issledovania*. Moscow: Institut Sotsiologii RAN.
- Shlapentokh, V. 1987 *The Politics of Sociology in the Soviet Union*. Boulder and London: Westview.
- Sorokin, P. A. 1920 *Sistema Sotsiologii*. Petrograd
- Yanitsky, O. 1993 *Russian Environmentalism*. Moscow: Mezhdunarodnye Otnoshenia.
- Zamoshkin, Iu. A. 1966 *Krizis Burzhuznogo Individualizma i Lichnost'*. Moskva: Mysl'.
- Zaslavskaja, T. (ed.) 1970 *Migratsia Sel'skogo Nasleniia*. Voskva: Mysl'.

- , and R. Ryvkina (eds.) 1980 *Metodologia Metodika Sistemnogo Izucheniia Derevni*. Novosibirsk Nauka.
- 1990 *The Second Socialist Revolution and Alternative Soviet Strategy*. London: I.B. Tauris.
- Zdravomyslov, A. G., ed. 1986 *Developments in Marxist Sociological Theory*. New York and London: Sage.
- 1998. "Becoming of Political Sociology in Russia: The First Steps." In P. Sztompka, ed., *Building Open Society in East-Central Europe*. London: JSE Allen.

IGOR S. KON  
VLADIMIR A. IADOV

## SPORT

People in all cultures have always engaged in playful physical activities and used human movement as part of their everyday routines and collective rituals (Huizinga 1955). The first examples of organized games in societies worldwide probably emerged in the form of various combinations of physical activities and religious rituals (Guttman 1978). Those games were connected closely with the social structures, social relations, and belief systems in their societies. Although they often recreated and reaffirmed existing systems of power relations and dominant ideologies, they sometimes served as sites for resistant or oppositional behaviors (Guttman 1994; Sage 1998). Variations in the forms and dynamics of physical activities and games indicate that they are cultural practices that serve different social purposes and take on different meanings from time to time and place to place. Research on these variations has provided valuable insights into social processes, structures, and ideologies (Gruneau 1999; Sage 1998).

The physical activities that most sociologists identify as "modern sports" emerged in connection with a combination of rationalization, industrialization, democratization, and urbanization processes in the eighteenth and nineteenth centuries. As various forms of physical activities and play were constructed as institutionalized, competitive, rule-governed challenges and games, they became associated with a range of processes and structures in societies. To varying degrees in different settings, "organized sports" were implicated in processes of social development and the structure of family life, socialization and education, identity

formation and government policy, commodification and the economy, and globalization and the media. Today, sports constitute a significant part of the social, cultural, political, and economic fabric of most societies.

As cultural practices, organized sports constitute an increasingly important part of people's lives and collective life in groups, organizations, communities, and societies. In addition to capturing individual and collective attention, they are implicated in power relations and ideological formation associated with social class, gender, race and ethnicity, sexuality, and physical ability. Because sports are social constructions, they may develop around particular ideas about the body and human nature, how people should relate to one another, expression and competence, human abilities and potential, manhood and womanhood, and what is important and unimportant in life. These ideas usually support and reproduce the dominant ideology in a society, but this is not always the case. Ideology is complex; therefore, the relationship between sports and ideological formation and transformation is sometimes inconsistent or even contradictory. Furthermore, sports come in many forms, and those forms can have many different associated social meanings.

Although sports continue to exist for the enjoyment of the participants, commercialized forms are planned, promoted, and presented for the entertainment of vast numbers of spectators. Sport events such as the Olympic Games, soccer's World Cup (men's and women's), the Tour de France, the tennis championships at Wimbledon, American football's Super Bowl, and championship boxing bouts capture the interest of billions of people when they are televised by satellite in over 200 countries around the world. These and other formally organized sports events are national and global industries. They are implicated in processes of state formation and capitalist expansion and are organized and presented as consumer activities for both participants and spectators. Although sport programs, events, and organizations may be subsidized directly or indirectly by local or national governments, support increasingly comes from corporations eager to associate their products and images with cultural activities and events that are a primary source of pleasure for people all over the world. Corporate executives have come to realize, as did Gramsci (1971) when he discussed

hegemony and consensus-generating processes, that sponsoring people's pleasures can be crucial in creating a consensus to support corporate expansion. At the same time, most sport organizations have sought corporate support.

People of all ages connect with sports through the media. Newspapers in many cities devote entire sections of their daily editions to sports, especially in North America, where the space devoted to sports frequently surpasses that given to the economy, politics, or any other single topic of interest (Lever and Wheeler 1993). Major magazines and dozens of specialty magazines cater to a wide range of interests among participants and fans. Radio coverage of sporting events and sports talk shows capture the attention of millions of listeners every day in some countries. Television coverage of sports, together with commentary about sports, is the most prevalent category of video programming in many countries. First the transistor radio and more recently satellites and Internet technology have enabled millions of people around the world to share their interest in sports. As Internet technology expands, these media-facilitated connections that revolve around sports will take new forms with unpredictable social implications.

Worldwide, many people recognize high-profile teams and athletes, and this recognition fuels everything from product consumption to tourism. Sports images are a pervasive part of life in many cultures, and the attention given to certain athletes today has turned them into celebrities, if not cultural heroes. In cultures in which there have been assumed connections between participation in sports and character formation, there has been a tendency to expect highly visible and popular athletes to become role models of dominant values and lifestyles, especially for impressionable young people. This has created a paradoxical situation in which athletes often are held to a higher degree of moral accountability than are other celebrities while at the same time being permitted or led to assume permission to act in ways that go beyond traditional normative boundaries.

People around the world increasingly talk about sports. Relationships often revolve around sports, especially among men but also among a growing number of women. Some people identify with teams and athletes so closely that what happens in

sports influences their moods and overall sense of well-being. In fact, people's identities as athletes and fans may be more important to them than their identities related to education, religion, work, and family.

Overall, sports and sports images have become a pervasive part of people's everyday lives, especially among those who live in countries where resources are relatively plentiful and the media are widespread. For this reason, sports are logical topics for the attention of sociologists and others concerned with social life.

### USING SOCIOLOGY TO STUDY SPORTS

Although play and games received attention from various European and North American behavioral and social scientists between the 1880s and the middle of the 20th century, sports received scarce attention in that period (Loy and Kenyon 1969). Of course, there were notable exceptions. Thorstein Veblen wrote about college sports in the United States in 1899 in *Theory of the Leisure Class*. Max Weber mentioned English Puritan opposition to sports in the 1904 and 1905 volumes of *The Protestant Ethic and the Spirit of Capitalism*, and William Graham Sumner discussed "popular sports" in his 1906 *Folkways*. Willard Waller devoted attention to the "integrative functions" of sports in U.S. high schools in *The Sociology of Teaching* in 1932.

The first analyst to refer to a "sociology of sport" was Theodor Adorno's student Heinz Risse, who published *Sociologie des Sports* in 1921. Sports received little or no further analytic attention from social scientists until after World War II. Then, in the mid-1950s, there was a slow but steady accumulation of analyses of sports done by scholars in Europe and North America (Loy and Kenyon 1969; Dunning 1971).

The origins of the sociology of sport can be traced to both sociology and physical education (Ingham and Donnelly 1997; Sage 1997). The field initially was institutionalized in academic terms through the formation of the International Committee for Sport Sociology (ICSS) and the publication of the *International Review for Sport Sociology* (IRSS) in the mid-1960s. The ICSS was a subcommittee of the International Council of Sport Science and Physical Education and the International Sociological Association, and it sponsored the publi-

cation of the IRSS. Other publications in the 1960s and 1970s provided examples of the research and conceptual issues discussed by scholars who claimed an affiliation with the sociology of sport (Kenyon 1969; Krotee 1979; Lüschen 1970). In addition to meeting at the annual conferences of the ICSS beginning in the mid-1960s, many scholars in the sociology of sport also met at the annual conferences of the North American Society for the Sociology of Sport (NASSS). This organization was founded in 1978. It has sponsored conferences every year since then, and its membership has been as high as 326 in 1998. In 1984, the *Sociology of Sport Journal* was published under the sponsorship of the NASSS.

Although the sociology of sport involves scholars from many countries and has its foundations in traditional academic disciplines, its early growth was fueled partly by the radical and reform-oriented work of social activists trained in a variety of social sciences. That work attracted the attention of a number of young scholars in both sociology and physical education. For example, in U.S. universities, many courses devoted to the analysis of sport in society in the 1970s highlighted sport as a social institution, but many also used sports as a focal point for critical analyses of U.S. society as a whole. Objections to the war in Vietnam inspired analyses of autocratic and militaristic forms of social organization in sports and other spheres of social life. Critiques of capitalism were tied to research on the role of competition in social life and the rise of highly competitive youth and interscholastic sports. Concern with high rates of aggression and violence in society was tied to an analysis of contact sports that emphasize the physical domination of opponents. Analyses of racial and civil rights issues were tied to discussions of racism in sports and to issues that precipitated the boycott of the 1968 Mexico City Olympic Games by some black American athletes (Edwards 1969). Analyses of gender relations were inspired by the widespread failure of U.S. high schools and universities to comply with Title IX legislation that, among other things, mandated gender equity in all sport programs sponsored by schools that received federal funds.

Today, those who are dedicated to studying sports as social and cultural phenomena constitute a small but active, diverse, and steadily expanding collection of scholars from sociology, physical edu-

cation and kinesiology, sport studies, and cultural studies departments. This has made the field unique because many of these scholars have realized that to maintain the field they must engage each other despite differences in the research questions they ask and the theoretical perspectives and methodologies they use.

Mainstream sociology has been slow at the institutional level to acknowledge the growing social and cultural significance of sports and sports participation. The tendency among sociologists to give priority to studies of work over studies of play, sports, or leisure accounts for much of this disciplinary inertia. Furthermore, sports have been seen by many sociologists as nonserious, nonproductive dimensions of society and culture that do not merit scholarly attention. Consequently, the sociology of sport has continued to exist on the fringes of sociology, and studying sports generally does not forward to a scholar's career in sociology departments. For example, in 1998–1999, only 149 (1.3 percent) of the 11,247 members of the American Sociological Association (ASA) declared "Leisure/Sport/Recreation" as one of their three major areas of interest, and over half those scholars focused primarily on leisure rather than sports. Only thirty-seven ASA members identified "Leisure/Sports Recreation" as their primary research and/or teaching topic (0.3 percent of ASA members), and only two Canadian and two U.S. sociology departments offer a graduate program in the sociology of sport, according to the 1998 Guide to Graduate Departments of Sociology. At the 1998 annual ASA meeting, there were approximately 3,800 presenters and copresenters, and only 20 dealt with sport-related topics in their presentations; only 2 of the 525 sessions were devoted to the sociology of sport. Patterns are similar in Canada, Great Britain, and Australia (Rowe et al. 1997).

In physical education and kinesiology, the primary focus of most scholars has been on motor learning, exercise physiology, biomechanics, and physical performance rather than the social dimensions of sports (see Sage 1997). Social and cultural issues have not been given a high priority in the discipline except when research has had practical implications for those who teach physical education, coach athletes, or administer sport programs. As the legitimacy and role of physical education departments have been questioned in many

universities, the scholars in those departments have been slow to embrace the frequently critical analyses of sports done by those who use sociological theories and perspectives. Therefore, studying sports as social phenomena has not earned many scholars high status among their peers in physical education and kinesiology departments. However, the majority of sociology of sport scholars with doctorates have earned their degrees and now have options in departments of physical education or kinesiology and departments of sport studies and human movement studies.

There have been noteworthy indications of change. For example, there are a number of journals devoted to social analyses of sports (*Sociology of Sport Journal*, *International Review for the Sociology of Sport*, *Journal of Sport & Social Issues*, *Culture, Sport, Society*). Many mainstream journals in sociology and physical education now accept and publish research that uses sociological perspectives to study sports. National and regional professional associations in sociology and physical education in many countries sponsor regular sessions in the sociology of sport at their annual conferences. Annual conferences also are held by a number of national and regional sociology of sport associations around the world, including those in Japan, Korea, and Brazil as well as the countries of North America and Europe. The International Sociology of Sport Association (ISSA, formerly the ICSS) holds annual conferences and meets regularly with the International Sociological Association. Attendance at many of these conferences has been consistent, and the quality of the programs has been impressive. The existence of such organizational endorsement and support, along with continued growth in the pervasiveness and visibility of sports in society, suggests that the discipline will continue to grow.

Among other indications of growth, articles in the *Sociology of Sport Journal* are cited regularly in social science literature. Scholars in the field are recognized as “public intellectuals” by journalists and reporters associated with the mass media. Quotes and references to sociology of sport research appear increasingly in the popular print and electronic media. Amazon.com, the world’s major Internet bookseller, listed over 260 books in its “Sociology of Sport” reference category in March 1999. Most important, major publishers such as McGraw-Hill estimate that every year nearly

30,000 university students take courses in the “sport in society” category.

Complicating the issue of future growth is the fact that scholars in this field regularly disagree about how to “do” the sociology of sport. Some prefer to see themselves as scientific experts who do research on questions of organization and efficiency, while others prefer to see themselves as facilitators or even agents of cultural transformation whose research gives a voice to and empowers people who lack resources or have been pushed to the margins of society. This and other disagreements raise important questions about the production and use of scientific knowledge, and many scholars in the sociology of sport are debating those questions. As in sociology as a whole, the sociology of sport is now a site for theoretical and paradigmatic debates that some scholars fear will fragment the field and subvert the maintenance of an institutionalized professional community (Ingham and Donnelly 1997). Of course, this is a challenge faced in many disciplines and their associated professional organizations.

#### CONCEPTUAL AND THEORETICAL ISSUES

Through the mid-1980s, most research in the sociology of sport was based on two assumptions. First, sport was assumed to be a social institution similar to other major social institutions (Lüschen and Sage 1981). Second, sports were assumed to be institutionalized competitive activities that involve physical exertion and the use of physical skills by individuals motivated by a combination of personal enjoyment and external rewards (Coakley 1990). These conceptual assumptions identified the focus of the sociology of sport and placed theory and research on sports within the traditional parameters of sociological theory and research.

Theory and research based on these assumptions were informative. However, many scholars in the field came to realize that when analytic attention is focused on institutionalized and competitive activities, there is a tendency to overlook the lives of people who have neither the resources to formally organize their physical activities nor the desire to make them competitive. Scholars became sensitive to the possibility that this tendency can reinforce the ideologies and forms of social organization that have disadvantaged certain categories and collections of people in contemporary

societies (Coakley 1998). This encouraged some scholars to ask critical questions about sports as contested activities in societies. Consequently, their research has come to focus more on the connections between sports and systems of power and privilege and the changes needed to involve more people in the determination of what sports can and should be in society.

These scholars used an alternative approach to defining sports that revolved around two questions: What gets to count as a sport in a group or society? and Whose sports count the most? These questions forced them to focus more directly on the social and cultural contexts in which ideas are formed about physical activities and the social processes that privilege some forms of physical activities. Those who have used this approach also note numerous cultural differences in how people identify sports and include them in their lives. In cultures that emphasize cooperative relationships, the idea that people should compete for rewards may be defined as disruptive, if not immoral, and for people in cultures that emphasize competition, physical activities and games that have no winners may seem pointless. These cultural differences are important because there is no universal agreement about the meaning, purpose, and organization of sports. Similarly, there is no general agreement about who will participate in sports, the circumstances in which participation will occur, or who will sponsor sports or the reasons for sponsorship. It is now assumed widely by scholars who study sports that these factors have varied over time from group to group and society to society and that sociological research should focus on the struggle over whose ideas about sports become dominant at any particular time in particular groups or societies. This in turn has highlighted issues of culture and power relations in theory and research in the sociology of sport.

Before the mid-1980s, most research and conceptual discussions in the sociology of sport were inspired or informed by structural functionalist theories and conflict theories (Lüschen and Sage 1981; Coakley 1990), and in parts of western Europe, figurational sociology was used by some scholars who studied sports (see Dunning 1992). Those with structural functionalist perspectives often focused on questions about sports and issues of socialization and character development, social integration, achievement motivation, and struc-

tural adaptations to change in society. The connections between sports and other major social institutions and between sports and the satisfaction of social system needs were the major topics of concern.

Those who used conflict theories viewed sports as an expression of class conflict and market forces and a structure linked to societal and state institutions. Their work was inspired by various interpretations of Marxist theory and research focused generally on connections between capitalist forms of production and consumption and social behaviors in sports and on the ways in which sports promote an ideological consciousness that is consistent with the needs and interests of capital. Specifically, they studied the role of sports in processes of alienation, capitalist expansion, nationalism and militarism, and racism and sexism (Brohm 1978; Hoch 1972).

Figurational, or “process,” sociology was and continues to be inspired by the work of Elias (Elias 1978; Elias and Dunning 1986; Jarvie and Maguire 1994). Figurational sociologists have focused on issues of interdependence and interaction in social life and have identified historical linkages between the structure of interpersonal conduct and the overall structure of society. Unlike other theoretical approaches, figurational sociology traditionally has given a high priority to the study of sport. Figurational analyses have emphasized sports as a sphere of social life in which the dichotomies between seriousness and pleasure, work and leisure, economic and noneconomic phenomena, and mind and body can be shown to be false and misleading. Before the mid-1980s, research done by figurational sociologists focused primarily on the historical development of modern sport and the interrelated historical processes of state formation, functional democratization, and expanding networks of international interdependencies. Their best known early work focused on linkages between the emergence of modern sports and the dynamics of civilizing processes, especially those associated with the control of violence in society (Elias and Dunning 1986).

Since the mid-1980s, the sociology of sport has been characterized by theoretical and methodological diversity. Fewer scholars use general theories of social life such as structural functionalism



and conflict theories. The theories more often used are various forms of critical theories, including feminist theories and hegemony theory; also used are interpretive sociology (especially symbolic interactionism), cultural studies perspectives, and various forms of poststructuralism (Rail 1998). Figurational sociology still is widely used, especially by scholars outside North America. A few scholars have done research informed by the reflexive sociology of Pierre Bourdieu (Laberge and Sankoff 1988; Wacquant 1995a, 1995b) and the structuration theory of Anthony Giddens (Gruneau 1999).

Methodological approaches also vary. Quantitative data and statistical analyses remain popular, although various qualitative methods and interpretive analyses have become increasingly popular, if not the dominant research approaches in the field (Donnelly 2000). Ethnography and in-depth interviewing, along with textual and discourse analysis, have emerged as common methodologies among many scholars studying sports and sport participation (Coakley and Donnelly 1999). Quantitative methods have been used most often to study issues and questions related to sport participation patterns, the attitudinal and behavioral correlates of participation, and the distribution of sports-related resources in society. Both quantitative and interpretive methods have been used to study questions and issues related to socialization, identity, sexuality, subcultures, the body, pain and injury, disability, deviance, violence, emotions, the media, gender relations, homophobia, race and ethnic relations, new and alternative sports forms, and ideological formation and transformation (Coakley and Dunning 2000).

#### FINAL NOTE

Sociologists study sports because they are prominent and socially significant cultural practices in contemporary societies. The sociology of sport contains an active, diverse, and slowly expanding collection of scholars united by professional organizations and academic journals. Continued growth of the field depends on whether these scholars continue to do research that makes meaningful contributions to the way people live their lives and recognized and visible contributions to knowledge in sociology as a whole.

#### REFERENCES

- Brohm, Jean-Marie 1978 *Sport—A Prison of Measured Time*, trans. I. Frasier. London: Ink Links.
- Coakley, J. 1990 *Sport in Society: Issues and Controversies*, (4th ed.). St. Louis: Mosby.
- , 1998 *Sport in Society: Issues and Controversies*, 6th ed. New York: McGraw-Hill.
- , and P. Donnelly, eds. 1999 *Inside Sports*. London: Routledge.
- eds. 2000 *Handbook of Sport and Society*. London: Sage.
- Donnelly, P. 2000 “Interpretive Approaches to the Sociology of Sport.” In J. Coakley and E. Dunning, eds., *Handbook of Sport and Society*. London: Sage.
- Dunning, E., ed. 1971 *The Sociology of Sport*. London: Cass.
- 1992 “Figurational Sociology and the Sociology of Sport: Some Concluding Remarks.” In E. Dunning, and C. Rojek, eds., *Sport and Leisure in the Civilizing Process*. Toronto: University of Toronto Press.
- Edwards, H. 1969 *The Revolt of the Black Athlete*. New York: Free Press.
- Elias, N. 1978 *The Civilizing Process*, vol. 1: *The History of Manners*. Oxford, UK: Blackwell.
- , and E. Dunning, eds. 1986 *Quest for Excitement: Sport and Leisure in the Civilizing Process*. Oxford, UK: Blackwell.
- Gramsci, A. 1971 *Selections from the Prison Notebooks*, trans. and ed. Q. Hoare and G. Smith. New York: International Publishers.
- Gruneau, R. 1999 *Class, Sports, and Social Development*. Champaign, Ill. Human Kinetics.
- Guttman, A. 1978 *From Ritual to Record: The Nature of Modern Sports*. New York: Columbia University Press.
- 1994 *Games and Empires: Modern Sports and Cultural Imperialism*. New York: Columbia University Press.
- Hoch, P. 1972 *Rip Off the Big Game: The Exploitation of Sports by the Power Elite*. Garden City, N.Y.: Anchor.
- Huizinga, J. 1955 *Homo Ludens: A Study of the Play Element in Culture*. Boston: Beacon Press.
- Ingham, A. G., and P. Donnelly 1997 “A Sociology of North American Sociology of Sport: Disunity in Unity, 1965–1996.” *Sociology of Sport Journal* 14(4):362–418.
- Jarvie, G., and J. Maguire 1994 *Sport and Leisure in Social Thought*. London: Routledge.
- Kenyon, G. S., ed. 1969 *Aspects of Contemporary Sport Sociology*. Chicago: Athletic Institute.
- Krotee, M., ed. 1979 *The Dimensions of Sport Sociology*. West Point, N.Y.: Leisure Press.

- Laberge, S., and D. Sankoff 1988 "Physical Activities, Body Habitus and Lifestyles." In J. Harvey and H. Cantelon eds., *Not Just a Game: Essays in Canadian Sport Sociology*. Ottawa: University of Ottawa Press.
- Lever, J., and S. Wheeler 1993 "Mass Media and the Experience of Sport." *Communication Research* 20(1):299-313.
- Loy, J. W., G. S. and Kenyon, eds. 1969 *Sport, Culture, and Society*. London: Collier-Macmillan.
- Lüschen, G. ed. 1970 *The Cross-Cultural Analysis of Sport and Games*. Champaign, Ill. Stipes.
- , and G. H. Sage 1981 "Sport in Sociological Perspective." In G. Lüschen and G. H. Sage, eds., *Handbook of Social Science of Sport*. Champaign, Ill. Stipes.
- Rail, G., ed. 1998 *Sport and Postmodern Times*. Albany: State University of New York Press.
- Rowe, D., J. McKay, and G. Lawrence 1997 "Out of the Shadows: The Critical Sociology of Sport in Australia, 1986-1996." *Sociology of Sport Journal* 14(4):340-361.
- Sage, G. H. 1997 "Physical Education, Sociology, and Sociology of Sport: Points of Intersection." *Sociology of Sport Journal* 14(4):317-339.
- 1998. *Power and Ideology in American Sport*. Champaign, Ill. Human Kinetics.
- Wacquant, L. J. D. 1995a "The Pugilistic Point of View: How Boxers Feel about Their Trade." *Theory and Society* 24:489-535.
- 1995b "Pugs at Work: Bodily Capital and Bodily Labour among Professional Boxers." *Body & Society* 1(1):65-93.

JAY COAKLEY  
JANET LEVER

## STANDARDIZATION

Standardization is a technique used in comparing indicators from two or more populations. The goal of the standardization procedure is to control for compositional differences between these groups that may influence the indicator that is being examined. This method allows a researcher to determine the extent to which differences in the rates of events between populations are due to differences in population characteristics. Often sociologists ask questions, that require comparisons between groups of people: Which city has a

higher crime rate? Which country has lower mortality? Which ethnic group is more likely to coreside with elderly family members? In making these comparisons, one usually calculates a summary measure: crimes per capita, crude death rate, or the proportion of elders living with family members. However, any two groups of people are likely to differ along several dimensions, such as age, educational level, race, and income. These dimensions, or factors, also may be related to the event being explored. As a result, the summary measure to some extent reflects the compositional differences in the groups being studied.

Standardization historically has been a central aspect of demographic methods (Bogue 1969; Hinde 1998; Murdock and Ellis 1991; Shryock and Siegel 1980), but its importance extends beyond that use to a way of thinking about summary or aggregate measures. While offering the advantage of conciseness, aggregate measures mask underlying compositional differences, and the use of standardization represents an acknowledgment that population characteristics influence the rate at which events occur in a population. Summary indicators are very useful; they provide a single number for comparison rather than a whole series of numbers, and they are easily calculated. However, comparisons among population groups or among subgroups in a population should account for the differing compositional makeup of those groups. Demographers have been led to standardization for several reasons. First, there is a natural desire to make comparisons between groups along demographic indicators: crude death rates, crude birthrates, marriage rates, and employment, among others. Standardization allows these comparisons to reflect differences in the underlying processes, rather than being confounded by the effects of composition. Standardization procedures can accommodate the effects of a single factor or many factors, leaving the technique bounded only by the available data. Standardization also allows the estimation of indicators for groups for which data are incomplete or of poor quality.

Many demographic measures are affected by the composition of the population, particularly the age distribution. Age composition is especially critical in considering crude death rates, since mortality rates have a very distinctive age-specific pattern: high at very young and very old ages.

Populations with a large proportion of persons in those age groups experience a large number of deaths, regardless of age-specific rates of mortality. Two populations with identical sets of age-specific rates of mortality but different age distributions will have different crude death rates. The removal of the “interference” of age distribution from the summary measure—the crude death rate—is the goal of the standardization procedure. In the rest of this article, the standardization procedure will be explained using mortality rates, and then several other examples of standardization will be presented.

The first step in a comparison is to calculate a crude rate or proportion. Crude rates or proportions are calculated by the formula

$$CR = \frac{E}{P} \quad (1)$$

where  $E$  refers to the number of events of interest in the population during the time period and  $P$  refers to the population during that period. If the population is measured at the middle of the year and the events occur throughout the year, this proportion can be interpreted as a rate. In cases where this proportion is small, for instance, mortality rates, the crude rate commonly is multiplied by 1,000 and reported as the number of events per 1,000 people.

Crude rates or proportions are used to represent a variety of characteristics of a population. These rates have an advantage over a comparison of absolute numbers, since they account for differences in size between two populations. Obviously, in a comparison of the annual number of homicides in Chicago versus that in Seattle, one must account for the fact that the population of Chicago is 2.8 million people compared to about one-half million in Seattle. Similarly, comparing the number of deaths in the United States (over 2 million) to those in Sweden (about 90,000) in 1994 would be unreasonable without knowing that the population of the United States is three times that of Sweden.

Despite the advantage of crude rates over absolute numbers, crude rates are influenced by the composition of the populations being compared. If the event of interest varies by some factor and the two populations have varying levels of that

factor, the crude rates will partly reflect this compositional variation rather than only a difference in the rate at which the event is occurring. If the populations being compared are standardized with respect to the factor, any remaining difference between the crude rates can be attributed to a true difference in rates of occurrence. If the difference in the crude rate disappears, one can conclude that the compositional variation rather than a difference in the underlying rates of occurrence led to a difference in the crude of events.

To understand the rationale of standardization, it is necessary to recognize that in essence, the crude rate is a weighted average of a set of factor-specific rates, where the weights are the distribution of the factor in the population. Thinking in this manner, one can rewrite the crude rate as

$$CR = \sum \frac{e_a}{p_a} \frac{p_a}{P} \quad (2)$$

where  $p_a$  is the population in group  $a$  and  $e_a$  is the number of events occurring in group  $a$ . The sum of all  $e_a$  equals the total number of events,  $E$ , and the sum of all  $p_a$  equals the total population,  $P$ . Note that this equation has two components. The first,  $e_a/p_a$ , represents the group-specific rate of events or the group-specific proportion, which sometimes is expressed as  $m_a$ . The second component of the rate calculation,  $p_a/P$ , represents the proportion of the population in each of the  $a$  groups. These are the two series of elements needed to apply the direct standardization technique. Using this notation, the crude rate can be rewritten as

$$CR = \sum m_a \cdot \frac{p_a}{P} \quad (3)$$

When the formula for the crude rate is written in this manner, it is easy to see how the composition of the population, that is, its distribution among the  $a$  groups, affects the crude rate. If the group-specific rate  $m_a$  is high when the proportion of the population in that group,  $p_a/P$ , is large, more events will be observed in the total population than will be observed if  $p_a/P$  is small. Similarly, if  $m_a$  is small when  $p_a/P$  is small, few events will occur.

A comparison of the crude death rates in Sweden and the United States provides an example of the use of standardization. Sweden has one of the world's highest life expectancies at birth,

approximately 76 years for men and 81.4 years for women in 1994. The crude death rate of Sweden, however, was about 10.4 deaths per 1,000 in that year. In contrast, life expectancy at birth in the United States was 72.2 years for men and 78.8 years for women in 1993, and the crude death rate was about 8.6 deaths per 1,000 in that year (United Nations 1997). It seems natural to expect that the country with the longest life expectancy would also have the lowest crude death rate, so what accounts for this discrepancy? To understand the reason for this difference in the crude rates, it is necessary to observe the differing age distributions of the two populations. In the United States about 13 percent of the population is over age of 65; while in Sweden over 17 percent of people are over that age. Since death rates are highest in this age range, the larger proportion of the Swedish population in old age creates more deaths, even with lower age-specific death rates. Standardization demonstrates the extent to which these differences in age distribution account for the difference in the crude death rate.

As was mentioned above, this method of standardization—direct standardization—requires a standard population distribution and a set of factor-specific rates for the populations being studied. Direct standardization uses this standard population to calculate new standardized crude rates for the populations of interest. In this case, the population distribution of the standard population replaces the observed population distribution. Since each population's crude rate will be calculated with the same distribution, the effect of the compositional differences will be eliminated and each population will have the same composition. To apply direct standardization, the formula

$$DSR = \sum \frac{e_a^j \cdot p_a^s}{p_a^j \cdot P^s} \quad (4)$$

is used, where  $e_a^j$  represents the number of events occurring in group  $a$  in population  $j$ ,  $p_a^j$  represents the population size of group  $a$  in population  $j$ ,  $p_a^s$  represents the number of people in group  $a$  in the standard population  $s$ , and  $P^s$  represents the standard population. Comparing equations (2) and (4) shows the similarities. The second term in equation (2), the compositional distribution of the population of interest,  $p_a/P$ , has been replaced with the compositional distribution of the stan-

dard population,  $p_a^s/P^s$ . The first term in the crude rate calculation remains the factor-specific rate in the population of interest, population  $j$ .

Returning to the example of the United States and Sweden, using the age distribution of the United States as the standard distribution and computing a standardized crude death rate for Sweden by applying the age-specific death rates of Sweden yields a standardized crude death rate of 7.6 deaths per 1,000 for Sweden. Instead of being higher than the crude death rate in the United States, Sweden's crude death rate falls below that of the United States. At least part of the difference in the crude rates therefore is due to Sweden's older population rather than to a difference in age-specific death rates. In general, populations with a relatively old age distribution tend to have higher crude death rates than do populations with similar age-specific mortality patterns, since death rates are higher at older ages.

The data demands for direct standardization, while not overwhelming, can be difficult to meet if there is limited information on factor-specific rates in one of the populations of interest. For example, in many studies of mortality in less developed countries or in a historical perspective, information on age-specific death rates may be missing or unreliable. In these cases, an alternative method referred to as indirect standardization can be used. Indirect standardization requires knowledge only of the composition of the population and the total number of events of interest. Direct standardization involves the application of population-specific sets of rates to a standard population; conversely, indirect standardization involves the application of a standard set of rates to individual population distributions. In indirect standardization, a set of standard rates is applied to the population and the expected number of events is compared to the actual number. This standardizing ratio is estimated by the formula

$$SR = \frac{E^j}{\sum m_a^s p_a^j} \quad (5)$$

where  $E^j$  is the actual number of events in the population  $j$ ,  $m_a^s$  is the factor-specific rate in the standard population  $s$ , and  $p_a^j$  is the number of people in population  $j$  who are in group  $a$ . The denominator of the ratio calculates the number of

events that would be expected in population  $j$  if the factor-specific rates of the standard population were applied to the population. When the event of interest is death, this ratio often is referred to as the standardized mortality ratio. To obtain the new indirectly standardized crude rate, this standardizing ratio is multiplied by the crude rate for the standard population:

$$ISR = SR \cdot CR^s \quad (6)$$

where  $CR^s$  is the crude rate in the standard population. These indirectly standardized crude rates then can be compared to each other. Obviously, when the standardizing ratio is greater than 1.0, the *ISR* will be larger than the crude rate for the standard population, and when the standardizing ratio is less than 1.0, the *ISR* will be smaller than the standard population's crude rate.

Indirect standardization does not control for composition as well as the direct standardization method does but should yield similar results in terms of direction and magnitude. Returning to the example of Sweden and the United States, the actual number of recorded deaths in Sweden would be greater than the observed number if U.S. age-specific death rates were applied to the Swedish population's age distribution. The resulting standardized mortality ratio would be 0.912, and when that was multiplied by the crude rate for the United States, the *ISR* for Sweden would be 7.8, very similar to the result obtained through direct standardization.

When indirect standardization is employed, there is no choice to be made about the standard population; this method is used when only one population distribution is available. The choice of the standard population for direct standardization should be considered carefully, but within reasonable bounds the choice of standard should not alter the conclusions radically. Researchers generally are interested in the direction and approximate size of differences between the groups, and these values are preserved with the choice of any of a number of reasonable standard populations. There are three general choices for the standard: use one of the populations being studied, use an average of the populations, or use a population outside those being studied. Each of these choices has advantages and disadvantages. Theoretically,

the choice of standard should be made to minimize the effects of that choice on the results.

Using one of the populations being studied eliminates the need to standardize that population and often makes the explication of comparisons easier. For instance, in comparing crime rates across several cities, choosing one city as the basis for comparison may be appropriate. When comparisons are made of a population over time, it is standard procedure to choose a distribution that is representative of the middle of the time period. For instance, in a study of mortality change between 1950 and 1990 in the United States, it would be appropriate to use the 1970 census for the standard age distribution. A drawback to using one of the study populations as the standard, however, can be that the population chosen has an unusual distribution of factors. This unusual distribution may skew the summary measures in a way that is inconsistent or difficult to interpret. Also, choosing one of the populations as a standard can carry implications that this distribution is the "ideal" or "correct" distribution and may place interpretational burdens on the results.

Using an average of the populations eliminates the problem of setting one population as the ideal and ameliorates the problem of unusual distributions. A comparison of racial differences in mortality in the United States, for example, might use the age distribution of the total U.S. population, an unweighted average of the distribution of each racial group, as the standard. This choice eliminates the assumption that any one population has a preferred distribution and allows for meaningful comparisons among groups. The use of an aggregate population as the standard is encountered frequently in comparisons of subgroups within a national population.

A third choice is to pick a population completely exogenous to the study as a standard. This choice most often involves an artificial population that is representative of a standard pattern of factor distributions. Several sources of standard populations exist. In the case of age, Coale and Demeny's (1983) set of regional model life tables contains sets of age distributions typical of a variety of mortality levels and patterns. The use of an external standard eliminates any value judgments associated with the choice of standard. An exter-

nal standard also can be chosen to minimize or eliminate extreme distributions of factors. The external standard also provides a way of comparing very diverse populations. Again, the choice of standard should match the populations being studied as closely as possible to minimize the effect of that choice on the results.

An exogenous standard also might be employed as a way to simulate the effects of a variety of changes in population composition on the crude rate. This use of the standardization technique highlights the underlying logic of the procedure by using the method to investigate the extent to which compositional changes influence aggregate comparisons. Here the technique is used as a methodological device to explore the effects of changes. For instance, a researcher might be interested in the effects on average wages of changing occupational structures among men and women. A testable hypothesis could be that as women approach men in terms of occupational distribution, the gender gap in wages will disappear. If a variety of simulated occupational structures are applied to a set of gender- and occupation-specific wage rates, the effect of occupational structure on the wage gap can be examined.

Since standardization developed in the field of demography, most applications involve the study of demographic phenomena. The example of the United States and Sweden involved comparisons of mortality rates. However, standardization is used widely in other areas as well. For example, the U.S. Census Bureau routinely reports the distribution of the American population aged 15 and older among marital states, and historical comparisons of this distribution are used to examine changes in marital behavior over time. However, the age composition of the population can greatly influence the distribution among marital states, particularly when the proportion of the population in the age range of 15 to 25 years is very large. In 1960, 65.6 percent of women aged 15 and older were married compared to 60.4 percent of similarly aged women in 1975 (United States Bureau of the Census 1976). At first glance, these comparisons seem to signal a retreat from marriage: A smaller proportion of women was married in 1975 than in 1960. However, when the age distribution of the population is standardized to the 1960 population, the proportion married in 1975 in-

creases to 63.5. While this is still a decline compared to 1960, the magnitude of the change is much less. The difference in the proportion married is due largely to a difference between 1960 and 1975 in the proportion of women just over the age of 15, the baby boomers, who were young teenage women who had not yet married.

Standardization can be used to control for characteristics other than age. Suppose, for instance, one is comparing the health status of two different groups: elderly white Americans and elderly African-Americans. If we compare the proportion of each group in poor health, we find that 34 percent of elderly whites and 50 percent of elderly African-Americans report their health as fair or poor. However, we know that health status varies by education and that the educational distributions of these two groups differ. Among elderly whites, about 12 percent have fewer than eight years of school, compared to 39 percent of elderly African-Americans. Clearly, since lower levels of education are associated with poorer health and elderly African-Americans have lower levels of educational attainment, some of the difference in observed health status between the groups can be expected to result from the different educational compositions.

It is desirable to compare these two groups without the influence of education. Using the educational distribution of the elderly white population as a standard and applying the observed education-specific rates of poor health among elderly African-Americans, one obtains an overall proportion of 42 percent in poor health, compared to the unstandardized proportion of 50 percent. Thus, if the African-American older population had an educational distribution similar to that of the more highly educated white elderly population, the expected health status of older African-Americans would improve.

Lichter and Eggebeen (1994) used standardization techniques to examine the effects of parental employment on rates of child poverty. In their work these researchers use direct standardization techniques in two different ways. In the first, they simulate the effects of a variety of assumptions about parental employment patterns on children's poverty rates. This is an illustration of using an "exogenous" or artificial population

distribution as a standard population. By changing the employment distribution of the parents of children in poverty, they determine that only modest declines in child poverty would result from increasing those levels of employment. Their second application of standardization compares the poverty rates of black children obtained by using the employment distribution of white parents as the standard to the rates directly observed. In this case, they have chosen one of the study populations as the standard and are interested in the extent to which differences in child poverty between blacks and whites are determined by factors other than parental employment distributions. They find in fact that parental employment differences among female-headed families account for a substantial portion of the observed differences in child poverty.

Standardization can control for more than one factor at a time and can be applied to more than two groups. Himes et al. (1996) standardize for age, sex, and marital status in an examination of the living arrangements of minority elderly in the United States. Living arrangements are known to be different for men and women, for married and unmarried, and for younger and older elderly. These factors—age, sex, and marital status—also are known to vary across racial and ethnic subgroups. Therefore, the observed differences in living arrangements are likely to be due in part to these underlying characteristics rather than being a reflection of differences in attitudes or beliefs. Standardization allows a comparison among groups without the influence of these compositional differences. In this research, the compositional distribution of the entire United States with respect to age, sex, and marital status was chosen as the standard. In this analysis, the standardization procedure had the greatest effect on comparisons of the African-American population and much smaller effects on the white, non-Hispanic, Hispanic, Asian, and Native American populations.

Standardization is widely used in a variety of sociological inquiries. While it originated in demographic analyses, it can be applied to a variety of questions in which a researcher wants to determine the extent to which compositional differences in population groups account for observed differences in summary measures. Standardization is also useful as a simulation technique, allow-

ing researchers to explore the effects of a variety of compositional changes on a summary indicator. Researchers should bear in mind, however, that the results of standardization are merely artificially constructed indicators; they do not represent a real population or circumstance.

## REFERENCES

- Bogue, Donald J. 1969 *Principles of Demography*. New York: Wiley.
- Coale, Ansely J., and Paul Demeny 1983 *Regional Model Life Tables and Stable Populations, 2nd ed.* New York: Academic Press.
- Himes, Christine L., Dennis P. Hogan, and David J. Eggebeen 1996 "Living Arrangements of Minority Elders." *Journal of Gerontology: Social Sciences* 51B:S42-S48.
- Hinde, Andrew 1998 *Demographic Methods*. New York: Oxford University Press.
- Lichter, Daniel T., and David J. Eggebeen 1994 "The Effect of Parental Employment on Child Poverty." *Journal of Marriage and the Family* 56:633-645.
- Murdock, Steve H., and David R. Ellis 1991 *Applied Demography: An Introduction to Basic Concepts, Methods, and Data*. Boulder, Colo.: Westview Press.
- Shryock, Henry S., and Jacob S. Siegel 1980. *The Methods and Materials of Demography*, 4th printing (rev.). Washington, D.C.: U.S. Government Printing Office.
- United Nations 1997 *Demographic Yearbook 1995*. New York: United Nations.
- United States Bureau of the Census 1976 *Social Indicators 1976*. Washington, D.C.: U.S. Government Printing Office.

CHRISTINE L. HIMES

## STATE, THE

The term "state" denotes the complex of organizations, personnel, regulations, and practices through which political power is exercised in a territory. In simple societies organized as bands of families, as tribes, or as chiefdoms, political power is not separated from power relationships rooted in kinship structures or religion. Those societies also lack organizations and specialized personnel (beyond the chief) for exercising political authority



and therefore have no real states. The state emerged only with the development of more complex societies, either cities or tribal confederations, which formed the bases for city-states, monarchies, and empires. Monarchies and empires in turn have given way to liberal states, modernizing dictatorships, and one-party states as the most widespread current forms of states.

The “state” is a rather abstract term. Over time and space, the concrete organizational forms, the kinds of personnel, the specific laws and regulations, and the practices of states have varied greatly with the historical development of societies and across different cultures and regions. The modern nation-state is a very particular kind of state that developed in Europe in the eighteenth and nineteenth centuries and currently is spreading across the world (Poggi 1990). However, like other forms of the state, this organizational form is likely to have its day and then fade; already various kinds of supranational and international bodies have begun to take over some of the political power formerly monopolized by nation-states.

The basis of the state is political power. This article examines the roots of that power and then explores the various forms taken by states from their beginnings to the present day.

### **POLITICAL POWER**

All forms of power involve the ability of powerholders to coerce others into giving up their property, their free choice of action, and even their lives. Political power, as opposed to economic power (based on money or other forms of wealth), religious power (based on relationships to transcendent forces), family power (based on sex, seniority, and kin relationships), and pure coercion (based on brute force), is rooted in the recognition of the rightful authority of the ruler (Weber 1968). That authority stems from the demands within a society for specialists with the ability to mediate and coordinate.

Any group of human beings in regular interaction among themselves is prone to conflict over possessions, decisions regarding group actions (to hunt or not, to camp here or there, to fight or flee from a threat), and individual actions that give offense (insults, injury, infidelity). In small groups,

such conflicts usually can be settled through the arbitration of respected family members or elders, but in larger groups or groups in which much interaction occurs among nonkin, those conflicts produce demands for justice that require a more broadly recognized form of mediation. Individuals who are particularly skilled at mediating such conflicts, who gain a reputation for wisdom and justice, can acquire the role of a specialist in settling conflicts. In addition, every group of human beings faces external threats from wild animals, the weather, and other human groups. Individuals who are particularly skilled at coordinating actions within a group for the purposes of attack, hunting, and defense can gain a reputation that translates into a calling as a specialist in coordinating group actions to meet threats.

The functions of mediation to produce internal justice and of coordination to deal with external threats are distinct; indeed Native American tribes sometimes had a “peace chief” and a “war chief” who specialized in those functions. Modern societies have legal-judicial systems and executive-military systems that show a similar division of functions. However, these functions tended to merge because in both cases it was necessary to have mechanisms to compel compliance with the arbitration decisions of the mediator or the action directives of the coordinator. Once a society develops regular means to compel compliance with those decisions and directives (generally armed warriors closely attached to or under the direct supervision of the mediator or coordinator), that society is on its way to developing a state. Political power is thus the authority given to a recognized leader (whether judge or general) to compel compliance with his or her decisions.

Political power, however, is a two-edged sword. On the one hand, the power of the leader must be sufficient to ensure that arbitration is enforced and that the coordination of military, hunting, or building activities is effective. The larger a society is, the more complex its economy is, the stronger its enemies are, and the more threatening and varied its environment is, the greater are the tasks facing the state. Thus, for a society to avoid turmoil and defend itself, it must grow in organizational size, complexity, and power along with the society of which it is a part. On the other hand, as the leaders acquire control of larger and richer



organizations and larger and more powerful coercive forces, there is a danger that that organizational and coercive force will be used to enrich and serve the desires of the ruler, not to meet the demands for justice and protection of the population (Mann 1986).

The history of the state is thus a history of balancing acts and often of overreaching. State rulers frequently use their organization and authority to expand their power and wealth. Some rulers invest heavily in conquest, acquiring power over new regions and peoples by brute force and then setting up organizations and laws to acquire and enforce political authority. Other rulers have sought to distinguish themselves primarily as lawgivers or (e.g., King Solomon) paragons of justice. Still others have simply taken their power as given and abused it. Sometimes they gain mightily from such abuse, but at other times—under very particular conditions—they may become the object of elite revolts or popular revolutions.

For sociologists, the key to understanding the state is knowledge about the shifting relationships between state rulers, their organizations and resources, and their societies. Much of the history of the development of state forms comes from the competition between rulers seeking to extend their control of political organizations and coercive force and elite and popular groups seeking to limit or channel political authority into socially acceptable goals and actions.

### **CITY-STATES, EMPIRES, AND FEUDALISM**

Although cities and states initially may have developed independently, with both gradually moving forward between 8000 and 3000 B.C., by the third millennium B.C., the conjunction between urbanization and state making was firmly established in the Middle East. Elsewhere—in sub-Saharan Africa and southeast Asia (especially Java and Cambodia)—states and even empires developed without true cities; those states operated through dense clusters of villages that often centered on great temple complexes. By contrast, in the Middle East and the New World, large cities grew up around the temple complexes that served as the headquarters and ceremonial centers of the new states. Several of those city-states had great success in

expansion and became the nucleus of larger empires, such as Sumer, Akkad, Assyria, and the empire of the Aztecs.

City-states continued to emerge throughout history, especially in periods of early settlement of new lands (such as the Greek city-states that spread throughout the Mediterranean in the second millennium B.C.) or after the breakup of large empires (as occurred in Italy and along the Rhine in Germany after the collapse of Charlemagne's empire in the ninth century A.D.). The legacy of these city-states is that they experimented with a wide array of state forms. At various times, the Greek and Roman city-states of the eighth through fourth centuries B.C.—including Athens, Sparta, Thebes, Corinth, and Rome—were ruled by a single monarch, pairs of kings (or consuls), oligarchies of the wealthy or well-born aristocrats, and popular assemblies. The modern forms of democracy and monarchy can be traced back to the Greek and Roman city-states of that period. However, city-states generally did not survive in any area for more than a few centuries before being swallowed up by large territorial empires.

Those large territorial empires became the dominant form of the state in much of the world for the next 5,000 years, from roughly 3000 B.C. to A.D. 1900 (Eisenstadt 1963). In the Middle East, the major empires included of Sumer, Akkad, Egypt, Assyria, Babylonia, and Persia and the Hellenistic empires founded by the generals of Alexander the Great. These empires were followed by the Roman Empire, the Byzantine Empire, and the Islamic empires founded by the followers of Mohammed. These empires were followed by the vast empires of the Mongols and the Turks, the last of which was the Ottoman Empire, which ruled large portions of north Africa, the Middle East, and southeastern Europe and lasted until 1923. In Europe, after the fall of the Roman Empire there followed the empires of Charlemagne and his sons. That empire left as a legacy the Holy Roman Empire, which eventually evolved into the Austro-Hungarian Empire, which survived until 1918. After roughly A.D. 1500, much of eastern Europe and central Asia was under the control of the Russian Empire, which lasted until 1917. In China and India, large empires emerged in the third and fourth centuries B.C. In China, the Qin and Han dynasties initiated a pattern of imperial rule that lasted until the birth

of the Chinese Republic in 1911; in India, the Maurya and Gupta dynasties briefly unified the subcontinent and were followed by the Mughal Empire, which lasted until India came under British domination in the eighteenth century.

In Africa, there also were large Empires, including the Aksum Empire in Ethiopia which was (founded around 300 B.C. and whose successor empires and dynasties lasted until 1974), the Ghana Empire and Mali Empire in west Africa, Great Zimbabwe and Mutapa in southern Africa, and the Zulu Empire, which ruled over much of southeastern Africa until it was defeated by the British in the late nineteenth century. In the Americas, three major indigenous empires developed: the Maya and the Aztecs in what is today Mexico and the Incas centered in modern-day Peru. After defeating the Aztecs and Incas in the sixteenth century, Spain established an empire in the Americas extending from Chile to California that it ruled for nearly 300 years.

The vast majority of these empires were conquest empires in which strong imperial centers acquired territory, troops, and resources to build ever-larger empires and thus conquer ever more territory. However, many imperial rulers also were famous lawgivers renowned for establishing justice and order in their empires; they included Hammurabi of Babylonia, Justinian of Rome, and Suleyman the Magnificent of the Ottoman Empire. Their lawcodes were established not to give "rights" to subjects but to produce order by making a clear list of crimes and the penalties that would be imposed.

Though powerful, these empires were not immune to decay and disintegration. Even the longest-lived empires, such as those of Egypt and China, had periods of civil war and broke up into multiple states. Population growth that created pressure on the capacity of the land to yield taxes, military defeat by powerful neighbors, and conflicts among elite factions could all produce disorganization and decay of the imperial state administration. In times of decay, a locally based form of rule known as feudalism often arose.

Feudalism in the strict sense is a pattern of allegiance by oath taking in which a lord gives control of land (a "fief") to a vassal in return for a promise of service. This pattern may have one dominant lord controlling many vassals, or there

may be many lords and many vassals, with some vassals dispensing fiefs and thus becoming lords themselves. In this sense, feudalism is not a state, for no centralized administration has full control of the territory. However, if a single lord manages to emerge as dominant over all the other lords and vassals in a territory and is able to expand his own household and personal administration to exert his will throughout the territory, one can then speak of a state, which usually is described as a kingdom or monarchy. Kingdoms were known throughout the world and generally appear in periods in which large empires have broken down or before they are established. In most of the world, empires continued to reestablish themselves, often building on the strongest kingdom in a region. However, in western and central Europe, no empire ever reestablished lasting control over the area that had been controlled by the Romans. Instead, the period of feudalism in Europe (roughly A.D. 600 through 1300) was followed by many centuries in which a number of competing kingdoms controlled major portions of the European continent.

#### **ABSOLUTISM AND BUREAUCRATIC-AUTHORITARIAN STATES**

The early empires and kingdoms all had rudimentary administrations and relatively undifferentiated elites. That is, the officers of the state were mainly family members of the ruler or personal favorites appointed at the ruler's pleasure; many were also high-ranking officials in the church. They gained much of their income from the control of personal properties or privileges granted by the ruler. The mingling of state and church was based on a strong connection between religious and state power; there was usually an official state religion that supported the state and was in turn supported by the ruler.

By around the sixteenth century A.D., however, most of the kingdoms and empires of Europe and Asia had begun to develop into more impersonal and bureaucratic states. State offices were fixed in a "table of ranks," and officers were expected to undergo rigorous academic training to qualify for their positions. The number of state offices multiplied greatly, and while favorites still were chosen for key positions, an increasing number were chosen and promoted for their merit and

services. States also began to diversify their sources of income. Most early empires relied on various forms of tribute collection or taxes paid “in kind,” such as set amounts of grain, cloth, or labor services. In contrast, by the sixteenth century, most states had begun to specify and collect taxes in cash, with which they paid regular salaries to state officials. In those states, subjects still had few rights and no participation in politics; rulers remained absolute in authority. However, those states became “bureaucratic-authoritarian” in the sense that authority increasingly was exercised through uniform rules enforced by bureaucratic officials rather than through local and customary practices enforced by fairly autonomous local notables.

Dependence on cash meant that many states also placed a greater emphasis on trade and on taxes on commerce as an alternative to taxes on land. For some states (e.g., the Netherlands and Great Britain), taxes on trade and industry soon exceeded revenues from traditional land taxes (Tilly 1990). In the period 1500–1900, the promotion of trade and commerce led to a vast expansion of long-distance trade, both ocean-borne and land-based, across the globe. European kingdoms, stymied in creating empires in Europe, created them overseas. Seeking natural resources and new markets, European states (and later Japan) invested in colonies and overseas companies and administrations to control them in the Americas, Africa, India, southeastern Asia, Korea, and along the Chinese coast.

While this period remained one of kingdoms and empires, bureaucratic-authoritarian states faced two extensive periods of challenge. From the late sixteenth to the mid-seventeenth century and again from the late eighteenth to the mid-nineteenth century, all of Eurasia experienced several trends that reshaped states. First, in those two periods the population grew dramatically, doubling or more, while in other periods the population declined or was stable. These periods of population growth were also periods of rising prices as a result of more extensive commerce and a rising demand for basic goods. Pay to laborers and available land for peasants, however, declined as the population grew faster than did the agricultural economy. Population growth also led to factional conflicts among elite groups competing for control of state offices and to greater demands on state administrations. However, states were run-

ning into financial trouble, for population growth was reducing the surplus available for taxation and the rapid growth of commerce was shifting more resources into areas where traditional tax collection was weak, leaving more resources in the hands of merchants, local landowners, and urban and regional elites. Toward the end of those two periods—roughly 1580–1660 and 1770–1860—conflicts between state rulers and elites over the rulers’ prerogatives and resources triggered worldwide waves of revolutions and rebellions in kingdoms and empires; these included the English, American, and French revolutions, the anti-Habsburg revolts, and the revolutions of 1848 in Europe; the collapse of the Ming Empire and the Taiping rebellion in China; and the janissary, Balkan, and Egyptian revolts in the Ottoman Empire (Goldstone 1991).

#### REVOLUTIONS, NATIONALISM, AND NATION-STATES

Those revolutions and rebellions all involved popular uprisings and elite rebellions against the ruler and loyal elements of the state but had different outcomes in different areas. In most societies, the elites were deeply frightened by popular uprisings and sought to reestablish state power more firmly by tightening the reins of state power and enforcing allegiance to the state-sponsored religion. This was the case in Catholic Spain, Italy, and Austria under the Counter-Reformation; in Confucian China under the Qing dynasty; and in the Islamic Ottoman Empire. However, in England 1689, America in 1776, and France in 1789, the elites were more concerned that excessive state power would damage their positions and fuel future revolutions. Reviving ideas and institutions from the days of democratic Greece and republican Rome, they attempted to place limits on state power and reserve specific rights to elites and even to ordinary workers and peasants. Those limits and rights were codified in a variety of documents, including “declarations of rights”, and especially in constitutions that became the basis for state power. Those constitutions marked a distinctively modern turn in the history of state. Previously, state authority

had always rested on coercion and demands for the dispensation of order and been supported by religious belief and tradition, but from the age of constitutions, the legitimacy of state authority rested on whether the ruler abided by the limits in the constitution and recognized the rights of the elites and popular groups that had established that constitution.

Constitutions meant that a new relationship was forged between the state and the population of the territory it ruled. Under empires, the state established order and most people were simply economic producers, not political actors. By contrast, under constitutions, the people, or at least those involved in creating and establishing constitutional rule, were the ultimate controllers and beneficiaries of state power. This new relationship led to new demands by various groups.

One demand was for greater and more regular political participation by groups that had been excluded: religious and ethnic minority groups, women, and the poor. Though frequently resisted by elites and state rulers, in many areas those groups gained elite allies and acquired rights to regular political participation, most notably through voting (Reuschmeyer et al. 1992). States where voting rights are widespread and the state's power over its subjects has significant limits are commonly described as democratic or "liberal" states. By the late nineteenth century, most of the states in Europe west of Russia and in North and South America were liberal states.

Another demand came from professionals, merchants, and sometimes military officers who lived under empires and wanted to take control of their positions and territories under something like the relationship that prevailed in constitutional regimes, where the state was identified as an instrument of the people rather than the reverse. Those elites argued that every ethnic group should be entitled to its "own" state and its own rulers. The resulting ideology was known as "nationalism" (Calhoun 1998), and it spread widely throughout the world. Nationalism fueled the revolutions of 1830 in Poland and Greece; those of 1848 in Hungary, Germany, Italy, and Romania; the effort

to expel the Austro-Hungarians and unify Italy under Italian rule in the 1860s; and the Serb liberation movement that helped start World War I. Nationalist sentiments also fueled revolts in Ireland throughout the nineteenth and twentieth centuries; the Chinese Republican Revolution of 1911; the anticolonial revolutions in India, Algeria, Indonesia, and Vietnam after World War II; and a host of other anticolonial revolts in Africa and Asia.

Nationalism fostered the ideal that states should be "national" states, reflecting the identity and promoting the aspirations of their inhabitants as a united community rooted in shared traditions and culture (Anderson 1991). In fact, to comply with this ideal, many traditions had to be invented and national languages had to be created. Even today, it often is ambiguous whether a given nation-state reflects a nation (Is there is British nation or only English, Welsh, and Scottish nations plus portions of Scot-settled Ireland sharing the state of Great Britain?). However, the ideal of the nation-state spread widely, even to older states, so that it became expected that modern nation-states would have a national language, a national flag and anthem, national systems of schooling and communications (newspapers, radio, and television), national systems of transportation (highways, railways, and airlines), and a national army.

Nonetheless, since almost all existing states included members of more than one ethnic, linguistic, or cultural group within their boundaries, most nation-states inevitably failed to satisfy to a greater or lesser degree the aspirations of subnational groups, which in turn often developed their own nationalist ambitions. A large number of the violent conflicts in the world in recent years are the result of nationalist movements within nation-states, such as the Chechens in Russia, the Basques in Spain, the Kurds in Turkey, the Uighurs in China, and the Albanian Kosovars in Yugoslavia.

While nationalism seemed poised to bring more liberal, constitutional states into being, things did not develop that way. The defeat of many early nationalist movements led nationalist leaders to conclude that above all else, a people needed a

strong state to protect them from control by others, whether multinational empires or other nations. As a result, many nationalist movements gave rise to authoritarian, populist dictatorships. Those dictatorships often promulgated constitutions and claimed to draw their legitimacy—in the modern fashion—from their service to and identification with the people of the territories they ruled, but in fact they operated in as absolute a manner as any older imperial state, only now they were backed by the latest industrial and military technology. Thus, while nationalism was destroying the old traditional empires and replacing them with modern states, those modern states were following divergent paths into democracy and dictatorship.

### DEMOCRACIES AND DICTATORSHIPS

The history of the state in the twentieth century has largely been one of a struggle between democracies and dictatorships. In the liberal states over the course of the twentieth century, the range of citizen rights has been expanding, the participation in politics of ordinary citizens (through rallies, financial contributions, petitions, and voting) has grown, and the obligations of the state to support its citizens (the modern “welfare state”) have been extended. A major result of these patterns is that women, the working class, and the poor are far more closely integrated into political life in liberal states as voters and direct recipients of state actions than ever before (O’Conner et. al, 1999). To accommodate and channel this political participation, most liberal states have a number of political parties that organize and control the competition for political power. At the end of the twentieth century, as a result of growing state obligations, the personnel and budget of modern liberal states has swollen to the point where state expenditures make up one-quarter to one-half of the entire national product of their societies.

However, the model of the liberal state did not triumph in every place where empires collapsed. In many regions, spurred by nationalist sentiment and the failure of liberal states to provide economic and military security under the chaotic

conditions that followed military defeat or economic crises, modern dictatorships emerged. Some of those dictatorships, such as those of Adolph Hitler and his Nazi Party in Germany and Benito Mussolini and his Fascist Party in Italy, did not outlast their founders. However, in Russia and China, Communist parties took on a dominant life of their own, and those countries became one-party states in which everything of economic, military, and political importance was controlled by the party-state. In other countries, notably in Africa (e.g., Nigeria), Latin America, and eastern Asia (e.g., Korea and Indonesia), military personnel seized power and held on for periods ranging from years to generations. For most of the twentieth century, such modern one-party and military dictatorships, all professing nationalist ideals and even staging (controlled) popular elections, controlled the vast majority of the states and peoples of the world.

In the last two decades of the twentieth century, however, the majority of those one-party states and military dictatorships collapsed (Walder 1995; Goldstone et al. 1991). Their extensive control of the economy stifled innovation and encouraged corruption, leading their revenues to fall well below those of the leading liberal states. Within dictatorial states, even the elites looked on the far greater material wealth and personal freedom of their counterparts in liberal states with envy. Efforts at reform in one-party and military states thus quickly turned into movements to establish liberal regimes. As a result, for the first time in history, it appears that humankind will enter a new millennium with a majority of its nations and populations living under liberal constitutional states (Huntington 1991).

### BEYOND THE NATION-STATE

While the twentieth century has closed with the national, liberal state seemingly triumphant, there is no assurance that this form of state will endure. Constitutional states often have been overthrown by dictatorships, both military and populist, when they encounter severe military or economic setbacks. The Great Depression led a host of democ-

racies to collapse into dictatorships, and struggles with economic development led many Latin American and African states into communist takeovers and military coups in the 1960s and 1970s. In most of the world outside Europe and North America, liberal states are not firmly established and may be vulnerable if another major economic trauma sweeps the globe. Thus, the past threats to the continuance of liberal states may reemerge.

In addition, new threats to the primacy of the nation-state have arisen in the form of supranational organizations with genuine sovereignty and military power. The most notable of these organizations are NATO (a military alliance with a unified command embracing the forces of most European nations and the United States) and the European Union (a supranational body ruled by representatives from most European nations with taxing and legislative authority over certain aspects of its member states). A variety of cooperative multinational organizations established by treaty, such as the United Nations, the International Monetary Fund, the International Court of Justice, and various environmental commissions and human rights organizations, also have impinged on state sovereignty. The future may see still greater transfers of state power to such supranational bodies as the problems of establishing human rights, safeguarding the global environment, and maintaining stable and sound financial institutions may grow beyond the capacity of any single state or ad hoc arrangement of states to resolve.

REFERENCES

Anderson, Benedict 1991 *Imagined Communities*. London: Verso.

Calhoun, Craig 1998 *Nationalism*. Minneapolis: University of Minnesota Press.

Eisenstadt, S. N. 1963 *The Political Systems of Empires*. London: Macmillan.

Goldstone, Jack A. 1991 *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.

—, Ted Robert Gurr, and Farrokh Moshiri 1991 *Revolutions of the Late Twentieth Century*. Boulder, Colo.: Westview.

Huntington, Samuel P. 1991 *The Third Wave: Democratization in the Late Twentieth Century*. Norman: Oklahoma University Press.

Mann, Michael 1986 *The Sources of Social Power*. Cambridge, U.K: Cambridge University Press.

O’Conner, Julia S, Ann Shola Orloff, and Sheila Shaver 1999 *States, Markets, Families: Gender, Liberalism, and Social Policy in Australia, Canada, Great Britain, and the United States*. Cambridge: Cambridge University Press.

Poggi, Gianfranco 1990 *The State: Its Nature, Development, and Prospects*. Stanford, Calif.: Stanford University Press.

Rueschemeyer, Dietrich, Evelyn Huber Stephens, and John D. Stephens 1992 *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Tilly, Charles 1990 *Coercion, Capital, and European States, AD 990–1990*. Oxford, UK: Basil Blackwell.

Walder, Andrew 1995 *The Waning of the Communist State*. Berkeley and Los Angeles: University of California Press.

Weber, Max 1968 *Economy and Society*, edited by Guenther Roth and Claus Wittich. New York: Bedminster.

JACK A. GOLDSTONE

**STATISTICAL GRAPHICS**

Statistical graphs present data and the results of statistical analysis, assist in the analysis of data, and occasionally are used to facilitate statistical computation. Presentation graphs include the familiar bar graph, pie chart, line graph, scatterplot, and statistical map. Data analysis employs these graphical forms as well as others. Computational graphs (“nomographs”) sometimes display data but usually show theoretical quantities such as power curves for determining sample size. Computational graphs are convenient when statistical tables would be unwieldy, but computer programs are even more convenient, and so nomographs are used with decreasing frequency. This article emphasizes the role of graphs in data analysis, although many of the considerations raised here also apply to graphical presentation.

Although it generally is recognized that the pictorial representation of information is a par-

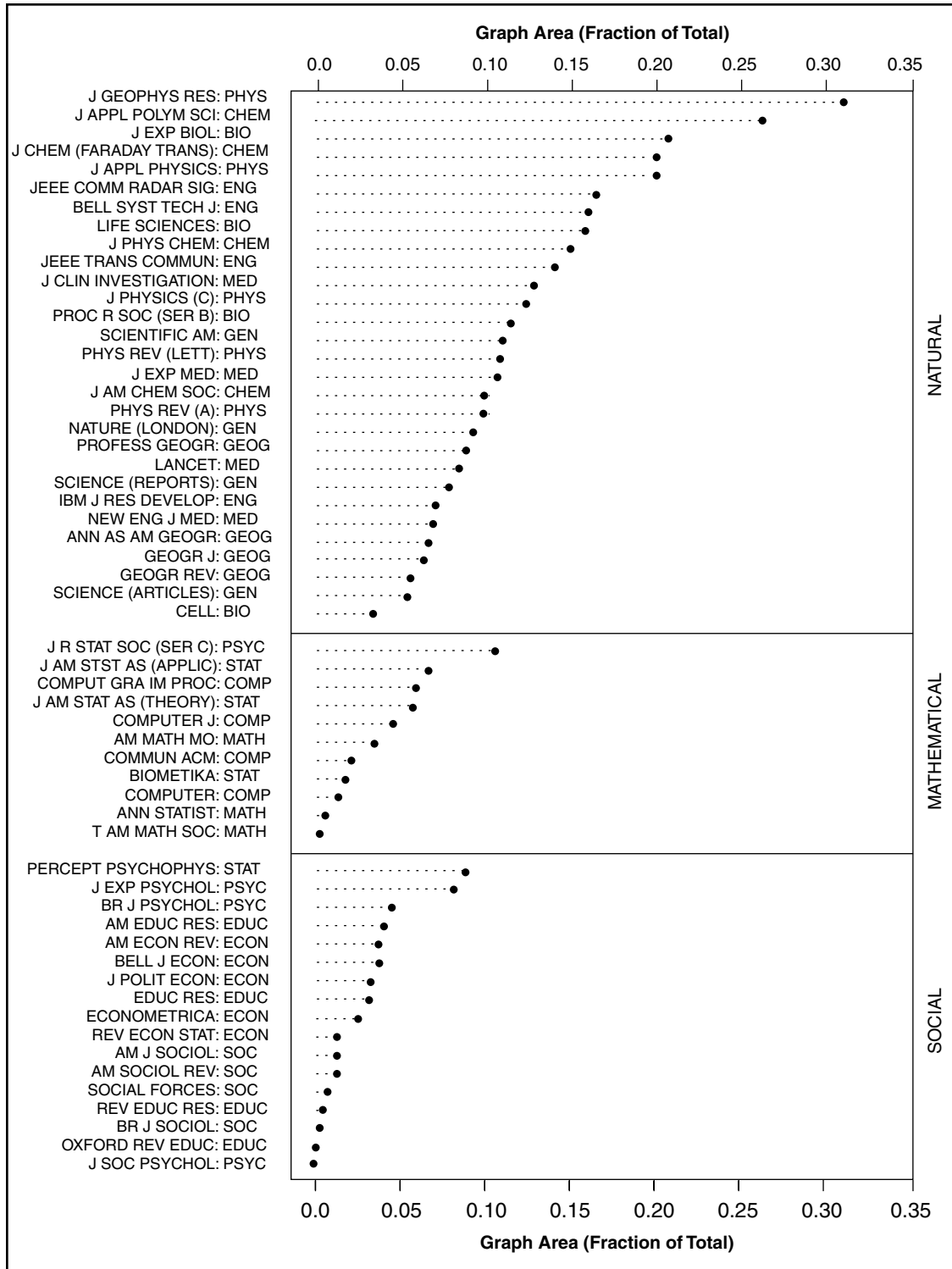


Figure 1. Dot graph showing the fractional area devoted to graphs in fifty-seven journals in the natural, mathematical, and social sciences. Four sociology journals appear near the bottom of the graph. To construct the graph, fifty articles were sampled from each journal in 1980 and 1981.

SOURCE: Reprinted from Cleveland (1984) with the permission of the American Statistical Association.

ticularly effective mode of communication, statistical graphs seldom appear in sociological publications. Figure 1, from Cleveland (1984), shows the relative space devoted to graphs in leading scientific publications, including four sociology journals. Sociology, of course, is not a wholly quantitative discipline. Nevertheless, even a cursory examination of publications in the field reveals that sociologists much more frequently report numerical information in tabular than in graphical form. Informal observation also suggests that sociologists usually analyze numerical data without the assistance of statistical graphs, a situation that may be changing.

### HISTORY

Broadly construed, graphic communication dates to the cave paintings of human prehistory and to the earliest forms of writing, which were pictorial or semipictorial. The first diagrams to communicate quantitative information—about location and distance—were maps: Egyptian cartographers employed coordinate systems in maps prepared 5,000 years ago, and cartography remains a relatively well developed area of graphical representation. Musical notation, which charts pitch as a function of time, also has an ancient origin and illustrates the spatial display of essentially nonspatial information. Rectilinear coordinate graphs are so familiar that it is easy to lose sight of the radical abstraction required to represent diverse quantities, such as pitch, as distances along an axis.

In the seventeenth century, the French mathematician and philosopher René Descartes established the relationship between algebraic equations and curves in a rectilinear coordinate space. The graphical representation of functions is not logically necessary for the display of empirical data as points in space, and there are isolated examples before Descartes of statistical graphs that employ abstract coordinate systems. Nevertheless, Descartes's analytic geometry no doubt provided the impetus for the development of statistical graphics, and the most common forms of statistical graphs evolved slowly over the subsequent three and a half centuries.

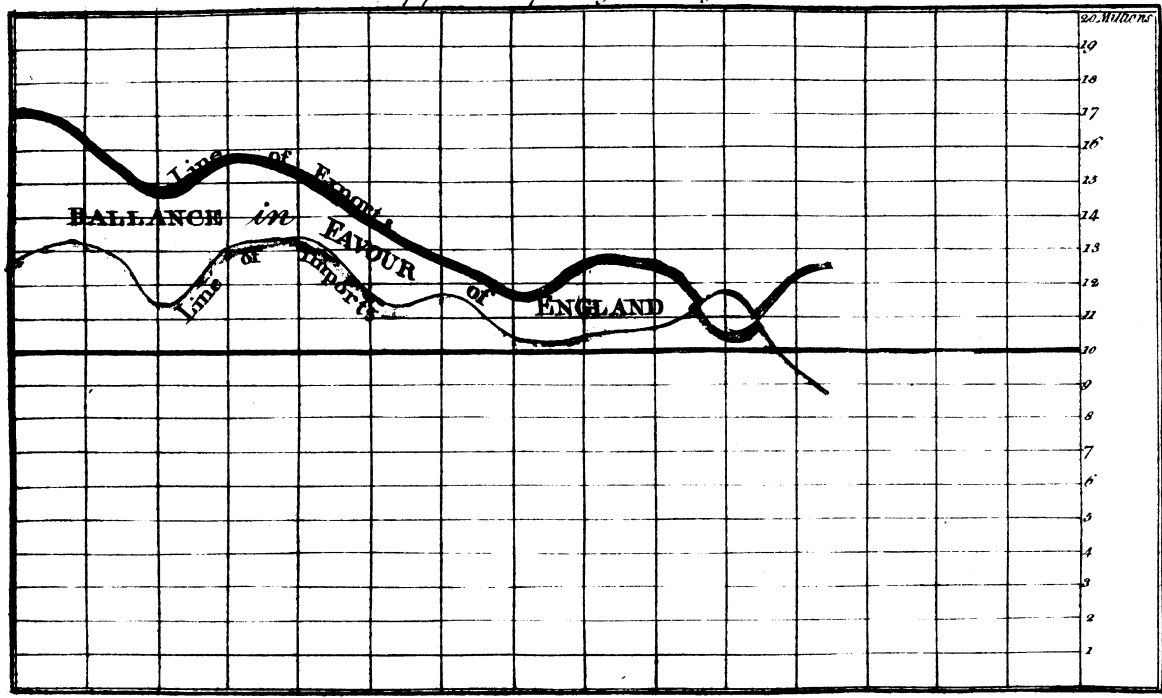
Among many individuals' contributions to this evolution, the work of William Playfair at the turn of the nineteenth century is of particular importance. First, Playfair either invented or popularized several common graphical forms, including the line graph, the bar graph, the pie chart, and the circle chart (in which areas of circles represent quantities). Second, Playfair employed statistical graphs to display social and economic data. Figure 2a, from Playfair's 1786 *Commercial and Political Atlas*, is a time series line graph of imports to and exports from England in the period 1771–1782. In the original graph, the space between the two curves is colored green when the balance of trade favors England (i.e., when the curve for exports is above that for imports) and red when the balance favors England's trading partners. Of the forty-two graphs in Playfair's atlas, all but one depict time series. The sole exception is a bar graph of imports to and exports from Scotland (Figure 2b), the data for which were available only for the year 1780–1781, precluding the construction of time series plots. Playfair's 1801 *Statistical Breviary* included a wider variety of graphical forms.

The first half of the nineteenth century was a period of innovation in and dissemination of statistical graphics, particularly in England and France. The ogive (cumulative frequency curve), the histogram, the contour map, and graphs employing logarithmic and polar coordinates all appeared before 1850. Later in the century, the British scientist Sir Francis Galton exploited an analogy to contour maps in his determination of the bivariate-normal correlation surface, illustrating the role of graphs in discovery.

The nineteenth-century enthusiasm for graphic representation of data produced many memorable and high-quality statistical graphs, such as those of Playfair, Florence Nightingale, E. J. Marey, and Charles Joseph Minard (several of which are reproduced in Tufte 1983). The same enthusiasm produced early abuses, however, including the graph from M. G. Mulhall's 1892 *Dictionary of Statistics* shown in Figure 3: The heights of the triangles indicate the accumulated wealth of each country, but their areas are wildly disproportionate to the quantities represented, conveying a misleading



*Chart of Imports & Exports to & from England for 12 Years*  
 From 1771 to 1782 by W. Playfair



*The Divisions at the Bottom express Years & those on the Right hand MILLIONS of POUNDS*  
 Published as the Act directs 20<sup>th</sup> Aug<sup>r</sup> 1783

Figure 2a

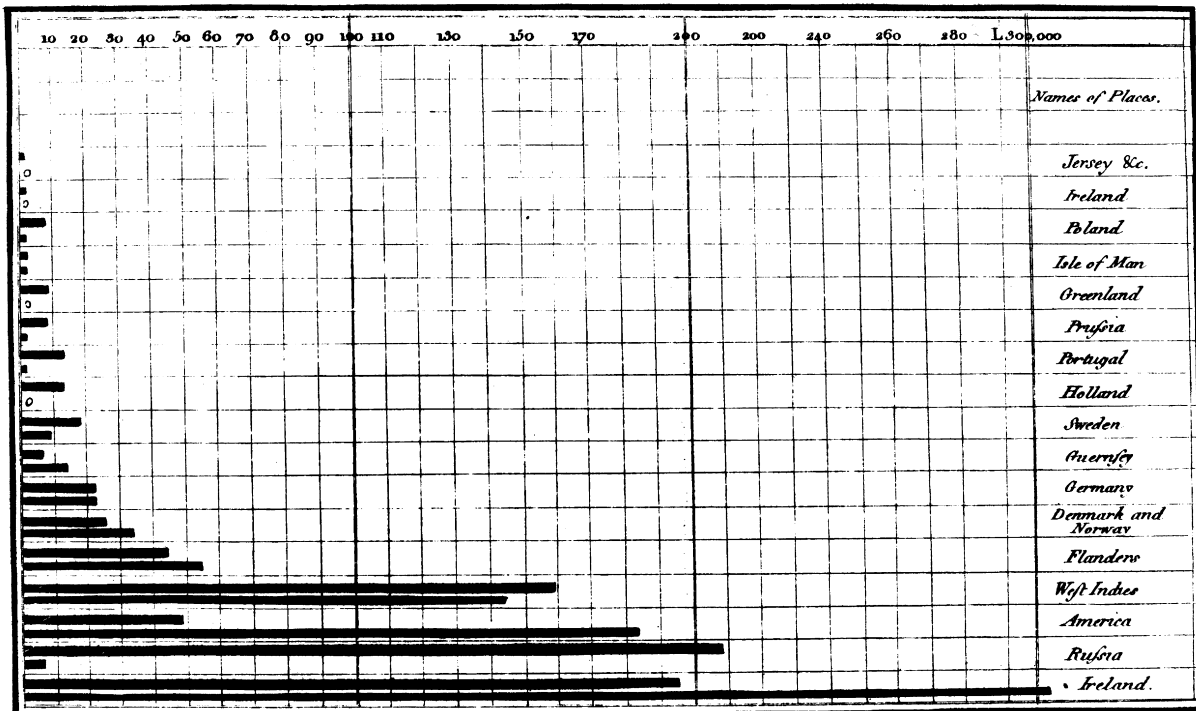
impression of the data. Furthermore, the horizontal arrangement of the countries bears no relationship to the purpose of the graph and apparently was done for artistic effect: It would be more natural to order the countries by wealth. Many modern graphs have similar problems, a situation that has motivated a substantial literature of graphic criticism (such as the works by Schmidt, Tufte, and Wainer discussed below).

The evolution of statistical graphics paralleled the general growth of statistical science well into the twentieth century. This relationship changed radically in the 1930s as statisticians such as R. A. Fisher emphasized the development of procedures for statistical inference. Fisher's influential *Statistical Methods for Research Workers*, first published in 1925, includes a brief chapter on "diagrams"; this chapter incorporates line graphs, scatterplots, and a histogram with a superimposed normal-density

curve. The remainder of the book, however, contains many numerical tables but just five additional figures, none of which presents empirical information. Fisher's 1935 *The Design of Experiments* includes just three graphs, all of which are theoretical.

The rebirth of interest in statistical graphics may be traced to John W. Tukey's work on exploratory data analysis, beginning in the 1960s and culminating in the publication of his text on this subject in 1977. Tukey's coworkers and students, most importantly the group at Bell Laboratories and its successors associated with William S. Cleveland, continue to contribute to the modern development of statistical graphics (see, in particular, Chambers et al. 1983; Cleveland 1993, 1994). Further information on the history of statistical graphics can be found in Funkhouser (1937), Tufte (1983), and Beninger and Robyn (1978), the last of which contains a useful chronology and bibliography.

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.



The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.

Published as the Act directs June 7<sup>th</sup> 1786 by W<sup>m</sup> Playfair

Next coup 1552 Strand, London.

Figure 2b. Two graphs from Playfair's 1786 Commercial and Political Atlas: (a) A time series line graph showing imports to and exports from England, 1771-1782. (b) A bar graph showing imports to and exports from Scotland for the year 1780-1781. The originals are in color.

SOURCE: Photographs courtesy of the William Clements Library, University of Michigan, Richard W. Ryan, curator of books.

### GRAPHIC STANDARDS

After several abortive efforts, the International Statistical Congresses held in Europe in the nineteenth century abandoned the attempt to formulate graphical standards. Since that time, many authors have proposed standards and principles for the construction of statistical graphs, but consensus on these matters remains elusive. Schmidt (1983, p. 17), for example, suggests that grid lines should always appear on rectilinear line graphs, while Tufte (1983, p. 112) maintains that grids "should usually be muted or completely suppressed," an instance of his more general principle that good graphs maximize the "dataink ratio" (the amount of ink devoted to the display of data as a proportion of all the ink used to draw the

graph) and eliminate "chartjunk" (extraneous graphical elements).

Disagreements such as this are due partly to the lack of systematic data on graphical perception (a situation that is improving), partly to differences in style and taste, and partly to the absence of adequate general theories of graph construction and perception (although there have been attempts, such as Bertin 1973). Also, good graphical display depends on the purposes for which a graph is drawn and on particular characteristics of the data, factors that are difficult to specify in advance and in a general manner.

Huff (1954, chap. 5), for example, argues that scales displaying ratio quantities should always start at zero to avoid exaggerating the magnitude

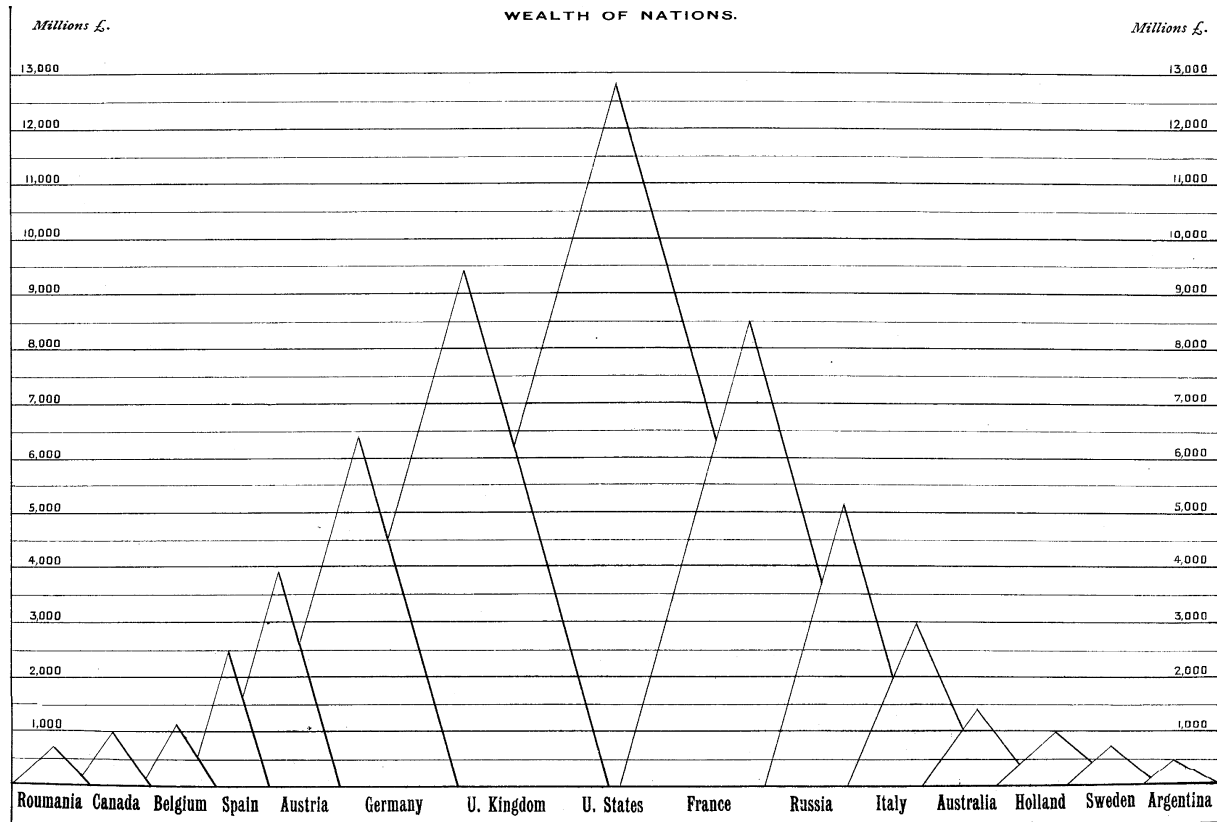


Figure 3. A modified bar graph from Mulhall's 1892 Dictionary of Statistics, substituting triangles with unequal bases for equal-width rectangular bars. The height of each triangle represents accumulated national wealth in 1888. The original is in color.

SOURCE: Photograph by University of Michigan Photographic Services.

of differences between data values. This principle, however, often disguises patterns in data that are revealed clearly by graphical magnification. Consider Figure 4, a and b, which shows the relative value of the Canadian and U.S. dollars in the eight weeks surrounding the June 23, 1990, deadline for the ratification of the ill-fated "Meech Lake" amendment to the Canadian constitution. This period was widely interpreted, both domestically and abroad, as one of constitutional crisis and uncertainty for Canada. Because in the short term the Canadian dollar traditionally trades in a narrow range against the U.S. dollar, Figure 4a is essentially uninformative, while Figure 4b reveals that the Canadian dollar fell slightly as the Meech deadline approached and rose afterward.

Despite some areas of disagreement, commentators on the design of statistical graphs, such

as Tufte (1983, 1990, 1997), Schmidt, and Wainer, offer a great deal of uncontroversially sound advice. In a tongue-in-cheek essay (reprinted in Wainer 1997: chap. 1), Wainer enumerates twelve rules to help the reader "display data badly." Several of these rules are illustrated in Figure 5a, which appeared in the *Miami Herald* in 1984: "Rule 7, Emphasize the trivial (ignore the important)"; "Rule 11, More is murkier: (a) more decimal places and (b) more dimensions"; and "Rule 12, If it has been done well in the past, think of a new way to do it." The graph in Figure 5a is meant to show the presumably negative relationship between the success of the twenty-six major league baseball teams in the 1984 season and the average salaries paid to the players on those teams. The lengths of the bars represent average players' salaries, while the teams' records of wins and losses are hidden in parenthe-

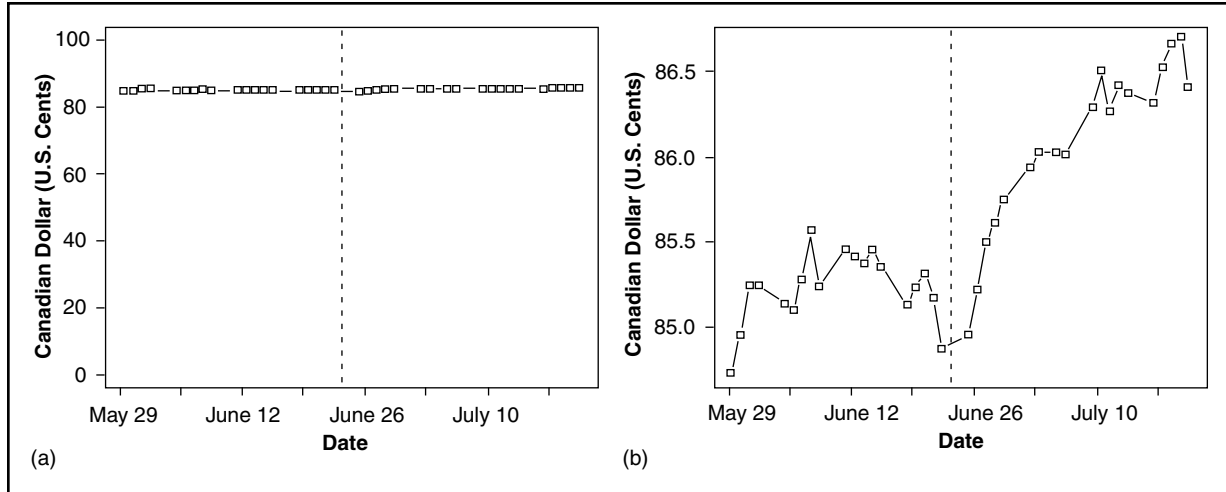


Figure 4. The relative value of the Canadian and U.S. dollar in an eight-week period in 1990 surrounding the failure of the Meech Lake amendment to the Canadian constitution. (a) Beginning the vertical axis at zero. Note that the upper end point of one is arbitrary, since the Canadian dollar can (at least in theory) trade above par with the U.S. dollar. (b) Scaling the vertical axis to accommodate the range of the data. The vertical line in each graph is drawn at the June 23 deadline for ratifying the Meech Lake accord.

SOURCE: Daily foreign exchange quotations in the *New York Times*.

ses within the bars, making it essentially impossible to tell whether the two variables are related—ostensibly the point of the graph. The bars are drawn in three-dimensional perspective, apparently for artistic effect, but the result is that the quantities represented are slightly distorted: For example, the average salary of the New York Yankees, \$458,544, appears to be about \$410,000. A standard representation of these data appears in the scatterplot in Figure 5b, revealing a slight *positive* relationship between salary and success.

**RESEARCH ON GRAPHIC PERCEPTION**

The earliest psychophysical research on perception of graphs, conducted in the 1920s, focused on the relative merits of pie charts and bar charts for displaying percentage data and was inconclusive. More recently, statisticians and psychologists have undertaken systematic experimentation on graphical perception. Spence and Lewandowsky (1990) review the literature in this area up to 1990.

Cleveland and McGill (1984), for example, conducted a series of experiments to ascertain the

relative accuracy of ten elementary perceptual tasks that extract quantitative information from graphs, as represented schematically in Figure 6. Ranked in order of decreasing average accuracy, these tasks involve judgment of position along a common scale; position along nonaligned scales; length, direction, or angle; area; volume or curvature; and shading or color saturation. Similarly, Spence (reported in Spence and Lewandowsky 1990) has shown in an experiment that categorical information differentiating points on a scatterplot is encoded most effectively by colors and least effectively by confusable letters (e.g., E, F, H); other coding devices, such as different shapes (circles, squares, triangles), degrees of fill, and discriminable letters (H, Q, X), were intermediate in effectiveness.

Cleveland (1993) demonstrates that slope judgments are most accurate for angles close to forty-five degrees and least accurate for angles near zero or ninety degrees. Cleveland therefore suggests that the aspect ratio of graphs (the relative lengths of the axes) be set so that average slopes are close to forty-five degrees, a procedure he terms “banking to forty-five degrees.” This process is illus-

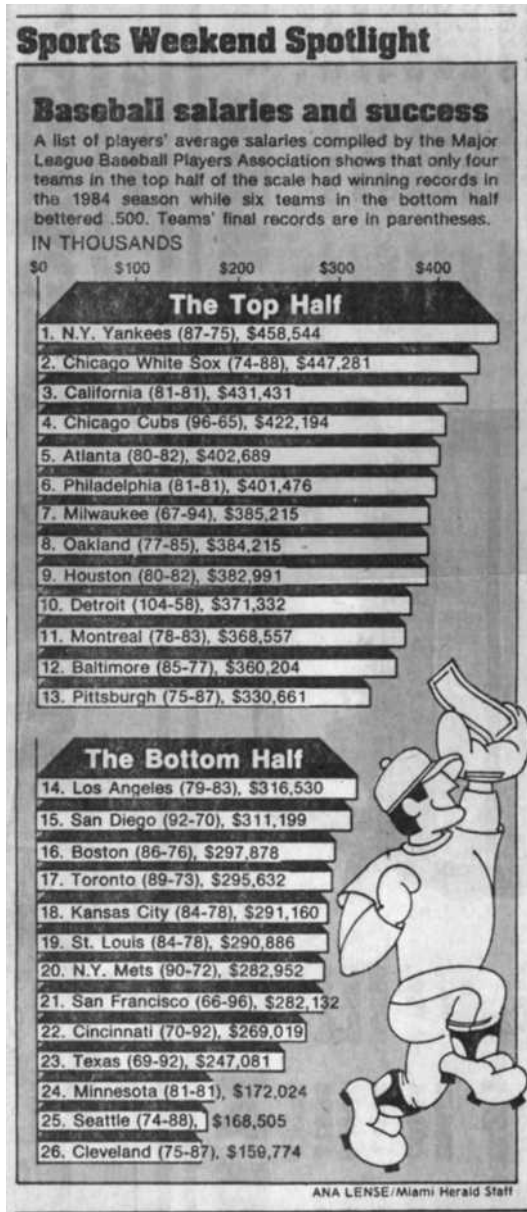


Figure 5a

trated in Figure 7. Both graphs in this figure plot the same data, but the periodic pattern of the data is nearly impossible to discern in Figure 7a because the average slope of the curve is too steep.

Cleveland and his colleagues have designed new graphical forms that apply these and similar findings by encoding important information through the employment of accurately judged graphic elements. One such form is the dot graph, an exam-

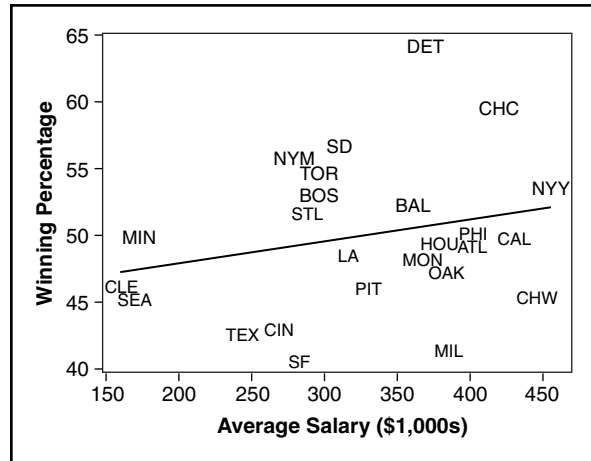


Figure 5b. Major League Baseball salaries and team success in the 1984 season. (a) As depicted in the Miami Herald. The lengths of the bars (slightly distorted) represent the average salaries paid to players from each team; the teams' won-lost records appear in parentheses within the bars. The apparent point of the graph is that there is a negative relationship between salaries and success. (b) The same data in standard scatterplot. The line on the plot, derived from a logistic regression of wins on average salaries, indicates a weak positive relationship between salaries and success.

ple of which appears in Figure 1. Similarly, Cleveland and McGill (1984) suggest the replacement of quantitative statistical maps that use shading or hue (e.g., Figure 8a) with maps that employ framed rectangles (Figure 8b), which exploit the more accurate judgment of position along nonaligned scales. Despite the inferiority of Figure 8a for judging differences in murder rates among the states, however, this map more clearly reveals regional variations in rates, illustrating the principle that the purpose for which a graph is drawn should influence its design.

The effectiveness of statistical graphs is rooted in the remarkable ability of people to apprehend, process, and remember pictorial information. The human visual system, however, is subject to distortion and illusion, processes that can affect the perception of graphs. Good graphical design can minimize and counteract the limitations of human vision. In Figure 9, for example, it appears that the difference between the hypothetical import and export series is changing when this difference

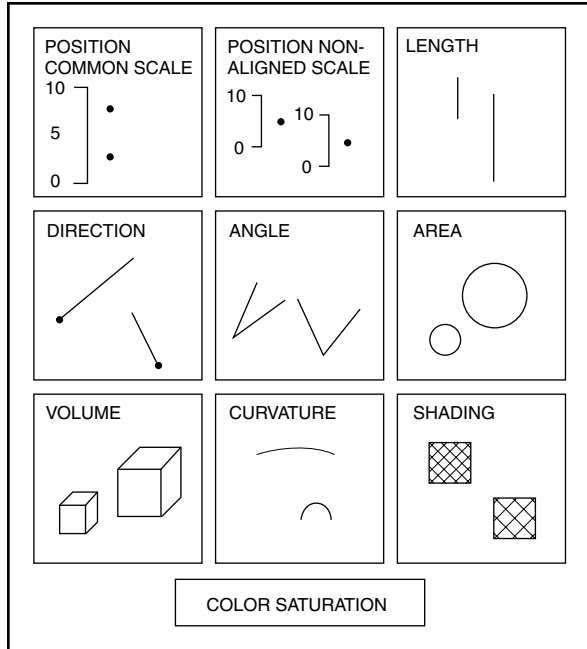


Figure 6. Ten elementary perceptual tasks for decoding quantitative information from statistical graphs.

SOURCE: Reprinted from Cleveland and McGill (1984) with the permission of the American Statistical Association.

actually is constant (cf., Playfair’s time series graph in Figure 2a). The source of the illusion is the tendency to attend to the least distance between the two curves rather than to the vertical distance. Thus, an alternative is to graph the difference between the two curves—the balance of trade—directly (cf. Figure 12, b and c, below), exploiting the relatively accurate judgment of position along a common scale, or to show vertical lines between the import and export curves, employing the somewhat less accurate judgment of position along nonaligned scales.

### GRAPHS IN DATA ANALYSIS

Statistical graphs should play a central role in the analysis of data, a common prescription that is most often honored in the breach. Graphs, unlike numerical summaries of data, facilitate the perception of general patterns and often reveal unusual, anomalous, or unexpected features of the data—characteristics that might compromise a numerical summary.

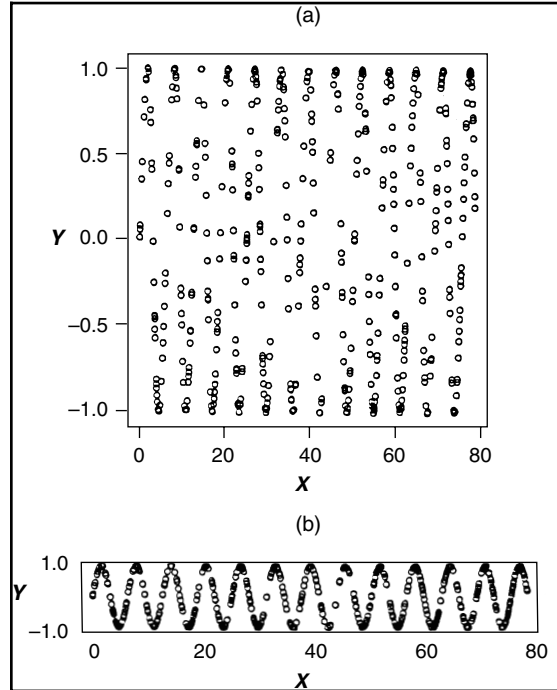


Figure 7. Two scatterplots of the same data. Five hundred X-values were randomly generated in the interval  $[0, 25\pi]$ , and  $Y = \sin X$ . The periodic pattern of the data is clear in (b), where the aspect ratio of the plot is adjusted so that the average slope of the curve is not too steep, but not in panel (a).

The four simple data sets in Figure 10, from Anscombe (1973) and dubbed “Anscombe’s quartet” by Tufte (1983), illustrate this point well. All four data sets yield the same linear least-squares outputs when regression lines are fitted to the data, including the regression intercept and slope, coefficient standard errors, the standard error of the regression (i.e., the standard deviation of the residuals), and the correlation, but—significantly—not residuals. Although the data are contrived, the four graphs tell very different imaginary stories: The least-squares regression line accurately summarizes the tendency of  $y$  to increase with  $x$  in Figure 10a. In contrast, the data in Figure 10b clearly indicate a curvilinear relationship between  $y$  and  $x$ , a relationship the linear regression does not capture. In Figure 10c, one point is out of line with the rest and distorts the regression. Perhaps the outlying point represents an error in recording the data or a  $y$ -value that is influenced by factors other than  $x$ . In Figure 10d, the ability to fit a line

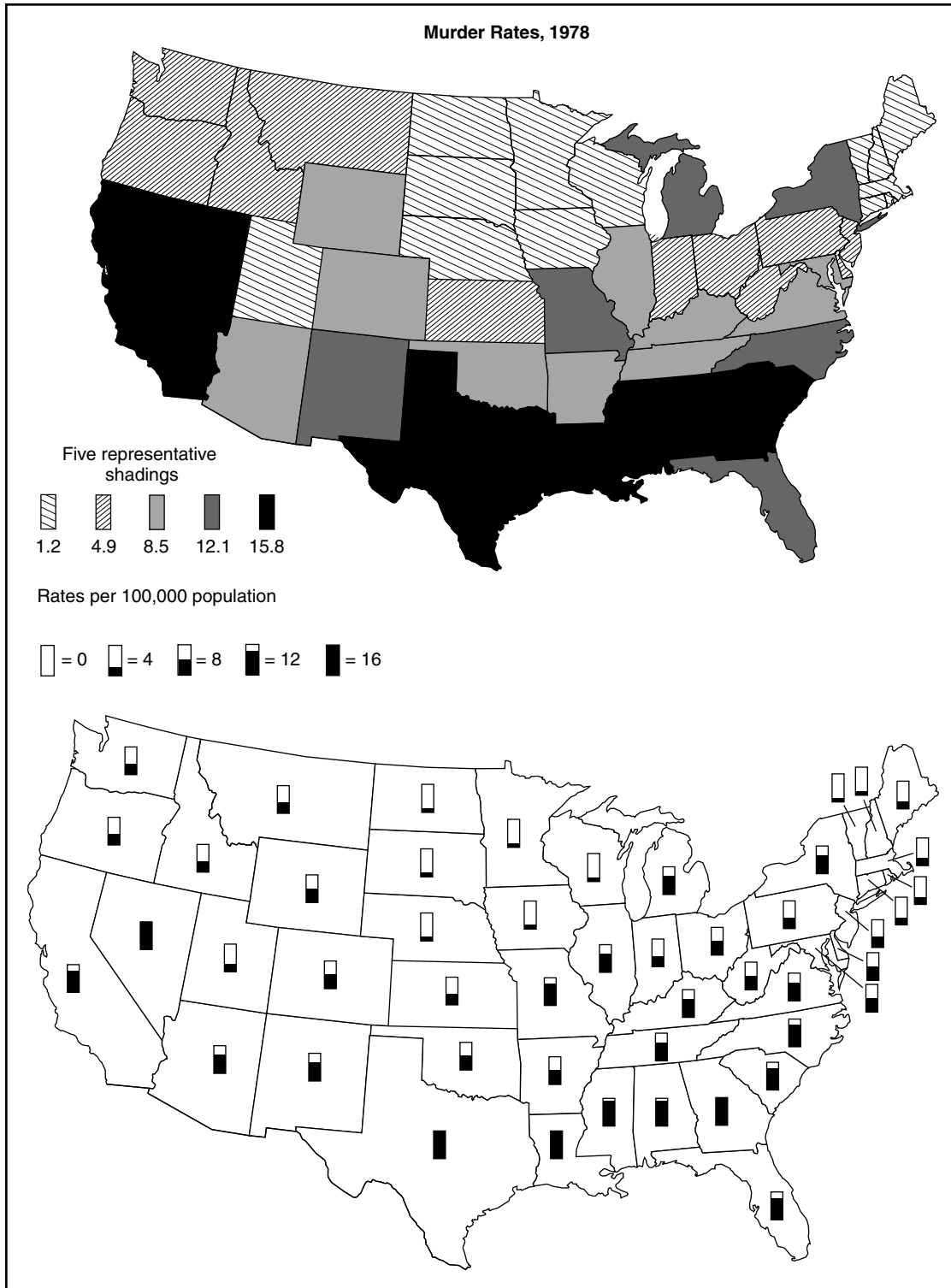


Figure 8. Statistical maps of state murder rates in 1978 employing (a) shading and (b) framed rectangles.

SOURCE: Reprinted from Cleveland and McGill (1984) with the permission of the American Statistical Association.

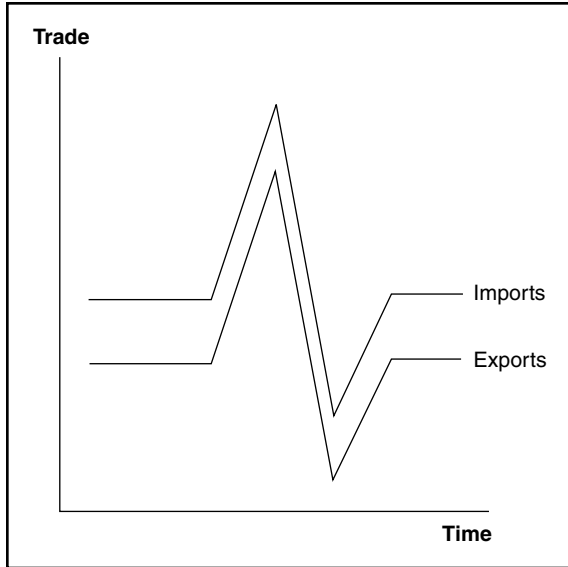


Figure 9. Despite appearances, the vertical separation between the curves for imports and exports is constant. The “data” are contrived.

and the line’s specific location depend on the presence of a single point.

Diverse graphical forms are adapted to different purposes in data analysis. Many important applications appear in the figures below, roughly in order of increasing complexity, including graphs for displaying univariate distributions, bivariate relationships, diagnostic quantities in regression analysis, and multivariate data.

Particularly useful for graphically screening data are methods for displaying the distributions of quantitative variables. Several univariate displays of the distribution of infant mortality rates for 201 countries are shown in Figure 11, using data compiled by the United Nations.

Figure 11a is a traditional histogram of the infant mortality data, a frequency bar graph formed by dissecting the range of infant mortality into class intervals or “bins” and then counting the number of observations in each bin; the vertical axis of the histogram is scaled in percent. Figure 11b shows an alternative histogram that differs from Figure 11a only in the origin of the bin system (the bars are shifted five units to the left). These graphs demonstrate that the impression conveyed by a histogram depends partly on the

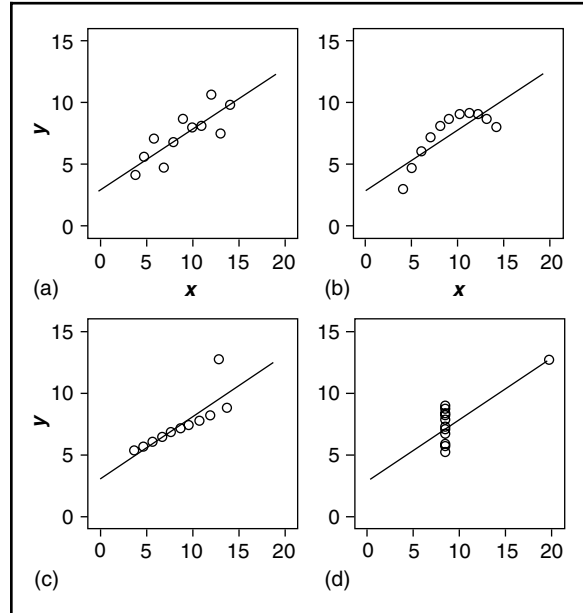


Figure 10. The four data sets have the same linear least-squares regression, including the regression coefficients, their standard errors, the correlation between the variables, and the standard error of the regression.

SOURCE: Redrawn from Anscombe (1973) with the permission of the American Statistical Association.

arbitrary location of the bins. Figure 11c is a stem-and-leaf display, a type of histogram (from Tukey) that records the data values directly in the bars of the graph, thus permitting the recovery of the original data. Here, for example, the values given as 1:2 represent infant mortality rates of 12 per 1,000.

Figure 11d is a kernel density estimate, or smoothed histogram, a display that corrects both the roughness of the traditional histogram and its dependence on the arbitrary choice of bin location. For any value  $x$  of infant mortality, the height of the kernel estimate is

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

where  $n$  is the number of observations (here, 201); the observations themselves are  $\chi_1, \chi_2, \dots, \chi_n$ ,  $h$  is the “window” half-width for the kernel estimate, analogous to bin width for a histogram; and  $K$  is



STATISTICAL GRAPHICS

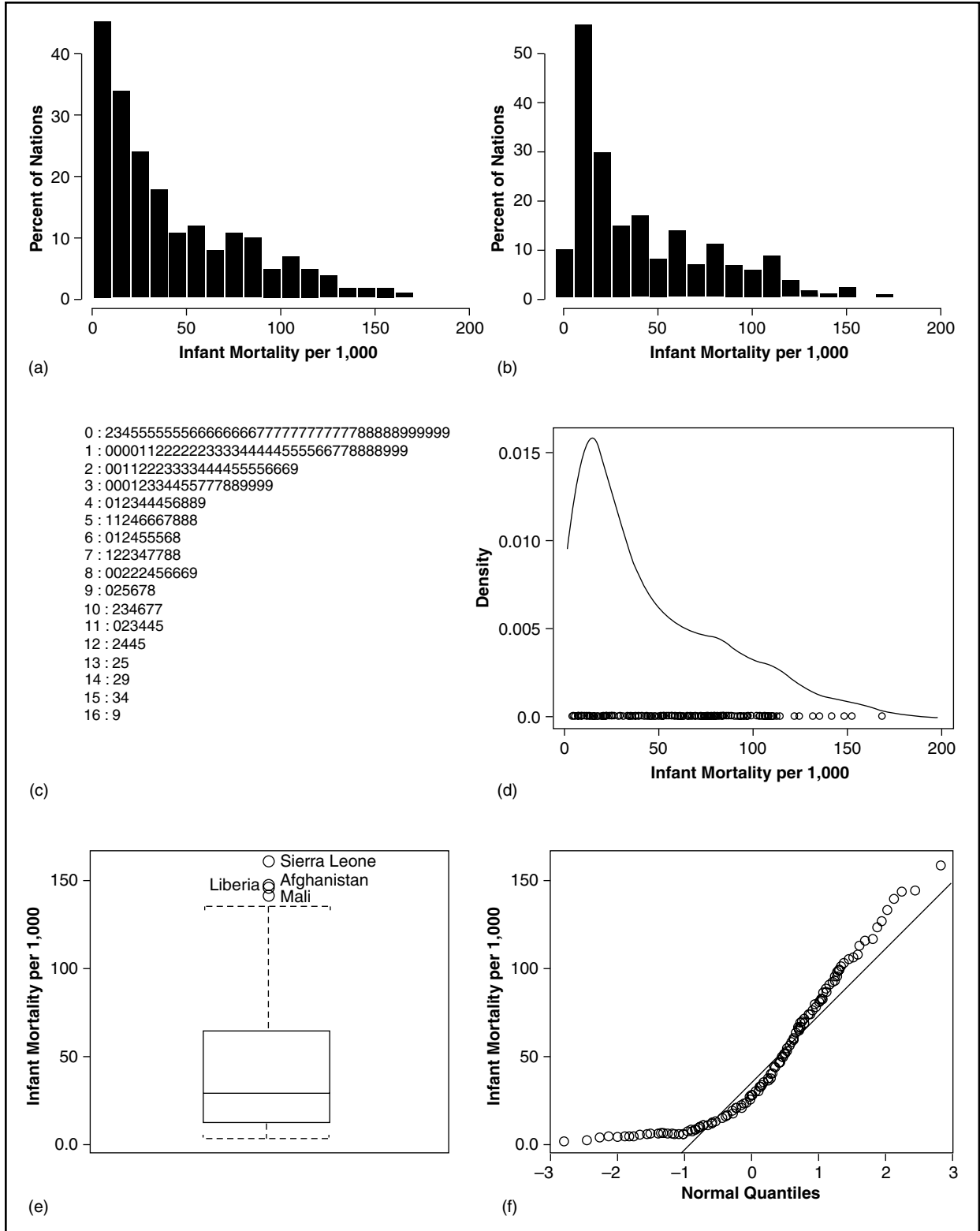


Figure 11. Six univariate displays of the distribution of infant mortality rates in 201 nations. The histograms (a) and (b) both have bins of width ten, but the bars of (b) are five units to the left of those of (a). A stem-and-leaf display is shown in (c), a kernel density estimate in (d), a boxplot in (e), and a normal quantile comparison plot in (f).

SOURCE: United Nations, <http://www.un.org/Depts/unsd/social/main.htm>.

some probability–density function, such as the unit-normal density, ensuring that the total area under the kernel estimate is one. A univariate scatterplot – another form of distributional display giving the location of each observation – is shown at the bottom of Figure 11d.

Figure 11e, a “boxplot” of the infant mortality data (a graphic form also from Tukey), summarizes a variety of important distributional information. The box is drawn between the first and third quartiles and therefore encloses the central half of the data. A line within the box marks the position of the median. The whiskers extend either to the most extreme data value (as on the bottom) or to the most extreme nonoutlying data value (as on the top). Four outlying data values are represented individually. The compactness of the boxplot suggests its use as a component of more complex displays; boxplots may be drawn in the margins of a scatterplot to show the distribution of each variable, for example.

Figure 11f shows a normal quantile comparison plot for the infant mortality data. As the name implies, this graph compares the ordered data with corresponding quantiles of the unit-normal distribution. By convention, the  $i$ th largest infant mortality rate, denoted  $\chi_{(i)}$ , has  $P_i = (i - 1/2)/n$  proportion of the data below it. The corresponding normal quantile is  $z_i$ , located so that  $Pr(Z \leq z_i) = P_i$ , where  $Z$  follows the unit-normal distribution. If  $X$  is normally distributed with mean  $\mu$  and standard deviation  $\sigma$ , then within the bounds of sampling error,  $x_{(i)} \cong \mu + \sigma z_i$ . Departure from a linear pattern therefore indicates nonnormality. The line shown in Figure 11f passes through the quartiles of  $X$  and  $Z$ . The positive skew of the infant mortality rates is reflected in the tendency of the plotted points to lie above the fitted line in both tails of the distribution.

While the skewness of the infant mortality data is apparent in all the displays, the possibly multimodal grouping of the data is clearest in the kernel density estimate. The normal quantile comparison plot, in contrast, retains the greatest resolution in the tails of the distribution, where data are sparse; these are the regions that often are problematic for numerical summaries of data such as means and regression surfaces.

Many useful graphs display relationships between variables, including several forms that ap-

peared earlier in this article: bar graphs (Figure 2b), dot graphs (Figure 1), and line graphs such as time series plots (Figures 2a and 4). Parallel boxplots are often informative in comparing the distribution of a quantitative variable across several categories. Scatterplots (as in Figure 10) are invaluable for examining the relationship between two quantitative variables. Other data-analytic graphs adapt these forms.

In graphing quantitative data, it is sometimes advantageous to transform variables. Logarithms, the most common form of transformation, often clarify data that extend over two or more orders of magnitude (i.e., a factor of 100 or more) and are natural for problems in which ratios of data values, rather than their differences, are of central interest.

Consider Figure 12, which shows the size of the Canadian and U.S. populations for census years between 1790 and 1990 in the United States and between 1851 and 1991 in Canada. The data are graphed on the original scale in Figure 12a and on the log scale in Figure 12b. Because the Canadian population is much smaller than that of the United States, it is difficult to discern the Canadian data in Figure 12a. Moreover, Figure 12b shows more clearly departures from a constant rate of population growth, represented by linear increase on the log scale, and permits a direct comparison of the growth rates in the two countries. These rates were quite similar, with the U.S. population roughly ten times as large as the Canadian population throughout the past century and a half. Figure 10c, however, which graphs the difference between the two curves in Figure 10b (i.e., the log population ratio), reveals that the United States was growing more rapidly than Canada was before 1900 and more slowly afterward.

Graphs also can assist in statistical modeling. Least-squares regression analysis, for example, which fits the model

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i \quad (2)$$

makes strong assumptions about the structure of the data, including assumptions of linearity, equal error variance, normality of errors, and independence. Here  $Y_i$  is the dependent variable score for the  $i$ th of  $n$  observations;  $\chi_{1i}, \chi_{2i}, \dots, \chi_{ki}$  are independent variables;  $\epsilon_i$  is an unobserved error that is assumed to be normally distributed with zero ex-

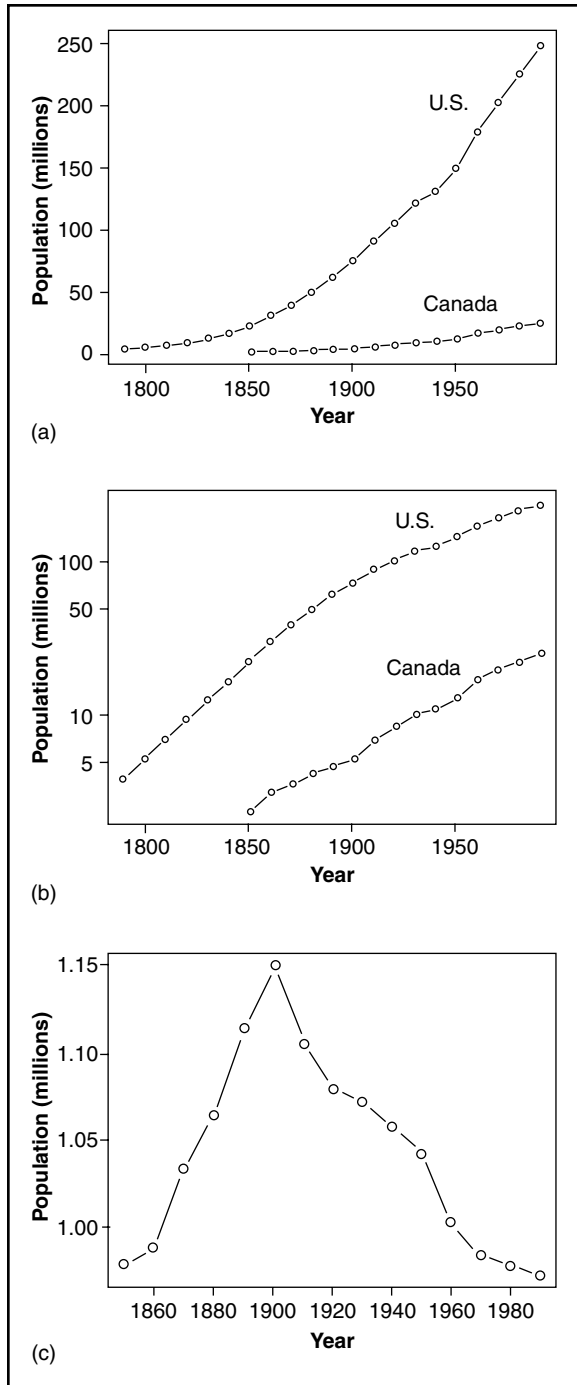


Figure 12. Canadian and U.S. population figures are plotted directly in (a) and on a log scale in (b). The difference between the two log series is shown in (c).

SOURCE: *Canada Yearbook 1994* and *Statistical Abstract of the United States: 1994*.

pectation and constant variance  $\sigma^2$ , independent of the  $x$ 's and the other errors; and the  $\beta$ 's are regression parameters, which are to be estimated along with the error variance from the data.

Graphs of quantities derived from the fitted regression model often prove crucial in determining the adequacy of the model. Figure 13, for example, plots a measure of leverage in the regression (the "hat values"  $h_i$ ) against a measure of discrepancy (the "studentized residuals"  $t_i$ ). Leverage represents the degree to which individual observations can affect the fitted regression, while discrepancy represents the degree to which each observation departs from the pattern suggested by the rest of the data. Actual influence on the estimated regression coefficients is a product of leverage and discrepancy and is displayed on the graph by Cook's  $D_i$ , represented by the areas of the plotted circles. The data for this graph are drawn from Duncan's (1961) regression of the rated prestige of forty-five occupations on the educational and income levels of the occupations. The plot suggests that two of the data points (the occupations "minister" and "conductor") may unduly affect the fitted regression.

Figure 14 is a scatterplot of residuals against fitted  $Y$ -values,

$$\hat{Y}_i = b_0 + b_1x_{1i} + b_2x_{2i} + \dots + b_kx_{ki} \quad (3)$$

where the  $b$ 's are sample estimates of the corresponding  $\beta$ 's. If the error variance is constant as assumed, the variation of the residuals should not change systematically with the fitted values. The data for Figure 14 are drawn from work by Ornstein (1976) relating the number of interlocking directorate and executive positions maintained by 248 dominant Canadian corporations to characteristics of the firms. The plot reveals that the variation of the residuals appears to increase with the level of the fitted values, casting doubt on the assumption of constant error variance.

Figure 15 shows a partial residual (also called a component plus residual) plot for the relationship between occupational prestige and income, a diagnostic useful for detecting nonlinearity in regression. The plot is for a regression of the rated prestige of 102 Canadian occupations on the gender composition, income level, and educational level of the occupations (see Fox and Suschnigg

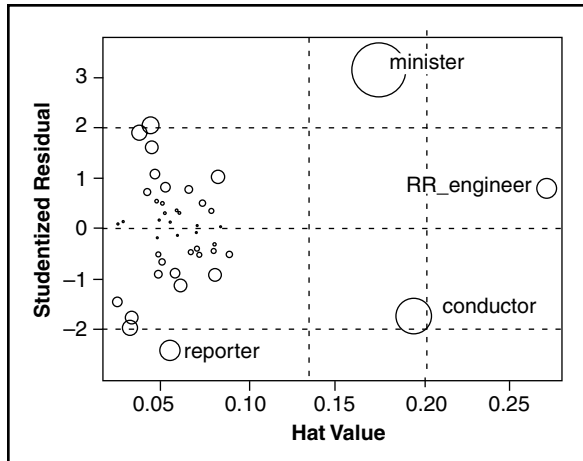


Figure 13. Influence plot for Duncan's regression of the rated prestige of forty-five occupations on their income and educational levels. The hat values measure the leverage of the observations in the regression, while the studentized residuals measure their discrepancy. The plotted circles have area proportional to Cook's  $D$ , a summary measure of influence on the regression coefficients. Horizontal lines are drawn at plus and minus 2; in well-behaved data, only about 5 percent of studentized residuals should be outside these lines. Vertical lines are drawn at two and three times the average hat value; hat values greater than two or three times the average are noteworthy. Observations that have relatively large residuals or leverages are identified on the plot.

1989). The partial residuals are formed as  $e_{1i} = b_1\chi_{1i} + e_i$ , where  $b_1$  is the fitted income coefficient in the linear regression,  $\chi_{1i}$  is the average income of incumbents of occupation  $i$ , and  $e_i$  is the regression residual. The nonlinear pattern of the data, which is apparent in the graph, suggests modification of the regression model. Similar displays are available for generalized linear models such as logistic regression. Further information on the role of graphics in regression diagnostics can be found in Atkinson (1985), Fox (1991, 1997), and Cook and Weisberg (1994).

Scatterplots are sometimes difficult to interpret because of visual noise, uneven distribution of the data, or discreteness of the data values. Visually ambiguous plots often can be enhanced by smoothing the relationship between the variables, as in Figure 15. The curve drawn through this plot was determined by a procedure from

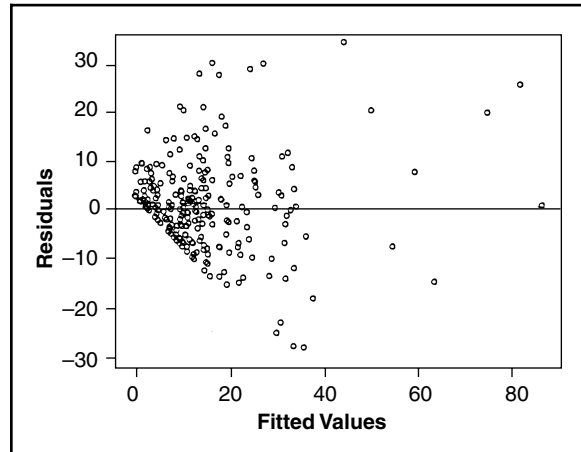


Figure 14. Plot of residuals by fitted values for Ornstein's regression on interlocks maintained by 248 dominant Canadian corporations on the characteristics of the firms. The manner in which the points line up diagonally at the lower left of the graph is due to the lower limit of zero for the dependent variable.

SOURCE: Personal communication from M. Ornstein.

Cleveland (1994) called *locally weighted scatterplot smoothing* ("lowess"). Lowess (also called "loess," for local regression) fits  $n$  robust regression lines to the data, with the  $i$ th such line emphasizing observations whose  $\chi$ -values are closest to  $\chi_i$ . The lowess fitted value for the  $i$ th observation,  $\hat{y}_i$ , comes from the  $i$ th such regression. Here  $x$  and  $y$  simply denote the horizontal and vertical variables in the plot. The curve plotted on Figure 15 connects the points  $(\chi_i, \hat{y}_i)$ . Lowess is one of many methods of nonparametric regression analysis, including methods for multiple regression, described, for example, in Hastie and Tibshirani (1990) and Fox (forthcoming *a* and *b*). Because there is no explicit equation for a nonparametric regression, the results are most naturally displayed graphically.

Scatterplots for discrete data may be enhanced by paradoxically adding a small amount of random noise to the data to separate the points in the plot. Cleveland (1994) calls this process "jittering." An example is shown in Figure 16a, which plots scores on a vocabulary test against years of education; the corresponding jittered plot (Figure 16b) reduces the overplotting of points, making the relationship much clearer and revealing other characteristics

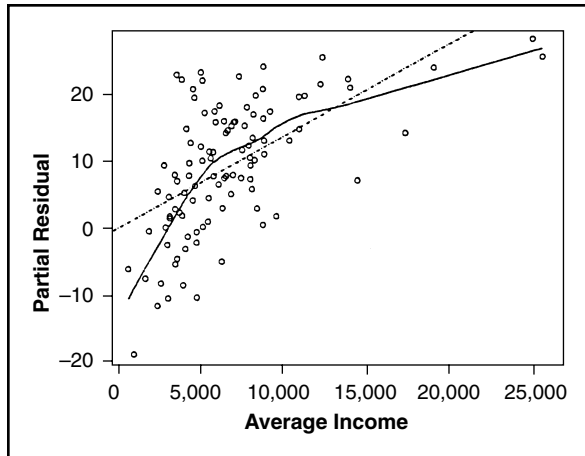


Figure 15. Partial residual (component+residual) plot for income in the regression of occupational prestige on the gender composition and income and education levels of 102 Canadian occupations in 1971. The broken line gives the linear least-squares fit, while the solid line shows the lowess (nonparametric regression) fit to the data.

SOURCE: Fox and Suschnigg (1989).

of the data, such as the concentration of points at twelve years of education.

Because graphs commonly are drawn on two-dimensional media such as paper and computer screens, the display of multivariate data is intrinsically more difficult than that of univariate or bivariate data. One solution to the problems posed by multivariate graphic representation is to record additional information on a two-dimensional plot. Symbols such as letters, shapes, degrees of fill, and color may be used to encode categorical information on a scatterplot, for example (see Figure 19, below). Similarly, there are many schemes for representing additional quantitative information, as shown in Figures 8 and 13.

A scatterplot matrix is the direct graphic analogue of a correlation matrix, displaying the bivariate relationship between each pair of a set of quantitative variables and thus providing a quick overview of the data. In contrast to a correlation matrix, however, a scatterplot matrix can reveal nonlinear relationships, outlying data, and so on. The scatterplot matrix in Figure 17 is for rates of seven different categories of crime in the thirty largest U.S. cities (excluding Chicago) in 1996.

The regression curve shown in each scatterplot was determined by the lowess procedure described above.

A limitation of the scatterplot matrix is that it displays only the *marginal* relationships between the variables, while *conditional* (or *partial*) relationships are more often the focus of multivariate statistical analysis. This limitation sometimes can be overcome, however, by highlighting individual observations or groups of observations and following them across the several plots (see the discussion of “brushing” in Cleveland 1994). These methods are most effective when they are implemented as part of an interactive computer system for graphic data analysis.

One approach to displaying conditional relationships is to focus on the relationship between the dependent variable and each independent variable fixing the other independent variable (or variables) to particular, possibly overlapping ranges of values. A nonparametric regression smooth then can be fitted to each partial scatterplot. Cleveland (1993) calls this kind of display a “conditioning plot” or “coplot.” The strategy breaks down, however, when there are more than two or three independent variables, or when the number of observations is small.

Many of the most useful graphical techniques for multivariate data rely on two-dimensional projections of the multivariate scatterplot of the data. A statistical model fitted to the data often determines these projections. An example of a display employing projection of higher-dimensional data is the partial residual plot shown in Figure 15. Another common application of this principle is the similarly named but distinct partial regression (or added-variable) plot. Here the dependent variable ( $Y$ ) and one independent variable in the multiple regression model (say,  $x_1$ ) are each regressed on the other independent variables in the model (i.e.,  $\chi_2, \dots, \chi_k$ ), producing two sets of residuals (which may be denoted  $y_{(1)}$  and  $\chi_{(1)}$ ). A scatterplot of the residuals (that is,  $y_{(1)}$  versus  $\chi_{(1)}$ ) is frequently useful in revealing high-leverage and influential observations. Implementation on modern desktop computers, which can exploit color, shading, perspective, motion, and interactivity, permits the effective extension of projections to three dimensions (see Monette 1990; Cook and Weisberg 1994; Cook 1998).

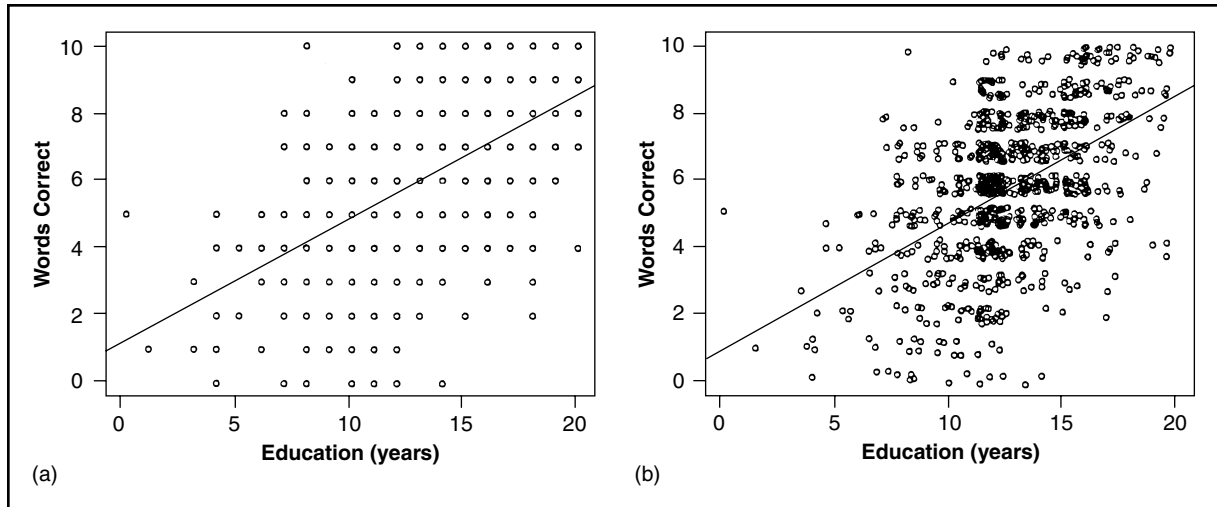


Figure 16. Randomly “jittering” a scatterplot to clarify discrete data. The original plot in (a) shows the relationship between score on a ten-item vocabulary test and years of education. The same data are graphed in (b) with a small random quantity added to each horizontal and vertical coordinate. Both graphs show the least-squares regression line.

SOURCE: 1989 General Social Survey, National Opinion Research Center.

When there are relatively few observations and each is of separate interest, it is possible to display multivariate data by constructing parallel geometric figures for the individual observations. Some feature of the figure encodes the value of each variable. One such display, called a “star plot,” is shown in Figure 18 for the U.S. cities crime rate data. The cities are arranged in order of increasing general crime rate.

Other common and essentially similar schemes include “trees” (the branches of which represent the variables), faces (whose features encode the variables), and small bar graphs (in which each bar displays a variable). None of these graphs is particularly easy to read, but judicious ordering of observations and encoding of variables sometimes can suggest natural clusterings of the data or similarities between observations. Note in Figure 18, for example, that Oklahoma City and Jacksonville have roughly similar “patterns” of crime, even though the rates for Oklahoma City are generally higher. If similarities among the observations are of central interest, however, it may be better to address the issue directly by means of clustering or ordination (also called multidimensional scaling); see, e.g., Hartigan (1975), and Kruskal and Wish (1978).

### THE PRESENT AND FUTURE OF STATISTICAL GRAPHICS

Computers have revolutionized the practice of statistical graphics much as they earlier revolutionized numerical statistics. Computers relieve the data analyst of the tedium of drawing graphs by hand and make possible displays—such as lowess scatterplot smoothing, kernel density estimation, and dynamic graphs—that previously were impractical or impossible. All the graphs in this article, with the exception of several from other sources, were prepared with widely available statistical software (most with *S-Plus*, the graphical and other capabilities of which are ably described by Venables and Ripley 1997). Virtually all general statistical computer packages provide facilities for drawing standard statistical graphs, and many provide specialized forms as well.

Dynamic and interactive statistical graphics, only a decade ago the province of high-performance graphics workstations and specialized software, are now available on inexpensive desktop computers. Figure 19 illustrates the application of Cook and Weisberg’s (1999) state-of-the-art *Arc* package to Duncan’s occupational prestige data.

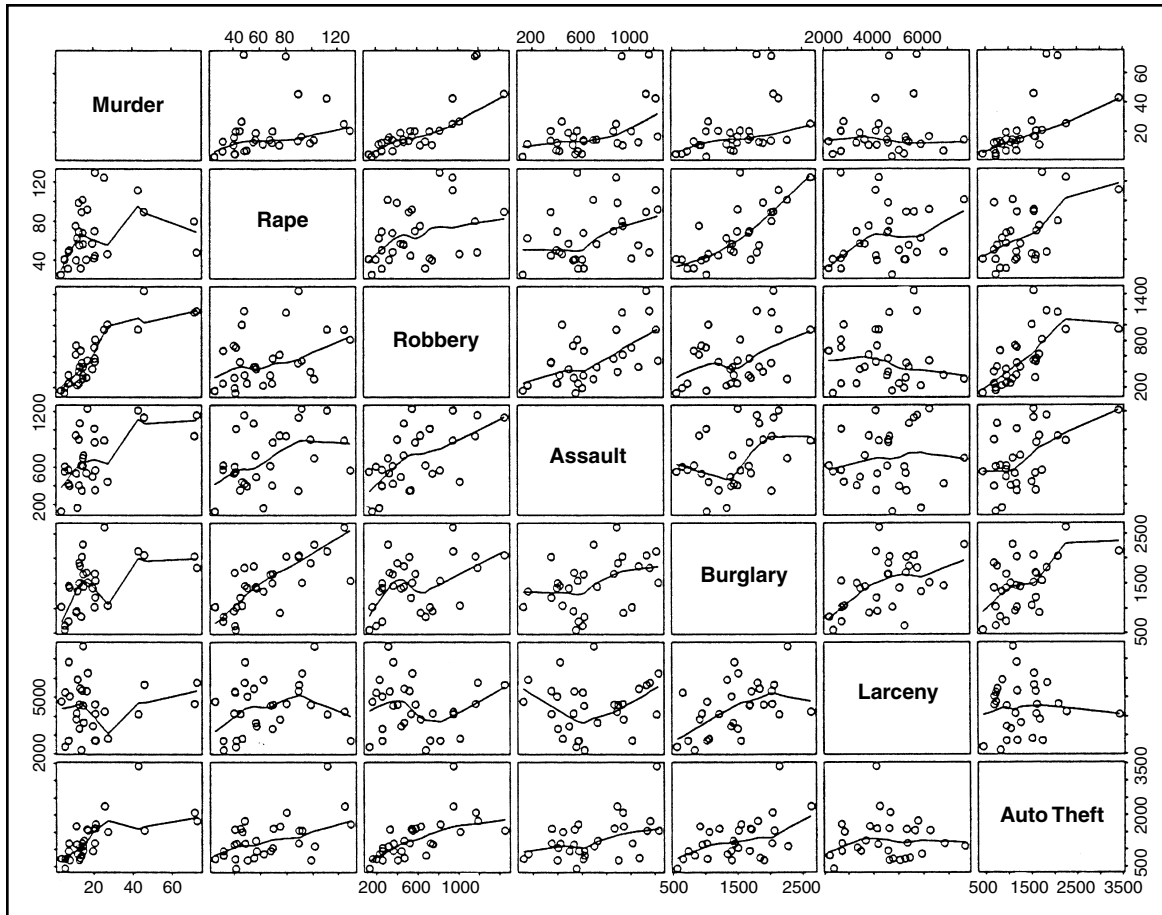


Figure 17. Scatterplot matrix for the rates of seven categories of crime in the thirty largest U.S. cities in 1996 (Chicago is omitted because of missing data). The rate labeled “Murder” represents both murder and manslaughter. The line shown in each panel is a lowess scatterplot smooth.

SOURCE: *Statistical Abstract of the United States: 1998*.

*Arc*, programmed in Tierney’s (1990) *Lisp-Stat* statistical computing environment, is freely available software that runs on Windows computers, Macintoshes, and Unix workstations. Standard statistical packages such as *SAS* and *SPSS* are gradually acquiring these capabilities as well.

The other edge of the computing sword cuts in the direction of ugly, poorly constructed graphs that obfuscate rather than clarify data: Modern software facilitates the production of competent (if not beautiful) statistical graphs. Nevertheless, a data analyst armed with a “presentation graphics” package can, with little effort or thought and less taste, produce elaborate, difficult to read, and misleading graphs.

REFERENCES

Anscombe, Frank J. 1973 “Graphs in Statistical Analysis.” *American Statistician* 27:17–22.

Atkinson, A. C. 1985 *Plots, Transformations, and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Oxford, UK: Clarendon Press.

Beninger, James R., and Dorothy L. Robyn 1978 “Quantitative Graphics in Statistics: A Brief History.” *American Statistician* 32:1–11.

Bertin, Jacques 1973 *Semiologie graphique*, 2nd ed. Paris: Mouton.

Chambers, J. M., William S. Cleveland, Beat Kleiner, and Paul A. Tukey 1983 *Graphical Methods for Data Analysis*. Belmont Calif.: Wadsworth.

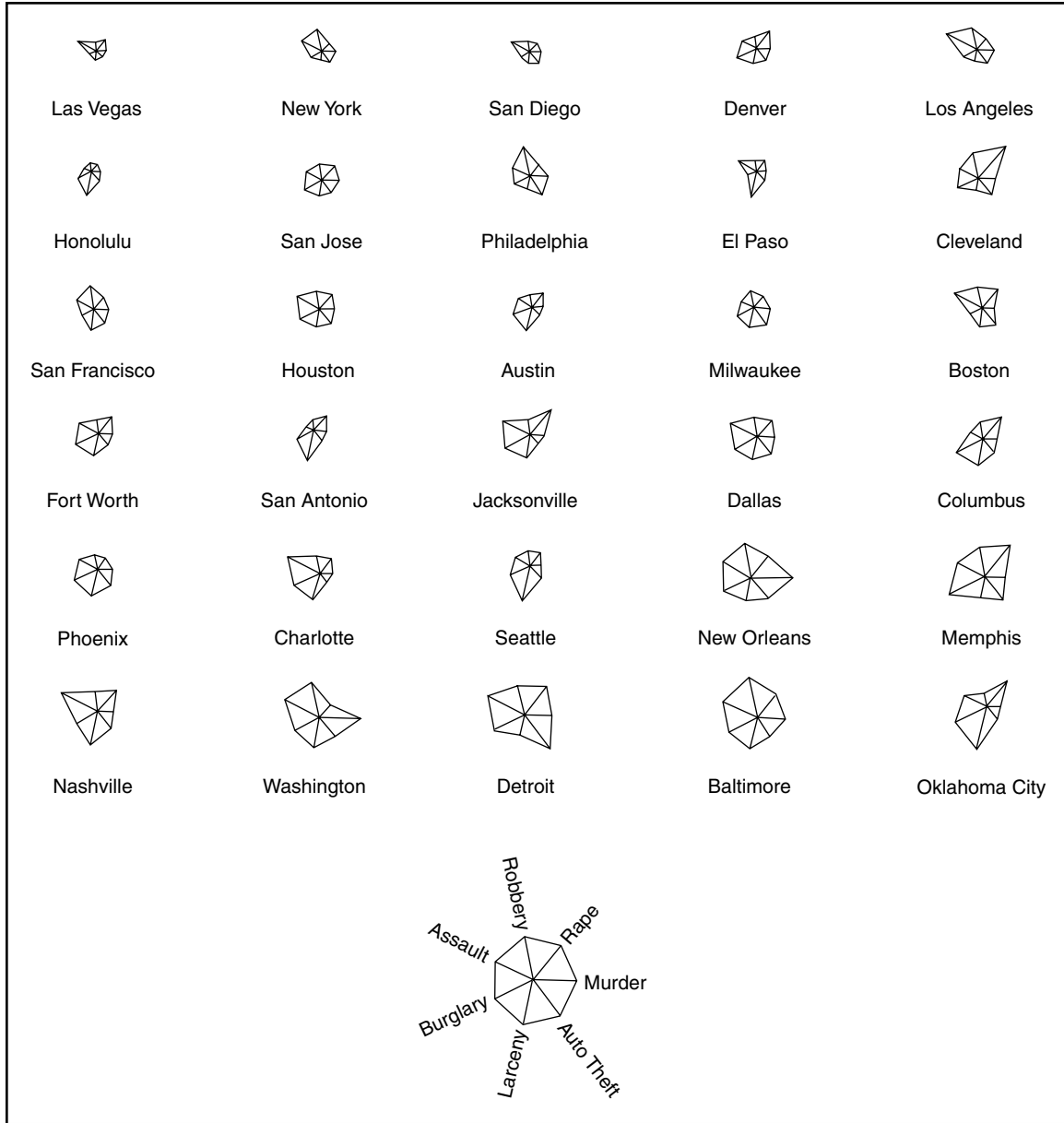


Figure 18. Star plot of rates of seven categories of crime in the thirty largest U.S. cities (Chicago is omitted because of missing data). The plot employs polar coordinates to represent each observation: Angles (the “points” of the star) encode variables, while distance from the origin (the center of the star) encodes the value of each variable. The crime rates were scaled (by range) before the graph was constructed. A key to the points of the star is shown at the bottom of the graph: “Murder” represents both murder and manslaughter.

SOURCE: *Statistical Abstract of the United States: 1998*.



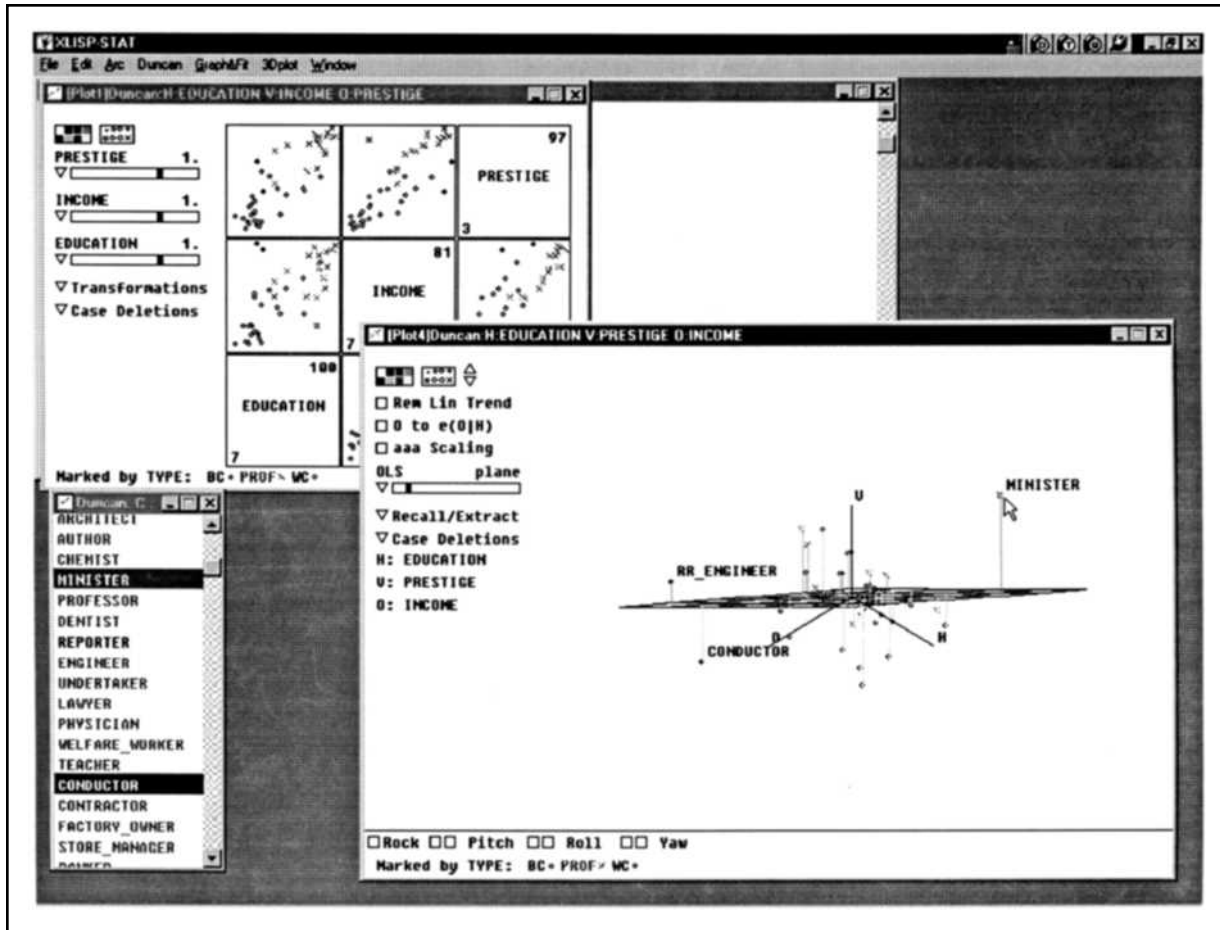


Figure 19. Modern statistical computer graphics: Cook and Weisberg's Arc. The window in the foreground contains a rotating three-dimensional scatterplot of Duncan's occupational prestige data. The points in the plot are marked by type of occupation; a regression plane and residuals to the plane also are shown. Several occupations have been identified with a mouse. (The mouse cursor currently points at the occupation "minister.") To the left and bottom of the window, a variety of controls for manipulating the plot appear. The small window at the bottom left of the screen contains the names of the observations; note that this window is linked to the three-dimensional scatterplot. At the upper left, partly hidden, is a window containing a scatterplot matrix of the data, which also is linked to the other windows. Plot controls for this graph include power-transformation sidebars at the left of the window.

Cleveland, William S. 1984 "Graphs in Scientific Publications." *American Statistician* 38:261-269.

— 1993 *Visualizing Data*. Summit N.J.: Hobart Press.

— 1994 *The Elements of Graphing Data*, rev. ed. Summit N.J.: Hobart Press.

—, and Robert McGill 1984 "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." *Journal of the American Statistical Association* 79:531-554.

Cook, R. Dennis 1998 *Regression Graphics: Ideas for Studying Regressions through Graphics*. New York: Wiley.

—, and Sanford Weisberg 1994 *An Introduction to Regression Graphics*. New York: Wiley.

— 1999 *Applied Regression Including Computing and Graphics*. New York: Wiley.

Duncan, Otis Dudley 1961 "A Socioeconomic Index for All Occupations." In Albert J. Reiss, Jr., Otis Dudley Duncan, Paul K. Hatt, and Cecil C. North, eds., *Occupations and Social Status*. New York: Free Press.

Fox, John 1991 *Regression Diagnostics*. Newbury Park, Calif.: Sage.

- 1997 *Applied Regression Analysis, Linear Models, and Related Methods*. Thousand Oaks, Calif.: Sage.
- forthcoming (a) *Nonparametric Simple Regression: Scatterplot Smoothing*. Thousand Oaks, Calif.: Sage.
- forthcoming (b) *Multiple and Generalized Nonparametric Regression*. Thousand Oaks, Calif.: Sage.
- , and Carole Suschnigg 1989 “A Note on Gender and the Prestige of Occupations.” *Canadian Journal of Sociology* 14:353–360.
- Funkhouser, H. Gray 1937 “Historical Development of the Graphical Representation of Statistical Data.” *Osiris* 3:267–404.
- Hartigan, John A. 1975 *Clustering Algorithms*. New York: Wiley.
- Hastie, Trevor J., and Robert J. Tibshirani 1990 *Generalized Additive Models*. London: Chapman and Hall.
- Huff, Darrell 1954 *How to Lie with Statistics*. New York: Norton.
- Kruskal, Joseph B., and Myron Wish 1978 *Multidimensional Scaling*. Beverly Hills, Calif.: Sage.
- Monette, Georges 1990 “Geometry of Multiple Regression and Interactive 3-D Graphics.” In John Fox and J. Scott Long, eds., *Modern Methods of Data Analysis*. Newbury Park, Calif.: Sage.
- Ornstein, Michael D. 1976 “The Boards and Executives of the Largest Canadian Corporations: Size, Composition, and Interlocks.” *Canadian Journal of Sociology* 1:411–437.
- Schmidt, Calvin F. 1983 *Graphics: Design Principles and Practices*. New York: Wiley.
- Spence, Ian, and Stephan Lewandowsky 1990 “Graphical Perception.” In John Fox and J. Scott Long, eds., *Modern Methods of Data Analysis*. Newbury Park, Calif.: Sage.
- Tierney, Luke 1990. *Lisp-Stat: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics*. New York: Wiley.
- Tufte, Edward R. 1983 *The Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press.
- 1990 *Envisioning Information*. Cheshire, Conn.: Graphics Press.
- 1997 *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Conn.: Graphics Press.
- Tukey, John W. 1977 *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley.
- Venables, W. N., and B. D. Ripley 1997 *Modern Applied Statistics with S-PLUS*, 2nd ed. New York: Springer-Verlag.

Wainer, Howard 1997. *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. New York: Springer-Verlag.

JOHN FOX

## STATISTICAL INFERENCE

Making an inference involves drawing a general conclusion from specific observations. People do this every day. Upon arising in the morning, one observes that the sun is shining and that the day will be nice. The news reports the arrest of a military veteran for child abuse, and a listener infers that military veterans have special adjustment problems. Statistical inference is a way of formalizing the process of drawing general conclusions from limited information. It is a way of stating the degree of confidence one has in making an inference by using probability theory. Statistically based research allows people to move beyond speculation.

Suppose a sociologist interviews two husbands. Josh, whose wife is employed, does 50 percent of the household chores; Frank, whose wife does not work for pay, does 10 percent. Should the sociologist infer that husbands do more housework when their wives are employed? No. This difference could happen by chance with only two cases. However, what if 500 randomly selected husbands with employed wives average 50 percent of the chores and randomly selected husbands with nonemployed wives average 10 percent? Since this difference is not likely to occur by chance, the sociologist infers that husbands do more housework when their wives are employed for pay.

Researchers perform statistical inferences in three different ways. Assume that 60 percent of the respondents to a survey say they will vote for Marie Chavez. The *traditional hypothesis testing* approach infers that Chavez will win the election if chance processes would account for the result (60 percent support in this survey) with less than some a priori specified statistical significance level. For example, if random chance could account for the result fewer than five times in a hundred, one would say the results are statistically significant. Statistical significance levels are called the *alpha*

(e.g.,  $\alpha = .05$  for the 5 percent level). If Chavez would get 60 percent support in a sample of the size selected less than 5 percent of the time by chance, one would infer that she will win. The researcher picked the 5 percent level of significance before doing the survey. (The test, including the  $\alpha$  level, must be planned *before* one looks at the findings.) If one would get this result 6 percent of the time by chance, there is no inference. Note that not making the inference means just that: One does not infer that Chavez's opponent will win.

A second strategy involves stating the *likelihood of the result occurring by chance* without an a priori level of significance. This strategy reports the result (60 percent of the sample supported Chavez) and the probability of getting that result by chance, say, .042. This gives readers the freedom to make their inferences using whatever level of significance they wish. Sam Jones, using the .01 level ( $\alpha = .01$ ) in the traditional approach would see that the results do not meet his criterion. He would not conclude that Chavez will win. Mara Jabar, using the .05 level, would conclude that Chavez would win.

The third strategy places a *confidence interval around a result*. For example, a researcher may be 95 percent confident that Chavez will get between 55 percent and 65 percent of the votes. Since the entire interval—55 percent to 65 percent—is enough for a victory, that is, is greater than 50 percent one infers that Chavez will win.

Each approach has an element of risk attached to the inference. That risk is the probability of getting the result by chance alone. Sociologists tend to pick low probabilities (e.g., .05, .01, and even .001), because they do not want to conclude that something is true when it is at all likely to have occurred by chance.

**TRADITIONAL TESTS OF SIGNIFICANCE**

Traditional tests of significance involve six steps. Three examples are used here to illustrate these steps: (1) A candidate will win an election, (2) mothers with at least one daughter will have different views on abortion than will mothers with only sons, and (3) the greater a person's internal political efficacy is, the more likely that person is to vote.

*Step 1:* State a hypotheses ( $H_1$ ) in terms of statistical parameters (characteristics such as means, correlations, proportions) of the population:

$H1$ :  $P(\text{vote for the candidate}) < .50$ . [Read: The mean for mothers with daughters is not equal to the mean for mothers with sons.]

$H2$ :  $\mu$  mothers with daughters  $\neq$   $\mu$  mothers with sons. [Read: The means for mothers with daughters is not equal to the mean for mothers with sons.]

$H3$ :  $\rho < 0.0$ . [Read: The population correlation  $\rho$  (rho) between internal political efficacy and voting is greater than zero.]

$H2$  says that the means are different but does not specify the direction of the difference. This is a two-tail hypothesis, meaning that it can be significant in either direction. In contrast,  $H1$  and  $H3$  signify the direction of the difference and are called one-tail hypotheses.

These three hypotheses are not directly testable because each involves a range of values. *Step 2* states a null hypothesis, which the researcher usually wishes to reject, that has a specific value.

$H1_0$ :  $P(\text{vote for the candidate}) = .50$ .

$H2_0$ :  $\mu$  mothers with daughters =  $\mu$  mothers with sons.

$H3_0$ :  $\rho = 0$ .

An important difference between one-tail and two-tail tests may have crossed the reader's mind. Consider  $H1_0$ . If 40 percent of the sample supported the candidate, one fails to reject  $H1_0$  because the result was in the direction opposite of that of the one-tail hypothesis. In contrast, whether mothers with daughters have a higher or lower mean attitude toward abortion than do mothers with sons, one proceeds to test  $H2_0$  because a difference in either direction could be significant.

*Step 3* states the a priori level of significance. Sociologists usually use the .05 level. With large samples, they sometimes use the .01 or .001 level. This paper uses the .05 level ( $\alpha = .05$ ). If the result would occur in fewer than 5 percent (corresponding to the .05 level) of the samples if the null hypothesis were true in the population, the null hypothesis is rejected in favor of the main hypothesis.

Suppose the sample correlation between internal political efficacy and voting is .56 and this would occur in fewer than 5 percent of the samples this size if the population correlation were 0 (as specified in  $H_0$ ). One rejects the null hypothesis,  $H_0$ , and accepts the main hypothesis,  $H_1$ , that the variables are correlated in the population. What if the sample correlation were .13 and a correlation this large would occur in 25 percent of the samples from a population in which the true correlation were 0? Because 25 percent exceeds the a priori significance level of 5 percent, the null hypothesis is not rejected. One cannot infer that the variables are correlated in the population. Simultaneously, the results do not prove that the population correlation is .00, simply that it could be that value.

Step 4 selects a test statistic and its critical value. Common test statistics include  $z$ ,  $t$ ,  $F$ , and  $\chi^2$  (chi-square). The *critical value* is the value the test statistic must exceed to be significant at the level specified in step 3. For example, using a one-tail hypothesis, a  $z$  must exceed 1.645 to be significant at the .05 level. Using a two-tail hypothesis, a  $z$ , must exceed 1.96 to be significant at the .05 level. For  $t$ ,  $F$ , and  $\chi^2$ , determining the critical value is more complicated because one needs to know the degrees of freedom. A formal understanding of degrees of freedom is beyond the scope of this article, but an example will give the reader an intuitive idea. If the mean of five cases is 4 and four of the cases have values of 1, 4, 5, and 2, the last case must have a value of 8 (it is the only value for the fifth case that will give a mean of 4, since  $1 + 4 + 5 + 2 + x = 20$ , only if  $x = 8$  and  $20/5 = 4$ ). Thus, there are  $n - 1$  degrees of freedom. Most test statistics have different distributions for each number of degrees of freedom.

Figure 1 illustrates the  $z$  distribution. Under the  $z$  distribution, an absolute value of greater than 1.96 will occur by chance only 5 percent of the time. By chance a  $z > 1.96$  occurs 2.5 percent of the time and a  $z < -1.96$  occurs 2.5 percent of the time. Thus, 1.96 is the critical  $z$ -score for a two-tail .05 level test. The critical  $z$ -score for a one-tail test at the .05 level is 1.645 or  $-1.645$ , depending on the direction specified in the main hypothesis.

Step 5 computes the test statistic. An example appears below.

Step 6 decides whether to reject or fail to reject the null hypothesis. If the computed test statistic

exceeds the critical value, one rejects the null hypothesis and makes the inference to accept the main hypothesis. If the computed test statistic does not exceed the critical value, one fails to reject the null hypothesis and make no inference.

**Example of Six Steps Applied to  $H_1$ .** A random sample of 200 voters shows 60 percent of them supporting the candidate. Having stated the main hypothesis (step 1) and the null hypothesis (step 2), step 3 selects an a priori significance level at  $\alpha = .05$ , since this is the conventional level. Step 4 selects the test statistic and its critical level. To test a single percentage, a  $z$  test is used (standard textbooks on social statistics discuss how to select the appropriate tests statistics; see Agresti and Finlay 1996; Loether and McTavish 1993; Raymondo 1999; Vaughan 1997). Since the hypothesis is one-tail, the critical value is 1.645 (see Figure 1).

The fifth step computes the formula for the test statistic:

$$z = \frac{p_s - p}{\sqrt{\frac{pq}{n}}}$$

- where  $p_s$  is the proportion in the sample
- $p$  is the proportion in the population under  $H_0$
- $q$  is  $1 - p$
- $n$  is the number of people in the sample.

Thus,

$$z = \frac{.6 - .5}{\sqrt{(.5 \times .5) / 200}}$$

$$z = 2.828$$

The sixth step makes the decision to reject the null hypothesis, since the difference is in the predicted direction and  $2.828 > 1.645$ . The statistical inference is that the candidate will win the election.

### REPORTING THE PROBABILITY LEVEL

Many sociological researchers do not use the traditional null hypothesis model. Instead, they report the probability of the result. This way, a reader knows the probability (say, .042 or .058) rather than the significant versus not significant status. Reporting the probability level removes the “magic

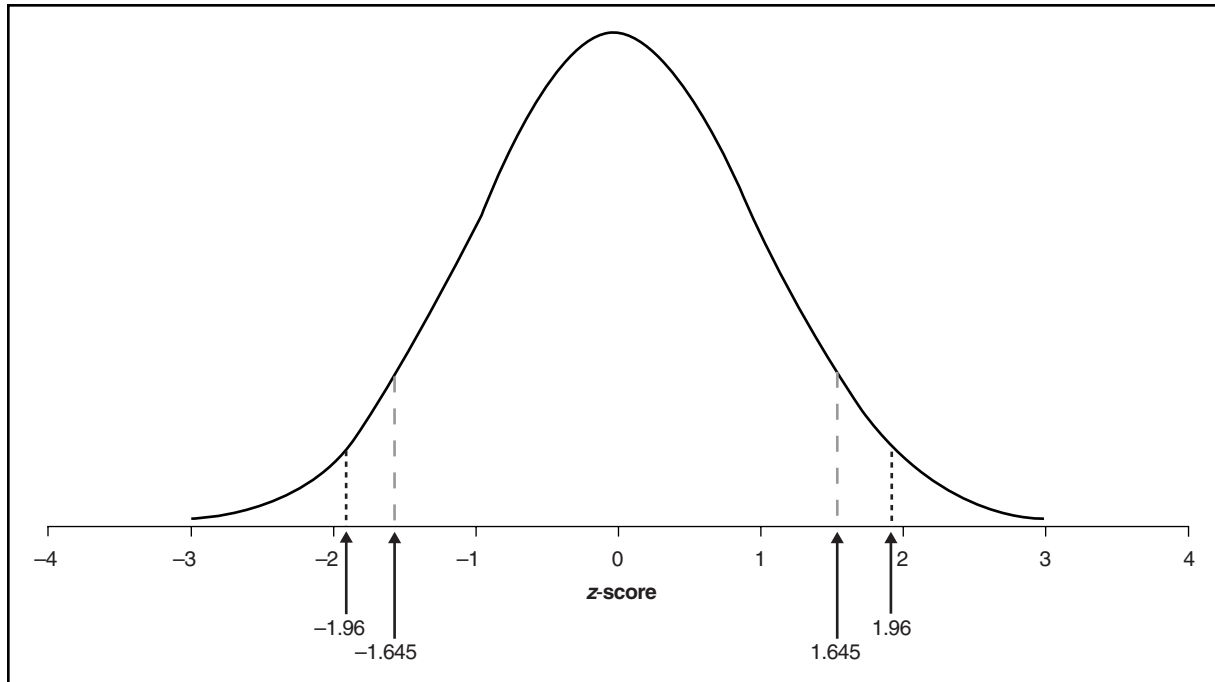


Figure 1. Normal deviate ( $z$ ) distribution.

of the level of significance.” A result that is significant at the .058 level is not categorically different from one that is significant at the .042 level. Where the traditional null hypothesis approach says that the first of these results is not significant and the second is, reporting the probability tells the reader that there is only a small difference in the degree of confidence attached to the two results. Critics of this strategy argue that the reader may adjust the significance level post hoc; that is, the reader may raise or lower the level of significance after seeing the results. It also is argued that it is the researcher, not the reader, who is the person testing the hypotheses; therefore, the researcher is responsible for selecting an a priori level of significance.

The strategy of reporting the probability is illustrated for *H1*. Using the tabled values or functions in standard statistical packages, the one-tail probability of a  $z = 2.828$  is .002. The researcher reports that the candidate had 60 percent of the vote in the sample and that the probability of getting that much support by chance is .002. This provides more information than does simply saying that it is significant at the .05 level. Results that could happen only twice in 1,000 times by chance

(.002) are more compelling than are results that could happen five times in 100 (.05).

Since journal editors want to keep papers short and studies often include many tests of significance, reporting probabilities is far more efficient than going through the six-step process outlined above. The researcher must go through these steps, but the paper merely reports the probability for each test and places an asterisk along those which are significant at the .05 level. Some researchers place a single asterisk for results significant at the .05 level, two asterisks for results significant at the .01 level, and three asterisks for results significant at the .001 level.

### CONFIDENCE INTERVALS

Rather than reporting the significance of a result, this approach puts a confidence interval around the result. This provides additional information in terms of the width of the confidence interval.

Using a confidence interval, a person constructs a range of values such that he or she is 95 percent confident (some use a 99 percent confi-

dence interval) that the range contains the population parameter. The confidence interval uses a two-tail approach on the assumption that the population value can be either above or below the sample value.

For the election example,  $H1$ , the confidence interval is

$$p_s \pm z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

where  $z_{\alpha/2}$  is the two-tail critical value for the alpha level

$p_s$  is the proportion in the sample

$p$  is the proportion in the population under  $H_0$

$q$  is  $1 - p$

$n$  is the number of people in the sample

$$\begin{aligned} &= .6 \pm 1.96 \sqrt{\frac{.5 \times .5}{200}} \\ &= .6 \pm 1.96 \times .03535 \\ &= .6 \pm .0693 \end{aligned}$$

upper limit	.669
lower limit	.531

The researcher is 95 percent confident that the interval, .531 to .669, contains the true population proportion. The focus is on the confidence level (.95) for a result rather than the low likelihood of the null hypothesis (.05) used in the traditional null hypothesis testing approach.

The confidence interval has more information value than do the first two approaches. Since the value specified in the null hypothesis ( $H_0: P = .50$ ) is not in the confidence interval, the result is statistically significant at the .05 level. Note that a 95 percent confidence level corresponds to a .05 level of significance and that a 99 percent confidence interval corresponds to a .01 level of significance. Whenever the value specified by the null hypothesis is not in the confidence interval, the result is statistically significant. More important, the confidence interval provides an estimate of the range of possible values for the population. With 200 cases and 60 percent support, there is confidence that the candidate will win, although it may be a close election with the lower limit indicating 53.1 per-

cent of the vote or a landslide with the upper limit indicating 66.9 percent of the vote. If the sample were four times as large,  $n = 800$ , the confidence interval would be half as wide (.565-.635) and would give a better fix on the outcome.

### COMPUTATION OF TESTS AND CONFIDENCE INTERVALS

Table 1 presents formulas for some common tests of significance and their corresponding confidence intervals where appropriate. These are only a sample of the tests that are commonly used, but they cover means, differences of means, proportions, differences of proportions, contingency tables, and correlations. Not included are a variety of multivariate tests for analysis of variance, regression, path analysis, and structural equation models. The formulas shown in Table 1 are elaborated in most standard statistics textbooks (Agresti and Finlay 1996; Blalock 1979; Bohrnstedt and Knoke 1998; Loether and McTavish 1993; Raymond 1999; Vaughan 1997).

### LOGIC OF STATISTICAL INFERENCE

A formal treatment of the logic of statistical inference is beyond the scope of this article; the following is a simplified description. Suppose one wants to know whether a telephone survey can be thought of as a random sample. From current census information, suppose the mean,  $\mu$ , income of the community is \$31,800 and the standard deviation,  $\sigma$ , is \$12,000. A graph of the complete census enumeration appears in Panel A of Figure 2. The fact that there are a few very wealthy people skews the distribution.

A telephone survey included interviews with 1,000 households. If it is random, its sample mean and standard deviation should be close to the population parameters,  $\mu$  and  $\sigma$ , respectively. Assume that the sample has a mean of \$33,200 and a standard deviation of \$10,500. To distinguish these sample statistics from the population parameters, call them  $M$  and  $s$ . The sample distribution appears in Panel B by Figure 2. Note that it is similar to the population distribution but is not as smooth.

One cannot decide whether the sample could be random by looking at Panels A and B. The distributions are different, but this difference might

**Common Tests of Significance Formulas**

What is Being Tested?	$H_1$	$H_0$	Test Statistic	Large-Scale Confidence Interval
Single mean against value specified as $\chi$ in $H_0$	1-tail: $\mu > \chi$ 2-tail: $\mu \neq \chi$	$\mu = \chi$	$t$ with $n - 1$ degrees of freedom $t = \frac{M - \chi}{s/\sqrt{n}}$	$M \pm t_{\alpha/2} \sigma / \sqrt{n}$
Single proportion against value specified as $\chi$ in $H_0$	1-tail: $P > \chi$ 2-tail: $P \neq \chi$	$P = \chi$	$z = \frac{P_s - \chi}{\sqrt{PQ/N}}$	$P_s \pm z_{\alpha/2} \sqrt{\frac{PQ}{n}}$
Difference between two means	1-tail: $\mu_1 > \mu_2$ 2-tail: $\mu_1 \neq \mu_2$	$\mu_1 = \mu_2$	$t$ with $n_1 + n_2$ degrees of freedom: $t = \frac{M_1 - M_2}{\sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$ where $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$M_1 - M_2 \pm t_{\alpha/2} \sigma_{M_1 - M_2}$ where $\sigma_{M_1 - M_2}$ is defined as the denominator of the $t$ -test in the cell to the immediate left
Difference between two proportions	1-tail: $P_1 > P_2$ 2-tail: $P_1 \neq P_2$	$P_1 = P_2$	$z = \frac{P_{s_1} - P_{s_2}}{\sqrt{PQ \sqrt{\frac{n_1 + n_2}{n_1 n_2}}}}$	$P_{s_1} - P_{s_2} \pm z_{\alpha/2} \sigma_{P_{s_1} - P_{s_2}}$ where $\sigma_{P_{s_1} - P_{s_2}}$ is defined as the numerator of the $z$ -test in the cell to the immediate left
Significance of contingency table	The level on one variable depends on the level on the second variable	No dependency between the variables	$\chi^2 = \frac{\sum(F_o - F_e)^2}{F_e}$	Not applicable
Single correlation	1-tail $\rho > 0$ 2-tail $\rho \neq 0$	$\rho = 0$	$F$ with 1 and $n - 2$ degrees of freedom: $F = \frac{r^2(n - 2)}{1 - r^2}$	Complex, since it is not symmetrical

Table 1

have occurred by chance. Statistical inference is accomplished by introducing two theoretical distributions: the sampling distribution of the mean and the  $z$ -distribution of the normal deviate. A theoretical distribution is different from the population and sample distributions in that a theoretical distribution is mathematically derived; it is not observed directly.

**Sampling Distribution of the Mean.** Suppose that instead of taking a single random sample of 1,000 people, one took two such samples and determined the mean of each one. With 1,000 cases, it is likely that the two samples would have means that were close together but not the same. For instance, the mean of the second sample might be \$30,200. These means, \$33,200 and \$30,200, are pretty close to each other. For a sample to have a mean of, say \$11,000, it would have to include a

greatly disproportionate share of poor families; this is not likely by chance with a random sample with  $n = 1,000$ . For a sample to have a mean of, say, \$115,000, it would have to have a greatly disproportionate share of rich families. In contrast, with a sample of just two individuals, one would not be surprised if the first person had an income of \$11,000 and the second had an income of \$115,000.

The larger the samples are, the more stable the mean is from one sample to the next. With only 20 people in the first and second samples, the means may vary a lot, but with 100,000 people in both samples, the means should be almost identical. Mathematically, it is possible to derive a distribution of the means of all possible samples of a given  $n$  even though only a single sample is observed. It can be shown that the mean of the

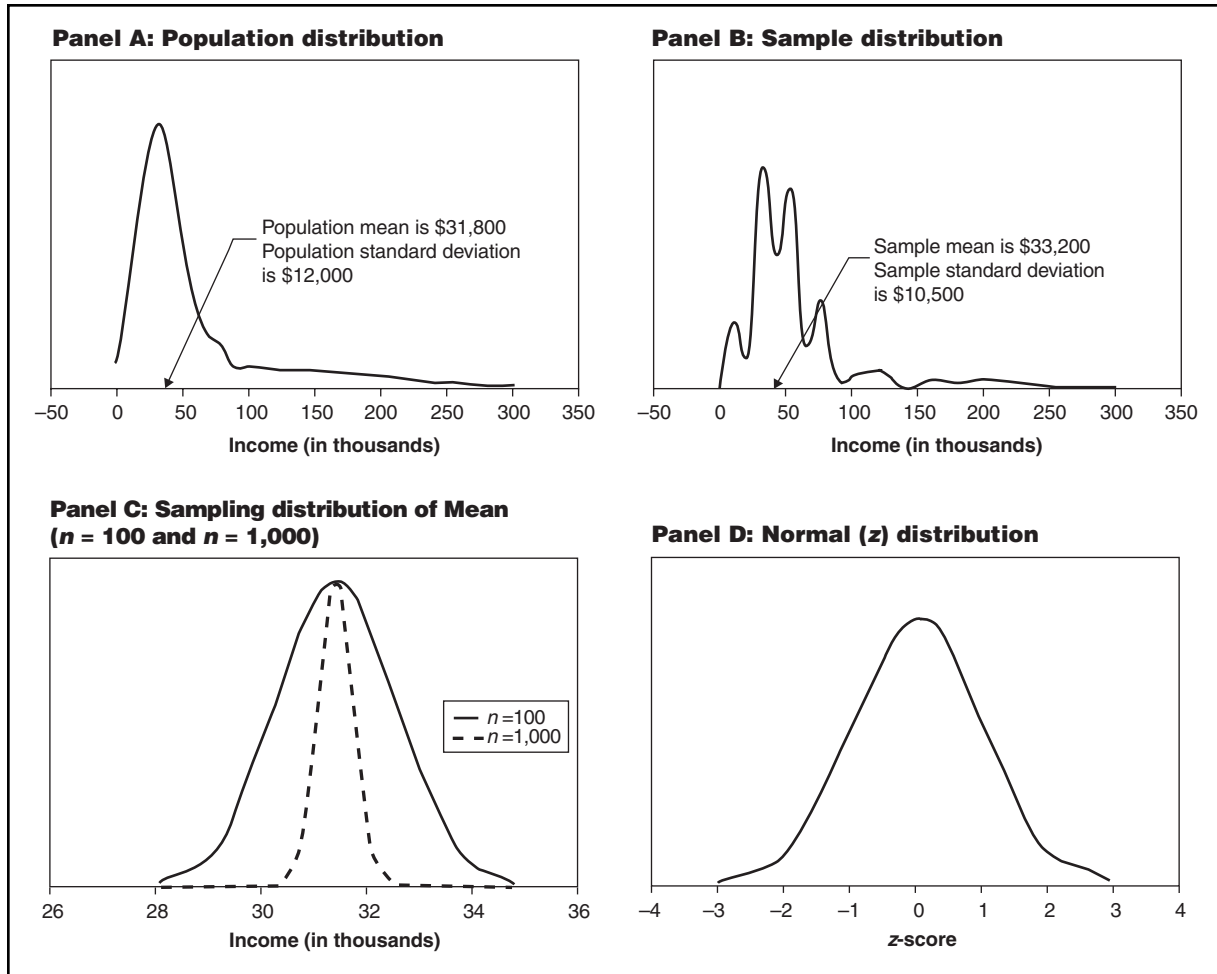


Figure 2. Four distributions used in statistical inference: (A) population distribution; (B) sample distribution; sampling distribution for  $n=100$  and  $n=1,000$ ; and (D) normal deviate ( $z$ ) distributions

sampling distribution of means is the population mean and that the standard deviation of the sampling distribution of the means is the population standard deviation divided by the square root of the sample size. The standard deviation of the mean is called the *standard error of the mean*:

$$\text{Standard error of the mean (SEM)} = \sigma_M = \frac{\sigma}{\sqrt{n}}$$

This is an important derivation in statistical theory. Panel C shows the sampling distribution of the mean when the sample size is  $n = 1,000$ . It also shows the sampling distribution of the mean for  $n = 100$ . A remarkable property of the sampling distribution of the mean is that with a large sample

size, it will be normally distributed even though the population and sample distributions are skewed.

One gets a general idea of how the sample did by seeing where the sample mean falls along the sampling distribution of the mean. Using Panel C for  $n = 1,000$ , the sample  $M = \$33,200$  is a long way from the population mean. Very few samples with  $n = 1,000$  would have means this far from the population mean. Thus, one infers that the sample mean probably is based on a nonrandom sample.

Using the distribution in Panel C for the smaller sample size,  $n = 100$ , the sample  $M = \$33,200$  is not so unusual. With 100 cases, one should not be surprised to get a sample mean this far from the population mean.



Being able to compare the sample mean to the population mean by using the sampling distribution is remarkable, but statistical theory allows more precision. One can transform the values in the sampling distribution of the mean to a distribution of a test statistic. The appropriate test statistic is the distribution of the normal deviate, or  $z$ -distribution. It can be shown that

$$z = \frac{M - \mu}{\sigma/\sqrt{n}}$$

If the  $z$ -value were computed for the mean of all possible samples taken at random from the population, it would be distributed as shown in Panel D of Figure 2. It will be normal, have a mean of zero, and have a variance of 1.

Where is  $M = \$33,200$  under the distribution of the normal deviate using the sample size of  $n = 1,000$ ? Its  $z$ -score using the above formula is

$$\begin{aligned} z &= \frac{33,200 - 31,800}{12,000/\sqrt{1,000}} \\ &= 3.689 \end{aligned}$$

Using tabled values for the normal deviate, the probability of a random sample of 1,000 cases from a population with a mean of \$31,800 having a sample mean of \$33,200 is less than .001. Thus, it is extremely unlikely that the sample is purely random.

With the same sample mean but with a sample of only 100 people,

$$\begin{aligned} z &= \frac{33,200 - 31,800}{12,000/\sqrt{100}} \\ &= 1.167 \end{aligned}$$

Using tabled values for a two-tail test, the probability of getting the sample mean this far from the population mean with a sample of 100 people is .250. One should not infer that the sample is nonrandom, since these results could happen 25 percent of the time by chance.

The four distributions can be described for any sample statistic one wants to test (means, differences of means, proportions, differences of proportions, correlations, etc). While many of the calculations will be more complex, their logic is identical.

## MULTIPLE TESTS OF SIGNIFICANCE

The logic of statistical inference applies to testing a single hypothesis. Since most studies include multiple tests, interpreting results can become extremely complex. If a researcher conducts 100 tests, 5 of them should yield results that are statistically significant at the .05 level by chance. Therefore, a study that includes many tests may find some “interesting” results that appear statistically significant but that really are an artifact of the number of tests conducted.

Sociologists pay less attention to “adjusting the error rate” than do those in most other scientific fields. A conservative approach is to divide the Type I error by the number of tests conducted. This is known as the Dunn multiple comparison test, based on the Bonferroni inequality. For example, instead of doing nine tests at the .05 level, each test is done at the  $.05/9 = .006$  level. To be viewed as statistically significant at the .05 level, each specific test must be significant at the .006 level.

There are many specialized multiple comparison procedures, depending on whether the tests are planned before the study starts or after the results are known. Brown and Melamed (1990) describe these procedures.

## POWER AND TYPE I AND TYPE II ERRORS

To this point, only one type of probability has been considered. Sociologists use statistical inference to minimize the chance of accepting a main hypothesis that is false in the population. They reject the null hypothesis only if the chances of it’s being true in the population are very small, say,  $\alpha = .05$ . Still, by minimizing the chances of this error, sociologists increase the chance of failing to reject the null hypothesis when it should be rejected. Table 2 illustrates these two types of error.

Type I, or  $\alpha$ , error is the probability of rejecting  $H_0$  falsely, that is, the error of deciding that  $H_1$  is right when  $H_0$  is true in the population. If one were testing whether a new program reduced drug abuse among pregnant women, the  $H_1$  would be that the program did this and the  $H_0$  would be that the program was no better than the existing one. Type I error should be minimized because it would be wrong to change programs when the new program was no better than the existing one. Type I

**Type I ( $\alpha$ ) and Type II ( $\beta$ ) Errors**

Decision Made by the Researcher	True Situation in the Population	
	$H_0$ , the null hypothesis, is true	$H_1$ , the main hypothesis, is true
$H_0$ , the null hypothesis is true	$1 - \alpha$	$\beta$
$H_1$ , the main hypothesis is true	$\alpha$	$1 - \beta$

Table 2

error has been described as “the chances of discovering things that aren’t so” (Cohen 1990, p. 1304). The focus on Type I error reflects a conservative view among scientists. Type I error guards against doing something new (as specified by  $H_1$ ) when it is not going to be helpful.

Type II, or  $\beta$ , error is the probability of failing to reject  $H_0$  when  $H_1$  is true in the population. If one failed to reject the null hypothesis that the new program was no better ( $H_0$ ) when it was truly better ( $H_1$ ), one would put newborn children at needless risk. Type II error is the chance of missing something new (as specified by  $H_1$ ) when it really would be helpful.

Power is  $1 - \beta$ . Power measures the likelihood of rejecting the null hypothesis when the alternative hypothesis is true. Thus, if there is a real effect in the population, a study that has a power of .80 can reject the null hypothesis with a likelihood of .80. The power of a statistical test is measured by how likely it is to do what one usually wants to do: demonstrate support for the main hypothesis when the main hypothesis is true. Using the example of a treatment for drug abuse among pregnant women, the power of a test is the ability to demonstrate that the program is effective if this is really true.

Power can be increased. First, get a larger sample. The larger the sample, the more power to find results that exist in the population. Second, increase the  $\alpha$  level. Rather than using the .01 level of significance, a researcher can pick the .05 or even the .10. The larger  $\alpha$  is, the more powerful the test is in its ability to reject the null hypothesis when the alternative is true.

There are problems with both approaches. Increasing sample size makes the study more costly. If there are risks to the subjects who participate,

adding cases exposes additional people to that risk. An example of this would be a study that exposed subjects to a new drug treatment program that might create more problems than it solved. A larger sample will expose more people to these risks.

Since Type I and Type II errors are inversely related, raising  $\alpha$  reduces  $\beta$  thus increasing the power of the test. However, sociologists are hesitant to raise  $\alpha$  since doing so increases the chance of deciding something is important when it is not important. With a small sample, using a small  $\alpha$  level such as .001 means there is a great risk of  $\beta$  error. Many small-scale studies have a Type II error of over .50. This is common in research areas that rely on small samples. For example, a review of one volume of the *Journal of Abnormal Psychology* (this journal includes many small-sample studies) found that those studies average Type II error of .56 (Cohen 1990). This means the psychologist had inadequate power to reject the null hypothesis when  $H_1$  was true. When  $H_1$  was true, the chance of rejecting  $H_0$  (i.e., power) was worse than that resulting from flipping a coin.

Some areas that rely on small samples because of the cost of gathering data or to minimize the potential risk to subjects require researchers to plan their sample sizes to balance  $\alpha$ , power, sample size, and the minimum size of effect that is theoretically important. For example, if a correlation of .1 is substantively significant, a power of .80 is important, and an  $\alpha = .01$  is desired, a very large sample is required. If a correlation is substantively and theoretically important only if it is over .5, a much smaller sample is adequate. Procedures for doing a power analysis are available in Cohen (1988); see also Murphy and Myous (1998).

Power analysis is less important for many sociological studies that have large samples. With a large sample, it is possible to use a conservative  $\alpha$  error rate and still have sufficient power to reject the null hypothesis when  $H_1$  is true. Therefore, sociologists pay less attention to  $\beta$  error and power than do researchers in fields such as medicine and psychology. When a sociologist has a sample of 10,000 cases, the power is over .90 that he or she will detect a very small effect as statistically significant. When tests are extremely powerful to detect small effects, researchers must focus on the substantive significance of the effects. A correlation of

.07 may be significant at the .05 level with 10,000 cases, but that correlation is substantively trivial.

### STATISTICAL AND SUBSTANTIVE SIGNIFICANCE

Some researchers and many readers confuse statistical significance with substantive significance. Statistical inference does not ensure substantive significance, that is, ensure that the result is important. A correlation of .1 shows a weak relationship between two variables whether it is statistically significant or not. With a sample of 100 cases, this correlation will not be statistically significant; with a sample of 10,000 cases, it will be statistically significant. The smaller sample shows a weak relationship that might be a zero relationship in the population. The larger sample shows a weak relationship that is all but certainly a weak relationship in the population, although it is not zero. In this case, the statistical significance allows one to be confident that the relationship in the population is substantively weak.

Whenever a person reads that a result is statistically significant, he or she is confident that there is some relationship. The next step is to decide whether it is substantively significant or substantively weak. Power analysis is one way to make this decision. One can illustrate this process by testing the significance of a correlation. A population correlation of .1 is considered weak, a population correlation of .3 is considered moderate, and a population correlation of .5 or more is considered strong. In other words, if a correlation is statistically significant but .1 or lower, one has to recognize that this is a weak relationship—it is statistically significant but substantively weak. It is just as important to explain to the readers that the relationship is substantively weak as it is to report that it is statistically significant. By contrast, if a sample correlation is .5 and is statistically significant, one can say the relationship is both statistically and substantively significant.

Figure 3 shows power curves for testing the significance of a correlation. These curves illustrate the need to be sensitive to both statistical significance and substantive significance. The curve on the extreme left shows the power of a test to show that a sample correlation,  $r$ , is statistically significant when the population correlation,  $\rho$  (rho), is .5. With a sample size of around 100, the power

of a test to show statistical significance approaches 1.0, or 100 percent. This means that any correlation that is this strong in the population can be shown to be statistically significant with a small sample.

What happens when the correlation in the population is weak? Suppose the true correlation in the population is .2. A sample with 500 cases almost certainly will produce a sample correlation that is statistically significant, since the power is approaching 1.0. Many sociological studies have 500 or more cases and produce results showing that substantively weak relationships,  $\rho = .2$ , are statistically significant. Figure 3 shows that even if the population correlation is just .1, a sample of 1,000 cases has the power to show a sample result that is statistically significant. Thus, any time a sample is 1,000 or larger, one has to be especially careful to avoid confusing statistical and substantive significance.

The guidelines for distinguishing between statistical and substantive significance are direct but often are ignored by researchers:

1. If a result is not statistically significant, regardless of its size in the sample, one should be reluctant to generalize it to the population.
2. If a result is statistically significant in the sample, this means that one can generalize it to the population but does not indicate whether it is a weak or a strong relationship.
3. If a result is statistically significant and strong in the sample, one can both generalize it to the population and assert that it is substantively significant.
4. If a result is statistically significant and weak in the sample, one can both generalize it to the population and assert that it is substantively weak in the population.

This reasoning applies to any test of significance. If a researcher found that girls have an average score of 100.2 on verbal skills and boys have an average score of 99.8, with girls and boys having a standard deviation of 10, one would think this as a very weak relationship. If one constructed a histogram for both girls and boys, one would find them almost identical. This difference is not substantively significant. However, if there was a sufficiently

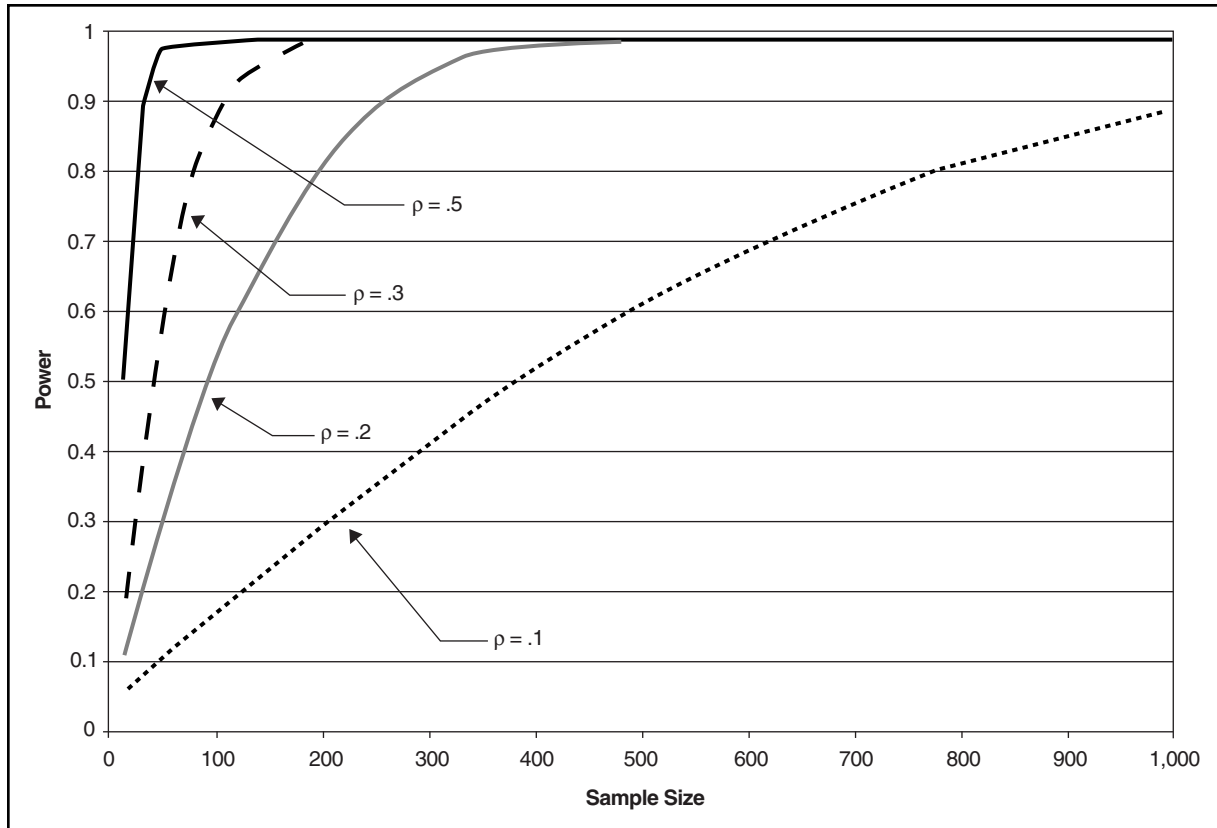


Figure 3. Power of test of  $r$ ,  $\alpha = .05$

large sample of girls and boys, say,  $n = 10,000$ , it could be shown that the difference is statistically significant. The statistical significance means that there is some difference, that the means for girls and boys are not identical. It is necessary to use judgment, however, to determine that the difference is substantively trivial. An abuse of statistical inference that can be committed by sociologists who do large-scale research is to confuse statistical and substantive significance.

### NONRANDOM SAMPLES AND STATISTICAL INFERENCE

Very few researchers use true random samples. Sometimes researchers use convenience sampling. An example is a social psychologist who has every student in a class participate in an experiment. The students in this class are not a random sample of the general population or even of students in a university. Should statistical inference be used here?

Other researchers may use the entire population. If one wants to know if male faculty members are paid more than female faculty members at a particular university, one may check the payroll for every faculty member. There is no sample—one has the entire population. What is the role of statistical inference in this instance?

Many researchers would use a test of significance in both cases, although the formal logic of statistical inference is violated. They are taking a “what if” approach. If the results they find could have occurred by a random process, they are less confident in their results than they would be if the results were statistically significant. Economists and demographers often report statistical inference results when they have the entire population. For example, if one examines the unemployment rates of blacks and whites over a ten-year period, one may find that the black rate is about twice the white rate. If one does a test of significance, it is unclear what the population is to which one wants

to generalize. A ten-year period is not a random selection of all years. The rationale for doing statistical inference with population data and nonprobability samples is to see if the results could have been attributed to a chance process.

A related problem is that most surveys use complex sample designs rather than strictly random designs. A stratified sample or a clustered sample may be used to increase efficiency or reduce the cost of a survey. For example, a study might take a random sample of 20 high schools from a state and then interview 100 students from each of those schools. This survey will have 2,000 students but will not be a random sample because the 100 students from each school will be more similar to each other than to 100 randomly selected students. For instance, the 100 students from a school in a ghetto may mostly have minority status and mostly be from families that have a low income in a population with a high proportion of single-parent families. By contrast, 100 students from a school in an affluent suburb may be disproportionately white and middle class.

The standard statistical inference procedures discussed here that are used in most introductory statistics texts and in computer programs such as SAS and SPSS assume random sampling. When a different sampling design is used, such as a cluster design, a stratified sample, or a longitudinal design, the test of significance will be biased. In most cases, the test of significance will underestimate the standard errors and thus overestimate the test statistic ( $z$ ,  $t$ ,  $F$ ). The extent to which this occurs is known as the “design effect.” The most typical design effect is greater than 1.0, meaning that the computed test statistic is larger than it should be. Specialized programs allow researchers to estimate design effects and incorporate them in the computation of the test statistics. The most widely used of these procedures are WesVar, which is available from SPSS, and SUDAAN, a stand-alone program. Neither program has been widely used by sociologists, but their use should increase in the future.

## REFERENCES

Agresti, Alan, and Barbara Finlay 1996 *Statistical Methods for the Social Sciences*. Englewood Cliffs, N.J.: Prentice-Hall.

Blalock, Hubert M., Jr. 1979 *Social Statistics*. New York: McGraw-Hill.

Bohrnstedt, George W., and David Knoke 1988 *Statistics for Social Data Analysis*, 2nd ed. Itasca, Ill.: F.E. Peacock.

Brown, Steven R., and Lawrence E. Melamed 1990 *Experimental Design and Analysis*. Newbury Park, Calif.: Sage.

Cohen, Jacob 1988 *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, N.J.: Erlbaum.

— 1990 “Things I Have Learned (So Far).” *American Psychologist* 45:1304–1312.

Loether, Herman J., and Donald G. McTavish 1993 *Descriptive and Inferential Statistics*. New York: Allyn and Bacon.

Murphy, Kelvin R., and Brentt Myous, eds. 1998 *Statistical Power Analysis: A Simple and Graphic Model for Traditional and Modern Hypothesis Tests*. Hillsdale, N.J.: Erlbaum.

Raymondo, James 1999 *Statistical Analysis in the Social Sciences*. New York: McGraw-Hill.

Vaughan, Eva D. 1997 *Statistics: Tools for Understanding Data in Behavioral Sciences*. Englewood Cliffs, N.J.: Prentice-Hall.

ALAN C. ACOCK

## STATISTICAL METHODS

In the 1960s, the introduction, acceptance, and application of multivariate statistical methods transformed quantitative sociological research. Regression methods from biometrics and economics; factor analysis from psychology; stochastic modeling from engineering, biometrics, and statistics; and methods for contingency table analysis from sociology and statistics were developed and combined to provide a rich variety of statistical methods. Along with the introduction of these techniques came the institutionalization of quantitative methods. In 1961, the American Sociological Association (ASA) approved the Section on Methodology as a result of efforts organized by Robert McGinnis and Albert Reiss. The ASA’s yearbook, *Sociological Methodology*, first appeared in 1969 under the editorship of Edgar F. Borgatta and George W. Bohrnstedt. Those editors went on to establish the quarterly journal *Sociological Methods and Research* in 1972. During this period, the National Institute of Mental Health began funding training

programs that included rigorous training in quantitative methods.

This article traces the development of statistical methods in sociology since 1960. Regression, factor analysis, stochastic modeling, and contingency table analysis are discussed as the core methods that were available or were introduced by the early 1960s. The development of additional methods through the enhancement and combination of these methods is then considered. The discussion emphasizes statistical methods for causal modeling; consequently, methods for data reduction (e.g., cluster analysis, smallest space analysis), formal modeling, and network analysis are not considered.

### THE BROADER CONTEXT

By the end of the 1950s, the central ideas of mathematical statistics that emerged from the work of R. A. Fisher and Karl Pearson were firmly established. Works such as Fisher's *Statistical Methods for Research Workers* (1925), Kendall's *Advanced Theory of Statistics* (1943, 1946), Cramér's *Mathematical Methods of Statistics* (1946), Wilks's *Mathematical Statistics* (1944), Lehman's *Testing Statistical Hypotheses* (1959), Scheffé's *The Analysis of Variance* (1959), and Doob's *Stochastic Processes* (1953) systematized the key results of mathematical statistics and provided the foundation for developments in applied statistics for decades to come. By the start of the 1960s, multivariate methods were applied routinely in psychology, economics, and the biological sciences. Applied treatments were available in works such as Snedecor's *Statistical Methods* (1937), Wold's *Demand Analysis* (Wold and Juréen, 1953), Anderson's *An Introduction to Multivariate Statistical Analysis* (1958), Simon's *Models of Man* (1957), Thurstone's *Multiple-Factor Analysis* (1947), and Finney's *Probit Analysis* (1952).

These methods are computationally intensive, and their routine application depended on developments in computing. BMD (Biomedical Computing Programs) was perhaps the first widely available statistical package, appearing in 1961 (Dixon et al. 1981). SPSS (Statistical Package for the Social Sciences) appeared in 1970 as a result of efforts by a group of political scientists at Stanford to develop a general statistical package specifically

for social scientists (Nie et al. 1975). In addition to these general-purpose programs, many specialized programs appeared that were essential for the methods discussed below. At the same time, continuing advances in computer hardware increased the availability of computing by orders of magnitude, facilitating the adoption of new statistical methods.

### DEVELOPMENTS IN SOCIOLOGY

It is within the context of developments in mathematical statistics, sophisticated applications in other fields, and rapid advances in computing that major changes occurred in quantitative sociological research. Four major methods serve as the cornerstones for later developments: regression, factor analysis, stochastic processes, and contingency table analysis.

**Regression Analysis and Structural Equation Models.** Regression analysis is used to estimate the effects of a set of independent variables on one or more dependent variables. It is arguably the most commonly applied statistical method in the social sciences. Before 1960, this method was relatively unknown to sociologists. It was not treated in standard texts and was rarely seen in the leading sociological journals. The key notions of multiple regression were introduced to sociologists in Blalock's *Social Statistics* (1960). The generalization of regression to systems of equations and the accompanying notion of causal analysis began with Blalock's *Causal Inferences in Nonexperimental Research* (1964) and Duncan's "Path Analysis: Sociological Examples" (1966). Blalock's work was heavily influenced by the economist Simon's work on correlation and causality (Simon 1957) and the economist Wold's work on simultaneous equation systems (Wold and Juréen 1953). Duncan's work added the influence of the geneticist Wright's work in path analysis (Wright 1934). The acceptance of these methods by sociologists required a substantive application that demonstrated how regression could contribute to the understanding of fundamental sociological questions. In this case, the question was the determination of occupational standing and the specific work was the substantively and methodologically influential *The American Occupational Structure* by Blau and Duncan (1967), a work unsurpassed in its integration

of method and substance. Numerous applications of regression and path analysis soon followed. The diversity of influences, problems, and approaches that resulted from Blalock and Duncan's work is shown in Blalock's reader *Causal Models in the Social Sciences* (1971), which became the handbook of quantitative methods in the 1970s.

Regression models have been extended in many ways. Bielby and Hauser (1977) have reviewed developments involving systems of equations. Regression methods for time series analysis and forecasting (often called Box-Jenkins models) were given their classic treatment in Box and Jenkins's *Time Series Analysis* (1970). Regression diagnostics have provided tools for exploring characteristics of the data set that is to be analyzed. Methods for identifying outlying and influential observations have been developed (Belsley et al. 1980), along with major advances in classic problems such as heteroscedasticity (White 1980) and specification (Hausman 1978). All these extensions have been finding their way into sociological practice.

**Factor Analysis.** Factor analysis, a technique developed by psychometricians, was the second major influence on quantitative sociological methods. Factor analysis is based on the idea that the covariation among a larger set of *observed* variables can be reduced to the covariation among a smaller set of *unobserved* or latent variables. By 1960, this method was well known and applications appeared in most major sociology journals. Statistical and computational advances in applying maximum-likelihood estimation to the factor model (Jöreskog 1969) were essential for the development of the covariance structure model discussed below.

**Stochastic Processes.** Stochastic models were the third influence on the development of quantitative sociological methods. Stochastic processes model the change in a variable over time in cases where a chance process governs the change. Examples of stochastic processes include change in occupational status over a career (Blumen et al. 1955), friendship patterns, preference for job locations (Coleman 1964), and the distribution of racial disturbances (Spilerman 1971). While the mathematical and statistical details for many stochastic models had been worked out by 1960, they were relatively unknown to sociologists until the

publication of Coleman's *Introduction to Mathematical Sociology* (1964) and Bartholomew's *Stochastic Models for Social Processes* (1967). These books presented an array of models that were customized for specific social phenomena. While these models had great potential, applications were rare because of the great mathematical sophistication of the models and the lack of general-purpose software for estimating the models. Nonetheless, the influence of these methods on the development of other techniques was great. For example, Markov chain models for social mobility had an important influence on the development of loglinear models.

**Contingency Table Analysis and Loglinear Models.** Methods for categorical data were the fourth influence on quantitative methods. The analysis of contingency tables has a long tradition in sociology. Lazarsfeld's work on elaboration analysis and panel analysis had a major influence on the way research was done at the start of the 1960s (Lazarsfeld and Rosenberg 1955). While these methods provided useful tools for analyzing categorical data and especially survey data, they were nonstatistical in the sense that issues of estimation and hypothesis testing generally were ignored. Important statistical advances for measures of association in two-way tables were made in a series of papers by Goodman and Kruskal that appeared during the 1950s and 1960s (Goodman and Kruskal 1979). In the 1960s, nonstatistical methods for analyzing contingency tables were replaced by the loglinear model. This model made the statistical analysis of multiway tables possible. Early developments are found in papers by Birch (1963) and Goodman (1964). The development of the general model was completed largely through the efforts of Frederick Mosteller, Stephen E. Fienberg, Yvonne M. M. Bishop, Shelby Haberman, and Leo A. Goodman, which were summarized in Bishop et al.'s *Discrete Multivariate Analysis* (1975). Applications in sociology appeared shortly after Goodman's (1972) didactic presentation and the introduction of ECTA (Fay and Goodman 1974), a program for loglinear analysis. Since that time, the model has been extended to specific types of variables (e.g., ordinal), more complex structures (e.g., association models), and particular substantive problems (e.g., networks) (see Agresti [1990] for a treatment of recent developments). As with regression models, many early applications appeared in the area

of stratification research. Indeed, many developments in loglinear analysis were motivated by substantive problems encountered in sociology and related fields.

### ADDITIONAL METHODS

From these roots in regression, factor analysis, stochastic processes, and contingency table analysis, a wide variety of methods emerged that are now applied frequently by sociologists. Notions from these four areas were combined and extended to produce new methods. The remainder of this article considers the major methods that resulted.

**Covariance Structure Models.** The covariance structure model is a combination of the factor and regression models. While the factor model allowed imperfect multiple indicators to be used to extract a more accurately measured latent variable, it did not allow the modeling of causal relations among the factors. The regression model, conversely, did not allow imperfect measurement and multiple indicators. The covariance structure model resulted from the merger of the structural or causal component of the regression model with the measurement component of the factor model. With this model, it is possible to specify that each latent variable has one or more imperfectly measured observed indicators and that a causal relationship exists among the latent variables. Applications of such a model became practical after the computational breakthroughs made by Jöreskog, who published LISREL (*linear structural relations*) in 1972 (Jöreskog and van Thillo 1972). The importance of this program is reflected by the use of the phrase "LISREL models" to refer to this area.

Initially, the model was based on analyzing the covariances among observed variables, and this gave rise to the name "covariance structure analysis." Extensions of the model since 1973 have made use of additional types of information as the model has been enhanced to deal with multiple groups, noninterval observed variables, and estimation with less restrictive assumptions. These extensions have led to alternative names for these methods, such as "mean and covariance structure models" and, more recently, "structural equation modeling" (see Bollen [1989] and Browne and

Arminger [1995] for a discussion of these and other extensions).

**Event History Analysis.** Many sociological problems deal with the occurrence of an event. For example, does a divorce occur? When is one job given up for another? In such problems, the outcome to be explained is the time when the event occurred. While it is possible to analyze such data with regression, that method is flawed in two basic respects. First, event data often are censored. That is, for some members of the sample the event being predicted may not have occurred, and consequently a specific time for the event is missing. Even assuming that the censored time is a large number to reflect the fact that the event has not occurred, this will misrepresent cases in which the event occurred shortly after the end of the study. If one assigns a number equal to the time when the data collection ends or excludes those for whom the event has not occurred, the time of the event will be underestimated. Standard regression cannot deal adequately with censoring problems. Second, the regression model generally assumes that the errors in predicting the outcome are normally distributed, which is generally unrealistic for event data. Statistical methods for dealing with these problems began to appear in the 1950s and were introduced to sociologists in substantive papers examining social mobility (Spilerman 1972; Sorensen 1975; Tuma 1976). Applications of these methods were encouraged by the publication in 1976 of Tuma's program RATE for event history analysis (Tuma and Crockford 1976). Since that time, event history analysis has become a major form of analysis and an area in which sociologists have made substantial contributions (see Allison [1995] and Petersen [1995] for reviews of these methods).

**Categorical and Limited Dependent Variables.** If the dependent variable is binary, nominal, ordinal, count, or censored, the usual assumptions of the regression model are violated and estimates are biased. Some of these cases can be handled by the methods discussed above. Event history analysis deals with certain types of censored variables; loglinear analysis deals with binary, nominal, count, and ordinal variables when the independent variables are all nominal. Many other cases exist that require additional methods. These methods are called quantal response models or models for categorical, limited, or qualitative dependent vari-



ables. Since the types of dependent variables analyzed by these methods occur frequently in the social sciences, they have received a great deal of attention by econometricians and sociologists (see Maddala [1983] and Long [1997] for reviews of these models and Cameron and Trivedi [1998] on count models).

Perhaps the simplest of these methods is logit analysis, in which the dependent variable is binary or nominal with a combination of interval and nominal independent variables. Logit analysis was introduced to sociologists by Theil (1970). Probit analysis is a related technique that is based on slightly different assumptions. McKelvey and Zavoina (1975) extend the logit and probit models to ordinal outcomes. A particularly important type of limited dependent variable occurs when the sample is selected nonrandomly. For example, in panel studies, cases that do not respond to each wave may be dropped from the analysis. If those who do not respond to each wave differ nonrandomly from those who do respond (e.g., those who are lost because of moving may differ from those who do not move), the resulting sample is not representative. To use an example from a review article by Berk (1983), in cases of domestic violence, police may write a report only if the violence exceeds some minimum level, and the resulting sample is biased to exclude cases with lower levels of violence. Regression estimates based on this sample will be biased. Heckman's (1979) influential paper stimulated the development of sample selection models, which were introduced to sociologists by Berk (1983). These and many other models for limited dependent variables are extremely well suited to sociological problems. With the increasing availability of software for these models, their use is becoming more common than even that of the standard regression model.

**Latent Structure Analysis.** The objective of latent structure analysis is the same as that of factor analysis: to explain covariation among a larger number of observed variables in terms of a smaller number of latent variables. The difference is that factor analysis applies to interval-level observed and latent variables, whereas latent structure analysis applies to observed data that are noninterval. As part of the American soldier study, Paul F. Lazarsfeld, Sam Stouffer, Louis Guttman, and others developed techniques for “factor ana-

lyzing” nominal data. While many methods were developed, latent structure analysis has emerged as the most popular. Lazarsfeld coined the term “latent structure analysis” to refer to techniques for extracting latent variables from observed variables obtained from survey research. The specific techniques depend on the characteristics of the observed and latent variables. If both are continuous, the method is called factor analysis, as was discussed above. If both are discrete, the method is called latent class analysis. If the factors are continuous but the observed data are discrete, the method is termed latent trait analysis. If the factors are discrete but the data are continuous, the method is termed latent profile analysis. The classic presentation of these methods is presented in Lazarsfeld and Henry's *Latent Structure Analysis* (1968). Although these developments were important and their methodological concerns were clearly sociological, these ideas had few applications during the next twenty years. While the programs ECTA, RATE, and LISREL stimulated applications of the loglinear, event history, and covariance structure models, respectively, the lack of software for latent structure analysis inhibited its use. This changed with Goodman's (1974) algorithms for estimation and Clogg's (1977) program MLLSA for estimating the models. Substantive applications began appearing in the 1980s, and the entire area of latent structure analysis has become a major focus of statistical work.

**Multilevel and Panel Models.** In most of the models discussed here, observations are assumed to be independent. This assumption can be violated for many reasons. For example, in panel data, the same individual is measured at multiple time points, and in studies of schools, all the children in each classroom may be included in the sample. Observations in a single classroom or for the same person over time tend to be more similar than are independent observations. The problems caused by the lack of independence are addressed by a variety of related methods that gained rapid acceptance beginning in the 1980s, when practical issues of estimation were solved. When the focus is on clustering with social groups (such as schools), the methods are known variously as hierarchical models, random coefficient models, and multilevel methods. When the focus is on clustering with panel data, the methods are referred to as models for cross-section and time series data, or simply

panel analysis. The terms “fixed and random effects models” and “covariance component models” also are used. (See Hsiao [1995] for a review of panel models for continuous outcomes and Hamerle and Ronning [1995] for panel models for categorical outcomes. Bryk and Raudenbush [1992] review hierarchical linear models.)

**Computer-Intensive Methods.** The availability of cheap computing has led to the rapid development and application of computer-intensive methods that will change the way data are analyzed over the next decade. Methods of resampling, such as the bootstrap and the jackknife, allow practical solutions to previously intractable problems of statistical inference (Efron and Tibshirani 1993). This is done by recomputing a test statistic perhaps 1,000 times, using artificially constructed data sets. Computational algorithms for Bayesian analysis replace difficult or impossible algebraic derivations with computer-intensive simulation methods, such as the Markov chain algorithm, the Gibbs sampler, and the Metropolis algorithm (Gelman et al. 1995). Related developments have occurred in the treatment of missing data, with applications of the EM algorithm and Markov chain Monte Carlo techniques (Schafer 1997).

**Other Developments.** The methods discussed above represent the major developments in statistical methods in sociology since the 1960s. With the rapid development of mathematical statistics and advances in computing, new methods have continued to appear. Major advances have been made in the treatment of missing data (Little and Rubin 1987). Developments in statistical graphics (Cleveland 1985) are reflected in the increasing number of graphics appearing in sociological journals. Methods that require less restrictive distributional assumptions and are less sensitive to errors in the data being analyzed are now computationally feasible. Robust methods have been developed that are insensitive to small departures from the underlying assumptions (Rousseeuw and Leroy 1987). Resampling methods (e.g., bootstrap methods) allow estimation of standard errors and confidence intervals when the underlying distributional assumptions (e.g., normality) are unrealistic or the formulas for computing standard errors are intractable by letting the observed data assume the role of the underlying population (Stine 1990). Recent work by Muthén (forthcom-

ing) and others combines the structural component of the regression model, latent variables from factor and latent structure models, hierarchical modeling, and characteristics of limited variables into a single model. The development of Mplus (Muthén and Muthén 1998) makes routine application of this general model feasible.

## CONCLUSIONS

The introduction of structural equation models in the 1960s changed the way sociologists viewed data and viewed the social world. Statistical developments in areas such as econometrics, biometrics, and psychometrics were imported directly into sociology. At the same time, other methods were developed by sociologists to deal with substantive problems of concern to sociology. A necessary condition for these changes was the steady decline in the cost of computing, the development of efficient numerical algorithms, and the availability of specialized software. Without developments in computing, these methods would be of little use to substantive researchers. As the power of desktop computers grows and the ease and flexibility of statistical packages increase, the application of sophisticated statistical methods has become more accessible to the average researcher than the card sorter was for constructing contingency tables in the 1950s and 1960s. As computing power continues to develop, new and promising methods are appearing with each issue of the journals in this area.

Acceptance of these methods has not been universal or without costs. Critiques of the application of quantitative methods have been written by both sympathetic (Lieberman 1985; Duncan 1984) and unsympathetic (Cosser 1975) sociologists as well as statisticians (Freedman 1987) and econometricians (Leamer 1983). While these critiques have made practitioners rethink their approaches, the developments in quantitative methods that took shape in the 1960s will continue to influence sociological practice for decades to come.

## REFERENCES

- Agresti, Alan 1990 *Categorical Data Analysis*. New York: Wiley.
- Allison, Paul D. 1995 *Survival Analysis Using the SAS® System: A Practical Guide*. Cary, NC: SAS Institute.

- Anderson, T. W. 1958 *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.
- Bartholomew, D. J. 1967 *Stochastic Models for Social Processes*. New York: Wiley.
- Belsley, David A., Edwin Kuh, and Roy E. Welsch 1980 *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.
- Berk, R. A. 1983 "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48:386-398.
- Bielby, William T., and Robert M. Hauser 1977 "Structural Equation Models." *Annual Review of Sociology* 3:137-161.
- Birch, M. W. 1963. "Maximum Likelihood in Three-Way Contingency Tables." *Journal of the Royal Statistical Society Series B* 27:220-233.
- Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland 1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT Press.
- Blalock, Hubert M., Jr. 1960 *Social Statistics*. New York: McGraw-Hill.
- 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- , 1971 *Causal Models in the Social Sciences*. Chicago: Aldine.
- Blau, Peter M., and Otis Dudley Duncan 1967 *The American Occupational Structure*. New York: Wiley.
- Blumen, I., M. Kogan, and P. J. McCarthy 1955 *Industrial Mobility of Labor as a Probability Process*. Cornell Studies of Industrial and Labor Relations, vol. 6. Ithaca, N.Y. Cornell University Press.
- Bollen, Kenneth A. 1989 *Structural Equations with Latent Variables*. New York: Wiley.
- Borgatta, Edgar F., and George W. Bohrnstedt, eds. 1969 *Sociological Methodology*. San Francisco: Jossey-Bass.
- , eds. 1972 *Sociological Methods and Research*. Beverly Hills, Calif.: Sage.
- Box, George E. P., and Gwilym M. Jenkins 1970 *Time Series Analysis*. San Francisco: Holden-Day.
- Browne, Michael W., and Gerhard Arminger 1995 "Specification and Estimation of Mean- and Covariance-Structure Models." In Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- Bryk, Anthony S., and Stephen W. Raudenbush 1992 *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage.
- Cameron, A. Colin, and Pravin K. Trivedi 1998 *Regression Analysis of Count Data*. New York: Cambridge University Press.
- Cleveland, William S. 1985 *The Elements of Graphing Data*. Monterey, Calif.: Wadsworth.
- Clogg, Clifford C. 1977 *MLLSA: Maximum Likelihood Latent Structure Analysis*. State College: Pennsylvania State University.
- Coleman, James S. 1964 *Introduction to Mathematical Sociology*. Glencoe, Ill.: Free Press.
- Coser, Lewis F. 1975 "Presidential Address: Two Methods in Search of Substance." *American Sociological Review* 40:691-700.
- Cramér, Harald 1946 *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton University Press.
- Dixon, W. J. chief ed. 1981 *BMD Statistical Software*. Berkeley: University of California Press.
- Doob, J. L. 1953. *Stochastic Processes*. New York: Wiley.
- Duncan, Otis Dudley 1966 "Path Analysis: Sociological Examples." *American Journal of Sociology* 72:1-16.
- 1984 *Notes on Social Measurement: Historical and Critical*. New York: Russell Sage Foundation.
- Efron, Bradley, and Robert J. Tibshirani 1993 *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- Fay, Robert, and Leo A. Goodman 1974 *ECTA: Everyman's Contingency Table Analysis*.
- Finney, D. J. 1952 *Probit Analysis, 2nd ed.* Cambridge, UK: Cambridge University Press.
- Fisher, R. A. 1925 *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Freedman, David A. 1987 "As Others See Us: A Case Study in Path Analysis." *Journal of Educational Statistics* 12:101-128.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin 1995 *Bayesian Data Analysis*. New York: Chapman and Hall.
- Goodman, Leo A. 1964 "Simple Methods of Analyzing Three-Factor Interaction in Contingency Tables." *Journal of the American Statistical Association* 58:319-352.
- 1972. "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables." *American Sociological Review* 37:28-46.
- 1974 "The Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable. Part I: A Modified Latent Structure Approach." *American Journal of Sociology* 79:1179-1259.
- , and William H. Kruskal 1979 *Measures of Association for Cross Classification*. New York: Springer-Verlag.

- Hamerle, Alfred, and Gerd Ronning 1995 "Panel Analysis for Qualitative Variables." In Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- Hausman, J. A. 1978 "Specification Tests in Econometrics." *Econometrica* 46:1251-1272.
- Heckman, James J. 1979 "Sample Selection Bias as a Specification Error." *Econometrica* 47:153-161.
- Hsiao, Cheng 1995 "Panel Analysis for Metric Data." In Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- Jöreskog, Karl G. 1969 "A General Approach to Confirmatory Maximum Likelihood Factor Analysis." *Psychometrika* 34:183-202.
- , and Marielle van Thillo 1972 *LISREL: A General Computer Program for Estimating a Linear Structural Equation System Involving Multiple Indicators of Unmeasured Variables*. Princeton, N.J.: Educational Testing Service.
- Kendall, Maurice G. 1943 *Advanced Theory of Statistics*, vol. 1. London: Griffin.
- 1946. *Advanced Theory of Statistics*, vol. 2. London: Griffin.
- Lazarsfeld, Paul F., and Neil W. Henry 1968 *Latent Structure Analysis*. New York: Houghton Mifflin.
- , and Morris Rosenberg, eds. 1955 *The Language of Social Research*. New York: Free Press.
- Leamer, Edward E. 1983 "Let's Take the Con Out of Econometrics." *American Economic Review* 73:31-43.
- Lehmann, E. L. 1959 *Testing Statistical Hypotheses*. New York: Wiley.
- Lieberson, Stanley 1985 *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California.
- Little, Roderick J. A., and Donald B. Rubin 1987 *Statistical Analysis with Missing Data*. New York: Wiley.
- Long, J. Scott 1997 *Regression Models for Categorical and Limited Dependent Variables*. Newbury Park, Calif.: Sage.
- Maddala, G. S. 1983 *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.
- McKelvey, Richard D., and William Zavoina 1975 "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *Journal of Mathematical Sociology* 4:103-120.
- Muthén, Bengt O. 1998 "Second-Generation Structural Equation Modeling with a Combination of Categorical and Continuous Latent Variables: New Opportunities for Latent Class/Latent Growth Modeling." In A. Sayer and L. Collins, eds., *New Methods for the Analysis of Change*. Washington D.C.: APA.
- Muthén, Linda K., and Bengt O. Muthén 1998 *Mplus: The Comprehensive Modeling Program for Applied Researchers*. Los Angeles: Muthén & Muthén.
- Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent 1975 *Statistical Package for the Social Sciences, 2nd ed.* New York: McGraw-Hill.
- Petersen, Trond 1995 "Analysis of Event Histories." In Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- Rousseeuw, Peter J., and Annick M. Leroy 1987 *Robust Regression and Outlier Detection*. New York: Wiley.
- Schafer, J. L. 1997 *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.
- Scheffé, H. 1959 *The Analysis of Variance*. New York: Wiley.
- Simon, Herbert 1957 *Models of Man*. New York: Wiley.
- Snedecor, George W. 1937 *Statistical Methods*. Ames: Iowa State University Press.
- Sorensen, Aage 1975 "The Structure of Intragenerational Mobility." *American Sociological Review* 40:456-471.
- Spilerman, Seymour 1971 "The Causes of Racial Disturbances: Tests of an Explanation." *American Sociological Review* 36:427-442.
- 1972 "The Analysis of Mobility Processes by the Introduction of Independent Variables Into a Markov Chain." *American Sociological Review* 37:277-294.
- Stine, Robert 1990 "An Introduction to Bootstrap Methods." In John Fox and J. Scott Long, eds., *Modern Methods of Data Analysis*. Newbury Park, Calif.: Sage.
- Theil, H. 1970 "On the Estimation of Relationships Involving Qualitative Variables." *American Journal of Sociology* 76:103-154.
- Thurstone, L. L. 1947 *Multiple-Factor Analysis*. Chicago: University of Chicago Press.
- Tuma, Nancy B. 1976 "Rewards, Resources, and the Rate of Mobility." *American Sociological Review* 41:338-360.
- , and D. Crockford 1976 *Invoking RATE*. Center for the Study of Welfare Policy. Menlo Park, Calif.: Stanford Research Institute.
- White, Halbert 1980 "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity." *Econometrica* 48:817-838.
- Wilks, S. S. 1944 *Mathematical Statistics*. Princeton, N.J.: Princeton University Press

Wold, Herman, and Lars Juréen 1953 *Demand Analysis*.  
New York: Wiley.

Wright, Sewall 1934 "The Method of Path Coefficients."  
*Annals of Mathematical Statistics* 5:161–215.

J. SCOTT LONG

## STATUS ATTAINMENT

Status attainment is the process by which individuals attain positions in the system of social stratification in a society. If one thinks of social stratification as referring to the rewards society offers and the resources individuals use to obtain those rewards, education, occupation, and income are the key factors. The amount and kind of education people attain determine the kinds of jobs they get. The kind of work people do is the main determinant of their income. Moreover, the education, occupation, and income of parents largely determine the kinds of advantages or disadvantages they create for their children. Sociologists usually think of education, occupation, and income as the main aspects of socioeconomic status, and the study of status attainment is therefore the study of how these attributes of people are related both within and across generations.

### ESTABLISHMENT OF THE FIELD

As a distinctive area of research, status-attainment research had its origins in the work of Otis Dudley Duncan in the 1960s. Duncan (1961) reconceptualized the study of intergenerational occupational mobility—which is concerned with the degree and pattern of association between the kinds of work done by parents and offspring (in practice, fathers and sons)—as the study of the factors that determine who gets what sort of job, with the father's occupation being only one of several determining factors. Other researchers extended Duncan's findings to take account of the factors that determine how much schooling people get and how much money they make.

Duncan's conceptual reformulation was accompanied by two important technical innovations. The first was the creation of a socioeconomic status scale for occupations. Unlike education and income, occupation has no intrinsic metric: No

natural ordering of occupations exists in terms of relative status. For many kinds of research, however, especially the study of status attainment, it is desirable to arrange occupations into some sort of status hierarchy, that is, a hierarchy of the relative socioeconomic advantage enjoyed by people in different occupations. Duncan created such an ordering of occupations for the categories of the 1950 U.S. Census classification by taking the weighted average of the education and income of typical incumbents, with the weights chosen to maximize the association between the resulting socioeconomic status scale and the relative prestige of occupations as measured by popular evaluations. He was able to do this because prestige and socioeconomic status are very highly correlated: Occupations that have high socioeconomic status (that is, that require a great deal of education and pay well) also tend to have high prestige, and jobs that require little education and pay poorly tend to have low prestige.

Second, Duncan introduced path analysis into sociology. Path analysis is a way of statistically representing the relative strength of different relationships between variables, both direct and indirect. For example, it is known that educated people tend to earn more than do uneducated people, but it is not clear whether this is the case simply because they have jobs of higher status or whether, among those who have jobs of similar status, the better educated earn more than do the less well educated. Path analysis provides a way of answering this question: Even among people doing the same sort of work, the better educated tend to earn more.

### SUBSTANTIVE ISSUES

Four central issues have dominated research on status attainment. The first issue is the extent of "social reproduction," the tendency for class and socioeconomic status position to be perpetuated, or "reproduced," from generation to generation. A value assumption underlies this question. "Open" societies, that is, societies with low rates of social reproduction or, to put it differently, high rates of intergenerational social mobility, are regarded as desirable since they are assumed to have relatively high equality of opportunity and to emphasize "achievement" rather than "ascription" as the basis for socioeconomic success.

The second issue is the factors other than the status of parents that affect education, occupation, and income. Of course, some factors may be correlated with the status of parents and also may have an independent effect. For example, there is a modest negative correlation between socioeconomic status and fertility—high-status people tend to have fewer children—and there is also a tendency for people from large families not to go as far in school as people from small families do. Thus, part of the reason the children of high-status people go further in school is that such people have smaller families. However, it is also true that at any given level of parental status (e.g., for families where both parents are college-educated professionals), people from smaller families go further in school. Therefore, the number of siblings has an *independent* effect on educational attainment apart from its correlation with parental status. Sorting out such effects is facilitated by the application of path analysis.

The third issue is the extent to which there are sex and racial (or ethnic) differences in patterns of status attainment. With respect to gender, the questions are: Do men and women from similar social origins go equally far in school? Do equally qualified men and women get jobs of equal status? Are women paid as well as men doing similar work? The same set of questions is asked with respect to differences between racial and ethnic groups.

The fourth issue is whether the process of status attainment operates the same way in different countries or in the same country in different historical periods. What follows is a summary of what is known about each of these four issues with respect to educational attainment, occupational attainment, and income attainment.

## EDUCATION

**Reproduction.** In regard to the extent of educational reproduction, the evidence in the United States in the late twentieth century is clear: America is an “open” society. Educational attainment (how far people go in school) is only weakly dependent on parental status. Only about 20 percent of the variability in years of school completed can be attributed to the level of education attained by one’s father or mother. When several different family background characteristics are taken into

account, the connection is not much stronger; at most, about one-third of the variability in educational attainment can be attributed to the status of the family one comes from. The rest is due to factors unrelated to social origins.

**Other Factors.** Apart from the social status of parents, the main factors that affect educational attainment are intelligence, the number of siblings (as was noted above, all else being equal, people from large families get less schooling), family stability (those from nonintact families, people whose parents have divorced or died—one or both—go less far in school), the influence of “significant others” (family members, friends, and teachers), and academic performance (the better people do in school, the longer they continue to go to school).

The question naturally arises as to why and how origin status and these other factors affect educational attainment. In a country such as the United States, where education up to the college level is free, parental wealth has relatively little effect on whether people stay in school. This claim is supported by the observation that the effect of social origins on educational attainment declines with each successive educational transition. That is, social origins have a stronger influence on whether people graduate from high school than on whether high school graduates go on to college and an even weaker influence on the graduation chances of those who begin college. If parental wealth is not important, what is?

There are two underlying factors: “Human capital” (sometimes called “cultural capital”) is the most important, but “social capital” is involved as well. Human capital refers to the knowledge, skills, and motivations of individuals. The basic argument here is that growing up in a high-status family enhances one’s human capital and that those with high human capital do better in school and therefore gain more education, which of course further enhances their human capital. The idea is that children who grow up in well-educated families or professional families learn the kinds of skills and acquire the kinds of motivations that enable them to do well in school. There are many books in such houses, and there are often computers. Schoolwork is familiar to these children because it is the same sort of thing they find at home.

Social capital refers to the social connections people have with others. Here the idea is that

people are strongly influenced by the company they keep. Young people whose friends drop out of high school are more likely to drop out of high school themselves than are others whose friends have a social background and academic performance level that encourage educational attainment. Similarly, those whose friends go to college are more likely to go themselves than are others whose friends go to work after high school, and those whose teachers encourage them to continue their education are more likely to do so than are others whose records are just as good. Since people with high-status origins tend to live in neighborhoods with others of similar origins, they tend to have greater social capital than do those with low-status origins.

**Sex and Racial Differences.** In the United States, there is little difference in the average amount of education attained by men and women, but more men than women tend to be very well educated or very poorly educated; that is, more men than women graduate from college, but more men than women drop out of high school. However, the effect of social origins and other factors on educational attainment is very similar for men and women. Race and ethnicity are a different story. Blacks are substantially less well educated than are whites and those of other races. In part, this is the case because the parents of blacks are poorly educated. However, blacks are also less able to convert whatever advantage they do have into a corresponding advantage for their children. In particular, blacks do not go as far in school as would be predicted from their parents' status. The sharp difference between blacks and other groups is a continuing legacy of slavery. While there are differences in the educational attainment levels of other ethnic groups, those differences are largely the result of differences among those groups in the average status of parents.

In nations such as South Africa, where until 1994 racial distinctions were embedded in law and social institutions (as in the American South before 1964), racial differences in educational attainment are much larger than they are in the United States. Whereas in the United States in 1990 whites averaged 13.1 years of schooling and blacks averaged 12.3 years, a difference of 0.8 year in South Africa in 1991 whites averaged 10.0 years of schooling and blacks averaged 4.5 years, a difference of 5.5 years, with the other racial groups falling be-

tween these values. This was a direct consequence of government policies that created separate and unequal school systems for South Africa's four "official" racial groups.

**Cross-Cultural and Cross-Temporal Variations.** Differences between countries in the educational attainment process are due both to general factors such as the level of industrialization and to specific differences in the way education is organized. In general, in places where the level of educational inequality in the parents' generation is high, educational attainment is more dependent on social origins than it is in countries where the level of educational inequality in the parents' generation is low. This is a consequence of the effect of human capital acquired at home. In a country such as the United States, where janitors have about ten years of school and high school teachers have about sixteen, the son of a janitor will be able to compete in school much more effectively with the son of a high school teacher than is the case in a society such as India where high school teachers also have about sixteen years of schooling but janitors have no schooling at all and are illiterate. Second, in highly industrialized countries schooling is less dependent on social origins than it is in less industrialized countries, in part because schooling tends to be free in industrialized countries. Third, in countries where the state provides not only free education but financial subsidies to students, as has been done in eastern Europe and in some western European countries, education tends to be less dependent on social origins than it is in countries without such subsidies.

There is a worldwide trend for educational attainment to become less "ascriptive" over time. That is, in almost all countries educational attainment has become less and less dependent on social origins throughout most of the twentieth century. The reason for this is straightforward. As was mentioned above, the effect of social origins on the probability that people will move from one level of education to the next declines with each higher level of education. Therefore, since the average *level* of educational attainment has been steadily increasing in most countries, it follows that more and more people are in educational categories where social origins matter relatively little.

An important distinction in educational systems is that between divided and unitary systems.

In the United States, there is, with only modest exceptions, a single path to educational attainment: primary school, to secondary school, to college or university, to graduate or professional school. Students achieve a certain level of education and then leave school to take up other pursuits. Thus, years of schooling is a very good indicator of educational attainment. In Europe and elsewhere, schooling tends to be divided into parallel tracks. In particular, a distinction is made between academic and vocational tracks, beginning in secondary school. Thus, in Europe, educational attainment must be measured not only by the amount of schooling but by the type of schooling a student has. In general, academic credentials have more value in the labor market than do vocational credentials in that they lead to jobs with higher status and higher income.

Among nations at a similar level of economic development, there often are substantial variations in the dependence of education on social origins. For example, in the 1970—the latest period for which there are systematic comparative data—55 percent of French male university graduates were the sons of managers or professionals, while in Great Britain this was true of only 35 percent. In general, at every selection point, social origins mattered more in France than they in Great Britain. In this sense, one can say that the British educational system was (and probably still is) substantially more egalitarian than the French system.

Finally, particular historical events can have a major impact on educational attainment. For example, the 1966–1977 Cultural Revolution in China caused massive disruptions in almost all aspects of social life. Secondary schools were closed from 1966 to 1968; universities were closed until 1972 and, when they reopened, accepted students on the basis of political status rather than academic merit until 1977. The results were twofold. First, the educational advantage of high-status origins—particularly growing up in a professional family—were very reduced substantially for those who would have entered secondary school or university during that period. Second, the quality of education declined because even when the schools remained open, they were devoted largely to political indoctrination rather than conventional studies. The evidence indicates that those educated during the Cultural Revolution read less well than do

those with the same amount of schooling who were educated before or after the Cultural Revolution.

## OCCUPATIONAL STATUS

**Reproduction.** Like educational status, occupational status is only weakly related to social origins. However, it is somewhat harder to pin this down than is true for education since, unlike education, which is completed by most people early in life, occupational status may vary over the life course, as people change jobs. The convention in most research on occupational attainment therefore is to restrict the analysis to men (since women not only change jobs but move in and out of the labor force for marriage or childbearing) and to compare the occupations held by men at the time they are interviewed with the occupations of their fathers when the interviewed men were teenagers, usually age 14. The relationship between fathers' and sons' occupational statuses turns out to be even weaker than the relationship between parents' and offspring's educational attainment. Thus, with respect to occupational statuses as well as educational attainment, America is an open society.

**Other Factors.** In the analysis of occupational attainment, an important issue has been to assess the relative importance of social origins (measured by the father's occupational status) and education as determinants of men's occupational status. The ratio of these two effects has been taken as an indicator of the degree of societal openness. In the United States and most industrial societies, education is by far the most important determinant of occupational status, while the direct effect of a father's occupational status is very limited. In the past, many people directly inherited their occupational position from their parents (for example, the sons of farmers were likely to take over their fathers' farms, the sons of shopkeepers to take over their fathers' shops, and so on), but in modern societies such as the United States, where people tend to work in large organizations, most jobs cannot be inherited directly. Instead, occupational status inheritance, insofar as it occurs at all, results mainly from the children of high-status people going further in school and those going further in school attaining better jobs. However, since, as was shown above, education is largely independent of social origins, the results is that education serves mainly as a vehicle of social mo-



bility rather than a mechanism of social reproduction or status inheritance.

**Sex and Racial Differences.** The most striking difference between men and women is that most men work most of the time once they complete their schooling, whereas the work lives of many women are interrupted for childbearing and child rearing. However, the labor-force participation rates of women and men are converging in the United States as more women remain in the labor force even when their children are very young. In general, men and women work at jobs of equal status, although the specific jobs held by men and women are very different. Most managers, skilled and unskilled manual workers, and farm workers are men; most clerical and service workers are women; and professional, sales, and semiskilled manual jobs tend to be performed by both men and women. The sex segregation of the labor force has important implications for income differences between men and women, as is discussed below.

Blacks tend to work at lower-status occupations than do whites and others. In part this is due to their lower levels of educational attainment, but in part it is due to the fact that blacks are not able to obtain jobs as good as those which can be obtained by equally well-educated members of other groups. Again, as in the case of education, differences in occupational status among nonblack ethnic groups are largely attributable to differences in educational attainment.

**Cross-Cultural and Cross-Temporal Variations.** In highly industrialized societies and in relatively egalitarian societies, there is little direct transmission of occupational status from one generation to the next; in those societies, occupational transmission is largely indirect, occurring through education. In less industrialized and less egalitarian societies, the importance of the father's occupation as a determinant of occupational status increases and the importance of education decreases, although education always remains more important than the father's occupation even in the least-developed societies.

Although the association between father's and son's occupational statuses has been declining over time and is weaker in industrialized societies, the *pattern* of intergenerational occupational mobility appears to be largely invariant, with only minor variations across societies caused by specific

historical circumstances, at least in industrialized societies and probably in nonindustrialized societies as well. That is, the relative chances that, say, the son of a professional and the son of a laborer will become professionals, rather than skilled workers, appear to be essentially similar in all societies.

Despite the commonality in intergenerational mobility patterns, there are substantial national variations in the strength of the linkage between schooling and career beginnings. In general, there is a tighter connection between education and the status of one's first job in countries, such as Germany and Switzerland, where there are separate vocational and academic tracks and assignment to one or the other is made early and where vocational secondary education provides occupation-specific skills than there is in countries, such as Great Britain, Japan, and the United States, where neither condition holds. Japan is, however, a special case. Japanese secondary schools and universities are highly stratified on the basis of prestige. Schools have close connections with large business firms and are able to place their students there. Students from the best schools go to the best firms, where they are trained by being rotated through a series of jobs. Thus, there is very tight schooling-first job connection in Japan, but of a kind not well captured by the association between the amount of schooling and the prestige of the first job.

There are also national differences in the sensitivity of career opportunities to the expansion or contraction of the economy, depending on institutional differences, particularly in welfare state policies and labor market structures. In the United States, for example, rates of job mobility show great sensitivity to structural change and to the labor market resources of individual workers, whereas in the Netherlands, jobs are largely insulated from structural forces.

Finally, careers can be strongly affected by specific historical events. The collapse of communism in eastern Europe in 1989 forced many political officials and administrators into early retirement. However, since the political transformation was accompanied by an economic collapse, with the economies of many former communist countries shrinking by about one-third in the early 1990s, unemployment increased and many women and older workers left the labor force. At the same time, there were substantial new opportunities,

particularly in the newly emerging private sector of the economy. Thus, there was a substantial increase in occupational mobility, at least in the short run.

As with education, the extent of reproduction of occupational status has been systematically decreasing over time in almost all societies. The reasons for this are not clear. There may be a worldwide shift toward an emphasis on achievement as opposed to ascription, although the likelihood that a shift in value orientations could have such a large and systematic effect does not seem great. More likely, the systematic increase in the average level of education in almost all countries is responsible, since it is known that the association between fathers' and sons' occupational statuses decreases for those who have obtained higher levels of education.

### INCOME

**Reproduction.** Little is known about the extent of income reproduction because it is very difficult to measure income in the parents' generation. Most data used in intergenerational analyses are obtained by asking people to report on their parent's characteristics. While people tend to know how much schooling their parents had and what sort of work their fathers were doing when the respondents were teenagers, few people have a very good idea of what their parents' income were. However, one major study has obtained such information: a study of the graduating class of 1957 from Wisconsin high schools conducted by Sewell and Hauser (1975). This cohort of graduates has been followed up in a number of surveys over the years, so that information has become available about its members' occupations and incomes at various stages after completing school. In addition, with careful arrangements to guard confidentiality, the researchers were able to obtain information from the Wisconsin Department of Taxation and the Social Security Administration regarding the incomes of the parents at the time the students were in high school. These data suggest that the intergenerational transmission of income is even weaker than is true for education or occupation. Other ways of indirectly estimating this relationship yield similar results.

One possible reason for this is that income (measured in real dollars, that is, adjusted for

inflation) is highly variable over the life cycle and, for some workers—particularly those who are self-employed or whose jobs are dependent on the weather—even from year to year. Moreover, age differences in earnings vary systematically for different occupational groups. The earnings of professionals tend to increase steadily over the course of their careers, while at the other extreme, the earnings of unskilled laborers do not change at all. Thus, when they first start working, unskilled laborers earn as much as or more than do professionals just beginning their careers, but by the time they near retirement, professionals earn several times as much as laborers of the same age earn. Incomes are also highly variable from place to place, reflecting differences in the cost of living, and even within cities, different firms pay different wages or salaries for the same job. All these factors make individual variations in income rather unpredictable.

**Other Factors.** Unlike parental education and occupational status, which affect educational attainment but have little direct effect on occupational attainment or income, parental income directly affects the income of offspring even when education and occupational attainment are taken into account. In fact, parental income is nearly as important as occupational status in determining income and is more important than education. Apparently, there is a propensity to earn money, and this propensity is transmitted from generation to generation. Whether this reflects differences in values that transmitted from parents to their children—with some people choosing jobs on the basis of how well they pay and others choosing jobs on the basis of their intrinsic interest, how secure they are, and so on—or in another factor is not known.

Other factors that affect income even when parental education and the respondent's own education and occupational status are taken into account include ability, the quality of the college attended, and the kind of work people do. Doctors earn more than professors do even though the jobs are of similar status, and garbage collectors earn more than ditch diggers earn. There is an extensive, although inconclusive, literature on differences in earnings across industrial sectors, and there is some evidence that earnings are higher in more strongly unionized occupations and industries.

**Sex and Racial Differences.** Gender is the big story here. In the United States, among full-time year-round workers, women earn about 60 percent of what men earn, and this ratio has remained essentially unchanged since the 1950s. Of the 40 percent gender difference, about 20 percent can be accounted for by the greater work experience of men, differences in the kinds of education received, and similar factors. The other 20 percent is due in part to the fact that the jobs performed mainly by women tend to pay less than do the jobs performed mainly by men even though many of these jobs are similar with respect to the skill required, the effort involved, and the responsibility entailed, and in part to the fact that women tend to earn less than do men in the same occupations. This state of affairs is possible because of the extreme gender segregation of the labor force. Most jobs tend to be performed either mostly by men or mostly by women, with relatively few jobs open to both sexes.

One consequence of this is that, at least in the United States, poverty is concentrated in female-headed households, especially where there are young children present. Not only do women in such situations find it difficult to work because of their child care responsibilities, even when they do work, their earnings tend to be low. Thus, the total income of such households is often below the poverty line.

In the United States, racial differences in income are somewhat smaller than gender differences and have been declining steadily for the last half century, as has occupational segregation by race. There is little evidence that the racial composition of jobs affects their pay levels. Instead, racial differences in income are attributable both to the fact that many blacks tend to be less educated and to work at lower-status jobs than most whites and others and to the fact that blacks get a lower return on their education and occupational status than do whites and others. Interestingly, there appears to be an across-the-board difference between the earnings of black and other males at any given level of education, occupational status, and so forth. However, the racial difference in the earnings of women is somewhat more complicated. At low levels of education and occupational status, black women earn much less than do other women, but at high levels of education and occupational

status, there is little or no difference in earnings among women of all races.

**Cross-Cultural and Cross-Temporal Variations.** While international comparisons of the determinants of personal or family income are scarce, probably because of the difficulty in measuring income in a comparable way across countries, differences in the *distribution* of income across nations and over time are well established. Income inequality is related to the level of economic development in a curvilinear way: It is low for the least developed nations, where most people are peasants; high for nations at medium levels of development, which often display large regional differences as a result of uneven economic development; and low for the most developed nations, where a combination of tax and welfare policies tends to ensure that most of the population enjoys at least a moderately adequate standard of living and constrains opportunities to become extremely rich. Because of restrictions on the accumulation of private property in communist regimes, income inequality tends to be smaller than it is in capitalist nations at a corresponding level of economic development. Finally, rampant inflation, such as that which occurred in eastern Europe after the collapse of communism, may cause dramatic reversals of fortune, impoverishing those on fixed incomes, such as government employees and pensioners, and enriching sellers of goods and services whose prices keep pace with inflation.

#### REFERENCES

- Allmendinger, Jutta 1989 "Educational Systems and Labor Market Outcomes." *European Sociological Review* 5:231-250.
- Blau, Peter, and Otis Dudley Duncan 1957 *The American Occupational Structure*. New York: Wiley.
- Deng, Zhong, and Donald J. Treiman 1997 "The Impact of the Cultural Revolution on Trends in Educational Attainment in the People's Republic of China." *American Journal of Sociology* 103:391-428.
- DiPrete, Thomas A., Paul M. de Graaf, Ruud Laijckx, Michael Tählin, and Hans-Peter Blossfeld 1997 "Collectivist versus Individualist Mobility Regimes? Structural Change and Job Mobility in Four Countries." *American Journal of Sociology* 103:318-358.
- Duncan, Otis Dudley 1961 "A Socioeconomic Index for All Occupations," In Albert J. Reiss, Jr., ed., *Occupations and Social Structure*. New York: Free Press.

- Erikson, Robert, and John H. Goldthorpe 1992 *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford, UK: Clarendon.
- Featherman, David L., and Robert M. Hauser, 1978 *Opportunity and Change*. New York: Academic Press.
- Ganzeboom, Harry B. G., Ruud Luijkx, and Donald J. Treiman 1989 "Intergenerational Class Mobility in Comparative Perspective." *Research in Social Stratification and Mobility* 8:3-84.
- , and Donald J. Treiman 1993 "Preliminary Results on Educational Expansion and Educational Achievement in Comparative Perspective." In Henk A. Becker and Piet L. J. Hermkens, eds., *Solidarity of Generations: Demographic, Economic and Social Change, and Its Consequences*. Amsterdam, thesis.
- , ———, and Wout C. Ultee 1990 "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277-302.
- Hout, Michael 1988 "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358-1400.
- Mare, Robert D. 1980 "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75:295-305.
- Müller, Walter, and Wolfgang Karle 1993 "Social Selection in Educational Systems in Europe." *European Sociological Review* 9:1-24.
- Roos, Patricia A. 1985 *Gender and Work: A Comparative Analysis of Industrial Societies*. Albany: State University of New York Press.
- Rosenbaum, James E., and Takehiko Kariya 1991 "Do School Achievements Affect the Early Jobs of High School Graduates in the United States and Japan?" *Sociology of Education* 64:78-95.
- Sewell, William H., and Robert M. Hauser 1975 *Education, Occupation, and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Shavit, Yossi, and Hans-Peter Blossfeld, eds., 1993 *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Colo.: Westview.
- , and Walter Müller, eds., 1998 *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford, UK: Clarendon.
- Treiman, Donald J. 1998. "Results from the Survey of 'Social Stratification in Eastern Europe after 1989': What We Have Learned and What We Need to Find Out." In *Transformation Processes in Eastern Europe, 1997*, proceedings of an NWO workshop, Amsterdam, March. The Hague: NWO.
- , and Kam-Bor Yip 1989 "Educational and Occupational Attainment in 21 Countries." In Melvin L. Kohn, ed., *Cross-National Research in Sociology*. Newbury Park, Calif.: Sage.
- , Matthew McKeever, and Eva Fodor 1996 "Racial Differences in Occupational Status and Income in South Africa, 1980 and 1991." *Demography* 33:111-132.

DONALD J. TREIMAN

## STATUS INCONGRUENCE

The phenomenon sociologists call "status incongruence" has equivalents in many languages. Expressions such as "*nouveau riche*," "*déclassé*," "*roturier*" and "*parvenu*" show that people in many societies perceive the incongruence between various statuses. The popular dictum "the heart on the left, the pocket on the right" expresses this incongruence between positions and feelings.

As a sociological concept, status incongruence is relatively recent. It was devised some time after the adoption of the notion of "status," following the discovery of Max Weber's writings on this subject by American sociologists in the late 1930s. In the 1950s, some twelve articles were published on "status inconsistency," most of them in the *American Sociological Review*. Those articles had a cumulative effect. At a certain point in the 1960s, it was felt that the debate on this topic had become saturated. In the absence of more empirical evidence, the theoretical discussion on status incongruence stagnated, but in the meantime the concept had been diffused in textbooks and compendiums.

After a period of neglect, the concept of status inconsistency has been reinvigorated over the last two decades as sociologists on both sides of the Atlantic have acknowledged a "decline of social classes." However despite the fact that the idea of social class has been dethroned, social inequalities persist.

The concept of status incongruence is a companion of the theory of cross-pressure. The first article focusing directly on status incongruence appeared in the same year (1944) as *The People's Choice* by Lazarsfeld, Berelson, and McPhee. The two notions nevertheless remain distinct in the

sociological literature because they respond to different analytic needs.

The incidence of status incongruence increases in times of social upheaval, such as the period of the Weimar Republic, the economic depression in the United States in the early 1930s, and that in Russia after the implosion of the Soviet regime. In the two decades before the French Revolution of 1789, the incidence of status incongruence was particularly high.

In emphasizing the revolutionary potential of downward mobility, which he called the "proletarianization of middle classes," Marx paid little attention to upward mobility and the effects of status incongruence. That neglect has been considered by some scholars to be one of his more glaring errors (Lopreato and Hazelrigg, 1972 p. 445). In contemporary sociological literature, the notion of status incongruence is related to role theory, rational theory, the theory of relative deprivation, and the theory of social movements.

This article considers only advanced Western societies, partly because the empirical evidence on status incongruence is available primarily for those countries and partly because social mobility and its impact on status incongruence are a less widespread phenomenon in developing countries.

#### **STATUS INCONSISTENCY AS A CORRECTION OF WEAK CORRELATIONS**

For a long time in sociological research, correlation between levels of social stratifications and other variables were rarely as significant as expected in light of the hypotheses and theoretical frameworks that had been adopted. Even when the rudimentary dichotomy of manual and nonmanual was abandoned and more categories were taken into consideration, the empirical results did not provide satisfactory explanations. Even when class as a rigid and restricted concept was largely replaced by the dimension of occupational status, the research strategy was not improved. Certainly, the emphasis on status groups is one of Weber and Pareto's chief corrections of Marx's theory (Lopreato and Hazelrigg 1972, p. 83). Nevertheless, an essential approach was missing until the 1950s, that of status inconsistency, which marked an advance in sociological thinking. It has been demonstrated that the consistency or

inconsistency of a person's status based on various criteria is a better predictor of social behavior than is the level of status based on a single criterion.

#### **FROM SOCIAL CLASS TO STATUS INCONSISTENCY**

Status incongruence is generated by gaps in income, occupation, education, and ethnic origin and other inconsistencies between a person's social position in one domain and that person's relatively lower status in another dimension. Status incongruence can be found in census results by cross-tabulating indicators such as education, income, professional hierarchical position, qualification, and racial origin. There is a logical relationship between the spread of status incongruencies and the weakening of social class consciousness.

Status inconsistency has become an essential aspect of social stratification in contemporary postindustrial society. It has been exacerbated by the growth of the middle classes and the decline of the peasantry and the industrial working class. Vertical mobility is the main source of status discrepancy. Most studies of social mobility have focused on upward mobility, particularly during the postwar period of economic development, but in more recent times, downward mobility has become equally important. Today, social mobility consists mostly in what Lipset and Zetterberg (1956, p. 563) called "the interchange of ranks." For every upward move, there must be a downward move. What was then only a hypothesis has been confirmed empirically: "[S]ome proportion of the children of the middle class fall in socio-economic status; some do not have the abilities to complete higher education or to get along in a bureaucratic hierarchy, and fall by the wayside. Whatever the reason that some persons of middle class origin move downward, they leave room for others of lower-class background to rise" (p. 570). Today, millions of Europeans and Americans born into the middle classes are in such incongruent situations. The downward move can be intragenerational or intergenerational.

Another source of status incongruence is liberation from primary social groups, particularly religious communities and families. More and more, through schooling, individual achievement negates the constraints of family background. For this

reason, status inconsistency is a fertile ground for individualistic tendencies.

The concept of status inconsistency raises the concept of status crystallization, which was proposed by Lenski (1954) as a nonvertical dimension of social status. Strong or weak status crystallization refers to the degree of incongruence or coherence of a person's ranking according to various criteria. A strong status crystallization implies that a person is rated consistently on all important criteria, whether the rating is high or low. Today, a large part of the population in Western societies finds itself in a situation of weak status crystallization. Solid social class can exist only if the majority of the population experiences strong status crystallization.

One of the most visible varieties of status incongruence occurs among schoolteachers, who are more numerous today than were workers in the heavy industry plants four decades ago. For many teachers there is a serious gap between the level of their education and their role in society and income level. The left-wing orientation of most teachers in European countries can be explained in terms of status incongruence, rather than class. Even some college professors experience this incongruence.

If one compares status incongruence today and in the past, two important categories have become prominent over the last two decades: the "intellectual proletariat" and the "ethnic achiever" (as opposed to the "skidder").

The spread of education in most advanced societies has highlighted the need of postindustrial economies for highly educated people. Today, two-thirds of people aged 18 are still in school. At the end of their college years, most of them do not find a job that corresponds to their expectations in terms of intellectual and economic rewards. It is in this category of the population—young educated people "with diplomas in their pockets"—that the rate of unemployment is the highest in most west European countries. This overabundance of graduates results from the incapacity of a highly technological society to absorb them in "interesting" occupations, with the existing jobs being protected by unions. This imbalance between the level of education, the quality of the job, and the amount of income generates status incongruences for "overeducated" young people. An advanced

postindustrial society in search of productivity replaces people with machines, producing a new kind of educated proletariat that was born into the middle class. In western Europe in the last decade (except in Germany), one of every four or five young people under age 25 was unemployed, and others were pushed down into "degraded" jobs. Those who accept jobs beneath their abilities, "degraded jobs," represent one of the most frequent varieties of status incongruence, a "reserve army" of alienated people.

The ethnic achiever is a new variety in western Europe and an old one in the United States. Status inconsistency can be found among ethnic and racial minorities in Great Britain, France, Germany, Belgium, the Netherlands, Switzerland, and Austria. Immigrants of European origin in Europe are integrated and assimilated in a single generation, with the best example being the eight million French citizens of Italian, Spanish, Portuguese, Polish, or Armenian origins. The children of these European immigrants are not normally in a position of status inconsistency. When language is combined with ethnicity and religion, as with immigrants from the southern rim of the Mediterranean, the integration process takes two generations and the younger generation often experiences status incongruence. When skin color is considered, the difficulties of integration are compounded. Many immigrants from southern Asia and Africa feel excluded from the host society. Nevertheless, a substantial minority are economically well integrated, and many climb the income ladder. They are ethnic achievers, more than completely assimilated immigrants. They are deeply rooted in status incongruences.

In Europe, these two varieties of status incongruence contrast with a social category of status crystallization at the bottom of society. According to a recent survey by the Organization for Economic Cooperation and Development (OECD), almost one-fourth of the adult population in Western advanced societies is functionally semi-illiterate and coexists with a high proportion of functionally overeducated younger adults. Strong status crystallization arises from the fact that these semi-illiterates are also those who receive the lowest salaries and perform the most menial work, and the large majority of them are of non-European origin. The status crystallization that occurs in

Europe has a similar and more deeply ingrained counterpart in the United States.

### MINORITY STATUS AND STATUS INCONGRUENCE

In many studies of electoral behavior (which are preferred because of the availability of statistics), particularly those conducted by means of survey research, the issue of social context has been neglected. Only the characteristics of individuals are taken into consideration, while the parameters of the social milieu are ignored. With some notable exceptions, too many sociologists have forgotten that the behavior of people is conditioned by their social context. This mistake has been denounced by the German sociologist Scheuch (1969) as the "individualistic fallacy," a complement to the "ecological fallacy." The direct consequence of "individualistic research" is the dismissal of the notion of a "minority" in spite of the fact that its importance has been demonstrated repeatedly. Examples are found in the contrasting behaviors of the same "unidimensional category," whether they are a frustrated minority or a dominant majority. Examples include Irish Catholics who vote for the leftist party in Britain versus "good" Catholics who vote conservative in France or Catholics in Germany who vote for the Christian democratic party and industrial workers who live in densely populated working-class areas versus the same kind of workers who live in middle-class districts. These notions of "minority context" and "majoritarian context" are directly related to the issue of status incongruence, because in many places minorities live in a more or less hostile environment. In such cases, three concepts are involved: status incongruence, minority complex, and cross-pressure.

Contradictory propositions have been suggested concerning the political effects of upward and downward mobility in terms of status incongruence. According to some scholars, upward mobility favors a conservative orientation, and downward mobility a liberal-leftist tendency. Others scholars have arrived at the opposite conclusion. This confusing situation can be explained by the neglect of the social context by those who extrapolate at the national level the results obtained at the local level. Most studies of status incongruence have been conducted in individual

cities (including Lenski's 1954 and 1966 studies). It is misleading to generalize from a series of local monographs that do not represent a truly national sample: "Consistency theory seeks to show that predictable effects result from the combination or interaction of statuses, and that these effects differ from the effects of several independent variables" (Rossides 1976, p. 87). However, in practice it is difficult to weigh the importance of each variable in the social context. In one case, it may be a question of race; in another, income; and in still another, professional position. Extrapolated at national level, these variables conceal important variations across local social contexts.

### STATUS INCONGRUENCE AND INDIVIDUALISM

Most frequently, status inconsistency refers to individuals, not to collectivities. Incongruence of status is a characteristic of a relationship between individuals. When an individual cannot raise the lower factors of the incongruence, he or she tends to avoid people who react to them (Malewski 1963, p. 306). He or she makes an individual move. If an individual can raise the lower factor, "he has a natural tendency to think of himself in terms of that status or rank which is highest, and to expect others to do the same, [but] others, who come in contact with him, have a vested interest in doing just the opposite, that is, in treating him in terms of the lowest status or rank" (Lenski 1966, p. 87). Even in this case, the relationship is between individuals. Vertical mobility separates ascending individuals from nonmobile peers who remain in their status of origin. A high rate of individual upward mobility breaks the unity of the social class by effectively promoting certain people and generating in the minds of others expectations of moving out of the class and into a better one. As Dahrendorf has noted, a high rate of upward mobility favors individualism to the detriment of class consciousness.

However, high rates of downward mobility may have the opposite effect, favoring, as Marx emphasized, the spread of class consciousness. In that case, the tendency is not to leave the group but to identify oneself with others in the same situation of incongruence of status. In some social contexts that aggregate individuals, such as large factories, mines, railways, working-class suburbs, and ghettos in large cities, the phenomenon of

individual status incongruence blooms into a collective social consciousness and a "minority complex."

### CONFIGURATIONS OF STATUS INCONGRUENCES

The amount of status inconsistencies depends on the configuration of three dimensions that may be dichotomized for analytic purposes.

1. Culturally homogeneous societies versus heterogeneous societies. In recent decades, immigration in Western countries has differed from that of former times. In most cases, immigrants coming from Western Europe to the United States and Canada require only two generations for complete assimilation into the dominant culture. More recent immigrants in Western countries came from the southern rim of the Mediterranean and Africa. Not only are their distinctive characteristics not only religious and linguistic, they also differ in skin color. Their integration requests more than two generations, and many of them manifest a preference for multiculturalism, that is, for a recognition and institutionalization of ethnic diversity. Such diversity is currently a source of status inconsistencies but may have different effects in the future.
2. Segmented versus fluid societies. Heterogeneous countries may be segmented or fluid. Segmented societies are divided into religious or linguistic communities, as in Belgium and Northern Ireland, or into "pillars," as in the Netherlands until the middle of 1980s (Lijphart 1977). In these societies, there is little room for ethnic status inconsistency. By contrast, in fluid societies, the crossing of vertical and transversal cleavages is relatively common and generates incongruences.
3. High versus low vertical mobility. Another dichotomy is related to the amount of vertical social mobility on the economic scale, which may be relatively high or relatively low. The fact that high vertical mobility, either upward or downward, increases the frequency of incongruence of statuses is well established.

These three factors have a cumulative effect on the proportion of people who experience incongruence of status.

### STATUS INCONGRUENCE AT THE ELITE LEVEL

What is missing in Pareto's "circulation of elites" is the concept of status incongruence. This is surprising in the writings of someone who emphasized the importance of upward and downward social mobility. If the concept of status incongruence was applied to the highest levels of society, elite studies would be enhanced. The psychological portrait of some of the world's most famous painters could be better understood in the light of status inconsistency. The biographies of masters such as Michelangelo, Bellini, Bosch, Goya, van Gogh, and Toulouse Lautrec could be enriched by an interpretation in terms of status incongruence. Many novelists, including Dostoevsky, Tolstoy, Stendhal, Balzac, de Lampedusa, Proust, and Dumas, have analyzed the psychological aspects of status inconsistency even if they have not used that sociological term. One of the main themes of *The Red and the Black* and *The Leopard* is status inconsistency. The most common case is that of the rich man's daughter who becomes enamored of a young man of lower status. No sociologist has explored the hundreds of cases of status incongruence described by famous writers, starting with Shakespeare's *Romeo and Juliet*.

The concept of status incongruence should be applied even to saints. The best analyses of the personality of the evangelist Paul have been written by theologians and religious historians, who have used the notion of status inconsistency implicitly. The subtitle of Dieter Hilbrand's *Saul-Paul: A Double Life* is significant. Baslez insists on the status incongruence of Saint Paul: Born as a Roman citizen but at the periphery, in Syria; he was a stranger in Ephesus; a polyglot Jew, an apostate, and the son of a Pharisee, he was rejected as a missionary in many communities. Paul accumulated many incongruences. Moses, as the nephew of the pharaoh, and Muhammad, as the poor husband of a rich wife, are examples of status inconsistencies.

The use of the concept of status incongruence is appropriate for a better understanding of political leaders from Spartacus to Robespierre and



from Trotsky to Castro. There are numerous examples of the status incongruence of athletes, clergymen, businessmen, politicians: poets, and movie stars, but this notion has been insufficiently used to explain the metamorphosis of labor leaders. The concept could even be applied to sociologists for a better understanding of the theories and motivations of scholars such as Pareto, Michels, Veblen, Sorokin, Mills, and Lazarsfeld.

### THE RELEVANCE OF STATUS INCONGRUENCE TODAY

The incidence of status incongruence in advanced societies today is many times higher than it was in earlier generations. This upsurge is a result of increasing upward and downward economic mobility, the increasing ethnic heterogeneity of Western societies (as a consequence of massive non-European immigration), and a better perception of inequalities and the spread of "multiculturalism" as opposed to the doctrine of the melting pot, particularly among the so-called second generation, which is composed of the sons and daughters of immigrants.

Four decades ago, status incongruence was usually a question of an imbalance between education, income, occupation, religion, and gender. Today it originates primarily in ethnic and racial intermingling. Religious differences have become less prominent.

In most Western societies on both sides of the Atlantic, a homogeneous majority no longer exists. Any conceivable majority is necessarily composed of multiple minorities of all kinds. An advanced society is a multidimensional society that includes many parallel hierarchies. The political game consists precisely in building coalitions of minorities to crystallize a temporary and unstable political-electoral majority. In almost all these countries, the leftist party has become the party of amalgamated minorities, of those who experience frustrations generated by status incongruences and the psychological complex of belonging to a minority. In the United States, the electorate of the Democratic Party is much more ethnically heterogeneous than is its adversary. It is a conglomerate of minorities. In France, the leftist coalition has officially adopted the label "plural majority." Without the concept of status incongruence, it would be difficult to explain its electoral success.

Projections of demographic trends suggest that Western societies are becoming increasingly diversified along a noneconomic axis and that the amount of status incongruence nourished by ethnic and racial characteristics will increase.

A mountain of statistics has been collected showing that objective inequality and social consciousness explain only a relatively small part of the variance in studies of social stratification. What must be added is an interpretation in terms of status congruence-incongruence.

#### REFERENCES

- Barber, Bernard 1957 *Social Stratification*. New York: Harcourt Brace.
- Bendix, Reinhard, and Seymour M. Lipset, eds. 1966 *Class, Status, and Power*. Free Press.
- Bourdieu, Pierre 1978 "Classement, Déclassement, Reclassement." *Actes de la Recherche in Sciences Sociales* 24:2-22.
- Clark, T. N., and S. M. Lipset 1991 "Are Social Classes Dying?" *International Sociology* 6(4):397-410.
- Dogan, Mattei 1995 "Erosion of Class Voting and of the Religious Vote in Western Europe." *International Social Science Journal* 146:525-538.
- 1999 "Marginality." In *Encyclopedia of Creativity*, vol. I. London: Academic Press.
- Esping-Andersen, Costa 1992 "Post-Industrial Class Structures: An Analytical Framework," working paper, Madrid: Juan March Institute.
- Goldthorpe, John H. 1996 "Class Analysis and the Reorientation of Class Theory." *British Journal of Sociology* 47(3):481-505.
- Feagin, Joe 1997 "The Future of U.S. Society in the Era of Racism, Group Segregation and Demographic Revolution." *The Heritage and Future of Sociology in the North American Region*. International Sociological Association.
- Lenski Gerhard E. 1954 "Status Crystallization: A Non-Vertical Dimension of Social Status." *American Sociological Review* 19:405-413.
- 1966 *Power and Privilege*. New York: McGraw-Hill.
- Lijphart, Arend 1977 *Democracy in Plural Societies*. New Haven, Conn.: Yale University Press.
- Lipset, S. M., and R. Bendix 1959 *Social Mobility in Industrial Society*. University of California Press.
- and H. Zetterberg 1956 "A Theory of Social Mobility." In Reinhard Bendix and Seymour M. Lipset, eds., *Class, Status, and Power*. New York: Free Press.

- Lopreato, Joseph, and Lawrence Hazelrigg 1972 *Class, Conflict, and Mobility*. San Francisco: Chandler.
- Malewski, Andrej 1963 "The Degree of Status Incongruence and its Effects." In Reinhard Bendix and Seymour M. Lipset, eds., *Class, Status, and Power*. New York: Free Press.
- Paugam, S. 1994 *La Disqualification Sociale: Essai sur la Nouvelle Pauvreté*. Paris: Presses Universitaires de France.
- Rossides, Daniel 1976 *The American Class System*. Boston: Houghton-Mifflin.
- Scheuch, Erwin 1969 "Social Context and Individual Behavior." in M. Dogan and S. Rokkan, eds., *Quantitative Ecological Analysis in the Social Sciences*. Cambridge, Mass.: M. I. T. Press.
- Thelat, Claude 1982 *Tel Père, Tel Fils: Position Sociale et Origine Familiale*. Paris: Dunod.
- Turner, Frederick 1992 *Social Mobility and Political Attitudes*. New Brunswick: Transaction Publishers.
- Wilenski, Harold L., and Hugh Edwards 1959 "The Skidders: Ideological Adjustments of Downwardly Mobile Workers." *American Sociological Review* 24:215-231.

MATTEI DOGAN

## STEREOTYPES

See Attitudes; Discrimination; Prejudice.

## STRESS

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

The theoretical interest in social epidemiology, the study of effects of social conditions on the diffusion of distress and diseases in the population, can be traced to Durkheim's study of suicide in 1897 (1951). Since then, theory and research have elaborated on the associations among the various forms of social integration and psychiatric disorder. Among the classic works are Faris and Dunham's study of the ecology of mental disorders in urban areas (1939), Hollingshead and Redlick's research on social class and mental illness in New Haven (1958), the midtown Manhattan studies (Srole et al. 1962; Langner and Michael

1962; Srole 1975), the Sterling County studies by the Leightons and their colleagues (A. H. Leighton 1959; C. C. Hughes et al. 1960; D. Leighton et al. 1963) and the British studies by Brown and his associates (Brown and Harris, 1978). Each study illuminates the linkage between social conditions and distress and advances theories, hypotheses and empirical evidence in the specification of the relationships.

A parallel theoretical development has also taken place, over the past thirty-five years, in the formulation of the life stress paradigm in social psychiatry. The birth of this paradigm can be dated to the work of Hans Selye (1956) whose study of the undifferentiated response (physiological and psychological) that is generated by diverse external stimuli (stressors) linked sociological constructs to the internal individualistic responses made by individuals to their environment. This stress-distress model provided impetus for a convergence between the earlier sociological concerns with consequences of social integration and the physiological modeling of internal responses to the external environment.

The stress research enterprise gained further momentum when Holmes and Rahe, and subsequently other researchers, developed measures of life experiences that require social adjustments, known as inventories of life events (Holmes and Rahe 1967; Dohrenwend and Dohrenwend 1974, 1981; Myers and Pepper 1972). The life events schedules provide a convenient instrument that can be applied to a wide range of populations and administered with ease. The instrument has shown a high degree of validity and reliability relative to many measures of distress across populations and time lags.

In general, the research shows that life stressors, as measured by the life events schedules, exert a significant but moderate influence on mental and physical well-being. In a simple zero-order correlation, the relationship between life stressors and well-being (e.g., depressive symptoms) ranges between .25 and .40 (Rabkin and Struening 1976). This figure is somewhat less for physical health (House 1981; Wallston et al. 1987; Ensel 1986). The magnitude of this relationship seems to hold up when other factors are taken into account (e.g., general socioeconomic status measures; age; gender; psychological resources such as self-esteem,

personal competence, and locus of control; physical health; and prior mental state).

### **MODIFICATIONS AND EXTENSIONS—THE MEDIATION PROCESSES**

Modifications of the stressors-distress paradigm have taken several directions. In one direction, the conceptualization of stress as undifferentiated response has been modified so that the nature of stressors entails further specification. For example, in the analysis of life events, desirability, controllability, and importance are identified as dimensions exerting differential effects on distress (Thoits 1981; Tausig 1986). Research has shown that when only self-perceived undesirable life events are considered, the effect of the stressor instrument on distress increases marginally but significantly. It has also been shown that when items pertaining to psychological states (sleeping and eating problems) or illnesses are deleted, the magnitude of its effect is only marginally reduced (Ensel and Tausig 1982; Tausig 1982, 1986).

Conceptualization and operationalization of stressors have also been extended to include role strains (Pearlin and Schooler 1978) and daily hassles (Lazarus and Folkman 1984). Generally speaking, these stressors have demonstrated consistent but moderate effects on mental health, with zero-order correlations with various measures of mental health ranging from .15 to .35.

Another direction focuses attention on factors mediating or buffering the stressors-distress relationship. Researchers have identified three major components involved in the stress process: stressors, mediating factors, and outcome variables. Pearlin et al. (1981) viewed these constructs as multifaceted. Mediators consist of both external coping resources (i.e., social support) and internal coping resources (i.e., mastery and self-esteem). Outcome factors consist of psychological and physical symptomatology.

Social support, for example, has been considered a major candidate variable, and the cumulative evidence is that it exerts both direct and indirect effects on mental health (Cobb 1976; Cassel 1974, 1976; Nuckolls, Cassel, and Kaplan 1972; Dean and Lin 1977; Lin et al 1979; Turner 1981; Barrera and Ainlay 1983; Aneshensel and Huba 1984; Sarason and Sarason 1985; Kessler

and McLeod 1985; Lin, Dean, and Ensel 1986; Berkman 1985; Cohen and Wills 1985; House, Umberson, and Landis 1988). Coping has also received substantial research attention and been found to be an effective mediator (Pearlin et al, 1981; Wheaton 1983, 1989; Lazarus and Folkman 1984). This type of research has served as the prototype for the sociopsychological study of stress in the 1980s (Pearlin, 1989). Emphasis has been placed on the mechanisms by which social resources, provided or called upon in the presence of a stressor, operate to alter the effect of the stressor (House, Umberson, and Landis 1988; Kessler, Price, and Wortman 1985; Thoits 1985).

### **DEVELOPMENT OF INTEGRATIVE AND TIME-LAGGED MODELS**

While conceptual analysis and research attention have been given to life stress, resources (social support and coping), and psychological stress for their potential effects on health and mental health, only recently have specific proposals emerged in integrating these elements into a coherent theoretical framework. Dohrenwend and Dohrenwend (1981) summarized various formulations of life stress processes involving stressors (life events) and the psychological and social contexts in which they occurred. These formulations were synthesized into six hypotheses, each of which was shown to provide viable conceptual linkages between stressors (life events) and health outcomes and to have received some empirical support. The hypotheses in these models share two common features: (1) The ultimate dependent variable is adverse health or adverse health change rather than mental health problems or disorders, and (2) each hypothesis delineates and explains the possible empirical association between life events and health. Some of the hypotheses affirm the primary role of life events as causing health problems, while others incorporate mediating factors to explain health problems. The Dohrenwends (1974) proposed that these hypotheses should be examined together for their relative merits. Golden and Dohrenwend (1981) outlined the analytic requirements for testing these causal hypotheses.

Further elaboration of these hypotheses formed the basis of an integrative life stress paradigm in which stressors and resources in three environments—social, psychological, and physiological—

are considered as the factors impinging on well-being (Lin and Ensel 1989). This model specifies the enhancing (resources) and detrimental (stressing) forces in each environment. These stressors and resources in the three environments interact in affecting one's physical and mental health. Empirical evidence suggests that social resources tend to mediate the stress process involving mental health, whereas psychological resources are more prominent in mediating the process involving physical health.

Another integrative attempt incorporates multidisciplinary and multilevel variables in the study of life stress. For example, Lazarus and Folkman (1984) and Trumbull and Appley (1986) have conceptualized cognitive mechanisms involved in the stress process. Lazarus and Folkman proposed a model in which three levels of analysis (social, psychological, and physiological) are conducted to understand the antecedent, mediating, and immediate as well as long-term effects on distress. Trumbull and Appley (1986) proposed the simultaneous assessment of the physiological system, psychological system, and social system functioning. These functionings have both intrasystem and intersystem reciprocal relationships and exert joint effects on distress. In the later paradigms, emphasis has been placed on personality factors and coping skills. Additionally, the importance of linking social, psychological, and physical factors in the study of the stress process has been noted. Causal antecedents of both depressive and physical symptomatology are viewed as coming from social, psychological, and physiological sources and are hypothesized to be mediated by a variety of coping factors and perceived social support.

Pearlin and Aneshensel have proposed a synthesized paradigm (Pearlin 1989; Pearlin and Aneshensel 1986) in which health behaviors and illness behaviors have been incorporated into the basic stress process and in which equal attention has been given to the potential mediating and moderating roles of social and psychological resources. Thus, in addition to mediating the effect of stressors on illness outcomes, coping and social support are viewed as having the potential to mediate health and illness behaviors. An important element of this synthesizing paradigm is the recognition that physical illness creates life problems that are reflected in an increase in undesir-

able life events—that is, in addition to stressors affecting physical illnesses, physical illness also has the potential to bring about the occurrence of stressors. In such a synthesized paradigm, stressors embedded in social structure (e.g., role strains and problems) interact with illness behavior and illnesses. These interactions are mediated by coping and social support.

Finally, growing attention has been given to the need for studying the stress process over time (Wheaton 1989). Not only have there been concerns with causal interpretations of cross-sectional data, but more importantly, a call for longer lags in the panel design to capture the stress process in the life course more realistically (Thoits 1982). Some of the earlier panel studies, such as the midtown Manhattan study (Srole and Fischer 1978), the Kansas City study (Pearlin et al. 1981), the New Haven study (Myers, Lindenthal, and Pepper 1975), and the Cleveland GAO study (Haug and Folmar 1986) have all made significant contributions to understanding the stress process in urban communities. More current efforts, incorporating prevailing models and variables, would substantially add to the knowledge about stress in the life course. Current panel studies, such as those mounted by Aneshensel in Southern California; House and his associates on a national sample; Berkman in New Haven; Murrell in Kentucky; and Lin, Dean, and Ensel in upstate New York have the potential to expand research programs into investigations of the life-course process of stress.

(SEE ALSO: *Mental Illness and Mental Disorders; Personality Theories*)

#### REFERENCES

- Aneshensel, Carol S. 1992 "Social Stress: Theory and Research." *Annual Review of Sociology* 18:15–38.
- Aneshensel, C. S., and G.J. Huba 1984 "An Integrative Causal Model of the Antecedents and Consequences of Depression over One Year." In James R. Greenley, ed., *Research in Community and Mental Health*. Greenwich, Conn.: JAI Press.
- Avison, William R. and Ian H Gotlib (eds.) 1994 *Stress and Mental Health: Contemporary Issues and Prospects for the Future*. New York: Plenum Press.
- Barrera, M., and S. L. Ainlay 1983 "The Structure of Social Support: A Conceptual and Empirical Analysis." *Journal of Community Psychology* 11:133–143.

- Berkman 1985 "The Relationship of Social Networks and Social Support to Morbidity and Mortality." In S. Cohen and S. L. Syme, eds., *Social Support and Health*. New York: Academic Press.
- Brown, G. W., and T. Harris 1978 *Social Origins of Depression: A Study of Psychiatric Disorder in Women*. New York: The Free Press.
- Cassel, J. 1974 "An Epidemiological Perspective of Psychosocial Factors in Disease Etiology." *American Journal of Public Health* 64:1040-1043.
- 1976 "The Contribution of the Social Environment to Host Resistance." *American Journal of Epidemiology* 104:107-123.
- Cobb, S. 1976 "Social Support as a Moderator of Life Stress." *Psychosomatic Medicine* 38:300-314.
- Cohen, S., and T. A. Wills 1985 "Stress, Social Support, and the Buffering Hypothesis." *Psychological Bulletin* 98(2):310-357
- Coyne, James C. and Geraldine Downey 1991 "Social Factors and Psychopathology: Stress, Social Support, and Coping Processes." *Annual Review of Psychology* 42:401-425.
- Dean, Alfred, and Nan Lin 1977 "The Stress Buffering Role of Social Support." *Journal of Nervous and Mental Disease* 165(2):403-13.
- Dohrenwend, B. S., and B. P. Dohrenwend 1974 *Stressful Life Events: Their Nature and Effect*. New York: Wiley.
- 1981 "Life Stress and Illness: Formulation of the Issues." In B. S. Dohrenwend and B. P. Dohrenwend, eds., *Stressful Life Events: Their Nature and Effects*. New York: Prodist.
- Durkheim, Emile 1951 *Suicide*. Glencoe, Ill.: The Free Press.
- Ensel, Walter M. 1986 "Measuring Depression: The CES-D scale." In Nan Lin, Alfred Dean, and Walter M. Ensel, eds., *Social Support, Life Events, and Depression*. Orlando, Fla.: Academic Press.
- , and Mark Tausig 1982 "The Social Context of Undesirable Life Events." Presented October 11-12 at the National Conference on Social Stress, Durham, N.H.
- and Nan Lin 1991 "The Life Stress Paradigm and Psychological Distress." *Journal of Health and Social Behavior* 32:321-341.
- Faris, Robert E. K., and H. Warren Dunham 1939 *Mental Disorders in Urban Areas*. Chicago: University of Chicago Press.
- Fernandez, Maria E., Elizabeth J. Mutran, and Donald C. Reitzes 1998 "Moderating the Effects of Stress on Depressive Symptoms." *Research on Aging* 20:163-182.
- George, Linda K. 1993 "Sociological Perspectives on Life Transitions." *Annual Review of Sociology* 19:353-373.
- Golden, R. R., and B. S. Dohrenwend 1981 "Testing Hypotheses about the Life Stress Process: A Path Analytic Method for Testing Causal Hypotheses." In B. S. Dohrenwend and B. P. Dohrenwend, eds., *Stressful Life Events: Their Nature and Effects*. NY: Prodist.
- Gotlib, Ian H. and Blair Wheaton, eds. 1997 *Stress and Adversity over the Life Course: Trajectories and Turning Points*. Cambridge: Cambridge University Press
- Haug, M. R., and S. J. Folmar 1986 "Longevity, Gender, and Life Quality." *Journal of Health and Social Behavior* 27:332-346.
- Hollingshead, August, and Fredrick Redlick 1958 *Social Class and Mental Illness*. New York: Wiley.
- Holmes, T., and R. Rahe 1967 "The Social Readjustment Rating Scale." *Journal of Psychosomatic Research* 11:213-218.
- House, James S. 1981 *Work Stress and Social Support*. Reading, Mass.: Addison-Wesley.
- , Karl R. Landis, and Debra Umberson 1988 "Social Relationships and Health." *Science* 241 (July 29):540-545.
- 1988 "Structures and Processes of Social Support." *Annual Review of Sociology* 14:293-318.
- Hughes, C. C., M. A. Tremblay, et al. 1960 *People of Cove and Woodlot*, vol. 2 of the Sterling County Study. New York: Basic Books.
- Kessler, R. C., and J. McLeod 1985 "Sex Differences in Vulnerability to Undesirable Life Events." *American Sociological Review* 49 (5):620-631.
- Kessler, R. C., R. H. Price, and C. B. Wortman 1985 "Social Factors in Psychopathology: Stress, Social Support, and Coping Processes." *Annual Review of Psychology* 36:531-572.
- Langner, T. S., and S. T. Michael 1962 *Life Stress and Mental Health*. New York: The Free Press.
- Lazarus, Richard S. 1991 "Psychological Stress in the Workplace." *Journal of Social Behavior and Personality* 7:1-13.
- , and S. Folkman 1984 *Stress, Appraisal, and Coping*. New York: Springer.
- Leighton, A. H. 1959 *My Name Is Legion*. New York: Basic Books.
- Leighton, D. C., et al. 1963 *The Character of Danger*. New York: Basic Books.
- Lin, Nan, Alfred Dean, and Walter M. Ensel 1986 *Social Support, Life Events, and Depression*. Orlando, Fla.: Academic Press.

- Lin, Nan, and Walter M. Ensel 1989 "Life Stress and Health: Stressors and Resources." *American Sociological Review* 54:382-399.
- Lin, Nan, Ronald Simeone, Walter M. Ensel, and Wen Kuo 1979 "Social Support, Stressful Life Events, and Illness: A Model and an Empirical Test." *Journal of Health and Social Behavior* 20 (1):108-119.
- Myers, J. K., and M.P. Pepper 1972 "Life Events and Mental Status: A Longitudinal Study." *Journal of Health and Social Behavior* 13:398-406.
- Myers, J. K., J. J. Lindenthal, and M. P. Pepper 1975 "Life Events, Social Integration, and Psychiatric Symptomatology." *Journal of Health and Social Behavior* 16:421-429.
- Nuckolls, C. G., J. Cassel, and B. H. Kaplan 1972 "Psychosocial Assets, Life Crises, and the Prognosis of Pregnancy." *American Journal of Epidemiology* 95:431-441.
- Pearlin, L. I. 1989 "The Sociological Study of Stress." *Journal of Health and Social Behavior* 30:241-256.
- , and C. Aneshensel 1986 "Coping and Social Supports: Their Function and Applications." In L. Aiken and D. Mechanic, eds., *Applications of Social Science in Clinical Medicine and Health*. New Brunswick, N.J.: Rutgers University Press.
- , M. A. Lieberman, E.G. Menaghan, and J. T. Mullan 1981 "The Stress Process." *Journal of Health and Social Behavior* 22:337-356.
- , and C. Schooler 1978 "The Structure of Coping." *Journal of Health and Social Behavior* 19 (1):2-21.
- Rabkin, J. G., and E. L. Struening 1976 "Life Events, Stress, and Illness." *Science* 194:1013-1020.
- Sarason, I. G., and B. R. Sarason 1985 *Social Support: Theory, Research, and Application*. The Hague: Martinus-Nijhoff.
- Scheck, Christine L., Angelo J. Kinicki, and Jeannette A. Davy 1995 "A Longitudinal Study of a Multivariate Model of the Stress Process Using Structural Equations Modeling." *Human Relations* 48:1481-1510.
- Seyle, Hans 1956 *The Stress of Life*. New York: McGraw-Hill.
- Srole, L. 1975. "Measurements and Classification in Sociopsychiatric Epidemiology: Midtown Manhattan Study I (1954) and Midtown Manhattan Restudy II (1974)." *Journal of Health and Social Behavior* 16: 347-364.
- , T. S. Langner, S. T. Michael, et al. 1962 *The Midtown Manhattan Study*. New York: McGraw-Hill.
- Tausig, Mark 1982 "Measuring Life Events." *Journal of Health and Social Behavior* 23 (March):52-64.
- 1986 "Measuring Life Events." In Nan Lin, Alfred Dean, and Walter M. Ensel, eds., *Social Support, Life Events, and Depression*. Orlando, Fla.: Academic Press.
- Thoits, Peggy A. 1981 "Undesirable Life Events and Psychophysiological Distress: A Problem of Operational Confounding." *American Sociological Review* 46 (1):97-109.
- 1982 "Conceptual, Methodological, and Theoretical Problems in Studying Social Support as a Buffer Against Life Stress." *Journal of Health and Social Behavior* 24:145-159.
- 1985 "Social Support Processes and Psychological Well-being: Theoretical Possibilities." In I. G. Sarason and B. R. Sarason, eds., *Social Support: Theory, Research, and Application*. The Hague: Martinus-Nijhoff.
- 1995 "Stress, Coping, and Social Support Processes: Where Are We? What Next?" *Journal of Health and Social Behavior*, extra issue: 53-79.
- Tijhuis, M. A. R., H. D. Flap, M. Foets, and P. P. Groenewegen 1995, "Social Support and Stressful Events in Two Dimensions: Life Events and Illness as an Event." *Social Science and Medicine* 40:1513-1526.
- Trumbull, R., and M. H. Appley 1986 "A Conceptual Model for the Examination of Stress Dynamics." In M. H. Appley and R. Trumbull, eds., *Dynamics of Stress: Physiological, Psychological, and Social Perspectives*. New York: Plenum.
- Turner, R. J. 1981 "Social Support as a Contingency in Psychological Well-being." *Journal of Health and Social Behavior* 22:357-367.
- Uhlenhuth, E. H., et al. 1982 "Symptom Checklist Syndromes in the General Population: Correlations with Psychotherapeutic Drug Use." *Archives of General Psychiatry* 40:1167-1173.
- Unger, Jennifer B., C. Anderson Johnson, and Gary Marks 1997 "Functional Decline in the Elderly: Evidence for Direct and Stress-Buffering Protective Effects of Social Interactions and Physical Activity." *Annals of Behavioral Medicine* 19:152-160.
- Wallston, B. S., et al. 1987 "Social Support and Physical Health." *Health Psychology* 2 (4):367-391.
- Wheaton, B. 1983 "Stress, Personal Coping Resources, and Psychiatric Symptoms: An Investigation of Interactive Models." *Journal of Health and Social Behavior* 24:208-229.
- 1989 "Life Transitions, Role Histories, and Mental Health." *American Sociological Review* 2:209-223.

NAN LIN

## STRIKES

*See* Labor Movements and Unions; Industrial Sociology.

## STRUCTURAL EQUATIONS

*See* Causal Inference Models; Correlation and Regression Analysis; Multiple Indicator Models.

## STRUCTURAL LAG

The concept of structural lag originally was suggested by the observation that in the late twentieth century there was a discrepancy between the growing number of older healthy people and the meaningful roles available to them. This simple empirical observation is only one instance of a more general phenomenon: a mismatch between the numbers and kinds of people of a given age and existing patterns in the social structures into which people must fit. This mismatch occurs because changes in people's lives and changes in social structures typically are not synchronic. When social structures fail to adapt to new cohorts with characteristics different from those of previous cohorts, there is a situation of structural lag (Riley et al. 1994).

### PREMISES ABOUT AGE AND SOCIETY

How and why structural lags emerge and how they are dealt with can be better understood by considering the underlying principles of age as both an individual and a social phenomenon and its relationship to social change.

1. Age not only is a characteristic of people but is built into social structures in the form of criteria for entering or leaving social roles, expectations about behavior in those roles, and resources and rewards for role performance. Formal and informal rules govern the age at which children enter and stay in school, age patterns in the family such as the appropriate age at first marriage, and age of entry into and retirement from the work force. Age norms influence behavior and orientations in these roles, and conformity is buttressed by material rewards and social approval.
2. Both the process of aging from birth to death and social structures related to age are subject to change. The aging process is not the same for all cohorts, since members of each cohort (those born in the same period) grow up and grow older under unique social, political, and environmental circumstances. For example, cohorts of people born around the beginning of the twentieth century differ from those born a half century later in level of education, exposure to illness, size of the family, job skills, the likelihood of being married and divorced, and attitudes and worldviews. Cohorts born at the end of the twentieth century will differ from their predecessors in still other ways. For example, the birth weight of newborns in the 1980s and 1990s was greater than that of babies born in earlier times, and this undoubtedly will influence the way that members of those cohorts develop. Further, as the more recent cohorts grow up, it is likely that they will benefit from new medical advances. Age patterns in social structures have changed as well. To cite two examples, schools have raised the school-leaving age, and government and corporate policies have encouraged a decline in the typical age of retirement.
3. Changes in patterns of aging and in social structures affect each other; they are interdependent. As an example, by altering long-standing employment practices, restructuring, downsizing, and mergers of large firms in the United States have led to a shift in the work lives of employees. Thus, compared to previous cohorts of workers, fewer workers now can look forward to lifetime employment in one firm; many have to make multiple career changes, and some have to make do with temporary and part-time employment at some time in their working lives. In turn, new patterns of careers over the life course are likely to have an impact on societal institutions. Firms are likely to reduce their commitment to training workers, with educational institutions as-

suming greater responsibility for training and retraining.

While changing aging processes and social structures affect each other, the two processes of change are distinct, with each following its own dynamics. The aging process, while varying across cohorts, has a distinct rhythm, as people are born and then proceed through childhood, adolescence, adulthood, and old age. Social structures do not have a similar rhythm. The economy has ups and downs; political shifts follow their own paths; and cataclysmic events such as wars, depressions, famines, and epidemics affect all social institutions. Consequently, in any period there is likely to be a poor fit between lives and structures: an imbalance between what people of given ages have to offer, what they need and expect in their lives, and their motivations versus what social structures can accommodate or demand.

4. Lags can occur in either direction. Sometimes people's lives lag behind changes in social structures. For example, many older people may be reluctant to learn and use new technologies, or adolescents may not be motivated to take the science courses that will prepare them for technological changes. At the end of the twentieth century, however, a key form of imbalance is the lag of social structures behind changes in people's lives. As the examples discussed below indicate, structural lag is pervasive, affecting people of all ages and many social institutions.

#### STRUCTURAL LAG AT THE MILLENNIUM

A focus on the fit or misfit between people and structures in three major age strata shows how these principles play themselves out and how structural lags have emerged.

**The Old.** The increase in the number and proportion of older people in the twentieth century is a dramatic example of a change in people's lives that has posed numerous problems for societal institutions. By the end of the century people 65 years old and over represented 13 percent of the U.S. population compared to about 4 percent at the beginning of the century. More than 70

percent of Americans now live to age 65, almost three times the proportion at the beginning of the century. These changes are in large part the result of public and private health care measures that reduced infant mortality decades ago, increasing the proportion of individuals who could survive to the later years. Reductions in infant and child mortality were followed by reductions in death rates among older people, partly as a result of public health and scientific and medical innovations and partly because of healthy practices in diet and exercise undertaken by individuals. Indeed, old people at the end of the twentieth century are a relatively healthy lot. Most report that they have no disabilities; even among those over age 85, about 40 percent report being able to function in daily life (Rowe and Kahn 1998). Thus, not only are more people growing old, they are aging well.

Social institutions have been slow to accommodate to the needs of this new kind of older population, a lag that represents not only lost opportunities for the old but a loss of the productive capabilities of older people to society. Consider the organization of work and retirement. Although 65 is the age of eligibility for full Social Security benefits, most people in recent years have been retiring before that age. This pattern of early retirement was facilitated by Social Security regulations, devised in an earlier period, that exact little or no cost for retiring before age 65. On their part, many employing organizations, driven by changing personnel requirements, offer financial incentives for early retirement. Also, in the process of restructuring their firms or merging with others, employers let many long-term employees go, many of whom retired early rather than face the uncertainties of the job market. If they are assured of financial security in retirement, some workers welcome the opportunity to retire early, perhaps because of their health or because of the onerous or stultifying nature of their jobs. Nevertheless, surveys find that a sizable proportion of older workers prefer not to sever their ties to the labor force completely (Burkhauser and Quinn 1994). However, few firms permit workers to continue at their old jobs under the more flexible working conditions many workers prefer. Available part-time or temporary jobs typically have few, if any, of the benefits of workers' former employment, and the pay is generally low. The



result is that many older workers withdraw from the labor force completely—often unwillingly. Employing organizations thus lose the benefits that experienced workers can bring to their firms.

Paid employment is not the only way older workers can make productive contributions to society. A sizable minority of older people do volunteer work that has social value—in religious, charitable, and civic organizations, for example. Structural lag and resistance to new ideas and values consistent with a changed society may explain why more older persons do not volunteer. Volunteer organizations often do not have recruiting mechanisms to draw on the large pool of potential older volunteers. With respect to societal values, volunteer work is not accorded the same respect as paid employment (Kahn 1994).

Societal institutions are lagging behind the needs of an unfortunate sector of the older population: those in poor health who need support. Especially among the oldest old, there is a need for long-term care either in the home or in a nursing facility, but affordable arrangements for such care are inadequate. As a result some older people are not getting the care they need, and the burden for caring for them falls on their elderly spouses or their middle-aged offspring or other relatives. In such ways, structural lag in care institutions affects both the old and the middle-aged. Unless there are relevant structural changes in these care institutions and/or government programs to shore them up, there will be problems for the baby boom cohorts when they reach their later years. They will have fewer kin available to provide the needed support, since the baby boom cohorts were followed by relatively small cohorts.

**Children.** Children's lives also have changed dramatically; today they have vastly different growing up experiences than did earlier cohorts. As a consequence, children now differ from their predecessors in attitudes, capabilities, motivations, behaviors, and the choices they make. These dispositions will affect their paths of future development: their school careers, job choices and opportunities, and marriage and family decisions.

One development that has altered the lives of children has been the increase in single-parent families. On average, these families are poorer than two-parent families, and there are long-term consequences for children raised under condi-

tions of poverty. Experiencing poverty as infants and young children (zero to 5 years of age) affects people's subsequent educational achievement and employability (Duncan, et al. 1998). The increase in single-parent families thus does not bode well for the future of their young offspring, many of whom will not be prepared to fill the roles available in a technologically advanced and constantly changing society.

There has also been a marked increase in families where both the father and the mother work outside the home. Unless they have high incomes, dual-earner families, share with single-parent families the problem of finding adequate child care arrangements. A small proportion of mothers cope with both work and taking care of children, and a similarly small percentage of fathers care for children while mothers work away from the home. In some cases, grandparents or other relatives care for children outside the children's home. The well off can afford paid babysitters, and about one-quarter of preschoolers of working mothers are in some form of organized day care. Many of these facilities, however, have been judged unsafe or unsanitary and do not offer a warm and intellectually stimulating environment.

While the long-term outcome of these new socialization environments for infants and children cannot be known yet, social structures outside the family clearly are not filling the gaps created by changed family arrangements. Most important, there has been no institutionalization of satisfactory nonparental child care arrangements; social structures outside the family are lagging behind changes in children's lives and changes in the family.

There is also a gap between the lives of school-age children and social structures. Consider institutions of public education. Among the many undertakings of public schools at the end of the twentieth century, there are two major tasks: educating students raised in changed family environments and preparing those students for a changed society in which people increasingly will need the ability to adapt to continual changes and more jobs will require high levels of conceptual thinking.

By and large, experimental programs and various changes in schools notwithstanding, schools are falling behind in meeting these challenges. For example, many teachers have inadequate knowl-

edge of subject matter; curricula often lag behind new knowledge and are too often shaped by the nature of national or statewide tests; frequently there is administrative inertia and resistance to change; and students often are not challenged sufficiently in terms of, for example, the amount and nature of homework. These patterns militate against the goal of inculcating the kinds of thinking and other skills that will be needed in the next decades, when today's students will enter the labor force and take on adult responsibilities.

**Adults.** Structural lags affecting the old and the young have an impact on people in their middle years. It is those people years who must undertake the care of the young and the old when there are no alternatives, but it is precisely when people are in their prime years that they have heightened job and career responsibilities. Social structures are lagging in providing arrangements—such as flexible hours or flexible workplace settings and respite care from social agencies—that would ease the multiple burdens of those with career and family care responsibilities.

It is not only work organizations and social agencies that are not helping workers undertake their multiple responsibilities: Family structures are not adapting to the new realities either. It is women who typically bear the brunt of multiple burdens of work and family care. While most married women with children work outside the home, they still have the main responsibility for tasks in the home. More husbands—when they are present—are “helping” with household work than was the case in earlier times, but the norm of equal sharing has not been institutionalized. There is a gap between the changed lives of women and the way most families are structured.

**Class, Gender, and Race Differences.** There are differences within the several age strata (layers of people who differ in age and confront structures differentially appropriate for particular ages) in the degree of fit between changed people and changed institutions. The match or mismatch between people of given ages and institutions often depends on the gender, ethnicity, race, or class of the people involved.

Thus, the impact of structural lag is uneven. In some instances it is the most disadvantaged segments of the different age strata that are most likely to feel the brunt of the lag.

### CLOSING THE GAP: PRESSURES FOR CHANGE

The gap between structures and the lives of the individuals in those structures creates tensions, inefficiencies, and other problems that are potent stimuli for change in both people and structures. Whatever the constraints, social structures tend to respond to these forces.

**Social Structural Responses.** Social structural responses to an increase in cohort size—a key source of pressure on structures—are illustrative. Beginning in midcentury, the pressure on social institutions came from the unexpectedly large number of people in the baby boom cohorts, whose large size created a lag in structures at every life stage. As Waring (1976) shows, social institutions coped with these large cohorts in myriad ways. When the baby boom cohorts were newborns, hospitals reduced the typical length of stay of new mothers and their babies to accommodate the flood of new births. When the baby boomers entered school, new schools were built and younger teachers were hired to compensate for the teacher shortage. In their college years, educational requirements were extended, with the result that entry into the labor force was delayed, helping to prevent a labor glut. These changes generally came piecemeal, but many of them, such as shorter maternal hospital stays, have turned out to be long-lasting.

As the number of old people has increased and as the baby boomers approach old age, there are signs that social institutions are making further changes. Indeed, as the changes made in many different institutional settings accumulate, new social meanings of age may arise. Age barriers to entry into a broad range of social roles are being relaxed and even breaking down. For individuals this means increased opportunities to intersperse periods of education, work, family time, and leisure over the life course, unlike the more rigid pattern of education in youth, work in adulthood, and retirement and leisure in old age that has been the typical shape of the life course (Riley et al. 1999). Within institutions, as more roles become available to people of all ages, cross-age interactions are likely to increase. Also, the pool of human capital available to varied social institutions will no longer be limited by rigid age norms.

**Mechanisms of Change.** Changes, whether piecemeal or encompassing, do not come about automatically. A number of processes operate singly or in combination to bring about change.

*Actions of policymakers.* Some individuals and agencies are in a good position to make and implement policy—government officials and company executives, for example—because they have an overview of their organizations and can propose or institute policies that will reduce the gap between people and structures.

One of the portents of more flexible age criteria has been the opening up of colleges and universities to older students. Educational administrators have played an important role in this development. They saw an opportunity to expand the pool of prospective students and felt a responsibility to offer access to their schools to older people. They have devised special degree programs for older people along with a wide range of nondegree classes. On a less formal level, other organizational leaders have developed elder hostel programs that give older people a chance to combine education and recreation, programs that have expanded beyond college settings and beyond the United States.

Undoubtedly, these policy initiatives were influenced by the actions of older people who were seeking avenues for enriching their lives or for filling in gaps in their education. Indeed, people inside social structures often act as agents of change, sometimes engaging in purposeful action with others and sometimes acting independently, as in the case of cohort norm formation, another mechanism of change.

*Cohort norm formation.* As formulated by Riley (1978), cohort norm formation is a process that occurs when the members of a cohort, reacting independently but in like fashion to changes in society, create new patterns of behavior and attitudes. These changes often spread to the succeeding cohorts, contributing to the establishment of new norms.

The centurylong increase in women's labor force participation is a prime example of this process. Since the early decades of the twentieth century, in successive cohorts, increasing proportions of women have worked outside the home. They were, of course, responding to broad social

forces that eased household burdens, facilitated control of family size, and introduced workplace technologies that women could manage efficiently. However, it was individual women who made the decision to enter and remain in the labor force after marriage. As they did so, norms for women's labor force participation changed. Early in the century the typical married woman did not hold a job outside the home and was not expected to, but as increasing numbers of women sought employment, it became acceptable for them to work for pay. At the end of the twentieth century, not only are high proportions of single and married women in the labor force, it is expected that women will have paying jobs regardless of their marital status. This norm was embodied in welfare legislation in the 1990s: Poor women were given time limits for welfare payments to support them and their dependent children. After that deadline, they were expected to be self-supporting.

Among older people at terminal stages of the life course, another set of norms has been forming (Riley 1990). Older people have been pressing for new norms that will help them avoid a prolonged, painful process of dying, provide palliative care, and give them more control of the way they die. New norms for the dying process and arrangements to implement them are being put in place. Hospice care that eschews heroic measures for those near death has been more widely accepted, more people are writing living wills detailing measures to be taken near the end, and hospitals have been forming medical ethics committees to deal with these issues. The "right to die" movement has gained power and, with it, some of the structural changes it supports. However, this movement has taken on a new cast as it has focused more on improving caretaking arrangements than on the psychosocial needs of patients (see "Death and Dying" in this encyclopedia).

What started out as individual but uncoordinated responses of numerous older people and others concerned about unacceptable practices in the American way of dying—an example of cohort norm formation—have taken on the shape of a social movement, a more organized effort to change customary practices.

*Social movements.* Social movements have played a role in effecting many age-related changes. These movements take several forms. They may involve

organized groups exerting pressure for change or may entail collective actions that arise more spontaneously. Whether organized or not, they bring issues of concern to public attention. Sometimes their actions lead to conflicts with groups that have different interests and agendas.

Organized social movements encompassing large segments particular age strata have emerged relatively infrequently. Most noteworthy was the Townsend movement in the 1930s, which organized older people to work for a publicly supported pension program for the elderly in the United States. At its height it had organized groups in almost every state. It played a role in the eventual enactment of Social Security legislation.

Although broad-based movements involving age-related issues are difficult to organize, more limited movements for structural change crop up. In the recent past there has been the "right to die" movement with its shifting emphases, as noted above, and organizations focusing on the problems of older women. At the other end of the age spectrum there is a children's rights movement concerned particularly about neglected children. By bringing problems to public attention and by lobbying policymakers directly, these social movements are able to stimulate at least piecemeal changes.

At times, social movements can trigger conflict among age strata. In the 1980s, for example, some pressure groups attempted to pit younger people against the old with dire predictions of intergenerational conflict in the future. Those groups argued that the elderly receive undue advantages from U.S. government programs and that younger adults are unfairly burdened with supporting those programs. Such age conflicts have not emerged, however, and challenges from younger adults that might provide the impetus for changes in government policies seem unlikely. The cross-age support for Social Security suggests why this is so. Younger adults want to maintain the Social Security program to safeguard their own future as well as to protect their parents' present status. Apart from affection for their parents, self-interest is involved. Without publicly supported pensions, those under age 65 would have to bear an increased financial burden to support their elderly parents. Moreover, many adults under age 65 benefit indirectly from gifts and ultimately

from inheritances made possible by the financial security afforded to the old by Social Security pensions. In short, the inevitability of growing old and the intergenerational bonds and exchanges within the family are powerful deterrents to conflicts between the young and the old (Foner 1974).

**Obstacles to Change.** Changing age-related components of social structures does not necessarily proceed smoothly. Long-held values, institutional rigidities, and the possible costs involved create impediments to effective change.

Proposals to make lifelong education a reality—not only educational opportunities for the old but also time off for retraining and sabbaticals for educational enrichment among those of working age—may seem simple to implement. However, many employers perceive that giving time off for sabbaticals is costly and often see no payoff from retraining mature workers when newly trained and cheaper young workers are available.

More equal sharing of household and child care responsibilities by young and middle-aged married adults can be thwarted by entrenched values about the appropriate roles of men and women.

Giving sick and dying patients increased autonomy in regard to their care does not comport with common practices such as rigid scheduling and beliefs of physicians and authorities in medical institutions that patients do not have the professional expertise needed to deal with their illness.

**Unintended Consequences.** Such obstacles notwithstanding, change does occur, but sometimes it has unforeseen results. For example, the spread of hospices that provide relief and palliative care for near-death patients has been a welcome alternative for people who do not want heroic—and often painful and expensive—measures. However, as the number of hospices and their patients has increased in the United States, federal regulatory agencies have found themselves hard put to monitor them and fraud has increased (New York Times 1998).

New norms about extending the work life will give healthy and eager older workers productive roles while at the same time addressing certain problems of the financial viability of the Social Security system. However, these new norms could undermine the right to retirement, discrediting those who are unable to work or those who need a

rest after long years of toil. Further, an increase in the number of older workers in the labor market looking for good part-time jobs could lead to competition with young and female workers also looking for such jobs. One result could be depressed wages in this sector of the labor market.

Viewed in their particulars, changes in social structures do not always neatly adjust social institutions to the changed lives of people. The change may be only partial, some changes may work out well for some people but not for others, and new problems may emerge, calling for additional changes. Viewed in the long run, however, structural lag turns out to be a frequently unrecognized but powerful force for change.

### STRUCTURAL LAG AND THEORIES OF CHANGE

The concept of structural lag as a force for change has a ring of familiarity. Many analysts have put forth theories about discrepancies among the several parts of the society that press for change.

Perhaps the idea seemingly most similar to structural lag is Ogburn's (1932) concept of "cultural lag." Ogburn conceived of culture as complex, consisting of interdependent parts. There is the material culture with its technology, raw materials, manufactured products, and the like, and the nonmaterial culture that includes folkways, mores, social institutions, beliefs, laws, and governments. Ogburn thought that changes in the nonmaterial culture generally were dependent on changes in the material culture and often lagged behind changes in the material culture, hence the notion of cultural lag.

While Ogburn's concept of culture includes many of the same components of social structure posited in the theory of structural lag, his approach to social change differs from the analysis of structural lag in a number of ways. Ogburn's emphasis is on the relationship among elements in the culture. The theory of structural lag introduces the lack of fit between *people* and structures. Moreover, not only are changes in the lives of people not synchronized with changes in social structures, people act as agents of change in trying to align structures with their changing lives. A related difference is that Ogburn views the motive power of change as residing in the material culture. By

contrast, the analysts of structural lag consider changing lives and changing social structures as interdependent, with no claim for the priority of one over the other.

Others also have proposed that inconsistencies in social structures create pressures for social change. For example, Marx and Engels (1848) discussed contradictions within capitalism, and Merton (1938, 1957) analyzed the disjunction between culturally defined goals and socially differential access to the opportunity structure for achieving those goals, but these theories differ from the analysis of structural lag in a number of ways.

For Marx, change has its source in a fundamental contradiction of capitalism between private ownership of the means of production and the social nature of the production process, a contradiction that results in the exploitation of wage laborers employed by and dependent on capitalist employers. As workers struggle to improve their working conditions, their actions lead to fundamental change in the social relations of production. However, Marx's analysis of capitalist contradictions focuses on the crucial role of the productive sphere, whereas the analysis of structural lag does not give preeminence to any particular institution. Structural lag can and does occur in all societal structures, and pressures for change can emanate from all of them. Nor is social conflict the major mechanism of social change posited in the theory of structural lag, where, as was noted above, other mechanisms of change are generally more important.

In Merton's theory, the disjunction between goals and means leads to several modes of deviant adaptation, of which one, "rebellion", clearly augurs social change. In rebellion, one segment of the population rejects both goals and means as socially defined and seeks to replace them with a "greatly modified social structure" (Merton 1957, p. 155). Thus, this theory suggests a mechanism for changing the existing cultural and social structure. It is not concerned with the continuous entry into society of new cohorts whose changing lives confront social structures with the need for change, a central focus of the theory of structural lag.

In summary, the theory of structural lag, while rooted in the special qualities of age and aging as social phenomena—seemingly a narrow focus—is a broad theory that links age and aging to both

social structures and social change. In its structural aspects, it views age as a key element with which social structures must cope. From a dynamic perspective, it sees social forces bringing about change in people's lives, with those changing lives in turn causing pressure for changes in social structures. Structural lag thus is both a consequence of social change and an impetus for further change.

## REFERENCES

- Burkhauser, Richard V., and Joseph F. Quinn 1994 "Changing Policy Signals." In Matilda White Riley, Robert L. Kahn, and Anne Foner, eds., *Age and Structural Lag: Society's Failure to Provide Meaningful Opportunities in Work, Family, and Leisure*. New York: Wiley.
- Duncan, Greg, J. W. Jean Yeung, Jeanne Brooks-Gunn, and Judith R. Smith 1998 "How Much Does Childhood Poverty Affect the Life Chances of Children?" *American Sociological Review* 63:406-423.
- Foner, Anne 1974 "Age Stratification and Age Conflicts in Political Life." *American Sociological Review* 39:187-196.
- Kahn, Robert L. 1994 "Opportunities, Aspirations, and Goodness of Fit." In Matilda White Riley, Robert L. Kahn, and Anne Foner, eds., *Age and Structural Lag: Society's Failure to Provide Meaningful Opportunities in Work, Family, and Leisure*. New York: Wiley.
- Marx, Karl, and Frederich Engels. (1848) 1978 "Manifesto of the Communist Party." In R. Tucker, ed., *The Marx-Engels Reader*, 2nd ed. New York: Norton.
- Merton, Robert K. 1938 "Social Structure and Anomie." *American Sociological Review* 3:672-682.
- 1957 "Social Structure and Anomie," In *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- New York Times*, May 10, 1998. "Hospice Boom Is Giving Rise to New Fraud" by Douglas Frantz, pp. 1, 8.
- Ogburn, William F. 1932 *Social Change*. New York: Viking.
- Riley, John, Jr. 1990 "Death and Dying." In E. Borgatta and M. Borgatta, eds., *The Encyclopedia of Sociology*. New York: Macmillan.
- Riley, Matilda White 1978 "Aging, Social Change, and the Power of Ideas." *Daedalus* 107:39-52.
- , Anne Foner, and John W. Riley, Jr. 1999 "The Aging and Society Paradigm: From Generation to Generation." In Vern L. Bengtson and K. Warner Schaie, eds., *Handbook of Theories of Aging*. New York: Springer.
- , Robert L. Kahn, and Anne Foner, eds. 1994. *Age and Structural Lag: Society's Failure to Provide Meaning-*

*ful Opportunities in Work, Family, and Leisure*. New York: Wiley.

Rowe, John W. and Robert L. Kahn 1998 *Successful Aging*. New York: Pantheon.

Waring, Joan 1976 "Social Replenishment and Social Change: The Problem of Disordered Cohort Flow." In Anne Foner, ed., *Age in Society*. Beverly Hills, Calif.: Sage.

ANNE FONER

## STUDENT MOVEMENTS

Student movements generally are thought of college student movements. These young adult movements have a long history in widely differing societies. Some have been characterized as direct student redress of situational grievances, such as the seventeenth-century sacking of the English Jesuit College of La Fleche to protest a rigid, strained regimen and the student protests led by African-American and Hispanic students on over a hundred campuses in the 1980s and 1990s to protest cutbacks in governmental aid and scholarships for lower-income students. Other student protest movements have been related to larger social movements. Examples are evident over time and space, including the nineteenth-century Russian revolutionary student movement, the American civil rights and antiwar student movements of the 1930s and 1960s, the 1970s Greek student Polytechnic protest that precipitated the downfall of that country's military dictatorship, and the ill-fated Chinese Tianenmen Square democratic movement in the late 1980s.

Student movements have the potential to generate major social change in the context of underlying economic, demographic, and other social forces. This makes student movement a strategic factor in assessing the nature of some consequential social change developments in society. The recent history of the United States exemplifies this idea. The far-reaching Civil Rights Act of 1964, the public shift from support of to opposition to the Vietnam War, and the pressure to diversify college student bodies and curricula racially and ethnically all involved student protest-induced changes that have affected the lives of people throughout American society and influenced student move-

ments in other societies, as movements in other societies have influenced American students.

Other examples of consequential student protest movements extend back a millennium or more. What is different about contemporary student movements is the combination of their frequency and their consequences for social change in society. This is a reflection of the central role of formal education in economic and social stability and development in both advanced technological and developing societies.

The massive growth of higher education, with the concomitant potential for student movements, is evident from the change in the proportion of young adults in their late teens and early twenties attending college. Before World War II, even in the advanced industrial nations of Japan, the United States, and Canada, as well as in Great Britain and western Europe, less than 10 percent of the young adult age cohort attended college. The figure was less than 1 percent in emerging, often formerly colonial, developing nations. In contrast, by the late twentieth century, close to half of young adults were in college in advanced technological societies, and the fastest growing student body in developing countries had become collegiate. Overall, instead of a few hundred or a few thousand students, major state universities in the United States now generally have between 20,000 and 40,000 or more students, with long-established private universities typically having over 10,000. Similarly, large national universities such as those in Beijing, Tehran, Madrid, Mexico City, and Moscow have student bodies larger than those of the largest U.S. state universities.

The growth of public education generally, and collegiate education in particular, has placed young adult students in a strategic position in respect to the potential of protest movements to induce social change. Prototypical examples at the end of the twentieth century range from the student protests that keyed the unexpected election of the former professional wrestler Jesse Ventura to the governorship of Minnesota and the overthrow of the authoritarian Suharto regime after over thirty years in Indonesia, the fourth most populous nation in the world.

The precipitating causes of these student protest activities were very different. In the case of

Minnesota, the economy was strong and played no discernible role. Key factors were the intense national political conflict between Republicans and Democrats over President Clinton's sexual scandal, charges of obstruction of justice, and partisan controversy over a presidential impeachment. This induced many students and other young people to change allegiance from the established national parties to a reform party candidate. In Indonesia, the Asian economic crisis of 1998 played a key role in the protest activity of students demanding a more democratic government and more equitable economic opportunities.

In terms of real or potential effects on the direction of society, it is not only that college students represent a high proportion of influential future economic, political, and social leaders. The growth of colleges since the middle of the twentieth century is also an international reflection of the general public's and democratic and authoritarian regimes' recognition of the importance of the educational training of students. This training represents a key element in the future of various societies in the modern cybernetic economic era.

There has been extensive research on what motivates consequential proportions of students to engage in protest movements. There is irony in the fact that students represent a relatively privileged and prospectively influential group in society. These characteristics generally are associated with support for the established social order, yet students are often in the activist forefront of protest movements aimed at changing that order.

This seeming contradiction has been addressed in intergenerational conflict terms since the time of Socrates and Plato. In sociology, Mannheim (1952) addressed specific attention to this phenomenon as part of his concern with the sociology of knowledge. Building on Mannheim's analyses, Feuer (1969) holds that the need for the emerging young to replace older adults in societies generates inherent intergenerational conflict that crystallizes in increasingly influential collegiate settings.

In this context, it is held that students act out their traditional intergenerational conflicts in a setting that is particularly conducive to challenging the older generation. Colleges, and to a lesser extent primary and secondary schools, remove students from familial and kinship settings. While

faculty members present an adult schooling influence, in the increasingly large school settings, students are placed in a peer-related situation removed from both direct familial influences and the later pressures of occupational positions.

This relatively separated, peer-influenced life pattern is evident in the precipitating protest actions of many student movements. Most sociological research has dealt with student participation in major protest movements involving civil rights, environmental protection, war, and other momentous public issues. However, a review of the student movement literature demonstrates that often the early motivation for student protest against university administrators and more general societal authorities has been related to specific student-experienced grievances over American-based situational concerns such as poor dormitory food in the 1950s and concerns among Italian and Chinese students in the 1960s and 1980s that growing numbers of college graduates were unemployed or were receiving lower pay than were undegreed manual laborers (Altbach and Peterson 1971; Lipset 1971).

Immediate student self-interest also can be seen in respect to student participation in larger social movements. This has been evident in respect to direct student concerns about conscription and being forced into combat situations. The 1860s Harvard University economic and social elite student anticonscription protests during the Civil War helped precipitate congressional modification of who was subject to the draft. Those with several hundred dollars were allowed to commute their draft status to the next young man called up who could not afford to commute being drafted. This was a central factor in the poor nonstudent Irish Catholic, conscription riots of 1863 in New York City that left several hundred dead.

Similar immediate self-interest was a part of the American Student Union antiwar movement before World War II in the 1930s as well as the anti-Vietnam War movement led by the Students for a Democratic Society (SDS) in the 1960s. These student protests partially reflected general public disagreement about war support, but the most common precipitating thread was immediate student interest. A particularly clear case of student self-interest was the high involvement of African-

American students in the civil rights movements of the 1960s, working for more openings and support for African-Americans, who had long been excluded from equal higher educational opportunity.

However, immediate self-interest does not explain the active involvement of most participants in student movements to support disadvantaged minority and low-income groups. This is seen in protest actions such as extensive involvement in the Student Non-Violent Coordinating Committee and other American civil rights organizations in the 1960s and the 1989 student effort to establish democracy in China. In this respect, students tend to activate parental ideals and values that are perceived to be falling short in their implementation (Davies 1969). What has become evident in the extensive empirical research on student movement participants since the 1960s is that the conflict of generations thesis advanced by Mannheim, Feuer, and others is less a conflict of generations than an active attempt among the student generation to realize the values to which they have been socialized by the parental generation.

Rather than challenging the values of the parental generation, student activists generally support those values and work to see them actualized (DeMartini 1985). A case in point is the background characteristics of students who were active in the politically liberal SDS, which was strongly against the Vietnam War, and that of those in the politically conservative Young Americans for Freedom (YAF), which was strongly supportive of that war. As Lipset (1971) reports, SDS students were mostly from high-status Protestant homes where secular, liberal values prevailed. In contrast, but in intergenerational concurrence, YAF student activists generally were drawn from strongly religiously observant and conservative homes in lower middle-class and working-class settings. Another example of this intergenerational confluence is Bell's (1968) documentation that the largest proportion of white student activists in the Congress of Racial Equality (CORE) were Jewish and were actively expressing their home-based familial values in support of minority rights.

Student movement concerns with actualizing ideals have been a dynamic aspect of those movements. The national student Free Speech Movement in 1964 was precipitated by University of



California at Berkeley students who protested a specific ban on allowing a CORE civil-rights-information table on the campus in an open mall area. While a relatively small number of students were actively involved with the CORE table, a large majority of students, first at Berkeley and then nationally, supported the First Amendment right of open expression, leading to the larger Free Speech Movement (Altbach and Peterson 1971).

Protest movements generally are time-delimited. Given the relatively short age dimensions of student status, student movements tend to have even shorter time spans. Even with time and leadership delimitations, student movements are sufficiently common and consequential that more systematic research is needed on not only who the student protestors are but also where they go after a student activist movement ends. Research is beginning to ascertain the extent to which former student protest activists' ideas and behaviors continue to reflect their protest values.

It is clear that most student activists enter into business and professional, high-socioeconomic-status positions. What is not as clear is the extent to which they continue to adhere to the values and related issues that motivated them to engage in student movements. Research in this area of student movements is suggestive of long-term consequences.

An analysis over time of 1960s student protest activists and nonactivists indicates that protest values continue to influence social, economic, and political behavior. Well over a decade after their civil rights and anti-Vietnam War activity, former activists continued to be more change-oriented than average, and given the nature of their protests, they were more liberal on issues of civil rights and civil liberties. Their orientation was to support more than did nonactivists government action to address a wide range of social problems and support specific policies such as abortion rights and affirmative action for minorities' and women's educational and employment opportunities (Sherkat and Blocker 1997). Further research may demonstrate additional social change consequences on society long after specific student movements have ended.

(SEE ALSO: *Protest Movements*; *Social Movements*)

## REFERENCES

- Altbach, Philip G., and Patti Peterson 1971 "Before Berkeley: Historical Perspectives on American Student Activism." *Annals of the American Academy of Political and Social Science* 395:1-14.
- Bell, Inge Powell 1968 *CORE and the Strategy of Non-Violence*. New York: Random House.
- Davies, James C. 1969 "The J-Curve of Rising Expectations and Declining Satisfactions as a Cause of Some Great Revolutions and Contained Rebellions." In H. Graham and T. Gurr, eds., *Violence in America*. New York: Bantam.
- DeMartini, Joseph R. 1985 "Change Agents and Generational Relationships: A Reevaluation of Mannheim's Problem of Generations." *Social Forces* 64:1-16.
- Feuer, Lewis S. 1969 *The Conflict of Generations: The Character and Significance of Student Movements*. New York: Basic Books.
- Lipset, Seymour M. 1971 *Rebellion in the University*. Boston: Little, Brown.
- Mannheim, Karl 1952 *Essays on the Sociology of Knowledge*. New York: Oxford University Press.
- Sherkat, Darren E., and T. Jean Blocker 1997 "Explaining the Political and Personal Consequences of Protest." *Social Forces* 75:1049-1076.

LEONARD GORDON

## SUBURBANIZATION

Suburbanization is one aspect of the more general process of the expansion and spatial reorganization of metropolitan settlements. Settled areas that are beyond the historical boundaries of what have been considered cities but still are clearly functionally linked to the cities or may not be considered suburban. What is suburban is a matter of social definition. For example, when small cities are enveloped by the expansion of larger cities, at what point should they be considered suburbs, if they should be called suburbs at all? As some cities extend their boundaries outward, will the newly settled areas not be considered suburban if they are within the new boundaries?

Many researchers in the United States have chosen to adopt conventions established by the U.S. Bureau of the Census. The term "suburban" refers to the portion of a metropolitan area that is not in the central city. This definition depends on what is defined as metropolitan and central city,

and those definitions change over the years. Such changes are not simply technical adjustments; they respond (among other criteria) to assumptions about what cities and suburbs are. For example, as many U.S. "suburbs" have become employment centers in the last two decades, altering traditional patterns of commuting to work, Census Bureau scientists have adjusted the definition of "central city" to include some of those peripheral areas.

For many purposes, it may be preferable to avoid these categories altogether. "Suburban" may be intended to reflect distance from the city center, recency of development, residential density, or commuting patterns—all of which can be measured directly. The main substantive rationale for accepting definitions tied to the juridical boundaries of cities is to emphasize the differences between cities and suburbs (and among suburbs) that are related to municipal governance. An important class of issues revolves around disparities in public resources: In what parts of the metropolis are taxes higher, where are better schools available, where is police protection greater? What are the effects of these differences on the opportunities available to people who live in different parts of the metropolis? Another dimension concerns local politics: How do localities establish land use and budget policies, and what are the effects of those policies on growth?

Because many suburban residents have worked in central cities while paying taxes in the suburbs, John Kasarda has described the city-suburb relationship in terms of "exploitation." Political scientists in particular have studied this issue in terms of arguments for the reform of structures of metropolitan governance. The normative implications of their arguments have explicit ideological underpinnings. Some, such as Dennis Judd, emphasize the value of equality of life chances and interpret differences between cities and suburbs as disparities; others (public choice theorists such as Elinor Ostrom) emphasize freedom of choice and interpret differences as opportunities for the exercise of choice.

Sociologists on the whole have been less willing to be proponents of metropolitan solutions and have shown more interest in the causes than in the consequences of suburbanization. Nevertheless, there are differences in theoretical perspective that closely parallel those in political science,

and they hinge in part on the importance of political boundaries and the political process. The main lines of explanation reflect two broader currents in sociological theory: Structural functionalism is found in the guise of human ecology and neoclassical economics, and variants of Marxian and Weberian theory have been described as the "new" urban theory.

Ecologists and many urban economists conceptualize suburbanization as a process of decentralization, as is reflected in Burgess's (1967) concentric-zone model of the metropolis. Burgess accepted the postulate of central place theory that the point of highest interaction and most valued land is naturally at the core of the central business district. The central point is most accessible to all other locations in the metropolis, a feature that is especially valuable for commercial firms. At the fringes of the business district, where land is held for future commercial development, low-income and immigrant households can compete successfully for space, though only at high residential densities. Peripheral areas, by contrast, are most valued by more affluent households, particularly those with children and a preference for more spacious surroundings.

The key to this approach is its acceptance of a competitive land market as the principal mechanism through which locational decisions are reached. More specific hypotheses are drawn from theories about people's preferences and willingness (and ability) to pay for particular locations or structural changes (e.g., elevators, transportation technology, and the need for space of manufacturers) that affect the value of a central location. Many researchers have focused on gradients linking distance from the center to various compositional characteristics of neighborhoods: population density (Treadway 1969), household composition (Guest 1972), and socioeconomic status (Choldin and Hanson 1982). Comparatively little research has been conducted on the preferences of residents or the factors that lead them to select one location or another.

Other sociologists have argued that growth patterns result from conscious policies and specific institutional interventions in the land and housing markets. Representative of this view is the study done by Checkoway (1980), who emphasizes the role of federal housing programs and institu-

tional support for large-scale residential builders in the suburbanization process of the 1950s. The move to suburbs, he argues, was contingent on the alternatives offered to consumers. The redlining of inner-city neighborhoods by the Federal Housing Administration, its preference for large new subdivisions, and its explicit discrimination against minority home buyers are among the major forces structuring these alternatives.

There have been few studies of the housing market from an institutional perspective, although the restructuring of real estate financing and the emergence of new linkages between large-scale developers and finance capital have begun to attract attention. More consideration has been given to the explicitly political aspects of land development (Logan and Molotch 1987). Following Hunter (1953), who believed that growth questions are the "big issue" in local politics, later studies found that the most powerful voices in local politics are the proponents of growth and urban redevelopment and, in this sense, that a city is a growth machine.

In applying this model to suburbs, most observers portray suburban municipalities as "exclusionary." Suburban municipalities have long used zoning to influence the location and composition of land development. Since environmentalism emerged as a formidable political movement in the early 1970s, it has become commonplace to hear about localities that exercise their power to preserve open space and historic sites by imposing restraints or even moratoriums on new development. The "no-growth movement" is a direct extension of earlier exclusionary zoning policies.

#### **SOCIOECONOMIC DIFFERENCES BETWEEN CITIES AND SUBURBS**

These two theoretical perspectives can be illustrated through their application to research on socioeconomic differences between cities and suburbs. It is well known that central cities in most metropolitan regions have a less affluent residential population than do the surrounding suburbs. There is much debate, however, whether this class segregation between cities and their suburbs represents a natural sorting out of social classes through the private market or whether its causes are political and institutional. Similar debate surrounds the phenomenon of differentiation *within* suburbia,

where there is great variation in economic function, class and racial composition, and other characteristics of suburbs.

Research from an ecological perspective has stressed a comparison between the older, larger, denser cities of the North and the more recently growing cities of the South and West. The principal consistent findings have been that (1) the pattern of low central city relative to suburban social status is more pronounced in older metropolitan regions but that (2) controlling for metropolitan age, there appears to be a universal generalization of this pattern over time (Guest and Nelson 1978). These sociologists propose that suburbs have natural advantages over central cities. For example, their housing stock is newer, their land is less expensive, and they are more accessible to freeways and airports. The socioeconomic differences between cities and suburbs reflect those advantages.

Others argue that disparities are generated primarily by political structures that allocate zoning control and responsibility for public services to local governments and require those governments to finance services from local sources such as taxes on real property. They propose that the typical fragmentation of metropolitan government creates the incentive and opportunity for suburbs to pursue exclusionary growth policies (Danielson 1976).

Seeking to test these theories, Logan and Schneider (1982) found greater disparities in metropolitan areas where central cities were less able to grow through annexation (thus where suburban municipal governments were more autonomous) and where localities were more reliant on local property taxes (hence had greater incentive to pursue exclusionary policies). They also found a significant racial dimension: Greater disparities were evident in both 1960 and 1970 in metropolitan areas in the North with a larger proportion of black residents. (This did not hold for the South and West, however.) This disparity is due both to the concentration of lower-income blacks in central cities and to a greater propensity of higher-status whites to live in suburbs in those metropolitan areas. This finding is reinforced by Frey (see Frey and Speare 1988; Shihadeh and Ousey 1996), who reported that the central-city proportion of black residents is a significant predictor of white flight, independent of other causes.

If suburbs follow exclusionary growth policies, it seems counterintuitive that suburbs experienced much more rapid growth than did cities in the postwar decades. The findings on city–suburb disparities, of course, indicate that exclusion has selective effects. Nevertheless, it is surprising that in a study of northern California cities, Baldassare and Protash (1982), found that communities with more restrictive planning controls actually had higher rates of population growth in the 1970s. Similarly, Logan and Zhou (1989) found that suburban growth controls had little, if any, impact on development patterns (population growth, socioeconomic status, and racial composition). In their view, the exclusionary policies of suburbs may be more apparent than real. The more visible actions, such as growth moratoriums, often are intended to blunt criticisms by residents concerned with problems arising from rapid development. Unfortunately, few studies have looked in depth at the political process within suburbs; there is as little direct evidence on the role of local politics as there is on the operation of the land market. Most research from both the ecological and the political-institutional perspectives has inferred the *processes* for controlling growth from evidence about the *outcomes*.

### SUBURBANIZATION OF EMPLOYMENT

A central problem for early studies of suburban communities was to identify the patterns of functional specialization among them. It was recognized that older industrial satellites coexisted with dormitory towns in the fringe areas around central cities. Both were suburban in the sense that they were integrated into a metropolitan economy dominated by the central city. Their own economic role and the nature of the populations they housed were quite distinct, however. The greatest population gains in the 1950s occurred in residential suburbs, communities that were wealthier, younger, newer, and less densely settled than the towns on the fringes of the region that had higher concentrations of employment. Schnore (see Schnore and Winsborough 1972) distinguished “suburbs” from “satellites” to acknowledge these different origins.

The metaphors of suburbs and satellites reflected the reality of early postwar suburbanization, a period when established towns and small cities

were surrounded by successive waves of new subdivisions. Those metaphors are no longer appropriate. Since the late 1950s, the bulk of new manufacturing and trade employment in the metropolis has been located in small and middle-sized cities in the suburban ring (Berry and Kasarda 1977, chap. 13). Downtown department stores compete with new suburban shopping malls. The highly developed expressway network around central cities frees manufacturing plants to take advantage of the lower land prices and taxes and the superior access to the skilled workforce offered by the suburbs. For the period 1963–1977, in the largest twenty-five metropolitan areas, total manufacturing employment in central cities declined by about 700,000 (19 percent), while their suburbs gained 1.1 million jobs (36 percent). At the same time, total central-city retail and wholesale employment was stagnant (dropping by 100,000). Trade employment in the suburbs increased 1.8 million (or 110 percent) in that period. Thus, total employment growth in the suburbs outpaced the growth of population (Logan and Golden 1986). This is the heart of the phenomenon popularized by Garreau (1991) as the creation of “Edge City.”

How has suburbanization of employment affected suburban communities? According to microeconomic and ecological models, locational choices by employers reflect the balance of costs and benefits of competing sites. New employment maintains old patterns because the cost–benefit equation is typically stable, including important considerations such as location relative to workforce, suppliers, markets, and the local infrastructure. In the terms commonly used by urban sociologists, this means that communities find their “ecological niche.” Stahura’s (1982) finding of marked persistence in manufacturing and trade employment in suburbs from 1960 to 1972 supports this expectation. Once it has “crystallized,” the functional specialization of communities changes only under conditions of major shifts in the needs of firms.

In this view, to the extent that changes occur, they follow a natural life cycle (Hoover and Vernon 1962). Residential suburbs in the inner ring, near the central city, tend over time to undergo two related transformations: to higher population density and a conversion to nonresidential development and to a lower socioeconomic status. Thus, inner suburbs that gain employment are—like older satellites—less affluent than residential suburbs.

By contrast, those who emphasize the politics of land development suggest very different conclusions. A growing number of suburbs perceive business and industry as a significant local resource. Once shunned by the higher-status suburbs, they now contribute to property values and the local tax base. Prestigious communities such as Greenwich, Connecticut, and Palo Alto, California, house industrial parks and corporate headquarters. The “good climate for business” they offer includes public financing of new investments, extensive infrastructure (roads, utilities, parking, police and fire protection), and moderate taxes (Logan and Molotch 1987).

Competition among suburbs introduces a new factor that has the potential to reshape suburban regions. Schneider (1989) reports that location of manufacturing firms is affected by the strength of the local tax base, suggesting that wealthy suburbs are advantaged in this competition. Logan and Golden (1986) find that newly developing suburban employment centers have higher socioeconomic status, as well as stronger fiscal resources, than other suburbs; this is a reversal of the pattern of the 1950s.

### MINORITY SUBURBANIZATION

The suburbanization process also increasingly involves minorities and immigrants, and the incorporation of those groups into suburban areas has become an important topic for research on racial and ethnic relations. As Massey and Denton (1987) document, the rate of growth of nonwhites and Hispanics in metropolitan areas is far outstripping the rate of growth of non-Hispanic whites. Much of this growth is occurring in suburbs. During the 1970s, for example, the number of blacks in the non-central-city parts of metropolitan areas increased 70 percent compared with just 16 percent in central cities, and the number of other nonwhites in those locations shot up 150 percent compared with approximately 70 percent in central cities. One reason for the rapidly increasing racial and ethnic diversity of suburbs may be that some new immigrant groups are bypassing central cities and settling directly in suburbs. Equally important is the increasing suburbanization of older racial and ethnic minorities, such as blacks (Frey and Speare 1988).

This phenomenon has encouraged researchers to study suburbanization as a mirror on the social mobility of minorities. Consistent with classical ecological theory, suburbanization often has been portrayed broadly as a step toward assimilation into the mainstream society and a sign of the erosion of social boundaries. For European immigrant groups after the turn of the century, residential decentralization appears to have been part of the general process of assimilation (Guest 1980).

Past studies have found that suburbanization of Hispanics and Asians in a metropolitan area is strongly associated with each group’s average income level (Massey and Denton 1987, pp. 819–820; see also Frey and Speare 1988, pp. 311–315). Further, again for Hispanics and Asians, Massey and Denton (1987) demonstrate that suburban residence typically is associated with lower levels of segregation and, accordingly, higher probabilities of contact with the Anglo majority. However, these and other authors report very different results for blacks. Black suburbanization is unrelated to the average income level of blacks in the metropolitan area and does not result in higher intergroup contact for blacks. The suburbanization process for blacks appears largely to be one of continued ghettoization (Farley 1970), as is indicated by high and in some regions increasing levels of segregation and by the concentration of suburban blacks in communities with a high incidence of social problems (e.g., high crime rates), high taxes, and underfunded social services (Logan and Schneider 1984; Reardon 1997).

These findings regarding black suburbanization have been interpreted in terms of processes that impede the free mobility of racial minorities: steering by realtors, unequal access to mortgage credit, exclusionary zoning, and neighbor hostility (Foley 1973). Home ownership indeed may be one of the gatekeepers for suburban living. Stearns and Logan (1986) report that blacks were less likely to live in suburban areas where higher proportions of the housing stock were owner-occupied.

Further evidence is offered by Alba et al (1999), who based their conclusions on an analysis of individual-level data from the 1980 and 1990 censuses. They find that suburban residence is more likely among homeowners and persons of higher socioeconomic status. There are strong effects of

family status (marriage and the presence of children in the household), but for many immigrant groups, measures of cultural assimilation (English language use, nativity, and period of immigration) have declining relevance for suburban location. Assimilation traditionally has been a major part of the suburbanization process for most groups, especially those arising out of immigration, but large pockets of relatively new immigrants have now appeared in the suburban ring.

Parallel results are found for the racial and ethnic sorting process within a suburban region (the New York–New Jersey suburban region, as reported by Logan and Alba 1993; Alba and Logan 1993). Two sorts of analyses were conducted. First, members of different racial and ethnic groups were compared on the average characteristics of the suburbs in which they resided. Second, regression models were estimated for members of each major racial or ethnic group to predict several of these indicators of place advantages or community resources.

There are important differences between whites, blacks, Hispanics, and Asians in regard to the kinds of suburbs in which they live. As some researchers have suspected, suburban Asians have achieved access to relatively advantaged communities that are similar in most respects to those of suburban non-Hispanic whites. Hispanics in the New York region have not. Suburban Hispanic by and large live in communities that are about the same as black suburbs: communities with low average income levels and low rates of home ownership and, perhaps more important, high crime rates (Alba et al. 1994).

Is the disadvantage of blacks and Hispanics attributable to individual qualities of group members, or do these groups face collective disadvantages? Analysis of individual characteristics that may predict the quality of the suburb in which one resides shows that the same location process does not apply equally to all minorities. The pattern for whites, who are relatively advantaged in terms of access to community resources, lends clear support to assimilation theory. Human capital and indicators of cultural assimilation are strongly associated with access to higher-status suburbs. The same can be said of Asians (who are relatively advantaged overall), with the exception that cul-

tural assimilation variables seem not to be important for Asians.

The results for blacks call attention to processes of racial stratification. Even controlling for many other individual characteristics, blacks live in suburbs with lower ownership and income levels than do non-Hispanic whites. Further, most human capital and assimilation variables have a smaller payoff for blacks than they do for whites. The findings for Hispanics are supportive of the assimilation model in several respects. Hispanics gain more strongly than whites do from most human capital characteristics; therefore, at higher levels of socioeconomic achievement and cultural assimilation, Hispanics come progressively closer to matching the community resources of whites. It should be noted, however, that Hispanics begin from a lower starting point and that black Hispanics face a double disadvantage that is inconsistent with an assimilation perspective.

### LOOKING TO THE FUTURE

Suburbanization continues to be a key aspect of metropolitan growth and is perhaps of growing importance in the global era (Muller 1997). The political boundaries between cities and suburbs accentuate interest in substantive issues of metropolitan inequality. They also create special opportunities for theories of urbanization to go beyond economic models and incorporate an understanding of the political process. Research on suburbanization has been most successful in describing patterns of decentralization and spatial differentiation. The movements of people and employment and the segregation among suburbs by social class, race, ethnicity, and family composition have been well documented. However, these patterns are broadly consistent with a variety of interpretations, ranging from those which assume a competitive land market (human ecology) to those which stress the institutional and political structuring of that market.

The principal gaps in knowledge concern the processes that are central to these alternative interpretations. Few sociologists have directly studied the housing market from the perspective of either demand (how do people learn about the alternatives, and how do they select among them?) or

supply (how does the real estate sector operate, how is racial and ethnic segmentation of the market achieved, how is the complex of construction industries, developers, and financial institutions tied to the rest of the economy?). Rarely have sociologists investigated government decisions (at any level) that impinge on development from the point of view either of their effects or of the political process that led to them. Of course, these observations are not specific to research on suburbanization. It is important to bear in mind that neither the theoretical issues nor the research strategies in this field distinguish suburbanization from other aspects of the urban process.

(SEE ALSO: *Cities; Community; Urbanization*)

## REFERENCES

- Alba, Richard and John R. Logan 1993 "Minority Proximity to Whites in Suburbs: An Individual-Level Analysis of Segregation" *American Journal of Sociology* 98:1388-1427.
- , —, and Paul Bellair 1994 "Living with Crime: The Implications of Racial and Ethnic Differences in Suburban Location" *Social Forces* 73:395-434.
- , —, Brian Stults, Gilbert Marzan, and Wenquan Zhang 1999 "Immigrant Groups and Suburbs: A Reexamination of Suburbanization and Spatial Assimilation." *American Sociological Review* 64:446-460.
- Baldassare, Mark, and William Protash 1982 "Growth Controls, Population Growth, and Community Satisfaction." *American Sociological Review* 47:339-346.
- Berry, Brian, and John Kasarda 1977 *Contemporary Urban Ecology*. New York: Macmillan.
- Burgess, Ernest W. 1967 "The Growth of the City." In R. E. Park, E. W. Burgess, and R. D. McKenzie, eds., *The City*. Chicago: University of Chicago Press.
- Checkoway, Barry 1980 "Large Builders, Federal Housing Programmes, and Postwar Suburbanization." *International Journal of Urban and Regional Research* 4:21-44.
- Choldin, Harvey M., and Claudine Hanson 1982 "Status Shifts within the City." *American Sociological Review* 47:129-41.
- Danielson, Michael 1976 *The Politics of Exclusion*. New York: Columbia University Press.
- Farley, Reynolds 1970 "The Changing Distribution of Negroes within Metropolitan Areas: The Emergence of Black Suburbs." *American Journal of Sociology* 75:512-529.
- Foley, Donald 1973 "Institutional and Contextual Factors Affecting the Housing Choices of Minority Residents." In Amos Hawley and Vincent Rock, eds., *Segregation in Residential Areas*. Washington, D.C.: National Academy of Sciences.
- Frey, William, and Alden Speare 1988 *Regional and Metropolitan Growth and Decline in the United States*. New York: Russell Sage Foundation.
- Garreau, Joel 1991 *Edge City: Life on the New Frontier*. New York: Doubleday.
- Guest, Avery M. 1972 "Patterns of Family Location." *Demography* 9:159-171.
- 1980 "The Suburbanization of Ethnic Groups." *Sociology and Social Research* 64:497-513.
- , and G. Nelson 1978 "Central City/Suburban Status Differences: Fifty Years of Change." *Sociological Quarterly* 19:723.
- Hoover, Edgar, and Raymond Vernon 1962 *Anatomy of a Metropolis*. Garden City, N.Y.: Doubleday.
- Hunter, Floyd 1953 *Community Power Structure*. Chapel Hill: University of North Carolina Press.
- Liska, Allen, John R. Logan, and Paul Bellair 1998 "Race and Violent Crime in the Suburbs." *American Sociological Review* 63:27-38.
- Logan, John R. and Richard Alba 1993 "Locational Returns to Human Capital: Minority Access to Suburban Community Resources." *Demography* 30:243-268.
- , and Reid Golden 1986 "Suburbs and Satellites: Two Decades of Change." *American Sociological Review* 51:430-437.
- , and Harvey L. Molotch 1987 *Urban Fortunes: The Political Economy of Place*. Berkeley: University of California Press.
- , and Mark Schneider 1982 "Governmental Organization and City-Suburb Income Inequality, 1960-1970." *Urban Affairs Quarterly* 17:303-318.
- 1984 "Racial Segregation and Racial Change in American Suburbs: 1970-1980." *American Journal of Sociology* 89:874-888.
- , and Min Zhou 1989 "Do Growth Controls Control Growth?" *American Sociological Review* 54:461-471.
- Massey, Douglas, and Nancy Denton 1987 "Trends in the Residential Segregation of Blacks, Hispanics, and Asians: 1970-1980." *American Sociological Review* 52:802-825.
- Muller, Peter O. 1997 "The Suburban Transformation of the Globalizing American City." *Annals of the American Academy of Political and Social Science* 551:44-58.

- Reardon, Kenneth M. 1997 "State and Local Revitalization Efforts in East St. Louis, Illinois." *Annals of the American Academy of Political and Social Science* 551:235–247.
- Schneider, Mark 1989 *The Competitive City: The Political Economy of Suburbia*. Pittsburgh: University of Pittsburgh Press.
- Schnore, Leo, and Hall Winsborough 1972 "Functional Classification and the Residential Location of Social Classes." In Brian Berry, ed., *City Classification Handbook: Methods and Application*. New York: Wiley.
- Shihadeh, Edward S., and Graham C. Ousey 1996 "Metropolitan Expansion and Black Social Dislocation: The Link between Suburbanization and Center-City Crime." *Social Forces* 75:649–666.
- Stahura, John 1982 "Determinants of Suburban Job Change in Retailing, Wholesaling, Service, and Manufacturing Industries: 1960–1972." *Sociological Focus* 15:347–357.
- Stearns, Linda, and John Logan 1986 "The Racial Structuring of the Housing Market and Segregation in Suburban Areas." *Social Forces* 65:28–42.
- Treadway, Roy C. 1969 "Social Components of Metropolitan Population Densities." *Demography* 6:55–74.
- Warner, Kee, and Harvey Molotch 1995 "Power to Build: How Development Persists Despite Local Limits." *Urban Affairs Review* 30:378–406.

JOHN R. LOGAN

## SUICIDE

To many people, suicide—intentional self-murder—is an asocial act of a private individual, yet sociology grew out of Durkheim's argument ([1897] 1951) that suicide rates are social facts and reflect variation in social regulation and social interaction. The concept of suicide derives from the Latin *sui* ("of oneself") and *cide* ("a killing"). Shneidman (1985) defines "suicide" as follows: "currently in the Western world a conscious act of self-induced annihilation best understood as a multidimensional malaise in a needful individual who defines an issue for which suicide is perceived as the best solution." Several conceptual implications follow from this definition.

Although suicidal types vary, there are common traits that most suicides share to some extent. (Shneidman 1985). Suicides tend to

- Seek a solution to their life problems by dying
- Want to cease consciousness
- Try to reduce intolerable psychological pain
- Have frustrated psychological needs
- Feel helpless and hopeless
- Be ambivalent about dying
- Be perceptually constricted and rigid thinkers
- Manifest escape, egression, or fugue behaviors
- Communicate their intent to commit suicide or die
- Have lifelong self-destructive coping responses (sometimes called "suicidal careers")

Completed suicides must be differentiated from nonfatal suicide attempts, suicide ideation, and suicide talk or gestures. Sometimes one speaks of self-injury, self-mutilation, accident proneness, failure to take needed medications, and the like—where suicide intent cannot be demonstrated—as "parasuicide." The most common self-destructive behaviors are indirect, such as alcoholism, obesity, risky sports, and gambling. There are also mass suicides (as in Jonestown, Guyana, in 1978 and in Masada in A.D. 72–73) and murder suicides. Individual and social growth probably require some degree of partial self-destruction.

Although most suicides have much in common, suicide is not a single type of behavior. Suicidology will not be an exact science until it specifies its dependent variable. The predictors or causes of suicide vary immensely with the specific type of suicidal outcome. Suicidologists tend to recognize three to six basic types of suicide, each with two or three of its own subtypes (Maris et al. 1992, chap. 4). For example, Durkheim ([1897] 1951) thought all suicides were basically anomic, egoistic, altruistic, or fatalistic. Freud (1917 [1953]) and Menninger (1938) argued that psychoanalytically, all suicides were based on hate or revenge (a "wish to kill"); on depression, melancholia, or hopelessness (a "wish to die"); or on guilt or shame (a "wish to be killed"). Baechler (1979)



added “oblativ” (i.e., sacrifice or transfiguration) and “ludic” (i.e., engaging in ordeals or risks and games) suicidal types.

**EPIDEMIOLOGY, RATES, AND PREDICTORS**

Suicide is a relatively rare event, averaging 1 to 3 in 10,000 in the general population per year. In 1996 (the most recent year for which U. S. vital statistics are available), there were 31,130 suicides, accounting for about 1.5 percent of all deaths. This amounts to an overall suicide rate of 11.6 per 100,000. Suicide is now the ninth leading cause of death, ranking just ahead of cirrhosis and other liver disease deaths and just behind human immunodeficiency virus (HIV) deaths. Suicide also has been moving up the list of the leading causes of death in this century (Table 1).

Suicide rates in the United States vary considerably by sex, age, and race (Table 2). The highest rates are consistently observed among white males, who constitute roughly 73 percent of all suicides. White females account for about 17 percent of all suicides. American blacks, especially females, rarely commit suicide (except for some young urban males). Some scholars have argued that black suicides tend to be disguised as homicides or accidents. In general, male suicides outnumber female suicides three or four to one. Generally, suicide rates gradually increase with age and then drop off at the very oldest ages. Female suicide rates tend to peak earlier than do those of males. Note in Table 3 that from about 1967 to 1977, there was a significant increase in the suicide rate of 15- to 24-year-olds and that suicide rates among the elderly seem to be climbing again.

Typically, marrying and having children protect one against suicide. Usually suicide rates are highest for widows, followed by the divorced and the never-married or single. Studies of suicide rates by social class have been equivocal. Within each broad census occupational category, there are job types with high and low suicide rates. For example, psychiatrists have high suicide rates, but pediatricians and surgeons have low rates. Operatives usually have low rates, but police officers typically have high rates.

The predominant method of suicide for both males and females in 1992 was firearms (Table 4).

**Ten Leading Causes of Death in the United States, 1996 (total of 2,314,690 deaths)**

Rank	Cause of Death	Rate per 100,000	No. of Deaths (all causes)
1	Disease of the heart	276.4	733,361
2	Malignant neoplasms	203.4	539,533
3	Cerebrovascular disease	60.3	159,942
4	Chronic obstructive pulmonary disease	40	106,027
5	Accidents	35.8	94,943
6	Pneumonia and influenza	31.6	83,727
7	Diabetes mellitus	23.3	61,767
8	HIV infection	11.7	31,130
9	Suicide	11.6	30,903
10	Chronic liver disease and cirrhosis	9.4	25,047

Table 1

SOURCE: Data from U.S. National Center for Health Statistics, 1998.

The second most common method among males is hanging, and among females it is a drug or medicine overdose. Females use a somewhat greater variety of methods than males do. Suicide rates tend to be higher on Mondays and in the spring-time (Gabennesch 1988).

Prediction of suicide is a complicated process (Maris et al. 1992). As is the case with other rare events, suicide prediction generates many false positives, such as identifying someone as a suicide when that person in fact is not a suicide. Correctly identifying true suicides is referred to as “sensitivity,” and correctly identifying true nonsuicides is called “specificity.” In a study using common predictors (Table 5) Porkorny (1983) correctly predicted fifteen of sixty-seven suicides among 4,800 psychiatric patients but also got 279 false positives.

Table 5 lists fifteen major predictors of suicide. Single predictor variables seldom correctly identify suicides. Most suicides have “comorbidity” (i.e., several key predictors are involved), and specific predictors vary with the type of suicide and other factors. Depressive disorders and alcoholism are two of the major predictors of suicide. Robins (1981) found that about 45 percent of all

**Rates of Completed Suicide per 100,000  
Population by Race and Gender, 1996**

Race and Gender Group	No. of Suicides	Percent of Suicides	Rate per 100,000
White males	22,547	73	20.9
White females	5,309	17.1	4.8
Black males	(1,389)	(4.5)	(11.4)
Black females	(204)	(0.8)	(2.0)
Nonwhite males*	2,451	8.0	11.3
Nonwhite females*	596	1.9	2.5
Totals	30,903	100.0	11.6

*Table 2*

NOTE: \*Includes American Indian, Chinese, Hawaiian, Japanese, Filipino, Other Asian or Pacific Islander, and Other.

SOURCE: Data from Centers for Disease Control, 1998.

completed suicides involved either depressed or alcoholic persons. Roughly 15 percent of all those with depressive illness and 18 percent of all alcoholics eventually commit suicide. Repeated depressive illness that leads to hopelessness is especially suicidogenic.

Nonfatal suicide attempts, talk about suicide or dying, and explicit plans or preparations for dying or suicide all increase suicide risk. However, for the paradigmatic suicide (older white males), 85 to 90 percent of these individuals make only one fatal suicide attempt and seldom explicitly communicate their suicidal intent or show up at hospitals and clinics. Social isolation (e.g., having no close friends, living alone, being unemployed, being unmarried) and lack of social support are more common among suicides than among controls. Suicide tends to run in families, and this suggests both modeling and genetic influences. Important biological and sociobiological predictors of suicide have been emerging, especially low levels of central spinal fluid serotonin in the form of 5-HIAA (Maris 1997).

#### **HISTORY, COMPARATIVE STUDIES, AND SOCIAL SUICIDOLOGISTS**

The incidence and study of suicide have a long history and were fundamental to the development

of sociology. The earliest known visual reference to suicide is Ajax falling on his sword (circa 540 B.C.). Of course, it is known that Socrates (about 399 B.C.) drank hemlock. In the Judeo-Christian scriptures there were eleven men (and no women) who died by suicide, most notably Samson, Judas, and Saul. Common biblical motives for suicide were revenge, shame, and defeat in battle. Famous suicides in art history include paintings of Lucretia stabbing herself (after a rape), Dido, and work by Edvard Munch and Andy Warhol.

Suicide varies with culture and ethnicity. Most cultures have at least some suicides. However, suicide is rare or absent among the Tiv of Nigeria, Andaman islanders, and Australian aborigines and relatively infrequent among rural American blacks and Irish Roman Catholics. The highest suicide rates are found in Hungary, Germany, Austria, Scandinavia, and Japan (Table 6). The lowest rates are found in several South American, Pacific Island, and predominantly Roman Catholic countries, including Antigua, Jamaica, New Guinea, the Phillipines, Mexico, Italy, and Ireland.

The sociological study of suicide started with Durkheim ([1897] 1951) and has continued to the present day primarily in the research and publications of the following sociologists: Short, (1954), J.P. Gibbs (1964), J.T. Gibbs (1988), Douglas (1967), Maris (1969, 1981), Phillips (1974), Phillips et al. (1991), Stack (1982), Wasserman (1989), and Pescosolido and Georgianna (1989). It is impossible in an encyclopedia article to do justice to the full account of the sociological study of suicide. For a more complete review, the reader is referred to Maris (1989).

Durkheim ([1897] 1951) claimed that the suicide rate varied inversely with social integration and that suicide types were primarily ego-anomic. However, Durkheim did not operationally define "social integration." Gibbs and Martin (1964) created the concept of "status integration" to correct this deficiency. They hypothesized that the less frequently occupied status sets would lead to lower status integration and higher suicide rates. Putting it differently, they expected status integration and suicide rates to be negatively associated. In a large series of tests from 1964 to 1988, Gibbs confirmed his primary hypothesis only for occupational statuses, which Durkheim also had said were of central importance.

**Rates of Completed Suicide per 100,000 Population by Year and Age in the United States**

Age*	YEAR				
	1957	1967	1977	1987	1992
5-14	0.2	0.3	0.5	0.7	0.9
15-24	4.0	7.0	13.6	12.9	12.9
25-34	8.6	12.4	17.7	15.4	14.6
35-44	12.8	16.6	16.8	15.0	15.1
45-54	18.0	19.5	18.9	15.9	14.7
55-64	22.4	22.4	19.4	16.6	14.9
65-74	25.0	19.8	20.1	19.4	16.6
75-84	26.8	21.0	21.5	25.8	23.1
>85	26.3	22.7	17.3	22.1	21.9
Total	9.8	10.8	13.3	12.7	12.0

*Table 3*

NOTE: Suicide not reported for individuals under 5 years of age.  
SOURCE: Data from Centers for Disease Control, 1995.

Short (Henry and Short 1954) expanded Durkheim's concept of external and constraining social facts to include interaction with social psychological factors of "internal constraint" (such as strict superego restraint) and frustration-aggression theory. Short reasoned that suicide rates would be highest when external restraint was low and internal restraint was high and that homicide rates would be highest when internal restraint was low and external restraint was high.

A vastly different sociological perspective on suicide originated with the work of ethnomethodologist Douglas. Douglas, in the tradition of Max Weber's subjective meanings, argued that Durkheim's reliance on official statistics (such as death certificates) as the data base for studying suicide was fundamentally mistaken (Douglas 1967). Instead, it is necessary to observe the accounts or situated meanings of individuals who are known to be suicidal, not rely on a third-party official such as a coroner or medical examiner who is not a suicide and may use ad hoc criteria to classify a death as a suicide. There are probably as many official statistics as there are officials.

Maris (1981) extended Durkheim's empirical survey of suicidal behaviors, but not just by measuring macrosocial and demographic or structural

variables. Instead, Maris focused on actual interviews ("psychological autopsies") of the intimate survivors of suicides (usually their spouses) and compared those cases with control or comparison groups of natural deaths and nonfatal suicide attempts. Maris claimed that suicides had long "suicidal careers" involving complex mixes of biological, social, and psychological factors.

Phillips (1974) differed with Durkheim's contention that suicides are not suggestible or contagious. In a pioneering paper in the *American Sociological Review*, he demonstrated that front-page newspaper coverage of celebrity suicides was associated with a statistically significant rise in the national suicide rate seven to ten days after a publicized suicide. The rise in the suicide rate was greater the longer the front-page coverage, greater in the region where the news account ran, and higher if the stimulus suicide and the person supposedly copying the suicide were similar. In a long series of similar studies, Phillips et al. (1991) expanded and documented the suggestion effect for other types of behavior and other groups. For example, the contagion effect appears to be especially powerful among teenagers. Nevertheless, contagion accounts only for a 1 to 6 percent increase over the normal expected suicide rates in a population.

**Percent of Completed Suicides in 1987 and 1992 by Method and Gender**

METHOD	Gender			
	Male		Female	
	1987	1992	1987	1992
Firearms (E955.0–955.4)	64.0	65	39.8	39
Drugs/medications (E950.0–950.5)	5.2	–	25.0	–
Hanging (E953.0)	13.5	16	9.4	14
Carbon monoxide (E952.0–952.1)	9.6	–	12.6	–
Jumping from a high place (E957)	1.8	–	3.0	–
Drowning (E954)	1.1	–	2.8	–
Suffocation by plastic bag (E953.1)	0.4	–	1.8	–
Cutting/piercing instruments (E965)	1.3	1	1.4	1
Poisons (E950.6–950.9)	0.6	–	1.0	–
Other*	2.5	18	3.2	45
Totals	100.0	100.0	100.0	99.0

*Table 4*

NOTE: \*Includes gases in domestic use (E951), other specified and unspecified gases and vapors (E952.8–952.9), explosives (E955.5), unspecified firearms and explosives (E955.9), and other specified or unspecified means of hanging, strangulation, or suffocation (E953.8–953.9).  
SOURCE: Data from National Center for Health Statistics, 1995.

Phillips’s ideas about contagion dominated the sociological study of suicide in the 1980s. Works by Stack (1982), Wasserman (1989), Kessler and Strip (1984), and others have produced equivocal support for the role of suggestion in suicide (Diekstra et al. 1989). Wasserman (1989) feels that the business cycle and unemployment rates must be controlled for. Some have claimed that imitative effects are statistical artifacts. Most problematic is the fact that the theory of imitation in suicide is underdeveloped.

The most recent sociologist to study suicide is the medical sociologist Pescosolido. She has claimed, contrary to Douglas, that the official statistics on suicide are acceptably reliable and, as Gibbs said earlier, are the best basis available for a science of suicide. Her latest paper (Pescosolido and Georgianna 1989) examined Durkheim’s claim that religious involvement protects against suicide. Pescosolido and Georgianna find that Roman Catholicism and evangelical Protestantism protect one against sui-

**Common Single Predictors of Suicide**

1. Depressive illness, mental disorder
2. Alcoholism, drug abuse
3. Suicide ideation, talk, preparation, religion
4. Prior suicide attempts
5. Lethal methods
6. Isolation, living alone, loss of support
7. Hopelessness, cognitive rigidity
8. Older white males
9. Modeling, suicide in the family, genetics
10. Work problems, economics, occupation
11. Marital problems, family pathology
12. Stress, life events
13. Anger, aggression, irritability, 5-HIAA
14. Physical illness
15. Repetition and comorbidity of factors 1–14, suicidal careers

*Table 5*

SOURCE: Maris et al. 1992, chap. 1.

cide (institutional Protestantism does not) and that Judaism has a small and inconsistent protective effect. Those authors conclude that with disintegrating network ties, individuals who lack both integrative and regulative supports commit suicide more often.

**ISSUES AND FUTURE DIRECTIONS**

Much of current sociological research on suicide appears myopic and sterile compared to the early work of Durkheim, Douglas, and Garfinkel. Not only is the scope of current research limited, there is very little theory and few book-length publications. Almost no research monographs on the sociology of suicide were written in the 1980s. Highly focused scientific journal articles on imitation have predominated, but none of these papers have been able to establish whether suicides ever were exposed to the original media stimulus. Since suicide does not concern only social relations, the study of suicide needs more interdisciplinary syntheses. The dependent variable (suicide) must include comparisons with other types of death and

SUICIDE

<b>Suicide Rates per 100,000 Population in 62 Countries, 1980-1986</b>		<b>Suicide Rates per 100,000 Population in 62 Countries, 1980-1986</b>	
COUNTRY	RATE	COUNTRY	RATE
1. Hungary	45.3	36. Trinidad and Tobago	8.6
2. Federal Republic of Germany	43.1	37. Guadeloupe	7.9
3. Sri Lanka	29.0	38. Ireland	7.8
4. Austria	28.3	39. Italy	7.6
5. Denmark	27.8	40. Thailand	6.6
6. Finland	26.6	41. Argentina	6.3
7. Belgium	23.8	42. Chile	6.2
8. Switzerland	22.8	43. Spain	4.9
9. France	22.7	44. Venezuela	4.8
10. Suriname	21.6	45. Costa Rica	4.5
11. Japan	21.2	46. Ecuador	4.3
12. German Democratic Republic	19.0	47. Greece	4.1
13. Czechoslovakia	18.9	48. Martinique	3.7
14. Sweden	18.5	49. Colombia	2.9
15. Cuba	17.7	50. Mauritius	2.8
16. Bulgaria	16.3	51. Dominican Republic	2.4
17. Yugoslavia	16.1	52. Mexico	1.6
18. Norway	14.1	53. Panama	1.4
19. Luxemborg	13.9	54. Peru	1.4
20. Iceland	13.3	55. Philippines	0.5
21. Poland	13.0	56. Guatemala	0.5
22. Canada	12.9	57. Malta	0.3
23. Singapore	12.7	58. Nicaragua	0.2
24. United States	12.3	59. Papua New Guinea	0.2
25. Hong Kong	12.2	60. Jamaica	0.1
26. Australia	11.6	61. Egypt	0.1
27. Scotland	11.6	62. Antigua and Barbuda	-
28. Netherlands	11.0		
29. El Salvador	10.8		
30. New Zealand	10.3		
31. Puerto Rico	9.8		
32. Uruguay	9.6		
33. Northern Ireland	9.3		
34. Portugal	9.2		
35. England and Wales	8.9		

*Table 6*

SOURCE: World Health Organization data bank, latest year of reporting as of July 1, 1988.

violence as well as more nonsocial predictor variables (Holinger 1987).

A second issue concerns methods for studying suicide (Lann et al. 1989). There has never been a truly national sample survey of suicidal behaviors in the United States. Also, most suicide research is retrospective and based on questionable vital statistics. More prospective or longitudinal research design are needed, with adequate sample sizes and comparison or control groups. Models of suicidal careers should be analyzed with specific and appropriate statistical techniques such as logistic regression, log-linear procedures, and event or hazard analysis. Federal funds to do major research on suicide are in short supply, and this is probably the major obstacle to the contemporary scientific study of suicide.

Most studies of suicide are cross-sectional and static. Future research should include more social developmental designs (Blumenthal and Kupfer 1990). There is still very little solid knowledge about the social dynamics or "suicidal careers" of eventual suicides (Maris 1990). For example, it is well known that successful suicides tend to be socially isolated at the time of death, but how they came to be that way is less well understood. Even after almost a hundred years of research the relationship of suicide to social class, occupation, and socioeconomic status is not clear.

A major issue in the study of suicide is rational suicide, active euthanasia, the right to die, and appropriate death. With a rapidly aging and more secular population and the spread of the acquired immune deficiency (AIDS) virus, the American public is demanding more information about and legal rights to voluntary assisted death (see the case of Nico Speijer in the Netherlands in Diekstra et al. 1989). The right to die and assisted suicide have been the focus of a few recent legal cases (Humphry and Wickett 1986; Battin and Maris 1983). Rosewell Gilbert, an elderly man who was sentenced to life imprisonment in Florida for the mercy killing of his sick wife, was pardoned by the governor of Florida (1990). However, in 1990, the U.S. Supreme Court (*Cruzon v. the State of Missouri*) ruled that hospitals have the right to force-feed even brain-dead patients. The Hemlock Society has been founded by Derek Humphry to assist those who wish to end their own lives, make living wills, or pass living will legislation in their states

(however, see the *New York Times*, February 8, 1990, p. A18). Of course, the state must assure that the right to die does not become the obligation to die (e.g., for the aged). These issues are further complicated by strong religious and moral beliefs.

Should society help some people to die, and if so, who and in what circumstances? All people have to die, after all, so why not make dying free from pain, as quick as is desired, and not mutilating or lonely? One cannot help thinking of what has happened to assisted death at the other end of the life span, when help has not been available, in the case of abortion. Women often mutilate themselves and torture their fetuses by default. The same thing usually happens to suicides when they shoot themselves in the head in a drunken stupor in a lonely bedroom or hotel room. Obviously, many abortions and most suicides are not "good deaths."

Euthanasia is not a unitary thing. It can be active or passive, voluntary or involuntary, and direct or indirect. A person can be against one type of euthanasia but in favor of another. "Active euthanasia" is an act that kills, while "passive euthanasia" is the omission of an act, which results in death. For example, passive or indirect euthanasia could consist of "no-coding" terminal cancer or heart patients instead of resuscitating them or not doing cardiopulmonary resuscitation after a medical crisis.

"Voluntary euthanasia" is death in which the patient makes the decision (perhaps by drafting a living will), as opposed to "involuntary euthanasia," in which someone other than the patient (e.g., if the patient is in a coma) decides (the patient's family, a physician, or a nurse).

"Direct euthanasia" occurs when death is the primary intended outcome, in contrast to "indirect euthanasia," in which death is a by-product, for example, of administering narcotics to manage pain but secondarily causes respiratory failure.

All the types of euthanasia have associated problems. For example, active euthanasia constitutes murder in most states. It also violates a physician's Hippocratic oath (first do no harm) and religious rules (does all life belong to God?) and has practical ambiguities (when is a patient truly hopeless?).

Passive euthanasia is often slow, painful, and expensive. For example, the comatose patient Karen Anne Quinlan lived for ten years (she survived even after the respirator was turned off) and seemed to grimace and gasp for breath. Her parents and their insurance company spent thousands of dollars on what proved to be a hopeless case. The U.S. Supreme Court ruled in *Cruzan* (1990) that hospitals cannot be forced to discontinue feeding comatose patients.

In a case in which the author served as an expert, Elizabeth Bouvia, a quadriplegic cerebral palsy patient in California, sued to avoid being force-fed as a noncomatose patient. Her intention was to starve herself to death in the hospital. The California Supreme Court upheld Bouvia's right to refuse treatment, but others called the court's decision "legal suicide."

A celebrated spokesperson for euthanasia in the form of assisted suicide has been Derek Humphry, especially in his best-selling book *Final Exit* (1996). Rational assisted suicide (Humphry assisted in his first wife's death and in the death of his father-in-law), even for the terminally ill within six months of death, has proved highly controversial, particularly to Catholics and the religious right. Basically, Humphry has written a "how-to" book on the practicalities of suicide for the terminally ill.

His preferred rational suicide technique is to ingest four or five beta-blocker tablets and 40 to 60 100-mg tablets of a barbituate (perhaps in pudding or Jell-O), taken with Dramamine (to settle the stomach), vodka (or one's favorite whiskey), and a plastic bag over the head loosely fixed by a rubber band around the neck. Humphry recommends against guns (too messy), cyanide (too painful), hanging (too graphic), jumping (one could land on another person), and other mutilating, violent, painful, or uncertain methods.

One of the big questions about *Final Exit* is its potential abuses, for example, by young people with treatable, reversible depression. Having the lethal methods for suicide described in such vivid, explicit details worries many people that suicide will become too easy and thus often will be inappropriate. Yet Humphry shows that it is hard to get help with self-deliverance without fear of penalties. He argues that laws need to be changed to permit and specify procedures for physician-as-

sisted suicide for the terminally ill under highly controlled conditions.

A few states have undertaken such reforms to permit legal assisted death. For example, Initiative 119 in the fall of 1991 in Washington and Proposition 161 in the fall of 1992 in California would have provided "aid in dying" for a person if (1) two physicians certified that the person was within six months of (natural) death (i.e., terminally ill), (2) the person was conscious and competent, and (3) the person signed a voluntarily written request to die witnessed by two impartial, unrelated adults. Both referenda failed by votes of about 45 percent in favor and 55 percent against.

Humphry waged a similar legal battle in Oregon, first as president of the Hemlock Society and later as president of the Euthanasia Research and Guidance Organization (ERGO) and the Oregon Right to Die organization. On November 4, 1994, Oregon became the first state to permit a doctor to prescribe lethal drugs expressly and explicitly to assist in a suicide (see Ballot Measure 16). The National Right to Life Committee effectly blocked the enactment of this law until-1997, when the measure passed overwhelmingly again. On March 25, 1998, an Oregon woman in her mid-eighties stricken with cancer became the first known person to die in the United States under a doctor-assisted suicide law (most, if not all, of Dr. Jack Kervorkian's assisted suicides have probably been illegal).

Physician-assisted suicide has been practiced for some time in the Netherlands. On February 10, 1993, the Dutch Parliament voted 91 to 45 to allow euthanasia. To be eligible for euthanasia or assisted-suicide in the Netherlands, one must (1) act voluntarily, (2) be mentally competent, (3) have a hopeless disease without prospect for improvement, (4) have a lasting longing (or persistent wish) for death, (5) have assisting doctor consult at least one colleague, and (6) have written report drawn up afterward.

The Dutch law opened the door for similar legislation in the United States, although the U.S. Supreme court seems to have closed that door shut in Washington and New York. Box 1 discusses reviews of Dr. Herbert Hendin's *Seduced by Death*, which opposes physician-assisted death the United States and the Netherlands. While the idea of legal assisted suicide will remain highly contro-



versial and devious, it is quite likely that bills similar to Oregon's Measure 16 will pass in other states in the next decade. A key issue will be safeguards against abuses (for example, Hendin argues that physicians in the Netherlands have decided on their own in some cases to euthanize patients).

### THE DUTCH CASE

*The following are excerpts from reviews of Dr. Herbert Hendin's *Seduced by Death, Doctors, Patients, and the Dutch Cure* (Norton 1997). See *Suicide and Life-Threatening Behavior* 28:2, 1998.*

*On June 26, 1997, the United States Supreme Court handed down a unanimous decision on physician-assisted suicide. All nine justices concurred that both New York and Washington's state bans on the practice should stand.*

*The picture [Hendin paints in the Netherlands] is a frightening one of excessive reliance on the judgment of physicians, a consensual legal system that places support of the physician above individual patient rights in order to protect the euthanasia policy, the gradual extension of practice to include administration of euthanasia without consent in a substantial number of cases, and psychologically naive abuses of power in the doctor-patient relationship.*

*[For example:] Many patients come into therapy with sometimes conscious but often more unconscious fantasies that cast the therapist in the role of executioner . . . It may also play into the therapist's illusion that if he cannot cure the patient, no one else can either." (Seduced by Death, p. 57)*

*Samuel Klagsburn, M.D., says of Hendin's argument: "He is wrong . . . suffering needs to be addressed as aggressively as possible in order to stop unnecessary suffering."*

*Hendin claims that in the Netherlands, "despite legal sanction, 60% of [physician-assisted suicide and death] cases are not reported, which makes regulation impossible."*

*Hendin goes on to argue that "a small but significant percentage of American doctors are now practicing assisted suicide, euthanasia, and the ending of patients' lives without their consent." But one also has to wonder: what*

*about all those patients being forced to live and suffer without the patients' consent?*

*Dr. Hendin is, after all, the former Executive Director and current Medical Director of the American Foundation for Suicide Prevention. What would really be news is if Hendin came out in favor of physician-assisted death. Certainly, there are abuses of any policy. But is that enough of a reason to fail to assist fellow human beings in unremitting pain to die more easily? Death is one the most natural things there is and often is the only relief.*

One of the most controversial advocates of physician-assisted suicide ("medicide") has been Dr. Kervorkian (Kevorkian 1991). Public awareness of assisted suicide and whether it is rational has focused largely on Kervorkian, the "suicide doctor." As of early 1999, Kervorkian had assisted in over 100 suicides.

Initially, with Janet Adkins, Kervorkian used a suicide machine, which he dubbed a "mercitron." This machine provided a motor-driven, timed release of three intravenous bottles; in succession, they were (1) thiopental or sodium pentathol (an anesthetic that produces rapid unconsciousness), (2) succinylcholine (a muscle paralyzer like the curare used in Africa use in poison darts to hunt monkeys), and (3) potassium chloride to stop the heart. The mercitron was turned on by the would-be suicide. Because of malfunctions in the suicide machine, almost all of Kervorkian's suicides after Atkins were accomplished with a simple facial mask hooked up to a hose and a carbon monoxide cannister, with the carbon monoxide flow being initiated by the suicide. For most nonnarcotic users or addicts, 20 to 30 milligrams of intravenous injected morphine would cause death.

All of Kervorkian's first clients were women, and most were single, divorced, or widowed. Almost all were not terminally ill or at least probably would not have died within six months. The toxicology reports at autopsy (by Frederick Rieders; the author spoke with Dr. Dragovic, the Oakland County, Michigan, medical examiner to obtain these data) showed that only two of the eight assisted suicides had detectable levels of antidepressants in their blood at the time of death. It could be concluded that Kervorkian's assisted suicides were for the most part not being treated for depressive disorders.



Given Kervorkian's zealous pursuit of active euthanasia, one suspects that at least his early assisted suicides were not adequately screened or processed, for example, in accordance with the Dutch rules (above) or other safeguards. Strikingly, Hugh Gale is reputed to have asked Kervorkian to take off the carbon monoxide mask and terminate the dying process and perhaps was ignored by Kervorkian.

It is difficult to be objective about assisted suicide. Paradoxically, Kervorkian may end up setting euthanasia and doctor-assisted suicide back several years. Not only has he lost (1991) his Michigan medical license (he was a pathologist) and been charged with murder (after videotaping the dying of an assisted suicide for a television program), but Michigan and many other states (including South Carolina) have introduced bills to make previously legal assisted suicide a felony, with concurrent fines and imprisonment.

These new laws may have a chilling effect on both active and passive euthanasia, even in the case of legitimate pain control ("palliative care") previously offered to dying patients by physicians and nurses. For example, in Michigan it is now a felony to assist a suicide. People who want self-deliverance from their final pain and suffering will be more likely to mutilate themselves, die alone and disgraced, and feel generally abandoned in their time of greatest need.

Kervorkian needs to be separated from the issue of assisted suicide. However, the issue of physician-assisted suicide or death itself is not silly and transitory.

Everyone has to die eventually, and many people will suffer machine-prolonged debilitating illness and pain that diminishes the quality of their lives. Suicide and death and permanent annihilation of consciousness (if there is no afterlife) are effective means of pain control. This refers primarily to physical pain, but psychological pain also can be excruciating. Pain cannot always be controlled short of death. Most narcotics risk respiratory death. Furthermore, narcotics often cause altered consciousness, nightmares, nausea, panic, long periods of disrupted consciousness and confusion, and addiction.

Pain control technology is progressing rapidly (e.g., spinal implant morphine pumps). There are

hospices that encourage the use classic painkilling drinks such as Cicely Saunder's "Brompton's cocktail" (a mixed drink of gin, Thorazine, cocaine, heroin, and sugar). It is also possible to block nerves or utilize sophisticated polypharmacy to soften pain.

However, some pain is relatively intractable (e.g., that from bone cancer, lung disease with pneumonia, congestive heart failure in which patients choke to death on their own fluids, gastrointestinal obstructions, and amputation). A few physicians have made the ludicrous death-in-life proposal to give hopeless terminally ill patients general anesthesia to control their pain. People do always get well or feel better. Sometimes they just need to die, not be kept alive to suffer pointlessly. Anyone deserves to be helped to die in such instances.

## REFERENCES

- Baechler, Jean 1979 *Suicides*. New York: Basic Books.
- Battin, Margaret P., and Ronald W. Maris, eds. 1983 *Suicide and Ethics*. New York: Human Sciences Press.
- Blumenthal, Susan J., and David J. Kupfer, eds. 1990. *Suicide over the Life Cycle: Risk Factor Assessment, and Treatment of Suicidal Patients*. Washington, D.C.: American Psychiatric Press.
- Diekstra, René F. W., Ronald W. Maris, Stephen Platt, Armin Schmidtke, and Gernot Sonneck, eds. 1989 *Suicide and Its Prevention: The Role of Attitude and Imitation*. Leiden: E.J. Brill.
- Douglas, Jack D. 1967 *The Social Meanings of Suicide*. Princeton, N.J.: Princeton University Press.
- Durkheim, Emile. (1897) 1951 *Suicide*. New York: Free Press.
- Freud, Sigmund (1917) 1953 "Mourning and Melancholia." In *Standard Edition of the Complete Works of Sigmund Freud*. London: Hogarth Press.
- Gabennesch, Howard 1988 "When Promises Fail: A Theory of Temporal Fluctuations in Suicide." *Social Forces* 67:129-145.
- Gibbs, Jack P., and W. T. Martin 1964 *Status Integration and Suicide*. Eugene: University of Oregon Press.
- Gibbs, Jewelle Taylor, ed. 1988 *Young, Black, and Male in America: An Endangered Species*. Dover, Mass.: Auburn House.
- Henry, Andrew F., and James F. Short 1954 *Suicide and Homicide*. New York: Free Press.

- Holinger, Paul C. 1987 *Violent Deaths in the United States: An Epidemiological Study of Suicide, Homicide, and Accidents*. New York: Guilford.
- Humphry, Derek 1996 *Final Exit*. New York: Dell.
- , and Ann Wickett 1986 *The Right to Die: Understanding Euthanasia*. New York: Harper & Row.
- Kessler, Ronald C., and H. Stripp 1984 “The Impact of Fictional Television Stories on U.S. Fatalities: A Replication.” *American Journal of Sociology* 90:151–167.
- Kervorkian, Jack 1991 *Prescription Medicine*. Buffalo, N.Y.: Prometens.
- Lann, Irma S., Eve K. Mościcki, and Ronald W. Maris, eds. 1989 *Strategies for Studying Suicide and Suicidal Behavior*. New York: Guilford.
- Maris, Ronald W. 1969 *Social Forces in Urban Suicide*. Chicago: Dorsey.
- 1981. *Pathways to Suicide: A Survey of Self-Destructive Behaviors*. Baltimore: Johns Hopkins University Press.
- 1989. “The Social Relations of Suicide.” In Douglas Jacobs and Herbert N. Brown, eds., *Suicide, Understanding and Responding: Harvard Medical School Perspectives*, Madison, Conn.: International Universities Press.
- 1990. The Developmental Perspective of Suicide.” In Antoon Leenaars, ed., *Life Span Perspectives of Suicide*. New York: Plenum.
- 1997 “Suicide.” In Renato Pulbecco, ed., *Encyclopedia of Human Biology*.
- , Alan L. Berman, John T. Maltzberger, and Robert I. Yufit, eds. 1992 *Assessment and Prediction of Suicide*. New York: Guilford.
- Menninger, Karl 1938 *Man against Himself*. New York: Harcourt, Brace.
- Pescosolido, Bernice A., and Sharon Georgianna 1989 “Durkheim, Suicide, and Religion: Toward a Network Theory of Suicide.” *American Sociological Review* 54:33–48.
- Phillips, David P. 1974 “The Influence of Suggestion on Suicide.” *American Sociological Review* 39:340–354.
- Pokorny, Alex D. 1983 “Prediction of Suicide in Psychiatric Patients.” *Archives of General Psychiatry* 40:249–257.
- , Katherine Lesyna, and David T. Paight 1991 “Suicide and the Media.” In Ronald W. Maris, et al., eds., *Assessment and Prediction of Suicide*. New York: Guilford.
- Robins, El: 1981 *The Final Months*. New York: Oxford University Press.
- Shneidman, Edwin S. 1985 *Definition of Suicide*. New York: Wiley Interscience.

- Stack, Stephen 1982 “Suicide: A Decade Review of the Sociological Literature.” *Deviant Behavior* 4:41–66.
- Wasserman, Ira M. 1989 “The Effects of War and Alcohol Consumption Patterns on Suicide: United States, 1910–1933.” *Social Forces* 67:129–145.

RONALD W. MARIS

## SUPERNATURALISM

*See* Religious Orientations.

## SURVEY RESEARCH

Survey research is the method most frequently used by sociologists to study American society and other large societies. Surveys allow sociologists to move from a relatively small sample of individuals who are accessible as carriers of information about themselves and their society to the broad contours of a large population, such as its class structure and dominant values. Surveys conform to the major requirements of the scientific method by allowing a considerable (though by no means perfect) degree of objectivity in approach and allowing tests of the reliability and validity of the information obtained.

Like many other important inventions, a survey is composed of several more or less independent parts: sampling, questioning, and analysis of data. The successful combination of those elements early in the twentieth century gave birth to the method as it is known today. (Converse 1987 provides a history of the modern survey).

### SAMPLING

The aspect of a survey that laypersons usually find the most mysterious is the assumption that a small sample of people (or other units, such as families or firms) can be used to generalize about the much larger population from which that sample is drawn. Thus, a sample of 1,500 adults might be drawn to represent the population of approximately 200 million Americans over age 18 in the year 2000. The sample itself is then used to estimate the extent to which numerical values calculated from it (for example, the percentage of the sample answering “married” to a question about marital status) are likely to deviate from the values that

would have been obtained if the entire population over age 18 had been surveyed. That estimate, referred to as “sampling error” (because it is due to having questioned only a sample, not the full population), is even stranger from the standpoint of common sense, much like pulling oneself up by one’s own bootstraps.

Although a sample of only 1,500 may be needed to obtain a fairly good estimate for the entire U.S. adult population, this does not mean that a much smaller sample is equally adequate for, say, a city of only 100,000 population. It is the absolute size of the sample that primarily determines the precision of an estimate, not the proportion of the population that is drawn for the sample—another counterintuitive feature of sampling. This has two important implications. First, a very small sample, for example, two or three hundred, is seldom useful for surveys, regardless of the size of the total population. Second, since it is often subparts of the sample, for example, blacks or whites, that are of primary interest in a survey report, it is the size of each subpart that is crucial, not the size of the overall sample. Thus, a much larger total sample may be required when the goal is to look separately at particular demographic or social subgroups.

All the estimates discussed in this article depend on the use of probability sampling, which implies that at crucial stages the respondents are selected by means of a random procedure. A nonprobability sampling approach, such as the proverbial person-in-the-street set of interviews, lacks scientific justification for generalizing to a larger population or estimating sampling error. Consumers of survey information need to be aware of the large differences in the quality of sampling that occur among organizations that claim to do surveys. It is not the case in this or other aspects of survey research that all published results merit equal confidence. Unfortunately, media presentations of findings from surveys seldom provide the information needed to evaluate the method used in gathering the data.

The theory of sampling is a part of mathematics, not sociology, but it is heavily relied on by sociologists and its implementation with real populations of people involves many nonmathematical problems that sociologists must try to solve. For example, it is one thing to select a sample of people according to the canons of mathematical

theory and quite another to locate those people and persuade them to cooperate in a social survey. To the extent that intended respondents are missed, which is referred to as the problem of nonresponse, the scientific character of the survey is jeopardized. The degree of jeopardy (technically termed “bias”) is a function of both the amount of nonresponse and the extent to which the nonrespondents differ from those who respond. If, for example, young black males are more likely to be missed in survey samples than are other groups in the population, as often happens, the results of the survey will not represent the entire population adequately. Serious survey investigators spend a great deal of time and money to reduce nonresponse to a minimum, and one measure of the scientific adequacy of a survey report is the information provided about nonresponse. In addition, an active area of research on the survey method consists of studies both of the effects of nonresponse and of possible ways to adjust for them. (for an introduction to sampling in social surveys, see Kalton 1983; for a more extensive classic treatment, see Kish 1965).

### QUESTIONS AND QUESTIONNAIRES

Unlike sampling, the role of questions as a component of surveys often is regarded as merely a matter of common sense. Asking questions is a part of all human interaction, and it is widely assumed that no special skill or experience is needed to design a survey questionnaire. This is true in the sense that questioning in surveys is seldom very different from questioning in ordinary life but incorrect in the sense that many precautions are needed in developing a questionnaire for a general population and then interpreting the answers.

Questionnaires can range from brief attempts to obtain factual information (for example, the number of rooms in a sample of dwelling units) or simple attitudes (the leaning of the electorate toward a political candidate) to extensive explorations of the respondents’ values and worldviews. Assuming that the questions have been framed with a serious purpose in mind—an assumption not always warranted because surveys are sometimes initiated with little purpose other than a desire to ask some “interesting questions”—there are two important principles to bear in mind: one

about the development of the questions and the other about the interpretation of the answers.

The first principle is the importance of carrying out as much pilot work and pretesting of the questions as possible, because not even an experienced survey researcher can foresee all the difficulties and ambiguities a set of questions holds for the respondents, especially when it is administered to a heterogeneous population such as that of the United States. For example, a frequently used question about whether “the lot of the average person is getting worse” turned out on close examination to confuse the respondents about the meaning of “lot,”—with some taking it to refer to housing lots. Of course, it is still useful to draw on expert consultation where possible and to become familiar with discussions of questionnaire design in texts, especially the classic treatment by Payne (1951) and more recent expositions such as that by Sudman and Bradburn (1987).

Pilot work can be done in a number of ways, for example, by having a sample of respondents think aloud while answering, by listening carefully to the reactions of experienced interviewers who have administered the questionnaire in its pretest form, and, perhaps best of all, by having investigators do a number of practice interviews. The distinction between “pilot” and “pretest” questionnaires is that the former refer to the earlier stages of questionnaire development and may involve relatively unstructured interviewing, while the latter are closer to “dress rehearsals” before the final survey.

The main principle in interpreting answers is to be skeptical of simple distributions of results often expressed in percentage form for a particular question, for example, 65 percent “yes,” 30 percent “no,” 5 percent “don’t know.” For several reasons, such absolute percentages suggest a meaningfulness to response distributions that can be misleading. For one thing, almost any important issue is really a cluster of subissues, each of which can be asked about and may yield a different distribution of answers. Responses about the issue of “gun control” vary dramatically in the United States depending on the type of gun referred to, the amount and method of control, and so forth. No single percentage distribution or even two or three distributions can capture all this variation, nor are such problems confined to questions about

attitudes: Even a seemingly simple inquiry about the number of rooms in a home involves somewhat arbitrary definitions of what is and is not to be counted as a room, and more than one question may have to be asked to obtain the information the investigator is seeking. By the same token, care must be taken not to overgeneralize the results from a single question, since different conclusions might be drawn if a differently framed question were the focus. Indeed, many apparent disagreements between two or more surveys disappear once one realizes that somewhat different questions had been asked by each even though the general topic (e.g., gun control) may look the same.

Even when the substantive issue is kept constant, seemingly minor differences in the order and wording of questions can change percentage distributions noticeably. Thus, a classic experiment from the 1940s showed a large difference in the responses to a particular question depending on whether a certain behavior was said to be “forbidden” rather than “not allowed”: To the question, “Do you think the United States should forbid public speeches against democracy?” 54 percent said yes, [Forbid], but to the question, “Do you think the United States should allow public speeches against democracy?” 75 percent said no (do not allow). This is a distinction in wording that would not make a practical difference in real life, since not allowing a speech would have the same consequence as forbidding it, yet the variation in wording has a substantial effect on answers. Experiments of this type, which are called “split-ballot experiments,” frequently are carried out by dividing a national sample of respondents in half and asking different versions of the question to each half on a random basis. If the overall sample is large enough, more than two variations can be tested at the same time, and in some case more complex “factorial designs” are employed to allow a larger number of variations (see Rossi and Nock [1982] for examples of factorial surveys).

The proportion of people who answer “don’t know” to a survey question also can vary substantially—by 25 percent or more—depending on the extent to which that answer is explicitly legitimized for respondents by mentioning it along with other alternatives (“yes,” “no,” “don’t know”) or omitted. In other instances, the location of a question in a series of questions has been shown to affect answers even though the wording of the question

is not changed. For example, a widely used question about allowing legalized abortion in the case of a married woman who does not want more children produces different answers depending entirely on its position before or after a question about abortion in the case of a defective fetus. Thus, the context in which a question is asked can influence the answers people give. These and a large number of other experiments on the form, wording, and context of survey questions are reported by Schuman and Presser (1981) (see Turner and Martin [1984] for several treatments of survey questioning, as well as more recent volumes by Schwarz and Sudman [1996] and Sudman et al. [1996] with a cognitive psychological emphasis).

### ANALYSIS

Although questioning samples of individuals may seem to capture the entire nature of a survey, a further component is vital to sociologists: the logical and statistical analysis of the resulting data. Responses to survey questions do not speak for themselves, and in most cases even the simple distribution of percentages to a single question calls for explicit or implicit comparison with another distribution, real or ideal. To report that 60 percent of a sample is satisfied with the actions of a particular leader may be grounds for either cheering or booing. It depends on the level of satisfaction typical for that leader at other times or for other individuals or groups in comparable leadership positions. Thus, reports of survey data should include these types of comparisons whenever possible. This is why for sociologists the collection of a set of answers is the beginning and not the end of a research analysis.

More generally, most answers take on clear meaning primarily when they are used in comparisons across time (for example, responses of a sample this year compared with responses of a sample from the same population five years ago), across social categories such as age and education, or across other types of classifications that are meaningful for the problem being studied. Moreover, since any such comparison may produce a difference that is due to chance factors because only a sample was drawn rather than to a true difference between time points or social categories, statistical testing is essential to create confidence that the difference would be found if the

entire population could be surveyed. In addition, individual questions sometimes are combined into a larger index to decrease idiosyncratic effects resulting from any single item, and the construction of this type of index requires other preliminary types of statistical analysis.

As an example of survey analysis, sociologists often find important age differences in answers to survey questions, but since age and education are negatively associated in most countries—that is, older people tend to have less education than do younger people—it is necessary to disentangle the two factors in order to judge whether age is a direct cause of responses or only a proxy for education. Moreover, age differences in responses to a question can represent changes resulting from the aging process (which in turn may reflect physiological, social, or other developmental factors) or reflect experiences and influences from a particular historical point in time (“cohort effects”). Steps must be taken to distinguish these explanations from one another. At the same time, a survey analyst must bear in mind and test the possibility that a particular pattern of answers is due to “chance” because of the existence of sampling error.

Thus, the analysis of survey data can be quite complex, well beyond, though not unrelated to, the kinds of tables seen in newspaper and magazine presentations of poll data. (The terms “poll” and “survey” are increasingly interchangeable, with the main difference being academic and governmental preference for “survey” and media preference for “poll.”) However, such thorough analysis is important if genuine insights into the meaning of answers are to be gained and misinterpretations are to be avoided. (A comprehensive but relatively nontechnical presentation of the logic of survey analysis is provided by Rosenberg [1968]. Among the many introductory statistical texts, Agresti and Finlay [1997] leans in a survey analytic direction.)

### MODE OF ADMINISTRATION

Although sampling, questioning, and analysis are the most fundamental components, decisions about the mode of administering a survey are also important. A basic distinction can be made between self-administered surveys and those in which interviewers are used. If it is to be based on probability sampling of some sort, self-administration, usually is carried out by mailing questionnaires to respon-

dents who have been selected through a random procedure. For instance, a sample of sociologists might be chosen by taking every twentieth name from an alphabetical listing of all regular members of the American Sociological Association, though with the recognition that any such listing would be incomplete (e.g., not everyone with an advanced degree in sociology belongs to the association).

The major advantage of mail surveys is their relatively low cost, which is limited to payments to clerical employees, stamps, and perhaps financial incentive for the respondents. One disadvantage of mail surveys is that they traditionally have produced low response rates; many obtain only 25 percent or less of the target sample. However, Dillman (1978) argues that designing mail surveys in accordance with the principles of *exchange theory* can yield response rates at or close to those of other modes of administration. Whether this is true for a sample of the U.S. population remains in doubt for the reason given below, although Dillman has implemented some of his strategies in government census-type surveys. It is clear from numerous experiments that the use of two specific features—monetary incentives (not necessarily large) provided in advance and follow-up “reminders”—can almost always improve mail questionnaire response rates appreciably. However, another important disadvantage of mail surveys in the United States is the absence of an available centralized national listing of households for drawing a sample; because of this situation, it is difficult to say what response rate could be obtained from a nongovernmental national mail sample in this country.

Mail surveys generally are used when there is a prior list available, such as an organization’s membership, and this practice may add the benefit of loyalty to the organization as a motive for respondent cooperation. Other disadvantages of mail surveys are lack of control over exactly who answers the questions (it may or may not be the target respondent, assuming there is a single target), the order in which the questionnaire is filled out, and the unavailability of an interviewer for respondents who cannot read well or do not understand the questions. One compensating factor is the greater privacy afforded respondents, which may lead to more candor, although evidence of this is still limited. Sometimes similar privacy is attempted an interview survey by giving a portion

of the questionnaire to respondents to fill out themselves and even providing a separate sealed envelope to mail back to the survey headquarters, thus guaranteeing that the interviewer will not read the answers. This strategy was used by Laumann et al. (1994) in a major national survey of sexual behavior, but no comparison with data obtained in a more completely private setting was provided. Tourangeau and Smith (1996) provide a different type of evidence by showing that respondents who answer directly into a computer appear more candid than do respondents who give answers to interviewers. Recently, the Internet has been investigated as a vehicle for self-administered surveys, although there are formidable problems of sampling in such cases.

Because of these difficulties, most surveys aimed at the general population employ interviewers to locate respondents and administer a questionnaire. Traditionally, this has been done on a face-to-face (sometimes called “personal”) basis, with interviewers going to households, usually after a letter of introduction has been mailed describing the survey. The sample ordinarily is drawn by using “area probability” methods: To take a simple example, large units such as counties may be drawn first on a random basis, then from the selected counties smaller units such as blocks are drawn, and finally addresses on those blocks are listed by interviewers and a randomly drawn subset of the listed addresses is designated for the actual sample, with introductory letters being sent before interviewing is attempted. In practice, more than two levels would be used, and other technical steps involving “stratification” and “clustering” would be included to improve the efficiency of the sampling and data collection.

A major advantage of face-to-face interviewing is the ability of the interviewer to find the target respondent and persuade her or him to take part in the interview. Face-to-face interviewing has other advantages: Graphic aids can be used as part of a questionnaire, interviewers can make observations of a respondent’s ability to understand the questions and of other behavior or characteristics of a respondent, and unclear answers can be clarified. The major disadvantage of face-to-face interviewing is its cost, since much of the time of interviewers is spent locating respondents (many are not at home on a first or second visit). For every actual

hour spent interviewing, five to ten hours may be needed for travel and related effort. Furthermore, face-to-face surveys require a great deal of total field time, and when results are needed quickly, this is difficult to accomplish and may add more expense. Another disadvantage is the need for an extensive supervisory staff spread around the country, and yet another is that survey administrators must rely on the competence and integrity of interviewers, who are almost always on their own and unsupervised during interviews. This makes standardization of the interviewing difficult.

Increasingly since the early 1970s, face-to-face interviewing has been replaced by telephone interviewing, usually from a centralized location. Telephone surveys are considerably less expensive than face-to-face surveys, though the exact ratio is hard to estimate because they also are normally shorter, usually under forty-five minutes in length; the expense of locating people for face-to-face interviews leads to hourlong or even lengthier interviews, since these usually are tolerated more readily by respondents who are interviewed in person. Telephone surveys can be completed more rapidly than can face-to-face surveys and have the additional advantage of allowing more direct supervision and monitoring of interviewers. The incorporation of the computer directly into interviewing—known as computer-assisted telephone interviewing (CATI)—facilitates questionnaire formatting and postinterview coding, and this increases flexibility and shortens total survey time. Still another advantage of telephone surveys is the relative ease of probability sampling: Essentially random combinations of digits, ten at a time, can be created by computer to sample any telephone number in the United States (three-digit area code plus seven-digit number). There are a variety of practical problems to be overcome (e.g., many of the resulting numbers are nonworking, account must be taken of multiple phones per household, and answering machines and other devices often make it difficult to reach real people), but techniques have been developed that make such samples available and inexpensive to a degree that was never true of the area sampling required for face-to-face interviewing. Perhaps the largest problem confronting survey research is the proliferation of telemarketing, which makes many potential respondents wary of phone calls and reluctant to devote time to a survey interview.

Because speaking on the telephone seems so different from speaking face to face, survey methodologists initially thought that the results from the two types of survey administration might be very different. A number of experimental comparisons, however, have failed to find important differences, and those which do occur may have more to do with different constraints on sampling (telephone surveys obviously miss the approximately 8 percent of the American households without telephones and produce somewhat higher levels of refusal by the intended respondents). Thus, the remaining reasons for continuing face-to-face surveys have to do with the need for longer interviews and special additions such as graphic demonstrations and response scales. (Groves [1989] discusses evidence on telephone versus face-to-face survey differences, and Groves et al. [1988] present detailed accounts of methodological issues involving telephone surveys.)

Face-to-face and telephone surveys share one important feature: the intermediate role of the interviewer between the questionnaire and the respondent. Although this has many advantages, as was noted above, there is always the possibility that some behavior or characteristic of the interviewer will affect responses. For example, as first shown by Hyman (1954) in an effort to study the interview process, a visible interviewer characteristic such as racial appearance can have dramatic effects on answers. This is probably the largest of all the effects discovered, no doubt because of the salience and tension that racial identification produces in America, but the possibility of other complications from the interview process—and from the respondent's assumption about the sponsorship or aim of the survey—must be borne in mind. This is especially true when surveys are attempted in societies in which the assumption of professional neutrality is less common than in the United States, and some recent failures by surveys to predict elections probably are due to bias of this type.

### THE SEQUENCE OF A SURVEY

Surveys should begin with one or more research problems that determine both the content of the questionnaire and the design of the sample. The two types of decisions should go hand in hand,

since each affects the other. A questionnaire that is intended to focus on the attitudes of different ethnic and racial groups makes sense only if the population sampled and the design of the sample will yield enough members of each group to provide sufficient data for adequate analysis. In addition, decisions must be made early with regard to the mode of administration of the survey—whether it will be conducted through self-administration or interviewing and, if the latter, whether in person, by telephone, or in another way—since these choices also influence what can be asked. Each decision has its trade-offs in terms of quality, cost, and other important features of the research.

After these planning decisions, the development of the questionnaire, the pretesting, and the final field period take place. The resulting data from closed, or fixed-choice, questions can be entered directly in numerical form (e.g., 1 = yes, 2 = no, 3 = don't know) into a computer file for analysis. If open-ended questions—questions that do not present fixed alternatives—are used and the respondents' answers have been recorded in detail, an intermediate step is needed to code the answers into categories. For example, a question that asks the respondents to name the most important problems facing the country today might yield categories for “foreign affairs,” “inflation,” “racial problems,” and so forth, though the words used by the respondents ordinarily would be more concrete. Finally, the data are analyzed in the form of tables and statistical measures that can form the basis for a final report.

#### **MODIFICATIONS AND EXTENSIONS OF THE SURVEY METHOD**

This discussion has concerned primarily the single cross-sectional or one-shot survey, but more informative designs are increasingly possible. The most obvious step now that surveys of the national population have been carried out for more than half a century is to study change over time by repeating the same questions at useful intervals. The General Social Survey (GSS) has replicated many attitude and factual questions on an annual or biennial basis since 1972, and the National Election Study (NES) has done the same thing in the political area on a biennial basis since the 1950s. From these repeated surveys, sociologists

have learned about substantial changes in some attitudes, while in other areas there has been virtually no change (see Niemi et al. [1989] and Page and Shapiro [1992] for examples of both change and stability). An important variant on such longitudinal research is the panel study, in which the same respondents are interviewed at two or more points in time. This has certain advantages; for example, even where there is no change for the total sample in the distribution of responses, there may be counterbalancing shifts that can be best studied in this way.

Surveys are increasingly being carried out on a cross-national basis, allowing comparisons across societies, though usually with the additional obstacle of translation to be overcome. Even within the framework of a single survey in one country, comparisons across different types of samples can be illuminating, for example, in an important early study by Stouffer (1955) that administered the same questionnaire to the general public and to a special sample of “community leaders” in order to compare their attitudes toward civil liberties. Finally, it is important to recognize that although the survey method often is seen as entirely distinct from or even opposite to the experimental method, the two have been usefully wedded in a number of ways. Much of what is known about variations in survey responses caused by the form, wording, and context of the questions has been obtained by means of split-ballot experiments, while attempts to study the effects of policy changes sometimes have involved embedding surveys of attitudes and behaviors within larger experimental designs.

#### **ETHICAL AND OTHER PROBLEMS**

As with other social science approaches to the empirical study of human beings, surveys raise important ethical issues. The success of survey sampling requires persuading individuals to donate their time to being interviewed, usually without compensation, and to trust that their answers will be treated confidentially and used for purposes they would consider worthwhile. A related issue is the extent to which respondents should be told in advance and in detail about the content and aims of a questionnaire (the issue of “informed consent”), especially when this might discourage their willingness to answer questions or affect the



kinds of answers they give (Singer 1993). The purely professional or scientific goal of completing the survey thus can conflict with the responsibility of survey investigators to the people who make surveys possible: the respondents. These are difficult issues, and there probably is no simple overall solution. There is a need in each instance to take seriously wider ethical norms as well as professional or scientific goals.

From within sociology, reliance on surveys has been criticized on several grounds. Sociologists committed to more qualitative approaches to studying social interaction often view surveys as sacrificing richness of description and depth of understanding to obtain data amenable to quantitative analysis. Sociologists concerned with larger social structures sometimes regard the survey approach as focusing too much on the individual level, neglecting the network of relations and institutions of societies. Finally, some see the dependence of surveys on self-reporting as a limitation because of the presumed difference between what people say in interviews and how they behave outside the interview situation (Schuman and Johnson 1976). Although there are partial answers to all these criticisms, each has some merit, and those doing survey research need to maintain a self-critical stance toward their own approach. However, the survey is the best-developed and most systematic method sociologists have to gather data. Equally useful methods appropriate to other goals have yet to be developed.

## REFERENCES

- Agresti, Alan, and Barbara Finlay 1997 *Statistical Methods for the Social Sciences*. Upper Saddle River, N.J.: Prentice Hall.
- Converse, Jean M. 1987 *Survey Research in the United States: Roots and Emergence, 1890–1960*. Berkeley: University of California Press.
- Dillman, Don A. 1978 *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- Groves, Robert M. 1989 *Survey Errors and Survey Costs*. New York: Wiley.
- Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, and Joseph Waksberg 1988 *Telephone Survey Methodology*. New York: Wiley.
- Hyman, Herbert H. 1954 *Interviewing in Social Research*. Chicago: University of Chicago Press.
- Kalton, Graham 1983 *Introduction to Survey Sampling*. Beverly Hills, Calif.: Sage.
- Kish, Leslie 1965 *Survey Sampling*. New York: Wiley.
- Laumann, Edward O., Robert T. Michael, John H. Gagnon, and Stuart Michaels 1994 *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
- Niemi, Richard, John Mueller, and Tom W. Smith 1989 *Trends in Public Opinion: A Compendium of Survey Data*. Westport, Conn.: Greenwood Press.
- Page, Benjamin I., and Robert Y. Shapiro 1992 *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.
- Payne, Stanley L. 1951. *The Art of Asking Questions*. Princeton, N.J.: Princeton University Press.
- Rosenberg, Morris 1968. *The Logic of Survey Analysis*. New York: Basic Books.
- Rossi, Peter H., and Steven L. Nock, eds. 1982 *Measuring Social Judgments: The Factorial Survey Approach*. Beverly Hills, Calif.: Sage.
- Schuman, Howard, and Michael P. Johnson 1976 "Attitudes and Behavior." *Annual Review of Sociology*, vol. 2. Palo Alto, Calif.: Annual Reviews.
- and Stanley Presser 1981 *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.
- Schwarz, Norbert, and Seymour Sudman, eds. 1996 *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass.
- Singer, Eleanor 1993 "Informed Consent in Surveys: A Review of the Empirical Literature." *Journal of Official Statistics*. 9:361–375.
- Stouffer, Samuel A. 1955 *Communism, Conformity, and Civil Liberties*. Garden City, N.Y.: Doubleday.
- Sudman, Seymour, and Norman M. Bradburn 1987 *Asking Questions: A Practical Guide to Question Design*. San Francisco: Jossey-Bass.
- , —, and Norbert Schwarz 1996 *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- Tourangeau, Roger, and Tom W. Smith 1996 "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Content." *Public Opinion Quarterly* 60:275–304.
- Turner, Charles, and Elizabeth Martin, eds. 1984 *Surveying Subjective Phenomena*, 2 vols. New York: Russell Sage Foundation.

HOWARD SCHUMAN

## SYMBOLIC INTERACTION THEORY

The term “symbolic interactionism” was invented by Blumer (1937) to describe sociological and social psychological ideas he presented as emanating directly from Mead, especially but not exclusively in *Mind, Self, and Society* (1934). “Symbolic interaction theory” is a term that is related to those ideas, though not necessarily in the specific forms presented by Blumer or Mead.

### FUNDAMENTAL IMAGERY

The fundamental character of symbolic interactionist ideas is suggested by the theoretical proposition that the self reflects society and organizes behavior and by related imagery that addresses the nature of society and the human being, the nature of human action and interaction, and the relationship between society and the person. That imagery begins with a vision of society as a web of communication: Society *is* interaction, the reciprocal influence of persons who, as they relate, take into account each other’s characteristics and actions, and interaction is communication. Interaction is “symbolic,” that is, conducted in terms of the meanings persons develop in the course of their interdependent conduct. The environment of human action and interaction is symbolically defined: It is the environment as it is interpreted that is the context, shaper, and object of action and interaction. Persons act with reference to one another in terms of symbols developed through interaction and act through the communication of those symbols. Society is a label aggregating and summarizing such interaction. Society does not “exist”; it is created and continuously re-created as persons interact. Social reality is a flow of events joining two or more persons. More than simply being implicated in the social process, society and the person derive from that process: They take on their meanings as those meanings emerge in and through social interaction.

Neither society nor the individual is ontologically prior to the other in this imagery; persons create society through their interaction, but it is society, a web of communication and interaction, that creates persons as social beings. Society and the individual presuppose each other; neither exists except in relation to the other. This concep-

tion of society implicitly incorporates a view of the human being as “minded” and that “mindedness” as potentially reflexive. That is, people can and sometimes do take themselves as the object of their own reflection, thus creating selves, and do this from the standpoint of the others with whom they interact. Selves are inherently social products, although they involve more than reflected appraisals of others in the immediate situation of interaction; in particular, selves involve persons as subjects responding to themselves as objects. Thinking takes place as an internal conversation that uses symbols that develop in the social process. Mind arises in both the evolutionary and individual senses in response to problems (interruptions in the flow of activities) and involves formulating and selecting from symbolically defined alternative courses of action to resolve those problems. Choice is an omnipresent reality in the human condition, and the content of choices is contained in the subjective experience of persons as that experience develops in and through the social process.

Following from this imagery is a view of human beings, both collectively and individually, as active and creative rather than simply responsive to environmental stimuli. Since the environment of human action and interaction is symbolic; because the symbols attaching to persons (including oneself), things, and ideas are the products of interaction and reflexivity and can be altered and manipulated in the course of that interaction; since thought can be used to anticipate the effectiveness of alternative courses of action in resolving problems; and because choice among alternatives is an integral feature of social conduct, one arrives at an image of social interaction as literally constructed, although not necessarily anew in each instance, in the course of interaction. One also arrives at an image that entails a degree of indeterminacy in human behavior in the sense that the course and outcome of social interaction cannot as a matter of principle (not uncertain knowledge) be completely predicted from conditions and factors existing before that interaction.

### THE SYMBOLIC INTERACTIONIST FRAMEWORK

Labeling the ideas of symbolic interactionism a “theory” is misleading. If one distinguishes between a systematic set of interrelated proposi-

tions about how a segment of the world is organized and functions and assumptions about and conceptualizations of the parts of that segment, symbolic interactionism has more the character of the latter than the former. That is, it is more a theoretical framework than a theory per se. While features of the framework appear to militate against attempts to formulate systematic theory by using it as a base and various proponents deny that possibility, a few sociologists have employed the framework in efforts to elaborate specific theories (e.g., Heise 1979; Stryker 1980, forthcoming; Stryker and Serpe 1982; Rosenberg 1984; Thoits 1983; MacKinnon 1994; Burke 1991). It is not possible to review such specific theories nor characterize the research that derives from that framework here. (For extensive references to classic literature and research literature before 1985, see Stryker and Statham [1985]. For more recent research, see one of the texts written from a symbolic interactionist perspective, such as Hewitt [1997], or *Symbolic Interaction*, a journal sponsored by the Society for the Study of Symbolic Interaction and devoted to work emanating from the framework.)

In the work of some (e.g., Blumer 1969), symbolic interaction “theory” is intended to be a general sociological frame that is applicable to the intellectual problems of sociology as a discipline from the most micro to the most macro levels. In the work of others (e.g., Stryker 1980), it is a frame restricted in utility to issues of social psychology. The first position does not seem defensible, because any framework brings into special focus particular variables and leaves unattended—at least relatively—other variables and because the symbolic interaction framework highlights interaction, social actors related through interaction, and subjective variables “internal” to those actors. It thus neglects features of the sociological landscape relating to large-scale social systems—the state, the economy, the “world system,” demographic variables, and so forth—and does not easily pose sociological questions involving interrelationships among those features of large-scale social systems. This neglect has led to criticism of the symbolic interactionist framework as lacking the social structural concepts needed for the analysis of power and consequently as an ideological apology for the status quo (see Meltzer et al. 1975; Stryker 1980; Reynolds 1990). Although many (e.g., Maines 1977) deny the validity of this criticism, pointing to work

by Hall (1972) and others, the criticism may be justified if the claim is that symbolic interactionism is a general sociological framework; it is not valid if the more restricted claim for its utility is made. However, there remains concern about the adequacy of the framework for problems of a distinctively *sociological* social psychology that centers on the reciprocal relationships of social units and social persons. There also is concern about whether the framework admits of and provides readily for the articulation of sociological and social psychology concepts. These concerns arise from the ways in which social structural concepts enter, or fail to enter, the symbolic interactionist frame (for presentations of an avowedly social structural version of the framework, see Stryker 1980, forthcoming). Whatever the intended coverage—from all of sociology to a limited social psychology—the framework traditionally has been conceived as knowing no cultural boundaries; that view, however, has been questioned (Hewitt 1990).

#### CENTRAL CONCEPTS

Implicated in the description of symbolic interactionist imagery provided above are many of the central concepts of the framework. The meaning of “meaning” is fundamental. By definition, social acts involve at least two persons taking each other into account in satisfying impulses or resolving problems. Since social acts occur over time, *gestures*—parts of an act that indicate that other parts are still to come—can appear. Vocal sounds, physical movements, bodily expressions, clothing, and so forth, can serve as gestures. When they do, they have meaning: Their meaning lies in the behavior that follows their appearance. Gestures that have the same meaning (implying the same future behavior) to those who make them and those who perceive them are *significant symbols*.

Things, ideas, and relationships among things and ideas can all be symbolized and enter the experience of human beings as objects; objects whose meanings are anchored in and emerge from social interaction constitute social reality. Although meanings are unlikely to be identical among participants, communication and social interaction presuppose significant symbols that allow meanings to be “sufficiently” shared. Because significant symbols anticipate future behavior, they entail plans of action: They organize behavior with

reference to what they symbolize. In the context of the ongoing social process, meanings must be at least tentatively assigned to features of the interactive situations in which persons find themselves; without the assignment of meanings, behavior in those situations is likely to be disorganized or random. The situation must be symbolized, as must its constituent parts; it must be defined or interpreted, and the products of that symbolization process are *definitions of the situation*. Those definitions focus attention on what is pertinent (satisfying impulses or resolving problems) in an interactive setting and permit a preliminary organization of actions appropriate to the setting. Tentative definitions are tested and may be reformulated through ongoing experience.

From the point of view of the actors involved, the most important aspects of a situation requiring definition are who or what they are in the situation and who or what the others with whom they interact are. Defining the others in the situation typically is accomplished by locating them as members of a socially recognized category of actors, one (or more) of the kinds of persons it is possible to be in a society (e.g., male or female, young or old, employed or unemployed). Doing this provides cues to or predictors of their behavior and permits the organization of one's own behavior with reference to them. When others are recognized as instantiations of a social category, behaviors are expected of them and actions that are premised on those expectations can be organized and directed toward them. Through this process, the introduction of early definitions of the situation can produce, although not inevitably, behavior that validates the definitions. This is an insight that underlies the notion of altercasting (Weinstein and Deutschberger 1963) and appears in the development of expectation states theory (Berger et al. 1974). When such behavior becomes routinized and organized, it also can serve to reproduce the existing social structure.

While some interactionists disdain the term, expectations attached to social categories—again, the kinds of persons it is possible to be in a society—are *roles*. Situations frequently allow one to locate others in multiple categories and open the possibility that conflicting expectations will come into play; in this circumstance, no clear means of organizing responses may be available. Defining oneself in a situation also involves locat-

ing oneself in socially recognized categories; to respond reflexively to oneself by classifying, naming, and defining who and what one is is to have a *self*. The self, conceived in this manner, involves viewing oneself as an object. The meaning of self, like that of any object, derives from interaction: To have a self is to view oneself from the standpoint of those with whom one interacts. The self, like any significant symbol, provides a plan of action. By definition, that plan implicates the expected responses of others.

People learn, at least provisionally, what they can expect from others through *role taking*, a process of anticipating the responses of the others with whom one interacts. In effect, one puts oneself in the place of those people to see the world as they do, using prior experience with them, knowledge of the social categories in which they are located, and symbolic cues available in interaction. On such bases, tentative definitions of others' attitudes are formulated and then validated or reshaped in interaction. Role taking permits one to anticipate the consequences of one's own and others' plans of action, monitor the results of those plans as they are carried out behaviorally, and sustain or redirect one's behavior on the basis of the monitoring. Because roles often lack consistency and concreteness while actors must organize their behavior as if roles were unequivocal, interaction is also a matter of *role making*: creating and modifying roles by devising performances in response to roles imputed to others (Turner 1962).

Many social acts take place within organized systems of action; consequently, both role taking and role making can occur with reference to a *generalized other*, that is, a differentiated but interrelated set of others (Mead's example involves baseball players anticipating the responses of other members of their team and those of their opponents). Not all others' perspectives are equally relevant to an actor; the concept of *significant other* indicates that some persons will be given greater weight when perspectives differ or are incompatible. It is implied here that meanings are not likely to be universally shared or shared in detail; if they are not, accuracy in role taking and difficulty in role making also will vary. It also is implied that smooth and cooperative interpersonal relations do not necessarily follow from accurate role taking: Conflict may result from or be sharpened by such accuracy.

The symbolic interactionist ideas reviewed here have a history. Many issue directly from Mead. Mead's ideas are part of a tradition of philosophical thought with roots in the Scottish moral philosophers Adam Smith, David Hume, Adam Ferguson, and Francis Hutcheson and, more proximately, in the American pragmatists Charles Peirce, William James, and John Dewey. They also contain important admixtures of evolutionary and dialectic premises. Mead's thought overlaps considerably with that of a number of sociologists who also wrote in the first decades of the twentieth century, in particular Charles Horton Cooley and William Isaac Thomas. Cooley's axiom that society and the person are two sides of the same coin (the coin, he added, is communication) and that of Thomas asserting that if humans define situations as real, the situations are real in their consequences, capture much of the essence of symbolic interactionism. A host of sociologists connect that past with the present; among others, Burgess, Blumer, Waller, Sutherland, Hughes, Shibutani, Kuhn, Cottrell, Hill, Lemert, Lindesmith, Mills, Miyamoto, and Stone are linked to a more contemporary set of persons that includes Goffman, Lofland, Becker, Lopata, Strauss, Geer, Weinstein, Farberman, Couch, Denzin, Bart, Maines, Reynolds, Turner, Daniels, Scheff, Wiseman, Heise, Stryker, Burke, Heiss, Fine, Hochschild, Weigert, McCall, Snow, and Hewitt. (For reviews of the history and literature of symbolic interactionism, see Stryker and Statham 1985; Meltzer et al. 1975; Reynolds 1990; Lewis and Smith 1980.) The presence of these researchers in a common listing does not indicate their adherence to a common credo; there may be much conceptual difference as well as similarity among them.

### COMMONALITIES AND VARIATIONS

Thus, no single version of symbolic interaction theory satisfies all who find its core ideas appealing and useful in conducting research and analyses. There appear to be three fundamental premises of a symbolic interactionist perspective that are shared by those who acknowledge their intellectual roots in this tradition of sociological thought (Stryker 1988). The first holds that an adequate account of human social behavior must incorporate the perspective of participants in interaction and cannot rest entirely on the perspective of the observer. The second is that the self, that is, persons' reflex-

ive responses to themselves, links the larger social organization or structure to the social interaction of those persons. The third asserts that processes of social interaction are prior to both self and social organization, both of which derive and emerge from social interaction.

Each of these premises leaves open issues of considerable importance with respect to the content, methods, and objectives of interactionist analyses on which symbolic interactionists can differ. Some sociologists for whom the three core premises serve as a starting point believe that social life is so fluid that it can be described only in process terms, that concepts purportedly describing social structures or social organization belie the reality of social life. Relatedly, some believe that actors' definitions, which theoretically are central and powerful as generators of lines of action, are reformulated continuously in immediate situations of interaction, making it impossible to use preexistent concepts to analyze social life (Blumer 1969). Others accept the "reality" of social structural phenomena, viewing the social structure as relatively stable patternings of social interaction that operate as significant constraints on actors' definitions. Social structure is thought to make for sufficient continuity in definitions to allow the use of concepts derived from past analyses of social interaction in the analysis of present and future interaction (Stryker 1980). The first premise hides, in the term "*account*," the important difference between those who seem to believe that given the constructed character of social behavior, only an "after the fact" understanding of past events is possible (Weigert 1981) and those who believe that sociology can build testable predictive explanations of social behavior (Kuhn 1964). Similarly, some argue that the perspective of a sociological observer of human social behavior is likely to distort accounts of that behavior and so must be abjured in seeking to capture the perspectives of those who live the behavior that is observed (Denzin 1970), and others argue directly or by implication that the requirement that accounts incorporate the perspective of the actors whose behavior is observed dictates only that actors' definitions be included in developed explanations, not that they constitute those explanations (Burke 1991). The first group tends to argue that the best, if not the only "legitimate," methods are naturalistic, primarily observational (Becker and Geer 1957); the

second group tends to be catholic with respect to methods, refusing to rule out categorically any of the full range of possible social science methods and techniques (Heise 1979).

With respect to the second premise, interactionists differ in the degree to which they assign an independent "causal" role to the self as the link between social organization or structure and social behavior. For many, self can and does serve as an independent source of that behavior (McCall and Simmons 1978). For others, social organization or structure (as the residue of prior interaction) builds selves in its image, thus making the self essentially a conduit through which these structures shape behavior, not an independent source of that behavior (Goffman 1959). Similarly, there is variation among symbolic interactionists in the degree to which the self is seen as the source of creativity and novelty in social life, the degree to which creativity and novelty in social life are seen as probable as opposed to simply possible (occurring only under a specific and limited set of social circumstances), and the degree to which social life is constructed anew rather than "merely" reconstructed in the image of prior patterns (Turner 1962; Hewitt 1997; Stryker and Statham 1985).

The third premise is interpreted by some as denying that social organization and selves have sufficient constancy to permit generalized conceptualization or the development of useful a priori theory on the basis of any investigation that can carry over reasonably to any new investigation (Glaser and Strauss 1967). In the view of others, this premise does not deny that there is in social life a reasonable constancy that implies a sufficient constancy in both selves and social organization to permit the elaboration of useful theories employing general concepts that potentially are applicable to wide instances of social behaviors (Heise 1986). Some emphasize the behavioristic elements in their intellectual heritage from Mead, concentrating on how concerted lines of social action are constructed (Couch et al. 1986; McPhail and Wohlstein 1986), while others adopt a stance that attends primarily to the phenomenological worlds of the actors (or interactors) they study (Denzin 1984).

Clearly, these possibilities for important variations in symbolic interactionist thought are not independent of one another: Those who subscribe to a view emphasizing the fluidity of social life and

the moment-to-moment situated character of definitions also are likely to emphasize the degree to which social order continuously emerges from fluid process, the self organizes social behavior in an unconstrained fashion, and creativity and novelty characterize human behavior. They also tend to insist that the point of view of the observer contaminates reasonable accounts of social interaction, that there is little utility in an analysis of conceptualizations and theory emanating from earlier analyses, and that understanding, not explanation, is the point of sociological efforts.

The set of views presented in the preceding paragraph identifies symbolic interactionism for many of its most passionate adherents and perhaps for a majority of its critics. Those approaching their work from symbolic interactionism so defined tended to present what they did in both conceptual and methodological opposition to available alternatives in sociology. For example, Blumer (1969) devoted much of his career to championing direct and participant observation aimed at accessing the interpretations of those whose ongoing interaction sociologists sought to understand as opposed to both statistical and structural analyses, whose categories, data, and mathematical manipulations seemed to him devoid of actors' meanings. Critics of symbolic interactionism attacked it and its adherents for being nonscientific and asociological. To circumscribe symbolic interactionism in the manner of these adherents and critics belies the diversity in views on key issues represented in the work of those who use the framework.

### CONTEMPORARY VITALITY

Interest in the symbolic interactionist framework within sociology has fluctuated. That interest was great from 1920 to 1950, reflecting in part the dominance of the University of Chicago in producing sociologists as well as the institutional structure of sociology. Through the 1950s and into the 1970s, interest waned, first as the structural functionalism of Parsons and Merton gained ascendance intellectually and Harvard and Columbia became institutionally dominant and later as Marxist and structuralist emphases on macro social processes swept the field. Symbolic interactionism, when not decried as reactionary or asociological, became the loyal opposition (Mullins 1973). Indeed, Mullins

predicted that it would disappear as a viable sociological framework.

More recent events contradict that prediction: Symbolic interactionism has had a remarkable revitalization in the past three decades (Stryker 1987), and there has been a corresponding resurgence of interest in the framework. The revitalization and resurgent interest reflect various sources. One is an emerging realization among sociologists with a structural orientation that their theories could benefit from the sociologically sophisticated theory of the social actor and action that symbolic interactionism can provide and the related increasing interest in linking micro to macro social processes. A second lies in a series of changing emphases in the work of contemporary symbolic interactionists. Although much recent work in a symbolic interactionist framework reflects traditional conceptual, theoretical, and methodological themes, on the conceptual level, newer work tends to adopt a "multiple selves" perspective, drawing on William James (Stryker 1989; McCall and Simmons 1978) rather than viewing the self as singular or unitary. Theoretically, there is greater attention to emotion, to affective dimensions of social life (Hochschild 1979; Thoits 1989; MacKinnon 1994; Ervin and Stryker, forthcoming), correcting for a "cognitive bias" in the framework; there is also greater appreciation for structural facilitators of and constraints on interaction and on self processes. While not yet prominent in contemporary interactionism, the groundwork has been laid (e.g., in Stryker and Statham 1985) for the reintroduction of the concept of habit, which was central in the writings of John Dewey and other forerunners of interactionism, in recognition that social life is not invariably reflexive and minded. Current symbolic interactionism is methodologically eclectic and tends to be more rigorous than it was in the past, whether the methods are ethnographic (Corsaro 1985) or involve structural equation modeling (Serpe 1987). Also contributing to the revitalization of symbolic interactionism is the attention to its ideas, often unacknowledged but sometimes recognized, paid by a psychological social psychology that is predominately cognitive in its orientation. For cognitive social psychology, concepts are mental or subjective structures formed through experience, and these structures affect recognizing, attending, storage, recall, and utilization of information impinging on the person; of prime significance

among concepts functioning in these ways are self-concepts. The link thus forged between cognitive social psychology and symbolic interactionism is mutually advantageous. Symbolic interactionism benefits from the "legitimacy" implicit in the attention given to its ideas and from the expanded pool of researchers focusing on those ideas; cognitive social psychology can benefit from understanding that cognitions are rooted in social structures and processes.

(SEE ALSO: *Identity Theory*; *Role Theory*; *Self-Concept*; *Social Psychology*)

## REFERENCES

- Becker, Howard S., and Blanche Geer 1957 "Participant Observation and Interviewing: A Comparison." *Human Organization* 16:28-32.
- Berger, Joseph, Thomas L. Connor, and M. Hamit Fisek, eds. 1974 *Expectation States Theory: A Theoretical Research Program*. Cambridge, Mass.: Winthrop.
- Blumer, Herbert 1937 "Social Psychology." In Emerson P. Schmidt, ed., *Man and Society*. New York: Prentice-Hall.
- 1969 *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, N.J.: Prentice-Hall.
- Burke, Peter J. 1991. "Attitudes, Behavior, and the Self." In Judith A. Howard and Peter L. Callero, eds., *The Self-Society Dynamic*. New York: Cambridge University Press.
- Corsaro, William A. 1985 *Friendship and Peer Culture in the Early Years*. Norwood, N.J.: Ablex.
- Couch, Carl J., Stanley L. Saxon, and Michael A. Katovich 1986 *Studies in Symbolic Interaction: The Iowa School*, 2 vols. Greenwich, Conn.: JAI Press.
- Denzin, Norman K. 1984 *On Understanding Emotion*. San Francisco: Josey Bass.
- 1970 *The Research Act*. New York: McGraw-Hill.
- Ervin, Laurie, and Sheldon Stryker forthcoming. "Self-Esteem and Identity." In Timothy J. Owens, Sheldon Stryker, and Norman Goodman, eds., *Extending Self-Esteem Theory and Research: Sociological and Psychological Currents*. New York: Cambridge University Press.
- 1984 "Toward a Phenomenology of Domestic Family Violence." *American Journal of Sociology* 3:483-513.
- Glaser, Barney G., and Anselm L. Strauss 1967 *The Discovery of Grounded Theory*. Chicago: Aldine.
- Goffman, Erving 1959 *The Presentation of Self in Everyday Life*. Garden City, N.Y.: Doubleday.

- Hall, Peter M. 1972 "A Symbolic Interactionist Analysis of Politics." *Sociological Inquiry* 42:35-75.
- Heise, David R. 1979 *Understanding Events*. New York: Cambridge University Press.
- 1986 "Modeling Symbolic Interaction." In Siegwart Lindenberg, James S. Coleman, and Stefan Nowak, eds., *Approaches to Social Theory*. New York: Russell Sage Foundation.
- Hewitt, John P. 1990 *Dilemmas of the American Self*. Philadelphia: Temple University Press.
- 1997. *Self and Society: A Symbolic Interactionist Social Psychology*, 7th ed. Boston: Allyn and Bacon.
- Hochschild, Arlie R. 1979 "Emotion Work, Feeling Rules, and Social Structure." *American Journal of Sociology* 85:551-575.
- Kuhn, Manfred H. 1964 "Major Trends in Symbolic Interaction Theory in the Past Twenty-Five Years." *Sociological Quarterly* 5:61-84.
- Lewis, J. David, and Richard J. Smith 1980 *American Sociology and Pragmatism: Mead, Chicago Sociology, and Symbolic Interaction*. Chicago: University of Chicago Press.
- McCall, George J., and J. L. Simmons 1978. *Identities and Interaction*, rev. ed. New York: Free Press.
- McPhail, Clark, and Ronald T. Wohlstein 1986 "Collective Locomotion as Collective Behavior." *American Sociological Review* 51:447-464.
- MacKinnon, Neil J. 1994. *Symbolic Interactionism as Affect Control*. Albany, NY: State University of New York Press.
- Maines, David 1977 "Social Organization and Social Structure in Symbolic Interactionist Thought." *Annual Review of Sociology* 3:235-259.
- Mead, George H. 1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Meltzer, Bernard N., John W. Petras, and Larry T. Reynolds 1975 *Symbolic Interactionism: Genesis, Varieties, and Criticism*. London: Routledge and Kegan Paul.
- Mullins, Nicholas 1973 *Theories and Theory Groups in Contemporary American Sociology*. New York: Harper and Row.
- Reynolds, Larry T. 1990 *Interactionism: Exposition and Critique*, 2nd ed. Dix Hills, N.Y.: General Hall.
- Rosenberg, Morris 1984 "A Symbolic Interactionist View of Psychosis." *Journal of Health and Social Behavior* 25:289-302.
- Serpe, Richard T. 1987 "Stability and Change in Self: A Structural Symbolic Interactionist Explanation." *Social Psychology Quarterly* 50:44-55.
- Stryker, Sheldon 1980 *Symbolic Interactionism: A Social Structural Version*. Menlo Park, Calif.: Benjamin/Cummings.
- 1987 "The Vitalization of Symbolic Interactionism." *Social Psychology Quarterly* 50:83-94.
- 1988 "Substance and Style: An Appraisal of the Sociological Legacy of Herbert Blumer." *Symbolic Interaction* 11:33-42.
- 1989 "Further Developments in Identity Theory: Singularity versus Multiplicity of Self." In Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds., *Sociological Theories in Progress: New Formulations*. Newbury Park, Calif.: Sage.
- 1997 "In the Beginning There Is Society: Lessons from a Sociological Social Psychology." In Craig McGarty and S. Alexander Haslam, eds., *The Message of Social Psychology*. Cambridge, Mass.: Blackwell.
- forthcoming "Identity Competition: Key to Differential Social Movement Participation?" In Sheldon Stryker, Timothy J. Owens, and Robert W. White, eds., *Self, Identity, and Social Movements*. Minneapolis: University of Minnesota Press.
- , and Richard T. Serpe 1982 "Commitment, Identity Salience, and Role Behavior." In William Ickes and Eric Knowles, eds., *Personality, Roles, and Social Behavior*. New York: Springer-Verlag.
- , and Anne Statham 1985 "Symbolic Interactionism and Role Theory." In Gardner Lindzey and Elliot Aronsen, eds., *The Handbook of Social Psychology*, 3rd ed. New York: Random House.
- Tamas Rudas, Clifford C. Clogg, Bruce G. Lindsay 1994. "A New Index of Fit Based on Mixture Methods for the Analysis of Contingency Tables" *Journal of the Royal Statistical Society. Series B (Methodological)* 56:623-639.
- Thoits, Peggy A. 1983 "Multiple Identities and Psychological Well-Being: A Reformulation and Test of the Social Isolation Hypothesis." *American Sociological Review* 48:174-187.
- 1989 "The Sociology of Emotions." *Annual Review of Sociology* 15:317-342.
- Turner, Ralph H. 1962 "Role-Taking: Process versus Conformity." In Arnold M. Rose, ed., *Human Behavior and Social Process*. Boston: Houghton-Mifflin.
- Weigert, Andrew J. 1981 *Sociology of Everyday Life*. New York: Longman.
- Weinstein, Eugene, and Paul Deutschberger 1963 "Some Dimensions of Altercasting." *Sociometry* 26:454-466.
- Wong, Raymond Sin-Kwok Wong 1990 "Understanding Cross-National Variation in Occupational Mobility." *American Sociological Review*, Vol. 55:560-573.



- 1992 “Vertical and Nonvertical Effects in Class Mobility: Cross-National Variations.” *American Sociological Review* 57:396–410.
- 1995 “Extensions in the Use of Log-Multiplicative Scaled Association Models in Multiway Contingency Tables.” *Sociological Methods and Research* 23:507–538.
- Xie, Yu 1992 “The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables” *American Sociological Review* 57:380–395.

SHELDON STRYKES

## SYSTEMS THEORY

**NOTE:** *Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.*

Systems theory is much more (or perhaps much less) than a label for a set of constructs or research methods. The term *systems* is used in many different ways (Boguslaw 1965; 1981, pp. 29–46). Inevitably this creates considerable confusion. For some it is a “way” of looking at problems in science, technology, philosophy, and many other things; for others it is a specific mode of decision making. In the late twentieth-century Western world it has also become a means of referring to skills of various kinds and defining professional elites. Newspaper “want ads” reflect a widespread demand for persons with a variety of “system” skills, for experts in “systems engineering,” “systems analysis,” “management systems,” “urban systems,” “welfare systems,” and “educational systems.”

As a way of looking at things, the “systems approach” in the first place means examining objects or processes, not as isolated phenomena, but as interrelated components or parts of a complex. An automobile may be seen as a system; a car battery is a component of this system. The automobile, however, may also be seen as a component of a community or a national transportation system. Indeed, most systems can be viewed as subsystems of more encompassing systems.

Second, beyond the idea of interrelatedness, systems imply the idea of control. This always includes some more or less explicit set of values. In some systems, the values involved may be as simple

as maintaining a given temperature range. The idea of control was implicit in Walter B. Cannon’s original formulation of the concept of homeostasis. Cannon suggested (Cannon 1939, p. 22) that the methods used by animals to control their body temperatures within well-established ranges might be adapted for use in connection with other structures including social and industrial organizations. He referred to the body’s ability to maintain its temperature equilibrium as *homeostasis*.

A third idea involved in the system way of looking at things is Ludwig von Bertalanffy’s search for a “general systems theory” (von Bertalanffy 1968; Boguslaw 1982, pp. 8–13). This is essentially a call for what many would see as an interdisciplinary approach. Von Bertalanffy noted the tendency toward increased specialization in the modern world and saw entire disciplines—physics, biology, psychology, sociology, and so on—encapsulated in their private universes of discourse, with little communication between any of them. He failed to note, however, that new interdisciplinary disciplines often quickly tend to build their own insulated languages and conceptual cocoons.

A fourth idea in the systems approach to phenomena is in some ways the most pervasive of all. It focuses on the discrepancy between objectives set for a component and those required for the system. In organizations this is illustrated by the difference between goals of individual departments and those of an entire organization. For example, the sales department wants to maximize sales, but the organization finds it more profitable to limit production, for a variety of reasons. If an entire community is viewed as a system, a factory component of this system may decide that short-term profitability is more desirable as an objective than investment in pollution-control devices to protect the health of its workers and community residents. Countless examples of this sort can be found. They all seem to document the idea that system objectives are more important than those of its subsystems. This is a readily understandable notion with respect to exclusively physical systems. When human beings are involved on any level, things become much more complicated.

Physical components or subsystems are not expected to be innovative. Their existence is ideal when it proceeds in a “normal” routine. If they wear out they can be replaced relatively cheaply,

and if they are damaged they can be either repaired or discarded. They have no sense of risk and can be required to work in highly dangerous environments twenty-four hours a day, seven days a week, if necessary. They do not join unions, never ask for increases in pay, and are completely obedient. They have no requirements for leisure time, cultural activities, or diversions of any kind. They are completely expendable if the system demands sacrifices. They thrive on authoritarian or totalitarian controls and cannot deal with the notion of democracy.

As a specific mode of decision making, it is this top-down authoritarianism that seems to characterize systems theory when it is predicated on a physical systems prototype. Computerization of functions previously performed by human beings ostensibly simplifies the process of converting this aspect of the theory into action. Computer hardware is presumably completely obedient to commands received from the top; software prepared by computer programmers is presumably similarly responsive to system objectives. Almost imperceptibly, this has led to a condition in which systems increasingly become seen and treated as identical to the machine in large-scale "man-machine systems." (The language continues to reflect deeply embedded traditions of male chauvinism.)

These systems characteristically have a sizable computerized information-processing subsystem that keeps assuming increasing importance. For example the U.S. Internal Revenue Service (IRS) obviously has enormous quantities of information to process. Periodically, IRS officials feel the necessity to increase computer capacity. To accomplish this, the practice has been to obtain bids from computer manufacturers. One bid, accepted years ago at virtually the highest levels of government, proposed a revised system costing between 750 million and one billion dollars.

Examination of the proposal by the congressional Office of Technology Assessment uncovered a range of difficulties. Central to these was the fact that the computer subsystem had been treated as the total system (perhaps understandably since the contractor was a computer corporation). The existing body of IRS procedures, internal regulations, information requirements, and law (all part of the larger system) was accepted as an immutable given. No effort had been made to

consider changes in the larger system that could conceivably eliminate a significant portion of the massive computer installation (Office of Technology Assessment 1972).

Almost two decades after attention had been called to these difficulties, system problems at the IRS continued to exist. A proposed Tax System Modernization was formulated to solve them. The General Accounting Office raised questions about whether this proposal, estimated to cost several billion dollars, was in fact "a new way of doing business" or simply intended to lower costs and increase efficiency of current operations. Moreover, the Accounting Office suggested that the lack of a master plan made it difficult to know how or whether the different component subsystems would fit together. Specifically, for example, it asked whether the proposal included a telecommunications subsystem and, if so, why such an item had not been included among the budgeted items (Rhile 1990).

To exclude the larger system from consideration and assume it is equivalent to a subsystem is to engage in a form of fragmentation that has long been criticized in related areas by perceptive sociologists (see Braverman 1974; Kraft 1977). Historically, fragmentation has led to *deskilling* of workers, that is, replacing craft tasks with large numbers of relatively simpler tasks requiring only semi-skilled or unskilled labor. This shields the larger system from scrutiny and facilitates centralization of control and power. It also facilitates computerization of work processes and even more control.

In the contemporary industrial and political worlds, power is justified largely on the basis of "efficiency." It is exercised largely through monopolization of information. Various forms of social organization and social structure can be used for the exercise of this power. Systems theory focuses not on alternative structures but, rather, on *objectives*, a subset of what sociologists think of as *values*. To hold power is to facilitate rapid implementation of the holder's values.

Fragmentation, in the final analysis, is an effort to divide the world of relevant systems into tightly enclosed cubbyholes of thought and practice controlled from the top. This compartmentalization is found in both government and private enterprises. The compartments are filled with those devoid of genuine power and

reflect the limitation of decisions available to their occupants. Those at the summit of power pyramids are exempt from these constraints and, accordingly, enjoy considerably more “freedom” (Pelton, Sackmann, and Boguslaw 1990).

An increasingly significant form of fragmentation is found in connection with the operation of many large-scale technological systems. Sociologist Charles Perrow has, in a path-breaking study, examined an enormous variety of such systems. He has reviewed operations in nuclear power, petrochemical, aircraft, marine, and a variety of other systems including those involving dams, mines, space, weapons, and even deoxyribonucleic acid (DNA). He developed a rough scale of the potential for catastrophe, assessing the risk of loss of life and property against expected benefits. He concluded that people would be better off learning to live without some, or with greatly modified, complex technological systems (Perrow 1984). A central problem he found involved “externalities,” the social costs of an activity *not* shown in its price, such as pollution, injuries, and anxieties. He notes that these social costs are often borne by those who do not even benefit from the activity or are unaware of the externalities.

This, of course, is another corollary to the fragmentation problem. To consider the technological system in isolation from the larger social system within which it is embedded is to invite enormous difficulties for the larger system while providing spurious profits for those controlling the subsystem.

Another interesting manifestation of the fragmentation problem arises in connection with two relatively new disciplines that address many problems formerly the exclusive province of sociology: operations research and management science. Each of these has its own professional organization and journal.

Operations research traces its ancestry to 1937 in Great Britain when a group of scientists, mathematicians, and engineers was organized to study some military problems. How do you use chaff as a radar countermeasure? What are the most effective bombing patterns? How can destroyers best be deployed if you want to protect a convoy?

The efforts to solve these and related problems gave rise to a body of knowledge initially

referred to as Operations Analysis and subsequently referred to as Operations Research. A more or less official definition of the field tells us Operations Research is concerned with scientifically deciding how to best design and operate man-machine systems usually under conditions requiring the allocation of scarce resources. In practice, the work of operations research involved the construction of models of operational activities, initially in the military, subsequently in organizations of all kinds. *Management science*, a term perhaps more congenial to the American industrial and business ear, emerged officially as a discipline in 1953 with the establishment of the Institute of Management Sciences.

In both cases, the declared impetus of the discipline was to focus on the entire system, rather than on components. One text points out that subdivisions of organizations began to solve problems in ways that were not necessarily in the best interests of the overall organizations. Operations research tries to help management solve problems involving the interactions of objectives. It tries to find the “best” decisions for “as large a portion of the *total system* as possible” (Whitehouse and Wechsler 1976).

Another text, using the terms *management science* and *operations research*, interchangeably defines them (or it) as the “application of scientific procedures, techniques, and tools to operating, strategic, and policy problems in order to develop and help evaluate solutions” (Davis, McKeown, and Rakes 1986, p. 4).

The basic procedure used in operations research/management science work involves defining a problem, constructing a model, and, ultimately, finding a solution. An enormous variety of mathematical, statistical, and simulation models have been developed with more or less predictable consequences. “Many management science specialists were accused of being more interested in manipulating problems to fit techniques than . . . (working) to develop suitable solutions” (Davis, McKeown, and Rakes 1986, p. 5). The entire field often evokes the tale of the fabled inebriate who persisted in looking for his lost key under the lamppost, although he had lost it elsewhere, because “it is light here.”

Under the sponsorship of the Systems Theory and Operations Research program of the National

Science Foundation, a Committee on the Next Decade in Operations Research (CONDOR) held a workshop in 1987. A report later appeared in the journal *Operations Research*. The journal subsequently asked operation researchers to comment on the report (Wagner et al. 1989). One of the commentators expressed what appears to be a growing sentiment in the field by pointing out the limitations of conventional modeling techniques for professional work. Criticizing the CONDOR report for appearing to accept the methodological status quo, he emphasized the character of models as “at best abstractions of selected aspects of reality” (Wagner et al. 1989). He quoted approvingly from another publication, “thus while exploiting their strengths, a prudent analyst recognizes realistically the limitations of quantitative methods” (Quade 1988).

This, however, is an unfortunate repetition of an inaccurate statement of the difficulty. It is not the limitations of quantitative methods that is in question but rather the recognition of the character of the situations to which they are applied. Sociologists distinguish between *established* situations, those whose parameters can be defined precisely and for which valid analytic means exist to describe meaningful relationship within them and *emergent* situations, whose parameters are known incompletely and for which satisfactory analytic techniques are not available within the time constraints of necessary action (Boguslaw [1965] 1981). In established situations mathematical or statistical models are quite satisfactory, along with other forms of rational analysis. In emergent situations, however, they can yield horrendous distortions. Fifty top U.S. corporation executives, when interviewed, recognized and acted upon this distinction more or less intuitively, although the situations presented to them were referred to as Type 1 and Type 2, respectively (Pelton, Sackmann, and Boguslaw 1990).

Individual persons, organizations, or enterprises may be viewed, on the one hand, as self-contained systems. On the other, they may be viewed as subsystems of larger social systems. Unfortunately, efforts are continually made to gloss over this dichotomy through a form of fragmentation, by treating a subsystem or collection of subsystems as equivalent to a larger system. It is this relationship between system and subsystem that

constitutes the core of the dilemma continuing to confront systems theory.

Achieving a satisfactory resolution of the discrepancy between individual needs and objectives of the systems within which individuals find themselves embedded or by which they are affected remains an unsolved problem as the twentieth century draws to a close.

(SEE ALSO: *Decision-Making Theory and Research; Social Dynamics; Social Structure*)

## REFERENCES

- Bernik, Ivan 1994 “Double Disenchantment of Politics: A Systems Theory Approach to Post-Socialist Transformation.” *Innovation* 7:345–356.
- Bivins, Thomas H. 1992 “A Systems Model for Ethical Decision Making in Public Relations.” *Public Relations Review* 18:365–383.
- Boguslaw, Robert (1965) 1981 *The New Utopians: A Study of Systems Design and Social Change*. Englewood Cliffs, N.J.: Prentice-Hall.
- 1982 *Systems Analysis and Social Planning: Human Problems of Post-Industrial Society*. New York: Irvington.
- Braverman, Harry 1974 *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review Press.
- Cannon, Walter B. 1939 *The Wisdom of the Body*, rev. ed. New York: Norton.
- Cohen-Rosenthal, Edward 1997 “Sociotechnical Systems and Unions: Nicety or Necessity.” *Human Relations* 50:585–604.
- Creedon, Pamela J. 1993 “Acknowledging the Infrasytem: A Critical Feminist Analysis of Systems Theory.” *Public Relations Review* 19:157–166.
- Davis, K. Roscoe, Patrick G. McKeown, and Terry R. Rakes 1986 *Management Science*. Boston, Mass.: Kent.
- Garnsey, Elizabeth 1993 “Exploring a Critical Systems Perspective.” *Innovation* 6:229–256.
- Janeksela, Galan M. 1995 “General Systems Theory and Structural Analysis of Correctional Institution Social Systems.” *International Review of Modern Sociology* 25:43–50.
- Kraft, Philip 1977 *Programmers and Managers: The Routinization of Computer Programming in the United States*. New York: Springer-Verlag.
- Office of Technology Assessment 1977 *A Preliminary Assessment of the IRS Tax Administration System*. Washington, D.C.: Office of Technology Assessment.

- Pelton, Warren, Sonja Sackmann, and Robert Boguslaw 1990 *Tough Chokes: Decision-Making Styles of America's Top 50 CEO's*. Homewood, Ill.: Dow Jones-Irwin.
- Perrow, Charles 1984 *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.
- Quade, E. S. 1988 "Quantitative Methods: Uses and Limitations" In H. J. Miser and E. S. Quade, eds., *Handbook of Systems Analysis: Overview of Uses, Procedures, Applications and Practice*, pp. 283-324. New York: North-Holland.
- Rhile, Howard G. (March 22) 1990 "Progress in Meeting the Challenge of Modernizing IRS' Tax Processing System." Testimony before the Subcommittee on Oversight, Committee on Ways and Means, House of Representatives. Washington, D.C.: General Accounting Office.
- Searight, H. Russell and William T. Merkel 1991 "Systems Theory and Its Discontents: Clinical and Ethical Issues." *American Journal of Family Therapy* 19:19-31.
- Stichweh, Rudolf 1995 "Systems Theory and Rational Choice Theory; Systemtheorie und Rational Choice Theorie." *Zeitschrift fur Soziologie* 24:395-406.
- Turner, Jonathan H. 1991 *The Structure of Sociological Theory*, 5th ed. Belmont, Calif.: Wadsworth.
- von Bertalanffy, Ludwig 1968 *General Systems Theory: Foundations, Development, Applications*. New York: George Braziller.
- Wagner, Harvey M., Michael H. Rothkopf, Clayton J. Thomas, and Hugh J. Miser 1989 "The Next Decade in Operations Research: Comments on the CONDOR Report," *Operations Research* 37:664-672.
- Warren, Keith, Cynthia Franklin, and Calvin L. Streeter 1998 "New Directions in Systems Theory: Chaos and Complexity." *Social-Work* 43 (4):357-372.
- Whitehouse, Gary E., and Ben L. Wechsler 1976 *Applied Operations Research*. New York: Wiley.

ROBERT BOGUSLAW

# T

---

## TABULAR ANALYSIS

In its most general form, tabular analysis includes any analysis that uses tables, in other words, almost any form of quantitative analysis. In this article, however, it refers only to the analysis of categorical variables (both nominal and ordered) when that analysis relies on cross-classified tables in the form of frequencies, probabilities, or conditional probabilities (percentages). In general, the use of such cross-tabulated data is practical only with variables that have a limited number of categories. Therefore, this article deals with some of the *analytic* problems of *categorical data analysis*. Although it sometimes is difficult to separate analysis from methods of data presentation, the emphasis here is decidedly on analysis (see Davis and Jacobs 1968).

Tabular analysis can take many different forms, but two methods deserve special attention. The first is known as *subgroup analysis*. The underlying logic of this type of analysis was codified under the name “elaboration paradigm” by Lazarsfeld and his colleagues (Kendall and Lazarsfeld 1950; Lazarsfeld 1955; Hyman 1955; Rosenberg 1968; Lazarsfeld et al., 1972; Zeisel 1985). Because of the simplicity of the method and the ease with which it can facilitate communication with others, subgroup analysis has been the mainstay of research reports dealing with categorical data.

The second is based on the use of *log-linear* and related models and has become increasingly popular (Bishop et al. 1975; Goodman 1978;

Haberman 1978, 1979; Fienberg 1980; Agresti 1984; 1990; Clogg and Shihadeh 1994). (For other related models, see McCullagh and Nelder 1983; Thiel 1991; Long 1997; Press and Wilson 1998). This method is flexible, can handle more complex data (with many variables), and is more readily amenable to statistical modeling and testing (Clogg et al. 1990). For this reason, the log-linear method is rapidly emerging as the standard method for analyzing multivariate categorical data. Its results, however, are not easily accessible because the resulting tabular data are expressed as multiplicative functions of the parameters (i.e., log-linear rather than linear), and the parameters of these models tend to obscure descriptive information that often is needed in making intelligent comparisons (Davis 1984; Kaufman and Schervish 1986; Alba 1988; Clogg et al. 1990).

These two methods share a set of analytic strategies and problems and are complementary in their strengths and weaknesses. To understand both the promises and the problems of tabular analysis, it is important to understand the logic of analysis and the problems that tabular analyses share with the simpler statistical analysis of linear systems. As a multivariate analysis tool, tabular analysis faces the same problems that other well-developed linear statistical models face in analyzing data that are collected under less than ideal experimental conditions. It therefore, is important to have a full understanding of this foundation, and the best way to do that is to examine the simplest linear system.

**STATISTICAL CONTROLS, CAUSAL ORDERING, AND IMPROPER SPECIFICATIONS**

$$Y = Xd_{yx} + e \tag{2}$$

Consider the simplest linear multivariate system:

$$Y = Xb_{yxz} + Zb_{yzx} + e \tag{1}$$

where all the variables, including the error term, are assumed to be measured from their respective means. When this equation is used merely to describe the relationship between a dependent variable  $Y$  and two other variables  $X$  and  $Z$ , the issue of misspecification—in other words, whether the coefficients accurately reflect an intended relationship—does not arise because the coefficients are well-known partial regression coefficients. However, when the linear model depicted in equation (1) is considered as a representation of an underlying theory, these coefficients receive meaning under that theory. In that case, the issue of whether the coefficients really capture the intended relationship becomes important. Causal relationships are not the only important relationships, but it is informative to examine this equation with reference to such relationships since this is the implicitly implied type of system.

Many different conceptions of causality exist in the literature (Blalock 1964, 1985a, 1985b; Duncan 1966, 1975; Simon 1954, 1979; Heise 1975; Mostetler and Tukey 1977; Bunge 1979; Singer and Marini 1987). However, the one undisputed criterion of causality seems to be the existence of a relationship between manipulated changes in one variable ( $X$ ) and attendant changes in another variable ( $Y$ ) in an ideal experiment. That is, a causal connection exists between  $X$  and  $Y$  if changes in  $X$  and  $X$  alone produce changes in  $Y$ . This is a very restrictive criterion and may not be general enough to cover all important cases, but it is sufficient as a point of reference. This definition is consistent with the way in which effects are measured in controlled experiments. In general, even in an ideal experiment, it is often impossible to eliminate or control all the variations in other variables, but their effects are made random by design. A simple linear causal system describing a relationship produced in an ideal experiment thus takes the following familiar form:

where  $e$  stands for all the effects of other variables that are randomized. The randomization makes the expected correlation between  $X$  and  $e$  zero. (Without loss of generality, it is assumed that all the variables [ $X$ ,  $Y$ , and  $e$ ] are measured as deviations from their respective means.) For the sake of simplicity, it is assumed for now that  $Y$  does not affect  $X$ . (For an examination of causal models dealing with reciprocal causation and with more complex systems in general, see Fisher 1966; Goldberger and Duncan 1973; Alwin and Hauser 1975; Duncan 1975; Blalock 1985a, 1985b.)

The coefficient  $d_{yx}$  measures the expected change in  $Y$  given a unit change in  $X$ . It does not matter whether changes in  $X$  affect other variables and whether some of those variables in turn affect  $Y$ . As long as all the changes in  $Y$  ultimately are produced by the manipulated initial changes in  $X$  and  $X$  alone,  $X$  receives total credit for them. Therefore,  $d_{yx}$  is a coefficient of *total causal effect* (referred to as an *effect coefficient* for short).

The customary symbol for a simple regression coefficient,  $b_{yx}$  is not used in equation (2) because  $b_{yx}$  is equivalent to  $d_{yx}$  only under these very special conditions. If one uses a simple regression equation in the form similar to equation (2) above and assumes that  $b_{yx}$  is equivalent to  $d_{yx}$ , the model is misspecified as long as the data do not meet all the assumptions made about the ideal experiment. Such errors in model specification yield biased estimates in general. Implications of some specification errors may be trivial, but they can be serious when one is analyzing nonexperimental data (see Kish 1959; Campbell and Stanley 1966; Leamer 1978; Cook and Campbell 1979; Lieberman 1985; Arminger and Bohrnstedt 1987).

Many underlying causal systems are compatible with the three-variable linear equation shown above. For the purpose at hand, it is enough to examine the simple causal systems shown in Figure 1. These causal systems imply critical assumptions about the error term and the causal ordering. If these assumptions are correct, there is a definite connection between the underlying causal parameters and the regression coefficients in equation (1). However, if some of these assumptions are wrong, equation (1) is a misrepresentation of

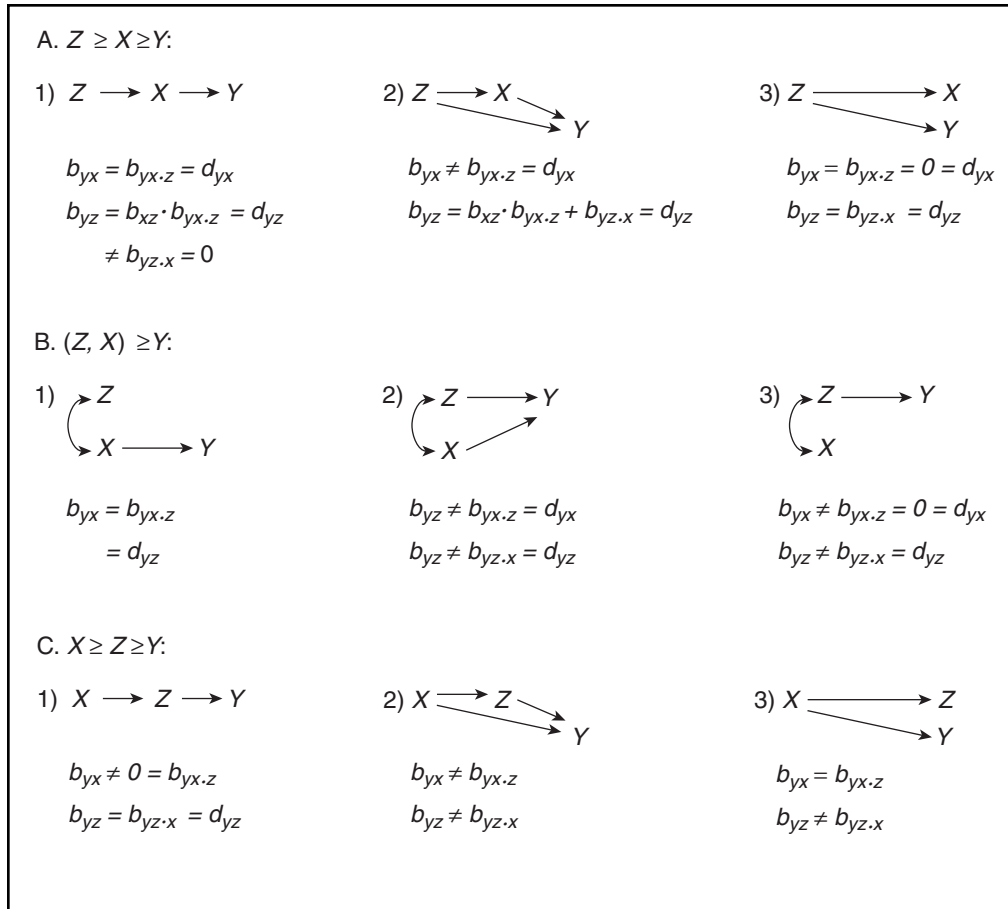


Figure 1. Some simplified linear causal systems

the assumed causal model (for a fuller description of other possible systems, see Duncan 1975).

The notation for causal hierarchy ( $\geq$ ) means that the preceding variable may affect the variables after it, but variables after ( $\geq$ ) may not affect the preceding variables. A connecting arrow between two variables indicates both the existence and the direction of effects; lack of a connecting arrow indicates no known effects. (For convenience, these diagrams do not show random errors, but their presence is assumed.)

For each causal system in Figure 1, the key relationships among simple regression coefficients, partial regression coefficients, and effect coefficients are listed below each causal diagram. Look at the simple causal chain (or a cascading system) shown in A1, for instance. The introduction of Z as a control variable has no effect on the observed

relationship between X and Y. Note also that the simple regression coefficient is equivalent to the effect coefficient ( $b_{yx} = b_{yx \cdot z} = d_{yx}$ ); similarly, the simple  $b_{yz}$  is equivalent to  $d_{yz}$ , but the partial  $b_{yz \cdot x}$  becomes zero. (If one were to control Y, the X-Z relationship would not change, but such control is superfluous given the assumptions about the causal ordering.) In fact, one could argue that these two conditions, given the assumptions about the causal hierarchy, uniquely define a simple causal chain. If the control variable Z enters the X-Y causal system only through X (or the effects of a set of variables are mediated completely through [an]other variable[s] in the system), there is no need to introduce Z (or a set of such variables) as a control to correctly specify the X-Y relationship.

In A2, the two partials ( $b_{yx \cdot z}$  and  $b_{yz \cdot x}$ ) are different from the respective bivariate coefficients ( $b_{yx}$



and  $b_{yz}$ ). The key point is that the partial  $b_{yxz}$  is equivalent to  $d_{yx}$ , while the partial between  $Z$  and  $Y$  ( $b_{yz\cdot x}$ ) simply reflects the portion of the causal effect from  $Z$  to  $Y$  that is not mediated by  $X$ .

In A3, there is no direct connection between  $X$  and  $Y$  once the effect of  $Z$  is controlled: The observed bivariate relation between  $X$  and  $Y$  is spurious or, more accurately, the observed association between  $X$  and  $Y$  is explained by the existence of a common cause. In this case, the introduction of  $Z$ , controlling its effects on both  $X$  and  $Y$ , is critical in ascertaining the true causal parameter of the system ( $d_{yx}$ ), which happens to be zero.

All the causal systems shown in B share similar patterns with A; the pattern of the relationship between the bivariate coefficients and the partials remains the same. For this reason, the  $X$ - $Y$  relationship in particular is examined in the same way by introducing  $Z$  as a control variable regardless of the specification of causal hierarchy between  $X$  and  $Z$ . Note in particular that introducing  $Z$  as a control variable in B1 and B3 is a misspecification of the model, but such misspecifications (including an irrelevant variable in the equation) do not lead to biased estimation (for a related discussion, see Arminger and Bohrnstedt 1987).

The systems shown in C do not require additional comments. Except for the changes in the order of the two variables  $X$  and  $Z$ , they are exact replicas of the systems in A. The resulting statistics show the same patterns observed in A. Nevertheless, the attendant interpretation of the results is radically different. For instance, when the partial  $b_{yxz}$  disappears, one does not consider that there is no causal relationship between  $X$  and  $Y$ ; instead, one's conviction about the causal relationship is reinforced by the fact that an intervening causal agent is found.

In summary, the assumptions about the causal ordering play a critical role in the interpretation of the coefficients of the linear model shown in equation (1). The assumptions about the order must come from outside knowledge.

There is one more type to note. All the systems examined so far are linear and additive. The partial coefficients reflect the expected change in the dependent variable given a unit change in a given independent variable while the other independent variables are kept constant. If two or more inde-

pendent variables interact, such simplicity does not exist. A simple example of such a system is given below:

$$Y = X_1 \cdot b_1 + X_2 \cdot b_2 + (X_1 \cdot X_2) \cdot b_3 + e \quad (3)$$

which is the same as equation (1) except for the simplification of labels for the variables and coefficients and the addition of a multiplicative term ( $X_1 \cdot X_2$ ).

The partial for  $X_i$  in such a system, for example, no longer properly represents the expected change in  $Y$  for a unit change in  $X_i$ , even if the assumptions about the causal order are correct. A partial differentiation of the equation with respect to  $X_i$  for instance, gives  $b_1 + X_2 \cdot b_3$ , which implies that the rate of change introduced by a change in  $X_i$  is also dependent on the values of the other causal variable ( $X_2$ ) and the associated coefficient ( $b_3$ ). One therefore cannot interpret the individual coefficients as measuring something independently of others. This point is important for a fuller understanding of the log-linear models introduced below, because a bivariate relationship is represented by interaction terms. The notion of control often invoked with *ceteris paribus* (other things being unchanged) also becomes ambiguous.

The logic of causal analysis for the additive systems can be extended easily to a system with more variables. If the assumptions about the causal order, the form of the relationship, and the random errors are correct, one can identify the causal parameters, such as  $d_{yx}$ , and decompose the linear connection between any set of variables into spurious (noncausal) and genuine (causal) components,  $d_{yx}$ , and the latter ( $d_{yx}$ ) into indirect (mediated) and direct (residual) components.

To identify  $d_{yx}$ , one must control all the potentially relevant variables that precede  $X$  in causal ordering but not the variables that might intervene between  $X$  and  $Y$ . Under this assumption, then, the partial  $b_{yx}(z\cdots)$ , where the variables in parentheses represent all such "antecedent" variables, is equivalent to  $d_{yx}$ . In identifying this component, one must not control the variables that  $X$  may affect; these variables may work as mediating causal agents and transmit part of the effect of  $X$  to  $Y$ .

The partial of a linear system in which both antecedent variables ( $Z$ s) and intervening vari-

ables ( $W$ s) are included ( $b_{yx} \cdot [x \cdots w \cdots]$ ) will represent the residual causal connection between  $X$  and  $Y$  that is not mediated by any of the variables included in the model. As more  $W$ s are included, this residual component may change. However, the linear representation of a causal system without these additional intervening variables is not misspecified. By contrast, if the introduction of additional  $Z$ s will change the  $X$ - $Y$  partial, an omission of such variables from the equation indicates a misspecification of the causal system because some of the spurious components will be confounded with the genuine causal components.

For nonexperimental data, the problems of misspecification and misinterpretation are serious. Many factors may confound the relationships under consideration (Campbell and Stanley 1966; Cook and Campbell 1979; Lieberman 1985; Arminger and Bohrnstedt 1987; Singer and Marini 1987). There is no guarantee that a set of variables one is considering constitutes a closed system, but the situation is not totally hopeless. The important point is that one should not ignore these issues and assume away potentially serious problems. Selection biases, contagion effects, limited variations in the data, threshold effects, and so on, can be modeled if they are faced seriously (Rubin 1977; Leamer 1978; Hausman 1978; Heckman 1979; Berk 1983, 1986; Heckman and Robb 1986; Arminger and Bohrnstedt 1987; Long 1988; Bollen 1989; Xie 1989). Furthermore, this does not mean that one has to control (introduce) every conceivable variable. Once a few key variables are controlled, additional variables usually do not affect the remaining variables too much. (This observation is a corollary to the well-known fact that social scientists often have great difficulty finding any variable that can substantially improve  $R^2$  in regression analysis.)

**FREQUENCY TABLES, CONDITIONAL PROBABILITIES, AND ODDS RATIOS**

To fix the ideas and make the following discussions concrete, it is useful to introduce basic notations and define two indicators of association for a bivariate table. Consider the simplest contingency table, one given by the cross-classification of two dichotomous variables. Let  $f_{ij}$  denote the observed frequencies; then the observed frequency distribution will have the following form:

		Observed Frequencies		
		Variable X		
		1	2	Total
Variable Y	1	$f_{11}$	$f_{12}$	$f_{1\cdot}$
	2	$f_{21}$	$f_{22}$	$f_{2\cdot}$
total		$f_{\cdot 1}$	$f_{\cdot 2}$	$N$

Note the form of marginal frequencies. Now let  $p_{ij}$  denote the corresponding observed probabilities:  $p_{ij} = f_{ij}/N$ . Let the uppercase letters,  $F_{ij}$  and  $P_{ij}$ , denote the corresponding expected frequencies and probabilities under same model or hypothesis.

If  $X$  and  $Y$  are statistically independent,

$$\frac{p_{ij}}{p_{\cdot j}} = \frac{p_i \cdot p_j}{p_{\cdot j}} = p_i$$

That is, the conditional probability of  $Y_i$  given  $X_j$  is the same as the marginal probability of  $Y_i$ . Thus, a convenient descriptive indicator of statistical independence is that  $b_{yx} = p_{11}/p_{\cdot 1} - p_{12}/p_{\cdot 2} = 0$ . The percentage difference is simply 100 times  $b_{yx}$ . The symbol  $b_{yx}$  is quite appropriate in this case, for it is equivalent to the regression coefficient. The fact that  $b_{yx} \neq 0$  implies a lack of statistical independence between  $X$  and  $Y$ .

Another equally good measure is the odds ratio or cross-product ratio:

$$\begin{aligned} \text{Odds ratio } (t) &= \frac{F_{11}/F_{12}}{F_{21}/F_{22}} \\ &= \frac{F_{11}/F_{21}}{F_{12}/F_{22}} \\ &= \frac{F_{11}/F_{22}}{F_{12}/F_{21}} \end{aligned}$$

The first line shows that the odds ratio is a ratio of ratios. The second line shows that it is immaterial whether one starts with odds (ratio) in one direction or the opposite direction. The final line indicates that the odds ratio is equivalent to the cross-product ratio. In general, if all the odds ratios in a table for two variables are 1, the two variables are statistically independent; the converse is also true. The fact that  $t$  equals 1 implies that  $X$  is independent of  $Y$ . Therefore, both the odds ratio ( $t$ ) and the percentage difference ( $b_{yx}$ ) can serve equally

well as descriptive indicators of association between variables.

Given that observed frequencies are unstable because of sampling variability, it is useful to test the null hypothesis that  $t = b_{yx} = 0$  in the population. Such a hypothesis is evaluated by using either the conventional chi-square statistic or the  $-2^*$ (likelihood ratio):

$$\begin{aligned} \chi^2 &= \sum \sum (f_{ij} - F_{ij})^2 / F_{ij} \\ L^2 &= -2 \sum \sum (f_{ij} \log(F_{ij}/f_{ij})) \\ &= 2 \sum \sum (f_{ij} \log(f_{ij}/F_{ij})) \end{aligned}$$

These values are evaluated against the theoretical distribution with the appropriate degrees of freedom. These two tests are equivalent for large samples. (For a related discussion, see Williams 1976; Tamas et al. 1994)

**ELABORATION AND SUBGROUP ANALYSIS**

The logic of linear systems that was presented earlier was introduced to social scientists through the elaboration paradigm and through an informal demonstration of certain patterns of relationship among variables (Kendall and Lazarsfeld 1950; Lazarsfeld 1955). Statistical control is achieved by examining relationships within each subgroup that is formed by the relevant categories of the control variable. The typical strategy is to start the analysis with an examination of the association between two variables of interest, say, *X* and *Y*. If there is an association of some sort between *X* and *Y*, the following two questions become relevant: (1) Is the observed relationship spurious or genuine? (2) If some part of the relationship is genuine, which variables mediate the relationship between the two? (The question of sampling variability is handled rather informally, relying on the magnitude of the percentage differences as a simple guide. Moreover, two variables that seemingly are unrelated at the bivariate level may show a stronger association after suppressor variables are controlled. Therefore, in some situations, applying such a test may be premature and uncalled for.)

To answer these questions adequately, one must have a fairly good knowledge of the variables under consideration and the implications of dif-

ferent causal systems. It is clear from the earlier examination of the linear causal systems that to answer the first question, one must examine the *X*-*Y* relationship while controlling for the factors that are antecedent to *X*(assuming that  $X \geq Y$ ). To answer the second question, one also must control factors that *X* may affect and that in turn may affect *Y*. Controlling for many variables is possible in theory but is impractical for two quite different reasons: (1) One runs out of cases very quickly as the number of subgroups increases, and (b) as the number of subgroups increases, so does the number of partial tables to examine and evaluate. Nevertheless, it is quite possible that one might find a strategically critical variable that might help explain the observed relationship either by proving that the observed relationship is spurious or by confirming a causal connection between the two variables.

To make the discussion more concrete, consider the hypothetical bivariate percentage table between involvement in car accidents (*Y*) and the gender of the driver (*X*). The percentage difference (10% = 30% - 20%) indicates that men are more likely to be involved in car accidents while driving than are women. Because there are only two categories in *Y*, this percentage difference ( $b_{yx}$ ) captures all the relationship in the table. Given the large sample size and the magnitude of the percentage difference, it is safe to assume that this is not an artifact of sampling variability.

Suppose a third variable (*Z* = amount of driving) is suspected to be related to both gender (*X*) and involvement in accidents (*Y*). It therefore is prudent to examine whether the *X*-*Y* relationship remains the same after the amount of driving is controlled or eliminated. Whether this conjecture is reasonable can be checked before one examines the three-variable subgroup analysis: There has to be some relationship between *X* and *Z* and between *X* and *Y*. Table 1b shows the bivariate relationship between gender (*X*) and driving (*Z*). Note that there is a very strong association:  $b_{yx} = .333$  (33.3%) difference between the genders.

The conditional tables may show one of the following four patterns: (1) The observed relationship between *X* and *Y* disappears within each subgroup:  $b_{yxz} = 0$ , (2) the relationship remains the same:  $b_{yxz} = b_{yx}$ , (3) the relationships change in magnitude but remain the same across the groups:

**Hypothetical Bivariate Tables**

	Men	Women
a) Car Accidents (Y) by Gender (X)		
Had at least one accident while driving	30%	20%
Never had an accident while driving	<u>70%</u>	<u>80%</u>
Total	100%	100%
(Number of cases)	(3,000)	(3,000)
b) Amount of Driving (Z) by Gender (X)		
More than 10,000 miles	67.7%	33.3%
Less than 10,000 miles	<u>33.3%</u>	<u>67.7%</u>
Total	100%	100%
(Number of cases)	(3,000)	(3,000)

Table 1

SOURCE: Adapted from Ziesel (1985), p. 146.

$b_{yxz(1)} = b_{yxz(2)} \neq b_{yx}$ , (4) the X-Y relationship in one group is different from the relationship in the other group:  $b_{yxz(1)} \neq b_{yxz(2)}$ . These examples are shown in Table 2. Compare these patterns with the corresponding causal systems shown in Figure 1.

Whether Z should be considered as antecedent or intervening depends on the theory one is entertaining. One's first interpretation might be that the original relationship has sexist implications in that it may mean that men are either more aggressive or less careful. Against such a hypothesis, the amount of driving is an extraneous variable. By contrast, one may entertain a social role theory stating that in this society men's roles require more driving and that more driving leads to more accidents. Then Z can be considered an intervening variable.

Pattern (1) will help undermine the psychological or biological hypothesis, and pattern (2) will enhance that hypothesis. Pattern (1) also will lend weight to the social role hypothesis. These patterns are the simplest to deal with but rarely are encountered in real life (see Lazarsfeld 1955; Rosenberg 1968; Zeisel 1985 for interesting examples). If one were lucky enough to come across such a pattern, the results would be considered important findings. Note that there are three causal systems in Figure 1 that share the same statistical pattern (the relationship between partials and origi-

nal coefficients) with each of these two. Of course, the choice must be dictated by the theory and assumptions about the causal ordering that one is willing to entertain.

Patterns (3) and (4) are more likely outcomes in real life. In (3), the magnitude of the X-Y relationship within each subgroup is reduced. (Sometimes the X-Y relationship may turn out to be even stronger.) This pattern is compatible with three causal systems—A2, B2, and C2—in Figure 1. Assume that one takes the causal order indicated in C; that is, one takes the gender role theory to account for the observed relationship. Part of the original relationship (.04 out of .10) is mediated by the amount of driving, but a greater part (.06) remains unexplained. If one believes that all the difference in the accident rate has nothing to do with psychological or biological differences between the genders, one has several other potential role-related connections to consider: Men may drive more during the rush hours than women do, men may drive during worse weather conditions than women do, and so on. One could introduce these variables as additional controls. By contrast, if one believes in the validity of the psychological explanation, one could collect data on the aggressiveness of each individual and introduce aggressiveness as a control variable.

Table 2d illustrates a pattern in which the effects of the two explanatory variables interact: X's effect on Y varies across the categories of Z, and Z's effect on Y varies across the categories of X. A corresponding example in linear systems was given by equation (3). One must consider both variables at the same time because the effect of one variable depends on the other.

In general, empirical data may exhibit patterns that are mixtures of 2c and 2d. In cross-tabulations of variables with more than two categories, it is often not easy, purely on the basis of eyeballing, to discern the underlying pattern. At this point, there is a need for more refined and systematic tools. Moreover, in some instances, an application of a log-linear model may indicate patterns that are different from what a linear model (such as using percentage tables) might indicate.

Before ending this section, it should be mentioned that some examples in the literature use the

TABULAR ANALYSIS

**Percent Ever Had Accident (Y) by Gender (X) by Amount of Driving (Z)**

a) Original X–Y Relationship Disappears

(Compatible with causal systems A3, B3, and C1)

Gender (X)	Amount of Driving (Z)		
	> 10,000 miles	< 10,000 miles	
Men	40% (2,000)	10% (1,000)	$b_{yxz} = 0$
Women	40% (2,000)	10% (1,000)	$b_{yzx} = .30$

b) Original X–Y Relationship Unchanged

(Compatible with causal systems A1, B1, and C3)

Gender (X)	Amount of Driving (Z)		
	> 10,000 miles	< 10,000 miles	
Men	30% (2,000)	30% (1,000)	$b_{yxz} = .10$
Women	20% (1,000)	20% (2,000)	$b_{yzx} = 0$

c) Original X–Y Relationship Diminishes

(Compatible with causal systems A2, B2, and C2)

Gender (X)	Amount of Driving (Z)		
	> 10,000 miles	< 10,000 miles	
Men	34% (2,000)	24% (1,000)	$b_{yxz} = .06$
Women	28% (1,000)	18% (2,000)	$b_{yzx} = .10$

d) X–Y Relationship Varies

Gender (X)	Amount of Driving (Z)		
	> 10,000 miles	< 10,000 miles	
Men	40% (2,000)	20% (1,000)	$b_{yxz(1)} = .20$
Women	20% (1,000)	20% (2,000)	$b_{yxz(2)} = 0$
			$b_{yzx(1)} = .20$
			$b_{yzx(2)} = 0$

Table 2

NOTE: Number of cases for the percentage base are in parentheses. Throughout these tables,  $b_{xz} = .40$  and  $b_{yx} = .10$  remain constant. Compare percents across the categories of that variable.

subgroup analysis as a full-fledged multivariate analysis tool. For instance, Davis (1984) shows how the logic of elaboration can be combined with the standardization technique to derive, among other things, the following decomposition of the relationship between the father’s and the son’s occupational statuses, where Zs represent the father’s education and the mother’s education and W represents the son’s education.

- a. Total observed relationship:  
 $b_{yx} = .256$
- b. Spurious connection resulting from environmental variables (Zs) (a-c) .052
- c. Total causal effect:  $b_{yxz...} = .204$

c1. Unmediated causal effect:

$$b_{yxz...w} = .138$$

c2. Effect mediated by education

$$(b_{yxz...} - b_{yxz...w}) = .066$$

The power of subgroup analysis comes mainly from the close analogy between the percentage differences and the coefficients of the linear system illustrated in Figure 1, but its uses need not be confined to the analysis of causal systems. There are various applications of this logic to survey data (Hyman 1955; Rosenberg 1968; Zeisel 1985). These accounts remain one of the best sources for learning the method as well as the art of pursuing research ideas through the use of percentage tables.

**ODDS RATIOS AND LOG-LINEAR MODELS**

A more formal approach to categorical data analysis is provided by the log-linear model and related models (Bishop et al. 1975; Goodman 1978; Haberman 1978, 1979; Fienberg 1980; Agresti 1984; Clogg and Shihadeh 1994; Long 1997). Some of these models are not even log-linear (Clogg 1982a, 1982b, Goodman 1984, 1990; Wong 1995; Xie 1992). Only the log-linear models are examined here.

By means of an ingenious device, the log-linear model describes the relationships among categorical variables in a linear form. The trick is to treat the logarithms of the cell frequencies as the (titular) dependent variable and treat design vectors as independent variables. The design vectors represent relevant features of the contingency table and hypotheses about them.

Once again consider a concrete example; the simplest bivariate table, in which each variable has only two categories. Such a table contains four frequencies. Logarithms of these frequencies (log-frequencies for short) can be expressed as an exact function of the following linear equation:

$$Y = b_0 + X_1 \cdot b_1 + X_2 \cdot b_2 + (X_1 \cdot X_2) b_3 \quad (4)$$

In this equation,  $Y$  stands for the log-frequencies ( $\log(F_{ij})$ ).  $X_1$  is a design vector for the first (row) variable, and  $X_2$  is a design vector for the second (column) variable. The last vector ( $X_1 \cdot X_2$ ) is a design vector for interaction between  $X_1$  and  $X_2$ , and it is produced literally by multiplying the respective components of  $X_1$  and  $X_2$ . It is important to note that the model is linear only in its parameters and that there is an interaction term. As is the case with linear models that contain interaction terms, one must be careful in interpreting the coefficients for the variables involved in the interaction term.

This type of model in which the observed frequencies are reproduced exactly also is known as a saturated model. (The model is saturated because all the available degrees of freedom are used up. For instance, there are only four data points, but this model requires that many parameters.) Of course, if one can reproduce the exact log-frequencies, one also can reproduce the actual frequencies by taking the exponential of  $Y - F_{ij} =$

$\exp(Y_{ij})$ . Note also the similarities between equations (3) and (4); both contain a multiplicative term as a variable. (For more general models, a maximum likelihood estimation requires an iterative solution, but that is a technical detail for which readers should consult standard texts (such as Nelder and Wedderburn 1972; Plackett 1974; Goodman 1978, 1984; Haberman 1978, 1979; Fleiss 1981; Agresti 1984). Many computer packages routinely provide solutions to these types of equations. Therefore, what is important is the logic underlying such analysis, not the actual calculation needed.)

It is no exaggeration to say that in more advanced uses of the model, what distinguishes a good and creative analysis from a mundane analysis is how well one can translate one's substantive research ideas into appropriate design vectors. Thus, it is worthwhile to examine these design vectors more carefully. Constructing a design matrix (the collection of vectors mentioned above) for a saturated model is easy, because one is not pursuing any specific hypothesis or special pattern that might exist in the relationship. Categories of each variable have to be represented, and there are many equivalent ways of doing that. This section will examine only the two most often used ones: effect coding and dummy coding. These design matrices for a  $2 \times 2$  table are shown in Table 3.

The first column ( $X_0$ ) in each coding represents a design vector for the constant term ( $b_0$ );  $X_1$  is for the row categories, and  $X_2$  is for the column categories. The last column ( $X_3$ ) is the product of the preceding two, needed to represent interaction between  $X_1$ , and  $X_2$ . Note the pattern of these design vectors. In the effect coding, except for the constant vector, each vector or column sums to zero. Moreover, the interaction vector sums to zero for each column and row of the original bivariate table. This pattern assures that each effect is measured as a deviation from its respective mean.

In dummy coding, the category effect is expressed as a deviation from one reference category, in this case, the category that is represented by zero. Whatever codings are used to represent the categories of each variable, the interaction design vector is produced by multiplying the design vector for the column variable by the design

**Design Vectors Used in Log-Linear Model for 2 × 2 Table**

a) Design Matrices for Saturated Model

Frequency	Effect Coding				Dummy Coding			
	$X_0$	$X_1$	$X_2$	$X_3$	$X_0$	$X_1$	$X_2$	$X_3$
$Y_{11}$	1	1	1	1	1	1	1	1
$Y_{12}$	1	1	-1	-1	1	1	0	0
$Y_{21}$	1	-1	1	-1	1	0	1	0
$Y_{22}$	1	-1	-1	1	1	0	0	0

b) Representation of Log-Frequencies in Terms of Parameter

$b_0+b_1+b_2+b_3$	$b_0+b_1-b_2-b_3$	$b_0+b_1+b_2+b_3$	$b_0+b_1$
$b_0-b_1+b_2-b_3$	$b_0-b_1-b_2+b_3$	$b_0+b_2$	$b_0$

c) Representation of Frequencies in Terms of Multiplicative Parameters, where  $t_i = \exp(b_i)$

$t_0^*t_1^*t_2^*t_3$	$t_0^*t_1/(t_2^*t_3)$	$t_0^*t_1^*t_2^*t_3$	$t_0^*t_1$
$t_0^*t_2/(t_1^*t_3)$	$t_0^*t_3(t_1^*t_2)$	$t_0^*t_2$	$t_0$

d) Parameters for Interaction in Log-Linear Model

$b_3$	$-b_3$	$b_3$	0
$-b_3$	$b_3$	0	0

Log (odds ratio)

$4*b_3$	$b_3$
---------	-------

e) Multiplicative Parameter for Interaction ( $t_3 = \exp(b_3)$ )

$t_3$	$1/t_3$	$t_3$	1
$1/t_3$	$t_3$	1	1

Odds ratio

$t_3^*t_3^*t_3^*t_3 = t_3^4$	$t_3$
------------------------------	-------

Table 3

vector for the row variable. Normally, one needs as many design vectors for a given variable as there are categories, minus one: (R-1) for the row variable and (C-1) for the column variable. In that case, there will be (C-1)(R-1) interaction design vectors for the saturated model. These interaction vectors are created by cross-multiplying the vectors in one set with those of the other set. There is only one vector for each of the three independent variables in equation (4) because both variables are dichotomous.

The names for these codings come from the fact that the first coding is customarily used as a convenient way of expressing factor effects in an

analysis of variance (ANOVA), while the second coding often is used in regression with dummy variables. As a result of coding differences in the representation of each variable, the constant term in each coding has a different meaning: In effect, coding it measures the unweighted grand mean, while in dummy coding, it measures the value of the category with all zeros (in this particular case,  $Y_{22}$ ). (For other coding schemes, see Haberman 1979; Agresti 1984; Long 1984.) Some parameter estimates are invariant under different types of coding, and some are not (Long 1984); therefore, it is important to understand fully the implications of a particular design matrix for a proper interpretation of the analysis results.



Panel (b) of Table 3 expresses each cell as a product of the design matrix and corresponding parameters. Since the particular vectors used contain 1, -1, or 0, the vectors do not seem to appear in these cell representations. However, when design vectors contain other numbers (as will be shown below), they will be reflected in the cell representation. Panel (c) is obtained by exponentiation of the respective cell entries in (b), the individual  $t$ -parameter also being the corresponding exponential of the log-linear parameter in panel (b).

Panel (d) isolates parameters associated with the interaction design vector. Panel (e) contains corresponding antilogs or multiplicative coefficients. These parameters play a critical role in representing the degree and nature of association between the row variables and the column variables. If all the odds ratios are 1, one variable is statistically independent from the other; in other words, information about the association between variables is totally contained in the pattern of odds ratios. Panels (d) and (e) show that the odds ratio in turn is completely specified by the parameter(s) of the interaction vector(s). In forming the odds ratio, all the other parameters cancel out (in logarithms, multiplication becomes addition and division becomes subtraction).

In short, this is an indirect way to describe a pattern of association in a bivariate table. Unfortunately, doing this requires a titular dependent variable and multiplicative terms as independent variables. Also, in effect coding, the log-odds ratio is given by  $4 \times b_3$ , but in dummy coding, it is given by  $b_3$ . This is a clear indication that one cannot assume that there is only one way of describing the parameters of a log-linear model. These facts make the interpretation of these parameters tricky, but the process is worth it for two reasons.

First, the advantage of this method for analyzing a  $2 \times 2$  table is trivial, but the model can be generalized and then applied to more complex contingency tables. Because of the ANOVA-like structure, it is easy to deal with higher-level interaction effects. Second, the parameters of the log-linear models (obtained through the likelihood procedure) have very nice sampling properties for large samples. Therefore, better tools for statistical testing and estimating are available. Without this second advantage, the fact that the approach allows the construction of ANOVA-like models

may not be of much value, for the log-linear models only indirectly and by analogy reflect the relationship between variables.

Consider the bivariate tables in Table 4. In all these tables, the frequencies are such that they add up to 100 in each column. Thus, one can take these frequencies as percentages as well. The first table shows a 20 percent difference and an odds ratio of 2.25. The second table shows only half the percentage difference of the first but the same odds ratio. The last table shows the same percentage difference as the second one, but its odd ratio is greater at 6.68. These descriptive measures indicate that there is some association between the two variables in each table.

Whether this observed association is statistically significant can be tested by applying a model in which the coefficient for the interaction design vector is constrained to be zero. (Here one is utilizing the properties of the log-linear model that were asserted earlier.) Constraining the interaction parameter to zero is the same as deleting the interaction design vector from the model. This type of a design matrix imposes the model of statistical independence (*independence* model for short) on the data. If such a log-linear model does not fit the data (on the basis of some predetermined criteria), the observed association is accepted as significant. For large samples, both the conventional chi-square test and the likelihood ratio ( $L^2$ ) test can be used for this purpose. The results of these tests are included in each table, and they indicate that all three associations are statistically significant at the conventional  $\alpha$  level of .05.

Thus, to describe fully the underlying pattern of the association in Table 4, one needs to introduce the interaction parameter, which in these cases is the same as it is using the saturated model. The right-hand tables show the multiplicative parameters ( $t$ -parameters) for the interaction term. (Here only the results of applying effect coding are included.) First, examine the patterns of these parameters. In each of the three tables, the  $t$ -parameters indicate that the main diagonal cells have higher rates than do the off-diagonal cells. This tendency is slightly higher in the last table than it is in the first two. This interpretation follows from the fact that to reproduce the observed frequency in each cell, the respective  $t$ -



TABULAR ANALYSIS

Odds Ratios (t) and Percentage Differences							
FREQUENCIES				MULTIPLICATIVE PARAMETERS			
a)		$X_1$	$X_2$	<i>Effect Coding</i>		<i>Dummy Coding</i>	
	$Y_1$	60	40	1.225	.816	2.25	1
	$Y_2$	<u>40</u>	<u>60</u>	.816	1.225	1	1
		100	100				
	$b_{yx} = .20$	$t = 2.25$					
	$L^2 = 8.05$	$p = .005$					
b)		$X_1$	$X_2$	<i>Effect Coding</i>		<i>Dummy Coding</i>	
	$Y_1$	20	10	1.225	.816	2.25	1
	$Y_2$	<u>80</u>	<u>90</u>	.816	1.225	1	1
		100	100				
	$b_{yx} = .10$	$t = 2.25$					
	$L^2 = 3.99$	$p = .046$					
c)		$X_1$	$X_2$	<i>Effect Coding</i>		<i>Dummy Coding</i>	
	$Y_1$	12	2	1.608	.622	6.68	1
	$Y_2$	<u>88</u>	<u>98</u>	.622	1.608	1	1
		100	100				
	$b_{yx} = .10$	$t = 6.68$					
	$L^2 = 8.46$	$p = .004$					

Table 4

parameter must be multiplied to whatever value may be implied by other parameters in the model. In the first and second tables, the frequencies in the main diagonal are about 22 percent higher (1.22 times) than they would be without the interaction effect. The frequencies in the off-diagonal cells are about 18 percent lower than they otherwise would be. If one were to examine only the statistics generated by log-linear models, however, it would be easy to overlook the fact that the percentage of the first cell in the last table is only 12 percent (see Kaufman and Schervish 1986 for a more extended discussion). This is one of the reasons why it is advisable to examine the percentage tables even if one is using the log-linear model almost exclusively.

There are other reasons, too. By the linear standard (percentage difference), the first table shows a greater degree of association than does the second or the third. By the standard of a log-linear model or odds ratio, the last table shows the

greatest degree of association. In most cases, where the percentages remain within the range of 20 to 80 percent, these two standards are roughly comparable, and the linear and log-linear models may produce similar results (see Goodman 1981). More important, in examining three-way interactions, if two subtables have the patterns shown in Table 4a and 4b, log-linear models will indicate no three-way interaction, while linear models will indicate it. There are models in which a particular standard is justified explicitly by the phenomenon under consideration, but one should not adopt a standard merely because a particular statistical model does so. It is important to understand the differences in the implicit standards that are used in different methods.

**SOME MODELS OF ASSOCIATION**

The flexibility of log-linear models is not obvious until one deals with several variables. However, even in a bivariate table, if there is an underlying

order in the categories of variables involved and the pattern of association, the model allows some flexibility for exploring this pattern. Consider the hypothetical table shown in Table 5a. The marginal totals are such that each column may be read as percentages. There is a definite pattern in the direction of the relationship, although the tendency is fairly weak. If one were to apply a test of independence, such a null hypothesis would not be rejected. ( $L^2 = 6.56$  with four degrees of freedom has a probability of .161.) Against an unspecified alternative hypothesis, the null hypothesis cannot be rejected at the conventional level of  $\alpha$ .

Knowing that almost every society values these two variables in the same order, one may expect that the underlying pattern of association reflects the advantages of the upper class over the lower class in obtaining valued objects. Both the pattern of percentage differences and the odds ratios seem to indicate such an ordering in the pattern: The advantage the upper class has over the lower class is greater than the one it has over the middle class. Furthermore, the upperclass does better in relation to educational levels that are farther apart (the odds ratio involving the corner cells is 2.87).

A conjecture or hypothesis like this can be translated into a design vector. Assign any consecutive numbers to the categories of each variable, but to be consistent with the effect coding, express them as deviations from the mean. One such scaling is to use  $(R+1)/2-i$  for the row variable and  $(C+1)/2-j$  for the column variable. (The mean and category values can be reversed, but this scheme assigns a higher value to a higher class and a higher educational level to be consistent with everyday language.) Recalling once again that only the interaction terms are relevant for the description of association, one needs to create such an interaction term by multiplying these two vectors component by component. An example is shown in Table 6.

The log-linear model, then, will include design vectors for the constant term, two vectors for the row and two vectors for the column, and one vector for the "linear-by-linear" interaction. This type of model is known as a *uniform association* model (for a fuller discussion of this and related models, see McCullagh 1978; Haberman 1979; Clogg 1982a, 1982b; Anderson 1984; Goodman 1984, 1985, 1987, 1990, 1991 Clogg and Shihadeh 1994). The results of applying such a model to

Table 5a are presented in Table 5b and 5c. First, this model fits the data extremely well. Moreover, the reduction of the  $L^2$  statistic ( $6.557 - .008 = 6.549$ ) with one degree of freedom is statistically significant. Therefore, the null hypothesis cannot be accepted against this specific alternative hypothesis (see Agresti 1984 for a fuller discussion of hypothesis testing of this type). Note the pattern of the expected frequencies and the interaction parameters. Both indicate that the odds ratio for every consecutive four cells is uniform. Moreover, the other odds ratios are exact functions of this basic odds ratio and the distances involved. For instance, the odds ratio for the four corner cells is  $2.87 = 1.303^{2 \times 2}$ , with each exponent indicating the number of steps between respective categories in each variable. A degree of parsimony has been achieved in describing the pattern of association and some statistical power has been gained in proposing a more definite alternative hypothesis than the general one that stipulates any lack of statistical independence (and hence uses up four degrees of freedom).

The introduction of different design matrices allows one to explore different patterns very easily. Just two are examined here. Consider the hypothetical tables shown in Table 7. In the first table, the odds ratios remain the same across the columns but vary across the rows, perhaps indicating that the order inherent in the row categories is not uniform, while that in the column category is. Differently stated, the *distance* between two consecutive row categories varies, while it remains constant for the column categories. Such an association pattern is known as the *row-effects association* model not because the column variable does not have any effect but because an equal-interval scale works well for it. In this case, one needs two design vectors to accommodate the unequal distances in the row categories. In general, the most one needs is the number of categories in the row minus one. As is shown in Table 6, these design vectors are obtained by cross-multiplying the linear distance vector of the column and the two vectors that already have been used to represent the row categories. (It works just as well to use the dummy coding.) The *column-effects* model is obtained if one reverses the role of these variables.

Table 7b is an example of the simplest possible *homogeneous row-column effects* model. The odds

TABULAR ANALYSIS

**A Hypothetical Table: Level of Educational Attainment (Y) by Social Class (X)**

a) Observed Table

Level of Education	Social Class		
	High	Middle	Low
College	17	12	8
High school	41	38	34
Less than high school	<u>42</u>	<u>50</u>	<u>58</u>
Total	100	100	100

b) Expected Frequencies under the Assumption of Independence

12.33	12.33	12.33	$L^2 = 6.56$
37.67	37.67	37.67	$df_1 = 4$
50.00	50.00	50.00	$p = .161$

c) Expected Frequencies under the Assumption of Uniform Association

16.90	11.90	8.20	$L^2 = .0082$
41.30	38.00	33.80	$df_2 = 3$
41.80	50.10	58.10	$p = .161$

d) Log-Linear and Multiplicative Parameters

.264	0	-.264	1.303	1	.768
0	0	0	1	1	1
-.264	0	.264	.768	1	1.303

$L_1^2 - L_2^2 = 6.48;$

$df_1 - df_2 = 1;$

$p = .0109$

Table 5

ratios change across the row and across the column, but the corresponding pair of categories in the row and in the column share the same odds ratio. In this particular example, there is a greater distance between the first two categories than there is between the second two. In general, a homogeneous row-column effects model can accommodate different intervals in each variable as long as the corresponding intervals are homogeneous across the variables. The design matrix for such a pattern is easily obtained by adding the row-effects model vectors and the column-effects model vectors. This is also how two variables are constrained to have equal coefficients in any linear model. Such a design matrix for a 3x3 table is also contained in Table 6. The examples shown in that

table should be sufficient to indicate strategies for generalizing to a larger table.

There are many other possibilities in formulating specific hypotheses. These relatively simple models are introduced not only for their intrinsic value but also as a reminder that one can incorporate a variety of specialized hypotheses into the log-linear model (for other possibilities, see Goodman 1984; Clogg 1982a, 1982b; Agresti 1983, 1984). Before ending this section, it should be noted that when design vectors such as the ones for the homogeneous row-column effects model are used, the connection between the parameters for linear models indicated in this article and the usual ANOVA notation used in the literature is not

TABULAR ANALYSIS

**Design Matrices for Row-Column Association Models for 3 × 3 Table**

<i>T</i>	<i>A</i>		<i>B</i>		<i>C</i>	<i>D</i>	<i>E</i> <i>C*D</i>		<i>F</i> <i>A*~D</i>		<i>G</i> <i>B*~D</i>		<i>H</i> <i>F+G</i>		<i>I</i> <i>(F~G)†</i>		<i>J</i> <i>A*~B</i>			
1	1	0	1	0	1	1	1	1	0	1	0	2	0	1	0	0	1	0	0	0
1	1	0	0	1	1	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0
1	1	0	-1	-1	1	-1	-1	-1	0	-1	-1	-2	-1	-1	0	-1	-1	0	-1	0
1	0	1	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	1	-1	-1	0	-1	0	0	-1	0	0	0	-1	0	-1	0	0	-1	0	-1
1	-1	-1	1	0	-1	1	-1	-1	-1	-1	0	-2	-1	-1	-1	0	-1	-1	0	0
1	-1	-1	0	1	-1	0	0	0	0	0	-1	0	-1	0	0	-1	0	0	-1	-1
1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	2	2	1	1	1	1	1	1	1

- T: design vector for the constant term.
- A: effect coding for row variable.
- B: effect coding for column variable.
- C: linear contrasts for row variable— $(R + 1)/2 - i$ ; any consecutive numbering will do; for variables with three categories, this is the same as the first code for the row variable.
- D: linear contrasts for column variable— $(C + 1)/2 - j$ .
- E: design for the linear-by-linear interaction or uniform association, obtained by multiplying the linear contrast vector for the row and for the column.
- F: design vectors for the row effects model, obtained by multiplying the design vectors for the row categories and the linear contrast vector for the column.
- G: design vectors for column effects model, obtained by multiplying the design vectors for the column variable and the linear contrast for the row variable.
- H: homogeneous row-column effects model, obtained by adding each vector in the matrix for the row and the corresponding vector in the matrix for the column.
- I: row and column effects model—concatenation of F and G minus the redundant linear-by-linear interaction vector.
- J: interaction vectors for saturated model, obtained by multiplying each vector in A with each vector in B.

Design matrix for each type of model is obtained by concatenating relevant vectors from above, and the degrees of freedom by number of cells in the table minus the number of columns in the design matrix.

	<i>Vectors</i>	<i>df</i>
Independence model	T~A~B	4
Uniform association model	T~A~B~E	3
Row-effects model	T~A~B~F	2
Column-effects model	T~A~B~G	2
Homogeneous row-column effects model	T~A~B~H	2
Row and column effects model	T~A~B~I	1
Saturated model	T~A~B~J	0

Note: ~ (Horizontal concatenation); \* (Multiplication); \*~ (Horizontal direct product); † (Excluding redundant vector).

Table 6

obvious. Those parameters pertaining to each cell, denoted by  $t_{ij}$ , are equivalent to the product of the relevant part of the design matrix and the corresponding coefficients.

**SOME EXTENSIONS**

There are several ways in which one can extend the basic features of the log-linear models examined so far. Among these, the following three seem important: (1) utilizing the ANOVA-like structure

**Hypothetical Tables Illustrating Some Association Models**

a) Row-Effects Association Model

		Frequency		Odds Ratio			
		X					
Y	400	400	50	4	4		
	200	800	400	2	2		
	100	800	800				
		Log Parameters			Multiplicative Parameter		
	1.155	0	-1.555	3.175	1	.315	
	-.231	0	.231	.794	1	1.260	
	-.924	0	.924	.397	1	2.520	

b) Homogeneous Row-Column Effects Model

		Frequency		Odds Ratio			
		X					
Y	400	100	100	4	2		
	100	100	200	2	1		
	100	200	400				
		Log Parameters			Multiplicative Parameters		
	.924	-.231	-.693	2.520	.794	.500	
	-.231	0	.231	.794	1	1.260	
	-.693	.231	.462	.500	1.260	1.587	

Table 7

of the log-linear model and the well-developed sampling theory to explore interaction patterns of multivariate categorical data, (2) manipulating the design matrices to examine more specific hypotheses and models, and (3) combining the strategic features of subgroup analysis and the flexibility and power of the log-linear models to produce more readily accessible analysis results. These three extensions are discussed below.

**General Extension of Log-Linear Models.** The most straightforward and widely used application of the log-linear model is to explore the interaction pattern of multivariate data by exploiting the ANOVA-like structure of the model. Given several variables to examine, especially when each variable contains more than two categories, it is almost impossible to examine the data structure in detail.

The ANOVA-like structure allows one to develop a convenient strategy to explore the existence of multiway relationships among the variables.

This strategy requires that one start with a design matrix for each variable (containing  $k-1$  vectors, where  $k$  is the number of categories in the variable). It does not matter whether one uses dummy coding or effect coding. To examine all the possible interrelationships in the data, one needs design matrices corresponding to each two-way interaction to  $m$ -way interaction where  $m$  is the number of variables. To construct a design matrix for a two-way interaction between variable  $A$  and variable  $B$ , simply cross-multiply the design vectors for  $A$  with those for  $B$ . (This method is illustrated in Table 6.) This general approach to design matrices is extended to  $m$ -way. For example, a three-way

interaction is handled by cross-multiplying each two-way vector with the basic design vectors for a third variable, and so on.

If one includes in the model all the vectors covering up to  $m$ -way interactions, the resulting model is saturated, and each frequency in the multiway table is completely described. In general, one wants to explore and, if possible, find a parsimonious way to describe the data structure. One general strategy, perhaps overused, is to examine systematically the hierarchical pattern inherent in the design constraints and serially examine a nested set of models. To illustrate, consider that there are three variables and that the basic design vectors for each variable are represented by  $A$ ,  $B$ , and  $C$ , respectively. Let  $T$  stand for the constant vector. Then an example of a nested set of models is illustrated below. The commas indicate concatenation, and two or more letters together indicate cross-multiplication of the basic design vectors for each variable.

- $H_1$ :  $T$
- $H_2$ :  $T, A, B, C$
- $H_{3a}$ :  $T, A, B, C, AB$
- $H_{3b}$ :  $T, A, B, C, AB, AC$
- $H_{3c}$ :  $T, A, B, C, AB, AC, BC$
- $H_4$ :  $T, A, B, C, AB, AC, BC, ABC$
- $(H_1)$  Equiprobability
- $(H_2)$  Total independence
- $(H_{3a})$  One two-way interaction
- $(H_{3b})$  Two two-way interactions
- $(H_{3c})$  No three-way interaction
- $(H_4)$  Saturated model

Each hypothesis is tested, using the appropriate degrees of freedom, which is given by the number of the cells in the frequency table minus the number of vectors contained in the design matrix and the  $\chi^2$  or  $L^2$  statistics associated with each model. The sequence from the hypotheses in set (3) is arbitrary; one may choose any nested set or directly examine 3c. One usually would accept the simplest hypothesis that is compatible with the data.

If variables contain many categories, even the simplest two-way interactions will use up many

degrees of freedom. This type of generic testing does not incorporate into the design matrix any special relationships that may exist between variables. Models of this type are routinely available in standard computer packages and therefore are quite accessible. For that reason, they are overused. Moreover, the sequential nature of the testing violates some of the assumptions of classical hypothesis testing. Nevertheless, in the hands of an experienced researcher, they become a flexible tool for exploring the multivariate data structure.

**The Uses of Constrained Models.** The flexibility and power of log-linear models are fully realized only when one incorporates a specific hypothesis about the data into the design matrices. There are virtually endless varieties one can consider. Some of the simple but strategic models of association were introduced in the preceding section.

Incorporating such models into a multivariate analysis is not difficult if one views the task in the context of design matrices. For instance, suppose one suspects that a certain pattern of relationship exists between  $X$  and  $Y$  (for instance, the social class of origin and destination in intergenerational mobility). Furthermore, one may have an additional hypothesis that these relationships vary systematically across different political systems (or across societies with different levels of economic development). If one can translate these ideas into appropriate design matrices, using such a model will provide a much more powerful test than the generic statistical models described in the previous section can provide. Many social mobility studies incorporate such design matrices as a way of incorporating a special pattern of social mobility in the overall design (for some examples, see Duncan 1979; Hout 1984; Yamaguchi 1987 and for new developments, see DiPrete 1990; Stier and Grusky 1990; Wong 1990, 1992, 1995; Xie 1992).

In general, there are two problems in using such design matrices. The first, which depends in part on the researcher's creative ability, is the problem of translating theoretically relevant models into appropriate design matrices. The second is finding a way to obtain a good statistical solution for the model, but this is no longer much of a problem because of the wide availability of computer programs that allow the incorporation of design matrices (see Breen 1984 for a discussion of

preparing design matrices for a computer program that handles generalized linear systems).

One of the general problems has been that researchers often do not make the underlying design matrices explicit and as a result sometimes misinterpret the results. A solution for this problem is to think explicitly in terms of the design matrices, not in analogy to a generic (presumed) ANOVA model.

**Use of Percentage Tables in Log-Linear Modeling.** Multivariate analysis is in general complex. Categorical analysis is especially so, because one conceptual variable has to be treated as if it were ( $k - 1$ ) variables, with  $k$  being the number of categories in the variable. Therefore, even with a limited number of variables, if each variable contains more than two categories, examining the multivariate pattern becomes extremely difficult. Therefore, the tendency is to rely on the general hypothesis testing discussed earlier.

It is useful to borrow two of the strategies of subgroup analysis: focusing on a bivariate relationship and using percentage distributions. After an acceptable log-linear model is identified, one therefore may display the relationship between two key variables while the effects of other variables are controlled or purged (Clogg 1978; Clogg and Eliason 1988a; Clogg et al. 1990; Kaufman and Schervish 1986). Furthermore, a percentage distribution for the bivariate distribution may be compared with the corresponding percentage distributions when different sets of variables are controlled in this manner. Fortunately, the log-linear modeling can provide a very attractive way in which the confounding effects of many variables can be purged from the relationship that is under special scrutiny (Clogg 1978; Kaufman and Schervish 1986; Clogg and Eliason 1988a; Clogg et al. 1990). Clogg et al. (1990) show a general framework under which almost all the known variations in adjustments can be considered a special case. Furthermore, they also describe statistical testing procedures for variety of statistics associated with such adjustments.

Tables 8 and 9 contain examples of traditional subgroup analysis, log-linear analysis, and the uses of standardization or purging methods. The upper panel of Table 8 contains a bivariate table showing (1) the race of the defendant ( $X$ ) and (2) the

**Death Penalty Verdict ( $Y$ ) by Defendant's Race ( $X$ ) and Victim's Race ( $Z$ )**

		Death Penalty		
Victim's Race	Defendant's Race	Yes	No	Percentage Yes
Total	White	53	430	11.0
	Black	15	176	7.9
White	White	53	414	11.3
	Black	11	37	22.9
Black	White	0†	16	0.0
	Black	4	139	2.8

Table 8

SOURCE: Radelet and Pierce (1991), p. 25, and Agresti (1996), p. 54.

NOTE: (1) For log-linear analysis, 0.5 is added to zero cell. (2) In the original data, there are two cases that involve both white and black victims. (3) The data do not consistently identify Spanish ancestry. Most defendants and victims with Spanish ancestry are coded as white. For detailed information, see Radelet and Pierce (1991).

verdict—death penalty versus other penalties ( $Y$ ), while the lower panel contains the result of traditional three-variables subgroup analysis, in which the original relationship between  $X$  and  $Y$  is reanalyzed within the categories of the third variable, the race of the victims ( $Z$ ). (These data are based on individuals who were convicted of multiple homicides in Florida. See Radelet and Peierce 1991; Agresti 1996.)

The original bivariate relationship seems to indicate that whites are more likely to receive the death penalty than are blacks. However, when the race of the victims is controlled, the partial relationship between  $X$  and  $Y$  is reversed: Within each category of victim, blacks are more likely to receive the death penalty than are whites. The underlying reasons for this reversal are two related facts: (1) There is a strong association between the race of the defendant ( $X$ ) and the race of the victims ( $Z$ ): white defendants are more likely to kill whites than blacks, while black defendants are more likely to kill blacks than whites, and (2) there is a strong relationship between the race of the victims and the death penalty: those who killed white victims are more likely to receive the death penalty than are those who killed blacks. Once these relation-

TABULAR ANALYSIS

**Design Matrix, Expected Frequencies, and Standardized Percentages under the Model without Three-Way Interaction**

a) Design Matrix and Coefficients for the Model without Three-Way Interaction

<i>I</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>XZ</i>	<i>YZ</i>	<i>XY</i>	Parameter	Coefficient	z-value
1	1	1	1	1	1	1	Constant (T)	2.959	21.0
1	1	1	-1	1	-1	-1	Z	1.039	6.9
1	1	-1	1	-1	1	-1	X	-0.135	-1.3
1	1	-1	-1	-1	-1	1	Y	-1.382	-11.1
1	-1	1	1	-1	-1	1	XZ	1.137	14.7
1	-1	1	-1	-1	1	-1	YZ	0.558	3.9
1	-1	-1	1	1	-1	-1	XY	-0.201	-2.2
1	-1	-1	-1	1	1	1	$L^2 = 0.284$	df = 1	Prob. = 0.594

b) Expected Frequencies under the Model without Three-Way Interaction

Victim's Race	Defendant's Race	Death Penalty		Percentage Yes
		Yes	No	
Total	White	53.5	430.0	11.1
	Black	15.0	176.0	7.9
White	White	53.3	413.7	11.4
	Black	10.7	37.3	22.3
Black	White	0.2	16.3	1.2
	Black	4.3	138.7	3.0

c) Direct Standardization

Total	White	35.87	447.62	7.4
	Black	28.16	162.84	14.7
White	White	33.58	260.67	11.4
	Black	25.91	90.33	22.3
Black	White	2.29	186.95	1.2
	Black	2.25	72.51	3.0

d) Purging XZ (= purging XZ and XYZ)

Total	White	17.7	183.2	8.8
	Black	34.9	160.9	17.8
White	White	17.0	132.3	11.4
	Black	33.5	116.6	22.3
Black	White	0.7	50.9	1.4
	Black	1.4	44.4	3.1

e) Purging XZ and YZ (= purging XZ, YZ, and XYZ)

Total	White	11.0	260.9	4.0
	Black	21.5	228.7	8.6
White	White	9.8	231.9	4.0
	Black	19.1	203.2	8.6
Black	White	1.2	29.0	4.0
	Black	2.4	25.4	8.6

Table 9



ships are taken into consideration, blacks receive a higher rate of the death penalty than do whites.

The log-linear analysis can supplement such a traditional subgroup analysis in several convenient ways. Panel (a) of Table 9 shows design matrix, coefficients, and standardized values under the model without three-way interaction. These statistics show several things that are not obvious in the conventional subgroup analysis: (1) The three-way interaction is not statistically significant, (2) all three bivariate relationships are statistically significant, and (3) in some sense, the association between  $X$  and  $Z$  is the strongest and that between  $X$  and  $Y$  is the weakest among the three bivariate relationships.

Panel (b) shows expected frequencies and relevant percentages under the model (where the three-way interaction is assumed to be zero). The pattern revealed in each subtable is very similar to that under the traditional subgroup analysis shown in Table 8. (This is as it should be, given no three-way interaction effect.) Within each category of victim's race, black defendants are more likely to receive the death penalty than are white defendants. The standardization or purging, then, allows one to summarize this underlying relationship between  $X$  and  $Y$  under the hypothetical condition that the effect of the third variable is controlled or purged. There are many different ways of controlling the effects of the third variable: (1) direct standardization, panel (c), (2) when the effects of  $XZ$  relationship are purged (in addition to the purging of the three-way interaction) panel (d), (3) when, in addition to the previous purging, the effects of the  $YZ$  relationship are also purged, panel (e). Although the percentage differences seem to vary, the underlying log-linear effect remains constant: Blacks are twice more likely to receive the death penalty than are whites when both kill a victim of the same race (see Clogg 1978; Clogg and Eliason 1988b; Clogg et al. 1990).

(SEE ALSO: *Analysis of Variance and Covariance; Causal Inference Models; Measures of Association; Nonparametric Statistics; Statistical Methods*)

REFERENCES

Agresti, Alan 1983 "A Survey of Strategies for Modeling Cross-Classifications Having Ordinal Variables." *Journal of the American Statistical Association* 78:184-198.

— 1984 *Analysis of Ordinal Categorical Data*. New York: Wiley.

— 1990 *Categorical Data Analysis*. New York: Wiley.

— 1996 *An Introduction to Categorical Data Analysis*. New York: Wiley.

Alba, Richard D. 1988 "Interpreting the Parameters of Log-Linear Models." In J. Scott Long, ed., *Common Problems/Proper Solutions*. Beverly Hills, Calif.: Sage.

Alwin, Dwane F., and Robert M. Hauser 1975 "The Decomposition of Effects in Path Analysis." *American Sociological Review* 40:37-47.

Anderson, J. A. 1984 "Regression and Ordered Categorical Variables." *Journal of the Royal Statistical Society* B46:1-30.

Arminger, G., and G. W. Bohrnstedt 1987 "Making It Count Even More: A Review and Critique of Stanley Lieberson's *Making It Count: The Improvement of Social Theory and Research*." *Sociological Methodology* 17:347-362.

Berk, R. A. 1983 "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48:386-398.

— 1986 "Review of *Making It Count: The Improvement of Social Research and Theory*." *American Journal of Sociology* 92:462-465.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland 1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT Press.

Blalock, Hubert M., Jr. 1964 *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

—, ed. 1985a *Causal Models in the Social Sciences*, 2nd ed. New York: Aldine.

—, ed. 1985b *Causal Models in Panel and Experimental Designs*. New York: Aldine.

Bollen, K. A. 1989 *Structural Equations with Latent Variables*. New York: Wiley.

Breen, Richard 1984 "Fitting Non-Hierarchical and Association Models Using GLIM." *Sociological Methods and Research* 13:77-107.

Bunge, Mario 1979 *Causality and Modern Science*, 3rd rev. ed. New York: Dover.

Campbell, D. T., and J. C. Stanley 1966 *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.

Clogg, Clifford C. 1978 "Adjustment of Rates Using Multiplicative Models." *Demography* 15:523-539.

— 1982a "Using Association Models in Sociological Research: Some Examples." *American Journal of Sociology* 88:114-134.

- 1982b “Some Models for the Analysis of Association in Multiway Cross-Classifications Having Ordered Categories.” *Journal of the American Statistical Association* 77:803–815.
- , and Scott R. Eliason 1988a “A Flexible Procedure for Adjusting Rates and Proportions, Including Statistical Methods for Group Comparisons.” *American Sociological Review* 53:267–283.
- 1988b “Some Common Problems in Log-Linear Analysis.” In J. Scott Long, ed., *Common Problems/Proper Solutions*. Beverly Hills, Calif.: Sage.
- , and Edward S. Shihadeh 1994 *Statistical Models for Ordinal Variables*. Thousand Oaks, Calif.: Sage.
- , James W. Shockey, and Scott R. Eliason 1990 “A General Statistical Framework for Adjustment of Rates.” *Sociological Methods and Research* 19:156–195.
- Cook, Thomas D., and Donald T. Campbell 1979 *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Davis, James A. 1984 “Extending Rosenberg’s Technique for Standardizing Percentage Tables.” *Social Forces* 62:679–708.
- , and Ann M. Jacobs 1968 “Tabular Presentations.” In David L. Sills, ed., *The International Encyclopedia of the Social Sciences*, vol.15. New York: Macmillan and Free Press.
- Duncan, Otis Dudley 1966 “Path Analysis: Sociological Examples.” *American Journal of Sociology* 72:1–16.
- 1975 *Introduction to Structural Equation Models*. New York: Academic Press.
- 1979 “How Destination Depends on Origin in the Occupational Mobility Table.” *American Journal of Sociology* 84:793–803.
- Fienberg, Stephen E. 1980 *The Analysis of Cross-Classified Data*, 2nd ed. Cambridge, Mass.: MIT Press.
- Fisher, F. M. 1966 *The Identification Problem in Econometrics*. New York: McGraw-Hill.
- Fleiss, J. L. 1981 *Statistical Methods for Rates and Proportions*, 2nd ed. New York: Wiley Interscience.
- Goldberger, Arthur S., and Otis Dudley Duncan (eds.) 1973 *Structural Equation Models in the Social Sciences*. New York and London: Seminar Press.
- Goodman, Leo A. 1978 *Analyzing Qualitative/Categorical Data: Log-Linear Analysis and Latent Structure Analysis*. Cambridge, Mass.: Abt.
- 1981 “Three Elementary Views of Loglinear Models for the Analysis of Cross-Classifications Having Ordered Categories.” In Karl F. Schuessler, ed., *Sociological Methodology*. San Francisco: Jossey-Bass.
- 1984 *The Analysis of Cross-Classified Categorical Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.
- 1985 “The Analysis of Cross-Classified Data Having Ordered and/or Unordered Categories: Association Models, Correlation Models, and Asymmetry Models for Contingency Tables with or without Missing Entries.” *Annals of Statistics* 13:10–69.
- 1987 “The Analysis of a Set of Multidimensional Contingency Tables Using Log-Linear Models, Latent Class Models, and Correlation Models: The Solomon Data Revisited.” In A. E. Gelfand, ed., *Contributions to the Theory and Applications of Statistics: A Volume in Honor of Herbert Solomon*. New York: Academic Press.
- 1990 “Total-Score Models and Rasch-Type Models for the Analysis of a Multidimensional Contingency Table, or a Set of Multidimensional Contingency Tables, with Specified and/or Unspecified Order for Response Categories.” In Karl F. Schuessler, ed., *Sociological Methodology*. San Francisco: Jossey-Bass.
- Haberman, Shelby J. 1978 *Analysis of Qualitative Data*, vol. 1: *Introductory Topics*. New York: Academic Press.
- 1979 *Analysis of Qualitative Data*, vol. 2: *New Developments*. New York: Academic Press.
- Hausman, J. A. 1978 “Specification Tests in Econometrics.” *Econometrica* 46:1251–1272.
- Heckman, J. J. 1979 “Sample Selection Bias as a Specification Error.” *Econometrica* 47 153–161.
- , and R. Robb 1986 “Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes.” In H. Wainer, ed., *Drawing Inferences from Self-Selected Samples*. New York: Springer-Verlag.
- Heise, David R. 1975 *Causal Analysis*. New York: Wiley.
- Hout, Michael 1984 “Status, Autonomy, Training in Occupational Mobility.” *American Journal of Sociology* 89:1379–1409.
- Hyman, Herbert 1955 *Survey Design and Analysis: Principles, Cases and Procedures*. Glencoe, Ill.: Free Press.
- Kaufman, Robert L., and Paul G. Schervish 1986 “Using Adjusted Crosstabulations to Interpret Log-Linear Relationships.” *American Sociological Review* 51:717–733.
- Kendall, Patricia L., and Paul Lazarsfeld 1950 “Problems of Survey Analysis.” In Robert K. Merton and Paul F. Lazarsfeld, eds., *Continuities in Social Research: Studies in the Scope and Method of the American Soldier*. Glencoe, Ill.: Free Press.
- Kish, Leslie 1959 “Some Statistical Problems in Research Design.” *American Sociological Review* 24:328–338.

- Lazarsfeld, Paul F. 1955 "Interpretation of Statistical Relations as a Research Operation." In Paul F. Lazarsfeld and Morris Rosenberg, eds., *The Language of Social Research*. Glencoe, Ill.: Free Press.
- , Ann K. Pasanella, and Morris Rosenberg, eds. 1972 *Continuities in the Language of Social Research*. New York: Free Press.
- Leamer, E. E. 1978 *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: Wiley Interscience.
- Lieberson, Stanley 1985 *Making It Count: The Improvement of Social Research and Theory*. Berkeley and Los Angeles: University of California Press.
- Long, J. Scott 1984 "Estimable Functions in Loglinear Models." *Sociological Methods and Research* 12:399–432.
- , ed. 1988 *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research*. Beverly Hills, Calif.: Sage.
- Mare, Robert D., and Christopher Winship 1988 "Endogenous Switching Regression Models for the Causes and Effects of Discrete Variables." In J. Scott Long, ed., *Common Problems/Proper Solutions*. Beverly Hills, Calif.: Sage.
- McCullagh, P. 1978 "A Class of Parametric Models for the Analysis of Square Contingency Tables with Ordered Categories." *Biometrika* 65:413–418.
- , and J. Nelder 1983 *Generalized Linear Models*. London: Chapman and Hall.
- Mosteller, F. 1968 "Association and Estimation in Contingency Tables." *Journal of the American Statistical Association* 63:1–28.
- , and John W. Tukey 1977 *Data Analysis and Regression*. Reading, Mass.: Addison-Wesley.
- Nelder, J. A., and R. W. M. Wedderburn 1972 "Generalized Linear Models." *Journal of the Royal Statistical Society* A135:370–384.
- Plackett, R. L. 1974 *The Analysis of Categorical Data*. London: Griffin.
- Press, S. L., and S. Wilson 1978 "Choosing between Logistic Regression and Discriminant Analysis." *Journal of the American Statistical Association* 73:699–705.
- Radelet, Michael I., and Glenn L. Pierce 1991 "Choosing Those Who Will Die: Race and the Death Penalty in Florida." *Florida Law Review* 43:1–34.
- Rosenberg, Morris 1968 *The Logic of Survey Analysis*. New York: Basic Books.
- Rubin, D. B. 1977 "Assignment to Treatment Group on the Basis of a Covariance." *Journal of Educational Statistics* 2:1–26.
- Simon, Herbert A. 1954 "Spurious Correlation: A Causal Interpretation." *Journal of the American Statistical Association* 49:467–479.
- 1979 "The Meaning of Causal Ordering." In Robert K. Merton, James S. Coleman, and Peter H. Rossi, eds., *Qualitative and Quantitative Social Research: Papers in Honor of Paul F. Lazarsfeld*. New York: Free Press.
- Singer, Burton, and Margaret Mooney Marini 1987 "Advancing Social Research: An Essay Based on Stanley Lieberson's *Making It Count*." In Clifford C. Clogg, ed., *Sociological Methodology*. Washington, D.C.: American Sociological Association.
- Thiel, Henri 1971 *Principles of Econometrics*. New York: Wiley.
- Williams, D. A. 1976 "Improved Likelihood Ratio Tests for Complete Contingency Tables." *Biometrika* 63:33–37.
- Xie, Yu 1989 "An Alternative Purging Method: Controlling the Composition-Dependent Interaction in an Analysis of Rates." *Demography* 26:711–716.
- Yamaguchi, Kazuo 1987 "Models for Comparing Mobility Tables: Toward Parsimony and Substance." *American Sociological Review* 52:482–494.
- Zeisel, Hans 1985 *Say It with Figures*, 6th ed. New York: Harper & Row.

JAE-ON KIM  
MYOUNG-JIN LEE

## TECHNOLOGICAL RISKS AND SOCIETY

See Society and Technological Risks.

## TERRITORIAL BELONGING

### DEFINITION

Belonging is defined as the state of being part of something. Territorial belonging implies being part of a territory. The definition of a territory, although it is conditioned by the morphology of space, is essentially a social operation that is connected with the factors that induce the perception of boundaries. These are complex factors that researchers in the "psychology of form" (*Gestaltpsychologie*) have attempted to specify (Reusch 1956, pp. 340–361). Campbell has identified seven

of these factors. Those analytically most relevant to social systems are similarity and shared destiny or “common fate” (Campbell 1958), to which the ecological, economic, and sociological traditions (Hawley 1950, p. 258) add interdependence, which is related to Campbell’s (1958) notion of internal diffusion.

Territorial belonging is therefore a form of social belonging (for a detailed treatment, see Pollini 1987) that is displayed by a spatially defined collectivity. Spatial definition more or less precisely and more or less sharply delimits (where the concept of a boundary refers to a line or zone) a territory to which a name is given. Belonging to a spatially defined collectivity thus may be related to the name given to a territory, so that it becomes simply territorial without ceasing to be social as well. To emphasize its twofold nature, it also may be called “socioterritorial belonging” (Pollini 1992, pp. 55–58).

#### THE MULTIPLICITY OF TERRITORIAL BELONGINGS

Like any form of social belonging, territorial belonging may relate to objective or subjective elements and may be defined by the self or by others (Merton 1963). Like social belonging, it may be largely exclusive or admit to multiplicity and may be ascribed or acquired (Simmel 1908).

Workers in human ecology, human geography, sociology, and land economics have long attempted to provide a definition of the most suitable territorial units for social purposes. They have oscillated between emphasizing the principle of similarity (the morphology of the territory, the physical and cultural features of the individuals who inhabit it, the predominant type of economic activity, etc.) and emphasizing the principle of interdependence (on the basis of gravitational flows for work and services, areas of relatively intense exchange, etc.) (see Galtung 1968). The implementation of the political function of governing human communities, moreover, has led to the fixing of territorial boundaries that express (and produce) a common fate (Hawley 1950, p. 258).

Reciprocal relations are among the main criteria used to define territorial units (similarity, interdependence, and common fate, taking the proximity of elements for granted). As studies of “nation building” have shown (see Deutsch [1953] 1966), similarity tends to create relations of interdependence and interdependence generates perceptions of similarity (Simmel 1890, p. 40; Shils 1975, p. 17); in the same manner, a common fate induces perceptions of similarity and interdependence and similarity and interdependence heighten the perception of a common fate.

The most enduring and significant spatial units are those with multi-confirmed boundaries (Campbell 1958), that is, those for which the criteria of similarity, functional interdependence, and common fate are congruent. The European nation-states of the nineteenth and twentieth centuries exemplify the successful achievement of this congruence (Eisenstadt 1973, pp. 231–235).

In traditional nomadic and agricultural societies, similarity, interdependence, and common fate are properties that relate substantially to a single socioterritorial unit that includes everyday life almost in its entirety, except when extraordinary occasions (great feasts and celebrations, great markets, wars, etc.) demonstrate the importance of a broader socioterritorial unit that is ethnic and/or tribal in nature or sometimes is of state dimension (kingdoms and empires, churches and great religious organizations). This is the social order that is called the “segmentary society of mechanical solidarity” by Durkheim (1893), *Gemeinschaft* by Toennies (1887), and the “independent community” by Hawley (1950, pp. 223 ff) and is exemplified by numerous contemporary societies (Dyson Hudson 1966).

As a significant division of territorial labor develops—induced by the reduced spatial friction brought about by advances in transport and communications that allow more frequent exchanges over longer distances (the “mobiletic revolution” of Russett 1967) and by technical progress, which requires the greater accumulation of capital and is not uniformly distributed across the territory (industrial revolution)—the areas of interdependence

expand and are structured into several levels (Hawley 1950, pp. 236–257). This has evident effects on areas of common fate. Increased interdependence facilitates temporary or permanent movements across the territory and thus alters similarities and differences as well as the criteria for their definition and perception (Sola Pool 1965).

In short, mainly as a result of these phenomena, a socioterritorial structure grows more complex. Important socioterritorial units proliferate, intersect, and are organized into larger (Parsons 1961, pp. 123 ff) and relatively fluid systems, while the congruence among the three main principals of sociospatial structuring diminishes. These changes have been interpreted as the decay of *Gemeinschaft* (Toennies 1887) and of the territorial state (Herz 1957), as stages in an ongoing evolution into cosmopolitanism, and as the onset of a single overarching socioterritorial unit: the world in its entirety. Parsons (1951) introduced in his theoretical pattern of variables for the definition of roles the dichotomy between particularism and universalism. Although of general analytic significance, this dichotomy also has been used to characterize the process of modernization (Parsons 1971), which, with reference to territorial belonging, includes the localism–cosmopolitanism dichotomy. The term “globalization,” which now has general currency although it dates from the early 1970s (Kaufman 1974), has been employed more recently to define this trend.

The increased complexity of socioterritorial belonging is obviously correlated to the complexity of the social structuring of the territory. All individuals are involved in relational networks that may be micro-local (habitation), local, regional, national, continental, and global, and they shift easily and rapidly from one level to another by virtue of the ease of communications and transportation or simply pass subjectively from one role to another (Webber 1963, 1964).

Owing to the ease of communications and transport, every individual may encounter and assimilate elements of other cultures. Cultural diversity has dwindled before the advance of modern culture as it is interpreted by Western society.

Intellectual elements (“scientific” criteria for the reliability of knowledge) and most emotional (aspirations and values) and evaluative ones (ethics and hierarchies of values) are widely shared by humankind (Inglehart 1997). Consequently, similarities and differences are difficult to define in territorial terms, and when they are thus definable, they are increasingly so only as symbols (languages, flags, cultural artifacts, physical resemblances, etc.), since knowledge-evoking and evaluative elements have been reduced to being options that are private, individual, and socially irrelevant (ethical and gnoseological relativism, individualism) (Thomas and Znaniecki 1918–1920; Halman, et al. 1987).

The state itself, which is called on to define the conditions of the collective control of the collective destiny, although it is still characterized by distinct territorial boundaries, is undergoing profound change as a result of the erosion of its sovereignty and power by the rise of supranational political organizations (e.g., the United Nations) and, in centralized states, infrastate political organizations (Galtung 1967), as well as by the growth of multinational economic enterprises and other noneconomic associations over which it can exert little or no control.

Consequently, if the continuing determination of the territorial boundaries of political and/or administrative units (Herz 1968) allows belongings to be related to them, those belongings grow increasingly less socially significant in regards to not only interdependencies and similarities but also common destiny. This occurs because social relevance is divided among several units organized into a system of relationships that need not be hierarchical and may indeed compete with each other.

#### **TERRITORIAL BELONGING SUBJECTIVELY DEFINED BY REFERENCE TO THE SELF (THE SENTIMENT OF TERRITORIAL BELONGING)**

**The localism–cosmopolitanism of territorial belonging: A single-or multidimensional concept?**  
The growing complexity of territorial belonging,

along with the hypothesis that it is the manifestation of an ongoing process whose final outcome is cosmopolitanism (or the erasure of any nonglobal, noncumenical sense of belonging), has prompted sociologists to study the phenomenon empirically by focusing on the subjective definition of belonging provided by individuals with reference to themselves. Taking the process of growing systemization at the “energetic” level for granted, attention has been focused on how the phenomenon is subjectively reflected in the subjective definition individuals give to their territorial belongings. The aim of these studies has been to single out the factors that induce a person to feel that she or he belongs primarily to one unit rather than to another and, more generally, why she or he expresses a primarily cosmopolitan sense of belonging rather than one anchored in a particular geographic unit. Other studies have explored attachment to units of a particular size (home and neighborhood, local community, region, nation, continent, etc.).

The data gathered in both types of studies have included highly modernized contexts. On the home and neighborhood, see Fried and Gleicher (1961), Galster and Hesser (1981), and Fried (1982). On the local community, see Kasarda and Janowitz (1974), Rojek et al. (1975), Taylor and Townsend (1976), Wasserman (1982), Fried (1982), Goudy (1990), and Beggs et al. (1996). On regional units, see Piveteau (1969) and Gubert (1997). On belongings and national pride, apart from studies of nationalism and ethnicity, see the European Values Study in Ashford and Timms (1992, pp. 89–91) and the World Values Study in Ingleheart (1997, pp. 303–305). On territorial belongings on the localism–cosmopolitanism continuum, see Treinen (1965), Gubert and Struffi (1987), Gubert (1992a), and Strassoldo and Tessarin (1992). These data confirm the complexity of the phenomenon of territorial belonging when it is defined subjectively (a sentiment of belonging). Subjectively felt belongings are multiple, and each has its own role to play; that is, they become socially important in accordance with the particular context, which may change rapidly. For example, the sense of national belonging is exalted during international sports events, but local and regional senses of belonging

emerge during sports events within a country. This multiplicity of territorial belongings therefore rules out their mutual exclusiveness; this emerges clearly when subjects are asked to declare the absolute level of attachment they feel to different territorial units (Gubert 1998).

It is probable that partly diversified belongings underlie this multiplicity. This diversity was not grasped in early empirical studies of the sentiment of belonging, which relied largely on relative measures of the strength of attachment to socioterritorial units arranged along a continuum, with the neighborhood and community at one extreme and cosmopolitanism at the other (Terhune 1965; Gubert 1972, p. 181). Relative measures also were used in later large-scale surveys such as the European Values Study and the World Values Study. It is difficult to imagine that a Pole’s identification with his or her nation is the same as his or her attachment to his or her place of residence. One may feel strongly Polish while also having close bonds with one’s local community, and this cannot be explained by a single-dimensional conception of localism–cosmopolitanism, which instead suggests a social experience in which the nation as a sociospatial unit has weaker emotional connotations.

Further evidence that feelings of belonging to diverse socioterritorial units differ is provided by analyses of the attitudes of subjects who declare that they do not feel attached to any particular territorial unit. These individuals may be called cosmopolitans, but their cosmopolitanism is not only the extreme position on the localism–cosmopolitanism continuum; it is also symptomatic of difficulties of social integration, or anomie (Bertelli 1992). The distribution of the strength of socioterritorial belonging therefore measures not only the territorial size of the main social collectivities of reference but also the intensity of social integration, which in the case of declared nonbelonging to any territorial unit is markedly diminished, perhaps more in some cases than in others; therefore, cosmopolitanism is internally differentiated or heterogeneous. Parsons’s assumption that there is social belonging (and therefore loyalty and attachment) if there is social conformity—if, that is,

the subject conforms with the institutional obligations of solidarity (Parsons 1951)—receives empirical support in that a lack of social integration is connected with a lack of belonging to any socioterritorial unit.

**Localism–cosmopolitanism as a single-dimensional continuum.** Taken for granted (or given) the multiplicity of socioterritorial belongings and their partly diverse nature a further finding of empirical surveys concerns the relative importance of each of these belongings with respect to the others. It is assumed here that, to some extent, they express the localism–cosmopolitanism dimension and therefore can be plotted along a continuum.

When asked about the matter, even interviewees in highly modernized contexts tend to assign more importance to local belongings than to national and supranational ones. Overall, the two units that predominate are the commune and the nation or state. The importance of the other subnational and supracomunal units (province, region) depends on the structure of the public powers (the federal or nonfederal structure of the state) (Gubert 1995) or on whether an ethnic and/or national minority constitutes the majority in a subnational unit (Gubert 1975, p. 305). By contrast, the emergence of subcommunal units (neighborhoods, districts) depends on the settlement pattern of a commune: When it is articulated into several settlements (districts), importance is more frequently ascribed to subcommunal units, while the neighborhood acquires more importance when it constitutes a “natural area” within the communal settlement and has weak links with the rest of the urban territory.

If, rather than proposing a predetermined set of spatial units largely defined in political-administrative terms, individuals who declare their attachment to a particular territory are asked to freely describe it, to give it a name or define its boundaries, the area of belonging is generally more circumscribed. Indeed, it is sometimes restricted to the domestic ambit and its immediate surroundings, with little consideration of supralocal spatial units (Gubert 1992b, pp. 266–277). One therefore may conclude that the immediate bond with a

territory is mainly local or microlocal and that only the mention of larger socioterritorial units prompts individuals to consider supralocal territorial attachments.

**Between negation of the hypothesis of an evolution toward cosmopolitanism and the reasons for the persistence of localism in a modern society.** Multiple regression analysis shows that the social conditions that orient people to cosmopolitan or supralocal belongings (e.g., higher educational level, greater geographic mobility, residence in a metropolis) attenuate local attachments, although they do not entirely eliminate them (Gubert 1992d, pp. 506–523). These conditions seeming to affect the intensity of territorial attachment, attenuating it all levels rather than eliminating its primacy at the local level. Residential mobility seems to multiply local attachments rather than creating a single cosmopolitan attachment (Rubinstein and Parmalee 1992, 1996; Gubert 1992b, pp. 326–330; Feldman 1996).

This phenomenon can be explained to some extent by the reasons adduced to account for the most important territorial attachments. These reasons mainly concern day-to-day living (Kasarda and Janowitz’s [1974] duration of residence) and the places of infancy and the family (Taylor and Townsend 1976); much less important are the physical characteristics of places (except for places which are morphologically very distinct, such as mountains and coastlines, partially confirming Fried’s [1982] argument) or utility and opportunity.

Therefore, the strongest territorial attachment is connected with strongly affective social relations or with affectively important individual experiences, such as those of a child who progressively establishes a relationship with his or her immediate surroundings.

It is evident that the geographic mobility for mainly utilitarian reasons (work, access to services, use of free time) typically induced by life in modern society does not affect this type of attachment greatly. For the same reason, little influence is exerted by educational level: Although it may extend relational ambits, it does so mainly for utilitarian reasons or ones tied to a person’s profes-

sional role (Webber 1964). This has evident consequences for the territorial area of matrimonial choice and the areas which contain a person's best friends or relatives; these areas have increased in size, but they are nevertheless of modest proportions. Also, as was found for residential mobility, the consequence of this extension is more to increase the number of places of particular attachment than to induce an attachment to broader spatial units. In other words, the elasticity of relational ambits to the reduced costs of overcoming distance is much greater for secondary instrumental relationships than it is for primary relationships, especially family ones, which are firmly anchored in residence (Parsons 1951, p. 180 ff; 1960, p. 250 ff).

Regression analysis not of the size of the area of main attachment but of the (absolute) intensity of the feeling of belonging to it reveals, as was already mentioned, that the social conditions most typical of modernity (residence in large cities, geographic mobility, residential mobility, higher levels of education, secularization, ethical relativism, individualism, etc.) tend to attenuate the intensity of local attachment but do not erase it. Another explanation for the continuing primacy of local attachments over cosmopolitan ones is therefore that modernity tends to reduce the intensity of territorial belongings, but not to such a varied extent across territorial units that it alters the hierarchy of subjective importance assigned to them by most individuals. This sheds light on why the indicators of localism and cosmopolitanism with regard to the territorial extension of main attachments are determined by a factor independent of the one that groups together the indicators of the intensity of attachment (in any case, the correlation between the indicators is weak).

In addition to intensity, the features that connote greater "modernity" of ecological social and cultural positions in individuals exert a certain amount of influence on the social importance of the sentiment of territorial belonging (Gubert 1992d). Perception of the boundaries that mark the zone of principal attachment are less sharp: The inside-outside boundary, unless it is marked by obvious physical barriers against communica-

tion and vision, is not connected with the perception of distinct sociocultural differences and tends to assume the features of a zone rather than those of a line. The identifying features of the collectivity on the other side of the boundary, except in the case of physical barriers against communication or vision, display not abrupt discontinuities but gradual variations. Stereotypes and the ingroup-outgroup opposition are not as marked as those which arise when racial or ethnic differences are involved.

Moreover, the multiplicity of territorial belongings and their dependence on the particular and contingent nature of the context prevent the onset of radical ingroup-outgroup conflicts. Factor analysis has shown that the intensity of territorial attachment and ingroup-outgroup opposition (which is measurable, for example, by acceptance or rejection of immigrants in the area of principal belonging) are independent factors. Feeling stronger attachment to a particular territory therefore is weakly correlated with greater hostility toward or less acceptance of outsiders. The explanation of variance in these attitudes has more to do with the sphere of interests than with that of territorial belongings.

Equally independent (or weakly correlated) are factors relative to the social significance of the sentiment of territorial belonging and indicators of the territorial extension of the most important units.

Expectations of a positive relation between modernity and cosmopolitanism are not borne out by the data, although cross-sectional analysis is not conclusive on this matter. Certainly contradicted are claims that cosmopolitan attachments predominate in contemporary societies. What the advance of modernity seems to have done is breed a plurality of territorial belongings and reduce their intensity and social significance.

#### **Distribution models of territorial attachments.**

Having ascertained that territorial attachment is felt mainly at the local level (the primary community of everyday life), one may inquire about the distribution pattern of the intensity of territorial attachment, extending the analysis to ambits that do not occupy the highest position in the hierarchy of spatial units.



Once again, the most detailed surveys have been carried out in Italy, a country where the dialectic among localism, nationalism, and cosmopolitanism is particularly lively. If one considers the first three positions in a decreasing scale of attachment from subcommunal units to the whole world, it is possible to identify the most common and significant rank orderings (Gubert 1992b, pp. 279–281). By far the most prevalent is a model one may call “lococentric:” Territorial units diminish in importance as they grow larger. In three-dimensional space, where a two-dimensional geographic plane constitutes the horizontal axes and the relative intensity of territorial attachment is the vertical axis, they assume a cone shape whose apex is the smallest spatial unit.

Two other patterns emerge: the opposite model (upside-down cone), where the importance of a territorial unit decreases as one moves from broader to narrower spatial units, and, more commonly, a volcano-shaped model, a variant of the cone where the greatest importance is attributed to intermediate units, the next greatest importance to smaller units, and the least importance to larger ones.

Despite the prevalence of the lococentric model, a model that might develop instead of the “cosmopolitan” one (upside-down cone) as a consequence of the continuing advance of modernity is the volcano model, which places greater emphasis on supralocal sociospatial units compared with smaller communal or subcommunal ones, but without the units assuming greater importance as they become more inclusive and extensive. This is a type of territoriality that reflects adaptation to the extension of relational spaces beyond the settlement of residence that assumes a form other than cosmopolitanism.

### CONCLUSIONS

Sociological research into territorial belonging as it is subjectively defined by self-reference contradicts the claim that the extension of relational spaces and the increased frequency of relations across larger distances (continental and global) have led to the superseding of local or at least

noncosmopolitan attachments. Attachment to local ambits still largely predominates. If anything, one discerns adaptations of a different kind, such as the increased complexity of the subjectively felt territorial bond and its closer dependence on changing contexts, the diminished intensity and social significance of various kinds of attachment, and a shift of primary attachment from the local level to one midway between localness and ecumene (the inhabited world). Rejection of any particular territorial attachment that might represent the outcome of cosmopolitan development is at least partly due to a lack of social integration.

Predictions that the demise of Toennies’s *Gemeinschaft* and Durkheim’s segmentary society would indicate the end of the overriding importance of attachment to particular places is not supported by empirical inquiry. Toennies’s *Gemeinschaft* has disappeared from modern society, but territorially restricted areas still have social relations of a communitarian nature, that is, relation in which community action in Weber’s sense (Weber [1922] 1972, p. 21) predominates.

However, this may not explain the facts unless one refers to Pareto’s assertion (1916, vol. II, sections 112–120, 1023–1041; Treinen 1965) that the sentiment of territorial belonging is a “residue” that can be included among those of the “persistence of aggregates.” A sentiment springs from the psychological association between emotionally significant experiences and the context in which they happen. For example, emotionally positive experiences tied to childhood and the family and relationships involving sexuality are positively associated with the places in which they have occurred. This gives rise to deep emotional bonds with those places, bonds that lie beyond reasoning or rationalization. In Pareto’s account, the attachment to places thus depends not on the communitarian or societal nature of the collectivity settled in the place of residence but instead on the fact that the most emotionally significant experiences in a person’s life necessarily occur in a territorially limited context, if only because of the limits imposed by the perceptive horizon. This also explains why the dispersion of these emotionally significant experiences tends more to multiply

the places of attachment than to create broad and inclusive attachments.

The technical progress of the means of communications and the means of transportation that facilitate the expansion of relational ambits, the individualization and secularization of culture, and the proliferation of the utilitarian businesslike relationships of *Gesellschaft* have had little or no effect on the processes that, according to Pareto's theory, cause the birth and persistence of attachment to place. As the negation of a particular bond with a particular place, cosmopolitanism unless it is not an ideological position that is deliberately assumed and declared to the interviewer, therefore may be a symptom of social marginalization caused by a lack of emotionally significant positive experiences. It is thus only rarely caused by the extreme territorial dispersion of those experiences and much more frequently caused by other personal and family events.

There are a number of reservations about the premises of the phenomenon known as globalization. Although relational ambits certainly extend over broader areas or, more precisely, relations have increased in intensity in even the largest of those areas, there is little evidence that the relative intensity of relations has increased. One cannot rule out the possibility that a diminution in the "friction of space" has intensified relations with all the territorial levels into which social life is structured, and it is likely that the intensification at medium and low levels is even greater for certain types of relation than it is at higher and broader ones. The increase in international trade, for example, does not mean that the system is becoming increasingly globalized; the increase may be greater at the national and regional levels, and so in relative terms, regional and national systems have become more self-contained (Deutsch 1960; Deutsch and Eckstein 1960–1961). The unproven assumption that the increase in the absolute density of relations has been accompanied by an increase in their relative density is credited to the existence of globalization processes that may not exist or may exist only for particular types of relations. Moreover, it is evident that in subjective perception (and not only in this area), the importance of the various levels

depends much more on their relative density than on their absolute density.

If more detailed empirical research confirms the validity of these remarks, the persisting primacy of local, or at least noncosmopolitan, belongings may be explained not only by Paretian hypotheses but also by the predominance at the relational level of local, regional, and national systems compared with continental and global ones.

Belonging as being part of is a phenomenon of central importance in sociological analysis. Territorial belonging is only one way to manifest social belonging. It may be associated with (ethnic and national belonging) or in competition with (membership in universal religions or international interest groups) other forms of social belonging. It also is a complex phenomenon, but it nevertheless seems to be characterized by features more durable than those of other social or group belongings. It resembles ethnic and national belongings but may be less exposed to change if the Paretian hypothesis is correct. It is a phenomenon that probably is grounded in enduring features of human experience, in what once might have been called human nature. Even the dissolution of Toennies's *Gemeinschaft* into a nonorganic assembly of communitarian relations or the reduction of these relations to simple human relations confined to interindividual space has not severed the bond felt by individuals with the places where they had their most emotionally significant and gratifying personal experiences. Theories of modernity should take account of this fact, and the historicist paradigms that hypothesize a progressive evolution toward cosmopolitanism should be revised. Empirical research must continue, enrich itself with longitudinal surveys, and increase the number of cases observed. Feeling part of a territory and feeling tied to places are still important phenomena in numerous areas of social life.

#### REFERENCES

- Ashford S., and N. Timms 1992 *What Europe Thinks: A Study of Western European Values*. Brookfield, Mass.: Dartmouth.
- Beggs, J. J., J. S. Hurlbert, and V. Haines 1996 "Community Attachment in a Rural Setting: A refinement and

- Empirical Test of the Systemic Model." *Rural Sociology* 61(3):407-426
- Bertelli, B. 1992 "Assenza e Debolezza del Legame con il Territorio: Anomia, Devianza Sociale o Forme Diverse di Appartenenza?" In R. Gubert, ed., *L'Appartenenza Territoriale tra Ecologia e Cultura*. Trent, Italy: Reverdito.
- Campbell, D. T. 1958 "Common Fate, Similarity and Other Indices of the Status of Aggregates of Persons as Social Entities." *Behavioral Science* 3:14-25
- Deutsch, K. W. 1960 "The Propensity to International Transactions." *Political Studies* 8(8):147-155.
- and A. Eckstein 1960-1961 "National Industrialization and the Declining Share of the International Economic Sector 1890-1959." *World Politics* 13.
- 1966 *Nationalism and Social Communication*. Cambridge, Mass.: MIT Press.
- Durkheim, E. 1893 *La Division du Travail Social*. Paris: F. Alcan.
- Dyson Hudson, N. 1966 *Karimjong Politics*. Oxford, UK: Clarendon.
- Eisenstadt S. N. 1973 *Tradition, Change and Modernity*. New York: Wiley.
- Feldman, R. M. 1996 "Constancy and Change in Attachments to Types of Settlements." *Environment and Behavior* 28(4):419-445.
- Fried, M. 1982 "Residential Attachment: Sources of Residential and Community Satisfaction." *Journal of Social Issues* 38(3):107-119.
- , and P. Gleicher 1961 "Some Sources of Residential Satisfaction in an Urban Slum." *Journal of the American Institute of Planners* 27:305-315.
- Galster G. C., and G. W. Hesser 1981 "Residential Satisfaction: Compositional and Contextual Correlates." *Environment and Behavior* 13(6):735-758.
- Galtung, J. 1967 "On the Future of the International System." *Journal of Peace Research* 4.
- 1968 "A Structural Theory of Integration." *Journal of Peace Research* 5(4).
- Goudy W. J. 1990 "Community Attachment in a Rural Region." *Rural Sociology* 55(2):178-198
- Gubert, R. 1972 *La Situazione Confinaria*. Treiste, Italy: Lint.
- 1975 "Entitività delle Etnie e Conflittualità Interetnica." *Studi di Sociologia*. 13n.(3-4):302-322.
- , and L. Struffi, (eds.) 1987 *Strutture Sociali del Territorio Montano*. Milan, Italy: Angeli.
- 1992a *L'appartenenza Territoriale tra Ecologia e Cultura*. Trent, Italy: Reverdito. English translation 1999: *The Territorial Belonging between Ecology and Culture*. Trent: University of Trent Press.
- 1992b "I Caratteri Generali degli Intervistati e del Loro Legame Socio-Territoriale. In R. Gubert, ed., *L'Appartenenza Territoriale tra Ecologia e Cultura*. Trent, Italy: Reverdito.
- 1992c *Gli Orientamenti Socio-Politici degli Italiani*. In R. Gubert ed., *Persistenze e Mutamenti dei Valori degli Italiani nel Contesto Europeo*. Trent, Italy: Reverdito.
- 1992d "Le Dimensioni dell'Appartenza Territoriale: Verso un Modello Causale." In R. Gubert, ed., *L'Appartenenza Territoriale tra Ecologia e Cultura*. Trent, Italy: Reverdito.
- 1995 "Analysis of Regional Differences in the Values of Europeans." In R. de Moor, ed., *Values in Western Societies*. Tilburg, The Netherlands: Tilburg University Press.
- (ed.) 1997 *Specificità Culturale di Una Regione Alpina nel Contesto Europeo*. Milan, Italy: Angeli.
- 1998 "Appartenenza e Mobilità Socio-Territoriali. Primi Risultati di un'Indagine su Operatori Turistici nel Nord.Est Italiano." In A. Gasparini, ed., *Nation, Ethnicity, Minority and Border: Contribution to an International Sociology*. Gorizia, Italy: ISIG.
- Halman, L., F. Heunk, R. de Moor and H. Zaunders, (eds.) 1987 *Traditie, Secularisatie en Individualisering*. Tilburg: The Netherlands: University of Tilburg Press.
- Hawley, A. 1950 *Human Ecology*. New York: Ronald Press.
- Herz, J. H. 1957 "Rise and Demise of the Territorial State." *World Politics* 9(4).
- 1968 "The Territorial State Revisited: Reflexions on the Future of the National State." *Polity, The Journal of Northeastern Political Science Association* 1(1):12-34.
- Inglehart, R. 1997 *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton, N.J.: Princeton University Press.
- Kasarda, J., and M. Janowitz 1974 "Community Attachment in Mass Society." *American Sociological Review* 39:328-339.
- Kaufman, G. 1974 *Il Sistema Globale*. Udine, Italy: Del Bianco.

- Merton R. K. 1963 *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Parsons, T. 1951 *The Social System*. Glencoe, Ill.: Free Press.
- 1960 *Structure and Process in Modern Societies*. Glencoe, Ill.: Free Press.
- 1961 “Order and Community in the International System.” In J. N. Rosenau, ed., *International Politics and Foreign Policy*. New York and London: Free Press and Collier Macmillan.
- 1971 *The System of Modern Societies*. Englewood Cliffs, N.J.: Prentice-Hall.
- Piveteau, J. L. 1969 “Le Sentiment d’Appartenance Régionale en Suisse.” *Revue de Géographie Alpine*, 57(3):361–386.
- Pollini, G. 1987 *Appartenenza e Identità*. Milan, Italy: Angeli.
- 1992 “L’Appartenenza Socio-Territoriale.” In R. Gubert, ed., *L’Appartenenza Territoriale tra Ecologia e Cultura*. Trent, Italy: Reverdito.
- Reusch, J. 1956 “Analysis of Various Types of Boundaries.” In R. R. Grinker, ed., *Toward a Unified Theory of Human Behavior*. New York: Basic Books.
- Rojeck D. G., F. Clemente and G. F. Summers 1975 “Community Satisfaction: A Study of Contentment with Local Services” *Rural Sociology* 40:177–192.
- Rubinstein, R., P. A. Parmelee 1992 “Attachment to Place and the Representation of the Life Course by the Elderly.” In I. Altman and S. Low, eds., *Place Attachment* New York: Plenum, pp. 139–164.
- Russet, B. M. 1967 “The Ecology of Future International Politics.” *International Studies Quarterly* 11.
- Shils, E. 1975 *Center and Periphery: Essays in Macrosociology*. Chicago: University of Chicago Press.
- Simmel, G. 1890 *Ueber Soziale Differenzierung, Soziologische und Psychologische Untersuchungen*. Leipzig, Germany: Dunker & Humblot.
- 1908 *Soziologie. Untersuchungen ueber die Formen der Vergesellschaftung*. Leipzig, Germany: Dunker & Humblot.
- Sola Pool I. 1965 “Effects of Cross-National Contact on National and International Images.” In H. C. Kelman, ed., *International Behavior: A Social-Psychological Analysis*. New York: Holt, Rinehart and Winston.
- Strassoldo, R., and N. Tessarin 1992 *Le Radici del Localismo: Indagine Sociologica sull’Appartenenza Territoriale in Friuli*. Trent, Italy: Reverdito.
- Taylor, C. C. and A. R. Townsend 1976 “The Social ‘Sense of Place’ as Evidenced in North-East England.” *Urban Studies* 13(2):133–146.
- Terhune, K. W. 1965 “Nationalistic Aspiration, Loyalty and Internationalism.” *Journal of Peace Research* 3.
- Thomas, W. I. and F. Znaniecki 1918–1920 *The Polish Peasant in Europe and America*. Boston: Badger.
- Toennies, F. 1887 *Gemeinschaft und Gesellschaft*. Leipzig, Germany: O. R. Reisland.
- Treinen, H. 1965 “Symbolische Ortsbezogenheit: Eine Soziologische Untersuchung zum Heimatproblem.” *Koelner Zeitschrift fuer Soziologie und Sozialpsychologie*, 17(1–2):73–97, 254–297.
- Wasserman, I. M. 1982 “Size of Place in Relation to Community Attachment and Satisfaction with Community Services.” *Social Indicators Research* 11:421–436.
- Webber, M. M. 1963 “Order in Diversity: Community without Propinquity.” In L. Wingo, ed., *Cities and Space*. Baltimore: Johns Hopkins University Press.
- 1964 *Culture, Territoriality, and the Elastic Mile*. *Regional Science Association Papers* 13:59–69
- Weber, M. (1922) 1972 *Wirtschaft und Gesellschaft*. Tuebingen, Germany: J.C.B. Mohr (Paul Siebeck).

RENZO GUBERT

## TERRORISM

Terrorism became an issue of worldwide concern in the last third of the twentieth century. Terrorist tactics were not new; they had been used for centuries before being defined as terrorism. The word “terror” entered the political lexicon during the French Revolution’s “reign of terror.” In the late nineteenth century, at the beginning of the twentieth, and again in the 1920s and 1950s—all periods between major wars on the European continents—terrorism became a technique of revolutionary struggle. Stalin’s regime in the 1930s and 1940s was called a reign of terror, but from the late 1940s to the 1960s the word was associated primarily with the armed struggles for independence waged in Palestine and Algeria, from which later generations of terrorists took their inspiration and instruction. After World War II, “terror” emerged

as a component of nuclear strategy; the fear of mutual destruction that would deter nuclear war between the United States and the Soviet Union was referred to as a balance of terror.

In the 1970s, "terrorism" became a fad word, promiscuously applied to a wide spectrum of conditions and actions. Bombs in public places were one form of terrorism, but some people asserted that oppression, poverty, hunger, racism, gang violence, spousal or child abuse, environmental destruction, and even medical malpractice were also forms of terrorism. Some governments labeled as terrorism all violent acts committed by their opponents, while antigovernment extremities claimed to be, and often were, the victims of government terror.

In an effort to get a firm hold on a slippery subject, those studying the phenomenon of terrorism were obliged to define it more precisely. Terrorism could be described simply as the use or threat of violence to create an atmosphere of fear and alarm and thus bring about a political result. But making this definition operative in political debate, rules of war, or criminal codes was anything but easy. Is all politically motivated violence terrorism? How does terrorism differ from ordinary crime? Should terrorism be considered a crime at all, or should it be seen as simply another form of armed conflict that is no less legitimate than any other form of war? Is the term properly reserved for those trying to overthrow governments, or can governments also be terrorists?

Definition was crucial because it ultimately determined the way in which terrorism has been studied. A major problem was that terrorism almost always has a pejorative connotation and thus falls in the same category of words as "tyranny" and "genocide," unlike such relatively neutral terms such as "war" and "revolution." One can aspire to objective and dispassionate research, but one cannot be neutral about terrorism any more than one can be neutral about torture. Thus, defining terrorism became an effort not only to delineate a subject area but also to maintain its illegitimacy. Even the most clinical inquiry was laden with values and therefore political issues. The very

study of terrorism implied to some a political decision.

Terrorism can be defined objectively by the quality of the act, not by the identity of the perpetrators or the nature of their cause. All terrorist acts are crimes, and many also would be war crimes or "grave breaches" of the rules of war if one accepted the terrorists' assertion that they wage war. All terrorist acts involve violence or the threat of violence, sometimes coupled with explicit demands. The violence is directed against noncombatants. The purposes are political. The actions often are carried out in a way that will achieve maximum publicity, and the perpetrators are usually members of an organized group.

Terrorist organizations are by necessity clandestine, but unlike other criminals, terrorists often but not always claim credit for their acts. Finally—the hallmark of terrorism—the acts are intended to produce psychological effects. This introduces a distinction between the actual victims of terrorist violence and the target audience. The connection between the victim and the target of terrorism can be remote. The identity of the victims may be secondary or even irrelevant to the terrorist cause. "Pure terrorism" is entirely indiscriminate violence.

Terrorism differs from ordinary crime in its political purpose and its primary objective. However, not all politically motivated violence is terrorism, nor is terrorism synonymous with guerilla war or any other kind of war.

Terrorist techniques can be used by governments or those fighting against governments: however, scholars generally use the term "terror" when discussing fear-producing tactics employed by governments and "terrorism" when referring to tactics used by those fighting against governments. The distinction is primarily semantic. Both groups may use threats, assassinations, or abductions, but government terror also may include arbitrary imprisonment, concentration camps, torture, mind-affecting techniques, and the use of drugs for political purposes. Antigovernment terrorists generally lack the infrastructure for such tactics. Government terror produces more victims than terrorism does. Terrorists tend to seek more publicity than do governments.

Although a prerequisite to empirical research, the attempt to define terrorism inevitably lent greater coherence to disparate acts of violence than did any analysis offered by the terrorists themselves, few of whom thought of assassinations, bombings, kidnappings, and airline hijackings as elements of a unified tactical repertoire, let alone the basis of a strategy. Ironically, in an effort to understand a phenomenon, researchers ran the risk of attributing to terrorists a level of strategic thinking they may not have possessed.

The term “international terrorism” refers to terrorist attacks on foreign targets or the crossing of national frontiers to carry out terrorist attacks. It was the dramatic rise in international terrorism—especially in the form of attacks on diplomats and commercial aviation in the late 1960s—that caused mounting alarm on the part of governments not directly involved in those local conflicts.

Although terrorist tactics were centuries old, contemporary terrorism, especially in its international form, emerged in the late 1960s from a unique confluence of political circumstances and technological developments. The political circumstances included the failure of the rural guerrilla movements in Latin America, which persuaded the guerrilla movements to take their armed struggles into the cities, where, through the use of dramatic actions such as kidnappings, they could be assured of attracting national and international attention. Their actions also provoked terrorist responses by governments that resorted to using “disappearances” of suspected guerrillas and their supporters, torture, and other tactics of terror.

In the Middle East, the failure of the Arab armies in the Six-Day War in 1967 caused the Palestinians to abandon dependence on Arab military power to achieve their aims and rely more heavily on the tactics of terrorism with the approval of some Arab governments. Israel retaliated with both military attacks and assassinations of suspected terrorist leaders.

The third political root of contemporary terrorism grew from widespread antigovernment demonstrations in universities in western Europe, Japan, and the United States that were provoked in

large measure but not exclusively by the war in Vietnam. By no stretch of the imagination could these antigovernment protests be called acts of terrorism, but the mass marches spawned extremist fringes that were inspired by third world guerrilla movements to carry on an armed struggle even after the student movements subsided.

Technological advances were equally important. Developments in communications—radio, television, communication satellites—made possible almost instantaneous access to global audiences, which was critical for a mode of violence aimed at publicity. Modern air travel provided terrorists with worldwide mobility and a choice of targets. Modern society’s dependence on technology created new vulnerabilities. Global weapons production guaranteed a supply of guns and explosives.

Once the tactics of terrorism were displayed worldwide, they provided inspiration and instruction for other groups, and terrorism became a self-perpetuating phenomenon.

The 1980s saw a new form of international terrorism: state-sponsored terrorism. Some governments began to use terrorist tactics or employ those tactics as a mode of surrogate warfare. Unlike government-directed terror, which is primarily domestic, state-sponsored terrorism is directed against foreign governments or domestic foes abroad. International diplomacy, economic sanctions, and in some cases, military actions brought about a reduction in this type of terrorism in the 1990s.

Despite great differences in political perspectives and outlook toward armed conflict, the international community gradually came to accept at least a partial definition of terrorism and prohibited certain tactics and attacks on certain targets. This approach reflected that of the academic community, focusing on the terrorist act and rejecting judgment based on the political objective or cause behind an act. Thus, by 1985, the United Nations General Assembly unanimously condemned international terrorism, including but not limited to acts covered by previous treaties against airline hijacking, sabotage of aircraft, attacks at civil air-

ports, attacks against maritime navigation and offshore platforms, attacks in any form against internationally protected persons (i.e., diplomats), and the taking of hostages.

By 1998, there were ten multilateral counterterrorism agreements that covered roughly half of all incidents of international terrorism but omitted primarily bombings of targets other than airlines or diplomatic facilities. One difficulty in delineating this type of terrorist act was distinguishing between terrorist bombings and aerial bombardment, which is considered a legitimate form of war. The rules of war prohibit indiscriminate bombing, thus providing at least a theoretical distinction between war and terrorism, although even with modern precision-guided munitions, collateral civilian casualties from aerial bombing in populated areas may vastly exceed casualties caused by the deliberate, indiscriminate bombs of terrorists.

In 1998, an International Convention for the Suppression of Terrorist Bombings attempted to address this lacuna. Unable to draw a clear distinction between terrorist bombings and other types of bombings, the treaty stated that the “activities of armed forces during an armed conflict . . . and the activities undertaken by military forces by a state in the exercise of their official duties, inasmuch as they are governed by other rules of international law, are not governed by this Convention.”

Terrorism is a subject matter, not a discipline. It has been approached by scholars from various academic perspectives with political scientists in the lead. Psychologists and psychiatrists have examined individual and group behavior, while jurists have formulated international legal approaches. Sociologists have not played a major role in research focusing specifically on terrorism but have addressed it in the broader context of deviance or social control.

Much research on terrorism has focused more narrowly on the topic. In part, this reflects the desire of researchers to avoid the murky, politically loaded area of underlying causes, where any discussion might be seen as condemnation or rationalization of terrorist violence. Nonetheless, there have been excellent case studies of individual groups and their tactics.

Defining terrorism in terms of the act has enabled researchers to maintain a theoretically objective approach and conduct at least some primitive quantitative analysis. Event-based analysis has enabled them to discern broad patterns and trends and chart the growth of terrorism and its diffusion around the globe. They have been able to demonstrate statistically that as terrorism has increased in volume, it has also become bloodier. Researchers were able to illustrate a clear trend toward incidents of large-scale indiscriminate violence in the 1980s and infer that terrorists tend to be more imitative than innovative in their tactics. Event-based analysis also has permitted researchers to distinguish the operational profiles of specific terrorist groups, and these profiles have been useful in identifying changes in a group’s *modus operandi*.

At the same time, event-based analysis has led the analysts into some methodological traps. An exclusive focus on terrorist actions, for example, resulted in terrorists being viewed first as if they were all part of a single entity and second as if they were almost extraterrestrial. While there are connections and alliances among some terrorist groups, the only thing the terrorists of the world have in common is a propensity for violence and certain tactics. Moreover, each group is rooted in its own social, political, and cultural soil, and cross-national comparisons are difficult. This has led to the question of whether there is such a thing as a terrorist-prone society.

It is, however, dangerous to attribute the actions of a few to perceived political defects or cultural flaws of a society as a whole, and researchers’ attempts to discern deeper causes or conditions that lead to high levels of terrorism in certain societies have produced meager results. Terrorism is not demonstrably a response to poverty or political oppression. The liberal democracies of western Europe have suffered high levels of terrorist violence, while totalitarian states are virtually free of terrorism. Overall, countries with perceived terrorist problems tend to be comparatively advanced politically and economically. They are more highly urbanized and have higher per capita incomes, larger middle classes, more university stu-

dents, and higher rates of literacy. One may ask whether political and economic advancement simply brings a more modern form of political violence.

One obstacle to linking high levels of terrorism with environmental factors is the problem of measuring terrorism. For the most part, this has been done by counting terrorist incidents, but international terrorism was narrowly and, more important, artificially defined to include only incidents that cause international concern, a distinction that has meant very little to the terrorists. Counting *all* terrorist incidents, both local and international, is better but still inadequate. Terrorist tactics, narrowly defined, represent most of what some groups, particularly those in western Europe, do but for other groups, terrorism represents only one facet of a broader armed conflict. In civil war situations, such as that in Lebanon in the 1970s, separating incidents of terrorism from the background of violence and bloodshed was futile and meaningless. And what about the extensive unquantified political and communal violence in the rural backlands of numerous third world countries? Broad statements about terrorist-prone or violence-prone societies simply cannot be made by measuring only a thin terrorist crust of that violence, if at all. The problem, however, is not merely one of counting. Although terrorists arise from the peculiarities of local situations, they may become isolated in a tiny universe of beliefs and discourse that is alien to the surrounding society. German terrorists were German, but were they Germany? In the final analysis, one is forced to dismiss the notion of a terrorist-prone society.

If terrorism cannot be explained by environmental factors, one must look into the mind of the individual terrorist for an explanation. Are there individuals who are prone to becoming terrorists—a preterrorist personality? Encouraged by superficial similarities in the demographic profiles of terrorists—many of them have been urban middle and upper class (not economically deprived) males in their early twenties with university or at least secondary school educations—researchers searched for common psychological features.

Behavioral analysts painted an unappealing portrait: The composite terrorist appeared to be a

person who was narcissistic, emotionally flat, easily disillusioned, incapable of enjoyment, rigid, and a true believer who was action-oriented and risk seeking. Psychiatrists could label terrorists as neurotic and possibly sociopathic, but they found that most of them were not clinically insane. Some behavioral analysts looked for deeper connections between terrorists' attitude toward parents and their attitudes toward authority. A few went further in claiming a physiological explanation for terrorism based on inner ear disorders, but these assertions were not given wide credence in the scientific community. The growing number of terrorists apprehended and imprisoned in the 1980s permitted more thorough studies, but while these studies occasionally unearthed tantalizing similarities, they also showed terrorists to be a diverse lot.

Much research on terrorism has been government-sponsored and therefore oriented toward the practical goal of understanding terrorism in order to defeat it. While social scientists looked for environmental or behavioral explanations for terrorism, other researchers attempted to identify terrorist vulnerabilities and successful countermeasures. They achieved a measure of success in several areas. Studies of the human dynamics of hostage situations led to the development of psychological tactics that increased the hostages' chances of survival and a better understanding (and therefore more effective treatment) of those who had been held hostage. In some cases, specific psychological vulnerabilities were identified and exploited. With somewhat less success, researchers also examined the effects of broader policies, such as not making concessions to terrorists holding hostages and using military retaliation. The conclusion in this area were less clear-cut.

Another area of research concerned the effects of terrorism on society. Here, researchers viewed terrorism as consisting of not only the sum of terrorist actions but also the fear and alarm produced by those actions. Public opinion polls, along with measurable decisions such as not flying and avoiding certain countries, provided the measure of effect.

Some critics who are skeptical of the entire field of terrorism analysis assert that the state and



its accomplice scholars have “invented” terrorism as a political issue to further state agendas through manipulation of fear, the setting of public discourse, preemptive constructions of “good” and “evil,” and the creation of deliberate distractions from more serious issues. “Terrorism,” a pejorative term that is useful in condemning foes, has generated a lot of fear mongering, and the issue of terrorism has been harnessed to serve other agendas, but one would have to set aside the reality of terrorist campaigns to see terrorism solely as an invention of the hegemonic state. While such deconstructions reveal the ideological prejudices of their authors, they nonetheless have value in reminding other analysts to be aware of the lenses through which they view terrorism.

Over the years, research on terrorism has become more sophisticated, but in the end, terrorism confronts people with fundamental philosophical questions: Do ends justify means? How far does one go on behalf of a cause? What is the value of an individual human life? What obligations do governments have toward their own citizens if, for example, they are held hostage? Should governments or corporations ever bargain for human life? What limits can be imposed on individual liberties to ensure public safety? Is the use of military force, as a matter of choice, ever appropriate? Can assassination ever be justified? These are not matters of research. They are issues that have been dictated through the ages.

(SEE ALSO: *International Law; Revolutions; Social Control; Violent Crime; War*)

## REFERENCES

- Barnaby, Frank 1996 *Instruments of Terror*. London: Satin.
- Hoffman, Bruce 1998 *Inside Terrorism*. London: Gollancz.
- Kushner, Harvey W. (ed.) 1998 *The Future of Terrorism: Violence in the New Millennium*. Thousand Oaks, Calif.: Sage.
- Laquer, Walter 1977 *Terrorism*. Boston: Little Brown.
- Lesser, Ian O. et al. 1999 *Countering the New Terrorism*. Santa Monica: Rand Corporation.
- Oliverio, Annamarie 1998 *The State of Terror*. Albany: State University of New York Press.
- O’Sullivan, Noel (ed.) 1986 *Terrorism, Ideology and Revolution*. Boulder, Colo: Westview.
- Schmid, Alex P., and Albert J. Jongman 1988 *Political Terrorism*. Amsterdam: North Holland.
- Simon, Jeffrey D. 1994 *The Terrorist Trap: America’s Experience with Terrorism*. Bloomington: Indiana University Press.
- Stern, Jessica 1999 *The Ultimate Terrorists*. Cambridge, Mass.: Harvard University Press.
- Thackrah, John Richard 1987 *Terrorism and Political Violence*. London: Routledge & Kegan Paul.
- Wilkinson, Paul, and Alasdair M. Stewart (eds.) 1987 *Contemporary Research on Terrorism*. Aberdeen, Scotland: Aberdeen University Press.

BRIAN MICHAEL JENKINS

## THEOCRACY

See Religion, Politics, and War; Religious Organizations.

## TIME SERIES ANALYSIS

Longitudinal data are used commonly in sociology, and over the years sociologists have imported a wide variety of statistical procedures from other disciplines to analyze such data. Examples include survival analysis (Cox and Oakes 1984), dynamic modeling (Harvey 1990), and techniques for pooled cross-sectional and time series data (Hsiao 1986). Typically, these procedures are used to represent the causal mechanisms by which one or more outcomes are produced; a stochastic model is provided that is presumed to extract the essential means by which changes in some variables bring about changes in others (Berk 1988).

The techniques called *time series analysis* have somewhat different intellectual roots. Rather than try to represent explicit causal mechanisms, the goal in classical time series analysis is “simply” to describe some longitudinal stochastic processes in summary form. That description may be used to inform existing theory or inductively extract new

theoretical notions, but classical time series analysis does not begin with a fully articulated causal model.

However, more recent developments in time series analysis and in the analysis of longitudinal data more generally have produced a growing convergence in which the descriptive power of time series analysis has been incorporated into causal modeling and the capacity to represent certain kinds of causal mechanisms has been introduced into time series analysis (see, for example, Harvey 1990). It may be fair to say that differences between time series analysis and the causal modeling of longitudinal data are now matters of degree.

### CLASSICAL TIME SERIES ANALYSIS

Classical time series analysis was developed to describe variability over time for a single unit of observation (Box and Jenkins 1976, chaps. 3 and 4). The single unit could be a person, a household, a city, a business, a market, or another entity. A popular example in sociology is the crime rate over time in a particular jurisdiction (e.g., Loftin and McDowall 1982; Chamlin 1988; Kessler and Duncan 1996). Other examples include longitudinal data on public opinion, unemployment rates, and infant mortality.

**Formal Foundations.** The mathematical foundations of classical time series analysis are found in difference equations. An equation “relating the values of a function  $y$  and one or more of its differences  $\Delta y, \Delta^2 y, \dots$  for each  $x$ -value of some set of numbers  $S$  (for which each of these functions is defined) is called a difference equation over the set  $S$ ” ( $\Delta y = y_t - y_{t-1}$ ,  $\Delta^2 = \Delta(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2}$ , and so on) (Goldberg 1958, p. 50). The  $x$ -values specify the numbers for which the relationship holds (i.e., the domain). That is, the relationships may be true for only some values of  $x$ . In practice, the  $x$ -values are taken to be a set of successive integers that in effect indicate *when* a measure is taken. Then, requiring that all difference operations  $\Delta$  be taken with an interval equal to 1 (Goldberg 1958, p. 52), one gets the following kinds of results (with  $t$  replacing  $x$ ):  $\Delta^2 y_t + k y_t = 2k + 7$ , which can be rewritten  $y_t - 2y_{t-1} + (1-k)y_{t-2} = 2k + 7$ .

Difference equations are deterministic. In practice, the social world is taken to be stochastic. Therefore, to use difference equations in time series analysis, a disturbance term is added, much as is done in conventional regression models.

**ARIMA Models.** Getting from stochastic difference equations to time series analysis requires that an observed time series be conceptualized as a product of an underlying substantive process. In particular, an observed time series is conceptualized as a “realization” of an underlying process that is assumed to be reasonably well described by an unknown stochastic difference equation (Chatfield 1996, pp. 27–28). In other words, the realization is treated as if it were a simple random sample from the distribution of all possible realizations the underlying process might produce. This is a weighty *substantive* assumption that cannot be made casually or as a matter of convenience. For example, if the time series is the number of lynchings by year in a southern state between 1880 and 1930, how much sense does it make to talk about observed data as a representative realization of an underlying historical process that could have produced a very large number of such realizations? Many time series are alternatively conceptualized as a population; what one sees is all there is (e.g., Freedman and Lane 1983). Then the relevance of time series analysis becomes unclear, although many of the descriptive tools can be salvaged.

If one can live with the underlying world assumed, the statistical tools time series analysis provides can be used to make inferences about which stochastic difference equation is most consistent with the data and what the values of the coefficients are likely to be. This is, of course, not much different from what is done in conventional regression analysis.

For the tools to work properly, however, one must at least assume “weak stationarity.” Drawing from Gottman’s didactic discussion (1981, pp. 60–66), imagine that a very large number of realizations were actually observed and then displayed in a large two-way table with one time period in each column and one realization in each row. Weak stationarity requires that if one computed the

mean for each time period (i.e., for each column), those means would be effectively the same (and identical asymptotically). Similarly, if one computed the variance for each time period (i.e., by column), those variances would be effectively the same (and identical asymptotically). That is, the process is characterized in part by a finite mean and variance that do not change over time.

Weak stationarity also requires that the covariance of the process between periods be independent of time as well. That is, for any given lag in time (e.g., one period, two periods, or three periods), if one computed all possible covariances between columns in the table, those covariances would be effectively the same (and identical asymptotically). For example, at a lag of 2, one would compute covariances between column 1 and column 3, column 2 and column 4, column 3 and column 5, and so on. Those covariances would all be effectively the same. In summary, weak stationarity requires that the variance-covariance matrix across realizations be invariant with respect to the displacement of time. Strong stationarity implies that the joint distribution (more generally) is invariant with respect to the displacement of time. When each time period's observations are normally distributed, weak and strong stationarity are the same. In either case, history is effectively assumed to repeat itself.

Many statistical models that are consistent with weak stationarity have been used to analyze time series data. Probably the most widely applied (and the model on which this article will focus) is associated with the work of Box and Jenkins (1976). Their most basic ARIMA (autoregressive-integrated moving-average) model has three parts: (1) an autoregressive component, (2) a moving average component, and (3) a differencing component.

Consider first the autoregressive component and  $y_t$  as the variable of interest. An autoregressive component of order  $p$  can be written as  $y_t - \Phi_1 y_{t-1} - \dots - \Phi_p y_{t-p}$ .

Alternatively, the autoregressive component of order  $p$  (AR[ $p$ ]) can be written in the form  $\Phi(B)y_t$ , where  $B$  is the backward shift operator—that is,  $(B)y_t = y_{t-1}$ ,  $(B^2)y_t = y_{t-2}$  and so on—and  $\phi(B) =$

$1 - \phi_1 B - \dots - \phi_p B^p$ . For example, an autoregressive model of order 2 is  $y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2}$ .

A moving-average component of order  $q$ , in contrast, can be written as  $\epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$ . The variable  $\epsilon_t$  is taken to be “white noise,” sometimes called the “innovations process,” which is much like the disturbance term in regression models. It is assumed that  $\epsilon_t$  is not correlated with itself and has a mean (expected value) of zero and a constant variance. It sometimes is assumed to be Gaussian as well.

The moving-average component of order  $q$  (MA[ $q$ ]) also can be written in the form  $\Theta(B)\epsilon_t$ , where  $B$  is a backward shift operator and  $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ . For example, a moving-average model of order 2 is  $\epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$ .

Finally, the differencing component can be written as  $\Delta^d y_t$  where the  $d$  is the number differences taken (or the degree of differencing). Differencing (see “Formal Foundations,” above) is a method to remove nonstationarity in a time series mean so that weak stationarity is achieved. It is common to see ARIMA models written in general form as  $\Theta(B)\Delta^d y_t = \Phi(B)\epsilon_t$ .

A seasonal set of components also can be included. The set is structured in exactly the same way but uses a seasonal time reference. That is, instead of time intervals of one time period, seasonal models use time intervals such as quarters. The seasonal component usually is included multiplicatively (Box and Jenkins 1976, chap. 9; Granger and Newbold 1986, pp. 101–114; Chatfield 1996 pp. 60–61), but a discussion here is precluded by space limitations.

For many sets of longitudinal data, nonstationarity is not merely a nuisance to be removed but a finding to be highlighted. The fact that time series analysis requires stationarity does not mean that nonstationary processes are sociologically uninteresting, and it will be shown shortly that time series procedures can be combined with techniques such multiple regression when nonstationarity is an important part of the story.

**ARIMA Models in Practice.** In practice, one rarely knows which ARIMA model is appropriate

for the data. That is, one does not know what orders the autoregressive and moving-average components should be or what degree of differencing is required to achieve stationarity. The values of the coefficients for these models typically are unknown as well. At least three diagnostic procedures are commonly used: time series plots, the autocorrelation function, and the partial autocorrelation function.

A time series plot is simply a graph of the variable to be analyzed arrayed over time. It is always important to study time series plots carefully to get an initial sense of the data: time trends, cyclical patterns, dramatic irregularities, and outliers.

The autocorrelation function and the partial autocorrelation function of the time series are used to help specify which ARIMA model should be applied to the data (Chatfield 1996, chap. 4). The rules of thumb typically employed will be summarized after a brief illustration.

Figure 1 shows a time series plot of the simulated unemployment rate for a small city. The vertical axis is the unemployment rate, and the horizontal axis is time in quarters. There appear to be rather dramatic cycles in the data, but on closer inspection, they do not fit neatly into any simple story. For example, the cycles are not two or four periods in length (which would correspond to six-month or twelve-month cycles).

Figure 2 shows a plot of the autocorrelation function (ACF) of the simulated data with horizontal bands for the 95 percent confidence interval. Basically, the autocorrelation function produces a series of serial Pearson correlations for the given time series at different lags: 0, 1, 2, 3, and so on (Box and Jenkins 1976, pp. 23–36). If the series is stationary with respect to the mean, the autocorrelations should decline rapidly. If they do not, one may difference the series one or more times until the autocorrelations do decline rapidly.

For some kinds of mean nonstationarity, differencing will not solve the problem (e.g., if the nonstationarity has an exponential form). It is also important to note that mean nonstationarity may be seen in the data as differences in level for

different parts of the time series, differences in slope for different parts of the data, or even some other pattern.

In Figure 2, the autocorrelation for lag 0 is 1.0, as it should be (correlating something with itself). Thus, there are three spikes outside of the 95 percent confidence interval at lags 1, 2, and 3. Clearly, the correlations decline gradually but rather rapidly so that one may reasonably conclude that the series is already mean stationary. The gradual decline also usually is taken as a sign autoregressive processes are operating, perhaps in combination with moving-average processes and perhaps not. There also seems to be a cyclical pattern, that is consistent with the patterns in Figure 1 and usually is taken as a sign that the autoregressive process has an order of more than 1.

Figure 3 shows the partial autocorrelation function. The partial autocorrelation is similar to the usual partial correlation, except that what is being held constant is values of the times series at lags shorter than the lag of interest. For example, the partial autocorrelation at a lag of 4 holds constant the time series values at lags of 1, 2, and 3.

From Figure 3, it is clear that there are large spikes at lags of 1 and 2. This usually is taken to mean that the  $p$  for the autoregressive component is equal to 2. That is, an AR[2] component is necessary. In addition, the abrupt decline (rather than a rapid but gradual decline) after a lag of 2 (in this case) usually is interpreted as a sign that there is no moving-average component.

The parameters for an AR[2] model were estimated using maximum likelihood procedures. The first AR parameter estimate was 0.33, and the second was estimate -0.35. Both had  $t$ -values well in excess of conventional levels. These results are consistent with the cyclical patterns seen in Figure 1; a positive value for the first AR parameter and a negative value for the second produced the apparent cyclical patterns.

How well does the model fit? Figures 4 and 5 show, respectively, the autocorrelation function and the partial autocorrelation function for the residuals of the original time series (much like

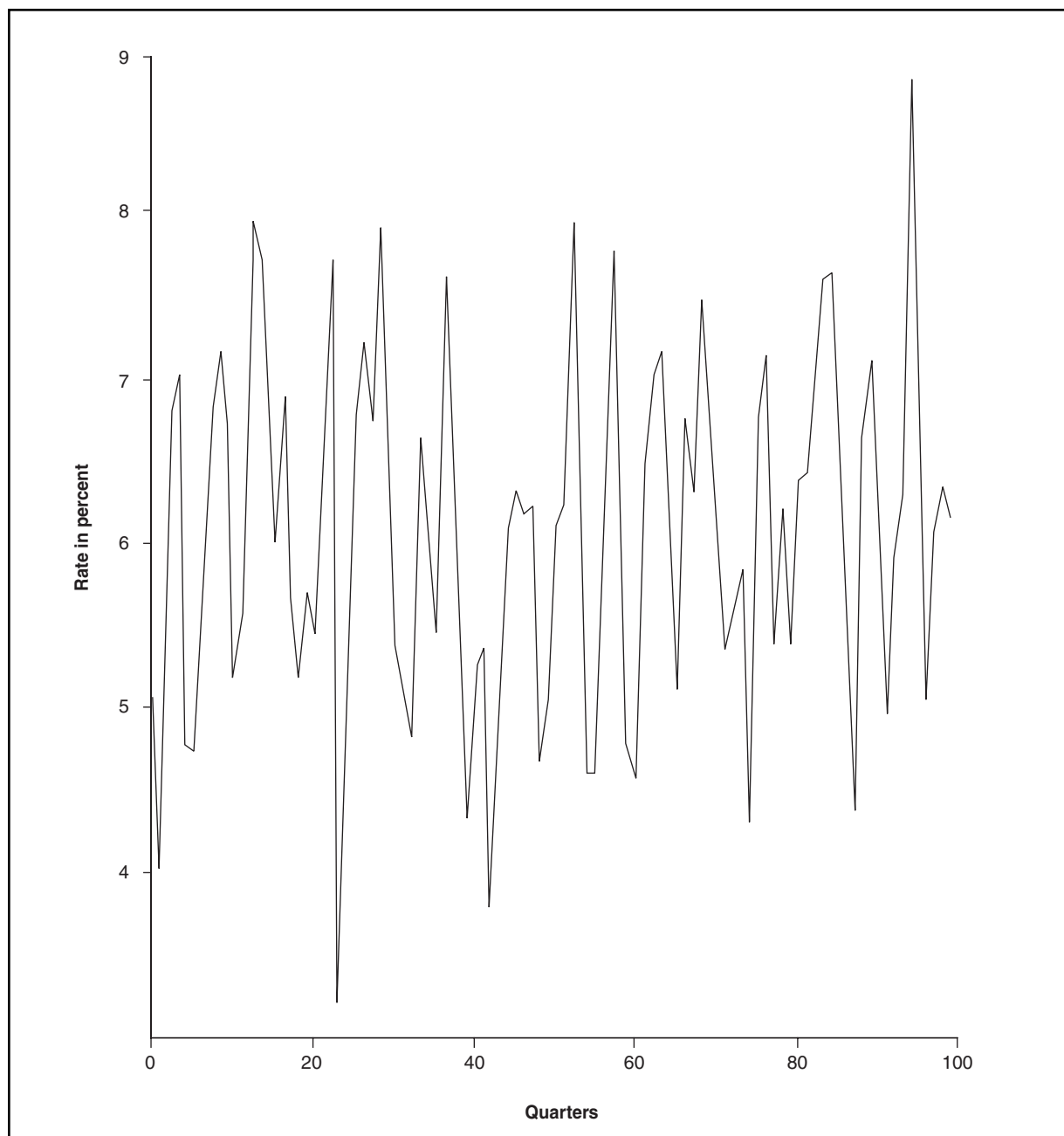


Figure 1. Unemployment Rate by Quarter

residuals in conventional regression analysis). There are no spikes outside the 95 percent confidence interval, indicating that the residuals are probably white noise. That is, the temporal dependence in the data has been removed. One therefore can conclude that the data are consistent with an underlying autoregressive process of order 2, with

coefficients of 0.33 and -0.35. The relevance of this information will be addressed shortly.

To summarize, the diagnostics have suggested that this ARIMA model need not include any differences or a moving-average component but should include an autoregressive component of

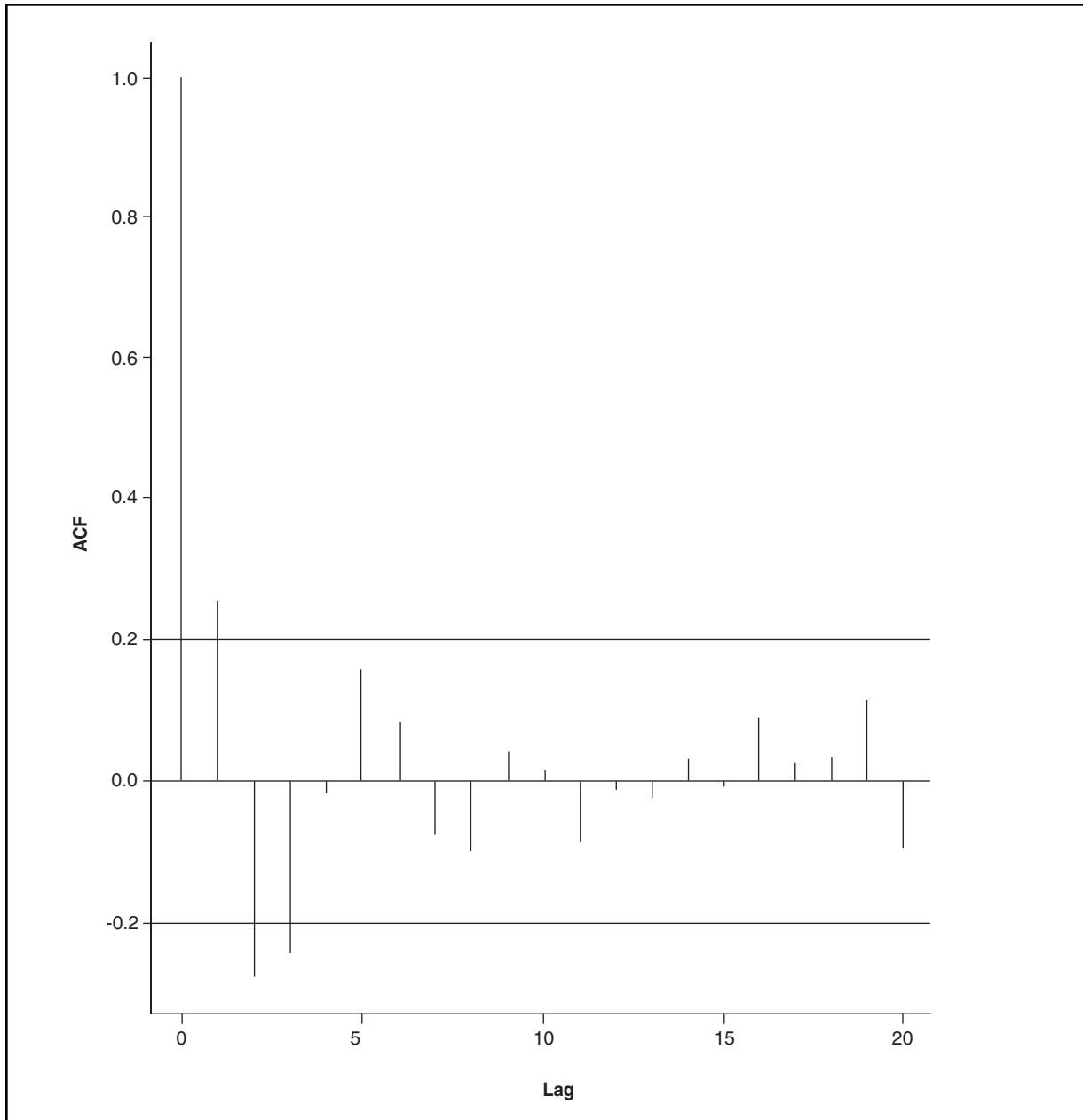


Figure 2. Unemployment Series: Autocorrelation Function

order 2. More generally, the following diagnostic rules of thumb usually are employed, often in the order shown.

1. If the autocorrelation function does not decline rather rapidly, difference the series one or more times (perhaps up to three) until it does.
2. If either before or after differencing the autocorrelation function declines very abruptly, a moving-average component probably is needed. The lag of the last large spike outside the confidence interval provides a good guess for the value of  $q$ . If the autocorrelation function declines rap-

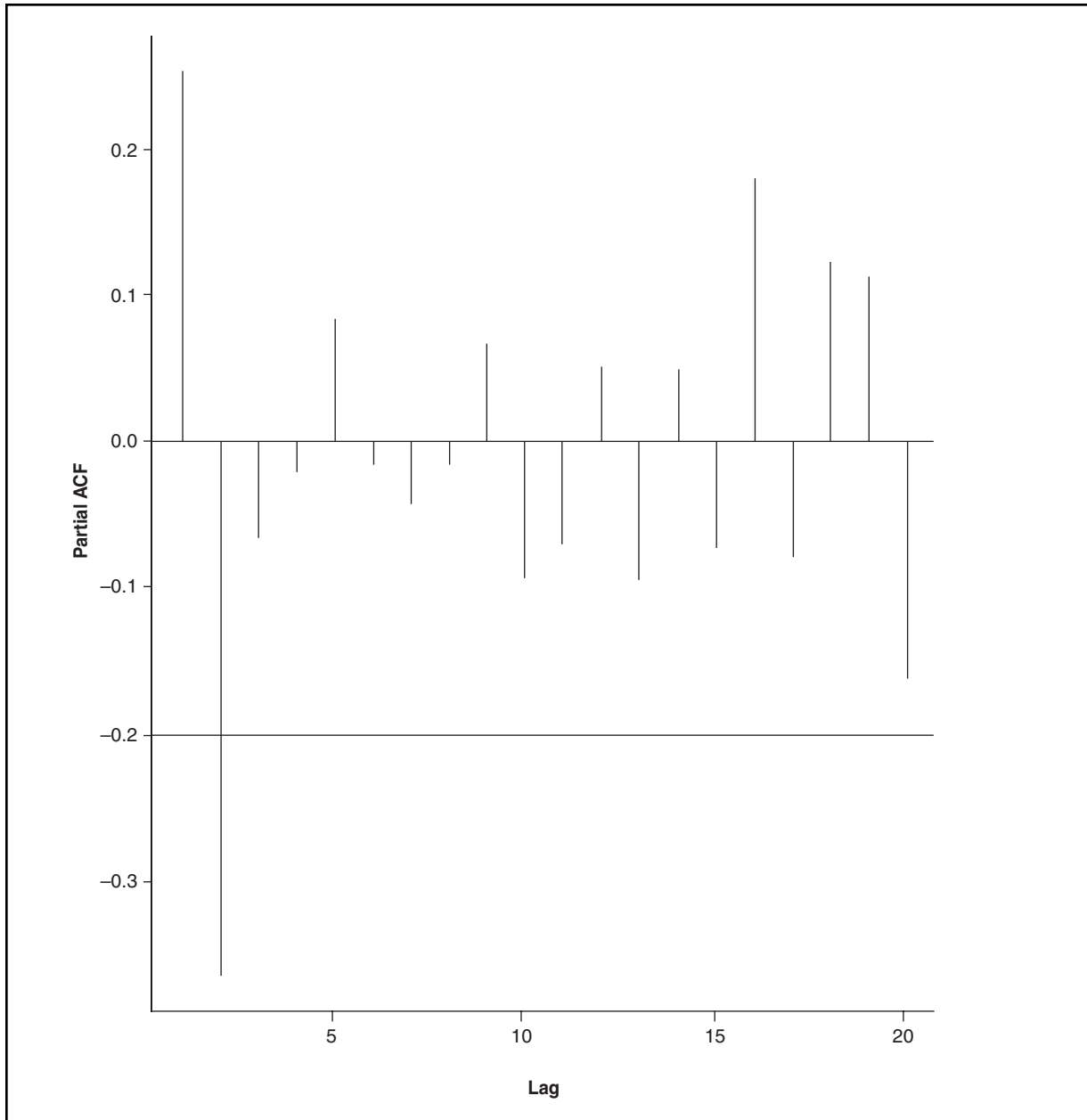


Figure 3. Unemployment Series: Partial Autocorrelation Function

idly but gradually, an autoregressive component probably is needed.

3. If the partial autocorrelation function declines very abruptly, an autoregressive component probably is needed. The lag of the last large spike outside the confidence interval provides a good guess for the value of  $p$ . If the partial autocorrelation

function declines rapidly but gradually, a moving-average component probably is needed.

4. Estimate the model's coefficients and compute the residuals of the model. Use the rules above to examine the residuals. If there are no systematic patterns in the residuals, conclude that the model

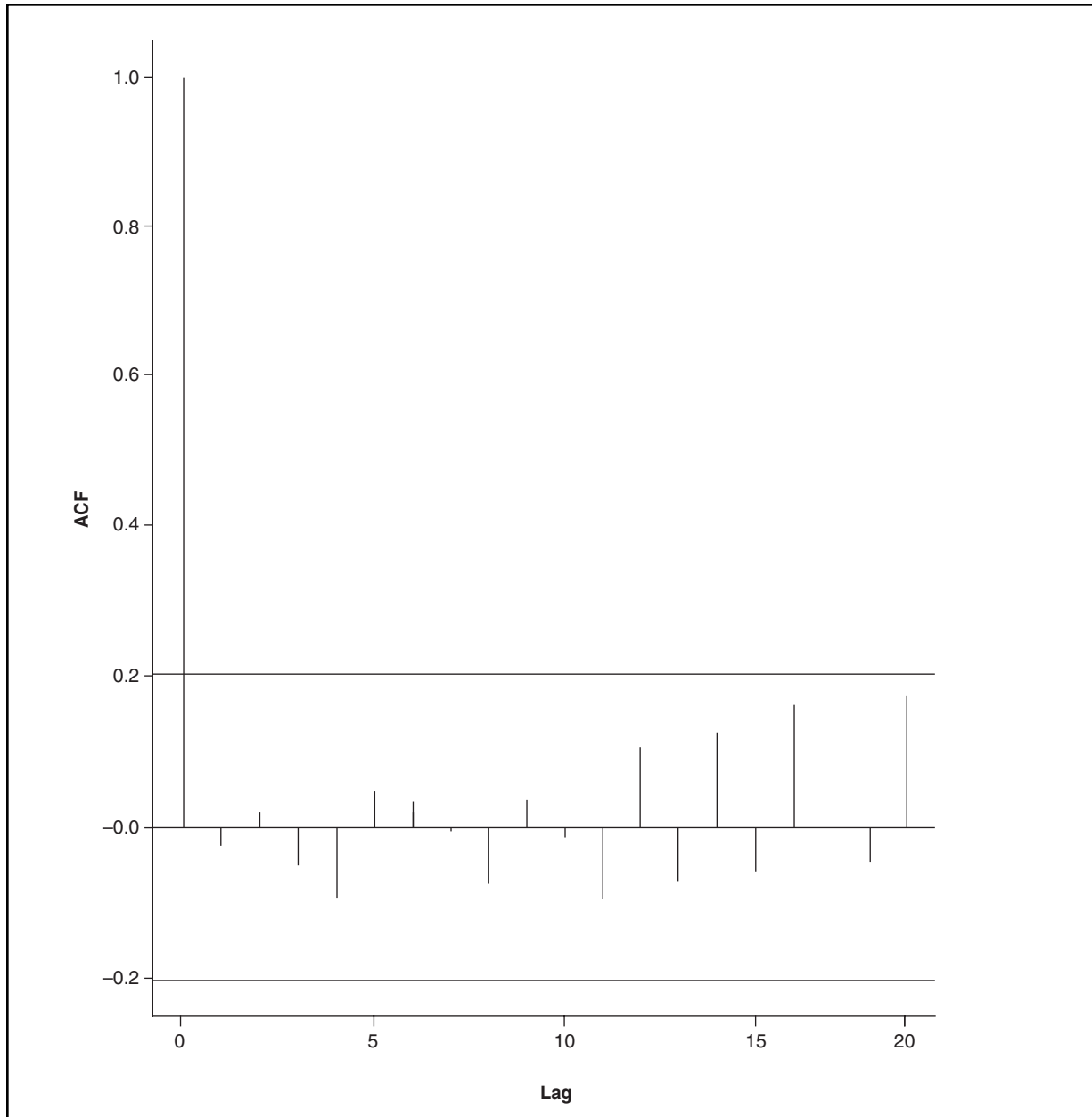


Figure 4. Residuals Series: Autocorrelation Function

is consistent with the data. If there are systematic patterns in the residuals, respecify the model and try again. Repeat until the residuals are consistent with a white noise process (i.e., no temporal dependence).

Several additional diagnostic procedures are available, but because of space limitations, they

cannot be discussed here. For an elementary discussion, see Gottman (1981), and for a more advanced discussion, see Granger and Newbold (1986).

It should be clear that the diagnostic process is heavily dependent on a number of judgment calls about which researchers could well disagree. Fortunately, such disagreements rarely matter. First, the disagreements may revolve around differences



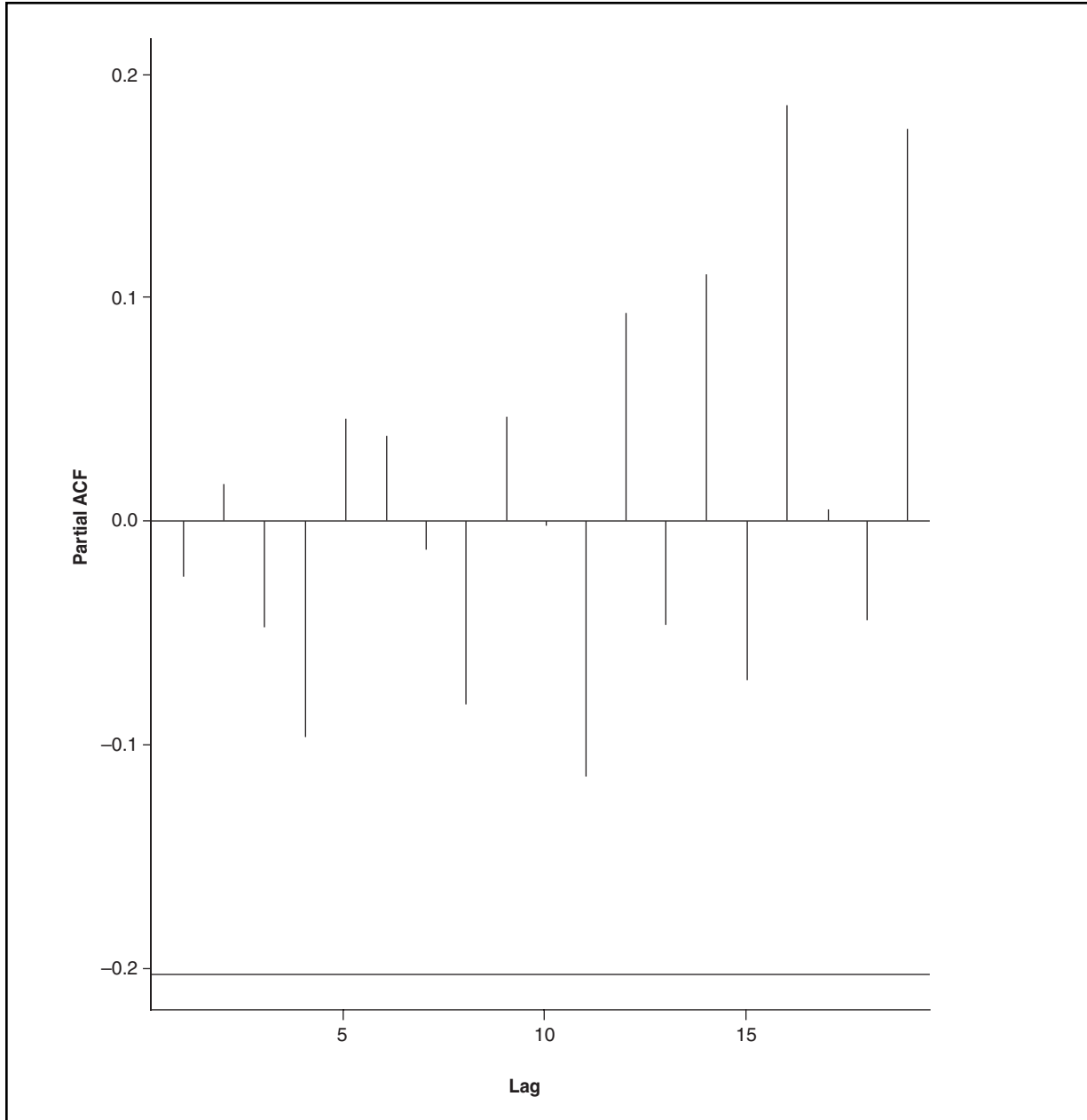


Figure 5. Residuals Series: Partial Autocorrelation Functions

between models without any substantive import. There may be, for instance, no substantive consequences from reporting an MA[2] compared with an MA[3]. Second, ARIMA models often are used primarily to remove “nuisance” patterns in time series data (discussed below), in which case the particular model used is unimportant; it is the result that matters. Finally and more technically, if

certain assumptions are met, it is often possible to represent a low-order moving-average model as a high-order autoregressive model and a low-order autoregressive model as a high-order moving-average model. Then model specification depends solely on the criteria of parsimony. That is, models with a smaller number of parameters are preferred to models with a larger number of parame-

ters. However, this is an aesthetic yardstick that may have nothing to do with the substantive story of interest.

**USES OF ARIMA MODELS IN SOCIOLOGY**

It should be clear that ARIMA models are not especially rich from a substantive point of view. They are essentially univariate descriptive devices that do not lend themselves readily to sociological problems. However, ARIMA models rarely are used merely as descriptive devices (see, however, Gottman 1981). In other social science disciplines, especially economics, ARIMA models often are used for forecasting (Granger and Newbold 1986). Klepinger and Weiss (1985) provide a rare sociological example.

More relevant for sociology is the fact that ARIMA models sometimes are used to remove “nuisance” temporal dependence that may be obstructing the proper study of “important” temporal dependence. In the simplest case, ARIMA models can be appended to regression models to adjust for serially correlated residuals ( Judge et al. 1985, chap. 8). In other words, the regression model captures the nonstationary substantive story of interest, and the time series model is used to “mop up.” Probably more interesting is the extension of ARIMA models to include one or more binary explanatory variables or one or more additional time series. Nonstationarity is now built into the time series model rather than differenced away.

**Intervention Analysis.** When the goal is to explore how a time series changes after the occurrence of a discrete event, the research design is called an interrupted time series (Cook and Campbell 1979). The relevant statistical procedures are called “intervention analysis” (Box and Tiao 1975). Basically, one adds a discrete “transfer function” to the ARIMA model to capture how the discrete event (or events) affects the time series. Transfer functions take the general form shown in equation (1):

$$(1 - \delta_1 B - \dots - \delta_r B^r) y_t = (\omega_0 - \omega_1 B - \dots - \omega_s B^s) x_{t-b}$$

If both sides of equation (1) are divided by the left-hand side polynomial, the ratio of the two polynomials in B on the right-hand side is called a transfer function. In the form shown in equation (1), *r* is the order of the polynomial for the “dependent variable” (*y<sub>t</sub>*), *s* is the order of the poly-

nomial for the discrete “independent variable” (*x<sub>t</sub>*), and *b* is the lag between when the independent “switches” from 0 to 1 and when its impact is observed. For example, if *r* equals 1, *s* equals 0, and *b* equals 0, the transfer function becomes  $\omega_0 / (1 - \delta_1 B)$ . Transfer functions can represent a large number of effects, depending on the orders of the two polynomials and on whether the discrete event is coded as an impulse or a step. (In the impulse form, the independent variable is coded over time as 0, 0, . . . 0, 1, 0, 0, . . . , 0. In the step form, the independent variable is coded over time as 0, 0, . . . 1, 1 . . . 1. The zeros represent the absence of the intervention, while the ones represent the presence of the intervention. That is, there is a switch from 0 to 1 when the intervention is turned on and a switch from 1 to 0 when the intervention is turned off.) A selection of effects represented by transfer functions is shown in Figure 6.

In practice, one may proceed by using the time series data before the intervention to determine the model specification for the ARIMA component, much as was discussed above. The specification for the transfer function in the discrete case is more ad hoc. Theory certainly helps, but one approach is to regress the time series on the binary intervention variable at a moderate number of lags (e.g., simultaneously for lags of 0 periods to 10 periods). The regression coefficients associated with each of the lagged values of the intervention will roughly trace out the shape of the time path of the response. From this, a very small number of plausible transfer functions can be selected for testing.

In a sociological example, Loftin et al. (1983) estimated the impact of Michigan’s Felony Firearm Statute on violent crime. The law imposed a two-year mandatory add-on sentence for defendants convicted of possession of a firearm during the commission of a felony. Several different crime time series (e.g., the number of homicides per month) were explored under the hypothesis that the crime rates for offenses involving guns would drop after the law was implemented. ARIMA models were employed, coupled with a variety of transfer functions. Overall, the intervention apparently had no impact.

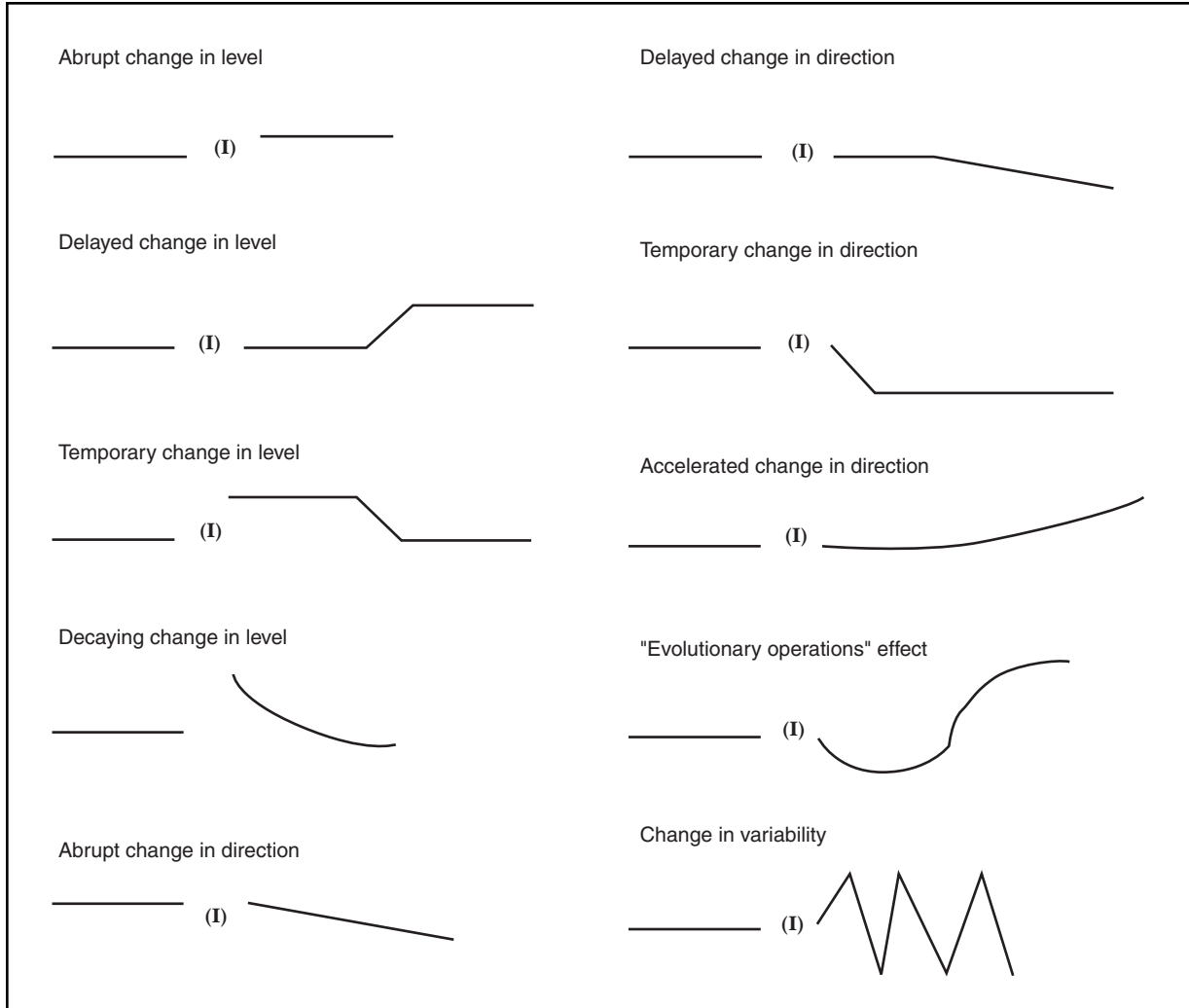


Figure 6. A Sampler of Intervention Effects. *I* = Intervention

**Multiple Time Series.** ARIMA models also may be extended to more than one time series (Chatfield 1996, Chap. 10) Just as the goal for the univariate case was to find a model that transformed the single time series into white noise, the goal for the multivariate case is to find a model that will transform a *vector* of time series into a white noise vector. In effect, each series is regressed simultaneously not only on lagged functions of itself and the disturbance term but on functions of all other time series and their disturbance terms. In practice, this sometimes reduces to building transfer function models that include a single response time series and several input time series, much as in multiple regression. For exam-

ple, Berk et al. (1980) explored how water consumption varied over time with the marginal price of water, weather, and a number of water conservation programs.

The mathematical generalization from the univariate case is rather straightforward. The generalization of model specification techniques and estimation procedures is not. Moreover, multivariate time series models have not made significant inroads into sociological work and therefore are beyond the scope of this chapter. Interested readers should consult Chatfield (1996) for an introduction or Granger and Newbold's (1986) for a more advanced treatment.

## CONCLUSIONS

Time series analysis is an active enterprise in economics, statistics, and operations research. Examples of technical developments and applications can be found routinely in a large number of journals (e.g., *Journal of the American Statistical Association*, *Journal of Business and Economic Statistics*, *Journal of Forecasting*). However, time series analysis has not been especially visible in sociology. Part of the explanation is the relative scarcity of true time series for sociological variables collected over a sufficiently long period. Another part is that time series analysis is unabashedly inductive, often making little use of substantive theory; time series analysis may look to some a lot like “mindless empiricism.” However, in many sociological fields, true time series data are becoming increasingly available. Under the banner of “data analysis” and “exploratory research,” induction is becoming more legitimate. Time series analysis may well have a future in sociology.

(SEE ALSO: *Longitudinal Research*; *Statistical Methods*)

## REFERENCES

- Berk R. A. 1988 “Causal Inference for Sociological Data.” In N. Smelser, ed., *The Handbook of Sociology*. Newbury Park, Calif.: Sage.
- , T. F. Cooley, C. J. LaCivita, S. Parker, K. Sredl, and M. Brewer 1980 “Reducing Consumption in Periods of Acute Scarcity: The Case of Water.” *Social Science Research* 9:99–120.
- Box, G. E. P., and G. M. Jenkins 1976 *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- , and G. C. Tiao 1975 “Intervention Analysis with Applications to Economic and Environmental Problems.” *Journal of the American Statistical Association* 70:70–79.
- Chamlin, Mitchel B. 1988 “Crime and Arrests: An Autoregressive Integrated Moving Average (ARIMA) Approach.” *Journal of Quantitative Criminology* 4(3):247–258.
- Chatfield, C. 1996 *The Analysis of Time Series: An Introduction*, 5th ed. New York: Chapman and Hall.
- Cook, T. D., and D. T. Campbell 1979 *Quasiexperimentation*. Chicago: Rand McNally.
- Cox, D. R., and D. Oakes 1984 *Analysis of Survival Data*. London: Chapman and Hall.
- Freedman, D. A., and David Lane 1983 “Significance Testing in a Nonstochastic Setting.” In P. J. Bickel, K. A. Doksum, and J. L. Hodges, Jr., eds., *A Festschrift for Erich L. Lehman*. Belmont, Calif.: Wadsworth International Group.
- Goldberg, Samuel 1958 *Introduction to Difference Equations*. New York: Wiley.
- Gottman, J. M. 1981 *Time-Series Analysis*. Cambridge, UK: Cambridge University Press.
- Granger, C. W. J., and P. Newbold 1986 *Forecasting Economic Time Series*. Orlando, Fla.: Academic Press.
- Harvey, A. C. 1990 *The Econometric Analysis of Time Series*, 2nd ed. London: Harvester Wheatshear.
- Hsiao, Cheng 1986 *Analysis of Panel Data*. Cambridge, UK: Cambridge University Press.
- Judge, D. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T.-C. Lee 1985 *The Theory and Practice of Econometrics*. New York: Wiley.
- Kessler, D. A., and S. Duncan 1966 “The Impact of Community Policing in Four Houston Neighborhoods.” *Evaluation Review* 6(20):627–669.
- Klepinger, J. D. H., and J. G. Weis 1985 “Projecting Crime Rates: An Age, Period, and Cohort Model Using ARIMA Techniques.” *Journal of Quantitative Criminology* 1:387–416.
- Loftin, C., M. Heumann, and D. McDowall 1983 “Mandatory Sentencing and Firearms Violence: Evaluating an Alternative to Gun Control.” *Law and Society Review* 17(2):287–318.
- , and D. McDowall 1982 “The Police, Crime, and Economic Theory.” *American Sociological Review* 47:393–401.

RICHARD A. BERK

## TIME USE RESEARCH

Time provides the organizational key to action at the level of individuals, groups, and institutions. It also defines a normative framework that regulates interpersonal relationships and allows synchronized operations in different parts of society. In the concept of time, structural as well as symbolic facets assume significance. In fact, as it is conceived in sociological theory, time is a means of social coordination as well as a dimension that assigns value to action schemes in a system assuring social order (Pronovost 1989; Sue 1994).

The empirical study of the temporal organization of human action reveals the functional charac-

teristics of social roles and the societal division of social tasks. For instance, indicators of inequality and social exclusion can be derived and compared by referring to estimates of the amount of time involved in gender-related activities such as market work and housework. The time use of populations or subpopulations is studied mostly by means of diary procedures that assess individual "time budgets:" the sequence, timing, and duration of activities performed by individuals over a specified period. It is misleading to think that the aim of time budget research is time as either a physical or a subjectively perceived entity. As stressed by the major postwar promoter of this area of study, the Hungarian sociologist Szalai, the object of studying time is to discover the use people make of their time, or "the arrangement and the fit of people's activities in a temporal frame of reference, the temporal order and structure of everyday life" (1984, p. 20).

From this point of view, time is a special kind of resource. As with material goods and more symbolic commodities such as money, people have a "fund" of time at their disposition and make decisions on how to use, "spend," or "invest" it. However, time is a far more democratic resource in that every person deals with the same basic "stock" of it, such as the twenty hours in a day and the seven days in a week. This means that a shared reference of differential time allocation patterns can facilitate coherent comparisons and meaningful interpretations. It is this particular aspect that has made time use an important topic in quantitative social research. Both academic scholars and national statistical offices have shown a growing interest in time-budget data because those data permit policy-oriented microanalyses of changing lifestyles at the individual or household level as well as macroanalyses of social and economic inequalities in the context of cross-national comparative studies.

In contrast to the physical notion that attributes equivalent temporal resources to all people and therefore facilitates systematic accounts, from a subjective point of view, the length of a day is not always the same. Some people seem to have more time than others. This phenomenon reveals the sociopsychological dimension of time in which concepts such as stress and alienation are relevant. In a world where the quality of life depends largely

on what one gets out of the time at one's disposal, coping with time and compressing multiple activities into the same time slot have become important skills. Furthermore, it is a universal observation that boring periods during a day, week, or year can seem long, whereas other, quantitatively equivalent but exciting periods of time are perceived as a passing moment. However, in studying social time, time-budget research adopts an activity-oriented approach and focuses only indirectly on subjective experiences. In substance, this means that observable patterns of behavior are selected as primary evidence and, in methodological terms, standard time units are chosen for measurement purposes.

### THE TIME USE RESEARCH TRADITION

Although social time is not intrinsically quantitative, the use of standard time units for the purposes of analyzing the structure of everyday life seems legitimate, since the transfer of human work from agriculture to artificially controlled industrial environments and the subsequent changes in civilization have largely transformed natural time (tied to seasonal conditions and biological needs) into conventional, rational time (Elias 1988). As a corollary to this process, social life has become dominated by timekeeping. This chronometric function is characterized by a universally accepted "time" language that coordinates rhythms of action in the public and private spheres (Zerubavel 1982). Therefore, time budgets, which are concerned with different kinds of schedules for structuring the flow of events, are a key to the systematic investigation of the complex interdependencies and trade-offs of modern life.

The historical origins of the study of social time lie in both sociological theory and empirical research. With regard to theory, the early French school of sociology was interested in this phenomenon from the point of view of the historical and anthropological dimensions of social change (Pronovost 1989). Around the turn of the century, Hubert, Mauss, Durkheim, and Halbwachs conceived of social time as an intrinsically qualitative phenomenon that was relevant for the characterization of the sacred-profane symbolic dichotomy in the evolution of the collective consciousness or

the formation of a collective memory. Their focus was a macro one, and their main interest was to explain long-term cultural change.

In contrast, Mead in the early 1930s chose an individualistic, micro-oriented approach that offered a philosophical rationale for studying time in the present. The present was conceived of as the context for the emergence and assimilation of various social time systems in interplay with the definition of different notions of "the self." Thus, the French and American approaches represent polar opposites, with one conceiving time within the matrix of historical societal relationships, and the other from the perspective of mutable configurations of symbolic interactions in small groups. The somewhat later work of Sorokin reflects both approaches. He presents the functionalist idea that the plurality of individual time schemes requires extensive synchronization to achieve social cohesion and that time expresses the sociocultural "pulsation" of a society.

In the 1960s, Gurvitch gave new impetus to the study of time after a considerable period of neglect. He was the first to conceive of time as an important source of contradictions and potential conflict. In particular, he stressed the hierarchically diversified aspects of the phenomenon (e.g., social time, time in organizations, time in special social groups) and raised the issues of power and legitimacy. More recently, Merton introduced the concept of "socially expected durations" that highlight the normative aspects of the embeddedness of time in social structures. In contrast, Elias stressed the role of time as a symbolic means of social regulation but also of increasingly unpredictable individual self-expression.

The historical origins of empirical investigations of social time are even older, going back to the middle of the nineteenth century. Three lines of research are significant: the research conducted by Friedrich Engels on the English working class, where the temporal organization of daily life was the issue; the studies undertaken by Frédéric LePlay, in which the economic "family-budgets" of workers in several European countries were assessed (similar to what "time budgets" do today); and the experimental work on time and motion performed by Frederick Taylor, based on carefully collected chronometric data. Taylor's aim was to introduce

strategies of scientific time management in industry. In regard to the adoption of time-budget methodologies, there were Bevans's pioneering studies (1913) of how workers spent their spare time and early Soviet inquiries by Strumilin into time use as the basis for rational social programming. Lundberg and Komarovsky's research into the organization of time within the realm of community research was conducted along the lines of American cultural anthropology.

Of more enduring interest, however, were two studies published in the 1930s. Jahoda et al.'s *Marienthal: The Sociography of an Unemployed Community* (1933) was a substantial contribution to the study of time use. It explores changes in the meaning of time for German working-class families when work, as a dominant regulating and legitimizing criterion for time use, has disappeared. Male workers tended to become severely disoriented and alienated after losing their work-based prestige, whereas their wives were successful in mastering that situation because they had much more positive attitudes based on more complex sources of social recognition. This research also shows how important it is to be aware that similar circumstances can assume diverse meanings for different groups.

The other study, whose importance lies in its methodological ideas, is Sorokin and Berger's *Time-Budgets and Human Behavior* (1939). Here the aim was to explore meaningful criteria for decision making conducive to different time structures. Information on motivations and the kinds of experiences associated with certain practices and future projects was collected to acquire a deeper understanding of how people deal with their time. Even more important, this research raised the crucial epistemological question of how to divide essentially continuous strings of behavior into activity segments that, beyond commonsense classifications, can be grouped into homogeneous and mutually exclusive categories (Kurtz 1984). The difficulty was that some activities that from an external viewpoint seemed identical could assume unequal functions in the eyes of those concerned, or conversely, that substantially different activities could assume similar functions. This means that any classification of activities presupposes an interpretive act. With this fundamental problem

in mind, these authors were forced to approach earlier and purely descriptive assessments of time allocation with skepticism and spell out the methodological issues in their research.

After World War II, research took different directions in accordance with divergent political ideologies. The assessment of living conditions, which involved obtaining background data for economic planning and monitoring centrally initiated social change, continued to be of pivotal importance in research in eastern Europe's communist countries. Prudensky's time-budget studies in the Soviet Union in the 1960s not only followed the direction of Strumilin's work but also reflected these ideological concerns. The need to broaden data gathering to obtain effective guidance for public policy at the national level led Hungary's statistical office to begin the first microcensus research in this field.

In capitalist societies, time use research was concerned principally with mass media and leisure culture. Pioneering time use studies of audiences were undertaken by the BBC as well as NHK, the Japanese radio and television system, using large-scale survey techniques. From 1960 onward, NHK conducted regular five-year follow-up rounds of research to obtain time series statistics that showed long-term longitudinal development. This was a useful strategy because it produced an interesting account of how, in terms of time use, traditional ways of life are supplanted by innovative, primarily television-centered styles.

This was the situation in 1963, when the idea of conducting a Multinational Comparative Time Budget Research Project emerged. This was an ambitious sociological initiative in light of the organizational and data-processing difficulties of those years. Launched by a group of scholars directed by Szalai, sponsored by UNESCO's International Social Science Council, and coordinated by the Vienna Centre, this project attempted to obtain an interculturally valid body of knowledge that would shed light on regularities or variations in the functioning of human societies with regard to time use. This information was to be derived from a database of twelve different countries by using methodological instruments that assured a high level of analytic precision (Szalai 1977). In organizing the initiative, the basic concern was to

avoid the emergence of a single central vantage point regarding the collection, elaboration, and interpretation of information. Therefore, research sites had considerable autonomy in studying the uniform data sets collected by means of strictly standardized survey instruments from probability samples of urban populations in the twelve countries under investigation.

From a positivist point of view, the focus on chronometric evidence and on an array of "hard" time use indicators enhanced the scientific character of the study and facilitated the collaboration of teams from such culturally and sociopolitically different environments as the United States and the Soviet Union. Of course, collaboration entailed the acceptance of common working hypotheses such as the expected influence of the major independent variables of industrialization and urbanization on the modalities of the division of market work and nonmarket work in households. By contrast, time for leisure was thought to be correlated with superior levels of modernization and democratization. These hypotheses clearly reflected the research traditions of the day, and so to connect the ideologically distant worlds of the 1960s, it was necessary to choose highly conventionalized and neutral time indicators as empirical evidence.

The Twelve Country Project, characterized by a strong belief in the "scientific *and* social import of cross-national comparative research" (Szalai 1977), did not go without criticism. Some thought that it was most important for cross-national research to contribute findings on general theoretical problems (Przeworski and Teune 1970). However the promoters were convinced that the discovery of the empirical peculiarities of cultural settings was at least as important as the verification of a priori hypotheses on common characteristics and trends. That the pragmatic point of view prevailed meant that the problem of a lack of reliable, relevant, and usable data had to be overcome.

In retrospect, it seems that this project did not contribute much to general theory, but it did produce an elaborate methodology whose essential lines are applied to basic and official survey research in many countries today. In fact, as Szalai hoped, the homogenization of time-budget methods now permits the drawing of "maps" of collec-

tive daily activity schemes at different levels of definition that have proved to be useful diagnostic elements for many policymakers and grassroots organizations. When economic indicators are insufficient, statistical information regarding the use of time can open up new policy perspectives and guide substantial change, especially when gaps in the quality of life manifest themselves and corrective action is needed to improve the conditions of disadvantaged social groups.

#### **METHODOLOGICAL ASPECTS OF TIME-BUDGET RESEARCH**

Since the work of Szalai, the methodology of time use research has been further refined. Under the auspices of the International Association of Time Use Research, in particular under the guidance of Harvey (1984, 1993), who has repeatedly codified the best practices, statistical bodies have reached a consensus on the format of official survey research.

The task of discovering the temporal order and structure of everyday life by means of time-budget methods involves far more complex activities than gathering simple answers to questions of who does what, when, where, and with whom. In the design of a time-budget study, methodological issues such as the scope and scale of the research, the population from which the sample is to be drawn, the format of the data-gathering instruments, the classification and coding of activities, the choice of basic indicators, and the validity and reliability of the data must be resolved.

Defining the scope of a study also means fixing its scale. In fact, a time use survey may deal with a special group of persons, such as working women or teenagers, or with comprehensive national populations, perhaps excluding preschoolers. Or it may focus on a daily activity such as housework and child care or leisure. Alternatively, it may attribute equal weight to all everyday pursuits. Finally, the study may be aimed at discovering how the time of a special kind of day is spent, may take into account the rhythm of the week by distinguishing workdays from Sundays, or may be interested in longer periods such as the year with its seasonal differences and, in the extreme case, the life cycle (which of course would have to be studied on the basis of long-term recollections). Theo-

retically grounded sociological research is, for economic reasons, more likely to have circumscribed objectives, whereas national statistical offices have all-encompassing multipurpose datasets available. However, in both cases, the predominant tendency is to focus on the twenty-four hours of one or more single days in the life of the respondent. When these days are distributed over the week, month, or year, the average profiles of the period can be synthetically reconstructed. Such profiles most often refer to uninterrupted sequences of nonoverlapping main or "primary" activities. When there is interest in "secondary" activities or, more precisely, in contemporaneous activity episodes, the respondent usually is asked to designate the elements that represent the principal flow, which typically covers the 1,440-minute arc of the day. In fact, leaving secondary activity out of focus furnishes an unduly simplified picture of what is going on, since it ignores efficiency strategies that enable people who are short of time to deal simultaneously with multiple jobs. This frequently criticized weakness is compensated for by the heuristically valid fact that strict 24-hour accounts produce agile descriptive models. With such models, whatever time is saved on one kind of activity is strictly accredited, using zero-sum logic, to one or more other activities. Therefore "time set free" and the equivalent "time gained" concept furnish concise indicators of social change. Tracing the balance of the two magnitudes gives a dynamic slant to the analysis of time use and sheds light on the spectrum of strategic options.

Depending on the scope of the study, populations and samples are variously defined. From this point of view, the most significant difference between time use studies regards the choice of the sampling unit, which may be the individual or the household. In earlier studies, individual time use was of primary concern, and so estimates were obtained by classifying persons by their sociodemographic characteristics. More recent research, however, has looked more into how different types of families, as molecular units, manage time allocation with regard to income generation as well as work in the sphere of home and child care. Statistical offices now use very large probability samples of households to be able to generate cross-tabulated data on specific territorial areas and particular social groups. In Italy, for instance,



the last national time use survey, conducted by Istat in 1988 and 1989, consisted of more than 38,000 persons belonging to almost 14,000 households. A survey conducted in Germany in 1991 and 1992 by the Statistisches Bundesamt included 7,200 households. One of the most difficult problems in time-budget research is the sample units' frequent refusal to respond once they see how much time is involved. In fact, nonresponse rates tend to be high and in some official surveys amount to almost 30 percent. This problem is easier to handle in smaller-scale studies, which often use quota sampling.

Data gathering in time use research begins with an interview (Scheuch 1972) to record the characteristics of the respondent and his or her family, contractual work arrangements, normal labor supply, and housing or other assets and to inquire into irregularities in the day designated for collecting the time-budget information. The time budget itself is registered in a protocol, a diary, or modular display where the beginning and the end of each activity can be indicated together with other information. The resulting datasets show for each day and respondent (1) the number of different activities performed and the frequency of each activity in separate episodes (for instance, the series of daily meals or the periods passed in front of the television set) and (2) the timing, duration, and sequence of activities or activity episodes. Most often, the interviewees register activities by using their own words. A grid of minimal time intervals is given (the "fixed interval" solution), where the task is to fill each interval with an activity, or the interviewee is asked to specify the exact time points of his or her schedule (the "open interval" solution). To obtain the essential elements of the interviewee's context, there is usually room to indicate contemporaneous activities (for instance, reading while using public transport or listening to music while doing homework); participation in activities with family members, neighbors, friends, and colleagues; and where the activity takes place. The least expensive method of data collection is the condensed telephone interview, which explores time use on the previous day. For field studies, there are other procedures, such as single face-to-face interviews and two personal interviews. In the first case, the person is asked to recall what he or she did the preceding day. This

procedure is complicated when a lot of detail is required. The second procedure involves two personal interviews. During the first, background information is collected and the time use diary is left behind, to be filled in the next day. During the second interview, on the day after the respondent's observation of his or her time use, the interviewer checks and refines the registrations. In Scandinavian countries, people were asked to return diaries by mail. This saves a second visit, but it is advisable only when intelligent and conscientious collaboration can be assumed.

The greatest methodological challenge in time use research is the choice of the scheme of classification of activities in terms of which the structure of everyday life is represented. Sorokin started to tackle this problem, but convincing theoretical or empirical criteria for constructing typological keys have not been found, and using conventional categories of ordinary language is not entirely satisfactory. Normative and/or contractual work arrangements suggest a fairly unambiguous specification of "market work," but there are some activities in the home that, according to circumstances, can be classified as either housework or leisure. This difficulty could be overcome if the respondent did the coding himself or herself, but usually the log of daily routines is described in the respondent's own words and codification is done by someone else, following criteria that exclude personal and/or subjective meanings. An even more fundamental issue is whether current classifications can be assumed to be meaningful in cross-cultural terms. Time use studies distinguish the minimal basic activity groupings of personal needs, formal work or education, household work, and leisure. The hidden dimension that is postulated by such groupings is obviously a reflection of the Western opposition between necessity and freedom of choice. It places market work immediately after biological needs and before domestic work, which is placed near leisure. This implicitly individualistic and work-oriented, contractual rationale probably is not well suited to representing the more solidarity-oriented temporal orders of everyday life in traditional societies (Bourdieu 1963).

Another difficulty concerns the level of specificity at which a common array of activities is reported at the collective level. The daily pursuits

of persons who lead a busy life can be meaningfully recorded in great detail, whereas those of persons tied to the home usually have much less texture. One way to approach this difficulty is to construct hierarchical coding frameworks in which the first column in a multiple-digit code divides the day in terms of major classes of activities. Additional columns focus on increasingly more complex but exhaustive time accounts. The time use project coordinated by Szalai identified in its time-budget protocols ninety-six activity categories. For some purposes, these were reduced to thirty-seven and, for others, to the following ten main groups: work, housework, child care, shopping, personal needs, education, organizational activity, entertainment, active leisure, and passive leisure. Today the coding schemes for official statistical surveys often include many more basic activity categories because they have to accommodate the heterogeneity of lifestyles across gender groupings, generations, occupational categories, and rural versus urban residential environments.

Once time-budget data have been collected and coded, decisions about data processing and indicator construction can be made. According to the complexity of statistical data elaborations, there are three different levels of analysis (Stone 1972, pp. 96-97). First, activity arrays in terms of frequencies or durations, possibly taking company or locations into account, are cross-tabulated with the sociodemographic characteristics of the actors. Second, single activities and their positioning during the course of the day are studied. Finally, stochastic activity sequences are analyzed, focusing on the structure and rhythm of chronological daily routines. Since the 1960s, the following set of indicators generally have been used in computations: (1) the generic average duration of an activity, where the numerator refers to the total sample, disregarding whether there was involvement in the activity, (2) the rate of participation, or the percentage of interviewees who were involved in an activity, and (3) the specific average duration of an activity, where the denominator includes only those who have engaged in it.

Another issue in these surveys is the validity and reliability of data sets. Questions of validity can be raised by difficulties in recall and incorrect identification of activities among respondents or

by possible alterations of spontaneous behavior after observation and the consequent distortions in reports or by research instruments that inadequately reflect the specificities of the observed sociocultural context. Research directed at data quality (Juster 1985; Niemi 1993) has shown that results obtained by means of time budgets present at the aggregate level a high correlation with those obtained by means of other forms of observation, such as interviews, workplace or school statistics, and telephone surveys. Moreover, the hypothesis that the desirability or social prestige of certain activities or lifestyles could influence time use reports has not been confirmed. In general, it seems that the twenty-four-hour frame of reference helps reduce such effects and brings informal and often undeclared work commitments to light. Certainly, activities that are assumed to be of secondary importance, such as conversations and listening to the radio, are under represented in current summary tables that restrict the attention to "primary" time allocations. However, this cannot be considered an invalidating shortcoming. Nevertheless, to assure validity, time budgets presuppose the concept of rational time. If a population does not live by the clock, any calculation of time budgets is meaningless.

#### CONTRIBUTIONS OF TIME USE RESEARCH

Time use research has gained momentum because of interest on the part of international agencies in comparing the functioning of societies in their national settings, the need to connect demographic change and social development, the need to focus on gender-related or generational variables to understand the changing role of the family, the awareness that economic variables reflect wealth and well-being only in very partial ways and that household and care activities must be brought into focus, and the need to construct articulated databases for decision making about social policy.

Since the late 1980s, nationally representative time use data sets have been available for several countries, but the evidence is not easy to compare because activity classifications do not always coincide. Therefore, with the hope that many European Union countries will participate in the very

**Daily Time Use, in Hours and Minutes, for Primary Activities in the United States, France and Hungary, by Gender and Employment Status (1965–1966)**

	United States (44 cities)	France (6 cities)	Hungary (Gyorz)
<b>Employed men</b>			
Personal needs	10:14	10:55	9:56
Work or study	7:36	7:29	8:47
Household work	1:17	1:33	1:48
Leisure	4:06	3:34	3:05
Non-work related travel	0:47	0:29	0:24
<b>Total</b>	<b>24:00</b>	<b>24:00</b>	<b>24:00</b>
<b>Employed women</b>			
Personal needs	10:20	10:53	9:30
Work or study	5:39	6:05	7:14
Household work	3:43	3:57	4:44
Leisure	3:33	2:40	2:03
Non-work related travel	0:45	0:25	0:29
<b>Total</b>	<b>24:00</b>	<b>24:00</b>	<b>24:00</b>
<b>Unemployed women</b>			
Personal needs	10:24	11:13	10:38
Work or study	0:35	0:09	0:52
Household work	7:12	8:13	9:19
Leisure	4:53	3:49	2:35
Non-work related travel	0:56	0:36	0:36
<b>Total</b>	<b>24:00</b>	<b>24:00</b>	<b>24:00</b>

Table 1

SOURCE: Adapted from A. Szalai, ed., *The Use of Time*, Statistical Appendix Table IV.4., 1972, p. 681.

expensive data-gathering process, Eurostat is preparing a standardized survey. Up to the present, only Japan has truly comparable five-year time series data to indicate trends and changes in life-style since 1970 (NHK 1991). In terms of Monday-to-Friday behavior, for instance, sleeping time and housework have decreased while market work has not. Over the years, leisure on all weekdays, especially hobbies, private lessons, and sports activities have grown steadily. Television viewing time reached a peak in 1975 (probably because of the advent of color television) but returned in 1990 to the 1970 level. The illustrations confirm the rule that trend-setting evidence of new life styles is found not so much in the main activity categories but in apparently marginal activities.

To demonstrate the interest of time-budget data in a comparative assessment of the different logics of time structurization, it is best to choose an example from the uniform data set gathered in the Twelve Country Study.

In Table 1, the patterns of daily urban time allocation ascertained in 1965–1966 for the United

States, France, and Hungary are presented in the most synthetic form. The data refer to an average weekday and compare time use for personal needs (mostly sleep and meals), market work or study, household work, leisure, and non-work-related travel among adults aged 18–65, subdivided by gender and employment status.

When one compares the starkly different research contexts of those days (market economy, welfare state, state-controlled system), a set of clear-cut differences in time use emerge from these different ways of life; at one extreme the United States and at the other Hungary, with France in between. The average duration of market work was much shorter in American cities than in Hungarian ones; this was due mainly to the fact that employed women were more likely to hold part-time jobs in the United States, while in Hungary they held full-time jobs. Across the three countries, household work and leisure show systematic secular trends. Employed American men enjoyed one hour more of leisure and contributed half an hour less to housework than did their

Hungarian counterparts. For employed American women, housework lasted one hour less, and leisure lasted one and a half hours more than was the case for employed Hungarian women. Finally, household work among unemployed American women required two hours less time and leisure benefits lasted two hours longer than was the case for the corresponding group in Hungary. The uneven availability of household appliances and unequal access to leisure amenities (in particular, television) were the causes of these differences in lifestyles. In addition, the table reveals well-known inequalities in gender and employment status. If one combines the time invested in market work and housework, it appears, though less significantly in the west than in the east, that women in the labor force contributed a considerably larger share of work and enjoyed much less leisure than did employed men.

Table 2 shows that gender differences are clearly implicated in the discrepancies in the number of hours of economic activity and housework per week between men and women, indicating that there are analogous patterns of inequality in developing countries such as Bangladesh, India, and Nepal. In these countries, women are more likely to spend their time in subsistence activities, whereas men tend to have a monopoly of paid jobs. In addition, women's time investment in housework is six times that of men in Bangladesh and three times that of men in the other two countries. However, the gap in overall work hours, though disadvantaging women, is less evident. In fact, in Bangladesh they contribute 54 percent of the total micro-productive time input, in India 55 percent, and in Nepal 58 percent.

Data from the French national survey (Insee 1989) on the effect of cumulative social roles among women are much more analytic. In France, for mothers with husbands under 45 years of age and at least one child younger than 25 years old, increasing from one child to three and more children means, if they are unemployed, an increase of one hour and fifteen minutes of house work and, if they are employed, an increase of fifty-two minutes. Where does this extra time come from? In the case of especially pressured employed mothers, the time investment in human capital in the form of caring for additional children implies a

reduction of one hour and thirteen minutes in the duration of market work, including a seven-minute reduction in free time and fourteen minutes less of sleep. Especially for women with more than one child, this negatively affects their competitive position in the professional world. Data such as these should be of interest to policymakers.

This example also shows that considerable caution is required in interpreting time use data. In general, estimates of differential time allocations for men and women in certain activities do not reflect only gender differences. Demographic variables, especially family composition, the structure of the labor force, and the availability of household help, intervene in causal links between gender and time use. Particular attention must be paid to these influences in longitudinal analyses such as that of Gershuny and Robinson (1988), which analyzed U.S. and British time-budget data over three decades. Statistically controlling for female labor force participation, male unemployment, and declining family sizes, those authors concluded that in the 1980s, women did substantially less housework while men did a little more than in the 1960s. In another study that analyzed data from repeated surveys in eight Western countries, a general reduction in time dedicated to all kinds of work was found, along with a convergence of time use models among males and females and a growing international similarity in the patterns of the division of time between work and leisure (Gershuny 1992).

In recent decades, considerable progress has been made in representing the multidimensionality of time use phenomena because official data, instead of regarding samples of randomly chosen individuals, have been collected from all members of households. This has made it possible to observe how husbands' time management affects their wives and vice versa and to determine what it means for families if both husband and wife are employed and if children come into the family nucleus.

In Table 3, pertinent data on couples from the national time use survey conducted in Germany in 1990-1991 are presented. It shows time use models for types of families defined by employment status and the presence of children. Assuming the operation of compensatory mechanisms, it also

**Time Use in Three Southern Asian Countries (1982–1992)**

Countries:	Social Groups:	Hours per week of economic activity			Hours per week of housework	Total work hours per week
		Paid	Subsistence	Total		
Bangladesh	Ages 5+ Women	14	8	22	31	53
	Men	38	3	41	5	46
India	Ages 18+ Women	28	7	35	34	69
	Men	43	4	47	10	57
Nepal	Ages 15+ Women	18	17	35	42	77
	Men	29	12	41	15	56

*Table 2*

SOURCE: United Nations, *The World's Women 1995: Trends and Statistics*, Chart 5.3, 1995, New York.

takes the weekly rhythm of time use into account by distinguishing workdays (Monday to Friday) from weekends (Saturday and Sunday).

One model regards more traditional couples in which only husbands are employed and wives do most of the housework. Husbands increase their market work when there are children, but regardless of the presence of children, they defend their daily leisure and contribute to housework mostly on weekends. The other model concerns couples in which both partners are employed. In this case, it is not surprising that the wives' market work is considerably shorter than the husbands', but what is important is that in the presence of children, both partners increase their market work by more than one hour each. The housework of mothers increases on all days, whereas fathers limit increases in their domestic chores to the weekends.

From a micro perspective, what the German example shows are the implications of decisions made within the family. Here the family is seen as the institutional arena where partners search for a suitable compromise in their interlocking role definitions. From a macro perspective, the implications of gendered time use arrangements for the changing division of labor and the growing interaction between the market sector and the household sector are important. Sociologists and economists have often criticized the fact that mostly female domestic and caring activities, mostly male "do-it-yourself" repair initiatives, and voluntary

and other socially useful work done by both men and women go unrecorded in labor statistics and national accounts. These are "productive activities" insofar as they can be delegated to persons other than those who benefit from them.

Table 4 shows a selection of the results of a United Nations Development Programme analysis of a posteriori standardized time budget data from the most recent national surveys conducted in fourteen different countries (Goldschmidt-Clermont and Pagnossin-Aligisakis 1995). This analysis distinguishes between market oriented System of National Accounts (SNA) activities considered in the UN System of National Accounts and non-SNA activities, and introduces the necessary controls for the demographic structures of the populations.

Despite the nonhomogeneous social structures and value systems of France, Germany, Great Britain, and the United States, everywhere statistically unrecorded (non-SNA) activities absorb about as much labor time as do recorded (SNA) activities. Furthermore, total economic time allocations (SNA plus non-SNA) tend to be equal among men and women. Although this demonstrates social equality in general terms, it can be seen that very strong gendered divisions of tasks prevail in all cases. In fact, in these four countries, women contribute only one-third of total market-oriented productive time, whereas they contribute two-thirds of total non-market-oriented productive time.

The availability of comparable time budget data is a prerequisite for official statistics that aim

**Workday and Week-End Time Use of Husbands and Wives in German Households defined by the Employment Status of the Couple and the Presence of Children in the Family (1990-1991).**

Activity categories	Husbands				Wives			
	Only husband employed		Both employed		Only husband employed		Both employed	
	Monday-Friday	Saturday-Sunday	Monday-Friday	Saturday-Sunday	Monday-Friday	Saturday-Sunday	Monday-Friday	Saturday-Sunday
<b>Without children</b>								
Personal needs	9:43	12:02	10:01	12:02	11:40	12:44	10:39	12:16
Market work and study	8:23	1:51	7:11	1:37	0:13	–	4:26	0:37
Household work	2:02	3:05	2:30	3:05	6:58	4:53	4:51	4:46
Free time	3:36	6:46	4:01	6:58	4:57	6:14	3:52	6:13
Other activities	0:16	0:16	0:17	0:18	0:12	0:09	0:12	0:08
<b>Total:</b>	24:00	24:00	24:00	24:00	24:00	24:00	24:00	24:00
<b>With children</b>								
Personal needs	9:25	11:43	9:22	11:39	10:24	11:31	9:46	11:42
Market work and study	8:44	1:16	8:49	1:23	0:24	–	5:36	0:45
Household work	2:16	3:59	2:23	3:49	8:55	6:18	5:23	5:45
Free time	3:22	6:37	3:15	6:47	3:57	5:51	3:06	5:42
Other activities	0:13	0:25	0:11	0:22	0:20	0:20	0:09	0:06
<b>Total:</b>	24:00	24:00	24:00	24:00	24:00	24:00	24:00	24:00

Table 3

SOURCE: Adapted from Statistisches Bundesamt, *Die Zeitverwendung der Bevoelkerung*, Tabellenband I, 1995, Wiesbaden.

to include the production value of nonmonetized activities, through “satellite accounts,” in their quantitative frameworks. The most recent national time use surveys have been conducted with these applications in mind.

**THE RELEVANCE OF TIME-BUDGET DATA SETS FOR SOCIAL POLICY**

Sociologists have often been skeptical about the utility of time-budget research. Time budgets are thought to provide data that are “broad but shallow” (Converse 1972, p. 46) and offer no more than static, tendentially commonsense descriptions of only the manifest aspects of everyday life. From a theoretical point of view, it is argued that there is a lack of explanatory hypotheses and relevant concepts that could bring norms, experiences, attitudes, and values to the fore. On the methodological side, the main criticism is that classification schemes of activities are imprecise, are unevenly general or detailed, and have barely changed since the 1920s (Pronovost 1989, pp. 78-80). Implicit in these observations is the dilemma Szalai faced earlier: whether priority should be assigned to the

testing of hypotheses or to multipurpose database construction. With the latter option comes the question of how to reconcile, in designing the studies, cross-national and longitudinal comparability and adherence to sociocultural settings and historically changing conditions.

Time-budget research is applied research that has increasingly been aimed at the design and evaluation of social policy. Its relevance in this context derives from the fact that in modern affluent societies, citizens often value scarce time more than material or monetary resources; thus, time use rationalization and efficient time management in the personal, family, and public sphere have become matters of general concern. Contingencies curtailing time use evidently are distributed unequally in the social world. For instance, health checks in public institutions often involve waiting times that private medical care does not, and not owning a means of transportation makes long commuting times unavoidable. Hence, social policies and their provisions try to make circumstances or opportunities more equal for everyone. Part-time employment, flexible worktimes, and compressed workweeks have been introduced mostly for pressured working mothers with small chil-

**Distribution of Economic Time in Four Countries Between SNA and Non-SNA Activities, by Gender. Indices Showing Unequal Participation of Men and Women in Each Group of Activities**

	Men (M)	Women (F)	Men and Women (M+F)	Inequality Index $I=F/(M+F)$
<b>France (1985–1986)</b>				
SNA activities	4:00	2:10	6:10	35%
Non SNA activities	2:28	4:59	7:27	67%
<b>Total</b>	6:50	7:09	13:37	52%
<b>Germany (1991–1992)</b>				
SNA activities	4:28	2:12	6:40	33%
Non SNA activities	2:53	5:08	8:01	64%
<b>Total</b>	7:21	7:20	14:41	50%
<b>Great Britain (1985)</b>				
SNA activities	4:39	2:34	7:13	36%
Non SNA activities	2:12	4:19	6:41	65%
<b>Total</b>	6:51	6:53	13:54	50%
<b>United States (1985)</b>				
SNA activities	4:31	2:47	7:18	38%
Non-SNA activities	2:37	4:46	7:23	65%
<b>Total</b>	7:08	7:33	14:41	51%

Table 4

SOURCE: Adapted from L. Goldschmidt-Clermont and E. Pagnossinn-Aligisakis, *Measures of Unrecorded Economic Activities in Fourteen Countries*, Occasional Papers no. 20, 1995, New York: UNDP.

dren. Shop-opening hours and office schedules have been changed to permit effective coordination and a better reconciliation of tasks. The effectiveness these measures can be monitored with the help of time-budget procedures.

The complexity of time-budget data sets has often been insufficiently exploited. Initially this was due to limitations in handling enormous data sets, but increased technological resources and new multivariate statistical techniques have opened new frontiers. Almost exclusive attention has been given to average durations and frequencies of primary activities, but the study of configurations emerging from an association of these activities with other contemporaneous activities promises a better understanding of modern time regimes. Until now, not much research has been done on routinized rhythms or the strategic sequencing of activities. Also, the collaborative or conflicting interface of the various schedules of family members, the reconstruction of networks of participative personal contacts (a topic of great significance in regard to lonely children, the ill, and the elderly), and the relationship of the use of urban spaces to

time use (a problem studied by human geographers) are all interesting areas for future research because of the greater availability of important data sets.

Activity classifications will have to undergo critical study to better reflect changes in the activity patterns of everyday life. Statistical offices are presently reconceptualizing their taxonomies. Paid work might be broken down into its constituent parts and examined analytically, but other activities need redefinition. For example, some kinds of domestic work have been absorbed by the market; care activities now regard the elderly more than children; dealing with service bureaucracies has become a time-consuming task; there is a new spectrum of voluntary forms of participation at the social, political, and cultural levels; and leisure behavior has changed in relationship with new media.

However, the problem is not just a technical one regarding exclusively descriptive coding schemes. Time use data assume importance only when they provide a valid epistemological key for the interpretation of social change. As has already been

pointed out, earlier lines of research identified two central components of time use: market work and leisure. Today, time use studies based on data from household samples may help identify other valid criteria of time use to better understand how families cope with growing structural unemployment and increasing social insecurity.

## REFERENCES

- Bevans, George Esdras 1913 *How Workingmen Spend Their Spare Time*. New York: Columbia University Press.
- Bourdieu, Pierre 1963 "La Société Traditionnelle: Attitude à l'Égard du Temps et Conduite Économique." *Sociologie du Travail*, 5:24-44.
- Converse, Philip E. 1968 "Time Budgets" In David L. Sills, ed., *The International Encyclopedia of Social Sciences*, vol. 16. New York: Macmillan-Free Press.
- Elias, Norbert 1988 *Time*. New York: Basil Blackwell.
- Gershuny, Jonathan 1992 "La répartition du temps dans les sociétés post-industrielles." *Futuribles*, Mai-Juin, 215-226.
- , and John P. Robinson 1988 "Historical Changes in the Household Division of Labor." *Demography* 25(4):537-552.
- Goldschmidt-Clemont, Luisella, and Elisabetta Pagnossin-Aligisakis 1995 *Measures of Unrecorded Economic Activities in Fourteen Countries*, Occasional papers no. 20. New York: UNDP.
- Harvey, Andrew S., Alexander Szalai, David Elliott, Philip J. Stone, and Susan M. Clark 1984 *Time Budget Research: An ISSC Workbook in Comparative Analysis*. Frankfurt: Campus.
- 1993 "Guidelines for Time Use Data Collection." *Social Indicator Research* 30:197-228.
- INSEE 1989 "Les Emplois du Temps des Français." *Economie et Statistique* 223.
- Jahoda, Marie, Paul Lazarsfeld, and Hans Zeisel 1971 *Marienthal: The Sociology of an Unemployed Community*. Chicago: Aldine (German orig., 1933).
- Juster, F. Thomas 1985 "The Validity and Quality of Time Use Estimates Obtained from Recall Diaries." In F. T. Juster and F. P. Stafford, eds., *Time, Goods, and Well-Being*. Ann Arbor: Institute of Social Research, University of Michigan.
- Kurtz M.-F. D. 1984 "Les Budget-Temps: Réflexion Épistémologique," *L'année Sociologique* 34:9-27.
- NHK 1991 *Japanese Time Use in 1990*. Japanese Broadcasting Corporation, Public Opinion Research Division.
- Niemi, Iris 1993 "Systematic Error in Behavioural Measurement: Comparing Results from Interview and Time Budget Studies." *Social Indicators Research* 30:229-244.
- Przeworski, Adam, and Henri Teune 1970 *The Logic of Comparative Inquiry*. New York-London: Wiley Interscience.
- Pronovost, Gilles 1989 "The Sociology of Time." *Current Sociology* 37(3):1-129.
- Scheuch, Erwin K. 1992 "The Time-Budget Interview." In Alexander Szalai, ed., *The Use of Time*. The Hague-Paris: Mouton.
- Sorokin, P. A., and C. Q. Berger 1939 *Time-Budgets and Human Behavior*. Cambridge, Mass.: Harvard University Press.
- Statistisches Bundesamt 1995 *Die Zeitverwendung der Bevölkerung*, Tabellenband I. Wiesbaden.
- Stone Philip J. 1972 "The Analysis of Time-Budget Data." In Alexander Szalai, ed., *The Use of Time*. The Hague-Paris: Mouton.
- Sue, Roger 1994 *Temps et Ordre Social*. Paris: Presses Universitaires de France.
- Szalai, Alexander, ed. 1972 *The Use of Time: Daily Activities of Urban and Suburban Populations in Twelve Countries*. The Hague-Paris: Mouton.
- 1977 "The Organization and Execution of Cross-National Survey Research Projects." In Alexander Szalai and Riccardo Petrella, eds., *Cross-National Comparative Survey Research*. Oxford, UK: Pergamon.
- 1984 "The Concept of Time Budget Research." In Andrew S. Harvey et al., eds., *Time Budget Research*. Frankfurt: Campus.
- United Nations 1995 *The World's Women 1995: Trends and Statistics*. No. E.95. XVII.2.
- Zerubavel, Eviatar 1982 "The Standardization of Time. A Sociohistorical Perspective." *American Journal of Sociology* 88 (1):1-23.

ELKE KOCH-WESER AMMASSARI

## TOURISM

Tourism is an economic phenomenon with important sociocultural implications that acquired a fundamental significance in the last decades of the twentieth century. It is one of the economic sectors with the highest rates of growth, together with transportation, communications, and the computer industry, with which it works in a synergistic way.



According to a classic definition, tourism can be identified in the complex of relations and manifestations that rise from the travel and stay of foreigners when the stay is temporary and is not motivated by a lucrative occupation (Hunziker and Krapf 1942). "Foreigners" are persons who do not reside habitually in the zone in which the tourist activity is carried out; depending on whether the zone of residence is in the same state, one can distinguish between internal tourism and international tourism.

Other elements have to be considered in distinguishing fully-fledged tourism from similar activities. First, two "fundamental actors" have to be dealt with. On one side, there are tourists (*active tourism*), who decide to undertake this activity because of several motivations. This is one of the primary topics in the sociopsychological analysis of tourism. On the other side, there is *passive* (or receptive) *tourism* constituted by the technical and socioeconomic structures that exist in the zones of reception with the aim of hosting tourists. In modern tourism, a third actor, consisting of agents of tourist intermediation (travel agencies, tour operators, carriers, etc.), has assumed greater importance by connecting the demand for and the supply of tourism. The tourism described here is essentially a mass phenomenon that exists alongside elite tourism, which was the first type to appear.

Merchant writers such as Marco Polo, traveler-explorers, and missionaries, often accompanied by anthropologists and ethnologists, were the forerunners of tourists, but only after the "Grand Tours" of the eighteenth and nineteenth centuries can one speak of the emergence of the tourist phenomenon. The grand tour was considered a fundamental stage in the formation of young aristocrats and later of the children of the emergent high bourgeoisie. It consisted of a visit to the more centers of the culture of the age, with a predilection for southern Europe, in particular Italy and its remnants of classic culture.

Toward the end of nineteenth century, with the inauguration of the first seaside resorts, tourism began to acquire mass characteristics, a phenomenon that was facilitated by the improvement of transportation systems, especially the extension of the railway network (Urry 1990).

The transformation of elite tourism into a phenomenon that involved wide strata of the popu-

lation did not occur until after the end of World War II and, in the more economically developed countries, the possession of the automobile as an individual and family means of transportation and the expansion of transcontinental and transoceanic flights.

Tourism is therefore facilitated, in addition to the elevation of individual incomes and better tariff conditions, by technical, political, and social factors. It also involves a psychological evolution in society, especially in the richer countries, where it provides an escape from the stresses of city life and the daily routine. A stronger desire for social intercourse has grown along with a desire for physical activity to compensate for a sedentary lifestyle.

Certain forms of travel have a demonstrative scope, since "to tour," in particular elite tourism but also mass tourism, may be thought of as an expression of one's prestige and social position.

Urry (1990) introduced the concept of the "tourist gaze," stating that "part at least of that experience is to gaze upon or view a set of different scenes, of landscapes or townscapes which are outside of the ordinary" (1990, p. 1). As Urry describes it:

1. "tourism is a leisure activity which presupposes its opposite, namely regulated and organized work, . . .
2. tourist relationships arise from a movement of people to, and their stay in, various destinations. This necessarily involves some movement through the space, that is the journey, and a period of stay in a new place or places,
3. the journey and stay are to, and in, sites which are outside the normal places of residence and work. Periods of residence elsewhere are of a short-term and temporary nature, . . .
4. the places gazed upon are for purposes which are not directly connected with the paid work and normally they offer some distinctive contrasts with work (both paid or unpaid);
5. a substantial proportion of the population of modern societies engages in such tourist practices, new socialized forms of

provision are developed in order to cope with the mass character of the gaze of tourist (as opposed to the individual character of “travel”)

6. places are chosen to be gazed upon because there is an anticipation, especially through daydreaming and fantasy, of intense pleasures, either on a different scale or involving different senses from those customarily encountered, . . .
7. the tourist gaze is directed to features of landscape and townscape which separate them off from everyday experience. Such aspects are viewed because they are taken to be in some sense out of the ordinary, . . .
8. the gaze is constructed through signs, and tourism involves the collection of signs, . . .
9. an array of tourist professionals develop who attempt to reproduce ever-new objects of the tourist gaze. . . (Urry 1990, pp. 2-3 passim).

**THE DIMENSIONS OF THE PHENOMENON**

The modern growth of tourism (Prosser 1994, p. 19) has been referred to ironically by Lodge (1992) as the “new global religion,” and a work on this phenomenon is entitled “*The Golden Hordes*” (Turner and Ash 1975). Prosser also supplies some data on a “phenomenon” that he describes by using the metaphor of the tsunami: By the mid-1990s, the tourism sector constituted 6 percent of world gross national product and 13 percent of the money spent for consumption and could be defined as the fastest-growing industry. According to the forecasts of the World Tourism Organization, in the year 2005, the tourist industry will involve 40 million persons.

Taking into account only persons who cross their state borders for tourism (perhaps equally important is the internal tourist movement), more recent data show, in approximately a decade, a near doubling of the phenomenon and also indicate that total revenues have increased to three times their original amount (Table 1).

The distribution of tourism in a wide range of countries has occurred in time span of only a few years. Tourism in the industrialized countries, while

**Arrivals and Revenues in International Tourism**

Year	Arrivals (\$ millions)	Revenue (\$ billions)
1986	339	142
1987	362	175
1988	395	203
1989	427	219
1990	458	266
1991	464	273
1992	503	311
1993	518	318
1994	547	348
1995	566	393
1996	592	423

Table 1

SOURCE: World Tourism Organization.

declining in relative terms, accounts for ever 50 percent of the total, but the share of the developing countries has grown and now accounts for almost a third of the total. The countries of central and eastern Europe still suffer from the backwardness of decades of relative inaccessibility, but after the fall of the Berlin wall, their proportion of the world tourism has grown (Table 2).

France is the leader in tourist presence, followed by the United States and two countries of Mediterranean Europe, Spain and Italy. In fifth place is China, a country only recently opened to international tourism that has a large potential that has been limited by a deficiency of infrastructure and receptive structures.

The forecasts of an increase of tourism are plausible because some countries of the former communist bloc already play an important role. Other developing countries (e.g., Brazil and South Africa), apart from China, have potential and will be able to play a more important role in international tourism if they stabilize their political and/or economic situation.

There currently is a remarkable concentration of tourist destinations in which the top ten countries together account for over 50 percent of tourism and top twenty countries account for over two-thirds (Table 3).

These figures suggest that the market for tourism will grow and become more differentiated, that there will be more specialization and segmentation of that market, and that organized travel

**Percentage Distribution of Tourist Presences**

	1990 (%)	1996 (%)
Industrialized countries	61.5	55.8
Developing countries	28.3	31.0
Central and eastern Europe	10.2	13.2

Table 2

SOURCE: World Tourism Organization.

packages will become more personalized to cope with the desire for greater individual freedom through a modular design of the product (Schwaninger 1989).

### THE NATURE OF TOURISM

Mass tourism is the main concern of this article inasmuch as the current growth of the tourism industry essentially has resulted from it. Mass tourism is a “fickle” market in which status-elevating motivations are important. “If people do not travel, they lose status: travel is the maker of status” (Urry 1990, p. 5). The concept of conspicuous consumption (Veblen [1899] 1970) is operative here because in choosing a vacation, one takes into account the attributions of status defined on the basis of the place one visits and the characteristics of the other visitors.

One therefore is dealing with a market that is very sensitive to fashion and changes in values. The relative loss of importance of seaside resorts, which were the preferred destinations at the beginning of mass tourism, can be cited in this regard:

*In the post-war period it has been the sun, not the sea, that is presumed to produce health and sexual attractiveness. The ideal body has come to be viewed as one that is tanned. This viewpoint has been diffused downwards through the social classes with the result that many package holidays present this as almost the reason for going on holiday. . . . Seaside resorts have also become less distinctive because of the widespread de-industrialization of many towns and cities so that there is less need to escape from them to the contrasting seaside. As the everyday has changed, as towns and cities have become de-industrialized and many have themselves become objects for the tourist gaze, with wave machines and other features of the*

*beach, so seaside resorts are no longer extraordinary. (Urry 1990, pp. 37–38)*

Tourism is therefore a fashion phenomenon that goes through all the typical phases of a product of that type, from discovery and emergence, to increasing popularity, saturation, attenuation of its appeal, and eventually decline. It is sensitive to the relationship between demand and supply, based on perceptions, expectations, attitudes, and values of people, and therefore is subject to cultural filters:

*The various contents and destinations of tourism, from the nineteenth Century to our days, seem to follow a standardized route . . . They are invented by individuals that live in conditions of originality and marginality in relationship to the ‘world.’ Subsequently they are consecrated by the notables: the monarchs and their families, followed by the artists and the celebrities . . . Finally they are diffused through the capillary imitation of the behavior of one social layer by the immediately inferior one. As soon as a place or a tourist fashion is known, there begins an emulation process that leads quickly to congestion; processes of distinction are then activated by groups that address to other places and invent other activities, reopening a new cycle. The succession of dissemination and invention cycles leads to the need for distinction to introduce more and more far and unusual goals. (Savelli 1998, pp. 92–93).*

One can speak of the “pleasure periphery,” as in the case of the increase of Antarctic tourism (Prosser 1994, p. 22). For this aspect, the model of Plog (1973) is relevant. Plog analyzes the personality of the tourist: Along a continuum, one can go from psycho-centered, expectant subjects preoccupied with the small daily problems and escaping to adventures, to subjects who are as allocentered, confident in themselves, curious, and adventurous. The places visited by these varied subjects are obviously very different. In the survey conducted by Plog among the inhabitants of New York, while the psychocentered subjects do not venture beyond Coney Island, the midcentered travel to Europe and the allocentered do not dare to face the Pacific or Africa.

In dealing with tourism from a socioeconomic point of view, the “positional goods” concept

**International Arrivals by Country, 1996**

Country	Thousands of tourists	Country	Thousands of tourists
France	61,500	Austria	16,641
United States	44,791	Germany	15,070
Spain	41,295	Hong Kong	11,700
Italy	35,500	Switzerland	11,097
China	26,055	Portugal	9,900
Great Britain	25,800	Greece	9,725
Mexico	21,732	Russian Federation	9,678
Hungary	20,670	Turkey	7,935
Poland	19,420	Malaysia	7,742
Canada	17,345	Total	430,801
Czech Republic	17,205	World total	591,864

Table 3

SOURCE: World Tourism Organization.

(Hirsh 1978) can be used. This term refers to social goods, services, jobs, positions, and other relations that are scarce or subject to congestion and/or crowding. The competition is zero-sum: When someone consumes these kinds of goods in excess, someone else is forced to consume less. The supply is limited because quality would lessen as a result of quantitative growth.

One also can trace a conflict of interest between the actors described in the first part of this article (tourists, agencies, and the tourism industry in the hosting countries) and environmentalists. Since natural and cultural resources may be irremediably spoiled, there is thus a conflict of interest between present and future generations (Mishan 1969).

Another peculiar characteristic of tourism is that “almost all the services provided to tourists have to be delivered at the time and place at which they are produced. As a consequence the quality of the social interaction between the provider of the service, such as the waiter, flight attendant or hotel receptionist, and the consumer, is part of the ‘product’ being purchased by the tourist. If aspects of that social interaction are unsatisfactory (the offhand waiter, the unsmiling flight attendant, or the rude receptionist), then what is purchased is in effect a different service product” (Urry 1990, p. 40). Production of services for the consumer, in fact, cannot be done entirely behind the scenes, far away from the tourist gaze. Moreover, tourists have high expectations about what they will receive, since the search for the extraordinary is an essential aspect of the choice to travel.

“Spatial fixity” is a crucial characteristic of tourist services (Bagguley 1987), and customers are more mobile and now consume tourist services on a global scale. This means that “part of what is consumed is in effect the place in which the service producer is located. If the particular place does not convey appropriate cultural meanings, the quality of the specific service may well be tarnished” (Urry 1990, p. 40).

Since the services offered are intrinsically labor-intensive, employers try to diminish the costs. However, this may undermine the extraordinary character of the tourist experience (Urry 1990, p. 41).

**TOURISTS AND THEIR MOTIVATIONS**

In an attempt to grasp the features that distinguish tourists from other kinds of travelers, Cohen (1974) singles out certain dimensions that are thought to be essential: duration of the travel, voluntariness, direction, distance, recurrence, and purpose. On the basis of these elements, a tourist may be defined as a traveler who moves voluntarily and for a limited period of time to obtain pleasure from the experience of novelty and change, following a relatively long and non-recurring route.

For the sake of clarity, distinctions are introduced in the form of a dichotomy. However, one can assume that in many cases there are different degrees of distance from “full-fledged tourism.”

When the *duration* of the travel and stay is short (less than twenty-four hours in the definition of the UN Conference on International Travel and

Tourism), there are trips and excursions. There is also an upper limit, more difficult to determine, beyond which one can speak of permanent travelers (wanderers, nomads).

When the element of *voluntariness* is lacking, one is dealing with the exile (sometimes voluntary), the slave, the prisoner of war, or the political refugee. The pilgrim also can be considered a type of traveler who differs from the full-fledged tourist inasmuch as in many cases there is a lack of voluntariness. This is the case because social expectations can determine the decision to travel and the stay (e.g., pilgrimages to Mecca by Muslim believers).

In terms of *direction*, tourists return to their countries of origin, while immigrants make a one-way trip. There are also intermediate categories that are less easy to classify, such as "tourist immigrants" and "permanent tourists." These people leave home as tourists but decide to stay for a longer time span in a foreign country. Persons such as the "expatriates" (e.g., the many foreign artists who reside in cities such as Paris) are also difficult to define. They decide to live in a foreign country for indefinite periods without completely cutting their ties with the country of origin.

If the *distance* is short, one can speak of excursionists and hikers, while if the distance is much longer, one could have spoken in the past of explorers. Today, nearly all the possible destinations on the face of the earth seem to be within the reach of the tourist. If the distance implies crossing a national border, there is the already mentioned distinction between internal tourism and international tourism.

When travel and stay have a season or weekend regularity (*recurrence*), one is dealing with the *habitué*, who often is the owner of a summer house. This person is not properly a tourist, because the elements of novelty and change are lacking.

Finally, the *purpose* for the tourist does not have to be instrumental but can involve the seeking of pleasure. If the purpose is instrumental or has another specific nature different from the search for novelty and change, one is dealing with students, old country visitors, conventioners, business travelers, tourist employees, and the like.

However, this criterion is not as precise as it might appear at a first glance. The noninstrumental character of the purpose and the search for novelty and change has to be considered from a social point of view. When an individual takes a vacation for reasons of prestige, this travel is socially defined as a pleasure trip even if that individual will not enjoy the experience. More likely, there will be the opposite case: The purpose is declared as instrumental, but other instrumental (and not) purposes are also relevant (Savelli 1998, p. 57).

Tourists' motivations also can be analyzed by distinguishing the *push* factors that lead to the desire to go on vacation from the *pull* factors that the various areas of attraction exercise on the tourist (Savelli 1986, p. 2269).

To show the "versatility" of the tourism phenomenon, a relationship can be seen between some of its forms and the fundamental needs listed by Maslow. Therapeutic tourism satisfies physiological needs, while the needs of security and belonging are satisfied by familiar and "identity" tourism. The need for social recognition is catered to by tourism *à raconter*, (The French expression *à raconter* refers to a tourist who leads you to extraordinary places where extraordinary things happen that one is very pleased to narrate to friends, thereby obtaining social status.) and people satisfy the need for self-esteem through sport and cultural tourism (Kovacshazy and people 1998, p. 58).

To describe the psychological and social situation experienced by the tourist, some authors propose an interesting analogy between the tourist and the pilgrim. Both move from a familiar place to a distant one and then come back. In faraway localities, they dedicate themselves—although in different ways—to the "worship" of sacred places. These can be described as "liminoid" situations in which daily obligations are suspended (Turner and Turner 1978): "There is license for permissive and playful 'non-serious' behavior and the encouragement of a relatively unconstrained 'communitas' or social togetherness" (Urry 1990, p. 10). The purpose of a vacation thus consists of overturning the daily routine: Middle-class tourists try to be a "peasant for a day," while tourists with a lower social rank try to be "king/queen for a day" (Gottlieb 1982).

In a survey carried out in Italy (Isnart 1997) by interviewing only persons who go on vacation habitually, only the expenses for food and daily living were judged "more necessary" than those for traveling. The expenses for car use and maintenance and those undertaken to dress were lower than those for the consumption of vacations.

There often exists a link among subjective motivations, perception of the visited localities, and the objective connotations of those localities. Some connotations are always valid (effectiveness and efficiency, a proper quality-price ratio, a satisfactory environmental quality, the hospitality and warmth of the residents). Other connotations assume a nearly cyclical course: They gain a special reputation for one or two seasons and then fade out.

However, five major categories of motivations more or less summarize what this article has described so far:

1. *Subjectivity*: the sense of curiosity, interest, discovery, opportunity, and "digression" of the vacation
2. *Security*: the sense of confidence that vacation places must transmit and the possibility of relaxing (nearly the opposite of the insecurity of large cities)
3. *Transgression*: the willingness to have a good time, to push the limits, to have "extraordinary" and "sensual" experiences
4. *Budget*: the search for something that does not divert too many resources from other needs and opportunities
5. *Status*: the idea that travel is first of all social gratification, something to show, a reached goal (Isnart 1997, p. 16)

Among these categories of motivations, subjectivity prevails, with status and transgression not far behind. Obviously, budget is much more a concern of the elderly (who also appreciate security) and young people (who do not care much about status). Some of these differences are related to socioeconomic class.

### THE IMPACT OF TOURISM

The tourist's role is a total one: "He cannot hide his own externality from the local population and all his relations are imprinted and denoted, in the

first place, by the tourist role. In the same way, he is recognized as such from other tourists, regardless, in some manner, of his social condition, nationality, origin and race" (Savelli 1998, pp. 129-130).

The tourist's presence therefore cannot pass unnoticed, and the increase of tourism can carry, besides the obvious economic advantages, some negative consequence in the countries that receive tourist flows. In this regard, there are pessimistic visions that are valid, especially for developing countries. These are the countries in which tourism can be expected to show steadily increasing rates of growth and in which there is more to earn from this development.

Tourist destinations are vulnerable, and one can even speak about economic colonialism, because investments and the largest part of demand are controlled by the developed countries. Exploitation can be not only economic but also social and environmental, inasmuch as community displacement, societal dislocation, and cultural transformation may occur (Ryan 1991): "Village farmland is appropriated, there is inter-generational stress as younger groups succumb to the 'demonstration effect' of tourist material wealth and behavior, intra-family stress as male-female role balance shifts, and community disharmony as religious ceremonies and artforms are commercialized" (Prosser 1994, p. 29).

Therefore, it is necessary to foster a sustainable tourism that tries "to sustain the quantity, quality, and productivity both of human and natural resources systems over time, while respecting and accommodating the dynamics of such systems" (Prosser 1994, pp. 31-32). This alternative form of tourism must "search for spontaneity, enhanced interpersonal relations, creativity, authenticity, solidarity, and social and ecological harmony" (Pearce 1989, p. 101).

The social relations between tourists and indigenous populations are complex and can lead to conflict as a result of several factors. Among the more important ones are the number of tourists who visit a place in relation to the size of the hosting population, the type of organization of the tourist industry, the effects of tourism on preexisting agricultural and industrial activities, economic and social differences between the visitors and the majority of the hosts, and the degree to which

visitors demand particular standards of lodging and service, that is, the expressed desire to be locked in an “environmental bubble” for protection from the “disappointing” characteristics of the hosting society (Urry 1990, p. 90).

As a counterbalance of these potential dangers, one has to consider that the cost of a new workplace in the tourist sector has been estimated at £4,000, compared with £32,000 in the manufacturing industry and £300,000 in mechanical engineering (Lumley 1988, cited by Urry 1990, p. 114). These are older figures, and therefore are not necessarily still valid, but the ratios probably continued to be valid. The “tourist prescription” therefore can be recommended particularly for countries that do not have many financial resources.

For tourism to be sustainable and respectful of the natural and social environment, the attitudes and behaviors of the three main actors must change:

- The attitudes of *tourists* must change. Tourists tend to believe that other tourists are the problem. Thus, their attitudes remain elitist and short-term.
- The *destination areas* must assume a longer-term attitude. An equilibrium between optimization of the revenues and protection of the resources must be found. Populations must be involved in all phases of development: ideation and planning, construction and implementation, conduction and management, and monitoring and modification.
- The *tourist industry* must find an equilibrium between opposing requirements. There is an unavoidable push for environmental control from foreign investors and operators in order to obtain greater profits that can be detrimental to local populations and governments. At the same time, the tourist industry feels the need to appear to be ecologically responsible (Prosser 1994, p. 32).

It has been proposed that tourism should be considered only a preliminary stage in which resources are obtained, that can be used later for “true” development through investment in other sectors. That is reasonable, because diversification is a key factor in economic security and stability, especially if tourism can be defined as a fashion

industry. However, one may question whether the impact of other industrial initiatives is less harmful and more sustainable than that of tourism. This opinion results from a dated attitude characterized by an ideologically rooted prejudice that is disappearing: “In the last few years in Britain many Labour councils have enthusiastically embraced local tourist initiatives, having once dismissed tourism as providing only ‘candy-floss jobs’” (Urry 1990, p. 115).

### POSTMODERN TOURISM

While the countries that receive tourist flows need to find a balance between the advantages and disadvantages and search for a sustainable “receipt,” the benefits for tourists seem to be without shortcomings. Krippendorf (1987) speaks about “travel” that represents recuperation and regeneration, compensation and social integration, escape, and communication, intellectual expansion, freedom and self-determination, self-realization, and happiness.

The fact that the tourist industry continues to grow indicates that it is able to give a satisfactory answer to tourists’ expectations; otherwise there would be frustration, and the phenomenon would recede. One can ask why tourists continue to travel and their numbers continue to increase in spite of the “alarm bells” that call attention to the problem of overcrowding and the relative nonauthenticity of the tourist experience.

This article has dealt with the problem of overcrowding in its characterization of the tourist product as a “positional good.” This pessimistic thesis has been criticized by Beckerman (1974), who raises two interesting issues. First, the concern about the effects of the mass tourism is basically a “middle-class” anxiety (like many other environmental concerns) because the really rich “are quite safe from the masses in the very expensive resorts, or on their private yachts or private islands or secluded estates” (Beckerman 1974, pp. 50–51). Second, most people who are affected by mass tourism benefit from it, including the “pioneers,” who, when they return to a place, find services that were not available when the number of visitors was small.

One also can criticize the applicability of the scarcity concept to the tourist industry. The im-



PLICIT scarcities in the tourist industry are complex, and strategies can be adopted that allow the enjoyment of the same object by a greater number of persons. Thus, one must distinguish between the “physical capacity” and “perceptive capacity” of a tourist place (Walter 1982).

One also has to consider that in addition to the “romantic” tourist gaze, which emphasizes solitude, privacy, and a personal, quasi-spiritual relation with the observed object, there is an alternative “collective” gaze with different characteristics. The collective gaze demands the participation of wide numbers of other people to create a particular atmosphere: “They indicate that this is *the* place to be and that one should not be elsewhere.” (Urry 1990, p. 46). This is the case for major cities, whose uniqueness lies in their cosmopolitan character: “It is the presence of people from all over the world (tourists in other words) that gives capital cities their distinct excitement and glamour” (Urry 1990, pp. 46).

Some people prefer to move around in compact formations because otherwise they will not enjoy themselves, while others prefer to travel in solitude. Therefore, Hirsh’s (1978) thesis on scarcity and positional competition should be applied mainly to tourism characterized by the romantic gaze. When the collective gaze is more important, the problem of crowding and congestion is less marked. Moreover, the scarcity thesis would be totally applicable only if one maintained that there are severe limits to the number of “objects” worthy of the admiration of the tourist. However, “if Glasgow can be remade as a tourist attraction, one might wonder whether there are in fact any limits to the tourist, or post-tourist, gaze” (Urry 1990, p. 156).

Another issue refers to the nonauthenticity of the tourist experience. Turner and Ash (1975) describe a tourist who is placed at the center of a rigorously circumscribed world (the “environmental bubble”). Travel agents, couriers, and hotel managers are described as surrogate parents who relieve the tourist of every responsibility, protect the tourist from harsh reality, and decide for the tourist which objects are worthy to be admired.

Various types of tourists exist, and they are pushed by various needs and motivations for which various means are available to realize the tourist experience. In an age that is being defined as postmodern, the posttourist also is being redefined.

*The post-tourist knows that they are a tourist and that tourism is a game, or rather a whole series of games with multiple texts and no single authentic tourist experience. The post-tourist thus knows that they will have to queue time and time again, that there will be hassles over foreign exchange, that the glossy brochure is a piece of pop culture, that the apparently authentic local entertainment is as socially contrived as an ethnic bar, and that the supposedly quaint and traditional fishing village could not survive without the income from tourism. (Urry 1990, p. 100).*

The post-tourist knows that “he is not a time-traveller when he goes somewhere historic, not an instant noble savage when he stays on a tropical beach, not an invisible observer when he visits a native compound. Resolutely ‘realistic,’ he cannot evade his condition of outsider” (Feifer 1985, p. 271). This means that many travelers appreciate the “not-authenticity” of the tourist experience and “find pleasure in the multiplicity of tourist games. They know that there is *no* authentic tourist experience, that there are merely a series of games or texts that can be played” (Urry 1990, p. 11).

## REFERENCES

- Beckerman, W. 1974 *In Defense of Economic Growth*. London: Jonathan Cape.
- Bugguley, P. 1987 *Flexibility, Restructuring and Gender: Changing Employment in Britain’s Hotels*. Lancaster: Lancaster Regionalism Group, Working Paper no. 24.
- Cohen, E. 1974 “Who Is a Tourist? A Conceptual Clarification.” *Sociological Review* 4:527–556.
- Feifer, M. 1985, *Going Places*. London: Macmillan.
- Gottlieb, A. 1982 “Americans’ Vacations.” *Annals of Tourism Research* 9:165–187.
- Hirsh, F. 1978 *Social Limits to Growth*. London: Routledge & Kegan Paul.
- Hunziker, W. and K. Krapf 1942 *Grundriss der allgemeine Fremdenverkehrslehre*. Zurich.
- Istituto Nazionale Ricerche Turistiche (National Institute for Tourism Research) 1997 *1997 Dove vanno in vacanza gli italiani*. Milan: Unioncamere.
- Kovacshazy, M. C. 1998 “Le tourisme des seniors en 2010.” *Futuribles* 233:47–64.
- Krippendorf, J. 1987 *The Holiday Makers*. London: Heinemann.
- Lodge, D. 1992 *Paradise News*. London: Penguin.



- Lumley, R. (ed.) 1988 *The Museum Time-Machine*. London: Routledge.
- Mishan, E. 1969 *The Costs of Economic Growth*. Harmondsworth, UK: Penguin.
- Pearce, D. 1989 *Tourist Development*. Harlow, UK: Longman.
- Plog, S. V. 1973 "Why Destination Areas Rise and Fall in Popularity." *Cornell Hotel and Restaurant Administration Quarterly*, November, pp. 13–16.
- Prosser, R. 1994 "Societal Change and the Growth in Alternative Tourism." In E. Carter and G. Lowman, eds., *Ecotourism: A Sustainable Option?* New York: Wiley.
- Ryan, C. 1991 *Recreational Tourism*. London: Routledge.
- Savelli, A. 1986 "Turismo." In F. Demarchi, A. Ellena, and B. Cattarinussi, eds., *Nuovo dizionario di sociologia*. Rome: Paoline.
- (1998), *Sociologia del turismo*. Milan: Angeli.
- Schwaninger, M. 1989 "Trends in Leisure and Tourism for 2000–2010." In S. F. Witt and L. Moutinho, eds., *Tourism Marketing and Management Handbook*. Hemel Hempstead, UK: Prentice-Hall.
- Turner, L., and J. Ash 1975 *The Golden Hordes*. London: Constable.
- Turner, V., and E. Turner 1978 *Image and Pilgrimage in Christian Culture*. New York: Columbia University Press.
- Urry, J. 1990 *The Tourist Gaze: Leisure and Travel in Contemporary Society*. Thousand Oaks, Calif.: Sage.
- Veblen, T. [1899] 1970 *The Theory of the Leisure Class*. London: Allen & Unwin.
- Walter, J. 1982 "Social Limits to Tourism." *Leisure Studies* 1:295–304.

GIOVANNI DELLI ZOTTI

## TRANSNATIONAL CORPORATIONS

A transnational corporation (TNC) is "any enterprise that undertakes foreign direct investment, owns or controls income-gathering assets in more than one country, produces goods or services outside its country of origin, or engages in international production" (Biersteker 1978, p. xii). Various terms multinational corporations (MNCs) and multinational enterprises (MNEs), transnational corporations are formal business organizations that have spatially dispersed operations in at least two countries. One of the most "transnational" major TNCs is Nestlé, the Swiss food giant; 91

percent of its total assets, 98 percent of its sales, and 97 percent of its workforce are foreign-based (UNCTAD 1998, p. 36).

### TNCS AND THE GLOBAL ECONOMY

Although TNCs existed before the twentieth century (colonial trading companies such as the East India Company, the Hudson's Bay Company, and the Virginia Company of London were precursors of the modern TNC), only since the 1960s have they become a major force on the world scene (World Bank 1987, p. 45). Table 1 corroborates this by listing the foreign direct investment (FDI) stock of corporations by country from the beginning of the century to 1997. In 1900, only European corporations were major transnational players, but by 1930, American TNCs had begun to make their presence felt. The year 1960 marks the beginning of a new era in corporate transnationalization. In each of the decades from 1960 to the present, world FDI stock has more than tripled, whereas it only doubled during the first half of the century.

The phenomenal increase in transnational corporate activity in the latter part of the twentieth century can be accounted for in large part by technological innovations in transportation, communication, and information processing that have permitted corporations to establish profitable worldwide operations while maintaining effective and timely organizational control. The actual difference in foreign direct investment up to and after 1960 is even greater than the figures in Table 1 indicate. FDI for 1960 and before includes foreign *portfolio* investment, which is undertaken mainly by individuals, as well as foreign *direct* investment, which almost always is made by TNCs. These two types of investment were not reported separately for most countries before 1970. Thus, total FDI stocks are inflated. For example, Wilkins (1974, pp. 53–54) reports that in 1929–1930, U.S. foreign portfolio and direct investments were almost equal. American direct investment abroad was only \$7.5 billion; the remaining \$7.2 billion recorded in Table 1 was foreign portfolio investment.

Table 1 reveals that TNCs from only eleven countries accounted for almost 85 percent of all FDI in 1997. American TNCs accounted for more than one-quarter of total foreign investment, and

TRANSNATIONAL CORPORATIONS

**FDI Outward Investment Stock by Country, 1900–1997 (billions of US\$)**

Country	1900*	1930*	1960*	1971	1980	1990	1997†
United States	0.5	14.7	31.8	82.8	220.2	435.2	907.5
United Kingdom	12.1	18.2	13.2	23.1	80.4	229.3	413.2
Germany	4.8	1.1	0.6	7.0	43.1	151.6	326.0
Japan	Negligible	Negligible	Negligible	4.3	19.6	201.4	284.6
France	5.2	3.5	2.2	9.2	23.6	110.1	226.8
Netherlands	1.1	2.3	1.7	3.5	42.1	109.0	213.2
Switzerland	Negligible	Negligible	Negligible	6.5	21.5	65.7	156.7
Canada	Negligible	1.3	3.0	5.7	22.8	84.8	137.7
Italy	Negligible	Negligible	Negligible	NA	7.3	56.1	125.1
Belgium and Luxembourg	Negligible	Negligible	Negligible	NA	6.0	40.6	96.4
Sweden	Negligible	0.5	0.5	3.3	5.6	49.5	74.8
Others	Negligible	Negligible	Negligible	13.5	32.4	171.2	579.4
Total‡	23.8	41.6	53.8	159.2	524.6	1,704.5	3,541.4

*Table 1*

SOURCE: Data for 1900–1971 adapted from Buckley (1985), p. 200. Data for 1980–1997 from UNCTAD (1998), pp. 379–384.

NOTE: \*Includes foreign portfolio investment as well as foreign direct investment.

†Estimates.

‡World total, excluding former Comecon countries, except for 1997.

corporations based in the Triad (United States, European Union, and Japan) were responsible for nearly four-fifths of world FDI stock (UNCTAD 1998, pp. 379–384). Clearly, TNCs largely operate out of and invest in the developed countries of the global economy.

The magnitude of FDI flow in the world is revealed by the fact that worldwide sales of foreign affiliates in 1997 totaled \$9.5 trillion, almost one and a half times more than world exports of goods and services of \$6.4 trillion (UNCTAD 1998, p. 5). Global sales of affiliates are considerably more important than exports in delivering goods and services to markets worldwide, underlining the importance of TNCs in structuring international economic relations. In 1997, 53,607 TNCs controlled nearly 450,000 foreign affiliates throughout the world (UNCTAD 1998, p. 4).

Table 2 presents the top 30 TNCs ranked by foreign assets. Although fewer than one-quarter of these corporations are American in origin, most names are well known in the United States. It is the

nature of transnational enterprise to generate this degree of familiarity. Among the top 100 TNCs in terms of foreign assets, 41 originate in the European Union, 28 in the United States, and 18 in Japan (UNCTAD 1998, p. 317). Most FDI inflows and outflows take place within the Triad. In 1996, approximately one-quarter of all foreign sales was accounted for by these top 100 firms. Among the major industries in which these TNCs operate, electronics and electrical equipment account for the largest number (17), followed by chemicals and pharmaceuticals (16), automotive (14), petroleum and mining (14), and food and beverages (12). In 1996, these transnational giants employed nearly 6 million foreign workers (UNCTAD 1998, pp. 35–43).

#### REASONS FOR BECOMING TRANSNATIONAL

The move toward integrated transnational investment can be seen as a logical and rational decision by business enterprises to adapt to their environ-

TRANSNATIONAL CORPORATIONS

**World's Leading Transnational Corporations by Foreign Assets, 1996 (billions of US\$)**

Corporation	Country	Industry	Foreign Assets	Total Assets
General Electric	United States	Electronics	82.8	272.4
Shell, Royal Dutch	United Kingdom/Netherlands	Petroleum	82.1	124.1
Ford Motors	United States	Automotive	79.1	258.0
Exxon	United States	Petroleum	55.6	95.5
General Motors	United States	Automotive	55.4	222.1
IBM	United States	Computers	41.4	81.1
Toyota	Japan	Automotive	39.2	113.4
Volkswagen	Germany	Automotive	—*	60.8
Mitsubishi	Japan	Diversified	—	77.9
Mobil	United States	Petroleum	31.3	46.4
Nestlé	Switzerland	Food	30.9	34.0
Asea Brown Boveri	Switzerland/Sweden	Electrical equipment	—	30.9
Elf Aquitaine	France	Petroleum	29.3	47.5
Bayer	Germany	Chemicals	29.1	32.0
Hoechst	Germany	Chemicals	28.0	35.5
Nissan	Japan	Automotive	27.0	58.1
FIAT	Italy	Automotive	26.9	70.6
Unilever	Neth/U.K.	Food	26.4	31.0
Daimler-Benz	Germany	Automotive	—	65.7
Philips Electronics	Netherlands	Electronics	24.5	31.7
Roche	Switzerland	Pharmaceuticals	24.5	29.5
Siemens	Germany	Electronics	24.4	56.3
Alcatel Alsthom Cie	France	Electronics	23.5	48.4
Sony	Japan	Electronics	23.5	45.8
Total	France	Petroleum	—	30.3
Novartis	Switzerland	Pharmaceuticals/ chemicals	21.4	43.4
British Petroleum	United Kingdom	Petroleum	20.7	31.8
Philip Morris	United States	Food/tobacco	20.6	54.9
ENI Group	Italy	Petroleum	—	59.5
Renault	France	Automotive	19.0	42.2

*Table 2*

SOURCE: UNCTAD (1998), p. 36.

NOTE: \*Data on foreign assets are suppressed to avoid disclosure or are not available. In case of nonavailability, they are estimated on the basis of the ration of foreign to total sales, the ratio of foreign to total employment, or similar ratios.

**Reasons for Corporations Becoming Transnational**

1. *Cost-Related Reasons*
  - a. To take advantage of differences in technological development, labor potential, productivity and mentality, capital market, and local taxes
  - b. Reduction of transport costs
  - c. Avoidance of high tariff barriers
  - d. To take advantage of local talents when establishing R&D overseas
2. *Sales Volume Reasons*
  - a. Foreign middlemen unable to meet financial demands of expanded marketing
  - b. For quicker adaptation to local market changes and better adaptation to local conditions
  - c. Following important customers abroad
  - d. Keeping up with competitors
  - e. Persuasion and coercion of foreign governments
  - f. To obtain a better international division of labor, larger production runs, and better utilization of available economies of scale
  - g. To avoid home country regulations, e.g., fiscal and antitrust legislation
3. *Reasons Related to Risk Factors*
  - a. To avoid exclusion from customers' and suppliers' markets, promoting forward and backward integration
  - b. To counter inflexibility and avoid country-specific recessions
  - c. To reduce risks of social and political disruption by establishing operations in a number of host countries

*Table 3*

SOURCE: Taylor and Thrift (1982), p. 21.

ment. Historically, there have been several distinct strategies: (1) expansion in the size of operations to achieve economies of scale, (2) horizontal integration, or the merging of similar firms to increase market share, (3) vertical integration, or the acquiring of firms that either supply raw materials (backward integration) or handle output (forward integration) to attain greater control, (4) spatial dispersion or regional relocation to expand markets, (5) product diversification to develop new markets, and (6) conglomeration or mergers with companies on the basis of their financial performance rather than what they produce (Chandler 1962, 1990; Fligstein 1990). Establishing an integrated TNC simply represents a new strategy in this evolutionary chain. Furthermore, depending on how a corporation is set up and with recent innovations in communications and information technology, a TNC can incorporate all these strategies so that the newly structured enterprise has far

greater control and a much less restricted market than it had previously.

Table 3 presents a list of reasons why it may be profitable for an organization to become transnational. First, direct costs for raw materials, labor, and transportation as well as indirect cost considerations such as tariff barriers and trade restrictions, local tax structures, and various government inducements obviously loom large in the decision to establish operations transnationally. Second, market factors may be equally important in that decision. Direct and easy access to local markets unfettered by foreign trade quotas and other legislative restraints can give TNCs an edge over their nontransnational competitors. Finally, the decision to become transnational may hinge on factors related to organizational control. Control over raw materials (backward integration) and markets (forward integration) and achieving sufficient regional and product diversification to withstand temporary economic downturns are other reasons for transnational relocation.

**TNCS, NATION-STATES, AND GLOBALIZATION**

Integrated TNCs traversing real-time electronic networks that span the global economy have produced a "borderless world" (Ohmae 1991). These technologically enhanced corporations also operate in the nonnationally controlled interstices of the planet (i.e., oceans, seabeds, airwaves, sky, and space), sometimes leaving toxic, life-threatening indicators of their presence. Existing in a sort of parallel world, they are responsible only to amorphous groups of shareholders. Gill and Law (1988, pp. 364–365) state that there is a "growing lack of congruence between the 'world economy,' with its tendencies to promote ever-greater levels of economic integration, and an 'international political system' comprised of many rival states." The rivalry between these two systems of world organization is revealed by the fact that 51 of the 100 largest economies in the world are TNCs (Karliner 1997).

The increasing domination of the world economy by TNCs directly challenges national sovereignty. Historically, the sovereignty and therefore the power of a nation-state lay in its ability to achieve compliance with whatever it commanded its territorially defined space. Borderlines physi-

cally defined what was territorially sovereign and what was not. If a state's sovereignty was challenged from outside its territory, it could resort to force to maintain control. However, as a result of various technological developments, the idea of a physically bounded and sealed state is now open to question. These developments underlie the transnational corporate threat to state sovereignty along the following three dimensions:

1. *Permeability of borders.* Borderlines between nation-states have been rendered permeable and porous in a number of innovative ways, erasing many of the traditional distinctions between "inside" and "outside." For example, what borders do electronic communications and atmospheric pollutants observe? Under whose borders do oil and gas reserves lie? Do space satellites invade territorial integrity? The new permeability of borders diminishes the capacity of nation-states to distinguish and determine what occurs "inside" their territory.
2. *Mobility across borders.* Developments in transportation, communication, and information technology not only have increased the rate of cross-border mobility among TNCs but also have increased the speed or velocity with which cross-border transactions take place. Concurrently measuring both the location and the velocity of TNC activity often produces "uncertain" results, generating "inderminacy" for a state.
3. *Border straddling.* To the extent that TNCs operate simultaneously in different sovereign jurisdictions, which jurisdiction has precedence over which corporate activities at what time? This complex issue blurs the legal boundaries between states. It also confuses the notion of "citizenship" and its attendant rights and responsibilities.

Through the use of these and other innovative strategies, TNCs have manipulated the concept of borders to their advantage. What exactly is the advantage that TNCs achieve through their cross-border flexibility? They gain between-border variability. The fact that different states have different laws and standards regarding all aspects of economic activity contributes to the power of TNCs that strategically play off one country's set of rules

against another's. For example, variations in national laws on tariffs, financing, competition, labor, environmental protection, consumer rights, taxation, and transfer of profits are all carefully weighed by TNCs in deciding where and how to conduct business. Together, these considerations form what has come to be known as "the policy environment" (UNCTAD 1993, pp. 173-175). In the international competition to attract foreign investment by creating a "favorable policy environment," between-border variability encourages a "race to the bottom" (Chamberlain 1982, p. 126), resulting in a continuing erosion of sovereignty. Whereas TNCs operate in a de facto borderless world created by technological ingenuity, de jure political and legal distinctions still mark the boundaries on a world map composed of nation-states. This represents the crux of the inherent conflict between TNCs and nation-states as they are currently structured.

Never before has there been a situation in which foreign organizations have been granted license almost as a matter of course to operate freely within the legally defined boundaries of a sovereign state. This, together with the fact that TNCs and nation-states are different organizational forms, established for different purposes, administered by different principles, and loyal to different constituencies, means that structural problems are bound to arise.

#### TNCS AND WORLD DEVELOPMENT

Although only 30 percent of FDI stock is in developing countries (UNCTAD 1998, p. 373), because of the immense power of many TNCs, great concern has arisen about the impact of TNCs on world development. Because the goals of transnational capitalist enterprise and indigenous national government are fundamentally different, many scholars have debated whether TNCs are an aid or a hindrance to world development. According to Biersteker (1978), the major points of contention in this debate are the degrees to which TNCs (1) are responsible for a net outflow of capital from developing countries, (2) displace indigenous production, (3) engage in technology transfer, (4) introduce capital-intensive, labor-displacing technologies, (5) encourage elite-oriented patterns of consumption, (6) produce divisiveness within local social structures owing to competing loyalties

to TNCs and nation-states, and (7) exacerbate unequal distributions of income.

In a study of many of these issues, Kentor (1998, p. 1025) analyzed a fifty-year data set consisting of seventy-five developing countries to determine whether the modernization thesis (i.e., FDI in developing countries promotes “economic growth by creating industries, transferring technology, and fostering a ‘modern’ perspective in the local population”) or dependency theory (i.e., FDI results in disarticulated economic growth, repatriation of profits, increased income inequality, and stagnation) better explains the long-term results of foreign direct investment. Kentor (p. 1042) summarizes his findings as follows:

*The results of this study confirm that peripheral countries with relatively high dependence on foreign capital exhibit slower economic growth than those less dependent peripheral countries. These findings have been replicated using different measures of foreign investment dependence, GDP data, countries, time periods, and statistical methods. This is a significant and persistent negative effect, lasting for decades. Further, a structure of dependency is created that perpetuates these effects. The consequences of these effects, as described in the literature, are pervasive: unemployment, overurbanization, income inequality, and social unrest, to name a few.*

Given current conditions, it would appear that overreliance on foreign investment by developing countries will widen the already huge global rift between rich and poor nations.

### TNCS AND REGULATION

In the late 1960s, the United Nations (UN) reached the opinion that “transnational corporations had come to play a central role in the world economy and that their role, with its transnational character, was not matched by a corresponding understanding or an international framework covering their activities” (UNCTC 1990, p. 3). In the 1970s, the UN produced a draft “Code of Conduct on Transnational Corporations.” However, twenty years later, after much political wrangling, UN delegates concluded in 1992 that “no consensus was possible on the draft Code,” and thus the process of trying to achieve some effective legal

reconciliation between the goals of TNCs and those of host governments was brought to “a formal end” (UNCTAD 1993, p. 33).

Currently, although several international voluntary guidelines monitor the activities of TNCs, generally they have not been very successful (Hedley 1999). As of 1997, 143 countries had legislation in effect that specifically governs foreign direct investment (UNCTAD 1998, p. 53). Although initially most of those laws were framed to control the entry and regulate the activities of TNCs, legislative changes increasingly have become *more* favorable to foreign investment. For example, from 1991 to 1997, of the 750 changes to foreign investment policy made by countries worldwide, 94 percent were in the direction of liberalization (UNCTAD 1998, 57). In 1997, in attempts to ease high debt loads and survive a worldwide economic downturn, seventy-six developed and developing countries introduced 135 legislative inducements along the following lines: more liberal operational conditions and frameworks (61), more incentives (41), more sectoral liberalization (17), more promotion (other than incentives) (8), more guarantees and protection (5), and more liberal entry conditions and procedures (3) (UNCTAD 1998, p. 57). In their competition to attract foreign investment by creating favorable policy environments, these countries are yielding ever more control to TNCs.

Given the increasing dominance of TNCs in the global economy, the reasons why corporations become transnational, the diminishing sovereignty of nation-states, and the long-term effects of FDI on world development, one may question whether the move toward liberalization is in the interests of the countries and people who are encouraging it. What is called for is nothing short of a revolution in world governance. To regulate *transnational* corporations, it is necessary to introduce *trans-* or *supranational* legislation. To maintain national sovereignty in a global economy, authority must be coordinated and shared across borders. Legislative harmonization, although entailing an initial loss of sovereignty for participating states, can restore their authority over TNCs operating within their jurisdictions. By these means, corporate accountability can be imposed according to the needs and wishes of civil society. Whether or when such legislative harmonization will occur is open to question. However, in the view of the U.S. Tariff

Commission, "It is beyond dispute that the spread of multinational business ranks with the development of the steam engine, electric power, and the automobile as one of the major events of economic history" (cited in Lall and Streeton 1977, p. 15).

REFERENCES

Biersteker, Thomas J. 1978 *Distortion of Development? Contending Perspectives on the Multinational Corporation*. Cambridge, Mass.: MIT Press.

Buckley, Peter J. 1985 "Testing Theories of the Multinational Enterprise." In Peter J. Buckley and Mark Casson, eds., *The Economic Theory of the Multinational Enterprise*. London: Macmillan.

Chamberlain, Neil W. 1982 *Social Strategy and Corporate Structure*. New York: Macmillan.

Chandler, Alfred D., Jr. 1962 *Strategy and Structure: Chapters in the History of Industrial Enterprise*. Cambridge, Mass.: MIT Press.

— 1990 *Scale and Scope: The Dynamics of Industrial Capitalism*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Fligstein, Neil 1990 *The Transformation of Corporate Control*. Cambridge, Mass.: Harvard University Press.

Gill, Stephen, and David Law 1988 *The Global Political Economy*. Baltimore: Johns Hopkins University Press.

Hedley, R. Alan 1999 "Transnational Corporations and Their Regulation: Issues and Strategies." *International Journal of Comparative Sociology* 40(2):215–230.

Karliner, Joshua 1997 *The Corporate Planet: Ecology and Politics in the Age of Globalization*. Sierra Club Books. Available at <http://www.sierraclub.org/books/>.

Kentor, Jeffrey 1998 "The Long-Term Effects of Foreign Investment Dependence on Economic Growth, 1940–1990." *American Journal of Sociology* 103(4):1024–1046.

Lall, Sanjaya, and P. Streeton 1977 *Foreign Investment, Transnationals, and Developing Countries*. London: Macmillan.

Ohmae, Kenichi 1991 *The Borderless World: Power and Strategy in the Interlinked Economy*. Hammersmith, UK: Fontana.

Taylor, M. J., and N. J. Thrift 1982 *The Geography of Multinationals: Studies in the Spatial Development and Economic Consequences of Multinational Corporations*. London: Croom Helm.

UNCTAD 1993 *World Investment Report 1993: Transnational Corporations and Integrated International Production*. New York: United Nations Conference on Trade and Development.

— 1998 *World Investment Report 1997: Trends and Determinants*. New York: United Nations Conference on Trade and Development.

UNCTC 1990 *The New Code Environment*, Series A, No. 16. New York: United Nations Centre on Transnational Corporations.

Wilkins, Mira 1974 *The Maturing of Multinational Enterprise: American Business Abroad from 1914 to 1970*. Cambridge, Mass.: Harvard University Press.

World Bank 1987 *World Development Report 1987*. New York: Oxford University Press.

R. ALAN HEDLEY

**TRANSSEXUALS**

See Sexual Orientation.

**TRANSVESTITISM**

See Sexual Orientation.

**TRIBES**

See Indigenous Peoples.

**TYPOLOGIES**

A typology is a multidimensional classification. The study of typological procedures is impeded by the use of a plethora of terms, some of which are used interchangeably. "Classification" can be defined as the grouping of entities on the basis of similarity. For example, humans can be classified into female and male. A related term is "taxonomy." According to Simpson (1961, p. 11), taxonomy "is the theoretical study of classification, including its bases, principles, procedures, and rules." Interestingly, the term "classification" has two meanings: One can speak of both the process of classification and its end product, a classification. The terms "classification," "typology," and "taxonomy" are all used widely and somewhat interchangeably in sociology.

Any classification must be mutually exclusive and exhaustive. This requires that there be only one cell for each case. For example, if humans are being classified by sex, this requires that every case

be placed in a cell (either male or female) but that *no* case be placed in more than one cell (no intermediate cases are allowed). It is assumed that the bases or dimensions for classification (such as sex) are clear and important (see Tiryakian 1968).

A type is one cell in a full typology. In sociology, emphasis often has been placed on one or a few types rather than on the full typology. The study of types developed largely as a verbal tradition in sociology and lately has been merged with a more recently developed quantitative approach.

In the verbal tradition, types were often defined as mental constructs or concepts, in contrast to empirically derived entities. Stinchcombe (1968, p. 43, original emphasis) says that “a *type concept* in scientific discourse is a concept which is constructed out of a *combination of the values of several variables*.” Lazarsfeld (1937, p. 120) says that “one is safe in saying that the concept of type is always used in referring to special compounds of attributes.” The variables that combine to form a type must be correlated or “connected to each other” (Stinchcombe 1968, pp. 44–45).

An important function of a type is to serve as a criterion point (for comparative purposes) for the study of other types or empirical phenomena. In this case, only a single type is formulated. The most famous single-type formulation is Weber’s ideal type:

*An ideal type is formed by the one-sided accentuation of one or more points of view. . . In its conceptual purity, this mental construct [Gedankenbild] cannot be found empirically anywhere in reality. It is a utopia. Historical research faces the task of determining in each individual case the extent to which this ideal-construct approximates to or diverges from reality, to what extent for example, the economic structure of a certain city is to be classified as a “city economy.”* (1947, p. 90, original emphasis)

This strategy has been criticized. Martindale is startled by the suggestion that “we compare actual individuals with the (admittedly imaginary) ideal typical individuals to see how much they deviate from them. This is nothing but a form of intellectual acrobatics, for actual individuals ought to deviate from the ideal type just as much as one made them deviate in the first place” (1960, p. 382).

Seizing on Weber’s statement that the pure ideal type “cannot be found empirically anywhere in reality,” critics view the ideal type as hypothetical and thus without a fixed position, rendering it useless as a criterion point. A more realistic interpretation is that the ideal type represents a type that *could* be found empirically; it is simply that the purest case is the one most useful as a criterion, and this case is unlikely to be found empirically. As an example, a proof specimen of a coin is the best criterion for classifying or grading other coins, but it is not found empirically in the sense of being in circulation. If it were circulated, its features soon would be worn to the extent that its value for comparison with other coins would be greatly diminished.

The strategy of the ideal type is a sound one. Its logic is simple, and the confusion surrounding it is unfortunate, perhaps being due in part to the translation of Weber’s work. The genius of the ideal type lies in its parsimony. Instead of using a large full typology (say, of 144 cells, many of which may turn out to be empirically null or empty), a researcher can utilize a single ideal type. Then, instead of dealing needlessly with many null cells, the researcher need only fill in cells for which there are actual empirical cases and only as those cases are encountered. The ideal type is an accentuated or magnified version (or purest form) of the type. Although rarely found empirically in this pure form, the ideal type serves as a good comparison point. It usually represents the highest value on each of the intercorrelated variables or the end point of the continuum. While one could use the middle of the continuum as a referent (just as one uses the mean or median), it is convenient and perhaps clearer to use the end point (just as one measures from the end of a ruler rather than from its middle or another intermediate point).

Another single type that is used as a criterion is the constructed type. McKinney (1966, p. 3, original emphasis) defines the constructed type as “a *purposive, planned selection, abstraction, combination, and (sometimes) accentuation of a set of criteria with empirical referents that serves as a basis for comparison of empirical cases*.” The constructed type is a more general form of the ideal type.

In addition to formulations that use a single type, there are formulations that use two or more types. One strategy involves the use of two “polar”



types (as in the North and South poles). These types serve as two bracketing criteria for the comparison of cases. A famous set of types is Tönnies's (1957) *Gemeinschaft* and *Gesellschaft* ("community" and "society"). Another is introvert and extrovert. Still others are primary and secondary groups and localistic and cosmopolitan communities (see McKinney 1966, p. 101, for these and other examples).

### SUBSTRUCTION

One problem with the common practice of using only a single type or a few types is that the underlying correlated dimensions on which they are based may not be clear. In some cases, it is possible to make these dimensions clear and extend them all to form a property space or attribute space; a set of axes representing the full range of values on each dimension. Then the existence of other potential related types that were not originally formulated, can be discerned. This process of extending the full property space and the resulting full typology from a single type or a few types is called *substruction* and was developed by Lazarsfeld (1937; Barton 1955). As an example, Barton (1955, pp. 51–52) performed a substruction in which the attributes underlying the four types of folkways, mores, law, and custom were extended to form a full property space. Barton found three underlying dimensions of the four types ("how originated," "how enforced," and "strength of group feeling") and combined them to form the property space.

### REDUCTION

The opposite of substruction is *reduction*. Reduction is used when one has a full typology that is unmanageable because of its size. The three basic forms of reduction presented by Lazarsfeld (1937, p. 127) are functional, arbitrary numerical, and pragmatic. Lazarsfeld's functional reduction consists of discarding from the typology all empirically null and thus unnecessary cells.

The second form of reduction is arbitrary numerical. Lazarsfeld (1937, p. 128) provides an example: In constructing an index of housing conditions, one might weight plumbing without central heat or a refrigerator as being equal to the other two without plumbing. Coding the existence of an attribute by 1 and the lack of it by 0 and

taking variables in this order (plumbing, central heat, refrigerator), Lazarsfeld is saying that  $(1, 0, 0) = (0, 1, 1)$ . Thus, two previously different three-dimensional cells are equated and reduced to one.

Lazarsfeld's third form of reduction is pragmatic reduction. It consists of collapsing contiguous cells together to make one larger (but generally more heterogeneous) cell. As Lazarsfeld (1937, p. 128) says, "in the case of pragmatic reduction, certain groups of combinations are contracted to one class in view of the research purpose." For examples of these three forms of reduction, see Bailey (1973).

With Lazarsfeld's rigorous work as a notable exception, it can be said that most work in the typological tradition has been qualitative. Blalock, commenting on McKinney's (1966) constructive typology, says:

*He [McKinney] also claims that there is nothing inherently anti-quantitative in the use of typologies. He notes that historically, however, researchers skilled in the use of typologies have not been statistically or mathematically inclined, and vice-versa. This may be one of the reasons for the existing gap between sociological theory and research. (1969, p. 33)*

A persistent problem in the qualitative typological tradition has been the confusion over the status of the type as a heuristic device, a mental construct, or an empirical entity. Winch (1947) distinguished between heuristic and empirical types. He said that heuristic types are conceptually derived and may not have empirical examples. Empirical types, in contrast, result solely from data analysis, without prior conceptualization. A persistent problem with the conceptual types, such as the ideal type, has been the problem of inappropriate reification. If a type is a construct, concept, or model, it may not be found empirically but is designed only to be heuristically used in developing theory. However, there is often a tendency over time to reify the type or act as though it were actually found empirically. Figure 1 shows that the qualitative tradition has both heuristic and empirical types, while the quantitative tradition (discussed below) has primarily empirical types, as its types are derived from data analysis.

In other cases in the qualitative typological tradition, types are meant as empirical phenom-

	<b>Qualitative</b>	<b>Quantitative</b>
<b>Heuristic</b>	Ideal type	Probably null
<b>Empirical</b>	Ethnographic types	Types derived from cluster analysis or numerical taxonomy

Figure 1. A Typology of Typologies

ena rather than heuristic devices. This is particularly true in the area of social ethnography or field research, where researchers eschew statistical analysis but analyze data resulting from field studies by developing typologies based on observations recorded in their field notes (see Spradley and McCurdy 1972). Typologies in this case take the form of tables with names or labels in the cells rather than frequencies of occurrence as in statistical tables. Here the labels or types are generally inductively or empirically derived through intensive study of groups in the field. However, even here there may be a distinction between the types derived by the researcher and the types actually used by the people being studied. For example, the types that tramps identify among themselves (mission stiff, bindle stiff) may be different from the types identified by researchers or the lay public (bums, winos, homeless persons). For a discussion of taxonomies in ethnographic research and a number of examples of actual taxonomies (inducting the tramp example), see Spradley and McCurdy (1972).

**EMPIRICAL DERIVATION**

Computerization has brought on a new era of quantitative typology construction, which now coexists with the older qualitative tradition. This new approach often is called numerical taxonomy, cluster analysis, or pattern recognition (see Sneath and Sokal 1973; Bailey 1974). In contrast to the earlier verbal approach, which largely dealt with concepts and mental constructs, the newer quantitative approach is largely empirical and inductive.

It begins with a data set and derives empirical types from the data through a variety of quantitative procedures, many of them computerized.

This newer statistical approach to classification can be elucidated through the monothetic-polythetic distinction. A typology is monothetic if the possession of a unique set of features is both necessary and sufficient for identifying a specimen as belonging to a particular cell in the typology. That is, each feature is necessary and the set is sufficient. Thus, no specimen can be assigned to a particular type unless it possesses all the features (and no others) required of that type. This means that all the specimens in a given type are identical in every way (at least in all the features specified).

In contrast, a polythetic typology is constructed by grouping together the individuals within a sample that have the greatest number of shared features. No single feature is either necessary or sufficient (Sokal and Sneath 1963, p. 14). The objects or specimens are grouped to maximize overall similarity within each group. In a polythetic type, each individual possesses a large number of the classifying properties and each property is possessed by a large number of individuals. In the case where *no* single property is possessed by every individual in the group, the type is said to be fully polythetic.

While a verbal type (such as the ideal type) may be purely homogeneous (i.e., monothetic), it is unlikely that an empirically constructed type will be monothetic (except for some divisively derived types), especially if it contains a large number of cases grouped on a large number of variables. Thus, most empirically constructed types are polythetic, and some may be fully polythetic, without even a single feature being common to all the members of the group.

A basic distinction for all empirical classification techniques is whether one groups objects or variables. The former is known as Q-analysis, and the latter as R-analysis (Sokal and Sneath 1963, p. 124). In R-analysis, one computes coefficients (either similarity or distance coefficients) down the columns of the basic score matrix, which includes objects and variables (see Table 1 in Bailey 1972). In Q-analysis, one correlates rows. The interior data cells are the same in any case, and one form is the simple matrix transposition of the other. The difference is that Q-analysis correlates the objects

(e.g., persons), while R-analysis correlates the variables (e.g., age). While Q-analysis is the most common form in biology (see Sneath and Sokal 1973), it rarely is used in sociology (for an example, see Butler and Adams 1966). One problem is that Q-analysis requires a small sample of cases measured on a large number of variables, while R-analysis requires a large sample of cases with a smaller number of variables. Biology has the former sort of data; sociology, the latter.

Most sociologists have had little experience with Q-analysis. Most statistical analysis in sociology is concerned with relationships between two or more variables, with few studies making inferences concerning individuals rather than variables. Thus, the very notion of correlating individuals is alien to many sociologists.

Once the researcher has decided whether to pursue Q-analysis or R-analysis, the next step is to decide which measure of similarity to use. A researcher can measure similarity either directly, with a correlation coefficient, or indirectly, with a distance coefficient. While similarity coefficients show how close together two objects or variables are in the property space, distance coefficients show how far apart they are in that space. For a discussion of these measures, see Bailey (1974).

The next task of empirical typology construction is to parsimoniously group the cases into homogeneous types. There are two chief ways to proceed. One can envision all  $N$  cases as forming a single type. This is maximally parsimonious but maximizes within-group or internal variance. Grouping proceeds “from above” by dividing the cases into smaller groups that are more homogeneous. This is called the *divisive strategy*. Divisive classification generally proceeds by dividing the group on the basis of similarity on one or more variables, either simultaneously or sequentially. According to Sokal and Sneath (1963, p. 16), divisive classification is “inevitably largely monothetic.”

The alternative strategy (the *agglomerative strategy*) is to envision the  $N$  cases as forming  $N$  separate groups of one case each. Then each group is homogeneous (including only a single case), but parsimony is minimal. The strategy here is “classification from below” by agglomerating or grouping the most similar cases together, yielding some loss of internal homogeneity but gaining parsimony (as  $N$  groups are generally too unwieldy).

Unlike divisively formed types, agglomeratively formed types are generally polythetic and often fully polythetic.

The basic typological strategy is very straightforward and logically simple for divisive methods. All one must do is partition the set of cases in all possible ways and choose the grouping that maximizes internal homogeneity in a sufficiently small number of clusters. The problem is that the computation is prohibitive even for a modest number of cases measured on a modest number of variables.

A basic problem with empirically derived typologies is that they are generally static because the measures of similarity or distance that are used are synchronic rather than diachronic. While this is a problem, it is not a problem unique to classification but is shared by almost all forms of sociological analysis. Further, it is possible to deal with this issue by using diachronic data such as change coefficients or time series data.

Despite procedural differences, there are clear congruences between the qualitative and quantitative typological approaches. The ideal type is essentially monothetic, as are some types produced by quantitative divisive procedures. Quantitative procedures produce types that are polythetic, even fully polythetic. The results of quantitative procedures are generally not full typologies but reduced form that include fewer than the potential maximum number of types. Such polythetic types can be seen as analogous to the result of subjecting full monothetic typologies to reduction (either pragmatic or arbitrary numerical). Thus, contemporary typologists meet the need for reduction by using quantitative methods. Any correlational method of typology construction is by definition a method of functional reduction.

Further, the method usually will perform pragmatic reduction along with the functional reduction. Remember that pragmatic reduction collapses monothetic cells. The correlation coefficients utilized in typological methods are never perfect. The lower the correlations are, the more diverse the individuals in a group are. Placing diverse individuals in one group is tantamount to collapsing monothetic cells by means of pragmatic reduction. Thus, there are two basic avenues for constructing reduced types: Begin with monothetic types (such as ideal types) and subject them to the various forms of reduction to yield polythetic types

or construct polythetic types directly by using quantitative methods. Thus, the qualitative and quantitative procedures can produce similar results.

Given the breadth and diversity of sociological typologies (for example, from quantitative to qualitative procedures and from heuristic to empirical types), it is not surprising that there have been a number of criticisms of typologies. Some alleged problems are that typologies are not mutually exclusive and exhaustive, are treated as ends in themselves rather than as means to an end, are not parsimonious, are based on arbitrary and ad hoc criteria, are essentially static, rely on dichotomized rather than internally measured variables, yield types that are subject to reification, and are basically descriptive rather than explanatory or predictive. All these factors can be problems but are relatively easy for a knowledgeable typologist to avoid. The ones that cannot be easily avoided (such as the problem of cross-sectional data) often are seen as general problems for sociology as a whole and are not specific to typology construction.

### MERITS

Even if pitfalls remain, the merits of carefully constructed typologies make them well worth the effort. One of the chief merits of a typology is parsimony. A researcher who is overwhelmed by thousands or even millions of individual cases can work comfortably with those cases when they are grouped into a few main types. A related merit is the emphasis on bringing simplicity and order out of complexity and chaos. A focus on the relative homogeneity of types provides an emphasis on order in contrast to the emphasis on diversity and complexity that is paramount in untyped phenomena. A third merit of a full typology is its comprehensiveness. There is no other tool available that can show not only all relevant dimensions but also the relationships between them and the categories created by the intersections. Such a typology shows the entire range of every variable and all their confluences. A fourth merit (as was noted above) is a typology's use of a type or types for comparative purposes. A fifth merit is a typology's use as a heuristic tool to highlight the relevant theoretical dimensions of a type. A sixth is a typology's ability to show which cells have empirical examples and which are empirically null. This can aid in hypothesis testing, especially when a

large number of variables have a small number of values that actually occur (Stinchcombe 1968, p. 47). A seventh merit is a typology's ability to combine two or more variables in such a way that interaction effects can be analyzed (Stinchcombe 1968, pp. 46-47).

### TYPOLOGIES AND CONTINUOUS DATA

A clear but sometimes unstated goal of scientific development is to move past simple, nominal-variable analysis to the use of complex continuous-data models by employing ratio or interval variables. This has clearly been the case in sociology, which now depends on sophisticated regression models that work best with ratio (or at least interval) variables. Thus, some might argue that as science moves away from types toward the use of variables, typology construction becomes secondary.

Although the logic of moving from a reliance on types to a reliance on interval and ratio variables may seem irrefutable, this transition is not as smooth as some might wish. In fact, a number of obstacles to the transition from types to variables have arisen. Some researchers feel that once they have adopted sophisticated statistical techniques that use ratio variables, typologies are no longer needed. The reasoning here is that typologies are chiefly descriptive, arise at an early level of scientific analysis, and are essentially crude or unsophisticated formulations. In contrast, later models focus on explanation and prediction rather than description.

This notion belies the fact that science must constantly develop new ideas and theories to regenerate itself. As it does so, it must repeat the process of providing sound typologies that facilitate research by aiding in concept development and clarification and provide a comprehensive overview. Thus, it is a dangerous myth to think that sociology has "outgrown" the need for typologies. In fact, new ideas, theories, and sociological areas of research continually require new typologies. Even researchers in older, more mature sociological areas that have based their theory and research on inadequate typologies may find that the foundations of their field are crumbling, requiring new attempts to provide sound typological reinforcements.

In addition to the constant need for typological renewal and rejuvenation, some sociologists find

that attempts to move past types to sophisticated statistical analyses of ratio variables are confronted with a bewildering array of obstacles. Contemporary sociological statisticians who wish to rely on ratio variables are faced with a classic paradox. On the one hand, their regression models assume (or even demand) at least interval, or ideally ratio, variables. On the other hand, sociological theory is dependent on empirically important concepts, many of which are found to be essentially nominal or ordinal in their measurement levels. These include central ascribed or achieved statuses such as gender, race, religion, geographic region, nationality, occupation, and political affiliation.

Other important variables, such as income, education, and age, are more suitable for sophisticated statistical models. However, even these variables often are utilized theoretically in a limited ordinal form (young–old, high income–low income, etc.). Thus, there may be an empirical disjuncture between the type of variable needed for regression analysis (or other modern statistical techniques) and the type required by empirical sociological theory. Theory needs concepts such as race, gender, and religion, and these concepts are more suited for typological analysis than for regression analysis.

This suggests two areas of future research. One is to modify regression models to accommodate categorical variables, and this has been done (Aldrich and Nelson 1984). However, such accommodation may be costly, as it is unclear whether modified models operate efficiently or significantly underestimate the degree of explained variance. The second avenue is to rely more heavily on typological analysis. Although this may not seem as “sophisticated,” it may prove more compatible with theory and thus facilitate theoretical development more than statistical models do.

#### **TYPOLOGIES IN THE AGE OF STATISTICS**

If one has to choose between a sophisticated statistical analysis with variables that are not central to sociological theory and a typological analysis that accommodates theoretically important variables, it is foolish to rule out the latter in the name of scientific progress. Such progress would be false if the use of sophisticated techniques proved theo-

retically vacuous. This would be a classic case of the statistical tale wagging the theoretical dog. A wiser course is to recognize the complementarity between typologies and statistics. Statistics need not be viewed as necessarily or inevitably supplanting typologies; instead, each can be used when it proves valuable.

The conclusion to this point is that sociological progress has not rendered typological analysis obsolete by emphasizing statistical techniques such as multiple regression analysis. Thus, it may prove useful to look further at the epistemological foundations of contemporary sociology to see what the role of typologies is in an era when statistical analysis dominates. Consider the gap between the language of theory construction and the language of statistical data analysis. Imagine that a sociologist is interested in the type concept of “underachiever” and defines it as a person who has the ability to achieve at a higher level than is actualized.

When one substructures this type, it is clear that it is formed from two dimensions: (1) individual ability and (2) individual achievement. The sociologist can then theorize that an affluent childhood results in a particular type of personality. Individuals with that personality feel no pressing psychological need to achieve at a high level, since their needs continue to be met. This is an intriguing and ideographically rich sociological hypothesis. It involves images of a living person who has a particular type of childhood that leads to a particular type of adulthood. Thus, an earlier type concept (“the rich kid”) evolves into a later type concept (“the underachiever”). Conversely, one could hypothesize that the type concept of “impoverished youth” leads to the subsequent adult concept of “overachiever.”

The most direct way to test the hypothesis that the rich kid evolves into the adult underachiever is to identify a group of rich kids, follow them until adulthood, and then measure their subsequent achievement rates over a period of time. However, this is both tedious and time-consuming and is not the typical approach in social science. The most common approach is to gather cross-sectional survey data and then conduct a statistical analysis on the data. It is simple to select the two salient variables of parental wealth and adult achievement. Suppose one finds a negative correlation between parental wealth and adult achievement.

Since a negative correlation of achievement with wealth is not synonymous with the type concept of underachievement, the data analysis is not adequate to test the hypothesis.

Even if the statistical analysis were sufficient to test the hypothesis, a mere correlation value (e.g.,  $r = .43$ ) is very sterile and is isolated from both sociological reality and the richness of sociological theory. It fails to convey the richness of the type-concept description. While the type referent for the type concept of “underachiever” is the holistic, living human individual, the referents for the statistical analysis are the variables of wealth and achievement, which seem artificially separated from the sociological reality the theory refers to and the type concept manages to capture.

The unfortunate aspect of this for sociological development is that it leaves theory construction and statistical analysis as two juxtaposed but separate entities with a clear disjuncture between them. This disjuncture results from the fact that theorizing is largely a conceptual undertaking. It involves both deductive and inductive reasoning, and its language is the holistic language of the individual actor. The prime theoretical referent is the object, not the variable. This object is usually the human individual but can be an alternative object, such as a group, city, or country. In any event, the primary focus is on the object, with variables receiving a secondary focus. However, even if variables have the primary focus, the focus remains on both object and the variables.

In statistics using R-analysis such as multiple regression, the epistemological focus is quite different. Here objects such as persons enter the analysis only as data carriers in the sample. As soon as the R-matrix of correlations among variables is established, it suffices for the remainder of the analysis. The result is that the individuals virtually disappear from the picture except in those rare instances in sociology where Q-correlations are used.

Thus, theory and statistical analysis remain two separate paradigms within sociology rather than two aspects of the same research process. This obviously hinders scientific progress sociology and stands in stark contrast to the physical sciences, where theory and method are not separated processes but are well integrated, enabling much swifter progress.

### INTEGRATING TYPES AND TAXA

One way to bridge the dichotomy between theory and statistical method is to link qualitative type concepts with the empirical clusters derived quantitatively through methods of numerical taxonomy. Following the lead of Bailey (1994), these are called taxa. As was noted above, types are generally conceptual, monothetic, and based on underlying R-dimensions (although the cell entries are empirical objects). In contrast, taxa tend to be empirical, polythetic, and Q-analytic (based on individuals). While the differences may seem to mimic the differences between theory and statistics discussed above, both types and taxa can be seen primarily as mirror images of each other and thus as having structural similarities that allow a bridge to be built from one to the other.

Since much theorizing is done in terms of types, a needed first step is to move from the realm of type concepts to the realm of empirical data analysis. While this traditionally is accomplished by turning from theory to statistical analysis, an alternative is to link conceptually formed types with statistically derived taxa.

The first task is to move from the conceptual to the empirical. As outlined in Bailey (1994, p. 66), this is rather straightforward and merely involves the identification of empirical cases for each conceptual cell. An example would be to locate an actual ethnographic type such as “bundle stiff” (Spradley and McCurdy, 1972). The empirical cases found for each cell in a typology (such as Figure 1) are equivalent to the taxa formed through cluster analysis. The second task in bridging the gap between types and taxa is converting from monothetic to polythetic. As was discussed above, this can be achieved in various ways, such as Lazarfeld’s (1937) process of pragmatic reduction. The third task is to connect the R-analysis of types with the Q-analysis of taxa. The easiest way to accomplish this is to use R-analysis for clustering.

In the other direction—from taxa to types—all the tasks are reversed and involve going from empirical to conceptual, from polythetic to monothetic, and from Q-analysis to R-analysis. Going from empirical to conceptual entails finding a concept to represent the statistically constructed group. For example, if the cluster analysis yields an empirical cluster composed primarily of people who scored very high on an exam, one

could formulate the type concept of "high achievers" to represent it.

The second task involves going from polythetic to monothetic. Technically speaking, this entails changing a heterogeneous empirical grouping to a homogeneous grouping and cannot be accomplished empirically except somewhat artificially. For example, Lockhart and Hartman (1963) constructed monothetic clusters by discarding all the characters that varied within the group. This is compensated for by the prior step, in which the conceptual type concept monothetically represents the empirical polythetic taxa. The third task is to achieve R-analytic clustering by using R-correlations rather than Q-correlations in the cluster analysis.

### CONCLUSION

A well-constructed typology can bring order out of chaos. It can transform the overwhelming complexity of an apparently eclectic congeries of numerous apparently diverse cases into a well-ordered set of a few homogeneous types clearly situated in a property space of a few important dimensions. A sound typology forms a firm foundation and provides direction for both theorizing and empirical research. No other tool has as much power to simplify life for a sociologist.

The task for the future is the further elaboration of this crucial nexus between the qualitative and statistical approaches. This requires effort from sociologists with both theoretical and statistical talents. McKinney (1966, p. 49) recognizes the "complementary relationship of quantitative and typological procedures" and advocates "the emergence of a number of social scientists who are procedurally competent in both typology and statistical techniques." Costner (1972, p. xi) also recognizes the basic unity of the qualitative and quantitative approaches to typology construction.

For further information on typologies, see Capecchi (1966), Sokal and Sneath, (1963), Sneath and Sokal (1973), Bailey (1973, 1974, 1983, 1989, 1993, 1994), Hudson et al. (1982), Aldenderfer and Blashfield (1984), and Kreps (1989).

(SEE ALSO: *Levels of Analysis*; *Tabular Analysis*)

### REFERENCES

- Aldenderfer, Mark S., and Roger K. Blashfield 1984 *Cluster Analysis*. Thousand Oaks, Calif.: Sage.
- Aldrich, John H., and Forrest D. Nelson 1984 *Cluster Analysis*. Thousand Oaks, Calif.: Sage.
- Bailey, Kenneth D. 1972 "Polythetic Reduction of Monothetic Property Space." In Herbert L. Costner, ed., *Sociological Methodology 1972*. San Francisco: Jossey-Bass.
- 1973 "Monothetic and Polythetic Typologies and Their Relationship to Conceptualization, Measurement, and Scaling." *American Sociological Review* 38:18-33.
- 1974 "Cluster Analysis." In David R. Heise, ed., *Sociological Methodology 1975*. San Francisco: Jossey-Bass.
- 1983 "Sociological Classification and Cluster Analysis." *Quality and Quantity* 17:251-268.
- 1989 "Taxonomy and Disaster: Prospects and Problems." *International Journal of Mass Emergencies and Disasters* 7:419-431.
- 1993 "Strategies of Nucleus Formation in Agglomerative Clustering Techniques." *Bulletin De Methodologie Sociologique* 38:38-51.
- 1994 *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, Calif.: Sage.
- Barton, Allen H. 1955 "The Concept of Property Space in Social Research." In Paul F. Lazarsfeld and Morris Rosenberg, eds., *The Language of Social Research*. New York: Free Press.
- Blalock, Herbert M. 1969 *Theory Construction: From Verbal to Mathematical Formulations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Butler, Edgar W., and Stuart N. Adams 1966 "Typologies of Delinquent Girls: Some Alternative Approaches." *Social Forces* 44:401-407.
- Capecchi, Vittorio 1966 "Typologies in Relation to Mathematical Models." *Ikon* Suppl. No. 58:1-62.
- Costner, Herbert L. 1972 "Prologue." In Herbert L. Costner, ed., *Sociological Methodology 1972*. San Francisco: Jossey-Bass.
- Hudson, Herschel C., and associates (eds.) 1982 *Classifying Social Data*. San Francisco: Jossey-Bass.
- Kreps, Gary A. (ed.) 1989 "The Boundaries of Disaster Research: Taxonomy and Comparative Research" (Special Issue). *International Journal of Mass Emergencies and Disasters* 7:213-431.
- Lazarsfeld, Paul F. 1937 "Some Remarks on the Typological Procedures in Social Research." *Zeitschrift für Sozialforschung* 6:119-139.

- Lockhart, W. R., and P. A. Hartman 1963 "Formation of Monothetic Groups in Quantitative Bacterial Taxonomy." *Journal of Bacteriology* 85:68-77.
- Martindale, Don 1960 *The Nature and Types of Sociological Theory*. Boston: Houghton Mifflin.
- McKinney, John C. 1966 *Constructive Typology and Social Theory*. New York: Appleton, Century, Crofts.
- Simpson, George G. 1961 *Principles of Animal Taxonomy*. New York: Columbia University Press.
- Sneath, Peter H. A., and Robert R. Sokal 1973 *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco: Freeman.
- Sokal, Robert R., and Peter H. A. Sneath 1963 *Principles of Numerical Taxonomy*. San Francisco: Freeman.
- Spradley, James P., and David W. McCurdy 1972 *The Cultural Experience: Ethnography in Complex Society*. Chicago: Science Research Associates.
- Stinchcombe, Arthur L. 1968 *Constructing Social Theories*. New York: Harcourt, Brace, and World.
- Tiryakian, Edward A. 1968 "Typologies." In David L. Sills, ed., *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press.
- Tönnies, Ferdinand 1957 *Gemeinschaft und Gesellschaft*, trans. and ed. C. P. Loomis. East Lansing: Michigan State University Press.
- Weber, Max 1947 *Theory of Social and Economic Organization*. A. R. Henderson and Talcott Parsons, trans., and Talcott Parsons, ed. New York: Free Press.
- Winch, Robert F. 1947 "Heuristic and Empirical Typologies: A Job for Factor Analysis." *American Sociological Review* 12:68-75.

KENNETH D. BAILEY



# U

---

## UNIONS

*See* Labor Movements and Unions; Industrial Sociology.

## URBAN SOCIOLOGY

Urban sociology studies human groups in a territorial frame of reference. In this field, social organization is the major focus of inquiry, with an emphasis on the interplay between social and spatial organization and the ways in which changes in spatial organization affect social and psychological well being. A wide variety of interests are tied together by a common curiosity about the changing dynamics, determinants, and consequences of urban society's most characteristic form of settlement: the city.

Scholars recognized early that urbanization is accompanied by dramatic structural, cognitive, and behavioral changes. Classic sociologists (Durkheim, Weber, Toinnes, Marx) delineated the differences in institutional forms that seemed to accompany the dual processes of urbanization and industrialization as rural-agrarian societies were transformed into urban-industrial societies (see Table 1).

Several key questions that guide contemporary research are derived from this tradition: How are human communities organized? What forces produce revolutionary transformations in human settlement patterns? What organizational forms accompany these transformations? What differ-

ences do urban living make, and why do those differences exist? What consequences does the increasing size of human concentrations have for human beings, their social worlds, and their environment?

Students of the urban scene have long been interested in the emergence of cities (Childe 1950), how cities grow and change (Weber 1899), and unique ways of life associated with city living (Wirth 1938). These classic treatments have historical value for understanding the nature of pre-twentieth-century cities, their determinants, and their human consequences, but comparative analysis of contemporary urbanization processes leads Berry (1981, p. xv) to conclude that "what is apparent is an accelerating change in the nature of change itself, speedily rendering not-yet-conventional wisdom inappropriate at best."

Urban sociologists use several different approaches to the notion of community to capture changes in how individual urbanites are tied together into meaningful social groups and how those groups are tied to other social groups in the broader territory they occupy. An interactional community is indicated by networks of routine, face-to-face primary interaction among the members of a group. This is most evident among close friends and in families, tribes, and closely knit locality groups. An ecological community is delimited by routine patterns of activity that its members engage in to meet the basic requirements of daily life. It corresponds with the territory over which the group ranges in performing necessary activities such as work, sleep, shopping, education,

**Classic Contrasts Between Urban and Rural Societies**

Institution	Urban-Industrial	Rural-Agrarian
Agreements	Contractual	Personal
Authority	Bureaucratic	Paternalistic
Communication	Secondary	Primary
Integrative mechanism	Specialization	Common experience
Normative standards	Universalistic	Particularistic
Normative structure	Anomic	Integrated
Problem solution	Rational	Traditional
Production	Manufacturing	Agriculture
Social control	Restitutive	Repressive
Social relations	Segmentalized	All encompassing
Socialization	Formal	Informal
Stratification	Achieved status	Ascribed status
Values	Money and power	Family
World views	Secular	Sacred

Table 1

and recreation. Compositional communities are clusters of people who share common social characteristics. People of similar race, social status, or family characteristics, for example, form a compositional community. A symbolic community is defined by a commonality of beliefs and attitudes among its members. Its members view themselves as belonging to the group and are committed to it.

Research on the general issue of how these forms of organization change as cities grow has spawned a voluminous literature. An ecological perspective and a sociocultural perspective guide two major research traditions. Ecological studies focus on the role of economic competition in shaping the urban environment. Ecological and compositional communities are analyzed in an attempt to describe and generalize about urban forms and the processes of urban growth (Hawley 1981).

Sociocultural studies emphasize the importance of cultural, psychological, and other social dimensions of urban life. These studies focus on the interactional and symbolic communities that characterize the urban setting (Wellman and Leighton 1979; Suttles 1972).

Early theoretical work suggested that the most evident consequence of the increasing size, density, and heterogeneity of human settlements was a breakdown of social ties, a decline in the family, alienation, an erosion of moral codes, and social disorganization (Wirth 1938). Later empirical research has clearly shown that in general, urbanites

are integrated into meaningful social groups (Fischer 1984).

The sociocultural tradition suggests that cultural values derive from socialization into a variety of subcultures and are relatively undisturbed by changes in ecological processes. Different subcultures select, are forced into, or unwittingly drift into different areas that come to exhibit the characteristics of a particular subculture (Gans 1962). Fischer (1975) combines the ecological and subcultural perspectives by suggesting that size, density, and heterogeneity are important but that they produce integrated subcultures rather than fostering alienation and community disorganization. Size provides the critical masses necessary for viable unconventional subcultures to form. With increased variability in the subcultural mix in urban areas, subcultures become more intensified as they defend their ways of life against the broad array of others in the environment. The more subcultures, the more diffusion of cultural elements, and the greater the likelihood of new subcultures emerging, creating the ever-changing mosaic of unconventional subcultures that most distinguishes large places from small ones.

Empirical approaches to urban organization vary according to the unit of analysis and what is being observed. Patterns of activity (e.g., commuting, retail sales, crime) and characteristics of people (e.g., age, race, income, household composition) most commonly are derived from government reports for units of analysis as small as city blocks and as large as metropolitan areas. These types of

data are used to develop general principles of organization and change in urban systems. General questions range from how certain activities and characteristics come to be organized in particular ways in space to why certain locales exhibit particular characteristics and activities. Territorial frameworks for the analysis of urban systems include neighborhoods, community areas, cities, urban areas, metropolitan regions, nations, and the world.

Observations of networks of interaction (e.g., visiting patterns, helping networks) and symbolic meanings of people (e.g., alienation, values, worldviews) are less systematically available because social surveys are more appropriate for obtaining this kind of information. Consequently, less is known about these dimensions of community than is desirable.

It is clear that territoriality has waned as an integrative force and that new forms of extralocal community have emerged. High mobility, an expanded scale of organization, and an increased range and volume of communication flow coalesce to alter the forms of social groups and their organization in space (Greer 1962). With modern communication and transportation technology, as exists in the United States today, space becomes less of an organizing principle and new forms of territorial organization emerge that reflect the power of large-scale corporate organization and the federal government in shaping urban social and spatial organization (Gottdiener 1985).

Hawley's (1950, 1981) ecological approach to the study of urban communities serves as the major paradigm in contemporary research. This approach views social organization as developing in response to basic problems of existence that all populations face in adapting to their environments. The urban community is conceptualized as the complex system of interdependence that develops as a population collectively adapts to an environment, using whatever technology is available. Population, environment, technology, and social organization interact to produce various forms of human communities at different times and in different places (Table 2). Population is conceptualized as an organized group of humans that function routinely as a unit; the environment is defined as everything that is external to the population, including other organized social groups.

Technological advances allow people to expand and redefine the nature of the relevant environment and therefore influence the forms of community organization that populations develop (Duncan 1973).

In the last half of the twentieth century, there were revolutionary transformations in the size and nature of human settlements and the nature of the interrelationships among them (Table 3). The global population "explosion" created by an unprecedented rapid decline in human mortality in less developed regions of the world after 1950 provided the additional people necessary for this population "implosion:" the rapid increase in the size and number of human agglomerations of unprecedented size. Urban sociology attempts to understand the determinants and consequences of this transformation.

The urbanization process involves an expansion in the entire system of interrelationships by which a population maintains itself in its habitat (Hawley 1981, p. 12). The most evident consequences of the process and the most common measures of it are an increase in the number of people at points of population concentration, an increase in the number of points at which population is concentrated, or both (Eldridge 1956). Theories of urbanization attempt to understand how human settlement patterns change as technology expands the scale of social systems.

Because technological regimes, population growth mechanisms, and environmental contingencies change over time and vary in different regions of the world, variations in the pattern of distribution of human settlements generally can be understood by attending to these related processes. In the literature on urbanization, an interest in the organizational forms of systems of cities is complemented by an interest in how growth is accommodated in cities through changes in density gradients, the location of socially meaningful population subgroups, and patterns of urban activities. Although the expansion of cities has been the historical focus in describing the urbanization process, revolutionary developments in transportation, communication, and information technology in the last fifty years expanded the scale of urban systems and directed attention toward the broader system of the form of organization in which cities emerge and grow.

**Comparative Urban Features of Major World Regions**

Basic Feature	Nineteenth Century North America	Twentieth Century North America	Third World	Postwar Europe
<b>Summary</b>	Concentrated	Spread out	Constrained	Planned
<b>Size</b>	1–2 million	14 million	19 million	8 million
<b>Density</b>	High	Low	Medium	High
<b>Timing</b>	250 years long period	Emergent no pressure	Very rapid since 1950s	Very slow stationary
<b>Scale</b>	Regional and local	Inter-metro and global	Global and local	National and local
<b>City system</b>	Rank size regional	Daily urban national	Primate national	Rank size national
<b>Occupations</b>	Secondary manufacture	Tertiary services	Family and corporate	Diverse mixture
<b>Spatial mix</b>	Zone-sector core focus	Multinodal mosaic	Reverse zonal	Overlaid mixed use
<b>Rural–urban differences</b>	Great in all areas	Narrow and declining	Medium and growing	Narrow except work
<b>Status mix</b>	Diverse hierarchical	High overall poor pockets	Bifurcated high % poor	Medium compacted
<b>Migration</b>	Heavy rural-urban and foreign	Inter-metro and foreign	Heavy rural-urban circulation	Foreign skilled
<b>Planning</b>	Laissez-faire capitalism	Decentral, ineffective	Centralized, ineffective	Decentral, effective

*Table 2*

SOURCE: Abstracted from Berry 1981.

Much research on the urbanization process is descriptive in nature, with an emphasis on identifying and measuring patterns of change in demographic and social organization in a territorial frame of reference. Territorially circumscribed environments employed as units of analysis include administrative units (villages, cities, counties, states, nations), population concentrations (places, agglomerations, urbanized areas), and networks of interdependency (neighborhoods, metropolitan areas, daily urban systems, city systems, the earth).

The American urban system is suburbanizing and deconcentrating. One measure of suburbanization is the ratio of the rate of growth in the ring to that in the central city over a decade (Schnore 1959). While some Metropolitan Statistical Areas (MSAs) began suburbanizing in the late 1800s, the greatest rates for the majority of places occurred in the 1950s and 1960s. Widespread use of the automobile, inexpensive energy, the efficient production of materials for residential infrastructure,

and federal housing policy allowed metropolitan growth to be absorbed by sprawl instead of by increased congestion at the center.

As the scale of territorial organization increased, so did the physical distances between black and white, rich and poor, young and old, and other meaningful population subgroups. The Index of Dissimilarity measures the degree of segregation between two groups by computing the percentage of one group that would have to reside on a different city block for it to have the same proportional distribution across urban space as the group to which it is being compared (Taeuber and Taeuber 1965). Although there has been some decline in indices of dissimilarity between black and white Americans since the 1960s, partly as a result of increasing black suburbanization, the index for the fifteen most segregated MSAs in 1990 remained at or above 80, meaning that 80 percent or more of the blacks would have had to live on different city blocks to have the same distribution in space as whites; thus, a very high

**Population of World's Largest Metropolises  
(in millions), 1950–2000 and Percent Change, 1950–2000**

Metropolis	1950	2000	% Change
Mexico City, Mexico	3.1	26.3	748
Sao Paulo, Brazil	2.8	24.0	757
Tokyo/Yokohama, Japan	6.7	17.1	155
Calcutta, India	4.4	16.6	277
Greater Bombay, India	2.9	16.0	452
New York/northeastern N.J., USA	12.4	15.5	25
Seoul, Republic of Korea	1.1	13.5	113
Shanghai, China	10.3	13.5	31
Rio de Janeiro, Brazil	3.5	13.3	280
Delhi, India	1.4	13.2	843
Greater Buenos Aires, Argentina	5.3	13.2	149
Cairo/Giza/Imbaba, Egypt	2.5	13.2	428
Jakarta, Indonesia	1.8	12.8	611
Baghdad, Iraq	0.6	12.8	2033
Teheran, Iran	0.9	12.7	1311
Karachi, Pakistan	1.0	12.1	1110
Istanbul, Turkey	1.0	11.9	1090
Los Angeles/Long Beach, Calif., USA	4.1	11.2	173
Dacca, Bangladesh	0.4	11.2	2700
Manila, Philippines	1.6	11.1	594
Beijing (Peking), China	6.7	10.8	61
Moscow, USSR	4.8	10.1	110
Total world population	2,500	6,300	152

*Table 3*

SOURCE: Adapted from Dogan and Kasarda (1988b) Table 1.2.

degree of residential segregation remains. Although there is great social status diversity in central cities and increasing diversity in suburban rings, disadvantaged and minority populations are overrepresented in central cities, while the better educated and more affluent are overrepresented in suburban rings.

A related process—deconcentration—involves a shedding of urban activities at the center and is indicated by greater growth in employment and office space in the ring than in the central city. This process was under way by the mid-1970s and continued unabated through the 1980s. A surprising turn of events in the late 1970s was signaled by mounting evidence that nonmetropolitan counties were, for the first time since the Depression of the 1930s, growing more rapidly than were metropolitan counties (Lichter and Fuguitt 1982). This process has been referred to as “deurbanization” and “the nonmetropolitan turnaround.” It is unclear whether this trend represents an enlargement of the scale of metropolitan organization to encompass more remote counties or

whether new growth nodes are developing in nonmetropolitan areas.

The American urban system is undergoing major changes as a result of shifts from a manufacturing economy to a service economy, the aging of the population, and an expansion of organizational scale from regional and national to global decision making. Older industrial cities in the Northeast and Midwest lost population as the locus of economic activity shifted from heavy manufacturing to information and residentiary services. Cities in Florida, Arizona, California, and the Northwest have received growing numbers of retirees seeking environmental, recreational, and medical amenities that are not tied to economic production. Investment decisions regarding the location of office complexes, the factories of the future, are made more on the basis of the availability of an educated labor pool, favorable tax treatment, and the availability of amenities than on the basis of the access to raw materials that underpinned the urbanization process through the middle of the twentieth century.

The same shifts are reflected in the internal reorganization of American cities. The scale of local communities has expanded from the central business district-oriented city to the multinodal metropolis. Daily commuting patterns are shifting from radial trips between bedroom suburbs and workplaces in the central city to lateral trips among highly differentiated subareas throughout urban regions. Urban villages with affluent residences, high-end retail minimalls, and office complexes are emerging in nonmetropolitan counties beyond the reach of metropolitan political constraints, creating even greater segregation between the most and least affluent Americans

Deteriorating residential and warehousing districts adjacent to new downtown office complexes are being rehabilitated for residential use by childless professionals, or "gentry." The process of gentrification, or the invasion of lower-status deteriorating neighborhoods of absentee-owned rental housing by middle- to upper-status home or condominium owners, is driven by a desire for accessibility to nearby white-collar jobs and cultural amenities as well as by the relatively high costs of suburban housing, which have been pushed up by competing demand in these rapidly growing metropolitan areas. Although the number of people involved in gentrification is too small to have reversed the overall decline of central cities, the return of affluent middle-class residents has reduced segregation to some extent. Gentrification reclaims deteriorated neighborhoods, but it also results in the displacement of the poor, who have no place else to live at rents they can afford (Feagin and Parker 1990).

The extent to which dispersed population is involved in urban systems is quite variable. An estimated 90 percent of the American population now lives in a daily urban system (DUS). These units are constructed from counties that are allocated to economic centers on the basis of commuting patterns and economic interdependence. The residents of a DUS are closely tied together by efficient transportation and communication technology. Each DUS has a minimum population of 200,000 in its labor shed and constitutes "a multinode, multiconnective system [which] has replaced the core dominated metropolis as the basic urban unit" (Berry and Kasarda 1977, p. 304). Less than 4 percent of the American labor force is engaged in agricultural occupations. Even

the residents of remote rural areas are mostly "urban" in their activities and outlook.

In contrast, many residents of uncontrolled developments on the fringes of emerging megacities in less developed countries are practically isolated from the urban center and live much as they have for generations. Over a third of the people in the largest cities in India were born elsewhere, and the maintenance of rural ways of life in those cities is common because of a lack of urban employment, the persistence of village kinship ties, and seasonal circulatory migration to rural areas. Although India has three of the ten largest cities in the world, it remains decidedly rural, with 75 percent of the population residing in agriculturally oriented villages (Nagpaul 1988).

The pace and direction of the urbanization process are closely tied to technological advances. As industrialization proceeded in western Europe and the United States over a 300-year period, an urban system emerged that reflected the interplay between the development of city-centered heavy industry and requirements for energy and raw materials from regional hinterlands. The form of city systems that emerged has been described as rank-size. Cities in that type of system form a hierarchy of places from large to small in which the number of places of a given size decreases proportionally to the size of the place. Larger places are fewer in number, are more widely spaced, and offer more specialized goods and services than do smaller places (Christaller 1933).

City systems that emerged in less industrialized nations are primate in character. In a primate system, the largest cities absorb far more than their share of societal population growth. Sharp breaks exist in the size hierarchy of places, with one or two very large, several medium-sized, and many very small places. Rapid declines in mortality beginning in the 1950s, coupled with traditionally high fertility, created unprecedented rates of population growth. Primate city systems developed with an orientation toward the exportation of raw materials to the industrialized world rather than manufacturing and the development of local markets. As economic development proceeds, it occurs primarily in the large primate cities, with very low rates of economic growth in rural areas. Consequently, nearly all the excess of births over deaths

in the nation is absorbed by the large cities, which are more integrated into the emerging global urban system (Dogan and Kasarda 1988a).

Megacities of over 10 million population are a very recent phenomenon, and their number is increasing rapidly. Their emergence can be understood only in the context of a globally interdependent system of relationships. The territorial bounds of the relevant environment to which population collectively adapts have expanded from the immediate hinterland to the entire world in only half a century.

Convergence theory suggests that cities throughout the world will come to exhibit organizational forms increasingly similar to one another, converging on the North American pattern, as technology becomes more accessible globally (Young and Young 1962). Divergence theory suggests that increasingly divergent forms of urban organization are likely to emerge as a result of differences in the timing and pace of the urbanization process, differences in the positions of cities in the global system, and the increasing effectiveness of deliberate planning of the urbanization process by centralized governments holding differing values and therefore pursuing a variety of goals for the future (Berry 1981).

The importance of understanding this process is suggested by Hawley (1981, p. 13): "Urbanization is a transformation of society, the effects of which penetrate every sphere of personal and collective life. It affects the status of the individual and opportunities for advancement, it alters the types of social units in which people group themselves, and it sorts people into new and shifting patterns of stratification. The distribution of power is altered, normal social processes are reconstituted, and the rules and norms by which behavior is guided are redesigned."

## REFERENCES

- Berry, Brian J. L. 1981. *Comparative Urbanization: Divergent Paths in the Twentieth Century*. New York: St. Martins.
- , and John D. Kasarda 1977 *Contemporary Urban Ecology*. New York: Macmillan.
- Childe, V. Gordon 1950 "The Urban Revolution." *Town Planning Review* 21:4–7.
- Christaller, W. 1933 *Central Places in Southern Germany*, transl. C. W. Baskin. Englewood Cliffs, N.J.: Prentice-Hall.
- Dogan, Mattei, and John D. Kasarda 1988a *The Metropolis Era: A World of Giant Cities*, vol. 1. Newbury Park, Calif.: Sage.
- 1988b. "Introduction: How Giant Cities Will Multiply and Grow." In Mattei Dogan and John D. Kasarda, eds., *The Metropolis Era: A World of Giant Cities*, vol. 1. Newbury Park, Calif.: Sage.
- Duncan, Otis Dudley 1973 "From Social System to Ecosystem." In Michael Micklin, ed., *Population, Environment, and Social Organization: Current Issues in Human Ecology* Hinsdale, Ill.: Dryden.
- Eldridge, Hope Tisdale 1956 "The Process of Urbanization." In J. J. Spengler and O. D. Duncan, eds., *Demographic Analysis*. Glencoe, Ill.: Free Press.
- Feagin, Joe R., and Robert Parker 1990 *Building American Cities: The Urban Real Estate Game*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Fischer, Claude S. 1975 "Toward a Subcultural Theory of Urbanism." *American Journal of Sociology* 80:1319–1341.
- 1984 *The Urban Experience*. San Diego: Harcourt Brace Jovanovich.
- Gans, Herbert J. 1962 "Urbanism and Suburbanism as Ways of life: A Reevaluation of Definitions." In A. M. Rose, ed., *Human Behavior and Social Processes*. Boston: Houghton Mifflin.
- Gottdiener, Mark 1985 *The Social Production of Urban Space*. Austin: University of Texas Press.
- Greer, Scott 1962 *The Emerging City*. New York: Free Press.
- Hawley, Amos H. 1950 *Human Ecology: A Theory of Community Structure*. New York: Ronald.
- 1981 *Urban Society: An Ecological Approach*. New York: Wiley.
- Kleniewski, Nancy 1997 *Cities, Change, and Conflict: A Political Economy of Urban Life*. Belmont, Calif.: Wadsworth.
- Lichter, Daniel T., and Glenn V. Fuguitt 1982 "The Transition to Nonmetropolitan Population Decentralization." *Demography* 19:211–221.
- Nagpaul, Hans 1988 "India's Giant Cities." In Mattei Dogan and John D. Kasarda, eds., *The Metropolis Era: A World of Giant Cities*, vol. 1. Newbury Park, Calif.: Sage.
- Palen, J. John 1997 *The Urban World*. New York: McGraw-Hill.
- Schnore, Leo F. 1959 "The Timing of Metropolitan Decentralization." *Journal of the American Institute of Planners* 25:200–206.

- Suttles, Gerald 1972 *The Social Construction of Communities*. Chicago: University of Chicago Press.
- Taeuber, Karl E., and Alma F. Taeuber 1965 *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago: Aldine.
- Weber, Adna F. 1899 *The Growth of Cities in the Nineteenth Century*. New York: Columbia University Press.
- Wellman, B., and B. Leighton 1979 "Networks, Neighborhoods and Communities: Approaches to the Study of the Community Question." *Urban Affairs Quarterly* 15:369-393.
- Wirth, Louis 1938 "Urbanism as a Way of Life." *American Journal of Sociology* 44:1-24.
- Young, Frank, and Ruth Young 1962 "The Sequence and Direction of Community Growth: A Cross-Cultural Generalization." *Rural Sociology* 27:374-386.

LEE J. HAGGERTY

## URBAN UNDERCLASS

No social science concept has generated more discussion and controversy in recent years than that of the urban underclass. Some argue that it is little more than a pithy and stigmatizing term for the poor people who have always existed in stratified societies (Gans 1990; Jencks 1989; Katz 1989; McGahey 1982). Others contend that the underclass is a distinct and recent phenomenon, reflecting extreme marginalization from mainstream economic institutions and aberrant behavior (drug abuse, violent crime, out-of-wedlock births), that reached catastrophic proportions in the inner cities by the early 1980s (Glasow 1980; Auletta 1982; Reischauer 1987; Nathan 1987; Wilson 1987, 1996). Among the multifaceted, subjective, and often ambiguous definitions of the urban underclass, most all include the notions of weak labor-force attachment and persistently low income (Jencks and Peterson 1991; Sjoquist 1990). Indeed, the first scholar who introduced the term "underclass" in literature characterized its members as an emergent substratum of the permanently unemployed, the unemployable and the underemployed (Myrdal 1962).

Widely differing interpretations of the causes of the presence of an underclass have been offered, ranging from Marxist to social Darwinist. The most influential contemporary analysis of the urban underclass is Wilson's (1996) *When Work*

*Disappears*. Building on his earlier treatise *The Truly Disadvantaged* (1987), Wilson links the origins and growth of the urban underclass to the structure of opportunities and constraints in American society. Its roots are hypothesized to lie in historical discrimination and the mass migration of African-Americans to northern cities in the first half of the twentieth century. Its more recent growth and experiences are posited to have resulted from industrial restructuring and geographic changes in metropolitan economies since the 1960s, in particular the economic transformation of major cities from centers of goods processing to centers of information processing and the relocation of blue-collar jobs to the suburbs. These changes led to sharp increases in joblessness among racially and economically segregated African-Americans who had neither the skills to participate in new urban growth industries nor the transportation or financial means to commute or relocate to the suburbs. Rapidly rising joblessness among inner-city African-Americans, together with selective outmigration of the nonpoor, in turn caused the high concentrations of poverty and related social problems that characterize the urban underclass (see also Kasarda 1985, 1989; Wilson 1991; Hughes, 1993).

Alternative views on the cause of the underclass appear in the works of Murray (1984), Mead (1988), and Magnet (1993). These conservative scholars view underclass behaviors as rational adaptations to the perverse incentives offered by government welfare programs that discourage work and a lack of personal responsibility among many for actions harmful to themselves and others. Abetted by well-intentioned but misguided public programs, joblessness and persistent poverty are seen more as the consequences of deviant behaviors than as the causes of those behaviors. For an elaboration of these competing views and a partial empirical assessment, see Kasarda and Ting (1996).

Measurement of the size of the underclass varies as much as explanations of its causes. A number of researchers have focused on individual-level indicators of persistent poverty, defined as those who are poor for spells from  $n$  to  $n + x$  years (Levy 1977; Duncan et al. 1984; Bane and Ellwood 1986) and long-term Aid to Families with Dependent Children (AFDC) recipients (Gottschalk and Danziger 1987). In an empirical study, Levy (1977) estimated that approximately eleven million Ameri-



cans were persistently poor for at least five years. When the underclass was defined as those who were not need persistently poor for eight or more years, six million people were found in that category (Duncan et al. 1984). This represented approximately one-fifth of the thirty-two million Americans living in poor households in 1988 (Mincy et al. 1990).

Another measurement strategy focuses on the geographic concentration of the poor. Using the U.S. Bureau of the Census tract-level definitions of local poverty areas, Reischauer (1987) reported that among the population living in such poverty areas, the central cities housed over half in 1985, up from one-third in 1972. Jargowsky (1997) documented that along with the growth of poverty populations in metropolitan areas, the number of high poverty areas (defined as census tracts containing at least 40 percent poor people) more than doubled between 1970 and 1990. The number of African-Americans living in high-poverty areas, mostly segregated ghettos, climbed from 2.4 million to 4.2 million in that period, far outpacing other minority groups. By 1990, 34 percent of poor African-Americans in metropolitan areas resided in high-poverty census tracts (see also Kasarda 1993).

Massey and Denton (1993) present an analysis and simulations that lead them to conclude that concentrated poverty can be explained largely by two basic factors: the degree of spatial segregation of a racial group and the group's overall poverty rate. Their analysis and conclusion sparked heated debates over racial versus economic segregation explanations (Jargowsky 1997).

As was noted above, the concept of the underclass typically is considered to entail more than poverty. It also is posited to incorporate geographically concentrated behavioral characteristics that conflict with mainstream values: joblessness, out-of-wedlock births, welfare dependency, dropping out of school, drug abuse, and illicit activities.

While considerable debate continues to surround definitions and even the existence of the underclass, attempts have been made to measure its size by using aggregated "behavioral" indicators derived from census tract data. Ricketts and Sawhill (1988) measured the underclass as people living in neighborhoods whose residents in 1980 exhibited disproportionately high rates of school

dropout, joblessness, female-headed families, and welfare dependency. Using a composite definition in which tracts must fall at least one standard deviation above the national mean on all four characteristics, they found that approximately 2.5 million people lived in those tracts in 1980 and that those tracts were disproportionately located in major cities in the Northeast and Midwest. They reported that in underclass tracts, on average, 63 percent of the resident adults had less than a high school education, 60 percent of the families with children were headed by women, 56 percent of the adult men were not regularly employed, and 34 percent of the households received public assistance. Their research also revealed that although the total poverty population grew only 8 percent between 1970 and 1980, the number of people living in underclass areas grew 230 percent, from 752,000 to 2,484,000.

Mincy and Wiener (1993) and Kasarda (1993) updated Ricketts and Sawhill's analysis by using 1990 census tract data. Both found that the number and concentration of persons living in tracts with disproportionately high rates of problem attributes continued to rise in the 1980s, although not nearly as much as it did in the 1970s.

These location-based aggregate measures of underclass populations have been criticized on the grounds that aside from race, most urban census tracts are quite heterogeneous along economic and social dimensions. Jencks (1989; Jencks and Peterson 1991), for example, observes that with the exception of tracts made up of public housing projects, there is considerable diversity in residents' income and education levels, joblessness, and public assistance reciprocity in even the poorest urban neighborhoods. Conversely, considerable numbers of urban residents who are poor, jobless, and welfare-dependent live in census tracts where fewer than 20 percent of the families fall below the poverty line.

Nevertheless, while most scholars concur that behaviors linked to underclass definitions and measurements are found throughout society, it is the concentration of these behaviors in economically declining inner-city areas that is said to distinguish the underclass from previously impoverished urban subgroups. Geographic concentration is argued to magnify social problems and accelerate their spread to *nearby households* through social

contagion, peer pressure, and imitative behavior (Wilson 1987, 1996). The members of economically stable households selectively flee the neighborhood to avoid these problems. Left behind in increasingly isolated concentrations are those with the least to offer in terms of marketable skills, role models, and familial stability. The result is a spiral of negative social and economic outcomes for those neighborhoods and the households that remain.

Incorporating the effects of neighborhoods and social transmission processes means that the future research agenda on the urban underclass will be qualitative as well as quantitative in approach. Ethnographic studies of underclass neighborhoods, family structures, and individual behaviors will complement growing numbers of surveys on and sophisticated statistical analyses of the persistence and intergenerational transfer of urban poverty (see Anderson 1990, 1994; Furstenberg et al. 1999). Additionally more comparative studies will assess similarities to and differences from the American case in European, Latin American, and Asian cities. The root of this work stretches deep, building on classic culture of poverty (Lewis 1966) and social and economic marginalization theses (Clark 1965).

(SEE ALSO: *Cities; Community; Poverty; Segregation and Desegregation; Urbanization; Urban Sociology*)

## REFERENCES

- Anderson, Elijah 1990 *Street Wise: Race, Class, and Change in an Urban Community*. Chicago: University of Chicago Press.
- 1994 “The Code on the Streets.” *Atlantic Monthly*, May, pp. 81–94.
- Auletta, Ken 1982 *The Underclass*. New York: Random House.
- Bane, Mary Jo, and David Ellwood 1986 “Slipping into and out of Poverty: The Dynamics of Spells.” *Journal of Human Resources* 21:1–23.
- Clark, Kenneth B. 1965 *Dark Ghetto: Dilemmas of Social Power*. New York: Harper & Row.
- Duncan, G. J., R. D. Coe, and M. S. Hill 1984 In *Years of Poverty, Years of Plenty*. Ann Arbor: Institute of Social Research, University of Michigan.
- Furstenberg, Frank F., Jr., Thomas D Cook, Jacquelynne Eccles, Glen H. Elder, Jr., and Arnold Sameroff 1999 *Managing to Make It: Urban Families an Adolescent* *Success*. Chicago and London: University of Chicago Press.
- Gans, Herbert J. 1990 “Deconstructing the Underclass: The Term’s Danger as a Planning Concept.” *Journal of the American Planning Association*, pp. 271–277.
- Glasgow, Douglas G. 1980 *The Black Underclass: Poverty, Unemployment, and Entrapment of Ghetto Youth*. San Francisco: Jossey-Bass.
- Gottschalk, P., and S. Danziger 1987 Testimony on poverty, hunger, and the welfare system, August 5, 1986. *Hearing before the Select Committee on Hunger, House of Representatives*, 99<sup>th</sup> Cong., 2nd Sess., Ser. No. 23. Washington, D.C.: U.S. Government Printing Office.
- Hughes, Mark Alan 1993 *Over the Horizon: Jobs in the Suburbs of Major Metropolitan Areas*. Philadelphia: Public/Private Ventures.
- Jargowsky, Paul A. 1997 *Poverty and Place: Ghettos, Barrios, and the American City*. New York: Russell Sage Foundation.
- Jencks, Christopher 1989 “What Is the Underclass—and Is It Growing?” *Focus* 12:14–31.
- , and Paul Peterson, eds. 1991 *The Urban Underclass*. Washington, D.C.: Brookings Institution.
- Kasarda, John D. 1985 “Urban Change and Minority Opportunities.” In Paul Peterson, ed., *The New Urban Reality*. Washington, D.C.: Brookings Institution.
- 1989 “Urban Industrial Transition and the Underclass.” *Annals of the American Academy of Political and Social Sciences* 501:26–47.
- 1993 “Inner City Concentrated Poverty and Neighborhood Distress: 1970 to 1990.” *Housing Policy Debate* 4(3):253–302.
- , and Kwok-Fai Ting 1996 “Joblessness and Poverty in America’s Central Cities: Causes and Policy Prescriptions.” *Housing Policy Debate* 7(2):387–419.
- Katz, Michael 1989 *The Undeserving Poor: From the War on Poverty to the War on Welfare*. New York: Pantheon.
- Levy, Frank 1977 “How Big Is the American Underclass?” Washington, D.C.: Urban Institute.
- Lewis, Oscar 1966 *La Vida: A Puerto Rican Family in the Culture of Poverty—San Juan and New York*. New York: Random House.
- Magnet, Myron 1993 *The Dream and the Nightmare: The Sixties’ Legacy to the Underclass*. New York: Morrow.
- Massey, Douglas S., and Nancy Denton 1993 *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- McGahay, R. 1982 “Poverty’s Voguish Stigma.” *New York Times*, March 12:29.

- Mead, Lawrence M. 1988 "The Hidden Jobs Debate." *Public Interest*, spring: 40–59.
- Mincy, Ronald B., Isabel V. Sawhill, and Douglas A. Wolf 1990 "The Underclass: Definition and Measurement." *Science* 248:450–453.
- , and Susan J. Wiener 1993 *The Under Class in the 1980s: Changing Concepts, Changing Realty*. Washington, D.C.: Urban Institute.
- Murray, Charles A. 1984. *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books
- Myrdal, Gunnar 1962 *Challenge to Affluence*. New York: Pantheon.
- Nathan, Richard P. 1987 "Will the Underclass Always Be with Us?" *Society* 24:57–62.
- Reischauer, Robert D. 1987 *The Geographic Concentration of Poverty: What Do We Know?* Washington, D.C.: Brookings Institution.
- Ricketts, Erol, and Isabel Sawhill 1988 "Defining and Measuring the Underclass." *Journal of Policy Analysis and Management* 7:316–325.
- Sjoquist, David 1990 "Concepts, Measurements, and Analysis of the Underclass: A Review of the Literature." Atlanta: Georgia State University, typescript.
- Wilson, William Julius 1996 *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.
- 1991 "Studying Inner-City Social Dislocations." *American Sociological Review* 56:1–14.
- 1987 *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

JOHN D. KASARDA

## UTILITY THEORY

See Decision-Making Theory and Research; Rational Choice Theory.

## UTOPIAN ANALYSIS AND DESIGN

**NOTE:** Although the following article has not been revised for this edition of the Encyclopedia, the substantive coverage is currently appropriate. The editors have provided a list of recent works at the end of the article to facilitate research and exploration of the topic.

"From the time of its first discovery, the island of King Utopus has been shrouded in ambiguity,

and no latter-day scholars should presume to dispel the fog, polluting utopia's natural environment with an excess of clarity and definition" (Manuel and Manuel 1979, p. 5).

But this ambiguity extends well beyond simple obscurity or murkiness; it reaches to unqualified contradiction. Many utopian visionaries have been denounced for their meticulous delineation of details as they constructed models of social worlds bearing no resemblance to existing, potential, or possible reality. Utopias, it would seem, suffer from the twin infirmities of ambiguity and excessive efforts to achieve clarity and definition. Our dictionaries tell us they are, on the one hand, ideally perfect places but, on the other hand, are simply impractical thought or theory. Utopians are customarily viewed as zealous but quixotic reformers. The books in which they describe their societies may be praised as fascinating, fanciful literature but not as scientific tomes.

It is quite possible as well as reasonable to view utopians as model builders. Models are quite different objects from what is being modeled and have properties not shared by their counterparts. "The aim of a model is precisely not to reproduce reality in all its complexity. It is, rather, to capture in a vivid, often formal way what is essential to understanding some aspect of its structure or behavior" (Weizenbaum 1976, pp. 149–150).

One occupational disability of model builders everywhere is a sort of pathological obsession with a single element, or at most a strictly circumscribed set of elements, of reality, along with an unwavering refusal to examine the larger milieu in which they are found.

In Sir Thomas More's *Utopia* (1655), a central value or societal goal is the concept of economic equality; but this does not include the notion of social equality. There exists in Utopia a large underclass of slaves who are assigned the more distasteful but necessary tasks of the society. This class is composed of war prisoners (More's society is not free of war), persons born into slavery (it is not free of slavery), condemned criminals from other countries who are purchased from foreign slave markets (crime has not been eliminated), and working-class foreigners (class distinctions persist) who volunteer for slavery in Utopia rather than suffer the unpleasant conditions in their home countries (ethnic and immigration difficulties con-

tinue to exist). All able-bodied persons in Utopia become part of its work force—slaves, male nonslaves, and even women! This is seen as an enormous augmentation of the work force. Within each household, however, male dominance prevails. Households are under the authority of the oldest free male. Women are specifically designated as “subordinate” to their husbands, as children are to their parents and younger people generally are to their elders. In Utopia, the applicability of equality is severely restricted.

In discussing utopias it is important to distinguish between analytic and design models. Analytic models purport to be summaries of existing empirical reality; design models are summaries or sketches of future, past, or alternative societies, social structures, or worlds.

Characteristically, utopian literature contains a critique of existing society along with a model of a different one. Frequently the design model incorporates a more or less indirect critique of an existing state of affairs. Plato's *Republic* (1941), the work that seems to have been the prototype of More's *Utopia*, was greatly influenced by the social conditions observed and experienced by Plato. He saw the Athens in which he lived as a very corrupt democracy and felt that in such a system politicians inevitably pandered to mobs. If the mob insisted upon venal demands, politicians found it necessary to agree with them or lose their own positions. Reform, he felt, was not possible in a corrupt society. In the *Republic* Socrates, voicing Plato's sentiments, concludes that “the multitude can never be philosophical. Accordingly, it is bound to disapprove of all who pursue wisdom; and so also, of course, are those individuals who associate with the mob and set their hearts on pleasing it” (1941, p. 201).

Interestingly, it has been suggested that Plato's hostility to democracy was, at least to some extent, shaped by his economic and social background. Members of his family were large landholders who, along with others in a similar position, saw the rise of commerce as a threat to their economic positions. Democratic government undermined their political preeminence, as did militant foreign policies. They had a great deal to lose through war because they were subject to heavy war taxes. Moreover, some had had their lands ravaged by Spartans during the Peloponnesian War; others

had retreated behind the walls of Athens. These conservative elements were not above attempting to subvert the democratic system (Klosko 1986, p. 10).

In any event, Plato's utopia is clearly elitist in nature. For a variety of reasons most utopian schemes seem to be controlled by elites of some sort. As one writer explains it:

*They begin with the proposition that things are bad; things must become better, perhaps perfect here on earth; things will not improve by themselves; a plan must be developed and carried out; this implies the existence of an enlightened individual, or a few, who will think and act in a way that many by themselves cannot think and act.* (Brinton 1965, p. 50)

For Plato, the elites were what he called philosophers. In a sense these were the theoreticians or model makers. The problem he saw was converting their models—their ideal worlds—into reality. Plato was very realistic about this matter of convertibility. He has Socrates ask, “Is it not in the nature of things that action should come less close to truth than thought?” (1941, p. 178). He is, however, concerned about trying to come as close as possible to having the real world correspond to the ideal one. The solution? To have philosophers become rulers or to have rulers become philosophers. In either case enormous, if not complete, power is to be held by a caste of elites.

In effect, social inequality is found even in the work of the triumvirate usually referred to as the “utopian socialists”: Claude Henri de Rouvroy de Saint Simon (1760–1825), Charles Fourier (1772–1837), and Robert Owen (1771–1858).

In his early work Saint Simon's elites were scientists, but later he tended to subordinate them or at least to keep them on a par with industrial chiefs. He evaded the problem of social equality by saying that each member of society would be paid in accordance with his or her “investment.” This referred to the contribution each made to the productive process. Since different people had different talents, these contributions would differ. Some people's contributions would be more important than others', and accordingly those people would be paid more. But although the rewards of different people would differ, there would not be wide discrepancies between the rewards of the

lowest- and highest-paid workers (Manuel and Manuel 1979, pp. 590–614).

Unlike Saint Simon, who never wrote a detailed description of a utopian society, Charles Fourier wrote thousands of pages of detailed descriptions of his “Phalanx,” including architectural specifications, work schedules and countless other details. The Phalanx was to be organized essentially as a shareholding corporation. Members were free to buy as many shares as they wished or could afford. Fourier stressed the fact that in his utopia there would be three social classes: the rich, the poor, and the middle. The condition of the poor would be enormously better than their condition in existing society, but the rich or upper class would be entitled to more lavish living quarters, more sumptuous food, and, in general, a more luxurious life-style than the others. During the last fifteen years or so of his life, most of Fourier’s efforts were devoted to the search for a wealthy person to subsidize a trial of his Phalanx (Beecher 1986).

Robert Owen insisted on what he regarded to be complete equality. Conceding that people were born with differing abilities, he contended that these abilities were provided by God and should not be the basis for differential rewards. Nevertheless, as a self-made man who became extremely successful and managed the most important cotton-spinning factory in Britain, he never seemed to lose the self-assurance that he knew best how to manage a community and that all members would understand the wisdom of his decisions. He has been characterized as a benevolent autocrat who acted somewhat like a military commander who has little direct contact with his troops (Cole 1969; Manuel and Manuel 1979, pp. 676–693).

In the United States, the most widely read utopian novel based on the assumption of absolute economic equality is undoubtedly Edward Bellamy’s *Looking Backward* (1887). Bellamy (1850–1898), influenced by the development of the large economic trusts in the United States, postulated that by the year 2000 only one enormous trust would remain: the United States government. He went to great pains to make it clear that his utopia was devoid of Marxist or other European influences. The principle of income or reward on which it was based was neither “From each according to his investment or product” nor the classic

“From each according to his ability, to each according to his need,” although it was much closer to the latter than to the former.

In Bellamy’s vision of the United States in the year 2000, each person received an equal share of the total national product. In effect, every inhabitant received a credit card showing his or her share of the product. The share could be spent in any manner. If too many individuals decided to buy a particular product, the price of that product would be raised. The point, however, is that people were entitled to a share of the national product *not* on the basis of their individual productivity but simply because they existed as human beings. In some telling passages Bellamy’s characters observe that members of families do not deny food or other needs to other family members because they have been unproductive. In effect, the entire country (and, presumably, ultimately the entire world) would resemble our more primitive notion of one family.

Bellamy’s work received widespread attention throughout the world. In England, William Morris (1834–1896) objected strenuously to the centralized control and bureaucratic form of organization in *Looking Backward*. Morris wrote his own utopian novel, *News from Nowhere* (1866). Unlike Bellamy’s utopia, which came into being through a process of evolution, a violent revolution has occurred in *Nowhere*. London has become a series of relatively small villages separated by flowers and wooded areas. There is no centralized government—no government at all—as we normally understand it. With the end of private property and domestic arrangements in which women are essentially the property of men, the underlying reasons for criminal behavior have been eliminated. Random acts of violence are regarded as transitory diseases and are dealt with by nurses and doctors rather than by jailers.

It has been argued that Morris was essentially an anarchist theorist, although Morris himself vigorously objected to such characterization of his work. It has been suggested that anarchism has two major forms: collectivist and individualist. Morris is seen as essentially a collectivist anarchist, although not an anarchosyndicalist—the form that stresses trade-union activity. He ridiculed conventional forms of individualism. Anarchism itself is defined as a social theory that advocates a commu-

nity-centered life with great amounts of personal liberty. It opposes coercion of its population (Sargent 1990, pp. 61–64).

Other commentators see *News from Nowhere* as an effort by Morris to present his arguments against anarchism (Holzman 1990, p. 99). It seems clear that his work does not fit neatly into any prefabricated ideological cubbyhole. Morris cherished aesthetic over intellectual values (he was an architect, artist, poet, designer, and craftsman). When one of his characters in *News from Nowhere* is asked how labor is rewarded, the reply is quite predictable: it is *not* rewarded. Work has become a pleasure—not a hardship. Each person does what he or she can do best; the quandary of extrinsic motivation has substantially disappeared.

Motivation, however, is the central concern in B. F. Skinner's *Walden Two* (1948). Burrhus Frederic Skinner (1904–1990) was a professional psychologist whose utopia was a product of his interest in behavioral engineering. His ideal community has been described as one of means rather than of ends—one in which technique has been elevated to utopian status (Kumar 1987, p. 349).

This is not completely accurate. It does capture the essence of how Skinner himself saw his utopia, but it omits direct consideration of the implicit values held by its designer.

Skinner himself was unquestionably a well-motivated, humanistic scientist, but he neglected his customary penetrating analysis when approaching the area of values held by the boss scientist. At one point in *Walden Two*, however, he does seem to have some insight into this difficulty. Frazier, the founder of the community, voices the unspoken criticism of one of the other characters by pointing to his own insensitivity to the effect he has on others, except when the effect is calculated; his lack of the personal warmth responsible in part for the success of the community; the ulterior and devious nature of his own motives. He then cries out, “But God *damn* it Burris . . . can't you see? *I'm-not-a-product-of-Walden-Two!*” (Skinner 1948, p. 233).

Economic and basic social equality exist in this community, but effective control is exercised through the built-in reinforcement techniques of its designer. When Frazier is challenged on this by one of the characters who observes that Frazier,

looking at the world from the middle of the twentieth century, assumes he knows the best course for humanity forever, Frazier essentially agrees. His defense is that the techniques of behavioral engineering currently exist (and presumably will continue to be used), but they are in the wrong hands—those of charlatans, salespeople, ward heelers, bullies, cheats, educators, priests, and others. Ultimately, Skinner's designer insists, human beings are never free—their behavior is determined by prior conditioning in the society in which they were raised. The belief in their own freedom is what allows human beings unwittingly to become conditioned by reinforcers in their existing environments.

Thus, in effect, *Walden Two* achieves its effects by changing the psychological characteristics of its inhabitants through environmental modification. Its final form is presumably an experimental question. The queries are simple enough and are stated explicitly at one point: What is the best behavior for the individual as far as the group is concerned? How can an individual be induced to behave in that way? The answer presumably can change over time, on the basis of experimental experience. The entire edifice would seem to depend upon the continuing moral superiority of the reinforcement designers over the charlatans they replace.

Quite a different sort of utopia has been proposed by the philosopher Robert Nozick, who outlines what he calls the framework for a utopia. In a word (or two), this framework is equivalent to what Nozick calls the minimal state (Nozick 1974, pp. 297–334). This is a state “limited to the narrow functions of protection against force, theft, fraud, enforcement of contracts, and so on . . . any more extensive state will violate persons' rights not to be forced to do certain things and is unjustified . . .” (Nozick 1974, p. ix).

Nozick is not concerned with modifying behavior or specifying social structures beyond this minimum state. He begins with the assumption that individual persons have certain rights that may never be violated by any other person or the state. These include the right *not* to be killed or attacked if you are not doing any harm; *not* to be coerced or imprisoned; *not* to be limited in the use of your property if that use does not violate the rights of others.

In arguing for a minimal state, Nozick, on the one hand, is arguing against anarchism (in which there is no state at all). On the other hand, he argues against all forms of the welfare state (in which some people with excessive wealth may be required to surrender some of their property to help others who are less fortunate) (Paul 1981).

As Nozick sees it, rights define a moral boundary around individual persons. The sanctity of this boundary takes priority over all other possible goals. Thus, it becomes readily understandable why he feels that nonvoluntary redistribution of income is morally indefensible:

*It is an extraordinary but apparent consequence of this view that for a government to tax each of its able-bodied citizens five dollars a year to support cripples and orphans would violate the rights of the able-bodied and would be morally impermissible, whereas to refrain from taxation even if it meant allowing the cripples and orphans to starve to death would be the morally required governmental policy.* (Scheffler 1981, p. 151)

Here again we see the clash of values that lie at the heart of utopian schemes and their critics. A serious and widely discussed effort to resolve these clashes was made late in the twentieth century by another social philosopher, John Rawls. *A Theory of Justice* (Rawls 1971) was not a utopian novel but a meticulously argued tome that has been compared with John Locke's *Second Treatise of Civil Government* and John Stuart Mill's *On Liberty*. The central question confronting his work has been expressed thus: "Is it possible to satisfy the legitimate 'leftist', 'socialist' critics of Western capitalism within a broadly liberal, capitalist and democratic framework?" (Goldman 1980, p. 431).

Unfortunately, Rawls has found himself increasingly caught between attacks from both the left and the right. The left feels he has not gone far enough in constraining property rights; the right feels he places too great an emphasis upon the value of equality, especially at the expense of the right to property (Goldman 1980, pp. 431–432).

A central point argued by Rawls is that there is no injustice if greater benefits are earned by a few, provided the situation of people not so fortunate is thereby improved (Rawls 1971, pp. 14–15).

As one commentator expressed it, for Rawls equality comes first. Goods are to be distributed equally unless it can be shown that an unequal distribution is to the advantage of the least advantaged. This would be a "just" distribution (Schaar 1980). One might add, parenthetically, that this justice would depend substantially upon the nature of the existing social and economic arrangements under which this inequality occurs. Would a different set of arrangements allow greater equality? For example, is capital available only through private sources? Would public sources serve similar ends with less inequality?

The central issue for utopian analysts from Plato through twentieth-century philosophers is how one constructs a "just" society. But there is no single definition of "just"; it all depends on what you consider to be important. Are you concerned exclusively with yourself? your immediate family? others in your community? in your country? in the world?

And so it is that utopian analysis and design ultimately begin with an implicit, if not explicit, value orientation. One school of thought begins with an overwhelming belief that elites of one sort or another must be favored in the new society. Elite status may be gained through existing wealth, birth, talent, skill, intelligence, or physical strength. Another school begins with what is, broadly speaking, the concept of equality. Here the implicit notion is not unlike Western ideas of the family: to each equally, irrespective of either productivity or need. Between these two polar positions lie a range of intermediate proposals that may provide greater amounts of compensation based upon some definition of need or elite status. In turn, compensation may or may not be linked directly to political or other forms of power.

Issues relating to the nation-state (its form, its powers, and even its very existence), ethnicity, and inequality became acute in the final decade of the twentieth century. Ethnic groups throughout the world grew militant in their demands for their own national entities. Many saw this as a path to a solution for their own problems of inequality. With the apparent easing, if not the elimination, of Cold War tensions between the Soviet Union and the United States, widespread controversies began relative to the shape of a "new world order." This posed unprecedented challenges to utopian thought.

To deal with these challenges, social scientists, as well as imaginative novelists and others, were confronted with the task of integrating value configurations, social structures, and psychological sets on levels that may well make all previous efforts at utopian analysis and design resemble the stumbling steps of a child just learning to walk.

(SEE ALSO: *Equity Theory*; *Social Philosophy*)

## REFERENCES

- Beecher, Jonathan 1986 *Charles Fourier: The Visionary and His World*. Berkeley: University of California Press.
- Brinton, Crane 1965 "Utopia and Democracy." In Frank E. Manuel, ed., *Utopias and Utopian Thought*. Boston: Beacon Press.
- Cole, Margaret 1969 *Robert Owen of New Lanark 1771-1858*. New York: August M. Kelley.
- Gil, Efraim 1996 "The Individual within the Collective: A New Perspective." *Journal of Rural Cooperation* 24:5-15.
- Goldman, Alan H. 1980 "Responses to Rawls from the Political Right." In H. Gene Blocker and Elizabeth H. Smith, eds., *John Rawls' Theory of Social Justice*. Athens: Ohio University Press.
- Hacohen, Malachi-Haim 1996 "Karl Popper in Exile: The Viennese Progressive Imagination and the Making of The Open Society." *Philosophy of the Social Sciences* 26:452-492.
- Hodgson, Geoffrey M. 1995 "The Political Economy of Utopia." *Review of Social Economy* 53:195-213.
- Holzman, Michael 1990 "The Encouragement and Warning of History: William Morris's *A Dream of John Ball*." In Florence S. Boos and Carole G. Silver, eds., *Socialism and the Literary Artistry of William Morris*. Columbia: University of Missouri Press.
- Klosko, George 1986 *The Development of Plato's Political Theory*. New York: Methuen.
- Kumar, Krishan 1987 *Utopia and Anti-Utopia in Modern Times*. New York: Basil Blackwell.
- Lowy, Michael 1997 "The Romantic Utopia of Walter Benjamin" (L'Utopie romantique de Walter Benjamin) *Raison Presente* 121:19-27.
- Maler, Henri. 1998. "An Pocyphal Testament: Socialism, Utopian and Scientific." *Science and Society* 62:48-61.
- Manuel, Frank E., and Fritzie P. Manuel 1979 *Utopian Thought in the Western World*. Cambridge, Mass.: Harvard University Press.
- Martensson, Bertil 1991 "The Paradoxes of Utopia: A Study in Utopian Rationalism." *Philosophy of the Social Sciences* 21:476-514.
- More, Sir Thomas (1516) 1965 *Utopia*. Paul Turner, trans. London: Penguin.
- Morris, William 1966 *News from Nowhere*. In *The Collected Works of William Morris*, vol. 16, pp. 3-211. New York: Russell and Russell.
- Nozick, Robert 1974 *Anarchy, State and Utopia*. New York: Basic Books.
- Oyzerman, Teodor Il'ich. 1998 "Marxism and Utopianism. Marxisms's Overcoming of Utopianism as an Unfinished Historical Process" (Marksizm i utopizizm. Preodolenie marksizmom utopizma kak nezavershennyi istoricheskiy protsess) *Svobodnya-Mysl* 2:76-83.
- Paul, Jeffrey (ed.) 1981 *Reading Nozick*. Totowa, N.J.: Rowan and Littlefield.
- Plato 1941 *The Republic of Plato*, Francis MacDonald Cornford, trans. and ed. New York and London: Oxford University Press.
- Prat, Jean-Louis 1995 "Utopian Utilitarianism" (L'Utilitarisme utopique) *Revue du MAUSS* 6:53-60.
- Rawls, John 1971 *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Sargent, Lyman Tower 1990 "William Morris and the Anarchist Tradition." In F. S. Boos and C. G. Silver, eds., *Socialism and the Literary Artistry of William Morris*. Columbia: University of Missouri Press.
- Schaar, John H. 1980 "Equality of Opportunity and the Just Society." In H. G. Blocker and E. H. Smith, eds., *John Rawls' Theory of Social Justice*. Athens: Ohio University Press.
- Scheffler, Samuel 1981 "Natural Rights, Equality and the Minimal State." In Jeffrey Paul, ed., *Reading Nozick*. Totowa, N.J.: Rowan and Littlefield.
- Skinner, B. F. 1948 *Walden Two*. New York: Macmillan.

ROBERT BOGUSLAW



# V

---

## VALIDITY

In the simplest sense, a measure is said to be valid to the degree that it measures what it is hypothesized to measure (Nunnally 1967, p. 75). More precisely, validity has been defined as the degree to which a score derived from a measurement procedure reflects a point on the underlying construct it is hypothesized to reflect (Bohrnstedt 1983). In the most recent *Standards for Educational and Psychological Testing* (American Psychological Association 1985), it is stated that validity “refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from . . . scores.” The emphasis is clear: Validity refers to the degree to which evidence supports the *inferences* drawn from a score rather than the scores or the instruments that produce the scores. Inferences drawn for a given measure with one population may be valid but may not be valid for other measures. As will be shown below, evidence for inferences about validity can be accumulated in a variety of ways. In spite of this variety, validity is a unitary concept. The varied types of inferential evidence relate to the validity of a particular measure under investigation.

Several important points related to validity should be noted:

1. Validity is a matter of degree rather than an all-or-none matter (Nunnally 1967, p. 75; Messick 1989).
2. Since the constructs of interest in sociology (normlessness, religiosity, economic conservatism, etc.) generally are not ame-

nable to direct observation, validity can be ascertained only indirectly.

3. Validation is a dynamic process; the evidence for or against the validity of the inferences that can be drawn from a measure may change with accumulating evidence. Validity in this sense is always a continuing and evolving matter rather than something that is fixed once and for all (Messick 1989).
4. Validity is the sine qua non of measurement; without it, measurement is meaningless.

In spite of the clear importance of validity in making defensible inferences about the reasonableness of theoretical formulations, the construct more often than not is given little more than lip service in sociological research. Measures are assumed to be valid because they “look valid,” not because they have been evaluated as a way to get statistical estimates of validity. In this article, the different meanings of validity are introduced and methods for estimating the various types of validity are discussed.

## TYPES OF VALIDITY

The *Standards* produced jointly by the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education distinguish between and among three types of evidence related to validity: (1) *criterion-related*, (2) *content*, and (3)

*construct evidence* (American Psychological Association 1985).

**Criterion-Related Evidence for Validity.** Criterion-related evidence for validity is assessed by the correlation between a measure and a criterion variable of interest. The criterion varies with the purpose of the researcher and/or the client for the research. Thus, in a study to determine the effect of early childhood education, a criterion of interest might be how well children perform on a standardized reading test at the end of the third grade. In a study for an industrial client, it might be the number of years it takes to reach a certain job level. The question that is always asked when one is accumulating evidence for criterion-related validity is: How accurately can the criterion be predicted from the scores on a measure? (American Psychological Association 1985).

Since the criterion variable may be one that exists in the present or one that a researcher may want to predict in the future, evidence for criterion-related validity is classified into two major types: predictive and concurrent.

Evidence for *predictive validity* is assessed by examining the future standing on a criterion variable as predicted from the present standing on a measure of interest. For example, if one constructs a measure of work orientation, evidence of its predictive validity for job performance might be ascertained by administering that measure to a group of new hires and correlating it with a criterion of success (supervisors' ratings, regular advances within the organization, etc.) at a later point in time. The evidence for the validity of a measure is not limited to a single criterion. There are as many validities as there are criterion variables to be predicted from that measure. The preceding example makes this clear. In addition, the example shows that the evidence for the validity of a measure varies depending on the time at which the criterion is assessed. Generally, the closer in time the measure and the criterion are assessed, the higher the validity, but this is not always true.

Evidence for *concurrent validity* is assessed by correlating a measure and a criterion of interest at the *same* point in time. A measure of the concurrent validity of a measure of religious belief, for example, is its correlation with concurrent attendance at religious services. Just as is the case for

predictive validity, there are as many concurrent validities as there are criteria to be explained; there is no single concurrent validity for a measure.

Concurrent validation also can be evaluated by correlating a measure of *X* with extant measures of *X*, for instance, correlating one measure of self-esteem with a second one. It is assumed that the two measures reflect the same underlying construct. Two measures may both be labeled self-esteem, but if one contains items that deal with one's social competence and the other contains items that deal with how one feels and evaluates oneself, it will not be surprising to find no more than a modest correlation between the two.

Evidence for validity based on concurrent studies may not square with evidence for validity based on predictive studies. For example, a measure of an attitude toward a political issue may correlate highly in August in terms of which political party one *believes* one will vote for in November but may correlate rather poorly with the *actual* vote in November.

Many of the constructs of interest to sociologists do not have criteria against which the validity of a measure can be ascertained easily. When they do, the criteria may be so poorly measured that the validity coefficients are badly attenuated by measurement error. For these reasons, sociological researchers have rarely computed criterion-related validities.

**Content Validity.** One can imagine a *domain of meaning* that a construct is intended to measure. *Content validity* provides evidence for the degree to which one has representatively sampled from that domain of meaning. (Bohrnstedt 1983). One also can think of a domain as having various facets (Guttman 1959), and just as one can use stratification to obtain a sample of persons, one can use stratification principles to improve the evidence for content validity.

While content validity has received close attention in the construction of achievement and proficiency measures psychology and educational psychology, it usually has been ignored by sociologists. Many sociological researchers have instead been satisfied to construct a few items on an ad hoc, one-shot basis in the apparent belief that they are measuring what they intended to measure. In

fact, the construction of good measures is a tedious, arduous, and time-consuming task.

Because domains of interest cannot be enumerated in the same way that a population of persons or objects can, the task of assuring the content validity of one's measures is less rigorous than one would hope. While an educational psychologist can sample four-, five-, or six-letter words in constructing a spelling test, no such clear criteria exist for a sociologist who engages in social measurement. However, some guidelines can be provided. First, the researcher should search the literature carefully to determine how various authors have used the concept that is to be measured. There are several excellent handbooks that summarize social measures in use, including Robinson and Shaver's *Measures of Social Psychological Attitudes* (1973); Robinson et al.'s *Measures of Political Attitudes* (1968); Robinson et al.'s *Measures of Occupational Attitudes and Occupational Characteristics* (1969); Shaw and Wright's *Scales for the Measurement of Attitudes* (1967); and Miller's *Handbook of Research Design and Social Measurement* (1977). These volumes not only contain lists of measures but provide existing data on the reliability and validity of those measures. However, since these books are out of date as soon as they go to press, researchers developing their own methods must do additional literature searches. Second, sociological researchers should rely on their own observations and insights and ask whether they yield additional facets to the construct under consideration.

Using these two approaches, one develops sets of items, one to capture each of the various facets or strata within the domain of meaning. There is no simple criterion by which one can judge whether a domain of meaning has been sampled properly. However, a few precautions can be taken to help ensure the representation of the various facets within the domain.

First, the domain can be stratified into its major facets. One first notes the most central meanings of the construct, making certain that the stratification is exhaustive, that is, that all major meaning facets are represented. If a facet appears to involve a complex of meanings, it should be subdivided further into substrata. *The more one refines the strata and substrata the easier it is to construct the items later and the more complete the coverage*

*of meanings associated with the construct will be.* Second, one should write several items or locate several extant indicators to reflect the meanings associated with each stratum and substratum. Third, after the items have been written, they should be tried out on very small samples composed of persons of the type the items will eventually be used with, using cognitive interviewing techniques, in which subjects are asked to "think aloud" as they respond to the items. This technique for the improvement of items, while quite new in survey research, is very useful for improving the validity of items (Sudman et al. 1995). For example, Levine et al. (1997) have shown how cognitive interviewing helped in the improvement of school staffing resources, as did Levine (1996) in describing the development of background questionnaires for use with the large-scale cognitive assessments. Fourth, after the items have been refined through the use of cognitive laboratory techniques, the newly developed items should be field-tested on a sample similar to that with which one intends to examine the main research questions. The field-test sample should be large enough to examine whether the items are operating as planned vis-à-vis the constructs they are putatively measuring, using multivariate tools such as confirmatory factor analysis (Joreskog 1969) and item response theory methods (Hambleton and Swaminathan 1985).

Finally, after the items are developed, the main study should employ a sampling design that takes into account the characteristics of the population about which generalizations are to be made (ethnicity, gender, region of country, etc.). The study also should be large enough to generate stable parameter estimates when one is using multivariate techniques such as multiple regression (Bohrnstedt and Knoke 1988) and structural equation techniques (Bollen 1989).

It can be argued that what the *Standards* call content validity is not a separate method for assessing validity. Instead, it is a set of procedures for sampling content domains that, if followed, can help provide evidence for *construct validity* (see the discussion of construct validity below). Messick (1989), in a similar stance, states that so-called content validity does not meet the definition of validity given above, since it does not deal directly with scores or their interpretation. This position can be better understood in the context of construct validity.

**Construct Validity.** The 1974 *Standards* state: “A construct is . . . a theoretical idea developed to explain and to organize some aspects of existing knowledge. . . It is a dimension understood or inferred from its network of interrelationships” (American Psychological Association 1985). The *Standards* further indicate that in developing evidence for construct validity,

*the investigator begins by formulating hypotheses about the characteristics of those who have high scores on the [measure] in contrast to those who have low scores. Taken together, such hypotheses form at least a tentative theory about the nature of the construct the [measure] is believed to be measuring.*

*Such hypotheses or theoretical formulations lead to certain predictions about how people. . . will behave. . . in certain defined situations. If the investigator’s theory. . . is correct, most predictions should be confirmed. (p. 30)*

The notion of a construct implies hypotheses of two types. First, it implies that items from one stratum within the domain of meaning correlate together because they all reflect the same underlying construct or “true” score. Second, whereas items from one domain may correlate with items from another domain, the implication is that they do so only because the constructs themselves are correlated. Furthermore, it is assumed that there are *hypotheses* about how measures of different domains correlate with one another. To repeat, construct validation involves two types of evidence. The first is evidence for *theoretical validity* (Lord and Novick 1968): an assessment of the relationship between items and an underlying, latent unobserved construct. The second involves evidence that the underlying latent variables correlate as hypothesized. If either or both sets of these hypotheses fail, evidence for construct validation is absent. If one can show evidence for theoretical validity but evidence about the interrelations among those constructs is missing, that suggests that one is not measuring the intended construct or that the theory is wrong or inadequate. The more unconfirmed hypotheses one has involving the constructs, the more one is likely to assume the former rather than the latter.

The discussion above makes clear the close relationship between construct validation and the-

ory validation. To be able to show construct validity assumes that the researcher has a clearly stated set of interrelated hypotheses between important theoretical constructs, which in turn can be measured by sets of indicators. Too often in sociology, one or both of these components are missing.

Campbell (1953, 1956) uses a *multitrait-multimethod matrix*, a useful tool for assessing the construct validity of a set of measures collected using differing methods. Thus, for example, one might collect data using multiple indicators of three constructs, say, prejudice, alienation, and anomie, using three different data collection methods: a face-to-face interview, a telephone interview, and a questionnaire. To the degree that different methods yield the same or a very similar result, the construct demonstrates what Campbell (1954) calls *convergent validity*. Campbell argues that in addition, the constructs must not correlate too highly with each other; that is, to use Campbell and Fiske’s (1959) term, they must also exhibit *discriminant validity*. Measures that meet both criteria provide evidence for construct validity.

### VALIDITY GENERALIZATION

An important issue for work in educational and industrial settings is the degree to which the criterion-related evidence for validity obtained in one setting generalizes to other settings (American Psychological Association 1985). The point is that evidence for the validity of an instrument in one setting in no ways guarantees its validity in any other setting. By contrast, the more evidence there is of consistency of findings across settings that are maximally different, the stronger the evidence for *validity generalization* is.

Evidence for validity generalization generally is garnered in one of two ways. The usual way is simply to do a nonquantitative review of the relevant literature; then, on the basis of that review, a conclusion about the generalizability of the measure across a variety of settings is made. More recently, however, meta-analytic techniques (Hedges and Olkin 1985) have been employed to provide quantitative evidence for validity generalization.

Variables that may affect validity generalization include the particular criterion measure used, the sample to which the instrument is adminis-

tered, the time period during which the instrument was used, and the setting in which the assessment is done.

**Differential predication.** In using a measure in different demographic groups that differ in experience or that have received different treatments (e.g., different instructional programs), the possibility exists that the relationship between the criterion measure and the predictor will vary across groups. To the degree that this is true, a measure is said to display *differential prediction*.

Closely related is the notion of *predictive bias*. While there is some dispute about the best definition, the most commonly accepted definition states that predictive bias exists if different regression equations are needed for different groups and if predictions result in decisions for those groups that are different from the decisions that would be made based on a pooled groups regression analysis (American Psychological Association 1985). Perhaps the best example to differentiate the two concepts is drawn from examining the relationship between education and income. It has been shown that that relationship is stronger for whites than it is for blacks; that is, education differentially predicts income. If education were then used as a basis for selection into jobs at a given income level, education would be said to have a predictive bias against blacks because they would have to have a greater number of years of education to be selected for a given job level compared to whites.

Differential prediction should not be confused with *differential validity*, a term used in the context of job placement and classification. Differential validity refers to the ability of a measure or, more commonly, a battery of measures to differentially predict success or failure in one job compared to another. Thus, the armed services use the battery of subtests in the Armed Services Vocational Aptitude Battery (U.S. Government Printing Office 1989; McLaughlin et al. 1984) in making the initial assignment of enlistees to military occupational specialties.

#### MORE RECENT FORMULATIONS OF VALIDITY

More recent definitions of validity have been even broader than that used in the 1985 *Standards*.

Messick (1989) defines validity as an evaluative judgment about the degree to which “empirical and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on . . . scores or other modes of assessment” (p. 13). For Messick, validity is more than a statement of the existing empirical evidence linking a score to a latent construct; it is also a statement about the evidence for the appropriateness of using and interpreting the scores. While most measurement specialists separate the use of scores from their interpretation, Messick (1989) argues that the value implications and social consequences of testing are inextricably bound to the issue of validity:

*[A] social consequence of testing, such as adverse impact against females in the use of a quantitative test, either stems from a source of test invalidity or a valid property of the construct assessed, or both. In the former case, this adverse consequence bears on the meaning of the test scores and, in the later case, on the meaning of the construct. In both cases, therefore, construct validity binds social consequences to the evidential basis of test interpretation and use.”* (p. 21)

Whether the interpretation and social consequences of the uses of measures become widely adopted (i.e., are adopted in the next edition of the *Standards*) remains to be seen. Messick’s (1989) definition does reinforce, the idea that although there are many facets to and methods for garnering evidence for inferences about validity, it remains a unitary concept; evidence bears on inferences about a single measure or instrument.

#### REFERENCES

- American Psychological Association 1985 *Standards for Educational and Psychological Testing*. Washington, D.C.: American Psychological Association.
- Bohrnstedt, G. W. 1983 “Measurement.” In Rossi, P. H., J. D. Wright, and A. B. Anderson, eds., *Handbook of Survey Research*. New York: Academic Press
- Bohrnstedt, G. W., and D. Knoke 1988, *Statistics for Social Data Analysis*. Itasca, Ill.: F.E. Peacock.
- Bohrnstedt, G. W. 1992 “Reliability.” In E. F. Borgatta (ed.) *Encyclopedia of Sociology*. 1st ed., New York: Macmillan.
- Bollen, K. A. 1989 *Structural Equations with Latent Variables*. New York: Wiley.

- Campbell, D. T. 1953 *A Study of Leadership among Submarine Officers*. Columbus: Ohio State University Research Foundation.
- 1954 "Operational Delineation of What Is Learned' via the Transportation Experiment." *Psychological Review* 61:167-174.
- 1956 *Leadership and Its Effects upon the Group*. Monograph no. 83. Columbus: Ohio State University Bureau of Business Research.
- , and D.W. Fiske 1959 "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81-105.
- Guttman, L. 1959 "A Structural Theory for Intergroup Beliefs and Action." *American Sociological Review* 24:318-328.
- Hambleton, R., and H. Swaminathan 1985 *Item Response Theory: Principles and Applications*. Norwell, Mass.: Kluwer Academic.
- Hedges, L. V., and I. Olkin 1985 *Statistical Methods for Meta-Analysis*. Orlando, Fla.: Academic Press.
- Jöreskog, K. G. 1969 "A general approach to confirmatory maximum likelihood factor analysis." *Psychometrika* 36:409-426.
- Levine, R. 1998 "What Do Cognitive Labs Tell Us about Student Knowledge?" Paper presented at the 28th Annual Conference on Large Scale Assessment sponsored by the Council of Chief State School Officers, Colorado Springs, Colo. Palo Alto, Calif.: American Institutes for Research
- Levine, R., J. Chambers, I. Duenas, and C. Hikido 1997 *Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Surveys (SASS)*. Palo Alto, Calif.: American Institutes for Research
- Lord, F. M., and M. R. Novick 1968 *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- McLaughlin, D. H., P. G. Rossmeissl, L. L. Wise, D. A. Brandt, and M. Wang 1984 *Validation of Current and Alternative Armed Services Vocational Aptitude Battery (ASVAB) Area Composites*. Washington, D.C.: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Messick, S. 1989 "Validity." In L. Linn, ed., *Educational Measurement*, 3rd ed. New York: Macmillan.
- Miller, D. 1977 *Handbook of Research Design and Social Measurement*, 3rd ed. New York: David McKay.
- Nunnally, J. C. 1967 *Psychometric Theory*. New York: McGraw-Hill.
- Robinson, J. P., R. Athanasiou, and K. B. Head 1969 *Measures of Occupational Attitudes and Occupational Characteristics*. Ann Arbor: Institute for Social Research.
- , J. G. Rusk, and K. B. Head 1968 *Measures of Political Attitudes*. Ann Arbor, Mich.: Institute for Social Research.
- Robinson, J. P. and P. R. Shaver 1973 *Measures of Social Psychological Attitudes*. Ann Arbor, Mich.: Institute for Social Research.
- Shaw, M., and J. Wright 1967 *Scales for the Measurement of Attitudes*. New York: McGraw-Hill.
- Sudburn, S., N. Bradburn, and N. Schwarz 1995 *Thinking about Answers, The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- U.S. Government Printing Office 1989 *A Brief Guide: ASVAB for Counselors and Educators*. Washington, D. C.: U.S. Government Printing Office.

GEORGE W. BOHRNSTEDT

## VALUE-FREE ANALYSIS

See Epistemology; Positivism; Scientific Explanation.

## VALUES THEORY AND RESEARCH

The study of values covers a broad multidisciplinary terrain. Different disciplines have pursued this topic with unique orientations to the concept of values. The classic conception of values in anthropology was introduced by Kluckhohn and Strodtbeck (1961). In this view, values answer basic existential questions, helping to provide meaning in people's lives. For example, Kluckhohn and Strodtbeck argue that Americans value individual effort and reward because of their fundamental belief in the inherent goodness of human nature and the capacity of individuals to obtain desired ends. Economists have considered values not in terms of the meaning they provide but as a quality of the objects used in social exchange (Stigler 1950). For economists, objects have value but people have preferences, and those preferences establish hierarchies of goods. It is the goods that have value, with those which are both scarce and highly desirable being the most highly valued.

Sociologists, particularly Parsons, have emphasized a different conception of values (see

Parsons and Shils 1951). In sociology, values are believed to help ease the conflict between individual and collective interests. Values serve an important function by enabling individuals to work together to realize collectively desirable goals. For example, while all the individual members of society may believe that public education is a good idea for themselves, their children, and/or the well-being of society in general, none of them is excited by the prospect of paying taxes to build schools and pay teachers. Even when people believe in the collective good, their private interest (keeping one's money for one's own use) conflicts with the necessities for keeping a society organized. Values such as being socially responsible, showing concern for others, and education encourage people to sidestep their own desires and commit themselves to the more difficult task of social cooperation. As Grube et al. (1994, p. 155) argue, "values play a particularly important role because they are cognitive representations of individual needs and desires, on the one hand, and of societal demands on the other."

Another way to understand the sociological conception of values is to examine when values become vital in social life. They do not matter much when everyone is in full agreement. For example, everyone values breathing over asphyxiation. Even though this value may be of life-and-death importance, it is not a particularly important object of social inquiry because no one disagrees about whether one should hold one's breath. The situation has been quite different with regard to abortion, affirmative action, the death penalty, same-sex marriage, environmental protection, and many other social issues that elicit conflicts in personal values. Values are important to understand when they conflict between individuals, groups, or whole societies. They provide a window through which one can view conflicts and variations within and between societies.

Although many formal definitions of values have been advanced by sociologists, one definition in particular captures the concept's core features well. Smith and Schwartz (1997, p. 80) observe five features:

1. Values are beliefs. But they are not objective, cold ideas. Rather, when values are activated, they become infused with feeling.

2. Values refer to desirable goals (e.g., equality) and to the modes of conduct that promote these goals (e.g., fairness, helpfulness).
3. Values transcend specific actions and situations. Obedience, for example, is relevant at work or in school, in sports or in business, with family, friends or strangers.
4. Values serve as standards to guide the selection or evaluation of behavior, people, and events.
5. Values are ordered by importance relative to one another. The ordered set of values forms a system of value priorities. Cultures and individuals can be characterized by their systems of value priorities.

Smith and Schwartz's conceptualization is consistent with the sociological view that values are abstract concepts, but not so abstract that they cannot motivate behavior. Hence, an important theme of values research has been to assess how well one can predict specific behavior by knowing something about a person's values. If someone claims to believe in protecting the environment, for example, how confidently can one assume that that person recycles, contributes to the Sierra Club, or supports proenvironmental legislation? Below, several empirical efforts to measure the link between values and behavior are discussed. However, some scholars are skeptical that such a link can be drawn (Hechter 1992, 1993).

The definition given above emphasizes the link between values and desired goals. In an earlier discussion, Schwartz (1992, p. 4) argued that values, when defined in this way, reflect three basic requirements of human existence: "needs of individuals as biological organisms, requisites of coordinated social interaction, and survival and welfare needs of groups." By understanding values, one can learn about the needs of both individuals and societies. Sociologists are especially concerned with how values facilitate action toward ends that enhance individual and collective outcomes or are perceived to do so by society's members. Research on values does not presuppose which values are best (social scientists are not preachers) but tries to discover what people believe in and how their beliefs motivate their behavior. A major part of the enterprise is concerned with strategies to measure

values: which ones people hold, how strongly they hold them, how their value priorities compare with those of others, how the value priorities of different groups or societies compare with one another.

Values research has a long and varied history in sociology. Important theoretical and empirical studies of values have been made by Parsons and Shils (1951), Kluckhohn (1951), Williams (1960), Allport et al. (1960), Scott (1965), Smith (1969), and Kohn (1969). Because the field is so broad, this article cannot cover all the ground but concentrates on recent empirical endeavors. Other reviews summarize the early studies in detail, such as Blake and Davis (1964), Williams (1968), Zavalloni (1980), Spates (1983), and Ball-Rokeach and Loges (1992).

Contemporary areas of research in values are not well integrated; each represents an active arena of social research that is empirically driven and theoretically informed. Below, will be summarized these areas, noting the unique contributions and insights of each one. The research reviewed here has been conducted by psychologists and political scientists as well as sociologists. However, all of it is premised on the sociological conceptual framework of values inherent in the definition given above.

### THE ROKEACH TRADITION

The most influential researcher on values in the last three decades is Rokeach. The focus of his work has been the development of an instrument to measure values that he believes are universal and transsituational (see especially Rokeach 1973). That is, Rokeach has tried to develop an instrument that can be used to compare individual commitment to a set of values wherever the researchers live and whenever they complete a survey. This instrument has been widely used in the measurement of values (Mayton et al. 1994).

The Rokeach Value Survey is an instrument made up of thirty-six value items that are ranked by survey subjects. The items are divided into two sets. The first ones are termed "instrumental values" and refer to values that reflect modes of conduct, such as politeness, honesty, and obedience. The second set refers to "terminal values" that reflect desired end states, such as freedom, equality, peace, and salvation. Each set of eighteen

value items is ranked by subjects according to the items' importance as guiding principles in their lives. The purpose of the procedure is to force subjects to identify priorities among competing values. In this model, the values are assumed to be universal; therefore, to some extent, each value is supported by every subject. The question is how subjects adjudicate between value conflicts. For example, the instrumental value "broad-minded" may conflict with the value "obedience." How would a person who is trying to conform to the expectations of racist parents maintain a broad-minded commitment to diversity? By requiring that values be rank-ordered, the Rokeach Values Survey helps disclose a person's value priorities.

One of the distinct advantages of the Rokeach Value Survey is that it is a fairly simple instrument that can be used by researchers in a variety of settings. Thus, it was possible to see if the value priorities of Michigan college students were similar to those of other subsamples of Americans, allowing comparisons of those with different demographic characteristics, such as age, sex, race, religion, and education. For example, in a national sample, Rokeach found that men and women tended to prioritize "a world at peace," "family security," and "freedom; however, men strongly valued "a comfortable life" while women did not, and women strongly valued "salvation" while men did not. Value priorities have been shown to be linked to a variety of attitudes about contemporary social issues. For example, as would be predicted, concern for the welfare of blacks and the poor is stronger among those who value equality.

The Rokeach Value Survey has been used by numerous researchers to explore many facets of values, such as the relationship between values and behavior, the role of values in justifying attitudes, and the extent to which people remain committed to particular values over time. An important early study of values employing the Rokeach model was conducted by Feather (1975), who measured the values of Australian high school and college students as well as those of their parents. One central finding demonstrated the importance of a close fit between the person and the environment in which that person is situated: Students were happiest when their values were congruent with those articulated by the schools they attended or the subjects they studied. Another finding was that parents were consistently more conservative, em-



phasizing values such as national security, responsibility, and politeness, while their children were more likely to emphasize excitement and pleasure, equality and freedom, a world of beauty, friendship, and broad-mindedness. It also was found that student activists were distinctive in their emphasis on humanitarianism, nonmaterialism, and social and political goals.

The Rokeach model underscores the potential conflicts between individuals with different value priorities. Different positions on important social issues may be traced to differential commitments to particular values. For example, Kristiansen and Zanna (1994) report that supporters of abortion rights emphasize values such as freedom and a comfortable life, whereas opponents place a high priority on religious salvation. Moreover, as they defend their positions, each group will justify its position by referring to its own value priorities; this, of course, may not be very convincing to people who do not share them. This may be one reason why the abortion debate seems intractable. Individuals also may be ambivalent about particular social issues because of their pluralistic commitment to two or more values that conflict in the public policy domain. This is the essence of Tetlock's (1986) "value pluralism model of ideological reasoning." For example, liberals tend to weight equality and freedom fairly equally, causing them to feel ambivalently about affirmative action policies (Peterson 1994).

Rokeach and Rokeach (1980) argue that values are not simply hierarchically prioritized but that each is interrelated in a complex system of beliefs and attitudes. Thus, a belief system may be relatively enduring, but changes in one value may lead to changes in others and in the whole system. When are personal values likely to endure, and when are they likely to change? Rokeach argues that individuals try to maintain a consistent conception of themselves that reflects their morality and competence. When their actions or beliefs contradict this self-conception, they feel dissatisfied and change is likely to occur to bring their actions or beliefs into line. Grube et al. (1994) review a number of studies in which researchers attempted to uncover contradictions in subjects'

values with the prediction that this conflict would lead to value change. These works have been called "self-confrontation" studies; they have found a significant degree of value change as a result of the method, even over long periods. However, the method is much less effective at inducing specific behavioral changes.

The central claim of values researchers is consistent with a commonsense understanding of values. Values are important because they guide people's behavior. At times they may be an even stronger motivation than is self-interest. For example, fear of arrest may not be as good an explanation for one's choice not to shoplift as is the more straightforward commitment to the value of right conduct. However, this central claim has been the most controversial in values research. The robust finding that values directly affect behavior has never surfaced in values research. The link does not exist, or several links in a long chain of causes intervene between these two crucial variables. This ambiguity has led Hechter (1992), for example, to suggest that social scientists stop using the term "values." Kristiansen and Hotte (1996, p. 79) observed that "although values, attitudes, and behavior are related, these relations are often small . . . one wonders why people do not express attitudes and actions that are more strongly in line with their values." Many people also wonder whether current measures of values are adequate. The Rokeach Value Survey, for example, may not be sufficiently complete or its definitions of values may be too abstract or vague to predict behavior accurately.

Kristiansen and Hotte (1996) argue that values researchers must pay much closer attention to the intervening factors in the values-behavior relationship. For example, those factors may include the way in which individuals engage in moral reasoning. Making a behavioral choice requires the direct application of very general values. How is this done? What do people consider in trying to make such a decision? Do they rely on ideological commitments to moral principles? Do they take into consideration the immediate context or circumstances? How much are they influenced by social norms? These questions are likely to guide

research on the values-behavior connection in the future.

### THE SCHWARTZ SCALE OF VALUES

A major evolution of the Rokeach Values Survey is found in the cross-cultural values research of Schwartz (see especially Schwartz 1992 and Smith and Schwartz 1997). Like Rokeach, Schwartz has focused on the measurement of values that are assumed to be universal. To that end, Schwartz has modified and expanded the Rokeach instrument. He also has proposed a new conceptual model that is based on the use of the new instrument in more than fifty countries around the world and more than 44,000 subjects (Smith and Schwartz 1997).

According to Schwartz (1992), values are arrayed along two general dimensions (Figure 1). In any culture, individual values fall along a dimension ranging from “self-enhancement” to “self-transcendence.” This dimension reflects the distinction between values oriented toward the pursuit of self-interest and values related to a concern for the welfare of others: “It arrays values in terms of the extent to which they motivate people to enhance their own personal interests (even at the expense of others) versus the extent to which they motivate people to transcend selfish concerns and promote the welfare of others, close and distant, and of nature” (1992, p. 43). The second dimension contrasts “openness to change” with “conservation”: “It arrays values in terms of the extent to which they motivate people to follow their own intellectual and emotional interests in unpredictable and uncertain directions versus to preserve the status quo and the certainty it provides in relationships with close others, institutions, and traditions” (1992, p. 43). This dimension indicates the degree to which individuals are motivated to engage in independent action and are willing to challenge themselves for both intellectual and emotional realization. Schwartz (1992, pp. 5–12) further postulates that within these two dimensions, there are ten motivational value types:

1. *Universalism*: “understanding, appreciation, tolerance, and protection for the welfare of all people and for nature”

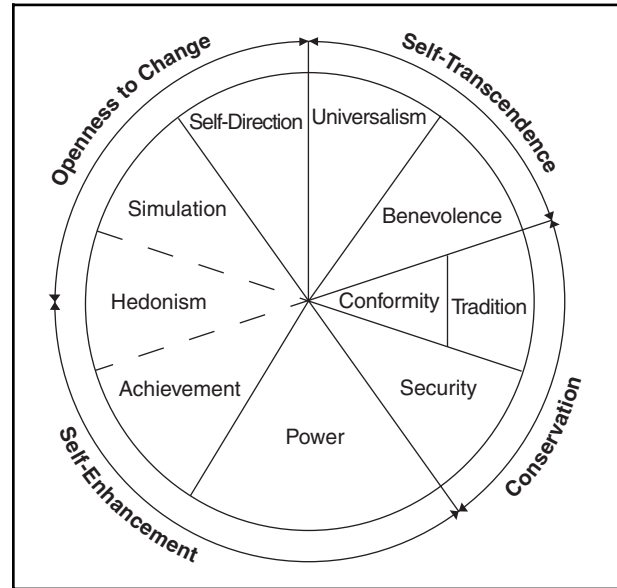


Figure 1. Structural relations among ten motivational types of values

SOURCE: Schwartz (1997), p. 87.

2. *Benevolence*: “preservation and enhancement of the welfare of people with whom one is in frequent personal contact”
3. *Conformity*: “restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms”
4. *Tradition*: “respect, commitment, and acceptance of the customs and ideas that one’s culture or religion imposes on the individual”
5. *Security*: “safety, harmony, and stability of society, of relationships, and of self”
6. *Power*: “attainment of social status and prestige, and control or dominance over people and resources”
7. *Achievement*: “personal success through demonstrating competence according to social standards”
8. *Hedonism*: “pleasure or sensuous gratification for oneself”
9. *Stimulation*: “excitement, novelty, and challenge in life”

10. *Self-direction*: “independent thought and action—choosing, creating, exploring”

Like Rokeach, Schwartz conceptualizes these motivational types as being dynamically interrelated, with those closest together being conceptually linked and having the greatest influence on one another. This model was not developed deductively but was derived from an empirical project of data collection in which the Schwartz Scale of Values was used. This instrument, which includes fifty-six Rokeach-style values items, is completed by subjects who rate each item on a ten-point scale of personal importance. Unlike the Rokeach instrument, this scale does not require the respondents to rank-order the items. Through the use of a multidimensional scaling technique (smallest-space analysis), statistical correlations of individual items in a survey sample are mapped in a two-dimensional space. Thus, each item is plotted on a graph, and clusters of those items constitute the domains identified in Figure 1. The major finding of the Schwartz project is that this basic visual model reappears in culture after culture. The system of values is essentially the same worldwide, although the emphasis given to particular domains varies from place to place.

Schwartz's dynamic model provides new insight into the values-behavior debate. Schwartz argues that the relationship of values to behavior (or any other variable) must be understood in the context of a multidimensional system. Voting for a particular political platform, for example, can be predicted on the basis of a person's value priorities. Given the interrelatedness of values in Schwartz's model, a person's values form a system: For example, a person who strongly endorses universalism is unlikely to endorse its distal correlate power while moderately endorsing values that are in closer proximity. Schwartz (1996) has used data from these values systems to predict political behavior.

Schwartz's research is especially important for distinguishing values at the individual and cultural levels. Individuals may differ in their values, but so too do cultures, with the members of one culture tending toward one set of priorities and the mem-

bers of another culture tending toward a different set. Cultural variation in values is of special interest to sociologists, while individual-level values are closer to the interests of social psychologists. Cultural values are important to sociologists because they reflect ways in which society balances conflicting concerns between individuals and groups and the dominant themes around which individuals are socialized. One issue is the provision of public goods; another is the extent to which individuals profess autonomy from the collectivity rather than identifying with it.

To obtain cultural-level values, Schwartz (1994) used the mean scores of values for each culture sample as the basis for plotting a new two-dimensional model. The data points thus are cultures rather than individual respondents. Among other findings, Schwartz discovered that east Asian nations emphasize hierarchy and conservatism, whereas west European nations emphasize egalitarianism and individual autonomy. Anglo nations, including the United States, fall between these extremes, emphasizing mastery and autonomy but also hierarchy; this may explain the greater tolerance for income inequality in countries such as the United States (Smith and Schwartz 1997).

Smith and Schwartz (1997) argue that values research should take two trajectories in the future. First, most studies now ask the respondents to report their own value priorities, whereas, especially for culture-level analyses, it would be useful to ask the respondents to report what they believe are the prevailing values of their culture. This may provide a better account of the normative milieu in which people evaluate their values and decisions. Second, most studies have examined the strength of individual commitment to particular values, but little research has dealt with the degree of value consensus in a culture. Because of the sociological concern about linking cultural values and the organization of societies, this is a crucial topic. One intriguing hypothesis is that socioeconomic development may enhance value consensus, while democratization may decrease it. These tendencies have broad implications for social stability and change in the future as countries pursue these goals.

## INDIVIDUALISM AND COLLECTIVISM

In cross-cultural research on values, no concepts have been explored in as much detail as individualism and collectivism. Consistent with the underlying theme of values research that private and communal interests may conflict, individualism and collectivism speak directly to the various ways in which cultures have balanced these competing goals.

The concept of individualism as a cultural construct has received much empirical attention, particularly since the publication of Hofstede's (1980) study of 117,000 IBM employees worldwide. In that study, fifteen items related to employment goals were subdivided into related clusters by using factor analysis, one of which Hofstede labeled individualism, inspiring this line of research. Theory and measurement in individualism and collectivism are associated primarily with Triandis (see especially Triandis 1989, 1995 and a review by Kagitcibasi 1997). In this tradition, the individualistic cultures of the West are typically contrasted with the collectivistic cultures of the East and Latin America. For example, Kim et al. (1994, pp. 6-7) argue that an individualistic ethos encourages individuals to be "autonomous, self-directing, unique, assertive, and to value privacy and freedom of choice." In contrast, "interdependency, succor, nurturance, common fate, and compliance" characterize a collectivistic ethos.

Triandis (1989, p. 52) defines collectivism in terms of in-groups and out-groups: "Collectivism means greater emphasis on (a) the views, needs, and goals of the in-group rather than oneself; (b) social norms and duty defined by the in-group rather than behavior to get pleasure; (c) beliefs shared with the in-group rather than beliefs that distinguish self from in-group; and (d) great readiness to cooperate with in-group members." Collectivism is characterized by two major themes that are consistent with the values dimensions of Schwartz's theory. First, collectivism is defined by conservation values: conformity, tradition, and security. The Japanese proverb "The nail that sticks up gets hammered down" illustrates the demand for conformity in the collectivistic Japa-

nese society. Second, collectivism is characterized by self-transcendent values. Individuals demonstrate a great willingness to cooperate in the pursuit of collective benefits, sacrificing their self-interest to do so. In conflicts between individual and collective interests, collectivists will subsume their individual interests in favor of those of the in-group. However, collectivists are not universally self-transcendent. Cooperation and self-sacrifice extend only to the boundaries of the in-group.

Individualism and collectivism are cultural constructs that define the values of societies, not those of individuals. Triandis argues that individuals vary in their adoption of the cultural ethos. To distinguish individualistic cultures from individualistic individuals, he uses the terms "idiocentrism" for the individual-level correlates of individualism and "allocentrism" for the individual-level correlates of collectivism. An individualistic culture is defined by having a majority of idiocentrics. These individuals identify primarily with the values of individualism, but not in every situation. Thus, individualistic cultures have both idiocentrics and allocentrics, and idiocentrics are collectivistic on occasion.

Triandis has developed a fifty-item scale to measure the various elements of individualism and collectivism. In addition, he advocates a multimethod approach to their study. For example, Triandis et al. (1990) used several measures, including the Schwartz Scale of Values. One of the measures is the Twenty Statements Test (Kuhn and McPartland 1954), which asks respondents to finish twenty sentences that begin with the words "I am . . ." This test is used to measure the degree of social identification or the "social content of the self" by disclosing the number and ordinal position of group membership references to the self relative to the number and ordinal position of individual references to the self. For example, "I am white" refers to group membership, whereas "I am kind" refers to a character trait. Collectivists are predicted to identify more closely with groups than are individualists. In Triandis et al.'s study, less than one-fifth of a U.S. sample's responses were social, whereas more than half of a mainland Chinese sample's responses were social. Using

another measure, individualists and collectivists were distinguished by attitude scales measuring the perceived social distance between in-group members and out-group members. Collectivists perceived in-group members as being more homogeneous than did individualist and also perceived out-group members as being more different from in-group members than did individualists.

Among the numerous findings of studies of the values-behavior relationship, one theme is particularly apparent. Individualists tend to emphasize competition, self-interest, and “free riding,” whereas collectivists tend to emphasize cooperation, conflict avoidance, group harmony, and group enhancement. Thus, in balancing individual and collective needs, collectivists favor the group more readily than do individualists. Collectivists also have been shown to favor equality in distributive outcomes, whereas individualists favor equity (Kagitcibasi 1997). Because this adjudication between the self and the collective is central to values research, this theme is replayed across research programs. Below, a line of research—“social values”—that provides a unique methodology for understanding these values will be examined.

Sociological research on values has long considered the relationship between values and social progress. For example, Weber ([1905] 1958) argued that an important factor in the rise of capitalism was the emergence of the Protestant Ethic, which encouraged hard work and self-control as a means of salvation. Thus, individuals were guided less by economic necessities or external coercion than by religious commitment. In values research, establishing a causal relationship between cultural values and social arrangements and outcomes is an ongoing endeavor. Triandis (1989), for example, suggests that individualism has two important structural antecedents: economic independence and cultural complexity. Independence enables individuals to pursue their own interests without fearing the economic consequences of deviation from the group. Cultural complexity, such as ethnic diversity and occupational specialization, fosters divergent interests and perspectives within a culture, increasing individualistic orientations. Another strand in values research has examined the

issues of cultural values and economic development. This line of research, which was initiated by Inglehart, is summarized below.

Future research on individualistic and collectivistic values is likely to proceed along three lines. First, these concepts may become more closely integrated with Schwartz’s general theory of values. Schwartz (1990, 1994) makes a case for this, and researchers are beginning to use measures of individualism/collectivism concurrently with the Schwartz Scale of Values (Triandis et al. 1990). Second, the overarching concepts of individualism and collectivism are becoming increasingly refined as specific relationships between values and other variables are examined. Triandis (1995) proposes that individualism and collectivism be further distinguished by horizontal and vertical dimensions in which “horizontal” refers to egalitarian social commitments and “vertical” refers to social hierarchies. Vertical collectivism may characterize the value structure of rural India, vertical individualism may characterize the structure of the United States, horizontal collectivism may characterize an Israeli kibbutz, and horizontal individualism may characterize Sweden’s value structure (Singelis et al. 1995). Third, another refinement has been proposed by Kagitcibasi (1997), who argues that “relational” individualism/collectivism be distinguished from “normative” individualism/collectivism. The normative approach emphasizes cultural ideals, such as an individualistic culture’s prioritization of rights and a collectivistic culture’s stress on group harmony and loyalty. The relational approach emphasizes differing concepts of the self in individualism and collectivism. In individualistic cultures, the self is perceived to be autonomous, with clear boundaries drawn between the self and others. In collectivist cultures, the self is perceived as more interdependent, with greater self-identification with the group.

### **SOCIAL VALUES**

The measurement of social values constitutes a unique approach in values research. More than any other approach, this one directly addresses the adjudication between individual and collective in-

terests. The basic issue in this research is how individuals prioritize allocations between themselves and anonymous others. How much are individuals willing to sacrifice their own interests for the good of the group?

Social values research is grounded in a larger paradigm of experimental gaming, the most famous example of which is the “prisoner’s dilemma.” Although game theory is quite complex, most experimental games have as a central theme the conflict between individual and collective outcomes. This is particularly true in “*n*-person” prisoner’s dilemma games and “commons” games, both of which are more generally called social dilemma games (for a general review of social dilemmas research, see Yamagishi 1994). The social values measure is a slight variation of these games, which always involve decisions that result in various payoffs to the self and others. These games are laboratory analogues of real-world situations in which values may play a significant role in behavioral choices. The example of supporting a tax levy for public education discussed at the beginning of the article constitutes a social dilemma because individual interests are in direct conflict with the common good. Another example is proenvironmental behavior such as not littering and recycling. The classic prisoner’s dilemma refers to a hypothetical situation involving the choice between exposing a coconspirator of a crime to obtain a lenient sentence and remaining loyal in spite of the greater personal risk in doing so.

In this research tradition, social values are measured through the administration of “decomposed games” to college students participating in social psychology experiments (Messick and McClintock 1968). Essentially, the subjects are presented with a series of payoffs that vary in consequence for both the self and a paired player. The subjects are asked to choose between two and sometimes three outcomes. For example, a subject may be asked which of the following outcomes would be preferable: receiving \$8 while the other person receives \$2 and receiving \$5 while the other person also receives \$5. The constellation of several choices with varying outcomes determines the subject’s social values. Primarily, the technique

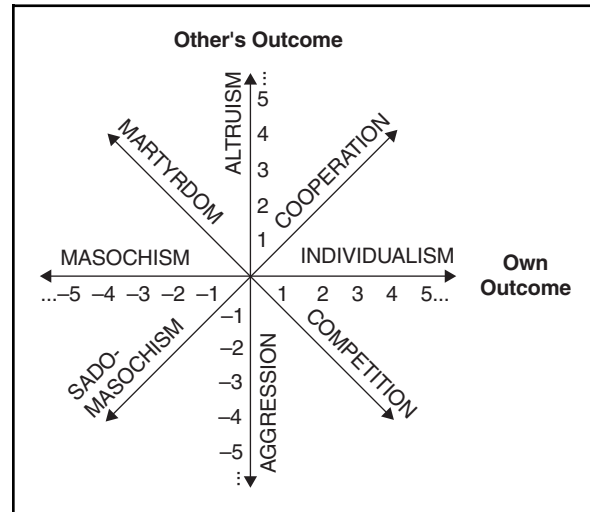


Figure 2. Vectors that define a subset of social values (Given a particular value orientation, an actor should select that combination of available own and others outcomes that has the greatest projection on the correspondent vector)

SOURCE: McClintock and Van Avermaet (1982), p. 49.

distinguishes between altruists, cooperators, individualists, and competitors, the most common classifications.

Each orientation is an indication of the preference that is given for the outcomes for both the self and the other. Subjects may attempt to maximize or minimize their own or others’ outcomes or may be indifferent to one or the other. Figure 2 displays the universe of social values in a two-dimensional representation of preferences for the self and the other. Altruists are defined by indifference to their own outcomes and a preference for maximizing others’ outcomes. Cooperators attempt to maximize both their own and others’ outcomes. Individualists maximize their own outcomes but are indifferent to those of others. Competitors are concerned with maximizing their own outcomes while minimizing others’ outcomes; that is, they attempt to maximize the difference between their own and others’ outcomes. Theoretically, other social values may exist, such as aggressors, who are indifferent to the self while minimizing others’ outcomes; sadomasochists, who minimize both self and others’ outcomes; masochists, who mini-

mize their own outcomes but are indifferent to others', and martyrs, who minimize their own outcomes while maximizing those of others. Except for occasional aggressors, these orientations have not been found empirically. Subjects who show no consistent pattern of choice are treated as unclassifiable.

Kuhlman and Marshello (1975) have shown that social values influence choices in prisoner's dilemma games, Liebrand (1986) and McClintock and Liebrand (1988) have demonstrated their influence in a variety of *n*-person games, and Kramer et al. (1986) have done the same thing in regard to a commons dilemma. In other words, values have been demonstrated to clearly affect behavioral choices in these laboratory situations.

Altruists and cooperators tend to cooperate, while individualists and competitors tend to defect (not cooperate). The essence of social values is the identification of individual differences regarding preferred outcomes in interdependent situations. One interpretation of social values is that "cooperators have internalized a value system in which satisfaction with interdependent relationships is directly proportional to the level of collective welfare they produce; competitors' rewards are directly proportional to how much more they receive than others; and individualists are relatively indifferent to others' outcomes, making them most similar to the traditionally conceived 'economic person'" (Kuhlman et al. 1986, p. 164).

Studies of social values have found that cooperators and noncooperators view social dilemmas differently. In general, decisions in social dilemmas are evaluated in terms of intelligence and morality. Players often are seen as making either "smart" or "good" decisions. Intelligence conforms to a player's social values. Cooperators tend to view cooperation as the intelligent choice, predicting that unintelligent others will defect. Noncooperators tend to view defection as the intelligent choice, predicting that unintelligent others will cooperate. This self-serving reversal does not occur with morality, however. Van Lange (1993) found that both cooperators and noncooperators view cooperation as moral. Both groups expect

more cooperation from highly moral others than from less moral others.

Although noncooperators see a link between morality and cooperation, they do not tend to view the social dilemma situation as being primarily moral. Cooperators are more likely to view cooperation as a moral act. Noncooperators frame the problem not in terms of morality but in terms of power: Cooperation is viewed as weak rather than moral. This is called the "might over morality hypothesis" (Liebrand et al. 1986). Viewing cooperation as both weak and unintelligent may provide the self-justification necessary for pursuing an egoistic goal ("Van Lange 1993). The might over morality hypothesis may overstate the case for noncooperators. Defectors have been found to assign more moral attributions to defection than do cooperators (Van Lange et al. 1990). The difference may be not only that cooperators view the dilemma as a moral situation more than defectors do but also that defectors may view their moral obligations differently. Both groups are likely to view self-enhancement as an important value.

Despite the fact that cooperators view social dilemmas as highly moral, their cooperation is not a matter of pure altruism. They are concerned with joint outcomes, with the self included. When they are exploited by noncooperators, they quickly defect (Kuhlman and Marshello 1975). In a study by Kuhlman et al. (1993), cooperators viewed cooperation as a partially self-interested act. That is, they recognized the self-beneficial outcomes of collective cooperation. By contrast, competitors and individualists did not do this. For cooperators and competitors, the difference may be explained by trust. Cooperators are high trusters, assuming that others will be cooperative. Competitors are low trusters, expecting others to defect as they themselves do (Kelley and Stahelski 1970). Competition therefore may be a result of a fear of exploitation or of losing in a competitive social arena. Individualists were found to be high trusters (expecting others to cooperate), unlike competitors. In this case, defection may be motivated more by greed than by fear.

Two studies suggest that social values discovered in the laboratory may have ecological validity,

that is, be relevant to real-world situations. Bem and Lord (1979) created a three-part strategy: First, they had experts list the personality characteristics of cooperators, competitors, and individualists. Second, they used decomposed games to measure the subjects' values. Third, they had the subjects' dormitory roommates describe the personality of the subjects. The personal descriptions of specific individuals correlated with both the personality templates created by the experts and subjects' social values as measured by the games. McClintock and Allison (1989) assessed the social values of subjects and, after several months, mailed them a request to donate their time to a charitable cause. Cooperators were more willing to donate time than were competitors and individualists.

Social values research describes differing motivational preferences and behaviors in social dilemmas. This line of research is fascinating because it has adopted the methodology (experimental games) of "rational choice" theorists, who argue that prosocial values always will be trumped by considerations of self-interest. Although the experimental paradigm is clearly artificial and perhaps contrived, it has fostered an accumulation of controlled evidence that supports the basic thesis of values research: Values are important determinants of behavioral choice.

### INGLEHART'S POSTMODERN THESIS

Values research as it is described in this article has followed two distinct strands represented by several schools of theory and research. The first is the micro-level strand. Values research at the microlevel has focused on individual values: what they are, how they are measured, how they vary, and how they affect behavior. The various methodologies for measuring values, from Rokeach's value survey, to the Schwartz scale, to Messick and McClintock's decomposed games, represent this strand. The second strand operates at the macro level, the level of cultures or societies. In this strand, one question concerns the distinct cultural variations in values priorities, such as Triandis's individualism versus collectivism. Another question follows from

Weber's work drawing a link between cultural values (Protestantism) and socioeconomic change (the emergence of capitalism). The contemporary work of Inglehart is concerned with the association of values and economic development and with how changes in economic conditions are reflected in very different value priorities. Important works in this tradition include Inglehart (1990) and Abramson and Inglehart (1995). A good summary is found in Inglehart (1995).

The starting point for this line of research is Weber's ([1905] 1958) classic association between Protestantism and the rise of capitalism in the West. Protestant Europe created a new value system that replaced several dogmatic restraints on the development of medieval European society. Weber was principally interested in the shift from traditional authority, best represented by the church, to what he called "rational-legal" authority, which endorsed individual achievement over ascriptive status and the preeminence of the impersonal state as an arbiter of conflicts. Crucial to modernization was secularization, which was reflected in an emerging scientific worldview, and bureaucratization, which was reflected in the rise of organizations driven by attempts at efficiency and explicit goal setting.

Inglehart argues that modernization has followed a fairly straightforward trajectory with economic growth and security at its epicenter. Correlated with modernization has been a coherent set of values such as industriousness, equity, thrift, and security. However, the achievement of economic security in the last twenty-five years in many countries around the world is fostering a change in the dominant values paradigm. Inglehart suggests that people may be experiencing a turn toward postmodern values that emphasize individualistic concerns such as friendship, leisure, self-expression, and the desire for meaningful, not just wealth-creating, work. In key ways, postmodern values follow a path similar to that of modernization values, especially in regard to secularization and individuation. However, they branch in other directions on several points. In societies in which major proportions of the members are economically secure, individuals seek to fulfill postmaterialistic



aims such as environmental protection and relational satisfaction. Individuals reject large institutions, whether religious or state-based, focusing instead on more private concerns. They seek new outlets for self-expression and political participation, particularly through local activism.

Some evidence for the postmodern shift comes from Inglehart and Abramson's (1994) analyses of the Euro-Barometer Surveys, which have measured values at frequent intervals since 1970 in all the European Community nations. These surveys have shown a general increase in postmaterialistic values.

Other evidence regarding the postmodern thesis is drawn from the 1990–1991 World Values Survey, which included data from representative samples from forty-three countries and more than 56,000 respondents. Using multiple indicators for the identification of modern and postmodern values, Inglehart tabulated mean scores for each country for forty-seven values. Those scores were employed in a factor analysis that disclosed two important dimensions. The first dimension contrasts traditional authority with rational-legal authority, and the second contrasts values guided by scarcity conditions with those guided by postmodern or security conditions. The distribution of these values in a two-dimensional space is illustrated in Figure 3. These distributions of values also correspond to countries, and so they can be plotted in a two-dimensional space (Inglehart 1995). For example, Inglehart places the United States, Great Britain, and Canada as well as the Scandinavian countries in the postmodern end of this dimension. China, Russia, and Germany ranked highest in the rational-legal domain. Nigeria stood out in its emphasis on traditional authority, while India, South Africa, and Poland fell between an emphasis on traditional authority and an emphasis on scarcity values.

These data do not suggest that once a country achieves a certain level of economic security, a sweeping change in values follows. The process is gradual, with segments of the population shifting from generation to generation. Hence, even in "postmodern" societies, many, if not most, of the members are likely to emphasize "modernist" val-

ues (Kidd and Lee 1997). These data do not suggest that those who adopt postmodern values score higher on various indicators of subjective well-being (Inglehart 1995). What changes is not their level of happiness per se but the criteria by which they evaluate their happiness.

Two issues will continue to receive attention in this line of research. First, there has been some debate about the role of environmentalism as a postmodern value. Does it indicate postmodern commitments, suggesting that it will be valued only by economically secure societies, or is it a more inclusive phenomenon? For a discussion of this issue, see Kidd and Lee (1997) and Brechin and Kempton (1997) along with other articles in that issue of *Social Science Quarterly*. More generally, the postmodern thesis must be tested with cross-national time-series data to identify values changes over time. These data also will provide insight into questions of causality (Granato et al. 1996): Do values affect economic development, or vice versa?

## CONCLUSION

Values research has been of interest to sociologists throughout the history of the discipline. Recently, the study of values has produced novel empirical research programs that carefully address core questions in this field of inquiry. Most fundamentally, values researchers ask what motivates behavior: Is it self-interest alone, self-interest and external coercion, or a combination of self-interest, coercion, and internalized values? A central issue in this line of questioning is the role of values in adjudicating conflicts between individual and collective pursuits.

Values researchers begin with the task of values measurement. What values do people hold? Which ones do they prioritize? How do values differ between members of society and between different cultures? Rokeach supplied the most common measure of values, and Schwartz expanded that measure. Messick and McClintock supplied a very different and innovative measure of social values within the paradigm of game theory research. Schwartz, Triandis, and Inglehart have

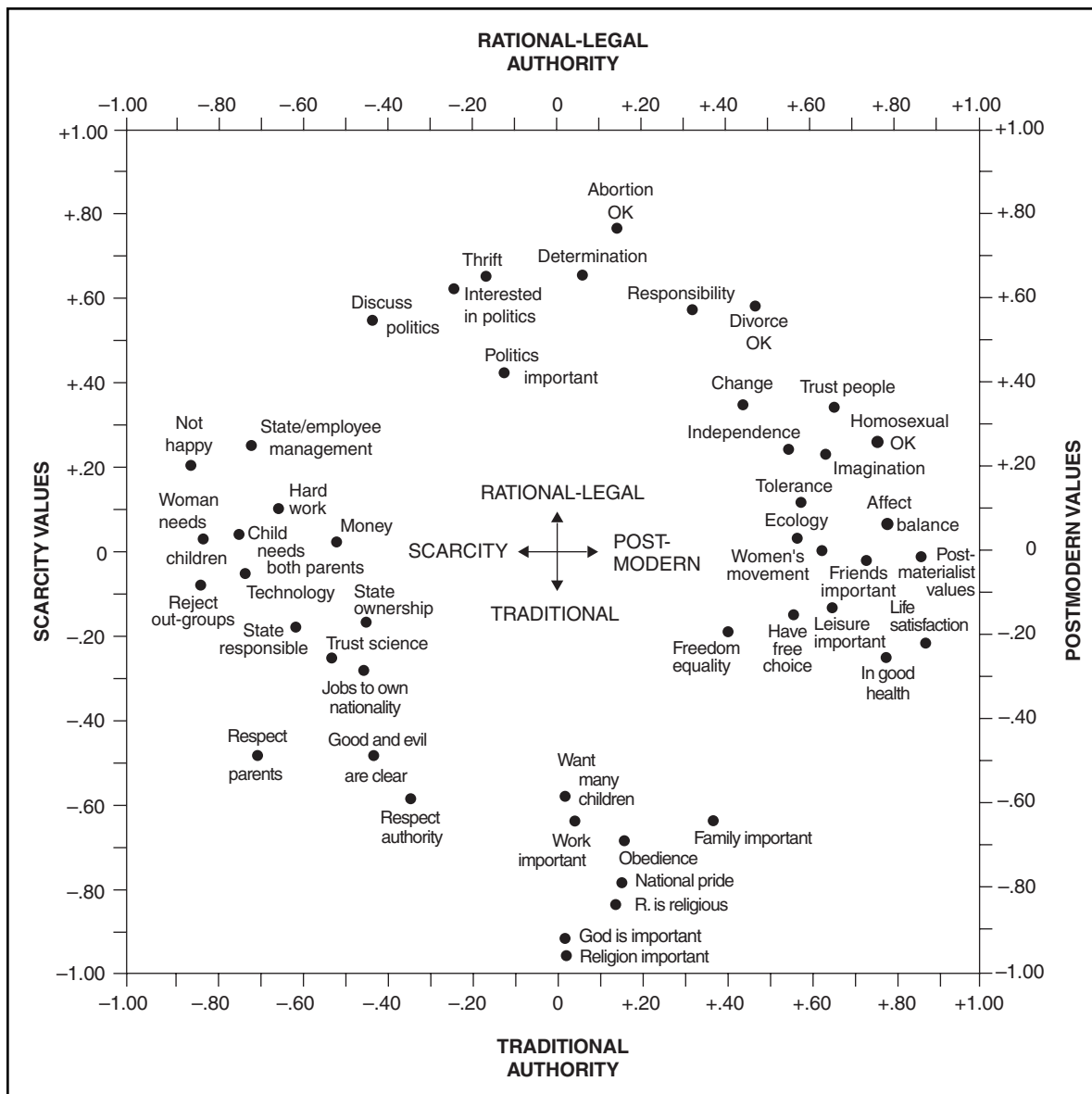


Figure 3. Variation in the values emphasized by different societies: traditional authority versus rational-legal authority and scarcity values versus postmodern values

SOURCE: Source: Inglehart (1995), p. 389.

made valuable contributions to the understanding of values cross-culturally. Of particular note is the apparent universality in the conceptual organization of values worldwide, while much variation in the cultural commitment to particular values has been observed.

Beyond measurement, values researchers have been concerned with the role of values in social

interaction. Do values motivate behavior? How are values related to other motivators of behavior? How do individuals increase or decrease their commitment to particular values? How do societies undergo values changes? How are conflicts between values adjudicated between individuals, between individuals and their communities, and between different cultures? Each of the research

traditions described in this article has made a contribution to an understanding of the complex values-behavior relationship. Rarely, however, has the question of values acquisition and retention been addressed. Given the enormous progress in cross-cultural values research, it is likely that this domain will garner a great deal of research attention in the next few years.

## REFERENCES

- Abramson, Paul R., and Ronald Inglehart 1995 *Value Change in Global Perspective*. Ann Arbor: University of Michigan Press.
- Allport, G. W., P. E. Vernon, and G. Lindsey 1960 *A Study of Values*. Boston: Houghton Mifflin.
- Ball-Rokeach, Sandra J., and William E. Loges 1992 "Value Theory and Research." In E. F. Borgatta and M. L. Borgatta, eds., *Encyclopedia of Sociology*, vol. 1. New York: MacMillan
- Bem, Daryl J., and Charles G. Lord 1979 "Template Matching: A Proposal for Probing the Ecological Validity of Experimental Settings in Social Psychology." *Journal of Personality and Social Psychology* 37:833-846.
- Blake, Judith, and Kingsley Davis 1964 "Norms, Values, and Sanctions." In R. E. L. Faris, ed., *Handbook of Modern Sociology*, Boston: Rand McNally.
- Brechin, Steven R., and Willett Kempton 1997 "Beyond Postmaterialist Values: National versus Individual Explanations of Global Environmentalism." *Social Science Quarterly* 78:16-20.
- Feather, Norman T. 1975 *Values in Education and Society*. New York: Free Press.
- Granato, Jim, Ronald Inglehart, and David Leblang 1996 "The Effect of Cultural Values on Economic Development: Theory, Hypotheses, and Some Empirical Tests." *American Journal of Political Science* 40:607-631.
- Grube, Joel W., Daniel M. Mayton, and Sandra J. Ball-Rokeach 1994 "Inducing Change in Values, Attitudes, and Behaviors: Belief System Theory and the Method of Value Self-Confrontation." *Journal of Social Issues* 50:153-174.
- Hechter, Michael 1992 "Should Values Be Written Out of the Social Scientist's Lexicon?" *Sociological Theory* 10:214-230.
- 1993 "Values Research in the Social and Behavioral Sciences." In M. Hechter, L. Nadel, and R. E. Michod, ed., *The Origin of Values*. New York: Aldine de Gruyter.
- Hofstede, G. 1980 *Culture's Consequences: International Differences in Work-Related Values*. Thousand Oaks, Calif.: Sage.
- Inglehart, Ronald 1990 *Culture Shift in Advanced Industrial Society*. Princeton, N.J.: Princeton University Press.
- 1995 "Changing Values, Economic Development, and Political Change." *International Social Science Journal* 145:379-403.
- and Paul R. Abramson 1994 "Economic Security and Value Change." *American Political Science Review* 88:336-354.
- Kagitcibasi, Cigdem 1997 "Individualism and Collectivism." In J. W. Berry, M. H. Segall, and C. Kagitcibasi, eds. *Handbook of Cross-Cultural Psychology*, vol. 3. Boston: Allyn and Bacon.
- Kelley, Harold H., and A. J. Stahelski 1970 "Social Interaction Basis of Cooperators' and Competitors' Beliefs about Others." *Journal of Personality and Social Psychology* 16:66-91.
- Kidd, Quentin, and Aie-Rie Lee 1997 "Postmaterialist Values and the Environment: A Critique and Reappraisal." *Social Science Quarterly* 78:1-15.
- Kim, Uichol, Harry C. Triandis, Cigdem Kagitcibasi, Sang-Chin Choi, and Gene Yoon 1994. "Introduction." In U. Kim, H. C. Triandis, C. Kagitcibasi S.-C. Choi, and G. Yoon, eds., *Individualism and Collectivism: Theory, Method and Applications*. Thousand Oaks, Calif.: Sage.
- Kluckhohn, Clyde 1951 "Values and Value Orientation in the Theory of Action." In T. Parsons and E. A. Shils, eds., *Toward a General Theory of Action*. Cambridge, Mass.: Harvard University Press.
- Kluckhohn, Florence R., and Fred L. Strodtbeck 1961 *Variations in Value Orientations*. Westport, Conn.: Greenwood Press.
- Kohn, Melvin L. 1969 *Class and Conformity: A Study in Values*. Homewood, Ill.: Dorsey.
- Kristiansen, Connie M., and Alan M. Hotte 1996 "Morality and the Self: Implications for the When and How of Value-Attitude-Behavior Relations." In C. Seligman, J. M. Olson, and M. P. Zanna, eds., *The Psychology of Values: The Ontario Symposium*, vol. 8. Mahwah, N.J.: Erlbaum.
- , and Mark P. Zanna 1994 "The Rhetorical Use of Values to Justify Social and Intergroup Attitudes." *Journal of Social Issues* 50:47-66.
- Kuhlman, D. Michael, Clifford Brown, and Paul Teta 1993 "Judgments of Cooperation and Defection in

- Social Dilemmas: The Moderating Role of Judges' Social Orientation." In W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke, eds., *Social Dilemmas: Theoretical Issues and Research Findings*. New York: Pergamon.
- , Curt R. Camac, and Denise A. Cunha 1986 "Individual Differences in Social Orientation." In H. A. M. Wilke, D. M. Messick, and C. G. Rutte, eds., *Experimental Social Dilemmas*. Frankfurt: Verlag Peter Lang.
- , and Alfred F. J. Marshello 1975 "Individual Differences in Game Motivation as Moderators of Preprogrammed Strategy Effects in Prisoner's Dilemma." *Journal of Personality and Social Psychology* 32:922-931.
- Kuhn, M. H., and R. McPartland 1954 "An Empirical Investigation of Self Attitudes." *American Sociological Review* 19:68-76.
- Liebrand, Wim B. G. 1986 "The Ubiquity of Social Values in Social Dilemmas." In H. A. M. Wilke, D. M. Messick, and C. G. Rutte, eds., *Experimental Social Dilemmas*. Frankfurt: Verlag Peter Lang.
- , Ronald W. T. L. Jansen, Victor M. Rijken, and Cor J. M. Suhre 1986 "Might Over Morality: Social Values and the Perception of Other Players in Experimental Games." *Journal of Experimental Social Psychology* 22:203-215.
- Mayton, Daniel M., Sandra J. Ball-Rokeach, and William E. Loges 1994 "Human Values and Social Issues: An Introduction." *Journal of Social Issues* 50:1-8.
- McClintock, Charles G., and Scott T. Allison 1989 "Social Value Orientation and Helping Behavior." *Journal of Applied Social Psychology* 19:353-362.
- , and Wim B. G. Liebrand 1988 "Role of Interdependence Structure, Individual Value Orientation, and Another's Strategy in Social Decision Making: A Transformational Analysis." *Journal of Personality and Social Psychology* 55:396-409.
- Messick, David M., and Charles G. McClintock 1968 "Motivational Bases of Choice in Experimental Games." *Journal of Experimental Social Psychology* 4:1-25.
- Parsons, Talcott, and Edward A. Shils 1951 "Values, Motives, and Systems of Action." In T. Parsons and E. A. Shils, eds., *Toward a General Theory of Action*. Cambridge, Mass., Harvard University Press.
- Peterson, Randall S. 1994 "The Role of Values in Predicting Fairness Judgments and Support of Affirmative Action." *Journal of Social Issues* 50:95-116.
- Rokeach, Milton 1973 *The Nature of Human Values*. New York: Free Press.
- , and Sandra J. Ball-Rokeach 1980 "Stability and Change in American Value Priorities, 1968-1981." *American Psychologist* 44:775-784.
- Schwartz, Shalom H. 1990 "Individualism-Collectivism: Critique and Proposed Refinements." *Journal of Cross-Cultural Psychology* 21:139-157.
- , 1992 "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries." *Advances in Experimental Social Psychology* 25:1-65.
- 1994 "Beyond Individualism/Collectivism: New Cultural Dimensions of Values." In U. Kim, H. C. Triandis, C. Kagitcibasi, S.-C. Choi, and G. Yoon, eds., *Individualism and Collectivism: Theory, Method, and Applications*. Thousand Oaks, Calif., Sage.
- 1996 "Value Priorities and Behavior: Applying a Theory of Integrated Value Systems." In C. Seligman, J. M. Olson, and M. P. Zanna, eds., *The Psychology of Values: The Ontario Symposium*, vol. 8. Mahwah, N. J.: Erlbaum.
- Scott, W. A. 1965 *Values and Organizations*. Chicago: Rand McNally.
- Singelis, T. M., H. C. Triandis, D. S. Bhawuk, and M. Gelfand 1995 "Horizontal and Vertical Dimensions of Individualism and Collectivism: A Theoretical and Measurement Refinement." *Cross-Cultural Research* 29:240-275.
- Smith, M. Brewster 1969 *Social Psychology and Human Values: Selected Essays*. Chicago: Aldine.
- Smith, Peter B., and Shalom Schwartz 1997 "Values." In J. W. Berry, M. H. Segall, and C. Kagitcibasi, eds., *Handbook of Cross-Cultural Psychology*, vol. 3. Boston: Allyn and Bacon.
- Spates, James L. 1983 "The Sociology of Values." *Annual Review of Sociology* 9:27-49.
- Stigler, George J. 1950 "The Development of Utility Theory (I)." *Journal of Political Economy* 58:307-327.
- Tetlock, Philip E. 1986. "A Value Pluralism Model of Ideological Reasoning." *Journal of Personality and Social Psychology* 50:819-827.
- Triandis, Harry C. 1989 "Cross-Cultural Studies of Individualism and Collectivism." In J. J. Berman, ed., *Nebraska Symposium of Motivation: Cross-Cultural Perspectives*. Nebraska: University of Nebraska Press.
- 1995 *Individualism and Collectivism*. Boulder, C. Col.: Westview.
- , Christopher McCusker, and C. Harry Hui 1990 "Multimethod Probes of Individualism and Collectivism." *Journal of Personality and Social Psychology* 59:1006-1020.

- Van Lange, Paul A. M. 1993 "Rationality and Morality in Social Dilemmas: The Influence of Social Value Orientations." In W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke, eds., *Social Dilemmas: Theoretical Issues and Research Findings*. New York: Pergamon.
- , B. G. Liebrand, and D. Michael Kuhlman 1990 "Causal Attribution of Choice Behavior in Three N-Person Prisoners Dilemmas." *Journal of Experimental Social Psychology* 26:34–48.
- Weber, Max 1958 [1905] *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner's.
- Williams, Robin M. 1960 *American Society: A Sociological Interpretation*. New York: Knopf.
- 1968 "The Concept of Values." In D. L. Shils, ed., *International Encyclopedia of the Social Sciences*, vol. 16. New York: Free Press.
- Yamagishi, Toshio 1994 "Social Dilemmas." In K. S. Cook, G. A. Fine, and J. House, eds., *Sociological Perspectives on Social Psychology*. Boston: Allyn and Bacon.
- Zavalloni, M. 1980 "Values." In H. C. Triandis and R. W. Brisman, eds., *Handbook of Cross-Cultural Psychology*, vol. 5. Boston: Allyn and Bacon.

DAVID R. KARP

## VIOLENCE

See Crime Rates; Criminology; Crowds and Riots; Family Violence; Sexual Violence and Exploitation; Terrorism.

## VOLUNTARY ASSOCIATIONS

The 1990s saw renewed interest in and concerns about voluntary associations and their roles in society. On the international level, countries that had been part of the Soviet Union and its power bloc continued to form and experiment with what they called "informal groups," which had the essential characteristics of voluntary associations. That is, those groups were independent of control from outside sources, people were free to join or leave them, and members established their own objectives and goals and developed means that might achieve them. Among the most important developments arising from these informal groups was the emergence of political parties as part of the struggle to establish democratic governments.

## ORIGIN OF THE IDEA OF VOLUNTARY ASSOCIATIONS

It is generally acknowledged that the origins of voluntary associations are in the writings of early Reformation leaders such as Martin Luther and John Calvin (Hooker 1997). Calvin taught that all believers should participate equally in church decisions. The way to accomplish this equality was to see the church as a free and voluntary association of members; at the same time, to become a member, an individual had to be approved by the congregation. An early expression of this democratic church model developed in New England towns, with the local Congregationalist church as the prototypical voluntary association.

When Alexis de Tocqueville based *Democracy in America* on his tour of the United States in the 1830s, he took particular note of the degree to which Americans formed groups to serve personal interests and solve problems from the mundane to the profound. Tocqueville (1956) was particularly impressed by New England small towns with their autonomous local church congregations, whose citizens gathered in "town meetings" and voted on projects, from building schools and roads to caring for the poor. Current American nostalgia for local control to preserve the moral order may owe much to the almost sacred aura given to the reading of Tocqueville's description of early American society.

## CHARACTERISTICS AND OBJECTIVES OF VOLUNTARY ASSOCIATIONS

Research on voluntary associations was limited until recently, with most people accepting their importance to a free society and concentrating on questions of demographic characteristics and the contributions they made to local communities (Irwin et al. 1997).

One of the most consistent findings about voluntary associations (Cutler 1976) was that individuals with higher socioeconomic status (SES) were more likely to participate in voluntary associations. Age, race, and gender (while influenced strongly by SES) also were identified as important factors in membership, with middle-aged persons, whites, and males more likely to be members.

Gender differences in voluntary association membership have been studied in terms of rates of participation as well as differences in the types of organizations to which each sex belongs. Historically, women's participation rates in voluntary associations (McPherson and Smith-Lowin 1986) were lower than men's. Furthermore, the groups to which women belonged tended to be smaller, single-sex, and expressive rather than instrumental. Still, in the 1980s, Knoke (1986) reported that the gender gap was narrowing as more women entered the professional ranks.

Studies of the effect of race on voluntary association membership provided inconsistent findings. For example, Hyman and Wright (1971) documented a sharp increase in membership among blacks between 1955 and 1962 (sharper than that among whites). However, blacks continued to be less likely to belong to a voluntary association other than the local church congregation and its Bible study groups. Knoke (1986, p.4) summarized more recent research with the statement that "researchers generally found that blacks' participation rates fell below whites' but disagreed on whether the gap could be traced to black SES disadvantages."

Researchers interested in the way nonpolitical voluntary associations influence political participation have found that individuals who are members of such organizations are more likely to vote and participate in politics (Sigelman et al. 1985; Wolfinger and Rosenstone 1980; Milbrath and Goel 1977; Rogers et al. 1975).

Voluntary associations range in size from groups of four or five persons to those with hundreds of thousands of members worldwide; structures vary from very informal with little leadership and few norms or guiding rules to highly structured with formal leadership, codes of conduct, and elected and appointed offices. These differences reflect different goals and the ability to influence civic and political affairs.

Some associations, such as the American Medical Association, labor unions, and churches that are hierarchic in structure or practice infant baptism, may have some of the characteristics of voluntary associations, but they are not seen as such in the definition adopted here.

## CURRENT RESEARCH

In the United States, Putnam (1996) developed the hypothesis that voluntary associations might well have run their course as he recounted the tale of "Bowling Alone," suggesting that the decline of voluntary associations was bringing with it a decline in the country's civic health. His hypothesis sparked renewed interest in voluntary associations and their place in American society.

While Putnam was suggesting the decline of voluntary associations, Wuthnow (1994) was reporting on the large and apparently growing number of Americans who were joining small groups that seemed to have the characteristics of voluntary associations. Wuthnow's national survey of American adults found that "exactly forty percent of the adult population of the United States claims to be involved in a small group that meets regularly and provides caring and support for those who participate in it" (1994, p. 45). Assuming that an American adult belonged only to one small group, Wuthnow estimated that at the time of his study, there were at least three million small groups active in the United States, with approximately one group for every eighty people, assuming group size averages of close to twenty-five. Drawing on a variety of sources, Wuthnow subdivided these small groups as follows:

Bible study and related religious groups:  
1.7 million

Self-help groups: 500,000

Special-interest groups (political, sports,  
book and/or discussion): 750,000  
(1994, p.76)

These figures contrast sharply with earlier attempts to estimate the number of voluntary associations active in American society. Rose (1967) estimated that there were over 100,000 such associations in the United States, and Hyman and Wright (1971) reported that 57 percent of the American adult population did *not* belong to a voluntary association. However, local and regional studies found higher participation rates. For example, Babchuk and Booth's Nebraska study (1969) found that 80 percent of the adult population belonged to at least one voluntary association. More recently, Knoke observed that "perhaps one third of U.S. adults belong to no formal voluntary organizations and only a third hold membership

in more than one (not including churches)” (1986, p. 3). Excluding churches may help account for much of the discrepancy in the figures provided by various scholars.

The highest figure provided for membership in voluntary associations among adult Americans came from the 1990–1991 World Values survey. Galston and Levine (1997, p. 2) reported that that survey showed that “82 percent of Americans belonged to at least one voluntary association, a rate exceeded only by Iceland, Sweden and the Netherlands.” This survey was carried out as Wuthnow was doing his study of small groups and Putnam was bemoaning the decline of at least some kinds of voluntary associations.

Evidence of concerns about growth and decline in voluntary associations can be found in events such as the agreement between Lions Clubs International and the Junior Chamber International (JCs) to form a global partnership to boost membership and encourage lifelong service to the community. Lions International (1998) reported 1.4 million members representing 43,700 clubs in 185 countries, while the Junior Chamber reported 322,000 members in 9,000 chapters in 123 nations. The members of JCs typically have been in the under-40 age bracket; the intent of the new collaboration is to have them join Lions Clubs as they move up the age ladder. Both groups would be encouraged to work more closely together in community service. In this way, they hope to stem the age creep that has brought stagnation and decline to many voluntary associations.

Skocpol and her colleagues in the Harvard Civic Engagement Project have begun to document the local, state, and national linkages of voluntary associations, in the process challenging the assumption that the strength of American civic life ever lay in the local focus of voluntary associations. In her historical overview, Skocpol (1997) identified events such as the Revolutionary War and the subsequent electoral politics, along with the development of an extraordinarily extensive and efficient national postal system, as key factors encouraging the activities of thousands of local and extralocal voluntary associations. Major growth spurts occurred between 1820 and 1840, from after the Civil War to the end of the century, and in the 1930s. These growth spurts seem to be related

to the great issues of the time: slavery and its moral dilemmas, industrialization, and economic crisis.

The Harvard group has so far tracked detailed life histories of some fifty-five voluntary associations that have enrolled 1 percent or more of American adults at some point in their history. Four-fifths of these associations still exist, with most of them paralleling the three-tiered government structure with local, state, and national branches. Although many groups have come and gone at the local level, a more balanced historical view sees voluntary associations as vital links between local and national civil life. The social historian Alexander Hoffman was cited as stating that “local institutions and organizations may best be understood as branch offices and local chapters . . . the building blocks of a ‘nation of joiners.’ . . . Americans enlisted in local church groups, fraternal lodges, clubs, and other organizations that belonged to nationwide networks” (Skocpol 1997, p. 3).

There is evidence of a decline in some types of voluntary associations even as new small groups emerge. For example, Skocpol (1996) noted that since the 1960s, the Christian Coalition has been one of the few cases of local to national federations growing, while some, such as the Lions, Rotary, and the Junior Chamber, have found themselves with an aging population and in a process of slow decline or even death. Thus, the new alliance between the Lions and the Junior Chambers mentioned above may be seen as an effort at revitalization.

Current research about voluntary associations has revealed a decline of same-sex organizations, growing numbers of college-educated and professional women members, and the replacement of family-oriented by professional associations. As Skocpol put it, “the best educated people are still participating in more groups overall, but not in the same groups as their less well-educated fellow citizens” (1997, p. 5).

At least in the short run, the educational gap, which is reflected in the occupational and income gaps, seems more of a threat to the well-being of civil society than does the so-called loss of the local group. Indeed, Wuthnow’s data, supported by recent research on small faith communities in the U.S. Catholic Church (D’Antonio 1997), suggests the opposite: Small associations are alive and booming at the local level, with more than a little sup-

port from national organizations that provide regional gatherings, bring together diverse racial and ethnic groups, and provide a wide range of literature that urges outreach as a part of their mission. Their members may be spending more time working in soup kitchens and other service activities. Meanwhile, other small groups are supporting local teenage sports clubs rather than participating in union-style bowling leagues.

Among Wuthnow's findings was that social support in these small groups tends to focus on the individual; the groups themselves revealed tendencies to see political and social issues in a conservative vein. Outreach seldom got beyond the soup kitchen state of concern for others.

To the extent that these new groups cut across class, gender, and age lines, they may be fulfilling the hope expressed by Skocpol (1997, p. 5) that Americans will find new ways to work together, "not just on 'helping the poor' but 'doing with' rather than 'doing for' if we want to revitalize the best traditions of American voluntarism."

#### NEW DIRECTIONS

Among the new directions in voluntary association activity is the development of an interdisciplinary relationship between social work and veterinary medicine. Built on the premise that "democracies are based on the value of the worth and dignity of each person, and the empowering of persons to take action in their own lives" (Granger and Granger 1998), Colorado State University has established a Human-Animal Bond Center (HABIC). Its goal is to provide animal-assisted therapy and activities in partnership with community health, mental health, education, and human service programs. The founders of this program extend the respect for humans to animals and to the environment as a crucial element in the survival of a democratic society. The essential factor in their vision is the linkage of voluntary associations with formal groups such as social work agencies and veterinary medicine societies.

At the international level, scholars in Finland are going beyond studying the functional role of voluntary associations in the stability and growth of democratic societies. They propose that "from the point of view of social constructionism voluntary associations can be seen as forums for the

production and transmitting of social meanings. This is an intersubjective process which may yield objectivated and taken-for-granted meanings. If internalized, these meanings become the source of personal identity and goal-formation of the association. Thus voluntary associations can be seen not only as part of the western culture heritage but also as cultures in themselves" (Raivio and Heikkala 1998, p. 1).

#### CONCLUSION

Voluntary associations generally are seen as central ingredients of a pluralist democratic society. Millions of Americans belong to hundreds of thousands of these associations, as do growing numbers of people worldwide. Among the more important research findings has been the multifaceted nature of voluntary associations' growth and impact: Some associations have remained strictly local; some have grown from local to state, national, and international levels; and not a few have grown from the top down, such as the American Legion and the PTA. Their influence on local, state, and national governments has led to much important legislation, such as the GI Bill fostered by the American Legion.

A review of the literature and current research challenges the nostalgic view that what is needed to restore vitality to American democracy is a return to localism and a shrinking national government. Instead, these findings suggest that the current trend toward the growth of a variety of types of voluntary associations, within and across national boundaries and working with rather than apart from governments, is the best formula for the revitalization of American political democracy. The limited evidence from other societies suggests that that same formula applies to all societies seeking to model Western democracies.

#### REFERENCES

- 1998 "Joining Forces to Promote Volunteerism, Membership Growth." *Association Management* 50:24.
- Babchuk, Nicholas, and Alan Booth 1969 "Voluntary Association Membership: A Longitudinal Analysis." *American Sociological Review* 34:31-45.
- Cutler, Stephen J. 1976 "Age Differences in Voluntary Association Membership." *Social Forces* 55:43-58.



- D'Antonio, William V. 1997 "Small Christian Communities in the Catholic Church: A Study in Progress" In Robert S. Pelton, ed., *Small Christian Communities: Imagining Future Church*. Notre Dame, Ind.: University of Notre Dame Press.
- Galston, William A., and Peter Levine 1997 "America's Civic Condition: A Glance at the Evidence." *Brookings Review*. 15:23-26.
- Granger, Ben P., and Georgia V. Granger 1998 "Voluntary Association Development and the Human-Animal Bond." <http://www/friends-partners.org/friends/audem/audem93/granger.htm>lopt-other-unix-english.
- Hooker, Richard 1997 "Voluntary Associations." World Cultures Home Page: <http://www.wsu.edu:8080/dee/GLOSSARY/VOLUNTA.HTM>.
- Hyman, Herbert H., and Charles R. Wright 1971 "American Adults: Replication Based on Secondary Analyses of National Sample Surveys." *American Sociological Review* 36:191-206.
- Irwin, Michael, Charles Tolbert, and Thomas Lyson 1997 "How to Build Strong Home Towns." *American Demographics* N.Y. Vol 19:42-47.
- Knocke, David 1986 "Associations and Interest Groups." *Annual Review of Sociology* 12:1-21.
- McPherson, J. Miller, and Lynn Smith-Lowin 1986 "Sex Segregation in Voluntary Associations." *American Sociological Review* 51:61-79.
- Milbrath, Lester W., and M. L. Goel 1977 *Political Participation*. Chicago: Rand McNally.
- Morgan, Patrick 1997 "Renovating Civil Society." *IPA Review*, March, pp. 29-30.
- Putnam, Robert D. 1996 "The Strange Disappearance of Civic America." *American Prospect* 24:34-49.
- Raivio, Risto, and Juha Heikkala 1998 "Voluntary Associations as Cultures." <http://www.uta.fi/laitokset/sosio/culture/firstassociat.htm>.
- Rogers, David L., Gordon L. Bultena, and Ken H. Barb 1975 "Voluntary Association Membership and Political Participation: An Exploration of the Mobilization Hypothesis." *Sociological Quarterly* 16:305-318.
- Rose, Arnold M. 1967 *The Power Structure*. New York: Oxford University Press.
- Sigelman, Lee, Philip W. Roeder, Malcolm E. Jewell, and Michael A. Baer 1985 "Voting and Nonvoting": A Multi-Election Perspective." *American Journal of Political Science* 29:749-765.
- Skocpol, Theda 1996 "Unravelling from Above." *American Prospect*. 25:20-25.
- 1997 "Building Community Top-Down or Bottom-Up? America's Voluntary Groups Thrive in a National Network." *Brookings Review* 15:16-19.
- Tocqueville, Alexis de [1836] 1956 *Democracy in America*, ed. and abridged by Richard D. Heffner. New York: Mentor.
- Wolfinger, Raymond E., and Steven J. Rosenstone 1980 *Who Votes?* New Haven, Conn.: Yale University Press.
- Wuthnow, Robert (ed.) 1991 *Between States and Markets: The Voluntary Sector in Comparative Perspective*. Princeton, N. J: Princeton, University Press.
- , 1994 *Sharing the Journey: Support Groups and America's New Quest for Community*. New York: Free Press.

WILLIAM V. D'ANTONIO

## VOTING BEHAVIOR

In addition to sociologists, scholars from many different fields, including history, political science, psychology, and geography, have studied elections and voting behavior. In current American sociology, however, these topics are largely neglected. Major advances have been made in related disciplines, yet as of one of the pioneers, the sociologist Rice (1928, p. p.vii) stated: "The phenomena of politics are functions of group life. The study of group life per se is a task of sociology." In general terms, despite variations in emphasis between different approaches, the sociological study of voting behavior is concerned with the way individuals obtain, select, and process information related to the political arena; the various forces that shape this process; the relevance individuals attribute to the political sphere; and how they decide to participate in or refrain from specific political actions. Elections provide a convenient focus, a point where the often elusive and latent processing of political information manifests behavioral correlates such as voting or abstaining and supporting one candidate or the other. In contrast, forecasting election returns is not a primary goal of the sociological study of voting behavior, although the general public, parties, and politicians are interested mostly in this aspect. Much applied research served these immediate needs and interests in the past and continues to do so. Still, in the field of voting behavior, pure (academic) and applied research peacefully coexist; cross-fertilization rather than mutual irreverence characterizes their relationship.

The study of voting behavior began in the late eighteenth century (Jensen 1969), although most of the very early work does not meet strict schol-

arly standards. In the course of its development as an academic discipline, two different strands that are still discernible have emerged. The first strand—*aggregate data analysis*—is characterized by the use of actual election returns compiled for geopolitical units such as wards, districts, and counties. Those returns are compared with census data, providing a sociodemographic profile of those areal units. Starting in the late nineteenth century, there developed a school of quantitative historiography that made extensive use of maps representing voting and/or census information by using different shades and colors (Frederick Jackson Turner in the United States and André Siegfried in France). The mere visual inspection and somewhat subjective interpretation of those maps by the Turner school were supplemented and then replaced by more vigorous statistical techniques, in particular correlation analysis, inspired by the sociologist Franklin Giddins at Columbia. One of Giddins's students, Rice (1928), demonstrated the utility of quantitative methods in politics. At the University of Chicago, interdisciplinary cooperation in the social sciences produced some of the most outstanding work of that time (e.g., Gosnell 1930). The advent of modern survey research in the 1930s and 1940s, however, obscured the aggregate approach for quite some time.

The second strand in the study of voting behavior—*analysis of survey data*—also had some early forerunners. Polling individuals about their voting intentions (“straw polls”) or past voting decisions started in the late nineteenth century. In one of the most extensive efforts, more than a quarter million returns from twelve midwestern states were tabulated by a Chicago newspaper for the 1896 presidential contest between McKinley and Bryant. In the 1920s, straw polls conducted by newspapers and other periodicals were quite common and popular. Their reputation was ruined, though, by the failure of the *Literary Digest* poll to foresee the landslide victory of Franklin D. Roosevelt in the 1936 election. By that time, however, pioneers of modern public opinion research such as George Gallup, Archibald Crossley, and Elmo Roper had started to use more rigorous sampling methods as well as trained interviewers to ensure a proper representation of all strata of the electorate (Gallup [1944] 1948).

Interest in voting and political behavior and concern with mass communication, marketing

strategies, and the public's attitude toward World War II stimulated the rapid development of modern survey research from about the mid-1930s through the 1940s and the establishment of survey research centers in both the academic and commercial sectors (Converse 1986). These centers include the Survey Research Center/Institute for Social Research (ISR) at the University of Michigan and the National Opinion Research Center (NORC) at the University of Chicago on the academic side and Gallup's American Institute of Public Opinion on the commercial side, to name just a few early organizations that are still leaders in the field.

Modern voting research based on the survey method typically uses small but randomly selected samples of about 1,000 (rarely more than 2,000) eligible voters. Information is collected through the use of standardized questionnaires that are administered by trained interviewers in person or increasingly over the telephone. Advances in modern communication technology such as the Internet are likely to change the face of scholarly survey research even more drastically in the very near future. “Standardized” means that a question's wording is predetermined by the researcher and that the interviewer is supposed to read questions exactly as stated and in the prearranged order. For the most part, the response alternatives also are predetermined (“closed questions”); sometimes, for select questions, verbatim answers are recorded (“open questions”) and subsequently sorted into a categorical scheme. In contrast to aggregate-level analysis and the use of official election returns, survey-based research on voting behavior relies on self-reports by individual citizens. Thus, it is subject to bias and distortion resulting from question wording, dishonest answers, memory failure, and unstable attitudes even if the sample is properly drawn. Its major advantage is the unequivocal linkage of demographic traits (e.g., age, sex, ethnicity, and social class) and political attitudes and behavior on the level of the individual.

### AGGREGATE DATA ANALYSIS

The use of aggregate data in studying voting behavior poses formidable methodological problems, yet it is the only approach available to study voting behavior before the mid-1930s. For example, the Germans voted Hitler into power in genuinely

democratic elections in the late 1920s and early 1930s. The voting behavior of Germans in the Weimar Republic has been subject to much debate and controversy in political sociology. The earlier consensus that Hitler's support came predominantly from the lower middle classes was challenged by later studies (e.g., Childers 1983; Falter 1991) that contended that his support had a much wider base cutting across all social groups.

Findings based on aggregate data analysis often depend heavily on seemingly technical details of preparing the database and the choice of specific statistical techniques. As a rule, findings are more reliable if the geopolitical units are small. However, even if the greatest care is exercised, there is always the danger of an "ecological fallacy." To use a contemporary example, if the vote for a white candidate increases with a rising percentage of white voters across voting districts, it is plausible to assume that people have voted along racial lines, yet this need not be the case. Perhaps ethnic minorities in predominantly white districts are more likely to vote for a white candidate than they are elsewhere. Therefore, they, not the white voters, may be responsible for the increased share of the white candidate.

A solution to the ecological inference problem has been proclaimed (King 1997), but despite some progress, that claim appears to be overstated. In spite of all the remaining shortcomings, though, aggregate data analysis is an indispensable tool for tracing patterns of voting behavior over time in a sociohistorical analysis (e.g., Silbey et al. 1978) or analyzing contemporary voting behavior when sufficiently detailed and reliable survey data are not available. Particularly for local or regional studies, there may not be sufficient funds to conduct appropriate surveys or the research interest may develop only after the elections have taken place.

### **SURVEY-BASED VOTING RESEARCH**

**The Columbia School.** Four landmark studies connected with the presidential elections of 1940, 1948, 1952, and 1956 mark the establishment of scholarly survey-based research on voting behavior (Rossi 1959). In essence, those studies provided the core concepts and models used in contemporary voting research. Reviewing those studies provides an introduction to present-day theories of voting behavior in U.S. presidential elections,

while congressional elections typically follow a very simple pattern: Incumbents are rarely defeated.

The first two studies were conducted by Lazarsfeld and his associates at Columbia University. Their main intention was to "relate preceding attitudes, expectations, personal contacts, group affiliations and similar data to the final decision" (Berelson et al. 1954, p.viii) and trace changes of opinion over the course of a campaign. Emphasizing the particular set of political and social conditions and its importance for this process, the Columbia group restricted its studies to one community (Erie County, Ohio, in 1940 and Elmira, New York, in 1948) and interviewed the same respondents repeatedly: up to seven times in 1940 and four times in 1948. Repeated interviews, or a "panel design," became a standard feature of more sophisticated voting studies, while the major studies to come abandoned the focus on one community in favor of nationwide representation.

Several major findings emerged from the Erie County study (Lazarsfeld et al. [1944] 1968). First, people tend to vote as they always have, in fact as their families have. In the Michigan school of voting behavior (see below), this attitude stability was conceptualized as "party identification," a stable inclination toward a particular party that for the most part develops during adolescence and early adulthood.

Second, attitudes are formed and reinforced by individuals' membership in social groups such as their social class, ethnic group, and religious group and by the associations they belong to. More concretely, the research team found that people of lower social status, people in urban areas, and Catholics tended to be Democrats while people of higher social class, people in rural areas, and Protestants were more likely to be Republicans. Subsequently, the alliance of particular segments of the population with specific parties was amply documented despite modifications in its particular form. More so than in the United States, voting behavior in the major European democracies (notably Britain and West Germany) could be explained largely by the links between social groups and particular parties (Lipset and Rokkan 1967), although those links have been weakening.

Third, change does occur, and people under cross-pressures are the most likely to change. A cross-pressure occurs when the set of different

group memberships provides conflicting stimuli. For example, in 1940, Protestant blue-collar workers experienced a pull toward the Republicans on the basis of their religious affiliation and a pull toward the Democrats because of their class position. In the United States today, the impact of religious affiliation is more complicated, but the general notion of cross-pressure remains important.

Fourth, Lazarsfeld and colleagues developed the concept of a “two-step flow of information.” According to this concept, most people are not directly persuaded by the mass media even if they are susceptible to change. Instead, they tend to follow opinion leaders, who are the informal leaders in the various social networks (family, friends, associates at the workplace) in which individuals are involved. These leaders pay close attention to the media; they disseminate and validate media messages. With the ever-increasing impact of mass media (television and more recently the Internet) over the last fifty years, this result of the 1940 study may not reflect the situation today. Unfortunately, it is very difficult to prove media effects conclusively, and the cumulative empirical evidence has not been able to settle a long-standing controversy about the extent of media effects.

The 1948 Elmira study was designed to test further and if necessary modify the findings of the earlier study and integrate the results into the body of existing knowledge (see Berelson et al. 1954, pp. 327–347, for a comparative synopsis of several major studies). As a matter of fact, its main contribution lies in the refinement of several aspects that were not covered sufficiently in the Erie County study. However, the Elmira study still failed to show systematically the links between the efforts of the various institutions in the community and the decisions of the voters. The focus on those links was the key rationale for limiting these studies to one community, a feature that invites doubt whether the findings can be generalized to all American voters.

**The Michigan School.** The sociological approach of the Columbia school was subsequently overshadowed by the social psychological model of the Michigan school that came to dominate survey-based voting research for many years. After a smaller study in 1948 (Campbell and Kahn 1952), the Michigan team, led by Campbell, conducted major studies in 1952 and 1956 (Campbell et al.

1954; Campbell et al. 1960). In contrast to Lazarsfeld and associates, their studies used national samples, thus expanding the geographic coverage, but only two interviews, one shortly before and one shortly after the elections. In addition, the Michigan group introduced far-reaching changes in the conceptualization of the voting process. On the basis of its national study of 1948, those researchers felt that social group memberships have little direct impact on the voting decision. Instead, they focused on “the psychological variables which intervene between the external events of the voter’s world and his ultimate behavior” (Campbell et al. 1954, pp. 85–86). In particular, they considered three concepts: party identification, issue orientation, and candidate orientation. *Party identification* refers to the sense of personal attachment an individual feels toward a party irrespective of formal membership or direct involvement in that party’s activities. It is thought of as a stable attitude that develops early in life. In contrast, both issue orientation and candidate orientation depend on the context of a particular election. *Issue orientation* refers to individuals’ involvement in issues they perceive as being affected by the outcome of an election. For example, if individuals are concerned about the economy and feel that it makes a difference whether the country has a Democratic or a Republican president, this will have an impact on their voting decisions. Similarly, *candidate orientation* refers to individuals’ interest in the personality of the candidates and to a possible preference that derives from the personal traits of the candidates. For example, Ronald Reagan portrayed himself as a firm and determined leader but also as a caring and understanding father. In that way, he was able to attract many voters otherwise attached to the Democrats.

The Michigan model posits a “funnel of causality.” The social factors emphasized by the Columbia school are not dismissed outright but are viewed as being at the mouth of the funnel, having an indirect effect only through the three central psychological variables, particularly party identification. Party identification in turn affects issue orientation and candidate orientation as well as having a direct effect on the voting decision. The simplicity of this model is both its strength and its weakness. It clearly marks the shift of emphasis to psychological processes of individual perception and evaluation, but it does not explicitly address

the social and political context. However, in *The American Voter* (Campbell et al. 1960), the Michigan group presents a much more comprehensive analysis of the 1956 elections, addressing topics such as the role of group membership, social class, and the political system without, however, explicitly expanding the basic model.

Additional concepts that have been used widely in subsequent research include the concept of a *normal vote* and the typology of elections as *maintaining*, *deviating*, or *realigning* (Campbell et al. 1966) and an assessment of mass belief systems (Converse 1964). The concept of a normal vote follows directly from the basic model: If all voters follow their long-standing inclinations (vote according to party identification), they produce a normal vote. Comparing actual election returns with the (hypothetical) normal vote allows one to assess the impact of contemporaneous, mostly short-term factors. In a maintaining election, the party with the larger number of partisans wins, but its vote share may be somewhat different from its normal share as a result of short-term factors. If short-term factors lead to the defeat of that party, the elections are considered deviating. Realigning elections mark a major shift in basic allegiances. Such shifts are rare and typically are not accomplished in a single election. In the 1930s, the American electorate shifted toward the Democrats as a consequence of economic depression and Roosevelt's New Deal, which promised a way out. However, given their long-term nature, processes of dealignment and realignment are difficult to determine in strict empirical terms (Dalton et al. 1984; Lawrence 1996).

With respect to the nature of mass belief systems, Converse's (1964) article triggered a long-lasting debate that has been settled. Converse asserted that the vast majority of the American people have little interest in politics, that their opinions on issues lack consistency and stability over time, and that those opinions are mostly "non-attitudes." Consequently, a large portion of the electorate does not vote at all; if those people do vote, their vote is based mostly on partisanship and/or a candidate's personality, not on an independent and careful evaluation of the issues.

**Critique of *The American Voter* and subsequent refinements.** Like other landmark empirical studies, *The American Voter* was not exempt

from sometimes radical critiques that can be grouped into three categories: challenges to the allegedly derogatory image of the American electorate and its implications for the democratic process, assertions that the findings are valid only for the 1950s, and a methodological critique of operationalization, measurement, and model specification. Most of the methodological critique is too technical to be discussed in this article (see, Asher 1983).

One of the earliest and most vocal critics was Key (1966). Using a reanalysis of Gallup data from 1936 to 1960, he developed a typology of "standpatters," "switchers," and "new voters" and asserted that the global outcome of those elections followed a rational pattern derived from an appraisal of past government performance. Hence, as a whole, the electorate acts responsibly despite the fact "that many individual voters act in odd ways indeed" (Key 1966, p. 7).

The most comprehensive effort to review American voting behavior over time, a critique of the second type, was presented by Nie et al. (1976), based on the series of Michigan election studies from 1952 to 1972. Still working within the framework of the Michigan model, those authors found significant changes in the relative importance of its three central factors: a steady decrease in the level of party identification, particularly among younger groups, and a much stronger relative weight of issue orientation and candidate orientation. In a turbulent period of internal strife and social change (civil rights, the Vietnam War, Watergate), the electorate became more aware of issues and much more critical of parties and the established political process. Nie et al. (1976) found a decomposition of the traditional support bases for both Democrats and Republicans, all adding up to an "'individuation' of American political life" (p. 347).

Still largely following the path of the Michigan school, much research in the 1980s was directed toward issue voting, which reflects a continuing decline in stable party attachments through political socialization and/or group memberships. In particular, the impact of economic conditions on electoral outcomes was investigated both in the United States and in other major Western democracies (Eulau and Lewis-Beck 1985; Lewis-Beck 1988; Norpoth et al. 1991). The findings were

diverse, contingent on the specification of the research question and the national context, yet one general pattern emerged and was confirmed beyond the United States in many national elections in the 1990s: Political actors perceived as better able to handle economic matters than their competitors have a significant advantage, and perceived economic competence is strongly related to an individual's voting decision.

Fiorina's (1981) concept of *retrospective voting* can be seen as a bridge between the classic Michigan model and the newer directions that have emerged in the last two decades. Fiorina posits that both party identification and issue orientation are largely dependent on the evaluation of past government performance. Party identification thus represents a sort of running tally of past experience. It is still a long-term influence, but it is subject to gradual change and is based more on cognition than on affection.

**More Recent Approaches.** With some simplification, one can discern three major directions in voting research in the last fifteen to twenty years, each anchored in a different discipline: Rooted in economic theory, rational choice models have been applied to voting (as well as to many other forms of social behavior); drawing on more general psychological theories, the subfield of political psychology studies in a comprehensive way how the individual perceives and processes political information, with voting being only one specific aspect; and the focus on reference groups on both the macro level such as class or religion (an important strand in European research on voting) and the micro level (social networks) has reemphasized the sociological perspective.

Rational choice models see the voter as carefully evaluating the pros and cons of each party or candidate, assessing its utility (consumer models as in Himmelweit et al. 1985) and proximity to the voter's own position (spatial models as in Enelow and Hinich 1984, 1990; directional models as in Rabinowitz and MacDonald 1989), and then voting for the closest or most useful party and/or candidate. It is doubtful, however, whether such models adequately portray the actual process of reaching a voting decision except among the small segment of highly informed and highly motivated citizens. The rational choice approach has been criticized more generally as conceptually inadequate

for the analysis of mass political behavior (e.g., Green and Shapiro 1994). Other criticism has come from within the rational choice camp, leading, for example, to an "expressive model of voting" (Brennan and Lomasky 1993) that shifts the focus away from the instrumental aspect of voting but maintains the conceptual and terminological framework.

The literature in political psychology is too extensive to be discussed in detail here; overviews are provided by Lau and Sears (1986), Sniderman et al. (1991), and Mutz et al. (1996), among others. However, the contribution of this approach to the discussion of two long-standing controversial topics must be mentioned explicitly: media effects and political belief systems. In regard to media effects, a number of studies using more refined concepts (agenda setting, framing, priming) have suggested a more direct and stronger impact of media than previously had been assumed (see Kinder 1998; Zaller 1996), but a generally accepted model of how the mass media influence the political process has not emerged. In regard to political belief systems or the lack thereof, the notion of a politically uninformed or ignorant and irrational electorate has strong implications for normative theories of democracy and, on a practically level, campaign strategists. There continues to be strong and largely undisputed empirical evidence that many Americans have rather limited factual knowledge of specific bills and policies, the makeup of political institutions, and even their elected representatives. However, it is controversial whether that lack of specific knowledge matters and whether voters are nevertheless able to make "correct" or "reasonable" choices on the basis of more implicit forms of information processing (Delli Carpini and Keeter 1996, Lau and Redlawsk 1997). In both areas, political psychology has not settled the controversy, but it has provided enhanced conceptualizations to guide more productive empirical analyses.

In regard to sociological perspectives, community studies (as in the Columbia school) have attempted to assess the impact of the local context on the decision making of the individual. Context information is gathered by using block-level census data or tracing and interviewing members of the social network of the primary respondents (Huckfeldt 1986; Huckfeldt and Sprague 1995).

On the macro level, the social cleavage approach, long dominant in European voting research (see below), has gained some attention in the United States (see Brooks and Manza 1997).

**Voting Behavior in Other National Contexts.**

The Michigan school of voting behavior has continued to have a major impact on the emergence of survey-based voting research in other Western democracies (see Beyme and Kaase 1978; Butler and Stokes 1976; Heath et al. 1985; Rose 1974) and electoral research today (e.g., Kaase and Klingemann 1998). A strict replication of the basic model, however, has rarely been feasible because of considerable differences in political systems, party organizations, and electoral rules. In particular, attempts to devise valid measures of the key concept of party identification have produced mixed results at best (e.g., Budge et al. 1976).

Despite considerable variation across European democracies and within specific countries, political parties articulate specific programmatic positions derived from basic ideological beliefs that bind all party members, and parliamentary votes typically follow party lines. Political parties, then, dominate the political contest, and most major parties have their roots in social cleavages relating to class, religion, region, or ethnicity. Consequently, such social group memberships have been powerful determinants of voting behavior (Lipset and Rokkan 1967). Much of the European debate in the last twenty years, however, has focused on evolving changes in the electorate that result in more volatile voters. First, the impact of social (class) origin on the individual life course in general (educational attainment, occupational opportunities, marriage and family, etc.) has declined, leading to a more idiosyncratic definition of self-interest and as a consequence a less predetermined voting pattern. Second, a growing disenchantment with established political parties and politicians in general has weakened the once highly internalized norm of political participation through voting and increased the propensity to vote for new and often extremist parties as a token of protest. Thus, the once stable alliances between parties and certain segments in the electorate along social cleavages have weakened (Crewe and Denver 1985; Franklin et al. 1991; Miller et al. 1990), though there is no consensus whether "class voting" has ended (e.g., Manza et al. 1995).

The emergence of new democracies after the demise of the Communist bloc in eastern Europe in the late 1980s and the continuing process of European integration, including the creation of European Union as a single political and not just economic entity, have created new challenges and opportunities for voting research in Europe. With the introduction of a common currency in 1999, a continuing reduction of national sovereignty, and a generally strengthened position of the European Parliament in exercising control over the executive branch (European Commission), European elections will lose their stigma as second-order elections. Studies of European elections in 1974, 1979, and 1994 suffered from limited funding and broke no new conceptual ground. However, they did establish the fact that national rather than pan-European issues dominated voting choices (Schmitt and Mannheim 1990; Van der Eijk and Franklin 1996). This is likely to change, but it is too early to tell the role these elections will play in the further process of European unification and whether voting behavior in those elections will follow patterns different from those in national elections.

The methodological problems of the European election studies in achieving functional equivalence of the instruments (questions) and ensuring uniform quality standards in sampling and interviewing are even more formidable in the new eastern European democracies. Apart from the technical aspects of conducting valid and reliable surveys, there may be inherent limits to exporting the survey method (Bulmer 1998). Rather than relying solely on survey data and taking them at face value, analysts must employ a more comprehensive approach that integrates quantitative and qualitative research, such as the one used by White et al. (1997) in their study of Russian voters. In the absence of a stable party structure and stable political institutions more generally and with the relative novelty of voting in free elections, it is unrealistic to expect the Michigan model or any other "reductionist" model to provide an adequate description of voting behavior in eastern Europe.

Beyond Europe, it is doubtful whether the concept of voting is functionally equivalent to voting in the United States or western Europe because of systemic differences and traditions. For example, even if one accepts the premise that Japan is a pluralist democracy (Richardson 1997),

an intricate net of mutual obligations governs much social behavior, including voting, and for many years relevant electoral competition occurred only within one dominant party. Still, the Michigan model has guided much Western research on Japanese voting behavior (Flanagan et al. 1991).

### OUTLOOK

Compared to the United States, research on voting behavior in western Europe has been tied more closely to the study of mass political behavior in general, satisfaction with democracy, the parties, the politicians, and, the stability of the political system despite a continuing orientation toward the Michigan model and its variants. What European research has to offer are not better micro models of voting behavior but detailed trend analyses of cross-national comparative data that put voting into a broader context of political behavior (e.g., Klingemann and Fuchs 1995). As a result of the systemic differences within Europe and the methodological limitations of the database, no unified theory of voting behavior has emerged. For each individual (national) election, there is a plausible explanation, at least in retrospect, drawing on well-recognized factors such as perceived economic competence and leadership image, but the relative weight of each factor varies from one election to the next. As information and communication behavior is undergoing drastic changes as a result of technological advances (Internet) and the ever-increasing presence of the media, it will become even more difficult to determine the relative weight of each factor in a voter's choice. The Columbia studies dealt with voters in a relatively contained world in which the amount of information and the number of transmission channels were limited; voters in the twenty-first century will be faced with an overload of information and will be as closely connected to the political world (or its competing representations) as they want to be. At best, more knowledgeable voters will be able to make better informed choices; at worst, more perplexed voters will succumb to the most skillful public relations managers. Somehow voters will have to reduce this complexity, and turning to social groups for guidance is a likely strategy. While it will be important to continue to study the processes of cognition and information processing, the real key to voters' "rationality" may lie in their social relations. More than ever, the study of vot-

ing behavior will have to build bridges between the various disciplines and incorporate the sociological perspective.

### REFERENCES

- Asher, Herbert B. 1983 "Voting Behavior Research in the 1980s." In Ada W. Finifter, ed., *Political Science*. Washington, D.C.: American Political Science Association.
- Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee 1954 *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.
- Beyme, Klaus von, and Max Kaase (eds.) 1978 *Elections and Parties*. London: Sage.
- Brennan, Geoffrey, and Loren Lomasky 1993 *Democracy and Decision: The Pure Theory of Electoral Preference*. New York: Cambridge University Press.
- Brooks, Clem, and Jeff Manza 1997 "Social Cleavages and Political Alignments: U.S. Presidential Elections, 1960 to 1992." *American Sociological Review* 62:937-946.
- Budge, Ian, Ivor Crewe, and Dennis Fairlie, eds. 1976 *Party Identification and Beyond*. London: Wiley.
- Bulmer, Martin (ed.) 1998 "Exporting Social Survey Research." *American Behavioral Scientist* (special issue) 42(2).
- Butler, David, and Donald Stokes 1976 *Political Change in Britain*, 2nd ed. New York: St. Martin's.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes 1966 *Elections and the Political Order*. New York: Wiley.
- 1960 *The American Voter*. New York: Wiley.
- , Gerald Gurin, and Warren E. Miller 1954 *The Voter Decides*. Evanston, Ill.: Row, Peterson.
- , and Robert L. Kahn 1952 *The People Elect a President*. Ann Arbor, Mich.: Institute for Social Research.
- Childers, Thomas 1983 *The Nazi Voter*. Chapel Hill: University of North Carolina Press.
- Converse, Jean M. 1986 *Survey Research in the United States: Roots and Emergence*. Berkeley: University of California Press.
- Converse, Philip 1964 "The Nature of Belief Systems in Mass Publics." In David Apter, ed., *Ideology and Discontent*. New York: Free Press.
- Crewe, Ivor, and David Denver (eds.) 1985 *Electoral Change in Western Democracies: Patterns and Sources of Electoral Volatility*. New York: St. Martin's.



- Dalton, Russell J., Scott C. Flanagan, and Paul Allen Beck, eds. 1984 *Electoral Change in Advanced Industrial Democracies: Realignment or Dealignment?* Princeton, N.J.: Princeton University Press.
- Delli Carpini, Michael X., and Scott Keeter 1996 *What Americans Know about Politics and Why it Matters*. New Haven, Conn.: Yale University Press.
- Enelow, James M., and Melvin J. Hinich 1984 *The Spatial Theory of Voting*. New York: Cambridge University Press.
- 1990 *Advances in the Spatial Theory of Voting*. New York: Cambridge University Press.
- Eulau, Heinz, and Michael Lewis-Beck (eds.) 1985 *Economic Conditions and Electoral Outcomes: The United States and Western Europe*. New York: Agathon Press.
- Falter, Jürgen W. 1991 *Hitler's Wähler*. Munich: Beck.
- Fiorina, Morris P. 1981 *Retrospective Voting in American National Elections*. New Haven, Conn.: Yale University Press.
- Flanagan, Scott C., Shinsaku Kohei, Ichiro Miyake, Bradley M. Richardson, and Joji Watanuki 1991 *The Japanese Voter*. New Haven, Conn.: Yale University Press.
- Franklin, Mark, Tom Mackie, and Henry Valen, eds. 1991. *Electoral Change: Responses to Evolving Social and Attitudinal Structures in Seventeen Democracies*. Cambridge, UK: Cambridge University Press.
- Gallup, George (1944) 1948 *A Guide to Public Opinion Polls*. Princeton, N.J. Princeton University Press.
- Gosnell, Harold 1930 *Why Europe Votes*. Chicago: University of Chicago Press.
- Green, Donald P., and Ian Shapiro 1994 *Pathologies of Rational Choice Theory*. New Haven, Conn.: Yale University Press.
- Heath, Anthony, Roger Jowell, and John Curtice 1985 *How Britain Votes*. New York: Pergamon Press.
- Himmelweit, Hilde, Patrick Humphreys, and Marianne Jaeger 1985 *How Voters Decide*. Milton Keynes, UK: Open University Press.
- Huckfeldt, Robert 1986 *Politics in Context: Assimilation and Conflict in Urban Neighborhoods*. New York: Agathon Press.
- , and John Sprague 1995 *Citizens, Politics, and Social Communication*. New York: Cambridge University Press.
- Jensen, Richard 1969 "American Election Analysis: A Case History of Methodological Innovation and Diffusion." In Seymour M. Lipset, ed., *Politics and the Social Sciences*. New York: Oxford University Press.
- Kaase, Max, and Hans-Dieter Klingemann (eds.) 1998. *Wahlen und Wähler: Analysen aus Anlass der Bundestagswahl 1994*. Opladen, Germany: Westdeutscher Verlag.
- Key, Valdimer Orlando, Jr. 1966 *The Responsible Electorate*. Cambridge, Mass.: Belknap Press, Harvard University Press.
- Kinder, Donald 1998 "Communication and Opinion." *Annual Review of Political Science* 1:167–197.
- King, Gary S. 1997 *A Solution to the Ecological Inference Problem*. Princeton, N.J. Princeton University Press.
- Klingemann, Hans-Dieter, and Dieter Fuchs 1995 *Citizens and the State* (vol. 1 of Max Kaase and Kenneth Newton, eds., *Beliefs in Government*). Oxford, UK: Oxford University Press.
- Lau, Richard, and David P. Redlawsk 1997 "Voting Correctly." *American Political Science Review* 91:585–598.
- , and David Sears 1986 *Political Cognition*. Hillsdale, N.J.: Erlbaum.
- Lawrence, David G. 1996 *The Collapse of the Democratic Presidential Majority*. Boulder, Colo.: Westview.
- Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet (1944) 1968 *The People's Choice*. 3rd ed. New York: Columbia University Press.
- Lewis-Beck, Michael 1988 *Economics and Elections: The Major Western Democracies*. Ann Arbor: University of Michigan Press.
- Lipset, Seymour Martin, and Stein Rokkan (eds.) 1967 *Party Systems and Voter Alignments*. New York: Free Press.
- Manza, Jeff, Michael Hout, and Clem Brooks 1995 "Class Voting in Capitalist Democracies since World War II." *Annual Review of Sociology* 21:137–162.
- Miller, William L., Harold D. Clarke, Lawrence Leduc, and Paul Whiteley 1990 *How Voters Change*. New York: Oxford University Press.
- Mutz, Diana C., Paul M. Sniderman, and Richard A. Brody (eds.) 1996 *Political Persuasion and Attitude Change*. Ann Arbor: University of Michigan Press.
- Nie, Norman H, Sidney Verba, and John R. Petrocik 1976 *The Changing American Voter*. Cambridge, Mass.: Harvard University Press.
- Norpoth, Helmut, Michael Lewis-Beck, and Jean-Dominique Lafay (eds.) 1991 *Making Governments Pay*. Ann Arbor: University of Michigan Press.
- Rabinowitz, George, and Stuart Elaine MacDonald 1989 "A Directional Theory of Issue Voting." *American Political Science Review* 83:93–121.
- Rice, Stuart A. 1928 *Quantitative Methods in Politics*. New York: Knopf.
- Richardson, Bradley 1997 *Japanese Democracy*. New Haven, Conn.: Yale University Press.

- Rose, Richard (ed.) 1974 *Electoral Behavior: A Comparative Handbook*. New York: Free Press.
- Rossi, Peter A. 1959 "Four Landmarks in Voting Research." In Eugene Burdick and Arthur Brodbeck, eds., *American Voting Behavior*. Glencoe, Ill.: Free Press.
- Schmitt, Hermann, and Renato Mannheimer (eds.) 1990. "The European Elections of 1989." *European Journal of Political Research* (special issue) 19(1).
- Silbey, Joel H., Allan G. Bogue, and William H. Flanigan, eds. 1978 *The History of American Electoral Behavior*. Princeton, N.J.: Princeton University Press.
- Snidermann, Paul M., Richard A. Brody, and Philip E. Tetlock 1991 *Reasoning and Choice*. New York: Cambridge University Press.
- Van der Eijk, Cees, and Mark N. Franklin 1996 *Choosing Europe?* Ann Arbor: University of Michigan Press.
- White, Stephen, Richard Rose, and Ian McAllister 1997 *How Russia Votes*. Chatham, N.J.: Chatham House.
- Zaller, John 1996 "The Myth of Massive Media Impact Revived: New Support for a Discredited Idea." In Diana C. Mutz, Paul M. Sniderman, and Richard A. Brody, eds., *Political Persuasion and Attitude Change*. Ann Arbor: University of Michigan Press.

MANFRED KUECHLER

# W

---

## WAR

The ubiquity and importance of war have made analyses of its causes a central concern of scholars for over two millennia. Many of the fundamental questions about the causes of war were raised by Thucydides in the fifth century B.C., but the vast amount of work on the topic since that time has produced ongoing debates instead of generally accepted answers. Studies of war can be divided into three broad categories (reviews of the literature using similar frameworks are provided by Waltz 1959; Bueno de Mesquita 1980; and Levy 1989). The first type takes the system as whole as the unit of analysis and focuses on how characteristics of the interstate system affect the frequency of war. States are the unit of analysis in the second type, which explores the relationships among the political, economic, and cultural features of particular states and the propensity of states to initiate wars. The third type analyzes war as an outcome of choices resulting from small group decision making.

Some debates focus on characteristics of the interstate system that are thought to increase or decrease the chance of war. Are wars more likely during a period of economic prosperity or one of economic contraction? Which is more likely to maintain peace, a balance of power in the international system or a situation in which one state is hegemonic? Has the increasing power of transnational organizations such as the United Nations changed the likelihood of war in the contemporary world?

Social scientists also disagree about the effects of political and economic factors within a state on

the possibility of war. Does a capitalist economy make a state more or less likely to initiate wars? Do democratic states start wars less often than autocracies do? Is increasing nationalism likely to cause more wars? Is the ethnic composition within and between states an important determinant of war?

There is also no consensus on which model of individual decision making is most appropriate for the study of war. Is the decision to go to war based on a rational calculation of economic costs and benefits, or is it an irrational outcome of distortion in decision making in small groups and bureaucracies? Are wars based on nationalist, ethnic, or religious conflicts generated more by emotions or values than by rational choices?

## THE INTERSTATE SYSTEM AND WAR

Most studies of war that use the interstate system as the unit of analysis begin with assumptions from the “realist” paradigm. States are seen as unitary actors in realist theories, and their actions are explained in terms of the structural characteristics of the system. The most important feature of the interstate system is that it is anarchic. Unlike politics within states, relations between states take place in a Hobbesian state of nature. Since an anarchic system is one in which all states constantly face actual or potential threats, their main goal is security. Security can be achieved in such a system only by maintaining power. In realist theories, the distribution of power in the interstate system is the main determinant of the frequency of war.

Although all realist theories agree on the importance of power distribution in determining war, they disagree about which types of power distributions make war more likely. Balance-of-power theories (Morgenthau 1967) suggest that an equal distribution of power in the system facilitates peace and that an unequal distribution leads to war. They argue that parity deters all states from aggression and that an unequal power distribution generally will result in the strong using force against the weak. When one state begins to gain a preponderance of power, a coalition of weaker states will form to maintain their security by blocking the further expansion of the powerful state. The coalitions that formed against Louis XIV, Napoleon, and Hitler seem to fit this pattern.

Hegemonic stability theory (Gilpin 1981) suggests exactly the opposite: that unequal power in the system produces peace while parity results in war. When one state has hegemony in the world system, it has both the incentive and the means to maintain order. It is not necessary for the most powerful state to fight wars, since its objectives can be achieved in less costly ways, and it is not rational for other states to challenge a state with overwhelming power. Gilpin notes that the periods of British and U.S. hegemony were relatively peaceful and that World Wars I and II occurred during intervening periods in which power was distributed more equally. Since balance-of-power and hegemonic stability theories seem to explain some but not all of the cases, what is needed is a theory specifying the conditions under which either parity or hegemony leads to war.

Balance-of-power and hegemonic stability arguments are not applicable to all wars, only those between great powers. A third attempt to explain great-power war is power transition theory (Organski 1968). This theory suggests that differential rates of economic growth create situations in which rising states rapidly catch up with the hegemonic state in the system and that this change in relative power leads to war. Organski argues that the rising state will initiate a war to displace the hegemonic state. This final part of the argument is questionable, since it seems at least as plausible that the hegemonic state will initiate the war against the rising challenger to keep the small advantage it still has (Levy 1989, p. 253).

Debates about power transitions and hegemonic stability are of much more than theoretical interest in the contemporary world. Although the demise of the Soviet Union has left the United States as an unchallenged military hegemon, American economic superiority is being challenged by the European Union (EU) and emerging Asian states (Japan in the short run, perhaps China in the long run). If power transition and hegemonic stability theories are correct, this shift of economic power could lead to great-power wars in the near future. If the main challenge is from the EU (the most likely scenario), it will be interesting to see if the cultural heritage of cooperation between the United States and most of Europe will be sufficient to prevent the great-power war that some theories predict.

Another ongoing debate about systemic causes of war concerns the effects of long cycles of economic expansion and contraction. Some scholars argue that economic contraction increases the chance of war, since the increased scarcity of resources leads to more conflict. Others have suggested the opposite: Major wars are more common in periods of economic expansion because only then do states have the resources necessary to fight. Goldstein's (1988) research suggests that economic expansion tends to increase the severity of great-power wars but that economic cycles have no effect on the frequency of war.

Two important changes in the last fifty years may make many systemic theories of war obsolete (or at least require major revisions). The first is technology. Throughout history, technological changes have determined the general nature of warfare. By far the most significant recent development has been the availability of nuclear weapons. Since the use of these weapons would result in "mutually assured destruction," they may have made war much less likely by making it irrational for both parties. Of course, the broadening proliferation of nuclear weapons raises serious problems, as does their existence in currently unstable states such as the Russian federation. A second technological change that may alter the nature of war is increasing dependence on computers. Although computers have increased the accuracy and precision of many types of military technology, they also leave the countries using them vulnerable to new kinds of attacks by "hackers" who could not only disarm military operations but

bring whole economies to a halt by disrupting the computer systems necessary for their operation.

The second significant change in the last half of the twentieth century has been the development and increasing power of transnational organizations such as the United Nations. Most theories of war begin with the assumption that the interstate system is anarchic, but this is no longer valid. If the military power of the United Nations continues to grow, that organization could become more and more effective at preventing wars and suppressing them quickly when they start. Of course, it remains to be seen whether powerful existing states will cede more power to such institutions.

Theoretical debates about the systemic causes of war have not been resolved, in part because the results of empirical research have been inconclusive. To take one example, equality of power in the interstate system decreased the number of wars in the nineteenth century and increased the number in the twentieth century. Proponents of each theory can point to specific cases that seem to fit its predictions, but they must admit that there are many cases it cannot explain. At least part of the problem is that systemic theories have not incorporated causal factors at lower levels of analysis, such as the internal economic and political characteristics of states. Since the effects of system-level factors on war are not direct but always are mediated by the internal political economy of states and the decisions made by individual leaders, complete theories of the causes of war must include these factors as well.

### **CAPITALISM, DEMOCRACY, AND WAR**

One of the longest and most heated debates about the causes of war concerns the effects of capitalism. Beginning with Adam Smith, liberal economists have argued that capitalism promotes peace. Marxists, by contrast, suggest that capitalism leads to frequent imperialist wars.

Liberal economic theories point to the wealth generated by laissez-faire capitalist economies, the interdependence produced by trade, and the death and destruction of assets caused by war. Since capitalism has increased both the benefits of peace (by increasing productivity and trade) and the costs of war (by producing new and better instru-

ments of destruction), it is no longer rational for states to wage war. The long period of relative peace that followed the triumph of capitalism in the nineteenth century and the two world wars that came after the rise of protectionist barriers to free trade often are cited in support of liberal economic theories, but those facts can be explained by hegemonic stability theorists as a consequence of the rise and decline of British hegemony.

In contrast to the sanguine views of capitalism presented by liberal economic theories, Marxists argue that economic problems inherent in advanced capitalist economies create incentives for war. First, the high productivity of industrial capitalism and a limited home market resulting from the poverty of the working class result in chronic "underconsumption" (Hobson [1902] 1954). Capitalists thus seek imperial expansion to control new markets for their goods. Second, Lenin ([1917] 1939) argued that capitalists fight imperialist wars to gain access to more raw materials and find more profitable outlets for their capital. These pressures lead first to wars between powerful capitalist states and weaker peripheral states and then to wars between great powers over which of them will get to exploit the periphery.

In contrast to the stress on the political causes (power and security) of war in most theories, the Marxist theory of imperialism has the virtue of drawing attention to economic causes. However, there are several problems with the economic causes posited in theories of imperialism. Like most Marxist arguments about politics, theories of imperialism assume that states are controlled directly or indirectly by dominant economic classes and thus that state policies reflect dominant class interests. Since states are often free of dominant class control and since many groups other than capitalists often influence state policies, it is simplistic to view war as a reflection of the interests of capitalists. Moreover, in light of the arguments made by liberal economists, it is far from clear that capitalists prefer war to other means of expanding markets and increasing profits.

With the increasing globalization of economies and the transition of more states to capitalist economies, the debates about the effects of capitalism, trade, and imperialism on war have become increasingly significant. If Adam Smith is right, the future is likely to be more peaceful than

the past, but if Marxist theorists are right, there will be an unprecedented increase in economically based warfare.

The form of government in a country also may determine how often that country initiates wars. Kant ([1795] 1949) argued that democratic states (with constitutions and separation of powers) initiate wars less often than do autocratic states. This conclusion follows from an analysis of who pays the costs of war and who gets the benefits. Since citizens are required to pay for war with high taxes and their lives, they will rarely support war initiation. Rulers of states, by contrast, have much to gain from war and can pass on most of the costs to their subjects. Therefore, when decisions about war are made only by rulers (in autocracies), war will be frequent, and when citizens have more control of the decision (in democracies), peace generally will be the result.

Empirical research indicates that democratic states are less likely than are nondemocratic states to initiate wars, but the relationship is not strong (Levy 1989, p. 270). Perhaps one reason for the weakness of the relationship is that the assumption that citizens will oppose war initiation is not always correct. Many historical examples indicate that in at least some conditions citizens will support war even though it is not in their economic interest to do so. Nationalism, religion, ethnicity, and other cultural factors often are cited as important causes of particular wars in journalistic and historical accounts, but there still is no general theory of the conditions in which these factors modify or even override economic interests. Many classical sociological arguments suggested that these "premodern" and "irrational" sources of war would decline over time, but the late twentieth century has demonstrated the opposite. Nationalist and ethnic wars have become more common and intense. This raises the general issue of the factors affecting the choices individuals make about war initiation: Can these factors be modeled as rational maximization of interests, or is the process more complex?

### DECISION MAKING AND WAR

Although the assumptions may be only implicit or undeveloped, all theories of war must contain some assumptions about individual decision making. However, few theories of war focus on the individual level of analysis. One notable exception

is the rational-choice theory of war developed and tested by Bueno de Mesquita (1981).

Bueno de Mesquita begins by assuming that the decision to initiate war is made by a single dominant ruler who is a rational expected-utility maximizer. Utilities are defined in terms of state policies. Rulers fight wars to affect the policies of other states, essentially to make other states' policies more similar to their interests. Rulers calculate the costs and benefits of initiating war and the probability of victory. War is initiated only when rulers expect a net gain from it.

This parsimonious set of assumptions has been used to generate several counterintuitive propositions. For example, common sense might suggest that states would fight their enemies and not their allies, but Bueno de Mesquita argues that war will be more common between allies than between enemies. Wars between allies are caused by actual or anticipated policy changes that threaten the existing relationship. The interventions of the United States in Latin America and of the Soviet Union in eastern Europe after World War II illustrate the process. Other counterintuitive propositions suggest that under some conditions a state may rationally choose to attack the stronger of two allied states instead of the weaker, and under some conditions it is rational for a state with no allies to initiate a war against a stronger state with allies (if the distance between the two is great, the weaker state will be unable to aid the stronger). Although these propositions and others derived from the theory have received strong empirical support, many have argued that the basic rational-choice assumptions of the theory are unrealistic and have rejected Bueno de Mesquita's work on those grounds.

Other analyses of the decision to initiate war focus on how the social features of the decision-making process lead to deviations from rational choice. Allison (1971) notes that all political decisions are made within organizations and that this setting often influences the content of decisions. He argues that standard operating procedures and repertoires tend to limit the flexibility of decision makers and make it difficult to respond adequately to novel situations. Janis (1972) focuses on the small groups within political organizations (such as executives and their cabinet advisers) that actually make decisions about war. He suggests

that the cohesiveness of these small groups often leads to a striving for unanimity that prevents a full debate about options and produces a premature consensus. Other scholars have discussed common misperceptions that distort decisions about war, such as the tendency to underestimate the capabilities of one's enemies and overestimate one's own. In spite of these promising studies, work on deviations from rational choice is just beginning, and there still is no general theoretical model of the decision to initiate war.

### CONCLUSION

The failure to develop a convincing general theory of the causes of war has convinced some scholars that no such theory is possible, that all one can do is describe the causes of particular wars. This pessimistic conclusion is premature. The existing literature on the causes of war provides several fragments of a general theory, many of which have some empirical support. The goal of theory and research on war in the future will be to combine aspects of arguments at all three levels of analysis to create a general theory of the causes of war.

(SEE ALSO: *Global Systems Analysis: Peace; Revolutions; Terrorism*)

#### REFERENCES

- Allison, Graham 1971 *Essence of Decision*. Boston: Little, Brown.
- Bueno de Mesquita, Bruce 1980 "Theories of International Conflict: An Analysis and Appraisal." In Ted Robert Gurr, ed., *Handbook of Political Conflict*. New York: Free Press.
- 1981 *The War Trap*. New Haven, Conn.: Yale University Press.
- Gilpin, Robert 1981 *War and Change in World Politics*. Cambridge, UK: Cambridge University Press.
- Golstein, Joshua 1988 *Long Cycles*. New Haven, Conn.: Yale University Press.
- Hobson, J. A. (1902) 1954 *Imperialism*. London: Allen and Unwin.
- Janis, Irving 1972 *Victims of Groupthink*. Boston: Houghton Mifflin.
- Kant, Immanuel (1795) 1949 "Eternal Peace." In C. J. Friedrich, ed., *The Philosophy of Kant*. New York: Modern Library.

Lenin, V. I. (1917) 1939 *Imperialism*. New York: International.

Levy, Jack S. 1989 "The Causes of War: A Review of Theories and Evidence." In Philip E. Tetlock, Robert Jarvis, Paul Stern, and Charles Tilly, eds., *Behavior, Society and Nuclear War*. Oxford, UK: Oxford University Press.

Morgenthau, Hans 1967 *Politics among Nations*. New York: Knopf.

Organski, J. F. K. 1968 *World Politics*. New York: Knopf.

Waltz, Kenneth 1959 *Man, the State, and War*. New York: Columbia University Press.

EDGAR KISER

### WELFARE

See Poverty; Public Policy Analysis; Social Security Systems.

### WELL-BEING

See Quality of Life.

### WHITE-COLLAR CRIME

In his 1939 presidential address at the annual meeting of the American Sociological Society, Edward H. Sutherland used, with great effect, the term "white-collar crime." In an interesting introduction to his discussion of Sutherland, Green (1990) noted that in 1901, 1907, and 1935, respectively, Charles Henderson, Edward Alsworth Ross, and Albert Morris had "anticipated" the ideas Sutherland had presented, after conducting much research, in 1939. Sutherland depicted a white-collar criminal as any person of high socioeconomic status who commits a legal violation in the course of his or her occupation (Green 1990). Later he defined white-collar crime as criminal acts committed by persons in the middle or upper socioeconomic groups in connection with their occupations (Sutherland 1949). Since that time, the concept has undergone some modification and "has gained widespread popularity among the public" but "remains ambiguous and controversial in criminology" (Vold and Bernard 1986). More specifically, some definitions have deleted the class of the offender as a consideration.

Edelhertz (1970) defines white-collar crime as an “illegal act or series of illegal acts committed by nonphysical means and by concealment or guile, to obtain money or property, to avoid payment or loss of money or property, or to obtain business or personal advantage.”

Others have attempted to refine the definition by differentiating between occupational and corporate contexts (Clinard and Quinney 1973; Clinard and Yeager 1980), clarifying the difference between and among government, corporate, organizational, and occupational crimes (Coleman 1994; Green 1990; Punch 1996) and avoiding the use of the term “crime” by substituting “law violations” that involve the violator’s position of power (Biderman and Reiss, 1980). Tappan (1947) said that white-collar crimes are not “crimes” if they are not included in legal definitions.

While various writers measure the extent of white-collar crimes in terms of the number of “violations” (Clinard and Yeager 1980) or the extent of harm done to the public, business, or the environment (Punch 1996), others focus on dollar costs. Reiman (1995) modified the cost estimates of white-collar crimes by the U.S. Chamber of Commerce (1974) and made those figures applicable to 1991. Among his categories were consumer fraud, credit card and check fraud, embezzlement and pilferage, insurance fraud, receiving stolen property, and securities theft and fraud. His total amounted to \$197.76 billion for 1991. Reiman noted that those figures compared favorably with Clinard’s estimate of \$200 billion for one year (1990) and with a similar figure reported in *U.S. News and World Report* (1985). Reiman notes that his own estimate, is on the conservative side but that it is “almost 6000 times the total amount taken in all bank-robberies in the US in 1991 and more than eleven times the total amount stolen in all thefts reported in the FBI Uniform Crime Reports for that year” (1995, p.111). Reiman’s figures apparently do not include any of the vast amounts of money lost in the savings and loan scandal. These issues are elaborated on below.

Green (1990) states that Sutherland’s three main objectives were to show that: (1) white-collar crime is real criminality because it is law-violative behavior (Sutherland asserted that civil lawsuits resulting in decisions against persons or corporations should be considered convictions and are

proof of violations of law), (2) poor people are not the only ones who commit crime, and (3) his theory of differential association constituted an approach that could explain a general process characteristic of all criminality.

Sutherland held that typical crime “statistics” picturing the criminal population as made up largely of lower-class, economically underprivileged people give a false impression of noncriminality on the part of the upper classes, including respected and highly placed business and political persons (Sutherland 1949). His white-collar criminality included some of the following: misinterpretation of the financial statements of corporations, manipulation of the stock exchange, bribery of public officials to obtain desirable contracts, misrepresentation in advertising and salesmanship, embezzlement and misuse of trust funds, dishonest bankruptcies, and price-fixing. He quoted the Chicago gangster Al Capone calling such practices “legitimate rackets” (Sutherland 1949) to differentiate them from the more violent rackets of the underworld.

Sutherland held that white-collar criminals are relatively immune because of the class bias of the courts and the power of their class to influence the administration of the law. As a result of this class bias, the crimes of the “respectable” upper class generally are handled differently than are the crimes of the lower class. To compensate for this class bias, Sutherland argued that official conviction statistics must be supplemented by evidence of criminal violations from other sources, such as hearings before regulatory commissions, civil suits for damages, administrative hearings, and various other procedures outside criminal court prosecutions (Vold and Bernard, 1986).

Coleman (1994) says that because “Sutherland’s work focused almost exclusively on business crimes and especially violations of federal economic regulations” and because he failed “to devote more attention to violent white-collar crimes,” a debate sprang up about whether white-collar crime is really crime. Since business offenses were handled as civil or administrative matters, Sutherland’s detractors suggested that white-collar criminals were not “real” criminals. Coleman states, “Had the argument focused on flammable clothes that burned helpless children,” Sutherland would have been in a stronger position. Still, Coleman



suggests, Sutherland would have encountered resistance. When he made his address in 1939, crime was seen as something that happened primarily among immigrants and poor people. The idea that business leaders should be considered criminals had an un-American sound to it. Moreover, corporate executives were not likely to support such ideas. Coleman holds, however, that Sutherland won the first round of this debate with the scholars who criticized him.

Tappan was one of the first to criticize Sutherland's position. Tappan, who was trained as both a lawyer and a sociologist, asserted that crime, if legally defined, was an appropriate topic of study for sociologists (1947). Accordingly, he felt that actions that were not against the law were not crimes and that persons who had not been convicted of criminal charges were not criminals. Sutherland, following Sellin (1951), held that Tappan's criterion of legal conviction was too far removed from the offense, which may go undetected, unprosecuted, and/or unconvicted (Green 1990). Burgess (1950) agreed with Tappan that Sutherland erred in failing to distinguish between civil and criminal law. However, Vold and Bernard (1986) suggest that a scientifically adequate theory of crime must explain all behaviors that have the same essential characteristics, whether or not the behavior has been defined as a crime by criminal justice agencies. Sutherland initiated his attempt to develop such a theory with his theory of differential association, which he felt could explain both lower-class and white-collar crime.

Biderman and Reiss (1980) considered white-collar crime to consist of "violations of law, to which penalties are attached, and that involve the use of a violator's position of significant power, influence, or trust in the legitimate economic or political institutional order for the purpose of illegal gain, or to commit an illegal act for personal or organizational gain." Coleman (1994) defines white-collar crime as a "violation of the law committed by a person or group of persons in the course of their otherwise respected and legitimate occupational or financial activity." Green (1990) points out that the terms "respectable" and "significant power or influence" employed in these definitions do not represent an improvement of Sutherland's definition because they are relative terms. Because of such problems, Green seems to suggest that pinpointing the white-collar criminal

(person) is considerably more problematic than delineating white-collar crime. He agrees with Sutherland that white-collar crime is inexorably limited to occupational opportunity.

In their analysis of corporate violations, Clinard and Yeager (1980) say that while corporate crime is white-collar crime, occupational crime is a different type of white-collar crime. Building on earlier work by Clinard and Quinney (1973), they suggest that "occupational crime is committed largely by individuals or by small groups of individuals in connection with their occupations." They include under this type "businessmen, politicians, labor union leaders, lawyers, doctors, pharmacists, and employees who embezzle money from their employees or steal merchandise and tools. Occupational crimes encompass income tax evasion; manipulation in the sale of used cars and other products; fraudulent repairs of automobiles, television sets, and appliances; embezzlement; check-kiting; and violations in the sale of securities" (1980, p. 18).

Clinard and Yeager, in agreement with Sutherland's opinion of what constitutes law violation, say that "a corporate crime is any act committed by corporations that is punished by the state, regardless of whether it is punished under administrative, civil, or criminal law" (1980, p. 16). They also state that Sutherland conducted the first research in this area and that his book *White Collar Crime* (1949) should have had the title of their book: *Corporate Crime*.

Green has extended the work of Clinard and Quinney and Clinard and Yeager and that of others in *Occupational Crime* (1990). He claims that corporate crime almost always occurs within the course of one's occupation, and thus the term "occupational crime" encompasses corporate offenses. To Green, occupational crime refers to any act punishable by law that is committed through an opportunity created in the course of an occupation that is legal. Green goes on to say that the criterion of a legal occupation is necessary, since otherwise the term could include all crimes. A legal occupation, he indicates, is one that does not in itself violate any laws. Thus, the term would exclude persons with occupations that are illegal to begin with, such as bank robbers and professional con men. He lists four types of occupational crime: (1) crimes for the benefit of an employing

organization, (2) crimes by officials through their state-based authority, (3) crimes by professionals in their capacity as professionals, and (4) crimes by individuals as individuals.

Coleman (1994) has noted both the tendency to redefine the concept of white-collar crime to include any nonviolent crime based on concealment or guile (Webster 1980) and the tendency to impose new terms such as "occupational crime," "corporate crime" (discussed above), "elite deviance," "corporate deviance," and "organizational crime" (not to be confused with "organized crime"). Coleman states that "the whole point behind most criminologists' concern with white collar crime is to give the same kind of attention to the crimes of the powerful and privileged that is given to common offenders." He states that the term "white-collar crime" best serves this purpose or goal and is too useful a conceptual tool to be thrown out: "Because it clearly identifies a specific problem of great concern to people around the world, 'white collar crime' has become one of the most popular phrases ever to come out of sociological research" (Coleman 1994, p. 5).

Coleman (1994) clearly acknowledges that some of the criticisms of Sutherland's original definition are valid; for example, (1) responsibility for some white-collar offenses is attributable to groups, and (2) many white-collar offenses are committed by persons from the middle levels of the status hierarchy. Coleman (1994) states that in a major respect, however, Sutherland's views seem relevant. One of the central issues in early debates about the definition of white-collar crime was whether the term should include violations of civil as well as criminal law. As the study of white-collar crime has developed over the last fifty years, many criminologists have sided with Sutherland's view that it should include both civil and criminal violations (Wheeler 1976; Schragger and Short 1978; Braithewaite 1979; Clinard and Yeager 1980; Hagar and Parker 1985).

Following the lines of other students, Coleman states that a typology of white-collar crimes is needed (1994, p. 10). He suggests that while a humanistic perspective might focus on the consequences of white-collar crimes for victims (property losses versus physical injury), a more useful typology might center on differences between offenders. Thus, he suggests a dichotomy between occupational crimes (consisting of offenses by in-

dividuals, whether employee or employer) and organizational crimes, which include both corporate and government crimes.

Poveda (1994, p. 70) asserts with some justification that the design of a typology depends on the theoretical biases of the researcher. He says that scholars need to be aware of the diversity of white-collar crime in considering explanations. Poveda suggests that there are two traditions, both of which can be traced to the initial Sutherland-Tappan debate. One school, the Sutherland school, focuses on the offender as the defining characteristic of white-collar crime. The other school emphasizes the offense as the central criterion of white-collar crime (Poveda 1994, p. 39). Both traditions are alive and well.

#### **TRENDS IN RESEARCH: CORPORATE, OCCUPATIONAL, AND ORGANIZATIONAL CASES**

Punch, Reiman, Clinard and Yeager, Green, and Coleman are unanimous that interest in white-collar crime emerged or gathered speed after the Watergate scandal in the 1970s. It does seem that while there was early excitement over Sutherland's initial 1939 address, interest quickly waned. In the late 1940s, interest in social order seemed to predominate. Clinard and Yeager (1980) reviewed the events of the 1960s and 1970s, describing the occasional corporate conspiracies and environmental abuses that came to the public's attention. They noted how the short and suspended sentences given to Watergate offenders contrasted "sharply with the 10-, 20-, 50-, and even 150-year sentences given to burglars and robbers." Reiman, beginning with the first edition *The Rich Get Richer and the Poor Get Prison* (1979), has been a consistent critic of the different ways in which justice has been meted out to the poor and the well to do.

Clinard and Yeager (1980), in what Punch (1996) calls the piece of research that is closest to Sutherland's legacy, conducted the first large-scale comprehensive investigation of the law violations of major firms since Sutherland's pioneering work. Sutherland (1949) conducted his study on seventy of the two-hundred largest U.S. nonfinancial corporations. The study of Clinard and Yeager (1980) involved a systematic analysis of federal administrative, civil, and criminal actions initiated or com-

pleted by twenty-five federal agencies against the 477 largest publicly owned manufacturing (Fortune 500) corporations and 105 of the largest wholesale, retail, and service corporations in the United States in the period 1975–1976. Thus, they had a total sample of 582 corporations. Clinard and Yeager found the following:

*A total of 1,553 federal cases were begun against all 582 corporations during 1975 and 1976 or an average of 2.7 federal cases of violation each. Of the 582 corporations, 350 (60.1 percent) had at least one federal action brought against them, and for those firms that had at least one actions brought against them, the average was 4.4 cases.* (1980, p. 113)

Six main types of violations were reported by Clinard and Yeager (1980): administrative, environmental, financial, labor, manufacturing, and unfair trade practices. Their evidence shows that “often decisions on malpractice were taken at the highest corporate levels, that records were destroyed or ingeniously doctored by executives and their accountants and that the outcomes of these decisions can have serious economic, financial, political, personal, and even physical consequences” (Punch 1996, p. 52).

Punch (1996) focused on selected cases, mainly from other countries, that bear on his hypothesis that business is criminogenic, i.e., it justifies or tolerates illegal behavior. This parallels Clarke’s (1990) contention that crime and misconduct are endemic to business and that the key to understanding them lies in recognizing the structure that the business environment gives to misconduct both in terms of opportunities and in terms of how misconduct is managed. Punch states that his underlying purpose is to use his cases to “unmask the underlying logic of business and the submerged social world of the manager” (1991, p. 213). Punch states, however, that he is not asserting that most companies are guilty of criminal behavior. Indeed, he says that “there are companies which explicitly set out to conform to the law; they maintain a record of no transgressions” (p. 214). In short, Punch focuses on companies that engage in corporate deviance.

If after reviewing the works of Clinard and Yeager (1980) and Sutherland (1949), a student of crime still is inclined to conclude that the term “white-collar crime” refers to harmless or small

mistakes in judgment rather than to harmful, selfish, profit-oriented motives followed by an intentional cover-up of brutal consequences, Punch sets the record straight. One of the most chilling of his cases is the Three Mile Island nuclear accident in Pennsylvania in March 1979. A series of incidents involved (1) a leaking seal, (2) leading to feedwater pumps that failed to come into operation and remove heat from the core (3) because of blocked pipes stemming from (4) valves having been left in the wrong position after routine maintenance some days earlier and (5) that could not be seen because the control switch indicator was obscured. Through further accidents, mounting tension, and delays, danger increased and a complete meltdown loomed. Then, “almost by coincidence, and perhaps because of a new shift supervisor checking the PORV [pilot-operated relief valve], the stuck relief valve was discovered. More as an act of desperation than understanding. . . the valve was shut” (p. 126). Punch then quotes Perrow (1984, p. 29):

*It was fortunate that it occurred when it did; incredible damage had been done with substantial parts of the core melting, but had it remained open for another thirty minutes or so, and HPI [high-pressure injector] remained throttled back, there would have been a complete meltdown, with the fissioning material threatening to breach containment.*

What Punch (1996) describes in his recounting of this case and others includes the following:

1. Earlier warnings by an efficient inspector about the inherent dangers in nuclear plant procedures had been brushed aside; the inspector then had been subjected to strong informal control in an attempt to deflect his message.
2. Profitability (based on the productivity of privately run industry) was the top priority, and this is a problem in most corporations.
3. Management responds to crises such as the one described (by Perrow, Punch, and others) by keeping things running.
4. Regulatory agencies tend to compromise rather than enforce the rules governing safety; indeed, they appear to “have a dual function both to regulate and promote an industry” (p. 255).

Among his conclusions, Punch writes:

*When cleared of radioactive debris, Unit Two will remain in quarantine for thirty years until it is dismantled as planned along with Unit One, which will then have reached the end of its working life. Until 2020, then, Unit Two at Three Mile Island in the middle of the Susquehanna River will remain a silent sentinel to a disturbing case of incompetence, dishonesty, complacency, and cover-up.* (pp. 134–135).

This conclusion is disturbingly similar to that made earlier by Clinard and Yeager (1980) as they noted the evasion of responsibility by many of the corporate managers in their study and the unethical climates that disregarded the public's welfare (p. 299). Punch (1996) found a similar criminal climate in cases that involved Thalidomide and its effects on pregnancy, originating in Germany; the Guinness affair in England involving illegal financial dealings; the Italian affair involving business, politics, organized crime, and the Vatican Bank; a case in the Netherlands involving deviance in a shipbuilding conglomerate; and the savings and loan scandal in the United States.

The savings and loan scandal, covering the period 1987–1992, represents one of the greatest fraud cases in the twentieth century (Mayer 1990; Pizzo et al. 1989; Thio 1998). Thio stated that this fraud cost taxpayers \$1.4 trillion. Coleman (1994) explained that the rise in interest rates in the 1980s created serious economic problems for many savings and loan associations (thrifts) as their inventories of low-interest fixed-rate mortgage loans became increasingly unprofitable. The thrifts had been restricted until then by government regulations and were losing money. Punch (1996) suggests that the election of Ronald Reagan to the presidency accelerated an ideology of deregulation that held that business should be freed from undue rules and restrictions and that market forces should be given free rein to enhance competition. Deregulation of the industry did take place, and this meant that interest rates were not rigid, financing could be offered with no down payment, and loans could be given for consumer and commercial purposes. Those developments opened up opportunities for unscrupulous businesspersons who moved in and exploited the thrifts for devious purposes. Deregulation not only loosened con-

trols but “raised the limit of federal protection on deposits to \$100,000; brokers could make commissions on them, investors received higher interest rates and the S&L attracted new funds while enjoying a federal safety net” (Punch 1996, p. 16). As quoted by Punch (1996), the Federal Home Loan Bank-Board reported the following to Congress in 1988:

*Individuals in a position of trust in the institution or closely affiliated with it have, in general terms, breached their fiduciary duties; traded on inside information; usurped opportunities for profits; engaged in self-dealing; or otherwise used the institution for personal advantage. Specific examples of insider abuse include loans to insiders in excess of that allowed by regulation; high-risk speculative ventures; payment of exorbitant dividends at times when the institution is at or near insolvency; payment from institutions' funds for personal vacations, automobiles, clothing, and art; payment of unwarranted commissions and fees to companies owned by a shareholder; payment of “consulting fees” to insiders or their companies; use of insiders' companies for association business; and putting friends and relatives on the payroll of the institutions.* (U.S. General Accounting Office 1989, p. 22).

Calavita and Pontell (1990) categorized these fraudulent practices as “unlawful risk-taking,” “looting,” and “covering-up.” Their work indicates that this type of corporate crime was unlike that reported by writers such as Clinard and Yeager (1980), which was designed to enhance corporate profits; rather, the savings and loan affair revealed premeditated looting for personal gain. They called this kind of crime “a hybrid of organizational and occupational crime”; this was crime by the corporation against the corporation, encouraged by the state, and with the taxpayer as the ultimate victim (Calavita and Pontell 1991).

Those studies of corporate, organizational, and occupational crimes appear to follow Sutherland's initial focus and concerns and to establish what might be called a Sutherland school of white-collar crime. Other researchers besides those mentioned earlier have made valuable contributions to this growing field, including Geis (1968), Braithewaite (1984), and Vaughn (1983).

### OFFENDERS AND OFFENSES

The focus on offenses has been referred to by Poveda (1994) as a second school of thought that follows Tappan's tradition and is distinct from Sutherland's tradition with its emphasis on the offender. Social class, of course, has always been at the center of this debate in terms of its relationship to white-collar crime. While Sutherland emphasized that many violations were committed by respectable persons in the course of their occupations, Tappan argued for the recognition of the law in defining crime. Poveda states that

*the law of course specifies which acts are to be criminalized regardless of who commits them. In this view the defining characteristic of white-collar crime is the offense rather than the offender. The problem of defining white-collar crime from this perspective becomes one of deciding which subset of crimes is "white collar." By separating white-collar crime from the characteristics of the offender, white-collar crime in the legal tradition ceases to be linked to any particular social class. (1994, p. 40)*

Edelhertz (1970) objected to Sutherland's assertion that white-collar crimes must occur in the course of the offender's occupation. He argued that that definition excludes offenses such as income tax evasion, receiving illegal Social Security payments, and other similar offenses, that he considered white-collar crime. Edelhertz's work suggests that the class of the offender need not be central to the concept of white-collar crime but that the offense should be the central consideration. In this respect, he reflected Tappan's stance. According to Poveda (1994), criminologists who focus on the offense have come to dominate views among workers in the justice system and have become more numerous among criminologists since Edelhertz modified Sutherland's definition in the 1970s.

Shapiro (1990) also suggests that the analysis of white-collar crime must shift from a focus on the offender to focus on the offense, particularly when violations of trust situations and norms are involved. She argues that the leniency shown to white-collar criminals accused of securities fraud has resulted from the social organization of their offenses and the problems of social control they posed rather than from class biases involving higher status.

It is relevant, then, to ask whether the laws on white-collar crime are applicable to corporations as well as to individuals. After reviewing the historical development of laws creating various white-collar crimes (embezzlement and theft, unfair competition, bribery and corruption, endangerment of consumers and workers, and environmental degradation), Coleman (1994) explained that the laws involving criminal intent have been extended to include corporations, stating that "it is now common for corporations to be charged with criminal violations of the regulatory statutes as well as more serious offenses" (p 125). As will be seen below, conviction and punishment are different matters.

### CONVICTIONS AND SENTENCING

In the 1980s, Wheeler directed a series of studies at Yale University that focused on both offenders and offenses as well as on a comparison of white-collar and conventional crimes. Wheeler et al. (1982, 1988) focused on eight white-collar offenses in the federal system, which they clustered into three types organized by complexity: (1) the most organized: antitrust and securities fraud, (2) intermediate: mail fraud, false claims, credit fraud, and bribery, (3) the least organized: tax fraud and bank embezzlement. Their studies included only convicted offenses, not violations of civil or administrative regulations. They found that white-collar criminals are better educated, older, and more likely to be white and well off financially compared with conventional criminals. They also found that female white-collar offenders are more similar to conventional criminals than to their male counterparts. In an analysis of the Yale data, Daly (1989) found substantial differences between male and female white-collar offenders. Daly raised questions about whether the term "white collar" should be applied to women since they seldom committed offenses as part of a group, made less money from their crimes, and were less educated. Box (1983) suggested that female workers had fewer criminal opportunities than men. In another important tangent to their findings, Weisbund et al. (1991) argued that much white-collar crime is engaged in by middle-class individuals, revealing an unexpected source of inequality in the Justice Department.

In the area of sentencing, Wheeler and colleagues (1982) found that higher-status defendants

charged with white-collar crimes were more likely to receive jail sentences than were lower-status defendants. In explanation, they suggested that because most cases against high-status defendants were never prosecuted, the few cases that were prosecuted were compelling. They also noted that the research was conducted shortly after the Watergate scandals, when judges were more attentive to misdeeds attributed to greed. In another study of sentencing and status, Hagar and Parker (1985), in an analysis of data on persons charged with criminal and noncriminal acts, observed differential sentencing of employers, managers, and workers. They found that compared with workers, managers were more likely and employers were less likely to be charged under the criminal code. Employers were instead more likely to be charged with Securities Act violations that carried less stigma and a shorter sentence. They noted that employers are in positions of power that allow them to be distanced from criminal events and obscure their involvement. Finally, they stated that Sutherland noted an "obfuscation as to responsibility" that accompanies corporate positions of power. Thus, employers face securities charges rather than criminal charges that require a demonstration of malice.

The Clinard and Yeager study (1980) did not include street criminals. The authors note that there is more leniency for corporate than for other white-collar offenders. In their study, it was found that only 4 percent of sanctions handed out for corporate violations involved criminal cases against individual executives. Of the fifty-six convicted executives in large corporations, 62.5 percent received probation, 21.4 percent had their sentences suspended, and 28.6 percent were incarcerated (p. 287). Among the latter defendants, two received six-month sentences and the remaining fourteen received sentences averaging only nine days (1980).

Relevant here are criticisms of legislation designed to control crime. Geis (1996) points out that legislation dealing with "three strikes" has given a "base on balls" to white-collar offenders and indicates an underlying class, ethnic, and racial bias that seeks to define criminals as "others" rather than confronting the more costly crimes of community leaders and the corporate world.

After reviewing a number of studies of sentencing, including studies by Shapiro, Wheeler and colleagues, Clinard and Yeager, Hagar and

Parker, and Mann, Coleman concludes that "the evidence leaves little doubt that white collar offenders get off much easier than street offenders who commit crimes of similar severity" (1994, p. 157). He notes, however, that "there is some evidence that the punishment for white-collar offenses has been slowly increasing in recent years" (p. 157). Coleman still maintains that while there is increasing severity of punishment, such prosecutors remain rare. He says that "Leo Barrile [1991] concluded that only 16 cases of corporate homicide have [ever] been charged, only 9 of those made it to trial, and in only three cases were corporate agents sentenced to prison" (p. 157).

### IMPLICATIONS FOR RESEARCH AND THEORY

Poveda (1994) notes that "in spite of the recent work on white collar offenses, our knowledge of white-collar crime is much more circumscribed than that of conventional crime and draws more heavily upon official crime statistics" (p. 79). He says that Wheeler and associates concluded that the vast majority of white-collar offenders are nonelite offenders who look like average Americans. Wheeler and colleagues argued that Sutherland's definition was narrowly focused on the upper class and ignored the middle group of offenders. Poveda suggests that "it is time to consider alternate approaches to gaining knowledge about white-collar crime, approaches that circumvent the official statistics." (1994, p. 79) because of the need to gather more information about crimes in large organizations. Poveda states that "large organizations have the power and resources to control public information about themselves to a much greater extent than other kinds of offenders" (1994, p. 79). He asserts that researchers will have to penetrate the curtain of secrecy that may enclose illegal behavior.

Poveda proposes that there is a need to study accidents and scandals. He cites the suggestions of Molotch and Lester (1974), who showed how routine news events are managed by political actors in society: corporations, labor unions, the president, members of Congress, and so on. These actors define issues for the public construct the news. Only accidents and scandals "penetrate this constructed reality of the news by catching these

major actors off guard. While accidents are unplanned, scandals involve planned events but they must typically be disclosed by an inside informer to an organization because they involve sensitive information (Poveda 1994, p. 80). Poveda suggests that these events often reveal the incidence of white-collar crime. He says that the *Challenger* disaster was an “‘accident’ that led to disclosures of questionable judgment by the National Aeronautics and Space Administration and the Morton Thiokol Corporation” (p. 81). Punch (1996) has undertaken research that focused on accidents and scandals. This research involved case studies of the savings and loan scandal, the Three-Mile Island nuclear accident, and many others. Braithwaite has written about the drug companies and their record of fraud in testing, price-fixing, and the provision of perks for medical practitioners (Braithwaite 1984).

It is important to note that research in Britain by Clarke and in Britain, Holland, and the United States by Punch and others suggests that corporate, organizational, and occupational crime in the industrial world have more common elements than differences. Moreover, cybercrime, relying heavily on the Internet, has increased greatly according to a report by the British Broadcasting Corporation. Fraud employing stolen credit cards and stolen identities would appear to have worldwide similarities as a result of the growing use of the World Wide Web.

Punch has noted that researchers are increasingly targeting accidents, disasters, and scandals in the business world. He has not noticed any slackening of the calculative nature of business, as evidenced in the transfer of technology and manufacturing to developing nations. Noting the existence of, if not an increase in, the number of shrewd players on the world scene and the growth of unregulated markets, Punch expects fresh scandals not only in the United States but in eastern Europe and the Far East (1996, pp. 268, 269). There is, then, growing agreement that a focus on accidents, disasters, and scandals will provide a growing database on white-collar crime and its various types and that such a database will lead to the development of a more adequate theory of white-collar crime. This would seem to be a prerequisite for the development of an adequate system of deterrence of white-collar crime.

There seems to be agreement that criminologists are some distance away from developing an adequate theory of white-collar crime. Wheeler and associates, Yeager and associates, Poveda, Braithwaite, Coleman, and others have contributed to concept development and theory building in this field. Thus, Reed and Yeager (1996) have emphasized the need to assess how notions of self-interest become merged with corporate interests and the conditions under which these socially constructed interests lead to socially harmful outcomes. Wheeler (1992) has addressed a similar concern with motivational and situational processes that drive individuals to risk involvement in white-collar crime. Coleman also asserts that the related theoretical problems of motivation and opportunity must be understood. He considers the neutralization of ethical standards by which white-collar criminals justify their pursuit of success, the secrecy that shields corporate actions, and the opportunities provided by the legal and judicial systems essential links in the development of an understanding of this type of crime.

Braithwaite (1984) also draws on the structure of opportunity in attempting to understand organizational crime but is perhaps best known for his concept of reintegrative shaming. In his approach, he utilizes control theory, specifying the processes by which corporate offenders are encouraged to strengthen their stake in conformity. He asserts that the other kind of shaming—stigmatization—has the effect of reinforcing offenders in their criminality.

There is an apparent need for a continued focus on occupational and corporate deviance and on individual and organizational offenders in the field of white-collar crime. Interest has been growing in the field, and there is reason to believe that academic researchers, government agencies, and legislatures must communicate with one another more if progress is to be made in the important matter of constructing better deterrents to white-collar crime. However, as suggested by Punch and others, investigators must search more closely for the dark, irrational side of organizations—incompetence, neglect, ambition, greed, power—as well as for motives and structures that allow managers to practice deviance against the organization. It is clear that while many researchers argue for a focus on the organization and others emphasize the need to study individual offenders, there is a grow-

ing acceptance of the need to explore all avenues that lead to white-collar crime. Indeed, there seems to be a clamor among researchers that not only must motivation and opportunity be studied but that the subcultures that facilitate immoral behavioral structures should be analyzed to understand white-collar or any other type of crime.

## REFERENCES

- Barrile, Leo 1991 "Determining Criminal Responsibility of Corporations." Paper presented at the American Society of Criminology, San Francisco, November.
- Biderman, Albert, and Albert J. Reiss 1980 *Data Sources on White Collar Lawbreaking*. Washington D.C.: National Institute of Justice.
- Box, Steven 1983 *Power, Crime, and Mystification*. New York: Tavistock.
- Braithwaite, John 1979 *Inequality Crime and Public Policy*. London: Routledge and Kegan Paul.
- 1984 *Corporate Crime in the Pharmaceutical Industry*. London: Routledge & Kegan Paul.
- Burgess, Ernest W. 1950 "Comment to Harting." *American Journal of Sociology* 56:25–34.
- Calavita, Kitty and Henry Pontell 1991 "Other People's Money Revisited: Collective Embezzlement in the Savings and Loan and Insurance Industries." *Social Problems* 38: 94–112.
- 1990 "Heads I Win, Tails You Lose: Deregulation, Crime, and Crisis in the Savings and Loan Industry." *Crime and Delinquency* 36:309–341.
- Chamber of Commerce of the United States 1974 *A Handbook on White Collar Crime*. Washington, D.C.: Chamber of Commerce of the United States.
- Clarke, M. 1990 *Business Crime* Cambridge, U.K.: Polity Press.
- Clinard, Marshall B., and Richard Quinney 1973 *Criminal Behavior Systems*, 2nd ed. New York: Holt, Rinehart & Winston.
- , and Peter C. Yeager. 1980 *Corporate Crime*. New York: Free Press.
- Coleman, James 1994 *The Criminal Elite: The Sociology of White Collar Crime*, 3rd ed. New York: St. Martin's.
- Daly, Kathleen 1989. "Gender and Varieties of White Collar-Crime." *Criminology* 27: 269–294.
- Edelhertz, Herbert 1970. *The Nature, Impact and Prosecution of White Collar Crime*. Washington D.C.: U.S. Government Printing Office.
- 1996 "A Base on Balls for White Collar Criminals." In David Shicor and Dale Seckrest, eds., *Three Strikes and You're Out: Vengeance as Public Policy*. Thousand Oaks, Calif.: Sage.
- Geis, Gilbert, ed. 1968 *White Collar Crime*. New York: Atherton Press.
- Green, Gary S. 1990 *Occupational Crime*. Chicago: Nelson Hall.
- Hagar, John, and Patricia Parker 1985 "White Collar Crime and Punishment: The Class Structure and Legal Sanctioning of Securities Violations." *American Sociological Review* 50:302–316.
- Mayer, Martin 1990 *The Greatest Ever Bank Robbery: The Collapse of the Savings and Loan Industry*. New York: Scribner.
- Molotch, Harvey, and Marilyn Lester 1974 "News as Purposive Behavior: On the Strategic Use of Routine Events, Accidents, and Scandals." *American Sociological Review* 39:101–112.
- Pizzo, S., M. Fricker, and P. Muolo 1989 *Inside Job: The Looting of America's Savings and Loans*. New York: McGraw-Hill.
- Poveda, Tony 1994 *Rethinking White-Collar Crime*. Westport, Conn.: Praeger.
- Perrow, Charles 1984 *Normal Accidents*. New York: Basic Books.
- Punch, Maurice 1996 *Dirty Business: Exploring Corporate Misconduct*. London: Sage.
- Reed, Gary E., and Peter Cleary Yeager 1996 "Organizational Offending and Neoclassical Criminology: Challenging the Reach of a General Theory of Crime." *Criminology* 34:357–382.
- Reiman, Jeffrey 1995 *The Rich Get Rich and the Poor Get Prison*, 4th ed. Boston: Allyn & Bacon.
- Sellin, Thorsten 1951 "The Significance of Records of Crime." *Law Quarterly Review* 67:489–504.
- Schrager, Laura S., and James P. Short 1978 "Toward a Sociology of Organizational Crime." *Social Problems* 25:407–419.
- Shapiro, Susan 1990 "Collaring the Crime, Not the Criminal: Reconsidering the Concept of White Collar Crime." *American Sociological Review* 55: 346–365.
- Sutherland, Edwin H. 1949 *White Collar Crime*. New York: Dryden Press.
- Tappan, Paul 1947 "Who Is the Criminal." *American Sociological Review* 12:96–102.
- Thio, Alex 1998 *Deviant Behavior*, 5th ed. New York: Addison-Wesley. *US News & World Report* 1985 May 20.
- Vaughn, Diane 1983 *Controlling Unlawful Organizational Behavior: Social Structure and Corporate Misconduct*. Chicago: University of Chicago Press.



- Vold, George, and Thomas J. Bernard 1986 *Theoretical Criminology*, 3rd ed. New York: Oxford University Press.
- Webster, Wilham 1980. "An Examination of FBI Theory and Methodology Regarding White Collar Crime Investigation and Prevention." *American Criminal Law Review* 17:275-286.
- Weisbund, D., S. Wheeler, E. Waring, and N. Bode 1991 *Crimes of the Middle Classes: White Collar Offenders in the Federal Courts*. New Haven, Conn.: Yale University Press.
- Wheeler, Stanton 1992 "The Problem of White Collar Crime Motivation." In Kip Schlegel and David Weisbund, eds., *White Collar Crime Reconsidered*. Boston: Northeastern University Press.
- , and E. Waring. 1988 "White Collar Crimes and Criminals." *American Criminal Law Review* 25:331-357.
- , D. Weisbund, and N. Bode 1982 "Sentencing the White-Collar Offenses." *American Sociological Review* 47:641-659.
- 1976 "Trends and Problems in the Sociological Study of Crime." *Social Problems*. 23:523-534

JAMES E. TEELE

## WIDOWHOOD

Marriages that do not end in divorce eventually dissolve through the death of a spouse. The stress of bereavement derives largely from the disorganization caused by the loss of the deceased from the social support system of the survivor. The death of a marital partner requires the development of alternative patterns of behavior so that the survivor can maintain satisfactory relations with the family, the kin group, and the community and sustain his or her personal equilibrium. Families exhibit considerable diversity in their attempts to accomplish these transitions. The difficult and sometimes devastating transition to widowhood or widowerhood necessitates a reintegration of roles suitable to a new status. If children are present, parental death precipitates a reorganization of the family as a social system. Roles and status positions must be shifted, values reoriented, and personal and family time restructured. The potential for role strains and interpersonal conflicts becomes evident as relationships are lost, added, or redefined (Pitcher and Larson 1989). Loneliness becomes a major problem. In many modern societies, this adaptive process proceeds

with few or no guidelines because the widowed person tends to be "roleless," lacking clear norms or prescriptions for behavior (Hiltz 1979).

## WIDOWHOOD ACROSS CULTURES

Human behaviors generally are guided by the dominant prescriptions and proscriptions embedded in particular societies, and this is reflected in wide cross-cultural variations among those who have lost a spouse through death (Lopata 1996). For example, the situation of Hindu widows in India has undergone numerous changes, ranging from extremely harsh treatment in the past to slow but steady improvement in the modern era. The custom of *suttee*—the wife's self-immolation on her husband's funeral pyre—has long been outlawed but periodically reappears, especially in rural areas. Even today, widows in that highly patriarchal, patrilineal, and patrilocal society experience isolation and a loss of status. Their remarriage rate is very low. Widows often face a difficult life that is influenced by vestiges of patriarchal and religious dogma and exacerbated by economic problems that force them to become dependent on sons, in-laws, and others. Widowers, by contrast, are encouraged to remarry soon and add progeny to the patriarchal line. Israel is another place where the society and religion are strongly patriarchal and women lose status in widowhood. Jewish mourning rituals "tend to isolate the widow and tie her to the past rather than providing means of creating a new life" (Lopata 1996). Moreover, women who lose husbands through civilian causes of death encounter greater difficulties than do those whose husbands are killed in the military. War widows and their families receive preferential treatment through government policies that give them special recognition, numerous benefits, and many more alternatives for improving their status and prestige than is possible in more traditional societies. Remarriage, for example, is a much more acceptable alternative for women in Israel than it is in India.

All societies are undergoing various degrees of transition. Korea is a society whose transitional problems are dramatically reflected in the situation of widows. Earlier in Korean history, widowhood resulted in a loss of status and remarriage generally was prohibited. Husbands tended to be much older than their wives and to have a higher

mortality rate, and a large number were killed in wars. Moreover, widowers remarried, whereas most widows remained single. All these factors contributed to a widening ratio of widows to widowers over the years. Under the impact of modernization, including increased urbanization and industrialization, Korean society is being transformed, and with it the conditions surrounding the status of widowhood. This transformation includes a shift from authoritarian societal and familial systems in a primarily rural environment toward systems based on more equalitarian norms. Widows began to move to the cities, and this had advantages and disadvantages. On the one hand, they could accompany their sons and take advantage of urban services and the possibility of new friendships. On the other hand, the move removed them from their extended families and neighborhood friends and the communal supports in their rural villages. Living with a son in the city often strained the daughter-in-law relationship. In addition, being distanced from the relatively stable and integrated life of their villages and lacking friendship networks in their new environment often left them vulnerable to loneliness, especially in the case of the elderly widowed. Presumably, succeeding generations with greater personal resources will encounter fewer adaptational requirements.

While survivors face certain common problems and role strains both within and outside the immediate family, it is difficult to specify a normative course of adjustment. This is the case because the widowed are a heterogeneous group characterized by wide differences in social and psychological characteristics. It also is due to the fact that spousal loss evokes a panorama of emotional and behavioral responses from the survivors, depending on factors such as the timing and circumstances of the spouse's death. For example, a wife whose husband was killed in a military battle will respond differently than she would if he had committed suicide or suffered a long terminal illness. Many other antecedent conditions, such as the quality of the marital relationship, affect the bereavement reactions and coping strategies of survivors.

### THE DEMOGRAPHICS OF WIDOWHOOD

Census data for the United States show that at the end of the 1990s there were more than 13.5 mil-

lion widowed persons, 85 percent of whom were women. However, people in the widowed category may leave it through remarriage. Hence, the number of people who have ever experienced spousal loss is much greater than is indicated by these data.

For some decades, the widowed female has outnumbered her male counterpart by an ever widening margin. Three factors account for this: (1) Mortality among females is lower than it is among males, and therefore, greater numbers of women survive to advanced years, (2) wives are typically younger than their husbands and consequently have a greater probability of outliving them, and (3) among the widowed, remarriage rates are significantly lower for women than for men. Other factors that contribute to the preponderance of widows include war, depressions, and disease pandemics.

For several reasons, widowhood has become largely a problem of aged women. Each year in the United States, deaths of spouses create nearly a million new widows and widowers. Among people 65 years of age or over, roughly half the women compared with about 14 percent of the men are widowed. (U.S. Bureau of the Census 1997). Advances in medical technology and the pervasiveness health programs have extended life expectancy. The probability of mortality before middle age has decreased, and for the most part widowhood has been postponed to the later stages of the life cycle. Gains in longevity have been more rapid for women than for men. Thus, the growing proportion of elderly females accents their higher rates of widowhood. About one-fourth of all married women will become widows by age 65, and one-half of the remaining women will be widowed by age 75. During that age span, only one-fifth of men will lose their wives. It is projected that the ratio of widows to widowers will increase dramatically from nearly six to one currently to ten to one over the next quarter century.

Because the large majority of the widowed are women, most studies have concentrated on them, while the social consequences for men who lose their spouses has remained a comparatively unexplored area since Berardo (1970) called attention to this gap three decades ago. Widowers, although fewer in number, face many of the same adjustments that confront their female counterparts. At the same time, there is ambiguous evidence that

suggests that widowers have greater vulnerability compared to their female counterparts, while other studies present the situation of widows as more problematic. This disagreement in findings results in part from the failure of many studies to control for the confounding influences of factors such as age, social class, income, health, and retirement.

### RESEARCH FINDINGS ON WIDOWHOOD

In making the transition from marriage to widowedhood, the bereaved often are confronted with a variety of personal and familial problems and are not always successful in adapting to those circumstances. This is reflected in the findings that compared with married persons, the widowed consistently have higher rates of mortality, mental disorders, and suicide (Balkwell 1981; Smith et al. 1988). While there is a consensus that bereavement is stressful, research on its effects on physical health has yielded inconsistent results. The evidence shows that the widowed experience poorer health than do the married, but the reasons for this difference are unclear.

Because widowhood is most likely to occur in the elderly, research has focused on that population. However, there is some evidence that the transition to widowhood varies by developmental stage. Older widows adapt more readily because losing a spouse at an advanced age is more the norm, making acceptance of the loss easier than it is for those who are young when widowed. Grieving over the death of a husband or wife at older ages can be exacerbated if additional significant others also die, requiring multiple grieving. This can cause *bereavement overload*, which makes it difficult for the survivor to complete the grief work and bring closure to the bereavement process (Berardo 1988). There is a consensus that the distress associated with conjugal bereavement diminishes over time. Grief becomes less intense as the years pass, but this is not a simple, linear process. The emotional and psychological traumas of grief and mourning may recur sporadically long after a spouse has died.

Gender differences in adaptation to widowhood have been widely debated. The evidence suggests a somewhat greater vulnerability for widowers (Stroebe and Stroebe 1983). Men are less likely to have same-sex widowed friends, are more likely to be older and less healthy, have fewer

family and social ties, and experience greater difficulty in becoming proficient in domestic roles (Berardo 1968, 1970). Higher mortality and suicide rates also suggest greater distress among widowers.

Continuous widowhood has been associated with a loss of income and an increased risk of poverty. Two-fifths of widows fall into poverty at some time during the five years after the death of their husbands. Female survivors, for example, have dramatically higher proportions in poverty than do their divorced counterparts, although both groups experience economic risk resulting from the ending of their marriages that may impede their and their families' adjustment to a new lifestyle (Morgan 1989). There is some evidence that widowers also suffer a decline in economic well-being, although to a lesser degree than do their female counterparts (Zick and Smith 1988). Poor adjustment to widowhood thus may be related to a lack of finances. Elderly individuals often have below-average incomes before the death of a spouse. They may be unwilling or unable to seek employment and are likely to face discrimination in the labor market (Morgan 1989). The younger widowed are more likely to have lost a spouse suddenly and therefore may be unprepared to cope with a lower financial status.

Life insurance has become a principal defense against the insecurity and risk of widowhood in urban industrial society with its nuclear family system. It is a concrete form of security that may help a bereaved family avoid an embarrassing dependence on relatives and the state in the case of an untimely death. However, the amount of insurance obtained is often insufficient to meet the needs of the survivors. Even in instances in which adequate assets have been accumulated, many surviving wives are not prepared to handle the economic responsibilities brought about by a husband's death (Nye and Berardo 1973). Presumably, in the future, a better educated and occupationally experienced population of widows, especially those who were involved in a more equalitarian marital relationship of shared responsibilities, will be better able to cope with their new single status.

Widowhood often leads to changes in living arrangements. Reduced income may force surviving spouses to seek more affordable housing. They also may choose to relocate for other reasons, such

as future financial and health concerns, a desire to divest themselves of possessions, and a desire to be near relatives or friends (Hartwigsen 1987). Most often, survivors living alone are women, usually elderly widows. Isolation and lack of social support can lead to deterioration in their physical and mental well-being. Compared with elderly couples, they are much more likely to live in poverty and less likely to receive medical care when it is needed (Kasper 1988).

### WIDOWHOOD AND DIVORCE

**Similarities and Differences.** Early epidemiological analyses suggested that more deleterious effects were associated with separation and divorce than with widowhood. However, later surveys found higher levels of physical and psychological distress among the widowed than among the divorced. (Kitson et al. 1989). These contradictory findings have not been reconciled. However, for many decades, researchers also have perceived a number of similarities in adjustment between the two groups. For both, there are accompanying disruptions in lifestyle related to changes in income, social interactions, definitions of self, lost emotional attachment, and general psychological well-being. For example, similarities in adjustment have been noted with respect to mode of death or cause of divorce, including the amount of prior warning or preparation a person has before either event, the degree of responsibility felt, and the cause of the event. The more unexpected the loss is, the more responsibility one feels for the loss and wonders whether he or she could have prevented it or helped the spouse and the more difficult the adjustment is.

Another similarity is that whether a spouse is lost through divorce or widowhood, the length of time for adjustment shows considerable variability. The degree of emotional attachment affects the degree of anxiety and depression associated with the loss of a partner, and in both cases the attachment declines as time passes. Emotional attachment is a normal outcome of the tendency for people to form strong affectional bonds to significant others and is not pathological. However, the accumulative changes that occur with the loss of a partner make those who are divorced or widowed more vulnerable to psychological and physical illness, suicide, accidents, and death. While

most partners return to their former level of functioning within a couple of years after the loss of a spouse, some never recover and continue to have poor levels of functioning.

There are also specific factors that make adjustment in widowhood or divorce more difficult, including age, gender, race, and socioeconomic status. Adjustment to the loss of a spouse in either case appears to be more difficult for younger women. Some analysts argue that age is a confounding factor because younger women are more studied as divorcees and older women more studied as widows and because divorce is more common among the young and widowhood more common among the old. The latter factor means that one's adjustment is somewhat dependent on those who have gone before and can help socialize a person to the new role. However, more recent research suggests that younger women still face more adjustment problems (Kitson et al. 1989; Gove and Shin 1989). Analyses suggest that the young and the old bereaved differ in both the intensity of grief and patterns of grief reactions, especially with respect to adverse health and psychological outcomes within the first two years after the demise of a husband (Sanders 1988). It appears that younger widows experience a different adjustment than do older widows, in part because they have fewer cohort friends who are also widows.

Younger survivors are developmentally "out of sync" with their cohorts, and this exacerbates their sense of loneliness and need for companionship (Levinson 1997). Their expectations may be different because they have more years ahead and more potentially eligible marital partners in the future than do older widows. Blacks appear to have an easier time adjusting to the loss of a spouse through divorce than do whites, and black females, who may receive more familial support than whites do, appear to adjust more easily than do white females. Finally, income and financial security play a major role in adjustment: Those near poverty have the most difficult time coping with the loss of a spouse. Female survivors have more problems coping with the loss of income than do their male counterparts, often because their incomes are tied to health insurance, retirement, and other benefits that accompanied the husband's occupation. Men have more difficulty than do women handling the household chores

that were often the responsibility of their wives. Future male cohorts may have less difficulty with this because of changes in the socialization of male children and the rising age at first marriage, and the fact that young men have to cope with household responsibilities on their own before marriage.

### **WIDOWHOOD AND REMARRIAGE**

The probability of remarriage is significantly lower for widows than for widowers, especially at the older ages. It appears that while a large majority of older widows remain attracted to and interested in men in terms of companionship, for a variety of reasons only a small minority report a favorable attitude toward remarriage (Talbot 1998). Some may feel they are committing psychological bigamy and therefore reject remarriage as an option (DiGiulio 1989). There is also a tendency to idealize the former partner, a process known as sanctification (Lopata 1979). This makes it difficult for widows to find a new partner who can compare favorably with the idealized image of the deceased (Berardo 1982). Widows also remarry less frequently than do widowers because of the lack of eligible men and the existence of cultural norms that degrade the sexuality of older women and discourage them from selecting younger mates. Many women manage to develop and value a new and independent identity after being widowed, leading them to be less interested in reentering the marriage market.

There are other barriers to remarriage for the widowed. Dependent children limit the opportunities of their widowed parents to meet potential mates or develop relationships with them. Older children may oppose remarriage out of concern for their inheritance. Widowed persons who cared for a dependent spouse through a lengthy terminal illness may be unwilling to risk bearing that burden again.

### **WIDOWHOOD AND MORTALITY**

The increased risk of mortality for widowed persons has been widely reported. Men are at a greater risk than women after bereavement. The causes of these differences are unknown. Marital selection theory posits that healthy widowers remarry quickly, leaving a less healthy subset that experiences premature mortality. Other factors, such as common

infection, shared environment, and lack of adequate daily care, also may influence the higher mortality rates of the widowed.

Studies of whether anticipatory grief or forewarning of the pending death of a spouse contributes to adjustment to bereavement have yielded conflicting results (Roach and Kitson 1989). Some suggest that anticipation is important because it allows the survivor to begin the process of role redefinition before the death, whereas unanticipated death produces more severe grief reactions. Survivors who have experienced unexpected deaths of their spouses report more somatic problems and longer adjustment periods than do those who anticipated the loss. Anticipatory role rehearsal does not consistently produce smoother or more positive adjustment among the bereaved. It appears that the coping strategies employed by survivors vary with the timing and mode of death, which in turn influence the bereavement outcome.

### **SOCIAL SUPPORT AND REINTEGRATION**

It has been suggested that social support plays an important role in the bereavement outcome and acts as a buffer for stressful life events, but the research is somewhat inconclusive, partly as a result of difficulties identifying those support efforts which produce positive outcomes and those which do not and the fact that support needs change over time. Nevertheless, there is evidence that the extent to which members of the social network provide various types of support to the bereaved is important in the pattern of recovery and adaptation (Vachon and Stanley 1988). Available confidants and access to self-help groups to assist with emotional management can help counter loneliness and promote a survivor's reintegration into society. The social resources of finances and education have been found to be particularly influential in countering the stresses associated with the death of a spouse. Community programs that provide education, counseling, and financial services can facilitate the efforts of the widowed and their families to restructure their lives.

For many older widows, a substantial period of future living alone remains: on average, another fourteen years or more. Borrowing from occupational career models, some researchers have suggested that adopting a "career of widowhood" orientation may facilitate the recovery and well-

being of these survivors: "That is, for most persons, widowhood need not be considered the end of productive life, but rather the beginning of a major segment of the life course, and one that should be pursued vigorously in order for it to be successful and fulfilling" (Hansson and Remondet 1988). In this perspective, the widowed are encouraged to seek control over their existence by actively construing their own life courses. The assumption is that they will adapt better if they plan for where they want to be at different potential stages during the entire course of widowhood. This plan might include the following phases: "a time for emotional recovery; a time for taking stock, reestablishing or restructuring support relationships, and formulating personal directions for the future; a time for discovering a comfortable and satisfying independent lifestyle, and for determining an approach to maintaining economic, psychological, and social functioning; perhaps a time for personal growth and change; and a time for reasoned consideration of one's last years and assertion of a degree of control over the arrangements surrounding one's own decline and death."

There is considerable heterogeneity among the survivor population and thus in their ability to implement a successful "career in widowhood." They differ, for example, in *relational competence*, that is, characteristics that help them acquire, develop, and maintain personal relationships that are essential for social support (Hansson and Remondet 1988). Establishing a new and satisfying autonomous identity after the loss of a spouse is never easy. The probability of achieving that goal can, however, be enhanced through counseling strategies designed for individual circumstances and programs that help survivors avoid desolation coupled with meaningful social and familial support systems.

(SEE ALSO: *Death and Dying*; *Filial Responsibility*; *Remarriage*; *Social Gerontology*)

## REFERENCES

- Arbuckle, Nancy Weber, and Brian de Vries 1995 "The Long-Term Effects of Later Life Spousal and Parental Bereavement on Personal Functioning." *The Gerontologist* 35:637-647.
- Aquilino, William S. 1994 "Later Life Parental Divorce and Widowhood: Impact on Young Adults' Assessment of Parent Child Relations." *Journal of Marriage and the Family* 56:908-922.
- Balkwell, Carolyn 1981 "Transition to Widowhood: A Review of the Literature." *Family Relations* 30:117-127.
- Bennett, Kate Mary 1997 "Longitudinal Study of Well-being in Widowed Women." *International Journal of Geriatric Psychiatry* 12:61-66.
- Berardo, Donna H. 1982 "Divorce and Remarriage at Middle-Age and Beyond." *Annals of the American Academy of Political and Social Science* 464:132-139.
- 1988 "Bereavement and Mourning." In Hannelore Wass, Felix M. Berardo, and Robert A. Neimeyer, eds., *Dying: Facing the Facts*, 2nd ed. New York: Hemisphere.
- Berardo, Felix M. 1968 "Widowhood Status in the United States: A Neglected Aspect of the Family Life Cycle." *Family Coordinator* 17:191-203.
- 1970 "Survivorship and Social Isolation: The Case of the Aged Widower." *Family Coordinator* 19:11-25.
- 1992 "Widowhood." In Edgar F. Borgatta and Marie L. Borgatta, eds., *Encyclopedia of Sociology*. New York: Macmillan.
- 1995 "Widowhood." In David Levinson, ed., *Encyclopedia of Marriage and the Family*. New York: Simon & Schuster, Macmillan.
- Campbell, Scott, and Phyllis R. Silverman 1996 *Widower: When Men Are Left Alone*. Amityville, New York: Baywood Publishing Company.
- Clark, Philip G., Robert W. Siviski, and Ruth Weiner 1986 "Coping Strategies of Widowers in the First Year." *Family Relations* 35:425-430.
- DiGiulio, R. C. 1989 *Beyond Widowhood*. New York: Free Press.
- Dimond, Margaret, Dale A. Lund, and Michael S. Caserta 1987 "The Role of Social Support in the First Two Years of Bereavement in an Elderly Sample." *Gerontologist* 27:599-604.
- Dykstra, Pearl A. 1995 "Loneliness Among the Never and Formerly Married: The Importance of Supportive Friendships." *Journal of Gerontology* 50B: S321-S329.
- Gove, Walter R., and Hee-Choon Shin. 1989 "The Psychological Well-Being of Divorced and Widowed Men and Women." *Journal of Family Issues* 10:122-144.
- Hansson, Robert O., and Jacqueline H. Remondet 1988 "Old Age and Widowhood: Issues of Personal Control and Independence." *Journal of Social Issues* 44:159-174.
- Hartwigsen, G. 1987 "Older Widows and the Transference of Home." *International Journal of Aging and Human Development* 25:195-207.

- Hiltz, Starr R. 1979 "Widowhood: A Roleless Role." In Marvin B. Sussman, ed., *Marriage and Family*. Collected Essay Series. New York: Hayworth Press.
- Hong, Lawrence K., and Robert W. Duff 1994 "Widows in Retirement Communities: The Social Context of Subjective Well-being." *The Gerontologist* 34:347-352.
- Kasper, Judith D. 1988 *Aging Alone—Profiles and Projections*. Baltimore: Commonwealth Fund.
- Kitson, Gay C., Karen Benson Babri, Mary Joan Roach, and Kathleen S. Placidi 1989 "Adjustment to Widowhood and Divorce." *Journal of Family Issue* 10:5-32.
- Levinson, Deborah S. 1997 "Young Widowhood: A Life Change Journey." *Journal of Personal and Interpersonal Loss* 2:277-291.
- Littlewood, Jane 1994 "Widows' Weeds and Women's Needs: The Re-feminization of Death, Dying and Bereavement." In Sue Wilkinson and Celia Kitzinger, eds., *Women and Health: Feminist Perspective*. New York: Taylor and Francis.
- Lopata, Helen Z. 1973 *Widowhood in an American City*. Cambridge, Mass.: Schenkman.
- 1979 *Women as Widows*. New York: Elsevier.
- 1996 *Current Widowhood*. Thousand Oaks, Calif.: Sage.
- Morgan, Leslie 1989 "Economic Well-Being Following Marital Termination: A Comparison of Widowed and Divorced Women." *Journal of Family Issues* 10:86-101.
- Murdock, Melissa E. et al. 1998 "Contribution of Small Life Events to the Psychological Distress of Married and Widowed Older Women." *Journal of Women and Aging* 10:3-22.
- Nye, F. Ivan, and Felix M. Berardo 1973 *The Family: Its Structure and Interaction*. New York: Macmillan
- Pitcher, Brian L., and Don C. Larson 1989 "Elderly Widowhood." In Stephen J. Bahr and Evan T. Peterson, eds., *Aging and the Family*. Lexington, Mass.: Heath.
- Roach, Mary J., and Gay T. Kitson 1989 "Impact of Forewarning and Adjustment to Widowhood and Divorce." In Dale A. Lund, ed., *Older Bereaved Spouses*. New York: Hemisphere.
- Sandell, Steven B., and Howard M. Iams 1997 "Reducing Women's Poverty by Shifting Social Security Benefits from Retired Couples to Widows." *Journal of Policy Analysis and Management* 16:279-297.
- Sanders, Catherine M. 1988 "Risk Factors in Bereavement Outcomes." *Journal of Social Issues* 44:97-111.
- Smith, Jack C., James A. Mercy, and Judith A. Conn 1988 "Marital Status and the Risk of Suicide." *American Journal of Public Health* 78:78-80.
- Stroebe, Margaret S., and Wolfgang Stroebe 1983 "Who Suffers More: Sex Differences in Health Risks of the Widowed." *Psychological Bulletin* 93:279-299.
- Talbott, Maria M. 1998 "Older Widows' Attitudes towards Men and Remarriage." *Journal of Aging Studies* 12:429-449.
- U.S. Bureau of the Census 1997 "Marital Status of the Population, by Sex and Age, 1996." *Current Population Reports*. Series P20-491.
- Vachon, Mary L.S., and Stanley K. Stylianos, 1988 "The Role of Social Support in Bereavement." *Journal of Social Issues* 44:175-190.
- Zick, Cathleen D., and Ken R. Smith 1988 "Recent Widowhood, Remarriage, and Changes in Economic Well-Being." *Journal of Marriage and the Family* 50:233-244.

FELIX M. BERARDO  
DONNA H. BERARDO

## WORK AND OCCUPATIONS

Work is the defining activity in people's lives. In most of the world, it is a matter of survival, but work also places people in stratification systems, shapes their physical and emotional well-being, and influences their chances for social mobility. Although the term "work" generally is used to denote the exertion of effort toward some end, economically it refers to activities oriented toward producing goods and services for one's own use or for pay. The conception of work as a means of generating income underlies most sociological scholarship on work and most of the available statistics. Unpaid productive work, including that done in the home (indeed, homemaking is the largest occupation in the United States) and volunteer work, tends to be invisible. This article focuses primarily on paid work.

### EVOLUTION OF WORK

Although contemporary work differs dramatically from work in the past, the evolution of the organization of production and people's attitudes toward work have important legacies for workers today. For much of human history, work and home lives were integrated: Most work was done at or near the home, and people consumed the products of their labor. The predecessors of the modern labor force were nonagricultural workers, in-

cluding servants and skilled artisans who made and sold products. The development of industrial work supplemented human effort with machines, introduced a division of labor that assigned specialized tasks to different workers, and ushered in a wage economy. In Europe, industrial work began as cottage industry, in which middlemen brought unfinished goods to cottagers—often women and children—who manufactured products. However, the exploitation of energy sources that could fuel large machines, the growing number of displaced peasants forced to sell their labor, and the expansion of markets for industrial goods made it more economical to shift industrial work to factories. The ensuing Industrial Revolution in the West laid the foundation for modern work and created the modern labor force. Some workers in developing countries continue to do agricultural or other subsistence work; others work in industrialized sectors, although seldom with the protections advanced industrial countries afford their workers.

### THE LABOR FORCE

In developed societies, the labor force—people who are employed or are seeking paid work—includes most adults. In Western industrialized nations, it ranged in the middle 1990s from less than half the adults in Ireland, Italy, and several Middle Eastern and north African countries to around 80 percent in Denmark, Cambodia, China, Iceland, Rwanda, Slovenia, and Burundi (United Nations 1999).

The composition of the labor force is in a continual flux. Although women and children were well represented in the earliest labor force in Western countries, as industrial labor replaced agricultural work, wage workers became increasingly male. However, as the growth of jobs labeled “women’s work” has drawn increasing numbers of women into the labor force worldwide, the labor force has become more sex-balanced. Women’s participation in the formal labor force varies cross-nationally, however. In the 1990s, according to the United Nations, women’s share of the labor force ranged from one in nine (Iran) to one in four workers (Turkey) in the Middle East and in north African countries. In Latin American countries, three to four in ten workers were female; as were 38 (Indonesia) to 48 percent (Cambodia) in southeast Asia, 38 (South Africa) to 50 percent (Burundi)

in sub-Saharan Africa, and 38 (Italy, Spain, Japan) to 48 percent (Norway, Denmark, Sweden) in the advanced industrial countries.

In the United States, more than 46 percent of the labor force was female in 1998. Just as the U.S. labor force has become more diverse in gender, it has become more diverse in its racial and ethnic composition. As the “baby bust” cohorts replace the baby boom cohorts, the U.S. labor force is aging. Smaller cohorts of young workers will lead employers to turn to other labor sources, such as immigrants, to fill low-wage, entry-level jobs.

Although child labor has all but disappeared in advanced industrial nations, children are a significant presence in the labor force in many developing countries. According to the International Labour Organization (ILO) (1996, number 16), in the mid-1990s, three to four of every ten sub-Saharan African children between ages 10 and 14 worked to help support themselves and their families. In some Asian countries, more than three children in ten are in the labor force (Bangladesh, Bhutan, East Timor, Nepal), and in several Latin American countries, at least one child in four works for pay (Bolivia, Brazil, Dominican Republic, Guatemala, Haiti, and Nicaragua). Child labor in Third World countries is partly a product of a global economy that makes impoverished children particularly attractive to Western-based multinational corporations in their worldwide search for cheap and docile workers.

**Extent of paid work.** The amount of time people spend at paid work has changed through the centuries. In the early decades of industrialization, adults and children often worked fourteen-hour days, six days a week. After labor organizations won maximum-hours laws and overtime pay, the average workweek shrank for European and American workers, although in some countries hours are increasing for some workers. In 1997, Japan’s workers logged more hours of work than did those in other countries for which records are available, averaging 1,990 hours annually, with U.S. workers second at 1,904 hours. Germany’s and Denmark’s workers average the fewest hours of paid work per year: 1,573 and 1,665, respectively (ILO 1998, number 25, p. 31). Declines in work hours mask a division in the extent of paid work, with growing numbers either putting in very long workweeks or working part-time. In many



industrialized countries, the proportion of workers employed part-time has doubled since the 1970s; indeed, the growth in part-time jobs almost entirely accounts for the growth of total employment in industrialized countries (ILO 1996, number 17, p. 28). This growth reflects both demand-side and supply-side forces. Teenagers and women with children disproportionately opt for part-time jobs to leave time for school or unpaid family work, and employers structure some jobs as part-time to avoid paying fringe benefits. The increased number of jobs structured as part-time has caused growth in the number of persons who work part-time involuntarily.

#### **Unemployment and underemployment.**

Throughout history, people seeking adequately paid employment usually have outnumbered jobs, leaving some would-be workers unemployed or underemployed. According to the ILO (1998, number 27, p. 6), one-third of the world's workers are underemployed (850 million persons) or unemployed (150 million persons), and unemployment is in the double digits in several countries, including Botswana, Spain, Finland, Puerto Rico, Barbados, and Poland (United Nations 1999). Globalization contributes to unemployment as multinational companies draw people in developing countries into the labor force and then put them out of work when they close plants in pursuit of cheaper labor (Dickinson 1997). In 1996, unemployment in the industrialized countries ranged from 4 percent in Norway to over 11 percent in Germany, with intermediate levels in Sweden, the United Kingdom, and the United States. (At the end of 1998, U.S. unemployment had fallen to 4.3 percent, although the rates for racial and ethnic minorities and youth were much higher.) In general, official statistics in industrialized countries underestimate unemployment by excluding "discouraged workers" who have stopped looking because they cannot find jobs for which they qualify.

**Preparing for jobs.** Workers' education and training affect the jobs they obtain. Schools teach vocational skills (including literacy and numeracy), inculcate traits that employers value (e.g., punctuality, ability to deal with bureaucracies), and provide credentials that signal the ability to acquire new skills. Vocational education provides skills and certification. In Germany, for example, vocational training is a major source of workers' skills. In the United States, in contrast, many workers—

especially those in traditionally male blue-collar jobs—acquire most of their skills on the job, whereas professionals and clerical workers acquire their skills largely before beginning employment. Jobs in advanced industrial societies—especially high-technology jobs—tend to require both more and different kinds of skills, such as precision and flexibility, as well as formal knowledge (Hodson and Parker 1988). In postindustrial societies, knowledge and technical expertise have become increasingly important for good jobs. As a growing number of jobs require at least some college, workers without a high school diploma face difficulties finding jobs that pay well and provide advancement opportunities. Moreover, workers displaced from production jobs need new skills for reemployment, and so refraining has become increasingly important.

**Job outcomes.** The processes that allocate workers to occupations, employers, and jobs are important because those elements strongly affect workers' earnings. Although thousands of distinct labor markets serve different locales and occupations, to understand the job-allocation process, it is necessary to distinguish primary markets that fill jobs characterized by high wages, pleasant working conditions, the chance to acquire skills, job security, and opportunities to advance from secondary markets that fill low-paid, dead-end, low-security jobs. Firms in the primary sector fill non-entry-level jobs through internal labor markets that provide employees with "ladders" that connect their jobs to related jobs higher in the organization. The failure of secondary-market jobs to provide job ladders that reward seniority, along with low pay and poor working conditions, encourage turnover (Gordon 1972). Both statistical discrimination and prejudice disproportionately relegate certain workers—the young, inexperienced, and poorly educated; racial and ethnic minorities; immigrants; and women—to jobs filled through secondary labor markets.

## **WORK STRUCTURES**

In classifying the paid work people do, social scientists refer to industries, occupations, establishments, and jobs. An industry is a branch of economic activity that produces specific goods or services. An establishment is a place where employees report for work, such as a firm or plant. An

occupation refers to a collection of jobs involving similar activities across establishments, whereas a job is a set of similar work activities performed at a specific establishment. In 1990, the U.S. Census Bureau distinguished 503 “detailed” occupations (for example, funeral director, meter reader, x-ray technician) that it grouped into six broad categories: managerial and professional specialties; technical, sales, and administrative-support occupations; service occupations; precision production, craft, and repair occupations; operators, fabricators, and laborers; and farming, forestry, and fishing occupations. The steady growth in the number of occupations since the Industrial Revolution reflects the increasing division of labor in complex societies. This elaboration of the division of labor is more visible at the job level. The U.S. Department of Labor’s *Dictionary of Occupational Titles* lists several thousand job titles, and the approximately 130 million employed Americans hold about a million different jobs.

**Occupational structure.** The distribution of workers across occupations in a society provides a snapshot of that society’s occupational structure. Comparing societies’ occupational and industrial structures at different times or across nations reveals a lot about their economic and technological development and the job opportunities available to their members. For example, in 1870, agriculture employed half of all American workers; in the 1990s, it provided jobs for about 2 percent. The effects of changing occupational and industrial structures—driven largely by the disappearance of smokestack industries and the explosion of service jobs in the United States—are expressed in the sharp decline in a worker’s chances of getting a unionized skilled production job. Hit hardest by the dwindling number of these jobs are the white men who once monopolized them. In contrast, the growing number of management jobs in the United States created a record number of managerial positions in the 1990s. This growth has helped to integrate managerial jobs by sex and race.

**Job segregation by sex and race.** One of the most enduring features of paid work is the differential distribution of male and female and white and minority workers across lines of work and places of employment, with minorities and white women concentrated in the less desirable jobs (Carrington and Troske 1998a, 1998b). In 1990, among all gainfully employed women in the United

States, 28 percent were concentrated in just 5 of the 503 detailed occupational categories—secretary, bookkeeper, manager, clerk, and registered nurse—and over half worked in just 19 of the 503 occupations distinguished by the Census Bureau. Men, in contrast, are spread more evenly across occupations: The top five—manager/administrator, production supervisor, truck driver, sales supervisor, and wholesale sales representative—accounted for 19 percent of all employed men. However, within-occupation sex segregation (many jobs share a single occupational title) means that job segregation is considerably more pervasive than is occupational segregation.

In every country, the sexes are segregated into different jobs, although the extent of occupational sex segregation varies sharply across nations: It is highest in Middle Eastern and African nations and lowest in Asian/Pacific nations (Anker 1998). In advanced industrial nations, it correlates positively with women’s labor force participation, paid maternity leave, and the size of the wage gap (Rosenfeld and Kalleberg 1990). Levels of sex segregation in European countries reflect both postindustrial economic structures that concentrate women in sales and service jobs and adherence to norms of gender equality. Independent of these forces, customarily male production jobs remain outside the reach of most women, and women continue to dominate clerical occupations (Charles 1998).

The last thirty years has witnessed worldwide declines in occupational sex segregation (Anker 1998). Integration occurs primarily through women’s entry into customarily male occupations rather than the reverse. Falling levels of occupational sex segregation can mask ongoing job-level segregation (Reskin and Roos 1990). Training workers for nontraditional jobs and enforcing antidiscrimination laws and affirmative-action regulations appear to be the most effective remedies for reducing sex segregation.

Occupations and jobs also are segregated by race. For example, before World War II, American blacks were concentrated in farming, service, and unskilled-labor jobs in the secondary sector of the economy, such as domestic worker, porter, and orderly. War-induced labor shortages opened the door to a wider range of jobs for blacks, and antidiscrimination regulations (especially Title VII of 1964 Civil Rights Act) further expanded blacks’

opportunities. As a result, racial segregation across occupations has declined sharply in the United States since 1940, especially among women. There is little systematic cross-national research on job segregation or job discrimination by race, although scattered studies document both around the world. For example, Moroccans are excluded from semi-skilled jobs in the Netherlands, West Indians face discrimination in Canada, and Vietnamese and the aboriginal populations encounter it in Australia (ILO 1995, number 12, pp. 29–30).

Workers' experience and preferences influence where they work and what they do, but at least as important are the operation of labor markets—the mechanisms that match workers to jobs and set wages—and employers' preferences and personnel practices. Sociologists have documented the importance of personal networks for workers' employment outcomes (e.g., Fernandez and Weinberg 1997). Employers favor the use of social networks to recruit workers because of their efficiency, low cost, and ability to provide information unavailable through formal sources. However, because people's acquaintances tend to be of the same sex and race, recruiting through employees' networks effectively excludes sex- and race-atypical workers.

**Layoffs and Displacement.** U.S. data for the 1980s and 1990s indicate that trends in job displacement rates roughly parallel those for unemployment. Between 1993 and 1995, 12 to 15 percent of workers lost a job because their companies closed, their jobs were cut, or work was slack. Depending on economic conditions, between 25 and 40 percent of displaced workers remain jobless one to three years later, and reemployed workers typically earn less than they did in their previous jobs (Economic Report of the President 1999).

## REWARDS FOR EMPLOYMENT

People seek jobs that maximize extrinsic rewards—income, prestige, the chance for promotion, and job security (Jencks et al. 1988)—as well as intrinsic rewards—satisfaction and autonomy. Earnings are the primary incentive for most workers. However, pay differs sharply across individuals and social groups. Substantial racial, sex, and ethnic inequality in pay characterize all industrial societies, although their extent depends on whether

countries permit unequal pay for equal work and the degree to which workers are segregated into unequally paying jobs on the basis of sex, race, or ethnicity. In the United States, the 1963 Equal Pay Act that outlawed wage discrimination by race, national origin, and sex and declining occupational segregation by race have reduced the racial gap in earnings among men and almost eliminated it among women. The disparity in earnings between the sexes has declined more slowly because of the resilience of sex segregation. Hence, in 1998, women employed full-time year-round earned 74 percent of the annual earnings of their male counterparts. The wage gap varies across nations (and across occupations and industries within countries). In the first half of the 1990s, pay inequality was lowest in Australia, Egypt, Kenya, Jordan, and New Zealand, where women averaged about 80 percent of what men earned, compared to a low of just 60 percent in Korea (ILO 1997, number 22). Factors that can reduce the wage gap among full-time workers include equalizing the sexes' educational attainment and labor-market experience, creating sex-integrated jobs, and implementing pay systems that compensate workers for the worth of a job without regard to its sex composition.

**Occupational prestige.** Social standing is conferred on persons partly on the basis of their jobs. In fact, social scientists have treated the distinction between blue-collar and white-collar jobs as a rough proxy for workers' social status. However, to capture the effects of one's type of work on one's social status, more sophisticated ways to measure occupational prestige are needed. The most commonly used is the Duncan Socioeconomic Index (SEI) (Duncan 1961), which assigns a score to each occupation on the basis of its incumbents' average educational and income levels. The occupational status hierarchy is quite stable over time and across cultures (Treiman 1977). Within societies, the occupational standing of workers is highly stratified. In the United States, for example, most workers have occupations with relatively low SEI scores.

**Intrinsic rewards: job satisfaction.** In advanced industrialized countries, many workers see a job as a place to find fulfillment, self-expression, and satisfaction. Workers in routine jobs try to imbue them with challenge or meaning, in part by creating a workplace culture. These adaptations contribute to the high levels of satisfaction Americans report with their jobs. Nonetheless, not all jobs are

satisfying, and not all workers are satisfied. On the assumption that dissatisfied workers are less productive, employers in the United States and other advanced industrialized countries have devised strategies such as workplace democracy, job-enrichment programs, and “quality circles” to enhance job satisfaction. According to Lincoln and Kalleberg (1990), however, Japanese and German workers, who show the lowest levels of satisfaction, are among the world’s most productive.

### WORK AND FAMILY

In expanding the factory system, the Industrial Revolution separated work and family, creating a division of labor that mandated domestic work for women and market work for men. Although women increasingly hold paid jobs, paid employment has not exempted them from primary responsibility for domestic work. Role overload and its concomitant stresses are risks for all workers, but especially for employed mothers, who accounted for 70 percent of married mothers and 60 percent of single mothers in the United States in 1996. Women have adapted by working part-time, sacrificing leisure time, renegotiating the domestic division of labor in their families, cutting out some domestic tasks, and purchasing more services. (The trend toward purchasing more services has fueled the growth of service jobs in fast-food chains, child care, and cleaning services and thus has increased the demand for low-wage workers.) What employed parents want most is flexible scheduling (Glass and Estes 1997), although organizational pressure prevents some from taking advantage of it when it is available (Hochschild 1997).

Just as paid work competes with workers’ domestic obligations, the demands of family life interfere with workers’ ability to devote themselves entirely to their jobs. Thus, employers have two incentives to reduce work–family conflicts: reducing absenteeism and turnover and increasing workers’ productivity and organizational commitment (Glass and Estes 1997). Many employers in advanced industrial societies have provided some of their employees with assistance with child care. The governments of most advanced industrialized countries have mandated programs such as parental leave, state-run nurseries, and guaranteed benefits for part-time workers. Among the 152 member nations of the ILO, only two advanced industrial

countries provide no paid maternity leave: New Zealand and the United States (ILO 1998, number 24, pp. 18–19). Employers’ increasing reliance on female workers and politicians’ desire for women’s support should bring more family-friendly policies and practices in the twenty-first century.

### TRENDS IN WORK

**Control of work.** As Marx recognized, whenever different actors control the tools of production and perform work, control over the work process is potentially a matter of contention. Employers have relied on a variety of tactics to control the labor process: paternalism, close supervision, embedding control into the technology of work, deskilling work, and bureaucratic procedures such as career ladders (Edwards 1979). Workers have resisted more or less effectively through collective action, including attempts to create a monopoly of their skills or the supply of labor. At the end of the twentieth century, several factors had given employers the upper hand in the struggle for control, including the decline of labor unions in Western industrialized societies (ILO 1997, number 22, p. 7), the disappearance of lifetime job protection in formerly communist societies, an increasing technological capacity to monitor workers electronically, access to a global “reserve labor army,” and the use of nonstandard employment relationships (see below).

**Technological change.** The history of work is a chronicle of technological innovation and its transformation of the production of goods and services. Employers invest in technology to increase productivity, contain labor costs, and control how work is done. According to some observers (e.g., Braverman 1974), employers seek technical advances in order to reduce workers’ control over the labor process and employ less skilled and thus cheaper labor. Some analysts see technological change as a threat to skilled jobs; others see technology as creating more of those jobs. The development of microelectronic technology has brought this debate to the fore.

Innovations in microprocessor technology have permitted advances in information processing and robotics that are revolutionizing the production of goods and services. Robots work around the clock, perform hazardous tasks, and have low operating costs. Although technical advances enhance jobs,

they also subject workers to technological control (an estimated 80 percent of U.S. workers are electronically monitored, for example [ILO 1998, no. 24, p. 25]) and, by improving productivity, lead to job losses. In industrialized nations, for example, microelectronic technology has eliminated some unskilled jobs and facilitated work transfers that shift tasks from paid workers to consumers such as banking transactions. By making it possible to export jobs to cheaper labor markets, technology has reallocated jobs from the workers who once performed them to lower-paid workers in other parts of the world. Although technological change has created jobs, it has eliminated more job—particularly less-skilled ones—than it has created and has eroded skills in middle- to low-skill jobs such as clerical work (Hodson and Parker 1988). Its creation of new highly skilled jobs has contributed to the economic polarization of workers and spurred the migration of well-educated workers from developing to advanced industrial countries (Hodson 1997).

**The globalization of work.** Although segments of the economy such as service work are organized locally, production work increasingly is organized in a global assembly line (Dickinson 1997). For jobs in which technology preempts skill, multinational corporations' worldwide pursuit of low-wage docile labor and microelectronic technology and cheap transportation reduce the friction associated with moving production around the globe. As a result, there has been a steady exportation of jobs from industrialized countries to the Pacific Rim, Latin America, and the Caribbean, where labor is cheap and tractable and labor laws are lenient. This redistribution of manufacturing jobs from advanced industrial nations to developing nations—fueled by the growth of multinational corporations—has given birth to an international division of labor in which the United States and other advanced industrial nations have become postindustrial societies that specialize in producing services rather than goods, while workers in less developed countries manufacture products, often under unsafe conditions. For example, between 1980 and 1993, semi-industrialized and industrialized countries in the Americas and Europe lost 30 to 70 percent of their jobs in the textile and footwear industries, while African and Asian countries have experienced astronomical job growth in those industries (ILO 1996, vol. number 18). Be-

tween 1970 and 1990, the number of textile and footwear jobs doubled, tripled, or more in Korea, Indonesia, Sri Lanka, Bangladesh, and Malaysia where production workers earn from one-fifth to one-half as much as do their counterparts in advanced industrial countries. Meanwhile job growth in developing countries leads to the disproportionate employment of teenagers and young adult women, who work for lower pay than do adult men.

Just as jobs move in search of cheaper workers, workers move in search of better-paying jobs. Often, however, the outcome of this migration is low-skilled, low-paid employment in domestic or service work. However, skilled technical and professional jobs also draw workers in global migration streams. In the mid-1990s, according to the ILO (1995, number 13), 70 million immigrants—most from the Third World—resided in countries other than their nations of birth. The globalization of competition among employers has made workers on different continents into competitors for jobs, held down wages, and militated against campaigns to improve working conditions in Third World establishments while eroding job security in First World production facilities (Hodson and Parker 1988; Dickinson 1997).

**The externalization of work and the erosion of jobs.** By the middle of the twentieth century, the normative employment relationship between employers and workers had become standardized in many industrialized societies. This standard employment arrangement typically involves the exchange of labor by a worker for a fixed rate of pay (hourly wages or a weekly, monthly, or annual salary) from an employer, with the labor performed on a preset schedule—usually full-time—at the employer's place of business, under the employer's control, and often with the shared expectation of continued employment. However, to cut costs and enhance flexibility, employers are increasingly "externalizing" work in terms of physical location administrative control and the duration of employment (Pfeffer and Baron 1988).

This externalization is seen in the increasing number of persons working for pay at home and the growth of nonstandard employment relationships (Barker and Christensen 1998). Neither home-work nor nonstandard employment relations are new. Only after unions won the right to bargain collectively and statutory rights protecting work-

ers did homework and nonstandard employment relations give way to standard employment relationship in advanced industrialized countries. In the 1990s, however, the trend seemed to have reversed. In 1995, for example, eight million Americans, at least two million Europeans, six million Filipinos, and one million Japanese worked for pay at home. Millions of these homeworkers telecommute (ILO 1998, number 27, p. 23). Many workers, especially women, opt for homework as a way to earn wages while supervising their children (Jurik 1998). However, part of the price of this flexibility is a lack of protection by health and safety regulations or maximum-hours rules, and these workers are outside the reach of organizing efforts. In addition, whether homework involves children is difficult to monitor even in countries strongly opposed to child labor (ILO 1995).

Work is externalized in a second way: Firms contract with individuals for specific duties (independent contractors) or with intermediary organizations that employ workers rather than directly employing all the persons who do work for them. Although contracting has long been common for some forms of work (e.g., agricultural labor), employers around the world are increasingly contracting out jobs formerly done by their own employees in everything from construction and manufacturing to human resources and security. Worldwide, more than one in four service workers are contract laborers. By outsourcing these functions to contract workers or independent contractors, employers avoid the obligation to provide long-term employment and short-circuit protective labor laws that apply to employees. Other nonstandard employment relationships include temporary work and part-time work, both of which disproportionately employ women, members of racial and ethnic minorities, and young workers.

The growth of nonstandard employment relationships has led some observers to predict an end to work organized through standard employment relations or the bifurcation of employment relations, with firms hiring core workers who enjoy the benefits of standard employment and creating explicitly temporary connections with peripheral workers who lack benefits and job security (Smith 1997; Leicht 1998). According to the U.S. Bureau of Labor Statistics, the proportion of U.S. workers in nonstandard work is slowly increasing (Barker

and Christensen 1998). In summary, employment relations must be seen as falling on a continuum from long-term attachments under bureaucratic control to weak connections of uncertain duration (Pfeiser and Baron 1988). Most research on work and occupations in industrial societies has dealt with the former end of the continuum. Technological change and globalization are shifting jobs—even in industrial countries—toward the latter end.

## REFERENCES

- Anker, Richard 1998 *Gender and Jobs: Sex Segregation of Occupations in the World*. Geneva: International Labour Office.
- Barker, Kathleen, and Kathleen Christensen 1998 *Contingent Work: American Employment Relations in Transition*. Ithaca, N.Y.: Cornell University Press.
- Carrington, William J., and Kenneth R. Troske 1998a "Sex Segregation across U.S. Manufacturing Firms." *Industrial and Labor Relations Review* 51:445–464.
- , and Kenneth R. Troske 1998b "Interfirm Segregation and the Black/White Wage Gap." *Journal of Labor Economics* 16:231–260.
- Charles, Maria 1998 "Structure, Culture, and Sex Segregation in Europe." *Research in Social Stratification and Mobility* 16:89–116.
- Dickinson, Torry D. 1997 "Selective Globalization: The Relocation of Industrial Production and the Shaping of Women's Work." In Randy Hodson, ed., *Research in the Sociology of Work—Globalization of Work*, vol. 6. Greenwich Conn.: JAI Press.
- Duncan, Otis Dudley 1961 "A Socioeconomic Index for All Occupations." In Albert J. Reiss, Otis Dudley Duncan, Paul K. Hatt, and Cecil C. North, eds., *Occupations and Social Status* New York: Free Press.
- Economic Report of the President 1999 Washington, D.C.: U.S. Government Printing Office.
- Edwards, Richard 1979 *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*. New York: Basic Books.
- Fernandez, Roberto M., and Nancy Weinberg 1997 "Sifting and Sorting: Personal Contacts and Hiring in a Retail Bank." *American Sociological Review* 62:883–902.
- Glass, Jennifer L., and Sarah Beth Estes 1997 "The Family Responsive Workplace." *Annual Review of Sociology* 21:289–313.
- Gordon, David M. 1972 *Theories of Poverty and Unemployment*. Lexington, Mass.: Lexington Books.

- Hochschild, Arlie 1997 *The Time Bind: When Work Becomes Home and Home Becomes Work*. New York: Metropolitan.
- Hodson, Randy 1997 "Introduction: Work from a Global Perspective." In Randy Hodson, ed., *Research in the Sociology of Work—Globalization of Work*, vol. 6. Greenwich Conn.: JAI Press.
- , and Robert E. Parker 1988 "Work in High Tech Settings—A Literature Review." In Richard L. Simpson and Ida Harper Simpson, eds., *Research in the Sociology of Work*, vol 4. Greenwich Conn.: JAI Press.
- International Labor Organization 1995 *Homework*. Report for the eighty-two second Session of the International Labour Conference, ILO, Geneva, Switzerland.
- 1995–1998 *World of Work: The Magazine of the ILO*. Selected issues as noted: Geneva, Switzerland.
- Jencks, Christopher, Lauri Perman, and Lee Rainwater 1988. "What Is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93:1322–57.
- Jurik, Nancy 1998 "Getting Away and Getting By: The Experiences of Self-Employed Homeworkers." *Work and Occupations* 25:7–35.
- Leicht, Kevin 1998 "Work (If You Can Get It) and Occupations (If There Are Any)? What Social Scientists Can Learn from Predictions of the End of Work and Radical Workplace Change." *Work and Occupations* 25:36–48.
- Lincoln, James R., and Arne L. Kalleberg 1990 *Culture, Control and Commitment: A Study of Work Organization and Work Attitudes in the U.S. and Japan*. New York: Cambridge University Press.
- Pfeffer, Jeffrey, and James N. Baron 1988 "Taking the Workers Back Out: Recent Trends in the Structuring of Employment." *Research in Organizational Behavior* 10:257–303.
- Reskin, Barbara F., and Patricia A. Roos 1990 *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Rosenfeld, Rachel A., and Arne L. Kalleberg 1990 "A Cross-National Comparison of the Gender Gap in Income." *American Journal of Sociology* 96:69–106.
- Smith, Vickie 1997 "New Forms of Work Organization." *Annual Review of Sociology* 21:315–339.
- Treiman, Donald J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
- United Nations 1999 Statistics Division Home Page, Social Indicators Home Page. [www.un.org/depts/unsd/social](http://www.un.org/depts/unsd/social).

BARBARA F. RESKIN

## WORK ORIENTATION

The sociology of work emerged as a specialty area in the 1980s, when the American Sociological Association prepared a compendium of course syllabi for the area and a number of textbooks appeared. The name of this sociological subfield is new, but the general area is not. The sociology of work represents an integration of two long-standing specialties: industrial sociology and occupations/professions. It also draws from industrial and organizational psychologists and sociologists' attempts to integrate stratification and organization literatures to better understand the employment relationship.

The study of the employment relationship encompasses a multitude of topics ranging from how the individual is initially matched to a job to all that happens on the job (being paid, becoming satisfied or dissatisfied, forming cliques, etc.) and to turnover (quitting or being dismissed). Considered important to these topics are the orientations employees have toward their work, the topic of this article.

Definitions of work abound, but most include the following features. First, although groups or collectivities may be viewed as actors involved in work (e.g., work groups, task groups, teams, or committees), the focus of attention, and therefore the unit of analysis, is usually the individual. Second, the individual is involved in physical or mental activity. Third, this activity usually involves some form of payment, but pay is not necessary for an activity to be considered work. This allows people involved in housekeeping activities to be included, along with family members who labor to support a family enterprise and volunteer helpers. Fourth, the activity involves the production or creation of something. Fifth, this usually is a good or service. Sixth, this good or service is valued by the individual or others and thus usually is consumed by either or both. Work thus is defined as the mental or physical activity of an individual directed toward the production of goods or services that are valued by that individual or others.

"Orientation to work," unfortunately, is a term without a clear or precise meaning. Generally, it is used to refer to two broad areas: (1) motivation to work and (2) responses to work. The first area covers why people work and for some time has

occupied the attention of industrial and organizational psychologists, who analyze need hierarchies, self-actualization, and intrinsic and extrinsic motivations. The second area has more often attracted the attention of sociologists. It takes the activity of work as given and addresses the ways in which individuals react to it. Job satisfaction and commitment have been given the most attention when sociologists study reactions to work.

This article is organized around work motivation and responses to work, but it places those topics the context of the social organization of the workplace. With only a few exceptions, work occurs in a social setting that has been called a "contested terrain" by Edwards (1979). Sociologists want to go beyond strictly individualistic portrayals of human behavior and are especially interested in understanding how this social setting, the workplace, affects an individual's work orientation. The explanations of this influence, often referred to as social control arguments, are discussed here. Finally, gender differences in work orientations need to be addressed. However, because the concept of alienation is related to job dissatisfaction and has been so prevalent in sociological accounts of work, it is considered first.

### ALIENATION

Sociologists continue to draw from Marx in referring to an alienated individual as being separated or estranged from certain aspects of work that give meaning and significance to that work and to life as a whole. For Marx, these aspects of work are control over the product, control over the work process, creative activity, and social relations with others. Clearly, a negative side of work is portrayed when alienation is the concept of interest.

A survey of journals and sociology of work texts over the past several decades suggests that sociologists have lost interest in this concept. For example, indexes for 1980s texts (e.g., Kalleberg and Berg 1987) do not include the term "alienation," and the Price and Mueller (1986) handbook on the measurement of major organization concepts does not devote a chapter to alienation. Even in 1990s texts and anthologies (e.g., Hodson and Sullivan 1995; Wharton 1998), alienation is given only limited attention.

This does not mean that interest in alienation is dead. Three things have happened. First, interest has shifted to conceptualizing and measuring positively worded concepts such as like job satisfaction. Second, scholars have moved away from the picture of capitalist work settings universally producing alienated workers and gone on to formulate a picture of multidimensional work settings and multimotivated employees who respond to work in varying ways. Third, out of this more pluralistic image of work, several concepts—for example, work motivation, self-actualization, job satisfaction and commitment—have emerged in an attempt to bring more precision to descriptions of how individuals are oriented to their work. Thus, alienation has been absorbed into several other concepts.

A particular line of research has implications for understanding alienation: Following more a Marxian picture of work, it has been assumed that the more formal and bureaucratic the workplace is, the more alienated (dissatisfied) the workers are. This assumption has been challenged with the argument that formal rules and regulations actually increase satisfaction in the workplace because they provide guidelines that apply to all and thus protect workers from arbitrary and unfair treatment. Although workers may not like the rules, the authority system is perceived as legitimate because all workers are treated according to the same formal rules. Research supports this more positive portrayal of formal rules and regulations.

### WORK MOTIVATION

Historically, sociologists have flirted with psychological concepts such as work motivation and work involvement and have disagreed about the relevance of those concepts to the study of social phenomena. For example, among the authors of the sociology of work textbooks in the past two decades, only Hall (1986) gives critical attention to the theoretical and empirical literature on the topic. Any treatment of work orientation must include this material, however, because most current literature is an offshoot of or a reaction to those theories.

Work motivation is the internal force that activates people to do the work associated with their jobs. Two theoretical traditions have been dominant. First, need theories argue that individu-



als are motivated by internal needs that usually develop early in life and often are not consciously recognized. Maslow (1954) identified a hierarchy of needs and claimed that higher-order needs (goals) cannot be met until lower-order needs are met sequentially. This hierarchy begins at the bottom with basic physiological needs and ends at the top with self-actualization. Others have modified Maslow's hierarchy into a continuum with fewer levels and with the idea that lower-order needs may reemerge at later stages as unmet. Herzberg (1966) was more interested in job satisfaction and argued that individuals are motivated by two types of factors: "Motivators" are the more intrinsic features of work, such as responsibility, advancement, and achievement, whereas "hygiene" factors characterize the workplace and include pay, job security, and working conditions. When motivators are present, employees are satisfied, but if they are absent, employees are not. When the hygiene factors are present, employees are neither dissatisfied nor satisfied, but when they are absent, employees are dissatisfied. McClelland (1961) argued that certain socialization environments produce a need for achievement and that individuals socialized in that manner strive for excellence in whatever they undertake. Management scholars were especially interested in this theory since it suggested who should be hired or promoted. Finally, McGregor (1960) argued that assumptions about human nature and motivation have resulted in two approaches to organizational design. Theory X is based on the assumption that individuals are basically lazy and are motivated primarily by extrinsic rewards such as pay. Theory Y assumes that humans act responsibly and contribute their skills and talents when their intrinsic needs, such as self-actualization, are met. This distinction is not unlike the classic dichotomy between functionalist and Marxian portrayals of society and human nature.

Overall, these need theories have lost favor. The empirical support is weak, the use in applied settings has proved difficult because of problems associated with measuring need levels and attempting to alter personality patterns that have developed in childhood, and the significance of the environment has been neglected.

The second dominant perspective—expectancy theory—comes from organizational and industrial psychologists. It bypasses the issue of needs

and emphasizes cognitive and rational processes. The underlying assumption is that motivations to work vary substantially from one individual to the next and are mutable across time and space (Vroom 1964; Lawler 1973). Motivations reflect the interplay of effort, expectations about outcomes, and the importance or value given to those outcomes. Put another way, a person's motivation to behave in a particular way is a function of the expected results and how valuable those results are to that person. Until recently, this theory has been dominant in studying work motivation in industrial and organizational psychology.

Sociologists are generally aware of these motivation theories and, like psychologists, now give less attention to need theories. However, unlike psychologists, they have not been overly interested in the theories per se of work motivation. In fact, psychologists have led the way in developing theories of motivation, and sociologists usually are a generation behind in adopting or rejecting those theories. For example, Smither (1988) mentions equity, behavioral, and goal-setting theories as receiving much attention in the psychological work motivation literature. Although equity theory has been explored for some time experimentally by sociologists, there is no evidence that sociologists have adopted in significant way any of these "newer" approaches to work motivation. What sociologists do in practice matches the expectancy model more closely. The picture is one in which "the fit" of an individual's characteristics and expectations with the actual work conditions forms the basis for whether that individual is motivated.

What sociologists have emphasized instead of motivation theory is socialization to work, that is, how individuals learn their work roles. This is not surprising given the long-standing interest of both sociologists and social psychologists in socialization processes. One stream of thought in this area concerns socialization into professional roles, where a popular strategy is to examine career stages. Another approach is represented by the work of Kohn and Schooler (1982), who not only argue for the intergenerational class-based transmission of work values but also propound and demonstrate reciprocal effects: An individual's work orientations (e.g., self-direction) are affected by job conditions, but those orientations also affect the kinds of jobs with which the individual is associated.

**RESPONSE TO WORK: JOB SATISFACTION**

Although the wording of definitions for “job satisfaction” has varied dramatically across disciplines and scholars, there is a near consensus on what the concept is. Smith et al. (1969) succinctly define it as the degree to which individuals like their jobs. The common element across definitions like this is the idea of the individual positively responding emotionally or affectively to the job.

The major issues in the study of job satisfaction are (1) What produces job satisfaction? (2) What are the consequences of differing levels of job satisfaction? and (3) Is it a global or unitary concept, or should facets (dimensions) of it be investigated?

Two dominant arguments exist regarding the determinants of job satisfaction. The first is that an individual’s job satisfaction is determined by the dispositions or “personality” traits that an employee brings to the workplace. In simple terms, individuals vary along a continuum from a negative to a positive orientation. These dispositions are reflected in a person’s responses to work conditions as well as to aspects of life such as family satisfaction and more general life satisfaction. The second argument is considered more “sociological” and emphasizes the importance of the work conditions an employee experiences. This approach is closer to a Marxian perspective in that it is the structural conditions of the workplace that make work rewarding or not rewarding; any individual dispositional differences that exist wane in importance in the face of these structural features.

Although sociologists give lip service to the disposition argument, the literature unequivocally documents a stronger interest in identifying the features of work that affect job satisfaction. Within this perspective, however, there is considerable disagreement about which features of work are important. One major debate concerns whether extrinsic (e.g., pay and fringe benefits) or intrinsic (e.g., self-actualization and task variety) features of work are more important. Following a needs framework or arguments from neoclassical economics about economic rationality leads one to argue that the extrinsic features must exist before the intrinsic features become important. In contrast, an expectancy argument would state that any of these features can be important and that it is the fit of what is found in the workplace with what the

individual expects and values that is crucial in determining the satisfaction level. A popular argument that has an expectancy logic associated with it comes from the justice literature. A theme common to all distributive justice theories is that an individual compares his or her actual reward with what is believed to be just or fair. Individuals expect a just reward and are dissatisfied if a reward is unjust. Another frequently used general perspective for understanding the effect of work conditions on job satisfaction is social exchange theory, which also relies on an expectancy logic. As developed initially by Homans (1958) in the study of small groups and extended to the study of organizations by Blau (1964), exchange theory argues that individuals enter social relations in anticipation of rewards or benefits in exchange for their inputs and/or investments in the relationship. Simply put, workers are satisfied with their jobs if the rewards they value and expect are given to them in exchange for their work effort and performance.

It is impossible to summarize here the thousands of studies conducted on the determinants of job satisfaction. Instead, a list of variables that have been found to have some relationship with job satisfaction is provided (the sign indicates the direction of the relationship with regard to satisfaction): variety (+), pay (+), autonomy (+), instrumental communication (+), role conflict (-), role overload (-), work group cohesion (+), work involvement (+), distributive justice (+), promotional opportunities (+), supervisory support (+), task significance (+), and external job opportunities (-). Spector (1997) provides a more complete account of the determinants and correlates of job satisfaction.

The debate over which work conditions affect job satisfaction continues to direct the research of sociologists, but a more interesting question involves the disposition versus situation debate. Sociologists devote much effort to cataloging and operationalizing the objective structural features of work, and little attention is given to identifying and measuring the dispositional traits of individuals. Evidence, however, continues to mount that individuals exhibit basic dispositional traits (e.g., negative and positive affectivity) that are relatively stable throughout their lifetimes and over different employment situations (Watson and Clark 1984). This research strongly suggests that work-

ers with positive dispositions usually are more satisfied with their jobs regardless of the work conditions, while those with negative dispositions seem not to be satisfied with anything.

Another issue concerns the consequences of job satisfaction. Two outcomes have received the most attention, primarily because of their practical significance to any business enterprise: job performance and withdrawal behavior, which includes absenteeism and voluntary turnover. The satisfaction–performance argument is of long-standing interest and thus has generated considerable empirical data. The hypothesis is that satisfaction is positively and causally related to productivity, and support is provided by meta-analyses showing a positive correlation of .25. In short, satisfied workers perform better, but the relationship is not a strong one. The weakness of this relationship could be due to the difficulties associated with measuring job performance, however.

With regard to the satisfaction–withdrawal relationship, the hypothesis is that the most satisfied employees will be the least often absent and the least likely to quit voluntarily. The meta-analyses for the satisfaction–absenteeism relationship suggest that the relationship is between  $-.10$  and  $-.15$ , which is weak at best. The findings for the satisfaction–turnover relationship are stronger (meta-analysis correlation of  $-.25$ ), but the conclusion is that job satisfaction serves more of a mediating function. That is, the structural features of work (e.g., promotional opportunities) and employee characteristics (e.g., education) directly affect job satisfaction (and commitment), which in turn affects turnover.

The final issue here is whether job satisfaction is a unitary concept or is a complex of many facets or dimensions. Since a fairly large number of work features are known to affect job satisfaction, it is logical to expect that individuals can be satisfied with some of these but not others. The data support this logic. In particular, there is evidence that for almost any distinct feature of the work situation—pay, autonomy, variety, work group cohesion, feedback—satisfaction scales can be developed that divide into distinct (but related) factors along these dimensions. This poses not only a theoretical problem but also a scale construction problem. As a simple example, a person may be satisfied with the pay but not satisfied with feed-

back about job performance. Combining scores for these two factors will show the person to be neither satisfied nor dissatisfied for the composite scale. In such situations, the rule of thumb is that scales developed to measure various satisfaction dimensions should not be combined. However, global job satisfaction scales—those which ask more generally about liking one's job—can be used to represent a person's general affective reaction to a job. Sociologists more often use these global scales and assume that work is experienced and responded to globally.

The facet approach clearly becomes more important in applied research. If an employer wishes to alter the work setting to increase job satisfaction, a global scale will be only somewhat helpful; a scale that captures satisfaction with pay, routinization, communication, and the like, will provide the information necessary to implement specific structural changes. Numerous established measures of job satisfaction, both global and facet-based, exist (see Cook et al. 1981; Price and Mueller 1986; Spector 1997).

#### **RESPONSE TO WORK: WORK COMMITMENT**

Although some concepts, such as Dubin's (1956) central life interest and Lodahl and Kejner's (1965) job involvement, go back more than three decades, most of the interest in work commitment has emerged fairly recently, to a large extent during a time when interest in job satisfaction has been diminishing. If employee commitment is defined as the level of attachment to some component or aspect of work, the door is opened to a large number of types of commitment. The most common strategy adopted for understanding various types of commitment is to differentiate between the components and the foci of commitment.

There are numerous potential foci of commitment, with those receiving the most attention being commitment to work, the career, the organization, the job, and the union. It is organizational commitment, however, that has received the most theoretical and empirical attention (Mueller et al. 1992). Considerable interest exists in how workers form and manage their commitments to multiple foci (Hunt and Morgan 1994; Lawler 1992; Wallace 1995). For example, if a worker is strongly committed to his or her career, will this translate

into a similarly strong commitment to his or her employer (organization)? Although some suggest that commitment is a zero-sum phenomenon by which commitment to an employer must decline if commitment to one's career increases, research consistently shows that most commitments to multiple foci are positively related.

Three components of commitment have received the most attention (Meyer and Allen 1997): affective commitment, continuance commitment, and normative commitment. *Affective* commitment refers to a worker's emotional attachment to an organization. Organizational and industrial psychologists are given credit for initiating interest in this concept. They argue that commitment intervenes between various features of work and individual characteristics and the outcomes of absenteeism and voluntary turnover. Sociologists (e.g., Lincoln and Kalleberg 1990) tend to see the structural conditions of work as the ultimate causes of affective commitment. The evidence generally is consistent with the claims from both disciplines (Hom and Griffeth 1995; Mueller and Price 1990). *Continuance* commitment treats a person's degree of attachment as a function of the costs associated with leaving an organization. In practice, it has been operationalized as the employee's stated intention to stay (or leave). This form of organizational commitment can be traced back to Becker's (1960) side-bet theory. Individuals are portrayed as making investments (e.g., seniority, a pension fund, coworkers as friends) when they are employed in a particular organization. These side bets accumulate with tenure and thus become costs associated with taking employment elsewhere. An employee will discontinue employment only when the rewards associated with another job outweigh the accumulated side bets associated with the current one. Although the evidence for the reasoning behind this theory has not been supported, research has consistently shown a relatively strong negative relationship (meta-analysis correlation of  $-.50$ ) between intent to stay and voluntary turnover. Much of the literature identifies intentions to stay or leave as intervening between affective commitment and turnover. *Normative* commitment refers to the felt obligation to stay with an employer. Remaining attached to an organization is what one should do even if one is not emotionally attached or has only a limited investment.

Without question, affective organizational commitment has dominated the scholarly interest of those who study organizational commitment. It is strongly positively related to job satisfaction and negatively related to absenteeism and turnover. These relationships indicate the importance of studying and understanding employee commitment not only to address the practical issues confronting human resource managers but also to address classical sociological concerns about the "glue" that holds social groups together.

### SOCIAL CONTROL IN THE WORKPLACE

This article began with a description of the workplace as a contested terrain, a social setting in which employer and employee struggle for control. The image that comes from most economists is that monetary rewards are what motivate both employers and employees: Employers want to maximize profits, and workers want high pay for their work. The implication of this for workers is that they will be satisfied and committed if their pay is high, and if it is not, they can quit to take another job. This argument and causal linkage have been challenged both empirically and theoretically in sociology. There are three issues here. First, as was alluded to above, pay is only one of many factors that affect satisfaction and commitment. Second, employers, not workers, historically have had the upper hand in controlling the workplace and establishing the employment relationship. Third, job satisfaction and commitment can and are manipulated by employers to increase productivity and retain employees. There have been several different historical accounts of how this employer control occurs (e.g., Clawson 1980; Edwards 1979; Jacoby 1985; Vallas 1993), but two basic models dominate the literature. They can be differentiated by whether the social control is direct or indirect and by the importance given to worker satisfaction and commitment in the control process.

The historically dominant model of the workplace portrays direct control of workers by the employer. Direct supervisory monitoring, "machine control," and strictly defined divisions of labor are used to control the behavior of employees. In such instances, job satisfaction and organizational commitment may emerge to increase performance, but they are viewed as secondary to the direct control that is essential to maximizing work-

ers' productivity. The other model relies much less on direct supervision and control by the production process and instead argues that high-performance employees are controlled *indirectly* by manipulating work structures that in turn produce satisfied and committed workers. It is the satisfied and committed workers, then, who will be the most productive. Lincoln and Kalleberg (1990) argue for this model (called the "corporatist" model) in their study of U.S. and Japanese workers. Concretely, they find that organizational structures that facilitate participation, integration, individual mobility, and legitimacy result in more satisfied and committed employees. This sociological interest in workplace control has practical implications. The same dichotomy is recognized in human resource management (HRM), where the direct strategy is called the control strategy and the indirect strategy is called the commitment strategy (Arther 1994). Similarly, in education, concern with low achievement scores among U.S. students has resulted in a debate over the organizational design of schools (Rowan 1990). The more direct approach, also called the control strategy, is based on an elaborate system of bureaucratic controls for regulating classroom teaching and standardizing student learning opportunities and outcomes. The more indirect approach, also called the commitment strategy, rejects bureaucratic controls and standards and argues instead for innovative working arrangements that support teachers' decision making and increase their involvement in the tasks of teaching. The claim for the second approach is that satisfied and committed teachers are critical to improving student performance. Without question, then, worker satisfaction and commitment still constitute a major component in the critical debates about social control in the workplace, worker productivity, and societal outcomes such as student achievement.

### GENDER DIFFERENCES

Associated with the increase in sociological interest in gender inequalities over the last three decades has been an increased concern with whether the work orientations of women and men are different. Two questions have received considerable attention. One concerns whether women and men have different work values, and the other refers to what is called the gender job satisfaction paradox.

Research consistently has shown that women are just as satisfied (and often more satisfied) with their jobs as their male counterparts are. This is viewed as a paradox because women's jobs are on the average "worse" jobs with lower pay, less autonomy, and fewer advancement opportunities. Several arguments have been offered to account for this paradox (Phelan 1994; Mueller and Wallace 1996). Justice-related arguments center on (1) women accepting their lower rewards because of their lower inputs, (2) women being socialized to accept the idea that lower rewards are all they are entitled to, and (3) women being satisfied because they are comparing their rewards to those of other women, who also receive less. The consensus seems to be that the "other women as referent" explanation best explains the paradox. The major competing explanation is that women and men value different aspects of work. This leads directly to the question of gender differences in work values.

Probably the most popular explanation for the gender satisfaction paradox is that men value extrinsic rewards (e.g., pay, benefits, and authority) more than women do, while women value intrinsic rewards (e.g., social support) more than men do. As a consequence, women are not less satisfied when they receive less pay and are promoted less often than are men. Research findings strongly reject this argument, however. Women and men hold essentially the same workplace values (Hodson 1989; Phelan 1994; Mueller and Wallace 1996; Rowe and Snizek 1995; Ross and Mirowski 1995).

These similar workplace values do not mean, however, that men have the same degree of work-family conflict as do women. Research shows that this conflict is greater for women (Glass and Estes 1997). This finding only adds to the paradox: If women have worse jobs and experience more work-family conflict, why are they so satisfied with their jobs?

### THE FUTURE

The last two decades in the United States have witnessed considerable change in the workplace. Organizations have downsized, hired more temporary (contingent) workers, and outsourced production tasks to become more flexible in competing in an increasingly global marketplace. In addition, the income gap between the top and

bottom segments of society has grown, labor union membership has declined to an all-time low, and although unemployment continues to be low, job expansion has occurred mainly in the service sector, where many jobs do not have advancement potential. All this suggests that in the future workers can expect to move from employer to employer more often. Also, workers can expect to find that their employers are less concerned with whether employees are satisfied and less interested in gaining a long-term commitment from them. As a consequence, occupational or career commitment may become a more important motivating factor for workers than is organizational commitment or job satisfaction. Without doubt, this changing landscape for the employment relationship will keep sociologists interested in studying and understanding work values, job satisfaction, and commitment.

## REFERENCES

- Arther, J. B. 1994 "Effects of Human Resource Systems on Manufacturing Performance and Turnover." *Academy of Management Journal* 37:670-687.
- Becker, Howard 1960 "Notes on the Concept of Commitment." *American Sociological Review* 66:32-40.
- Blau, Peter M. 1964 *Exchange and Power in Social Life*. New York: Wiley.
- Clawson, Dan 1980 *Bureaucracy and the Labor Process*. New York: Monthly Review.
- Cook, John, Susan Hepworth, Toby Wall and Peter Warr 1981 *The Experience of Work*. New York: Academic Press.
- Dubin, Robert 1956 "Industrial Workers' Worlds: A Study of the Central Life Interests of Industrial Workers." *Social Problems* 3:131-142.
- Edwards, Richard 1979 *Contested Terrain*. New York: Basic Books.
- Glass, Jennifer L., and Sarah Beth Estes 1997 "The Family Responsive Workplace." *Annual Review of Sociology* 23:289-313.
- Hall, Richard 1986 *Dimensions of Work*. Beverly Hills, Calif.: Sage.
- Herzberg, Frederick 1966 *Work and the Nature of Man*. Cleveland: World.
- Hodson, Randy, and Teresa Sullivan 1995 *The Social Organization of Work*. Belmont, Calif.: Wadsworth.
- 1989 "Gender Differences in Job Satisfaction: Why Aren't Women More Dissatisfied?" *Sociological Quarterly* 30:385-399.
- Hom, Peter W., and Rodger W. Griffeth 1995 *Employee Turnover*. Cincinnati: South-Western College Publishing.
- Homans, George C. 1958 "Human Behavior as Exchange." *American Journal of Sociology* 63:597-606.
- Hunt, Shelby D., and Robert M. Morgan 1994 "Organizational Commitment: One of Many Commitments or Key Mediating Construct?" *Academy of Management Journal* 37:1568-1587.
- Jacoby, Sanford 1985 *Employing Bureaucracy*. New York: Columbia University Press.
- Kalleberg, Arne, and Ivar Berg 1987 *Work and Industry*. New York: Plenum.
- Kohn, Melvin, and Carmi Schooler 1982 "Job Conditions and Personality: A Longitudinal Assessment of Their Reciprocal Effects." *American Journal of Sociology* 87:1257-1286.
- Lawler, Edward III. 1973. *Motivation in Work Organizations*. Monterey, Calif.: Brooks/Cole.
- Lawler, Edward J. 1992 "Affective Attachments to Nested Groups: A Choice-Process Theory." *American Sociological Review* 57:327-339.
- Lincoln, James, and Arne Kalleberg 1990 *Culture, Control and Commitment: A Study of Work Organization and Work Attitudes in the United States and Japan*. Cambridge, UK: Cambridge University Press.
- Lodahl, Thomas, and Mathilde Kejner 1965 "The Definition and Measurement of Job Involvement." *Journal of Applied Psychology* 49:24-33.
- Maslow, Abraham 1954 *Motivation and Personality*. New York: Van Nostrand Rheinhold.
- McClelland, David 1961 *The Achieving Society*. New York: Van Nostrand.
- McGregor, Douglas 1960 *The Human Side of Enterprise*. New York: McGraw-Hill.
- Meyer, John P., and Natalie J. Allen 1997 *Commitment in the Workplace: Theory, Research and Application*. Thousand Oaks, Calif.: Sage.
- Mueller, Charles W., and James Price 1990 "Economic, Psychological, and Sociological Determinants of Voluntary Turnover." *Journal of Behavioral Economics* 26:2181-2199.
- , and Jean E. Wallace 1996 "Justice and the Paradox of the Contented Female Worker." *Social Psychology Quarterly* 59:338-349.
- , ———, and James L. Price 1992 "Employee Commitment: Resolving Some Issues." *Work and Occupations* 19:211-236.
- Phelan, Jo 1994 "The Paradox of the Contented Female Worker: An Assessment of Alternative Explanations." *Social Psychological Quarterly* 57:95-107.

- Price, James, and Charles W. Mueller 1986 *Handbook of Organizational Measurement*. Cambridge, Mass.: Ballinger.
- Ross, Catherine, and John Mirowski 1996 "Economic and Interpersonal Work Rewards: Subjective Utilities of Men's and Women's Compensation." *Social Forces* 75:223-246.
- Rowan, B. 1990. "Commitment and Control: Alternative Strategies for the Organizational Design of Schools." In C. Cazden, ed., *Review of Research in Education*. Washington, D.C.: American Educational Research Association.
- Rowe, Reba, and William Snizek 1996 "Gender Differences in Work Values: Perpetuating the Myth." *Work and Occupations* 22:215-229.
- Smith, Patricia, Lorne Kendall, and Charles Hullin 1969 *The Measurement of Satisfaction in Work and Retirement*. Chicago: Rand McNally.
- Smither, Robert 1988 *The Psychology of Work and Human Performance*. New York: Harper and Row.
- Spector, Paul E. 1997 *Job Satisfaction: Application, Assessment, Causes, and Consequences*. Thousand Oaks, Calif.: Sage.
- Vallas, Steven P. 1993 *Power in the Workplace*. Albany: State University of New York Press.
- Vroom, Victor 1964 *Work and Motivation*. New York: Wiley.
- Wallace, Jean E. 1995 "Professionals in Bureaucracies: A Case of Proletarianization or Adaptation?" *Administrative Science Quarterly* 40:228-255.
- Watson, David, and Lee Clark 1984 "Negative Affectivity: The Disposition to Experience Aversive Emotional States." *Psychological Bulletin* 96:465-490.
- Wharton, Amy 1998 *Working in America*. Mountain View, Calif.: Mayfield.

CHARLES W. MUELLER

## WORLD RELIGIONS

Religious life throughout the world, regardless of the specific tradition, exhibits both personal-psychological and communal-social aspects. Of course, persons within the diverse religious traditions of the world perceive the spiritual dimension of their faith as transcending both the individual psychological and emotional as well as the corporate and social aspects of their faith's expressions. Nonetheless, two major academic strands of religious studies over the last century have focused primarily on either the psychological (e.g., James

1961; Freud 1928; Jung 1938) or the social (e.g., Weber 1963; Durkheim 1965; Wach 1958) dimensions of religion. An Oglala Lakota's ("Sioux" in Algonquian) vision reveals these two interactive aspects of religion.

The Plains Indians in America were noted for their vision quests, and periods of fasting and life-cycle rituals often were associated with those quests. However, the vision of Black Elk, a Lakota shaman, occurred spontaneously when he was 9 years old and was stricken by fever and other physical maladies (Neidardt 1972, pp. 17-39). His vision began with two men dressed in traditional garb but shaped like slanting arrows coming from the sky to get him. As a little cloud descended around him, the young Black Elk rose into the sky and disappeared into a large cloud bank. He saw an expansive white plain across which he was led by a beautiful bay horse. As he looked in the four directions, he saw twelve black horses in the West, twelve white horses in the North, twelve sorrel horses in the East, and twelve buckskin horses in the South. After the arrival of Black Elk, the horses formed into lines and formations to lead him to the "Grandfathers." As this heavenly equine parade proceeded, horses appeared everywhere, dancing and frolicking and changing into all types of animals, such as buffalo, deer, and wild birds. Ahead lay a large teepee.

As Black Elk entered the rainbow door of the teepee, he saw six old men sitting in a row. As he stood before the seated figures, he was struck by the fact that the old men reminded him of the ancient hills and stars. The oldest spoke, saying, "Your grandfathers all over the world are having a council, and they have called you here to teach you." Black Elk later remarked of the speaker, "His voice was very kind but I shook all over with fear now, for I knew that these were not old men but the Powers of the World and the first was the Power of the West; the second, of the North; the third, of the East; the fourth, of the South; the fifth, of the Sky; the sixth, of the Earth."

The spokesman of the elders gave Black Elk six sacred objects. First, he received a wooden cup full of water, symbolizing the water of the sky that has the power to make things green and alive. Second, he was given a bow that had within it the power to destroy. Third, he was given a sacred name, "Eagle Wing Stretches," which he was to embody in his role as shaman (healer and diviner)

for his tribe. Fourth, he was given an herb of power that would allow him to cleanse and heal those who were sick in body or spirit. Fifth, he was given the sacred pipe, which had as its purposes a strengthening of the collective might of the Lakota tribe and a healing of the divisions among the Lakota, to allow them to live in peace and harmony. Finally, Black Elk received a bright red stick that was the “center of the nation’s circle” or hoop. This stick symbolized a sacred focusing of the Lakota nation and linked the Lakota to their ancestors as well as to those who would follow them.

Black Elk’s vision ended with a flight into a foreboding future in which the Lakota would encounter white-skinned “bluecoats” who would threaten the sacred hoop of the Lakota nation. Many years later, as Black Elk reflected on his vision, he realized that even in the devastating upheaval caused by the wars between his nation and the “bluecoats,” his people had been given the sacred objects and rituals that would allow them to rise above mundane exigencies and to heal the nation and restore the hoop in times of trouble.

The vision of Black Elk makes it clear that what sometimes appear to be perfunctory religious rituals, fantastic myths, or arcane ethical injunctions often have their roots in a deep sense of the contact between human beings and that which they have experienced as a divine power. This article emphasizes the social aspects of world religions, but it is important to keep in mind that the religious experiences codified in the social institutions of the world’s religions are not fully captured by psychological or sociological explanations alone. There has been a tendency in the academic study of religion to interpret religious experiences and behavior by reducing them to psychological or social causes or antecedents. For example, Sigmund Freud (1928) reduces religious experiences to unconscious projections of human needs that he likens to infantile fantasies that rational humans should grow beyond. A contemporary of Freud, Emile Durkheim (1965, p. 466), has a tendency to reduce religions to their social functions: “If religion has given birth to all that is essential in society, it is because the idea of society is the soul of religion.”

While the pioneering work of Max Weber and Durkheim laid the groundwork for much of contemporary social analysis of religion, comparative

sociologists of religion such as Joachim Wach (1958) have tempered earlier tendencies toward sociological reductionism. Wach sought to understand the nature of religion by examining traditions throughout the world and noting the primary elements they shared. He identified religious experience as the basic and formative element in the rise of religious traditions around the world and then investigated the expression of this experience in thought, action, and community.

Wach said that there is a symbiotic relationship between religion and society. On the one hand, religion influences the form and character of social organizations or relations in the family, clan, or nation as well as develops new social institutions such as the Christian church, the Buddhist *sangha*, and the “Lakota nation.” On the other hand, social factors shape religious experience, expression, and institutions. For example, in Black Elk’s vision, the role of the warrior in Lakota society is expressed through the two men who come to escort Black Elk into the sky, and in his later mystical venture into the future, Black Elk as Lakota shaman (*wichash wakan* is one who converses with and transmits the Lakota’s ultimate spiritual power, or *Wakan*) becomes the ultimate warrior who battles a “blue man” (perhaps representing personified evil or the dreaded “bluecoats”). Lakota social conventions that name the natural directions as four (North, South, East, West) are modified by Black Elk’s vision to include Sky and Earth, making six vision directions that influence the number of elders Black Elk encounters in the heavenly teepee and the number of sacred objects he is given. Here the shaman’s vision modifies social conventions even as it creates a social subconvention for other visionaries who also name the directions as six. The objects are conventional implements of Black Elk’s culture that are empowered to serve symbolically as multivocal conveyors of sacred knowledge and wisdom. Finally, Black Elk’s vision can be viewed sociologically as confirming the corporate sacredness (the sacred hoop) of the nation of the Lakota. For example, a Lakota’s vision was powerful and meaningful only to the extent that the tribe accepted it. In this sense one can understand why Durkheim would say that religion, in this case the Lakota’s, is society writ large in the sky.

However, for Wach and for scholars, such as Niman Smart (1969), who follow his lead, the



forms and expressions of religious life are best understood as emanating from religious experience. Smart identifies six dimensions that all religions share: (1) ritual, (2) mythological, (3) doctrinal, (4) ethical, (5) social, and (6) experiential. The author of this article has provided an interpretative framework for understanding the necessary interdependence of these six elements of religious traditions in *Two Sacred Worlds: Experience and Structure in the World's Religions* (Shinn 1977). These dimensions of the religious life form the structure of this analysis of the social aspect of world religions.

### RELIGIOUS EXPERIENCE

Building on the insights of William James and Rudolph Otto (1946), more recent scholars such as Wach, Smart, and Mircea Eliade (1959) seek the origin of religion in the religious experience of a founder or religious community. These scholars assert that genuinely religious experiences include an awareness or an immediate experience of an ultimate reality or sacred power, whether a theistic divinity as in the case of the God(s) of Judaism, Christianity, and Islam or a nontheistic transcendental reality as in the case of the Buddhists' Nirvana or the Hindus' Brahman/Atman. James suggests that transcendental or mystical experiences are immediate apprehensions of the divine that are marked by ineffability, a noetic quality, transiency, and passivity. From one perspective, ineffability can be understood as the inability of language to relay the emotional and cognitive content of a peak religious experience; it also may be described as a failure of language to capture the divine subject of such an experience, that is, the ultimate reality itself. Nonetheless, religious experiences inevitably are understood as providing new states of knowledge that cannot be grasped fully by the discursive intellect. This noetic dimension of religious experience often is described as the revelation of new knowledge (i.e., illumination) that is provided by religious experiences. In fact, it is precisely an awareness of an encounter with a sacred reality in religious experiences that differentiates these experiences from nonreligious peak experiences (e.g., an aesthetic peak experience of a piece of music). Religious experiences also tend to be marked by brevity (i.e., transiency) and the passivity of the person having the experience. While aesthetic, political, and erotic peak experiences may be characterized by ineffability,

transiency, and passivity, only religious experiences bring with them a consciousness of an encounter with a "holy other" sacred reality.

Whether a founding religious experience is immediate and direct, such as the Buddha's nontheistic enlightenment experience of Nirvana, or cumulative and indirect, as was the lengthy exodus journey of the Hebrews, religious experiences are, in Wach's terms, "the most powerful, comprehensive, shattering, and profound experience" of which human beings are capable (1958, p. 35). Wach concludes that a necessary criterion of genuine religious experience "is that it issues in action. It involves imperative; it is the most powerful source of motivation and action" (1958, p. 36). Consequently, religious experiences may be viewed as the wellspring of religion both in the formation of a new religious tradition and in the origin of the faith of the later generations.

Even if one accepts the primacy of religious experience, it is important to note that founding religious experiences are deeply immersed in the social and cultural realities of their time and place. For example, whether immediate and direct or cumulative and indirect, religious experiences inevitably are expressed in the language and concepts of the persons and culture in which they arose. Black Elk's vision of *Wakan* in the form of the six Grandfathers clearly reflects the Lakotas' social and political structure as well as their idealized notions of nation and nature. The Thunder Beings and Grandfathers who are the personifications of *Wakan Tanka* ("Great Power") obviously arise from the natural, linguistic, and social environments of the Lakota. So does the conception of *Wakan* itself as a pervasive power that permeates animal and human life as well as that of nature. A contemporary Lakota has said, "All life is *Wakan*." So also is everything which exhibits power whether in action, as in the winds and drifting clouds, or in passive endurance, as the boulder by the wayside.

Religious experiences occur to persons who have already been socialized. The most obvious social tool is the language used to express even the most profound religious experiences. The ineffable nature of religious experiences requires the use of metaphors or extensions of everyday language, as in the case of Black Elk, and to some extent, the experience itself is shaped by the language in which it is expressed.

Divine names usually are borrowed from the social and linguistic environment of the founder or founding community. For example, the exodus experience of the Hebrew people was interpreted by them as a liberating religious experience fostered by the God of Abraham, Isaac, and Jacob. This God, whose name is given in the Book of Exodus as Yahweh (“I am who I am”), is also called El Elyon (“God most high”), El Shaddai (“God of the mountain”), and Elohim (usually translated as “God”). Moses probably borrowed the name “Yahweh” from the Midianites. El Elyon was the high god of Salem (later called Jerusalem) and was worshiped by King Melchizedek. It also is known that the Canaanite high god of the same period was named El and appears in different cultic sites throughout the ancient Near East. Although it is clear that the Hebraic religious texts understand Yahweh and El quite differently than do their known local counterparts, the Hebrew high god embraced the local deity nomenclatures while modifying their meanings.

In a similar fashion, the divinity of the man Jesus is acknowledged in early Christian texts through references to earlier Jewish apocalyptic language and expectations. In the Jewish apocalyptic literature (e.g., I Enoch), the “Son of Man” appears as a righteous judge who will come to earth to signal the beginning of the heavenly kingdom and God’s rule. As an eternal savior, the Son of Man will come to save the righteous followers of God and destroy all those who ignore him. In those linguistic borrowings, however, significant modifications of the original conceptions are made to adjust the titles and expectations to the man Jesus as perceived by his followers. For example, Jesus comes as the Son of Man not primarily as a stern and vengeful judge but as a savior who is himself the sacrifice. This linguistic and conceptual transformation reflects the dependence of language on experience as much as it reveals the social and linguistic dimensions of religious experience.

Similar examples of borrowed—and transformed—god names abound in religious literature and history throughout the world. In Saudi Arabia in the sixth century, Mohammed elevated a local polytheistic Meccan god, Allah, to the status of an international deity. In tenth-century Indian Puranic literature, devotees of the god Vishnu promote his *avatar*, called Krishna, to a supreme

theistic position as the god above all gods. Although the *Bhagavata Purana* recounts the *lilas*, or play, of Krishna as though the author were describing historic figure, it is clear to textual scholars that there are two essentially distinct and dynamic story traditions arise from the Brahminical Krishna of *Bhagavad Gita* fame and from the indigenous cowherd Gopala Krishna associated with the western Indian Abhira tribes.

Although devotees of either Allah or Krishna now perceive their divinity and his name as having been “from the beginning,” there is little doubt that the local social and linguistic environments provided both content and context for the names of the divinities in these two traditions. Perhaps the most radical example of theistic amalgamation is that of the Indian goddess Kali. Described in medieval Indian texts as being synonymous with literally dozens of local and regional goddess names and traditions, Kali is a latecomer to the Indian theistic scene as one who is given the primary attributes of many gods and goddesses. The mythological tale of the birth of Kali reveals an amalgamation process that gave birth to this great goddess now worshiped by millions in India as the “Supreme Mother.” Finally, the concept and expressions of the nontheistic Nirvana experienced by the Buddha were fundamentally shaped by the notion of reincarnation or rebirth and other metaphysical assumptions common to most religious traditions in India in the fifth century B.C.

These examples show that while religious experience of the sacred may be the initiating point of the world’s religious traditions or an individual’s faith, that experience is given shape and substance by the linguistic and social context out of which it arises. It is also true, however, that life-altering religious experiences such as those described above shape the language and traditions through which they are expressed. This symbiotic relationship occurs in the other dimensions of religious life that are shared by the world’s religions.

## MYTH AND RITUAL

Formative religious experiences contain within them impulses to expression (myth) and re-creation (ritual) that later become routinized and then institutionalized. Core myths and rituals, therefore, attempt to convey and re-create the experi-

ence of the founder or religious community. Both myths and rituals rely on symbols whose content must be shared in order for them to have meaning for the religious group that uses them. Symbols have not only shared cognitive meanings but also common emotional significance and value. That is, symbols do not simply convey intellectual understanding but also engender an emotive response. Furthermore, religious symbols are integrative and transforming agents in that they point to realities that have been encountered but are hidden from everyday vision and experience. Paul Ricoeur (1972) says that symbols yield their meaning in enigma, not through literal or direct translation. Symbols, therefore, suggest rather than explicate; they provide “opaque glimpses” of reality rather than definitive pictures. Understood in this fashion, the journey from symbol to myth is a short one for Ricoeur, who takes myth to be a narrative form of the symbol. Put simply, myths are narratives or stories of the sacred and of human encounters with it.

As stories of sacred powers or beings, myths fall into two basic categories: expressive and reflective. Expressive myths are sacred narratives that attempt to relate the founding or codifying religious experiences of a religious tradition, while reflective mythic narratives are composed subsequently to integrate the sacred experience into everyday life. For example, Black Elk’s “re-telling” of his vision experience becomes an expressive myth or sacred narrative for the Oglala Lakota to which they refer again and again in reflective stories of the Thunderbeings or the Grandfathers, wherein the Lakota attempt to extend the lessons of this experience to later problems they encounter. Nearly every extant religious tradition tells and retells its sacred narrative of the founder’s or founders’ encounter with the sacred reality. Black Elk’s vision becomes such a story for the Oglala Lakota.

The story of the exodus of the Hebrews is recounted as a symbolic and founding narrative of God’s liberation for Jewish people of all times. The stories of the life, death, and resurrection of Jesus form the core myths of Christians when one understands a myth to mean “sacred narrative” rather than “untrue story.” Likewise, the story of the Buddha’s arduous meditative journey culminating in the attainment of Nirvana inspires religious thought and behavior throughout all Buddhist

lands even today. Similarly, Muhammad’s auditory experience of Allah on Mount Hira, which resulted in his recording of the Qur’an, constitutes the sacred history of millions of Muslims on all the continents. Finally, even though scholars are confident in their judgment that the life of Krishna as told in the tenth-century *Bhagavata Purana* is really an anthology of stories borrowed from earlier Krishna traditions, these *lilas*, or “playful episodes,” told as a single life of Krishna have inspired religious experiences, poetry, and rituals that still enliven the lives of millions of Hindus throughout the world. From even this selective set of examples of founding myths, it is clear how deeply they drink from the social, linguistic, and institutional wellsprings of their time and place.

The generative function of core myths is shared by certain rituals that attempt to “represent” in a spatial and physical context the core experience of a religious tradition. From one perspective, core rituals are those that emerge from sacred narratives or myths as their active component. From a second perspective, core rituals represent repetitive, institutionalized behavior and clearly are immersed in the social sphere of religious life. For example, the Christian narrative that relates the Last Supper of Jesus as a sacramental event (e.g., Mark 14:12–26) is physically presented in the early Christian love feast that becomes the Lord’s Supper (Eucharistic ritual or Mass) of later Christian churches.

The work of Victor Turner (1969) in a traditional African religious context provides a vocabulary for the religious and social transactions that take place in core myths and rituals. Turner describes three phases in ritual reenactments that attempt to (1) separate or detach the participant from everyday consciousness and social position, (2) provide a moment of *liminality* and *communitas* of shared experience with participants in rituals, and then (3) reintegrate ritual participants back into everyday life with its social roles and structure. *Liminality* is the neutral psychological and social state of transition between one’s former social roles and consciousness and the new status one assumes beyond the ritual. *Communitas* for Turner is a mode of social relationship that is marked by an egalitarianism that is uncommon in the stratified roles and relationships of the everyday world. Consequently, Turner would argue

that religious rituals may provide an in-between, or liminal, moment of social and psychological experience that religious devotees often assert includes an encounter with their sacred power or reality.

The Passover narrative in the Book of Exodus provides a good example of a core myth that is later enacted, in this case in a Passover meal. In its literal meaning, the Passover myth refers to the tenth plague, when the angel of death killed Egyptian firstborn children while sparing the Hebrew children just before the exodus journey. In its symbolic sense, the Passover story that is “re-presented” in the Passover sacrificial meal symbolizes Yahweh’s power of liberation. To the extent that the story of the exodus reveals the beginning of Yahweh’s covenantal relationship to the Hebrew people, the Passover ritual attempts to re-create or revivify that relationship.

Beyond the community’s social embodiment of the sacred story of Israel’s encounter with Yahweh in a festive and communal sacrificial ritual of the Passover, the social aspects of both the myth and the ritual are evident. Sacrifices were the common mode of worship for the pre-Mosaic tribal religions as well as for the contemporary cults in Moses’ day. It is very likely that the Passover ritual described in Exodus 12 derives from a combination of a nomadic animal sacrifice and an agricultural feast of unleavened cakes, both of which predate the exodus event. While the Hebrews’ experience of Yahweh in the exodus journey reshapes both the story and the ritual as a liberation event, both the Hebrew myth and the ritual have antecedents in the social and religious world of which they were a part.

Similarly, the baptism and Eucharist rituals in the Christian faith are core rituals that stem from the religious narratives that gave birth to them. Likewise, traditional nontheistic Theravada forms of Buddhist meditation appear to stem directly from the Buddha’s spiritual struggle and release but draw on Jain and Hindu forms that predate them. Among the Oglala Lakota, the horse dance ritual was taught by Black Elk to his tribe in a fashion that replicated as closely as possible the vision he received. Therefore, the six old Grandfathers, the horses representing the four cardinal directions, and the various sacred implements he

was given all become central elements of the horse dance ritual.

In Islam, the Hajj is one of the five pillars of faith that is incumbent on all Muslims to honor and embody. The Hajj is a pilgrimage that reenacts the spiritual journey of Muhammad with periods of fasting, prayer, and meditation that culminate with ritual circumambulations of the Ka’ba, the black stone in the central mosque of Mecca that is the seat of Allah’s throne. In the Hindu devotional traditions, it is common for dramatic performances, stylized ritual dance forms such as Bharata Natyam, and temple dramatic readings to convey episodes of the encounter of devotees with the divine. Consequently, theatrical dramatic productions of the *lilas*, or playful pastimes, of the cowherd god Krishna are enjoyed by villagers throughout India not simply as theatrical events but as representations of Krishna’s delightful divine play. The daily ritual reenactment that occurs before the shrines of Krishna, Kali, and other Indian divinities is called *puja* and is a ritual ceremony that probably emanates from the stylized honorific behavior one accords to a royal guest. Here the social precursors to religious ritual are evident, even though they are transformed by the religious narrative and ritual context into which they are placed.

Scholars across a variety of disciplines and perspectives have asserted the interconnection of myth, ritual, and the religious community. Perhaps the most clear summary of this relationship is given by Bronislaw Malinowski, who says, “An intimate connection exists between the word, the mythos, the sacred tales of a tribe, on the one hand and their ritual acts, their moral deeds, their social organization, and even their practical activities on the other” (1954, p. 96). Malinowski indicates that while core myths and rituals may have their origin in founding religious experiences, they also serve as social “warrants” for the primary beliefs of the society out of which they arise and which they help shape. From this perspective, myths and rituals serve primarily as vehicles that legitimate social institutions. Core myths and rituals appear to be charged with the difficult task of representing and re-creating founding religious experiences. They also reflect and embrace their social and cultural contexts. Furthermore, not all myths and rituals serve this primary and essentialist function; certain myths, rituals, and religious behaviors diverge

considerably from the impetus the core narrative seems to suggest.

### ETHICS

Malinowski and Wach make clear that ethics arise partly as a result of religious experience but also participate fully in social processes. While religious experiences may give rise to immediate expression (core myths) and reenactments (core rituals), they also give impetus to new attitudes and intentions, which are reflected in norms for behavior. In the Christian context, such behavior is claimed to be the mark of a “reborn” person whose conduct manifests the tangible effects of an experience of God. Conversely, the ethical norms and traditions that arise within a religious institution may reflect as much the mores of the surrounding culture and society as they do the experience upon which the institution was founded. Social factors such as language, family roles, and social customs play a role in the process of the externalization of the religious life in ethical laws. James says simply that behavior is the empirical criterion for determining the quality and validity of a religious experience. The distinction he makes between the person who has a religious experience and the person who undergoes a religious conversion is the distinction between having a highly charged peak experience and living a new life born of that experience.

It appears that all religious traditions evidence an interdependent and necessary relationship of conduct to experience so that what is experienced as an ecstatic encounter with the divine is expressed as a new and integrated mode of living. The committed ethical life of a devotee, then, is ideally understood as an active extension of religious experience expressed through communal or shared norms. While an immediate religious experience may provide a core religious impulse (e.g., to love God and one’s neighbor in the Christian context or to fear Allah in the Muslim context), that impetus becomes manifest in the concrete situations of social behavior. For example, the nontheistic enlightenment experience of the Buddha resulted in a sense of detachment from the world that was linked to enduring traditions of *metta* and *karuna* (love and compassion) and resulted in “detached compassion” as the complex ethical norm the Buddha modeled for his disciples.

The most obvious intrusion of social norms and processes into the religious life occurs in moral decision making. The natural and social worlds in which people live provide challenges and problems that require an ethical response. Consequently, life in the world poses many situations not anticipated in the religious texts and routinized ethical norms of religious traditions. As a result, over time, ethical systems often come to reflect the surrounding secular culture and social norms as much as they do the basic religious impulse from which they are supposed to derive their direction. This process is mediated during the life of the founder whose authority and behavior provides a model for action. In subsequent generations, however, individuals and institutions such as the Pope, the Buddhist *sangha* (community of elders), and the Lakota tribal council often determine the ethical norms of a community. When ethical statements and positions stray too far from their initial impulses, they are in danger of mirroring the society they intend to make sacred. Put simply, while ethical impulses may originate in religious experiences, the ethical laws, norms, and traditions that are constituted in scriptures and institutional pronouncements often distort the moral imperative by including rationalizations that conform to social, not religious, expectations.

An example of the difference between ethical impulse and moral law can be found in the Hebrew notion of a covenantal relationship with God. Moses and the exodus tribes experienced a compassionate, mighty, jealous, and demanding God. The laws of the early Hebrews, therefore, were viewed not only as commandments arising from a stern leader or group of legalistic lawmakers but also as expressions of an appreciative and liberating relationship with God. The Sinai story of the transmission of the Ten Commandments is intended to reveal the Hebrews’ ethical relationship with Yahweh. It was on that holy mountain that the covenant between Yahweh and his people was given concrete expression. However, this relationship was marked by infidelity on the part of Yahweh’s people. Therefore, for many of them, the codes of conduct contained in the Ten Commandments and the Levitical Code were experienced as the oppressive laws of a judgmental God.

Jesus summarized the essence of ethical behavior in a twofold commandment to love God and love one’s neighbor that was enjoined on all

who would count themselves as disciples of God. However, the teachings of Jesus and the commandment of love have led over the centuries to disputes about whether Christians should engage in war, permit abortions, treat homosexuals as equals, and allow divorces. Institutionalized Christian churches in their many forms have decreed what proper ethical conduct is with regard to such issues, and those norms vary and even contradict each other within and across Christian religious traditions. This is the difference between the imperative to love God and love one's "neighbor" and ethical laws that must express divine love in complex and rapidly changing social contexts and situations. Seemingly universal laws such as "Do not kill" mean something quite different to a Lakota warrior who may kill (and sometimes scalp) his enemy (but not a fellow tribesman) than they do to a Muslim who is encouraged to kill an infidel who defames Allah or to a Buddhist who is enjoined not to kill *any* living being.

Even among seemingly similar traditions, such as the Hindu devotional sects, ethical norms can vary immensely. In the Kali goddess tradition, animal sacrifice is still commonly practiced as a way of returning to the goddess the life-giving force she has bestowed on her creation. Some devotees of Kali have interpreted her mythological destruction of demons as a model for their own behaviors and have followed suit as thieves and murderers in the Indian Thuggi tradition. By contrast, Kali devotees such as Rahmakrishnan understand Kali to be a transcendent "ocean of bliss" who engenders peacefulness and nonviolence in her disciples.

What is true of all these religious traditions around the world is that persons usually are taught what constitutes proper or ethical behavior, and in that context, ethics are learned conceptions born of the social process and its experiences. Consequently, ethical norms and their expression often reflect the social environment in which religious traditions arise. A clear expression of this fact is found in the Hindu religious tradition's embrace of the caste system that sacralizes a socially elitist and patriarchal social system that predates Hinduism. Caste distinctions that are sacralized in the mythical and theological texts of the Hindu tradition serve as warrants for social roles and norms that undergird not only the Hindu traditions but also those of the Buddhists and Jains in India.

## THEOLOGY AND DOCTRINES

Just as religious experience may result in the formation of a religious movement that tells the founding story of contact with a sacred power (core myth), tries to re-create that experience for the beginning and subsequent communities (core rituals), and impels new believers to act in accordance with this vision or revelation (ethical impulse leading to institutionalized ethics), so it is that even very early in a religious tradition's history questions and criticisms arise that must be answered. Religious reflection takes a variety of forms that touch the total corporate life of a religious community. Sacred scriptures often encompass expressive myths that relate in narrative form the founder's or founders' contact with the sacred core rituals in outline or in full, ethical injunctions and moral codes, and reflective myths, doctrines, and explications that attempt to answer believers' questions and unbelievers' skepticism. Almost inevitably, members of a religious community are provoked from without and within to explain how their sacred reality is related to the origin of the community and perhaps even to the origin of the world. Consequently, reflective myths that represent second-level or posterior reflection are incorporated to explain those beginnings.

Three distinct but interrelated purposes and functions of reflective myths are to (1) explain origins, (2) rationalize aspects of core beliefs, and (3) provide an apologetic defense of the faith to disbelieving insiders or outsiders. A good example of reflective theologizing is the development of the biography of the Buddha. The oldest Pali texts essentially begin the life of the Buddha with his disillusionment with the world at age 29, when he was already a husband and a father. The early texts indicate that his name was Siddhartha and that his father, Suddhodana, ruled a small district in the north Indian republic of the Sakyas. This early story indicates that Siddhartha was married at the age of 16 or 17, had a son, and then became disillusioned with the human suffering he saw around him and renounced the world to seek spiritual liberation while leaving his family behind.

Approximately five hundred years after the death of the Buddha, two separate "biographies" were written that contained accumulated legends not only about the miraculous birth of the Buddha but also about the great renunciation. The birth

story describes the descent of the Buddha from the heavens as a white elephant who miraculously enters his mother's side and is born nine months later as a fully functioning adultlike child. These biographies describe the Buddha's physical features (captured in religious images and icons) as including the lengthened ears of an aristocrat, a smoothly shaped conical bump on the top of the head indicating his intelligence, and other marks that foretell his later enlightenment.

These latter-day scriptures recount his renunciation of the world in a full-blown, theologized story of encounters with an ill man, a decrepit old man, a dead man, and a religious ascetic. The story of the Buddha's four visions provides a fuller explication of the reasons for his renunciation. Both the birth story confirming the Buddha's sacred origins and the story of the four visions of the Buddha (a rationalization of his renunciation) represent reflective myths that fill in biographical gaps in earlier stories of his life in light of his later enlightened status.

Parallels to the biographical history of the Buddha can be found in the scriptural stories of the miraculous births of Jesus, Mahavira (founder of the Jains), Krishna, Kali, and Muhammad, among others. A similar genre of reflective myths can be found in the creation stories that often are added dozens of years or even centuries after the founding experience. Good examples of this process are the Hebrew creation stories told in Genesis 1 and Genesis 2. God's creation in seven days is the youngest creation story (the priestly story of the seventh century B.C. that is told in Genesis 1:1–2:4a) and is placed at the beginning of the book of Genesis. It is likely that the Akkadian myth of Tiamat served as a model for this story of the creation of the world out of a watery chaos.

The older Yahwist creation story, found in Genesis 2:4b ff., is set in a desert environment instead of a primeval ocean and very likely goes back to the tenth century B.C. A decidedly more anthropomorphic story, the Yahwist Garden of Eden story, was added at least three to four hundred years after the exodus experience. Neither the priestly story nor the Yahwist story received its present form until the sixth or seventh century B.C., when both were called upon to explicate the creative power of their Hebrew God set against the Canaanites' theology of nature's seasonal birth,

death, and rebirth that the Hebrews encountered in Palestine. For the Palestinian farmer, Canaanite or Israelite, the question was, "Is it Yahweh or Baal to whom one should offer sacrifices and give allegiance if one's crops are to prosper?" The two Genesis creation stories explain not only who is responsible for the origin of life on earth but also how one can explain human illness, suffering, and death in the context of the God who led the Hebrews out of Egypt. In Africa and India, the numerous and sometimes contradictory creation stories one finds in a single religious tradition reveal less about the illogical nature of some reflective myths than they do about the human need to have questions of birth, death, suffering, social relationships, and the founding of the tribe placed in the context of a tradition's ultimate reality.

When religious traditions develop full-fledged social institutions, it is common for sacred texts and other interpretative theological texts to explain the necessity of those religious organizations and their officials. Whether it is the early church fathers' explanations of the seat of Peter on which the Pope sits in the Roman Catholic tradition or a Lakota visionary myth that explains the role of the shaman in the community, reflective myths and theologies develop as intellectual and institutional rationalizations for the extension of the founding experiences and tradition into all aspects of life and society. Religious councils, theological traditions, sectarian disputes, and doctrinal formulas all arise as socialized institutions that attempt to explicate, defend, and provide an apology for a religious faith firmly embedded in the personal and social lives of its adherents. For example, Islamic theology extends the influence of the Qur'anic faith into the economic, political, and social lives of the Muslim people. Likewise, from birth and family relationships through wars and death, the Lakota's life was experienced within the sacred hoop.

The extension of religious faith into all aspects of life is justified in scriptures and doctrinal tracts by the reflective process of mythmaking and theologizing. Peter Berger (1969) calls such activity the construction of a *nomos*. A theological *nomos* is essentially a socially constructed worldview that attempts to order all of human experience in the context of a sacred reality, whether theistic (e.g., Krishna or Allah) or nontheistic (e.g., Nirvana). Such theological reflection is determined to a

great extent by the social and human circumstances that give rise to the questions that must be answered as well as the language and social conventions through which the reflections are expressed. However, Wach reminds us that the prophetic function of religious traditions often shapes the social environment to a religious vision and not simply vice versa. Puritan society in colonial America is an example of religious faith shaping social mores and institutions.

### INSTITUTIONS

Religious institutions arise as the fullest and most obvious social expression of a religious faith. They are equally the home for the core myths and rituals to be enacted and the loci of the religious communities whose individual and collective needs must be met. Religious institutions vary from formal collectivities such as the Christian church, the Muslim mosque, the Hindu temple, and the Buddhist *sangha* to their extended representations in festivals and ceremonial events such as weddings and funerals. It is within the social institution that *communitas* understood as a spiritual leveling of religious adherents exists alongside a religious community in which social differentiation and hierarchies usually persist. Religious institutions are usually the most deeply embedded social aspect of religion, since it is their task to control the external conduct of their members through rites, rituals, and ethical norms while providing an economic and political power base through which they can compete with other social institutions. Simply put, religious institutions are to a great extent socially constructed realities that provide for the habituation and rationalization of religious thought and behavior.

James (1961) viewed the church, synagogue, or other religious organization as a “secondhand” extension of the religious life. In terms of institutional leadership, Abraham Maslow (1970) distinguishes between “prophets” (i.e., those who found the religion) and “legalists” (those who regulate, systematize, and organize religious behavior in institutional forms). Even from this brief discussion of the interrelationships of the primary aspects of the religious life, one can see why Michael Novak says, “Institutions are the normal, natural expression of the human spirit. But that spirit is self-transcending. It is never satisfied with its own

finite expressions” (1971, p. 156). According to Novak, the basic conflict is between the human spirit and all institutions.

No religious institution has escaped criticism of its creeds, dogmas, ethics, and authoritative pronouncements from those within the tradition who insist that the essential faith demands revisions of the institution’s expressions of that faith. These criticisms give rise not only to reform movements but also to schisms and new sects that emerge as a result of the clash between the received faith in its textual and social forms and the religious experiences and impulses of a reformer or critic within the organization. Martin Luther was a reformer whose critique of his received Roman Catholic heritage was both personal and theological. Similarly, the numerous Buddhist sects that arose in the first hundred years after the death of the Buddha gained their impetus from quarrels over doctrine, lifestyle, and interpretations of the essential nature of the faith. The Sunni and Shi’a (also called Shi’ite) branches of Islam have dozens of contemporary expressions that emanate from a fundamental split in the tradition that occurred shortly after the death of Muhammad and focused on the source of authority for future proclamations in Islam. Typical of other religious traditions, Islam gave early birth to a pietistic mystical tradition, known as Sufism, which has consistently criticized both major theological branches of that religion for their legalistic and worldly focus to the detriment of the nourishment of the spiritual life. The Kabbala is a similar type of mystical reform tradition within Judaism. From one perspective, sectarian and schismatic movements are attempts to recapture the original experience and spirit of a religious tradition in response to institutionalized forms of worship and expression that appear devoid of the core spirit that gave birth to them. Nonetheless, in those cases where the new movement or sect survives its charismatic beginning, it necessarily develops the same institutional forms (religious community, rituals, ethics, etc.) that it rejected in its predecessor and that are experienced by some faithful later generations as too distant from its spiritual foundation and in need of reform. This pattern of dissatisfaction with institutional codifications of religious experience, a time of spiritual innovation or reform, and then institutionalization of the reform is one that continues in all the major religious traditions in the world,



producing new sects or, in rare cases, altogether new religious traditions.

### NEW RELIGIOUS MOVEMENTS

The attempt to reform a traditional religion in a given cultural setting sometimes has produced a new religious movement (NRM) that threatens the established norms and values of the host society, not just the established religious institution. Often an NRM emanates from an established religion as a reform or even extension of that tradition. An example is early Christianity, which some Jewish and non-Jewish converts saw as fulfilling Jewish prophecy and others regarded as a dangerous and heretical sect that threatened both the Jewish and the Roman institutions of Jesus' time. When the connection with the established tradition is more tenuous, the new revelation and resulting behavior distance themselves almost immediately from traditional institutional forms. For example, Joseph Smith's discovery of lost tablets of scripture not only "completed" the Christian revelation and scriptures but essentially replaced them. Smith's Mormonism promoted theological (e.g., preeminence of the Book of Mormon), ethical (e.g., polygamy), and other views and practices that were at odds not only with traditional Christian norms and institutions but also with those of American society. Such NRMs often generate considerable opposition from both religious and political authorities who perceive a threat to their worldview and the norms that come from that *nomos*. In the first century after the death of Jesus, his followers were martyred by Roman authorities who considered them members of an NRM outside the protection of law afforded Jews in the Roman Empire. Likewise, by the end of the nineteenth century in America, the Mormons not only were attacked by their Christian neighbors as heretical "cult" but were for a time denied the legal right to hold property and to marry.

Approximately one hundred years after groups such as the Mormons, the Seventh-Day Adventists, and the Theosophical Society were considered "cults" to be suppressed, a new wave of NRMs (also called "cults") flooded America. Some of those NRMs were essentially splinter groups of Christians (e.g., Jesus movements) whose evangelical fervor and communitarian lifestyle set them

apart from more established Christian churches. Other NRMs, such as Scientology, were the imaginative offspring of idiosyncratic founders such as L. Ron Hubbard, a science fiction writer who promised "total freedom" to all who would practice his strict regimen of psychological and spiritual "clearing." Still other NRMs were imports from Asia with gurus such as Maharishi Mahesh Yogi (Transcendental Meditation) and Guru Maharaj Ji (Divine Light Mission) who taught their own particular Hindu meditational paths to enlightenment. One NRM, the Unification Church of Sung Mung Moon, was essentially a syncretistic blend of Christian missionary and Korean folk religious traditions. The Reverend Moon claims to have had a special revelation on Easter Sunday in 1936, when Jesus appeared to him and asked that Moon complete the messiah's work. Moon's revelation led to a new scripture called *The Divine Principle*, new rituals, and a worldwide mission to unify all Christian and world faiths.

Finally, some of the NRMs of the 1960s in America were not "new" at all but instead were traditional faiths of other cultures seeking converts in an American mission field. One such NRM was the International Society for Krishna Consciousness (ISKCON), more commonly called the Hare Krishnas. While lumped together with other NRMs, the Hare Krishnas practice what is more properly understood as a traditional form of devotional (*bhakti*) Hinduism centering on the god Krishna. This devotional Hindu faith was brought to America in 1965 by the Hindu sage A. C. Bhaktivedanta Prabhupada. Prabhupada was an *acharya*, or spiritual teacher, whose lineage traces back to the Krishna reformer Chaitanya in the sixteenth century and whose own guru asked him to bring the Hare Krishna faith to English-speaking people. While adapting his teachings to a foreign culture as all missionaries must, Prabhupada taught the same Indian scriptures (e.g., *Bhagavata Purana*), rituals (e.g., worship before Krishna images and chanting Krishna's name), religious dress (e.g., saffron robes), and ethics (e.g., vegetarianism and ritual cleanliness) that had been taught by Indian masters for centuries. While the Krishna faith originated over 2,000 years ago, part of what made this religion seem so new and different to American youths and religious institutions was its evangelical missionary and ecstatic devotional elements (e.g., public chanting and dancing), which

were innovations of Chaitinya's reform nearly 400 years ago (see Shinn 1987a).

Whatever the origin or character of NRMs, they represent external challenges to established religions in much the same way that sectarian reforms represent internal challenges. From the point of view of formative religious experiences, NRMs offer alternative spiritual paths to religious seekers who do not find spiritual satisfaction in their natal or traditional religious institutions (Ellwood 1973; Richardson 1985; Shinn 1993). The host society's response to NRMs often reveals the extent to which that society's secular or religious institutions satisfy the needs of its populace (Robbins and Anthony 1981; Barker 1982; Wilson 1981). When religious institutions have stagnated or strayed from their spiritual source, challenges and alternatives arise from within. Likewise, evangelical and missionary ventures from religions around the world take whatever opportunity they are given to provide alternative paths to spiritual fulfillment.

### INTERSECTION OF WORLD RELIGIONS

One tendency of insitutionalized religious traditions is to seek to become world religions. The impetus to spread a religion throughout the world sometimes comes from the exclusivistic theological claims that assert the superiority of one faith over another (e.g., Christianity and Islam). Some religious traditions actively seek less to convert others than to assimilate other religions into their own theology and practice (e.g., Hinduism). Still others spread to other lands and cultures after being forced out of their homelands (e.g., Judaism and Buddhism). The broad reach of world religions has resulted in multifaith societies such as India (e.g., Hinduism, Islam, Sikhism, and Jainism), China (e.g., Confucianism, Buddhism, and Taoism), and the United States (e.g., Christianity, Judaism, and Islam), where different religions have coexisted for centuries. What can one expect of the interaction of world religions as rapid communications and travel bring people and their religious faiths face to face in ever greater numbers in the twenty-first century?

First, it should be expected that wherever religious institutions are interwoven with political and cultural institutions, resistance to or rejection of other world faiths will occur. This tendency will be exacerbated in areas where religious funda-

mentalism is the dominant voice. Islamic states such as Pakistan, Iran, and Afghanistan reveal how religious institutions are interwoven with political institutions in ways that suppress tolerance of other faiths. Adding tribal or ethnic loyalties to the mix only increases the difficulty of achieving interreligious tolerance and harmony. The Catholics and Protestants in Northern Ireland, the Bosnian Muslims and Serbian Christians in Bosnia-Herzegovina and the Tamil Hindus and Singhalese Buddhists in Sri Lanka all represent inseparable blends of political, ethnic, and religious exclusivity. Therefore, one mode of interaction of world religions will be intolerance of and sometimes violence toward other faiths created to a great extent by the socialization of religious institutions by the nationalistic and ethnic norms of the people and cultures they intend to save.

Second, in areas where religions have coexisted for a long time, it is common for accommodations and even assimilation to occur that reflect the common home. For example, Hinduism, Buddhism, and Islam have coexisted for more than nine hundred years in India, and in spite of their sometimes violent interactions, remarkable innovations have occurred. Leaders from the Muslim King Akbar to the Hindu sage Gandhi have sought to bring about mutual respect among the religions of India and all the world. Likewise, devotional Hinduism historically has often bridged religious divides by inviting people of all faiths and castes to join in its worship. In the case of Sikhism, Guru Nanck blended devotional Hindu traditions with certain Islamic tenets to form a syncretistic new faith in the sixteenth century. A similar phenomenon occurred in Iran, where Zoroastrian and Islamic roots gave rise in the nineteenth century to the Baha'i faith, which incorporates the scriptures and symbols of all the major world religions into a new syncretistic religion. While the birth of such new syncretistic world religions is rare, what does occur often—and probably will increase—is the adoption of ideas (e.g., reincarnation and impersonal divinity) and practices (e.g., vegetarianism and meditation) from one faith by persons of another faith.

Third, some religious individuals and institutions will continue to seek dialogue with and understanding of persons of other faiths while maintaining their own religious ideas and practices. For example, Mahatma Gandhi was deeply influenced

by the Christian and Muslim scriptures and near the end of his life sought peace between Hindus and Muslims when few others could rise above communal loyalties. Still, when shot by an assassin, Gandhi uttered the name of his Hindu family divinity, Rama. Gandhi appreciated the teachings and practices of other world religions but died a Hindu. In a similar fashion, the Buddhist Sarvodaya Movement in Sri Lanka borrows liberally from Gandhi's ideas and disciples even as it embeds its work in Buddhist ideas and practices. So too the Reverend Dr. Martin Luther King, Jr., learned the rudiments of nonviolent action from Gandhi's teachings while situating them within his Christian theology and faith. Thus, even when certain ideas are transferred from one faith to another out of respectful dialogue and interaction, it is common for one's native tradition to remain at the core of one's thought and action.

On a more formal level, there have been many attempts at interfaith dialogue in which the formulation of a common theology (i.e., "perennial philosophy") or practice for all religions has been sought (see Shinn 1987b). The Christian Trappist monk Thomas Merton spent many of the last years of his life reading about and having a dialogue with persons of other faiths. He was accidentally killed in Bangkok, Thailand, during an interfaith conference with Christian, Buddhist, and other monks from Asia. Most efforts at interfaith dialogue arise when individuals seek to understand their own faith better and to transcend the institutional reflections of a limited time and place. Both formal and informal dialogues are certain to increase as "the global village" becomes a reality and world religions become increasingly familiar in all lands.

### CONCLUSION

Clifford Geertz argues that each world religion is essentially "(1) a system of symbols which acts (2) to establish powerful, pervasive, and long-lasting moods and motivation in men by (3) formulating conceptions of a general order of existence and (4) clothing these conceptions with such an aura of factuality that (5) the moods and motivations seem uniquely realistic" (1968, p. 1). This socio-anthropological definition of religion embraces in a clear and simple fashion most of the interpretation of underlying relationships that this article has described. Any religion, whether established or new,

is a system of symbols that simultaneously attempts to express and reveal dimensions of sacred experience beyond that of the everyday by using socially conditioned language and conceptions. Likewise, the general order of existence (*nomos*) that is formulated in the myths, rituals, and ethical norms of a religious tradition emerges from the social consciousness, communal norms, and shared conceptions of the community which give rise to those elements. Finally, what Berger calls "legitimation" and Geertz calls "factuality" represent nothing other than broad-based social acceptance of certain religious beliefs. Consequently, from their inception in religious experience to their full social expression in concrete institutions, religious traditions involve an interplay between personal and social forces. No aspect—experiential, mythical, ritual, ethical, doctrinal, or institutional—of any of the world's religious traditions escapes some social conditioning, and no culture or society is left unchallenged by its religious expressions and lifestyles.

### REFERENCES

- Barker, Eileen (ed.) 1982 *New Religious Movements: A Perspective for Understanding Society*. New York: Edwin Mellon Press.
- Berger, Peter 1969 *The Sacred Canopy: Elements of a Sociological Theory of Religion*. Garden City, N.Y.: Doubleday.
- , and Thomas Luckmann 1967 *The Social Construction of Reality*. New York: Doubleday.
- Durkheim, Emile 1965 *The Elementary Forms of the Religious Life*, trans. Joseph Ward Swain. New York: Free Press.
- Eliade, Mircea 1959 *The Sacred and the Profane*, trans. Willard R. Trask. New York: Harper and Brothers.
- Ellwood, Robert S., Jr. 1973 *Religious and Spiritual Groups in Modern America*. Englewood Cliffs, N.J.: Prentice-Hall.
- Freud, Sigmund 1928 *The Future of an Illusion*, trans. W. D. Robson-Scott, Horace Liveright, and the Institute of Psychoanalysis. London: Hogarth Press.
- Geertz, Clifford 1968 "Religion as a Cultural System." In Michael Banton, ed., *Anthropological Approaches to the Study of Religion*. London: Tavistock.
- James, William 1961 *The Varieties of Religious Experience: A Study in Human Nature*. New York: Collier.
- Jung, Carl 1938 *Psychology and Religion*. New Haven, Conn.: Yale University Press.

- Malinowski, Bronislaw 1954 "Myth in Primitive Psychology." In Bronislaw Malinowski, *Magic, Science and Religion*. Garden City, N.Y.: Doubleday.
- Maslow, Abraham H. 1970 *Religions, Values, and Peak Experiences*. New York: Penguin.
- Neidardt, John G. 1972 *Black Elk Speaks*. New York: Pocket Books (from field notes contained in Raymond J. DaMillie 1984 *The Sixth Grandfather: Black Elk's Teachings given to John Neihardt*. Lincoln: University of Nebraska Press).
- Novak, Michael 1971 *Ascent of the Mountain, Flight of the Dove*. New York: Harper & Row.
- Otto, Rudolph 1946 *The Idea of the Holy*, trans. J. W. Harvey. London: Oxford University Press.
- Richardson, James T. 1985 "The Active vs. Passive Convert: Paradigm Conflict in Conversion/Recruitment Research." *Journal for the Scientific Study of Religion* 24:163–179.
- Ricoeur, Paul 1972 "The Symbol Gives Rise to Thought." In Walter H. Capps, ed., *Ways of Understanding Religion*. New York: Macmillan.
- Robbins, Thomas and Dick Anthony 1981 *In Gods We Trust: New Patterns of Religious Pluralism in America*. New Brunswick: Transaction Books.
- Shinn, Larry D 1977 *Two Sacred Worlds: Experience and Structure in the World's Religions*. Nashville, Tenn.: Abingdon.
- 1987a *The Dark Lord: Cult Images and the Hare Krishnas in America*. Philadelphia: Westminster Press.
- 1987b "Inside the Mind of the Infinite: Dialogue and Understanding in Interfaith Encounters." In Larry D. Shinn, ed., *In Search of the Divine: Some Unexpected Consequences of Interfaith Dialogue*. New York: Paragon.
- 1993 "Who Gets to Define Religion? The Conversion/Brainwashing Controversy." *Religious Studies Review* 19(3): 195–207.
- Smart, Ninian 1969 *The Religious Experience of Mankind*. New York: Scribners.
- Turner, Victor 1969 *The Ritual Process: Structure and Anti-Structure*. Chicago: Aldine.
- Wach, Joachim 1958 *The Comparative Study of Religions*. New York: Columbia University Press.
- Weber, Max 1963 *The Sociology of Religion*, trans. Ephraim Fischhoff. Boston: Beacon.
- Wilson, Bryan (ed.) 1981 *The Social Impact of New Religious Movements*. New York: Rose of Sharon Press.

LARRY D. SHINN

# Index

---

*Note: Page numbers in **boldface** indicate article titles.*

## A

- Aarhus University, 2450  
Aaronson, Neil K., 2301, 2305  
AARP. *See* American Association of Retired Persons  
Abacha, Sani, 2134  
Abbagnano, Nicola, 1468–1469  
Abbott, Andrew, 1636, 2262, 2263, 2264, 2297  
ABCD. *See* Americans for Better Care of the Dying  
Abdullah (Saudi Arabian crown prince), 2132  
Abel, Theodore, 1074  
Abelard, 1179  
Abell, Peter, 2297  
Abelson, R. P., 335  
Åberg, Rune, 2452  
Åbo Academy, 2450, 2451  
Abolafia, Mitchell, 738–739  
Abolitionism, 2270, 2725  
Aborigines, 1069, 3079  
Abortion, 2238–2241  
    antiabortion activity, 2240, 2266, 2277, 2717, 2723  
    criminalization effects, 2240–2241  
    and fertility determinants, 1006  
    legal challenges to, 2240  
    legalization, 2240–2241  
    and legislation of morality, 1579, 1580  
    liberal/conservative perspectives, 1602  
    policy, 938, 943  
    procedures, 2238–2240  
    public funding restraints, 315  
    public opinion on, 2274, 2277  
    reasons offered for, 2239  
    spontaneous, 2233–2234, 2238  
    teenage rates, 2235  
    in United States, 2180  
    *See also* Family Planning; Pregnancy and pregnancy termination  
Abraham, Karl, 650  
Abrahamian, Ervand, 1871  
Abramowski, Edward, 2118  
Abrams, Philip, 576, 1195, 1197  
Abramson, Lauren, 651, 652  
Abramson, Paul, 3222, 3223  
Abridged Life Table, 613, 615  
Abruzzese, Alberto, 1473  
Abruzzese, Salvatore, 1473  
Abscam, 2127–2128  
Absolute judgment, 597–598  
Absolute monarchy, 2356  
Absolute properties of collectives, 1592  
Absolutism. *See* Dictatorship  
Abstract groups, 2298  
Abstracting and indexing services, 1608–1609  
    citation indexes, 1610–1611  
Academic and Professional Affairs Program (ASA), 150  
Academic freedom, 276–278  
Academic libraries. *See* Library resources and services for sociology  
Academic medical center hospitals, 1820  
Academy of Sciences (Ukraine), 2982  
Accelerated failure-time models, 870–871  
Accidents. *See* Disaster research; Society and technological risks  
Accornero, Aris, 1467  
Accountants and auditors, 2259  
Acculturation, 842  
    African, 64–65  
    Mexican, 1858, 1861  
Accumulative advantage (Merton concept), 2691  
Accusatory vs. inquisitorial legal model, 479  
Achen, Christopher H., 1595  
Achievement syndrome, 848  
Achievement-oriented society, 485  
ACLS. *See* American Council of Learned Societies  
ACLU. *See* American Civil Liberties Union  
Acquaintance rape, 2558–2559, 2577  
Acquaviva, Sabino, 1469, 1470  
Acquired immune deficiency syndrome. *See* AIDS/HIV

- Acta Sociologica* (journal), 2451
- Actes de la recherche en sciences sociales*, 1027
- ACTH (adrenocorticotrophic hormone), 652
- Action  
   values as transcending, 3213  
   Weber's four typologies of, 2519-2520
- Action, collective. *See* Collective behavior
- Action theory, 1080
- Activities of daily living (ADL), 1653, 1654, 1655, 1667
- Adams, Bert N., 485-486
- Adams, John, 273, 584
- Adaptive behavior  
   by religious organizations, 2379  
   personal dependency and, 2066
- Addiction. *See* Alcohol; Drug abuse
- Adkins, Janet, 3085
- ADL (activities of daily living) scales, 1653, 1654, 1655, 1667
- Adler, Alfred, 1714, 1715, 1717, 1718, 2084, 2088
- Adler, Nathan, 460
- Administration, Weber's three types of, 230
- Adolescence, **1-18**  
   acquaintance rape, 2577  
   abortion rates, 2235  
   acquaintance rape, 2588  
   AIDS/HIV risk, 2585, 2586-2587  
   alcohol use, 94, 95  
   anomie and strain theories, 166  
   attitude formation, 185  
   changing norms of, 1, 3  
   drug abuse concerns, 710-711, 712-713  
   duration of, 5, 6  
   and education, 10-11  
   educational attainment  
     predictors, 2784  
   fertility determinants, 1005, 1010  
   health assessment in, 1131  
   interdisciplinary study of, 2  
   life course focus, 2, 3-9, 12-13, 1620, 2861  
   and music, 1926  
   and parents' relationship, 2037  
   peer pressure in, 1057  
   pregnancy rates, 488, 2032, 2235  
   and self-esteem, 2513, 2534  
   sexual behavior patterns, 2551-2553  
   sexual orientation, 2552-2553  
   social historical perspective on, 3, 4-6  
   socialization, 10-11, 34, 2852-2860, 2861  
   socioeconomic status and health behavior, 1129  
   and suicide, 3078, 3080  
   and temporary work, 1724  
   transition markers in, 6, 7, 8, 9  
   work experience, 8-9, 11-12, 13, 34  
   and youth subculture, 512, 514  
   *See also* Gangs; Juvenile delinquency, theories of; Juvenile delinquency and juvenile crime
- Adoption (of children), 597-598  
   by gays and lesbians, 315
- Adoption (of innovation), 677, 678-679  
   diffusion research, 86-91  
   rural sociology on, 2430
- Adorno, Theodor, 822, 1027, 2986  
   on authoritarian personality, 317, 334, 540  
   and critical theory, 539, 540-541, 542, 544  
   and genocide theory, 1070  
   and German sociology, 1075, 1076, 1077, 1078, 1079  
   on mass culture, 173, 1645
- ADPSS. *See* Istituto Superiore di Sociologia
- Adreotti, Giulio, 2129
- Adult day care, 1657
- Adult education, **18-25**  
   acceleration of, 22, 23, 3064  
   contemporary structures, 22-24  
   participatory research, 2039, 2040, 2041  
   providers of, 18, 19, 21  
   and structural lag, 3064  
   types of, 18-20
- Adultery, 2541-2545  
   as divorce ground, 701, 1737
- Adulthood, **25-41**  
   accommodation and, 33  
   age norms, 29, 30, 31  
   alcohol consumption rates, 95  
   assumption of, 25  
   attachment styles, 2068  
   childfree, 109-111  
   delayed onset, 26, 27  
   divorce effects, 706, 707-709  
   education continuation. *See* Adult education  
   and family formation, 27, 28, 34, 35  
   and financial independence, 29  
   and gender roles, 30, 33, 34  
   goals of, 33  
   and labor force participation, 3262-3263  
   and moral development, 1900, 1903-1904  
   obstacles to, 27  
   and occupational and career mobility, 1892-1894  
   parental relations, 2037-2038  
   parental roles, 27, 28, 2031-2038  
   and personal dependency, 2063-2064  
   prerogatives of, 25  
   psychological orientations in, 32-33  
   race and transition to, 35  
   and residence, 26, 29  
   rituals of, 29  
   role adaptation, 32-36  
   and self-concept, 2508  
   and social class, 33  
   socialization in, 2860, 2861-2862  
   and structural lag, 3063  
   subjective indicators of, 25  
   subjective transition to, 28-30  
   transition markers in, 7, 25, 27, 28-32, 36, 2861  
   transition variations in, 26-28
- Advance directives, 585-586, 587, 3064, 3083
- Advanced degrees, in postindustrial society, 2196-2197
- Advanced industrial societies. *See* Postindustrial society

- Advanced Theory of Statistics (Kendall), 3035
- Advances in Field Theory* (Wheelan et al.), 1013
- Adventures of Huckleberry Finn, The* (Twain), 277
- Aesthetics, 1925
- Aetna (health insurance provider), 1822
- Affect Balance Scale, 2303, 2304, 2306
- Affect, Cognition, and Stereotyping* (Mackie and Hamilton eds.), 2244
- Affect control theory and impression formation, **41–47**
- applications of, 45–46
- cultural norms and, 2528
- emotions and, 43, 44–45, 780, 2527–2528
- impression formation and management, 42–44, 2506, 2776
- mathematical formulation, 1790
- sentiments and, 2521
- social perception and, 2750–2753
- social stimuli and, 2774
- Affective aggression, 68–69
- Affective-cognitive consistency model, 335, 337
- Affines, 1509
- Affinity, 2632
- Affirmative action, **47–53**
- admissions policies, 50, 51
- backlash against, 2140, 2267, 2499, 2706
- bans on, 51
- categories affected, 48–49
- civil rights and, 47, 48, 49, 2496, 2497
- civil service examinations and, 49
- consent decrees and, 49
- controversy in, 47, 49
- as discrimination remedy, 693, 694
- enforcement of, 47
- equality of opportunity and, 49, 833
- gender and, 50, 1063
- goals and timetables, 48
- government contracts and, 49, 50
- legal meaning of, 47
- political correctness and, 2140, 2141
- quotas and, 49
- requirements of, 48
- as reverse discrimination, 2706
- salary differentials and, 690, 692
- set-aside programs in, 49, 50
- social justice beliefs and, 2706
- Supreme Court cases, 50–51
- Affluence. *See* Wealth
- Afghanistan
- Islamic fundamentalism, 2371, 3288
- sociodemographic profile, 2938
- AFL, 1530
- AFL-CIO, 1532, 2148–2149
- Africa
- abortion illegality, 2240, 2241
- age-set societies, 1623
- AIDS/HIV epidemic, 2585, 2591, 2592–2593
- art, 173
- colonization, 1934
- Demographic and Health Surveys, 633
- dependency theory and, 643, 644
- ethnonationalism, 1934, 1948
- fertility determinants, 1006
- fertility rates, 219–220
- fertility transitions, 627–628, 1008, 2178
- health-care systems, 381
- historical empires, 2999
- legal ethnography, 1549, 1550
- life expectancy, 623
- military dictatorships, 3002, 3003
- multilingualism, 2909
- national boundaries, 1934
- political and governmental corruption, 2132–2134
- population factors, 2182
- poverty in, 2216
- slave trade from, 320, 321
- slavery and involuntary servitude in, 2501–2503, 2604–2608
- sub-Saharan population growth, 628, 1008
- See also* African studies; *specific countries*; Sub-Saharan Africa
- African American studies, **53–59**
- African studies inclusion, 67
- AIDS/HIV risks, 2587, 2590
- alcohol consumption rates, 94–95
- alienation, 101–102
- assimilation, 843–844
- attributional patterns, 196
- births per woman, 2032
- case studies, 243, 244, 245
- census undercounts, 286
- childhood poverty rates, 127
- city-suburb disparities, 3072
- clinical sociology clinics, 326
- direct and indirect
- discrimination, 54–57, 143, 689, 691, 692, 693, 2333
- divorce effects, 704–705, 707
- divorce potential, 112, 126
- educational attainment, 2931, 2932–2933
- environmental equity, 789
- equality of opportunity, 826, 845
- experiential communities, 245
- feminist theory, 545
- filial responsibility, 1020
- genocide charges, 1071–1072
- historical marriage and family structures, 121–122
- homelessness, 1204
- household income, 1279, 1280
- illegitimacy, 1259
- illicit drug convictions, 714
- income inequality, 3048
- infant and child mortality, 1328
- intelligence testing, 1360, 2330
- interracial marriage
- demographics, 124, 1776
- Islam adherents, 2950
- job segregation, 3264–3265
- juvenile crime rates, 1490, 1491
- language and dialects, 2901
- life expectancy, 1631
- marriage demographics, 124, 1775, 1776
- military service, 1879, 1880
- music, 1926
- nonmarital childbearing, 125

- per capita income, 1280  
 political alienation, 102–103  
 poverty level, 2215  
 poverty theories, 1287, 2333–2334  
 prejudice, 2243–2246  
 professional representation, 2260  
 protest movements, 2266  
*See also* Civil rights movement  
 remarriage rates, 126, 2388  
 residential segregation indices, 57, 2500–2505  
 retirement, 2407  
 rural conditions, 2428–2429  
 segregation and desegregation, 2491–2499, 2601, 2608  
 slave experience, 54–58, 64, 121–122, 320, 321, 2491–2492, 2599–2601  
 social deprivations, 322  
 sociolinguistics, 2901, 2908–2909  
 status attainment, 3044, 3046  
 stereotypes, 64, 2243  
 subculture of male violence, 664, 665, 666  
 suburbanization, 3074–3075  
 suicide rates, 3078, 3079  
 underemployment, 1721, 1722  
 urban migration, 532  
 urban riots, 555–556, 557, 558  
 urban underclass, 3198–3200  
 white immigrant experience comparison, 2333–2334  
 white-black relations theories, 53–54, 56  
 and Wisconsin model, 2784  
 women in labor force, 123  
*See also* Class and race; Race
- African National Congress (ANC), 2047  
 African primitive art, 173  
 African slave trade, 54, 320, 321, 2597, 2598, 2599, 2600, 2601  
 African studies, **60–68**  
     American scholarship, 66–67  
     cultural and social institutions, 63–65  
     and geography, 60  
     participatory research, 2040  
     and peoples, 61–63  
     religious core myths and rituals, 3281–3282  
     unifying themes, 65–66
- Afrocentricity, 54, 2140  
*Against Our Will: Men, Women, and Rape* (Brownmiller), 2587, 2589–2590, 2592  
 Agarwal, Priya, 180  
 Age appropriateness, life cycle and, 1623, 1625, 1626  
 Age Discrimination in Employment Act, 1397  
 Age effects, definition of, 80, 81  
 Age frequency distribution, 659  
 Age of Reason, 2206  
 Age patterns  
     cohabitation, 1750  
     crude death rate, 610–611  
     divorce rates, 1742, 1743, 1745–1747  
     first marriage rates, 620, 1425, 1742, 1743, 1744–1745  
     life-course variations, 1618  
     remarriage rates, 1742, 1743, 1748  
     spousal, 1744  
     underemployment, 1721, 1724  
     *See also* Birth and death rates; Cohort perspectives
- Age pyramids, 609–612, 635  
     as bar graph, 610, 611, 612  
     cohort perspective, 345, 609
- Age stratification theory, 79–80  
 Agency concept, 221, 1854, 2217, 2220, 2221  
     Marxist, 2645  
     and revolutions, 2414  
 Agency for Health Care Policy and Research, 1157, 2399–2400, 2401  
 Agency for International Development, 932  
 Agenda-control power, 2165–2166  
 Agersnap, Torben, 2451  
 Age-set societies, 1623, 1624  
 Age-standardized crude death rate (ASCDR), 611, 613  
 Ageton, Suzanne, 1492  
 Aggravated assault  
     absolute rates, 497  
     Uniform Crime Reports definition, 492  
     victimization rates, 499
- Aggregate data analysis, voting behavior, 3232–3233  
 Aggregate (ecological) level relationships, vs. individual level effects, 1594–1595  
 Aggregative fallacies, 1592  
     statistical analysis of, 1593–1594
- Aggression, **68–78**  
     affective vs. instrumental, 68–69  
     as sexual violence factor, 72, 2576–2583  
     basic definitions of, 68–69  
     causes of, 69–75, 2774  
     collective, 349, 2777  
     cues, 74  
     escalation cycle, 74–75  
     frustration leading to, 73, 349, 2670  
     and genocide, 1070  
     inhibiting factors, 2774  
     intervention, 75–76  
     macro-level theories of deviance on, 664, 671  
     media violence effects on, 272, 1762–1763, 2858  
     meta-analysis of sex differences in, 2531–2532  
     modeling theory and, 2858  
     removal of self-regulatory inhibitors to, 74  
     rewarding outcomes of, 2774  
     self-esteem and, 72, 2516–2517, 2777  
     self-presentation and, 2506  
     as sexual violence factor, 72, 2591  
     sibling, 331  
     social exchange theory on, 2670  
     social learning theory on, 70–71, 75, 1716  
     *See also* Crime, theories of; Crime rates; Juvenile delinquency and crime; Sexual violence and exploitation; Violence
- AGIL (Adaptation, Goal attainment, Integration, Latent Pattern Maintenance) system, 1554–1555, 1559–1560, 1978, 2005



- Aging. *See* Aging and the life course; Cohort perspectives; Filial responsibility; Intergenerational relations; Intergenerational resource transfer; Long-term care; Long-term care facilities; Retirement; Widowhood
- “Aging and Dying” (Marshall and Levy), 583
- Aging and Society* (Riley et al.), 1618
- Aging and the life course, **78–86**
- age-period-cohort effects, 80–81
  - alcohol consumption patterns, 94
  - American family patterns, 120, 129
  - case studies, 244–245
  - cohort perspectives, 342–347
  - as demographic factor, 636
  - demography of, 1159
  - depression risk, 652, 653, 656, 1839
  - effect of childlessness on, 110–111
  - effects of early life decisions on, 83
  - effects of early traumas and deprivations on, 82–83
  - eldercare provisions, 129–130, 1019
  - family structure trends and, 129, 925–926
  - female population, 2177
  - filial responsibility and, 1018
  - gender issues and, 83
  - health assessment, 1131
  - health behavior, 1139
  - health care, 1139
  - health promotion and, 1168–1169
  - heterogeneity research, 81–82
  - history-personal biography interrelationship, 82
  - increased numbers of elderly, 2180
  - intergenerational relations, 1018, 1386, 2037–2038, 2707
  - Japanese sociology on, 1482
  - labor-force participation and, 1524
  - leisure and, 1585–1587
  - in less developed countries, 931
  - and life cycle transitions, 1624
  - life endurancy and, 1631–1632
  - life span and, 1631
  - life-course trajectories and, 83–85, 1614–1621
  - long-term care and care facilities, 1652–1678, 1826
  - mental illness rate decrease, 1838–1839
  - moral judgment and, 1903–1904
  - morbidity and, 1137
  - never-married population, 125
  - “new” old people, 345, 3061
  - person-centered research, 84–85
  - phenomenological investigation, 2103
  - primary/secondary, 1137
  - quality-of-life research, 2300, 2304–2305
  - religion and, 2965
  - remarriage rate, 1779
  - retirement and, 2401–2410
  - sexual behavior patterns, 2555–2556
  - social definitions of, 1315
  - social security systems and, 2795, 2798–2799, 2802–2803
  - social structural changes and, 3064–3065
  - social structures and, 345–346, 3063–3064
  - socioeconomic status and, 1138
  - structural lag concept, 3060–3062
  - suicide rate, 3078
  - welfare-state convergence theory and, 426–427
  - widowhood and, 3256–3257
- See also* Life course
- Agnew, Robert, 166, 664
- Agnew, Spiro, 2126
- Agnosticism, 2383
- Agrarian political parties, 2155
- Agrarian society, 2810, 2811
- transition from, 2176
- Agribusiness, 2433
- Agricultural innovation, **86–92**
- adoption-diffusion of, 86–91, 677, 678, 2460
  - agribusiness and, 1222
  - capital requirements, 1222
  - and development of cities, 305–306, 634
  - in Indonesia, 2976–2977
  - Marxist theory on, 2460–2461
  - in Mexico, 1858
  - and modernization theory, 1885
  - in newly industrializing countries, 1317
  - and rural sociology, 86–91, 677, 678, 2429, 2431–2433
- Agriculture Department, U.S., 88, 148, 2214, 2235, 2283
- AHCPR. *See* Agency for Health Care Policy and Research
- AHEAD. *See* Asset and Health Dynamics Among the Oldest Old
- Ahlquist, Karen, 1925
- AHS. *See* Association for Humanist Sociology
- Aid to Families with Dependent Children, 967, 1284, 1286, 1288, 2799, 2803, 2804, 2961, 3198
- AIDS/HIV, 2585–2592
- and Asian sex trade, 2607
  - bisexuality as risk factor, 2559, 2586, 2590
  - and civil liberties, 316, 318
  - and condom protection, 957, 2559, 2560, 2586, 2587, 2590, 2592, 2593
  - and courtship practices, 487, 489
  - death rates, 222, 224, 2592
  - depression in patients, 656
  - epidemiology, 814, 2576
  - gay men and, 112, 2555, 2559, 2570, 2585–2586, 2587, 2588, 2590, 2591
  - international situation, 2591–2593
  - intravenous (IV) drug use and, 712, 2576, 2577, 2578, 2579, 2580, 2582
  - legislation of morality and, 1579
  - prostitution and, 2559, 2560, 2561, 2582, 2607
  - rape and, 2576, 2585
  - research history, 2586–2690
  - risk factors in United States, 2586–2588
  - risk reduction in United States, 2588–2591

- as sexual behavior research  
impetus, 2550, 2561,  
2585–2587
- and sexually risky behavior,  
2559, 2586
- Ajax (Trojan hero), 3079
- Ajzen-Fishbein model, 2785
- Akan, 54
- Akbar (king of India), 3287
- Ake, Claude, 641
- Akers, Ronald, 667
- Alamán, Lucas, 1856
- Alan Guttmacher Institute, 2239
- Alapuro, Risto, 2452
- Alaskan Natives, 95
- Alba, Richard, 843
- Albania, sociodemographic  
profile, 2938
- Albanians (ethnic), 2362, 2608
- Alberoni, Francesco, 1469, 1470
- Albert, Hans, 822, 1077
- Albigensian heresy, 2968
- Albrecht, Milton, 172–173
- Albright, Madeleine, 1064, 1244
- Albrow, Martin, 227
- Alcohol, **92–99**
- aggression linked with, 73–74,  
984–985
  - bankruptcy linked with, 205
  - community assessment  
process, 366
  - consumption decline, 95
  - and criminalization of deviance,  
521, 525, 526
  - death rates from, 1640
  - and drunk driving, 93, 164–165,  
1640, 1641
  - and drunk-driving reform  
movements, 2722, 2725, 2877
  - ethnicity and, 94–95
  - Finnish studies, 2451, 2453
  - gender and, 94, 1640–1641
  - homelessness and, 1204
  - incest and, 1274
  - and legislation of morality,  
1576–1577
  - peer influence and use of, 667
  - physical effects of, 92–93,  
1640, 1642
  - Prohibition/organized crime  
relationship, 2019, 2127
  - as rape excuse, 2577
  - religion and, 94–95
  - sexually transmitted diseases and,  
2587, 2588, 2589
  - social characteristics in use of,  
93–96
  - See also* Alcoholism; Drug abuse
- Alcohol, Drug Abuse, and Mental  
Health Administration, 1157
- Alcohol myopia, 73–74
- Alcoholics Anonymous, 96, 715, 853
- Alcoholism, 95–97
- depression and, 655
  - disease concept, 96–97, 1816
  - as divorce factor, 1737
  - genetic factors, 97
  - health risks, 93, 94, 1640
  - and life course, 1839
  - medical treatment of, 521
  - medicalization of, 1816
  - racial, ethnic, and religious  
factors, 94–95, 135
  - as self-destructive behavior, 3077
  - stigmatization of, 1815
  - as suicide predictor, 3078–  
3079, 3081
- Alexander II, Pope, 2967–2968
- Alexander, Jeffrey, 1704, 1710
- Alexis Carrel Foundation, 1026
- Alford, Robert, 2166
- ALFs. *See* Assisted living facilities
- Alger, Chadwick F., 1948
- Algeria, 1865, 1866
- anticolonial revolution, 3001
  - corruption in, 2132
  - fertility decline, 628
  - interethnic hate, 2529
  - sociodemographic profile, 2938
  - and terrorism, 3137
  - and wartime rape, 2580
- Alienation, **99–106**
- critical theory and, 540
  - definitions of, 99–100
  - Marxist theory of, 100, 697, 1705
  - measurement of, 100–102
  - political, 100–104
  - postmodern cultural theory  
and, 2173
  - stable measurement of, 2346,  
2350–2352
  - and work orientation, 3270
- Alinsky, Saul, 325
- Allah, 3280, 3284
- Allardt, Erik, 2450
- Allen, Carolyn, 1648
- Allgeier, Elizabeth Rice, 2558
- Allied Irish Bank, 1042
- Allison, Graham, 3244
- Allison, Scott T., 3222
- Allocation theory. *See*  
Credentials theory
- Allport, Floyd, 679
- Allport, Gordon, 1400, 2084, 2085,  
2087, 3214
- All-Russian Public Opinion Research  
Center, 2982
- All-volunteer force (AVF), 1877–  
1881, 1882
- Almond, G., 2917
- Altercasting, 3097
- Alternative dispute resolution, 1550
- Alternative lifestyles, **106–114**
- childfree adults, 109–111, 634,  
1506, 1625, 2035
  - and companionship family  
concept, 1506
  - countercultures, 459–462
  - as demographic research area,  
636–636
  - divorce and remarriage, 112–113,  
126, 2387–2393
  - and family size, 977
  - gay and lesbian, 111–112,  
2545–2547, 2567
  - See also* Cohabitation; Singlehood
- Alternative medicines, 654–655
- Althausser, Ronald P., 1985–  
1986, 1987
- Althusser, Louis, 226, 721, 1645,  
1753, 1784
- Altruism, **114–120**
- definition of, 114
  - during disaster, 684
  - family care, 1657–1658
  - history of, 114–115
  - parental, 1508

- reciprocal, 115, 1657–1658, 2882, 2883–2884
- research on, 116–117, 2774
- self-presentation and, 2506
- situational variables, 116–117
- as socialist ideal, 2847
- sociobiological view of, 115, 118, 2882–2884
- sociological context of, 117–118
- theories of, 115, 2774
- and values research methods, 3220, 3221
- “Altruism and Prosocial Behavior” (Batson), 118
- Alwin, Duane F., 2704, 2706
- Alzheimer’s disease, 1839
- AMA. *See* American Medical Association
- Ambedkar, B. R., 251
- Ambiguity
- authoritarian personality’s intolerance of, 334
  - conformity experiments, 401, 402
  - and emergence of crowd behavior, 558
  - replication and, 2396
- Ambrosini, Maurizio, 1472
- Amerasia* (journal), 179
- American Academy of Arts and Sciences, 1039, 2682
- Fundamentalism Project, 2372, 2945
- American Anti-Slavery Society, 2270
- American Association of Retired Persons, 2148
- American Betrayed* (Grodzin), 181
- American Board of Internal Medicine, 588
- American Cancer Society, 588
- American Civil Liberties Union, 315–316, 318
- American College of Physicians, 588
- American Committee for the Outlawry of War, 2270
- American Community Survey*, 284–285
- American Council of Christian Churches, 2370
- American Council of Learned Societies, 406
- American Demographics Index of Well-Being*, 2687
- American Economic Review, The*, 2920
- American Educational Association, 3207
- American Enterprise Institute, 1601
- American Evaluation Association, 864, 866
- American Evasion of Philosophy, The* (West), 2220
- American Express Company, 2402
- American families, **120–133**, 142
- African American, 121–122, 2333
  - American Indian, 120–121, 134
  - blended, 112–113, 126, 2390–2391
  - childbearing patterns, 125, 2032–2033
  - cross-national comparisons, 130
  - division of labor in, 1, 122, 2034
  - effects of long-term trends on, 127–130
  - elder members, 120, 129–130
  - historical gender roles, 122
  - historical overview, 120–124
  - immigration effects, 122–123
  - kinship system, 1502–1503, 1504, 1511, 1514–1515
  - labor-force participation, 1524, 1525–1526
  - life-cycle perspective, 1625
  - marital laws on affines and cousins, 1509
  - marriage and, 1734, 1735, 1738
  - Middletown study findings, 364
  - and military sociology, 1882
  - monolithic model, 106
  - parental coalitions, 331
  - parental rewards and costs, 2034–2035
  - parental subsidies, 2035
  - single-parent. *See* Single-parent households
  - structural modifications, 113
  - two-child norm, 2182
  - values, 106
  - widowhood, 126
- See also* Alternative lifestyles; Divorce; Family and household structure; Filial responsibility; Intergenerational relations; Intergenerational resource transfers; Kinship and family types
- American Federation of Labor, 1530
- American Federation of Labor-Congress of Industrial Organization, 1532, 2148–2149
- American Federation of State, County, and Municipal Employees, 372
- American Fertility Association, 2588
- American Hospital Association, 588
- American Indian Movement, 137, 1298
- American Indian Policy Review Commission (1976), 136
- American Indian studies, **133–140**
- alcohol abuse rates, 95, 135
  - anthropological, 2888, 2890
  - demographic, 133–134
  - and dependency theory, 134–135
  - discrimination, 57, 58, 135
  - divorce rates, 126
  - family structures, 120–121, 134
  - interracial marriage demographics, 1776
  - life expectancy, 1169
  - pan-Indian movement, 136–137
  - political and legal status, 135–136
  - potlatch reciprocity, 2883–2884
  - religious experience and symbols, 3277–3278, 3279, 3281, 3282
  - reservation conditions, 135
  - social and economic status, 134–135
  - tribal genocidal massacres, 1070
  - tribal sovereignty doctrine, 136
  - underemployment, 1721
  - urbanization, 136–137
- See also* Indigenous peoples
- American Journal of Sociology*, 176, 325, 583, 735, 1528, 1606
- Middle Eastern studies articles, 1864, 1870
- on research funding sources, 2398, 2399
- sociology of literature articles, 1646

- American Legion, 3230
- American Lives* (Clausen), 1616, 1620
- American Medical Association,  
1148, 2261  
lobbying activity, 1826  
Patient Self-Determination Act,  
586–587, 588
- American Medical International,  
1821
- American National Elections  
Studies, 2476
- American Occupational Structure, The*  
(Blau and Duncan), 260–261,  
2713, 3035–3036
- American Pain Society, 588
- American Prospect, The* (Winter), 368
- American Psychiatric Association,  
111, 1832–1833
- American Psychological Association,  
1360, 1876  
conference on death (1956), 581  
validity standards, 3207,  
3210, 3211
- American Revolution, 473–474,  
2138, 2865, 3000, 3229
- American society, **140–148**  
abortion policy and procedures,  
2180, 2238, 2239–2240  
African American studies, 53–59  
age-specific fertility rates, 2180  
age-stratification system, 1623  
aging population, 3061  
agricultural innovation, 87–90  
AIDS/HIV risk factors,  
2586–2588  
AIDS/HIV risk reduction,  
2588–2591  
alcohol control laws, 1577  
alcohol-consumption decline, 95  
alienation theory, 99, 101–104  
American Indian studies,  
133–140  
anomie and, 166  
anthropological studies, 2888,  
2890, 2893  
assimilation and, 143, 178  
bankruptcy sources, 206–207  
birth and death rates, 2180  
birth rate decline, 2032  
blue laws, 1560  
as capitalistic, 99, 143–144  
case studies, 243–248  
censorship and regulation of  
expression in, 268, 270–279  
census, 282–287, 493  
characteristics of, 140–141  
childhood sexual abuse data, 289  
childless marriages in, 2035  
church-state separation, 146,  
2356, 2357, 2358  
city systems, 306–307  
class and race in, 319, 321–323  
class-based voting, 604  
and clinical sociology history,  
323–326  
and Cold War-era triad, 332  
common law system, 465–  
472, 476, 477  
communitarianism, 361  
community studies and reform,  
364–366  
comparative health-care systems,  
374, 375, 376, 377, 379  
complex organizations, 143–144  
corporate organizations, 442–444  
corruption in, 2124–2128  
countercultures, 459–462, 2366  
courtship in, 484–485  
crime data sources, 492–498  
crime surveys, 498–500  
criminological theory, 503–505  
cultural anthropology, 2892  
daily time use, 3160  
as democracy, 603, 606  
and demographic training,  
636–637  
denominational affiliations,  
2376–2377  
direct and indirect  
discrimination, 143, 689–692  
disaster research, 682, 686, 687  
divorce history, 700–703  
divorce patterns, 125–126,  
127–128  
divorce rates. *See subhead*  
marriage divorce rates *below*  
drug abuse rate and treatment,  
710–718  
education and status attainment,  
1987–1988, 2713–2714,  
2783–2784, 2929–2934,  
3043, 3045  
educational institutions, 145,  
762–770  
and endogamy, 634  
“equal pay for equal work”  
right, 372  
and ethnic minority  
socialization, 2862  
and ethnic status  
incongruence, 3051  
fertility decline study, 626  
fertility transitions, 1007,  
2178, 2180  
foreign-controlled pharmaceutical  
companies, 1827  
“frontier mentality,” 1933  
fundamentalism in, 2361,  
2368–2373  
and globalization, 140  
government publications, 1613  
and governmental division of  
power, 1954  
health care industry, 1818–1829  
hegemonic stability theory  
and, 3242  
high infant mortality rate,  
130, 140  
higher education focus, 1180  
homicide rate, 68  
immigrants and population  
changes, 636  
immigration waves, 122–123  
income distribution, 130, 140,  
142, 1278–1290, 2705–  
2706, 3048  
indirect and direct discrimination  
in, 689–692  
individualism characterizing, 142  
industrialization and, 141, 143  
infant and child mortality, 633  
and intergenerational mobility,  
2712, 2713–2714  
Iranian sociologists’ presence in,  
1868–1869, 1870, 1873  
Islamic presence in, 2380–2381,  
2950–2951  
Jewish affiliations, 2377

- Jewish population's success in, 2332, 2333
- juvenile violence, 1484-1485, 1487-1489
- labor movements, 143, 1528, 1529-1531, 1532-1534
- labor-force participation trends, 1524, 3262
- and language issues, 2908-2909
- lawyers in, 468-471, 477-478
- legal system, 473-474, 476, 478-480
- legislation of morality, 1575-1577, 1579-1580
- liberalism/conservatism, 1596-1599
- life expectancy, 114, 196, 199, 1628-1631, 2180
- life expectancy ranking, 1631
- life histories and narratives, 1633, 1634, 1635
- life tables, 612, 614, 615, 1630
- life-cycle patterns, 1626
- longitudinal research, 1684
- long-term care and care facilities, 1653, 1655, 1656, 1659, 1661, 1664-1671
- long-term care funding, 1658-1659, 1663
- major institutions, 142-146
- marginal employment, 1719-1725
- marriage and divorce rates, 112, 125-126, 140, 700-706, 1738, 1741-1749
- See also subhead* remarriage *below*
- marriage rates by regions and states, 1749-1750
- Marxist sociology and, 1754
- mass media research, 1761-1766
- and mass society theory, 1773-1774
- medical sociology funding sources, 1814
- Mexican relations, 1856-1857, 1858, 1859, 1860, 1936, 1937
- Middle Eastern migrants to, 1865-1866, 2380-2381, 2950-2951
- military establishment, 144
- military sociology, 1875-1882
- and Moral Majority campaign, 462, 1580, 2371, 2717, 2719, 2723
- national surveys, 578
- and nationalist self-determination policy, 1945
- new religious movements, 2366
- occupation and status attainment, 2785-2786, 3045-3046, 3265
- occupational prestige scores (1964-1989), 1999
- organizational demographics, 395, 396
- organizational restructuring, 2011-2012
- organized crime in, 2018-2021
- overlapping identities in, 1939-1940
- participatory research, 2040
- peace mediation by, 2048
- penal policies, 2054-2057
- pluralism vs. political elite debate, 2624-2625
- police force development, 2111-2113
- political correctness in, 2139-2142
- political organizations in, 2148-2149
- political party system in, 2154, 2164
- and popular culture, 2168-2169, 2170-2172
- population, 2180
- pornography in, 2184-2188
- as postindustrial, 2197, 2199
- poverty in, 2213-2215, 2715
- power elite consolidation in, 2624
- pragmatism development in, 2217
- probation and parole in, 2252-2259
- as pronatalist, 2034
- protest movements, 2264-2271
- publishing industry, 1648
- quality of life, 2300-2301
- racial categories, 2332
- and rape, 2576-2577, 2579-2580
- religious organizations in, 145-146, 2376-2381, 2485
- religious political influence in, 2361-2362
- remarriage in, 112-113, 2387-2393
- retirement practices, 2402-2407
- romantic love complex in, 1698-1699
- rural sociology and, 2426-2429, 2430-2432
- and science-based technologies, 2461
- secondary data analysis and data archives, 2473-2481
- secularization and, 2484
- segregation and desegregation in, 2491-2499
- sexual behavior in, 2537-2548, 2549-2561
- sexual harassment issue in, 1880-1881
- smoking deaths, 1639-1640
- social problems, 2760-2764
- Social Science Data Archives, 575, 576, 579-580
- Social Security system, 2795-2799, 2800, 2802, 2803-2805
- social stratification, 142-143, 2815-2816
- social surveys, 577-578
- socialization agents, 2858, 2862
- sociomoral politicization in, 2361-2362
- sport sociology in, 2987
- and status incongruence, 3054
- suburbanization and, 311, 3070-3076
- suicide rates, 3078
- time use research, 3164
- and tourism, 3167, 3169
- and transnational corporations, 3174-3175, 3176
- unemployment in, 3263
- Uniform Commercial Code, 474, 476
- urban land use in, 311
- and urban system, 3194-3196
- and urban underclass, 513, 3198-3200
- utopian designs and, 2849, 3203-3204
- voluntary associations, 3227-3231

- voting behavior research, 3233–3238
- welfare system beliefs, 2705–2706
- and white-collar crime, 3245–3254
- widowed persons in, 3256–3257
- and woman suffrage, 703
- See also* American family; *specific aspects of society*
- American Sociological Association and other sociological associations, **148–157**
- aligned associations, 153
- American Sociological Society, 1422, 1424
- applied sociology programs, 155–156
- and *Code of Ethics*, 836–840
- computer access survey, 406
- Culture Section, 562, 1646
- diversity commitment, 156
- electronic access to journals, 1606
- founding of, 1422–1423
- goals of, 148
- health-policy analysis, 1159
- interdisciplinary ties, 153
- key governance changes, 151–152
- and literary sociology, 1646
- Mathematical Sociology Section, 1791
- Medical Sociology Section, 1158, 1813, 1814
- membership trends, 148–149
- Methodology Section, 3034
- and Middle Eastern sociology, 1869
- Peace and War Section, 1876
- political pressures and activism, 153–155, 156
- publications and programs, 149–150, 1646
- See also* *American Sociological Review*
- Rational Choice Section, 2375
- regional and state associations, 152, 153, 154
- Section on Microcomputing, 407
- sectors of, 1426, 2913
- social survey questionnaires, 578
- and Society for the Study of Social Problems formation, 2759
- sociology of art program, 172
- and Soviet sociology, 2980
- Teaching Services Program, 150
- Web sites, 413, 1606
- Yinger address on counterculture to, 460
- See also* International associations in sociology; International Sociological Association
- American Sociological Review*, 153, 156, 1528, 1606
- “Concepts of Culture and Social Systems,” 565
- paper on coalition formation theories, 330
- paper on suicide, 3080
- papers on formal models, 2028
- papers on Middle East, 1864, 1870
- papers on social incongruence, 3049
- papers on sociology of literature, 1646
- on research funding sources, 2398, 2399
- American Sociological Society. *See* American Sociological Association
- American Sociologist* (journal), 2042
- American Soldier, The* (Stouffer), 1876, 1881
- American Statistician, The* (journal), 407, 409
- American Statistics Index*, 1613
- American Student Union, 3069
- American Voter, The* (Campbell et al.), 3235
- Americans for Better Care of the Dying, 588
- Americans with Disability Act, 1133
- America’s Children: Key National Indicators of Well-Being* (journal), 2685
- Amin, Idi, 2133
- Amin, Samir, 641, 1706
- Aminzade, Ron, 1707
- Amish, 89, 460, 461
- Ammerman, Nancy, 2378
- Amnesty International, 2723
- AMOS (computer software), 1914
- Amphetamines, 713
- Amplification theory of criminalization, 525
- Amsterdam Treaty (1997), 1935
- Amsterdamska, Olga, 2459
- Anal sex, 2567, 2581
- Analysis, levels of. *See* Levels of analysis
- “Analysis of Propaganda: A Clinical Summary” (Lee), 325
- Analysis of variance and covariance, **157–164**
- adjusting for covariates, 161–162
- applicability, 163–164
- attribution theory and, 192–193
- basic concepts and procedures, 158–160
- for bivariate relationship, 661
- categorical and limited dependent variables, 3037–3038
- covariance structure models, 3037
- decomposing sums of squares, 159–160
- definition of variance, 158, 659–660
- experimental design and analysis, 157–158
- factor analysis and, 905–921, 3036
- general linear model, 162–163
- interrupted time series, 1691–1692
- less than normally distributed variables, 1796–1800
- log-linear model, 3116
- longitudinal research and, 1685–1686, 1689, 1691–1692
- in mass media research, 1764
- measures of association and, 1804–1812, 1966
- multivariate models, 2028
- nonparametric statistics and, 1962–1963
- regression with dummy variables, 162–163
- retrospective data collection and, 1685–1686

- sample selection bias and, 2437–2444
- scientific explanation and, 2467–2468
- specification of variables and, 1803
- statistical graphics and, 3015, 3018
- structural equation modeling (SEM) and, 1922
- summary measures, 160–161
- tabular, 3107–3126
- and typologies, 3181–3182
- validity generalization and, 3210–3211
- See also* Covariance; Variables
- Analysis of Variance, The* (Scheffé), 3035
- Analytic induction, 2297
- Analytic metatheory, 1852, 1853–1854
- Analytical properties of collectives, 1591
- Analytical psychology, 1714
- Ancestor reverence, 1514–1515
- Anchoring phenomenon, 594
- Ancient Law* (Maine), 1545–1546
- Andaman islanders, 3079
- Andersen, Ronald, 1152, 1814
- Anderson, Benedict, 2978
- Anderson, Bo, 2673, 2702
- Anderson, C. A., 73
- Anderson, Elijah, 244
- Anderson, Gerard F., 374, 376
- Anderson, Irina, 193
- Anderson, K. B., 73
- Anderson, Malcolm, 1932, 1934, 1936
- Anderson, Nels, 1203
- Anderson, T. W., 3035
- Andrews, Bernice, 196
- Andrews, Frank M., 2300, 2303, 2683–2684
- Androgyny, 999–1000, 1002
- Aneshensel, C., 3057
- Anfossi, Anna, 1470
- Anger
- social norms for, 2528
- See also* Aggression
- Anglican Church, 701
- Angrosino, Michael, 1636
- Angyll, Andreas, 2084
- Ani, Marimba, 54
- Animal Farm* (Orwell), 2139
- Animal rights movement, 2722, 3230
- Animism, 65
- Anisogamy, 2884–2886
- Annals of Epidemiology* (Roueché), 814
- “Année School” (Durkheim group), 1032
- Année Sociologique, L’* (journal), 1024, 1025, 2917–2918
- “Anniversary effect” (death and dying), 584
- Annual Review of Sociology* (journal), 1606, 2762
- Annual Time Series Statistics for the United States, 2477
- Annual Vital Health Statistics Report* (HEW Department), 497
- Annulment, 701, 947
- Anomaly, definition of, 2025
- Anomie, **164–168**
- alienation and, 100
- broadened meaning of, 165
- criminological theory of, 532, 633
- critical theory and, 544
- death and dying and, 581
- division of labor and, 698
- as Durkheim concept, 164–165, 533, 581, 698, 1493, 1772
- institutional theory and, 503
- juvenile delinquency theory and, 166, 1491, 1493–1494
- mass society concept and, 1772
- rising expectations and, 1491
- social psychological conceptions of, 166–167, 1024
- strain theory and, 166
- suicide and, 165, 3079
- ANOVA. *See* Analysis of variance and covariance
- Ansari-Bradley type tests, 1960
- Anscombe’s quartet, 3011
- Anthony, Susan B., 989
- Anthropological Literature*, 1611
- Anthropology
- and age-set societies, 1624
- applied, 2892
- boundary maintenance concept, 1931
- and courtship study, 483
- and cross-cultural analysis, 547–548, 550, 2893
- cultural approaches, 563, 564, 567
- cultural diffusion theories, 675–676
- culture and personality studies, 2080–2081
- definition of, 2888
- “emic-etic” frame, 550, 564, 2091–2092, 2889
- and ethnography, 852, 853
- and feminist theory, 990
- and functionalism, 1030
- on generative religious movements, 2367
- history of field, 2888–2892
- and Latin American studies, 1537
- legal systems comparisons, 1549–1550
- linguistics approach, 2894–2895
- methods, 2892–2893
- and Mexican studies, 1858–1859, 1861
- money theories, 1890–1891
- Polish sociology and, 2117–2120
- rape explanations, 2579–2580
- sexual behavior in children research, 2551
- and social exchange theory, 2670
- social network studies, 2727, 2729
- socialization concept, 2855
- sociocultural. *See* Sociocultural anthropology
- values concept, 3212
- See also* Ethnography; Ethnology; Ethnomethodology
- Antibiotics, 677
- Antidepressants, 654, 717
- Anti-drug programs. *See* Drug abuse, prevention and treatment
- Antigua, 2600, 3079
- Antioch College, “Foundations of Clinical Sociology” course, 325
- Anti-Semitism, 540

- Nazi policies, 1066, 1067, 1070, 1384, 2332
- Anti-Slavery International, 2603–2604, 2607
- Antisocial behavior. *See* Deviance theories
- Anti-toxics movements, 789–790
- Antitrust actions, 444
- Antonovsky, A., 689
- Anxiety
- as childhood sexual abuse sequel, 290
  - depression with, 655
  - and life course, 1839
  - women's employment and, 1838
- Aoi, Kazuo, 1480
- APA. *See* American Psychiatric Association; American Psychological Association
- Apartheid, 62, 1940, 2047, 2146
- abolishment of, 2725
  - and caste system, 250
- Apathy, 103, 2627
- Appeals court, 471–472, 476
- Applebaum, Barbara, 1904
- Applied behavior analysis, 215–216
- Applied science, 2460–2461
- Applied sociology, **168–171**
- American Sociological Association conference, 156
  - associations, 155
  - basic sociology vs., 168, 2845
  - boundaries of, 169–170
  - clients of, 169–170
  - internships, 171
  - rural sociology and, 2427
  - and sociological practice, 326, 2845
- Apprentice system, 27, 697
- “Approach to Clinical Sociology, A” (McDonagh), 326
- Apuzzo, Gian Matteo, 2231
- Aquino, Corazon, 2131
- Arab League, 1944
- Arab Magreb Union, 1944
- Arab studies. *See* Middle Eastern studies; Islamic societies; Sociology of Islam
- Araji, S., 2582
- Arbetslivsinstitutet (Sweden), 2451
- Arbitrators, 465
- Arbuthnot, John, 1957
- Arcadian tradition, 2426
- Arch, Joan, 1815
- Archaeology, 2889, 2891, 2893
- Archetypes, 1714
- Architecture, 2259
- postmodern, 2207
- Archives. *See* Library resources and services for sociology
- Archives de Sociologie des Religions*, 1026
- Ardigo, Achille, 1468, 1469, 1470
- Ardigo, Roberto, 1464, 1465
- Arditi, Jorge, 2958, 2959
- Arendt, Hannah, 1773
- Argentina
- demographic characteristics, 1535, 1536
  - economic liberalization, 1539, 1540, 1541
  - fertility decline, 627
  - gross national product, 1535
  - health-care system, 381
  - labor movement, 1532
  - political and social conditions, 1536, 1537
- Argonauts of the Western Pacific* (Malinowski), 2890
- Ariès, Philippe, 4, 2090, 2861
- Ariga, Nagao, 1477
- ARIMA (autoregressive moving average) models, 2679, 3143–3153
- Aristocracy, 415–416
- Aristotle, 54, 1545, 1587, 2882
- on communitarians, 355
  - on elites, 2623
  - on human character differences, 1717
  - on leaders, 1564
  - on passions, 2519
  - on personality, 2086
  - on social justice, 2697, 2698
- Arjomand, Said A., 1870
- Armed forces. *See* Military sociology; War
- Armed Forces and Society* (journal), 1876
- Armed Forces Qualifying Test (AFQT), 1877
- Armenia, 2982
- genocide, 1070, 1384
- Armingier, Gerhard, 1694
- Armstrong, Barbara, 2403, 2406
- Army Aberdeen Proving Grounds scandal (1996), 1880
- Army Research Institute, 1876
- Arnett, Jeffrey Jensen, 29
- Aron, Raymond, 1025, 1027, 1073, 1356
- Aronowitz, Stanley, 542
- Aronson, E., 339, 340
- Arranged marriages, 1698, 1775
- Arrestees Drug Abuse Monitoring program, 713
- Arrighi, Giovanni, 1264
- Arrow, Kenneth, 735, 2335, 2339, 2920
- Arrowood, A. John, 2650, 2651
- Arson, Uniform Crime Reports definition, 492
- Art and society, **171–174**
- African art, 64
  - African-American art, 64
  - and culture, 567
  - emotional depression incidence, 655–656
  - marginal categories, 173
  - mass vs. elite culture issue, 173
  - postmodernity and, 2207
  - suicide depictions, 3079
- Art Worlds* (Becker), 1647
- Artificial intelligence, 410, 1234
- Artisans, 3262
- Aruga, Kizaemon, 1478
- ASA. *See* American Sociological Association and other sociological associations
- ASA Teaching Services Program, 150
- Asabiyya* concept (social solidarity), 2941–2942
- Asai, 356
- Asante, Molefi Kete, 54
- ASCDR (age-standardized crude death rate), 610–611
- Asch, Solomon, 857



- conformity experiment, 401–402, 403, 404, 2094, 2616  
 marijuana use experiment, 523
- Ascher, William, 2678
- ASEAN (Association of Southeast Asian Nations), 2975–2976
- Ash, J., 3167
- Ashanti, 64, 65
- Ashraf, Ahmad, 1870
- Asia  
     abortion legality, 2240  
     AIDS/HIV demographics, 2591–2592  
     authoritarian  
         communitarianism, 356  
     child labor, 3262  
     Demographic and Health Surveys, 633  
     demographic transitions, 627  
     dependency theory and, 642  
     economic crisis, 729–730  
     fertility determinants, 1006  
     fertility rate decline, 220, 1008, 2178  
     governmental and political corruption, 2130–2131  
     labor force composition, 3262  
     labor movements, 1532  
     life expectancy, 623  
     preference for male babies, 627  
     racial traits, 2331  
     sex trade, 2607  
     slavery and involuntary servitude, 2602–2603, 2604–2608  
     Southeast Asia studies, 2974–2978  
     *See also specific countries*
- Asian-American studies, **174–184**  
     as academic discipline, 179–182  
     AIDS/HIV risk, 2590  
     assimilation model, 177, 178–179, 181  
     attribution and, 194  
     demographic characteristics, 175–177  
     discrimination, 58, 174, 175  
     discrimination consequences, 175–177  
     divorce potential, 126  
     household structures, 127  
     immigrant newcomers, 180  
     immigration and restrictions, 123, 143, 174–177  
     interracial marriage  
         demographics, 124, 1776  
     occupational clustering, 181–182  
     research areas, 177–182  
     single-gender communities, 123, 176  
     suburbanization, 3074, 3075  
     *See also* Chinese Americans; Japanese Americans; Korean Americans; Vietnamese Americans
- Asian-American Studies Center (University of California), 179
- Asksun Empire, 2999
- Asplund, Johan, 2452
- Asquith, H. H., 225
- Assault. *See* Aggression; Aggravated assault; Sexual violence and exploitation; Violence
- Assembly line, 697, 699
- Assertiveness, aggression vs., 68
- Asset and Health Dynamics Among the Oldest Old, 344
- Assimilation  
     in American society, 143, 178  
     Asian-American studies, 177, 178–179, 181  
     ethnicity and, 178–179, 842–844  
     of homosexual subculture, 2571  
     and immigrant  
         suburbanization, 3075  
     politically correct diversity vs., 2139–2140  
     power-conflict analysis vs., 53–54  
     religions and, 3288  
     subcultures resisting, 459  
     *See also* Multiculturalism
- Assisted living facilities, 1663, 1826
- Assisted suicide, 585, 586–587, 3083, 3084–3086
- Associated-dependent development, 642
- Association. *See* Measures of association
- Association for Health Services Research, 1158
- Association for Humanist Sociology, 326, 1247
- Association Internationale des Sociologues de Langue Française, 328
- Association of American University Graduates, 1868
- Association of Russia (Russian Sociological Society), 2982
- Assumption drag, 2678
- Assumption of homoscedasticity, 449
- Assumption of linearity, 447
- Assumption of rectilinearity, 447
- Assyrian Empire, 1069, 2998
- Asthma, 1640
- Aston University, 232
- Asymptotic distribution theory, 1957–1958
- Atal, Yogesh, 1292
- Ataturk, 2411, 2485
- Atheism, public tolerance for, 316, 317
- Athletics. *See* Sport
- Atkins, John R., 1513
- Atkinson, A. B., 2501
- Atkinson, Robert, 1636
- Atkinson Index, 2501
- Atlantic slave trade. *See* African slave trade
- ATLAS/*ti* (computer software), 420
- Atomist materialism, 1780–1781
- Attachment  
     and social belonging, 2631, 2632  
     as social control, 2658  
     types of, 2632
- Attachment theory, 2066–2067, 2068, 2090
- Attainment model, 1692, 2817
- Attention deficit hyperactivity disorder, 70
- Attenuated correlations, 1909
- Attica prison (New York State), 2054
- Attitude Interest Analysis Test, 999
- Attitudes, **184–192**  
     of aggression-prone people, 72  
     balance theory and, 187, 335  
     behavior relationship, 189  
     belief and, 184  
     communication-persuasion paradigm and, 188  
     congruency theory and, 336

- construct validity and, 190  
 definition of, 184  
 dissonance theory and, 187–188, 339  
 as employment evaluation influence, 1994  
 toward extramarital sex, 2542  
 toward premarital sex, 2568  
 factors in changing, 188–189  
 formation of, 184–185  
 functions of, 185  
 as indicators, 190–191  
 measurement of, 185–186, 190  
 organization of, 186–188  
 persuasion and, 2094–2098  
 phenomenological theory on, 2101–2102  
 pragmatist study of, 2220  
 reasoned action theory and, 189  
 role theory and, 2418  
 sentiments and, 2525  
 on singlehood, 107  
 situational constraints of, 189  
 social institutions and, 185  
 source of, 189  
 stability of, 188–189  
 stereotypes and, 184, 185, 189  
 survey research on, 3087–3094  
 values differentiated from, 2828–2829  
 voting behavior research on, 3233–3234  
 white racial, 2245–2246  
*See also* Compliance and conformity; Liberalism/conservatism; Prejudice; Public opinion; Social values and norms; Stereotypes
- Attorney General's Commission on Obscenity and Pornography, 2184, 2185
- Attractiveness, mate selection and, 1777
- Attribution theory, **192–200**  
 context of, 195–196  
 covariational model of, 193  
 definition of, 192  
 fundamental attribution error, 194, 2751  
 intentional/unintentional behavior and, 192, 194  
 intergroup, 197  
 interpersonal, 196–197, 2751–2752  
 major theories of, 192–196  
 measurement and methodology, 194–195  
 and prejudice, 2244  
 on self-concept development, 2507, 2856  
 societal, 197–198  
 sociological significance of, 196–198  
 stereotypes and, 197–198, 2244
- Attrition, in panel studies, 1691
- Aubert, Vilhelm, 2450  
 “Audience effect,” 2615
- Audio information-capturing technologies, 418–419
- Augustine of Hippo, Saint, 1507
- Augustus, John, 2253
- Austin, John, 1428
- Australia  
 aboriginal art, 173  
 aborigines, 3079  
 common law system, 465, 471  
 educational status attainment, 2784–2785  
 fertility transitions, 1007  
 foreign-controlled pharmaceutical companies, 1827  
 indigenous genocide, 1069  
 long-term care and care facilities, 1661  
 Muslim minorities, 2950, 2951  
 racial conflict, 321  
 Social Science Data Archive, 576, 2477  
 social security system, 2797, 2800  
 woman suffrage, 703
- Australian Social Trends* (report), 2685
- Austria  
 civil law system, 477  
 ethnic status incongruence, 3051  
 high suicide rate, 3079, 3082  
 labor market structure, 1987, 1988  
 political party system, 2159
- Social Science Data Archive, 576  
 social security system, 2797
- Austro-Hungarian Empire, 2362, 2998, 3001
- Ausubel, D. P., 2088, 2090
- Authoritarian personality  
 Adorno theory, 317, 334, 540  
 cognitive consistency theory, 334  
 object relations theory, 2064
- Authoritarianism  
 communitarian, 356  
 dictatorship, 2163, 2356, 3002  
 kingdoms and empires, 2999–3000  
 and persuasion, 2096  
 and protest movement violence, 2270  
 as reaction to nationalist movements, 3002  
 structural view of, 2163  
 and systems theory, 3103  
*See also* Autocracy; Patriarchy
- Authority  
 compliance with, 404  
 conflict theory on, 415–416  
 cultural values variation on, 3224  
 hierarchical division of labor and, 696  
 as perception of just power, 2165, 2997  
 political crime and, 2143–2146  
 revolution and, 415–416  
 sentiments and, 2519–2521  
 as social capital element, 2638–2639  
 and social exchange, 2673  
 Weber's typologies, 229–230, 2519–2520
- Autobiographies, 1636–1637  
*See also* Autoethnography
- Autobiography of a Drug Addict, The* (Hughes), 243
- Autocracy  
 democracy vs., 605, 606  
 leadership effectiveness in, 1565  
 and war initiation, 3244
- Autoethnography, 852, 1636–1637, 2291  
 and case studies, 245

- reflexive, 2293
- Automation, 239, 699, 3266
- Automobiles
- accident risk, 2877
  - assembly line production, 696, 699
  - driver education classes, 677
  - and drunk driving, 93, 164–165, 1640, 1641, 2877
  - and drunk-driving reform movements, 2722, 2725, 2877
  - labor unions and, 1531
  - Mexican industry, 1860
  - theft, 492, 497–498, 499
- Autonomy. *See* Personal autonomy
- Autopoiesis, 2088
- legal system, 1548, 1557–1559
- Autoregressive moving average (ARIMA) models, 2679, 3143–3153
- Auty, Richard M., 644
- Availability, decision processing, 593
- Average. *See* Mean (statistical)
- Aversive consequences, dissonance from, 339–340
- Aversive racism, 2245
- Aversive stimulation, aggression and, 73
- Avoidance, as childhood sexual abuse sequel, 290–291
- Avolio, B., 1566
- Axelrod, Robert M., 596, 1791
- Axiomatic theory, 1787–1788
- Axis Rule in Occupied Europe* (Lemkin), 1066
- Azerbaijan, 2362
- nationalism, 1871, 1945
- Aztec Empire, 2999
- B**
- Babeuf, Gracchus, 1066
- baby boom
- alienation and, 1, 100
  - cohort effects, 80, 100, 345, 1625, 2678
  - demographic assumption drag and, 2678
  - divorce rate and, 1747
  - fertility rate anomaly and, 1525
  - marriage rate decline and, 1741
  - social security funding and, 2803, 2804
  - and social structural responses, 3063
- Babylon, destruction of, 1069
- Babylonian Empire, 2998, 2999
- BAC (blood alcohol content), 93, 1641
- “Back to Africa” movements, 66
- Bacon, Francis, 2456
- Baechler, Jean, 3077–3078
- Baglioni, Guido, 1467, 1469
- Bahaism, 3287–3288
- Bahr, Howard M., 364, 367
- Bahrain, 1866, 1867
- Bahrtdt, Hans Paul, 1076, 1078
- Bahujan Samaj (Dalit political party), 252
- Bailey, Kenneth D., 3187
- Bainbridge, William Sims, 2375, 2381
- Bakan, David, 2058
- Baker, C. E., 269
- Baker, David, 1515, 2929
- Baker, Patrick, 1212
- Baker, Paul, 583
- Bakhtin, Mikhail, 1758
- BAL (blood alcohol level), 93
- Balance of power systems, 332–333, 3242
- Balance theory, 335–336
- attitudes and, 187
  - in family and kinship, 1511–1512
  - in history of the state, 2998
  - and social networks, 2731
- Balanced Budget Act of 1997, 1144, 1146
- Balbo, Laura, 1472
- Baldwin, James Mark, 2089
- Bales, Kevin, 228
- Bales, Robert F., 432, 696, 1014, 1565, 2666
- and observation systems, 1974, 1975, 1976, 1979, 1980
  - and small groups research, 2613, 2614, 2620
- Balkans, 1934, 1948
- See also* Eastern Europe; *specific countries*
- Balkwell, James W., 1790
- Ball-Rokeach, Sandra J., 3213, 3214
- Baltes, Margaret M., 1617
- Baltes, Paul B., 1617
- Baltic states, 1934, 1945, 2362
- post-communist transition in, 2136
  - sociology as discipline, 2982
- Balzer, Wolfgang, 2296
- Bandura, Albert, 2084
- social learning theory, 70, 1716–1717, 1895, 2858
- Banfield, Edward, 1467
- Bangladesh, 1941
- corruption, 2132
  - fertility decline, 628, 2179
  - slavery and slave-like practices, 2604, 3262
  - sociodemographic profile, 2938
  - time use research, 3161, 3162
- Banishment, 515, 527
- Bankruptcy and credit, **201–208**
- causes of, 204–205
  - consumer debt and, 206–207
  - international credit issues, 207
  - Latin American debt crisis (1980s), 1540
  - life cycle and, 1625
  - myths about, 205–206
  - regional variations, 203–204
  - trends in, 203–204
  - types of, 202–203
  - voluntary/involuntary, 202
  - white-collar crime and, 3246, 3250, 3253
- Banks, Arthur, 2917
- Banks, J. A., 225
- Banks, Olive, 225
- Banton, Michael, 225, 2114, 2415, 2417
- Baptists, 95
- Bar graphs, 659, 3008–3009, 3010
- Baran, Paul, 641, 1087
- Barbados, 2600, 3263
- Barbalet, Jack, 786–787
- Barbano, Filippo, 1469, 1470
- Barbe, Carlos, 1472
- Barbieri Masini, Eleonora, 1037, 1038, 1040
- Barbituates, 713

INDEX

- 
- Barbur, B. R., 269
- Bardes, Barbara, 2280–2281
- Bargaining  
     distribution of winnings and, 330  
     in group decision making, 597
- Barjaba, Kosta, 1472
- Barkey, Karen, 1872
- Barnby, Goodwyn, 355
- Barnard, Chester, 1080
- Barnes, J. A., 2729
- Barnett, James, 172–173
- Baron, Lawrence, 118, 503, 505
- Barotse, 1549
- Barreda, Gabino, 1857
- Barrile, Leo, 3252
- Barry, Marion, 2126
- Barth, F., 1931
- Barthel, D. W., 1653
- Barthes, Roland, 2168
- Bartholomew, David J., 2668, 3036
- Barton, Allen, 683, 3182
- Base rate information, 591
- Basel Convention on the Control  
     of Transboundary Movements of  
     Hazardous Waste and their  
     Disposal (1989), 794
- Bashkar, Roy, 823
- Basques, 3001
- BASS (Belgium data archive), 576
- Bass, B. M., 1566, 1567
- Batelle Institute, 1041
- Bates, Frederick L., 2416
- Bateson, Gregory, 824
- Batista, Fulgencio, 2134
- Batson, C. Daniel, 118
- “Battered Child Syndrome, The”  
     (Kempe et al.), 288
- Battered women  
     case studies of, 247  
     and family violence, 981
- Battisti, Francesco, 328
- Baudrillard, Jean, 1307, 2199, 2200,  
     2206, 2207
- Bauer, Raymond A., 2299, 2682
- Baum, Rainer C., 424
- Bauman, Zygmunt, 1533, 2119,  
     2207, 2983
- Baumeister, R. F., 2513, 2514,  
     2515, 2516
- Baumgartner, F., 2085
- Baumol, William J., 2668
- Bavelas, Alex, 1034
- Bayesian equilibrium, 330
- Bayesian statistical model, 2249,  
     3039
- Bayes’s theorem, 590–591, 598
- Bayley, David, 2114
- Bayma, Todd, 1648
- BCS. *See* British Crime Survey
- BDI. *See* Beck Depression Inventory
- BDSP (Grenoble, France), 576
- Beach, Frank, 2570
- Beattie, Geoffrey, 193
- Beauvoir, Simone de, 988, 989, 990
- Beccaria, Cesare, 528
- Beck, Aaron T., 651, 653
- Beck, E. M., 2481
- Beck Depression Inventory, 654
- Becker, David G., 643
- Becker, G. S., 181, 689
- Becker, Gary, 722, 735, 940, 1009
- Becker, Howard S., 173, 243, 244,  
     535, 536, 669, 853, 1647,  
     1925, 2297
- Becker, Marshall, 1814
- Beckford, James, 2378
- Beckman, J. H., 2301, 2305
- Beer, 92, 95
- Beery, Richard G., 2859
- Beethoven, Ludwig van, 1925, 1926
- Beezley, Dana A., 2557
- Behavior, deviant. *See* Deviance  
     theories
- Behaviorism, **208–217**  
     altruism theory, 115  
     applications in education,  
         215–216  
     choice and preference in,  
         212–214  
     concurrent schedules of  
         reinforcement, 212–213  
     conformity basis, 2616  
     depression theories and, 650–651  
     deviance theories and, 666–667  
     dissonance phenomena and, 339  
     evolutionary theory and,  
         2880–2881  
     and German sociology, 1080
- leadership and, 1565–1569  
     matching theory in, 214  
     operant reinforcement in,  
         209–210, 214, 1716, 2085,  
         2670  
     quantitative law of effect  
         and, 208, 214  
     response bias, 213  
     social exchange theory and, 2670  
     social interaction theory and,  
         2085  
     social psychology research and,  
         2769–2770  
     social values and norms and,  
         2838–2839  
     symbolic interactionism and,  
         2856  
     terrorism analysis, 3141  
     variable-ratio schedules in, 211  
     voluntary action and, 209
- Behaviorist psychology, 722, 1015
- Beijing Conference on Women  
     (1995), 932, 1768
- Bekhterev, V. M., 2979
- Belarus, 938, 2982
- Belgium  
     African colonization by, 60  
     and ethnic status  
         incongruence, 3051  
     fertility transition, 626  
     legal system, 471, 477  
     multilingualism, 2909  
     political and governmental  
         corruption, 2130  
     political party system, 2159  
     Social Science Data Archive, 576  
     social security system, 2800  
     social surveys, 577  
     transnational corporations, 3175
- Belief  
     and conformity, 2616  
     as generalized in collective  
         behavior, 353  
     and perception of justice, 2697  
     as social control, 2658  
     tensions from feelings  
         inconsistent with, 337  
     values and, 3213  
     *See also* Social values and norms
-

INDEX

- Bell, Daniel, 847, 1039, 1345, 1348, 1357, 1774, 2484, 2818  
 human ecology and  
 environmental analysis, 1218, 1220, 1223, 1224  
 as neoconservative, 1601  
 postindustrial theory, 2196, 2197, 2198, 2200, 2205
- Bell, Derrick, 58
- Bell, Inge Powell, 3069
- Bell, Wendell, 1037, 1038, 2501, 2502, 2677
- Bell and Howell Information and Learning, 1607, 1608
- Bell Curve, The* (Murray and Herrnstein), 2330
- Bell Laboratories, 3006
- Bell Telephone System, 444
- Bellah, Robert, 356, 360, 1197, 2080, 2484, 2713, 2918
- Bellamy, Edward, 3203
- Bellasi, Pietro, 1468
- Bell-curved distribution, 661, 2330, 2869
- Belloc, Hilaire, 225
- Belohradsky, Vaclav, 1472
- Belonging. *See* Social belonging; Territorial belonging
- Belsky, Jay, 1737
- Bem, Daryl J., 339, 2750, 3222
- Bem Sex Role Inventory, 1000
- Ben-David, Joseph, 2456
- Bender, Thomas, 362, 367
- Bendix, Reinhard, 383, 384, 387, 1704, 1886
- Bendor, S., 1511
- Benedict, Ruth, 178, 563, 2855, 2883, 2890
- Benedikt, Moritz, 2087
- Benevolence, 3216
- Bengtson, Vern, 1387, 1388, 1389, 2707
- Benin, 2604
- Benjamin, Lois, 57
- Benjamin, Walter, 539, 541, 542
- Benjamin Rose Institute (Cleveland), 2943
- Bennett, W. Lance, 1766
- Bennett, William, 1650
- Benson, Herman, 1533
- Benson, Michael, 2660
- Bentham, Jeremy, 355, 528, 1599, 1600, 2087, 2963
- Bentler, P. M., 2513
- Bereavement, 582, 584, 649, 650  
 and widowhood, 3255, 3257, 3259
- Bérégovoy, Pierre, 2129
- Berezin, Mabel, 569, 1647
- Berger, Bennett, 460
- Berger, Brigitte, 1886
- Berger, C. Q., 3155–3156
- Berger, Gaston, 1038
- Berger, Joseph, 880, 1790, 2029, 2702, 2704
- Berger, Mark C., 2303
- Berger, Monroe, 1868
- Berger, Paul, 1886
- Berger, Peter L., 226, 606, 2483, 2487, 2756, 2957–2958, 3285, 3289
- Berger, Raymond, 2943
- Berger, Ronald, 1636
- Berghe, Pierre van den, *See* Van den Berghe, Pierre
- Bergin, Allen E., 2189–2190
- Bergson, Henri, 1423
- Bergsträsser, Arnold, 1075
- Berk, Richard, 349, 350, 2439, 3038
- Berkeley, George, 2217
- Berkowitz, Leonard, 70, 73, 2700
- Berle, Adolph, 443
- Berleant, Daniel, 421
- Berlin Conference (1884–1885), 60, 1934
- Berlusconi, Silvio, 2129
- Berman, Harold J., 1555
- Bernard, Jessie, 1067, 1727, 1736
- Bernard, Thomas J., 1490–1491, 3247
- Berne, Eric, 2084
- Berners-Lee, Tim, 1445
- Bernstein, Basil, 225, 2894
- Bernstein, Edward, 2846, 2847, 2848
- Bernstein, Richard, 2219
- Berreman, Gerald, 254
- Berscheid, Ellen, 2700
- Bertalanffy, Ludwig von, 1558, 3102
- Bertaux, Daniel, 1635, 2662
- Berthelot, J.-M., 1028
- Bertilsson, Margareta, 2452
- Bertocchi, Graziella, 644
- Bertranou, Fabio M., 381
- Best, Joel, 2764
- Best, Steve, 1757
- Best alternatives to nonagreement (BATNA), 1955
- Beteille, Andre, 1292
- Betteridge, Anne, 1868
- “Between Universal and Native: The Case of Polish Sociology” (Kwasniewicz), 2119
- Bevans, George Esdras, 3155
- Beyond Self-Interest* (Mansbridge), 118
- Beyond the Classics?* (Glock and Hammond), 2373
- Bhagavad Gita* (Hindu scriptures), 3280
- Bhagavata Purana* (Hindu scriptures), 3280, 3281, 3287
- Bharatiya Janata Party (BJP; India), 2486
- Bhopal disaster, 683, 805, 2875, 2877–2878
- Bhutan, 3262
- Bhutto, Benazir, 2132
- BIA. *See* Bureau of Indian Affairs
- Biafra, 1941
- Bias  
 attribution, 194, 196, 197, 2751–2752  
 behavioral reinforcement deviation and, 213  
 invalid and unreliable indicators and, 1909–1910  
 moral development theory and, 1900, 1902–1904  
 panel studies and, 1691  
 predictive, 3211  
 sample selection, 2437–2444  
 in self-report survey voting research, 3232  
*See also* Prejudice; Stereotypes
- Bible  
 communitarian issues in, 355  
 on descent, 1513–1514  
 fundamentalist inerrancy of, 2368, 2369, 2370

INDEX

- and Hebrew religious experience, 3280, 3281, 3282, 3283, 3285  
on marital and filial responsibilities, 1508, 1511, 1513  
suicide instances in, 3079  
*Bibliography of Japanese Sociological Literature in Western Languages*, 1482  
Biddle, Bruce J., 2418  
Bideau, A., 633  
Biderman, Albert D., 2682, 3247  
Biedenkopf, Kurt, 362  
Bielby, William T., 3036  
Bienenstock, Elisa I., 2673–2674  
Big Five (personality traits), 2079–2080, 2085  
Biko, Steve, 2146  
bilateral kinship systems, 1507–1508, 1699–1700  
Bilingual education, 123, 1861, 2140, 2908  
Bill of Rights (U.S.), 270, 315, 359  
Billson, Janet Mancini, 245  
Bimodal distribution, 661  
Binary opposition, 1032, 1033  
Binet, Alfred, 1360, 2330  
Binstock, Robert, 1388  
Binswanger, Ludwig, 2084  
Biocolonization, 1824  
Bioethics, 585, 1824, 3064  
Biographies, 1636–1637  
Biology  
as link to sociology, 1029  
*See also* Evolution: biological, social, cultural;  
Sociobiology, human  
Bion, Wilfred, 1979–1879  
Biopiracy, 1824  
Biostructuralism, 1033  
Biotechnology industry, 1824  
Bipolar disorder, 649, 650, 1838  
Birch, M. C., 3036  
Birdwhistell, Ray, 1978  
Birmingham Centre for Contemporary Cultural Studies (Great Britain), 226  
Birmingham School, 1646  
Birth. *See* Birth and death rates; Childbearing; Fertility determinants; Fertility rate; Pregnancy and pregnancy termination  
Birth and death rates, **217–224**  
age-specific death rate, 223  
age-specific death rates and life expectancy, 1629  
age-specific fertility rates, 192, 193, 218–219  
American Indian death rates and causes, 133, 135  
birth rate calculation, 217–220  
birth rate decline, 110, 122, 125, 2032, 2182  
births to teenagers, 2235  
comparison of crude rates, 610–612  
crossover death rate, 1631  
death rate calculation, 220–224  
death rate for term pregnancy vs. abortion, 2238  
death rates comparison for African Americans and whites, 1631  
death rates from AIDS, 222, 224, 2592  
death rates from illegal abortions, 2241  
death rates from life-style risks, 1639–1642  
death rates from suicide, 3078, 3082  
demographic transitions, 425, 621–623, 633, 2177–2178  
ethnic census data, 259  
historical population overview and, 2175–2177  
and leading causes of death, 222, 224  
in less developed countries, 931  
life expectancy and, 1628, 1629, 2177  
life tables, 611–615, 1629–1630  
maternal deaths, 2236  
mortality modeling, 619  
mortality sex-specific rates, 222  
mortality transitions, 622–623, 624–625, 626  
mortality-fertility transition interrelationship, 628, 629  
natural fertility populations, 2176  
population projection and, 2180–2183  
prehistoric patterns, 2175–2176  
standardization, 2992–2995  
in United States, 2180  
*See also* Demographic transition; Fertility determinants; Fertility rates; Infant and child mortality; Life expectancy  
Birth cohort. *See* Cohort perspectives  
Birth control. *See* Family planning  
Birth defects, 1640  
“Birthday dip,” in dying, 584–585  
Bisexuality. *See* Sexual orientation  
Bishop, Yvonne M. M., 3036  
Bismarck, Otto von, 2797  
Bittner, Egon, 2108–2109, 2114  
Bivariate relationship, 661  
correlation and regression, 457  
hypothetical tables, 3113  
Biya, Paul, 2133  
Bjarnason, Thoroddur, 2452, 2453  
Black, Cyril E., 1947  
Black, D., 2108  
Black, Donald, 332, 1544, 2114  
Black, Duncan, 2920  
Black Elk (Lakota shaman), 3277–3278, 3279, 3281, 3282  
Black English, 2908–2909  
Black feminist theory, 545  
Black market, 2019, 2020, 2021, 2136  
Black Panthers, 460, 2267  
Black power movement, 66–67  
Black studies. *See* African American studies  
Black Vernacular English, 2909  
Black-box modeling, 2679  
Blackness, concepts of, 56, 62  
Blair, Tony, 228, 361  
Blake, Judith, 974–975, 1006, 3214  
Blalock, Hubert M., Jr., 847, 3035, 3036  
Blankenburg, Erhard, 1559  
Blascovich, J., 2512  
“Blaseization” (Simmel concept), 2528  
Blasphemy, 269  
Blau, Judith, 536

INDEX

- Blau, Peter M., 260–261, 262, 536, 2703  
 and *cui bono* criterion, 2004  
 on education and mobility, 260–261, 2713, 2714, 2716, 2927, 2929  
 and landmark division of labor theory, 698–699  
 and model of attainment, 1692, 2782, 2817  
 and model of social structure, 2029, 2825, 2826, 3035–3036  
 social exchange theory, 2670, 2671–2672, 2674, 2731, 3272  
 on social inequality, 2690–2691
- Blau, Zena Smith, 1510
- Blauner, Robert, 54, 582, 646
- Blau-Duncan model, 260–261, 262
- Blended families, 112–113, 126, 2390–2391  
 incest potential, 2583
- Blenker, Margaret, 2943
- Bloch, Marc, 384, 1891
- Block, Fred, 2163, 2198
- Block modeling, 1790
- Blocked exchanges, 727
- Blood alcohol content, 93, 1641
- Blood, R., 696
- Bloom, Martin, 2943
- Blue Cross/Blue Shield, 1825, 1826
- Blue laws, 1560
- Blues music, 1926
- Blum, Lawrence, 118
- Blumberg, Abraham S., 2960–2961
- Blumen, Isadore, 1692, 1789
- Blumer, Herbert, 179, 350, 352, 1254  
 social contagion theory, 679  
 and symbolic interactionism, 2221, 3095, 3096, 3098
- Blumstein, P., 2539, 2540, 2542, 2555
- BMD (Biomedical Computing Programs), 2035
- Board on Natural Disasters, U.S., 686
- Boas, Franz, 563, 675, 2888, 2890, 2892, 2893
- Bobo, Lawrence, 317, 2245
- Boccacio, 2185
- Bochner, Arthur, 248, 1636–1637, 2291–2292
- Bockstaele, Jacques van, 328
- Bockstaele, Maria van, 328
- Bode, N., 3251
- Boden, Deirdre, 2902–2903
- Body language. *See* Nonverbal cues
- Body sensations and types, personality theory of, 1717–1718
- Bogardus, Emery S., 2027–2028
- Bogart, Leo, 2273
- Bohannan, Paul, 1545, 1560
- Bohemian Beats, 460–461
- Bohrnstedt, George W., 3034
- Bohstedt, John, 2270
- Boissevain, Jeremy, 2729
- Bokassa, Jean-Bedel, 2129, 2133–2134
- Bokszanski, Zbigniew, 2119
- Boldrini, Marcello, 1465
- Boli, John, 427–428
- Bolivia  
 demographic characteristics, 1535  
 drug trafficking, 2135  
 economic liberalization, 1539  
 revolution (1952), 2414
- Bolshevik Revolution. *See* Russian Revolution
- Bolte, Karl Martin, 1075
- Bombings. *See* Terrorism
- Bonacich, Phillip, 2672, 2673–2674
- Bonald, Louis de, 1770
- Bonded labor. *See* Slavery and involuntary servitude
- Bonds, social. *See* Social belonging
- Bonger, William, 534
- Bongo, Omar, 2133
- Book of Mormon, 3287
- Book of Virtues* (Bennett), 1650
- Books: The Culture and Commerce of Publishing* (Powell), 1647
- Boolean operators, 386, 1608, 1610, 2296
- Booth, Alan, 1737
- Booth, Charles, 853
- Bootstrapping, 2397, 2449, 2678, 3039
- Borderline personality disorder, 291, 293
- Borders. *See* National border relations
- Borduain, C. M., 76
- Borel, E., 1045
- Borgatta, Edgar, 1213, 1226, 1227, 1228, 1468, 1974–1975, 2085, 3034
- Borkenau, Franz, 1075
- Born-leader concept, 1564
- Bornschiefer, Volker, 643
- Bornstein, R. F., 2063, 2064
- Boserup, Ester, 1708
- Bosnia, 2362, 2947, 3288  
 genocide, 68, 1944, 1946, 2529  
 rape of women and children in, 2579
- Boss, Medard, 2084
- Bossa nova, 1927
- Bossard, James, 1776–1777
- Boston Symphony, 2171
- Boswell, Terry, 645
- Botswana  
 fertility decline, 627–628  
 unemployment, 3263
- Bott, Elizabeth, 696, 2729
- Bottomore, Tom, 1752, 1753, 2027
- Bouchard, Thomas, 2089
- Boudon, Raymond, 2825, 2826
- Bouglé, C., 1024, 1025
- Boulding, Elise, 1039
- Boundaries  
 coalitions and, 333  
 conformity as enforcer of, 400  
 shared organizational maintenance of, 394  
*See also* National border relations
- Bourdieu, Pierre, 172, 227, 268, 416, 823, 2642, 2892, 2983, 2990  
 on class-based differences in taste, 1648–1649  
 cultural capital theory, 2626  
 definition of capital by, 2637  
 Japanese sociology and, 1479  
 on legal systems autonomy, 1548, 1549  
 popular culture studies, 2170–2171  
 and social philosophy, 2757
- Bourgeoisie

- and capitalism, 238  
 Marx's assessment of, 2847, 2848  
 and mass society theory,  
 1772–1773  
 and revolution, 1198, 2410, 2412  
 structuralist theory of state  
 and, 2163
- Bourgeois-Pichat, Jean, 631, 1326
- Bouthoul, Secondo, 1424
- Bouvia, Elizabeth, 3084
- Bovone, Laura, 1473
- Bower, Raymond, 677
- Bowlby, John, 650, 2066–2067,  
 2068, 2090
- Bowler, Anne, 173
- Bowling Green University, 2168
- Box, George E. P., 3036, 3143, 3144
- Box, Steven, 3251
- Box plot, 3012, 3015  
 definition of, 661
- Box-Jenkins model, 3036
- Boyatzis, Richard E., 420
- Boyce, James K., 645
- Boyle, Robert, 2456
- Bozorgmehr, Mehdi, 1872–1873
- Bracero program, 88
- Bracketing, 2100
- Bradshaw, York W., 644
- Braga, Giorgio, 1469
- Brahman, 250
- Brainstorming, 2618
- Brainwashing  
 effects of, 897–898  
 as extreme form of  
 persuasion, 2094
- Braithwaite, John, 530, 1491–1492,  
 3250, 3253
- Brandeis, Louis, 476
- Brando, Marlon, 2185
- Brandt, Willy, 1316
- Brant, Joseph, 136
- Braudel, Fernand, 1935, 2168, 2892,  
 2917–2918
- Brave New World* (Huxley), 1505
- Braverman, Henry, 542
- Brazil  
 clinical sociology, 328  
 demographic characteristics, 1535
- dependency theory and,  
 642–643, 1538
- fertility decline, 627
- gender role changes in, 941
- labor movement, 1532
- political and social conditions,  
 1536, 1537
- political corruption, 2135–2136
- racial continuum in, 2331–2332
- regional stratification structures,  
 2871–2874
- slavery and slave-like practices,  
 2599, 2600, 2601, 2604
- Breaching, 2100
- Breast cancer, 1640, 1641
- Breast-feeding, 2238  
 as fertility determinant, 1006  
 and Freudian dependency  
 theory, 2063
- Bregantini, Luca, 1472, 1473
- Breiger, Ronald L., 1787, 2298
- Brenner, Neil, 1088, 1093
- Brentano, Franz, 1423
- Brewin, Chris, 196
- Breyer, Stephen, 1099–1100
- Brezhnev, Leonid, 2136
- Bribery  
 and governmental and political  
 corruption, 2123, 2127, 2128,  
 2129, 2131, 2134, 2137  
 international convention  
 against, 2138  
 as white-collar crime, 3251
- Bricmont, Jean, 2208
- Bridge, The* (periodical), 179
- Bridges, George S., 671
- Bridges, Jeffrey, 1636
- Bridgman, Percy, 2218
- Brierley, John E. C., 1555
- Briffault, Robert, 1271
- Brint, Steven, 2263, 2264
- Britain. *See* British sociology;  
 United Kingdom
- British Association of  
 Economists, 370
- British Crime Survey, 498–499
- British Empire. *See* United Kingdom
- British school of social  
 anthropology, 563
- British Sociological Association,  
 1425
- British Sociological Society, 1423
- British sociology, **224–229**  
 bureaucracy study, 232  
 and case studies writing, 248  
 law, 1578–1579  
 life histories and narratives,  
 1633, 1635  
 literature, 1649  
 Marxist cultural studies, 1756  
 mathematical, 1791  
 medical, 1814  
 participatory research, 2039–  
 2040, 2041  
 on popular culture, 2169–  
 2170, 2171  
 professions research, 2261–2262  
 on scientific knowledge, 2459  
 and secularization, 2484, 2485
- British structuralism, 1034
- British West Indies, 2600
- Broad, Kendal, 2764
- Broadcast industry. *See* Mass media;  
 Mass media research; Television
- Broadrick, Juanita, 2581
- “Broken heart” syndrome, 584
- Bromley, David G., 2379
- Brook Farm (Massachusetts utopian  
 community), 2849
- Brookings Institution, 1876, 2162
- Brooks, Roy, 58
- Brothels. *See* Prostitution
- Brother-sister bonds, 1509
- Broverman, D. M., 2190
- Broverman, Inge K., 2089, 2190
- Brown, Jonathan, 2190
- Brown, Richard Harvey, 1033,  
 2200, 2207
- Brown, Roger, 2894
- Brown, W., 2348
- Brown v. Board of Education* (1954),  
 2493, 2962
- Brownmiller, Susan, 2576,  
 2578–2579, 2580, 2581
- Brundtland, Gro, 1222
- Brundtland Report (1987),  
 1222–1223
- Brunei, 2974



INDEX

- Bruner, Jerome, 857  
 Bruun, Kettil, 2451, 2453  
 Bryan, William Jennings, 2369, 2426  
 Bryce, Lord, 2125  
 BSRI inventory, 1000  
 Buber, Martin, 355–356  
 Buchanan, James M., 2273, 2920  
 Buck-Morss, Susan, 541, 542  
 Buddhism, 2366, 2974, 3279, 3281, 3283, 3284–3285, 3286, 3288, 3289  
 Bueno de Mesquita, Bruce, 3241, 3244  
 Buenos Aires, Argentina, 2179  
 Bukharin, Nikolay, 1753  
 Bulgaria, 2117, 2136, 2137  
 Bulgarian Sociological Association, 2117  
 Bumpass, Larry L., 708, 2393  
 Bundy, Edgar, 2370  
*Burakumins*, 253–254  
 Burawoy, Michael, 246, 1757  
 Bureau of American Ethnology, 2890  
 Bureau of Indian Affairs, U.S., 135, 136, 137  
 Bureau of Labor Statistics, 1724  
 Bureau of the Census, U.S. *See* Census  
 Bureaucracy, **229–236**  
   agrarian, 1264  
   and alienation, 3270  
   authoritarian state, 2999–3000  
   as corporate aspect of social life, 518  
   corporations as, 442–443  
   corruption and, 2125  
   and criminal sanctions, 518, 526–527  
   critical theory and, 544  
   cultural systems of  
     administration and, 229  
   definition of, 229  
   dysfunctions in, 233–234, 235  
   elites and nonelites in, 2627  
   formal characteristics of, 230–231  
   formal legal rationality and, 1546  
   fragmentation and, 3103–3104  
   historical sociology studies, 1198  
   individualism and, 1303  
   industrial, 1312  
   and modernization theory, 1885  
   and negotiation of power, 1951  
   pervasiveness of, 234–235  
   professionals and, 2263  
   semiprofessionals and, 2261  
   and social networks, 2728  
   in socialist societies, 2850  
   specialization and, 697, 698  
   theory of legal, 1497  
   as unresponsive, 603  
   Weber theory of, 229–230, 231, 232–233, 603, 697, 698, 2163, 2623, 2627  
 Burger, Warren, 461  
 Burgess, Ernest W., 176, 2883, 3098  
   and Chicago School, 324, 325, 363  
   city development theory, 308, 1502  
   on collective behavior, 2265  
   on family structure evolution, 1505  
   human ecology hypothesis, 1209  
   marital adjustment measure, 1726, 1727  
   residential segregation survey, 2500  
   on white-collar crime, 3247  
 Burgess, Robert, 667  
 Burglary  
   rate calculation, 498  
   Uniform Crime Reports definition, 492  
   victimization rates, 499  
 Burke, Edmund, 1351, 1598, 1599  
 Burke, Peter J., 696, 2902  
 Burke, Thomas, 773  
 Burlson, Brant R., 1736  
 Burma, 2362, 2974, 2975  
   peasant rebellion study, 2977  
   slavery and slave-like practices in, 2604, 2607, 2608  
 Burnout, 2528  
 Burns, J. M., 1566  
 Burr, Wesley, 1729  
 Burrow, Trigant, 2084  
 Bursik, Robert J., 2658  
 Burt, M., 2592  
 Burt, Ronald S., 737, 738, 2693, 2732, 2753, 2827  
 Burton, M., 1998  
 Burty, Cyril, 2458  
 Burundi  
   genocide, 68, 107, 1066, 1070, 2629  
   women in labor force percentage, 3262  
 Bus segregation, 2494–2495  
 Bush, George, 1825  
 Bushman, B. J., 73  
 Bushnell, Horace, 9  
 Business Roundtable, 2162  
 Buttel, Frederick G., 1214, 1217, 1228, 2429, 2431  
 Bystander inaction, 115  
 Byzantine Empire, 2998  
**C**  
 CA. *See* Conversation analysis  
 Cab problem (judgment consistency), 591  
 Cabet, Etienne, 2847, 2849  
 Cable television, 271, 1768  
 Cacioppo, John, 188  
 Caesaro-papist regimes, 2357  
 Caffeine, 654  
*Cahiers du Musée Social*, 1025  
*Cahiers Internationaux de Sociologie*, 1026  
 Cain, Glen, 2213  
 Cairo Conference on Population and Development (1994), 932, 2233  
 Calamities. *See* Disaster research; Society and technological risks  
 Calavita, Kitty, 3250  
*Calculus of Consent, The* (Buchanan and Tullock), 2920  
 Caldwell, John, 628, 1009  
 Calí cartel, 2135  
 California, Asian-American studies in, 179  
 California Bar Association, 468  
 California Psychological Inventory, 2076  
 California School, 853

- Callahan, Daniel, 588
- Calment, Jeanne, 1631
- Calvin, John, 3227
- Calvinism, 2211, 2520
- Cambodia, 2974, 2975, 2978  
 fertility decline, 627  
 genocide, 68, 1069, 1070, 1384, 2975  
 slavery and slave-like practices, 2604, 2606–2607  
 women in labor force percentage, 3262
- Cambodian Americans, 175, 180
- Cambridge-Somerville Study, 1488
- Cameroon, 2133  
 affinity of religion and family in, 936  
 and family size, 977  
 slavery and slave-like practices, 2604
- Camp David agreements (1978), 2048
- Campbell, Angus, 2299, 2300, 2303–2304, 2683, 3234, 3235
- Campbell, D. T., 2324–2325, 2326, 2327, 3128–3129, 3210
- Campbell, Donald, 865–866, 2282
- Campbell, John Creighton, 380
- Campbell, John L., 2164
- Campbell, Richard T., 1692
- Canada  
 church-state relations, 2359–2361, 2362  
 clinical psychology, 327  
 cohabitation, 109  
 divorce law reforms, 703  
 divorce rate, 112, 706  
 equal pay for work of equal value, 372  
 family violence, 981  
 health-care system, 374, 377, 378, 379, 380, 1827  
 legal system, 465, 471  
 long-term care and care facilities, 1660–1661  
 multilingualism, 2909  
 Muslim minorities, 2950  
 occupational status attainment, 2786  
 political system, 2359, 2360  
 prohibition legalization, 1577  
 publishing industry, 1648  
 racial conflict, 321  
 Social Science Data Archive, 576  
 social security system, 2800  
 tourism in, 3169  
 transnational corporations, 3175
- Canadian Human Rights Act of 1977, 372
- Canadian Mobility Study, 2786
- Cancer  
 alcohol abuse and, 93, 1640  
 dietary factors, 1641  
 quality of life and, 2301  
 smoking and, 1639
- Cancian, Frank, 87
- Canon law, 473, 1513, 1514, 1516  
 Montesquieu's comparative study of, 1545
- Canova, Fabio, 644
- Cantril, Hadley, 2303
- Canvases and Careers* (White and White), 172
- CAPI (Computer-Assisted Personal Interviewing), 410
- Capital  
 accumulation of, 1089, 1093, 1099, 1103  
 definition of, 2637  
 financial, 1264  
 flow of, 1088  
 foreign, 1087, 1088, 1268  
 Marx on contradictions of, 3066  
 surplus, 1265  
 transfer of, 1265  
*See also* Cultural capital; Human capital; Social capital  
*Capital* (Marx). *See Kapital, Das*
- Capital punishment. *See* Death penalty
- Capitalism, **237–243**  
 agricultural, 2429, 2432  
 and alienation, 100, 697  
 in American society, 99, 143–144  
 and anomie, 164  
 and class conflict, 2692  
 and class-race relationship, 320–323  
 and colonialism, 240, 320  
 and communitarian views, 360  
 comparative historical analysis of, 383  
 as conducive condition for democracy, 605  
 conflict theory on, 415  
 contradictions in, 1311  
 and crime theories, 504  
 and criminal and delinquent subcultures, 511–512  
 and critical theory, 540, 541, 542, 544  
 and dependency theory, 639–642  
 diffusion of, 1085  
 and division of labor, 697, 1782  
 and economic determinism, 723  
 and economic institutions, 724, 728  
 and economic sociology, 733  
 expansion of, 1197, 1215  
 and feminist theory, 989, 990  
 and globalization, 1084, 1085, 1091  
 historical sociology on, 1197, 1199  
 and human ecology, 1215  
 and imperialism, 1265  
 and individualism, 1303  
 and intellectualism, 1356  
 and kinship systems, 1502  
 and labor movement, 1528, 1531–1533  
 and Latin America, 1540–1541  
 law and origins of, 1553–1554, 1576  
 liberalism roots of, 1597  
 marginal utility analysis of, 2698  
 Marx case studies of, 245–246  
 Marx on, 237, 774, 1783, 2196, 3066  
 Marxist class structure theory and, 2814, 2847  
 Marxist deviance theory and, 669–670  
 Marxist historical materialism and, 543, 1704, 1782  
 Marxist historical specificity and, 2645–2646

- Marxist legal theory and, 1553  
 Marxist leisure theory and, 1583  
 Marxist revolution theory and, 2410, 2411  
 Marxist sociology and, 238, 531  
 Marxist structural theory of the state and, 2163  
 mass culture attributed to, 1645–1646  
 materialist analyses of new phase of, 1784–1785  
 mature states of, 1309  
 and modernization theory, 1885, 1886  
 monopoly stage of, 1264, 1754  
 neo-Marxist view of, 1078–1079  
 patriarchy and, 1708  
 postcommunist transitions to, 2851–2852  
 and postindustrial society, 2194–2202, 2205  
 and postmodern society, 2206, 2207  
 and power elite, 2162  
 power-conflict analysis in context of, 53  
 and proletariat struggle, 238–239  
 Protestant Ethic and, 2483, 2942, 2943, 3222  
 and racism, 319–321  
 and retirement patterns, 2402  
 slavery linked with, 238, 239, 321  
 and social justice theories, 2697–2698  
 and social security systems, 2800  
 and societal stratification, 2865  
 and structuralist theory of the state, 2163  
 and time use research, 3156  
 and transnational expansion, 239–240, 241, 242, 1085  
 and war, 3243–3244  
*See also* Corporate organizations  
*Capitalism and Modern Social Theory* (Giddens), 226  
*Capitalism, Socialism and Democracy* (Schumpeter), 733  
 Caplow, Theodore, 331, 332, 364, 367  
 Capone, Al, 3246  
 Car industry. *See* Automobiles  
 Carbonaro, Antonio, 1467, 1470  
*Cardinal Principles of Secondary Education* (1918), 764  
 Cardoso, Fernando H., 642, 1087, 1538  
 Cardozo, Benjamin, 476  
 Care  
     as moral reasoning basis, 1900, 1902–1903  
     mother's role and, 2036  
     personal dependency and, 2062  
 Career line  
     definition and concept of, 1982, 1984  
     *See also* Occupational and career mobility  
 Caregiver burden, 1658  
 Cargo cults, 2367  
 Caribbean News Agency, 1767  
 Caribbean region  
     African studies, 66  
     Demographic and Health Surveys, 633  
     demographic transition, 622, 627  
     governmental and political corruption, 2134  
     replacement-level fertility, 220  
     slavery in, 2599, 2600–2601  
 Carlsmith, M., 338, 339  
 Carlson, Richard, 677  
 Carlsson, N. Gösta, 2450  
 Carmichael, Stokely (Kwame Ture), 53  
 Carmines, E. G., 2345, 2347, 2350  
 Carnap, Rudolf, 821, 2756  
 Carr, E. G., 214  
 Carroll, Lewis, 2339  
 Carroll, Michael P., 2965  
 Carrying capacity, 1219  
 Carson, Rachel, 789, 803  
 Carter, Jimmy, 712, 2048  
 Carter, Rosalyn, 588  
 Cartesianism. *See* Descartes, René  
 Carthage, 1066, 1069  
 Cartwright, Dorwin, 335–336, 1034, 2415, 2417, 2611  
 Case frame grammar, 2297–2298  
 Case law, 467–468, 476  
 Case studies, **243–249**  
     Chicago School tradition of, 243–244  
     of childhood sexual abuse, 288, 293  
     in clinical sociology, 327  
     of counterculture, 461–463  
     functionalist-structural theories of, 244  
     as life histories, 245, 247–248, 1633  
     methodology of, 327  
     politics and poetics of writing, 247–248  
     radical, 245–246  
     of reality construction, 246–247  
     retirement models, 2405–2406  
     of revolutions, 1198  
 Cassubians, 2268  
 Caste and inherited status, **249–255**  
     Comte on, 1029  
     elites and, 2623  
     in Hindu society, 250–253, 3284  
     in Japan, 253–254  
     in Rwanda, 254  
     stratification parameters, 2810, 2811  
     *See also* Status attainment  
 Castell, Manuel, 1758  
 Castellano, Vittorio, 1465, 1470  
 Castro, Fidel, 2851  
 Castro, Luis J., 620  
 Catalano, Richard, 366  
 Catastrophes. *See* Disaster research; Society and technological risk  
 Categorical models, 1849, 3037–3038  
     tabular analysis, 3107–3126  
 Catell, Raymond, 2084  
 Catherine the Great (Russian czarina), 1435  
 Catholic Church. *See* Roman Catholic Church  
 Catholic schools. *See* Parochial schools  
 CATI (Computer-Assisted Telephone Interviewing), 410, 1802, 3092  
 Cattarinussi, Bernardo, 1468

- Catton, William, 1214, 1219  
*Caudillo-cacique* power structure (Mexico), 1856  
 Causal inference models, **255–267**  
   and attribution theory, 193–194, 195  
   and comparative historical analysis, 387–388  
   and criminology, 502–504, 536  
   and decision-making theory, 591–592  
   and deviance theories, 662  
   and factor analysis, 908, 917–918  
   and longitudinal research, 1685, 1688–1689  
   and measurement errors, 256–257, 264–266, 1908–1909  
   and multiple indicators, 1907–1923  
   and path analysis, 259–260, 455–456  
   of personal autonomy, 2059  
   and quasi-experimental research designs, 2310–2327  
   scenarios, 1041–1042, 2678  
   and simultaneous equation models, 261–264  
   in social psychology research, 2771–2778  
   and statistical analysis, 3035  
   and structural equation modeling, 1922–1923  
   and tabular analysis, 3108–3126  
   and voting behavior research, 3234–3235  
   *See also* Correlation and regression analysis; Experiments; Scientific explanation  
*Causal Inferences in Nonexperimental Research* (Blalock), 3035  
*Causal Models in the Social Sciences* (Blalock), 3036  
 Causality orientations theory, 2059  
 Causation, definition of, 255  
 Cause and effect. *See* Causal inference models  
 Cavalli, Luciano, 1468, 1470  
 C.C.S. (Certified Clinical Sociologist), 326  
 CCRCs. *See* Continuing care retirement communities  
 CDC. *See* Centers for Disease Control  
 CD-ROMs, 409, 420, 1605  
   data archives, 580, 2476  
   secondary data sets, 2479  
   and sexually explicit material, 2185  
 Cell frequency, 658, 661  
 Cella, Gian Primo, 1467  
 Censorship and the regulation of expression, **267–281**  
   civil liberties and, 315  
   communitarian view of, 360  
   hate speech and, 2140  
   judicial treatment of, 270–272  
   mass media, 1762–1763  
   national security and, 273–274  
   political correctness and, 2140  
   pornography and, 274–275, 2184, 2185–2186  
   rhetoric of, 268–269  
   and school curricula, 276–278  
   and social psychology, 278–279  
   sociolinguistics and, 2908  
   and student movements, 3070  
   *See also* Free speech  
 Census, **281–287**  
   accuracy of, 285–287  
   of agriculture, 2432  
   and China studies, 300, 301  
   County and City Data Books, 2480–2481  
   crime survey, 1488–1489  
   demographers and, 637  
   demographic data, 631  
   divorce statistics, 701  
   document depositories, 1606–1607, 2477  
   of family size, 971  
   and first automated data processing machine, 406  
   labor force concept and measurement, 1521–1523  
   marriage and divorce rates, 1744  
   National Crime and Victimization Survey, 494  
   National Longitudinal Surveys data, 2476  
   nursing home residents profile, 1667  
   occupational categories, 697, 1997, 3264  
   occupational prestige scale, 1997  
   population sampling, 2444  
   poverty data, 2214–2215  
   race and ethnicity data, 284–285, 286  
   remarriage rates, 2387, 2388  
   rotating panel sample, 1687  
   sampling procedure, 2448  
   as secondary data analysis source, 2475–2476, 2480–2481  
   social indicator publication, 2685  
   social surveys, 578  
   sociologist careers and, 148  
   standardization, 2995  
   suburban definition, 3070  
   urban underclass measurement, 2212, 3199  
   widowhood data, 3256–3257  
 Census Bureau, U.S. *See* Census  
 Center for Advanced Study in Political Science, 172  
 Center for Coordination of Research on Social Indicators, 2683  
 Center for Epidemiological Studies Depression Scale, 654  
 Center for Human Resource Research, 2475–2476  
 Center for Political Studies, 101  
 Center for Sociological Studies, 1026  
 Center for Substance Abuse Prevention, 715  
 Center for Women Policy Studies, 2577  
 Center to Improve Care of the Dying, 588  
 Centers, R., 1778  
 Centers for Disease Control, 1162, 2585, 2587, 2588  
 Central African countries, 60  
 Central African Republic, 2129, 2133–2134  
 Central America. *See* Latin America; Latin American studies; *specific countries*  
 Central economy, 1313, 2849, 2850–2851

- Central Europe. *See* Eastern Europe
- Central Intelligence Agency, 274, 898, 2144
- Central tendency of distribution, measures of, 659
- Centrifugal system of kinship, 1511–1512, 1513, 1514
- Centripetal system of kinship, 1510–1511, 1512, 1513–1514
- Centuries of Childhood* (Ariès), 2090
- Cerebral arteriosclerosis, 139
- Cerezo, Vincio, 2135
- Ceri, Paolo, 1467
- Certainty effect, 591–592
- CES. *See* Committee on Economic Security
- Césaire, Aimé, 66
- Cesarean section, 2237
- CES-D (depression scale), 1834, 1835
- CESSDA (Council of European Social Science Data Archives), 575, 576–577
- CFA. *See* Confirmatory factor analysis
- Chadwick, Bruce A., 364, 367
- Chain migration, 177
- Chain of being, 1599
- Chains of Opportunity* (White), 2662
- Chaitanya, 3287–3288
- Chajanov, A. V., 2979
- Chalasiniski, Jozef, 2119
- Challenger* disaster, 2875, 2878, 3253
- Chambliss, William J., 529, 530, 1497, 2961
- Champs de la sociologie française, Les* (Verret and Mendras), 1026
- Chancery courts, 477
- Chandler, Alfred, 738
- Chaney, James, 2495
- Change. *See* Social change
- Change measurement. *See* Experiments; Longitudinal research; Quasi-experimental research design; Measurement
- Chaos theory, 1212–1213
- Chapple, Elliot D., 1974, 1975, 1976
- Chapter 7 bankruptcy liquidation, 202, 203, 205
- Chapter 11 bankruptcy reorganization, 203
- Chapter 12 bankruptcy reorganization, 203
- Chapter 13 bankruptcy repayment plan, 202–203, 205
- Charismatic authority, 229–230  
emotional factors in, 2519–2520  
irrational legal order and, 1546  
as nationalist movement leadership, 1943
- Charity. *See* Philanthropy
- Charity Organizations Societies, 2841
- Charlemagne, 2940, 2998
- Charmaz, Kathy, 583
- Charny, Israel, 1072
- Charter schools, 765–766
- Charts  
descriptive statistics, 658–659  
graphic, 3003–3022
- Chateaubriand, François-Auguste-René, 1771
- Chattel slave systems, 2596–2597, 2601, 2603
- Chattopadhyay, K. P., 1291
- Chaves, Mark, 2379, 2484
- Chavez, Hugo, 2135
- Chavis, David, 363
- Chayanov, A. V., 2432
- Chechnya, 1945, 2136, 2947, 3001
- Checkoway, Barry, 3071–3072
- Chemers, M. M., 1571
- Chen Da, 298
- Cherlin, Andrew J., 702, 705, 706, 1390, 1391, 2393
- Chernobyl disaster, 683, 686, 805, 2875, 2876, 2877
- Chernomyrdin, Viktor, 2136
- Chesler, Pat, 2089
- Chess, S., 2088, 2090
- Cheung, Chau-Kiu, 1901
- Chi square, 1808–1809
- Chiapas rebellion (1994), 1861, 2271
- Chicago Eight, 2146
- Chicago, Illinois  
first juvenile court, 1485  
machine politics, 2126  
Muslim immigrants, 2950  
settlement houses and social reform, 365–3363  
slum life study, 365  
social disorganization study, 1495  
sociological child guidance clinics, 324
- Chicago School, 431, 532–533, 821, 843  
Asian-American research, 176, 177  
case studies tradition, 243–244  
clinical sociology, 324, 325  
collective behavior studies, 2265  
community studies, 363, 365  
deviance theory, 664–665, 2658  
and ethnography, 852, 853  
and functionalism and structuralism, 1031  
and human ecology, 2822  
life histories and narratives, 1633  
neighborhood focus of, 664–665  
pragmatist roots of, 2220–2221  
Simmel as influence on, 1772  
social control perspective, 2657  
and social problems, 2759
- Chicago School of Civics and Philanthropy, 366
- Child abuse and neglect  
African and Asian forced labor, 2605–2606  
and family violence, 981  
learned aggression from, 71  
low-birth-weight babies as, 221  
treatment and prevention, 75  
*See also* Childhood sexual abuse
- Child care, 128–129, 2032  
costs of, 2035  
employer-provided, 3266  
structural lag and, 3062
- Child care centers, 129, 359
- Child custody, 702, 707  
and sexual abuse charges, 2582  
and social justice, 2707  
and support payments, 128, 708, 947, 2707
- Child fingerprinting, civil liberties and, 316, 318
- Child guidance clinics, 324, 325
- Child labor, 2036, 2605–2606, 3262
- Child pornography, 274
- Child Study Movement, 2

- Child support payments, 128, 708  
 and family law, 947  
 and social justice, 2707
- Childbearing  
 age at first birth, 125  
 American family patterns, 125, 2032–2033  
 average number of births per woman, 2031–2032  
 birth rates, 217–220  
 by single women, 125, 634, 708, 1506, 1744, 2033  
 comparative health-care systems, 374, 378  
 as demographic research area, 635  
 gender preference in, 628  
 historical decline in, 2032  
 and life-cycle demographic model, 1625  
 as life-cycle transition, 1616, 1623, 1625  
 and lifestyle risks, 1640  
 and marital quality, 1729, 1730–1731, 1737, 2035, 2037  
 marriage as legitimization of, 1734  
 medicalization of, 1816  
 morbidity and mortality, 2236  
 rewards and costs, 2034–2033  
 sociobiological law of anisogamy on, 2884–2885  
 Total Fertility Rate, 627, 628, 629  
*See also* Family planning; Family size; Fertility determinants; Infant and child mortality; Pregnancy and pregnancy termination
- Childfree adults, 109–111, 634, 1506, 1625  
 rising rate of, 2035
- Childhood  
 aggressio-learning factors in, 70, 75, 76  
 altruism in, 115  
 attitude formation in, 185  
 censorship and, 272  
 cross-cultural socialization analysis, 550  
 dependency theory and, 2063–2064, 2066–2067  
 developmental studies, 1617, 1624, 1686  
 as distinct life stage, 122, 2681  
 family roles and, 122, 696, 2032  
 family trends and, 127–128, 142, 2033  
 Freudian drive theory and, 1713  
 historical views of, 2035–2036, 2090–2091  
 homuncularism and, 2090–2091  
 mass media influences in, 1762–1764  
 moral development stages, 1895–1897  
 nonmaternal caretakers, 128–129, 359  
 parental costs during, 2034–2035  
 parental divorce effects on, 127–128, 705–707, 1747, 1749, 1750, 2033, 2392  
 parental prescriptive altruism and, 1508  
 parental remarriage effects on, 2391–2392  
 parental roles and, 2035–2037  
 as personality influence, 2090  
 phenomenological investigation of, 2102–2103  
 poverty rates, 127, 1287  
 self-esteem development in, 2508, 2512–2513  
 sexual behavior in, 2550–2551  
 socialization, 2856–2860  
 and structural lag, 3062–3063  
*See also* Adolescence; Human rights/children's rights; Infant and child mortality
- Childhood sexual abuse, **288–297**, 2581–2582  
 and Asian sex trade, 2607  
 cross-cultural issues in, 292–293  
 definitions of, 288, 292, 2581  
 explanations of, 2582  
 and family violence victim attributions, 196  
 incidence and prevalence of, 288–289  
 intrafamilial, 2594  
 long-term effects of, 289–292, 293  
 by other children, 2582  
 and pornography, 274  
 rape theories and, 2591  
 treatment of, 293, 521  
*See also* Incest
- Childhood: A Global Journal of Child Research*, 550
- Childlessness. *See* Childfree adults
- Child-rearing styles. *See* Family and household structure; Parental roles; Socialization
- Children. *See* Childhood
- Children of the Great Depression* (Elder), 1618, 2662
- Children's Bureau, U.S., 366
- Children's Defense Fund, 130
- Children's Health Insurance Program, 1146
- Children's rights, 1239–1240, 1242–1244
- Chile  
 dependency theory and, 642  
 economic liberalization, 1539, 1540, 1541  
 economic, social, and political conditions, 1536, 1537  
 fertility decline, 627  
 health-care system, 381  
 labor movement, 1532  
 lawyers in, 478  
 protest movements, 2266
- Chin, Ko-lin, 512
- Chin, Vincent, 1410
- China studies, **297–304**  
 AIDS/HIV cases, 2591–2592  
 and Cold War-era triad, 332  
 communist dictatorship, 3002  
 conditions counter-conducive to democracy, 605  
 delinquent subcultures, 512  
 education and mobility, 3045  
 ethnonationalist movements, 1944, 1946  
 family and population policy, 930, 931–932, 972  
 fertility control, 220  
 fertility decline, 627  
 governmental and political corruption, 2137  
 health-care system, 380–381  
 historical empires, 2998–2999

- and historical sociology study  
of, 1198
- labor movement, 1532
- mass media, 1767
- nationalist movement, 3001
- occupational mobility, 1988, 1993
- protest movement and  
countermovement, 333, 2268,  
2270, 2718, 2721–2722, 3067
- revolutions, 3000, 3001
- secularization, 2485
- sex trade, 2607
- sexually transmitted diseases,  
2592
- slavery and slave-like  
practices, 2604
- social change patterns, 2643,  
2645
- social gerontology, 301
- socialism and communism, 298,  
2849, 3002
- socialist economic  
modifications, 2851
- tourism in, 3167, 3169
- women's equality, 990
- Chinese Americans
- chain migrants, 177
- crime rates, 531
- demographic characteristics, 176
- economic system of, 182
- immigration restrictions, 123,  
175, 176, 321
- income and status  
attainment, 181
- opium-use stereotype, 714
- social isolation study, 176, 178
- youth gangs, 512
- Chinese Communist Party  
(CCP), 298
- Chinese Democracy Movement,  
2721–2722
- Chinese Exclusion Act of  
1882, 123, 175
- Chinese Laundryman, The: A Study in  
Social Isolation* (Siu), 176
- Chinese Sociological Association,  
297, 300
- Chipendale (instructional  
package), 411
- Chirac, Jacques, 2129
- Chiropractors, 221, 2259
- Chivot, Daniel, 645, 2978
- Chi-square test, 1965, 1966, 1967
- Chiswick, Barry, 181–182
- Chivalry, 483–484
- Chlamydia, 2587, 2592
- Chmielewski, Piotr, 2119
- Chodorow, Nancy, 992–993, 994,  
998, 1058, 2172
- Choice* (book review  
publication), 1606
- Choice in Dying, 588
- Cholera, 814
- Cholesterol levels, 1642
- Chomsky, Noam, 438, 1033,  
1233, 2899
- Choron, Jacques, 582
- Chretien, James K., 2589
- Christal, R., 2085
- Christian Coalition, 3229
- Christian Democratic Party (Italy),  
2128, 2129
- Christian Science, 2366
- Christianity
- Engels's view as true  
communism, 2968
- and ethical norms, 3283–3284
- and fundamentalism, 2368–2372
- interfaith dialogue and, 3289
- Islamic parallels, 2937, 2939,  
2943
- kinship and family typologies,  
1507, 1516–1517
- and new religious  
movements, 3287
- origins, 3287
- religious experiences and  
symbols, 3280, 3281
- view of human nature, 2086
- See also* Protestantism; Religious  
movements; Religious  
organizations; Roman  
Catholic Church
- Christie, Nils, 2453
- Christine de Pisan, 988
- CHRR. *See* Center for Human  
Resource Research
- Chubais, Anatoly, 2136
- Chun Doo Hwan, 2131
- Chuprov, A. A., 2979
- Church and state, 2356–2358
- interpenetration model, 2356,  
2357, 2358, 2359–2360
- separation of, 146, 2356–  
2357, 2358
- two-powers model, 2357–2358
- Church groups. *See* Religious  
organizations
- Church of England, 701
- Church of Scientology, 900,  
2366, 3287
- Churches. *See* Denominations;  
Religious organizations
- Churchill, Winston, 1599
- Churching of America, The* (Finke and  
Stark), 2485
- Church-sect typology, 2365,  
2366–2367, 2373, 2378
- CIA. *See* Central Intelligence  
Agency, 898
- CICD. *See* Center to Improve Care  
of the Dying
- Cigarette smoking. *See* Smoking
- Cigna (health insurance  
provider), 1822
- Ciller, Tansu, 2132
- CIO. *See* Committee on Industrial  
Organizations
- Cipolla, Costantino, 824
- Cipriani, Roberto, 1473
- Circular reaction, 350
- Circulation mobility, 2712
- Cirrhosis, 93, 94, 1640
- Cisneros, Henry, 2128
- Citation indexes, 1610–1611
- Cities, **305–314**
- African-American segregation  
indices, 2500–2504
- alcohol consumption patterns, 94
- case studies, 243–244
- deconcentration, 3195
- definition of, 305
- ethnic enclaves in, 2498
- as fertility transition factor, 624
- juvenile gangs in, 1485
- kinship systems, 1502, 1503
- mass society theory on,  
1772, 1773
- megacities, 310–311, 312, 3197

- mental illness incidence, 1840, 1841
- metropolitan areas. *See* Metropolitan Statistical Areas
- Native American population, 136
- police forces, 2110–2112
- political corruption in, 2125–2126
- population distribution, 634
- population of world's largest metropolises, 3195
- population shifts and, 2179–2180, 3195–3196
- See also* Urbanization
- primate systems, 3196–3197
- quality of life ranking, 2302–2303
- rural societies contrasted with, 3192
- secondary data analysis, 2480–2481
- settlement houses, 2841
- social reform, 365–366
- and suburbanization, 311, 3070–3076, 3194
- underclass. *See* Urban underclass
- underemployment in, 1721
- zoning against “adult” businesses, 2186
- See also* Community; Urban sociology
- Citizens Clearing House for Hazardous Waste, 791
- Citizens for Excellence in Education, 770
- Citizens groups. *See* Public interest groups
- Citizenship
- discrimination based on, 692, 693
- and social security systems, 2797, 2800
- “City, The: Suggestions for Investigation of Human Behavior in the Urban Environment” (Park), 1772
- City of God* (Augustine), 1507
- City Police* (Rubenstein), 2114
- City-states, 2998
- Civic Forum (Czech Republic), 1532
- Civic society, 360
- Civil disobedience. *See* Protest movements; Student movements
- Civil law cases, 471–472, 479–480
- Civil law system, 464, 465–466, 471, 472–476, 480–481, 1554
- judges in, 477
- kinship distance computation, 1514
- lawyers in, 478
- legal rules vs. social norms, 1546
- Civil liberties, **314–319**
- democracy associated with, 605, 606
- liberalism's emphasis on, 355
- political correctness and, 2140
- responsive communitarianism and, 357, 361
- See also* Human rights/children's rights
- Civil rights
- affirmative action and, 47, 48, 49
- civil liberties distinguished from, 315
- and critical theory, 539
- and crowd behavior, 557, 558
- and feminist theory, 989
- pornography seen as issue of, 275
- See also* Censorship and regulation of expression; Civil rights movement; Discrimination; Segregation and desegregation
- Civil Rights Act of 1875, 2491
- Civil Rights Act of 1964, 275–276, 830, 3067
- affirmative action and, 48
- as civil rights movement outcome, 2266, 2267, 2268
- continuing racial discrimination and, 56, 2497
- as enforcement device, 2496
- equal pay for equal work provision, 372, 2706
- and expanded job opportunities, 3264–3265
- sexual harassment coverage, 2580
- Civil Rights Act of 1990, 275–276
- Civil rights movement, 58, 64, 136, 273, 558, 1315, 2266, 2267, 2493–2496, 2723
- and countermovement, 2267
- desegregation as focus of, 2494–2495
- emergence theory and, 2718
- ethnic studies and, 178–179
- Montgomery bus boycott as impetus for, 2493
- nonviolence tactic, 2269
- participant characteristics, 2145, 2268–2269, 2495, 3067–3070
- political alienation and, 103
- public opinion effects of, 2277
- as religious movement, 2365, 2374, 2377, 2493, 2719
- social stratification effects of, 142
- students and, 3067–3068, 3069–3070
- and urban race riots, 2270, 2495
- Civil service, 603, 2127, 2130, 2131, 2132
- Civil Service Retirement Act of 1920, 2402
- Civil War (U.S.), 1876, 2127
- Irish Catholic draft riots, 2269, 3069
- veteran pensions, 2797
- Claes, Willy, 2130
- Clans. *See* Indigenous peoples; Kinship systems and family types
- Clark, Burton, 1182, 1183
- Clark, Candace, 782
- Clark, Colin, 1218
- Clarke, J., 1578
- Clarke, M., 3249
- Clarkson, F. E., 2190
- “Clash of civilizations” hypothesis, 2940–2941
- Class. *See* Class and race; Marxist sociology; Social class; Social and political elites; Social structure; Status attainment
- Class and Class Conflict in Industrial Society* (Dahrendorf), 225–226
- Class and race, **319–323**
- academic achievement, 2933–2934
- adolescent sexual behavior patterns, 2551–2552
- African American studies, 53, 56–57
- as all-volunteer military force issue, 1878–1879, 1880
- ascriptive stratification, 2819



- British sociological studies, 226  
 caste inherited status, 250  
 city-suburb socioeconomic differences, 3072–3073  
 contemporary relations, 322–323  
 courtship endogamy, 485  
 criminal deviance, 526, 530, 531, 534, 536  
 criminal sanctions, 517  
 divorce correlates, 704–705, 707–708  
 education and mobility, 756, 757  
 environmental equity, 796  
 ethnic stratification, 845  
 fertility transitions, 1007  
 filial responsibility, 1020  
 historical sociology, 1198  
 identity politics, 1580  
 illicit drug use/convictions, 713  
 income inequality, 3048  
 inequality theories, 2692  
 juvenile delinquency and crime, 1490–1492  
 leisure, 1584–1585, 1588  
 Marxist criminology, 534–535  
 popular culture transmission, 2170–2172  
 prejudice, 2242–2246  
 status attainment, 3044  
 stratification models, 2817–2819  
 whiteness vs. blackness, 56  
*See also* Racism; Segregation and desegregation
- Class Structure in the Social Consciousness* (Ossowski), 2119
- Class struggle, 415, 601, 697, 1221, 1782  
 as basis of Marxist socialist ideal, 2847  
 and industrial sociology, 1310  
 oppositional framework for, 2692  
 religious orientation and, 2385, 2968  
 revolutions and, 2410  
 societal stratification and, 2865  
 theory of state and, 2163
- Classes* (Wright), 723
- Classes, Strata, and Power* (Wesolowski), 2120
- Classical elite theory. *See* Social and political elites
- Classical music. *See* Music
- Classical school of criminology, 528, 535, 536
- Classification. *See* Tabular analysis; Typologies
- Clausen, John A., 1616, 1620, 2860
- Clear, Todd, 360
- “Clear and present danger” designation, 318
- Clegg, Stewart R., 2201
- Clemens, Elisabeth S., 604
- Clemmer, Donald, 2051–2052, 2053, 2055
- Cleveland, Harlan, 1346
- Cleveland, William S., 3006, 3009–3010, 3017, 3018
- Client-centered therapy, 1715
- Clinard, Marshall B., 3247, 3248–3249, 3250, 3252
- Clinic for the Social Adjustment of the Gifted (New York University), 324
- Clinical depression (depressive disorder), 649–650, 651, 656
- Clinical Methods in Interracial and Intercultural Relations” (Haynes), 326
- Clinical psychology, 325, 326
- Clinical sociology, **323–329**  
 first formal definition of, 326  
 first known use of term, 323  
 first use of term in print, 325  
 history of, 323–326, 327  
 in international settings, 327–328  
 linear models, 592–593  
 professional association, 155, 326  
 and sociological practice, 326, 327  
 theories, methods, and intervention strategies, 327
- “Clinical Sociology” (Wirth), 325
- Clinical Sociology Association, 155, 326
- Clinical Sociology Review*, 326, 328
- “Clinical Study of Society, The” (Lee), 325
- Clinton, Bill, 361–362, 791, 1144, 1147, 1159, 1239, 1244, 1531, 1820, 1825
- administration scandals, 2128, 2276–2277, 2581, 3068  
 and family leave act, 2033–2034  
 and gays in the military issue, 1878, 1881  
 and national health care plan, 2804  
 and public opinion, 2276–2277, 2278
- Clinton, Hillary Rodham, 361–352
- Clogg, Clifford, 1720, 3038
- Cloning, 879, 1824, 2091
- Closed population, 616
- Cloth making, 64
- Clothing. *See* Fashions
- Cloward, Richard, 101, 513  
 on democracy, 603, 604, 606  
 on juvenile delinquency, 1945  
 on penology, 2052–2053  
 on urban gangs, 663  
 on welfare system, 2661
- CLS. *See* Critical legal studies (CLS) movement
- Club of Rome, 1038
- Clubs. *See* Voluntary associations
- Cluster analysis. *See* Correlation and regression analysis; Factor analysis
- Cluster sampling, 2447–2448
- CNN (Cable News Network), 1766
- CNN syndrome, 684
- Coale, Ansley J., 425–426, 619, 620, 624, 625, 626, 629, 631, 632, 633, 634
- Coalition governments, 2159
- Coalitions, **329–334**  
 choice of partners in, 331, 332  
 in conflict theory, 414–416  
 modern empirical work in, 330–331  
 and network exchanges, 2673  
 position of weaker party in, 331, 332  
 real-life applications of, 330–333  
 triads and, 329, 330, 331–332, 465  
 and war, 3242
- Coase, Ronald, 735, 2340
- Coase theorem, 2340
- Cobb, Jonathan, 542

- Cocaine, 711, 712, 713, 714, 715  
 Latin American traffic in, 2135  
*See also* Crack cocaine
- Cocaine Anonymous, 715
- Cocaine cartels, 2135
- Cochran's *Q* test, 1969
- Cockburn, Alexander, 1244
- Code Napoléon*, 474, 475, 1513
- Code of Ethics* (American Sociological Association [1989] 1997), 836–840
- Code of Ethics* (National Association of Social Workers), 2842
- Code of Jewish Law, The (Shulkham Arukh)*, 1510
- Codes  
 civil law, 474–476, 481  
 imperial, 2999
- Coding  
 computer-assisted content analysis, 419–420  
 in qualitative research, 2289, 2290
- Coefficient of correlation, 447, 449, 457  
 differential, 1693–1694  
 multiple, 452, 454  
 partial, 451–453, 457
- Coercive persuasion. *See* Extreme influence: thought reform, high control groups, interrogation, and recovered memory psychotherapy
- Coercive power. *See* Authority
- Cognitive anthropology, 2891
- Cognitive consistency theories, **334–342**  
 attitudes and, 187–189  
 decision processing, 595  
 personality theory, 2084
- Cognitive dissonance theory, 335, 337–340, 893  
 alternate versions of, 339–340  
 attitudes and, 187–188  
 belief adjustment and, 2701  
 history of, 338  
 situations promoting, 187–188
- Cognitive distortions  
 from childhood sexual abuse, 290  
 as depression model, 651  
 self-esteem and, 2514–2517
- Cognitive evaluation theory, 2059–2060
- Cognitive institutions, 395
- Cognitive psychology  
 aggressive personality theory, 70  
 altruism theory, 115  
 developmental invariant sequence, 1900, 1901  
 medical sociology and, 1814  
 moral development stages, 1894, 1895–1907  
 role theory, 2417–2418, 2424  
 social values and norms, 2828–2839
- Cognitive resource theory, leadership model, 1568
- Cognitive theories of depression, 651–652, 653–654  
 treatments, 655
- Cohabitation, 108–109, 484, 487, 489  
 attitudes toward, 106, 109, 2568  
 by divorced persons, 1747  
 courtship and, 108, 484, 487, 1779  
 and declining marriage rate, 1744  
 and declining remarriage rates, 1749  
 and extradyadic sex, 2541–2545  
 gay male and lesbian relationships, 2545–2547  
 higher divorce rate linked with, 705  
 increased rate of, 131, 1506, 1750  
 population composition and, 634, 635  
 premarital, 705  
 and remarriage rates, 1749, 2388  
 and sexual patterns, 2539, 2540
- Cohen, Albert, 510–511, 534, 1494
- Cohen, Bernard P., 2702
- Cohen, E., 3169
- Cohen, G. A., 1784
- Cohen, Gershon D., 2969
- Cohen, J., 1848–1849
- Cohen, Jack C., 2500
- Cohen, Lawrence, 506
- Cohen, Morris, 2218
- Cohen, Percy, 225
- Cohen, Ronald L., 2698
- Cohen, Stanley, 1578, 1579
- Cohen, Stephen P., 2614
- Cohen, William, 1244
- Cohn, Norman, 1071, 2367, 2968
- Cohort analysis. *See* Cohort perspectives; Longitudinal research: Quasi-experimental research design
- Cohort perspectives, **342–348**  
 age and period interrelations, 80–81  
 age pyramid, 345, 609  
 aging and, 342–344  
 alienation and, 100  
 as analysis tool, 346–347  
 childlessness and, 110–111  
 and cohort norm formation, 3064  
 compositional, 344–345  
 conceptual framework, 342, 343  
 contextual characteristics, 343–344  
 cross-sectional fallacy and, 1685  
 and cumulative fertility rates, 220  
 definitions of, 80  
 and disordered cohort flow, 345  
 divorce rate and, 1747  
 on economic well-being/fertility theory, 635  
 and fallacy of cohortcentrism, 343–344  
 intercohort, 344  
 intracohort, 343  
 on Japanese controlled immigration, 176  
 on juvenile crime, 189, 1488  
 labor force and, 3262  
 and life course, 1614, 1615, 1618, 1619, 2861–2862  
 and life table, 612, 614–615  
 and life-cycle transitions, 1623–1624, 1625, 1626  
 and longevity of successive cohorts, 345  
 population projection and, 616–617, 2181  
 population renewal theory and, 631–632  
 replacement fertility and, 2181

- research methods, 346–347  
social forecasting and, 2678  
social inequality and, 2691  
social security benefits and, 2803  
social structures and, 345–346, 3063–3066  
structural lag and, 3060–3061  
on student movements, 3068  
survey research, 2475  
*See also* Longitudinal research
- Colby, Anne, 1620–1621
- Colclough, Glenna, 2027
- Cold War  
Asian immigration and, 175  
and “clash of civilizations” hypothesis, 2940–2941  
and collapse of international systems, 2362  
democracy and, 601  
Iron Curtain of, 1934  
Latin American economic/political liberalization at end of, 1540  
military sociology and, 1876, 1878, 1882  
new nation-state formation at end of, 2362  
and nuclear annihilation fears, 3138  
peacemaking studies, 2047  
and peacetime conscription, 1876  
Polish Roman Catholic Church and, 2357  
and postcommunist transition, 2851–2852  
and power elite, 262  
protest movements and, 2271  
“strategic triangle” of, 331–332
- Cole, Charles L., 1727
- Colegio de Mexico, 1858, 1859
- Coleman, James S., 410, 606, 677  
on conceptualizing change, 1689–1690  
on corporate control, 443  
on crowd behavior, 559  
differential equation models, 1693  
on equality of opportunity, 758, 829, 830–831  
on labor union structure, 1533  
on money and social exchange, 1890  
as neoconservative, 1601  
and rational choice theory, 1791, 2335, 2452  
and representational models, 2029  
social capital concept, 366, 2637, 2638, 2639, 2640, 2733–2734  
social structural analysis, 2825, 2826  
on stochastic models of change, 2668, 2674  
and stochastic processes, 3036  
on white-collar crime, 3246–3247, 3248, 3250, 3251, 3252, 3253  
on youth culture, 512
- Coleman Report, 70, 830–831, 2931–2932, 2935
- Collaterality models, 1515
- Collective behavior, **348–354**  
altruism and, 114  
art production theory, 173  
best known forms of, 348–349  
bystander inaction and, 115  
conflict theory and, 415–416  
conformity and, 400–404, 2615–2617  
contagion theories of, 350–351  
crowds and riots as variants of, 349, 350–351, 553–560  
decision making and, 595–597  
differential association theory and, 665–666  
diffusion processes, 679  
disaster research on, 687  
emergent norm theory, 351, 354  
as facilitating individual desires, 349  
free rider and, 604  
group conflict resolution as, 1111–1115  
identity and, 1940  
interest groups and, 604  
intergroup attributions in, 197  
interpersonal identities in, 2221  
law as moral consciousness of, 1575, 1576  
leadership effectiveness and, 1571–1572  
leadership emergence and, 1567, 1570  
macro-level structural theories, 352–354  
micro-level convergence theories, 349–350  
micro-level interaction theories, 350–352  
nationalism and, 1940–1941  
nonconformity allowances in, 403–404, 606  
political participation and, 2339  
precipitating factors, 349, 353, 1940  
protest movements as, 2265–2271  
public opinion and, 2272  
rational choice theory and, 349–350, 2335, 2338–2339, 2341  
six determinants for, 353–354  
social belonging and, 2635  
social movements as, 2717–2725  
social networks and, 2734  
social psychology studies on, 2771  
and social values and norms, 2341  
values theory and, 352, 353, 3213, 3218–3219, 3222  
*See also* Complex organizations; Corporate organizations; Small groups
- Collective gaze, 3173
- Collective memory, 1636
- Collective representation, 2762
- Collective unconscious, 1714
- Collectivity. *See* Small groups; Social belonging
- Colleges and universities. *See* Higher education; *specific institutions*
- Collingridge, David, 2461
- Collins, Harry M., 2459
- Collins, Patricia Hill, 545
- Collins, Randall, 416, 777, 785–786, 1708, 1710
- Collor de Mello, Fernando, 2135–2136
- Cologne School, 1077, 1078

- Colombia  
 demographic transition, 626, 2178  
 drug cartels, 2135  
 drug crop-control program, 714  
 health-care system, 381  
 lawyers in, 478  
 political and social tradition, 1536
- Colonialism. *See* Imperialism, colonialism, and decolonization
- Color blindness (racial), 2246
- Colorado State University, 3230
- Colorectal cancer, 1640, 1641
- Colozzi, Ivo, 1472
- Columbia University, 681, 2193, 2756, 3099  
 voting behavior research, 3232, 3233–3234, 3238
- Column frequency and percentage, 658
- Colvin, Mark, 1498
- Coming White Minority, The* (Maharidge), 1580
- Command economy. *See* Central economy
- Commentary* (periodical), 1601
- Commercial and Political Atlas* (Playfair), 3005, 3006
- Commercial law, 473, 476
- Commercial sex workers. *See* Prostitution
- Commercial speech, restrictions on, 271
- Commission on National Goals, 2299
- Commission for Racial Justice of the United Church of Christ, 795
- Commission on Applied and Clinical Sociology, 326
- Commission on Human Rights, 2607
- Commission on Marijuana and Drug Abuse, 712
- Commission on Obscenity and Pornography, 274
- Commission Toward the Year 2000, 1039
- Commitment, as social control, 2658
- Committee on Economic Security, 2403
- Committee on Industrial Organizations, 1530
- Committee on Political Education, 1530, 1533–1534
- Committee on the Next Decade in Operations Research (CONDOR), 3105
- Commodity riots, 555, 556
- Common factor model, 908–910
- Common law system, 464, 465–472, 474, 475, 477–478, 480–482, 1554  
 and family violence, 981, 983  
 and marriage, 948  
 Montesquieu comparative study of, 1545  
 rape definition, 2576  
 social conditions conducive to, 1547
- Common Market, 548–549
- Common sense, 2100–2102, 2959
- Commoner, Barry, 1221, 1222
- Commons, J. R., 1310
- Communal riots, 555, 556
- Communality, 909, 912
- Communes, 459, 460
- Communication  
 and attitude change, 188  
 censorship and regulation of expression, 267–281  
 conversation analysis, 431–439  
 disaster reporting, 683–684  
 electronic, 408, 1321  
 emotions and, 2522  
 gendered differences in expression of love, 1700  
 Habermas theory, 543  
 history of graphic, 3005–3007  
 innovation diffusion research, 678  
 innovations in, 1346  
 instrumentalist theory of, 2219  
 intelligence and, 1380–1381  
 Japanese sociology studies, 1481  
 as marital adjustment factor, 1730, 1735–1736, 1737  
 model as vehicle for, 2029  
 as nationalism factor, 1940, 1942  
 nonverbal cues, 1976, 1978, 1980, 2061, 3096  
 persuasive, 188, 2776  
 in small-group problem solving, 2619  
 and social networks, 2734  
 symbolic interaction theory of, 2767–2768  
*See also* Internet; Language; Mass media; Mass media research; Sociolinguistics
- Communication of Innovations* (Rogers and Shoemaker), 678
- Communication-persuasion paradigm (attitude change), 188
- Communism. *See* Socialism and communism
- Communism, Conformity, and Civil Liberties* (Stouffer), 314, 316–317
- Communist Church, 355
- Communist Manifesto* (Marx), 2846
- Communist Party (China), 2137
- Communist Party (Italy), 2128, 2129
- Communist Party (Russia), 2136, 2357, 2850  
 and Soviet sociology, 2981
- Communitarian Challenge to Liberalism, The* (Kukathas), 356
- Communitarianism, **354–362**  
 as alternative to alienation, 104  
 authoritarian, 356  
 on civic society, 360  
 community and, 359–361  
 of countercultures, 460–461  
 critics and responses, 360–361  
 history of, 355–356  
 impact of, 361–362  
 kinship systems and, 1509, 1511–1512  
 new school of responsive, 356–361  
 political theory of, 356, 2337  
 primary contribution of, 361  
 utopian designs and, 3202–3203  
 values and, 3218–3219, 3222
- Community, **362–369**  
 agenda-control power in, 2165–2166  
 alienation from, 100  
 as anachronistic, 360  
 anomie and, 164–165

- belonging and, 2520, 2521  
 case studies of, 243, 244–245  
 and cities, 305  
 civil liberties balanced with  
   needs of, 318  
 classic perspective on, 362–363  
 communitarian view of, 358–361  
 as community of  
   communities, 359, 363  
 countercultures as, 459–463  
 and criminal sanctions, 515–  
   516, 537  
 definitions of, 362–363  
 disaster planning and reaction,  
   683, 684, 685  
 “disorganization” sociological  
   perspective on, 665  
 drug-abuse prevention programs,  
   715–717  
 education and, 2927  
 expressive needs of, 272  
 gay and lesbian, 2570  
*Gemeinschaft* concept vs., 2630  
 Japanese sociology studies, 1480,  
   1481, 1483  
 leisure and, 1584–1585  
 long-term care facilities,  
   1656–1657, 1658  
 macro-level deviance theories  
   and, 662, 663–666  
 and mass society theory, 1770,  
   1772, 1773–1774  
 moral infrastructure of, 357–  
   358, 359  
 and organizational functioning,  
   2006–2007  
 participation in decision  
   making, 605  
 participatory research and,  
   2038–2042  
 policing program, 716, 2114  
 political communitarianism  
   and, 356  
 power structure in India, 252–253  
 public health campaigns, 1640  
 responsive communitarianism  
   and, 358–361  
 rural sociology studies, 2428–  
   2429, 2431  
 “saved” argument, 367  
 and school curricula control,  
   277–278  
 settlement houses, 365–366, 2841  
 social capital and, 366  
 social disorganization theory  
   of, 1495  
 social mobility and, 2714  
 social networks and, 664–  
   665, 716  
 in social reform context, 365–366  
 social theory and transformation  
   of, 366–368  
 as socialist ideal, 2847, 2849  
 and “strength of weak ties”  
   hypothesis, 2693, 2731–2732,  
   2791, 2792, 2827  
 studies of, 363–365, 2220–2221  
 and territorial belonging,  
   3129–3132  
 urban sociology and, 3191–3197  
*Community and Power* (Nisbet), 356  
*Community and Society* (Tönnies),  
   2520  
 Community health. *See* Comparative  
   health-care systems; Health  
   promotion and health status;  
   Medical sociology  
 Community hospitals. *See* Hospitals  
 Community of limited liability  
   thesis, 367  
 Community Partnership Program,  
   716, 717  
 Community service, 2254  
   as criminal sanction, 515  
 Community supervision, as criminal  
   sanction, 515  
 Community surveys, 363–365, 1834  
 Companionship family, 1502, 1506  
 Comparable worth, **369–373**, 2141,  
   2706, 3048, 3265  
 Comparative health-care systems,  
   **373–383**  
   frameworks for comparison,  
     376–377  
   key characteristics, 374–375  
   long-term care, 375, 378, 1655,  
     1659–1661  
   review of selected systems,  
     377–381  
   and U.S. medical-industrial  
     complex, 1826–1829  
 Comparative judgment, 597–598  
*Comparative Politics* (Almond and  
   Powell), 2917  
 Comparative properties of  
   collectives, 1592  
 Comparative-historical sociology,  
   **383–392**  
   advantages of, 2918–2919  
   analytic types of, 385–388  
   civil vs. common law tradition,  
     472–481  
   court systems and law, 473–481  
   holistic comparisons, 387  
   legal systems, 1545–1551  
   and macrosociology, 1709  
   and origins of political parties,  
     2154–2155  
   on revolutions, 383, 384, 2413,  
     2414, 3000–3001  
   urban features, 3194  
   *See also* Convergence theories  
 Compensatory control, 518  
 Competition  
   coalitions and, 330–331  
   and sex selection theory,  
     2885–2886  
   and social exchange theory, 2671  
   and social values research, 3220,  
     3222–3223  
 Complementarity. *See* Mate selection  
   theories  
 Complete life table, 613  
 Complex organizations, **392–400**  
   in American society, 143–144  
   and bureaucracy, 229–235  
   corporate, 244, 441–446  
   definitions of organizations,  
     393–394  
   demographics of, 395–397  
   division of labor in, 696  
   environments of, 394–395  
   internal diversity sources among,  
     397–398  
   language use in, 2901–2902  
   large law firms as, 469–470  
   and political power, 2163,  
     2997–2998  
   structure of, 2002–2014  
   survival predictors, 397

- transnational, 3174–3180
- Complexity theory, 2753
- Compliance and conformity, **400–406**
- anomie concept and, 165–166
  - belief and behavioral typologies, 2616
  - classic experiments, 401–402
  - collectivism and, 3218
  - control theory and, 535–536
  - cultural conformity vs. social belonging, 2630
  - definitions of, 400
  - deviant typologies vs., 669
  - factors increasing, 402–403, 2776
  - Frankfurt School on sources of, 540
  - and helping behavior, 2774
  - and intolerance of deviance, 317
  - normative influences, 523–524, 2094, 2658
  - normative sanctions, 515, 2341
  - persuasion and, 2094
  - political correctness and, 2138–2142
  - self-presentation and, 2506
  - in sentiments, 2529
  - small group, 2615–2617, 2776
  - social belonging and, 3131–3132
  - social exchange theory and, 2670–2671
  - social networks and, 2732
  - “stakes in conformity” theory, 665, 667
  - values and, 3216
  - See also* Deviance theories; Nonconformity; Social values and norms
- Composite scale, 1909
- Comprehensive Drug Abuse and Control Act of 1970, 713
- Compulsory education, 742, 2056
- Computational graphics (“nomographs”), 3003
- Computer applications in sociology, **406–414**
- affect control theory, 45–46
  - bootstrapping and jackknifing, 2397, 2449, 2678, 3039
  - and capitalism, 240
  - causal modeling, 266
  - census data, 283
  - content analysis, 418–421
  - covariance structure models, 3037
  - data banks and depositories, 575, 580
  - disaster planning and management, 684, 685, 686
  - mathematical simulations, 1790
  - for measures of association of more than two variables, 1812
  - nonparametric test software, 1971
  - and observation systems, 1981
  - qualitative, 409–410
  - and replication, 2397
  - sampling standard error detection, 2449
  - and secondary data analysis, 2481
  - SEM (structural equation modeling), 908, 1692, 1910, 1914–1915, 1918–1923, 2346–2347, 3039
  - statistical graphics, 3003, 3003–3022, 3019–3020, 3022
  - statistical packages, 3035, 3038, 3039
  - systems theory, 3103
  - telephone interviewing, 410, 1802, 3092
  - typologies, 3183–3185
  - See also* Information society; Internet
- Computer Assisted Telephone Interviewing (CATI), 410, 1802, 3092
- Computerized Self-Administered Questionnaires, 410
- “Computing in the Social Sciences” (annual conference), 407
- Comte, Auguste, 734, 818, 1024, 1028, 1031, 1423, 1465, 1466, 1857, 2265
- on altruism, 114, 2882
  - as clinical sociology precursor, 327, 1029
  - on historical progress, 2644
  - as Japanese sociology early influence, 1477
  - macro-level phenomena concerns, 1704
  - as Polish sociology influence, 2118
  - on positivism, 2192, 2193, 2217
  - and secularization, 2483
- “Concepts of Culture and Social System, The” (Kroeber and Parsons), 565
- Conciliatory control, 518
- Concubines, 2601, 2602
- Concurrent schedules of reinforcement, 212–213
- Concurrent validity, 3208
- Condition of Education, The* (report), 265
- Conditional distributions, 2250
- Conditional mean, 447
- Conditional probabilities, 3111–3112
- Condom use, 957, 2559, 2560, 2586, 2587, 2590, 2592, 2593
- See also* Family planning
- CONDOR (Committee on the Next Decade in Operations Research), 3105
- Condorcet, Marquis de, 2339
- Conference for Security and Cooperation in Europe, 2047
- Confidence intervals. *See* Statistical inference
- Confirmatory factor analysis, 907, 915–918, 1920–1921
- Conflict
- definition of, 1111, 1451
  - global security and, 1222
  - industrial, 1312, 1314
  - See also* Conflict theory; Violence: War
- Conflict management, 1015
- group resolution, 1111–1116, 1400, 1401
  - leadership strategies, 1569
  - negotiation of power, 1950–1955
- Conflict subcultures, 509–510, 512
- and criminal sanctions, 516–517
- Conflict theory, **414–417**
- coalitions and, 330–332
  - cognitive consistency theories and, 334–335

- collective behavior theories and, 351–352
- competing interests and, 415
- criminological, 516–517, 535, 536, 2659
- and educational mobility, 756, 758, 2928
- of emotion, 2522–2523
- family and religion and, 939
- and German sociology, 1078
- intergroup and  
interorganizational relations  
and, 1401
- and juvenile delinquency, 1497–1498
- macro-level reactions to  
deviance and, 670
- as major social order  
school, 2337
- and penology, 2054–2055
- and Polish sociology, 2119–2120
- and political power, 2997
- on sexual behavior, 2537, 2538
- and social control, 2660
- and social problems, 2762
- sociolinguistics and, 2905–2906
- sports sociology and, 2989
- war and, 415, 3243–3244
- See also* Game theory and  
strategic interaction; Marxist  
sociology
- Conformity. *See* Compliance and  
conformity
- Confucianism, 1477, 2941
- Confucius, 1564
- Conger, J. A., 1566
- Congleton, Roger D., 2273
- Congo River Basin, 60
- Congregationalism, 3227
- Congress of Racial Equality, 2269, 2495, 3069, 3070
- Congress of the United States, 1954  
corruption scandals, 2127–2128  
direct election of members, 602  
and gays in the military  
issue, 1881  
and human rights issues, 2607  
power shift to executive branch  
from, 2624  
and public opinion, 2276–  
2277, 2278
- Congress of Vienna (1815), 1933
- Congress Party (India), 2486
- Congruency theory, 335, 336–337
- Conjugal love, 1699
- Conjunction fallacy, 595
- Conklin, Agnes, 324
- Connor, Walker, 1941, 1942
- Consciousness. *See* Phenomenology
- Conscription (military draft), 1876–1877  
protests against, 2269–2270, 3069
- Consensus model  
communitarian, 359  
of criminal law, 516  
of religious organizations,  
2379–2380
- Conservation movement, 802–803
- Conservatism. *See* Liberalism/  
conservatism
- Conservative Party (Great  
Britain), 2130
- Consistency, as reliability  
component, 2343, 2347–2350,  
2352–2353
- Consistency theory. *See* Cognitive  
consistency theories
- Consolidated metropolitan statistical  
area, 307
- Consolidated Omnibus Budget  
Reconciliation Act of 1985, 1143
- Consortium of Social Science  
Associations, 153
- “Conspicuous consumption”  
concept, 3168
- Constantine, John, 1729
- Constitution, U.S.  
banning of slavery, 2601  
Bill of Rights, 270, 315, 359  
census directive, 282  
civil liberties guarantees, 315, 316  
freedom of expression  
guarantees, 268, 270–273,  
275–276  
and protest movements,  
2265, 2269  
public ignorance of rights  
under, 318  
as shared value, 359
- Constitutional monarchy, 2356
- Constitutional personality theory,  
1717–1718
- Constitutional Revolution of  
1905–1911 (Iran), 1871
- Constitutions, 3000–3001
- Construct validity  
attitudes as indicators, 190  
in long-term longitudinal  
studies, 1690  
quasi-experimental research  
design, 2324–2325  
reliability and, 2346  
validity and, 3208, 3210
- Construction de la sociologie, La*  
(Berthelot), 1028
- Constructionist perspective  
and collective action, 354  
and emotions, 2523–2524  
of environmental sociology,  
810–811  
of epistemology, 823–824  
and ethnicity, 849  
of holistic personality theory,  
2088  
and medical sociology, 1813–  
1814, 1815–1816  
of popular culture, 2169,  
2172–2173  
and scientific knowledge, 2459  
and sexual behavior, 2537–2538,  
2566–2567  
and sexual orientation,  
2566–2567, 2569  
and social problems, 2762–2763  
and socialization, 2855
- Consumer debt, 206–207
- Consumer Price Index, 2213–2214
- Consumer Reports* (periodical), 864
- Consumerism  
and health care industry, 1826  
media messages of, 1699  
and postmodernism, 2200, 2205,  
2206, 2208–2209  
pressure against products made  
by child labor, 2607  
survey data collection, 575  
tourism as, 3168
- Consumption  
and economic sociology, 736

INDEX

- and environmental sociology, 800, 807
- Contagion theories, 350–351, 3080, 3081
- Contemporary Sociology* (journal), 1606, 1869–1870
- Contempt of court, 478
- Content analysis, **417–422**, 1978–1979, 1981, 2468
- Content validity, 3207, 3208–3209
- Contested Knowledge* (journal), 2207–2208
- Contextual properties of collectives, 1592
- Contingency of reinforcement, 210–212
- Contingency tables. *See* Tabular analysis; Typologies
- Continuing care retirement communities (CCRCs), 1664
- Continuing quality improvement (CQI), 1665
- Continuous National Surveys (CNS), 577, 578
- Contraceptives. *See* Family planning
- Contract with America, 2278
- Contract workers, 321, 2597
- Contracultures. *See* Countercultures
- Contribution to the Critique of Political Economy, A* (Marx), 722–723
- Control balance theory of deviance, 535
- Control theory. *See* Social control
- Convenience samples, 2444
- Convention for the Prevention and Punishment of the Crime of Genocide, 1066, 1072
- Convention on Combating Bribery of Foreign Public Officials (1999), 2138
- Convention on Psychotropic Substances (1971), 713
- Convention on the Rights of the Child (1989), 1243
- Convergence theories, **422–431**
- crowd behavior, 679
  - demographic transition, 425–426
  - family patterns, 426
  - globalization and, 427–428
  - in industrial sociology, 422–423, 424–425, 428
  - microlevel behavior, 349–350
  - modernization and, 422–423
  - social security systems, 2800–2801
  - urban organization and, 3197
  - welfare state, 426–427
- Convergent validity, 3210
- Conversation analysis, **431–441**
- case studies of, 247
  - content analysis, 1979
  - and gender identity, 1002
  - sequence organization, 433–435
  - sociolinguistics and, 2895, 2902–2903, 2904–2905
  - SYMLOG observation system, 1976–1977
  - on turn-taking, 247, 435–437, 438, 439
- Converse, Philip E., 2683, 3234, 3235
- Cook, Karen S., 2672, 2674, 2733
- Cook, T. D., 2324–2325, 2326, 2327
- Cook, Thomas, 2282
- Cook County, Illinois, 1485
- Cooley, Charles Horton, 783, 1254, 1313, 2089, 2218, 3098
- “looking-glass self” concept, 2089, 2344, 2512, 2750, 2856
  - pragmatism theory, 2218, 2219, 2221, 2423
  - on primary groups, 2610, 2611
- Cooper, Anna Julia, 66
- Cooper, J., 339–340
- Cooperation
- coalitions and, 329
  - among complex organizations, 393
  - within in-groups, 3218
  - and social values research, 3220, 3221–3222
- Cooperative Extension Service, 2416, 2429
- COPE. *See* Committee on Political Education
- Cope, Edward, 1–2
- Coping, 1814
- personal dependency and, 2062–2063, 2066
  - role theory and, 2415
  - self-destructive responses, 3077
- CORE. *See* Congress of Racial Equality
- Cornfield, Daniel B., 1533
- Cornish, Edward, 1038
- Corporal punishment, 76, 315
- as criminal sanction, 515
  - and family violence, 982–983, 985
- Corporate organizations, **441–446**
- career advancement in, 1984–1985, 1986
  - case studies of, 244
  - case studies of multinational, 246
  - CEOs of, 1564
  - democratic decision making in, 601
  - dominance in American society, 143–144
  - European labor movements, 1529, 1530–1532
  - as governing class, 604
  - health care industry, 1810–1829
  - and interest groups, 2151
  - legal rights of, 441
  - Mexican political, 1858, 1859
  - owner vs. managerial control of, 443
  - social control over, 443–444
  - and social networks, 2727–2728
  - transnational, 3174–3180
  - and white-collar crime, 3246–3254
- See also* Complex organizations
- Corporatist welfare state, 377, 379
- Corpus Juris Civilis*, 465, 473
- Corrections systems. *See* Criminology; Criminal sanctions; Penology
- Correlation and regression analysis, **446–459**
- attenuated correlations, 1909
  - differential equation models, 1693–1694
  - intercorrelations among items, 2345, 2348, 2354
  - linear regression, 162–163, 2251
  - and measurement, 1795, 2345
  - and measurement error, 1917
  - misleading results, 457
  - multiple indicator models, 1907–1923



- multiple regression, 452–456  
 nonparametric statistics and, 1957, 1969  
 statistical methods and, 3035–3036, 3039  
 in time series, 2679  
*See also* Causal inference models
- Correlation coefficient  
 for bivariate relationship, 661  
 definition of, 661  
 measures of association and, 1907–1908  
 and multiple indicator models, 1907–1908  
 and nonparametric statistics, 1957  
 scatter plot shape and direction, 661  
 statistical graphics, 3015–3016
- Correspondent inference theory, 193–194
- Corruption. *See* Political and governmental corruption
- Corsaro, William, 2297, 2855, 2901
- Corse, Sarah M., 1647, 1648, 1649, 1650
- Corsini, Raymond J., 1718, 2084
- Cortese, Charles F., 2500
- Cosby, William, 273
- Coser, Lewis A., 415, 1356, 1646, 2522–2523, 2817, 3039
- Cosmopolitanism, 3132–3133
- COSSA. *See* Consortium of Social Science Associations
- Costa, Paul T., 2087
- Costa Rica, 1536
- Cost-benefit analysis  
 criminal sanctions and, 519–520  
 public policy and, 2283
- Cost-effectiveness analysis, 2283
- Costner, Herbert L., 3188
- Cottage industry, 3262
- Cottrell, Leonard, 324, 1237, 1726, 1727, 3098
- Couch, Arthur, 1976
- Couch, Carl, 2222
- Coughlin, Richard, 360
- Council of European Social Science Data Archives, 575, 576–577
- Council on Social Work Education, 2845
- Countercultures, **459–463**, 2366, 2367, 2374  
 and mysticism, 460–461, 2969, 3287–3288  
 outcomes of movements, 2724  
 student movements, 3070
- Countermovements, 2717–2718
- Counter-urbanization, 311
- Country Life Movement, 2426
- Counts, Dorothy and David, 583
- County of Washington v. Gunther* (1981), 372
- Coups d'état, definition of, 2410
- Cours de philosophie positive* (Comte), 2192
- Course in General Linguistics* (Saussure), 1032
- Court systems and law, **464–483**  
 accusatory vs. inquisitorial models, 479–480  
 appeals court, 471–472, 476  
 civil and criminal procedures, 479–480  
 determinate sentencing, 2056  
 and division of powers, 1953  
 jurisprudential vs. sociological perspectives, 2961  
 juvenile, 1485  
 and legal justice, 2696  
 legal precedent, 476  
 litigation comparison by country, 471  
 and nonintervention in family law, 949–950, 951  
 probation and, 2254  
 triadic dilemma, 465–466  
*See also* Law and legal systems; Sociology of law
- Courtly love, 1697
- Courts of Indian Offenses, 137
- Courtship, **483–490**  
 advice books, 488  
 after divorce, 1779  
 and cohabitation, 108, 484, 487, 1779  
 history of, 483–484  
 love and, 484, 488, 489, 1698–1699  
 and mate selection theories, 1775  
 research trends, 488–489  
*See also* Sexual behavior patterns
- Cousins, 1509, 1513
- Covariance, 449  
 adjustments for, 161–162  
 attribution model, 193  
 for bivariate relationship, 661  
 definition of, 661  
 in event history models, 1693  
 structure models, 3037  
*See also* Analysis of variance and covariance; Factor analysis
- Covington, Martin V., 2859
- Cox, David, 872
- Cox, Oliver C., 53
- Cox regression. *See* Proportional hazards models
- Cox-Stuart change, 1969
- CPI. *See* California Psychological Inventory; Consumer Price Index
- Crack cocaine, 711, 714  
 civil liberties issues, 318  
 and sexually transmitted diseases, 2588
- Craft guilds, 697
- Craft unions, 1533
- Crámer, Harald, 3035
- Cramer's *V* (measure of association), 1809
- Crane, Diana, 567, 568, 569, 677, 2915
- Crane, Robert, 173
- Craven, Avery, 1217
- Craxi, Bettino, 2129
- Creation stories, 3285
- Creativity  
 depression linked with, 655–656  
 goal-relevant stimuli and, 2060  
 pragmatism and, 22220
- Credentials theory, 2929–2930, 2933–2934
- Credit. *See* Bankruptcy and credit
- Credit cards, 206–207
- Crenson, Matthew, 2165–2166
- Crespi, Irvin, 2274
- Cressey, Donald, 503, 507, 2020, 2052–2053
- Cresson, Edith, 2130

- Crews, Frederick, 2023, 2026
- Crime, theories of, **502–509**, 527–537  
 anomie and, 166  
 communitarian, 359  
 community studies, 365  
 and consensus view, 516  
 control balance, 535  
 criminal behavior approach, 506–507, 513, 528–529, 1575  
 definitions of, 529–528  
 deterrence and, 2659  
 deviance and, 664–672, 2658  
 differential association, 666  
 and distinct forms of crime, 505–506  
 and functional perspective, 516  
 low self-control vs. diminished social control, 667  
 macro-level origins, 662, 663–666, 672  
 macro-level reaction, 670–671  
 micro-level origins, 662, 663, 666–670  
 micro-level reaction, 663, 668–670  
 penal policy and, 2056–2057  
 pornography and, 505, 2186, 2187  
*See also* Criminal and delinquent subcultures; Criminal sanctions; Criminology; Juvenile delinquency, theories of
- Crime deterrence. *See* Criminal sanctions; Criminology; Social control
- Crime in the United States* (FBI publication), 493–494, 503
- Crime rates, **490–502**  
 in cities, 311, 531  
 criminological studies of, 530–533  
 criminological theories on, 503–504  
 international data of, 498–500, 549  
 measuring absolute vs. relative rates, 497–498
- Crimes against humanity, 1429
- “Crimes of obedience,” 404
- Criminal and delinquent subcultures, **509–515**  
 characterizations of criminals, 530–531, 534, 1517  
 deviance theory, 664, 665–668, 670  
 drug users, 713–714  
 factors in, 1494  
 labeling process, 668  
 “prisonization” and, 2052  
 social control approach to, 521  
 stable vs. retreatist, 1494
- Criminal justice system  
 criminal sanctions, 515–521, 526–527, 537, 2056  
 data from, 530  
 defendant’s rights and, 317, 318  
 determinate sentencing, 2056  
 due process and, 520–521  
 probation and parole, 2242–2258  
 and rape prosecution and defense, 2577  
 selection bias in disposition of cases, 2439  
 as social control, 2660  
 sociology of law and, 2960–2963  
 trials, 479–480  
*See also* Penology
- Criminal law  
 determinate sentencing, 2056  
 discrimination in sentencing, 2962  
 evolution of, 516–519  
 and family violence, 981  
 procedure, 479, 480  
 and subgroup variation, 527  
*See also* Criminal sanctions; Criminology
- Criminal sanctions, **515–522**  
 deterrence theory of, 519, 525, 529, 536, 537, 667, 2056–2057, 2341, 2650, 2659  
 deviance theories on, 666–667, 667, 670–671, 1575  
 discrimination in, 2962  
 and due process, 520–521  
 evolution of, 518, 526, 527–528  
 illicit drug use and, 712, 714  
 and labeling theory, 534  
 probation and parole and, 2252–2258  
 retribution and, 2056  
 sample selection bias, 2439  
 three dimensions of, 2659  
 for white-collar crime, 3251–3252  
*See also* Criminalization of deviance; Police
- Criminal Victimization in the United States* (report), 494
- Criminality and Economic Conditions* (Bonger), 534
- Criminalization of deviance, **523–527**  
 decriminalization trend, 521  
 drug traffic and use, 711, 713–714, 1577  
 Durkheim on, 1575, 1577  
 historical theories, 528–529, 1575  
 labeling theory, 243, 520, 525, 534–535, 1577  
 macro-level reactions, 669–670  
 micro-level reactions, 662, 668–669  
 politicization of, 524–526  
 prostitution studies, 2559–2560  
 sanctions and, 667, 670–671  
 sexual behaviors, 2567–2568  
 sociology of sanctions, 516–518
- Criminology, **527–539**  
 Chicago School, 532–533  
 Classical School, 528–529, 535, 536  
 and correlates of crime, 530–531  
 and cross-border crime control, 1935–1936  
 and cross-cultural analysis, 549  
 and definition of crime, 529  
 deterministic, 521  
 mainstream, 504  
 and organized crime, 2017–2021  
 penology, 2051–2057  
 radical-Marxist, 504–505, 534–535  
 and rape, 2576–2581  
 and white-collar crime, 530, 3245–3254  
*See also* Crime, theories of; Criminal sanctions; Juvenile delinquency, theories of; Juvenile delinquency and

- juvenile crime; Political crime; Terrorism
- Crimmins, E. M., 1632
- Crippen, T., 2881, 2883, 2884, 2886
- Crisis, definition of, 2025
- Crisis responses, collective, 348
- Crisis theory, 1755  
on political party origins, 2154
- Critcher, C., 353, 1578
- Criterion-related validity, 3207, 3208
- Critical legal studies (CLS)  
movement, 1548, 1556–1557
- Critical naturalism, 823
- Critical realism (Bashkar concept), 823
- Critical sociology, 1081
- Critical theory, **539–546**  
criminology, 504–505  
Frankfurt School, 539–542, 543, 1752, 1754, 1757, 1758  
of Habermas, 539, 542–544, 545  
human ecology and, 1214  
legal studies, 2961  
leisure and, 1583  
Marxist sociology and, 539, 545, 1754, 1757, 1758, 2760  
materialist theory and, 1785–1786  
medical sociology and, 1815–1816  
and participatory research, 2040  
political elite/pluralist debate, 2624–2626, 2627  
on popular culture, 2169  
on postindustrial society, 2205  
post-Marxist, 1757, 1758  
and poststructuralism, 544–545  
pragmatism and, 2219  
and rural sociology, 2431  
on scientific explanation, 2472  
social problems paradigm, 2760  
and sports sociology, 2990
- Crittenden, Kathleen, 196
- Croatia, 2362  
genocide, 68  
national movement, 1941
- Croce, Benedetto, 1465
- Crohn, Joel, 1411, 1412
- Croker, Richard, 2125
- Cronbach, Lee, 2348
- Cronbach's alpha, 2348–2350, 2351
- Cross-cultural analysis, **546–553**  
of affective responses, 42  
of altruism, 118  
anthropological, 547–548, 550, 2888, 2893  
of attribution, 194, 198  
borrowings in, 675  
challenges and problems in, 549–550  
changes theories, 674–675  
of childhood sexual abuse, 292–293  
of cognition, 2891  
of communitarian moral judgments, 361  
of conformity levels, 400, 404  
of depression manifestation, 656  
disaster research, 685  
of educational attainment, 3044–3045  
of family trends, 130, 1502  
future of, 550  
of health-care systems, 375–376  
of homosexual behavior, 111  
of income determinants, 3048  
of kinship structures, 1502, 1509  
of legal systems, 1549–1550  
of love, 1697–1698  
of mate selection, 2885  
methodological techniques and sources in, 547–549  
of moral development, 1901, 1902–1903  
of occupational status attainment, 3046–3047  
prestige evaluations similarities, 1998–1999  
of rape-prone societies, 2580–2581  
of religious ethics and moral law, 3284  
of religious myths and symbols, 3282–3283  
secondary data analysis, 2478–2479  
of socialization, 2862  
Weber's bureaucracy study, 229–230  
of widowhood, 3255–3256
- Cross-dressing, 2572, 2573
- Cross-gender sexual contacts. *See* Heterosexuality
- Cross-impact matrices, 2678
- Crossley, Archibald, 3232
- Cross-modality fallacies, 1593
- Cross-Polity Survey, A* (Banks and Textor), 2917
- Cross-pressure theory, 3049–3050
- Cross-sectional fallacy, 1593, 1685
- Cross-sectional surveys  
cohort perspectives, 344  
logic model, 2297  
longitudinal, 1687, 1688–1689  
marginal employment, 1722
- Crowds and riots, **553–562**  
characteristic features, 554–557  
collective behavior theory, 349, 350–351  
convergence theory, 679  
coordination of behavior of, 559–560  
crowd activities, 554–555  
diffusion theories, 679  
emergence conditions, 557–558  
participation units in, 556–557  
predictors of participation in, 560  
social psychology studies on, 2771  
spectator-bystander differentiation, 556  
*See also* Draft riots; Urban riots
- Crowley, M., 2533
- Crude rates, 610–612, 1740–1741  
standardization, 2992–2995  
*See also* Birth and death rates; Marriage and divorce rates
- Crump, Edward (“Boss”), 2126
- Crusades, 2967–2968
- Crutchfield, Robert, 665
- Cruzan, Nancy, 586
- Cruzan v. the State of Missouri* (1990), 3083, 3084
- CS. *See* Cluster sampling
- CSA. *See* Clinical Sociology Association
- CSAQ (Computerized Self-Administered Questionnaires), 410
- Cuba, 2134

- and criminalization of deviance, 524  
 revolution (1959), 2411, 2412  
 as socialist state, 1536, 2851
- Cuber, John, 1731
- Cui bono* criterion, 2004
- Cults  
 as collective behavior, 349  
 counterculture, 461  
 definition of, 2366  
 and recovered memory syndrome, 901  
 religious, 3287
- Cultural analysis, 2762, 2891–2892
- Cultural anthropology. *See* Anthropology; Sociocultural anthropology
- Cultural capital, 2626, 2714  
 and educational mobility, 2928, 3043  
 human capital vs., 2928  
 as stratifying force, 2812
- Cultural conformity, social belonging vs., 2630
- Cultural differentiation, 510
- Cultural diversity, 1407
- Cultural evolution. *See* Evolution: biological, social, cultural
- Cultural hegemony theory, 1753
- Cultural identity, 1199, 1402
- Cultural imperialism, 1322, 1767
- Cultural institutions, 1063
- Cultural integration, 1223
- Cultural lag, 3066
- Cultural markers, 1932
- Cultural pluralism. *See* Multiculturalism
- Cultural production. *See* Art and society; Literature and society; Music
- Cultural psychology, 548
- Cultural renewal, 1295
- Cultural resistance studies, 2170
- Cultural Revolution (China; 1966–1977), 3045
- Cultural selection, 1234
- Cultural specificity, 1368
- Cultural studies, 569  
 Marxist sociology, 1756, 1757, 1758  
 on popular culture, 2169–2170, 2172  
 pragmatism and, 2219  
*See also* Ethnography; Sociocultural anthropology
- Cultural Studies: A Research Annual* (journal), 2293
- Culturalists, 563
- Culture, 562–572**  
 African institutions, 63–65  
 age appropriateness and, 1623  
 alienation and, 100  
 American Indian, 137–138  
 art and, 172–174  
 Asian-American personality studies, 178  
 British sociological studies, 226, 227–228  
 coethnic, 2329  
 conformity levels and, 400  
 criminal sanctions and, 518  
 definitions of, 562–563, 566–567  
 as depression influence, 656  
 development of, 1199  
 deviant behavior linked with, 664  
 differing concepts of corruption, 2124, 2126–2127  
 diffusion theories, 675–676, 679, 1085  
 disaster reaction and, 686  
 emic-etic frame, 550, 564, 2091–2092, 2889, 2891, 2892  
 emotions and, 774, 781–785  
 and ethnic-group resources, 847–848  
 ethnoscience and, 2891  
 Eurocentric, 55  
 global, 1085, 1091, 1092  
 high vs. mass culture debate, 565–566, 1645–1646  
 historical perspective, 568  
 of honor, 2528  
 and Latin American studies, 1537  
 legal systems as reflection of, 472  
 leisure as component of, 1583  
 marginalization, 2367, 2634–2635  
 Marxist sociology and, 562, 568, 1755, 1756, 1757  
 mass media mainstreaming of, 1766  
 and mental health concepts, 2190  
 metatheory and, 1853–1854  
 modernization theory and, 1885–1886  
 moral development theory, 1900–1902  
 music and, 1924–1927  
 nonmaterial, 1210  
 order-deficit model, 153  
 personal, 1305  
 postindustrial, 2200  
 postmodernist theories of, 1756, 1784, 2173, 2206–2209  
 production of, 567, 1647–1648, 1925  
 recorded, 568  
 regulatory, 1103  
 religious orientation and, 2384–2385  
 role theory and, 2422  
 sexual orientation and, 2566–2567, 2569  
 of sexuality, 1304  
 and social change, 2643, 2644  
 social interaction definition, 2767  
 and social movement emergence, 2719–2720  
 and social structure debate, 563–565  
 and socialization, 2855, 2862  
 as spousal role influence, 696  
 subcultures definition, 509  
 and suicide variation, 3079  
 as symbolic aspect of social life, 518  
 technological, 1210  
 as tool kit, 1102  
 and values and norms, 2830, 3217  
 and values in economy, 736, 737  
 and variants of values, 3217, 3218, 3222, 3224  
*See also* Countercultures; Cross-cultural analysis; Ethnicity; Evolution: biological, social, cultural; Popular culture; Sociocultural anthropology
- Culture* (Czarnowski), 2118
- Culture and Evolution* (Chmielewski), 2119

- Culture and Personality Approach in American Anthropology, The* (Mach), 2119–2120
- Culture circles hypothesis, 675
- Culture exchanges, 675
- Culture lag, 2644
- Culture of poverty, 1288, 2211–2212
- Culture war, 1580
- Culture: Sociological Perspectives* (Hall and Neitz), 567
- Cumulative distribution function, 2250
- Cumming, Elaine, 2300
- Cumulative advantage/disadvantage theory, 84
- Cumulative scale analysis, 1801
- Cunnilingus, 2553
- Cunningham, P. B., 76
- Currency. *See* Money
- Current Concepts of Positive Mental Health* (Jahoda), 2188–2189
- Current life table, 612, 614
- Current Population Survey, 284, 1521, 1687, 1720, 1722, 1747, 2215
- Current Research in Social Psychology* (journal), 413
- Curry, Theodore R., 2961
- Curtis, R. F., 2345
- Cuvillier, A., 1024
- Cybercrime, 3253
- Cyberspace. *See* Internet; Web sites
- Cyclical theories, 2644–2645
- Cyprus, 1945
- Czarniawska, Barbara, 2902
- Czarnowski, Stefan, 2118
- Czech Republic
  - divorce rate, 706
  - ethnonationalism and creation of, 1934, 1946, 1947
  - labor movement, 1532
  - post-communist transition in, 2136
  - sociology in, 2117
  - tourism in, 3169
- Czech revolution (1968), 2414
- Czyzewski, Marek, 2119
- D**
- Dadrian, Vahakn N., 1071
- Dahl, Robert A., 603–604, 1456, 2624
- Dahlström, Edmund, 2453
- Dahrendorf, Ralf, 414, 1753, 2337, 2415, 2814
  - and British sociology, 224, 225–226
  - and German sociology, 1078, 1079, 1080
- Daily time use. *See* Time use research
- Daley, Richard J., 2126, 2496–2497
- Dalit (Hindu oppressed), 252–253
- Daly, Kathleen, 3251
- Damaska, Mirjan, 1548, 1549
- D'Amato, Alfonse, 2125
- Damle, Y.B., 1291
- D.A.R.E. (Drug Abuse Resistance Education), 716, 717
- Darkness Visible: A Memoir of Madness* (Styron), 651
- Darley, John M., 115, 116–117
- Darrow, Clarence, 2369
- Dartmouth conference, 2047
- Darwin, Charles, 1031, 1209, 1466, 1636, 2218, 2522
  - evolution by natural selection theory, 876, 878, 2330, 2334, 2369, 2418, 2880–2881, 2885, 2889
  - evolution theory statements, 2880–2881
  - as influence on Marx, 573, 1782
  - as Polish sociology influence, 2118
- Data analysis
  - computing technology, 408–409
  - content analysis, 417–421
  - cross-cultural, 547–549
  - demographic, 609–610, 632
  - general linear model, 457
  - longitudinal, 1684, 1685–1686, 1689–1694
  - missing data treatment, 3039
  - multiple indicator models, 1907–1923
  - multiple regression, 451–457
  - replication, 2396–2397
  - sampling procedures, 2444–2449
  - secondary, 2473–2481
  - of social networks, 1789–1790
  - sources of personality data, 2076–2078
  - statistical distribution, 658–659, 661
  - statistical graphics, 3003, 3011–3019, 3039
  - statistical methods, 3034–3039
  - statistical models, 2028
  - of survey research, 3090
  - tabular, 3107–26
  - of time use research, 3157–3159
  - of variance and covariance, 157–164
  - of voting behavior, 3232–3236
- Data banks and depositories, **573–581**
  - access procedures and use, 579–580
  - age-period-cohort effects, 81
  - census, 282–287
  - computer-assisted, 409–411, 413
  - concept and history, 573–575
  - data sharing, 576
  - demographic data, 609, 631
  - information collection measures, 1802–1803
  - intercohort comparisons, 344
  - replication and, 2397
  - secondary data analysis from, 2473–2481
  - for social research, 2769
  - Social Science Data Archives, 575–580
  - social survey instruments, 577–579
  - World Wide Web, 406, 413
  - See also* Library resources and services for sociology
- Data collection. *See* Data banks and depositories; Public opinion; Survey research
- Data reduction, 906–907
- Date rape, 2558–2559, 2577, 2583
- Datenreport* (publications), 2685
- Dating. *See* Courtship
- David (king of Israel), 1508
- David, Rene, 1555
- Davie, Grace, 2966
- Davies, James C., 2270

- Davis, Beverly, 1594  
 Davis, Fred, 460  
 Davis, Gerald F., 2668  
 Davis, James, 411, 2684  
 Davis, Keith, 1778  
 Davis, Kingsley, 1006, 1030, 2690, 2813, 3214  
 Davis-Moore hypothesis, 2813  
 Davy, Georges, 1026, 1425  
 Dawes, R. M., 592-593  
 Dawkins, Richard, 2207, 2882  
 Day, Randall D., 1729  
 Day-care facilities  
     for adults, 1657  
     for children, 129, 359  
 DDA (Danish data archive), 576  
 DDT insecticide, 88  
 De Christoforo, Violet, 181  
*De Civitate Dei* (Augustine), 1507  
 De facto censorship, 269  
 De Greef, G., 1932  
*De la division du travail social* (Durkheim), 1024  
 De Marchi, Bruna, 1473  
 Death and dying, **581-590**  
     anniversary effect, 584-585  
     assisted suicide, 585, 586-587, 3083-3086  
     bereavement and, 582, 584, 649  
     causes of, 1137-1138  
     demographic data, 609-611  
     depression and, 656  
     end-of-life preferences, 586-587  
     and event history analysis, 869  
     funerals and, 582  
     Kübler-Ross's stages of, 582  
     leading causes of, 222, 224, 1639, 1641  
     lifestyles and, 1639-1642  
     and living wills, 585-586, 587, 3064, 3083  
     new norms for, 3064, 3065  
     nursing home hospice units, 1671  
     planning for, 583  
     and quality of dying, 585, 587-588  
     and "right to die" movement, 585, 586, 3064, 3065, 3083  
     risk of, 1127  
     self-motivation and, 584-585  
     seminal book on, 583  
     social stressors and, 584  
     widowhood mortality risk, 3259  
     *See also* Life expectancy; Suicide; Widowhood  
*Death and Identity* (Fulton), 581, 582  
*Death and the Right Hand* (Hertz), 1032  
 Death penalty  
     as civil liberties issue, 315  
     as criminal sanction, 515  
     as retribution, 2056  
 Death rates. *See* Birth and death rates; Infant and child mortality  
 Debs, Eugene, 273, 2146  
 Debt. *See* Bankruptcy and credit  
 Debt bondage, 2597, 2602, 2604, 2605, 2606, 2607  
*Decameron, The* (Boccaccio), 2185  
 Decency, 2527  
 Deci, E. L., 2059-2060  
 Decision-making power, 2165  
 Decision-making theory and research, **590-601**  
     absolute vs. comparative judgment, 597-598  
     and adolescence, 12  
     coalitions, 329-333  
     decision processing, 593-595  
     decision theories, 590-593, 598  
     democratic process, 603-606  
     general judgment and, 598  
     and group decision making, 595-597  
     on helping behavior, 115-116  
     individual differences, 595  
     innovation-decision process, 678  
     rational decision theory, 349-350  
     systems theory, 3102-3105  
     and war, 3244-3256  
     and widowhood, 126  
 Declaration of Independence (U.S.), 2267, 2365  
 Declaration of Rights of Man and the Citizen (France), 475  
*Déclassé*. *See* Status incongruence  
 Decolonization. *See* Imperialism, colonialism, and decolonization  
 Deconstructionism, postmodern, 2200  
 Deduction/induction. *See* Experiments; Quasi-experimental research designs; Scientific explanation; Statistical inference  
 Deep ecology, 803  
 Deep South Center for Environmental Justice, 791  
*Deep Throat* (film), 2185  
 Defendant's rights, 317, 318  
 Defense Department, U.S., 144, 1876, 1880, 1881  
 Defenses (emotional), 1713-1714  
 Defining Issues Test, 1898-1899, 1901, 1902  
 Degree of association, 1806-1811  
 Dehumanization, aggression and, 74  
 Deinstitutionalization, 1657, 1841, 2660  
 DeJanvry, Alain, 641  
 Delacroix, Jacques, 1709-1710  
 Delayed gratification, 668  
*Delinquent Boys* (Cohen), 534  
 Delinquent subcultures. *See* Criminal and delinquent subcultures  
 Delinquents, juvenile. *See* Juvenile delinquency, theories of; Juvenile delinquency and juvenile crime  
 Della Fave, Richard, 198  
 Della Porta, Donatella, 1472  
 Delli Zotti, Giovanni, 1468  
 Delphi method, 2618, 2678  
 Demagoguery, 1771  
*Demand Analysis* (Wold and Juréen), 3035  
 Demarchi, Franco, 1469  
 Dementia, 1839  
 Demeny, Paul, 619, 632  
 Democracy, **601-609**  
     and capitalism, 241-242  
     and caste system in India, 252  
     and censorship and regulation of expression, 267, 268  
     common understanding of, 602  
     consociational model, 2159  
     critical theory on, 545  
     critiques of, 603, 604-605  
     democratization process and, 2159-2160

- dictatorships vs., 3002–3003  
 and division of powers, 1952–1953  
 and educational mobility, 2927  
 effectiveness of, 605, 606  
 elites and, 2164, 2623, 2624–2626, 2627  
 equality and, 606  
 exclusions in, 602–603, 604  
 and French School of Sociology, 1025  
 historical sociology studies of, 1198  
 individual rights and, 1238–1239  
 interest groups and, 604  
 and Japanese sociology, 1479  
 in labor movements, 605, 606, 1532–1533  
 Latin American movement, 1537, 1538, 1539  
 leadership and, 1565, 2164  
 and Mexican studies, 1861  
 origins of, 605–606  
 participatory, 605, 2164, 2627  
 and peace, 2045, 2046, 2049  
 Plato’s hostility to, 3202  
 pluralist theory of, 604–605, 2164, 2165, 2624–2626  
 political alienation and, 103–104  
 political corruption and, 2124, 2125  
 in political organizations, 2149  
 and political participation, 2339  
 and political party fragmentation, 2157–2158  
 and political party systems, 2153, 2154, 2159–2160  
 public opinion and, 2273  
 rational choice theory of, 2339  
 religious interests’ representation in, 2358–2362  
 renewed definitions of, 605  
 representative, 602–603, 2154, 2156–2157, 2164, 2627  
 and socialism, 2848  
 Tocqueville’s study of, 601, 606, 1515, 1704, 2966, 3227  
 and voluntary associations, 3227, 3230  
 voting behavior research, 3232–3238  
 and war initiation, 605, 3244  
*Democracy in America* (Tocqueville), 1515, 2966, 3227  
 Democratic National Committee, 2128  
 Democratic National Convention (1968), 556  
 Democratic Party  
   labor movement and, 1530, 1531, 1532, 1533–1534  
   liberalism and, 1596  
   machine politics, 2125–2126  
   protest movements within, 2267  
   voter classification, 3233, 3234  
 Democratic Party of the Left (Italy), 2129  
 Democratic Republic of Congo (formerly Zaire), 68, 2133  
 Democratic Revolutionary Party (PRD; Mexico), 2135  
 Demographic and Health Survey, 548, 633  
 Demographic methods, **609–621**  
   birth and death rates, 217–224  
   crime rate calculation, 491–492  
   cross-cultural research, 548–549  
   crude rates comparison, 610–612  
   data description, 609–610, 632  
   descriptive statistics, 658–659  
   fertility determinants, 1005, 1008  
   hazards models, 617  
   indirect estimation, 617–618  
   mortality modeling, 619  
   nonparametric statistics, 1956–1971  
   organizational knowledge as secondary data, 574  
   population projection, 615–616  
   standardization, 2991–2996  
   unemployment rate measurement, 1521, 1720  
   *See also* Population  
 Demographic transition, **621–630**  
   birth and death rates calculations, 217–224  
   childbearing rate decline, 110, 122, 125, 2032  
   and city systems, 3196–3197  
   convergence theories, 425–426  
   divorce rate relationship, 112, 125–126  
   and historical population trends, 2175–2177  
   and long-term care needs, 1653, 1654, 1655  
   marital age, 620, 1425  
   marriage squeeze factors and, 1775–1776  
   men’s economic position and, 1525–1526  
   in Mexico, 1859–1860  
   in Middle Eastern countries, 1866–1868  
   1950s as anomaly, 124, 1525  
   population size and, 632–633, 635, 1219, 2176–2183  
   post-World War II, 704, 1525  
   replacement fertility and, 2181  
   and social change, 703–704, 2642–2643  
   *See also* Birth and death rates; Fertility determinants; Infant and child mortality; Marriage and divorce rates  
 Demography, **630–639**  
   of AIDS/HIV, 2586–2588, 2591–2593  
   alienated groups and, 102  
   American Indian studies, 133–134  
   of American marriages, 124–126  
   of Asian Americans, 175–177  
   and census, 281–287  
   of child sexual abuse, 2581  
   and China studies, 298  
   of cohabitators, 109  
   data archives for secondary analysis, 2476  
   definitions of, 630, 632  
   dependency theory and, 645  
   divorce correlates, 634, 704, 704–705  
   drug users, 710–711  
   and equality of opportunity, 826–827  
   and family size, 970–972  
   formal, 631–632  
   of labor force, 1525–2526

- of legal profession, 468–469  
and life cycle, 1625  
of nursing home residents,  
1667–1668  
organizational, 395–397  
parenting trends, 2031–2033  
population policy, 635  
populations included in, 630–631  
of poverty in less developed  
countries, 2216  
of poverty in United States,  
2214–2215  
as profession, 636–637  
of rape, 2576–2577  
of remarriage, 2387–2388  
research areas, 635–636  
and rural sociology, 2428  
social, 632–635, 2678–2679  
suicide rates by country, 3082  
U.S. suicide rates, 3078  
of underemployment, 1720–1722  
of widowhood, 126, 3256–3257  
*See also specific countries*
- Demographic and Health Surveys  
Project (Macro International),  
2178
- Deng Ziaoping, 2643
- Denisoff, R. Serge, 1927
- Denmark  
family violence, 981  
health-care system, 374, 375, 377  
legal system, 471  
long-term care and care facilities,  
1652, 1653, 1655, 1661  
pornography legalization, 2188  
same-sex marriage legalization,  
111  
Social Science Data Archive,  
575, 576  
social security system spending,  
2797, 2800  
sociology, 2450  
women in labor force  
percentage, 3262
- Dennis, Norman, 225
- Denominations  
decline of liberal mainline, 2379  
definition of, 2365  
members in United States,  
2376–2377
- Dentistry, 2259
- Denton, Herbert, 2497
- Denton, Nancy, 366, 2500, 2501,  
2502–2503, 2504, 3199
- Denton, Wayne H., 1736
- Denzin, Norman, 569, 1633,  
1635, 2220
- Dependency (personal). *See* Personal  
dependency
- Dependency theory, 134–135,  
**639–648**, 1084, 1087–1089, 1091,  
1214, 2922  
critics and defenders of, 644–645  
Latin American studies, 641–643,  
1535, 1538, 2922  
modernization theory vs.,  
639–646, 1706–1707
- Dependent variable (statistical), 3038
- Depression, **648–657**  
aging and, 652, 653, 656, 1839  
in artists, 655  
case study of, 245  
childhood sexual abuse and,  
290, 293  
concomitants, 655–656  
divorce and, 705  
measurement of, 653–654, 1834  
risk factors, 652–653, 1836  
and self-esteem, 2514  
as suicide predictor, 650, 656,  
3078–3079, 3081  
theories of, 650–652, 1838  
treatment of, 654–655  
types of, 649–650  
women's employment and, 1838
- Depression (economic). *See* Great  
Depression
- Depression Adjective Check  
List, 654
- Depressive disorder, 648–649, 650
- Deprivation. *See* Poverty; Relative  
deprivation
- Derrida, Jacques, 2207, 2757
- Dertaux, Daniel, 1633
- DES cattle feed, 88
- Desai, A.R., 1291, 1293
- Desai, I.P., 1291
- DeSapio, Carmine, 2126
- Descartes, René, 1302, 1781, 2218,  
2219, 3005
- Descent. *See* Inheritance
- Descriptive statistics, **657–662**  
data distribution, 658–659, 661  
summary statistics, 659–661
- Desegregation. *See* Segregation and  
desegregation
- Design effect, sampling procedure,  
2447, 2448
- Design of Experiments, The*  
(Fisher), 3006
- Desktop computers. *See* Computer  
applications in sociology
- Desmond, Adrian, 2460
- Determinate sentencing, 2254
- Determined behavior, 528, 529, 539
- Determinism, 2218, 2226
- Deterministic criminology, 521
- Deterrence theory  
of criminal sanctions, 519, 525,  
529, 536, 537, 2056–2057,  
2659  
of criminalization, 525, 529, 536  
general and specific  
processes, 2659
- Detroit Area Studies, 2476
- Deurbanization, 3195
- Deutsch, Karl W., 1940, 1942
- Deutsch, Morton, 402, 1012, 1978,  
2620, 2917  
on social justice, 2699, 2701
- Deutsche Gesellschaft für Soziologie,  
1074–1075, 1077, 1080
- Deutscher, Irwin, 2222
- DeVault, Marjorie L., 1648
- Developing countries  
adult education, 23  
AIDS/STDs, 2591–2593  
case studies of global economic  
changes, 246  
child labor, 3262  
dependency theory, 639–646,  
1706  
and environmental problems,  
793–794, 932  
ethnic political conflict, 847  
family planning, 959–960, 972  
family and religion in, 944  
feminist perspectives, 1708  
fertility declines, 2178–2179



- fertility transitions, 627–629, 633, 1005–1010
- Green Revolution, 89–90
- infant mortality rates, 221
- liberation movement, 541
- and modernization theory, 1706, 1883–1887
- mortality decline, 2177–2178
- mortality levels, 622–623
- mortality modeling, 619
- participatory research, 2040
- population, 2177–2178
- post-disaster impact, 685
- revolutions, 2412–2413
- rural sociology, 2429–2430
- social security systems, 2799–2801, 2803
- time use research, 3161, 3162
- total fertility rates, 219–220
- transnational corporations and, 697, 3178–3179
- UN New World Information and Communication Order, 1767
- See also* Industrialization in less developed countries; *specific countries*
- Development. *See* Education and development; Modernization theory
- Developmental psychology, 1617, 1624
- Freudian, 1713–1714, 2088
- on homosexuality and bisexuality, 2565–2566
- invariant sequence in, 1900, 1901
- moral development stages, 1894–1897
- personality theories, 2088, 2090, 2092
- Piaget stages theory, 1894, 1895–1896, 2085, 2092, 2855
- quasi-experimental design, 1686
- on self-esteem, 2512–2513
- on socialization, 2855, 2856–2862
- Deviance. *See* Alienation; Anomie; Criminalization of deviance; Deviance theories; Legislation of morality
- Deviance theories, **662–674**
- behavior typology, 669
- cross-cultural analysis, 549
- defiant vs. evasive deviance, 523–524
- definition problems, 2659
- ecological factors, 2658
- on illness and disability, 1815–1816
- and labeling, 520, 669–670, 1496–1497
- and legislation of morality, 1575–1576
- liberal vs. conservative, 357
- macro-level origins, 662, 663–666
- macro-level reactions, 663, 670–671, 2660
- on mental illness, 669, 1836–1837, 1840
- micro-level origins, 662, 663, 666–670
- micro-level reactions, 663, 668–670
- music and, 1926
- neutralization theory on, 1496
- pragmatism theory and, 2221
- pure deviance/falsely accused/secret deviance distinctions, 668
- Scandinavian sociology, 2453
- and social control, 2657–2659
- and social learning theory, 666–667, 668, 672
- social stimuli effects on, 2775
- and socialization, 2858
- subcultural approach to, 664, 2775
- and urban underclass, 2497–2498
- See also* Compliance and conformity; Criminalization of deviance; Juvenile delinquency, theories of
- Deviation from the mean (statistics). *See* Mean absolute deviation
- Dewey, John, 1014, 1249, 1250, 1783, 3100
- and pragmatism, 2218–2219, 2220, 2423, 2955, 3098
- Dezalay, Yves, 1550–1551
- DHS. *See* Demographic and Health Survey
- Di Palma, Giuseppe, 2159
- Di Pietro, Antonio, 2129
- Diagnostic and Statistical Manual of Mental Disorders (DSM-I, DSM-II, DSM-III, DSM-IV)*, 649, 650, 653, 1832–1833, 1834, 1838, 1840, 1841
- Diagnostic Interview Schedule (DIS), 1834
- Diagnostic Products, 1824
- Dialectic
- pragmatism and, 2217, 2218
- protest and counterprotest movements, 2267
- Dialectic of enlightenment, 541, 542
- Dialectical Fairy Scene. See* *Passagen-Werk*
- Dialectical materialism, 1781–1782
- definition of, 1782
- and French School of Sociology, 1026–1027
- “Dialectics of Systemic Constraint and Academic Freedom: Polish Sociology under Socialist Regime” (Kwasniewicz), 2119
- Dialects, 2901
- Dianetics* (Hubbard), 2366
- Diaphragm, 626
- Diaries, life histories, 1633
- Diarrheal diseases, 623
- Díaz, Porfirio, 1857
- Dickens, Charles, 1309
- Dictatorship, 1067, 1070, 2163, 2356, 3002–3003
- Dictionary of Occupational Titles*, 3264
- Dictionary of Sociology* (Fairchild ed.), 326
- Dictionary of Statistics* (Mulhall), 3005–3006, 3007, 3008
- Dictionnaire critique de la sociologie* (Boudon and Bourricaud), 1028
- Diderot, Denis, 1976
- Diederich, Charles, 899
- Diet. *See* Eating and diet
- Differential association theory, 507, 531, 533, 534, 537, 666, 668, 2775
- and juvenile delinquency, 1495–1496
- Differential equation models, 1693–1694
- Differential parental investment (DPI) theory, 2884–2885

- Differential prediction, differential validity vs., 3211
- Differentiation, and functionalism, 1031, 2484
- “Differentiation and the Principle of Saving Energy” (Simmel), 697
- Diffuse collective behavior, 554
- Diffusion of Innovations* (Rogers), 678
- Diffusion theories, **674–681**, 1030
- agricultural innovation, 86–91, 2429–2430, 2431
  - collective behavior, 679
  - cultural diffusion, 675–676, 679
  - definition of, 1885
  - earliest social scientific use of term, 675
  - of innovations, 674, 676–679
  - and mass media research, 1763–1764
  - and modernization theory, 1885
  - and popular culture, 2168–2172
  - and rural sociology, 2429–2430
  - and social change, 2643
  - and social networks, 2732
- Diggins, John, 2219
- Digman, John, 2085
- DiIulio, John J., 2053–2054
- Dill, K. E., 73
- Dill, William S., 1005
- Dillman, Don A., 3091
- Dilthey, Wilhelm, 819
- DiMaggio, Paul, 173, 737, 1649–1650
- sociology of music study, 1925, 1926, 2171
- DiMaggio, Paul J., 1925, 1926
- Diop, Cheikh Anta, 66
- DiPrete, Thomas J., 1986, 1987, 1992
- Direct democracy. *See* Participatory democracy
- Directions in Sociolinguistics* (Gumperz and Hymes eds.), 2894
- Dirsmith, Mark, 2221
- Disability
- and eldercare, 1021–1022
  - stigmatization of, 1815
- Disaggregative (ecological) fallacies, 1592
- statistical analysis of, 1593–1594
- Disaster research, **682–688**
- conceptualization of “disaster,” 682–683
  - relationship to sociology, 687
  - technological risk, 2874–2879
  - transemergency period behavior, 684
  - universality of generalizations, 685–686
  - white-collar crime issues, 3252–3253
- Disaster Research Center, 681–682
- Disaster Research Group, 681–682
- Discouraged workers, definition of, 1720
- Discourse, political, 271, 279
- Discrete Multivariate Analysis* (Bishop et al.), 3036
- Discrete-time methods, and event history analysis, 874–875
- Discrimination, **688–695**
- affirmative action and, 47–52, 2496, 2706
  - African Americans and, 54–58, 67, 143, 250, 845, 2333, 2491–2499, 2500–2504, 2601
  - aging and, 79
  - American Indians and, 57, 133–134, 143
  - in American society, 142, 143
  - Asian Americans and, 143, 174, 175
  - comparable worth remedy for, 370–372
  - in criminal sentencing, 2962
  - decomposition approach to, 690–691
  - direct/indirect distinction, 688–694
  - effective remedial intervention, 692
  - ethnicity and, 844, 848
  - homosexuality and, 2570
  - institutional racism and, 53–54
  - noncitizen workers and, 2608
  - occupational, 1312
  - pornography and, 275
  - prejudice and, 2243
  - recognition as social problem, 2761
  - and reference group perceptions, 2753
  - remedy complexities, 691–692
  - sexism and, 988, 989, 1838
  - sexual harassment and, 2591
  - social justice beliefs and, 2706
  - sociological definition of, 688–689
  - “vicious circle” of, 694
  - See also* Equality of opportunity; Segregation and desegregation
- Disease. *See* Epidemiology; Health and illness behavior; Public health; Sexually transmitted diseases; *specific diseases*
- Disengagement theory, 2300
- Dishaw, Nancy B., 1868
- Disordered cohort flow, 345
- Dispersion of distribution, measures of, 659–660, 1964
- Displaced aggression, 73
- Displaced workers, 1722–1723
- Dispute resolution
- coalition process of, 332–333, 465
  - court systems and law, 464–481
  - in group, 1111–1116
  - legal ethnographies, 1549–1550
  - mediation and arbitration, 465, 479
- Disraeli, Benjamin, 1599
- Dissociative identity disorder, 291
- Dissonance theory. *See* Cognitive dissonance theory
- Distribution. *See* Joint distribution; Conditional distribution; Uniform distribution
- Distribution of mean, 3028–3029
- Distribution-free statistics. *See* Nonparametric statistics
- Distributive justice. *See* Human rights/children’s rights; Social justice
- Disturbance terms, 257
- DIT (Defining Issues Test), 1898–1899, 1901, 1902
- Divergence. *See* Convergence theories
- Diversity. *See* Multiculturalism
- Divination. *See* Futures studies as human and social activity
- Divine Light Mission, 3287
- Divine names, 3280

- Division of labor, **695–700**  
 anomie and, 164  
 Blau's landmark theory of, 699  
 capitalism and, 697, 1782  
 coleadership in small groups, 696  
 Durkheim on, 698, 734, 1553, 1575, 1576, 1704, 2647  
 and exchange theory, 2673  
 family and household, 1, 122, 696, 706–707, 1708, 2034, 2568  
 feminist theory on, 1708  
 gender changes in, 1502, 1624  
 heterosexuality images and, 2568  
 life-cycle earnings pattern and, 1624  
 in marriage, 1734, 1736  
 Marxist thought on, 697–698, 1754, 1782  
 and media portrayals of family life, 1699  
 in medical profession, 1816  
 as nationalism factor, 1940  
 occupational specialization as, 696–697, 3264  
 and oppositional class theory, 2692  
 in organizational structure, 696–697, 699, 2002, 2003–2004  
 professions and, 2259  
 radical case studies on, 246  
 reasons for, 112  
 in remarriages, 2390  
 routinization vs. expert specialization, 698  
 small group, 2620  
 societal differentiations, 697  
 subordination of women and, 1708  
 theoretical approaches, 697–699
- Division of Labor in Society, The* (Durkheim), 698, 734, 1553, 1575, 1576
- Divorce, **700–710**  
 acceptability of, 112  
 American patterns, 125–126, 127–128, 700–703  
 barriers to, 1737  
 changes in, 1305  
 child custody and, 702, 707, 2707  
 child support payments and, 128, 708, 947  
 in China, 302  
 cohabitor rates, 109, 705, 922  
 consequences of, 705–708  
 correlates of, 126, 705  
 courtship after, 1779  
 demographic effects, 634, 635, 703  
 determinants in, 1737  
 and event history analysis, 869  
 factors contributing to increases in, 702  
 and family law, 947, 949  
 and family policy in less developed countries, 929  
 and family policy in Western societies, 962, 967  
 and family size, 977  
 and gender, 112, 126, 701–703, 707–808, 1526, 1747  
 as grandparental role factor, 696  
 grounds for, 701, 702–703  
 historical in America, 700–703  
 in Indonesia, 938  
 intergenerational relations and, 696, 1388–1392  
 life course effect of, 82–83, 127–128  
 and life cycle, 1625  
 life expectancy and, 126  
 marital adjustment and, 1726, 1731, 1737  
 median duration of marriage prior to, 1747  
 1950s anomaly, 702, 704  
 no-fault laws, 703, 704, 1305, 1560  
 and non-cohabiting frequency of sex, 2539  
 parental roles and, 2033  
 premarital cohabitation as risk factor, 705  
 rate effect on courtship, 489  
 rate leveling, 112, 125, 1525  
 rates by age, 1742, 1743, 1745–1747  
 recent rate declines, 701  
 remarriage and, 112–113, 125, 126, 488, 708, 922, 923, 1744, 1746, 1748, 2387–2388, 2390–2393  
 Roman Catholic ban on, 1516  
 social justice and, 2706–2707  
 suicide rate, 3078  
 table of rates (1970–1990), 1742  
 widowhood adjustment compared with, 3258–3259  
 women's labor-force participation and, 1525  
*See also* Marriage and divorce rates; Single-parent household
- Divorce Registration Area (DRA), 1743, 1744
- Dixon, William J., 645
- Do not resuscitate (DNR) orders, 585
- Dobbelaere, Karel, 2484
- Dobbins, Frank, 738
- Dobyns, Henry, 133
- Dobzhansky, T., 2885
- Doctors. *See* Physicians
- Documentary method of interpretation (Mannheim concept), 857
- Doeringer, Peter B., 1985
- Dogmatism, persuasion and, 2096
- Dohrenwend, B. S. and B. P., 3056–3057
- Dolby-Stahl, Sandra, 1635
- Dolci, Danilo, 1467
- Dollard, John, 73, 2084, 2670
- Domanski, H., 2121
- Domar, Evesy D., 2596, 2597–2598, 2607–2608
- Domasio, Antonio, 2088
- Domestic partners. *See* Cohabitation; Same-sex marriage
- Domestic partnership certificates, 2546
- Domestic violence. *See* Family violence
- Domhoff, G. William, 604, 2162, 2166
- Dominant stratification ideology, 2245
- Dominican Republic  
 fertility decline, 627

- forced labor, 2603  
 government and political corruption, 2134
- Donati, Pierpaolo, 1472
- Donnan, H., 1933
- Doob, J. L., 73, 3035
- Dooley, Kevin, 410–411
- Dopamine, 652, 654
- Dore, Ronald, 224
- Dornbusch, Sanford M., 1595, 1596
- Dorso, Guido, 1466
- Dos Santos, Theotonio, 640–642
- Dot graph, 3004
- Double standard, 701
- Douglas, Ann, 2171, 2172
- Douglas, Jack D., 459–460, 3079, 3080, 3081
- Douglas, Mary, 1890, 1891, 2484, 2762, 2891
- Dovidio, John F., 118
- Downing, Brian M., 605
- Downs, Anthony, 2278, 2335, 2339, 2920
- Downward mobility, 3050, 3052–3053
- Draft. *See* Conscription
- Draft riots, 2269–2270, 3069
- Drake, St. Clair, 55
- Dramaturgical school, 2768
- Dramaturgy. *See* Symbolic interaction theory
- DRC. *See* Disaster Research Center
- Dream analysis, 1714
- Dred Scott v. Sandford* (1857), 62
- Dreeben, Robert, 2927
- Dress. *See* Fashions
- DRG. *See* Disaster Research Group
- Drive theory (psychological), 1713, 2087
- Driver education classes, 677
- Droegmuller, W., 288
- Dror, Yehezkel, 1040
- Drug abuse, **710–719**  
 aggression linked with, 73  
 aging and declining rates of, 1838  
 AIDS/HIV risk, 712, 2559, 2586, 2587–2591  
 in cities, 311  
 civil liberties issues, 318  
 community assessment process, 366  
 concomitant depression, 655  
 control measures, 713–714  
 and criminal and delinquent subcultures, 509, 512  
 and criminalization of deviance, 523, 525, 526, 1577  
 cross-border trafficking, 1935, 1936  
 decriminalization efforts, 711–712, 717  
 deviance theories on, 664, 667, 672  
 as divorce factor, 1737  
 epidemiology, 711–712  
 harm-reduction policy, 712, 718, 1651–1642  
 Latin American drug trafficking, 2135  
 and legislation of morality, 1577  
 medical treatment of, 521  
 moral and social decline linked with, 359  
 North Korean trade, 2137  
 organized crime profits, 2017  
 policy, 712–713  
 political correctness on, 2140  
 prevention and treatment, 711, 712, 714–717  
 and rehabilitation, 899  
 religious practices issue, 137  
 research, 717–718  
 and sexually transmitted diseases, 2578–2579  
 subcultures, 510, 512–513  
 as victimless crime, 1577
- Drug companies. *See* Pharmaceutical companies
- Drug Enforcement Administration, U.S., 714, 2135, 2143–2144
- Drug testing, 315, 714
- Drugs (therapeutic). *See* Medications
- Drunken driving, 93, 1640, 1641  
 anomie concept, 164–165  
 reform movements, 2722, 2725, 2877
- DSM-I, II, III, IV. See Diagnostic and Statistical Manual of Mental Disorders*
- D'Souza, Dinesh, 2140
- Du Bois, W. E. B., 55, 56, 58, 66, 1071, 2428–2429
- Dualism, 1781  
 pragmatic critiques of, 2218–2219
- Dube, S.C., 1292, 1293
- Dubin, Robert, 3273
- Dubnick, Melvin, 2280–2281
- Duby, Georges, 1516
- Due process, and criminal sanctions, 520–521
- Duhem, Pierre, 821
- Dühring, Eugen, 1066
- Duke Longitudinal Study, 2555–2556
- Dumas, Roland, 2129
- Dummy variable analysis, 454
- Dumont, Louis, 253, 1292
- Duncan, Beverly, 2500
- Duncan, Otis Dudley, 632, 1594, 2682, 2929, 3039  
 education and mobility study, 260–261, 262, 2713, 2716, 2927  
 model of attainment, 1692, 2781, 2782, 2817, 3042  
 occupational structure model, 3035–3036  
 and path analysis, 3035, 3036  
 residential segregation indices, 2500  
 scales of reward or status, 2867  
 and social inequality, 2690, 2867, 2868  
 Socioeconomic Index (SEI), 1997, 2000, 3265
- Duncan's regression, 3017
- Dunham, H. Warren, 1833, 3055
- Dunlap, John T., 422
- Dunlop, John, 1985
- Dunning, E., 2989
- Dunphy, Dexter, 1979
- Duration analysis, 869
- Duration-dependent phenomenon. *See* Life tables
- Durig, Alex, 2298
- Durkheim, Émile, 237, 531, 568, 724, 732, 734, 773, 806, 1074, 1173, 1223, 1233, 1235, 1247, 1272, 1274, 1303, 1314, 1423,

- 1537, 1932, 2193, 2639, 2640, 2645, 2889  
 anomie theory, 164–165, 533, 581, 698, 1493, 1772  
 and British sociology, 225  
 as clinical sociology precursor, 327  
 coining of term “sociology” by, 327  
 on collective social conscience, 526  
 on communitarianism, 355, 2337  
 on community, 362, 366, 1772, 3129  
 comparative historical analysis by, 383  
 on consequences of criminalization, 526  
 on criminalization of deviance, 533, 1575, 1577  
 on division of labor, 698, 699, 1575, 1704, 2633, 2647  
 and ethnography, 852  
 as founder of French School of Sociology, 1024–1025, 1026, 1028  
 and functionalism and structuralism, 131, 1030, 1035  
 and historical sociology, 1196, 1197, 2917, 2917–2918  
 as Japanese sociology early influence, 1477  
 on laicization, 2484  
 legal theory, 1546, 1553  
 macro and micro sociological themes of, 1704  
 and modernization theory, 1885  
 on necessity for definition, 1854  
 organic solidarity concept, 2458  
 as Polish sociology influence, 2117, 2118  
 and positivism, 2192  
 on religious experience function, 775–776, 1032, 2385, 3278  
 representation collectives concept, 197  
 on rising expectations, 1491  
 on rural society, 2426  
 secondary data use by, 574  
 and secularization tradition, 2483  
 segmentary society view, 3134  
 on sentiments, 2519  
 and small group research foundations, 2611  
 social group definition, 2634  
 on societal reaction to crime, 1575, 1576, 1577  
 and sociology of education, 2927  
 as sociology of knowledge basis, 2954, 2955  
 and sociology of religion, 2373  
 and sociology’s supremacy over psychology, 2921  
 and structuralism, 1537  
*Suicide as paradigm*, 2024  
 suicide study, 165, 574, 581, 584, 1595, 3055, 3077, 3079, 3080  
 Dutoix, Marc, 2130  
 Dutt, B. N., 1291  
 Duval, S., 2509  
 Duvalier, François, 2134  
 Duvalier, Jean-Claude, 2134  
 Duverger, Maurice, 2164  
 DVD-ROM, 409  
 DWI (driving while intoxicated). *See* Drunken driving  
 Dworkin, Andrea, 2186–2187  
 Dworkin, Ronald, 466  
 Dwyer, James, 1689, 1691  
 Dyadic Adjustment Scale (marital), 1727–1728  
 Dynamic models. *See* Social dynamics  
 Dynamic sample panel, 1687  
 Dynamic sociology, 1027, 1028  
*Dynamic Sociology: or Applied Social Science* (Ward), 168  
*Dynamic Theory of Personality, A* (Lewin), 1012  
 Dysphoric mood, 649, 650, 651  
 Dysthymia, 649
- E**  
 Eagle Forum, 770  
 Eagly, Alice, 2418, 2530, 2533  
 Early retirement  
   incentive programs (ERIPs), 2407  
   male labor-force participants, 1524  
   structural lag and, 3061–3062  
 East African countries, 60  
 East Asian demographic transition, 626  
 East India Company, 3174  
 East Timor, 2975, 3262  
   genocide, 1066  
 Easterlin, Richard, 635, 2684  
 Eastern Europe  
   cross-border crime, 1936  
   democratization process in, 2159, 2160  
   and equality for women, 990  
   legal system, 1549, 1550  
   political corruption, 2136  
   political party systems, 2159  
   postcommunist revolution, 2851  
   protest movement strength, 2267, 2270, 2271  
   social security systems, 2796  
   socialist economic modifications, 2851–2852  
   sociology studies, 2116–2117  
   voluntary associations, 3227  
   *See also* Soviet and post-Soviet sociology; *specific countries*  
 Eastern religions, 2969, 3287  
 Eating and diet  
   Food Stamp program, 2283  
   health relationship, 1641  
   for pregnant woman, 2235–2236  
   significance in Judaism, 1510  
 Eating disorders, 655  
 Ebaugh, Helen Rose Fuchs, 245  
 Ebonics, 2908, 2909  
 ECI. *See* Ethics of Care Interview  
 Eckstein, Susan, 2414  
 ECLA. *See* Economic Commission for Latin America  
 Ecofeminism, 803  
 Ecological democracy, and environmental equity, 791–793  
 Ecological fallacies. *See* Disaggregative (ecological) fallacies  
*Ecological Imperialism: The Biological Expansion of Europe 900–1900* (Crosby), 2922  
 Ecological interaction., 2630, 2632  
 Ecological paradigm, 824–825  
 Ecological regression (levels of analysis), 1594–1595

- Ecology. *See* Demography;  
 Environmental equity;  
 Environmental sociology; Human  
 ecology and environmental  
 analysis
- Eco-Marxism, 792–793
- EconLit*, 1611
- Econometrics, 256, 3059
- “Economic Action and Social  
 Structure: The Problems of  
 Embeddedness” (Granovetter),  
 735
- Economic and Social Research  
 Council Data Archive (University  
 of Essex), 575, 576, 580
- Economic Behavior Program (Survey  
 Research Center), 2476
- Economic Commission for Latin  
 America, 1537
- Economic determinism, **721–724**,  
 1309
- Economic development. *See*  
 Industrialization in less-developed  
 countries; Modernization theory;  
 Rural Sociology
- Economic Dynamics* (Baumol), 2668
- Economic institutions, **724–731**  
 in American society, 143–144  
 authoritarian state  
 development, 3000  
 capital and, 2637  
 Chinese reforms, 302–303, 2851  
 coalitions and, 331  
 ethnicity and, 845–846  
 family and religion as, 938–941  
 and geographical scope, 728–730  
 and informal economy,  
 1334–1343  
 interdependence, 1211, 1225  
 and leisure activities, 1589  
 and market capitalism, 724  
 and market organization,  
 725–727  
 and rational choice theory, 2335,  
 2336–2337  
 and slave-created wealth, 55–56  
 and Smith’s wealth of nations  
 theory, 2340  
 and social justice, 2697–2698  
 socialist central planning, 2849,  
 2850–2851  
 varieties of, 727–728  
 war and, 3242, 3243  
*See also* Capitalism; Caste and  
 inherited status; Economic  
 sociology; Globalization and  
 global systems analysis
- Economic Opportunity Act of  
 1965, 2266
- Economic sociology, **731–741**, 1309,  
 2919–2921  
 agricultural innovation and,  
 89–91  
 career mobility and, 1989–1990  
 on childbearing rewards and  
 costs, 2034–2035  
 convergence theory, 422–428  
 criminal sanctions and, 518  
 dependency theory, 639–646  
 on direct and indirect  
 discrimination, 689–694  
 divorce effects, 705, 706–707  
 divorce rate relationship, 701  
 and economic justice, 2694  
 education/economic impact  
 relationship, 2934  
 and European Union free trade  
 movement, 1034  
 and experiments, 891  
 and French School, 1024  
 and German sociology, 732–734  
 globalization and, 606  
 on governmental corruption,  
 2124  
 historical, 733–734  
 Latin American studies,  
 1539–1543  
 liberalism/conservatism and,  
 1597  
 macro-level reactions to  
 deviance and, 670  
 modernization theory, 1883–1887  
 and money, 1888–1893  
 new, 734–739  
 organizational knowledge as  
 secondary data, 574  
 on political power vs.  
 economic, 2997  
 population growth and, 633  
 and postindustrial society,  
 2196–2203  
 rational choice theory and, 732,  
 939, 2335–2336, 2338  
 and social capital, 2637–2638  
 social indicators, 576  
 and social security systems,  
 2801–2803, 2804–2805  
 social surveys, 578  
 and Soviet sociology, 2979  
 and tourism, 3165–3172  
 wage inequalities and, 2691  
 Weber’s world religions analysis  
 in context of, 2942–2943  
 and white-collar crime, 3248–  
 3254  
 and work and occupations,  
 3261–3268
- Economic Theory of Democracy, An*  
 (Downs), 2920
- Economics, relationship between  
 sociology and, 2919–2921
- Economics of Non-Human Societies, The*  
 (Tullock), 2920
- Economist, The* (periodical), 794
- Economy and Society* (Weber),  
 721, 732, 733
- ECT (electroconvulsive therapy), 655
- ECTA (software program), 3038
- Ectomorphy, 1717, 1718
- Ectopic pregnancy, 2234
- Ecuador border dispute, 1934
- Ecumenical movement, 2365
- Edelhertz, Herbert, 3246, 3251
- Edelstein, J. David, 1533
- Edgeworth, F. Y., 370, 371, 1045
- Education  
 in adolescence, 10–11  
 of African-American students,  
 2497, 2931–2932, 2933  
 age-graded schooling, 1623  
 allocation and, 11  
 American Indian, 135  
 behavioral applications, 215–216  
 bilingual, 123, 1861, 2140, 2908  
 communitarian view of, 358, 359  
 compulsory, 742, 2056  
 computer-assisted teaching and  
 learning, 411  
 and criminal sanctions, 517  
 curricula control battles, 276–278

- desegregation/integration results, 297, 2493, 2498-2499
- divorce and, 126
- educational access and, 145
- equality of opportunity, 756-757, 759, 826
- as family planning factor, 957-958
- as family size factor, 973, 975, 1009
- and gender identity, 997, 998
- of Hispanic Americans, 1190-1192, 1193
- income distribution and, 1279-1280
- labeling theory and, 2859
- Mexican bilingual, 1861
- Middle Eastern demographics, 1867
- planful competence and, 13, 32
- race and, 145
- reformist projects, 2934-2935
- right to, 1242-1243
- school effectiveness factors, 2933
- and school vouchers, 315
- and schools as socialization agents, 2858-2859
- and schoolteacher status incongruence, 3051
- and self-concept/school performance relationship, 2508-2509
- similarity of male and female math and science performance, 2532
- sociological definition of, 2926
- sociology of, 2926-2935
- and sports, 2986-2987
- structural lag and, 3062-3063
- teacher expectations/achievement relationship, 2932
- and teaching of evolution, 2369
- tracking and, 10, 2932
- training standards, 1180
- as underemployment factor, 1721, 1722
- and violence in schools, 1484-1485, 1487-1488, 1491
- and vocational skills, 3263
- of women, 1219
- See also* Adult education
- Education and development, **741-755**
- convergence theories, 426
- findings, 750-754
- measurement of, 748-750
- and modernization theory, 746-747, 754
- and recent socioeconomic development, 742-746
- theoretical background of, 746-748
- Education and mobility, **755-760**
- barriers to, 830-831
- career mobility relationship, 1987-1988, 1991-1992
- cross-cultural comparison, 3044-3045
- cross-national comparison, 2716
- empirical studies, 2929-2934
- factors in, 3043-3045
- family status and, 2929-2930
- functionalist view of, 757, 2927, 2928
- intergenerational, 2690
- mate selection and, 1776
- occupational prestige relationship, 2000
- socialist quotas and, 2850
- and societal stratification, 2927
- status attainment model, 758, 2713, 2782, 2783-2785, 2817, 2867, 2868, 3042, 3043-3045
- status conflict approach, 2928
- and status incongruence, 3050-3051, 3051
- tolerance of nonconformists linked with, 317
- and transition to work, 2714
- underemployment and, 1721
- See also* Equality of Opportunity
- Education Department, U.S., 766
- Education Morale, L'* (Durkheim course), 1025
- Educational and Psychological Measurement* (journal), 407, 409
- Educational organization, **761-772**
- in American society, 145, 762-770
- and bureaucratic power, 768-769
- case studies of, 246
- and charter schools, 765-766
- diffusion of theories, 677
- drug-abuse prevention programs, 716
- elementary and secondary schools, 763-766
- and ideological formation, 769-770
- and market competition, 767-768
- Marxist sociology of, 1755
- modes of influence over, 766-770
- and national school systems, 760-762
- and political authority, 767
- and professional authority, 769
- racial integration effects, 2932-2933
- reformist projects, 2934-2935
- school choice plans, 2935
- school resources/learning relationship, 2931-2932, 2933
- school shootings and, 1484-1485, 1487-1488, 1491
- school success evaluations, 2934-2935
- and school vouchers, 315
- and structural lag, 3062-3063
- and tracking, 10, 2932
- Edwards, J. E., 1566
- Edwards, Jonathan, 2086
- EEOC. *See* Equal Employment Opportunity Commission
- EFA. *See* Exploratory factor analysis
- Effectance motivation, 2060-2061
- Efficacy, trust and, 101-102
- Egalitarian family system, 1490, 1734
- Egalitarianism. *See* Equality of opportunity
- Egbaugh, Helen Rose Fuchs, 245
- Eggan, Fred, 289
- Ego, 1713, 1714
- Egypt, 1865, 1866, 1867
- fertility rate decline, 220
- government and political corruption, 2132
- Israel peace accord, 2048
- pan-Arab nationalism, 1944

- social change in, 941  
sociodemographic profile, 2938  
sociological article count, 1869
- Eheart, Brenda Krause, 1636
- Ehrenhalt, Alan, 356
- Ehrenreich, Barbara, 1818
- Ehrenreich, John, 1818
- Eighteenth Brumaire of Louis Bonaparte, The* (Marx), 723, 2163
- Einstein, Albert, 1012, 2196
- Eisenhower, Dwight D., 2127, 2299–2300
- Eisenstadt, S. N., 1198, 2690
- Ejido*, 1859
- El Salvador  
protest movements, 1170, 2266  
wartime rape, 2579
- Elaboration paradigm, 3107
- Elavil, 654
- Elazar, Daniel, 2, 2126–2127, 2377
- Elder, Glen H. Jr., 31–32, 360, 1618, 1619, 2662, 2855, 2861
- Eldercare, 129–130  
and filial responsibility, 1019  
*See also* Long-term care; Long-term care facilities
- Elderly people. *See* Aging and the life course
- Election polling, 575, 2278–2279, 3232–3236  
panel design, 1686  
presidential race fiascoes, 2273–2274, 3232
- Elections. *See* Voting behavior
- Electoral system types, 2154, 2156–2157
- Electra complex, 332, 1713
- Electroconvulsive therapy, 655
- Electronic journals, 413
- Electronic mail, 406, 407, 408, 1607, 1768
- Electronic networks, 408
- Electronic-text form, 407–408
- Elementary and secondary education, 763–766
- Elementary and Secondary Education Act of 1965, 2266
- Elementary Forms of the Religious Life* (Durkheim and Mauss), 1032
- Elementary Structures of Kinship, The* (Lévi-Strauss), 1032, 1033
- Éléments de sociologie* (Bouglé and Raffault), 1025
- Éléments de sociologie* (Davy), 1026
- Eliade, Mircea, 3279
- Elias, Norbert, 4, 782, 1075, 2521–2522, 2989, 3155
- Elite paradigm. *See* Intellectuals; Social and political elites
- Elizabethan Poor Laws, 2840–2841
- Elling, Ray H., 376
- Elliott, Delbert, 1492
- Ellis, Carolyn, 248, 1636–1637, 2291–2292
- Ellis, Havelock, 2566
- Ellison, Curtis W., 1925
- Ellwood, Robert, 2380
- Elmira (New York), voting behavior study, 3233, 3234
- Elohim, 3280
- Elster, Jon, 1234, 2341, 2452, 2701
- EM algorithm, 3039
- E-mail (electronic mail), 406, 407, 408, 1607, 1768
- Embarrassment, 2523
- Embezzlement, 3246, 3247, 3251
- Embracing the Other* (Oliner et al.), 118
- Emergent norm theory, 351, 354
- Emerson, R., 208, 2670, 2671–2673, 2674
- Emerson, Ralph Waldo, 2218
- Emic-etic frame, 550, 564, 2091–2092, 2889, 2891, 2892
- Emler, Nicholas, 1901, 1902
- Emolins, E., 1025
- Emotional attachments. *See* Attachment; Attachment theory; Social belong
- Emotional energy, 786
- Emotions, **772–788**  
affect control theory, 43, 44–45, 780, 2527–2528  
attitude change and, 189  
and communication, 2522  
cultural approach to, 774, 781–785  
divorce effects on, 707–708
- impression formation and, 43  
and macroprocesses, 786–787  
masking of. *See subhead* affect control theory *above*  
and nineteenth-century sociology, 773–776  
norms and, 2527–2528  
and reason, 785–786  
regulation of. *See subhead* affect control theory *above*  
and self-esteem, 2514  
sentiments vs., 2518–2519, 2523  
social contagion theory, 679  
and social movements, 2722  
social structural approach, 776–781  
sociohistorical approach, 2521–2522  
sociological case studies of, 245  
values and, 3213  
*See also* Depression; Stress
- Empathy, 116
- Empedocles, 2086
- Emphysema, 1640
- Empire Strikes Back, The* (Solomos, Findlay, Jones, and Gilroy), 226
- Empires, 2998–2999, 3001
- Empirical analytic metatheory. *See* Metatheory
- Empirical sociology, 1079
- Empiricism, 1249
- Employee Retirement Income Security Act of 1974, 2404
- Employment. *See* Labor force; Work and occupations
- Employment Act* (report), 576
- Employment relationship. *See* Work orientation
- Employment Retirement Income Security Act of 1974 (ERISA), 1826
- Empty-nest stage, 1729
- Enclave theory (Portes concept), 848–849
- Enculturation, 2855
- Encyclopedia of Social Sciences*, 583
- Encyclopedia of Statistical Sciences* (Kotz and Johnson), 1957
- Endler, Norman, 651



- Endo, Ryukichi, 1477
- End-of-ideology argument, 2626
- End-of-life preferences. *See* Death and dying
- Endogamy  
     historical kinship effects, 1507, 1515  
     mate selection and, 1776, 1779  
     population studies, 634
- Endogenous variables, 262, 2251
- Endomorphy, 1717, 1778
- Endowment effect, and decision processing, 594
- Engels, Friedrich, 415, 698, 988, 1196, 1234, 1290, 2640, 3066  
     and dialectical materialism, 1782  
     on family structure, 1505  
     and historical materialism, 1751, 1752, 1782  
     and industrialization, 2196  
     on religion, 2965, 2967, 2968  
     scientific socialism concept, 1753  
     socialist ideal of, 2847  
     and sociology of law, 1576  
     theories of the state, 2162, 2163  
     time use study, 3155  
     on women's inequality, 2692
- Engerman, Stanley L., 2596
- Engineering, 2259, 2460, 2461  
     women careerists, 2532, 2533, 2785
- England. *See* British sociology; United Kingdom
- English Poor Law of 1601, 2840
- Enlightenment, 603, 2178  
     critical theory on, 541, 544  
     egalitarian ideals of, 2811  
     on free expression, 268  
     humanism and, 1247-1248  
     and materialism, 1781, 1782  
     and personality theories, 2086-2087  
     postmodernism and, 2205, 2206, 2207, 2208  
     and rational choice theory, 2335, 2340  
     scientific secularism of, 2142  
     and secularization, 2482, 2965-2966
- Ennis, Carolyn Z., 2089
- Ennis, Philip, 1925
- Enriquez, Eugene, 328
- Ensel, Walter M., 2792
- Enterprise labor movements, 1529
- Entitlements, 1221-1222, 2699
- Environment  
     definition of, 1228  
     *See also* Environmental equity; Environmental movement; Environmental sociology; Human ecology and environmental analysis
- Environmental Defense Fund, 803
- Environmental equity, **788-800**, 809  
     and ecological democracy, 791-793  
     and environmental justice movement, 789-790, 791, 803, 809  
     and epidemiology, 817  
     and global issues, 793-794  
     and grassroots protests, 791  
     and mainstream environmental groups, 790-791  
     and materialist theory, 1785  
     and racism, 789, 803, 809, 1159  
     research finding on, 794-795  
     and urban development, 308-309  
     *See also* Environmental sociology
- Environmental justice. *See* Environmental equity
- Environmental Justice Resource Center, 791
- Environmental movement, 2717, 2719, 2724-2725, 2877  
     and suburban exclusionary zoning, 3072
- Environmental Protection Agency, 791, 802, 804, 2461
- Environmental racism, 789, 803, 809, 1159
- Environmental security, 1222
- Environmental sociology, **800-813**  
     birth of, 804-805  
     and conservation movement, 802-803  
     constructivist/interpretative orientation, 810-811  
     and cross-cultural analysis, 549  
     current trends in, 810-811  
     and evolution of problems, 801-802, 804-805, 807-809  
     and mainstream sociology, 805-807, 811  
     and Mexican sociology, 1860  
     and modern movements, 803-804  
     as proactive, 1226-1228  
     solutions to problems of, 809-810  
     and urbanization, 311-312  
     *See also* Human ecology and environmental analysis
- EPA. *See* Environmental Protection Agency
- EPA (evaluation, potency, and activity) responses, 42, 43, 45-46
- Epidemiologic transition theory, 1325-1326, 1327
- Epidemiology, **813-818**  
     drug abuse, 710-711  
     and healthy life expectancy, 1632  
     and lifestyle risks and health, 1639-1642  
     and medical sociology, 1813  
     mental illness and disorders, 1833-1840  
     methodology, 815-817  
     origins of, 814  
     sexually transmitted diseases, 814, 2585-2593  
     suicide, 3078-3079  
     and theories of crime, 503  
     typical paradigm in, 814-815
- Episcopalians, 95
- Epistemology, **818-826**  
     and evaluation research, 867  
     and feminist theory, 994-996  
     and German Historical School, 819-820  
     and materialism, 1780  
     and metatheory, 1852-1854  
     postmodern, 2205, 2206-2208, 2757  
     and scientific explanation, 2469-2472  
     and social philosophy, 2756, 2757  
     and sociological positivism, 818-819

- and sociology of knowledge, 2954, 2957  
 and twentieth-century sociology, 820–824
- Epstein, Cynthia Fuchs, 995
- EQS (computer software), 1914
- Equal Employment Opportunity Commission, 49, 2591  
 sexual harassment guidelines, 2580  
 sexual harassment policy, 2581
- “Equal Pay to Men and Women for Equal Work” (Edgeworth), 370
- Equal Rights Amendment (ERA), 579, 2267, 2724
- Equality of Educational Opportunity* (Coleman et al.), 758, 830–831
- Equality of opportunity, **826–835**  
 African-American historical denial of, 2333–2334, 2491–2499  
 comparable worth and, 369–372  
 democracy’s relationship with, 606  
 and demographic trends, 826–827  
 direct/indirect discrimination as bar to, 689–692  
 division of labor as threat to, 698  
 in education, 756–757, 759, 826  
 and family policy in Western societies, 963  
 and feminist theory, 988  
 as industrial era’s defining feature, 2811–2812, 2820  
 in less developed countries, 932  
 meritocracy model and, 2626  
 policy and practices of, 833–834  
 political correctness and, 2140, 2141  
 prejudice and, 2244, 2245  
 process differences in, 829–830  
 residual differences in, 827–828  
 social justice and, 2704–2707  
 as Swedish health-care principle, 377  
*See also* Civil rights; Class and race; Discrimination; Segregation and desegregation; Social class; Social inequality
- Equilibrium theory. *See* Cognitive consistency theories; Social dynamics
- Equity law, 477–478
- Equity theory, 2700–2701  
 and just society, 3205–3206
- Equivalence consistency, 2348–2349
- ERA. *See* Equal Rights Amendment
- Erbakan, Necmetin, 2132
- Erie County (Ohio) voting behavior study, 3233–3234
- Erikson, Erik, 5, 460, 1617, 2084, 2092  
 eight stages of life concept, 1624, 2861  
 moral development theory, 2089
- Erikson, Robert, 425, 2452
- ERISA. *See* Employment Retirement Income Security Act of 1974
- Eritrea, 1941
- Erotica, pornography vs., 2184–2185
- Error. *See* Measurement; Reliability; Validity
- Eskola, Antti, 2452
- Esping-Andersen, Gosta, 377, 379, 2452, 2797
- Espionage Act of 1917, 2146
- Espy, Mike, 2128
- Essed, Philomena, 57
- Essentialist sexual behavior theory, 2537, 2566, 2567
- Estes, Richard J., 2302
- Estimation techniques, 1958  
 sample-based, 2444
- Estonia, 2362  
 sociological research, 2982–2983
- Ethanol alcohol. *See* Alcohol
- Ethic of Care Interview, 1903
- Ethics  
 and cloning, 1824  
 and death and dying, 585, 3064  
 and economic institutions, 727  
 in religious experiences and social processes, 3283–3284  
 in survey research, 3093–3094  
*See also* Ethics in social research; Moral development
- Ethics in Government Act of 1978, 2127
- Ethics in medicine. *See* Bioethics
- Ethics in social research, 542, **835–840**, 853  
*Code of Ethics*, 836–840  
 confidentiality, 836  
 disclosure of purpose, 836–837  
 government regulation, 838–839  
 life histories and narratives, 2293  
 moral relativism, 840  
 scientific research fraud, 2458  
 utilitarianism, 839–840
- Ethics of Care Interview, 1900, 1902–1903
- Ethiopia, 2602, 2604, 2999
- Ethnic cleansing. *See* Genocide
- Ethnic myth, 2332
- Ethnic niche, 849
- Ethnic succession, 532, 558  
 in organized crime, 2019
- Ethnic violence, 1199
- Ethnicity, **840–852**  
 African sociocultural institutions, 63–65  
 AIDS/HIV risks, 2587, 2590  
 alcohol use and, 94–95  
 assimilation model, 178–179, 842–844  
 and capitalism, 239  
 case studies research, 244  
 and caste and inherited status, 253  
 and census, 284–285, 286  
 and civil rights movement, 178–179  
 and collective mobility, 2715–2716  
 and courtship, 486  
 and crime rates, 531–532  
 and criminal deviance, 526, 530  
 definition of, 841  
 demographic factors, 636  
 differentiated from race, 2329  
 diversity in, 1296  
 and economic conflict, 845–846  
 and ethnic-group resources, 847–849  
 and ethnocentrism, 1274, 1400, 1580, 2882  
 evolutionary perspective on, 2882–2883

- and family policy in Western societies, 966
- and family structure, 123
- and forecasts of American white minority, 1580
- and genocide, 1066–67, 1069–1070, 1071, 2529
- and health care, 1152
- and informal economy, 1340–1341, 1349–1341
- and Japanese sociology, 1483
- language varieties, 2909
- and Latin American societies, 1536
- markers of, 23239
- and Mexican social change, 1857, 1858–1859, 1861
- and nationalism, 1939–1949, 3001–3002
- in neighborhoods, 532, 558
- and organized crime, 2018–2019
- and political correctness, 2140, 2141
- and popular culture transmission, 2170–2171
- poverty level, 2215
- and prejudice, 1600
- and race, 841
- and reactions to deviance, 670
- and relative deprivation, 1940
- and religious intolerance, 3288
- remarriage rates, 2388
- and social constructionism, 849
- social surveys of, 578
- socialization in subcultures, 2862
- and status attainment, 3044
- and status incongruence, 3051, 3052, 3054
- and stratification, 844–847, 2818–2819
- and suburbanization, 3074–3075
- suicide variation, 3079
- and underemployment, 1721, 1722, 1724
- and urban life, 307, 308, 532–533, 558, 2498
- and war, 3244
- See also* Cross-cultural analysis; Discrimination; Equality of opportunity; Multiculturalism; Race; Social mobility; *specific groups*
- Ethno (computer program), 409
- Ethnocentrism, 1274, 1400, 1580, 2882
- Ethnogenesis, 1296
- Ethnographic Atlas*, 548
- Ethnography, **852–856**
- autoethnography, 245, 852, 1636–1637, 2291, 2293
- British studies, 2890
- case studies, 243–244, 245, 246–247
- Chinese studies, 297–298, 301
- and community studies, 364–365
- cross-cultural analysis, 547–548
- and cultural studies, 564
- definition of, 2888
- of drug abuse, 717
- emic-etic frame, 550, 564, 2091–2092, 2889, 2892
- field research, 853–854
- legal, 1549–1550
- money theories, 1890
- and new religious movements, 2367
- postmodern, 855, 2893
- and publication, 854–855
- sociocultural anthropology and, 2888–2894
- sociolinguistics and, 2891
- of underclass neighborhoods, 3200
- of urban area deviance, 665
- writing of, 247–248
- Ethnology
- China studies, 298, 300
- culture definitions, 563
- definition of, 2888
- Ethnomethodology, 852, **856–861**, 2756
- background and development of, 856–858, 2099
- and case studies of reality construction, 246–247
- contemporary research initiatives in, 858–860
- conversation analysis, 247, 431–439
- definition of, 2768
- phenomenology and, 2099
- and social institutions and worlds, 859–860
- and social structures, 858–859
- and sociolinguistics, 2895
- and workplace studies, 860
- Ethnonationalism. *See* Nationalism
- Ethnoscience, 547, 2891
- Etic. *See* Emic-etic frame
- Etzioni, Amitai, 316, 318, 1234, 2379
- on alienation, 104
- “I-We” paradigm, 1601
- and new communitarianism school, 356
- EU. *See* Expected utility (EU) theory
- Eucharist, 3281, 3282
- Eugenics, 320, 1272
- and family law, 948–949
- and fertility determinants, 1005, 1007
- and genocide, 879
- See also* Social Darwinism
- Eunuchs, 2602
- Eurobarometers, 549, 577, 578–579, 3223
- Eurocentrism, 55, 57–58
- Europe
- abortion legality, 2240
- borderlessness, 1934–1938
- boundaries development, 1932, 1939
- colonialism in Africa, 60–61, 1934
- countercultures, 462–463
- data archives, 2477
- discrimination practices, 692–693
- divorce policy, 703–704
- drug abuse policy, 712
- empires, 2998
- evolution of paid work, 3262
- fertility transitions, 622–623, 625–627, 633, 1006–1010, 1007, 2178
- highest life expectancies in, 623
- imperialist class and race effects, 320–322
- interest groups, 2150–2151
- kinship systems, 1515, 1516–1517

- labor market structure, 1988  
labor movement, 1528, 1529, 1530–1531  
legal codification, 475–476  
legal system, 1546–1547  
literary sociology, 1649  
long-term care and care facilities, 1652–1654  
marriage and divorce rates, 1749  
mathematical work in sociology, 1791  
migrant Arab workers, 1865  
mortality transition, 621, 632  
Muslim minorities, 2950  
occupational mobility, 1987, 1988  
opinion studies, 578–579  
political corruption, 2128–2130  
political party systems, 2157–2158  
population factors, 636, 2177, 2182  
rape incidence, 2576  
retirement comparisons, 2407  
Roman law as legal system basis, 465, 1545  
sex segregation of occupations, 3264  
social anthropology studies, 2888, 2890  
Social Science Data Archives, 575, 576–577, 579–580  
state system's emergence in, 1933, 2356, 2362  
and status incongruence, 3051–3052  
supranational bodies, 3003  
*See also* European Union; North Atlantic Treaty Organization  
terrorism, 3137  
voting behavior research, 3237, 3238  
woman suffrage, 703  
*See also* Eastern Europe; *specific countries*
- European Commission  
corruption scandal, 2130  
General Directorate for Science and Technology, 580
- European Community. *See* European Union
- European Economic Community. *See* European Union
- European Fertility Project (Princeton), 625–627  
*European Household Panels*, 577  
European Lifelong Learning Initiative, 24  
European Parliament, 2130  
European Sociological Institute, 2452  
European Union, 395, 428, 464, 476, 480, 1948  
borderlessness of, 1934, 1936  
Eurobarometer surveys, 549, 577, 578–579, 3223  
and hegemonic stability theory, 3242  
interest groups and, 2151  
lifelong education, 23–24  
minority unemployment, 692–693  
multinational legal system, 1550  
opinion studies, 578–579  
political corruption, 2130  
powers of, 3003  
time use research, 3159–3160  
values surveys, 3223
- Eurostat time use survey, 3160
- Euthanasia, 585, 2719, 3088  
passive, 3084  
voluntary vs. involuntary, 3083  
*See also* Assisted suicide
- Euthanasia Research and Guidance Organization, 3084
- EV. *See* Expected value (EV) theory
- Evaluation research, **861–869**  
accountability in, 862–863  
definition of, 861–862  
development of, 862  
future of, 866–868  
practice of, 868  
quantitative vs. qualitative, 864–865  
syntheses, 865–866  
and utilization of findings, 863–864  
and value theory, 867–868, 2828
- Evangelicalism, 2370, 2965, 2966–2967, 3081
- Evans, Peter, 642, 643, 1706
- Evans-Pritchard, E. E., 563, 2890
- Event frame, 2298
- Event history analysis, **869–875**  
accelerated failure-time models, 870–871  
data, 870  
discrete-time methods, 874–875  
mobility research, 2817  
models, 1692–1693, 1790  
multiple events, 873–874  
problems with conventional methods, 870  
proportional hazards models, 871–873  
repeated events, 874  
statistical method, 3037  
and terrorism, 3140
- Event management, 43–44
- Event sequence analyses, 2297
- Everday life. *See* Time use research
- Evers, Medgar, 2495
- Evolution: biological, social, cultural, 1, 563, **875–880**, 1029, 1030, 1074, 1228, 1272  
altruism theory, 115, 118, 2882–2883  
and anthropological origins, 2889–2890  
fitness principle, 2881–2882, 2885  
fundamentalist opposition to theory of, 2369  
gender role theory and, 2418  
and generative religious movements, 2367  
genocide theory and, 1068  
macro perspective, 1704, 1705  
Marxist historical materialism and, 1782  
modernization theory and, 1885, 1886  
and Polish sociology, 2118  
pragmatism and, 2218, 2219  
and progress concept, 2644–2645  
racial theories and, 2330, 2334  
self-esteem theory, 2513  
on sex differences, 2530  
and sexual behavior, 2537, 2567, 2884–2885  
social organization and, 1271

- sociobiology and, 2880–2887, 2892
- statements of Darwin’s theory, 2880–2881
- “Evolution of Autopoietic Law” (Teubner), 1558–1559
- Excellence: Can We Be Equal and Excellent Too?* (Gardner), 1384–1385
- Exchange and Power in Social Life* (Blau), 2670
- Exchange media. *See* Money
- Exchange mobility, 2712
- Exchange theory. *See* Social exchange theory
- Excluded variables, 2251
- Existential personality theory, 1717, 2084, 2085
- Existentialism, 989, 2756
- Exner, John E., Jr., 2077, 2078
- Exodus, 3053, 3280, 3282, 3283, 3285
- Exogamy
- kinship, 1516
  - mate selection and, 1776
- Exogenous variables, 2251
- Expectation states theory, **880–887**
- in academic achievement, 2932
  - application of, 884–885
  - and gender, 882–883, 2530
  - interpersonal relations and, 2774
  - and perception of justice, 2696–2697
  - and performance expectations, 880, 882
  - relative deprivation
    - expression, 349
  - research in, 881–883
  - role theory and, 2418, 2423
  - and social networks, 2731
  - and socialization, 2859
  - and status characteristics, 880, 882
  - and work orientation, 3271, 3272
- Expected utility (EU) theory, 591–592, 597, 598
- Expected value (EV) theory, 591, 592, 2673
- Experiential personality theory, 1718
- Experimental Economics* (journal), 891
- Experiments, **887–892**
- analysis of variance and covariance, 157–164
  - compliance and conformity, 400–401, 404
  - cost of, 890
  - and economic sociology, 891
  - multiple-indicator models as manipulation checks, 1922
  - quasi-experimental design, 2309–2327
  - replication and, 2395–2397, 2471
  - scientific method, 2463–2472
  - social psychology, 2769
  - See also* Scientific explanation
- Expert specialization, 699
- Expert systems, 410
- Exploratory factor analysis (EFA), 907, 911, 912, 1920–1921
- Exponential smoothing, 2679
- Expressive order, in social relations, 46
- Expulsion, 527
- Extended families. *See* Kinship systems and family types
- Extramarital/extradyadic sex, 2541–2545
- Extreme influence: thought reform, high control groups, interrogation, and recovered memory psychotherapy, **892–904**
- and effects of brainwashing, 897–898
  - and group-awareness training, 898–900
  - and hypnosis, 898–899
  - and physical abuse, 897
  - and police interrogation and false confessions, 900–901
  - and recovered memory psychotherapy, 901–902
  - and reeducation programs, 892–898
  - research on, 902
  - Schein definition of, 893
- Extremism, political, 2160
- Extroversion, 1565
- Exxon-Valdez* oil spill, 2877
- Eyadema, Gnassingbe, 2133
- Eyerman, Ron, 1357–1358, 1927
- Eysenck, Hans, 70, 2084, 2087
- Ezzy, Douglas, 1636
- F**
- F scale (authoritarian personality measurement), 317
- F tests (variance and covariance), 160–161
- Faberman, Harvey, 2220
- Face-to-face interviewing, 3091–3092
- Facial expressions, 1976
- Facilitating effect, 349, 2615
- Factions
- kinship systems and, 1509–1511
  - pivotal role of weak, 330
- Factor analysis, **905–922**
- basis of, 3036
  - and causal modeling, 908, 917–918
  - and common factor model, 908–910
  - and communality estimation, 912
  - confirmatory, 907, 915–918, 1920–1921
  - and data reduction, 906–907
  - early development of, 906
  - and estimation and testing, 910–911
  - exploratory, 907, 911, 912, 1920–1921
  - and intercorrelatedness matrix, 2346
  - and invariance, 914
  - latent structure analysis vs., 3038, 3039
  - and matrix notation, 911–912
  - and measurement-concept problem, 1788
  - and multiple indicator causal models, 908
  - and multiple-indicator, multiple-cause models, 917
  - and multitrait-multimethod models, 917
  - and personality trait measurement, 2078
  - principal components analysis vs., 913

- as quantitative methods influence, 3036
- and reliability, 2346, 2347
- as research tool, 906
- and rotation problem-correlated factors, 913–914
- and score estimation, 914–915
- and standardization, 2996
- for status crystallization levels, 2869
- and true-score models, 917
- Factories
- and corporate organization, 441–442
  - labor movement in, 1528
  - and Marxist theory, 2196
  - See also* Industrial Revolution; Industrialization
- Fads
- as collective behavior, 348, 554
  - conformity and, 402
  - See also* Fashions
- FAFO (Trade Union Movement Research Foundation; Norway), 2451
- Failure-time analysis, 869
- Fairchild, H. P., 326
- Fairness. *See* Social justice
- Faletto, Enzo, 1538
- Falk, R. Frank, 2500
- Fall, Albert, 2127
- Fallacies
- of cohortcentrism, 343
  - disaggregative (ecological), 1592, 1593–1594
  - gambler's, 594–595
  - individualistic, 3052
  - life-course, 344–345, 1614
  - longitudinal, 1593, 1614
- Fallers, Lloyd, 1549
- Fals Borda, Orlando, 2040, 2041, 2042
- False memories, 2083
- Falwell, Jerry, 2370, 2371
- Familism. *See* Filial responsibility; Family and household structure; Intergenerational relations; Intergeneration resource transfers
- Family allowance programs, 2795, 2798, 2799, 2803
- Family and household structure, **922–928**
- adulthood and, 27, 28, 34, 35
  - of African Americans, 121–122, 2333
  - alternative lifestyles and, 106–113, 1506
  - American family trends and, 126–129
  - of American Indians, 120–121, 134
  - of Asian Americans, 176–177
  - blended families, 112–113, 126, 2391–2392
  - childhood sexual abuse and, 291, 292
  - childless, 109–111, 634
  - child-rearing style/juvenile crime relationship, 1490
  - in China, 302
  - coalitions within, 331–332
  - in colonial America, 121
  - convergence theories, 426
  - cross-national comparisons, 130
  - and culture-of-poverty viewpoint, 2211–2212
  - declining size of, 487
  - demographic transition and, 625, 628–629, 634
  - deviance theories and, 666, 667, 668
  - disaster planning and behavior and, 683, 684
  - division of labor and, 1, 122, 696, 706–707, 1708, 2034, 2568
  - divorce and, 704, 705–708, 922
  - drug-abuse prevention and, 716, 717
  - dual-income families, 127, 1524, 1525–1526, 3062
  - educational attainment and, 2930, 2931
  - eldercare and, 129–130, 1019–1020
  - eschelon authority structure, 706
  - grandparents and, 131, 696, 1390–1391
  - health behavior and, 1129
  - heterosexuality construct and, 2568
  - homelessness and, 1204
  - income and, 1281
  - informal long-term caregiving and, 1657–1658
  - institutional-to-companionship shift, 1502, 1506
  - intergenerational relations and, 1388–1392
  - intergenerational resource transfers and, 1393–1395
  - interpersonal conflict resolution and, 1454
  - Islamic patriarchy, 2949–2950
  - Japanese sociology and, 1478, 1480, 1481
  - juvenile delinquency theories and, 1490, 1498
  - leisure and, 1584–1585
  - liberal vs. communitarian view of, 359
  - life-cycle concept of, 1615–1616, 1617–1618, 1620, 1624–1625
  - long-term care needs and, 1654
  - marital adjustment and, 1725–1732
  - marital satisfaction and, 1729
  - marital trends and, 487, 488, 922–923
  - Mexican studies, 1860–1861
  - Middletown study findings, 364
  - modernity and, 1501–1507
  - nonmaternal childcare, 128–129
  - nuclear. *See* Nuclear family
  - object relations theory and, 2063–2064
  - patriarchal, 1009, 1271, 1490, 1498, 1503, 1579, 1708, 1734
  - population composition studies, 634
  - and poverty level, 2215
  - remarriage and, 2390–2393
  - as repetitive cycles, 1506–1507
  - and social learning theory, 665–666, 667–668
  - and social mobility, 2714
  - social surveys of, 577, 578
  - socialization and, 2856–2858, 2862

- strengthening of, 1243–1244  
 structural lag and, 3062–3063  
 synchronization of lives and, 1615  
 time use research, 3160–3163  
 work relationship, 3266, 3275  
*See also* American families;  
 Intergenerational relations;  
 Intergenerational resource  
 transfers; Kinship systems and  
 family types; Parental roles;  
 Single-parent household
- Family and Medical Leave Act of  
 1992, 2033–2034
- Family and population policy in less  
 developed countries, **928–934**  
 definition of, 929, 933
- Family and religion, **934–947**  
 interfaith marriage, 911,  
 1411, 1776  
 in private sphere, 935, 942  
 in public sphere, 935, 942–944  
 secularization of, 937–938  
 and social change, 936–942
- Family Assistance Plan, 2281
- Family bereavement, 582, 584
- Family Coordinator, The* (journal), 106
- Family farming, 2431–2432
- Family law, **947–952**  
 and common law marriage, 948  
 and contraception use, 950  
 and court nonintervention,  
 949–959  
 and divorce, 947, 949, 950  
 and filial responsibility,  
 1018–1019  
 and interracial marriage, 949, 950  
 justice analysis, 2707  
 and marriage, 947, 948–949
- Family life cycle, 1625, 1729,  
 1737–1738
- Family planning, **952–962**  
 in China, 220, 303, 930,  
 931–932, 972  
 conditions for contraceptive use,  
 2178–2179  
 contraceptive availability/lowered  
 abortion rate ratio, 2241  
 and contraceptive effectiveness,  
 956–957  
 and contraceptive provision,  
 958–959, 2178  
 and contraceptive use, 954–956  
 demographic research on, 635  
 and education and information,  
 957–958  
 and eugenics, 1272  
 and family law, 950  
 fertility declines and, 626, 627,  
 628, 1005, 1008, 1010, 2032,  
 2176, 2178  
 international comparisons of,  
 959–960  
 Malthusian model and, 633  
 mortality rates affecting, 624,  
 2177  
 national programs, 2179  
 as population policy in less  
 developed countries, 928–  
 929, 2179  
 and rights of minors, 950  
 and risk of pregnancy, 952–953  
 and unintended pregnancy,  
 953–954, 2180, 2234  
*See also* Family policy in Western  
 societies; Family size
- Family policy in Western societies,  
**962–970**  
 definition of family, 964–965  
 definition of policy, 965–966  
 goals of, 966  
 and poverty, 967–968
- Family roles. *See* Alternative life  
 styles; American families; Family  
 and household structure;  
 Parental roles
- Family size, **970–981**  
 coalitions and, 332  
 definition of, 970  
 demographic trends in, 970–972,  
 1008, 1525, 2032, 2177  
 educational attainment and,  
 974–975  
 family structure and, 977  
 gender preference and, 628  
 labor-force participation and,  
 1525  
 only children, 974  
 parental attitudes and, 976  
 parenting style and, 2773  
 physical/social-psychological  
 development and, 974  
 and population projections, 2182  
 social change linked with, 2641  
 and socioeconomic attainment  
 and mobility, 975–976, 3043  
 two-child norm, 2182
- Family Size and Achievement* (Blake),  
 974–975
- Family Support Act of 1988, 967,  
 1261, 1988
- Family therapy, case studies of, 247
- Family Time and Industrial Time*  
 (Hareven), 2662
- Family violence, **981–988**  
 case studies of, 247  
 common law on, 981, 983  
 and corporal punishment,  
 982–983  
 and drinking, 984–985  
 explanations of, 983–985  
 and female victim's  
 attributions, 196  
 and homicide, 981–982  
 and marital rape, 950, 2577–2588  
 and nonfamily violence, 985  
 and pregnancy, 2234  
 statistics, 982–983  
 as statutory crime, 981  
*See also* Child abuse and neglect;  
 Childhood sexual abuse
- Fanon, Frantz, 66
- Fanshel, David, 2900
- Fantastic Lodge* (Hughes), 243
- Fararo, Thomas J., 1790, 2029
- Farber, Bernard, 1514–1515
- Farias, Paul César, 2135
- Faris, Ellsworth, 1833, 2220
- Faris, Robert E. K., 3055
- Farming. *See* Agricultural innovation;  
 Rural sociology
- Fascism, 539, 3002  
 conditions conducive to, 605  
 labor movement subordination  
 under, 1529  
 political party systems and, 2154  
 postmodernism and, 2206  
 structural theory on rise of, 2163
- Fashionable Nonsense* (Sokal and  
 Bricmont), 2208

- Fashions  
 as collective behavior, 348  
 compliance and conformity to, 403  
 counterculture, 461, 462  
 diffusion theories, 679  
 innovations and transmission of, 2170  
 mass tourism and, 3268  
 normative consensus and, 524  
 “Father figure” leader, 1566  
 Father-daughter incest, 1275–1276, 2583  
 Fatherhood  
 changed concept of, 2036–2037  
*See also* Family and household structure; Gender roles  
 Fauçonnet, P., 1024  
 Fazio, R. H., 339–340  
 FBI. *See* Federal Bureau of Investigation  
 FCC. *See* Federal Communications Commission; Federal Council of Churches  
 Feagin, Joe R., 57, 2211  
 Fear  
 authority and, 2520  
 of rape, 2576  
 as safeguarding mechanism, 2522, 2525–2526  
 Feather, Norman T., 3214–3215  
 Featherman, David L., 1685  
 Featherman, Jones, Jauser (FJH) hypothesis, 2712  
 Febvre, Lucien, 1933  
 Fecundity, 1006  
 Federal Bureau of Investigation, 493–494, 503, 530, 1486–1487  
 and desegregation protests, 2494–2495  
 juvenile rape arrests study, 1489  
 and political crime, 2143, 2146  
 on rape incidence, 2587  
*See also* Uniform Crime Reports  
 Federal Communications Commission, 242, 271–272  
 Federal Council of Churches, 2370  
 Federal Council of the Churches of Christ in America, 326  
 Federal Election Campaign Act of 1971, 2127  
 Federal Housing Administration, 845, 3072  
 loan program, 56  
 Federal Interagency Forum on Child and Family Statistics, 2685  
*Federalist Papers, The* (Madison), 721  
 Federici, Nora, 1465  
 Feeley, Malcolm M., 2961  
*Feeling for the Organism, A* (Keller), 994  
 Fei Xiaotong, 298, 300  
 Feifel, Herman, 581, 582  
 Fein, Helen, 1071  
 Feinleib, Manning, 1691  
 Feld, S., 2303  
 Feldman, Alison, 1872  
 Fellatio, 2553  
 Felson, Marcus, 506  
 Femininity/masculinity, **997–1005**  
 androgyny, 999–1000, 1002  
 anisogamy theory, 2885–2886  
 attributional style and, 196  
 development of, 998–999  
 feminist theory on, 2886  
 heterosexuality images and, 2568  
 learning theories of, 999  
 literary portrayals of, 2171, 2172  
 moral development theory on, 1900, 1902–1903  
 and personal dependency, 2064  
 psychoanalytic theory of, 998  
 psychological measure of, 999–1001  
 rape theories and, 2576, 2577–2578, 2579, 2581  
 research on, 1002–1003  
 roots of, 997–998  
 and sex role socialization, 2857  
 sociological view of, 1001–1002  
 and transsexuality, 2554, 2572–2573  
*See also* Sex differences; Sex stereotypes; Sexual orientation  
*Feminism: The Essential Historical Writings* (Schneir), 989  
 Feminist movement. *See* Women’s movement  
*Feminist Papers, The* (Rossi), 989  
 Feminist sociology. *See* Feminist theory  
 Feminist theory, **988–997**  
 American Sociological Association and, 153–155  
 as anti-anisogamy, 2886  
 and critical theory, 541, 542, 544, 545  
 on incest, 991, 1275–1276  
 and macrosocial inquiry, 1708–1709  
 and Marxist sociology, 989, 990, 1754–1755  
 on medical profession, 1816  
 and nature of knowledge, 994–996  
 and parental roles, 992–993  
 and participatory research, 2040, 2041  
 on peacekeeping, 2045  
 on personality development, 2089  
 and political correctness, 2141–2142  
 on popular culture, 2171–2172  
 pornography opponents, 274–275, 2186–2187  
 on prostitution, 2560  
 on psychiatric classification, 1833  
 and radical case studies, 246  
 of rape, 2576, 2578–2579, 2581, 2583  
 and redefinition of gender, 996  
 on science and scientific method, 2460, 2471, 2958  
 on social attachment, 1307  
 on social inequality of gender, 2692  
 and sociology of knowledge, 2958  
 and symbolic estates, 1512  
 and theories of crime, 505, 536  
 Third World, 1708  
 variations of, 993–994  
*Feminization of American Culture* (Douglas), 2171  
 Fenn, Richard, 2484  
 Fenollosa, E. F., 1477  
 Ferguson, Adam, 3098  
 Fernand Braudel Center, 1869, 1872  
 Ferracuti, Franco, 664  
 Ferrarotti, Franco, 1466, 1469, 1470



- Ferrero (Italian sociologist), 2520
- Fertility determinants, **1005–1012**  
 cohort perspective, 2678  
 definitions of, 1005–1006  
 demographic models, 620  
 demographic transitions, 425,  
 621–622, 624–629, 632–635,  
 704, 2178–2179  
 and education and development,  
 745–746  
 factors in, 623–625, 635, 1525,  
 2032, 2182  
 and fertility transitions,  
 1007–1010  
 and life-cycle stages, 1625  
 in pretransition societies,  
 1006–1007  
 replacement fertility, 2181  
 and social change, 2462  
 socioeconomic status as, 3043  
 Total Fertility Rate, 627
- Fertility rate  
 age-specific, 192, 193, 218–219  
 American Indian, 134  
 in China, 303  
 control measures, 1219  
 cumulative, 220  
 definition of, 2234  
 downward transition of, 1219  
 and family and household  
 structure, 924  
 global, 1219  
 “natural” populations, 2176  
 1950s anomaly, 1525  
 total, 192, 193, 194  
 in United States, 122, 125,  
 2032, 2180  
*See also* Birth and death rates;  
 Family size
- Fertility transitions  
 in Australia, 1007  
 downward, 1219  
 in Europe, 1007  
 and fertility determinants,  
 1007–1010  
 in France, 1007  
 in Middle East, 628, 1008, 1867  
 in Southeast Asia, 2976  
 in Third World countries, 1010  
 in United States, 1007
- Festinger, Leon, 2120, 2615  
 cognitive dissonance theory, 188,  
 335, 337–340, 2701  
 social comparison theory, 402,  
 2507, 2649, 2650–2651,  
 2652, 2653
- Fetal alcohol syndrome, 1640
- Fetus  
 AIDS/HIV transmission to,  
 2587, 2591  
 prenatal lifestyle risks to, 1640  
 viability, 2236  
*See also* Abortion
- Feudalism, 1932–1933  
 characteristics of, 2999  
 as conservative doctrine,  
 1600, 1601  
 revolution theory and, 2410, 2414  
 serfs under, 2597  
 stratification parameters,  
 2810–2811, 2812
- Feuer, Lewis S., 3068, 3069
- Feyerabend, Paul K., 823
- FHA. *See* Federal Housing  
 Administration
- Fiction. *See* Literature and society
- Fiedler, Fred, 1568, 1571, 2620
- Field, Mark G., 376
- Field, P., 2512
- Field research  
 in anthropology, 2893  
 description of, 2461  
 in ethnography, 852, 853–  
 854, 2893  
 and experiments, 887  
 quasi-experimental research  
 designs, 2309–2327  
 in social psychology, 2769
- Field research methods. *See* Case  
 studies; Ethnography;  
 Ethnomethodology; Sociocultural  
 anthropology; Qualitative  
 methods; Participatory research
- Field theory, **1012–1018**  
 current state of, 1013–1014  
 and group dynamics, 2611  
 and interpersonal power,  
 1458–1459
- Field Theory in the Social Sciences*  
 (Lewin), 1013
- Fienberg, Stephen E., 3036
- Fighting Back (drug abuse  
 program), 716, 717
- Figurational analysis, sport sociology,  
 2989, 2990
- Filial responsibility, **1018–1023**  
 affection vs. obligation in,  
 1020–1021  
 attitudes toward, 1019  
 and direct parent care,  
 1019–1020  
 and kinship systems, 1508–1509,  
 1514–1515  
 and legal mandates for financial  
 support, 1018–1019  
 and social policy, 1021–1022  
*See also* Long-term care
- Filiation, 1508
- Filipino Americans, 176, 177  
 academic studies of, 181  
 household structure, 127  
 immigration quotas, 175  
 occupations, 181, 182
- Film industry, 2172  
 X-rated films, 2185
- Final Exit* (Humphry), 585, 3084
- Financial bubbles, 726
- Financial panics, 726
- Findlay, Bob, 226
- Fine, Gary Alan, 1648, 2221
- Fine art, 173
- Finegold, Kenneth, 2165
- Finke, Roger, 2375, 2485
- Finkelhor, David, 292, 293, 1276,  
 2580, 2581, 2582
- Finland  
 health-care system, 374, 375, 377,  
 378, 379, 380  
 multilingualism, 2909  
 retirement patterns, 2407  
 social security system, 2800  
 unemployment, 3263  
 woman suffrage, 703  
*See also* Finnish sociology
- Finney, D. J., 3035
- Finnish Foundation for Alcohol  
 Studies, 2451

INDEX

- Finnish sociology, 2449, 2450–2451, 2452, 2453
- Finnish State Alcohol Monopoly, 2453
- Fiorina, Morris P., 3236
- Firsov, Boris, 2982
- First Amendment  
and peaceful protest right, 2265, 2269  
and regulation of expression, 268, 270–276  
*See also* Free speech
- First-cousin marriages, 1272, 1273, 1509, 1513
- First-order partial, 452
- Firth, Raymond, 289
- Fischer, Charles S., 3192
- Fischer, Joel, 2843
- Fischer, John L., 2896
- Fischer, Michael, 2221
- Fisher, B. M., 1566
- Fisher, G. A., 2299–2300
- Fisher, R. A., 454, 3006, 3035
- Fisher, S., 1718
- Fisher-Pitman test, 1957
- Fishman, Joshua, 2894
- Fiske, D., 2085, 3210
- Fisse, Brent, 530
- Fitzpatrick, Mary Anne, 1735
- 5-HIAA, 3079
- Five-Factor Model of Personality, 2079–2080, 2085
- Fixed-sample panel, 1687
- FJH hypothesis, 2712
- Flacks, Richard, 460
- Flashpoint model of collective behavior, 353
- Flechtheim, Ossip, 1038
- Fleming, Jacqueline, 2497
- Flexner, Abraham, 325, 2841
- Flick industrial group, 2123
- Fligstein, Neill, 738, 739
- Flinn, Kelly, 1880
- Flis, Andrzej, 2119
- Flowers, Gennifer, 2581
- Flynn, Charles, 1247
- Flynn, Edward, 2126
- Focus groups, online, 408
- Focus on the Family, 770
- Folic acid, 2235, 2238
- Folk Devils and Moral Panics* (Cohen), 1578
- Folk Islamic tradition, 2943
- Folk music, 1927
- Folkhälsoinstitutet (Sweden), 2451
- Folkloristics, literary, 1635
- Folkman, S., 2066, 3057
- Folkways* (Sumner), 2986
- Foner, Anne, 1618
- Food. *See* Agricultural innovation; Eating and diet; Eating disorders
- Food and Drug Administration, U.S., 88
- Food Stamp program, 2283
- For Marx* (Althusser), 721
- Foran, John, 1871
- Forbes, Charles, 2127
- Forced labor. *See* Slavery and involuntary servitude
- Forces of Order: Police Behavior in Japan and the United States* (Bayley), 2114
- Ford, Henry, 697, 699
- Ford, Henry Jones, 2124
- Ford, R., 1653
- Ford Foundation, 636, 637, 2398, 2401  
research funding in China by, 302
- Fordham Index of Social Health, The*, 2687
- Forecasting. *See* Futures studies as human and social activity; Social forecasting
- Foreign Corrupt Practices Act of 1977, 2127
- Foresight. *See* Futures studies as human and social activity
- Forest Service, U.S., 802
- Forging Industrial Policy* (Dobbins), 738
- Form, William, 423, 424
- Formal employment. *See* Labor force; Work and occupations
- Formal models, 2028
- Formality  
and historical sociology, 386  
and positivism, 2194
- Formes élémentaires de la vie religieuse, Les* (Durkheim), 1024
- Forrester, Jay W., 2662
- Forsyth, Donelson, 2418
- Forsyth, Elaine, 2729
- Fortes, Meyer, 289, 1507, 1508
- Forza Italia Party, 2129
- Foster, John, 1707
- Foucault, Michel, 227, 1199, 1303, 1707, 2173, 2892  
as Japanese sociology influence, 1479  
on nature of knowledge, 2757  
on social change, 2647–2648  
on social control, 2660
- Foucher, M., 1936
- Foundation Health Systems, 1822
- Foundation research grants, 2400–2401  
*See also specific foundations*
- Four-function paradigm (Parsons concept), 1554–1555, 1559–1560, 1978, 2005  
applied to organizational structure, 2005–2014
- Fourier, Charles, 2846, 2847, 2849, 3203
- Fourteenth Amendment, 270, 283, 587
- Fox, Renee, 583
- Fragmentation  
bureaucratic, 3103–3104  
political party, 2157–2158, 2159
- Frampton, Merle E., 1503–1504, 1505, 1506–1507
- France  
abortion policy, 2238, 2239  
African colonization, 60, 61  
age pyramid, 610, 612  
civil law system, 467, 474, 475, 476, 477, 478–479  
clinical psychology literature, 327  
comparative remarriage rates, 1749  
conditions conducive to democracy, 605, 606  
daily time use, 3160  
education and status attainment, 3045  
ethnic immigration, 636

INDEX

- and ethnic status  
 incongruence, 3051
- family planning, 2178
- family policy, 966
- fertility decline, 626, 2178
- fertility transitions, 1007
- governmental division of  
 power, 1953
- health-care system, 375, 377,  
 379, 380
- in Indochina, 2974, 2975
- job discrimination against  
 immigrants, 693
- juvenile violence, 1487
- labor market structure, 1987,  
 1988
- labor movement, 1529
- long-term care and care facilities,  
 1652, 1653, 1655, 1661
- migrant Arab workers in, 1865
- occupational mobility, 1987, 1988
- origination of secularization as  
 concept in, 2482–2483
- political and governmental  
 scandals, 2129, 2134
- political party system, 2154, 2159
- Political Systems Performance  
 Data, 2477
- and postmodernism, 2207
- relative marriage rate, 1749
- retirement practices, 2407, 2408
- social movement emergence,  
 2719
- Social Science Data Archive,  
 575, 576
- social security system  
 spending, 2800
- social surveys, 577
- sociocultural anthropology,  
 2891, 2892
- status incongruence, 3054
- time use research, 3161, 3164
- tourism in, 3167, 3169
- transnational corporations,  
 3175, 3176
- See also* French School of  
 Sociology, The
- Francis, Emmerich, 1075
- Franco, Francisco, 2129
- Frank, Andre Gunder, 640, 1706
- Frank, David J., 2662
- Frank, Jerome, 2961
- Frankfurt Institute of Social  
 Research. *See* Frankfurt School
- Frankfurt School  
 on art and culture, 173, 2173
- British sociology and, 226
- on conformity sources, 540
- critical theory tradition, 539–542,  
 543, 1752, 1754, 1757, 1758
- epistemology debate, 822, 1027
- German sociology and, 1075,  
 1076, 1077, 1078, 1079, 1081
- high culture/mass culture debate,  
 1645–1646
- Marxist sociology and, 540, 543,  
 1076–1077, 1732, 2756
- mass culture theories, 2168, 2169
- mass society theory, 1772
- Frankl, Viktor, 5, 460, 1617, 2084
- Franklin, Benjamin, 2456, 2483
- Franzosi, Roberto, 1199
- Fraser, Nancy, 545
- Fraternal lodges. *See* Voluntary  
 associations
- Fraud. *See* White-collar crime
- Fraudulent research. *See* Ethics in  
 social research
- Frazer, James, 1271, 2889
- Frederick the Great, king of  
 Prussia, 1435
- Frederickson, George, 55
- Free enterprise system. *See*  
 Capitalism
- Free market. *See* Capitalism;  
 Economic institutions
- Free rider concept, 604
- Free speech, 269, 280, 2908
- civil liberties and, 315
- civil rights movement and, 273
- hate speech and, 275–276
- rise of democracy and, 267–268
- student movements and,  
 3069–3070
- Vietnam War and, 273–274
- See also* Censorship and the  
 regulation of expression
- Free Speech Movement, 3069–3070
- Free will, 527–528, 1248, 1278, 2218
- Freedman, David A., 3039
- Freedman, Ronald, 425
- Freedom of expression. *See*  
 Censorship and freedom of  
 expression; Free speech
- Freedom of Information Act, 2401
- Freedom Riders, 2494, 2495
- Freedom Summer Project  
 (Mississippi), 2495
- Freeman, Alan D., 2961
- Freeman, Howard, 156, 1160
- Freeman, John, 2029
- Freeman, Linton, 2734
- Frege, Gottlob, 821
- Freidmann, Harriet, 2433
- Freidson, Eliot, 226, 1813–1814,  
 1815, 2263, 2264
- Freire, Paulo, 2040
- French, Hilary, 1230
- French language, 327–328
- French Revolution, 606, 1771, 2138,  
 2206, 2412, 2414, 2846,  
 2865, 3000
- and civil law system, 473–  
 474, 475
- historical sociology study of, 1198
- Marxist perspective on, 2410,  
 2411
- and mass murder and terror,  
 1066, 3137
- and origins of political parties,  
 2154–2155
- social dynamics of, 2664
- and status incongruence, 3050
- French School of Sociology, The,  
**1024–1029**
- and clinical sociology, 327–328
- contemporary, 1026–1028
- and economic sociology, 734
- and life histories and  
 narratives, 1633
- popular culture studies, 2170
- principal branches of, 1024–1025
- See also* France
- French structuralism, 563, 1027,  
 1032–1034, 1035
- Frenk, Julio, 374
- Frequency distribution, 658
- Frequency polygons, 659
- Freud, Sigmund, 1304, 1902, 2069
- on crowd behavior, 553, 559

- dependency theory, 2063, 2064  
depression description, 650  
family triads, 332  
and feminist theory, 990, 991  
identity formation theory, 2856  
on incest taboo, 1273, 1274, 1275  
moral development theory, 2089  
personality theory, 540,  
1713–1714, 2084, 2087, 2088,  
2090, 2092  
on primal horde, 1273, 1576  
on religious experience, 2367,  
2373, 2965  
sexuality theory, 1273, 1275,  
2537, 2565  
suicide theory, 3077  
“tender years doctrine,” 702  
Freyer, Hans, 1075, 1076  
Fried, Morton, 2891  
Friedkin, Noah E., 2673  
Friedland, Roger, 2166  
Friedman, Edward, 606  
Friedman, G., 1026  
Friedman, Milton, 722  
Friedman two-way analysis of  
variance, 1962–1963  
Friendship. *See* Interpersonal  
attraction; Social networks  
Friis, Henning, 2450  
Frith, Simon, 1925, 1927  
Fromm, Erich, 539  
Frontiers. *See* National border  
relations  
Frude, Neil, 1276  
Frustration, collective expressions  
of, 349  
Frustration-aggression hypothesis,  
73, 2670  
Fuchs, Josef, 1038  
Fukutake, Tadashi, 1479  
Fulton, Robert, 581, 582, 583  
Functional analysis, 1030, 1031  
Functionalism and structuralism,  
**1029–1037**  
and British structuralism,  
1034, 2890  
case studies, 244  
challenges to functionalism, 2416  
convergence theories and, 423  
criminal law and, 516  
definition of, 1885  
differentiation and, 2484  
educational mobility theories,  
757, 2927, 2928  
and French School of Sociology,  
1026–1027, 1032–1034  
and German sociology, 1078,  
1080, 1081  
and Japanese sociology, 1479  
juvenile delinquency theories,  
1493–1495, 1497–1498  
and kinship systems, 1502–  
1503, 1507  
legal systems comparisons, 472  
macro themes, 1704, 1705  
materialism and, 1784  
modernization theory and, 1885  
and penology, 2051–2054  
and Polish sociology, 2119  
political party origins and,  
2154–2155  
rational choice theory and,  
2341–2342  
role theory and, 2415, 2416,  
2417, 2421–2423, 2425  
and sentiments, 2522  
as sexual behavior approach,  
2537  
and social and political  
elites, 2626  
social control studies, 2657  
and social exchange theory, 2670  
and social inequality, 2690  
and social problems perspective,  
2759–2760  
and socialization theory, 2856  
social-structural model of law  
and, 1559–1560  
stratification theories, 2626,  
2813–2814  
*See also* Social structure;  
Structuralism  
Functionalist hypothesis, 2866  
*Functions of Police in Modern Society,*  
*The* (Bittner), 2108–2109, 2114  
Fundamental attribution error,  
194, 2751  
Fundamentalism, 2361, 2368–2373  
and censorship, 277  
coining of term, 2369  
comparative analysis of,  
2372–2373  
counterculture, 462  
definitions of, 2372, 2944–2945  
different meanings of, 2368,  
2368–2371  
global, 2371–2372, 3288  
Islamic, 2371, 2374, 2940–2941,  
2943, 2944–2945  
and legislation of morality,  
1577, 1580  
as response to secularization,  
2487–2488, 2966–2967  
social class and, 2378  
tensions with other religious  
orientations, 2386  
Fundamentalism Project, 2372–  
2373, 2945  
*Fundamentals, The* (anthology),  
2368–2369  
Funding of research. *See* Research  
funding in sociology  
Funerals, 582, 584  
*Furman v. Georgia* (1972), 2056  
Furstenberg, Frank, 1390, 1391,  
2393  
*Future of an Illusion, The*  
(Freud), 2965  
*Future of Marriage, The*  
(Bernard), 1736  
*Future of Religion, The* (Stark and  
Bainbridge), 2375  
Futures studies as human and social  
activity, **1037–1043**  
characteristics of, 1039–1040  
extrapolative, 1040–1041  
history of, 1038–1039  
limits of, 1040  
normative, 1040–1041  
prediction and, 2224–2231  
preferable futures concept, 2677  
and scenarios, 1041–1042  
and social forecasting, 2676–2681  
Futures Studies Internet Society,  
2231–2232  
Futuribles/Futuristics/Futurology.  
*See* Futures studies as human  
and social activity; Prediction and  
futures studies; Social forecasting

**G**

- Gabba, Carlo Francesco, 1464  
 Gabon, 2133  
 Gabor, Istvan, 2117  
 Gagnon, John H., 2539, 2550, 3091  
 Gaiser, Ted J., 408  
 Galanter, Marc, 471, 2961  
 Galbraith, Jay, 2011  
 Galbraith, John Kenneth, 724, 2921  
 Gale, Hugh, 3086  
 Gale Group, 1608  
 Galen, 1717, 2086  
 Galileo Galilei, 1781  
 Gall, Franz Joseph, 528, 529, 1717  
 Gallagher, John, 1266  
 Galli, Maria Callari, 1467  
 Gallier, Xavier, 1587  
 Gallino, Luciano, 1467  
 Gallup, George, 3232  
 Gallup polls. *See* Election polling; Public opinion; Survey research  
 Galpin, Charles J., 2428  
 Galston, William A., 356  
 Galton, Francis, 446–447, 550, 879, 2091–2092, 3005  
     eugenics theory, 1272  
     intelligence theory, 1360, 1361, 1364  
 Galton's problem, 550  
 Galtung, Johan, 639, 1467  
 Gambler's fallacy, 594–595  
 Gambling  
     and criminalization of deviance, 525  
     organized crime operations, 2017, 2019, 2021  
     as self-destructive behavior, 3077  
 Game theory and strategic interaction, **1045–1056**  
     altruism and, 118  
     characteristic function form, 1046  
     coalition formation and, 329–331, 332  
     conflict theory and, 414–416  
     differential games, 1049  
     dynamic games, 1048–1049  
     and economic sociology, 735, 2340  
     and equity theory, 2700  
     evolutionary games, 1049  
     and exchange networks, 2673–2674  
     experimental studies, 1047–1048  
     institutional analysis of, 1049–1051  
     mathematical sociology and, 1791  
     rational choice theory and, 2335, 2336–2337, 2338, 2419  
     representational forms of, 2337  
     and role theory, 2419  
     and social dynamics, 2666  
     and social values research, 3220–3222, 3223  
     solution concepts, 1046–1047  
     strategic form, 1046  
     supergames, 1049  
     theoretic concepts, 1045–1047  
     typology, 1046  
 Gamio, Manuel, 1858  
 Gamson, William A., 101–102, 645–646, 2270  
     on social movement successes, 2725  
 Gandhi, Indira, 2132  
 Gandhi, Mohandas, 1230, 2269, 3287–3288  
 Gang rape, 2580  
 Gangs  
     as counterculture, 460, 461  
     criminological research, 530, 533  
     cultural values and, 2171  
     as delinquent subculture, 509, 511, 512–513, 514  
     group norms and structure study of, 244, 363, 364, 365, 2611, 2613  
     macro-level deviance theories and, 663, 664  
     nineteenth-century urban, 1485  
 Gans, Herbert, 845  
 Ganzeboom, H. B. G., 2787  
 García, Carlos, 2131  
 Gardner, Howard, 1368, 1369  
 Gardner, John, 1384–1385, 2140  
 Garelli, Franco, 1473  
 Garfield, Eugene, 1610  
 Garfield, James, 2127  
 Garfinkel, Harold, 226, 856–857, 859, 2756  
     and conversation analysis, 431, 432  
     as founder of ethnomethodology, 246, 431  
 Garment industry, 2333  
 Garmon, Lance C., 1903  
 Garth, Bryant G., 1550–1551  
 Garvey, Marcus, 66  
 Garza, Gustav O., 1859  
 Gasparini, Alberto, 1468, 1472, 1473, 2288  
 Gaston, Berger, 1037  
 Gates, Bill, 1285  
 Gates Commission, 1877  
 GATT (General Agreement on Tariffs and Trade), 794, 1859  
 Gaulejac, Vincent de, 328  
 Gauset, Jessie Redmon, 66  
 Gauss-Markov theorem, 2251  
 Gaventa, John, 2040  
 Gay community. *See* Sexual orientation  
 Gay Liberation Front, 111  
 Gay rights movement, 315, 2719  
 Gaza, 1866  
 Gaze theory  
     collective gaze, 3173  
     male gaze, 2172  
     tourist gaze, 3166–3167, 3172–3173  
 Gdansk shipyard (Poland), 2268  
 GDOS (Group Development Observation System), 1979–1980  
 GDP. *See* Gross Domestic Product  
 Geddes, Patrick, 1290, 1291  
 Geen, R. G., 70  
 Geertz, Clifford, 387, 547, 852, 2855, 2891, 2906  
     Indonesian studies, 2976–2977  
     on Islamic society, 2941, 2944  
     semiotic theory, 2958  
     on world religions, 3289  
 Gehlen, Arnold, 1076, 1077, 1234  
 Geiger, Theodor, 2450  
 Geis, Gilbert, 3250, 3252  
 Gellner, Ernest, 227, 2941, 2942  
     theory of Muslim society, 2943–2944, 2945, 2946, 2947

- Gemeinschaft* (belonging), 2630, 3129, 3134
- Gemeinschaft-Gesellschaft* shift (Tönnies concept), 355, 362, 363, 364, 367, 1502, 1506, 2426, 2483, 3130, 3135, 3182
- Gemeinschaftsgefühl* (Adler concept), 1714
- Gender, 1057–1066**
- affirmative action and, 50, 1063, 2706
  - African-American education and, 2499
  - aggression and, 1452
  - aggressive tendencies and, 72
  - aging and, 83
  - AIDS/HIV contraction, 1641, 1642, 2589–2590
  - alcohol use differences, 94, 1640–1641
  - altruism and, 117
  - American Sociological Association and issues of, 153–155, 156
  - ascriptive character of, 2565
  - attitude formation toward, 185, 997
  - bankruptcy and, 204
  - and changing attitudes, 1063–1064
  - and childhood sexual abuse, 289, 291, 1275–1276, 2581
  - and childlessness, 109, 110
  - and collective mobility, 2716
  - and comparable worth, 369–372, 3048, 3265
  - conformity propensity, 403
  - and courtship, 486, 487, 489
  - and crime rates, 498, 534
  - and criminal sanctions, 517
  - and criminological theory, 504, 505, 511, 530, 536, 984
  - and cultural norms, 2528
  - differences between gay and lesbian relations, 111–112
  - differences between widows and widowers, 1749, 3255, 3257, 3258
  - differences in depression diagnosis, 649, 652–653, 1837, 1838
  - differences in diet and obesity, 1641
  - differences in language use, 2909
  - differences in marijuana use, 523
  - differences in marital communication, 1735–1736
  - differences in marriage, 1736
  - differences in reasons for marrying, 488
  - differences in suicide rate, 3078, 3079, 3081
  - differences in work orientation, 3275
  - direct and indirect discrimination, 689–691
  - as distinct from sex, 1057–1058, 2529, 2886
  - divorce petitioners and, 1747
  - divorce practices and, 112, 126, 620, 701–703, 707–708, 1425, 1526
  - double standard, 701
  - drug use breakdown, 710–711
  - earnings pattern and, 1624–1625
  - and economic determinism, 723
  - and economic sociology, 736
  - equality and, 107, 142, 302
  - See also subhead inequality below*
  - equality of opportunity and, 830, 832
  - ethnography and, 854
  - expectation states theory and, 882–883
  - feminist definition of, 2886
  - Finnish studies, 2453
  - hate speech and, 275–276
  - health promotion/health status and, 1165, 1166
  - helping behavior and, 117
  - historical variations in, 1199
  - identity and, 1001, 1061, 1256
  - impression formations and, 43
  - income disparity and, 645
  - inequality, 1062–1063, 2692
  - inequality and theories of crime, 504, 984
  - intergenerational relations and, 1389, 1391
  - interpersonal conflict resolution and, 1452
  - interracial marriage and divorce rates, 1750
  - in Islamic societies, 2948–2950
  - job satisfaction and, 3275
  - juvenile delinquency and, 1489–1490
  - kinship system changes and, 1502
  - labor-force composition and, 1524–1526, 3262
  - See also* Labor force, women's participation in
  - leadership image and, 1570
  - legislation of morality and, 1579–1580
  - leisure and, 1584, 1585, 1589
  - life cycle and, 1623–1625
  - life expectancy and, 1058, 1165
  - lifestyle risk-taking differences, 1639, 2559
  - literary sociology and, 1648, 2171–2172
  - love and, 1697, 1700–1701
  - marital age and, 124
  - marital equality and, 107
  - marital expectation differences, 1736
  - marital quality measurement by, 1730–1732, 1737
  - Marxist sociology and, 1754–1755
  - mate selection factors and, 1775–1776
  - medical profession and, 1816
  - mental health and illness and, 1837–1838
  - mental illness rates, 1837–1838
  - Middle Eastern studies, 1867
  - and military service, 1879–1880
  - moral development theories and, 1900, 1902–1903, 2089
  - music and, 1926
  - neutrality, 1057
  - Norwegian sociological studies, 2453
  - novel authorship and, 1648
  - nursing home residents and, 1667
  - as object of sexual orientation, 2565–2567, 2570
  - occupational aspiration levels, 2785

- and occupational segregation, 379, 2012, 3046, 3262, 3264–3265
- and occupational status attainment, 2000, 2785, 3046
- and ordination of women, 2379
- peer group segregation by, 2859
- personality theory and, 2089
- persuasion receptivity and, 2096–2097
- political alienation and, 102
- political correctness and, 2140, 2141–2142
- popular culture and, 2170, 2171–2172
- poverty and, 1288, 2033, 2215, 3048
- professions and, 2259, 2262–2263
- redefinition of, 996
- relationship changes, 704
- religion and, 1057
- as remarriage factor, 1748–1749
- remarriage satisfaction and, 2390
- retirement and, 2406–2407
- and rural sociology, 2430
- schematization, 1000
- and scientific explanation bias, 2471
- and self-esteem development, 2513
- self-serving attributions and, 196
- sex differences studies, 2529–2535
- sexual harassment and, 2580–2581
- singlehood and, 107–108, 124–125
- single-parent households and, 127, 128
- smoking and, 1640
- and social comparison bases, 2654
- social construction of, 1057–1058–1059
- and social security system inequities, 2802
- and social stratification, 2817–2819
- Southeast Asian studies, 2976
- and status attainment, 3044, 3046
- stereotypes. *See* Sex stereotypes
- and structural lag, 3063
- Swedish studies, 2453
- symbolic interaction theory and, 1001, 1002
- time use research, 3160–3163, 3164
- and transvestism and transsexuality, 2572–2573
- underemployment inequality and, 1720–1721, 1722
- in unilineal vs. bilateral kinship systems, 1507–1508
- and voluntary association membership, 3228
- and white-collar crime offenders, 3251
- widow/widower differences, 3255, 3257, 3258
- widow/widower remarriage rates, 1749
- widowhood demographics, 126, 3256–3257
- and women's high status in Southeast Asia, 2976
- work and, 122, 1059–1062, 3275
- See also* Occupational segregation
- See also* Femininity/masculinity; Feminist theory; Gender roles; Sex stereotypes; Sexual behavior patterns; Women
- Gender identity. *See* Femininity/masculinity
- Gender preference
- family planning and, 628
- family size and, 2182
- Southeast Asian neutrality on, 2976
- Gender roles
- adulthood, 30, 33, 34
- alternative lifestyles, 111
- American family, 122
- case studies, 246
- child custody awards, 702
- division of labor, 696, 1057, 1062, 1064
- divorce effects on, 703, 704–705
- fatherhood, 2036–2037
- femininity/masculinity concepts, 997
- fertility transition effects on, 625, 627, 634
- and filial responsibility, 129, 1020, 1022
- in Islamic society, 2949–2950
- mate selection theories and, 1775
- media stereotypical portrayals of, 1699
- in Mexico and Brazil, 941, 1860–1861
- motherhood, 2036
- parental roles and, 2035
- and personal dependency, 2064
- role theory and, 2418
- sex differences and, 2418
- socialization process and, 2857, 2886
- Gene therapies, 1824
- Genealogical mapping, 1512–1516
- Genentech, Inc., 1824
- General Accounting Office, 838
- General Agreement on Tariffs and Trade, 794, 1859
- General Directorate for Science and Technology (European Commission), 580
- General Economic History* (Weber), 733
- General Inquirer, The* (Stone et al.), 1979
- General linear model, 162–163, 457
- General Social Survey (GSS), 166–167, 317, 548, 576–577, 1479, 2475
- attitude items, 190, 3093
- cross-sectional samples, 1684
- on family size, 973, 977
- on sexual behavior, 2538, 2541, 2545
- and social indicators, 2682, 2684
- General systems theory, 1558, 1931
- General theory of action, 1197
- General Theory of Secularization, A* (Martin), 2485
- Generalized belief, 353
- Generalized other, 3097
- Generation life table, 612, 614
- Generational equity, 1388, 1398
- Genesis, 1511, 2369, 3285
- Genetic psychology, 1274

INDEX

- Genetic structuralism, 1027, 1028
- Genetics
- as aggression influence, 69, 70
  - as alcoholism factor, 97
  - altruism theory, 115
  - as depression risk factor, 652, 1836
  - as educational attainment factor, 2931
  - and incest taboos, 1273, 1274
  - as intelligence factor, 1369–1373, 2090, 2140, 2330
  - Mendelian, 2880, 2881
  - and personality traits, 2088, 2089–2090, 2092
  - and puberty, 6
  - and public policy concerns, 1824
  - as schizophrenia factor, 1836
  - and self-esteem, 2513
  - and sexual selection theory, 2885
  - as suicide tendency factor, 3079
  - See also* Nature vs. nurture
- Geneva Conference on Vietnam (1954), 332
- Geneva Convention (1949), 1244
- Genital intercourse. *See* Heterosexuality; Sexual behavior patterns
- Genocide, **1066–1073**
- contemporary, 1070, 1071
  - definitions of, 1066, 1071–1072
  - early historical, 1068–1071
  - from environmental destruction, 1221
  - as ethnic cleansing, 1944, 1946, 1948
  - ethnic hatred and, 2529
  - and ethnonationalism, 1944, 1946, 1948
  - and eugenics, 879
  - “genocidal massacre” distinction, 1072
  - human rights and, 1244
  - ideological, 1067, 1069–1070
  - ideological vs. instrumental, 1067, 1070
  - and intelligence of targeted groups, 1384
  - and international law, 1244, 1429
  - Nazi Holocaust, 1066, 1067, 1070, 1384, 2206
- Genovese, Eugene, 54
- Genovese, Kitty, 115
- Gentile, Giovanni, 1465
- “Gentleman’s Agreement” of 1907–1908, 175, 176
- Gentrification, 3196
- Genuine Progress Indicator, 2687
- Geographical Information System, 411, 580
- Geography
- applications to sociology, 2921–2923
  - and economic institutions, 728–730
  - and population changes, 2179–2180
- Geopolitics, 1933
- George Washington Medical School, 588
- George Washington University, 1422
- Georgia (republic), 2362, 2982
- Georgia State University, content-analysis Web site, 421
- Georgianna, Sharon, 3079, 3081
- Gerard, Harold, 402
- Gereffi, Gary, 643
- Gergen, Kenneth J., 117
- Gergen, Mary M., 117
- German idealism, 1248–1249
- German Sociological Association, 1076, 1078, 1081
- German Sociological Society, 1423
- German sociology, **1073–1084**
- action theory, 1080
  - behaviorism, 1080
  - classical period of, 1074
  - Cologne School, 1077, 1078
  - consolidation, 1080–1081
  - critical theory, 1076, 1077–1078, 1079
  - divisions in, 1076
  - in East Germany, 2117
  - economic, 732–734
  - expansion of, 1077–1080
  - functionalism, 1080
  - historical materialism, 1080
  - idealism, 1248–1249
- as Japanese sociology early influence, 1477–1478
  - life histories and narratives, 1633, 1635
  - literary sociology, 1648
  - medical sociology, 1814
  - and Nazi regime, 1074, 1075
  - and New Left, 1079
  - as “normal science,” 1081–1082
  - positivism controversy, 1077
  - reconstruction era, 1075–1077
  - Schelsky school, 1077
  - Social Science Data Archive, 575, 576, 580
  - and sociology of knowledge, 2953–2954
  - structural functionalism, 1078
  - systems theory, 1080
  - time use research, 3162, 3163, 3164
  - traits of, 1073
  - Weimar period, 1074–1075
  - Young Turks in, 1075, 1076, 1077, 1078, 1080
  - See also* Frankfurt School; Germany
- Germani, Gino, 1537
- Germany
- African colonization by, 60
  - apprenticeship system, 27
  - civil law system, 471, 474, 475–476, 477, 478, 479, 480
  - communitarianism, 361, 362
  - conditions conducive to democracy, 605, 606
  - data archive, 2477
  - dictatorship, 3002
  - disaster research, 687
  - divorce laws, 703
  - divorce rate, 706
  - East German centrally managed privatization, 2851
  - epistemology, 819–820
  - ethnic immigration, 636, 693
  - ethnic status incongruence, 3051
  - family “legacy of silence” in, 1512
  - family policy, 966
  - governmental division of power, 1953



INDEX

- health-care system, 374–375, 376, 377, 379–380
- high suicide rate, 3079, 3082
- historical city-states, 2998
- job discrimination against immigrants, 693
- kinship mapping priority, 1513, 1515
- labor movement, 1529, 1531
- lawyers in, 477, 478
- Lebensraum* concept, 1933
- legal system, 1550
- long-term care and care facilities, 1653, 1659, 1660, 1661
- migrant Arab workers in, 1865
- Nazi era. *See* Nazis
- and Nazi protests in 1920s, 2268
- occupational mobility, 1987
- organizational demographics, 395
- political party system, 2154, 2157, 2159, 2163
- political scandals, 2130
- relative marriage rate, 1749
- research university model, 1180
- retirement practices, 2407–2408
- revolution, 3001
- social security system, 2796, 2797, 2800
- social surveys, 577
- transnational corporations, 3175, 3176
- unemployment, 3263
- Weimar Republic voting behavior research, 3233
- woman suffrage, 703
- See also* Germany sociology
- Gerontology. *See* Aging and the life course; Filial responsibility; Intergenerational relations; Intergenerational resource transfers; Long-term care; Long-term care facilities; Retirement; Widowhood
- Gerth, Hans, 2069
- Gesell, Arnold, 1894–1895, 2092
- Gesellschaft*. *See* *Gemeinschaft-Gesellschaft* shift
- Gesellschaft und Demokratie in Deutschland* (Dahrendorf), 1078
- Gestalt psychology
- cognitive consistency theories and, 334
- and group cohesiveness, 2614
- personality theory, 1718, 2085
- Getting a Job* (Granovetter), 736
- Getting into Print* (Powell), 1647
- Ghana, 66
- family size, 978
- health-care system, 381
- Ghana Empire, 2999
- Ghetto-specific cultural traits. *See* Urban underclass
- Ghost Dance, 137
- Ghurye, G. S., 1290, 1291, 1292
- Gibbs, Jack, 504, 3079
- Gibbs, James L., Jr., 464
- Gibbs, Jewelle Taylor, 3079
- Gibbs, John, 194, 1899
- Gibbs sampler, 3039
- Giddens, Anthony, 226, 227, 228, 1033, 1086, 1090, 1212, 1224, 2642, 2983
- Japanese sociology and, 1479
- on money's symbolism, 1890
- on postindustrial society, 2201, 2202
- theory of structuration, 1710
- Giddings, Franklin, 1424, 3232
- as Japanese sociology early influence, 1477
- and positivism, 2192, 2193
- Gideon v. Wainwright* (1963), 2960
- Gieryn, Thomas F., 2455, 2458, 2459
- Gift, The* (Mauss), 103, 734
- Gifted children, 324
- Gilbert, G. Nigel, 2459
- Gilbert, Rosewell, 38083
- Gilded Age, 2125
- Gilder, Georges, 723
- Gilligan, Carol, 993–994, 1307, 1900, 1902, 1903, 2089, 2172
- Gillmore, Samuel, 1853–1854
- Gilman, Albert, 2894
- Gilman, Charlotte Perkins, 852, 989
- Gilpin, Robert, 3242
- Gilroy, Paul, 226
- Gini, Corrado, 1423, 1424, 1465
- Ginkgo, 654
- Ginsberg, Benjamin, 2273
- Gioja, Melchiorre, 574
- Giorio, Giuliano, 1472
- GIS (Geographic Information Systems), 411, 580
- Giscard d'Estaing, Valéry, 2129, 2134
- Gillow v. New York* (1925), 270
- Glaser, Barney, 582, 2290
- Glaser, Daniel, 510
- Glaser, W. A., 410
- Glasnost, 2981
- Glass, G. V., 1843
- Glazer, Nathan, 847, 1595, 1601, 2333, 2818
- Glendon, Mary Ann, 356, 472
- Glenn, Jerome C., 1039
- Glick, Clarence E., 324
- Global environmental change (GEC), 810–811
- Global environmental problems, 801, 805
- and futures studies, 1041
- Global Media Monitoring Project, 1768
- Global properties of collectives, 1591
- “Global village” concept, 428
- Globalization and global systems analysis, **1084–1098**
- agricultural innovation and, 87–88
- American society and, 140
- boundaries and, 1084
- British sociology and, 227–228
- capitalism and, 242, 1084, 1085, 1091
- censorship and, 280
- child labor and, 3262
- class and race relationship and, 322
- competition and, 1085, 1091
- and conditions for future social changes, 2646
- convergence theories, 427–428
- core/periphery/semiperiphery structure in, 1089, 1265
- criminal and delinquent subcultures and, 513
- cross-cultural analysis, 549

INDEX

- current systems analysis and, 1090–1094
- democracy and, 606
- dependency theory and, 1087–1089
- development issues and, 1706
- diffusion and, 1085
- disembedding and, 1086
- division of labor and, 697
- economic institutions and, 729, 1197, 1225, 1313
- feminist perspective on, 1708
- fundamentalism and, 2371–2372
- and Islamic militancy effects, 2947–2948
- Japanese sociology and, 1481–1484
- labor-force participation and, 1088, 1526, 3267, 3275
- legal systems and, 1550–1551
- leisure and, 1589
- Marxist theory and, 1084–1086
- and Marxist theory rethinking, 1757, 1758
- mass media research, 1767–1768
- materialist theory on, 1784–1785
- medical-industrial complex and, 1827
- in Mexican studies, 1861
- migration and, 1092
- modernization theory and, 1084–1086, 1886–1887
- music and social structure link and, 1924–1925
- nationalism and, 1944–1945, 1948
- peacekeeping and, 2045–2046
- postindustrial theory and, 2202
- postmodernism and, 2206–2207
- radical case studies, 246
- risk in, 1086
- rural sociology and, 2432, 2433
- sex industry and, 2560–2561, 2607
- social dynamics of, 2665–2666
- and social inequality, 1087, 1088, 2691, 2705–2708
- and social movement emergence, 2719
- state system as basis of, 1092, 1093, 2362
- and territorial belonging, 3135
- and terrorist tactics, 3139–3140
- trade and, 1087, 1088
- transnational corporations and, 3174–3180
- urbanization and, 306, 309, 311
- and war causes, 3243–3244
- world system theory, 1089–1090, 1197, 1199, 1706, 1758, 1876, 2646
- Globe, Le* (French periodical), 2846
- Glock, Charles Y., 2373
- Gluckman, Max, 289, 1549, 1550
- Glue sniffing, 713
- Gnoseology. *See* Epistemology
- GNP. *See* Gross National Product
- Goals
- anomie and, 165, 1494
  - effectiveness of democracy in achieving, 606
  - group size and, 1118–1121
  - group vs. individual, 595–596
  - intrinsic/extrinsic motivation and, 2059–2060
  - organizational orientation toward, 393–394, 2008–2011
  - strain theory of deviance and, 664, 1494–1495
  - values and, 3213
  - values differentiated from, 2829
- Godet, Michel, 1037, 1041
- Goffman, Erving, 707, 776–777, 852, 853, 995, 1577, 1932, 2173, 2901, 2902, 2907, 2908
- affect control theory, 41, 46
  - and Chicago School, 244
  - on face-to-face interaction, 431, 432
  - on interaction rituals, 416, 2523
  - on medicine as social control, 1815–1816
  - pragmatism theory, 2423
  - role theory, 2417, 2423, 2425, 2506
  - sociological theory of personality, 2080
  - on total institution, 1673, 1674
- Gold, Martin, 1013
- Gold Coast and the Slum, The: A Sociological Study of Chicago's Near North Side* (Zorbaugh), 324
- Golden Hordes, The* (Turner and Ash), 3167
- Goldhagen, Daniel, 387
- Goldman, Emma, 989
- Goldmann, Lucien, 1645
- Goldschmidt, Walter, 88
- Goldstein, Joshua, 3242
- Goldstein, Kurt, 1714, 1718, 2087, 2088
- Goldstone, Jack, 390, 1707, 1708, 2413
- Goldthorpe, John H., 224, 425
- Goleman, Daniel, 1368
- Golod, Serguei, 2982
- Gomez, Juan Vicente, 2134
- Gonorrhea, 2582, 2583
- Gonzalez, Felipe, 2129
- “Good neighbors” studies, 115
- Good society model, 357, 360
- Goode, Patrick, 1752, 1753
- Goode, William J., 1258–1259, 1503, 1504, 1615, 2823
- Goodman, Andrew, 2495
- Goodman, Leo A., 1594–1595, 1811, 3036, 3038
- Goodman's gamma, 661
- Goodman's tau, 661
- Goodness-of-fit test, 1964–1966
- Goodwin, Charles, 2905
- Goodwin, William, 1233
- Goody, Jack, 1515–1517, 2906
- Gorbachev, Mikhail, 2981
- Gordon, Leonard, 2270
- Gordon, M. M., 2883
- Gordon, Milton, 53, 842–843
- Gordon, Robert, 1382
- Gordon, Steven, 782
- Gordon, Theodore J., 1039
- Gore, Albert, Jr., 2686
- Gore, M. S., 1293
- Gospel music, 1926
- Gotay, Carolyn C., 2301, 2305–2306
- Gotham, Kevin Fox, 1636
- Gottfredson, Michael, 505, 507, 535, 667

- Gottman, J. M., 3143, 3149, 3151
- Gould, Stephen Jay, 995
- Gouldner, Alvin W., 233–234, 325, 327
- Governing elite. *See* Social and political elites
- Governing Prisons* (DiIulio), 2053–2054
- Government corruption. *See* Political and governmental corruption
- Government division of powers, 1952–1953
- Government regulation, **1098–1111**  
 censorship as, 267–281  
 and comparative health-care systems, 374, 375, 378–381  
 definition of, 1098–1100  
 deregulation and reregulation, 1103–1105  
 deregulation/white-collar crime relationship, 3250  
 economic, 1100–1101, 1311, 1337  
 effectiveness of, 1102  
 empirical studies of, 1106–1109  
 and ethics in social research, 838–839  
 and family policy in Western societies, 963–964  
 origins of, 1101  
 reform of, 1104  
 theories of, 1100–1103  
*See also* Public policy analysis; Social security systems
- Grabb, Edward G., 1754
- Graebner, Fritz, 675
- Graen, G., 1567
- Graft, 2124, 2125, 2126, 2127, 2137
- Graham, Billy, 2370
- Graig, Laurene A., 376, 377
- Grammar analysis, 438–439, 2297–2298  
 of social interaction, 2897–2899
- Gramsci, Antonio, 1753, 1758, 1772  
 British sociology and, 226  
 power elites/popular culture relationship, 2169, 2985–2986  
 and structuralist theory of state, 2163
- Grand Tour, 3166
- Grandparents, 131, 1390–1391  
 cultural influences on roles of, 696  
*See also* Filial responsibility; Intergenerational resource transfers
- Granovetter, Mark, 349, 350, 735–736, 2029, 2419  
 “strength of weak ties” theory, 2693, 2731–2732, 2791, 2827
- Grant Foundation, 2398
- Grant requests. *See* Research funding in sociology
- Graphics theory and systems  
 application, 1790  
 computer, 411  
 descriptive statistics, 658–659  
 history of, 3005–3007  
 perception research, 3009–3011  
 relationships between summary statistics and, 661  
 and social network analysis, 1789–1790, 2729  
 statistical, 3003, 3007–3022  
 values ranking, 3217
- Grassroots protests, 791, 803
- Gratian, 1507
- Gratification, delayed vs. short-term, 668
- Grattet, Ryken, 2961
- Gray, Louis N., 1790
- Gray market, 2019, 2021
- Great Britain. *See* United Kingdom
- Great Depression  
 agricultural innovation and, 88–89  
 and American Sociological Association membership decline, 148  
 and birth rate decline, 2032  
 and census analysis, 283  
 cohort effects, 80, 1625, 2861  
 collapse of democracies during, 3002–3003  
 community studies on social effects of, 363–364  
 conservative policies blamed for, 1597  
 and divorce rate, 125, 702  
 and industrial sociology, 1310–1311  
 and labor movements, 1530  
 and labor-force measurement concept, 1521  
 life-course study of children born in, 1618  
 and pension plans, 2402–2403  
 and rural sociology research support, 2427  
 and social security benefits, 2798  
 and voter realignment, 3235
- Great Migration, 532
- Great refusal (Marcuse concept), 541
- Great Society, 2299
- Great Zimbabwe Empire, 2999
- Great-man theory, 1564–1565
- Greece  
 clinical sociology, 328  
 Cyprus conflict, 1945  
 educational status attainment, 2784–2785  
 government and political corruption, 2130  
 long-term care and care facilities, 1652, 1653  
 organizational demographics, 395  
 revolution, 3001  
 student protest, 3067
- Green, Gary S., 3245, 3246, 3247
- Green Revolution, 89–90, 252–253, 1222
- Greenberg, David F., 1692, 2961
- Greenberg, Jerald, 2698
- Greenfeld, Liah, 173
- Greening of America, The* (Reich), 1773
- Greenpeace, 2149, 2725
- Greenwood, Ernest, 2841
- Greer, Scott, 1223
- Gresham’s Law, 729
- Greve, Heinrich C., 2668
- Grief  
 of bereavement, 582, 648  
 depression compared with, 650
- Grief, Esther B., 1898
- Griff, Mason, 172–173
- Griffin, Larry, 1636
- Griffin, Monica D., 1648, 1649

- Griffin, S., 2576
- Griggs v. Duke Power Co.* (1971), 691
- Grimshaw, Allen D., 2298, 2896,  
2901, 2903, 2904, 2905,  
2906, 2907
- Griswold, Wendy, 173, 1644,  
1646–1647, 1650
- Groat, H. T., 2346, 2351
- Grodzin, Morton, 181
- Gross, Feliks, 1467
- Gross, Michael L., 1901
- Gross, Neal, 87, 677, 2415,  
2416, 2429
- Gross Domestic Product (GDP)  
and comparative social security  
system spending, 2799, 2800  
and dependency theory, 641,  
642, 643  
and family policy, 928  
in Middle Eastern countries, 1865
- Gross National Product (GNP)  
developing country disaster  
effects on, 685  
and education and development,  
745, 748–749  
and modernization theory, 1883  
table of selected Muslim  
countries, 2938
- Grotius, Hugo, 1427, 1559
- Grounded theory, 2290, 2468
- Group behavior. *See* Collective  
behavior; Crowds and riots;  
Group conflict resolution; Small  
groups; Social networks
- Group cohesiveness. *See*  
Interpersonal attraction; Small  
groups
- Group conflict resolution,  
**1111–1117**  
contact theory and, 1114–1115  
in decision making, 595–597  
mediation and, 1115  
methods of, 1112, 1113  
peacekeeping and, 1115  
restorative justice and, 1116  
social identity theory and,  
1112, 1113  
third-party intervention and, 1115
- Group Development Observation  
System, 1979–1980
- Group Dynamics* (Cartwright and  
Zander), 2611
- Group dynamics (Lewin concept),  
1014, 2611
- Group homes, 1656
- Group majority. *See* Compliance and  
conformity
- Group problem solving. *See*  
Decision-making theory and  
research; Group conflict  
resolution; Group process;  
Small groups
- Group process  
field theory and, 1014  
interaction and problem solving,  
2617–2619  
intergroup relations and,  
1399–1406  
membership and nonmembership  
groups, 2634–2635  
norms and controls, 2617–2620,  
2774–2775  
participatory research, 2038,  
2039, 2613  
personal dependency and, 2065  
reference group perception,  
2752–2753  
role concepts and, 2422–2423,  
2774–2775
- Group size, **1117–1124**  
coalition triad, 329–332, 335–336  
communication and, 1122–1124  
as conformity factor, 403  
division of labor and, 676  
dyad-triad, 1117  
goal orientation and, 1118–1121  
individual effort related to, 2619  
social dilemmas and, 1121–1122  
task structure and, 1118,  
1119, 1120  
*See also* Small groups
- Group-awareness training, 898–900
- Groups  
sociological definition of  
concept, 2610  
*See also* Collective behavior;  
Group conflict resolution;  
Group process; Group size;  
Small groups
- Groupthink, 400–401, 2615
- Grube, Joel W., 3213, 3214
- Gruder, Charles L., 2651–2652
- Grushin, Boris, 2982
- GSS. *See* General Social Survey
- GSSDIRS interactive system, 409
- Guadalajara, Mexico, 1859
- Guaranteed income, 2213
- Guatemala  
demographic characteristics,  
1535, 1536  
drug trafficking, 2135  
fertility decline, 627  
poverty, 2216
- Gubert, Renzo, 1468
- Gubrium, Jaber, 245, 247, 855,  
1636, 2304  
interviews of nursing home  
residents, 1673–1674
- Gudjonsson, G., 70
- Guerra, Alfonso, 2129
- Guerra, Juan, 2129
- Guerrilla warfare, 1537, 2362,  
3138, 3139
- Guest, Avery, 368, 626
- Guggenheim Foundation, 2398
- Guichard, P., 1515, 1517
- Guide to Resources and Services* (ICPSR  
data archives), 2474
- Guidicini, Paolo, 1468, 1472
- Guild Socialism, 2196
- Guilt, 2773
- Guinea, 2938
- Guinea-Bissau, 2216
- Guinness affair, 3250
- Guinier, Lani, 58
- Gulf Co-operation Council, 1944
- Gulf War. *See* Persian Gulf War
- Gullestad, Marianne, 1512
- Gulliver, P. M., 1549
- Gumperz, John J., 2894, 2904
- Gumplowicz, Ludwig, 1074,  
1423, 2117
- Gunnlaugsson, Helgi, 2452, 2453
- Gupta dynasty, 2999
- Gurin, Gerald, 2303
- Gurvitch, G., 1026, 3155
- Gusfield, Joseph, 855, 1576–1577,  
1578, 1580, 2173, 2222
- Gutierrez Rebollo, Jesús, 2135
- Gutman, Herbert, 1707, 2333

- Guttman, Louis, 906, 912, 913,  
1801–1802, 2298, 3038
- Guttman scaling, 1801–1802, 2297
- Gypsies, 1067, 1384
- H**
- Haavio-Mannila, Elina, 2453
- Habenstein, Robert W., 483–  
484, 584
- Haberman, Shelby, 3036
- Habermas, Jurgen, 2484, 2647, 2983  
and British sociology, 226  
critical theory, 539, 542–544, 545  
and epistemology debate, 822  
evolutionary model, 1705  
and German sociology, 1077,  
1079, 1080, 1081  
and Japanese sociology, 1479  
postmodernist rejection of, 2206  
and social philosophy, 2756–2757  
on state censorship, 268
- Habibie, B. J., 2131
- Habit, 3100
- Hacienda* system (Mexico), 1857
- Hadden, Jeffrey K., 2372, 2375
- Haeckel, Ernst, 1, 1209
- Hagan, John, 1490, 1498, 2961, 3252
- Hagen, Everett E., 1886
- Hagestad, Gunhild, 1390
- Hagiwara, Shigeru, 198
- Hague, Frank, 2126
- Hague Opium Convention of  
1912, 713
- Haiti  
demographic characteristics,  
1535, 1536  
fertility decline, 627  
Gross National Product, 1535  
political corruption, 2134  
slave rebellion, 2600  
vodun, 65  
wartime rape, 2579
- Hajj, 3282
- Halas, Elzbieta, 2119
- Halbwachs, Maurice, 1024, 1424
- Haley, Alex, 67
- Hall, Edward T., 1978
- Hall, G. Stanley, 1, 2
- Hall, Jerome, 2961
- Hall, John, 224, 567, 568, 606
- Hall, Peter, 2221, 2222, 3096
- Hall, Richard, 232, 3270
- Hall, S., 1578
- Hall, Stuart, 226, 1756, 2958
- Hall, Thomas D., 645
- Halle, David, 1649
- Halle, Morris, 2899
- Haller, Archibald O., 2781,  
2782–2783, 2788, 2867
- Hallinan, Marianne, 197
- Hallinger, Philip, 2418
- Halmos, Peter, 2845
- Halsey, A. H., 225
- Hamilton, Charles, 53
- Hamilton, David L., 2244, 2245
- Hamilton, Gilbert, 1726
- Hamilton, Harry, 677–678
- Hamilton, V. Lee, 198
- Hammen, Constance, 654
- Hammond, Phillip, 2373
- Hammurabi's code (2270 B.C.), 1485
- Han dynasty, 2998–2999
- Hanassab, Shideh, 1872, 2999
- Handbook of Economic Sociology*  
(Smelser and Swedberg),  
735, 2921
- Handbook of Experimental Economics*  
(Kagel and Roth), 891
- Handbook of Research Design and  
Social Measurement* (Miller), 3209
- Handbook of Sociology*, 583
- Hankiss, Elmer, 2117
- Hannan, Michael T., 1693, 1694,  
2029, 2296–2297, 2668
- Hans, S., 2098
- Hansa-Rhine-Italy trade routes, 2922
- Happiness. *See* Life Satisfaction  
Index; Quality of life
- Harary, Frank, 336, 2415, 2417
- Harbison, Frederick H., 422
- Hardin, G. R., 595–596, 1220
- Harding, Warren, 2127
- Hare, A. Paul, 1977–1978
- Hare Krishna, 460, 461–462, 2366,  
3287–3288
- Hareven, Tamara, 2662
- Hargis, Billy James, 2370
- Harper, Charles L., 2887
- Harray, Frank, 1034
- Harré, Rom, 823
- Harris, Lou, 2371
- Harris, Louis, 1395
- Harris, Marvin, 1705, 2891
- Harris v. Forklift Systems, Inc.*  
(1993), 2580
- Harrison, David, 2887
- Harrison, Gualtiero, 1467
- Harrison Act of 1914, 712
- Harroff, Peggy, 1731
- Hart, Keith, 1337
- Harter, S., 2512–2513
- Hartford Seminary, 2377
- Hartman, H. I., 370
- Hartman, P. A., 3188
- Hartmann, Heidi, 1708
- Hartshorne, H., 114–115, 2083
- Harvard Business School, 1978
- Harvard Civic Engagement  
Project, 3229
- Harvard III Psychosocial  
Dictionary*, 1979
- Harvard Medical School, 582  
death and dying study, 586  
depression in artists study, 655
- Harvard University, 2193, 2766,  
3099  
leader behavior study, 1565  
observation laboratory, 1979  
student movement, 3069
- Harvey, Andrew S., 3157
- Harvey, Clyde C., 2416
- Hasan, Mohamme (“Bob”), 2131
- Hashimoto, Ryutaro, 2131
- Hastings Center, 585, 586, 587–588
- Hate, 2529
- Hate crimes, 2761, 2764, 2908
- Hate speech  
and censorship, 275–276  
codes, 2140  
communitarian view of, 360  
and political correctness, 269
- Hatt, Paul K., 198, 1997
- Haug, Marie, 1815
- Hausa, 1900
- Hauser, Philip, 632, 1720, 2354
- Hauser, Robert, 366, 2481, 3036

- Haushofer, Karl, 1933
- Havel, Vaclav, 1230
- Hawaii, 177
- Hawkins, David, 366
- Hawkins, Gordon, 2054
- Hawkins, James L., 1727
- Hawley, Amos H., 1457, 3129  
     human ecology theory, 1210,  
         1211, 1212, 1213, 1215, 1218,  
         1225, 1226, 1228  
     and urban sociology, 3193,  
         3197, 3198
- Hawthorne effect, 889, 2325–2326
- Hayano, David, 245
- Haynes, George Edmund, 326
- Hayward, M. D., 1632
- Hazan, C., 2068
- Hazard, John N., 1555
- Hazards Reduction and Recovery  
     Center (Texas A&M), 682
- HBO and Company, 1822
- HCA. *See* Hospital Corporations  
     of America
- HCA Columbia, 1821
- Health and health care. *See* Medical-  
     industrial complex
- Health and Human Services  
     Department, U.S., 838,  
     2398–2399, 2576
- Health and illness behavior,  
     **1127–1136**  
     applications of research in, 1134  
     depression and, 655  
     health behaviors, 1127–1130  
     health belief model, 1127,  
         1128, 1814  
     illness behaviors, 1130–1134  
     longevity and, 1131  
     medical care decisions,  
         1133–1134  
     pregnancy and, 2241  
     preventative practice, 1129  
     sick role concept, 1813  
     social comparison process, 2654  
     social constructions of,  
         1813–1814, 1815–1816  
     social norms and values in,  
         1133–1134  
     socioeconomic status and, 1129
- Health and Retirement Study, 344
- Health and the life course,  
     **1136–1139**  
     acute/chronic distinctions, 1137  
     aging and, 79, 81, 1137  
     alcohol use and, 93  
     conceptions of life course, 1137  
     concepts of health, 1136–1137  
     demographic transitions, 621–622  
     disease vs. illness, 1136–1137  
     end-of-life decisions, 585–587  
     functional model in, 1136  
     healthy life expectancy and, 1632  
     issues and implications,  
         1138–1139  
     life expectancy variations,  
         1137–1138  
     lifestyles and, 1639–1642  
     medical model of, 1136  
     medical model of long-term  
         care, 1665  
     mortality and, 1137–1138  
     quality of life and, 2301,  
         2305–2306
- Health Belief Model, 1127,  
     1128, 1814
- Health care financing. *See* Health  
     care utilization and expenditures;  
     Health policy analysis
- Health Care Financing  
     Administration (Medicare and  
     Medicaid), 588, 815, 1157,  
     1670, 1828
- Health care industry. *See* Medical-  
     industrial complex
- Health care providers  
     assisted-living companies, 1826  
     and family planning, 959  
     *See also* Managed-care  
         organizations
- Health care utilization and  
     expenditures, **1140–1156**  
     access and barriers to care,  
         1151–1152  
     by older people, 79, 1144  
     and capitalism, 241  
     drug abuse treatment, 714–715  
     drug costs, 1141, 1148  
     fee-for-service plans, 1143, 1144  
     growth in expenditures,  
         1140–1142
- health maintenance  
     organizations. *See* HMOs
- home health care, 1148, 1149
- hospital care, 1147–1148
- indemnity health plans, 1141
- long-term care funding,  
     1148–1149, 1658–1661, 1665
- minorities and, 1150, 1152
- point-of-service plans, 1143
- political economy perspective,  
     1153
- preferred provider  
     organizations, 1144
- private payers, 1142–1145
- provider-sponsored  
     organizations, 1144
- public payers, 1145–1147
- rate of growth, 1827
- resource allocation, 1147–1148
- sociological models of access to  
     service, 1152–1153
- Supplemental Security Income,  
     1145, 1146
- Temporary Assistance for Needy  
     Families program, 1145, 1146,  
     1288, 1395
- uninsured individuals, 1149–1151  
     *See also* Managed-care  
         organizations; Medicaid;  
         Medical-industrial complex;  
         Medicare
- Health care workers, 1818
- Health, Education, and Welfare  
     Department, U.S., 497
- Health industry. *See* Medical-  
     industrial complex
- Health insurance  
     fee-for-service plans, 1143, 1144  
     government policy, 283–284,  
         1827, 2798  
     *See also* Medicaid; Medicare  
     private/third-party, 1819–1820,  
         1825, 1826  
     uninsured population, 1828  
     *See also* Health care utilization  
         and expenditures; HMOs
- Health Insurance Portability and  
     Accountability Act of 1996,  
     1143, 1826
- Health maintenance organizations.  
     *See* HMOs

- Health of Regionville, The* (Koo), 1813
- Health policy analysis, **1157–1162**  
 comparative systems, 130, 373–381  
 definition of, 1157  
 engineering model, 1159, 1160  
 enlightenment model, 1159, 1160  
 limited sociological involvement, 1158  
 and medical sociology, 1158, 1813–1818  
 of medical-industrial complex, 1818–1829  
 politics and, 1159  
 privatization and, 1160  
 quality of life and, 2301  
 reform and, 1159  
 scope of topics under, 1159  
 social policy and, 1159
- Health promotion and health status, **1162–1173**  
 aging and, 1168–1169  
 assessment of, 1169–1171  
 barriers to, 1167–1168  
 diffusion of new practices, 677  
 disease control and, 1163  
 and environmental threats, 790–791  
 functional ability and, 1171  
 gender and, 1165, 1166  
 individual role in, 1166–1167  
 medical care systems and, 1163  
 minorities and, 1169  
 personal behavior and. *See* Lifestyles and health  
 physical environment and, 1163  
 political/economic institutions and, 1163  
 prenatal care and, 2236  
 primary prevention and, 1163–1164  
 quality of life indicators and, 1171  
 range of activities, 1163–1166  
 secondary prevention and, 1164–1165  
 social environment and, 1163  
 social networks and, 2733  
 tertiary prevention and, 1165
- Healthy life expectancy, definition of, 1632
- Heart attack, 139
- Heart disease, 93, 1640, 1641
- Heartburn, 2235
- Hebdige, Dick, 2170, 2171
- Hebrew Bible. *See* Bible
- Hechter, Michael, 1940, 3215
- Heck, Ronald, 2418
- Heckman, James J., 1684, 3038
- Heckman's estimator, 2441–2442
- Hecllo, Hugh, 603
- Hedbridge, Dick, 245
- Hedonic contingency thesis, 189
- Hegel, Georg W. F., 539, 774, 1066, 1781, 1782  
 and German idealism, 1248–1249
- Hegemonic stability theory, 3242
- Heidegger, Martin, 1783, 2756, 2757
- Heidenheimer, Arnold, 2124
- Heider, Fritz, 187, 335–336, 1034, 2702  
 attribution theory, 192, 194, 2244  
 and social psychology of status attainment, 2781
- Height measurement, 2343–2344
- Heimer, Carol A., 2876–2877
- Heirich, Max, 351–352
- Heise, David R., 409, 780, 1790, 2297, 2298
- Held, David, 606
- Helle, Jürgen, 1080
- Hellenistic Empire, 2998
- Hell's Angels, 460, 461, 1577, 1579
- Helms, Jesse, 276, 2585
- Helping behavior. *See* Altruism
- Helping professions, 1525
- Helsinki Accords (1975), 2047
- Helsinki University, 2450, 2451
- Heman's estimator, 2441–2442
- Hemlock Society, 585, 3083, 3084
- Hempel, Carl, 2464
- Hendin, Herbert, 3084–3085
- Henggeler, S. W., 76
- Hennion, Antoine, 1925
- Henry, Andrew F., 3080
- Henry, Louis, 620
- Henry, Neil W., 3038
- Henry, William E., 2300
- Henshel, Richard L., 2676, 2677
- Hepatitis, 93
- Heraud, Brian, 2844
- Herbal remedies, 654–655
- Heredity. *See* Genetics; Nature vs. nurture
- Heresies, 2968
- Heritage Foundation, The, 1601
- Hermeneutics, 543, 2219, 2472
- Hernes, Gudmund, 2452, 2453
- Hernes, Helga, 2453
- Herodotus, 852
- Heroin, 526, 711, 712, 714, 715
- Herrnstein, Richard, 214, 506, 1382, 2056, 2140, 2330
- Hertz, Robert, 11032
- Herz, J. H., 3130
- Herzberg, Frederick, 3271
- Herzer, Manfred, 2567
- Hesse, Mary, 823
- Heteroscedascity, 3036
- Heterosexuality, 2567–2569  
*See also* Sexual behavior in marriage and close relationships; Sexual behavior patterns
- Heuristics, 593–595
- Hewitt, John P., 3096
- Hewlett Foundation, 637
- Hewstone, Miles, 193
- Hickman, Lauren C., 1595, 1596
- Hicks, Alexander, 2166
- Hierarchical linear models, **1173–1178**  
 drug-abuse programs assessment, 71  
 estimation of, 1176–1177  
 levels of analysis and, 1594  
 longitudinal data in, 1177  
 model development, 1174  
 multilevel variables in, 1173  
 random effects in, 1176–1177  
 specification in, 1174–1176  
 statistical software for, 1177  
 statistical tests and, 1177  
 structural equation modeling and, 1922  
 variance-covariance components in, 1176

- High culture vs. mass culture debate, 565–566, 1645–1646
- High Islam, 2943
- High School and Beyond Survey, 761, 2480
- High-energy physics, 2470
- Higher education, **1178–1186**
- academic freedom, 276–278
  - academic libraries, 1604–1605, 1606–1607
  - academic organization, 1180–1182
  - affirmative action, 50, 51
  - African Americans and, 2497–2499
  - African studies, 66–67
  - authority structure, 1181
  - chair systems, 1181
  - communications departments, 678
  - contemporary pressures in, 1184–1185
  - demographers, 636
  - and equality of opportunity, 826
  - first sociology courses, 323–324
  - funding for, 1182
  - government connections, 1180
  - growth of, 1183–1184
  - historical development, 1179–1180
  - and historically black institutions, 2499
  - Humboldtian principles and, 1179–1180
  - interdisciplinary programs, 1181
  - land-grant universities, 1180
  - national systems of, 1182–1183
  - and occupational hierarchy, 2929–2930
  - personalized instruction, 215
  - and political correctness, 2142
  - professional training, 1179, 2259
  - reform movements, 1179–1180
  - research universities, 1179–1180
  - rural sociology departments, 2425–2427
  - sexual harassment and, 2591
  - social class and, 1179
  - and student movements, 3067–3070
  - tenure, 1182
  - theoretical knowledge and, 2196–2197
  - in United States, 766
  - “Wisconsin Idea,” 1180
  - See also* Education and mobility
- Highlander Research and Education Center, 2040
- High-reliability organizations (HROs), 2878
- Hightower, James, 88
- Hilbrand, Dieter, 3053
- Hilferding, Rudolf, 1265
- Hill, Anita, 2581
- Hillary, George, 363
- Hilton, Jeanne, 1391
- Hinduism, 2366, 2486, 3282, 3288–3289
- caste system, 250–253, 2811, 3284
  - Gandhi and, 3288–3289
  - new movements, 3287
  - religious experience, symbols, and ethical norms, 3280, 3281, 3284
  - widowhood and, 3255
- Hinin*, 253
- Hippies, 459–460
- Hippocrates, 2086
- Hippocratic oath, 3083
- Hiroshima bombing, 582
- Hirsch, F., 3173
- Hirschi, Travis, 505, 507, 535, 667, 1495, 2658
- Hirschman, Charles, 843
- Hispanic Americans, **1186–1195**
- AIDS/HIV risks, 2587, 2590
  - alcohol consumption rates, 94
  - average number of births per woman, 2032
  - case studies of, 244
  - census, 259
  - citizenship rate, 1193
  - demographic research, 636
  - demographic trends, 1190–1192
  - diversity of, 1186, 1187, 1188–1190
  - divorce rates, 126
  - economic profile, 1190–1192
  - education attainment, 1190–1192, 1193
  - ethnic resilience, 1192
  - and experiments, 887–890
  - family structure, 123
  - fertility rates, 1186
  - and filial responsibility, 1020
  - household income, 1279, 1280
  - household structure, 127
  - immigration and, 143, 1186, 1188, 1189
  - income per capita, 1280
  - labels and, 1187–1188
  - labor-force participation, 1191, 1193
  - language proficiency, 1193
  - life expectancy, 1169, 1631
  - marriage demographics, 124
  - modes of incorporation, 1194
  - political influence of, 1187
  - population growth, 1186
  - poverty rates, 1191–1192, 1287, 2215
  - racial discrimination against, 322
  - remarriage rates, 2388
  - segmented integration of, 1187
  - singlehood attitudes, 107
  - social integration of, 1187, 1188, 1190, 1192–1194
  - and suburbanization, 3074, 3075
  - and underemployment, 1721, 1722
  - voting patterns, 1193–1194
- Histograms, 659, 661, 3013–3014
- Historical analysis. *See* Comparative-historical analysis; Event history analysis; Historical sociology
- Historical counterfactuals, 386
- Historical hypotheticals, 386
- Historical institutionalism, 2164–2165
- Historical materialism, 543, 1217, 1704
- central tenet of, 1783
  - and cumulative social change, 2644, 2645, 2646, 2647
  - definition of, 1782
  - and French School of Sociology, 1026–1027
  - and German sociology, 1080



- and Marxist sociology, 1751–1752, 1758  
 as Marx's main metatheory, 1782  
 and materialism, 1781–1782, 1784  
 and progress concept, 2644, 2645  
 and Soviet sociology, 2980  
*Historical Materialism: A System of Sociology* (Bukharin), 1753  
 Historical sociology, **1195–1202**  
   comparative, 383–392, 2918–2919  
   cultural perspective, 569  
   and cumulative social change, 2643–2648  
   demographic data, 609, 621, 632  
   history/sociology differentiation, 1195–1196  
   on Islamic society, 2940, 2948  
   and life course, 1619, 2861–2862  
   Marxist, 1754, 1756  
   and national boundaries  
     development, 1932–1933, 1934–1935, 1939  
   organizational knowledge as  
     secondary data, 574  
   origins of democracy and, 605  
   postmodernism and, 2206  
   and social history, 2917–2919  
 Historically black colleges and universities, 2499  
 Historicism, 2953  
*History, Culture, and Region in Southeast Asian Perspectives 1999* (Wolters), 2978  
*History of Sociological Thought* (Szacki), 2120  
 Hitler, Adolf, 1310, 1357, 2332, 2356, 3002, 3233, 3242  
*Hitler's Willing Executioners* (Goldhagen), 387  
 HIV. *See* AIDS/HIV  
 HLM. *See* Hierarchical linear models  
 HMOs (health maintenance organizations), 1158, 1818, 1819, 1820, 1821  
   health utilization and  
     expenditures, 1142, 1143, 1144, 1145  
   mergers, 1822  
   physician resistance to, 1826  
   women physicians and, 2262  
 Ho, Hsiu-Zu, 194  
 Hobbes, Thomas, 527, 1029, 1302, 1576, 1781, 2456, 2882  
 Hobhouse, L. T., 1599, 1600  
 Hobson, John, 1265  
 Hochfeld, Julian, 2119  
 Hochschild, Arlie Russell, 244–245, 777, 781–782  
 Hodgson, Marshall G., 2943, 2946  
 Hoebel, E. Adamson  
   definition of law, 1544–1545  
   legal system study, 1549–1550  
 Hoem, Jan M., 2453  
 Hoffman, Alexander, 3229  
 Hofstadter, Richard, 1356, 1357  
 Hofstede, G., 3218  
 Hoggart, Richard, 226  
 Holden, Constance, 1226  
*Holiday of Darkness* (Endler), 651  
 Holiness movement, 2369  
 Holistic personality theory, 1718, 2084, 2085, 2088–2089  
 Hollander, Arie den, 1425  
 Hollander, E. P., 1567  
 Hollander, Edwin, 404  
 Hollandsworth, James G., 2305  
 Hollerith, Herman, 283, 406  
 Hollerith code (IBM cards), 575  
 Hollingshead, August B., 1813, 1834, 2927, 3055  
 Holm, Sverre, 2451  
 Holmes, Oliver Wendell, 476, 949  
 Holmes, T., 3055  
 Holocaust, 1066, 1067, 1070, 1384, 2206  
 Holstein, James, 855, 1636  
 Holter, Harriet, 2453  
 Holtzman test, 2077  
 Holy Roman Empire, 2998  
 Homans, George C., 208, 432, 722, 2335  
   on distributive justice, 2700, 2701, 2702, 2703  
   on group behavior, 2611  
   social exchange theory, 2670–2671, 2672, 3272  
 Home care services, 1656–1658, 1661–1662, 1663, 1677  
   as medical-industrial complex component, 1818  
 Home workers, 1339, 1341, 3267–3268  
 Homelessness, **1202–1208**  
   contemporary, 1203  
   deinstitutionalization and, 1841  
   estimates of, 1203–1204  
   historical condition of, 1202  
   illicit drug use and, 711  
   poverty and, 2215  
   research and remedies, 1202, 1203  
   urban renewal and, 1203  
 Homemaking, 3261  
 Homicide  
   as affective aggression, 68, 69, 72, 73  
   American Indian rates, 135  
   by juveniles, 1487–1488, 1489–1490, 1498  
   consistency in international definitions of, 499–500  
   and family violence, 981–982  
   international rates, 500  
   male jealousy and rivalry and, 2886  
   rate calculation, 491, 497–498  
   statistical map of U.S. rates, 3012  
   subcultural deviance theory on, 664  
   U.S. rates, 68, 3012  
   Uniform Crime Reports definition, 492  
 Homophobia, 2567, 2569–2570  
 Homoscedasticity, 449, 2251  
 Homosexuality. *See* Sexual orientation  
 Homuncularism, 2090–2091  
 Honduras  
   demographic characteristics, 1536  
   poverty, 2216  
   wartime rape, 2579  
 Hong Kong  
   dependency theory, 642  
   and family size, 977, 1008  
   labor movement, 1532  
   long-term care and care facilities, 1656  
 Honor, 2527–2528

- aggression and, 72, 2579  
 crimes of, 2566  
 slavery as denial of, 2596
- Hoogvelt, Ankie M. M., 1886–1887
- Hook, Sidney, 2218
- Hooligan: A History of Respectable Fears* (Pearson), 1579
- Hoover, Herbert, 2682
- Hoover, J. Edgar, 2494
- Hoover Institution, 1601
- Hopeless theory of depression, 651, 652
- Hopi, 138
- Horan, Patrick, 2027, 2481
- Horkheimer, Max, 539, 541, 542, 544, 1075, 1076, 1077, 1079, 1645
- Hormones, 70, 652
- Horney, Karen, 540, 1713–1714, 2084, 2089
- Horowitz, Donald, 847
- Horowitz, Ruth, 244
- Horticultural society. *See* Agrarian society
- Horton, Myles, 2040
- Hospices, 1671, 1818, 3064, 3086
- Hospital Corporations of America, 1821
- Hospitals  
 abortions in, 2240  
 ambulatory care centers, 1821  
 comparative systems, 375  
 death and dying in, 582–583, 585, 586, 587  
 Medicaid and Medicare and, 1825  
 as medical-industrial complex component, 1818, 1819  
 mergers, 1821–1822  
 multihospital systems, 1819–1820  
 ownership and expenditure rates, 1819, 1820, 1821–1822  
 ownership shifts, 1820–1821  
 utilization and expenditures, 1147–1148
- Hostility biases, 71–72
- Hotte, Alan M., 3215
- House, James S., 2070
- House, R. J., 1566, 1568–1569
- Household structure. *See* Family and household structure
- Houseknecht, Sharon K., 1942
- Housing  
 African-American discrimination, 56, 57, 143, 250, 845  
 African-American segregation indices, 2500–2504  
 and family policy in Western societies, 965  
 federal projects/move to suburbs correlation, 3072  
 as fertility transition factor, 625  
 maintenance/deviance control theory, 666  
*See also* Homelessness
- Housing and Urban Development Department, U.S., 838, 2128
- Housing for Senior Citizens, 1656
- Howard, Judith, 198, 1648
- HPA axis (hypothalamic-pituitary-adrenal), 652
- Hrdy, S. B., 2886
- HRO theory, 2878
- HRS. *See* Health and Retirement Study
- Hsiung, James, 332
- HTML (Hypertext Markup Language), 408
- Huang, Jie, 644
- Hubbard, L. Ron, 2366, 3287
- Hubert, H., 1024
- HUD. *See* Housing and Urban Development Department, U.S.
- Hudson, Walter, 2943
- Hudson's Bay Company, 3174
- Huesmann, L. R., 70
- Hughes, Everett C., 244, 1813, 2415, 2417
- Hughes, Helen MacGill, 243, 244
- Hughes, Langston, 66
- Hughes, Michael, 2300, 2307
- Hulbert, Jeanne, 2792
- Hull, Raymond, 234
- Hull House (Chicago), 365–366
- Hull House Maps and Papers* (1895), 366
- Human agency, 1620
- Human awareness. *See* Phenomenology
- Human capital theory, 513, 747  
 Asian Americans and, 181
- cultural capital vs., 2928  
 and educational investment, 2927  
 and social capital, 2637  
 and social stratification, 2812, 3043  
 as stratifying force, 2812  
 wage trajectories and, 1989–1990, 1991
- Human Development Report* (UN), 2917
- Human ecology and environmental analysis, 800, **1209–1233**, 2922  
 in American cities, 308  
 Burgess hypothesis, 1209  
 carrying capacity and, 1219  
 centripheity process, 1212–1213  
 change and, 1211  
 classical, 1209–1210  
 critical theory on, 544  
 and crowds and riots, 557  
 definition of environment, 1228  
 and deviance theory, 2658  
 dominance and succession in, 1209  
 ecological demography and, 1213  
 ecosystem and, 1210  
 and environmental degradation, 1214, 1216  
 environmental sociology and, 1214  
 evaluation of, 1213–1214  
 expansionist perspective and, 1217–1218  
 factors as sociological influence, 2921–2922  
 functionalist structuralist approach to, 1029, 1031  
 and futures studies, 1041  
 and global problems, 801, 805, 810–811  
 neoclassical, 1210–1212  
 neo-Malthusian perspective, 1219–1220, 1228  
 and new urban sociology, 1215–1216  
 POET variables, 1210, 1214  
 and political correctness, 2140–2141  
 political economy approach, 1214–1215, 1221–1222

- population and, 1216–1222  
 proactive environmental sociology and, 1226–1228  
 pro-growth perspective, 1217–1218  
 social chaos and, 1223–1224  
 social structure and, 2822  
 socioeconomic differences between cities and suburbs and, 3072–3073  
 sustainable development and, 1222–1223, 1225  
 and territorial belonging, 3129  
 and transnational problems, 1220  
*See also* Environmental sociology; Rural sociology
- Human emancipation, 539
- Human exemptionalism paradigm, 806
- Human Genome Project, 1824
- Human immunodeficiency virus. *See* AIDS/HIV
- Human involvement. *See* Social belonging
- Human Meaning of Social Change, The* (Campbell and Converse), 2683
- Human nature, **1233–1236**, 1301  
 altruism and, 114–120  
 conservative vs. liberal view of, 357, 1598–1599  
 constitutional personality theory and, 1717–1718  
 dualistic nature of, 1233  
 optimistic vs. pessimistic perspectives, 1233  
 plastic theory of, 1233, 1234  
 positivist view of, 528  
 pragmatist view of, 2217–2218  
 religious views of, 2086–2087  
 responsive communitarianism view of, 357–358  
 self-actualization theory, 2087–2088  
 selfishness and, 2882  
 sociobiological theory of, 2881–2882  
*See also* Compliance and conformity; Deviance theories; Personality theories
- Human potential movement, 2717
- Human Relations* (journal), 2942
- Human Relations Area Files, 548, 2893
- Human Resources Research Organization, 1876
- Human rights, children's rights, and democracy, **1236–1246**  
 British sociological study, 228  
 children's rights, 1239–1240, 1242–1244  
 communitarian view of, 361  
 democracy and individual rights, 1237–1238  
 incest as violation of, 1274–1275  
 international accords, 2047  
 and international campaign against human bondage, 2607  
 international law, 1429–1430  
 liberalism's emphasis on, 355  
 nationalist movements' abuses of, 1944, 1948  
 political crime and, 2146  
 relevance of sociology to, 1237–1238  
 religious movement leadership and, 2374  
 Universal Declaration of Human Rights, 1240–1242  
 wartime rape and, 2579
- Human Side of Enterprise, The* (McGregor), 1014
- Human Societies* (Davis), 1030
- Human sociobiology. *See* Sociobiology, human
- Humana, 1821, 1822
- Human-Animal Bond Center (HABIC), 3230
- Humanism, **1246–1251**  
 Enlightenment and, 1247–1248  
 flexibility and, 1250  
 German idealism and, 1248–1249  
 pragmatism and, 1249–1250  
 and writing case studies, 248
- Humanistic coefficient (Znaniecki concept), 2118
- Hume, David, 818, 1247, 1302, 1684, 2335, 2338, 2943, 2946, 3098
- Hummel, Hans-Joachim, 1080
- Hummel, Raymond C., 1901
- Humor  
 persuasion and, 2095
- political correctness and, 2140
- Humphrey, Hubert, 1597
- Humphreys, Laud, 839
- Humphries, Drew, 2961
- Humphry, Derek, 585, 3083, 3084
- Hungary  
 daily time use, 3160  
 divorce rate, 706  
 high suicide rate, 3079, 3082  
 legal system, 474  
 post-communist transition, 2136  
 protest movement, 2267  
 revolution, 2414, 3001  
 Social Science Data Archive, 576  
 social surveys, 577  
 socialist economic modifications, 2850–2851  
 sociology in, 2116, 2117  
 tourism in, 3169  
 and World War I, 2362
- Hunt, J. G., 1571
- Hunt, Sonia, 2306
- Hunter, Floyd, 2624
- Hunting and gathering society, 1068, 2810
- Huntington, Ellsworth, 2921
- Huntington, Samuel P., 1936, 1940, 2124, 2159, 2363  
 “clash of civilizations” hypothesis, 2940–2941
- Huron, 1070
- Hurston, Zora Neale, 66, 1648, 2890
- Husserl, Edmund, 1080, 2099, 2756
- Hutcheson, Francis, 3098
- Hutchinson, Anne, 2146
- Hutchinson, Ray, 1214
- Hutu, 254
- Huxley, Aldous, 1505
- Huxley, J., 2880
- Hybrid corn innovation, 87, 677, 2429
- Hyman, Herbert H., 2479, 3092
- Hymes, Dell, 2894
- Hyperactivity, 1816
- Hyperemesis, 2235
- HyperResearch* (computer software), 419, 420
- Hypertension, 93
- Hypertext, 408

- Hypnosis, 898–899
- Hypodermic needle model of diffusion, 679
- Hypodescent rule, 2331–2332
- Hypothesis testing, 3023–3024  
*See also* Scientific explanation; Statistical inference
- Hypothetical cohort, 614
- I**
- Iadov, Vladimir, 2980, 2981
- Iannaccone, Laurence, 939, 2375, 2485
- Ianni, Francis A. J., 2018
- Ibarar, Peter R., 2763
- IBM cards (Hollerith code), 575
- Ibn Khaldun, Abd-al-Rahman, 327, 1564, 2941–2942, 2943, 2945, 2946
- Ibo, 54, 1384
- Icaria (Illinois utopian community), 2849
- Iceland  
 long-term care and care facilities, 1654, 1655, 1661  
 sociology in, 2451, 2452, 2453
- ICPSR (Inter-University Consortium for Political and Social Research, Ann Arbor, Michigan), 409, 575, 576, 577–578, 579, 1606  
 resources and functions, 2474–2475, 2476–2477, 2481
- ICSS (International Committee for Sport Sociology), 2986–2987
- Id, 1713
- Ideal type formulation, 3181
- Idealism  
 German, 1248–1249  
 materialism vs., 1780  
 and pragmatism, 2218, 2219  
 and symbolic interactionism, 2856
- Idealization, 1699
- Ideation. *See* Sociology of knowledge
- Idée de droit social, L'* (Gurvitch), 1026
- Identity. *See* Identity theory; Individualism; Self-concept; Social identity
- Identity politics, 1580, 1784
- Identity theory, **1253–1258**  
 adolescence and, 4  
 adulthood and, 29, 32  
 affect and emotion in, 1257  
 attitudes and, 44  
 cognitive focus of, 1257  
 and collectivity, 2632–2633  
 commitment in, 1255, 1256  
 definition of, 2784  
 educational aspirations and, 2784  
 ethnicity and, 1939–1940  
 gender and, 1001, 1061, 1256  
 groups and, 45  
 impression formation and, 42–44, 2776  
 individuals and, 100, 2221, 2856  
 Japanese Americans and, 180–181  
 nationalism and, 1939–1940  
 parental role and, 2034  
 personality theory and, 2083  
 Polish sociology and, 2119  
 possibility of choice in, 1253  
 role theory and, 2423–2425  
 sexual orientation and, 2567, 2570–2571  
 social belonging and, 2633–2634  
 social perception and, 2750  
 socialization and, 2856, 2861  
 sociolinguistics and, 2906–2908  
 symbolic interactionism and, 2423, 3095–3100
- Ideology  
 and functionalism, 1031  
 and genocide, 1067, 1069–70  
 and Iranian revolution, 1871  
 Mannheim's theory of, 2955–2956, 2957, 2958  
 Marxist sociological research on, 1755, 2955, 2956  
 and nationalism, 1940, 1944  
 political correctness and, 2138–2142  
 and protest movements, 2267  
 and revolutions, 2413, 3001  
 as social capital influence, 2639  
 and sociology of knowledge, 2955–2956, 2958  
*Ideology and Utopia* (Mannheim), 2955–2956, 2959
- “Ideology as a Cultural System” (Geertz), 2958
- Idiosyncrasy credits, 404
- IFDO (International Federation of Data Organizations), 577
- IIS. *See* International Institute of Sociology
- I-It and I-Thou relations (Buber concept), 355
- Ikegami, Naoki, 380
- Iker, Howard, 420
- Illegal migrants, 1436, 1936
- Illegitimacy, **1258–1264**  
 consequences of, 1260–1261  
 economic issues of, 1260  
 public policy and, 1261  
 rise in, 1259  
 tolerance for, 1259, 1260  
 trends in, 1259–1260
- Illiberal Education* (D'Souza), 2140
- Illicit drug use. *See* Drug abuse
- Illinois State Penitentiary, 325
- Illiteracy. *See* Literacy
- Illness. *See* Health and illness  
 behavior; Health and the life course; Health promotion and health status; Medical sociology; Medical-industrial complex; *specific conditions*
- Illusions, and positive mental health, 2190
- ILO. *See* International Labour Organization
- Image analysis (Guttman concept), 906
- Image of the Future, The* (Polak), 1038
- Imagined Communities* (Anderson), 2978
- Imbalanced states, 335
- IMF. *See* International Monetary Fund
- Imitation  
 as crowd behavior mechanism, 679  
 and role theory, 2415  
 and social learning theory, 2768–2769
- Immanent critique (Frankfurt School concept), 539–540, 542, 543
- Immigration. *See* International migration

- Immigration Act of 1924, 174
- Immigration Act of 1965, 143, 175, 177
- Immigration Act of 1990, 143
- Immigration Reform and Control Act of 1986, 143, 1189, 1436–1437
- “Immorality of Being Softhearted” (Hardin), 1220
- Immune system, 656
- Immunization. *See* Vaccines
- Impeachment, 1239, 2128, 2136  
public opinion polling on, 2274, 2276, 2278
- Imperialism: A Study* (Hobson), 1265
- Imperialism, colonialism, and decolonization, **1264–1266**  
in Africa, 60–61, 63, 64, 66, 1550, 2132, 2602  
and antiblack ideology, 55  
causes of decolonization, 1267–1268  
as class and race factor, 320–322  
consequences of decolonization, 1268–1269  
decolonization, 1265–1269  
definitions of, 1264, 1266  
dependency theory and, 640–646  
ecology and, 1214  
economic analysis of, 1264  
effects of twentieth-century wars on, 2362  
eras of, 1266–1267  
European political contraction and, 1266–1267  
European political expansion and, 1264–1265  
exploitation and, 1214, 1221  
genocide and, 1069, 1071  
global political discourse and, 1268  
historical overview, 1264–1270  
imperialism theories, 1265–1266  
indigenous movements and, 1267  
and Latin American studies, 1536–1537, 1538  
legal systems layering (transplant) and, 465, 1550, 1556  
Marxist-Leninist view of, 3243  
mass media and, 1767  
metropolitan vs. indigenous traditions, 1265  
national borders and, 1933–1934  
nationalism and, 1944–1945, 3001  
power-conflict perspective on, 53, 54, 55  
and resulting corruption, 2132, 3234  
revolutions and, 1267, 3001  
and slavery in the Americas, 2598–2601  
social Darwinism and, 2330  
in Southeast Asia, 2974–2975  
systemic factors, 1268  
violence and, 1267
- Imperialism: The Highest Stage of Capitalism* (Lenin), 1265
- Impression formation and management. *See* Affect control theory and impression formation; Social perception
- Improper coalition, 332
- Improper linear models, 592–593
- In a Different Voice* (Gilligan), 993–994
- Inca Empire, 2999
- Incarceration. *See* Penology
- Incentive-based aggression, 73
- Incest, **1270–1278**, 2582–2583  
affine and cousin marriages and, 1272, 1273, 1509, 1513  
as child abuse, 1270, 1273, 1275–1276  
feminist theory on, 991, 1275–1276  
incidence of, 2557, 2559  
and mate selection taboos, 1270–1274, 1509, 1513, 1776  
as prohibited coitus, 1270  
stepfamilies and, 1272, 2583
- Incest: The Nature and Origin of the Taboo* (Durkheim), 1032, 1272
- Inclusion, 2633
- Income distribution in the United States, **1278–1290**  
Asian Americans and, 181, 182  
comparable worth and, 369–373  
definition of, 1279  
democracy and, 606  
discrimination in, 689–690  
education and, 1279–1280  
employee benefits and, 1282  
and equality of opportunity, 826–827  
household structure and, 127, 1279  
inequalities in, 130, 140, 142, 689–690, 1281–1282, 2691, 3048  
labor-market factors, 1283  
legal profession and, 470  
lifestyle factors, 1283  
per capita, 1279, 1281  
and poverty demographics, 1285–1288, 2214–2215  
size distribution changes, 2869  
skilled vs. unskilled worker gap, 1281  
social justice and, 2705–2706  
social security systems and, 2802  
status attainment and, 3047–3048  
Survey of Income and Program Participation, 1284–1285, 1288, 1722, 2475  
taxes and transfers, 1282–1283  
and theories of crimes, 506, 536  
underemployment and, 1721–1722  
War on Poverty and, 1286, 1494, 1882, 2404, 2760  
wealth and, 1283–1285  
women’s earnings gap, 370, 984  
*See also* Poverty; Wages and salaries
- Income generation. *See* Wages and salaries; Work and occupations
- Incommensurability thesis, 2025, 2026
- Inconsistency, tension state from, 334, 337–340
- Indentured servants, 2597, 2598, 2599, 2600, 2601
- Independence Day, 584
- Independent contractors, 3268
- Independent counsels, 145, 2127, 2128
- Independent practice associations. *See* IPAs

- Independent Public Opinion  
Research Service Vox Populi  
(Russia), 2982
- Indeterminate sentencing, 2254,  
2255
- Index, Durkheim as concept  
formulator, 1553
- Index crime rate, 503, 1486–1487
- Index of Forbidden Books, 267
- Index of General Affect, 2303–2304
- Index of Social Progress, 2302
- Indexing services, sociological  
periodicals, 1608–1609
- India  
anticolonial revolution, 3001  
Bhopal disaster, 683, 805, 2875,  
2877–2878  
and British Empire's  
dissolution, 2362  
caste and inherited status,  
250–253, 2811, 3284  
common law system, 465  
divinity names, 3280  
fertility rate decline, 220  
genocide, 1070  
Green Revolution, 90  
historical empire, 299, 2998  
and Kashmir nationalism, 1945  
moral development study, 1901  
multilingualism, 2909  
Muslim minority in, 2950  
nonviolent protest, 2269  
Pakistan border fighting, 2362  
political and governmental  
corruption, 2132  
poverty in, 2216  
religious coexistence, 3288  
religious movements, 2366,  
2485–2486, 3281, 3284  
secularization, 2486  
sexually transmitted diseases  
in, 2593  
slavery and slave-like practices,  
2602–2603, 2604–2606, 2607  
time use research, 3161, 3162  
widowhood, 3252  
*See also* Indian sociology
- Indian Americans, 175, 180
- Indian Council of Social Science  
Research, 1292
- Indian Self-Determination and  
Educational Assistance Act of  
1975, 136
- Indian Sociological Association, 1291
- Indian Sociological Society, 1293
- Indian sociology, **1290–1295**  
areas of study, 1290, 1291, 1293  
dependency theory and, 641  
development trends, 1291  
leader behavior study, 1566  
relevance issue, 1292, 1293  
university offerings, 1290,  
1291, 1294  
Western influence on, 1291–1292
- Indians, American. *See* American  
Indian studies
- Indicator  
definition of, 1907  
random error in individual, 1909  
single, 1910–1912  
standardized, 2991–2996  
*See also* Multiple indicator models
- Indigenismo* movement (Mexico),  
1858, 1861
- Indigenous peoples, **1295–1301**  
alcohol consumption, 95, 135  
contemporary issues, 1299  
decolonization, 1267–1268  
definition of, 1295–1296  
demographic issues, 1297–1298  
economic development, 1298  
genocide and, 1069  
in Mexico, 1857, 1858,  
1861, 2271  
and new religious  
movements, 2367  
New World empires, 299  
political activism by, 1298–1299  
population estimates, 1297–1298  
self-determination issue, 1299  
sociological significance of,  
1296–1297  
Spanish enslavement of, 2600  
terminology of, 1295–1296  
tourism and, 3171–3172  
*See also* American Indian studies;  
Imperialism, colonialism, and  
decolonization; *specific groups*
- Indirect demographic estimation,  
617–618
- Indirect discrimination, 689–691,  
692
- Individual behavior  
collective action as facilitator, 349  
role theory and, 2423–2425  
*See also* Individualism
- Individual differences  
aggression and, 71–72  
in American society, 142  
in career mobility, 1991–1993  
decision processing, 595  
disaster behavior, 683, 684  
innovation adoption, 87  
intelligence, 1369–1372  
life course transitions, 1626  
micro-level deviance theories and,  
662, 663, 666–670
- Individual psychology, 1714
- Individual rights, 1303, 1598, 1602  
democracy and, 1238–1239  
*See also* Civil liberties; Civil rights
- Individualism, **1301–1308**  
affective, 1303  
alternative life styles, 113  
art and, 1302  
community and, 359  
conformity vs., 400  
as cultural construct, 3218  
current issues, 1306–1307  
divorce rate linked with, 112, 113  
existential personality theory  
and, 1717  
identity and, 2221  
liberalism's focus on, 355  
methodological, 1234, 1306  
Pandectist doctrine, 476  
and persuasion receptivity, 2098  
political-economic organization  
and, 1303  
racial prejudice and, 2245  
responsive communitarian  
view of, 357  
rights and, 1303  
status incongruence and,  
3052–3053  
values and, 3218–3219, 3221  
*See also* Identity theory; Individual  
differences; Self-concept

- Individualistic fallacy, 3052
- Individualizing interpretive analysis, 387
- Indochina. *See* Vietnam
- Indochinese Americans, 180, 181, 182
- Indonesia, 2362, 2974, 2975, 2978
- agricultural innovation, 2976–2977
  - anticolonial revolution, 3001
  - and British Empire's dissolution, 2362
  - coalition triad in 1998 uprising, 333
  - divorce law, 938
  - Dutch conquest of, 2975
  - fertility rate decline, 220, 2178, 2179, 2976
  - genocide, 1070, 1071
  - Islamic society study, 2944
  - labor movement, 1532
  - legal system, 479
  - multilingualism, 2909
  - political and governmental corruption, 2124, 2130–2131
  - protest movements, 2271, 3068
  - slavery and slave-like practices, 2605, 2606–2607
  - sociodemographic profile, 2938
  - women in labor force percentage, 3262
- Inductive logic, 2466
- Industrial capitalism. *See* Capitalism
- Industrial ecology, 811
- Industrial feudalism, 2118
- Industrial parks, 3074
- Industrial prisons, 518
- Industrial relations, 1310
- Industrial Revolution, 818, 1219, 1316, 1346, 2759
- agricultural innovation and, 90
  - alienation linked with, 100
  - American family effects of, 122
  - anomie and, 164
  - criminal and delinquent subcultures and, 511, 531
  - demographic transition, 633
  - effect on Southeast Asian economy, 2975
  - and growth of cities, 306
  - historical sociology on, 1197
  - labor movement and, 1528, 1533
  - and modern labor force, 3262
  - and political party origins, 2154
  - rural sociology and, 2426
  - and social change, 2644, 2646
  - and work-family separation, 3266
  - See also* Industrialization
- Industrial Revolutionary Party (Mexico). *See* PRI
- Industrial sociology, **1308–1316**
- business cycles, 1310
  - and community transformation, 366–368, 664, 665, 698, 2658
  - contemporary, 1312–1315
  - convergence theory, 422–423, 424–425, 428
  - division of labor, 696–697
  - Durkheim's anomie theory, 698
  - fertility rate factors, 624
  - German sociology and, 1076, 1081
  - global economy and, 1310
  - and historical sociology, 1197
  - Japanese sociology and, 1480
  - on social networks, 2728
  - and work orientation, 3269–3276
  - See also* Industrialization; Labor movements and unions
- Industrialism and Industrial Man* (Kerr et al.), 422–423
- Industrialization
- adolescence and, 5
  - agricultural innovation and, 90, 2432
  - alienation and, 99
  - of American society, 141, 143
  - anomie and, 164, 698
  - and capital, 2637
  - characteristics of, 1884, 2196
  - correlates of, 1318–1319
  - definitions of, 1316–1318, 1884, 2196
  - “disorganization” of community from, 366–368, 664, 665, 698, 2658
  - distributive effects of, 1309
  - division of labor consequences, 697–698, 699
  - ecology and, 1209
  - family life and, 122
  - and fear of machinery, 2526–2627
  - and fertility transitions, 1007, 2176
  - historical sociology on, 1197, 2297–98
  - individual brands of, 1313
  - industrial sociology and, 1308–1316
  - inheritance and, 1394
  - Japanese sociology and, 1480
  - macrosociological studies of, 1706–1707
  - material-processing systems and, 1346
  - measures of, 1318
  - and modernization theory, 1883, 1884, 1885
  - paid work evolving from, 3262
  - and population growth, 2177
  - postindustrial differentiation, 2195–2203
  - rural sociology and, 2425, 2426, 2427
  - and social change, 2643, 2644, 2646
  - and social inequality, 2690
  - and social problems, 2759–2764
  - stratification parameters, 2810, 2811–2812, 2816
  - and work and family separation, 3266
  - and work orientation, 3269–3276
  - See also* Marxist sociology; Technology and society
- Industrialization in less developed countries, **1316–1324**
- and agrarian transition, 1317
  - child labor and, 3262
  - communications technology and, 1321–1322
  - convergence theory, 422–423, 425, 427
  - correlates of, 1318–1319
  - demographic transition, 633
  - dependency theory, 639–646, 1706
  - development goals, 1319–1321

INDEX

- economic activity and, 1317  
 education and, 741–754, 760, 762  
 and environmentalism, 932,  
 1219–1220  
 family policy and, 928–933, 2179  
 feminist perspectives on, 1708  
 human development and, 1318  
 indirect demographic estimation,  
 617–618  
 information gap and, 1322  
 infrastructure and, 1318  
 Latin American studies,  
 1537–1538, 1539–1540  
 Mexican studies, 1857–1858  
 Middle Eastern studies,  
 1867–1868  
 model life tables, 618–619  
 modernization theory and, 1706,  
 1883, 1884, 1886–1887  
 mortality levels, 622  
 population growth and, 2179  
 and postindustrialism, 2202  
 poverty and, 2215–2216  
 radical case studies, 246  
 redistribution of manufacturing  
 jobs, 3267  
 rural sociology and, 2429–2430  
 and social change, 2646  
 Southeast Asian socioeconomic  
 development, 2975–2976  
 traditional manufacturing and,  
 1317  
 transnational corporations and,  
 3178–3179  
 work and occupations, 3262,  
 3267  
 World Bank survey, 549  
*See also* Developing countries;  
*specific countries and regions*
- Ineffability, 3279  
 Inefficacy, 101, 102, 103, 104  
 Inequality. *See* Discrimination;  
 Segregation and desegregation;  
 Social inequality; Social justice  
*Inequality and Heterogeneity* (Blau),  
 698–699  
 Infant and child mortality, 221–222,  
**1324–1337**  
 African trends, 1332–1333  
 from AIDS, 2593  
 American families, 122, 130, 140  
 Asian trends, 1331–1332  
 Children's Bureau community  
 studies, 366  
 comparative health-care  
 systems, 377  
 contemporary trends, 1327–1328  
 current, 1138  
 data sources, 1324–1325  
 declines in, 1327, 1328,  
 1330, 2177  
 declines/life expectancy  
 relationship, 224  
 in developing countries, 744,  
 1330–1333  
 differentials in, 1333–1335  
 Eastern European trends,  
 1328–1330  
 epidemiologic transition and,  
 1325–1326  
 and fertility determinants, 1009  
 and fertility rates, 219, 1326  
 genetically transmitted defects  
 and, 1273  
 graphic representations, 3013,  
 3014, 3015  
 health care utilization and, 1151  
 health status and, 1170  
 historical view, 1326–1327  
 influences on, 1333, 1334  
 Latin American trends,  
 1330–1331  
 measures of, 1324–1325  
 minority, 1152  
 mortality modeling, 619  
 mortality transition, 621,  
 624–625, 633, 2177  
 neonatal, 1325, 1326, 2236  
 patterns of, 1324  
 perinatal, 1325, 2236  
 postneonatal, 1325, 1326  
 rate calculation, 221–222  
 rates, 1324–1325  
 sanitary revolution and, 1327  
 socioeconomic indicators of,  
 1334  
 total fertility rate and, 219  
 U.S. ranking, 1827
- Infantile sexuality theory (Freudian),  
 1274, 1275, 2090  
 Inference. *See* Statistical inference  
 Infidelity. *See* Adultery;  
 Extramarital/extra dyadic sex  
 Influence peddling, 2127, 2128  
 Informal economy, **1337–1344**  
 conceptual development of,  
 1337–1340  
 definition of, 1337–1338  
 forms of activity in, 1339–1340  
 measurement of, 1341–1343  
 Informal knowledge, 2959  
*Information Age, The* (Castell), 1758  
 Information flow  
 models of, 679  
 as social capital component, 2638  
 Information society, **1344–1348**  
 laser technology in, 1345  
 laws governing, 1348, 1349–1350  
 library resources, 1605–1613  
 Marxist sociology and, 1758  
 and “New Class” of knowledge  
 workers, 2626  
 photonics in, 1345  
 postindustrialism and, 2195,  
 2196–2197, 2205  
 and social change, 2645  
 technology in, 1344  
*See also* Internet  
*Information Society, The* (journal), 414  
 Infrapolitics (surreptitious  
 resistance), 254  
 Inglehart, Ronald, 3222–3224  
 INGOs (international  
 nongovernmental organizations),  
 1046, 2050  
 Ingroup/outgroup categorization,  
 2244  
 Inhalants, 712  
 Inheritance, **1348–1355**  
 caste and status system,  
 249–255, 2811  
 and class mobility, 757  
 contemporary issues, 1352–1354  
 equity and, 1353  
 family structure and, 926, 1349,  
 1350, 1353  
 and filial responsibility, 1018,  
 1508–1509



- generational distribution  
 patterns, 1351–1352
- impartible, 1348–1349, 1350
- industrialization and, 1394
- inter vivos, 1349, 1353
- intestate succession in, 1348
- kinship systems descent rules,  
 1507–1509, 1510, 1513, 1516
- modern, 1350
- partible, 1349, 1350
- patterns, 1349–1350
- preindustrial retirement as, 2402
- primogeniture system, 1350–1351
- soundness of mind and, 1351
- of status, 250, 251, 1239–1240,  
 2810, 2811, 2869
- testamentary freedom and,  
 1350–1351
- See also* Intergenerational  
 resource transfers
- Inheritance (biological). *See* Genetics;  
 Nature vs. nurture
- Inhibiting effect, 2615
- Initial Review Group, 2399, 2401
- Initiating-structure leadership, 1566
- Inkeles, Alex, 1085, 1885, 2069
- on convergence and divergence,  
 423–424, 425, 426
- on relationship between social  
 psychology and sociology,  
 2921
- In-laws. *See* Affines
- Inner Circle, The* (Useem), 736
- Inner-circle thesis, 2162
- Inner-city communities. *See* Urban  
 underclass
- Innis, Harold, 1765
- Innovation
- leadership management and,  
 1567
- legal system and, 1560
- mass media spread of, 1763
- Merton's anomie theory and,  
 165–166, 1494
- popular culture and, 2169–2170
- in postindustrial society,  
 2196–2197
- scientific research and  
 development, 2460–2461
- self-concepts in resisting, 2509
- and social change, 2643, 2644
- See also* Agricultural innovation;  
 Diffusion theories
- Inquisition, 268
- Inquisitorial vs. accusatory legal  
 model, 479
- Institut für Sozialforschung,  
 1075, 1076
- Institute for Advanced Study  
 (Princeton), 325, 1012
- Institute for Scientific  
 Information, 1610
- Institute for Social Research  
 (Copenhagen), 2452
- Institute for Social Research  
 (Frankfurt), 539
- See also* Frankfurt School
- Institute for Social Research  
 (Norway), 2451, 2452
- Institute for Social Research  
 (Sweden), 2451, 2452
- Institute for Social Research  
 (University of Michigan), 578,  
 2297, 2300
- Institute for Sociological Research  
 (USSR), 2981
- Institute of Concrete Social Research  
 (USSR), 2980, 2981
- Institute of Human Relations  
 (Yale), 325
- Institute of Management  
 Sciences, 3104
- Institute of Medicine, 588
- Institute of Muslim Minority Affairs,  
 2950–2951
- Institute of Social and Political  
 Problems (Russia), 2982
- Institute of Sociology (Ukraine),  
 2983
- Institute of Sociology (USSR),  
 2981, 2982
- Institute of Statistical Mathematics  
 (Tokyo), 1479, 1482
- Institutional dominance, 941
- Institutional environments, 394–395
- nursing home residents,  
 1672–1677
- Institutional family, 1502
- Institutional review board, 838, 853
- Institutional theories
- of anomie, 166, 503
- of political party systems,  
 2153–2254, 2164–2165
- and rational choice analysis,  
 2340, 2342
- Institutional trust, 2525
- Instituto Superiore di Sociologia  
 (Milan), 576
- Institutt for Samfunnforskning,  
 2451, 2452
- Instrumental aggression, 69
- Instrumental rationality, 541,  
 542–543, 544
- Instrumental values, 3214
- Instrumentalism
- language theory, 2219
- and Marxist sociology, 2162
- Insurance
- life, 3257
- long-term care, 1659–1660, 1677
- See also* Health insurance
- Integration, 79
- and academic achievement,  
 2932–2933
- desegregation vs., 2495–2496
- failure of, 2496–2499
- See also* African American studies;  
 Segregation and desegregation
- Intellectuals, **1355–1358**
- critical theory on, 541
- definition of, 1356
- political correctness and,  
 2139–2142
- Intellectuals on the Road to Class  
 Power, The* (Konrad and  
 Dzelenyi), 2117
- Intelligence, **1359–1386**
- artificial, 410, 1234
- chronometric approach, 1360–  
 1361
- communication and, 1380–1381
- consequences of differences in,  
 1373–1385
- crystallized, 1363
- daily life and, 1377–1379
- definitions of, 1361–1362
- education and, 1375
- as educational and occupational  
 attainment factor, 2784, 2817,  
 2931, 3043

- elementary cognitive tasking in, 1361, 1371  
 emotional, 1369  
 employment and, 1377  
 environmental and, 1369–1373, 2330  
 experimental strategy, 1359, 1360–1361  
 and family size, 972, 973–974  
 fundamental nature of, 1361  
 group differences, 1372–1373  
 heritability controversy, 1369–1373, 2090, 2140, 2330  
 individual differences in, 1369–1372  
 interpersonal context, 1380–1382  
 meaning of, 1361–1369  
 measurement of, 1359–1361  
 metacognitive skills and, 1367  
 multiplicity theories and, 1368–1369  
 psychometric testing, 1359–1360, 1362–1368, 1375, 1376, 2330  
 racial differences argument, 2330  
 self-esteem and, 2516  
 social pathology and, 1379–1380  
 specificity theories and, 1368  
 trainability and, 1363–1364, 1375  
 Intentionality. *See* Rational choice theory  
 Interaction chronography, 1974  
 Interaction Process Analysis, 1974, 1975, 1976, 1977, 1978, 1979, 1980  
     for small groups, 2613–2614  
*Interaction Process Analysis* (Bales), 2666  
 Interaction Process Scores (IPS), 1974–1975  
 Interaction theories. *See* Social exchange theory; Social interaction; Symbolic interaction theory  
 Intercorrelations, 2345–2346, 2348, 2354  
     *See also* Factor analysis  
 Interest groups, 144, 145, 604, 2148, 2162  
 Interfaith dialogue, 3289  
 Interfaith marriage, 911, 1411, 1776  
 Intergenerational relations, **1386–1393**  
     age-stratification theory, 1387  
     altruism and, 2882  
     changing dependencies in, 1391–1392  
     family focus on, 1388–1392  
     filial responsibility and, 1018–1022  
     gender and, 1389, 1391  
     generation gap and, 1386  
     generational equity and, 1388, 1398  
     grandparenthood, 131, 696, 1390–1391  
     interactions with nursing home residents, 1676  
     justice issues, 2707  
     long-term caregiving, 1653, 1657–1658  
     macrosociological perspectives, 1386–1388  
     microsociological approaches, 1388–1392  
     parental roles and, 2033, 2037–2038  
     reciprocal altruism and, 1391, 1394, 1657–1658  
     and social security in premodern societies, 2796  
     and social security services, 3065  
     and student movement concerns, 3069–3070  
 Intergenerational resource transfers, **1393–1399**, 1617–1618  
     and educational attainment, 2929  
     equity issue in, 1397–1398  
     exchange theory in, 1394  
     and life cycle, 1615–1616, 1620  
     life-cycle model of, 1394  
     and occupation and status, 2785–2786, 2787, 3042, 3046–3047  
     and political and economic status, 2787  
     reciprocal altruism in, 1394  
     and social mobility, 1982–1984, 2690, 2711–2714  
     socialist elimination of, 2850  
     on societal level, 1395–1397  
     *See also* Caste and inherited status; Status attainment  
 Intergroup and interorganizational relations, **1399–1407**  
     ethnic-racial, 1400  
     homogeneity/heterogeneity in, 1404  
     personality determinants and, 1400  
     prejudice and, 2243–2245  
     systems theory in, 1405  
 Interitem correlations, 190  
 Interlock studies, 736–737  
 Intermarriage, **1407–1415**  
     Asian-American, 1410–1411  
     contemporary, 1408–1410  
     and courtship, 484  
     factors in, 1411–1413  
     history of, 1407–1408  
     identity and, 1411  
     interfaith, 911, 1411, 1776  
     law and, 949, 1776  
     marriage and divorce rates, 1750  
     and occupational achievement, 2691  
     racial demographics, 124  
     and racial hypodescent rule, 2331–2332  
     rate increase, 1776  
     stereotypes and, 1411  
     tolerance for, 1412  
 Internal labor markets, 1985–1987  
 Internal migration, 929, **1415–1422**  
     by African Americans, 2334, 2491  
     Chinese restrictions, 302–303  
     definition of, 1415–1417  
     differential, 1418  
     and ethnonationalism, 1946  
     and growth of cities, 306, 309, 2431  
     international comparisons of, 1420–1421  
     Mexican pattern, 1859–1860  
     Middle Eastern, 1865–1866  
     mobility and, 1415–1419  
     population distribution and, 1419  
     push/pull factors, 1418  
     reasons for, 1418  
     for retirement, 1419–1420

INDEX

- return migration, 1417  
 rural-to-urban, 2431  
 and social networks, 2728  
 and urban underclass, 2212
- Internal Revenue Service, 3103
- International Association for the Empirical Study of Literature, 1649
- International Association of Chiefs of Police, 493
- International Association of French Language Sociologists, 328
- International Association of Time Use Research, 3157
- International associations in sociology, **1422-1426**  
 clinical sociology, 327-328  
 cross-cultural analysis, 550-551  
 life histories and narratives, 1633-1634, 1635  
 rural sociology, 2433  
*See also* International Sociology Association
- International Bibliography of Sociology*, 1606
- International Committee for Sport Sociology, 2986-2987
- International Conference on Population and Development (Cairo; 1994), 932, 2233
- International Convention for the Suppression of Terrorist Bombings (1998), 3140
- International Council of Sport Science and Physical Education, 2986
- International Court of Justice, 1429-1430, 1944, 1945, 3003
- International Covenant on Civil and Political Rights (1966), 1242
- International Crime (Victim) Survey (IC(VS)), 549
- International Encyclopedia of Statistics* (Kruskal and Tanur), 1957
- International Encyclopedia of the Social Sciences*, 746, 1888
- International Federation of Data Organizations, 577
- International Institute of Sociology (IIS), 1025, 1422, 1483
- International Journal of Japanese Sociology*, 1483
- International Journal of Mass Emergencies and Disasters*, 682
- International Journal of Small Group Research*, 1980
- International Labour Office, 1531, 2795, 2800
- International Labour Organization, 2607, 3262, 3263, 3266
- International Labour Review*, 1532
- International law, **1426-1431**  
 arbitration and, 1427  
 censorship and, 279  
 custom and, 1428  
 dispute settlement, 1427  
 enforcement of, 1429  
 on genocide, 1244, 1429  
 individuals and, 1427  
 multinational and transnational systems, 1550-1551  
 natural law and, 1427  
 sources, 1428  
 universal jurisdiction and, 1429  
*See also* International Court of Justice
- International migration, **1431-1438**  
 Asian, 58, 174-177, 713  
 assimilation and, 53, 143  
 bankruptcy and, 204  
 basis of individual group's success in United States, 2332-2333  
 black slave experience vs. immigrant, 54  
 Canadian immigrant occupational status attainment, 2786  
 chain migration, 177  
 citizenship and, 1436  
 consequences of, 1433-1435  
 crime rates and, 532  
 demographic models, 620, 636  
 determinants of volume of, 1432  
 differentials, 1432  
 ethnicity and, 841  
 European discrimination, 691-693  
 and event history analysis, 869  
 family structure and, 122-123  
 fertility determinants and, 1005  
 and globalization of work, 3267  
 Hispanic Americans, 143, 1186, 1188, 1189
- illegal, 1436, 1936  
 and immigrant settlement houses, 2841  
 and immigrant suburban communities, 3074, 3075  
 immigrants as threat to African American employment, 2496  
 Irish American Civil War draft riots, 2269  
 Jewish success factors in United States, 2332, 2333  
 legislation, 1435  
 legislation reform, 178  
 mass society theory on, 1771, 1773  
 Mexican, 1860  
 Middle Eastern, 1865-1866, 1869, 1872-1873  
 of Muslims, 2950-2951  
 national borders and, 1936, 1937  
 negative values of, 1432  
 and Polish life-history project, 1616, 1618, 1634  
 and population redistribution, 2180, 2181-2182  
 racial consciousness and, 2329  
 racial constraints and, 321  
 racial privileges and, 56, 58  
 reasons for, 1432  
 refugee vs. immigrant distinction, 180  
 and religious organizations, 2380-2381  
 and status incongruence, 3051, 3052, 3054  
 stream/counterstream, 1431  
 territorial distribution, 633-634  
 U.S. increased openness to, 143  
 and underemployment, 1721, 1724  
 Western European discrimination against, 692-693  
 of workers, 636, 1338, 1436, 1858, 1865, 2496, 2498, 2608
- International Monetary Fund, 730, 2138  
 dependency theory and, 642, 644  
 and state sovereignty, 3003
- International nongovernmental organizations (INGOs), 1046, 2050

- International Political Science Association (IPSA), 2916
- International Prestige Scale, 1997
- International relations  
 balance of power systems, 332  
 coalitions, 332  
 national border relations, 1931–1939  
 peacekeeping and, 2044–2050  
 and social networks, 2728  
*See also* War
- International Review for Sport Sociology* (journal), 2986
- International Rural Sociological Association, 2433
- International Social Security Association, 2795
- International Social Survey Program, 548, 577, 2477
- International Society for Krishna Consciousness. *See* Hare Krishna
- International Society for Quality of Life Research, 2301
- International Society for Quality of Life Studies, 2686
- International Sociological Association, 328, 550–551, 736, 805, 1423, 1424, 1425, 1426, 2913, 2986  
 Biography and Society Section, 1633–1634, 1635  
 hybrid specialties, 2924  
 Japan Sociological Society and, 1479, 1482
- International Sociological Association of the Research Committee on Disasters, 682
- International Sociological Library, 1025
- International Sociology of Sport Association, 2987, 2988
- International Statistical Congress, 283, 3007
- International Survey of Economic Attitudes, 548
- International system. *See* Globalization and global systems analysis
- International terrorism. *See* Terrorism
- International tourism. *See* Tourism
- International Typographical Union, 1533
- International Workers of the World, 1530
- Internet, **1438–1447**  
 AlohaNet, 1439  
 America-on-Line, 1440  
 applications in sociology, 406–407, 409  
 censorship and regulation of expression issues, 279, 280  
 and census taking, 283  
 chat rooms, 1443–1444  
 communication among sociologists via, 1607  
 content-analysis sites, 421  
 cybercrime and, 3253  
 as dating aid, 489  
 e-mail communication, 406, 407, 408, 1441–1442, 1607, 1768  
 Ethernet, 1439–1441  
 flaming and, 1442  
 free expression and, 271  
 futures studies, 2231–2232  
 as globalization convergence factor, 428  
 hardware for, 1438–1439  
 history of, 1438–1439  
 and information society, 1345, 1346  
 information technology, 1605  
 interactive data analysis Web sites, 409  
 interest groups and, 2150  
 local area networks, 1439–1441  
 mailing lists, 413, 414  
 and mass media research, 1761  
 modalities, 1441–1446  
 multi-user domains, 1444–1445  
 as new media technology, 1768–1769  
 on-line focus groups, 408  
 on-line library catalogs, 1611  
 search engines, 1446  
 search techniques, 1607–1608  
 service providers, 1440  
 and sexually explicit material, 2185  
 sociology sites, 413–414, 1606–1611  
 spamming and, 1441, 1442  
 UseNet and, 1442–1443  
 UseNet Newsgroups, 414  
 World Wide Web, 1346, 1441, 1445–1446  
*See also* Web sites
- Interpersonal attraction, **1447–1450**  
 assumed similarity in, 1447  
 friendship and, 1449  
 mate selection and, 1777–1778  
 play and work issues, 1448–1449  
 propinquity and, 1447  
 relationship formation, 1448  
 romantic love and, 1449–1450, 1696–1701  
 sexual behavior and, 1449–1450, 2537–2548  
 theoretical explanations for, 1447–1448
- Interpersonal attribution, 196–197
- Interpersonal behavior. *See* Interpersonal conflict resolution; Interpersonal power; Personal relations; Small groups
- Interpersonal conflict resolution, **1451–1456**  
 age factors, 1452  
 friends vs. nonfriends, 1453–1454  
 gender and, 1452  
 negotiation strategies model, 1452  
 social information processing and, 1451–1452  
 theoretical models of, 1451–1452  
 within family, 1454
- Interpersonal diffusion networks, 677, 678, 679
- Interpersonal power, **1456–1464**  
 bases of, 1457  
 behavioral exchange theory, 1460–1462  
 case studies, 247  
 coercion and, 1458  
 concepts of, 1457–1458  
 conflict theory, 414–416  
 and emotions, 777–778  
 exchange theories, 1459–1463, 2674  
 expected utility theory, 1460  
 field theory, 1458–1459

- means of, 1457  
 power-dependence theory, 1461  
 rational choice theories, 1459  
 resistance theory, 1462–1463  
 self-presentation and, 2776
- Interpersonal relations. *See* Personal relations
- INTERPOL, 500
- Interpower negotiations. *See* Interpersonal power; Negotiation of power
- Interpretation interaction, 350
- Interpretive research, 2292–2293  
 comparative-historical analysis, 386–388  
 and sociology of education, 2928–2929
- Interpretive sociology (Weber concept), 820
- Interquartile range (IQR), definition of, 660–661
- Interracial marriage. *See* Inter-marriage
- Inter-University Consortium for Political and Social Research, 409, 575, 576, 577–578, 579, 1606, 1684  
 resources and functions, 2474–2475, 2476–2477, 2481
- Interval scales, 1793
- Interview  
 computer-assisted, 410, 1802, 3092  
 election polling, 2274  
 life histories and narrative, 1633, 1635, 1636  
 multistage sampling, 2448  
 survey research, 418–419, 1802, 3091–3092, 3092, 3232
- Intolerance. *See* Discrimination; Prejudice; Stereotypes
- Intrauterine device (IUD), 627, 2178, 2180
- Intravenous (IV) drug use, 712, 1641, 1642, 2586–2592
- Introduction to Mathematical Sociology* (Coleman), 2668, 3036
- Introduction to Multivariate Statistical Analysis, An* (Anderson), 3035
- Introduction to Social Psychology* (McDougall), 114
- Introduzione alla Sociologia clinica* (Luison), 328
- Invariant sequence, 1900, 1901
- Inventions. *See* Technology and society
- Inverting symbiosis, 357
- Involuntary part-time workers, definition of, 1720
- Involuntary servitude. *See* Slavery and involuntary servitude
- Iowa State University, 677
- IPA (Interaction Process Analysis), 1974, 1975, 1976, 1977, 1978, 1979, 1980  
 for small groups, 2613–2614
- IPAs (independent practice associations), 1818, 1819
- IPAT (environmental impact equation), 808
- IPS system, 1974–1975
- Ipsen, Gunter, 1076
- IPUMS (interactive data access system), 409
- IQ. *See* Intelligence
- IQR. *See* Interquartile range
- Iran, 1865, 1866, 1867  
 historical-sociological analyses, 1869, 1870–1871, 1872  
 immigrants in United States, 1872–1873  
 interpenetration church-state model, 2357  
 Islamic fundamentalism, 2371, 2374, 3288  
 Islamic revolution. *See* Iranian revolution  
 leadership behavior, 1564, 1566  
 and nationalist movement, 1945  
 sociodemographic profile, 2938  
 wartime rape, 2579  
 women in labor force percentage, 3262
- Iran-Contra affair, 2128
- Iranian revolution (1977–1979), 1868, 1869, 1870–1871, 1873, 2357, 2371, 2412
- Iranian Studies* (journal), 1872
- Iraq, 1866  
 ethnonationalism, 1944, 1945  
 fertility decline, 627
- Ireland  
 common law system, 465  
 fertility decline, 2178  
 kinship mapping priority, 1515  
 low suicide rate, 3079  
 nationalist movement, 2717, 3001
- IRG. *See* Initial Review Group
- Irish Americans  
 alcohol consumption rates, 95  
 Civil War draft riots, 2269–2270, 3069
- Irish Republican movement, 2717
- Irnerius, 1179
- Iron Curtain, 1934
- “Iron Law of Oligarchy” (Michels), 2623–2624
- Iroquois Confederation, 1070
- Irrigation, 88
- Irwin, John, 2054–2055
- IS. *See* Institute of Sociology, 2982
- ISA. *See* International Sociological Association
- ISA Research Committee on Stratification, 550–551
- ISEA. *See* International Survey of Economic Attitudes
- Iser, Wolfgang, 1648
- ISIG. *See* Istituto di Sociologia Internazionale di Gorizia
- Islam. *See* Islamic societies; Sociology of Islam; World religions
- Islamic fundamentalism, 2371, 2374, 2940–2941  
 emergence and influence of, 2943, 2944–2945, 3288  
 three competing theories of, 2945–2946  
 and women’s status, 2949–2950
- Islamic holy war, 2365, 2371, 2717
- Islamic law, 464, 1546, 1548, 1550, 1554, 2357, 2601  
 fundamentalist emphasis on, 2945
- Islamic societies  
 and African slave trade, 2601  
 and family planning, 2179  
 fertility transition, 628  
 fundamentalist influences, 2371, 2940–2941, 2945  
 gender issues, 2948–2950

- historical empires, 2998  
 and institutional  
   differentiation, 937  
 and interpenetration church-state  
   model, 2357  
 Iranian revolution analysis,  
   1870–1871  
 kinship systems, 1515, 1517  
 migrants to the United  
   States, 1872  
 militancy in, 2946–2948  
 as minorities in other countries,  
   2950–2951  
 “new Arab social order” and,  
   1867–1868  
 political movements and,  
   2365, 2366  
 religious renaissance in, 2946  
 religious traditions and, 3285  
 slavery practices, 2601–2602  
 “social project” of Islam and,  
   2939–2940  
 sociodemographic profile,  
   2937, 2938  
 sociology of, 2937–2951  
 in Southeast Asia, 2974  
 state-religion relationship, 2948  
 typology of, 2947  
 in United States, 2380–2381,  
   2950–2951  
 and women rape victims, 2579  
*See also* Middle Eastern studies
- Isolation, 176, 178  
 Isomorphism, definition of, 2024  
 Isomura, Eiichi, 1480  
 Israel  
   Arab peace accord, 2048  
   demographics, 1865, 1866, 1867  
   educational status attainment,  
     2785  
   and family size, 978  
   kinship mapping priority, 1513  
   long-term care and care  
     facilities, 1659  
   migrants to United States, 1872  
   occupational status  
     attainment, 2787  
   relative marriage rate, 1749  
   religion and politics linkage in,  
     2358–2359  
   Social Science Data Archive, 576  
   and terrorism, 3139  
   widow remarriage, 3255  
   women’s employment, 1867  
 ISSA. *See* International Sociology of  
   Sport Association  
 ISSP. *See* International Social  
   Survey Program  
 Istituto di Sociologia Internazionale  
   di Gorizia, 2231, 2288  
*It Takes a Village* (H. Clinton),  
   361–362  
 Italian Americans  
   alcohol-consumption patterns, 95  
   case studies of, 244  
 Italian Society of Sociology, 1423,  
   1465–1466  
 Italian sociology, **1464–1475**  
   academic institutionalization of,  
     1468–1469  
   clinical, 328  
   life histories and narratives, 1635  
   origins of, 1464–1469  
   planning and, 1468  
   postwar, 1466–1469  
   social change and, 1471–1473  
   Social Science Data Archive, 575  
   university study, 1469–1471  
   urban, 1468  
 Italy  
   civil law system, 471, 473, 474,  
     475, 476, 477, 478, 480  
   dictatorship, 3002  
   disaster research, 687  
   divorce rate, 112  
   fertility transition, 626  
   foreign-controlled pharmaceutical  
     companies, 1827  
   governmental division of  
     powers, 1954  
   health-care system, 374, 377  
   labor movement, 1529  
   long-term care and care facilities,  
     1655, 1661  
   low suicide rate, 3079  
   Mafia in, 2018, 2128  
   marijuana decriminalization, 712  
   organizational demographics, 395  
   political and governmental  
     corruption, 2128–2129  
   political party system, 2157  
   retirement policy, 2407, 2408  
   revolution, 3001  
   social security system  
     spending, 2800  
   tourism in, 3169  
   transnational corporations,  
     3175, 3176  
   women in labor force  
     percentage, 3262  
   *See also* Italian sociology  
 IUD (intrauterine device), 627,  
   2178, 2180  
 IV-drug use. *See* Intravenous  
   (IV) drug use  
 “I-We” paradigm, 1601  
 IWW (International Workers of the  
   World), 1530
- J**  
 Jackknifing, 2397, 2449, 3039  
 Jacklin, C. N., 2530–2531  
 Jackson, Andrew, 2127  
 Jackson, B. A., 1653  
 Jackson, Don, 540, 1713–1714, 2084  
 Jackson, E. F., 2345  
 Jacobs, Herbert, 467  
 Jaeger, Gertrude, 566  
 Jaffe, M. W., 1653  
 Jaffee, David, 643, 2887  
 Jahoda, Marie, 2188–2189, 2190,  
   3155  
 JAI Press, 407  
 Jains, 3282, 3284, 3285  
 Jakobson, Roman, 1032, 2891  
 Jakubs, John F., 2502  
 Jamaica  
   African culture, 65  
   African slavery in, 2600  
   fertility decline, 627  
   HIV/STD control program, 2593  
   low suicide rate, 3079  
 James, David R., 2501  
 James, William, 2, 1255, 1783,  
   2082, 2083  
   on global self-esteem, 2512  
   and pragmatism, 2218, 2219,  
     2423, 3098  
   on religious experience, 2373,  
     3279, 3283, 3286

- and role theory, 2423, 2424, 3100
- Jameson, Fredric, 2200
- Jamison, Andrew, 1927
- Janet, Pierre, 540, 1713–1714, 2084
- Janis, Irving, 400–401, 2512, 3244–3245
- Janissary rebellion, 3000
- Janowitz, Morris, 555, 1876
- Jantsch, Eric, 1037, 1040
- Japan
- authoritarian communitarianism, 356, 3218
  - caste and inherited status, 253–254
  - conditions conducive to democracy, 605, 606
  - convergence theory and, 424
  - divorce rate, 112, 706
  - health-care system, 374, 375, 376, 377, 379–380, 1827
  - hegemonic stability theory and, 3242
  - high suicide rate, 3079
  - hours of work, 3262
  - labor movement, 1531
  - legal system, 464, 469, 471, 475, 479, 1550
  - life expectancy rate at birth, 223, 1631
  - long-term care and care facilities, 1654, 1655, 1659–1660, 1661
  - Meiji restoration as “elite” revolution, 2411
  - political corruption, 2131
  - population growth, 2179
  - postindustrialism, 2201
  - rape incidence, 2587
  - retirement policy, 2407, 2408
  - social movement emergence, 2719–2720
  - social security system spending, 2800
  - time use survey, 3160
  - transnational corporations, 3175, 3176
  - voting behavior research, 3237–3238
  - women in labor force percentage, 3262
  - World War II conquest of Southeast Asia, 2975
  - World War II war crimes trials, 1429
  - See also* Japanese sociology
  - Japan Institute of Sociology, 1423, 1477
  - Japan Society for the Promotion of Science, 1482
  - Japan Sociological Society, 1423, 1477, 1479, 1480–1481, 1483
  - Japanese American Evacuation and Resettlement Study, 180–181
  - Japanese Americans
    - academic studies of, 178, 179, 180–181
    - demographic characteristics, 176
    - immigration, 123, 175
    - income and status attainment, 181
    - World War II relocations, 123, 177, 178, 180–181
  - Japanese Prime Minister’s Office, 1482
  - Japanese Sociological Review*, 1480, 1482
  - Japanese sociology, **1477–1484**
    - diversification stage (1960s–19920s), 1479–1481
    - globalization state (1990s and beyond), 1481–1484
    - leader behavior study, 1566
    - and mathematical sociology, 1791
    - postwar state (1946–1960s), 1478–1789
    - pre-World War II stage (1893–1945), 1477–1478
  - Japanese War Crimes Trials, 1429
  - Jargowsky, Paul A., 2502
  - Jarley, Paul, 606
  - Jasso, Guillermina, 1790, 2703, 2704, 2705
  - Jauss, Hans Robert, 1648
  - Java. *See* Indonesia
  - Jayaram, Jayalitha, 2132
  - Jazz, 1925, 1926
  - J-curve thesis, 2270
    - of revolution, 349
  - Jefferson, Gail, 247, 431, 435–437, 438
  - Jefferson, T. (sociologist), 1578
  - Jefferson, Thomas, 273, 584, 2265, 2427, 2483
  - Jehovah’s Witnesses, 2366
  - Jellinek, E. M., 96
  - JEMCO Workshop Study, 1979
  - Jencks, Christopher, 1375
  - Jenkins, Craig, 2166
  - Jenkins, Gwilym M., 3036, 3143, 3144
  - Jenness, Valerie, 2764, 2961
  - Jennings, Helen Hall, 2729
  - Jensen, Arthur, 2140
  - Jensen, Jay P., 2189–2190
  - Jensen, Richard, 3231
  - Jersey City (New Jersey) political machine, 2126
  - Jesus, 3280, 3281, 3283–3284, 3285, 3287
  - Jesus movements, 3287
  - Jet* magazine, 2493
  - Jews. *See* Judaism and Jews
  - Jihad. *See* Islamic holy war
  - Jiji Press, 1482
  - Jim Crow laws, 54, 62, 2491, 2761
  - Jipemoyo Project, 2040
  - Joas, Hans, 2220
  - Job cluster, 1985
  - Job description, 2422
  - Job discrimination. *See* Work and occupations
  - Job displacement, 1722–1723
  - Job orientation. *See* Work orientation
  - Job satisfaction, 3270, 3272–3373
    - gender and, 3275
  - Job segregation. *See* Occupational segregation
  - Job shifts. *See* Occupational and career mobility
  - Job trajectory
    - definition and concept of, 1982, 1984
    - See also* Occupational and career mobility
  - Job-evaluation, as comparable worth remedy, 371–372
  - Johansen, Robert, 1981
  - Johansson, Sten, 2452
  - John Birch Society, 462

- John M. Olin Foundation, 1601
- John Paul II, Pope, 2268
- John XXIII, Pope, 2365
- Johns Hopkins University, 1180, 1423
- Johnson, L. (sociologist), 2581
- Johnson, Lyndon B., 1286, 1435, 2682
- Civil Rights Act of 1965, 2496
- Great Society, 2299
- public opinion and, 2275, 2276, 2277
- War on Poverty, 1286, 1494, 1882, 2404, 2760
- Johnson, Marilyn E., 1618
- Johnson, Richard, 1756
- Johnson, Virginia E., 2554
- Johnson Publications, 2493
- Johnston, Michael, 2125
- Johnstown, Pennsylvania, 366
- Join Together (drug-abuse prevention program), 716
- Joint distribution, 658, 661, 2250
- Joint Information Bureau (Department of Defense), 278
- Jonassohn, Kurt, 1068, 1069, 1070, 1071
- Jonathan (biblical figure), 1508
- Jones, Edward, 193
- Jones, Gavin, 633
- Jones, J. R., 1771–1772
- Jones, Jim, 900
- Jones, Karen, 353
- Jones, Lois Mailou, 64
- Jones, Paula, 2581
- Jones, Richard S., 245
- Jones, Simon, 226
- Jordan, 1866
- fertility decline, 628
- sociodemographic profile, 2938
- Jöreskog, Karl, 908, 1923, 3037
- Josephs, R. A., 73–74, 2513
- Jourard, S. M., 2189
- Journal of Asian American Health*, 179
- Journal of Asian American Studies*, 179
- Journal of Comparative Studies in Society and History*, 1197
- Journal of Contemporary Ethnography*, 854
- Journal of Educational Sociology*, 326
- Journal of Experimental Social Psychology*, 2650
- Journal of Family Issues*, 1738
- Journal of Health and Social Behavior*, 1814
- Journal of Marketing Research*, 407, 409
- Journal of Marriage and Family*, 106
- Journal of Mathematical Sociology*, 1787, 1790, 2028
- Journal of Narrative and Life History*, 1635–1636
- Journal of Social Issues*, 2042
- Journal of World Systems Research*, 413
- Journals. *See* Publications; *specific titles*
- Jouvenel, Bertrand de, 1037, 1038
- JSTOR program, 1606, 1607
- Juarez, Benito, 1857
- Judaism and Jews
- American affiliation figures, 146
- and anti-Semitism, 540
- and black civil rights movement participation, 2268–2269
- ethical impulse/moral law relationship, 3283
- interfaith marriage rate upswing, 1776
- Islamic parallels, 2937, 2939, 2943
- and Israeli politics, 2359
- Kabbalah mysticism, 2969, 3286
- kinship and family typology, 1507, 1510–1511, 1512, 1513, 1514, 1517
- liberalism and neoconservatism, 1603
- low alcohol abuse rates, 95
- medieval messianic movements, 2968–2969
- Nazi genocide. *See* Holocaust
- and new religious movements, 3287
- and racial categorization, 2332
- religious experience, symbols, and theology, 3279, 3280, 3281, 3282, 3284
- religious organization perspectives, 2377
- success in American society, 2332, 2333
- widowhood, 3255
- Judd, Dennis, 3071
- Judd, Walter, 175
- Judges, 476–478
- common law vs. civil law system, 465, 467, 468, 472, 476–478, 480–481
- historical power of, 474
- inquisitorial vs. accusatory legal model, 479
- and negotiation of power, 1953
- numbers by countries (table), 471
- treatment of forms of expression, 270–272
- Judgment
- absolute vs. comparative, 597
- consistency studies, 591
- general, 598
- knowledge effect on, 593
- moral. *See* Moral development
- Judgments. *See* Decision-making theory and research
- Judicial power, 465
- Judicial review, 467
- Judicial sociology, 1024
- Jung, Carl G., 1714, 1718, 2076, 2077, 2084, 2087, 2090
- Junior Chamber of Commerce International, 3229
- Junior colleges, 1180
- Juppé, Alan, 2129
- Juréen, Lars, 3035
- Jüres, Ernst-August, 1078
- Jus commune*, 473
- Just reward, 2703, 2704
- Just society. *See* Social justice; Utopian analysis and design
- Justice. *See* Court system and law; Criminal justice system; Social justice
- Justice Department, U.S., 714, 982
- Justice without Trial* (Skolnick), 2108, 2114
- Just-identified models, 191, 1915–1917
- Justinian, emperor of Rome, 465, 473, 476, 2999



INDEX

- Juvenile delinquency, theories of, 3, 502, 507, **1493–1499**  
 anomie, 1493–1494  
 anomie and rising expectations theory, 1491  
 anomie and strain theories, 166  
 conflict theory, 1497–1498  
 control theory, 667, 1495  
 cultural influence, 663, 668, 1494  
 differential association theory, 666, 667  
 labeling theory, 1496–1497  
 macro-level origins, 663, 664–665  
 micro-level origins, 667–668  
 neutralization theory, 1496  
 social class, 1491–1492, 1494  
 and social controls, 2658  
 social disorganization theory, 1495  
 structural functionalism theory, 1493–1495  
 symbolic interactionism theory, 1495–1497
- Juvenile delinquency and juvenile crime, 521, **1484–1493**  
 AIDS/HIV risks and, 2586  
 community studies, 365, 366  
 and criminal and delinquent subcultures, 510–513, 534  
 differing legal definitions of, 1485–1486  
 extent and trends, 1486–1489  
 factors in, 1489–1492  
 gangs, 460, 512, 1485  
 historical, 1485  
 life history, 1616  
 treatment for violent offenders, 76  
*See also* Criminal sanctions; Gangs
- K**  
*Kabbalah* (Jewish mystical writings), 2969, 3286  
 Kabo, V. E., 2979  
 Kagan, Jerome, 2090  
 Kagitcibasi, Cigden, 3218, 3219  
 Kahn, Herman, 1041, 1218  
 Kahn, Robert, 2415, 2416  
 Kahneman, D. J., 591–592, 595, 598
- Kaigo Hoken (long-term care insurance), 1659–1660  
 Kairys, David, 1556  
 Kali (Hindu goddess), 3280, 3284, 3285  
 Kalish, Richard, 582  
 Kalleberg, Arne L., 1985–1986, 3275  
 Kalton, Graham, 190–191  
 Kamo, Y., 696  
 Kanka, Megan, 2582  
 Kansas City (Missouri) political machine, 2126  
 Kansas Marital Satisfaction (KMS) scale, 1728  
 Kant, Immanuel, 1302, 1303, 1304, 3244  
 and German idealism, 1248, 1249  
 Kanter, Rosabeth Moss, 244, 2262  
 Kantorowicz, Hermann, 1074  
 Kanungo, R. A., 1566  
 Kapadia, K.M., 1291  
*Kapital, Das* (Marx), 539, 2814  
 dedicated to Darwin, 573  
 Kaplan, Abraham, 2023  
 Kaplan, S., 1978  
 Kareev, N. I., 2979  
*Karl Marx's Theory of History* (Cohen), 1784  
 Karmarck, Andrew W., 2921–2922  
 Karp, David, 245, 360  
 Kasaba, Resat, 1871–1872  
 Kasarda, John, 307, 3071, 3198  
 Kashmi, 1945  
 Kashmir, 2366  
 Katovich, Michael, 2222  
 Katz, Elihu, 677–678, 679  
 Katz, Jack, 507  
 Katz, Leon, 2729  
 Katz, Michael B., 2211  
 Katz, S., 1653  
 Kauffman, Kelsey, 2053  
 Kautsky, Karl, 1753, 2432  
 Kazakhstan, 1946, 2362  
 sociodemographic profile, 2938  
 sociology as discipline, 2982  
 Kearl, Michael C., 583  
 Keating, Charles, 2128  
 Keddie, Nikki R., 1871, 2948  
 Kefauver hearings (1951), 2126
- Keith, Jennie, 245  
 Kejner, Matilde, 3273  
 Keller, Evelyn Fox, 994–995, 2460  
 Keller, Fred, 215  
 Kelles-Kraux, Kazimierz, 2118  
 Kelley, H. H., 2065, 2670  
 Kelley, Harold, 193, 2419  
 Kellner, Douglas, 545, 1757  
 Kellner, Hansfried, 1886  
 Kellogg-Briand Pact (1928), 2270  
 Kelly, Edward, 2126  
 Kelly, George A., 1715–1716, 1717, 2076  
 personality theory, 2084  
 Kemp, Jack, 362  
 Kempe, C. H., 288  
 Kemper, Theodore, 777–778, 780, 785, 2522  
 Kempny, Marian, 2119  
 Kendall, Maurice G., 1796, 3035  
 Kendall, Patricia L., 3112  
 Kendall's coefficient of concordance, 1957  
 Keniston, Kenneth, 459  
 Kennedy, Gail, 2217  
 Kennedy, John F., 48, 1435, 1494  
 Kennedy, Paul, 1708  
 Kenny, John V., 2126  
 Kent State University, 2270  
 Kenya  
 anti-AIDS/HIV campaign, 2593  
 fertility rate decline, 220, 627–628  
 poverty in, 2216  
 Kepler, Johannes, 2465  
 Kerckhoff, Alan, 1778  
 Kerr, Clark, 422–423, 424, 1180  
 Kerr, S. C., 1566  
 Kessler, Ronald C., 1692, 3081  
 Kesting, Hanno, 1078  
 Ketkar, S. V., 1291  
 Kevorkian, Jack, 585, 587, 3084, 3085–3086  
 Kewus, Oscar, 2211  
 Key, V. O., Jr., 2124  
 Key, Valdimer Orlando, Jr., 3235  
*Key Problems in Sociological Theory* (Rex), 226  
 Keyder, Caglar, 1871–1872, 1876

- Keyfitz, Nathan, 631
- Keynes, John Maynard, 1265
- Keyton, Joann, 2418
- “Khadi justice,” 464, 1546, 1548
- Khaldun. *See* Ibn Khaldun, Abd-al-Rahman
- Khalil, Mohammad, 643
- Khomeini, Ayatollah Ruhollah, 2357, 2371
- Khrushchev, Nikita, 2980
- Kidd, Quentin, 3223
- Kiecolt, R. Jill, 2478
- Kiev International Institute of Sociology, 2982
- Kiev-Mohyla Academy, 2982
- Kikuchi, Charles, 181
- Kikuchi Diary*, 181
- Kilakowski, Leszek, 1356
- Killian, Lewis, 351, 352, 553, 554, 558, 559, 560
- Kim Hyun Chul, 2131
- Kim Young Sam, 2131
- Kimball, Peter, 352
- Kimura, D., 2884, 2885
- Kindergartens, 677
- Kindleberger, Charles, 2823
- Kinesiology, 2987, 2988
- King, Alexander, 1038
- King, Gary, 1595
- King, Martin Luther, Jr., 178, 556, 2146, 2269, 2494, 2495, 2496
- “Letter from a Birmingham Jail,” 2495
- nonviolence philosophy, 3289
- and plight of Northern urban blacks, 2496–2497
- “Kinkeeper” role, 696
- Kinsey, Alfred, 111, 2538, 2541, 2549, 2553, 2554, 2559, 2560, 2561, 2566, 2569, 2572
- Kinsey Institute for Sex Research, 2187
- Kinship systems and family types, **1501–1519**
- African American, 122, 705, 927
- African slave masters and, 2601
- alliance theories of, 1507–1508
- altruism and, 2882–2883
- American systems, 142
- centrepetal vs. centrifugal, 1510–1512
- classificatory, 2298
- coalitions within, 331
- courtship and mate selection and, 1698, 1699–1700
- divorce-related, 709
- ethnicity as extended kinship group, 2329
- evolution theory and, 2882–2883
- and family and household structure, 926–927
- and family and population policy in less developed countries, 965
- high fertility rates and, 628
- historical typologies, 1507–1517
- historical typologies critiques, 1505–1507
- illegitimacy and, 1261
- incest taboos, 1270–1277
- Japanese American, 123
- and life course, 1615, 1618, 1620
- and life cycle, 1625
- mapping priorities, 1512–1515
- and marital eligibility limitations, 484, 1270–1277, 1509, 1513, 1776
- marital unity vs. sibling, 1509
- marriage as central to, 1733–1734
- Mexican American, 123
- nationalist movements and, 1942
- parental role and, 2034
- religion and, 935, 937
- slavery as severance of, 2596, 2597
- social structure and, 1509–1512
- socialization and, 2858
- See also* Nuclear family
- Kirk, Russell, 1598
- Kissinger, Henry, 2048
- Kitschelt, Herbert, 2160
- Kitson, Gay, 1737
- Kitsuse, John I., 2763
- Kittrie, Nicholas N., 2143, 2146
- KKK. *See* Ku Klux Klan
- Klages, L., 2085
- Klandermans, Bert, 2763
- Klapper, Joseph, 1762
- Klare, Karl E., 2961
- Klein, Melanie, 650
- Klein, Rudolof, 378
- Kline, R. B., 1923
- Klineberg, Otto, 1425
- Kloppenburg, Jack R., Jr., 2460
- Kloskowska, Antonina, 2121
- Kluckhohn, Clyde, 564, 2890
- Kluckhohn, Florence R., 3212, 3214
- Knapp, Peter, 2919
- Knights of Labor, 1530
- Knoke, David, 604, 606, 2165, 2380
- Knorr Cetina, Karin, 2469–2470
- Knowledge
- attitudes as function of, 185
- common sense, 1377–1379, 2100–2102, 2959
- effect on decision-making, 593, 678
- Foucault’s “archaeology,” 2647–2648, 2757
- “new sociology” of, 2958–2959
- organizational, 584
- as participatory research emphasis, 2040
- postindustrial focus on theoretical, 2196, 2205
- social learning theory and, 70, 75
- sociology of, 2953–2959
- sociology of scientific, 2458–2460
- See also* Epistemology; Scientific explanation
- “Knowledge workers,” 2626
- Kogan, M., 1692
- Kohl, Helmut, 2123
- Kohlberg, Lawrence
- and gender identity development, 998–999
- moral judgment theory, 114, 993, 1894, 1895–1904, 2089, 2092
- stages of moral development, 1896–1897
- Kohli, Martin, 1635
- Kohn, Melvin L., 1619, 2070–2071, 2072, 3214, 3271
- Kolata, Gina, 2550
- Kolmogrov-Smirnov test, 1965–1966
- Komitet Gosudarstvennos Bezopasnoti, 2137
- Kondratjev, N. D., 2979

- König, René, 1074, 1075, 1077, 1425  
 Konrad, Gyorgy, 2117  
 Konvitz, Milton, 2217  
 Koos, Earl L., 1813  
 Koran, 328, 1550, 2939, 2940, 2950, 3281, 3285  
     on women's status, 248–249, 250  
 Korea. *See* South Korea  
 Koreagate, 2127–2128  
 Korean Americans, 175  
     entrepreneurs and  
         professionals, 180, 182  
     household structure, 127  
 Korman, A. K., 1566, 1567–1568  
 Kornai, Janos, 2117  
 Kornhauser, William, 356, 1773  
 Korsch, Karl, 539  
 Kosaka, Kenji, 2298  
 Kosovo, 1944, 1945, 1946, 1948, 2080, 2143, 2362, 2529, 2579, 2608, 3001  
 Koss, Mary, 2557  
 Kotz, Hein, 472, 1545, 1547  
 Kovalevsky, Maxim, 2979  
 Kovel, Joel, 2245  
 Krafft-Ebing, Richard von, 2566  
 Kramer, P. D., 2513  
 Krasnodebski, Zdzislaw, 2119  
 Kraus, V. E. O., 1998  
 Krauss, Robert M., 2620  
 Krebs, Dennis L., 118, 1902–1903  
 Kretschmer, E., 1717  
 Krieger, Nancy, 57  
 Kriesi, Hanspeter, 462  
 Krippendorf, J., 3172  
 Krishna (Hindu deity), 3280, 3281, 3282, 3285  
 Kristiansen, Connie M., 3215  
 Kristol, Irving, 1601  
 Kroeber, Alfred L., 133, 564–565, 566, 675, 679, 2890  
 Krohn, Marvin D., 671  
 Krukskal, William A., 1811  
 Kruskal, William H., 1811, 3036  
 Kruskal-Wallis test, 1960, 1962  
 Krysan, Maria, 317  
 Krzeminski, Ireneusz, 2119  
 Krzywicki, Ludwik, 2118  
 Kshatriya (Hindu warrior-chief), 250  
 Ku Klux Klan, 460, 461, 462, 2266  
 Kubitschek de Oliveira, Juscelino, 2135  
 Kübler-Ross, Elisabeth, 582, 583  
 Kubovy, M., 593  
 Kuhlman, D. Michael, 3221  
 Kuhn, Thomas S., 164, 574, 823, 3098  
     model of scientific change, 2024–2027, 2193, 2374, 2458–2459, 2756  
     on paradigm, 2023  
 Kukathas, Chandran, 356  
 Kumina (Ashanti religion), 65  
 Kuper, Leo, 1068, 1071, 1072  
 Kurczewski, J., 2121  
 Kurdistan nationalism, 1871, 1942, 1945, 3001  
 Kurian, George T., 2677  
 Kurtines, William, 1898  
 Kurzman, Charles, 1871  
 Kuwait, 1865, 1866, 1867, 1944  
     fertility decline, 628  
     sociodemographic profile, 2938  
 Kvalevsky, M. M., 2979  
 Kwaanza, 67  
 Kwakiutl, 1032  
 Kwasniewicz, Wladyslaw, 2119
- L**  
 La Barre, Weston, 2367  
 La Gory, Mark, 1215, 1226  
 Labeling theory  
     attribution theory integration, 198  
     and criminalization of deviance, 243, 520, 525, 534–535, 2659  
     and deviant behavior, 520, 669–670  
     and ethnography, 852  
     and juvenile delinquency, 1496–1497  
     and legislation of morality, 1576, 1577  
     and medical sociology, 1815–1816  
     and mental illness, 669, 1836–1837, 1838, 1840  
     and social problems, 2760  
     and socialization, 2859  
 Labor, forced. *See* Slavery and involuntary servitude  
 Labor Department, U.S., 2398  
 Labor force, **1521–1527**  
     adequate-to-underemployment transition, 1722  
     affirmative action and, 47–52, 2496  
     African Americans and, 55–56, 2491–2492, 2495–2496  
     alienation theory, 100, 104  
     Asian Americans and, 181  
     and capitalism, 237, 238, 239, 240, 320  
     case studies, 246  
     change determinants and consequences, 1523–1526  
     in China, 302–303  
     comparable worth and, 369–372  
     composition of, 3262  
     conditions conducive to democracy, 605  
     and conflict analysis of crime, 536  
     contingent workers, 1723–1725  
     control of work in, 3266  
     convict labor and, 518  
     and crowd behavior, 558  
     definition of participation rate, 1521  
     and divorce pattern changes, 703, 705  
     dual and segmented market theory, 1985  
     educational attainment and, 2930–2931  
     emergence of, 3262  
     and equality of opportunity, 827–828, 830  
     and event history analysis, 869, 3262  
     extent of paid work, 3262–3263  
     externalization and job erosion, 3267–3268  
     and family size, 972, 973  
     and fertility determinants, 624, 635, 1009  
     globalization and, 1088, 1526, 3267, 3275  
     and Hindu caste system, 252–253

- Hispanic-American participation  
in, 1191, 1193
- home workers, 1339, 1341,  
3267–3268
- immigrant vs. black workers,  
2496, 2498
- and income distribution, 1283
- industrial sociology and, 1312
- informal economy and, 1339
- internal vs. external market,  
1985–1987
- job displacement in, 1722–1723
- job preparation for, 3263
- job segregation in, 379, 2012,  
3046, 3262, 3264–3265
- marginal employment, 1719–1725
- measurement of, 1521–1523
- noncitizen workers, 2608
- occupational and career mobility,  
1982–1994
- occupational opportunities  
barriers, 832
- older participants, 2407
- and organizational structure,  
2002, 2003–2004, 2012–2014
- participation of men and women  
(1976–1998), 2405
- part-time workers, 3262–3263
- predecessors of modern,  
3261–3262
- reentry into, 2406
- retirement from, 2403–2408,  
3061
- rural analysis, 2430
- segmented, 2715
- segregation and, 2491–2492
- structural lag effects on,  
3061–3062
- structure of, 1984–1991
- technological change and,  
3266–3267
- urban underclass and, 3198
- vacancy chain models, 2691
- wage labor system and, 2596
- women’s disadvantage in,  
1720–1721
- women’s participation in, 123,  
126, 127, 142, 239, 424–425,  
512, 624, 625, 703, 705, 962,  
972, 981, 1009, 1219,  
1523–1526, 1579, 1729, 1837,  
2032, 2262, 2404, 2406, 2532,  
3046, 3064, 3262, 3266
- worker characteristics, 1991–1993
- young males and, 1524, 1526
- See also* Division of labor; Labor  
movements and unions;  
Migrant workers; Slavery and  
involuntary servitude; Work  
and occupations
- Labor movements and unions,  
**1527–1535**
- in American society, 143, 1528,  
1529–1531, 1553
- and crowd behavior, 558
- and economic institutions, 727
- hours and pay laws, 3262
- internal labor market and, 1987
- materialist theory on, 1785
- membership decline, 3276
- nineteenth-century roots of, 1309
- oligarchical leadership in, 2624
- origin of, 1528, 1533
- participatory democracy in, 605,  
606, 1532–1533
- pension negotiations and, 2403
- Polish Solidarity, 2268
- as political organizations,  
2148–2149
- protests and demonstrations,  
2266
- trends in, 1531–1533
- types of, 1529
- Labor theory of value, 1754,  
2697, 2698
- Labor Utilization Framework, 1720
- Labov, Teresa G., 2906–2907
- Labov, William, 2900, 2901, 2908
- Lachmann, Richard, 2414
- Laclau, Ernesto, 545
- La Guardia, Fiorello, 2125
- Laicization (Durkheim term), 2484
- Laissez-faire leadership, 1565
- Lakota, 3277–3278, 3279, 3281, 3282
- Lalonde, R. J., 2442
- Lamarck, Jean Baptiste, 876
- Lamarckism, 1714, 2460
- Land, Kenneth C., 2301, 2302
- Land-grant colleges, 88, 1180
- Landolt, Patrica, 2640
- Landrecht* (Prussia; 1794), 475
- Landsberg, Martin, 642
- Landshut, Siegfried, 1075
- Landsman, T., 2189
- Lane, David, 224
- Lane, William, 2849
- Lang, Gladys Engel, 173, 1764
- Lang, Kurt, 173, 1764
- Language
- African families, 63–64
- Asian-American proficiency in,  
181–182
- bilingual education, 123, 1861,  
2140, 2908
- describing universal personality  
traits, 2092
- grammar analysis, 438–439,  
2297–2298
- homogenization policies, 1946
- instrumentalist theory of, 2219
- postmodernism on, 2205–2206,  
2207, 2757
- poststructural theory, 2206
- as religious experience  
expression, 3279–3280
- See also* Conversation analysis;  
Sociolinguistics
- Lantenari, Vittorio, 2367
- Laos, 2974, 2975, 2978
- Laotian Hmong Americans, 175, 180
- LaPalombara, Joseph, 2154
- Larceny, Uniform Crime Reports  
definition, 492
- Large-scale multiwave surveys, 1686
- Larson, Magali Sarfatti, 2261
- Larson, Reed, 2300, 2304
- Lasch, Christopher, 2200
- Laschi, Rodolfo, 1464
- Laser systems, 1345
- Lasswell, Harold, 1039, 1761,  
1766, 1940
- “Last Acts” program, 588
- Last Chapters* (Marshall), 583
- Last Supper, 3281
- Last Tango in Paris* (film), 2185
- Latané, Bibb, 115, 116–117
- Latent constructs, 1907
- Latent homosexuality concept, 2566
- Latent structure analysis, 3038, 3039

- Latent Structure Analysis* (Lazarsfeld and Henry), 3038
- Latin America
- abortion illegality, 2240, 2241
  - AIDS/HIV incidence, 2592
  - audencia*, 474
  - child labor, 3262
  - corruption in, 2134–2136
  - court and legal system, 465, 474
  - democratization process in, 1539–1543, 2159, 2160
  - Demographic and Health Surveys, 633
  - demographic transition, 627
  - dependency theory, 641–643, 1535, 1538, 1761, 2922
  - drug trafficking, 2135
  - economic liberalization, 1539–1543
  - fertility rate decline, 220, 1008, 2178
  - frontiers and borders, 1933–1934
  - health-care systems, 381
  - labor movements, 1532
  - life expectancy, 623
  - military dictatorships, 3002, 3003
  - population factors, 2182
  - poverty, 2216
  - protest movements, 2266
  - women in labor force percentage, 3262
  - See also specific countries*
- Latin American studies, **1535–1544**
- Marxist influence, 1757
  - participatory research, 2040
- Latinos/Latinas. *See* Hispanic Americans
- Latour, Bruno, 2459
- Latvia, 2362
- Laub, John H., 535, 668, 1616
- Laudan, Larry, 2026
- Laumann, Edward O., 2165, 2539, 2541, 2542, 2545, 2550, 2554, 2571, 2579, 2693, 3091
- Laurence, J. R., 2083
- Law and legal systems, **1544–1552**
- alternative dispute resolution, 1550
  - antidiscrimination legislation, 692, 693
  - autopoietic, 1548, 1557–1559
  - civil law tradition, 464, 465–466, 471, 472–476, 480–481
  - common law system, 464, 465–472, 474, 475, 477–478, 480–482
  - court systems, 464–481
  - critical legal studies movement, 1548, 1556–1557
  - definitions of law, 144–145
  - delinquency definitions, 1485–1486
  - discrimination-testing litigation, 689, 691, 692–693
  - formal vs. substantive, 1546
  - formalist theories vs. Marxist theories, 1548
  - historical imperial lawgivers, 2998, 2999
  - international, 1426–1430, 1550–1551
  - intestacy law, 1513
  - judiciary. *See* Court systems and law; Judges
  - jurisprudential vs. sociological perspectives, 2960–2961
  - layering (transplant) in, 465, 1550, 1556
  - legal autonomy theory, 1555–1556, 1557
  - legal profession. *See* Lawyers
  - legal systems comparison, 464, 1545–1551, 1554
  - legal tradition concept, 1547–1548
  - macrocomparisons, 1547–1549, 1550
  - microcomparisons, 1549–1551
  - public opinion and, 2277
  - social impact of legal changes, 2962
  - social origins of laws, 2961
  - social stratification and, 2961–2962
  - social systems interaction, 1559–1561
  - social-structural interactions, 1559–1561
  - sociology of law relationship, 1555–1559, 2960–2963
  - Weber typology, 464
- Law and society, **1552–1563**
- alcohol prohibition laws, 1576–1577
  - American Indian status, 135–136
  - civil liberties, 317
  - cohabitation, 109
  - corporations, 443–444
  - divorce impediments, 701
  - divorce revisions, 701, 703, 704, 709
  - drug regulation, 711, 712–713
  - and evolution of criminal sanctions, 516
  - family law, 947–951
  - filial responsibility, 1018–1019
  - homosexual activity, 111
  - immigration restrictions, 143, 175
  - kinship mapping, 1513–1515
  - legal theories, 1555–1559
  - legislation of morality, 1560, 1575–1581
  - marital law for affines and cousins, 1509
  - marriage and inheritance, 1509, 1513
  - marriage restrictions, 1272
  - obscenity laws, 2185–2186
  - personal conduct laws, 1560
  - physician-assisted suicides, 3084–3086
  - protest movement results, 2266–2267
  - rape revisions, 2576, 2677, 2678
  - right to die cases, 585, 586–587
  - sexual harassment policies, 2580–2581
  - social-structural model, 1559–1561
  - sociological theories, 1554–1555, 2960–2963
  - suppression of socially threatening speech, 273–274
  - victimless crimes, 1576, 1577–1578
  - Weber's definition of law, 464
  - white-collar crime, 530, 3245–3255
  - See also* Court systems and law; Criminal sanctions

- Law enforcement. *See* Criminal justice system; Criminology; Penology; Police
- Law of regression. *See* Correlation and regression analysis
- Lawrence, Bruce, 2372
- Lawrence, Joseph J., 2660
- Laus* (Plato), 2086
- Laws, Curtis Lee, 2369
- Lawson, Kay, 2155
- Lawyers  
     dominance in common law systems, 466, 468–472  
     national comparisons, 478–479  
     numbers by countries (table of), 471  
     paraprofessionals, 2260  
     as profession, 2259  
     women’s status as, 468, 2262, 2263
- Layoffs and displacement, 3265
- Lazarsfeld, Paul, 574, 679, 1160, 1425, 1591, 2193, 2755–2756, 3049, 3054  
     and latent structure analysis, 3038  
     and panel design, 1686, 3036  
     and tabular analysis, 3107, 3112  
     time use study, 3165  
     and typologies, 3181, 3182, 3187  
     and voting behavior research, 3, 3233, 3234
- Lazarus, Richard S., 2066, 3057
- Lazzarini, Guido, 1472
- Leach, Edmund, 289, 2891
- Leacock, Eleanor, 991, 2891
- Leader Behavior Description Questionnaire, 1566
- Leader Opinion Questionnaire, 1566
- Leadership, **1563–1575**  
     behavioral approach, 1565–1569  
     by elite, 2163  
     contingency approaches, 1567–1569  
     effectiveness measurement, 1571–1572  
     great-man vs. situation approach, 1564–1565  
     historical review, 1564–1565  
     management vs., 1564  
     and nationalist movements, 1943  
     oligarchical, 603  
     political elites, 2623–2624  
     power management by, 2997–2998  
     public opinion and, 2278  
     research challenge, 1571–1572  
     situational studies, 1567  
     in small groups, 695, 2619–2620  
     social exchange theory and, 2671  
     of state, 2997–2998  
     traits theory, 1565, 1570–1571
- League of Nations, 1945  
     Convention to Suppress the Slave Trade and Slavery, 2602
- Leamer, Edward E., 3039
- Learned helplessness model of depression, 651, 652
- Learning. *See* Education; Knowledge; Sociology of education; Sociology of knowledge
- Learning theories. *See* Behaviorism; Social learning theory; Socialization
- Leary, M. R., 2514
- Leary, Timothy, 1975–1976, 1978
- Least interest, principle of, 1701
- Least Preferred Coworker scale, 1568
- Least-squares regression analysis, 3015–3016
- Leavitt, Harold, 1034
- Lebanon, 1866, 1867  
     sociodemographic profile, 2938
- Lebensraum* policy, 1933
- Le Bon, Gustave, 553, 559, 679, 1770, 2117, 2265
- Le Bras, Gabriel, 1425
- Lectures on Jurisprudence* (Smith), 2340
- Lectures on Sociology* (Levada), 2980
- Lederer, Emil, 1773
- Lee, Aie-Rie, 3223
- Lee, Alfred McClung, 326, 1247, 1250
- Lee, Barrett A., 1772
- Lee, Elizabeth McClung, 1247
- Lee Kuan Yew, 356
- Legal autonomy theory, 1555–1556
- Legal bureaucracy theory, 1497
- Legal formalism, 2960, 2961
- “Legal honoratores” (Weber concept), 1554
- Legal layering (transplant) concept, 465, 1550, 1556
- Legal profession. *See* Lawyers
- Legal realism, 476
- Legal systems. *See* Court systems and law; Law and legal systems
- Legionnaires’ disease, 814
- Legislation of morality, 1560, **1575–1581**, 2156–2157
- Legislative branch, 1952, 1955
- Legislative coalitions, 331
- Legitimacy. *See* Interpersonal power; Organizational structure
- Legitimate Violence Index, 985
- Lehman, Edward W., 360
- Lehman, Herbert, 2125
- Lehmann, E. L., 3035
- Leibniz, Gottfried, 1045
- Leicht, Kevin, 2166
- Leigh, Geoffrey K., 1729
- Leighton, Alexander, 178
- Leighton, Barry, 367
- Leisure, **1581–1591**  
     nature of, 1587–1589  
     tourism and, 3166–3173  
     work and time studies, 1582–1584, 3159–3160
- Lemert, Charles, 2200
- Lemert, Edwin, 669, 1496
- Lemkin, Raphaël, 1066, 1067, 1070, 1071
- Lenin, V. I., 1772, 1782, 1945, 2411, 2431, 2432, 2485, 2847  
     on capitalist imperialistic war, 3243  
     economic study of imperialism, 1265  
     and revised Marxist socialism, 2848–2849  
     and Soviet sociology, 2979  
     *See also* Marxism-Leninism
- Leningrad State University, Laboratory of Concrete Social Research, 2980, 2981
- Lenski, Gerhard E., 416, 1704, 1705, 1886, 2645, 2646, 2690  
     social stratification theory, 2813–2184, 2815

- societal stratification theory, 2866, 2867
- status incongruence theory, 3051, 3052
- Leo, Richard, 901
- Leonard, Kenneth E., 2557
- Leonardi, Franco, 1470
- Leopold II, king of Belgium, 60
- Lepenius, Wolf, 2757
- LePlay, Frédéric, 1025, 1503–1504, 3155
- Lepître, Alain, 2306
- Lepsius, Rainer, 1075, 1079, 1080
- Lerman, David, 194
- Lerner, Daniel, 1885
- Lesbianism. *See* Sexual orientation
- Lesotho, 2216
- Lesourne, Jacques, 1041
- Less developed nations. *See* Developing countries; Industrialization in less developed countries
- Lester, Marilyn, 3252
- Letourneau, Charles, 1424
- “Letter from a Birmingham Jail” (King), 2495
- Letters, life histories, 1633
- Lev, Daniel S., 472
- Levada, I. A., 2980
- Levada, Yuri, 2982, 2983
- Levels of analysis, **1591–1596**
- aggregate level vs. individual level effects, 1594
  - cross level inferences, 1594–1596
  - cumulative scale analysis, 1801–1802
  - deviance theories, 662–668
  - inferences from one level to another, 1592–1593
  - and levels of measurement, 1793, 1794–1796
  - of role theory, 2421–2425
  - types of variables, 1591–1592
  - typologies, 3180–3188
  - values scale, 3214–3215
- Levine, M. L., 677–678
- Levine, R., 3209
- LeVine, Robert A., 1900
- Levinson, Daniel, 2069
- Levinson, Randy, 198
- Levi-Strauss, Claude, 563, 990, 1034, 1271, 1273
- and exchange theory, 2670
  - and incest taboos, 1271, 1273
  - and kinship system, 1507
  - and structural analysis, 1027, 1032, 1033, 2891
- Levitical Code, 3283
- Levy, Frank, 3198–3199
- Levy, Jack S., 3241
- Levy, Judith, 583
- Levy, Marion, 2417
- Lévy-Bruhl, H., 1024
- Lévy-Bruhl, L., 1024
- Lewin, Kurt, 1012–1016, 2069, 2085, 2611, 2614, 2619
- and social psychology of status attainment, 2781, 2784
- Lewin Legacy, The* (Stivers and Wheelan), 1013
- Lewin-Epstein, Noah, 643
- Lewinsky, Monica, 2581
- Lewis, C. I., 2218
- Lewis, Helen Block, 784
- Lewis, Robert, 1508, 1726, 1728, 1778
- Lewis, Sinclair, 2369–2370
- Lewis, W. A., 2921–2922
- Lexical studies, personality trait measurement, 2079
- Liang, Jersey, 2304
- Liberal Democratic Party (Japan), 2131
- Liberal welfare state, 377
- Liberalism/conservatism, **1596–1604**
- common denominator of liberalism, 1597
  - communitarian influences on, 361
  - communitarianism contrasted with liberalism, 355, 356, 358, 359
  - communitarianism contrasted with social conservatism, 359
  - conservative ethos, 1598–1599, 1600
  - contemporary differences, 1601–1602
  - culture war and, 1580
  - and culture-of-poverty view, 2212
  - and environmental sociology, 802–803
  - and fundamentalism, 2371
  - legislation of morality, 1576–1579
  - liberal ethos, 1597, 1598
  - linguistic labels for, 1602
  - and mass society theory, 1770–1774
  - in Mexican history, 1856, 1857
  - national democracies vs. dictatorships, 3002–3003
  - and neoconservatism, 1600–1601, 1603, 1758
  - and neo-Marxism, 1758
  - nineteenth-century tenets, 1599–1600, 1602
  - and political correctness, 2139–2142
  - and political party system, 2155, 2159, 2160
  - and protest movements, 2267
  - reform and counterreform movements, 2717–2718, 2723
  - and religious organizations and orientations, 2379, 2386
  - secularization and, 2484
  - socioeconomic status and, 1602–1603
  - twentieth-century liberalism, 1600
  - views of human nature, 357, 1598
- Liberation theology, 937, 944, 2371
- Libertarianism, 526
- entitlement theory, 2699
  - political correctness and, 2140
- Library of Congress call number, 1612
- Library resources and services for sociology, **1604–1614**
- academic libraries, 1604–1605
  - data archives, 575–577, 580, 2474–2477
  - kinds of materials covered, 1612–1613
  - literature search strategies, 1607–1608, 1611–1613
  - on-line catalogs, 1611
  - periodical literature search strategies, 1608–1611
  - secondary data analysis, 2474–2481

- social psychology research  
and, 2769
- Libya, 1866
- sociodemographic profile, 2938
- Lichterman, Paul, 1648
- Lieberson, Stanley, 843, 1400, 1684,  
1685, 2819, 3039
- Liebig, Justus von, 1180
- Lieblich, Amia, 1636
- Liebow, Elliott, 243
- Liebrand, Wim B. G., 3221
- Liem Sioe Liong, 2131
- Life course, **1614–1622**
- accelerated, 7–9, 35
  - adolescence and, 2, 3–9,  
12–13, 2861
  - adulthood and, 25–35
  - aging and, 78–79, 80, 82–  
85, 1137
  - alcohol consumption rates, 95, 96
  - childbearing patterns and,  
125, 2032
  - childhood as distinct phase of, 4,  
122, 2861
  - clarification of concepts,  
1615–1617
  - cohort perspective and, 346–347,  
1618, 1619, 2861–2862
  - continuity and change in, 1586
  - in demographic models, 1625
  - depression risk factors, 653
  - divorce and, 82–83, 125–126
  - divorce's significance in, 700–701
  - early life decisions and, 83
  - education and, 20–21
  - effect of early trauma on, 82–83
  - effect of parental divorce on,  
82–83, 127–128
  - fallacies, 344–345, 1614
  - female fertility and, 2233
  - Finnish sociology on, 2453
  - health inequalities in, 1138
  - health practice and, 1129–1130,  
1136–1139
  - historical events and, 82,  
2861–2862
  - human agency and, 1620
  - individualization of, 7
  - intragenerational mobility over,  
2714–2715
  - juvenile delinquency and, 1489
  - labor-force participation data  
and, 1523
  - leisure and, 1582–1583,  
1585–1587
  - life cycle and, 1615–1616,  
1617–1618, 1625–1627
  - life history and, 1616
  - life-span study, 1616–1617
  - marital age and, 124, 130, 1742,  
1743, 1744–1745
  - marriage and, 1733
  - mental illness and, 1838–1839
  - mobility factors, 2714
  - occupational mobility and  
earnings, 1982–1994, 1983
  - paradigmatic principles, 1619–  
1621
  - personal biography and, 82
  - personal values and norms and,  
2837–2838
  - planful competence and, 12,  
13, 32
  - relationships across, 80
  - retirement and, 2401–2410
  - role transitions during, 2424
  - sexuality during, 2550–2556
  - as social creation, 1137
  - socialization across, 2860–2862
  - socioeconomic status and, 1138
  - stability across, 80
  - standardization of, 7, 27
  - structural lag and, 3060–3067
  - theory emergence, 1617–1619
  - timing patterns, 1619–1620
  - transitions in, 1615–1620, 1626,  
1627, 1635, 1733, 2714, 2861
  - and widowhood response,  
3257, 3258
  - work life during, 1614–1615
- Life cycle, **1623–1627**
- as analogy for rise and fall of  
whole societies, 2644
  - definitions of, 1615–1616, 1623
  - government regulation and, 1100
  - intergenerational resource  
transfer and, 1394
  - Jungian personality theory  
and, 1714
  - linked lives in, 1617–1618, 1620
  - and marital satisfaction pattern,  
1729, 1730, 1737–1738
  - occupational advancement and  
earnings, 1624–1625, 1983,  
1984, 1990
  - parenthood and, 2037–2038
  - as research tool, 1624–1625
  - rites, 937
  - structural lags and, 3061
  - widowhood and, 3256–3257
- Life endurancy, definition of,  
1631–1632
- Life expectancy, **1627–1633**
- calculation of, 223
  - changes over time, 1628–1629
  - comparative health-care systems,  
377–378, 1827
  - comparative statistics, 375
  - current, 1138
  - definition of, 1627–1628
  - differentials in, 1630–1631
  - divorce rate relationship, 126
  - as factor in American family  
changes, 122, 126, 131
  - gains in, 621, 622–623, 624, 627,  
628, 2177
  - gender and, 1165
  - and health and illness behaviors,  
1131, 1137–1138
  - historical increases in, 2176–2177
  - infant and child mortality  
and, 1335
  - Japan as world leader, 1631
  - life table, 614–617, 1628
  - long-term care needs and,  
1653, 1654
  - marriage duration and, 122
  - Middle Eastern countries, 1866
  - minorities and, 1169
  - Muslim countries (table of  
selected), 2938
  - pattern in, 624
  - population effects of, 2177
  - race and, 125
  - related concepts, 1631–1632
  - socioeconomic status and, 1138
  - Total Fertility Rate  
relationship, 628, 629
  - trends in, 223–224



- in United States, 114, 122, 126, 131, 196, 199, 1628–1631, 1827, 2180  
 widowhood and, 126  
 for women, 1058, 1151
- Life histories and narratives, **1633–1639**  
 as artistic endeavors, 2291–2293  
 case studies, 245, 247–248, 1633  
 comparative narrativists, 387–388  
 definition of, 1616, 1633  
 pragmatism and, 2222  
 sociocultural anthropology and, 2891  
 on voluntary associations, 3229
- Life insurance, 3257
- Life Satisfaction Index, 2304, 2306, 2541, 2683–2684
- Life span  
 definition of, 1631  
 study, 1616–1617
- Life stress paradigm. *See* Stress
- Life tables, 612, 614–616  
 definitions and interpretations of  
 functions of, 1629–1630  
 event history analysis and, 1790  
 hazards models, 616–617  
 life endurancy rates, 1631–1632  
 life expectancy and, 1628  
 population projection, 615–616
- Life-course fallacy, 344–345
- Lifelong learning. *See* Adult education
- Lifestyles and health, **1639–1643**  
 health promotion and, 1164  
 marital status and, 1750  
 sexual risk-taking and, 1641–1642, 2559  
 sexually transmitted diseases and, 2585–2593  
*See also* Alcohol; Alcoholism; Drug abuse; Health promotion and health status; Smoking
- Life-sustaining procedures, 585
- Lifton, Jay, 581–582
- Lifton, Robert, 895
- Light, Donald W., 376
- Light, Ivan, 848
- Lijphart, Arend, 1946, 2154, 2159
- Likert scales, 186, 1565, 2094, 2346
- Liking, 779–780
- Lilienfeld, Paul de, 1423
- Liljeström, Rita, 2453
- Liminality, 3281–3282
- Limits of the Criminal Sanction, The* (Packer), 2114
- “Limits to Growth, The” (King project), 1038–1039
- Limits to Growth, The* (Meadows and Meadows), 1041
- Lin, Nan, 2732, 2791, 2792
- Lincoln, J. R., 603
- Lincoln, James, 3275
- Lind, Georg, 1899
- Lindemann, Eric, 582
- Lindenberg, Siegwart, 1080
- Line graphs, 659
- Linear models  
 causal system, 3108–3111  
 decision-making theory, 592–593  
 elaboration and subgroup analysis, 3112–3114  
 hierarchical, 1173–1178  
 and nonlinear models, 1788–1789  
 probability, 2251–2252  
 structural, 1788, 1922
- Linear regression, 162–163, 2251  
 regression line, 447–450  
 sample selection bias, 2437
- Linguistics. *See* Language; Sociolinguistics
- Link, Bruce G., 2190
- Linked triads, 331
- Linton, Ralph, 2415, 2419
- Linz, Juan, 1941
- Lions Clubs International, 3229
- Lippmann, Walter, 2273
- Lipset, Seymour M., 103, 425, 605, 606, 3069  
 labor union study, 1533
- Latin American entrepreneurship analysis, 1537, 1538  
 as neoconservative, 1601  
 on political party origins, 2153–2154, 2154–2155  
 and political sociology, 2917
- Lipsitz, George, 2170, 2918
- Liquor. *See* Alcohol
- Liska, Allen E., 671, 2658, 2660
- Lisp-Stat* (computer software), 3020
- LISREL (computer software), 190, 266, 409, 1691, 1692, 1694, 1914, 3037, 3038
- List Servers (Listserv), 413, 414
- Literacy, 21, 2909  
 and fertility determinants, 1009, 2178
- Literary Digest*, 2273  
 presidential election poll, 3232
- Literature and society, **1644–1652**  
 on death and dying, 582  
 fundamentalist caricatures in, 2369–2370  
 gender theory and, 2171–2172  
 international approaches, 1649  
 life histories and narratives, 1635  
 personality trait diversity in, 2092  
 popular culture studies, 2168, 2169, 2171–2172  
 postmodernism and, 2207  
 pragmatism and, 2219  
 reception theory, 1648  
 reflection theory, 1644–1645, 1646, 2906  
 romantic love portrayals, 1699, 2171–2172  
 social incongruence as theme, 3053  
 sociological advances in study of, 1646–1648  
 as sociological reflection, 21, 1646, 2171–2172, 2906  
 stratification systems, 1648–1649  
 as structural reflection, 1645  
 utopias, 320105
- Lithuania, 2362
- Litigation. *See* Court systems and law; Law and legal system
- Little, Steven, 197
- Littleton (Colorado) school shooting, 1485, 1491
- Litwak, Eugene, 1389, 1503
- Liver cirrhosis, 93, 94, 1640
- Liverpool, Lord, 1771
- “Live-world” concept, 2756
- Livi, Livio, 1465
- Living and the Dead, The* (Warner), 364

INDEX

- Living Standards Measurement Study, 549
- Living wills, 585–586, 587, 3064, 3083
- Llewellyn, Karl, 1549–1550
- Lloyd, David, 656
- Lobbying. *See* Interest groups
- Locke, Harvey J., 1502, 1727
- Locke, John, 268, 355, 527, 528, 1247, 1248, 1302, 2337
- Lockhart, W. R., 3188
- Lockheed company, 2131
- Lockwood, David, 225–226
- Lodahl, Thomas, 3273
- Lofland, John, 554
- Loftin, Colin, 2284
- Log normal curve, 2870
- Logan, John R., 2501, 3072
- Logarithms, 3015
- Loges, William E., 3214
- Logic models, 2296–2297  
deductive vs. inductive, 2465–2466
- Logic of Collective Action, The* (Olson), 2149, 2920
- Logical inconsistency, cognitive dissonance theory vs., 338
- Logical positivism, 2193, 2756
- Logique sociale* (Tarde), 1025
- Logistic regression, 454–455
- Logit analysis, 3038
- Log-linear models, 1970, 2817, 3036–3037, 3107  
design vectors, 3116  
tabular analysis, 3115–3118, 3122–3126
- Lois de l'imitation* (Tarde), 1025
- Lok Dal (Indian political party), 90
- Lolita* (Nabokov), 1648
- Lomax, John, 1921, 1926
- Lombroso, Cesare, 502, 528, 529, 1465, 1575, 1717
- London, Harvey, 2078
- London Co-Operative Magazine*, 2846
- London New Police, 2110, 2112–2112
- London School of Economics, 225, 226, 228
- Loneliness, 2525, 3255
- Long, Elizabeth, 1647–1648
- Long, Norman, 2432–2433
- Longevity  
health and illness behavior and, 1131  
of successive cohorts, 345
- Longitudinal fallacies, 1593  
on life course, 1614
- Longitudinal research, **1683–1696**  
on aging in single cohort, 343–344  
and causal inference models, 1685, 1688–1689  
data analysis approaches, 1691–1694  
data analysis problems, 1689–1694  
data types, 1685–1686  
design of, 1687–1688  
divorce effects on children, 706  
divorce effects on women, 708  
of educational attainment, 2783–2784  
hierarchical linear models, 1177  
life course, 1614, 1617, 1618  
measurement protocols, 1687–1688, 1690–1691  
method artifacts in, 2351–2352  
multiple-indicator measurement, 1922  
need for childhood sexual abuse studies, 293  
outcome prediction, 1691  
population definition, 1687  
quasi-experimental and descriptive approaches, 1686  
rationale for, 1684–1685  
retirement models, 584, 2405–2406  
sample consistency, 1687  
secondary analysis and data archives, 2475–2476  
social surveys, 577  
and stability, 2343  
time series, 3142–3153
- Long-term care, **1652–1663**  
comparative systems, 375, 378, 1655, 1659–1661  
definitions of, 1652  
funding of, 1658–1661, 1663, 1668–1669  
health provider organizations and, 1826  
insurance for, 1657–1660  
models of, 165–167  
needs assessment, 1653, 1654  
place of residence, 1655  
policy and practice issues, 1661–1662  
quality issues, 1661  
respite care and, 1658  
services and providers. *See* Long-term care facilities  
utilization and expenditures, 1148–1149
- Long-term care facilities, **1663–1683**  
American families and, 129  
assisted living, 1663, 1826  
community-based, 1656–1657  
continuing care retirement communities, 1664  
death and dying in, 584, 586  
depression of residents of, 656  
home-based, 1656–1658, 1661–1662, 1663, 1677  
as medical-industrial complex component, 1818, 1819, 1826  
nursing homes, 1148, 1149, 1653–1654, 1661, 1664–1678, 1819  
residential homes, 1654–1655  
sheltered housing, 1655–1656
- Looking Backward* (Bellamy), 3203
- “Looking-glass self” concept, 2089, 2344, 2507, 2512, 2750, 2856
- Lopata, Helen, 2417, 3255
- Lopez Portillo, José, 2135
- Lopreato, Joseph, 1467, 2881, 2883, 2884, 2886
- Lord, Charles G., 3222
- Loria, Achille, 1423, 1464, 1465
- Los Angeles, California  
Iranian immigrants in, 1872  
as megacity, 312  
Mexican immigration to, 1860  
Muslim immigrants in, 2950
- Loseke, Donileen, 247
- “Lost causes,” 2615

- “Lost letter” technique, 186
- Louis Dirn (French sociologist group), 1028–1029
- Louis Harris Data Center, University of North Carolina, Chapel Hill, 2477
- Louis Philippe, king of France, 1771
- Louis XIV, king of France, 3242
- Louisiana, civil law code, 465, 1513, 2127
- Love, **1696–1701**  
 altruism and, 114  
 courtship and, 484, 488, 489  
 cultural differences in, 1697–1698  
 emotions and, 778–780  
 family violence and, 984  
 gender differences and, 1697, 1700–1701  
 historical conceptions of, 1697  
 interpersonal attraction and, 1449–1450  
 marital adjustment and, 1730  
 marital redefinition and, 705  
 mate selection and, 1775, 1778  
 parental role and, 2034  
 personal dependency and, 2065  
 person-centered theory and, 1715  
 physical symptoms of, 1701  
 romantic love complex, 1698–1700, 2171–2172  
 universal aspects of, 1698
- Love Canal, 789, 791, 801, 805, 2875
- Lovejoy, Arthur S., 1599
- Low, J. O., 364
- Low-birthweight babies, 2236
- Lowen, Alexander, 2085
- Lowenthal, Leo, 539, 565, 1075, 1645
- Lower quartile, definition of, 659
- “Lower-class value stretch” concept, 2212
- Lowie, Robert, 675, 2890
- Low-skill workers. *See* Marginal employment
- LSMS. *See* Living Standards Measurement Survey
- LTC. *See* Long-term care
- Lubbock, John, 1423
- Luce, R. Duncan, 2729
- Luckenbill, David F., 407
- Luckmann, Thomas, 226, 2756, 2957–2958
- Luddites, 2525
- Luhmann, Niklas, 1080, 1234, 1479, 1548, 1558, 1560, 1704, 2484
- Luisson, Lucio, 328
- Lukács, Georg, 541, 1645, 2756
- Lukes, Steven, 2165
- Lumumba, Patrice, 66
- Lunch counter sit-ins, 2494, 2495
- Lund, Paul S., 364
- Lundeen Bill, 2402–2403
- Lung cancer, 1639, 1640, 1641
- Lusinchi, Jaime, 2135
- Luther, Martin, 3227, 3286
- Lutheranism, 95
- Lutyński, Jan, 2121
- Lutz, Burkhardt, 1075
- Luvox, 654
- Luxembourg  
 social surveys, 577  
 transnational corporations, 3175
- Lyall, Kathryn, 2282
- Lyman, Stanford, 2220
- Lynchings, 349, 2266
- Lynd, Helen, 363–364, 367, 2521
- Lynd, Robert, 363–364, 367, 1160
- Lynott, Robert J., 2304
- Lytotard, Jean-François, 2206, 2207, 2648, 2757
- M**
- Ma, Hing-Keung, 1901
- Maastricht Treaty (1992), 1935
- Macari, Daniel, 1391
- MacBride, Sean, 1767
- MacBride Commission, 1767
- McCarthy, Joseph, 1238–1239
- Macchiavelli, Niccolo, 1464
- Maccoby, E. E., 2530–2531
- Macedonia, 2362
- Macfarlane, Alan, 1504
- Mach, Ernst, 821, 2192–2193
- Mach, Zdzislaw, 2119–2120
- Machada, Bernardino, 1423
- Machiavelli, Niccolò, 1077, 1564, 2520, 2623
- Machine politics, 2125–2126, 2135
- Machinery. *See* Industrialization; Technology and society
- MacIver, Robert, 1424, 2521  
 as Japanese sociology early influence, 1478
- Macke, Anne Stratham, 2417, 2423
- MacKenzie, Donald A., 2461
- Mackie, Diane M., 2244
- MacKinnon, Catharine, 274, 276, 2186–2187
- Macklin, Eleanor D., 113
- MacLeod, Jay, 665–666, 2171
- Macpherson, C. B., 1302
- Macroeconomics, 747
- Macrosociology, **1703–1712**  
 collective behavior theories, 352–354  
 comparative legal systems, 1547–1549, 1550  
 comparative-historical inquiry, 390  
 criminal and delinquent subcultures theories, 513–514  
 cross-cultural research, 547  
 deviance origins, 662, 663–666  
 deviance reactions, 670–671  
 distributional inequality, 2690–2691  
 economic impact of education, 2934  
 of emotions, 787  
 four paradigms of, 2027  
 future of, 1710  
 historical background, 1704–1705  
 institutional anomie theory, 166  
 intergenerational relations, 1386–1388  
 of law and society, 1552–1561  
 longitudinal research, 1684  
 microsociology vs., 1703–1705, 2671  
 and music, 1924–1925, 1927  
 research methodology, 1709–1710  
 risk research, 2877–2880  
 Scandinavian sociology, 2452  
 social control studies, 2660  
 and social movements emergence, 2718–2719

- and social networks, 2727–2728  
 and social psychology studies, 2777–2778  
 theory and research themes, 1705–1709  
 units of analysis, 1703–1704  
*See also* Evolution: biological, social, and cultural;  
 Functionalism and structuralism;  
 Phenomenology; Structuralism
- MAD. *See* Mean absolute deviation; Mutually assured destruction
- Madagascar, 2216
- MADD (Mothers Against Drunk Driving), 2722, 2877
- MADD (multiattribute dynamic decision model), 1015
- Madison, James, 721, 2265
- Madsen, Richard, 2080, 2484
- Mafia, 2018–2019, 2128, 2129, 2130
- Magatti, Mauro, 1472
- Magnet, Myron, 3198
- Magnitude effect (meta-analysis), 1848–1849
- Maguire, Patricia, 2039, 2040
- Mahar, David, 1227
- Maharaj Ji, 3287
- Maharidge, Dale, 1580
- Mahathir Bin Mohamad, 356
- Mahavira, 3285
- Mahdi, Ali Akbar, 1872
- Mahoney, F. I., 1653
- Mahoney, James, 1636
- Mahrishi Mahesh Yogi, 3287
- Mail surveys, 3091
- Mailing lists, Internet, 413, 414
- Main Trends in the Social and Human Sciences* (UNESCO), 2917
- Maine, Henry, 1501–1502, 1506  
 on legal systems development, 1545–1546
- Maines, David, 1636, 2220
- Mainframe computers, 575
- Mainichi Press, 1482
- Mainstream criminology, 504
- Maintenance production (MP) behaviors, 1566
- Mair, Lucy, 289
- Maistre, Joseph de, 1770
- Majone, Giandomenico, 1098, 1100, 1103, 1104, 1107, 1108
- Major, John, 2130
- Major depression. *See* Clinical depression
- Major personality theories, **1712–1719**  
*See also* Personality and social structure; Personality theories
- Majority. *See* Compliance and conformity
- Majors, Richard, 245
- Mäkelä, Klaus, 2451
- Making Markets* (Abolafia), 738–739
- Making of a Counter Culture, The* (Roszak), 459
- Malawai, 2591
- Malaysia, 356, 2974, 2975, 2978  
 authoritarian  
 communitarianism, 356  
 fertility decline, 627, 2976  
 labor movement, 1532  
 sociodemographic profile, 2938
- Malcolm X, 58
- Male gaze, 2172
- Malewski, Andrzej, 2119, 2120
- Mali, 2133  
 sociodemographic profile, 2938
- Mali Empire, 2999
- Malinowski, Bronislaw, 563, 853, 1030, 1258, 1259, 1271, 2118–2119, 2373, 2855, 2889, 2890, 2892  
 on core myths and rituals, 3282–3283  
 and exchange theory, 2670
- Malnutrition, 622
- Malthus, Thomas R., 632–633, 928, 1525
- Malthusian marriage system, 1504
- Malthusian theory, 632–633, 1005, 1008, 1525  
 human ecology and environmental analysis, 1219–1220
- Maltoni, Bruno, 2231
- Man and Society in Calamity* (Sorokin), 681
- Managed-care organizations  
 elite paradigm and, 2628  
 failures of, 1828  
 financial status and profits, 1822–1826  
 as medical-industrial complex component, 1818, 1819–1822  
 utilization and expenditures, 1143–1145  
*See also* HMOs
- Management science, 3104
- Managerial revolution thesis, 2628
- Managers, 442, 443  
 leaders vs., 1564  
 meritocracy and, 2626  
 retirement patterns, 2406  
 and white-collar crime, 3252
- Mandela, Nelson, 2047
- Manic-depressive disorder. *See* Bipolar disorder
- Mann, Michael, 224, 1704, 1707, 2823
- Mann, Richard, 1978
- Manners and Customs of the Police* (Black), 2114
- Mannheim, Karl, 582, 583, 857, 2273  
 and intelligence theory, 1386–1388  
 as Japanese sociology early influence, 1478  
 and mass society theory, 1773  
 and sociology of knowledge, 2953–2954, 2955–2956, 2957, 2958, 2959  
 on student movements, 3068, 3069
- Manning, Peter, 2114
- Mann-Whitney test, 1960
- Mann-Whitney-Wilcoxon test, 1963–1964
- MANOVA. *See* Multivariate Analysis of Variance
- Mansbridge, Jane A., 118
- Mansour, Ali H., 643
- Manuel de sociologie* (Cuvillier), 1024
- MAO inhibitors, 654
- Mao Zedong, 298, 300, 2136, 2485, 2643, 2849
- MAP (Multiple Analysis Project), 2903

- Marable, Manning, 54
- March Current Population Surveys, 971
- Marchus, Philip H., 1533
- Marcia, J. E., 1900
- Marcos, Ferdinand, 645, 2131
- Marcuse, Herbert, 539, 540, 541, 1075, 1305, 1772, 1773
- Marcus-Newhall, A., 73
- Mare, R. D., 2439
- Marey, E. J., 3005
- Marginal employment, **1719–1725**  
 definition of, 1719  
 distribution by selected characteristics (table of), 1721  
 dynamics of, 1722  
 and job displacement, 1721–1723  
 as underemployment, 1720–1722
- Marginal frequency, 658
- Marginal utility analysis, 733, 2698
- Marginalization, 2367, 2634–2635
- Marienthal: The Sociography of an Unemployed Community* (Jahoda et al.), 3155
- Marijuana, 711  
 decriminalization efforts, 712, 713, 718  
 legalization movement, 523  
 stereotypes, 713
- Marijuana Tax Act of 1937, 711
- Marini, Margaret A., 1689
- Maris, Ronald W., 3079, 3080
- Marital adjustment, **1725–1733**  
 age of first marriage and, 1744  
 cohabitation effect on, 486  
 consequences of, 1731–1732  
 factors in, 1726–1728, 1735–1738, 2035  
 marital quality and, 1726, 1736–1738  
 marital satisfaction vs., 1727  
 parenthood and, 2037  
 prediction factors, 1728–1731  
 and violence, 984  
*See also* Divorce; Marital quality
- Marital Adjustment Test, 1727
- Marital cohesion, definition of, 1726
- Marital Comparison Level Index, 1728
- Marital happiness, definition of, 1727
- Marital quality, 1736–1738  
 definition of, 1726  
 remarriage and, 2390  
 and sexual satisfaction, 2540
- Marital rape, 950, 2577–2578
- Marital satisfaction  
 communication and, 1735–1736  
 definition of, 1726–1727
- Marital stability, 962
- Market organization  
 and economic institutions, 725–727  
 and economic sociology, 732  
 labor-force participation and, 1522  
*See also* Capitalism
- Market research  
 adoption of new products, 678  
 survey data collection, 575
- Marketing and Sociology of Books Group (Netherlands), 1649
- Markets and Hierarchies* (Williamson), 735
- Marketti, James, 55
- Markiewicz, Wladyslaw, 2119
- Markoff, John, 2414
- Markoulis, Diomedes, 1901
- Markov chain, 1692, 1789, 1790, 1983, 3039
- Markovsky, B. D., 2672, 2673, 2674
- Markus, H. R., 2513
- Marmor, Theodore, 2281
- Marriage, **1733–1740**  
 as adulthood marker, 26  
 adjustment to. *See* Marital adjustment  
 in African-American slave communities, 121–122  
 age at first marriage, 124, 130, 1525, 1526, 1737–1739  
 age distribution model, 620  
 alliance vs. descent conflicts in, 1508–1509  
 alternative lifestyles effects on, 107, 108, 113  
 American family trends, 123–126, 1525–1526  
 arranged, 1698, 1775  
 balanced reciprocity in, 1508  
 changing trends in, 487–488, 1525  
 childbearing demographics, 125, 1006, 2031–2032  
 childbearing effects on, 2035, 2037  
 childless rates, 2035  
 cohabitation prior to, 108–109, 705, 923  
 collateral prohibitions, 1272, 1273, 1509, 1513, 1776  
 in colonial America, 121  
 common law, 948  
 communitarian vs. religious conservative view of, 359–360  
 consanguineous, 1273  
 couple unity vs. sibling group unity, 1509  
 courtship, 483–489  
 decline in rates of, 1259  
 delayed, 1526, 1738, 1744, 2178, 2182  
 demographic hazards model, 616–617  
 differences between first and remarriages, 2390–2393  
 division of labor in, 696  
 divorce predictions, 701, 1738  
 dual-earner couples, 127, 1524, 1525–1526, 3062  
 and economic determinism, 723  
 elderly spousal caregiving, 1657  
 endogamous, 1270–1271, 1272, 1273, 1776, 1779  
 exogamous, 1271, 1272, 1776  
 extramarital sex and, 2541–2545  
 and family and household structure, 922  
 and family law, 947, 948–949  
 and family planning, 635, 952–960  
 and family policy in Western societies, 962  
 and family violence, 983  
 fertility transition and, 625, 1008, 2178, 2182  
 first marriage rates by age, 1744–1745

- functions of, 1258  
gender/mental illness  
  relationship, 1837–1838  
higher expectations for, 702,  
  703–704  
incest proscriptions, 1270–1273  
as institution, 1733–1734  
interfaith, 911, 1411, 1776  
interracial. *See* Intermarriage  
Japanese-American “picture  
  brides,” 176  
and kinship and family systems,  
  1507–1517  
kinship prohibitions, 1270–1273,  
  1509, 1513, 1776  
labor-force participation and,  
  1524, 1525–1526  
legitimacy and, 1258  
as life-cycle transition, 1615,  
  1616, 1623  
love and, 1697, 1698, 1699,  
  1775, 1778  
Malthusian system, 1504  
mate selection theories,  
  1774–1779  
median duration before  
  divorce, 1747  
monogamy in, 1271  
population relationship to, 2176  
as process, 1733, 1735–1738  
and quality of life, 2303  
rates, 1970–1990 (table of), 1742  
redefinitions of, 705  
as rite or ritual, 1733, 1734–1735  
same-sex, 111, 131, 315, 489,  
  1506, 1776, 2546  
satisfaction in, 1726–1728,  
  1735–1738  
sexual behavior in, 2531–2541  
shifts in normative, 635,  
  700–701, 702  
social psychology studies of, 2770  
spousal bereavement as death  
  “cause,” 584  
time use research, 3161,  
  3162, 3163  
traditional distribution of  
  power in, 1734  
typologies of long-lasting, 1731  
variations in, 1258  
wife’s roles in, 1736  
and women’s property rights, 702  
women’s role conflict in, 2417  
*See also* Divorce; Family and  
  household structure;  
  Intermarriage; Remarriage;  
  Widowhood  
Marriage and divorce rates,  
  **1740–1751**  
  alternative life styles and, 112  
  American patterns, 125–126,  
  1738, 1741–1749  
  crude rates, 1741  
  in developed countries, 1749  
  divorce rate predictions, 708,  
  1738  
  divorce rates by age, 1742, 1743,  
  1745–1747  
  divorce rates in remarriages, 2393  
  divorce rates in selected  
  countries, 706  
  divorce rates in United States,  
  701, 702, 703, 704, 705, 1742,  
  1743, 1745–1747  
  divorce rates with minor  
  children, 1749  
  first marriage rates, 1742, 1743,  
  1744–1745  
  interracial and interfaith marriage  
  upswing, 1776  
  marriage rate decline, 1738  
  marriage rate upswing, 1741  
  1950s patterns, 124  
  refined divorce rate, 1746–1747  
  refined marriage rates, 1741  
  remarriage, 708, 2387–2388  
  remarriage rates by age, 1742,  
  1743, 1748  
  selected variables and, 1750  
  significance of divorce, 700–  
  701, 709  
  table of divorce rates  
  (1970–1990), 1742  
  table of first marriage rates  
  (1970–1990), 1742  
  table of remarriage rates  
  (1970–1990), 1742  
  total marriage rates, 1748–1749  
*See also* Divorce; Marriage;  
  Remarriage  
Marriage market theory, 1775–1777  
  and women’s remarriage  
  potential, 2388–2389  
Marriage Registration Area (MRA),  
  1743, 1744, 1747  
Marriage squeeze, definition of,  
  1775–1776  
Marro, Antonio, 1  
Marsden, George, 2369  
Marsden, Peter, 2792  
Marsella, A. J., 1718  
Marselli, Gilberto Antonio, 1467  
Marsh, H. W., 2512  
Marshall, Alfred, 1423, 2335  
Marshall, John, 136, 476  
Marshall, Susan E., 646  
Marshall, T. H., 2210  
Marshall, Thomas, 2277  
Marshall, V. W., 581, 582, 583  
Marshello, Alfred F. J., 3221  
Martelli, Stefano, 1473  
Martin, Alfred von, 1075  
Martin, David, 2485  
Martin, Teresa Castro, 2393  
Martin, W. T., 3079  
Martindale, Don, 2217, 3181  
Martineau, Harriett, 852  
Martinelli, Alberto, 1468  
Martinez, O. J., 1935  
Martinotti, Guido, 1468  
Martinson, Robert, 2056  
Martinussen, Willy, 2452  
Marty, Martin E., 940, 2372  
Marwell, Gerald, 604  
Marx, Karl, 822, 1028, 1304, 2069,  
  2193, 2640, 2889  
  alienation concept, 100, 104,  
  1705, 3270  
  on bureaucracy, 229  
  on capitalism. *See under*  
  Capitalism  
  class theory, 757, 2623, 2692,  
  2812, 2814, 2819, 2927  
  class theory vs. classical elite  
  theory, 2623  
  and comparative historical  
  analysis, 383, 1196  
  on conflict sources, 425, 517,  
  2865, 3243

- data collection by, 573  
on division of labor, 697–698, 1754, 1782  
economic theory, 722–723, 724  
and ethnography, 852  
and Frankfurt School, 540  
and Hegelian philosophy, 1248–1249  
and historical materialism, 543, 1217, 1704, 1751, 1752, 1781–1782, 1784, 2644, 2645, 2646, 2647  
and industrialization, 2196  
labor theory of value, 1754, 2697, 2698  
on leadership stemming from *zeitgeist*, 1564  
and legal theory, 1548, 1553, 1576, 1577, 2961  
as macrosociologist, 1704–1705  
Marxist sociology emergence and, 1753  
mass society theorists' revisions of, 1772–1773  
and modern materialism, 1781–1782  
on money's significance, 1888, 1889, 1890  
postmodernist rejection of, 2206  
on religion, 2373, 2385, 2483, 2965, 2967, 2968  
on revolution, 2338, 2410–2411, 2412, 2413, 2414  
on science and technology, 1785  
on social inequality, 2690, 2692  
and socialist ideal, 2846, 2847  
on societal organization, 773–774, 2865, 3066  
as sociology of knowledge antecedent, 2953, 2954, 2955, 2957  
and status incongruence, 3050  
structural theory of state of, 2162, 2163  
*See also* Marxism-Leninism; Marxist sociology; Neo-Marxist theory  
Marxism. *See* Marx, Karl; Marxism-Leninism; Marxist sociology  
Marxism-Leninism, 1751, 1782, 2717, 2848–2849  
and destratification, 2813–2814  
imperialism theory, 1265, 1266  
on imperialist wars, 3243  
and political correctness, 2139  
postmodernist rejection of, 2206  
and Soviet sociology, 2116  
Marxist sociology, **1751–1760**  
ambiguity of term, 1751–1752  
British sociology and, 226  
and capitalism, 238, 531  
and case studies, 245–246  
class analysis in, 1753, 1754–1755  
on class and race interrelationship, 319  
on class struggle, 415, 601, 697  
*See also* Class struggle  
and community, 362  
and conflict theory, 414–416, 1401, 2767  
contemporary themes in, 1754–1755, 1757–1758  
and court systems and law, 467  
and criminal and delinquent subcultures, 511–512  
and criminology, 504–505, 534–536  
crisis of, 1755–1756  
and critical theory, 539–545, 1754, 2760  
and cultural theory, 562, 568, 1755, 1756, 1757, 1758  
current key shifts in, 1756–1759  
on data collection, 603  
decline of, 1756–1758  
delinquency and deviance theory, 511–512, 670, 1498  
on democracy, 601–602  
and dependency theory, 639–640, 644, 1087, 1088  
on division of labor, 697–698, 1754, 1782  
and economic determinism, 722–723  
and economic sociology, 736, 806  
and ethnicity, 179, 845  
and family and religion, 937  
and feminist theory, 989, 990, 1754–1755  
Frankfurt School and, 540, 543, 1076–1077, 1732, 2756  
French sociology and, 1027  
German sociology and, 1074, 1076, 1076–1077, 1078, 1080, 1082  
and globalization, 1084–1086, 1757, 1758  
on historical development stages, 2645  
Japanese sociology and, 1479  
Latin American studies and, 1538, 1539  
legal systems theory, 1548, 1553, 1576, 1577  
leisure critique, 1583  
literary, 1645, 1649  
macro themes, 1704  
macro-level reactions to deviance, 670  
and Marxism as “science” or “critique” approaches, 1752  
and modernization theory, 1084–1086  
origins of, 1752–1754  
and participatory research, 2040  
Polish sociology and, 2118, 2119, 2121  
on postindustrialism/postmodernism, 1757–1758, 2200, 2205  
on progress, 2644, 2645  
and radical criminological theory, 504–505  
on religious orientation, 2385  
on rural society, 2426, 2429  
Scandinavian sociology and, 2452  
scientific applied research and, 2460–2461  
on social change, 2647  
social problems paradigm, 2760  
social stratification emphasis of, 2162–2163  
and Soviet and post-Soviet sociology, 2116, 2979  
Soviet bloc disintegration effects on, 1757  
stratification theory, 2814–2815  
structuralism and, 1030, 1753, 1754, 1755, 1756, 1784, 2162–2163  
on war incentives, 3243, 3244

- and world system theory, 1197, 1706  
*See also* Historical materialism; Marx, Karl; Marxism-Leninism
- Masaryk, Thomas, 1423
- Masculinity. *See* Femininity/masculinity
- Masini, Eleonora, 1473
- Maslow, Abraham, 1014, 2085, 2087, 2088, 3271, 3286
- Mass culture  
 Frankfurt school studies of, 2168, 2169  
 high culture debate, 565–566, 1645–1646  
 power elite theory vs., 2624, 2625  
 vs. popular culture, 173  
*See also* Popular culture
- Mass media  
 and agenda control, 2166  
 censorship/regulation of expression issues and, 271–272, 280  
 corruption investigations, 2127–2128  
 definition of, 1761  
 as disaster communication, 684, 685  
 drug-abuse prevention programs, 716  
 and education, 760  
 Frankfurt School's conformity theory and, 540–541, 545  
 fundamentalists' utilization of, 2372  
 and innovation diffusion, 678  
 and legislation of morality, 1577, 1578–1579  
 mass society theory and, 1772–1773  
 materialist analysis and, 1785  
 new technologies, 1768–1769  
 and normalization of homosexuality, 2571  
 and political process, 145, 1764–1766  
 and popular culture, 2168, 2170, 2171–2172, 2173  
 and postmodern society, 2200, 2206  
 and public opinion, 2273, 2276  
 romantic love portrayals, 1698–1700  
 and sexually explicit material, 2185–2186, 2187  
 and social movements, 2722  
 as socialization agent, 2858  
 and sociology of art, 173  
 and sociology of culture, 563, 566, 568–569  
 sporting events coverage, 2986  
 and status aspirations, 2781  
 violence effects issue, 1762–1763, 2858  
*See also* Internet; Television
- Mass media research, **1761–1770**  
 on aggregate effects, 1763–1766  
 on aggression cues, 74, 75, 76, 272  
 on audience responses, 1761–1765  
 on Internet potential, 1768–1769  
 on persuasion, 2094, 2097–2098  
 on stereotypical portrayals, 1767–1768  
 Swedish sociology and, 2453  
 on time use, 3156, 3160  
 on voting behavior, 3234  
*See also* Mass media; Television
- Mass murder. *See* Genocide
- Mass society, **1770–1774**, 2428
- Massachusetts, first statewide probation law, 2253
- Massé, Pierre, 1038
- Massey, Douglas, 366, 2500, 2501, 2502–2503, 2504, 3199
- Masterman, Margaret, 2026
- Masters, William H., 2554
- Masters and Johnson sex survey, 111
- Masters of Polish Sociology* (Sztompka), 2120
- Masturbation, 2567–2568
- Matching theory, 214
- Mate selection theories, **1774–1780**  
 complementarity and, 1777  
 cross-cultural analysis, 1885  
 eligibility pool, 484–485, 1698, 1775–1776  
 gay and lesbian couples, 2546  
 incest taboos, 1270–1274, 1509, 1513, 1776  
 and kinship systems, 484, 1509, 1513, 1698, 1776  
 levels of courtship, 485–486  
 process theory of, 1777–1778  
 propinquity and, 1776–1777, 1779  
 and remarriage, 1779, 2388–2390  
 romantic love complex and, 1698–1699  
 and sexual behavior, 2537  
 sociobiological theory of, 1274, 2884–2885, 2886  
*See also* Courtship; Interpersonal attraction
- Materialism, **1780–1786**  
 atomist philosophy and, 1780–1781  
 decline of, 1784  
 deviance and, 663–664  
 and globalization, 1784  
 historical vs. dialectical, 1781–1782  
*See also* Historical materialism
- Maternal deaths, 2236
- Maternity leave, 2033–2034, 3266
- Mathematical achievement  
 sex differences in, 2532–2533  
 women's mathematical careers and, 2786
- Mathematical Methods of Statistics* (Crámer), 3035
- Mathematical sociology, **1786–1792**  
 intergenerational mobility modelity, 1983  
 models, 2028, 2029, 2296–2298  
 probability theory, 2248–2252  
 and social networks, 2728, 2729  
 and social structure analysis, 2826–2827  
 statistical methods, 3003–3039  
 structural role theory and, 2417  
*See also* Statistical graphics; Statistical methods
- Mathematical Statistics* (Wilks), 3035
- Mathiesen, Thomas, 2453
- Matrimonial Causes Act of 1857, 948
- Matrix algebra, 1789



INDEX

- 
- Matrix notation  
     and factor analysis, 911–912  
     scatterplot, 3018, 3020
- Matrix, scatterplot, 3018
- Matsueda, Ross L., 533
- Matsumoto, Jyun'ichiro, 1478
- “Matthew effect” (Merton concept), 2691
- Maturana, Humberto R., 1557, 2088
- Maturational development theory, 1894–1895
- Matza, David, 1496
- Mauritania, 1866  
     slavery and slave-like practices, 2602, 2604, 2605  
     sociodemographic profile, 2938
- Mauro, Paolo, 2124
- Maurya dynasty, 2999
- Mauss, Armand L., 2763, 2764
- Mauss, Marcel, 734, 1024, 1032, 1273, 2670
- Max Planck Institute (Berlin), 1617
- Maximillian, emperor of Mexico, 1856–1857
- Maximization principle, 2881
- May, M. A., 114–115, 2083
- May, Rollo, 2085
- Maya Empire, 2999
- Mayer, Thomas, 1983
- Mayhew, L. H., 693
- Mayne, Sir Richard, 2110, 2113
- Mayntz, Renate, 1075, 1076
- Mayton, Daniel M., 3213
- Mazur, Joanna, 2122
- MBTI (Myers-Briggs Type Indicator), 2076, 2079
- McCall, Michael, 1633
- McCarran-Walter Act of 1952, 175
- McCarthy, E. Doyle, 2958, 2959
- McCarthy, John D., 1941, 1942
- McCarthy, P. J., 1692
- McCarthy era, 314, 316, 1238–1239, 2278
- McClelland, David, 1885–1886, 3271
- McClintock, Barbara, 994
- McClintock, Charles G., 3221, 3222, 3223
- McCloskey, Donald, 2767
- McCrae, Robert R., 2087
- McDonagh, Edward, 326
- McDonald, Dwight, 565–566
- McDonaldization, 2208–2209, 2529
- McDonaldization of Society, The* (Ritzer), 2208–2209
- McDougall, William, 114, 2085, 2087
- McDowell, David M., 2284
- McDowell, J. J., 214
- McGann, Anthony J., 2160
- McGill, Robert M., 3009–3010
- McGillis, Daniel, 193
- McGinnis, Robert, 3034
- McGinty, Patrick, 2222
- McGrath, Joseph, 180, 1979, 2610
- McGregor, Douglas, 1014, 3271
- McHale, John, 1037, 1039
- McIntire, Carl, 2370
- McKay, Claude, 66
- McKay, Henry, 665, 1495
- McKelvey, Richard D., 3038
- McKenzie, Roderick, 1209, 1210
- McKinley, John, 1815
- McKinney, John C., 3181–3182, 3188
- McKinney, William, 2484
- McKinney Act of 1987, 1203, 1206, 1207
- McLemore, Clinton, 1976
- McLennan, John, 1270, 1271
- McLuhan, Marshall, 428, 1769
- McMahan, Eva, 1636
- McMaster Health Index Questionnaire, 2306
- McMichael, Philip, 387, 2433
- McMillan, David, 363
- McMillen, Curtis, 195
- McPhail, Clark, 352, 555, 559, 560
- McPhee, W. N., 410
- Mead, George Herbert, 355, 781, 1254, 1303, 1313, 1423, 2069, 2417, 2630  
     on altruism, 114, 117–118  
     personality theory, 2085, 2089, 2611  
     and pragmatism, 1249, 1250, 2218, 2219, 2220, 2222, 2423, 2955  
     role theory, 2415, 2507–2508
- self-esteem concept, 2344, 2507–2508, 2512, 2750  
     and social psychology of status attainment, 2781  
     and symbolic interactionism, 2856, 3095, 3098  
     time use research, 3155
- Mead, Lawrence M., 3198
- Mead, Margaret, 1, 563, 998, 1003, 2418, 2855, 2890, 2892
- Mean (statistical), 661  
     computation of, 659  
     sampling distribution of, 3028–3029  
     standard error of, 3029  
     weighted mean effect size (in meta-analysis), 1848–1849
- Mean absolute deviation (MAD), 564  
     definition of, 660–661
- Meaning  
     existential personality theory and, 1717, 2084  
     in symbolic interactionism, 3096–3097
- Means, Gardiner, 443
- Means of production, 1782
- Means over time, 1691–1692
- Measurement, **1792–1804**  
     composite scale, 1909  
     consequences of less than normally distributed variables, 1796–1800  
     construction of measures, 1801–1802  
     of education and development, 748–750  
     error as excluded variable, 2251–2252  
     errors in causal inference models, 256–257, 264–266, 1909, 1917  
     errors in correlations among indicators, 1917  
     errors in longitudinal research, 1690–1691  
     of global self-esteem, 2512  
     of height, 2343–2344  
     of infant and child mortality, 1324–1325  
     of informal economy, 1337–1343, 1341–1343
-

- and information collection, 1802–1803
- of intelligence, 1359–1361
- interval in, 1792–1800
- longitudinal research protocols, 1687–1688
- of marital adjustment, 1727–1728
- models, 2027–2028
- of moral judgment, 1897–1900
- multiple indicator models, 1907–1923
- of nonrandom error, 1917–1920
- of personality, 2073–2081
- of poverty, 2213–2214
- power and Type I and Type II errors, 3030–3032, 3033
- precision of variables, 1803
- protocols of, 2344–2347
- of quality of life, 2302–2304, 2683–2684
- random error, 1908–1909
- reliability, 2343–2355
- of retirement, 2406
- sampling procedure, 2250, 2444–2449, 3088
- in scientific explanation, 2467
- of self-esteem, 2344–2350, 2512
- of social mobility, 2712–2713
- of social values, 3219–3222
- of societal stratification, 2870–2874
- standard error, 2449
- standardization, 2991–2996, 3034–3035
- statistical models, 3003–3039
- of urban underclass, 3198–3200
- validity, 3207–3211
- of values, 3214–3223
- See also* Analysis of variance and covariance; Factor analysis; Levels of analysis; Quasi-experimental research designs; Statistical analysis; Survey research
- Measures of association, **1804–1812**
- ambiguity of meaning, 1809
- applications of, 1811–1812
- bivariate relationships, 661
- correlation and regression analysis, 446–457
- degree of, 1806–1811
- multiple indicators, 1907–1908
- and nonparametric statistics, 1966–1969
- and relative reduction in prediction error, 1810–1811
- tabular analysis, 3118–3126
- two major types of, 1966
- Measures of Occupational Attitudes and Occupational Characteristics* (Robinson et al.), 3209
- Measures of Political Attitudes* (Robinson et al.), 3209
- Measures of Social Psychological Attitudes* (Robinson and Shaver), 3209
- Mecca, 2939, 3282
- Mechanic, David, 376, 1814
- Mechanical solidarity (Durkheim concept), 165
- Medellín cartel, 2135
- Media. *See* Mass media; Mass media research; Popular culture; Television
- Median (statistical), 661
- computation of, 659
- Mediation, 465, 1115
- Medicaid, 79, 130, 1352, 2799
- and contraceptive service, 958
- and filial responsibility, 1019
- Health Care Financing Administration, 588, 815, 1157, 1670, 1828
- inception of, 2798
- income distribution and, 1284, 1286
- and medical-industrial complex, 1814, 1824–1825
- and nursing home care, 1663, 1668, 1669–1670, 1819
- utilization and expenditures, 1140, 1145, 1146–1147, 1148, 1149, 1150
- Medical Industrial Complex, The* (Wohl), 1818
- Medical profession. *See* Nurses; Physicians; Medical-industrial complex
- Medical sociology, **1813–1818**
- areas of study, 1814, 1816
- and bioethics, 585
- and criminal sanctions, 521
- critical theory in, 1815–1816
- and death and dying, 582–583, 585
- definition of field, 1813
- and health and illness behavior, 1813
- and health policy analysis, 1158, 1813–1818
- on health system changes, 1814–1815
- and healthy life expectancy, 1632
- and lifestyles and health, 1639–1642
- and medical-industrial complex, 1813, 1814–1815, 1828–1829
- quality-of-life studies 23011, 2305–2306
- sociologists *of* vs. sociologists *in* medicine, 1815
- stress studies and, 3055–3057
- Medical supply companies, 1818, 1823, 1824
- Medical-industrial complex, **1818–1832**
- criticism of, 1827–1829
- diversification, restructuring, and growth, 1821–1822
- financial status and profits, 1822–1824, 1827–1828
- gender theory on, 1816
- government financing, 1814
- and health promotion and status, 1163
- introduction of concept, 1818
- medical sociology and, 1813, 1814–1815, 1828–1829
- and postindustrial society, 2197
- professional occupations, 2259, 2260
- ranking in global context, 1826–1829
- regulatory needs, 1824–1826
- structure of, 1818–1822
- See also* Health care utilization and expenditures
- Medicare, 1352, 1663, 1814
- added coverage, 2803–2884
- and eldercare, 79, 1021, 1828

- Health Care Financing  
 Administration, 588, 815,  
 1157, 1670, 1828  
 inception of, 2798  
 and income distribution, 1284  
 and medical-industrial complex,  
 1818, 1824–1825  
 and nursing home care, 1663,  
 1668–1669, 1669–1670, 1819  
 public policy analysis, 2281  
 utilization and expenditures,  
 1140, 1141, 1144, 1145–  
 1146, 1148
- Medications  
 abortion-induction, 2238–2239  
 AIDS/HIV, 2590–2591, 2593  
 antidepressants, 654–655  
 comparative systems of  
 prescription drugs, 378  
 costs of, 1141, 1148  
 diffusion of new antibiotic,  
 677–678  
 for drug abuse, 712, 715  
 legal use of mood-altering  
 prescription drugs, 718  
 pharmaceutical companies, 1818,  
 1824, 1827, 2263  
 psychopharmacological, 1840–  
 1841  
*See also* Drug abuse
- Medici family, 2298
- Medieval era. *See* Middle Ages
- Meech Lake amendment, 3008, 3009
- Meertens, Roel W., 2244, 2245
- Megachurch, 2379
- Megacities, 310–311, 312, 3197
- Megan's Law, 317, 2582
- Mehrabian, Albert, 1978
- Meiji restoration (1868), 2411
- Melby, Jeffrey, 197
- Melotti, Umberto, 1472
- Melucci, Alberto, 354
- Memorial Societies of America, 588
- Memory  
 cognitive consistency theories  
 and, 334  
 life histories and narratives,  
 1636–1637
- Memphis, Tennessee, 2126
- Menarche, 2233
- Mencius, 1564
- Mencken, H. L., 2369–2370
- Mendelian genetics, 2880, 2881
- Mendras, H., 1026
- Menger, Carl, 733, 1423
- Menger, Pierre-Michel, 1925
- Menninger, Karl, 3077
- Menopause, 2233
- Mental health  
 medical sociology studies, 1814  
 positive, 2188–2191  
 religion and, 2965  
 self-esteem and, 2511  
 women's employment and,  
 1837–1838
- Mental illness and mental disorders,  
**1832–1843**  
 classification and diagnosis of,  
 1832–1833, 1834–1835, 1840  
 definitions of, 1832–1833  
 deinstitutionalization and,  
 1205, 1841  
 as divorce effect, 707  
 epidemiology of, 1833–1840  
 gender and, 1838–1839  
 homelessness and, 1204  
 illicit drug use and, 711  
 institutionalization and, 2660–  
 2661  
 insurance coverage, 1150–1151  
 labeling theory and, 669,  
 1836–1837, 1838, 1840  
 life course and, 1839  
 long-term care services, 1656  
 micro-level reaction theories, 669  
 psychopharmacological  
 treatment, 1840–1841  
 rural/urban incidence of,  
 1840, 1841  
 social causation theories, 1835  
 social control and, 521, 1841,  
 2660–2661  
 social integration studies, 3055  
 social selection and drift  
 theory, 1836  
 stigmatization of, 1815  
 suicide predictors, 3078–3079,  
 3081
- See also* Depression
- Menzel, Herbert, 677, 1591
- Mercury poisoning, 2724–2725
- Meriam Report (1928), 135
- Meritocracy, 756–757, 758,  
 2626–2627  
 and educational mobility, 2927  
 and intergenerational mobility,  
 2712  
 and social stratification, 2812  
 in tribal societies, 2809–2810  
*See also* Affirmative action;  
 Equality of opportunity
- Merryman, John Henry, 475, 476,  
 478, 479, 480  
 legal tradition concept,  
 1547–1548
- Merton, Robert, 244, 342–343, 2193,  
 2678, 2755–2756, 2915  
 accumulative advantage  
 concept, 2691  
 and analytical paradigms, 2023  
 and anomie theory, 165–166,  
 533, 1493–1494  
 on bureaucratic structure, 233  
 collectivity definition, 2632  
 on cultural conformity vs. social  
 belonging, 2630  
 on functional analysis, 1030  
 on goals and means  
 disjunction, 3066  
 and macro-level deviance  
 theory, 663  
 and “marginal man” concept,  
 2634, 2635  
 and medical sociology, 1813  
 role theory, 2415, 2634, 2635  
 on scientific ethos, 2456  
 on social forecasting, 2676  
 and social inequality concept,  
 2690  
 and social philosophy, 2757–2758  
 on sociology of knowledge, 2956  
 structural analysis, 2825–2826,  
 3066, 3099  
 “theories of the middle range,”  
 2825–2826  
 and time use study, 3155
- Merton, Thomas, 3289

INDEX

- MESA. *See* Middle East Studies Association of North America
- Mesomorphy, 1717, 1718
- Messedaglia, Angelo, 1465
- Messianic movements, 2367, 2968–2969
- Messick, David M., 3222, 3223
- Messick, S., 3211
- Messner, Steven, 503, 671
- Mestizo* culture (Mexico), 1857, 1858
- Meta-analysis, **1843–1852**  
 definition of, 1843  
 drug abuse programs, 718  
 of four “established” sex differences, 2531–2534  
 human sociobiology, 2880  
 narrative reviewing and, 1844–1845  
 replication and, 2397  
 reviews of evidence, 1845–1849  
 seven steps in, 1845–1849  
 trends in practice of, 1849–1850
- Metaforecasting, 2678
- Metatheory, **1852–1855**  
 Marx’s historical materialism as, 1782  
 Polish sociology and, 2120  
 two types of, 1852–1853  
 underlying constituents of, 1852
- Methadone, 712, 715
- Method of agreement/method of difference, 386
- Methodological anarchism (Feyerabend concept), 823
- Methodological individualism, 1027
- Methotrexate, 2238, 2239
- Metropolis algorithm, 3039
- “Metropolis and Mental Life, The” (Simmel), 1772
- Metropolitan Statistical Areas (MSAs), 307, 2478, 3194–3195, 3196  
 mobility patterns, 1416  
 suburbanization and, 3070–3071, 3072–3073  
 world comparison, 3194
- Mexican Americans, 636, 1860  
 family structure, 123  
 household structure, 127
- Mexican Revolution of 1910, 1856, 1857, 1858, 2411, 2414
- Mexican studies, **1855–1864**  
 age pyramid, 610, 611  
 age-standardized crude death rate, 611, 613  
 AIDS/HIV incidence, 2592  
 clinical sociology, 328  
 demographic characteristics, 1535, 1536  
 divorce rate, 706  
 drug crop control program, 713–714  
 drug trafficking, 2135  
 economic liberalization, 1539  
 family size, 977  
 fertility decline, 627, 2178  
 gender role changes, 941, 1861  
 health-care system, 375, 380–381  
 immigrant discrimination, 58  
 indigenous empires, 2999  
 labor movement, 1531, 1532  
 life table, 612  
 low suicide rate, 3079  
 Mexican history, 1855–1858  
 political and social tradition, 1536  
 political corruption, 2135  
 protest movement, 2271  
 U.S. border tensions, 1936, 1937
- Mexican tourism, 3169
- Mexican-American War, 1856
- Mexico. *See* Mexican studies
- Mexico City, Mexico  
 as megacity, 312  
 population size, 1859, 1860, 2179
- Meyer, Adolf, 2085
- Meyer, John W., 427–428, 738, 1085, 1086
- Meyer, Katherine, 2379
- Meyer, Leonard B., 1925
- Mezzogiorno, 1466–1467
- Mianamata disease, 2724–2725
- Michael, Robert T., 2539, 2550, 3091
- Michaels, Stuart, 2539, 3091
- Michels, Robert, 3054  
 on development of oligarchies, 229, 603, 606, 1533, 2163, 2377, 2623–2624, 2627
- and German sociology, 1074  
 and Italian sociology, 1464, 1465  
 and political sociology, 2162, 2163
- Michaelson, William, 309
- Michigan, right-to-die issue, 587
- Michigan Panel Study of Income Dynamics, 2480
- Michigan school of voting behavior, 3233, 3234–3236, 3237
- Michigan State University, 681
- MicroCase (instructional package), 411
- Microeconomics, 725, 2670, 2671
- Microprocessor, 1345
- Microsociology  
 crime theories, 662, 663, 666–670  
 of institutions, 2340  
 macrosociology vs., 1703–1705
- Mid Atlantic Medical Services, 1822
- Middle Ages  
 Arab-Muslim conquests, 2940  
 boundaries, 1932  
 civil law system, 473  
 corporate organization, 441–442  
 marriage and kinship systems, 1516–1517  
 religion and heresies, 2968–2969  
 university development, 1179
- Middle class  
 African American, 2492, 2493, 2496, 2497, 2498, 2499  
 alienation among, 103  
 and downward mobility, 3050  
 industrial sociology and, 1309  
 and new religions, 2380  
 and race relations, 321–322  
*See also* Bourgeoisie; Social class; Social stratification
- Middle East peace accords (1974, 1978), 2048
- Middle East Studies Association of North America, 1868–1869
- Middle Eastern Diaspora Communities in America* (Bozorgmehr and Feldman eds.), 1872
- Middle Eastern studies, **1864–1875**  
 ancient city-states and empires, 2998

- countries covered, 1865
- Demographic and Health Surveys, 633
- fertility transition, 628, 1008, 1867
- Islamic fundamentalism, 2371
- modernization theory, 1885
- “new Arab social order,” 1867–1868
- pan-Arab nationalism, 1944
- political corruption, 2132
- slavery and involuntary servitude, 2601–2602, 2603, 2604–2608
- terrorism, 3137, 3139
- theoretical and methodological contributions, 1870–1873
- women in labor force percentage, 3262
- See also* Islamic societies; Sociology of Islam; *specific countries*
- Middletown* (Lynd and Lynd), 363–364
- Middletown in Transition* (Lynd and Lynd), 363–364, 367
- Middletown studies (Lynd and Lynd), 363–364
- Middletown III study (Caplow, Bahr, and Chadwick), 364
- Midlife crisis, 2861
- Midnight Cowboy* (film), 2185
- Midtown Manhattan study, 1834
- Midwest Prevention Project, 716
- Midwives, 375, 378
- abortion services, 2240
- Mifepristone (RU 486), 2238–2239
- Migrant labor, 636, 1338, 1436, 1858, 1865, 2496, 2498, 2608
- Migration. *See* Internal migration; International migration
- Mikhailovsky, Nikolai, 2979
- Milbank Quarterly*, 1815
- Milbrath, L. W., 2302
- Miles, Catherine Cox, 999
- Miles, Robert, 226
- Milgram, Stanley, 404, 838
- Miliband, Ralph, 1755
- Miliband-Poulantzas debate, 1755
- Military sociology, **1875–1883**
- and all-volunteer force, 1877–1881, 1882
- and American defense establishment, 144
- and American police uniform, 2112–2113
- and democracy, 605
- and dictatorships, 3002
- and dynamics of revolutions, 416
- and supranational organizations, 3003
- and widow status, 3255
- See also* War
- Mill, John Stuart, 268, 818, 988, 990, 1577
- as Japanese sociology early influence, 1477
- and liberalism, 355, 1600
- method of agreement/method of difference, 386
- on subjection of women, 2692
- on women’s wage gap, 371
- Millenarianism, 2368
- Millennium Project, 1039
- Miller, Arthur, 103, 104
- Miller, Austin, 54
- Miller, D., 3209
- Miller, David, 559, 560
- Miller, G. Tyler Jr., 1219
- Miller, Gale, 247
- Miller, Joan, 194, 198
- Miller, Neal, 73, 2084, 2120, 2670
- Miller, S. M., 1601–1602
- Miller, Walter, 513, 534
- Miller, Warren E., 3234, 3235
- Miller v. California* (1973), 274, 2186
- Millet, Kate, 2576
- Milling, 350
- Mills, C. Wright, 821, 859, 1027, 1635, 1753, 1772, 2069, 2205, 2980, 3054
- and humanistic sociology, 1250
- “power elite” theory, 2162, 2624–2626, 2627
- and sociology of knowledge, 2954
- Milton, John, 268
- Milward, A., 1934–1935
- Min, Pyong Gap, 182
- Minard, Charles Joseph, 3005
- Mincy, Ronald B., 3199
- Mind, Self, and Society* (Mead), 3095
- Ming Empire, 3000
- Mingione, Enzo, 1338
- Minnesota, 3068
- Minnesota Conference (1967), 581, 582, 585
- Minnesota Medical School, 582
- Minnesota Multiphasic Personality Inventory (MMPI), 2076, 2078
- Depression Scale, 654
- Minorities
- comparable worth, 369–372
- cultural innovation
- transmission by, 2170
- in legal profession, 468
- rights ideology, 1945
- See also* Discrimination; Ethnicity; Race; *specific groups*
- Minority Fellowship Program (ASA), 156
- Minority (within group)
- influence, 400, 403
- Minority parties, 2157
- Mintz, Beth, 443
- Mirowsky, John, 1834–1835
- Mirsepassi, Ali, 1870
- Mischel, Walter, 2074, 2080, 2083
- Mishler, Elliot, 1636, 1816
- Mishra, Joya, 2166
- Missing data, 3039
- Misumi, J., 1566
- Mitnick, Barry, 1098, 1099, 1100
- Mitroff, Ian, 2456
- Mitterrand, François, 2129
- Mixed-motive games, 2337
- Miyamoto, Frank, 179, 180, 181
- Mizokawa, Donald, 194
- MKULTRA (CIA brainwashing program), 898
- MLLSA (software), 3038
- MMPI. *See* Minnesota Multiphasic Personality Inventory
- MOA (monoamine oxidase) inhibitors, 654
- Moaddel, Mansoor, 1871
- Mobility (geographical). *See* Internal migration; International migration

- Mobility (social). *See* Education and mobility; Occupational and career mobility; Social mobility; Social stratification
- Mobutu Sese Seko, 2133
- Modarres, Ali, 1872
- Mode (statistical), 661  
     computation of, 659
- Model T automobiles, 699
- Models. *See* Paradigms and models
- Models of Man* (Simon), 3035
- Moderate diffusionism, 675
- Modern Sociological Theory in Continuity and Change* (Becker and Boskoff eds.), 2980
- Modern synthesis theory, 2880
- Modern World System, The* (Wallerstein), 1197
- Moderne Kapitalismus, Der* (Sombart), 733
- Modernity (philosophical), 1884  
     fundamentalist reaction and, 2368–2369, 2370  
     postmodernist theorists on, 2648  
     and sociology of knowledge, 2954–2955
- Modernization theory, **1883–1888**  
     contributors to, 1885, 1886  
     convergence theory and, 422–423  
     critical theory and, 543–544  
     critiques of, 1886  
     and cumulative social changes, 2645–2647  
     dependency theory vs., 639–646, 1706–1707  
     ecology and, 1211  
     education and, 746–747, 754  
     factors in modernization, 2154  
     family structure and, 1389, 1501–1507  
     fundamentalism and, 2372  
     globalizing nature of, 1084–1086, 1886–1887  
     historical sociology and, 1199  
     influences on, 1885  
     Islamic society and, 2943–2944, 2946  
     Japanese sociology and, 1478  
     Latin American studies and, 1537–1538, 1539
- macrosociology and, 1705–1707
- major explicit tenets of, 1884
- Marxist sociology and, 1084–1086, 1758
- migration and, 1418
- nationalism and, 1940–1941, 1946–1948
- and political party system origins, 2154–2155
- political sociology and, 2162
- pragmatism and, 2219
- religion and, 934–935, 942, 2966
- rural sociology and, 2429–2430, 2431–2432
- secularization vs. sacralization and, 2487–2488, 2489, 2966
- social and political elites and, 2626–2627
- and values theory, 3222–3223
- vs. development, 1884
- vs. industrialization, 1884
- vs. philosophical concepts of modernity and postmodernity, 1884
- See also* Postindustrial society; Postmodernism
- Moe, Terry, 1100, 1101, 1106
- Moerman, Michael, 433
- Moghaddam, Fathali, 196
- Mohammed. *See* Muhammad
- Mohammed and Charlemagne* (Pirenne), 2940
- Mohr, John, 1199, 1649
- Mokrzycki, Edmund, 2120
- Mola, Fortuny Loret de, 941
- Moldova, 1934
- Molina Enríquez, Andrés, 1857
- Molitor, Grahm, 2677
- Molm, L. D., 2673
- Molotch, Harvey, 3252
- Momomorphic genes, 2089
- Money, **1888–1894**  
     and evolution theory, 1885  
     “new money”/happiness correlation, 2684  
     *See also* Wealth
- Money market, 1885, 1889–1890
- Mongol Empire, 1070, 2998
- Monitoring the Future survey, 717, 2784
- “Monkey Trial” (Scopes trial), 2369, 2370
- Monoamines, 652, 654
- Monopoly capitalism, 1754
- Monothetic-polythetic distinction, 3183–3185
- Monson, Ingrid, 1925
- Monte Carlo techniques (Markov chain), 3039
- Montenegro, 2362
- Montesquieu, 1024, 1031, 2520  
     legal systems comparison, 1545
- Montgomery, James, 2419
- Montgomery (Alabama) bus boycott (1955), 2269, 2493, 2494
- Monthly Catalog of United States Government Publications*, 1613
- Monthly Review* (journal), 1754
- Mood  
     and decision processing, 593  
     depression, 648–656  
     marital happiness and, 1727  
     persuasive attempts and, 2096
- Mood-modifying substances. *See* Drug abuse
- Moon, Sun Myung, 2366  
     *See also* Unification Church
- Mooney, James, 133
- Moore, Barrington, 386, 605, 1886, 2154  
     historical sociology study, 1198, 2918  
     and macrosociology, 1704, 1706, 1707
- Moore, Erin, 1901
- Moore, Wilbert E., 423, 1030, 2676–2677, 2690, 2813
- Moore, William, 1352
- Moot legal system, 464
- Mora, José María Luis, 1856
- Moral Commonwealth, The* (Selznick), 356
- Moral development, **1894–1906**  
     in adults, 1900, 1903–1904  
     altruism and, 114–118  
     cognitive stages doctrine, 1895–1897  
     communitarian view of, 357–358, 359, 422

- and criminal transgression, 519–520
- criticisms and further research, 1900–1904
- cultural influences on, 1901
- family and religion and, 943
- gender bias argument, 1900, 1902–1903, 2089
- idealism and, 1248
- measurement of, 1897–1900
- religious orientation and, 2384
- and social justice, 2700–2701
- theoretical foundation of, 1894–1895
- See also* Ethics; Ethics in social research; Morality
- Moral Economy of the Peasant: Subsistence and Rebellion in Southeast Asia* (Scott), 2977
- Moral infrastructure, 357–358, 359
- Moral judgment. *See* Moral development
- Moral Judgment Interview (MRJ), 1897–1898
- Moral Judgment Test (MJT), 1899
- Moral Majority, 462, 1580, 2371, 2717, 2719, 2723
- Moral panics. *See* Legislation of morality
- Moral relativism, and ethics in social research, 840
- Moral sociology, 1024
- Morale et la science des moeurs, La*, 1024
- Morality
- as aggression justification, 74
  - communitarian obligation viewpoint, 355
  - definition of, 1894
  - legislation of, 1575–1581
  - male-biased double standard, 701
  - postmodernity and, 2207, 2208
  - See also* Moral development
- Morality of care, 1900, 1902–1903
- Morawska, Ewa, 569
- Morcellini, Mario, 1473
- More, Sir Thomas, 54, 3201–3202
- Moreno, Jacob, 1027, 1034, 1977, 2415, 2425
- on group structure, 1611
  - on social networks, 2728–2729
  - sociometric research approach, 2613
- Morgan, Barrie S., 2502
- Morgan, Lewis Henry, 1271, 1272, 2889
- Morgan, Robin, 2187
- Morgenstern, Oskar, 591, 1049
- n*-person game theory, 329, 2335, 3221
- Morin, Edgar, 824
- Morioka, Kiyomi, 1480
- Mormon Freeman, 462
- Mormonism
- affinity of religion and family in, 936
  - low alcohol consumption rates, 95
  - as new religious movement, 2366, 3287
- Moroccans (ethnic), 692
- Morocco, 1865, 1866, 1867
- fertility decline, 628
  - Islamic society study, 2944
  - political corruption, 2132
  - slavery and slave-like practices, 2604, 2605
  - sociodemographic profile, 2938
- Morphine, 713
- Morra, Gianfranco, 1470
- Morrill Acts of 1862 and 1890, 1180
- Morris, Albert, 3245
- Morris, Aldon D., 2270
- Morris, Charles, 2218
- Morris, David G., 2053
- Morris, Michael, 2211
- Morris, William, 3203–3204
- Morselli, Enrico, 1465
- Mortality. *See* Death and dying; Life expectancy
- Mortality modeling, 619
- Mortality rates. *See* Birth and death rates; Infant and child mortality
- Mortality transitions, 621–623, 633–634
- as fertility transition factor, 624, 625, 627, 628, 629, 2178
  - historical markers in, 2176–2177
  - in Middle East, 1867
  - population projections and, 2181–2183
  - rate decline, 2177–2178
- Morton Thiokol Corporation, 3253
- Mosaic law, 2840
- Mosca, Gaetano, 603, 1464, 1465, 2163, 2164, 2622–2623, 2626, 2690
- Moscovici, Serge, 197, 350, 400, 403
- Moscow University, 2980
- Moses, 3053, 3280, 3283
- Moskos, Charles, 1876
- Moskowitz, W., 1653
- Moslem law. *See* Islamic law
- Moslems. *See* Islamic societies; Sociology of Islam
- Mosteller, Frederick, 3036
- Mother-father-child triad, 331
- Motherhood
- child custody and, 702
  - divorce effects on, 706, 707–708
  - emotional depression and, 655
  - role of, 2036
  - unmarried, 125, 128, 484, 488, 634, 1258–1264, 1506, 1626, 1744, 2033
  - See also* Childbearing; Parental roles; Pregnancy and pregnancy termination; Single-parent households
- Mothers Against Drunk Driving (MADD), 2722, 2877
- Motivation
- “audience effect” and, 2615
  - behavioral, 1128
  - control research, 2060–2061
  - and high vs. low self-esteem, 2515
  - intrinsic and extrinsic, 2059–2060
  - leadership and, 1568–1569
  - personal autonomy and, 2059–2060
  - rational choice theory and, 2338–2339, 2340–2341
  - social comparison process and, 2651–2652
  - and social values research, 3221–3222
  - and value types, 3216–3217, 3223

- values vs., 2829  
work orientation and, 3269–3270, 3270–3271
- Motor vehicle accidents. *See*  
Drunken driving
- Motor vehicle theft  
rate calculation, 497–498  
Uniform Crime Reports  
definition, 492  
victimization rates, 499
- Motorcycle gangs, 460
- Mottura, Giovanni, 1467
- Mouffe, Chantal, 545
- Mourning. *See* Bereavement
- Movies. *See* Film industry
- Mowrer, O. Hobart, 73, 2087
- Moynihan, Daniel Patrick, 847, 2281, 2333, 2818
- M-plus model, 3039
- MRA. *See* Marriage Registration Area
- Mucha, Janusz, 2119
- Muckrakers, 2124
- Muhammad (Mohammed), 2937, 2939, 3053, 3280, 3281, 3285, 3286
- Muir, John, 802
- Muir, William, Jr., 2114
- Mukherjee, Ramkrishna, 1290, 1291, 1292, 1293, 2301–2302, 2307
- Mukherji, D. P., 1291
- Mulhall, M. G., 3005–3006, 3007, 3008
- Mulkay, Michael, 2459
- Muller, John H., 1927
- Müller, Karl Valentin, 1076
- Muller, Kate, 1927
- Mullet, Etienne, 1903–1904
- Mullins, Nicholas, 1012
- Multiattribute dynamic decision model (MADD), 1015
- Multi-Attribute Utility Analysis, 2618
- Multiculturalism  
accommodation and, 1402  
ethnicity and, 843, 2049  
leadership and, 1570  
in Mexico, 1861  
Moral Majority view of, 1580  
political correctness and, 2139–2140  
and status incongruence, 3054
- Multidisciplinary studies. *See*  
Sociology among the Social Sciences
- Multihospital corporations, 1819–1820, 1821
- Multilingualism, 2909
- Multinational Comparative Time Budget Research Project, 3156
- Multinational corporations. *See*  
Transnational corporations
- Multinational legal system, 1550
- Multinational organizations, 3003  
peacekeeping and, 1429, 2046, 2048  
*See also* International Court of Justice; International Monetary Fund; United Nations; World Bank
- Multiparty system, 2154
- Multiphasic Personality Inventory  
Depression Scale (MMPI-D), 653
- Multiple Analysis Project (MAP), 2903
- Multiple factor analysis, 905–906
- Multiple Factor Analysis*  
(Thurstone), 3035
- Multiple indicator models, **1907–1924**  
definition of, 1907, 1912  
and factor analysis, 908, 1920–1921  
and good measurement protocols, 2345  
with nonrandom measurement error, 1917–1920  
overidentified, 191  
of quality of life, 2303–2304, 2306, 2307  
reliability of, 1908–1909, 1921  
and single-indicator models, 1910–1912  
standardization and, 2991–2996  
strengths and weaknesses of, 1920–1923  
validity of, 1908, 1909–1910, 1917, 1921
- Multiple personalities, 2087–2088
- Multiple personality disorder, 901
- Multiple regression analysis, 452–456, 1812  
nonadditive effects, 456  
statistical methods, 3035–3036
- Multiple time series design, 2318–2319, 3152
- Multiple-influence model of leadership, 1568
- Multiple-linkage model of leadership, 1568
- Multiplism, critical (evaluation research concept), 866
- Multistage sampling, 2448, 2449
- Multi-step flow, 670
- Multisystemic Therapy, 76, 1917
- Multitrait-multimethod matrix, 3210
- Multiuniversity, 1180
- Multivariate analysis of variance, 1922  
models for, 2028  
statistical methods, 3003–3039, 3108
- Multivocality, 2206
- Multiway tables, 3036
- Mulvey, E. P., 76
- Munch, Edvard, 3079
- Muncie, Indiana (Middletown studies), 363
- Muqiddimah* (Ibn Khaldun), 1564
- Muraoka, Miles Y., 2301, 2305–2306
- Murder. *See* Homicide
- Murdock, George P., 548, 1502, 1506, 2893
- Murphy, Charles F., 2125
- Murphy, John M., 1903
- Murphy, Lois, 114
- Murray, Charles, 968, 1382, 2213, 2330, 3198
- Murray, Henry, 2085
- Murstein, Bernard, 1775, 1778
- Muscat, 2602
- Musée social, Le* (monograph collection), 1025
- Music, **1924–1930**  
African and African-American, 64–65  
innovations and transmission of, 2170  
notation, 3005  
punk rock counterculture, 460, 461



- song lyric analysis, 1927  
*Muslim Society* (Gellner), 2943  
 Muslims. *See* Islamic society;  
     Sociology of Islam  
 Musolf, Gil Richard, 2221  
 Mussolini, Benito, 1310, 1357, 3002  
 Mutapa Empire, 2999  
 Muthén, Bengt O., 3039  
 Mutran, Elizabeth, 1692  
 Mutually assured destruction,  
     3138, 3242  
 Myanmar. *See* Burma  
 Myers, Charles A., 422  
 Myers, Martha, 670  
 Myers-Briggs Type Indicator,  
     2076, 2079  
 Myrdal, Gunnar, 2451, 2692  
 Mysticism, 2968–2969  
     countercultures, 460–461, 2969,  
         3287–3288  
     religious movements, 3286  
 Mythology and ritual  
     and core religious myths,  
         3280–3283, 3284, 3285  
     and French School of  
         Sociology, 1025  
     and French structuralism, 1032
- N**
- Nabokav, Vladimir, 1648  
 NAC. *See* Native American Church  
 Nadel, S. F., 1034, 2417  
 Nader, Laura, 1549  
 NAE. *See* National Association of  
     Evangelicals  
 Naficy, Hamid, 1872  
 NAFTA (North American Free  
     Trade Agreement), 794, 2433  
     labor movement opposition, 1531  
     Mexican participation, 1859, 1861  
 Nagai, Michio, 1480  
 Nagel, Stuart, 2280  
 Naimark, Norman M., 1948  
 Naive art, 173  
 “Naive” psychology, 192  
 Nako, K., 1998  
 Nalewajko, Ewa, 2121  
 Namboodiri, Krishnan, 1213  
 Nanck, Guru, 3287–3288  
 Napoléon, emperor of France,  
     1179, 3242  
 Napoleonic code (*Code Napoléon*),  
     474, 475, 1513  
 Napoleonic Wars, 1933  
 Narcissism, aggression and, 72  
 Narcotics. *See* Drug abuse  
 Narcotics Anonymous, 715  
 Narrative reviewing, 1844–1845,  
     2291, 2902  
 Narrativists. *See* Life histories and  
     narratives  
 NASA. *See* National Aeronautics and  
     Space Administration  
 Nash equilibrium, 330  
 NASS. *See* North American Society  
     for the Sociology of Sport  
 Nassau County (New York) political  
     machine, 2125  
 NAT theory, 2878  
 Nathan, Laura E., 2478  
*Nation at Risk, A* (education  
     report), 769  
 Nation of Islam, 2950  
 National Abortion Federation, 2240  
 National Academy of Sciences,  
     1226, 2461  
     Disaster Research Group,  
         681, 686  
 National Action Party (PAN;  
     Mexico), 2135  
 National Advisory Commission on  
     Civil Disorders, 2270  
 National Advisory Council for  
     Health Care Policy, Research,  
     and Evaluation, 2399  
 National Aeronautics and Space  
     Administration, 2682, 2878, 3253  
 National AIDS Behavioral Survey,  
     2541, 2550  
 National Association of Evangelicals,  
     2370, 2380  
 National Association of Social  
     Workers, 2842  
 National Audubon Society, 780, 802  
 National border relations, **1931–**  
     **1939**  
     and borderless Europe,  
         1934–1937  
     cross-border crime and  
         trafficking, 1935–1936  
     four interaction types, 1935  
     and guest workers, 2608  
     historical boundary setting,  
         1933–1934, 1939, 2356  
     nationalism and, 1944–1945  
     self-determination doctrine  
         and, 1945  
     transnational corporations  
         and, 3178  
     *See also* Nationalism  
 National Cancer Institute, 2399,  
     2687  
 National Center for Health Statistics,  
     701, 816, 970, 1157  
     marriage and divorce rates, 1743,  
         1744, 1748, 1749, 1750  
 National Center for Juvenile  
     Justice, 1489  
 National Center for Prevention and  
     Control of Rape, 2576, 2577  
 National Center for Social Research  
     on Alcohol and Drugs, 2451  
 National Center of Scientific  
     Research, 1026  
 National Center on Child Abuse and  
     Neglect, 2581  
 National Commission on Excellence  
     in Education, 769  
 National Commission on the Causes  
     and Prevention of Violence in the  
     United States, 2269  
 National Committee on Pay  
     Equity, 372  
 National Conference of Charities  
     and Corrections, 2841  
 National Congress of American  
     Indians, 137, 1298  
 National Council for Crime  
     Prevention (Sweden), 2451  
 National Council of Churches,  
     2370, 2380  
 National Council of Senior Citizens,  
     2148–2149  
 National Council on Measurement  
     in Education, 3207  
 National Council on the Aging, 1395  
 National Crime and Victimization  
     Survey, 491–500, 530, 982  
 National Crime Survey, 1488–  
     1489, 2475

INDEX

- and social indicators, 2682  
on rape, 2576
- National differences. *See*  
Comparative-historical sociology;  
Cross-cultural analysis
- National Election Study, 190, 3093
- National Endowment for the  
Arts, 276
- National Endowment for the  
Humanities, 2398
- National Environmental Policy Act,  
803–804
- National Family Violence Surveys,  
982–983, 984, 985
- National Health and Social Life  
Survey, 2538, 2550
- National health insurance,  
comparative perspectives, 375,  
2804
- National Health Service (Great  
Britain), 378
- National Hospice Organization, 588
- National Household Survey on Drug  
Abuse, 711, 717, 2559
- National Incidence Study of Child  
Abuse and Neglect, 289
- National Income and Product  
Accounts, 2682
- National Indigenist Institute  
(Mexico), 1858, 1861
- National Institute for Drug  
Abuse, 2398
- National Institute of Anthropology  
and History (Mexico), 1858
- National Institute of Child Health  
and Disease, 2399
- National Institute of Child Health  
and Human Development,  
637, 2398
- National Institute of Demographic  
Studies, 1026
- National Institute of Education, 2398
- National Institute of Mental Health,  
1237, 1840, 3034–3035  
Catchment Area Program, 1834  
depression treatment study, 655  
research funding, 2398, 2399
- National Institute on Alcoholism and  
Alcohol Abuse, 97, 1640
- National Institute on Aging,  
53, 1157
- cohort studies, 344  
research funding, 2398, 2399
- National Institute on Drug  
Abuse, 710, 717
- National Institutes of Health,  
636, 717  
AIDS research, 2585
- National Labor Relations Act of  
1935, 1530, 1553, 2266
- National Labor Relations Board,  
2403
- National Liberation Front  
(Algeria), 2132
- National list system (of  
representation), 2154, 2157
- National Longitudinal Surveys of  
Labor Market Experience,  
2475–2476, 2479
- National Medical Enterprises,  
1820, 1821
- National Nursing Home Surveys  
(1995), 1667
- National Opinion Research Center  
anomia items, 167–168  
civil liberties views, 316, 317  
as data bank, 577, 578  
on family size, 973  
NLS data sets, 2475–2476  
occupational prestige study, 1997  
peacetime disaster research, 681  
quality of life research,  
2299–2300  
secondary data analysis and, 2478  
and social indicators, 2682, 2684  
University of Chicago, 1479
- National Organization for  
Women, 2267
- National Origins Act of 1924, 175
- National Research and Development  
Centre for Welfare and Health  
(Finland), 2451, 2452
- National Research Council, 2461  
population policy report,  
633, 635  
racial discrimination  
perception, 56
- National Research Council/National  
Academy of Sciences report  
(1981), 371
- National Right to Life  
Committee, 3084
- National school systems. *See*  
Educational organization
- National Science Foundation, 637,  
2398, 2399, 2475, 2476, 2532,  
2682, 2685  
Systems Theory and Operations  
Research, 3104–3105
- National security, as censorship  
rationale, 273–274
- National Socialism. *See* Nazism
- National Study of Organizations,  
2476–2477
- National Survey of Families and  
Households, 1729, 1748, 2538
- National Survey of Family  
Growth, 923
- National Survey of Health and  
Development (Great Britain),  
1488
- National Surveys of Adolescent  
Males, 2550
- National Surveys of Men, 2550
- National Toxics Campaign  
Environmental Justice Project,  
791
- National Training Laboratory  
(Bethel, Maine), 1013
- National University of Mexico, 1858
- National Urban League, 326
- National Violence Against Women  
(1998 survey), 2557
- National Wildlife Federation, 790
- National Youth Survey, 494–495,  
496, 497, 1488, 1489  
on criminal delinquency and  
crime, 1488, 1489
- Nationalism, **1939–1949**  
and civil law development, 473  
following decolonization,  
1267–1268  
definition of, 1930  
demographic factors, 636  
genocide and, 1067, 1070  
identity and, 1939–1940  
in Iran, 1871  
and Japanese sociology, 1483  
and Latin American  
corruption, 2134  
in Mexico, 1856, 1858  
movements, 1941–1944, 3001–  
3002

- peripheral, 1267–1268  
 and political party origins, 2154  
 regional and global influences, 1944–1945, 1948  
 social constructions of, 1940–1941  
 in Southeast Asia, 2975, 2978  
 in Soviet former republics, 1199, 1934  
 and state development, 3001–3002  
 state reaction to, 1945–1946  
 and violence and terrorism, 1199, 1947–1948, 3137  
 and war, 3244  
*See also* National border relations
- Nation-states. *See* National border relations; Nationalism; State
- Native American Church, 137
- Native American Rights Fund, 1298
- Native Americans. *See* American Indian studies
- Nativistic movements, 2367
- NATO (North Atlantic Treaty Organization), 1944, 1948, 2130, 2143, 2362, 2608, 3003
- Natural disasters. *See* Disaster research
- Natural Hazards Research Center, 682
- Natural law, 1427
- Natural Resources Defense Council, 803
- Natural selection  
 altruism and, 2882  
 behaviorism and, 209  
 Darwin's theory of evolution by, 876, 878, 2330, 2334, 2369, 2418, 2880–2881, 2885  
 and maximization principle, 2881–2882  
 racial theories and, 2330, 2334  
 sexual selection and, 2885–2886
- Nature vs. nurture, 998  
 intelligence and, 1369–1373, 2090, 2140, 2330  
 personality traits and, 2090  
 sex differences and, 2530
- Navajo, 138
- Nazism  
 as authoritarian  
     communitarianism, 356  
 compliance with authority and, 404  
 as dictatorship, 3002  
 divorce laws, 703  
 and ethnonationalism, 1944  
 and eugenics, 1272  
 expansionist policy, 1933  
 forced labor camps, 2608  
 Frankfurt School study of, 2169  
 genocide policy, 1066, 1067, 1070, 1384, 2206  
 and German families' "legacy of silence," 1512  
 and German sociology, 1074, 1075  
 as 1920s protest movement, 2268  
 party corruption and patronage, 2130  
 racial views, 1272, 2332  
 structuralist view of rise of, 2163  
 Total War concept, 1067  
 use of judicial discretion by, 477  
 war crimes trials, 1429
- NCHS. *See* National Center for Health Statistics
- NCS. *See* National Crime Survey
- NCVS. *See* National Crime and Victimization Survey
- Neal, A. G., 2346, 2350–2351
- Nebbia, Giorgio, 1473
- Need-Achievement Dictionary*, 1979
- Needle (IV) exchange programs, 712, 1642, 2588
- Needs, values differentiated from, 2829
- Neff, Ronald, 1276
- Negative identity, 460
- Negative Income Tax, 2213, 2282, 2284
- Negative moods. *See* Depression
- Negative self-schema, 651
- Negotiation of power, **1950–1956**  
 group conflict resolution, 1111–1117  
 group decision-making, 597  
 interpersonal conflict resolution, 1452  
 and interpersonal power, 1456–1464  
 pragmatic theory and, 2221  
*See also* State, The
- Negritude movement, 66
- Neighborhood structure  
 community studies, 364–366  
 ethnic succession, 532  
 and territorial belongings, 3131  
 and underclass, 2212, 3200  
*See also* Community
- Neitz, Mary Jo, 567, 568
- Nemeroff, Charles, 652
- NEO Personality Inventory, 2079
- Neo-Chicago School, 853
- Neoclassical economics, 725, 726–727
- Neoclassical price theory, 2340
- Neoconservatism, 1600–1601, 1603, 1758
- Neofunctionalism, 1031, 2484  
 and feminist theory, 996
- Neoliberalism. *See* Neoconservatism
- Neolithic period, 2175
- Neo-Machiavellians, 2623
- Neo-Malthusianism. *See* Malthusian theory
- Neo-Marxist theory, 1078, 1214, 2814–2815, 2928  
 "late capitalism" terminology, 1078–1079  
*See also* Conflict theory
- Neo-Nazis, 462
- Neo-Weberians, 2815–2816
- NEP (new ecological paradigm), 806–807
- Nepal  
 child labor, 3262  
 poverty, 2216  
 slavery and slave-like practices, 2605, 2606–2607  
 time use research, 3161, 3162
- Nepotistic favoritism, 2822
- Nested games, 330
- Nestlé company, 3174
- Netherlands  
 "clean" government reputation, 2130  
 cohabitation, 109

INDEX

- divorce rate, 706  
 drug policy, 711–712, 713  
 ethnic status incongruence, 3051  
 family policy, 966  
 health-care system, 374, 375, 376, 377  
 job discrimination against immigrants, 693  
 long-term care and care facilities, 1653, 1655, 1659  
 marijuana decriminalization, 712  
 occupational mobility, 3046  
 physician-assisted suicide, 3085–3086  
 political party system, 2157  
 retirement practices, 2407, 2408  
 Social Science Data Archive, 575, 576  
 social security system spending, 2800  
 sociology of literature, 1649  
 Southeast Asia colonialism, 2975  
 transnational corporations, 3175, 3176  
 Network analysis. *See* Social networks  
 Neugarten, Bernice, 1618, 1619–1620  
 Neuman, W. Russell, 2273  
 Neumann, Franz, 2163  
 Neumann, John von, *See* Von Neumann, John  
 Neumann, Sigmund, 1773  
 Neural network models, 410, 2753  
 Neurath, Otto, 821  
 Neurotransmitters, 652, 654  
 Neutralization theory, 1496  
 Never-married adults, 107–108, 124–125, 487–488, 1738  
 New Age movements, 2717  
 New Australia (Paraguayan utopian community), 2849  
 New China News Agency (Xinhua), 1767  
 New Deal, 862, 2125, 2126, 2427  
   and labor unions, 2148  
   liberalism, 1601  
   and voter realignment, 3235  
   white favoritism, 56  
 New Delhi, India, 2179  
 New ecological paradigm (NEP), 806–807  
 New England  
   black slavery in, 54  
   historical divorce grounds, 701  
*New England Journal of Medicine*, 1818  
 New Guinea, 2569  
   low suicide rate, 3079  
*New Handbook of Political Science, A* (Goodin and Klingemann), 2921  
 New Harmony (Indiana utopian community), 2849  
 New Home Economics, 1009  
*New Ideas in Sociology*, 2979  
 New institutionalism, 737–738  
 New Left, 542, 1784, 2452, 2723  
   and German sociology, 1075  
*New Left Review* (journal), 1757  
 New Orleans Child Guidance Clinic, 324, 325  
 New religious movements (NRMs), 2366–2368, 2374, 2380, 3287–3288  
 New Right, 1580  
 New School of Social Research, 582  
 New World Information and Communication Order, 1767  
 New World Information Order, 280  
 New York City  
   Muslim immigrants in, 2950  
   police force, 2111  
   political machines, 2125–2126  
 New York Longitudinal Survey, 2088  
 New York State  
   black slavery in, 54  
   political machines, 2125  
   right-to-die issue, 587  
 New York Stock Exchange, 726  
 New York University, 2207  
   early clinical sociology course, 324  
*New York Weekly* (periodical), 273  
 New Zealand  
   common law system, 465, 471  
   Social Science Data Archive, 576  
   social security system, 2797  
   woman suffrage, 703  
 Newburyport, Massachusetts (Yankee City project), 364–365  
 Newby, Howard, 2429, 2431  
 Newcomb, M. D., 2513  
 Newcomb, T. M., 334, 335, 336  
 Newcomb, Theodore, 1034, 2415  
*News from Nowhere* (Morris), 3203–3204  
 News media. *See* Mass media; Mass media research; Television  
 Newton, Sir Isaac, 1248, 1781, 1783, 2465  
 Neysmith, Sheila M., 646  
 Nezam Mulq Tussi, 1564  
 NGOs (nongovernmental organizations), 1085, 1827, 1948, 2151–2152, 2461, 2607  
 NHS. *See* National Health Service  
 NHLS. *See* National Health and Social Life Survey  
 NIA. *See* National Institute on Aging  
 Nicaragua  
   Iran-Contra affair, 2128  
   political corruption, 2134  
   protest movements, 2266, 2270  
   revolution study, 1871, 2411, 2412  
 Niceforo, Alfredo, 1465  
 Niche density, 1031  
 Nichiren Shoshu/Sokagakki Buddhism, 2719–2720  
 Nichols, Elizabeth, 1870  
*Nicomachean Ethics* (Aristotle), 2697  
 Nie, Norman H., 2282, 3235  
 Nieboer, Herman J., 2596, 2597–2598, 2607–2608  
 Nieboer-Domar hypothesis, 2596, 2597–2598, 2607–2608  
 Niebuhr, H. Richard, 2378  
 Nielsen, Francois, 847  
 Nielson, Margaret, 2943  
 Nietzsche, Friedrich, 2092, 2756, 2757  
   as sociology of knowledge antecedent, 2953  
 Niger, 2133  
   poverty in, 2216  
   sociodemographic profile, 2938  
 Nigeria, 3079  
   corruption, 2134

- Ibo genocide, 1384  
 literary sociology, 1646, 1650  
 moral development study, 1900  
 slavey and slave-like practices, 2605  
 sociodemographic profile, 2938
- Nightingale, Florence, 3005
- NIH. *See* National Institutes of Health
- NIMH. *See* National Institute of Mental Health, 2398
- 1960s, social changes of, 703
- Nineteenth Amendment, 2266
- Nippon Hoso Kyokai, 1482
- Nirvana, 3279, 3281
- Nisan, Mordecai, 1904
- Nisbet, Robert, 356, 1599
- NIT. *See* Negative Income Tax
- Nixon, Richard, 1764, 1936, 2281, 2284  
 drug treatment program, 711  
 modern retirement policies, 2404  
 Watergate scandal, 2127, 2274, 2276, 2277
- Nkrumah, Kwame, 66
- NLS. *See* National Longitudinal Surveys of Labor Market Experience
- NLS Handbook*, 2475, 2476
- NLSY79 survey, 2475, 2480
- NLSY97 survey, 2475
- Nodes (graph theory), 1789–1790
- Noelle-Neumann, Elisabeth, 2277–2278
- No-fault divorce laws, 703, 704, 708  
 as innovative legal norm, 1560
- Nolan, Patrick, 1886
- Noller, Patricia, 1735
- Nolo contendere* pleas, 480
- Nominal Group Technique, 2618
- Nominal scales, 1793
- Nonconformity  
 criminalization of, 524–526  
 high status as permission for, 404  
 as minority influence, 400–401, 403  
 public tolerance of, 316–317, 358, 606
- Nonconsequentialism, 516
- Nondirective therapy, 1715
- Nongovernmental organizations (NGOs), 1085, 1827, 1948, 2151–2152, 2461, 2607
- Non-identity thinking (Adorno concept), 539–540
- Nonlinear models, 1788–1789, 2252
- Nonmarital childbearing. *See* Unwed childbearing
- Nonmarital cohabitation. *See* Cohabitation
- Nonparametric statistics, **1956–1971**  
 advantages and criticisms of, 1956–1957  
 definition of, 1956  
 literature on, 1957–1958  
 and measurement, 1795  
 new developments in, 1970–1971  
 runs and randomness, 1969  
 tests and techniques, 1958–1970  
*See also* Tabular analysis
- Nonprobability samples, 2444–2445, 3088
- Nonrandom measurement error, 1917–1920
- Nonreductionism, 1780
- Nonstable population model, 618–619
- Nonverbal cues, 1976, 1978, 1980, 2061, 3096
- Nonviolent protest, 2269, 3289
- Non-zero-sum games, 330
- NORC. *See* National Opinion Research Center
- Nordlinger, Eric, 1946
- Norepinephrine, 652, 654
- Noriega, Manuel Antonio, 2135
- Normal accident theory, 2878
- Normalization of danger, 2878
- Norman Conquest, 465
- Normative influence. *See* Social values and norms
- Normative institutions, 395
- Normative sanctions, 515, 523, 537
- Normatology, 2192
- Norms. *See* Social values and norms
- Norpramin, 654
- North, Cecil C., 1997, 1998
- North, Douglass, 735, 2335, 2340
- North Africa, 1865
- North America  
 population factors, 2182  
*See also* American society;  
 Canada; Mexican studies
- North American Free Trade Agreement (NAFTA), 794, 2433  
 labor movement opposition, 1531  
 Mexican participation, 1859, 1861
- North American Industry Classification System, 395
- North American Man Boy Love Association, 1276, 2581
- North American Society for the Sociology of Sport, 2987
- North Atlantic Treaty Organization (NATO), 1944, 1948, 2130, 2143, 2362, 2608, 3003
- North Carolina Manumission Society, 2270
- North Central Rural Sociology Committee, 677
- North Korea, 2137
- Northcote, Lord, 2130
- Northern Ireland, 1947, 2362, 2366, 3288
- Northwestern University, 1876
- Norway  
 age pyramid, 610, 611  
 age-standardized crude death rate, 611, 613  
 legal system, 471  
 life table, 612  
 occupational mobility, 1987, 1988  
 same-sex marriage legalization, 111  
 Social Science Data Archive, 575, 576, 580  
 social security system, 2800  
 unemployment, 3263  
 woman suffrage, 703  
 women in labor force percentage, 3262  
*See also* Norwegian sociology
- Norwegian Data Service (Bergen), 575, 576, 580
- Norwegian Social Science Data Services, 2477
- Norwegian sociology, 2449, 2450, 2451, 2452–2453
- Nouveau riche*. *See* Status incongruence

- NOVA (Norwegian Institute for Research on Childhood and Adolescence, Welfare, and Aging), 2451, 2452
- Novak, Michael, 3286
- Novels. *See* Literature and society
- Novosibirsk Institute of Industrial Economics and Organization, 2980
- Nowak, Stefan, 2119, 2120, 2121
- Nozick, Robert, 2699, 3204–3205
- N*-person games, 329, 2335, 3221
- NRMs. *See* New religious movements
- NSD. *See* Norwegian Data Service
- NSF. *See* National Science Foundation
- Nuclear disasters, 2875
- Chernobyl power plant, 683, 686, 805, 2875, 2876, 2877
  - planning and research, 682, 686
- Nuclear family, 1058, 1259
- American historical, 120, 121, 142, 1502–1503, 1504
  - coalitions within, 331–332
  - communitarian support for, 359
  - divorce effects on, 706–707
  - economic and kinship factors in, 1502–1503
  - fertility decline and, 628
  - and intergenerational relations, 1389
  - “provider” vs. “homemaker” role, 696
- Nuclear weapons, 2363, 3242–3243
- deterrence coalitions, 331
  - disarmament movement, 2723
  - terrorism as strategy component, 3138
- Nuer, The* (Evans-Pritchard), 2890
- Null hypothesis, 456
- Numerosity heuristic, 594
- Nunes, Rosado, 941
- Nunn, Sam, 1881
- Nunnally, J. C., 1794
- Nuremburg Trials, 1429
- Nurses
- increase in, 1819
  - as medical-industrial complex component, 1818, 1822
  - terminal patient care, 582, 587–588
- Nursing homes, 1664–1678
- characteristics and distribution of, 1668
  - definitions of, 1653–1654
  - depression of residents of, 656
  - hospice units, 1671
  - increased numbers of, 1819
  - institutionalization effects on, 1672–1676
  - international variations in, 1654
  - as medical-industrial complex component, 181
  - models of, 1665–1667
  - percentage of population in, 1653
  - personal interactions in, 1676
  - profile of residents, 1667–1668
  - proprietary ownership, 1821
  - quality of care in, 1661
  - recent trends in, 1671–1677
  - special needs units, 1671
  - standards, 1669–1671
  - utilization and expenditures, 1148, 1149
- Nurturance
- and care-based moral reasoning, 1900, 1902–1903
  - gendered roles and, 1058
  - Judaic emphasis on, 1510
  - motherhood and, 2036
- Nutrition. *See* Eating and diet
- Nyden, Philip W., 603
- Nye, F. Ivan, 1488
- Nye, Joseph, 2124
- Nyerere, Julius, 1319
- NYS. *See* National Youth Survey
- O**
- Obedience. *See* Compliance and conformity
- Oberschall, A., 1942
- Obesity, 1641, 3077
- Object relations theory
- adult dependency and, 2063–2064
  - personality development and, 2090
- Objective self-awareness, 2509
- Objectivity. *See* Epistemology; Measurement; Scientific explanation
- Obligations, 2638
- O’Boyle, Ciaran A., 2302
- OBRA. *See* Omnibus Budget Reconciliation Act
- O’Brien, John, 2125
- O’Brien, Robert M., 1594
- Obscenity. *See* Pornography
- Observation systems, **1973–1982**
- categorical, 1973–1975
  - dimensional, 1975–1978
  - future of, 1980–1981
  - recent developments, 1979–1980
  - social psychology, 2769
  - sociocultural anthropology, 2893
- Obshechesvennoe mnenie* (Russia), 2982
- Occupational and career mobility, **1982–1996**
- collective, 2716
  - components in, 2714–2717, 3045–3047
  - and criminal sanctions, 517
  - economic explanations, 1989–1990
  - and educational credentials theory, 2929–2930
  - and elite theory, 2626
  - employer perceptions and, 1993–1994
  - ethnic vs. African-American, 2498
  - factors in, 2715–2717, 2929
  - higher education and, 1178–1186
  - in Hindu caste system, 250–251
  - income and, 3047–3048
  - and intergenerational mobility, 1982–1984, 2711–2712
  - and labor market structure, 1984–1991
  - mobility tables, 2711–2716
  - path model, 1692
  - prestige and. *See* Occupational prestige
  - role change and, 2424
  - social inequality and, 2691
  - socialist quotas and, 2850
  - societal stratification and, 2865–2866, 2867

- and status attainment, 2782, 2783, 2785–2787, 2817, 2867, 3042, 3045–3049
- status attainment  
measurement, 2867
- vacancy chain models, 1990, 2691
- worker characteristics and, 1991–1993
- See also* Caste and inherited status; Education and mobility
- Occupational Crime* (Green), 3247–3248
- Occupational prestige, **1996–2002**  
professions and, 2259–2260, 2263–2264  
ranking stability over time, 1998, 3265  
scale developments, 1997–1998  
social stratification theories on, 1996–1997  
and socialization values, 2773  
status measurement, 3265  
suicide relationship, 3079  
and women professionals, 2262, 2532–2533  
and women’s remarriage potential, 2390
- Occupational segregation (by sex and race), 379, 2012, 3046, 3262, 3264–3265
- Occupations. *See* Professions; Work and occupations
- Oceania, 623
- Ochs, Elinor, 439, 2904
- OCR (optical character recognition), 408
- Odaka, Kunio, 1479
- Odds ratios, 3115–3118
- O’Donnell, Guillermo, 1537–1538
- OECD. *See* Organization for Economic Cooperation and Development
- Oedipus complex, 332, 991, 1273, 1713, 2089
- Oeser, Oscar, 2415, 2417
- Of Human Bonding* (Rossi and Rossi), 1618
- Offe, Claus, 1936, 2163
- Office of Censorship (Department of Defense), 278
- Office of Drug Control Policy, 712
- Office of Economic Opportunity, 2398
- Office of Environmental Equity, 791
- Office of Federal Contract Compliance Programs, 49
- Office of Management and Budget, 259
- Office of Population Research (Princeton), 637
- Office of Scientific Review, 2399
- Office of Technology Assessment, 2461, 3105
- Office of Technology Assistance, 585
- Ofshe, Richard, 901
- Ogburn, William F., 576, 1039, 1424, 2644, 2676, 2682, 2684, 2685  
cultural lag theory, 3066
- Ogives, 659
- Oglala Lakota, 3277–3278, 3279, 3281, 3282
- O’Hare, William P., 2215
- Ohio State University, 681  
Center for Human Resource Research, 2475, 2476  
leader behavior study, 1565, 1566, 1569
- Ohlin, Lloyd, 513, 663, 1494
- Oil spills, 2877
- Okonkwo, Rachel U. N., 1901
- Old-age benefits. *See* Social security systems
- Older adults. *See* Aging and the life course
- Older Americans Act of 1965, 2404
- Oligarchy  
development of, 229, 603, 2163–2164, 2623–2624  
in labor movements, 1532–1533  
in political organizations, 2149, 2163–2164  
in post-Soviet Russia, 2136  
in religious organizations, 2377
- Oliner, Pearl M., 118
- Oliner, Samuel P., 118
- Oliver, Pamela E., 604
- Olivetti, Adriano, 1467
- Ollman, Bertell, 1705
- Olson, Mancur, 604, 2149, 2335, 2338, 2683, 2685, 2920
- Olympic Games, 2985, 2987
- Oman, 1866  
abolishment of slavery, 2602  
sociodemographic profile, 2938
- Omi, Michael, 319
- Omnibus Budget Reconciliation Act of 1981, 1825
- Omnibus Budget Reconciliation Act of 1989, 2399
- Omnibus Budget Reconciliation Act of 1997, 1665, 1670–1671
- Omnibus Reconciliation Act of 1980, 1825
- Omran, A.R., 1325, 1326
- On Crimes and Punishments* (Beccaria), 528
- On Death and Dying* (Kübler-Ross), 582, 583
- “On the Human Mind” (Aristotle), 2086
- On the Jewish Question* (Marx), 2961
- On the Take* (Chambliss), 529
- One Nation After All* (Wolfe), 358
- O’Neill, John, 2208
- One-parent households. *See* Single-parent households
- Only children, 974
- Ontology of Socialism, The* (Staniszki), 2121
- “Open and Closed Relationships” (Weber), 2816
- Opera, 1925
- Operant conditioning. *See* Operant reinforcement
- Operant reinforcement, 209–210, 214, 1716, 2085, 2670
- Operation “Gatekeeper,” 1936
- Operation “Wetback,” 1936
- Operations Research, 3104–3105
- Operations Research* (journal), 3105
- Opinion polls. *See* Election polling; Public opinion; Survey research
- Opium, 714
- Opp, Karl-Dieter, 1080
- Oppenheimer, Franz, 1074, 1075
- Optical character recognition, 408
- Optimal matching techniques, 2297
- Optimism, 1233
- Oral contraceptive, 2178, 2180
- Oral histories, 1633
- Oral sex, 2553, 2567, 2569

- Orbuch, Terri, 1636
- Order-deficit theories, 53  
*See also* Assimilation
- Ordinal scales, 1793
- Ordinary least squares (OLS)  
 estimates, 2251
- Ordinary least squares (OLS)  
 regression, 1922, 2437
- Ordination of women, 2379
- Oregon, right-to-die issue, 587, 3084
- Organic farming, 88
- Organization for Economic  
 Cooperation and Development,  
 375, 794, 1041, 2137–2138
- Organization of African Unity, 1934
- Organizational behavior, 2326
- Organizational culture, 2013–2014
- Organizational development,  
 1014, 1015
- Organizational effectiveness. *See*  
 Bureaucracy; Complex  
 organizations; Industrial  
 sociology; Organizational  
 structure; Social organization
- Organizational knowledge  
 data banks and, 574  
 demographic data, 631  
 leaders vs. managers, 1564
- Organizational structure, **2002–2017**  
 and adaptation to environment,  
 2005–2008  
 Blau's theory of, 698, 2004  
 bureaucracy and, 229–235  
 coalitions within, 331, 332, 333  
 in complex organizations,  
 392–399  
 democratic, 601–602, 603,  
 605, 606  
 distributive justice in, 1003–1004,  
 2003–2004, 2014  
 division of labor in, 696–697,  
 699, 2002, 2003–2004  
 and economic sociology, 736,  
 737–738  
 effectiveness and efficiency of,  
 2004–2005, 2010, 2013–2014  
 elite paradigm and, 2627–2628  
 and globalization, 1085, 1091  
 and goal attainment, 393–394,  
 2008–2011  
 governing class, 604  
 and industrial sociology, 1312,  
 1313, 1314  
 and intergroup/  
 interorganizational relations,  
 1399–1407  
 internal integration, 2011–2012  
 and job mobility, 1984–1985,  
 1986, 1988–1989, 1991–1992  
 latent factors in strategic  
 planning, 2012–2014  
 and meritocracy, 2626–2627  
 oligarchical, 229, 603, 2163–2164,  
 2623–2624  
 participatory decision making,  
 605  
 participatory research, 2041  
 and political power, 2997–2998  
 and political sociology, 2162  
 postindustrial, 2196  
 postmodern, 2201  
 of prisons, 2053, 2056–2057  
 professions and, 2263, 2264  
 and religious organizations,  
 2378–2379  
 role theory and, 2416–2417, 2422  
 scientists and, 2460–2461  
 and social capital, 2639  
 social movement organizations,  
 2722–2723  
 of sociological funding  
 organizations, 2398–2401  
 systems theory and, 3102–3105  
 and theory of the state, 2162,  
 2163–2164  
 urban societies, 3192–3197  
 and work orientation, 3269–3276  
*See also* Corporate organizations;  
 Educational organization;  
 Political organizations;  
 State, The
- Organizations  
 character of, 1314  
 definitions of, 393–394, 441  
 disaster planning and reaction,  
 683, 684, 685  
 theory of, 1091  
*See also* Bureaucracy; Complex  
 organizations; Corporate  
 organizations; Organizational  
 structure
- Organized crime, 505, 530,  
**2017–2022**, 2186  
 in Italy, 2128, 2129  
 and political corruption, 2127  
 in Russia, 2136, 2137  
 white-collar crime and, 3246
- Organized religion. *See* Religious  
 organizations; World religions;  
*specific religions*
- “Organizing Religious Work for the  
 21st Century” (study), 2377
- Organski, J. F. K., 3242
- Orientation to work. *See* Work  
 orientation
- Origins of Private Property, the Family,  
 and the State* (Engels), 988
- Ornstein's regression, 3017
- Oros, Cheryl J., 2557
- Orshansky Index, 2213
- Orwell, George, 2139
- Osgood, C. E., 335, 336–337
- Ossipov, Gennady, 2980, 2982
- Ossowska, Maria, 2119
- Ossowski, Stanislaw, 2119, 2121
- Osterberg, Dag, 2452
- Ostow, Mortimer, 2965
- Ostrom, Elinor, 3071
- O'Sullivan, Lucia, 2558
- “Other” (feminist theory  
 concept), 989
- Otto, Rudolph, 3279
- Ottoman Empire. *See* Turkey
- Outsider art, 173
- Ouvriers des deux mondes, Les*  
 (monograph collection), 1025
- Ouvriers européens, Les* (monograph  
 collection), 1025
- Overidentified model, 1916–1917
- Overpopulation. *See* Human ecology  
 and environmental analysis;  
 Environmental sociology;  
 Population
- Owen, Robert, 2846, 2847,  
 2849, 3203
- Oxford Health Plan, 1822
- Oxford Mobility Studies, 226

**P**

Paci, Massimo, 1468



- Pacific Islanders. *See* Asian-American studies
- Pacificare, 1822
- Packer, Herbert, 2020, 2114
- PACO (Panel Comparability Project), 577
- PACS (political action committees), 144, 145, 2150
- Pagani, Angelo, 1470
- Pagani, Enrico, 1464, 1466, 1469
- Page, Benjamin I., 2276
- Paige, Jeffrey, 390
- Pakistan, 1945
  - and British Empire's dissolution, 2362
  - fertility rate, 2178, 2179
  - genocide, 1070
  - Indian border fighting, 2362
  - political and governmental corruption, 2132
  - religious movements, 2366, 2485, 3288
  - slavery and slave-like practices, 2604, 2606
  - sociodemographic profile, 2938
- Palazzo, Agostino, 1470
- Paleoanthropology, 2889
- Palestine Liberation Organization (PLO), 2048, 2717
- Palestinians, 2048, 2787, 2947
  - and terrorism, 3137, 3139
- Pali texts, 3284
- Paluch, Andrzej, 2119
- Pan Guangdan, 298
- Pan-African Congress, 66
- Panama
  - demographic characteristics, 1535
  - drug trafficking, 2135
- Pandectists, 476
- Panel Comparability Project, 577
- Panel study, 1686, 1687, 1694, 3036
  - attrition in, 1691
- Panel Study of Income Dynamics (PSID), 576, 1685, 2475, 2480
- Pan-Hellenic Socialist Movement (Greece), 2130
- Panic, 348, 2876
- Pan-Indianism, 136–137
- Pankhurst, Emmeline, 989
- Pannuti, Alessandro, 1472
- Papandreou, Andreas, 2130
- Pappi, Franz U., 2693
- Paradigm shift, definition of, 2025
- Paradigms and models, **2023–2031**
  - basic models and applications, 1787–1788
  - Bayesian statistical model, 2249
  - computer assisted, 410
  - definition of model, 2023–2024, 2028
  - definition of paradigm, 2023, 2024
  - differential equation, 1693–1694
  - of drug use, 716
  - event history, 1692–1693, 1871–1873
  - formal models, 2028
  - functions of models, 2028–2029
  - hierarchical linear development, 1174
  - isomorphism of model, 2024
  - in mathematical sociology, 1786–1791
  - in meta-analysis, 1849
  - nonlinear models, 1788–1789
  - paradigm vs. theory, 2024
  - qualitative, 2296–2298
  - religious shift in, 2964–2971
  - scientific, 2458–2459
  - simulation, 2678–2679
  - sociological applications, 2027–2028, 2248–2249
  - statistical graphics, 3015–3017
  - stochastic, 2249, 3036
  - structural equation application, 1692
- Paraguay, 1539, 2134
- Paralinguistic cues. *See* Nonverbal behavior
- Paraprofessionals, 2260
- Parental leave policy, 967, 990, 2033–2034
- Parental roles, **2031–2038**
  - adulthood and, 27, 28
  - changing responsibilities of, 2035–2037
  - child custody awards and, 702
  - and childhood self-esteem development, 2513, 2856
  - and coalition triads, 331–332
  - communitarian view of, 358–359
  - and dependency theory, 2064
  - differential investments theory, 2884–2885
  - and division of labor, 695–696
  - divorce effects on, 706
  - in drug-abuse prevention, 717
  - and educational attainment, 2930
  - and family law, 947
  - and family size, 976
  - and family violence, 983
  - and feminist theory, 992–993, 1028
  - as fertility transition factors, 625, 628
  - and gender concept, 1058
  - Jewish family norms and, 1510–1511
  - life-cycle transition points and, 1616
  - marital roles vs., 1508
  - marital satisfaction pattern and, 1729, 1737
  - in middle and later years, 2037–2038
  - and personality development, 2090
  - power-control theory and, 1490, 1498
  - prescriptive altruism and, 1508
  - and remarriage, 2389, 2390–2392
  - rewards and costs of, 2034–2035
  - shared, 1058
  - and socialization, 1028, 2772–2773, 2856–2858, 2862
  - and support and control, 2857
  - transition to, 2037, 2856–2862
  - and upwardly mobile activities, 2714, 2785–2786, 2787
  - and working mothers, 3266*See also* Filial responsibility; Intergenerational relations
- Parentela orders model, 1513–1514, 1515
- Pareto, Vilfredo, 2632, 3054
  - elite theory, 2625, 2626, 2627, 2644, 3050, 3053
  - on ethnic group formation, 2882
  - on income distribution, 2690

- and Italian sociology, 1464, 1465  
and political sociology, 2163, 2164  
and rational choice theory, 2335, 2338–2339  
on “residuo” of displaying sentiments, 2519, 2528  
on territorial belonging, 3134–3135
- Paris arcade, 541
- Parity distribution of winnings, 330
- Parity progression analysis, 1625
- Park, Peter, 2040
- Park, Robert E., 842, 853, 1401, 2883  
assimilation studies, 176, 177  
and Chicago School, 243–244, 363, 365, 1772  
and collective behavior, 356, 2265  
and human ecological approach, 1209, 1210  
life histories and narratives, 1633  
on marginalization, 2634  
and positivism, 2192  
and pragmatism, 2220–2221
- Park Chung-hee, 2131
- Parker, Patricia, 3252
- Parkin, Frank, 2815, 2818
- Parks, Rosa, 2493, 2494
- Parlements* (France), 474, 476
- Parliamentary system, 1953, 2361  
in Canada, 2360  
in Israel, 2359  
and origins of political parties, 2153, 2154
- Parochial schools, 2861, 2931, 2935
- Parole. *See* Probation and parole
- Parry, J., 1891
- Parsa, Misagh, 1871
- Parsons, Talcott, 244, 352, 423, 733, 783, 856, 878, 991–992, 1233, 1457, 1558, 2193, 2450, 2639, 2819, 3130  
on American kinship system, 1502–1503, 1504  
and antimaterialism, 1783–1784  
boundary maintenance concept, 1931  
British sociology and, 226  
culture and social systems concept, 564–565, 566, 2337  
death and social action theory, 581, 582, 583, 584  
and evolutionary theory, 878, 1705, 1883  
family relations analysis, 1389  
four-function paradigm, 1554, 1559–1560, 1978, 2005  
and functionalism, 1014, 1025, 1030, 1031, 1537, 1704, 1705, 2341, 2415, 2416, 2421, 2484, 3099  
on *Gemeinschaft*, 2630  
general theory of action, 1197  
and German sociology, 1076, 1078, 1080  
on incest taboos, 1273–1274  
as Japanese sociology influence, 1478, 1479  
on law and society, 1554–1555, 1558, 1559–1560  
macro themes of, 1704  
on money’s symbolism, 1890  
on music as social structure, 1924  
neo-evolutionary modernization theory, 1885, 1886  
observation system, 1974, 1978  
and post-World War II sociology, 1783–1784  
and role theory, 2415, 2425  
sick role concept, 1813, 1815  
and social belonging/social conformity relationship, 3131–3132  
and social definition of health, 1136  
and social philosophy, 2057, 2755, 2756  
social system analysis, 2825  
on sociology of knowledge, 2956  
and structure of social belonging, 2630–2631, 2633, 2634  
on territorial belonging, 2629–2630  
values theory, 3212–3213, 3214
- Partial correlation coefficient, 451–452, 453, 457
- Partial residual, 3016, 3017, 3018
- Participant observation. *See* Ethnography; Observation
- systems; Qualitative methods; Sociocultural anthropology
- Participatory democracy, 605, 2164, 2627
- Participatory research, 852, **2038–2044**  
issues in, 2041–2042  
methods, 2038  
relationship to other fields, 2040–2041  
small groups, 2612–2613
- Participatory Research Network (Toronto), 2040
- Particle physics, 2459, 2470
- Particularism, 355
- Partin, Harry, 2380
- Partnership, 441, 442
- Partnership for a Drug-Free America, 716
- Partnership for Organ Donation, 588
- Part-time workers, 1720, 1725, 3262–3263, 3268
- Party system. *See* Political party systems
- Parvenu*. *See* Status incongruence
- Passagen-Werk* (Benjamin), 541
- Passas, Nikos, 166
- Passeron, Jean-Claude, 2642
- Passions, 2518, 2519, 2524, 2529
- Passive static-group design, 2321–2322
- Passive-observational designs, 2321
- Passover narrative, 3282
- Pasteur, Louis, 1327
- “Past-in-present” discrimination. *See* Discrimination
- Pastoralism, 2426
- Patent system, 2461
- Paternalistic leadership style, 1566
- Paternity (family law), 947
- Path analysis, 259–260, 266, 455–456, 975, 1692, 3035–3036  
models, 1788, 2782, 3035  
on status attainment, 3042  
and structural equation modeling, 1910
- “Path Analysis: Sociological Examples” (Duncan), 3035, 3042
- Path coefficient, 1908

- Path goal theory, 1568–1569  
*Pathways to Social Class* (Bertaux and Thompson), 2662  
 Patient Self-Determination Act of 1990, 586, 587  
 Patient-responsive care model, 1666  
 Patient's bill of rights, 1826  
 Patriarchy  
   child sexual abuse theory, 2581  
   family system, 1009, 1271, 1490, 1498, 1503, 1579, 1708, 1734  
   “father figure” as leader, 1566  
   and feminist gender inequality theory, 2692, 2886  
   and feminist rape theory, 2576–2677, 2581  
   and incest theory, 1275  
   individual autonomy seen as, 1307  
   Islamic, 2949–2950  
   and male prison rape victims, 2578  
   object relations theory and, 2064  
   sociobiological view of, 2886  
   and widow status loss, 3255  
 Patrilineal descent, 1515  
 Patronage, 2125, 2127, 2128, 2131, 2134, 2135  
 Pattern theory of culture, 564  
 Patterson, Orlando, 2596, 2597  
 Pattison, Philippa E., 2298  
 Patton, T., 2672, 2673  
 Paul (apostle and saint), 3053  
 Pauly, John, 1498  
 Pauperism, 2840–2841  
 Pawluch, Dorothy, 2763  
 Pawnship, 2597, 2602  
 Paxil, 654  
 Pay equity. *See* Comparable worth  
 PC. *See* Political correctness  
 PD. *See* Prisoner's dilemma  
 PDIA. *See* “Project Death in America”  
 Peace, **2044–2050**  
   coalition formation and, 332  
   conscripted during, 1876–1877  
   and group conflict resolution, 1115  
   and international law, 1429  
   peacetime disaster research, 681–682  
   social movements, 2717, 2719  
   UN peacekeeping troops, 1429  
   *See also* Military sociology; War  
 Peace of Westphalia (1648), 1933, 2363, 2366  
 Pearlin, Leonard, 656, 3057  
 Pearson, Geoffrey, 1579  
 Pearson, Karl, 447, 449, 3035  
 Pearson product moment correlation, 1795  
 Peasant rebellions, 2269, 2412  
   Burma and Vietnam studies, 2977  
 Peasantries, 605, 1857, 1859  
   Chinese socialism's glorification of, 2849  
   decline of, 2431  
 Peccei, Aurelio, 1038  
*Pedagogy of the Oppressed* (Freire), 2040  
 Pedersen, Trond, 2452  
 Pedriana, Nicholas, 1102, 1106, 1107  
 Peel, Robert, 2110  
 Peer groups  
   as deviant behavior influence, 666–667, 668  
   and drug abuse, 717  
   and individual learning, 2932  
   and interpersonal conflict resolution, 1453  
   and self-esteem, 2513  
   as socialization agents, 2858, 2859–2860  
   and student movements, 3069  
 Peer marriage, 359–360  
 Peer pressures. *See* Compliance and conformity  
 Peirce, Charles S., 2218, 2219, 2423, 3098  
 Peiss, Kathy, 2170  
 Pelat, Roger-Patrice, 2129  
 Pelham, B. W., 2512  
 Pelley, William Dudley, 2370  
 Pellicciari, Gianni, 1468  
 Pellizzi, Camillo, 1464, 1470  
 Pellizzoni, Luigi, 1473  
 Penal slavery, 518  
 Pendergast machine, 2126  
 Pendleton Act of 1881, 2127  
 Penitentiary movement, 528  
 Pennati, Eugenio, 1470  
 Penner, Louis A., 118  
 Penology, **2051–2057**  
   American prison population increase, 140  
   and case studies of prison life, 245  
   and criminal sanctions, 4, 515, 517, 528, 535, 2055–2057  
   and criminalization of deviance, 525  
   and deterrence theory, 2056–2057  
   and Norwegian sociology, 2453  
   parole procedure, 2252, 2255–2256  
   prison population, 714, 2056, 2660  
   and prison rape victims, 2578  
   and prisoner labor, 2603, 2608  
   retribution and rehabilitation policies, 2055–2057  
 Pension systems. *See* Retirement; Social security systems  
 Penty, Arthur, 2196  
 Peons, 2597  
 People of Color Environmental Leadership Summit, 792  
*People's Choice, The* (Lazarsfeld, Berleson, and McPhee), 3049  
 People's Liberation Army (China), 2137  
 People's Temple, 900  
 Percentile, definition of, 659  
 Perception. *See* Social perception  
 Perception of graphs, 3009–3011  
 Perestroika, 2981  
 Pérez, Carlos Andres, 2135  
 Pérez Jiménez, Marcos, 2134  
 Performance  
   and “audience effect,” 2615  
   and expectation states theory, 880, 882  
   and role theory, 2422  
   and small group interaction, 2617–2620  
 Performance art, and case studies, 248

- Perinatal death, 2236
- Period effects  
 and changes in social values and norms, 2838  
 definition of, 80, 81  
 means over time, 1691–1692
- Period life table, 612, 614, 1628
- Periodicals. *See* Publications
- Peron, Juan, 1537
- Perron, Charles, 3104, 3249
- Perruci, Robert, 2166
- Perry, Albert D., 2729
- Perry, W. J., 675
- Persian Empire, 2998
- Persian Gulf War, 1766, 1875, 1878, 1944, 2720  
 and international law, 1429
- Personal adjustment, marital  
 adjustment and, 1731
- Personal Attributes Questionnaire (PAQ), 1000
- Personal autonomy, 1945, **2058–2069**  
 definition of, 2058  
 in leisure roles, 1584–1585  
 persuasion and, 2098  
 potential for, 1303  
 radical individual, 1307  
 and scientific research, 2461
- Personal conduct laws. *See* Legislation of morality
- Personal constructs theory, 1715–1716, 2084
- Personal dependency, **2062–2069**  
 integrative view of, 2067–2068  
 negative views of, 2063–2065  
 positive views of, 2065–2067
- Personal efficacy, 560
- Personal hygiene, 2177
- Personal relations  
 affect control in, 45  
 aging and, 83  
 dependency and, 2062  
 identity theory and, 1256  
 individualism and, 1303  
 interpersonal attraction and, 1447–1450  
 interpersonal power and, 1456–1464  
 life course and, 1620  
 observation systems, 1973–1981  
 personality theory and, 2083  
 role theory and, 2415  
 self-concept development and, 2506–2508  
 self-esteem and, 2515–2517  
 sentiments and, 2520  
 social capital accumulation from, 2637–2640  
 and social inequality, 2693  
 social psychology research on, 2770–2771, 2775–2777  
 symbolic interaction perspective, 2767–2768
- Personal resources. *See* Social resources theory
- Personal responsibility and Work Opportunity Reconciliation Act of 1996, 1288, 1719
- Personal rulership, 2133
- Personalities Theories, Research, and Assessment* (Corsini and Marsella), 1712–1713
- Personality, definitions of, 2073–2075, 2082–2083
- Personality and social structure, **2069–2075**  
 and social belonging, 2633  
*See also* Personality theories
- Personality disorders, 1839
- Personality measurement, **2073–2082**  
 across-time stabilities, 2076  
 Big Five (five-factor), 2079–2080  
 of consistencies, 2075–2076  
 definitions of terms, 2073–2074  
 sources of data, 2076–2078  
 traits worth measuring, 2078–2080
- Personality Research Form, 2079
- Personality theories, **2082–2093**  
 aggression and, 70, 75  
 altruism and, 116  
 approaches to, 2069  
 Asian-American culture studies, 178  
 attribution theory and, 193–194  
 criminal behavior theories and, 506–507  
 definitions of personality, 2073–2074, 2082–2083  
 developmental psychology and, 2088  
 feminist theory and, 2089  
 Frankfurt School critical theory use of, 540  
 gender bias and, 1902  
 on genocide-relevant traits, 1070  
 historical precursors, 2086–2087  
 and holism, 2088–2089  
 and homuncularism, 2090–2091  
 intolerance and, 317  
 leadership and, 1564, 1565, 1570–1571  
 major theorists and theories, 1712–1718, 2084–2085  
 mate selection and, 1777  
 and modernization theory, 1885–1886  
 moral development and, 1894–1904, 2089  
 neuropsychological influences in, 2089–2890  
 overview, 2083–2086  
 and personal dependency, 2063–2068  
 of positive mental health, 2188–2191  
 and profession research, 2260–2261  
 prosocial behavior and, 115  
 relational, 2083  
 and self-esteem, 2512  
 and socialization, 2855, 2860  
 tables of, 1713, 2084–2085  
 tourism and, 3168  
 and work satisfaction, 3272–3273
- Personalized instruction, 215
- Person-centered theory, 1715
- Person-years concept, 615–616
- Persuasion, **2093–2099**  
 definition of, 188, 2094  
 extreme influence, 892–902  
 factors affecting effectiveness of, 2094–2097  
 in innovation-decision process, 678  
 overt vs. covert, 2776  
 research approaches to, 2097

INDEX

- susceptibility vs. resistance to, 2097–2098
- Peru  
border dispute, 1934  
demographic characteristics, 1536  
dependency theory, 643  
drug crop control program, 714  
indigenous empire, 2999  
political and social conditions, 1536, 1537  
poverty in, 2216
- Pescosolido, Bernice A., 3079, 3081
- Pessimism, 1233
- Pestello, Frances, 2221, 2222
- Pestello, Fred, 2222
- Pesticides and pesticide control, 88
- Peter, Laurence E., 234
- Peter the Great, emperor of Russia, 1435
- Peters, Robert, 554
- Peterson, Richard A., 173, 567, 1647, 1925, 1926, 1927
- Petrażycki, Leon, 2117–2118, 2979
- Petrocik, John R., 3235
- Petrograd University, 2979
- Petroleos Mexicanos, 2135
- Pettigrew, Thomas F., 2244, 2245, 2246
- Petty, Richard, 188
- Peyote, 137, 713
- Pfeffer, Karl-Heinz, 1076
- “Phalanx” (Fourier), 3203
- Pharmaceutical companies, 1818, 1824, 1827, 2263
- Pharmacists, 2259, 2260, 2261, 2263
- Phenomenology, **2099–2107**  
British sociology and, 226  
ethnomethodology and, 856–860, 2099  
German idealism and, 1248  
German sociology and, 1080  
historical background of, 2099  
implications, 2103–2104  
Jungian stress on, 1714  
life histories and, 1635  
personality theory, 2084  
practice, 2102–2103  
pragmatism and, 2219  
social, 246  
and sociology of knowledge, 2957  
techniques, 2099–2101  
theory, 2101–2102  
in United States, 2756
- Phenomenology of the Social World, The* (Schutz), 226
- Phenotype, 2887
- Philadelphia Geriatric Morale Scale (PGCMS), 2304
- Philadelphia (Pennsylvania), machine politics, 2126
- Philanthropy  
as altruism, 114, 115  
Judaic linkage of food and, 1510  
and social welfare, 2841  
*See also* Voluntary associations
- Philippines, 2974, 2975, 2978  
development strategy, 645  
family size, 977  
guerrilla warfare, 2362  
low suicide rate, 3079  
political corruption, 2131
- Phillips, David P., 3079, 3080
- Philosophical-Economical Manuscripts* (Marx), 822
- Philosophie des sciences sociales* (Worms), 1025
- Philosophy. *See* Social philosophy; *specific philosophies*
- Philosophy and Sociology: From the Methodological Doctrine to Research Practice* (Mokrzycki), 2120
- Philosophy of history, 2644
- Philosophy of Money, The* (Simmel), 733
- Philosophy of science. *See* Scientific explanation
- Phipps, Polly, 2263
- Phoenix House, 715
- Phrenology, 1717
- PHS. *See* Public Health Service
- Phylogenetic unconscious theory (Freud), 2090
- Physical abuse. *See* Aggression; Child abuse and neglect; Corporal punishment; Family violence; Violence
- Physical activities. *See* Sport
- Physical capital, 2637
- Physical education, 2986–2987, 2987–2988
- Physical Quality of Life Index (PQLI), 2302
- Physicians  
assisted suicide by, 585, 586–587, 3083, 3084–3086  
comparative health-care systems, 374–375, 377–378, 379, 380  
comparative income ratio, 1827  
dealing with death by, 582, 587–588  
medical sociology studies on, 1813–1815  
as medical-industrial complex component, 1818, 1826  
practice changes (1970s–1990s), 1819, 1826  
professionalization of, 2259, 2261  
suicide rate among, 3078  
utilization and expenditures, 1148  
women as, 2262–2263
- Physiognomica* (Aristotle), 1717
- Piaget, Jean, 1353, 2824  
altruism studies, 114, 2087  
developmental stages theory, 1894, 1895–1896, 2085, 2092, 2855  
moral development theory, 1902, 1903
- Piazzini, Giuliano, 1468
- Pickering, Andrew, 2459
- “Picture brides,” 176
- Pieretti, Giovanni, 1472
- Pigs for the Ancestors* (Rappaport), 2891
- Pilgrimages, 3282
- Piliavin, Irving, 2943
- Piliavin, Jane Allyn, 118
- Pinchot, Gifford, 802
- Pinna, Luca, 1467
- Piore, M. J., 1985
- Pirenne, Henry, 2940
- Piven, Frances Fox, 101, 603, 604, 606, 2661
- Pizzorno, Alessandro, 1467, 1468, 1470
- Plains Indians, 3277–3278

- Planful competence, 12, 13, 32, 1620, 2860
- Planned behavior, theory of, 1128
- Planned Parenthood, 958, 2240
- Plantation system, 2600–2601
- Plate tectonics, 2459
- Platen, August von, 1066
- Plato, 1644, 2086, 2337, 2519, 2520, 2689  
     elite theory, 2623, 2627  
     utopian design, 3292
- Playfair, William, 3005, 3006
- Plea bargaining, 466, 480, 2960
- Pleasure principle, 2087
- Pleasure-pain spectrum, 667
- Plekhanov, Georgy, 1782
- Plessner, Helmuth, 1075, 1234
- Plessy v. Ferguson* (1896), 2491, 2492, 2493
- PLO (Palestine Liberation Organization), 2048, 2717
- Plog, S. V., 3168
- Plomin, Robert, 2089
- Plummer, Ken, 1633, 1636
- Pluralistic democracy, 604–605, 2164, 2165, 2360  
     political elites vs., 2624–2626, 2627
- Pocock, D. F., 1292
- Podgorecki, Adam, 2121
- Podhoretz, Norman, 1601
- Poetics* (journal), 1649
- Poincaré, Jules-Henri, 821
- Poincare, Raymond, 1423
- Pol Pot, 1069
- Polak, Fred, 1038, 1039
- Poland  
     post-Communist transition, 2136  
     protest movemnt, 2267  
     revolution, 3001  
     social surveys, 577  
     Solidarity labor movement, 1531, 1532, 2121, 2268, 2365–2366  
     tourism in, 3169  
     two-power church-state relations, 2357–2358  
     unemployment, 3263  
     *See also* Polish and Eastern European sociology
- Polanyi, Karl, 384, 387, 724–725, 1224, 1890
- Polanyi, Michael, 2261, 2471, 2756
- Polar types, 3182–3183
- Police, **2107–2116**  
     avocational, 2108–2109  
     British model, 2110–2111, 2113  
     community-oriented, 716, 2114  
     corruption scandals, 2126  
     crime rate records, 499, 500  
     drug-abuse prevention programs, 716  
     interrogation and false confessions, 900–901  
     military analogy, 2112–2113  
     preventive patrol by, 2113  
     private-sector forms, 2109  
     racial profiling by, 1491  
     reorientation of, 2113–2114  
     and riots, 555, 556  
     significance of uniforms and arms, 2111–2112, 2113  
     as social control agents, 537, 556–557, 2108–2109, 2660  
     sociological definition of, 2108–2109
- Police: Streetcorner Politicians* (Muir), 2114
- Police Work: The Social Organization of Policing* (Manning), 2114
- Policeman in the Community, The* (Banton), 2114
- Policy agendas. *See* Political organizations
- Polish and Eastern European sociology, **2116–2123**
- Polish Peasant in Europe and America, The* (Thomas and Znaniecki), 1616, 1618, 1634, 2118, 2220
- Polish Society* (Szczepanski), 2119
- Political action, participatory research and, 2039
- Political action committees (PACs), 144, 145, 2150
- Political and governmental corruption, **2123–2138**  
     in communist and post-communist states, 2136–2137  
     culturally different conceptions of, 2124, 2126–2127  
     in East Asia, 2130–2131  
     and forced labor and sexual prostitution, 2606–2607  
     international measures against, 2137–2138  
     in Middle East, 2132  
     in South Asia, 2132  
     in sub-Saharan Africa, 2132–2134  
     at U.S. federal government level, 2127–2128  
     in U.S. local and state government, 2127  
     in Western Europe, 2128–2130  
     white-collar crime and, 3248, 3251, 3252–3253
- Political correctness, **2138–2142**  
     contemporary, 2139–2142, 2908  
     effects of, 2142  
     history of, 2138–2139
- Political crime, **2142–2147**  
     definition problems, 2143–2144  
     deviance perspectives and, 524–525, 527  
     *See also* Terrorism
- Political discourse, restrictions on, 271
- Political elites. *See* Social and political elites
- Political extremism, 2160
- Political frontiers. *See* National border relations
- Political machines, 2125–2126, 2135
- Political Man* (Lipset), 2917
- Political organizations, **2147–2153**  
     alienation and, 100–104  
     in American society, 144–145  
     elites and, 2623–2624  
     in Latin America, 1536, 1537–1538  
     party machines, 2125  
     political sociology approach to, 2162, 2163–2164  
     types of, 2148
- Political participation. *See* Voting behavior
- Political Parties* (Michels), 2623–2624
- Political party systems, **2153–2162**  
     alignments and realignments, 2156, 2157  
     American trends, 144, 145

- corruption and, 2125–2126, 2127, 2128–2129, 2132, 2135
- definitions of, 2153
- extremism and, 2160
- fragmentation and, 2157–2158, 2159
- institutional theories of, 2153–2154
- and Israeli religious parties, 2359
- in Italy, 2128–2129
- liberal vs. conservative orientations, 1597
- in Mexico, 1857–1858, 1859, 1861, 2135
- and minority parties, 2157
- modernization theories on origins, 2154–2155
- oligarchical leadership in, 2163–2164, 2623–2624, 2624
- organizational approach to, 2162, 2163–2165
- origin theories, 2153–2155
- party loyalty in, 2156
- party types and, 2155
- polarization and, 2160
- political organizations vs., 2148
- representation and, 2156–2157, 2164
- research in 1990s, 2159–2160
- social cleavages and, 2155–2156
- structural features of, 2156–2158
- in Turkey, 2132
- in Venezuela, 2135
- volatility and, 2157
- voting behavior and, 3234, 3235, 3236
- Political revolutions. *See* Revolutions
- Political science
- on coalition formation, 329, 330–332
  - economics interaction with, 2921
  - relationship between sociology and, 2916–2917
  - on uses of corruption, 2124
- Political sociology, **2162–2168**, 2916–2917
- and caste and inherited status, 249–254
  - and class interests, 2628
  - and communitarianism, 356
  - and critical theory, 543
  - definitions of, 2162
  - and democracy, 601–606
  - on elite leadership, 2622–2628
  - and French School of Sociology, 1024
  - and Islamic society, 2941–2942
  - and Latin American studies, 1537–1538
  - and liberalism/conservatism, 1596–1603
  - and macro-level reactions to deviance, 669–670
  - and mass media research, 1764–1766
  - multidimensional theory of, 2165–2166
  - and negotiation of power, 1950–1955
  - and neoconservatives, 1601
  - organizational approach to, 2162, 2163–2165
  - and pluralism, 604–605, 2164, 2165, 2624–2626, 2627
  - and political expression of religious interests, 2358–2362
  - and political justice, 2696
  - and political organizations function, 2150–2151
  - and public opinion, 2272–2280
  - and Scandinavian sociology, 2453
  - and state development, 2996–3003
  - and status incongruence, 3053–3054
  - structural approach to, 2162–2163
  - survey data collection methods, 575
  - and survivalist counterculture, 461–462
  - terrorism research, 3140–3142
  - two converging traditions in, 2162
  - and voting behavior, 3231–3238
  - and war initiation, 3244
- See also* Marxist sociology; Organizational structure; Religion, politics, and war
- Political Systems Performance Data, 2477
- Politics* (Aristotle), 1564
- Politics of Punishment, The* (Wright), 2054
- Polk, Kenneth, 505
- Polkinghorne, Donald, 1637
- Polley, Richard B., 1980
- Polling. *See* Election polling; Public opinion; Survey research
- Pollini, G., 3129
- Pollock, Jackson, 656
- Pollution. *See* Environmental sociology
- Polygyny, 1515, 2885
- Polytheism, 65
- Polythetic-monothetic distinction, 3183–3185
- Pontell, Henry, 3250
- Pontiac (Ottawa leader), 136
- Pooled time series analysis, 2680
- Poor laws, 1018, 1019, 1021, 2840
- Poor People's Campaign, 2497
- Poor population. *See* Poverty
- Popenoe, David, 106, 360
- Popitz, Heinrich, 1078
- Popline*, 1611
- Popper, Karl, 822, 1027, 1077, 2193, 2464, 2756
- Popular culture, **2168–2175**
- cohort perspectives, 342–347
  - critical theory on, 541–542, 2169
  - high culture debate, 565–566, 1645–1646
  - music studies, 1925, 1926
  - social differences in consumption, 2171
  - sociology of art and, 173
  - sociology of culture and, 563
  - sociology of literature and, 1645–1646
  - sporting events, 2986
  - vs. mass culture, 173
- See also* Mass culture; Mass media; Mass media research; Television
- Population, **2175–2184**
- aging, 636, 1018
  - American Indian, 133–134
  - and causal modeling, 259
  - Chinese trends, 303

- composition changes, 636
- demographic methods, 609–620
- demographic transition, 620–629, 633, 2176–2183
- density/crowd behavior  
relationship, 557
- density/"stakes in conformity"  
relationship, 665
- distribution changes, 2179–2180
- dynamics in less developed  
countries, 928–929
- ecology of, 444, 737, 738
- economic growth and, 1218
- and environmental sociology,  
801, 807, 1216–1222
- and epidemiology, 813
- expansive nature of, 1218
- and family and household  
structure, 924
- food supply and, 1219
- growth as disorganizing  
influence, 531, 1229
- growth rate, 749, 1005,  
1007–1008
- historical overview of changes in,  
2175–2177
- human ecology and  
environmental analysis,  
1210–1220
- labor-force changes and, 1521,  
1523–1524
- Malthusian model, 632–633,  
1005, 1008, 1525
- Mexican growth rate, 1859
- Middle Eastern countries, 1866
- neo-Malthusian perspective,  
1219–1220
- nonparametric statistics on,  
1956–1971
- projections, 616–617, 2180–2183
- redistribution, 1945, 2179–2180,  
2181–2182
- regional highest growth rates,  
627–628
- and replacement fertility, 2181
- rural declines, 2428, 2432
- sampling, 283–287, 1687,  
2444–2449
- shifts in, 3195–3196
- size range, 609, 632–633
- socioeconomic characteristics,  
634
- Southeast Asian growth rate,  
2974
- stability of, 1231
- stable vs. nonstable models,  
618–619
- territorial distribution, 633–634
- of United States, 2180
- in urban societies, 3193
- urbanization and, 310, 311–312
- world figures, 2177
- of world's largest  
metropolises, 3915
- See also* Census; Cities;  
Demography
- Population control. *See* Family  
planning
- Population Council, 636
- Population parameters. *See*  
Population, sampling
- Population Reference Bureau,  
623, 627, 628
- Populist Party, 2426
- Porkorny, Alex D., 3078
- Pornography, **2184–2188**  
censorship and regulation of,  
271, 274–275, 2185
- child, 274
- erotica vs., 2184–2185
- feminist opposition to, 274–275,  
2186–2187
- and legislation of morality,  
1577, 1579
- and theories of crime, 505,  
2186, 2187
- Portes, Alejandro, 182, 848, 1538,  
2639–2640, 2640, 2782
- Portocarero, Lucienne, 425
- Portugal  
African colonization, 60
- Latin American colonization by,  
536–1637
- Southeast Asian influences, 2974
- Positive mental health, **2188–2191**  
normality vs., 2191
- Positivism, 540, 722, 1465,  
**2192–2195**  
on criminal behavior, 528, 529
- critical theory rejection of,  
539, 541, 543
- and cross-cultural analysis, 548
- and culture debate, 565
- and epistemology, 818–819, 822
- and German sociology, 1077
- Hegelian rejection of, 1248–1249
- humanism vs., 1246, 1249–1250
- and Italian sociology, 1464, 1465
- and Japanese sociology,  
1477, 1479
- main tenets of, 2194–2195
- postpositivism, 2193–2194, 2220
- pragmatism as response to, 2217
- scientific explanation and, 2468
- and secularization, 2483
- and sentiments, 2522
- uses of terms, 2192
- Posse Comitatus, 462
- Postadolescence, 26
- Postcolonialism, 545, 1708, 1758
- Postindustrial society, **2195–2205**  
characteristics of, 2195,  
2196–2199, 2205
- and convergence theory, 422
- global division of labor, 697
- and new industrial elites, 2626
- and postmodern society,  
2199–2201
- and status incongruence, 3051
- Postmodernism, **2205–2209**  
British sociological studies,  
227–228
- characterization, 2200
- and convergence theory, 422
- and cultural theory, 1756, 1784,  
2173, 2206–2209
- definition of, 1884
- definitions of, 2200
- disaster conceptualization, 682
- epistemology and, 2207–2208
- family and religion and, 935, 942
- and feminist theory, 995
- interpretative biography, 1635
- legal anthropology, 1550
- Marxist sociology as target of,  
1756, 1757–1758, 1784
- and medical sociology, 1816
- modernization theory vs., 1884



INDEX

- and new criminologies, 505  
 postindustrialism and, 2199–2201, 2202, 2203  
 radically modern, 2291  
 revised pragmatism and, 2220  
 science and, 2459  
 secularization and, 2489  
 sentiments and, 2528–2529  
 on social change, 2648  
 and social structure, 2819  
 and sociology, 2757, 2758  
 and Sokal hoax, 2207  
 and symbolic estates, 1512  
 and theories of crime, 504  
 “tourist gaze” concept, 3166–3167, 3172–3173  
 values research thesis, 3222–3223
- Postone, Moishe, 1785
- Postpositivism, 2193–2194, 2220
- Poststructuralism, 1027, 1034, 1707  
 characterization of, 2206  
 critical theory and, 544  
 cultural theory and, 1758, 2172–2173  
 and historical sociology, 1199  
 Marxist sociology and, 1756, 1757, 1758  
 on words/symbols, 2206
- Posttest-only control group design, 2322–2323
- Posttraumatic stress disorder (PTSD)  
 and childhood sexual abuse, 290, 293  
*DSM* inclusion as mental disorder, 1833
- Potlatch, 1032, 2883–2884
- Poulantzas, Nicos, 1755, 2163, 2814
- Pound, Roscoe, 2961
- Poveda, Tony, 3248, 3251, 3252–3253
- Poverty, **2209–2217**  
 African-American factors, 2333–2334, 2496–2499  
 aggression linked with, 71, 984–985  
 capitalism and, 241  
 centrifugal kinship system and, 1512  
 children and, 1287, 3062  
 in cities, 311  
 community studies, 365  
 conceptions of, 2210–2211  
 and crime rates, 531, 536  
 “culture of poverty” explanation, 2211–2212  
 debates over, 1288  
 and decision-making voice, 603–604  
 declines in, 1286–1287  
 definition problems, 2210  
 definitions of, 1286–1288  
 demography of, 2214–2215  
 depression and stress resulting from, 656  
 “deserving vs. undeserving poor” distinction, 2210–2211  
 earned income tax credit and, 1287  
 ecology and, 1214  
 environmental state of “strain” and, 663  
 epidemiology of, 817  
 and family planning, 955–956  
 and family policy in Western societies, 967–968  
 feminization of, 1288, 2033, 2215, 3048  
 food supply and, 1219  
 gender and, 1288, 2033, 2215, 3048  
 health-care deficiencies, 130  
 Hispanic-American rates, 1191–1192, 1287, 2215  
 illegitimacy and, 1260  
 income distribution and, 1285–1288, 2214–2215  
 individualistic attributions and, 2705  
 informal economy and, 1338  
 international, 1289  
 and life course mobility, 2715  
 macro-level deviance theories and, 663–666, 672  
 measures of, 2213–2214  
 minimum wage and, 1287  
 official level of, 2213, 2214  
 old-age, 2403, 2404  
 population and, 1217, 1221  
 relative vs. absolute, 2210–2211  
 rising expectations and, 1491  
 and rural areas, 2431  
 single-parent household, 127, 128, 2215, 3062  
 and slavery and forced labor, 2608  
 and social justice, 2705–2706  
 and social work, 2840–2841  
 Supplemental Security Income and, 1286  
 theory and policy, 2211–2213  
 thresholds, 1286, 1720  
 as transitory state, 1288  
 underemployment and, 1720–1721  
 and urban underclass, 2497–2498, 3198–3200  
 War on Poverty program, 1286, 1494, 1882, 2404, 2760  
 and welfare programs. *See* Social welfare system  
 widowhood and, 3257  
*Poverty of Postmodernism, The* (O’Neill), 2208  
 Powell, G. Bingham, 2917  
 Powell, Walter W., 1647  
 Power, Martha Bauman, 1636  
 Power analysis (statistical), 3030–3032, 3033  
*Power Elite, The* (Mills), 1773  
 Power theory  
 elites and, 2622  
 as exchange theory basis, 2669–2670, 2671–2672  
 interpersonal, 1456–1464  
 and juvenile delinquency, 1490, 1498  
 and nationalism, 1940  
 political vs. other types, 2997  
 and profession research, 2261–2262  
 and rape, 2587, 2589, 2590  
 and social inequality concept, 2690  
 and societal stratification, 2864–2868  
 and status attainment, 2781  
 three-dimensions of power, 2165–2166  
 on U.S. racial relations, 53–54  
 and values research, 3221

- and war, 3242  
 Weber's classic theory, 2165, 2865  
*See also* Negotiation of power
- Power threat hypothesis (Blalock concept), 847
- Power transition theory, 3242
- PPOs (preferred provider organizations), 1818, 1819
- PQLI (Physical Quality of Life Index), 2302
- Prabhupada, A. C.  
 Bhaktivedanta, 3287
- Prabhupada, Srila, 462
- Practical sociology. *See* Sociological practice
- Pragmatism, **2217–2224**  
 and humanism, 1249–1250  
 influence on social science, 2219–2222  
 main ideas and variations, 2217–2219  
 and social problems, 2759  
 and sociology of knowledge, 2955  
 and symbolic interactionism, 2423, 2896, 3098
- PRE measures (proportional reduction in error), 1810, 1811
- Prebisch, Raul, 1087
- Preconscious, 1713
- Predatory property crime, 506
- Predestination (Calvinist doctrine), 774–775
- Predicting Success or Failure in Marriage* (Burgess and Cottrell), 1726
- Prediction and futures studies, **2224–2233**  
 methods of prediction, 2226–2231  
 social forecasting, 2676–2680  
 social indicators, 2685  
*See also* Futures studies as human and social activity
- Predictive validity, 3208, 3211
- Preference reversals, 598
- Preferred provider organizations (PPOs), 1818, 1819
- Pregnancy and pregnancy termination, **2233–2242**  
 age demographics, 125
- AIDS/HIV transmission, 2587, 2591  
 comparative health-care systems, 374, 378  
 conduct of, 2235–2236  
 and family planning, 952–954, 2180  
 fetal effects of alcohol, 1640  
 fetal effects of smoking, 1640  
 lifestyle risks, 1640  
 maternity leave plans, 2033–2034  
 maternity leave provisions, 2033–2034, 3266  
 neonatal deaths, 1325, 1326, 2236  
 nonmarital, 484, 488, 1744  
 outcomes of, 2236–2238  
 pregnancy and birthrates, 2234–2235  
 pregnancy rate definition, 2234  
 pregnancy trimesters, 2233–2234  
 premarital, 1744  
 prenatal drug use incidence, 711  
 prenatal tets, 2237  
 rape fears and, 2587  
 response to pregnancy, 2234  
 and sexually risky behavior, 2559  
 sociological view of, 2884  
 termination of, 2238–2241  
*See also* Abortion
- unmarried, 125, 128, 484, 488, 634, 1258–1264, 1506, 1626, 1744, 2033  
 unplanned, 953–954, 2180, 2234  
 viable vs. nonviable, 2238  
*See also* Birth and death rates; Childbearing; Fertility rate
- Prehistoric population, 2175–2176
- Prejudice, **2242–2248**  
 conservatism and, 1600  
 ethnic, 844–846  
 homophobic, 2567, 2569–2570  
 racial, 2242–2246  
 social distance concept, 177  
*See also* Discrimination; Stereotypes
- “Preliminary Concepts for a Theory of Religious Movements” (Stark and Bainbridge), 2375
- Premenstrual syndrome, 1816, 1833
- Prenatal tests, table of, 2237
- Prescription drugs. *See* Medications
- Presidential elections. *See* Election polling; Voting behavior
- President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry, 1826
- President's Commission on Country Life (1908), 2426, 2427
- President's Commission on Law Enforcement and Administration of Justice, 2055
- President's Commission on Obscenity and Pornography (1971), 2184, 2815
- Press. *See* Mass media; Mass media research
- Press freedom, 273
- Pressure groups. *See* Interest groups
- Prestige, 416  
 social stratification theories, 1996–1997  
*See also* Occupational prestige; Status attainment
- Pretest observations, 2325
- Pretransition societies, fertility determinants in, 1006–1007
- Prevision. *See* Futures studies as human and social activity
- PRI (*Partido Revolucionario Institucional*; Mexico), 1857–1858, 1859, 1861, 2135
- Primal horde (Freudian concept), 1273, 1576
- Primary data, 574
- Primary groups. *See* Small groups
- Primary jobs, 1985
- Primary Mental Abilities* (Thurstone), 905
- Primary metropolitan statistical area, 307
- Primary sampling units, 2448
- Primitive Classification* (Durkheim and Mauss), 1032
- Primogeniture, 1350–1351
- Primordial sentiments, 2906
- Prince, Samuel, 681
- Prince, The* (Machiavelli), 1564
- Princeton University, 1868  
 fertility transition research, 625–627

- Office of Population Research, 637
- Princeton University Press, 2917
- Principal components analysis, factor analysis vs., 913
- Principes historiques du droit* (Vinogradoff), 1025
- Principles of Biology* (Spencer), 1029
- Principles of Environmental Justice*, 792
- Principles of Psychology* (James), 2218
- Principles of Sociology* (Spencer), 697, 1029
- Principles of Topological Psychology* (Lewin), 1012
- Printing press, 267, 275
- Prison, The: Policy and Practice* (Hawkins), 2054
- Prison, The: Studies in Institutional Organization and Change* (Cressey), 2052
- Prison Community, The* (Clemmer), 2051–2052
- Prison Officers and Their World* (Kauffman), 2053
- Prison population, 714, 2056, 2660
- Prison riots, 352
- Prison system. *See* Penology
- Prisoner labor, 2603, 2608
- Prisoner's dilemma game, 596, 2337, 2338, 2419, 2620–2621, 2700, 3220, 3221
- Prisons in Turmoil* (Irwin), 2054–2055
- Private funding organizations, 2400–2401
- Private sphere
- feminist theory on, 990, 991–993, 2601
  - health care industry, 1819–1826, 1827–1828
  - long-term care funding, 1659
  - religion and family in, 935, 942
- Proactive aggression, 69
- Probability sample surveys, 301, 2444–2448
- survey research, 3088–3094
- Probability theory, **2248–2252**
- decision-making theories, 590–592
  - life tables, 615
  - main concepts, 2249–2250
  - sampling, 2250
  - scientific explanation, 2465, 2466
  - statistical inference, 3025–3026, 3030–3032
- Probation and parole, **2252–2258**
- Probit analysis, 3038
- Probit Analysis* (Finney), 3035
- Problem invariance, 598
- Problem solving
- small group interaction in, 2617–2620
  - small group vs. individual, 2618–2619
- Procedural democracy, 545
- Procedural justice. *See* Social justice
- Proceedings of the Sixty-Fourth Annual Congress of the American Prison Association*, 325
- Process-produced data, 574
- Production of Culture, The* (Crane), 568
- Production system model, 2297
- Production-oriented leadership, 1566
- Productivity
- labor-market experience and, 1989–1990
  - postindustrial, 2198
  - technology and, 3267, 3268
- Profession of Medicine* (Freidson), 1813–1814
- Professional associations. *See* American Sociological Association and other sociological associations; International associations in sociology
- Professions, **2259–2265**
- changes in, 2263–2264
  - information-based elites, 2626
  - marginal, 2259–2260
  - paraprofessionals, 2260
  - and postindustrial society, 2197, 2198, 2199, 2202
  - professionalization process, 2260
  - ranking of, 2259–2260
  - research study approaches, 2260–2262
  - retirement patterns, 2406
  - semiprofessionals, 2259–2260, 2261, 2262
  - sex differences, 2532–2533
  - sex segregation, 2259, 2262–2263
  - social work, 2841–2842, 2845
  - sociologist certification program, 155–156
  - suicide rates, 3078
  - system of, 2262
  - and voluntary associations, 3229
  - See also* Occupational prestige
- Prognosis. *See* Futures studies as human and social activity
- Prognosis: A Science in the Making* (Polak), 1038
- Progress, theories of, 2644–2645
- Progressive era, 365–366, 2127
- Prohibition, 526, 1576–1577, 1580, 2019, 2127
- Project Camelot, 838, 839
- “Project Death in America,” 588
- Project Metropolitan (Stockholm), 2454
- Project Northland, 715
- Project 100,000 (War on Poverty), 1882
- Proletariat
- and capitalism, 238–239
  - and revolution, 2410, 2848–2849
  - See also* Class struggle
- Property
- conservatism and defense of, 1598, 1601
  - crimes of, 506, 509–510, 556, 1576
  - liberalism and defense of, 1597
  - slavery tied to, 2596–2598
- Prophecy. *See* Prediction and future studies; Social forecasting
- Proportion tests, 1970
- Proportional hazards models, 617, 871–873
- Proportional reduction in error (PRE) measures, 1810–1811
- Proportional representation, 2154, 2157, 2164, 2359, 2360
- Proportionate random sampling, 2447
- Proprietorship, 441, 442
- Prosocial behavior. *See* Altruism
- Prospect Theory, 592
- Prospective. *See* Futures studies as human and social activity

INDEX

- Prosser, R., 3167
- Prostate cancer, 1641
- Prostitution, 2186, 2559–2561  
 child, 2607  
 deviance theories on, 664  
 global sex industry, 2560–2561, 2607  
 homosexual, 2559, 2560, 2561  
 and legislation of morality, 1577  
 and sexually transmitted diseases, 2585, 2591–2593  
 studies of male clients, 2560
- Protection of Children against Sexual Exploitation Act of 1977, 2581
- Protest movements, **2265–2271**  
 collective behavior factors, 349, 350–354, 558  
 consequences, 2267–2268, 2270–2271  
 and countermovement resistance, 2266, 2267–2268, 2717, 2718  
 and crowd behavior, 557–558  
 extremist terrorists, 3139  
 Iranian studies, 1870–1871  
 music and, 1927  
 Native American, 136–137  
 participants and methods, 2268–2270  
 participatory decision-making in, 605  
 political alienation and, 101, 103  
 and political criminality, 2144–2145  
 popular vs. unpopular, 2269  
 and riots, 555–556  
 and triad coalitions, 333  
*See also* Revolutions; Social movements; Student movements
- Protestant ethic, 1577, 1885, 2211, 2456, 2483, 2943  
 values theory and, 3219, 3222
- Protestant Ethic and the Spirit of Capitalism, The* (Weber), 237, 389, 722, 774, 2483, 2942, 2986
- Protestantism  
 American affiliation figures, 146  
 and American sociomoral political issues, 2361  
 and capitalism, 2942–2943  
 and censorship, 267  
 as conducive to democracy, 605  
 decline of liberal mainline denominations, 2379  
 ecumenical movement, 2365  
 evangelical expansion, 2966  
 expansion of, 937  
 fundamentalist, 2361, 2368–2372  
 and moral taint of poverty, 2211  
 and negative taint to poverty, 2211  
 schisms in, 2364, 2370  
 social activism of, 2365  
 and voluntary associations origin, 3227  
 and wars of religion, 2363
- Provo Study, 1488
- Provocation  
 aggression and, 72, 73  
 by politically alienated, 101–102
- Prozac, 654, 718
- Przygotzki, Nathalie, 1903–1904
- PSDA. *See* Patient Self-Determination Act of 1990
- PSID (Panel Study of Income Dynamics), 576, 1685, 2475, 2480
- Psychiatry, 1832–1833
- Psychoanalysis, 1713–1714  
 and adult dependency, 2063, 2067  
 critical theory and, 543  
 femininity/masculinity theory, 998  
 personality theories, 540, 1713–1714, 2084, 2087, 2089, 2090, 2860
- Psycholinguistics, 2895
- Psychological Abstracts (database), 409, 1015
- Psychological Abstracts/PsycInfo*, 1611
- Psychological casualties. *See* Extreme influence: thought reform, high control groups, interrogation, and recovered memory psychotherapy
- Psychology  
 Chinese study of, 298  
 and criminology, 528–529  
 and cross-cultural analysis, 548  
 depression theories, 650  
 evolutionary, 1234  
*See also* Social psychology  
*Psychology of Being, The* (Maslow), 2087  
*Psychology of Helping and Altruism* (Schroeder et al.), 118  
*Psychology of Sex Differences, The* (Maccoby and Jacklin), 2530–2531  
 Psychology of supply and demand, 278  
 Psychometrics, 906, 908, 1909, 3039  
 intelligence testing, 1359–1360, 1362–1368, 1375, 1376, 2330  
 Psychosis. *See* Mental illness and mental disorders
- Psychotherapy  
 as drug abuse treatment, 715  
 family therapy case studies, 247  
 personality theory and, 1713–1715  
 recovered memory, 901–902
- Psychotropic agents, 717
- PTSD. *See* Posttraumatic stress disorder
- Puberty, 6
- Public administration, 1951, 1952
- Public data, 574–575
- Public funding organizations, 2398–2400
- Public goods theory, 2338
- Public health  
 demographic transitions and, 622–623  
 diffusion of new drugs, 677–678  
 drug abuse control, 712–713  
 healthy life expectancy and, 1632  
 lifestyle risks, 1639–1642  
 and mortality rate decline, 2177–2178  
 as social control, 521
- Public Health Service, 2398–2399
- Public housing, 1656
- Public interest group, 2148, 2149
- Public opinion, **2272–2280**  
 on alcoholism as disease, 96  
 on American political and social issues, 145  
 attitude as indicators, 190–191  
 attribution theory, 190

- on civil liberties, 316–318  
 collective participants, 348  
 on drug abuse, 713  
 dynamics of, 2275–2276  
 Eurobarometers, 549, 577, 578–579, 3223  
 on governmental unresponsiveness, 101  
 Internet as polling potential, 1768–1769  
 on liberal/conservative beliefs, 1602–1603  
 low rating of lawyers, 468  
 policy effects of, 2276–2277, 2282  
 polling approaches, 575, 1686, 2273–2275, 2277–2279, 3232  
 polling quota samples, 2444–2445  
 polling survey data collection methods, 575  
 on rape, 2583  
 on sexual harassment, 2581  
 social surveys, 821, 2769  
 survey instruments, 576–579  
 survey research, 3087–3094  
*See also* Voting behavior
- Public policy analysis, 2280–2285**  
 adult education, 23  
 biotechnology advances, 1824  
 conceptual development, 2280–2281  
 data banks and depositories, 574–577  
 deviance theories, 666, 667–668, 669–670  
 drug abuse, 711–717  
 evaluation research, 2282–2283  
 family policy, 963–964  
 filial responsibility, 1021–1022  
 health care financing, 1814, 1827  
 health policy, 1159  
 illegitimacy, 1261  
 implementation analysis, 2284  
 interest group lobbying, 2115, 2150  
 language policy, 2909  
 long-term care funding, 1658–1659  
 marital adjustment, 1731  
 medical-industrial complex, 1824–1826  
 methodological research, 2281  
 outcome analysis, 2283–2284  
 policy explanation, 2281–2282  
 pro-family programs, 2033  
 public opinion and, 2276–2277, 2282  
 and quality of life concept, 2686  
 response to structural lag, 3064  
 social indicators, 2687–2688  
 and social problem objectivist paradigms, 2760, 2764  
 social security systems, 2797–2805  
 sociological research funding, 2398–2401  
 time use research, 3163–3165  
 utilization of, 2284–2285  
*See also* Government regulation
- Public schools. *See* Educational organization**
- Public spaces, communitarian view of, 359**
- Publications**  
 American Sociological Association, 150, 153, 154, 155  
 Asian-American studies, 179  
 clinical psychology, 326  
 electronic, 413  
 ethnography, 854–855  
 first use of term “clinical sociology,” 325  
 historical sociology, 1197  
 Japanese sociology, 1481–1482  
 library resources, 1606  
 library search techniques, 1608–1611  
 life histories and narratives, 1635–1636  
 Middle Eastern studies, 1864, 1868–1869  
 neoconservative, 1601  
 on-line indexes access, 1607  
 on qualitative methods, 2293  
 rural sociology, 2427  
 Scandinavian sociology, 2451  
 on small group research, 1980, 1981  
 on social indicators, 2685  
 of social work, 2845  
 sociological computing, 407, 409, 411  
 Soviet and post-Soviet sociology, 2981, 2983  
 sport sociology, 2987, 2988  
 Web sites for nonelectronic, 414  
*Public Interest, The* (periodical), 1601  
 Publishing industry, 1647–1648, 2172  
 Puerto Ricans  
 divorce rates, 126  
 household structure, 127  
 Puerto Rico, 3263  
 Pugliese, Enrico, 1467  
 Pugwash meetings, 2047  
 Pukumina (Jamaican religion), 65  
 Punch, Maurice, 3248, 3249–3250, 3253  
 Punched-card system, 283, 406, 420, 575, 580  
 Punishment  
 and behavioral conformity, 2616  
 as deviance prevention, 667, 2341  
 of nonconformity, 525–526  
 parental style/child’s deviant behavior relationship, 2858  
 as social capital factor, 2638  
 social learning theory on, 1717  
 sociology of, 517–518  
 three dimensions of, 2659  
*See also* Capital punishment; Corporal punishment; Criminal sanctions; Penology; Rewards; Social exchange theory
- Punk rock counterculture, 460, 461, 462  
 Pure-conflict games, 2337  
 Puritanism, 2211, 2456  
 divorce grounds, 701  
 work ethic, 1577, 1885, 2211, 2456  
 Purnell Act of 1925, 2426  
 Purposive action model, 2838–2839  
*Pursuit of the Millennium, The* (Cohn), 2367  
 Putnam, Robert, 359, 368, 3228, 3229  
 Pyschodynamic theories of depression, 649

Psychogenic theory of criminal behavior, 506

**Q**

*Qabus-Nameh* (Unsuru'l-Ma'ali), 1564

Q-analysis, 3183–3185, 3187, 3188

Qatar, 1865, 1866

QCA (qualitative comparative analysis), 386

Qin dynasty, 2998–2999

Qing dynasty, 3000

QL. *See* Quality of life

QOL. *See* Quality of life

QPL (Questionnaire Programming Language), 410–411

QSERVE (interactive data access system), 409

Quadagno, Jill S., 483–484

Qualitative comparative analysis (QCA), 386

*Qualitative Health Research* (journal), 2293

*Qualitative Inquiry* (journal), 2293

Qualitative methods, **2287–2296**  
 artistic aspects, 2291–2293  
 comparative-historical sociology, 386

computer analysis, 410

content analysis, 417–421

criminal analysis, 530

cross-cultural analysis, 547

ethnomethodological case studies, 247–248

macrosociology, 1709–1710

middle-group approaches, 2289–2291

phenomenological research, 2100

scientific explanation, 2468

social forecasting, 2677–2678

*See also* Data analysis

Qualitative models, **2296–2299**

*Qualitative Sociology* (journal), 854, 2293

*Quality of American Life: Perceptions, Evaluations, and Satisfaction* (Campbell et al.), 2300

Quality of dying, 585, 587–588

Quality of life, **2299–2309**

definitions of, 2301–2304

health promotion and, 1171

human ecology and, 1229

measurement of, 2302–2304, 2683–2684

nursing home improvement interventions, 1672

retirement and, 2404

as social indicator, 2683–2684, 2685, 2686–2687

summary indices of, 2686–2687

theories of crime and, 503

widowhood as diminishment of, 3257–3258

*Quality of Life and Pharmacoeconomics in Clinical Trials* (Spilker), 2301

*Quality of Life Newsletter*, 2301

*Quality of Life of Cancer Patients, The* (Aaronson and Beckman eds.), 2301

*Quality of Life Research* (journal), 2301

*Quality of Life Valuation in Social Research, The* (Mukherjee), 2301–2302

Quantitative law of effect, 214

Quantitative methods, 1633

of criminal analysis, 530

of cross-cultural analysis, 547

macrosociology, 1709–1710

meta-analysis, 1843–1850

of prediction, 2227–2231

statistical, 2028, 3034–3039

statistical graphics, 3011–3019

statistical models, 2028

tabular analysis, 3107–3126

typologies, 3182–3185

Quantity principle, and decision processing, 593–594

Quarks, 2459, 2470

Quasi-experimental research designs, **2309–2328**

construct validity threats, 2324–2325

external validity threats, 2323–2324

internal validity threats, 2312–2315

measurement and, 1792–1803

nonexperimental designs, 2321–2322

panel study as, 1686

public policy, 2282

statistical conclusion validity threats, 2326–2327

Quebec, Canada

church-state relations, 2360

legal system, 464

national movement, 1941, 1947, 2719

separatist political violence, 2145, 2271

Queen, Stuart A., 483–484

Queens College-CUNY, interactive database access, 409

Questionnaire Programming Language (QPL), 410–411

Questionnaires

computerized self-administered, 410–411

and secondary data analysis, 2478

survey research, 578, 3088–3090, 3093

Quill, Timothy, 587

Quine, Willard Van Ormand, 823

Quinlan, Karen Anne, 586, 3084

Quinnry, Richard, 3247

Quint, Jeanne, 582

Quirino, Elpidio, 2131

Quota samples, 2444–2445

Qur'an. *See* Koran

**R**

Race, **2329–2335**

adulthood transition and, 35

affirmative action and, 47–52, 2706

African peoples and, 61–63

AIDS/HIV risks, 2587–2590

alcohol use and, 94–95

alienation and, 102, 103

altruism and, 117

in American society, 142, 143

attributional patterns and, 196, 197

biological conceptions of, 2329–2332

census and, 284–285, 286

childbearing and, 125, 1010, 2032

childlessness and, 109

as Chinese immigration exclusion law basis, 175

INDEX

- classification systems, 62, 2330–2332
- collective mobility of, 2715–2716
- as conflict theory factor in deviance, 670
- and crime rates, 498, 530–532, 534, 536
- and critical theory, 539, 541
- cycle theory, 177
- and direct and indirect discrimination, 689–694
- diversity in American cities, 308
- divorce and, 112, 126, 705, 707, 708
- drug abuse policy and, 714
- education and, 145
- educational achievement and, 2931, 2932–2933
- equality of opportunity and, 830
- and ethnicity, 841
- ethnicity differentiated from, 2329
- European imperialist view of, 320–321
- and family households, 127, 925
- and family planning, 955–956, 966
- genocide and, 1067, 1070, 1071–1072
- and hate speech, 275–276
- helping behavior and, 117
- hypodescent rule and, 2331–2332
- income inequality and, 2691, 3048
- infant and child mortality and, 130, 1334
- informal economy and, 1340
- intelligence and, 1372–1373, 2330
- intermarriage and, 1407–1415
- intermarriage and divorce rates, 1750
- intermarriage demographics, 124
- intermarriage/occupational achievement intersections, 2691
- interracial courtship, 485
- interracial marriage laws, 949, 1776
- interrelationship with class. *See* Class and race
- life expectancy and, 125, 1631
- manipulation of classifications, 62–63
- marital age and, 124, 1750
- and mental illness diagnosis, 1838
- military sociology and, 1879, 1880
- music ascriptions and, 1926
- and nursing home residents profile, 1667
- and occupational segregation, 379, 2012, 3046, 3264–3265
- “one-drop rule,” 1296–1297
- political correctness and, 2140, 2141
- population composition studies, 633
- poverty level and, 2215
- prejudice and, 2243–2246
- protest movements. *See* Civil rights movement
- public opinion shifts on, 2275
- quality of life and, 2300
- remarriage rates, 126, 2388
- retirement and, 2406–2407
- segregation and desegregation, 2491–2499
- segregation indices, 2500–2504
- singlehood and, 107–108
- social conceptions of, 2332–2334
- social Darwinism and, 2330, 2334
- and social stratification, 253, 558, 2817–2819
- South African categories of, 62
- and status attainment, 3044, 3046
- and status incongruence, 3054
- stereotypes of, 64, 197–198, 2243
- and structural lag, 3063
- suburbanization and, 3074–3075
- suicide rates, 3078, 3079
- and underemployment, 1721, 1722, 1724
- and uneven power allocation, 1940
- and urban riots, 555–556, 557, 558, 2269, 2270, 2495, 2661
- voluntary association membership and, 3227, 3228
- See also* African American studies; American Indian studies; Asian-American studies; Class and race; Racism
- Race Matters* (West), 2220
- Race relations theories, 53–54
- resentments and, 2245
- See also* Class and race; Prejudice; Racism
- Racial prejudice. *See* Prejudice
- Racial profiling, 1491
- Racial resentment, 2245
- Racial unconscious theory (Jung), 2090
- Racism
- antiblack ideology, 55, 56, 57–58, 66
- apartheid, 62, 1940, 2047, 2146
- capitalism and, 319–321, 322
- collective aggression and, 349
- environmental, 789, 803, 809, 1159
- environmental equity and, 789, 790, 795, 796
- environmental sociology and, 803
- ethnicity and, 844–845
- eugenics and, 879
- feminist theory and, 988, 989
- genocide and, 1067
- imperialist roots of, 320
- institutional, 53–54
- prejudice and, 2242–2246
- resistance movements, 58
- Social Darwinism and, 2330, 2334
- stereotypes, 64, 2243
- U.S. historical, 57–58
- of white working class, 321
- See also* Class and race; Discrimination; Segregation and desegregation
- Radcliffe-Brown, A. R., 563, 564, 1030, 1031, 1034, 2890
- Radical activist countercultures, 460
- Radical democracy, 545
- Radical individualism, 316
- Radical positivism, 2468
- Radical sociology, case studies, 245–246
- Radical-Marxist criminology, 504–505, 534–535
- Rado, Sandot, 650
- Radojkovic, Mroľjub, 1472

- Radway, Janice, 1647, 1648, 1650, 2171–2172
- RAE. *See* Research Assessment Exercises
- Raffault, J., 1025
- Ragin, Charles, 243, 386, 547, 1709–1710, 2297
- Rahe, R., 3055
- Rahim, M. A., 1569
- Rahmann, F., 2939–2940
- Rainwater, Lee, 2452
- Rama, 3288
- Ramirez, Francisco O., 427–428, 2662
- Ramos, Fidel, 2131
- Ramsay, Natalie R., 2451
- R-analysis, 3183–3185, 3187, 3188
- Rand Corporation, 1876, 1877, 1881  
 Medical Outcome's Study, 2306
- Random measurement error, 1908–1909, 1917
- Random net theory, 1790
- Random sampling, 2274, 2288, 2324  
 procedures, 2446–2447  
 replication, 2397
- Randomness  
 causal relationship experiment design, 2312, 2322–2323  
 hierarchical linear model effects, 1176–1177  
 probability theory and, 2249, 2252
- Range (statistical), definition of, 659
- Rank, Otto, 2058, 2085, 2087
- Rank-dependent theory, 592, 598
- Ranke, Leopold, 1180
- Rao, M. S. A., 1291
- Rap music, 1926
- Rape, 2556–2559, 2587–2594  
 AIDS/STDs and, 2576, 2585  
 and blaming the victim, 2579  
 British definitions of, 499  
 child sexual abuse similarities, 2581  
 common law definition of, 2587  
 criminological theories of, 503–504  
 explanations for, 2579–2580  
 gang, 2580  
 juvenile age and arrest rate, 1489  
 law revisions, 2576, 2578  
 legal definition, 2577  
 male honor and, 2579, 2580  
 of male victims, 2578  
 marital, 950, 2588–2589  
 rapist typologies, 2588  
 and sexually explicit media content, 2187  
 stranger vs. nonstranger, 2556–2557, 2558–2559, 2588  
 treatment and prevention, 2583  
 Uniform Crime Reports definition, 492, 499  
 and war, 2578–2579
- Rape in Prison* (Scacco), 2578
- “Rape: The All American Crime” (Griffin), 2587
- Rapoport, Anatol, 1791
- Rappaport, Roy, 1226, 1228, 2891
- Rasler, Karen, 1871
- Rastafarianism, 64, 65
- RATE (computer software), 3037, 3038
- Rating scales, attitude measurement, 185–186
- Ratio level measurement, 1793
- Ratio scales, 1793
- Rational choice theory, **2335–2343**  
 base rate information, 591  
 and coalitions, 330  
 and collective behavior, 349–350, 2335, 2338–2339  
 criminological theory and, 535–536  
 of democratic voting behavior, 2339  
 deterrence theory and, 519, 529  
 differential association theory and, 533  
 and economic sociology, 732, 939, 2335–2336, 2338  
 and exchange networks, 2674  
 feminist theory and, 996  
 functional explanation, 2341–2342  
 game theory and, 2335, 2336–2337, 2419  
 and human nature, 1234  
 and institutions, 2340, 2342  
 and interest group formation, 604  
 and interpersonal power, 1459  
 and mathematical sociology, 1791  
 and new paradigm of religion, 2375  
 and postmodernism, 2206  
 and public interest group support, 2149  
 and Scandinavian sociology, 2452  
 and social networks, 2731  
 and social values and norms, 2340–2341, 2835–2836, 3222  
 and status incongruence, 3050  
 and voting behavior, 3236  
 and war, 3244  
*See also* Social exchange theory
- “Rational Choice Theory and Religion” (1994 conference), 2375
- Rational legal systems, 464
- Rationalism. *See* Scientific rationalism
- Rationality  
 individualism and, 1306  
 postmodern theory on, 2206  
 secularization and, 2483  
 variability of judgment and, 598
- Rationality and Society* (journal), 1791, 2375
- Rationalization, theory of, 541
- Rational-legal authority, 230
- Ratzel, Friedrich, 1933
- Ratzenhofer, Gustav, 1074
- Rawls, John, principles of justice, 2698–2699, 2700, 3205
- Reactance theory, 2325–2326  
 and aggression, 69  
 and motivation, 2060
- Reactive arrangements, 2325–2326
- Reading  
 literary reception theory and, 1648  
 as virtue, 1650  
*See also* Literacy
- Reading the Romance* (Radway), 1648
- Reagan, Ronald, 101, 1150, 1159, 1825  
 and deregulation, 3250  
 drug abuse policy, 712



- and fundamentalism, 2371  
 and Iran-Contra affair, 2128  
 and Social Security system, 2799  
 and voting behavior, 3234
- Reality construction, 2099  
 case studies, 246–247
- Reason, 1303  
 and emotion, 773, 785–786  
*See also* Rational choice theory;  
 Rationality
- Reasoned action theory, 189
- Rebellions. *See* Revolutions
- Recent Social Trends* (study),  
 576, 2682
- Reception theory (literary), 1648
- Reciprocal altruism, 115, 2882,  
 2883–2884  
 and family care, 1391, 1394,  
 1567–1568, 1657–1658
- Reciprocal causation, 261  
 social learning theory and,  
 1716–1717
- Reciprocity  
 Lévi-Strauss principle of, 1032  
 marital balancing of, 1508  
 in Native American potlatch,  
 1032, 2883–2884  
 as social capital basis, 2637,  
 2638, 2640  
 and symbolic interactionism,  
 2423–2424  
 and territorial belonging, 3129
- Recorded culture, 568–569
- Recovered memory psychotherapy,  
 901–902
- Rectilinear coordinate graphs, 3005
- Recursive models, and causal  
 inference models, 256, 257,  
 258, 261
- Redfield, Robert, 367, 1506,  
 2428, 2483
- Redlich, Frederick, 1813, 1834, 3055
- Redlining, 3072
- Reductionism, 1780
- Reed, Gary E., 3253
- Reed, Mark, 2658
- Reeducation programs, 892–898
- Reference group theory, 2702–2704,  
 2752–2753  
*See also* Role theory; Self-concept
- Reference services. *See* Data banks  
 and depositories; Library  
 resources and services for  
 sociology
- Reflected appraisal model (self-  
 esteem), 2515
- Reflections on the Revolution in France*  
 (Burke), 1599
- Reflex arc (Dewey concept), 1014
- Reflexive thinking. *See* Self-concept
- Reform movements, 853, 2717
- Reformation, 2154, 2211, 3227  
*See also* Protestantism
- Reformatories, 518
- Réforme sociale, La*, 1025
- “Reforms as Experiments”  
 (Campbell), 866
- Refugee Act of 1980, 143
- Refugees  
 demographic factors, 636  
 distinguished from  
 immigrants, 180  
 Indochinese, 180, 181, 182  
 interviews with, 299
- Regime, definition of, 2356
- Regional sociological  
 organizations, 153
- Règles de la méthode sociologique, Les*  
 (Durkheim), 819, 1024
- Regression analysis. *See* Correlation  
 and regression analysis
- Regression coefficients, 259, 262,  
 451, 1693
- Regression line. *See* Linear  
 regression models
- Regression-discontinuity design,  
 2319–2320
- Reher, D., 633
- Rehnquist, William, 587
- Reich, Charles, 460, 1773
- Reich, Robert, 1091, 1347
- Reichel-Dolmatoff, G., 1225
- Reid, Anthony, 2978
- Reid, Whitelaw, 1357
- Reification theory, 541, 543
- Reigrotzki, Erich, 1076
- Reil, Johann Christian, 2087–2088
- Reiman, Jeffrey, 3246, 3248
- Reinforcement theory  
 cognitive dissonance, 338  
 operant conditioning, 209–210,  
 214, 1716, 2085, 2670  
 social learning, 1717, 2768
- Reischauer, Robert D., 3199
- Reiss, Albert J., Jr., 2108, 2682,  
 3034, 3247
- Reiss, I. L., 2537
- Reiss, Ira, 1778
- Relational power theory, 2165  
 and social inequality, 2693
- Relational properties of  
 collectives, 1592
- Relative deprivation  
 collective protest and, 349, 1940  
 and distributive justice, 2701  
 and status incongruence, 3050
- Relative mobility, 2712
- Relativism  
 and liberalism, 355  
 and scientific knowledge,  
 2459, 2955  
 and sociology of knowledge,  
 2953, 2954, 2955
- Reliability, **2343–2356**  
 consistency and stability as  
 components of, 2343  
 multiple indicator tests of,  
 1909–1910, 1917  
 in narrative research, 2292  
 of observation systems, 1977  
 in personality measurement,  
 2075–2076  
 qualitative data criteria,  
 2288–2289  
 replication and, 2395–2397  
 of single-indicator measures, 1921  
 and specificity of variables, 1803  
 time use research, 3159  
 validity vs., 1909, 2080
- Religion  
 African traditions, 65  
 and alcohol use, 94–95  
 and canon law, 473, 1513, 1514,  
 1516, 1545  
 and caste and inherited  
 status, 250  
 and censorship, 267  
 and cohabitation, 109  
 and conservatism, 1598, 1599  
 divine experience and, 3283

- and divorce bans, 700  
 Durkheim on function of, 775–776, 1032, 2385, 3278  
 experience vs. conversion, 3283  
 family and. *See* Family and religion  
 gender and, 997, 1057, 2379  
 genocide and, 1069, 1071  
 hate speech and, 275–276  
 historical sociology and, 1198, 1199  
 on human nature, 2086–2087  
 and incest taboo, 1272  
 and interfaith dialogue, 3289  
 and interfaith marriage, 911, 1411, 1776  
 and legislation of morality, 1560, 1580  
 Native American, 137–138, 3277–3278, 3279, 3281, 3282  
 new paradigm of, 2367–2368, 2374–2375  
 and ordination of women, 2379  
 paradigm shift in, 2964–2971  
 personality theories and, 2086–2087  
 political correctness and, 2142  
 political expression of, 2358–2362  
 as political orthodoxy in premodern states, 2138  
 and professional clergy, 2259  
 sectarian, 1303  
 and secularization, 2482–2489  
 secularization of, 935, 937–938, 2373, 2374  
 sociological defining feature of, 2382–2383  
 and sociology of culture, 563  
 subjective, 2384–2385  
 suicide and, 3081  
 theology and doctrines, 3284–3286  
 world, 3277–3289  
*See also* Church and state; Judaism and Jews; Protestantism; Religious movements; Religious organizations; Religious orientations; Roman Catholic Church; Sociology of Islam; *specific denominations and sects*
- Religion in Secular Society* (Wilson), 2483
- Religion, politics, and war, **2356–2364**  
 decline of liberal mainline denominations, 2379  
 and ethnicity, 3288  
 and fundamentalism and, 2361, 2369, 2371, 2378, 2943, 2944–2945, 3288  
 and secularization, 2485–2486
- Religious fundamentalism. *See* Fundamentalism
- Religious movements, **2364–2376**  
 categories of, 2364–2368  
 and counterculture, 460, 461, 2366, 2367, 2374, 2969  
 fundamentalism, 277, 462, 1577, 1580, 2361, 2368–2373, 2378, 2386, 2966–2967  
 new movements, 2366–2368, 2374, 2380, 3286–3288  
 new paradigms in, 2966–2967  
 and social conservative vs. communitarian view of moral decline, 359  
 and social movements, 2719, 2969  
 theoretical understanding of, 2373–2375  
*See also* World religions
- Religious organizations, **2376–2387**  
 in American society, 145–146, 2376–2377  
 changes in, 2364  
 civil liberties challenges, 315  
 and civil rights movement leadership, 315, 2365, 2374, 2377, 2493  
 and consensus formation, 2379–2380  
 corporate, 441  
 current trends, 2379–2380  
 denominational politics, 2377  
 Jewish, 2377  
 secularization and, 2482–2489  
 in sociological context, 2377–2379  
 special-purpose, 2380  
 theological texts and, 3285  
 umbrella, 2380  
 voluntary associations, 3227, 3228–3230  
 in world religions, 3286–3287
- Religious orientations, **2382–2387**  
 democracy and, 605  
 and interfaith dialogue, 3288–3289  
 pluralism and, 2969–2970  
 and world religions, 3277–3289  
*See also* Religious movements
- Religious sociology. *See* Sociology of religion
- Relman, Arnold, 1818
- Remarriage, **2387–2395**  
 blended families, 112–113, 126  
 comparative rates in developed countries, 1749  
 division of labor in, 2390  
 divorce and, 488, 2885  
 divorce potential, 125  
 divorce rate, 1747, 2393  
 first marriages vs., 2390–2393  
 following divorce, 708–709, 1746, 2387, 2388  
 and life cycle, 1625  
 marital quality and, 2390  
 mate selection and, 1779, 2388–2390  
 parental roles and, 2033, 2389, 2390–2392  
 predictors of, 2388  
 of previously married couples, 1748  
 rate decline, 2387  
 rate for divorced women, 1746  
 rate upswing, 1744  
 rates, 1970–1990 (table of), 1742  
 rates by age, 1742, 1743, 1748, 1779, 2388  
 stability of, 2392–2393  
 variables in, 1750  
 after widowhood, 126, 1749, 2387, 2388, 3255, 3259
- Remick, Helen, 371, 372
- Remmling, Gunter, 2956
- Rempel, J. K., 335
- Repeated cross-sectional surveys, 1687
- Replication, **2395–2397**, 2471  
 secondary data analysis, 2481

- Report to the Nation on Crime and Justice*, 2684
- Reports on the World Social Situation*, 2916–2917
- Reppucci, D., 76
- Representative democracy, 602–603  
and political elites, 2627  
and political party systems, 2154, 2156–2157, 2164  
and religious interests, 2358–2362
- Representativeness  
and decision processing, 594–595  
and electoral system, 2156–2157
- Repression (emotional), 1713–1714
- Reproduction. *See* Childbearing;  
Birth and death rates;  
Demographic transition; Fertility determinants; Pregnancy and pregnancy termination
- Reproduction of Mothering, The* (Chodorow), 992–993
- Republic, The* (Plato), 1644, 3202
- Republican Party  
conservatism and, 1603  
factors in emergence of, 2154  
machine politics, 2125, 2126  
and political polls, 2278  
protest movements within, 2267  
voter classification, 3233, 3234
- Resampling methods. *See*  
Bootstrapping; Sampling methods
- Research and Development Corporation (RAND), 1038, 1041
- Research Assessment Exercises (Great Britain), 227
- Research Center for Group Dynamics (University of Michigan), 1013
- Research Committee 38 (ISA), 1633–1634
- Research Committee on Social Trends, 1039, 2682
- Research Committee on Stratification (ISA), 550
- Research funding in sociology, **2397–2401**, 2475
- Research methods. *See* Case studies; Comparative-historical sociology; Ethnomethodology; Ethnography; Evaluation research; Qualitative methods; Survey research; Statistical methods
- Research on Social Work Practice*, 2845
- Research replication. *See* Replication
- Research synthesis. *See* Meta-analysis
- Research Working Group on the Ottoman Empire, 1876
- Residential homes  
for drug abuse, 715  
long-term care, 1652, 1654–1655
- Residential segregation. *See* Housing; Segregation indices
- Resolving Social Conflicts* (Lewin), 1013
- “Resource curse,” 643–644
- Resource dependency, 737–738  
and environmental sociology, 800, 805  
and social exchange theory, 2674
- Resource mobilization theory, 791, 1871  
of collective behavior, 353–354  
and social movement organizations, 2723
- Resources. *See* Social resources theory
- Resources, natural. *See* Human ecology and environmental analysis
- Respite care, 1658
- Response bias, 213
- Responsibility  
cultural variants in attributions of, 198  
as prosocial behavior trait, 116
- Responsive Communitarian Platform, The: Rights and Responsibilities*, 357
- Responsive communitarians, 355, 356–360
- Responsive Community, The* (journal), 356
- Rest, James B., 1899
- Restitution, as criminal sanction, 515
- Restlessness, circular reaction and, 350
- Rethinking Marxism* (journal), 1757
- Retirement, **2401–2410**  
age of, 3061  
community case studies, 245  
continuing care communities, 1664  
as death “cause,” 584  
and decreased male labor-force participants, 1524  
determinants of, 2405–2406  
early, 2404, 2406, 2407, 2408  
first pension plan, 2402  
gender and race issues, 2406–2407  
historical development of, 2402–2404  
international comparisons, 2407–2408  
measurement of, 2406  
social factors of, 79  
social security payments and, 2403, 2796, 2803, 2804  
women’s income, 84
- Retirement Research Foundation, 588
- Retributivism, 516
- Rettig, S., 2350
- Reuel, Denney, 1595
- Reurbanization, 311
- Revelle, Roger, 1218
- Revenge of the Past, The* (Naimark), 1948
- Revenue Act of 1921, 2402
- Reverse discrimination, 2706
- Review* (journal), 1869, 1872
- (Re)vision (Rich pun), 991–992
- Revised Feelings of Inadequacy Scale, 2512
- Revitalization movements, 2367
- Revolutionary Armed Forces of Colombia (FARC), 2135
- Revolutions, **2410–2415**  
analysis of Iranian, 1870–1871, 1873  
Chinese socialist, 2849  
civil law effects of, 473–474  
coalitions within, 331, 332, 333  
collective behavior factors, 349  
comparative historical analysis of, 383, 384, 2413, 2414, 3000–3001  
conflict theory on, 416  
contemporary theory of, 1707–1708

- counterculture groups, 460  
 creole decolonization and, 1267  
 enforcement coalitions and, 415–416  
 historical case studies of, 1198  
 J-curve hypothesis of, 349  
 Lenin's theory of, 2848–2849  
 Mao's theory of, 2849  
 Marx's theory of, 2339, 2410–2411, 2412, 2413, 2414  
 Merton's anomie concept and, 165  
 nationalism and, 3001–3002  
 Old Regime vs. Third World, 2412  
 outcomes and, 2414  
 postcommunist, 2851  
 proletarian, 2848–2849  
 rebellions vs., 2410  
 social movements and, 2717  
 socialist, 2848–2849  
 Southeast Asia studies, 2977  
 state and, 3000–3001  
 and status incongruence, 3050  
*See also* American Revolution; French Revolution; Mexican Revolution of 1910; Russian Revolution
- Revolutions of 1848, 3000, 3001  
*Revue française de sociologie, La*, 1026  
*Revue internationale de sociologie*, 1025
- Rewards  
 and behavioral conformity, 2616  
 and compliance, 400  
 as external social control, 2657  
 group cohesiveness and, 2615  
 helping behavior and, 115, 2774  
 and intrinsic motivation, 2059–2060  
 of parenting, 2034  
 and social justice, 2697, 2701–2704  
 and work satisfaction, 3272, 3275  
*See also* Reinforcement theory; Social exchange theory
- Rex, John, 226  
 “Rhapsodien” (Reil), 2087–2088  
 Rheaume, Jacques, 328  
 Rheinstein, Max, 1554
- Rhetoric of censorship, 268–269  
 Ricardo, David, 1597  
 Rice, Stuart A., 3231  
 Rich, Adrienne, 991  
*Rich Get Richer and the Poor Get Prison, The* (Reiman), 3248  
 Richards, Audrey, 289  
 Richardson, Laurel, 1637  
 Richmond, Mary, 2843  
 Rickert, Heinrich, 819–820, 2957  
 Ricketts, Erol, 3199  
 Rickman, David, 194  
 Ricoeuer, Paul, 3281  
 Riemann, Gerhard, 1635  
 Riemer, Svend, 1275  
 Riesman, David, 4, 1582, 1595, 2529, 2624  
*Riggs v. Palmer*, 466  
 “Right to die” movement, 585, 586, 3064, 3065, 3083  
 Rights. *See* Civil rights; Human rights, children's rights, and democracy  
 Right-wing extremism  
 political parties and, 2160  
 survivalist counterculture, 461–462  
 Riker, William, 331, 2920  
 Riley, Matilda White, 150, 1618  
 Rinaldi, A., 583  
 Rinde, Erik, 1425  
 Ringer, Benjamin, 54  
 Riots. *See* Crowds and riots  
 Rising expectations theory, 1491, 2270, 2493
- Risk  
 control theory on, 667  
 lifestyle and health, 1639–1642  
 macro-level research, 2877–2880  
 micro-level research, 2876–2877  
 rational decision theory and, 350  
 sexual behavior and, 2559, 2577  
 technological, 2874–2579
- Risse, Heinz, 2986  
 Rita, Lidia de, 1467  
 Ritalin, 718  
 Rites of passage, 1624, 2861  
 Ritzer, George, 1703–1704, 2027, 2208–2209
- Robbery  
 deviance theories on, 664  
 juvenile rate, 1498  
 rate calculation, 497, 498  
 Uniform Crime Reports definition, 492  
 victimization rates, 499  
*See also* Burglary; Theft
- Robert Wood Johnson Foundation, 587, 588, 715, 2400  
 Robertoland, 227  
 Roberts, Brian, 1338, 1578  
 Robertson, Roland, 1199  
 Roberty, Evgeny de, 2979  
 Robeson, Paul, 1071  
 Robespierre, Maximilien, 1066  
 Robinson, Dawn T., 1790  
 Robinson, J. P., 3209  
 Robinson, Ronald, 1266  
 Robinson, William S., 1592  
 Robotics, 699, 3266  
 Rochberg-Halton, Eugene, 2217  
 Roche Holding, Ltd., 1824  
 Rocheblave, Anne-Marie, 2415  
 Rockefeller Foundation, 637, 716, 2398  
 Rockwell, R. C., 2300  
 Rodman, Hyman, 2212  
*Roe v. Wade* (1973), 2274, 2277  
 Roemer, John, 2815  
 Roemer, Milton I., 376  
 Rogers, Andrei, 620  
 Rogers, Carl, 1714, 1715, 1717, 1718, 2085, 2087, 2088  
 Rogers, Everett, 86, 87, 89, 676, 678–679  
 Rogers, Kim Lacy, 1636  
 Rogers, Mary F., 1648  
 Roh Tae Woo, 2131  
 Rohypnol, 488, 2577  
 Rokeach, Milton, 2120, 3214–3216, 3217, 3222, 3223  
 Rokeach Value Survey, 3214–3216, 3217, 3222, 3223  
 Rokkan, Stein, 574, 576, 2154–2155, 2450, 2452, 2453  
 comparative political analysis, 2922  
 Role change, 2424–2425, 2838

INDEX

- Role choice behavior, 1253, 1255
- Role conflict, 2416–2417
- Role playing, 2423
- Role theory, **2415–2420**
- altruism and, 117–118
  - cognitive perspectives, 2417–2418, 2424
  - disagreements and confusions on, 2415–2416, 2419
  - functionalist, 2416, 2421–2423, 2425
  - history of, 2415–2416
  - and marginality, 2634–2635
  - organizational analysis, 2416–2417
  - pragmatism and, 2221
  - recent trends in, 2418–2419
  - and self-concept, 2507–2508
  - and situational self-concept, 2506
  - small-group division of labor, 696
  - small-group role differentiations, 695, 2422–2423, 2774–2775
  - as social psychology perspective, 2415, 2417–2418, 2768, 2772
  - and socialization, 2855–2856, 2860, 2861
  - and status incongruence, 3050
  - structural, 2417
  - symbolic interactionist, 2417, 2423–2424, 2856, 3097
  - traditional, 2423
  - See also* Gender roles; Parental roles; Social roles
- Role theory: foundations, extensions, and applications, **2420–2425**
- Role-Construct Repertory Test, 2076
- Romagnoli, Guido, 1467
- Roman Catholic Church
- adaptability study, 2379
  - alcohol consumption rates, 95
  - American affiliation figures, 146
  - in Canada, 2360
  - canon law, 473, 1513, 1514, 1516, 1545
  - case study of nun, 245
  - censorship by, 267, 268
  - communitarian elements in social teachings of, 355
  - divorce ban, 1516
  - historical antimaterialism, 1781
  - kinship and marriage teachings, 1514, 1516–1517
  - labor movement and, 1529
  - low suicide rates and, 3079, 3081
  - parochial schools, 2861, 2931, 2935
  - in Poland, 2357, 2365–2366
  - religious movement incorporation by, 2364–2365
  - and social movements, 2365–2366
  - U.S. membership, 2376
  - and wars of religion, 2363
- Roman law, 465, 473, 475, 476, 1513
- Montesquieu comparative study of, 1545
- Roman Republic and Empire, 2998, 2999
- city-states in, 2998
  - civil law antecedents in, 465, 473
  - frontier line, 1932, 1935
  - mass destruction of Carthage, 1066, 1069
  - slavery in, 2811
- Romance novels, 1699, 2171–2172
- Romania
- illegal abortion death rates, 2241
  - labor movement, 1532
  - political and government corruption, 2136, 2137
  - protest movement, 2267
  - revolution, 3001
- Romanian, sociology in, 2117
- Romantic fiction, 1648, 1699
- Romantic love complex, 1698–1700
- Romanticism, 1697, 1700
- Rommetveit, Ragnar, 2415
- Roof, Wade Clark, 2484
- Room, Robin, 2451
- Roos, J. P., 2453
- Roosevelt, Franklin D., 1286, 1311, 1596, 2125, 2126, 2273, 2277, 2403, 3232, 3235
- Roosevelt, Theodore, 1007, 2426
- Root, Elihu, 1427
- Roots* (Haley), 67
- Roots and Counterpoint* (papers collection), 179
- Roper Center for Public Opinion Research, 2477
- Roper, Elmo, 3232
- Roper Public Opinion Research Center, 317–318, 576
- Rorschach ink-blot test, 2077
- Rosaldo, M. Z., 992
- Rosen, Bernard, 848
- Rosen, Lawrence, 464
- Rosenbaum, James L., 1989
- Rosenberg, Gerald N., 2962
- Rosenberg, M., 2344, 2346, 2347, 2350
- Rosenberg, M. J., 335, 337
- Rosenberg, Neil, 1927
- Rosenberg, Stanley, 2512
- Rosenfeld, Richard, 503
- Rosengren, Karl-Erik, 1648, 2453
- Rosenkranz, P. S., 2190
- Rosenstock, Irwin M., 1814
- Rosensweig, Mark, 633
- Rosenthal, R., 1849–1850
- Rosenthal, Sandra, 2217
- Rosenzweig, Roy, 2170
- Ross, Catherine E., 1834–1835
- Ross, Edward, 1074, 1423
- Ross, Edward Alsworth, 1738, 3245
- Ross, Marc, 1070, 1071
- Rossi, Alice S., 360, 989, 1618
- Rossi, Peter, 156, 425, 1206, 1618, 2215, 2281, 2282
- Rossiter, Clinton, 1598
- Rostow, Walt W., 1705–1706, 1885, 1886
- Rozzak, Theodore, 459
- Rotating panels, 1687
- Rotation problem-correlated factors, 913–914
- Rothchild, John, 460
- Rothko, Mark, 656
- Rothman, Ellen K., 483
- Roturier*. *See* Status incongruence
- Roueché, Berton, 814
- Rousseau, Jean-Jacques, 527, 528, 1024, 1029, 1233, 1247, 1248, 1249, 1310
- Rovine, Michael, 1737
- Row frequency and percentage, 658
- Rowan, Charles, 2110
- Rowland, M. D., 76
- Rowntree, B. S., 2210

INDEX

- Royal Society of London, 1423  
 RU 486 (mifepristone), 2238–2239  
 Rubakin, N. A., 2979  
 Rubenstein, Jonathan, 2114  
 Rubin, Beth A., 2961  
 Rubin, Gayle, 990–991, 992  
 Rucker, Darnell, 2217  
 Rueschemeyer, Dietrich, 389  
 RUGMARK program, 2607  
*Ruiz v. Estelle* (1980), 2054  
 Rule of law, 1549  
 Rules, bureaucratic enforcement of, 230–231  
*Rules, The* (Fein and Schneider), 488  
 Ruling class. *See* Social and political elites  
 Rumania. *See* Romania  
 Rumiantsev, A. M., 2980  
 Rumor, 348  
 Run (statistical), definition of, 1969  
 Ruocco, Maria Ricciardi, 1467  
 Rural sociology, **2425–2436**  
   agricultural innovation, 86–91, 677, 678, 2429, 2431–2433  
   American Indian demographics, 134  
   declining number of agricultural workers, 3264  
   Japanese sociology and, 1478, 1479, 1480, 1481, 1483  
   kinship structure and, 1502  
   mental illness incidence, 1840, 1841  
   Mexican studies, 1858–1860  
   “new rural sociology” theories, 2429, 2431–2433  
   research agenda, 2427–2431  
   and social networks, 2728  
   underemployment, 1721, 1722  
   urban societies contrasted with, 3192  
   Wageningen School approach, 2432–2433  
*Rural Sociology* (journal), 2427  
 Rural Sociological Society, 2427, 2433  
 Rusconi, Gian Enrico, 1471  
 Rush, F., 2581  
 Rushdie, Salman, 279  
 Rushton, Philippe, 2140  
 Russell, Bertrand, 225, 821, 1457  
 Russell, D., 2583  
 Russell, Dan, 194  
 Russell, Diana, 1275–1276  
 Russell Sage Foundation, 2299, 2398  
 Russett, B., 2917  
 Russia  
   caesaro-papist church-state relations, 2357, 2358  
   disaster research, 687  
   ethnonationalist movements, 1934, 1945, 1946, 1948  
   family size, 977  
   health-care system, 380  
   labor movement, 1532  
   legal system, 1549, 1550  
   life expectancy, 623  
   nationalist movements, 1945, 3001  
   nuclear weapons, 3242  
   organized crime, 2136, 2137  
   political and governmental corruption, 2136–2137  
   post-communist conditions, 2136–2137  
   protest movements, 2271  
   sexually transmitted diseases, 2591  
   sociology, 2116–2117, 2979–2984  
   *See also* Soviet and post-Soviet sociology; Soviet Union  
 Russian Empire, 2998  
 Russian Orthodox Church, 938, 2357  
*Russian Public Opinion Monitor* (journal), 2983  
 Russian Revolution (1917–1923), 2411, 2412, 2643, 2664, 2849, 2851  
   historical sociology study of, 1198  
   and Soviet sociology, 2979  
 Russian Sociological Society, 2979, 2982  
 Rutkevich, Mihail, 2980  
 Rutter, Michael, 2090  
 Rwanda  
   caste and inherited status, 254  
   colonial boundaries, 1934  
   genocide, 68, 107, 1066, 1069, 1070, 1072  
   peace mediation, 2048  
   poverty in, 2216  
 Ryan, Bryce, 87, 677, 2429  
 Ryan, R. M., 2059–2060  
 Rybicki, Pawel, 2120  
 Ryder, Norman B., 346–347, 1618, 2678  
 Rytina, Steven L., 1853–1854
- S**  
 S&L scandal, 3250  
 Saalfeld, Thomas, 2157  
 Sabatelli, Ronald, 1728  
 Sabatier, Paul, 1099, 1108  
 Sachs, Wolfgang, 1225  
 Sacks, Harvey, 247, 431, 432, 433, 435–437, 438  
 Sacred area preservation (Native American), 137–138  
*Sacred Canopy, The* (Berger), 2483  
 Sahlins, Marshall, 1033, 2891  
 Sahrer, Heinz, 1081  
 St. John’s wort, 654–655  
 St. Simon, Henri, 1309  
 Saint-Simon, Claude-Henri de Rouvroy, 1028–1025, 3202–3203  
 Saito, Y., 1632  
 Saksena, R. N., 1291, 1293  
 Salaries. *See* Wages and salaries  
 Salinas de Gortari, Carlos, 2135  
 Salinas de Gortari, Raul, 2135  
 Salvemini, Gaetano, 1465, 1466  
 Sambians, 2569  
 Same-sex attractions. *See* Sexual orientation  
 Same-sex coalitions, 331  
 Same-sex marriage, 111, 131, 315, 489, 1506  
   and endogamy norms, 1776  
   legal barriers, 2546  
 Samper, Ernesto, 2135  
 Sample selection bias, **2437–2444**  
 Sampling procedures, **2444–2449**  
   basic probability procedures, 2446–2448  
   and census, 283, 284–285, 286–287  
   cluster sampling, 2447–2448  
   estimators, 2440–2442

- General Social Survey, 578  
longitudinal research, 1687–1688  
models of selection, 2439–2442  
multistage, 2448, 2449  
and nonparametric statistics, 1956–1971  
and online surveys, 408  
probability theory, 2249, 2250, 3088  
and public opinion polls, 2273–2274  
random samples, 2274, 2288, 2324, 2397, 2446–2447, 3232  
and replication, 2397  
resampling methods, 3039  
sampling frame, 2445–2446  
and standard error, 2449, 3088  
survey research, 2444–2449, 3087–3088, 3232  
variations and tests against null hypothesis, 456–457
- Sampson, Robert, 365, 366, 514, 535, 536, 668, 1616
- Samuelson, Paul, 2338
- Samurai*, 253
- San Antonio, Texas, 2126
- Sanchez-Jankowski, Martí, 530
- Sanctions. *See* Criminal sanctions; Punishment
- Sandel, Michael, 104, 356
- Sanitation, 2177
- Sanneh, Kanja B. A. S., 381
- Sanskritization, 250–251
- Santer, Jacques, 2130
- Santeria, 65
- Sapir, Edward, 2890
- Sapir-Whorf hypothesis, 2092
- Saraceno, Chiara, 1472
- Saran, A. K., 1224
- Saris, William E., 2684
- Sarkar, B. K., 1291
- Sarney, José, 2135
- Sartori, Giovanni, 1470, 2153
- Sartre, Jean-Paul, 989, 1071
- Sarvodaya Movement, 3289
- SAS. *See* Society for Applied Sociology
- Satanic Verses, The* (Rushdie), 279
- SATs (Scholastic Aptitude Tests), 2439, 2497  
male vs. female mathematics scores, 2532
- Saud, king of Saudi Arabia, 2132
- Saudi Arabia, 1866  
abolishment of slavery in, 2602  
royal corruption, 2132  
sociodemographic profile, 2938
- Saul-Paul: A Double Life* (Hilbrand), 3053
- Saussure, Ferdinand de, 1032, 2205–2206, 2891
- Savigny, Friedrich Karl von, 475
- Savings and loan scandal, 3250, 3253
- Sawhill, Isabel, 3199
- Saxton, Stanley, 2221
- Sayrs, Lois W., 2680
- Scacco, A. M., Jr., 2578
- Scaife Foundation, 1601
- Scales. *See* Factor analysis; Levels of analysis; Measurement
- Scales for the Measurement of Attitudes* (Shaw and Wright), 3209
- SCAN (Sociology and Computers: a Newsletter)*, 407
- Scandals, political. *See* Political and governmental corruption
- Scandinavia  
“clean” government reputation, 2130  
cohabitation, 109  
comparative remarriage rates, 1749  
feminist theory, 990  
high suicide rate, 3079  
homosexual marriage legalization, 111  
labor movement, 1529, 1531  
life expectancy, 1631  
relative marriage rate, 1749  
*See also* Denmark; Finland; Iceland; Norway; Scandinavian sociology; Sweden
- Scandinavian Sociological Association, 2451
- Scandinavian sociology, **2449–2455**
- Scanning for optical character recognition, 408
- Scapegoating, 696, 1311
- Scarce, Rik, 839
- Scartezzini, Riccardo, 1472
- Scatterplots, 659, 3017–3019  
jittered, 3017–3018, 3019  
locally weighted, 3017  
matrix, 3018, 3020  
of residuals, 3016  
shape and direction of, 661  
univariate, 3015
- Scenarios (futures studies), 1041–1042, 2678
- Schaeffer, Nora Cate, 2707
- Schaeffle, Albert, 1074, 1423
- Scharping, Rudolf, 361
- Schattschneider, E. E., 2273
- Scheff, Thomas, 782, 783–784
- Scheffé, H., 3035
- Schegloff, Emanuel A., 247, 431, 432, 435–437, 438–439, 2904
- Schein, Edgar, 893, 895
- Scheler, Max, 2520–2521, 2953–2954, 2955, 2956, 2957
- Schelling, Thomas, 735
- Schelsky, Helmut, 1076, 1077, 1079
- Schelsky school, 1076, 1077, 1079
- Schemas, 2509, 2751
- Schematic processing, 2424
- Schengen Agreement (1985), 1934
- Schengen Convention (1990), 1934, 1935
- Scheuch, Erwin, 576, 1075, 1079, 1080, 3052
- Schilder, P., 1718
- Schizophrenia, 1836, 1838, 1839
- Schlesinger, Arthur M., 1580
- Schlick, Moritz, 821
- Schmid, Thomas, 245
- Schmidt, Calvin F., 3008
- Schmoller, Gustav von, 733, 1423
- Schnaiberg, Allan, 1214, 1229
- Schneider, David, 2891
- Schneider, H. K., 2670
- Schneider, Joseph, 2762
- Schneider, Mark, 3072, 3074
- Schneir, Miriam, 989
- Schoeck, Helmut, 1079
- Schoenwald, S. K., 76
- Schofield, R., 633

- Scholastic Aptitude Tests (SATs), 2439, 2497  
 male vs. female mathematics scores, 2532
- School systems. *See* Educational organization
- School violence, 1484–1485, 1487–1488, 1491
- School vouchers, 315
- Schooler, Carmie, 656, 1619, 2070–2071, 2072, 3271
- Schools. *See* Education; Educational organization; Education and mobility; Sociology of education
- School-to-Work Act of 1994, 763
- Schrecker, John, 2645
- Schroeder, David A., 118
- Schudson, Michael, 2173
- Schuessler, Karl F., 2299–2300
- Schultz, T. P., 633
- Schuman, Howard, 190–191, 317
- Schumpeter, Joseph, 474, 724, 733, 1266, 1356, 1424, 2164, 2273  
 on elite rule and mass apathy, 2627  
 and rational choice theory of democracy, 2339
- Schutz, Alfred, 226, 246, 856, 1479, 1932, 2099, 2756  
 “stranger” concept, 2635
- Schütze, Fritz, 1633, 1635
- Schwartz, B., 1998
- Schwartz, Barry, 1636
- Schwartz, Gary, 513
- Schwartz, Michael, 443
- Schwartz, Pepper, 359, 2539, 2540, 2542, 2555
- Schwartz, Shalom H., 3213, 3216–3217, 3218, 3222, 3223
- Schwartz, Sharon, 2190
- Schwartz Scale of Values, 3216–3217, 3218, 3219, 3222, 3223
- Schwarz, Frederick C., 2370
- Schwendinger, Herman and Julia, 511–512
- Science, **2455–2463**  
 applied research and development, 2460–2461  
 case studies, 2455  
 controversial aspects of, 2461  
 disciplines, 2458  
 economic determinism and, 722  
 historical, 2456–2457  
 Kuhn’s model of change in, 2024–2027, 2193, 2458–2460, 2756  
 meta-analysis of sex differences in achievements in, 2532–2533  
 normal vs. revolutionary, 2458–2460  
 normative ethos, 2456  
 philosophy of, 1077  
 and postindustrial society, 2197  
 and postmodern society, 2206, 2207–2208  
 pragmatic view of, 2217, 2218, 2219  
 research fraud and, 2458  
 social stratification in, 2457  
 sociological definitions of, 2455  
 and sociology of culture, 563  
 and sociology of scientific knowledge, 2458–2460  
 women’s careers in, 2786  
 work groups, 2457–2458  
*See also* Scientific explanation
- Science Indicators* (journal), 2685
- Science sociale, La*, 1025
- Scientific communities, 2025–2026
- Scientific explanation, **2463–2473**  
 building on existing data, 2465  
 and causal inference models, 255–256  
 critiques of hypothetico-deductive model, 2468–2472  
 deductive vs. inductive logic in, 2465–2466  
 deterministic laws and, 2465  
 Enlightenment, 2206, 2207, 2208  
 falsification and, 2464, 2466, 2470–2471  
 German sociology and, 1077  
 materialism and, 1781, 1785  
 and mathematical sociology, 1786–1791  
 measurement and, 2467  
 and metatheory, 1852–1854  
 objectivity in, 2467  
 and philosophy of science, 1077  
 positivism and, 2194–2195  
 postmodern, 2207–2208  
 and prediction, 2226–2233  
 replication and, 2395–2397, 2471  
 and social philosophy, 2756, 2757  
 stochastic laws and, 2465  
 theory testing in, 2466–2467
- Scientific management theory, 697
- Scientific method. *See* Epistemology; Scientific explanation
- Scientific rationalism, 1085, 1086, 1247
- Scientific realism, 822–823
- Scientism, 544
- Scientology, 900, 2366, 3287
- Sciolla, Loredana, 1472
- SCLC. *See* Southern Christian Leadership Conference
- Scope conditions, 671
- Scopes, John, 2369, 2370
- Score construction, 2346
- Score estimation, and factor analysis, 914–915
- Scores. *See* Factor analysis; Levels of analysis; Measurement
- Scotland  
 civil law, 465  
 divorce law reforms, 703  
 Enlightenment in, 2335, 2340
- Scott, James C., 2977
- Scott, Joan, 1708–1709
- Scott, Kesho Y., 57
- Scott, W. A., 3214
- Scott, W. Richard, 2004
- Scottish Enlightenment, 2335, 2340
- Scripps Foundation, 637
- SDA Archive (interactive data access system), 409
- SDS. *See* Students for a Democratic Society
- Seal, B. N., 1291
- Sears, R. R., 73
- Seaside resorts, 3168
- Seasonal Variations of the Eskimos: A Study in Social Morphology* (Mauss), 1032
- Seattle-Denver Income Maintenance Experiment (SIME-DIME), 2213
- Second International (Marxist Workers Congress), 723



- Second Sex, The* (Beauvoir), 988, 989, 990
- Second Vatican Council (1962–1965), 2365
- Secondary data analysis and data archives, **2473–2482**  
 advantage and disadvantage, 2477–2478, 2480–2481  
 data archives, 2473–2477  
 and Durkheim’s suicide study, 574, 1595, 3080  
 nature of secondary analysis, 2477–2481  
 replication, 2481
- Secondary jobs, 1985
- Secondhand smoke, 1640
- Second-order partial, 451
- Sects, formation of, 2364, 2365, 2366–2367, 2378
- Secularization, **2482–2491**  
 early and recent concepts of, 2482–2489  
 Enlightenment and, 2965–2966  
 myths of, 2484–2485, 2486, 2966  
 paradoxes of, 2487–2489  
 political correctness and, 2142  
 of religion, 935, 937–938, 2373, 2374  
 values theory and, 3222
- Security, 3216
- Sedition Act of 1798, 273
- Sedition Act of 1917, 273
- Seditious libel laws, 273
- Seduced by Death* (Hendin), 3084–3085
- Segal, David and Mady, 1876
- Segerstedt, Torgny T., 2450, 2451
- Segregation and desegregation, 2334, **2491–2500**  
 African American studies, 54, 56, 57, 58, 66  
 in American cities, 308, 532  
 in American society, 143, 145  
 black-initiated return to segregated programs, 2498–2499  
 and coercive labor practices, 2608  
 and crime rates, 532  
 desegregation vs. integration, 2495–2496  
 and fears about interracial sex, 1408  
 health care access and, 1152  
 in housing, 1152  
 Jim Crow laws, 54, 62, 2491, 2761  
 occupational, 3264–3265  
 occupational by sex, 370  
 protest movements, 2269  
 recognition as social problem, 2761  
 and residential inequality, 56, 57, 143, 250, 845, 2500–2504  
 and social impact of legal changes, 2962  
 South African apartheid, 62, 250, 1940, 2047, 2146, 2725  
 Southern desegregation initiatives, 2493, 2494–2495  
 Supreme Court rulings, 145, 2491, 2493, 2494  
 from urban population shifts, 3072, 3196  
*See also* Civil rights movement
- Segregation indices, 57, 2500–2505
- SEI (Socioeconomic Index), 1997, 2000, 3265
- Seidler, John, 2379
- Seidman, Robert, 1497
- Seidman, Steven, 2207–2208, 2220
- Seigfried, Charles Haddock, 2220, 2221
- Seik, Eikichi, 1478
- Selection models. *See* Sample selection bias; Sampling procedures
- Selective fallacy, 1593
- Selective serotonin inhibitors (SSRIs), 654
- Self. *See* Identity theory; Self-concept
- Self-actualization, 1304, 1714, 1715, 1718, 2085, 2087–2088
- Self-appraisal. *See* Self-concept
- Self-concept, **2505–2510**  
 attitudes reinforcing, 185, 2508  
 and childhood sexual abuse, 290  
 cognitive dissonance theory and, 339  
 cognitive-distortion model of depression, 651  
 consequences of, 2508–2509  
 development of, 2506–2508  
 German idealism and, 1248–1249  
 hierarchy of roles in, 2505–2506  
 identity theory, 1254–1255  
 inaccuracies in, 2750  
 and internalized values and norms, 2837  
 interpersonal relations and, 2506–2508, 2773, 2774  
 James (William) on, 2083  
 “looking-glass self” concept, 2089, 2344, 2507, 2512, 2750, 2856  
 performance relationship, 2508–2509  
 personality overlap with, 2082, 2083  
 and relationship with community, 357  
 role theory on, 2423–2424, 2506, 2507–2508  
 self-esteem and, 2506, 2508, 2511, 2512  
 self-esteem vs., 2511  
 self-perception and, 2750  
 situational, 42–44, 2506  
 social comparison theory on, 2507, 2650, 2651, 2653, 2856  
 and social self, 2218  
 socialization as formation of, 2856, 2860, 2862  
 stable, 2505–2506  
 symbolic interaction theory on, 2423, 2856, 3095–3100
- Self-confidence, self-esteem vs., 2511
- Self-control  
 aggression and, 74  
 crime theory and, 667  
 as illusion in positive mental health, 2190  
 socialization and, 2856
- Self-determination. *See* Personal autonomy
- Self-efficacy, 116, 1895  
 definition of, 2750  
 self-esteem vs., 2511  
*See also* Alienation
- Self-employment, Social Security coverage, 2798
- Self-esteem, **2511–2518**

- aggression and, 72, 2516–2517, 2777
- consequences of high or low, 2514–2517
- consistency of, 2347–2350
- definition of, 2511, 2750
- depression and, 651
- global, 2511–2512
- “looking-glass self” concept, 2089, 2344, 2512
- measurement protocols, 2344–2350, 2512
- multidimensional trait model, 2511
- peer group as source of, 2513, 2860
- persuasion and, 2096
- role theory and, 2424
- school performance and, 2859
- and self-concept, 2506, 2508, 2511
- self-expectations adherence and, 2837
- sex differences in, 2534
- social comparison and, 2654
- sources and functions of, 2512–2514
- Self-fulfilling prophecy
- and academic achievement, 2932
  - and dangers of forecasting, 2676
  - deviant labels and, 1577
  - sex roles and, 2530
- Self-help movements, 2717
- Self-identity theory, 2750
- Self-image. *See* Self-concept
- Self-interest
- prejudice and, 2245–2246
  - rational choice theory and, 2338, 2339
  - social vs. self-control and, 667
- “Selfish gene” metaphor, 2882
- Selfishness, altruism vs., 2882, 2883–2884
- Self-perception, 893
- definition of, 2750
  - dissonance theory and, 339
- Self-presentation behavior. *See* Affect control theory and impression formation
- Self-regulation. *See* Self-control
- Self-report measures
- depression measurement, 653, 654
  - juvenile delinquency and crime, 1488, 1490, 1492
  - life history, 1616
  - self-esteem, 2516
  - voting behavior, 3232
- Self-schemas, 2509
- Seligman, Brenda, 1271
- Seligman, Martin., 651, 652
- Sellin, Thorsten, 664, 3247
- Selye, Hans, 3055
- Selznick, Philip, 356, 566
- SEM. *See* Structural equation modeling
- Semantic differential technique, 186
- Semigroup concept, 2298
- Semiotics, 2958
- Semmelweis (Hungarian physician), 2464
- Semyonov, Moshe, 643
- Sen, Amartya, 1220
- Seneca Falls convention (1848), 989
- Senegal
- poverty in, 2216
  - sociodemographic profile, 2938
- Senghor, Léopold Sédar, 66
- “Senior boom,” 345
- Sennacherib, 1069
- Sennet, Richard, 542
- Senofonte, 2520
- Sensenbrenner, Julia, 2639–2640
- Sentimental novels, 2171, 2172
- Sentiments, **2518–2529**
- affect control theory and, 43–44, 45
  - attitudes and, 187
  - norms of, 2527–2528
  - postmodernism and, 2528–2529
  - social impact of, 2524–2526
  - See also* Emotions
- “Separate but equal” ruling, 2491, 2492
- Separation of powers, 473–474, 1952, 2624
- Sequential games, 330
- Serbia, 2362, 2363
- genocide by, 68
- nationalism, 3001
- use of wartime rape, 2579
  - See also* Kosovo
- Serendipitous model of kinship mapping, 1514–1515
- Serfdom, 1811, 2597, 2602
- Serotonin, 70, 652, 654, 3079
- Servants, 3262
- Service, Elman R., 2891
- Service employment
- independent contractors, 3268
  - in Mexico, 1860
  - postindustrial shift to, 2196, 2197, 2199, 2430, 3276
  - working mothers as factor, 3266
  - See also* Informal economy
- Servile labor. *See* Slavery and involuntary servitude
- Servile marriage, 2602
- Set theory, 2296
- Settlement houses, 365–366, 2841
- Seventh-Day Adventism, 2366, 3287
- Sevigny, Robert, 328
- Sewell, William, 1854, 2427, 2714, 2782
- Sex and Temperament in Three Primitive Societies* (Mead), 998
- Sex differences, **2529–2537**
- contextual influences, 2534–2535
  - feminist theory denial of, 2886
  - gender roles and, 2418
  - See also* Gender roles
  - gender terminology and, 1057–1058, 2529, 2886
  - infant mortality rates, 222
  - life expectancy, 126, 1630–1631, 1827, 2177, 2180
  - meta-analysis of four established beliefs on, 2530–2532
  - mortality rate, 222–223
  - reproduction potential, 2884
  - sociobiological law of anisogamy and, 2884–2886
  - theoretical accounts of, 2530
  - See also* Femininity/masculinity; Gender; Sex stereotypes
- Sex discrimination
- occupational, 370
  - sexism and, 988, 989, 1838

- sexual harassment as, 1063, 1880–1881, 2580–2581
- UNESCO report on global conditions of, 1767–1768
- See also* Comparable worth; Occupational segregation; Sex stereotypes; Women’s rights movement
- Sex in America* (Michael et al.), 2550
- Sex offenses
- civil liberties and, 317
  - sexually explicit materials and, 2187
- Sex ratio
- Asian immigrants, 123, 176
  - marriage squeeze, 1775
- Sex research. *See* Sexual behavior patterns
- Sex Roles* (journal), 2418
- Sex stereotypes
- attribution theory, 197–198
  - conformity and, 403
  - femininity/masculinity, 997, 2568
  - gay and lesbian sexual behavior, 2554–2555, 2556
  - mass media, 1699, 1767–1768
  - and mental health theories, 2190
  - rape-prone societies and, 2591
  - sex differences and, 2530, 2534
  - and sexual behavior, 2569
- Sex telephone services, 2185
- Sex tourism, 2560–2561, 2607
- Sexism. *See* Sex discrimination
- Sex-role models. *See* Gender; Gender roles; Femininity/masculinity; Role theory; Socialization
- Sexual abuse. *See* Childhood sexual abuse
- Sexual assault. *See* Rape; Sexual violence and exploitation
- Sexual behavior in marriage and close relationships, **2537–2549**
- gay men and lesbian relationships, 2545–2547, 2554–2555
  - infidelity studies, 2541–2545
  - interpersonal attraction and, 2537–2548
  - love and, 1697
  - rape and, 950, 2577–2578
  - theoretical frameworks, 2537–2538
- Sexual behavior patterns, **2549–2564**
- approaches to, 2537
  - in close relationships, 2537–2548
  - consent and coercion, 2556–2559, 2577–2580
  - See also* Rape
  - descriptive characteristics of, 2538–2541
  - extramarital/extradynamic, 2541–2545
  - frequency of sex, 2539, 2553, 2555
  - in gay and lesbian adolescents, 2552–2553
  - in gay male and lesbian relationships, 2539–2540, 2545–2547, 2554–2555, 2561, 2570
  - incest and, 1270–1277
  - and Islamic control over women, 2949–2950
  - in life course, 2550–2556
  - love and, 1697
  - number of partners, 2540, 2553
  - oral sex prevalence, 2553, 2569
  - outside of marriage, 487, 1734, 2568
  - peer group socialization, 2859
  - premarital, 2568
  - premarital rates in China, 302
  - risk reduction, 2579–2581
  - risk-taking, 1641–1642, 2559, 2585–2593
  - same-sex relationships not identified as “homosexual/bisexual,” 2554
  - serial monogamy and, 2553
  - sexual satisfaction and, 2540
  - in single adults, 2553–2555
  - sociobiological law of anisogamy on, 2884–2885
  - See also* Prostitution; Sexual orientation
- Sexual development
- Freudian stages of, 1713
  - of gay men and lesbians, 2570
  - See also* Infantile sexuality theory
- Sexual harassment, 1063, 2580–2581
- Hill-Thomas case, 2581
- and political correctness, 2141
  - suit against Clinton, 2128, 2581
  - two forms of, 2581
  - and women in the military, 1880–1881
- Sexual material. *See* Pornography
- Sexual orientation, **2564–2575**
- and adolescent sexual attraction, 2552–2553
  - AIDS/HIV and, 1641, 1642, 2559, 2570, 2585–2586, 2587, 2588, 2590, 2591
  - AIDS/HIV risk reduction, 2589
  - bisexuality, 2571–2572, 2577, 2581
  - case studies, 245
  - character of relationships, 2545–2547
  - and childhood sexual attractions, 2551
  - civil liberties issues, 315
  - conceptualization changes, 2564–2565, 2570
  - current concepts of, 2567–2573
  - declassification as mental disorder, 1833
  - disease model, 2567
  - family arrangements, 1064
  - gay and lesbian relationships, 111–112, 489, 2539–2540, 2545–2547, 2554–2555, 2570
  - gay men vs. lesbian frequency of sex, 2539, 2555
  - and gender, 2565–2567
  - and hate speech, 275–276
  - heterosexuality, 2567–2569
  - homosexuality, 1234, 2569–2571
  - and legislation of morality, 1580
  - and life-course sexual behavior, 2551, 2552–2553
  - military policies on, 1878, 1879, 1881
  - political correctness and, 2140
  - and prostitution, 2559, 2560, 2561
  - rape and, 2578
  - religious organizations and, 2380
  - same-sex marriage trend, 111, 131, 315, 489, 1506, 1776

INDEX

- sexual behavior patterns, 2539–2540, 2545–2547, 2554–2555, 2556, 2561
- sexual coercion studies, 2557
- sexual risk-taking, 2559
- singlehood demographics and, 125
- transvestism and transsexuality, 2554, 2572–2573
- See also* Femininity/masculinity
- Sexual Politics* (Millet), 2587
- Sexual revolution, 484, 487, 2585
- Sexual selection (SS) corollary, 2885
- Sexual strategies theory, 2537
- Sexual violence and exploitation, 1064, 2557, 2559, **2576–2584**
- African-American slave experience, 54, 62, 1407–1408
- aggression-prone attitudes and, 72, 2591
- censorship and, 275
- consent and coercion, 2556–2559
- and criminalization of deviance, 523
- feminist theory on, 991
- incest and, 991, 2557, 2559, 2594
- males as victims, 2578
- prostitution and, 2560–2561
- sex industry and tourism, 2560–2561
- treatment and prevention, 2583
- women as victims, 2587–2589
- See also* Childhood sexual abuse; Incest; Rape; Sexual harassment
- Sexuality. *See* Sexual behavior patterns; Sexual orientation
- Sexually explicit content. *See* Pornography
- Sexually transmitted diseases, **2585–2595**
- effect on courtship, 487, 489
- and fertility determinants, 1010
- as lifestyle risk, 1641–1642, 2559
- See also* AIDS/HIV
- Shagari, Shehu, 2134
- Shakespeare, William, 2091
- Shalin, Dimitri, 2217
- Shame, 2521, 2529, 2773
- Shame of the Cities, The* (Steffens), 2124, 2126
- Shaming, as social control, 519, 520
- Shanahan, Suzanne, 2662
- Shanas, Ethel, 1389
- Shango, 65
- Shannon, Thomas, 1706
- Shapin, Steven, 2456
- Shapiro, Robert Y., 2276
- Shapiro, Susan, 2962, 3251, 3252
- Sharecroppers, 2492
- Shared good. *See* Communitarianism
- Shared-values school, 2337
- Sharia. *See* Islamic law
- Sharif, Nawaz, 2132
- Sharot, Stephen, 2968–2969
- Shaver, P. R., 2068, 3209
- Shavitt, S., 2098
- Shaw, Clifford, 324, 665, 1495, 1633
- Shaw, M., 3209
- Sheatsley, Paul, 2243
- Sheldon, Eleanor Bernert, 2682
- Sheldon, William, 1717–1718
- Shell International Petroleum Company, 1041–1042
- Sheltered housing, 1655–1656
- Sherif, Muzaffer, 2415
- conformity experiment, 401, 402, 404, 2611, 2616
- Shi'a Islam, 3286
- Shils, Edward A., 225, 836, 2632, 2954, 3213, 3214
- Shilts, Randy, 2585
- Shimmei, Masamichi, 1478
- Shine, M., 1628
- Shinn, Larry D., 3279
- Shiva, Vandana, 1230
- Shively, W. Phillips, 1595
- Shlapentokh, V., 2980
- Shneidman, Edwin S., 3077
- Shoemaker, Floyd, 678
- Shogunate, 253
- Short, James, 1488, 3079, 3080
- Short-Form 36 Health Survey, 2306
- Shostak, Arthur B., 107–108
- Shott, Susan, 782–783
- Shubkin, Vladimir, 2980
- Shudras (Hindu servants), 250
- Shupe, Anson, 2372, 2379
- Shweder, Richard, 1900, 1902
- Siblings
- in blended families, 2391–2392
- in coalitions, 332
- incest, 1273, 1275
- unity vs. marital unity, 1509
- See also* Family size
- Sicily, 2128
- Sick role (Parsons concept), 1813, 1815
- Siddhartha Buddha, 3284–3285
- Sidney, Stephen, 57
- Siegel, Fred, 316
- Siegel, Sidney, 1957
- Siegfried, André, 3232
- Sierra Club, 790, 802, 803
- Sighele, Scipio, 1464, 1465
- Sigler, Jay A., 1945
- Sign test (nonparametric), 1961
- Signs and symbols, postmodern theory on, 2205–2206
- Sikes, Melvinn, 57
- Sikhism and Sikhs, 2366, 2486, 3287–3288
- Silent Spring* (Carson), 789, 803
- Silpa-Archa, Banharn, 2131
- Silver, A. K., 288
- Silverman, David, 247
- Silverman, F. N., 288
- SIME-DIME study, 2213
- Simiand, François, 734, 1024, 1891, 2917
- Similarity hypothesis, 2650
- Simmel, Georg, 178, 581, 733, 820, 852, 1028, 1074, 1116, 1117, 1305, 1313, 1423, 2693
- “blaseization” concept, 2528
- as Blau social exchange theory influence, 2671, 2672
- as Chicago School influence, 1772
- coalitions perspective, 329, 2611
- conflict theory, 415
- frames concept of, 1932
- on individual identity and social belonging relationship, 2633–2634
- as Japanese sociology early influence, 1477, 1478

- on money's significance, 1888, 1889
- music theory, 1924
- on role differentiation, 696, 698
- on sentiments, 2520
- and social inequality concept, 2690
- as sociology of knowledge basis, 2954, 2955
- "stranger" concept, 2635
- on territorial belonging, 3129
- on unstability of triads, 465
- Simon, Herbert, 3035
- Simon, Theodore, 2330
- Simple random sampling (SRS), 2446-2447
- Simpson, Ada Harper, 150-151, 156
- Simpson, Elizabeth L., 1900
- Simpson, Richard, 150-151, 156
- Simpura, Jussi, 2451
- Simulation and Games* (journal), 407
- Simulation models
- computer-assisted, 410, 3039
  - forecasting and, 2678-2679
  - standardization and, 2996
- Simultaneous equation model, 261-264, 2251, 3035
- Sinclair, Upton, 245
- SINET: Social Indicators Network News* (quarterly), 2685
- Singapore, 2130, 2974
- authoritarian
  - communitarianism, 356
  - dependency theory, 642
  - fertility transition, 1008, 2976
  - labor movement, 1532
- Singer, Burton, 1684, 1689, 1691
- Singer, Milton, 564
- Single Convention on Narcotic Drugs (1961), 713
- Single European Act (1987), 1550, 2151
- Single Market (EU; 1993), 1934
- Single mothers. *See* Single-parent households; Unwed childbearing
- Singlehood, 107-108
- childbearing and, 125, 128, 484, 488, 634, 1258-1264, 1506, 1626, 1744, 2033
  - demographics, 124-125
  - divorce and, 112
  - gender and, 107-108, 108, 124, 125
  - homosexuality and, 111, 124
  - marriage squeeze factor and, 1775-1776
  - never-married adults, 107-108, 124-125, 487-488, 1738
  - race and, 107-108
  - sexual behavior patterns, 2553-2555, 2556
  - suicide rate, 3078
  - voluntary/involuntary, 107-108
- Single-indicator models, 1910-1912, 1922
- Single-item scale
- for attitude, 185-186
  - for quality of life, 2303, 2307
- Single-member representational districts, 2154, 2157, 2164
- Single-parent households, 142, 359, 487, 488, 1625, 2033
- African-American, 122
  - as choice, 1506, 1626
  - from divorce, 112, 128, 708, 1747, 1749, 2033
  - effects on children of, 705, 706, 707-708, 3062
  - nonmarital births, 125
  - from out-of-wedlock births, 634, 708, 2033
  - poverty level and, 2215
  - poverty rate of, 127, 128, 3062
  - rise in rate, 127
  - See also* Unwed childbearing
- Sinha, J. B. P., 1566
- Sirowy, Larry, 423
- Situational analysis. *See* Life histories and narratives
- Siu, Paul, 176, 178
- Six-Day War (1967), 3139
- Sixteenth Street Baptist Church (Birmingham, Ala.) bombing, 2495
- Siyassat Nameh* (Nezam Mulk Tussi), 1564
- Skadden, Arps, Slate, Meagher & Flom, 469
- Skills standardization, 398
- Skinheads, 460, 461
- Skinner, B. F., 722
- depression postulate, 650
  - operant reinforcement theory, 209, 210, 1716, 2085
  - utopian design, 3204
- Skinner, Denise A., 2417
- Sklair, Leslie, 227, 1085
- Skocpol, Theda, 416, 426, 1102
- and comparative-historical sociology, 383, 384, 385
  - historical sociology study, 1196, 1198
  - on Iranian revolution, 1870
  - and macrosociology, 1704, 1707, 1708, 1709
  - and metatheory, 1852
  - political sociology theory, 2163, 2164-2165
  - on social dynamics, 2664
  - and voluntary associations, 3229, 3230
- Skoe, Eva E., 1900, 1903
- Skolnick, Jerome, 2108, 2114
- Skvoretz, John, 2029
- Slater, P. E., 696
- Slater, Philip J., 1734, 1974
- Slavery and involuntary servitude, **2596-2610**
- and abolitionist movement's success, 2725
  - in Africa and Middle East, 2501-2503, 2604-2608
  - by African Americans, 54-56, 57-58, 64, 121-122, 320, 321, 2333, 2491-2492, 2599-2601
  - in the Americas, 2598-2601
  - in Asia, 2602-2603, 2604-2608
  - Asian sex trade as, 2607
  - as background to contemporary African American problems, 2333
  - British sociological study of modern, 228
  - capitalism and, 238, 239, 321
  - definition of, 2596
  - examples in 1998 (table of), 2604-2608
  - forced child labor, 2605-2607, 3262
  - as human rights issue, 1238

- imperialism and, 320–321  
international efforts against, 2607  
interracial marriage and, 1407–1408  
and land-labor ratio hypothesis, 2596, 2597–2598, 2600–2601, 2607–2608  
marriage and family structures, 121–122  
patterns since 1990, 2603–2608  
penal, 518, 2603, 2608  
slavery vs. involuntary servitude, 2596–2597  
stratification parameters, 2810–2811  
and systemic power, 2166  
Sleeper, Jim, 53  
“Slippery slope” argument  
    assisted suicide and, 587  
    civil liberties and, 318  
Slit-half consistency, 2347–2348  
Sloan Kettering Hospital, 588  
Slomczynski, Kazimierz M., 2070  
Slovak Republica, 1934, 1946, 1947  
Slovenia, 1941, 2136, 2362  
Slum neighborhoods, community studies, 365  
Small, Albion, 1423, 1424, 2192  
*Small Group Behavior* (journal), 1980  
*Small Group Research* (journal), 1979, 1980  
Small groups, **2610–2622**  
    case studies, 244–245  
    censuses, 283–284  
    coalitions within, 330  
    cohesion and productivity, 2516, 2614–2615, 2619  
    conformity in, 2615–2617  
    cooperation and competition within, 2620–2621  
    division of labor within, 696  
    games and, 330  
    group dynamics and, 1014, 2611  
    influence in, 2615–2617  
    interaction and performance, 2617  
    interpersonal systems, 2775–2777  
    leadership in, 2619–2620  
    mathematical formulations, 1790  
    minority influence in, 2616–2617  
    and nonmember responses, 2777  
    observation methods, 1974–1981, 2612–2613  
    perceptions of inequity in, 2777  
    persuasion in, 2776  
    polarization, 2617  
    primary and secondary, 2610–2611  
    research approaches, 2612–2614  
    role assignments, 695, 2422–2423, 2774–2775, 2776–2777  
    and social comparison theory, 2650, 2652–2654  
    and social dynamics, 2666  
    social psychology studies on, 2770–2771, 2775–2777  
*Small Social Structures: An Introduction to Structural Microsociology* (Szmatka), 2120–2121  
Smart, Niman, 3278–3279  
Smeeding, Tim, 2452  
Smelser, Neil, 553, 733, 1885, 2069, 2914  
    collective behavior theory, 352–353, 2635  
    and economic sociology, 2921  
    on industrialization, 1197  
    on nationalist movements, 1941–1942, 1943  
Smelser, William T., 2069  
Smith, Adam, 2335, 3098, 3243  
    on division of labor, 697, 1782  
    economic exchange theory, 721, 722, 2337  
    and individualism, 1302, 1305, 1306  
    industrial sociology and, 1308, 1309  
    and liberalism, 355, 1597  
    mass society theorists and, 1771  
    social development theory of, 2340  
    on social justice, 2698  
    on sympathy, 782  
Smith, Anthony, 227, 1941, 1943  
Smith, Brian T., 2961  
Smith, Cyril, 225  
Smith, David H., 1885–1886  
Smith, Dorothy E., 246, 2958  
Smith, Dwight C., Jr., 2018  
Smith, Elliot, 675  
Smith, Gerald L. K., 2370  
Smith, J. H., 225  
Smith, Joseph, 3287  
Smith, Louis, 1636  
Smith, Michael, 1214, 1215  
Smith, Peter B., 3213, 3214, 3216  
Smith, Thomas, 2243  
Smith, Tom W., 3091  
Smith Richardson Foundation, 1601  
Smith-Lever Act of 1914, 24126  
Smith-Lovin, Lynn, 1790  
Smithsonian Institution, 1039  
    Native American artifact repatriation, 137  
Smoking  
    illicit drug use correlation, 711  
    peer influence and, 667  
    as preventable health hazard, 1639–1640  
    smokers’ cognitive dissonance on, 338  
Smolenska, M. Zuzanna, 118  
Smolla, R. A., 268  
SMOs. *See* Social movement organizations  
Smuts, Jan Christian, 2088  
SNCC. *See* Student Nonviolent Coordinating Committee  
Snedecor, George W., 3035  
Sniderman, Paul M., 2245  
Snow, David, 554, 1237  
Snow, John, 814  
Snowball samples (public opinion polls), 2445  
Snyder, David, 2270  
Sober, Elliott, 118  
Social action, 1306  
    components of, 352  
    participatory research and, 2038–2040  
    theory of social resources and, 2790–2791  
*Social and Cultural Dynamics* (Sorokin), 2662  
*Social and Cultural Report* (Netherlands), 2685  
Social and political elites, **2622–2629**

INDEX

- and authoritarian state  
development, 2999–3000
- bases of, 2163
- classical elite theory, 2623–2624
- conflict mobilization, 415
- as conflict theorist deviance  
factor, 670, 2660
- as conformity influence, 403–404
- conservative ethos and,  
1598–1599
- critical/radical theorists,  
2624–2626
- cultural transmission by, 2171
- definition of elites, 2622
- democracy and, 2164, 2627
- and fertility, 1007
- as governing class, 604
- and hierarchical inequality, 2693
- Jewish liberal ethos and, 1603
- limitations of elite paradigm,  
2627–2628
- as Marxist deviance theory  
factor, 670
- mass media control by, 2166
- mass society theory on,  
1772–1773
- and pluralistic representation,  
2164, 2624–2626, 2627
- and political correctness,  
2139–2142
- political power development by,  
2162–2163, 2163–2164
- and popular culture, 2168–  
2169, 2171
- public opinion polling and,  
2278–2279
- religious orientation and, 2385
- revolution and, 2411–2412, 2414,  
3000–3001
- and social movements, 2717,  
2725
- Soviet Communist Party  
*nomenklatura*, 2850
- and state development,  
2997–3002
- and status attainment, 2787
- and status incongruence,  
3053–3054
- structuralist political sociological  
view of, 2162–2163
- in tribal societies, 2809–2810
- utopian analysis and design,  
3202–3203
- voting behavior and, 604  
*See also* Status attainment
- Social anthropology. *See*  
Sociocultural anthropology
- Social anxiety, 2512
- Social attributions, 197–198
- Social behavior, 194, 1254, 1255,  
1274, 1824
- exchange theorists on, 2669–2675
- observation systems, 1973–1981
- and other components of human  
action, 2632–2634
- social comparison and,  
2653–2654
- sociobiology and, 288087
- “Social Behavior as Exchange”  
(Homans), 2670
- Social Behavior: Its Elementary Forms*  
(Homans), 722, 2670
- Social belonging, **2629–2637**
- community and, 2520, 2521
- conformity and, 403–404, 2630
- definition of, 2630
- deviant behavior in absence of,  
535, 667, 668
- inclusion and, 2365, 2633
- marginalization and, 2367,  
2634–2635
- membership and nonmembership  
groups, 2634–2635
- participation definition, 2635
- structure of, 2630–2632
- territorial belonging and,  
2629–2630, 2632, 3129,  
3131–3132
- Social capital, **2637–2641**
- in adulthood, 35
- and career advancement, 1993
- community social reform  
and, 366
- definitions of, 366, 2637, 2639
- delinquent subcultures’ lack  
of, 513
- and educational attainment, 2930
- forms of, 2638–2639
- negative effects of, 2640
- social networks as, 1993, 2637,  
2639, 2640, 2732, 2733–2734
- and upward mobility, 2714,  
3043–3044
- Social casework, 2841, 2842, 2843
- Social categorization, 197
- Social change, **2641–2649**
- adolescence and, 14
- anomie and, 164
- attitudes as indicators of,  
190–191
- and British sociology, 225
- communitarian view of, 359–360
- conservative view of, 1599
- convergence theories, 421–428
- countercultures and, 461
- culture lag and, 2644
- cumulative effects, 2643–2648
- demographic change in  
relationship with, 632,  
703–704, 2642, 2643
- discontinuous, 2643, 2676–2677
- family religious orientation and,  
936–942
- globalization and, 1084
- and Italian sociology, 1471–1473
- in kinship structure, 1502–1503
- Kuhn’s model of, 2024–2027,  
2756
- macrosociological theory and  
research, 1705
- main sequence theory of, 1502
- Marxist view of, 3066
- Mexican studies on, 1857,  
1858–1859, 1861
- model for directed, 2687–2688
- no-fault divorce and, 703–704
- parental characteristics and,  
2032–2033
- race and class issues and,  
322–323
- religious organizations and,  
2379–2380
- social dynamics vs., 2663–2665
- social forecasting and, 2676–2680
- social indicators analysis,  
2681–2688
- and social movement  
emergence, 2718

INDEX

- social movements and, 2717–2725
- social work and, 2842
- structural lag and theories of, 3066–3067
- technology and, 2680
- values and openness to, 3216
- violence reduction suggestions, 76
- Social Change in the Industrial Revolution* (Smelser), 1197
- Social Choice and Individual Values* (Arrow), 2920
- Social citizenship, 1304
- Social class
- adulthood and, 27, 33
  - aging and, 1138
  - alcohol use and, 94
  - alienation and, 102–103
  - American families and, 122
  - anomie and, 1491, 1494
  - attributional patterns, 196
  - authoritarian personality and, 317
  - caste and inherited status, 249–255
  - church-sect typology and, 2378
  - community studies findings, 363–364
  - conflict theory, 414–416, 536–537
  - convergence theory, 424–425
  - courtship endogamy, 484, 486 and criminal and delinquent subcultures, 513
  - cultural studies and, 2169–2173
  - cultural taste differences, 1648–1649
  - depression/stressor relationship and, 656
  - education and, 742–746
  - educational attainment and, 2930–2931
  - emotional depression and, 655
  - family size and, 975
  - fertility change and, 1009, 3043
  - fundamentalism and, 2378
  - gendered differences on romantic love and, 1700
  - gradational groupings, 2816
  - health behavior and, 1129
  - higher education and, 1179
  - historical variations in, 1198, 1199
  - illegitimacy and, 1261
  - incongruence factor, 3049–3055
  - industrial psychology and, 1309
  - inequalities and, 1218, 2689–2694
  - interrelationship with race. *See* Class and race
  - juvenile crime and, 1491–1492, 1494
  - kinship and family systems and, 1511–1512
  - labor movement orientation and, 1529–1530
  - language use and, 2901, 2908
  - legal autonomy theory and, 1556 and legislation of morality, 1576, 1577, 1579
  - leisure and, 1584–1585, 1588
  - liberal/conservative orientation and, 1597, 1602–1603, 1603
  - life expectancy and, 1138, 1631
  - marital communication patterns and, 1736
  - Marxist analysis of, 1753, 1754–1755, 1756, 2162–2163, 2692, 2812
  - Marxist legal theory and, 1553, 1576
  - mate selection and, 1775–1776
  - measurement devices, 2866–2867
  - mental disorder epidemiology and, 1835–1840
  - mobility and, 2711–2716
  - music and, 1926
  - and “new middle class” emergence, 2815
  - and occupational prestige ranking, 2000, 3265
  - oppositional frameworks for, 2692
  - Polish sociology and, 2120
  - political organization membership and, 2150
  - popular culture transmission and, 2168–2173
  - poststructuralist view of, 1756
  - professions and, 2259–2265
  - religious organizations and, 2378, 2380
  - religious orientation and, 2384–2385
  - self-concept of adults and, 2508
  - socialization emphasis and, 2772–2773
  - societal stratification and, 2866
  - sociological concept of, 1198
  - status allocation and, 2780–2781, 2812
  - status attainment and, 759, 3042–3049
  - status crystallization and, 3051
  - status incongruity and, 3049–3054
  - structural lag and, 3063 and structure research, 2807–2820
  - tourism and, 3166
  - voluntary associations and, 3227
  - voting behavior and, 604, 3233, 3234
  - white-collar crime and, 3246, 3251–3252
- See also* Middle class; Social and political elites; Social stratification; Status attainment; Working class
- Social Class and Mental Illness* (Hollingshead and Redlich), 1813
- Social cognition. *See* Social perception
- Social cohesion, 1080
- Social comparison processes, **2649–2657**
- conformity and, 402
  - and justice evaluation, 2654, 2701–2702, 2705
  - and reference group comparison, 2702–2704, 2752–2753
  - self-concept and, 2507, 2650, 2651, 2653, 2856
  - and social comparison theory, 402, 2507, 2649, 2650–2651
- Social conformity. *See* Compliance and conformity
- Social conservatism, 357, 359
- Social Construction of Reality, The* (Berger and Luckmann), 226, 2957–2958
- Social constructionism. *See* Constructionist perspective



- Social contagion theory, 679
- Social contract, 528, 535, 1310, 1933
- Social control, **2657–2662**  
 conflict theory reactions to  
 deviance and, 669–670  
 control theory, 535, 536,  
 537, 1495  
 over corporations, 443–444  
 criminal sanctions as, 518–519,  
 521, 528, 2659  
 of crowds and riots, 556–557  
 internal vs. external, 2657  
 marriage licenses as, 948  
 medical profession as form of,  
 1815–1816  
 micro-level deviance theories and,  
 662, 666–670, 672  
 motivation and, 2060–2061  
 neighborhood stability as, 664  
 personal autonomy vs., 2061  
 police as agents of, 537, 556–557,  
 2107–2114  
 psychopharmacological drugs  
 used as, 1841  
 public opinion as, 2272  
 rape as, 2576, 2580  
 risk and, 667  
 sanctions and, 666–667, 671  
 and social disorganization theory,  
 1495, 2657  
 social order distinguished from,  
 2657, 2661  
 social welfare and, 2841  
 socialization as internalization  
 of, 2856  
 types of, 518  
 workplace, 3274–3275  
*See also* Law and legal system;  
 Law and society
- Social credit, 2734
- Social Darwinism, 63, 66, 423, 2332  
 city development theories, 308  
 and eugenics, 1272  
 evolutionary perspective, 878–879  
 and generative religious  
 movements, 2367  
 racial theory of, 2330, 2334
- Social Democratic Party (SPD;  
 Germany), 1753, 2163, 2623
- Social democratic systems, 377
- Social demography. *See* Demography
- Social determination theory,  
 2956–2957, 2957, 2958
- Social Diagnosis* (Richmond), 2843
- Social Differentiation* (Svalastoga),  
 2867
- Social dilemmas, 595–596, 3220  
 group size and, 1121–1122
- Social disorganization theory, 366,  
 532, 1495, 2657, 2759–2760
- Social distance, 177, 2693  
 model of, 2027–2028
- Social dynamics, **2662–2669**  
 of collective behavior, 352–354  
 of conformity, 400–404  
 dynamic sociology and, 1027,  
 1028  
 dynamic vs. static models,  
 2666–2667  
 French School of Sociology and,  
 1027, 1028  
 main contexts of, 2603  
 social change vs., 2663–2665  
 social organization and,  
 2735–2745  
 systems theory, 3102–3105  
 types of dynamic models,  
 2667–2668
- Social Dynamics: Models and Methods*  
 (Tuma and Hannan), 2668
- Social ecology, 792–793
- Social equilibrium, 1210
- Social evolution. *See* Evolution:  
 biological, social, cultural
- Social exchange theory, **2669–2776**  
 altruism and, 115  
 applications of, 2674–2675  
 behaviorism and, 208  
 economic determinism and,  
 722, 723  
 and emotions, 786  
 and exchange connections, 2672  
 and intergenerational resource  
 transfers, 1394  
 of interpersonal power,  
 1459–1463, 2674  
 and job satisfaction, 3272, 3275  
 and justice evaluation model,  
 2703–2704
- and mail survey design, 3091  
 major proponents, 2670–2674  
 as major social order school,  
 2337–2338  
 marital adjustment scale, 1728,  
 1729–1730  
 marital satisfaction and, 1729  
 marriage market and, 1777  
 money and, 1890  
 negotiation processes and, 1950  
 parity and, 330  
 and Polish sociology, 2119  
 rational choice theory and, 2335  
 and Scandinavian sociology, 2452  
 on sexual behavior, 2537, 2538  
 social capital and, 2637–  
 2639, 2640  
 social networks and, 2672–2674,  
 2730, 2731, 2733  
 as social psychology perspective,  
 2670, 2768  
 and social resources theory,  
 2790–2794
- Social Forces* (journal), 1528,  
 1606, 1870
- Social forecasting, **2676–2681**  
 future trends, 2680  
 futures studies, 1037–1042  
 judgmental and qualitative  
 models, 2677–2678  
 pragmatic statistical analysis of  
 time series, 2679  
 prediction and futures studies,  
 2224–2233  
 social demography and,  
 2678–2679  
 social indicators for, 2685
- Social gerontology. *See* Aging and  
 the life course; Filial  
 responsibility; Intergenerational  
 relations; Intergenerational  
 resource transfers; Long-term  
 care; Long-term care facilities;  
 Retirement; Widowhood
- Social identity theory, 1057, 1111,  
 1113, 2750, 2778  
 and group conflict resolution,  
 1112, 1113  
 language and, 2906–2908
- Social imitation. *See* Behaviorism;  
 Social learning theory; Social

- psychology; Social learning theory; Socialization
- Social impact assessment, 2619, 2687
- Social imperatives, life course and, 1619
- Social indicators, **2681–2689**  
 criterion indicators, 2683  
 data banks and, 576  
 definitions and examples of, 2301, 2681–2682  
 descriptive, 2684  
 historical developments in, 2682–2683  
 life satisfaction/happiness indicators, 2683–2684  
 policy analysis function of, 2687–2688  
 public enlightenment function of, 2685–2686  
 of quality of life, 2299–2307, 2685, 2686–2687  
 and secondary data analysis, 2473–2481  
 time use as, 3153–3154
- Social Indicators* (Bauer), 2682
- Social Indicators of Well-being: Americans' Perceptions of Life Quality* (Andrews and Witty), 2300
- Social Indicators Research* (journal), 2683–2684, 92682–92683
- Social Indicators Research: An International Journal of Quality of Life Measurement*, 2301
- Social inequality, **2689–2695**  
 beliefs about, 2704–2705, 2808  
 changes in forms of, 2820  
 and criminal sanctions, 2962  
 degree of dispersion, 2868–2869  
 dichotomous categories of, 2692–2693  
 distributional theories of, 2690–2691, 2694, 2696  
 globalization and, 1087, 1088, 2691, 2705–2708  
 as hierarchical relationship, 2693, 2694, 2781  
 human ecology and, 1215  
 and income distribution in the United States, 130, 140, 142, 689–690, 1281–1282, 2691, 3048
- intelligence and, 1384–1385  
 and social justice, 2696, 2701, 2705–2707  
 and Soviet socialism, 2850  
 and status allocation, 2780–2789, 2807–2820, 2813  
 and time-use indicators, 3154  
 in utopian designs, 3202–3203  
 as vertical classification, 2691–2693, 2694  
*See also* Discrimination; Equality of opportunity; Social and political elites; Social stratification; Societal stratification
- Social instability, causes of, 1230
- Social insurance, long-term care, 1659–1660, 1677
- Social integration, 79, 1223–1224  
 as suicide protection, 1595
- Social interaction, 44, 117, 196, 208, 772, 1027, 1253, 1347, 2085  
 conflict resolution, group, 1111–1117  
 conflict resolution, interpersonal, 1451–1456  
 conversation analysis, 431–439  
 differential association theory, 533  
 Durkheim on, 1030  
 and emotions, 2523  
 functionalism and structuralism, 1030  
 grammars of, 2897–2899  
 interpersonal power, 1456–1464  
 marital adjustment, 1729–1730  
 microlevel collective behavior, 350–352  
 nursing home adaptation, 1675–1676  
 personal dependency, 2062–2068  
 and Polish sociology, 2119  
 small groups, 2613–2614, 2617–2620  
 and social exchange theory, 2669–2775  
 symbolic interactionist modeling, 2298  
*See also* Personal relations
- Social Interaction Systems: Theory and Measurement* (Bales), 1014
- Social isolation, 176, 178
- Social justice, **2695–2711**  
 beliefs about inequality, 2704–2706  
 definition of, 2696  
 distributive justice theories, 2701–2704  
 distributive vs. procedural justice, 2699–2700  
 entitlement theory, 2699  
 equity theory, 2700–2701  
 and group conflict resolution, 1116  
 income inequality and, 2705–2706  
 mathematical models of distributive, 1790  
 as moral judgment orientation, 1902–1903  
 in organizational structure, 2003–2004, 2014  
 perception of, 2696–2697  
 philosophical roots of, 2697–2699  
 principles of, 2701  
 proportionality principle, 2697  
 Rawls principles of, 2698–2699, 2700, 3205  
 referential comparisons, 2702–2704  
 research areas, 2705–2707  
 social comparison process and, 2654, 2701–2702, 2705  
 social exchange theory and, 2671, 2703–2704  
 utopian designs and, 3201–3206  
 and work orientation, 3272  
*See also* Affirmative action; Comparable worth; Court systems and law; Discrimination; Gender
- Social Lamarckism, 879
- Social learning theory, 1716–1717  
 on aggression, 70–71, 75, 1716  
 on altruism, 115  
 on attitudes, 185  
 deviance and, 666–667, 668, 672  
 on femininity/masculinity, 999  
 moral development and, 1895

- on overdependency, 2067  
on personal dependency, 2064  
as personality theory, 1716–1717  
on sexual behavior, 2537, 2538  
as social psychology perspective, 2768–2769  
and socialization theory, 1715–1717, 2855–2856  
subcultural explanations of crime and, 534  
*Social Life of a Modern Community, The* (Warner and Lund), 364  
Social loafing effect, 2619  
Social mobility, **2711–2717**  
collective, 2715–2716  
convergence theories and, 425  
education and, 756, 758  
illegitimacy and, 1260  
industrialization and, 1311  
intergenerational, 1982–1984, 2711–2713  
minority suburbanization and, 3074–3075  
Scandinavian sociology studies, 2452  
and status incongruence, 3049–3054  
three traditions of, 2817  
*See also* Occupational and career mobility; Social stratification; Status attainment  
Social morphology, 515, 518, 1024, 1031  
Social movement organizations (SMOs), 2722–2724, 2725, 2763  
Social movements, **2717–2727**  
and age-strata changes, 3064–3065  
black civil rights. *See* Civil rights movement  
characteristics of successful, 2725  
collective behavior theories and, 349, 350–352  
countercultures, 459–463  
and criminal law construction, 1576  
effectiveness of, 2724–2725  
emergence of, 2718–2720  
information-based, 1347  
interest group size and, 604  
Internet as, 406–407  
labor unionization as, 1527–1534  
life history data, 1635  
Mexican opposition party as, 1861  
music and, 1927  
nationalist, 1940–1948  
organizational approach to, 2162, 2163–2164  
organizations, 2722–2724, 2725  
outcomes, 2724–2725  
participation in, 2720–2722  
participatory research, 2038–2042  
political power and, 2164, 2165  
religious movements as subcategory of, 2364–2375  
revolution as, 2412  
as significant sources of democracy, 606  
social problems and, 2763  
and status incongruence, 3050  
and structural lag, 3064–3065  
technological risk and, 2877  
theory, 939  
types of, 2717–2718  
*See also* Protest movements; Student movements  
Social networks, 1255, **2727–2735**  
in adulthood, 35  
as career advancement tool, 1993  
community integration/deviant behavior relationship, 664–665, 716  
and community of limited liability thesis, 367  
and consensus formation, 2379  
consequences of, 2731–2734  
and crowd participation, 560  
diagrams of, 2729, 2730  
divorce as weakening effect on, 706, 707  
and economic sociology, 736–737  
and exchange theory, 2672–2674, 2730, 2731, 2733  
functions of, 2727  
and innovation diffusion models, 678  
and interpersonal diffusion, 677  
mathematical modeling of, 1789–1790, 2729  
as meaningful work source, 665  
and measurement of social organization, 2738–2739  
models, 2029, 2298  
and peer deviant behavior, 666  
and Polish sociology, 2122  
precursors, 2728–2730  
and rational choice, 2731  
as social capital, 1993, 2637, 2639, 2640, 2714, 2732, 2733–2734, 3043–3044  
and social movement participation, 2721  
and social perception, 2753  
and social resources theory, 2790–2794  
as spousal division of labor influence, 696  
and status attainment, 3044  
“strength of weak ties” hypothesis, 2693, 2731–2732, 2791, 2792, 2827  
structural properties/structuring of, 1034–1035, 2729–2731, 2827  
urban interactions, 3193  
and widowhood support, 3259–3260  
and workers’ employment outcomes, 3265  
*Social Networks* (journal), 1790, 2734  
Social norms. *See* Social values and norms  
*Social Order of the Shum, The* (Suttle), 363, 365  
Social organization, **2735–2748**  
adolescence and, 3  
age stratification and, 79–80  
aging and, 79  
American society and, 146  
anomie and, 166  
attitudes as links within, 184  
Blau’s landmark theory of, 699  
childhood and adolescence peer groups, 2859  
classification of units, 2739–2741  
collective agency, 2744  
commitment and trust, 2744–2745  
complexity and, 2739

- definition of, 2735–2736  
 economic sociology and, 732  
 endurance and, 2737–2738  
 evolution and, 1271  
 form and, 2736–2737  
 globalization and, 1092  
 hybrid forms of, 2738–2739  
 industrial sociology and, 1308, 1309  
 information society and, 1347  
 logic of, 1092  
 major schools of, 2337  
 marriage as institution within, 1734  
 military sociology and, 1882  
 pluralism and plasticity, 2742–2743  
 pragmatic principles and, 221  
 rational choice theory and, 2338  
 of religion, 2376–2386  
 rural sociology and, 2425–2433  
 scale and, 2737–2738  
 segmentation and, 2737, 2743  
 sentiments and, 2524–2525  
 as social capital, 2639  
 social movements and, 2722–2724  
 specialization and, 2737, 2743  
 standardization and, 2743–2744  
 stratification and, 2737  
 and technological risk, 2878–2880  
*See also* Division of labor
- Social Organization of Sexuality: Sexual Practices in the United States* (Laumann et al.), 2550
- Social Origins of Dictatorship and Democracy* (Moore), 2918
- Social perception, 2748–2755**  
 accuracy of, 2751–2752  
 attribution theory, 192, 2751  
 cognitive consistency theories, 334–341  
 group perception, 2752–2753  
 leadership and, 1570  
 person perception, 2750–2752  
 racial prejudice and, 2243–2246  
 self-perception, 2750
- Social phenomenology. *See* Phenomenology
- Social philosophy, **2755–2758**  
 communitarianism, 354–362  
 postmodernism, 2205–2209, 2757, 2758  
 pragmatism, 2217–2224  
 religious orientation, 2382–2386
- Social physics (Comte concept), 818
- Social planning, 539
- Social Planning/Policy and Development Abstracts*, 1609
- Social policy analysis. *See* Public policy analysis
- Social power. *See* Social and political elites; Social control
- Social prestige. *See* Occupational prestige; Social and political elites: Social mobility; Status Attainment
- Social problems, 2759–2766**  
 and community reform, 365–366  
 deviance theories, 662–671  
 objectivist paradigm, 2759–2760, 2764  
 social movements defining, 2717, 2719, 2724–2725, 2763  
 subjectivist paradigm, 2760–2764
- Social Problems* (journal), 2759, 2762
- Social psychology, 2766–2780**  
 anomie conceptions, 166–167, 1024  
 attribution theory and, 192–198  
 censorship and, 278  
 coalition formation theories, 330  
 cognitive dissonance theory, 335, 337–340  
 counterculture implications, 460  
 and courtship study, 483  
 cross-cultural analysis method, 549–550  
 on fashion behavior, 679  
 femininity/masculinity measurement, 999–1000  
 first experiment (1897), 2615  
 on gender-mental health equation, 1838  
 on genocide, 1070  
 individual-level concepts, 2769–2770  
 of language, 2895
- major personality theories, 1712–1718  
 and Marx’s alienation concept, 1705  
 observation systems, 1973–1991  
 and pragmatism, 2217–2222  
 relationship with sociology, 2915, 2921  
 as religious studies perspective, 3277, 3278  
 role theory, 2415, 2417–2418, 2768, 2772  
 and self-concept, 2505–2509  
 and self-esteem, 2511–2518  
 and small groups, 2610–2621  
 and social exchange theory, 2670, 2768  
 and social mobility, 2714  
 and social structural effects, 2772–2775  
 and socialization, 2855  
 and stress, 3055–3057  
 symbolic interactionism perspective, 2767–2768, 3095–3101  
 theory and method, 2767–2769  
 voting behavior research, 3234–3238  
 and work orientation, 3271  
*See also* Affect control theory and impression formation; Attitudes; Behaviorism; Cognitive consistency theories; Collective behavior; Emotions; Field theory; Personality theories
- Social psychology of status allocation, **2780–2789**  
 history of theoretical field, 2781–2782  
*See also* Social stratification
- Social reality. *See* Phenomenology; Sociology of knowledge
- Social Reality of Death, The* (Charmaz), 583
- Social reform  
 community in context of, 365–366  
 liberal vs. conservative view of, 1602

INDEX

- Social relations. *See* Personal relations
- Social reorganization, 225
- Social representations, 197–198
- Social reproduction, 756, 758
- Social resources theory, **2790–2795**  
aging and, 79  
network vs. contact resources, 2792
- Social revolutions. *See* Revolutions
- Social roles  
aging and, 79  
altruism and, 117  
life cycle and, 1623  
sex differences and, 2530  
and stratification systems, 2807–2808  
*See also* Gender roles; Parental roles; Role theory
- Social Science Computer Review* (journal), 407, 409, 411, 414
- Social Science Computing Association, 407
- Social Science Data Archives (SSDA), 575–577, 579–580  
common goals and tasks, 579–580
- Social Science Research Council, 227, 576, 1634, 2274  
Center for Coordination of Research on Social Indicators, 2682  
Committee on Sociolinguistics, 2903  
Political Development project, 2154
- Social Science Research Council Fellowship, 324
- Social Science Research Council Sexuality Research Fellowship, 2550
- Social Sciences Citation Index* (SSCI), 1607, 1608, 1610–1611, 2923
- Social Sciences Index*, 1608
- Social Security Act of 1935, 2403, 2798
- Social Security Act of 1962, 1524
- Social Security Administration, 2398, 2795
- Social Security Amendments of 1983, 1524, 2407, 2799
- Social security systems, **2795–2807**  
age at receiving benefits, 2180, 2404, 3061  
comparative perspectives, 2796, 2799–2801  
consequences of, 2801–2802  
coverage and benefits distribution, 2796  
establishment of, 2403, 3065  
funding of, 2798–2799, 2803–2805  
historical background, 2797–2799  
and intergenerational resource transfers, 1395, 1396  
male labor-force participants and, 1524  
means-testing and, 2796–2797, 2798  
older people's benefits, 79, 925–926, 965, 967, 2403, 2404, 2406, 2795  
pay-as-you-go system, 2799  
policy problems and issues, 2802–2803  
recent trends, 2803–2805  
spousal and survivor benefits, 2406  
Supplemental Security Income, 1145, 1146, 1286  
*See also* Social welfare systems; Welfare state
- Social Security throughout the World*, 2795
- Social Statistics* (Blalock), 3035
- Social status. *See* Social class; Status attainment; Status incongruence
- Social stratification, **2807–2821**  
American system, 142–143, 2815–2816, 3043–3044  
ascriptive process of, 2809, 2817–2819, 3042, 3043, 3045  
attribution theory and, 198  
basic concepts of, 2808–2809  
convergence theories, 425  
and criminal sanctions, 515, 518  
and cross-cultural analysis, 550  
education and, 756, 758, 2927, 2929–2931, 3043–3044  
in eight ideal-typical systems, 2810  
elites and, 2622–2628  
and emotions, 777  
and ethnicity, 844–847, 2818–2819  
and expectation states theory, 880, 882  
forms of, 2809–2814  
and gender, 2818–2819  
and German sociology, 1076, 1078  
gradational groupings, 2816  
hierarchical continuum, 2781  
and industrial sociology, 1309, 1311, 1314  
legal equity and, 2961–2962  
levels of, 518, 2690–2694  
literary sociology and, 1648–1649  
as Marxist emphasis, 1753, 1754–1755, 2162, 2163, 2410–2411, 2813, 2814–2815  
and meritocracy, 2626–2627  
multidimensional approach to, 2818–2819  
nationalism and, 1942  
and occupational prestige, 2996–2000, 3265  
and personality, 2070–2071  
and Polish sociology, 2120  
and popular culture, 2170–2171  
professions and, 2259  
and race, 253, 558, 2818–2819  
revolution theory and, 2410–2411  
rigidity of system, 2809  
Scandinavian sociology studies, 2453  
in science, 2457, 2460  
as social organization, 2737  
social psychology of, 2780–2789  
socialist countermeasures, 2850  
and societal stratification, 2864–2874  
sources of, 2812–2814  
and status attainment, 3042–3049  
and status crystallization, 2809, 3051  
and status determination, 2781  
and status incongruence, 3049–3054  
and theories of the state, 2162, 2163

- types of assets, resources, and valued goods underlying, 2808
- urban underclass and, 3198–3200
- views of prestige, 1996–1997
- voting behavior relationship, 3233, 3234
- See also* Caste and inherited status; Social and political elites; Social class
- Social structure, **2822–2828**
- age effects on, 79–80, 3060–3061
- agricultural innovation adoption and, 89–91, 2431–2432
- anomie and, 165, 166
- and career advancement, 1984–1985
- city vs. urban, 3072–3073
- cohort perspectives, 342–343, 345–346
- culture debate, 563–565
- definitions of, 2069, 2771, 2822
- disaster symposium, 687
- division of labor, 124–127
- and economic sociology, 732
- empirical analytic metatheory on, 1853–1854
- and ethnomethodology, 858–859
- fertility transition and, 626–627, 628
- helping behavior and, 117
- identity theory and, 1253, 1254
- individualism and, 1301
- and inequality, 2705
- intelligence and, 1383
- kinship systems and, 1509–1512
- legal systems interaction, 1559–1561
- and legislation of morality, 1578–1579
- literary reflection of, 1645
- macro approach to, 710, 1708–1709
- Marxist views of, 775, 2162–2163
- mobility in, 2711–2716
- models of, 2029
- modernization theory and, 1885–1887
- music and, 1924, 1926
- and nationalist patterns, 1940–1942
- Parson's components of, 1560
- personality and, 2069–2074
- Polish sociology and, 2120
- and political party systems, 2155–2158
- and poverty theory, 2213
- regularities in, 2822–2823
- and religious orientation theories, 2384–2385, 2386
- revolutions and, 2410–2414
- role theory and, 2221
- shift from mechanical to organic solidarity, 697–698
- social capital and, 2638
- social exchange theory and, 2671
- and social inequality, 2689–2694
- social network models and, 2029
- social organization and, 2735–2745
- social psychology studies on, 2771, 2772–2773
- and social resources theory, 1793, 2790
- and socialization, 2862
- sociolinguistics and, 2901–2902
- and status attainment, 2782
- and structural lag, 3060–3067
- and symbolic interactionism, 2768
- and theories of state, 2162–2163
- and theory of emotions, 780–781
- values and norms in, 2836–2837
- See also* Social networks; Structuralism
- Social Structure and Disaster*, 687
- Social support
- emotional depression and perceived lack of, 653
- in multivariate analysis of stress, 1814
- social networks and, 2732–2733
- and social resources theory, 2790–2794
- and widowhood, 3259–3260
- Social surveys. *See* Data banks and depositories; Public opinion; Survey research
- Social System of the Modern Factory, The* (Warner and Low), 364
- Social System, The* (Parsons), 1813
- Social systems
- coalitions, 329–334
- culture and, 564–565, 566
- four functional requirements for survival, 2005
- human ecology and, 1211
- leisure and, 1583
- organizational, 394, 2005–2014
- psychological influences on, 2775–2778
- role theory analysis, 2421–2423, 2423
- Social Systems of American Ethnic Groups, The* (Warner and Srole), 364
- Social Text* (journal), 2207
- Social time. *See* Time use research
- Social traps, 2621
- Social Trends* (reports), 2684
- Social values and norms, **2828–2839**
- adolescent changes in, 1, 3
- age and, 29, 30, 31, 3064
- aggression and, 2774
- of agricultural communities, 164
- altruism and, 115
- in American society, 146
- anomie as breakdown of, 164–165, 533
- attitudes and, 185, 189, 2828–2829
- as behavior explanation, 2838–2839
- cohort differences in, 345–346
- collective behavior and, 351, 353, 2616
- communitarian, 354–356, 358–361
- community-based drug abuse prevention programs and, 716
- concept of a norm, 2829–2830
- concept of a value, 2828–2829
- as conformity influence, 402–403, 523
- countercultures, 460–462
- cultural differences and, 663–664
- death and dying, 582, 585
- deviance theories, 663–672, 2657–2659

- dimensions of importance, 2832–2836
- emergent norm theory, 350–351, 354
- of emotions, 2527–2528, 3213
- evaluative criteria, 2831–2832
- and family violence, 983
- and functionalism, 1031
- in health and illness behavior, 1133–1134
- institutions, 394
- interrelationships, 2836
- liberalism’s core view of, 355
- love and, 1698
- mass media and, 1766
- as mate selection factor, 1775–1776
- normative sanctions, 515, 518–519, 523
- object units of, 2830–2831
- origin of, 2836–2838
- and positive mental health, 2190
- rational choice theory and, 2340–2341
- responsive communitarian, 357–359
- role theory and, 2418
- sanctions stemming from, 515, 523, 537
- as social capital, 2638, 2639–2640
- as social controls, 2657, 2661
- social networks and, 2732
- socialization and, 2855–2864
- sources of change in, 2838
- structural lags and changes, 3064
- and student movement concerns, 3069–3070
- subjective, 189
- susceptibility to persuasion and, 2098
- utopian designs and, 3202–3205
- value-added theory of collective action, 352, 353
- values theory and research, 3212, 3219–3222
- vectors of subsets, 3220
- violations of, 165
- work group, 2617–2618
- Social welfare systems
- changes in, 2804, 3064
- entitlements and, 1221–1222, 2699
- first institutionalization of, 2840–2841
- marginal employment and, 1719–1720
- military service and, 1882
- neo-Malthusian arguments, 1220
- and poverty theories, 2213
- privatization of, 963
- as social control, 2660, 2661
- and social justice beliefs, 2705–2706
- Temporary Assistance for Needy Families program, 1145, 1146, 1288, 1395
- See also* Social security systems
- Social work, 2840–2846**
- casework, 2841, 2842, 2853
- degrees and training, 2845
- knowledge base of, 2843
- and Lewin’s legacy, 1015
- practice of effectiveness movement, 2843
- practitioner-researcher, 2844
- professional roots of, 2840–2841
- professionalization of, 2841–2842
- research design and sampling, 2843, 2845
- sociology vs., 2840, 2844–2845
- Social Work abstracts*, 1611
- Social Work Research and Abstracts*, 2845
- Social-area analysis. *See* Cities
- Socialism and communism, 2846–2852**
- authoritarian communitarianism and, 356
- capitalism vs., 237, 238
- Chinese vs. Russian system, 2849
- church-state relations and, 2357
- classical elite theory vs., 2623
- convergence theories and, 422, 428
- demise in Eastern Europe, 428, 1199, 1757, 2851
- democracy and, 2848
- and dialectical materialism, 1782
- dictatorships, 3002
- future of, 2852
- governmental corruption, 2136–2137
- and historical sociology, 1199
- ideal of, 2847–2849
- inadequate worker motivation and, 2813
- and Italian politics, 2128–2129
- labor movement and, 1529, 1530
- law and legal system, 1548, 1554
- legacies of, 2851–2852
- liberal/conservative tags in, 1598
- Marxist sociology’s identification with, 1752
- Marx’s two-stage evolution into communism, 2847–2849
- origin of term “socialism,” 2846
- Polish sociology under, 2116
- politically correct “party line” in, 2139
- public tolerance of, 316, 317
- reality of socialism, 2849–2851
- revisions of Marxist theory, 2848–2849
- in South East Asia, 2975, 2978
- Soviet and Soviet bloc collapse of, 428, 1119, 1757, 2851
- and Soviet Marxism-Leninism, 1751, 1782
- Soviet sociology under, 2116
- Soviet-type, 2849–2850
- stratification parameters, 2810, 2812, 2813–2815
- utopian, 3202–3203
- See also* Cold War; Marx, Karl; Marxism -Lenism; Marxist sociology
- Socialist Party (France), 2129
- Socialist Party (Germany), 1533
- Socialist Party (Italy), 2128, 2129
- Socialization, 2855–2864**
- in adolescence, 10–11, 34, 2852–2860, 2861
- age appropriateness and, 1624
- altruism and, 115
- attitude formation and, 185
- in childhood, 1057
- communitarian vs. liberal views of, 358

- contemporary issues and themes, 2862–2863
- content and contexts of, 2856–2860
- cross-cultural analysis of childhood, 550
- and death and dying, 583
- deviance theories and, 662
- diverse meanings of, 2855
- and family violence, 983–984
- and gender identity and roles, 625, 997, 998, 1275, 2529, 2886
- genocide and, 1071
- as identity formation, 2856
- illness behavior and, 1130
- as internal social control, 2657, 2661
- Judaic aims in, 1510
- throughout life course, 2860–2862
- of music performers and audiences, 1925
- personal values and norms stemming from, 2837–2838
- “prisonization” as, 2052
- racial resentment and, 2245
- resocialization and, 2860
- and self-concept, 2856, 2860, 2862
- and self-esteem development, 2512–2513
- sex differences and, 2529, 2530
- and sexual behavior, 2537–2538
- social class and, 2773
- and social learning theory, 1716–1717, 2856
- social psychology research on, 2771–2773
- and social structure, 2822
- and structural lag, 3062–3063
- and youth subcultures, 514
- Societal stratification, **2864–2874**
- classical theory, 2865–2866
- distribution profiles, 2869–2874
- educational mobility and, 2927
- empirical tradition, 2865, 2866–2867
- hierarchically ordered power relationships, 2864–2865
- and intergenerational resource transfers, 1395–1397
- levels of structural dimensions, 2868–2869
- measurement and comparison data, 2870–2874
- and status crystallization, 2869
- utopian variants, 3202–3206
- See also* Social stratification
- Societe de Sociologie de Paris, 1422
- Society and technological risks, **2874–2880**
- disaster planning and research, 683–684
- systems fragmentation and, 3104
- terrorism and, 3138, 3139
- war and, 3242–3243
- “Society and the Imperative of Death” (Fulton), 583
- Society for Applied Sociology, 155, 326
- Society for the Advancement of Field Theory, 1013
- Society for the Scientific Study of Religion, 2373
- Society for the Study of Social Problems, 155, 326, 2759, 2763
- Society for the Study of Symbolic Interaction, 3096
- Society of Captives, The* (Sykes), 2052
- Society of Sociology of Paris, 1025, 1026
- Sociobiology, human, **2880–2888**
- aggression mechanisms, 69–70, 72
- altruism theory, 115, 2882–2884
- conceptions of race and, 2329–2332, 2334
- crime and criminal behavior theories, 502, 506, 528, 529
- and culture, 563
- depression theories and treatment, 652, 654–655
- elements of theory, 2881–2886
- evolution theory and, 876–877, 2880–2881, 2892
- and genocide, 1068
- and incest taboos, 1274
- and law of anisogamy, 2884–2886
- maximization principle, 2881
- origination of term, 2880–2881
- and self-esteem determination, 2513
- as sexual behavior approach, 2537, 2566
- social Darwinism and, 66, 2330
- suicide predictors, 3079
- See also* Evolution: biological, social, cultural
- Sociocultural anthropology, **2888–2894**
- anthropological method and theories vs., 2889
- current issues, 2893–2894
- ethnography and, 852–856, 2888
- history of, 2888–2892
- on Islamic society, 2944
- linguistic analysis, 2891–2892
- Malinowski’s work, 2118–2119
- methods and organization, 2892–2893
- and Polish sociology, 2117–2118, 2119–2120
- rape explanations, 2579–2580
- theory, 2889–2892
- See also* Anthropology; Ethnography
- Sociocultural ecology
- and criminology, 531
- and urban life, 308–309, 531–532
- See also* Human ecology and environmental analysis
- Socioeconomic Index (SEI), 1997, 2000, 3265
- Socioeconomic status. *See* Social class; Status attainment
- Sociogenic theory of criminal behavior, 506
- Sociograms, 2728–2729
- Sociolinguistics, **2894–2912**
- collaborative work across disciplines, 2903–2904
- conversation analysis and, 431–440, 2895, 2902–2903, 2904–2905
- mutual embeddedness proposition, 2896–2897, 2906–2907
- and personal and social identities, 2906–2908



- sociocultural anthropology and, 2891–2892
- Sociologia* (Danish journal), 2451
- Sociological Abstracts*, 409, 1015, 1606, 1607, 1608–1609, 1610
- citations for Middle Eastern studies, 1869
- quality of life entry, 2300
- Sociological Imagination, The* (Mills), 1027
- Sociological Institute (Russia), 2979
- Sociological Methodology* (ASA yearbook), 409, 414, 1611, 3034
- Sociological Methods and Research* (journal), 3034
- Sociological Practice: A Journal of Clinical and Applied Sociology*, 326, 328
- Sociological Practice Association, 325, 326, 328
- Sociological Practice Review*, 155, 156
- Sociological Research Online* (quarterly), 413
- Sociological Society (Great Britain), 225
- “Sociological Technique in Clinical Criminology, A” (Alinsky), 325
- Sociological Theory of Law, A* (Luhmann), 1558
- Sociologie allemande contemporaine* (Aron), 1025
- Sociologie des Sports* (Risse), 2986
- Sociologie du travail* (Friedman), 1026
- Sociologisk forskning* (Swedish journal), 2451
- Sociologisk Tidsskrift* (Norwegian journal), 2451
- Sociologists for Women in Society, 153–155
- Sociology: Theory, Methods, Marketing* (Ukrainian journal), 2983
- Sociology
- and affirmative action, 47, 48, 49
  - of aging, 52–85
  - American Sociological Association, 147–156
  - applied, 155–156, 168–171, 1237, 2845
  - British, 223–228
  - case studies, 244–248
  - clinical, 323–328
  - communitarianism and, 361
  - comparative historical analysis, 383
  - of culture, 172
  - definition and characteristics of
    - community problems, 362, 363
  - disaster research and, 687
  - Durkheim’s coining of term, 327
  - of education, 2926–2935
  - of emotions, 773
  - environmental, 800–812
  - French, 1024–1029
  - German, 1073–1084
  - globalization of, 1090–1091
  - of helping, 117–118
  - historical, 1195–1200
  - and human rights, 1237–1238
  - humanistic orientation, 1246–1251
  - hybridization, 2923–2925
  - inception as response to social changes, 2644
  - Indian, 1290–1295
  - indigenization of, 1292
  - industrial, 1308–1316
  - international associations, 1422–1426
  - Internet resources, 413–414
  - Italian, 1464–1475
  - of language, 2894–2912
  - library resources and services for, 1604–1613
  - life histories and narratives, 1633–1638
  - literature reflecting, 1646
  - macro vs. micro, 1703–1705
  - Marxist, 1752–1759
  - and materialist theory, 1783–1784
  - mathematical, 1786–1792
  - medical, 1813–1818
  - metatheory in, 1852–1854
  - Middle Eastern studies and, 1868–1873
  - military, 1875–1882
  - models in, 2027–2028, 2029
  - of money, 1883–1893
  - as moral enterprise, 1246
  - music and, 1924–1927
  - Native American, 136
  - pantheon of founders, 1704
  - paradigm use in, 2027
  - phenomenology applied to, 2099–2106
  - of police, 2114–2115
  - Polish and Eastern European, 2116–2123
  - political, 2162–2168
  - popular culture studies, 2169–2175
  - and positivism, 2192–2195
  - postmodernist, 2205–2209
  - and pragmatism, 2220–2222
  - of prisons, 2051–2057
  - professional associations, 148–156
  - relevance of, 1237–1238
  - of religion, 2964–2974
  - research funding in, 2397–2401
  - rural, 2425–2436
  - social capital and, 2641
  - and social exchange theory, 2670–2675
  - and social philosophy, 2755–2758
  - and social problems theorists, 2759–2764
  - social work vs., 2840, 2844–2845
  - socialization orientations in, 2855–2856
  - sociocultural anthropology shared interests, 2893–2894
  - sociology of law relationship, 1553–1555
  - Soviet and post-Soviet, 2979–2985
  - specialized domains, 2913
  - of sport, 2986–2990
  - statistical methods since 1969, 3035–3039
  - structuralist, 1034, 1559–1560
  - suicide studies, 3079–3081
  - survey research use, 3087–3094
  - teaching of, 150
  - urban, 3191–3197, 3191–3198
  - UseNet Newsgroups, 414
  - values theory, 3213–3225
  - of work, 3269
- Sociology* (Japanese journal), 1480
- Sociology among the social sciences, **2913–2926**

- new school of communitarians  
and, 356–359
- Sociology and Social Research* (journal),  
326, 1864
- Sociology of agriculture. *See* Rural  
sociology
- Sociology of art. *See* Art and society
- “Sociology of Art, The”  
(Barnett), 173
- Sociology of Art and Literature, The*  
(Albrecht, Barnett, and Griff  
eds.), 172–173
- Sociology of Childhood (ISA  
research committee), 550
- Sociology of culture. *See* Culture;  
Mass culture
- Sociology of Culture* (Kloskowska),  
2121
- Sociology of Culture, The: Emerging  
Theoretical Perspectives* (Crane  
ed.), 568, 569
- Sociology of democracy. *See*  
Democracy
- Sociology of development, 300, 754
- Sociology of Diasters: Contributions of  
Sociology to Disaster Research*, 687
- Sociology of education, **2926–2937**  
and development, 754  
empirical studies, 2929–2934  
home vs. school influences, 2933  
reformist projects, 2934–2935
- Sociology of emotions. *See* Emotions
- Sociology of environmental issues.  
*See* Environmental sociology
- Sociology of Islam, **2937–2953**  
Gellner’s theory of Muslim  
society, 2943–2944, 2946,  
2947  
and gender issues, 2948–2950  
and Islamic state, 2948  
Muslim minorities, 2950–2951  
and religious experience,  
symbols, and theology, 3280,  
3281, 3282, 3284, 3285  
and social theory, 2941–2948  
Sunni and Shi’a sects, 3286
- Sociology of knowledge, **2953–2960**  
ideation change, 1009–1010  
and “new sociology of  
knowledge,” 2958–2959
- Sociology of Knowledge, The* (Stark),  
2956–2957
- Sociology of law, **2960–2964**  
comparative legal systems,  
1545–1551  
definition of, 1576  
and international law, 1426–30  
legislation of morality, 1560,  
1575–1581  
relationship to general sociology,  
1552–1555  
relationship to legal theory,  
1555–1559  
role of general theory in, 2963  
social-structural model, 1559–  
1560  
*See also* Law and society
- Sociology of literature. *See* Literature  
and society
- Sociology of religion, **2964–2974**  
on defining feature of religion,  
2382–2383  
family and, 936, 937  
French School and, 1024  
new paradigm in, 2367–2368,  
2374–2375  
religious movements and,  
2364–2375  
religious organizations in context  
of, 2377–2379  
sect formation and, 2378–2379  
Weber’s economics-based  
analysis, 2942–2943  
world traditions, 3277–3289
- Sociology of science. *See* Science
- Sociology of Sport Journal*, 2988
- Sociology of Teaching, The* (Waller),  
2927, 2986
- Sociometry, 1014, 1027, 2613, 2729
- Sociomoral Reflection Measure  
(SRM), 1899–1900
- Sociomoral Reflection Measure-Short  
Form (SRM-SF), 1903
- Socioterritorial belonging. *See*  
Territorial belonging
- Socrates, 3079, 3202
- Sodomy laws, 315
- SOFI. *See* Institute for Social  
Research (Sweden)
- Software. *See* Computer applications  
in sociology; *specific packages*
- Soga, 1549
- Sohrabi, Nader, 1871
- “Sojourner” concept, 178
- Sokal, Alan, 2207–2208
- Solidarity  
collective, 2631, 2632–2633, 2635  
and social capital, 2640
- Solidarity movement (Poland), 1532,  
2121, 2268, 2365–2366
- Solipsism, 2217
- Solis, Leopoldo, 1858
- Solomon, king of Israel, 2998
- Solomos, John, 226
- Somalia, 1865, 1866, 1867, 1934  
slavery and forced labor in,  
2603, 2606
- Somatotype personality, 1717–1718
- Sombart, Werner, 384, 733, 1074,  
1075, 2711
- “Some Principles of Stratification”  
(Davis and Moore), 1030
- Some Remarks on the Social System*  
(Lockwood), 225–226
- Song lyrics analysis, 1927
- Sons of Liberty, 2125
- Sorensen, Aage, 1349, 2452
- Sorokin, Pitirim, 115, 681, 821,  
1027, 1081, 1424, 1932, 2921  
on cyclical dynamics of societal  
change, 1705, 2644  
on education and mobility, 2927  
on family structure, 1505  
functionalist hypothesis, 2866  
and rural sociology, 2426,  
2427, 2428  
and Russian sociology, 2116  
on social dynamics, 2662  
and social inequality concept,  
2690  
on societal stratification,  
2865–2866, 2867  
and Soviet sociology, 2979, 3054  
and status attainment, 2781  
stratification profile, 2869–2870  
stratification profile, 2869–2870  
time use research, 3155–3156
- Soros, George, 588, 2983
- Sosiologia* (Finnish journal), 2451
- Sotsiologicheskie Issledovania* (Soviet  
journal), 2981

INDEX

- 
- Soule, Sarah A., 2961
- Souls of Black Folk, The* (Du Bois), 55
- South, U.S.
- and interracial sexual relations, 1408
  - low alcohol consumption rate, 94
  - segregation policy, 54, 62, 2491–2492
  - subculture of violence in, 664
- South Africa
- apartheid, 62, 250, 1940, 2047, 2146
  - apartheid abolishment, 2725
  - clinical sociology, 328
  - fertility rate decline, 220
  - involuntary servitude, 2602, 2608
  - political crime, 2146
  - post-apartheid peacemaking, 2047–2048
  - protest movement, 2270
  - racial classification, 2332
  - racial conflict, 321
  - social change, 941
  - status attainment, 3044
  - women in labor force percentage, 3262
- South America. *See* Latin American; Latin American studies; *specific countries*
- South Commission (1990), 1319–1321
- South Korea
- dependency theory, 642
  - fertility decline, 627, 2178
  - labor movement and unions, 1531, 1532
  - political corruption, 2131
  - rapid economic expansion/population growth, 2179
  - widowhood, 3255–3256
- Southeast Asia in the Age of Commerce* (Reid), 2978
- Southeast Asia studies, **2974–2979**
- fertility transition, 622
  - history, 2974–2976
  - women in labor force percentage, 3262
- Southeast Asian Refugee Studies Project, 182
- Southern Africa, countries of, 60
- Southern Baptist Convention, 2370, 2376
- Southern Christian Leadership Conference, 2492, 2494, 2495
- Southern Organizing Committee, 791
- Southwest Network for Environmental and Economic Justice, 791
- Southwood, Kenneth, 2703
- Soviet and post-Soviet sociology, 2116–2117, **2979–2985**
- bureaucratic development, 2163
  - Marxist sociology and, 1756, 1757
  - suppression and revival, 2979–2980
  - time use research, 3155, 3156
  - See also* Russia; Soviet Union
- Soviet Sociological Association (SSA), 2980, 2981, 2982
- Soviet Union
- in Cold War-era triad, 332
  - Collins's predicted collapse of, 1708
  - communist dictatorship, 3002
  - cross-border crime upswing, 1936
  - equality for women, 990
  - forced labor, 2608
  - Marxism-Leninism, 1751, 2982
  - Marxist view of democracy, 601–602
  - political corruption, 2136–2137
  - secularization, 2485
  - social movement emergence, 2718
  - socialism, 28, 2846, 2849, 2849–2851
  - See also* Russia
- Sowell, Thomas, 1233
- Sozanski, Tadeusz, 2122
- Soziologie heute* (König), 1077
- Space program, 2682
- Spain
- divorce laws reforms, 703
  - divorce rate, 706
  - Latin American colonization by, 1536–1637, 1934
  - legal system, 471, 474
  - long-term care and care facilities, 1652, 1653, 1661
  - marijuana decriminalization, 712
  - multilingualism, 2909
  - nationalist movement, 3001
  - and New World empire, 2999
  - and New World slavery, 2600
  - political and governmental corruption, 2129
  - retirement patterns, 2407
  - Southeast Asian influences, 2974
  - tourism in, 3169
  - unemployment, 3263
  - women in labor force percentage, 3262
- Spanier, Graham, 1508, 1726, 1727, 1728, 2393
- Spann, Othmar, 1075
- Spates, James L., 3214
- Spatial ability, sex-differences study, 2531
- Spearman, Charles, 905–906, 908, 909–910, 1364, 1365, 2348
- Spearman rank correlation, 1795, 1957, 1969
- Spearman-Brown Prophecy, 2348–2349
- Specialization, 697, 698, 699
- Speech. *See* Conversation analysis; Language; Sociolinguistics
- Speech-recognition programs, 1981
- Spencer, Herbert, 297, 581, 582, 819, 1028, 1271, 1423, 1465, 1466, 2889
- evolutionary “fitness” theory, 2881
  - evolutionary racial theory, 2330
  - and evolutionary theory, 879
  - and functionalism and structuralism, 1029, 1030, 1031
  - on historical progress, 2644–2645
  - as Japanese sociology early influence, 1477
  - macro-level phenomena concerns, 1704
  - and modernization theory, 1885
  - as Polish sociology influence, 2117
  - and positivism, 2192, 2193
  - and secularization, 2483
  - social Darwinism of, 66, 2330
-

- on specialization and  
routinization, 697, 698
- Spengler, Oswald, 877, 2644
- Spenner, Kenneth L., 270
- Spickard, James, 940
- Spilerman, Seymour, 557, 1614–  
1615, 1991
- Spirit of the Laws, The*  
(Montesquieu), 1545
- Spitzer, Steven, 1498
- Split panels, 1687
- Split-half reliability, 2347–2348
- Spock, Benjamin, 2036–2037
- Spohn, Willfried, 569
- Spoils system, 2127
- Spontaneous abortion, 2233–  
2234, 2238
- Sport, **2985–2991**, 3077
- Spousal division of labor, 695–696
- Spragens, Thomas, Jr., 356
- Spreadsheet software, 418–419
- SPSS (Statistical Package for the  
Social Sciences), 1607, 1923, 2035
- Spurzheim, J. C., 1717
- Sri Lanka, 2362  
high suicide rate, 3082  
political corruption, 2132  
religious movement, 2366, 3288
- Srinivas, M. N., 1291, 1292, 1293
- SRM. *See* Sociomoral Reflection  
Measure
- SRM-SF. *See* Sociomoral Reflections  
Measurement Short Form
- Srole, Leo, 364
- SRS (simple random sampling),  
2446–2447
- SSA. *See* Soviet Sociological  
Association
- SSCI. *See* Social Sciences Citation Index
- SSDA. *See* Social Science Data  
Archives
- S-shaped curve (innovation  
adoption/diffusion), 87, 677
- SSI. *See* Supplementary Security  
Income
- SSRIs (selective serotonin  
inhibitors), 654
- SSRS (Systematic-simple random  
sampling), 2446–2447
- SSSP. *See* Society for the Study of  
Social Problems
- Stability, as reliability component,  
2343, 2346, 2350–2354
- Stable criminal subculture, 1494
- Stable population model, 618
- Stack, Stephen, 3079, 3081
- Staffen, Lisa R., 2876–2877
- Stakeholders, 602
- STAKES. *See* National Research and  
Development Centre for Welfare  
and Health (Finland)
- Stalin, Joseph, 298, 2136, 2608,  
2812, 2979, 3137
- Standard American model of kinship  
mapping, 1514–1515
- Standard Cross Cultural Sample*  
(Murdock and White), 548
- Standard deviation (statistical)  
definition of, 659  
variance compared with, 660
- Standard error, 457, 2449
- Standard error of mean, 3029
- Standard Metropolitan Statistical  
Areas (SMSAs), 2478
- Standard of living  
divorce effects, 706, 707–708  
industrialization and, 1217, 1219
- Standardization, **2991–2996**
- Standardized partial regression  
coefficient, 453–454
- Standards for Educational and  
Psychological Testing*, 3207,  
3210, 3211
- Stanford Political Dictionary*, 1979
- Stanford Research Institute, 1041
- Stanford-Binet Intelligence Scale,  
1362
- Staniszki, Jadwiga, 2121
- Stanley, J. C., 2325, 2327
- Staples, William, 1636
- Star plots, 3021
- Stare decisis*, 476
- Stark, Oded, 633
- Stark, Rodney, 665, 2367,  
2374–2375, 2381, 2484–2485,  
2967
- Stark, Werner, 1958, 2956–2957
- Starovoitova, Galina, 2981
- Starr, Bernard D., 2555–2556
- Starr, Paul, 375
- Starr-Weiner Report, 2555–2556
- State, The, **2996–3003**  
bureaucratization in, 1198  
child support and, 1261  
church-state relations models,  
356–358  
constitutions and, 3000–3001  
definition of term, 2996  
ecology and, 1213  
emergence in Europe, 1933,  
2356, 2362  
expansion of, 1266  
formation and breakdown studies  
of, 1707–1708  
formation history, 2362–2363,  
2998–3002  
globalization and, 1092,  
1093, 2362  
historical sociological studies  
of, 1198  
individualism and, 1303  
Islamic society relationship, 2948  
liberal model, 3002  
*See also* Welfare state  
macrosociological approach to,  
1707–1708  
Marxist theory of, 1755,  
2162, 2163  
materialist theory on, 1785  
national, 1198  
national boundaries and,  
1931–1938, 1939  
nationalist movements and, 1941,  
3001–3002  
nation-state definition, 2997  
negation of power by, 1951–1955  
and organizational  
functioning, 2006  
political power of, 2997–2998  
reaction to nationalist movements  
by, 1945–1946, 3002  
regulatory function, 1099, 1228  
role in globalization, 1092, 1093  
role in higher education,  
1184–1185  
structuralist theory of, 2162–2163  
and territorial belonging, 3129

- war and twentieth-century nation-state formations, 2362–2363  
*See also* Democracy
- State, Culture, and Society* (journal), 1869
- State Department, U.S., Human Rights Practices report, 2607
- State Self-Esteem Scale, 2512
- State socialism, 728
- State system. *See* National border relations; State, The
- Static-group comparison, 2316
- Stationary population, 616
- Stationers' Company, 267
- Statistical Analysis of the Social Organism (Comte concept), 1029
- Statistical Breviary* (Playfair), 3005
- Statistical graphics, **3003–3023**  
 in data analysis, 3003, 3011–3019, 3039  
 history of, 3005–3007  
 multivariate data, 3018–3019  
 nonparametric, 1795  
 perception research, 3009–3011  
 standards, 3007–3009  
*See also* Content analysis
- Statistical inference, **3023–3034**  
 confidence intervals, 3026–3027  
 logic of, 3027–3030  
 models, 2028  
 nonrandom samples and, 3033–3034  
 power and Type I and Type II errors, 3030–3032, 3033  
 probability level, 3025–3026  
 and replication attempts, 2396  
 and statistical and substantive significance, 3032–3033  
 traditional tests, 3024–3025  
 variance and standard deviation, 659, 660
- Statistical methods, **3034–3039**  
 analysis of variance and covariance, 158–164  
 categorical and limited dependent variables, 3037–3038  
 comparative-historical sociological analysis, 385–386, 389, 390  
 computer programs, 409  
 computer-intensive methods, 3039  
 conclusion validity, 2326–2327  
 contingency table analysis, 3036–3037  
 correlation and regression analysis, 447–457, 2251, 3035–3036  
 covariance structure models, 3037  
 crime rate calculation, 491–492, 503–504  
 and cross-cultural analysis, 547  
 demographic, 608–620  
 descriptive, 657–661  
 distributions, 2869–2870  
 event history analysis, 869–874, 3037  
 factor analysis, 905–921, 1788, 3036  
 graphics, 3003–3022  
 inference. *See* Statistical inference  
 latent structure analysis, 3038  
 linear models, 592–593  
 measures of association, 1804–1812  
 models, 1787, 2028, 3015–3017  
 multilevel and panel models, 3038–3039  
 multiple tests of significance, 3030  
 nonparametric, 1795, 1956–1971  
 quasi-experimental research design errors, 2326–2327  
 sampling theory, 2444–2449  
 social indicators, 2684  
 standardization, 2991–2996  
 statistical and substantive significance, 3032–3034  
 stochastic processes, 3036  
 tabular analysis, 3108–26  
 time series analysis, 2679, 3142–3153  
 Type I and Type II errors, 3030–3032  
 typologies, 3186–3187  
*See also* Content analysis; Crime rates; Measurement; Measures of association
- Statistical Methods* (Snedecor), 3035  
*Statistical Methods for Research Workers* (Fisher), 3006, 3035
- Statistical Package for the Social Sciences (SPSS), 1607, 1923, 2035
- Statistisk Sentrabryrå, 2451
- Status attainment, **3042–3049**  
 aspiration level formation, 2781, 2782–2783  
 and capitalism, 238  
 conformity and, 404  
 crystallization levels, 2869  
 definers and models, 2782  
 and delinquent subcultures, 511  
 and education, 758, 2713, 2782, 2783–2785, 2928, 2929–2931  
 and emotions, 777–778  
 establishment of field, 3042  
 and expectation states theory, 880, 882  
 gradational groupings, 2816  
 income and, 3047–3048  
 incongruence concept, 3049–3054  
 inherited, 250, 251, 1239–1240, 2810, 2811, 2869  
 leisure displays and, 1583  
 marriage and, 1734–1735  
 in mate selection theory, 1775–1776  
 measurement devices, 2866  
 models of, 2782–2783, 2817  
 nonformist/innovative behavior allowed by, 403–404  
 occupational prestige and, 2000–2001, 2259–2260, 3045–3047  
 professions and, 2259–2260  
 social network position and, 2733  
 social psychology studies of, 170–171, 2780–2789  
 and social reproduction, 3042, 3043, 3045  
 social surveys, 578  
 tourism and, 3166, 3168  
 upward structural mobility and, 2868  
 and urban life, 307  
 Weber's studies, 1704

- See also* Social and political elites;  
Social class; Social stratification
- Status crystallization, 3051
- Status incongruence, **3049–3055**  
configurations of, 3053  
at elite level, 3053–3054  
and relative deprivation, 349,  
1940, 2701, 3050  
widowhood and, 3255
- Status inconsistency. *See* Status incongruence
- Status quo  
conservative vs. liberal ethos,  
1598, 1600  
endowment effect and, 594
- Status System of a Modern Community, The* (Warner and Lund), 364
- Status-value formulation, 2702
- Stavisky affair (1934), 2129
- STDs. *See* Sexually transmitted diseases
- Stearns, Linda B., 2501
- Stebbins, Robert, 1587
- Steele, B. F., 288
- Steele, C. M., 73
- Steen, Sara, 671
- Stefano, Antonino, 2968
- Steffen, Gustaf Fredrik, 2449
- Steffens, Lincoln, 2124, 2126
- Stein, J. A., 2513
- Steinberg, Stephen, 845, 2332–2333
- Steiner, Ivan D., 2617
- Steinmetz, Devora, 1511
- Steinmetzarchief (Amsterdam Arts and Sciences Real Academy), 575, 576
- Steinmo, Sven, 375
- Stem family, 1503–1504
- Stem-and-leaf plots, 659
- Stepfamilies. *See* Blended families
- Stephens, Evelyne Huber, 389
- Stephens, John, 389, 390
- Stepney, Bishop of, 225, 1706
- Stepwise regression analysis, 456
- Stereotypes  
African-American, 64, 2243  
on aging and sexual activity, 2556  
as attitude subtype, 184, 189  
as attribution shaping, 197–198, 2244  
conformity and, 403  
functions of, 185  
of fundamentalism, 2368,  
2369–2370, 2371  
gender. *See* Sex stereotypes  
and intermarriage, 1411  
as learned, 185  
of marijuana use, 713  
mass media and, 1767–1768,  
1773  
and mental illness diagnosis, 1838  
and reference group perceptions,  
2752, 2753  
of whiteness, 56  
*See also* Discrimination; Prejudice
- Sterilization  
and eugenics, 1272  
and family planning, 954, 2178
- Sterling, Robert, 197
- Sterling, Theodore D., 2396
- Sternberg, R. J., 1850
- Sternberg, Robert, 1368
- Stevens, S. S., 1793, 1800
- Steward, Julian, 2891
- Steward, Lyman and Milton, 2368
- Stewart, John A., 2459
- Stigler, George, 1101
- Stigmatization  
concept of, 1815  
as criminal sanction, 520  
of homosexuality, 2570  
of urban underclass, 3198
- Stimulus diffusion, 676
- STIRPAT (environmental impact equation), 808
- Stochastic Models for Social Processes* (Bartholomew), 2668, 3036
- Stochastic processes, 2249,  
2465, 3036
- Stochastic Processes* (Doob), 3035
- Stockholm University, 2452
- Stogdill, R. M., 1565, 1567–1568
- Stoics, 2519
- Stokes, Donald E., 3234, 3235
- Stone, Christopher, 443
- Stone, Gregory, 2221
- Stone, Lawrence, 1504, 1505
- Stone, Philip J., 1978, 1979
- Stonequist, Everett, 2634
- Stonich, Susan, 1221
- Storytelling. *See* Life histories and narratives
- Stouffer, Samuel A., 314, 316–317,  
1876, 1881, 2193, 3038, 3093
- Strain theory, 166  
of collective behavior, 353  
of deviance, 664, 1494–1495,  
2775
- Strain toward symmetry model,  
335, 336
- “Strange Disappearance of Sick America, The” (Putnam), 368
- “Stranger” concept (Simmel and Schutz), 2635
- Strangers from a Different Shore: A History of Asian Americans* (Tataki), 180
- Strassoldo, Raimondo, 1468
- Strategic narrative concept, 388
- Stratham, Anne, 2417, 2423
- Stratification. *See* Social stratification
- Stratified random sampling, 2447
- Straus, Murray, 503, 505
- Straus, Robert, 1815
- Strauss, Anselm, 582, 1813, 2221
- Strauss, B., 2290
- Straw polls. *See* Election polls
- Street Corner Society* (Whyte), 244,  
363, 364, 365, 2611
- Street gangs. *See* Gangs
- Streib, Gordon, 1388
- “Strength of weak ties” hypothesis,  
2693, 2731–2732, 2791,  
2792, 2827
- Strength-of-position hypothesis, 2791
- Stress, **3055–3059**  
aggression and, 73  
as caregiver burden, 1658  
collective situations, 683  
depression and, 649, 651,  
653, 656  
of divorce, 705, 707  
homelessness and, 1205  
integrative and time-lagged  
models, 3056–3057

- mediating factors, 3056  
 medical sociology studies, 1814  
 personal dependency and, 2066  
 social comparison process  
   and, 2654  
 social networks counteracting,  
   2732, 2733
- Strikes. *See* Labor movements and  
 unions; Industrial sociology
- Stripp, H., 3081
- Strodtbeck, Fred L., 3212
- Stroessner, Alfredo, 2134
- Stroke, 139, 1641
- Structural analysis, 1027  
   in macrosociology, 1703–1704
- Structural assimilation, 842
- Structural conduciveness, 353
- Structural equation modeling (SEM),  
   908, 1692, 1910, 1914–1915,  
   1918–1923  
   benefits and use limitations,  
     1922–1923, 2346–2347  
   hierarchical linear models, 1922  
   and mobility research, 2817  
   significance of, 3039
- Structural functionalism. *See*  
 Functionalism and structuralism
- Structural Holes* (Burt), 737
- Structural lag, **3060–3067**  
   cultural lag concept and, 3066  
   and social structural responses,  
     3063–3066  
   and theories of change,  
     3066–3067
- Structural properties of  
 collectives, 1591
- Structural regulation of speech, 269
- Structuralism, 226, 1031, 1034, 1707  
   anthropological, 563, 564,  
     2891–2892  
   British, 568  
   French, 563, 1027, 1032–  
     1034, 1035  
   and German sociology, 1078  
   and Latin American studies, 1537  
   and Marxist sociology, 1753,  
     1754, 1755, 1756, 1784,  
     2162, 2163  
   and prison sociology, 2051
- and public opinion, 2273  
 and revolutions, 2413  
 and role theory, 2417  
 and social mobility, 2712, 2713  
 and social networks, 1034–1035,  
   2729–2731  
 theory of collective behavior,  
   352–354  
 theory of personality, 1717–1718  
 theory of poverty, 2211–2212  
 theory of state, 2162–2163  
*See also* Functionalism and  
 structuralism; Social structure
- Structurally unemployed. *See* Job  
 displacement
- Structuration theory, 226
- Structure of the Scientific Revolution,*  
*The* (Kuhn), 2023, 2374,  
 2458–2459
- Structure of the Social World, The*  
 (Rybicki), 2120
- Structured strain theory. *See*  
 Strain theory
- Strumilin, S. G., 2979, 3155
- Stryker, Robin, 388, 390, 1102, 1106,  
 1107, 1108, 1109
- Stryker, Sheldon, 1253, 1256, 1257,  
 2221, 2417, 2423, 2505, 2508
- Student movements, **3067–3070**  
   anti-Vietnam War, 2269–2270  
   Chinese Tianenmen Square,  
     2268, 2718, 2721–2722, 3067  
   civil rights participants, 2268,  
     2495  
   countercultures, 460, 461  
   and crowd behavior, 558  
   and German sociology, 1079  
   interaction theories, 351–352  
   life history data, 1635  
   and political criminality,  
     2144–2145  
   popular culture critique by, 2169  
   resource mobilization theories,  
     353–354
- Student Nonviolent Coordinating  
 Committee, 2495, 3069
- Students for a Democratic Society,  
 460, 3069
- Studies in Ethnomethodology*  
 (Garfinkel), 226
- Studies on Crime and Crime Prevention*  
 (journal), 2452
- Study of Sociology, The* (Japanese  
 journal), 1480
- Study of Sociology, The* (Spencer), 297
- Study to Understand Prognoses and  
 Preferences in Risks of  
 Treatment (SUPPORT), 587–588
- Study-findings comparisons. *See*  
 Meta-analysis
- Sturzo, Luigi, 1424
- Styron, William, 651
- Subcultures  
   countercultures vs., 459, 460  
   criminal and delinquent,  
     509–514, 534, 1494  
   cultural approach and, 566, 567  
   definition of, 509  
   stratification, 2818
- Subgame perfect equilibrium, 330
- Subjection of Women, The* (Mill),  
 988, 990
- Subjective well-being measurement,  
 2683–2684
- Subordinate Behavior  
 Description, 1566
- Sub-Saharan Africa  
   AIDS epidemic, 2591–2593  
   involuntary servitude, 2602, 3262  
   political and governmental  
     corruption, 2132–2134  
   population growth, 628  
   slave trade from, 2598, 2599  
   women in labor force, 3262
- Substance abuse. *See* Alcohol;  
 Drug abuse
- Substance Abuse and Mental Health  
 Administration, 715
- Substantive legal systems, 464
- Subtle prejudice, 2245
- Subunemployed, definition of, 1720
- Suburbanization, **3070–3077**  
   African American integration  
     and, 2498  
   alcohol consumption patterns  
     and, 94  
   in American urban system,  
     3070–3071, 3194  
   cities and, 311, 3070–3071, 3194

INDEX

- definition of, 3070–3071  
 exclusionary zoning and, 3072, 3073  
 manufacturing and trade employment, 3073–3074  
 minorities and, 3074–3075  
 and no-growth movement, 3072  
 and satellite developments, 3073–3074
- Suction curettage, 2238, 2239
- Sudan, 1866, 1941  
 slavery and forced labor, 2603–2604, 2606  
 sociodemographic profile, 2938
- Sudden infant death syndrome, 1640
- Suffrage. *See* Voting behavior; Voting rights
- Sufism, 3286
- Suharto, 2130–2131, 2978, 3068
- Suicide, 3077–3087**  
 American Indian rates, 135  
 analysis of prevention center calls, 431  
 anomic, 165, 3079  
 assisted suicide, 585, 586–587, 3083–3086  
 common traits, 3077  
 concept of, 3077  
 contagion theory, 3080  
 depressive disorder and, 650, 656, 3078–3079  
 Durkheim study, 574, 581, 584, 1575, 1595, 1736, 2024, 3055, 3077, 3079, 3080  
 Durkheim's study use of secondary data, 1595, 3080  
 euthanasia as, 585, 2719, 3083, 3088  
 gender and, 3078, 3079, 3081  
 Hindu widow self-immolation, 3255  
 history of, 3079  
 mass, 3077  
 methods, 3078, 3081  
 national differences in, 3079  
 physician-assisted, 585, 586–587, 3083, 3084–3086  
 predictors, 3078–3079, 3081  
 rational, 3084  
 self-destructive behaviors vs., 3077  
 social integration and, 584  
 sociological studies of, 3079–3081  
 study suggestions, 3082–3085  
 U.S. rates of completed, 3080
- Suicide, Le* (Durkheim), 574, 1024, 1575, 1736  
 as paradigm, 2024
- Suleyman the Magnificent, sultan of Ottoman Empire, 2999
- Sullivan, Albert, 582
- Sullivan, Harry Stack, 2085, 2088, 2092
- Sullivan, Mercer L., 510, 513, 665
- Sullivan, William M., 2080, 2484
- Sultanistic regime, 2133
- Sumer Empire, 2998
- Summary statistics, 659–661
- Summer learning research, 2933
- Summers, Lawrence, 794
- Sumner, William Graham, 581, 1400, 2192, 2330, 2882, 2986
- Sums of squares  
 calculation of, 158–159  
 decomposing, 159–160
- Sun Dance, 137
- Sundbärg, A. Gustav, 2450
- Sundt, Eilert, 2449
- Sunni Islam, 3286
- Super Bowl, 2985
- Superego, 1713
- Supernaturalism. *See* Religious orientations
- Supinski, Jzoeff, 2117
- Supplemental Security Income (SSI), 1145, 1146, 1286, 2799
- Supplementary Convention on the Abolition of Slavery, the Slave Trade, and Institutions and Practices Similar to Slavery, 2602, 2603
- SUPPORT program, 587–588
- Suppressor variable, 452
- Supreme Court, U.S.  
 and division of powers, 1953, 1954  
 function of, 1953  
 high status of justices, 476  
 numbers of cases heard, 472
- Supreme Court rulings  
 affirmative action, 50–51  
 antimiscegenation, 950  
 capital punishment, 2056  
 census sampling, 286–287  
 comparable worth, 372  
 desegregation, 2494  
 employment discrimination, 691  
 eugenics, 949  
 Internet First Amendment protections, 2185  
 obscenity definitions, 2186  
 peyote use, 137  
 pornography, 274  
 public opinion on, 2277  
 regulation of expression, 268, 270, 271, 272, 274  
 right-to-die cases, 585, 586, 587, 3083, 3084  
 school desegregation, 2493  
 segregation, 145, 2491  
 sexual harassment, 2580  
 sociological analysis of, 2960–2961, 2962
- Surgical abortion procedure, 2238, 2239, 2240
- Survey of Consumer Attitudes and Behavior, 578
- Survey of Income and Program Participation, 1284–1285, 1288, 1722, 2475
- Survey research, **3087–3094**  
 analysis, 3090  
 complex sample designs, 3034  
 content analysis, 418–421  
 cross-cultural analysis, 548  
 data collection methods, 575, 2475  
 of dating behavior, 3232–3236  
 distinctive characteristic of, 2468  
 elicitation/preference reversals in, 598  
 ethical and other problems, 3093–3094  
 factors in development of, 3232  
 measurement of, 1802–1803  
 mode of administration, 3090–3092



INDEX

- nonreactive studies, 2243  
online, 408–409  
questions and questionnaires, 3088–3090, 3093  
sampling procedures, 2444–2449, 3087–3088, 3232  
secondary data analysis, 2474–2481  
sex research problems, 2550  
social psychology, 2769  
statistical inference, 3023–3034  
survey sequence, 3092–3093  
voting behavior, 3203–3238, 3232  
*See also* Public opinion
- Survey Research Center (University of Michigan), 578, 2448, 2474  
Economic Behavior Program, 2476
- Surveys of Consumer Finances*, 578
- Survival analysis, 869
- Survivalists, 460, 461, 462
- Sussman, Marvin, 1389, 1391, 1737
- Sustainable development, 1222–1223, 1225
- Sutherland, Edwin H., 1203, 3098  
on moral crusade influences on law enactment, 1576  
theory of differential association, 507, 533, 666, 667, 1495–1496  
on white-collar crime, 3245, 3246–3247, 3248, 3249, 3250, 3251
- Suttee* (self-immolation), 3255
- Suttle, Gerald, 363, 365
- Svalastoga, Kaare, 2450, 2781, 2867, 2868, 2869
- Svoretz, John, 1790
- Swann, W. B., Jr., 2511
- Swanson, Guy, 1198
- Sweat lodge, 137
- Swedberg, Richard, 2921
- Sweden, 2449  
affinity of religion and family, 936, 940, 941  
antidiscrimination legislation, 693  
cohabitation as commonplace, 109  
divorce rate, 706  
family policy, 966–967  
health-care system, 374, 375, 377–378, 379  
legal system, 471  
long-term care and care facilities, 1654, 1655, 1661  
marital age model, 620  
Political Systems Performance Data, 2477  
postindustrialism, 2201  
retirement policy, 2408  
same-sex marriage legalization, 111  
Social Science Data Archive, 576  
social security system spending, 2796, 2800  
social surveys, 577  
transnational corporations, 3175, 3176  
unemployment, 3263  
woman suffrage, 703  
women in labor force percentage, 3262  
*See also* Swedish sociology
- Swedish Data Act of 1973, 2453
- Swedish Personal Data Protection Act of 1998, 2454
- Swedish sociology, 2450, 2451, 2452–2454
- Sweethearting, 1698
- Swerner, Michael, 2495
- Swidler, Ann, 2080, 2173, 2484, 2958, 2959
- Swiss code, 1550
- Switzerland  
“clean” government reputation, 2130  
divorce rate, 706  
drug policy, 712  
ethnic status incongruence, 3051  
family policy, 966  
health-care system, 375  
multilingualism, 2909  
social movement emergence, 2719  
Social Science Data Archive, 576  
social security system, 2800  
transnational corporations, 3175, 3176
- SWS. *See* Sociologists for Women in Society
- Sydney S. Spivack Program in Applied Social Research and Social Policy, 150
- Sykes, Gresham M., 1496, 2052, 2055
- Symbiosis, communitarianism and, 357
- Symbol systems, religious, 2382–2383, 3280–3283
- Symbolic anthropology, 2891–2892
- Symbolic Crusade* (Gusfield), 1576–1577
- Symbolic estates concept, 1510–1511, 1512
- Symbolic Interaction* (journal), 2293, 3096
- Symbolic interaction theory, **3095–3102**  
altruism attribution, 114  
basic premise of, 2423, 2767–2768  
on biological emotions vs. social sentiments, 2523  
central concepts, 3096–3098  
commonalities and variations, 3098–3099  
conversation analysis, 432  
critical theory and, 541  
description of, 2221  
ethnology and, 852  
and feminist theory, 996  
fluctuating interest in, 3099–3100  
fundamental imagery and framework, 3095–3096  
and gender identity, 1001, 1002  
on human nature, 1234  
identity theory and, 1253, 1254, 1255  
and juvenile delinquency, 1495–1497  
on marital satisfaction, 1729, 1731  
and medical sociology, 1815–1816  
on popular culture, 2173  
revitalization of, 3100  
and role theory, 2417, 2423–2424, 2856, 3097  
on sexual behavior, 2537–2538  
social interaction model, 2298

- social learning theory and, 1716–1717  
 as social psychology perspective, 2767–2768  
 and socialization, 2856, 2857  
 and sociology of sports, 2990  
 SYMLOG (System for the Multiple Level Observation of Groups), 1976–1978, 1979, 2614  
   coding sample, 1978  
 Sympathy, 782  
 Synanon, 899  
 Syncretism, African, 64, 65  
 Syntactical positioning, 2298  
 Synthetic cohort, 614  
 Synthetic metatheory, 1852–1853, 2880  
 Synthetic power theory, 2166  
 Syphilis, 2585, 2591, 2592  
 Syria, 1866  
   fertility decline, 628  
   Kurdistan nationalism, 1945  
   pan-Arab nationalism, 1944  
   sociodemographic profile, 2938  
 System for the Multiple Level Observation of Groups (SYMLOG), 1976–1978, 1979, 2614  
   coding sample, 1978  
*System of Logic, Ratiocinative and Inductive*, A (Mill), 818  
 System of National Accounts (SNA) activities, 3162, 3164  
*System of Sociology* (Sorokin), 2979  
 Systematic-simple random sampling (SSRS), 2446–2447  
 Systemic power, 2166  
 Systems theory, 1080, 1229, **3102–3106**  
   and intergroup/  
   interorganizational relations, 1405  
 Systems Theory and Operations Research, 3104–3105  
 Szacki, Jerzy, 2120  
 Szalai, Alexander, 3154, 3155, 3156–3157, 3159  
 Szasz, Andrew, 1103, 1106  
 Szczepanski, Jan, 2119  
 Szelenyi, Ivan, 2117  
 Szmata, Jacek, 2120–2121, 2122  
 Sztompka, Piotr, 2119, 2120, 2644, 3121  
**T**  
 Tabular analysis, **3107–3128**  
   census and, 283  
   contingency tables, 3036–3037  
   of data distribution, 658  
   elaboration and subgroup analysis, 3112–3114  
   measures of association, 3118–3126  
   multiway tables, 3036  
   odds ratios and log-linear models, 3115–3118  
   typologies, 3180–3188  
 Taeuber, Karl E., 2501  
 Tafarodi, R., 2511, 2513  
 Tailhook incident (1991), 1880  
 Taiwan  
   dependency theory, 642  
   fertility decline, 627, 2178  
   labor movement, 1532  
   occupational status attainment, 2786–2787  
 Tajfel, Henri, 2243  
 Takahashi, Akira, 1480  
 Takata, Yasuma, 1478  
 Takebe, Tongo, 1423, 1477  
 Takhtarev, K. M., 2979  
 Talarico, Suzette, 670  
 Tallman, Irving, 1790  
*Tally's Corner* (Liebow), 243  
 Tamils, 2366, 3288  
 Tammany Hall (New York City), 2125  
 Tamney, Joseph, 2946  
 Tampere University, 2450, 2451  
 Tanaka, Kakuei, 2131  
 Taney, Roger, 62  
 Tangri, S., 2580–2581  
 Tannen, Deborah, 2909  
 Tannenbaum, Franklin, 1496  
 Tannenbaum, P. G., 335, 336–337  
 Tanzania, 2040, 2583  
 Tape recording, 1974  
 Tappan, Paul, 3247, 3248, 3251  
 Tarde, Gabriel, 677, 679, 1025, 1423, 1477, 2117  
 Task activities (of crowds), 555–556, 557, 559  
 Task Force on Women in the Military (1988), 1880, 1881  
 Task specialization. *See* Division of labor  
 Task-focused leadership, 1566  
 Tasmania, 1069  
 TAT (Thematic Apperception Test), 2077  
 Tataki, Ronald, 180  
 Tate, D. C., 76  
 Tax revolts, 103  
 Taxation, 3000  
   and income distribution, 1282–1283  
 Taxonomy, 3181, 3183  
 Taxpayer Relief Act of 1997, 1282  
 Taylor, Charles, 356  
 Taylor, Edward, 1423  
 Taylor, Frederick, 697, 1014, 3155  
 Taylor, Ian, 1578  
 Taylor, Paul, 1858  
 Taylor, Shelley, 2190  
 Tchouprov, Alexandre, 1423  
 Teachers. *See* Education; Educational organization; Sociology of education  
 Teaching Quality Assessment (Great Britain), 227  
 Teaching Services Program, 150  
*Teaching Sociology* (journal), 411  
*Teaching Sociology with Fiction* (ASA publication), 1646  
 Team leadership, 1567  
 Teapot Dome scandal (1922–1924), 2127  
 Technological forecasting, 1038  
 Technological risk. *See* Society and technological risks  
 Technology and society  
   agricultural innovation, 88–89, 91  
   computer applications, 406–411  
   disaster planning and management, 684, 686  
   and evolution of paid work, 3262  
   fear of machines and, 2525–2526

- fragmentation and, 3104  
fundamentalist utilization of, 2372  
human ecology and, 1210  
information systems, 1344, 1345, 1346, 1347  
latent effects, 2680  
as materialist theory focus, 1785  
modernization theory on, 1885  
new institutional elites in, 2626  
organizational structure and, 2011–2012  
postindustrialization and, 2196–2203  
postmodernism and, 2200  
postmodernism on, 2205–2209  
and routinized vs. specialized division of labor, 698–699  
rural sociology and, 2430  
scientific research and development and, 2460–2461  
social change and, 2643, 2645  
social forecasting and, 2676–2677  
and social movement emergence, 557, 2719  
structural lag and, 3061, 3062  
and technological risks, 2874–2879  
in United States, 141  
urban sociology on, 3193  
urbanization and, 3196  
war and, 3242–3243  
and work and occupations, 3262, 3266–3267  
*See also* Industrial sociology; Industrialization; Industrialization in less developed countries; Internet
- Tecumseh (Shawnee leader), 136  
Teenagers. *See* Adolescence; Juvenile delinquency, theories of; Juvenile delinquency and juvenile crimes  
Teitelbaum, M. S., 425–426  
Telecommunications Act of 1996, 1762–1763  
Telephone  
computer-assisted interviewing system, 410, 1802  
conversation analysis, 431  
free expression and, 271  
survey interviews, 418–419, 578, 1802, 3092, 3232  
“Telephone law” (legal autonomy issue), 1549  
Televangelists, 2371, 2372  
Television  
aggression cues and, 74, 75  
American social effects of, 141, 1762, 1766  
as attitude influence, 185  
and innovation diffusion, 1763–1764  
and regulation of expression issues, 271, 272  
*Roots* miniseries, 67  
as socialization agent, 2858  
time-spent-viewing survey, 3160  
violence portrayed on, 1762–1763, 2858  
*Television and Social Behavior* (U.S. Surgeon General report), 1762  
Tellia, Bruno, 1468  
Temperament, definition of, 2074  
Temperance movement, 1577, 1580, 2717  
*Tempest, The* (Shakespeare), 2091  
TEMPO (time-by-event-by-member pattern observation), 1015, 1979  
Temporary Assistance for Needy Families program, 1145, 1146, 1288, 1395, 2804  
Temporary gatherings, 352  
Temporary help (THS) supply firms, 1724  
Temporary workers, 1724, 1725, 3268, 3275  
Ten Commandments, 943, 3283  
Tenbruck, Friedrich, 1075, 1079  
*Tendances et volontés de la société française* (essay collection), 1026  
“Tender years doctrine,” 702  
TenHouten, Warren, 1636  
Tension state, from inconsistency, 334–335, 337  
Tentori, Tullio, 1467  
Terman, Lewis, 999  
Terminal patients. *See* Death and dying  
Terminal values, 3214  
Territorial belonging, **3128–3137**  
definition of, 3128–3129  
distribution models of, 3133–3134  
localism-cosmopolitanism continuum, 3132  
multiplicity of, 3129–3130, 3133  
and social belonging, 2629–2630, 2632, 3129, 3131–3132  
subjectivity and, 3130–3134  
Territorial boundaries. *See* National border relations  
Terrorism, **3137–3142**  
counterculture groups, 460  
definition of, 3138  
escalation of, 2145  
genocide as, 1068–1069  
international, 3139–3140  
in Latin America, 1537  
religion and, 2362  
*Testing Statistical Hypotheses* (Lehmann), 3035  
Testosterone, 70  
Tetlock, Philip E., 3215  
Tetracycline, 677  
Teubner, Gunther, 1558–1559, 1560  
Texas A&M University, 682  
Texas War of Independence (1835–1836), 1856  
Text production  
computer-assisted, 407–408  
content analysis, 418–419  
*Textanalyst* (computer software), 420, 421  
Textor, Robert, 2917  
*TextSmart* (software), 419–420  
TFR (Total Fertility Rate), 627, 628, 629, 1867  
Thailand, 2974, 2975, 2978  
AIDS/STD control program, 2592  
family size, 978  
fertility rate decline, 2178, 2976  
labor movement, 1532  
legal system, 479  
political and governmental corruption, 2131  
sex traffic in, 2606, 2607  
slavery and slave-like practices, 2606–2607  
Thalidomide, 3250

INDEX

- Thamm, Robert, 784–785
- Thatcher, Margaret, 228, 2039, 2130
- Theft, 1576, 3251  
*See also* Burglary; Robbery
- Theil, Henri, 2501
- Theil's entropy index, 2501
- Their Eyes Were Watching God* (Hurston), 1648
- Theistic amalgamations, 3280
- Thematic Apperception Test (TAT), 2077
- Theocracy, 1245, 2356, 2357, 2359
- Theology. *See* World religions
- Theoretical construct models, 2029
- Theoretical knowledge, 2196–2197
- Theoretical Studies in Social Organization of the Prison* (Cloward), 2052–2053
- Theories of Cognitive Consistency: A Sourcebook* (Abel et al.), 335
- Theory construction  
 causal inference models, 266  
 definition of theory, 2463  
 falsification problem, 2464  
 paradigm vs., 2024  
*See also* Scientific explanation; *specific theories*
- Theory of Committees and Elections, The* (Black), 2920
- Theory of demographic transition, 1219
- Theory of global economy, 1197
- Theory of Justice, A* (Rawls), 2698
- Theory Of Moral Sentiments* (Smith), 782
- Theory of Political Coalitions, The* (Riker), 2920
- Theory of Psychological Units* (Abramowski), 2118
- Theory of Religion, A* (Stark and Bainbridge), 2375
- Theory of Social Becoming* (Sztompka), 2121
- Theory of the Leisure Class* (Veblen), 2986
- Theosophical Society, 3287
- Theosophy, 2366
- Therapeutic community (TC) movement, 2087
- Therapeutic control, 518
- Theravada Buddhism, 3282
- Therborn, Göran, 2452
- There Ain't No Black in the Union Jack* (Solomos, Findlay, Jones, and Gilroy), 226
- Thesaurus of Sociological Index Terms*, 1609, 1611–1612
- Thibaut, John, 2419, 2670
- Third party. *See* Triads
- Third Way, The: The Renewal of Social Democracy* (Giddens), 228
- Third World. *See* Developing countries
- Third World debt, 2433
- Third World feminism, 1708
- Thirteenth Amendment, 2601
- Thirty Years' War, 1933, 2356, 2362, 2363
- Thoits, Peggy, 782, 787
- Thoma, Stephen J., 1901–1902
- Thomas, A., 2088, 2090
- Thomas, Clarence, 2581
- Thomas, George M., 427–428
- Thomas, K. W., 1569
- Thomas, Mary Margaret, 1502
- Thomas, Melvin E., 2300, 2307
- Thomas, William I., 2332, 2498, 3098  
 life histories and narratives, 1616, 1618, 1633, 1634, 2118, 2220, 2221, 2222  
 on perception and reality, 2676, 2697
- Thompson, E. P., 226, 1198, 1707, 1756, 2169
- Thompson, Hunter, 1579
- Thompson, Paul, 2662
- Thompson, Sandra A., 439
- Thörnberg, E. H., 2451
- Thorndike, Edwin, 1
- Thornton, Sarah, 245
- Thought reform, 892–900
- Thoughtful vs. thoughtless aggression, 69
- Thrasymachus, 2337
- Three Mile Island nuclear power plant accident, 789, 805, 2875, 3249–3250, 3253
- Three Worlds of Welfare Capitalism, The* (Esping-Andersen), 377
- Thrill seeking, 664
- Thrupp, Sylvia, 1197
- Thucydides, 3241
- Thurrow, Lester, 2966
- Thurstone, L. I., 905, 914, 3035
- Tianenmen Square protest (1989; China), 2268, 2718, 2721–2722, 3067
- Tice, D. M., 2515
- Tidsskrift for Samfunnsforskning* (Norwegian journal), 2451
- Tilburg University, 1649
- Till, Emmett, 2493–2494
- Tilly, Charles, 385, 555, 606, 1707, 1942, 2270, 2271, 2819  
 on durable inequality, 2692  
 and historical sociology, 2918, 2919  
 and social structure analysis, 2823, 2827
- Tilly, Louise, 385, 1708–1709
- Tilly, Richard, 385
- Time and motion studies, 697
- Time budgeting. *See* Time use research
- Time Budgets and Human Behavior* (Sorokin and Berger), 3155–3156
- Time, interaction, and performance (TIP), 1015, 1979
- Time of measurement effects. *See* Period effects
- Time series analysis, **3142–3153**  
 classical, 3143–3151  
 comparative historical, 385  
 forecasting and, 2679, 2680  
 longitudinal research, 1691–1692  
 multiple time series design, 2318–2319, 3152  
 pooled, 2680  
 regression, 2679  
 of social indicators, 2681–2688
- Time Series Analysis* (Box and Jenkins), 3036
- Time series design, 2315–2316
- Time use research, **3153–3165**  
 contributions of, 3159–3160  
 daily time use (table of), 3160  
 leisure and work, 1582, 1584, 3262–3263  
 methodology, 3157–3159

- for social policy, 3163–3165  
tradition of, 3154–3155
- Time Warner, 1768
- Ting, Kwok-Fai, 3198
- Tingstrom, Daniel, 197
- TIP (time, interaction, and performance), 1015, 1979
- Tipton, Steven M., 2080, 2484
- Tiryakian, Edward A., 1886
- Title VII, Civil Rights Act of 1964, 372, 2591, 2706, 3264
- sexual harrasment, 2580
- Tittle, Charles R., 535, 668, 1491
- Tiv (people), 3079
- TM (Transcendental Meditation), 1015, 3287
- TNCs. *See* Transnational corporations
- Tobacco Movement (1890–1892; Iran), 1871
- Tobacco use. *See* Smoking
- Tobit model, 2439–2440
- Toby, Jackson, 1491
- Tocqueville, Alexis de, 229, 383, 1028, 2125, 2521  
on decay of religious faith, 2966  
democracy study, 601, 606, 1515, 1704, 2966, 3227  
and historical sociology, 1196, 1197  
and individualism as concept, 1301  
on intelligence as social equalizer, 1385  
on voluntary associations, 3227
- Toda, Teizo, 1478
- Todorsky, A. I., 2979
- Toennies, Ferdinand. *See* Tönnies, Ferdinand
- Tofranil, 654
- Togo, 2606
- Tokar, Brian, 1230
- Tolbert, Charles W., II, 2481
- Tolerance  
and democracy, 606  
for illegitimacy, 1259, 1260  
for intermarriage, 1412  
of nonconformist views, 317  
political correctness and, 2140  
of variety of social behaviors, 358
- Tolnay, Stewart, 626
- Tomaka, J., 2512
- Tominaga, Ken'ichi, 1480
- Toniolo, Ginseppe, 1464
- Tönnies, Ferdinand, 225, 356, 1423  
on community, 2520  
*Gemeinschaft* definition, 2630, 3129, 3134  
on *Gemeinschaft-Gesellschaft* shift, 355, 1074, 1502, 1506, 1772, 2426, 2483, 3130, 3135, 3182  
and German sociology, 1074, 1075  
as Japanese sociology early influence, 1477, 1478  
on public opinion, 2273
- Torrance, John, 224
- Tosi, Michelina, 328
- Total Fertility Rate, 627, 628, 629, 1867
- Total institution concept, 1673
- Total Quality Management (TQM), 605, 1665
- Total War (Nazi concept), 1067
- Totalitarianism, 1212, 1238  
and French School of Sociology, 1025  
labor movement subordination under, 1529  
personality type and, 317  
responsive communitarianism vs., 357  
studies of social sources of, 1310–1311
- Tour de France, 2985
- Touraine, Alain, 1533, 2205
- Tourangeau, Roger, 3091
- Tourism, **3165–3174**  
active vs. passive, 3166  
classic definition of, 3166  
fashions and trends in, 3168–3169  
impact of, 3171–3172  
leisure and, 1589  
motivations for, 3169–3171  
“Tourist gaze,” 3166–3167, 3173  
*Toward a Social Report* (Olson, 1969), 2682  
*Toward an Aesthetic of Reception* (Jauss), 1648
- “Toward the Proletarianization of Physicians” (McKinley and Arch), 1815
- Townsend movement, 2402, 2798, 3065
- Toxic threats, 2875
- Toyama, Masakazu, 1477
- Toynbee, Arnold J., 877, 2644
- TQM (Total Quality Management), 605, 1665
- Trade unions. *See* Labor movements and unions
- Tradition, 3216  
communitarianism and, 355
- Traditional authority, 229
- Tragedy of the commons, 595–596, 1220, 1223, 1229
- Trait theory. *See* Race; Personality theories
- Traité de sociologie générale* (Gurvitch), 1026
- Transaction cost analysis, 2340
- Transactional leadership, 1566
- Transatlantic slave trade. *See* African slave trade
- Trans-border cities, 311
- Transcendental Meditation, 1015, 3287
- Transformation of Corporate Control, The* (Fligstein), 738
- Transformational Grammar (Chomsky), 438
- Transformational leadership behavior, 1566, 1570
- Transitions. *See* Life course, transitions in; Life cycle
- Transnational corporations, **3174–3180**  
agribusiness, 2433  
and cities, 309–310  
corporate organization, 443–444  
division of labor in, 697  
globalization/unemployment relationship, 3263  
health care industry, 1827  
imperialism and, 1264  
materialist theory and, 1784  
regulation of, 3179–3180  
structural-functionalist case study of, 244

INDEX

- world leaders (table of), 3176  
 Transparency International, 2137  
 Transvestism and transsexuality, 2554, 2572–2573  
 Traoré, Moussa, 2133  
 Trappist monks, 3289  
 Travel. *See* Tourism  
 Treadway, Roy, 625, 626  
 Treas, J., 1998  
 Treasury Department, U.S., 714  
*Treatise on the Family* (Becker), 2920  
 Treaty of Rome (1957), 1550, 1934  
 Treaty of Utrecht (1713), 1933  
 Treaty of Versailles (1919), 1945  
 Treece, Davide, 1927  
 Treiman, Donald J., 370, 425  
     International Prestige Scale, 1997, 1999  
 Trevelyan, Lord, 2130  
 Treves, Renato, 1470  
 Triad societies, 512  
 Triadic dilemma, 464–465  
 Triads  
     balance theory, 335–336  
     coalitions within, 329, 331–332, 465  
     conformity tendency, 403  
     as court system model, 465–468, 474, 479  
     group size and, 1117  
     strain toward symmetry model, 336  
 Trials. *See* Court systems and law; Criminal justice system  
 Triandis, Harry C., 3218, 3222, 3223–3224  
 Tribal systems, 2809, 3180  
     *See also* Indigenous peoples  
 Tribute collection, 3000  
 Trichomoniasis, 2583  
 Tricyclic drugs, 654  
 Trinidad, 2600  
 Triplett, N., 2615  
 Trivers, R. L., 2883, 2884, 2885  
 Troeltsch, Ernst, 2366, 2378  
 Tropp, Asher, 225  
 Trotter, William Monroe, 66  
 Trow, Martin, 606, 1533  
 Troyer, Ronald J., 2763  
 Trubetskoi, Nikolai, 1032  
 “Trucking game” study, 2620  
 Trujillo, Rafael, 2134  
*Truly Disadvantaged, The* (Wilson), 3198  
 Truman, David R., 2624  
 Truman, Harry, 2126, 2127, 2274  
 Truncated sample, 2437  
 Trussell, James, 620, 634  
 Trust  
     efficacy and, 101–102, 103  
     personal dependency and, 2062  
     as social capital, 2640, 2734  
     as social order support, 2524–2525  
     in social organization, 2744–2745  
 Truth, 1249, 1250, 2957  
 Tshandu, Zwelakhe, 644  
 Tuchman, Gaye, 1648, 2172  
 Tufte, Edward R., 3008, 3011  
 Tukey, John W., 3006  
 Tulane University, 324  
     early clinical sociology course, 324  
 Tullio-Altan, Carlo, 1472  
 Tullock, Gordon, 2920  
 Tuma, Nancy Brandon, 1693, 1694, 2668, 3037  
 Tunisia, 1865, 1866, 1867  
     political corruption, 2132  
     sociodemographic profile, 2938  
 Tupes, E., 2085  
 Ture, Kwame (Stokely Carmichael), 53  
 Turkey, 1865, 1866  
     Armenian genocide by, 1070, 1384  
     and Cyprus conflict, 1945  
     drug crop control program, 713–714  
     fertility decline, 628, 1867  
     gender gap in education, 1867  
     health-care system, 375  
     historical-sociological analyses, 1869, 1870, 1871–1872, 1873  
     Kurdistan nationalism, 1945  
     layered legal system, 1550  
     nationalist movement, 3001  
     as Ottoman Empire, 2998, 2999  
     political corruption, 2132  
     rebellions, 3000  
     revolution (1919), 2411  
     secularization, 2485  
     sociodemographic profile, 2938  
     women in labor force percentage, 3262  
     World War I, 2362  
 Turks (ethnic), 692  
 Turku University, 2451, 2452  
 Turner, Frederick Jackson, 1933, 2918, 3232  
 Turner, Jay, 656  
 Turner, Jonathan, 416, 1237  
 Turner, L., 3167  
 Turner, Ralph, 351, 352, 460, 2069  
     on crowd behavior, 553, 554, 558, 559, 560  
     role theory, 2417  
 Turner, Victor, 2891, 3281–3282  
 Tuss, Paul, 194  
 Tutsi, 254  
 Tuval-Mashiach, Rivka, 1636  
 Tversky, A., 591–592, 595, 598  
 Twain, Mark, 277  
 Tweed, William (“Boss”), 2125  
 Twelve Country Project, 3156–3157  
 Twelve Tables of Rome, 464, 1513  
 Twelve-Step Programs, 714  
 “Two Field Theories” (Gold), 1013  
*Two Sacred Worlds: Experience and Structure in the World’s Religions* (Shinn), 3279  
 Two-party political systems, 2164  
 Two-person games, 329  
 Tydings-McDuffie Act of 1935, 175  
 Tylor, Edward Burnett, 563, 565, 1271, 2888, 2889  
     cultural diffusion theory, 675, 679  
 Typologies, **3180–3189**  
     constructed type, 3181  
     and continuous data, 3185–3186  
     empirical derivation, 3183–3185  
     ideal type, 3181  
     merits, 3185  
     polar types, 3182–3183  
     reduction, 3182–3183

- substruction, 3182
- U**
- U.S. Chamber of Commerce, 2162
- U.S. departments and agencies. *See* *key word*
- U.S. Government Document Depositories, 2477
- U.S. Health, 1822
- U.S. National Academy of Sciences. *See* National Academy of Sciences
- U.S. National Surveys*, 577
- U.S. Surgeon General report, 1762
- UCR. *See* Uniform Crime Reports
- Udy, Stanley, 232
- Uganda, 2133
- AIDS/HIV epidemic, 2591, 2592
  - poverty in, 2216
  - and Rwandan genocide, 68, 1069
  - slavery and forced labor in, 2603, 2606
- Uighurs, 3001
- Ukraine, 2137, 2362, 2982
- Ulmer, Jeffrey, 2221–2222
- Ulrich's International Periodicals Directory*, 1606
- “Ultimate” bargaining game, 597
- Ultimate judgement, 597
- “Ultimatum game,” 597
- Unconscious, 1713, 1714
- Uncover Reveal* (journal article-alerting service), 1611
- Underclass. *See* Urban underclass
- Underdeveloped countries. *See* Dependency theory; Developing countries; Industrialization in less developed countries;
- Underemployment. *See* Marginal employment
- Underground economy, 1339
- Underidentified model, 1916
- Understanding Prediction: Essays in Methodology of Social and Behavioral Sciences* (Nowak), 2120
- Unemployment
- global rates, 3263
  - layoffs and displacement, 3265
  - programs, 2795, 2798
  - rate in rural areas, 2430
- rate measurement, 1521, 1720
  - urban underclass and, 3198
- UNESCO. *See* United Nations Educational, Scientific and Cultural Organization
- Ungaro, Daniele, 1473
- UNICEF, 2607
- Unification Church, 900, 2379
- affinity of religion and family in, 936
  - as new religious movement, 2366, 3287
- Uniform Code of Military Justice, 1881
- Uniform Commercial Code, 475, 476
- Uniform Crime Reporting Handbook* (FBI), 493
- Uniform Crime Reports (UCR), 491–500, 530
- arrests by race, 1490
  - Census Bureau survey as supplement, 1489
  - and family violence, 982
  - general definitions of offenses, 492
  - and juvenile crime, 1486–1487, 1489
  - and sex crimes, 2557
- Uniform distribution, shape of, 661
- Unikel, Luis, 1859
- Unilineal kinship systems, 1502–1503, 1507
- Union Carbide, 2877
- Unions. *See* Labor movements and unions; Industrial sociology
- United Arab Emirates, 1865, 1866, 1867
- sociodemographic profile, 2938
- United Arab Republic, 1947
- United Furniture Workers, 1533
- United HealthCare, 1822
- United Kingdom
- African colonization, 60, 61
  - colonial racial conflicts, 321
  - common law system, 465, 474–475, 477–478, 480, 1545
  - communitarianism, 361, 362
  - conditions conducive to democracy, 605–606
  - crime rate surveys, 498–499
  - divorce reform laws, 703
  - drug policy, 712, 713
  - education and status attainment, 3045
  - Elizabethan Poor Laws, 2840–2841
  - ethnic immigrants, 636
  - ethnic status incongruence, 3051
  - fertility transitions, 2178
  - first modern police force, 2110–2112, 2113
  - foreign-controlled pharmaceutical companies, 1827
  - governmental division of power, 1953
  - health-care system, 374, 375, 376, 378–379, 380, 1827
  - juvenile violence, 1487
  - kinship systems and family type, 1504
  - labor movement, 1529, 1533
  - long-term care and care facilities, 1652, 1653, 1655, 1659, 1661
  - organizational demographics, 395
  - political party system, 2154, 2164
  - political scandals, 2130
  - postcolonial India and, 641
  - racial discrimination, 692, 693
  - and rebellion, 3000
  - retirement patterns, 2407
  - slave trade suppression by, 2600, 2602, 2607
  - social anthropology, 2890, 2893
  - Social Science Data Archive, 575, 576, 580
  - social security system, 2796, 2797
  - social surveys, 577
  - solicitor-barrister distinction, 478
  - Southeast Asian influences by, 2975
  - time use research, 3164
  - tourism in, 3169
  - transnational corporations, 3175, 3176
  - unemployment, 3263
  - utopian literature, 3203–3204
  - woman suffrage, 703
  - World War II, 2362

INDEX

- United Methodist Church, 2376
- United National Development Programme, 3162
- United Nations, 240  
 beneficiaries of human development programs, 646  
 and census, 282  
 Conference on Adult Education, 23  
 creation of, 1427  
 Crime and Justice Network, 499  
 Development Program, 2215  
 Economic Commission for Latin America, 639–640, 1537  
 genocide definition, 1066, 1071–1072  
*Human Development Report*, 2917  
 and human rights abuses, 1944, 1948  
 Human Rights Commission, 2607  
 human rights declaration, 1240–1242  
*See also* Universal Declaration of Human Rights  
 International Labour Office, 1531, 2795, 2800  
 and international law, 1429  
 and international terrorism, 3139–3140  
 life expectancy ranking by, 1631  
 natural disaster research, 686  
 New World Information and Communication Order, 1767  
 peacekeeping activities, 1429, 2046, 2048  
 population policy, 635  
 population projections, 2181, 2182  
 and poverty in low-income countries, 2215–2216  
 research funding in China, 302  
 slavery and involuntary servitude convention, 2602, 2603, 2607  
 social statistics reports, 2916–2917  
 and state sovereignties, 3003  
 time use research, 3162  
 transnational corporation regulation, 3179  
 and urban agglomeration, 310  
 as war prevention force, 3243  
 World Conference on Women, 932, 1768
- United Nations Educational, Scientific and Cultural Organization (UNESCO), 23, 280, 426, 1038, 1767, 1768  
*Main Trends in the Social and Human Sciences*, 2917  
 time use research, 3156
- United Nations University, 1039
- United States. *See* American families; American society
- United States departments and agencies. *See key word, e.g.*, Justice Department, U.S.
- Univac I (computer), 406
- Universal Communitarian Association, 355
- Universal cultural order (Znaniecki theoretical system), 2118
- Universal Declaration of Human Rights (1949), 742, 1240–1242, 1245, 1319
- Universal health care, 377
- Universal pragmatics (Habermas concept), 543
- Universalism, 1301, 3216
- Universities. *See* Higher education; *names of specific institutions*
- Universities: American, English, German* (Flexner), 325
- University of California, as multiuniversity, 1180
- University of California, Berkeley  
 Asian-American research, 179, 180–181  
 Free Speech Movement, 3069–3070  
 interactive database access, 409  
 student protest theory, 351–352
- University of California, Los Angeles, Asian-American Studies Center, 179
- University of Chicago  
 Asian-American studies, 176, 177, 178  
 Department of Sociology. *See* Chicago School  
 deviance theories, 664–665  
 ethnographic tradition, 243–244  
 first clinical sociology course, 324, 325  
 and Japanese sociology, 1477, 1478  
 leisure study, 1582  
 National Opinion Research Center, 167, 577, 1479, 2300  
 peacetime disaster research, 681  
 social change studies, 2682  
 and symbolic interactionism, 2856  
 voting behavior analysis, 3232
- University of Chicago Press, 176, 2372
- University of Cologne  
 Cologne School, 1077, 1078  
 Zentralarchiv fur Empirische Sozialforschung, 575, 576, 579, 580, 2477
- University of Colorado, 682
- University of Connecticut, Roper Center for Public Opinion Research, 2477
- University of Copenhagen, 2450
- University of Delaware, 681
- University of Essex, 227  
 Economic and Social Research Council Data Archive, 576, 580  
 Summer School in Social Science Data Analysis and Collection, 421
- University of Illinois, 1979  
 Survey Research Laboratory, 2445
- University of Kharkiv, 2982
- University of Kiev, 2982
- University of Koln. *See* University of Cologne
- University of Lancaster, 227
- University of Leeds, 226, 2039
- University of London. *See* London School of Economics
- University of Lviv, 2982
- University of Michigan  
 Center for Political Studies, 101  
 hate speech code, 276  
 Institute for Social Research, 578, 2299, 2300  
 interactive database access, 409



- Inter-University Consortium for Political and Social Research (ICPSR), 409, 575, 576, 577-578, 579, 1606, 2474-2475, 2476, 2477, 2481
- leader behavior study, 1565
- Survey Research Center, 578, 2448, 2474, 2476
- voting behavior research, 3233, 3234-3236, 3237
- University of Minnesota
- “Death Education” program, 581, 582, 585
- interactive database access, 409
- Southeast Asian Refugee Studies Project, 182
- University of North Carolina, Chapel Hill
- Louis Harris Data Center, 2477
- Social Science Data Archive, 576
- University of Odessa, 2982
- University of Oklahoma, 681
- University of Texas, 681
- University of Tokyo, 1477, 1478
- University of Washington, Asian-American studies, 179
- University of West Virginia, 1617
- University of Wisconsin, Madison, Social Science Data Archive, 576
- University of Maryland, 1876
- Unresponsive Bystander, The: Why Doesn't He Help?* (Latané and Darley), 115
- Unruh, Jesse, 2397
- Unscheduled Events* (newsletter), 682
- Unsuru'l-Ma'ali, 1564
- Unto Others* (Sober and Wilson), 118
- Untouchables, 250
- Unwed childbearing, 125, 128, 484, 488, 634, 1258-1264, 1506, 1626, 1744, 2033
- Upper class. *See* Social and political elite; Social class; Status attainment
- Upper quartile, definition of, 659
- Uppsala University, 2452
- Uprisings. *See* Protest movements; Revolutions
- Upward mobility, 2714, 2868, 3043-3044
- See also* Status attainment
- Urban II, Pope, 2967
- Urban agglomeration, 310
- Urban ecology, 843, 2922, 3071
- Urban political economy, 1214
- Urban riots, 555-556, 557, 558, 2269, 2270, 2495, 2661
- Urban sociology, **3191-3198**
- community breakdowns and, 664
- community studies, 363, 365
- and crisis of new urban theory, 1215-1216
- deviance theories, 664-665
- drug abuse prevention programs and, 714-717
- economics and economic geography influences on, 2922
- ethnic succession and, 532-533
- fertility transition and, 625
- gang behavior and, 663, 665
- human ecology and, 843, 1215-1216, 2922
- and industrial sociology research, 1309
- and Italian sociology, 1468
- kinship systems and, 1502
- macro-level deviance theories and, 663, 671
- Mexican studies, 1860-1861
- political economy and, 1214-1215
- and population distribution, 633
- and population redistribution, 2179-2180
- and social networks, 2728
- and social reform, 365
- See also* Cities; Rural sociology; Urban underclass; Urbanization
- Urban underclass, **3198-3201**
- alternate causal views of, 3198
- case studies of, 243
- and centrifugal kinship system, 1512
- class and race issues, 319
- crack violations, 713
- and criminal and delinquent subculture theory, 513, 1494
- culture-of-poverty concept vs., 2212
- definition of, 2212, 2497-2498
- deviant behavior theories, 663-665, 669-670, 671, 1494
- divergent language varieties and, 2901
- and ethnicity, 845
- failure of integration and, 2496-2498
- family structure, 122
- and global economy effects, 311
- measurement approaches, 3198-3200
- and organized crime, 2019
- and poverty theory, 2212-2213
- underemployment and, 1721
- and violent crime, 1490
- “Urbanism as a Way of Life” (Wirth), 367, 1772
- Urbanization
- cities and, 306, 309
- and community of limited liability thesis, 367-368
- comparative features of major world regions, 3194
- and convergence theory, 3197
- and crime rates, 532
- and deurbanization, 3195
- and deviance theories, 2658-2659
- divergent patterns of, 1211
- ecology and, 1209
- and education and development, 744, 749
- and environmental sociology, 805-806
- and expansion in interrelationships, 3193
- and fertility determinants, 1009
- and globalization, 306, 309, 311
- homelessness and, 1203
- industrialization and, 1318-1319
- kinship systems and, 1502, 1503
- mental illness incidence and, 1840, 1841
- in Mexico, 1857, 1859-1860
- in Middle Eastern countries, 1866
- and modernization theory, 1885
- and population growth, 2179
- population of world's largest metropolises, 3195
- rural societies contrasted with, 3192

- sociology of law on, 1576  
 technological advances and, 3196  
 and urban sociology, 3191–3197  
*See also* Cities; Urban underclass
- Urine testing. *See* Drug testing
- Urry, J., 3166, 3168, 3169
- Uruguay  
 clinical sociology, 328  
 demographic characteristics, 1535, 1536  
 gross national product, 1535  
 political, economic, and social conditions, 1536, 1537, 1540
- Useem, Bert, 352
- Useem, John, 2901
- Useem, Michael, 443, 736, 2162
- UseNet, 1442–1443
- UseNet Newsgroups, 414
- USSR. *See* Soviet and post-Soviet sociology; Soviet Union
- USSR Academy of Sciences (Moscow), 2980
- Utilitarianism  
 and criminal sanctions, 515, 516  
 and ethics in social research, 839–840  
 and money, 1892  
 and nineteenth-century liberalism, 1599–1600  
 pleasure-pain principle, 2087  
 and social exchange theory, 2670
- Utility theory. *See* Decision-making theory and research; Rational choice theory
- Utopia* (More), 3201–3202
- Utopian analysis and design, **3201–3206**  
 communitarianism, 355, 460–461  
 community failures, 2849  
 counterculture, 460–461, 462  
 Mannheim's theory of, 2955–2956  
 nostalgia and, 2847  
 prediction and, 2224–2225  
 social forecasting and, 2677  
 socialist, 2847, 2849, 3202–3203
- Uusitalo, Hannu, 2451
- Uzbekistan, 2982  
 sociodemographic profile, 2938
- V**
- Vacancy chain models, 2691
- Vaccines, 622, 878, 2177
- Vaillant, Caroline O., 1729
- Vaillant, George E., 1729
- Vaishya (Hindu trader), 250
- Vajpayee, Atal Bihari, 2132
- Valachi, Joseph, 2018
- Valanides, Nicolas, 1901
- Valente, Thomas, 676
- Validity, **3207–3212**  
 concurrent, 3208  
 convergent, 3210  
 definition of, 3207  
 differential prediction, 3211  
 generalization, 3210–3211  
 of hermeneutic explanations, 2472  
 important points about, 3207  
 in life narrative, 2292  
 in moral judgment measurement, 1898  
 and multiple indicator models, 1908, 1909–1910, 1917, 1921–1922  
 in narrative history, 2292  
 in personality measurement, 2080  
 predictive, 3208  
 proportional reduction in error and, 1810–1811  
 qualitative data criteria, 2288–2289  
 quasi-experimental research designs, 2312–2315, 2317, 2323–2327  
 reliability vs., 1909, 2080  
 replication and, 2395–2397  
 in social comparison process, 2641  
 time use research, 3159  
 types of, 3207–3210  
*See also* Construct validity
- Valkonen, Tapani, 2453
- Value-added theory, 352
- Value-free analysis. *See* Epistemology; Positivism; Scientific explanation
- Values theory and research, **3212–3227**  
 and collective behavior, 352, 353, 3213, 3218–3219, 3222  
 conceptualization of values, 3213  
 and evaluation research, 867–888  
 individualism and collectivism, 3218–3219  
 Inglehart's postmodern thesis, 3222–3223  
 ordered priorities of values, 3213  
 rational choice theory and, 2336, 2338  
 Rokeach Value Survey, 3214–3216  
 Schwartz Scale of Values, 3216–3217, 3218, 3219  
*See also* Social values and norms
- Van den Berg, Axel, 2220
- Van den Berghe, Pierre, 841, 842, 1412, 2883
- Van der Ploeg, Jan Douwe, 2432–2433
- Van Lange, Paul A. M., 3221
- Van Maanen, John, 247–248
- Van Rees, Cees, 1649
- Vander Mey, Brenda, 1276
- Vanni, Icilio, 1464
- Varela, Francisco J., 1557, 2088
- Variables  
 bivariate relationship measures, 660  
 categorical and limited dependent, 3037–3038  
 definition of, 2249  
 in effective persuasion, 2094–2097  
 frequency distribution table, 658  
 graphic representation, 3015  
 hierarchical linear multilevel, 1173  
 less than normally distributed, 1796–1800  
 and level of analysis, 1591–1592  
 measures of association and, 1804–1812, 1966  
 models with excluded, 2251–2252  
 multiple indicators, 1907–1923  
 in narrative reviewing, 1845  
 precise specification of, 1803  
 reliability of measurement, 2343–2355  
 scatter plot shape and direction, 661

INDEX

- and secondary data analysis, 2478  
time series regressions, 2679  
time-varying independent, 1691  
validity generalization, 3210–3211  
*See also* Analysis of variance and covariance
- Variance (statistical)  
definition of, 158, 659–660  
*See also* Analysis of variance and covariance
- Varieties of Police Behavior* (Wilson), 2108, 2114
- Vasudev, Ivotsna, 1901, 1902
- Vatican II. *See* Second Vatican Council
- Vatican Bank, 3250
- Vaughan, Diane, 2878, 3250
- Vaughn, John C., 2792
- Vaupel, J., 1629
- V-chip (electronic device), 1762–1763, 2185–2186
- VCRs (videocassette recorders), 2185
- VDL model, 1567
- Veblen, Thorstein, 724, 735, 1423, 1576, 2209, 2520, 2986, 3054  
conspicuous consumption concept, 3168
- Veditz, C. W. A., 1422
- Veiling of women (Islamic), 2949–2950
- Vemer, Elizabeth, 2390
- Vendee, The* (Tilly), 2918
- Venezuela  
economic liberalization, 1539, 1541  
fertility decline, 627  
political corruption, 2134, 2135
- Ventura, Jessie, 3068
- Verba, Sidney, 2282, 3235
- Verbal ability, sex differences study, 2531
- Verbal behavior. *See* Conversation analysis
- Verkko, Veli, 2450
- Veroff, J., 2303
- Verret, M., 1026
- Vertical-dyad linkage (VDL) model, 1567
- Veterans Administration, 374–375, 582
- Veterans Health Administration, 588
- Viagra, 2555
- Victimization surveys, 491–490, 530, 535, 549  
and feminist theory, 989
- Victimless crimes, 1576, 1577–1578
- Victorian era, 2125  
children's status, 2036  
women's status, 1697
- Videocassette recorders (VCRs), 2185
- Vienna Centre, 3156
- Vienna Circle, 821, 1027, 1028, 2192–2193, 2756
- Vierkandt, Alfred, 1074, 1075
- Vietnam, 2974, 2978  
anticolonial revolution, 3001  
family size, 978  
French colonization in, 2975  
independence movement, 2975  
peace conference triad (1954), 332  
peasant rebellion study, 2977
- Vietnam War, 1875, 1877, 1878, 2975  
African American conscription and casualties in, 1879, 2270  
antiwar protest movements, 2266, 2269–2270, 3067, 3069, 3139  
free expression and, 273, 274  
heroin use, 711  
marriage rate and, 1741, 1744  
political alienation and, 101, 103  
public opinion polling, 2275, 2276  
sociology of combat and, 1881  
student movements and, 3069
- Vietnamese Americans, 175  
household structure, 127  
as refugees, 180, 181
- Vigilante groups, 2109
- Village life. *See* Rural sociology
- Vincelli, Guido, 1467
- Vinogradoff, P., 1025
- Violence  
as aggression subtype, 68  
in cities, 311  
civil rights movement, 2495  
community studies, 365  
facilitators of collective, 349  
genocide and, 1066–1072  
juvenile, 1484–1485, 1487–1488, 1490–1491, 1494, 1498  
kin selection and ethnic, 2882–2883  
literary sociology and, 1650  
media cues to, 74, 75, 272, 1762–1763, 2858  
nationalist movements and, 1199, 1947–1948  
political, 2145  
protest movements and, 2266, 2269–2270  
riots and, 555–556  
self-esteem and, 2516  
sexual jealousy and, 2886  
societal reduction measures, 76  
subcultures of, 534, 663–664  
terrorist, 3137–3142  
treatment of, 521  
of war, 3241–3245  
*See also* Aggression; Crime, theories of; Crime rates; Criminology; Family violence; Homicide; Revolutions; Sexual violence and exploitation
- Violence Against Women Act of 1994, 985
- Violence and the Police* (Westley), 2114
- Virginia Company, 3174
- Vishnu (Hindu deity), 3280
- Vision quests, 3277–3278
- Vital registration system, 632
- Vocational training, 1488, 3263
- Vodun, 65
- Vogel, S. R., 2190
- Vogel, Steven, 1104, 1105, 1106, 1109
- Vögelin, Eric, 1075
- Voice of America, 274
- Voice recognition software, 408  
survey research analysis, 418
- Vold, George, 1497, 3247
- Volkgeist* (folk spirit), 475
- Volpato, Mario, 1469
- Voltaire, 2482–2483
- Voluntary associations, **3227–3231**  
altruism and, 114, 115, 118

- characteristics and objectives, 3227–3228
- communitarian view of, 358–359, 360
- current research, 3228–3229
- origin of, 3227
- political organizations, 2147–2152
- Voluntary childlessness. *See* Childfree adults
- Volunteering, 114, 115, 118
- Von Beyne, Klaus, 2153, 2154, 2155
- Von Neumann, John, 329, 591, 1045, 1049, 2335
- Voting behavior, **3231–3240**
- aggregate data analysis, 3232–3233
  - American trends, 144
  - issue voting, 3235–3236
  - liberal/conservative, 1602–1603
  - mass media research on, 1765–1766
  - normal vote vs. realignment, 3235
  - party identification and, 3234, 3236
  - political alienation factors, 101, 103
  - political elites and, 2627, 3233, 3234
  - polls, 575, 1686, 2273–2274, 2278–2279
  - rational choice theory, 2339
  - social class relationship, 604, 3233, 3234
  - survey-based research, 3232, 3233–3238
- Voting rights
- African American, 58
  - exclusions, 602
  - historical development of, 3001 and political party systems, 2154 and representativeness, 2156–2157
  - for women, 703
  - women’s protest movements, 2266
- Voting Rights Act of 1965, 58
- Vroom, V. H., 1568, 1569
- Vroom’s expectancy model of motivation, 1568
- Vujacic, Veljiko, 1199
- W**
- W. K. Kellogg Foundation, 2400
- Wach, Joachim, 3278, 3283, 3286
- Waddington, David, 353
- Waerness, Kari, 2453
- Wage gap. *See* Comparable worth
- Wage labor system. *See* Labor force; Labor movement and labor unions; Work and occupations
- Wagenfeld, Morton O., 1840
- Wageningen School, 2432–2433
- Wages and salaries
- comparable worth and, 369–372, 2141, 2706
  - discrimination in, 689–691, 3265
  - global inequalities in, 2691, 2705–2706
  - human capital theory on, 1989–1990, 1991
  - inequality from marginal employment, 1721–1722
  - job shifts and, 1983
  - in legal profession, 470
  - life-cycle pattern in, 1624–1625, 1983
  - as occupational prestige factor, 2000
  - percentage of household income, 1719
  - and status attainment, 3047–3048
  - of temporary workers, 1724
  - women’s earnings gap, 370–372, 984, 1720, 3265
  - work generating, 3261
  - and work hours, 3262–3263
  - worker characteristics and, 1991–1993
- See also* Income distribution in the United States
- Wahid, Abdurrahman, 2131
- Wainer, Howard, 3008
- Waite, Linda, 360
- Waitzkin, H., 2288–2289
- Wakenhut, R., 1899
- Waldegrave, James, 1045
- Walden Two* (Skinner), 3204
- Waldinger, Roger, 1872
- Walesa, Lech, 2268
- Walker, Lawrence J., 1902
- Walker Report (1968), 555
- Wallace, Anthony F. C., 2966
- Wallace, Karl M., 1727
- Wallace, Michael, 2961
- Wallace, Samuel, 1203
- Wallas, Graham, 225
- Waller, Willard, 2927, 2986, 3098
- Wallerstein, Immanuel, 383, 387, 645, 697, 728–729, 1265, 1785 and historical sociology, 2297
- world system theory, 1089–1090, 1706, 1758, 1876, 2646, 2690
- Walster, Elaine, 1508, 2700, 2701
- Walster, G. William, 1508, 2700, 2701
- Walton, John, 1706–1707
- Waltz, Kenneth B., 3241
- Walzer, Michael, 356
- War, **3241–3245**
- as aggression, 68
  - balance of power coalitions and, 332–333, 3242
  - censorship and regulation of expression and, 274
  - combat study, 1881
  - conflict theory on, 415, 3243–3244
  - contemporary conflict patterns, 2363
  - decision making and, 3244–3245
  - decolonization and, 1267
  - democracy in relationship with, 605, 3244
  - divorce rates and, 701, 702
  - ecology and, 1214
  - and enslavement of prisoners, 2601, 2603
  - evolutionary perspective on, 878, 1029
  - genocide and, 1066, 1067, 1068–1069
  - international law on, 1429–30
  - interstate system and, 3241–3243
  - rape and, 2578–2579
  - religion and, 2356, 2362–2363
  - social evolution and, 898, 1029
  - societal socialization and, 1068–1069

- sociology of combat and, 1876–1877, 1881, 1882  
systemic theories of, 3242  
terrorism and, 3140  
total, 1067  
widow status from, 3255  
*See also* Peace; Revolutions
- War crimes, 1429, 1948  
War on Crime, 2112–2113  
War on Drugs, 2760  
War on Poverty, 1286, 1494, 1882, 2404, 2760  
Ward, Frank Lester, 2192  
Ward, Lester, 168, 1423, 1424  
Warhol, Andy, 3079  
Waring, E., 3251  
Waring, Joan, 345  
Waris, Heikki, 2451  
Wark, Gillian R., 1902–1903  
Warner, Lloyd W., 363, 364–365, 582, 583  
Warner, Malcolm, 1533  
Warner, Stephen, 2367, 2374–2375, 2485, 2964  
Warner Brothers, 1768  
Warning labels, 1640  
Warren, B., 1085  
Warren, Carol A. B., 245  
Warren, Robert, 367  
Wars of religion, 2363  
Warsaw Pact, 2362  
Washington, right-to-die issue, 586–587  
Washington, Booker T., 66  
Washington, George, 285, 2125  
Washington University, 1834  
Wasilewski, Jacek, 2121  
Wasserman, Ira M., 3079, 3081  
Watanuki, Joji, 1480  
Watergate scandal (1972–1974), 1764, 1766, 2127  
    public opinion and, 2274, 2276, 2277  
    as white-collar crime research impetus, 3248  
Waters, Mary, 843  
Watson, Alan, 1157, 1555–1556, 1560  
Watson, John B., 1249, 1505  
Watt, W. Montgomery, 2939, 2945  
Watts, Jon, 375  
Watts (Los Angeles) riots, 2495  
“Weak ties” theory. *See* “Strength of weak ties” hypothesis
- Wealth  
    distribution of, 1283–1285, 1311  
    and distribution of power, 416  
    and family size, 975–976  
    materialist theory on inequality of, 1784–1785  
    money as material representation of, 1890  
    slave-created, 55–56  
    and social justice, 2705–2706  
    and social origins of deviance, 663  
    *See also* Economic institutions
- Wealth and Poverty* (Gilder), 723  
*Wealth of Nations, The* (Smith), 697, 721, 1771, 2340
- Weapons of mass destruction, 3242, 3243
- Web. *See* Internet; Web sites
- Web sites  
    content-analysis, 421  
    futures studies, 2231–2232, 2957  
    quality-of-life studies, 2686  
    religious movements literature, 2375  
    research grants, 2400  
    social indicators, 2685  
    sociological journals, 1606  
    sociology-related, 406–407, 413–414  
    survey sampling specialists, 2445
- Webb, Beatrice, 225, 853, 1310, 1423  
Webb, G. L., 2053  
Webb, Sydney, 1310, 1423  
Weber, Alfred, 1074, 1075, 1235  
Weber, Max, 542, 568, 939, 1028, 1116, 1234, 1754, 2193, 2208, 2265, 2889  
    British sociology and, 226  
    on bureaucracy, 229–230, 231, 232–233, 603, 697, 698, 1312, 2163, 2623, 2627  
    causal interpretation, 387  
    church-sect typology, 2365, 2366–2367, 2373, 2378  
    class definition, 2163  
    and community, 362  
    and comparative-historical analysis, 383, 386, 389  
    conflict model, 414, 415  
    on death, 581  
    on democracy, 603  
    and economic sociology, 721, 722, 732, 733–734, 806, 2640, 2690  
    and emotions, 774–775, 785  
    and epistemology, 820, 822  
    and ethnography, 852  
    on *Gemeinschaft*, 2630  
    and German sociology, 1074, 1075, 1077  
    and historical sociology, 1196, 1197, 1199  
    ideal type definition, 3181  
    Indian sociology and, 1290, 1291  
    on industrialization, 2643  
    on interpersonal power, 1456, 1458  
    on Islamic society, 2941, 2942–2943, 2946  
    as Japanese sociology early influence, 1478  
    law definition by, 1544–1545  
    legal conceptualizations, 1553–1554  
    legal system typology, 464, 1546–1547, 1548, 1549, 1554  
    macro and micro sociological themes of, 1704  
    and modernization theory, 1885  
    on musical notation, 1924  
    on nations and nationalism, 1199  
    on origins of capitalism, 237, 238, 542  
    and phenomenology, 2099  
    philosophical influence on, 2756  
    and political organization, 2162, 2163–2164  
    on political party system, 2153  
    power conception, 2165, 2865  
    Protestant ethic concept, 237, 389, 722, 774, 2483, 2986, 3219, 3222

- and secularization tradition, 2483  
on social belonging, 2632  
on social change, 2645, 2647  
and social inequality concept, 2690  
and social status, 3049, 3050  
social stratification theory, 2815–2816, 2927  
societal stratification theory, 2865, 2868, 2967  
and sociological significance of family therapy, 247  
and sociology of knowledge, 2954, 2955, 2957  
and sociology of religion, 2373  
on state as institutionalization of power, 2162  
on sultanistic regime, 2133  
and territorial belonging, 2629–2630  
and theory of action, 2099  
and theory of rationalization, 541  
on types of action, 2519–2520  
and values theory, 3219, 3222  
warning against monocausal theory, 724
- Wechsler Adult Intelligence Scale, 1362–1363, 1376
- Wedding ceremony, function served by, 1734–1735
- Wedlock, Eldon D., 2143, 2146
- “We-feeling,” 2632
- Wegener, Bernd, 1997, 2705
- Weibull model, 871–873
- Weick, Karl E., 1975
- Weill, Felix, 1075
- Weimar Republic (Germany), 606, 2268  
German sociology during, 1074–1075  
voting behavior research, 3233
- Weinberg, Ian, 424
- Weinberg, S. Kirson, 1275
- Weiner, Bernard, 194, 195
- Weiner, Marcella B., 2555–2556
- Weiner, Myron, 2154
- Weis, D. L., 2538
- Weisbund, D., 3251
- Weisman, Avery, 582
- Weiss, Carol H., 2284–2285
- Welfare Reform Act of 1996, 1261
- Welfare state  
in American society, 144  
comparative research, 376–377, 930  
convergence theories, 426–427  
countercultures and, 460  
fertility decline and, 624  
growth of, 3002  
health care industry and, 1827  
health-care and social services, 378  
Krzywicki’s “industrial feudalism” as precursor, 2118  
liberalism/conservatism and, 1601  
long-term care funding, 1658–1659  
social security systems, 2795–2805
- Welfare system. *See* Social welfare system
- Well-being. *See* Quality of life
- Wellman, Barry, 367
- Wellness Councils of America, 588
- Wellpoint Health Networks, 1822
- Wells, H.G., 1038
- Weltanschauung*, 2953, 2955
- Wesolowski, Wlodzimierz, 2120, 2122
- West, Cornel, 2220, 2221
- West Africa, countries of, 60
- West Bank, 1866
- West Germany. *See* Germany
- Westergaard, John, 224, 225
- Westermarck, Edward, 1274, 2449–2450
- Western, Bruce, 385
- Western Electric Company, 2325
- Western Marxism. *See* Marxist sociology
- Westley, William, 2114
- Westmarck Society (Finland), 2451
- Westphalian system (1648), 1933, 2356, 2363
- WFS. *See* World Fertility Survey
- What is a Case?* (Ragin and Becker eds.), 243
- Wheaton, Blair, 656
- Wheelan, Susan, 1979–1980
- Wheeler, Stanton, 3251–3252, 3253
- Wheeler-Howard Indian Reorganization Act of 1934, 136
- When Work Disappears* (Wilson), 3198
- Whiskey. *See* Alcohol
- White, Cynthia, 172
- White, Douglas, 548
- White, Harrison, 172, 739, 1789, 1984, 2298, 2662  
vacancy chain model, 2691
- White, Leslie, 2891
- White, Lynn K., 1737
- White, Martin, 225
- White, Michael J., 2501, 2502
- White Aryan Resistance, 462
- White Citizen’s Council, 2266
- White-collar crime, 530, **3245–3255**  
convictions and sentencing, 3251–3252  
definitions of, 3245–3246, 3247, 3248, 3251  
neutralization theory on, 1496  
offenders, 3251  
research and theory implications, 3252–3254  
violation types, 3249–3250
- White-Collar Crime* (Sutherland), 3247
- White-collar jobs  
in labor market hierarchy, 1988  
and postindustrial society, 2197, 2198, 2199
- Whitehead, Alfred North, 821
- Whitehead, Barbara Dafoe, 106
- Whiteness, concepts of, 56, 57–58, 62
- WHO. *See* World Health Organization
- Who Shall Survive?* (Moreno), 2728–2729
- Whorf, Benjamin, 2890
- Whyte, Martin, 360, 1734
- Whyte, William F., 244, 363, 364, 365, 510, 696, 2041  
street-corner gangs study, 2611, 2613
- Wiarda, Howard, 1537
- Wiatr, Jerzy, 2119
- Wicklund, R. A., 2509
- Widaman, K. F., 1899
- Widgery, Alban, 1292
- Widowhood, **3255–3261**

- case studies of, 244–245  
cross-cultural, 3255–3256  
demographics, 126, 3256–3257  
divorce adjustment compared  
with, 3258–3259  
as life-cycle transition point, 1616  
as longer duration for  
women, 2177  
and non-cohabiting frequency of  
sex, 2539  
remarriage, 126, 1744, 1749,  
2387, 2388, 3255, 3259  
research findings, 3257–3258  
Social Security benefits,  
2406, 2799  
and social support, 3259–3260  
suicide rate, 3078
- Wiederman, M. W., 2541  
Wiener, Susan J., 3199  
Wierzbick, Susan, 368  
Wiese, Leopold von, 1074,  
1076, 1077  
Wigmore, John Henry, 1555  
Wilcoxon matched-pairs signed-rank  
test, 1960, 1961  
Wildavsky, Aaron, 2762  
Wilensky, Harold L., 426, 427,  
2260, 2261  
Wiley, Mary Glenn, 196  
Wiley, Norbert, 2219–2220  
Wilks, S. S., 3035  
Willer, David, 2672, 2673  
Willet, David, 362  
Willey, Kathleen, 2581  
Williams, Dale E., 2270  
Williams, Eric, 321  
Williams, Patricia, 54  
Williams, R. M., Jr., 689  
Williams, Raymond, 1756, 2169,  
2171  
Williams, Robin M., 3214  
Williams, Terry, 510  
Williams & Connolly, 469  
*Williams v. Employers Liability  
Assurance Corporation, Limited*  
(1961), 465  
Williamson, John B., 2211  
Williamson, Oliver, 726, 735  
Willing to pay (WTP) vs. willing to  
accept (WTA), 594
- Willis, Cynthia, 197  
Willis, Paul, 2171  
Wilson, Bryan, 2483–2484  
Wilson, David S., 118  
Wilson, Edward O.  
on genocide, 1068  
and sociobiology, 1234, 2881,  
2882, 2892  
Wilson, J. M., 1933  
Wilson, James Q., 506, 603, 2056,  
2108, 2114  
Wilson, Robert N., 172  
Wilson, Warner, 2303  
Wilson, William Julius, 536, 830  
on class and race, 319, 1490  
on government services, 964, 968  
on permanent underclass, 513,  
845, 2212–2213, 2497–  
2498, 3198  
social disorganization theory, 366  
Wilson, Woodrow, 1423, 1945  
Wimbledon, 2985  
Winant, Howard, 319  
Winch, Robert, 1777  
Windelband, Wizhelm, 819–820  
Wine, 92, 95, 1642  
“Winner-take-all” elections, 2164  
Winnings, division of, 330, 332  
Winrod, Gerald B., 2370  
Winship, C., 2439  
Winternitz, Milton C., 324, 325  
Wippler, Reinhard, 1080  
*Wirtschaft und Gesellschaft*  
(Weber), 1075  
Wirth, Louis, 324, 325, 367, 1209,  
1425, 1502, 1772, 2959  
Wisconsin Idea, 1180  
Wisconsin Longitudinal Study of the  
High School Class of 1975, 972  
Wisconsin model, 757–758, 2714,  
2782–2783, 2784, 2785–2786,  
2788  
Wish, Myron, 1978  
*Wissenssoziologie*, 2953, 2954, 2956  
Wissler, Clark, 675  
*Witchcraft, Oracles and Magic among  
the Azande* (Evans-Pritchard), 2890  
Withey, Stephen B., 2300, 2303,  
2683–2684  
Wittfogel, Karl August, 1075
- Wittgenstein, Ludwig, 821, 1783,  
2756, 2757  
Wohl, Stanley, 1818  
Wohlin, Nils R., 2450  
Wohlstein, Ronald, 555  
Wold, Herman, 3035  
Wolf, Eric R., 2939  
Wolf, Susan Berns, 460  
Wolfe, Alan, 358  
Wolfe, D., 696  
Wolfensohn, James, 1321  
Wolff, Janet, 173, 1647  
Wolfgang, Marvin, 534, 664,  
1488, 1489  
Wolpe, Joseph, 2085  
Wolters, O. W., 2978
- Women  
adolescent sexual behavior  
patterns, 2551–2552  
AIDS/HIV risk, 2586, 2587,  
2589–2590  
as caregivers of elderly, 129  
comparable worth issue, 369–372,  
645, 2706, 3048  
conformity inclination, 403  
depression incidence, 649,  
652–653  
direct and indirect discrimination  
against, 689–690  
divorce and, 701–705, 707–708  
divorce factors, 112  
educational and occupational  
status attainment factors,  
2785–2786, 3046, 3048  
educational opportunities, 962  
employment/mental health  
link, 1837  
family and household roles,  
695–696  
family law rights of, 951  
as family violence victims,  
196, 247, 981  
fertility transition factors, 624,  
627, 634, 635, 2181  
income disparity. *See subhead*  
comparable worth issue *above*  
juvenile delinquency rate  
rise, 1489  
kinship systems and, 1515

- in labor force, 123, 126, 127, 142, 239, 424–425, 512, 624, 625, 703, 705, 962, 972, 973, 981, 1009, 1219, 1523–1526, 1579, 1729, 1837–1838, 2032, 2262, 2404, 2406, 2532, 2706, 3046, 3064, 3262, 3266
- in legal profession, 468, 2262, 2263
- and leisure, 1584, 1585
- life expectancy, 1630, 1631
- life-cycle events, 1623
- and literary sociology, 1648
- love expectations of, 1700
- male leadership image and, 1570
- marriage statistics, 1738
- media stereotypes, 1767–1768
- mental health status, 1837–1838
- Mexican studies, 1861
- military service by, 1879–1880
- morality of care and, 1900, 1902–1903
- motherhood role, 2036
- noneconomic domestic roles of, 122
- occupational segregation, 370–372, 379, 1060–1062, 1716, 2012, 3046, 3262, 3264–3265
- ordination of, 2379
- percentage in top professions, 2259
- poverty and, 1288, 2033, 2215, 3048
- in professions, 2259, 2262–2263
- prostitution and, 2560–2561
- relatively high status in Southeast Asia, 2976
- remarriage potential, 2388–2389
- retirement income, 84
- semiprofessional occupations, 2261, 2262
- sexual activity norms, 2568–2569
- sexual harassment and, 2580–2581
- sexual violence toward, 2576–2583
- sexually transmitted disease risks, 2578, 2583, 2584
- as single-parent household heads, 2033
- as slaves in Africa and Asia, 2601
- Sociologists for Women in Society, 153–155
- status in Muslim societies, 2948–2950
- suffrage, 703
- as temporary workers, 1724
- Victorian-era view of, 1697
- voluntary association participation, 3228
- “women’s work” categories, 1525, 3264
- See also* Childbearing; Pregnancy and pregnancy termination; Widowhood
- Women, Infants and Children Program, 1334
- Women’s movement, 558
- and Equal Rights Amendment, 2267
- and family violence, 981
- and mass media research, 1768
- nineteenth-century, 989
- outcomes evaluation, 2724
- and rape awareness, 2576
- rejection of masculinity-femininity scale, 999
- as social movement, 2717, 2719
- See also* Feminist theory; Gender; Gender roles
- Women’s work. *See* Occupational segregation
- Wonderlic Personnel Test, 1375, 1376
- Wood, Ellen Meiksins, 1785
- Wood, James, 2380
- Wood, W., 2530
- Woodiwiss, Anthony, 227–228
- Woods, Cindy J. P., 1903
- Woolcock, Michael, 2641
- Woolf, Virginia, 989
- Woolgar, Steve, 2459, 2763
- Word processing, 407–408
- Work and occupations, **3261–3269**
- access to, 35
- in adolescence, 8–9, 11–12, 13, 34–35
- in adulthood, 26, 27
- affirmative action and, 47
- African American job categories, 56–57
- in American society, 141
- Asian-American clustering, 181–182
- authority in industry study, 383
- commitment and, 3273–3274
- comparable worth issue, 369–372, 2706
- corporate organizations and, 395–396, 442–443
- daily time use, 3160
- definitions of work, 3261, 3269
- deviant behavior/limited opportunity linkage, 665
- direct and indirect discrimination, 689–693
- See also* Occupational segregation
- division of labor, 696–697, 698–699
- employee benefits, 1282
- employment relationship, 3269–3276
- ethnomethodological workplace studies, 860
- evolution of, 3261–3262
- externalization of work, 3267–3268
- family life separated from, 122
- gender and, 122, 1059–1062, 1720–1721
- gender differences in job satisfaction, 3275
- ghettoization in, 1061
- glass ceiling/glass escalator in, 1061
- group performance, 2617–2618
- from home, 1339, 1341, 3267–3268
- homosexuality and, 111
- industrialization and, 122
- of Jewish immigrants, 2333
- job preparation, 3263
- job satisfaction, 3270, 3272–3373, 3275
- job-allocation process, 3263
- labor force emergence, 3262–3263
- See also* Labor force
- lawyer status, 468–471
- leisure studies, 1582–1584
- in life course, 1614–1615



INDEX

- links to school, 28–29
- marginal employment, 1719–1725
- Mexican categories, 1860
- mobility path model, 1692, 1982–1994
- motivation and, 3270–3271
- organizational demographics, 395–397
- and organizational structure, 696–697, 699, 2002, 2003–2004
- orientation to, 3269–3276
- part-time workers, 1720, 1725, 3262–3263, 3268
- personality as factor in, 2071–2073
- planful competence and, 13, 32
- popular culture transmission and, 2170
- in postindustrial society, 2197, 2198, 2202
- prestige studies, 1997–2001
- professions distinction, 2259–2260
- race and, 35, 1059, 1060, 2716
- retirement and, 2401–2410
- segregation by sex, 370–372, 1060–1062, 1716, 3264–3265
- See also* Occupational segregation
- sexual harassment and, 2591–2592
- smokefree workplaces, 1640
- social security coverage, 2798, 2799
- as socialization agents, 2860
- sociological conception of, 3261
- and sociology of culture, 563
- specialization of roles, 696, 698–699
- standardization of skills, 398–399
- stratification and, 1996–2000, 2813, 3265
- structural lag and, 3060, 3061–3062
- structures, 3263–3265
- suburbanization of, 3073–3074
- systems theory and, 3102–3205
- temporary workers, 1724, 1725, 3268, 3275
- transition from school to, 2714
- transition to adulthood and, 30–32
- white-collar crime and, 3247–3248, 3252, 3253
- “women’s work” categories, 1525, 3264
- worker participation and control, 605
- workplace changes, 3275–3276
- workplace group conflict resolution, 1111
- workplace social control, 3274–3275
- See also* Equality of opportunity; Labor force; Occupational prestige; Wages and salaries
- Work ethic. *See* Protestant ethic
- Work of equivalent value. *See* Comparable worth
- Work orientation, **3269–3277**
- Working class
- British cultural studies of, 2169–2170, 2171
- historical sociology on, 1197
- industrial sociology on, 1309
- marital communication problems, 1736
- poststructural challenge to Marxist theory of, 1756, 1757
- racial division in, 321
- and voice in decision-making, 604
- See also* Class struggle; Proletariat; Social class; Social stratification
- Working poor
- definition of, 1720
- See also* Marginal employment
- Workplace. *See* Work and occupations
- Worland, Stephen T., 2697, 2698
- World Bank, 240, 685, 742, 793, 930
- anticorruption program, 2138
- cross-cultural survey, 549
- dependency theory and, 642
- international comparisons of poverty, 2216
- population policy report, 641
- research funding in China, 302
- and state sovereignty, 3003
- World Development Report*, 2937
- World Tables of Economic and Social Indicators, 2476
- World Commission on Environment and Development, 1222
- World Conference on Women (Beijing; 1995), 932, 1768
- World Congress of International Sociological Associations (1970), 2117
- World Congress of Sociology (1894), 1423
- World Congress of Sociology (1986), 1292
- World Congress of Sociology (1991), 1482–1483
- World Congress of Sociology (1998), 682
- World Cultures data, 548
- World Cup, 2985
- World Development Report*, 2937
- World Dynamics* (Forrester), 2662
- World economy. *See* Globalization and global systems analysis
- World Fertility Survey, 548, 633
- World Food Organization, 2922
- World Future Society, 1038
- World Futures Studies Federation, 1038, 1039
- World Handbook of Political and Social Indicators* (Russett et al.), 2917
- World Health Organization, 930, 2152, 2922
- Quality of Life group, 2302
- sexually transmitted diseases statistics, 2591, 2593
- World Modernization* (Moore), 423
- World population estimates, 2181
- World religions, **3277–3290**
- institutions, 2376–2386, 3286–3287
- intersection of, 3288–3289
- myth and ritual, 3280–3283, 3284, 3285
- theology and doctrines, 3284–3286
- See also* Religion; Religious organizations; Religious orientations; Sociology of religion; *specific religions*
- World system model, 747, 1084, 1085, 1089–1090, 1197, 1199,

- 1214, 1309, 1535, 1538, 1539,  
1706, 1758, 1876, 2646, 2690  
applied to Ottoman Empire and  
Turkey, 1871–1872, 1873
- World systems. *See* Globalization and  
global systems analysis
- World Tables of Economic and  
Social Indicators, 2476
- World Trade Organization, 22433
- World Values Survey (WVS), 549,  
3223, 3229
- World War I  
Armenian genocide, 1070  
and collapse of international  
systems, 2362  
conscription, 1876  
demographic results, 610  
intelligence testing, 2330  
and nationalist revolutions, 3001  
self-determination policy  
following, 1945
- World War II  
and African American  
opportunity, 2493, 2494, 3264  
cohort effects, 2861  
and collapse of international  
systems, 2362  
conscription, 1876  
demographic results, 610, 622,  
703–704, 1625, 2032  
fertility transition, 622, 626  
Japanese-American relocations,  
123, 177, 178, 180–181  
life course effects, 1619  
life-cycle perspective, 1625  
marriage rate upswing, 1741  
military sociology, 1876, 1881  
modernization theory  
following, 1706  
nationalism following, 1944  
Nazi genocide policy. *See*  
Holocaust; Nazism  
and power elite, 2624  
in Southeast Asia, 2975  
student antiwar movement, 3069  
and survey research  
development, 3232  
university-government  
cooperation, 1180  
war crimes trials, 1429
- World Wide Web. *See* Internet;  
Web sites
- Worldwatch Institute Report  
(1990), 1222
- Worms, René, 1025, 1422, 1424
- Wounded Knee occupation  
(1973), 137
- Wright, Erik O., 723, 1754, 2054,  
2692, 2814–2815
- Wright, J., 3209
- Wright, James D., 102–103
- Wright, Jim (politician), 2128
- Wright, Sam, 555
- Wright, Sewell, 259, 908  
path analysis, 259, 455–456,  
1788, 3035
- Wrong, Dennis, 1223
- Wu Jingchao, 298
- Wundt, Wilhelm, 820, 1074, 1423
- Wuthnow, Robert, 1707, 2380, 2484,  
3228–3229
- WVS (World Values Survey), 549,  
3223, 3229
- WWW. *See* Internet
- X**  
X-rated films, 2185
- Y**  
Yahweh, 3280, 3282, 3283, 3285
- Yale School of Medicine, 324, 325  
Institute of Human  
Relations, 325
- Yale University, 3251
- Yale University *Bulletin*, 325
- Yamaguchi, K., 2672, 2674
- Yamasee War, 134–135
- Yankee City studies (Warner), 363,  
364–365
- Yasemin, Soysal, 2662
- Yasuda, Saburo, 1480
- Yeager, Peter C., 1102, 1103, 1106,  
1107, 3247–3248, 3248–3249,  
3250, 3252, 3253
- Year in Civil Liberties, The*  
(ACLU), 315
- Yeltsin, Boris, 2136, 2981
- Yemen, 1865, 1866, 1867, 1947  
sociodemographic profile, 2938
- Yetton, P. W., 1569
- Yilmaz, Mesut, 2132
- Yin, P., 1628
- Yinger, Milton, 459, 460–461,  
462, 463
- Yoneda, Shotaro, 1477, 1478
- Yoruba, 54
- Yoshida, Tamito, 1480
- Young, Donald, 1237
- Young, Lawrence, 2375
- Young, M. Crawford, 1944
- Young Americans for Freedom  
(YAF), 3069
- Young Generation of Peasants*  
(Chalasiniski), 2119
- Youth crime. *See* Juvenile  
delinquency, theories of; Juvenile  
delinquency and juvenile crime
- Youth culture, 460, 512, 514  
innovation and transmission of,  
2170, 2171  
and legislation of morality,  
1578–1579  
and religious movements, 2366,  
2367, 2374
- Youth curfews, 315
- Youth gangs. *See* Gangs
- Yugoslavia (former), 1, 1934,  
1946, 1948  
conflicts in, 1199, 2362–2363  
nation-state formation from, 2362  
socialist economic  
modifications in, 2850  
sociology in, 2117  
wartime rape, 2579
- Yule's Q phi, 661
- Z**  
ZA. *See* Zentralarchiv fur Empirische  
Sozialforschung
- Zagare, Frank C., 332
- Zago, Moreno, 1472, 1473, 2231
- Zaire  
AIDS/HIV infection, 2583  
genocide, 68  
sultanistic ruler, 2133
- Zald, Mayer N., 1941, 1942
- Zambia  
AIDS/HIV infection, 2591  
family and population policy, 931  
poverty in, 2216

---

INDEX

---

- Zander, Alvin, 2611
- Zanna, M. P., 335, 339
- Zaslavskaja, Tatiana, 2980, 2981, 2983
- Zavala, Lorenzo da, 1856
- Zavalloni, M., 3214
- Zeigarnik, B., 2784
- Zeigarnik effect, 2784
- Zeisel, Hans, 3155
- Zeitgeist, great-man theory vs., 1564–1565
- Zeitlin, Maurice, 443, 1088
- Zeitz, G., 603
- Zelditch, Maurice, 2702
- Zeleny, Milan, 1557–1558
- Zelizer, Viviana, 737
- Zeller, R. A., 2345, 2346, 2347, 2350, 2351, 2352
- Zenger, John Peter, 273
- Zentralarchiv für Empirische Sozialforschung, 575, 576, 579, 580, 2477
- Zermelo, E., 1045
- Zero-sum games, 330, 332
- Zetterberg, Hans, 425, 1787–1788, 2452
- Zhou, M., 182
- Zhou, Xueguang, 2662
- Zhu Rongji, 2137
- Zilber, Tamar, 1636
- Zimbabwe, 321
- family policy, 932–933
- fertility rate decline, 220, 628
- Zimbardo, Philip, 901
- Zimmer, Jules, 194
- Zimmerman, Carle C., 362–363
- kinship and family structure, 1503–1504, 1505, 1506–1507
- rural sociology, 2426, 2427, 2428
- Zimmermann, Ekkart, 2157
- Ziolkowski, Marek, 2119, 2121
- Zionism, 2359
- Znaniecki, Florian, 1616, 1618, 1634, 2118, 2220, 2221, 2498
- and analytic induction method, 2297
- and sociology of knowledge, 2954
- Zola, Irving K., 1816
- Zolberg, Vera, 173, 174
- Zolof, 654
- Zoning laws
- against “adult” ventures, 2186
- and suburban exclusionary policies, 3072, 3073
- Zorbaugh, Harvey Warren, 324
- Zoroastrian, 3287–3288
- Zuckerman, Harriet, 342–343, 2458
- Zulu Empire, 2999
- Zung Self-Rating Depression Scale (SDS), 654
- Zuravin, Susan, 195
- Zurcher, Louis, 554
- Zvoina, William, 3038
- Zweigert, Konrad, 472, 1545, 1547
- Zwelakhe, Tshandu, 643
- Zylan, Yvonne, 2961